JOURNAL OF CHEMICAL EDUCATION

Article

pubs.acs.org/jchemeduc

Using Free Computational Resources To Illustrate the Drug Design Process in an Undergraduate Medicinal Chemistry Course

Ricardo P. Rodrigues,[†] Saulo F. Andrade,[‡] Susimaire P. Mantoani,[†] Vera L. Eifler-Lima,[‡] Vinicius B. Silva,[§] and Daniel F. Kawano*[,‡,∥]

[†]Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, Avenida do Café s/n, 14040-903 Ribeirão Preto, São Paulo, Brazil

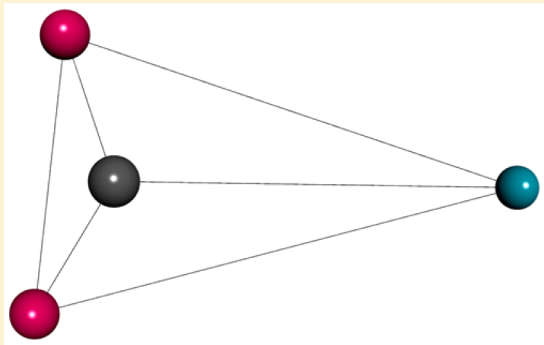
[‡]Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, Avenida Ipiranga 2752, 90610-000 Porto Alegre, Rio Grande do Sul, Brazil

[§]Departamento de Biomedicina e Farmácia, Pontifícia Universidade Católica de Goiás, Rua 232 no. 128, 74605-140 Goiânia, Goiás, Brazil

[∥]Departamento de Química Orgânica, Instituto de Química, Universidade Estadual de Campinas, Rua Josué de Castro s/n, 13083-970 Campinas, São Paulo, Brazil

Supporting Information

ABSTRACT: Advances in, and dissemination of, computer technologies in the field of drug research now enable the use of molecular modeling tools to teach important concepts of drug design to chemistry and pharmacy students. A series of computer laboratories is described to introduce undergraduate students to commonly adopted *in silico* drug design methods, such as molecular geometry optimization, pharmacophore modeling, protein−ligand docking simulations, homology modeling, virtual screening, and pharmacokinetics/toxicity predictions. Freely available software and web servers are selected to compose this pedagogical resource, such that it can be easily implemented in any institution equipped with an Internet connection and Windows OS computers. This material is an illustration of a drug discovery pipeline, starting from the structure of known drugs to obtain novel bioactive compounds, and, therefore, is a valid pedagogical instrument for educating future professionals in the field of drug development.



KEYWORDS: *Graduate Education/Research, Laboratory Instruction, Computer-Based Learning, Drugs/Pharmaceuticals, Molecular Modeling*

INTRODUCTION

The use of computer technology in the field of drug research has increased over the past decades. The availability of faster and cheaper computers in association with the development of more accurate software makes it possible to hold information about the known properties of molecules, to perform drug docking calculations, and to aid in the recognition of protein binding sites and their interactions via 3D visualization systems.[1] At the same time, the use of these computer-based tools for teaching chemistry, and particularly drug design, has been much less explored[2]—a peculiar fact considering that the education of undergraduates and professionals is a crucial factor in the success of drug discovery projects.[3]

It has been observed that "the core knowledge required of organic and medicinal chemists for careers in drug synthesis is surely knowing how to synthesize quickly and efficiently a wide range of candidate structures and at the same time knowing what molecules are more likely to have the desired biological properties,"[4] the latter being a field of computer-aided drug design. Therefore, by introducing modern concepts of *in silico* drug design to chemistry and pharmacy students, academic institutions can provide an opportunity for their students to gain some experience with techniques that have been used to design drugs that reached the market, such as the antiglaucoma agent dorzolamide, the antiviral saquinavir, or the antihypertensive aliskiren.[5]

In this context, a series of 4 h computer laboratories was envisioned to provide pharmacy and chemistry students a practical experience with some important issues regarding the drug design process, such as the study of target−ligand interactions, geometry optimization, molecular docking, pharmacophore, and homology modeling. All the software and online tools used here are free of charge to academic groups, a fact that allows for easy implementation in any institution equipped with Internet connection and Windows OS computers. These protocols were introduced to pharmacy students of Universidade Federal do Rio Grande do Sul, Brazil, in a first-semester medicinal chemistry course to illustrate the

Published: February 24, 2015

DOI: 10.1021/ed500195d
J. Chem. Educ. 2015, 92, 827−835

process of drug design and to foster discussion of issues concerning structure−activity relationships for nonsteroidal anti-inflammatory drugs (NSAIDs).

NSAIDs are among the most widely prescribed medications in the world. Although many patients may benefit from their anti-inflammatory and analgesic effects, these drugs may increase the risk of gastrointestinal and cardiovascular complications compared with non-NSAID users.[6] NSAIDs act via inhibition of prostaglandin endoperoxide H synthase (PGHS), a bifunctional hemoprotein endowed with both cyclooxygenase and peroxidase activities. Because the cyclooxygenase (and not the endoperoxide) activity of PGHS is inhibited by NSAIDs to achieve the clinical effect, this enzyme is often simply referred to as cyclooxygenase (COX).[7] Considering the recommendation that teachers must emphasize not only the core science that underpins their discipline but also the relevance of that science to the contemporary practice of the profession,[8] the use of COX, a clinically relevant drug target, is proper to illustrate a drug discovery pipeline.

BACKGROUND

Course Description

With almost 7 million students in 2012, Brazil has the largest higher education system in Latin America.[9] In addition to private universities, there are also federal, state, and municipal universities in Brazil, which are autonomous public institutions that offer education of excellence, free of charge. State plus federal institutions account for fundamentally almost all the research output and the PhD programs in the country.[10] The Universidade Federal do Rio Grande do Sul (UFRGS) is a Brazilian Federal University founded in 1895 and is ranked among the top 10 universities in Latin America according to the ARWU, Webometrics, EduRoute, SCImago, URAP, CWTS Leiden, Research Gate, and QS World University Rankings.[11]

Students come to the Medicinal Chemistry course at UFRGS having had formal coursework in pharmacology, pharmacokinetics, immunology, pathology, physiology, anatomy, biochemistry, physics, organic, inorganic, and analytical chemistry. The first semester of the Medicinal Chemistry course (6 credit hours) focuses on the fundamental aspects and current methodologies involved in the drug discovery process (*in silico* methods, prodrugs, bioisosteres, and the primary exploration of structure−activity relationships), and in the study of drugs that act in the cardiovascular, respiratory, and central nervous systems. Concurrent with the lectures, laboratory experiments are performed to study the physicochemical properties of drugs, the strategies for the synthesis of classical drugs and prodrugs, and *in silico* approaches for drug design.

The computer laboratories of the Medicinal Chemistry course at UFRGS were updated in the second semester of 2012, and, since then, about 28−38 students/semester (total = 157) have been exposed to the protocols reported here. Students are typically divided into four groups (10 individuals max per group) where they work individually for 4 h/class under the supervision of the same professor who previously introduced the underlying concepts of the methods in lectures. Although we did not perform a formal quantitative analysis, most students were receptive to these initiatives and expressed a view that these computer laboratories helped them to improve their understanding of a drug discovery pipeline.

Evidence of Student Learning through Computer Simulations

Several papers suggest that manipulating physical or computer models in laboratory and lecture courses may enhance student understanding of chemistry concepts.[12] Carvalho et al.[13] and Oliveira et al.[14] described the use of molecular modeling tools to increase students' perception of molecular recognition and ligand binding interactions. Tsai[15] used Web sites and freeware programs to introduce the principles of pharmaceutical chemistry in drug discovery and to solve pharmaceutical chemistry problems. Manallack et al.[16] and Sutch et al.[17] reported case studies in which students designed potential drugs through the use of software, and Simpson et al.[18] used geometry optimizations and optical rotation calculations to illustrate to students the role of the solvent in stabilizing the zwitterionic form of an amino acid. Recently, Price et al.[19] reported a Python script to assist undergraduates in file conversion, minimization, and docking experiments, and Hayes[20] reported a considerably positive response from students after the implementation of a 3D model visualization and basic molecular modeling laboratory suitable to an introductory medicinal chemistry course.

Simulations are more effective when the educational goal is to transfer and apply knowledge to real-world problems rather than simply to memorize facts or procedures. Additionally, low fidelity simulations may help students to learn the basics more quickly than high fidelity simulations because the more closely a simulation models a complex dynamic system, the more difficult the simulation is for someone to learn and understand how to use.[21] Therefore, these protocols are not used to teach students how to perform high accuracy molecular docking simulations using ArgusLab.[22] Rather, the intent is to use ArgusLab as a simple, free and user-friendly molecular docking tool that allows observation on a computer screen of how docking algorithms work. Furthermore, many students have a great deal of difficulty understanding and learning abstract concepts, such as pharmacophore. Computers allow demonstration of complex abstract ideas and provide multiple examples in seconds, enhancing student learning by self-resolution of alternative conceptions: informal ideas that have the general characteristics of being poorly articulated, internally inconsistent, and highly dependent on context.[23,24]

COMPUTER LABORATORIES OVERVIEW AND DISCUSSION

Lab 1. Conversion of 2D Structures to 3D, Geometry Optimization, and Construction of a Pharmacophore Model

Chemistry is a 3D subject, as illustrated by topics such as stereochemistry, chirality, conformation, metal coordination geometry, and molecular symmetry.[25] However, 2D structural representations are usually preferred to teach and study chemistry because they are easier to interpret, allow the complete visualization of a molecule at once, and provide a simple way to recognize chemical patterns. Furthermore, only a few 3D structures of small molecules determined by X-ray crystallography are freely available. For example, the Cambridge Structural Database contains the crystal structures for over 600,000 organic and organometallic compounds:[26] a tiny fraction of the more than 88 million organic and inorganic substances registered by Chemical Abstracts Service.[27] Therefore, the accurate conversion of 2D to 3D structures represents

DOI: 10.1021/ed500195d
J. Chem. Educ. 2015, 92, 827−835

an important issue for *in silico* drug design, since, for example, 3D structures of ligands are necessary for molecular docking or pharmacophore modeling.

Many software packages are able to convert 2D representations to 3D structures directly, but this is usually an error-prone process that rarely may be used alone to produce highly accurate 3D structural representations. One example error is the generation of deformed nonplanar aromatic rings after conversion of 2D representations to 3D structures,[28] a structural feature that is certainly far from what would be observed in crystal structures. Therefore, it is very convenient to use a geometry optimization method (also known as energy minimization) to refine 3D structures in order to adjust bond length, angle, and torsion.

In the first lab class, students draw the 2D representations of eight NSAIDs using ChemSketch 12.0[29] and save each in separate files (interchangeable mol format) to be opened in ArgusLab 4.0.[22] These eight NSAIDs are nonselective COX inhibitors (inhibit both the COX-1 and COX-2 isoforms) and belong to two different subclasses: propionic acids (ibuprofen, ketoprofen, naproxen, and flurbiprofen) and aryl/heteroaryl acetic acid derivatives (tolmetin, etodolac, indomethacin, and sulindac).

Except for indomethacin and tolmetin, all these NSAIDs have two stereoisomers or geometrical isomers (e.g., sulindac). In this process, students will draw only the most active enantiomer/geometrical isomer to define a correct pharmacophore model at the end of the lab. However, they must be aware that different enantiomers and geometrical isomers may have not only different affinities for the pharmacological target but also different pharmacokinetic and toxicity profiles, as exemplified by the preference of human albumin by the active *S*-enantiomer of ketoprofen.[30]

The 2D representations are automatically converted using the 3D optimization feature in ChemSketch, a 3D optimization algorithm modified from a molecular mechanics package (CHARMM) that takes into account bond stretching, angle bending, internal rotation, and van der Waals nonbonded interactions.[31]

Once in ArgusLab, proper stereochemistry of the NSAIDs can be easily confirmed by simply rotating the lowest priority group toward the back (Cahn–Ingold–Prelog rule set) and observing whether the other three branches are ordered in clockwise (*R*) or counterclockwise (*S*) fashion. Each structure is submitted for geometry optimization using the PM3 (parametrized model number 3), a semiempirical force field based on the neglect of differential diatomic overlap (NDDO) integral approximation.[32] Besides observing in "real time" the adjustments to bond length, angle, and torsion on the computer screen, students may confirm the efficiency of the method in optimizing the 3D geometry of the NSAIDs by calculating the percentage change between the initial and final energy of each compound after energy minimization.

The 3D optimized NSAID structures are used to generate a pharmacophore model to screen potential inhibitors of COX. A pharmacophore can be viewed as the maximum common denominator of a group of molecules that exhibit similar pharmacological profiles by acting at the same site of a target.[33] A pharmacophore relies on a 3D point of view of molecules since it reflects the way medicinal chemists characterize the binding ability of molecules for a given target.[34] Therefore, students must keep in mind that the goal of geometry optimization here is to calculate bond lengths that are similar to the corresponding ones in the crystallographic pivot, a minimum requirement to achieve a reliable pharmacophore model.

For the generation of the pharmacophores, the PharmaGist web server allows assignment of a "pivot" molecule on which the other target ligands are aligned.[35] Students use the NSAID diclofenac in its bioactive conformation, as found in the cyclooxygenase active site of the enzyme (PDB code: 3N8Y), as the "pivot" structure. This is an opportunity to discuss with students that, based on the induced fit theory, the bioactive conformation of a flexible ligand does not necessarily correspond to the lowest energy conformer, as it would be intuitive to think.[36] The link with the resulting pharmacophores generated by the PharmaGist web server is sent by e-mail to students to be analyzed in the next computer lab.

By starting with simplistic two-dimensional drawings of chemical structures to study their corresponding stereochemistry, bond lengths, angles, and torsions in more complex 3D models and to finally build and analyze a very abstract concept such as a pharmacophore, a progressive learning approach is provided by which students can move toward more sophisticated understanding of important scientific concepts. The move toward expertise requires building a more complex idea upon the understanding of some underpinning knowledge, and incorporating more ideas and connecting to ideas of other related topics.[37]

Lab 2. Pharmacophore Screening and Homology Modeling

The pharmacophore models generated in the first class are analyzed by students, and the top-scoring pharmacophore hypothesis in which all nine structures are aligned is selected. This pharmacophore model is composed of three features: two hydrogen bond acceptors (represented by the two oxygen atoms of the carboxylate), a negative ion (carboxylate), and the aromatic ring of the phenylacetic acid moiety of diclofenac (Figure 1). To confirm the validity of this model, students

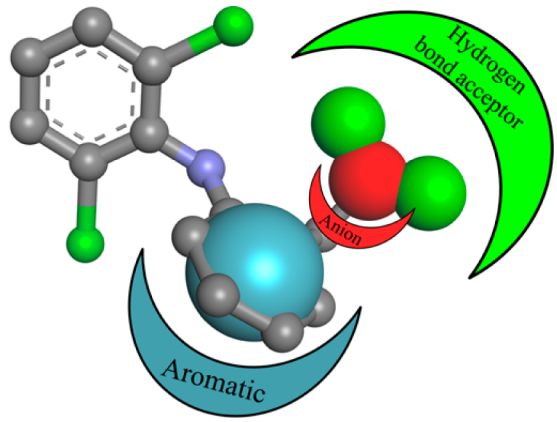


Figure 1. Pharmacophore model generated in PharmaGist and visualized in Discovery Studio Visualizer. The selected model comprises two hydrogen bond acceptors, one negative ion, and one aromatic feature.

analyze the interactions of the amino acid residues in the cyclooxygenase site of the enzyme with the corresponding pharmacophore groups for the aryl acetic acid NSAID diclofenac (PDB code: 3N8Y, chain A) and for the propionic acid NSAID ibuprofen (PDB code: 1EQG, chain A) using Ligand Explorer, a very accessible Java-based program available from RCSB Protein Data Bank, that allows visualizing the

DOI: 10.1021/ed500195d
J. Chem. Educ. 2015, 92, 827–835

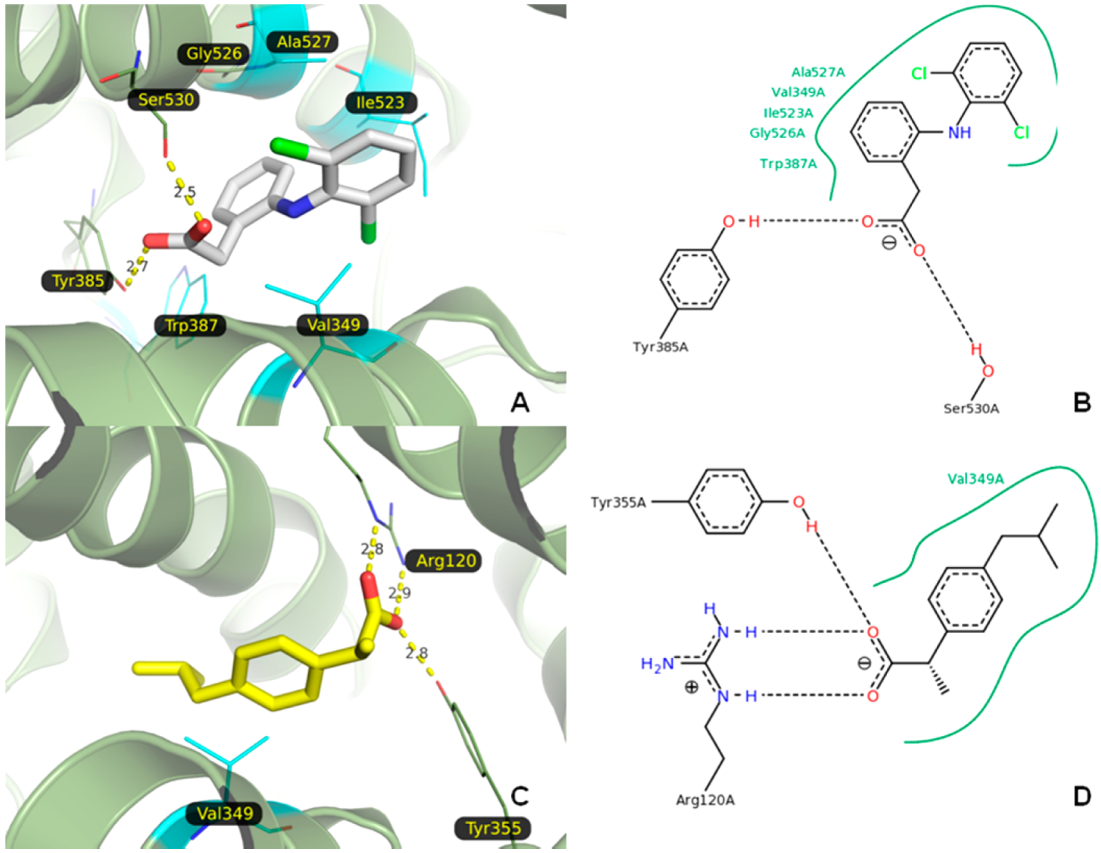


Figure 2. (A) 3D view of the binding site of the crystallographic structure of COX-1 with diclofenac (PDB code: 3N8Y) showing two hydrogen bonds (Tyr-385 and Ser-530, 2.7 and 2.5 Å, respectively) and the hydrophobic pocket in cyan (Ala-527, Ile-523, Val-349, and Trp-387). Although a glycine amino acid does not have a hydrophobic side chain, it contributes to the stabilization of the hydrophobic pocket. (B) 2D Jmol representation of the main interactions of diclofenac. (C) 3D view of the binding site of the crystallographic structure of COX-1 (PDB code: 1EQG) with ibuprofen, showing the hydrogen bonds between the carboxylate of ibuprofen with Arg-120 and Tyr-355 (2.8−2.9 and 2.8 Å, respectively) and also the hydrophobic interactions with Val-349/Ala-527. (D) 2D Jmol representation of the main interactions of ibuprofen.

interactions of bound ligands in protein and nucleic acids structures.[38]

By comparing the two protein−ligand complexes, students will notice the hydrogen bonding interactions involving the carboxylate of diclofenac and the Tyr385/Ser530 residues of COX-1 and the hydrogen bonds between the same group of ibuprofen with Arg-120 and Tyr-355, which are reported to be critical for ligand binding.[39,40] At the same time, they will confirm that the aromatic feature detected by PharmaGist is also important for binding due to the establishment of hydrophobic interactions with Trp-387, Gly-526, and Ala-527 for diclofenac and with Val-349/Ala-527 for ibuprofen (Figure 2). Therefore, the pharmacophore model is considered viable for the screening of potential inhibitors of COX-1.

By understanding the concepts behind the construction of a pharmacophore, students will realize that compounds exerting similar activities at the same enzyme or receptor possess, in most cases, closely related binding properties due to similar chemical features in sterically consistent locations.[34] Additionally, this is an opportunity to illustrate to students that a pharmacophore model, prepared on the basis of only the structural information on ligands, may successfully predict the most important features for the interaction of a ligand within the active site of a drug target.

The pharmacophore-based screening is performed using ZINCPharmer, a free pharmacophore search web server for screening purchasable compounds from the ZINC database.[41] Students recover the 3D information (x, y, and z coordinates) of the pharmacophore features generated by PharmaGist using Discover Studio Visualizer 4.0[42] and perform a search in ZINCPharmer in order to select the 50 top-scoring compounds for use in molecular docking simulations. The notion that this pharmacophore screening is based purely on the presence and arrangement of pharmacophore features and does not account for steric effects is discussed with students. Consequently, this method alone does not definitely ensure complementarity at the ligand binding site.[43] The advantages of choosing a strategy that relies on a pharmacophore model as a prescreening step prior to docking-based virtual screening include providing additional information to improve docking results and significantly shortening the amount of computer time required for docking.[44]

Although this virtual screening is only for educational purposes, students must keep in mind that the selected compounds could be purchased from ZINC and tested *in vitro* for their potential to inhibit COX. Consequently, this process simulates a feasible drug design pipeline that could be used in academia or small pharmaceutical companies. Additionally, by simulating scientific activities, cognitive skills tend naturally to shift from slower, effortful, and consciously controlled rational operations to more intuitive operations, which are faster, automatic, effortless, and associative.[45] Therefore, computer simulations demonstrate great potential for providing creative

DOI: 10.1021/ed500195d
J. Chem. Educ. 2015, 92, 827−835

learning environments that act as very useful intermediates between the classroom and the real world.

The importance of 3D protein structures for molecular docking studies is introduced. However, since there is no experimental 3D structure available for the human COX-1 enzyme, a homology model must be derived using the ovine homologue. Therefore, students use the amino acid sequence of the human COX-1 to retrieve the ten top-scoring sequences of COX homologues with solved 3D structures using the NCBI Blast server.[46] Considering that the COX-1 homologues, with just over 600 amino acids, share about 85−90% sequence identity among different species,[47] the homology of structure can be inferred since these values are significantly higher than the homology cutoff of 25% sequence identity for long alignment lengths (>250 residues).[48] For practical purposes, a minimum of 40% of sequence identity must be assumed to build satisfactory models using the most popular modeling packages.[49]

This is a discussion topic to help students remember important concepts regarding homology modeling, such as (1) homology between amino acid sequences suggests structural and functional similarity, (2) homologous proteins contain conserved internal domains (composed mainly of secondary structure elements: α-helices and β-sheets), and (3) the loop regions are typically much more divergent.[50] Another important point is that the conserved protein regions include the catalytic portion of the enzyme where docking simulations will be performed.

The SWISS-MODEL server is used to build the 3D model of the human COX-1 enzyme.[51] Although the server allows the construction of a model by providing the amino acid sequence of the query protein and the PDB code of the homologous protein 3D template (Automate Mode), the model is constructed on the basis of multiple sequence alignment composed of the query sequence (human COX-1) plus the ten sequences previously retrieved in the Blast search (Alignment Mode). Although one can theoretically construct a homology model based on only the pairwise sequence alignment between the query and the 3D template, the use of multiple sequence alignment of structurally related proteins tends to minimize the probability of errors in the process.[52] Using the multiple alignment of the sequences in the EMBL-EBI ClustalW web server,[53] the underlying theory of sequence homology, i.e., which sites share a common evolutionary history,[54] is discussed with students.

The homology model is built in SWISS-MODEL on the basis of the 3D structure of the ovine COX-2 enzyme cocrystallized with ibuprofen (PDB code: 1CQE). Although backbone atom positions can be more easily assigned by SWISS-MODEL when the aligned residues are identical, regions of insertions or deletions (gaps) in the target-template alignment require the use of a scoring scheme that accounts for force field energy, steric hindrance, and favorable interactions, such as hydrogen bond formation, to select the best loop. Regarding side-chain modeling, the most likely conformation is selected by the use of a backbone-dependent rotamer library in association with a scoring function that determines favorable interactions (hydrogen bonds, disulfide bridges) and unfavorably close contacts. Finally, deviations in the protein structure geometry are regularized in the last modeling step by a steepest descent energy minimization using the GROMOS96 force field.[51]

Model validation is performed using built-in scoring functions for the estimation of the protein structure model quality (QMEAN4 score and QMEAN4 *Z*-score) and by the Ramachandran plot generated online.[55] QMEAN4 (qualitative model energy analysis) is a composite scoring function that describes the major geometrical aspects of a single protein structure and, consequently, reflects the general reliability of the model.[56] When the QMEAN4 score of a model is compared to distributions obtained from high-resolution structures solved by X-ray crystallography, a resulting scoring function called QMEAN4 *Z*-score is generated. Therefore, the QMEAN4 *Z*-score is an estimate of the "degree of nativeness" of the structural features observed in a model by describing the likelihood that a model is of comparable quality to high-resolution experimental structures.[57]

The Ramachandran plot relies on the principle that, due to steric hindrance, the main chain of a polypeptide usually assumes preferred, energetically favorable conformations. Therefore, deviations from the preferred conformations can then be used as indicators of potential errors in the model. In the Ramachandran plot, two torsion angles are used to describe the rotations of the polypeptide backbone around the bonds between N−C_α (called phi, φ) and C_α−C (called psi, ψ). Consequently, the plot provides an overview of the allowed and disallowed regions of the torsion angle values and serves as an important factor in the assessment of the quality of protein three-dimensional structures.[58]

Regardless of the validation method used, students must understand the importance of the model since the generated model must be sufficiently accurate to provide detailed structural information necessary for molecular docking. The ligand ibuprofen from the ovine COX1 structure (PDB code: 1EQG) is inserted into the binding site of the modeled human COX1 structure for use with ArgusLab software for molecular docking since the software uses the ligand coordinates to define the binding site.[22]

In addition to building and validating a model, this is also an opportunity for students to manipulate protein models in order to visualize important protein features, such as α-helices, β-sheets, loops, and binding sites. Although a simple exercise, this may have a very positive impact on learning, since the mental picture of three-dimensional objects is one of the crucial problems students face in structural chemistry, biochemistry, and cell biology.[59] Also, students may use both instructor- and self-guided visualizations of protein structures complexed with drug molecules to gain understanding in the process of protein−drug interactions, understanding the importance of functional groups in the drug and the noncovalent interactions performed with the target.[60]

Lab 3. Docking-Based Screening and Pharmacokinetic/Toxicity Predictions

An emerging viewpoint of cognition suggests that the body has a central role in shaping the mind and that cognitive processes are deeply rooted in the body's interaction with the world that "embodied cognition or learning."[61] If so, the documented difficulties for learners to grasp and to engage in molecular sciences might, at least in part, be explained by the lack of direct experience of the micro world. Although molecular modeling simulations are important for the analysis of biomolecular structures and to understand molecular interactions better, they also provide new ways of teaching that can be powerful cognitive aids for the understanding of the structural require-

DOI: 10.1021/ed500195d
J. Chem. Educ. 2015, 92, 827−835

ments of ligands necessary for their interaction with pharmacological targets.[62]

Molecular docking approaches are useful to model interactions between small molecules and proteins at the atomic level: both for characterizing the behavior of these small molecules at the binding site and for the elucidation of fundamental biochemical processes.[63] Via molecular docking using ArgusLab, the 50 compounds selected in the pharmacophore screening are rescored. This is a very popular molecular modeling package within academia because of its user-friendly interface and intuitive calculation menus.[64] It is important to note that users may experience some problems with stability, especially when running ArgusLab under 64 bit versions of MS Windows. For these situations, the use of Autodock Vina[65] or web-based docking services such as DockThor[66] is recommended.

Before starting the virtual screening protocol, an initial validation of the docking procedure is performed by comparison of the conformation, position, and orientation (which are referred to as pose) of the docked ligand with the corresponding pose of the ligand cocrystallized with the target. The accurate docking of the crystallographic ligand at the protein binding site, a process known as redocking, is considered to be a minimum requirement to determine whether the selected docking parameter is appropriate for the proposed task or not.[67] This comparison is made through the root-mean-square deviation (rmsd) value, a measure of the average distance between atoms of the reference and the docked ligand.

Although an rmsd threshold value of 2.0 Å is widely accepted as distinguishing between success and failure in reproducing a known binding mode,[68] this criterion must be viewed with care depending on the number of rotatable bonds of the ligand since docking programs may have trouble predicting the correct positioning of very flexible ligands.[69] In these simulations, redocking of ibuprofen in the active site of human COX-1 using the GADock docking engine yielded an rmsd of 1.99 Å, a value considered appropriate for academic purposes (Figure 3).

Conceptually, the docking process may be subdivided into two basic steps: sampling conformations of the ligand in the active site of the protein (i.e., reproduce the experimental binding mode), and then ranking these conformations via a scoring function (assessment of the binding affinity).[70] Docking of the 50 compounds derived from the ZINC database is performed using the GADock docking engine, handling the ligand as flexible. Ideally, the flexibilities of both the ligand and receptor must be considered in docking simulations since both the ligand and receptor change their conformations to form a minimum energy perfect-fit complex.[70] However, the computational cost is very high when the receptor is also flexible, and this feature is not available for ArgusLab.

GADock is a built-in genetic sampling algorithm of ArgusLab.[29] Genetic sampling algorithms are inspired by Darwin's theory of evolution. Here, the degrees of freedom of the ligand are encoded as binary strings called "genes" that make up a "chromosome", thus actually representing the pose of the ligand.[70] Therefore, genetic algorithms demand that the fittest "individuals" (poses) are carried on to the next generation, and random or biased mutations can be made to increase genetic (conformational) diversity.[71]

Through the use of its built-in scoring function, AScore,[29] ArgusLab ranks the compounds in order of the suggested binding affinity (kcal/mol). Binding affinity energies tend to be

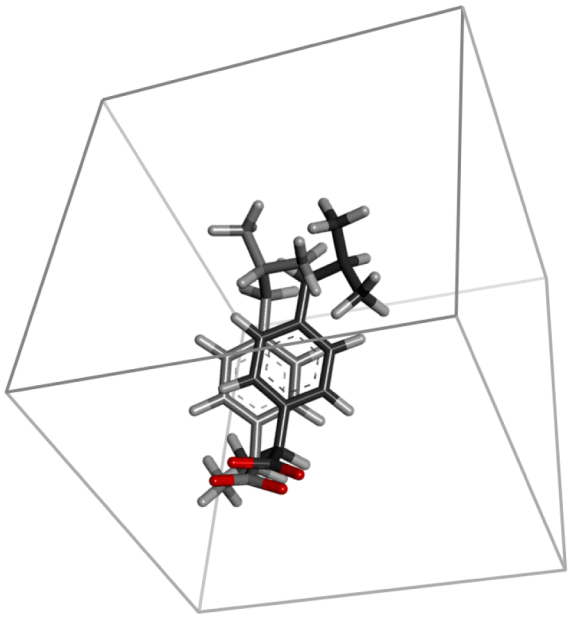


Figure 3. ArgusLab's docking grid. In light gray, the crystallographic conformation of ibuprofen; in dark gray, the top ranked docked solution. Docking was performed using GADock docking engine, handling the ligand as flexible (rmsd = 1.99 Å).

expressed as negative values because they are experimentally quantified through the Gibbs free energy for binding, ΔG_{bind}, which is related to a binding constant (eq 1):

$$\Delta G_{bind} = -RT \ln K_b \quad (1)$$

where R is the gas constant, T is the temperature in kelvins, and K_b is the binding constant. As binding energy indicates how well a protein and a ligand can bind to each other, then, the more negative the binding energy, the stronger the binding between the two molecules.[72] As an empirical scoring function, the AScore decomposes the predicted binding energy into several energy components, such as hydrogen bond, van der Waals interactions, hydrophobic effect, and deformation effect upon binding.[29,73] For an overview of docking concepts, sampling algorithms, and scoring functions, see Meng et al.[70]

Recently, pharmaceutical companies have worked to reduce the number of projects interrupted due to toxicity or pharmacokinetics issues by predicting and optimizing the absorption, distribution, metabolism, excretion, and toxicity (ADMETox) properties of chemical compounds throughout the drug discovery process, rather than at the final stages.[74] To illustrate that ADMETox properties can be predicted from chemical structures, the Osiris Property Explorer tool is used to analyze the ten top-scoring compounds from the docking-based virtual screening.[75] Osiris Property Explorer allows the estimation of varied toxicity effects (mutagenicity, tumorigenicity, skin irritation, and reproductive effects) through the identification of potentially hazardous fragments in the query compound based on comparison to known toxic molecules. The server also calculates druglikeness scores and allows the prediction of other parameters that affect druglikeness, such as LogP (i.e., the ratio of concentrations of un-ionized compound between lipophilic and hydrophilic phases), solubility, and molecular weight. After the analysis of all these properties, students are able to decide which compound(s) would be the

DOI: 10.1021/ed500195d
J. Chem. Educ. 2015, 92, 827−835

most promising to be synthesized and tested for the inhibition of the COX enzyme.

CONCLUSIONS

A series of computational laboratories was introduced to give graduate students a practical experience with some important issues regarding the drug design process, such as the study of target−ligand interactions, geometry optimization, molecular docking, pharmacophore, and homology modeling. The use of freely available software and web servers makes these pedagogical resources useful and easily implementable in almost any institution. Furthermore, these laboratories are effective for illustrating to students the drug discovery approach, starting from the structure of known drugs to obtain potentially bioactive and safer compounds.

ASSOCIATED CONTENT

Supporting Information

Computer lab protocols, suggested literature, and notes for the instructor. This material is available via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: dkawano@ufrgs.br.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank all the companies and developers for the free use of the software and web servers described in this paper. The authors also thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS), and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for the financial support.

REFERENCES

(1) Dalkas, G. A.; Vlachakis, D.; Tsagkrasoulis, D.; Kastania, A.; Kossida, S. State-of-the-art technology in modern computer-aided drug design. *Briefings Bioinf.* **2013**, *14* (6), 745−752.

(2) Gledhill, R.; Kent, S.; Hudson, B.; Richards, W. G.; Essex, J. W.; Frey, J. G. A computer-aided drug discovery system for chemistry teaching. *J. Chem. Inf. Model.* **2006**, *46* (3), 960−970.

(3) Testa, B.; Pedretti, A.; Vistoli, G. Reactions and enzymes in the metabolism of drugs and other xenobiotics. *Drug Discovery Today* **2012**, *17* (11), 549−560.

(4) Ganellin, C. R.; Mitscher, L. A.; Topliss, J. G. Education of medicinal chemists in the departments of medicinal chemistry (USA). *Med. Res. Rev.* **1998**, *18* (2), 121−137.

(5) Talele, T. T.; Khedkar, S. A.; Rigby, A. C. Successful applications of computer aided drug discovery: moving drugs from concept to the clinic. *Curr. Top. Med. Chem.* **2010**, *10*, 127−141.

(6) Sostres, C.; Gargallo, C. J.; Arroyo, M. T.; Lanas, A. Adverse effects of non-steroidal anti-inflammatory drugs (NSAIDs, aspirin and coxibs) on upper gastrointestinal tract. *Best Pract. Res., Clin. Gastroenterol.* **2010**, *24* (2), 121−132.

(7) Tosco, P. A Mechanistic Hypothesis for the Aspirin-Induced Switch in Lipid Mediator Production by Cyclooxygenase-2. *J. Am. Chem. Soc.* **2013**, *135* (28), 10404−10410.

(8) Alsharif, N. Z.; Destache, C. J.; Roche, V. F. Teaching Medicinal Chemistry to Meet Outcome Objectives for Pharmacy Students. *Am. J. Pharm. Educ.* **1999**, *63* (1), 34−40.

(9) Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, INEP. Resumo Técnico: Censo da Educação Superior 2013. http://www.inep.gov.br (accessed Jun 2014).

(10) Steiner, J. E. Qualidade e diversidade institucional na pós-graduação brasileira. *Estudos Avançados* **2005**, *19* (54), 341−365.

(11) Universidade Federal do Rio Grande do Sul, UFRGS. Posição da UFRGS nos rankings de avaliações internacionais. Secretaria de Avaliação Institucional da UFRGS. http://www.ufrgs.br/sai/dados-resultados/a-ufrgs-nos-rankings/arquivos/. *UFRGSnosRankingsInternacionais.pdf* (accessed Jun 2014).

(12) Springer, M. T. Improving students' understanding of molecular structure through broad-based use of computer models in the undergraduate organic chemistry lecture. *J. Chem. Educ.* **2014**, *91* (8), 1162−1168.

(13) Carvalho, I.; Borges, A. D. L.; Bernardes, L. S. C. Medicinal chemistry and molecular modeling: an integration to teach drug structure-activity relationship and the molecular basis of drug action. *J. Chem. Educ.* **2005**, *82* (4), 588−596.

(14) Oliveira, M. T.; Santos, M. A.; Silva, T. H. A. Computer-based conformational analysis of acetylcholine and muscarine combined with an overview of the receptor-ligand interaction. *J. Chem. Educ.* **2006**, *83* (5), 780−781.

(15) Tsai, C. S. Using computer applications and online resources to teach and learn pharmaceutical chemistry. *J. Chem. Educ.* **2007**, *84* (12), 2019−2023.

(16) Manallack, D. T.; Chalmers, D. K.; Yuriev, E. Using the β2-adrenoceptor for structure-based drug design. *J. Chem. Educ.* **2010**, *87* (6), 625−627.

(17) Sutch, B. T.; Romero, R. M.; Neamati, N.; Haworth, I. S. Integrated teaching of structure-based drug design and biopharmaceutics: a computer-based approach. *J. Chem. Educ.* **2012**, *89* (1), 45−51.

(18) Simpson, S.; Autschbach, J.; Zurek, E. Computational modeling of the optical rotation of amino acids: an 'in silico' experiment for physical chemistry. *J. Chem. Educ.* **2013**, *90* (5), 656−660.

(19) Price, G. W.; Gould, P. S.; Marsh, A. Use of freely available and open source tools for in silico screening in chemical biology. *J. Chem. Educ.* **2014**, *91* (4), 602−604.

(20) Hayes, J. M. An integrated visualization and basic molecular modeling laboratory for first-year undergraduate medicinal chemistry. *J. Chem. Educ.* **2014**, *91* (6), 919−923.

(21) Holton, D. L. How people learn with computer simulations. In *Handbook of Research on Human Performance and Instructional Technology*; Song, H., Kidd, T., Eds.; Information Science Reference: Hershey, 2010; Vol. *1*, pp 485−504.

(22) Thompson, M. A. Molecular docking using ArgusLab: An efficient shape-based search algorithm and an enhanced XScore scoring function. *228th ACS National Meeting*, **2004**, Philadelphia, PA, CINF 42.

(23) Windschitl, M.; Andre, T. Using computer simulations to enhance conceptual change: the roles of constructivist instruction and student epistemological beliefs. *J. Res. Sci. Teach.* **1998**, *35* (2), 145−160.

(24) Garfield, J. B.; Ben-Zvi, D.; Chance, B.; Medina, E.; Roseth, C.; Zieffler, A. Using technology to improve student learning of statistics. In *Developing Students' Statistical Reasoning: Connecting Research and Teaching Practice*; Garfield, J. B., Ben-Zvi, D., Eds.; Springer: Dordrecht, 2008; Vol. *1*, pp 91−114.

(25) Henderson, S.; Battle, G.; Allen, F. Teaching chemistry in 3D using crystal structure data. *Educ. Chem.* **2011**, *48* (6), 175−178.

(26) Allen, F. H. The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **2002**, *B58* (3/1), 380−388.

(27) Chemical Abstracts Service, CAS. Database counter (updated daily). CAS database. http://www.cas.org/content/at-a-glance (accessed Jun 2014).

DOI: 10.1021/ed500195d
J. Chem. Educ. 2015, 92, 827−835

(28) Patrick, G. L. *An Introduction to Medicinal Chemistry*, 4th ed.; Oxford University Press: Oxford, 2009; pp 333−334.
(29) *ACD/ChemSketch, version 12.01*; Advanced Chemistry Development, Inc.: Toronto, Canada, 2009.
(30) Smith, S. W. Chiral toxicology: it's the same thing... only different. *Toxicol. Sci.* **2009**, *110* (1), 4−30.
(31) *ACD/ChemSketch, version 12.0; Reference manual: comprehensive interface description*; Advanced Chemistry Development, Inc.: Toronto, Canada, 2010; p 189.
(32) Rappe, S.; James, J. P. Optimization of parameters for semiempirical methods I. Method. *J. Comput. Chem.* **1989**, *10* (2), 209−220.
(33) *Pharmacophores and Pharmacophore Searches*, 1st ed.; Langer, T., Hoffmann, R. D., Eds.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, 2006; pp 3−16.
(34) *The Practice of Medicinal Chemistry*, 3rd ed.; Wermuth, C. G., Ed.; Academic Press: 2008; pp 572−586.
(35) Schneidman-Duhovny, D.; Dror, O.; Inbar, Y.; Nussinov, R.; Wolfson, H. J. PharmaGist: a web server for ligand-based pharmacophore detection. *Nucleic Acids Res.* **2008**, *36* (Suppl. 2), W223−W228.
(36) Perola, E.; Charifson, P. S. Conformational analysis of drug-like molecules bound to proteins: an extensive study of ligand reorganization upon binding. *J. Med. Chem.* **2004**, *47* (10), 2499−2510.
(37) Stevens, S. Y.; Delgado, C.; Krajcik, J. S. Developing a hypothetical multi-dimensional learning progression for the nature of matter. *J. Res. Sci. Teach.* **2010**, *47* (6), 687−715.
(38) Moreland, J. L.; Gramada, A.; Buzko, O. V.; Zhang, Q.; Bourne, P. E. The Molecular Biology Toolkit (MBT): a modular platform for developing molecular visualization applications. *BMC Bioinf.* **2005**, DOI: 10.1186/1471-2105-6-21.
(39) Selinsky, B. S.; Gupta, K.; Sharkey, C. T.; Loll, P. J. Structural analysis of NSAID binding by prostaglandin H2 synthase: time-dependent and time-independent inhibitors elicit identical enzyme conformations. *Biochemistry* **2001**, *40* (17), 5172−5180.
(40) Sidhu, R. S.; Lee, J. Y.; Yuan, C.; Smith, W. L. Comparison of cyclooxygenase-1 crystal structures: cross-talk between monomers comprising COX-1 homodimers. *Biochemistry* **2010**, *49* (33), 7069−7079.
(41) Koes, D. R.; Camacho, C. J. ZINCPharmer: pharmacophore search of the ZINC database. *Nucleic Acids Res.* **2012**, *40* (W1), W409−W414.
(42) *Discovery Studio Modeling Environment, Release 4.0*; Accelrys Software Inc.: San Diego, CA, 2013.
(43) Toba, S.; Srinivasan, J.; Maynard, A. J.; Sutter, J. Using pharmacophore models to gain insight into structural binding and virtual screening: an application study with CDK2 and human DHFR. *J. Chem. Inf. Model.* **2006**, *46* (2), 728−735.
(44) Peach, M. L.; Nicklaus, M. C. Combining docking with pharmacophore filtering for improved virtual screening. *J. Cheminf.* **2009**, DOI: 10.1186/1758-2946-1-6.
(45) Riva, G.; Mantovani, F. Being there: understanding the feeling of presence in a synthetic environment and its potential for clinical. In *Virtual Reality in Psychological, Medical and Pedagogical Applications*; Eichenberg, C., Ed.; InTech: New York, 2012; pp 3−34.
(46) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215* (3), 403−410.
(47) Rao, P. N. P.; Knaus, E. E. Evolution of Nonsteroidal Anti-Inflammatory Drugs (NSAIDs): cyclooxygenase (COX) inhibition and beyond. *J. Pharm. Pharm. Sci.* **2008**, *11* (2), 81s−110s.
(48) Krieger, E.; Nabuurs, S. B.; Vriend, G. Homology Modeling. In *Structural Bioinformatics*; Bourne, P. E., Weissig, H., Eds.; John Wiley & Sons: 2003; Vol. *44*, pp 509−524.
(49) Nayeem, A.; Sitkoff, D.; Krystek, S., Jr. A comparative study of available software for high-accuracy homology modeling: from sequence alignments to structural models. *Protein Sci.* **2006**, *15* (4), 808−824.
(50) Branden, C.; Tooze, J. *Introduction to Protein Structure*, 1st ed.; Garland: New York, 1991; pp 347−372.
(51) Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* **2003**, *31* (13), 3381−3385.
(52) Flower, D. R. *Bioinformatics for Vaccinology*, 1st ed.; John Wiley & Sons: Oxford, 2008; pp 217−256.
(53) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. ClustalW and ClustalX version 2. *Bioinformatics* **2007**, *23* (21), 2947−2948.
(54) *Sequence Alignment: Methods, Models, Concepts, and Strategies*, 1st ed.; Rosenberg, M. S., Ed.; University of California Press: 2009; pp 1−22.
(55) Gopalakrishnan, K.; Sowmiya, G.; Sheik, S. S.; Sekar, K. Ramachandran Plot on The Web (2.0). *Protein Pept. Lett.* **2007**, *14* (7), 669−671.
(56) Benkert, P.; Tosatto, S. C.; Schomburg, D. QMEAN: A comprehensive scoring function for model quality assessment. *Proteins* **2008**, *71* (1), 261−277.
(57) Benkert, P.; Biasini, M.; Schwede, T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **2011**, *27* (3), 343−350.
(58) Kleywegt, G. J.; Jones, T. A. Phi/psi-chology: Ramachandran revisited. *Structure* **1996**, *4* (12), 1395−1400.
(59) Span, E. A.; Goodsell, D. S.; Ramchandran, R.; Franzen, M.; Herman, T.; Sem, D. S. Protein structure in context: the molecular landscape of angiogenesis. *Biochem. Mol. Biol. Educ.* **2013**, *41* (4), 213−223.
(60) Satyanarayanajois, S. D. Active-learning exercises to teach drug-receptor interactions in a medicinal chemistry course. *Am. J. Pharm. Educ.* **2010**, *74* (8), 1−9.
(61) Zangaladze, A.; Epstein, C.; Grafton, S.; Sathian, K. Involvement of visual cortex in tactile discrimination of orientation. *Nature* **1999**, *401* (6753), 587−590.
(62) Tibell, L.; Ainsworth, S.; Persson, P. B.; Höst, G. Haptic influences on reasoning and learning in protein education. In *Proceedings of the 9th Nordic Research Symposium on Science Education*; Macdonald, A., Ed.; Science Education Research Group, School of Education, University of Iceland: Reykjavík, 2008; pp 165−168.
(63) McConkey, B. J.; Sobolev, V.; Edelman, M. The performance of current methods in ligand-protein docking. *Curr. Sci.* **2002**, *83* (7), 845−855.
(64) Joy, S.; Nair, P. S.; Hariharan, R.; Pillai, M. R. Detailed comparison of the protein-ligand docking efficiencies of GOLD, a commercial package and ArgusLab, a licensable freeware. *In Silico Biol.* **2006**, *6* (6), 601−605.
(65) Trott, O.; Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455−461.
(66) Magalhães, C. S.; Barbosa, C. H. S.; Almeida, D. M.; Dardenne, L. E. Improving differential evolution accuracy for flexible ligand docking using a multi-solution strategy. In *Intelligent Data Engineering and Automated Learning—IDEAL 2012*; Yin, H., Costa, J. A. F., Barreto, G., Eds.; Springer-Verlag: Berlin, 2012; pp 688−697.
(67) Arooj, M.; Kim, S.; Sakkiah, S.; Cao, G. P.; Lee, Y.; Lee, K. W. Molecular modeling study for inhibition mechanism of human chymase and its application in inhibitor design. *PLoS One* **2013**, *8* (4), e60470 DOI: 10.1371/journal.pone.0060470.
(68) Yusuf, D.; Davis, A. M.; Kleywegt, G. J.; Schmitt, S. An alternative method for the evaluation of docking performance: RSR vs RMSD. *J. Chem. Inf. Model.* **2008**, *48* (7), 1411−1422.
(69) Thomsen, R.; Christensen, M. H. MolDock: a new technique for high-accuracy molecular docking. *J. Med. Chem.* **2006**, *49* (11), 3315−3321.
(70) Meng, X. Y.; Zhang, H. X.; Mezei, M.; Cui, M. Molecular docking: a powerful approach for structure-based drug discovery. *Curr. Comput.-Aided Drug Des.* **2011**, *7* (2), 146−157.

DOI: 10.1021/ed500195d
J. Chem. Educ. 2015, 92, 827−835

(71) Brooijmans, N.; Kuntz, I. D. Molecular recognition and docking algorithms. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 335−373.
(72) Murphy, K. P. Predicting binding energetics from structure: looking beyond ΔG°. *Med. Res. Rev.* **1999**, *19* (4), 333−339.
(73) Wang, R.; Lai, L.; Wang, S. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *J. Comput.-Aided Mol. Des.* **2002**, *16* (1), 11−26.
(74) Van de Waterbeemd, H.; Gifford, E. ADMET in silico modelling: towards prediction paradise? *Nat. Rev. Drug Discovery* **2003**, *2* (3), 192−204.
(75) Sander, T. *OSIRIS Property Explorer*; Actelion Pharmaceuticals Ltd.: Allschwil, Switzerland, 2011.

DOI: 10.1021/ed500195d
J. Chem. Educ. 2015, 92, 827−835