JOURNAL OF CHEMICAL EDUCATION

Article

pubs.acs.org/jchemeduc

Drug Design Workshop: A Web-Based Educational Tool To Introduce Computer-Aided Drug Design to the General Public

Antoine Daina,[†,▽] Marie-Claude Blatter,[‡,§,∥,▽] Vivienne Baillie Gerritsen,[‡,§] Patricia M. Palagi,[‡,∥] Diana Marek,[‡,∥,⊥] Ioannis Xenarios,[§,⊥] Torsten Schwede,[#] Olivier Michielin,[†] and Vincent Zoete*[,†]

[†]Molecular Modeling Group, SIB Swiss Institute of Bioinformatics, Bâtiment Génopode, Quartier Sorge, 1015 Lausanne, Switzerland

[‡]Outreach Team, SIB Swiss Institute of Bioinformatics, Bâtiment Génopode, Quartier Sorge, 1015 Lausanne, Switzerland

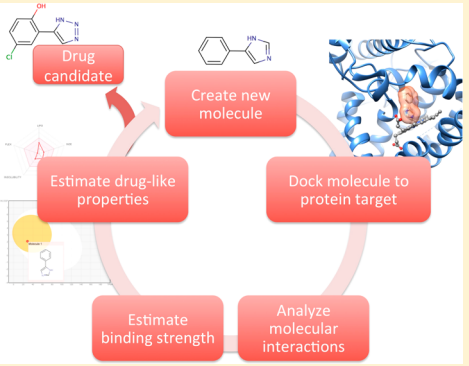
[§]Swiss-Prot Group, SIB Swiss Institute of Bioinformatics, CMU, 1 rue Michel Servet, 1211 Geneva 4, Switzerland

[∥]Training Group, SIB Swiss Institute of Bioinformatics, Bâtiment Génopode, Quartier Sorge, 1015 Lausanne, Switzerland

[⊥]Vital-IT, SIB Swiss Institute of Bioinformatics, Bâtiment Génopode, Quartier Sorge, 1015 Lausanne, Switzerland

[#]Computational Structural Biology, SIB Swiss Institute of Bioinformatics & Biozentrum, Universität Basel, Klingelbergstrasse 50/70, 4056 Basel, Switzerland

ABSTRACT: Due to its impact on society, the design of new drugs has the potential to interest a wide audience, and provides a rare opportunity to introduce several concepts in chemistry and biochemistry. Drug design can be seen as a multiobjective cyclic optimization process. Indeed, it is important to develop the understanding not only that a drug is generally an effective ligand for a protein of therapeutic interest, but also that these molecules need to have drug-like properties. Computer-aided drug design and bioinformatics approaches play a fundamental role in addressing these different challenges. Here we introduce a new freely available integrated web-based educational tool, Drug Design Workshop, which presents the basics of drug design and provides anyone with access to computational methods and resources to conceive and evaluate molecules for their potential to become actual drugs. We provide 3 examples of drug design targets for the discovery of common or state-of-the-art drugs, which can be used by educators to introduce easily the different concepts related to drug discovery: anti-inflammatory agents and drugs for immunotherapy or targeted cancer therapy. Since 2015, the workshop has been successfully given to over 1,500 people. The Web site was optimized on the basis of the positive and constructive comments from high-school teachers and students (15−19 years old).



KEYWORDS: *General Public, Biochemistry, Chemoinformatics, Public Understanding/Outreach, Interdisciplinary/Multidisciplinary, Internet/Web-Based Learning, Collaborative/Cooperative Learning, Drugs/Pharmaceuticals, Molecular Modeling, Molecular Recognition, Medicinal Chemistry, Physical Properties*

INTRODUCTION

The design and production of drugs is a field in which chemistry has had a favorable impact on life expectancy and quality over the past century.[1,2] As such, this field provides a rare opportunity to introduce several concepts in chemistry and biochemistry to a large audience.

It is widely known that the design and development of a new drug generally costs more than 1 billion dollars in total and takes at least 10 years,[3,4] while, despite all these efforts, only a very limited number of drug discovery projects will lead to the actual release of a new drug.[5,6] Several technologies have been developed to rationalize the process by reducing duration, cost, and attrition rate, one of which is computer-aided drug design (CADD).[7−10] CADD uses computing resources, algorithms, and 3D-visualization to help generate rational ideas about how to create or modify molecules, and to make decisions in the execution of the drug design process.

Whereas the general audience is aware of the overall concept and global cost of drug discovery and development, usually little is known about the actual challenges and the role played by CADD. To address this, we present a new freely available web-based educational tool, which introduces the basics of drug design and provides anyone with access to simple computational methodologies to conceive and evaluate molecules for their potential to become actual drugs.[11]

Although macromolecular entities, such as antibodies, can act as therapeutic agents, in this report we will consider that drugs are small organic molecules that activate or inhibit the function of a biomolecule, generally a protein, which in turn results in a therapeutic or prophylactic benefit to the patient.

Received: August 8, 2016
Revised: January 26, 2017
Published: February 27, 2017

DOI: 10.1021/acs.jchemed.6b00596
J. Chem. Educ. 2017, 94, 335−344

Nature has been the most important source of medicinal agents for centuries. Many useful drugs were developed from plant products, including for instance morphine from Poppy *Papaver* for pain management, quinine from *Cinchonae* tree's bark as an antimalarial drug and muscle relaxant, or paclitaxel (also known as taxol) from the Pacific yew tree *Taxus brevifolia* for cancer therapy, to name a few. While natural molecules are still a major source of inspiration for drug design, only 6% of the small-molecule drugs developed over the past decades are purely natural products, unmodified in structure. Other compounds comprise natural product derivatives containing synthetic modifications (27%), synthetic molecules inspired by natural products (32%), and brand new structure synthetic compounds (35%).[12] In other words, 94% of the newly released drugs have, at the very least, necessitated chemical modifications either to increase affinity and selectivity for the protein target, to correct absorption distribution metabolism or excretion (ADME) and toxicity problems, or to circumvent an intellectual property (IP) issue. Although serendipity has had an important role in many therapeutic advances, rational design including CADD has become a major factor in producing new treatments.[13] The vast majority of the drugs developed recently have benefited to various extents from computer-aided approaches as introduced below.[7]

BASIC PRINCIPLES OF COMPUTER-AIDED DRUG DESIGN

CADD technologies can be classified into two main categories: structure-based and ligand-based approaches.

Structure-based approaches make use of the three-dimensional structure of the protein target when it is available, or can be reliably modeled.[14] We generally consider that molecular docking is the cornerstone of structure-based drug design, of which anti-influenza drugs zanamivir (Relenza) and oseltamivir (Tamiflu) are among the most salient successful applications.[15,16]

The first and most basic objective in structure-based drug design is indeed to predict whether a given small molecule will bind to a chosen protein target and, if so, what will be the strength of this molecular recognition. The first goal can be achieved using a so-called docking program, whose aim is to predict the most probable geometry and position of a small molecule at the surface of a protein by optimizing the interactions between both molecular partners.[8] Many docking programs are freely available and can be used for educational purposes, including web-based tools such as SwissDock.ch,[17] or downloadable programs such as Autodock[18] and Autodock Vina.[19] The concept of molecular docking is very intuitive and can be easily introduced to the general audience. For example, we obtained good results using scaled 3D-printed models of cyclooxygenase (COX) and ibuprofen, a well-known anti-inflammatory drug that binds to COX. The COX model can be opened, exposing the binding site that accommodates ibuprofen. Then, students (even youngsters) are invited to manually perform the docking of ibuprofen into COX (Figure 1). Although this manual positioning neglects the difficulties encountered when accounting for the flexibility of both the protein and the ligand, it shows the challenges of the process, and the necessity to automatize the approach using computer algorithms to possibly treat a large number of molecules to dock.

The second goal, i.e., determining the strength of binding of the small molecule to the protein, can be achieved using a binding free energy estimator. Several computer-aided approaches are available for this purpose.[8] They are generally based on high-level methods involving concepts in physical chemistry and statistical physics. However, this theoretical complexity can be hidden behind the simple notions of fitness or scoring. Docking software usually provides a crude estimation of this binding free energy, which can be presented simply to users as a score (without physical or chemical meanings) to optimize.

Figure 1. As a clear example of the molecular docking concept, educators can let participants dock a drug in the binding site of the protein target manually. Both the drug and the protein must be printed in 3D at the same scale. Here, we selected an example related to the first online workshop: an anti-inflammatory drug (ibuprofen in yellow) to be placed inside the cyclooxygenase 1 protein (COX1 in white). The printed protein model can be opened, exposing the ibuprofen binding site. The protein file, retrieved from the protein databank (PDB),[43] can be converted into a 3D-object and saved as STL, VMRL, or X3D files compatible with most 3D-printers, by using a molecular visualization software such as UCSF Chimera.[20]

Basically, drug design consists of the conception of molecules that are complementary to the protein target in terms of 3D-shape and charge distribution, to optimize molecular recognition and binding. Through prediction of molecular recognition and binding affinity, molecular docking opens the road to *in silico* design and optimization of virtual compounds.

On the contrary, ligand-based approaches rely on the knowledge implicitly contained in the chemical structure or physical properties of other molecules that bind to the biological target of interest. Typically, this knowledge can be extracted, analyzed, and used to create predictive models using machine-learning technologies, under the name quantitative structure−activity relationships (QSAR) if the objective is to create new ligands and/or predict their activity, or quantitative structure−properties relationships (QSPR) if the objective is to

DOI: 10.1021/acs.jchemed.6b00596
J. Chem. Educ. 2017, 94, 335−344

predict molecular properties in relation to lipophilicity,[21] drug likeness, or pharmacokinetics[22] (PK), for example. These molecular properties are fundamental in drug design. Indeed, although a high affinity for the protein target is essential, it is not sufficient for the designed small molecule to become a drug: to obtain a therapeutic effect, the molecule needs to reach its target in the body, and stay there long enough for the expected biological events to occur. Therefore, to support efficiently the design of new drugs, it is important to predict their PK behaviors with computer-aided approaches.

In addition to QSAR and QSPR, another set of ligand-based approaches rely on the commonly accepted assumption that very similar small molecules are more likely to be active on the same target. Such approaches can be used to perform molecular screening, i.e., searching for molecules similar to known active compounds and potentially also active on the same target, or reverse screening, i.e., deducing the potential protein targets of a given molecule by identifying similar existing compounds for which the activity is experimentally known.[23−25] This reverse screening can be of high interest to predict potential secondary targets of small molecules, i.e., proteins to which a small molecule will be able to bind although it was developed to target another macromolecule. These secondary targets can be at the origin of the negative side effects of small molecules, but on the contrary can also open the way to positive drug repurposing, i.e., finding another possible application to an existing drug.[26,27]

PEDAGOGICAL OBJECTIVES

Approaches, methodologies, and technologies involved in computational chemistry, chemoinformatics, bioinformatics, and molecular modeling, and generally in CADD, have a wide scope of application, but their teaching remains limited, even at an advanced academic level.[28] Several remarkable educational protocols have been proposed to achieve the objective of teaching how to properly perform drug discovery tasks with existing computational tools to future professionals in pharmaceutical research. Of note is the well-structured, thorough course by Tsai,[29] which includes different modules, lectures, and practical sessions encompassing many facets of CADD. More recently, Rodrigues et al. provided a comprehensive technical course of drug design.[30] Other educational programs emphasis more on virtual screening[31] or ADME[32] aspects. Moreover, some studies have demonstrated the positive impact of using tridimensional molecular graphics visualization for the perception of complex molecular properties.[33,34] All of these excellent teaching materials imply multiple methodologies based on a combination of web and standalone software. Whereas this has the merit to make the student face the technical hurdles of the discipline (e.g., incompatibility of file formats, irreproducibility of implementations, instability of multiple computer platforms), we believe this can prevent reaching the pedagogical goals of teaching the global concepts of the drug design process, especially at the high-school level. Not surprisingly, the above-mentioned courses and sessions are merely dedicated to upper-level undergraduate students. A noticeable exception is the e-malaria project, which was used to introduce high-school students to drug design in a real-life context.[35] Unfortunately, this latter endeavor faced licensing and confidentiality issues that required a complex hardware and network infrastructure along with an account login procedure. Together with significant computational time, this lack of flexibility hinders the trial-and-error cyclic process, which we consider key to appreciating the basis of molecular design.

By leveraging our expertise in designing expert CADD services,[17,24,36−39] we took advantage of today's opportunities provided by web technologies and open-source resources to develop the fully integrated, flexible educational tools described below for a broader audience, including high-school students, high-school teachers, undergraduate students, and the public at large.[11] This web-based teaching environment reduces technical difficulties to the minimum, allowing several pedagogical objectives to be reached.

First, it is useful to remind or inform the general audience that, in conventional medicine, a drug is a small molecule, most of the time synthesized by organic chemistry, which interacts with a biological macromolecule to generate the therapeutic effect. This concept is the essence of drug design and must be understood at the beginning of the Drug Design Workshop. In contrast, it can also serve as a starting point for a discussion regarding the differences between conventional (also known as allopathic) and homeopathic medicine.

Second, we consider it important to state that the design of new drugs is a collective effort necessitating the close collaboration of several different scientific backgrounds including not only biology, medicine, pharmacy, and biochemistry, but also chemistry, molecular modeling, and bioinformatics.

Third, we would like to introduce the key concept that drug design is a multiobjective optimization process whose aim is to create compounds with not only a high affinity for the target but also optimal pharmacokinetics properties. This requires access to professional scientific web-based tools and necessitates the guidance of an expert in the field or an educator trained in these specific topics.

A fourth objective is to explain that computer-aided drug design activities consist of the usage of several diverse structure-based and ligand-based approaches, to predict and evaluate all fundamental characteristics necessary for a molecule to become a drug: e.g., complementarity and affinity for the target, fate inside the organism, along with possible side effects and toxicity.

Finally, we would like to state that drug design is generally a cyclic optimization process, which gathers knowledge obtained from the first molecules in order to design better compounds in the next rounds and converge toward drug-candidates. From our experience and from teacher feedback, high-school students are very smoothly engaged with and attracted by the challenged of iteratively generating a molecule with the highest score on a given target. Positive emulation and competition in classrooms were frequently observed and reported.

DESIGN OF THE WORKSHOP AND NEW EDUCATIONAL TOOLS

Short Movie Describing the Role of Bioinformatics in Drug Design

First, with the help of professional graphic designers from Studio KO,[40] we produced a short movie to introduce the concepts mentioned in the pedagogical objectives, but also to recall the nature and definition of a protein. The movie also links diseases to possible over- or underexpression of proteins, or to protein mutations leading to malfunction. This allows the introduction of the notion that a drug is generally a small molecule able to bind such proteins and lead to the therapeutic

DOI: 10.1021/acs.jchemed.6b00596
J. Chem. Educ. 2017, 94, 335−344

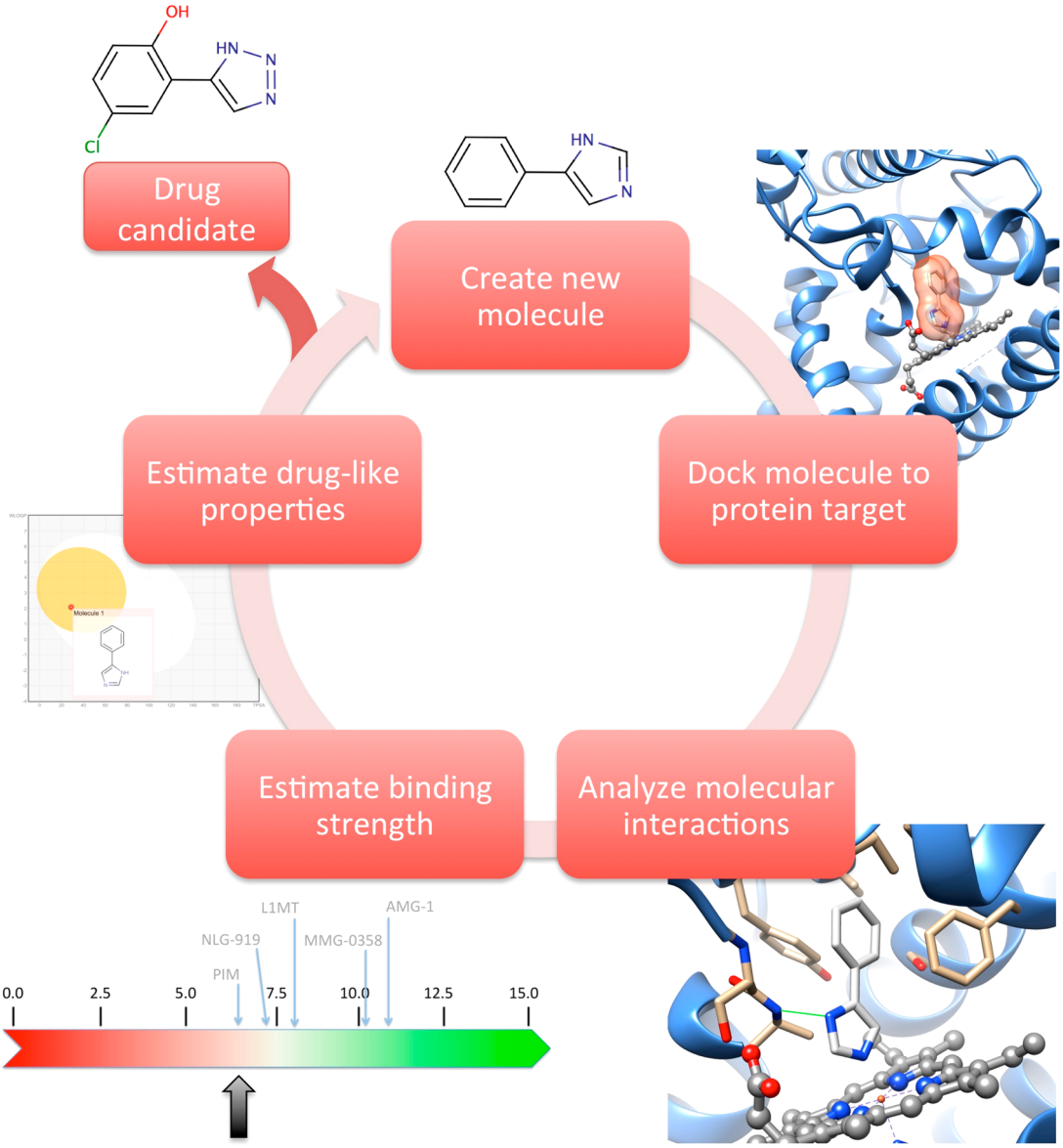


Figure 2. General principle of the online Drug Design Workshop exemplified in the context of the inhibition of indolamine 2,3-dioxygenase (IDO1) by the optimization of a newly discovered inhibitor (PIM) to obtain a drug candidate (MMG-0358). Several cycles of optimization can be performed, during which the molecules are drawn by the users, automatically docked into the protein, scored for molecular complementarity, and analyzed for some ADMET properties and possible secondary targets. All technical aspects have been simplified and can be performed by one-click or drag-and-drop actions. NLG-919, L1MT, and AMG-1 are other known ligands of IDO1 used as examples in the workshop and defined in the Web site.

effect. This movie is available online at the main Drug Design Workshop URL.[11]

Online Drug Design Workshop: Selecting Didactic Drug Design Targets

Second, we created a simple and integrated web interface to perform the basic steps of CADD. As for real drug design, this online tool allows for performance of multiple iterative cycles of molecular optimization, taking into account the complementarity of the designed molecule for the target. In a second step, different properties regarding ADME, toxicity, and secondary targets (Figure 2) may be considered.

For this, we have selected three relevant protein targets for drug design: the cyclooxygenase (two isoforms: COX1 and COX2), B-Raf, and indoleamine 2,3-dioxygenase 1 (IDO1).

COX1/2 are the targets of the very well-known nonsteroidal anti-inflammatory drugs (NSAID) like ibuprofen or diclofenac, which are commonly used to treat inflammation, pain, and fever. COX1/2 can be used to make a link between the concepts introduced during the workshop and a medication that everyone has already used. COX1/2 is also a good model to introduce the notion of selectivity for the target. Indeed, COX1 plays an important role in blood coagulation and in protecting the gastric lining, while COX2 is produced locally in the inflamed tissue, and is directly responsible for the sensation of pain. Recent efforts led to the design of ligands specific for COX2, which once targeted becomes responsible for the therapeutic effect, thus avoiding COX1 which is responsible for the side effects of the classical NSAIDs. Users are invited to design selective ligands by following the example of celecoxib, a selective COX2 inhibitor.

DOI: 10.1021/acs.jchemed.6b00596
J. Chem. Educ. 2017, 94, 335−344

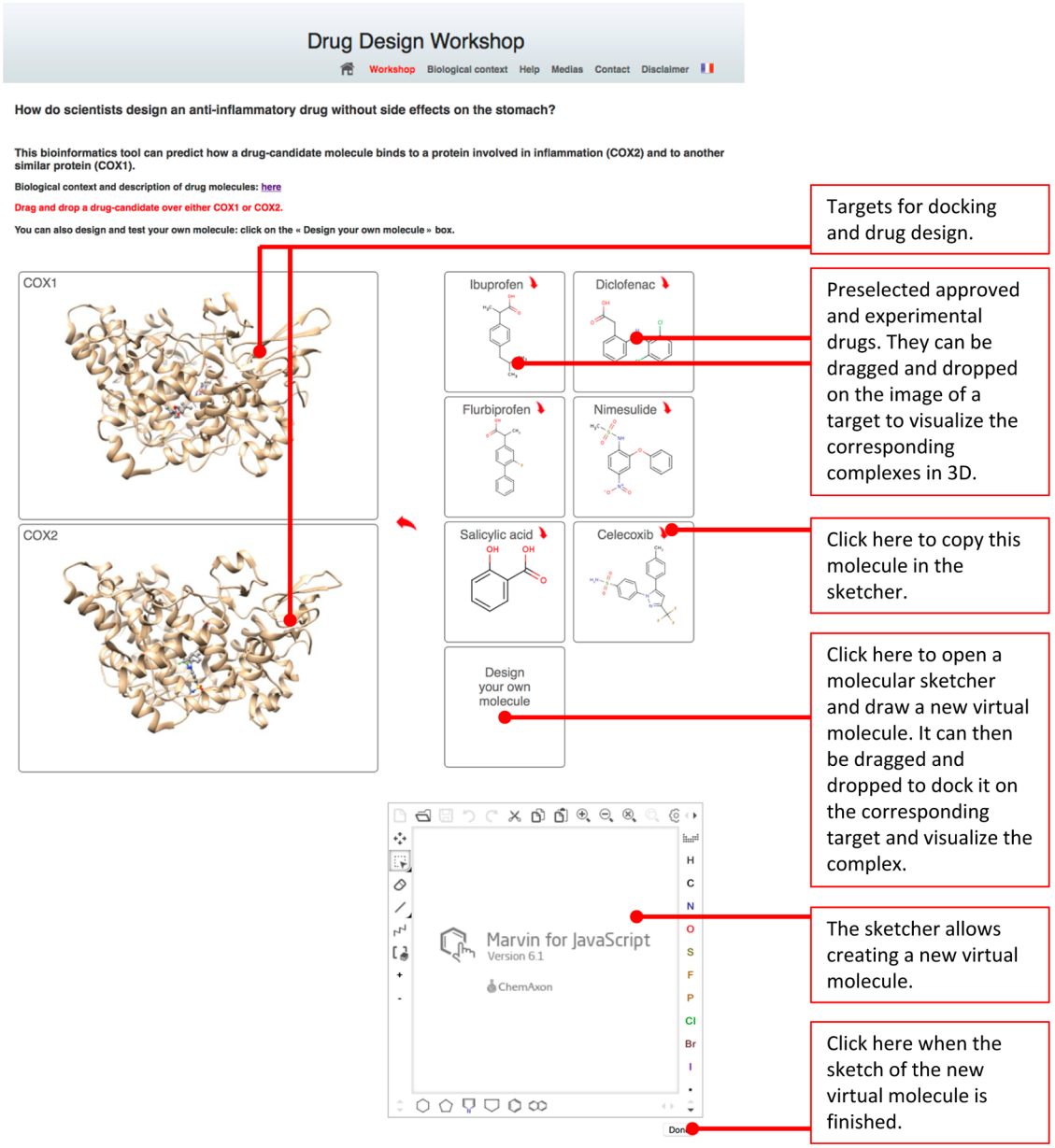


Figure 3. Input page of the Drug Design Workshop Web site.

B-Raf is a kinase whose mutants commonly cause cancer by excessive stimulation of cell growth. Inhibitors of the V600E B-Raf mutant, a form often found in melanoma cells, were recently introduced for the treatment of late-stage melanoma. Molecules such as vemurafenib, a specific inhibitor of V600E B-Raf, were among the first drugs to trigger an efficient response against this type of skin cancer. This target protein thus provides an example of a recent success story of drug design in the targeted therapy of cancer. Possibly, it can also be used to open a discussion on personalized medicine. Indeed, in case of melanoma cells not bearing the V600E mutation of B-Raf, vemurafenib was proven to be deleterious as the drug favors tumor growth.[41] Therefore, its prescription can only be made upon sequencing the BRAF gene of the patient's cancer cells to ascertain the presence of this sequence alteration.

IDO1 is an enzyme that catabolizes tryptophan, and is used by cancer cells to shun the immune system. Therefore, inhibitors of IDO1 could be of major interest for cancer immunotherapy,[42] and evaluated for coadministration with other agents that inhibit immune escape of cancer cells (e.g., monoclonal antibodies ipilimumab or nivolumab). IDO1 is thus an elegant opportunity to delineate the relationship between CADD and the latest state-of-the-art discoveries in cancer treatment.

The biological contexts corresponding to the above-mentioned targets are introduced and summarized online in our Web site.[11]

Online Drug Design Workshop: Experiencing the Drug Design Process

For each target, we selected representative well-known approved or experimental drugs, whose binding modes are available in the Protein Databank,[43] or were precalculated using the Autodock Vina docking program.[19] Figure 3 shows the input page of the Drug Design Workshop Web site. Images representing the 3D-structure of target proteins are displayed on the left, and the 2D chemical structures of typical drugs are

DOI: 10.1021/acs.jchemed.6b00596
J. Chem. Educ. 2017, 94, 335−344

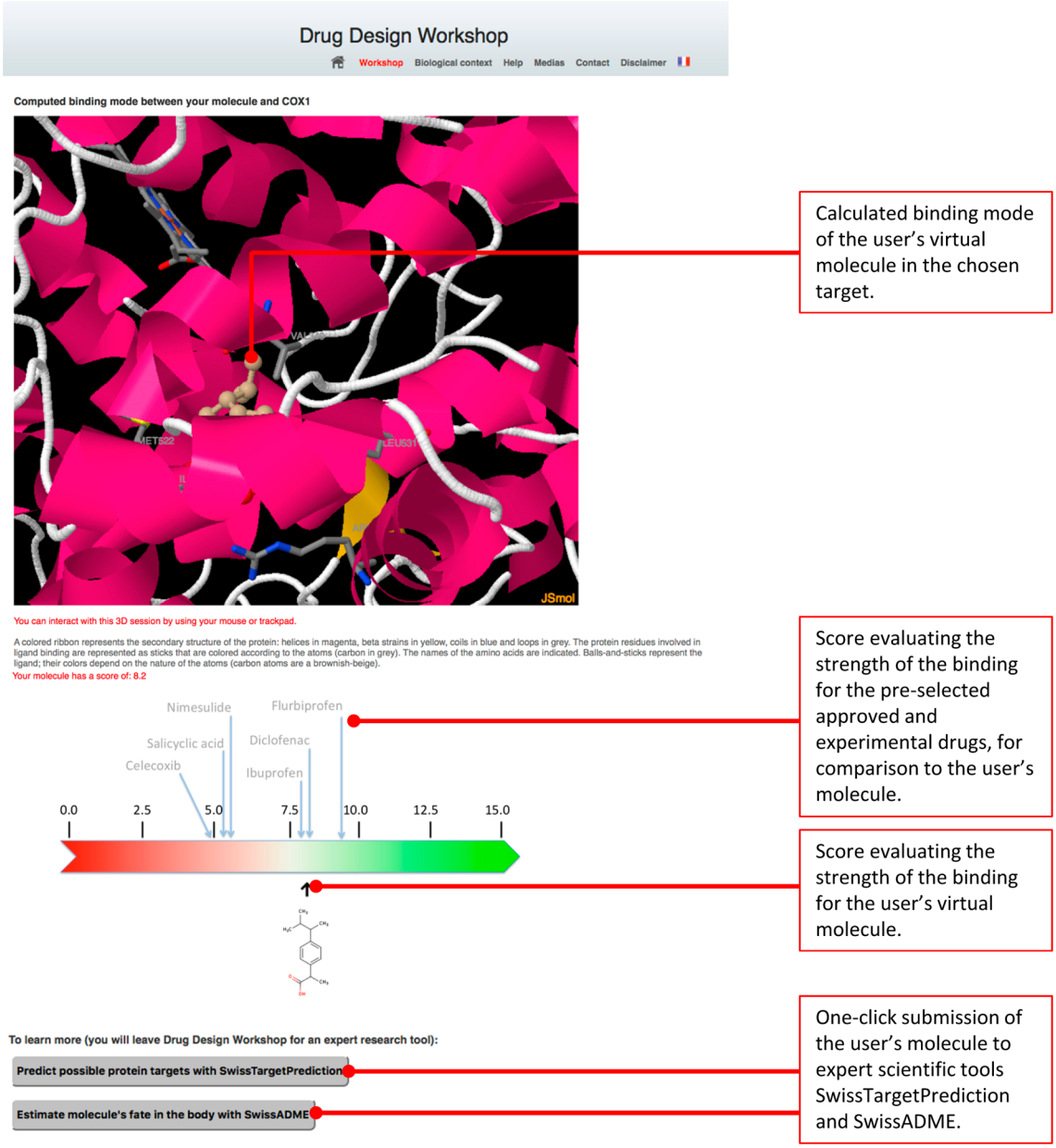


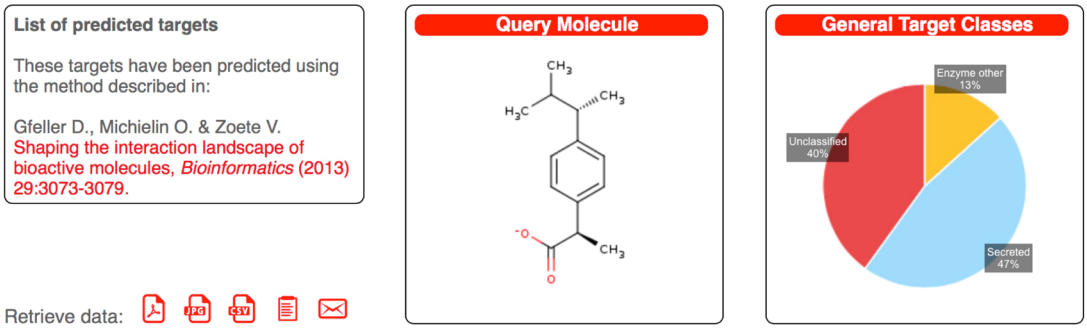
Figure 4. Output page of the Drug Design Workshop Web site.

available in boxes, on the right. To visualize the complex between the protein target and one preselected drug, the user simply needs to drag the drug image and drop it on the protein image. The corresponding complex will be immediately displayed as an interactive 3D-session in the user's web browser thanks to the JSmol molecular visualization applet.[44]

To design a molecule, the user is invited to click on the "Design your own molecule" box. This will open the Marvin4JS molecular sketcher[45] to draw a new *virtual* molecule. To simplify the process for users with little experience in organic chemistry, the sketcher can be filled automatically with one of the preselected drugs, by clicking on the corresponding "down" red arrow. Then, the user can modify these molecules within the sketcher. Thanks to this simplified process and the ability of the sketcher to indicate inconsistencies in chemical structures, we have experienced that even users without any knowledge of organic chemistry are capable of drawing relevant molecules.

Once the new molecule has been drawn and the "Done" button clicked, its image will appear on the corresponding box, and it is available for drag-and-drop. If the user requires the visualization of the complex between a target and a designed ligand, the molecule is automatically docked with Autodock Vina, after determination of the most probable microspecies (protonation state and tautomer) at physiological conditions. For the sake of simplicity, these steps are performed without any intervention from the user. Docking calculations can be time-consuming. Therefore, to allow the workshop to be executed on any computer, calculations are not performed on the user's workstation (desktop or laptop), but on one of our multicore machines, managed by a queuing system that allows several docking calculations in parallel. Users are informed about the waiting list and duration of the calculation. In our practice, this setup allows up to 15 users to perform basic operation of drug design at the same time and in good conditions.

DOI: 10.1021/acs.jchemed.6b00596
J. Chem. Educ. 2017, 94, 335−344



Target	Common name	Uniprot ID	ChEMBL ID	Probability*	# sim. cmpds (3D / 2D)	Target Class
Prostaglandin G/H synthase 1	PTGS1	P23219	CHEMBL221		68 / 69	Enzyme
Prostaglandin G/H synthase 2	PTGS2	P35354	CHEMBL230		68 / 69	Enzyme
IL-8(8-77)	IL8	P10145	CHEMBL2157		13 / 8	Secreted
Muscleblind-like protein 1	MBNL1	Q9NR56	CHEMBL1293317		22 / 4	Unclassified
Solute carrier family 22 member 6 *(by homology)*	SLC22A6	Q4U2R8	CHEMBL1641347		2 / 1	Unclassified
Solute carrier family 22 member 10 *(by homology)*	SLC22A10	Q63ZE4			2 / 1	Unclassified
Solute carrier family 22 member 25 *(by homology)*	SLC22A25	Q6T423			2 / 1	Unclassified
Solute carrier family 22 member 9 *(by homology)*	SLC22A9	Q8IVM8	CHEMBL2073721		2 / 1	Unclassified
Solute carrier family 22 member 24 *(by homology)*	SLC22A24	Q8N4F4			2 / 1	Unclassified
Serum albumin *(by homology)*	ALB	P02768	CHEMBL3253		1 / 1	Secreted
Alpha-fetoprotein *(by homology)*	AFP	P02771			1 / 1	Secreted
Connective tissue-activating peptide III(1-81) *(by homology)*	PPBP	P02775			13 / 8	Secreted
GRO-alpha(4-73) *(by homology)*	CXCL1	P09341			13 / 8	Secreted
GRO-beta(5-73) *(by homology)*	CXCL2	P19875			13 / 8	Secreted
GRO-gamma(5-73) *(by homology)*	CXCL3	P19876			13 / 8	Secreted

* Probabilities have been computed based on a cross-validation. They may therefore not represent the actual probability of success for any new molecule (see FAQ).

Figure 5. Output page of SwissTargetPrediction, obtained upon one-click query from the Drug Design Workshop Web site. Target names, common names, Uniprot ID, ChEMBL ID, and Target classes are those defined in the ChEMBL database,[46] which was used to build the predictive model of protein targets for small molecules.[47]

Binding modes of the existing or virtual compounds in the protein target are displayed on a dedicated page, which also provides a score that evaluates the strength of the binding (Figure 4). This score is in fact the opposite of the binding free energy estimated by the Autodock Vina docking software. We chose to use this score rather than the actual binding free energy since it is easier for the user to follow the idea that "the larger the score, the better the ligand" rather than the more confusing notion that "the more negative the binding free energy, the better the ligand". Of course, this modification can be explained to more advanced students. For each compound, the calculated score of the designed molecule is compared to those of the preselected drugs, allowing the user to compete with "real drugs" in terms of affinity for the therapeutic target. We have experienced that this score creates a powerful incentive for the users to create better molecules, and thus to enter naturally and seamlessly into the typical iterative optimization cycle, which is one of the fundamental processes of CADD. It is noteworthy that the docking engine used involves a stochastic algorithm, which is necessary for having docking results quickly enough for true interactivity (approximately between 30 s to 3 min, depending on the size of molecule and binding site). Whereas reproducibility cannot be ensured, we have set parameters to maximize convergence. As a result, docking with Drug Design Workshop returned a significantly different binding mode in only 5−10% of the runs (related mainly to molecule flexibility and binding site size). To gain confidence in the predicted binding mode, it is advised to run the same docking several times or in some cases compare the results obtained by each student in the classroom.

The usage of the Web site is supported by help pages and FAQs providing technical guidance.

Online Drug Design Workshop: Introducing the Multiobjective Character of Drug Design

The above-mentioned cyclic optimization process is limited to the enhancement of the affinity of the ligand for the protein by using a structure-based approach. As we discussed above, one pedagogical objective is also to introduce the multiobjective nature of the optimization process in (computer-assisted) drug design. This implies that, besides affinity, the pharmacokinetic and the pharmacodynamic properties of the small molecule are also to be optimized. To this end, we enable a seamless one-click submission of the user's molecule from the Drug Design Workshop to SwissTargetPrediction[24] or to SwissADME.[22] Both these online tools are in-house research-grade web services developed to predict possible targets, ADME, or

DOI: 10.1021/acs.jchemed.6b00596
J. Chem. Educ. 2017, 94, 335−344

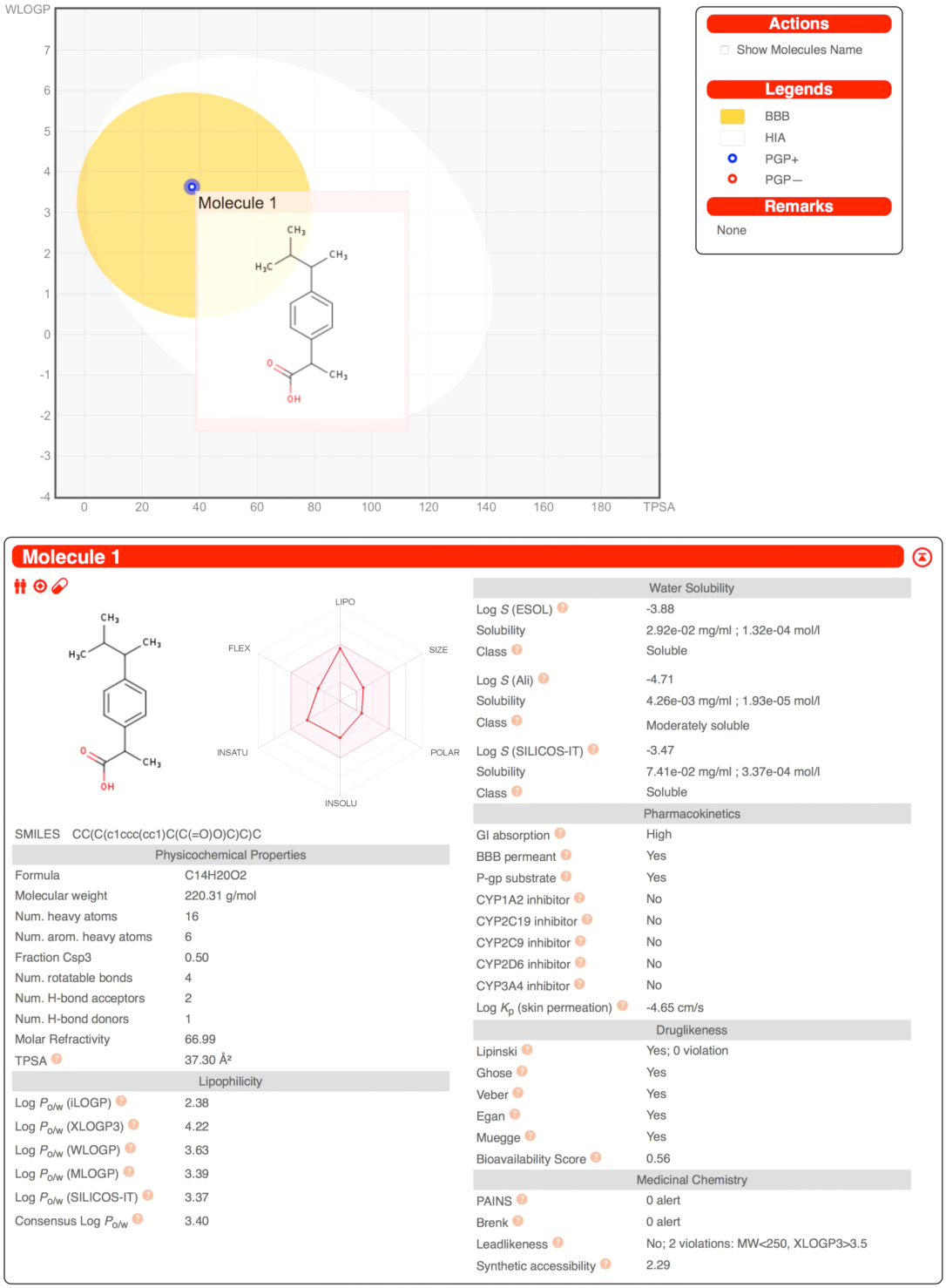


Figure 6. Output page of SwissADME, obtained upon one-click query from the Drug Design Workshop Web site. The upper panel shows the BOILED-Egg, a graphical classification model to predict gastrointestinal absorption (HIA, white ellipse) and permeation through the blood−brain barrier (BBB, yolk).[22] The position of the molecule on this panel is shown as a dot, whose color reflects the prediction for the molecule to be the substrate of the multidrug resistance protein "P-glycoprotein 1" (PGP). The lower panel compiles all predicted ADME parameters for the molecule under study.[48]

toxicity properties of small drug-like molecules. Both rely on ligand-based approaches, allowing an introduction to this type of technology for the most advanced users. SwissTargetPrediction provides a list of the 15 most probable protein targets for the small molecule under consideration, giving an estimate of the selectivity of the molecule and a prediction of potential side effects (Figure 5). SwissADME calculates numerous molecular properties related to, e.g., pharmacokinetics, drug-likeness, physicochemistry, and synthetic accessibility (Figure 6). Of note, SwissADME models include the BOILED-Egg that we developed to predict the propensity of small molecules to be absorbed by the gastrointestinal tract or to access the brain.

DOI: 10.1021/acs.jchemed.6b00596
J. Chem. Educ. 2017, 94, 335−344

These predictions are made simply by plotting the small molecule on a 2D graph based on the lipophilicity and polarity of molecules, where the regions containing molecules able to cross specific biological barriers (gastrointestinal wall or blood-brain barrier) are delineated by ellipses (Figure 6). Thanks to simplicity and speed, the BOILED-Egg model is of great support for the users to apprehend the concepts of absorption and distribution, and to figure out what type of chemical modifications must be made to the small molecule to obtain the desired absorption and distribution, in an intuitive and iterative way.

The results from SwissTargetPredicition and SwissADME, notably the BOILED-Egg model, can be useful for Drug Design Workshop users through the guidance of experts in the field or educators previously trained in the topic. This information can be fruitfully taken into account during the global optimization process of one's own molecule, providing a better overview of the multiobjective character of CADD.

DISCUSSION

During the last two years, approximately 900 high-school students, 15−19 years old, attended this computer-aided drug design workshop, in about 50 different sessions. We proposed an anonymous online feedback form to the students with the aim of regular improvement. Student feedback was very satisfactory, with an overall appraisal of 5.11/6.00 based on 209 evaluations by pairs of users. Of note, students particularly appreciated the opportunity to use "professional" bioinformatics tools (5.30/6.00) and said they had learned a lot about drug design and CADD during the session (5.23/6.00). In view of these encouraging experiences, we decided to provide training to teachers who could use this material as a support for biology or chemistry classes.

In our experience, thanks to the simplicity of the user-friendly Web site and molecular sketcher, but also the incentive provided by the affinity score, the younger users also appreciate the workshop, even if they have little or no experience in organic chemistry. Due to the versatility of the approach, which provides an opportunity to introduce a large number of different concepts of CADD in relation to the users' backgrounds, we also successfully used this workshop as a rapid and simple hands-on introduction to the cyclic iterative optimization process in drug design for students from the doctoral school of pharmaceutical sciences from the University of Geneva or to bachelor and master students in biochemistry from the University of Fribourg, Switzerland.

Since the workshop requires very little material, i.e., a few standard computers connected to the Internet, it is easy to give it not only directly in schools but also during scientific exhibitions: hundreds of visitors have had the opportunity to successfully experience the early stages of designing drugs during various science fairs. In addition, the workshop has also been implemented into public laboratories and educational platforms, such as the Chimiscope of the University of Geneva[49] and l'Eprouvette of the University of Lausanne,[50] Switzerland.

Although high-school students and teachers are certainly our main target, those who have participated in our workshop, from families to doctoral students, have always shown great interest and enthusiasm whatever their scientific background and level. The subject not only is timely, but also concerns each and every one. Our workshop provides a simplified view of complex notions and allows a wide audience to discover the key stages in drug discovery as well as the importance of bioinformatics in life science today.

AUTHOR INFORMATION

Corresponding Author

*E-mail: Vincent.Zoete@sib.swiss.

ORCID

Vincent Zoete: 0000-0002-2336-6537

Author Contributions

[∇]A.D. and M.-C.B. authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

A short movie introducing the concepts mentioned in the pedagogical objectives, and recalling the nature and definition of a protein, is available in ref 11 under CC-BY-ND-NC license.

ACKNOWLEDGMENTS

This work was supported by the Swiss National Science Foundation through the Agora grant CRAGP3_151515, and by the SIB Swiss Institute of Bioinformatics.

REFERENCES

(1) Bunker, J. P. The Role of Medical Care in Contributing to Health Improvements Within Societies. *Int. J. Epidemiol* **2001**, *30* (6), 1260−1263.

(2) Lichtenberg, F. R. The Impact of New Drug Launches on Longevity: Evidence From Longitudinal, Disease-Level Data From 52 Countries, 1982−2001. *Int. J. Health Care Finance Econ* **2005**, *5* (1), 47−73.

(3) Munos, B. Lessons From 60 Years of Pharmaceutical Innovation. *Nat. Rev. Drug Discovery* **2009**, *8* (12), 959−968.

(4) Paul, S. M.; Mytelka, D. S.; Dunwiddie, C. T.; Persinger, C. C.; Munos, B. H.; Lindborg, S. R.; Schacht, A. L. How to Improve R&D Productivity: the Pharmaceutical Industry's Grand Challenge. *Nat. Rev. Drug Discovery* **2010**, *9* (3), 203−214.

(5) Arrowsmith, J.; Miller, P. Trial Watch: Phase II and Phase III Attrition Rates 2011−2012. *Nat. Rev. Drug Discovery* **2013**, *12*, 569−569.

(6) Waring, M. J.; Arrowsmith, J.; Leach, A. R.; Leeson, P. D.; Mandrell, S.; Owen, R. M.; Pairaudeau, G.; Pennie, W. D.; Pickett, S. D.; Wang, J.; Wallace, O.; Weir, A. An Analysis of the Attrition of Drug Candidates From Four Major Pharmaceutical Companies. *Nat. Rev. Drug Discovery* **2015**, *14* (7), 475−486.

(7) Jorgensen, W. L. *Science* **2004**, *303* (5665), 1813−1818.

(8) Zoete, V.; Grosdidier, A.; Michielin, O. Docking, Virtual High Throughput Screening and in Silico Fragment-Based Drug Design. *J. Cell. Mol. Med.* **2009**, *13* (2), 238−248.

(9) Schneider, G. From Theory to Bench Experiment by Computer-Assisted Drug Design. *Chimia* **2012**, *66* (3), 120−124.

(10) Woltosz, W. S. If We Designed Airplanes Like We Design Drugs···. *J. Comput.-Aided Mol. Des.* **2012**, *26* (1), 159−163.

(11) Drug Design Workshop. http://www.drug-design-workshop.ch (accessed Dec 2016).

(12) Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs From 1981 to 2014. *J. Nat. Prod.* **2016**, *79* (3), 629−661.

(13) Seddon, G.; Lounnas, V.; McGuire, R.; van den Bergh, T.; Bywater, R. P.; Oliveira, L.; Vriend, G. *J. Comput.-Aided Mol. Des.* **2012**, *26* (1), 137−150.

(14) Schmidt, T.; Bergner, A.; Schwede, T. Modelling Three-Dimensional Protein Structures for Applications in Drug Design. *Drug Discovery Today* **2014**, *19* (7), 890−897.

(15) von Itzstein, M.; Wu, W. Y.; Kok, G. B.; Pegg, M. S.; Dyason, J. C.; Jin, B.; Van Phan, T.; Smythe, M. L.; White, H. F.; Oliver, S. W.; et al. Rational Design of Potent Sialidase-Based Inhibitors of Influenza Virus Replication. *Nature* **1993**, *363* (6428), 418−423.

DOI: 10.1021/acs.jchemed.6b00596
J. Chem. Educ. 2017, 94, 335−344

(16) Kim, C. U.; Lew, W.; Williams, M. A.; Liu, H.; Zhang, L.; Swaminathan, S.; Bischofberger, N.; Chen, M. S.; Mendel, D. B.; Tai, C. Y.; Laver, W. G.; Stevens, R. C. Influenza Neuraminidase Inhibitors Possessing a Novel Hydrophobic Interaction in the Enzyme Active Site: Design, Synthesis, and Structural Analysis of Carbocyclic Sialic Acid Analogues with Potent Anti-Influenza Activity. *J. Am. Chem. Soc.* **1997**, *119* (4), 681−690.
(17) Grosdidier, A.; Zoete, V.; Michielin, O. SwissDock, a Protein-Small Molecule Docking Web Service Based on EADock DSS. *Nucleic Acids Res.* **2011**, *39*, W270−W277.
(18) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30* (16), 2785−2791.
(19) 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.
(20) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera–a Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25* (13), 1605−1612.
(21) Daina, A.; Michielin, O.; Zoete, V. iLOGP: a Simple, Robust, and Efficient Description of N-Octanol/Water Partition Coefficient for Drug Design Using the GB/SA Approach. *J. Chem. Inf. Model.* **2014**, *54* (12), 3284−3301.
(22) Daina, A.; Zoete, V. A BOILED-Egg to Predict Gastrointestinal Absorption and Brain Penetration of Small Molecules. *ChemMedChem* **2016**, *11* (11), 1117−1121.
(23) Gfeller, D.; Michielin, O.; Zoete, V. Shaping the Interaction Landscape of Bioactive Molecules. *Bioinformatics* **2013**, *29* (23), 3073−3079.
(24) Gfeller, D.; Grosdidier, A.; Wirth, M.; Daina, A.; Michielin, O.; Zoete, V. SwissTargetPrediction: a Web Server for Target Prediction of Bioactive Small Molecules. *Nucleic Acids Res.* **2014**, *42* (W1), W32−W38.
(25) Gfeller, D.; Zoete, V. Protein Homology Reveals New Targets for Bioactive Small Molecules. *Bioinformatics* **2015**, *31* (16), 2721−2727.
(26) Oprea, T. I.; Bauman, J. E.; Bologa, C. G.; Buranda, T.; Chigaev, A.; Edwards, B. S.; Jarvik, J. W.; Gresham, H. D.; Haynes, M. K.; Hjelle, B.; Hromas, R.; Hudson, L.; Mackenzie, D. A.; Muller, C. Y.; Reed, J. C.; Simons, P. C.; Smagley, Y.; Strouse, J.; Surviladze, Z.; Thompson, T.; Ursu, O.; Waller, A.; Wandinger-Ness, A.; Winter, S. S.; Wu, Y.; Young, S. M.; Larson, R. S.; Willman, C.; Sklar, L. A. Drug Repurposing From an Academic Perspective. *Drug Discovery Today: Ther. Strategies* **2011**, *8* (3−4), 61−69.
(27) Bertolini, F.; Sukhatme, V. P.; Bouche, G. Drug Repurposing in Oncology−Patient and Health Systems Opportunities. *Nat. Rev. Clin. Oncol.* **2015**, *12* (12), 732−742.
(28) Wild, D. J. Cheminformatics for the Masses: a Chance to Increase Educational Opportunities for the Next Generation of Cheminformaticians. *J. Cheminf.* **2013**, *5* (1), 32.
(29) Tsai, C. S. Using Computer Applications and Online Resources to Teach and Learn Pharmaceutical Chemistry. *J. Chem. Educ.* **2007**, *84* (12), 2019.
(30) Rodrigues, R. P.; Andrade, S. F.; Mantoani, S. P.; Eifler-Lima, V. L.; Silva, V. B.; Kawano, D. F. Using Free Computational Resources to Illustrate the Drug Design Process in an Undergraduate Medicinal Chemistry Course. *J. Chem. Educ.* **2015**, *92* (5), 827−835.
(31) 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.
(32) 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.
(33) Carvalho, I.; Borges, Á. 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.
(34) 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.
(35) 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.
(36) Zoete, V.; Cuendet, M.A.; Grosdider, A.; Michielin, O. SwissParam: A Fast Force Field Generation Tool for Small Organic Moleules. *J. Comput. Chem.* **2011**, *32* (11), 2359−2368.
(37) Gfeller, D.; Michielin, O.; Zoete, V. SwissSidechain: a Molecular and Structural Database of Non-Natural Sidechains. *Nucleic Acids Res.* **2013**, *41* (D1), D327−D332.
(38) Wirth, M.; Zoete, V.; Michielin, O.; Sauer, W. H. B. SwissBioisostere: a Database of Molecular Replacements for Ligand Design. *Nucleic Acids Res.* **2013**, *41* (D1), D1137−D1143.
(39) Zoete, V.; Daina, A.; Bovigny, C.; Michielin, O. SwissSimilarity: a Web Tool for Low to Ultra High Throughput Ligand-Based Virtual Screening. *J. Chem. Inf. Model.* **2016**, *56* (8), 1399−1404.
(40) Studio KO. http://www.studio-ko.ch (accessed Dec 2016).
(41) Bollag, G.; Tsai, J.; Zhang, J.; Zhang, C.; Ibrahim, P.; Nolop, K.; Hirth, P. Vemurafenib: the First Drug Approved for BRAF-Mutant Cancer. *Nat. Rev. Drug Discovery* **2012**, *11* (11), 873−886.
(42) Röhrig, U. F.; Majjigapu, S. R.; Vogel, P.; Zoete, V.; Michielin, O. Challenges in the Discovery of Indoleamine 2,3-Dioxygenase 1 (IDO1) Inhibitors. *J. Med. Chem.* **2015**, *58* (24), 9421−9437.
(43) Berman, H. M. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28* (1), 235−242.
(44) JSMol. https://sourceforge.net/projects/jsmol (accessed Dec 2016).
(45) ChemAxon. http://www.chemaxon.com (accessed Dec 2016).
(46) Bento, A. P.; Gaulton, A.; Hersey, A.; Bellis, L. J.; Chambers, J.; Davies, M.; Krüger, F. A.; Light, Y.; Mak, L.; McGlinchey, S.; Nowotka, M.; Papadatos, G.; Santos, R.; Overington, J. P. The ChEMBL Bioactivity Database: an Update. *Nucleic Acids Res.* **2014**, *42* (D1), D1083−D1090.
(47) Gfeller, D.; Grosdidier, A.; Wirth, M.; Daina, A.; Michielin, O.; Zoete, V. SwissTargetPrediction: a Web Server for Target Prediction of Bioactive Small Molecules. *Nucleic Acids Res.* **2014**, *42* (W1), W32−W38.
(48) SwissADME. http://www.swissadme.ch (accessed Dec 2016).
(49) Chimiscope. http://www.chimiscope.ch/thematiques/drug-design (accessed Dec 2016).
(50) L'Eprouvette. http://wp.unil.ch/mediationscientifique/activites/eprouvette (accessed Dec 2016).

DOI: 10.1021/acs.jchemed.6b00596
J. Chem. Educ. 2017, 94, 335−344