NEW FRONT. CHEM. (**2015**)
Volume 24, Number 2, pp. 101-107
ISSN 2393-2171; ISSN-L 2393-2171

Former: Ann. West Univ. Timisoara – Series Chem.
ISSN: 1224-9513

Article

TEACHING BIOINFORMATICS: ONLINE MOLECULAR DOCKING

Narcisa Crista[1], Adriana Isvoran[2,♣]

[2]University of Agricultural Sciences and Veterinary Medicine of Banat "Regele Mihai I al Romaniei" from Timisoara, Romania
[2]Department of Biology-Chemistry, West University of Timisoara, 16 Pestallozi, 300115, Timisoara, Romania, e-mail: adriana.isvoran@e-uvt.ro

ABSTRACT

This study describes the use of molecular docking technique applied on MTiAutoDock online server in teaching bioinformatics for students in chemistry, biology and biochemistry. It also illustrates the use the online sever to predict the cytochrome P450 2C9 enzyme interaction with the drug ibuprofen. The benefits of using the online docking server for teaching molecular docking to the students in chemistry, biology and biochemistry and in a country where the powerful computing facilities are missing are also discussed.

Keywords: teaching, bioinformatics, MTiAutoDock.

1. INTRODUCTION

In the universities of Timisoara the students in chemistry, biology and biochemistry do not have solid informatics skills as there are not topics of basic programming and statistical knowledge in their university curricula. Also, with a very few exceptions, they have not any knowledge concerning Unix or Linux operating systems. Taking into consideration these features, the online bioinformatics tools are very useful and inexpensive teaching and learning instruments allowing both the teachers and students to perform bioinformatics analysis. Teachers have the opportunity to illustrate how the bioinformatics tools are applied and students have the possibility to perform computational tasks in their own rhythm, to repeat the tasks as many times as they need in order to have a detailed view of the process they analyze.

♣ Correspondent author: Tel: +40256592634, Fax:+40256592620, E-mails: adriana.isvoran@e-uvt.ro

Last decades are characterized by the development of the internet and increase of the on-line resources. Also, the bioinformatics topics are now present in the university curricula for biology, biochemistry and computer science, especially at the master degree level. In order to better prepare students for master degrees and research activities there is a real need to teach bioinformatics in the undergraduate classes of biology and biochemistry [1, 2] and even to the high school level [3, 4]. Starting to 2011, PLOS Computational Biology's Education section introduced a new collection "Bioinformatics: Starting Early" (http://www.ploscollections.org/article/browseIssue.action?issue=info:doi/10.1371/issue.pcol.v03.i09). This collection is dedicated to bioinformatics teaching in secondary schools and it proves the great interest in introducing bioinformatics early in the study programs of pupils and students.

Another characteristic of the last decades is the massive accumulation of data concerning the sequences and structures of the biological macromolecules. These data are now structured and deposited in data bases, most of them being free accessible for educational purposes. The bioinformatics training help the students to access these databases, to efficiently use the information and the proposed analysis tools [5]. All these illustrate the need of teaching bioinformatics at least at the university level.

On-line facilities for teaching and performing bioinformatics analysis increase continuously. Forming competences for using on-line facilities to the students in chemistry, biology and biochemistry corresponds to the actual trend in modern chemical and biological sciences teaching. Once the students become familiar to using the databases and a few basic bioinformatics tools, such as to visualize and analyze the spatial structures of the biological molecules and to simulate their interactions with different ligands through molecular docking technique, they are able to explore new on-line bioinformatics facilities. Usually, tutorials and illustrations are available online to help students to learn how to use every resource and gather data that may answer their initial questions or can be use in further studies.

Besides the great diversity of the bioinformatics tools, the molecular docking is one of the computational tools that help the students to understand the interaction between biologically relevant molecules (proteins, nucleic acids, carbohydrates, lipids) at the atomic level, as well as to clarify fundamental biochemical processes, especially those based on the protein-protein and protein–ligands interactions. There are many servers allowing online protein-protein (http://bioinformatictools.blogspot.ro/2011/12/protein-protein-docking-servers.html) and protein-ligand (http://bip.weizmann.ac.il/toolbox/structure/binding.htm) molecular docking**,** a comprehensive list of docking web services, databases and computer-aided drug design tools being found on the Click2Drug server (http://www.click2drug.org/)**.**

Molecular docking tools are an integral part of many actual structure-based drug discovery studies. From this point of view, it becomes important for students to be able to use docking programs for a particular study.

Studying science means that students learn to use the scientific methods, understand how to develop a scientific theory, evaluate different data, test hypotheses and make predictions. They must develop ways of thinking and acting in the practice of science and use the correct language and different scientific methodologies to collect, organize, interpret, calculate and communicate information. For example, to explain the correlations between the structure and function of biologically important macromolecules and their involvement into cell processes, the students must to understand first the specific characteristics of chemical substances and

macromolecules utilized by living systems. By using bioinformatics tools, students have the opportunity to broaden their competences in studying, understanding and communicating science.

This paper illustrates the use the MTiAutoDock online server [6] to perform molecular docking tasks in teaching bioinformatics for students in chemistry, biology and biochemistry at master level. The illustrated molecular docking computation concerns the prediction of the interactions of the cytochrome P450 2C9 (CYP2C9) enzyme with the drug ibuprofen.

2. METHOD

The topics of the bioinformatics course cover the followings: analysis of the protein's sequences and sequence alignments, analysis of the global and local structural properties of proteins (protein's backbone properties, surface properties such as electrostatics and hydrophobicity, surface roughness, surface pockets, cavities and protrusions) and prediction of the protein-ligand interactions based on molecular docking method. There are 28 hours of teaching and 28 hours of practical activities for the bioinformatics course. For every student, a distinct project concerning a bioinformatics study of the interaction between a protein and a specific ligand is asked for the final examination. Information concerning the protein's sequences is retrieved using the UniProt database [7] and the facilities therein (http://www.uniprot.org/). Information concerning the ligand structure and properties may be retrieved from ZINC database (http://zinc.docking.org/) [8]. Sequence alignment is illustrated using the Basic Local Alignment Search Tool [9] on-line (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and multiple sequences alignment is also performed on-line using CLUSTALW tool [10]. The structures of the proteins and protein-ligand complexes are retrieved from the Protein Data Bank (PDB) [11] and the interactive visualization and analysis of structures is performed using UCSF Chimera computational tool [12]. Fpocket on-line tool is used for identification and characterization of pockets and binding sites present on the protein surface [13]. The hydrophobicity distribution on the protein surface is visualized and analyzed using UCSF Chimera tool. Prediction of the protein-ligand interactions is performed using the molecular docking implemented under the on-line docking servers. In the followings, the use of 1-Click Docking server is illustrated.

Molecular docking is a computational technique that predicts the binding orientation of one small molecule at the binding site of a target macromolecule, usually a protein, and estimates the binding affinity [14]. Molecular docking studies use one of the two approaches: the geometry-based approach considering the surface complementarity of the ligand and protein [15] and the energetic-based approach computing ligand-protein pairwise interactions energies. There are both advantages and limitations for every approach.

MTiAutoDock is an online docking tool predicting the binding affinity of a single or multiple ligands to a protein based on the energetic approach and allows the download and/or visualization of the binding poses (http://bioserv.rpbs.univ-paris-diderot.fr/services/MTiOpenScreen/). There also is a video tutorial for using it (http://bioserv.rpbs.univ-paris-diderot.fr/services/MTiOpenScreen/#usage) that students must follow before to use the server.

3. RESULTS AND DISCUSSIONS

This part of our study illustrates the use of the MTiAutoDock online tool for predicting the binding affinity of ibuprofen, an analgesic and anti-inflammatory drug, to the CYP2C9 enzyme. CYP2C9 enzyme is known to metabolize approximately 15% of the clinically used drugs, such as antidiabetics, anticonvulsants, anti-inflammatories, anti-hypertensives, proton pomp inhibitors and anxiolytics [17]. The three dimensional structures both for the unbound protein and for its complex with the drug warfarine has been determined [18] and their codes entry in the PDB are 1OG2 and 1OG5 respectively. There also is a crystallographic structure of mutated CYP2C9 in complex with the drug flurbiprofen [19], PDB code entry 1R9O. In our study we use the structural file 1OG5 without the ligands and prepared for docking using Chimera computational tool [12].

The structural file for the drug ibuprofen is extracted from ZINC database [8]. The two files are uploaded of the web page of MTiAutoDock tool (figure 1).

Figure 1: Upload the structural files on MTiAutoDock web page

It is possible to perform a blind docking or to specify the binding site by its residues or by the coordinates of a selected atom belonging to the binding site. Also, the user may upload a ligand or may try one of the ligands that are found in a predefined compound library. The ligand is taken randomly or selected by the physico-chemical criteria defined by the user.

After running the docking, the poses can be visualized or download for further analysis (figure 2) and the binding energy is also delivered (figure 3). The best binding mode of ibuprofen to the CYP2C9 has a binding affinity of -8.32 kcal/mol. Considering the crystallographic structures of CYP2C9 in complex with flurbiprofen and warfarin, we also perform molecular docking predictions for these drugs binding to the enzyme as a control.

The use of MTiAutoDock tool to analyze the binding of flurbiprofen to CYP2C9 predicts the binding energy of -9.20 kcal/mol and the predicted binding energy of warfarin to CYP2C9 is -9.63 kcal/mol. These results suggest that ibuprofen binds to CYP2C9, but the interactions with flurbiprofen and warfarin are stronger

Figure 2: The best pose visualization for the binding of ibuprofen to CYP2C9

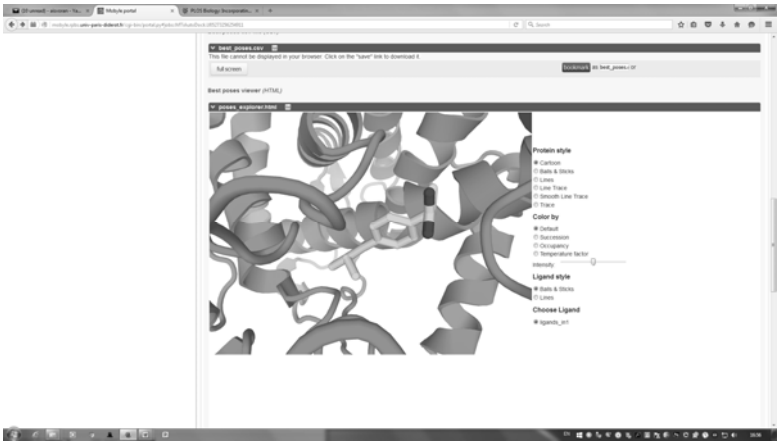
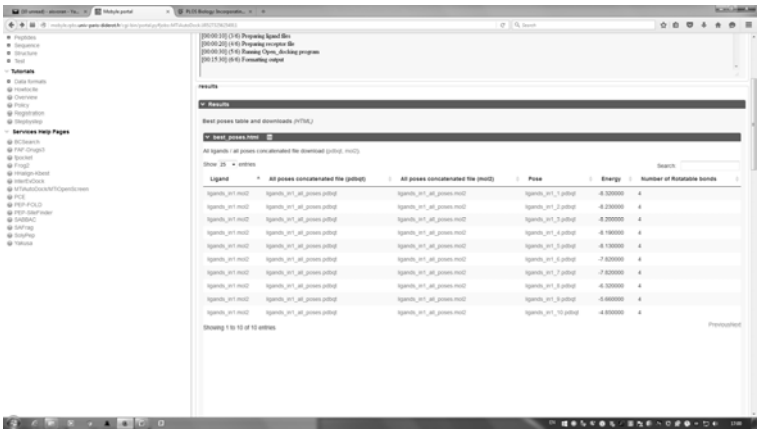


Figure 3: The interaction energies for all predicted poses for the binding of ibuprofen to CYP2C9



As the crystal structures of CYP2C9 in complex with the anti-inflammatory drug flurbiprofen and to the anticoagulant drug warfarin are solved, the residues identified as interacting to ibuprofen may be compared to those interacting to warfarin and flurbiprofen and it informs about the structural basis of drug binding to CYP2C9 .

4. CONCLUSION

In our faculty, the topic of bioinformatics is present in the study program of the master students in chemistry, biology and biochemistry only since 4 years. During this period, about 80 students have benefited from mentoring to use online tools to study biological processes at the molecular level. At the end of the semester students complete anonymous questionnaires about their personal benefits obtained by completing the course. Most students appreciate the totally new information and competencies regarding the use of bioinformatics methods to study chemistry, biology and biochemistry, especially the use of the biological databases and

of the bioinformatics tools for analyzing the biological interactions at the molecular level. They also consider that a bioinformatics course at the license level would be beneficial allowing them a better understanding of the structural levels of biological macromolecules and of their structure-function relationships.

The results presented in this study illustrates that online molecular docking facilities can be successfully employed as effective teaching tools to demonstrate the use of molecular docking to beginners in the field.

Free available on-line bioinformatics tools may be accessed by students from home, they only need their computers to be connected to the internet and they have the opportunity to repeat the steps they performed in classroom as many time as they need and in their own rhythm in order to obtain a complete understanding of both the bases of the bioinformatics tool they use and the explored biochemical processes. This opportunity contributes to a better understanding of the applications of bioinformatics tool and also may increase the interest of students to explore other tools and to perform other kinds of bioinformatics research without any cost.

The on-line teaching bioinformatics resources are important for both teachers and students because they develop interest to the students by facilitating understanding advanced studies and they promote quality human resource to undertake challenging research in the field of modern chemistry, biology and biochemistry. They also contribute to the interdisciplinary approach of the study of the physiological processes at the molecular level.

REFERENCES

1. Pevzner, P.; Shamir, R.. Computing has changed biology–biology education must catch up. *Science*, 2009, 325, (541–542), doi: 10.1126/science.1173876.
2. Ditty, J.L; Kvaal, C.A.; Goodner, B.; Freyermuth, S.K.; Bailey, C. Incorporating genomics and bioinformatics across the life sciences curriculum. *PLOS Biol,* 2010, 8, (e1000448), doi: 10.1371/journal.pbio.1000448.
3. Form, D.; Lewitter, F.. Ten Simple Rules for Teaching Bioinformatics at the High School Level. *PLOS Comp Biol*, 2011, 7, (e1002243), doi: 10.1371/journal.pcbi.1002243.
4. Lewitter, F.; Bourne P.E. Teaching Bioinformatics at the Secondary School Level. *PLOS Comp Biol*, 2011, 7, (e1002242), doi: 10.1371/journal.pcbi.1002242.
5. Luo, J. Teaching the ABCs of bioinformatics: a brief introduction to the Applied Bioinformatics Course. *Briefings in bioinformatics*, 2014, 15, (1004-1013), doi: 10.1093/bib/bbt065.
6. Labbé, C.M.; Rey, J.; Lagorce, D.; Vavruša, M.; Becot, J., Sperandio, O., Villoutreix, BV.O.; Tufféry, P; Miteva, M.A. MTiOpenScreen: a web server for structure-based virtual screening. *Nucl Acids Res,* 2015, (1-7), doi: 10.1093/nar/gkv306.
7. The UniProt Consortium, The Universal Protein Resource (UniProt), *Nucl Acids Res*, 2008, 36, (D190–D195), doi: 10.1093/nar/gkl929.
8. Irwin, J.J.; Shoichet, B.K.; Zinc – a free database of commercially available compounds for virtual screening. *J Chem Inf Mod*, 2005, 45, (177–182), doi: 10.1021/ci049714.

9. Altschul, S., Gish, W., Miller, W., Myers, E., Lipman, D. Basic local alignment search tool. *J Mol Biol,* 1990, 2153, (403–410), PMID: 2231712.
10. 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. Clustal W and Clustal X version 2.0. *Bioinformatics,* 2007, 23, (2947-2948), 10.1093/bioinformatics/btm404.
11. Bernstein, F.C.; Koetzle, T.F.; Williams G.J.; Meyer, E. E. Jr.; Brice M.D.; Rodgers, J.R.; Kennard, O.; Shimanouchi, T.; Tasumi, M.. The Protein Data Bank: A Computer-based Archival File For Macromolecular Structures. *J Mol Biol,* 1977, 112, (535-542), PMID: 875032.
12. 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 Comp Chem*, 2004, 25, (1605-1612), doi:10.1002/jcc.20084.
13. Schmidtke, P.: Le Guilloux, V.; Maupetit, J.; Tufféry, P. Fpocket: online tools for protein ensemble pocket detection and tracking. *Nucl Acids Res*, 2010, 38, (W582-W589), doi:10.1093/nar/gkq383.
14. Morris, G.M.; Lim-Wilby, M.; Molecular docking. *Meth Mol Biol* , 2008, 443, (365-382), doi: 10.1007/978-1-59745-177-2_19.
15. Shoichet, B.K.; Kuntz, I.D.; Bodian, D.L. Molecular docking using shape descriptors. *J Comp Chem,* 2004, 13, (380–397), doi: 10.1002/jcc.540130311.
16. Meng, X-Y.; Zhang, H-X.; Mezei, M.; Cui, M. Molecular Docking: A powerful approach for structure-based drug discovery. *Curr Comp Aided Drug Des,* 2011, 7, (146–157), PMCID: PMC3151162.
17. Yiannakopoulou E.C. Pharmacogenomics for individualized therapies, In *OMICS-Applications in Biomedical. Agricultural and Environmental Sciences*.; Bahr D., Zambare V., Azevedo V., Eds.; CRC Press, Boca Raton USA, 2013; 95-120.
18. Williams, P.A.; Cosme, J.; Ward, A.; Angove, H.C.; Vinkovic, D.M.; Harren, J. Crystal structure of human cytochrome P450 2C9 with bound warfarin. *Nature,* 2003, 424, (464-468), doi:10.1038/nature01862.
19. Wester, M.R.; Yano, J.K.; Schoch, G.A.; Yang, C.; Griffin, K.J.; Stout, C.D.; Johnson, E.F. The Structure of Human Cytochrome P450 2C9 Complexed with Flurbiprofen at 2.0 A Resolution. *J Biol Chem*, 2004, 279, (35630-35637), doi: 10.1074/jbc.M405427200.