Programming a Web Interface
for Protein-Ligand Docking

Aldo Hiram Tovar-Domínguez[1], Adrián Arturo Ayala-Sánchez[1],
Mauricio Carrillo-Tripp[2], José Marcos Falcón-González[1]

[1] Instituto Politécnico Nacional,
Unidad Profesional de Ingeniería Campus Guanajuato,
Mexico

[2] Instituto Politécnico Nacional,
Centro de Investigación y de Estudios Avanzados Unidad Monterrey,
Mexico

`jfalcong@ipn.mx`

Abstract. Molecular docking has been widely studied because it seeks to predict the energies and binding modes between ligand and protein, with applications in drug development, the study of diseases, protein design, etc. In the present work, the aim is to automate the docking results through an intuitive and flexible interface, with a view to virtual screening, such that predictive systems are made more efficient, giving a shorter response time. Thus, it is not necessary to install different packages, nor to intervene in each of the intermediate processes, such as the addition of Gasteiger and Kollman charges, and the removal of external molecules to the system, giving a more efficient response time.

Keywords: Docking, interface, affinity.

1 Introduction

Over the years, structural biologists and bioinformaticians have shown interest in the way molecules interact with one another due to its implications for drug design, the study of diseases [7], among other applications. There has been a need to better understand how these complex interactions arise from the topology of the molecules, their physicochemical properties, their environment, and many other factors. Some molecular modeling tools include molecular docking (docking), a very useful technique for predicting the structure of intermolecular complexes established between two or more molecules [9].

Enantiomers are stereoisomers that cannot be superimposed, the importance of which lies in the fact that in biological systems only one of the two possible enantiomers are present for a chiral molecule [3]. Some studies have reported enantiomers can behave as potent and selective inhibitors of matrix metalloproteinases (MMPs), which could be useful for the treatment of diseases such as cancer and arthritis [5], AR, IR, SIRT-6 have also been reported to be used as prominent target proteins to study the

Table 1. Software used in the development of the interface.

Software/Packages	Use
Open Babel	Geometry optimization for each ligand [10]
Chimera	Removing of all molecules except the chains of interest from .pdb
AutodockFR	Addition of missing hydrogen atoms and Gasteiger (Ligand) and Kollman (Receptor) charges; Keeping of polar hydrogen atoms
P2Rank [2]	Search of pocket coordinates for the first (coarse) docking
Awk	Processing of atom coordinates for the computation of the box
Autodock Vina [7]	Ligand conformers, affinities, and distance from the best conformer (RMSDs upper bound and lower bound)
Python and Visual Studio Code	Python is used as the programming language; VSC is used as a viewer linking to the web server.
Django	Used to make virtual environments and databases.
PHP	In charge of receiving the documents and creating the folder that will later be called with javaScript.
HTML and CSS	They are used to give structure, format, and declaration of variables to run the script and the documents.
JavaScript	Start the php docs and start the file with the script.
GitHub	Web server where the page is located.

interaction of selected antidiabetic compounds isolated from various medicinal plants through *in silico* detection [4].

The rigid coupling of proteins and ligands is based on Emil Fischer's classic lock and key model. This approach assumes that the ligand and the protein are rigid during their interaction [6]. While this is not the most accurate way to describe these events [1], it makes for computationally inexpensive methods and a starting point for more detailed studies on the subject.

In this case, the preprocessing generally includes three major steps: the deletion of molecules that may interfere with the docking (e.g., buffer molecules and the ligand itself), the addition of charges (since the ligand and the receptor may behave differently at a given pH), and the adjustment of a box which delimits the space for the proposal of several possible ligand locations.

When it comes to docking protein ligand pairs, the process of manual work, which involves working directly with files and issuing commands one by one, can be tedious and error prone. A first approach is to automate these tasks with a shell script or Python and creating a web interface.

2 Methods

When testing the software, all protein and ligand structures were downloaded from RCSB PDB and PubChem, respectively. P2Rank is a tool that predicts pockets in a protein using a machine learning approach. This helps to quickly find places where it
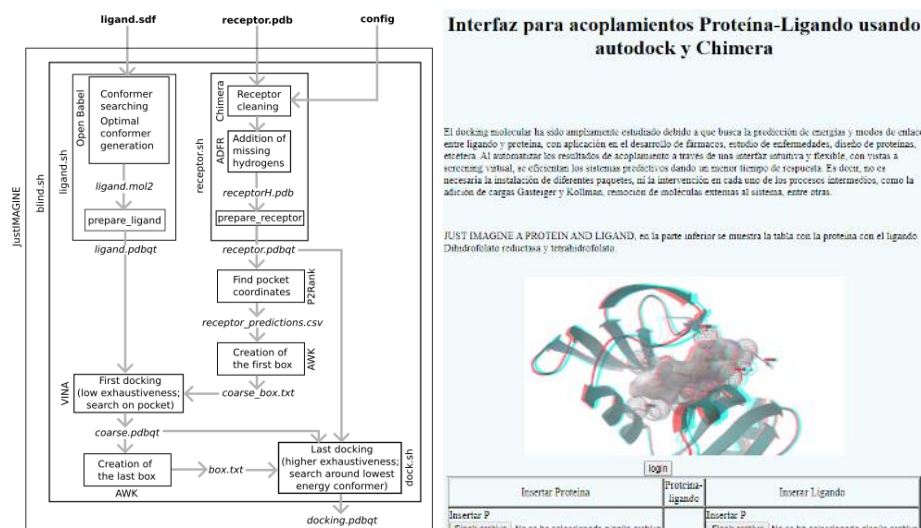
Fig.11. Flowchart of the docking software on the left. On the top right is the main web page and at the bottom is the docking processing.

is more likely for the ligand to bind the protein and thus it cuts down the time required when using Vina with a higher exhaustivity value.

The molecular docking and its previous processes were put together in a Bash script. Figure 1 illustrates the flow of information through the script. In bold format are the input files (given by the user) and in italics are the files given as the output.

3 Results

For the visualization of the coupling, it was decided to use free hosting (GitHub), the GitHub software is used mainly for the creation of source code of computer programs. The script can be downloaded [11] to be executed by simply running ./blind.sh from the terminal. Before using it this way, however, the user must install the necessary tools (Open Babel, Chimera, AutodockFR, P2Rank and Awk); it is also advisable to work on a UNIX like environment for this.

Protein ligand coupling presents efficient waiting times of less than 10 min (on a machine with 4 GB of RAM and a quad core processor running at 2.10 GHz). For correct operation as shown in Figure 1 in the upper right corner shows the start of the web page, at the bottom of it there is a blue link with the word coupling. In the lower right image is the page to make a link, in which the ligand and protein files must be inserted, later they are uploaded by pressing the upload button and clicking on "click me" to execute the link.

The files are sent to the server which processes them and returns the result (only .pdb and .sdf files with terminal are accepted). The output given by Autodock Vina consists of a summary of the newfound conformers, their affinities and distances from the best mode (lowest affinity), measured as RMSD lower bound and RMSD upper bound. A short example is given in Table 2.

Table 12 Output given by Autodock Vina.

Mode	Affinity (Kcal/mol)	Dist from best mode Rmsd l.b	Dist from best mode Rmsd u.b.
1	-7.5	0	0
2	-7.3	1.939	2.436
3	-7.1	3.801	6.552

4 Conclusions

The molecular docking automation was achieved in such a way that the user should hardly do more than provide the files representing the ligand and its receptor. The process is standardized and takes considerably less time to complete, which makes it more efficient and lets the user concentrate on any further parts of their investigation.

References

1. Fan, J., Fu, A., Zhang, L.: Progress in molecular docking. Quantitative Biology, vol. 7, no. 2, pp. 83–89 (2019) doi: 10.1007/s40484-019-0172-y
2. Krivák, R., Hoksza, D.: P2Rank: machine learning based tool for rapid and accurate prediction of ligand binding sites from protein structure. Journal of cheminformatics, vol. 10, no. 1, pp. 1-12 (2018) doi: 10.1186/s13321-018-0285-8
3. Rascón, M. A.: Obtención de alcoholes quirales mediante catálisis asimétrica por complejos de rodio (III) con ligantes tipo base de Schiff (2014)
4. Singh, P., Singh, V. K., Singh, A. K.: Molecular docking analysis of candidate compounds derived from medicinal plants with type 2 diabetes mellitus targets. Bioinformation, vol. 15, no. 3, pp. 179 (2019) doi: 10.6026/97320630015179
5. Pochetti, G., Gavuzzo, E., Campestre, C., Agamennone, M., Tortorella, P., Consalvi, V., Gallina, C., Hiller, O., Tschesche, H., Tucker, P. A., Mazza, F.: Structural insight into the stereoselective inhibition of MMP-8 by enantiomeric sulfonamide phosphonates. Journal of medicinal chemistry, vol. 49, no. 3, pp. 923–931 (2006) doi: 10.1021/jm050787+
6. Prieto-Martínez, F. D., Medina-Franco, J. L.: Diseño de fármacos asistido por computadora: cuando la informática, la química y el arte se encuentran. TIP Revista Especializada en Ciencias Químico-Biológicas, vol. 21. no. 2, pp. 124–134 (2019)
7. Trott, O., Olson, A. J.: AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of computational chemistry, vol. 31, no. 2, pp. 455–461 (2010)
8. Vargesson, N.: Thalidomide- induced teratogenesis: History and mechanisms. Birth Defects Research Part C, Embryo Today: Reviews, vol. 105, no. 2, pp. 140–156 (2015)
9. Velásquez, M., Drosos, J., Gueto, C., Márquez, J., Vivas-Reyes, R. A.: Autodock-PM6 método para seleccionar una mejor pose en estudios de acoplamiento molecular. Revista Colombiana de Química, vol. 42, no. 1, pp. 245–261 (2013)
10. Yoshikawa, N., Hutchison, G. R.: Fast, efficient fragment-based coordinate generation for Open Babel. Journal of cheminformatics, vol. 11, no. 1, pp. 1–9 (2019)
11. https://athas1.github.io/A/Docking/Prime.html