v. 3, n. 5, 2023

International Journal of Health Science

# MOLECULAR DOCKING AS A TOOL FOR OPTIMIZING TESTS TO SEARCH FOR TARGETS FOR THE TREATMENT OF FUNGAL INFECTIONS

***Benedito Rodrigues da Silva Neto***
Instituto de Patologia Tropical e Saúde Pública, Universidade Federal de Goiás, Goiânia, Goiás, Brazil

**Abstract: Background**: The optimization of in vitro tests using the in silico methodology is promising because it favors the accuracy and value of experimental tests. **Methods**: In this work we employed MOLDOCK program in all molecular docking simulations. Accuracy of enzyme-ligand docking was validated on a set of CYP51-ligand (punicalagin) complexes for which crystallographic structure were available, generating root-mean square deviation below 2.0 Å. **Results:** Application of the default MOLDOCK protocol generated the structure with an RMSD of 1.3 Å. Clearly, the default protocol worked for this complex. Here we described an efficient molecular docking protocol, which was able to recover crystallographic position of a punicalagin present in the active site of the CYP51. **Conclusions**: The in silico tests suggest that punicalagin is a promising plant-derived compound that can be used for in vitro analyzes aimed at the development of new antifungals.
**Keywords:** Molecular Docking, Natural compounds, Fungal infections.

## INTRODUCTION

Fungal infections in immunocompromised patients have substantially increased in number and severity over the past five decades [1]. Emergence of candidemia cases have been observed, with this infection becoming a severe problem in all over the world, causing a mortality rate around 50% [2-4]. Candidiasis is considered as the 4th most common nosocomial infection in the world, and in Brazil according to Doi et al. [5], the 7th most prevalent [6,7].

Natural compounds from plants have garnered increasing attention among the scientific community for their lower cost, higher bioavailability, and less toxicity compared to synthetic pharmaceutical agents [8]. As an important source of anthocyanins and hydrolysable tannins, pomegranate is consumed as a fruit and is also used for its antioxidant and anti-inflammation potential on disease prevention and treatment. Pomegranate peel extract contains high amounts of bioactive compounds, mainly phenolic acids, flavonoids and tannins. Among all the polyphenols in pomegranate, punicalagin [9]

Molecular Docking, a computer simulation methodology that can predict the conformation of a protein-drug complex, with relatively high accuracy when compared with experimental structures was previously analyzed. Analyses of the interactions between a protein target and a drug can be simulated by these evolutionary algorithms [10]. The molecular docking approach can be used to model the interaction between a small molecule and a protein at the atomic level, which allow us to characterize the behavior of small molecules in the binding site of target proteins as well as to elucidate fundamental biochemical processes. The docking process involves two basic steps: prediction of the ligand conformation as well as its position and orientation within these sites (usually referred to as *pose*) and assessment of the binding affinity. These two steps are related to sampling methods and scoring schemes, respectively, which will be discussed in the theory section [11].

In this scenario, plants and their derivatives are known to be important in pharmacological research. New compounds from natural products provide a source of chemical scaffolds that have a variety of biological activities in the drug development.

## MATERIALS AND METHODS

### STRUCTURE AND INFORMATION

The 2D and 3D structure of punicalagin was obtained from the public bank (https://pubchem.ncbi.nlm.nih.gov/compound/

Puni-calagin), as well as general information regarding its chemical properties, toxicity, pharmacology, biochemistry and classification.

## PRE-DOCKING ANALYSIS

The main goal of the pre-docking analysis was to investigate the overall quality of the structures in the sterol 14α-demethylase (CYP51) dataset (PDB access code: 5FSA, resolution 2.86 Å), allowing us deciding which structure in the dataset has the most reliable crystallographic information (Figure 1).

## MOLECULAR DOCKING SIMULATIONS

One of the fundamental questions in structural biology is the study of protein-ligand interactions, particularly considering the pharmacological applications of such study in the design of drugs based on structure [12]. To simulate the interaction of with a library of ligands, we used the Molegro Virtual Docker (MVD) that presents two EAs to carry out position search in MDSs. Recent evaluation of MOLDOCK, program [13] an implementation of a variant of the evolutionary algorithm (EA), strongly indicates that it is capable of finding the right position of a ligand. In the present work, all simulations were performed in a MacBook Air (Intel Processor Core i5 Duo, 1.4 GHz, 4 GB SDRAM DDR3 1600 MHz).

## RE-DOCKING ANALYSIS

In molecular docking simulations, the best binary complex (protein-ligand) is the one closer to the crystallographic structure. For that reason, we must establish a methodology that assesses the distance from the computer-generated solution (pose) to the crystallographic structure. This distance can be calculated using the root-mean-square deviation (RMSD), which is a measure of the differences between values predicted by a model and the values actually observed from the object being modeled or estimated (protein-ligand complex). The RMSD is calculated between two sets of atomic coordinates, in this case, one for the crystallographic structure (xctal, yctal, zctal; the object being modeled) and another for the atomic coordinates obtained from the docking simulations (xpose, ypose, zpose; predicted model) [14].

In docking simulations, it is expected that the best results generate RMSD values less than 2.0 Å compared with crystallographic structures. This procedure of obtaining the crystallographic position of the ligand is often called "re-docking," which is fundamentally a validation method that determines whether the molecular docking algorithm is able to recover the crystallographic position using computer simulation. In this work, all RMSD calculations were calculated for non-hydrogen atoms [15].

We used the CYP51 crystallographic coordinates available at the protein data bank (PDB), under the access code PDB: 5FSA, resolution 2,86 Å [16]. We performed the docking simulation against the active site of 5FSA and compared the docked poses with the crystallographic structure.

## SELECTION OF RESULTS WITH LOWEST SCORES

MOLDOCK program is the workhorse of the present protocol. It was used in all docking simulations described here. During a typical docking simulation several orientations can be obtained for each ligand. Here we selected the one with the lowest scoring function. The scoring function used by MOLDOCK improves accuracy of scoring functions with a new hydrogen bonding term and new charge schemes. Four scoring functions are implemented in the MOLDOCK, including

MOLDOCK score and PLANTS score [17,18]. These two functions offer grid-based versions, in which hydrogen bond directionality is not considered. In the present protocol we employed grid-based MOLDOCK score since it offers approximately four-fold greater speed by performing a precalculus of potential-energy values on an equally spaced cubic grid.

In order to better visualize the intermolecular interactions between target protein and ligand residues, with access to protein residues that interact with the small molecule atoms by hydrogen bonds and hydrophobic interactions, the Ligplot program was used [19]. Thus, it is possible to evaluate the amino acid residues of the protein interacting with the small molecule atoms under test, in addition to understanding the nature of these interactions.

### Statistical analysis

Statistical analysis was performed using GraphPad Prism version 5.01 software for windows (GraphPad Inc., San Diego, CA). Data are expressed as mean ± standard deviation (SD) of three independent assays. One-way analysis of variance (ANOVA) followed by Bonferroni's post-test was used for inter group variation analysis. Statistical significance was established as $p<0.05$.

## RESULTS

### CRYSTALLOGRAPHIC INFORMATION ABOUT STEROL 14A-DEMETHYLASE (CYP51)

Sterol 14α-demethylase (CYP51, EC 1.14.13.70) is the most functionally conserved cytochrome P450 (CYP) monooxygenase (Figure 1) [20]. Azoles block sterol biosynthesis by inhibiting fungal sterol 14α-demethylase, the membrane-bound enzyme in the endoplasmic reticulum that removes the 14 -methyl group from the first cyclized sterol precursor and thus initiates the advancement of the pathway toward its final products (ergosterol in fungi) [21].

A search in the protein data bank (PDB) for the keyword sterol 14α-demethylase (CYP51) resulted in 53 entries (april, 2020). As we can see, most of the CYP51 crystallographic structures were obtained from *Trypanosoma cruzi* (19 structures) *and Trypanosoma brucei* (9 structures) 28 out of 53 entries (52,8%). Another interesting key aspect of the available CYP51 structures; they were all solved by X-ray diffraction crystallography technique (molecular replacement method).

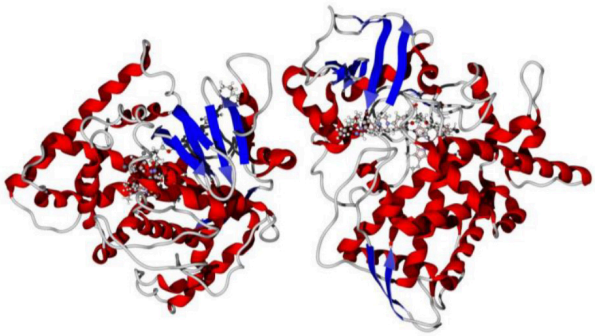

Figure 1 - Crystal structure of sterol 14-alpha demethylase (CYP51) from a pathogenic yeast *Candida albicans* in complex with the antifungal drug posaconazole (PDB access code: 5FSA), resolved in the MVD program.

### DOCKING

A search for the best molecular docking protocol was performed. The structure of CYP51 in complex with the antifungal drug posaconazole was used for re-docking simulations. The key criterion describing the quality of an MDS is the RMSD. In molecular docking applications, the best binary complex is the one closer to the structure determined by x-ray crystallography. Analysis of the re-docking results for the combination of 4 scoring functions (MolDock Score, MolDock Score [GRID], PLANTS Score and PLANTS Score [GRID]) and 4 search algorithms (MolDock Optimizer, MolDock SE, Interated

Simplex, GPU Screening [CUDA]) (a total of 16 different docking protocols).

Molecular docking simulations of the crystallographic structure solved at high-resolution for CYP51 in complex with the antifungal drug posaconazole PDB: 5FSA, resolution 2,86 Å [10], is presented here to illustrate the application of RMSD to the analysis of the docking results. Application of the default MOLDOCK protocol generated the structure shown in (Figure 2), with an RMSD of 1.3 Å. Clearly, the default protocol worked for this complex. In addition, we performed the docking simulation against the active site of CYP51 and compared the docked poses with the crystallographic structure.

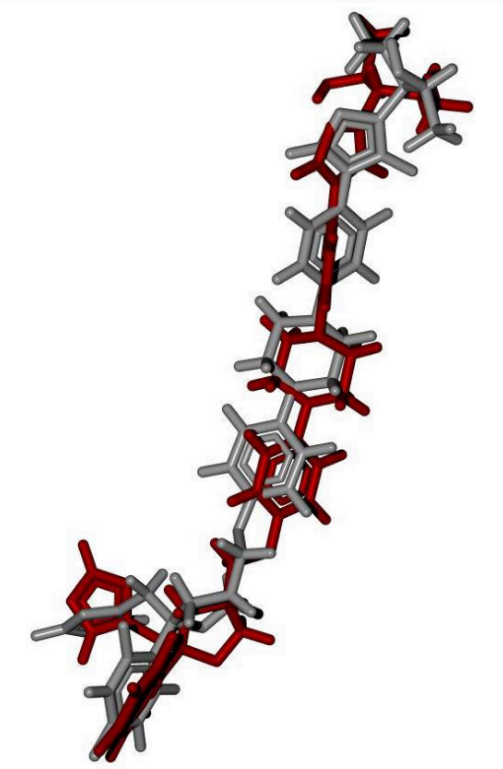

Figure 2 - Superposition of the best docked structure. Pose and crystallographic structure of posaconazole in the active site of sterol 14α-demethylase (PDB access code: 5FSA)

The best results were obtained for the following search engines: MOLDOCK Score and MOLDOCK optimizer (RMSD 1.89 Å). The parameters for docking, especially the search engine features, were optimized by running several MDSs on the complex structure. The following parameters and their combinations were varied: radius of the docking sphere, number of runs, maximum number of interactions, and maximum population size.

The optimized parameters for the docking are the following:

Scoring function - Empirical scoring function: Re-rank score (used for ranking the MDS results). Binding site - Origin: x= (190.93); y= (2.93) and z= (39.39) Å, and a docking sphere radius of 15 Å. Search algorithm- Algorithm: MOLDOCK Optimizer; Number of runs: 20; Constrain poses to cavity: Enabled. Parameter settings - Max iterations: 2000; Max population size = 50. Pose generation - Energy threshold: 100.00. Simplex evolution - Max steps: 300; Neighbour distance factor: 1.00. Figure 3 shows the docking sphere used in the re-docking simulations. Since the combination of MOLDOCK Score and MOLDOCK optimizer generated very low RMSD, we chose this docking protocol and used it in all further MDSs. These tests indicated that the docking simulation was successful, and that the protocol is good enough to be used for the virtual screening process.

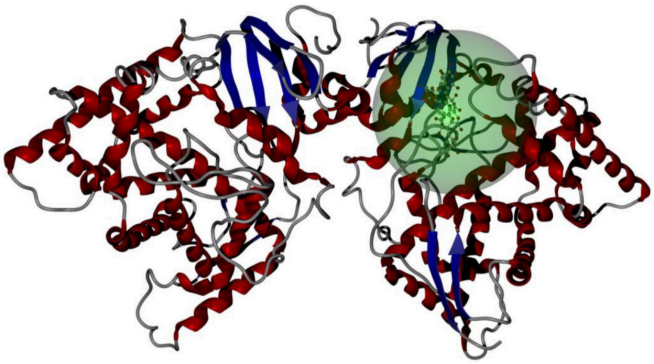

Figure 3 - Search space sphere (green) defined for molecular docking simulations.

## INTERMOLECULAR INTERACTIONS

In order to better understand the interactions of punicalagin with CYP51, we used the program LIGPLOT [22] to access the atoms of both, the small molecules and the protein ones that are responsible to make hydrogen bonds and van der Waals contacts. We could observe, only, that among the selected compounds the best scores mean a greater potential to interact with CYP51 binding cavity (Figures 4 and 5).

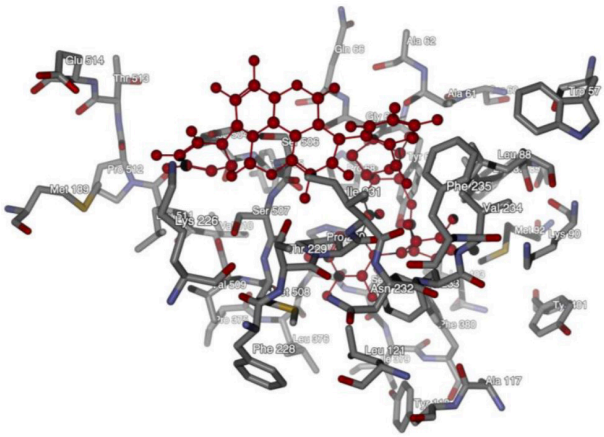


Figure 4 - CYP51-binding pocket with main residues found in intermolecular interactions with punicalagin.

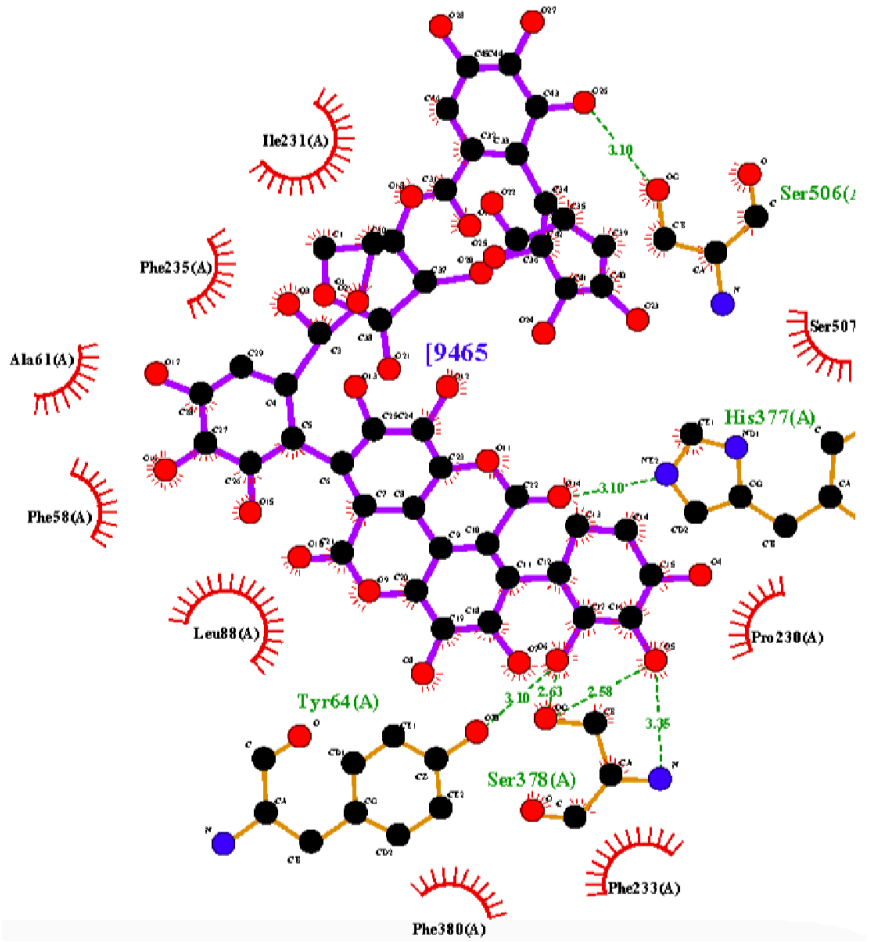


Figure 5 - Map of interactions occurring in the protein-binding complex. (Purple: ligand-bond (Punicalagin); Yellow: Non-ligand bond; Green dotted: Hydrogen bond and its length; Medium circles with red scratches: Non-ligand residues involved in hydrophobic contact (s) (van der Waals)). Black spheres with scratches: Corresponding atoms involved in hydrophobic contact (s).

## DISCUSSION

Identifying and validating molecular targets represent the first step in finding new drugs. With increasing technological development, it became possible to identify key enzymes, receptors, and blockers from action targets in the fungal cell [23].

In this way, the docking simulation results corroborate the importance of some CYP51-active site residues as responsible to establish intermolecular interactions with the substrate as well as with the tested ligand. The binding of punicalagin to its cavity, presents pivotal residues that make protein-ligand interactions possible, as shown in Fig. 5. These residues are essential to the ligand binding and, to the reaction catalyzed by the enzyme [24-26].

For the docked structure of CYP51-punicalagin, there are four intermolecular hydrogen bond involving residue Ser506, His377, Ser378 and Tyr64. We observe five intermolecular hydrogen bonds involving residues, where two of them interactions for Ser378 [27]. Analysis of the van der Waals interactions between CYP51 and punicalagin at this site indicated the participation of residues Ile231, Phe235, Ala61, Phe58, Leu88, Phe380, Phe233, Pro230 and Ser507 as the major contact points with the ligands.

## CONCLUSION

Advanced molecular docking algorithms available nowadays make it possible to undertake larger virtual screening studies focused on small-molecules libraries up to millions of compounds. Here we described an efficient molecular docking protocol, which was able to recover crystallographic position of a ligand present in the active site of the CYP51. Re-docking simulations generated RMSD results below 2 Å. Besides, we identified the CYP51 binding-cavity residues that are essential to make possible the interactions of this enzyme with punicalagin.

## REFERENCES


1. Pfaller MA, Diekemaa DJ. Epidemiology of invasive mycoses in North America. Crit Rev Microbiol. 2010;36(1):1–53.

2. Colombo AL, Guimarães T, Sukienik T, Pasqualotto AC, Andreotti R, Queiroz-Telles F, et al. Prognostic factors and historical trends in the epidemiology of candidemia in critically ill patients: an analysis of five multicenter studies sequentially conducted over a 9-year period. Intensive Care Med. 2014;40(10):1489–98.

3. Magill SS, Edwards JR, Bamberg W, Beldavs ZG, Dumyati G, Kainer MA, et al. Multistate Point-Prevalence Survey of Health Care–Associated Infections. N Engl J Med [Internet]. 2014 Mar 27;370(13):1198–208. Available from: http://www.nejm.org/doi/10.1056/NEJMoa1306801

4. Bassetti M, Righi E, Ansaldi F, Merelli M, Trucchi C, Cecilia T, et al. A multicenter study of septic shock due to candidemia: outcomes and predictors of mortality. Intensive Care Med [Internet]. 2014 Jun;40(6):839–45. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24807083

5. Doi AM, Pignatari ACC, Edmond MB, Marra AR, Camargo LFA, Siqueira RA, et al. Epidemiology and microbiologic characterization of nosocomial candidemia from a Brazilian National Surveillance Program. PLoS One [Internet]. 2016;11(1):e0146909. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgiartid=4726651&tool=pmcentrez&rendertype=abstract

6. Pfaller MA, Diekema DJ. Epidemiology of Invasive Candidiasis: a Persistent Public Health Problem. Clin Microbiol Rev [Internet]. 2007 Jan 1;20(1):133–63. Available from: http://cmr.asm.org/cgi/doi/10.1128/CMR.00029-06

7. Lortholary O, Renaudat C, Sitbon K, Madec Y, Denoeud-Ndam L, Wolff M, et al. Worrisome trends in incidence and mortality of candidemia in intensive care units (Paris area, 2002-2010). Intensive Care Med. 2014;40(9):1303–12.

8. Atanasov A.G., Waltenberger B., Pferschy-Wenzig E.-M., Linder T., Wawrosch C., Uhrin P., Temml V., Wang L., Schwaiger S., Heiss E.H., et al. Discovery and resupply of pharmacologically active plant-derived natural products: A review. *Biotechnol. Adv.* 2015;33:1582–1614. doi: 10.1016/j.biotechadv.2015.08.001. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

9. Kaderides K., Kyriakoudi A., Mourtzinos I., Goula A.M. Potential of pomegranate peel extract as a natural additive in foods. *Trends Food Sci. Technol.* 2021;115:380–390. doi: 10.1016/j.tifs.2021.06.050. [CrossRef] [Google Scholar]

10. de Azevedo Jr WF and Heberlé G. Bio-Inspired Algorithms Applied to Molecular Docking Simulations. Current Medicinal Chemistry 2011; 18:1339-1352.

11. McConkey BJ, Sobolev V, Edelman M. The performance of current methods in ligand-protein docking. *Current Science.* 2002;83:845–855. [Google Scholar]

12. De Azevedo Jr WF, Canduri F, Da Silveira NJF (2002) Biochem Biophys Res Commun 293:566–571.

13. Thomsen R, Christensen MH (2006) MolDock: a new technique for high-accuracy molecular docking. J Med Chem 11:3315–3321

14. ViannaCP, & De Azevedo Jr WF, (2012) Identification of new potential Mycobacterium tuberculosis shikimate kinase inhibitors through molecular docking simulations. J Mol Model 18:755–764

15. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shaw DE, Shelley M, Perry JK, Francis P, Shenkin PS (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 47:1739–1749

16. Hargrove TY, Friggeri L, Wawrzak Z, Qi A, Hoekstra WJ, Schotzinger RJ, York JD, Guengerich FP, and Lepesheva GI (2017) Structural analyses of *Candida albicans* sterol 14 -demethylase complexed with azole drugs address the molecular basis of azole-mediated inhibition of fungal sterol biosynthesis. *J. Biol. Chem.* (2017) 292(16) 6728 –6743

17. Thomsen R, Christensen MH (2006) MolDock: a new technique for high-accuracy molecular docking. J Med Chem 11:3315–3321

18. De Azevedo Jr WF (2010) MolDock applied to structure-based virtual screening. Curr Drug Targets 11:327–334

19. Wallace AC, Laskowski RA, Thornton JM. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng. 1995; 8(2):127-134.

20. Lepesheva, G. I., and Waterman, M. R. (2011) Structural basis for conservation in the CYP51 family. *Biochim. Biophys. Acta* **1814,** 88 –93

21. Van den Bossche, H. (1988) in *Sterol Biosynthesis Inhibitors* (Berg, D., and Plempel, M., eds) pp. 79 –119, Ellis Horwood, Chichester, UK

22. Wallace AC, Laskowski RA, Thornton JM (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Prot Eng 8:127–134

23. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. Nat Rev Drug Discov [Internet]. 2004 Nov;3(11):935–49. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15520816

24. Vasconcelos LCDS, Sampaio MCC, Sampaio FC, Higino JS. Use of Punica granatum as an antifungal agent against candidosis associated with denture stomatitis. Mycoses [Internet]. 2003 Jun;46(5–6):192–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12801361

25. Westbrook J, Feng Z, Chen L, Yang H, Berman HM. The Protein Data Bank and structural genomics. Nucleic Acids Res 2003; 31(1): 489-91.

26. Berman HM, Battistuz T, Bhat TN, *et al*. The Protein Data Bank. Acta Crystallogr D Biol Crystallogr 2002; 58(Pt 6 No 1): 899-907.

27. Berman HM, Westbrook J, Feng Z, *et al*. The Protein Data Bank. Nucleic Acids Res 2000; 28: 235-42.