DocVSP (Docking-based Virtual Screening Perl-script) for Automating and Integrating AutoDock and SBDD

Shivam Chauhan, Mani Srivastava, Jaikee Singh

Jaipur National University, Jaipur, India

Corresponding author: Shivam Chauhan, Email: shivamchauhan5120420@gmail.com

Molecular-docking-based virtual screening is a significant tool in investigation ligand-receptor relatedness and recognizes possible drug candidates in drug discovery. In Silico method of drug discovery screen of millions compound in a reasonable time with appropriate scoring function is a challenge. A widely used docking program is AutoDock but some practical difficulties arise, while other hand reasonable time also other technical challenge to screen of millions of compounds. So, to meet this, we use high performance computing platforms and techniques. We have developed a fully automated Perl script to perform all the step of structure base drug designing (SBDD) in such a way to overcome the limitations of AutoDock while performing virtual screening, a clear advantage over AutoDock and single-stop platform for all protocol of SBDD. The docking tool and Perl script executes in alongside on a Linux system. Furthermore, the docking period is totally automated combined with a Linux system for simultaneously computing, the consequence can be evoked and extracted. This method separates many complication and collective troubles related with virtual screening, and come up with a clear advantage and well organized outcome for AutoDock users on a Linux assemble programme.

Keywords: Virtual screening, Perl script, Fully automated integrated, SBDD, AutoDock.

Shivam Chauhan, Mani Srivastava, Jaikee Singh

1   Introduction

In the last several years, the protein structures attainability lead the way to become molecular docking an increasingly important tool for drug discovery [1]. The principal objective in novel drug designing & development is to permeate novel chemical structures that have a peak probability of interacting partnership with targeted protein to obtain the appropriate biological function. To improve the speed and efficiency of the virtual screening is being increasingly used in drug discovery and development process [2]. Widely used High Throughput virtual screening tools are LigandFit[3], Dock[4], FlexX[5], Gold[6], Glide[7] and Fred[8]. The bulk of them are private enterprise packages whereas Fred & Dock are special cases that are freely accessible for educational cause.

Although, treaties that can irreproachable dock a large number of molecules are not standard for high-scoring ligands. Two crucial conditions for such a protocol running on a Linux System:

- The potential to initiate alongside docking tasks constantly a queuing models; and
- the potential to procedure millions of chemical structures in a sensible time.

Since the later demand may cry out for many of CPUs (central processing units) operating concurrently, the protocol should constructively handle the correlated data flow along the file system without influencing the production of the cluster.

An extensively applied docking program evolved is AutoDock at the Scripts Research Institute which needs some different pre-docking steps previous the genuine docking; include receptor & ligand preparation, and grid map computation [9]. AutoDock for high throughput virtual screening have shortfall to many causes like: Due to its approximately accurate docking algorithm and scoring function it consumes relatively more docking time. Few practical troubles also appear mainly for graphical user interface users. In AutoDock website [10], tutorial which teach user to perform virtual screening with help AutoDock using UNIX commands. From this tutorial, we manually develop various discrete pre-docking steps. One and all time the user has to replicate the protocol for docking parameter file generation & ligand preparation.

At a time, we report a Docking-based Virtual Screening Perl program, here shell & Perl code are used from AutoDock Tool, OpenBabel[11] and AutoDock to integrate executable. As consequence the docking procedure can't be left ignore by the client for any virtual screening assessment with the subsist attributes of AutoDock. In this current task each process are completely automated by a UNIX script. In addition to this characteristic like automated file conversion and automated all pre-docking step are also comprise. The principal points are to find a sensible solution to the numerous troubles normally connected with AutoDock and reduce user involvement during the docking procedure so that the process can run neglected.

2   Strategy

Tools Requirement:

This Perl code executes on both or window/linux stand in bash shell and require few pre-installed software such as python, MGLTools1.5.0, AutoDock4.0, AutoGrid4.0, OPENBABEL 2.2.0, DG-Ammos,& Perl programming platform.

Script Coding:

UNIX based Shell is used Perl programming in this present task. To make this benefit further convenient and easily to operate, a Dialog command is used to design the several dialog boxes at any moment needed. According to the MAIN DIALOG BOX (see Fig.1), client enter the option as

per situation and after complete task program return to main dialog box automatically. Where if you want to go head then select option otherwise exit.

1. Pre-docking

In this step, all the ligand (N no.) extracted (in sdf) form adequate data of curiosity molecules (from downloaded single file from pubchem) with the help of a python script use in this scripting (option 1 see fig. 1) which saved lots of time to make ligand molecule library. Further ligand converted to mol format (by Open Babel) (option 2) and optimize all the structure using "Dg-ammos.py"[12]. Next all ligands transformed to the Auto-Dock native form either mol or pdb into pdbqt by using a Python code, "prepare_ligand4.py" from Autodock Tools (option 3).The allowable input file format form for ligands cover sdf, mol & pdb. If the ligand file format is in the mol2 format, then transformed it into pdb by utilize tool OpenBabel. The receptor molecule is also transformed into pdbqt format by using a Python program, "prepare_receptor4.py" from Autodock Tools(option 4). It also create conf file for docking (option 5). The pre-docking step save reasonable time for further analysis. The consistent programs and work flow used to construct the program showed in figure (see Fig.2).

2. Docking

In this step if user has all pre-docking files and need only docking then it perform automatic docking for all the compounds.

3. Post-Docking

After completed docking, it generate energy file containing list of energy of all docked molecules for comparative study of docking molecules.
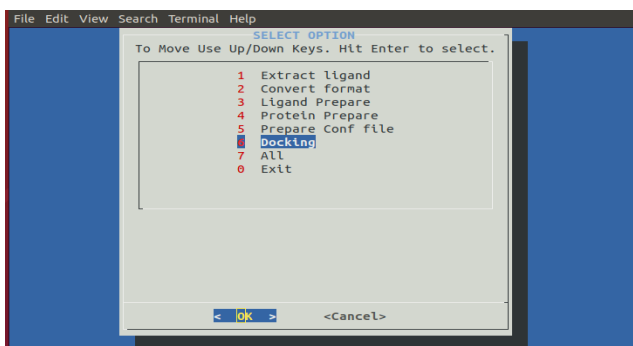

Fig. 1. How the script looks after execution.

3 Discussion

To validate Script, We executed a well ordered docking study with docking demonstration on a raw data of 608 compounds (Pueraria) obtained from Pubchem database against Peroxisome Proliferator-Activates Y (PPAR-Y) a potential target for ant diabetic drug. DocVSP showed a satisfactory performance when virtual screening performed on a single system locally. From the results, it is understandable that the precision of docking using protocol is approximate similar to the validity of docking using standard protocol of AutoDock vina tool. The docking trials illustrate that our protocol is algorithmically rapid and effortless. The alongside accomplishment of the protocol of SBDD form drug discovery extra smooth, rapidly and effortless. Nonetheless our protocol contain of collective docking potency, less algorithmically time is consume in every individual operation and this outcomes in bottommost docking hour generally.

Shivam Chauhan, Mani Srivastava, Jaikee Singh

DocVSP can efficiently perform all step of Drug Discovery and screen thousand of ligand per day and also generate Docking file which have list of docking energy to filter ligand molecular based on energy. DocVSP is an integrated and automated which removes most of complication and collective troubles related with extensive high outturn virtual screening and provides a convenient well organized outcome for SBDD user. However, the UGI and other enhancement to make single-stop platform for Drug Discovery are under development process.

4 Conclusion

DocVSP is a fully automated Perl script to perform all the step of structure base drug designing (SBDD) in such a way to overcome the limitations of AutoDock while performing virtual screening. This method separate many complication and collective troubles related with virtual screening and come up with a clear advantage and well organized outcome for AutoDock users on a Linux assemble programme. This script is appropriate for managing exhaustive high throughput virtual screening on Linux system. Furthermore, Integrating of Gromacs for molecular dynamics simulation and minimization energy before and after docking to complete the bridge between theory and experiments of Drug Discovery are under development phase.

- Availability: http://bio-kit.co/portfolio-details.html
- Tutorial link: https://www.youtube.com/watch?v=EA392-_n-Uw

Acknowledgment
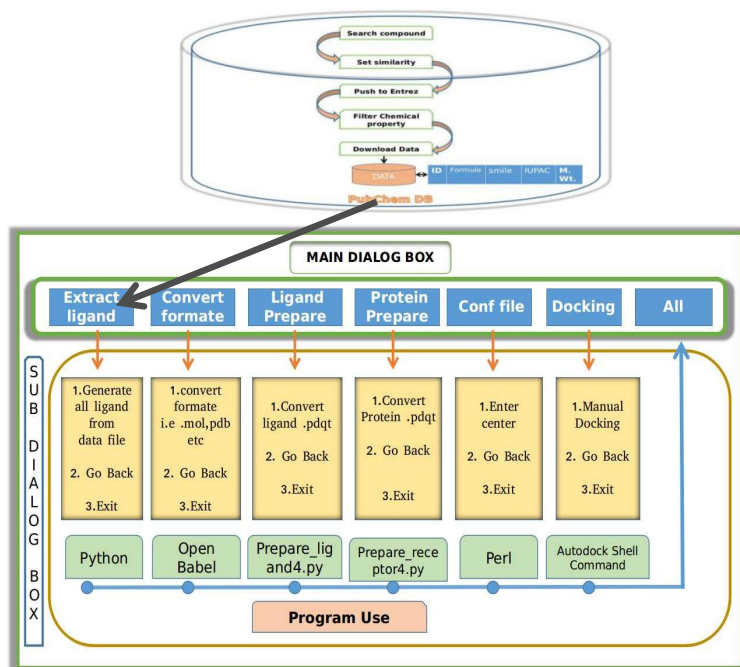
Fig.2. Flowchart indicating the several relative programs that used to build DocVSP program and Dialog boxes illustrates the work flow.

References

[1] Xuan, Y. M. et al. (2011). Molecular docking: a powerful approach for structure-based drug discovery. *Current Computer-Aided Drug Design*, 7(2):146-157.

[2] Sutapa, G. et al. (2006). Structure-based virtual screening of chemical libraries for drug discovery. *Current Opinion in Chemical Biology*, 10(3):194-202.

[3] Venkatachalam, C. M. et al. (2003). LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *Journal of Molecular Graphics and Modelling*, 21(4):289-307.

[4] Demetri, T. M. et al. (2006). Development and validation of a modular, extensible docking program: DOCK 5. *Journal of Computer-Aided Molecular Design*, 20(10-11):601–619.

[5] Kramer, B., Rarey, M. and Lengauer, T. (1999). Evaluation of the FLEXX incremental construction algorithm for protein–ligand docking. *Proteins*, 37(2):228–241.

[6] Jones, G. et al. (1997). Development and validation of a genetic algorithm for flexible docking. *Journal of Molecular Biology,* 267(3):727-748.

[7] Halgren, T. A. et al. (2004). Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *Journal of Medicinal Chemistry*, 47(7):1750–1759.

[8] Mark, R. M. et al. (2003). Gaussian docking functions. *Biopolymers*, 68(1):76-90.

[9] Forli, S. et al. (2016). Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nature Protocols*, 11:905–919.

[10] Morris, G. M. et al. (2014). User Guide AutoDock Version 4.2 Automated Docking of Flexible Ligands to Flexible Receptors.
https://autodock.scripps.edu/wp-content/uploads/sites/56/2021/10/AutoDock4.2.6_UserGuide.pdf

[11] O'Boyle, N. M. et al. (2011). Open Babel: An open chemical toolbox. *Journal of Cheminformatics*, 3:33.

[12] Lagorce, D. et al. (2009). DG-AMMOS: A New tool to generate 3D conformation of small molecules using Distance Geometry and Automated Molecular Mechanics Optimization for in Silico Screening. *BMC Chemical Biology*, 9:6.