

VSDK: Virtual screening of small molecules using AutoDock Vina on Windows platform

Natsumi Baba & Eiichi Akaho*

Faculty of Pharmaceutical Sciences, Kobe Gakuin University, 1-1-3, Minatojima, Chuo-ku, Kobe, Japan 650-8586; Eiichi Akaho - Email: akaho@pharm.kobegakuin.ac.jp; *Corresponding author

Received June 08, 2011; Accepted July 12, 2011; Published August 02, 2011

Abstract:

Screening of ligand molecules to target proteins using computer-aided docking is a critical step in rational drug discovery. Based on this circumstance, we attempted to develop a virtual screening application system, named VSDK Virtual Screening by Docking, which can function under the Windows platform. This is a user-friendly, flexible, and versatile tool which can be used by users who are familiar with Windows OS. The virtual screening performance was tested for an arbitrarily-selected receptor, FGFR tyrosine kinase (pdb code: 1agw), by using ligands downloaded from ZINC database with its grid size of x,y,z = 30,30,30 and run number of 10. It took 90 minutes for 100 molecules for this virtual screening. VSDK is freely available at the designated URL, and a simplified manual can be downloaded from VSDK home page. This tool will have a more challenging scope and achievement as the computer speed and accuracy are increased and secured in the future.

Availability: VSDK is freely available at VSDK home page for non-commercial use at: http://www.pharm.kobegakuin.ac.jp/~akaho/english_top.html

Background:

An increasing number of three-dimensional structures of macromolecules such as proteins and DNAs have been reported in recent years. As of July 5, 2011 there are 74,297 structures accumulated in RSCB Protein Data Bank and they are available to use for the scientific community. At the same time, structure databases in which numerous numbers of small molecules are stored and supplemented by physico-chemical properties, have been made available. One of those notable examples is ZINC (<http://zinc.docking.org/choose.shtml>). On the other hand, docking application tools, in which the ligand-macromolecule interaction conformation is predicted and scored in terms of binding free energy between the ligand and macromolecule, have made a remarkable progress in their accuracy and speed, although molecular docking is an extremely demanding task for computer resources. There is no question about the fact that the CAMD (computer-aided molecular design) approach is becoming a common practice in drug discovery and development [1]. Owing to all these advancements in structure data bases and application software, a high-through-put virtual screening technology, in which one can screen thousands of compounds for their binding affinity against a target macromolecule, has been developed and utilized to aid in the drug discovery. Indeed, virtual screening tools have been reported in the literature [2, 3]. However, there exist various constraints such as strict computer hardware requirements, demands for highly-technical computer knowledge, complex operation steps, and user-unfriendly procedures. Herein, we have developed a virtual screening application system which functions under the Windows environment. This is a user-friendly, flexible, and versatile tool which can be used by anybody who is familiar with Windows OS.

Computer environments:

VSDK is designed to run on any version of MS Windows in addition to Linux platform. The system is a console application for MS Windows platform.

Required input files and directories:

AutoDock is one of the most widely used docking application tool, and its use requires a set of preparation steps for general screening [4]. Included in the process are preparations of acceptable ligands and a receptor macromolecule, calculation of maps, creation of folders for each ligand, and so on. AutoDock Vina is a new program for molecular docking and virtual screening and achieved an approximately two orders of magnitude speed up compared to AutoDock4 [5]. At this point users face several difficulties to execute AutoDock Vina, and so we developed a user-friendly and flexible application tool for virtual screening based on AutoDock Vina. VSDK needs two preparation steps only: preparations of the receptor and ligands, and config file in which a grid center, a grid box size, and a docking run number are assigned (Figure 1). The virtual screening with a new receptor can simply be repeated by changing the receptor *.pdbqt file and modifying the config file accordingly. Users create a working directory (we use, as an example, c:/my_document/VS01) in which all necessary files will be saved. Users download a target macromolecule (*.pdb format) from the Protein Data Bank, and identify the grid center by using AutoDock Tools (ADT). Then the *.pdb format of the macromolecule should be converted to *.pdbqt format by using ADT, and the *.pdbqt format will be saved in the VS01 directory. For the ligands, users search and obtain small molecules from molecular databases such as ZINC. Finally, users create a conf.txt file which includes a receptor in *.pdbqt format, a grid center with x, y, z coordinates in Å, a grid box size in Å, and a docking run number, usually 10 or more. After converting ligands from *.mol2 format to *.pdbqt format, the ligands with *.pdbqt format will be saved in the VS01 directory. Users access VSDK, go to "Application Tool", select (copy) an appropriate bash file from the "bash file list", and paste it in the VS01 directory.

Details of algorithm and the execution of virtual screening:

The detailed flow chart of VSDK algorithm is shown in **Figure 1**. All pertinent ligand and receptor (macromolecule) files required for virtual screening along with conf.txt file will be saved in VS01 directory, and executed by a bash file (VS01.bash was used as an example). For the virtual screening execution, users open Cygwin, go to c/my_document/VS01, and input [./VS01.bash]. Upon the execution, new directory named "Data" is automatically created as an output file and the virtual screening result will be saved in it.

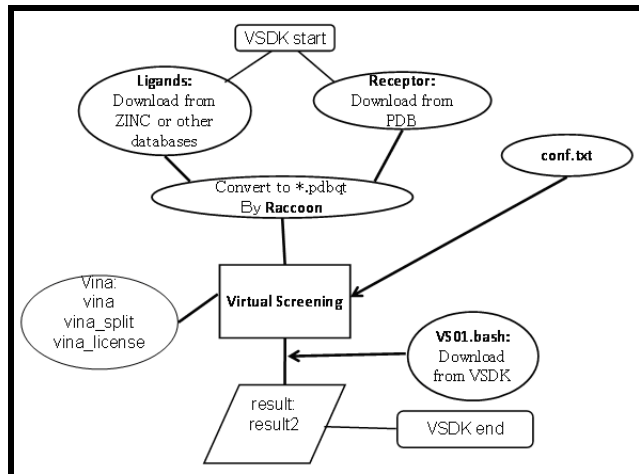


Figure 1: A flowchart showing VSDK algorithm

Output file processing:

A "result" file is automatically created in "Data" directory, and converted to Excel format which sorts and evaluates the result automatically.

Testing of VSDK:

A virtual screening was performed for small molecules (MW = 200s) downloaded from ZINC against FGFR tyrosine kinase (pdb code: 1agw). The specifications of Windows computer used are: model; NEC Value star, and cpu; Intel Pentium 4. It took one minute when a virtual screening for one molecule was performed with its grid size of x,y,z = 30 and run number of 10. Further investigation revealed that it took 10 minutes for 10 molecules, 90 for 100, and 451 for 1000. Another virtual screening was performed against EGFR tyrosine kinase (pdb code: 1m17) on a different computer whose specifications are: model; HP compaq, and cpu; AMD Sempron. It took about

700 minutes for 103 molecules with its grid size of x,y,z = 30 and run number of 10. Prakhov *et al.* reported that they achieved an average throughput virtual screening of about 420 molecules/CPU/day [3]. This means that the outcome of our system in the former case executed about 3.5 times more molecules in one day than their system. Since any virtual screening tools handles thousands of compounds obtained from huge databases of small molecules which may include flawed, or unusual datum, the virtual screening by this system on the Windows platform encounters an unexpected transaction which ends up with the aboard incidence, although this incidence rarely happens. This type of aboard incidence does not occur on the Linux platform. This is an operating system problem and we hope that an elaborated improvement of Windows OS will solve this kind of problem in the near future.

Discussion:

AutoDock Vina certainly contains the function of virtual screening. However, the procedure is highly complicated. It requires typing-in of the complicated 27 command lines, while VSDK needs only one line to type. Moreover, VSDK automatically create the file which sorts the virtual screened result in the descending order with the lowest ΔG at the top. The high through-put virtual screening method has been increasingly utilized along with molecular databases for the drug discovery and development [6, 7]. It is important for those tools to be user-friendly and easy to use. VSDK enables any user of a Windows computer to perform virtual screening of thousands of small molecules against a macromolecule of interest. It is an open access, flexible application tool, and users can modify it as needed. It was found that the more molecules for screening we have in this system the less time it is required per molecule. It is a challenge to combat against the aboard incidence and to evade it.

Acknowledgement:

We are grateful to Dr. Arthur J. Olson and Dr. Oleg Trott of the Scripps Research Institute, USA, for generous consideration and kindness toward this paper.

References:

- [1] Adane L & Bharatam PV. *J Mol Model*. 2011 **17**: 657 [PMID: 20524021]
- [2] Zang S *et al. BMC Bioinformatics* 2008 **9**: 126 [PMID: 18304355]
- [3] Prakhov ND *et al. Bioinformatics* 2010 **26**: 1374 [PMID: 20378556]
- [4] Morris GM *et al. J Comput Chem*. 2009 **30**: 2785 [PMID: 19399780]
- [5] Trott O & Olson AJ. *J Comput Chem*. 2010 **31**: 455 [PMID: 19499576]
- [6] Chang MW *et al. PLoS one*. 2010 **5**: 1 [PMID: 20694138]
- [7] Irwin JJ & Shoichet BK. *J Chem Inf Model*. 2005 **45**: 177 [PMID: 15667143]

Edited by P Kanguane

Citation: Baba & Akaho. *Bioinformation* 6(10): 387-388 (2011)
provided the original author and source are credited.

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes,