BIOINFORMATICS APPLICATIONS NOTE Vol. 26 no. 10 2010, pages 1374–1375
doi:10.1093/bioinformatics/btq149

Structural bioinformatics

Advance Access publication April 8, 2010

VSDocker: a tool for parallel high-throughput virtual screening using AutoDock on Windows-based computer clusters

Nikita D. Prakhov[1,2,*], Alexander L. Chernorudskiy[2,†] and Murat R. Gainullin[2,3]
[1]Faculty of Biology, Nizhny Novgorod State University, 23 Gagarin Avenue, 603950 Nizhny Novgorod, [2]Postgenomic Technology Group, Institute of Applied and Fundamental Medicine, Nizhny Novgorod State Medical Academy, 10/1 Minin Square, 603005 Nizhny Novgorod, Russia and [3]Centre of Ophthalmology and Visual Sciences, Biomedical Institute for Research in Light and Image (IBILI), Faculty of Medicine, University of Coimbra, Azinhaga de Santa Comba, 3000-354 Coimbra, Portugal

Associate Editor: Anna Tramontano

ABSTRACT

Summary: VSDocker is an original program that allows using AutoDock4 for optimized virtual ligand screening on computer clusters or multiprocessor workstations. This tool is the first implementation of parallel high-performance virtual screening of ligands for MS Windows-based computer systems.

Availability: VSDocker 2.0 is freely available for non-commercial use at http://www.bio.nnov.ru/projects/vsdocker2/

Contact: nikita.prakhov@gmail.com

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received on 27 December 2009; revised on 18 March 2010; accepted on 6 April 2010

1 INTRODUCTION

Virtual screening is a process of searching for small biologically active molecules (ligands) from large compound databases by means of computer-assisted techniques (Guido *et al.*, 2008). Virtual screening is based on molecular docking which currently plays an essential role both in the study of macromolecular structure and interaction and in drug design (Huey *et al.*, 2007). However, molecular docking is extremely demanding for computing resources; on the other hand, modern research often requires analysis of a large set of small ligands at acceptable time. This reason determines a wide use of high-throughput computing clusters and algorithms. AutoDock4 is one of the most widely used instruments for molecular docking (Morris *et al.*, 2009), but it is not adapted for parallel working on a multiprocessor cluster. An example of AutoDock adaptation for Linux cluster system is DOVIS (Zhang *et al.*, 2008). However, there was no similar utility for MS Windows systems to date. To address this problem, we have developed a new software tool termed VSDocker. This software possesses several features ensuring the increased performance of a screening process. VSDocker provides automation of all virtual screening steps, including ligand and receptor preparation, docking itself and analysis of results. Balanced distribution of cluster resources allows parallel screening of a large number of ligands. VSDocker works both on multiprocessor computing clusters and multiprocessor workstations operated by Windows. This makes possible an execution of virtual screening tasks even on a single high-performance multicore desktop that can be found virtually in every laboratory.

*To whom correspondence should be addressed.
†Present address: IFOM, the FIRC Institute for Molecular Oncology Foundation, Via Adamello 16, 20139 Milan, Italy.

2 DESCRIPTION

2.1 System requirements

VSDocker 2.0 is a virtual ligand screening utility using parallel execution set of AutoGrid4, AutoDock4 and original algorithms. The utility is a console application for Microsoft Windows platform. It is written in C++ with the use of parallel programming library mpi.h. VSDocker is developed for working on multiprocessor computing cluster or desktop that is operated by MPICH2 or MSMPI implementation of message passing interface.

VSDocker 2.0 is designed to run on any version of Windows NT operating system starting with Windows 2000, but it was tested only on Windows XP, Windows 2003 Server and Windows HPC Server.

2.2 Input

Use of AutoDock4 requires a set of preparation steps for general screening. This process includes preparation of ligands and a receptor, calculation of maps and creation of folders for each ligand. In contrast, VSDocker needs two preparation steps only: preparation of the receptor and AutoGrid parameter files.

As an input VSDocker uses *.pdb and *.pdbqt for receptor files, and *.gpf (AutoGrid Parameter Files) for the library. The screening with a new receptor can be simply repeated by changing 'input/Receptor' access path and preparing AutoGrid parameter files for the new receptor. The docking parameters of AutoDock4 can be changed by using specific command line keys described in VSDocker user manual.

2.3 Algorithm details

VSDocker 2.0 implements a special algorithm to use computer cluster resources (processor time, file system operation and network connection) with a maximum efficiency (Fig. 1). This algorithm loads all cluster components equally during the whole program operation time and includes several steps:

(i) Creation of receptor maps;
(ii) unpacking of ligands from multiligand containing *.mol2 format into separate files;
(iii) converting of *.mol2 to *.pdbqt;

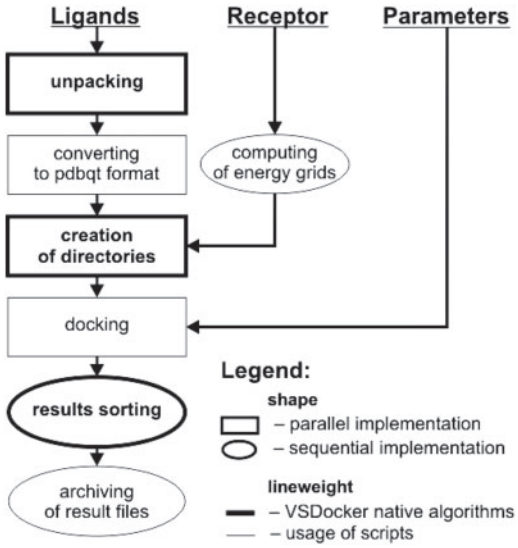

Fig. 1. Schematic diagram of VSDocker algorithm.

(iv) making of separate directories for the work of AutoDock;
(v) docking; and
(vi) analysis of results.

Each of these steps has its own requirements to the system resources, e.g. unpacking of ligands considering their large quantity and creation of directories may load the file system heavily, leaving majority of processors not involved. To avoid that situation, our algorithm executes steps 2, 3, 4 and 5 in parallel (see Supplementary Material Figure S1 for more details). Each active process sends a report about a completed subtask to the main process, which plays a coordinator role. Steps 1 and 6 are executed before and after main procedures, respectively, and do not take much time.

2.4 Output

AutoDock4 scripts create separate report file containing information about respective complexes for each ligand. VSDocker 2.0 algorithm analyses these files and combines data into a single complex list, sorted by the binding-free energy of complexes. The file named 'summarize_sort.rtf' containing the sorted results of virtual screening is created in the 'Output' directory after the finish of the screening. The resulting *.dlg files may be also found in the respective directories.

2.5 Testing

To test the technical performance and speed of VSDocker, we performed a virtual screening calculation of 86 775 ligands from the ZINC database (Irwin and Shoichet, 2005) against a 145-amino acid protein target (see Supplementary Material for details). The size of the binding site was $28 \times 40 \times 24$ Å. The related AutoDock parameters were set as follows: 10 ga runs, each with population size of 150, 250 000 energy evaluations and a maximum of 27 000 generations per ga run. The parameters were equal to those used by the developers of DOVIS 2.0, an implementation of for parallel screening using AutoDock4 on Linux-operated clusters (Jiang *et al.*, 2008). We achieved an average throughput of about 420 ligands/CPU/day, which is comparable to the performance of DOVIS in the analogous test.

3 CONCLUSIONS

VSDocker 2.0 is an utility for virtual screening using AutoDock4, suitable for parallel screening of large number of ligands, that significantly reduces analysis time. It is the first implementation for high-throughput virtual screening using AutoDock4 on Microsoft Windows-based computer systems. Program features include:

- Comprehensive automation of screening process;
- ability to use large sets of ligand structures (*.mol2 files), the number of ligands is practically unlimited;
- ability to set up AutoDock4 parameters via command line during VSDocker launch; and
- balanced algorithm of cluster resources distribution that allows achieving the maximum performance on any suitable computer system.

VSDocker works both on multiprocessor computing clusters and multiprocessor workstations operated by MPICH2 or MSMPI. It may be very helpful for variety of research tasks based on virtual screening, especially for time consuming applications.

We should note that molecular docking software is currently under constant development and improvement. This suggests great opportunities for further VSDocker modification, in order to use it in tandem with new software like AutoDock Vina, the latest docking tool from Olson's laboratory (Trott and Olson, 2010). This latter tool demonstrates very good performance in docking but cannot perform virtual screening in its current realization (see Supplementary Material for details), making it worthy to adapt VSDocker for 'team-work' with AutoDock Vina.

In conclusion, we should stress that VSDocker is an original valuable tool for virtual screening, as far as it is a first implementation for virtual screening on Windows-operated clusters, and there are no direct analogs to date.

ACKNOWLEDGEMENTS

We are grateful to metacluster program developer Andrey Senin for timely and competent support.

Funding: Russian Foundation for Basic Research (grant 08-04-01816a).

Conflict of Interest: none declared.

REFERENCES

Guido,R.V.C. *et al.* (2008) Virtual screening and its integration with modern drug design technologies. *Curr. Med. Chem.*, **15**, 37–46.

Huey,R. *et al.* (2007) A semiempirical free energy force field with charge-based desolvation. *J. Comput. Chem.*, **28**, 1145–1152.

Irwin,J.J. and Shoichet,B.K. (2005) ZINC — a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.*, **45**, 177–182.

Jiang,X. *et al.* (2008) DOVIS 2.0: an efficient and easy to use parallel virtual screening tool based on AutoDock 4.0. *Chem. Cent. J.*, **2**, 18.

Morris,G.M. *et al.* (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.*, **30**, 2785–2791.

Trott,O. and Olson,A.J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.*, **31**, 455–461.

Zhang,S. *et al.* (2008) DOVIS: an implementation for high-throughput virtual screening using AutoDock. *BMC Bioinformatics*, **9**, 126.