Nucleic Acids Research, 2013, Vol. 41, Web Server issue **W329–W332**
doi:10.1093/nar/gkt406

CovalentDock Cloud: a web server for automated covalent docking

Xuchang Ouyang[1,2], Shuo Zhou[1], Zemei Ge[1], Runtao Li[1,]* and Chee Keong Kwoh[2,]*

[1]State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Beijing, 100191, China and [2]BioInformatics Research Centre, School of Computer Engineering, Nanyang Technological University, Singapore, 639798, Singapore

Received February 16, 2013; Revised April 22, 2013; Accepted April 23, 2013

ABSTRACT

Covalent binding is an important mechanism for many drugs to gain its function. We developed a computational algorithm to model this chemical event and extended it to a web server, the CovalentDock Cloud, to make it accessible directly online without any local installation and configuration. It provides a simple yet user-friendly web interface to perform covalent docking experiments and analysis online. The web server accepts the structures of both the ligand and the receptor uploaded by the user or retrieved from online databases with valid access id. It identifies the potential covalent binding patterns, carries out the covalent docking experiments and provides visualization of the result for user analysis. This web server is free and open to all users at http://docking.sce.ntu.edu.sg/.

INTRODUCTION

The prediction of the interaction between the proteins and small molecules, often through molecular docking, serves as an important step in rational drug discovery pipeline (1,2). Despite the fact that these interactions can be either non-covalent or covalent, few of the current molecular docking programs can account the formation of covalent bond during docking properly (3–6), if not incapable of doing it at all. However, covalent binding is important and common. It is reported that 3 of the 10 top-selling drugs in USA in 2009 are covalent drugs (7).

To formulate and model covalent docking more accurately and to enable large-scale covalent virtual screening, we developed the software package CovalentDock (8). The package has an empirically calibrated energy model to estimate the contribution of the covalent binding. It can also automatically detect and process structural patterns to determine which covalent linkage is most probable. Although best effort for user-friendliness has been observed for the CovalentDock package, many users still experienced difficulties in compiling, installing or using the software owing to the diverse variation of their machines, especially for those who lack in such technical expertise. Moreover, owing to the fact that there are multiple options for the molecular structures to consider owing to chirality when forming a covalent linkage, the performance of running CovalentDock on an average workstation is not desirable.

With such limitation and desire to make the computational tool available to more users, we present a fully functional web server, the CovalentDock Cloud, for automated covalent docking. It allows the users to carry out covalent docking experiments with their own data and analyze the results via a web interface, without any configuration effort on their local machines. This web server is publicly available at http://docking.sce.ntu.edu.sg/. It is free and open to all users, and no login is required.

MATERIALS AND METHODS

Covalent docking with empirical energy estimation of covalent linkage formation

The CovalentDock Cloud web server uses the CovalentDock in its backend. CovalentDock (8) is a software package inherently considering covalent docking, which is capable of the automatic identification of the chemical patterns suitable for covalent linkage formation, pre-processing the structures by altering them to reflect the nature of covalent binding, taking into account of the chirality during structure change and molecular geometry constrains.

One unique contribution we developed for CovalentDock is the energy estimation for the formation of the covalent linkage. The interaction between the ligand and its receptor is still modeled the same as in conventional molecular dockings through non-covalent interactions such as van der Waals, electrostatics, solvent

---

*To whom correspondence should be addressed. Tel: +65 6790 6057; Fax: +65 6792 6559; Email: asckkwoh@ntu.edu.sg
Correspondence may also be addressed to Runtao Li. Tel/Fax: +86 10 82801504; Email: lirt@bjmu.edu.cn

effect or so. The additional energy contribution from covalent linkage formation is estimated by a newly formulated model:

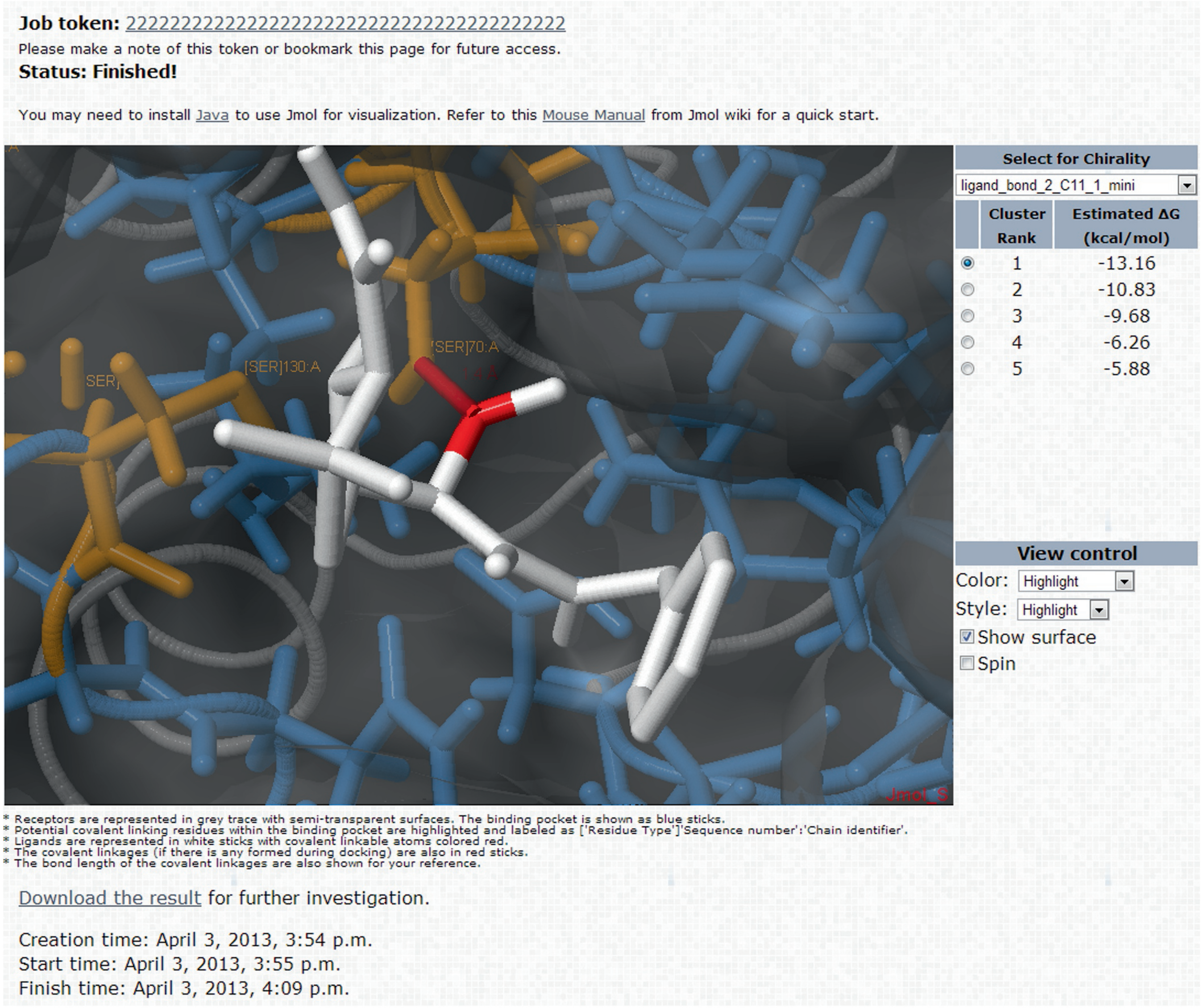
$$E = \begin{cases} D(e^{-2\alpha(r-r_0)} - 2e^{-\alpha(r-r_0)}) - T\Delta S + C, & r < r_m \\ 0, & r \geq r_m \end{cases}$$

where D is the dissociation energy, α is a parameter controlling the well width, r is the bond length, and r_0 is ideal equilibrium bond length, T is the system temperature, ΔS is the conformation entropy estimated by quantum mechanics simulation, C is an empirical correction constant to make it compatible with the energy contribution from non-covalent interaction and r_m is the maximum bond length without disassociation. The parameterization of this equation is subjected to the specified type of covalent bonding pairs.

This energy model, together with the automation of preprocessing, made the fully automated covalent docking and covalent virtual screening possible. Readers can refer to the CovalentDock article (8) for the description on the technical details and how it works.

Input

CovalentDock Cloud has a simple and easy-to-use interface for new job submission with user-specified data.



Job token: 2222222222222222222222222222222222222222

Please make a note of this token or bookmark this page for future access.

Status: Finished!

You may need to install Java to use Jmol for visualization. Refer to this Mouse Manual from Jmol wiki for a quick start.

Select for Chirality

ligand_bond_2_C11_1_mini

	Cluster Rank	Estimated ΔG (kcal/mol)
●	1	-13.16
○	2	-10.83
○	3	-9.68
○	4	-6.26
○	5	-5.88

View control

Color: Highlight

Style: Highlight

☑ Show surface

☐ Spin

* Receptors are represented in grey trace with semi-transparent surfaces. The binding pocket is shown as blue sticks.
* Potential covalent linking residues within the binding pocket are highlighted and labeled as ['Residue Type']'Sequence number':'Chain identifier'.
* Ligands are represented in white sticks with covalent linkable atoms colored red.
* The covalent linkages (if there is any formed during docking) are also in red sticks.
* The bond length of the covalent linkages are also shown for your reference.

Download the result for further investigation.

Creation time: April 3, 2013, 3:54 p.m.
Start time: April 3, 2013, 3:55 p.m.
Finish time: April 3, 2013, 4:09 p.m.

Figure 1. The screenshot of a typical result page from CovalentDock Cloud. The structure of the covalent docking result is visualized with Jmol plugin. The receptor structure is presented as gray backbone trace with the binding pocket highlighted in blue sticks, whereas the ligand is in white. All potential covalent linking residues in the binding pocket are specially colored by their residue type. The covalent linkage formed during docking between the ligand, and the receptor is emphasized in red. The length of this covalent bond is also marked on the canvas for reference. A simple control panel is attached on the right to allow the users to choose which configuration of the ligand to display and to show the rank and estimated free energy of each result cluster, as well as some options to alter the visualization for better inspection. The results are also available for downloading in case further investigation is necessary. Some other information about the covalent docking job such as the status and timing information is also available on this page. The all-two token in the screenshot is the result of deliberate manipulation to make it a special live sample. Real tokens are much more unpredictable and secure from competent forgery.

When creating a new covalent docking job, the user needs to specify the structures of both the receptor and the ligand. For receptors, user can submit them in as PDB files, while MOL2 files are accepted for ligands. The web site also provides an option for the users to retrieve structures from online databases. With valid access IDs, receptor structures can be fetched from RCSB PDB (9), and ligand structures can be retrieved from ZINC (10). As the underlying covalent docking protocol is not a blind-docking protocol, knowledge about the binding site is also required. The user only needs to specify the position of the binding site by the target residue identifier or more directly by the Cartesian coordinates of its center, and the size of the binding site will be automatically determined according to the ligand size.

The web server searches the uploaded structures for supported chemical patterns before proceeding to perform covalent docking. This mainly involves the identification of electrophile-nucleophile pairs with one in the ligand and the other in the receptor. Currently, the backend CovalentDock features the mechanism of the Michael additions and of the ring opening of β-lactams, both of which are prominent representatives of covalent binding. More chemical patterns will be available in near future as the CovalentDock package evolves.

Output

On successful submission of a new job, a unique token will be given. The token is generated as a salted 160-bit SHA-1 digest, which is proven to be secure from competent forgery. It is also the algorithm standardized and recommended by US government for collision resistance applications (11). The covalent docking job will be accessible if and only if a valid token is presented, preserving the confidentiality of the user data submitted to the server. All data submitted to the server will be kept on the server for a limited period (15 days for successful finished jobs and 24 h for jobs terminated due to errors) and destroyed afterward.

The backend CovalentDock will perform 10 independent covalent docking experiments for each job. The results will be clustered by their structural similarity. On successful job finishing, the results will be displayed in a Jmol plugin, with a visualization emphasis on the newly formed covalent linkage, if there is any. A control panel is also presented alongside the visualization window to allow the user the choice on which of the configuration results to be displayed. The results are also available for download, in case further investigation is necessary. The screenshot of a typical covalent docking result page is shown in Figure 1.

Web server implementation

The backend CovalentDock package is implemented in C++ and expanded from the source code of Autodock4.2 (3), accompanied with Python and bash scripts for automation. The web site is implemented with the Django web framework and MySQL database. The web server will perform input validation on job creation and send the data to the Nanyang Analytics HPC Cluster in Nanyang Technological University for actual docking experiments. Only the web server is open to the Internet, and the cluster is secured by the enterprise firewall for maximum security of the user data.

RESULTS

The backend CovalentDock package has been validated on a data set of 76 complexes discussed in literature, all of which are experimentally observed with covalent linkage formed between the ligand and the receptor. Among the 76 complexes, 13 are with Michael acceptors, and the rest are in β-lactam family.

The benchmark results suggest that CovalentDock is highly capable of producing accurate covalent docking results. It achieved an average root-mean-square deviation (RMSD) of 1.68 Å when comparing the predicted and the native ligand structure, significantly outperforming the average RMSD of 2.49 Å and 3.69 Å achieved by the covalent docking methods from Autodock (3) and GOLD (4), respectively (8). CovalentDock also achieved a better hit rate (the percentage of results with an RMSD within a given threshold) when compared with the other two covalent docking protocols, as shown in Figure 2 where the hit rate is calculated with different thresholds (8).

In addition, detailed analysis also suggests that CovalentDock has a strong chirality and target selectivity during covalent docking, and it preserves the molecular geometry of the structures very well. Readers can refer to the CovalentDock article (8) for detailed description on validation and case studies.

CONCLUSION

The prediction and quantification of covalent linkage formation in molecular docking is of great interest and

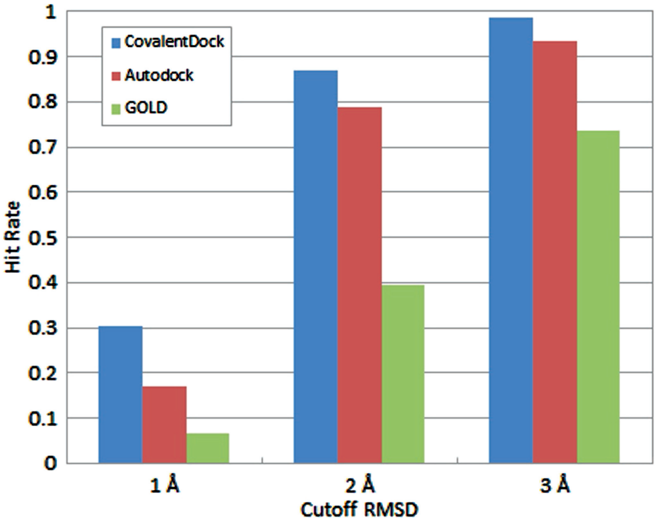


Figure 2. The hit rate of CovalentDock, Autodock and GOLD on the data set of 76 covalently bound complexes under different threshold RMSD. The hit rate under a given threshold RMSD is calculated as the percentage of the results with an RMSD smaller than the threshold.

potential to the systematic discovery of covalent drugs. The CovalentDock Cloud provides a simple yet user-friendly portal to carry out covalent docking experiments and to analyze the result online via a web browser. With the powerful backend docking engine, we believe that this web server will offer a more accessible way for simulation and prediction of more accurate covalent docking.

In the future, it is planned to fine-tune the web site for better user-experience and to continuously develop and enhance the CovalentDock package. It is also planned to make more reactions adopting covalent binding mechanism available and allow users to specify the covalent linking pattern by themselves.

ACKNOWLEDGEMENTS

The authors thank the High Performance Computing Centre in Nanyang Technological University for providing and maintaining the computing facility to support this web server. The gratitude also goes to the users of this web server, especially to those who reported bugs or submitted feedback to us.

FUNDING

Jardine OneSolution (2001) Pte Ltd [4060]; the National Natural Science Foundation of China [21172011]. Funding for open access charge: Nanyang Technological University and Peking University.

Conflict of interest statement. None declared.

REFERENCES

1. Kitchen,D.B., Decornez,H., Furr,J.R. and Bajorath,J. (2004) Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discov.*, **3**, 935–949.
2. Sousa,S.F., Fernandes,P.A. and Ramos,M.J. (2006) Protein-ligand docking: current status and future challenges. *Proteins*, **65**, 15–26.
3. Morris,G.M., Huey,R., Lindstrom,W., Sanner,M.F., Belew,R.K., Goodsell,D.S. and Olson,A.J. (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.*, **30**, 2785–2791.
4. Verdonk,M.L., Cole,J.C., Hartshorn,M.J., Murray,C.W. and Taylor,R.D. (2003) Improved protein-ligand docking using GOLD. *Proteins*, **52**, 609–623.
5. De Cesco,S., Deslandes,S., Therrien,E., Levan,D., Cueto,M., Schmidt,R., Cantin,L.D., Mittermaier,A., Juillerat-Jeanneret,L. and Moitessier,N. (2012) Virtual screening and computational optimization for the discovery of covalent prolyl oligopeptidase inhibitors with activity in human cells. *J. Med. Chem.*, **55**, 6306–6315.
6. Fradera,X., Kaur,J. and Mestres,J. (2004) Unsupervised guided docking of covalently bound ligands. *J. Comput. Aided Mol. Des.*, **18**, 635–650.
7. Singh,J., Petter,R.C., Baillie,T.A. and Whitty,A. (2011) The resurgence of covalent drugs. *Nat. Rev. Drug Discov.*, **10**, 307–317.
8. Ouyang,X., Zhou,S., Su,C.T., Ge,Z., Li,R. and Kwoh,C.K. (2013) CovalentDock: automated covalent docking with parameterized covalent linkage energy estimation and molecular geometry constraints. *J. Comput. Chem.*, **34**, 326–336.
9. Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235–242.
10. Irwin,J.J., Sterling,T., Mysinger,M.M., Bolstad,E.S. and Coleman,R.G. (2012) ZINC: a free tool to discover chemistry for biology. *J. Chem. Inf. Model.*, **52**, 1757–1768.
11. US Department of Commerce. (1995) Secure hash standard. *Federal Information Processing Standards Publication 180-1.*