Pakistan J. Zool., vol. 54(4), pp 1577-1582, 2022 DOI: https://dx.doi.org/10.17582/journal.pjz/20210109170113

Let's Dock: A Smart Platform for Molecular Docking

Yaruq Jabeen[1], Nida Ansari[1], Haroon Rasheed[1], Muhammad Asif Rasheed[1,*], Muhammad Awais[2], Muhammad Ibrahim[1], Sumaira Kanwal[1], Aqsa Khalid[3], Manzoor Ahmad Zahid[4] and Farrukh Jamil[1,*]

[1]*Department of Biosciences, COMSATS University, Sahiwal Campus, Punjab*
[2]*Department of Biochemistry and Molecular Biology, Faculty of Sciences, University of Sialkot, Sialkot*
[3]*National University of Sciences and Technology, H-12, Islamabad*
[4]*Department of Mathematics, COMSATS University, Sahiwal Campus, Punjab*

ABSTRACT

In the past few years, several docking tools have been developed for molecular dockings such as LeDock, rDock, AutoDock Vina, AutoDock, UCSF DOCK, GOLD, Glide, Surflex-Dock, LigandFit and MOE-Dock. Among these, AutoDock Vina has been widely used by academia for molecular docking. This tool performs docking of a ligand into a protein by seven steps. For docking analysis of several ligands, AutoDock Vina is a time-consuming tool. In order to make AutoDock Vina more efficient and smart way, a platform has been developed in this study. It is designated "Let's Dock". By using this platform, we can perform a docking analysis of several ligands into a protein with AutoDock Vina by just submitting ligands and protein pdb files. Apart from thermodynamic data analysis this platform also calculates descriptors and Lipinski's rule parameters (Rule of Five) of a ligand by a single click.

Article Information
Received 09 January 2021
Revised 03 March 2021
Accepted 18 March 2021
Available online 10 September 2021 (early access)
Published 07 April 2022

Authors' Contribution
YJ, NA, HR, MAR and AK Performed the experiment. MA, MI, SK, MAZ and FJ reviewed the manuscript. FJ designed and supervised the study.

Key words
Molecular docking, AutoDock Vina, Virtual screening, Lipinski rule.

INTRODUCTION

Molecular docking is a technique which is widely used in pharmacological target identification by virtual screening of different available natural and organic compounds for a specific protein (Forli *et al*., 2016; Gupta *et al*., 2018). The docking analysis is based on the identification of different ligands conformations in the active site of the protein and the ranking of these conformations according to binding affinity (Meng *et al*., 2011). It provides us valuable thermodynamic data of protein-ligand interactions for suitable ligand identification against target protein (Alvarez-Garcia and Barril, 2014; Uehara and Tanaka, 2016).

Studies have shown that molecular docking can be classified into two classes. First is targeted docking in which the binding sites are known for a ligand. This way expressively increases the docking efficacy (Meng *et al*., 2011). Second is blind docking in which there is no information on binding sites available for receptor and ligand. So, a ligand is docked on the entire surface of a protein (Lee and Zhang, 2012). There are also many online servers like GRID (Goodford, 1985; Kastenholz *et al*., 2000), POCKET (Levitt and Banaszak, 1992), SurfNet (Glaser *et al*., 2006; Laskowski, 1995), PASS (Brady and Stouten, 2000) and MMC (Mezei, 2003) through which putative active sites can be identified in a protein 3D structure (Meng *et al*., 2011).

Over the last two decades, there are almost more than 60 different docking programs have been developed for academic and commercial use. Some important programs are GOLD (Jones *et al*., 1997), DOCK (Venkatachalam *et al*., 2003), LigandFit (Venkatachalam *et al*., 2003), MOE-Dock (Corbeil *et al*., 2012), AutoDock Vina (Trott and Olson, 2010), AutoDock (Österberg *et al*., 2002), LeDock (Zhao and Caflisch, 2013). These docking programs are tested with several different complexes and have been validated (Bodian *et al*., 1993; Debnath *et al*., 1999; Friesner *et al*., 2004; Jones *et al*., 1997; Österberg *et al*., 2002; Rarey *et al*., 1996; Shoichet *et al*., 1993; Trott and Olson, 2010; Venkatachalam *et al*., 2003). Analysis of top scored programs have shown that performance of the academic programs follows this order: LeDock (57.4%) > rDock (50.3%) - AutoDock Vina (49.0%) > AutoDock (PSO) (47.3%) > UCSF DOCK (44.0%) > AutoDock

* Corresponding authors: farrukh@cuisahiwal.edu.pk; farrukhccb@gmail.com; asif.resheed@cuisahiwal.edu.pk
0030-9923/2022/0004-1577 $ 9.00/0

(LGA) (37.4%). The performance of the commercial programs follow this order: GOLD (59.8%) > Glide (XP) (57.8%) > Glide (SP) (53.8%) > Surflex-Dock (53.2%) > LigandFit (46.1%) > MOE-Dock (45.6%) (Pagadala *et al*., 2017). These results showed AutoDock Vina, GOLD, and MOE-Dock considered highest with best scores of docking (Wang *et al*., 2016).

For academic purpose, mostly used docking tool is AutoDock Vina which is an open-source tool and perform docking through seven number of steps (Gaillard, 2018; Helgren and Hagen, 2017; Topaz *et al*., 2019). These steps are: preparing ligand coordinate file, reading atomic coordinates, preparation of pdbqt file, preparing receptor coordinate file, preparing a configuration file for Vina, run Vina to start docking and visualization of results. This method interprets with single ligand and single protein molecule (Forli *et al*., 2016). In order to analyze several ligands with a protein, we need considerable time for repeating seven-step procedure for each pair of protein and ligand. To reduce the analysis time and human efforts, a docking platform is developed in this study. It is designated as "Let's Dock". Apart from thermodynamic data this platform also calculates descriptors and Lipinski's rule parameters (Rule of Five) of a ligand by a single click.

MATERIALS AND METHODS

Standard inputs required for Let's Dock are: co-ordinate file (.pdb) for protein, co-ordinate file (.sdf) for ligands and grid parameters. Grid parameters manually set by the user to provide ease, as every user didn't prefer blind docking.

Integration and docking

Integration of MGL Tools and Auto Dock Vina has been implemented by using python subprocess module (Dale *et al*., 2011). A python script, consist of three sub pieces, has been written which automate Open Babel, MGL Tools and Auto Dock Vina. The first piece of the script runs MGL Tools in the background and commands it to take a protein input file and ligand input files to convert into pdbqt format, and place them into the user-defined folder. This script later runs Auto Dock Vina in the background to retrieve pdbqt files of ligand and protein from the folder and dock them. After docking, it will place results files *i.e.* conf file, log file and out file in the same folder.

Formation of best protein-ligand complex

The second piece of this script generates a protein-ligand complex file by taking the best model of ligand, having the lowest binding affinity, from out file (having ligand's conformations model) and merge the best ligand model co-ordinates with protein co-ordinates. These all files are placed in a single folder which is named as same as of ligand file.

Table I.- Physical properties of the 14 ligands docked by Let's Dock in NS3-NS2B. The ligands are arranged in ascending order of binding energies.

Ligands	Binding affinity (kcal/mol)	Molecular weight (Da) <= 500.0	Rotatable bonds <= 3.0	H-bond donor <= 5.0	H-bond acceptor <= 10.0	Log P <= 5.0
1111153	-8.8	433.284	3.0	0.0	5.0	4.622
1111008	-8.5	471.313	5.0	0.0	6.0	4.382
42877561	-8.5	436.345	4.0	0.0	5.0	5.27
17021741	-8.5	431.292	5.0	2.0	4.0	4.387
17021740	-8.3	450.29	5.0	2.0	4.0	4.526
5126841	-8.2	478.3	4.0	0.0	6.0	4.525
4434656	-7.7	487.312	5.0	0.0	7.0	4.395
68956077	-7.7	407.27	2.0	0.0	5.0	3.769
4768896	-7.6	364.25	5.0	2.0	4.0	3.474
44624874	-7.4	379.256	4.0	3.0	1.0	4.294
67132806	-7.4	433.284	6.0	1.0	5.0	4.358
68956356	-7.2	442.723	2.0	0.0	5.0	4.422
67133299	-7.1	423.269	6.0	1.0	6.0	3.638
24631700	-6.5	314.243	5.0	2.0	4.0	2.239

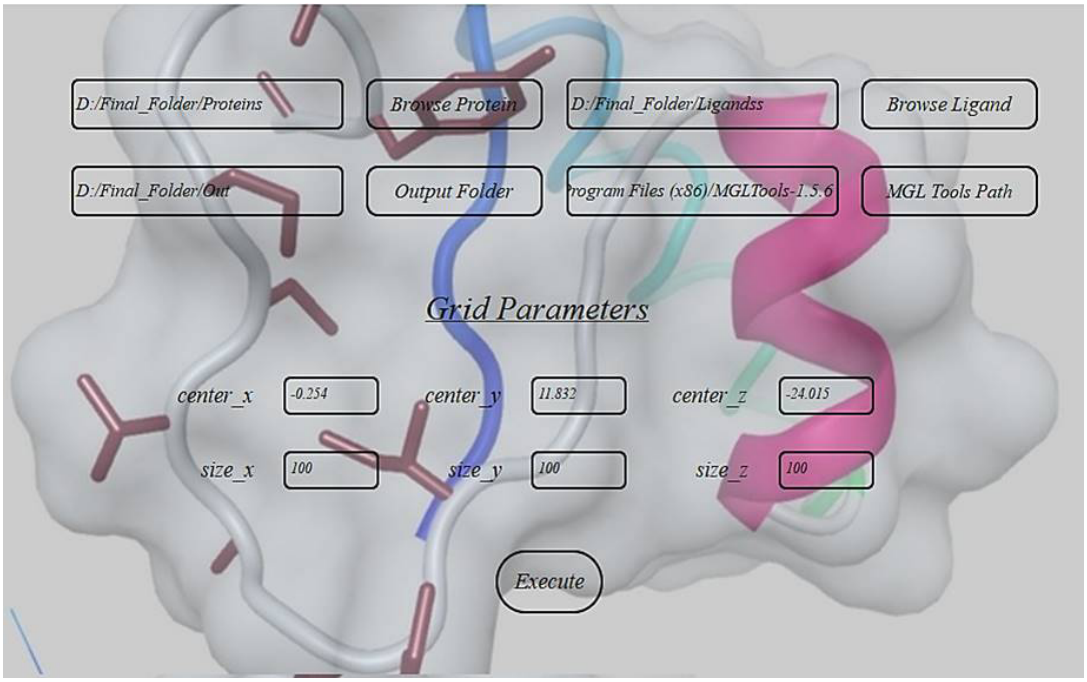

Fig. 1. Input interface of Let's Dock. "Browse Protein" is to select the folder of the protein of interest, "Browse ligands" is to select the folder of ligands and "Output Folder" is to select the folder to save the results.

Descriptor calculation

The third piece of script use built-in library PYDPI (Cao *et al*., 2013) to calculate descriptors for Rule of Five of ligands. Before calculation, a module runs Open Babel automatically in the background to convert the molecular file format (O'Boyle *et al*., 2011). Moreover, it determines five properties that cover the "Rule of Five" and highlight them green if the properties fall within the threshold values defined by Lipinski *et al*. (2001).

For a case study a pdb file NS3-NS2B protease of Zika virus and 14 ligands, PubChem IDs are shown in Table 1, has been submitted to Let's Dock as an input. Grid parameters are set with dimension of 100×100×100 Å and center x = -0.254, y = 11.832 and z = -24.015. By executing, we obtain the list of descriptors which is shown in Table I.

RESULTS AND DISCUSSION

A platform has been developed to perform docking and to calculate Rule of Five from the protein pdb file and Ligand mol files. The input interface of Let's Dock is illustrated in Figure 1. This figure shows different buttons that take inputs for the docking process. On click: Browse protein takes folder having protein of interest in pdb format, Browse ligands takes folder having single or multiple ligands of interest in mol format, Output folder takes the folder in which user want to store results, MGL Tools path takes the path of MGL Tools folder which will create when it is installed. MGL Tools used to convert input files into .pdbqt by adding different charges. Grid parameters take the set parameters against a defined pocket of the protein. Then, by clicking on execute button docking will start.

Docking and calculations

We submitted a zika virus NS3-NS2B protease and 14 ligands to Let's Dock. Its results are generated showing the binding affinities of the ligands with the NS3-NS2B docking site in terms of kcal/mol and five parameters of Rule of Five as shown in Table I.

The first column shows the PubChem IDs of the ligands and the second column shows the binding affinity; thermal stability and strength of the interaction between the binding protein and ligand. The resultant binding affinities range from -8.8 to -6.5 Kcal/mol. This tool also arranges the result according to the ascending order of binding affinity. The first molecule with the lowest binding affinity of -8.8 kcal/mol is the most stable lead compound out of the given input. A previous study on the same protein also reported a score of -9.9 kcal/mol for the strongest interacting molecule (Fatima *et al*., 2018).

The next five columns (3-7) enlist the five parameters

of Lipinski's rule. This is based on observations that most orally active drugs are relatively small and lipophilic molecules (Lipinski *et al*., 2001). These rules are a set of guidelines for the structural properties of ligand molecules that have the ability of well absorbed after intake. The third column shows the molecular weight which according to Rule of Five should be less than or equal to 500 Da. An increase in molecular weight results in the formation of a large cavity in water to solubilize the ligand molecule and then solubility decreases. The fourth column shows the rotatable bonds, a single bond bound to a nonterminal heavy atom, which should be less than and equal to 3. The fifth column displays the number of hydrogen donors which should be less than or equal to 5. The sixth column displays the number of hydrogen acceptors which should be less than or equal to 10. An increasing number in hydrogen bonds decreases partitioning from the aqueous phase into lipid bilayer membrane for permeation by passive diffusion. The last column shows LogP, partition coefficient, which should be less than or equal to 5. Increasing LogP decreases aqueous solubility, which reduces absorption. Let's Dock has colored green to these parameters which are in the allowed range (Table I). Due to this feature, the user can easily analyze which ligand is optimum. Analysis of the data showed that 1st, 8th and 12th ligand fulfill all the requirements.

Table II.- Comparison of binding affinities calculated by Let's Dock and conventional method.

Ligands	Let's Dock	Conventional method
1111153	-8.8	-8.8
1111008	-8.5	-8.4
42877561	-8.5	-8.3
17021741	-8.5	-8.2
17021740	-8.3	-8.3
5126841	-8.2	-8
4434656	-7.7	-7.9
68956077	-7.7	-7.8
4768896	-7.6	-7.8
44624874	-7.4	-7.5
67132806	-7.4	-7.5
68956356	-7.2	-7.3
67133299	-7.1	-7
24631700	-6.5	-6.3

We validated the software working by comparing the docking results of the same 14 ligands docking on the same protein (zika virus NS3-NS2B protease) by using a conventional method involving each step to perform manually. A comparison of resultant binding affinities calculated by the conventional method of docking and Let's Dock is given in Table II. Analysis of comparison showed that percentage difference ranges from -2.63 to 3.53 %. This difference is due to two reasons. First, our software prepare the protein and ligands files by default parameters using MGL Tools. Secondly, Vina is predicting software that is run by Let's Dock for docking in back ground. Vina can perform analysis with a single ligand where Let's Dock can perform multiprocessing with reliability.

CONCLUSION

In this study we have created a platform for several ligand docking in a protein that can perform docking analysis faster than conventional AutoDock Vina. Most importantly, it calculates binding energies and Rule of Five parameters of the ligand.

Statement of conflict of interest

The authors have declared no conflict of interests.

REFERENCES

Alvarez-Garcia D. and Barril, X., 2014. Molecular simulations with solvent competition quantify water displaceability and provide accurate interaction maps of protein binding sites. *J. Med. Chem*., **57**: 8530-8539. https://doi.org/10.1021/jm5010418

Bodian, D.L., Yamasaki, R.B., Buswell, R.L., Stearns, J.F., White, J.M. and Kuntz, I.D., 1993. Inhibition of the fusion-inducing conformational change of influenza hemagglutinin by benzoquinones and hydroquinones. *Biochemistry*, **32**: 2967-2978. https://doi.org/10.1021/bi00063a007

Brady, G.P. and Stouten, Jr. P.F., 2000. Fast prediction and visualization of protein binding pockets with PASS. *J. Computer-Aided Mol. Design*, **14**: 383-401. https://doi.org/10.1023/A:1008124202956

Cao, D.S., Liang, Y.Z., Yan, J., Tan, G.S., Xu, Q.S. and Liu, S., 2013. PyDPI: freely available python package for chemoinformatics, bioinformatics, and chemogenomics studies. *J. Chem. Inform. Model*, **53**: 3086-3096. https://doi.org/10.1021/ci400127q

Corbeil, C.R., Williams, C.I. and Labute, P., 2012. Variability in docking success rates due to dataset preparation. *J. Computer-Aided Mol. Design*, **26**: 775-786. https://doi.org/10.1007/s10822-012-9570-1

Dale, R.K., Pedersen, B.S. and Quinlan, A.R., 2011. Pybedtools: a flexible Python library for manipulating genomic datasets and annotations. *Bioinformatics*, **27**: 3423-3424. https://doi.org/10.1093/bioinformatics/btr539

Debnath, A.K., Radigan, L. and Jiang, S., 1999. Structure-based identification of small molecule antiviral compounds targeted to the gp41 core structure of the human immunodeficiency virus type 1. *J. Med. Chem*., **42**: 3203-3209. https://doi.org/10.1021/jm990154t

Fatima, M., Khan M.S., Rashid, H., Mehmood, A., Kanwal, S., Rasheed, M.A. and Jamil, F., 2018. Structure based virtual screening and molecular docking studies for identification of allosteric inhibitors against zika virus. *Pakistan J. Zool*., **50**: 1709-1715. https://doi.org/10.17582/journal.pjz/2018.50.5.1709.1715

Forli, S., Huey, R., Pique, M.E., Sanner, M.F., Goodsell, D.S. and Olson, A.J., 2016. Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nat. Protoc*., **11**: 905-919. https://doi.org/10.1038/nprot.2016.051

Friesner, R.A., Banks, J.L., Murphy, R.B., Halgren, T.A., Klicic, J.J., Mainz, D.T., Repasky, M.P., Knoll, E.H., Shelley, M., Perry, J.K., Shaw, D.E., Francis, P. and Shenkin, P.S., 2004. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem*., **47**: 1739-1749. https://doi.org/10.1021/jm0306430

Gaillard, T., 2018. Evaluation of AutoDock and AutoDock Vina on the CASF-2013 Benchmark. *J. Chem. Inform. Model*, **58**: 1697-1706. https://doi.org/10.1021/acs.jcim.8b00312

Glaser, F., Morris, R.J., Najmanovich, R.J., Laskowski, R.A. and Thornton, J.M., 2006. A method for localizing ligand binding pockets in protein structures. *Proteins*, **62**: 479-488. https://doi.org/10.1002/prot.20769

Goodford, P.J., 1985. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.*, **28**: 849-857. https://doi.org/10.1021/jm00145a002

Gupta, M., Sharma, R. and Kumar, A., 2018. Docking techniques in pharmacology: How much promising? *Comput. Biol. Chem*., **76**: 210-217. https://doi.org/10.1016/j.compbiolchem.2018.06.005

Helgren, T.R. and Hagen, T.J., 2017. Demonstration of AutoDock as an educational tool for drug discovery. *J. Chem. Edu*., **94**: 345-349. https://doi.org/10.1021/acs.jchemed.6b00555

Jones, G., Willett, P., Glen, R.C., Leach, A.R. and Taylor, R., 1997. Development and validation of a genetic algorithm for flexible docking. *J. mol. Biol*., **267**: 727-748. https://doi.org/10.1006/jmbi.1996.0897

Kastenholz, M.A., Pastor, M., Cruciani, G., Haaksma, E.E. and Fox, T., 2000. GRID/CPCA: a new computational tool to design selective ligands. *J. Med. Chem*., **43**: 3033-3044. https://doi.org/10.1021/jm000934y

Julia, R.K, Webe, K., Roberts, R., Gehm, M. and Craig, P.A., 2019. Predicting and testing enzyme function in the undergraduate lab using computational and wet lab tools. *Biophys. J*.., **116**: 450a. https://doi.org/10.1016/j.bpj.2018.11.2426

Laskowski, R.A., 1995. SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J. Mol. Graph*., **13**: 323-330. https://doi.org/10.1016/0263-7855(95)00073-9

Lee, H.S. and Zhang, Y., 2012. BSP-SLIM: a blind low-resolution ligand-protein docking approach using predicted protein structures. *Proteins*, **80**: 93-110. https://doi.org/10.1002/prot.23165

Levitt, D.G. and Banaszak, L.J., 1992. POCKET: a computer graphics method for identifying and displaying protein cavities and their surrounding amino acids. *J. Mol. Graph*., **10**: 229-234. https://doi.org/10.1016/0263-7855(92)80074-N

Lipinski, C.A., Lombardo, F., Dominy, B.W. and Feeney, P.J., 2001. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev*., **46**: 3-26.

Meng, X.Y., Zhang, H.X., Mezei, M. and Cui, M., 2011. Molecular docking: a powerful approach for structure-based drug discovery. *Curr. Computer-Aided Drug Design*, **7**: 146-157. https://doi.org/10.2174/157340911795677602

Mezei, M., 2003. A new method for mapping macromolecular topography. *J. Mol. Graph. Model*., **21**: 463-472. https://doi.org/10.1016/S1093-3263(02)00203-6

O'Boyle, N.M., Banck, M., James, C.A., Morley, C., Vandermeersch, T. and Hutchison, G.R., 2011. Open Babel: An open chemical toolbox. *J. Cheminform*., **3**: 33. https://doi.org/10.1186/1758-2946-3-33

Osterberg, F., Morris, G.M., Sanner, M.F., Olson, A.J. and Goodsell, D.S., 2002. Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in AutoDock. *Proteins*, **46**: 34-40. https://doi.org/10.1002/prot.10028

Pagadala, N.S., Syed, K. and Tuszynski, J., 2017. Software for molecular docking: a review. *Biophys. Rev*., **9**: 91-102. https://doi.org/10.1007/s12551-016-0247-1

Rarey, M., Kramer, B., Lengauer, T. and Klebe, G., 1996. A fast flexible docking method using an incremental construction algorithm. *J. mol. Biol*., **261**: 470-489. https://doi.org/10.1006/jmbi.1996.0477

Shoichet, B.K., Stroud, R.M., Santi, D.V., Kuntz, I.D. and Perry, K.M., 1993. Structure-based discovery of inhibitors of thymidylate synthase. *Science*, **259**: 1445-1450. https://doi.org/10.1126/science.8451640

Topaz, R.G., Sosa, H., Edward, E., Anderson, R. and Stieglitz, K., 2019. High throughput virtual screening and mutational studies using docking and molecular dynamics on a shared memory GPU Supercomputer. *FASEB J*., **33**: 468. https://doi.org/10.1096/fasebj.2019.33.1_supplement.468.8

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. https://doi.org/10.1002/jcc.21334

Uehara, S. and Tanaka, S., 2016. AutoDock-GIST: Incorporating thermodynamics of active-site water into scoring function for accurate protein-ligand docking. *Molecules*, **21**: 1604. https://doi.org/10.3390/molecules21111604

Venkatachalam, C.M., Jiang, X., Oldfield, T. and Waldman, M., 2003. LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *J. Mol. Graph. Model*., **21**: 289-307. https://doi.org/10.1016/S1093-3263(02)00164-X

Wang, Z., Sun, H., Yao, X., Li, D., Xu, L., Li, Y., Tian, S. and Hou, T., 2016. Comprehensive evaluation of ten docking programs on a diverse set of protein-ligand complexes: the prediction accuracy of sampling power and scoring power. *Phys. Chem. Chem. Phys*., **18**: 12964-12975. https://doi.org/10.1039/C6CP01555G

Zhao, H. and Caflisch, A., 2013. Discovery of ZAP70 inhibitors by high-throughput docking into a conformation of its kinase domain generated by molecular dynamics. *Bioorg. Med. Chem. Lett*., **23**: 5721-5726. https://doi.org/10.1016/j.bmcl.2013.08.009