International Journal of Computational Bioinformatics and In Silico Modeling
Vol. 5, No. 2 (2016): 787-792
Research Article
Open Access

ISSN: 2320-0634

Molecular docking analysis and screening of plant compounds against lung cancer target EGFR T790M mutant

Priyanka Maiti[1#], Mahesha Nand[2#], Hema Joshi[1] and Subhash Chandra[1]*

[1] *Department of Botany, Kumaun University, S.S.J Campus, Almora, Uttarakhand, India.*
[2] *Department of Botany, Kumaun University, S.S.J Campus, Almora, Uttarakhand, India.*

*Corresponding author: Subhash Chandra; email: subhashku@kunainital.ac.in
# Authors contributed equally

Received: 21 January 2016 Accepted: 16 February 2016 Online: 01 March 2016

ABSTRACT

The aim of present study is to identify anti-lung cancer compounds against EGFR 696-1022 T790M mutant from five medicinal plants. In house library of 70 phytochemicals was screened by molecular docking using AutoDock Vina and by CDRUG server. 11 phytochemicals were screened with efficient range of negative binding energy and high probability to act as anticancer drugs. Finally 8 phytochemicals were obtained as hit with accepted results in terms of drug likeness (FAF-Drug3 server) and cytotoxicity on lung cancer cell lines (CLC-Pred). Thus, it can be concluded that out of 70, 8 hits has potential therapeutic activity against lung cancer.

Keywords: Lung cancer inhibitors, Docking, CDRUG, FAF-Drug3, CLC-Pred.

1. INTRODUCTION

Lung cancer has become a major threat to human health worldwide accounting for approximately 1.59 million deaths per year. Among the two subtypes of lung cancer, occurrence of small cell lung cancer (SCLC) is 10-15% and non-small cell lung cancer (NSCLC) is 85-90% [1]. Mutations like in-frame deletions or amino acid substitutions, clustered around the ATP-binding pockets of the tyrosine kinase domain of the epidermal growth factor receptor (EGFR) are the prime cause of NSCLC [2]. Studies show that mutations in the EGFR gene occur more often in women (37.5%) than in men (13.0%), in never- smokers (50.8%) than in former/current smokers (9.0%), and in patients of East Asian origin (29.1%) than in patients from the United States (9.5%) [3]. Two most common region of mutation in the EGFR gene during lung cancer are exon 19 and an L858R point mutation in exon 21. Both of these mutations affect protein structure near the adenosine triphosphate cleft of the tyrosine kinase domain of EGFR [4]. Tyrosine kinase inhibitors (TKIs) such as gefitinib and erlotinib are widely used in clinical treatment of NSCLC [5]. The epidermal growth factor receptor family of tyrosine kinases consists of four forms: EGFR (ErbB1, HER1), ErbB2 (HER2), ErbB3 (HER3) and ErbB4 (HER4). The binding of specific ligands leads to homo- and hetero-dimerization, with subsequent auto phosphorylation of the intracellular receptor TK domain. HKI-272 is a second-generation TKI that uses both the strategies of covalent binding and multi- targeting; it is an irreversible inhibitor of the EGFR and HER-2 receptors [6]. Nevertheless, the effectiveness of these inhibitors is restricted due to emergence of drug resistance. Therefore targeting EGFR is a promising approach to treat NSCLC.

Together with the swift development of bioinformatics, computational methods have become more efficient and popular for designing of new drugs [7]. Present paper aims to identify the natural inhibitors of mutant EGFR from 5 plant sources namely *Acorus calamus* [8], *Asparagus racemosus* [9], *Moringa oleifera* [10],

Withania somnifera [11] and *Rhododendron Arboreum* [12] using computational techniques. Hit compounds obtained from the study could play an important role in designing personalized therapy and development of innovative drugs against NSCLC.

2. MATERIALS AND METHODS

2.1 Library preparation

A library of 70 phytochemicals from selected plants mentioned above was built in-house by extracting information from scientific literatures and Herbal Net Digital Repository database [13]. The selected plants are well reported for their anti-cancer properties and traditionally known for their various ethanopharmacological activities against several diseases. Plant compounds with known structures were retrieved and used to create the library in mol2 format. Several structures of the phytochemicals were retrieved from PubChem [14] and Chemspider database [15]. The Structures of plant compounds which were not available were drawn by using ChemSketch [16] software in .mol2 format. All the molecules in the library were annotated by molecular weight, number of rotatable bonds, logP, number of H-bond donors, number of H-bond acceptors which are important parameters for designing potent drugs. Molecular format conversion was done by using Openbabel [17] and Avogadro software [18].

2.2 Molecular Docking

The subjected receptor was EGFR 696-1022 T790M mutant. The structure of the receptor was obtained in .pdb file from the Brookhaven Protein Data Bank (www.pdb.org). The above receptor was bound with native ligand WZ4002 covalently bind with the protein (PDB NO: 3IKA). PyMOl, a visualization tool was used as to investigate structural details of binding of ligands with receptor [19].

The grid parameters of native ligand were calculated by using Chimera 1.8 to use for docking of receptor and ligands. Virtual screening by docking were performed using AutoDock Vina program in PyRx graphic user interface (GUI) version 0.8 that run on a PC containing Windows7 [20]. The binding energy (kcal/mol) of receptor and ligand was calculated using Lamarckian genetic algorithm. AutoDock Vina used a measure of distance between the experimental and predicted structures – RMSD (root-mean-square deviation) to compare the accuracy of the predictions of the experimental structure. RMSD lower bound (rmsd/lb) and RMSD upper bound (rmsd/ub) were the expressions for RMSD metrics differing in how the atoms are matched in the distance calculation. The grid parameters used for docking analysis were:

Center coordinates: X= -12.704, Y=12.770, Z= -33.538

Dimensions (Å): X = 28.86; Y = 27.79; Z = 53.78

2.3 Anticancer efficacy

All molecules of the in house library were again screened for their anticancer efficacy using CDRUG online server [21]. Required smiles formats of all molecules ware merged in to a single .sdf molecule by using Open babel and then converted into smiles format. CDRUG server calculates molecular descriptors (relative frequency-weighted fingerprint) to implement query compounds. The similarity between the query and active compounds were measured in terms of hybrid score. At final step, a confidence level (p value) was calculated automatically which predicted the anticancer activity of compounds.

2.4 Drug likeness

Drug likeness of the 16 hit molecules was checked by using FAF-Drug3 [22] server which is originally based on the free chemoinformatics toolkit Frowns

2.5 Cell-Line Cytotoxicity Prediction

Finally hit molecules were analyzed for their cytotoxic effect on tumor and normal cell-lines using CLC- Pred web services [23]. Based on the structural formula this server predicts cytotoxicity of molecules for tumor and normal cell-lines. Prediction of the result was based on PASS (Prediction of Activity Spectra for Substances) technology. Here predictions were based on both mRNA- based training set and protein-based training set. The mRNA based training set allows to predict drug-induced changes of gene expression for 952 genes (500 up- and 475 down regulations).The protein based training set allows to predict drug-induced changes of gene expression for 129 genes (85 up- and 51 down regulations).

3. RESULTS AND DISCUSSION

Identification of anticancer compounds from phytochemicals is an important step in modern drug discovery field. In present study an in house library of 70 phytochemicals were screened initially by molecular docking. Docking is very useful computational tool to identify or prove mechanism of action of ligands that have shown some specific biological properties. All of them were docked in the active site of EGFR protein. WZ4002 was used as reference ligand for analyzing the docking results. Total 32 out of 70 natural compounds were screened by docking process. It can be concluded from the results [Table-1] that the compounds may have mechanism of action like the reference ligand [Figure-1].

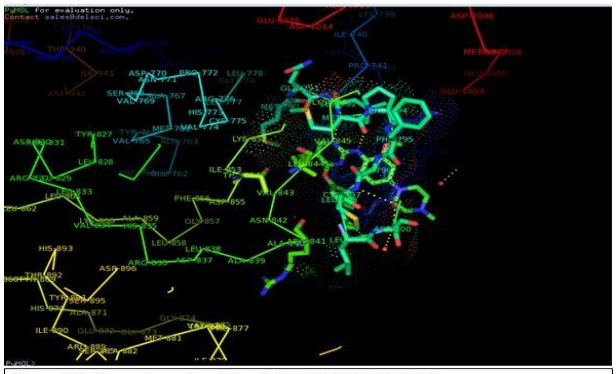

Figure -1 Binding conformation of ligand with EGFR 696-1022 T790M

Acorus calamus		*Withania somnifera*		*Rhododendron arboreum*		*Moringa oleifera*		*Asparagus racemosus*	
Name of Phytochemicals	Docking Score kcal/mol	Name of Phytochemicals	Docking Score kcal/mol	Name of Phytochemicals	Docking Score kcal/mol	Name of Phytochemicals	Docking Score kcal/mol	Name of Phytochemicals	Docking Score kcal/mol
Isoeugenol methyl ether	-4.7	1-Isobutyl-3-methylcyclopentane	-4.7	Lupeol	-4.8	4-(4'-*O*-acetyl-α-L-rhamnopyranosyloxy)benzyl isothiocyanate	-5.5	Aspargamine A	-6.0
Acoramone	-4.8	Diazene	-2.7	Betulin	-4.7	4-(-L-rhamnopyranosyloxy)benzyl isothiocyanate	-4.7	Racemosol	-5.2
Asarylaldehyde	-3.9	Withaferin A	-5.5	3-B-acetoxyurs-11-en-13-B	-4.4	Niazimicin	-5.9	Racemofuran	
Shyobunone	-4.3	Viscosalactone B	-5.2	Betulinic acid	-4.8	Pterygospermin	-5.2	Quercitin	-5.0
Epishyobunone	-4.1	Withanolide D	-6.0	3-o-acetylbetulinic acid	-4.5	Benzyl isothiocyanate	-5.9	Shatavarin V	-5.2
Isocalamendiol	-4.2	Lanosterin	-5.2	Isopropyl-beta-D-thiogalactopyranoside acetate	-3.6	4-(α-L-rhamnopyranosyloxy)benzyl glucosinolate	-6.1	Sarsasapogenin	-6.4
Acoragermacrone	-4.1	Anaferine	-4.7	Arbutin	-4.4	m-Diazine	-4.7		
Preisocalamendiol	-4.9	Amyl Nitrite	-3.4	Rutin	-6.2				
Thujane	-4.9	Ashwagandhanolide	-4.5	Alnulin	5.0				
Terpinolene	-5.4			Ursolic acid	-4.7				
Galangin	-5.3								
Beta-sitosterol	-4.9								

Table-1 Molecular Docking results of phytochemicals

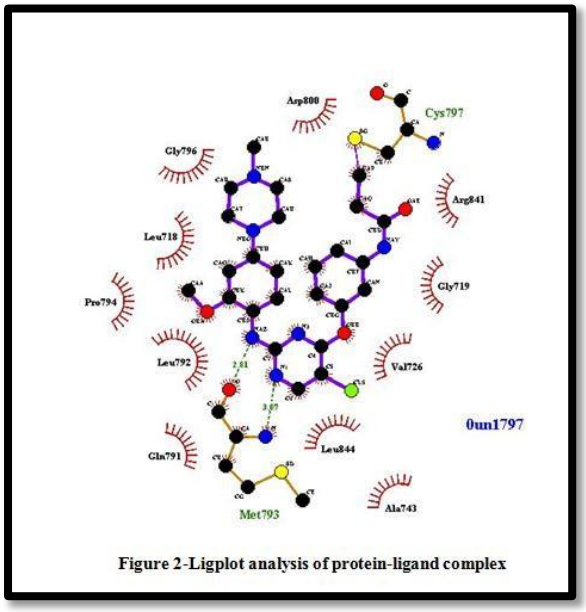

Figure 2-Ligplot analysis of protein-ligand complex

Table 2: FAF Drug and CLC-Pred Results, Pa=Probability to be active, Pi=Probability to be in active

S. No	Name Hit Compounds	FAF Drug Results	Cell-Line Cytotoxicity Prediction by CLC-Pred					
		Accepted Or Rejected	Pa	Pi	Cell- line	Cell-line full name	Tissue	Tumor type
1	**Beta-sitosterol**	**Accepted**	0.994	0.006	A-427	Lung carcinoma cells	**Lung**	Carcinoma
			0.965	0.010	NCI-H1975	Bronchoalveolar carcinoma cells	**Lung**	Carcinoma
			0.925	0.018	NCI-H1299	Non-small cell lung carcinoma	**Lung**	Carcinoma
			0.910	0.012	NSCLC	Non-small cell lung carcinoma cells	**Lung**	Carcinoma
			0.923	0.034	NCI-H128	Small cell lung cancer	**Lung**	Carcinoma
2	**Lanosterin**	**Accepted**	0.996	0.006	A-427	Lung carcinoma cells	**Lung**	Carcinoma
			0.975	0.023	NCI-H128	Small cell lung cancer	**Lung**	Carcinoma
			0.908	0.015	NCI-H1975	Bronchoalveolar carcinoma cells	**Lung**	Carcinoma
3	**Lupeol**	**Accepted**	0.996	0.006	A-427	Lung carcinoma cells	**Lung**	Carcinoma
			0.971	0.013	NCI-H1299	Non-small cell lung carcinoma	**Lung**	Carcinoma
			0.979	0.021	NCI-H128	Small cell lung cancer	**Lung**	Carcinoma
4	**Preisocalamendiol**	**Accepted**	0.923	0.018	NCI-H1299	Non-small cell lung carcinoma	**Lung**	Carcinoma
5	**Viscosalactone B**	**Accepted**	0.967	0.007	NSCLC	Non-small cell lung carcinoma cells	**Lung**	Carcinoma
			0.968	0.013	NCI-H1299	Non-small cell lung carcinoma	**Lung**	Carcinoma
			0.965	0.010	NCI-H1975	Bronchoalveolar carcinoma cells	**Lung**	Carcinoma
			0.949	0.005	NCI-H460	Non-small cell lung carcinoma	**Lung**	Carcinoma
			0.923	0.010	A-427	Lung carcinoma cells	**Lung**	Carcinoma
			0.929	0.021	HOP-92	Non-small cell lung carcinoma cells	**Lung**	Carcinoma
6	**Withaferin A**	Accepted	0.974	0.005	NCI-H460	Non-small cell lung carcinoma	**Lung**	Carcinoma
			0.961	0.007	NSCLC	Non-small cell lung carcinoma cells	**Lung**	Carcinoma
			0.949	0.014	HOP-92	Non-small cell lung carcinoma cells	**Lung**	Carcinoma
7	**Withanolide D**	Accepted	0.994	0.004	NCI-H460	Non-small cell lung carcinoma	**Lung**	Carcinoma
			0.955	0.013	HOP-92	Non-small cell lung carcinoma cells	**Lung**	Carcinoma

			0.943	0.010	NSCLC	Non-small cell lung carcinoma cells	**Lung**	Carcinoma
			0.943	0.017	EKVX	Non-small cell lung carcinoma cells	**Lung**	Carcinoma
8	**Quercitin**	Accepted	0.946	0.012	NCI-H1975	Bronchoalveolar carcinoma cells	**Lung**	Carcinoma

High scored results obtained from docking were visualized by making Ligplot of the docked complex [Figure-2]. The Ligplot analysis was done to evaluate the in-depth interaction pattern between docked ligand and active site residues in the protein. Ligplot analysis was necessary to examine the hydrophobic interactions as well as hydrogen bonding pattern. Subsequent level of annotation was done by using CDRUG server. In the results of this server the molecules were categorized as highly possible, possible, and less possible depending on the p-value, and are colored by green, black, and gray, respectively showing their various drug likeness properties. 11 phytochemicals were screened which showed high possibility to act as anticancer drug. Afterwards these 11 ligands were tested for Lipinski's rule of five and approval for drug likeness by using FAF Drug 3 server. 8 phytochemicals showed accepted result in this context and were screened as hits. Finally these 8 hits were subjected to CLC-Pred server and evaluated for their cytotoxicity on cell lines. The results [Table 2] showed all of them have cytotoxic effects on several lung cancer cell lines with greater than 0.9 probability to be active. This result proves that hits might be efficient therapeutic candidates to treat lung cancer Further experimental studies on eight phytochemicals would delineate the possible anticancer properties of these hits to treat NSCLC. This *in silico* study of signifies the medicinal importance of phytochemicals at molecular level. Plants containing these hits may be a better source of production of these phytochemicals facilitating industrial process of extraction and manufacturing of these medicinally important phytochemicals.

4. CONCLUSION

The problem of systemic toxicity and drug resistance (MDR) in lung cancer shows the vital necessity of discovering new anti-lung cancer drug. From this work, we identified several novel compounds from plants sources which could be very effective to treat lung cancer in human being. Using different computational validation methods, four natural compounds are identified in present work that may have potential as novel EGFR 696-1022 T790M inhibitors. As natural compounds constitute a higher percentage of marketed drugs, indicating a significantly higher rate of return per molecule against NSCLC.

5. Acknowledgements

The authors are thankful to "Prof. Hema Joshi" Department of Botany, SSJ Campus, Kumaun University, Almora, for providing necessary facilities to carry out the present study.

6. REFERENCES

1. Petersen I (2011).The morphological and molecular diagnosis of lung cancer. Dtsch Arztebl Int. 108:525-531
2. Zhang Z. F (2012). Activation of the AXL kinase causes resistance to EGFR-targeted therapy in lung cancer. Nat Genet. 44:852–860
3. El-Azab A S, Al-Omar M A,Abdel-Aziz A et al .(2010).Design. synthesis and biological evaluation of novel quinazoline derivatives as potential antitumor agents: molecular docking study. Eur Journal MedChem. 2 45:4188–4198
4. Franklin WA,Veve R,Hirsch F R, et al.(2002).Epidermal growth factor receptor family in lung cancer and premalignancy. Semin Oncol. 29:3–14
5. Hou T. J. Zhang W. Wang J. & Wang W (2009). Predicting drug resistance of the HIV-1 protease using molecular interaction energy components. Proteins.74 :837–846
6. Linggi B, Carpenter G.(2006). ErbB receptors: new insights on mechanisms and biology. Trends Cell Biology. 16: 649–656
7. Loo L H, Wu L F, Altschuler S J.(2007). Image-based multivariate profiling of drug responses from single cells. Nat Methods. 4: 445–453
8. Arasan E R,Vijayalakshmi M, Garikapati D. (2009). Acoruscalamus linn. Chemistry and Biology. Research J. Pharm. 2: 39-41
9. Gangwar K K, Deepali, Gangwar R S.(2010). Ethno medicinal Plant Diversity in Kumaun Himalaya of Uttarakhand. India. Nature and Science. 8: 66-78.
10. Jed W, Fahey D. (2005).Moringa oleifera: A Review of the Medical Evidence for Its Nutritional. Therapeutic. and Prophylactic Properties. Trees for Life Journal. 5:9-13
11. Kumar S M,Kumar D V,Kumar S,Aslam S (2011)The Phytochemical Constituents of Withania somnifera and Withania obtusifolia by GCMS Analysis. International Journal of Pharmacognosy and Phytochemical Research. 3:31-34
12. Srivastava P. (2012) Rhododendron arboretuman overview. Journal of Applied Pharmaceutical Science. 2:158-162
13. Herbal Net Digital Repository. http://herbalnet.healthrepository.org/handle/123456789/2418. Accessed 8th December. 2014
14. The Pub Chem Project. https://pubchem.ncbi.nlm.nih.gov. Accessed 15th December.2015
15. Chem Spider. http://www.chemspider. Com. Accessed 16th December. 2015
16. ACD/Structure Elucidator. version 15.01. Advanced Chemistry Development. Inc.. Toronto. ON. Canada. www.acdlabs.com. 2015
17. O'Boyle. N. M, Banck M, James C A et al.(2011). Open Babel: An open chemical tool box. Journal of Chemo informatics. 3: 33
18. Hanwell M D, Curtis D E,David C. Lonie T et al. (2012).Avogadro: an advanced semantic chemical editor. visualization. and analysis platform. J Chemo informatics. 4:17-21
19. De Lano W L. (2002).The pymol molecular graphics system. DeLano Scientific San Carlos. CA. USA
20. Trott O. Olson A J. (2010).AutoDock Vina Improving the speed and accuracy of docking with a new scoring function. efficient optimization. and multithreading. J Computed Chem. 31:455–461
21. Li G H. Jing F H. (2012).CDRUG a web server for predicting anticancer activity of chemical Compounds. Advance Access Oxford University Press. Bioinformatics. 6: 1-3

22. Miteva M A,Violas S et al.(2006). FAF-Drugs free ADME/tox filtering of compound collections. J Nucleic Acids Res. 7: 738-744
23. CLC-Pred. Cell-Line Cytotoxicity Predictor. http://www.pharmaexpert.ru/PASS online. the training set created on the basis of data on cytotoxicity retrieved from ChEMBLdb. version 17. https://www.ebi.ac.uk/chembldb.

*****