INTERNATIONAL RESEARCH JOURNAL OF PHARMACY

Research Article

IN SILICO IDENTIFICATION OF NOVEL EGFR TYROSINE KINASE INHIBITORS ASSOCIATED WITH NON-SMALL CELL LUNG CANCER FROM PHYTOCHEMICAL LIBRARY

Mahesha Nand [1], Priyanka Maiti [2], Subhash Chandra [3]*, Vena Pande [4]

[1]Research scholar, Department of Biotechnology, Kumaun University, Bhimtal Campus Bhimtal, Uttarakhand, India
[2]Research scholar, Department of Botany, Kumaun University, S.S.J Campus, Almora, Uttarakhand, India
[3]Assistant Professor, Department of Botany, Kumaun University, S.S.J Campus, Almora, Uttarakhand, India
[4]Associate Professor, Department of Biotechnology, Kumaun University, Bhimtal Campus Bhimtal, Uttarakhand, India
*Corresponding Author Email: subhashku@kunainital.ac.in

ABSTRACT

Inhibition of EGFR by targeting its tyrosine kinase domain is a worldwide accepted treatment of non-small cell lung cancer (NSCLC). Plant derived compounds show excellent potential as cancer growth inhibitors. Present study aims to identify potent plant derived inhibitors against EGFR tyrosine kinase ATP binding domain. 50 phytochemicals from different plant sources were screened by machine learning model and molecular docking. After that admetSAR server was used to check toxicity profile of screened phytochemicals followed by pharmacophore analysis. Finally six phytochemicals were obtained as potent natural inhibitors of EGFR tyrosine kinase which could be used as drug candidate against NSCLC in future.

Key words: Non-small cell lung cancer, Phytochemical inhibitor, Machine learning, Molecular docking, Pharmacophore.

INTRODUCTION

Lung cancer is a group of diseases which is characterized by presence of abnormal growth of cells. Current researches show that deaths caused by lung cancer have surpassed breast cancer death cases as the leading cause of cancer deaths in women for the past 25 years. Lung cancer is also the main cause of cancer death among the men. Lung cancer is mainly grouped into two subtypes, small cell lung cancer (SCLC) and non-small cell lung cancer NSCLC accounting 10-15% and 85-90% of all cases respectively. In case of metastasized NSCLC, the five-year overall survival of patients is less than 5%.[1] In NSCLC, a major problem has been found the overexpression and/or mutation in the protein sequence of epidermal growth factor receptor family of tyrosine kinases like HER1 (epidermal growth factor receptor; EGFR), HER2, HER3 and HER4. It is very noticeable fact that in case of 90 % solid tumors at least one member of the HER family of receptors has been found to be overexpressed [2]. The Gain -of- function mutation in EGFR has been recorded 10-15% in Caucasian and 20-30% in Asian NSCLC patients [3]. The first-generation EGFR tyrosine kinase inhibitors (TKIs) which are in use for treatment of lung cancer are Gefitinib (Iressa®; AstraZeneca Pharmaceuticals, Wilmington, DE) and erlotinib (Tarceva®; Genentech, Inc., South San Francisco, CA). These drugs have significant effect on NSCLC survival when given as salvage treatment after chemotherapy. However several clinical reports indicate that the patients have developed the acquired resistance against first-generation EGFR TKIs which necessitates discovery of second generation EGFR TKIs. Currently many second generation of EGFR TKIs are being introduced into clinical trials from the drug developmental pipelines. Some drugs like Tarceva and Gilotrif work very well as antagonist to EGFR4 [4-7]. Hence, targeting EGFR may be promising approach to discover new drugs against NSCLC. To discover noble drugs, phytochemicals are important source to provide basis of molecular framework since long time. It has been reported that many phytochemicals when are used in combination with chemo therapeutic drugs, they reduce unwanted side effects and synergize the therapeutic effects. Some anticancer drugs derived from plants namely *taxanes* (including paclitaxel isolated from *Taxus brevifolia* Nutt., *Taxaceae*) and *vinca* alkaloids (*Catharanthus* alkaloids) (including vinblastine and vincristine, isolated from *Catharanthus roseus* (L.) G. Don, *Apocynaceae*) are currently in clinical use. The proposed work focuses on to identify the phytochemicals which could be used as EGFR inhibitors. In this context an in house library of 50 phytochemicals was built. In present study, we have applied machine leaning, molecular docking (Autodock vina), pharmacophore analysis (PharmaGist) and toxicity prediction (admetSAR) to find out lead phytochemicals against lung cancer which further could be optimized and synthesized to develop potential drugs to manage and treat lung cancer efficiently in future.

MATERIALS AND METHODS

Protein structure retrieval

The three dimensional crystal structure of target co-crystal structure of 1.85 angstrom structure of EGFR kinase domain with Gefitinib **(**PDB id**:** 4WKQ**)** protein was retrieved from Protein Data Bank[8]. Miscellaneous ligands and other hetero-atoms such as water, etc. were removed from the protein structure to prepare for docking using Chimera 1.8 software[9].

Ligand structure preparation

The three dimensional structure of ligand molecules were retrieved from ChemSpider and Zinc database. After that .mol format of the ligand structures were converted into different molecular formats like .mol2, .sdf, etc. by using Avogadro software [10].

Calculation of Molecular descriptors and development of machine learning Model

To develop machine learning model a set of active and inactive compounds from PubChem bioassay (AID: 256664) was retrieved. Molecular descriptors were calculated using PaDEL-Descriptor v2.20software [11]. Total 34 molecular descriptors were calculated which included Ghose-Crippenoctanol-water partition coefficient (ALogP), H bond donor (nHBDon_Lipinski), H bond acceptor (nHBAcc_Lipinski), molecular weight (MW) etc. A machine learning model of training set and test set was prepared in Weka software[12] using Random Forest classifier.

Molecular Docking

The computational docking[13]of all ligands were performed into the active site of EGFR kinase domain using AutoDockVina program in PyRx graphic user interface (GUI) version 0.8. Before docking, polar hydrogen's were added to protein and thereafter, Kollman united atom partial charges were assigned. Docking was performed using Lamarckian Genetic Algorithm[14] with standard docking protocols. The coordinates of the docked protein along with the ligand was visualized using UCSF chimera and interactions of protein-ligand complex were studied using LigPlot+[15].

ADMET

The pharmacological properties such as absorption, distribution, metabolism, excretion and toxicity of hit compounds were predicted using admetSAR[16] server. In this server the latest and most comprehensive manually curated data for diverse chemicals associated with known ADMET profiles were compared with query compounds.

Pharmacophore elucidation

Common pharmacophores for the ligands were analyzed by using PharmaGist web server. All nine active molecules obtained from bioassay (AID: 256664) were used as reference for analyzing the pharmacophore in hit molecules. After that common pharmacophore for the hit ligands that are screened from above three processes were also analyzed[17].
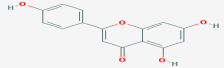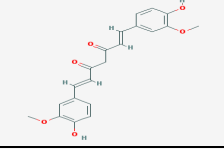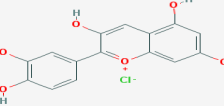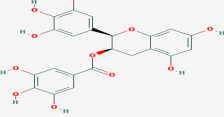
RESULTS

Several Millions are being affected and dying each year because of the NSCLC. Many drugs exist commercially that are being used for treatment of this disease. But still the investigation for new le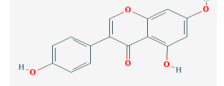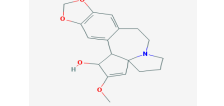ad continues because of the side-effects posed by the existing drugs; emergence of the drug resistant and the need to develop drug with better efficacy and potency. In this study, 50 phytochemicals were targeted against EGFR protein to analyze their inhibitory effect in NSCLC and the results were compared with the inhibitory efficiency of standard inhibitor (Gefitinib). A machine learning model was prepared to screen active molecules. 28 molecules out of 50 were found as active. Further docking analysis revealed [Figure 1] that 18 compounds have very high affinity towards the target protein and was similar to show binding affinity comparing to standard molecule [Figure 1]. ADMET & Toxicity predictions were finally performed to evaluate the pharmacokinetics and the toxicity of the15 phytochemicals. 6 hit ligands were screened finally by analyzing ADMET results as they do not have any mutagenicity and carcinogenicity properties [Table 1]. They were also found to follow Lipinski's rule of five which allow the evaluation of pharmacokinetics of the drug including absorption, distribution, metabolism and excretion (ADMET). All of them could cross blood brain barrier and be absorbed by human intestine which allows them to achieve high concentration at the target site.

Afterward pharmaGist server was used to elucidate pharmacophores for the 9 compounds from PubChem bioassay and 6 hit compounds**.** Pharmacophores were detected by extracting common chemical features from 3D structures of the active ligand set that are representative of essential interactions between the ligands and EGRF**.** The results predict physico-chemical property namely, hydrogen-bond acceptor/donor atom; a set of atoms of an aromatic ring and adjacent hydrophobic atoms**.** The active compounds showed four hydrogen bond acceptor, one hydrogen bond donor, one hydrophobic group, two aromatic rings whereas the common pharmacophore from 6 hit molecules showed three hydrogen bond acceptor, one hydrogen bond donor, one hydrophobic group, zero aromatic rings**.**

Table 1: Molecular docking and admetsar results

S.N	Name of Phytochemicals	Docking Score (kcal/mol)	admetSAR Results			
			Human Intestinal Absorption	**Caco-2 Permeability**	**Blood Brain Barrier**	**Carcinogens**
i.	Reference	-4.1	HIA+	Caco2+	BBB+	Non-carcinogen
ii.	Apigenin	-8.2	HIA+	Caco2+	BBB+	Non-carcinogen
iii.	Curcumin	-5.0	HIA+	Caco2+	BBB+	Non-carcinogen
iv.	Cyanidin	-5.6	HIA+	Caco2+	BBB+	Non-carcinogen
v.	Epigallocatechin gallate	-5.0	HIA+	Caco2+	BBB+	Non-carcinogen
vi.	Genistein	-4.9	HIA+	Caco2+	BBB+	Non-carcinogen
vii.	Gingerol	-3.9	HIA+	Caco2-	BBB+	Non-carcinogen
viii.	Glucosinolate	-3.2	HIA+	Caco2-	BBB-	Non-carcinogen
ix.	Daidzein	-6.6	HIA-	Caco2-	BBB+	Non-carcinogen
x.	Isoflavone	-6.7	HIA+	Caco2-	BBB+	Non-carcinogen
xi.	Phenoxodiol	-5.9	HIA+	Caco2-	BBB+	Non-carcinogen
xii.	Benzopyran	-4.9	HIA+	Caco2-	BBB-	Non-carcinogen
xiii.	Cephalotaxine (A)	-5.9	HIA+	Caco2+	BBB+	Non-carcinogen
xiv.	Resveratrol	-5.8	HIA-	Caco2-	BBB+	Non-carcinogen
xv.	Rohitukine	-3.1	HIA+	Caco2-	BBB+	Non-carcinogen
xvi.	Flavopiridol	-4.1	HIA-	Caco2-	BBB+	Non-carcinogen

Table 2: Sources and structures of six hit phytochemicals

Sl. NO	Name of Compounds	Sources		Structure
1	Apigenin	*Moringa oleifera*		
2	Curcumin	*Curcuma longa L*		
3	Cyanidin	Grapes, blackberry, cranberry, raspberry, apples, plums, red cabbage. And red onion.		
4	Epigallocatechin gallate	Green tea		
5	Genistein	lupine, fava beans, Soybeans, kudzu, psoralea, *Flemingia vestita*, and coffee.		
6	Cephalotaxine	Cephalotaxus		

Figure 1-Binding conformation of ligand with EGFR tyrosine kinase ATP binding domain

DISCUSSION

The complete study was conducted *in silico* to explore the therapeutic potentials of the naturally derived phytochemicals against EGFR. The overall analysis suggests that 6 phytochemicals [Table 2] show a higher affinity towards the EGFR protein. The study provides a clue for the development of new lead molecules with therapeutic importance that could be used for the prevention or cure of NSCLC.

ACKNOWLEDGMENT

The authors are thankful to the Head, Department of Botany, SSJ Campus, Kumaun University, Almora for providing necessary facility to carry out this work.

REFERENCES

1. Julian RM, Ping Y, Stephen D, Cassivi, Steven E. Schildand AA, Adjei. Non–Small Cell Lung Cancer Epidemiology Risk Factors Treatment and Survivorship. Mayo Clinic Proceedings 2008; 83:584–594.
2. Quaresma M, Rachet B, Coleman MP. 40-year trends in an index of survival for all cancers combined and survival adjusted for age and sex for each cancer in England and Wales. Lancet 2014; 6736:61396-61399.
3. Dorsey BD, Levin RB, Mcdaniel SL, Vacca JP, Guare JP, Darke PL, Zugay JA, Emini EA, Schleif WA, Quintero JC. The design of a potent and orally bioavailable HIV protease inhibitor. Journal of Medicinal Chemistry 1994; 37:3443–3451.
4. Fukuoka M, Yano S, Giaccone G. Multi-institutional randomized phase II trial of gefitinib for previously treated patients with advanced non-small- cell lung cancer (The IDEAL 1 Trial). Journal of Clinical Oncology 2003; 21:2237–2246.
5. Kris MG, Natale RB, Herbst RS. Efficacy of gefitinib, an inhibitor of the epidermal growth factor receptor tyrosine kinase, in symptomatic patients with non-small cell lung cancer. The Journal of the American Medical Association 2003; 290: 2149 –2158.
6. Perez SR, Chachoua A, Hammond LA. Determinants of tumor response and survival with erlotinib in patients with non–small-cell lung cancer. Journal of Clinical Oncology 2004; 22:3238–3247.
7. Shepherd FA, Rodrigues PJ, Ciuleanu T. Erlotinib in previously treated non-small-cell lung cancer. Journal of Medicinal Chemistry 2005; 353:123–132.
8. Rcsb.org. Protein Data Bank [updated July 2015 13; cited July 2015 15]. Available from: http://www.rcsb. org/pdb/home/home.do/.
9. Pettersen EF. UCSF Chimera-A visualization system for exploratory research and analysis, Journal of Computational Chemistry 2004; 25:1605.
10. Hanwell MD, Curtis DE, David C, Lonie T, Vandermeersch, Zurek E, Geoffrey R. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. Journal of Cheminformatics 2012; 4:17.
11. Liew HY, Sharma CYN, Woo SK, Chau YT, Yap CW. PaDEL DD Predictor Open-source software for PD-PK-T prediction. Journal Computer 2012; 4:1-15.
12. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P. The WEKA data mining software. ACM SIGKDD Explore News 2009; 11:10-11.
13. Dewi Yuliana, Mursalin, Ahmad Najib. In silico screening of chemical compounds from Sweet flag (Aracus calamus L) as α-Glucosidase inhibitor. Int. Res. J. Pharm. 2013; 4(3):110-112 DOI: 10.7897/2230-8407.04320
14. Trott O, Olson AJ. AutoDockVinaImproving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry 2010; 31:455–461.
15. Wallace AC, Laskowski RA. LIGPLOT a program to generate schematic diagrams of protein-ligand interactions. Journal of Protein Engineering 1995, 8:127-134.
16. Cheng F, Li W, Zhou Y, Shen J, Wu Z, Liu G, Lee P, Tang W. admetSAR A Comprehensive Source and Free Tool for Assessment of Chemical ADMET Properties. Journal of Chemical Information and Modeling 2012; 52:3099-3105.
17. Schneidman DD, Dror O, Inbar Y, Nussinov R, Wolfson HJ. PharmaGist: a webserver for ligand-based pharmacophore detection. Nucleic Acids Research 2008; 36:223-228.

Cite this article as:

Mahesha Nand, Priyanka Maiti, Subhash Chandra, Vena Pande. *In silico* identification of novel EGFR tyrosine kinase inhibitors associated with non-small cell lung cancer from phytochemical library. Int. Res. J. Pharm. 2016;7(3):22-25 http://dx.doi.org /10.7897/2230-8407.07323