Exploring Molecular network and docking analysis of Osimertinib in cancer Cascade: Need for Drug Repurposing

Shweta More, Y.B. Chavan
College of Pharmacy, India

Abstract

Background: The network-target-based network pharmacology is a promising approach for the next generation of drug research and development for traditional as well as synthetic medicine. We report on the major in-depth molecular analysis of Osimertinib and provide new insight into molecular events involved in the progression of Non-small cell lung cancer.

Objective: Network pharmacology uses computational biology to develop our understanding of drug actions and to advance drug discovery. Here we apply network pharmacology to generate testable hypothesis about multi-target mechanism of Osimertinib against Non-small-cell lung cancer (NSCLC).

Methods: We reconstructed drugtarget pathways and network to predict the protein targets of Osimertinib and interactions between targets and the drug. Then we validated our prediction of five candidate targets (c-MET, EGFR, FGFR, VEGFR, ERR) by performing docking studies with Osimertinib.

Results: The results suggest that Osimertinib acts against NSCLC by regulating function of signaling proteins, including CREB, CDK2, EGFR, TNF, BRPF-1, ErbB, P13K/AKT, HIF-1, NFIB, CAMP, and HGF, which regulates the functions of various biological, molecular and cellular responses in NSCLC. Osimertinib is predicted to affect networks involved mainly in cancer viz., renal cell carcinoma, endometrial cancer, prostate cancer, and bladder cancer.

Conclusion: This approach of repurposing may be useful for multi-target drugs against complex diseases.

Keywords: Network pharmacology, Osimertinib, non-small cell lung cancer (NSCLC), and Molecular Docking.

Introduction

Epidermal Growth Factor Receptor (EGFR) is a membrane-bound glycoprotein which constitutes one of the four members of ErbB family of tyrosine kinase receptors (1). Uncommon activation of the EGFR is a dynamic force for tumor progress in a sub-group of non-small cell lung cancer (NSCLC) patients (2). Autophosphorylation of receptor tyrosine kinase occurs by activation of EGFR which leads to initiate a cascade of downstream signaling pathways involved in regulating cellular proliferation, differentiation, and survival. EGFR is activated by mechanisms which are associated with the development of variety of human cancers like over-expression, mutation, ligand-dependent receptor dimerization, and ligand-independent activation. Currently, the most established therapeutic target is the EGFR which also includes HER-2, HER-3, and HER-4. One of the vital targets for tumor chemotherapy is EGFR inhibition. Four EGFR-TK inhibitors (EGFR-TKIs) (Gefitinib, Erlotinib, Afatinib, and Icotinib, the latter to be used China only) have consistently demonstrated with superior efficacy as compared with platinum-based chemotherapy in phase 3 trials of EGFR mutation positive advanced NSCLC, thus emerging as standard first-line treatment in this context (3). Gefitinib and Erlotinib were the earliest EGFR-TKIs to be developed for clinical use, and are generally referred to as first-generation or reversible EGFR-TKIs. Afatinib is a second-generation EGFR-TKI, which blocks EGFR in an irreversible manner, also inhibiting other members of the EGFR family (HER-2 and HER-4). Osimertinib is a third-generation, orally available, irreversible, mutant-selective, EGFR inhibitor, with potential antineoplastic activity (4). Oncogenic mutations in EGFR usually increase the kinase activity of EGFR, thus leading to hyper activation of the pro-survival signaling pathways (5). Signaling...
through EGFR can be divided in two main categories: kinase-dependent and kinase-independent signaling. Kinase-dependent signaling results in the activation of several downstream pathways. A first activated pathway is Ras-Raf-MEK-ERK and JNK signaling, that promotes cell survival and proliferation. A second signaling pathway is the activation of PI3K/AKT through the association of Her-3. Thirdly, activation of EGFR also leads to STAT-3 dimerization and translocation to the nucleus. Lastly, PLC signaling is activated through Src. Kinase-independent functions includes the stimulation of DNA synthesis, activation of mitogen-activated protein kinase (MAPK) signaling through hetrodimerization, and anti-autophagic effects (6).

In the present study, we have focused on prediction and analysis of the potential novel targets associated with anticancer of the EGFR kinase domain. Here, SAR, molecular docking, virtual screening, and ligand-based pharmacophore screening and computational system biology tools were used as rational strategies for the elucidation of molecular basis similarity of structural analogs of Osimertinib. Additionally, increasing demands on to reduce animal testing, cost, and time and laboratory wastages are another reason to develop in silico models. Therefore, most of the regulators and stakeholders include pharmaceutical industries found to use these programs to evaluate structural analogs for specific therapeutic targets to potentiate them as for further in vitro or in vivo approaches (7); so that, our findings could provide crucial viewpoints and give a platform for researchers to design new methods for novel scaffolds envisaging towards EGFR models for moderating the NSCLC and other human malignant diseases. Thus, to reveal the mechanism and to study the structural aspect of Osimertinib, a brief synthetic approach has been highlighted along with the SAR. Our findings may give a fruitful aspect to synthetic chemists and biologists to design and synthesize new and potent compounds as the initial virtual screening process.

2.DESIGNING AND STRUCTURAL ACTIVITY RELATIONSHIP OF OSIMERTINIB

Osimertinib Mesylate chemically belongs to mono-anilino-pyrimidine small molecule with log P value of 4.49. The basic scaffold of Osimertinib is N-(3-pyrimidin-2-ylamino) phenyl acrylamide and the pyrimidine forms two hydrogen bonds with the main-chain nitrogen and carbonyl of Met-793 in the hinge region. The indole group is present at the adjacent of the gatekeeper residue, the amine moiety placed in the solvent channel and the covalent bond formed to Cys-797 via the acrylamide group of Osimertinib. The mechanism of action against EGFR is determined by irreversible binding through a covalent bond with the C797 amino acid which is ligand-induced EGFR activation in NOX-2-dependent sulfenylation of a cysteine residue. Binding of EGFR to its ligand leads to hetrodimerisation with its family member’s viz., Her-2, Her-3, and Her-4. As a result of EGFR inhibition, different pathways, in particular RAS/RAF/MAPK and PI3K/AKT, involved in DNA synthesis and proliferation, are inhibited (8) (Figure 1).

3.MATERIALS AND METHODS

3.1. Data preparation

NSCLC genes were obtained from Gene card human gene database (https://www.genecards.org/) which provides information related to all annotated and predicted human genes. Among the GeneCards categories only protein-coding class was selected for the interactions. Information about the drug i.e., Osimertinib were obtained through Pubchem (https://pubchem.ncbi.nlm.nih.gov/). Protein-protein interaction (PPI) is important aspect to study the involvement of proteins in various biochemical processes as well as to understand the cellular organization, bioprocess, and functions. This can be done by using the virtual screening database called STRING (9).

3.2. Construction of PPI of selected genes

The genes of the selected components were uploaded to STRING (https://string-db.org/cgi/network.pl) to get the information about PPIs. The setting for generating the PPI network was in accordance with ‘Homo Sapiens’ and the confidence in the interaction between the target protein was set to the highest confidence data >0.9. The network nodes represent proteins whereas the edge represents associated protein-protein (10).

3.3. Prediction of Compound-target pathway

Once the protein-protein interaction was carried, the next step is to understand the molecular mechanism which is achieved by constructing the compound-target network using Cytoscape visualization software v.3.7.1. The compound-target network helps to understand and analyze the mechanism of the components with target as well as the pathways involved.

Enrichment Analysis of NSCLC Target Gene Ontology (GO) Enrichment

In order to find the ontology terms associated with molecular, cellular, biological, and KEGG pathway enrichment analysis was carried out. The analysis was performed by Cytoscape plugin called ClueGO and Cluepedia. Pathways and networks were ranked according to the amounts of the molecules participating in pathways and networks, respectively. Pathways and networks shared by targets related to NSCLC and the potential Osimertinib targets were identified (11).

3.4. Docking of Osimertinib

Docking study was carried out to find the affinity as well as orien-
tation of the selected active components by docking them against the selected receptors as c-MET, EGFR, FGFR-κ/κ, VEGF, and ERR using Glide v_7.6 program interfaced with Maestro v_11.3 of Schrödinger 2017 (Schrödinger, LLC, New York, NY, USA). The crystal structure for c-MET (PDB Id: 5EYD), EGFR (PDB Id: 1XKK), FGFR-κ/κ (PDB Id: 4QQC), VEGFR (PDB Id: 1Y6B), and ERR (PDB Id: 1ERR) were taken from RCSB Protein Data Bank (https://www.rcsb.org/). The structures of compounds were built using Maestro build panel and optimized to lower energy conformers using Ligprep v_3.3.

4. RESULTS

4.1. NSCLC-related gene pathways and networks

A total of 17363 human genes associated with NSCLC were identified in the Gene Bank database. Genes obtained from Swiss target prediction was found to be 101 for Osimertinib. After filtering out 56 genes were used for constructing the network of Osimertinib with respect to NSCLC. The results for PPI were as: number of nodes: 56; number of edges: 191; average node degree: 6.82; average clustering coefficient: 0.783 and PPI enrichment value: <1.0 e-16 (Figure 2). Among these, CREB, CDK2, EGFR and BRPF-1 is located in the network which indicates the role of protein in the pathogenesis of NSCLC. Basically, ERBB and EGFR activation is a forthcoming starting point to discover the mechanisms of non-small cell lung cancer. So, PPI network analysis and pathway analysis of novel genes were carried out for the recognition of critical genes related to the NSCLC.

4.2. Compound-target Network and network analysis

The signaling pathways and functions of the selected target genes were examined by importing data to Cytoscape in order to construct compound-target network. The compound and target disease interaction network was constructed which elucidate the mechanisms of action of drugs in the treatment of NSCLC (Figure 3).

Figure 2: Protein-protein interaction network of Osimertinib in non-small cell lung cancer (NSCLC) targets obtained from STRING v_11.0 database.

Figure 3: Compound-target-NSCLC network constructed by Cytoscape v_3.7.

This fact inferred that the Osimertinib might influence these targets synergistically; as a result it may have therapeutic effects on other disease or disorder in addition to NSCLC. The details of three topological parameters i.e. Betweenness Centrality (the shortest path between other nodes), Closeness Centrality (the shortest paths between all nodes), and Degree (is a count of the number of unique edges that are connected) are given in Table 1 which gives an important role of each target in the network structure. The connectivity of a drug in the networks represents the drug’s effects, either therapeutic, in the context of the pathway networks, or toxic, in the case of the ADR network. In the case of drug-pathway networks, node degree is a measure of connectivity; high-degree drugs in these networks affect more functional pathways. Some drugs were found to have no significant effects on pathways, appearing as drugs with a degree of 0. Markedly, this does not mean that these drugs have no pharmacological effects; rather, it is indicative of no observable activity at the gene expression level, making them comparatively less likely to exhibit other system-wide effects than other more connected NSCLC drugs. In the context of ADRs, there are two complementary connectivity measures for each drug: degree indicates the number of possible side effects of a given drug, while node strength is a measure of relative risk for any side effect. The use of network-based metrics allows for a simple and generalizable categorization of these drugs based on their possible biological effects.

Table 1: The degree information of compound-target network obtained with network analysis.
4.3. Gene Ontology (GO) Enrichment and Network Analysis

GO enrichment analysis was carried out to analyze the target proteins. Three criteria were applied to analyze the target genes for GO biological (Table 2), Go molecular (Table 3), and GO cellular (Table 4) and the most important parameter as KEGG pathway (Figure 4).

The GO term fusion was restricted to pVs 0.005 that is based on the false discovery rate (Benjamini-hochberg) (12). The Ras-Raf-MAPK and P13K/AKT pathway is a major signaling route for the ErbB family, which leads to increased cell proliferation and inhibition of apoptosis.

Table 2: GO Biological process

Table 3: GO Molecular Function

Table 4: GO Cellular Component
INTRODUCTION

Genetically Engineered Biomaterials

The techniques used to develop new materials and to modify the properties of existing materials, are subjected to recombinant DNA techniques in the 1970s [6-9]. The area of large leap forward with the invention and development of ventures [5]. Scientists within the San Francisco Bay Area took a lead on NSCLC but via KEGG analysis it was found responsible in many cancers as colorectal, breast, renal cell carcinoma, pancreatic, endometrial, prostate, melanoma, bladder, hepatocellular carcinoma, and gastric cancer. Owing to the facts and visualization Osimertinib may be used as novel drug for the treatment of various diseases and disorder.

As far as NSCLC is considered, over-expression of ErbB receptor may lead to Ras activation (13). GO and KEGG analysis determined the signaling pathway as MAPK, ErbB, Ras, cAMP, HIF-1, FoxO, sphingolipid, PI3K-Akt, AMPK, VEGF, JAK-STAT, TNF, insulin, GnRH, estrogen signaling pathway, prolactin signaling pathway, thyroid hormone signaling pathway, and relaxin signaling pathway. Osimertinib can be used for the treatment of other conditions also as hepatitis B, influenza, Epstein-Barr virus infection, Kaposi sarcoma-associated herpes virus infection. As the constituents of Osimertinib was selected to examine the effects on NSCLC but via KEGG analysis it was found responsible in many cancers as colorectal, breast, renal cell carcinoma, pancreatic, endometrial, prostate, melanoma, bladder, hepatocellular carcinoma, and gastric cancer. Owing to the facts and visualization Osimertinib may be used as novel drug for the treatment of various diseases and disorder.

4.4. Binding mode

The mechanism of Osimertinib was reflected by interaction of compound and target. The crystal structure for c-MET (PDB Id: 5EYD), EGFR (PDB Id: 1XKK), FGFR (PDB Id: 1ERR) were taken from RCSB Protein Data Bank and prepared for docking using ‘protein preparation wizard’ (Table 5 and Figure 5). Water molecules were removed then the receptor was prepared for docking. In order to enhance the docking score the protein constraints were defined using protein preparation wizard. We found hydrogen bonding and π-π stacking was the main forms of interaction. For instance, the amine modification and pyrimidine groups of Osimertinib formed hydrogen bonds with the proteins, while with the aromatic ring and benzene ring of Osimertinib engaged in π-π stacking.

Table 5: Docking results of Osimertinib with selected tyrosine kinase

<table>
<thead>
<tr>
<th>St. No.</th>
<th>Tyrosine Kinase with PDB ID</th>
<th>Docking Score</th>
<th>Interacting Amino Acids</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>EGFR (PDB ID: 1XKK)</td>
<td>-8.53</td>
<td>NET 93, ASP 800</td>
</tr>
<tr>
<td>2</td>
<td>c-MET (PDB ID: 3YED)</td>
<td>-7.391</td>
<td>TYR 1250, GLU 1082</td>
</tr>
<tr>
<td>3</td>
<td>VEGFR (PDB ID: 1QOC)</td>
<td>-8.176</td>
<td>PHD 119, ARA 981</td>
</tr>
<tr>
<td>4</td>
<td>FGFR (PDB ID: 4QQC)</td>
<td>-9.122</td>
<td>GLY 680, ALA 159</td>
</tr>
<tr>
<td>5</td>
<td>ERR (PDB ID: 1ERR)</td>
<td>-9.432</td>
<td>LEU 753, VAL 838</td>
</tr>
</tbody>
</table>

Figure 5: Docking analysis of targets i.e., (A) EGFR (PDB Id: 1XKK) (B) FGFR (PDB Id: 4QQC) (C) ERR (PDB Id: 1ERR) (D) VEGFR (PDB Id: 1Y6B) (E) c-Met (PDB Id: 5EYD) against Osimertinib

5. DISCUSSION

Identification of drug target is taken into account to be a crucial principle in any new drug discovery. This will be achieved through...
system pharmacological approach by analyzing network interaction between biological entities (14). Network pharmacology approach helps to know the mechanism of medicine by considering its action and side effects in context with regulatory networks, within which drugs targets and disease gene products function well. Moreover, it focuses to know complex organism by studying genes and protein through network pharmacology that participate in gene-gene, protein-protein, and gene-protein interaction (15). This interaction of biological entities is often studied through network visualizing software like Cytoscape. Osimertinib were reported to inhibit cancer activity, but its efficiency and mechanism remain unclear till today. Mutations or over-expression of EGFR results in proliferation and abnormal fusion of ALK. EGFR plays a crucial role in carcinogenesis. ALK activation, which ultimately causes cell proliferation, invasion, and inhibition of apoptosis as an example, the epidermal protein receptor (EGFR) may be a transmembrane receptor that belongs to the family of receptor tyrosine kinases (RTK) (16). The mechanisms of E2/ERα interactions mainly involve ligand binding, receptor dimerization and activation of estrogen response elements (ERES) in target genes (17). EGFR binds to the ligands of EGF family and activates several signaling cascades to convert extracellular signal into appropriate cellular responses. Estrogen receptor beta class of nuclear hormone receptor which binds to estrogens with an affinity almost like that of ESR1 resulting in the activation expression of genes containing estrogen response elements (ERE) in an estrogen-dependent manner (18). Melanocortin 4 receptors (MC4R) activated by melanocortins like –melanocyte-stimulating hormone ( –MSH). MC4R are shown to switch the activity of multiple kinases like PKA, AMP-activated kinase, c-jun kinase, phosphatidylinositol-3-kinase and protein kinase C15 (19). Down-stream of those kinases, MC4R signaling regulates ion channel activity and organic phenomenon (20). Effects of MC4R on organic phenomenon have thus far been attributed to cAMP-mediated PKA activation resulting in subsequent phosphorylation of the transcription factor CREB and CRE-dependent transcription (21). Indeed, within the first tumor microenvironment, the stromal cells provide potent oncogenic signals, like TGFβ, HGF, epidermal growth factor (EGF), Wnt, and b-fibroblast growth factor (FGF), which stimulate cancer cell proliferation, survival, and invasion, thus facilitating metastasis (22). Therefore, in context, we have evaluated Osimertinib and studied their interaction with NSCLC protein through network pharmacology via molecular docking analysis.

In this study, we’ve used Cytoscape to explore and assess anticancer activity of Osimertinib with reference to NSCLC. Indeed, our drug-target-pathway network supported the target prediction and GO/KEGG pathway analysis showed that Osimertinib possesses multiple protein targets that involve in an array of interactive pathways. Basically the pathways involved within the current studies are pathways in cancer, calcium signaling pathway, cAMP, p-53, NF-KB, ErbB, EGFR, VEGFR, Jak-STAT, MAPK, HIF-1, FoxO, GnrR, hormone signaling pathway, RAS, RAP-1, estrogen, TNF, and cGMP-PKG signaling pathways (Figure 6).

The genes were selected consistent with the centrality within the constructed network. The obtained mentioned pathways aren’t only important in NSCLC but also found in other cancers as renal, CRC, gastric, prostate, bladder, and carcinoma. Thus it are often put into limelight that Osimertinib isn’t useful in cancer except for other diseases and disorders as hepatitis B, influenza, Epstein-Barr virus infection, and Kaposi’s sarcoma-associated herpes infection. Lastly, we’ve utilized both the docking software to predict and confirmed more reliable and versatile ligand/drug binding mode and affinity towards the cancer. The docking results may provide an important aspect for analyzing the structural relationship with respect to the residues of the drug targets. Hence, we’ve reasons to believe that Osimertinib is associated with a spread of pharmacological activities against many diseases. Thus, our methods help to improvise the present knowledge of drug proteins interaction, making it possible to relate pharmacological space with genomic space so as to better the line of treatment cancer disease.

CONCLUSION

In this current investigation, we accessed a computational approach to predict targets for cancer by exploring network pharmacology, integrating information from space analysis, docking of Osimertinib with NSCLC proteins. The binding mode and affinity of Osimertinib was also confirmed from Schrodinger docking program. Our present study suggested that the network target-based approach can be used for elucidating inter-relationship between complex disease and drugs intervention, which in turn can be used as booming and safe permutation of drugs for cancer management in the upcoming years.
LIST OF ABBREVIATIONS

ADME: Absorption, distribution, metabolism, excretion
AML: Acute myeloid leukemia
AMPK: AMP-activated protein kinase
ATP: Adenosyl triphosphate
Bcl-2: B-cell leukemia/lymphoma-2
CAMP: Cyclic adenosine 3', 5'-monophosphate
CDK: Cyclin-dependent kinase
cGMP: cyclic GMP; CML, Chronic myeloid leukemia
CRC: Colorectal cancer
DL: Drug-likeness
DNA: Deoxyribonucleic acid
EC: endothelial cell
EGFR: Epidermal growth factor receptor
ESR: Estrogen Signaling Receptor
ERR: Estrogen-Related Receptor
ERE: Estrogen Response Element
FAK: Focal adhesion kinase
FGFR: Fibroblast growth factor receptor
FoxO: Forkhead box O
GnrH: Gonadotropin-releasing hormone
GO: Gene ontology
HGF: Hepatocyte growth factor
HGFR: hepatocyte growth factor receptor
HIF-1: Hypoxia-inducible factor 1
IGF: Insulin-like growth factor
IL: Interleukin
IR: Insulin receptor
JAK: Janus Kinase
KDR: Kinase domain-containing receptor
KEGG: Kyoto Encyclopaedia of Genes and Genomes
MAPK: Mitogen-activated protein kinase
MET: Mesenchymal-epithelial transition
mTOR: Mammalian target of Rapamycin
NF-κB: Nuclear factor kappa B
Nrf2, nuclear factorE2-related factor 2
NRTKs: non receptor tyrosine kinases
NSCLC: Non-small cell lung cancer
PDB: Protein Data bank
PDGFR: Platelet-derived growth factor receptors
PI3K/AKT: phosphoinositide-3-kinase-protein kinase B/Akt
PIGF: Placental growth factor
PPI: Protein-protein interaction
PKA: protein kinase A
PKG: cGMP-dependent protein kinase
Rap-1: Ras-proximate-1 or Ras-related protein 1
RTKs: Receptor Tyrosine Kinase
STAT: signal transducer and activator of transcription proteins
TKIs: Tyrosine kinase inhibitors
TNF: Tumor necrosis factor.
VEGF: Vascular endothelial growth factor
VEGFR: Vascular endothelial growth factor receptor.

DECLARATIONS

ETHICS APPROVAL AND CONSENT TO PARTICIPATE
This research did not require an ethical approval as it does not involve any human or animal experiment.

HUMAN AND ANIMAL RIGHTS
No Animals/Humans were used for studies that are base of this research.

CONSENT FOR PUBLICATION
Not applicable.

AVAILABILITY OF DATA AND MATERIALS
Not applicable.

FUNDING
This research received no grant from any funding agency.

CONFLICT OF INTEREST
The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS
The authors thankfull to the Mrs. Fatima Rafiq Zakaria, Chairman Maulana Azad Educational Trust, Dr. Rafiq Zakaria Campus, Aurangabad 431001(MS), India for providing the laboratory facility. This work was supported by the Department of Science and Technology (DST), New Delhi, India [Project File No. SR/FT/LS-132/2012 & Sanction Diary No. SERB/F/4302/2012-13].

SUPPLEMENTARY MATERIAL
Supplementary Table 1: Input for construction of “component-disease-target network” through Cytoscape v_3.2.1.

ORCID
Santosh N Mokale https://orcid.org/0000-0002-9860-8895

REFERENCES

FIGURE LEGENDS

Figure 1: Designing and structural activity relationship of Osimertinib

Figure 2: Protein-protein interaction network of Osimertinib in non-small cell lung cancer (NSCLC) targets obtained from STRING_v_11.0 database

Figure 3: Compound-target-NSCLC network constructed by Cytoscape_v_3.7.1

Figure 4: Kyoto encyclopaedias of genes and genomes pathway and Gene ontology enrichment analysis

Figure 5: Docking analysis of components and targets i.e., (A) EGFR (PDB Id: 1XKK) (B) FGFR (PDB Id: 4QQC) (C) ERR (PDB Id: 1ERR) (D) VEGFR (PDB Id: 1Y6B) (E) c-Met (PDB Id: 5EYD) against Osimertinib

Figure 6: A comprehensive pathway map of epidermal growth factor receptor signaling pathway.

Table 1: Important node with network analyzer results

Table 2: GO Biological process

Table 3: GO Molecular function

Table 4: GO Cellular component

Table 5: Docking Score of Tyrosine Kinase (Interacting amino acid)