Research in Pharmaceutical Sciences, April 2023; 18(2): 121-137
Received: 12-04-2022
Peer Reviewed: 06-06-2022
Revised: 24-08-2022
Accepted: 25-12-2022
Published: 19-01-2023

School of Pharmacy & Pharmaceutical Sciences
Isfahan University of Medical Sciences

Original Article

Identifying potential ligand molecules EGFR mediated TNBC targeting the kinase domain-identification of customized drugs through *in silico* methods

Hima Vyshnavi[1] and Krishnan Namboori[1,*]

[1]Computational Chemistry Group (CCG), Amrita Molecular Modeling and Synthesis Research Lab, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, India

Abstract

Background and Purpose: Triple-negative breast cancer (TNBC) is an aggressive subtype of breast cancer in which three hormone receptors are negative. This work aimed at identifying customized potential molecules inhibiting epidermal growth factor receptor (EGFR) by exploring variants using the pharmacogenomics approaches.

Experimental approach: The pharmacogenomics approach has been followed to identify the genetic variants across the 1000 genomes continental population. Model proteins for the populations have been designed by including genetic variants in the reported positions. The 3D structures of the mutated proteins have been generated through homology modeling. The kinase domain present in the parent and the model protein molecules has been investigated. The docking study has been performed with the protein molecules against the kinase inhibitors evaluated by the molecular dynamic simulation studies. Molecular evolution has been performed to generate the potential derivatives of these kinase inhibitors suitable for the conserved region of the kinase domain. This study considered variants within the kinase domain as the sensitive region and remaining residues as the conserved region.

Findings/Results: The results reveal that few kinase inhibitors interact with the sensitive region. Among the derivatives of these kinase inhibitors molecules, the potential kinase inhibitor that interacts with the different population models has been identified

Conclusions and implications: This study encompasses the importance of genetic variants in drug action as well as in the design of customized drugs. This research gives way to designing customized potential molecules inhibiting EGFR by exploring variants using the pharmacogenomics approaches.

Keywords: Conserved region; EGFR; Kinase domain; Sensitive region; TNBC.

INTRODUCTION

Breast carcinoma or breast cancer has been reported as the second leading cause of death, next to lung cancer among women (1). Among breast cancer subtypes, triple-negative breast cancer (TNBC) is one of the lethal forms of cancer, where the three hormone receptors, estrogen, progesterone, and human epidermal growth factor receptor 2 (HER2) are absent (2). Though TNBC accounts for only 10-15 % of all breast cancer cases, most of the patients affected with the disease seem to be having complex metastases, leading to a minimum of 60% reduction in the 5-year survival rate of young women and making the disease into a matter of concern (3). The TNBC is due to the overexpression of the epidermal growth factor receptor (EGFR) gene, which is a kinase receptor, from the EGFR family (4). The missense substitutions; L858R (leucine of the 858th position of the protein to arginine) and T849I (threonine of the 849th position of the protein to isoleucine) of EGFR protein were identified in TNBC cell lines (5).

*Corresponding author: K. Namboori
Tel: +422-2685592, Fax: +422-2686274
Email: n_krishnan@cb.amrita.edu

Access this article online

Website: http://rps.mui.ac.ir

DOI: 10.4103/1735-5362.367792

The variation of a nucleotide base in a specific genomic position constitutes the single nucleotide variation (SNV) (6), which plays a crucial role in studying the individual-specific behavior of disease susceptibility and drug response.

Correlating SNVs to breast cancer subtypes has become a complex task (7). Population study shows a variation in the frequency of occurrence of SNVs within and across population groups. SNVs within coding regions of the genome may lead to a change in amino acids causing functional changes in the protein. Non-coding variants influence gene regulation if present in the regulatory regions (8). Hence both coding and non-coding variants are equally crucial for the analysis. Among various variant annotation platforms 'Ensemble Variant Predictor' provide an integrated web interface to carry out the analysis, prioritization, and annotation of variants within coding and non-coding regions (9-12). Homology modeling could generate a 3D structure of a variant protein (whose 3D structure is unavailable) from a known structure (13).

Kinase inhibitors (KI) are available in the market, but due to the tricky nature of kinases, they develop resistance to inhibitors (14). The protein-drug interaction (PDI) network could be used to study the interaction between protein and drug molecules. The network analysis focuses on studies with applications in cancer research (15). A molecular docking study could be employed to evaluate the PDI network (16). In 2017, Zeeshan Yousuf *et al*. identified potential molecules that could inhibit multi-targets of breast cancer through the *in-silico* approach (17). The effect of furanocoumarins in controlling breast cancer by targeting multiple targets has been studied through molecular docking studies (18).

To enhance the quality of scoring functions of docking, advanced computational techniques such as molecular mechanics-Poisson Boltzmann surface area or molecular mechanics-generalized born surface area (MM-GBSA)' can be incorporated. This helps in identifying the most accurate binding pose of the ligand within the binding site of the target and the corresponding binding energy (19). The hydrogen bond stability of the ligand-target complex can be further evaluated through molecular dynamic (MD) simulation (20). Evolutionary analysis of chemical structures would generate evolutionary derivatives of parent compounds which could be analyzed using physicochemical, absorption distribution metabolism, excretion, and toxicity (ADMET) properties (21).

The drug's effect on any population depends upon the variation of the target gene and the respective protein from the reference gene and protein kept in the repositories. Hence, the variants have to be incorporated into the gene and protein levels to identify the population-specific gene and protein targets. The population-specific variant percentage may also be considered while designing the model proteins (22).

In the present work, kinase domain-specific variants alone have been considered while making the target model protein molecules. The possibility of designing population-specific (customized) potential ligand molecules, inhibiting the EGFR-mediated TNBC and targeting the conserved region of the kinase domain, has been excavated in this work.

MATERIALS AND METHODS

The EGFR gene sequence has been collected from NCBI (23). Gene ontology has been studied using Uniprot and Gene cards (24,25). The genetic variations of EGFR have been collected from the database of single nucleotide polymorphisms (26). The variants were further analyzed using an ensemble variant effect predictor (VEP) (27). The functional domains of the protein have been identified using Expasy Prosite (28). The protein data bank (PDB) structures within the domain region have been identified from Uniprot (24). The identified variants have been included in the kinase domain of the target protein molecule.

Molecular modeling and evaluation of models

The 3D structures of the protein target molecule have been designed through homology modeling with the help of the online tool, SWISS-MODEL (29). Homology modeling includes the following steps: (i) identification of structural template(s): the

template structure has been identified based on the PDB structures present within the domain region; (ii) alignment of the target sequence and template structure(s): this has been done by doing sequence similarity between the template and the target sequence; (iii) model-building: the template was then subjected to model building where a 3D structure for target protein sequence have been generated based on the template protein structure; (iv) model quality evaluation. The quality estimation of modeled structure has been carried out using the Ramachandran plot and the ERRAT plot using the SAVES server (30).

PDI network

The list of approved kinase inhibitors has been collected from the literature and the National Institute of Health (NIH) database (31). The drugs included in the list have been compared with National Comprehensive Cancer Network (NCCN) guidelines (32). These drugs were further screened for the EGFR gene by biomolecular networking analysis using the PDI network with the help of Cytoscape 3.9.0 (). The network obtained is characterized with the help of degree, closeness centrality, and betweenness centrality using Cytoscape (33). The control drugs selected from biomolecular networking were further subjected to an interaction study with the target protein using Glide (34).

Molecular docking studies

The 3D structure of molecules that passed the screening has been retrieved from PubChem (35). Molecular docking studies were carried out using the Glide module of Schrodinger in the windows operating system (34).

Ligand preparation

The ligand minimization was carried out using the LigPrep module with the optimized potentials for liquid simulations (OPLS3e)' force field by adding the hydrogen atoms to the molecules and assigning bond orders.

Protein preparation

The protein preparation was carried out using the protein preparation wizard module. During the pre-processing step of protein preparation, the bond orders were assigned to the protein using the CCD database, adding hydrogen atoms, creating zero-order bonds to metal, creating disulfide bonds, and generating Het states using EPIK at pH 7.0. The pre-processed structure is refined through optimization by assigning hydrogen bonds using the option PROPKA at pH 7.0. The water molecules were removed where the Hets are beyond 3.0 A. The protein structure was then subjected to minimization using an OPLS4 forcefield.

Receptor grid generation

The binding site was defined based on the amino acid residues from the domain region. The grid was generated around the defined site.

Ligand docking

A flexible ligand docking was carried out using standard precision. Here, sample nitrogen inversions, sample ring conformations, and bias sampling of torsions for all predefined functional groups were followed, along with the addition of Epik state penalties to the docking score. Ten poses were generated for each ligand molecule. Post-docking minimization was performed for all the poses.

MD simulation

The MD simulation was carried out using the NAMD tool. A system was built using the ligand-target complex, where an orthorhombic simulation box was created. The complex was solvated within the simulation box using explicit water modeling with the CHARMM65 force field. The solvated model was then exposed to 0.15 M salt concentration. NVT (number, volume, and temperature) and NPT (number, pressure, and temperature) ensembles were used by setting the temperature as 310 K, pressure as 1 Pa, number of runs as 20,000, and simulation time as 100 ns. The results were analyzed using root mean square deviation (RMSD) and root mean square fluctuation (RMSF) plots (36).

MM-GBSA analysis

The ligand's binding affinity within the target's active site was further evaluated through the MM-GBSA method using the

Prime module of Schrodinger. The OPLS4 force field has been used with an implicit solvent model for intramolecular hydrogen bonding, hydrophobic, and pi-pi interactions. During the process, VSGB 2.0 solvation model was implemented (37,38).

RESULTS

The EGFR gene has been considered the target gene for TNBC. The gene was found to be involved in mechanisms including epidermal growth factor-activated receptor activity, nitric-oxide synthase regulator activity, nitric-oxide synthase activity, mutagen-activated protein kinase activity, protein serine/threonine kinase activity, ATPase, and protein tyrosine kinase activity.

Variant annotation

Among 43223 protein-coding SNVs identified, only 2027 were somatic. While analyzing the variants, it has been found that the SNVs rs55959834, rs41420046, rs2229066, rs1140475, rs41396448, rs2293347, and rs55737335 are synonymous; rs138240620, rs371228501, rs538497054, rs575565383, rs201830126, rs144496976, rs17290699, and rs542967903 are missense variants, and rs55959834 is within the splice region and is synonymous (Table S1).

The population analysis of variants gave the 1000 genome continental allele frequencies for the African (AFR), American (AMR), East Asian (EAS), European (EUR), and South Asian (SAS) populations with the help of the tool, VEP. Altogether, 16 variants are found to be in the kinase domain region. Ten variants among them are identified in the AFR population, six are found in the AMR, three are found in the EAS, six in the EUR, and five in the SAS. The frequency of occurrence of these variants across the population is included in Table 1.

Identification and design of targets

The kinase domain of EGFR protein is in the 712^{th} to 979^{th} positions of P00533. Among the 3D protein molecules included in the repository, 1XKK has been identified as the target EGFR protein keeping the kinase domain with a resolution of 2.4 Å. The population-specific protein models have been designed by including the variants within the domain region. The variants keeping part of the domain are considered the sensitive (highly mutable) region and the remaining part is considered a relatively conserved region (Fig. 1). Five mutant protein models have been generated corresponding to AMR, EUR, AFR, EAS, and SAS populations. The 3D target protein models have been designed by homology modeling. The quality of the models has been assessed by the RMSD, the qualitative model energy analysis (QMEAN), the ERRAT plot and the Ramachandran plot (Table 2).

Table 1. The 1000 genome continental allele frequency score for epidermal growth factor receptor.

Single nucleotide variation ID	African	American	East Asian	European	South Asian
rs55959834	0.0000	0.0000	0.005	0.0000	0.0010
rs138240620	0.0000	0.0014	0.0000	0.0000	0.0000
rs41420046	0.0136	0.0000	0.0000	0.0000	0.0000
rs371228501	0.0008	0.0000	0.0000	0.0000	0.0000
rs2229066	0.0008	0.0058	0.0000	0.0169	0.0051
rs530256683	0.0000	0.0029	0.0000	0.0000	0.0000
rs1140475	0.9281	0.8963	0.9196	0.8907	0.9397
rs538497054	0.0000	0.0000	0.0000	0.0000	0.0031
rs41396448	0.0023	0.0000	0.0000	0.0000	0.0000
rs575565383	0.0015	0.0000	0.0000	0.0000	0.0000
rs201830126	0.0000	0.0000	0.0000	0.0010	0.0000
rs144496976	0.0008	0.0000	0.0000	0.0020	0.0000
rs17290699	0.0340	0.0000	0.0000	0.0000	0.0000
rs542967903	0.0008	0.0000	0.0000	0.0000	0.0000
rs2293347	0.0129	0.1542	0.2470	0.1113	0.2301
rs55737335	0.0000	0.0058	0.0000	0.0030	0.0000

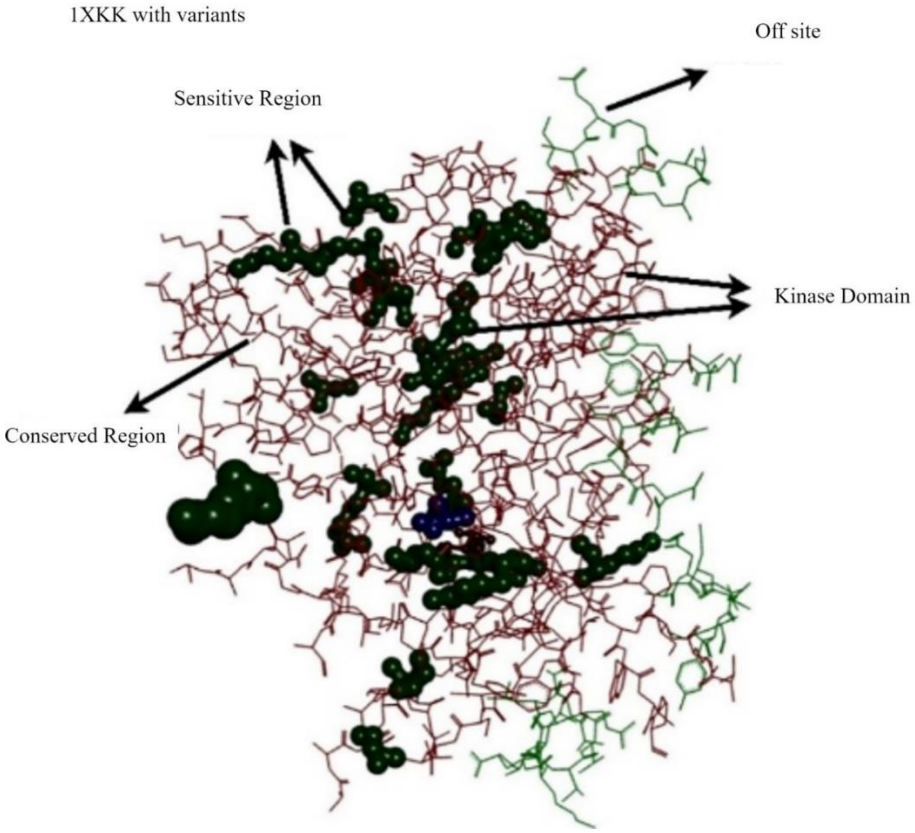


Fig. 1. 1XKK with variants incorporated in its domain region.

Table 2. Evaluation of model proteins.

Model quality check	American	African	East Asian	South Asian	South Asian
Qualitative model energy analysis	-3.43	-1.49	-0.84	-2.22	-2.10
Ramachandran plot					
Most favored regions	83.10%	90.00%	90.00%	88.20%	85.30%
Additional allowed regions	13.10%	9.50%	9.50%	10.70%	12.20%
Generously allowed regions	3.30%	0.50%	0.50%	1.10%	2.50%
Disallowed regions	0.50%	0.00%	0.00%	0.00%	0.00%
ERRAT					
Overall quality factor	91.2568	94.8718	94.8718	94.9721	92.8426
Root mean square deviation	1.4288	2.5680	1.3250	1.9740	1.2972

Identification of control drugs

The popular KI derived from 4-aminoquinazoline core pharmacophores such as afatinib, gefitinib, erlotinib, lapatinib, dacomitinib, sapatinib, sunitinib, icotinib, poziotinib, etc. has been considered for the analysis (Table S2). The biomolecular network with the derivatives connected to EGFR, covering the PDI is included in Fig. 2. All the kinase inhibitors interacting with EGFR have been selected and their closeness centrality and betweenness centrality were calculated (Table S3).

Molecular interaction study

Lapatinib, sapitinib, sunitinib, icotinib, and poziotinib are found to be interacting with the kinase domain with appreciably high affinity (Table 3). From docking results, lapatinib was found to be highly interacting with the template protein and AMR model; afatinib was found to be suitable for AFR and EAS models; neratinib showed a good binding affinity with the SAS model, and gefitinib showed a good binding affinity with EUR model. The results obtained by docking have been further evaluated by MMGBSA and MD simulation studies.

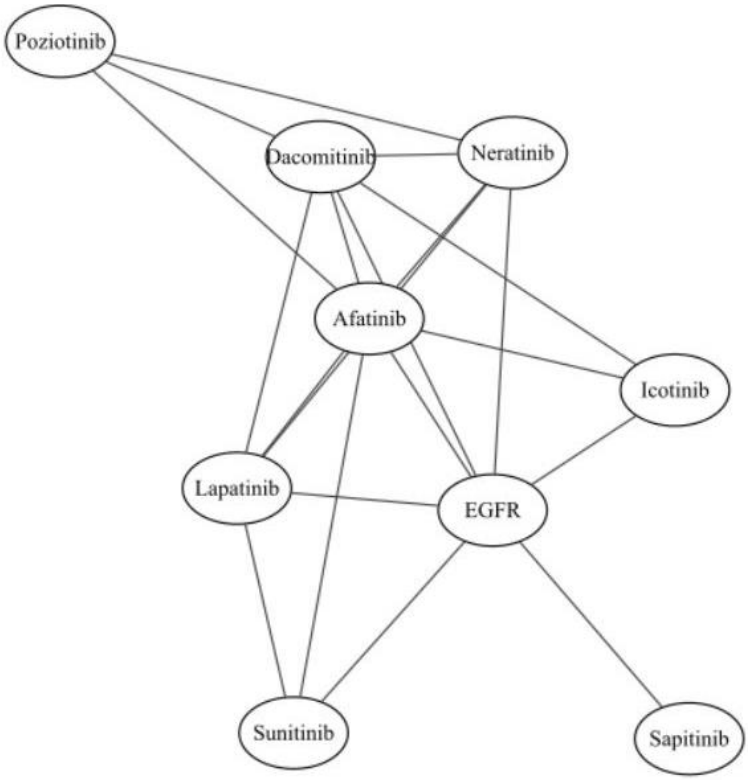


Fig. 2. The protein-drug interaction network of 4-aminoquinazoline derivatives and EGFR. EGFR, Epidermal growth factor receptor.

Table 3. The binding affinity (kcal/mol) of kinase inhibitors with the target.

Drug	1XKK	American	African	East Asian	South Asian	European
Afatinib	-6.995	-6.3127	-6.3481	-7.7268	-5.6299	-5.0180
Dacomitinib	-5.933	-5.1468	-5.5975	-5.2547	-5.5596	-5.9527
Erlotinib	-7.454	-7.8763	-5.3144	-6.2418	-5.5981	-4.9913
Gefitinib	-7.447	-7.1597	-6.1544	-5.9375	-5.1464	-5.9705
Lapatinib	-8.683	-7.9189	-5.0576	-5.6943	-4.9109	-5.9557
Neratinib	-7.992	-5.9925	-5.0136	-5.1994	-6.8993	-5.8290
Poziotinib	-8.389	-4.8763	-5.5216	-4.9908	-6.3311	-5.6902
Sapitinib	-6.351	-5.6874	-5.2546	-4.8792	-4.8268	-5.7038
Sunitinib	-7.871	-5.3445	-5.5387	-4.8088	-5.9046	-5.0794

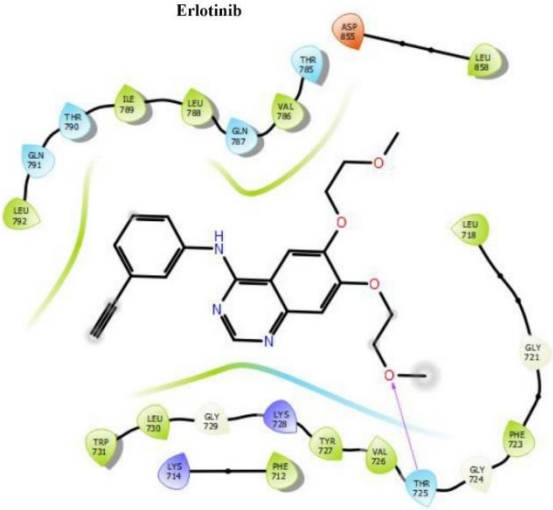
Evaluation of ligand-target complex

MD simulation

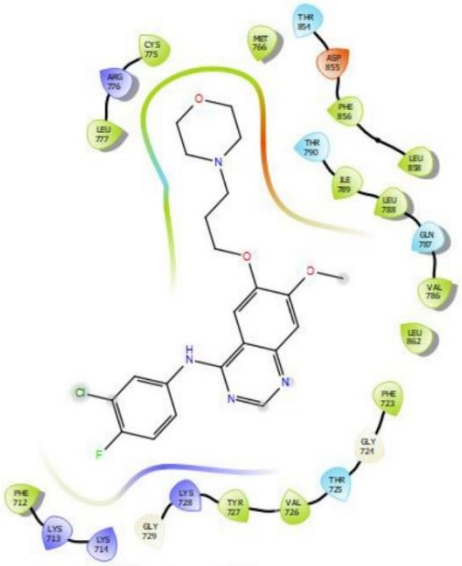
MD simulation showed that lapatinib and gefitinib showed good interaction with the kinase domain of 1XKK and showed good hydrogen bond stability during simulation. Lapatinib showed hydrogen bond stability with the AMR model, sunitinib showed H-bond stability with AFR and SAS models, neratinib for EAS, and sapitinib for the EUR model. For the ligand-target complexes (Fig. 3) when subjected to MD simulation, few of them lost their hydrogen bond, and few of them retained their bond.

The RMSF plots of the target models are included in Fig. S1. Fluctuations were found in the RMSF plots of model proteins. The residues within the position 726, 750, 798, 870, and 970 of 1XKK template protein showed the highest fluctuations; in the AMR model, residues within 728, 754, 806, 810, 915, and 970 positions showed high fluctuations; the residues within the region 806, 883, and, 962 of AFR model showed the highest fluctuations. The EAS model's residues within positions 885, 915, and 979 fluctuated the most. Similarly, for the EUR model, residues within positions 884, 910, 936, and 962 showed fluctuations. SAS model residues within the region 806, 860, and 910 showed the highest fluctuations among the residues.

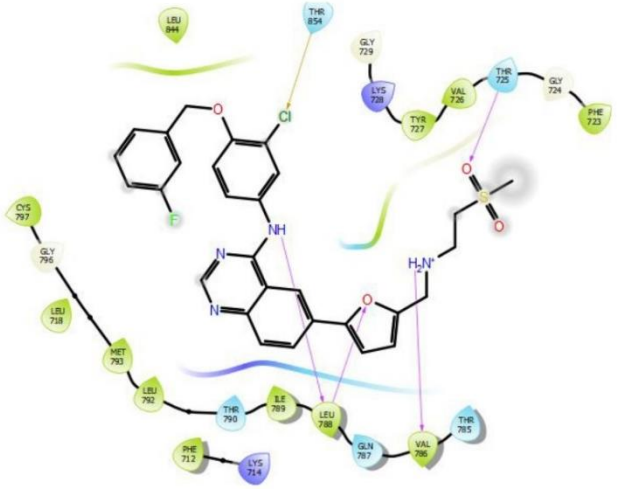
1XKK (Template protein) - Erlotinib



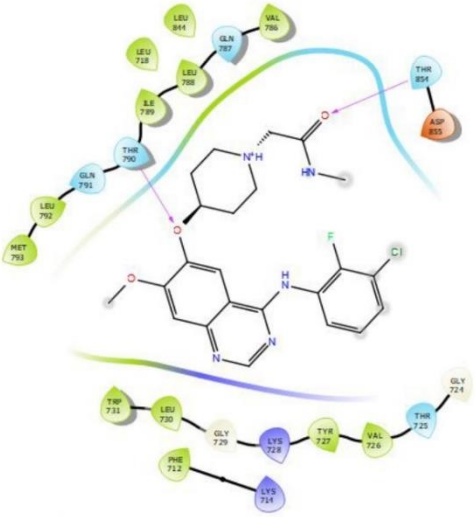
1XKK (AMR protein model) - Gefitinib



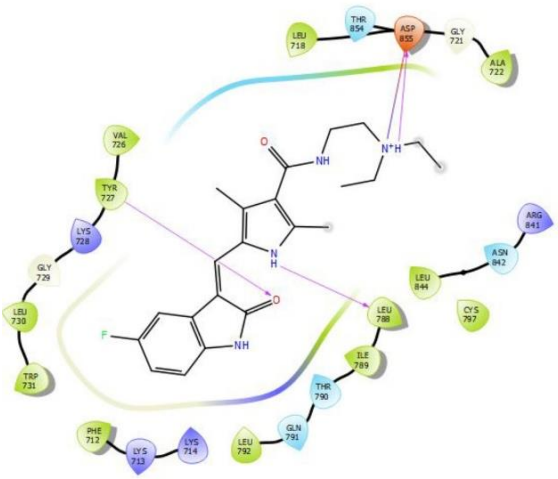
1XKK (AMR protein model) -Lapatinib



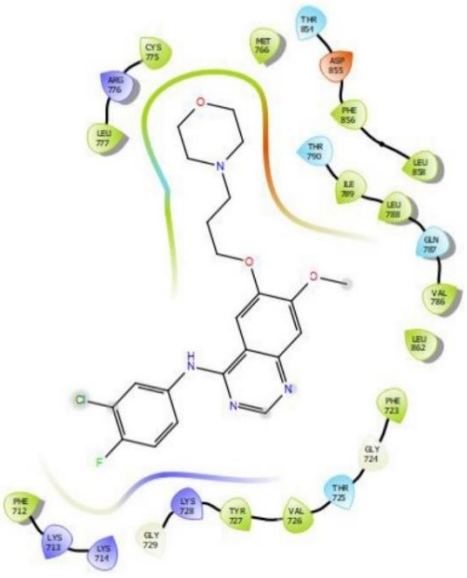
1XKK (EUR protein model) -Sapitinib



1XKK (AFR protein model) -Sunitinib



1XKK (EAS protein model) - Neratinib



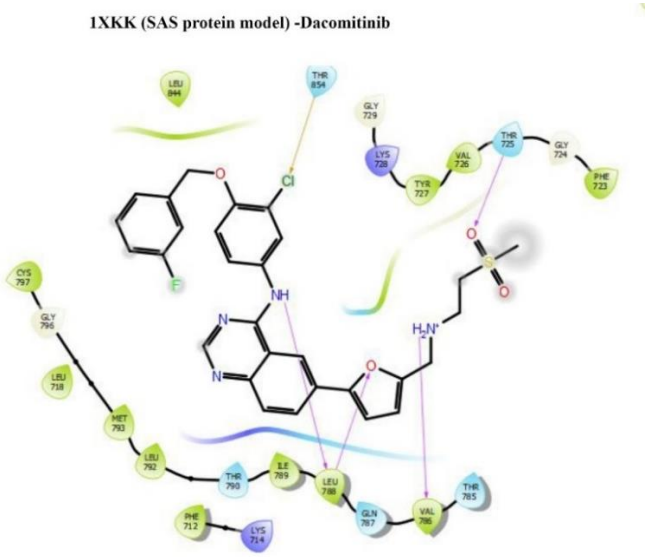


Fig. 3. Ligand interaction diagram of complexes.

Table 4. The MMGBSA (ΔG bind) of kinase inhibitors with model proteins and its difference from template protein 1XKK (ΔΔG) in kcal/mol

Drug	1XKK	American-model		African-model		East Asian-model		South Asian-model		European-model	
	ΔGbind	ΔGbind	ΔΔG	ΔGbind	ΔΔG	ΔGbind	ΔΔG	ΔGbind	ΔΔG	ΔGbind	ΔΔG
Afatinib	-51.18	-12.56	38.62	-46.65	4.53	-29.35	21.83	-24.80	26.38	-22.35	28.88
Dacomitinib	-50.70	-20.46	30.24	-11.50	39.20	-19.60	31.10	-37.24	13.46	-20.10	30.60
Erlotinib	-59.28	-32.56	26.72	-37.53	21.75	-10.25	49.03	-20.12	39.16	-28.95	30.33
Gefitinib	-50.71	-49.96	0.75	-29.14	21.57	-21.10	29.61	-20.28	30.48	-15.21	35.50
Lapatinib	-60.77	-53.48	7.29	-38.68	22.09	-30.25	30.52	-19.25	41.52	-20.77	40.00
Neratinib	-53.52	-32.34	21.28	-35.36	18.16	-40.00	13.52	-29.41	24.11	-32.54	20.98
Poziotinib	-60.27	-20.16	40.11	-20.13	40.14	-18.29	41.98	-21.53	38.74	-28.24	32.08
Sapitinib	-45.37	-40.69	4.32	-19.49	15.88	-11.19	24.18	-20.56	14.81	-45.32	-9.95
Sunitinib	-48.35	-21.00	27.35	-51.23	-2.88	-21.68	26.67	-30.54	17.81	-28.39	19.96

The variant amino acids within the region 859 and 904 of the AMR model showed RMSF values of 2.71 Å and 1.882 Å; the residues within 786, 831, 903, 904, 910, and 917 of AFR-model showed RMSF values of 0.708 Å, 0.194 Å, 0.846 Å, 1.182 Å, 1.287 Å, and 0.801 Å. The residues within position 727 of the EAS model had an RMSF value of 0.516 Å; the residues within the 727 and 904 positions had RMSF values of 0.484 Å and 3.198 Å. The residues within the 904th and 910th positions of the EUR model had RMSF values of 4.655 Å and 1.908 Å. The pharmacophoric properties used for the primary screening of ligands, such as the number of H-bonding donor sites, number of H-bonding acceptor sites, logP, polar surface area, etc. of the top interacting molecules are included in Table S4. The simulation diagram of the drugs that showed good interaction results with the population model is shown in Fig. 4.

MMGBSA

The MMGBSA method computes the relative binding-free energy of each ligand molecule with all the population models and the template protein (Table 4). The ΔG value of the drug is a measure of IC_{50} using the following equation:

$$\Delta G = -RT \ln(pIC_{50})$$

where R is the universal gas constant, T, temperature, and pIC_{50} equals -log10 (IC_{50}).

The variation of ΔG_{bind} is found to be in accordance with the experimental PIC_{50} (Fig. 5). Moreover, the calculated entropy values support dimensionality and probable steric hindrance provided by the inhibitors (Table 5). The difference in binding affinity of kinase inhibitors with the template and the protein models (ΔΔG) was computed for each population (Table 5).

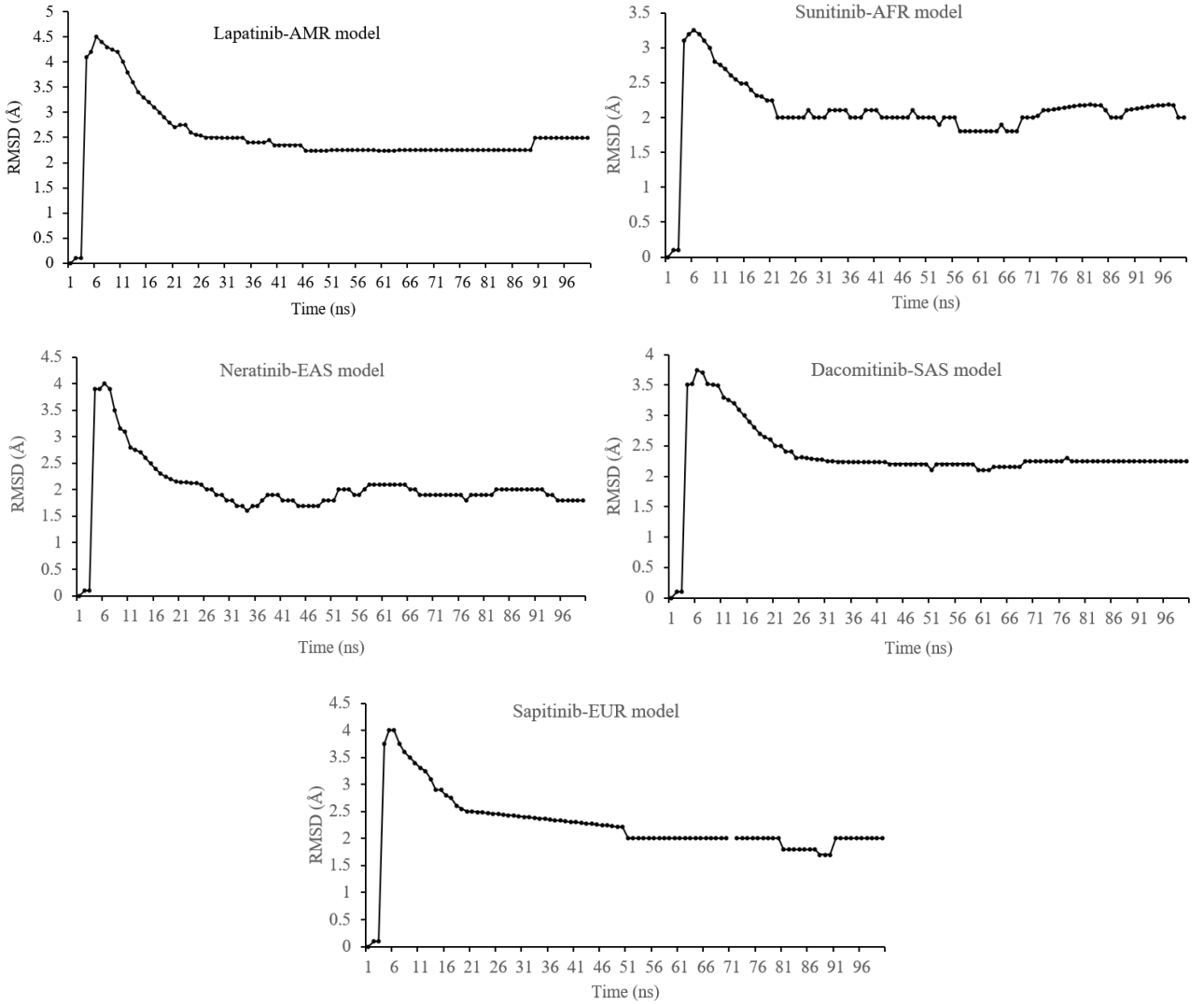


Fig. 4. RMSD plot of ligand-target complexes during molecular dynamic simulation. RMSD, Root mean square deviation.

Table 5. The PIC_{50}, the calculated entropy of ligand molecules, expected bind free energy, and binding free energy obtained from docking ΔG_{bind}(docking) in kcal/mol.

Kinase inhibitors	PIC_{50}	$\Delta G_{bind\ (expected)}$	$\Delta G_{bind\ (docking)}$	$T\Delta S_{Calculated}$
Lapatinib	9.70	-13.850242	-8.683	-5.167242
Poziotinib	8.66	-12.3652676	-8.389	-3.9762676
Erlotinib	9.00	-12.85074	-7.454	-5.39674
Neratinib	8.66	-12.3652676	-7.992	-4.3732676
Afatinib	7.85	-11.208701	-6.994	-4.214701
Dacomitinib	7.36	-10.5090496	-5.933	-4.5760496
Gefitinib	7.12	-10.1663632	-7.447	-2.7193632
Sunitinib	6.76	-9.6523336	-7.871	-1.7813336
Sapatinib	7.92	-11.3086512	-6.351	-4.9576512

The non-strain MMGBSA_ΔG_{bind} has been studied for the complexes. It is the binding/interaction energy without considering the receptor and ligand conformational changes required to form the complex. While studying the MMGBSA_ΔG_{bind} (NS), dacomitinib showed the highest energy for the template model, icotinib showed the highest energy for the AMR model, sunitinib showed high energy for the EAS model, poziotinib for SAS model, and sunitinib for EUR model, respectively (Table S5).

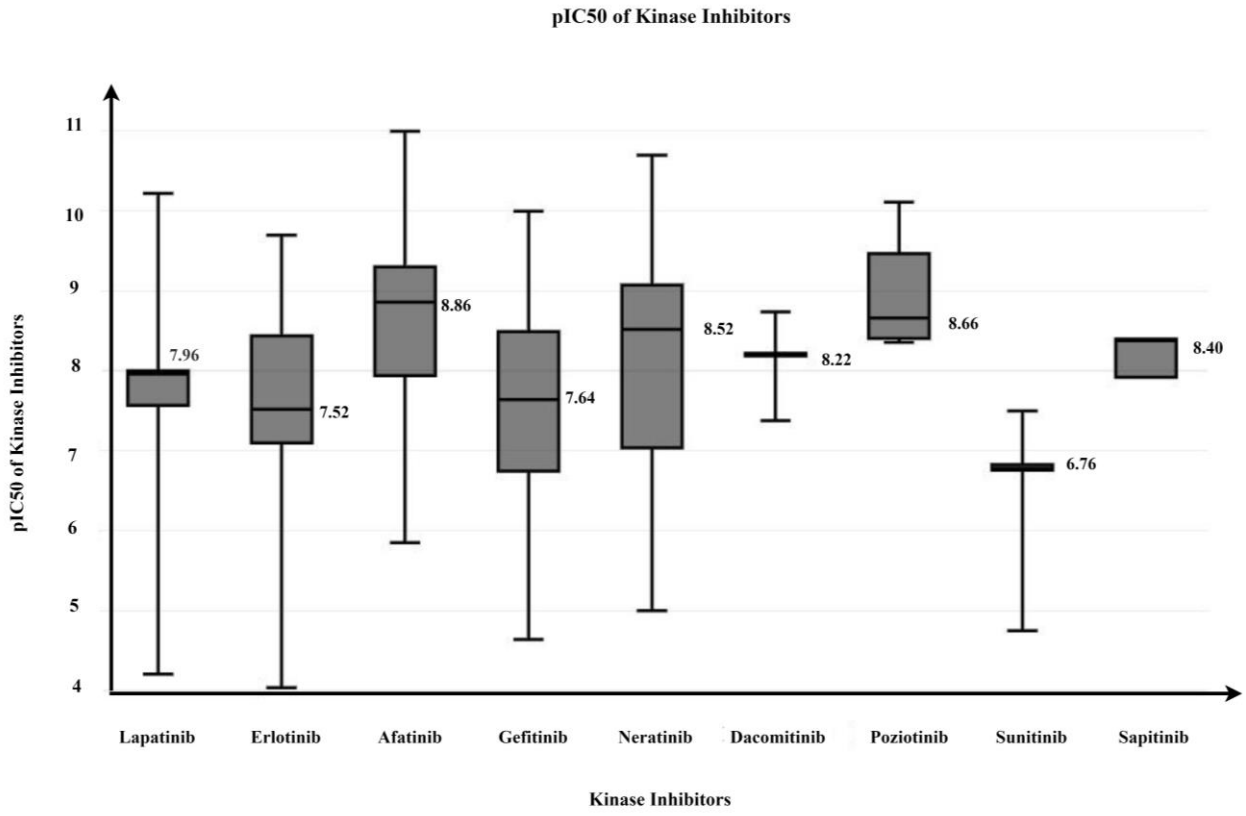


Fig. 5. pIC50 v/s IC50 plot of kinase inhibitors.

The clinical trial reports were analyzed for lapatinib, gefitinib, sunitinib, sapitinib, and neratinib. It has been identified that the drug action of lapatinib was high for AMR and comparatively low for SAS and EAS. The adverse events and death due to disease progression were high for SAS and EAS and minimum for AMR when lapatinib was administered. When gefitinib was administered to the patients, 50% died due to disease progression among EAS and 19% suffered severe adverse events. The drug showed 15% adverse side effects among AMR. The disease progression, as well as the death, was not reported for AMR. When sunitinib was administered to AMR and EUR populations, the patients completed the course without any serious adverse events or death. When it was given to the EAS, there were 19.23% death, 11.6% disease progression, and 40% serious adverse events. When sapitinib was given to the AMR and EUR populations, 0.67% of patients among AMR showed disease progression and 25% of them showed serious adverse events. Whereas there was no disease progression among EUR patients and there were 50% side effects.

DISCUSSION

TNBC, also known as basal-like breast cancer, has been characterized by a deficiency in targeted therapies, aggressive history, and discrete molecular profile (39). The molecular profile states that a high expression of CK5, CK14, caveolin-1, caix, p63, and EGFR/HER1 influences the mammary gland (40,41). The pathway analysis of breast cancer and its subtypes revealed that TP53 mutations, PI3K, and MEK pathway activation, mutagen-activated protein kinase pathway, the Akt pathway, and the poly ADP-ribose polymerase pathway play a key role in TNBC (42). It was found that the overexpression of genes EGFR, KIT, IGF1R, Notch1, Notch4, and LRP6 are the initial phase in the TNBC pathway.

If the phase 1 genes are inhibited, then the activation of other pathways could be downregulated, thereby preventing the progression, proliferation, and translocation of TNBC cells. The genome-wide association study analyzes the genetic variants in different individuals, thus correlating variants and phenotypes (43). Researchers have demonstrated that about 85% of variants occur

within populations and 15% occur across the population (30). This might be the reason behind the variation in drug action and drug response across and within the population. Various *in-silico* approaches are available to identify existing variants' phenotypic and disease correlation.

The variant annotation showed that variants rs2227983 and rs371228501 were reported to involve malignant breast neoplasm and breast carcinoma, whereas rs2293347, rs35918369, rs2072454, and rs2227983 were involved in lung carcinoma. The variants rs55959834, rs41420046, rs2229066, rs1140475, rs41396448, rs2293347, and rs55737335 are synonymous; rs138240620, rs371228501, rs538497054, rs575565383, rs201830126, rs144496976, rs17290699, and rs542967903 are missense variants and rs55959834 is both splice region and synonymous variant. Among 43223 variants, 98 were completely annotated for identifying 1000 genome continental allele frequencies. It has been identified that 45 variants belonged to the AFR, 32 variants belonged to the AMR, 47 variants belonged to EAS, 46 variants belonged to the EUR, and 38 variants belonged to SAS. Among these 172 variants, 7 variants including rs2072454, rs2227983, rs17290169, rs2227984, rs10258429, rs1140475, and rs2293347 were found commonly among all the population class. The positions of coding variants have been retrieved.

The P00533 was identified as the EGFR protein. The protein kinase domain has been identified within the 719^{th}-979^{th} position of P00533. Amino acids within the 718-726 region are nucleotide phosphate binding in nature within the kinase domain. K745 is identified as the binding site, and D837 is found as the active site involved in enzyme catalysis. The protein kinase ATP binding domain lies within residues from the 718^{th}-745^{th} position, and the protein tyrosine kinase domain falls within the region 833-845.

1XKK has been identified as the template protein of EGFR as it possesses all the amino acids present in the kinase domain. The 11 most frequent variations within kinase domain are 780 G/S, 785 L/P, 797 S/C, 859 P/S, 868 D/H, 874 P/A, 885 R/Q, 894 P/A, 940 A/D, 945 Y/C, and 950 D/N. The mutant protein sequence was generated by incorporating these variants into the protein sequence.

The predicted models generated by homology modeling were validated using various validation techniques. The best model for each population has been selected based on the RMS deviation, Ramachandran plot, ERRAT plot, and QMEAN values. The biomolecular interaction network helps us to predict the interaction possibility of the selected molecules. As biofunctionality is closely related to molecular interactions, this sort of networking and characterization are found to be useful in the functional enrichment of the molecules. Usually, molecules with a top degree, closeness centrality, and betweenness centrality will be selected as functionally relevant. Here, from Fig. 1 and Table S3, afatinib, canertinib, dacomitinib, erlotinib, gefitinib, lapatinib, icotinib, neratinib, poziotinib, sapitinib, sunitinib, and vandetanib are found to be interacting with EGFR and are with sufficiently high closeness centrality and betweenness centrality. These molecules are found to be functionally relevant and interactionally predominant kinase inhibitors of EGFR. These inhibitors were subjected to molecular docking with the parent target template, 1XKK and the models designed for the populations within the kinase domain binding site.

During the evaluation of PDI, the O4 atom of sapitinib formed hydrogen bond interactions with the N atom of VAL884 and CD PRO885 within 3.31 Å and 3.33 Å, respectively. Sunitinib developed a pi-H bond between its 6-ring and CD1 atom of LEU785 within 3.72 Å. O5 and O6 atoms of lapatinib formed hydrogen bonding interactions between CD and N atoms of PRO 761 within 3.93 Å and 2.98 Å, respectively. Br1 atom of vandetanib formed a hydrogen bond with the O atom of VAL842 within 3.62 Å. O5 and O6 atoms of lapatinib formed hydrogen bonding interactions between CD and N atoms of PRO761 within 3.93 Å and 2.98 Å, respectively. It also developed a pi-H bond between its 5-ring and CB atom of ALA763 within 3.97 Å.

The O3 atom of icotinib formed a hydrogen bond with NH2 of ARG844 within 2.92 Å. O5

atom of poziotinib shared hydrogen bonding interactions with NE and NH1 of ARG897 within 3.26 Å and 3.52 Å. Its N9 atom formed a hydrogen bond with the O atom of LEU866 3.28 Å. These drugs interact with the sensitive region of the kinase domain.

The O3 atom of afatinib shared a hydrogen bond interaction with the NE atom of the ARG844 target within 3.25 Å. The N6 and N10 atoms of canertinib formed a hydrogen bond with NH1 atom of ARG844 within 3.05 Å and CA atom of ALA763 within 3.8 Å. It also formed a pi-H bond between the 6-ring of the drug and the CB atom of ALA763 within 3.93 Å. Dacomitinib formed a hydrogen bond between N7 of the drug and OE1 of GLU770 within 3.20 Å.

O3 of erlotinib developed an H-bond with NH1 atom of ARG844 within 2.97 Å. N7 and C19 atoms of gefitinib showed hydrogen bond interaction with OD1 atom of ASP863 within 3.19 Å and 3.37 Å. It also formed a hydrogen bond between the N8 atom of the drug and the CB atom of ALA751 within 3.38 Å. CL1 and N5 of neratinib shared hydrogen bonds with O atoms of LYS762 within 3.19 Å and LEU866 within 2.91 Å. The complexes were formed within the conserved region of the kinase domain.

The ligand-target complexes were then subjected to MD simulation to check the KI-target complex stability. It was observed that the drug, neratinib, tends to lose its interaction after 0.5 ns. At the same time, the O3 atom of lapatinib showed hydrogen bonding interactions with NH1 and NH2 atoms of ARG844 within 3.14 Å and 3.20 Å, respectively. Sapitinib developed an ionic bond with ASP840 and a hydrogen bond with ARG836 of the sensitive region within the kinase domain of 1XKK. Sunitinib had both ionic bonds with ASP863 and THY725 residue of the conserved region and LYS753 of the sensitive region within the domain. Vandetinib interacted with the residues VAL842 and VAL859 of the sensitive region.

The evaluation of ligand affinity within the active site of the target could be studied through MMGBSA. MMGBSA calculates the free energy state by considering three energy terms such as Ebond (bond, angle, and dihedral), Eel (electrostatic), and EvdW (van der Waals) interactions, Gpol (polar contribution), Gnp (non-polar contribution), and the last term is T (absolute temperature) multiplied by S (entropy). Here the non-polar solvation energy is in linear relation to the solvent-accessible surface area. Coulomb's law was used to calculate the electrostatic term. In order to calculate the entropy term, all the water molecules and residues are > 8 Å from the ligand. The ligand-target complexes such as lapatinib-1XKK, gefitinib-AMR model, sunitinib-AFR model, sunitinib-SAS model, and sapitinib-EUR model were found with good MMGBSA_ΔG_{bind} energies. Lapatinib and gefitinib showed interaction with the AMR variant; sunitinib interacted with the AFR variant, dacomitinib interacts with the SAS variant, neratinib interacted with the EAS variant, sapitinib interacted with the EUR variant.

CONCLUSION

The study elucidates the involvement of single nucleotide variants within the domain of EGFR in drug action. An increased expression of EGFR was found among 'Triple Negative Breast Cancer cell lines. Thus, EGFR mutant protein has been considered as the target. The EGFR variant annotation has been carried out. Based on the presence of SNVs, the kinase domain is classified as the sensitive region (the region with variants) and the conserved region (the region without variants). The remaining region of the Uniprot ID P00533 is identified as the offsite region. PDI explains the approved kinase inhibitors that are reported to interact with EGFR protein. Few drugs were identified to interact with the sensitive region and are considered sensitive drugs. This may be the reason for the variation in drug action and response. This has been illustrated by designing mutant model proteins of EGFR for populations AMR, EUR, EAS, AFR, and SAS. The variation of drug interaction and its thermodynamic as well as kinetic stability were studied. The results were cross-checked with the existing clinical trial reports. The observed results complemented the clinical trial reports.

Acknowledgments

The authors would like to acknowledge the Ministry of Electronics and Information Technology, Government of India' for meeting the research fellowship under theVisvesvaraya Ph.D. Scheme for Electronics and IT. The authors express their gratitude to Biopharma Solutions, the industry partner of AMMAS research lab for their support and help in completing the project.

Conflict of interest statement

The authors declared no conflict of interest in this study.

Authors' contribution

H. Vyshnavi contributed to the investigation and data interpretation and analysis, and also wrote the manuscript; K. Namboori contributed to the concept, design, and data interpretation and analysis. The finalized article was approved by all authors.

REFERENCES

1. Alkabban FM, Ferguson T. Breast Cancer. Treasure Island (FL): StatPearls Publishing; 2022. pp. 1-29. Available from: https://www.ncbi.nlm.nih.gov/books/NBK482286/.
2. Selase A, Cynthia AD, Newman O, Williams A, Michael O. Palmatine sensitizes chemoresistant triple negative breast cancer cells *via* efflux inhibition of Multidrug resistant protein 1. Sci Afr. 2021;14:(e01022),1-8. DOI: 10.1016/j.sciaf.2021.e01022.
3. Al-Mahmood S, Sapiezynski J, Garbuzenko OB, Minko T. Metastatic and triple-negative breast cancer: challenges and treatment options. Drug Deliv Transl Res. 2018;8(5):1483-1507. DOI: 10.1007/s13346-018-0551-3.
4. Sepahdar Z, Miroliaei M, Bouzari S, Khalaj V, Salimi M. Surface engineering of *Escherichia coli*-derived OMVs as promising nano-carriers to target EGFR-overexpressing breast cancer cells. Front Pharmacol. 2021;12:719289,1-16. DOI: 10.3389/fphar.2021.719289.
5. You KS, Yi YW, Cho J, Park JS, Seong YS. Potentiating therapeutic effects of epidermal growth factor receptor inhibition in triple-negative breast cancer. Pharmaceuticals (Basel). 2021;14(6): 589,1-76. DOI: 10.3390/ph14060589.
6. Sebastiani P, Timofeev N, Dworkis DA, Perls TT, Steinberg MH. Genome-wide association studies and the genetic dissection of complex traits. Am J Hematol. 2009;84(8):504-515. DOI: 10.1002/ajh.21440.
7. Fragomeni SM, Sciallis A, Jeruss JS. Molecular subtypes and local-regional control of breast cancer. Surg Oncol Clin N Am. 2018;27(1):95-120. DOI: 10.1016/j.soc.2017.08.005.
8. Riera C, Padilla N, de la Cruz X. The complementarity between protein-specific and general pathogenicity predictors for amino acid substitutions. Hum Mutat. 2016;37(10):1013-1024. DOI: 10.1002/humu.23048.
9. Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. SIFT missense predictions for genomes. Nat Protoc. 2016;11(1):1-9. DOI: 10.1038/nprot.2015.123.
10. Lu Guanting, Ma Liya, Xu Pei, Xian Binqiang, Wu Lianying, Ding Jianying, *et al*. A *de novo* ZMIZ1 pathogenic variant for neurodevelopmental disorder with dysmorphic facies and distal skeletal anomalies. Front Genet. 2022;13:840577,1-14. DOI: 10.3389/fgene.2022.840577.
11. Rentzsch P, Schubach M, Shendure J, Kircher M. CADD-splice-improving genome-wide variant effect prediction using deep learning-derived splice scores. Genome Med. 2021;13(1):31,1-12. DOI: 10.1186/s13073-021-00835-9.
12. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, *et al*. The ensembl variant effect predictor. Genome Biol. 2016;17(1):122,1-14. DOI: 10.1186/s13059-016-0974-4.
13. Bhattacharya R, Rose PW, Burley SK, Prlić A. Impact of genetic variation on three-dimensional structure and function of proteins. PLoS One. 2017;12(3):e0171355,1-22. DOI: 10.1371/journal.pone.0171355.
14. Bhullar KS, Lagarón NO, McGowan EM, Parmar I, Jha A, Hubbard BP, *et al*. Kinase-targeted cancer therapies: progress, challenges and future directions. Mol Cancer. 2018;17(1):48,1-20. DOI: 10.1186/s12943-018-0804-2.
15. Szklarczyk D, Santos A, von Mering C, Jensen LJ, Bork P, Kuhn M. STITCH 5: augmenting protein-chemical interaction networks with tissue and affinity data. Nucleic Acids Res. 2016;44(D1):D380-D384. DOI: 10.1093/nar/gkv1277
16. Muhseen ZT, Kadhim S, Yahiya YI, Alatawi EA, Aba Alkhayl FF, Almatroudi A. Insights into the binding of receptor-binding domain (RBD) of SARS-CoV-2 wild type and B.1.620 variant with hACE2 using molecular docking and simulation approaches. Biology (Basel). 2021;10(12):1310,1-15. DOI: 10.3390/biology10121310.
17. Yousuf Z, Iman K, Iftikhar N, Mirza MU. Structure-based virtual screening and molecular docking for the identification of potential multi-targeted inhibitors against breast cancer. Breast cancer (Dove Med Press). 2017;9:447-459. DOI: 10.2147/BCTT.S132074.
18. Acharya, R, Chacko, S, Bose P, Lapenna A, Pattanayak SP. Structure based multitargeted molecular docking analysis of selected furanocoumarins against breast cancer. Sci Rep. 2019;9(1):15743,1-13. DOI: 10.1038/s41598-019-52162-0.

19. Maffucci I, Hu X, Fumagalli V, Contini A. An efficient implementation of the Nwat-MMGBSA method to rescore docking results in medium-throughput virtual screenings. Front Chem. 2018;6:43,1-14. DOI: 10.3389/fchem.2018.00043.
20. Fatriansyah JF, Rizqillah RK, Yandi MY, Fadilah, Sahlan M. Molecular docking and dynamics studies on propolis sulabiroin-A as a potential inhibitor of SARS-CoV-2. J King Saud Univ Sci. 2022;34(1):101707,1-9. DOI: 10.1016/j.jksus.2021.101707.
21. Zafar F, Gupta A, Thangavel K, Khatana K, Sani AA, Ghosal A, *et al*. Physicochemical and pharmacokinetic analysis of anacardic acid derivatives. ACS Omega. 2020;5(11):6021-6030. DOI: 10.1021/acsomega.9b04398.
22. PK. Design and development of a pharmacogenomic model for breast cancer to study the variation in drug action and side effects. Int J Appl Pharm. 2022;14(3):61-68. DOI: 10.22159/ijap.2022v14i3.44356.
23. National Center for Biotechnology Information (NCBI). Bethesda (MD): National Library of Medicine (US). Updated to 2022. Available from: https://www.ncbi.nlm.nih.gov/gene/1956.
24. The UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res. 2021;49(D1):D480-D489. DOI: 10.1093/nar/gkaa1100.
25. Safran M, Rosen N, Twik M, BarShir R, Iny Stein T, Dahary D, *et al*. The genecards suite. In: Abugessaisa I, Kasukawa T, editors. Practical guide to life science databases. Springer;2022. pp: 27-56. DOI: 10.1007/978-981-16-5812-9_2.
26. Sherry ST, Ward MH, Kholodov, M Baker, J Phan, L Smigielski, *et al*. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. 2001;29(1):308-311. DOI: 10.1093/nar/29.1.308.
27. Martina M, Acquadro A, Barchi L, Gulino D, Brusco F, Rabaglio M, *et al*. Genome-wide survey and development of the first microsatellite Markers database (AnCorDB) in *Anemone coronaria* L. Int J Mol Sci. 2022;23(6):3126,1-17. DOI: 10.3390/ijms23063126.
28. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, *et al*. The protein data bank. Nucleic Acids Res. 2000;28(1):235-242. DOI: 10.1093/nar/28.1.235.
29. Bienert S, Waterhouse A, de Beer TA, Tauriello G, Studer G, Bordoli L, *et al*. The SWISS-MODEL repository-new features and functionality. Nucleic Acids Res. 2017;45(D1):D313-D319. DOI: 10.1093/nar/gkw1132.
30. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 1993;2(9):1511-1519. DOI: 10.1002/pro.5560020916.
31. National Institute of Mental Health. (2011). Borderline personality. DHHS Publication No. 11-7790. Washington, DC: U.S. Government Printing Office. Available from: www.cancer.gov/about-cancer/treatment/drugs/breast
32. National Comprehensive Cancer Network (NCCN) guidelines. Genetic/familial high-risk assessment: breast, ovarian and pancreatic. Version 2.2022. 2022. Available from: https://www.nccn.org/guidelines/guidelines-detail?category=1&id=1419.
33. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, *et al*. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498-2504. DOI: 10.1101/gr.1239303.
34. Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, *et al.* Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. J Med Chem. 2006;49(21):6177-6196. DOI: 10.1021/jm051256o.
35. Kim S, Thiessen PA, Bolton EE, Chen, J, Fu G, Gindulyte A, *et al*. PubChem substance and compound databases. Nucleic Acids Res. 2016;44(D1):D1202-D1213. DOI: 10.1093/nar/gkv951.
36. Phillips JC, Hardy DJ, Maia JDC, Stone JE, Ribeiro JV, Bernardi RC, *et al*. Scalable molecular dynamics on CPU and GPU architectures with NAMD. J Chem Phys. 2020;153(4):044130,1-34. DOI: 10.1063/5.0014475.
37. Genheden S, Ryde U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opin Drug Discov. 2015;10(5): 449-461. DOI: 10.1517/17460441.2015.1032936.
38. Brai A, Riva V, Clementi L, Falsitta L, Zamperini C, Sinigiani V, et al. Targeting DDX3X helicase activity with BA103 shows promising therapeutic effects in preclinical glioblastoma models. Cancers. 2021;13(21):5569,1-26. DOI: 10.3390/cancers13215569.
39. Alluri P, Newman LA. Basal-like and triple-negative breast cancers: searching for positives among many negatives. Surg Oncol Clin N Am. 2014;23(3):567-577. DOI: 10.1016/j.soc.2014.03.003.
40. Petrelli F, Cabiddu M, Ghilardi M, Barni S. Current data of targeted therapies for the treatment of triple-negative advanced breast cancer: empiricism or evidence-based? Expert Opin Investig Drugs. 2009;18(10):1467-1477. DOI: 10.1517/13543780903222268.
41. Chakrabarty A, Chakraborty S, Bhattacharya R, Chowdhury G. Senescence-induced chemoresistance in triple negative breast cancer and evolution-based treatment strategies. Front Oncol. 2021;11:674354,1-14. DOI: 10.3389/fonc.2021.674354.
42. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27-30. DOI: 10.1093/nar/28.1.27.
43. MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, *et al*. The new NHGRI-EBI catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res. 2017;45(D1): D896-D901. DOI: 10.1093/nar/gkw1133.

SUPPLEMENTARY MATERIALS

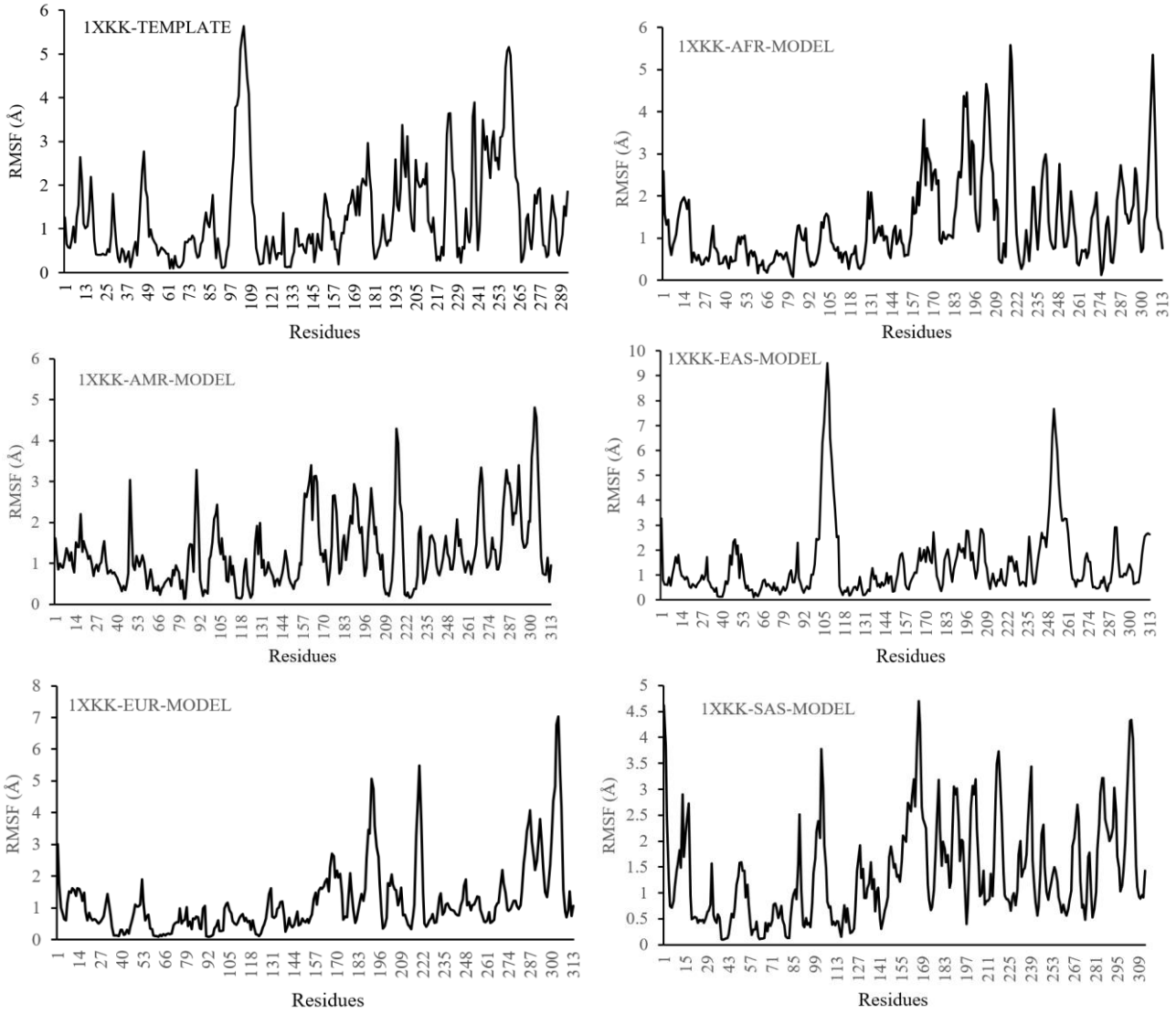


Fig. S1. The RMSF plots of ligand-target complexes after molecular dynamic simulation. RMSF, Root mean square fluctuation.

Table S1. The variant annotation.

Single nucleotide variant	Location	Allele	Protein position	Amino acids	Codons
rs55959834	7:55174034-55174034	A	725	T	acG/acA
rs138240620	7:55174039-55174039	G	727	Y/C	tAt/tGt
rs41420046	7:55191736-55191736	A	784	E	gaG/gaA
rs371228501	7:55191740-55191740	T	786	R/C	Cgt/Tgt
rs2229066	7:55191757-55191757	T	791	R	cgC/cgT
rs530256683	7:55192840-55192840	T	855	Y	taC/taT
rs1140475	7:55198724-55198724	C	858	T	acT/acC
rs538497054	7:55198725-55198725	A	859	V/I	Gtt/Att
rs41396448	7:55198763-55198763	T	871	D	gaC/gaT
rs575565383	7:55198815-55198815	T	889	P/S	Cct/Tct
rs201830126	7:55200330-55200330	A	910	A/T	Gca/Aca
rs144496976	7:55200352-55200352	A	917	R/H	cGt/cAt
rs17290699	7:55201204-55201204	C	943	H/P	cAt/cCt
rs542967903	7:55198857-55198857	A	948	V/I	Gtc/Atc
rs2293347	7:55201223-55201223	T	949	D	gaC/gaT
rs55737335	7:55201256-55201256	G	960	E	gaA/gaG

Table S2. List of kinase inhibitors**.**

Afatinib	Neratinib	Sorafenib	Regorafenib	Ripasudil	Ribocicilib
Aloitinib	Poziotinib	Nilotinib	Dabrafenib	Alectinib	Abemaciclib
Dacomitinib	Sapitinib	Crizotinib	Trametinib	Cobimetinib	Baricitinib
Canertinib	Soratinib	Axitinib	Ibrutinib	Lenvatinib	
Erlotinib	Sunitinib	Tofacitinib	Nintedanib	Palbociclib	Simotinib
Lapatinib	Vandetinib	Bosutinib	Idelalisib	Radotinib	Binimetinib
Gefitinib	Tucatinib	Cabozantinib	Ceritinib	Osimertinib	Lorlatinib
Icotinib	Valertinib	Ponatinib	Apatinib /Rivoceranib	Olmutinib	Flumatinib
Peficitinb	Selumetinib	Tirabrutinib	Filgotinib	Orelabrutinib	Trilaciclib
Avapritinib	Ripretinib	Almonertinib	Tirbanibulin	Tepotinib	Zanubrutinib

Table S3. The network analysis.

Node	Betweenness centrality	Closeness centrality
Afatinib	8.43	9.50
Erlotinib	2.10	9.00
Gefitinib	2.10	9.00
Dacomitinib	5.13	9.00
Neratinib	3.33	8.50
Lapatinib	0.80	8.50
Icotinib	0.33	7.50
Sunitinib	0.33	7.50
Poziotinib	0.33	6.33
Sapitinib	0.33	5.33

Table S4. Pharmacophoric properties of kinase inhibitors.

Poroperties	Afatinib	Dacomitinib	Erlotinib	Gefatinib	Lapatinib	Neratinib	Poziotinib	Sapitinib	Sunitinib
Molecular weight	485.946	469.947	393.442	446.909	581.066	557.052	491.349	473.935	398.48
cLogP	3.6354	4.7187	3.0713	3.9851	4.7281	4.3389	5.2946	3.4467	1.836
cLogS	-5.476	-6.017	-3.527	-5.062	-8.061	-6.074	-6.716	-5.256	-3.471
H-Acceptors	8	7	7	7	8	9	7	8	6
H-Donors	2	2	1	1	2	2	1	2	3
Total surface area	362.94	355.7	319.97	333.75	424.92	437.56	352.68	350.2	312.95
Drug likeness	-0.79682	-0.60745	-5.9718	0.47937	-4.2297	-4.0813	-4.6966	4.6824	8.335
LE from molecule name	0.08037	0.080603	0.17751	0.17294	0.12621	0.06872	0.066509	0.080639	0.10754
LLE from molecule name	-1.6433	-2.7798	0.68105	-0.07723	-1.0481	-2.3352	-3.6948	-1.507	0.43735
LELP from molecule	45.229	58.542	17.302	23.043	37.461	63.139	79.607	42.742	17.072
Mutagenic	None	none	none	none	none	high	high	none	none
Tumorigenic	None	none	none	none	none	none	low	none	none
Reproductive effective	None	none	none	none	none	none	low	none	none
Irritant	None	none	none	none	none	none	low	none	high
Nasty functions							polyhalo aromatic ring		
Shape index	0.5	0.57576	0.55172	0.58065	0.6	0.55	0.54545	0.54545	0.55172
Non-H atoms	34	33	29	31	40	40	33	33	29
Non-C/H atoms	10	9	7	9	11	10	10	10	7
Metal-atoms	0	0	0	0	0	0	0	0	0
Electronegative atoms	10	9	7	9	11	10	10	10	7
Stereo center	1	0	0	0	0	0	0	0	0
Rotatable bonds	8	7	10	8	11	11	6	7	7

LE, Ligand efficiency, LLE, ligand-lipophilicity efficiency; LELP, lipophilic efficiency.

Table S5. MMGBSA_ΔG_{bind} (NS) of kinase inhibitors and target protein molecules in kcal/mol.

Kinase inhibitor	1xkk	American model	African model	East Asian model	South Asian model	European model
Afatinib	-62.00	-31.12	-57.3	-54.6	-52.21	-37.25
Dacomitinib	-57.00	-29.26	-24.25	-38.68	-48.24	-35.12
Erlotinib	-65.89	-48.57	-54.45	-23.8	-38.93	-39.50
Gefitinib	-54.87	-57.86	-48.57	-30.56	-27.55	-28.15
Lapatinib	-68.1	-65.24	56.14	-42.96	-26.18	-27.85
Neratinib	-62.15	-58.32	-44.98	-58.01	-34.15	-39.21
Poziotinib	-70.18	-41.85	-60.15	-54.65	-31.63	-35.14
Sapitinib	-48.98	-50.15	-30.78	-27.86	-32.90	-53.38
Sunitinib	-56.45	-33.45	-61.84	-33.54	-37.84	-37.64