RESEARCH PAPER

Discovery of Efficient EGFR Inhibitors with the Assistance of Cheminformatics Approaches: Structure based Drug Design, Computational docking and ADME Parameters Analysis and Molecular dynamics simulation

Mohan Anbuselvam[1], Sridhar Muthusami[1, 2*]

[1]Department of Biochemistry, Karpagam Academy of Higher Education, Coimbatore- 641021, Tamil Nadu, India.
[2]Centre for Cancer Research, Karpagam Academy of Higher Education, Coimbatore- 641021, Tamil Nadu, India.

***Corresponding Author;** Dr. Sridhar Muthusami,
*Associate Professor and Head, Centre for Cancer Research, Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India.E-mail ID: sridharuniv@gmail.com ; sridhar.m@kahedu.edu.in

Abstract
Cervical Cancer (CC), the fourth deadliest gynecologic malignancy among women globally. CC has been found to be primarily driven by cellular growth factors, thus, making then highly potential targets for therapy. Although, targeting the dysregulated kinase activity of the epidermal growth factor receptor (EGFR) with small molecules remains challenging in C treatment. Therefore, in the present investigation, we performed structure based virtual screening of EGFR to identify pharmacologically efficient EGFR inhibitors from large chemical database. At the end of this investigation, theoretical ADME parameters were evaluated for the top three filtered molecules to assess the drug likeness. The binding stability of the Protein-ligand complexes was examined by molecular dynamics simulation. These computational investigations suggest that the three small molecules could be opting for further experimental studies to ascertain the inhibitory effect on EGFR kinase over expression.

Keywords: Cervical cancer, EGFR, virtual screening, molecular dynamics simulation
How to cite this article: Anbuselvan M, Muthusami S. Discovery of Efficient EGFR Inhibitors with the Assistance of Cheminformatics Approaches: Structure based Drug Design, Computational docking and ADME Parameters Analysis and Molecular dynamics simulation. Int J Drug Deliv Technol. 2026;16(41s): 968-976. DOI: 10.25258/ijddt.16.41s.102

Introduction

Cervical Cancer (CC) is the fourth leading cause of cancer-related mortality among women globally. According to the American Cancer Society, in 2025, around 6, 60, 000 cases of cervical cancer were diagnosed and about 3, 42, 000 deaths were occured globally; furthermore, around 90% of cervical cancer incidence occurs in low and middle-income countries [1, 2]. The Epidermal Growth Factor Receptor (EGFR) is a transmembrane glycoprotein belongs to the largest family of receptor tyrosine kinase and it is located in 7P12-14 at chromosome 7 and consists of 1186 amino acids, the molecular mass of the receptor 170 KDa encoded by HER1 and acts as a proto-oncogene [3]. The EGFR consists of four diverse family members including EGFR/HER1/ErbB-1, ErbB-2/HER2/neu, ErbB3 (HER3) and ErbB4 (HER4) [4]. Structurally, the tyrosine kinase domain EGFR 1 possesses 82% conserved with C-erbB-2 proteins. Even though, it's has distinctive biological mechanisms especially tyrosine kinase activity. Structurally, EGFR is composed of an extracellular ligand binding domain, a single alpha-helical transmembrane gap, and an intracellular carboxyl-terminal domain with tyrosine kinase activity. EGFR is mainly activated by some crucial ligands, including Epigen (EPG), Amphiregulin (AREG), transforming growth factor-α (TGF-α), Heparin-binding EGF-like growth factor (HB-EGF), Epiregulin, Betacellulin (BTC), and Neuregulins (NRG). Among these, epidermal growth factor and transforming growth factor-α serveas the endogenous ligands [5]. After ligand binding, this leads to receptor dimerization and autophosphorylation at the carboxyl domain of EGFR, triggering the activation of several crucial intracellular downstream signaling pathways including RAS/RAF/MAPK pathway, STAT pathway, PI3K/ AKT/mTOR pathway and also protein kinase C (PKC). The activation of downstream signaling pathways triggers normal biological mechanism of EGFR, including cell proliferation, survival, migration, adhesion, and differentiation. Mutations occurs in the kinase domain region of EGFR, the normal biological mechanisms of EGFR results inover expressed [6, 7]. The over expression of EGFR has been frequently associated with a wide range of malignancies in humans, including bladder, ovary, colorectal, breast, neck, lungs, kidney, prostate, brain and cervical cancer [8]. Therefore, it is of interest to report the Moleculardocking and dynamic simulation analysis of EGFR with compounds from the ZINC database.

Materials and Methods:

Structure based Virtual screening:

In order to find efficient pharmacologically active small molecules with desired Biopharmaceutical features. We employed the most promising

*Author for Correspondence: sridharuniv@gmail.com

computational approach receptor based virtual screening. In the current *insilico* screening investigation, we selected the appropriate crystal structure of target receptor EGFR using PDB accession code: 1M17 with co-crystallized erlotinib was extracted from macromolecule structural database protein data bank [9]. The "protein preparation " panel was utilized to refine the receptors topology file to get the appropriate refined receptor file for docking. In the preparation process, drug molecule erlotinib was eliminated from target receptor along with water molecules, simultaneously, bond orders was rectified and appropriate charges was assigned to get refined topology file, force field OPLS3e was applied to execute the receptor energy minimization until, the RMSD threshold value reached 0.30Å [10]. Furthermore, the Receptor Grid Generation panel was applied to generate grid box for refined target receptor. Key functional residues (M69, L694A, G772A, L820A, and L768A) were selected for grid generation, and the docking grid was centered on these active-site residues.. The size of the grid box was set at coordinates X=20.5209Å, Y=-1.2884Å, Z=55.4471 within 14 Å radius [11]. After that, the 500,000 small molecules topology from ZINC chemical library were obtained in the SDF format, after that, all the structure were further imported and prepared using ligprep panel to get the appropriate ligand topology file for docking. During the ligand refinement process, the tautomers were generated, proper atomic charges were added for each ligand and force field OPLS3e was applied to executed energy minimization process. Subsequently, we applied ligand docking panel which is having three distinct funnel type scrutinization steps. Namely HTVS, SP and XP and all other parameters were kept at default settings through the docking process. At the end of the screening process, the key interactions between the receptor and ligands were analyzed with the help of Glide XP visualizer panel along with binding affinity score [12, 13].

Theoretical ADME Profile Analysis

The pharmaceutical sector endeavors to bring out potential drug candidates to the market at an affordable cost for clinical use. Nowadays, vast number of drug discovery and development process ended at final stages of clinical trials, due to toxicity and inappropriate biocompatibility profiles. The early assessment of drug-like properties of chemical entities is an important and integral part of drug discovery process to significantly reduce the use of experimental animals and other resources. The experimental analysis engulfs considerable amount as well as other resources to determine the pharmacokinetics, safety and efficacy of drug candidate [14]. Computational pharmacology is a rapidly growing field and gained on much attention in modern research and mainly contributes their role to develop innovative tool to generate and analyze the biological data. Therefore, in this analysis, the simple, fast, reliable *insilico* tool Qikprop [15] was applied to investigate the theoretical ADME profile of the selected lead molecules. The following drug like profile were investigated: these are aqueous solubility, MDCK permeability (QPPMDCK), Pfizer rule, blood brain barrier, Log IC_{50} value for K+ channels (QP log HERG) and overall human oral absorption percentage.

Molecular Dynamics simulation

MD simulation analysis was carried out to investigate the conformational changes and binding stability of the protein-ligand complexes. For the molecular dynamics simulation studies, we employed most reliable computational package Desmond [16]. Initially, the protein-ligand complexes were imported into Desmond panel, subsequently we utilized system "builder panel" to prepare the system for dynamics. Then, we set the simulation box size at 10Å. Subsequently, the counter ions Na^+ and Cl^- were added and neutralize the system. The NPT ensemble with the constant temperature 300K and a pressure 1 bar was maintained for the entire simulation system, furthermore, the OPLS3e force field were applied for entire simulation [17]. Finally, the simulation interaction diagram tool was employed to examine the structural behavior of intricate interactions between protein-ligand complexes.

Results and Discussion

Selection of Lead Molecules:

For the identification of potential small molecules inhibitors for EGFR, we applied a promising computer assisted strategies protein based virtual screening. The prepared 500,000 molecules were permit to bind with biding cavity region of EGFR with the help of Glide HTVS, SP and XP docking mode and we filtered three lead molecules with binding affinity score -8.13 Kcal/mol comparable to or better than the reference molecules erlotinib. Docking results indicate that the lead molecules interact with same experimentally determined conserved functional residues. Three lead molecules were selected based on the binding affinity, which efficiently binded with binding cavity of the EGFR with adequate pharmacokinetic profile for human oral use qualifying as a potential drug-like molecule from the database. The results showed that the three lead molecules were potentially interacted with the active site of EGFR were shown in figure 1 and their 2Dstructure were presented in figure 2. The docking investigation revealed that the hit Zinc14618306 formed only two hydrogen bond interactions with EGFR. The residues MET 769 was involved in hydrogen bond interaction. The dock score was observed for Zinc14618306 -8.27Kcal/mol and reported in table 1**.** Two hydrogen bond interactions were formed between Zinc14638402 into EGFR, that the residues are participated in hydrogen bond interactions MET 769 and ASP 831. The ASP 831

residues formed four hydrogen bond interactions with Zinc14638402. The binding free energy was noted -8.18Kcal/mol. The docking simulation of Zinc14610013 into EGFR formed only one hydrogen bond interaction. The binding free energy was calculated glide score of-8.13Kcal/mol. The conclusion of the docking results was MET 769, is a common interacting residue across all docked complexes, suggesting that MET 769 plays a critical rose in the function of EFGR.

Table 1: Glide extra-precision (XP) results for the three lead molecules, by use of Schrodinger 10.2.

S. No	ZINC ID	Glide score	Hydrogen bond interaction
1.	Zinc14618306	-8.27	MET 769 (2) , LYS 721 ,ASP 831 (2)
2.	Zinc14638402	-8.18	MET 769, LYS 721, ASP 831 (4)
3.	Zinc14610013	-8.13	GLU 738, ASP 776, MET 769 (2)

The compound IDs are labeled as such under the ZINC database; Glide score (Kcal/mol), Interacting residues.

Table 2: Assessment of drug-like properties of the three lead molecules as verified by QikProp.

Compound ID	MW	HBD	HBA	QPLogPo/w	QPlogS	QPlogKhsa	MDCK	Caco-2	QPLogHERG	QPLogB/B	% Oral absorption
Zinc14618306	341.251	3	5	0.885	0.885	-0.221	21	55	-4.644	-1.759	63
Zinc14638402	298.251	3	5	0.939	-2.997	-0.242	43	106	-4.760	-1.447	69
Zinc14610013	298.251	3	5	1.005	-2.955	-0.247	55	131	-4.772	-1.349	71

Molecular weight (<500 Da). Hydrogen bond donors (<5). Hydrogen bond acceptors (<10). Predicted octanol/water partition co-efficient log p (acceptable range: 22.0 to 6.5). Predicted Caco-2 cell permeability in nm/s (acceptable range: 25 are poor and 500 is great). Predicted IC_{50} value for blockage of HERG K+ channels (concern: below −5). QP log BB for brain/blood (-3.0 to 1.2). Predicted percentage of human oral absorption (25 % is poor).

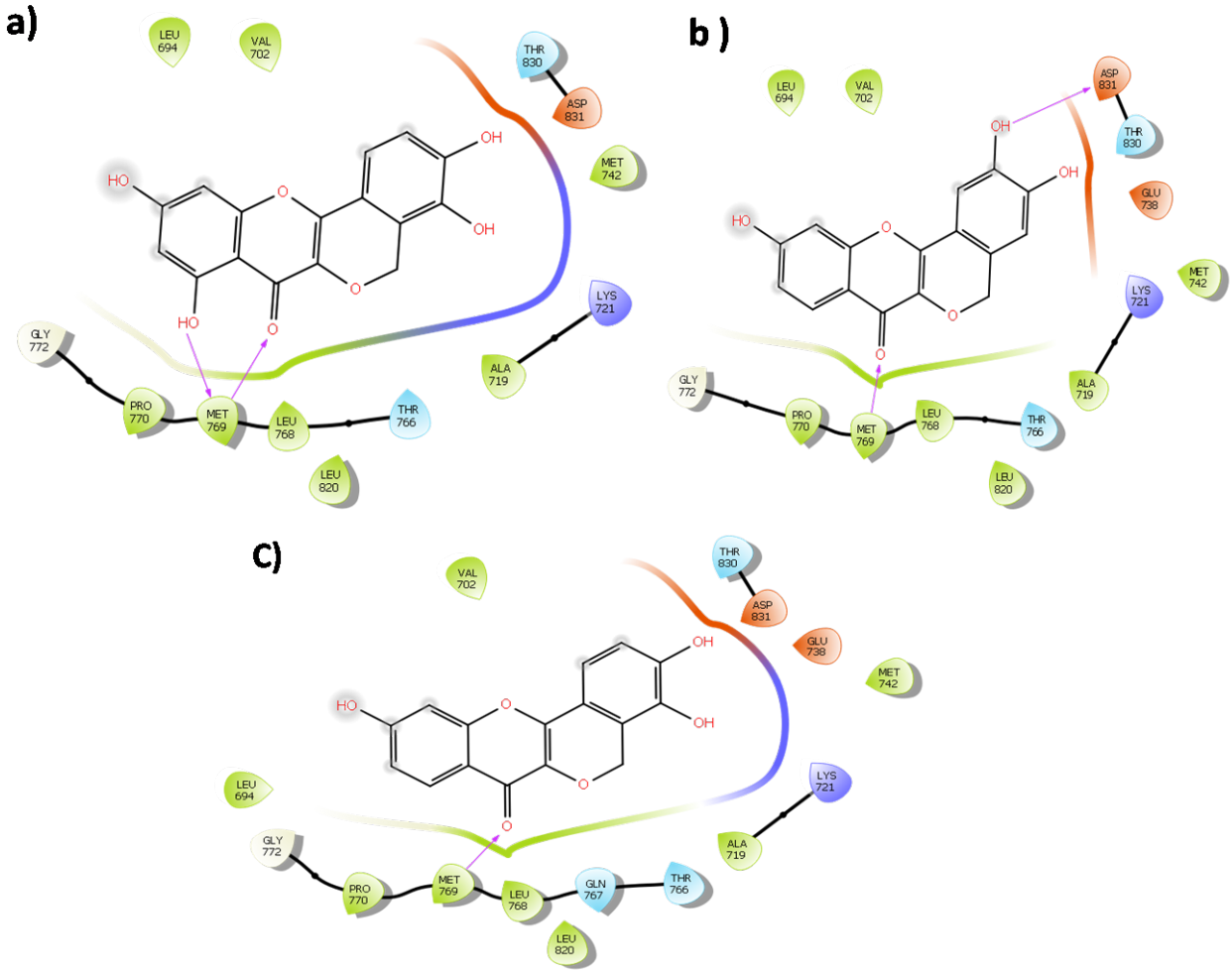

Figure 1: 2-D interactionof the three lead molecules studied. a) Zinc14618306, b) Zinc14638402, c) Zinc14610013.

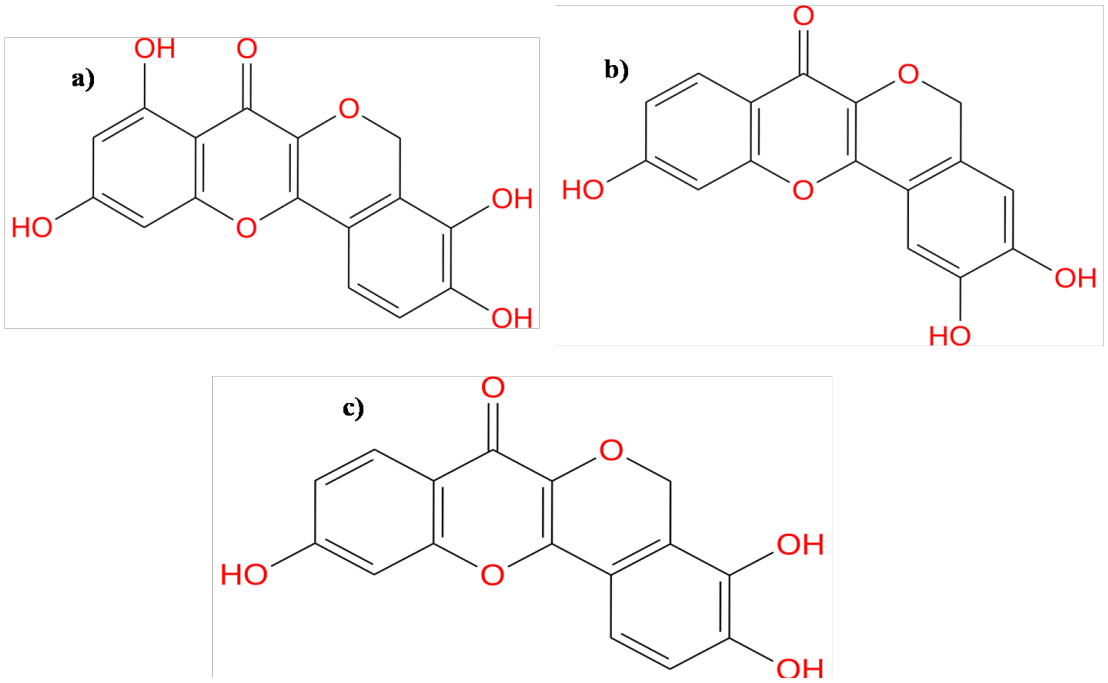

Figure 2: Chemical structures of the three lead molecules studied. a) Zinc14618306, b) Zinc14638402, c) Zinc14610013.

ADME properties Analysis

The pharmacokinetic properties of the three newly discovered lead compounds were assessed by the use of QikProp. The above three lead molecules are proved with drug-like properties based on Lipinski's rule of five. The molecular weight of the lead molecules are less than 500kDa, number of hydrogen bond donors is >5 and hydrogen bond acceptor are>10 and the predicted octanol/water partition coefficient (QPlogPo/w) > 5. All these properties fit well within the acceptable range of the Lipinski rule for drug-like molecules. The lead compounds were further evaluated for their drug-like behavior through analysis of pharmacokinetic parameters required for absorption, distribution, metabolism, excretion and toxicity (ADMET) by use of QikProp. For the three lead molecules, the aqueous solubility (QPlogS) critical for estimation of absorption and distribution of drug within the body range between -2.997 to -2.955 respectively. Cellular permeability (QPPCaco2) responsible for drug metabolism and its access to the biological membrane is within the acceptable range 55 to 131. Overall, the percentage of human oral absorption for the compounds ranged from 63% to 71%. The predicted IC_{50} value for the blockage of HERG K+ channels (QPlogHERG) are in the acceptable range of below -5. The predicted value of binding to human serum albumin (QPksha) fit well in the acceptable range ~-0.221 to ~-0.247. The cell permeability (QPPMDCK) of the molecules ranged from ~21 to ~55. The predicted Brain/Blood barriers are under acceptable range ~ -1.759 to 1.349. All the pharmacokinetics parameters are fit well with the acceptable range defined for use of human. The results retrieved are also listed in the Table 2.

Molecular dynamics simulation

Totally, 100ns MD simulation analysis was carried out to examine the dynamics nature and real time conformational stability of the identified three best active molecules by virtual screening within the binding cavity of the therapeutic target EGFR. The simulation event analysis panel was applied to retrieve the detailed knowledge of protein-ligand complexes to understand the conformational stability of protein-ligand complexes by analyzing RMSD, RMSF and protein-ligand contacts analysis. The RMSD of the target protein in the presence of active molecules Zinc14618306 was evaluated and reported in Figure 3a, we noted the RMSD in the ranges 1.5 to 2.8 Å and the RMSD was observed 3.3 Å. Likewise, we monitored the ligand RMSD to examine the ligand stability for the 100ns simulation period, in the ligand RMSD analysis, we observed, there was no fluctuation was noted for the first 55ns, but exponential deviation was observed after 55ns to end of the simulation and the RMSD was noted 4.6 Å, but it this deviation do not cause any impact on the ligand stability, because the ligand RMSD was found in appropriate range 4.6Å, these RMSD analysis emphasis, the there was no conformational changes after ligand binding during the course of the 100ns simulation. For the active molecules Zinc14638402, the RMSD was noted within the binding pocket of the

EGFR. Initially, the protein RMSD was analyzed and it found in the ranges 1.6Å to 2.4Å, in this protein RMSD analysis, we observed no deviation for first 58ns, by gradual deviation was showed after 58ns to end of the simulation period and the RMSD was observed 3.4 Å, but these gradual changes in protein RMSD cause no major impact on the protein stability. we noted the RMSD of the ligand and it found in the ranges 0.4 Å to 2.2 Å. Ligand showed slight deviation for first 58ns, after that, the ligand maintain their steady binding stability for the remaining simulation period within the binding pocket of the target protein EGFR and the RMSD was noted at the end of the simulation 2.3Å and reported in figure 3b. The RMSD of the active molecules Zinc14610013-EGFR was analyzed and reported in figure 3c. Initially, we examine the protein RMSD, there were no conformational changes in protein RMSD for the first 30 ns, but we noted major deviation for the remaining simulation and the RMSD reached 4.3 Å. In the ligand RMSD analysis, the RMSD showed in the ranges 1.1 Å to 1.3 Å. We noted a slight fluctuation at 5ns; suddenly, it retained its binding stability for the remaining simulation. The analysis of protein-ligand RMSD analysis clearly indicated the best three active molecules having consistent binding stability within the active site of the target protein EGFR throughout the 100ns simulation.The analysis of flexibility and motility of each and every backbone residues target protein EGFR by using RMSF values and it depicted in figure 4a-c. Greater RMSF value indicates more flexibility, while lower RMSF value indicates strong binding affinity. The RMSF values of protein backbone were noted in the ranges 5.6Å to 1.2Å. Likewise, in the presence of the lead molecule Zinc14638402, the protein backbone residues RMSD showed in the ranges 1.6Å to 1Å. In the lead molecules Zinc14610013, the protein backbone residues RMSD noted in the ranges 6.7 Å to1.2 Å. The protein and ligand RMSF values were observed in the appropriate ranges and the analysis strongly demonstrated, the ligand does not induce any major impact in their conformation. Analysis of atomic level interaction among the protein-ligand is a one of the vital steps in to evaluate the binding efficiency of protein –ligand during the course of the MD simulation and reported in figure 5a-c. Therefore, to ascertain the binding stability of the ligand with active site of the protein, we thoroughly analyzed various vital intermolecular interactions including hydrogen bond, hydrophobic contact, ionic contacts, and salt bridge contacts. For the molecule Zinc14618306, we observed the polar residues THR 766 and hydrophobic residue MET 769 held hydrogen bond contacts with maximum simulation occupancy 50%, 77%, ASP 776 and CYS 773 held water bridge interactions with maximum occupancies of 38% and 39%, respectively, more ever active molecules 1 held hydrophobic contact with the residue LEU 694. For the molecule Zinc14638402, the crucial binding site residues GLU 738, ASP 831, LYS 721, MET 769 and GLN 767 having hydrogen bond contacts with highest simulation occupancy 52%, 39%, 70%, 42%, 97% and 37% respectively. Interestingly, only one water bridgecontact with GLN 767. Among other interactions, the hydrogen bond contacts occupied more interactions to enhance the stability of the protein-ligand complexes. For the molecule Zinc14610013, we observed residue MET 769 held the hydrogen bond interaction with maximum simulation occupancy 98%. The key residues LEU 820, ALA 719 involved in hydrophobic interaction with active molecule Zinc14610013 to maintain their consistent binding stability.

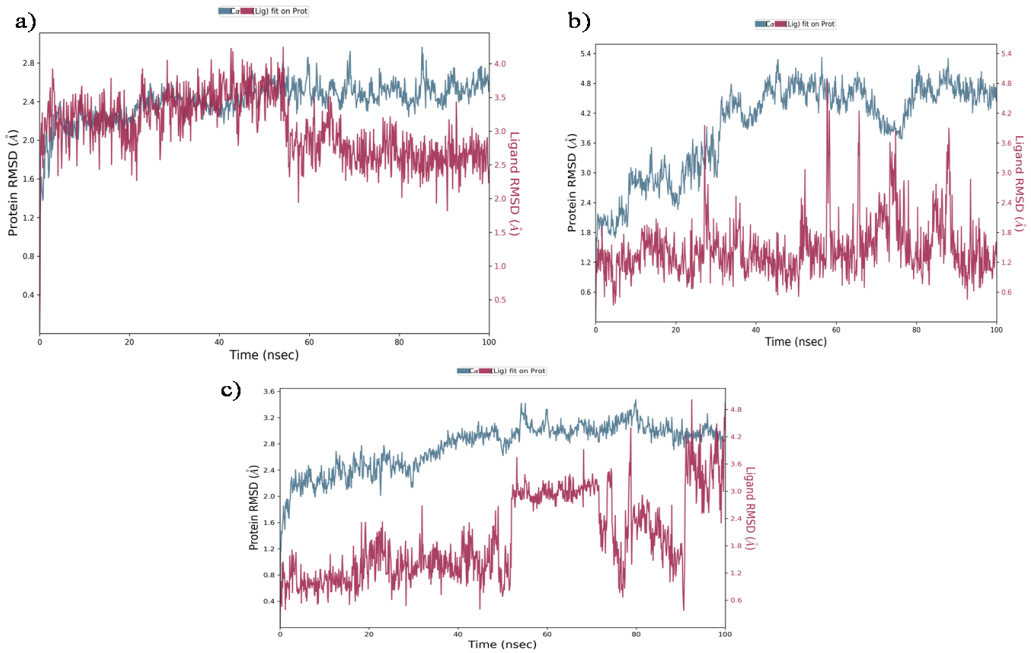

Figure 3: a) The RMSD plot of Zinc14618306–EGFR complex during 100ns simulations. b) The RMSD plot of Zinc14638402-EGFR contacts during 100ns simulations. c) The RMSD plot of Zinc14610013- EGFR complex at end of the 100 ns simulations.

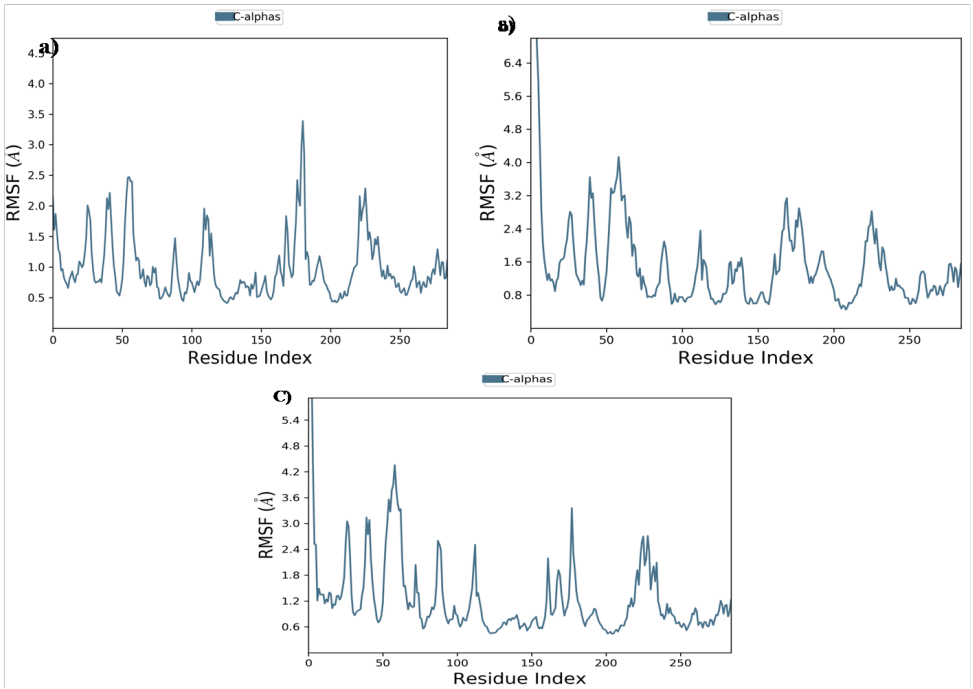

Figure 4: **a)** The RMSF analysis of Zinc14618306–EGFR complex during 100ns simulations. b) The RMSF analysis of Zinc14638402-EGFR contacts during 100ns simulations. c) The RMSF analysis of Zinc14610013-EGFR complex at end of the 100 ns simulations.

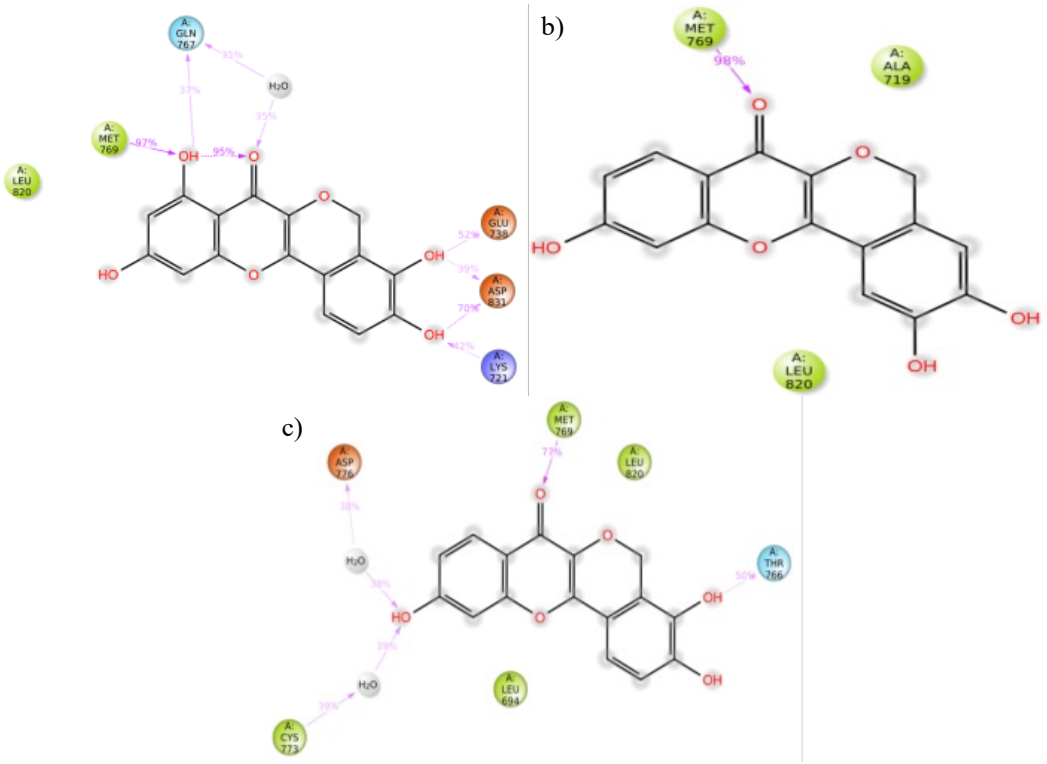

Figure 5: a) The 2Ddiagrams of Zinc14618306–EGFR complex during 100ns simulations. b) The 2Ddiagrams of Zinc14638402-EGFR contacts during 100ns simulations. c) The 2Ddiagram of Zinc14610013-EGFR complex at end of the 100 ns simulations.

Conclusion:

Targeting the inhibition of EGFR over expression using small molecules is recognized as a significant therapeutic target for oncology, especially CC. Three lead compounds ZINC14618306, ZINC14638402, and ZINC14610013 have optimal binding features with EGFR and are recommended for further investigation.

Acknowledge:

The authors thank the Management of Karpagam academy of higher education. We thank our Managing director Dr. R. Vasanthakumar, CEO, Vice-Chancellor and Registrar for providing funding support and completeness of all this work. .

Funding statement:

This research received no external funding.

Conflicts of interest: None

References:

1. http://www.cancer.org/research/ cancer-facts-statistics.
2. Li Z, Liu P, Yin A, Zhang B, Xu J, Chen Z, Zhang Z, Zhang Y, Wang S, Tang L, Kong B, Song K. Global landscape of cervical cancer incidence and mortality in 2022 and predictions to 2030: The urgent need to address inequalities in cervical cancer. Int J Cancer. 2025 Jul 15;157(2):288-297. doi: 10.1002/ijc.35369.
3. Kaufman NEM, Dhingra S, Jois SD, Vicente MDGH. Molecular Targeting of Epidermal Growth Factor Receptor (EGFR) and Vascular Endothelial Growth Factor Receptor (VEGFR). Molecules. 2021 Feb 18;26(4):1076. doi: 10.3390/molecules26041076.
4. Tito C, Masciarelli S, Colotti G, Fazi F. EGF receptor in organ development, tissue homeostasis and regeneration. J Biomed Sci. 2025 Feb 19;32(1):24. doi: 10.1186/s12929-025-01119-9.
5. Amelia T, Setiawan AN, Kartasasmita RE, Ohwada T, Tjahjono DH. Computational Prediction of Resistance Induced Alanine-Mutation in ATP Site of Epidermal Growth Factor Receptor. Int J Mol Sci. 2022 Dec 13;23(24):15828. doi: 10.3390/ijms232415828.
6. Dickerson H, Diab A, Al Musaimi O. Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors in Cancer: Current Use and Future Prospects. Int J Mol Sci. 2024 Sep 17;25(18):10008. doi: 10.3390/ijms251810008.
7. Martin-Fernandez ML, Clarke DT, Roberts SK, Zanetti-Domingues LC, Gervasio FL. Structure and Dynamics of the EGF Receptor as Revealed by Experiments and Simulations and Its Relevance to

Non-Small Cell Lung Cancer. Cells. 2019 Apr 5;8(4):316. doi: 10.3390/cells8040316.
8. Halder S, Basu S, Lall SP, Ganti AK, Batra SK, Seshacharyulu P. Targeting the EGFR signaling pathway in cancer therapy: What's new in 2023? Expert Opin Ther Targets. 2023 Apr-May;27(4-5):305-324. doi: 10.1080/14728222.2023.2218613.
9. Stamos J, Sliwkowski MX, Eigenbrot C. Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. J Biol Chem. 2002 Nov 29;277(48):46265-72. doi: 10.1074/jbc.M207135200.
10. Lenselink EB, Louvel J, Forti AF, van Veldhoven JPD, de Vries H, Mulder-Krieger T, McRobb FM, Negri A, Goose J, Abel R, van Vlijmen HWT, Wang L, Harder E, Sherman W, IJzerman AP, Beuming T. Predicting Binding Affinities for GPCR Ligands Using Free-Energy Perturbation. ACS Omega. 2016 Aug 31;1(2):293-304. doi: 10.1021/acsomega.6b00086. Epub 2016 Aug 30.
11. Anbuselvam M, Easwaran M, Meyyazhagan A, Anbuselvam J, Bhotla HK, Sivasubramanian M, Annadurai Y, Kaul T, Pappusamy M, Balasubramanian B. Structure-based virtual screening, pharmacokinetic prediction, molecular dynamics studies for the identification of novel EGFR inhibitors in breast cancer. J Biomol Struct Dyn. 2021 Aug;39(12):4462-4471. doi: 10.1080/07391102.2020.1777899.
12. Irwin JJ, Shoichet BK. ZINC--a free database of commercially available compounds for virtual screening. J Chem Inf Model. 2005 Jan-Feb;45(1):177-82. doi: 10.1021/ci049714
13. Mohan A, Krishnamoorthy S, Sabanayagam R, Schwenk G, Feng E, Ji HF, Muthusami S. Pharmacophore based virtual screening for identification of effective inhibitors to combat HPV 16 E6 driven cervical cancer. Eur J Pharmacol. 2023 Oct 15;957:175961. doi: 10.1016/j.ejphar.2023.175961.
14. Mohan A, C Ji K, Chitra B, Jeeva A, Ji HF. Molecular docking analysis of N-myristoyl-transferase with small molecules from the ZINC database for screening potential anti-malarial drugs. Bioinformation. 2025 Jun 30;21(6):1397-1403. doi: 10.6026/973206300211397.
15. Schrodinger Release–2: Qikprop, Schrodinger, LLC, New York, 2020.
16. Poochi SP, Easwaran M, Balasubramanian B, Anbuselvam M, Meyyazhagan A, Park S, Bhotla HK, Anbuselvam J, Arumugam VA, Keshavarao S, Kanniyappan GV, Pappusamy M, Kaul T. Employing bioactive compounds derived from *Ipomoea obscura* (L.) to evaluate potential inhibitor for SARS-CoV-2 main protease and ACE2 protein. Food Front. 2020 Jun;1(2):168-179. doi: 10.1002/fft2.29.
17. Mohan A, Rendine N, Mohammed MKS, Jeeva A, Ji HF, Talluri VR. Structure-based virtual screening, in silico docking, ADME properties prediction and molecular dynamics studies for the identification of potential inhibitors against SARS-CoV-2 M^{pro}. Mol Divers. 2022 Jun;26(3):1645-1661. doi: 10.1007/s11030-021-10298-0.