



Design and investigation of interactions of novel peptide conjugates of purine and pyrimidine derivatives with EGFR and its mutant T790M/L858R: an in silico and laboratory study

Hannah L. Hunt¹ · Beatriz G. Goncalves¹ · Mary A. Biggs¹ · Mia I. Rico¹ · Molly E. Murray¹ · Charlotta G. Lebedenko¹ · Ipsita A. Banerjee¹

Received: 17 July 2023 / Accepted: 17 November 2023
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2024

Abstract

Peptide-based therapeutics have been gaining attention due to their ability to actively target tumor cells. Additionally, several varieties of nucleotide derivatives have been developed to reduce cell proliferation and induce apoptosis of tumor cells. In this work, we have developed novel peptide conjugates with newly designed purine analogs and pyrimidine derivatives and explored the binding interactions with the kinase domain of wild-type EGFR and its mutant EGFR [L858R/ T790M] which are known to be over-expressed in tumor cells. The peptides explored included WNWKV (derived from sea cucumber) and LARFFS, which in previous work was predicted to bind to Domain I of EGFR. Computational studies conducted to explore binding interactions include molecular docking studies, molecular dynamics simulations and MMGBSA to investigate the binding abilities and stability of the complexes. The results indicate that conjugation enhanced binding capabilities, particularly for the WNWKV conjugates. MMGBSA analysis revealed nearly twofold higher binding toward the T790M/L858R double mutant receptor. Several conjugates were shown to have strong and stable binding with both wild-type and mutant EGFR. As a proof of concept, we synthesized pyrimidine conjugates with both peptides and determined the KD values using SPR analysis. The results corroborated with the computational analyses. Additionally, cell viability and apoptosis studies with lung cancer cells expressing the wild-type and double mutant proteins revealed that the WNWKV conjugate showed greater potency than the LARFFS conjugate, while LARFFS peptide alone showed poor binding to the kinase domain. Thus, we have designed peptide conjugates that show potential for further laboratory studies for developing therapeutics for targeting the EGFR receptor and its mutant T790M/L858R.

Keywords Tumor targeting · Peptide conjugates · Pyrimidine · Purine · Tyrosine kinase domain · Molecular dynamics

Introduction

Currently, most cancer treatment options include chemotherapy, radiation, and surgery. However, these treatments may cause complications and side effects due to invasiveness and relatively low specificity [1]. Although immunotherapy has been gaining traction, it is also associated with side effects and resistance and the effectiveness can vary depending upon genetic mutations [2]. Likewise, resistance to common chemotherapeutic drugs also may result in their

failure [3, 4]. Thus, various strategies are being developed for more targeted therapeutics [5]. One approach that has been gaining importance is peptide-based tumor targeting due to its greater specificity toward tumor cells [6]. This approach involves targeting specific receptors that are expressed by tumor cells which can bind to distinct peptides and elicit signaling pathways involved in tumor cell proliferation, migration, and invasion [7, 8]. Many of these receptors become over-expressed in cancer and therefore are ideal targets for tumor therapeutics [9]. For example, the sequence VSWFSRHRYSPPFAVS obtained from a phage display library was found to be highly specific for binding to integrin $\alpha_6\beta_1$ receptor [10]. In a separate study, it was demonstrated that the neurohormone peptide Bombesin functionalized with gold nanoparticles showed higher uptake and specificity toward gastrin-releasing receptors that are often

✉ Ipsita A. Banerjee
banerjee@fordham.edu

¹ Department of Chemistry, Fordham University, 441 East Fordham Road, Bronx, NY 10458, USA

over-expressed in breast and small-cell lung cancers [11]. In addition, small molecule-peptide-drug conjugates containing bioactive peptide sequences have been developed [12]. For example, the low-density lipoprotein receptor (LRP-1) binding peptide, Angiopep-2, was conjugated with the cell penetrating TAT peptide and the drug Paclitaxel for targeting glioma cells and was found to be highly efficacious [13].

Among the various types of receptors, growth factor receptors, G-protein coupled receptors, and integrins are considered ideal targets for therapeutics due to their role in tumor cell growth and metastasis [14]. In particular, the epidermal growth factor receptor (EGFR), which belongs to the tyrosine kinase (RTK) family of receptors, is an attractive target as it is over-expressed in several types of cancers [15–17]. In addition to the overexpression of wild-type EGFR, various mutations of EGFR have been observed [18]. These mutations often alter the dimerization and phosphorylation threshold of the receptors and thus increase their activity, leading to chemoresistance, cell proliferation, and more aggressive tumors [19]. For example, the expression of the exon 19 deletion (19-del) and exon 21 deletion (21-del) L858R point mutation has been found to be highly prevalent in non-small cell-lung cancers [20]. Other known common mutations include T790M and C797S [21, 22]. More recently, some uncommon mutations of EGFR have also been shown to occur, and these include G719C, G719A, L861Q, and S768I [23]. Additionally, the L858R and T790M mutations often coexist in tumor cells and their combination can cause increased drug resistance toward multiple EGFR inhibitors [24]. A majority of the EGFR mutations occur in the kinase domain. The T790M mutation is called the “gatekeeper” mutation due to its position within the ATP binding cleft [25, 26], while the L858R lies in the activation loop (A-loop) region [27]. In addition, the conserved DFG motif plays an important role as it is part of the A-loop that forms a cleft that binds to the substrate and controls access of ligands and drugs to the active site of the kinase domain [28]. Thus, the DFG motif is also a popular target for drug design [29]. On the basis of molecular modeling studies, the mono-anilino pyrimidine-based drug AZD9291 was designed for targeting the EGFR T790M mutant kinase domain [30]. Since then, several modifications of the structure have been attempted, and it was found that an indole ring derivative with 5,6-dihydro-4H-pyrrolo-[3,2,1-ij] quinolone motif inhibited the activity of the EGFR T790M/L858R mutant with higher potency and lower cardiotoxicity. Molecular docking studies demonstrated that an amino group can act as an H-bond donor with M793 in the EGFR T790M, while the indole ring aided in hydrophobic interactions with the hinge region of the kinase domain [31]. Several 2-anilino-pyrimidine compounds have been synthesized as analogs of the drug WZ4002 [32, 33] to enhance interactions with T790M and C797S EGFR

mutants by incorporating different hydrophilic or hydrophobic moieties. Interestingly, a pyrrolo[3,2-d] pyrimidine drug has been shown to specifically inhibit the kinase activity of both wild-type EGFR and EGFR [L858R/T790M] [34]. Potent RTK kinase inhibitors including Gefitinib, Erlotinib, and Osimertinib have also been widely studied [35, 36]. Recently, through in silico and laboratory studies, thioeno [2,3-d] pyrimidine derivatives have also been developed as EGFR inhibitors [37].

Peptide-based therapeutics are also being studied for targeting not only the wild-type EGFR, but also its mutants. For example, in a recent study, the cyclic peptide sequences CHVPGSYLC and CVNAMQSYC, where the -N and -C terminal cysteines were connected through a disulfide bond, were conjugated with the drug camptothecin. The drug conjugates were found to specifically target EGFR expressed in H1299 cells and its mutant, EGFRvIII-expressed in DKMG cell lines [38]. Additionally, substituted purines and their bioisosteres have been designed and synthesized as potential EGFR kinase domain inhibitors [39].

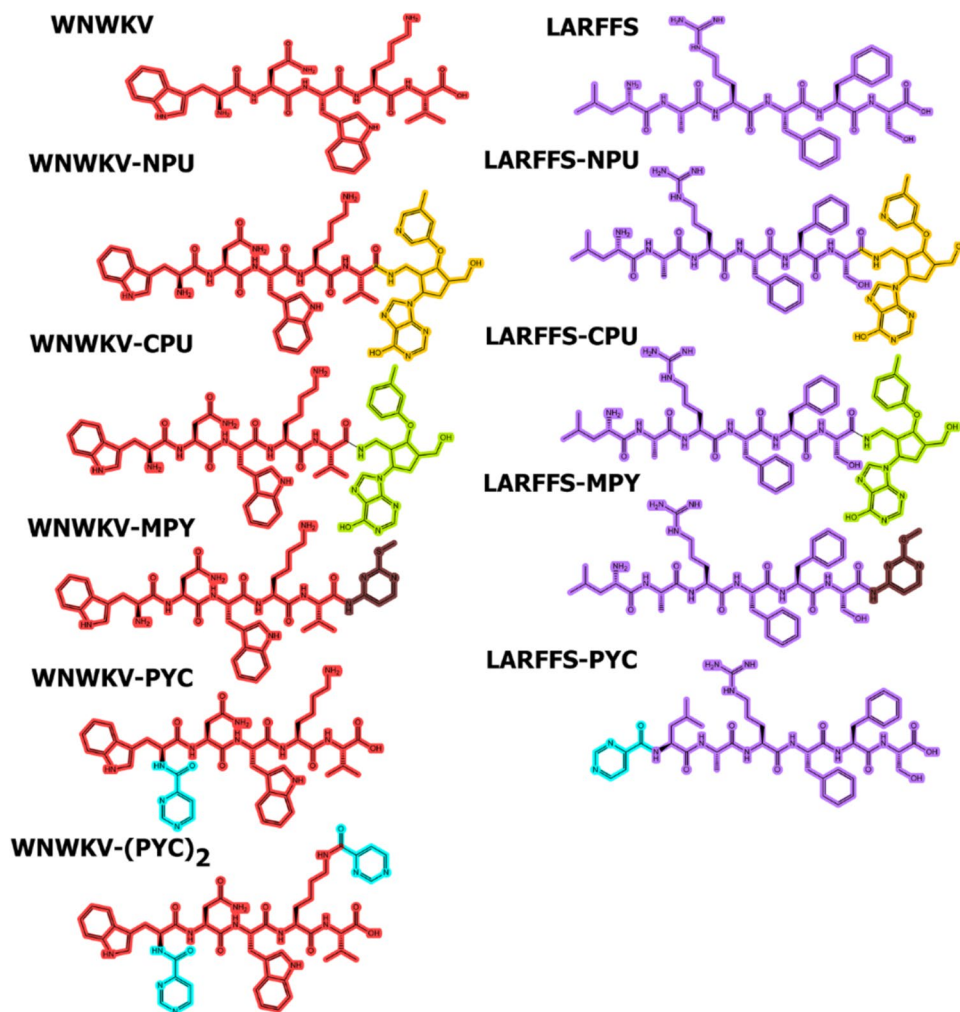
In this work, we have designed new peptide-based purine and pyrimidine derivatives by conjugating them with the peptide sequences WNWKV and LARFFS and explored their efficacy in binding to the kinase binding domain of the EGFR and its mutant T790M/L858R. It is well known that kinase activity increases due to T790M/L858R mutation and the receptor shows an increased affinity toward AMP, which leads to higher stabilization of the double mutant EGFR [40] in the active state. Therefore, developing novel therapeutics to target this double mutant EGFR kinase domain would be advantageous. Specifically, the sequence WNWKV is a bioactive peptide containing multiple indole ring systems derived from sea cucumber and has been shown to have antioxidant properties [41]. In previous work, it has been shown that several antioxidant peptides also displayed anticancer effects [42]. On the other hand, LARFFS was designed for targeting EGFR over-expressed tumor cells, and was shown through peptide library design and docking analysis to bind to domain I of the EGFR [43]. Given the hydrophobic moieties present in both peptides, we hypothesized that upon conjugating to purine or pyrimidine moieties, the conjugates may show potential for binding within the hydrophobic pocket of the EGFR kinase domain. Thus, we utilized a target hopping approach to explore the binding interactions of the peptides and their conjugates with the EGFR kinase domain. Specifically, we explored the binding interactions with the mutant T790M/L858R and the wild-type EGFR kinase domain. Target hopping approach has been found to be fruitful in identifying several drug molecules. For example, this approach was utilized to show that lithocholic acid, a physiological ligand of the nuclear FXR receptor and the TGR5 receptor, could also behave as an antagonist toward the EphA2 receptor [44].

The purine derivatives designed include 9-((3R,4R)-2-(aminomethyl)-4-(hydroxymethyl)-3-(m-tolulyloxy) cyclopent-1-en-1-yl)-9H-purin-6-ol, abbreviated as CPU, and its bioisostere NPU, where we replaced the tolulyloxy group with a 5-methyl pyridin-3-yl moiety, leaving all other structural aspects of the molecule the same. Thus, we designed 9-((3R,4R)-2-(aminomethyl)-4-(hydroxymethyl)-3-((5-methylpyridin-3-yl) oxy) cyclopent-1-en-1-yl)-9H-purin-6-ol, abbreviated as NPU. CPU was designed as a mimic of the molecule (2R,3S,5R)-5-(6-Amino-2-chloro-9H-purin-9-yl)-3-(benzoyloxy) tetrahydrofuran-2-yl methyl 2-fluorobenzoate, which is an intermediate in the preparation of clofarabine, a purine nucleoside analog with antineoplastic and antiviral properties [45, 46]. We replaced the furan ring with a cyclopentynyl ring that was then attached to a $-\text{CH}_2\text{OH}$ group at position 4, a $-\text{CH}_2-\text{NH}_2$ group at position 1, and tolulyloxy group at position five. The chloro and amino group side chains attached to the purine moiety were replaced by hydrogen and hydroxyl groups at positions 2 and 6, respectively, in an effort to enhance hydrophobic and H-bond interactions with the kinase binding domain of the

receptors. In previous work, it has been shown that a bioisostere of cabozantinib containing trimethylpyridine instead of benzene in the center of the structure demonstrated enhanced antiproliferative effects toward multiple tumor cell lines and greater specificity for hepatocellular carcinoma compared to cabozantinib [47]. We therefore created the bioisostere of CPU, namely NPU. Each of these molecules was attached to the peptide sequences LARFFS and WNWKV separately.

In a study conducted by Xiao and co-workers, pyrimidine derivatives containing 5-(methylthio) group were found to be potent against EGFR T790M/L858R mutants [48]. Thus, for designing the pyrimidine derivatives, we utilized two small molecules, 2-methylthio pyrimidin-4-amine, abbreviated as MPY, and pyrimidine-4-carboxylic acid, abbreviated as PYC. Both of these compounds were also conjugated with LARFFS and WNWKV. In the case of WNWKV, a second conjugate was also created with PYC, where the lysine moiety was also attached to a second PYC molecule. Thus, in total, nine conjugates were designed. The chemical structures of all conjugates evaluated in this study are shown in Fig. 1 We

Fig. 1 Chemical structures of designed conjugates and peptides studied. WNWKV moiety is represented in red, LARFFS moiety in purple, PYC is shown in light blue, MPY is shown in brown, CPU is represented in green; NPU is represented in yellow



also explored the binding interactions with the neat CPU, NPU, MPY, and PYC molecules as well as the neat peptides individually. The binding interactions with both EGFR T790M/L858R and EGFR wild-type were studied using molecular docking and molecular dynamics studies as well as MMGBSA analysis. Results of molecular docking studies revealed that the binding interactions with WNWKV were enhanced upon conjugation with the purine or pyrimidine derivatives with both the wild-type and the mutant receptor. However, with the exception of LARFFS-PYC conjugate, the LARFFS conjugates showed marginal changes in binding affinities compared to the peptide alone for the wild-type receptor, while lesser binding affinities were seen for the mutant.

As proof of concept, we synthesized two of the peptide conjugates, WNWKV-(PYC)₂ and LARFFS-PYC for laboratory studies and explored the binding interactions with both the wild-type and mutant EGFR receptors using SPR. Additionally, we also carried out viability studies to determine if the conjugates showed cytotoxicity toward EGFR T790M/L858R and wild-type over-expressed EGFR cells as well as FACS analysis to determine if apoptosis was involved. Overall these studies reveal that the conjugates were able to bind to both receptors, though higher binding was seen for the double mutant receptor. In the case of WNWKV, critical H-bond interactions occurred with the CYS797 residue, along with ASP 800 and PHE 795 in the hinge region, and therefore, it may be further studied for targeting the triple mutant T790M/L858R/C797S in future work. The conjugates were found to form vital H-bond interactions within the active site residues which were part of activation loop, glycine-rich loop, or the hinge region. The peptide LARFFS, however, appeared to bind further into the C-terminal lobe away from the binding cleft. Conjugation with the pyrimidine derivatives, however, changed the interactions and the LARFFS conjugates were found to interact within the binding pocket, particularly with the hinge region residues.

SPR analysis revealed that WNWKV peptide had the lowest KD value for the wild-type EGFR receptor, while WNWKV-(PYC)₂ had the lowest KD value with the double mutant. Furthermore, the conjugates were found to induce higher apoptosis and cell blebbing in T790M/L858R expressing cells compared to the wild-type EGFR expressing cells. In particular, LARFFS peptide and its conjugates did not show significant induction of apoptosis in the wild-type cells. Additionally, the purine conjugates, which were studied computationally, also showed promising results, particularly with the double mutant receptor and may be potentially developed as well for future therapeutics for targeting the EGFR mutant receptor.

Methods

Computational methods

Peptide analysis

The Anti-CP 2.0 web server [49, 50] was used to determine if the peptides utilized in this study were predicted to demonstrate anticancer activity. The web server employs support vector machine (SVM) models using amino acid composition and the SVM scores indicate anticancer characteristics. The web server also predicts the physicochemical properties of the peptide based on its amino acid composition, such as the hydrophobic/hydrophilic character of the peptide sequence.

Structure design

All nucleotide derivatives, peptides, and peptide conjugates were designed using ChemDraw (20.1.1). To prepare the conjugates, the free amine groups of the designed purine or pyrimidine analog were attached to the carboxylic group of WNWKV or LARFFS. For WNWKV with PYC, two conjugates were designed, where the lysine group was also attached to the PYC as a second conjugate through its carboxylic groups. The designed structures were then transferred to ChemDraw 3D (20.1.1), where the energy minimization was carried out, producing stable 3D conformations, and then saved as .pdb files. These files were then opened on PyMoL (2.5.2) [51] to check for structural errors and add hydrogens to the ligand.

Sigma profiles and surfaces

Sigma profiles and sigma surfaces were generated in order to investigate the physicochemical properties of the conjugates and the peptides. Each .pdb ligand file was first prepared on PyMoL by adding hydrogens prior to starting the runs using the software Turbomole, which utilizes ab initio quantum chemical calculations [52]. The resulting .cosmo files were then opened on COSMOTHERM (2020) to determine the sigma profiles of each ligand [53]. The sigma profiles provide a probability distribution of the charge density of the surface of the designed compounds, ranging from -0.03 to 0.03 $e/\text{\AA}^2$. The hydrogen-bond donor region corresponds to $\sigma < -0.0082$ $e/\text{\AA}^2$, the nonpolar region corresponds to -0.0082 $e/\text{\AA}^2 < \sigma < +0.0082$ $e/\text{\AA}^2$, and the hydrogen-bond acceptor region corresponds to $\sigma > 0.0082$ $e/\text{\AA}^2$. Sigma surfaces were also generated to visualize the surface charge densities of each ligand.

Receptor processing

The PDB files of wild-type EGFR kinase domain (PDB ID: 4JQ8) [54] and EGFR kinase domain [L858R/T790M] (PDB ID: 6S9C) [55] were downloaded from the RCSB Protein data bank [56]. Any pre-attached ligands and water molecules were removed in PyMOL (2.5.2). The structures were exported as .pdb files in preparation for docking studies.

Receptor binding pocket analysis

The .pdb files of the receptors were then uploaded to the Pocket-Cavity Search Application (POCASA) (1.1) [57] a web server which performed binding pocket analysis of each receptor. The standard parameters were used for all receptors: a probe radius of 2 Å, an SPF of 16, a PDF of 18, a grid size of 1 Å, and an unlimited number of cavities. POCASA obtained the volume for each surface cavity by inserting spheres of various radii between atoms in the receptor and then adding a rolling probe to roll along the receptor's surface. The volume of each surface cavity was filled in with markers. The POCASA results also included information about the analysis parameters and depth centers for each surface cavity. The surface cavity information obtained from POCASA was used to determine the most probable docking regions to which the ligands could bind for each receptor. To visualize the H-bond donor/ acceptor sites within the binding pockets of the receptor, we utilized Desmond in Schrodinger Suites version 2023-2 SiteMap Panel [58, 59]. To begin SiteMap calculation, the cleaned receptor, which was prepared using the protein preparation wizard, was displayed on the workspace. The minimum number of site points required for the initial site-finding stage to define a site was set to 15 site points. The size of the grid used in the displayed site maps was set to standard, corresponding to 0.70 Å. The default distance from the nearest site point at which to crop the individual site maps for display was set to 4 Å. The option to detect shallow binding sites was selected. When the job was completed, the protein was displayed on the workspace showing different map types which correspond to hydrophobic, hydrophilic, hydrogen-bond donor, hydrogen-bond acceptor, and metal-binding regions, which were color coded. SiteMap not only aided in further confirming the binding pockets, but also provided imaging data showing the areas of the residues involved.

Molecular docking studies

Two separate programs were utilized to conduct docking studies, namely Autodock Vina v.1.2.0 and DockThor, thereby allowing further validation of the results. For Autodock Vina, first the clean .pdb file of the receptor was uploaded to Autodock Tools (1.5.6.) [60], the water

molecules were removed, and polar hydrogens and Kolman charges were added before generating a .pdbqt file. On a separate workspace, the .pdb file of the ligand was also uploaded to Autodock Tools (1.5.6.) and converted to a .pdbqt file. Then, both .pdbqt files were added to the same workspace on Autodock Tools (1.5.6.) where a grid box was generated around the receptor. The dimensions of the docking grids were set based on the location of the binding pockets determined by POCASA, with wild-type EGFR having grid dimensions of (92 Å × 110 Å × 100 Å) and EGFR [L858R/T790M] having grid dimensions of (98 Å × 126 Å × 50 Å). The coordinates of the grid box for wild-type EGFR were (141.183, 120.587, 200.952) and those of the grid box for EGFR [L858R, T790M] were (− 24.959, − 57.578, 2.169). Parameters such as exhaustiveness and energy range were kept at their default values of 8 and 4, respectively. Each protein–ligand complex was then inputted on Autodock Vina v. 1.2.0 [61], which gave a table of optimal binding affinities and an output.pdbqt file of the ideal binding configuration for each complex based on RMSD values. The .pdbqt file was opened in Pymol (2.5.2) along with the receptor to visualize the binding conformation. DockThor utilizes a hybrid scoring function developed using a mixture of force field-based, contact-based, and knowledge-based descriptors, such as DockTScore from the DockThor program provided through the webserver [62]. For docking studies, using the DockThor webserver, the cleaned receptors and ligands were uploaded to the webserver. The grid center parameter for the wild-type EGFR was (148.081 × 123.146 × 201.531 Å) and that of the T790M/L858R double mutant was (− 24.93 × − 58.91 × − 4.80 Å). The corresponding grid size for the wild-type receptor complex was (36 × 36 × 40 Å), and the T790M/L858R receptor was (40 × 40 × 40 Å). The results obtained for the highest ranked models were downloaded and visualized on PyMOL and the corresponding binding affinities were recorded.

Binding interactions using protein–ligand interaction profiler (PLIP)

The binding interactions occurring within each protein–ligand complex were determined using the Protein–Ligand Interaction Profiler (PLIP) online interface [63]. The output .pdbqt file generated from Autodock Vina v. 1.2.0 and the .pdb file of the receptor were opened in PyMOL (2.5.2) and exported as a single .pdb file to be uploaded to the PLIP web server. The web server generated results in .txt format and in .pdb format, both of which illustrated the residues and distances involved in forming non-covalent interactions with the ligand. The results were tabulated and heatmaps were created to illustrate the most common residues involved in binding across the various ligands.

Molecular dynamics studies

Molecular dynamics studies were performed on the receptor–ligand complexes using Desmond through Maestro software 2023-2 [64] from Schrodinger. For each protein–ligand complex, the output .pdbqt file generated from Autodock Vina v. 1.2.0 and the .pdb file of the clean receptor were opened in PyMoL (2.5.2) and exported as a single .mae file. In the case of the apo receptors, the .pdb file of the receptor without the ligand was opened in PyMOL and exported as a single .mae file. The .mae file was then opened in Maestro 2023-2 to prepare them for molecular dynamic studies. The structure was checked for any errors, including missing aromatic rings, and edited accordingly. In the Protein Preparation Wizard application in Maestro, hydrogens and missing side chains were added to the receptors and disulfide bonds were created. Hydrogen bonds were optimized with a pH setting of 7.0 and restrained minimization was carried out. Heavy atoms were converged to 0.30 Å RMSD. Using the System Builder application in Maestro, a 10 Å × 10 Å × 10 Å grid box was formed around the entire protein–ligand complex with the SPC solvent model and OPLS4 force field. This particular force field offers improved parameters for proteins leading to enhanced structural stabilization during MD simulations, improved binding predictions, conformational analysis and binding free energies resulting in an enhanced model accuracy system [65]. To mimic physiological conditions, the complex was surrounded by water and the system was neutralized with sodium and chloride ions depending on the charge of the complex. In the Molecular Dynamics panel, the run time was set for 250 ns and 1000 frames were set to be produced for the trajectory images. An NPT ensemble class was selected to equilibrate the system at 310 K and 1.01325 bar followed by relaxation of the system through a series of minimizations prior to starting the run. Once simulations were completed through Desmond, the resulting out.cms files were analyzed in Maestro 2023-2 using the Simulation Interactions application. This provides data on the stability of the protein–ligand complexes and more information about the nature of the binding interactions. Trajectory images at various time points in molecular dynamics simulation were analyzed to visualize the binding interactions over time.

MMGBSA studies

Molecular mechanics generalized Born surface area (MMGBSA) energy calculations for each ligand–receptor complex to determine the theoretical free binding energies were carried out. For these calculations, trajectory files generated from each molecular dynamics simulations were analyzed using the script thermal_mmbgsa.py. and the average free energy (as well as the average electrostatic energy, average

H-bond energy, average lipophilic energy, solvation, and average van der Waals energy) for the entire 250 ns run was determined using the Prime module of Schrodinger Suite 2023-02 [66, 67]. The major components contributing to the total energy across the 250 ns simulation for three separate simulations were averaged, tabulated, and reported. This includes coulomb energy, covalent, Van der Waals, lipophilic, solv GB (generalized born electrostatic solvation energy) H-bond, and packing (pi–pi interaction) [68].

Pharmacokinetics predictions

Absorption, distribution, metabolism, and excretion (ADME) studies were performed on all ligands to predict their pharmacological properties. For this analysis, we used ADMETlab 2.0 [69] which utilizes quantitative structure–property relationship (QSPR) models employed by a robust multi-task graph attention (MGA) framework to construct accurate prediction models. Properties predicted included hERG blocker, MDCK cell permeability, ability to act as a Pgp substrate or inhibitor, and reactivity with cytochrome P450 enzymes, which are important pharmacokinetic factors to consider when designing a drug for use as a chemotherapeutic.

Laboratory methods

Materials

The peptide sequences, WNWKV and LARFFS, were custom ordered from Genscript. N-hydroxysuccinimide (NHS), 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDAC), 11-mercaptopundecanoic acid, ethanol (95%), and Pyrimidine-4-Carboxylic acid were purchased from Sigma Aldrich. Dulbecco's Modified Eagle's Medium (DMEM) was purchased from VWR. Fetal bovine serum (FBS), Phosphate Buffered Saline (PBS), Antibiotic–Antimycotic, and Penicillin–Streptomycin–Amphotericin B Solution as well as Human Lung fibroblast cells, human NCI-H1975 lung carcinoma cells (CRL-5908), and Human epithelial lung carcinoma cells (A549) were purchased from ATCC (Manassas, VA, USA). EGFR (kinase domain) and EGFR (T790M/L858R) proteins were purchased from Sino Biological or Thermo Fisher Scientific. FACS buffer was acquired from BD Bioscience and Annexin V-FITC/PI Apoptosis Kits were purchased from Elabscience Biotechnology Inc. (Houston, TX, US). WST-1 assay reagents were bought from Cayman Chemical. Dasatinib was purchased from Selleck Chemicals. Gold biosensor chips (SF-10 glass, index = 1.72) were ordered from Platypus Technologies and index fluid was purchased from Cargille.

Synthesis

We synthesized the peptide conjugates WNWKV-(PYC)₂ and LARFFS-PYC using standard peptide coupling methods [70]. For the LARFFS conjugates, Pyrimidine-4-Carboxylic acid (PYC) (0.0322 M) was dissolved in dimethyl formamide (DMF). Once dissolved, *N*-Hydroxysuccinimide (NHS) (0.05 M) and 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDAC) (0.05 M) were added to the solution, which was shaken at 200 rpm for 1 h at 4 °C to activate the carboxyl group. Then, LARFFS (0.03 M) was added to the reaction mixture and shaken at 200 rpm for 48 h at 4 °C. After the incubation period, rotary evaporation was used to remove the solvent. The product was recrystallized from acetone and water, and then dried using SpeedVac Vacuum Concentrator. For the WNWKV conjugate, a similar method was used, except that the molar concentration of PYC utilized was 0.062 M to ensure conjugation with both amino groups of the WNWKV peptide. The formation of the products was confirmed by ¹H NMR spectroscopy using a Bruker 400 MHz NMR spectrometer. Samples were prepared in the solvent DMSO-d₆ with 0.03% TMS. The ¹H NMR spectrum showed the following peaks for the LARFFS-PYC conjugate. δ 9.4 (1H, d); δ 9.3 (1H, s); δ 9.1 (1H, s); δ 8.3 (1H, s); δ 8.1 (5H, s); δ 7.9 (1H, s); δ 7.2 (8H, d); δ 7.0 (2H, s); δ 6.2 (2H, s); δ 5.1 (1H, s); δ 4.8 (2H, t); δ 4.5 (1H, q); δ 4.3 (1H, t); δ 4.1 (2H, d); δ 3.9 (1H, t); δ 3.4 (2H, t); δ 3.2 (4H, d); δ 1.8 (4H, m); δ 1.6 (m, 2H); δ 1.5 (m, 1H); δ 1.4 (3H, d); δ 0.8 (d, 6H).

The ¹H NMR spectrum showed the following peaks for the WNWKV-(PYC)₂ conjugate. δ 12.1 (1H, s); δ 10.5 (2H, s); δ 9.4 (2H, s); δ 9.2 (1H, s); δ 9.1 (2H, d); δ 9.0 (1H, s); δ 8.3 (2H, d); δ 8.5 (4H, s); δ 7.5 (2H, s); δ 7.4 (2H, s); δ 7.3 (2H, s); δ 7.1 (2H, d); δ 7.0 (2H, s); δ 6.8 (2H, d); δ 4.8 (2H, t); δ 4.7 (1H, t); δ 4.5 (1H, t); δ 4.3 (1H, d); δ 3.3 (d, 4H); δ 2.9 (2H, t); δ 2.6 (2H, d); δ 2.1 (1H, m); δ 1.8 (2H, q); δ 1.5 (2H, m); δ 1.3 (2H, m); δ 1.1 (6H, d).

Surface plasmon resonance studies (SPR)

Surface plasmon resonance (SPR) analysis is useful for examining live-binding of peptides or antibodies to specific receptors, proteins, or antigens [71]. SPR was used to examine the binding interactions of the conjugates and neat peptides with the kinase domain of EGFR (Sino Biological) EGFR Protein, Human, Recombinant (aa 668–1210, His & GST Tag), and EGFR (T790M, L858R) protein (Thermo Fisher Scientific). Each study was carried out in triplicate. Gold chips (Platypus technologies) were functionalized according to previously established methods [72]. Briefly, the chips were washed in a 70% ethanol solution and irradiated with UV light for 10 min. Then, 11-mercaptoundecanoic acid (1 M) was used to fully coat

and functionalize the chips. After an hour, the chips were coated with NHS (0.01 M) followed by EDAC (0.01 M). The chips were then incubated for two hours at 4 °C. Then, either the wild-type EGFR or the double mutant EGFR solution was allowed to incubate on the coated chips for 4 °C before use. Before beginning binding analysis, the functionalized chip was placed with gold side down on the flow sensor and a drop of Cargille's 7.21 index fluid was added to the opposite side of the chip before placing the prism. The system was calibrated and allowed to stabilize with 1X PBS buffer. Once the calibration run had been performed, 1X PBS was allowed to run through the system for 500 s before switching it out for the analyte solution (peptide or PYC conjugate). The binding interactions of analyte solutions ranging from 50 nM to 100 μM were analyzed at room temperature. The analyte was allowed to run through the system for 2500 s before the 1X PBS solution was then switched back in and allowed to circulate for at least 500 s. The flow rate was kept constant at 30 μL/min. The data from the SPR were then input into GraphPad Prism 8 (Graphpad Software Inc., San Diego, CA, USA) to perform a non-linear regression analysis in order to determine the K_D value of each sample. The values obtained for three separate runs for each sample were averaged and reported. Statistical Analysis was carried out using Student's *T* tests.

Cell studies

As a proof of concept, cell studies were carried out to examine cytotoxic effects of the pyrimidine-4-amide conjugates of both LARFFS and WNWKV and the neat peptides. For consistency, cell lines from the same organ (human lungs) were used to characterize the effects of the constructs and evaluate specificity. A549 cells (lung carcinoma epithelial cells which are known to overexpress wild-type EGFR) [73], and NCI-H1975 lung cells (ATCC CRL-5908) expressing EGFR [T790M/ L858R] were tested. In addition, primary lung fibroblasts (HLF) (ATCC PCS 201-013) which are non-cancer cells were also compared. All cancer cell types were grown to confluence in DMEM, which was supplemented with 10% FBS, 2 mL of 1X Penicillin–Streptomycin–Amphotericin B Solution (0.004%), and 20 μL of 100X antibiotic–antimycotic mixture. In the case of the lung fibroblasts, those cells were grown in fibroblast basal medium (ATCC PCS 201-030), which was supplemented with fibroblast growth kit (PCS-201-041). All cell types were grown as monolayers in a humidified incubator set to 37 °C and 5% CO₂ and monitored on a daily basis. Media were changed every 2–3 days and cells were split once confluence was reached, which ranged in time frame from 2 days to a week.

Cell viability and morphology studies

To examine cell viability, colorimetric tetrazolium-based WST-1 assays [74] were conducted after treatment with PYC conjugates or peptides. After removing the media and washing with PBS buffer, the adherent cells from the culture flasks were first detached using trypsin for three minutes at 37 °C after which it was then neutralized with culture media. The contents of the flask were then spun at 500 gs for 5 min at which point the pellet was obtained and the media were removed and replaced with fresh media and mixed with the media. The cells were then plated at a density of 1×10^5 cells/well on a 96-well Falcon polystyrene plate and incubated for 3 h at 37 °C in a 5% CO₂ to allow for attachment to the well plates prior to adding the constructs. For the assays, six different concentrations of the constructs were tested (50 nM; 100 nM; 1 µM; 5 µM; 50 µM; and 100 µM) to examine the cytotoxic effects of the constructs on the cells. Each construct was prepared in an aqueous solution containing 2% sterile DMSO. Each sample was added to the wells and allowed to incubate with the cells for 24 h. DMSO-water (containing 2% DMSO) was added to the control cells in place of the constructs. Dasatinib (5 µM) was utilized as a positive control. Prior to use, the WST-1 developer reagent and the electron mediator solution were thawed and equal volumes were mixed to prepare the WST-1 mixture. After 24 h, 10 µL of WST-1 reagent was added to each well, followed by an incubation period of 3 h at 37 °C. Then, the contents within the wells of the plates were gently mixed and the absorbance at 450 nm was read using a BioTek Eon microplate reader. The mean absorbance of the wells only containing media was subtracted from all samples. The percentage cell viability was then determined by subtracting the absorbance of treated cells from untreated cells and then divided by the absorbance of untreated cells. The results obtained were then multiplied by 100. All studies were carried out in triplicate and the results reported were average values of three independent experiments. Cells were imaged using an inverted Amscope IN480TC-20MB13 microscope at various magnifications (10X, 20X, and 40X) before and after treatment with constructs to examine the morphology of the cells after exposure to the constructs. IC-50 values were obtained by inputting the cell viability data obtained into Graphpad 9.5 software and by utilizing dose–response calculations. Statistical analysis was carried out using Student's *T* tests and *p* values * < 0.05 and ** < 0.01 were considered statistically significant.

Apoptosis assay

Apoptosis assays were conducted in order to examine if the constructs induced apoptosis. For each cell line, cells

were plated at a density of 1×10^5 cells/well in a 6-well Falcon polystyrene tissue culture plate and allowed to adhere for 24 h at 37 °C in a 5% CO₂ incubator. Then constructs (5 µM) were added to each well and incubated for 24 h before performing an Annexin V FITC-Propidium Iodide Assay [75]. This concentration was chosen based on the results obtained from viability studies. After 24 h, the cells were removed from the culture well plates using trypsin and were centrifuged at 400 gs for five minutes. The supernatant was then decanted and the pellet was washed with 1X Binding Buffer, followed by another round of centrifugation. The pellet was then resuspended with 50 µL of Annexin V FITC/Propidium Iodide Staining Solution and incubated in the dark at room temperature for 10 min. Then, 150 µL of 1X Binding Buffer was added along with FACS buffer (50 mL of 1 × PBS, 1% bovine serum albumin, and 0.05% sodium azide) which were then filtered through filter caps into FACS test tubes to be loaded into the instrument for Flow Cytometry [76–78]. A BD FACS Melody flow cytometer was used. Each sample was read at excitation of 488 nm and emission of 525 nm for the Annexin V FITC staining solution and 655–730 emission for the propidium iodide staining solution. The total number of events was kept at 10,000 for every sample. Results were then analyzed using the software FlowJo v10.9. Each scatter plot was gated on the FSC-A/SSC-A control pseudo color sample plot, followed by the FSC-A/FSC-W and SSC-A/SSC-W plots. The population was then plotted as Annexin FITC-A/Propidium Iodide-A with a biexponential scale to incorporate all events in each window of the quadrant gate. The gates were then all copied to the remaining samples to quantify the frequency of each event.

Characterization

Surface plasmon resonance

As proof of concept, the binding affinities of the compounds WNWKV-(PYC)₂, and the neat peptides were studied for both wild-type and T790M/L858R double mutant EGFR using a GWC Horizon SPRImager II instrument.

FACS

Fluorescence-Activated Cell Sorting (FACS) Flow Cytometry was conducted using a BD FACSMelody flow cytometer. The results of the flow cytometry experiments were analyzed using FlowJo v10.9 software.

Results and discussion

Peptide analysis

We utilized the antiCP web server 2.0 to predict the anticancer potential of the peptides WNWKV and LARFFS. The WNWKV sequence displayed an SVM (support vector machine) score of 0.63 (63%) and LARFFS showed a SVM score of 0.79 (79%). Both peptides were predicted to be anticancer peptides. Additionally, the hydropathicity score for LARFFS was found to be 0.98 with a hydrophilicity of -0.67 , while the hydropathicity score for WNWKV was -1.00 , and the hydrophilicity score was -1.02 . Furthermore, the amphipathicity score was higher for WNWKV (0.73) compared to LARFFS (0.43). This study provided confirmation that the peptides had anticancer properties.

Sigma profiles

To determine the physicochemical properties of the peptides and their conjugates, we utilized COSMO-RS. Conductor-like Screening Model with Real Solvents (COSMO-RS) utilizes a statistical thermodynamics methodology to determine the solubility properties, chemical potential, and activity coefficients of molecules in a mixture based on quantum chemical calculations [79].

Furthermore, COSMO-RS has been used to generate molecular surface charge distributions based on the 3D distribution of charges on the surface of molecules in the form of sigma profile plots. The sigma profiles provide comprehensive information about the polarity distribution of molecules and therefore provide vital information about the H-bond donor and acceptor capabilities of the molecules [80]. The sigma surfaces and sigma profiles for each group of peptide conjugates and the unconjugated molecules are shown in Fig. 2. Results for WNWKV and its conjugates are shown in Fig. 2a which indicated that the neat peptide had a hydrophobic character due to the presence of tryptophan and valine. The carbonyl groups (from amide linkages) contribute to the H-bond acceptor region peak at $(0.015 \text{ e}/\text{\AA}^2)$ [81]. Notable peaks in the hydrophobic region of the sigma profile of WNWKV were seen at $-0.003 \text{ e}/\text{\AA}^2$ and $0.008 \text{ e}/\text{\AA}^2$.

The sigma profiles of WNWKV-CPU and WNWKV-NPU, while relatively similar to each other, indicated greater hydrophobic character than the other WNWKV conjugates. Since both CPU and NPU contain purine moieties, as well as a cyclopentynyl ring with an additional toluloxo (CPU) or pyridyl moiety (NPU), higher hydrophobicity is displayed for those conjugates. It is well known that purine moieties can alternate between H-bond

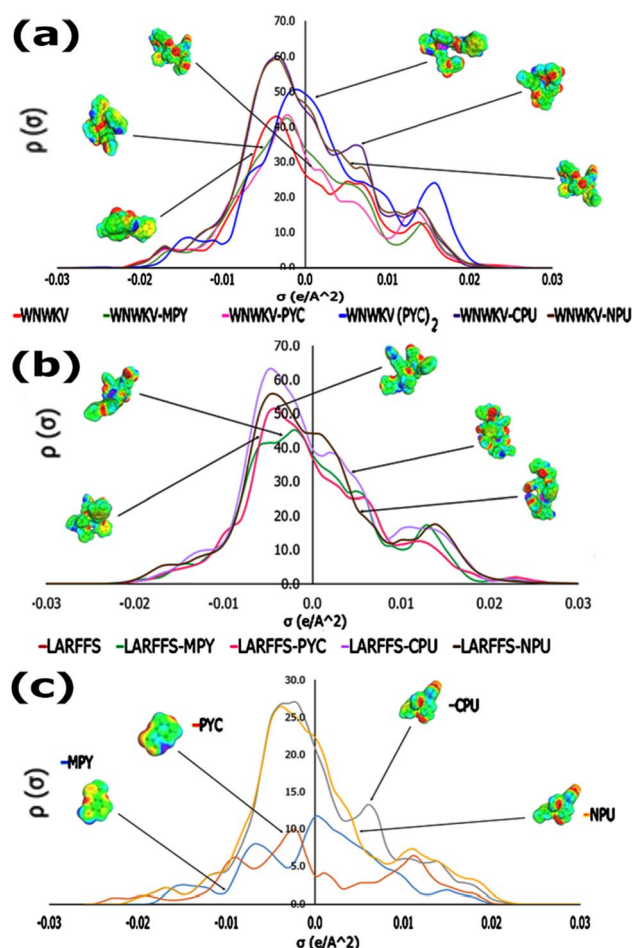


Fig. 2 Comparison of sigma profiles and sigma surfaces of **a** WNWKV and its conjugates with NPU, CPU, MPY and PYC; **b** LARFFS and its conjugates with NPU, CPU, MPY and PYC; **c** Unconjugated compounds NPU, CPU, MPY, and PYC

donor and H-bond acceptor capabilities [82]. Thus, the relatively strong peak displayed at $0.008 \text{ e}/\text{\AA}^2$ in the hydrogen-bond accepting region is attributed to the NPU and CPU moieties due to the additional H-bond accepting capabilities of the purine ring systems. The results for WNWKV-MPY and WNWKV-PYC were relatively similar to that of WNWKV alone, though WNWKV-(PYC)₂, where both the free amino groups of WNWKV are conjugated with PYC, was found to be more hydrophobic due to the presence of two pyrimidine groups. WNWKV-(PYC)₂ displayed a strong peak in the hydrogen acceptor region at $0.018 \text{ e}/\text{\AA}^2$ due to the two pyrimidine groups attached to the WNWKV. In addition, all conjugates also displayed a short peak in the H-bond donating region due to the presence of $-\text{NH}$ from amide groups of the peptide, the contribution from the side chain amino group of lysine, and the free carboxyl groups. The sigma surfaces allow for visualization of the charge densities of a molecule's

surface. The red regions of the sigma surfaces indicate the location of hydrogen-bond acceptor surfaces, the blue regions indicate the location of hydrogen donating surfaces, and the green and yellow regions indicate the location of nonpolar and neutral regions. Thus, higher green and red regions are seen for the WNWKV-CPU and WNWKV-NPU conjugates, which corroborated with the sigma profiles, while higher blue region was seen for the WNWKV-MPY conjugate.

The results of LARFFS and its conjugates are shown in Fig. 2b. A similar trend as that of the WNWKV conjugates was seen, where the LARFFS-CPU and the LARFFS-NPU conjugates showed the highest peaks in the hydrophobic region. The presence of the guanidino group from arginine residue also contributes to the peak within the hydrogen accepting region as well as higher hydrophobicity [83]. LARFFS-MPY showed slightly less hydrophobic character than LARFFS alone. In fact, LARFFS-MPY displayed a split peak in the hydrophobic region with one of the peaks leaning toward H-bond donor region due to the presence of the thio-methyl group being attached to the pyrimidine ring [84]. Interestingly, LARFFS-PYC had virtually the same sigma profile as LARFFS alone. Thus, while conjugation with MPY generated notable differences in the sigma profiles, conjugation with PYC had relatively less effect on the physicochemical properties.

We also compared the properties of the unconjugated pyrimidine and purine derivatives. Overall, the hydrophobicity was markedly reduced in all cases compared to the peptide conjugates due to the lack of the peptide moieties. As shown in Fig. 2c, CPU and NPU showed greater hydrophobic character than either MPY or PYC, both of which are significantly smaller structures. Both CPU and NPU shared a similar peak in the hydrophobic region and similar sigma surfaces. For CPU, however, the hydrogen acceptor peak was shifted more toward the left closer to the hydrophobic region. This is expected as CPU has a nonpolar toluene ring system compared to the more polar pyridyl ring system seen in NPU. Both MPY and PYC showed comparatively short peaks in the hydrophobic region due to smaller size. However, PYC also showed small peaks in both the hydrogen donating and hydrogen accepting region, as would be expected due to the free carboxylic group and the pyrimidine ring system. Thus, overall, the conjugates showed higher H-bond donor/acceptor capabilities as well as higher hydrophobicity due to incorporation of the peptides and therefore may potentially lead to greater interactions with the receptors.

Receptor analysis

The potential locations and volumes of the binding pockets of the kinase domain of the receptors were first determined

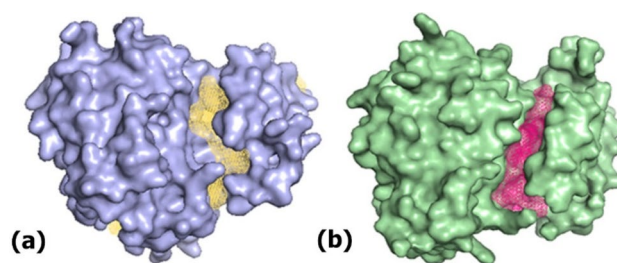


Fig. 3 Comparison of binding pockets obtained from POCASA. **a** Wild Type EGFR; **b** Double Mutant T790M/L858R EGFR

Table 1 Binding Pocket Analysis from POCASA

Rank number	Pocket number	Volume (Å ³)	VD value
Wild type EGFR			
1	190	426	1345
2	122	183	502
3	57	85	215
4	374	93	201
5	12	37	85
[L858R/ T790M] EGFR			
1	261	363	1025
2	37	182	508
3	19	47	145
4	27	43	116
5	32	33	84

through the POCASA web server. Results are seen in Fig. 3 and Table 1. For wild-type EGFR, the pocket ranked number 1 by POCASA had a volume distribution (VD) value of 1345. The next highest ranked pocket had a VD value of 508 and the fifth highest ranked pocket had a VD value of 85. Thus, the wild-type EGFR is predicted to have one large binding pocket, which is seen in Fig. 3a. For EGFR [L858R, T790M], the binding pocket ranked number 1 by POCASA had a VD value of 1025. The pocket ranked number 2 had a VD value of 508, and the pocket ranked number 5 had a VD values of 84. As with wild-type EGFR, the number 1 ranked pocket was significantly bigger than the other binding pockets (Fig. 3b). However, the VD value of the highest ranked pocket for EGFR [L858R/ T790M] was lower than that of wild-type EGFR.

To gain further insight about the binding sites of the wild-type and the double mutant (T790M/L858R) kinase domain of the EGF receptors and to locate binding sites, we utilized SiteMap from Schrodinger Suite. Specifically, SiteMap can display regions within the binding site appropriate for occupancy by hydrophobic groups or by ligand hydrogen-bond donors, H-bond acceptors, or metal-binding functional groups. This allows for evaluation of a

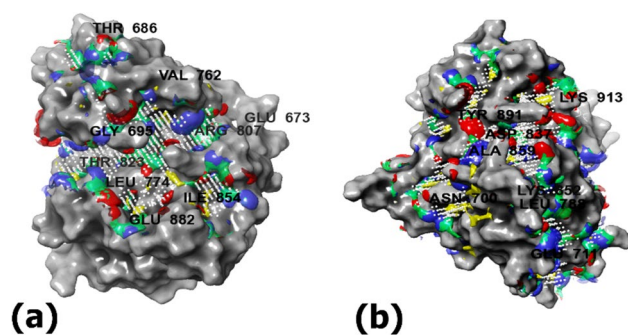


Fig. 4 Comparison of predicted binding sites within the **a** wild-type EGFR receptor; **b** Double mutant receptor (T790M/L858R). White dots are indicative of the binding pocket regions, some of the residues within the pockets are labeled. The specific regions within the binding pocket are color-coded as follows. Hydrophilic regions—green; hydrophobic—yellow; Hydrogen bond donor region—blue; H-bond acceptor region is indicated in red

potential drug molecule's complementarity with the receptor. The results obtained are shown in Fig. 4.

As can be seen, there are several H-bond donor and acceptor regions as well as hydrophobic regions which encompass the binding pocket of the receptors. These involve residues from the activation loop, Glycine rich P-loop, and the hinge region [85, 86]. In the case of the double mutant, residues in the region deeper into the C-lobe region (residue numbers higher than 900) are also seen. Overall, there appears to be relatively more hydrophobic regions in the case of the double mutant receptors, while both receptors displayed H-bond donor and H-bond acceptor regions within the receptor. These results imply that it is likely that the conjugates may interact more favorably through hydrophobic interaction with the double mutant receptor, while also interacting with the H-bond donor–acceptor groups of the receptors through complementary regions of the ligands.

Molecular docking studies

In order to examine the binding interactions of the designed conjugates, the individual pyridine and purine-based compounds, and the peptides with EGFR (wild) and T790M/L858R EGFR, we conducted molecular docking studies using two separate docking methodologies, AutoDock Vina 1.1.2 and DockThor. AutoDock Vina utilizes algorithms that search for the conformation of flexible ligands by a stochastic global optimization of the scoring function. The binding score is evaluated by an empirical scoring function, which comprises the sum of six weighted energy terms as follows:

$$\Delta G = W_{s_1} \times \text{gauss1} + W_{s_2} \times \text{gauss2} + W_R \times \text{Repulsion} \\ + W_{hp} \times \text{Hydrophobic} + W_{hb} \times \text{HB} + W_{\text{rotNrot}},$$

where the first three terms describe steric and Van der Waals interactions, while the last three terms indicate hydrophobic, H-bonds, and changes in torsional entropy upon binding [87]. On the other hand, the DockThor program is primarily utilized to examine docking of highly flexible ligands [88, 89] such as peptides. Because we are exploring relatively short peptides and small molecule-peptide conjugates, we compared the results obtained from AutoDock Vina with the DockThor program for examining protein–ligand docking. The DockThor program utilizes the MMFF94S force field, and the scoring function for the binding energy is a sum of various intermolecular interactions including Van der Waals forces and electrostatic potentials; torsional entropy, protein–ligand lipophilic interaction, and polar solvation that takes into consideration the loss of polar interactions of the charged groups following binding and also a nonpolar solvation term, that is proportional to the solvent-accessible surface [90]. A comparison of the results of binding affinities obtained from the docking studies using both methods is shown in Table 2. In general, a common trend seen in most cases for both the double mutant and wild-type receptor was that the binding affinities were higher for the WNWKV conjugates and peptides compared to the LARFFS conjugates, the exception being the LARFFS-PYC conjugate result seen using AutoDock Vina for the wild-type receptor. In addition, the binding affinities of the neat peptides LARFFS and WNWKV were comparable for the wild-type using both docking methods, while LARFFS displayed a slightly higher docking score using the DockThor program for the double mutant receptor. Overall, in most cases, the numerical values of the binding affinities obtained were slightly higher for the DockThor program compared to AutoDock Vina, which is expected given that the methodologies utilized for scoring function are different for the two software as described earlier. Overall, the highest binding scores with the double mutant and the wild-type EGFR kinase domain were obtained for the purine derivatives when conjugated to WNWKV. Among the pyrimidine derivatives, higher binding affinities with comparable scores for both receptors were seen for WNWKV(PYC)₂. The LARFFS-PYC conjugate, however, showed a relatively higher binding toward the wild-type receptor compared to the double mutant. As expected, the unconjugated PYC and MPY molecules displayed the lowest binding, affinities, indicating that conjugating with the peptides enhanced binding affinities toward both receptors.

Binding interaction analysis

To further explicate these results, we conducted PLIP analysis of the docked conjugates and peptides with both receptors. Results of the interactions are shown in Tables 3

Table 2 Binding affinities of peptides, peptide conjugates, and pyrimidine/purine derivatives with EGFR and EGFR [L858R/ T790M] (Kcal/mol) obtained from Molecular Docking

Compounds	AutoDock Vina	DockThor	AutoDock Vina	DockThor
	EGFR	EGFR	EGFR [L858R/T790M]	EGFR [L858R/T790M]
WNWKV	− 6.0	− 8.7	− 7.0	− 8.2
WNWKV-CPU	− 10.3	− 9.5	− 8.8	− 8.7
WNWKV-NPU	− 8.9	− 9.7	− 9.0	− 10.0
WNWKV-MPY	− 7.3	− 8.1	− 7.5	− 7.9
WNWKV-PYC	− 6.7	− 8.5	− 8.9	− 8.1
WNWKV-(PYC) ₂	− 8.0	− 8.2	− 8.0	− 8.2
LARFFS	− 6.1	− 8.7	− 6.6	− 8.9
LARFFS-CPU	− 7.4	− 8.0	− 6.5	− 8.5
LARFFS-NPU	− 7.1	− 8.9	− 6.9	− 9.7
LARFFS-MPY	− 5.7	− 8.2	− 5.8	− 7.8
LARFFS-PYC	− 8.4	− 8.2	− 5.0	− 7.3
CPU	− 7.8	− 7.6	− 6.9	− 6.6
NPU	− 6.5	− 7.9	− 7.1	− 7.7
MPY	− 4.7	− 7.0	− 4.2	− 6.5
PYC	− 5.0	− 6.1	− 4.6	− 5.9

and 4 which highlight H-bond interactions and hydrophobic interactions, respectively. As can be seen in the case of the wild-type EGFR kinase domain, while both LARFFS and WNWKV and their respective conjugates showed a few commonalities in interactions within the binding pocket, some of the residues involved were different.

Interestingly LARFFS-PYC, LARFFS-NPU, WNWKV-NPU, WNWKV-CPU, and WNWKV-(PYC)₂ showed H-bond interactions with the DFG motif residue ASP 831 [91]. This is a critical interaction found in a number of EGFR inhibitors and is therefore promising. Additionally, WNWKV-CPU, WNWKV-NPU, WNWKV-MPY, and LARFFS-CPU displayed H-bond interactions with MET 769 in the region just following the α C-helix, which has also been implicated in interacting with several EGFR inhibitors [92]. The most prominent residue involved in hydrophobic interactions was PHE 699 from the gly-rich P-loop region [93] which interacted with all of the conjugates except WNWKV-MPY; neat LARFFS and LARFFS-MPY. Other common residues involved in interacting with multiple conjugates included LYS 721, VAL 702 CYS 773, and ASP 776. In addition, LEU 694 also displayed multiple interactions with WNWKV-CPU, WNWKV-NPU, WNWKV-MPY, LARFFS-PYC, and LARFFS-CPU. The neat peptide WNWKV and LARFFS-PYC displayed H-bond interactions with catalytic loop residues ASP 813, ARG 817, and ASN 818, while neat LARFFS showed several H-bond interactions with SER 888, ARG 779, and LYS 889. In addition, interactions with PRO 853, which is also part of the activation loop are also seen [94] in the case of WNWKV-PYC and WNWKV-CPU. While one of the indole rings in neat

WNWKV did participate in a salt-bridge interaction with LYS 855 wild-type EGFR, the tryptophan side chains of WNWKV and the purine derivative conjugates generally participated in H-bond interactions rather than hydrophobic interactions. These H-bonds were formed with the -NH in the indole rings of tryptophan (Supplementary Information Figures S1 and S2). However, the indole rings in the pyrimidine derivative conjugates tended to form more hydrophobic interactions than H-bonds. A salt bridge was also observed between WNWKV-PYC and ARG 817. One notable exception to this trend was seen for WNWKV-MPY, which showed the Trp moiety forming multiple H-bond interactions with MET 769.

For the T790M/L858R double mutant, the most prominent H-bond interaction was seen with the activation loop residue ARG 841, which was common for WNWKV and all of its conjugates with the exception of WNWKV-(PYC)₂, with which it showed a hydrophobic interaction. ASP 800 and CYS 797 from the hinge region were also found to form H-bonds across most of the WNWKV conjugates and the WNWKV peptide. Exceptions included WNWKV-CPU and WNWKV-MPY which did not display any interaction with CYS 797. WNWKV-MPY also did not display H-bond interactions with ASP 800. Interestingly, ASP 855 from the activation loop also formed H-bonds with WNWKV-CPU, WNWKV-MPY, and WNWKV-(PYC)₂. Other residues that showed interactions with most of the WNWKV conjugates included LEU 718, which not only formed H-bonds but also displayed hydrophobic interactions. Remarkably, VAL 726 and PHE 723 displayed hydrophobic interactions with WNWKV and all of its conjugates implying their important

Table 3 Hydrogen bond interactions of WNWKV (W) and LARFFS (L) and its conjugates with wild-type EGFR and double mutant EGFR kinase domain

W-CPU	W	W-NPU	W-MPY	W-PYC	W-(PYC) ₂	W-CPU	W	W-NPU	W-MPY	W-PYC	W-(PYC) ₂
ALA719	ASP813	LEU694	MET769	ASP776	LYS721	LEU 718	LEU 718	MET 793	ARG 803	LEU 718	ALA 722
LYS721	ARG817	LYS721	MET769	ASP813	ASP831	SER 720	PHE 795	GLY 796	ARG 803	MET 793	CYS 797
LYS721	ASN818	MET769	CYS773	ARG817	ASP831	SER 720	CYS 797	CYS 797	ASP 837	CYS 797	ASP 800
GLU734	LYS851	MET769			ASP831	ASP 800	ASP 800	ASP 800	ARG 841	ASP 800	ARG 803
GLU738	LYS855	CYS773			ASP831	ARG 841	ARG 841	ASP 800	ARG 841	ARG 841	ASP 855
THR766		ASP776			GLY833	ARG 841		ARG841	ASN842	ASP 855	
MET769		GLU780				ARG 841		ILE878	ASN842		
MET769		THR830				ASP 855		LYS 879	ASP855		
CYS773		ASP831						LYS 913			
CYS773								ASP 916			
ASP776								ASP 916			
ASP831											
L-CPU	L	L-NPU	L-MPY	L-PYC		L-CPU	L	L-NPU	L-MPY	L-PYC	
PHE771	GLN958	ARG817	ASN802	LYS721		LYS716	ALA722	ALA722	LYS754	LEU 703	
CYS773	LYS889	ASP831	ASN802	GLU734		LYS728	ALA722	PHE723	GLU758	LEU 703	
ASP776	SER888	ASP831	GLU805	GLU734		PRO794	PHE723	ARG803	ARG836		
GLU780	SER888	GLY833	ARG808	ARG812		LYS846	SER752	ASP837	ARG836		
GLU780	SER888	ILE854	ARG808	ASP813		HIS850	SER752	ARG841	ARG841		
ARG817	SER888		ARG808	ARG817		GLU1015	LYS875	ARG841	ASN842		
	THR885		ALA840	ASN818				VAL876	ASN842		
	ARG779		HIS869	ASP831				ILE878	LYS875		
	ARG779		LYS936	ASP831					LYS875		
				LYS836							

*Residues that appear only once are shown in white background. Color coded residues indicate that those residues were involved in interacting more than once either with the same or different ligands

role in binding with those conjugates. Moreover, PHE 723 also promoted π - π stacking interactions with WNWKV and its conjugates. Additionally, ALA 722 was found to form H-bond interactions with WNWKV-(PYC)₂, LARFFS, and with LARFFS-NPU, while ALA 743 from the α C-helix region displayed hydrophobic interactions with WNWKV-CPU, WNWKV-NPU, WNWKV-(PYC)₂, and WNWKV. Of note is the fact that LARFFS-CPU showed a unique H-bond interaction with GLU 1015 and hydrophobic interaction with VAL 1010, which are part of the C-terminal tail region of the kinase domain, while also interacting with the hinge region residues PRO 794 and PHE 795 as well as the activation loop residue HIS 850 and with β 3/ α C region residues LYS 716 and LYS 728. LARFFS-NPU and LARFFS-MPY also showed interactions with the activation loop residue ARG 841. In addition, both LARFFS-NPU and WNWKV-NPU

formed H-bonds with ILE 878. Overall, LARFFS-PYC showed the least number of interactions (four hydrophobic and two H-bonds) with the double mutant receptor, while the peptide WNWKV showed the highest number of hydrophobic interactions (twelve) and WNWKV-NPU formed the highest number of H-bonds. These results are promising given that the conjugate appears to make critical interactions with residues within the ATP binding cleft of the kinase domain that were found to interact with previous drugs that were designed. The binding interactions are likely enhanced due to the presence of the indole moiety in the WNWKV peptide and its conjugates, which has been shown to be effective in previous work in binding to EGFR [95].

We also compared the binding interactions of both receptors with the four individual molecules without conjugation (CPU, NPU, PYC, and MPY). Results are shown in

Table 4 Hydrophobic interactions of WNWKV (W) and LARFFS (L) and its conjugates with NPU, CPU, MPY, and PYC with wild-type EGFR and double mutant EGFR kinase domain

W-CPU	W	W-NPU	W-MPY	W-PYC	W-(PYC)2	W-CPU	W	W-NPU	W-MPY	W-PYC	W-(PYC)2
LEU694	PHE699	LEU694	LEU694	ALA698	PHE699	LEU718	LEU718	PHE723	PHE723	LEU718	LEU718
LEU694	PHE699	LEU694	LEU694	PHE699	PHE699	LEU718	LEU718	PHE723	PHE723	LEU718	PHE723
LEU694	PRO853	LEU694	VAL702	VAL702	PHE699	LEU718	PHE723	PHE723	VAL726	PHE723	PHE723
PHE699	ILE854	PHE699	ALA719	ARG817	VAL702	PHE723	PHE723	PHE723	LEU799	VAL726	VAL726
VAL702	LYS855	VAL702	PHE771	LEU820	LYS721	VAL726	VAL726	VAL726	LEU844	VAL726	ALA743
VAL702	TRP856	LYS721	LEU820	PRO853	LEU723	LYS728	VAL726	VAL726	LEU844	VAL726	MET793
TYR777	ALA896	GLU734		PRO853	ALA731	ALA743	ALA743	ALA743		LYS913	ARG841
GLU780					GLU734	LEU792	MET793	ASP855			LEU844
PRO853					ILE735	PRO794	LEU799	PRO877			
					ARG817		LEU844				
					LEU834		LYS879				
							TRP880				
L-CPU	L	L-NPU	L-MPY	L-PYC		L-CPU	L	L-NPU	L-MPY	L-PYC	
LEU694	ILE914	PHE699	GLU805	LEU694		LEU730	LEU747	PHE723	ALA722	GLN701	
LEU694	PRO913	LEU723	LYS843	PHE699		PRO794	LEU747	PHE723	LEU747	LEU703	
PHE699	PRO913	GLU734	HIS869	PHE699		PHE795	ILE759	LEU799	LEU747	LEU703	
VAL702	PRO910	ILE375	GLN870	PHE699		PHE795	ILE759	ARG858	GLU758	LEU1017	
LEU768	PHE886	ILE375		VAL702		GLU1005	ARG858	TRP880	ARG858		
	TRP881	LEU834		VAL702		VAL 1010	LYS875	LYS913	VAL876		
	TRP881	ILE854		VAL702		VAL1010	LYS875	LYS913			
	LYS782	TRP856		ARG817			LYS875				
				LEU820							
				LEU834							
				LYS851							

*Residues that appear only once are shown in white background. Color coded residues indicate that those residues were involved in interacting more than once either with the same or different ligands

Supplementary Information Figure S3. As expected, relatively lesser number of interactions were seen with the individual molecules. With the wild-type EGFR, (Supplementary Information Table S1) CPU formed hydrophobic interactions close to the P-loop region with LEU 694 in addition to β 1 and C-helix residues. One hydrophobic interaction was seen with LEU 820 in the helix-2 region. Interactions with NPU, on the other hand, included those with the DFG motif residue ASP 831, as well as with the A-loop residues GLY 833 and LEU 835. An H-bond interaction with the catalytic loop residue ASP 813 was also seen in addition to PHE 699 from the Gly-rich P-loop. Thus, NPU seemed to have more critical binding interactions within the binding pocket compared to CPU implying that the pyridyl moiety

has higher interactions compared to the toluene containing CPU. Compared to NPU and CPU, MPY and PYC displayed fewer interactions. MPY only showed interactions with three residues (THR 766, LEU 764, and LYS 721), while PYC showed one interaction with the DFG motif residue ASP 831 in addition to LYS 721 and GLU 738. With the T790M/L858R receptor, (Supplementary Information Table S2) CPU showed one interaction with the A-loop residue ARG 841, while other interactions such as LEU 747, GLU 758, and ILE 759 were seen within the C-helix region. ALA 722 from β 2 region was also involved in one H-bonding interaction. NPU, on the other hand, displayed H-bond interactions with the hinge region residue MET 793, in addition to other interactions from β 1- β 2 region residues and one interaction

with LEU 844 of the activation loop. MPY interestingly only formed four H-bonds with ASN 771, VAL 774, LYS 852, and ASP 1014. No other interactions were observed. PYC showed even fewer interactions (three H-bonds). Although the interactions were closer to the catalytic loop residues (SER 811 and GLN 812), one interaction was observed further into the C-lobe with GLN 976.

Overall, conjugation with the peptides generally increased the number of interactions with the receptors. In particular, the LARFFS-CPU conjugate showed interactions away from the binding pocket toward the C-terminal domain with the double mutant EGFR, while it showed interactions within the binding pocket hinge region for the wild-type receptor. Furthermore, WNWKV and its conjugates were found to make higher number of binding interactions within the binding pocket residues of A-loop, P-loop, hinge region, and in some cases, the catalytic loop regions compared to those seen for LARFFS and its conjugates. The LARFFS-PYC conjugate, however, showed the highest number of interactions within the binding pocket among the LARFFS conjugates with the wild-type EGFR compared to that seen for the double mutant. Interestingly, it also showed the least number of interactions with the double mutant receptor.

Molecular dynamics studies

Because overall docking studies indicated higher binding affinities with the peptides and conjugates, MD simulations were conducted with the conjugates and the neat peptides in order to gain further insight into the stability of the protein–ligand complexes. All MD simulations were performed over 250 ns and the results reported are an average of three independent simulations for each complex. Results of the root mean square deviations for the receptor C α and that of ligand bound receptors are shown for Fig. 5. In general, as can be seen in the case of the wild-type receptor, the C α RMSD (Fig. 5a) attained stability after 50 ns and remained stable without any significant deviations for the rest of the simulation. The apo receptor displayed the lowest RMSD values as expected (0.8 nm), while the C α for LARFFS-NPU stabilized at 1.2 nm, which was the highest. All others displayed values in the range of 0.9 nm to 1.1 nm. The protein–ligand complex with the wild-type kinase domain of EGFR RMSD plots (Fig. 5b) displayed stability and very minimal difference with the C α RMSD for most of the WNWKV conjugates when complexed with the wild-type EGFR. In particular, WNWKV-PYC, LARFFS-PYC, WNWKV-NPU, WNWKV-(PYC)₂, and WNWKV-MPY showed RMSD values between 0.6 nm and 1.2 nm at the end of the simulation. In the case of WNWKV-CPU, for the first 40 ns, the RMSD was low (0.4 nm). However, it continued to increase over time, reaching 1.7 nm by the end of the simulation. This indicates that, compared to some of

the other WNWKV conjugate complexes, the WNWKV-CPU complex is relatively less stable. Interestingly, compared to the conjugates, WNWKV peptide by itself showed relatively less stability, particularly for the first 60 ns where it showed deviations, after which the RMSD value came down and remained fairly constant at 2.3 nm for the rest of the simulation.

The LARFFS-MPY conjugate formed the least stable complex with the wild-type receptor and the RMSD value remained high for most of the simulation. However, a gradual decrease was observed after 150 ns, at which point the RMSD did not show significant deviations. The LARFFS-NPU conjugate, on the other hand, starts off stable and showed very little deviations up to 100 ns. However, a gradual increase is seen between 100 and 150 ns, after which it stabilized at 2.1 nm. This is likely due to a conformational change occurring within the receptor complexed with LARFFS-NPU. Thus, in general, for the wild-type receptor, most WNWKV conjugates formed more stable complexes compared to LARFFS conjugates. The corresponding trajectory snapshots at different time points over the 250 ns simulations are shown in Supplementary Information Figures S4 and S5. The trajectory images corroborate with the RMSDs. In particular, WNWKV-(PYC)₂, WNWKV-MPY, LARFFS-PYC, and WNWKV-NPU remain attached within the ATP binding cleft throughout the simulation, making key interactions within the ATP binding pocket with LEU 694, PHE 699 as well as with LYS 721, ARG 817, MET 769, and ASP 831. Thus, the presence of a methyl-pyridyl ring in NPU in place of the toluene ring located in CPU may have allowed WNWKV-NPU to form stronger, more stable interactions with ASP 776, ASP 831, and GLU 780 residues in wild-type EGFR. In particular, the DFG motif residue, ASP 831, is implicated in the binding to several EGFR targeting chemotherapeutic drugs, including a series of pyrazole derivatives [96]. Previous research has also shown that ARG 817 from the catalytic loop [97, 98] is also a critical residue in binding to EGFR kinase inhibitors such as Midostaurin [99]. Thus, these results are promising. On the other hand, LARFFS-NPU and LARFFS-MPY barely remain attached to the receptor by the end of the simulation accounting for the high RMSD values. Interestingly, unconjugated LARFFS and WNWKV peptides, as well as WNWKV-PYC, appear to attach below the ATP binding cleft toward the C-lobe, making key contacts with SER 888, ARG 908, TYR 891, and LYS 889. Additionally, while WNKV did initially appear to bind within the binding pocket of wild-type EGFR, over the course of the simulation, the peptide moved away from the binding pocket while still showing interaction with the C-lobe residue ASP 892. Out of all of the LARFFS conjugates, LARFFS-PYC also formed stable complex within the active site of wild-type EGFR. The conjugate displayed interactions with several residues, including ASP 813, ARG

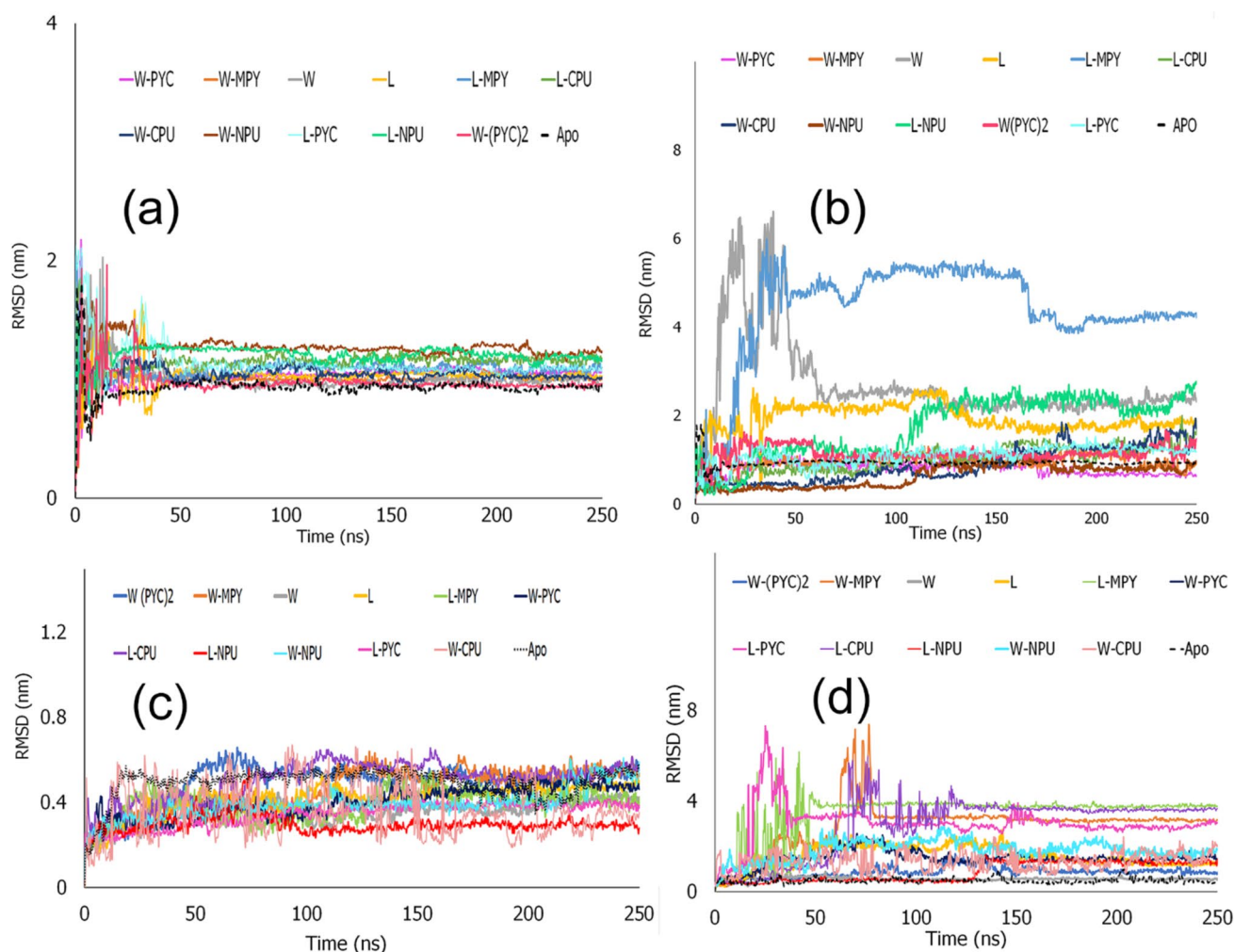


Fig. 5 RMSD plots for EGFR kinase domain complexes with peptides and conjugates. **a** C α RMSD for wild-type EGFR; **b** protein–ligand complex RMSD for wild-type receptor; **c** C α RMSD for

T790M/L858R EGFR; **d** protein–ligand complex RMSD for T790M/L858R receptor. Dotted lines in all cases are indicative of apo receptors

817, PHE 699, and ASP 831. In previous work, EGFR targeting inhibitors, including quinazoline derivatives containing pyrimidine moieties, form interactions with PHE 699, indicating the importance of its participation in binding with these complexes [100]. Overall, the trajectories indicate that WNWKV conjugates formed more stable complexes with the wild-type EGFR than the LARFFS conjugates. Compounds with average PL-RMSD values below 0.7 nm were seen to remain stably bound within the main active site region of the kinase domain of wild-type EGFR. While WNKWV-NPU, WNWKV-MPY, and WNKWV-(PYC)₂ all met these criteria, LARFFS-PYC was the only LARFFS conjugate to do the same.

The C α RMSD values were found to be lower in all cases for the double mutant receptor (Fig. 5c) compared to the wild-type receptor. In general, very little deviations were observed and RMSD values ranged from 0.25 to 0.5 nm

across all peptides and conjugates. In the case of the ligand bound receptor complexes (Fig. 5d), WNWKV and WNWKV-(PYC)₂ formed the most stable complexes and showed lowest RMSD values (ranging from 0.4 to 0.5 nm). The RMSDs of WNWKV-CPU, WNWKV-NPU, and WNWKV-PYC showed deviations in the first 100 ns, but gradually stabilized at 1.1 nm by the end of the simulation. Interestingly, once again WNWKV-MPY, LARFFS-MPY, LARFFS-PYC, and LARFFS-CPU showed high RMSD values with significant deviations for the first 120 ns, after which the complexes were found to stabilize between 3.0 nm and 3.6 nm indicating those formed less stable complexes with the T790M/L858R kinase domain of the receptor. The trajectory images indicate that WNKWV and its conjugates are mostly interacting with residues within the ATP binding pocket, encompassing the C-helix, hinge region, and the Gly-rich loop residues as well as activation loop residues

in some cases. Most common residues that were found to interact during the course of the simulation included ARG 748, ASN 756, SER 720, PRO 794, and LEU 718 (Supplementary Information Figure S6). The LEU 718 residue is known to form hydrophobic interactions with kinase inhibitors, including erlotinib, especially with the presence of the L858R mutation [101–104] which is encouraging. However, in the case of the conjugate WNWKV-MPY, the ligand appears to be constantly changing positions throughout the simulation and part of the conjugate remains outside of the binding pocket by the end of the simulation which accounts for the high deviations seen. Interestingly, while the WNWKV-(PYC)₂ conjugate remains firmly attached, making several contacts with the ATP binding pocket residues including with the activation loop residue ARG 841, the corresponding monoconjugate WNWKV-PYC constantly changes conformation within the binding pocket though at the end of the simulation, it is found to interact with residues ARG 841, CYS 797, and ASN 756. In particular, the interaction with CYS 797 is highly promising as it has been implicated the C797S mutation and can cause drug resistance to kinase inhibitors, including pyrimidine derivatives [105, 106]. Compared to the WNWKV conjugates, most LARFFS conjugates were found to change positions during the course of the simulation as seen in Supplementary Information Figure S7. In particular, LARFFS-CPU, which was mostly outside of the binding pocket, moves further inwards and interacts with the C-lobe residues SER 912, ASP 916 toward the end of the simulation. Very few contacts are seen with residues within the ATP binding pocket with the exception of LYS 806 which occurs between 150 and 200 ns of the simulation. Interestingly, however, the LARFFS-NPU conjugate remains stable throughout the simulation making critical contacts within the ATP binding pocket once again showing that the pyridyl ring in place of the toluoxy ring aids in binding within the ATP binding pocket. Like the WNWKV-MPY conjugate, the LARFFS-MPY conjugates are also constantly changing positions within the binding pocket, thus accounting for its high RMSD. Likewise, LARFFS-PYC also was found to move away from the binding pocket, and only the peptide part of the conjugate remained attached to the double mutant receptor accounting for the deviations seen during the simulation. Thus, both the MD simulations revealed that overall the WNWKV peptide conjugates formed more stable complexes with the double mutant receptor.

Analysis of root mean square fluctuations

We analyzed the RMSF (root mean square fluctuation) of the protein backbone of both the wild-type and double mutant receptors upon binding to the various peptide conjugates and peptides over the 250 ns trajectory. Each simulation

was run independently three times and the data reported are the mean of the backbone values (Fig. 6) RMSF provides information about the flexibility of residues when complexed with the ligands [107]. As can be seen, in both the wild-type and the double mutant, the highest fluctuations occurred in the *-N* and *C*-terminal regions, which is expected given that those regions are known to display higher flexibility [108, 109]. The noteworthy observations during MD simulations were that residues forming the P-loop, *C*-terminal loop of the α C-helix region (A767–G779) [110], A-loop, and the β 3-alpha-C loop showed higher flexibility as indicated by high degree of fluctuation. Thus, overall, the loop regions displayed high degree of flexibility compared to the beta sheets and alpha-helices. In general, the regions containing residues showing higher RMSF values suggest more flexibility which indicates higher plausibility to interact with the peptide conjugates and the peptides involved [111]. In the case of the double mutant, WNWKV-CPU showed fluctuations throughout the simulation, implying that the ligand is likely highly mobile within the binding pocket which was also seen in the trajectory images. Specifically, higher and broader fluctuations are observed in the distal hydrophobic pockets C-lobe residue region encompassing TRP 880 through PRO-934 implying their impact on binding with the ligand. Previous research has shown that the C-lobe residues can be a potential target for EGFR inhibitors as they can bind to certain endogenous ligands and suppress the over-expression of EGFR [112]. In addition, high fluctuations are also seen in the α C-helix region residues, which is significant as α C-helix plays an important role in EGFR kinase domain as it generally forms a docking site on the receiver kinase domain [113]. The RMSF values in the case of the wild-type receptor did not show significant variations, though the C-loop region of the α C-helix region and the C-lobe residues PRO 910 through MET 921 showed slightly higher fluctuations with the LARFFS-PYC conjugate compared to the other conjugates.

MMGBSA studies

The free energies of the binding interactions of both the wild-type and the double mutant receptors with the peptides and conjugates were estimated by molecular mechanics with generalized Born and surface area solvation (MM/GBSA) [114]. The generalized Born continuum solvent model takes into account a continuum representation of the solvent along with an all atom force field (in this case OPLS4) and is acquiescent to calculation of gradients that are needed for minimizations. An optimized implicit solvent model system (VSGB 2.1) was utilized, given that the samples were subjected to a homogenous environment, containing water/ions and all simulations were done at constant temperature (310°K). Additionally, implicit solvent model

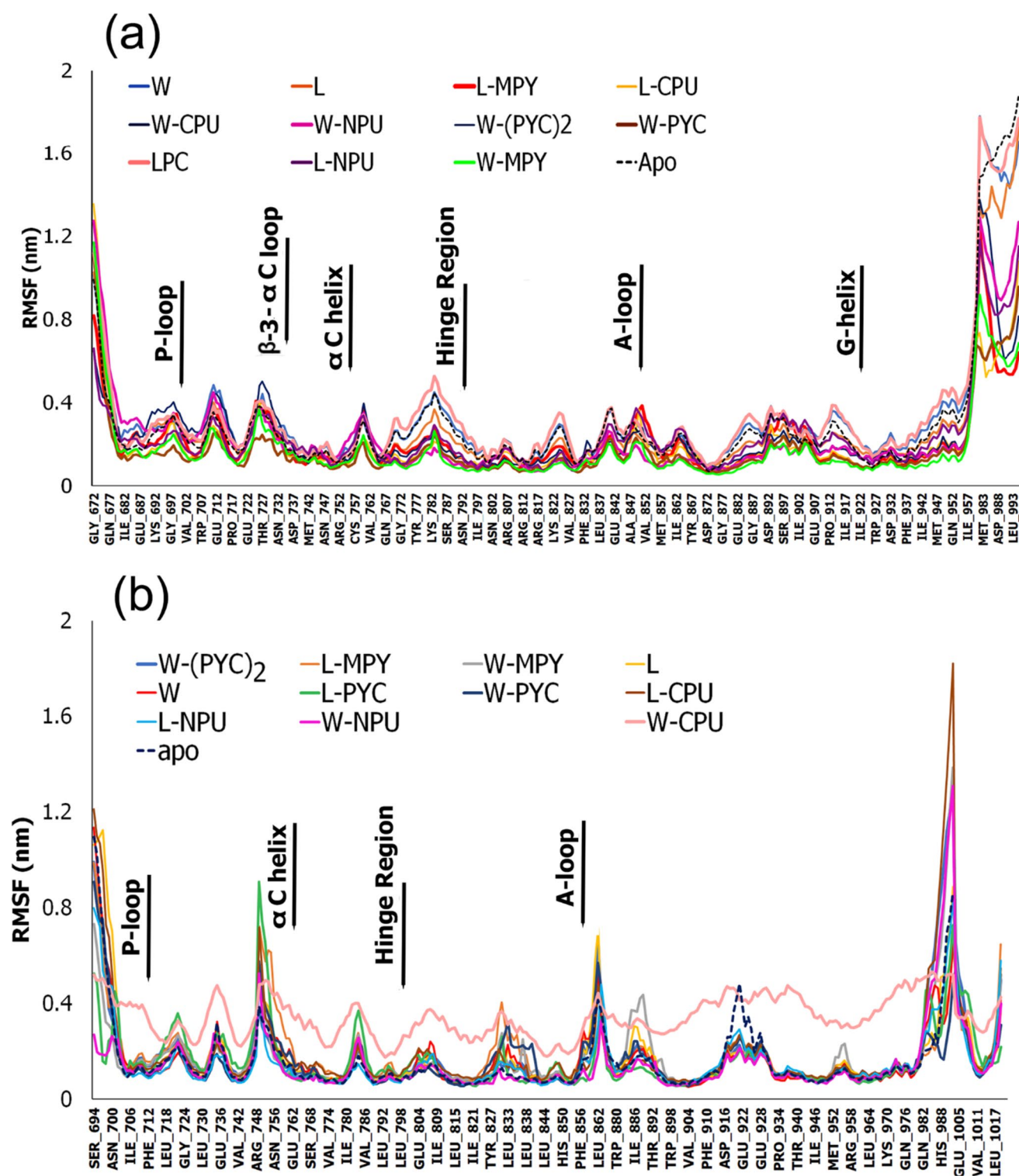


Fig. 6 Comparison of RMSF of protein backbone of **a** Wild Type EGFR kinase domain upon binding to conjugates and peptides designed; **b** Comparison of RMSF of protein backbone of double mutant T790M/L858R EGFR upon binding to conjugates and peptides

systems provide a significantly high algorithm flexibility and provide improved sampling, and the samples can explore the available conformational space relatively faster [115]. The

solvation energies (Solv GB) obtained are indicative of the electrostatic solvation energy. The results shown (Tables 5 and 6) are the averages obtained for three separate MD

simulations. Overall, the negative ΔG bind values obtained for the double mutant receptor were approximately twice that of those obtained for the wild-type receptor, which indicates that all conjugates had a higher binding energy toward the double mutant receptor. Interestingly, both in the case of the wild-type and the double mutant, the pyrimidine conjugates displayed slightly higher binding energies. In particular, WNWKV-MPY and WNWKV-(PYC)₂ showed the highest binding energy toward the wild-type receptor with values of -36.67 kcal/mol and -34.87 kcal/mol, respectively. For the double mutant, the WNWKV-PYC conjugate displayed the highest ΔG bind value at -74.65 kcal/mol followed by LARFFS-PYC at -70.19 kcal/mol. The results of the components of the total energy contributors are also shown, which indicated that coulomb energy followed by Van der Waals energy played the most significant roles in contributing to the binding energies (see Table 6).

Prediction of pharmacokinetic properties

The web server ADMETlab2.0 was used to predict the pharmacokinetic properties of the peptides, peptide conjugates, and unconjugated compounds. Results are shown in Table 7. The table includes $\log P$ scores, which are an indication of lipophilic character of a drug candidate. Specifically, it is the partition coefficient between the aqueous and lipophilic phases [116]. $\log P$ scores for the peptides and conjugates ranged from 0.318 to 2.013, indicating that they display drug like properties [117]. The $\log P$ scores were generally greater for the conjugates than the individual peptides, particularly for the purine derivative conjugates as they contain more hydrophobic ring systems. The unconjugated compounds generally had low $\log P$ scores, with MPY and PYC displayed negative $\log P$ values, indicating that these compounds are more hydrophilic. None of the tested ligands

were indicated to be hERG blockers, signifying that the drugs will not cause cardiotoxicity related to hERG channel inhibition [118]. Additionally, all of the compounds were shown to be negative for AMES toxicity, indicating that they are not likely to be mutagenic substances. [119] Madin-Darby canine kidney (MDCK) cells are commonly used for studying cellular interactions with drugs, including cell permeability and cellular uptake [120]. The MDCK cell permeability values indicate that the ligands are expected to permeate the cellular membrane. Pgp, an efflux pump that is often over-expressed in tumor cells, is involved in the transport and absorption of drugs and may be involved in the development of multi-drug resistance [121].

Interactions of drugs with pgp are important to examine due to the possibility of drug-interactions, particularly since co-administration of substrates and inhibitors can lead to unwanted side effects [122]. While none of the ligands were determined to be Pgp inhibitors, most of the conjugates were determined to be Pgp substrates with the exception of LARFFS-PYC. CPU and NPU were also categorized as Pgp substrates. This indicates that the drug efficiency of these compounds may be affected due to interactions with Pgp. Cytochrome P450 (CYP) is a group of enzymes involved in the metabolism of many drugs, so interactions with CYP are important to examine when developing drugs [123]. While LARFFS and the LARFFS conjugates were classified as neither CYP substrates nor inhibitors, many of the WNWKV conjugates were categorized as CYP inhibitors. Specifically, WNWKV-CPU, WNWKV-NPU, and WNWKV-MPY were predicted to inhibit the activity of CYP3A4, which is mainly involved in phase I metabolism. While inhibition can increase the bioavailability of certain drugs, it may result in side effects including cholestasis and disruption of endocrine system signaling, so further study would have to be conducted to determine safe administration of such drugs [124].

Table 5 MMGBSA of each peptide and conjugate with wild-type EGFR kinase domain

Peptide/conjugate	ΔG Bind Kcal/mol	Coulomb Kcal/mol	H-bond Kcal/mol	Lipophilic Kcal/mol	Solv GB Kcal/mol	van der Waals Kcal/mol	Pi-Pi Packing Kcal/mol
WNWKV	-26.13	-133.74	-1.22	-5.20	133.61	-21.27	-0.0006
WNWKV-CPU	-28.55	-145.22	-2.65	-4.23	148.04	-25.29	-0.42
WNWKV-NPU	-30.82	-162.20	-2.88	-5.35	168.18	-31.04	-0.083
WNWKV-MPY	-36.67	-143.18	-0.31	-1.77	142.27	-28.90	0
WNWKV-PYC	-33.44	-139.54	-1.76	-5.71	139.47	-26.44	-0.001
WNWKV-(PYC) ₂	-34.87	-152.79	-2.09	-8.52	152.45	-24.89	-0.001
LARFFS	-32.80	-135.72	-2.22	-6.54	135.54	-25.00	0
LARFFS-CPU	-26.95	-130.29	-2.50	-2.48	132.75	-24.47	-0.001
LARFFS-NPU	-32.51	-113.38	-1.02	-6.76	112.52	-24.41	-0.31
LARFFS-MPY	-23.79	-117.84	-1.59	-3.99	116.51	-17.63	0
LARFFS-PYC	-21.85	-120.87	-2.27	-3.12	121.03	-18.24	-0.015

Table 6 MMGBSA of each peptide and conjugates with T790M/L858R EGFR kinase domain

Peptide/conjugate	ΔG Bind Kcal/mol	Coulomb Kcal/mol	H-bond Kcal/mol	Lipophilic Kcal/mol	Solvent GB Kcal/mol	Van der Waals Kcal/mol	Pi-Pi packing Kcal/mol
WNWKV	− 64.99	− 301.25	− 5.75	− 13.27	301.05	− 49.13	− 0.25
WNWKV-CPU	− 69.04	− 307.26	− 6.39	− 14.07	351.73	− 54.11	− 0.21
WNWKV-NPU	− 67.91	− 291.84	− 5.33	− 15.24	293.80	− 52.62	− 0.22
WNWKV-MPY	− 65.99	− 303.28	− 7.04	− 13.32	300.73	− 46.54	− 0.54
WNWKV-PYC	− 74.65	− 301.20	− 5.92	− 15.62	302.01	− 59.41	− 0.23
WNWKV-(PYC) ₂	− 61.48	− 299.92	− 5.34	− 12.56	298.89	− 45.83	− 0.43
LARFFS	− 67.93	− 296.36	− 4.91	− 15.83	298.53	− 54.17	− 0.02
LARFFS-CPU	− 66.64	− 325.10	− 6.65	− 13.77	325.03	− 50.43	− 0.39
LARFFS-NPU	− 59.42	− 313.19	− 5.66	− 11.56	312.91	− 42.89	− 0.40
LARFFS-MPY	− 64.23	− 306.48	− 6.43	− 13.31	304.54	− 46.28	− 0.41
LARFFS-PYC	− 70.19	− 282.50	− 5.81	− 15.16	287.51	− 57.82	− 0.06

Table 7 ADME studies of peptides, peptide conjugates, and nucleotide derivatives

Compound	LogP	MDCK cell permeability	hERG blocker	CYP inhibitor/substrate	Pgp substrate/inhibitor	AMES toxicity
WNWKV	0.318	3×10^{-6}	No	No/No	No/No	No
WNWKV-CPU	2.013	1×10^{-6}	No	Yes (for CYP3A4)/No	Yes/No	No
WNWKV-NPU	1.377	1×10^{-6}	No	Yes (for CYP3A4)/No	Yes/No	No
WNWKV-MPY	1.797	1×10^{-6}	No	Yes (for CYP3A4)/No	Yes/No	No
WNWKV-PYC	0.642	1×10^{-6}	No	No/Yes (for CYP2C9)	Yes/No	No
WNWKV-(PYC) ₂	1.051	2×10^{-6}	No	Somewhat (for CYP2C9) /Yes (for CYP2C9)	Yes/No	No
LARFFS	0.29	0.000137	No	No/No	No/No	No
LARFFS-CPU	1.67	2×10^{-6}	No	No/No	Yes/Somewhat	No
LARFFS-NPU	1.134	1×10^{-6}	No	No/No	Yes/No	No
LARFFS-MPY	1.366	4×10^{-6}	No	No/No	Yes/no	No
LARFFS-PYC	0.48	7×10^{-6}	No	No/No	No/No	No
CPU	0.549	2×10^{-6}	No	No/No	Yes/no	No
NPU	0.008	2×10^{-6}	No	Yes (for CYP3A4)/Yes (for CYP2C9)	Yes/no	No
MPY	− 1.854	1.1×10^{-5}	No	No/Yes (for CYP1A2)	No/no	No
PYC	− 0.469	9×10^{-6}	No	No/Yes (for CYP1A2)	No/no	No

WNWKV-PYC and WNWKV-(PYC)₂ were characterized as substrates of CYP2C9, which metabolizes many drugs including sulfonyleureas and anticoagulants [125]. Thus, the effectiveness of these drugs may be altered due to interactions with CYP2C9.

SPR analysis

In order to validate the computational results, as a proof of concept, we synthesized and examined the binding interactions of the neat peptides as well as the LARFFS-PYC and WNWKV-(PYC)₂ conjugates with both receptor proteins. Results are shown in Table 8. The corresponding SPR

sensograms are shown in Supplementary Information Figure S8. The concentration of each peptide or conjugate was varied between 50 nM and 100 μ M. Each sample was run three times. The average KD values were determined. Overall, all conjugates showed higher binding with the T790M/L858R EGF receptor. In the case of the wild-type receptor, compared to the two neat peptides, both the LARFFS-PYC and WNWKV-(PYC)₂ showed higher binding, indicating once again that conjugation enhanced binding interactions. The highest binding (lowest KD) was found to be for WNWKV-(PYC)₂ with both receptors, though the KD was significantly lower for the T790M/L858R receptor. These results corroborate with the computational studies which also indicated that

the conjugates demonstrated higher binding with the double mutant receptor.

Cell studies

Having confirmed binding interactions, we explored the cytotoxicity effects of WNWKV-(PYC)₂ and LARFFS-PYC as well as the neat peptides by conducting in vitro cell studies with A549 lung carcinoma epithelial cells known to overexpress wild-type EGFR and NCI-H1975 lung cancer cells expressing EGFR [L858R/ T790M] [126]. The morphologies of the cells before and after treatment with the peptides and the constructs were examined (Fig. 7). The EGFR T790M/L858R double mutant expressing untreated NCI-H1975 cells showed healthy growth with epithelial-like morphology, exhibiting focal adhesion points and cell-to-cell contacts [127] as expected. Upon treatment with the LARFFS peptide, we observed a change in morphology. Cells appeared to round up, while cell ruffling was also observed in a few cells. This resulted in cell membrane re-organization. Upon treatment with LARFFS-PYC and WNWKV peptide, however, the cells appeared to be rounded up and cell blebbing was observed. In the case of the positive control, Dasatinib, some cells showed blebbing, while overall fewer live-cells were observed. In the case of the WNWKV-(PYC)₂ conjugate, few cells displayed epithelial morphology while most appeared to round up indicating that cell proliferation was disrupted.

The untreated A549 cells demonstrated cobble-shaped morphology and were well spread out, making cell-to-cell contacts, indicating that the cells were healthy and proliferating. After exposure to LARFFS and LARFFS-PYC, the cells appear to have fewer cell-to-cell contacts and are relatively smaller in size; however, a significant loss of morphology was not observed. The cells exposed to Dasatinib, on the other hand, appeared to round up and displayed a loss of morphology indicating that proliferation was stunted. In

comparison, upon treatment with WNWKV peptide, several cells showed complete loss of morphology and fewer cobble-shaped cells were observed. Few live cells however did appear to form cytoskeletal extensions and lamellopodia in the case of the WNWKV-(PYC)₂ treated cells. These results indicate that it is likely that the conjugates are more potent toward the double mutant expressing cells compared to the wild-type EGFR expressing cells.

To further explore specificity, we also examined the effect of treatment of the conjugates on primary lung fibroblast non-cancer cells that do not express EGFR. As seen in the figures, upon treatment with LARFFS or the LARFFS-PYC conjugate, no change in morphology was observed. The fibroblasts showed long spindle-shaped morphology [128] and were well spread out throughout the well plate. This indicated that LARFFS and its PYC conjugate did not have detrimental effects on the fibroblasts. Treatment with Dasatinib, however, resulted in complete loss of morphology, indicating that Dasatinib not only targets tumor cells, but also non-cancer cells. This is likely because Dasatinib is a tyrosine kinase inhibitor that is known to target PDGF-receptor tyrosine kinase activity in fibroblasts [129]. The WNWKV treated fibroblasts appeared to be smaller with less spindle-shaped morphology, while WNWKV-(PYC)₂ by and large maintained the spindle-shaped structures and few cells appeared to round up. Thus, compared to LARFFS and LARFFS-PYC, the WNWKV-treated fibroblasts seem to be to some extent affected upon treatment. However, compared to Dasatinib, all of the conjugates demonstrated significantly lesser cytotoxic effects toward fibroblasts.

In order to quantitatively assess the effects of the conjugates and the peptides on the cell lines, IC₅₀ values (50% viability of cells) were determined for the wild-type EGFR and the double mutant EGFR expressing cells. To calculate IC₅₀, a series of dose–response data from 50 nM to 100 µM concentration (log of concentrations of peptides or conjugates vs % viability) were plotted and calculations were carried out using GraphPad Prism 9.5.0. Results are shown in Table 9. The results of the mean of three independent viability studies for each construct are presented.

As shown in the table, results indicated that the WNWKV-(PYC)₂ conjugate was most potent toward the double mutant EGFR expressing cells, while the WNWKV peptide alone was more potent toward the wild-type EGFR expressing cells. In previous work, it has been shown that Dasatinib displays an IC₅₀ of approximately 2.2 µM against A-549 cells; [130] therefore, these results indicate that the IC₅₀ is relatively higher compared to Dasatinib for the wild-type EGFR expressing cells. The LARFFS peptide and its conjugates were effective against the double mutant cell line; however, the IC₅₀ was significantly higher for the wild-type treated cells, indicating that LARFFS and LARFFS-PYC were less effective against the EGFR wild-type cells.

Table 8 KD values obtained from SPR analysis of PYC conjugates and peptides with EGFR receptors

Compound	Average KD (µM)
Wild type EGFR	
WNWKV	132.5 ± 2.1
WNWKV(PYC) ₂	16.2 ± 0.5
LARFFS	110.5 ± 2.3
LARFFS-PYC	48.41 ± 1.8
T790M/L858R EGFR	
WNWKV	49.6 ± 3.7
WNWKV(PYC) ₂	4.7 ± 2.5
LARFFS	21.2 ± 1.2
LARFFS-PYC	14.3 ± 2.7

Fig. 7 Interactions of Peptides and PYC conjugates with T790M/L858R expressing NCI-H1975 tumor cells (**a–f**). **a** Untreated cells; **b** cells treated with LARFFS; **c** cells treated with LARFFS-PYC; **d** cells treated with Dasatinib; **e** cells treated with WNWKV; **f** cells treated with WNWKV-(PYC)₂. Interactions of Peptides and PYC conjugates with Wild Type EGFR expressing A549 tumor cells (**g–l**). **g** Untreated cells; **h** cells treated with LARFFS; **i** cells treated with LARFFS-PYC; **j** cells treated with Dasatinib; **k** WNWKV; **l** cells treated with WNWKV-(PYC)₂. Interactions of Peptides and PYC conjugates with primary lung fibroblasts (**m–r**). **m** Untreated fibroblasts; **n** cells treated with LARFFS; **o** cells treated with LARFFS-PYC; **p** cells treated with Dasatinib; **q** cells treated with WNWKV; **r** cells treated with WNWKV-(PYC)₂. Scale bar = 25 μ m. All images shown show treatment with 5 μ M constructs after 24 h of incubation. Dastanib was used as a positive control

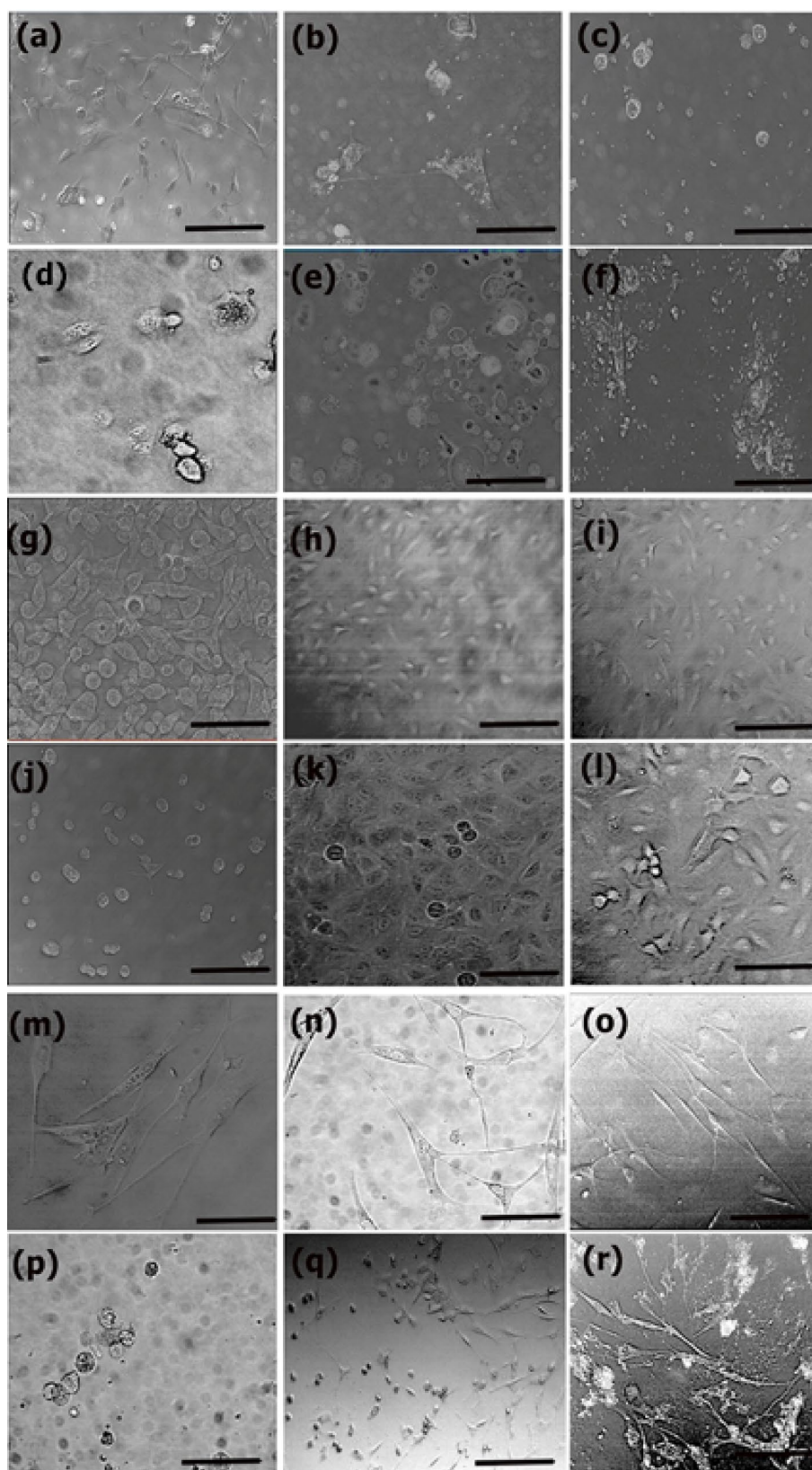


Table 9 Comparison of the effects of the peptides and conjugates on tumor cell lines

Cell line	IC-50 value of WNWKV (μ M)	IC-50 of WNWKV-(PYC) ₂ (μ M)	IC-50 of LARFFS (μ M)	IC-50 of LARFFS-PYC (μ M)
NCI-H1975 (expressing T790M/L858REGFR)	4.6	2.8	3.3	3.7
A-549 (expressing WT EGFR)	4.7	10.3	18.2	20.6

*Each data was calculated as a mean of three separate studies

Apoptosis studies

To determine if the constructs induced cytotoxicity through apoptosis in the wild-type EGFR expressing cells and the double mutant expressing cells, an Annexin-FITC assay using FACS was performed. As seen in Fig. 8, in the case of the double mutant expressing cells, apoptosis was observed in all cases. Specifically, the known positive control Dasatinib showed 58.9% early apoptotic cells, while the WNWKV-(PYC)₂ conjugate exhibited 67.9% early apoptotic cells demonstrating its efficacy in inducing apoptosis in the double mutant cells. In previous work, it has been shown that Dasatinib induces TRAIL-mediated apoptosis in tumor cells [131].

In addition, LARFFS-PYC conjugate also displayed 58.3% early apoptotic cells, while the peptide LARFFS alone showed 30.9% apoptotic cells. These results imply that WNWKV-(PYC)₂ is a strong candidate for reduction of cell proliferation and induction of apoptosis in these cells.

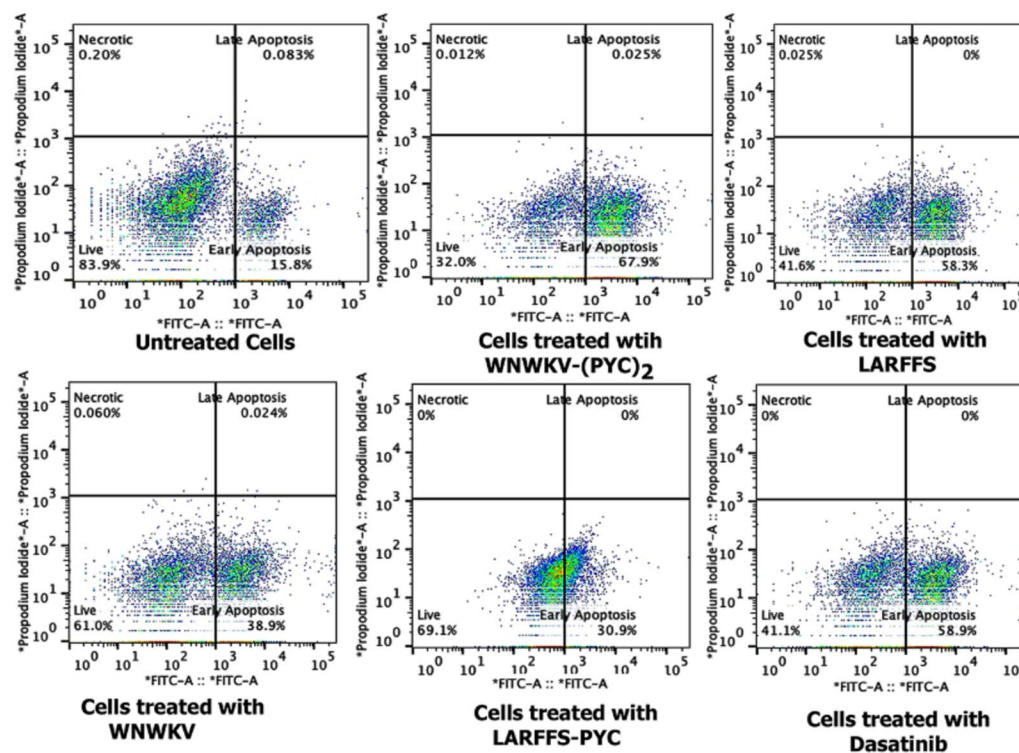
In comparison, while the wild-type EGFR expressing cells demonstrated apoptosis upon treatment with the various conjugates, it was to a much lesser extent compared to the double mutant cells. While the positive control demonstrated significant apoptosis (49%), WNWKV-(PYC)₂ showed only 24.8% early apoptotic cells; while the LARFFS-PYC conjugate showed 26.5% early apoptotic cells, these percentages are only about 10% lesser than the % of apoptotic cells seen in the case of the control. The least effect was seen upon treatment with LARFFS, which only showed a 4% increase in early apoptotic cells compared to the control. WNWKV peptide was the lone candidate among the designed peptides and candidates that showed a relative percentage of early apoptotic cells. Overall, these results indicate that the designed peptides and conjugates are more effective against the double mutant expressing cells. These results are in agreement with the computational studies and SPR which showed stronger binding with the double mutant compared to that observed for the wild-type EGFR expressing cells. We also carried out apoptosis studies with non-cancer cells (lung fibroblasts). Results are shown in Supplementary Information Figure S9. As can be seen, the known drug Dasatinib induced the highest apoptosis in the fibroblast cells, while very little effect was observed upon treatment

with LARFFS and LARFFS-PYC (1.9% and 11.7%, respectively). WNWKV-(PYC)₂ induced 14.3% apoptosis, while the peptide WNWKV showed slightly higher number of early apoptotic cells at 26.3%. These results indicate that the LARFFS peptide and its conjugate induces minimal apoptosis to fibroblasts. Additionally, the degree of apoptosis induced by WNWKV and WNWKV-(PYC)₂ conjugate is much lesser compared to the results seen for the double mutant. Overall, these results indicate the conjugates may have higher selectivity toward the double mutant expressing tumor cells.

Conclusions

In this work, we have designed new purine and pyrimidine derived peptide conjugates and explored their binding interactions with the kinase domain of wild-type EGFR and EGFR T790M/L858R double mutant receptor using molecular docking and molecular dynamics studies. This is the first time where a sea cucumber-derived peptide with antioxidant properties, WNWKV, was utilized and its ability to bind to the kinase domain of these receptors was examined to potentially target over-expressed EGFR receptors in tumor cells. The peptide LARFFS, which had been shown through phage display libraries in previous work to bind to domain I of the EGFR, was shown to have moderate binding toward the kinase domain of the wild-type receptor and poor binding stability with the double mutant. Many of the conjugates and the WNWKV peptide were shown to interact with the activation loop region of the receptors, hinge region, as well as the Gly-rich loop which is promising. The new purine derivatives were designed by preparing bioisosteres of an intermediate utilized in the preparation of the antineoplastic drug clofarabine. Our results indicated that the 5-methyl methyl-pyridyl group side chain showed more stable binding with the wild-type receptor compared to the toluene group when conjugated with WNWKV. The pyrimidine derivatives were prepared by conjugating small molecules pyrimidine 4-carboxylic acid and 2-methylthio pyrimidin-4-amine with the peptides. Molecular dynamics and MMGBSA analysis showed that the binding energies were higher for the double mutant receptor, compared to

NCI-H1975 Cells



A549 cells

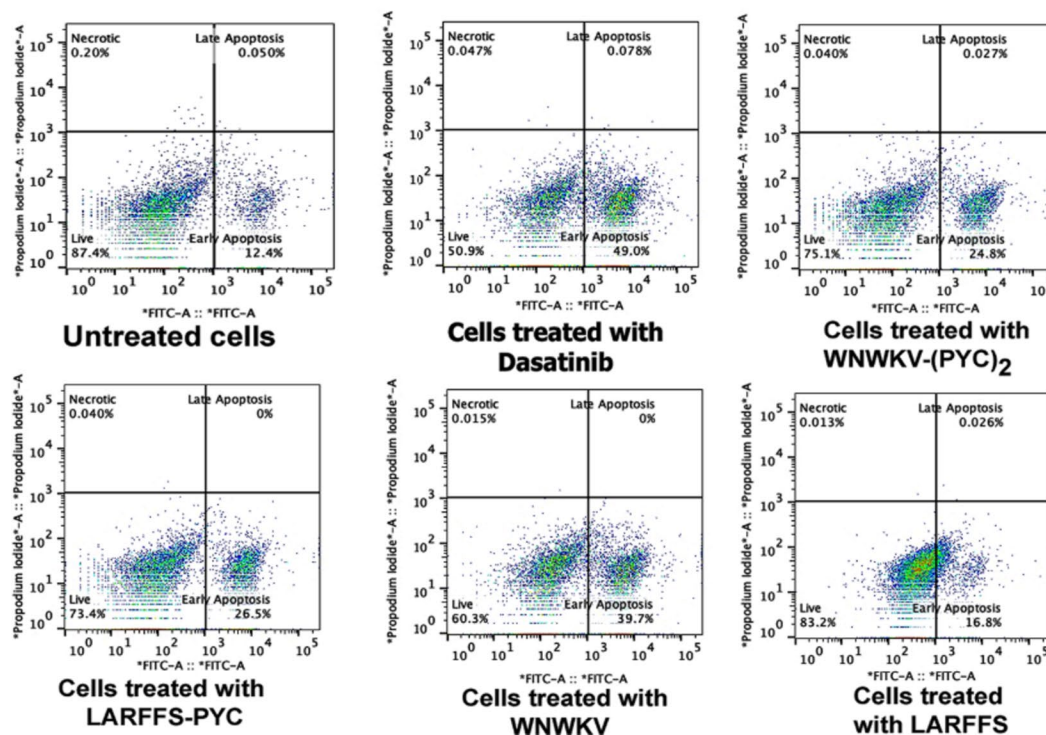


Fig. 8 Annexin V/propidium iodide apoptosis studies using flow cytometry results showing effects of treatment with various constructs after 24 h incubation period both NCI-H1975 lung tumor

cells (expressing T790M/L858R EGFR) and A549 lung cancer cells (expressing wild-type EGFR)

the wild-type. Additionally, the conjugates enhanced binding with the receptor in most cases. Both purine and pyrimidine derivatives showed higher binding upon conjugation. As a proof of concept, the pyrimidine 4-amide conjugates with WNWKV and LARFFS were tested for binding with the receptors using SPR analysis and the results corroborated with computational analysis. In vitro cell studies indicated that the WNWKV-(PYC)₂ conjugate was more potent compared to the LARFFS peptide and its conjugate toward the T790M/L858R EGFR expressing cells. However, the WNWKV neat peptide successfully induces apoptosis in the wild-type cells. In contrast, LARFFS and its PYC conjugates showed little effect on the wild-type EGFR expressing cells. Additionally, most of the conjugates showed very little effect upon treatment with fibroblast non-cancer cells. Overall, these studies reveal the utilization of new peptides and their conjugates through target hopping approach for binding to over-expressed EGFR receptors, particularly the double mutant T790M/L858R, which has been known to cause chemoresistance. The pyrimidine and purine conjugates shown here (particularly the NPU) as well as the WNWKV peptide and WNWKV-(PYC)₂ conjugate may also be considered for future laboratory studies for development of therapeutics against EGFR over-expressed tumor cells.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11030-023-10772-x>.

Acknowledgements HH, BG, MM, and MB thank the Henry Luce foundation for the Clare Boothe Luce Scholarship program for financial support of this work. The authors thank Ms. Chau Anh Phan for her help with running some of the surface plasmon resonance experiments. IB thanks the Fordham University Research Office for their support. The authors acknowledge the Advanced Research Computing, Education Technologies and Research Computing department at Fordham University, a division of the Office of Information Technology for providing their assistance and access to research computing resources that have contributed to the results reported here.

Author contributions HLH: contributed toward in silico modeling, interpretation of data, data curation (both laboratory studies and in silico studies), and writing the first draft of the manuscript. BGG: contributed toward in vitro studies, data analysis, and writing parts of the manuscript; MAB: contributed toward data curation, in silico modeling, and in vitro studies; MIR: contributed toward data curation (cell viability studies); MEM: contributed toward data curation (SPRs); CGL: performed some of the molecular dynamics studies; IAB: contributed toward conceptualization, supervision, data analysis, editing, and writing the final version of the manuscript.

Funding This study was funded by Fordham University Research office. IB thanks NSF-MRI grant # 2117625 for the acquisition of the Fluorescence-Activated Cell Sorter instrument.

Declarations

Competing interest The authors declare no competing interests.

References

- Vagace JM, Gervasini G (2011) Chemotherapy Toxicity in Patients with Acute Leukemia. In: Antica M (Eds), *Acute Leukemia - The Scientist's Perspective and Challenge*. Intech Open. doi:<https://doi.org/10.5772/841>.
- Sharma P, Hu-Lieskovan S, Wargo JA, Ribas A (2017) Primary, adaptive, and acquired resistance to cancer immunotherapy. *Cell* 168:707–723. <https://doi.org/10.1016/j.cell.2017.01.017>
- Alfarouk KO, Stock C-M, Taylor S, Walsh M, Muddathir AK, Verduzco D, Bashir AH, Mohammed OY, Elhassan GO, Harguindey S, Reshkin SJ, Ibrahim ME, Rauch C (2015) Resistance to cancer chemotherapy: failure in drug response from ADME to P-GP. *Cancer Cell Int.* <https://doi.org/10.1186/s12935-015-0221-1>
- Zhu X, Zhu H, Luo H, Zhang W, Shen Z, Hu X, Zhu X (2016) Molecular mechanisms of cisplatin resistance in cervical cancer. *Drug Des Devel Ther* 10:1885–1895. <https://doi.org/10.2147/DDDT.S106412>
- Ma Y, Yu S, Ni S, Zhang B, Kung ACF, Gao J, Lu A, Zhang G (2021) Targeting strategies for enhancing paclitaxel specificity in chemotherapy. *Front Cell Dev Biol.* <https://doi.org/10.3389/fcell.2021.626910>
- Xiao Y-F, Jie M-M, Li B-S, Li B-S, Hu C-J, Xie R, Tang B, Yang S-M (2015) Peptide-based treatment: a promising cancer therapy. *J Immunol Res* 2015:1–13. <https://doi.org/10.1155/2015/761820>
- Feldmann G, Dhara S, Fendrich V, Bedja D, Beaty R, Mullendore M, Karikari C, Alvarez H, Lacobuzio-Donahue C, Jimeno A, Gabrielson KL, Matsui W, Maitra A (2007) Blockade of hedgehog signaling inhibits pancreatic cancer invasion and metastases: a new paradigm for combination therapy in solid cancers. *Cancer Res* 67:2187–2196. <https://doi.org/10.1158/0008-5472.CAN-06-3281>
- Cheetham AG, Keith D, Zhang P, Lin R, Su H, Cui H (2016) Targeting tumors with small molecule peptides. *Curr Drug Targets* 16:489–508. <https://doi.org/10.2174/1568009616666151130214646>
- Akhtar MJ, Ahamed M, Alhadlaq HA, Alrokayan SA, Kumar S (2014) Targeted anticancer therapy: overexpressed receptors and nanotechnology. *Clin Chim Acta* 436:78–92. <https://doi.org/10.1016/j.cca.2014.05.004>
- Murayama O, Nishida H, Sekiguchi K (1996) Novel peptide ligands for Integrin $\alpha 6 \beta 1$ selected from a phage display library. *J Biochem* 120:445–451. <https://doi.org/10.1093/oxfordjournals.jbchem.a021431>
- Chanda N, Kattumuri V, Shukla R, Zambre A, Katti K, Upen-dran A, Kulkarni RR, Kan P, Fent GM, Casteel SW, Smith CJ, Boote E, Robertson JD, Cutler C, Lever JR, Katti KV, Kannan R (2010) Bombesin functionalized gold nanoparticles show in vitro and in vivo cancer receptor specificity. *Proc Natl Acad Sci USA* 107:8760–8765. <https://doi.org/10.1073/pnas.1002143107>
- He R, Finan B, Mayer JP, DiMarchi RD (2019) Peptide conjugates with small molecules designed to enhance efficacy and safety. *Molecules* 24:1855. <https://doi.org/10.3390/molecules24101855>
- Li Y, Zheng X, Gong M, Zhang J (2016) Delivery of a peptide-drug conjugate targeting the blood brain barrier improved the efficacy of paclitaxel against glioma. *Oncotarget* 7:79401–79407. <https://doi.org/10.18632/oncotarget.12708>
- Tesauro D, Accardo A, Aloj L, Aurilio M, Morelli G, Tesauro D (2014) Receptor binding peptides for target-selective delivery of nanoparticles encapsulated drugs. *Int J Nanomed* 9:1537–1557. <https://doi.org/10.2147/IJN.S53593>

15. Sigismund S, Avanzato D, Lanzetti L (2017) Emerging functions of the EGFR in cancer. *Mol Oncol* 12:3–20. <https://doi.org/10.1002/1878-0261.12155>
16. Kim JW, Kim YT, Kim DK, Song CH, Lee JW (1996) Expression of epidermal growth factor receptor in carcinoma of the cervix. *Gynecol Oncol* 60:283–287. <https://doi.org/10.1006/gyno.1996.0039>
17. Rimawi MF, Shetty PB, Weiss HL, Schiff R, Osborne CK, Charness GC, Elledge RM (2010) Epidermal growth factor receptor expression in breast cancer association with biologic phenotype and clinical outcomes. *Cancer* 116:1234–1242. <https://doi.org/10.1002/cncr.24816>
18. Sarabipour S (2017) Parallels and distinctions in FGFR, VEGFR, and EGFR mechanisms of transmembrane signaling. *Biochemistry* 56:3159–3173. <https://doi.org/10.1021/acs.biochem.7b00399>
19. Bethune G, Bethune D, Ridgway N, Xu Z (2010) Epidermal growth factor receptor (EGFR) in lung cancer: an overview and update. *J Thorac Dis* 2:48–51
20. Li AR, Chitale D, Riely GJ, Pao W, Miller VA, Zakowski MF, Rusch V, Kris MG, Ladanyi M (2008) EGFR mutations in lung adenocarcinomas. *J Mol Diagn* 10:242–248. <https://doi.org/10.2353/jmoldx.2008.070178>
21. Toyooka S, Kiura K, Mitsudomi T (2005) EGFR mutation and response of lung cancer to gefitinib. *N Engl J Med* 352:2136–2136. <https://doi.org/10.1056/NEJM200505193522019>
22. Grabe T, Lategahn J, Rauh D (2018) C797S resistance: The undruggable EGFR mutation in non-small cell lung cancer? *ACS Med Chem Lett* 9:779–782. <https://doi.org/10.1021/acsmedchemlett.8b00314>
23. Zhang T, Wan B, Zhao Y, Li C, Liu H, Lv T, Zhan P, Song Y (2019) Treatment of uncommon EGFR mutations in non-small cell lung cancer: new evidence and treatment. *Transl Lung Cancer Res* 8:302–316. <https://doi.org/10.2137/tlcr.2019.04.12>
24. Yan F, Liu X, Zhang S, Su J, Zhang Q, Chen J (2018) Effect of double mutations T790M/L858R on conformation and drug-resistant mechanism of epidermal growth factor receptor explored by molecular dynamics simulations. *RSC Adv* 8:39797–39810. <https://doi.org/10.1039/C8RA06844E>
25. Tamirat MZ, Koivu M, Elenius K, Johnson MS (2019) Structural characterization of EGFR exon 19 deletion mutation using molecular dynamics simulation. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0222814>
26. Suda K, Onozato R, Yatabe Y, Mitsudomi T (2009) EGFR T790M mutation: a double role in lung cancer cell survival? *J Thorac Oncol* 4:1–4. <https://doi.org/10.1097/JTO.0b013e3181913c9f>
27. Yun C-H, Boggon TJ, Li Y, Woo MS, Greulich H, Meyerson M, Eck MJ (2007) Structures of lung cancer-derived EGFR mutants and inhibitor complexes: mechanism of activation and insights into differential inhibitor sensitivity. *Cancer Cell* 11:217–227. <https://doi.org/10.1016/j.ccr.2006.12.017>
28. Thomson RJ, Moshirfar M, Ronquillo Y (2023) Tyrosine kinase inhibitors. StatPearls Publishing. <https://www.ncbi.nlm.nih.gov/books/NBK563322/>. Accessed 10 Jul 2023
29. Zhao Z, Xie L, Bourne PE (2018) Structural insights into characterizing binding sites in epidermal growth factor receptor kinase mutants. *J Chem Inf Model* 59:453–462. <https://doi.org/10.1021/acs.jcim.8b00458>
30. Jänne PA, Yang JC-H, Kim D-W, Planchard D, Ohe Y, Ramalingam SS, Ahn M-J, Kim S-W, Su W-C, Horn L, Haggstrom D, Felip E, Kim J-H, Frewer P, Cantarini M, Brown KH et al (2015) AZD9291 in EGFR inhibitor-resistant non-small-cell lung cancer. *N Engl J Med* 372:1689–1699. <https://doi.org/10.1056/NEJMoa1411817>
31. Singh PK, Silakari O (2018) Molecular dynamics guided development of indole based dual inhibitors of EGFR (T790M) and c-MET. *Bioorg Chem* 79:163–170. <https://doi.org/10.1016/j.bioorg.2018.04.001>
32. Ayati A, Moghimi S, Toolabi M, Foroumadi A (2021) Pyrimidine-based EGFR TK inhibitors in targeted cancer therapy. *Eur J Med Chem* 221:113523. <https://doi.org/10.1016/j.ejmech.2021.113523>
33. Romu AA, Lei Z, Zhou B, Chen Z-S, Korlipara V (2017) Design, synthesis and biological evaluation of WZ4002 analogues as EGFR inhibitors. *Bioorg Med Chem* 27:4832–4837. <https://doi.org/10.1016/j.bmcl.2017.09.048>
34. Sogabe S, Kawakita Y, Igaki S, Iwata H, Miki H, Cary DR, Takagi T, Takagi S, Ohta Y, Ishikawa T (2012) Structure-based approach for the discovery of pyrrolo[3,2-d] pyrimidine-based EGFR T790M/L858R mutant inhibitors. *ACS Med Chem Lett* 4:201–205. <https://doi.org/10.1021/ml300327z>
35. Tiefenbacher A, Pirker R (2017) EGFR tyrosine kinase inhibitors as first-line therapy in advanced EGFR mutation-positive non-small cell lung cancer: Strategies to improve clinical outcome. *J Thorac Dis* 9:4208–4211. <https://doi.org/10.21037/jtd.2017.10.02>
36. Wood ER, Truesdale AT, McDonald OB, Yuan D, Hassell A, Dickerson SH, Ellis B, Pennisi C, Horne E, Lackey K, Alligood KJ, Rusnak DW, Gilmer TM, Shewchuk L (2004) A unique structure for epidermal growth factor receptor bound to GW572016 (lapatinib). *Cancer Res* 64:6652–6659. <https://doi.org/10.1158/0008-5472.CAN-04-1168>
37. Sayed MT, Halim PA, El-Ansary AK, Hassan RA (2023) Design, synthesis, anticancer evaluation, and in silico studies of some thieno [2,3-d]pyrimidine derivatives as EGFR inhibitors. *Drug Dev Res*. <https://doi.org/10.1002/ddr.22088>
38. Furman O, Zaporozhets A, Tobi D, Bazylevich A, Firer MA, Patsenker L, Gellerman G, Lubin BCR (2022) Novel cyclic peptides for targeting EGFR and EGFRVIII mutation for drug delivery. *Pharmaceutics* 14:1505. <https://doi.org/10.3390/pharmaceutics14071505>
39. Gavriil E-S, Doukatas A, Karampelas T, Myrianthopoulos V, Dimitrakis S, Mikros E, Marakos P, Tamvakopoulos C, Pouli N (2019) Design, synthesis and biological evaluation of novel substituted purine isosteres as EGFR kinase inhibitors, with promising pharmacokinetic profile and in vivo efficacy. *Eur J Med Chem* 176:393–409. <https://doi.org/10.1016/j.ejmech.2019.05.029>
40. Zhou H, Fu H, Liu H, Shao X, Cai W (2022) Uncovering the mechanism of drug resistance caused by the T790M mutation in EGFR kinase from absolute binding free energy calculations. *Front Mol Biosci*. <https://doi.org/10.3389/fmolb.2022.922839>
41. Zhang Y, He S, Bonneil É, Simpson BK (2020) Generation of antioxidative peptides from Atlantic sea cucumber using alcalase versus trypsin: in vitro activity, de novo sequencing, and in silico docking for in vivo function prediction. *Food Chem* 306:125581. <https://doi.org/10.1016/j.foodchem.2019.125581>
42. Yaghoubzadeh Z, Peyravii Ghadikolaii F, Kaboosi H, Safari R, Fattahi E (2019) Antioxidant activity and anticancer effect of bioactive peptides from rainbow trout (*Oncorhynchus mykiss*) skin hydrolysate. *Int J Pept Res* 26:625–632. <https://doi.org/10.1007/s10989-019-09869-5>
43. Song S, Liu D, Peng J, Deng H, Guo Y, Xu LX, Miller AD, Xu Y (2009) Novel peptide ligand directs liposomes toward EGF-R high-expressing cancer cells in vitro and in vivo. *FASEB J* 23:1396–1404. <https://doi.org/10.1096/fj.08-117002>
44. Tognolini M, Incerti M, Pala D, Russo S, Castelli R, Hassan-Mohamed I, Giorgio C, Lodola A (2013) Target hopping as a useful tool for the identification of novel EPHA2 protein-protein antagonists. *Chem Med Chem* 9:67–72. <https://doi.org/10.1002/cmdc.201300305>

45. Ghanem H, Kantarjian H, Ohanian M, Jabbour E (2012) The role of clofarabine in acute myeloid leukemia. *Leuk Lymphoma* 54:688–698. <https://doi.org/10.3109/10428194.2012.726722>
46. Daly MB, Roth ME, Bonnac L, Maldonado JO, Xie J, Clouser CL, Patterson SE, Kim B, Mansky LM (2016) Dual anti-HIV mechanism of clofarabine. *Retrovirology* 13:20. <https://doi.org/10.1186/s12977-016-0254-0>
47. Karmacharya U, Guragain D, Chaudhary P, Jee J-G, Kim J-A, Jeong B-S (2021) Novel pyridine Bioisostere of cabozantinib as a potent C-met kinase inhibitor: synthesis and anti-tumor activity against hepatocellular carcinoma. *Int J Mol Sci* 22:9685. <https://doi.org/10.3390/ijms22189685>
48. Xiao Q, Qu R, Gao D, Yan Q, Tong L, Zhang W, Ding J, Xie H, Li Y (2016) Discovery of 5-(methylthio) pyrimidine derivatives as L858R/T790M mutant selective epidermal growth factor receptor (EGFR) inhibitors. *Bioorg Med Chem* 24:2673–2680. <https://doi.org/10.1016/j.bmc.2016.04.032>
49. Tyagi A, Kapoor P, Kumar R, Chaudhary K, Gautam A, Raghava GPS (2013) In silico models for designing and discovering novel anticancer peptides. *Sci Rep* 3:2984. <https://doi.org/10.1038/srep02984>
50. Agarwal P, Bhagat D, Mahalwal M, Sharma N, Raghava GPS (2021) AntiCP 2.0: an updated model for predicting anticancer peptides. *Brief Bioinform*. <https://doi.org/10.1093/bib/bbaa153>
51. Schrödinger L, DeLano W (2020) PyMOL. Accessed <http://www.pymol.org/pymol>.
52. Balasubramani SG, Chen GP, Coriani S, Diedenhofen M, Frank MS, Franzke YJ, Furche F, Grotjahn R, Harding ME, Hattig C, Hellweg A, Helmich-Paris B, Holzer C, Hunair U, Kaupp M et al (2020) TURBOMOLE: modular program suite for ab initio quantum-chemical and condensed-matter simulations. *J Chem Phys* 152:184107. <https://doi.org/10.1063/5.0004635>
53. Mullins E, Oldland R, Liu YA, Wang S, Sandler SI, Chen C-C, Zwolak M, Seavey KC (2006) Sigma-profile database for using cosmo-based thermodynamic methods. *Ind Eng Chem Res* 45:4389–4415. <https://doi.org/10.1021/ie060370h>
54. Peng Y-H, Shiao H-Y, Tu C-H, Liu P-M, Hsu J, Amancha PK, Wu J-S, Coumar MS, Chen C-H, Wang S-Y, Lin W-H, Sun H-Y, Chao Y-S, Lyu P-C, Hsieh H-P, Wu S-Y (2013) Protein kinase inhibitor design by targeting the asp-phe-gly (DFG) motif: the role of the DFG motif in the design of epidermal growth factor receptor inhibitors. *J Med Chem* 56:3889–3903. <https://doi.org/10.1021/jm400072p>
55. Engelhardt H, Böse D, Petronczki M, Scharn D, Bader G, Baum A, Bergner A, Chong E, Döbel S, Egger G, Engelhardt C, Ettmayer P, Fuchs JE, Gerstberger T, Gonnella N, Grimm A, Grondal E, Haddad N, Hopfgartner B, Kousek R et al (2019) Start selective and rigidify: the discovery path toward a next generation of EGFR tyrosine kinase inhibitors. *J Med Chem* 62:10272–10293. <https://doi.org/10.1021/acs.jmedchem.9b01169>
56. Zardecki C, Dutta S, Goodsell DS, Lowe R, Voigt M, Burley SK (2021) PDB-101: educational resources supporting molecular explorations through biology and medicine. *Protein Sci* 31:129–140. <https://doi.org/10.1002/pro.4200>
57. Yu J, Zhou Y, Tanaka I, Yao M (2009) Roll: a new algorithm for the detection of protein pockets and cavities with a rolling probe sphere. *Bioinformatics* 26:46–52. <https://doi.org/10.1093/bioinformatics/btp599>
58. Halgren TA (2009) Identifying and characterizing binding sites and assessing druggability. *J Chem Inf Model* 49:377–389. <https://doi.org/10.1021/ci800324m>
59. Halgren TA (2007) New method for fast and accurate binding-site identification and analysis. *Chem Biol Drug Des* 69:146–148. <https://doi.org/10.1111/j.1747-0285.2007.00483.x>
60. Harris R, Olson AJ, Goodsell DS (2008) Automated prediction of ligand-binding sites in proteins. *Proteins* 70:1506–1517. <https://doi.org/10.1002/prot.21645>
61. Eberhardt J, Santos-Martins D, Tillack AF, Forli S (2021) AutoDock Vina 1.2.0: new docking methods, expanded force field, and python bindings. *J Chem Inf Model* 61:3891–3898. <https://doi.org/10.1021/acs.jcim.1c00203>
62. dos Santos KB, Guedes IA, Karl ALM, Dardenne L (2020) Highly Flexible Ligand Docking: benchmarking of the DockThor program on the LEADS-PEP protein-peptide dataset. *J Chem Inf Model* 60:667–683. <https://doi.org/10.1021/acs.jcim.9b00905>
63. Adasme MF, Linnemann KL, Bolz SN, Kaiser F, Salentin S, Haupt VJ, Schroeder M (2021) PLIP 2021: expanding the scope of the protein–ligand interaction profiler to DNA and RNA. *Nucleic Acids Res* 49:W530–W534. <https://doi.org/10.1093/nar/gkab294>
64. Bowers K, Chow E, Xu H, Dror RO, Eastwood MP, Gregersen BA, Klepeis JL, Kolossvary I, Moraes MA, Sacerdoti FD, Salmon JK, Shan Y, Shaw DE (2006) Scalable algorithms for molecular dynamics simulations on commodity clusters SC06: Proceedings of the 2006 ACM/IEEE conference on Supercomputing. <https://doi.org/10.1145/1188455.1188544>
65. Lu C, Wu C, Ghoreishi D, Chen W, Wang L, Damm W, Ross GA, Dahlgren MK, Russell E, Von Bargen CD, Abel R, Friesner RA, Harder ED (2021) OPLS4: Improving force field accuracy on challenging regimes of chemical space. *J Chem Theory Comput* 17:4291–4300. <https://doi.org/10.1021/acs.jctc.1c00302>
66. Jacobson MP, Pincus DL, Rapp CS, Day T, Honig B, Shaw DE (2004) A hierarchical approach to all-atom protein loop prediction. *Proteins Struct Funct* 55:351–367. <https://doi.org/10.1002/prot.10613>
67. Pattar SV, Adhoni SA, Kamanavalli CM, Kumbar SS (2020) In silico molecular docking studies and MM/GBSA analysis of coumarin-carbonodithioate hybrid derivatives divulge the anti-cancer potential against breast cancer. *Beni-Suef Univ J Basic Appl Sci* 9:36. <https://doi.org/10.1186/s43088-020-00059-7>
68. Li J, Abel R, Zhu K, Cao Y, Zhao S, Friesner RA (2011) The VSGB 2.0 model: a next generation energy model for high resolution protein structure modeling. *Proteins* 79:2794–2812. <https://doi.org/10.1002/prot.23106>
69. Xiong G, Wu Z, Yi J, Fu L, Yang Z, Hsieh C, Yin M, Zeng X, Wu C, Lu A, Chen X, Hou T, Cao D (2021) ADMETlab2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties. *Nucleic Acids Res* 49(W1):W5–W14. <https://doi.org/10.1093/nar/gkab255>
70. Nam K, Kimura T, Kishida A (2008) Controlling coupling reaction of EDC and NHS for preparation of collagen gels using ethanol/water co-solvents. *Macromol Biosci* 8:32–37. <https://doi.org/10.1002/mabi.200700206>
71. Yang HM, Teoh JY, Yim GK, Park Y, Kim YG, Kim J, Yoo D (2020) Label-free analysis of multivalent protein binding using bioresponsive nanogels and surface plasmon resonance (SPR). *ACS Appl Mater Interfaces* 12:5413–5419. <https://doi.org/10.1021/acsami.9b17328>
72. Harpaz D, Koh B, Marks RS, Seet RCS, Abdulalim I, Tok AIY (2019) Point-of-care surface plasmon resonance biosensor for stroke biomarkers NT-proBNP and S100β using a functionalized gold chip with specific antibody. *Sensors* 19:2533. <https://doi.org/10.3390/s19112533>
73. Minakata K, Takahashi F, Nara T, Hashimoto M, Tajima K, Murakami A, Nurwidya F, Yae S, Koizumi F, Moriyama H, Seyama K, Nishio K, Takahashi K (2012) Hypoxia induces gefitinib resistance in non-small-cell lung cancer with both mutant and wild-type epidermal growth factor receptors. *Cancer Sci* 103:1946–1954. <https://doi.org/10.1111/j.1349-7006.2012.02408.x>

74. Ngamwongsatit P, Banada PP, Panbangred W, Bhunia AK (2008) WST-1 based cell cytotoxicity assay as a substitute for MTT-based assay for rapid detection of toxigenic bacillus species using CHO cell line. *J Microbiol Methods* 73:211–215. <https://doi.org/10.1016/j.mimet.2008.03.002>
75. Chan A, Reiter R, Wiese S, Fertig G, Gold R (1998) Plasma membrane phospholipid asymmetry precedes DNA fragmentation in different apoptotic cell models. *Histochem* 110:553–558. <https://doi.org/10.1007/s004180050317>
76. Lakshmanan I, Batra S (2013) Protocol for apoptosis assay by flow cytometry using annexin v staining method. *Bio protoc* 3:374. <https://doi.org/10.21769/bioprotoc.374>
77. Rieger AM, Nelson KL, Konowalchuk JD, Barreda DR (2011) Modified annexin V/propidium iodide apoptosis assay for accurate assessment of cell death. *J Vis Exp*. <https://doi.org/10.3791/2597>
78. Catitti G, De Fabritiis S, Brocco D, Simeone P, De Bellis D, Vespa S, Veschi S, Lellis LD, Tinari N, Verginelli F, Marchisio M, Cama A, Patruno A, Lanuti P (2022) Flow cytometry detection of anthracycline-treated breast cancer cells: an optimized protocol. *Curr Issues Mol Biol* 45:164–174. <https://doi.org/10.3390/cimb45010013>
79. Klamt A (2011) The cosmo and cosmo-RS solvation models. *Wiley Interdiscip Rev Comput Mol* 1:699–709. <https://doi.org/10.1002/wcms.56>
80. Aissou M, Chemat-Djenni Z, Yara-Varón E, Fabiano-Tixier A-S, Chemat F (2017) Limonene as an agro-chemical building block for the synthesis and extraction of bioactive compounds. *C R Chim* 20:346–358. <https://doi.org/10.1016/j.crci.2016.05.018>
81. Modro AM, Modro TA (1999) The phosphoryl and the carbonyl group as hydrogen bond acceptors. *Can J Chem* 77:890–894
82. Osifová Z, Šála M, Dračinský M (2023) Hydrogen-bonding interactions of 8-substituted purine derivatives. *ACS Omega*. <https://doi.org/10.1021/acsomega.3c03244>
83. Milic M, Targos K, Tellez Chavez M, Thompson MAM, Jennings JJ, Franz AK (2021) NMR quantification of hydrogen-bond-accepting ability for organic molecules. *J Org Chem* 86:6031–6043. <https://doi.org/10.1021/acs.joc.0c02876>
84. Paul A, Thomas R (2022) Evidences for sulfur centered hydrogen bond with sulfur atoms as a donor in aromatic thiols and aliphatic thiols in aqueous solution. *J Mol Liq* 348:118078. <https://doi.org/10.1016/j.molliq.2021.118078>
85. Shan Y, Eastwood MP, Zhang X, Kim ET, Arkhipov A, Dror RO, Jumper J, Kuriyan J, Shaw DE (2012) Oncogenic mutations counteract intrinsic disorder in the EGFR kinase and promote receptor dimerization. *Cell* 149:860–870. <https://doi.org/10.1016/j.cell.2012.02.063>
86. Kaufman NE, Dhingra S, Jois SD, da Vicente M (2021) Molecular targeting of epidermal growth factor receptor (EGFR) and vascular endothelial growth factor receptor (VEGFR). *Molecules* 26:1076. <https://doi.org/10.3390/molecules26041076>
87. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 31:455–461. <https://doi.org/10.1002/jcc.21334>
88. de Magalhães CS, Almeida DM, Barbosa HJC, Dardenne LE (2014) A dynamic niching genetic algorithm strategy for docking highly flexible ligands. *Information Sci* 289:206–224. <https://doi.org/10.1016/j.ins.2014.08.002>
89. de Magalhães CS, Barbosa HJC, Dardenne LE (2004) A Genetic algorithm for the ligand-protein docking problem. *Genetics Mol Biol* 27:605–610
90. Guedes IA, Barreto AMS, Marinho D, Kremser E, Kuenemann MA, Sperandio O, Dardenne LE, Miteva MA (2021) New machine learning and physics-based scoring functions for drug discovery. *Sci Rep* 11:3198
91. Treiber DK, Shah NP (2013) Ins and Outs of kinase DFG motifs. *Chem Biol* 20:745–746. <https://doi.org/10.1016/j.chembiol.2013.06.001>
92. Al-Warhi T, El Kerdawy AM, Said MA, Albohy A, Elsayed ZM, Aljaeed N, Elkaeed EB, Eldehna WM, Abdel-Aziz HA, Abdelmoaz MA (2022) Novel 2-(5-Aryl-4,5-dihydropyrazol-1-yl) thiazol-4-one as EGFR inhibitors: synthesis, biological assessment and molecular docking insights. *Drug Des Devel Ther* 16:1457–1471. <https://doi.org/10.2147/DDDT.S356988>
93. Stamos J, Sliwkowski MX, Eigenbrot C (2002) Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. *J Biol Chem* 277:46265–46272. <https://doi.org/10.1074/jbc.M207135200>
94. Kumar A, Petri ET, Halmos B, Boggon TJ (2008) Structure and clinical relevance of the epidermal growth factor receptor in human cancer. *J Clin Oncol* 26:1742–1751. <https://doi.org/10.1200/JCO.2007.12.1178>
95. Ezelarab H, Ali T, Abbas S, Sayed AM, Beshr EA, Hassan HA (2023) New antiproliferative 3-substituted oxindoles inhibiting EGFR/VEGFR-2 and tubulin polymerization. *Mol Divers*. <https://doi.org/10.1007/s11030-023-10603-z>
96. Nawaz F, Alam O, Perwez A, Rizvi MA, Naim MJ, Siddiqui N, Pottoo FH, Jha M (2020) 3'-(4-(Benzyloxy)phenyl)-1'-phenyl-5-(heteroaryl/aryl)-3,4-dihydro-1'H,2H-[3,4'-bipyrazole]-2-carboxamides as EGFR kinase inhibitors: synthesis, anticancer evaluation, and molecular docking studies. *Arch Pharm* 353:1900262. <https://doi.org/10.1002/ardp.201900262>
97. Ohkura K, Tabata A, Uto Y, Hori H (2020) Effect of isomerization of TX-2036 derivatives on the interaction with tyrosine kinase domain of EGF receptor. *Anticancer Res* 40:4675–4680. <https://doi.org/10.21873/anticancer.14466>
98. Jura N, Shan Y, Cao X, Shaw DE, Kuriyan J (2009) Structural analysis of the catalytically inactive kinase domain of the human EGF receptor 3. *Proc Natl Acad Sci USA* 106:21608–21613. <https://doi.org/10.1073/pnas.0912101106>
99. Mirza Z, Schulten H-J, Farsi HM, Al-Maghrabi JA, Gari MA, Chaudhary AG, Abuzenadah AM, Al-Qahtani MH, Karim S (2015) Molecular interaction of a kinase inhibitor midostaurin with anticancer drug targets, S100A8 and EGFR: transcriptional profiling and molecular docking study for kidney cancer therapeutics. *PLoS ONE* 10:1–17. <https://doi.org/10.1371/journal.pone.0119765>
100. Zare S, Emami L, Faghih Z, Zargari F, Zeinab F, Khabnadideh S (2023) Design, synthesis, computational study and cytotoxic evaluation of some new quinazoline derivatives containing pyrimidine moiety. *Sci Rep* 13:14461. <https://doi.org/10.1038/s41598-023-41530-6>
101. Chan C, Gill GN (1996) Mutational analysis of the nucleotide binding site of the epidermal growth factor receptor and V-src protein-tyrosine kinases. *J Biol Chem* 271:22619–22623. <https://doi.org/10.1074/jbc.271.37.22619>
102. Yaqeen Alhaqq FG, Mahdi M, Dawood A (2021) Theoretical drug design, molecular docking and ADME study of new 1,3,4-oxadiazole derivatives: promising anticancer agents against both breast and lung cancers. *Egypt J Chem* 64:6269–6283. <https://doi.org/10.21608/ejchem.2021.75663.3735>
103. Lemmon MA, Schlessinger J, Ferguson KM (2014) The EGFR family: not so prototypical receptor tyrosine kinases. *Cold Spring Harb perspect biol* 6:a020768. <https://doi.org/10.1101/cshpe rspect.a020768>
104. Mirtavoos-mahyari H, Rismani E, Lotfabadi AS, Dezfouli AA, Sheikhy K, Dezfouli MM, Hesmatania J (2021) Primary erlotinib resistance in a patient with non-small cell lung cancer carrying simultaneous compound EGFR L718A, Q849H, and L858R

- mutations. *Biomol Concepts* 12:164–174. <https://doi.org/10.1515/bmc-2021-0018>
105. Ercan D, Choi HG, Yun C-H, Capelletti M, Xie T, Eck MJ, Gray NS, Jänne PA (2015) EGFR mutations and resistance to irreversible pyrimidine-based EGFR inhibitors. *Clin Cancer Res* 21:3913–3923. <https://doi.org/10.1158/1078-0432.CCR-14-2789>
 106. Pawara R, Ahmad I, Surana S, Patel H (2021) Computational identification of 2,4-disubstituted amino-pyrimidines as L858R/T790M-EGFR double mutant inhibitors using pharmacophore mapping, molecular docking, binding free energy calculation, DFT study and molecular dynamic simulation. *In Silico Pharmacol* 9:54. <https://doi.org/10.1007/s40203-021-00113-x>
 107. Egieyeh S, Egieyeh E, Malan S, Christofells A, Fielding B (2021) Computational drug repurposing strategy predicted peptide-based drugs that can potentially inhibit the interaction of SARS-COV-2 spike protein with its target (human ACE2). *PLoS ONE* 16(1):e0245258. <https://doi.org/10.1371/journal.pone.0245258>
 108. Benson NC, Faggett V (2008) Dymeomics: large-scale assessment of native protein flexibility. *Protein Sci* 17:2038–2050. <https://doi.org/10.1110/ps.037473.108>
 109. Kalita MM, Fischer WB (2016) Asymmetric dynamics of ion channel forming proteins—hepatitis C virus (HCV) P7 bundles. *Biochim Biophys Acta* 1858:1462–1470. <https://doi.org/10.1016/j.bbame.2016.04.004>
 110. Robichaux JP, Le X, Vijayan RSK, Hicks JK, Heeke S, Elamin YY, Lin HY, Udagawa H, Skoulidis F, Tran H, Vargese S, He J, Zhang F, Nilsson MB, Hu L, Poteete A, Rinsurongkawong W, Zhang X, Ren C, Liu X, Hong L, Raymond V, Fang B, Wang J, Ha MJ, Cross JB, Gray JE, Heymach JV (2021) Structure-based classification predicts drug response in EGFR-mutant NSCLC. *Nature* 597:732–737. <https://doi.org/10.1038/s41586-021-03898-1>
 111. Sharma J, Bhardwaj VK, Singh R, Rajendran V, Purohit R, Kumar S (2021) An in-silico evaluation of different bioactive molecules of tea for their inhibition potency against non structural protein-15 of SARS-COV-2. *Food Chem* 346:128933. <https://doi.org/10.1016/j.foodchem.2020.128933>
 112. Moonrin N, Songtawee N, Rattanabunyong S, Chunsriviro S, Mokmak W, Tonsima S, Choowongkamon K (2015) Understanding the molecular basis of EGFR kinase domain/ MIG-6 peptide recognition complex using computational analysis. *BMC Bioinformatics* 16:103. <https://doi.org/10.1186/s12859-015-0528-x>
 113. Jura N, Shan Y, Cao X, Shaw D, Kuriyan J (2009) Structural analysis of the catalytically inactive kinase domain of the human EGF receptor 3. *PNAS USA* 106:21608–21613. <https://doi.org/10.1073/pnas.0912101106>
 114. Kumar CK, Das S, Subramanian PT, Murali P, Isaac AE, Ram-anathan K, Balamurli MM, Chanda K (2022) In-silico molecular modelling, MM/GBSA binding free energy and molecular dynamics simulation study of novel Pyrdo fused imidazo [4,5-c] quinolines as potential anti-tumor agents. *Front Chem*. <https://doi.org/10.3389/fchem.2022.991369>
 115. Onufriev AV, Case DA (2019) Generalized Born Implicit Models for Biomolecules. *Annu Rev Biophys* 48:275–296. <https://doi.org/10.1146/annurev-biophys-052118-115325>
 116. Avdeef A (2001) Physicochemical profiling (solubility, permeability and charge state). *Curr Med Chem* 1:277–351. <https://doi.org/10.2174/1568026013395100>
 117. Benet LZ, Hosey CM, Ursu O, Oprea TI (2016) BDDCS, the rule of 5 and drugability. *Adv Drug Deliv Rev* 101:89–98. <https://doi.org/10.1016/j.addr.2016.05.007>
 118. Mitcheson JS (2008) HERG potassium channels and the structural basis of drug-induced arrhythmias. *Chem Res Toxicol* 21:1005–1010. <https://doi.org/10.1021/tx800035b>
 119. Vijay U, Gupta S, Mathur P, Suravajhala P, Bhatnagar P (2018) Microbial mutagenicity assay: Ames test. *Bio protoc* 8:e2763. <https://doi.org/10.21769/BioProtoc.2763>
 120. Irvine JD, Takahashi L, Lockhart K, Cheong J, Tolan JW, Selick HE (1999) MDCK (Madin-Darby canine kidney) cells: a tool for membrane permeability screening. *J Pharm Sci* 88:28–33. <https://doi.org/10.1021/js9803205>
 121. Sharom FJ (2011) The P-glycoprotein multidrug transporter. *Essays in Biochem* 50:161–178. <https://doi.org/10.1042/bse0500161>
 122. Montanari F, Ecker GF (2015) Prediction of drug–ABC-transporter interaction—recent advances and future challenges. *Adv Drug Deliv Rev* 86:17–26. <https://doi.org/10.1016/j.addr.2015.03.001>
 123. McDonnell AM, Dang CH (2013) Basic review of the cytochrome P450 system. *J Adv Pract Oncol* 4:263–268. <https://doi.org/10.6004/jadpro.2013.4.4.7>
 124. Guttman Y, Kerem Z (2022) Dietary inhibitors of CYP3A4 are revealed using virtual screening by using a new deep-learning classifier. *J Agric Food Chem* 70:2752–2761. <https://doi.org/10.1021/acs.jafc.2c00237>
 125. Zhou Y, Nevošadová L, Eliasson E, Lausckhe VM (2023) Global distribution of functionally important CYP2C9 alleles and their inferred metabolic consequences. *Hum Genomics* 17:15. <https://doi.org/10.1186/s40246-023-00461-z>
 126. Sordella R, Bell DW, Haber DA, Settleman J (2004) Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. *Science* 305:1163–1167. <https://doi.org/10.1126/1535-7163.MCT-05-0325>
 127. Zhang H, Shao H, Golubovskaya VM, Chen H, Cance W, Adjei AA, Dy GK (2016) Efficacy of focal adhesion kinase inhibition in non-small cell lung cancer with oncogenically activated MAPK PATHWAYS. *Br J Cancer* 115:203–211. <https://doi.org/10.1038/bjc.2016.190>
 128. Ichim TE, O’Heeron P, Kesari S (2018) Fibroblasts as a practical alternative to mesenchymal stem cells. *J Translational Med* 16:212. <https://doi.org/10.1186/s12967-018-1536-1>
 129. Virakul S, Dalm VAS, Paridaens D, van den Bosch W, Hirankarn N, van Hagen P, Dik W (2014) The tyrosine kinase inhibitor dasatinib effectively blocks PDGF-induced orbital fibroblast activation. *Graefes Arch Clin Exp Ophthalmol* 252:1101–1109. <https://doi.org/10.1007/s00417-014-2674-7>
 130. Wang M, Chi CA (2018) Molecular mechanism of action and potential biomarkers of growth inhibition of synergistic combination of afatinib and dasatinib against gefitinib-resistant non-small cell lung cancer cells. *Oncotarget* 9:16533–16546. <https://doi.org/10.18632/oncotarget.24814>
 131. Wang X, Xue Q, Wu L, Wang B, Liang H (2018) Dasatinib promotes TRAIL mediated apoptosis by upregulating CHOP-dependent death receptor 5 in gastric cancer. *FEBS Open Bio* 8:732–742. <https://doi.org/10.1002/2211-5463.12404>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.