Research Article

Open Access

Hanifeh Mirtavoos-mahyari*, Elham Rismani, Alireza Sarkar Lotfabadi, Azizollah Abbasi Dezfuli, Kambiz Sheikhy, Mojtaba Mokhber Dezfuli, Jalal Heshmatnia

Primary Erlotinib Resistance in a Patient with Non-Small Cell Lung Cancer Carrying Simultaneous Compound EGFR L718A, Q849H, and L858R Mutations

<https://doi.org/10.1515/bmc-2021-0018>

received November 2, 2021; accepted December 3, 2021.

Abstract: Nowadays, mutations in the epidermal growth factor receptor (EGFR) kinase domain are studied in targeted therapy of non-small cell lung cancer (NSCLC) with EGFR tyrosine kinase inhibitors including gefitinib and erlotinib. The present study reports a rare case of a patient harboring three simultaneous EGFR mutations (L718A, Q849H, and L858R). The development of erlotinib resistance was detected in the subsequent treatment. Using a computational approach, the current study investigated the conformational changes of wild-type and mutant EGFR's kinase domains in the interaction with erlotinib. Their binding modes with erlotinib were elucidated during molecular dynamics simulation, where higher fluctuations were detected in the mutated forms of the EGFR tyrosine kinase domain. Prediction of stability and functional effect of mutations revealed that amino acidic substitutions have decreased the protein stability as well as the binding affinity to erlotinib. These results may be useful for a recommendation of EGFR mutational analysis for patients with NSCLC carcinoma.

Keywords: EGFR; erlotinib; molecular dynamics; non-small cell lung cancer.

Introduction

One of the leading causes of cancer-related mortality worldwide is lung cancer, with an overall five-year survival rate of 15% [1]. Approximately 85% of all lung cancers include Non-small cell lung carcinoma (NSCLC) [2]. Significant advances in overall survival (OS) and progression-free survival (PFS) in NSCLC patients were recently achieved by targeted therapies and genetic alterations detection.

Several genetic alterations in genes such as Kirsten ras (KRAS), anaplastic lymphoma kinase (ALK), and repressor of silencing 1 (ROS1), in NSCLC have been studied, but the most important mutations happen in epidermal growth factor receptor (EGFR) [3]. EGFR is a transmembrane protein that has a tyrosine kinase domain located at exon 18-21, where EGFR mutations occur. EGFR somatic mutation leads to the constant activation of the receptor and causes uncontrolled cell division and proliferation, tumor progression, and tumoral cell migration (metastasis). Single-nucleotide polymorphisms (SNPs) in the EGFR gene are associated closely with the response of cells to the 4-anilinoquinazoline (AQ4) based reversible EGFR tyrosine kinase inhibitors (TKIs), including erlotinib and gefitinib [4]. The classic or common EGFR mutations include L858R substitution in exon 21, exon 20 insertions, and in-frame deletions in exon 19, [5]. Besides the frequent EGFR mutations, many other types of EGFR gene mutations occur in exons 18 to 21, but the frequency of such mutations is relatively low with unclear clinical significance.

*Corresponding author: Hanifeh Mirtavoos-mahyari, Lung Transplantation Research Center, National Research Institute of Tuberculosis and Lung Diseases (NRITLD), Shahid Beheshti University of Medical Sciences, Tehran, Iran, E-mail: hanifehmirmah@yahoo.com
Elham Rismani, Molecular Medicine Department, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran
Alireza Sarkar Lotfabadi, Department of Bioengineering and Biotechnology, Faculty of Life Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran
Azizollah Abbasi Dezfuli, Kambiz Sheikhy, Mojtaba Mokhber Dezfuli, Jalal Heshmatnia, Lung Transplantation Research Center, National Research Institute of Tuberculosis and Lung Diseases (NRITLD), Shahid Beheshti University of Medical Sciences, Tehran, Iran

It is essential to understand how protein mutations alter the main functions of proteins for obtaining a suitable perception of various medical issues, particularly in oncology. Changes in proteins due to mutations can affect diverse protein properties, such as stability, catalytic activity, or the ability to interact with other molecules.

This study presents a rare case of lung cancer that harbors three simultaneous EGFR substitution mutations, which was identified subsequent to the administration of erlotinib treatment to the patient. Herein, this study reports the structural effect of three mutations (L718A, Q849H and L858R) located within or near the interacting region of EGFR to erlotinib. In this work, we assess whether these substitutions can directly influence the protein structure stability and efficiency of protein interactions with drugs.

Case Presentation and Methods

Case Presentation

A 41-year-old man was referred to the hospital in November 2018 after a lung nodule was detected on chest computed tomography (CT) during a physical examination, and no particular family medical history was reported for this patient. The patient was presented with progressive cough and sputum, no history of alcohol intake and cigarette. On primary examination, the patient was cachectic, tachypneic (respiratory rate 24 times/min), and saturating index of oxygen was at 96% on ambient room air. Other vital signs were within normal range. On percussion of chest and auscultation breath, sounds were normal. Supraclavicular lymph nodes were impalpable but 2 axillary lymph nodes were detectable. Other cardiovascular, abdominal, and neurological examinations were noncontributory. The patient was operated and the acinar predominant adenocarcinoma (pathologic tumor-node-metastasis stage: T3N2M0) was diagnosed by a postoperative pathological examination. EGFR mutations in extracted DNA of tumor tissue biopsy were evaluated by nested-PCR followed by sequencing with ABI 3500xl DNA Sequencer. The nested-PCR was performed using primers of exons 18-21 of the EGFR gene. It was an in-house method for PCR assay, and total reaction volume was 25 μ L include 12.5 μ L master mix (Ampliqon), 0.5 μ L for each forward and reverse primer, 6.5 μ L dH₂O, and 5 μ L DNA sample. The denaturation stage of PCR was performed once at 94°C for 3 minutes. There were 35 cycles for the annealing stage at 94°C, 60°C, and 72°C temperatures with 30 seconds for each.

The extension stage was also performed once at 72°C for 5 minutes. Finally, the novelty of the mutations was evaluated in dbSNP and the literature. Patient started the treatment with the erlotinib 150 mg daily. The patient did not respond to erlotinib and progressed rapidly after 6 weeks of treatment in CT scan, so the case was considered as primary resistance, and treatment with erlotinib was stopped (after 6 weeks). According to the patient's poor performance status, second line chemotherapy was not started, and the patient died after 2 months.

Informed consent: Informed consent has been obtained from the patient included in this study.

Ethical approval: The research related to human use has been complied with all the relevant national regulations, institutional policies and in accordance the tenets of the Helsinki Declaration, and has been approved by the Shahid Beheshti Medical University's ethics and scientific committees (IR.SBMU.NRITLD. REC.1399.200).

Protein Sequence and Modeling

The sequence of epidermal growth factor receptor (EGFR) was extracted from the UniprotKB database (ID: P00533) and blasted against Protein Data Bank (PDB) using NCBI Blastp to find homologous proteins.

Investigating the Functional Consequence of SNPs by Protein Variation Effect Analyzer (PROVEAN)

PROVEAN (<http://provean.jcvi.org/index.php>) predicts whether an amino acid substitution or indel affects the phenotypic and biological function of a protein. By recruiting the sequence homology-based method and clustering of BLAST hits with a parameter of 75% global sequence identity, PROVEAN introduces a delta alignment score that finally generates the PROVEAN score. The PROVEAN score equal to or below a predefined threshold (e.g., -2.5) predicts a "deleterious" effect of the protein variant, whereas the score above the threshold indicates a "neutral" effect of the variant on the protein function [6]. The simultaneous identified variants of L718A, Q849H, and L858R were presented as input to determine the amino acid substitutions consequences on the functional alterations of EGFR.

Investigating the Functional Impacts of SNPs by Nonacceptable Polymorphisms Screening

SNAP2 (<https://roslab.org/services/snap2web/>) server was used to find the functional effects of the above-mentioned mutations on EGFR. The prediction done by SNAP2 is formed on the information of secondary structure, multiple sequence alignment, and solvent accessibility of the protein sequence using the neural network method [7]. The output, including the expected accuracy, score (ranges from -100 strong neutral prediction to +100 strong effect prediction), and, prediction (Effect or neutral) are summarized in Table 1.

Prediction of Protein Stability by I-Mutant 2.0

To predict the protein stability changes of EGFR upon the above-mentioned mutations, I-Mutant 2.0 (<http://folding.biofold.org/i-mutant/i-mutant2.0.html>) was used as a support vector machine-based (SVM) tool. The EGFR sequence, temperature (37°C), and pH [7] were defined as input parameters. The server predicted the influence of mutation on protein stability by calculating the free energy change value ($\Delta\Delta G$) of protein, where positive $\Delta\Delta G$ value indicated higher stability and a negative value indicated a decrease in the stability of the protein [8]. The output as $\Delta\Delta G$ value and reliability index (RI) value ($0 < RI < 10$) are in Table I.

Modeling the Molecular Effects of SNPs on Protein 3D Structure

The crystal structure of the kinase domain of EGFR protein (EGFRK) in complex with a 4-anilinoquinazoline (AQ4) inhibitor erlotinib is accessible in Protein Data Bank (PDB ID: 1M17) [9]. The EGFR contains 1210 amino acids, from which 333 amino acids of the kinase domain (696-988 of EGFR) have been resolved in crystal structure with a resolution of 2.60 Å. To generate the mutated (L718A, Q849H, and L858R) 3D models of EGFR protein, Chimera (<http://www.cgl.ucsf.edu/chimera>) (version 1.8) was used [10]. The atomic clashes of the structure were removed by 100 steps energy minimization using AMBER f12SB force field parameters and steepest descent and conjugate gradient algorithms (step size = 0.02 Å).

Molecular Dynamics Simulation

The molecular dynamics simulations (MD) of native and mutant forms of EGFR protein in complex with AQ4 were implemented using GROMACS version 2020.1 [11]. The topology file of AQ4 was generated utilizing the PRODRG server, and the united-atom GROMOS 96 43A1 force field was hired for EGFR protein topology files [12]. Solvation of system was with SPC (simple point charge) water model in a 10 Å cubic box. The system was neutralized with Na⁺ and Cl⁻ ions, and energy was minimized in several steps to relax the system. Subsequently, system equilibration was conducted by the 500 ps of NVT ensemble (constant number of particles, volume, and temperature) followed by the 500 ps of isothermal-isobaric NPT ensemble (constant number of particles, pressure, and temperature). MD simulations were carried out for a period of 100 ns with two fs time steps. MD trajectories were analyzed in terms of time-dependent structural properties by calculating the root mean square deviation (RMSD) and root-mean-square fluctuation (RMSF). The graphs were visualized using the Xmgrace software. The binding affinity of complexes was calculated using the PRODIGY server [13].

Retrieval of Variants

Analysis of gene sequencing of the tumor confirmed three simultaneous mutations, L718A in exon 18 and Q849H and L858R in exon 21 of EGFR. The dbSNP-NCBI database was searched for retrieving the Reference SNPs (rs) in the human EGFR gene (Gene ID: 1956). The somatic mutation of L858R corresponds to variant rs121434568. However, there was no reference ID for L718A and the Q849H.

Results

Analysis of Phenotypic Impacts by PROVEAN and SNAP2

PROVEAN scores indicated that variants L718A and L858R (score of -4.213 and -5.202, respectively) were deleterious, though Q849H (score of -0.698) was predicted as neutral. The results from the SNAP2 server predicted L718A and L858R as effective functional mutations with an expected accuracy of 75% and 95% and a score of 59 and 96, respectively, whereas Q849 was considered as neutral (expected accuracy of 78% and a score of -60) (Table I). Protein function is affected by amino acid substitutions,

Table 1: Evaluation of mutation effects on EGFR.

Variant	PROVEAN score	Prediction (cutoff= -2.5)	SNAP2 score	I-Mutant stability prediction	$\Delta\Delta G$ (Kcal/mol)	RI ¹
L718A	-4.213	Deleterious	59	Decrease	-3.32	9
Q849H	-0.698	Neutral	-60	Decrease	-1.08	7
L858R	-5.202	Deleterious	96	Decrease	-2.26	8

¹ Reliability index

and its impact is calculated as a score ranges from -100 as a strong neutral prediction to +100 as a strong effect prediction that reflects the likelihood of the mutation to alter the native protein function.

Prediction of Protein Stability Changes due to Mutation Using I-Mutant 2.0 server

The prediction of stability changes of EGFR due to the mutations by I-Mutant 2.0 is given in Table 1. The $\Delta\Delta G$ results showed negative values that were interpreted as a decrease in the stability of the protein owing to the substitutions in amino acid sequence. All three mutants were discovered to be reduced in protein stability with a reliability index (RI) above 5.

Structural Analysis of Mutants and Native Structures Using MD Simulation

To investigate the structural deviations as well as stability differences between wild-type and mutant forms of EGFR protein, structural analysis was done. Backbone Root Mean Square Deviation (RMSD) of protein compared to the initial structure was used to investigate the conformational behavior of EGFR complexes during 100 ns simulation. The RMSD plots are depicted in Figure 1. A. The RMSD values of the complexes were in the range of 0.3-0.4 nm, while there was a slightly higher fluctuation in the triple mutant complex (EGFR-AHR) from 70 to 90 ns of simulation compared to the native complex. Also, the RMSD plot of EGFR-L858R complex showed overlapping with the triple mutant complex during the simulation, though the structural deviation of EGFR-L858R complex was slightly higher in the final 20 nanoseconds.

The structure compactness of complexes during simulations was calculated by the radius of gyration (Rg). The EGFR-L858R complex showed higher fluctuations compared to the triple mutant EGFR. The Rg value of the

triple mutant EGFR and EGFR-L858R complexes was in the range of 1.95-2 nm, and 1.9-2.1 nm, respectively. Comparing to the native EGFR, both mutant forms indicated higher Rg values. Though, the Rg plots of native EGFR and EGFR-L858R complexes were fluctuated, the Rg values for all complexes were very close to each other during the last 30 ns of simulation (Figure 1. B).

Analysis of root means square fluctuations (RMSFs) displayed that all the residues in protein structures fluctuated between 0.05 and 0.35 nm in both complexes (Figure 1. C). The zoom-in of regions 700-730 and 840-870 of native, the triple mutant EGFR, and EGFR-L858R is illustrated in Figure 1.D and Figure 1.E, respectively. There was no difference in the residue fluctuations between native, and EGFR-L858R in region 700-730, though a slightly higher fluctuation was seen for the triple mutant EGFR (the mutated residue L719A is indicated). Despite relatively similar fluctuation of L858R in the triple mutant EGFR, and EGFR-L858R plots, higher fluctuation was detected in Q849H of the triple mutant EGFR compared to native, and EGFR-L858R complexes. The cartoon representation of wild-type and mutant residues of all complexes is depicted in Figure 1. F, G and H.

Distinct Binding Pattern of Wild-type and Mutant EGFR to Erlotinib After Simulation

The interaction model of EGFR with erlotinib was evaluated before and after MD simulation to understand the mutation effect on the binding affinity. LigPlot visualized the hydrogen bonds of erlotinib with the surrounding residues in both forms of mutant protein (AHR and L858R) with regard to the wild-type (Figure 2 and 3). Initially, the hydrogen binding pattern of erlotinib to EGFR in both mutant forms was remained similar to wild-type structure, where M793 showed hydrogen binding interaction with erlotinib in both cases. Analysis of H-bond in the snapshots of wild-type and mutant EGFR complexes during simulation revealed that wild-

Figure 1: Analysis of MD simulations for wild-type and mutant forms of EGFR. (A) RMSD and (B) Rg plot for EGFRs. The plots of the wild-type, the triple mutant form (EGFR-AHR), and EGFR-L858R are colored in black, blue, and pink, respectively. (C) RMSF values of EGFRs' residues during simulations. (D) Zoom-in the RMSF values of residues 700-730. (E) Zoom-in the RMSF values of residues 840-870. Cartoon view of the superimposition of the wild-type, EGFR-AHR, and EGFR-L858R 3D structures for variants including (F) L718A, (G) Q849H, and (H) L858R. Wild-type, EGFR-AHR, and EGFR-L858R residues are colored in red, blue, and cyan, respectively.

type EGFR and EGFR-L858R had preserved the H-bond of M793 to erlotinib, whereas the triple mutant form showed different H-bond interaction of erlotinib (Q791) (Figure 4). Further, evaluation of binding affinity of the wild-type, EGFR-L858R, and the triple mutant form of EGFR indicated that affinity of protein and erlotinib as a

small ligand in the initial complexes was closely similar (-8.7, -8.7 and -8.5 Kcal/mol, respectively), whereas energy minimization during simulation revealed the difference in binding affinity of the wild-type, EGFR-L858R, and the triple mutant EGFR to erlotinib (-10, -9.5, and -8.6 Kcal/mol, respectively).

Figure 2: Ligplot analysis of erlotinib binding with the triple mutant and wild-type EGFR. **A)** Wild-type, and **B)** EGFR-AHR form. EGFR residue in H-bond is depicted in orange (chain A) and AQ4 is (erlotinib) in blue (chain B). The green dashed lines refer to hydrogen bonds. The non-stick residues in red correspond to hydrophobic contacts of EGFR. Mutations have changed the hydrophobic interactions of erlotinib compared to the wild-type.

Discussion

Even though EGFR-TKI harbors a better clinical benefit in the treatment of NSCLC patients with EGFR mutations compared with traditional platinum-based combination chemotherapy, patients will inevitably develop acquired EGFR-TKIs resistance [14-15]. Co-occurring mutations in

EGFR or combined mutations with other genes have been identified as one of the various mechanisms of resistance to TKIs [16]. In order to understand the mutation effect on the inhibition potency of erlotinib, we followed the changes of the EGFR kinase domain in wild-type and mutant forms bearing L858R+Q849H+L718A and L858R alone in complex with erlotinib. The L718A and L858R

Figure 3: Ligplot analysis of erlotinib binding with the triple mutant and EGFR-L858R. **A)** EGFR-L858R, and **B)** EGFR-AHR form. EGFR residue in H-bond is depicted in orange (chain A) and AQ4 is (erlotinib) in blue (chain B). The green dashed lines refer to hydrogen bonds. The non-stick residues in red correspond to hydrophobic contacts of EGFR. L718A, and Q849H in EGFR-AHR have changed the hydrophobic interactions of erlotinib compared to EGFR-L858R.

mutations were predicted as effective functional mutations by SNAP2 and deleterious by PROVEAN. Although Q489H was predicted as a neutral mutation by PROVEAN, L858R, Q849H, and L718A were observed to decrease the protein stability by I-Mutant 2.0.

The hydrophobic clamp structure of the kinase engages the inhibitory compound through hydrophobic interactions between side-chains of Leu 718, Val 726, and Leu 844. It is concluded that the binding affinity of the compound-EGFR complex may be influenced by mutations that occurred in the hydrophobic clamp.

Figure 4: The binding pattern of wild-type, and EGFR variants to erlotinib before and after simulation. **(A)** Superimposition of wild-type and EGFR-L858R/erlotinib complexes (colored in orange, and green, respectively), **(B)** Superimposition of wild-type and EGFR-AHR/erlotinib complexes (colored in orange, and yellow, respectively), and **(C)** Superimposition of EGFR-L858R and EGFR-AHR/erlotinib complexes (colored in green, and yellow, respectively). **(D)** Superimposition of wild-type, EGFR-L858R, and EGFR-AHR/erlotinib complexes before MD (colored in light blue, light and dark pink, respectively). Stick formats of residues 718, 849 and 858 of wild-type, EGFR-L858R, and EGFR-AHR are colored in red, yellow and blue, respectively. **(E)** Superimposition of wild-type, EGFR-L858R, and EGFR-AHR/erlotinib complexes after MD (colored in orange, green and yellow, respectively). Before MD simulation, L858R had not caused a detectable difference in protein compared to wild-type (the light blue color of wild type is not distinct from the light pink color of the mutant). Stick formats of residues 718, 849 and 858 of wild-type, EGFR-L858R, and EGFR-AHR are colored in red, cyan and blue, respectively. **(F)** The binding pattern of wild-type EGFR to erlotinib (M793). **(G)** The binding pattern of EGFR-L858R to erlotinib (M793). **(H)** Binding pattern of EGFR-AHR to erlotinib (Q791). Erlotinib is colored in pink, cyan and green in complex with wild-type, EGFR-L858R, and EGFR-AHR, respectively.

Moreover, in some cases, these mutations result in drug resistance (e.g., models harboring L718Q, L844V, and C797S have rendered EGFR resistance to pyrimidine-based EGFR kinase inhibitors (WZ4002 and CO-1686) *in-vitro*) [17]. It is clear that Ala718, as a hydrophobic residue, has a shorter side-chain compared to leucine, which may result in a shift in hydrophobic interaction around the erlotinib. Although in the analysis of amino acid fluctuations, a slight difference was observed between the triple mutation, EGFR-L858R, and wild type in position 718, the amount of fluctuations in amino acids before and after this position (A718) was higher compared to EGFR-L858R and wild type (both L718). It is possible that the presence of alanine in this position has led to changes in the spatial arrangement of adjacent amino acids.

Conformational changes of all kinases during the transition from the inactive to the active form include unfolding the short α C-helix (residues G857–G863 of EGFR) and adapting an extended conformation. Kannan et al. showed by structural modeling and binding assays that the presence of Arg at the 858 position of EGFR destabilizes the inactive state by the opening of the allosteric pocket transiently during the transition from the inactive to the active state [18]. Consistent with the role of L858R, slight fluctuations of residues were detected in the RMSF plot of mutant EGFR during MD simulation. Comparing data analysis of the EGFR-AHR and EGFR-L858R revealed that the residue fluctuations in the range of 840–870 were more consistent with the triple mutation. However, the magnitude of the Q849H effect appears to be structurally smaller than that seen for L858R / L718A mutations.

Erlotinib binds to EGFR in the ATP-binding cleft that tightly associates to the hinge of the kinase and forms a hydrogen bond with Met at the 793 position. The wild-type and mutant forms of EGFR initially showed a similar binding mode to erlotinib, where the erlotinib receives a hydrogen bond from Met793 in all cases. Then, the binding mode of wild-type and the mutant forms of EGFR to erlotinib was studied during simulation. Evaluation of all EGFR forms by aligning before the simulation showed similar orientation of erlotinib towards the EGFR binding site. After molecular dynamic simulation, wild type and EGFR-L858R displayed relatively similar binding pattern (M793) to erlotinib but binding mode of EGFR-AHR was different in terms of hydrogen bond and hydrophobic interactions. Analysis of snapshots during simulations revealed that erlotinib was nearly stable in the bound conformation to wild-type throughout a 100 ns MD run, though there were some fluctuations that had changed the residues involved in hydrogen bond formation of mutant form. In

the snapshots of wild-type and mutant EGFR complexes during simulation, erlotinib bound preferentially to M793 in wild-type and EGFR-L858R, whereas EGFR-AHR showed different H-bond interaction to erlotinib (Q791). Likewise, the binding affinity of the wild-type and mutant forms of EGFR to erlotinib was different after energy minimization during molecular dynamic simulation (-10, -9.5 and -8.6 Kcal/mol, respectively).

In the study of the binding mechanism of ATP to EGFR, it has been found that the association of L858R mutation in EGFR with T790M has led to drug resistance to gefitinib and erlotinib, and this effect, in addition to the role of mutations in altering the binding affinity to TKIs, was due to increased affinity of mutants for ATP [19]. In the present study, the mechanism of ATP binding was not investigated, but it is possible that altered ATP affinity, along with differences in protein binding to erlotinib due to mutations, is effective in drug resistance. This issue can be studied in the future.

A search of articles published in PubMed was conducted using the terms “NSCLC,” “EGFR,” “cancer,” “uncommon mutations,” “L718X,” “Q849H,” “rare mutations,” “L858R,” “exon 20,” “atypical mutations,” and “compound mutations”. It had so far found no reports of co-occurrence of L858R + Q849H + L718A in EGFR, neither in NSCLC nor in other cancers. Prevalent uncommon mutations were detected in the P-loop (L718–V726) and the C-terminal loop of the α C-helix that account for 10–20% of all EGFR mutations [20]. L718Q/N occurred primarily in exons 18 (1.3% of EGFR mutations) and has been reported that L718Q is nearly resistant to osimertinib, especially when in association with the L858R mutation [21–22]. Despite sensitivity of most uncommon EGFR mutations (e.g. G719X, S768I, and Q861) to afatinib (second generation TKIs), patients with NSCLC containing uncommon and compound EGFR mutations show reduced and heterogeneous responses to EGFR inhibitors including osimertinib and they are significantly less sensitive to first generation TKIs. However, there are no clear guidelines for EGFR TKI treatment for such patients and they experience shorter time to treatment failure (TTF) compare to other patients [22].

This study reported a rare case of lung cancer harboring a compound L858R+Q849H+L718A EGFR substitution mutations, which was identified subsequent to the resistance to erlotinib treatment. Advances in understanding mutation effects on the EGFR protein functions may provide promising insights to enhance the conductance for optimal treatment schemes in individual patients. Despite the high cost of comprehensive genome sequencing, structural evaluation of genetic alterations

could identify the genetic causes of drug resistance at diverse points during the clinical course.

Abbreviations

EGFR; Epidermal growth factor receptor, TK; tyrosine kinase, SNP; Single nucleotide polymorphism, NSCLC; Non-small cell lung carcinoma, OS; overall survival, KRAS; Kirsten ras, ALK; anaplastic lymphoma kinase, ROS1; repressor of silencing 1, AQ4; 4-anilinoquinazoline, TKIs; tyrosine kinase inhibitors, PDB; Protein Data Bank, MD; Molecular Dynamics, RMSD; Root Mean Square Deviation, RMSF; Root Mean Square Fluctuation, SPC; simple point charge, RI; Reliability Index.

Author contributions: H.M., J.H., K.S., A.A.D.; Contributed to conception and design. A.A.D.; Contributed to the experimental work; E.R., A.S.L., M.M.D Contributed to in silico analysis, data analysis, and interpretation of data. H.M., J.H., K.S.; Were responsible for overall supervision. E.R.; Drafted the manuscript, which was revised by H.M., A.S.L., and M.M.D. All authors read and approved the final manuscript.

Conflicts of interests: The authors state no conflict of interest

Data Availability Statement: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Funding: There is no financial support in this study.

Acknowledgement: The authors wish to thank the Lung Transplantation Research Center of Masih Daneshvari Hospital, where this research was conducted.

References

1. Luo YH, Luo L, Wampfler JA, Wang Y, Liu D, Chen YM, et al. 5-year overall survival in patients with lung cancer eligible or ineligible for screening according to US Preventive Services Task Force criteria: a prospective, observational cohort study. *Lancet Oncol*. 2019 Aug;20(8):1098–108.
2. Duma N, Santana-Davila R, Molina JR. Non-Small Cell Lung Cancer: Epidemiology, Screening, Diagnosis, and Treatment. *Mayo Clin Proc*. 2019 Aug;94(8):1623–40.
3. Pikor LA, Ramnarine VR, Lam S, Lam WL. Genetic alterations defining NSCLC subtypes and their therapeutic implications. *Lung Cancer*. 2013 Nov;82(2):179–89.
4. Gaughan EM, Costa DB. Genotype-driven therapies for non-small cell lung cancer: focus on EGFR, KRAS and ALK gene abnormalities. *Ther Adv Med Oncol*. 2011 May;3(3):113–25.
5. Riess JW, Gandara DR, Frampton GM, Madison R, Peled N, Bufill JA, et al. Diverse EGFR exon 20 insertions and co-occurring molecular alterations identified by comprehensive genomic profiling of NSCLC. *J Thorac Oncol*. 2018 Oct;13(10):1560–8.
6. Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*. 2015 Aug;31(16):2745–7.
7. Hecht M, Bromberg Y, Rost B. Better prediction of functional effects for sequence variants. *BMC Genomics*. 2015; 16 Suppl 8(Suppl 8): S1. <https://doi.org/10.1186/1471-2164-16-S8-S1>.
8. Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Res*. 2005 Jul;33(Web Server issue suppl_2):W306–10.
9. Stamos J, Sliwkowski MX, Eigenbrot C. Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. *J Biol Chem*. 2002 Nov;277(48):46265–72.
10. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem*. 2004 Oct;25(13):1605–12.
11. Abraham M J, Murtola T, Schulz R, Páll S, Smith J C, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. 2015; 1-2(19-25).
12. Schüttelkopf AW, van Aalten DM. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr*. 2004 Aug;60(Pt 8):1355–63.
13. Xue LC, Rodrigues JP, Kastritis PL, Bonvin AM, Vangone A. PRODIGY: a web server for predicting the binding affinity of protein-protein complexes. *Bioinformatics*. 2016 Dec;32(23):3676–8.
14. Ahn MJ, Sun JM, Lee SH, Ahn JS, Park K. EGFR TKI combination with immunotherapy in non-small cell lung cancer. *Expert Opin Drug Saf*. 2017 Apr;16(4):465–9.
15. Thress KS, Paweletz CP, Felip E, Cho BC, Stetson D, Dougherty B, et al. Acquired EGFR C797S mutation mediates resistance to AZD9291 in non-small cell lung cancer harboring EGFR T790M. *Nat Med*. 2015 Jun;21(6):560–2.
16. Bean J, Brennan C, Shih JY, Riely G, Viale A, Wang L, et al. MET amplification occurs with or without T790M mutations in EGFR mutant lung tumors with acquired resistance to gefitinib or erlotinib. *Proc Natl Acad Sci USA*. 2007 Dec;104(52):20932–7.
17. Ercan D, Choi HG, Yun CH, Capelletti M, Xie T, Eck MJ, et al. EGFR mutations and resistance to irreversible pyrimidine-based EGFR inhibitors. *Clin Cancer Res*. 2015 Sep;21(17):3913–23.
18. Kannan S, Venkatachalam G, Lim HH, Surana U, Verma C. Conformational landscape of the epidermal growth factor receptor kinase reveals a mutant specific allosteric pocket. *Chem Sci (Camb)*. 2018 May;9(23):5212–22.
19. Saldaña-Rivera L, Bello M, Méndez-Luna D. Structural insight into the binding mechanism of ATP to EGFR and L858R, and

- T790M and L858R/T790 mutants. *J Biomol Struct Dyn*. 2019 Oct;37(17):4671–84.
20. Evans M, O'Sullivan B, Smith M, Hughes F, Mullis T, Trim N, et al. Large-Scale EGFR Mutation Testing in Clinical Practice: Analysis of a Series of 18,920 Non-Small Cell Lung Cancer Cases. *Pathol Oncol Res*. 2019 Oct;25(4):1401–9.
 21. Li H-S, Yang G-J, Wang Y. Case Report: Dacomitinib May Not Benefit Patients Who Develop Rare Compound Mutations After Later-Line Osimertinib Treatment. *Frontiers in oncology*. 2021; 11(649843-649843).
 22. Robichaux JP, Le X, Vijayan RS, Hicks JK, Heeke S, Elamin YY, et al. Structure-based classification predicts drug response in EGFR-mutant NSCLC. *Nature*. 2021 Sep;597(7878):732–7.