Genotype-determined EGFR-RTK heterodimerization and its effects on drug resistance in lung Cancer treatment revealed by molecular dynamics simulations

Mengxu Zhu[*] ⓘ, Debby D. Wang and Hong Yan

Abstract

Background: Epidermal growth factor receptor (EGFR) and its signaling pathways play a vital role in pathogenesis of lung cancer. By disturbing EGFR signaling, mutations of EGFR may lead to progression of cancer or the emergence of resistance to EGFR-targeted drugs.

Results: We investigated the correlation between EGFR mutations and EGFR-receptor tyrosine kinase (RTK) crosstalk in the signaling network, in order to uncover the drug resistance mechanism induced by EGFR mutations. For several EGFR wild type (WT) or mutated proteins, we measured the EGFR-RTK interactions using several computational methods based on molecular dynamics (MD) simulations, including geometrical characterization of the interfaces and conventional estimation of free energy of binding. Geometrical properties, namely the matching rate of atomic solid angles in the interfaces and center-of-mass distances between interacting atoms, were extracted relying on Alpha Shape modeling. For a couple of RTK partners (c-Met, ErbB2 and IGF-1R), results have shown a looser EGFR-RTK crosstalk for the drug-sensitive EGFR mutant while a tighter crosstalk for the drug-resistant mutant. It guarantees the genotype-determined EGFR-RTK crosstalk, and further proposes a potential drug resistance mechanism by amplified EGFR-RTK crosstalk induced by EGFR mutations.

Conclusions: This study will lead to a deeper understanding of EGFR mutation-induced drug resistance mechanisms and promote the design of innovative drugs.

Keywords: Epidermal growth factor receptor (EGFR), Drug resistance, EGFR signaling, Signaling crosstalk, Molecular dynamics (MD) simulations, Geometric properties

* Correspondence: mengxuzhu2-c@my.cityu.edu.hk
Department of Electrical Engineering, City University of Hong Kong,
Kowloon, Hong Kong

Background

Lung cancer is the leading cancer killer globally [1, 2]. Accounting for 85% of all cases, non- small cell lung carcinoma (NSCLC) is a main type of lung cancer and has thus become an active research topic in pathogenesis [3]. Epidermal growth factor receptor (EGFR), belonging to the ErbB family, plays an important role in the molecular pathology of NSCLC and other cancers [4]. Upon activation by a ligand, EGFR will form a homodimer or heterodimer with a partner, which stimulates its tyrosine kinase (TK) activity and switches on the downstream signaling [5, 6]. Over-expression of EGFR appear in nearly 60% of NSCLC patients [7], who frequently harbor EGFR mutations (like L858R) that will lead to aberrant singling, enhanced cell proliferation and progression of lung cancer. Tyrosine kinase inhibitors (TKIs) to EGFR, such as Gefitinib or Erlotinib, are generally effective in treating EGFR-mutated NSCLC patients. However, the treatment will encounter failures later due to the emergence of drug resistance [8, 9]. Well-explained reasons for such resistance lie in the decreased EGFR-TKI binding affinity induced by a second EGFR mutation (like T790M, accounting for 50% of resistant cases) [10–13] and the associated alterations in cell signaling pathways caused by crosstalk between EGFR and other receptor tyrosine kinases (RTKs) [14, 15].

Crosstalk between EGFR and other RTKs, such as ErbB-family members [16], c-Met [15] and Insulin-like growth factor 1 receptor (IGF-1R) [17, 18], are vital in ascertaining the mechanisms of drug resistance in cancer treatment. Previous research showed that both direct and indirect connections exist, including direct heterodimerization with EGFR to participate in the cell signaling [19–22]. For c-Met specifically, Tanizaki et.al have revealed that Met can form heterodimers with ErbB family members with immunoprecipitation [22]. Wheeler et.al also showed that EGFR heterodimerized with cMet [23]. Harwardt et al. and Knowles et al. have found the presence of heteromeric complex of EGFR and cMet by microscopy [24] or antibody array method [25]. Besides, Ortiz-Zapater et al. studied EGFR-Met dimerization on cell lines [21], and Lee et al. applied co-immunoprecipitation and Forster resonance energy transfer (FRET) FLIM to study the EGFR-Met dimers [26]. For IGF-1R, Nahta et al. have studied the influence of IGF-1R/HER-2(ErbB-2) dimers to drug resistance by experiments on cells [27]. Iyer et al. studied the heterodimerization of EGFR with IGF-1R on head and neck cancer cells [28]. Additionally, Oliveira et al. proposed that both direct and indirect crosstalk exists between EGFR and IGF-1R, from association through signal network to direct dimerization [29]. Morgillo et.al studied the drug resistance related to IGFR/EGFR heterodimer on lung cancer cell lines and mice [30]. Becker et.al also

investigated EGFR/IGFR dimers in non-small-cell lung cancer (NSCLC) cell lines [31]. EGFR-targeted TKIs cannot block the interactions (heterodimerization) between EGFR and other RTKs, and such unblocked interactions can even be amplified due to EGFR mutations. The mutation-induced amplified crosstalk or signaling will potentially lead to drug resistance. Assisted by a c-Met TKI (SGX523), Ortiz-Zapater et al. have assessed cell proliferation in vitro, tumor growth and EGFR-c-Met dimerization for lung cancer cell lines with different EGFR mutations. They showed significantly reduced cell proliferation, tumor growth and EGFR-c-Met dimerization caused by SGX523 in the resistant mutation case (L858R-T790M) than the other cases [21]. This demonstrated an amplified heterodimerization between c-Met and the resistant EGFR mutants than non-resistant ones, which further implies that EGFR-c-Met heterodimerization is determined by EGFR genotype [21]. Dual inhibitors to both EGFR and c-Met can mitigate the drug-resistant cases [32, 33], which also demonstrates the enhanced EGFR-c-Met heterodimerization as a potential resistant mechanism. Such mechanism also applies to the crosstalk between EGFR and IGF-1R [29, 30, 34]. Clinical studies have shown that IGF-1R-targeted TKIs can inhibit NSCLC cells that are resist to EGFR-targeted drugs [30], implying the significant role of IGF-1R in drug resistance mechanisms. In this study, we aim to explore the EGFR-mutation induced drug resistance mechanism through the perspective of EGFR-RTK crosstalk (heterodimerization), and we considered ErbB2 (amplification frequently occurs in NSCLC), c-Met and IGF1-R as partners for mutated EGFRs.

Molecular structural analysis, dynamics simulations and computational characterization have facilitated a great number of studies of EGFR-mutated lung cancer [16, 35–41]. Specifically, we studied the interactions between mutated EGFRs and an RTK partner (ErbB2, c-Met or IGF1-R) in heterodimerization using such computational techniques. Wild type (WT) EGFR, EGFR with the TKI-sensitive mutation L858R (point mutation at exon 21) [42] and that with co-expressed L858R and T790M mutations (TKI-resistant) [43] were comparably investigated in this study. Molecular dynamics (MD) simulations, as a widely used method in researching about lung cancer [40, 41, 44], were implemented on EGFR-RTK heterodimer structures, and three- dimensional Alpha Shape modeling was subsequently applied to each MD trajectory frame for extraction of geometrical properties [45, 46]. Based on the simulation and modeling results, we characterized the EGFR-RTK interactions in the heterodimers by geometrical properties of the interaction sites, including matching rate of atomic solid angles in these sites and center-of- mass distances between interacting atoms. Computationally estimating

the free energy of binding has been broadly used in revealing molecular binding affinities in previous studies [40, 41, 47]. As a benchmark method, it was also applied in this study to validate the geometrical characterization. These characterization methods will shed light on the correlations between the genotypes and the EGFR-RTK heterodimerization, further providing a deeper understanding of the mutation-induced drug resistance mechanisms.

Results

Modeling of EGFR mutants, formation of EGFR-RTK heterodimers and MD simulations of heterodimers

To derive the structures of EGFR mutants (L858R and L858R-T790M), we adopted *Rosetta* to model such structures according to homology modeling techniques. WT EGFR was used in this modeling as the structural template, whose unresolved residues were generated using *Rosetta* prior to the mutant modeling. The modeling results are shown in Fig. 1a and b, where the WT protein is superimposed on the mutant structures to shown the mutation sites. To form the EGFR-RTK heterodimer structures, we used the WT EGFR-EGFR homodimer structure as a template and aligned an EGFR WT/mutant (WT, L858R or L858R-T790M) and an RTK (c-Met, ErbB2 or IGF-1R) to the two positions in the dimer. An example of such heterodimers, namely the WT EGFR-c-Met dimer, is presented in Fig. 1c.

For each heterodimer, we simulated its dynamics in explicit-solvent environment using *Amber* software suite. A series of procedures were sequentially imposed on the solvated system, including a short energy minimization, a heating process, and a number of equilibration steps. An equilibrated system can guarantee a reliable production MD simulation for our analysis.

Therefore, the equilibration of each system was verified prior to the production MD step, using the root-mean-square-deviation (RMSD) curve of the EGFR-RTK dimer in the equilibration phase. Such curves for the

heterodimers (250 frames at time interval of 2 ps) are shown in Fig. 2, where each equilibration can be verified by a stable curve. Our production MD simulations for each heterodimer lasted for 3 ns, resulting in a trajectory of 5000 frames (interval of 10 ps).

Extraction of geometrical properties of EGFR-RTK interfaces in heterodimers and investigation of drug resistance mechanisms

As we are more interested in the EGFR-RTK interactions in each heterodimer in the dynamics, we first extracted their interfaces using weighted Alpha Shape modeling for all the MD trajectory frames based on Eq. (5). As an example, procedures to generate interfacial atoms of the WT EGFR-c-Met dimer in one trajectory frame (randomly selected) are showed in Fig. 3.

Sub-Fig. A is the original wild type EGFR-cMet dimer. We applied Alpha Shape Modeling to reconstruct the surface of the dimer (subfigure B). Similarly, we reconstructed the surfaces of the two proteins in the dimer (subfigure C: EGFR, subfigure D: cMet). The surface atoms of EGFR that are not on the surface of the dimer (comparing subfigures B and C) are regarded as the interfacial atoms on EGFR, and similarly we can capture the interfacial atoms on cMet (comparing subfigures B and D). Subfigures E and F show the main idea of this process.

The procedures are shown in the following diagram, as in Fig. 4. Information about the resulting interfacial atoms in this example is listed in Table 1.

For each heterodimer, we kept the pairs of interfacial atoms sharing an edge in any tetrahedron that connects the interaction sites of the two chains, and calculated the average solid angles of these interfacial atoms of each trajectory frame (Eq. (6)). Subsequently, the matching rate, namely the ratio of matched solid angle pairs over all pairs, was derived according to Eq. (7). Comparing the WT EGFR and mutant L858R, the matching rate
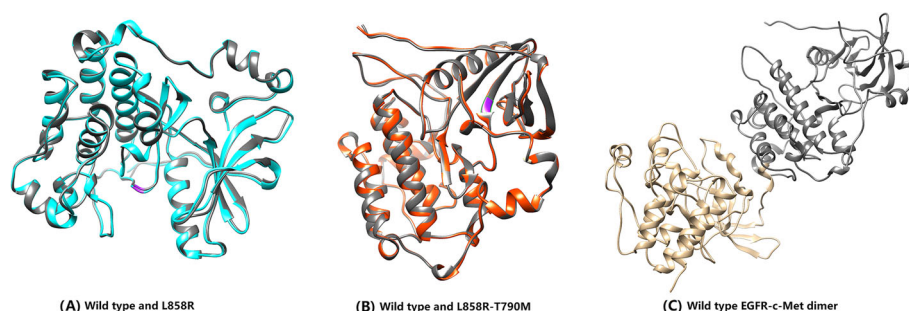

Fig. 1 (**a**) Modeled L858 (blue) mutant and wild type (grey) EGFR structures, with the mutation site colored purple. (**b**) Modeled L858R-T790M (red) mutant and wild type (grey) EGFR structures, with the mutation sites colored purple. (**c**) WT EGFR-c-Met heterodimer structure (WT EGFR: grey, c- Met: gold)

Zhu *et al. BMC Molecular and Cell Biology*   (2021) 22:34
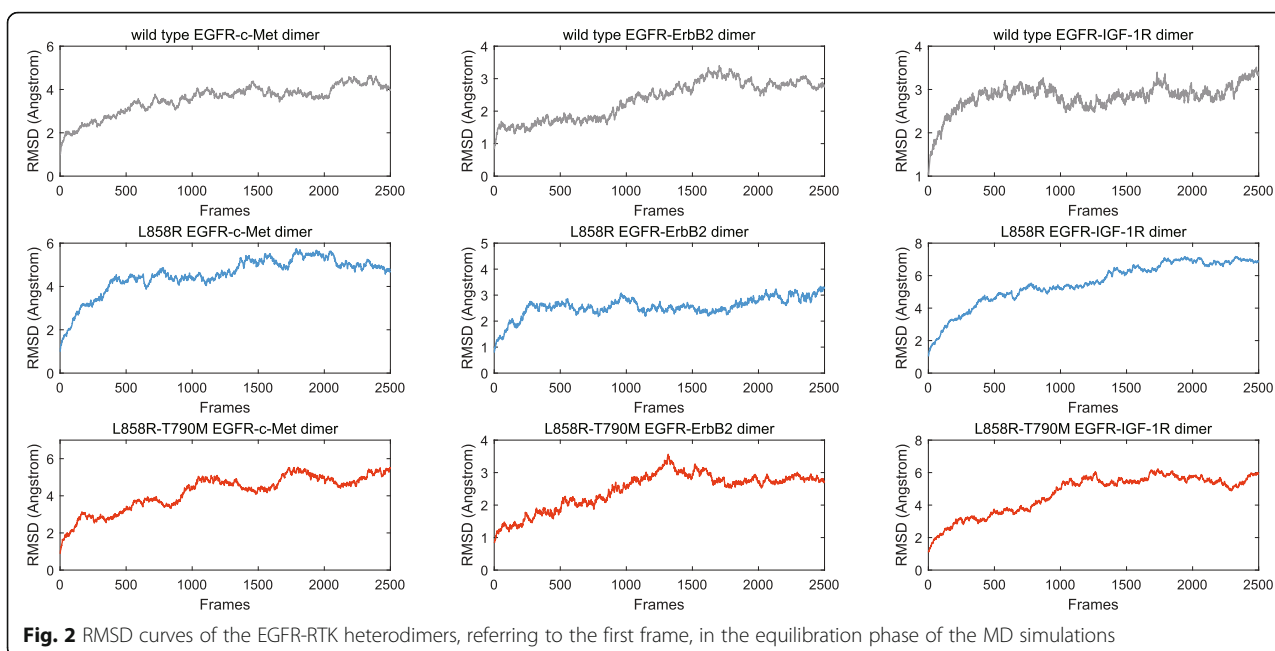
Page 4 of 13

Fig. 2 RMSD curves of the EGFR-RTK heterodimers, referring to the first frame, in the equilibration phase of the MD simulations

along the MD trajectory for each of the EGFR-RTK heterodimers is displayed in Figs. 4a (EGFR-c-Met), 5B (EGFR- ErbB2) and 5C (EGFR-IGF-1R) respectively.

Furthermore, we calculated the average COM distance of interfacial atom pairs, each of which share an edge in any boundary tetrahedron, for each EGFR-RTK heterodimer along the MD trajectory.

Comparisons of the COM distance trajectories are displayed in Fig. 5.

Finally, we estimated the binding free energy within each EGFR-RTK heterodimer using MM/GBSA in *Amber*. For each RTK partner (c-Met, ErbB2 or IGF-1R), WT EGFR, L858R and L858R-T790M were compared as a group in Fig. 6.
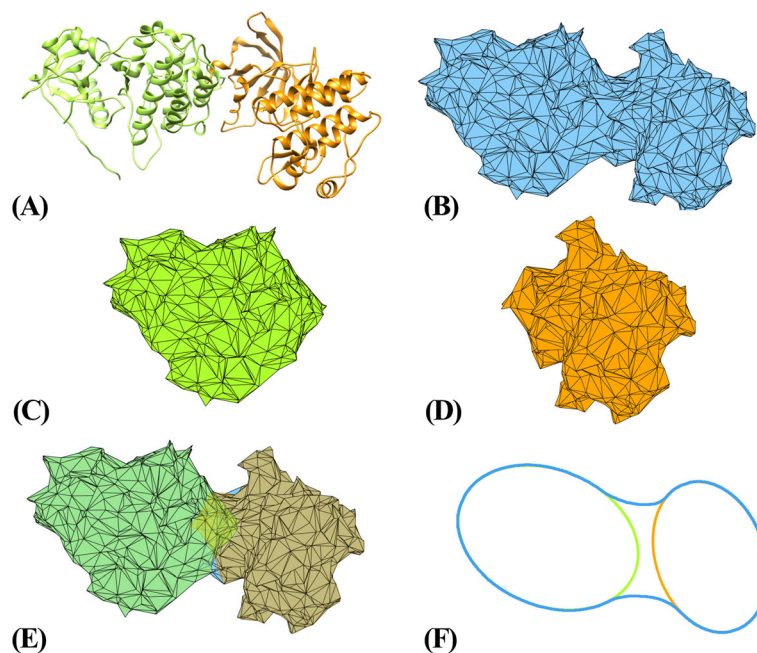

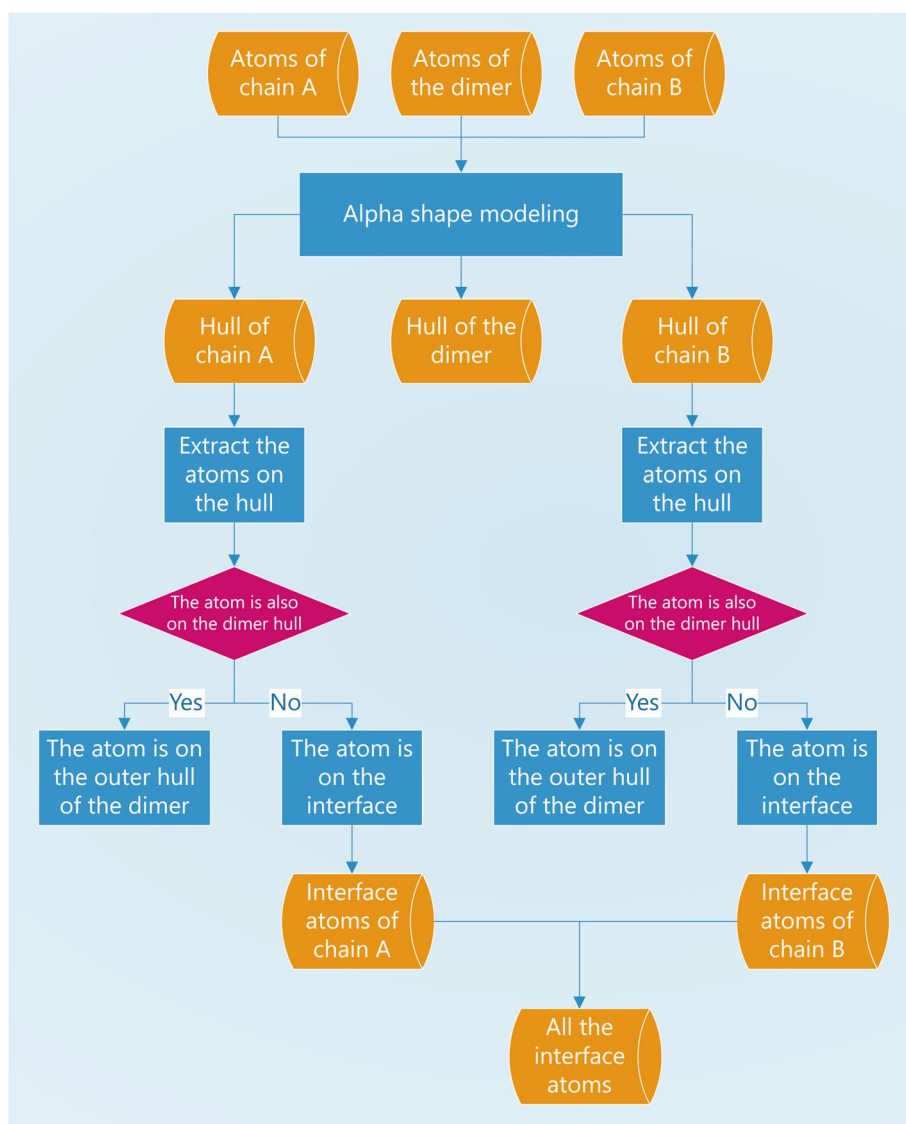**Fig. 3** Procedures to capture the interfacial atoms

Fig. 4 Flow chart of the algorithm to obtain the interface atoms

Discussions

The methods revealed properties of the dimers. As shown in Fig. 5, L858R-RTK dimers generally have a lower matching rate of interfacial atoms than WT EGFR-RTK dimers. Such differences indicate a looser crosstalk between L858R and downstream RTKs, than that between WT EGFR and RTKs. It demonstrates that for TKI blocked EGFR with L858R mutation there is a looser crosstalk between EGFR and RTKs, guaranteeing the effectiveness of EGFR-TKIs in treating such mutated cases in the perspective of signaling crosstalk. The comparison results of matching degree for L858R-RTK and L858R-T790M- RTK heterodimers are presented in Figs. 5d-f. Compared to L858R, mutation L858R- T790M corresponds to a higher matching rate for all the EGFR-

RTK dimers, implying a tighter crosstalk between L858R-T790M and these RTKs than that between L858R and RTKs. This amplified crosstalk for mutant L858R-T790M probably induces the drug resistance, compared to L858R.

The COM distance method shows similar results. For WT EGFR and mutant L858R, distance trajectory for each of the EGFR-RTK heterodimers is presented in Figs. 6a-c respectively. L858R-RTK dimers generally have a longer interfacial distance than the WT EGFR-RTK dimers, implying a higher instability and a looser crosstalk. Similarly, comparisons between L858R-RTK dimers and L858R-T790M- RTK dimers in Figs. 6d-f show a shorter interfacial distance and a tighter crosstalk between L858R-T790M and RTKs. This proposes the effectiveness of TKIs
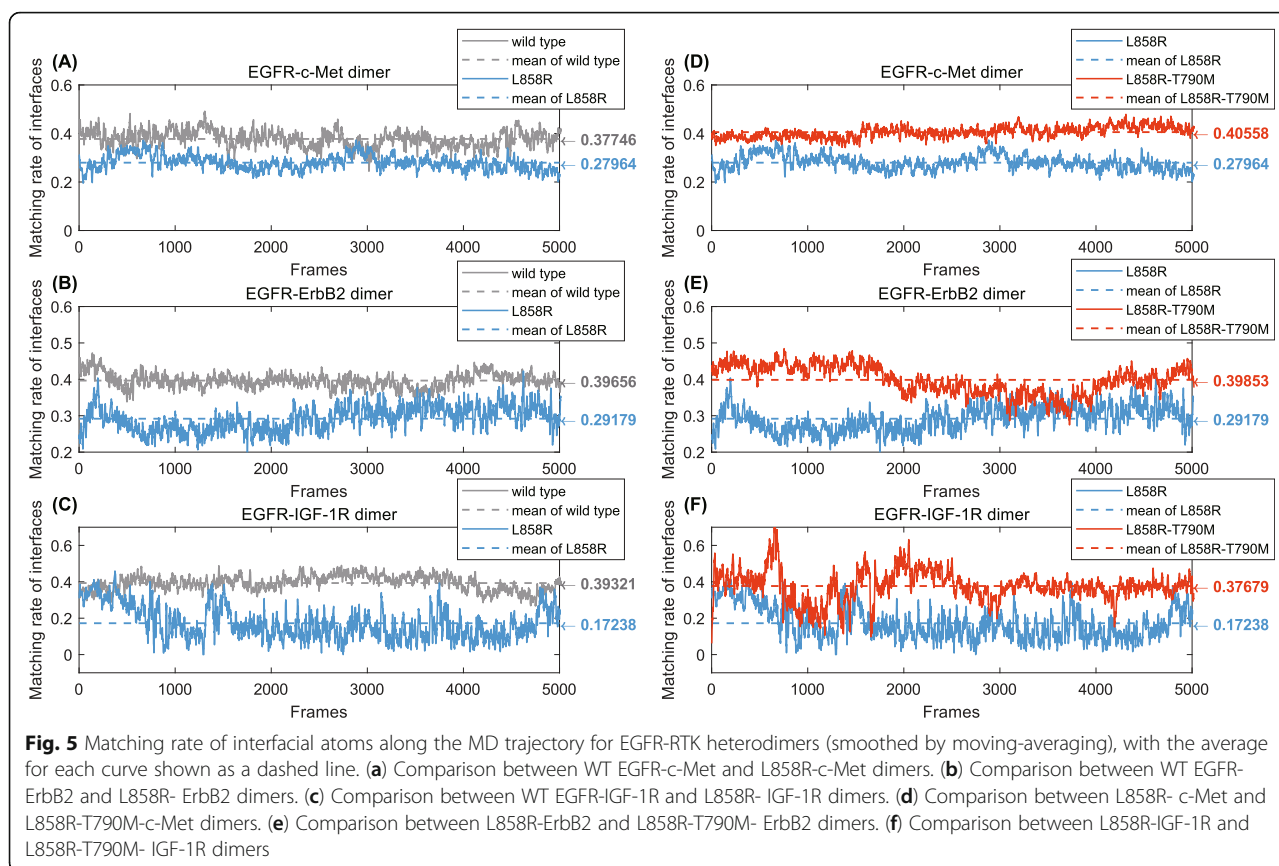
Table 1 Information of the interfacial atoms. Indices of the atoms and residues here are different form the numbers in the original pdb file, as they were renumbered sequentially from 1.

Chain label	Residue number	Residue name	Atom number	atom type	interacting partners
A	1	GLN	3	C	3862,4094,4098
A	1	GLN	8	C	4098,4229
A	2	ALA	11	C	4097,4098
A	2	ALA	13	C	4200
A	5	HIE	34	C	3838,3860,3861
A	7	VAL	50	C	3838,4222,4224
A	56	GLN	427	C	4221
A	56	GLN	428	O	4221
A	56	GLN	431	C	
A	59	THR	456	C	4199
A	60	GLU	458	N	4199
A	60	GLU	459	C	4199,4176
A	60	GLU	461	C	4199
A	62	ILE	474	C	4176,4199,4200
A	62	ILE	475	C	4194
A	62	ILE	477	C	
A	62	ILE	478	O	4194
A	80	LEU	610	C	4073
B	481	VAL	3838	C	34,50
B	484	TRP	3860	C	34
B	484	TRP	3861	C	34
B	484	TRP	3862	C	3
B	511	ARG	4073	C	610
B	512	LEU	4080	C	
B	514	GLN	4094	C	3
B	514	GLN	4097	C	11
B	514	GLN	4098	O	3,8,11
B	523	TYR	4171	O	
B	524	MET	4176	S	459,474
B	526	MET	4194	C	475,478
B	527	VAL	4196	N	
B	527	VAL	4199	C	456,458,459,461,474
B	527	VAL	4200	C	13,474
B	530	TRP	4221	C	427,428
B	530	TRP	4222	C	50
B	530	TRP	4224	C	50
B	530	TRP	4229	C	8

to mutation L858R and the resistance mechanism of TKIs to a second mutation T790M, from the perspective of heterodimer stability and EGFR-RTK crosstalk.

Results of binding free energy method provides another evidence. As shown in Fig. 7, for all RTK partners, L858R corresponds to the lowest binding affinity with them (highest binding free energy), showing the effectiveness of TKIs to this mutation. Comparing to L858R, a second mutation T790M (co-expressed L858R and T790M mutations) results in a higher binding affinity (lower binding free energy) with the RTKs, promisingly indicating the drug resistance mechanism of a

Fig. 5 Matching rate of interfacial atoms along the MD trajectory for EGFR-RTK heterodimers (smoothed by moving-averaging), with the average for each curve shown as a dashed line. (**a**) Comparison between WT EGFR-c-Met and L858R-c-Met dimers. (**b**) Comparison between WT EGFR-ErbB2 and L858R- ErbB2 dimers. (**c**) Comparison between WT EGFR-IGF-1R and L858R- IGF-1R dimers. (**d**) Comparison between L858R- c-Met and L858R-T790M-c-Met dimers. (**e**) Comparison between L858R-ErbB2 and L858R-T790M- ErbB2 dimers. (**f**) Comparison between L858R-IGF-1R and L858R-T790M- IGF-1R dimers
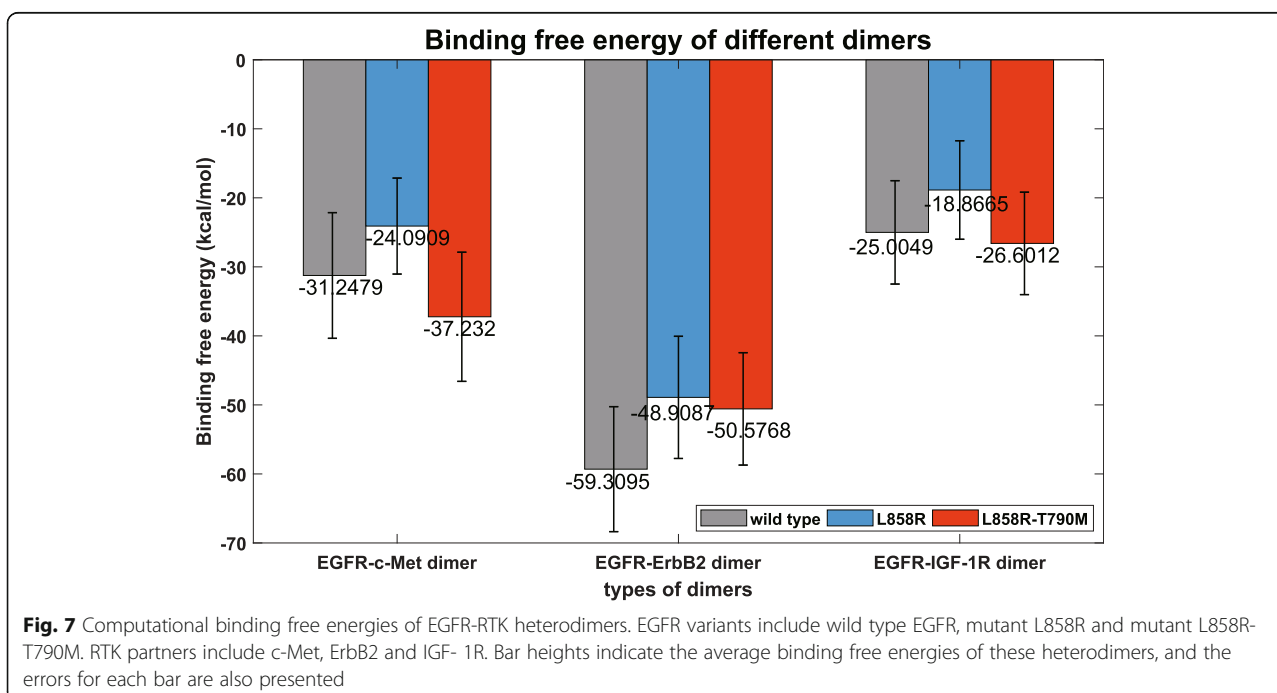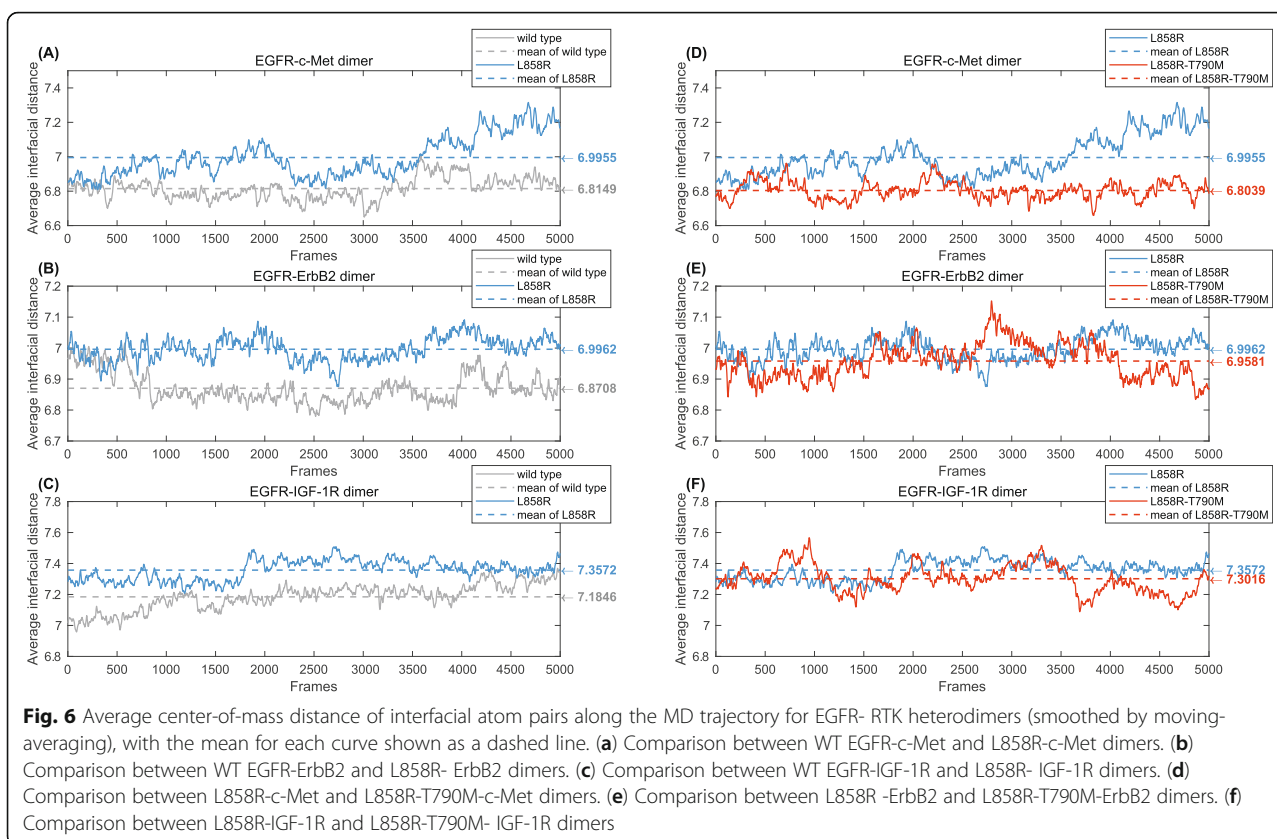
tighter L858R-T790M-RTK crosstalk. These results are consistent to those revealed by the interfacial geometrical properties that are extracted in the preceding sections.

Conclusion

Protein mutations will result in deformation of protein surfaces and changes in geometrical property, which will alter their interactions with other partners and thus modify their functions. In this study, we examined how mutation influences the crosstalk between EGFR, who plays a significant role in lung cancer progression, and its RTK partners in downstream signaling pathways. WT EGFR, a TKI-effective mutation L858R and a TKI-resistant mutation L858R-T790M were considered and paired up with RTK partners c-Met, ErbB2 and IGF-1R to form EGFR-RKT heterodimers. MD simulations were implemented for each dimer to uncover its dynamics, after which Alpha Shape modeling was applied to reveal the geometrical properties of dimer interfaces. Specifically, matching rate of atomic solid angles in the interaction sites and center-of-mass distances between interacting atoms of each dimer were calculated based on the extracted Alpha Shapes. Results show that L858R is the most alienated from the RTK partners while a second mutation (T790M) will make it closer to the RTKs.

This demonstrates a potential drug resistance mechanism which is resulted from the amplified EGFR-RTK crosstalk induced by EGFR mutations. It also shows that such EGFR-RTK heterodimerization is determined by EGFR genotypes. Conventional estimation of binding affinity for a molecular binding system based on MD simulations, namely calculating the free energy of binding, was also implemented to support the study. Consistent results were derived according to such conventional methods, which guaranteed the finding revealed by the geometrical characterization.

This study will lead to a deeper understanding of EGFR mutation-induced drug resistance in lung cancer treatments and will promote the design of genotype-determined therapies or drugs. Our work also has limitations. Although our results are self-consistent and agree with earlier experiments by other researchers, this work is based on computer simulations only. In further research, experimental methods are necessary to further validate the results and reveal the mechanism of drug resistance. We hope this work can provide theoretical guidance for future experiments and clinical applications. Moreover, the geometrical characterization of molecular interfaces proposed in this study will facilitate other studies for feature extraction or interface modeling.

Fig. 6 Average center-of-mass distance of interfacial atom pairs along the MD trajectory for EGFR- RTK heterodimers (smoothed by moving-averaging), with the mean for each curve shown as a dashed line. (**a**) Comparison between WT EGFR-c-Met and L858R-c-Met dimers. (**b**) Comparison between WT EGFR-ErbB2 and L858R- ErbB2 dimers. (**c**) Comparison between WT EGFR-IGF-1R and L858R- IGF-1R dimers. (**d**) Comparison between L858R-c-Met and L858R-T790M-c-Met dimers. (**e**) Comparison between L858R -ErbB2 and L858R-T790M-ErbB2 dimers. (**f**) Comparison between L858R-IGF-1R and L858R-T790M- IGF-1R dimers

Fig. 7 Computational binding free energies of EGFR-RTK heterodimers. EGFR variants include wild type EGFR, mutant L858R and mutant L858R-T790M. RTK partners include c-Met, ErbB2 and IGF- 1R. Bar heights indicate the average binding free energies of these heterodimers, and the errors for each bar are also presented

Methods

Modeling EGFR mutants and constructing EGFR-RTK heterodimers

The crystal structure of a homodimer of WT EGFR tyrosine kinase (TK) domains was collected from the Protein Data Bank (PDB) [48, 49]. PDB:2GS2, an X-ray structure with resolution of 2.80 Å was selected as template. In what follows, we abbreviated the TK domain to simplify the presentation. Residues at positions 672 to 994 on chain A of 2GS2 in original PDB file were selected and represented as indices 1−323 for following procedures. *Rosetta* was used to model the L858R and L858R-T790M mutant structures using the monomer WT EGFR as a template [50]. Some regions on the structure are not resolved in the X-ray diffraction, and residues in these regions do not have detailed coordinates in the original PDB file. For PDB:2GS2, residues at positions 723−725 and 967−981 in the original PDB file are unresolved in the crystal structure. Thus, prior to the modeling, we generated the coordinates of the unresolved residues in the crystal structure according to *Rosetta comparative modeling* (CM) protocol. To achieve this, we prepared the fragment files based on the fragment library in *Rosetta*, [51] where proteins can decompose into 9mer and 3mer fragment files. *Rosetta* will rank the generated structures according to the energy, a lower energy corresponds to a more stable state. Our refined template structure (with coordinates of unresolved residues generated) was derived based on this principle (*extract_pdb* protocol in *Rosetta*). Then we use the high-resolution *ddg_monomer* (HRDM) protocol to generate the structures of mutants L858R and L858R-T790M, where L858R means a point mutation at the 858th position from Leucine (L) to Arginine (R) and T790M represents a replacement from Threonine (T) to Methionine (M) at the 790th position. In such modeling, the Harmonic distance constraints were applied for pre-minimization, and all the C-alpha atoms in both backbone and side chains were restricted within 9 Angstroms. Default parameters in *score12* of *Rosetta* were used. Weights of repulsive term were increased to 100% with three rounds of minimization executed.

Early research by Zhang et al. [49] showed that a damaged/blocked EGFR can be re-activated in the signaling pathway if there are an intact C-lobe in EGFR and an RTK-partner with an intact N-lobe/ATP site. Accordingly, we hypothesize that one possible drug-resistance mechanism for a drug-blocked EGFR is providing its C-lobe face to interact with the N-lobe faces of other RTK partners (re-activation of EGFR signaling). Driven by this idea, we attempted to form the EGFR-partner dimers first in our simulation studies. As crystal structures of these dimers are not available, we modeled them using a template of similar structures and the homology modeling techniques. The computational methods applied here could bring more uncertainty. Thus, we used mature techniques that have played an important role in virtual screening and drug design. The structures produced using proper templates and these techniques have been broadly applied in molecular and structural biology research, and proved to be reliable.

Referring to the asymmetric EGFR-EGFR dimer structure, the "Align" function of Chimera [52] was employed. Based on sequence-alignments and coordinate-superposing methods, we superpose the structure on the template, and generate the new coordinates to achieve spatial proximities. In this study, we used the ErbB2 homodimer PDB ID: 3PP0 [53], which provide a crystal structure with X-ray diffraction at the resolution of 2.25 Å as the template. In PDB:3PP0, N-lobe of chain A faces the C-lobe of chain B. Thus, the RTK structures were superposed to chain A of the 3PP0 template using the "align" function to provide the N-lobe for interactions, and the EGFR variants (wild-type and mutants) structures were superposed to chain B of 3PP0 to provide the C-lobe. The C-Met structure applied here is from PDB ID:4GG5(2.42 Å resolution) [54], and the IGF-1R structure is from PDB ID:5FXQ (2.30 Å resolution) [55]. Both crystal structures were obtained from X-ray diffraction. Each modeled EGFR-RTK dimer structure was then retained for subsequent energy-minimizations and MD simulations.

Molecular dynamics (MD) simulations

MD simulations are a popular method for studying the physical movements of molecules at the atomic level. The MD simulator considers the initial states, force fields and potential energies in the system, and solves the Newton's equations of motions for every atom based on the known forces and energy [56]. Positions of atoms are updated and recorded at every time step. The position trajectories can provide valuable information about the behavior and functions of molecules. In this study, we performed MD simulations on each EGFR-RTK heterodimer structure using Amber software suite [57], based on the explicit-solvent model. *Ff99SB* and *gaff* force fields were adopted in the simulations. The dimers were computationally solvated within periodic water boxes of TIP3P water model. The minimal distance between each dimer and the walls of its belonging water box was set to 10 Å to achieve efficient simulations. The solvated system was neutralized using Na + and Cl- ions. Such settings were adopted by the tleap module of Amber. After setting up the environment, each system was first energy-minimized by 20,000 cycles prior to the simulations. Then each system underwent the heating (100 ps), density equilibration (100 ps) and constant-pressure equilibration (5 ns) processes before the

production simulation. The ultimate production simulations on the equilibrated structures lasted for 50 ns (ns), each resulting a trajectory of 5000 frames for subsequent analysis.

Characterization of EGFR-RTK interaction in a dimer
Alpha shape modeling and matching rate of interfacial atoms
Alpha Shape modeling is actually a linear approximation method, which reconstructs the surface of an object and reveals its geometrical properties effectively [58]. The theory of Alpha Shape modeling originates from triangulation algorithms. The basic criterion of a triangulation for a set of point is that no point in the set will be in the circumcircle of any triangles from the triangulation. Figure 8 shows two simple examples of triangulation for a point set {**A**, **B**, **C**, **D**}, where one of them fulfils the criterion and the other does not.

Algorithms are designed to accomplish the idea. One of the most popular methods is the Delaunay triangulation, whose core idea is to maximize the minimum of all angles in the simplexes based on following determinant [59]:

$$\begin{vmatrix} x_A & y_A & x_A^2 + y_A^2 & 1 \\ x_B & y_B & x_B^2 + y_B^2 & 1 \\ x_C & y_C & x_C^2 + y_C^2 & 1 \\ x_D & y_D & x_D^2 + y_D^2 & 1 \end{vmatrix} > 0 \tag{1}$$

where (x_i, y_i) represent the coordinates of a point, **A**, **B**, **C** or **D**. A positive value of this determinant indicates that point **D** is in the circumcircle of triangle **ABC**. After implementing 3D Delaunay triangulation, the circumscribed sphere of every tetrahedron can be calculated by Alpha Shape algorithm. The algorithm checks the circumscribed spheres of the tetrahedrons so that squared

radii of the spheres are smaller than or equal to a predefined value α and the spheres contain no points in the target set. Ultimately, the corresponding tetrahedrons of qualified spheres form an alpha shape.

In this study, we consider all the heavy atoms in a dimer as the point set to be modeled. To simulate the differences of mass between various kinds of atoms, we adopted the weighted Alpha Shape method that takes both location and weight into consideration [45]. Here we define two points, $p_1 = (p_1', p_1'')$ and $p_2 = (p_2', p_2'')$, for instance, where p_1' and p_2' represent the locations, and and stand for weights of the points. We define $dist(p_1', p_2')$ as the Euclidean distance between the two points, then the algorithm compares this distance and the sum of weights as follows:

$$dist\left(p_1', p_2'\right) = p_1'' + p_2'' \tag{2}$$

$$dist\left(p_1', p_2'\right) > p_1'' + p_2'' \tag{3}$$

Equation 2 guarantees two orthogonal points and eq. 3 two sub-orthogonal points. For a predefined real value α, a point can be defined as:

$$P_\alpha = \left(p', p'' + \alpha\right) \tag{4}$$

For a satisfied weighted Alpha Shape, each tetrahedrons has a point that is orthogonal to the points that are relevant to the vertices of the tetrahedron, and sub-orthogonal to all the other points. In this study, we employed the Computational Geometry Algorithm Library (CGAL) to compute weighted Alpha Shape models of each MD frame [60]. Value of α was set as 0, and weights of atoms were defined as the square values of their Van der Waals radii.
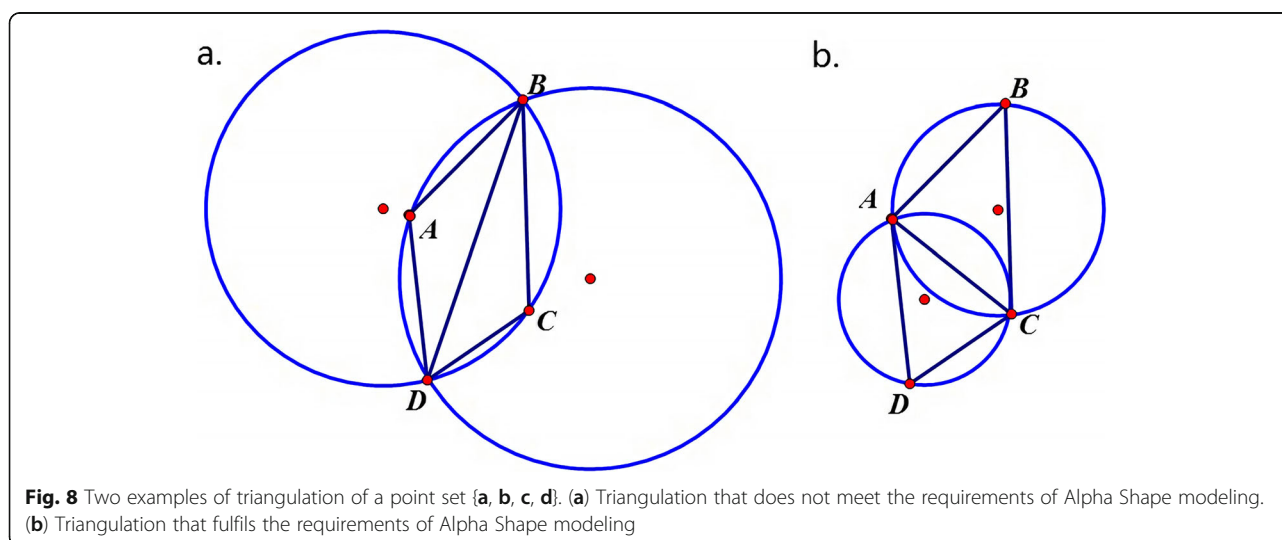

Fig. 8 Two examples of triangulation of a point set {**a**, **b**, **c**, **d**}. (**a**) Triangulation that does not meet the requirements of Alpha Shape modeling. (**b**) Triangulation that fulfils the requirements of Alpha Shape modeling

Zhu *et al. BMC Molecular and Cell Biology*   (2021) 22:34

Page 11 of 13

We are more interested in the atoms at the interaction sites in a dimer. Alpha Shape modeling reveals the atoms on the geometrical surface of a protein or complex, and we derived the interfacial atoms as follows. First, we extracted the surface atoms of a dimer, defined as point set M, using the method. Then we calculated the Alpha Shapes of each monomer EGFR or RTK individually, which resulted surface point sets A and B. The interfacial atom set X, coupled with those for the individual chains (XA and XB), can be derived according to the following equation:

$$\begin{cases} X = (A \cup B) - M \\ X_A = X \cap A \\ X_B = X \cap B \end{cases} \qquad (5)$$

To characterize the interactions in a dimer based on the extracted Alpha Shapes, the convex/concave degree of interfacial atoms, which can simply be indicated by solid angles, were used [61]. For a tetrahedron with vertices **A**, **B**, **C** and **D**, the solid angle of **A** can be expressed by the dihedral angle between triangle ABC and ABD (ϕ), that between triangle ABC and ACD (ϕ_{AB}), and that between triangle ABD and ACD (ϕ_{AB}). The overall solid angle is then expressed in Eq. (8), where index i represents any tetrahedron that is topped with vertex **A** and the sum of values for all tetrahedrons is finally rescaled into the rage of [− 1, 1]. Ω indicates the average curvature at vertex **A**, with a positive value showing a convex shape and a negative value a concave shape, defined as

$$\Omega_A = \cos\left(\frac{\sum_i \left(\phi_i^{AC} + \phi_i^{AB} + \phi_i^{AD} - \pi\right)}{4}\right) \qquad (6)$$

Subsequently, the interactions in dimers can be estimated by the matching rate of such atomic solid angles at the interaction sites of the two chains. For a pair of atoms (**A** and **B**) sharing an edge in any tetrahedron that connects the two interaction sites, we calculated their solid angles (Ω_A and Ω_B). Then whether the pair of atoms are matched or not, (A, B), can be inferred according to the following equation:

$$f(A, B) = \begin{cases} 1, \Omega_A \times \Omega_B < 0 \\ 0, otherwise \end{cases} \qquad (7)$$

A matched pair means a pair of convex-concave complementary atoms, while an unmatched pair means the two atoms both have either convex solids or concave ones. For each EGFR- RTK dimer, we calculated the matching rate of the two chains, namely the ratio of matched pairs over all pairs $\sum_{(i,j)} f(A_i, B_j)/N$, along the MD trajectory to show the interactions.

Center-of-mass distances between interfacial atoms

As stated above, each pair of atoms (**A** and **B**) sharing an edge in any tetrahedron that connects the two interaction sites were considered in this study. Specifically, we calculated the center- of-mass (COM) distance of each pair ($d_{COM}(A, B)$) . Locations of mass centers of the atoms are provided by *cpptraj* tool in *Amber* [62]. The center of mass for a series of particles (masses of x_1, …, x_N) along a specific coordinate axis (   ) can be simply expressed as follows:

$$X_{COM} = \frac{\sum_{i=1}^N m_i x_i}{\sum_{i=1}^N m_i} \qquad (8)$$

For all pairs of atoms connecting the two chains in a dimer, we averaged their COM distances $\overline{(d_{COM}(A, B))}$ and monitored such averaged distance among the MD trajectories. Such simple metric can measure the interaction between the two chains in the dimer and thus reveal the stability of the whole binding system.

Free energy of binding for EGFR-RTK dimers

Conventional computational methods to measure the binding strength of a system include the calculation of free energy of binding, and here we also applied such methods in our study. Free energy of binding is normally calculated based on the thermodynamic cycle, where the energy itself in a solvent environment is obtained indirectly as follows [63]:

$$\Delta G_{bind} = \Delta G_{sol,bind} - \Delta G_{vac,bind} = \Delta G_{sol,dimer} - \Delta G_{sol,chainA} - \Delta G_{sol,chainB}$$

$$(9)$$

Here, $\Delta G_{sol,\ bind}$ is the energy difference between bounded and unbounded states in a solvent environment, while $\Delta G_{vac,\ bind}$ is the difference in the vacuum environment. $\Delta G_{sol,\ dimer}$, $\Delta G_{sol,\ chainA}$ and $\Delta G_{sol,\ chainB}$ are solvation free energies of the dimer, chain A and chain B respectively. A lower value (negative) of binding free energy represents a tighter binding affinity within the binding system. Molecular mechanics generalized Born and surface area (MM/GBSA) continuum solvation is a typical approach for estimating the binding free energy, and we applied such protocol in *Amber* to estimate the interaction strength in different EGFR- RTK heterodimers and to fertilize the geometrical characterization in preceding sections.

Zhu *et al. BMC Molecular and Cell Biology*   (2021) 22:34

Page 12 of 13

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References
1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424. https://doi.org/10.3322/caac.21492.
2. Paez JG, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science. 2004;304(5676):1497–500. https://doi.org/10.1126/science.1099314.
3. Gallardo E, Navarro A, Viñolas N, Marrades RM, Diaz T, Gel B, et al. miR-34a as a prognostic marker of relapse in surgically resected non-small-cell lung cancer. Carcinogenesis. 2009;30(11):1903–9. https://doi.org/10.1093/carcin/bgp219.
4. Normanno N, de Luca A, Bianco C, Strizzi L, Mancino M, Maiello MR, et al. Epidermal growth factor receptor (EGFR) signaling in cancer. Gene. 2006;366(1):2–16. https://doi.org/10.1016/j.gene.2005.10.018.
5. Herbst RS. Review of epidermal growth factor receptor biology. Int J Radiat Oncol Biol Phys. 2004;59(2):S21–6. https://doi.org/10.1016/j.ijrobp.2003.11.041.
6. Rikova K, Guo A, Zeng Q, Possemato A, Yu J, Haack H, et al. Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. Cell. 2007;131(6):1190–203. https://doi.org/10.1016/j.cell.2007.11.025.
7. Sharma SV, Bell DW, Settleman J, Haber DA. Epidermal growth factor receptor mutations in lung cancer. Nat Rev Cancer. 2007;7(3):169–81. https://doi.org/10.1038/nrc2088.
8. Herbst RS, Maddox AM, Rothenberg ML, Small EJ, Rubin EH, Baselga J, et al. Selective oral epidermal growth factor receptor tyrosine kinase inhibitor ZD1839 is generally well-tolerated and has activity in non–small-cell lung cancer and other solid tumors: results of a phase I trial. J Clin Oncol. 2002; 20(18):3815–25. https://doi.org/10.1200/JCO.2002.03.038.
9. Sordella R, Bell DW, Haber DA, Settleman J. Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. Science. 2004; 305(5687):1163–7. https://doi.org/10.1126/science.1101637.
10. Kobayashi S, Boggon TJ, Dayaram T, Jänne PA, Kocher O, Meyerson M, et al. EGFR mutation and resistance of non–small-cell lung cancer to gefitinib. N Engl J Med. 2005;352(8):786–92. https://doi.org/10.1056/NEJMoa044238.
11. Ghosh A, Yan H. Hydrogen bond analysis of the EGFR-ErbB3 heterodimer related to non-small cell lung cancer and drug resistance. J Theor Biol. 2019; 464:63–71. https://doi.org/10.1016/j.jtbi.2018.12.035.
12. Yun C-H, Mengwasser KE, Toms AV, Woo MS, Greulich H, Wong KK, et al. The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. Proc Natl Acad Sci. 2008;105(6):2070–5. https://doi.org/10.1073/pnas.0709662105.
13. Zou B, Lee VH, Yan H. Prediction of sensitivity to gefitinib/erlotinib for EGFR mutations in NSCLC based on structural interaction fingerprints and multilinear principal component analysis. BMC Bioinformatics. 2018;19(1):1–13.
14. 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. 2007;104(52): 20932–7. https://doi.org/10.1073/pnas.0710370104.
15. Engelman JA, Zejnullahu K, Mitsudomi T, Song Y, Hyland C, Park JO, et al. MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. science. 2007;316(5827):1039–43. https://doi.org/10.1126/science.1141478.
16. D. D. Wang, L. Ma, M. P. Wong, V. H. Lee, and H. Yan, "Contribution of EGFR and ErbB-3 heterodimerization to the EGFR mutation-induced gefitinib-and erlotinib-resistance in non-small-cell lung carcinoma treatments," *PloS one*, vol. 10, no. 5, p. e0128360, 2015.
17. Ahmad T, Farnie G, Bundred NJ, Anderson NG. The mitogenic action of insulin-like growth factor I in normal human mammary epithelial cells requires the epidermal growth factor receptor tyrosine kinase. J Biol Chem. 2004;279(3):1713–9. https://doi.org/10.1074/jbc.M306156200.
18. Jones HE, Goddard L, Gee JMW, Hiscox S, Rubini M, Barrow D, et al. Insulin-like growth factor-I receptor signalling and acquired resistance to gefitinib (ZD1839; Iressa) in human breast and prostate cancer cells. Endocr Relat Cancer. 2004;11(4):793–814. https://doi.org/10.1677/erc.1.00799.
19. F. Cappuzzo et al., "Increased MET gene copy number negatively affects survival of surgically resected non–small-cell lung cancer patients," *Journal of Clinical Oncology*, vol. 27, no. 10, p. 1667, 2009.
20. Jo M, Stolz DB, Esplen JE, Dorko K, Michalopoulos GK, Strom SC. Cross-talk between epidermal growth factor receptor and c-met signal pathways in transformed cells. J Biol Chem. 2000;275(12):8806–11. https://doi.org/10.1074/jbc.275.12.8806.
21. E. Ortiz-Zapater et al., "MET-EGFR dimerization in lung adenocarcinoma is dependent on EGFR mtations and altered by MET kinase inhibition," *PLoS One*, vol. 12, no. 1, p. e0170798, 2017.
22. Tanizaki J, Okamoto I, Sakai K, Nakagawa K. Differential roles of trans-phosphorylated EGFR, HER2, HER3, and RET as heterodimerisation partners of MET in lung cancer with MET amplification. Br J Cancer. 2011;105(6):807–13. https://doi.org/10.1038/bjc.2011.322.
23. Wheeler DL, Huang S, Kruser TJ, Nechrebecki MM, Armstrong EA, Benavente S, et al. Mechanisms of acquired resistance to cetuximab: role of HER (ErbB) family members. Oncogene. 2008;27(28):3944–56. https://doi.org/10.1038/onc.2008.19.
24. M.-L. I. Harwardt et al., "Single-Molecule Super-Resolution Microscopy Reveals Heteromeric Complexes of MET and EGFR upon Ligand Activation," *International Journal of Molecular Sciences*, vol. 21, no. 8, p. 2803, 2020.
25. J. Knowles and Z. Gechtman, "Probing EGFR, HER2, and c-Met Protein-Protein Interactions Using an Antibody Array," in *Molecular Cancer Therapeutics*, 2013, vol. 12, no. 11: AMER ASSOC CANCER RESEARCH 615 CHESTNUT ST, 17TH FLOOR, PHILADELPHIA, PA … .
26. R. Lee et al., "T6 MET targeted therapy in lung adenocarcinoma: does 'resistant'EGFR make a MET-responsive dimer?," ed: BMJ publishing group ltd, 2015.
27. Nahta R, Yuan LX, Zhang B, Kobayashi R, Esteva FJ. Insulin-like growth factor-I receptor/human epidermal growth factor receptor 2 heterodimerization contributes to trastuzumab resistance of breast cancer cells. Cancer Res. 2005;65(23):11118–28. https://doi.org/10.1158/0008-5472.CAN-04-3841.
28. Iyer G, Price J, Bourgeois S, Armstrong E, Huang S, Harari PM. Insulin growth factor 1 like receptor (IGF-1R). BMC Cancer. 2016;16(1):1–11.
29. Oliveira S, Schiffelers R, Storm G, Henegouwen P, Roovers R. Crosstalk between epidermal growth factor receptor-and insulin-like growth factor-1 receptor signaling: implications for cancer therapy. Curr Cancer Drug Targets. 2009;9(6): 748–60.
30. Morgillo F, Woo JK, Kim ES, Hong WK, Lee H-Y. Heterodimerization of insulin-like growth factor receptor/epidermal growth factor receptor and

Zhu *et al. BMC Molecular and Cell Biology*   (2021) 22:34

Page 13 of 13

induction of survivin expression counteract the antitumor action of erlotinib. Cancer Res. 2006;66(20):10100–11. https://doi.org/10.1158/0008-5472.CAN-06-1684.

31. M. A. Becker and D. Yee, "Crosstalk Between Insulin-like Growth Factor (IGF) and Epidermal Growth Factor (EGF) Receptors," in *EGFR Signaling Networks in Cancer Therapy*: Springer, 2008, pp. 147–160.

32. Tang Z, du R, Jiang S, Wu C, Barkauskas DS, Richey J, et al. Dual MET–EGFR combinatorial inhibition against T790M-EGFR-mediated erlotinib-resistant lung cancer. Br J Cancer. 2008;99(6):911–22. https://doi.org/10.1038/sj.bjc.6604559.

33. Xu H, Stabile LP, Gubish CT, Gooding WE, Grandis JR, Siegfried JM. Dual blockade of EGFR and c-met abrogates redundant signaling and proliferation in head and neck carcinoma cells. Clin Cancer Res. 2011;17(13):4425–38. https://doi.org/10.1158/1078-0432.CCR-10-3339.

34. Berasain C, Ujue Latasa M, Urtasun R, Goñi S, Elizalde M, Garcia-Irigoyen O, et al. Epidermal growth factor receptor (EGFR) crosstalks in liver cancer. Cancers. 2011;3(2):2444–61. https://doi.org/10.3390/cancers3022444.

35. A. Dixit and G. M. Verkhivker, "Computational modeling of allosteric communication reveals organizing principles of mutation-induced signaling in ABL and EGFR kinases," *PLoS computational biology,* vol. 7, no. 10, p. e1002179, 2011.

36. Ma L, Wang DD, Huang Y, Wong MP, Lee VH, Yan H. Decoding the EGFR mutation-induced drug resistance in lung cancer treatment by local surface geometric properties. Comput Biol Med. 2015;63:293–300. https://doi.org/10.1016/j.compbiomed.2014.06.016.

37. L. Ma, B. Zou, and H. Yan, "Identifying EGFR mutation-induced drug resistance based on alpha shape model analysis of the dynamics," *Proteome science,* vol. 14, no. 1, p. 12, 2016.

38. Shan Y, Arkhipov A, Kim ET, Pan AC, Shaw DE. Transitions to catalytically inactive conformations in EGFR kinase. Proc Natl Acad Sci. 2013;110(18):7270–5. https://doi.org/10.1073/pnas.1220843110.

39. Shan Y, Eastwood MP, Zhang X, Kim ET, Arkhipov A, Dror RO, et al. Oncogenic mutations counteract intrinsic disorder in the EGFR kinase and promote receptor dimerization. Cell. 2012;149(4):860–70. https://doi.org/10.1016/j.cell.2012.02.063.

40. M. Z. Tamirat, K. J. Kurppa, K. Elenius, and M. S. Johnson, "Deciphering the structural effects of activating EGFR somatic mutations with molecular dynamics simulation," *JoVE (Journal of Visualized Experiments),* no. 159, p. e61125, 2020.

41. M. Z. Tamirat, M. Koivu, K. Elenius, and M. S. Johnson, "Structural characterization of EGFR exon 19 deletion mutation using molecular dynamics simulation," *PLoS one,* vol. 14, no. 9, p. e0222814, 2019.

42. Ji H, Li D, Chen L, Shimamura T, Kobayashi S, McNamara K, et al. The impact of human EGFR kinase domain mutations on lung tumorigenesis and in vivo sensitivity to EGFR-targeted therapies. Cancer Cell. 2006;9(6):485–95. https://doi.org/10.1016/j.ccr.2006.04.022.

43. L. V. Sequist *et al.*, "Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors," *Science translational medicine,* vol. 3, no. 75, pp. 75ra26-75ra26, 2011.

44. Ni Z, Wang X, Zhang T, Jin RZ. Molecular dynamics simulations reveal the allosteric effect of F1174C resistance mutation to ceritinib in ALK-associated lung cancer. Comput Biol Chem. 2016;65:54–60. https://doi.org/10.1016/j.compbiolchem.2016.10.005.

45. H. Edelsbrunner, Weighted alpha shapes. University of Illinois at Urbana-Champaign, 1992.

46. Edelsbrunner H, Mücke EP. Three-dimensional alpha shapes. ACM Transactions on Graphics (TOG). 1994;13(1):43–72. https://doi.org/10.1145/174462.156635.

47. Ma L, Wang DD, Huang Y, Yan H, Wong MP, Lee VH. EGFR mutant structural database: computationally predicted 3D structures and the corresponding binding free energies with gefitinib and erlotinib. BMC Bioinformatics. 2015;16(1):1–10.

48. Song Y, DiMaio F, Wang RYR, Kim D, Miles C, Brunette T, et al. "High-resolution comparative modeling with RosettaCM". Structure. 2013;21(10):1735–42.

49. Zhang X, Gureasko J, Shen K, Cole PA, Kuriyan J. An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. Cell. 2006;125(6):1137–49. https://doi.org/10.1016/j.cell.2006.05.013.

50. Simons KT, Kooperberg C, Huang E, Baker D. Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions. J Mol Biol. 1997;268(1):209–25. https://doi.org/10.1006/jmbi.1997.0959.

51. Robetta. "Full-chain Protein Structure Prediction Server." (accessed 12-June, 2019).

52. 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;25(13):1605–12. https://doi.org/10.1002/jcc.20084.

53. Aertgeerts K, Skene R, Yano J, Sang BC, Zou H, Snell G, et al. Structural analysis of the mechanism of inhibition and allosteric activation of the kinase domain of HER2 protein. J Biol Chem. 2011;286(21):18756–65. https://doi.org/10.1074/jbc.M110.206193.

54. Wu K, Ai J, Liu Q, Chen TT, Zhao A, Peng X, et al. Multisubstituted quinoxalines and pyrido [2, 3-d] pyrimidines: synthesis and SAR study as tyrosine kinase c-met inhibitors. Bioorg Med Chem Lett. 2012;22(20):6368–72. https://doi.org/10.1016/j.bmcl.2012.08.075.

55. Degorce SBL, et al. Discovery of a potent, selective, orally bioavailable, and efficacious novel 2-(pyrazol-4-ylamino)-pyrimidine inhibitor of the insulin-like growth factor-1 receptor (IGF-1R). J Med Chem. 2016;59(10):4859–66. https://doi.org/10.1021/acs.jmedchem.6b00203.

56. D. Frenkel and B. Smit, Understanding molecular simulation: from algorithms to applications. Elsevier, 2001.

57. R. M. B. D.A. Case, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. Legrand, P. Li C. and T. L. Lin, R. Luo, B. Madej, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao and P.A. Kollman, "AMBER 2016", 2016.

58. F. Bernardini and C. L. Bajaj, "Sampling and reconstructing manifolds using alpha-shapes," 1997.

59. Guibas L, Stolfi J. Primitives for the manipulation of general subdivisions and the computation of Voronoi. ACM transactions on graphics (TOG). 1985;4(2):74–123. https://doi.org/10.1145/282918.282923.

60. CGAL. "Computational Geometry Algorithms Library. 1996. [online] Available: https://www.cgal.org/.

61. Zhou W, Yan H, Hao Q. Analysis of surface structures of hydrogen bonding in protein–ligand interactions using the alpha shape model. Chem Phys Lett. 2012;545:125–31. https://doi.org/10.1016/j.cplett.2012.07.016.

62. Roe DR, Cheatham TE III. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. J Chem Theory Comput. 2013;9(7):3084–95. https://doi.org/10.1021/ct400341p.

63. Srivastava HK, Sastry GN. Molecular dynamics investigation on a series of HIV protease inhibitors: assessing the performance of MM-PBSA and MM-GBSA approaches. J Chem Inf Model. 2012;52(11):3088–98. https://doi.org/10.1021/ci300385h.