

Cite this: *Mol. BioSyst.*, 2016,  
12, 1552

## Selectivity profile of afatinib for EGFR-mutated non-small-cell lung cancer†

Debby D. Wang,<sup>ab</sup> Victor H. F. Lee,<sup>\*c</sup> Guangyu Zhu,<sup>d</sup> Bin Zou,<sup>b</sup> Lichun Ma<sup>b</sup> and Hong Yan<sup>b</sup>

EGFR-mutated non-small-cell lung cancer (NSCLC) has long been a research focus in lung cancer studies. Besides reversible tyrosine kinase inhibitors (TKIs), new-generation irreversible inhibitors, such as afatinib, embark on playing an important role in NSCLC treatment. To achieve an optimal application of these inhibitors, the correlation between the EGFR mutation status and the potency of such an inhibitor should be decoded. In this study, the correlation was profiled for afatinib, based on a cohort of patients with the EGFR-mutated NSCLC. Relying on extracted DNAs from the paraffin-embedded tumor samples, EGFR mutations were detected by direct sequencing. Progression-free survival (PFS) and the response level were recorded as study endpoints. These PFS and response values were analyzed and correlated to different mutation types, implying a higher potency of afatinib to classic activation mutations (*L858R* and *deletion 19*) and a lower one to *T790M*-related mutations. To further bridge the mutation status with afatinib-related response or PFS, we conducted a computational study to estimate the binding affinity in a mutant–afatinib system, based on molecular structural modeling and dynamics simulations. The derived binding affinities were well in accordance with the clinical response or PFS values. At last, these computational binding affinities were successfully mapped to the patient response or PFS according to linear models. Consequently, a detailed mutation–response or mutation–PFS profile was drafted for afatinib, implying the selective nature of afatinib to various EGFR mutants and further encouraging the design of specialized therapies or innovative drugs.

Received 18th January 2016,  
Accepted 26th February 2016

DOI: 10.1039/c6mb00038j

[www.rsc.org/molecularbiosystems](http://www.rsc.org/molecularbiosystems)

## 1 Introduction

Somatic mutations in the epidermal growth factor receptor (EGFR) have been identified in a subset of non-small-cell lung cancer (NSCLC),<sup>1–4</sup> defining a novel NSCLC subgroup that exquisitely depends on the EGFR signaling network.<sup>5–7</sup> The EGFR signaling, which plays a central role in cancer progression and malignancy,<sup>5,8</sup> is briefly outlined in Fig. 1a. A multitude of EGFR mutations can induce aberrant signaling<sup>9</sup> by promoting or stabilizing the catalytic activity of the tyrosine kinase (TK) domain,<sup>1,2</sup> and this is well validated as a cancer-prompting mechanism.<sup>1,2,8</sup> Such mutations normally correspond to exons

18 to 21 in the EGFR gene and cluster around the ATP-binding pocket of the TK domain.<sup>6,10</sup> The most frequently occurring ones, termed classical activating mutations, include in-frame deletions around the residues 746 to 750 of exon 19 (45–50% of all), and the Leu858Arg (*L858R*) point mutation in exon 21 (40–45% of all).<sup>11–13</sup>

A number of therapeutic agents, which reversibly inhibit the TK domain of EGFR (Fig. 1b), have been developed and broadly used in NSCLC treatment.<sup>8–10,14</sup> First-generation TK inhibitors (TKIs) are represented by gefitinib and erlotinib,<sup>1–3,9,15</sup> and showed high anti-tumor activity in patients with the EGFR-mutated NSCLC.<sup>14,16</sup> In particular for those with classical activating mutations, the use of EGFR TKIs led to a high objective response rate, and prolonged progression-free survival (PFS) and overall survival (OS).<sup>2,16,17</sup> However, these patients eventually developed an acquired resistance to the TKIs, commonly characterized by the gatekeeper mutation *T790M* (corresponding to the acquired TKI resistance in 50–60% of patients).<sup>14,16,18,19</sup> A well-established explanation is that steric blocking of the binding with TKIs has been caused by this mutation,<sup>20</sup> which directly results in a low binding affinity between this mutant and the TKIs.<sup>18–20</sup> In this regard, the binding affinity between an EGFR TK mutant and a TKI can be a valuable indicator of the sensitivity/resistance to TKIs.<sup>2,10,18,20</sup>

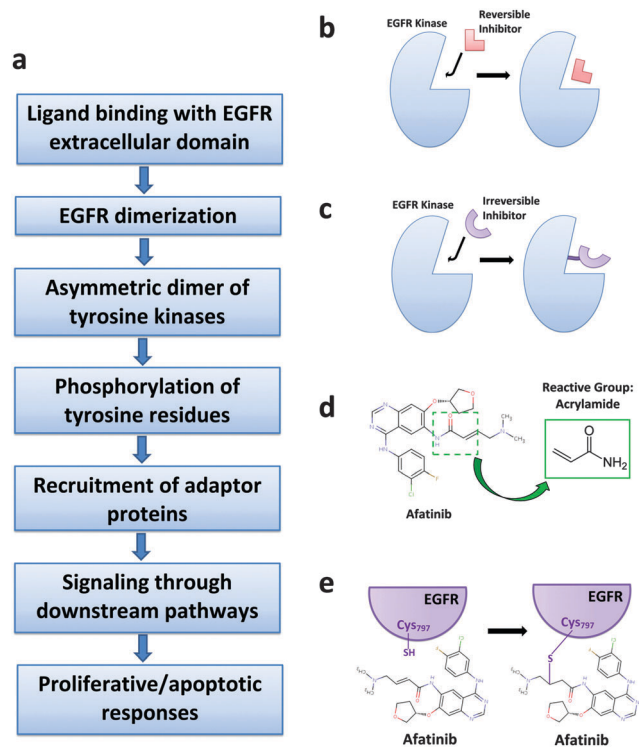
<sup>a</sup> Caritas Institute of Higher Education, 18 Chui Ling Road, New Territories, Hong Kong, China. E-mail: damwang6-c@my.cityu.edu.hk; Tel: +852-96462175

<sup>b</sup> Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong, China

<sup>c</sup> Department of Clinical Oncology, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Pokfulam, Hong Kong, China. E-mail: vhflee@hku.hk; Tel: +852-22554352

<sup>d</sup> Department of Biology and Chemistry, City University of Hong Kong, Kowloon, Hong Kong, China. E-mail: guangzhu@cityu.edu.hk; Tel: +852-34426857

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c6mb00038j



**Fig. 1** EGFR signaling and EGFR-targeted TKIs. (a) A profile of EGFR signaling. (b) EGFR-targeted reversible TKIs. (c) EGFR-targeted irreversible TKIs. (d) The composition of a second-generation TKI (afatinib) and its primary reactive group (acrylamide). The chemical composition structure is from DrugBank (DB08916). (e) The Michael reaction between afatinib and the cysteine (Cys797) residue of an EGFR TK domain. The chemical composition structure is from DrugBank (DB08916).

Based on the binding affinity, both experimental (direct binding assays and surface plasmon resonance (SPR) techniques)<sup>8,19,20</sup> and computational (molecular modeling and simulations)<sup>10,21–25</sup> studies have contributed to a substantially improved understanding of TKI-resistant mechanisms in cancer promotion. The knowledge, meanwhile, triggers the rapid development of innovative agents for the treatment of EGFR-mutated NSCLC.<sup>8,10,14</sup>

Afatinib (BIBW 2992; *N*-[4-[(3-chloro-4-fluorophenyl)amino]-7-[[[(3*S*)-tetrahydro-3-furanyl]oxy]-6-quinazoliny]-4-(dimethylamino)-2-butenamide]<sup>14,16,26,27</sup> is a second-generation EGFR-family inhibitor, which covalently binds to an EGFR-family receptor and irreversibly blocks the downstream signaling (Fig. 1c).<sup>5,8,14</sup> The composition of afatinib and its primary reactive group (acrylamide),<sup>16,28,29</sup> capable of Michael addition to the conserved cysteine residue within an EGFR TK domain (Cys797),<sup>16,30,31</sup> are presented in Fig. 1d and e. The preclinical *in vitro* and *in vivo* activity profiles of afatinib guaranteed its high potency against NSCLCs harboring the classic activation mutations (*L858R* and *deletion 19*),<sup>8,14,16</sup> and its potency, albeit lower, against those having the gatekeeper mutation (*T790M*).<sup>14,26,32</sup> Similarly, the binding affinity of a mutant with afatinib can be a simplified estimate of afatinib potency.<sup>8,33,34</sup> The studies of mutant–afatinib affinities can also be supported by experimental assays (SPR and biomolecular interaction analysis (BIA) techniques)<sup>8,26</sup> or computational simulations.<sup>10,23</sup> Importantly, these studies further

prompt the investigation of afatinib in the treatment of patients with EGFR-mutated NSCLC.

In order to achieve an optimal use of afatinib in EGFR-mutated NSCLC, an exploration of its selectivity and binding mode with an EGFR mutant is of paramount importance.<sup>8,14,16</sup> Specifically, the correlation, which is between the binding mode in a mutant–afatinib complex and the PFS/response in patient treatments, should be dissected. In our study, for a cohort of patients with the EGFR-mutated NSCLC, their tumor progression, characterized by PFS or objective responses, was evaluated after a prescription of oral afatinib. These observations were then analyzed in detail. Moreover, computational modeling and simulations<sup>10,14</sup> were implemented to investigate the involved EGFR TK mutants and their binding modes with afatinib, from a molecular perspective. Importantly, such binding modes correlated well with the afatinib-related PFS/responses.

## 2 Results and discussion

### 2.1 Patient characteristics

This study was performed on a cohort of 18 patients, who were diagnosed as having advanced (stage IV) NSCLC, in the Queen Mary Hospital in Hong Kong. These patients were treated and followed up during December 2010 to October 2014, and their demographic characteristics are described in Table 1. Specifically, eight (44%) of them were females and ten (56%) were males, with their age varying from 48 to 68. Merely four patients (22%) were heavy smokers, and the others (78%) were non-smokers. All of them suffered from adenocarcinoma. Furthermore, among these patients, fifteen (83%) were prescribed with afatinib as a first-line treatment, and three (17%) as a third-line treatment with confirmed *T790M* and *exon 19 deletion* mutations by tumor re-biopsy immediately before starting afatinib. For those three patients, gefitinib or elotinib was used as their first-line TKI and systemic platinum-based chemotherapy was employed as the second-line treatment.

### 2.2 EGFR mutation screening and afatinib inhibition profiles

For each individual patient, DNA was extracted from the formalin-fixed paraffin-embedded (FFPE) tumor tissues.<sup>35,36</sup> Based on the polymerase chain reaction (PCR) amplification of exons 18 to 21 in the EGFR TKI-binding domain, direct sequencing was implemented to reveal the EGFR somatic mutations.<sup>14,35</sup> Table 2 presents the mutation profile of our patient group. To provide a thorough analysis, we examined the mutations in terms of positions, subtypes and their distributions. As shown in Fig. 2a, the patients primarily harbor exon mutations in positions 19 to 21. Furthermore, single-point (one exon) mutations occupy the majority, among which *deletion 19* and *L858R* mutations are the most important (Fig. 2b). In particular, all the multi-point mutations are *T790M*-involved (Fig. 2b). For individual mutation types, the detailed distribution is displayed in Fig. 2c, in which the frequently-occurring types are *delE746\_A750* and *L858R*.

**Table 1** Characteristics of our advanced-NSCLC patients including age, sex, smoking history, performance status, histology and the TKI used in the first-line treatment

Patient no.	Age	Sex	Smoking history	Performance status	Histology	First-line TKI
1	60	Male	Non-smoker	1	Adenocarcinoma	Erlotinib
2	50	Male	Heavy smoker	2	Adenocarcinoma	Erlotinib
3	65	Male	Non-smoker	1	Adenocarcinoma	Afatinib
4	64	Female	Non-smoker	1	Adenocarcinoma	Afatinib
5	68	Male	Non-smoker	1	Adenocarcinoma	Afatinib
6	57	Female	Non-smoker	1	Adenocarcinoma	Afatinib
7	52	Female	Non-smoker	1	Adenocarcinoma	Afatinib
8	63	Male	Non-smoker	2	Adenocarcinoma	Afatinib
9	65	Male	Heavy smoker	1	Adenocarcinoma	Afatinib
10	48	Female	Heavy smoker	2	Adenocarcinoma	Afatinib
11	63	Male	Non-smoker	1	Adenocarcinoma	Afatinib
12	51	Female	Non-smoker	1	Adenocarcinoma	Afatinib
13	62	Male	Non-smoker	1	Adenocarcinoma	Afatinib
14	50	Female	Non-smoker	2	Adenocarcinoma	Afatinib
15	65	Male	Heavy smoker	2	Adenocarcinoma	Afatinib
16	60	Female	Non-smoker	2	Adenocarcinoma	Afatinib
17	61	Female	Non-smoker	2	Adenocarcinoma	Afatinib
18	60	Male	Non-smoker	2	Adenocarcinoma	Gefitinib

**Table 2** The profile of EGFR mutations and afatinib efficacies. EGFR mutation types of our patients, respectively, at the DNA level (exons 19 to 21) and the protein level (residues 746 to 861), are listed in this table. Patient objective responses to afatinib and the PFS values (in months) are summarized as well

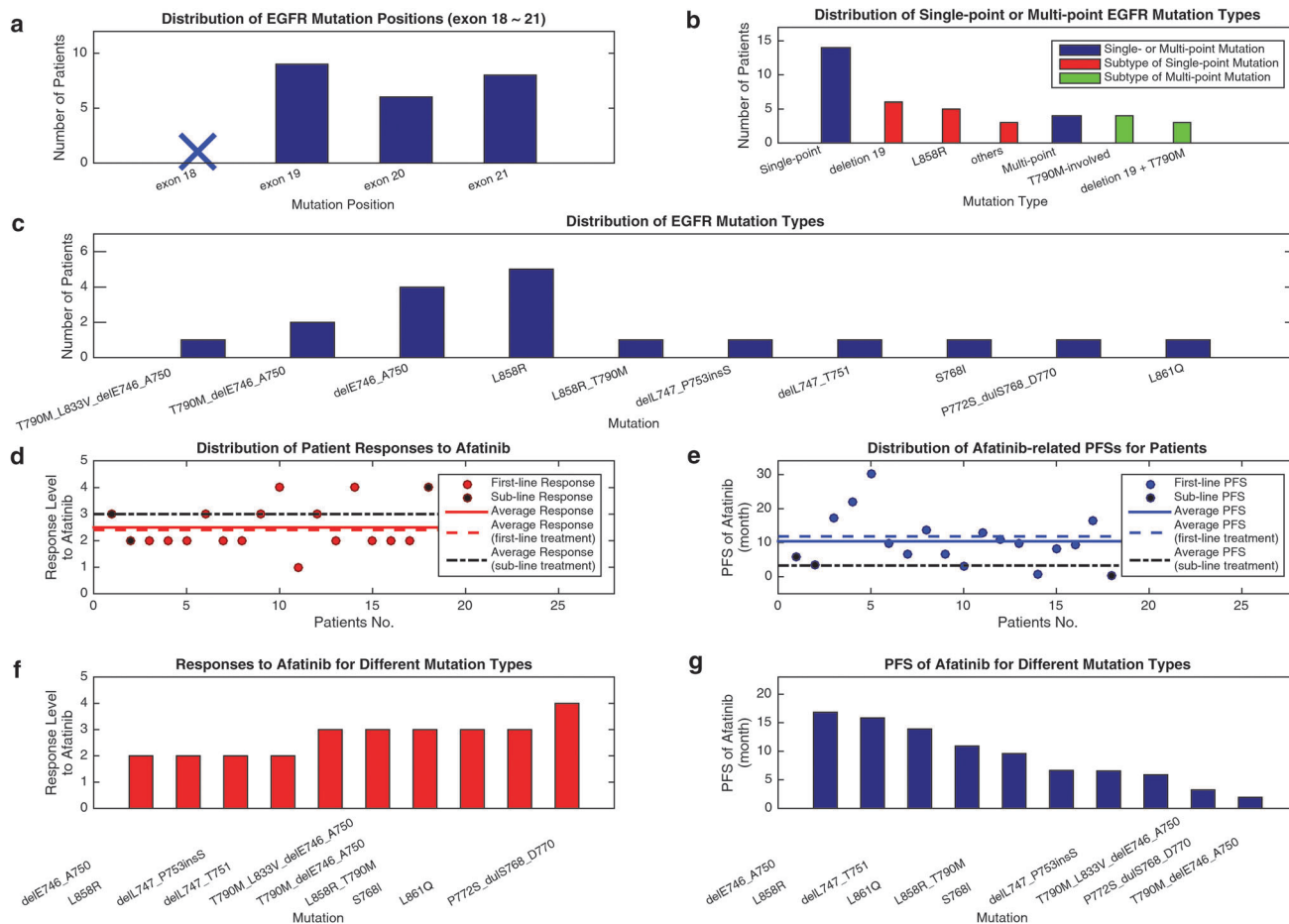
Patient no.	EGFR DNA mutation	EGFR mutation	Best response to afatinib	PFS (month)
1	exon 19 + exon 20 + exon 21	<i>T790M_L833V_delE746_A750</i>	3	5.98
2	exon 19 + exon 20	<i>T790M_delE746_A750</i>	2	3.58
3	exon 19	<i>delE746_A750</i>	2	17.31
4	exon 19	<i>delE746_A750</i>	2	21.82
5	exon 21	<i>L858R</i>	2	30.23
6	exon 20 + exon 21	<i>L858R_T790M</i>	3	9.59
7	exon 19	<i>delL747_P753insS</i>	2	6.57
8	exon 19	<i>delL747_T751</i>	2	13.9
9	exon 20	<i>S768I</i>	3	6.64
10	exon 20	<i>P772S_dulS768_D770</i>	4	3.25
11	exon 21	<i>L858R</i>	1	12.81
12	exon 21	<i>L861Q</i>	3	10.94
13	exon 19	<i>delE746_A750</i>	2	9.92
14	exon 21	<i>L858R</i>	4	0.85
15	exon 21	<i>L858R</i>	2	8.18
16	exon 21	<i>L858R</i>	2	9.2
17	exon 19	<i>delE746_A750</i>	2	16.39
18	exon 19 + exon 20	<i>T790M_delE746_A750</i>	4	0.3

All of our patients received afatinib at their own expense. Majority of them received it at a dose of 40 mg daily except for one patient who started at 50 mg daily, all at the discretion of treating oncologists. Afatinib would be suspended if severe grade 3 or above drug-related adverse events emerged, and be resumed at a reduced dose when a better symptom was observed. Specifically for our patients, the common adverse events included rash, diarrhea, electrolyte disturbance and impaired liver function. Merely three patients (17%) received dose reduction due to severe adverse events.

During the afatinib therapy, PFS was recorded for each individual as the primary endpoint, and the objective response and disease control as the secondary endpoints. The detailed PFS or response list is presented in Table 2, and distribution analyses of such PFS/response values are shown in Fig. 2d–g.

In Fig. 2d, the responses of various patients to afatinib, with an average of 2.5, are shown. Response levels were stratified into four groups (1 = complete response [CR], 2 = partial response [PR], 3 = stable disease [SD], and 4 = progressive disease [PD]), where a lower value represents a better objective response. For those patients who received afatinib as a third-line treatment, the average response level is larger than that of patients having the first-line afatinib. This is possibly due to the *T790M*-involved mutations that all patients with sub-line afatinib harbored, resulting in a lower efficacy of afatinib. Similarly, PFS values concerning different patients are scattered in Fig. 2e, where a higher value denotes a better efficacy of afatinib. Also, patients prescribed with afatinib as a third-line treatment had shorter PFS values than the others who received it as first-line therapy.

To further evaluate the correlation between the afatinib-related response and PFS with an EGFR mutation type, we derived the response or PFS for each mutation. For each patient subgroup harboring the same mutation type, their median response level to afatinib was regarded as the response to this mutation. The responses sorted in an ascending manner are displayed in Fig. 2f, where *deletion 19* and *L858R* mutations correspond to better responses. Meanwhile, PFS values for various mutation types are presented in Fig. 2g, in a descending order. Owing to a large variation of PFS among patients, the PFS or averaged PFS, corresponding to the median response in a patient subgroup, was defined as the PFS for a mutation. As exhibited in Fig. 2g, *delE746\_A750* and *L858R* possess longer PFSs, while *T790M*-involved mutations are lower ranked. Such results are in good agreement with early studies of the EGFR-mutated NSCLC.<sup>8,14,16,26,32</sup>



**Fig. 2** EGFR mutation profile and afatinib efficacy analysis of the patients. (a) EGFR mutation positions (exons 18 to 21) of the NSCLC patients. (b) EGFR mutation subtypes of the NSCLC patients. (c) Specific EGFR mutation types of the NSCLC patients. (d) Responses to afatinib during the TKI therapies, with first-line and sub-line treatments considered. (e) PFS values of afatinib during the therapies, with first-line and sub-line treatments considered. (f) Responses to afatinib regarding different mutation types. (g) PFS values of afatinib regarding different mutation types.

### 2.3 Computational modeling results of mutant-afatinib binding affinities

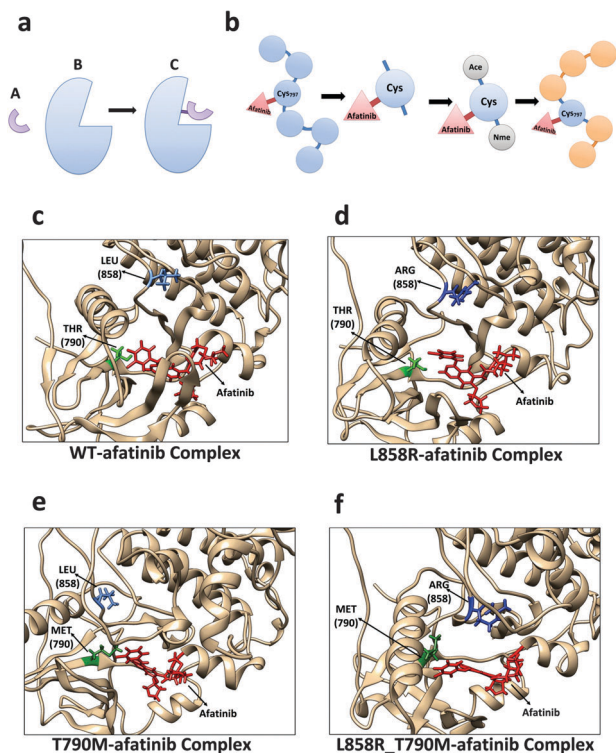
An advanced computational study was carried out to model the mutant-afatinib binding affinities for our patients, and correlate these affinities to the above-mentioned responses or PFS values. This study can assist the construction of a detailed mutation-response or mutation-PFS profile for afatinib. *AMBER* MD simulations<sup>21</sup> and structural modeling were the primary techniques used in this study. Each mutant-afatinib binding affinity can be computationally estimated by the free energy difference between the mutant-afatinib complex (C) and the two isolated molecules (A and B) in solvent environments (Fig. 3a).

#### 2.3.1 Modeling results of EGFR mutant-afatinib complexes.

For each mutant, the structural modeling of its bound complex (C) with afatinib should be accomplished in advance. Our mutant structures were collected from various literature reports<sup>10,37</sup> for preparation. Based on a structural template (PDB:2ITY), computational modeling software, scap,<sup>38</sup> loopy<sup>39</sup> and ROSETTA,<sup>40</sup> were adopted for the structural determination in these studies.

Similarly, a structural template (PDB:4G5P) was downloaded from the PDB for the modeling of EGFR mutant-afatinib complexes. This template contains an EGFR kinase mutant (*T790M*) that is covalently bound with afatinib, and each prepared mutant was aligned to the template prior to the modeling. As afatinib binds to the conserved cysteine residue at position 797 (Cys797) of an EGFR kinase, the afatinib-cysteine block can be regarded as a building block for each mutant-afatinib complex (Fig. 3b). This is well in accordance with *AMBER*, which treats residues as building blocks for protein systems and builds force fields based on pair-wise charges. After deriving the atomic charges of the afatinib-cysteine block, we can insert this block into the corresponding position of each mutant with the assistance of *AMBER*.

For the afatinib-cysteine fragment, capping at its head (Ace) and tail (Nme) was first implemented (Fig. 3b). R.E.D. was subsequently adopted for RESP charge derivation of this capped fragment. Geometry optimization (*GAMBLESS*), MEP calculation and charge fitting (RESP) comprised this calculation. Specifically in the geometry optimization, a strict density convergence criterion ( $1.0 \times 10^{-6}$ ) was used in the direct



**Fig. 3** Computational modeling of EGFR mutant-afatinib complexes and the binding affinities. (a) A covalently bound system ( $A + B \rightarrow C$ ). (b) The modeling of a mutant-afatinib complex based on a structural template. (c) Modeling result of the WT-afatinib complex, with sites 790 (THR), 797 (afatinib-binding site) and 858 (LEU) labeled. (d–f) Several instances of the modeled mutant-afatinib complexes, regarding *L858R*, *T790M* and *L858R\_T790M*, with the mutated residues and afatinib-cysteine blocks labeled.

self-consistent field (SCF) calculation, to achieve a refined geometry. Depending on the derived charges, those of the fragment were extracted, after which the fragment was inserted into each aligned mutant to form its complex with afatinib (Fig. 3b). Such complexes were used in later AMBER simulations.

In Fig. 3c–f, several modeled WT- or mutant-afatinib complexes are comparably presented, regarding wild type (WT) EGFR, *L858R*, *T790M*, and *L858R\_T790M*, with the afatinib-cysteine blocks and mutated residues colored. Before the binding affinity calculation, each complex or mutant was carefully checked for missing residues, and such residues were added based on templates PDB:3GOP/PDB:1M17 as a refinement. These rough models with equal lengths were further structurally optimized using AMBER simulations (30 000 steps), with a QM/MM setting on the afatinib-cysteine blocks.

**2.3.2 MD simulations and parameterization results.** Once we obtained the structures of those mutants and mutant-afatinib complexes, we estimated the mutant-afatinib binding affinities based on MD simulations. As previously elucidated, this binding affinity can be roughly expressed by the free energy difference between the complex (C) and the two isolated molecules (A and B) in a solvent environment.

The relative free energy calculation adopted by AMBER can provide more accurate results for binding systems. In our

studies, the relative free energies of each mutant and the corresponding mutant-afatinib complex were calculated; meanwhile, those of each mutant and its ligand (afatinib) were derived. The binding affinity can be then estimated based on these relative energies.

Prior to the relative free energy calculation, the dynamics of involved systems should be simulated. First, for a mutant and its afatinib-binding complex, the template PDB:4RIW (EGFR-ErbB3 TK heterodimer) was used to combine them as a heterodimer system. Specifically, the mutant and the complex were aligned to the two monomers in the TK heterodimer template, with the assistance of *Chimera*.<sup>41</sup> The system was then computationally solvated in a solvent environment, represented by a truncated octahedron TIP3P water box (10-Å buffer).<sup>21</sup> Two minimization steps, with or without position constraints on the system, were implemented for structural optimization. The convergence of each minimization was examined to guarantee a reliable simulation scenario. In particular, for the afatinib-cysteine block, a QM/MM setting was utilized for a careful treatment of the drug-binding residue. Before the production MD simulations, a series of equilibration operations were adopted for each system. In detail, these operations include, a 50 ps heating from 100 K to 300 K, a 25 ps density equilibration and a 250 ps constant pressure equilibration at 300 K.<sup>10,21</sup> The simulation parameters were selected because of the stable equilibrations that the density, temperature and total energy achieved. Such equilibrations of the *L858R*-involved system are displayed in Fig. 4a–c. Successively, a 2.5 ns production MD simulation was carried out for later free energy calculation. The QM/MM setting of the afatinib-cysteine block remained during the simulation. In order to derive a promising production simulation, we implemented several consecutive short simulations (500 ps for each) and observed the backbone RMSD curve of the involved solute. Trajectory frames at every picosecond, namely 500 frames in each short production simulation, were collected. An example of the backbone RMSD curves corresponding to these short simulations (involving *L858R*) is presented in Fig. 4d, which indicates good stabilization at 2.5 ns. Other examples of the systems, respectively, involving *L858R\_T790M*, *T790M* and WT are presented in Fig. 4e–h, i–l, and m–p.

The calculation of relative free energies of each mutant and afatinib can be similarly implemented. PDB:4G5P was adopted as a template, where afatinib was untied from the covalent association with EGFR. The untied afatinib was structurally refined by adjusting the hydrogen atoms, adding missing atoms, and optimizing its geometry (*GAMMESS* with a strict convergence criterion of  $1.0 \times 10^{-8}$ ). Additionally, R.E.D. was used for the charge derivation of this small molecule, for later AMBER simulations. Each mutant was then aligned to the template to form a non-covalent association with the isolated afatinib. After the computational solvation of such a system, a whole simulation procedure includes two short minimizations, a 50-ps heating, a 25-ps density equilibration, a 250-ps constant pressure equilibration and a 1.5-ns production simulation. The parameter selection was also based on the stabilization of each system (density, temperature, total energy and backbone RMSD of solute).

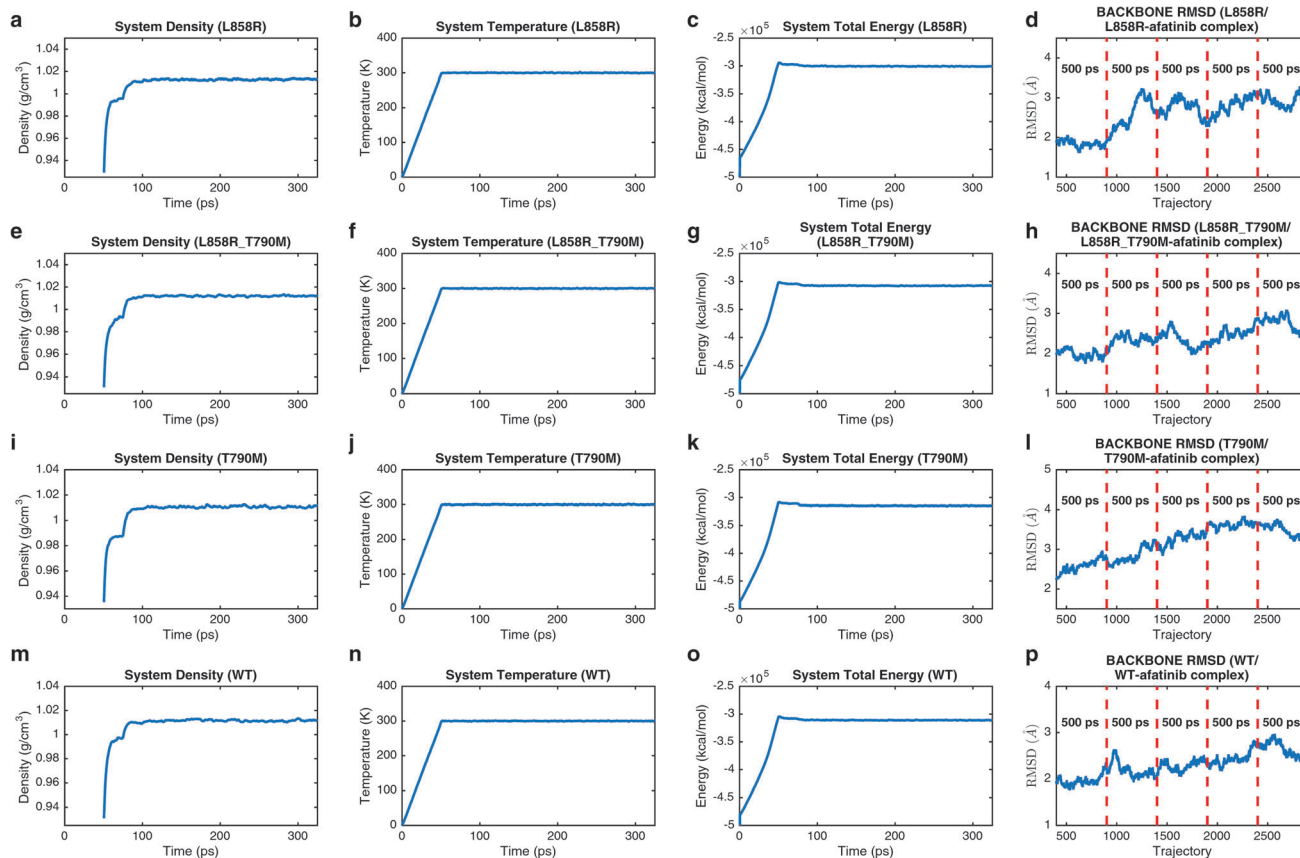


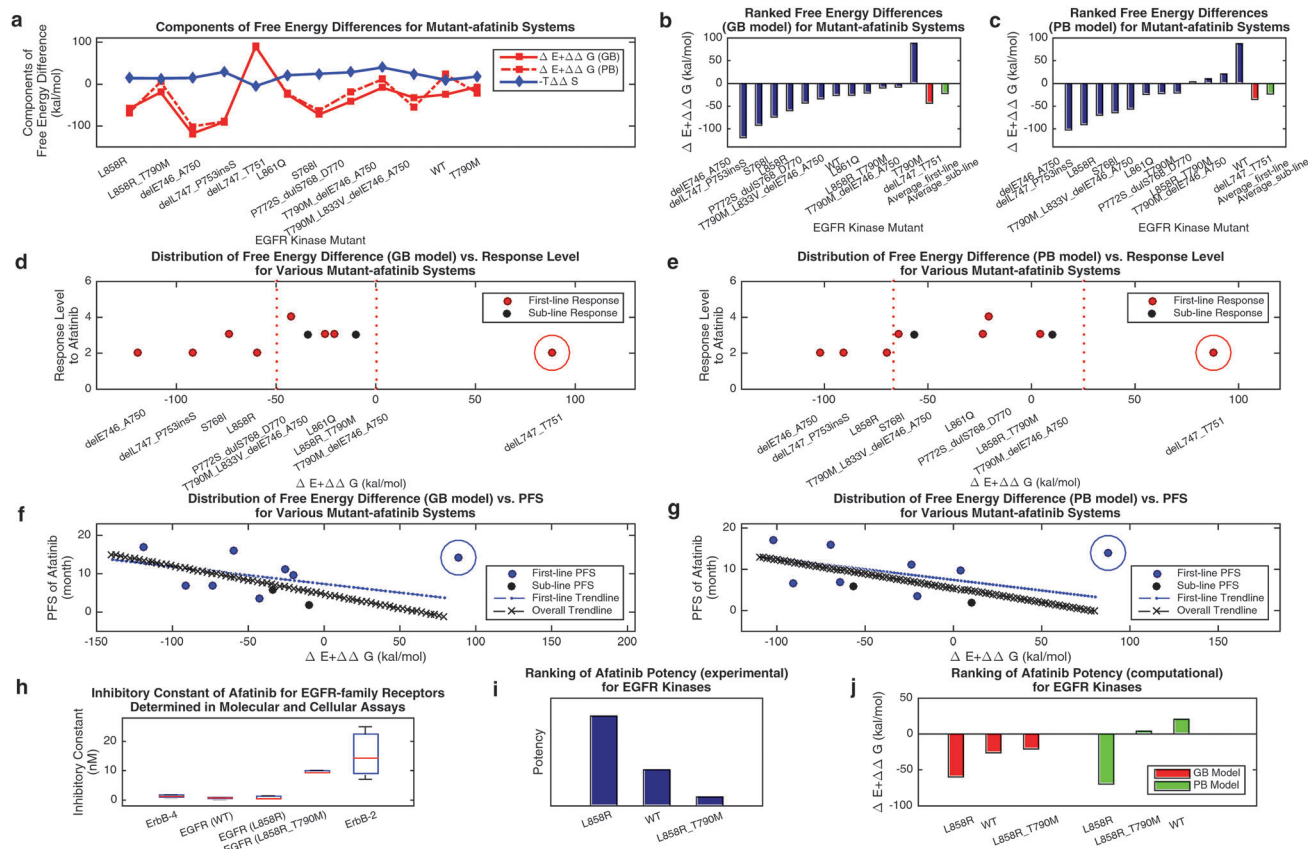
Fig. 4 Simulation stabilization of several systems each containing an EGFR mutant and the mutant–afatinib complex. (a–c) Density, temperature and total energy curves of the system involving *L858R* and its complex with afatinib, before the production simulation. (d) The backbone RMSD curve of the system involving *L858R* and its complex with afatinib, in the 2.5 ns production simulation phase. (e and h) Stabilization curves of the system containing *delE746\_A750* and the *delE746\_A750*–afatinib complex. (i and l) Stabilization curves of the system containing *L858R\_T790M* and the *L858R\_T790M*–afatinib complex. (m and p) Stabilization curves of the system involving *dell747\_P753insS* and the *dell747\_P753insS*–afatinib complex.

**2.3.3 Binding affinity calculation results.** Using the simulation trajectory of each system ( $\{B,C\}$  or  $\{A,B\}$ ), the MM/PB(GB)SA module<sup>26</sup> of *AMBER* calculates the relative free energies of the solvated solutes. Accordingly, both the PB and GB models were implemented in our work for a comparable study. The entropy contribution ( $-T\Delta S$ ) was ignored to ease the computational load in our calculations. In this regard, the objective free energy difference ( $\Delta G_{\text{bind,solv}}$ ) can be simplified to  $\Delta E + \Delta\Delta G_{\text{solv}}$  (Table S1, ESI<sup>†</sup>). In Fig. 5a, such free energy differences and the entropy contributions regarding our mutant–afatinib systems are displayed. Here the entropy terms are similar for these systems, thus they can be reasonably eliminated.

In Fig. 5b and c, we further rank the free energy differences (GB or PB model) in an ascending order. As interpreted previously, a lower free energy difference indicates a better binding affinity. In this ranking list, *L858R* and *deletion 19* mutants are the top ones, while the *T790M*-involved mutants are comparably lower-ranked. Here *dell747\_T751* can be regarded as an outlier due to its high positive value, which may result from the inaccurate structural modeling. This should be improved in our future studies. Moreover, the free energy differences, regarding the mutations harbored by patients with first-line afatinib, were averaged, and so were

those corresponding to sub-line afatinib. As displayed, the first-line average is lower than the sub-line one for both models. These results are in good agreement with our clinical observations of the response levels and PFS values for the corresponding mutations.

Next, Fig. 5d presents the distribution of free energy difference (GB model) *vs.* the response level for all mutation types that emerged in our patients. Regardless of *dell747\_T751*, a lower free energy difference basically corresponds to a better response level (1 = [CR], 2 = [PR], 3 = [SD], and 4 = [PD]). Besides, boundaries can be estimated for different groups of response levels, constructing a rough mutation–response profile for afatinib. Furthermore, we scatter the mutations for a free energy difference–PFS distribution in Fig. 5f, where the points corresponding to first-line PFS values and sub-line PFS values are separately colored. Based on this distribution with the outlier ignored, a linear first-line trendline and an overall one were resolved. These two trendlines (first-line:  $\text{PFS} = -0.0454 \times \text{free energy difference} + 7.1358$  with  $p$ -values of [0.4782, 0.1482] for the two coefficients, overall:  $\text{PFS} = -0.0737 \times \text{free energy difference} + 4.6907$  with  $p$ -values of [0.1552, 0.1534] for the two coefficients) are quite similar and draft the mutation–PFS profile for afatinib. The results regarding the PB model are



**Fig. 5** Analyses of the computationally resolved binding affinities in mutant-afatinib systems. (a) Simplified free energy differences ( $\Delta E + \Delta\Delta G_{\text{solv}}$ ) solved by the GB or PB model and the entropy term ( $-T\Delta S$ ) for various mutant-afatinib systems. (b and c) Ranked free energy differences from (a). Here averaged free energy differences, respectively, of the mutations regarding first-line-afatinib treatments and those corresponding to sub-line-afatinib treatments, are displayed. (d) Distribution of the free energy difference (GB model) vs. response level, with an outlier marked. (e) Distribution of the free energy difference (PB model) vs. response level, with an outlier marked. (f) Distribution of the free energy difference (GB model) vs. PFS, with trendlines presented and an outlier marked. (g) Distribution of the free energy difference (PB model) vs. PFS, with trendlines presented and an outlier marked. (h) Experimentally resolved inhibitory constants of afatinib for different EGFR-family receptors. (i) Ranking of afatinib potency for EGFR kinases with reference to (g). (j) Ranking of computationally resolved afatinib potency for EGFR kinases, with both the GB and PB models considered.

listed in Fig. 5e and g (first-line: PFS =  $-0.0513 \times$  free energy difference +  $7.2505$  with  $p$ -values of  $[0.3667, 0.0811]$  for the two coefficients, overall: PFS =  $-0.0687 \times$  free energy difference +  $5.4418$  with  $p$ -values of  $[0.1390, 0.0628]$  for the two coefficients). Here the PB model is more confident than the GB one, and in future studies more data should be collected to improve the modeling confidence.

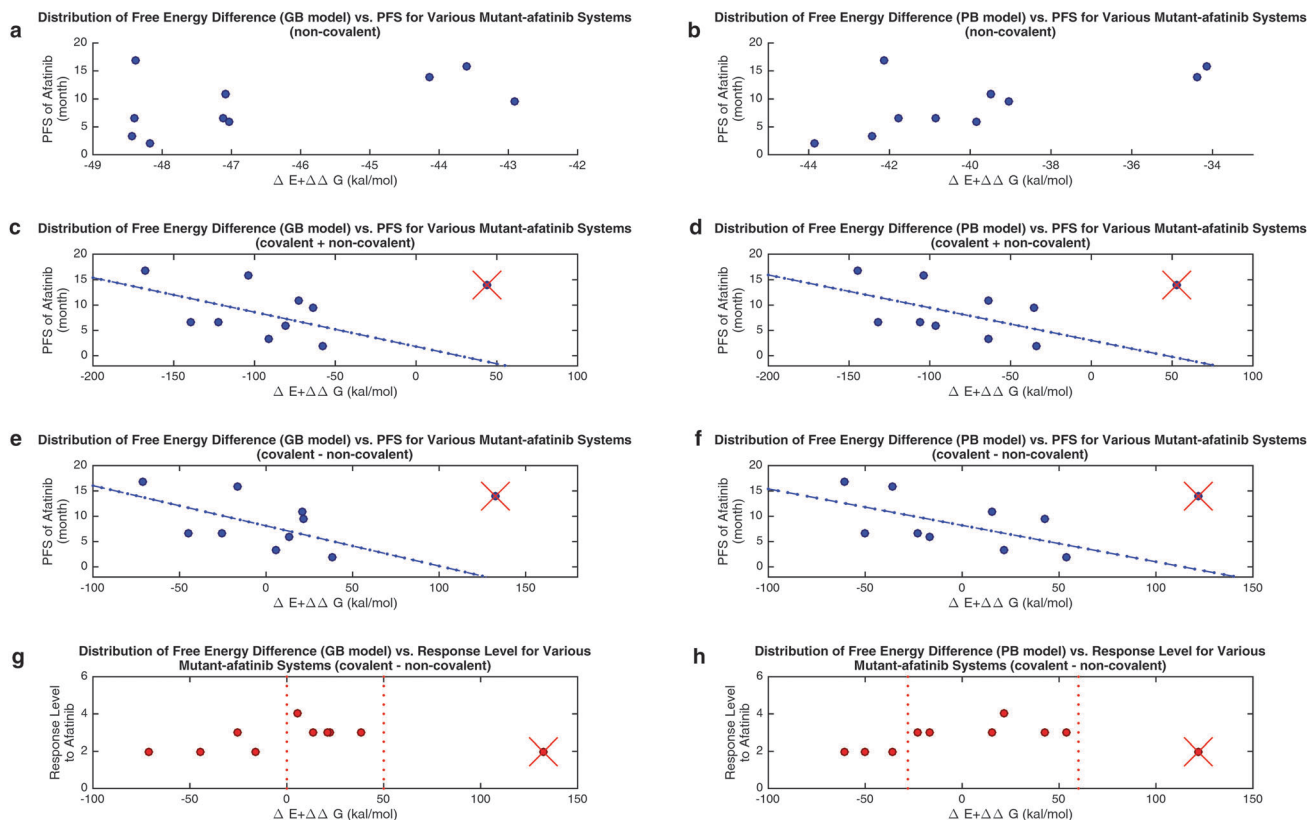
At last, we compared the computational afatinib potency with the experimental one, to verify its reliability. The inhibitory constant (nM), which evolved from biologically resolved  $EC_{50}$  or  $IC_{50}$  values, is broadly used to measure the potency of an inhibitor.<sup>8,26</sup> Collected from a number of literature studies,<sup>8,26,42</sup> the inhibitory constants of afatinib to several representative EGFR-family receptors (Table S2, ESI<sup>†</sup>) are shown in Fig. 5h. A lower inhibitory constant represents a higher potency, resulting in the potency ranking of several well-acknowledged EGFR kinases ( $L858R > WT > L858R\_T790M$ ) in Fig. 5i. Our computational afatinib potencies ( $-(\Delta E + \Delta\Delta G_{\text{solv}})$ ) for these kinases, based on the GB model (Fig. 5j), are well in accordance with the experimental ones. This can partly validate

the reliability of our computations. The results concerning the PB model should be further polished in future studies due to the occurrence of small positive values of free energy difference.

## 2.4 Conclusion and discussion

EGFR-mutated NSCLC has become an innovative and appealing target in cancer research and drug discovery.<sup>2,5,6</sup> Aside from the first-generation reversible inhibitors of EGFR-family receptors,<sup>1-3</sup> new-generation inhibitors, such as afatinib,<sup>14,16,26,27</sup> that irreversibly bind to those receptors were explored and developed. Decoding the correlations between afatinib potency and EGFR mutation status can provide a thorough support to the optimal use of such inhibitors. Recently, computational biology has been growing rapidly and serves as a creditable tool in these studies.<sup>10,21,23</sup>

In this work, a group of patients with EGFR-mutated NSCLC were studied. Their detailed EGFR mutation types were biologically screened prior to their treatment with oral afatinib, and the cancer progression of each patient was carefully followed up. Specifically, PFS and response level were two



**Fig. 6** Supplementary analyses of mutant-afatinib non-covalent systems. (a and b) Distribution of free energy difference (GB or PB model) vs. PFS for various mutant-afatinib non-covalent systems. (c and d) Combination of the two kinds of free energy differences (covalent and non-covalent) in the PFS analysis using a summation operation, based on either the GB or PB model. The trendlines for majority of the points are shown. (e and f) Combining the two kinds of free energy differences (covalent and non-covalent) in the PFS analysis using a subtraction operation, based on either the GB or PB model. The trendlines for majority of the points are shown. (g and h) Combining the two kinds of free energy differences (covalent and non-covalent) in the response analysis using a subtraction operation, based on either the GB or PB model. Rough boundaries between different response groups are displayed.

major indicators of drug effectiveness. Supported by the collected patient data, we detected that the classic activating mutations (*L858R* and *exon 19 deletion*) normally resulted in better afatinib-related responses or PFS values, while those involving the *T790M* mutation mostly perform worse during the treatment. To further bridge the EGFR mutation status and the response/PFS, we conducted a computational study. In advance, all EGFR mutant-afatinib covalent complexes were structurally modeled. Depending on MD simulations, the free energy difference between each complex system and the corresponding two isolated molecules was estimated as the mutant-afatinib binding affinity. This binding affinity was then mapped to the afatinib-related response or PFS, to briefly construct a mutation-selectivity profile for afatinib.

As a dynamic process, the binding kinetics of a mutant-afatinib system can also be an important aspect for measuring the potency/escape of afatinib. For additional discussion, the free energy differences for mutant-afatinib non-covalent systems (each containing the unbound afatinib and a mutant) were calculated and further investigated in Fig. 6. Fig. 6a and b present the distributions of such free energy differences for the two GB/PB models. In Fig. 6c and d, we combined the two kinds

of free energy differences (covalent and non-covalent) using a summation operation, and the *p*-values for the trendline coefficients are [0.1870, 0.7233] and [0.1738, 0.4786]. Depending on a subtraction operation, the combination of the two kinds of free energy differences is displayed in Fig. 6e and f, with the *p*-values of [0.1273, 0.0013] and [0.1113, 0.0011] for the trendline coefficients, respectively. Compared to Section 2.3.3, the subtraction operation improves the confidence of the trendlines. This subtraction operation was also implemented in the response analysis (Fig. 6g and h), resulting in a marginal improvement of the boundary determination. In future studies, this computational model should be further improved by taking into account more factors and operations. Besides, more data should be collected to increase the credibility and accuracy of this model. These studies can greatly promote optimal drug application and new drug exploration.

### 3 Experimental

All clinical data were collected from the Queen Mary Hospital in Hong Kong, and our study was implemented anonymously with



the consent of all participated patients. This study was approved by the ethics committee – Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (HKU/HA HKW IRB) – before commencement.

### 3.1 Patients and EGFR mutation screening

Incipiently, a cohort of patients diagnosed with stage IV metastatic NSCLC was identified between December 2010 and October 2014 in the hospital. Tumor stage was judged according to the TNM classification system by an independent review.<sup>14</sup> All these patients harbored EGFR mutations, which were evaluated prior to TKI therapy. Assisted by formalin-fixed paraffin-embedded (FFPE) tumor biopsy samples, the screening of EGFR mutations for individual patients was performed.<sup>35,36</sup> Specifically, tumor cells of each patient were first enriched by microdissection, followed by the extraction of the genetic DNA (QIAmp DNA FFPE tissue kit).<sup>35</sup> EGFR exons 18 to 21 were amplified with intron-based primers<sup>1,2,35,36</sup> by polymerase chain reaction (PCR), leading to a bi-directional sequencing of the PCR products.<sup>35,36</sup> Rare or multi-point mutations were verified by repeated PCR and sequencing, with the normal DNA sequenced as in ref. 36.

### 3.2 Patient treatment and PFS/response analysis

All patients were recommended to receive self-financed afatinib in their treatments, until disease progression, intolerable adverse events, or self-withdrawal.<sup>14,35</sup> All patients received a starting dose of 40 mg daily, except for one patient who started the medication at 50 mg daily, at the discretion of treating oncologist. Adverse events were assessed every 4 weeks and graded according to the National Cancer Institute-Common Terminology Criteria for Adverse Events (CTCAE) version 3.0.<sup>35</sup> If grade 3 or higher drug-related adverse events emerged, afatinib would be suspended for up to 14 days and be resumed until recovery to grade 1 or lower. Upon resuming the drug, dose reduction to one dose-level lower was implemented. Afatinib would be permanently discontinued if grade 3 or higher adverse events developed at a dose of 30 mg daily. Tumor assessment was implemented by computed tomography (CT) scan or positron emission tomography with integrated CT (PET-CT) scan every 12 weeks, and the objective response was determined by Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1.<sup>14</sup>

The primary study endpoint was progression-free survival (PFS), defined as the time from TKI-commencement to disease progression or death.<sup>5,35</sup> Our secondary endpoints were objective response (complete response [CR] or partial response [PR]) and disease control (stable disease [SD] or progressive disease [PD]).<sup>5,14,35</sup> PFS and objective responses among various EGFR mutation-subgroups were comparably investigated in our study.

### 3.3 Computational modeling of binding affinity in EGFR mutant–afatinib systems

Besides the PFS- or response-related distribution analysis, a computational study was implemented to decode the binding affinity between afatinib and each mutation emerged in our patients.

Furthermore, this affinity or binding mode was subsequently mapped to the recorded PFS or response, to build a mutation-selectivity profile for afatinib.

**3.3.1 Modeling of EGFR mutant–afatinib complexes.** Prior to the calculation of binding affinity, structural modeling of mutant–afatinib covalent complexes should be accomplished. The mutant structures were derived in advance from previous computational studies,<sup>10,37</sup> where side-chain prediction<sup>38,40</sup> and loop modeling<sup>39,40</sup> were implemented to determine each molecular structure. Such mutant structures were adopted to form the covalent complexes with afatinib. Specifically, our modeling relied on molecular structural data from the protein data bank (PDB),<sup>43</sup> and the complex PDB:4G5P<sup>8</sup> where afatinib covalently binds to mutant *T790M* was adopted as a major template. As a preparation, each mutant was first aligned to the template.

It is well acknowledged that afatinib binds to a cysteine residue of an EGFR kinase, and mutations rarely occur in this binding site. Accordingly, treating the afatinib–cysteine block in the template as a building block is the main idea to model the mutant–afatinib complexes. Later, such rough models can be optimized by *AMBER*<sup>10,21</sup> and can support the binding affinity calculation. *AMBER* is a software suite designed for biomolecular simulations and analyses,<sup>10,21</sup> and it strongly depends on a specified force field ( $V$ ) that combines different energy components,

$$V = E_{\text{bonded}} + E_{\text{nonbonded}} = \left( \sum_{\text{bonds}} + \sum_{\text{angles}} + \sum_{\text{dihedrals}} \right) + \left( \sum_{\text{electrostatic}} + \sum_{\text{vanderwaals}} \right) \quad (1)$$

*AMBER* force fields are built upon the concept of pair-wise charges, with each amino acid residue regarded as the underlying charge unit for proteins. This charge model allows individual residues to be the building blocks for large protein systems, without the requirement of charge refitting for each system.<sup>21</sup> In this regard, the afatinib–cysteine block can be treated as a novel building block, and importantly, its atomic charges should be derived for modeling the mutant–afatinib complexes. To guarantee the compatibility with later *AMBER* simulations, the Restrained Electrostatic Potential (RESP) charges<sup>44</sup> should be derived. This charge fitting model adds a restraint to the original ESP fitting to ensure a better approximation. Two general restraint forms are as follows:

$$\chi_{\text{rstr}}^2 = a \sum_j \left( q_j^0 - q_j \right)^2 \quad (2)$$

$$\chi_{\text{rstr}}^2 = a \sum_j \left( \sqrt{q_j^2 + b^2} - b \right) \quad (3)$$

Here  $a$  and  $b$  are scale factors, and  $q_j^0$  and  $q_j$ , respectively, represent the target and fitted charges of atom  $j$ . Once atomic charges of the afatinib–cysteine building block are derived, we can insert this block into the corresponding site of each aligned mutant to form the covalent complex.

Recently, R.E.D. (RESP ESP charge Derive) is a favorable option for RESP charge derivation.<sup>45,46</sup> It automatically derives RESP charges and builds *AMBER* force field libraries for innovative molecular fragments, with high reproducibility.<sup>45</sup> Briefly, this charge model can be summarized in three steps, namely, (i) geometry optimization, (ii) molecular electrostatic potential (MEP) calculation based on the optimized geometry, and (iii) fitting the atom-centered charges to MEP.

R.E.D. relies on a quantum mechanics (QM) program, such as the *GAMESS* package<sup>47</sup> and the RESP program,<sup>44</sup> to sequentially perform the three steps. *GAMESS* is the software for *ab initio* molecular quantum chemistry,<sup>47</sup> and it implements self-consistent field (SCF) calculations for all Hartree–Fock (HF) wavefunctions. The core of SCF calculations is to approximate the solution of the following Hartree equation:

$$\hat{F}(1)\phi_j(1) = \varepsilon_j\phi_j(1) \quad (4)$$

Here  $\hat{F}(1)$  is the one-electron Fock operator,  $\phi_j(1)$  represents a set of one-electron wavefunctions (molecular orbitals) for atom  $j$ , and  $\varepsilon_j$  denotes the orbital energies.

Unfortunately, the afatinib–cysteine block is a fragment, thus a capping procedure should be performed before R.E.D. calculation to better restore the atomic charges. Normally, the acetyl (Ace) or amine (Nme) groups are used in *AMBER* to cap proteins, and such caps will be removed in actual simulations. R.E.D. was applied to reproduce the atomic charges of the capped fragment, after which the atomic charges of the afatinib–cysteine fragment can be extracted. In detail, a direct SCF calculation of *GAMESS*, with a density convergence criterion of  $1.0 \times 10^{-6}$ , was adopted in the geometry optimization. Depending on the optimized geometry, the rigid-body re-orientation algorithm in R.E.D. assisted the molecular orientation and MEP calculation, resulting in highly reproducible RESP charges.<sup>45,46</sup> With the generated charges and force field library of the fragment, mutant–afatinib complexes were successfully modeled.

Finally, afatinib was untied from the original template and carefully added missing atoms, after which *GAMESS* was adopted to optimize its geometry (convergence criterion of  $1.0 \times 10^{-8}$ ) and R.E.D. was applied to derive its charges and the force field library. These products were used in the subsequent binding affinity calculation.

**3.3.2 Molecular dynamics (MD) simulations and mutant–afatinib binding affinities.** Affinity in a target–inhibitor system is an important indicator of the inhibitory efficacy. For a covalently bound system, the free energy difference  $\Delta G_{\text{bind,solv}}$  between the complex (C) and the isolated molecules (A and B) can roughly evaluate the binding affinity (AFF) of the system. In our study, B and A, respectively, represent an EGFR kinase mutant and afatinib, and C denotes the mutant–afatinib covalent complex. Accordingly, the approximated binding affinity (AFF) is formulated as follows:

$$\text{AFF} \approx \Delta G_{\text{bind,solv}} = G_{\text{C}} - G_{\text{A}} - G_{\text{B}} \quad (5)$$

where  $G$  indicates the free energy of a solvated molecular system, and a lower  $\Delta G_{\text{bind,solv}}$  value implies a better affinity.

In *AMBER*,<sup>21</sup>  $G$  is computed using the Molecular Mechanics/Poisson Boltzmann (Generalized Born) Surface Area (MM/PB(GB)SA) module. Three major portions, namely free energy in vacuum ( $E$ ), entropy contribution ( $-T\Delta S$ ) and solvation free energy ( $\Delta G_{\text{solv}}$ ), comprise this computation.

$$G = E + \Delta G_{\text{solv}} - T\Delta S \quad (6)$$

Here  $E$  can be expressed using various energy terms in eqn (7). The entropy term  $-T\Delta S$ , measuring the disorder of a system, can be evaluated using normal mode analysis (NMA). The solvation free energy  $\Delta G_{\text{solv}}$  typically encompasses the polar and nonpolar contributions in eqn (8), where  $\Delta G_{\text{polar}}$  is normally approximated by a PB or GB model. Since *AMBER* calculates the relative free energies for binding systems in a more accurate way, we adopted such calculations in our study. Normally, the entropy terms are ignored to ease the computational load.

$$E = E_{\text{bonds}} + E_{\text{angles}} + E_{\text{dihedrals}} + E_{\text{electrostatic}} + E_{\text{van der Waals}} \quad (7)$$

$$\Delta G_{\text{solv}} = \Delta G_{\text{polar}} + \Delta G_{\text{nonpolar}} \quad (8)$$

In this regard, the free energy difference  $\Delta G_{\text{bind,solv}}$  can be further expressed as follows:

$$\begin{aligned} \Delta G_{\text{bind,solv}} &\approx \Delta E + \Delta \Delta G_{\text{solv}} \\ &= (E_{\text{C}} - E_{\text{A}} - E_{\text{B}}) + (\Delta G_{\text{solv,C}} - \Delta G_{\text{solv,A}} - \Delta G_{\text{solv,B}}) \end{aligned} \quad (9)$$

In order to accomplish the above free energy calculation, MD simulations of each solvated system (A, B or C) should first be generated using *AMBER*. *AMBER* can accurately calculate the relative free energies of two molecules in a binding system, thus we adopted two systems, {B,C} and {A,B}, for the calculation of relative free energies. {B,C} was regarded as an EGFR hetero- or homo-dimer system (template PDB:4RIW), and {A,B} was treated as a common protein–ligand non-covalent system (template PDB:4G5P). For each system, a general solvent environment was established in advance, as a truncated octahedron TIP3P water box with a 10-angstrom (Å) buffer around the solute in each direction. *AMBER ff12SB* (upgraded version of *AMBER ff99SB*) and *gaff* force fields were selected in the construction of molecular topologies due to their broad applications. Each solvated system was subsequently neutralized.

Prior to the key MD simulations, a solvated system should be minimized and equilibrated, to guarantee a stable simulation. The following series of operations were adopted in this phase for the system {B,C}.

- Short minimization (10 000 steps) to remove bad contacts with weak position restraints on the solute.
- 20 000-step minimization without restraints on the system, while with a QM/MM setting on the afatinib–cysteine block.
- Heating of 50 picoseconds (ps) from 100 K to 300 K with weak restraints on the solute.
- 25 ps density equilibration with weak restraints on the solute.
- 250 ps constant pressure equilibration at 300 K.

Owing to a better structural optimization, we adopted two successive minimizations. The time step in each operation was confined to 1 femtosecond (fs) for a favorable equilibration. Each system was validated for its equilibration through the stabilized density, temperature and total energy curves. On an equilibrated system, a production MD simulation of 2.5 nanoseconds (ns) was implemented, with the QM/MM setting on the afatinib-binding residue remained. Here, the stabilization can be verified by the backbone RMSD curve of each solute. All our computations relied on a high performance computer cluster (HPCC), consisting of a master node and 44 computer nodes with 16 cores for each node. Regarding the equilibration and production simulations, a 5-node computation led to an approximately 39 hour run.

Trajectory frames were collected every picosecond, and a total of 2500 frames were derived. Based on these frames, MM/PB(GB)SA calculated the relative free energies of {B,C}. Both the GB and PB models were applied in the calculations for a comparable study. Using the 5-node computer cluster, each calculation merely cost 3 hours.

For the system {A,B}, similar equilibration and production simulations were implemented. Due to the relatively small size of such systems, a 1.5 ns production simulation was selected to ease the computations. All parameters passed the stabilization test, which guaranteed a reliable simulation. The overall procedure cost 10 hours for a 5-node computation.

Based on the relative free energies derived above, binding affinities can be roughly estimated (eqn (5):  $G_C^{B\text{-relative}} - G_A^{B\text{-relative}} - G_B$ ). Such binding affinities can be adopted to characterize the sensitivity of afatinib to each mutant. Apparently, a lower negative value of the free energy difference represents a larger binding affinity between a mutant and afatinib, thus implying a better inhibitory efficacy of afatinib. Specifically for each mutation that emerged in our patients, its corresponding binding affinity with afatinib was calculated. These affinity values were comparably investigated and connected to the PFS or response, supporting the construction of a mutation-selectivity profile for afatinib.

## Acknowledgements

This work was supported by the Hong Kong Health and Medical Research Fund (Project 01121986).

## References

- 1 T. J. Lynch, D. W. Bell, R. Sordella, S. Gurubhagavatula, R. A. Okimoto, B. W. Brannigan, P. L. Harris, S. M. Haserlat, J. G. Supko and F. G. Haluska, *et al.*, *N. Engl. J. Med.*, 2004, **350**, 2129–2139.
- 2 J. G. Paez, P. A. Jänne, J. C. Lee, S. Tracy, H. Greulich, S. Gabriel, P. Herman, F. J. Kaye, N. Lindeman and T. J. Boggon, *et al.*, *Science*, 2004, **304**, 1497–1500.
- 3 W. Pao, V. Miller, M. Zakowski, J. Doherty, K. Politi, I. Sarkaria, B. Singh, R. Heelan, V. Rusch and L. Fulton, *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 13306–13311.
- 4 S. V. Sharma, D. W. Bell, J. Settleman and D. A. Haber, *Nat. Rev. Cancer*, 2007, **7**, 169–181.
- 5 L. V. Sequist, J. C.-H. Yang, N. Yamamoto, K. O'Byrne, V. Hirsh, T. Mok, S. L. Geater, S. Orlov, C.-M. Tsai and M. Boyer, *et al.*, *J. Clin. Oncol.*, 2013, **31**, 3327–3334.
- 6 H. Yasuda, S. Kobayashi and D. B. Costa, *Lancet Oncol.*, 2012, **13**, e23–e31.
- 7 J. Dengjel, I. Kratchmarova and B. Blagoev, *Mol. BioSyst.*, 2009, **5**, 1112–1121.
- 8 F. Solca, G. Dahl, A. Zoepfel, G. Bader, M. Sanderson, C. Klein, O. Kraemer, F. Himmelsbach, E. Haaksma and G. R. Adolf, *J. Pharmacol. Exp. Ther.*, 2012, **343**, 342–350.
- 9 J. D. Haley and W. J. Gullick, *EGFR signaling networks in cancer therapy*, Springer Science & Business Media, 2009.
- 10 D. D. Wang, W. Zhou, H. Yan, M. Wong and V. Lee, *Sci. Rep.*, 2013, **3**, 2855.
- 11 L. V. Sequist, D. W. Bell, T. J. Lynch and D. A. Haber, *J. Clin. Oncol.*, 2007, **25**, 587–595.
- 12 H. Shigematsu, L. Lin, T. Takahashi, M. Nomura, M. Suzuki, I. I. Wistuba, K. M. Fong, H. Lee, S. Toyooka and N. Shimizu, *et al.*, *J. Natl. Cancer Inst.*, 2005, **97**, 339–346.
- 13 M. Tokumo, S. Toyooka, K. Kiura, H. Shigematsu, K. Tomii, M. Aoe, K. Ichimura, T. Tsuda, M. Yano and K. Tsukuda, *et al.*, *Clin. Cancer Res.*, 2005, **11**, 1167–1173.
- 14 J. C.-H. Yang, J.-Y. Shih, W.-C. Su, T.-C. Hsia, C.-M. Tsai, S.-H. I. Ou, C.-J. Yu, G.-C. Chang, C.-L. Ho and L. V. Sequist, *et al.*, *Lancet Oncol.*, 2012, **13**, 539–548.
- 15 J. Zhang, J. Jia, F. Zhu, X. Ma, B. Han, X. Wei, C. Tan, Y. Jiang and Y. Chen, *Mol. BioSyst.*, 2012, **8**, 2645–2656.
- 16 V. A. Miller, V. Hirsh, J. Cadranel, Y.-M. Chen, K. Park, S.-W. Kim, C. Zhou, W.-C. Su, M. Wang and Y. Sun, *et al.*, *Lancet Oncol.*, 2012, **13**, 528–538.
- 17 D. M. Jackman, B. Y. Yeap, L. V. Sequist, N. Lindeman, A. J. Holmes, V. A. Joshi, D. W. Bell, M. S. Huberman, B. Halmos and M. S. Rabin, *et al.*, *Clin. Cancer Res.*, 2006, **12**, 3908–3914.
- 18 S. Kobayashi, H. Ji, Y. Yuza, M. Meyerson, K.-K. Wong, D. G. Tenen and B. Halmos, *Cancer Res.*, 2005, **65**, 7096–7101.
- 19 G. R. Oxnard, M. E. Arcila, C. S. Sima, G. J. Riely, J. Chmielecki, M. G. Kris, W. Pao, M. Ladanyi and V. A. Miller, *Clin. Cancer Res.*, 2011, **17**, 1616–1622.
- 20 C.-H. Yun, K. E. Mengwasser, A. V. Toms, M. S. Woo, H. Greulich, K.-K. Wong, M. Meyerson and M. J. Eck, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 2070–2075.
- 21 D. A. Case, T. A. Darden, T. E. Cheatham, III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang and K. M. Merz, *et al.*, *AMBER 12*, University of California, San Francisco, 2012.
- 22 M. W. Sneddon and T. Emonet, *Nat. Methods*, 2012, **9**, 239–242.
- 23 S. Wan, D. W. Wright and P. V. Coveney, *Mol. Cancer Ther.*, 2012, **11**, 2394–2400.
- 24 Q. Bai, D. Shi, Y. Zhang, H. Liu and X. Yao, *Mol. BioSyst.*, 2014, **10**, 1958–1967.
- 25 C. G. Gadhe and M.-H. Kim, *Mol. BioSyst.*, 2015, **11**, 618–634.

- 26 D. Li, L. Ambrogio, T. Shimamura, S. Kubo, M. Takahashi, L. Chirieac, R. Padera, G. Shapiro, A. Baum and F. Himmelsbach, *et al.*, *Oncogene*, 2008, **27**, 4702–4711.
- 27 T. A. Yap, L. Vidal, J. Adam, P. Stephens, J. Spicer, H. Shaw, J. Ang, G. Temple, S. Bell and M. Shahidi, *et al.*, *J. Clin. Oncol.*, 2010, **28**, 3965–3972.
- 28 P. de Antonellis, *Naunyn-Schmiedeberg's Arch. Pharmacol.*, 2014, **387**, 503–504.
- 29 S. Kumar and R. Agrawal, *Recent Pat. Anti-Cancer Drug Discovery*, 2014, **9**, 382–393.
- 30 C. Carmi, M. Mor, P. G. Petronini and R. R. Alfieri, *Biochem. Pharmacol.*, 2012, **84**, 1388–1399.
- 31 M. H. Johansson, *Mini-Rev. Med. Chem.*, 2012, **12**, 1330–1344.
- 32 M. L. Sos, H. B. Rode, S. Heynck, M. Peifer, F. Fischer, S. Klüter, V. G. Pawar, C. Reuter, J. M. Heuckmann and J. Weiss, *et al.*, *Cancer Res.*, 2010, **70**, 868–874.
- 33 P. Hoffknecht, A. Tufman, T. Wehler, T. Pelzer, R. Wiewrodt, M. Schütz, M. Serke, J. Stöhlmacher-Williams, A. Märten and R. M. Huber, *et al.*, *J. Thorac. Oncol.*, 2015, **10**, 156.
- 34 H. Modjtahedi, B. C. Cho, M. C. Michel and F. Solca, *Naunyn-Schmiedeberg's Arch. Pharmacol.*, 2014, **387**, 505–521.
- 35 V. H. Lee, V. P. Tin, T.-S. Choy, K.-O. Lam, C.-W. Choi, L.-P. Chung, J. W. Tsang, P. P. Ho, D. K. Leung and E. S. Ma, *et al.*, *J. Thorac. Oncol.*, 2013, **8**, 1148–1155.
- 36 I. Y. San Tam, L. P. Chung, W. S. Suen, E. Wang, M. C. Wong, K. K. Ho, W. K. Lam, S. W. Chiu, L. Girard and J. D. Minna, *et al.*, *Clin. Cancer Res.*, 2006, **12**, 1647–1653.
- 37 L. Ma, D. D. Wang, Y. Huang, H. Yan, M. P. Wong and V. H. Lee, *BMC Bioinf.*, 2015, **16**, 85.
- 38 Z. Xiang and B. Honig, *J. Mol. Biol.*, 2001, **311**, 421–430.
- 39 Z. Xiang, C. S. Soto and B. Honig, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 7432–7437.
- 40 K. W. Kaufmann, G. H. Lemmon, S. L. DeLuca, J. H. Sheehan and J. Meiler, *Biochemistry*, 2010, **49**, 2987–2998.
- 41 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605–1612.
- 42 W. Zhou, D. Ercan, L. Chen, C.-H. Yun, D. Li, M. Capelletti, A. B. Cortot, L. Chirieac, R. E. Iacob and R. Padera, *et al.*, *Nature*, 2009, **462**, 1070–1074.
- 43 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235–242.
- 44 C. I. Bayly, P. Cieplak, W. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269–10280.
- 45 F.-Y. Dupradeau, A. Pigache, T. Zaffran, C. Savineau, R. Lelong, N. Grivel, D. Lelong, W. Rosanski and P. Cieplak, *Phys. Chem. Chem. Phys.*, 2010, **12**, 7821–7839.
- 46 A. Pigache, P. Cieplak and F. Dupradeau, *Abstracts of Papers of The American Chemical Society*, 2004, p. U1011.
- 47 M. S. Gordon and M. W. Schmidt, *Theory and applications of computational chemistry: the first forty years*, 2005, pp. 1167–1189.