J Mol Model (2013) 19:497–509
DOI 10.1007/s00894-012-1559-0

ORIGINAL PAPER

Computational study of EGFR inhibition: molecular dynamics studies on the active and inactive protein conformations

Napat Songtawee · M. Paul Gleeson · Kiattawee Choowongkomon

Received: 24 November 2011 / Accepted: 2 August 2012 / Published online: 7 September 2012

Abstract The structural diversity observed across protein kinases, resulting in subtly different active site cavities, is highly desirable in the pursuit of selective inhibitors, yet it can also be a hindrance from a structure-based design perspective. An important challenge in structure-based design is to better understand the dynamic nature of protein kinases and the underlying reasons for specific conformational preferences in the presence of different inhibitors. To investigate this issue, we performed molecular dynamics simulation on both the active and inactive wild type epidermal growth factor receptor (EGFR) protein with both type-I and type-II inhibitors. Our goal is to better understand the origin of the two distinct EGFR protein conformations, their dynamic differences, and their relative preference for Type-I inhibitors such as gefitinib and Type-II inhibitors such as lapatinib. We discuss the implications of protein dynamics from a structure-based design perspective.

Keywords EGFR · Kinase · Molecular dynamics · AMBER force field · GROMACS · Gefitinib · Lapatinib

Abbreviations

EGFR	Epidermal growth factor receptor
G-loop	Glycine-rich loop
A-loop	Activation loop
R-spine	Regulatory spine
H-cluster	Hydrophobic cluster
DFG motif	Asp-Phe-Gly conserved motif
HRD motif	His-Arg-Asp conserved motif
PDB	Protein data bank
MD	Molecular dynamics
RMSD	Root average square deviation
RMSF	Root average square fluctuation
SD	Standard deviation

Electronic supplementary material The online version of this article (doi:10.1007/s00894-012-1559-0) contains supplementary material, which is available to authorized users.

N. Songtawee · K. Choowongkomon
Department of Biochemistry, Faculty of Science,
Kasetsart University,
50 Phaholyothin Rd, Chatuchak,
Bangkok 10900, Thailand

M. P. Gleeson (✉)
Department of Chemistry, Faculty of Science,
Kasetsart University,
50 Phaholyothin Rd, Chatuchak,
Bangkok 10900, Thailand
e-mail: paul.gleeson@ku.ac.th

K. Choowongkomon (✉)
Center for Advanced Studies in Tropical Natural Resources,
NRU-KU, Kasetsart University,
50 Phaholyothin Rd, Chatuchak,
Bangkok 10900, Thailand
e-mail: fsciktc@ku.ac.th

Introduction

Protein kinases are an important class of therapeutic targets in drug discovery. At present, eight kinase inhibitors are currently marketed as anti-cancer treatments [1], and it has been estimated that approximately one-third of all pharmaceutical research projects are dedicated to such targets [2]. Three of the eight marketed kinase drugs target the epidermal growth factor receptor (EGFR), also known as ErbB1 kinase. A wealth of biochemical and structural information has been generated on this target, offering us considerable insight into the structure, function and inhibition of this important therapeutic target class [3–5].

Over 160 unique protein kinase X-ray structures have been deposited in the RCSB (http://www.thesgc.org/resources/kinases), offering a great deal of information to aid in the design of new or improved kinase-directed therapies. The protein kinase domain of EGFR is comprised of two lobes: a

smaller N-terminal lobe consisting mainly of β-strands and a single large α-helix; and a larger C-terminal lobe, which is almost exclusively α-helical. The ATP-binding site is located at the hinge region between the lobes, meaning the active site is dynamic in size and shape. The structure of EGFR kinase can be further divided into a number of structural regions, as highlighted in Fig. 1a,b. These include the glycine-rich loop (G-loop), the C α-helix on the N-lobe, the activation loop (A-loop), and conserved DFG and HRD motifs on the C-lobe. Several important features for EGFR activation include: (1) reorientation of the C α-helix closer to the ATP-binding site, resulting in the formation of a salt bridge between E738 on the helix and the conserved K721 residue on the β5-strand. The latter also interacts with the α- and β-phosphates of ATP (see supporting information Fig. S5A). (2) The positioning of DFG-D831 and HRD-D813 residues to interact with the ATP phosphate groups and the peptide substrate, respectively; and (3) extension of the A-loop, and translation away from the active site; (4) the formation of the regulatory spine (R-spine) by three hydrophobic (M742, L753, F832) and one polar residue (H811), leading to a H-bond between the HRD-R812 and the DFG+1-L834, which is proposed to help maintain the active kinase conformation (Fig. 1c) [6, 7]. Several other residues, including L723, M742, L764, D831, L834 and L837, are proposed to form a small hydrophobic cluster (H-cluster) between the C α-helix and the A-loop, which is also believed to be important for stabilizing the inactive conformation of EGFR kinase (Fig. 1d) [7, 8].

The structural diversity observed across protein kinases, resulting in subtly different active site cavities, as well as the often distinctly different protein conformations, is highly desirable in the pursuit of selective inhibitors, yet it is also can be a hindrance from a structure-based design perspective. For example, analysis of the active EGFR-gefitinib crystal structure (PDB accession code: 2ITY) would suggest that the addition of the substituent 1-methoxy,3-F-phenyl to the quinazoline template would not be tolerated. However, not only is this substituent tolerated, it is believed that the resultant complex, along with increased ErbB2 activity, give lapatinib its improved efficacy (PDB accession code: 1XKK).

Kinase inhibitors can be classified into two to three distinct categories [1, 9, 10]. Type-I inhibitors target the

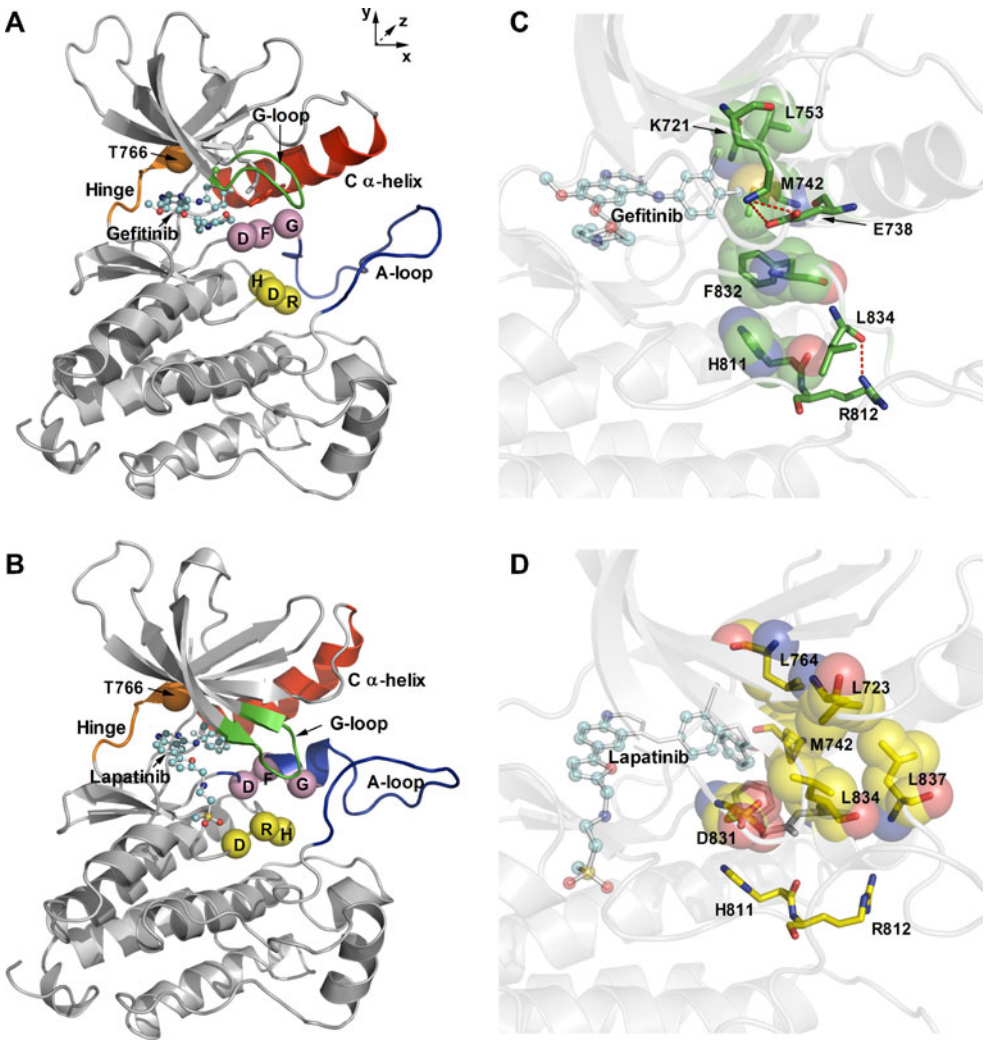

Fig. 1 Ribbon representations of active (**a**) and inactive (**b**) epidermal growth factor receptor (EGFR) kinase structures (PDB codes 2ITY and 1XKK, respectively). Key secondary structural elements are colored (*green* glycine-rich loop; *red* C α-helix; *blue* activation loop). The ligands are shown in *ball and stick* notation (C-atoms in *cyan* for both gefitinib and lapatinib). Gatekeeper (T766) and DFG and HRD motifs are shown in *space-filled balls*. Conserved interactions and residue clusters differentiating the active (**c**) and inactive (**d**) EGFR conformational states are indicated. Regulatory spine and hydrophobic cluster are represented as *transparent spheres* (C-atom in *green* and *yellow*, respectively). The salt bridge of K721-E738 and the H-bond of R812-L834 are shown in *red dashed lines*. The figure was made using PYMOL (DeLano Scientific, San Calos, CA)

ATP-binding site of the active or inactive protein. As these inhibitors target the more generic, evolutionary conserved, ATP-binding pocket, undesirable activity towards other members of the approximately 500 strong protein kinase families is frequently observed [11]. Type-II inhibitors target both the ATP-binding and allosteric pockets formed within an inactivated protein. This includes the DFG-out conformation in c-Abl [12] and the allosteric binding pocket of MEK that lies adjacent to the ATP-binding site [13]. Type-I½ inhibitors can be considered a hybrid of the previous two, targeting the ATP-binding site of the DFG-in conformation of the inactive protein, as well as a rather large back-pocket as exemplified in the EGFR-lapatinib complex (Fig. 2) [4]. There is particular interest in Type-I½ and Type-II inhibitors from the point of view of selectivity as these regions will be under reduced evolutionary pressure to remain constant. These differences are therefore more likely to be exploited to produce a selective kinase inhibitor. However, problems pursuing this type of inhibitor also exist. Mutations at or around allosteric pockets are more likely to occur than at evolutionary conserved regions, potentially leading to problems associated with drug-resistance [14].

Molecular dynamics (MD) studies have in the past been used to elucidate dynamic aspects associated with protein kinases, including EGFR [7, 15–18]. Such simulations offer additional insight beyond the static, but nonetheless critical, snapshot as represented by an X-ray crystal structure. MD has proved particularly insightful for EGFR from a drug resistance perspective, as a dynamic assessment of the effect of mutations, including L834R, G695S and L834R and T766M, on protein structure can be assessed [7, 15–17]. Liu et al. [15] studied the origin of resistance for the Type-I EGFR inhibitor gefitinib (Iressa®), noting the implications each mutation had on the ATP-binding pocket and on inhibitor binding. Balius et al. [16] studied the effect of EGFR mutations on Type-I inhibitors; erlotinib (Tarceva®), gefitinib and AEE788, using models of the wild type (WT) and three different active EGFR mutants. They were able to explain the majority of the fold resistance changes in the different mutants from the calculated binding free energy, as well as giving an explanation for their physical origin. Recently, Wan et al. [17] also studied the changes in drug-binding affinities due to the cancer-related mutations of EGFR using multiple short MD simulations, which provide significantly enhanced conformational sampling. Also worth mentioning is a study by Papakyriakou et al. [7], who investigated EGFR protein dynamics in the absence of inhibitors. The focus of their study was understanding the

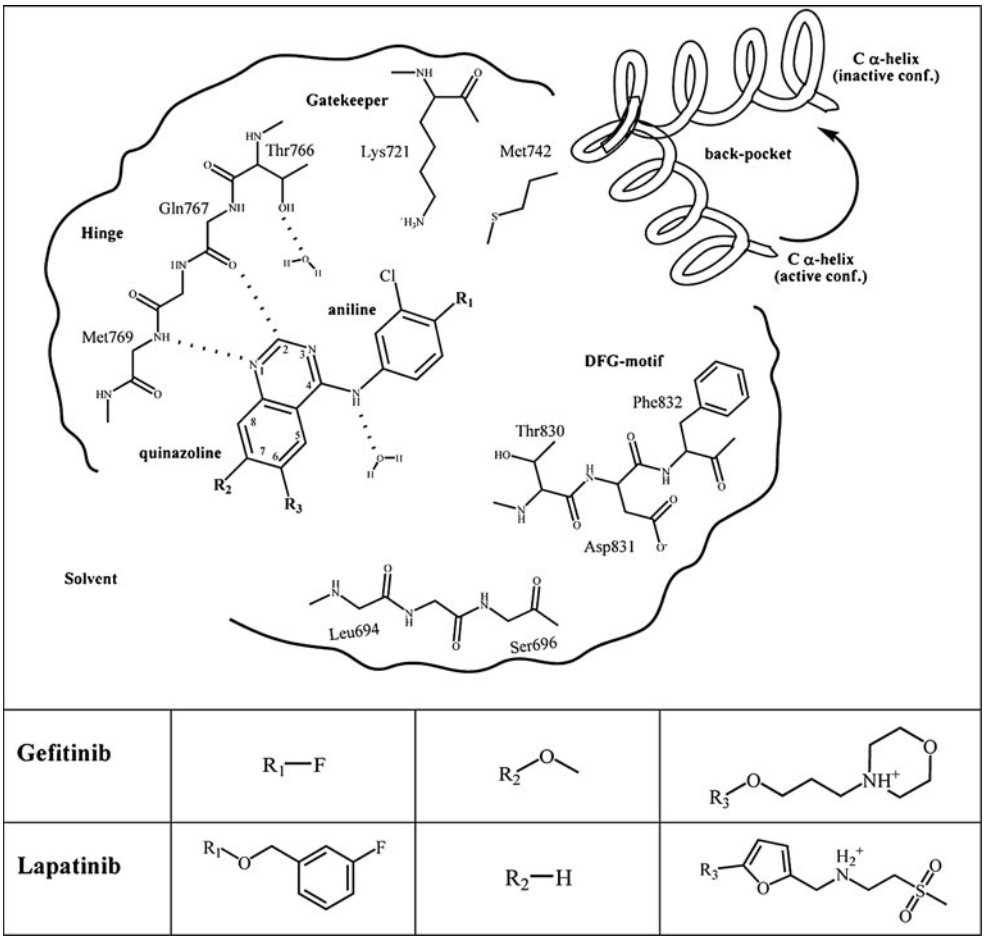

Fig. 2 Illustration of the EGFR kinase binding site for potential kinase inhibitors (PDB codes 2ITY and 1XKK). Important amino acid residues located in the binding site and the chemical structure of ligands (gefitinib and lapatinib) are shown. The main interactions of the EGFR kinase hinge with the quinazoline moiety are indicated. The LIGPLOT diagram [40] for all hydrophobic and H-bond interactions from the PBD files are shown in Supporting Information Fig. S6

dynamic differences between the active and inactive protein conformations of EGFR, and the transition between them. Using 5 ns of targeted MD to drive the transition between the two different states, they concluded that the timescales needed for the formation the back-pocket in inactive EGFR protein are beyond the timescales of conventional MD.

An important challenge in structure-based design is to better understand the dynamic nature of protein kinases, and the underlying reasons for the different protein conformational preferences observed with different inhibitors. In this novel study, we perform MD simulations on both active and inactive protein complexes of wild type EGFR, with both type-I and type-II inhibitors. Our goal is to try to understand the origin of the distinct EGFR conformations, and the relative preference of these protein conformations for the Type-I inhibitor gefitinib and Type-I½ inhibitor lapatinib. An improved understanding of inhibitor binding to the inactive conformation is highly desirable given that inhibitors of this conformation, rather than the active form, appear to be more efficacious [9]. To this end, we employed MD simulations using the AMBER force field within GROMACS to simulate the active and inactive forms of the protein. We assessed the drug molecules gefitinib and lapatinib to try and decipher the relative contribution of the inhibitor to the stability of the two protein conformations. We also simulate the APO forms of both protein conformations, and the case where gefitinib bound to the inactive conformation. The value of this information is that the contribution of the various structural elements, or individual residues, to inhibitor binding and protein stability can be better understood, potentially allowing the more focused direction of chemistry resources to target the area most likely to give rise to higher affinity, tighter binding inhibitors.

Methods and materials

Protein preparation

The EGFR protein coordinates for the active and inactive conformations were obtained from the Protein Data Bank (http://www.pdb.org). The active and inactive coordinates used in this study correspond to the PDB structures with accession codes 2ITY [5] and 1XKK [4], respectively. EGFR-ligand protein structure models were created by removing all ions, and all water molecules except those found within the binding site (three molecules in 2ITY and eleven in 1XKK). We retained the water molecules found in the active sites of each protein for all the simulations used here as they have been shown to be critical in protein–ligand simulations in the past [19, 20]. Water molecules have been shown to be important for rationalizing dynamic phenomena from MD simulations [21] as well as docking and scoring results [22, 23].

The amino acid sequences used for our simulations begin with A678 and finish at G959, based on the 2ITY numbering (or A702 to G983 using the alternative EGFR numbering system). Missing loops in both PDB structures (E842 to K851 in 2ITY and E710 to K713, A726 to S728, E844 to K851 in 1XKK) were built using MODELER 9v4 [24]. The stereochemical quality of the resultant models was assessed using PROCHECK v3.5.4 [25]. The N- and C-terminal ends of both models were capped with acetyl (ACE) and methyl amino (NME) groups, respectively. The 2ITY and 1XKK based models are henceforth referred to as gefitinib-$EGFR_{(active)}$ and lapatinib-$EGFR_{(inactive)}$, respectively.

A model of the inactive EGFR protein with gefitinib bound was also generated for the purpose of comparison. This model was created by replacing lapatinib with gefitinib in the inactive 1XKK based EGFR model. Superposition of the ligands was performed based on the 2ITY-1XKK Cα-atom alignment. Due to differences in the G-loop conformations between active and inactive proteins, the orientation of the 7-methoxy-6-(3-morpholin-4-ylpropoxy) moiety of gefitinib was altered subtly to avoid steric clashes, adopting a conformation resembling that of the 5-{[2-(methylsulfonyl)ethyl] amino} methyl) group of lapatinib. This structure is termed gefitinib-$EGFR_{(inactive)}$. No model of lapatinib bound to the active EGFR conformation was generated as the absence of the back-pocket II prevents the inhibitor from binding. APO structures of the active and inactive EGFR protein were created by removing the ligands from the gefitinib-$EGFR_{(active)}$ and lapatinib-$EGFR_{(inactive)}$ models. These models are termed the APO-$EGFR_{(active)}$ and APO-$EGFR_{(inactive)}$ models, respectively.

The AMBER-99SB force field was used to simulate all protein structures and the ionization state of amino acid residues was set according to the standard protocol [26]. All models were solvated in a triclinic box of TIP3P water, keeping a distance of 10 Å between the protein and the sides of the solvent box. Chloride ions were added to neutralize the charge of the system.

Ligand preparation

Gefitinib and lapatinib were simulated in their protonated state in line with their predicted basic pKa according to MarvinView v5.3.1 (ChemAxon, Budapest, Hungary) at pH 7.0 (Fig. 2). This protonation state is further favored due to the close proximity of a number of H-bond acceptors (either the negatively charge side chain of D776 or the carbonyl oxygen atom of L694) at the entrance to the ATP-binding pocket. GROMACS topology files were generated using ACPYPE script [27]. GAFF force field parameters [28] were used for both inhibitors. Partial charges were calculated using the AM1-BCC method [29] as implemented in QUACPAC 1.3.1 (OpenEye Scientific Software, Santa Fe, NM).

Simulation conditions

Simulations were carried out using GROMACS v4.0.2 [30, 31] with the AMBER force field ports [32, 33]. All simulations used isobaric-isothermal (NPT) conditions at standard temperature (300 K) and pressure (1 bar), using the Berendsen coupling method [34]. The linear constraint (LINCS) algorithm was applied to fix all hydrogen related bond lengths, facilitating the use of a 2-fs time step [35]. A short-range nonbonded interaction cut-off distance of 10 Å was used. The particle mesh Ewald (PME) method with a 0.12 nm Fourier grid spacing was used to account for long-range electrostatics [36, 37].

A three-step procedure was used for MD simulations. First, each of the EGFR models was energy-minimized using the steepest descent method (until the maximum force was less than 100 kJ mol^{-1} nm^{-1} on any atom) to reduce undesirable van der Waals contacts, and to optimize H-bond interactions present. In the second step, each model was subjected to 500 ps of a position-restrained MD in which heavy atom positions of each protein were restrained harmonically using a force constant of 1000 kJ mol^{-1} nm^{-2}. Water molecules, counterions and inhibitors, if present, were not restrained. The systems were then heated from 0 K to 300 K over the first 50 ps, followed by 450 ps of equilibration. The third step involved unrestrained MD for a period of 20 ns. Coordinates were archived every 1 ps. The simulations for the APO-$EGFR_{(inactive)}$ and gefitinib-$EGFR_{(inactive)}$ models were subsequently extended to 50 ns to assess the conformational characteristics of their C α-helices.

Analyses

All MD analyses were performed using tools available within the GROMACS suite. The tool "*g_rms*" was used to evaluate the root mean square deviation (RMSD) of heavy atoms in MD trajectories from those of original structures obtained before energy minimizations. The tool "*g_rmsf*" was used to compute the root mean square fluctuation (RMSF) of heavy-atom positions with respect to their time-averaged position and was used to calculate a theoretically derived B-factor (temperature factor).

The statistical significance of any reported differences in either the means or standard deviations (SD) in the RMSD or RMSF have been confirmed using an unpaired Student's *T*-test or F-test, respectively. All reported differences are significance above the commonly used 95 % confidence level unless otherwise stated. All statistics were computed in Microsoft Excel 2007.

Results and discussion

MD simulations were performed on five separate EGFR models that differ in terms of the bound inhibitor [gefitinib, lapatinib and no inhibitor (i.e. APO)], or the protein conformation (active or inactive). Analysis of the RMSD of the protein heavy-atoms (i.e., compared to the initial X-ray structures obtained before energy minimizations) showed that all simulations had reached equilibrium well before $t=10$ ns. The protein structures remained stable throughout the simulation; with the overall heavy-atom RMSD remaining within 3.0 Å of the original X-ray coordinates (Fig. 3a–b). In addition, the total energy of each model remained essentially constant over the course of the simulation, giving further confirmation of its stability (Supporting Information Fig. S1). We report all structural parameter analyses between $t=10$ to 20 ns unless otherwise stated.

The reliability of such simulations can be assessed qualitatively by comparing the experimental C_α-atom B-factor values to those computed from the MD simulation (Fig. 3c, d). As shown in Fig. 3d, the MD-predicted B-factors of the gefitinib-$EGFR_{(active)}$ and APO-$EGFR_{(active)}$ models are in good qualitative agreement with the corresponding experimental data. Deviation to some degree is expected since the X-ray data is obtained in a dynamically restricted crystalline phase. This suggests the MD results are physically representative of the protein in general.

Dynamic characteristics of the active and inactive protein-inhibitor complexes

In the following sections, we consider the dynamic characteristics of the gefitinib-$EGFR_{(active)}$ and lapatinib-$EGFR_{(inactive)}$ models, and whether the structural differences observed between the models have arisen due to (1) the effect of the ligand, (2) the protein, or (3) a combination of both. To ascertain their origin, we contrast the results to simulation data obtained using models of gefitinib bound to the inactive EGFR conformation and models of both active and inactive APO protein conformations.

From Fig. 3a,b and Table 1, it can be seen that the average RMSD and SD of the gefitinib-$EGFR_{(inactive)}$ model is greater than that of the lapatinib-$EGFR_{(inactive)}$ model over the course of the simulation. The differences were found to be statistically highly significant ($P<0.0001$). From the APO simulations, we can see that the inactive protein conformation (APO-$EGFR_{(inactive)}$) behaves similarly to the lapatinib-$EGFR_{(inactive)}$ model in that the average RMSD is roughly comparable (the SD of the RMSD is only moderately lower by 0.01, Table 1). It is somewhat different for the active EGFR model in that the RMSD increases by 0.39 Å in the APO structure compared to the ligand bound structure. This suggests that the ligand plays a more important role in stabilizing this conformation (Table 1).

To investigate these differences further, we computed the average energy of the proteins over the course of the simulation as the sum of the bonded (internal) and

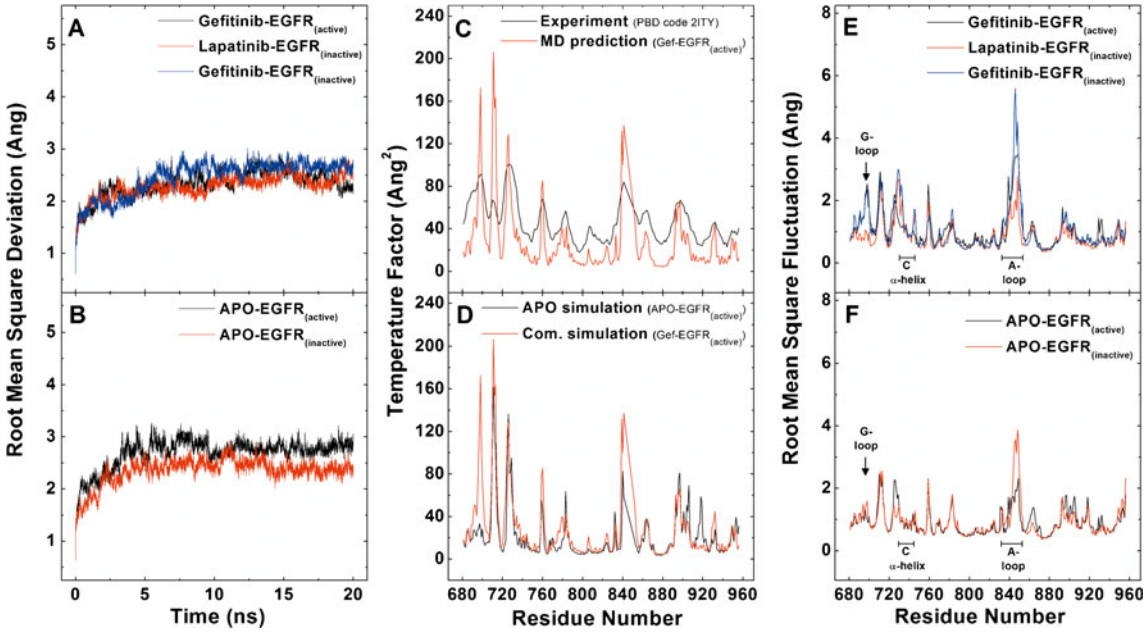

Fig. 3 Heavy-atom protein RMSD plots of the complex (**a**) and apo (**b**) simulations. Comparisons of experimental versus predicted B-factor values (**c**) and predicted B-factors of complex versus apo simulations (**d**). Heavy-atom RMSF/residue plots of the complex (**e**) and apo (**f**) EGFR kinase structures. Key secondary structural elements (glycine-rich loop, C α-helix and activation loop are indicated

nonbonded (electrostatic and van der Waals) energies. This energy neglects energetic components of the solvent and the inhibitor (apart from the interactions with the protein, which are considered) as a means of comparing, albeit approximately, their energy profile over the course of the simulations. We observed the following statistically significant trends in the computed energy: APO-EGFR$_{(active)}$< lapatinib-EGFR$_{(inactive)}$<APO-EGFR$_{(inactive)}$<gefitinib-EGFR$_{(inactive)}$<gefitinib-EGFR$_{(active)}$. It is possible to compare these energies qualitatively since all models have an identical number of atoms. The most energetically favorable protein conformation over the course of the simulation was the APO protein in the active conformation. Surprisingly, the binding of gefitinib to the active protein conformation resulted in the protein with the highest overall energy, more so even than when bound to the inactive EGFR conformation. The APO form of the inactive protein is found to be less energetically favourable than the active form but binding of lapatinib stabilizes this conformation. Gefitinib was found to destabilize the inactive protein conformation. One might therefore conclude that the active protein is destabilized to a degree by inhibitor binding, whereas the inactive protein conformation is stabilized, at least by inhibitors such as lapatinib, which possess a back-pocket binding group.

These observations do not represent the complete picture since many of the individual structural elements found within the five different protein models will be dynamically dissimilar over the course of the simulation. To assess these differences, the heavy-atom RMSF value of each residue (that which is related to the crystallographic B-factor or thermal motion) was calculated to understand how the different structural elements behave between t=10 to 20 ns

Table 1 The heavy-atom root mean square deviation (RMSD) average and standard deviation (SD) values (Å) (in parenthesis) for the overall protein structure, key secondary structural elements conserved interactions and ligands bound over the course of the simulation (values calculated from t=10 to 20 ns)

	gefitinib-EGFR$_{(active)}$	lapatinib-EGFR$_{(inactive)}$	gefitinib-EGFR$_{(inactive)}$	APO-EGFR$_{(active)}$	APO-EGFR$_{(inactive)}$
Overall structure	2.39 (0.17)	2.34 (0.15)	2.55 (0.24)	2.78 (0.15)	2.42 (0.14)
Glycine-rich loop	1.73 (0.21)	0.66 (0.11)	0.81 (0.15)	1.68 (0.14)	1.52 (0.39)
C α-helix	1.84 (0.28)	2.99 (0.61)	2.72 (0.54)	1.80 (0.19)	2.55 (0.30)
Activation loop	3.76 (0.72)	3.14 (0.23)	4.01 (0.74)	3.92 (0.20)	2.82 (0.30)
Hydrophobic cluster	1.72 (0.19)	1.66 (0.15)	2.15 (0.26)	1.84 (0.13)	1.77 (0.27)
Regulatory spine	1.58 (0.27)	0.97 (0.14)	1.43 (0.20)	1.24 (0.19)	1.07 (0.26)
Ligand	1.26 (0.28)	1.08 (0.11)	1.14 (0.26)	–	–

(Fig. 3e–f). We have also assessed the heavy-atom RMSD (compared to the initial structures obtained before energy minimizations) of three important conserved secondary structural elements, C α-helix (P729 to A743), glycine-rich loop (G-loop; G695 to T701) and activation loop (A-loop; D831 to V852) and the bound ligands. In addition, we have assessed the differences in the EGFR specific regions, termed the regulatory spine (R-spine) and hydrophobic cluster (H-cluster), which are believed to be important in differentiating the active and inactive protein conformations, respectively. The data are summarized in Table 1.

Ligand binding

The average RMSD of the inhibitors in the gefitinib-$EGFR_{(active)}$ and lapatinib-$EGFR_{(inactive)}$ complexes were comparable over the course of the simulation, yet gefitinib was found to fluctuate to a much greater extent than lapatinib since the RMSD SD of the ligand in the former is 0.17 Å greater than in the latter (Table 1, Fig. 4). The differences are due primarily to the solvent-exposed tails of the two inhibitors, which is evident given that the RMSD of the quinazoline and anilino-based substituents are rather small (Supporting Information Fig. S4). The results from the gefitinib-$EGFR_{(active)}$ simulation are consistent with a recent study in that the solvent-exposed tail of the inhibitors exhibits greater movement than the central scaffold [17]. This is expected given the importance of the hinge interaction between the quinazoline acceptor and the M769 donor (Fig. 2 and Supporting Information Fig. S6). The differences in flexibility at the solvent-exposed region could be a result of differences in the intrinsic flexibility of adjacent protein structural elements, or a reflection of the differing binding strengths/characteristics of the substituents in question. Analysis of the interactions present in Fig. 5 and Supporting Information Table S1 indicates that the basic nitrogen of gefitinib interacts strongly with the carbonyl oxygen of L694 located on the G-loop, considerably more so compared to lapatinib. In contrast, lapatinib forms stronger interactions with the C-lobe via residues such as C773 and its furan ring, D776 and its basic nitrogen, and R817 and its sulfone group. The three strong interactions formed by lapatinib contrast with just one strong interaction of gefitinib, and help to explain the latter's larger RMSD. Analysis of the results from the gefitinib-$EGFR_{(inactive)}$ simulation, where gefitinib was placed in the active site of the inactive protein conformation, reveals that the ligand also

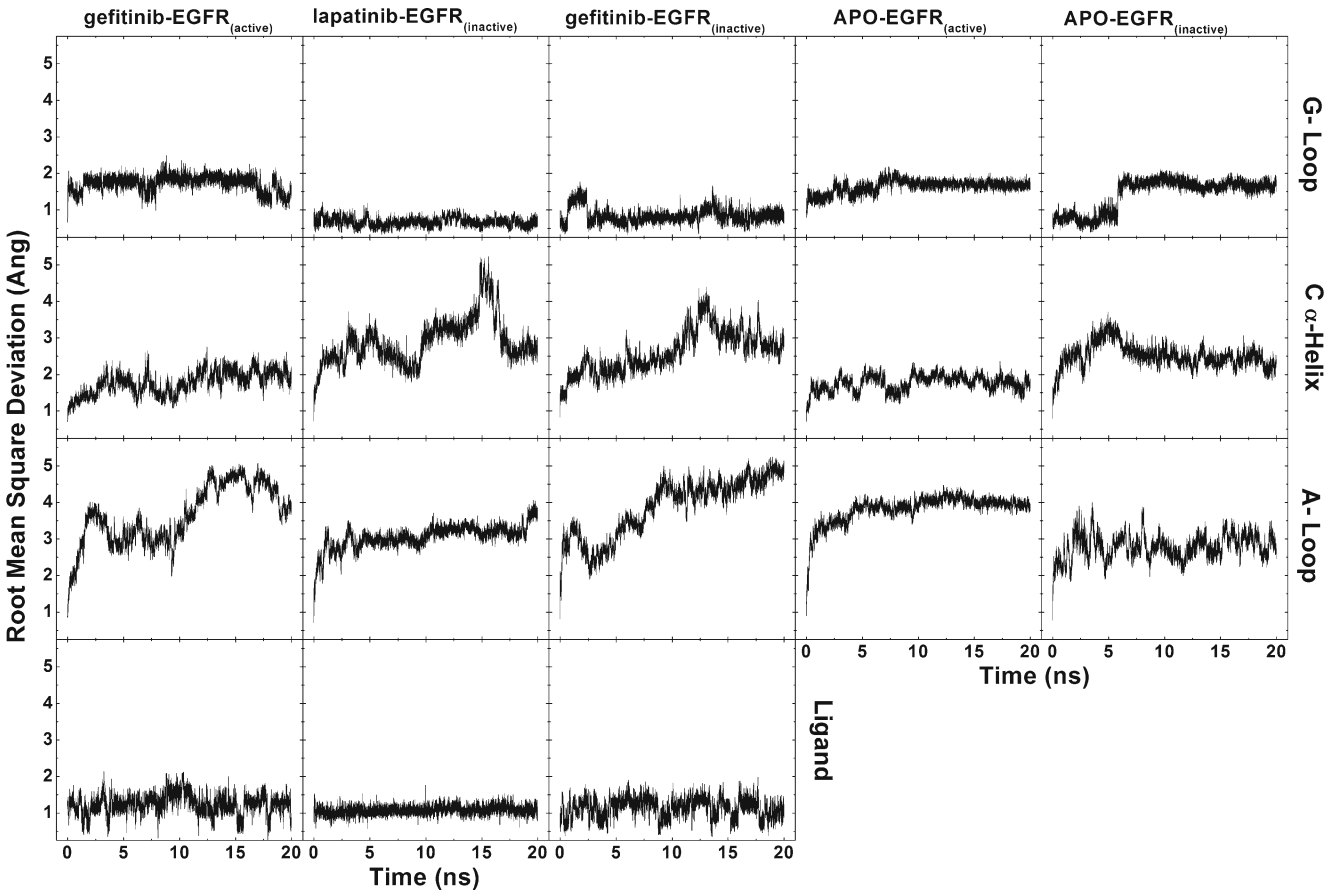

Fig. 4 Heavy-atom RMSD plots of the residues in the key secondary structural elements; glycine-rich loop, C α-helix, activation loop and ligands for all molecular dynamics (MD) simulations

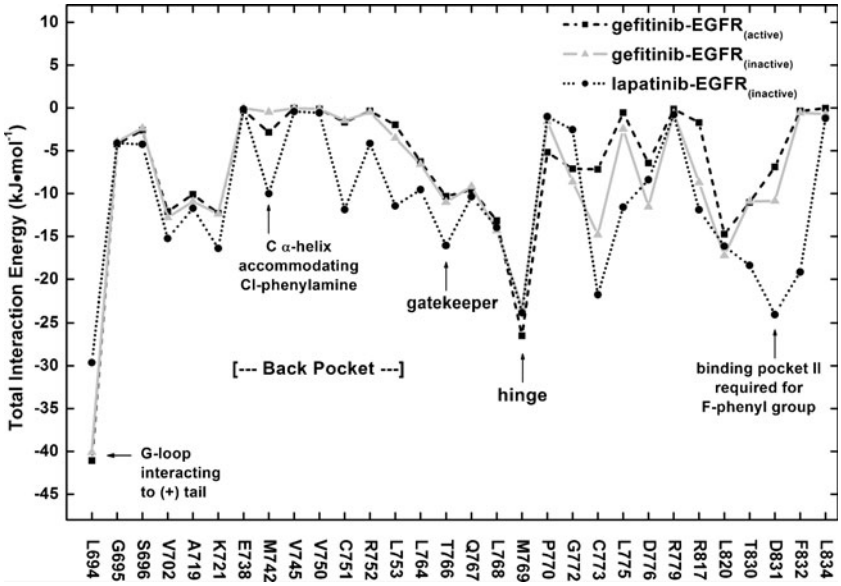

Fig. 5 Plots for the interaction energy of selected residues located in the binding site for the three complex simulations. The important residues are indicated

fluctuates to a larger degree also in this cavity (Table 1, Fig. 4). This suggests that the behavior of gefitinib is a characteristic of the molecule itself and not an effect of the protein conformation.

Analysis of the interaction energies (the sum of electrostatic and van der Waals interactions) between the inhibitors and adjacent active site residues was subsequently considered (Fig. 5 and Supporting Information Table S1). A key difference between the two inhibitors is the presence of a large back-pocket binding group of lapatinib, which can make a number of strong interactions not made by gefitinib. These include significant interactions by the methoxy-3-F-phenyl substituent with T830, L753, M742, D831, and F832 (see also Supporting Information Fig. S6). Common interactions made by both inhibitors are those between the 3-Cl-phenylamine portion and T830 and K721, although these are noticeably stronger in the case of lapatinib. Strong interactions are made with the hinge M769 residues in all cases as can be seen in Fig. 5 and Supporting Information Table S1.

It is also possible to look at the atomic fluctuation of individual active site residues as this can shed light on the nature of the binding site interactions (Fig. 6). An analysis of the RMSF of these residues shows noticeable differences between the different EGFR complexes. For example, L694 and S696 of the glycine-rich loop fluctuate considerably more in the gefitinib-based complexes. Gefitinib forms a favorable interaction with the G-loop residue L696 and, over the course of the simulation, this interaction is maintained even as the protein undergoes significant fluctuation. Residues E738 to V745 are located in the back-pocket and fluctuate along with the inhibitor back-pocket substituents

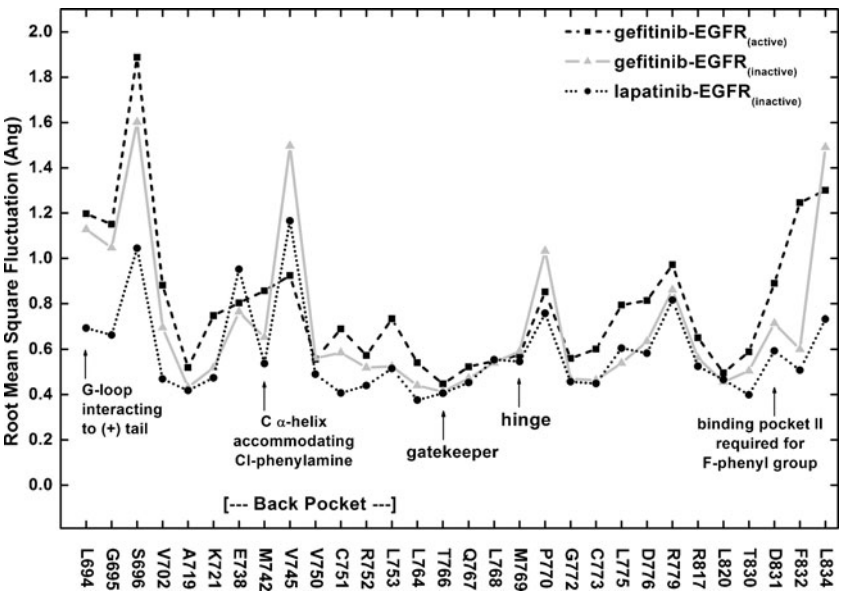

Fig. 6 Plots for the RMSF for the selected residues located in the binding site for the three complex simulations. The important residues are indicated

(or lack of the extended back-pocket binding group of gefitinib). The hinge binding region, along with the gatekeeper (T766 to M769), show rather low RMSF values as might be expected given the importance of the interaction made by the inhibitors and M769 (although for gefitinib the values are generally slightly larger). Residues D776–L834 in Fig. 6 correspond to the floor of the ATP-binding site, as well as part of the back-pocket. It appears that the RMSF values of these are also slightly larger for gefitinib-$EGFR_{(active)}$ than lapatinib-$EGFR_{(inactive)}$, which would suggest that lapatinib forms the more tightly bound complex of the two inhibitors.

We assessed the binding free energies of the inhibitors (ΔG_{bind}) using the linear interaction energy (LIE) method [38, 39]. We found that the gefitinib ΔG_{bind} to the active and inactive proteins were comparable, but that these were≈ 7 kJ mol^{-1} higher in energy than that observed for lapatinib binding to the inactive protein. Although these energies are not considered as precise as alternate methods such as FEP or even MM-PBSA, the results obtained are consistent with the reports from Woods et al. [4] who found that lapatinib has a much slower off rate than other EGFR inhibitors such as gefitinib. That said, the results are not in agreement with the experimental IC_{50} values. IC_{50} values are generally determined to the active protein conformation in biochemical assays, and not the inactive form (requiring longer equilibration times), which might help explain the discrepancy.

C α-helix

The C α-helix of EGFR kinase is the principle region that differs between the active and inactive protein conformations. The formation of the inactive EGFR conformation requires translation of the C α-helix in the z-direction relative to the rest of the protein (Fig. 1). This movement then leads to the formation of an additional hydrophobic pocket (back-pocket II), which is occupied by the 3- F-phenyl group of lapatinib.

The C α-helix in the lapatinib-$EGFR_{(inactive)}$ structure was found to have a larger average RMSD and RMSF over the course of the simulation than those observed in the gefitinib-$EGFR_{(active)}$ simulation (Fig. 3e, Fig. 4 and Table 1). This appears to be due, in part, to helix–coil transitions at the N-terminus of the helix (P729–I735) in simulations of the $EGFR_{(inactive)}$ structures, even though those residues tended to be helical at the end of simulation. The higher degree of flexibility and conformational transitions are in fact consistent with the ambiguous electron density found for residues A726–P729 in the 1XKK structures [4] and are consistent with the simulation results reported by others [7]. This observation was assumed initially to be an artifact of the modeled loop, consisting of three amino acid residues at the top of the C α-helix in 1XKK. However, a similar effect is observed in 2ITY, where these residues have been resolved experimentally. Additionally, upon simulation of the APO protein derived from 2ITY (APO-$EGFR_{(active)}$), we found the helix remained essentially intact over the course of the simulation, suggesting the transitioning was, at least in part, ligand induced. We also observed a dramatic drop in the flexibility of the C α-helix going from the inactive ligand-bound structure (lapatinib-$EGFR_{(inactive)}$) to the APO structure (APO-$EGFR_{(inactive)}$). The RMSD SD dropped from (0.61 to 0.30 Å) indicating that the ligand plays an important role in inducing this instability. Additional evidence for this is the drop in flexibility, albeit smaller, going from lapatinib-$EGFR_{(inactive)}$ to gefitinib-$EGFR_{(inactive)}$. This is because the latter lacks the back-pocket II binding group created by the C α-helix movement. As noted by Papakyriakou et al. [7], it is also likely that differences between crystal stacking forces and those of the simulated water will be responsible for some of these differences.

Of additional interest to us was whether the active and inactive structures had converged to any degree, particularly when the simulations were extended to 50 ns for the APO-$EGFR_{(inactive)}$ and gefitinib-$EGFR_{(inactive)}$ models. The N-terminus of the C α-helix (P729 to I735) of the $EGFR_{(inactive)}$ structures is unstable and shows the helix to coil transitioning. In the APO-$EGFR_{(inactive)}$ simulation, we also found helical bending leading to the N-terminal segment of the helix drifting toward that of the active conformation (Supporting Information Fig. S5C and S5E). In contrast, the C-terminal segment of the helix does not change its position, consistent with the long distance between K721 and E738 (Supporting Information Fig. S2). This helical bending was seldom observed in the gefitinib-$EGFR_{(inactive)}$ and lapatinib-$EGFR_{(inactive)}$ complexes (Supporting Information Fig. S5B, S5D and S5F).

Comparing the C α-helices from structures obtained from the five different simulation structures, to that of the original active X-ray structure (2ITY), reveals some interesting trends (Table 2). The APO-$EGFR_{(active)}$ structure displays the RMSD of the helix of≈3.7 Å, compared to 3.3 Å for the gefitinib-$EGFR_{(active)}$. RMSD analysis of the $EGFR_{(inactive)}$ simulations using the initial active conformation as the reference structure showed that the C α-helix of the $EGFR_{(inactive)}$ complexes have the RMSD of≈7 Å to the active conformation and increase over time to 8.6 A for gefitinib-$EGFR_{(inactive)}$ after 50 ns of simulation; however, in simulation of the $EGFR_{(inactive)}$ protein in the APO form, the RMSD decreases from 7.03 Å to 5.44 Å after 20 ns and to 5.01 after 50 ns (Table 2 and Supporting Information Fig. S7). This suggests that the C α-helix of both the active and inactive APO structure start to converge towards a similar minimum (RMSD decreasing from 7.0 to 3.3 Å at t=20 ns). However,

Table 2 Heavy-atom RMSD (Å) of the C α-helices after overall superimposition of the $EGFR_{(active)}$ and $EGFR_{(inactive)}$ structures at t=20 ns and t=50 ns when compared to the $EGFR_{(active)}$ structure at t=0 ns

	APO-$EGFR_{(active)}$	APO-$EGFR_{(inactive)}$	gefitinib-$EGFR_{(active)}$	lapatinib-$EGFR_{(inactive)}$	gefitinib-$EGFR_{(inactive)}$
t=20 ns	3.69	5.44	3.31	7.33	7.15
t=50 ns	–	5.01	–	–	8.64

while the C α-helix of the inactive protein did appear to move towards a more active-like conformation, we did not observe the formation of the salt bridge between K721 and E738, the most critical element associated with the EGFR activation [3].

Analysis of the regulatory (R-) spine (residues, M742, L753, H811 and F832) [6] shows that the conformation adopted by these residues in the gefitinib-$EGFR_{(inactive)}$ and APO-$EGFR_{(inactive)}$ structures are not close to those found in the active conformation. (Table 1 and Supporting Information Fig. S3). The H-bond between HRD R812 and DFG+1 L834, which is an important feature in the active conformation, was measured over the course of simulation (Supporting Information Fig. S2). Although at the end of the simulations, the R812–L834 distance is shorter in the gefitinib-$EGFR_{(inactive)}$ structure, when compared to the lapatinib-$EGFR_{(inactive)}$ and APO-$EGFR_{(inactive)}$ structures, this key H-bond is still unlikely to form during the gefitinib-$EGFR_{(inactive)}$ simulation since it is still found to be>4.0 Å.

Glycine-rich loop

The G-loop consists of a set of flexible residues that are located on the N-lobe of protein kinases. These residues are important in defining the size and shape of the ATP-binding pocket, as well as its dynamic characteristics, and should have major implications for inhibitor binding. Indeed, the opening of the active site pocket to solvent in the inactive protein containing lapatinib is slightly smaller than in the active protein with gefitinib due to the conformation adopted by the G-loop.

Over the course of the 20 ns simulation, it can be seen that the G-loop of gefitinib-$EGFR_{(active)}$ model deviates further than that of lapatinib-$EGFR_{(inactive)}$, and fluctuates to a greater extent (Fig. 4, Table 1). Analysis of the atomic coordinates reveals that this movement is primarily in the y-dimension as defined in Fig. 1, corresponding to the expansion and contraction of the entrance into the ATP-binding site.

The dynamics characteristics of the G-loop in gefitinib-$EGFR_{(active)}$ and lapatinib-$EGFR_{(inactive)}$ appear to correlate with that of the bound inhibitors (Table 1). For example, the average RMSD and SD of the inhibitor and G-loop in gefitinib-$EGFR_{(active)}$ simulation are almost double those obtained from the lapatinib-$EGFR_{(inactive)}$ simulation. This is expected given that gefitinib interacts more strongly with the G-loop than lapatinib, as discussed previously. From the results of the separate gefitinib-$EGFR_{(inactive)}$ simulation, we observed that the G-loop and inhibitor have an average RMSD between the gefitinib-$EGFR_{(active)}$ and lapatinib-$EGFR_{(inactive)}$ values. However, while the RMSD SD of the G-loop is also intermediate in value, that of gefitinib is the same as the original gefitinib-$EGFR_{(active)}$ simulation. From a consideration of the APO simulation results (Fig. 4), it appears that the presence of an inhibitor significantly stabilizes the G-loop of the inactive protein. Removal of the inhibitor leads to a dramatic increase in the RMSD and RMSF compared to the gefitinib-$EGFR_{(inactive)}$ and lapatinib-$EGFR_{(inactive)}$ structures (Table 1, Fig. 3e–f and Fig. 4). In contrast, the RMSD and RMSF values of the APO-$EGFR_{(active)}$ structure do not deviate dramatically from that of the gefitinib-EGFR(active) structure, suggesting the G-loop conformation in the active protein is intrinsically more stable.

Activation loop

The A-loop is an important structural element found in protein kinases. It contains amino acid residues that are critical to achieve their catalytic function of phosphorylation. The key portion of the A-loop is the so called DFG motif, which is found in the "*in*" conformation in known EGFR structures. A key interaction between residues in the C and N-lobes are mediated through the DFG D831 residue of the A-loop, and K721 of the β3-strand. There also exists a short α-helical segment towards the N-terminus of the A-loop (L834 to L837) in the inactive EGFR structure, but is not present in the active structure. This may also affect the conformational characteristics in this region.

The A-loop showed the largest RMSD and RMSF values among the three key structural elements over the course of the simulation for both EGFR conformations (Table 1, Figs. 3e, f; 4). The largest movement of the A-loop is in agreement with the work of others [17] and this is in line with the experimental X-ray data in that the residues E844 to K851 are disordered in both 2ITY and 1XKK structures [4, 5].

The A-loop found in the lapatinib-$EGFR_{(inactive)}$ structure was found to fluctuate considerably less than that found in gefitinib-$EGFR_{(active)}$ (RMSD s.d. of 0.23 Å vs 0.72 Å, respectively) (Table 1). It appears that the short α-helical

segment towards the N-terminus of the A-loop in the inactive protein limits the degree of flexibility. In addition, the interaction between D831 and K721 over the course of the MD simulations was somewhat weaker in gefitinib-EGFR$_{(active)}$ than lapatinib-EGFR$_{(inactive)}$ simulations, consistent with the original X-ray structure (Supporting Information Fig. S2).

The simulation of gefitinib bound to the inactive EGFR protein structure (gefitinib-EGFR$_{(inactive)}$) helps to shed light on the characteristics of the A-loop. The average RMSD and RMSF from this simulation are roughly comparable to that of the gefitinib-EGFR$_{(active)}$ simulation, suggesting that the effect of gefitinib on the A-loop is similar to that observed for the G-loop (Table 1, Figs. 3e,f; 4).

From the simulations of the APO-EGFR$_{(active)}$ and APO-EGFR$_{(inactive)}$ structures, it appears that the A-loop in the active protein deviates to a greater extent (RMSD average of 3.9 Å vs 2.8 Å, respectively), but fluctuates to a lesser degree (RMSD SD of 0.20 Å vs 0.30 Å, respectively). However, the binding of lapatinib to the inactive protein leads to a drop in the RMSD and RMSF suggesting that it help stabilize the A-loop conformation. In contrast, gefitinib binding to the active protein conformation leads to dramatically increased RMSD SD values in particular (0.20 Å vs 0.72 Å, respectively) suggesting it has the opposite effect. The binding of gefitinib to the inactive protein conformation also results in the larger RMSF and RMSD SD, suggesting that the destabilizing effect is due to the fact it makes no interactions with residues on the C-lobe (in contrast to the three interactions made by lapatinib) (see also Table 1, Figs. 3e,f; 4; 5).

Hydrophobic cluster

A network of several residues (L723, M742, L764 and D831), including two on the activation loop (L834 and L837), form a small hydrophobic (H-) cluster, and are reported to be important for the stabilization of the inactive EGFR conformation [7, 8]. Indeed, the cancer-related mutation L834R, has been known to be involved in either the disruption of the hydrophobic packing of the inactive EGFR kinase structure [8] or the introduction of an intermediate state in the active-inactive transformation pathway, adjusting the relative stability of both states [18], subsequently inducing EGFR activation.

From our simulation data, we find that the H-cluster does not vary dramatically over the course of the simulation. Although the H-cluster does not exist in the active conformation, we show the RMSD values for the purpose of comparison (Table 1 and Supporting Information Fig. S3). The average RMSD of the H-cluster in both APO proteins was roughly comparable; however, the RMSD SD in the APO-EGFR$_{(active)}$ protein is considerably lower than that in the APO-EGFR$_{(inactive)}$ (0.13 vs 0.27 Å), suggesting that it does play some form of stabilizing role in the active conformation. The observation that the lapatinib-EGFR$_{(inactive)}$ simulation displays a very low RMSD (both average and SD values) is not surprising given that these residues can reorientate, with interactions with lapatinib in the back-pocket region being made. The larger RMSD of the H-cluster in the gefitinib-EGFR$_{(inactive)}$ structure is consistent with the partial unfolding of the short α helix located close to L834 and L837. This helical transition was not observed in the APO-EGFR$_{(inactive)}$ and lapatinib-EGFR$_{(inactive)}$ structures.

Conclusions and future directions

In this study, we performed MD simulations on both the active and inactive protein complexes of the wild type EGFR, with both type-I and type- 1½ inhibitors. Our goal was to better understand the origin for the two distinct protein conformations and their relative preference for the Type-I inhibitor gefitinib and Type-II inhibitor lapatinib. We also simulated both APO forms and gefitinib bound to the inactive conformation to decipher the relative contribution of the inhibitor to stability of the two EGFR conformations.

We find that binding of gefitinib to the active protein appears somewhat destabilizing when compared to the APO simulation of the same conformation. The major cause of destabilization is increased fluctuation of the G-loop, A-loop and, to a lesser extent, the C α-helix. In contrast, the binding of lapatinib to the inactive conformation helps to lower the energy of the protein. Lapatinib binding leads to lower fluctuation in the G-loop and A-loop but a dramatic increase for the C α-helix.

Calculation of binding free energies suggests that lapatinib also binds more strongly to the inactive protein than gefitinib does to the active protein. While this would appear to contradict their reported experimental pIC_{50} values, it is in agreement with the experimentally determined slower off rate displayed by lapatinib [4], which is one of the proposed reasons for its better efficacy (along with its ErbB2 inhibition). The more favorable binding energy can be explained by the more extensive π type interactions it can make in the EGFR back-pocket by virtue of its additional 3-Cl-phenyl substituent. In addition, lapatinib makes three moderately strong interactions with the C-lobe (via its sulfone, furan and basic nitrogen) in contrast to gefitinib, which makes a single strong interaction with the G-loop (via its basic nitrogen). In addition, an analysis of the RMSF shows that the G-loop in the gefitinib-EGFR$_{(active)}$ structure fluctuates dramatically, but in the process also maintains the interaction already present. The lack of any interaction between gefitinib and the C-lobe also explains the large RMSF values of the A-loop. In contrast, the three relatively strong interactions that lapatinib makes with the C-lobe appears to restrain

the A-loop, while the weaker interaction with the G-loop leads to a higher RMSF in this region.

The C α-helix of the inactive EGFR conformation is intrinsically more mobile than that of the active conformation. This is unsurprising given its more extended position based on the known X-ray structure. It is found that the presence of an inhibitor in either protein conformation increases the flexibility in this region compared to the equivalent APO structure as a result of interactions mediated with residues in the back-pocket. Interestingly, from an analysis of the APO simulations, it appears that the C α-helix conformation in the active and inactive proteins begins to converge to a similar minimum after 50 ns of MD. However, we did not observe the formation of a salt bridge between K721 and E738—a critical element associated with EGFR activation [3].

The principle value of inhibitor binding information from MD is that the contribution of the various structural regions or individual residues to inhibitor binding and protein stability can be better understood. This potentially allows the more focused direction of chemistry resources to target regions in the active site most likely to give rise to higher affinity, tighter binding inhibitors. For example, rather than searching for any H-bond interactions to increase inhibitor affinity, an analysis of inhibitor binding, and the resultant change this has on receptor flexibility, may help to determine which H-bonds will contribute best to inhibitor binding, since the formation of such interactions may induce either greater stability or instability in the protein.

Acknowledgments This work was supported by the Higher Education Research Promotion and National Research University Project of Thailand, Office of the Higher Education Commission and the Thailand Research Fund grants RMU5180032 (KC) and RSA5480016 (MPG). We wish to express our gratitude for the use of Laboratory for Computational and Applied Chemistry (LCAC) research facilities at Kasetsart University provided by the National Center of Excellence in Petroleum, Petrochemical Technology and Advanced Materials. M.P.G. is grateful for the support of the Faculty of Science at KU and Associate Professor Supa Hannongbua in particular. N.S. was supported by a grant under the program Strategic Scholarships for Frontier Research Network for the Join PhD Program Thai Doctoral degree from the Office of the Higher Education Commission, Ministry of Education, Thailand.

References

1. Zuccotto F, Ardini E, Casale E, Angiolini M (2009) Through the "Gatekeeper Door": exploiting the active kinase conformation. J Med Chem 53:2681–2694
2. Weinmann H, Metternich R (2005) Editorial: drug discovery process for kinease inhibitors. Chem Bio Chem 6:455–459
3. Stamos J, Sliwkowski MX, Eigenbrot C (2002) Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. J Biol Chem 277:46265–46272
4. Wood ER, Truesdale AT, McDonald OB, Yuan D, Hassell A, Dickerson SH, Ellis B, Pennisi C, Horne E, Lackey K, Alligood KJ, Rusnak DW, Gilmer TM, Shewchuk L (2004) A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib). Cancer Res 64:6652–6659
5. Yun CH, Boggon TJ, Li Y, Woo MS, Greulich H, Meyerson M, Eck MJ (2007) Structures of lung cancer-derived EGFR mutants and inhibitor complexes: mechanism of activation and insights into differential inhibitor sensitivity. Cancer Cell 11:217–227
6. Kornev AP, Haste NM, Taylor SS, Ten Eyck LF (2006) Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. Proc Natl Acad Sci USA 103:17783–17788
7. Papakyriakou A, Vourloumis D, Tzortzatou-Stathopoulou F, Karpusas M (2009) Conformational dynamics of the EGFR kinase domain reveals structural features involved in activation. Proteins 76:375–386
8. Zhang X, Gureasko J, Shen K, Cole PA, Kuriyan J (2006) An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. Cell 125:1137–1149
9. Liu Y, Gray NS (2006) Rational design of inhibitors that bind to inactive kinase conformations. Nat Chem Biol 2:358–364
10. Liao JJL (2007) Molecular recognition of protein kinase binding pockets for design of potent and selective kinase inhibitors. J Med Chem 50:409–424
11. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S (2002) The protein kinase complement of the human genome. Science 298:1912–1934
12. Schindler T, Bornmann W, Pellicena P, Miller WT, Clarkson B, Kuriyan J (2000) Structural mechanism for STI-571 inhibition of abelson tyrosine kinase. Science 289:1938–1942
13. Tong L, Pav S, White DM, Rogers S, Crane KM, Cywin CL, Brown ML, Pargellis CA (1997) A highly specific inhibitor of human p38 MAP kinase binds in the ATP pocket. Nat Struct Biol 4:311–316
14. Bikker JA, Brooijmans N, Wissner A, Mansour TS (2009) Kinase domain mutations in cancer: implications for small molecule drug design strategies. J Med Chem 52:1493–1509
15. Liu B, Bernard B, Wu JH (2006) Impact of EGFR point mutations on the sensitivity to gefitinib: insights from comparative structural analyses and molecular dynamics simulations. Proteins 65:331–346
16. Balius TE, Rizzo RC (2009) Quantitative prediction of fold resistance for inhibitors of EGFR. Biochemistry 48:8435–8448
17. Wan S, Coveney PV (2011) Rapid and accurate ranking of binding affinities of epidermal growth factor receptor sequences with selected lung cancer drugs. J R Soc Interface 8:1114–1127
18. Wan S, Coveney PV (2011) Molecular dynamics simulation reveals structural and thermodynamic features of kinase activation by cancer mutations within the epidermal growth factor receptor. J Comput Chem 32:2843–2852
19. Amadasi A, Surface JA, Spyrakis F, Cozzini P, Mozzarelli A, Kellogg GE (2008) Robust classification of "relevant" water molecules in putative protein binding sites. J Med Chem 51:1063–1067
20. Barillari C, Taylor J, Viner R, Essex JW (2007) Classification of water molecules in protein binding sites. J Am Chem Soc 129:2577–2587
21. Treesuwan W, Hannongbua S (2009) Bridge water mediates nevirapine binding to wild type and Y181C HIV-1 reverse transcriptase—evidence from molecular dynamics simulations and MM-PBSA calculations. J Mol Graph Model 27:921–929
22. Santos R, Hritz J, Oostenbrink C (2009) Role of water in molecular docking simulations of cytochrome P450 2D6. J Chem Info Model 50:146–154
23. Huang N, Shoichet BK (2008) Exploiting ordered waters in molecular docking. J Med Chem 51:4862–4865

24. Šali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779–815
25. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26:283–291
26. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins 65:712–725
27. Sousa da Silva AW, Vranken WF (2012) ACPYPE—AnteChamber PYthon Parser interfacE. BMC Res Notes 5:367
28. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. J Comput Chem 25:1157–1174
29. Wang J, Wang W, Kollman PA, Case DA (2006) Automatic atom type and bond type perception in molecular mechanical calculations. J Mol Graph Model 25:247–260
30. Van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC (2005) GROMACS: fast, flexible, and free. J Comput Chem 26:1701–1718
31. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theor Comput 4:435–447
32. Sorin EJ, Pande VS (2005) Exploring the helix-coil transition via all-atom equilibrium ensemble simulations. Biophys J 88:2472–2493
33. DePaul AJ, Thompson EJ, Patel SS, Haldeman K, Sorin EJ (2010) Equilibrium conformational dynamics in an RNA tetraloop from massively parallel molecular dynamics. Nucleic Acids Res 38:4856–4867
34. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81:3684–3690
35. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: a linear constraint solver for molecular simulations. J Comput Chem 18:1463–1472
36. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: An N • log(N) method for Ewald sums in large systems. J Chem Phys 98:10089–10092
37. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. J Chem Phys 103:8577–8593
38. Åqvist J, Medina C, Samuelsson JE (1994) A new method for predicting binding affinity in computer-aided drug design. Protein Eng 7:385–391
39. Hansson T, Marelius J, Åqvist J (1998) Ligand binding affinity prediction by linear interaction energy methods. J Comput Aided Mol Des 12:27–35
40. Wallace AC, Laskowski RA, Thornton JM (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng 8:127–134