BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2009,** *48,* 8435–8448 8435
DOI: 10.1021/bi900729a

Quantitative Prediction of Fold Resistance for Inhibitors of EGFR†

Trent E. Balius‡ and Robert C. Rizzo*,‡,§

‡*Department of Applied Mathematics and Statistics, and* §*Institute of Chemical Biology and Drug Discovery, Stony Brook University, Stony Brook, New York 11794*

Received April 28, 2009; Revised Manuscript Received July 17, 2009

ABSTRACT: Clinical use of ATP-competitive inhibitors of the epidermal growth factor receptor (EGFR) kinase domain can lead to an acquired drug resistant mutant L858R&T790M which dramatically reduces binding affinity relative to a prevalent cancer causing mutation L858R. In this study, we have used molecular dynamics (MD) computer simulations, free energy calculations (MM-GBSA method), and per-residue footprint analysis to characterize binding of three inhibitors (erlotinib, gefitinib, and AEE788) with wildtype EGFR and three mutants. The goal is to characterize how variation in structure and energy correlate with changes in experimental activities and to deduce origins of drug resistance. For seven fold resistance values, each computed from the difference of two independent computer simulations, excellent agreement was obtained with available experimental data ($r^2 = 0.84$). Importantly, the results correctly predict that affinity will increase as a result of L858R and decrease due to L858R&T790M. Per-residue analysis shows an increase in favorable packing at the site of the methionine mutation reaffirming that a steric clash hypothesis is unlikely; however, large losses in van der Waals, Coulombic, and H-bond interactions strongly suggest that resistance is not due solely to changes in affinity for the native substrate ATP as recently proposed. Instead, the present results indicate that drug resistance more likely involves disruption of favorable interactions, including a water-mediated H-bond network between the ligands and residues T854, T790, and Q791, which could have important implication for guiding rational design of inhibitors with improved resistance profiles.

Cancer is the second highest cause of death within the United States led by lung and bronchial cancers for which an estimated 215 000 new cases and 161 000 deaths were reported in 2008 (*1*). Nonsmall cell lung cancer (NSCLC)[1] comprises the largest subset of lung cancers (*2*). A major oncogene that drives tumorigenesis in NSCLC, as well as other types of cancer, is the membrane receptor tyrosine kinase epidermal growth factor receptor (EGFR). Overexpression of EGFR is observed in 62% of NSCLC tumors (*3*), and its role in mediating tumor cell growth and survival for NSCLC, as well as many other types of cancer, has been well described (*4*−*6*). The importance of EFGR has been clinically validated, and within the past several years, inhibitors of EGFR have been approved for treatment of NSCLC, pancreatic, colorectal, head and neck, and breast cancers (*4*, *6*, *7*).

Structurally, EGFR can be divided into five distinct regions: an extracellular ligand binding domain, a transmembrane domain, an intracellular juxtamembrane domain, an intracellular tyrosine kinase domain (TKD), and a C-tail region where phosphorylation occurs (*8*, *9*). Normally, EGFR is a monomer. However, extracellular ligand binding of endogenous epidermal growth factor (EGF) promotes dimerization with another protein from the ErbB family such as EGFR (ErbB1 or HER1), ErbB2 (HER2), ErbB3 (HER3), or ErbB4 (HER4) (*4*, *10*). The homo- or heterodimerization event induces a conformational shift in the TKD from an inactive to active form (*8*, *9*, *11*). Activation results in binding of ATP, phosphorylation, and signal transduction through a number of downstream pathways (*4*, *10*, *12*). Normally, signaling activity is under tight regulatory control. However, cancer causing mutations can result in constitutive activation of EGFR (*9*). ATP-competitive inhibitors have been described that preferentially bind the active or inactive conformation (*13*, *14*). The primary structural differences between the active and inactive forms is a conformational shift in the TKD activation loop and movement of the N-lobe helix, both of which are located near the ATP binding site (Figure 1).

There are two classes of inhibitors of EGFR: (i) monoclonal antibodies such as Cetuximab (IMC-C225) which target the extracellular domain and block binding of native EGF ligand to the receptor, and (ii) small molecules that compete with ATP in the intracellular TKD and block activity, regardless of endogenous ligand binding (Table 1) (*4*−*6*, *15*). Focusing on ATP competitive inhibitors, approved small molecules of the TKD domain include erlotinib (Tarceva, OSI Pharmaceuticals), gefitinib (Iressa, AstraZeneca), and lapatinib (Tykerb, GlaxoSmithKline) (*7*). A fourth compound called AEE788 (Novartis) is in development (*16*). Although erlotinib and gefitinib primarily target EGFR, multireceptor inhibition is possible given the high structural homology of the TKD (*4*, *13*, *16*, *17*). Lapatinib is a dual inhibitor of EGFR and ErbB2 (*13*, *15*) and AEE788 binds EGFR, ErbB2, and the related VEGF receptor (*4*, *16*). Erlotinib is label-approved for use against NSCLC and pancreatic cancer,

†This work was funded in part by the Stony Brook University Office of the Vice President for Research and the School of Medicine (Carol M. Baldwin Breast Cancer Research Award), and NIH Grants F31CA134201 (to T.E.B.) and R01GM083669 (to R.C.R.).

*Corresponding author. E-mail: rizzorc@gmail.com. Phone: 631-944-2891. Fax: 631-632-8490.

[1]EGFR, epidermal growth factor receptor; FR, fold resistance; MD, molecular dynamics; MM-GBSA, molecular mechanics generalized born surface area; NSCLC, non-small cell lung cancer; rmsd, root-mean-square-deviation; TKD, tyrosine kinase domain.

© 2009 American Chemical Society
Published on Web 07/23/2009

while lapatinib is approved to treat patients with advanced or metastatic breast cancer whose tumors also overexpress HER2 (*7*). Gefitinib was originally approved to treat NSCLC; however, the FDA has limited its usage given that no significant effect on patient survival was found (*7*). AEE788 is being evaluated as therapy for brain and central nervous system cancers (*16*).

Several cancer causing mutations in EGFR have been reported which map to either the extracellular ligand binding domain (e.g., in glioblastoma) (*19*, *22*) or the TKD region (e.g., in NSCLC) (*23*−*27*), which cause activation of EGFR independent of EGF-ligand binding. For the TKD domain, such cancer causing mutations can occur at positions L858R or G719S (point mutations), and E746-A750 or E746-S752 (referred to as exon 19 deletions) (*23*−*27*). L858R and exon 19 deletions are the most frequent mutations in NSCLC (*28*, *29*). Figure 1 shows the TKD domain cancer causing mutations mapped to EGFR in red. Although patients with wild type EGFR benefit from low molecular weight inhibitors, patients whose tumors harbor activating L858R or deletion mutations are especially responsive to erlotinib and gefitinib treatment (*25*). Interestingly, in contrast to most systems in which mutations lead to a decrease in binding, studies have shown that affinity is enhanced for TKD ligands for L858R over wild type (Table 1). As these ligands bind preferentially to the active conformation this could explain their enhanced binding affinity toward the mutants. In contrast, decreases in affinity for gefitinib and AEE788 have been reported relative to wildtype for the G719S point mutant (Table 1).

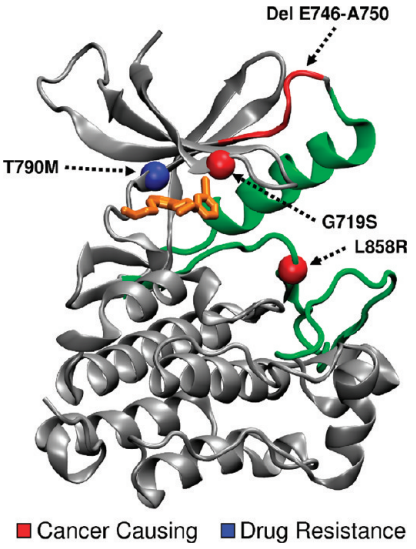
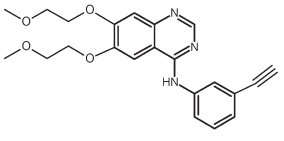
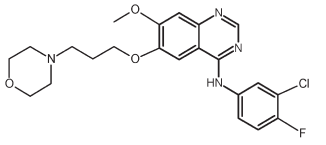
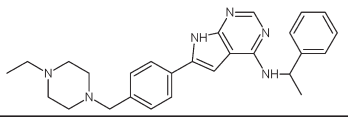


FIGURE 1: Ribbon diagram showing EGFR complexed with the ATP-competitive inhibitor erlotinib. Regions which change conformation (N-lobe helix and activation loop) upon receptor activation are shown in green. Locations of cancer causing mutations (deletion or point) which cause receptor activation are in red. The secondary T790M drug resistance mutation is shown in blue. Coordinates from pdb code 1M17 (*14*).

As with many chemotherapeutics, acquired resistance to current EGFR inhibitors can occur with continued use (*30*). A T790M resistance-mutation is commonly observed in patients treated with erlotinib and gefitinib for those tumors which also harbor the primary cancer causing point mutation at position L858R or exon 19 deletions (*30*). The location of T790M is shown mapped in blue to the TKD site on EGFR in Figure 1. As shown in Table 1, the double mutant L858R&T790M shows marked decrease in inhibition when compared to the active L858R mutant alone for all three ligands. Erlotinib in particular shows a large >800 fold resistance (FR = ratio of activities). Similar secondary drug resistance mutations have been described previously for other molecular targeted therapeutics such as Gleevec. The well-known BCR-Abl kinase resistance mutation at position T315I, referred to as the "gate keeper", arises from treatment with Gleevec (*31*, *32*). The gate keeper is important in modulating selectivity and affinity for BCR-Abl inhibitors and analogous here to the point mutant in EGFR at position T790M.

An improved understanding of the molecular determinates that drive ligand binding for EGFR is critical for development of improved inhibitors. Prior computational studies of this system have included use of homology and molecular modeling (*33*), comparative molecular field analysis (CoMFA) (*34*), virtual

Table 1: Experimental Fold Resistance (FR) Values for ATP-Competitive Inhibitors with EGFR

Inhibitor	Structure	Experimental Fold Resistance[a]		
		L858R / WT	L858R&T790M / L858R	G719S / WT
erlotinib		6.25 / 17.5 nM[b] 0.36 FR −0.61 $\Delta\Delta G_{FR}$	>10000 / 12.5 nM[c] >800 FR >3.96 $\Delta\Delta G_{FR}$	–
gefitinib		2.4 / 35.3 nM[d] 0.068 FR −1.59 $\Delta\Delta G_{FR}$	10.9 / 2.4 nM[d] 4.54 FR 0.90 $\Delta\Delta G_{FR}$	123.6 / 53.5 nM[e] 2.31 FR 0.50 $\Delta\Delta G_{FR}$
AEE788		1.1 / 5.3 nM[d] 0.21 FR −0.92 $\Delta\Delta G_{FR}$	18.6 / 1.1 nM[d] 16.9 FR 1.68 $\Delta\Delta G_{FR}$	11.3 / 10.9 nM[e] 1.04 FR 0.02 $\Delta\Delta G_{FR}$

[a] Fold resistance (FR) = ratio of experimental activities. $\Delta\Delta G_{FR}$ exptl ≈ RT ln(FR) at 298.15 K in kcal/mol. [b] K_i values (nM) from Carey et al. (*18*). [c] IC_{50} values (nM) from Ji et al. (*19*). [d] K_d values (nM) from Yun et al. (*20*). [e] K_d values (nM) from Yun et al. (*21*).

screening (*35*), and molecular dynamics (*33*, *34*, *36*, *37*). Use of MM-PBSA methods, similar in principle to the calculations employed in the present manuscript, were reported by Hou et al. (*34*) for refinement of docked ligand poses, and by Liu et al. (*36*) to study the impact of point mutations on binding for gefitinib. Surprisingly, there have been few all-atom molecular dynamics studies reporting quantitative binding energy comparisons between theory and experiment for ligands with EGFR. In this report, we have carried out simulations of the TKD of EGFR in complex with three ATP-competitive inhibitors to investigate the effects of clinically relevant point mutations on ligand binding. Studies to address deletion mutations are in progress. Specifically, goals of the present project are 3-fold: (i) Develop robust quantitative computational models to study EGFR-ligand binding for wildtype, L858R, G719S, and the drug resistant double mutant L858R&T790M. (ii) Determine how variation in structural and energetic results correlate with variation in reported experimental activities. (iii) Deduce the origins of drug resistance. Characterization of FR at the molecular level will ultimately enable development of next generation compounds with improved resistance profiles.

METHODS

Binding Free Energies. Accurate calculation of protein−ligand binding energies remains an important and challenging problem. In this report, we employ the molecular mechanics Generalized Born solvent accessible surface area (MM-GBSA) method (*38*, *39*) to computationally estimate binding free energies (ΔG_b) for inhibitors with EGFR. Although considered to be an approximate free energy theory, the benefits of MM-GBSA include relative ease of set up, low computational overhead, and systems with dissimilar topologies can be more readily examined in comparison to other methods such as free energy perturbation. Projects in our laboratory employing similar protocols to that reported here, which successfully used MM-GBSA to study protein−ligand binding, include HIVgp41 (*40*), neuraminidase (*41*), and MMP-13 (*42*).

The method relies on the thermodynamic relationship shown in Figure 2 to estimate the free energy of binding (ΔG_b calcd/exptl) which occurs in the condensed phase. The computed free energy of binding is estimated as the sum of nonbonded gas-phase (ΔG_{gas}) interactions modulated by the overall change in hydration free energy ($\Delta\Delta G_{hyd}$) for the complexation event

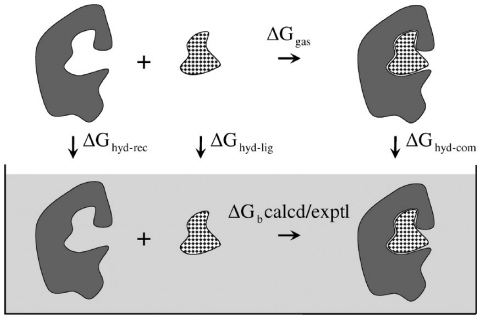


FIGURE 2: Schematic representation of the thermodynamic cycle used to calculate free energies of binding (ΔG_b calcd) for comparison with experiment (ΔG_b exptl). The cycle highlights the relationship between ΔG_b exptl occurring in condensed phase with the free energy of interaction in the gas-phase (ΔG_{gas}) modulated by three terms representing the free energy of hydration (ΔG_{hyd}) for the transfer from vacuum to water for each separate species (com = complex, rec = receptor, lig = ligand).

(*38*, *39*). The hydration term accounts for important desolvation penalties, which include changes in entropy due to the hydrophobic effect, that occur as a result of unbound solvated species coming together to form a complex. Additional terms to include estimates for changes in solute entropy were not included in the present study.

A molecular dynamics trajectory of each protein−ligand complex is performed in explicit solvent with system energies, as well as root-mean-square-deviations (rmsd), being monitored for stability and convergence. For MM-GBSA analysis, the explicit solvent is stripped off and coordinates are separated into three individual species (complex, receptor, and ligand) with eqs 1−3 being used to compute the total binding affinity (*38*, *39*). The relevant individual energy terms include van der Waals (ΔE_{vdw}), Coulombic (ΔE_{coul}), polar (ΔG_{polar}), and nonpolar ($\Delta G_{nonpolar}$) contributions. Generalized Born (ΔG_{polar}) and solvent accessible surface area ($\Delta G_{nonpolar}$) calculations are used to estimate ΔG_{hyd} for each individual species (*43*).

$$\Delta G_b \text{ exptl} \approx \Delta G_b \text{ calcd} = \Delta G_{gas} + \Delta G_{hyd-com} - (\Delta G_{hyd-rec} + \Delta G_{hyd-lig}) \quad (1)$$

$$\Delta G_{gas} = \Delta E_{vdw} + \Delta E_{coul} \quad (2)$$

$$\Delta G_{hyd-species} = \Delta G_{polar} + \Delta G_{nonpolar} \quad (3)$$

Interaction Signatures: Molecular Footprints. To identify important binding site residues and characterize how interactions may change as a result of mutation, structural and energetic molecular "footprints" were computed for each MD trajectory. Footprints represent the per-residue decomposition of interactions, averaged over the production simulations, between each EGFR residue and the inhibitors. Our laboratory has successfully used such footprints to deduce origins of resistance conferred by a R292K mutation for sialic acid-based inhibitors of neuraminidase (*41*), and to show that the hydrophobic pocket region on HIVgp41 is an important drug target site for modulating binding affinity (*40*). Separate footprints for Coulombic and van der Waals energy, as well as hydrogen bonds, were computed. Difference footprints were also computed, using results from the L858R&T790M − L858R simulations, and represent the change in energy (or H-bonds) at each residue due to mutation.

System Setups. A single set of receptor coordinates (pdb code 1M17) (*14*), of EGFR in the active form, was used as the basis for construction of all simulation setups. All solvent was removed from the 1M17 structure, and only the TKD of EGFR (defined as a.a. numbers 710−983) were retained. Initial geometries for ligands erlotinib, gefitinib, and AEE788 were obtained from 1M17 (*14*), 2ITY (*21*), and 2J6M (*21*) pdb codes, respectively, and placed into the master 1M17 reference frame through alignment of receptor backbone atoms in common with each pdb. Mutant forms of EGFR (L858R, L858R&T790M, and G719S) were obtained through manual modification of 1M17 to the desired residue(s). Starting rotameric states for modified side chains were made using energetic packing consideration subject to visual inspection to ensure there were no intermolecular clashes as a result of model building. The MOE (*44*) program was used for initial preparation of ligand (mol2 format) and receptor (pdb format) files for subsequent processing. The AMBER8 (*45*) program modules *leap* and *antechamber* were used to assemble, solvate (10 Å buffer), and assign force-field parameters for each complex consisting of FF99SB (*46*)

(protein), TIP3P (*47*) (solvent), and GAFF (*48*) (ligand). For the ligands, partial atomic charges were obtained at the HF/6-31G*// HF/6-31G* level of theory via the ChelpG (*49*) method using Gaussian98 (*50*). All ligands were modeled as having a net zero charge. Unless otherwise stated, system setups employed default input parameters for each program. The size of the complete model was 274 receptor residues plus one ligand residue solvated in a TIP3P periodic solvent box of ca. size 77 × 89 × 75 Å^3 containing ca. 14050 waters.

Simulation Protocols. A nine step equilibration protocol, consisting of short energy minimizations and molecular dynamics (MD), was used to eliminate any unfavorable interactions which may have occurred as a result of model building and to gently adjust the starting structure to the molecular mechanics force field prior to production MD. Heavy atoms of the complex were initially restrained to the crystallographic coordinates using a harmonic restraint force constant of 5 kcal/mol Å^2 with water molecules and hydrogen atoms free to move during 1000 steps of steepest decent energy minimization (step 1). Each subsequent equilibration step used the last set of coordinates from the previous step as the restraint reference structure. Next, the same restraint mask and coefficient were used for 50 ps of MD in which waters and hydrogens were further able to adjust (step 2). This was followed by three minimizations of 1000 steps each in which the restraints were reduced from 2, to 0.1, to 0.05 kcal/mol Å^2, respectively (steps 3−5). Three additional MD runs of 50 ps were run where weights were reduced from 1 to 0.5 (steps 6−7) followed by 0.1 kcal/mol Å^2 (step 8) with only backbone receptor atoms at Cα, C, and N being restrained in the latter step. The final 50 ps MD equilibration step used the 0.1 kcal/mol Å^2 receptor backbone weight, but only the last three residues on each N- and C-terminus were restrained (step 9). Production MD employed the same weak restraints as the final equilibration step and was extended to 5 ns with the restraint reference updated every ns. A time step of 1 fs was used for equilibration and increased to 2 fs for production MD which concurrently required use of the SHAKE (*51*) algorithm. Coordinates for post processing were saved every 1 ps. Long-range electrostatics were computed using the particle mesh Ewald (PME) (*52*) with a real-space cut off of 8 Å. A constant temperature of 298.15 K and pressure of 1 bar was maintained during the simulations through Berendsen schemes (*53*) with heat bath coupling and pressure relaxation time constants of 1.0 ps. These calculations employed the AMBER8 *sander* module.

Analysis. Binding free energies, fold resistance, and molecular footprints were obtained from postprocessing of each protein−ligand complex MD trajectory. All waters were removed unless otherwise stated. Individual snapshots (N = 5000) were split into coordinates representing separate ligand, receptor and complex and single point calculations using *sander* were performed to obtain the energy terms (ΔE_{coul}, ΔE_{vdw}, ΔG_{polar}, $\Delta G_{nonpolar}$) used to compute free energies of binding (Figure 2, eqs 1−3). As in previously reported studies from our laboratory (*40*−*42*), the GB model implemented into AMBER8 described by Onufriev et al. (*54*) (type igb = 5) was used to estimate polar energies (ΔG_{polar}) with mbondi2 radii and dielectric constants of 1 and 78.5. Nonpolar energies ($\Delta G_{nonpolar}$) were obtained from solvent accessible surface area calculations via $\Delta G_{nonpolar} = \gamma\,\mathrm{SASA} + \beta$ using standard constants of γ = 0.00542 kcal/mol Å^2 and β = 0.92 kcal/mol (*38*, *55*). GBSA calculations for molecules containing fluorine and chlorine employed radii of 1.50 Å (F) and 1.70 Å (Cl) which required modification to the AMBER8 distribution file src/sander/mdread.f. In-house scripts were used to compute the per-residue decomposition (molecular footprints) for intermolecular H-bonds, Coulombic, and van der Waals interactions for which the sum over all the EGFR residues is equivalent to the total value (i.e., ΔE_{coul}, ΔE_{vdw}, H-bond). Hydrogen bonds were defined as a structural interaction between three atoms X_D−H_D----X_A with a distance less than or equal to 2.5 Å and angle between 120 and 180 degrees. The NAMD program (*56*) was used to compute and gauge the importance of two highly populated waters involved in a network of water-mediated ligand H-bonds through calculation of the pairwise Coulombic and van der Waals energies between the waters ($2H_20$ = species 1) and the interactions partners (T790/M790, Q791, T854, or ligand = species 2).

For the analysis of fold resistance (FR), the experimental free energies are estimated as $\Delta\Delta G_{FR}$ exptl $\approx RT\ln(\mathrm{FR})$ at 298.15 K using FR ratios in Table 1. It should be emphasized that FR is defined as the ratio of two activities (mutant/wildtype); thus it is important that both measurements be made under the same conditions, which usually implies that data be obtained from the same laboratory to minimize artifacts resulting from different experimental protocols (i.e., assay conditions). Computationally, fold resistance energies ($\Delta\Delta G_{FR}$ calcd) are defined as the difference in predicted free energies of binding from two independent simulations (e.g., $\Delta G_{L858R} - \Delta G_{WT}$) as obtained in each case via eqs 1−3.

RESULTS AND DISCUSSION

Simulation Stability. To assess the behavior of the MD simulations, and gauge the robustness of results, structural root-mean-square-deviations (rmsds) and system energies were examined as a function of time. As demonstrated in Figure 3, which is representative, simulations of erlotinib complexed with EGFR for wild type, L858R, and L858R&T790M show reasonable stability in plots for the estimated free energies of binding (ΔG_b calcd) and for rmsds. Block-smoothed energies (black line) indicate good behavior when computed from the running average of the previous 100 instantaneous snapshot energies (gray points) as shown in Figure 3 left. On the right in Figure 3 are shown instantaneous rmsds values for the EGFR backbone at Cα, C, N, and O (blue line), erlotinib heavy atoms (green line), and erlotinib quinazoline ring atoms (red line). Here, the reported rmsd values are obtained after each individual snapshot is fit to the first frame of the production MD runs using receptor Cα backbone atoms as the match criteria. Thus, rmsd values for the ligands reflect variation in both internal geometry as well as rigid body movement relative to the protein. At some points in the trajectories, erlotinib reveals larger than expected rmsd values (> 2 Å) which could be of concern (Figure 3, green lines). However, examination of rmsds for only the central fused-ring quinazoline scaffold (see Table 1) shows much lower values (Figure 3, red lines), which indicates erlotinib remains anchored in the binding pocket and it is only the solvent-exposed flexible r-groups extending off the quinazoline ring which fluctuate significantly. Other simulations behave similarly, with gefitinib (4-aminoquinazoline) and AEE788 (pyrrolopyrimidine) scaffolds showing less movement than the overall ligand. In all cases, rmsds results for the receptor backbone are well behaved and low (< 2 Å), which additionally indicate robust simulation behavior (Figure 3, blue lines).

Comparison with Crystallographic Structures. Only a single crystallographic structure (*14*) of active form EGFR

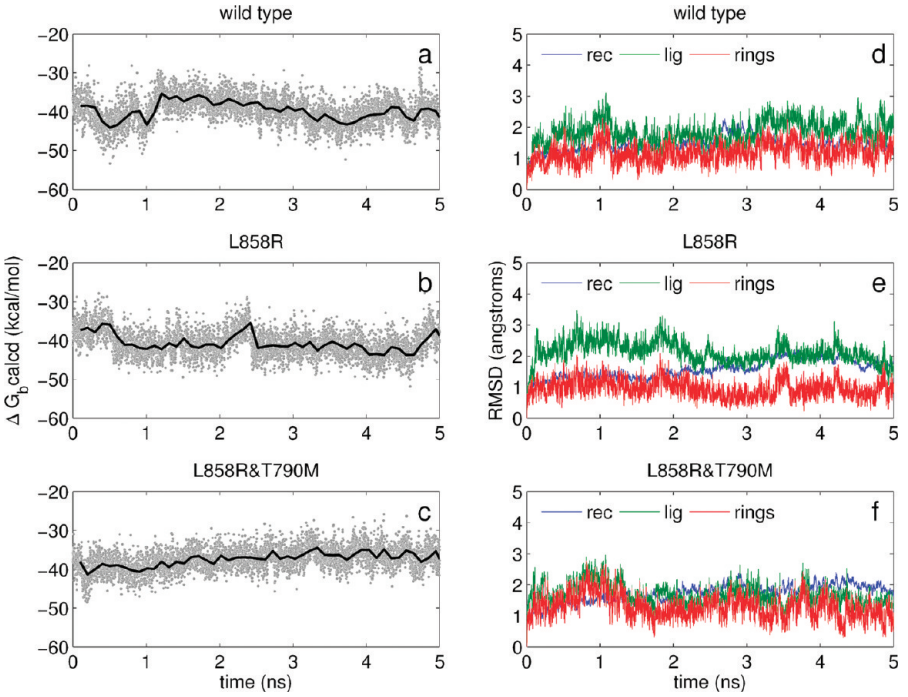


FIGURE 3: Fluctuations in computed free energies of binding (ΔG_b calcd) and root-mean-square-deviation (rmsd) for erlotinib with wild type EGFR (panels a and d), L858R (panels b and e), and L858R&T790M (panels c and f) vs time. Left panels show instantaneous energies (gray dots) and block-running averages over 100 frames (black line). Right panels show rmsds for receptor backbone (blue line), ligand (green line), and ligand quinazoline scaffold atoms (red line).

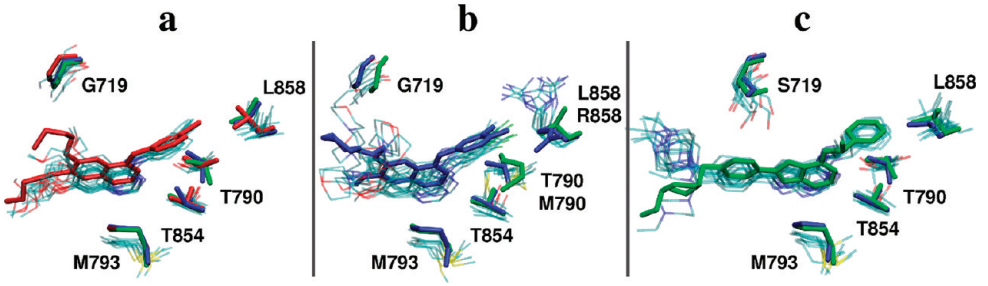


FIGURE 4: Representative snapshots from MD simulations of ligands with EGFR showing side-chain sampling of key residues (thin lines, $N = 10$ each) vs crystallographic conformations (bold lines) for erlotinib (a), gefitinib (b), and AEE788 (c). Pdb codes for **4a**: red = 1M17 (erlotinib with wildtype), blue = 2ITY (gefitinib with wildtype), green = 2J6M (AEE788 with wildtype). Pdb codes for **4b**: blue = 2ITZ (gefitinib with L858R), green = 2JIU (AEE788 with T790M). Pdb codes for **4c**: blue = 2ITO (gefitinib with G719S), green = 2ITP (AEE788 with G719S).

(pdb code 1M17) was available at the time of our initial mutant setups for erlotinib. All subsequent simulations employed the same set of protein coordinates originally derived for this ligand. However, other EGFR structures, including those with several of the mutations studied here, have been reported (*20*, *21*). To structurally compare the theoretical and experimental results, as well as assess computer sampling during the simulations, evenly spaced MD snapshots were individually fit to available crystallographic structures again using Cα backbone atoms as the match criteria. As shown in Figure 4 for three representative simulations, sampling of ligand positions is consistent with the experimentally observed conformations. And as expected, solvent-exposed regions of the ligands visually show greater movement than for scaffolds, which is in agreement with the rmsd results plotted in Figure 3 (panels d−f). Notably, Figure 4 highlights how key crystallographic positions of important side chains including G719 (or S719 mutant), T854, L858, T790 (or M790 mutant), and M793 are well-sampled during the simulations and consistent with experiment. An anomaly is a difference in the rotameric states sampled for L858R vs the crystal structure (Figure 4b, blue line). Here, the MD simulations sample a solvent-exposed Arg conformation as opposed to an intramolecular H-bond as seen in the crystal structure. Although longer MD simulations might be required to sample the experimental L858R rotamer, since both conformations appear to be physically reasonable, an alternative would be to begin simulations using crystallographic coordinates of L858R instead of those based on the 1M17 models. However, available EGFR structures of L858R mutants show disorder in the residue range spanning 867−875. Additionally, the activation loop region in these structures (defined as 855−876) (*14*) adopts a unique conformation which is dependent on the choice of crystallographic buffer conditions. (*21*) Thus, given the considerable ambiguity in how to model nine missing residues (a.a. 867−875) into an unknown activation loop conformation, we have elected to retain the models originally constructed using 1M17 containing the complete loop. As described below, the good agreement between computational and experimental activities obtained using the 1M17-derived coordinates suggests these are reasonable models to study EGFR-ligand binding in the kinase active form.

Table 2: Experimental versus Calculated Fold Resistance (FR) Energies ($\Delta\Delta G_{FR}$) and Energy Components for Ligands with EGFR

inhibitor	$\Delta\Delta E_{vdw}$	$\Delta\Delta E_{coul}$	$\Delta\Delta G_{polar}$	$\Delta\Delta G_{nonpolar}$	$\Delta\Delta G_{FR}$ calcd[a]	$\Delta\Delta G_{FR}$ exptl[b]
	A	B	C	D	E = (A+B+C+D)	F
			L858R − WT			
erlotinib	−0.86 ± 0.06	−0.34 ± 0.13	0.21 ± 0.11	0.06 ± 0.003	−0.97 ± 0.07	−0.61
gefitinib	−0.99 ± 0.06	−0.72 ± 0.07	−0.58 ± 0.06	−0.01 ± 0.004	−2.30 ± 0.07	−1.59
AEE788	−2.41 ± 0.06	−0.48 ± 0.07	0.36 ± 0.06	−0.30 ± 0.005	−2.84 ± 0.06	−0.92
			L858R&T790M − L858R			
erlotinib	2.30 ± 0.06	7.42 ± 0.11	−6.56 ± 0.10	0.09 ± 0.003	3.30 ± 0.06	> 3.96
gefitinib	−0.10 ± 0.05	−0.06 ± 0.07	0.49 ± 0.06	−0.06 ± 0.004	0.27 ± 0.06	0.90
AEE788	3.39 ± 0.07	3.15 ± 0.09	−4.33 ± 0.07	0.20 ± 0.004	2.40 ± 0.08	1.68
			G719S − WT			
erlotinib	−2.08 ± 0.06	−0.05 ± 0.12	−0.24 ± 0.11	0.04 ± 0.003	−2.38 ± 0.07	not reported
gefitinib	0.74 ± 0.07	−0.85 ± 0.07	1.59 ± 0.07	0.04 ± 0.004	1.50 ± 0.08	0.50
AEE788	−0.65 ± 0.06	−0.78 ± 0.06	0.55 ± 0.05	0.08 ± 0.005	−0.81 ± 0.07	0.02
r^2 =	0.70	0.47	0.19	0.30	0.84	7 data points[c]

[a] $\Delta\Delta G_{FR}$ calcd derived from the difference of two independent simulations (eg L858R − WT) computed using eqs 1−3. [b] $\Delta\Delta G_{FR}$ exptl values from Table 1. Correlations coefficients (r^2 values) obtained from fitting the change in each energy component to $\Delta\Delta G_{FR}$ exptl. All energies in kcal/mol ± standard errors of the mean from 5000 MD snapshots. [c] Data point for erlotinib with double mutant (> 3.96) excluded from r^2 calculations given ambiguity in the experimental $\Delta\Delta G_{FR}$ measurement.

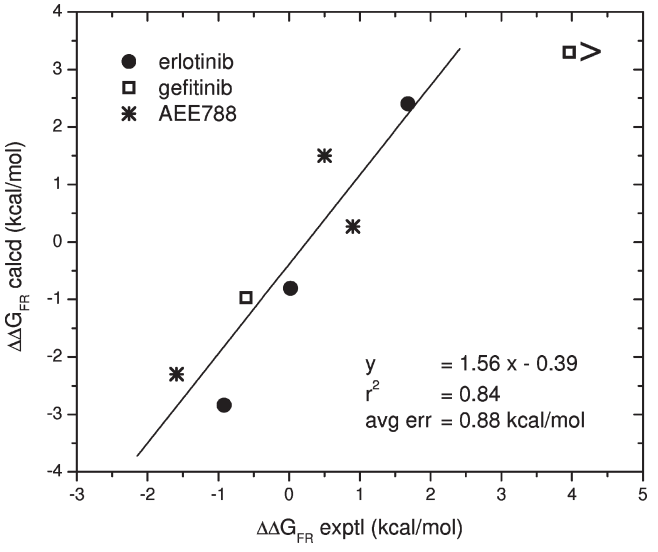


FIGURE 5: Predicted FR energies ($\Delta\Delta G_{FR}$ calcd) vs experimental FR energies ($\Delta\Delta G_{FR}$ exptl) for inhibitors with EGFR. Each point is the difference between results from two independent MD simulations (16 simulations total) from 5000 MD snapshots each. Data point for erlotinib with double mutant (> 3.96) excluded from r^2 calculations given ambiguity in the experimental $\Delta\Delta G_{FR}$ measurement.

Correlation with Experimental Fold Resistance. Overall, the computational results are strongly correlated with the experimental fold resistance values as shown in Table 2 and graphically plotted in Figure 5. Calculated values represent average quantities obtained over 5000 MD snapshots. Low standard errors of the mean (sem) indicate the energetic results are converged. Notably, the computational results correctly predict that affinity is always enhanced (negative $\Delta\Delta G_{FR}$ values) for all three ligands with the cancer causing L858R EGFR mutation relative to wildtype (Table 2 columns E vs F). Further, results for the drug resistant double mutant (L858R&T790M) correctly predict that decreases (positive $\Delta\Delta G_{FR}$ values) will occur in binding relative to L858R alone (Table 2 columns E vs F). Compellingly, the magnitudes for the energetic changes which occur across the inhibitor series in Table 2 are in excellent agreement with experiment. For example, results for erlotinib ($\Delta\Delta G_{FR}$ calcd = 3.30 vs $\Delta\Delta G_{FR}$ exptl > 3.96 kcal/mol) and AEE788 ($\Delta\Delta G_{FR}$ calcd = 2.40 vs $\Delta\Delta G_{FR}$ exptl = 1.68 kcal/mol) both show much larger computational and experimental FR values for the double mutant relative to gefitinib ($\Delta\Delta G_{FR}$ calcd = 0.27 vs $\Delta\Delta G_{FR}$ exptl = 0.90 kcal/mol) which is less affected. Despite the fact that the simulations correctly predict AEE788 to bind more tightly to L858R, a minor discrepancy is the improper rank ordering for L858R−WT relative to gefitinib. In terms of sign, the sole outlier in Table 2 is for AEE788 for which the G719S/WT fold resistance yields essentially no energetic change experimentally but our calculations show enhanced affinity. Interestingly, a prediction for the effect of G719S on binding of erlotinib also shows enhanced affinity (Table 2). FR calculations for gefitinib with G719S yield the correct experimental trend. Despite the one outlier, there is excellent accord overall, and a linear fit between the data points shows a strong correlation coefficient of $r^2 = 0.84$ (Figure 5, Table 2), which indicates the simulations well reproduce trends in the experimental FR energies.

Examination of the individual terms which comprise $\Delta\Delta G_{FR}$ calcd along with calculation of correlation coefficients (r^2 values) for each term with $\Delta\Delta G_{FR}$ exptl was done to pinpoint which term(s) best explain experimental variation and thus resistance. It should be noted that due to ambiguities in the experimental FR measurement for erlotinib with the double mutant (> 3.96 kcal/mol, Table 2) all fittings excluded this data point. For L858R relative to wildtype EGFR, all three inhibitors show more favorable van der Waals and Coulombic interactions which lead to an overall stronger computed $\Delta\Delta G_{FR}$ in agreement with experiment (Table 2 columns A and B). For the drug resistant mutant (L858R&T790M − L858R), the most dramatic losses observed experimentally correlate with the large computed losses in van der Waals and Coulombic energy for erlotinib ($\Delta\Delta G_{FR}$ > 3.96, $\Delta\Delta E_{vdw}$ = 2.30, $\Delta\Delta E_{coul}$ = 7.42 kcal/mol) and AEE788 ($\Delta\Delta G_{FR}$ = 1.68, $\Delta\Delta E_{vdw}$ = 3.39, $\Delta\Delta E_{coul}$ = 3.15 kcal/mol). For

gefitinib with the double mutant the less deleterious effect on binding ($\Delta\Delta G_{FR}$ = 0.90 kcal/mol) appears to be solely from changes in desolvation ($\Delta\Delta G_{polar}$ = 0.49 kcal/mol) given the minor changes computed in the other terms ($\Delta\Delta E_{coul}$ = −0.06, $\Delta\Delta E_{vdw}$ −0.10 kcal/mol). For the G719S mutation relative to wildtype, binding losses for gefitinib again appear to be a result of increased desolvation ($\Delta\Delta G_{polar}$ = 1.59 kcal/mol) as any gains computed in Coulombic energy are offset by reduction in steric packing ($\Delta\Delta E_{coul}$ = −0.85 vs $\Delta\Delta E_{vdw}$ = 0.74 kcal/mol). For AEE788 with G719S, the previously noted disagreement between computed and experimental affinities for this data point renders component analysis here indeterminate. Affinity for erlotinib with G719S is predicted to be enhanced primarily as a result of increased van der Waals interactions.

Notably, the most correlated term in Table 2 with experiment is for $\Delta\Delta G_{FR}$ calcd (r^2 = 0.84) indicating that for these systems a balance of energetic terms is most important for describing changes in FR. Of the individual components, changes in van der Waals energy ($\Delta\Delta E_{vdw}$ r^2 = 0.70) show the largest r^2 value followed by Coulombic ($\Delta\Delta E_{coul}$ r^2 = 0.47), nonpolar $\Delta\Delta G_{nonpolar}$ (r^2 = 0.30), and polar desolvation energies ($\Delta\Delta G_{polar}$ = 0.19). The low r^2 value of 0.04 obtained for the sum of $\Delta\Delta E_{coul}$ and $\Delta\Delta G_{polar}$ vs experiment suggests that steric packing probably contributes more to variation in FR as opposed to changes in solvent mediated electrostatics. Interestingly, visually plotting changes in energy components vs $\Delta\Delta G_{FR}$ exptl reveals grouped data in the $\Delta\Delta E_{coul}$ plot which do not appear to lie on the trend line. A fit of this cluster alone leads to an even poorer correlation (Figure 6a dashed line). In contrast, Figure 6b shows how changes in $\Delta\Delta E_{vdw}$ are more closely associated with changes in $\Delta\Delta G_{FR}$ across the entire data set.

Energetics of Binding: What Drives Association? To further characterize how terms contribute to molecular recognition, results from the underlying free energy of binding (ΔG_b) used to determine $\Delta\Delta G_{FR}$ were examined (Table 3). Overall, inhibitor binding appears to be most strongly driven by van der Waals interactions. Values for ΔE_{coul} are always less favorable than ΔE_{vdw} and not sufficient to overcome the competing unfavorable polar desolvation terms (ΔG_{polar}) which suggests steric packing dominates association. For the EGFR variants studied, gefitinib shows stronger ΔE_{vdw} interactions relative to either erlotinib or AEE788. A plot of ΔE_{vdw} vs ΔG_b exptl highlights the separation between gefitinib and AEE788 and additionally shows how changes in van der Waals interactions may track for individual ligands (Figure 7). Although the combined correlation with ΔG_b exptl is poor (ΔE_{vdw} r^2 = 0.004), van der Waals energies for gefitinib (r^2 = 0.83) or AEE788 (r^2 = 0.82) when plotted separately show strong correlation with experiment (Figure 7).

Despite the importance of steric packing, electrostatics in this system appear to play critical roles in mediating affinity. For example, differences in intermolecular H-bonding, as illustrated graphically in Figure 8, likely contribute to enhanced Coulombic interactions for AEE788 and erlotinib relative to gefitinib (Table 3 column B). Average number of H-bonds (Table 3 column G) shows 2.02 interactions for AEE788 with wildtype EGFR followed by erlotinib at 1.82 and gefitinib at 1.16. All the inhibitors show highly populated and significant H-bonding with the backbone amide hydrogen at position M793. A second interaction at M793 for AEE788 largely accounts for the greater number of H-bond relative to the other inhibitors (Figure 8). For erlotinib, an additional significant H-bond is observed between

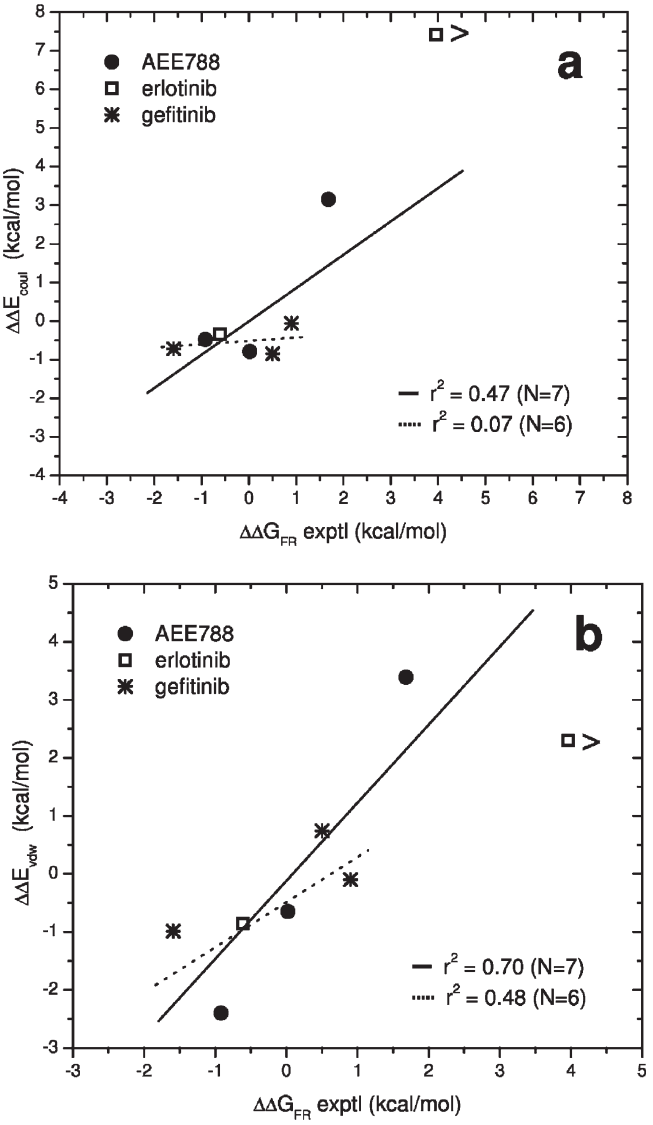


FIGURE 6: Predicted changes in Coulombic ($\Delta\Delta E_{cou}$ panel a) and van der Waals ($\Delta\Delta E_{vdw}$ panel b) energy components versus experimental fold resistance energy ($\Delta\Delta G_{FR}$ exptl) for inhibitors with EGFR. Data point for erlotinib with double mutant (>3.96) excluded from r^2 calculations given ambiguity in the experimental $\Delta\Delta G_{FR}$ measurement.

the backbone at C797 and a terminal O atom for which the other inhibitors have no spatial equivalent (Figure 8). A less populated yet quantifiable interaction for erlotinib includes a unique pi-type H-bond made between the ligands para-alkyne and T790@OH. Pi-type interactions for erlotinib were counted by simply defining the centroid of the alkyne C≡C bond as an H-bond acceptor. Interestingly, the unique H-bond acceptor in erlotinib is replaced by a spatially analogous interaction in gefitinib between the meta-chlorine and T790@OH. AEE788 also shows a weak H-bond at position T790, although this was only observed in the simulation of L858R. Here, a slightly different positioning of AEE788 in the binding pocket relative to the other inhibitors allows for a third H-bond with the pyrrolopyrimidine scaffold (Figure 8).

Origins of Resistance. In order to gauge the relative importance that specific amino acids may contribute to binding, the number of intermolecular H-bonds, van der Waals energy, and Coulombic energy were computed on a per-residue basis. Examinations of H-bond footprint plots (Figure 9) show consistency in overall shape from simulation to simulation, which

Table 3: Absolute Free Energies and Component Decomposition for Inhibitors with EGFR

system	ΔE_{vdw}	ΔE_{coul}	ΔG_{polar}	$\Delta G_{nonpolar}$	ΔG_b calcd	ΔG_b exptl[a]	H-bond
	A	B	C	D	E = A + B + C + D	F	G
				erlotinib			
wildtype	−49.01 ± 0.04	−24.71 ± 0.09	39.73 ± 0.08	−6.05 ± 0.002	−39.69 ± 0.05	−10.58[b]	1.82
L858R	−49.86 ± 0.04	−25.04 ± 0.09	39.94 ± 0.07	−5.99 ± 0.002	−40.66 ± 0.05	−11.19[b]	2.17
L858R&T790M	−47.57 ± 0.05	−17.62 ± 0.07	33.38 ± 0.06	−5.89 ± 0.002	−37.36 ± 0.04	> −6.82[c]	0.99
G719S	−51.09 ± 0.04	−24.76 ± 0.08	39.49 ± 0.07	−6.01 ± 0.003	−42.07 ± 0.05	not reported	1.95
				gefitinib			
wildtype	−53.50 ± 0.05	−14.02 ± 0.05	28.80 ± 0.04	−6.30 ± 0.003	−45.01 ± 0.06	−10.17[d]	1.16
L858R	−54.49 ± 0.04	−14.74 ± 0.04	28.22 ± 0.04	−6.31 ± 0.003	−47.32 ± 0.05	−11.76[d]	1.24
L858R&T790M	−54.59 ± 0.04	−14.80 ± 0.05	28.71 ± 0.05	−6.37 ± 0.003	−47.05 ± 0.05	−10.86[d]	1.05
G719S	−52.76 ± 0.04	−14.87 ± 0.06	30.39 ± 0.05	−6.26 ± 0.002	−43.51 ± 0.05	−9.42[e]	1.08
				AEE788			
wildtype	−50.08 ± 0.05	−21.77 ± 0.04	31.97 ± 0.03	−5.93 ± 0.004	−45.81 ± 0.05	−11.29[d]	2.02
L858R	−52.49 ± 0.04	−22.26 ± 0.06	32.33 ± 0.05	−6.24 ± 0.003	−48.65 ± 0.04	−12.22[d]	2.19
L858R&T790M	−49.10 ± 0.06	−19.11 ± 0.07	28.00 ± 0.05	−6.03 ± 0.003	−46.25 ± 0.07	−10.55[d]	2.48
G719S	−50.73 ± 0.04	−22.56 ± 0.04	32.52 ± 0.03	−5.85 ± 0.003	−46.62 ± 0.04	−10.86[e]	1.99

[a] ΔG_b exptl ≈ RT ln(activities) at 298.15 K in kcal/mol. [b] K_i values (nM) from Carey et al.(*18*) [c] IC_{50} values (nM) from Ji et al. (*19*). [d] K_d values (nM) from Yun et al. (*20*). [e] K_d values (nM) from Yun et al. (*21*).

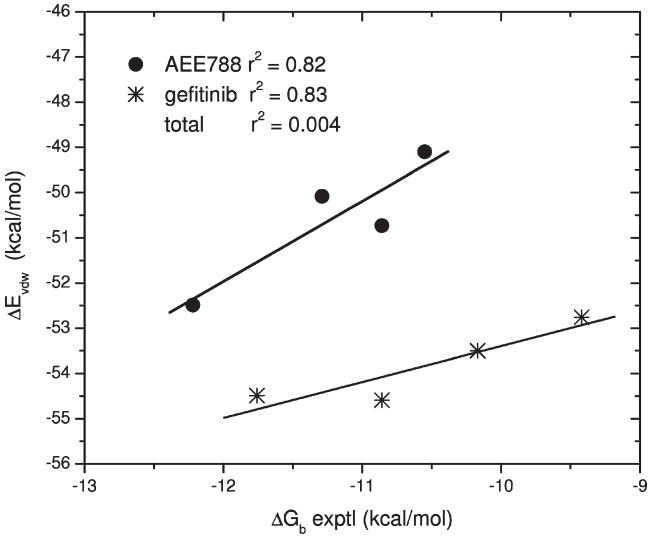


FIGURE 7: Correlation of the van der Waals energy (ΔE_{vdw}) component with ΔG_b exptl. Energies in kcal/mol.

provides additional support that results obtained from averaging 5000 MD frames are converged and well-behaved. As an example, greater number of H-bonds are consistently obtained for AEE788 (2 key scaffold H-bonds) versus other inhibitors (1 key scaffold H-bond) across the various simulations (Figures 8 and 9). Although L858R&T790 does not appear to affect the number of H-bonds at this key backbone position, the resistant mutant clearly results in abolishment of weaker H-bond interactions for all inhibitors at the site of the T790 mutation relative to L858R or wildtype alone (Figure 9c vs 9a,b). In addition, for erlotinib, the more significant H-bond at position C797 is also lost as a result of the double mutant (Figure 9c vs 9b, red line). Here, the loss at C797 is the result of only a slight shift in the binding pocket; otherwise, erlotinib appears well accommodated in the double mutant (Figure 10). The similarity in binding obtained here between L858R&T790M vs L858R suggests a steric clash mechanism of resistance is unlikely and consistent with recent crystallographic evidence from Yun et al. (*20*) (see discussion below). Although no clear reason was identified, the slight increases in H-bonding computed at position C797 for erlotinib with the single mutants in Figure 9 relative to wildtype may contribute to both the experimental (L858R) and predicted (G719S) increase in affinity for this compound (Table 2).

Energetic footprints representing van der Waals and Coulombic per-residue contributions were also plotted to quantify changes as a result of drug resistant EGFR relative to L858R. Focusing on the key residues, Figure 11 shows the contiguous region between Q787 and N808 (shaded) and residues for which favorable interactions are computed to be ca. > 1 kcal/mol. Again, the strong similarity in the general shape of the footprints, as well as the similar magnitudes in ΔE_{vdw} at specific positions (i.e., L718, A743, K745, T790, L792, L845 (Figure 11a) suggest that the computational results are sensitive enough to highlight both regions with conserved interaction as well as reflect differences which may prove useful in understanding affinity.

Consistent with the H-bond patterns described in Figures 8 and 9, in which ligands show high population of H-bonds between M793 and the central scaffolds, the most favorable ΔE_{coul} interactions for all ligands occur with residue M793 (Figure 11b). As before, the strongest interactions are computed for AEE788 (−11.5 kcal/mol, green line) versus gefitinib (−5.1 kcal/mol, blue line) or erlotinib (−3.9 kcal/mol, red line) which mirrors the fact that AEE788's scaffold makes two H-bonds versus one for the other inhibitors (Figure 8). Less populated, but "standard" H-bonds between T790 and AEE788, and C797 and erlotinib are also visible in the ΔE_{coul} footprints but as expected are weaker than those with M793 (ca. −1 to −1.5 kcal/mol). The more unique erlotinib (pi-type), or gefitinib (chlorine-type) interactions with T790 depicted in Figure 8 are not readily apparent in the ΔE_{coul} footprints but instead are presumably reflected in the favorable ΔE_{vdw} energies which occur at this position (Figure 11a vs 11b).

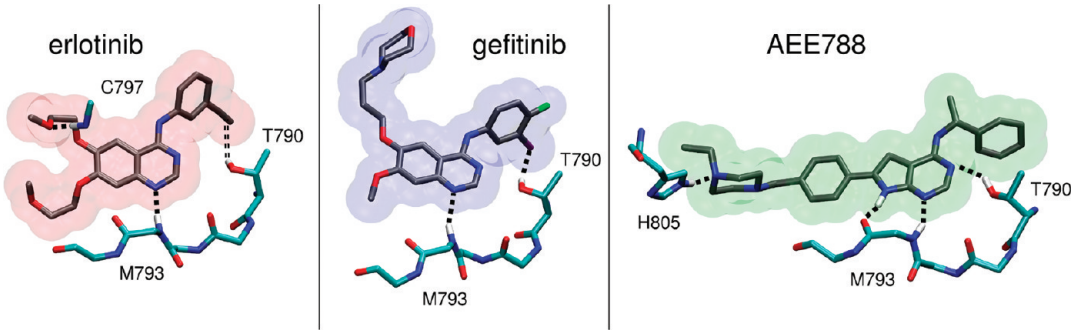


FIGURE 8: Primary H-bonding (dashed lines) interactions for inhibitors with EGFR.

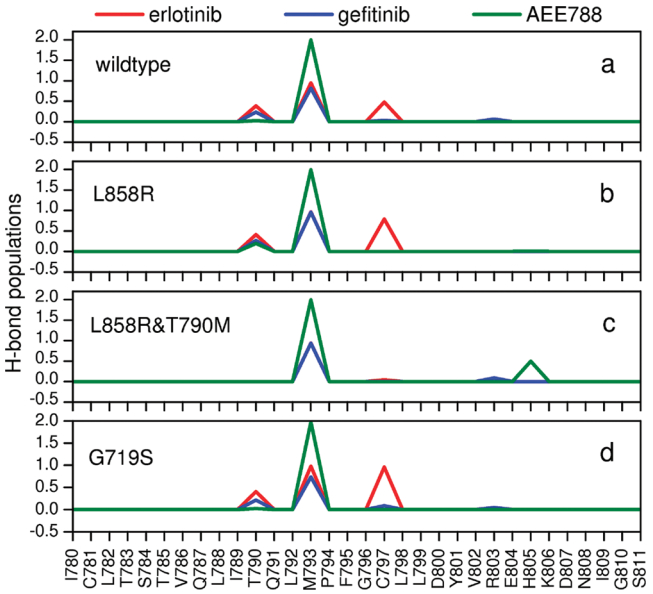


FIGURE 9: Comparison of per-residue H-bond footprints for erlotinib (red), gefitinib (blue), and AEE788 (green) with wildtype (panel a), L858R (panel b), L858R&T790M (panel c), and G719S (panel d) EGFR variants (N = 5000 MD snapshots).

Examination of difference footprints ($\Delta\Delta E_{vdw}$ and $\Delta\Delta E_{coul}$) computed from the L858R&T790M − L858R breakdowns show that erlotinib and AEE788 lose significant interactions, on a residue-by-residue basis, as a result of the deleterious mutation relative to L858R (Figure 11c, d). In contrast, and in agreement with the fact that gefitinib is experimentally the least affected by the resistance mutation, the $\Delta\Delta E_{coul}$ footprint is flatter, shows no overall reduction in total Coulombic energy (Table 2), and changes on a per-residue basis show negligible losses at all positions (Figure 11d, blue line). A prior study from our laboratory of neuraminidase inhibitors also revealed that the most robust compound had an overall flatter $\Delta\Delta E_{coul}$ and ΔH-bond profile (*41*). The most significant $\Delta\Delta E_{coul}$ energy losses (ca. > +1 kcal/mol) occur for erlotinib (Figure 11d, red line) at positions C797 and D800, and for AEE788 at positions T790M and D800 (Figure 11d, green line). Losses in $\Delta\Delta E_{coul}$ for erlotinib at position C797 are expected to be a result of the previously described H-bond disruption (Figures 8−10). For AEE788, the significant loss in $\Delta\Delta E_{coul}$ at T790M is likely due in part to disruption of the third H-bond with the pyrrolopyrimidine scaffold as a result of the resistance mutation (Figures 8 and 9). No straightforward structural explanation was found to explain reductions in $\Delta\Delta E_{coul}$ at position D800 for erlotinib or AEE788

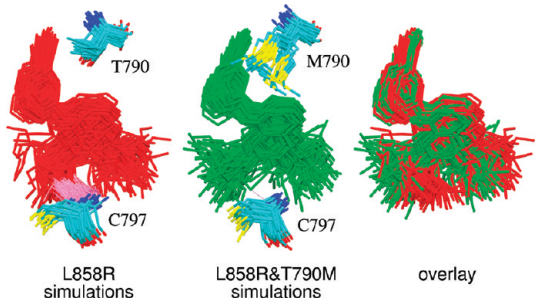


FIGURE 10: Comparison of erlotinib binding poses (N = 100) from EGFR simulations for L858R (left, red), L858R&T790M (middle, green), and overlaid (red vs green). Intermolecular H-bonds at position C797 shown in purple (N = 5000).

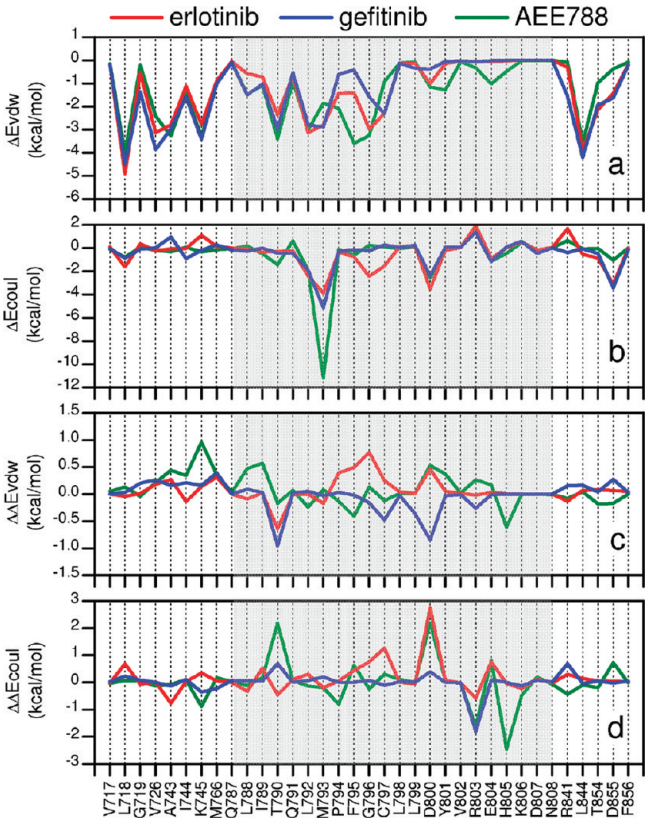


FIGURE 11: Per-residue footprints for inhibitors with EGFR for cancer causing (L858R, panels a,b) and drug resistance (L858R&T790M, panels c,d) variants from a reduced set of amino acids in the contiguous range Q787-N808 (shaded region) or for which any ligand shows ΔE > 1 kcal/mol.

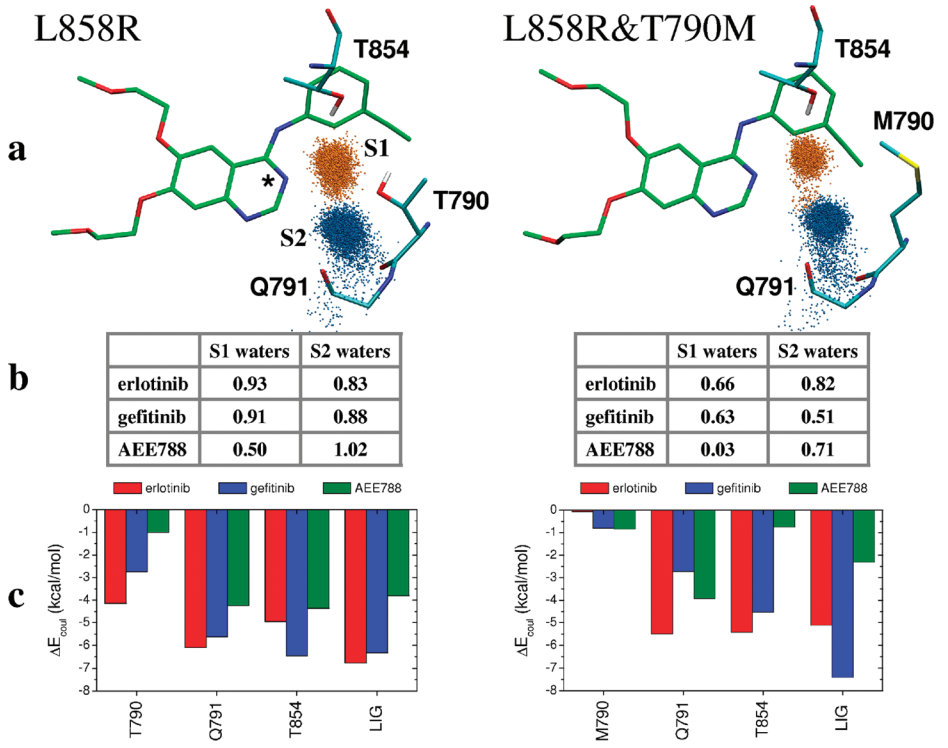


FIGURE 12: Water-mediate H-bonds for inhibitors with EGFR for L858R (left) and L858R&T790M (right). (a) Visually shows population of waters at site 1 (S1 orange) and site 2 (S2 blue) over all 5000 simulation frames for representative erlotinib simulations. Site waters defined if water hydrogens are within 2.5 Å of each ligand at N* (S1) or residue Q791 at O (S2). (b) Shows for all three inhibitors the average number (count/5000) of waters at S1 and S2. (c) Shows for all three inhibitors the average pairwise Coulombic interaction energies between the two waters closest to each ligand at N* with residues T790 (or M790), Q791, T854, and the ligands.

but as a charged residue this could be a long-range and nonspecific effect. For AEE788 with the double mutant, increases in $\Delta\Delta E_{coul}$ at position H805 (Figure 11c) are traceably to the formation of a new piperazine ring H-bond (Figures 8 and 9c). However, despite the fact that gains in interaction energy occur at this position, changes overall in $\Delta\Delta E_{coul}$ and $\Delta\Delta E_{vdw}$ for AEE788 are still unfavorable (Table 2).

Prior studies by Daub et al. (*57*) and Kobayashi et al. (*58*) hypothesized that a steric clash was the likely mechanism of drug resistance for T790M. And Liu et al. (*36*) reported MD simulation results of gefitinib with either T790M or L858R&T790M which led to "ligand escape from the binding pocket" which could also be consistent with a steric clash. However, a recent study reported by Yun et al. (*20*) suggests this is not a likely mechanism as a cocrystal structure of AEE788 with a T790M single mutant shows essentially the same binding pose as wildtype. Results from the present study similarly suggest that a threonine/methionine swap in the double mutant will not result in a steric clash given that wildtype, L858R, and L858R&T790M simulations show an overall consistent binding pose (Figure 10). In addition, an examination of the van der Waals differences footprints ($\Delta\Delta E_{vdw}$) shows that for all inhibitors a methionine at position 790 is energetically accommodated in the pocket and steric packing interactions localized to this position in fact become more favorable as a result of the double mutation (Figure 11c). Increased packing as a result of T790M is physically reasonable and occurs as a result of the hydrophilic to hydrophobic substitution. Although other van der Waals changes are less readily explained, the H805 increase with AEE788 coincides with the previously noted piperazine H-bond. Compellingly, erlotinib (positions F795, G796, D800) and AEE788 (positions K745, I789, D800) show significant losses in $\Delta\Delta E_{vdw}$ in contrast to gefitinib, which likely contributes to these compounds being more affected by the double mutations (Figure 11c).

Water-Mediated Interactions. Examination of the underlying explicit solvent TIP3P-MD trajectories, used subsequently for continuum-based free energy calculations, revealed water molecules which appear to be important for positioning of ligands in the binding pocket. High water occupancy is observed at two primary positions, termed site 1 (S1) and site 2 (S2), as shown in Figure 12a for erlotinib with L858R and the double mutant, which are representative. Figure 12b quantifies S1 and S2 populations for all six inhibitor simulations with averages = total count/5000 frames. Site waters were defined as present if a water hydrogen was within 2.5 Å of each ligand's relevant nitrogen acceptor (S1) or residue Q791 at O (S2). Importantly, the MD simulations reproduce the crystallographically observed water at S1 for all ligands (*14*, *21*). The waters at both sites are observed in the crystal structure of AEE788 with EGFR (2J6M) (*21*). For all ligands with L858R (Figure 12b left), waters are present 50−90% at S1 and > 80% at S2 which indicates these are long-lived significant interactions. As shown in Figure 12a, these waters are involved in a quadrifurcated H-bonding network involving the ligands with three nearby residues (T790, Q791, and T854), including the site of the known drug resistance mutation T790M. Notably, in all cases, occupancy at S1 and S2 is reduced as a result of L858R&T790 (Figure 12b right).

As an alternative metric, energy calculations reveal favorable Coulombic interactions between pocket waters and amino acids in the H-bond network including the ligands (Figure 12c left). Here, the two waters closest to each ligand at N* (Figure 12a left) were used define key pocket waters. Interestingly, the L858R&T790M mutant leads to changes in bridging water interactions with each ligand that roughly mirror trends in the

experimental FR data with erlotinib ($\Delta\Delta E_{coul}$ = +1.7 kcal/mol) and AEE788 ($\Delta\Delta E_{coul}$ = +1.5 kcal/mol) both being adversely affected compared to gefitinib ($\Delta\Delta E_{coul}$ = −1.1 kcal/mol). Favorable electrostatic interactions between these waters and residue 790 (Figure 12c right) are similarly reduced as a result of the double mutant, particularly for erlotinib (red bar), and thus expected to lead to weaker protein−ligand binding. Further, despite the fact that some water-mediated H-bonding with M790 is observed, an overall weaker network would be expected due to the fact that sulfur is a weaker H-bond acceptor than oxygen (*59*). Overall, the energetic description (Figure 12c) is consistent with the reduced population counts (Figure 12b) suggesting weaker interactions in the drug resistant mutant.

H-bonding between quinazoline-based inhibitors and binding site waters were previously predicted by Wissner et al. (*33*) and Hou et al. (*34*), although interestingly the two studies came to different conclusions as to whether residue T854 or T790 was involved. Here, calculations indicate that both T854 and T790 residues make significant water-mediated ligand interactions. Stamos et al. (*14*) noted the T790 bridging water in the erlotinib-EGFR crystal structure but suggested it was not significant, citing data reported by Rewcastle et al. (*60*) in which only a minor effect on affinity was seen for related ligands where the H-bond acceptor was substituted for carbon. However, examination of the original activities (see Rewcastle et al., Table 1, compounds 15 vs 20) show > 5000 fold loss between compounds that differ only by a nitrogen at the T790 acceptor position which suggests the water is in fact important (*60*). And, a recent docking study by Cavasotto et al. (*35*) notes that inclusion of this bridging water was necessary to correctly reproduce the binding pose of the EGFR inhibitor AG1478.

In conjunction with their proposed steric clash mechanism, Kobayashi et al. (*58*) also hypothesized that disruption of water-mediated binding would be a factor in resistance. For the water-mediated interactions at Q791 (Figure 12), the H-bonds primarily involve the backbone carbonyl oxygen; thus, any alteration of side chains at this site would be expected to be less detrimental, particularly since there is little direct van der Waals contact or favorable Coulombic interactions with the ligand at Q791. However, the simulation results strongly suggest that a mutation at position T854 would disrupt the quadrifurcated network and, in a manner analogous to T790, disrupt water-mediated ligand binding. This hypothesis is consistent with results recently reported by Bean et al. (*61*) in which a novel T854A resistance mutation was identified from a patient with reduced affinity for erlotinib. A combination of mutations involving T790 and T854, if biologically viable, would likely lead to further disruption of the H-bond network involving inhibitors and an increase in unfavorable fold resistance. While our current studies cannot rule out the recent hypothesis by Yun et al. (*20*) that T790M resistance is caused primarily by increased affinity for ATP, based on the present simulations, it is reasonable to propose that disruption of water-mediated H-bond networks involving the inhibitors (Figure 12) is a contributing factor. Additionally, given the fact that our calculations yield quantitative energetic agreement with experiment, yet involve only inhibitors and EGFR (and not ATP), it strongly suggests that differences in affinity for ATP are not the sole cause of experimentally observed drug resistance. Additional studies are needed to more fully address this issue.

An examination of the network shown in Figure 12 indicates the possibility of designing alternative H-bonding involving residues T854, T790, and Q791. EGFR inhibitors based on a 4,6-dianilinopyrimidine scaffold have been reported (*62*) which are proposed to make direct H-bonds with both M793 and T790. However, as expected, the T790M mutant showed resistance against a representative compound in the series presumably due to the loss of a direct H-bond between the pyrimidine N3 and the OH at position 790. Wissner et al. (*33*) has reported an inhibitor in which the bridging nitrogen atom was replaced by a cyano group and proposed to displace the site 1 water. However, the cyano compound was also proposed to make a direct H-bond with T854; thus, the recently reported T854A mutation would likely lead to a loss in binding. An alternative strategy to address resistance, provided that sufficient specificity could be achieved, would be the design of inhibitors with additional protein backbone H-bonds (direct or water mediated). In any event, due to the entropically favorable process of displacing bound waters (*63*), analogues which replace the water-mediated interactions seen here may show enhanced affinity. Alternative binding patterns are likely to result in unique resistance profiles which may prove useful.

CONCLUSION

In this study, all-atom explicit solvent molecular dynamics followed by free energy calculations were employed to compute fold resistance energies for three ATP-competitive inhibitors (erlotinib, gefitinib, and AEE788) with epidermal growth factor receptor (EGFR) for wildtype, and L858R, G719S, and L858R&T790M mutants. The primary purpose of this study was development of robust quantitative computational models to compute EGFR-ligand binding, characterize how variation in structural and energetic results correlate with variation in reported experimental activities, and determine origins of drug resistance. System stability and overall convergence of results were carefully monitored through comparisons with crystallographic structures (Figure 4), and by plotting instantaneous and running block averages for free energies of binding and root-mean-square deviations (Figure 3). Fluctuations in energy and structure show the simulations are well-behaved, comparable with other studies from our laboratory (*40*−*42*), and low standard errors of the mean (Tables 2 and 3) indicate the results are reasonably converged.

Notably, computed fold resistance energies, which represent a ratio of activities and are obtained from the difference in results of two independent MD simulations, show excellent agreement with available experimental data (r^2 = 0.84). Importantly, the magnitudes of the experimental and theoretical FR results are similar (Table 2, Figure 5). For all inhibitors, the simulations correctly predict that affinity for EGFR will increase as a result of the cancer causing L858R mutation relative to wildtype and decrease as a result of a drug resistant double mutant (L858R&T790M) relative to L858R (Table 2 columns E vs F). Affinity predictions for gefitinib with a second cancer causing mutation at position G719S also yield the correct experimental trend. The sole outlier in the study is for AEE788 with G719S in which the computational results incorrectly predict the mutation to be slightly favorable (Table 2).

Decomposition of the contributing components to $\Delta\Delta G_{FR}$, and the underlying absolute ΔG_b values used to compute FR, reveal modest gains in favorable van der Waals and Coulombic energies for all three inhibitors as a result of the cancer causing mutation L858R and large losses for erlotinib and AEE788 for the drug

resistance double mutant L858R&T790M (Table 2). Losses for gefitinib appear to be a result of increased desolvation penalties. Values for ΔE_{vdw} are computed to be more favorable than other ΔG_b calcd terms (Table 3), which suggest that steric packing is the dominant driving force for association. In general, van der Waals interactions are stronger for gefitinib relative to other compounds and changes in ΔE_{vdw} track especially well with ΔG_b exptl for gefitinib and AEE788 (Figure 7). Relative $\Delta\Delta E_{vdw}$ energies are modestly correlated with $\Delta\Delta G_{FR}$ across the series (Figure 6b).

Despite the obvious importance of van der Waals terms, the simulation results indicate that electrostatic interactions are critical for specificity and correct positioning of ligands in the ATP binding pocket (Figures 8, 9, and 11). Examination of per-residue H-bonding, and Coulombic energy, reveal changes at key amino acids which are important for understanding origins of fold resistance (Figures 9−11). Two highly populated H-bonds for AEE788, and one for erlotinib and gefitinib, are observed between inhibitors and the EGFR backbone at position M793 (Figures 8 and 9). Coulombic energy footprints (Figures 11) mirror the H-bond trends with M793 showing stronger interaction energies formed with AEE788 (−11.5 kcal/mol) versus gefitinib (−5.1 kcal/mol) or erlotinib (−3.9 kcal/mol). The resistance mutation L858R&T790 does not change interactions localized to M793; however, all inhibitors lose a less populated H-bond at the site of the T790 mutation (Figure 9 panels b vs c). Losses at 790 are traced to unique H-bonds (Figures 8 and 9) involving the acetylene group of erlotinib (pi-type) and a chlorine atom in the case of gefitinib (chlorine-type). For erlotinib, the loss of an additional H-bond at position C797 leads to an overall reduction (−1.18 H-bonds), which likely contributes to the larger FR energy compared with the other inhibitors (gefininb −0.19 H-bonds, AEE788 +0.29 H-bonds).

The simulations additionally reveal a significant network of water-mediated H-bonds involving a spatially equivalent nitrogen atom on each inhibitor, residues T854, T790, Q791, and two bridging waters which become disrupted as a result of the L858R&T790M drug resistance mutation (Figure 12a,b). The bridging waters interact favorably with residues in the binding pocket and the double mutation leads to reduced Coulombic energies, especially for erlotinib and gefitinib (Figure 12c), and reduced overall occupancy (Figure 12b). The calculations suggest that resistance likely involves changes in water-mediated H-bonds, in contrast to prior reports, which hypothesize that EGFR resistance is primarily a function of either a steric clash (*57*, *58*) involving methionine 790 or due to increased affinity for the native substrate ATP (*20*). In agreement with recent crystallographic evidence (*20*), per-residue footprint calculations (Figure 11c) and structural analysis (Figure 10) reveal favorable van der Waals energies with T790M, which indicates a steric clash mechanism of resistance is unlikely. Finally, while increased affinity for ATP (*20*) may be a contributor to resistance, the present results suggest that disruption of favorable interactions, including changes in H-bonding, are likely to be as important and thus should be considered when designing next-generation compounds.

The growing problem of drug resistance, arising from clinical use of EGFR molecular targeted therapeutics, highlights the need for continued studies to elucidate how binding affinity is modulated by mutations and how ligands could be modified to circumvent deleterious changes. The present study has participated in these aims through prediction of fold resistance energies for inhibitors of EGFR which show quantitative agreement with experiment thereby providing a framework to probe origins of resistance. The simulations correctly predict the effects of the L858R cancer causing mutation and the L858R&T790 drug resistance mutant for three inhibitors. Residue-based structural and energetic analysis was used to identify how key side chains are involved in binding, how water molecules mediate affinity through an intricate network of H-bonding, and how interactions change as a result of the mutations. Similar to the growing arsenal of antivirals required to effectively combat HIV, design of multiple kinase inhibitors is likely to be an important long-term strategy to address issues of drug resistance.

ACKNOWLEDGMENT

Gratitude is expressed to Todd Miller for helpful discussions and to the Computational Science Center at Brookhaven National Laboratory for computational support.

REFERENCES

1. Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., Murray, T., and Thun, M. J. (2008) Cancer statistics. *CA Cancer J. Clin. 58*, 71–96.
2. Travis, W. D., Travis, L. B., and Devesa, S. S. (1995) Lung cancer. *Cancer 75*, 191–202.
3. Hirsch, F. R., Varella-Garcia, M., Bunn, P. A., Jr., Di Maria, M. V., Veve, R., Bremmes, R. M., Baron, A. E., Zeng, C., and Franklin, W. A. (2003) Epidermal growth factor receptor in non-small-cell lung carcinomas: correlation between gene copy number and protein expression and impact on prognosis. *J. Clin. Oncol. 21*, 3798–3807.
4. Hynes, N. E., and Lane, H. A. (2005) ERBB receptors and cancer: the complexity of targeted inhibitors. *Nat. Rev. Cancer 5*, 341–354.
5. Mendelsohn, J., and Baselga, J. (2003) Status of epidermal growth factor receptor antagonists in the biology and treatment of cancer. *J. Clin. Oncol. 21*, 2787–2799.
6. Rocha-Lima, C. M., Soares, H. P., Raez, L. E., and Singal, R. (2007) EGFR targeting of solid tumors. *Cancer Control 14*, 295–304.
7. Drugs@FDA website. http://www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm (accessed Jan 22, 2009)
8. Riese, D. J., 2nd, Gallo, R. M., and Settleman, J. (2007) Mutational activation of ErbB family receptor tyrosine kinases: insights into mechanisms of signal transduction and tumorigenesis. *Bioessays 29*, 558–565.
9. Zhang, X., Gureasko, J., Shen, K., Cole, P. A., and Kuriyan, J. (2006) An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell 125*, 1137–1149.
10. Yarden, Y., and Sliwkowski, M. X. (2001) Untangling the ErbB signalling network. *Nat Rev Mol. Cell Biol. 2*, 127–137.
11. Hubbard, S. R., and Miller, W. T. (2007) Receptor tyrosine kinases: mechanisms of activation and signaling. *Curr. Opin. Cell Biol. 19*, 117–123.
12. Scaltriti, M., and Baselga, J. (2006) The epidermal growth factor receptor pathway: a model for targeted therapy. *Clin. Cancer Res. 12*, 5268–5272.
13. Wood, E. R., Truesdale, A. T., McDonald, O. B., Yuan, D., Hassell, A., Dickerson, S. H., Ellis, B., Pennisi, C., Horne, E., Lackey, K., Alligood, K. J., Rusnak, D. W., Gilmer, T. M., and Shewchuk, L. (2004) A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells. *Cancer Res. 64*, 6652–6659.
14. Stamos, J., Sliwkowski, M. X., and 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.
15. Fry, D. W. (2003) Mechanism of action of erbB tyrosine kinase inhibitors. *Exp. Cell Res. 284*, 131–139.
16. Traxler, P., Allegrini, P. R., Brandt, R., Brueggen, J., Cozens, R., Fabbro, D., Grosios, K., Lane, H. A., McSheehy, P., Mestan, J., Meyer, T., Tang, C., Wartmann, M., Wood, J., and Caravatti, G. (2004) AEE788: a dual family epidermal growth factor receptor/ErbB2 and vascular endothelial growth factor receptor tyrosine kinase inhibitor with antitumor and antiangiogenic activity. *Cancer Res. 64*, 4931–4941.
17. Karaman, M. W., Herrgard, S., Treiber, D. K., Gallant, P., Atteridge, C. E., Campbell, B. T., Chan, K. W., Ciceri, P., Davis, M. I.,

Edeen, P. T., Faraoni, R., Floyd, M., Hunt, J. P., Lockhart, D. J., Milanov, Z. V., Morrison, M. J., Pallares, G., Patel, H. K., Pritchard, S., Wodicka, L. M., and Zarrinkar, P. P. (2008) A quantitative analysis of kinase inhibitor selectivity. *Nat. Biotechnol. 26*, 127–132.
18. Carey, K. D., Garton, A. J., Romero, M. S., Kahler, J., Thomson, S., Ross, S., Park, F., Haley, J. D., Gibson, N., and Sliwkowski, M. X. (2006) Kinetic analysis of epidermal growth factor receptor somatic mutant proteins shows increased sensitivity to the epidermal growth factor receptor tyrosine kinase inhibitor, erlotinib. *Cancer Res. 66*, 8163–8171.
19. Ji, H., Zhao, X., Yuza, Y., Shimamura, T., Li, D., Protopopov, A., Jung, B. L., McNamara, K., Xia, H., Glatt, K. A., Thomas, R. K., Sasaki, H., Horner, J. W., Eck, M., Mitchell, A., Sun, Y., Al-Hashem, R., Bronson, R. T., Rabindran, S. K., Discafani, C. M., Maher, E., Shapiro, G. I., Meyerson, M., and Wong, K. K. (2006) Epidermal growth factor receptor variant III mutations in lung tumorigenesis and sensitivity to tyrosine kinase inhibitors. *Proc. Natl. Acad. Sci. U. S. A. 103*, 7817–7822.
20. Yun, C. H., Mengwasser, K. E., Toms, A. V., Woo, M. S., Greulich, H., Wong, K. K., Meyerson, M., and Eck, M. J. (2008) The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc. Natl. Acad. Sci. U. S. A. 105*, 2070–2075.
21. Yun, C. H., Boggon, T. J., Li, Y., Woo, M. S., Greulich, H., Meyerson, M., and Eck, M. J. (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.
22. Wikstrand, C. J., McLendon, R. E., Friedman, A. H., and Bigner, D. D. (1997) Cell surface localization and density of the tumor-associated variant of the epidermal growth factor receptor, EGFRvIII. *Cancer Res. 57*, 4130–4140.
23. Lynch, T. J., Bell, D. W., Sordella, R., Gurubhagavatula, S., Okimoto, R. A., Brannigan, B. W., Harris, P. L., Haserlat, S. M., Supko, J. G., Haluska, F. G., Louis, D. N., Christiani, D. C., Settleman, J., and Haber, D. A. (2004) Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med. 350*, 2129–2139.
24. Paez, J. G., Janne, P. A., Lee, J. C., Tracy, S., Greulich, H., Gabriel, S., Herman, P., Kaye, F. J., Lindeman, N., Boggon, T. J., Naoki, K., Sasaki, H., Fujii, Y., Eck, M. J., Sellers, W. R., Johnson, B. E., and Meyerson, M. (2004) EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science 304*, 1497–1500.
25. Pao, W., Miller, V., Zakowski, M., Doherty, J., Politi, K., Sarkaria, I., Singh, B., Heelan, R., Rusch, V., Fulton, L., Mardis, E., Kupfer, D., Wilson, R., Kris, M., and Varmus, H. (2004) EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc. Natl. Acad. Sci. U. S. A. 101*, 13306–13311.
26. Sordella, R., Bell, D. W., Haber, D. A., and Settleman, J. (2004) Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. *Science 305*, 1163–1167.
27. Riely, G. J., Pao, W., Pham, D., Li, A. R., Rizvi, N., Venkatraman, E. S., Zakowski, M. F., Kris, M. G., Ladanyi, M., and Miller, V. A. (2006) Clinical course of patients with non-small cell lung cancer and epidermal growth factor receptor Exon 19 and Exon 21 mutations treated with gefitinib or erlotinib. *Clin. Cancer Res. 12*, 839–844.
28. Sequist, L. V., Bell, D. W., Lynch, T. J., and Haber, D. A. (2007) Molecular predictors of response to epidermal growth factor receptor antagonists in non-small-cell lung cancer. *J. Clin. Oncol. 25*, 587–595.
29. Pao, W., and Miller, V. A. (2005) Epidermal growth factor receptor mutations, small-molecule kinase inhibitors, and non-small-cell lung cancer: current knowledge and future directions. *J. Clin. Oncol. 23*, 2556–2568.
30. Pao, W., Miller, V. A., Politi, K. A., Riely, G. J., Somwar, R., Zakowski, M. F., Kris, M. G., and Varmus, H. (2005) Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. *PLoS Med. 2*, e73.
31. Branford, S., Rudzki, Z., Walsh, S., Grigg, A., Arthur, C., Taylor, K., Herrmann, R., Lynch, K. P., and Hughes, T. P. (2002) High frequency of point mutations clustered within the adenosine triphosphate-binding region of BCR/ABL in patients with chronic myeloid leukemia or Ph-positive acute lymphoblastic leukemia who develop imatinib (STI571) resistance. *Blood 99*, 3472–3475.
32. Gorre, M. E., Mohammed, M., Ellwood, K., Hsu, N., Paquette, R., Rao, P. N., and Sawyers, C. L. (2001) Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. *Science 293*, 876–880.
33. Wissner, A., Berger, D. M., Boschelli, D. H., Floyd, M. B., Jr., Greenberger, L. M., Gruber, B. C., Johnson, B. D., Mamuya, N., Nilakantan, R., Reich, M. F., Shen, R., Tsou, H. R., Upeslacis, E., Wang, Y. F., Wu, B., Ye, F., and Zhang, N. (2000) 4-Anilino-6,7-dialkoxyquinoline-3-carbonitrile inhibitors of epidermal growth factor receptor kinase and their bioisosteric relationship to the 4-anilino-6,7-dialkoxyquinazoline inhibitors. *J. Med. Chem. 43*, 3244–3256.
34. Hou, T., Zhu, L., Chen, L., and Xu, X. (2003) Mapping the binding site of a large set of quinazoline type EGF-R inhibitors using molecular field analyses and molecular docking studies. *J. Chem. Inf. Comput. Sci. 43*, 273–287.
35. Cavasotto, C. N., Ortiz, M. A., Abagyan, R. A., and Piedrafita, F. J. (2006) In silico identification of novel EGFR inhibitors with anti-proliferative activity against cancer cells. *Bioorg. Med. Chem. Lett. 16*, 1969–1974.
36. Liu, B., Bernard, B., and Wu, J. H. (2006) Impact of EGFR point mutations on the sensitivity to gefitinib: insights from comparative structural analyses and molecular dynamics simulations. *Proteins 65*, 331–346.
37. Liu, Y., Purvis, J., Shih, A., Weinstein, J., Agrawal, N., and Radhakrishnan, R. (2007) A multiscale computational approach to dissect early events in the Erb family receptor mediated activation, differential signaling, and relevance to oncogenic transformations. *Ann. Biomed. Eng. 35*, 1012–1025.
38. Massova, I., and Kollman, P. A. (2000) Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Perspect. Drug Discovery Des. 18*, 113–135.
39. Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D. A., and Cheatham, T. E.3rd. (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res. 33*, 889–897.
40. Strockbine, B., and Rizzo, R. C. (2007) Binding of antifusion peptides with HIVgp41 from molecular dynamics simulations: quantitative correlation with experiment. *Proteins: Struct. Funct. Bioinformatics 67*, 630–642.
41. Chachra, R., and Rizzo, R. C. (2008) Origins of resistance conferred by the R292K neuraminidase mutation via molecular dynamics and free energy calculations. *J. Chem. Theor. Comput. 4*, 1526–1540.
42. Carrascal, N., and Rizzo, R. C. (2009) Calculation of binding free energies for non-zinc chelating pyrimidine dicarboxamide inhibitors with MMP-13. *Bioorg. Med. Chem. Lett. 19*, 47–50.
43. Still, W. C., Tempczyk, A., Hawley, R. C., and Hendrickson, T. (1990) Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc. 112*, 6127–6129.
44. (2007) MOE, Chemical Computing Group, Montreal, Canada.
45. Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Wang, B., Pearlman, D. A., Crowley, M., Brozell, S., Tsui, V., Gohlke, H., Mongan, J., Hornak, V., Cui, G., Beroza, P., Schafmeister, C., Caldwell, J. W., Ross, W. S., Kollman, P. A. (2004) AMBER 8, University of California, San Francisco.
46. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins 65*, 712–725.
47. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926–935.
48. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) Development and testing of a general amber force field. *J. Comput. Chem. 25*, 1157–1174.
49. Breneman, C. M., and Wiberg, K. B. (1990) Determining atom-centered monopoles from molecular electrostatic potentials - the need for high sampling density in formamide conformational-analysis. *J. Comput. Chem. 11*, 361–373.
50. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Zakrzewski, V. G., Montgomery, J. A., Jr., Stratmann, R. E., Burant, J. C., Dapprich, S., Millam, J. M., Daniels, A. D., Kudin, K. N., Strain, M. C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Petersson, G. A., Ayala, P. Y., Cui, Q., Morokuma, K., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Cioslowski, J., Ortiz, J. V., Stefanov, B. B., Liu, G., Liashenko, A., PisVkorz, P., Komaromi, I., Gomperts, R., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Gonzalez, C., Challacombe, M., Gill, P. M. W., Johnson, B. G., Chen, W., Wong, M. W., Andres, J. L., Head-Gordon, M.,

Replogle, E. S., Pople, J. A. (1998) Gaussian 98, revision A.9 , Gaussian Inc., Pittsburgh PA.
51. Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comp. Phys. 23*, 327–341.
52. Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N [center-dot] log(N) method for Ewald sums in large systems. *J. Chem. Phys. 98*, 10089–10092.
53. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys. 81*, 3684–3690.
54. Onufriev, A., Bashford, D., and Case, D. A. (2004) Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins 55*, 383–394.
55. Sitkoff, D., Sharp, K. A., and Honig, B. (1994) Accurate calculation of hydration free-energies using macroscopic solvent models. *J. Phys. Chem. 98*, 1978–1988.
56. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem. 26*, 1781–1802.
57. Daub, H., Specht, K., and Ullrich, A. (2004) Strategies to overcome resistance to targeted protein kinase inhibitors. *Nat. Rev. Drug Discovery 3*, 1001–1010.
58. Kobayashi, S., Boggon, T. J., Dayaram, T., Janne, P. A., Kocher, O., Meyerson, M., Johnson, B. E., Eck, M. J., Tenen, D. G., and Halmos, B. (2005) EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med. 352*, 786–792.
59. Rablen, P. R., Lockman, J. W., and Jorgensen, W. L. (1998) Ab initio study of hydrogen-bonded complexes of small organic molecules with water. *J. Phys. Chem. A 102*, 3782–3797.
60. Rewcastle, G. W., Denny, W. A., Bridges, A. J., Zhou, H., Cody, D. R., McMichael, A., and Fry, D. W. (1995) Tyrosine kinase inhibitors. 5. Synthesis and structure-activity relationships for 4-[(phenylmethyl)amino]- and 4-(phenylamino)quinazolines as potent adenosine 5′-triphosphate binding site inhibitors of the tyrosine kinase domain of the epidermal growth factor receptor. *J. Med. Chem. 38*, 3482–3487.
61. Bean, J., Riely, G. J., Balak, M., Marks, J. L., Ladanyi, M., Miller, V. A., and Pao, W. (2008) Acquired resistance to epidermal growth factor receptor kinase inhibitors associated with a novel T854A mutation in a patient with EGFR-mutant lung adenocarcinoma. *Clin. Cancer Res. 14*, 7519–7525.
62. Zhang, Q., Liu, Y., Gao, F., Ding, Q., Cho, C., Hur, W., Jin, Y., Uno, T., Joazeiro, C. A., and Gray, N. (2006) Discovery of EGFR selective 4,6-disubstituted pyrimidines from a combinatorial kinase-directed heterocycle library. *J. Am. Chem. Soc. 128*, 2182–2183.
63. Dunitz, J. D. (1994) The entropic cost of bound water in crystals and biomolecules. *Science 264*, 670.