BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

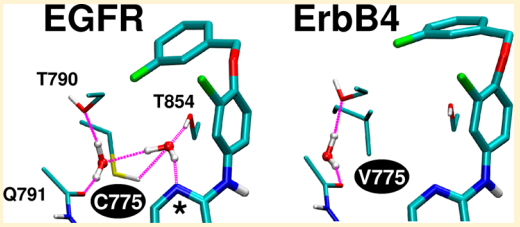
pubs.acs.org/biochemistry

A Water-Based Mechanism of Specificity and Resistance for Lapatinib with ErbB Family Kinases

Yulin Huang[†,‡] and Robert C. Rizzo*[,‡,§]

[†]Graduate Program in Biochemistry and Structural Biology, [‡]Department of Applied Mathematics and Statistics, and [§]Institute of Chemical Biology and Drug Discovery, Stony Brook University, Stony Brook, New York 11794, United States

ABSTRACT: The dual kinase inhibitor lapatinib has a high affinity for EGFR and HER2 but a weak affinity for ErbB4, although the factors driving specificity for these highly homologous members of the ErbB family of receptor tyrosine kinases are not well understood. In this report, homology modeling, molecular dynamics simulations, and free energy calculations are employed with the goal of uncovering the energetic and structural molecular basis of lapatinib specificity and resistance. The results reveal a distinct network of three binding site water molecules that yield strikingly similar hydration patterns for EGFR and HER2 in contrast to that of ErbB4, which shows a different pattern with a reduced occupancy at one of the positions. The primary cause was traced to a single amino acid change in the binding site (EGFR position 775), involving a swap from C or S (EGFR and HER2) to V (ErbB4), for which the side chain is bulkier, is hydrophobic, and lacks the ability to form a H-bond with water. Notably, excellent quantitative agreement with experimental activities is obtained across the series (EGFR > HER2 > ErbB4) when key waters are included in the calculations. Quantitatively, Coulombic interactions and H-bond counts between network waters and species involved in the network are less favorable in ErbB4 by ∼40% relative to those in EGFR or HER2. Additional simulations with clinically relevant EGFR (C775F, T854A, and T790M) and HER2 (T790I) mutants demonstrate that resistance can also be understood in terms of changes that occur in the binding site water network. Overall, the results of this study have yielded a physically reasonable water-based mechanism for describing lapatinib specificity and resistance.



Excluding skin cancer, breast cancer is the most common type of cancer for women in the United States, and one in eight women is expected to develop breast cancer.[1] Members of the highly homologous ErbB family of receptor tyrosine kinases (EGFR, HER2, HER3, and ErbB4) play key roles in breast cancer,[2,3] and EGFR[4] and HER2[5] are important prognostic markers. The design of agents targeting EGFR or HER2 alone, or in combination, has been a major therapeutic focus.[2,3] Successful examples include the HER2 neutralizing antibody trastuzumab (Herceptin) and the small molecule inhibitor lapatinib (Tykerb) (Table 1), which targets EGFR and HER2.[6] Unfortunately, less than one-third of patients with HER2 amplification benefit from trastuzumab.[3] Importantly, lapatinib has shown activity against trastuzumab-treated breast cancer cell lines[7] and in patients previously treated with trastuzumab.[8] However, acquired resistance to lapatinib involving alternative signaling pathways[9−13] or potentially deleterious point mutations[14,15] will likely hamper long-term clinical utility. Uncovering mechanisms associated with sensitivity and resistance to ErbB family inhibitors is an important long-term therapeutic goal. The focus of this work is an improved understanding of the atomic- and molecular-level details driving lapatinib dual specificity for EGFR and HER2 (strong) versus ErbB4 (weak)[16,17] (Table 1) and what leads to a loss of binding with specific point mutations.[14,15]

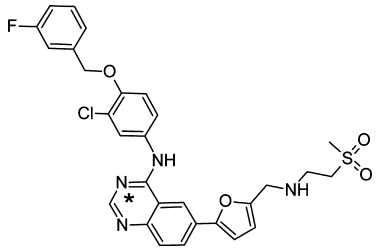
Members of the ErbB family share a similar overall structural architecture comprising (i) an extracellular ligand binding domain, (ii) a transmembrane domain, (iii) an intracellular juxtamembrane domain, (iv) an intracellular tyrosine kinase domain, and (v) a C-terminal regulatory region where phosphorylation occurs.[18] Interestingly, no exogenous ligands that bind to HER2 have been identified, and HER3 lacks kinase activity.[2,3] Figure 1 highlights the similarities in sequence (EGFR numbering), including the lapatinib binding site, across the kinase domains of the proteins studied in this report (EGFR, HER2, and ErbB4). Binding of a ligand to the extracellular domain of EGFR, HER3, or ErbB4 can promote homo- or heterodimerization with another ErbB family member, resulting in activation of the intracellular kinase domain.[2,3] Activation (inactive → active form) is mediated by a structural reorganization that primarily involves conformational changes in the C-helix and activation loop regions.[18] Active homodimeric−heterodimeric complexes can, in turn, cause activation of a series of downstream signaling cascades. Aberrant signal transduction can promote cell proliferation, promote differentiation, and eventually lead to malignant transformation.[2,3]

From a therapeutic standpoint, small molecule kinase inhibitors that preferentially target either the inactive [lapatanib (Table 1)][17] or the active (erlotinib[19] and gefitinib[20]) kinase forms have been developed. These compounds bind competitively at the ATP-binding site, blocking phosphorylation

Received: November 2, 2011
Revised: February 2, 2012
Published: February 21, 2012

   dx.doi.org/10.1021/bi2016553 | *Biochemistry* 2012, 51, 2390−2406

Table 1. Experimental Activities for Lapatinib with ErbB Family Members



	EGFR	HER2	ErbB4
IC_{50}[a] (nM)	10.8 ± 0.53	9.2 ± 0.75	367.0 ± 4.2
K_i[b] (nM)	3.0 ± 0.2	13 ± 1	347.0 ± 16
ΔG_{bind}[c] (kcal/mol)	−11.62	−10.75	−8.81

[a]From ref 16. [b]EGFR and HER2 (K_i) and Erb4 (calculated K_i) values from ref 17. [c]$\Delta G_{bind}(exptl) \approx RT \ln(K_i)$ at 298.15 K. Note that N* indicates a key ligand atom involved in water-mediated H-bonding (see the text for a discussion).

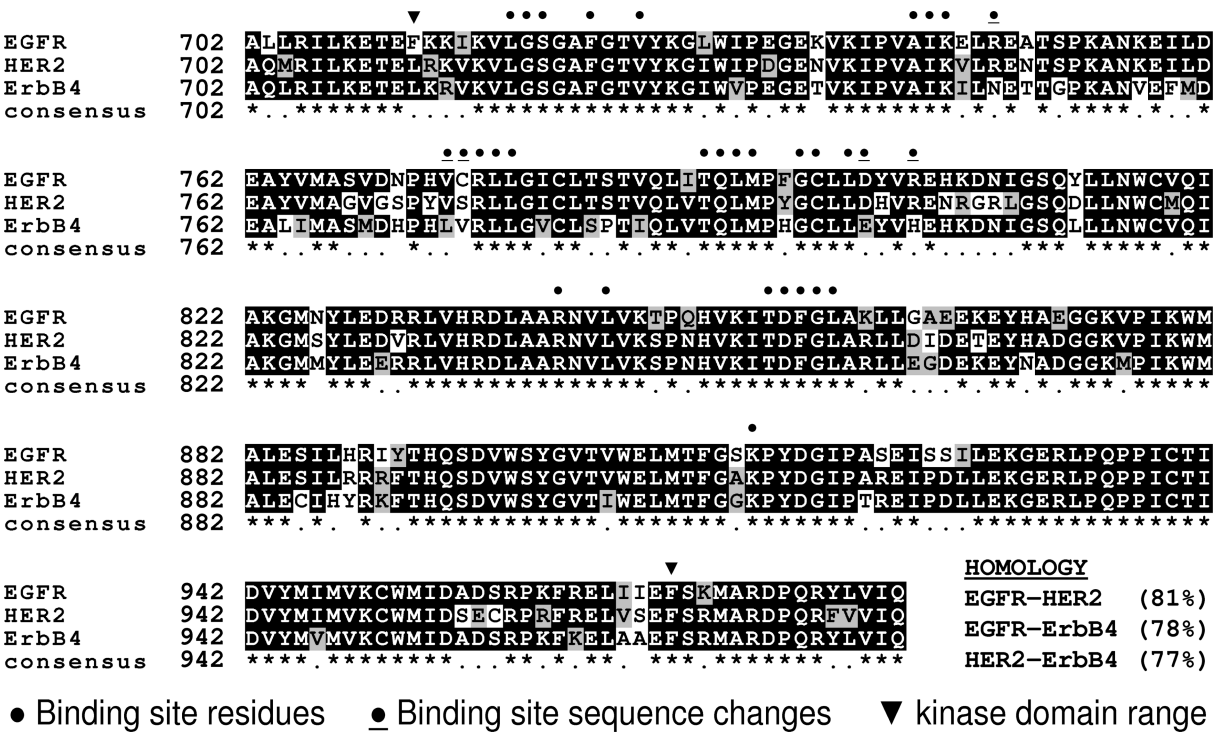


Figure 1. Sequence comparisons among the kinase domains of EGFR, HER2, and ErbB4. The numbering is that of EGFR (PDB entry 1XKK).

and subsequent downstream signaling. Lapatinib is approved for the treatment of HER2-overexpressing breast cancer, while erlotinib and gefitinib are approved for the treatment of non-small cell lung cancer.[6] Unlike erlotinib and gefitinib that primarily target EGFR, lapatinib is a dual kinase inhibitor[21] targeting both EGFR and HER2. Interestingly, despite the high degree of sequence homology (Figure 1), lapatinib only weakly inhibits ErbB4 (Table 1). This fact may be clinically relevant as it has been reported that ErbB4 may lead to a potentially beneficial antiproliferative response in human breast cancer cells,[22] and overexpression of ErbB4 correlates with reduced recurrence of breast cancer.[23] Prior studies, including crystallographic reports of lapatinib in complex with inactive form EGFR[17] or ErbB4,[24] have not offered clear reasons why lapatinib preferentially binds EGFR and HER2 but not ErbB4.

There is growing concern that drug resistance, as has been observed with the related kinase inhibitors erlotinib and gefitinib, may occur upon long-term treatment with lapatinib.[14,15,25] Known EGFR mutations associated with acquired resistance to erlotinib or gefitnib include T790M and T854A.[26−28] Here, the so-called gate keeper mutation at position 790 (EGFR or HER2) is similar to the T315I mutation arising from treatment with imatinib (Gleevec) in the BCR-ABL kinase system.[29] Avizienyte et al.[14] and Trowe et al.[15] have reported that several ErbB family mutations, including those in EGFR (C775F, T854A, and T790M) and HER2 (T790I), also negatively affect lapatinib. Proposed resistance mechanisms include (i) a steric clash for EGFR C775F,[14] EGFR T790M,[25] or HER2 T790I,[15] (ii) increased affinity for ATP for EGFR T790M,[30] and (iii) a loss of contact for EGFR T854A.[28] An improved understanding of what contributes to lapatinib binding at the structural level will be important in explaining the impact of observed mutations as well as the design of next-generation EGFR and HER2 inhibitors with improved resistance profiles.

dx.doi.org/10.1021/bi2016553 | Biochemistry 2012, 51, 2390−2406

Examples of prior computational work addressing ligand binding in ErbB systems have employed homology and molecular modeling,[31−33] comparative molecular field analysis,[34,35] docking/virtual screening,[34−36] molecular dynamics,[31,32,34,35,37−39] and Monte Carlo simulations.[40] Surprisingly, few studies have addressed binding across multiple members of the ErbB family. Exceptions include the work of Kamath et al.,[32] who explored ATP selectivity between EGFR and HER2 using molecular dynamics, and Scaltriti et al.,[33] who explored binding of lapatinib with EGFR, HER2, and HER3 using homology modeling in combination with energy minimization. Prior calculations addressing drug resistance include reports by Liu et al.[37] for gefitinib in complex with EGFR (T790M, L858R, and T790M) and Balius et al.[39] for erlotinib, gefitinib, and AEE788 with EGFR (wild type, L858R, L858R, and T790M). Focusing on the latter study, Balius et al.[39] compellingly showed that drug resistance in EGFR, in contrast to proposed mechanisms primarily involving a steric clash[27,41] or altered affinity for ATP,[30] more likely involves disruption of favorable interactions, including an important water-mediated H-bond network. Free energy calculations for inhibitors with EGFR, reported by Michel et al.,[40] similarly revealed a local water network that mediates binding. In general, hydration is known to be an important factor in drug design.[40,42−45] In particular, specific waters can help to mediate protein−ligand binding.[43] Examples of other FDA-approved drugs[6] that make water-mediated interactions include erlotinib,[19] zanamivir,[46] and nevirapine.[47] For the lapatinib simulations presented here, bridging waters also appear to play a key role in modulating both specificity and resistance among ErbB family tyrosine kinases.

Our long-term goal is the development of improved anticancer small molecule drugs. Goals of this study are threefold: (i) to construct robust all-atom computational models to quantify lapatinib binding with inactive forms of wild-type EGFR, HER2, and ErbB4 and deleterious mutations, including EGFR C775F, EGFR T854A, EGFR T790M, and HER2 T790I, (ii) to determine how receptor sequence and associated structural changes lead to variation in computed lapatinib activities for comparison with experimental results, and (iii) to elucidate the primary factors controlling lapatinib specificity and resistance. The design of inhibitors with tailored selectivity and resistance profiles will ultimately be enabled through use of well-tested models for quantifying and predicting molecular recognition at the atomic level.

METHODS

Model Construction. Simulation-ready models for three different ErbB systems (EGFR, HER2, and ErbB4) were constructed using one of two crystallographic structures termed here template 1 (PDB entry 1XKK containing EGFR with lapatinib[17]) and template 2 (PDB entry 2R4B containing ErbB4 with a lapatinib-like covalent inhibitor[48]). 1XKK was chosen as a template because it was the only inactive structure available for EGFR. 2R4B was chosen as a template for inactive ErbB4 because there was only a short section missing (six residues) in comparison to an alternative structure 3BBT[24] that was missing 14 residues in the activation loop region. No inactive structures were available for HER2, although an active form structure was recently published.[49] Structural changes in the binding site reveal lapatinib is compatible only with inactive conformations. Modeller9v6[50] was employed to construct homology models using both inactive templates. Figure 2 outlines the overall workflow schematically.

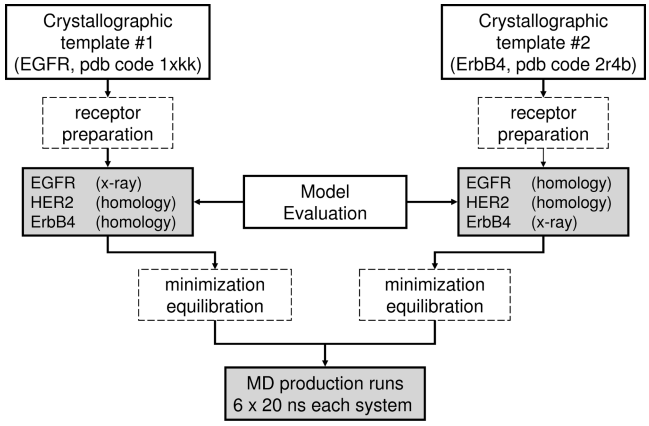


Figure 2. Overall workflow schematic showing model construction for EGFR, HER2, and ErbB4 using two crystallographic templates.

For lapatinib, the ligand conformation observed in template 1XKK was used as the initial set of coordinates for all simulations. Template 2R4B contained a lapatinib-like covalent inhibitor, and there was incomplete ligand electron density in 3BBT. Interestingly, 3BBT contained an alternative rotamer amounting to an ∼180° ring flip of the fluorophenyl moiety relative to 1XKK or 2R4B. However, average ligand *B* factors in 3BBT were more than double (100.2) compared to those in 1XKK (41.8), which suggests substantial uncertainty. In addition, preliminary studies showed an energetic preference for the 1XKK conformation, and some simulations starting from the 3BBT conformation flipped to 1XKK but never vice versa. Additional studies to investigate the discrepancy observed in the two crystallographic structures are in progress.

On the basis of sequence alignments using ClustalW,[51] the kinase receptor domains share very high levels of sequence identity: EGFR and HER2 (81%), EGFR and ErbB4 (78%), and HER2 and ErbB4 (77%). Thus, homology models for the receptors can be constructed with high degrees of confidence. To assess the effects of homology modeling, two complete sets of three systems each were constructed (EGFR, HER2, and ErbB4) using both templates (1XKK and 2R4B). Missing regions in 1XKK (residues 734−737, 750−753, and 868−875) were completed on the basis of PDB entry 2GS7[18] containing inactive EGFR with ligand AMP-PNP, while the missing region in 2R4B (residues 755−760) was completed using Modeler. Otherwise, conformations for identical residues in the alignments were kept the same as those in the relevant template.

For each complex, 10 homology models were generated and the model with the lowest discrete optimized protein energy (DOPE)[52] score was selected as the initial structure for molecular dynamics simulations. Models were evaluated using PROCHECK,[53] which showed no residues in disallowed regions and ∼90% of the residues in the most favored regions. These metrics help confirm the structures are of overall good quality. In general, the six completed receptors, based on template 1XKK (EGFR, HER2, and ErbB4) or template 2R4B (EGFR, HER2, and ErbB4), were structurally similar except for some differences in the activation loop, C-helix, and N-lobe regions (Figure 3, arrows).

To probe how point mutations would affect ligand binding, additional setups for mutant forms of EGFR (C775F, T854A,

dx.doi.org/10.1021/bi2016553 | *Biochemistry* 2012, 51, 2390−2406

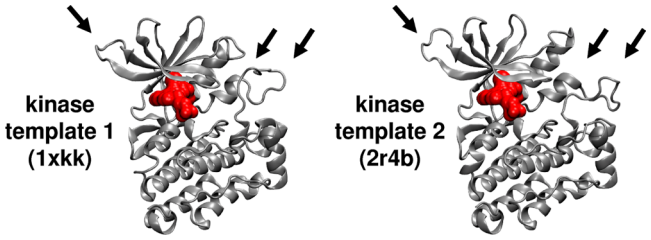


Figure 3. Homology modeling templates derived from crystallographic coordinates of inactive EGFR (PDB entry 1XKK) or inactive ErbB4 (PDB entry 2R4B). Arrows denote regions with structural differences between the templates.

and T790M), HER2 (T790I), and ErbB4 (V775C) were constructed through modification of the relevant wild-type structures using MOE.[54] Starting rotamers for mutant side chains were chosen to overlap well with the wild-type side chains subject to visual inspection to ensure there were no steric clashes. The same equilibration and production protocols described below for wild-type systems were employed for the mutants.

Simulation Setups and Molecular Dynamics Details. Molecular dynamic simulations were performed with NAMD2.6[55] using input files prepared using the AMBER8[56] suite of programs that was used to assign force field parameters and assemble and solvate each complex in a periodic box containing ~45000 waters (~80 Å^3). No crystallographic waters were employed. The following force field parameters were used: FF99SB[57] for the protein, GAFF[58] for the ligand, and TIP3P[59] for water. Ligand charges were obtained at the HF/6-31G* level of theory using the ChelpG[60] method as implemented in Gaussian98.[61] On the basis of the environment, residue His 803 in the ErbB4 binding site was modeled as protonated. Lapatinib was modeled as neutral. Prior to the production runs, a nine-step equilibration procedure was used to relax the protein and solvent in a sequential way and consisted of three minimization steps (1000 steps of steepest decent) and six molecular dynamics runs (50 ps each). The production runs employed weak restraints only on short stretches of the N-termini (10 amino acids) and C-termini (14 amino acids). Slightly different equilibration protocols were used for crystallographic versus homology models. Table 2 summarizes the various equilibration steps.

Molecular dynamics simulations were performed in the *NPT* ensemble using Langevin dynamics[62] at a specified constant temperature of 298.15 K and a pressure of 1.01325 bar. MD equilibrations employed a 1 fs time step, while production runs employed a 2 fs time step that required the use of SHAKE.[63] Additional key input parameters include use of the particle mesh Ewald[64] to compute long-range electrostatics (1.0 Å grid spacing) and a 12 Å direct space cutoff (10 Å smoothing switch) for nonbonded interactions. MD snapshots were saved every 1 ps during the productions runs for subsequent analysis. For each wild-type system, six 20 ns production simulations were performed (different random seeds), while for each mutant system, one 20 ns production simulation was performed.

Calculation of Binding Free Energies. Free energies of binding [ΔG_b(calcd)] were estimated using the well-described MM-GBSA method[65,66] with AMBER, which our laboratory has employed to successfully characterize a number of systems,[67−70] including active form EGFR.[39] For this work, free energies were estimated using a four-term equation [$\Delta G_b(\text{calcd}) = \Delta E_{vdw} + \Delta E_{coul} + \Delta\Delta G_{polar} + \Delta\Delta G_{nonpolar}$] consisting of the intermolecular van der Waals energy (ΔE_{vdw}), the intermolecular Coulombic energy (ΔE_{coul}), changes in polar hydration energy ($\Delta\Delta G_{polar}$), and changes in nonpolar hydration energy ($\Delta\Delta G_{nonpolar}$). No additional entropic terms were included. Polar hydration energies were obtained using the GB model described by Onufriev et al.[71] (igb = 5) with mbondi2 radii and interior and exterior dielectric constants of 1 and 78.5, respectively. Nonpolar hydration energies were obtained from solvent accessible surface areas through the relationship $\Delta G_{nonpolar} = \gamma \times \text{SASA} + \beta$, where γ = 0.00542 kcal mol^{-1} Å^{-2} and β = 0.92 kcal/mol.[72] The AMBER distribution file src/sander/mdread.f was modified to include radii of fluorine (1.50 Å) and chlorine (1.70 Å) required by GBSA calculations for lapatinib. Free energies of binding were computed using three different protocols (0WAT, 3WAT, and 1WAT) depending on the number of explicit waters retained from the fully solvated MD trajectories. 0WAT includes no waters and is the default MM-GBSA protocol; 3WAT includes the three waters closest to N* (labeled in Table 1), and 1WAT includes one water if it is within 3 Å of N*.

RESULTS AND DISCUSSION

System Behavior and Convergence. *Structural Stability (crystallographic versus homology starting coordinates).* To assess the behavior of the simulations and, in particular, to assess stability of the homology models, we examined root-mean-square-deviations (rmsds) as a function of

Table 2. Equilibration Protocol for Crystallographic and Homology Models

step[a]	crystallographic restraints[b] (weight)[c]	homology restraints (weight)
1 min	wat O (5), pro all (5), lig all (5)	wat O (5), pro homo all (5), pro nonhomo hev (5), lig all (5)
2 min	pro all (5), lig all (5)	pro hom all (5), pro nonhomo back (5), pro nonhomo side (1), lig all (5)
3 min	pro hev (5), lig hev (5)	pro homo hev (5), pro nonhomo back (5), lig hev (5)
4 md	pro hev (5), lig hev (5)	pro homo hev (5), pro nonhomo back (5), lig hev (5)
5 md	pro hev (1), lig hev (1)	pro homo hev (1), pro nonhomo back (1), lig hev (1)
6 md	pro Cα (1), pro hev (0.1), lig hev (0.1)	pro Cα (1), pro homo hev (0.1), pro nonhomo back (0.1), lig hev (0.1)
7 md	pro Cα (1)	pro Cα (1)
8 md	pro Cα (0.5)	pro Cα (0.5)
9 md	pro Cα (0.1), 702−711, 969−982	pro Cα (0.1), 702−711, 969−982

[a]Abbreviations: min, energy minimization; md, molecular dynamics. [b]Abbreviations: wat, water; pro, protein; lig, ligand; hev, heavy atoms; homo, residues homologous to those of the template; nonhomo, residues not homologous to those of template; back, backbone Cα, C, N, O; side, side chain. [c]Restraint weights in parentheses in kilocalories per mole per square angstrom.

dx.doi.org/10.1021/bi2016553 | *Biochemistry* 2012, 51, 2390−2406

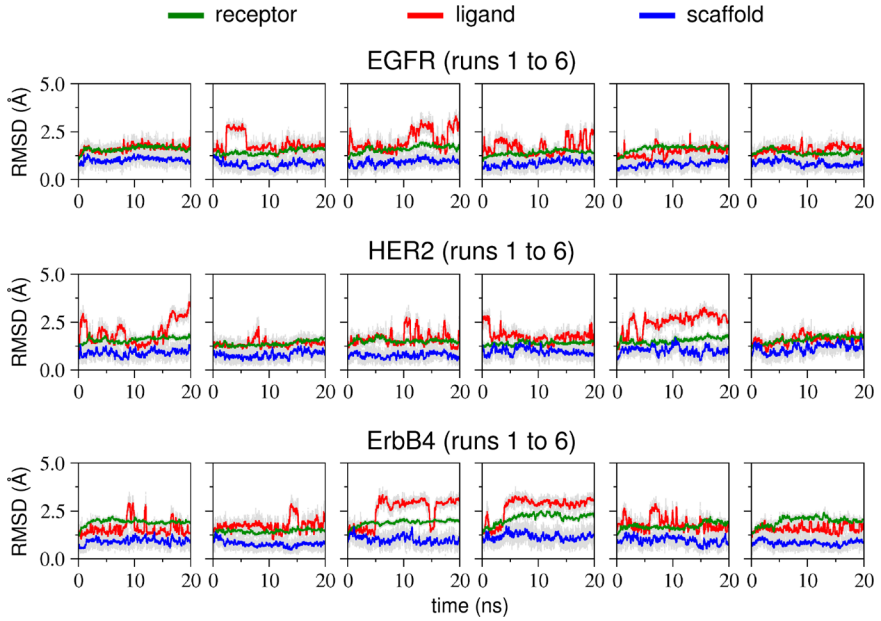


Figure 4. Root-mean-square-deviation (rmsd) vs time (nanoseconds) for lapatinib with EGFR (top), HER2 (middle), and ErbB4 (bottom) for six MD runs based on template 1XKK. The rmsds are shown for the receptor backbone (green line), ligand (red line), and ligand scaffold quinazoline plus aniline rings (blue line).

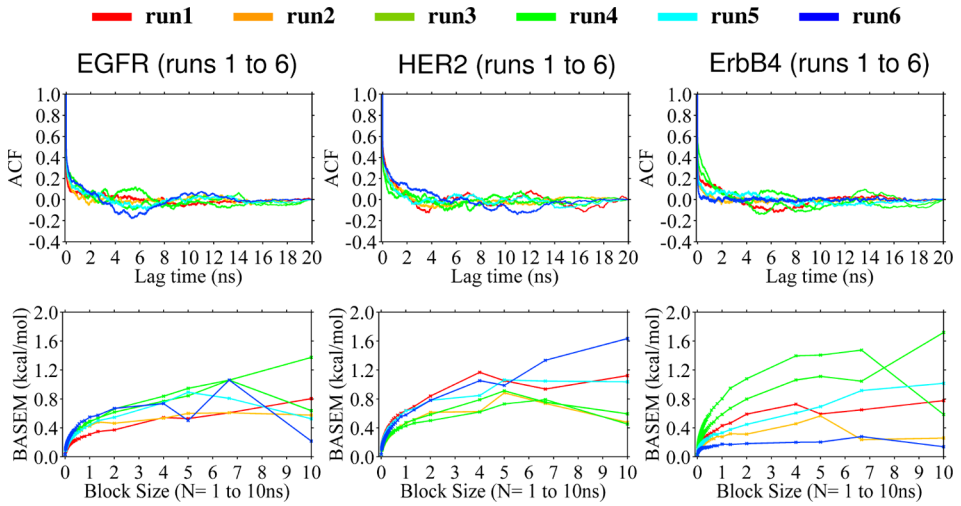


Figure 5. Autocorrelation functions (ACFs) (top) for free energies of binding of lapatinib with EGFR (left), HER2 (middle), and ErbB4 (right) computed from six independent 20 ns MD simulations each. Block-averaged standard errors of the mean (BASEM) (bottom) in kilocalories per mole as a function of block size. Free energies of binding calculated with three binding site waters (3WAT protocol).

time for each wild-type setup. To more accurately estimate binding free energies, six simulations using different random seeds were performed. Figure 4 plots results initiated from the crystallographic structure of EGFR (1XKK) and the two homology models of HER2 and ErbB4 constructed using 1XKK as the template. Here, light gray points show instantaneous rmsds with colored lines representing the running average of the previous 100 snapshots for receptor Cα, C, N, and O backbone atoms (green line), lapatinib heavy atoms (red lines), or lapatinib quinazoline plus aniline rings defined as the scaffold (blue lines). The rmsds were computed after each MD production snapshot was fit to the initial model coordinates using receptor Cα backbone atoms as the match criteria. Importantly, for each of the six replicas across all systems, the receptor backbone rmsds are within 2 Å of each other, which is an indication that both the crystallographic and the homology models have structural integrity. While ligand total rmsds, in some cases, show fluctuations of >2 Å, the ligand scaffold rmsds always remain low (1−1.5 Å), which suggests it is only the solvent-exposed methyl sulfone tail (see Table 1) that fluctuates significantly. Similar ligand rmsd results were obtained in an earlier study of inhibitors that target the active form of EGFR.[39]

Binding Free Energy (crystallographic versus homology starting coordinates). To help gauge convergence of the computed free energies of binding [ΔG_b(calcd)], autocorrelation functions (ACFs) and block-averaged standard error of the mean (BASEM) analysis were also performed.[73,74] For a given time series, as demonstrated in Figure 5 for the group of six simulations based on template 1XKK, ACFs provide a means of assessing the extent to which data are correlated, while BASEM plots can be used to quantify statistical noise. The ΔG_b(calcd)

dx.doi.org/10.1021/bi2016553 | *Biochemistry* 2012, 51, 2390−2406

Table 3. Autocorrelation Function Percent Uncorrelated Data (ACF %) and Block-Averaged Standard Errors of the Mean (BASEM) for ΔG_b(calcd) Values from Simulations of Lapatinib with EGFR, HER2, and ErbB4 for Various Block Lengths Based on 1XKK and 2R4B Templates

	N = 1 ps		N = 100 ps		N = 1000 ps		N = 2000 ps	
	ACF %[b]	BASEM[c]	ACF %[b]	BASEM[c]	ACF %[b]	BASEM[c]	ACF %[b]	BASEM[c]
				Template 1[a]				
EGFR	47.55	0.03	26.79	0.18	10.11	0.45	3.68	0.56
HER2	52.68	0.03	31.78	0.21	11.24	0.54	6.43	0.68
ErbB4	44.50	0.03	23.64	0.17	8.93	0.43	4.22	0.57
average	48.24	0.03	27.40	0.19	10.09	0.47	4.78	0.60
				Template 2[a]				
EGFR	43.36	0.03	22.91	0.16	7.84	0.41	1.51	0.52
HER2	48.75	0.03	27.44	0.19	10.81	0.48	6.29	0.63
ErbB4	46.71	0.03	25.26	0.18	8.95	0.43	6.77	0.55
average	46.27	0.03	25.20	0.18	9.20	0.44	4.86	0.57

[a]Template 1 based on PDB entry 1XKK, template 2 based on PDB entry 2R4B. [b]ACF as a percentage of correlated data. [c]BASEM energies in kilocalories per mole. Free energies of binding calculated with three binding site waters (3WAT protocol).

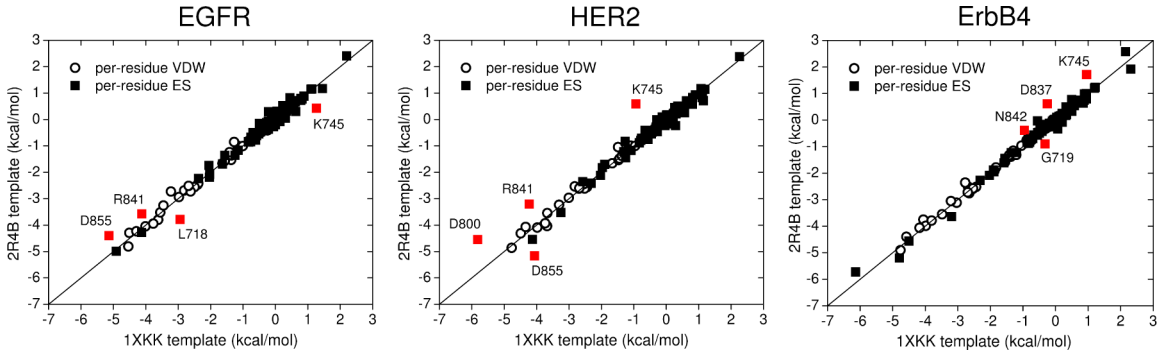


Figure 6. Per-residue van der Waals (VDW) and Coulombic (ES) interaction energies for lapatinib with EGFR (left), HER2 (middle), and ErbB4 (right) in which the receptors were constructed using different homology modeling templates based on PDB entry 1XKK (EGFR) or 2R4B (ErbB4). Average energies computed from 120000 MD snapshots (see the text for details). For each plot, red squares denote the four residues with the largest variation between the two templates. Diagonal line $Y = X$.

values in Figure 5 were obtained using a three-water protocol (3WAT), as described further below; however, two other protocols (0WAT and 1WAT) led to similarly good convergence. Focusing on the EGFR results, Figure 5 reveals the ACFs drop quickly, and by the ~1 ns lag time, the data are largely uncorrelated (80−90%). In addition, the accompanying BASEM results show the expected monotonical increase as block averaging size increases (from 1 ps to 10 ns), which begins to reach a plateau indicating the error estimates are converging. The same general trend observed for EGFR is maintained for the other systems, including the simulations based on the alternative template 2R4B or other protocols. Interestingly, one HER2 simulation yielded ACF and BASEM results that showed poor convergence behavior in comparison to that of others simulations. A simulation generated using an alternative random seed that was better behaved was substituted for the unconverged run. For comparison, Table 3 quantifies ACF and BASEM values computed at four different lag time and block size values (1, 100, 1000, and 2000 ps). Overall, by using a 1 ns block size, uncertainties in ΔG_b(calcd) values on the order of 0.5 kcal/mol may be considered to be a reasonable estimate. ACF data at this interval are relatively uncorrelated (~90%), and a sufficient number of blocks (N = 20) can be employed for BASEM calculations.

Influence of Homology Model Templates on Interaction Energy Variability. To further explore how the use of different templates might affect the energetic results, protein−ligand van der Waals and Coulombic interaction energies, on a per-residue basis, were plotted against each other as shown in Figure 6 for the range from −7 to 3 kcal/mol. Notably, both homology templates yield very similar energetics. In all cases, per-residue van der Waals interactions are essentially identical [Figure 6 (open circles)], and only a few residues (D855, N842, R841, D837, D800, K745, G719, and L718) show variability in their gas-phase Coulombic energies [Figure 6 (red squares)]. It is interesting that many of the residues showing Coulombic differences are charged and/or in loop regions;[19] thus, larger variation may not be unexpected. Overall, the strong similarity in per-residue interaction starting from two different homology modeling templates is striking. We attribute the good correspondence to (i) the fact that the homology models are based on a very high level of sequence homology of ~80% and (ii) the fact that the EGFR, HER2, and ErbB4 kinase domains are all of the same length (no insertions or deletions), which minimized the actual homology modeling required (i.e., side chains in common were kept in their respective crystallographic template position).

Importance of Binding Site Waters. In general, water molecules are known to play important roles in molecular recognition.[40,42−45] Our laboratory's prior study of the active form of EGFR identified an important network of waters involved in mediating resistance to the anti-EGFR drugs erlotinib and gefitinib.[39] For the inactive forms studied here, examination of explicit solvent MD trajectories of EGFR,

dx.doi.org/10.1021/bi2016553 | *Biochemistry* 2012, 51, 2390−2406

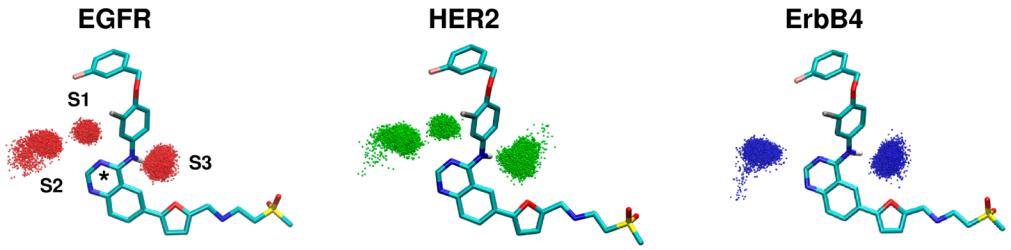


Figure 7. Representative MD trajectories showing primary hydration sites labeled S1, S2, and S3 from simulations of lapatinib with EGFR (left, red), HER2 (middle, green), and ErbB4 (right, blue) based on template 1XKK. Images show overlaid waters (O atoms) within 3 Å of ligand atom N* (S1), kinase carbonyl oxygen at residue 791 (S2), and ligand aniline hydrogen (S3) from 10000 evenly spaced MD snapshots (20 ns trajectories).

Table 4. Calculated versus Experimental Free Energies of Binding

absolute	0WAT[a] ΔG_b(calcd)[d]	3WAT[b] ΔG_b(calcd)	1WAT[c] ΔG_b(calcd)	ΔG_b(exptl)[e]
EGFR avg[f]	−61.73	−68.23	−65.33	−11.62
EGFR 1[g]	−62.19	−68.80	−65.84	
EGFR 2[h]	−61.27	−67.66	−64.81	
HER2 avg	−61.08	−67.71	−64.57	−10.75
HER2 1	−61.39	−67.82	−64.53	
HER2 2	−60.77	−67.60	−64.61	
ErbB4 avg	−62.96	−66.91	−63.20	−8.81
ErbB4 1	−62.93	−67.19	−63.35	
ErbB4 2	−62.99	−66.63	−63.05	
difference (avg)	0WAT $\Delta\Delta G_b$(calcd)	3WAT $\Delta\Delta G_b$(calcd)	1WAT $\Delta\Delta G_b$(calcd)	$\Delta\Delta G_b$(exptl)
EGFR − HER2	−0.65	−0.52	−0.76	−0.87
HER2 − ErbB4	1.88	−0.80	−1.37	−1.94
EGFR − ErbB4	1.23	−1.32	−2.13	−2.81

[a]0WAT protocol computed with zero waters. [b]3WAT protocol computed with the three waters closest to N* (labeled in Table 1). [c]1WAT protocol computed using one water if it is within 3 Å of N*. [d]ΔG_b(calcd) values in kilocalories per mole from six 20 ns simulations. [e]ΔG_b(exptl) ≈ RT ln(activity) at 298.15 K using K_i values from Table 1. [f]Average of templates 1 and 2. [g]Computed using template 1 (1XKK). [h]Computed using template 2 (2R4B).

HER2, and ErbB4 with lapatinib reveals a similar network of waters. As illustrated in Figure 7, which plots a representative MD trajectory for each system, the analysis identified that for EGFR and HER2 there are three distinct sites with high water occupancies (defined here as S1, S2, and S3) in contrast to ErbB4 for which the S1 site, on average across all six runs, shows a much lower occupancy. The definitions are based on waters within 3 Å of ligand atom N* (S1), kinase carbonyl oxygen at residue 791 (S2), or ligand aniline hydrogen (S3). On the basis of such solvation patterns, we hypothesized that ligand binding would be influenced by differences in hydration that could occur as a result of differences in the primary amino acid sequence. It is important to note that these water positions are a result of MD sampling as no crystallographic waters were included in the initial setups. Comparison with available crystallographic data is interestingly mixed. For EGFR with lapatinib (1XKK), the water at site S1 is observed. For HER2, no inactive structures have yet been published. For ErbB4 with lapatinib (3BBT), strangely, no waters are included in the binding site, although the surrounding environment contains water. For ErbB4 with a related covalently bound analogue (2R4B), water at site S1 is observed. In an attempt to quantify the effects of hydration at these sites, as described below, we computed binding free energies using different numbers of key explicit waters.

Correlation with Experimental Binding Trends. Table 4 shows free energies of binding [ΔG_b(calcd)] obtained using one of three calculation protocols that includes zero explicit waters (0WAT), three explicit waters (3WAT), or one explicit water (1WAT). 0WAT represents the default MM-GBSA protocol, 3WAT is designed to capture hydration effects at sites S1−S3, and 1WAT is designed to isolate effects primarily due to site S1. It is important to note that for any given simulation a single solvated trajectory is postprocessed to derive 0WAT, 3WAT, and 1WAT protocols, the underlying protein−ligand ensembles in each case are identical, and it is only the numbers of explicit waters retained that are different. A number of prior studies[75−77] have also included key binding site waters using related free energy calculation protocols. For each receptor setup, ΔG_b(calcd) values are presented for the combined average (avg) of templates 1XKK (template 1) and 2R4B (template 2) as well as each individual template. Each template group represents the average of six well-converged 20 ns simulations. The discussions that follow employ the template-averaged data. In terms of magnitude, the overestimation of absolute ΔG_b(calcd) values in Table 4 relative to experiment is a well-known occurrence when explicit solute entropic terms are omitted, as was the case in this study. Focusing on the differential energies ($\Delta\Delta G_b$), for which neglected entropic terms may be reasonably assumed to cancel, we found the template-averaged results show remarkable agreement with the experimental trends provided that key explicit waters (3WAT or 1WAT protocol) are included in the calculations. Calculations in which no explicit waters are retained (0WAT

dx.doi.org/10.1021/bi2016553 | Biochemistry 2012, 51, 2390−2406

Table 5. Absolute (Δ) Binding Energy Components (kilocalories per mole)

	ΔE_{vdw}	ΔE_{coul}	ΔG_{polar}	$\Delta G_{nonpolar}$	ΔG_b(calcd)	ΔG_b(exptl)
			0WAT			
EGFR avg	−75.91	−31.40	53.17	−7.56	−61.73	−11.62
HER2 avg	−75.83	−32.89	55.22	−7.58	−61.08	−10.75
ErbB4 avg	−77.00	−35.59	57.22	−7.60	−62.96	−8.81
			3WAT			
EGFR avg	−76.86	−40.61	56.67	−7.44	−68.23	−11.62
HER2 avg	−76.54	−42.50	58.78	−7.45	−67.71	−10.75
ErbB4 avg	−77.74	−42.61	60.93	−7.49	−66.91	−8.81
			1WAT			
EGFR avg	−76.22	−35.64	54.07	−7.53	−65.33	−11.62
HER2 avg	−76.07	−36.92	55.95	−7.53	−64.57	−10.75
ErbB4 avg	−77.00	−35.94	57.33	−7.59	−63.20	−8.81

protocol) incorrectly predict lapatinib to bind most tightly to ErbB4.

As shown in Table 4, use of the 3WAT protocol increases the averaged absolute binding energies, relative to those derived with the 0WAT protocol, nearly identically for EGFR (−6.50 kcal/mol) and HER2 (−6.63 kcal/mol) in contrast to that of ErbB4 (−3.95 kcal/mol), which is smaller. The smaller energetic gain for ErbB4 makes physical sense given the lack of S1 water observed in the hydration patterns (Figure 7). For the 1WAT protocol, again nearly identical increases in binding energy are observed for EGFR (−3.60 kcal/mol) and HER2 (−3.49 kcal/mol) in comparison to that of ErbB4, which is significantly smaller (−0.24 kcal/mol). Here, the effect for ErbB4 is more pronounced given that only the S1 water is considered. Notably, the 1WAT protocol yields remarkable numerical agreement with the experimental $\Delta\Delta G_b$ differences: (i) EGFR − HER2 = −0.76 for calcd versus −0.87 for exptl, (ii) HER2 − ErbB4 = −1.37 for calcd versus −1.94 for exptl, and (iii) EGFR − ErbB4 = −2.13 for calcd versus −2.81 for exptl. Overall, the solvated results underscore the importance of considering binding site waters in these systems. In particular, as discussed further below, different amounts of solvent in the S1 site appear not only to drive the specificity of lapatinib for the three different receptors but also to play a role in how clinically relevant point mutations affect ligand binding.

Component Contributions to Affinity. To gauge which of the underlying energy terms comprising ΔG_b(calcd) change most as a result of including explicit waters, component analysis was performed as shown in Table 5 for the different protocols. Here, as was observed in our earlier study of active form EGFR, the affinity for lapatinib with the inactive kinases also appears to be most strongly driven by favorable intermolecular van der Waals interactions (Table 5). Favorable intermolecular ΔE_{coul} terms are roughly half of the accompanying favorable ΔE_{vdw} terms, and the unfavorable desolvation terms (ΔG_{polar}) are significant. Taken together, the results suggest molecular association in these systems is primarily driven by steric packing. However, selectivity across the series (EGFR > HER2 > ErbB4) does not appear to correlate with variation in packing. In every case, ΔE_{vdw} terms for ErbB4 are more favorable ($\leq$ −77 kcal/mol) compared with the other two receptors (> −77 kcal/mol).

As shown in Table 5, addition of binding site waters (3WAT or 1WAT) affects electrostatic contributions (ΔE_{coul} and ΔG_{polar}) much more dramatically than the packing terms (ΔE_{vdw} and $\Delta G_{nonpolar}$). As water-mediated hydrogen bonding is primarily electrostatic in nature, this observation is reasonable. The largest changes in ΔE_{vdw} are <1 kcal/mol. In sharp contrast, for the 3WAT protocol relative to the 0WAT protocol, ΔE_{coul} favorably increases by approximately −9 kcal/mol for EGFR and HER2 versus −7 kcal/mol for ErbB4, and for the 1WAT protocol, the differences are approximately −4 kcal/mol for EGFR and HER2 versus −0.4 kcal/mol for ErbB4. The negligible ΔE_{coul} change for ErbB4 is a function of less favorable hydration at site S1. In support of this conclusion, all the energetic terms (ΔE_{vdw}, ΔE_{coul}, ΔG_{polar}, and $\Delta G_{nonpolar}$) for ErbB4 from the 0WAT or 1WAT results are nearly identical. In fact, the smaller increases in ΔE_{coul} (favorable) and ΔG_{polar} (unfavorable) for ErbB4 versus the other two systems are the primary reasons why the correct experimental ordering (EGFR > HER2 > ErbB4) is obtained using solvated ΔG_b(calcd) protocols. Finally, it is interesting to note that of all the terms, ΔG_{polar} most closely tracks the experimental trend. This provides additional evidence that selectivity for these receptors is a function of differential solvation in the bound state.

Footprint Contributions to Affinity. Although structure−activity relationships can often be explained using per-residue energetic decompositions (molecular footprints), an examination here suggests lapatinib specificity is not governed directly by changes in residue-specific intermolecular interactions. As shown in Figure 8, which plots results based on template 1XKK simulations, the remarkably high degree of overlap in the three van der Waals (ΔE_{vdw}) footprints derived from the EGFR, HER2, and ErbB4 results reveals no specific residue(s) that might play a role. While not a particularly useful indicator of ErbB family specificity, the tight ΔE_{vdw} overlap in Figure 8 does provide additional evidence that the homology model results are robust and well-converged. In contrast, Coulombic energy (ΔE_{coul}) footprints do show variability at approximately five positions (KKK 745, CSV 775, DDE 800, RRH 803, and DDD 855), and these sites were examined in greater detail. This nomenclature refers to the amino acid present at any given position following the EGFR−HER2−ErbB4 receptor order. At positions KKK 745, CSV 775, and RRH 803, the interactions between lapatinib and the accompanying residue in ErbB4 are in fact less favorable than those in EGFR, HER2, or both; thus, these residues could possibly play a role in the weakened binding. However, these ErbB4 losses are not enough to compensate for gains in favorable ΔE_{coul} energies at positions DDE 800 and DDD 855. While a localized increase at position 800 with ErbB4 may make physical sense, in terms of a DDE change contributing to a more favorable global ΔE_{coul} term (see the 0WAT results in Table 5), the net effect is in the wrong direction if ligand−

dx.doi.org/10.1021/bi2016553 | Biochemistry 2012, 51, 2390−2406

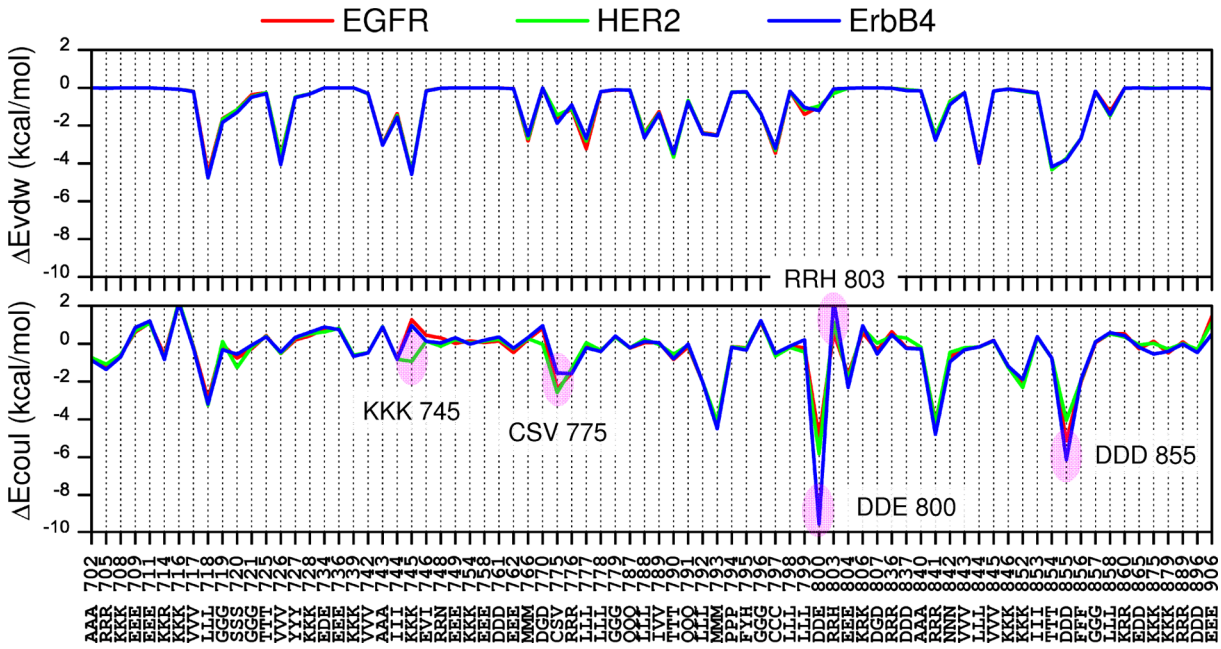


Figure 8. Average per-residue van der Waals (ΔE_{vdw}) and Coulombic (ΔE_{coul}) footprints for lapatinib with EGFR, HER2, and ErbB4 from six 20 ns simulations derived from the 1XKK template. Single-letter amino acid codes on the *x*-axis indicate, at any given position, the side chain present for EGFR (bottom), HER2 (middle), or ErbB4 (top).

protein intermolecular Coulombic interactions alone were driving specificity. Similarly, the favorable increase at residue 855 with ErbB4 is counterintuitive, although it is interesting that this position involves the DFG motif (residues 855−857) that can adopt an inactive (DFG-out) or active (DFG-in) conformation.[78] Although the footprints do not provide a direct route for understanding specificity, indirectly, the relatively small energetic change at position 775 involving a swap from a polar (C or S) residue in EGFR or HER2 to a nonpolar (V) residue in ErbB4 has very large consequences in terms of changing ligand hydration.

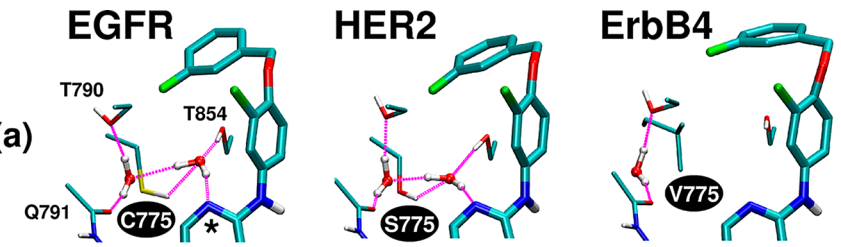
Primary Sequence Differences Alter the Solvation That Drives Specificity. Examination of MD trajectories with explicit solvent reveals hydration at sites S1 and S2 to be mediated primarily by four residues close to the ligand: CSV 775, TTT 790, QQQ 791, and TTT 854 (Figure 9). Of these positions, only position 775 involve an amino acid change. The representative snapshots shown in Figure 9a highlight the quadrifurcated H-bonding network involving sites S1 and S2, while Figure 9b quantifies differences in terms of H-bond counts and Coulombic energy for the three waters closest to N* with nearby functionality (ligand N* atom, all ligand heavy atoms, or residue CSV 775, TTT 790, QQQ 791, or TTT 854). H-Bond counts employed a 3 Å acceptor−donor pair cutoff and an X_D-H_D----X_A angle between 120° and 180°. Values represent averaging over 120000 MD snapshots using the structures derived from template 1XKK. Importantly, the results show high-occupancy waters of >70% with all four residues being examined, which indicate long-lived interactions, with the notable exception that for simulations of lapatinib with ErbB4 the H-bound counts and Coulombic energies are dramatically lower.

The binding site graphic in Figure 9a visually highlights why a bulkier side chain at position 775 in ErbB4 (V), without the same H-bonding capability as in EGFR (C) or HER2 (S), is the most likely cause of reduced water occupancy. Quantitatively (Figure 9b), there are zero waters H-bonding at this position in ErbB4 compared to 0.71 and 0.85 in EGFR and HER2, respectively, and the change to valine reduces the number of counts involving the N* atom of lapatinib with ErbB4 (0.11) compared with EGFR (0.88) and HER2 (0.75). Counts for the total ligand show a similar trend (1.65 for ErbB4 vs 2.30 for EGFR vs 2.15 for HER2). It is probable these differences are the primary contributors to the larger desolvation penalties in ErbB4 (larger ΔG_{polar} terms in Table 5). Interestingly, water H-bonds involving residue TTT 790 do not appear to be reduced in simulations of ErbB4 relative to EGFR; however, at QQQ 791, and in particular for TTT 854, there are reductions (Figure 9b). In the latter case for T854, the number of interactions ErbB4 (0.11) versus EGFR (0.89) or HER2 (0.76) is essentially the same as those obtained from calculations involving only the ligand N* atom. The correspondence here confirms a highly coupled water site. Interestingly, counts with the ligand track with the experimental ordering: EGFR (2.30) > HER2 (2.15) > ErbB4 (1.65). Overall, the summed H-bond counts for waters with the ligand and the four nearby residues are ∼40% lower in ErbB4 (3.51) than in EGFR (6.03) or HER2 (5.79).

Examination of Coulombic energies (ΔE_{coul}) reveals a similar trend with averaged water−residue interactions being significantly less favorable in the ErbB4 binding site than in EGFR or HER2 (Figure 9b). The fact the summed ΔE_{coul} interactions (ligand and four nearby residues) are again reduced by roughly 40% for ErbB4 (−18.91 kcal/mol) versus EGFR (−29.28 kcal/mol) or HER2 (−31.88 kcal/mol) indicates that in these systems Coulombic losses scale linearly with changes in hydrogen bonding. As before, the most significant losses occur at CSV 775 and TTT 854, and similar to the H-bond results, variation in ΔE_{coul} for waters with the ligand also tracks with experiment: EGFR (−9.26 kcal/mol) > HER2 (−9.19 kcal/mol) > ErbB4 (−7.21 kcal/mol). Interestingly, the ∼2 kcal/mol lower ΔE_{coul} for ErbB4 versus those of the other receptors is roughly similar in magnitude to $\Delta\Delta G_b$(exptl)

dx.doi.org/10.1021/bi2016553 | *Biochemistry* 2012, 51, 2390−2406

(a)



(b)

Residues	Population (count) EGFR	HER2	ErbB4	ΔE_{coul} energy (kcal/mol) EGFR	HER2	ErbB4
ligand atom N*	0.88	0.75	0.11			
ligand	2.30	2.15	1.65	-9.26	-9.19	-7.21
CSV 775	0.71	0.85	0	-1.63	-4.47	-0.90
TTT 790	0.76	0.90	0.77	-3.87	-4.55	-3.36
QQQ 791	1.37	1.13	0.98	-6.33	-6.77	-5.63
TTT 854	0.89	0.76	0.11	-8.19	-6.90	-1.81
sum (excluding N*)	6.03	5.79	3.51	-29.28	-31.88	-18.91

Figure 9. Water-mediated H-bonds involving lapatinib and nearby residues in EGFR, HER2, and ErbB4. (a) Representative snapshots showing the quadrifurcated H-bonding network involving S1 and S2 waters. (b) Averaged H-bond counts and pairwise Coulombic interaction energies (N = 120000 frames each) for the three waters closest to N* with N*, the entire ligand, or residue CSV 775, TTT 790, QQQ 791, or TTT 854. Results are based on template 1XKK.

(ErbB4 − EGFR = 2.81 kcal/mol, and ErbB4 − HER2 = 1.94 kcal/mol).

Effects of Mutation. At least two mechanisms of resistance impact anticancer drugs that target receptor tyrosine kinases: (i) redundant survival pathways activated as a consequence of inhibition and (ii) point mutations that directly alter ligand binding. With a focus on the latter mechanism, although point mutations in EGFR or HER2 are not commonly associated with lapatinib resistance as a result of clinical use to treat breast cancer, there is a concern that mutation sites observed as a result of treatment with related inhibitors for other cancers could eventually impact lapatinib.[14,15,25] Using in vitro screening, Avizienyte et al.[14] and Trowe et al.[15] both reported numerous ErbB family mutations that negatively affect lapatinib activity. Compellingly, several mutations identified (C775F, T854A, T790M, S775P, and T790I) map here to sites involved in water-mediated ligand binding. Analogous to the EGFR > HER2 > ErbB4 selectivity arguments discussed above, it is reasonable to propose that changes affecting water-mediated binding would also lead to drug resistance. To test this hypothesis, additional MD simulations and analysis (Figure 10 and Table 6) for lapatinib in complexes with the following mutant kinases were performed: EGFR (C775F, T854A, and T790M), HER2 (T790I), and ErbB4 (V775C). Simulations were based on template 1XKK, and for each mutant system, a single 20 ns MD run was performed.

ErbB4 V775C. The first mutation studied, ErbB4 V775C (Figure 10h and Table 6h), is a hypothetical mutation that was primarily introduced to reaffirm the conclusions reached above that the observed differences in experimental activity for the EGFR > HER2 > ErbB4 series are primarily driven by the character of the side chain at position 775. It was hypothesized that mutating the native nonpolar valine back to a polar cysteine (as in EGFR) would restore the water network, and the affinity of lapatinib for ErbB4 would improve. As expected, the simple swap at position 775 indeed restores the bridging water at S1 as illustrated visually by a comparison of the patterns before and after hydration (Figure 10g vs Figure 10h). Further, the ErbB4 V775C patterns (Figure 10h) are remarkably similar to that of wild-type EGFR (Figure 10a) or wild-type HER2 (Figure 10e). Quantitatively (Table 6), water counts for the ErbB4 V775C mutant localized to site 775 show an increase (0 → 0.80) that yields remarkable accord with values observed in wild-type EGFR (0.71) or HER2 (0.85) simulations. Good agreement is also seen for the increase occurring at position 854: ErbB4 V775C (0.11 → 0.97) versus EGFR (0.89) and HER2 (0.76). Ligand−water values also increase (1.65 → 2.57) to values that are similar to that observed in the EGFR (2.30) and HER2 (2.15) wild-type systems. Only minor changes are observed at positions 790 and 791, which make physical sense, given that occupancies at site S2 are not expected to be as drastically altered by this mutation.

Energetically, it is gratifying that the accompanying ΔG_b(calcd) values for the ErbB4 V775C mutant also show favorable increases relative to that of wild-type ErbB4 (Table 6h vs Table 6g) as originally hypothesized, using either the 3WAT or 1WAT protocol, and values for the ErbB4 V775C mutant (−69.21 kcal/mol for 3WAT and −66.18 kcal/mol for 1WAT) are close to that of wild-type EGFR, which the V → C mutation was designed to mimic (−68.80 kcal/mol for 3WAT and −65.84 kcal/mol for 1WAT). Overall, the simulation results strongly suggest that a valine at position 775 is a primary factor that contributes to the selectivity of lapatinib for EGFR and HER2 relative to ErbB4.

EGFR C775F. Avizienyte et al.[14] proposed that a direct steric clash with the lapatinib aniline group was a likely cause of reduced activity with the EGFR C775F variant identified using in vitro screening. However, on the basis of simulation results for this mutant, positional sampling of lapatinib in the binding site appears similar to that in wild-type EGFR, which suggests unfavorable intermolecular interactions are not introduced as a result of the amino acid swap. In fact, a comparison of protein−ligand van der Waals interaction energies, localized to position 775, actually shows enhanced interactions for the C775F mutant relative to those of the wild type (−3.24 kcal/mol for EGFR C775F vs −1.56 kcal/mol for the wild type), which indicates a clash at this position is unlikely. This is not

dx.doi.org/10.1021/bi2016553 | Biochemistry 2012, 51, 2390−2406

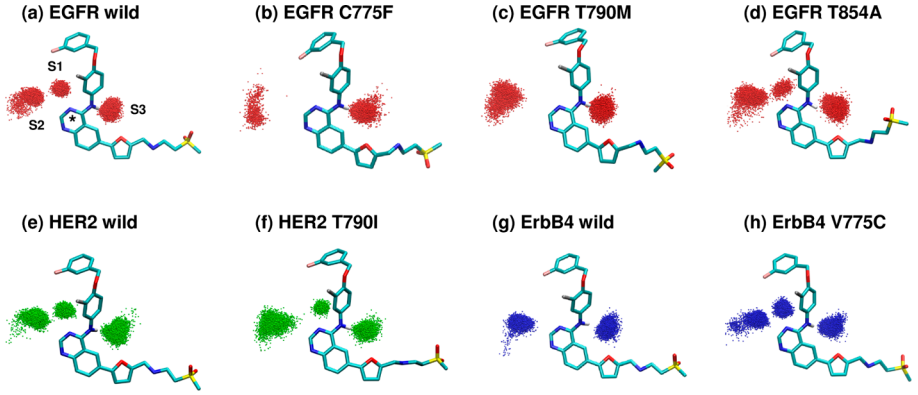


Figure 10. Water-mediated H-bonding patterns (sites S1−S3) for lapatinib with wild-type and mutant forms of EGFR (red), HER2 (green), and ErbB4 (blue). Images show overlaid waters (O atoms) within 3 Å of ligand atom N* (S1), kinase backbone oxygen at residue 791 (S2), and ligand aniline hydrogen (S3) from 10000 evenly spaced MD snapshots (20 ns trajectories).

Table 6. Average Numbers of Water-Mediated H-Bonds and Binding Free Energies for Lapatinib with Wild-Type and Mutant Kinases

residue (wild type)	(a) EGFR wild[a]	(b) EGFR C775F[b]	(c) EGFR T790M[b]	(d) EGFR T854A[b]	(e) HER2 wild[a]	(f) HER2 T790I[b]	(g) ErbB4 wild[a]	(h) ErbB4 V775C[b]
				H-Bonds with Water (counts)[c]				
ligand (N*)	0.88	0	0	0.22	0.75	0.54	0.11	0.95
ligand	2.30	1.37	1.40	1.76	2.15	2.29	1.65	2.57
CSV 775	0.71	0	0.10	0.22	0.85	0.36	0	0.80
TTT 790	0.76	0.09	0.06	0.87	0.90	0	0.77	0.75
QQQ 791	1.37	0.27	0.99	1.05	1.13	0.62	0.98	0.91
TTT 854	0.89	0	0	0.03	0.76	0.54	0.11	0.97
sum (exclude N*)	6.03	1.73	2.55	3.93	5.79	3.81	3.51	6.00
				ΔG_b(calcd) (kcal/mol)[d]				
3WAT protocol	−68.80	−67.61	−67.15	−66.60	−67.82	−65.25	−67.19	−69.21
1WAT protocol	−65.84	−63.56	−63.28	−62.57	−64.53	−62.10	−63.35	−66.18

[a]Wild-type values averaged from six MD runs (120000 frames). [b]Mutation values averaged from one MD run (20000 frames). [c]Averaged H-bond counts employ the three waters closest to N* (defined in Table 1) with N*, the entire ligand, or residue 775, 790, 791, or 854. [d]Binding energies computed using the 3WAT protocol (three closest waters to N*) or the 1WAT protocol (one water if it is within 3 Å of N*). Simulations based on template 1XKK (see Methods).

unexpected as the related C to V swap is tolerated at this position in ErbB4.

What does dramatically change, as a result of the increased hydrophobicity and bulk of phenylalanine relative to cysteine, is the water occupancy at both sites S1 and S2 (Figure 10b and Table 6b). Relative to the wild type, the EGFR C775F mutant (Table 6a vs Table 6b) leads to the total or nearly total loss of H-bonding with F775 (0.71 to 0), T790 (0.76 to 0.09), T854 (0.89 to 0), or the ligand N* (0.88 to 0). Counts at Q791 (1.37 to 0.27) or with the total ligand (2.30 to 1.37) are also significantly reduced. Decreases in overall computed binding energies between 1.19 (3WAT) and 2.28 kcal/mol (1WAT) again suggest a resistance mechanism involving network waters.

Interestingly, the EGFR C775F mutation was identified only in lapatinib in vitro resistance screenings but not in erlotinib.[14] Although differences in bound inactive versus active EGFR could be a factor, an examination of EGFR crystal structures with lapatinib (1XKK) and erlotinib (1M17) did not reveal any obvious reason why, in principle, C775F could not negatively impact erlotinib. Structurally, an EGFR C775F mutant should negatively affect S1 water occupancy in both active and inactive forms. It should be noted that EGFR T790M and EGFR T854A do negatively affect both lapatinib and erlotinib.[14] Although not yet reported, the related EGFR C775V mutation would also be expected to be detrimental to both classes of inhibitors.

The analogous mutation for HER2 is S775P. Trowe et al.[15] suggest this mutation may act by direct steric interference, by analogy to the spatially homologous V299L mutation associated with resistance to imatinib in BCR-ABL. However, the V299L mutation in this system involves a nonpolar-to-nonpolar change that would not directly affect changes in H-bonding with water in contrast to the S775P polar-to-nonpolar change introduced into HER2. Thus, while a steric mechanism might be appropriate for describing the effects of the V299L mutation in BCR-ABL, for lapatinib with HER2 S775P, a water-based mechanism is more likely. Additional simulations would be needed to more fully explore this issue.

EGFR T790M. The gatekeeper mutation T790M, associated with clinical resistance to gefitnib and erlotinib with EGFR,[26,27] has also been identified as a mutation affecting lapatinib.[14] Although it has been suggested[25] that the T790M mutation sterically hinders binding of lapatinib to EGFR, again our results indicate changes involving water are much more likely. Results here, from simulations of EGFR T790M, reveal that H-bond counts localized to M790 (0.06), the ligand N* (0), or T854 (0) are essentially zero relative to the wild-type counts (Table 6c vs Table 6a). The increased bulk in going from T to M and the fact that sulfur is a weaker H-bond acceptor than

dx.doi.org/10.1021/bi2016553 | Biochemistry 2012, 51, 2390−2406

oxygen are the most likely causes.[39] Visually, the accompanying T790M graphic (Figure 10c) show S1 waters are no longer present, which is reflected well in the total H-bond counts with ligand, which goes from 2.30 to 1.40.

Using analysis similar to that reported here, Balius et al.[39] showed that mutations at the T790 site (and T854 as described below) would disrupt water-mediated binding. The prior study not only resolved the ambiguity surrounding which residue (T790 or T854) was primarily involved in water-mediated interactions (both are) but provided convincing quantitative evidence of why a T790M steric clash mechanism is unlikely. Given the positional similarity for high-occupancy water sites in both active and inactive kinase forms, and on the basis of our results that show disruption of the water network, lapatinib resistance due to EGFR T790M in the inactive form is also not likely driven by a steric clash. The favorable increase in van der Waals energy observed at the site of the mutation (−5.85 kcal/mol for EGFR T790M vs −3.56 kcal/mol for the wild type) reaffirms this conclusion. The fact that the mutant also yields less favorable computed binding energies [ΔG_b(calcd)] relative to that of the wild type (Table 6c vs Table 6a), with losses of between 1.65 and 2.56 kcal/mol, when solvent is included (3WAT and 1WAT protocols, respectively), provides additional evidence of a water-based mechanism.

HER2 T790I. The analogous gatekeeper mutation in HER2 is T790I, which was the most frequent mutation identified from in vitro screening by Trowe et al.[15] and is correlated with a high level of lapatinib resistance. Although direct steric interference was again suggested,[15] on the basis of our computational results, a water-mediated mechanism is more likely. In agreement with results from the EGFR T790M simulations, occupancy at S1 in the HER2 T790I mutant becomes reduced and the increase in "favorable" van der Waals interactions occurring at residue 790 (−3.72 kcal/mol for the mutant vs −3.67 kcal/mol for the wild type) indicate no clash. Quantitatively, counts for S775, T790, and T854 become reduced. As expected, the reduced S1 occupancy also affects S2 waters that are reflected in reduced counts at Q791. Interestingly, counts with the total ligand slightly increase, which is somewhat counterintuitive but could be a function of slight differences occurring at, for example, the solvent-exposed methyl sulfone tail. In any event, the reduction at the key ligand N* atom (from 0.75 to 0.54) indicates a lack of S1 water, which is consistent with a water-based resistance mechanism.

EGFR T854A. The final mutation studied, EGFR T854A, can confer resistance both to lapatinib and to erlotinib,[14,28] which for the latter case was identified from a patient with lung adenocarcinoma.[28] Possible resistance mechanisms proposed by Bean et al.[28] for erlotinib include a loss of contact with the inhibitor, altered specificity for ATP, or conformational changes in the protein. Balius et al.[39] hypothesized mutations at this site could disrupt erlotinib binding through alteration of the water-mediated network. For lapatinib, our results suggest both a contact and a water-based mechanism of T854A resistance in the inactive kinase system. Per-residue decomposition shows that van der Waals interactions, localized to position 854, become less favorable in simulations of the EGFR T854A mutant (−2.61 kcal/mol) relative to the wild type (−4.29 kcal/mol). Water H-bond counts with nearby residues also become reduced (Table 6d). For example, counts with N* (0.88 → 0.22), the ligand (2.30 → 1.76), residue 775 (0.71 → 0.22), residue 791 (1.37 → 1.05), and residue 854 (0.89 → 0.03) are all lower than that of the wild type. Interestingly, a slight increase is observed at residue 790 (0.76 → 0.87), which is counterintuitive. A possible explanation involves the fact that, compared to the other mutations studied, the T854A mutation involves a swap to a smaller amino acid. As a result, although waters may be less favorably accommodated, they are not sterically blocked from site S1 (see Figure 10d). This could allow water to more freely occupy site S2, thereby increasing the possibility of it interacting favorably with T790 and thus explaining the observed increase. In any event, from a quantitative standpoint (Table 6), summed H-bond counts are significantly reduced in simulations of T854A (3.93) versus the wild type (6.03).

Future Directions. These results suggest additional studies (both computational and experimental) that could be used to further characterize molecular recognition in these systems. For example, the prediction made here that ErbB4 V775C would improve lapatinib binding was verified computationally; however, the mutation should be tested experimentally. The related mutation, ErbB4 V775S, as a surrogate for HER2, should also be examined. Other avenues include the use of more quantitative computational methods (free energy perturbation and thermodynamic integration) and alternative force fields (polarizable and quantum-based) to probe the energetic effects of including bound waters at specific locations in the binding site or investigating resistance mutations in finer detail. Additional mutations to study, which are induced by lapatinib, include L747S, R776P, L777Q, L788I/V, K860T, G863S, and R889S.[14]

A growing body of evidence also suggests that the use of kinase domain inhibitors leads to upregulation of a fourth ErbB member, HER3, which in turn causes drug resistance through alternative signaling pathways.[79−81] Thus, HER3 could be considered as a potential therapeutic target. Interestingly, HER3 has historically been classified as an inactive kinase, although recent studies[82] indicate there is weak catalytic activity that could be clinically relevant. On the basis of our simulation results, and comparison with available inactive form structures of HER3,[82,83] we hypothesize that lapatinib would bind HER3 only weakly. Similar to ErbB4, the HER3 catalytic domain contains a valine at position 775 and thus should have weaker water-mediated H-bonding ability compared to EGFR or HER2. Weaker binding and weaker catalytic activity could both contribute to HER3 upregulation. Additional simulations to assess the binding of lapatinib to HER3 are planned for the future.

Finally, from the point of view of development, strategies for achieving improved ErbB inhibitors include modifications that yield tighter binding to wild-type enzymes, restore binding lost as a result of resistance mutations, or both. These results suggest that exploiting differences in the number of binding site waters could be important. As a first step, we performed simulations involving a straightforward modification of lapatinib at the N* position from N* to C* (Figure 11a). We hypothesized the swap would yield a lower affinity for wild-type EGFR as a result of water occupancy becoming reduced at site S1 caused by a lack of H-bonding capability at C*. Indeed, a simulation of the analogue with wild-type EGFR revealed a lower water occupancy at site S1 and a less favorable ΔG_b(calcd) value of −63.59 kcal/mol (analogue) versus a value of −65.84 kcal/mol (lapatinib) computed via the 1WAT protocol (Figure 11a) from a single representative run.

On the other hand, we reasoned the analogue would not be as detrimentally affected as lapatinib by mutations that alter S1

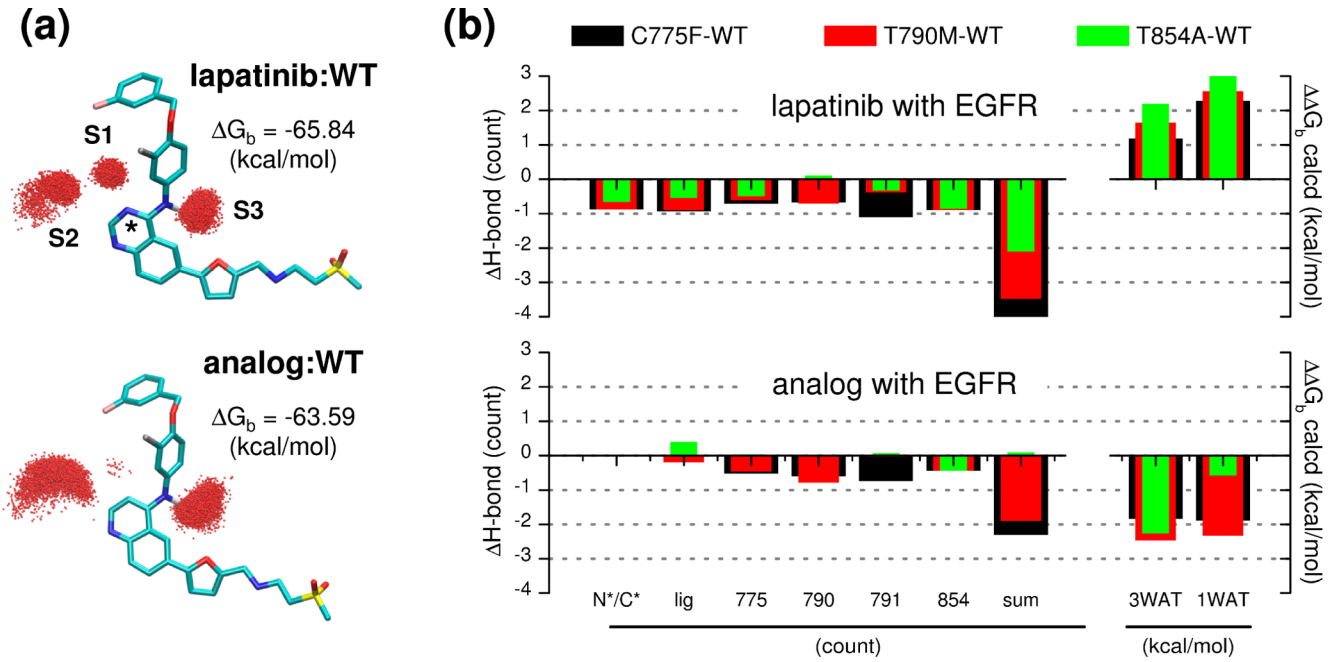


Figure 11. (a) Hydration patterns (sites S1−S3) and absolute binding energies [ΔG_b(calcd)] from wild-type EGFR simulations of lapatinib (top) and a N* → C* analogue (bottom). Images show overlaid waters (O atoms) within 3 Å of ligand atom N*/C* (S1), the kinase backbone oxygen at residue 791 (S2), and the ligand aniline hydrogen (S3) from 10000 evenly spaced MD snapshots (20 ns trajectories). (b) ΔH-bond and $\Delta\Delta G_b$(calcd) values (mutant − wild type) from simulations of wild-type EGFR and C775F, T790M, and T854A mutants with lapatinb (top) and a N* → C* analogue (bottom). Relative H-bonds (counts) computed among the three waters closest to N*/C* with N*/C*, the entire ligand, residues 775, 790, 791, and 854, and the overall sum excluding N*/C*. Relative binding free energies (kilocalories per mole) computed using the 3WAT (three waters closest to N*/C*) or 1WAT protocol (one water if it is within 3 Å of N*/C*).

occupancy. As illustrated in Figure 11b, which shows fold resistance profiles for both ligands, additional simulations of C775F, T854A, and T790M mutants with the analogue confirmed the hypothesis. In all three cases (C775F − WT, T790M − WT, and T854A − WT), and across both calculation protocols (3WAT and 1WAT), simulations with lapatinib yielded unfavorable fold resistance values [mutant minus wild-type ΔG_b(calcd)] in contrast to the analogue that interestingly showed enhanced binding (Figure 11b, bottom vs top graphs). The increase in binding affinity computed for the analogue with the mutants appears to involve overall more favorable hydrophobic (ΔE_{vdw} + $\Delta G_{nonpolar}$) interactions for T790M and T854A (Figure 11b, red and green bars) and a combination of more favorable hydrophobic and hydrophilic (ΔE_{coul} + ΔG_{polar}) interactions for C775F (Figure 11b, black bar). As expected, the simulations of lapatinib showed losses larger than those for the analogue for H-bonding of key waters in the pocket (three waters closest to N*/C*) with the ligand at N*/C*, the entire ligand, residues 775, 790, 791, and TTT 854, and the overall sum (Figure 11b, top vs bottom). Provided sufficient affinity to wild-type domains could be achieved, analogues that rely less on water at site S1 could have utility in combination therapies given the potential for orthogonal resistance profiles. This strategy might also benefit the development of HER3 inhibitors given the expected lower water occupancy at site S1 due to the valine at position 775. The design of analogues capable of displacing water at site S1, while mimicking the network seen between lapatinib and EGFR or HER2, should also be pursued.

CONCLUSION

In this report, all-atom molecular dynamics, free energies of binding, and energy decomposition analyses were performed for the kinase inhibitor lapatinib in complex with wild-type EGFR, HER2, and ErbB4, and relevant mutants, including C775F, T854A, T790M (EGFR), T790I (HER2), and V775C (ErbB4). The primary goal was to develop robust computational models consistent with experimental activity data (Table 1) and determine how variation in receptor sequence (Figure 1) and structure contributes to binding specificity and drug resistance. The simulations employed both crystallographic and crystallographically derived homology models. Convergence and stability were carefully evaluated using multiple MD runs (N = 6) for each wild-type system for which rmsds (Figure 4), autocorrelation functions (Figure 5), and block-averaged standard errors of the mean (Figure 5 and Table 3) demonstrate good system behavior. The remarkable numerical agreement obtained using homology models, derived from one of two different crystallographic templates, is particularly notable (Figure 6 and Table 4).

Analysis of MD trajectories with explicit solvent revealed three high-occupancy sites for water termed S1, S2, and S3 (see Figure 7) that mediate ligand binding in EGFR and HER2. In contrast, ErbB4 showed low occupancy at site S1. The results suggested hydration differences could play a role in determining receptor specificity, and subsequent free energy calculations in which explicit waters were included confirmed the hypothesis. Relative free energies of binding [$\Delta\Delta G_b$(calcd)], using template-averaged data, in which three key waters (3WAT protocol) or one key water (1WAT protocol) is included in the calculations, correctly predict the experimental trend: EGFR > HER2 > ErbB4 (Table 4).

dx.doi.org/10.1021/bi2016553 | *Biochemistry* 2012, 51, 2390−2406

Calculations in which water was not included (0WAT protocol) incorrectly predict the order. Quantitatively, the 1WAT protocol leads to excellent numerical agreement for $\Delta\Delta G_b$: (i) EGFR − HER2 = −0.76 for calcd versus −0.87 for exptl, (ii) HER2 − ErbB4 = −1.37 for calcd versus −1.94 for exptl, and (iii) EGFR − ErbB4 = −2.13 for calcd versus −2.81 for exptl, which underscores the importance of hydration and in particular site S1.

Examination of the underlying ΔG_b(calcd) energy components with (3WAT and 1WAT) or without (0WAT) explicit waters reveals that electrostatic contributions (ΔE_{coul} and ΔG_{polar}) change much more dramatically than the accompanying steric packing terms (ΔE_{vdw} and $\Delta G_{nonpolar}$), especially for EGFR and HER2 (Table 5). Smaller changes for ErbB4 are a function of occupancy of site S1 and the primary reason why solvated protocols yield the correct experimental ordering. Occupancies at site S1 appear to be a function of sequence changes at position 775 (Figure 9a) involving a swap from C or S (EGFR and HER2) to V (ErbB4), for which the side chain is bulkier, is hydrophobic, and has no H-bonding capability. Otherwise, no specific residues that might explain specificity through direct modulation of protein−ligand interactions were identified (Figure 8). The fact that ΔG_{polar}, and not ΔE_{coul} or ΔE_{vdw}, most closely tracks with experiment provides additional support for the suggestion that selectivity is a function of differential hydration in the bound state (Table 5).

At position 775, an important quadrifurcated H-bonding network that involves the ligand, S1 and S2 waters, and residues CSV 775, TTT 854, TTT 790, and QQQ 791 was identified (Figure 9a). A similar network was observed in an earlier study of active form EGFR.[39] Quantitatively, water H-bond counts and Coulombic interactions with species in the network are reduced by ∼40% for ErbB4 relative to EGFR or HER2 (Figure 9b). Water H-bond counts with the ligand N* atom or the total ligand compellingly track with the experimental order: EGFR > HER2 > ErbB4 (Figure 9b). The nearly identical counts across all three systems, for waters interacting with TTT 854 versus the ligand N* atom, confirm a highly coupled S1 site (Figure 9b). Other groups[27,31,34,36,40] have also discussed the importance of binding site waters in these and related systems, although to the best of our knowledge, differential hydration as the primary mechanism for ErbB family specificity has not yet been proposed. Additional studies to more precisely quantify the energetic impact of including water molecules in these sites would be worthwhile.

To determine if drug resistance could also be influenced by changes in water-mediated binding, a series of additional simulations were performed (Figure 10 and Table 6) for lapatinib with mutant kinases identified by in vitro screening. In good qualitative agreement with experiment, for four deleterious mutations studied (EGFR C775F, EGFR T854A, EGFR T790M, and HER2 T790I), computed free energies of binding using solvated ΔG_b(calcd) protocols become less favorable relative to that of the wild type (Table 6). In all cases, S1 occupancies are significantly reduced, which suggests specificity and resistance share a common mechanism (Table 6). In support of this argument, results from a hypothetical mutant (ErbB4 V775C), designed to restore S1 hydration, show a remarkable increase in S1 occupancy and a favorable increase in computed binding free energy (Table 6). The computational results also provide a testable prediction for future experimental work. Importantly, in contrast to previously proposed steric clash mechanisms for EGFR C775F,[14] EGFR T790M,[25] and HER2 T790I,[15] our results indicate a water-based mechanism of resistance. This conclusion is based on several observations. (i) Visualization of MD trajectories and ligand rmsds reveals no identifiable changes in sampling suggestive of a steric clash. (ii) Localized per-residue interactions show favorable increases in the favorable van der Waals energy at the sites of mutation. (iii) Water occupancies at site S1 become significantly reduced. (iv) H-Bond counts are lower, and Coulombic interactions are less favorable for binding site waters with nearby species. (v) Solvated ΔG_b(calcd) protocols yield less favorable energies. Somewhat differently, for EGFR T854A, the results are consistent with both a water-based mechanism[39] and a previously proposed loss-of-contact mechanism.[28]

An improved understanding of molecular factors that drive ligand binding for important anticancer targets such as EGFR and HER2 will ultimately allow the design of improved drugs with greater clinical utility. This work has contributed to this goal through the development of robust simulation models for lapatinib with inactive form kinases and by providing a physically reasonable water-based mechanism for understanding what drives ErbB family specificity and resistance. The simulations demonstrate that in these systems water is as important as any specific residue interaction for binding. The fact that the affinity of lapatinib for ErbB4 is disfavored because of a reduction in binding site water is, perhaps fortuitously, likely to be of beneficial outcome.[22,23] Future work could try to exploit this mechanism to further enhance selectivity as a result of hydration differences at site S1. However, analogues that further rely on the water network would also be vulnerable to mutations that alter the network. Alternatively, as suggested by the results from simulations of a lapatinib analogue (Figure 11), inhibitors that do not rely as strongly on water may be less vulnerable. The observation that the network can be altered in different ways (i.e., steric blockage of water, reduced H-bonding capability with water, or both) suggests that orthogonal binding motifs will likely be required to effectively combat drug resistance in the long term.

AUTHOR INFORMATION

Corresponding Author

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

Funding

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, Walk-for-Beauty Foundation Research Award).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Gratitude is expressed to Trent E. Balius and Sudipto Mukherjee for advice and computational assistance and to Kenneth Foreman and William J. Allen for helpful discussions. This research utilized resources at the New York Center for Computational Sciences at Stony Brook University/Brookhaven National Laboratory, which is supported by the U.S. Department of Energy under Contract DE-AC02-98CH10886 and by the State of New York.

ABBREVIATIONS

ACF, autocorrelation function; BASEM, block-averaged standard errors of the mean; BCR-ABL, tyrosine kinase drug target for chronic myelogenous leukemia; EGFR, human epidermal growth factor receptor 1; ErbB4, human epidermal growth factor receptor 4; HER2, human epidermal growth factor receptor 2; MD, molecular dynamics; MM-GBSA, Molecular Mechanics Generalized Born Surface Area; PDB, Protein Data Bank; rmsd, root-mean-square deviation.

REFERENCES

(1) American Cancer Society: Breast Cancer Facts & Figures 2011–2012. http://www.cancer.org (accessed Oct 26, 2011).

(2) Yarden, Y., and Sliwkowski, M. X. (2001) Untangling the ErbB signalling network. *Nat. Rev. Mol. Cell Biol. 2*, 127–137.

(3) Hynes, N. E., and Lane, H. A. (2005) ERBB receptors and cancer: The complexity of targeted inhibitors. *Nat. Rev. Cancer 5*, 341–354.

(4) Klijn, J. G. M., Berns, P. M. J. J., Schmitz, P. I. M., and Foekens, J. A. (1992) The Clinical Significance of Epidermal Growth Factor Receptor (EGF-R) in Human Breast Cancer: A Review on 5232 Patients. *Endocr. Rev. 13*, 3–17.

(5) Pegram, M. D., Pauletti, G., and Slamon, D. J. (1998) HER-2/neu as a predictive marker of response to breast cancer therapy. *Breast Cancer Res. Treat. 52*, 65–77.

(6) Drugs@FDA website. http://www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm (accessed Oct 26, 2011).

(7) Konecny, G. E., Pegram, M. D., Venkatesan, N., Finn, R., Yang, G., Rahmeh, M., Untch, M., Rusnak, D. W., Spehar, G., Mullin, R. J., Keith, B. R., Gilmer, T. M., Berger, M., Podratz, K. C., and Slamon, D. J. (2006) Activity of the dual kinase inhibitor lapatinib (GW572016) against HER-2-overexpressing and trastuzumab-treated breast cancer cells. *Cancer Res. 66*, 1630–1639.

(8) Blackwell, K. L., Pegram, M. D., Tan-Chiu, E., Schwartzberg, L. S., Arbushites, M. C., Maltzman, J. D., Forster, J. K., Rubin, S. D., Stein, S. H., and Burstein, H. J. (2009) Single-agent lapatinib for HER2-overexpressing advanced or metastatic breast cancer that progressed on first- or second-line trastuzumab-containing regimens. *Ann. Oncol. 20*, 1026–1031.

(9) Xia, W., Bacus, S., Hegde, P., Husain, I., Strum, J., Liu, L., Paulazzo, G., Lyass, L., Trusk, P., Hill, J., Harris, J., and Spector, N. L. (2006) A model of acquired autoresistance to a potent ErbB2 tyrosine kinase inhibitor and a therapeutic strategy to prevent its onset in breast cancer. *Proc. Natl. Acad. Sci. U.S.A. 103*, 7795–7800.

(10) Liu, L., Greger, J., Shi, H., Liu, Y., Greshock, J., Annan, R., Halsey, W., Sathe, G. M., Martin, A.-M., and Gilmer, T. M. (2009) Novel Mechanism of Lapatinib Resistance in HER2-Positive Breast Tumor Cells: Activation of AXL. *Cancer Res. 69*, 6871–6878.

(11) Eichhorn, P. J. A., Gili, M., Scaltriti, M., Serra, V., Guzman, M., Nijkamp, W., Beijersbergen, R. L., Valero, V., Seoane, J., Bernards, R., and Baselga, J. (2008) Phosphatidylinositol 3-Kinase Hyperactivation Results in Lapatinib Resistance that Is Reversed by the mTOR/Phosphatidylinositol 3-Kinase Inhibitor NVP-BEZ235. *Cancer Res. 68*, 9221–9230.

(12) Huang, C., Park, C. C., Hilsenbeck, S. G., Ward, R., Rimawi, M. F., Wang, Y. C., Shou, J., Bissell, M. J., Osborne, C. K., and Schiff, R. (2011) β1 integrin mediates an alternative survival pathway in breast cancer cells resistant to lapatinib. *Breast Cancer Res. 13*, R84.

(13) Rexer, B. N., Ham, A. J. L., Rinehart, C., Hill, S., de Matos Granja-Ingram, N., Gonzalez-Angulo, A. M., Mills, G. B., Dave, B., Chang, J. C., Liebler, D. C., and Arteaga, C. L. (2011) Phosphoproteomic mass spectrometry profiling links Src family kinases to escape from HER2 tyrosine kinase inhibition. *Oncogene 30*, 4163–4174.

(14) Avizienyte, E., Ward, R. A., and Garner, A. P. (2008) Comparison of the EGFR resistance mutation profiles generated by EGFR-targeted tyrosine kinase inhibitors and the impact of drug combinations. *Biochem. J. 415*, 197–206.

(15) Trowe, T., Boukouvala, S., Calkins, K., Cutler, R. E., Fong, R., Funke, R., Gendreau, S. B., Kim, Y. D., Miller, N., Woolfrey, J. R., Vysotskaia, V., Yang, J. P., Gerritsen, M. E., Matthews, D. J., Lamb, P., and Heuer, T. S. (2008) EXEL-7647 Inhibits Mutant Forms of ErbB2 Associated with Lapatinib Resistance and Neoplastic Transformation. *Clin. Cancer Res. 14*, 2465–2475.

(16) Rusnak, D. W., Lackey, K., Affleck, K., Wood, E. R., Alligood, K. J., Rhodes, N., Keith, B. R., Murray, D. M., Knight, W. B., Mullin, R. J., and Gilmer, T. M. (2001) The Effects of the Novel, Reversible Epidermal Growth Factor Receptor/ErbB-2 Tyrosine Kinase Inhibitor, GW2016, on the Growth of Human Normal and Tumor-derived Cell Lines in Vitro and in Vivo. *Mol. Cancer Ther. 1*, 85–94.

(17) 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.

(18) 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.

(19) 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.

(20) 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.

(21) Lackey, K. E. (2006) Lessons from the drug discovery of lapatinib, a dual ErbB1/2 tyrosine kinase inhibitor. *Curr. Top. Med. Chem. 6*, 435–460.

(22) Sartor, C. I., Zhou, H., Kozlowska, E., Guttridge, K., Kawata, E., Caskey, L., Harrelson, J., Hynes, N., Ethier, S., Calvo, B., and Earp, H. S. III (2001) HER4 Mediates Ligand-Dependent Antiproliferative and Differentiation Responses in Human Breast Cancer Cells. *Mol. Cell. Biol. 21*, 4265–4275.

(23) Barnes, N. L., Khavari, S., Boland, G. P., Cramer, A., Knox, W. F., and Bundred, N. J. (2005) Absence of HER4 expression predicts recurrence of ductal carcinoma in situ of the breast. *Clin. Cancer Res. 11*, 2163–2168.

(24) Qiu, C., Tarrant, M. K., Choi, S. H., Sathyamurthy, A., Bose, R., Banjade, S., Pal, A., Bornmann, W. G., Lemmon, M. A., Cole, P. A., and Leahy, D. J. (2008) Mechanism of Activation and Inhibition of the HER4/ErbB4 Kinase. *Structure 16*, 460–467.

(25) Gilmer, T. M., Cable, L., Alligood, K., Rusnak, D., Spehar, G., Gallagher, K. T., Woldu, E., Carter, H. L., Truesdale, A. T., Shewchuk, L., and Wood, E. R. (2008) Impact of Common Epidermal Growth Factor Receptor and HER2 Variants on Receptor Activity and Inhibition by Lapatinib. *Cancer Res. 68*, 571–579.

(26) 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.

(27) Kobayashi, S., Boggon, T. J., Dayaram, T., Jänne, 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.

(28) 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.

(29) 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.
(30) 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.
(31) 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.
(32) Kamath, S., and Buolamwini, J. (2005) Asp746 to Glycine Change May have a Greater Influence than Cys751 to Serine Change in Accounting for Ligand Selectivity between EGFR and HER-2 at the ATP Site. *J. Comput.-Aided Mol. Des. 19*, 287−291.
(33) Scaltriti, M., Verma, C., Guzman, M., Jimenez, J., Parra, J. L., Pedersen, K., Smith, D. J., Landolfi, S., Ramon y Cajal, S., Arribas, J., and Baselga, J. (2008) Lapatinib, a HER2 tyrosine kinase inhibitor, induces stabilization and accumulation of HER2 and potentiates trastuzumab-dependent cell cytotoxicity. *Oncogene 28*, 803−814.
(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) Chen, H. F. (2008) Computational study of the binding mode of epidermal growth factor receptor kinase inhibitors. *Chem. Biol. Drug Des. 71*, 434−446.
(36) Cavasotto, C. N., Ortiz, M. A., Abagyan, R. A., and Piedrafita, F. J. (2006) In silico identification of novel EGFR inhibitors with antiproliferative activity against cancer cells. *Bioorg. Med. Chem. Lett. 16*, 1969−1974.
(37) 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.
(38) 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.
(39) Balius, T. E., and Rizzo, R. C. (2009) Quantitative Prediction of Fold Resistance for Inhibitors of EGFR. *Biochemistry 48*, 8435−8448.
(40) Michel, J., Tirado-Rives, J., and Jorgensen, W. L. (2009) Energetics of Displacing Water Molecules from Protein Binding Sites: Consequences for Ligand Optimization. *J. Am. Chem. Soc. 131*, 15403−15411.
(41) Daub, H., Specht, K., and Ullrich, A. (2004) Strategies to overcome resistance to targeted protein kinase inhibitors. *Nat. Rev. Drug Discovery 3*, 1001−1010.
(42) Cozzini, P., Fornabaio, M., Marabotti, A., Abraham, D. J., Kellogg, G. E., and Mozzarelli, A. (2004) Free energy of ligand binding to protein: Evaluation of the contribution of water molecules by computational methods. *Curr. Med. Chem. 11*, 3093−3118.
(43) Lu, Y., Wang, R., Yang, C.-Y., and Wang, S. (2007) Analysis of Ligand-Bound Water Molecules in High-Resolution Crystal Structures of Protein-Ligand Complexes. *J. Chem. Inf. Model. 47*, 668−675.
(44) Young, T., Abel, R., Kim, B., Berne, B. J., and Friesner, R. A. (2007) Motifs for molecular recognition exploiting hydrophobic enclosure in protein−ligand binding. *Proc. Natl. Acad. Sci. U.S.A. 104*, 808−813.
(45) de Beer, S. B. A., Vermeulen, N. P. E., and Oostenbrink, C. (2010) The role of water molecules in computational drug design. *Curr. Top. Med. Chem. 10*, 55−66.
(46) Varghese, J. N., Epa, V. C., and Colman, P. M. (1995) Three-dimensional structure of the complex of 4-guanidino-Neu5Ac2en and influenza virus neuraminidase. *Protein Sci. 4*, 1081−1087.
(47) Ren, J., Esnouf, R., Garman, E., Somers, D., Ross, C., Kirby, I., Keeling, J., Darby, G., Jones, Y., Stuart, D., et al. (1995) High resolution structures of HIV-1 RT from four RT-inhibitor complexes. *Nat. Struct. Biol. 2*, 293−302.
(48) Wood, E. R., Shewchuk, L. M., Ellis, B., Brignola, P., Brashear, R. L., et al. (2008) 6-Ethynylthieno[3,2-d]- and 6-ethynylthieno[2,3-d]pyrimidin-4-anilines as tunable covalent modifiers of ErbB kinases. *Proc. Natl. Acad. Sci. U.S.A. 105*, 2773−2778.
(49) Aertgeerts, K., Skene, R., Yano, J., Sang, B. C., Zou, H., Snell, G., Jennings, A., Iwamoto, K., Habuka, N., Hirokawa, A., Ishikawa, T., Tanaka, T., Miki, H., Ohta, Y., and Sogabe, S. (2011) Structural analysis of the mechanism of inhibition and allosteric activation of the kinase domain of HER2 protein. *J. Biol. Chem. 286*, 18756−18765.
(50) Sali, A., and Blundell, T. L. (1993) Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol. 234*, 779−815.
(51) Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) Clustal W: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Res. 22*, 4673−4680.
(52) Shen, M.-Y., and Sali, A. (2006) Statistical potential for assessment and prediction of protein structures. *Protein Sci. 15*, 2507−2524.
(53) Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr. 26*, 283−291.
(54) *MOE* (2008) Chemical Computing Group Inc., Montreal.
(55) 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.
(56) Case, D. A., Darden, T. A., Cheatham, T. E., III, Simmerling, C. L., Wang, J., et al. (2004) *AMBER8*, University of California, San Francisco.
(57) 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: Struct., Funct., Bioinf. 65*, 712−725.
(58) 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.
(59) 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.
(60) 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.
(61) Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., et al. (1998) *Gaussian 98*, revision A.9, Gaussian Inc., Pittsburgh, PA.
(62) Feller, S. E., Zhang, Y., Pastor, R. W., and Brooks, B. R. (1995) Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys. 103*, 4613−4621.
(63) 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. Comput. Phys. 23*, 327−341.
(64) Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N log(N) method for Ewald sums in large systems. *J. Chem. Phys. 98*, 10089−10092.
(65) Srinivasan, J., Cheatham, T. E. III, Cieplak, P., Kollman, P. A., and Case, D. A. (1998) Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate−DNA helices. *J. Am. Chem. Soc. 120*, 9401−9409.
(66) Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S. H., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D. A., and Cheatham, T. E. III (2000) Calculating structures and free energies of complex molecules:

Combining molecular mechanics and continuum models. *Acc. Chem. Res. 33*, 889−897.
(67) Strockbine, B., and Rizzo, R. C. (2007) Binding of antifusion peptides with HIVgp41 from molecular dynamics simulations: Quantitative correlation with experiment. *Proteins: Struct., Funct., Bioinf. 67*, 630−642.
(68) 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. Theory Comput. 4*, 1526−1540.
(69) 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.
(70) McGillick, B. E., Balius, T. E., Mukherjee, S., and Rizzo, R. C. (2010) Origins of Resistance to the HIVgp41 Viral Entry Inhibitor T20. *Biochemistry 49*, 3575−3592.
(71) 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: Struct., Funct., Bioinf. 55*, 383−394.
(72) 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.
(73) Grossfield, A., and Zuckerman, D. M. (2009) Quantifying uncertainty and sampling quality in biomolecular simulations. *Annu. Rep. Comput. Chem. 5*, 23−48.
(74) Hess, B. (2002) Determining the shear viscosity of model liquids from molecular dynamics simulations. *J. Chem. Phys. 116*, 209−217.
(75) Spackova, N., Cheatham, T. E. III, Ryjacek, F., Lankas, F., Van Meervelt, L., Hobza, P., and Sponer, J. (2003) Molecular dynamics simulations and thermodynamics analysis of DNA-drug complexes. Minor groove binding between 4′,6-diamidino-2-phenylindole and DNA duplexes in solution. *J. Am. Chem. Soc. 125*, 1759−1769.
(76) Masukawa, K. M., Kollman, P. A., and Kuntz, I. D. (2003) Investigation of Neuraminidase-Substrate Recognition Using Molecular Dynamics and Free Energy Calculations. *J. Med. Chem. 46*, 5628−5637.
(77) Wong, S., Amaro, R. E., and McCammon, J. A. (2009) MM-PBSA Captures Key Role of Intercalating Water Molecules at a Protein-Protein Interface. *J. Chem. Theory Comput. 5*, 422−429.
(78) Liu, Y., and Gray, N. S. (2006) Rational design of inhibitors that bind to inactive kinase conformations. *Nat. Chem. Biol. 2*, 358−364.
(79) Sergina, N. V., Rausch, M., Wang, D., Blair, J., Hann, B., Shokat, K. M., and Moasser, M. M. (2007) Escape from HER-family tyrosine kinase inhibitor therapy by the kinase-inactive HER3. *Nature 445*, 437−441.
(80) Hsieh, A. C., and Moasser, M. M. (2007) Targeting HER proteins in cancer therapy and the role of the non-target HER3. *Br. J. Cancer 97*, 453−457.
(81) Garrett, J. T., Olivares, M. G., Rinehart, C., Granja-Ingram, N. D., Sanchez, V., Chakrabarty, A., Dave, B., Cook, R. S., Pao, W., McKinely, E., Manning, H. C., Chang, J., and Arteaga, C. L. (2011) Transcriptional and posttranslational up-regulation of HER3 (ErbB3) compensates for inhibition of the HER2 tyrosine kinase. *Proc. Natl. Acad. Sci. U.S.A. 108*, 5021−5026.
(82) Shi, F., Telesco, S. E., Liu, Y., Radhakrishnan, R., and Lemmon, M. A. (2010) ErbB3/HER3 intracellular domain is competent to bind ATP and catalyze autophosphorylation. *Proc. Natl. Acad. Sci. U.S.A. 107*, 7692−7697.
(83) Jura, N., Shan, Y., Cao, X., Shaw, D. E., and Kuriyan, J. (2009) Structural analysis of the catalytically inactive kinase domain of the human EGF receptor 3. *Proc. Natl. Acad. Sci. U.S.A. 106*, 21608−21613.