In: HER2 and Cancer
Editors: Sophie I. Williams et al., pp. 47-79
ISBN: 978-1-61122-650-8

Chapter II

Investigating Molecular Mechanisms of Specificity in Regulation of the HER2 Receptor Tyrosine Kinase through Molecular Modeling and Simulation

Shannon E. Telesco*, Andrew Shih, Yingting Liu and Ravi Radhakrishnan
Department of Bioengineering, University of Pennsylvania,
Philadelphia, Pennsylvania, USA

Abstract

Human epidermal growth factor receptor 2 (HER2)/ErbB2 is a receptor tyrosine kinase belonging to the EGFR/ErbB family and is overexpressed in 20-30% of human breast cancers. Since there is a growing effort to develop pharmacological inhibitors of the HER2 kinase for the treatment of breast cancer, it is clinically valuable to rationalize how specific mutations impact the molecular mechanism of receptor activation. Although several crystal structures of the ErbB kinases have been solved, the precise mechanism of HER2 activation remains unknown, and it has been suggested that HER2 is unique in its requirement for phosphorylation of Y877, a key tyrosine residue located in the activation loop (A-loop). In our studies, discussed here, we have investigated the mechanisms that are important in HER2 kinase domain regulation and compared them with the other ErbB family members, namely EGFR and ErbB4, to determine the molecular basis for HER2's unique mode of activation. We apply computational simulation techniques at the atomic and electronic structure (quantum mechanical) levels to elucidate details of the mechanisms governing the kinase domains of these ErbB members. Through analysis of our simulation results, we have discovered potential regulatory mechanisms common to EGFR, HER2, and ErbB4, including a tight coupling between the A-loop and catalytic loop that may contribute to alignment of residues required for catalysis in the active

*Department of Bioengineering, University of Pennsylvania, 210 S. 33rd Street, 240 Skirkanich Hall, Philadelphia, PA 19104 USA.

kinase. We further postulate an autoinhibitory mechanism whereby the inactive kinase is stabilized through sequestration of catalytic residues. In HER2, we also predict a role for phosphorylated Y877 in bridging a network of hydrogen bonds that fasten the A-loop in its active conformation, suggesting that HER2 may be unique among the ErbB members in requiring A-loop tyrosine phosphorylation for its functionality. In EGFR, HER2, and ErbB4, we discuss the possible effects of activating mutations. Delineation of the activation mechanism of HER2 in the context of the other ErbB members is crucial for understanding how the activated kinase might interact with downstream molecules and couple to signaling cascades that promote cancer. Our comparative analysis furthers insight into the mechanics of activation of the HER2 kinase and enables us to predict the effect of an identified insertion mutation on HER2 activation. Further understanding of the mechanism of HER2 kinase activation at the atomic scale and how it couples to downstream signaling at the cellular scale will elucidate predictive molecular phenotypes that may indicate likelihood of response to specific therapies for HER2-mediated cancers.

1. Introduction

HER2 (ErbB2/Neu) is a member of the ErbB family of transmembrane receptor tyrosine kinases, which also includes the epidermal growth factor receptor (EGFR/ErbB1), ErbB3 (HER3), and ErbB4 (HER4). ErbB receptors are composed of an extracellular ligand-binding domain, a transmembrane segment, an intracellular protein tyrosine kinase domain, and a C-terminal tail harboring tyrosine phosphorylation sites [1]. In the ErbB system, ligand binding to the extracellular domain promotes homo- or heterodimerization of the receptors and activation of their cytoplasmic domains [2, 3]. Dimerization of the kinase domains induces phosphorylation of tyrosine residues in the C-terminal tail segments, which serve as docking sites for signaling molecules containing SH2 or PTB domains [4]. The major signaling networks activated by the ErbB family include the Ras-Raf-MEK-MAPK and PI3K-Akt pathways, both of which result in transcription of genes involved in cellular proliferation, differentiation, and migration [5]. Aberrant activation of the ErbB network is correlated with oncogenesis and development of lung, gastric, and breast cancers [6-12]. In particular, mutation or overexpression of HER2 results in uncontrolled, ligand-independent activation of kinase signaling and is correlated with 20–30% of human breast cancers [13, 14].

Binding of ligand to the extracellular domain of the HER2 receptor induces conformational changes within the cytoplasmic domain. In order to catalyze transfer of the γ-phosphate of ATP to tyrosine residues on substrate peptides, several important sub-regions within the kinase domain must be appropriately organized [15]. The catalytic loop (C-loop) comprises residues 844–850 in HER2, and is crucial in promoting the phosphoryl transfer reaction. The αC helix (residues 761–775) and the nucleotide-binding loop, or N-loop (residues 727–732), help to coordinate the ATP molecule and substrate tyrosine residue. The activation loop (A-loop) includes residues 863–884 and regulates kinase activity by controlling access of the catalytic site to peptide binding. The A-loop undergoes a pronounced conformational change upon activation, extending from its inactive, furled state to uncover the catalytic machinery and promote binding of the substrate peptide to the C-loop [16].

In the majority of known protein kinases (insulin receptor kinase or IRK, for example), the A-loop assumes its catalytically competent conformation only if it is first phosphorylated on a regulatory tyrosine residue [17]. The regulatory tyrosine residue is Y877 in HER2 (Y845

in EGFR). Phosphorylation of EGFR on Y845 has been demonstrated experimentally; however, phosphorylation does not seem to be a criterion for catalytic activity, as EGFR containing a Y845F mutation retains its activity [18]. In HER2, by contrast, several reports have illustrated the significance of Y877 phosphorylation for kinase activity [19, 20]. Xu et al. [21] demonstrated that mutation of Y877 to phenylalanine in COS-7 cells results in decreased autophosphorylation of Y1248, a tyrosine located in the C-terminal tail of HER2. It remains an open question whether or not phosphorylation of Y877 is required for full HER2 kinase activity.

Although the crystal structure of the HER2 kinase domain has not yet been resolved, several crystal structures of the extracellular and kinase domains of the other ErbB family members (EGFR, ErbB3, ErbB4) have provided insights into potential mechanisms of regulation and activation [16, 22-30]. Recent structural studies demonstrate that EGFR and ErbB4 are activated through an asymmetric dimerization mechanism analogous to that of a cyclin bound to an activated cyclin-dependent kinase (CDK) [31, 32]. The C-lobe of the activating kinase interacts with the N-lobe of the kinase undergoing activation, promoting full activation through allosteric contacts. Hydrophobic interactions between the C-lobe of the activator kinase and the αC helix of the kinase undergoing activation constitute the majority of the dimer interface. A sequence alignment of the ErbB kinase domains reveals that the specific amino acids participating in the dimer interface are conserved across the ErbB family members, suggesting that HER2 may be activated through a similar asymmetric dimerization mechanism.

Uncontrolled, ligand-independent activation of the HER2 receptor results in aberrant induction of cellular proliferation pathways and ultimately in tumorigenesis [14]. This constitutive activity is typically a result of HER2 overexpression that stems from a gene duplication event. Thus therapeutic approaches to treat HER2-mediated cancers have focused on inhibition of extracellular dimerization, and hence activation, using monoclonal antibodies. Trastuzumab (Herceptin), a humanized monoclonal antibody, binds to domain IV of the HER2 extracellular domain, and mediates its therapeutic effects through a variety of mechanisms, including antibody-dependent cellular cytotoxicity (ADCC) and disruption of receptor dimerization [33-35]. Pertuzumab (Omnitarg) [36, 37] binds to domain II of the HER2 extracellular domain and mediates its antiproliferative effects partially through inhibition of HER2 heterodimerization. Pertuzumab and trastuzumab have complementary modes of action and have demonstrated enhanced antitumor activity when employed as a combined therapy [38]. However, development of drug resistance in certain patients treated with trastuzumab [39] is an emerging challenge and is thought to be mediated by a variety of mechanisms, including signaling through alternative pathways and failure to activate the cellular ADCC response. Thus alternative pharmacological approaches are necessary.

In particular, the intracellular kinase domain presents an alternative candidate for HER2-mediated cancer therapeutics. Indeed, the development of small-molecule tyrosine kinase inhibitors (TKI), which target the catalytic site of the kinase domain, has been an area of intense focus in cancer therapeutics. TKIs act as ATP analogues to block ATP-binding in the kinase active site, thereby inhibiting phosphorylation of target substrates. Lapatinib (Tykerb) [40-42], a TKI that binds to both EGFR and HER2, has been approved for the treatment of HER2-overexpressing breast cancers. Analogous TKIs that are specific for EGFR include erlotinib [43, 44] and gefitinib [45, 46], both of which have been approved for the treatment

of non-small cell lung cancer (NSCLC). A specific class of mutations in EGFR, including point mutations in the A-loop, have been reported to correlate with the favorable clinical response of patients to gefitinib and erlotinib [47, 48]. The identification of these mutations and their correlation with the drug response has motivated further investigation into the effects and specificity of TKIs as first-line cancer drugs.

In light of the potential for development of next-generation tyrosine kinase inhibitors targeting HER2, our work is focused on elucidating mechanisms of HER2 kinase domain regulation and activation through molecular modeling and simulation. In Section 2 we describe some of the commonly employed modeling and simulation techniques that are applied in our HER2 studies. Section 3 details some of our recent results in regard to HER2 kinase domain structure and function [29, 49-51]. Specifically, we investigate the dimer-mediated allosteric activation mechanism of the HER2 kinase through dynamics simulations of a HER2/EGFR heterodimeric system and quantify the effect of phosphorylation of the A-loop tyrosine residue, Y877, on HER2 activity. We also compare the regulatory mechanisms identified in HER2 with those in the other ErbB family members through simulation of both the EGFR and ErbB4 kinases, to determine the molecular basis for HER2's unique mode of autoinhibition and activation. Further understanding of the mechanism of HER2 kinase activation at the atomic scale will elucidate predictive molecular phenotypes that may indicate likelihood of response to specific therapies for HER2-mediated cancers.

2. Computational Methods

The utility and continuously improving methodology of multiscale simulation techniques has helped bridge the gap between theory and experiment. Electronic structure (quantum level or *ab initio*) simulations can demonstrate how biomolecules of interest assume stable geometrical configurations and charge distributions when subject to certain molecular environments. By examining the charge distributions and structure it is possible to quantify and predict structural properties as well as functional activity, such as mechanisms of enzyme catalysis. However, quantum simulations are limited to approximately a few hundred atoms at most. Therefore, one must turn to a simulation method suitable for a larger spatial and temporal scale: molecular dynamics (MD) simulations, which are based on classical (empirical) force-fields, can model hundreds of thousands of atoms for nanoseconds and in some cases, up to a microsecond in time. Since MD simulations can resolve proteins and biomolecules at the atomic level, they are able to provide detailed information regarding molecular mechanisms of protein activation, conformational changes in biomolecules, and the free energy of ligand- or inhibitor-binding. In the following sections, we describe some of the most commonly employed molecular modeling and simulation techniques that are applied to the study of biomolecular systems.

2.1. Molecular Dynamics

Molecular dynamics (MD) simulation methods are commonly used to model systems of biomolecules [52, 53] because they can monitor individual atoms and therefore provide a high

degree of resolution to understand molecular properties and mechanisms of interest. To perform MD simulations, the initial coordinates and velocities of the atoms comprising the molecular system must first be specified. The system typically includes the desired biomolecule in addition to water molecules or other solvent, ions and membranes. The atomic coordinates of the biomolecule of interest can usually be found as X-ray or NMR data included in the protein data bank (PDB) [54] (www.pdb.org). If the coordinates have not been deposited in the PDB, the biomolecule geometry may be constructed using homology modeling techniques (see section 2.2). The potential of interactions of each of the atoms are calculated using a force field, which parameterizes the non-bonded and bonded interaction terms of each atom depending on its constituent atom connectivity and includes terms for bond, angle, dihedral, improper dihedral, non-bonded Lennard-Jones, and electrostatic energies. The potential interactions are summed across all of the atoms contained in the system, to produce an overall potential energy function for the system [55-57]:

$$U(\vec{R}) = \sum_{bonds} K_b(b - b_0)^2 + \sum_{angles} K_\theta(\theta - \theta_0)^2 + \sum_{dihedrals} K_\chi(1 + \cos(\eta\chi - \delta)) + impropersK\phi\phi - \phi 02 + nonbondedEijRminijrij12 - Rminijrij6 + qiqjerij$$

The total force acting on the system is derived by computing the derivative of the potential energy function. Computation of the derivative results in a set of 3N-coupled second-order ordinary differential equations that must be solved numerically. Numerical integration of the equations of motion results in the advancement of the positions and velocities of each atom by one time step. To generate a complete MD trajectory, this procedure is repeated iteratively. Several commonly-used force fields include CHARMM27 [58] (www.charmm.org), AMBER [59] (www.ambermd.org), and GROMOS [60] (www. gromacs.org), as well as dynamic simulation packages and visualization tools such as NAMD [61] (www.ks.uiuc.edu/Research/namd/) and VMD [62] (www.ks.uiuc.edu/Research/ vmd/).

To analyze the results of a molecular dynamics simulation, hydrophilic interactions are typically assessed through a hydrogen bond analysis of the system. In this type of analysis, the hydrogen bonds that persist throughout a given trajectory can be identified in order to differentiate between true stabilizing bonds and transient interactions [50]. In addition, hydrophobic interactions can be quantified using the solvent accessible surface area (SASA) calculation. The SASA calculation maps the surface area created using a probe sphere of 1.4 Å and provides a quantitative metric of hydrophobic stabilization effects.

A complementary approach to analyze protein dynamics trajectories is principal component analysis (PCA) [63, 64], which projects out independent motions in an MD trajectory and sorts them in the order of their dominance. This is accomplished through diagonalization of the variance-covariance matrix of atomic fluctuations along the MD trajectory. The eigenvalue equation: $[\sigma - \lambda I]\xi = 0$ is then solved to project out independent modes of atomic fluctuations captured in an MD trajectory and sort them by their variance. Here σ is a two dimensional variance-covariance matrix of atomic fluctuations about the trajectory average, with elements $\sigma_{ij}=\langle(x_i-\langle x_i\rangle)(x_j-\langle x_j\rangle)\rangle$ (i,j =1,…,3N, N being the total number of atoms with position given by Cartesian coordinates x). $\xi=(\xi_1,\xi_2,\ldots,\xi_{3N})$ are the 3N independent (uncorrelated) eigenvectors (PC) with eigenvalues $\lambda=(\lambda_I, \lambda_2,\ldots,\lambda_{3N})$ sorted in descending order i.e. $\lambda_1>\lambda_2\ldots,\lambda_{3N-7}>\lambda_{3N-6}$. The eigenvectors represent the principal

components (PCs) and the eigenvalues reflect their relative magnitudes. Application of PCA to identify pair wise correlations between motions of atoms in a region of interest such as the active site of an enzyme can provide significant insights in terms of relating structure to function through dynamics captured in the MD trajectory [64].

A challenge involved in performing MD simulations is the requirement to perform sufficient sampling of the combinatorially large number of potential conformations of the biomolecule of interest [65, 66]. In order to ensure adequate conformational sampling, the simulation must be run for long periods of time, for micro- or even milliseconds. However, the fastest atomic vibrations occurring in the simulation limit the time step to approximately 2 fs, which restricts the entire simulation length to the nanosecond range. For typical solvated system sizes of 40,000-500,000 atoms, simulation times extending into the microsecond range are challenging to achieve, and require improved algorithms as well as high-performance computing hardware infrastructure. An example of an algorithmic improvement is the particle mesh Ewald (PME) algorithm [67], which allows for calculation of long-range electrostatic interactions by replacing the summation of interaction energies in real space with an equivalent summation in Fourier space. An example of improved hardware infrastructure is the availability of supercomputing resources such as the US National Science Foundation's TeraGrid Initiative (founded in 2001) for academic researchers. TeraGrid resources (www.teragrid.org) currently possess more than a petaflop of computing capability and more than 30 petabytes of online and archival data storage, with rapid access and retrieval over high-performance networks.

2.2. Homology Modeling for Protein Structure Prediction

Commonly used experimental methods for resolving protein structure through X-ray crystallography or NMR spectroscopy are challenging and time-intensive and thus can solve only a small fraction of proteins that have been sequenced in the past couple of decades. Currently, more than 6,800,000 protein sequences have been deposited in the protein sequence database (NR; accessible through the National Center for Biotechnology Information: ftp://ftp.ncbi.nlm.nih.gov/blast/db/), while the Protein Data Bank (PDB; http://www.rcsb.org/pdb/) contains fewer than 50,000 protein structures [68]. Thus protein structure prediction is an important technique for resolving the discrepancy between protein sequence and structure. In particular, the homology modeling method, or prediction of an unknown structure by using a related protein with a known structure as a template, has been one of the most successfully implemented methods for protein structure prediction [69-73]; several computer programs and web servers that perform homology modeling include: Swiss-Model server (http://www.expasy.ch/swissmod/); CPH models (http://www.cbs.dtu.dk/services/CPHmodels/); and MODELLER [70] (http://salilab.org/modeller/). Homology modeling typically consists of the following steps: search for homologous protein structures, selection of an appropriate template, target-template alignment, model construction, and model quality assessment.

The identification of highly related protein structures to be used as potential templates for homology modeling typically involves searching the PDB to compare the target sequence with the sequence of every structure in the database. In order to select the best template for comparative modeling, several factors must be considered. First, the higher the degree of

sequence identity between the target and template, the better the quality of the template. In addition, the molecular environment of the template and target (i.e., the type of solvent, pH, presence of ligands) should be as similar as possible [70]. Upon selection of the template structure, a target-template alignment must be performed, which is typically achieved using standard sequence alignment methods [74-76]. If the degree of sequence identity between the target and template is below 40%, user intervention is required to correct any gaps or misaligned residues generated in the alignment. Several homology modeling programs, including MODELLER, apply a combination of sequence and structure information in the alignment algorithm to produce a more accurate alignment.

Upon completion of the target-template alignment, a 3D model of the target protein may be constructed using various model-building algorithms [77, 78]. A list of spatial (geometric) restraints is specified based on the template geometry as well as on average geometries derived from a representative set of proteins contained in the model-building database [79-81]. The MODELLER program, for instance, applies spatial restraints that are derived from the bond distances and angles in the target structure, as well as stereochemical restraints on bond distance and dihedral angle values that are obtained from the average of a set of known protein structures. Energy minimization and equilibration of the homology-modeled structure are then applied to optimize the model geometry.

Once the homology model has been constructed, its overall quality must be assessed. Several model analysis methods exist, most of which assess the stereochemical properties of the model, including bonds, bond angles, dihedral angles, and non-bonded atom-atom distances. Programs such as PROCHECK (www.ebi.ac.uk/thornton-srv/software PRO CHECK [82]) perform this type of analysis. The homology model can also be analyzed by constructing an energy profile for each protein residue in the structure, where high-energy peaks in the profile may represent poorly modeled regions. A caveat of this method is that a segment of the homology structure may appear to be erroneous, when in fact it is only interacting with a poorly modeled region. Therefore, the use of energy profiles should not be the only means of model validation.

Although the homology modeling method has successfully been applied to many different types of biomolecules, several challenges exist. For target molecules with low percent sequence identity relative to the template structure (ie, <20% target-template sequence identity), approximately half of the residues in the model may be misaligned [83]. This challenge can be overcome by incorporating more than one template into the target-template alignment. Another issue in homology modeling is optimization of residue side-chain conformation. Typically, rotamer libraries are used [84], which contain preferred side-chain torsional angles for specific conformations of residue side chains. As the number of rotamers increases, the ability to sample all potential conformations becomes a combinatorial problem. Xiang et al. [85] have recently shown that the use of a rotamer library based on Cartesian coordinates of known structures, rather than optimal bond lengths and angles, can successfully predict side-chain conformation. Despite these current issues, homology modeling remains a widely-applicable method for accurate protein structure prediction. As the accuracy of the technique continues to improve, homology modeling will further close the gap between the number of known sequences and the number of available structures.

2.3. Free Energy of Biomolecular Systems

The second law of thermodynamics states that natural systems seek a state of minimum free energy at equilibrium. Therefore the calculation of the free energy of a system can be a useful metric for comparing the results of simulation and experiment. A variety of computational techniques have been developed for the calculation of the free energy of biological systems, and here we will discuss two of the more commonly applied techniques: the free energy perturbation (FEP) method [86] and umbrella sampling [87].

Free Energy Perturbation (FEP): In a biomolecular system, it is often useful to compute the free energy difference, ΔF, between two defined thermodynamic states, for example, a ligand-bound versus unbound molecule. The free energy difference between the two states is represented as [88]:

$$\Delta F = -\frac{1}{\beta} ln\langle \exp[-\beta \Delta v(x)]\rangle_0; \beta = \frac{1}{k_B T},$$

where the subscript zero indicates configurational averaging over the ensemble of configurations representative of the initial state of the system, k_B is the Boltzmann constant, T is the temperature, and $v(x)$ is the potential energy function that depends on the Cartesian coordinates of the system, $[x]$. ΔF can also be computed by the reverse integration:

$$\Delta F = -\frac{1}{\beta} ln\langle \exp[-\beta \Delta v(x)]\rangle_1,$$

where the subscript one indicates averaging over the ensemble of configurations representative of the final state of the system. The most straightforward implementation of the FEP method involves defining the potential energy function for each state and performing a molecular dynamics simulation for the initial state of the system, hence calculating the ensemble average in Equation 1. In order to estimate the statistical uncertainty in ΔF, both forward and backward integrations may be performed. This uncertainty arises from conformations sampled in the ensemble representative of the initial state but not the final state and vice versa, and is typically small when the initial and final states of the system are very similar (i.e., the free energy difference between the initial and final states is on the order of $2k_BT$, or 1.5 kcal/mol) [89]. For systems in which the free energy difference is significantly larger, a series of intermediate states must be defined and must differ by no more than $2k_BT$. The total ΔF can then be computed by summing the ΔFs between the intermediate states:

$$\Delta F = -\frac{1}{\beta} \sum_{i=1}^{M+1} ln\langle \exp[-\beta [v(x;\lambda_{i+1}) - v(x;\lambda_i)]]\rangle_{\lambda_i},$$

where M indicates the number of intermediate states and λ is the coupling parameter, a continuous parameter that marks the extent of the transition from the initial to the final state. As λ is varied from 0 (initial state) to 1 (final state), the potential energy function $v(x; \lambda)$ passes from v_0 to v_1.

A limitation of the FEP method is that the end points of the transformation, at which sets of atoms are created or destroyed, are subject to VDW clashes that result in end-point catastrophes [90, 91]. Therefore, to compute the free energy at the diverging end points, the

extent of sampling can be increased at the beginning and end of the FEP simulations to collect data from multiple windows with λ values close to zero or one. Despite this limitation, the FEP method has been widely applied in the calculation of the free energy of many biomolecular systems [92-96].

Umbrella Sampling: The umbrella sampling procedure provides a method for calculation of the potential of mean force (PMF) (free energy density) along a specified set of reaction coordinates (order parameters), with the end goal of computing the free energy change along the specified reaction coordinate(s) (see for example, [97]). For the free energy calculation, the probability distribution $P(\chi_i)$ is calculated by dividing the range of order parameter χ_i into several windows. The histograms for each window are collected by harvesting trajectories in that window, from which the PMF $\Lambda(\chi_i)$ is calculated; the PMF $\Lambda_i(\chi_i)$ is given by [98, 99],

$$\Lambda_i(\chi_i) = -k_B T \ln(P(\chi_i)) + \text{Constant; Then, } \exp(-\beta \Delta F) = \int \exp(-\beta \Lambda_i(\chi_i))\, d\chi_i$$

The functions $\Lambda_i(\chi_i)$ in different windows are pieced together by matching the constants such that the Λ_i function is continuous at the boundaries of the windows. The standard deviation in each window of the PMF calculations is estimated by dividing the set of trajectories into two blocks and collecting individual histograms. To compute the PMF along multiple reaction coordinates, the weighted histogram analysis method (WHAM) [100] can be applied, which helps to recombine the various windows of sampling and ultimately to compute ΔF [99-103].

2.4. Electronic Structure Methods

Electronic structure methods [104-106] are typically applied to map chemical reaction pathways and catalytic mechanisms occurring in enzymes and other types of biomolecular systems. Although electronic structure methods are computationally intensive, recent improvements in mixed quantum mechanics molecular mechanics (QM/MM) methods combine electronic structure methods with less labor-intensive molecular mechanics [107-118]. In this section we discuss the application of QM/MM methods to the study of catalytic reaction pathways.

In the QM/MM simulations, the system is sub-divided into two sub-regions, the quantum mechanical (QM) sub-region, where the reactive events take place, and the molecular mechanical sub-region (which provides the complete environment around the reactive chemistry) [108, 110]. The QM sub-region is focused on a small number of atoms of the total system, as electronic structure methods are computationally intensive and therefore restricted in terms of the number of atoms to be included. In an enzymatic system, the quantum region might include Mg^{2+} ions, water molecules within 3 Å of the Mg^{2+} ions, and segments of the substrate molecules and the catalytic protein residues (such as aspartic acids). The molecular mechanical (MM) sub-region consists of the remaining protein and solvent molecules, which are treated using the classical force-field (such as CHARMM27).

In QM/MM simulations, wave function optimizations are typically performed in the quantum (or QM) sub-region of the system using an electronic structure method such as density functional theory (DFT) [105]. In this step, the electrostatic coupling between the QM and the MM sub-regions is accounted for: i.e., the charges in the MM sub-region are allowed

to polarize the electronic wave functions in the QM sub-region. The forces in the quantum sub-region are calculated using DFT on-the-fly assuming that the system moves on the Born-Oppenheimer surface [110, 119]: we assume a clear timescale of separation between the electronic and nuclear degrees of freedom and the electronic degrees of freedom are in their ground state around the instantaneous configurations of the nuclei. The forces on the classical region are calculated using a classical force-field. A mixed Hamiltonian (energy function) accounts for the interaction between the classical and the quantum sub-regions. Since the QM/MM boundary often cuts across covalent bonds, one can use a link atom procedure [113] to satisfy the valences of broken bonds in the QM sub-region. Also, bonded and electrostatic terms between the atoms of the QM region and those of the classical region are typically included [111].

QM/MM methods are typically implemented based on existing interfaces between the electronic structure and the molecular dynamics programs; one implementation is between GAMESS-UK (www.cfs.dl.ac.uk [120]) (an *ab-initio* electronic structure prediction package) and CHARMM [58]. The system of interest then undergoes energy minimization and equilibration runs using the numerical integration procedures generally applied to pure MM systems. The significance of the QM/MM simulations is that the sequence of reaction events occurring in the active site can be mapped in the context of the MM region (solvent, remaining protein). The major challenge in the successful implementation of QM/MM simulations lies in the choice of the QM sub-region and the system size. In addition, high resolution structures are usually required for performing QM/MM simulations, as the trajectories sample a limited conformational space. These challenges are currently being addressed by enhanced methods for electronic structure and molecular mechanics simulations [121, 122].

2.5. Limitations of Computational Modeling

A potential limitation of computational modeling is the treatment of poorly- or unresolved regions in protein crystal structures. This issue can partially be overcome through homology modeling methods that allow for loop-modeling of the unresolved regions. Several challenges also exist in molecular dynamics simulations; these limitations include force-field uncertainties, solvent approximations, limited sampling, and finite size effects. To warrant the choice of one force field over another, several force-fields should be compared, such as CHARMM and AMBER. In regard to solvent approximations, the major issue lies in the determination of the protonation states of certain residues. This can be addressed by using Debye-Huckel calculations for titratable side chain residues and by Poisson-Boltzmann (MMPBSA) [123] evaluation of the relative free energies for residues participating in the catalytic reaction. QM/MM calculations can also be applied to compare the relative stabilities of the different protonation states. The QM/MM methods, however, also involve inherent approximations, such as the choice of the QM/MM boundary. As the boundary between the MM and QM regions cuts through covalent bonds, the single link atom procedure satisfies valences of broken bonds. However, several applications [111] have shown that the double link atom method produces more accurate results than the other popular approaches using the single link atom method [113].

The issue of limited conformational sampling explored in molecular dynamics simulations can be partially overcome by enhanced-sampling algorithms as discussed in section 2.3 (Umbrella Sampling). However, these enhanced-sampling methods must be assessed in terms of the effects of multiple low-energy pathways and the choice of reaction coordinates. The PCA method [124] (discussed in section 2.1) provides an alternative means of extracting the most significant and dominant modes of protein dynamics from the MD trajectory, despite the finite simulation time [125, 126]. While PCA is not useful in analyzing the slow modes of the system beyond what is captured in the dynamics trajectory, it does provide an approximate description of the slow modes faster than the time scale of the trajectory (t). Despite the approximations and limitations inherent in molecular modeling and simulation, the predictions and results produced from such simulations are invaluable for providing mechanistic explanations for biological processes and informing *in vitro* and *in vivo* experiments.

3. Applications of Computational Modeling and Simulation to Investigate Regulatory Mechanisms of Activation in the HER2 Kinase

In this section we describe the application of the computational methods discussed in Section 2 to the study of the mechanisms that are important in HER2 kinase domain regulation and activation. Our molecular dynamics simulations provide us with a framework for studying the atomistic behavior of the HER2 kinase domain in both monomeric and dimeric contexts, and for comparing HER2 with the other ErbB family members, EGFR and ErbB4, to determine the molecular basis for HER2's unique mode of activation. As mentioned in the Introduction, aberrant HER2 signaling is implicated in at least 30% of human breast cancer cases, and so elucidation of the molecular regulatory mechanisms will help establish structure-function relationships in the wild-type HER2 kinase, as well as predict mutations with a propensity for constitutive activation. Such molecular variants in HER2 receptors are known to profoundly affect specific therapies targeting HER2-mediated cancers.

3.1. Molecular Dynamics Simulations of the HER2 Kinase Identify an Orchestrated Set of Interactions among the Catalytic Domains to Promote Phosphoryl Transfer

In our recent studies [51], we performed molecular dynamics (MD) simulations of inactive and active HER2 structures to elucidate details of the mechanism by which the HER2 kinase domain is regulated and activated. The models of the HER2 kinase domain were constructed based on homology to EGFR, as the crystal structure of HER2 kinase has not yet been resolved. Each solvated system was subjected to MD simulation for 10 ns and the trajectories were analyzed for conformational shifting as well as for key bonding patterns. The inactive and active systems were stable for the duration of the simulation, as we observed by monitoring the root-mean square deviation (RMSD) of the backbone atoms of each structure over time. The αC helix demonstrated larger fluctuations in the active systems but

overall the RMSD remained steady and neither structure shifted toward an alternative conformation during the simulations.

Upon completion of the MD simulations, a principal component analysis (PCA) was performed to identify the most significant global motions of the HER2 kinase. We hypothesized that the atomic fluctuations in the inactive and active forms of the HER2 kinase would differ dramatically, as conformational rearrangement of the kinase domain is expected to correlate with significant changes in the dynamical behavior of the protein. PCA of the HER2 systems was focused on an active site region that includes the C_α atoms of all domains critical for catalysis, including the A-, C-, and N-loops and the αC helix. We observed that the first three eigenmodes account for the majority of the atomic fluctuations in each trajectory, and the top 10 eigenmodes account for 75–85% of the total variation in each system.

In the active HER2 system, the first principal motion demonstrated movement of the αC helix into the active site, coupled with a tilting of the N-loop. Simultaneously, the A-loop exhibited a slight bending and the C-loop and N-loop shifted toward each other. Motion in the second eigenmode also corresponded to coupled fluctuations in all four sub-domains. The first and second principal motions for the inactive structure, by contrast, were dominated by A-loop motion with dampened fluctuations in the other sub-domains, suggesting that the catalytic sub-domains are less coupled in the inactive system (Figure I). Our results support the hypothesis that the interactions among the A-, C-, and N-loops and the αC helix in the active systems is important for alignment of the catalytic sub-regions in preparation for phosphoryl transfer. As the A-loop and C-loop contain the catalytic aspartate, D845, and the coordinating aspartate, D863, respectively, the coupling of these two loops is crucial for ensuring proper orientation of the aspartate residues for catalysis.

In order to place the regulatory mechanisms in HER2 in context by comparing them with those in the other ErbB family members (EGFR and ErbB4), we constructed molecular systems for and applied a similar PCA analysis to both EGFR and ErbB4 [49, 50]. By analyzing HER2 in relation to EGFR and ErbB4, we can relate sequence similarities to specific bond-interaction networks, both hydrophilic as well as hydrophobic, and dynamical modes, in order to explain HER2's unique mode of kinase regulation and activation. Therefore we constructed fully solvated systems for EGFR and ErbB4, in both inactive and active conformations, and subjected the systems to MD simulation using the same methods applied to our HER2 simulations. Principal component analysis (PCA) of the molecular dynamics simulations revealed that motions occurring within the active sites of the inactive and active EGFR and ErbB4 monomer systems differ significantly, particularly in the A-loop and αC-helix. In similarity to the HER2 kinase domain, the inactive EGFR and ErbB4 monomers exhibited large-amplitude motion in both the αC-helix and the A-loop, with smaller fluctuations in the N-loop and C-loop (Figure I). By contrast, the active ErbB monomers demonstrated a uniform level of motion across all four sub-domains of the active site with only low-amplitude fluctuations (2-3Å), and showed no significant local deformations. This implies that the motions are more tightly coordinated across the catalytic site in the active conformation compared to the inactive conformation, a result that is in agreement with our PCA results for HER2.

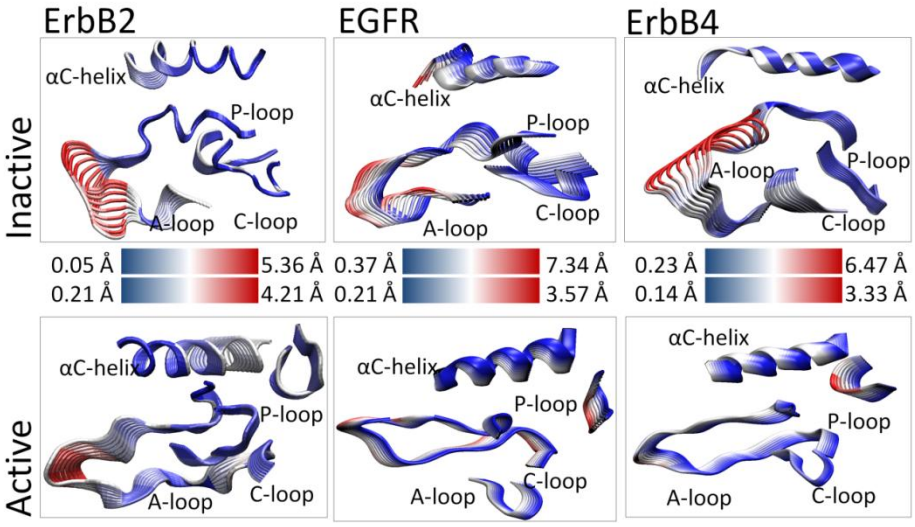

Figure I. Structures representing motion along the first principal mode for HER2, EGFR and ErbB4. The structures are color-coded according to the RMSD, where red regions indicate larger fluctuations and blue regions indicate smaller motions. The four catalytic sub-domains move in a concerted manner in the active systems, and these interactions may be important in alignment of catalytic residues.

3.2. Hydrogen Bonding Analysis of the HER2 Simulation Systems Reveals Distinctive Bonding Networks in the Inactive and Active Kinases

To determine the mechanistic basis for the coupling among the catalytic sub-domains in the active HER2 system, individual hydrogen bonds were recorded for each system through a hydrogen-bonding analysis of the MD trajectories (Table I). We observed six hydrogen bonds in the C-loop of the active HER2 structure; interestingly, most of these bonds couple the C-loop and A-loop, including: R844-L866, N850-T862, V842-R868, L852-K860 and L846-W888. In the inactive HER2 trajectory, we recorded only two hydrogen bonds that link the C-loop and the A-loop. The extensive network of hydrogen bonds and salt bridges that bridge the A- and C-loops in the active HER2 system helps to explain the concerted PCA fluctuations between these sub-domains. Specific bonds that are recorded in the hydrogen-bonding tables, such as R844-L866 and Y877-R844, also appear as correlated residue pairs in the PCA analysis.

The pattern of hydrogen bonds in the A-loop and αC helix of the inactive and active HER2 systems revealed results similar to that of the A-loop and C-loop (Table I). The major difference that we found between the inactive and active conformations was that there were significantly more bonds coupling the A-loop and αC helix in the active system compared to the inactive system. There are four hydrogen bonds and two salt bridges in the A-loop of the inactive structure, compared with nine hydrogen bonds and two salt bridges in the active A-loop. Similarly, there are two salt bridges in the αC helix of the active structure, whereas no salt bridges can be found in the inactive system. Most of the hydrogen bonds that we observed in the A-loop and αC helix of the active system bridge the catalytic sub-domains of the kinase. Two such bonds, E766-K883 and D769-R868, couple the A-loop and αC helix. The A-loop and C-loop reveal additional coupling bonds: R844-L866, N850-T862, V842-

R868, L852-K860, and L846-W888. The extensive network of hydrogen bonds that bridge the catalytic sub-domains (A-loop, C-loop, and αC helix) of the active HER2 kinase provides a mechanistic explanation for the concerted motions of these sub-domains that we observed in the PCA of the active system. As mentioned earlier, this tight coupling among the sub-domains may be significant in the pre-organization of the catalytic site for successful orchestration of the phosphoryl transfer reaction.

Table I. Hydrogen bonds that persist throughout the MD trajectories for HER2, EGFR, and ErbB4. The active systems contain a greater number of bonds that couple the catalytic sub-domains of the kinase, in comparison to the inactive systems

ErbB2 active	EGFR active	ErbB4 active	ErbB2 inactive	EGFR inactive	ErbB4 inactive	ErbB2 active	EGFR active	ErbB4 active	ErbB2 inactive	EGFR inactive	ErbB4 inactive
aC-helix A-loop bonds						A-loop C-loop bonds					
--	--	--	--	--	**E739,R841**	--	--	--	G865,V842	--	--
E766,K883	**E734,K851**	**E739,K856**	--	--	--	--	--	--	--	--	G838,R817
D769,R868	D737,K836	--	--	--	D742,R841	L866,R844	L834,R812	L839,R817	--	--	--
--	E738,F832	E743,F837	--	--	--	--	--	--	--	L834,D813	--
--	--	--	--	**E738,K836**	**E743,R841**	R868,V842	K836,V810	R841,V815	--	--	--
aC-helix C-loop bonds						--	**E848,R812**	--	--	--	--
--	--	--	--	--	E743,R817	--	--	--	--	K851,R812	--
aC-helix bonds						A-loop bonds					
A763,S760	--	--	--	--	--	--	--	D836,K726	**D863,K753**	--	**D836,K726**
E766,R756	--	--	--	--	--	--	--	D836,T835	--	--	--
E770,K753	**E738,K721**	**E743,K726**	--	--	--	L870,R840	L838,R808	L843,R813	--	--	--
--	--	--	M774,L785	M742,L753	M747,L758	D871,R840	--	--	--	--	--
--	A743,L679	A748,Q684	--	--	--	--	A840,G672	--	--	--	--
--	--	--	--	--	A748,R757	--	--	--	D873,R897	--	--
C-loop C-loop bonds						--	--	K848,T873	--	--	--
H843,D845	--	--	--	--	--	--	**K843,D932**	K848,D937	--	--	--
--	--	--	R844,D845	**R812,D813**	--	E876,R898	--	--	--	--	--
D845,R849	**D813,R817**	**D818,R822**	--	--	--	--	--	E849,K871	--	--	--
--	D813,N818	--	--	--	--	--	Y845,Y867	Y850,F872	--	--	--
A847,N850	A815,N818	A820,N823	A847,N850	A815,N818	--	--	--	--	--	H846,R865	--
A848,V851	--	--	--	--	--	--	--	A852,R870	--	--	--
						D880,R897	--	**D853,R870**	**D880,R897**	**E848,R865**	**D853,R870**
						--	--	--	--	--	G855,E730
						--	--	--	K883,E757	--	--
						--	--	--	--	--	**K856,E844**

In cataloging the interactions, we compared the HER2 bonding network to that of EGFR and ErbB4 (Table I). Several bonds were conserved across all members of the ErbB family in the active state (HER2 numbering is used in this discussion): two salt bridges: E766-K883 and E770-K753, three H-bonds: L866-R844, K868-V842, and L870-R840, as well as the bond D845-R849 which is an H-bond in HER2, but a salt bridge in EGFR and ErbB4. These bonds span the four key sub-domains, coordinate the sub-domain movements and ensure that the catalytically important residues are spatially localized. In particular, the E770-K753 salt bridge is highly conserved across all active RTKs and helps coordinate the α and β phosphates of ATP bound in the active site. The E766-K883 salt bridge connects the A-loop and the αC-helix, while the three conserved H-bonds link the A-loop and the C-loop, coupling the motions of these two loops. These can be regarded as "fastening" H-bonds that maintain the N-terminal side of the A-loop open in its active state – the alternative (in the inactive state) being steric hindrance to the binding of ATP and peptide substrates. The conserved

D845-R849 bond positions the sidechain of the catalytic aspartate D845 in the active site and likely facilitates the preorganization of the catalytic site in the active kinase system.

In contrast with the case for the active configurations, few intramolecular bonds that stabilize the inactive kinase conformation are conserved across the ErbB family. Instead, the bonds present in the inactive systems sequester key residues seen in the activating bond-network. ErbB4 is similar to HER2 in that the homologous K753 residue sequesters the coordinating aspartate D863, preventing the formation of the conserved E770-K753 salt bridge. In EGFR and ErbB4 , the other half of the salt bridge is sequestered: the homologous E770 residue is salt bridged to R868 (though in EGFR the arginine is replaced by a lysine). The similarity in the pattern of specific interactions that preferentially bridge the catalytic sub-domains in the active conformations of the EGFR, HER2, and ErbB4 kinases suggests that the tight coupling of the A- and C-loops may be a general feature of the architecture of ErbB receptor kinases that likely aids in the assembly of a catalytically competent active site.

3.3. Quantification of the Effects of A-Loop Phosphorylation in HER2 Reveals a Potential Role in Maintaining the Extended (Active) State of the A-Loop

In most protein kinases, the A-loop assumes its catalytically competent conformation only if it is first phosphorylated on a regulatory tyrosine residue within the A-loop [17]. The regulatory tyrosine residue is Y877 in HER2 (Y845 in EGFR). Although phosphorylation of EGFR on Y845 has been observed experimentally, phosphorylation does not seem to be required for catalytic activity, as EGFR possessing a Y845F mutation is fully active [18]. By contrast, the function of A-loop phosphorylation in HER2 is controversial, as multiple studies have reported the importance of Y877 phosphorylation for kinase activity [19, 20]. Hence it is possible that phosphorylation of Y877 potentiates HER2 kinase activity.

To investigate the effects of phosphorylation of Y877 on HER2 kinase activation, we simulated a Y877-phosphorylated version of the HER2 kinase in both inactive and active conformations. In the active Y877-phosphorylated HER2 system, we discovered a network of hydrogen bonds that preserve the A-loop in its extended, active conformation. The hydrogen bonds essentially fasten the A-loop to a segment of the αF helix (residues 896–901) and to the region between the αE helix and the C-loop (residues 840–844) to maintain the A-loop in its open state. Three hydrogen bonds (L866-R844, V842-R868, and R840-L870) fasten the A-loop at its N-terminal end and three bonds (Y877-F899, A879-R897, and E876-R898) fasten the A-loop at its C-terminal end. Interestingly, several of the interactions, such as L866-R844 and Y877-R844, link residues in the A-loop and C-loop, highlighting the prominent bridging of these regions in the active state. However, the coordinating aspartate residue, D863, and the catalytic aspartate, D845, do not appear in the hydrogen-bonding network and therefore remain available for catalysis. Our results support an activation model in which residues neighboring D863 and D845 form hydrogen bonds that maintain the A-loop in the active conformation while freeing the catalytic aspartate residues to remain poised for the phosphoryl transfer reaction.

In the active HER2 system, we identified a role for phosphorylated Y877 in bridging the network of hydrogen bonds that fasten the A-loop in its open conformation [51]. The

phosphoryl group on Y877 interacts with R844, K883, and R868, thereby bridging the C- and N-terminal ends of the A-loop. The main-chain oxygen of Y877 hydrogen-bonds with F899, further contributing to the network of fastening bonds in the A-loop. By contrast, fewer hydrogen bonds occur in the A-loop of the Y877-unphosphorylated active system, revealing that the absence of the phosphoryl group limits the number of intra-A-loop bonds that can form.

To quantify the effect of Y877-phosphorylation on HER2 activity, we computed the Helmholtz free energy difference between the Y877-unphosphorylated and Y877-phosphorylated states in the NVT ensemble using the FEP method (see Section 2.3). Four different simulations were performed for comparison, including the transformation of Y877 to pY877 in the unphosphorylated structures (inactive and active), and the transformation of pY877 to Y877 in the respective phosphorylated systems. The transformation of Y877 to phosphorylated Y877 produced a free energy change of -385.1 ± 1.2 kcal/mol for the active structure and -384.0 ± 0.8 kcal/mol for the inactive structure, yielding a $\Delta\Delta F$ value of -1.1 ± 1.4 kcal/mol. The ΔF values for the reverse transformation, pY877 to Y877, resulted in a $\Delta\Delta F$ value of 1.2 ± 1.5 kcal/mol [51]. The $\Delta\Delta F$ values suggest that phosphorylation of Y877 provides a small increase in stability of the active conformation relative to the inactive state, although it is insufficient to significantly lower the kinase activation barrier. Therefore, molecular stimuli in addition to the phosphorylation of Y877 are likely required for the complete conformational activation of the HER2 kinase.

3.4. The αC-β4 Loop Region May Contribute to Autoinhibition of the HER2 Kinase

In a recent study, Fan et al. [127] reported that HER2 is strongly autoinhibited relative to EGFR and ErbB4, and that a mechanism for the autoinhibition involves sequence variation in a loop connecting the αC helix and the β4 sheet. The HER2 kinase domain shares 83% sequence identity with EGFR; in the αC-β4 loop, however, five of the eight residues in HER2 differ from those in EGFR. In particular, the polar residues in the αC-β4 loop of EGFR are replaced by nonpolar residues in HER2, which form a hydrophobic patch that interacts with another segment of hydrophobic residues positioned in the A-loop. The hydrophobic patch in HER2 includes the following residues: V773, M774, G776, V777, G778, and V782 in the αC-β4 loop, and I861, T862, F864, L866, and L869 in the A-loop. Several groups have studied this hydrophobic region in terms of its interactions with various molecules, including the molecular chaperone Hsp90 [128, 129]. Fan et al. [127] hypothesized that the hydrophobic interactions between the αC-β4 loop and the A-loop stabilize the HER2 kinase in the inactive state, resulting in lower constitutive catalytic activity relative to EGFR and ErbB4.

To investigate the role of the αC-β4 region in HER2, we applied our hydrogen-bonding analysis to the αC-β4 loop in HER2 and compared to the bonding networks in EGFR and ErbB4. Our results revealed that both the inactive and active HER2 systems include only two hydrogen bonds in the αC-β4 region. The S783-I861 bond is shared by both systems and links the αC-β4 loop to the A-loop. EGFR and ErbB4, by contrast, reveal a stronger hydrogen

bonding network in the αC-β4 loop. The active EGFR and ErbB4 systems contain 10 and eight hydrogen bonds, respectively. Many of the bonds observed in active EGFR, including H749-I829, C751-I829, and H749-V827, connect residues in the αC-β4 loop and the A-loop. The active HER2 system lacks these hydrogen bonds because hydrophobic interactions, rather than hydrophilic contacts, predominate in the αC-β4 region. The H749-I829 and H749-V827 bonds are absent in HER2 because a relatively nonpolar tyrosine residue is located in the position analogous to the positively charged H749 in EGFR.

The network of hydrogen bonds identified in the αC-β4 region of the inactive HER2 system demonstrates a trend similar to the active HER2 system. The αC-β4 region of the inactive EGFR and ErbB4 systems contain four and six hydrogen bonds, respectively, whereas the inactive HER2 system contains only two bonds. The weak hydrogen-bonding network and the predominance of hydrophobic interactions in the αC-β4 loop of HER2 may be important in its association with the molecular chaperone Hsp90 [129]. Only HER2 among the members of the ErbB family interacts with Hsp90; therefore it has been hypothesized that binding of Hsp90 to the αC-β4 region in HER2 provides an inhibitory mechanism for regulation of HER2 activity by preventing dimerization of the HER2 kinase [128]. Interestingly, HER2 gene mutations have been identified in a cohort of non-small cell lung cancers that involve in-frame duplications/insertions within exon 20, a region that corresponds to the αC-β4 loop in the kinase domain. The most common mutation is the in-frame YVMA insertion at residue G776 ($G776^{YVMA}$), which produces significantly higher levels of tyrosine phosphorylation compared to wild-type HER2, resulting in uncontrolled cell growth [14]. These mutations may operate by weakening the hydrophobic interactions surrounding the αC-β4 loop and forming a hydrogen-bonding network similar to those we have observed in EGFR and ErbB4, which may perturb the HER2-Hsp90 interaction that normally maintains HER2 in an inhibited state.

3.5. Dimerization of HER2 Provides an Impetus for Kinase Activation by Disrupting the Network of Inactivating Bonds

The results that have been discussed in the previous sections focus on the monomeric HER2 kinase systems, but it has recently been shown that the ErbB kinases undergo activation through the formation of an asymmetric dimer. Therefore we constructed a HER2-EGFR heterodimer system in which HER2 is the kinase undergoing activation and EGFR is the activating kinase. The heterodimer was modeled on the structure of the EGFR dimer generated by Zhang et al. [32]. A second dimer was constructed in which HER2 was phosphorylated on Y877, to determine the effect of A-loop phosphorylation (if any) on dimerization and activation. We then solvated the dimer systems and performed a 20 ns MD simulation to determine whether the dimer interface acts as a stimulus for activation of HER2. The residues comprising the dimeric interface for HER2 included P707, Q711, M712, I714, L768, L790, and V794; for EGFR they included I917, Y920, M921, V924, M928, I929, and V956.

As was done for the monomer systems, PCA was performed for the dimer trajectories to extract the most significant atomic motions of the systems. The PCA was focused on an

active site region that included the C_α atoms of the A-, C-, and N-loops and the αC helix of HER2, the kinase that is undergoing activation. As we observed for the inactive monomeric simulations, the first and second eigenmodes are dominated by A-loop motion. However, there are also significant fluctuations in the αC helix of HER2. We also observe a slight conformational rearrangement of the αC helix toward the active state during the MD simulation, as it transitions from an RMSD of 2–6 Å relative to the inactive state. Thus the HER2 kinase (kinase undergoing activation) responds to the dimer interface by shifting toward the active state.

To provide a mechanistic explanation for our observations regarding the effects of the dimer interface on the HER2 structure, we focused our analysis on the pattern of hydrogen bonds and salt bridges that persist throughout the dimer trajectories. To identify hydrogen bonds that are weakened or broken as a result of the dimer interface, we performed a hydrogen-bonding analysis of the 20 ns dimer trajectories and compared the results to those of our monomeric systems. Our analysis of the bonding network in the heterodimeric systems showed that many interactions present in the monomeric systems were either profoundly weakened or severed throughout the course of the dimer simulations. In the Y877-unphosphorylated HER2 dimer, several inactivating bonds (ie, bonds that serve to maintain the inactive state) are broken, including N764-S760 and Y772-G776 in the αC helix and G865-V842 and D873-R897 in the A-loop. In the Y877-phosphorylated HER2 dimer, several bonds have also been perturbed, including T759-E874, L785-M774, and Y772-G776 in the αC helix and G865-H843, R868-R840, and V884-K887 in the A-loop. It is not entirely surprising that these interactions have been disrupted, as they are positioned in the region of the HER2 kinase (mostly the αC helix) that directly participates in the dimer interface. In both dimer systems, our results demonstrate that the dimerization interface directly perturbs the network of hydrogen bonds and salt bridges responsible for maintaining the kinase in the inactive state. This result is consistent with the allosteric activation mechanism proposed by Zhang et al. [32].

However, several interactions remain unperturbed by the dimerization interface and hence may pose significant free energy barriers to the conformational change accompanying activation. The D863-K753 interaction, which must break in order for the activating salt bridge, K753-E770, to assemble and also for the coordinating aspartate, D863, to be freed so that it can facilitate rearrangement of the substrate tyrosine and the ATP molecule during catalysis, persists throughout the dimer simulations. The K883-E757 bond, which is conserved among all three inactive ErbB kinases, is also important in maintaining the inactive structure, as it must break in order for the K883-E766 salt bridge to form in the active state. Thus we performed umbrella sampling simulations (see Section 2.3) of the HER2 dimer in order to force breakage of these unperturbed inactivating bonds and guide the dimer along potential activation pathways. The RMSD of the A-loop and αC-helix was applied as the reaction coordinate, as these regions undergo the most significant conformational changes upon kinase activation. Figure II details the results of our targeted simulations; the αC-helix shifted into the active site, transitioning from an RMSD of approximately 7 to 3 Å (see rightmost plot in Figure II). During the transition, the D863-K753 bond was weakened, although the activating salt bridge, K753-E770, has yet to form. Overall, the pattern of hydrogen bonds in the dimer has shifted toward the active conformation, although we

hypothesize that formation of the K753-E770 salt bridge may be one of the final events to occur along the complete activation pathway.

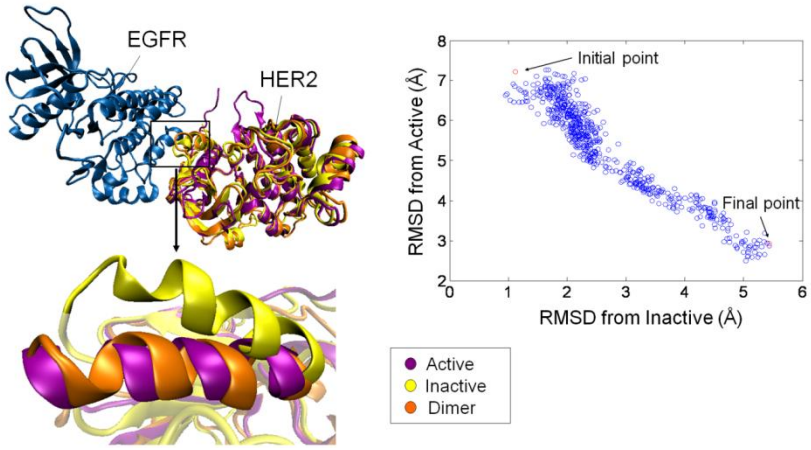

Figure II. Umbrella sampling to drive the HER2 heterodimer from the inactive to the active conformation. The αC helix is driven to the active configuration (purple). On the right is shown the progression of the simulation for the αC helix; the helix moves further from the inactive conformation and closer to the active conformation.

To determine the extent to which the molecular events occurring in the HER2 dimer are conserved across the ErbB family, homodimer systems of the EGFR and ErbB4 kinases were also constructed and simulated, and the results of the trajectory analysis were compared to those of the HER2 dimer. In both the EGFR and ErbB4 dimers, the C-lobe of the activator kinase was positioned in an asymmetric fashion relative to the N-lobe of the kinase undergoing activation, inducing a conformational shift of the αC-helix toward the active conformation. The conformational shift is correlated with perturbations to the hydrogen bonding networks of the dimer system. In EGFR, the interactions disrupted in the inactivating bond network surrounding the A-loop and the αC-helix are Y740-S744, L834-D813, H846-R865, and K851-R812. In ErbB4, a list of bonds disrupted upon dimerization includes E739-R841, D742-R841, E743-R817, G838-R817, G855-E730, and K856-E844. In both EGFR and ErbB4, the sequestering E738-K836 salt bridge has weakened significantly with the survival percentage (fraction present in the trajectory) of the bond in EGFR decreased from >90% in the inactive monomer trajectory to ~60% and from >90% in the monomer trajectory to ~70% in the dimer trajectory for ErbB4. Thus in all three dimer trajectories (EGFR, HER2, ErbB4), the pattern of bonds in the αC helix is profoundly changed, suggesting that the dimers are moving away from the inactive conformation and are more vulnerable to additional perturbations that would promote the active conformation.

3.6. Hydrophobic Analysis of HER2 and Comparison to EGFR and ErbB4

Recent structural studies have revealed highly conserved hydrophobic “spines” within kinases that are considered important for defining their catalytic state [130, 131]. The regulatory spine (R-spine) consists of four hydrophobic side chains (M774, L785, H843,

F864 in HER2) anchored by an aspartic acid in the αF-helix (D904 in HER2). The R-spine spans several key regulatory sub-domains, and coordinates the motion of the N- and C-lobes of the kinase [130]. The catalytic spine (C-spine) involves eight hydrophobic side-chains (V734, A751, L806, V851, L852, V853, T911, L915 in HER2) that help support and coordinate the adenine ring of ATP in the active state [131]. Similarly, in the inactive state there is a small hydrophobic 'core' formed between the αC-helix and the A-loop, which maintains the kinase in the inactive conformation. Disruption of this hydrophobic core by single point mutations has been shown to activate EGFR and ErbB4 [132-135].

Two analyses to measure the relative stability of the hydrophobic sub-regions include the mean Solvent Accessible Surface Area (SASA) method as well as a method for assessing water compressibility within a 5 Å volume of the relevant sub-regions as outlined by Garde and colleagues [136, 137]. The SASA analysis is a reliable measure of hydrophobicity for smooth hydrophobic interfaces but does not correlate perfectly with free energies of solvation of hydrophobic groups near irregular hydrophobic surfaces, which are often found in proteins.

To investigate the effect of hydrophobic interactions on the ErbB kinase conformations, we analyzed the SASA of relevant hydrophobic sub-regions, namely, the C-spine, R-spine, hydrophobic core, and the αC-β4 region. The four regions have a high percentage of hydrophobic residues; however, some minor differences exist among members of the ErbB family , particularly in the αC-β4 region. The role of hydrophobic stabilization is ambiguous with respect to the catalytically important C-spine and R-spine regions: while SASA analysis suggests that the active conformations benefit from a larger buried surface area (less exposure to solvent), the water density analysis suggests that the active and inactive conformations show very similar density fluctuations. Thus a differential effect of hydrophobic stabilization may not differentiate between the two kinase conformations with respect to the defined spine regions.

We extended our hydrophobic analysis to the asymmetric dimer interface for EGFR, HER2 and ErbB4, as the interface consists largely of hydrophobic side-chains in the N-lobe of the kinase undergoing activation (M712, I714, L768, L790, and V792 for HER2) and the C-lobe of the activating kinase (I949, Y952, M953, V956, and M960). Both SASA and fluctuation analyses of local water densities imply that the inactive monomer conformations of the ErbB kinases are preferentially stabilized through hydrophobic interactions associated with the dimer interface. Moreover, as observed consistently across the ErbB family, kinase domain dimerization further reduces the SASA of the dimer interface residues, implying that hydrophobic interactions provide a dominant impetus for dimerization.

The αC-β4 region is an unstructured span of residues between the αC-helix and the β4 sheet in RTKs. Only in HER2 is the αC-β4 region predominantly hydrophobic. Reflecting this difference, the αC-β4 region in the HER2 systems is shown to be uniquely stabilized in both the SASA and water density analysis, particularly in the inactive conformation. This feature of HER2 is thought to be responsible for its preferential association with the molecular chaperone Hsp90.

There are several small in-frame mutations found in the ErbB family kinase domains that are associated with an oncogenic cell fate. In HER2 the currently identified mutations are focused upon the αC-β4 region, which is not implicated as a part of the active site in RTKs. Instead, the activating clinical mutations in HER2 are in the αC-β4 region [14, 138], shown here to be uniquely stabilized through hydrophobic interactions. Disruption of the

hydrophobicity of the αC-β4 region could alter the binding to Hsp90 [139] and increase heterodimerization and thus, activity of HER2.

In EGFR and ErbB4, the oncogenic mutations tend to increase the basal activity level of the kinase. The two most common mutations seen in the EGFR kinase domain are L834R (L866 in HER2) and a small deletion near the N-terminus of the αC-helix. The L834R mutation, apart from directly disrupting the hydrophobic core in-lieu of the hydrophilic Arg, leads to additional coupling of the A-loop to the C-loop and the A-loop to the C-lobe . The deletion mutation directly alters the αC-helix to induce a shift toward its active conformation, although it also disrupts the normal dynamics and potential extension of the αC-helix. Most of the clinically identified mutations alter the activity in a similar fashion, though some are situated near the asymmetric dimer interface and are likely to alter kinase activity either by increasing the dimerization affinity or by reconfiguring the monomer by partially mimicking the formation of the asymmetric dimer interface. For a more complete list of mutations in EGFR and ErbB4 and their specific effects upon kinase activity, please see Shih et. al [49].

To summarize the results of the hydrophobic analysis, the active kinase conformations have a greater number of conserved intramolecular stabilizing bonds (hydrophilic specific interactions) than the inactive systems, whereas the inactive kinase conformations appear to be preferentially stabilized by hydrophobic interactions of the dimer interface region and the hydrophobic core. This observation is fully consistent with the view that, for ErbB kinases, the dominant stimuli that activate them operate by modulating the dimer interface and the hydrophobic core regions to destabilize the inactive conformation. For the wildtype systems, this finding provides strong support for the allosteric mechanism associated with formation of the asymmetric dimer interface reported by Zhang et al. [32] and Qiu et al. [31], and the recent studies of juxtamembrane domain associations [140, 141]. In this respect, the emerging view on allosteric ErbB regulation posits that, rather than forcing the protein to a new conformation, the allosteric interface guides the kinase domain to the active state through pre-existing conformational ensembles [142].

Conclusion

Given the involvement of the HER2 kinase in a variety of human malignancies, including lung, gastric and breast cancer [8, 9, 143], it is necessary to investigate the molecular mechanisms that govern kinase regulation and activation. Our MD simulations provide a framework for elucidating the behavior of the HER2 kinase domain at the atomic level, in both monomeric and dimeric forms. It is also useful to compare the molecular properties and behavior of HER2 to EGFR and ErbB4, to determine the mechanistic basis for HER2's unique mode of activation. Our computational studies of HER2 kinase characterized the principal motions of the HER2 systems and interpreted the differences in atomic fluctuations in terms of specific salt bridges and hydrogen bonds. In the active HER2 kinase, we identified an extensive bonding network among the four catalytic sub-domains which was correlated with the coupled principal motions of the sub-domains. The coupling between the A-loop and C-loop was especially pronounced, which suggests the role of the A-loop/C-loop interaction in alignment of the catalytic residues for phosphoryl transfer. In comparing HER2 to EGFR and ErbB4, we identified several hydrogen bonds that are conserved among the three ErbB

kinases in the A-loop, C-loop, and αC helix of the active systems. However, the bonding pattern in the αC-β4 loop differed considerably in HER2. As the αC-β4 region in HER2 is primarily hydrophobic, it lacks many of the hydrogen bonds that occur in the αC-β4 loop in EGFR and ErbB4. This hydrophobic patch of residues in HER2 is important in its association with the molecular chaperone Hsp90, which may regulate HER2 activity by preventing RTK dimerization. This additional layer of regulation by Hsp90 is unique to HER2; Hsp90 does not associate with either EGFR or ErbB4, as they do not possess the hydrophobic patch of residues. Unlike the other ErbB family members, the extracellular domain of HER2 is poised for dimerization in the absence of ligand binding [24], rationalizing the requirement for additional regulatory mechanisms that are unique to HER2. The weakening of hydrophobic interactions and concomitant increase in the degree of hydrogen bonding in the αC-β4 region of several clinically identified HER2 mutants alters the mutant HER2 bonding patterns in similarity to those of EGFR and ErbB4. Such mutations are expected to result in constitutive HER2 activity by perturbing the HER2-Hsp90 interaction and thus enabling uncontrolled HER2 dimerization and activation.

Since the effect of phosphorylation of the regulatory tyrosine residue, Y877, on HER2 kinase activation remains controversial, we compared simulation results for unphosphorylated versus phosphorylated forms of HER2. Our quantitative FEP simulations suggested that phosphorylation of Y877 is insufficient to promote complete kinase activation; there are likely additional molecular perturbations that are required for full activity. However, this does not exclude the possibility that phosphorylated Y877 plays a role in promoting or maintaining the active conformation. Indeed, our hydrogen-bonding analyses of the MD trajectories strongly suggest that phosphorylated Y877 contributes to the network of fastening bonds in the A-loop by bridging the hydrogen bonds at the C- and N-terminal ends of the loop, thereby considerably altering the conformational environment surrounding the A-loop. Consistent with these findings, Ishizawar et al. [144] postulated a role for phosphorylated Y877 in enhancing HER2-ErbB3 heterodimer formation by potentially changing the conformation of the kinase and engaging other molecules. Therefore it is plausible that the alteration in the A-loop conformational environment resulting from the bridging of A-loop residues by pY877 may affect recruitment of signaling mediators involved in cell proliferation and other downstream signaling pathways. The residues involved in maintaining the A-loop in its extended, open conformation (including R897, R898, and F899 in the αF helix) warrant further investigation because mutation of these residues may disrupt the fastening bonding network, and this effect may further elucidate the role of Y877-phosphorylation in HER2. These computational predictions can be experimentally tested using mutagenesis assays.

The results that we have summarized thus far apply to the HER2 kinase in its monomeric form, but it is proposed that HER2 is activated in the context of an asymmetric dimer. Thus our simulations of the HER2-EGFR heterodimer enabled us to predict the effect of the dimer interface on HER2 activation. The analysis of principal motions and hydrogen-bonding patterns in the heterodimeric system revealed the weakening of several bonds responsible for maintaining the inactive state, including bonds in the αC helix, as the αC helix directly participates in the dimerization interface. As we observed for the monomeric structures, the pattern of global motions of the kinase (in the case of the dimer, this motion included a shifting of the αC helix) could be explained in terms of individual interactions identified in our hydrogen bond analysis (ie, breakage of important inactivating bonds). However, we also

identified several interactions that persist throughout the dimer simulations and therefore may represent barriers to activation. Mapping of the activation pathway of the HER2 kinase would enable the prediction of mutations that might result in constitutive activity.

The protein kinase genes represent the most frequently mutated genes in human cancers, and several HER2 kinase domain mutants have been identified. Our computational simulations provide insight into the effect of these mutations in the αC-β4 loop region of the HER2 kinase through analysis of structural dynamics and hydrogen-bonding patterns in the kinase domain. We are also able to make predictions regarding the effects of less-studied regions, such as the αF helix, in regulation of kinase activation. Elucidation of the molecular regulatory mechanisms will help characterize structure-function relationships in the wild-type HER2 kinase, as well as predict mutations with potential for constitutive activation. Such mutations in HER2 receptors have been shown to profoundly impact the efficacy of therapies targeting HER2-mediated cancers.

Acknowledgments

We thank the lab of Mark A. Lemmon for helpful discussions. We acknowledge funding from National Science Foundation CBET-0853539 and CBET-0853389, National Institutes of Health NIBIB-1R01EB006818 and NHLBI-1R01HL087036, and computational support from National Partnerships for Advanced Computational Infrastructure NPACI-MCB060006. S.E.T. acknowledges funding through a GAANN fellowship from the UPenn Bioengineering department and an NSF graduate research fellowship.

References

[1] Schlessinger, J. (2000). Cell signaling by receptor tyrosine kinases. *Cell* 103:211-225.

[2] Linggi, B., and G. Carpenter. (2006). ErbB receptors: new insights on mechanisms and biology. *Trends in Cell Biology* 16:649-656.

[3] Yarden, Y., and M. X. Sliwkowski. (2001). Untangling the ErbB signalling network. *Nature Reviews Molecular Cell Biology* 2:127-137.

[4] Schulze, W. X., L. Deng, and M. Mann. (2005). Phosphotyrosine interactome of the ErbB-receptor kinase family. *Mol. Syst. Biol.* 1.

[5] Citri, A., and Y. Yarden. (2006). EGF-ERBB signalling: towards the systems level. *Nat. Rev. Mol. Cell Biol*. 7:505-516.

[6] Choi, S. H., J. M. Mendrola, and M. A. Lemmon. (2007). EGF-independent activation of cell-surface EGF receptors harboring mutations found in gefitinib-sensitive lung cancer. *Oncogene* 26:1567-1576.

[7] Ji, H., D. Li, L. Chen, T. Shimamura, S. Kobayashi, K. McNamara, U. Mahmood, A. Mitchell, Y. Sun, R. Al-Hashem, L. R. Chirieac, R. Padera, R. T. Bronson, W. Kim, P. A. Janne, G. I. Shapiro, D. Tenen, B. E. Johnson, R. Weissleder, N. E. Sharpless, and K. K. Wong. (2006). The impact of human EGFR kinase domain mutations on lung tumorigenesis and in vivo sensitivity to EGFR-targeted therapies. *Cancer Cell* 9:485-495.

[8] Lee, J. W., Y. H. Soung, S. Y. Kim, S. W. Nam, W. S. Park, Y. P. Wang, K. H. Jo, S. W. Moon, S. Y. Song, J. Y. Lee, N. J. Yoo, and S. H. Lee. (2006). ERBB2 kinase domain mutation in the lung squamous cell carcinoma. *Cancer Lett*. 237:89-94.

[9] Lee, J. W., Y. H. Soung, S. Y. Kim, W. S. Park, S. W. Nam, S. H. Kim, J. Y. Lee, N. J. Yoo, and S. H. Lee. (2005). ERBB2 kinase domain mutation in a gastric cancer metastasis. *APMIS* 113:683-687.

[10] Shigematsu, H., T. Takahashi, M. Nomura, K. Majmudar, M. Suzuki, H. Lee, Wistuba, II, K. M. Fong, S. Toyooka, N. Shimizu, T. Fujisawa, J. D. Minna, and A. F. Gazdar. (2005). Somatic mutations of the HER2 kinase domain in lung adenocarcinomas. *Cancer Res*. 65:1642-1646.

[11] Willmore-Payne, C., J. A. Holden, and L. J. Layfield. (2006). Detection of epidermal growth factor receptor and human epidermal growth factor receptor 2 activating mutations in lung adenocarcinoma by high-resolution melting amplicon analysis: correlation with gene copy number, protein expression, and hormone receptor expression. *Hum. Pathol*. 37:755-763.

[12] Willmore-Payne, C., J. A. Holden, and L. J. Layfield. (2006). Detection of EGFR- and HER2-activating mutations in squamous cell carcinoma involving the head and neck. *Mod. Pathol*. 19:634-640.

[13] Slamon, D. J., W. Godolphin, L. A. Jones, J. A. Holt, S. G. Wong, D. E. Keith, W. J. Levin, S. G. Stuart, J. Udove, and A. Ullrich, et al. (1989). Studies of the HER2/neu proto-oncogene in human breast and ovarian cancer. *Science* 244:707-712.

[14] Wang, S. E., A. Narasanna, M. Perez-Torres, B. Xiang, F. Y. Wu, S. Yang, G. Carpenter, A. F. Gazdar, S. K. Muthuswamy, and C. L. Arteaga. (2006). HER2 kinase domain mutation results in constitutive phosphorylation and activation of HER2 and EGFR and resistance to EGFR tyrosine kinase inhibitors. *Cancer Cell* 10:25-38.

[15] Huse, M., and J. Kuriyan. (2002). The conformational plasticity of protein kinases. *Cell* 109:275-282.

[16] Stamos, J., M. X. Sliwkowski, and C. Eigenbrot. (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.

[17] Hubbard, S. R., and J. H. Till. (2000). Protein tyrosine kinase structure and function. *Annu. Rev. Biochem*. 69:373-398.

[18] Tice, D. A., J. S. Biscardi, A. L. Nickles, and S. J. Parsons. (1999). Mechanism of biological synergy between cellular Src and epidermal growth factor receptor. Proceedings Of The National Academy Of Sciences Of The United States Of America 96:1415-1420.

[19] Bose, R., H. Molina, A. S. Patterson, J. K. Bitok, B. Periaswamy, J. S. Bader, A. Pandey, and P. A. Cole. (2006). Phosphoproteomic analysis of Her2/neu signaling and inhibition. *Proc. Natl. Acad. Sci .U S A* 103:9773-9778.

[20] Zhang, H. T., D. M. O'Rourke, H. Zhao, R. Murali, Y. Mikami, J. G. Davis, M. I. Greene, and X. Qian. (1998). Absence of autophosphorylation site Y882 in the p185neu oncogene product correlates with a reduction of transforming potential. *Oncogene* 16:2835-2842.

[21] Xu, W., X. Yuan, K. Beebe, Z. Xiang, and L. Neckers. (2007). Loss of Hsp90 association up-regulates Src-dependent ErbB2 activity. *Mol. Cell Biol*. 27:220-228.

[22] Bouyain, S., P. A. Longo, S. Li, K. M. Ferguson, and D. J. Leahy. (2005). The extracellular region of ErbB4 adopts a tethered conformation in the absence of ligand. *Proc. Natl. Acad. Sci. U S A* 102:15024-15029.

[23] Cho, H. S., and D. J. Leahy. (2002). Structure of the extracellular region of HER3 reveals an interdomain tether. *Science* 297:1330-1333.

[24] Cho, H. S., K. Mason, K. X. Ramyar, A. M. Stanley, S. B. Gabelli, D. W. Denney, Jr., and D. J. Leahy. (2003). Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. *Nature* 421:756-760.

[25] Ferguson, K. M., M. B. Berger, J. M. Mendrola, H. S. Cho, D. J. Leahy, and M. A. Lemmon. (2003). EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. *Mol. Cell*. 11:507-517.

[26] Garrett, T. P., N. M. McKern, M. Lou, T. C. Elleman, T. E. Adams, G. O. Lovrecz, M. Kofler, R. N. Jorissen, E. C. Nice, A. W. Burgess, and C. W. Ward. (2003). The crystal structure of a truncated ErbB2 ectodomain reveals an active conformation, poised to interact with other ErbB receptors. *Mol. Cell* 11:495-505.

[27] Garrett, T. P., N. M. McKern, M. Lou, T. C. Elleman, T. E. Adams, G. O. Lovrecz, H. J. Zhu, F. Walker, M. J. Frenkel, P. A. Hoyne, R. N. Jorissen, E. C. Nice, A. W. Burgess, and C. W. Ward. (2002). Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha. *Cell* 110:763-773.

[28] Wood, E. R., A. T. Truesdale, O. B. McDonald, D. Yuan, A. Hassell, S. H. Dickerson, B. Ellis, C. Pennisi, E. Horne, K. Lackey, K. J. Alligood, D. W. Rusnak, T. M. Gilmer, and L. Shewchuk. (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.

[29] Shi, F., S. E. Telesco, Y. Liu, R. Radhakrishnan, and M. A. Lemmon. (2010). ErbB3/HER3 intracellular domain is competent to bind ATP and catalyze autophosphorylation. *Proc. Natl. Acad. Sci. U S A* 107:7692-7697.

[30] Jura, N., Y. Shan, X. Cao, D. E. Shaw, and J. Kuriyan. (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.

[31] Qiu, C., M. K. Tarrant, S. H. Choi, A. Sathyamurthy, R. Bose, S. Banjade, A. Pal, W. G. Bornmann, M. A. Lemmon, P. A. Cole, and D. J. Leahy. (2008). Mechanism of activation and inhibition of the HER4/ErbB4 kinase. *Structure* 16:460-467.

[32] Zhang, X., J. Gureasko, K. Shen, P. A. Cole, and J. Kuriyan. (2006). An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell* 125:1137-1149.

[33] Bartsch, R., C. De Vries, U. Pluschnig, P. Dubsky, Z. Bago-Horvath, S. P. Gampenrieder, M. Rudas, R. M. Mader, A. Rottenfusser, C. Wiltschke, M. Gnant, C. C. Zielinski, and G. G. Steger. (2009). Predicting for activity of second-line trastuzumab-based therapy in her2-positive advanced breast cancer. *BMC Cancer* 9:367.

[34] Bartsch, R., A. Rottenfusser, C. Wenzel, K. Dieckmann, U. Pluschnig, G. Altorjai, M. Rudas, R. M. Mader, R. Poetter, C. C. Zielinski, and G. G. Steger. (2007). Trastuzumab prolongs overall survival in patients with brain metastases from Her2 positive breast cancer. *J. Neurooncol*. 85:311-317.

[35] Valabrega, G., F. Montemurro, and M. Aglietta. (2007). Trastuzumab: mechanism of action, resistance and future perspectives in HER2-overexpressing breast cancer. *Ann. Oncol.* 18:977-984.

[36] Agus, D. B., M. S. Gordon, C. Taylor, R. B. Natale, B. Karlan, D. S. Mendelson, M. F. Press, D. E. Allison, M. X. Sliwkowski, G. Lieberman, S. M. Kelsey, and G. Fyfe. (2005). Phase I clinical study of pertuzumab, a novel HER dimerization inhibitor, in patients with advanced cancer. *J. Clin. Oncol*. 23:2534-2543.

[37] Badache, A., and N. E. Hynes. (2004). A new therapeutic antibody masks ErbB2 to its partners. *Cancer Cell* 5:299-301.

[38] Baselga, J., K. A. Gelmon, S. Verma, A. Wardley, P. Conte, D. Miles, G. Bianchi, J. Cortes, V. A. McNally, G. A. Ross, P. Fumoleau, and L. Gianni. (2010). Phase II trial of pertuzumab and trastuzumab in patients with human epidermal growth factor receptor 2-positive metastatic breast cancer that progressed during prior trastuzumab therapy. *J. Clin. Oncol*. 28:1138-1144.

[39] Pohlmann, P. R., I. A. Mayer, and R. Mernaugh. (2009). Resistance to Trastuzumab in Breast Cancer. *Clin. Cancer Res*. 15:7479-7491.

[40] Burris, H. A., 3rd, C. W. Taylor, S. F. Jones, K. M. Koch, M. J. Versola, N. Arya, R. A. Fleming, D. A. Smith, L. Pandite, N. Spector, and G. Wilding. (2009). A phase I and pharmacokinetic study of oral lapatinib administered once or twice daily in patients with solid malignancies. *Clin. Cancer Res*. 15:6702-6708.

[41] Gomez, H. L., D. C. Doval, M. A. Chavez, P. C. Ang, Z. Aziz, S. Nag, C. Ng, S. X. Franco, L. W. Chow, M. C. Arbushites, M. A. Casey, M. S. Berger, S. H. Stein, and G. W. Sledge. (2008). Efficacy and safety of lapatinib as first-line therapy for ErbB2-amplified locally advanced or metastatic breast cancer. *J. Clin. Oncol*. 26:2999-3005.

[42] Tevaarwerk, A. J., and J. M. Kolesar. (2009). Lapatinib: a small-molecule inhibitor of epidermal growth factor receptor and human epidermal growth factor receptor-2 tyrosine kinases used in the treatment of breast cancer. *Clin. Ther*. 31 Pt 2:2332-2348.

[43] Smith, J. (2005). Erlotinib: small-molecule targeted therapy in the treatment of non-small-cell lung cancer. *Clin. Ther*. 27:1513-1534.

[44] Smith, N. F., S. D. Baker, F. J. Gonzalez, J. W. Harris, W. D. Figg, and A. Sparreboom. (2008). Modulation of erlotinib pharmacokinetics in mice by a novel cytochrome P450 3A4 inhibitor, BAS 100. *Br. J. Cancer* 98:1630-1632.

[45] Argiris, A., and N. Mittal. (2004). Gefitinib as first-line, compassionate use therapy in patients with advanced non-small-cell lung cancer. *Lung Cancer* 43:317-322.

[46] Ariyama, H., B. Qin, E. Baba, R. Tanaka, K. Mitsugi, M. Harada, and S. Nakano. (2006). Gefitinib, a selective EGFR tyrosine kinase inhibitor, induces apoptosis through activation of Bax in human gallbladder adenocarcinoma cells. *Journal of Cellular Biochemistry* 97:724-734.

[47] Lynch, T. J., D. W. Bell, R. Sordella, S. Gurubhagavatula, R. A. Okimoto, B. W. Brannigan, P. L. Harris, S. M. Haserlat, J. G. Supko, F. G. Haluska, D. N. Louis, D. C. Christiani, J. Settleman, and D. A. Haber. (2004). Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med*. 350:2129-2139.

[48] Paez, J. G., P. A. Janne, J. C. Lee, S. Tracy, H. Greulich, S. Gabriel, P. Herman, F. J. Kaye, N. Lindeman, T. J. Boggon, K. Naoki, H. Sasaki, Y. Fujii, M. J. Eck, W. R.

Sellers, B. E. Johnson, and M. Meyerson. (2004). EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 304:1497-1500.

[49] Shih, A., S. E. Telesco, S. H. Choi, M. A. Lemmon, and R. Radhakrishnan. (2010). Molecular Dynamics Analysis of Conserved Hydrophobic and Hydrophilic Bond Interaction Networks in ErbB Family Kinases Submitted to Molecular Biosystems.

[50] Shih, A. J., J. Purvis, and R. Radhakrishnan. (2008). Molecular systems biology of ErbB1 signaling: bridging the gap through multiscale modeling and high-performance computing. *Mol. Biosyst*. 4:1151-1159.

[51] Telesco, S. E., and R. Radhakrishnan. (2009). Atomistic insights into regulatory mechanisms of the HER2 tyrosine kinase domain: a molecular dynamics study. *Biophys. J*. 96:2321-2334.

[52] Karplus, M., A. T. Brunger, R. Elber, and J. Kuriyan. (1987). Molecular dynamics: applications to proteins. *Cold Spring Harb. Symp. Quant. Biol*. 52:381-390.

[53] Karplus, M., and J. Kuriyan. (2005). Molecular dynamics and protein function. *Proc. Natl. Acad. Sci. U S A* 102:6679-6685.

[54] Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. (2000). The Protein Data Bank. *Nucleic Acids Res*. 28:235-242.

[55] Ferrara, P., J. Apostolakis, and A. Caflisch. (2000). Computer simulations of protein folding by targeted molecular dynamics. *Proteins* 39:252-260.

[56] MacKerell, A. D., Jr., D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher III, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* 102:3586-3616.

[57] Wang, W., O. Donini, C. M. Reyes, and P. A. Kollman. (2001). Biomolecular simulations: Recent developments in force fields, simulations of enzyme catalysis, protein-protein, and protein-nucleic acid noncovalent interactions. *Annu. Rev. Biophys. Biomol. Struct.* 30:211-243.

[58] Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry* 4:187--217.

[59] Weiner, P. W., and P. A. Kollman. (1981). AMBER: assisted model building with energy refinement. *J. Comput. Chem*. 2:287-303.

[60] Scott, W. R. P. (1999). The GROMOS biomolecular simulation program package. *J. Phys. Chem. A* 103:3596-3607.

[61] Phillips, J. C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* 26:1781-1802.

[62] Humphrey, W., A. Dalke, and K. Schulten. (1996). VMD - Visual Molecular Dynamics. *Journal of Molecular Graphics* 14:33--38.

[63] Amadei, A., A. B. M. Linssen, and H. J. C. Berendsen. (1993). Essential Dynamics of Proteins. *Proteins-Structure Function and Genetics* 17:412-425.

[64] Amadei, A., A. B. M. Linssen, B. L. deGroot, D. M. F. v. Aalten, and H. J. C. Berendsen. (1996). An Efficient Method for Sampling the Essential Subspace of Proteins. *J. Biomol. Struct. Dynam*. 13:615-625.

[65] McCammon, J. A., B. R. Gelin, and M. Karplus. (1977). Dynamics of Folded Proteins. *Nature* 267:585-590.

[66] McCammon, J. A., and S. C. Harvey. 1987. Dynamics of Proteins and Nucleic Acids. Cambridge University Press, Cambridge, MA.

[67] Essmann, U. (1995). A smooth particle mesh Ewald method. *J. Chem. Phys*. 103:8577-8593.

[68] Kryshtafovych, A., and K. Fidelis. (2009). Protein structure prediction and model quality assessment. *Drug Discov. Today* 14:386-393.

[69] Dunbrack, R. L., Jr. (2006). Sequence comparison and protein structure prediction. *Curr. Opin. Struct. Biol*. 16:374-384.

[70] Fiser, A., and A. Sali. (2003). Modeller: generation and refinement of homology-based protein structure models. *Methods Enzymol*. 374:461-491.

[71] Ginalski, K. (2006). Comparative modeling for protein structure prediction. *Curr. Opin. Struct. Biol.* 16:172-177.

[72] Moult, J. (2008). Comparative modeling in structural genomics. *Structure* 16:14-16.

[73] Schueler-Furman, O., C. Wang, P. Bradley, K. Misura, and D. Baker. (2005). Progress in modeling of protein structures and interactions. *Science* 310:638-642.

[74] Blake, J. D., and F. E. Cohen. (2001). Pairwise sequence alignment below the twilight zone. *J. Mol. Biol*. 307:721-735.

[75] Dewey, T. G. (2001). A sequence alignment algorithm with an arbitrary gap penalty function. *J. Comput. Biol.* 8:177-190.

[76] Sanchez, R., and A. Sali. (1998). Large-scale protein structure modeling of the Saccharomyces cerevisiae genome. *Proc. Natl. Acad. Sci. U S A* 95:13597-13602.

[77] Bajorath, J., R. Stenkamp, and A. Aruffo. (1993). Knowledge-based model building of proteins: concepts and examples. *Protein Sci.* 2:1798-1810.

[78] Marti-Renom, M. A., A. C. Stuart, A. Fiser, R. Sanchez, F. Melo, and A. Sali. (2000). Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* 29:291-325.

[79] Aszodi, A., and W. R. Taylor. (1996). Homology modelling by distance geometry. *Fold Des.* 1:325-334.

[80] Havel, T. F., and M. E. Snow. (1991). A new method for building protein conformations from sequence alignments with homologues of known structure. *J. Mol. Biol.* 217:1-7.

[81] Srinivasan, S., C. J. March, and S. Sudarsanam. (1993). An automated method for modeling proteins on known templates using distance geometry. *Protein Sci*. 2:277-289.

[82] Laskowski, R. A., M.W. MacArthur, D.S. Moss, and J.M. Thornton. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 26:283-291.

[83] Xiang, Z. (2006). Advances in homology protein structure modeling. *Curr. Protein Pept. Sci.* 7:217-227.

[84] Ponder, J. W., Richard, F.M. (1987). Tertiary templates for protein use of packing criteria in the enumeration of allowed sequences for different structural classes. *J. Mol. Biol.* 193:775-791.

[85] Xiang, Z. W. (2001). Extending the accuracy limit of side-chain prediction. J. *Mol. Biol.* 311:421-430.

[86] .Zwanzig, R. W. (1954). High-temperature equation of state by a perturbation method. I. Nonpolar gases. *J. Chem. Phys*. 22:1420-1426.

[87] Frenkel, D., and B. Smit. 1996. Understanding Molecular Simulations. From Algorithms to Applications. Academic Press, San Diego, CA.

[88] Beveridge, D. L., and F. M. DiCapua. (1989). Free energy via molecular simulation: applications to chemical and biomolecular systems. *Annu. Rev. Biophys. Biophys. Chem.* 18:431-492.

[89] Straatsma, T. P., Berendsen, H. J. C., Postma, J. P. M. (1986). *J. Chem. Phys*. 85:6720.

[90] Beutler, T. C., A.E. Mark, R.C. van Schaik, P.R. Gerber, and W.F. van Gunsteren. (1994). Avoiding singularities and neumerical instabilities in free energy calculations based on molecular simulations. *Chem. Phys. Lett*. 222:529-539.

[91] Pitera, J. W., and W.F. van Gunsteren. (2002). A comparison of non–bonded scaling approaches for free energy calculations. *Mol. Sim*. 28:45-65.

[92] Florian, J., M. F. Goodman, and A. Warshel. (2003). Computer simulation of the chemical catalysis of DNA polymerases: Discriminating between alternative nucleotide insertion mechanisms for T7 DNA polymerase. *Journal of the American Chemical Society* 125:8163-8177.

[93] Radmer, R. J., and P. A. Kollman. (1997). Free energy calculation methods: A theoretical and empirical comparison of numerical errors and a new method for qualitative estimates of free energy changes. *Journal of Computational Chemistry* 18:902-919.

[94] Wang, J. Y., Y. Q. Deng, and B. Roux. (2006). Absolute binding free energy calculations using molecular dynamics simulations with restraining potentials. *Biophysical Journal* 91:2798-2814.

[95] Zhou, R., P. Das, and A. K. Royyuru. (2008). Single Mutation Induced H3N2 Hemagglutinin Antibody Neutralization: A Free Energy Perturbation Study. *J. Phys. Chem. B*, In Press: PMID 18975892.

[96] Suenaga, A., M. Hatakeyama, A. B. Kiyatkin, R. Radhakrishnan, M. Taiji, and B. N. Kholodenko. (2009). Molecular Dynamics Simulations Reveal that Tyr-317 *Phosphorylation Reduces Shc Binding Affinity for Phosphotyrosyl Residues of Epidermal Growth Factor Receptor*. 96:2278-2288.

[97] Chandler, D. 1987. Introduction to modern statistical mechanics. Oxford University Press, New York, NY.

[98] Chandler, D. (1978). Statistical mechanics of isomerization dynamics in liquids and the transition state approximation. *J. Chem. Phys*. 68:2959-2970.

[99] Bartels, C., and M. Karplus. (1998). Probability distribution for complex systems: Adaptive umbrella sampling of the potential energy. *J. Phys. Chem. B* 102:865-880.

[100] Roux, B. (1995). The Calculation of the Potential of Mean Force Using Computer-Simulations. *Computer Physics Communications* 91:275-282.

[101] Radhakrishnan, R., and T. Schlick. (2004). Biomolecular free energy profiles by a shooting/umbrella sampling protocol, "BOLAS". *Journal of Chemical Physics* 121:2436-2444.

[102] Radhakrishnan, R. (2007). Coupling of Fast and Slow Modes in the Reaction Pathway of the Minimal Hammerhead Ribozyme Cleavage. *Biophys. J.* 93:2391-2399.

[103] Mezei, M. (1989). Evaluation of the Adaptive Umbrella Sampling Method. *Molecular Simulation* 3:301 - 313.

[104] Szabo, A., and N. S. Ostlund. 1996. Modern Quantum Chemistry. Dover Publications, Mineola, New York.

[105] Parr, R. G., and W. Yang. 1989. Density-functional theory of atoms and molecules. Oxford University Press, Oxford.

[106] Jensen, F. 2007. Introduction to Computational Chemistry. John wiley and sons, Chichester.

[107] Warshel, A., and W. W. Parson. (2001). Dynamics of Biochemical and Biophysical Reactions: Insight from Computer Simulations. *Quart. Rev. Biophys*. 34:563-679.

[108] Warshel, A. 1989. Computer modeling of chemical reactions in enzymes and solution. John Wiley and Sons, New York.

[109] Shurki, A., and A. Warshel. (2003). Structure/function correlations of proteins using MM, QM/MM, and related approaches: Methods, concepts, pitfalls, and current progress. *Protein Simulations* 66:249-313.

[110] Senn, H. M., and W. Thiel. 2007. QM/MM methods for biological systems. In Atomistic Approaches in Modern Biology: from Quantum Chemistry to Molecular Simulations. Springer-Verlag Berlin, Berlin. 173-290.

[111] Das, D., K. P. Eurenius, E. M. Billings, P. Sherwood, D. C. Chatfield, M. Hodoscek, and B. R. Brooks. (2002). Optimization of quantum mechanical molecular mechanical partitioning schemes: Gaussian delocalization of molecular mechanical charges and the double link atom method. *Journal of Chemical Physics* 117:10534-10547.

[112] Reuter, N., A. Dejaegere, B. Maigret, and M. Karplus. (2000). Frontier bonds in QM/MM methods: A comparison of different approaches. *Journal of Physical Chemistry A* 104:1720-1735.

[113] Field, M. J., P. A. Bash, and M. Karplus. (2002). A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem*. 11:700-733.

[114] Zhang, Y., and W. Yang. (1999). A pesudobond approach to combining quantum mechanical and molecular mechanical methods. *J. Chem. Phys*. 110:46-54.

[115] Garcia-Viloca, M., and J. Gao. (2004). Generalized Hybrid Orbital for the treatment of boundary atoms in combined quantum mechanical and molecular mechanical calculations using the semiempirical parameterized model 3 method. *Theoretical Chemistry Accounts* 111:280-286.

[116] Pu, J., D. G. Truhlar, and J. Gao. (2004). The Generalized Hybrid Orbital (GHO) method for ab initio combined QM/MM calculations. *Journal of Physical Chemistry A* 108:632-650.

[117] Friesner, R. A., M.-H. Baik, V. Guallar, B. F. Gherman, M. Wirstam, R. B. Murphy, and S. J. Lippard. (2003). How iron-containing proteins control dioxygen chemistry: a detailed atomic level description via accurate quantum chemical and mixed quantum

mechanics/molecular mechanics calculations. *Coordination Chemistry Reviews* 238-239:267-290.

[118] Rega, N., S. S. Iyengar, G. A. Voth, H. B. Schlegel, T. Vreven, and M. J. Frisch. (2004). Hybrid ab initio empirical molecular dynamics: combining the ONIOM scheme with the atom-centered density matrix propagation (ADMP) approach. *J. Phys. Chem.* B 108:4210-4220.

[119] Mulholland, A. J. (2005). Modelling enzyme reaction mechanisms, specificity and catalysis. *Drug Discovery Today* 10:1393-1402.

[120] Schmidt, M. W., K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. J. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery. (1993). General atomic and molecular electronic-structure system. *J. Comput. Chem.* 14:1347-1363.

[121] Shimojo, F., R. K. Kalia, A. Nakano, and P. Vashishta. (2005). Embedded divide-and-conquer algorithm on hierarchical real-space grids: parallel molecular dynamics simulation based on linear-scaling density functional theory. *Computer Physics Communications* 167:151-164.

[122] Zhou, R., and B. J. Berne. (1995). A New Molecular Dynamics Method Combining the Reference System Propagator Algorithm with a Fast Multipole Method for Simulating Proteins and Other Complex Systems. *J. Chem. Phys*. 103:9444-9459.

[123] Kollman, P. A., I. Massova, C. Reyes, B. Kuhn, S. H. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, and T. E. Cheatham. (2000). Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Accounts of Chemical Research* 33:889-897.

[124] Glykos, N. M. (2006). Software news and updates. Carma: a molecular dynamics analysis program. *J. Comput. Chem*. 27:1765-1768.

[125] Balsera, M. A., W. Wriggers, Y. Oono, and K. Schulten. (1996). Principal component analysis and long time protein dynamics. J. *Phys. Chem.* 100:2567-2572.

[126] Arora, K., and T. Schlick. (2004). In Silico evidence for DNA Polymerase beta's Substrate-Induced Conformational Change. *Biophysical J*. 87:3088-3099.

[127] Fan, Y. X., L. Wong, J. Ding, N. A. Spiridonov, R. C. Johnson, and G. R. Johnson. (2008). Mutational activation of ErbB2 reveals a new protein kinase autoinhibition mechanism. *J. Biol. Chem*. 283:1588-1596.

[128] Citri, A., J. Gan, Y. Mosesson, G. Vereb, J. Szollosi, and Y. Yarden. (2004). Hsp90 restrains ErbB-2/HER2 signalling by limiting heterodimer formation. *EMBO Rep*. 5:1165-1170.

[129] Xu, W., X. Yuan, Z. Xiang, E. Mimnaugh, M. Marcu, and L. Neckers. (2005). Surface charge and hydrophobicity determine ErbB2 binding to the Hsp90 chaperone complex. *Nat. Struct. Mol. Biol*. 12:120-126.

[130] Kornev, A. P., N. M. Haste, S. S. Taylor, and L. F. Ten Eyck. (2006). Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proceedings of the National Academy of Sciences* 103:17783-17788.

[131] Kornev, A. P., S. S. Taylor, and L. F. Ten Eyck. (2008). A helix scaffold for the assembly of active protein kinases. *Proceedings of the National Academy of Sciences* 105:14377-14382.

[132] Choi, S. H., J. M. Mendrola, and M. A. Lemmon. (2007). EGF-independent activation of cell-surface EGF receptors harboring mutations found in gefitinib-sensitive lung cancer. *Oncogene* 26:1567-1576.

[133] Lynch, T. J., D. W. Bell, R. Sordella, S. Gurubhagavatula, R. A. Okimoto, and B. W. Brannigan. (2004). Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* 350:2129-2139.

[134] Paez, J. G., P. A. Janne, J. C. Lee, S. Tracy, H. Greulich, and S. Gabriel. (2004). EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 304:1497-1500.

[135] Pao, W., V. Miller, M. Zakowski, J. Doherty, K. Politi, and I. Sarkaria. (2004). EGF receptor gene mutations are common in lung cancers from 'never smokers' and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc. Natl. Acad. Sci. USA* 101:13306-13311.

[136] Godawat, R., S. N. Jamadagni, and S. Garde. (2009). Characterizing hydrophobicity of interfaces by using cavity formation, solute binding, and water correlations. *Proceedings of the National Academy of Sciences* 106:15119-15124.

[137] Acharya, H., S. Vembanur, S. N. Jamadagni, and S. Garde. (2010). Mapping hydrophobicity at the nanoscale: Applications to heterogeneous surfaces and proteins. Faraday Discuss. In Press DOI: 10.1039/b927019a.

[138] Stephens, P., C. Hunter, G. Bignell, S. Edkins, H. Davies, J. Teague, C. Stevens, S. O'Meara, R. Smith, A. Parker, A. Barthorpe, M. Blow, L. Brackenbury, A. Butler, O. Clarke, J. Cole, E. Dicks, A. Dike, A. Drozd, K. Edwards, S. Forbes, R. Foster, K. Gray, C. Greenman, K. Halliday, K. Hills, V. Kosmidou, R. Lugg, A. Menzies, J. Perry, R. Petty, K. Raine, L. Ratford, R. Shepherd, A. Small, Y. Stephens, C. Tofts, J. Varian, S. West, S. Widaa, A. Yates, F. Brasseur, C. S. Cooper, A. M. Flanagan, M. Knowles, S. Y. Leung, D. N. Louis, L. H. Looijenga, B. Malkowicz, M. A. Pierotti, B. Teh, G. Chenevix-Trench, B. L. Weber, S. T. Yuen, G. Harris, P. Goldstraw, A. G. Nicholson, P. A. Futreal, R. Wooster, and M. R. Stratton. (2004). Lung cancer: intragenic ERBB2 kinase mutations in tumours. *Nature* 431:525-526.

[139] Xu, W., X. Yuan, Z. Xiang, E. Mimnaugh, M. Marcu, and L. Neckers. (2005). Surface charge and hydrophobicity determine ErbB2 binding to the Hsp90 chaperone complex. *Nat. Struct. Mol. Biol.* 12:120-126.

[140] Red Brewer, M., S. H. Choi, D. Alvarado, K. Moravcevic, A. Pozzi, M. A. Lemmon, and G. Carpenter. (2009). The Juxtamembrane Region of the EGF Receptor Functions as an Activation Domain. *Molecular Cell* 34:641-651.

[141] Jura, N., N. F. Endres, K. Engel, S. Deindl, R. Das, M. H. Lamers, D. E. Wemmer, X. Zhang, and J. Kuriyan. (2009). Mechanism for Activation of the EGF Receptor Catalytic Domain by the Juxtamembrane Segment. *Cell* 137:1293-1307.

[142] Tsai, C.-J., A. d. Sol, and R. Nussinov. (2009). Protein allostery, signal transmission and dynamics: a classification scheme of allosteric mechanisms. *Molecular Biosystems* 5:207-216.

[143] Lee, J. W., Y. H. Soung, S. H. Seo, S. Y. Kim, C. H. Park, Y. P. Wang, K. Park, S. W. Nam, W. S. Park, S. H. Kim, J. Y. Lee, N. J. Yoo, and S. H. Lee. (2006). Somatic mutations of ERBB2 kinase domain in gastric, colorectal, and breast carcinomas. *Clinical Cancer Research* 12:57-61.

[144] Ishizawar, R. C., T. Miyake, and S. J. Parsons. (2007). c-Src modulates ErbB2 and ErbB3 heterocomplex formation and function. *Oncogene* 26:3503-3510.