Summer 8-12-2011

Multiscale Modeling of the ErbB Receptor Tyrosine Kinase Signaling Network Through Theory and Experiment

Shannon E. Telesco
University of Pennsylvania, shannone@seas.upenn.edu

Multiscale Modeling of the ErbB Receptor Tyrosine Kinase Signaling Network Through Theory and Experiment

Abstract
The biochemical processes occurring within a living cell span a spectrum of scales in space and time, ranging from the nano- to the macro-scale. We note that a single cellular process often operates on multiple spatial and temporal scales, and thus it becomes necessary to combine modeling techniques in multiscale approaches, in which different levels of theory are synergized to describe a system at a number of scales or resolutions. In this work we apply a multiscale modeling framework to investigate the molecular regulatory mechanisms governing the activation of the ErbB receptor tyrosine kinases, a family of kinases which are commonly over-expressed or mutated in human cancers, with a focus on the HER3 and HER4 kinases. Our multiscale model of HER3, a kinase which, until recently, has been considered kinase-dead, presents evidence of HER3 catalytic activity and demonstrates that even a weak HER3 signal can be amplified by other cellular signaling mechanisms to induce drug resistance to tyrosine kinase inhibitors *in silico*. Thus HER3, rather than the commonly-targeted EGFR and HER2 kinases, may represent a superior therapeutic target in specific ErbB-driven cancers. In the second major study, we construct a multiscale model of activity in the HER4 kinase, which has been shown to perform an anti-cancer role in certain tumor cells, by steering the cell toward a program of cellular differentiation and away from a program of uncontrolled proliferation. Our HER4 model, which applies a combined computational and experimental approach, elucidates the molecular mechanisms underlying this HER4-mediated 'switch' to the cellular differentiation program, with the ultimate aim of exploiting or modulating the HER4 pathway as a potential therapy in specific ErbB-driven cancers. Furthermore, the model provides structural insights into the effects of several HER4 somatic mutations which have recently been discovered in a subset of cancer patients, and which abrogate the anti-cancer effects of HER4 activity. We have illustrated that multiscale modeling provides a powerful and quantitative platform for investigating the complexity inherent in intracellular signaling pathways and rationalizing the effects of molecular perturbations on downstream signaling events and ultimately, on the cell phenotype.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Bioengineering

First Advisor
Ravi Radhakrishnan

Keywords
ErbB, receptor tyrosine kinase, multiscale modeling, cancer

MULTISCALE MODELING OF THE ERBB RECEPTOR TYROSINE KINASE SIGNALING NETWORK THROUGH THEORY AND EXPERIMENT

Shannon E. Telesco

A DISSERTATION

in

Bioengineering

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2011

_____

Ravi Radhakrishnan, Associate Professor of Bioengineering and Dissertation Supervisor

_____

Beth A. Winkelstein, Professor of Bioengineering and Graduate Group Chairperson

Dissertation Committee

Scott Diamond, Professor of Chemical and Biomolecular Engineering

Mark A. Lemmon, Professor and Chair, Biochemistry and Molecular Biophysics

David Meaney, Professor of Bioengineering

Rajanikanth Vadigepalli, Assistant Professor, Pathology, Anatomy & Cell Biology, Thomas Jefferson University

MULTISCALE MODELING OF THE ERBB RECEPTOR TYROSINE KINASE SIGNALING NETWORK THROUGH THEORY AND EXPERIMENT

DEDICATION

To my Mother and Father, for your unconditional love and dedication for the past 27 years. I only hope that I can lead a life as generous, fulfilling, and special as you have.

ACKNOWLEDGMENTS

I am profoundly grateful to my advisor, Ravi Radhakrishnan, for his excellent mentorship, encouragement, and insights. His gift for balancing leadership and mentorship with encouragement of independent, creative thinking in his students, and for providing broad perspective not just on the scientific process but also on the Ph.D. process, has left an indelible mark on my Ph.D. journey. He has been the most significant influence in my growth as an independent scientist, and it is a great privilege to have him as a mentor.

I would like to thank Mark Lemmon, whose dedication and energy throughout our collaboration has been infectious. During our collaboration, he has displayed his commitment to developing new avenues for harnessing the synergy between computational and experimental methods, at a time when many experimentalists are still resistant or uncertain of the best approach for incorporating computational techniques into their work. His scientific rigor and gift for teaching are an inspiration.

I am extremely grateful to Raj Vadigepalli, who has played a significant role in my Ph.D. His generosity in opening his lab to me and spending time to teach and share with me an array of tools which have significantly complemented my Ph.D. work are greatly appreciated. He is a true role model for combining computational and experimental approaches in creative ways, and has been instrumental in my scientific development.

I would also like to thank my committee members, David Meaney and Scott Diamond, for reading my thesis and providing me with helpful comments and suggestions.

I would like to thank Jason Burdick and his lab for their generosity in providing me with space and support for completing the experimental assays presented in this work.

I also wish to gratefully acknowledge those people who have contributed to my Ph.D. work or experience: Andrew Shih, Jeremy Purvis, Yingting Liu, Neeraj Agrawal, Fei Jia, Fumin Shi, Ravindra Venkatramani, Ryan Bradley, Peter Huwe, Jin Liu, Jessamine Winer, Jamie Ifkovits, Sudhir Khetan, Pam Barendt, and Santhosh Palani.

I wish to gratefully acknowledge Sharvan Kumar, who has been an incredible professor and mentor, and has gone above and beyond in providing insights and encouragement throughout many stages of my Ph.D.

I am extremely grateful to Drew, for your unconditional support during this entire process.

Finally, I am profoundly grateful to my family, Mom, Dad, Carolyn, and Kathleen, for your unconditional love and encouragement, not just during my Ph.D., but for the past 27 years. Your commitment and sacrifice in fostering successful, happy, and healthy lives for us is one of the most difficult jobs in the world, and you have gone above and beyond what any parents and sisters ever could. You are the most significant influence not only on my career, but on my life, and none of this would have been possible without you.

ABSTRACT

MULTISCALE MODELING OF THE ERBB RECEPTOR TYROSINE KINASE SIGNALING NETWORK THROUGH THEORY AND EXPERIMENT

Shannon E. Telesco

Ravi Radhakrishnan

The biochemical processes occurring within a living cell span a spectrum of scales in space and time, ranging from the nano- to the macro-scale. We note that a single cellular process often operates on multiple spatial and temporal scales, and thus it becomes necessary to combine modeling techniques in multiscale approaches, in which different levels of theory are synergized to describe a system at a number of scales or resolutions. In this work we apply a multiscale modeling framework to investigate the molecular regulatory mechanisms governing the activation of the ErbB receptor tyrosine kinases, a family of kinases which are commonly over-expressed or mutated in human cancers, with a focus on the HER3 and HER4 kinases. Our multiscale model of HER3, a kinase which, until recently, has been considered kinase-dead, presents evidence of HER3 catalytic activity and demonstrates that even a weak HER3 signal can be amplified by other cellular signaling mechanisms to induce drug resistance to tyrosine kinase inhibitors *in silico*. Thus HER3, rather than the commonly-targeted EGFR and HER2 kinases, may

represent a superior therapeutic target in specific ErbB-driven cancers. In the second major study, we construct a multiscale model of activity in the HER4 kinase, which has been shown to perform an anti-cancer role in certain tumor cells, by steering the cell toward a program of cellular differentiation and away from a program of uncontrolled proliferation. Our HER4 model, which applies a combined computational and experimental approach, elucidates the molecular mechanisms underlying this HER4-mediated 'switch' to the cellular differentiation program, with the ultimate aim of exploiting or modulating the HER4 pathway as a potential therapy in specific ErbB-driven cancers. Furthermore, the model provides structural insights into the effects of several HER4 somatic mutations which have recently been discovered in a subset of cancer patients, and which abrogate the anti-cancer effects of HER4 activity. We have illustrated that multiscale modeling provides a powerful and quantitative platform for investigating the complexity inherent in intracellular signaling pathways and rationalizing the effects of molecular perturbations on downstream signaling events and ultimately, on the cell phenotype.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Multiscale modeling of cellular processes

The biochemical processes occurring within a cell span a spectrum of scales in space and time, ranging from the nanoscale to the macroscale. As a result, many different modeling techniques are commonly used to represent the system of interest, several examples of which are highlighted in Figure 1.1. Fig. 1.1 illustrates the different levels of theory available to simulation, where the x-axis displays the range of time scales and the y-axis the various length scales. In the lower-most corner of Figure 1, we find the most accurate, yet most computationally demanding, method: electronic level, or quantum mechanics/molecular mechanics (QM/MM), simulations. At each time step of a QM/MM simulation, the interactions between the nuclei of a selected subset of atoms in a protein are computed using a quantum chemical electronic structure calculation, and therefore this type of simulation is typically used to investigate chemical reactions occurring during enzyme catalysis. At a lower level of resolution, molecules can be represented at the atomic scale in molecular dynamics (MD) simulations (Fig. 1.1), which are employed to study conformational changes in proteins. The largest MD simulations to date comprise more than one million atoms and can be run for hundreds to thousands of nanoseconds on current super-computers. The third level of theory represented in Fig. 1.1 focuses on cellular-level simulations to investigate specific signaling networks in biological systems of interest, and is most relevant for events occurring on the order of milliseconds to

hours. In this type of model it is typical for the molecular reactions to be represented as a system of ordinary differential equations (ODEs), or, if spatial information is relevant, as a system of partial differential equations, and solved numerically to produce time courses for expression levels of various model species. In this way, the interactions among various signaling components can be quantitatively investigated. The final modeling scale represented in Figure 1.1 is the genetic or transcriptional scale, which is used to analyze the interactions between transcription factors (TFs) and genes which produce the ultimate celluar phenotype. Of course, one can imagine extending the axes in Figure 1.1 to include other scales, such as the tissue or whole organism levels, but the range illustrated here represents the scales which are investigated in this thesis.

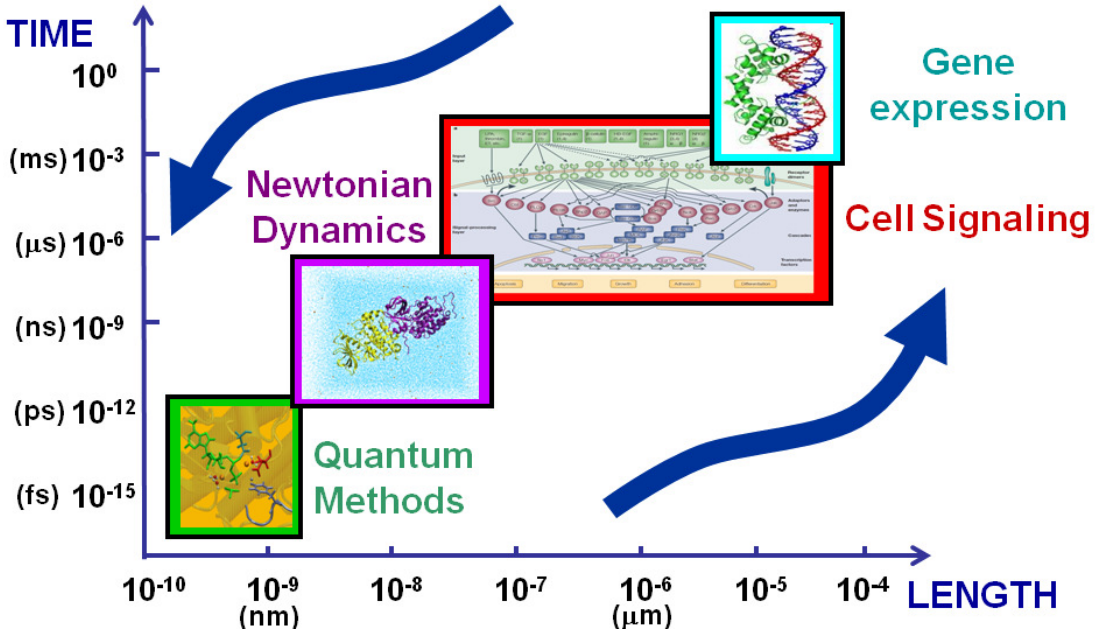

Figure 1.1. Multiscale modeling hierarchy.
The highest level of resolution applied in this thesis is the quantum, or electronic-level scale, which is used to map catalytic pathways in enzymes and proteins. The largest scale employed is that of gene expression and transcriptional activation.

We note that a single cellular process often operates on multiple spatial and temporal scales, and thus it becomes necessary to combine modeling techniques in multiscale approaches, in which different levels of theory are synergized to describe a system at a number of scales or resolutions. Multiscale modeling has been applied to a variety of biological systems [46, 143, 165, 193] to help quantify the complexity inherent in intracellular signaling networks and to provide a bridge betweent the microscopic and meso- or macroscopic time and length scales (Fig. 1.2). Multiscale modeling provides a powerful and quantitative methodology for studying the effects of molecular perturbations, such as mutation or phosphorylation of proteins, on downstream signaling events, including gene expression and the ultimate cellular phenotype.

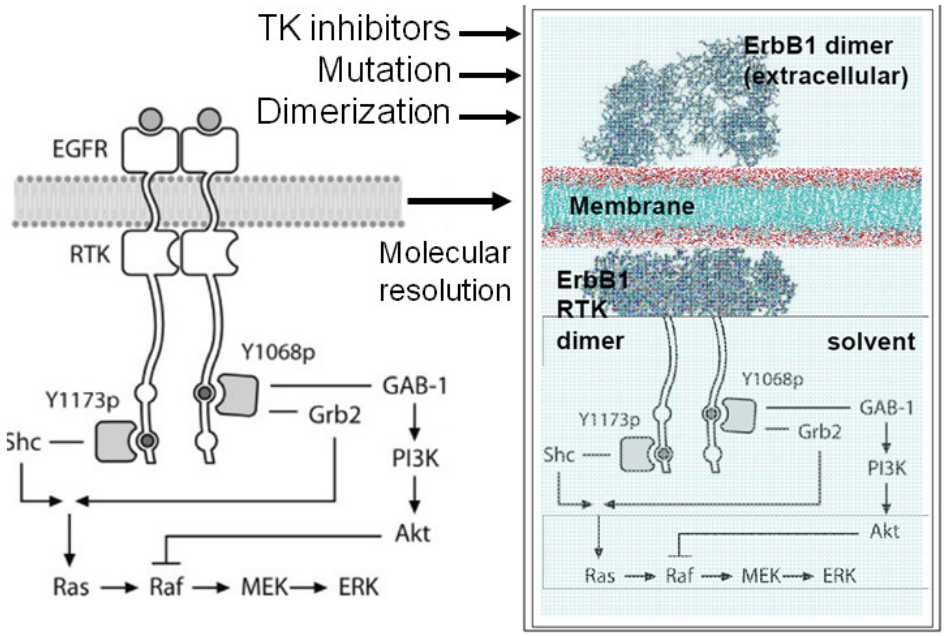

Figure 1.2. Multiscale modeling example.
On the left is a schematic of a typical signaling model; on the right, the model is displayed at increased (atomic-level) resolution to better define the effects of molecular perturbations on the selection of downstream signaling cascades.

1.2 The ErbB family of receptor tyrosine kinases

The ErbB family of kinases is a set of four homologous receptor tyrosine kinases (RTKs): EGFR/ErbB1/HER1, ErbB2/HER2/Neu, ErbB3/HER3, and ErbB4/HER4. RTKs are single-pass transmembrane proteins which play an important role in intracellular signaling by translating extracellular cues (ligands or growth factors) into activation of specific cell signaling cascades [73, 152]. In humans, there are currently 58 known RTKs divided into 20 families. The ErbB family of RTKs 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 [151]. Under physiological conditions, ligand binding promotes homo- or heterodimerization of the receptors and activation of their cytoplasmic domains [106, 206]. Dimerization results in auto- or trans-phosphorylation of tyrosine residues in the C-terminal tail segments, which serve as docking sites for signaling molecules containing SH2 or PTB domains[156]. The major signaling cascades activated by the ErbB family include the Ras-Raf-MEK-MAPK and the PI3K-Akt pathways, both of which result in transcription of genes involved in cellular proliferation, differentiation, and migration[31]. Aberrant activation of the ErbB network is frequently associated with cellular transformation and clinical malignancies such as lung, gastric, and breast cancers [29, 79, 102, 117, 164, 200]. Overexpression of HER2 results in constitutive, ligand-independent activation of kinase signaling and is found in 20-30% of human breast cancers, where it is correlated with an aggressive tumor phenotype [168, 195].

The ErbB signaling network presents a “bow-tie” architecture, in which multiple inputs and outputs are linked through a set of core processes [31]. The ErbB kinases are capable of binding to a variety of ligands, improving the flexibility and robustness of the ErbB signaling network by allowing for cross-talk between the ErbB kinase dimers and

compensating for any reduced signaling of individual ErbB members in a given cell type. For example, HER2 does not currently have a known ligand, and the kinase activity of HER3 is weaker than that of the other ErbB RTKs[18]. However, HER2/HER3 heterodimers are extremely relevant in cellular signaling, and their over-expression is correlated with oncogenic transformation in many types of cancer [2, 68]. While the response to extracellular signals occurs only through the four ErbB kinases, and is then channeled through a conserved, relatively small collection of core biochemical interactions, the ErbB network again expands through the activation of transcription factors and positive as well as negative feedback mechanisms, culminating in gene expression and the ultimate cellular phenotype (proliferation, differentiation, migration).

1.3 Structure of the ErbB RTK

Activation of the ErbB kinases initiates with the binding of ligand to the extracellular domain of the ErbB RTK. In the ErbB family, the extracellular domain is locked in an auto-inhibited state, which is then released through the ligand binding event. There are several mechanisms through which ligand binding facilitates the dimerization of RTKs, including "ligand-mediated" dimerization, in which the ligands bridge the two receptors without the receptors making direct contact, and "receptor mediated" dimerization, in which the ligands make no direct contribution to the dimer interface[104]. The ErbB family represents an extreme case of "receptor mediated" dimerization[56, 128]. The ErbB extracellular domain consists of four sub-domains, with auto-inhibitory interactions occurring between domains II and IV in a tethered conformation (Figure 1.3A)[19, 20, 26, 49]. Ligands in the ErbB family are bivalent and bind to Domains I and III, which induce a

conformational change to release the tethered conformation and expose a dimerization arm in domain II, allowing the dimerization arm to contact another ErbB RTK molecule.

Activation of the ligand-binding domain of the ErbB receptor triggers conformational changes within its cytoplasmic domain. In order to effect phosphoryl transfer of the γ-phosphate of ATP to tyrosine residues on target substrates, several key loops within the kinase domain must be appropriately positioned [75]. Residues 812 to 818 (EGFR numbering) comprise the catalytic loop (C-loop), which is crucial in facilitating the phosphoryl transfer reaction. The αC helix (residues 729-742) and the nucleotide-binding loop, or N-loop (residues 695 to 700), are responsible for coordination of the ATP and substrate tyrosine. The activation loop (A-loop) comprises residues 831 to 852 and governs activation of the kinase by regulating accessibility of the target substrate to the C-loop. The A-loop undergoes a significant conformational extension upon activation (Fig. 1.3B), uncovering the catalytic machinery and enabling binding of the tyrosine substrate to the C-loop[172].

In most protein kinases, the activation loop assumes its catalytically competent conformation only if it is first phosphorylated on a regulatory tyrosine residue within the A-loop[73], which corresponds to Y845 in EGFR (Y877 in HER2). Members of the EGFR family appear to be unique in not requiring A-loop phosphorylation for their activity[186]. However, several biochemical studies have demonstrated that phosphorylation of Y845 in EGFR significantly alters downstream signaling events, such as activation of STAT5b and EGF-induced DNA synthesis, potentially by altering the molecular environment in the A-loop and influencing recruitment of signaling mediators involved in downstream processes[15, 90]. Furthermore, Xu et al.[203] have reported 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. Therefore, it is possible that phosphorylation of the A-loop tyrosine residue augments ErbB kinase activity.

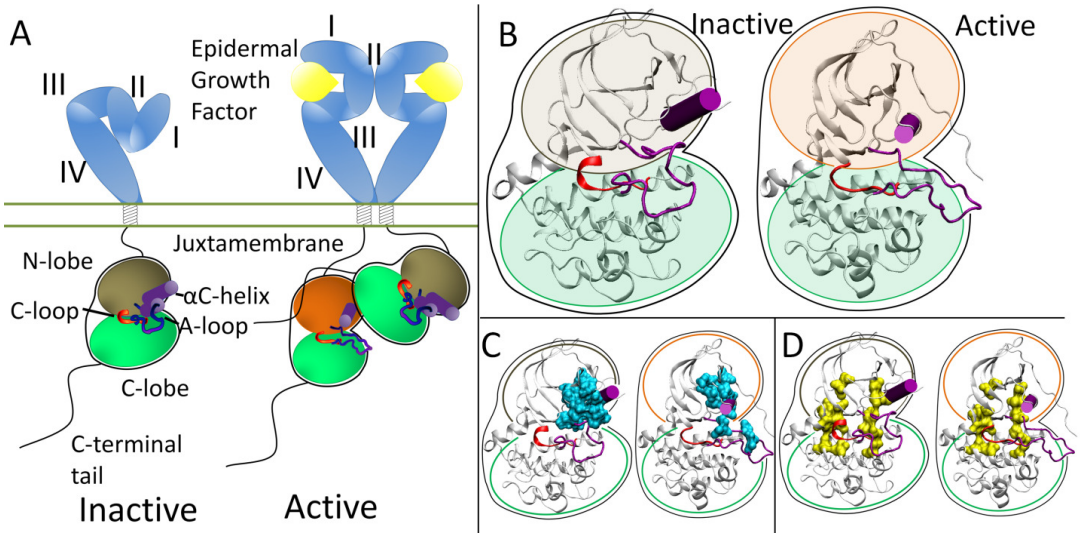

Figure 1.3. Activation scheme for the ErbB family.
(A) The inactive kinase (brown N-lobe) is auto-inhibited through the A-loop and αC-helix (purple). Introduction of the asymmetric dimer interface rotates the αC-helix to the active state (orange N-lobe). (B) Enhanced view of the inactive and active kinase domains, where the A-loop and αC-helix are shown in purple. (C) Hydrophobic core (cyan) in the inactive and active conformations. (D) C-spine (left, yellow) and R-spine (right, yellow) in the inactive and active conformations.

Recent structural studies have revealed highly conserved hydrophobic "spines" within kinases, which are considered to be important for defining their catalytic state [93, 94] (Fig. 1.3D). The regulatory spine (R-spine) consists of four hydrophobic side chains (M742, L753, H811, F832 in EGFR) anchored by an aspartic acid in the αF-helix (D872 in EGFR). The R-spine spans several key regulatory sub-domains, and coordinates the motion of the N- and C-lobes of the kinase[93]. The catalytic spine (C-spine) involves eight hydrophobic side-chains (V702, A719, L774, V819, L820, V821, T879, L883 in EGFR) which support and coordinate the adenine ring of ATP in the active state[94]. 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 (Fig. 1.3C). Disruption of this hydrophobic core by single point mutations has been shown to activate EGFR[29, 111, 131, 132, 172].

1.4 Allosteric activation of the ErbB kinase domain

Many protein kinase domains are inhibited through steric hindrance, in which a protein segment blocks the kinase active site to reduce the efficacy of the kinase. Dimerization positions two kinase domains in close proximity to each other, and, although the kinase efficiency is greatly reduced, it is theorized that each kinase exhibits sufficient activity to phosphorylate its dimerization partner. Phosphorylation prevents the protein segment from binding into the active site and allows the kinase to fully function. In the insulin[71] and FGFR[118] family of kinases, the A-loop serves as the inhibitory segment, while the juxtamembrane domain serves a similar auto-inhibitory role in MuSK [188], Flt3 [61], KIT [119], and the Eph [201] family. In the Tie2[162] kinase, a segment of the C-terminal tail acts as the auto-inhibitor.

The ErbB family of receptor tyrosine kinases employs an alternative method of auto-inhibition of catalytic activity, in that phosphorylation of the A-loop or any other protein segment does not activate the kinase. Rather than operating via steric hindrance of the active site, there are collective auto-inhibitory interactions preventing the proper coordination among key sub-domains in the kinase[91, 208]. Activation of the kinase domain is achieved through dimerization, though in this case, the dimer interface itself serves as the activating mechanism. Recent structural studies demonstrate

that EGFR and HER4 are activated through an asymmetric dimerization mechanism analogous to that of a cyclin bound to an activated cyclin-dependent kinase (CDK)[141, 208]. In this asymmetric arrangement, the C-lobe of the activator kinase interacts with the N-lobe of the CDK-like kinase, activating the CDK-like kinase through allosteric contacts. The dimer interface is dominated by hydrophobic interactions between the C-lobe of the activator kinase and the αC helix of the kinase undergoing activation. A sequence alignment of the ErbB kinase domains reveals that the residues involved in the N- and C-lobe faces of the dimer are essentially invariant among the family members, suggesting that the other ErbB kinases are likely to be activated by a similar asymmetric dimerization scheme.

Specifically, the αC-helix in the inactive ErbB kinase is rotated away from the catalytic site, preventing key interactions from forming. Introduction of the activating asymmetric dimer interface (Fig. 1.3A) forces the αC-helix to sample a different conformational space, steering the helix toward the active state. Furthermore, the juxtamembrane domain serves as a latch to facilitate the formation of the asymmetric dimer interface[84, 144, 185].

1.5 Regulation of ErbB signaling

Following the activation of RTKs, there are several mechanisms, including receptor-mediated endocytosis and cellular phosphatase activity, which modulate the length of time the kinase is active on the cell surface. Upon ligand-induced activation, RTKs are internalized, a process which removes the active RTK as well as the ligand from the cell surface[39, 142, 170]. The predominant pathway for internalization in RTKs is clathrin-

mediated endocytosis, in which the RTKs are rapidly endocytosed through clathrin-coated pits. One member of the ErbB family, HER4, utilizes an alternative method of internalization, namely, proteolytic cleavage[194], which constitutes a biochemical switch and is involved in proper cardiac and neural development [58, 65].

Protein tyrosine phosphatases (PTPs) counterbalance RTK activity by removing the phosphate group on the phosphotyrosine residues. The interplay between the RTKs and PTPs serves as an important cellular switch to regulate the full activation of the RTKs and thereby the cellular fate decisions [189]. Prior to RTK activation, PTPs are in constant activity to reduce any residual phosphorylation resulting from receptor cross-talk. Given a sufficient concentration of ligand, RTK activation overcomes local PTP activity to initiate signal propagation[145]. In some cases, ligand binding stimulates recruitment of PTPs which bind and dephosphorylate target RTKs, hence stabilizing the inactive RTK conformation at the cell surface and inhibiting further signaling [179]. The bivalent relationship between PTPs and RTKs therefore constitutes a versatile regulatory unit in cellular signaling.

Another family of cytoplasmic proteins which play a role in the activation and regulation of RTKs are the recently discovered cytohesins in EGFR [12] (and the proteins Dok7 in MuSK [77]). These proteins modulate RTK activity in both a positive and negative fashion, depending on their concentration. Increased levels of such proteins activate the RTKs without any ligand binding events, while low levels of cytohesins prevent RTK activation, even in the presence of ligand. An emerging view is that cytohesins are important in the scheme of ErbB dimerization and activation, although their specific regulatory role remains an open question.

1.6 Mutation of the ErbB kinase domain and relevance to cancer

Deregulation and mutation of RTKs have been correlated with cancer almost immediately after their discovery and purification in the early 1980s. The v-erbb oncogene in the avian erythroblastosis virus that was capable of inducing acute leukemia encoded a constituently active form of the homologous ErbB kinase protein[42]. With the growing research focus on RTKs, the correlation between deregulation of RTKs and a variety of ailments, especially cancer, has only grown stronger. Deregulation of RTKs in cancer can occur at several stages: (1) increased ligand production through enhanced local autocrine activation; (2) specific gene translocations to produce kinase fusions with altered signaling profiles; (3) RTK over-expression at the cell surface; (4) mutation of the RTK protein to modulate activity; and (5) deregulation of phosphatase and endocytosis mechanisms to increase RTK signal propagation.

Clinically identified activating kinase domain mutations have been discovered in many types of cancer (Table 1.1). The results presented in Table 1.1 are curated from the Catalog of Somatic Mutations In Cancer (COSMIC [52]), which contains a comprehensive listing of all known cancer mutations. The oncogenic mutations cluster near the kinase sub-domains known to play a role in kinase activation (Table 1.1). In Kit, the predominant clinically identified activating mutations are focused in the juxtamembrane domain and the A-loop, both of which alter the auto-inhibitory function of the juxtamembrane domain. In the FGFR family, the kinase mutations are centered on the A-loop, which serves as the auto-inhibitory segment. EGFR is *cis*-inhibited through auto-

	Lung	Colon	Skin	Breast	Prostate
EGFR	PL [111, 131] αC [111, 131, 132, 169] αCβ4 [70, 95] HC [111, 131, 132, 169] AD1 [169]	–	αCβ4 [200]	αC [35]	αC:[28]
ErbB2	αC [103] αCβ4 [21, 164, 175]	αCβ4:[103]	–	αCβ4:[103] αC:[158]	–
ErbB4	–	AL: [171]	AL:[140]	AL: [171]	–
PDGFRα	JM:[37] CT:[36]	–	–	–	–
CSF1R/Fms	–	–	–	–	–
Kit/SCFR	JM:[16]	–	JM:[9, 34, 190, 198] αC:[34, 190] AL:[9, 34, 190]	–	–
Flt3/Flk2	–	–	–	–	–
VEGFR2/ KDR	CT:[37]		–	TM:[5] AL: [5]	–
FGFR1	AL:[36]	–	–	–	–
FGFR2	–	–	JM:[57]* αC:[57]* AL:[57]*	–	–
FGFR3	–	–	AL:[64, 63, 109]	–	–
FGFR4	AL:[37]	–	–	–	–
Met	JM:[37, 92]	αC:[110]	–	–	–

Table 1.1. Clinically identified mutations in RTKs with the most lethal tumor types.

* denotes a loss-of-function mutation, JM: juxtamembrane domain, PL:P-loop, αC: αC-helix, αC-β4: αC-β4 loop, HC: Hydrophobic Core, AL: A-loop, AD1: Asymmetric Dimer Interface in ErbB family, AD2: Asymmetric Dimer Interface in Ret, CT: C-terminal Tail.

inhibitory interactions in the αC helix, and is released by the asymmetric dimer interface. The activating kinase mutants observed in EGFR-driven cancers are dominated by two mutations, which account for ~4500 of the 5000 total mutations (Figure 1.4): a point

mutation (L834R) within the hydrophobic core as well as a small in-frame deletion involving residues 747 to 751 in the αC helix. The HER2 RTK is prevented from forming heterodimers due to its association with the molecular chaperone protein Hsp90 through the uniquely hydrophobic αC-β4 region[30, 202], which is where the majority of the activating HER2 mutations occur (Figure 1.4). HER4 is not as well studied as EGFR and HER2; however, it has recently come under scrutiny as a potential therapeutic target. It is worth noting that few kinase mutations occur in the critical catalytic residues of the C-loop, as they would be likely to abolish kinase activity altogether. The increased kinase activity in the ErbB cancer mutants increases the dependency of the tumor upon the ErbB RTK; in effect, the tumor becomes "oncogenically addicted"[196] and thus inhibition of the RTK represents a viable route for cancer therapeutics. EGFR and one of its small molecule inhibitors, Gefitinib, is a canonical example of the relationship among RTKs, cancer, and targeted therapeutics. The initial discovery of Gefitinib in 1994 was met with much excitement as a potential cancer therapeutic since it would be a low-dose targeted oral cancer drug. In two phase II clinical trials of Gefitinib for advanced non-small cell lung cancer (following progression of the cancer with chemotherapy), patients exhibited symptom improvement rates of approximately 40% and 1-year survival rates of 25–35% [54, 97]. The favorable results from the phase II trials gained FDA approval for Gefitinib in 2003, prior to phase III clinical trials. However, the phase III clinical trials of Gefitinib *versus* placebo as a second-line therapy did not show any statistical significance in survival in the overall population, although there was a therapeutic benefit to a sub-group of Asian non-smokers [184]. Examination of the tumors revealed sets of mutations in the EGFR

tyrosine kinase domain [111, 131, 132]. The subset of tumors harboring these EGFR mutations are exceptionally sensitive to inhibition through Gefitinib, to the extent that Gefitinib has equal or greater efficacy than standard chemotherapy treatments in EGFR mutation-positive patients [41, 126]. There are several other small molecule tyrosine kinase inhibitors (TKIs) as well as antibodies already approved by the FDA and in use in the clinical setting (Table 1.2).

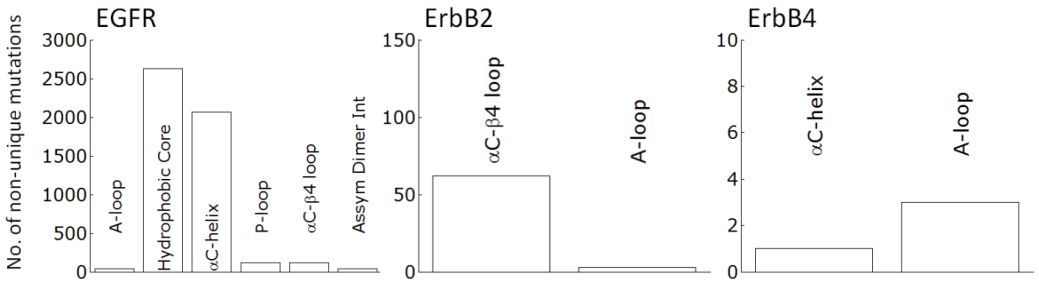

Figure 1.4. Non-unique mutations cataloged in cancer samples for the ErbB family.

Name	Target	Company	Class
Bevacizumab (Avastin)	VEGF	Genentech	Monoclonal antibody
Cetuximab (Erbitux)	EGFR	Imclone/Bristol-Meyers Squib	Monoclonal antibody
Panitumumab (Vectibix)	EGFR	Amgen	Monoclonal antibody
Ranibizumab (Lucentis)	VEGF	Genentech	Monoclonal antibody
Trastuzumab (Herceptin)	Erb2	Genentech	Monoclonal antibody
Pegaptanib (Macugen)	VEGF	OSI/Pfizer	RNA Aptamer
Dasatinib (Sprycel)	Src/Bcr-Abl	Bristol-Meyers Squib	Small molecule
Erlotinib (Tarceva)	EGFR	Genentech/OSI	Small molecule
Gefitinib (Iressa)	EGFR	AstraZeneca	Small molecule
Imatinib (Gleevec)	Bcr-Abl	Novartis	Small molecule
Lapatinib (Tykerb)	EGFR/Erb2	GSK	Small molecule
Nilotinib (Tasigna)	Bcr-Abl	Novartis	Small molecule
Pazopanib (Votrient)	VEGFR1/2/3 PDGFR/c-kit	GlaxoSmithKline	Small molecule
Sorafenib (Nexavar)	RAF/VEGFR2/PDGFRB	Onyx/Bayer	Small molecule
Sunitinib (Sutent)	VEGFR2/PDGFRB c-kit/FLT3	Pfizer	Small molecule

Table 1.2. FDA approved RTK inhibitors and antibodies currently in use.

Given the importance of the ErbB RTKs in human cancer, it is necessary to understand their activation mechanisms at the molecular level to motivate the design of therapeutics with higher specificity. This is especially important since, after sustained use of TKIs, the cancers tend to adapt through resistance mutations. In EGFR, the predominant mutation observed after extended treatment with Gefitinib is the T766M mutation[133]; the T766M resistance mutation was correctly predicted in EGFR through homology with resistance mutations seen in BCR-ABL, which was verified *in vitro* [14] and discovered several years later in patients [7, 43, 96]. Computational methodologies offer a powerful, quantitative, and complementary alternative for the study of intracellular kinase domains which, if utilized correctly, can predict resistance mutations [53].

1.7 Anticancer role of HER4 kinase

Although deregulation of the ErbB kinases is associated with many types of human cancer, one member of the ErbB family, HER4 kinase, is thought to play a beneficial role in certain cancers. In particular, HER4 expression, which is required for the differentiation of mammary epithelial cells during pregnancy, is associated with an anti-carcinogenic function and a favorable outcome in certain cellular contexts[48]. One mechanism by which HER4 is thought to impede tumor progression in mammary cells is through the activation of genes that promote cellular differentiation and inhibit proliferation, in effect, steering the cell away from a program of uncontrolled growth and instead toward a program of differentiation [80, 187]. HER4 is unique from the other ErbB receptors in that binding of the ligands neuregulin (NRG) or heparin-binding epidermal growth factor (HB-EGF) induces proteolytic cleavage of the 80 kDa kinase domain, termed the s80 or soluble cleavage product, and translocation of s80 to the nucleus, where

it complexes with the transcription factor (TF) STAT5a to regulate expression of genes involved in mammary cell differentiation pathways, including the milk protein genes β-casein and whey acidic protein (WAP)[117, 197]. Hence studies are underway to determine the molecular pathways that are stimulated by the soluble HER4 protein, particularly the network of transcriptional regulatory elements that are activated upon nuclear translocation of HER4 and STAT5a. Delineation of the transcriptional regulatory network associated with HER4/STAT5a activity would enable the exploitation of the pathway for targeting of malignant cells. Specifically, activation of HER4 signaling in aggressive breast tumors would present a novel therapeutic approach to suppress growth of these malignancies.

The HER4-mediated proliferation-to-differentiation switch in mammary cells may be encoded in terms of differential spatial and temporal regulation of HER4 activity. The phenotypic response of the mammary cell to HER4 stimulation is associated with a specific subcellular context, and the shuttling of HER4 among various compartments may determine the cellular decision to proliferate or differentiate. For instance, the membrane-bound HER4 kinase may dimerize with HER2 to produce a mitogenic response in the mammary cell, whereas the soluble HER4 domain (s80) translocates to the nucleus to effect cellular differentiation. Hence a comparison of HER4 activity at each relevant spatio-temporal scale would aid in delineation of the mechanisms by which the proliferation-to-differentiation switch is regulated.

1.8 Application of multiscale modeling to the ErbB signaling network

In this dissertation, we apply a multiscale modeling framework to investigate the molecular regulatory mechanisms governing the activation of the ErbB RTKs, with a focus on the HER3 and HER4 kinases. Chapter 2 presents a multiscale model of activity in the HER3 RTK, a kinase which, until recently, has been considered kinase-dead (see Section 2.1). In Chapter 2, we present evidence of HER3 activity and, through our multiscale model, we investigate the implications of HER3 activity for ErbB signaling dynamics and development of TKI resistance. Chapter 3 extends our analysis of HER3 to the other members of the ErbB family (EGFR, HER2, HER4) to facilitate a comparison of the activation and regulatory mechanisms across the ErbB family members. Chapters 4 and 5 present a multiscale model of activity in the HER4 RTK, beginning with molecular simulations of the HER4 dimer activation mechanism and extending the model to the cellular scale through a signaling model of the HER4 differentiation pathway in mammary cells. The HER4 multiscale model includes both a computational modeling and an experimental component. Chapter 6 presents our main conclusions and also summarizes future directions in which our work may be expanded.

Our multiscale models of ErbB kinase activity, which investigate the molecular regulatory mechanisms in the ErbB kinases, will help to elucidate structure-function relationships in drug-resistant cell lines and motivate the development of more efficacious TKI therapies targeting ErbB-mediated cancers.

Chapter 2

Multiscale Modeling of Pseudokinase Activation and Drug Resistance in the HER3/ErbB3 Receptor Tyrosine Kinase Signaling Network

Adapted from: Telesco SE, Shih AJ, Jia F, Radhakrishnan R. A Multiscale Modeling Approach to Investigate Molecular Mechanisms of Pseudokinase Activation and Drug Resistance in the HER3/ErbB3 Receptor Tyrosine Kinase Signaling Network. (2011) *Molecular BioSystems* 7(6):2066-80.

Shi F, Telesco SE, Liu Y, Radhakrishnan R, Lemmon M. The ErbB3/HER3 Intracellular Domain is Competent to Bind ATP and Catalyze Autophosphorylation. (2010) *Proc Nat Acad Sci* 107(17):7692-7.

2.1 Introduction

Approximately 10%, or 48 of the 518 protein kinases encoded by the human genome, lack at least one of three catalytic amino acid motifs in the kinase domain, including the Val-Ala-Ile-Lys (VAIK) motif in subdomain II, in which the lysine residue facilitates the positioning of the ATP molecule, the His-Arg-Asp (HRD) motif in subdomain VIb, which contains the catalytic aspartic acid residue responsible for deprotonating the substrate hydroxyl group, and the Asp-Phe-Gly (DFG) motif in the A-loop, which contains the Mg^{2+}-coordinating aspartic acid residue that orients the β and γ phosphates of ATP [18, 113]. These kinases, termed 'pseudokinases', are therefore commonly regarded as catalytically inactive. However, not all kinase domains that lack one or more of the conserved catalytic motifs are inactive. The kinase WNK1 is catalytically active despite lacking the crucial lysine in its VAIK motif, as a lysine residue in subdomain I performs the function

of the missing amino acid [116]. In a recent structural and biochemical study, Mukherjee et al. [121] resolved the crystal structure of the pseudokinase domain of Ca^{2+}/calmodulin-activated serine threonine kinase (CASK), which lacks both of the canonical Mg^{2+}-coordinating residues, and determined that CASK employs an unusual phosphorylation mechanism in which the kinase preferentially binds unchelated nucleotides, and is hence capable of catalyzing phosphotransfer despite lacking the conserved DFG motif. It is therefore plausible that other pseudokinases are not truly inactive, but may employ alternative modes of ATP binding and phosphoryl transfer.

In the ErbB family of receptor tyrosine kinases (RTKs), which includes epidermal growth factor receptor (EGFR/ErbB1), ErbB2 (HER2), ErbB3 (HER3) and ErbB4 (HER4), ligand binding followed by receptor homo- or heterodimerization performs a crucial role in regulating critical cellular processes including migration, differentiation, and proliferation [13, 31, 82, 106, 206]. EGFR, HER2, and HER4 all observe the canonical mechanism of phosphate transfer. HER3, however, is the only member of the ErbB family that has long been regarded as an inactive pseudokinase due to amino acid substitutions in the conserved kinase domain. Until recently, HER3 activity has been largely undermined, as it contains an asparagine residue (N815) in place of the catalytic aspartic acid residue that is responsible for deprotonating the substrate hydroxyl group [138]. In addition, a crucial glutamate residue in the αC helix, which indirectly participates in ATP-binding, is replaced by a histidine (H740) in HER3. Hence it has been widely believed that HER3 is dependent on interactions with the other active ErbB receptors for its biological activity.

However, it was recently demonstrated that HER3 does, in fact, exhibit robust

residual kinase activity that may be crucial for ErbB signaling. Two recent studies by Jura et al. [85] and Shi et al. [163] resolved the crystal structure of the HER3 kinase domain, revealing an inactive-like conformation. In addition, Shi et al. [163] demonstrated that, when clustered at a membrane surface, the purified HER3 intracellular domain is capable of robust autophosphorylation, albeit at a level 1000-fold weaker than the active EGFR kinase, and that the kinase domain binds ATP with an affinity similar to that of other active kinases. In the same study [163], our quantum mechanics molecular mechanics (QM/MM) simulations suggested that HER3 may catalyze phosphoryl transfer from its inactive-like kinase conformation via an alternative pathway, which explains the 1000-fold slower rate of phosphoryl transfer in HER3 compared to EGFR: the alternative pathway involves the migration of the substrate tyrosyl –OH proton directly to the O1γ oxygen of ATP, hence obviating the requirement for the catalytic aspartate residue. Our study suggests the intriguing possibility that therapeutic targeting of the robust activity of HER3, rather than the routinely-targeted EGFR and HER2 kinases, may provide a more effective treatment strategy for specific ErbB-driven cancers. However, a major question arising from our work is: what is the physiological relevance, if any, of weak HER3 activity to ErbB signaling in the cell?

This question is clinically pertinent as over-expression and activating mutations of the ErbB kinases are implicated in cellular transformation and clinical malignancies including lung and breast cancers [29, 79, 102, 117, 125, 164, 199]. Recent studies have demonstrated that mechanisms of resistance to tyrosine kinase inhibition (TKI) of EGFR and HER2 in the treatment of certain human malignancies are mediated by HER3 [44, 66, 107, 157]. In this scenario, it is hypothesized that incomplete inhibition of HER2 catalytic

activity restores phosphorylation of HER3 in the context of HER2/3 heterodimers. As HER3 contains six PI3K-binding sites in its C-tail, leaky HER3 phosphoryation induces proliferative signaling through the PI3K-AKT pathway [157]. However, the detailed molecular mechanism of this process remains an open question. Indeed, it is possible that HER3 catalytic activity is involved in the TKI resistance mechanism.

To explore the implications of HER3 activity for ErbB signaling and TKI resistance, a multiscale modeling approach is advantageous. Multiscale computational modeling has been applied to a variety of biological systems [46, 108, 143, 165, 193] to help quantify the complexity inherent in intracellular signaling networks. As the biochemical processes within a cell occur on multiple spatial and temporal scales, a multiscale modeling approach is necessary to represent a hierarchy of interactions ranging from the molecular (nm, ns) to cellular signaling (μm, ms) length and time scales. Multiscale modeling provides a powerful and quantitative methodology for studying the effects of molecular perturbations, in our case, HER3 catalytic activity, on downstream signaling events, *i.e.*, the ErbB signaling network and development of TKI resistance.

In light of the implication of HER3 in TKI resistance mechanisms, the recent interest in catalytic mechanisms of pseudokinases and particularly in the potential for HER3 activity demonstrated by Shi et al.[163], we pursue a multiscale modeling study of the HER3 kinase system at the molecular and cellular levels (Fig. 2.2). A variety of modeling techniques, ranging from atomic-level molecular dynamics (MD) simulations to systems-level modeling, are applied to investigate the non-canonical catalytic mechanism employed by the HER3 kinase and the physiological relevance of this activity to mechanisms of drug resistance in an ErbB-driven tumor cell *in silico*. Specifically, we

construct and assess several homology models of the HER3 kinase to provide a framework for comparing the ErbB kinase active sites and identifying the molecular features that distinguish HER3 from the other nodes (EGFR and HER2) in our multiscale model of HER3 activity. The results of our molecular-scale simulations support the characterization of HER3 as a weakly active kinase that, in contrast to its fully-active ErbB family members, depends upon a unique hydrophobic interface to coordinate the alignment of specific catalytic residues required for its activity. Translating our molecular simulation results of the uniquely active behavior of the HER3 kinase into a physiologically relevant environment (Fig. 2.2), our HER3 signaling model demonstrates that even a weak level of HER3 activity may be sufficient to induce AKT signaling and TKI resistance in the context of an ErbB signaling–dependent tumor cell, and therefore therapeutic targeting of HER3 may represent a superior treatment strategy for specific ErbB-driven cancers.

2.2 Materials and Methods

2.2.1 Homology modeling of the HER3 kinase domain

Several crystal structures of the ErbB kinases in the inactive conformation were selected as templates for modeling of the HER3 kinase domain. The templates included the crystal structures of EGFR (PDB code 2GS7) [208], HER4 (PDB code 3BBT) [141], as well as a combination of the EGFR and HER4 structures. The coordinates of the inactive conformations of the kinases were downloaded from the Protein Data Bank and a sequence alignment between HER3 and each template was performed in MODELLER [51, 150]. Missing residues were built using the loop modeling algorithm in MODELLER and the structures were further refined by performing additional energy minimization steps.

The hydrogen atoms were added by employing the hbuild routine in CHARMM27 [112]. The final models for the inactive HER3 kinase domain were assessed for stereochemical quality using PROCHECK [99]. The program VMD was used to display and animate all structural models [74].

2.2.2 Molecular dynamics (MD) simulations

The HER3 kinase crystal structure (PDB code 3LMG) solved by Shi et al. [163] was prepared for molecular dynamics simulation according to the procedure outlined in our previous studies[165, 181]. Briefly, the structure was explicitly solvated using the TIP3P model for water [81] and with the buffering distance set to 15 Å. Sodium (Na^+) and chloride (Cl^-) ions were added to achieve net electroneutrality of the system and an ionic strength of 75 mM. The ions were positioned at points of electrostatic extrema using a Debye-Huckel potential calculated within the program Solvate 1.0 [62]. All simulations were performed with NAMD [136] using CHARMM27 force-field parameters. To prepare the system for MD simulation, the solvated structure was energy-minimized using a conjugate gradient algorithm to remove unfavorable contacts. The system was then heated to 300 K using the temperature reassignment method in NAMD. Constant pressure and temperature (NPT) simulations were performed at 300 K and 1 atm to equilibrate the volume of the solvation box. Temperature and pressure were maintained using a Langevin piston coupling algorithm [47]. Following the NPT simulations, constant volume and temperature (NVT) simulations were performed in NAMD. Finally, a 10 ns production run was completed using the same parameters as in the NVT simulations.

2.2.3 Hydrogen bonding analysis

CHARMM was used to analyze the hydrogen bonds present in the 10 ns trajectory for

each system according to the procedure outlined in our previous studies [165, 181]. Hydrogen bonds were defined by a bond length cutoff of 3.4 Å and an angle cutoff of 150º. Bonds that fulfilled these criteria and were present in at least 60% of the trajectory were tabulated in CHARMM. Salt bridges were defined as hydrogen bonds occurring between an acidic and a basic residue and satisfying a bond length cutoff of 1.6 Å. All hydrogen bonds and salt bridges were also visualized in VMD [74] for the duration of the 10 ns simulation.

2.2.4 Principal component analysis

A principal component analysis (PCA) was applied to the 10 ns trajectory to identify the main eigenvectors (3*N* directions) along which the majority of the complex motion is defined. The calculation is based on the diagonalization of the variance-covariance matrix of the atomic fluctuations along each MD trajectory to yield the set of eigenvectors (PCs) and associated eigenvalues. The eigenvectors represent the independent modes of atomic motion, and the eigenvalues reflect the contribution of the corresponding eigenvectors to the global fluctuation of the protein. PCA computes the covariance matrix:

$$\sigma_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle$$

where (*i,j*=1,…,3*N*), and *N* is the total number of atoms with positions given by Cartesian coordinates *x*. The resulting matrix is diagonalized to compute the 3*N* independent (uncorrelated) eigenvectors, $\{\xi_i\}$, and the eigenvalues, $\{\lambda_i\}$, sorted in descending order. The angle brackets denote the time average over the entire trajectory. PCA was performed on the entire protein backbone (*i.e.* all C_α atoms), and also an active site region that comprises all catalytic sub-domains, including the A-, C-, and N-loops and the αC helix. The program CARMA [59] was used to project the atomic fluctuations along the

MD trajectory.

2.2.5 Solvent accessible surface area (SASA) and water density fluctuation analysis

SASA values were calculated in VMD [74] using the measure SASA module with a probe radius 1.4 Å larger than the van der Waals radius. The SASA was calculated for each step in the trajectory, from which the mean and standard deviation were computed. As an alternative measure of hydrophobicity in heterogeneous environments, following the procedure established in [1, 60], normalized water density fluctuations were computed by recording the ratio of $vol_N * \sigma_N^2/\langle N \rangle^2$, where vol_N, σ_N and $\langle N \rangle$ are the control volume of interest, the standard deviation, and mean associated with the number of water molecules within 5 Å of a specified hydrophobic sub-region in the HER3 kinase. Although results are presented for a cutoff of 5 Å, other cutoffs ranging from 3-15 Å were investigated and similar trends in $vol_N * \sigma_N^2/\langle N \rangle^2$ were recorded.

2.2.6 Systems model of ErbB signaling

The computational systems model was derived from that of Schoeberl et al. [154]. Mass-action reactions describing ligand-induced ErbB receptor homo- and heterodimerization, receptor internalization and degradation, constitutive dimerization, and activation of the PI3K-AKT signaling pathway were included. The HER3-specific ligand neuregulin-1β (NRG-1β) was assumed to induce EGFR-HER3, HER2-HER3 and HER3-HER3 dimers. To represent HER3 catalytic activity, we implemented HER3 homodimerization and *trans*-autophosphorylation, resulting in induction of PI3K-AKT signaling. HER3-mediated kinase phosphorylation was assumed to be 1000-fold slower than that of EGFR, based on previous experiments and QM/MM simulations [163]. The ErbB kinase inhibitor lapatinib was implemented in the model according to Schoeberl et al. [154]; lapatinib was

assumed to inhibit activation but not dimerization or ligand binding of the EGFR and HER2 kinases. All simulations were performed in MATLAB 7.10 (MathWorks, Natick, MA) and sensitivity analysis was conducted using SBToolbox 2.1 [153] and SBML-SAT [211]. Normalized, time-integrated sensitivities of each nonzero species with respect to pAKT were calculated by varying each species (one at a time) and simulating the perturbed model response. The sensitivity analysis was performed according to the procedure implemented in SBML-SAT[211]. Briefly, the normalized sensitivity, S_{ij}, was computed:

$$S_{ij} = \frac{\partial O_i / O_i}{\partial p_j / p_j}$$

where O_i is the time-integrated response of the *i*th model output (such as pAKT level) and p_j is the *j*th parameter (rate constant or initial condition). The SBML version of the model has been submitted to the BioModels Database[100, 105]. The signaling process diagram (Figure 2.2C) was constructed using the Systems Biology Graphical Notation (SBGN)[101].

2.2.7 Quantum mechanics molecular mechanics (QM/MM) simulations

ATP and a 7-residue peptide representing the Y1068 phosphorylation site in EGFR were placed in the TKD based on its superposition with the EGFR TKD structure (PDB code 2GS6) determined in complex with a bisubstrate analogue of ATP and peptide [208]. Two Mg^{2+} ions were then placed in complex with ATP based on the structure of protein kinase A (PKA) in its active conformation (PDB code 1ATP) [209]. The ground state of each ternary complex was obtained from energy minimizations of these model structures to ensure uniformity in reactant states across all systems prior to the QM/MM simulations.

The resulting models were then prepared for QM/MM simulations in which the quantum region contained the two Mg^{2+} ions, water molecules within 5 Å of the Mg^{2+} ions, segments of the ATP and peptide substrate and two catalytic residues (N815 and D833 in HER3, D813 and D831 in EGFR). The system was subjected to 1200 steps of the adopted basis Newton-Raphson minimization. For QM/MM molecular dynamics simulations, the system was first heated to 300 K and subjected to constant temperature dynamics using a Langevin thermostat for 10 ps using a 1 fs step of integration. We describe the reaction pathways in terms of several simple artificial reaction coordinates, χ_j, as described [147]. For phosphoryl-transfer through the associative mechanism, χ_1 is the distance between the reactive tyrosyl oxygen and the ATP γ-phosphate; χ_2 is the distance between the tyrosyl O- and ATP O2/3β, and χ_3 is the ATP P_γ- ATP O2/3β distance. For phosphoryl-transfer through dissociative mechanisms, in addition to these three distances, χ_4 includes the coordinate for proton abstraction, namely the distance between the tyrosyl OH- and D830:Oδ_2. Restrained minimization as well as restrained sampling simulations were performed along χ using the QM/MM Hamiltonian with a potential bias term. Reaction paths were computed by energy minimizations in the presence of restraints and recalculating single-point energies in their absence.

2.3 Results

2.3.1 QM/MM simulation of the HER3 phosphoryl transfer mechanism

Based on the experimental observations of HER3 phosphorylation[163], we hypothesized that the weak kinase activity of HER3 might utilize a mechanism distinct from that seen in other kinases, and might be carried out by the 'inactive-like' state. Activity from the

configuration seen in the HER3 TKD crystal structure is suggested by the finding that a V836A mutation prevented ATP binding and HER3 ICD autophosphorylation, in contrast with the activating effects of analogous mutations in EGFR and HER4. The lack of sequence conservation with other ErbB kinases in the N-lobe region, which would make 'receiver' contacts in the asymmetric dimer described by Zhang et al. [208], also argues that HER3 regulation is unique in this family. We therefore undertook quantum mechanics molecular mechanics (QM/MM) simulations to delineate the energy landscape of phosphoryl transfer from ATP to a substrate tyrosine catalyzed by the HER3 structure, and compared the outcomes with those seen in parallel computational studies of EGFR.

Experimental and theoretical studies have demonstrated that, in phosphoryl transfer by kinases and polymerases, nucleophilic attack on the target phosphate proceeds via a conformation that resembles a trigonal-bipyramidal transition state [98, 174], and that phosphoryl transfer can occur through either an associative or a dissociative mechanism. A conserved aspartate (D813 in EGFR, D166 in protein kinase A) is proposed to function as a base acceptor for proton transfer from the hydroxyl group of the substrate, as depicted by the red arrows in Figure 2.1A (pathway I). The presence of an asparagine at this location in HER3 (N815) precludes this pathway, and requires that proton transfer occurs via an alternative mechanism (Figure 2.1B). Indeed, our QM/MM simulations reveal an alternative pathway in HER3 and EGFR (pathway II) that is characterized by migration of the substrate tyrosyl –OH proton to the $O1_{\gamma}$ oxygen of ATP, and subsequently to the ATP $O2_{\beta}$ oxygen (green arrows in Figures 2.1A and B). Phosphoryl transfer in HER3 coincides with this proton transfer step, and occurs exclusively through

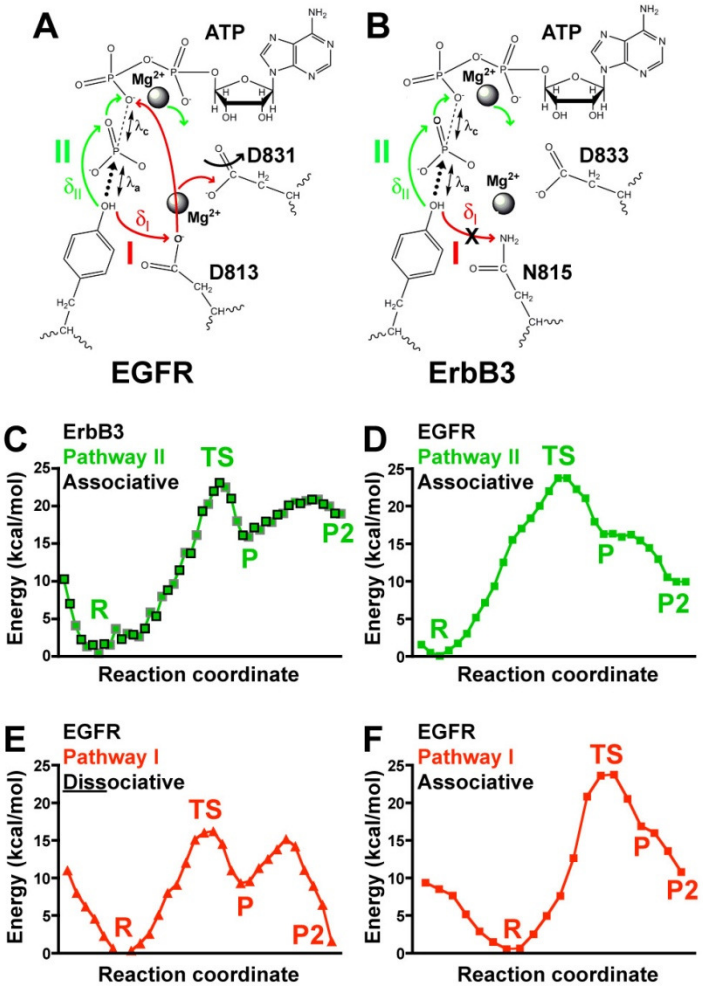

Figure 2.1. Schematic QM/MM pathway for phosphoryl transfer in EGFR and HER3. Mg^{2+} ions are marked, as are the catalytic aspartates (D831 in EGFR, D833 in HER3), the proposed catalytic base in EGFR (D813) and its replacement in ErbB3 (N815). Two potential pathways for proton migration are shown. Pathway I (red) involves proton abstraction from the substrate tyrosine –OH group by D813 of EGFR, and pathway II (green) involves proton migration to the $O1_\gamma$ oxygen of ATP. δ_I and δ_{II} denote distances between the proton of the substrate tyrosine hydroxyl and either the $O_{\delta 2}$ oxygen of D813 in EGFR δ_I or the $O1_\gamma$ oxygen of ATP (δ_{II}). The nucleophilic attack distance, λ_a (between the tyrosine oxygen and the ATP P_γ), and the bond cleavage distance, λ_c (distance between the ATP P_γ and ATP $O2/3_\beta$), are noted.

(*C-F*) Energy changes along the reactions involving Pathway II (green) or Pathway I (red). States correspond to 'R' (reactant); 'TS' (transition state with trigonal bipyramidal geometry around P_γ); 'P' (product representing the completion of phosphoryl-transfer with proton bound to $O1_\gamma$ of ATP; and 'P2' (product with proton transferred to $O2_\beta$ of ATP). (*C*) Energy changes for ErbB3. Symbols bounded by black squares represent the forward scan; gray squares represent the reverse scan. (*D*) Energy changes for EGFR utilizing pathway II. (*E*) Phosphoryl-transfer catalyzed by EGFR via the dissociative mechanism (utilizing pathway I for proton migration – via D813). (F) Associative phosphoryl-transfer comcomitant with pathway I for EGFR has an estimated E_a of 24 kcal/mol.

an associative mechanism, with characteristic values of 1.9 Å for the formation (λ_a) and cleavage (λ_c) of bonds in the transition state, and an estimated activation energy (E_a) of 23 kcal/mol (Figure 2.1C). Parallel QM/MM simulations of EGFR in its active configuration showed that pathways I and II (which are mutually exclusive) can both be utilized for transfer of the tyrosyl –OH proton, and compete with one another. Phosphoryl transfer concomitant with proton transfer through pathway II in EGFR occurs only through an associative transition state, with an estimated E_a of 24 kcal/mol (Figure 2.1D) – equal to that seen for HER3. When proton transfer occurs through pathway I, which involves proton abstraction by D813, the estimated E_a is significantly lower for the dissociative mechanism (Figure 2.1E), at 16 kcal/mol, but is 24 kcal/mol for the associative mechanism (Figure 2.1F).

Thus, our QM/MM simulations argue that phosphoryl transfer can occur in HER3 (and indeed EGFR) without abstraction of the tyrosine –OH proton by the catalytic base aspartate. The proton instead migrates to the $O1_\gamma$ oxygen of ATP, through pathway II in Figure 2.1B. Phosphoryl transfer concomitant with this pathway can be catalyzed by HER3 in the 'inactive-like' conformation, although it is predicted to be several orders of magnitude slower than the most favorable reaction channel in EGFR (*i.e.,* for pathway I through a dissociative mechanism), which is consistent with our previous experimental observations. The finding in our QM/MM studies that this mechanism can also operate in EGFR further suggests that mutating D813 in this receptor may not completely abolish kinase activity. Indeed, a D813A-mutated variant of EGFR has been reported to retain its ability to promote EGF-dependent DNA synthesis and MAP kinase activation despite exhibiting greatly reduced receptor autophosphorylation [33]. A low level of

autophosphorylation in D813A-mutated EGFR, similar to that shown for HER3, may therefore be sufficient to mediate certain key aspects of its signaling.

2.3.2 Topology of the multiscale model of HER3 activity

Our experimental and computational study[163], which elucidated the robust activity of the HER3 kinase and suggested an alternative catalytic pathway in HER3 (see discussion of QM/MM simulations in Section 2.3.1), raised the possibility that HER3 may play an active role in ErbB signaling dynamics. To investigate the relevance of our results to ErbB activity in a cellular context, we propose a pathway model of the HER3 signaling network (Figure 2.2A), in which ligand stimulation of the HER3, EGFR, and HER2 RTKs results in induction of the PI3K-AKT cascade. The ligand-induced coupling of the EGFR, HER2, and HER3 nodes to the PI3K-AKT pathway has been extensively validated computationally and experimentally[13, 22-24, 146, 176, 178], although thus far HER3 has been postulated to play a passive role in the ErbB-AKT signaling network, in that its phosphorylation (and hence, recruitment of PI3K/AKT) depends upon the catalytic activities of the EGFR and HER2 RTKs. Here we propose that HER3 can activate independently of its ErbB family members, a hypothesis which is reflected in the topology of our HER3 signaling pathway (Fig. 2.2A), and is based on a combination of recent *in silico*, *in vitro*, and *in vivo* evidence of HER3 activity. Namely, the *in vitro* kinase assays performed by Shi et al. [163] demonstrated that, in the absence of its ErbB family members, the purified HER3 kinase is capable of robust autophosphorylation, a result that was further supported by our QM/MM simulations of an alternative catalytic pathway in HER3. Additional *in vitro* evidence of HER3 activity derives from

observations of tyrosine phosphorylation of ligand-induced HER2 and HER3 RTKs in cells that express only these two ErbB receptors[146]. Given current models of ErbB receptor activation, in which one kinase domain in an ErbB dimer allosterically activates its neighbor and itself becomes *trans*-phosphorylated[83, 208], it is not clear how HER2 could be *trans*-phosphorylated in HER2-HER3 heterodimers through this mechanism unless HER3 has kinase activity of its own. Furthermore, previous experimental studies have demonstrated that HER3 is a key mediator of resistance to various tyrosine kinase inhibitors (TKIs) currently in use [8, 44, 76, 115, 157], although the resistance mechanism remains an open question. In particular, Sergina et al. [157] reported observations of HER3-mediated resistance and pAKT signaling in various TKI-treated tumor cell lines as well as *in vivo*.

In the current study, we aim to further substantiate the inclusion of HER3 as a uniquely and independently-active node in our ErbB pathway model (Fig. 2.2A), in order to quantitatively evaluate the implications of HER3 activity for ErbB signaling dynamics and mechanisms of HER3-mediated drug resistance in an ErbB-driven tumor cell *in silico* (Fig. 2.2C). Our multiscale model of HER3 activity begins with atomic-level simulations of the HER3 kinase crystal structure (Fig. 2.2B), in order to identify the molecular features which distinguish HER3 from the other nodes (EGFR, HER2) in our proposed ErbB signaling network. Moreover, our molecular simulations address clinically-relevant questions regarding the mechanism and specific types of interactions that enable HER3 to maintain its robust activity.

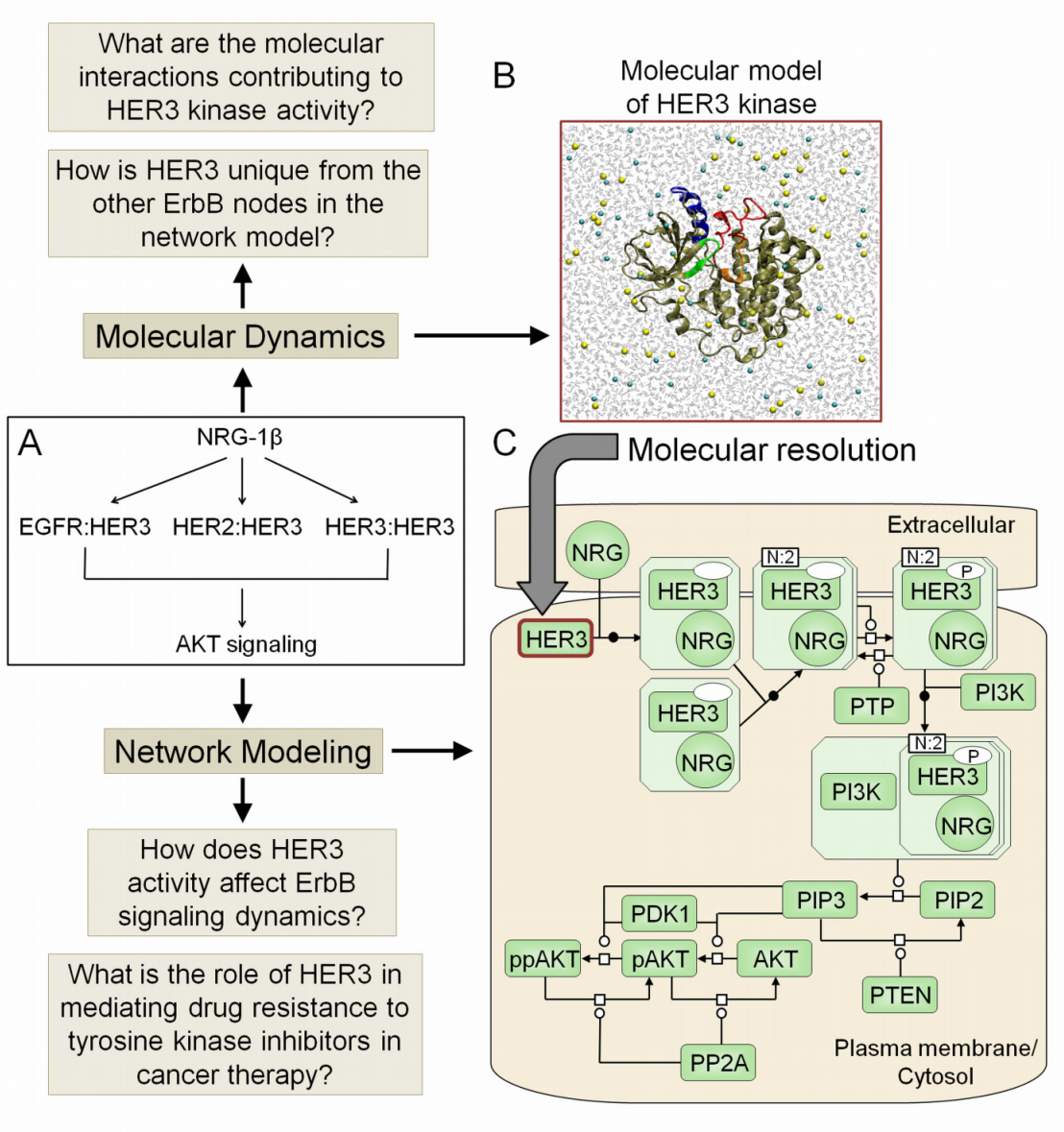

Figure 2.2. Representation of the multiscale model of HER3 activity.
(A) Schematic of the HER3 network model topology, in which ligand stimulation of EGFR:HER3, HER2:HER3, and HER3:HER3 dimers induces the AKT cascade. (B) The HER3 node in (A) is examined at molecular resolution. The molecular model comprises two parts: homology modeling to refine the HER3 kinase crystal structure, and molecular dynamics simulations of the refined HER3 structure to identify the molecular features which distinguish HER3's unique mechanism of activity from that of the EGFR and HER2 nodes in the HER3 network model. (C) Process diagram of the HER3 network model in SBGN notation (see Methods). The aim of the HER3 network model is to investigate the implications of HER3 activity for ErbB signaling dynamics and mechanisms of HER3-mediated drug resistance in an ErbB-driven tumor cell *in silico*. Note that, for clarity, only HER3 dimers are illustrated in (C), although EGFR:HER3 and HER2:HER3 dimers are also present in the network model.

2.3.3 Homology modeling of the HER3 kinase domain

As the recently-resolved HER3 crystal structure[85, 163] is missing several amino acid residues in catalytically important sub-domains, we applied a homology modeling algorithm to build in these residues and construct a complete HER3 kinase structure for our molecular simulations[182]. An additional advantage of our homology modeling analysis is that it provides a means of deriving insights into a structural comparison of the ErbB kinase active sites and the molecular features which distinguish HER3 from the other nodes (EGFR and HER2) in our HER3 pathway model (Fig. 2.2A). As the HER3 kinase domain shares a relatively high level of sequence identity with its ErbB family members (54% and 63% identity with EGFR and HER4, respectively) (Fig. 2.3), and the crystal structures of these kinases have been determined [141, 172, 208], we selected EGFR (PDB code 2GS7) [208] and HER4 (PDB code 3BBT) [141] as templates for homology modeling of the HER3 kinase in the inactive-like conformation. In addition, we applied the multiple templates (MT) algorithm in MODELLER[150] to generate a HER3 kinase model based on a combination of the EGFR and HER4 crystal structures. To verify that EGFR and HER4 were the most suitable templates available for modeling HER3, a search was performed in MODELLER's internal database for other proteins that might have high sequence identity with HER3. The ErbB receptor kinases were confirmed to be the most homologous structures for modeling of the HER3 kinase domain.

The protein sequence selected for alignment of the kinase domains included residues 678-957 (EGFR) and 683-962 (HER4); we opted to exclude the flexible C-tail from the alignment, as its sequence is highly variable among the ErbB kinases. A total of 50 models were generated from each of the templates (EGFR, HER4, and multiple

templates) by satisfying a set of static and dynamic spatial restraints in MODELLER. In addition, we applied the loop-modeling algorithm in MODELLER[51] to remodel the flexible A-loop (residues 833-855 in HER3), as the HER3 crystal structure is lacking these amino acids. The top models (shown in Fig. 2.4) were selected based on stereochemical quality and the Discrete Optimized Protein Energy (DOPE) method [161], which is an atomic distance-dependent statistical potential optimized for model assessment in MODELLER.

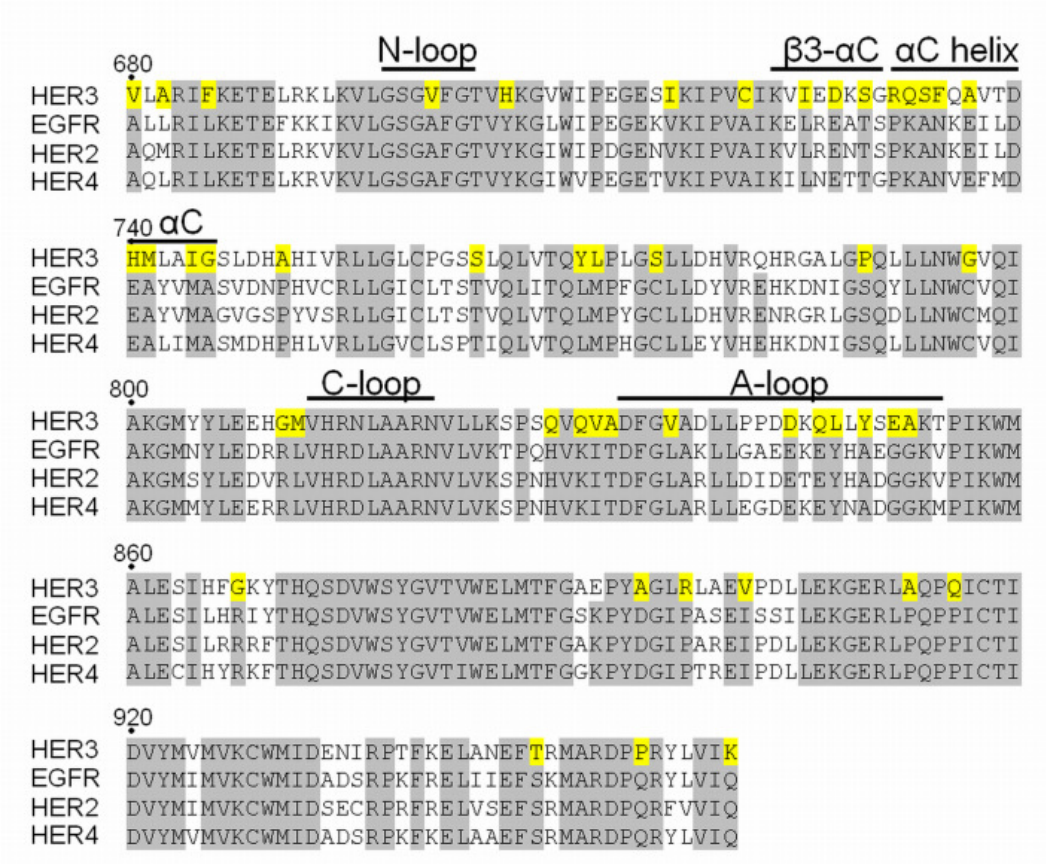

Figure 2.3. Multiple sequence alignment of the kinase domains of the ErbB RTK family. Residues highlighted in gray are conserved among all four ErbB kinases, whereas residues highlighted in yellow are unique to HER3. The HER3 kinase shares a relatively high level of sequence identity with its family members (54%, 55% and 63% identity with EGFR, HER2, and HER4, respectively). The key catalytic sub-domains (A-loop, C-loop, N-loop, αC-helix, and β3-αC loop) are labeled.

A comparison of the most energetically favorable models derived from each template revealed several similarities in structure and in DOPE energies: overall, the C-lobes of the kinase models are similar, as are the DOPE profiles, although the DOPE scores in the A-loop region are lowest for the HER4- and MT-based models (Fig. 2.4B). However, in comparing the top homology-modeled structures to the HER3 crystal structure, we observed several conformational differences, which can be better understood by first describing the major structural features that distinguish the HER3 crystal structure from the EGFR and HER4 structures. A unique feature of the catalytic site in the HER3 crystal structure is the truncated N-terminus of the αC helix (Fig. 2.4A); in the EGFR and HER4 kinase crystal structures, the helix is fully formed. The molten terminus of the HER3 αC helix forms a loop, referred to as the β3-αC loop, which interacts with specific sub-regions of the catalytic site, including the short 3_{10} helix in the A-loop, to form a hydrophobic interface that maintains the weakly-active conformation of HER3 and is not observed in EGFR or HER4 [180]. F734, which is located in the β3-αC loop and corresponds to an Asn in EGFR and HER4, appears to nucleate the hydrophobic interface (Fig. 2.4A) [180], and V836, L839 and L840, which are located in the 3_{10} helix, form hydrophobic contacts with the truncated αC helix. The extensive set of residues stabilizing the hydrophobic interface and the 3_{10} helix are unique to HER3 and function to maintain the distinctive catalytic site conformation resolved in the HER3 crystal structure.

In comparing our top homology-modeled structures to the HER3 crystal structure, we focused on the unique set of interactions in the β3-αC loop, αC helix and 3_{10} helix (Fig. 2.5A). The side-chain conformations of the aliphatic residues in the 3_{10} helix,

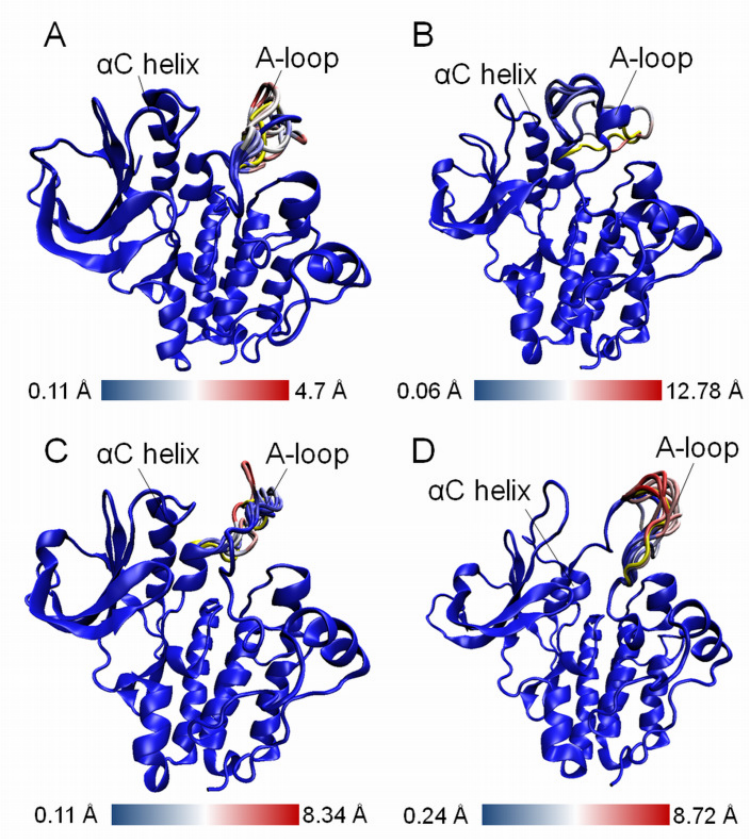

Figure 2.4. The top HER3 homology models.
The structures were based on (A) the EGFR template, (B) the HER4 template, (C) Multiple templates and (D) the loop-modeled HER3 crystal structure. The structures are color-coded according to the RMSD, where red regions indicate large RMSD values and blue regions represent small RMSD values (closely aligned structures). The original unrefined model for each HER3 structure is shown in yellow. The top structures form a dominant cluster of conformations in each HER3 model, indicating a pronounced energy minimum and a higher level of accuracy in the best structural prediction.

including V836, L839, and L840, were correctly predicted in the models derived from EGFR, HER4, and MT. This result is not entirely surprising, as the inactive EGFR and HER4 templates also contain the 3_{10} helix in the A-loop. However, F734, which serves to nucleate the hydrophobic interface in the HER3 crystal structure, points away from the hydrophobic pocket in the homology models. The reason for the altered F734 conformation is depicted in Fig. 2.5A, which illustrates that the homology models contain a fully formed αC helix in place of the truncated helix and extended β3-αC loop (which

contains F734) in the HER3 crystal structure. In addition, the H740 ring in the αC helix, which stabilizes the 3_{10} helix in the HER3 crystal structure via contacts with G835 in the DFG motif, is flipped away from G835 in the homology models. The D838-R814 bond, which bridges the 3_{10} helix and the C-loop in the HER3 crystal structure, is conserved in the HER4- and MT-based models, whereas the residues are unbonded in the EGFR-based model. The HER4-based model more closely mimics the extended β3-αC loop in the HER3 crystal structure, as the apex of the loop projects away from the catalytic site.

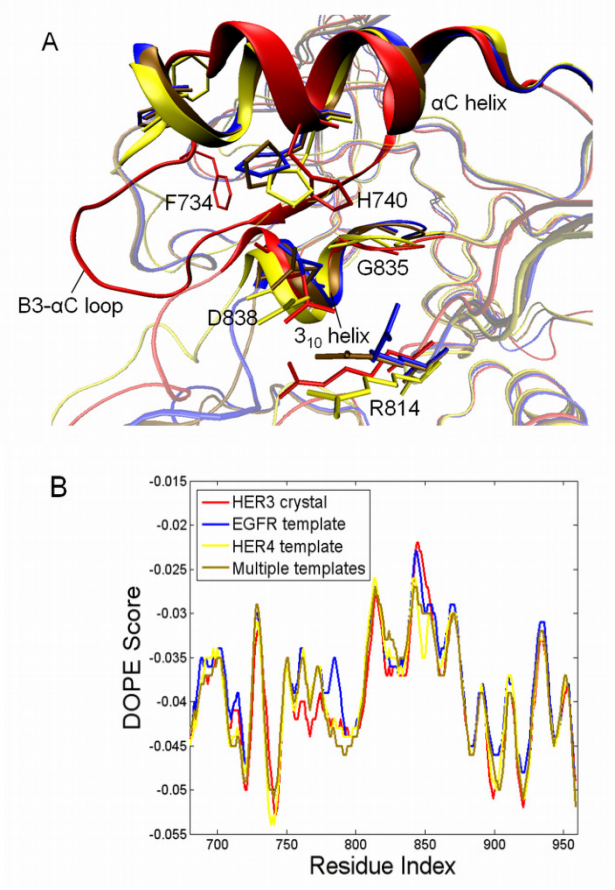

Figure 2.5. Superposition of the top HER3 models constructed from the ErbB templates. The following templates were used: EGFR (*blue*), HER4 (*yellow*), multiple templates (*brown*), as well as the HER3 crystal structure (*red*) (PDB code 3LMG). Key residues contributing to the hydrophobic interface in HER3 are highlighted. (B) DOPE energy plots for the top HER3 models based on each ErbB template after A-loop refinement. The HER4- and MT-based models exhibit lower DOPE energies in certain regions, including the sequence spanning residues 775-800. The DOPE energy plot for the HER3 crystal structure (PDB code 3LMG) is illustrated for reference.

Although none of the homology models correctly predict the molten αC helix, several features of the HER4- and MT-based models, including the D838-R814 bond, approximate the HER3 crystal structure more closely than does the model derived from the EGFR template. Additionally, the DOPE plots for the top structures (Fig. 2.5B) reveal smaller (more favorable) DOPE energies for the HER4- and MT-based models, especially for residues 775-800, which exhibit an energy peak in the EGFR-based model, and residues 840-860 in the A-loop. These results may be attributed to the closer evolutionary ties between HER3 and HER4, as HER3 and HER4 are thought to have emerged from a gene duplication event separate from the one that led to EGFR and HER2 [173]. Indeed, Jura et al. [85] postulate that the N-lobe dimer observed in their resolved HER3 and HER4 crystal structures but not in their EGFR structure may be a consequence of the evolutionary homology between HER3 and HER4. In addition to providing a complete HER3 kinase structure for our molecular simulations, our homology modeling analysis presents a framework for identifying the molecular features that contribute to HER3's unique catalytic conformation, which further validates the treatment of HER3 as a distinctive node in our proposed ErbB pathway model (Fig. 2.2A).

2.3.4 Molecular dynamics simulations of the HER3 kinase

While our homology modeling analysis of the HER3 crystal structure provides insights into specific amino acid motifs and conformational features that distinguish HER3 from its fully-active ErbB family members, in order to identify the most significant molecular interactions that contribute to potential mechanisms of HER3 activity, it is necessary to provide a dynamic picture of the HER3 kinase. To address the dynamic behavior of the HER3 kinase, we performed 10 ns molecular dynamics (MD) simulations of the complete

HER3 crystal structure (PDB code 3LMG) [163] in an explicitly solvated state (see Methods). The HER3 structure was stable for the duration of the simulation, as indicated by the time-course plots of the root mean square deviation (RMSD) of the backbone atoms. We then performed a series of analyses of the MD trajectory to answer the following questions: does HER3 behave similarly to an active kinase (*i.e.*, fully-active EGFR, HER2 and HER4) at the atomic level? Furthermore, what types of molecular interactions enable HER3 to maintain its weak yet robust activity? These questions are important for justifying, at molecular resolution, the representation of HER3 as a uniquely active node in our ErbB pathway model (Fig. 2.2A).

PCA reveals that the dynamic behavior of HER3 at the molecular level diverges from that of an inactive ErbB kinase

Principal component analysis (PCA) was applied to the 10 ns MD trajectory of the HER3 crystal structure to characterize the most significant global motions of the HER3 kinase. The PCA calculation is based on the diagonalization of the variance-covariance matrix of the atomic fluctuations along the MD trajectory to yield the set of eigenvectors (PCs) and associated eigenvalues. The eigenvectors represent the independent modes of atomic motion, and the eigenvalues reflect the contribution of the corresponding eigenvectors to the global fluctuation of the protein.

Motion along the first eigenmode for the complete HER3 kinase was compared to PCA trajectories for the inactive and active conformations of EGFR, HER2 and HER4 that had been previously simulated by us [181]; overall the global motions of the kinases were conserved across the ErbB family members. To determine whether the PCA pattern

was also conserved in HER3's unique catalytic site, we applied PCA to an active site region which included the A-, C-, and N-loops and the αC helix. Figure 2.6 illustrates that motion along the first eigenmode in the inactive EGFR, HER2 and HER4 systems is dominated by A-loop movement, with smaller fluctuations in the other catalytic sub-domains, whereas the active ErbB systems exhibit a uniform level of motion across the sub-domains, with lower-amplitude fluctuations. Despite the crystallization of the HER3 kinase in an inactive-like conformation, the PCA pattern in HER3 appears to diverge

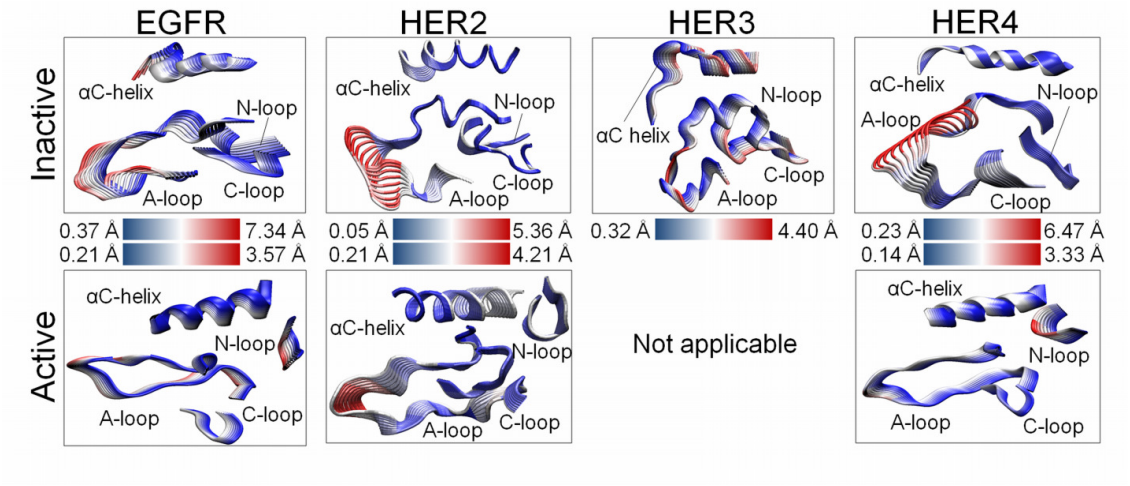

Figure 2.6. Motion along the first PC of the MD trajectory for the ErbB systems.
The structures are color-coded according to the RMSD, where red regions indicate large-amplitude fluctuations and blue regions indicate small-amplitude fluctuations. Despite the crystallization of the HER3 kinase in an inactive-like conformation, the HER3 PCA pattern resembles that of the fully-active ErbB kinases (EGFR, HER2, HER4), in terms of the concerted and low-amplitude fluctuations of its catalytic sub-domains (A-loop, N-loop, C-loop, αC helix).

from that of the other inactive ErbB kinases and instead resembles that of the active ErbB kinases, in terms of its concerted and low-amplitude sub-domain fluctuations. The normalized PCA variance-covariance matrices for vector displacements of atoms further emphasize and quantify the coupling among the catalytic sub-domains in the HER3 system, for example, between the C-loop (residues 814-820 in HER3) and A-loop/3_{10}

helix (residues 828-840). We reason that the interactions among the A-, C-, and N-loops and the αC helix in the active ErbB systems (*i.e.*, active EGFR, HER2 and HER4, and weakly-active HER3) are crucial for alignment of the key sub-domains for catalysis. Consistent with this view, our PCA demonstrates that at the atomic level, the sub-domain motions of the HER3 system deviate from those of an inactive ErbB kinase.

Hydrogen bonding analysis indicates that the molecular mechanism of catalysis in HER3 is distinct from that operating in its fully-active ErbB family members

To identify specific interactions that could be contributing to the distinct pattern of global motion in the HER3 kinase in the inactive-like conformation, individual salt bridges and hydrogen bonds were tabulated for the HER3 system through a hydrogen-bonding analysis of the 10 ns MD trajectory (Table 2.1) and compared to the bonds present in the other ErbB systems [165, 183, 181]. The major difference that we observed between the inactive and active ErbB systems is a significantly greater number of bonds that couple the catalytic sub-domains of the kinases in the active systems compared to the inactive systems. Table 2.1 highlights several of these bonds (EGFR numbering will be used here): E734-K851 is a key salt bridge which couples the A-loop and αC helix and is conserved across the active conformations of EGFR, HER2 and HER4, and the L834-R812 and K836-V810 bonds bridge the A-loop and C-loop to maintain the A-loop in its extended, active conformation. We hypothesize that the tight coupling of the sub-domains in the active ErbB systems may help to correctly position the catalytic residues during assembly of the active site for phosphoryl transfer [181].

EGFR active	HER2 active	HER4 active	EGFR inactive	HER2 inactive	HER4 inactive	HER3 inactive
αC-helix A-loop bonds						
—	—	—	—	—	E739,R841	—
E734,K851	**E766,K883**	**E739,K856**	—	—	— —	—
D737,K836	D769,R868	—	—	—	D742,R841	—
E738,F832	—	E743,F837	—	—	— —	—
—	—	—	E738,K836	—	E743,R841	—
αC-helix C-loop bonds						
—	—	—	—	—	E743,R817	—
αC-helix bonds						
—	A763,S760	—	—	—	—	—
—	E766,R756	—	—	—	—	—
E738,K721	**E770,K753**	**E743,K726**	—	—	—	—
—	—	—	M742,L753	M774,L785	M747,L758	—
A743,L679	—	A748,Q684	—	—	—	—
—	—	—	—	—	A748,R757	—
C-loop C-loop bonds						
—	H843,D845	—	—	—	—	H813,N815
—	—	—	R812,D813	R844,D845	—	—
D813,R817	**D845,R849**	**D818,R822**	—	—	—	—
D813,N818	—	—	—	—	—	N815,N820
A815,N818	A847,N850	A820,N823	A815,N818	A847,N850	—	A817,N820
—	A848,V851	—	—	—	—	—
A-loop C-loop bonds						
—	—	—	—	G865,V842	—	—
—	—	—	—	—	G838,R817	—
L834,R812	**L866,R844**	**L839,R817**	—	—	—	—
—	—	—	L834,D813	—	—	—
K836,V810	**R868,V842**	**R841,V815**	—	—	—	—
—	—	—	—	—	—	D838,R814
E848,R812	—	—	—	—	—	—
—	—	—	K851,R812	—	—	—
A-loop bonds						
—	—	D836,K726	—	D863,K753	D836,K726	D833,K723
—	—	D836,T835	—	—	—	—
L838,R808	**L870,R840**	**L843,R813**	—	—	—	—
—	D871,R840	—	—	—	—	—
A840,G672	—	—	—	—	—	—
—	—	—	—	D873,R897	—	—
—	—	—	—	—	—	D844,K853
—	—	K848,T873	—	—	—	—
K843,D932	—	K848,D937	—	—	—	—
—	E876,R898	—	—	—	—	—
—	—	E849,K871	—	—	—	—
Y845,Y867	—	Y850,F872	—	—	—	—
—	—	—	H846,R865	—	—	—
—	D880,R897	D853,R870	E848,R865	D880,R897	D853,R870	—

Table 2.1. Hydrogen bonding analysis of the HER3 MD trajectory.

By contrast, the bonding pattern in HER3 is distinct from the active and inactive states of its ErbB family members in its scarcity of hydrogen bonds (Table 2.1). HER3 lacks many of the bonds present in the other ErbB kinases due to substitutions in its amino acid sequence: E738 (EGFR numbering), which coordinates the α and β phosphates of ATP by forming a salt bridge with K721 in the αC helix, is mutated to a His in HER3. In addition, E734, which couples the A-loop and αC helix by bonding to K851, is an Ala in HER3. The few bonds that are present in HER3 help to maintain its distinctive active site: the D838-R814 bond, which bridges the 3_{10} helix and the C-loop in HER3, does not appear in the other ErbB kinases. HER3's unique hydrogen bonding network (or lack thereof) suggests that the molecular mechanism of catalysis in HER3 is distinct from that operating in its fully-active ErbB family members (EGFR, HER2, HER4).

Hydrophobic analysis of the ErbB kinases reveals that HER3 critically depends upon a coordinated set of hydrophobic interactions for its weak yet robust activity

In the active conformation, EGFR, HER2 and HER4 rely on an extensive network of hydrophilic interactions to maintain their active sites in a catalytically competent state, whereas the inactive conformations exhibit characteristic hydrophobic interactions, a feature that helps to differentiate the inactive and active states. In ErbB-mediated signaling, the introduction of the ErbB dimer interface alters the hydrophobicity of the region and, in conjunction with allosteric effects, perturbs the conformational sampling space of the protein to induce a shift toward the active conformation [141, 166, 208].

To quantify these hydrophobic interactions, we computed the solvent accessible surface area (SASA) and water density fluctuations (see Methods) for the HER3 MD

trajectory and compared the results to the EGFR, HER2 and HER4 trajectories previously simulated by us [167]. In particular, we focused on several highly conserved hydrophobic regions which are considered to be important for defining the catalytic state of the ErbB kinases[93, 94]: the catalytic spine (C-spine), regulatory spine (R-spine), hydrophobic core, and β3-αC loop. Figure 2.7 displays the correlation between the mean SASA and the normalized water density fluctuations for each hydrophobic region in the inactive and active conformations of each ErbB kinase. The SASA analysis is a reliable measure of

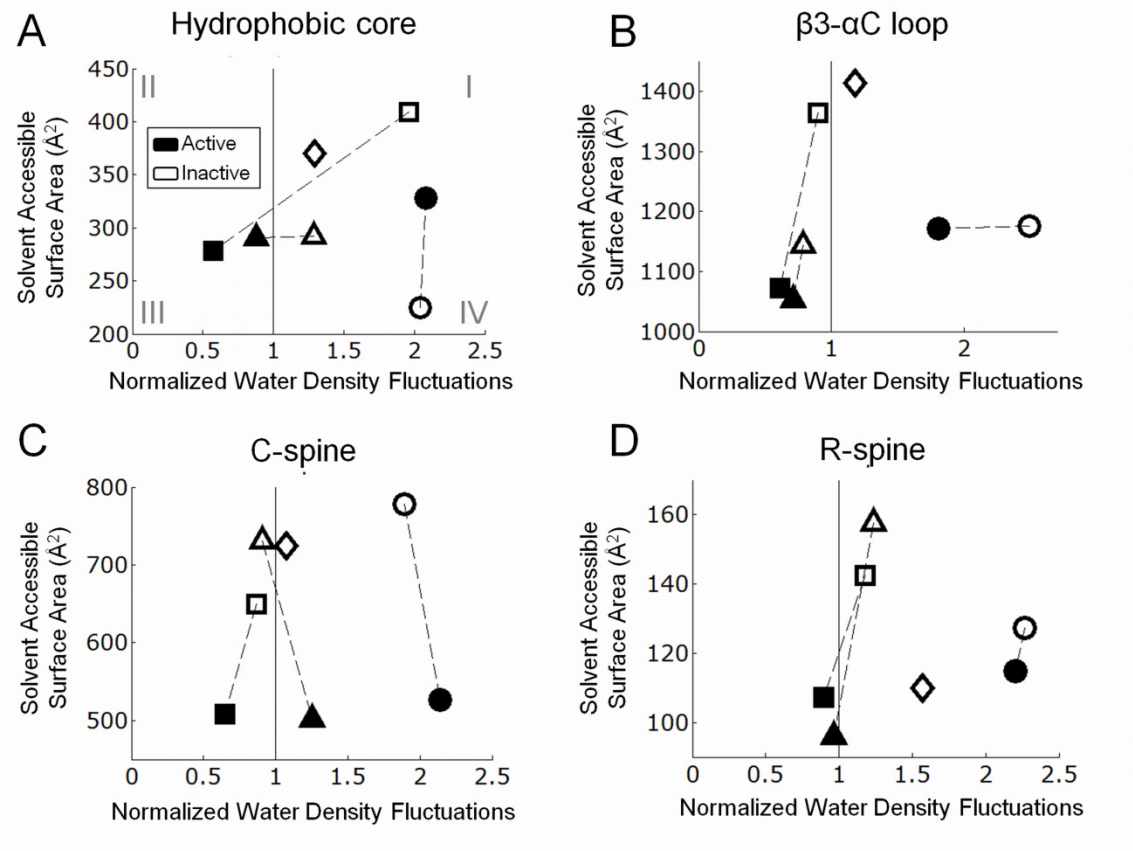

Figure 2.7. Correlation between the SASA and water density fluctuations.
(A) hydrophobic core, (B) β3-αC loop, (C) C-spine and (D) R-spine. Dark and light markers represent the active and inactive ErbB systems, respectively. The ErbB kinase systems are represented as follows: EGFR (square), HER2 (circle), HER3 (diamond), and HER4 (triangle). Quadrant I represents a perturbation-sensitive region, quadrant II defines a hydrophilically favorable region, and quadrant IV represents a hydrophobically favorable region.

hydrophobicity for smooth hydrophobic interfaces but does not always correlate perfectly with free energies of solvation of hydrophobic groups near irregular hydrophobic surfaces, which are often found in proteins. The normalized water density fluctuation calculation (see Methods), which quantifies the hydrophobicity of heterogeneous surfaces, provides an alternative to the SASA analysis and proposes that increased normalized water density fluctuations signify a more hydrophobic region, where a normalized value of 1 indicates a neutral surface [1, 60]. Thus the correlation between the SASA and water density fluctuations produces additional insights into region hydrophobicity that are not revealed by either analysis independently. Specifically, we have designated four 'quadrants' in each correlation subplot (Fig. 2.7): quadrant II represents a hydrophilically favorable region, quadrant IV defines a hydrophobically favorable region, and quadrant I, which is characterized by high SASA (more hydrophilic) and large water density fluctuations, represents a fragile or perturbation-sensitive region. The quadrants facilitate the identification of key patterns that may emerge from a comparison of the hydrophobic interactions in the ErbB systems.

The HER2 system, in general, lies within the hydrophobically favorable region (quadrant IV), which may be rationalized in terms of HER2's unique ability to bind the molecular chaperone Hsp90 through a hydrophobic interface, a point that is elaborated in our previous studies [166, 183, 181]. In the hydrophobic core, which includes segments of the αC helix and the A-loop and helps to maintain the ErbB kinases in the inactive conformation [93, 94], HER3 demonstrates characteristic hydrophobic interactions that are similar to the inactive EGFR and HER4 systems. Fig. 2.7A illustrates that HER3, as well as the inactive EGFR and HER4 systems, lies within the 'perturbation-sensitive'

quadrant, indicating that the hydrophobic core is a potential fragile point of the protein, in which single point mutants would be expected to disrupt the local interactions. Indeed, it has been reported that mutations in the hydrophobic core of EGFR and HER4 result in activation of the kinase, due to destabilization of the inactive state[141, 160, 208]. By contrast, an analogous mutation in HER3 abolishes ATP-binding and phosphorylation activity [163], indicating that hydrophobic interactions in the core promote HER3 activity, rather than maintain an autoinhibited state as they do in EGFR and HER4. Furthermore, the β3-αC loop, which is uniquely extended in HER3 and lies proximal to the hydrophobic core, represents a perturbation-sensitive node in HER3 (Fig. 2.7B). Thus mutations in this region would also be expected to inactivate the HER3 kinase. Our results support the conclusion that HER3 critically depends upon a coordinated set of hydrophobic interactions for its weak yet robust activity from the inactive-like conformation, whereas EGFR and HER4 employ similar interactions to maintain their autoinhibited status from the inactive-like conformation. We propose that these hydrophobic contacts, including the interactions in the β3-αC loop, account for the tightly coordinated sub-domain motions observed in our PCA analysis of the HER3 system, in contrast to the hydrogen bond-mediated mechanism underlying the coordinated sub-domain motion in the active ErbB systems.

The hydrophobicity of the C-spine and R-spine, whose function is to coordinate the motions of the N- and C-lobes of the EGFR and HER4 kinases in the active conformation [93, 94], was also quantified (Fig. 2.7C, D). In the C-spine and R-spine hydrophobicity plots, the inactive and active conformations of EGFR and HER4 are clearly delineated with respect to the SASA, with the active systems exhibiting a

preferential hydrophobic stabilization (low SASA). The SASA of the HER3 C-spine falls within range of the inactive EGFR and HER4 systems, reflecting that, despite its weak activity, there is no corresponding 'fully-active' state for HER3 as for the other ErbB kinases. This inability to 'fully' activate can be attributed to the lack of the crucial hydrogen bonding network identified earlier, which is required to stabilize the active-like kinase conformation. The SASA of the HER3 R-spine deviates from the values for the inactive EGFR and HER4 systems, and instead demonstrates low SASA (high hydrophobicity). This result can be rationalized by the increased hydrophobicity of the R-spine, which includes segments of the truncated αC helix in HER3, hence positioning the HER3 system in the hydrophobically-favorable quadrant of the R-spine hydrophobicity plot (Fig. 2.7D).

The results of our molecular-scale simulations support the characterization of HER3 as a weakly active kinase that, in contrast to its fully-active ErbB family members, depends upon a unique hydrophobic interface to coordinate the alignment of specific catalytic residues required for its activity. Thus the MD simulations substantiate, at a molecular level of resolution, the inclusion of the weakly-active HER3 node in our ErbB pathway model (Fig. 2.2A). Moreover, our results highlight the most significant molecular interactions that contribute to potential mechanisms of HER3 activity, which will be clinically relevant in terms of informing the design of small-molecule inhibitors targeted to the HER3 kinase. Now that we have provided support for the *topology* of our ErbB pathway model (Fig. 2.2A), it is necessary to quantitatively evaluate the model by incorporating appropriate kinetic parameters (Fig. 2.2C). This is an important step in our multiscale modeling scheme, as the ErbB signaling model represents the translation of

our observations of the robust activity of the HER3 kinase into a physiologically relevant context, with the aim of rationalizing the clinical implications of HER3 catalytic activity for elucidating mechanisms of drug resistance in ErbB-driven tumor cells.

2.3.5 Systems model of ErbB signaling defines a mechanism for HER3-mediated TKI resistance

To investigate the implications of HER3 activation for ErbB signaling dynamics in a cellular context, we constructed a systems-level model (Fig. 2.2C) derived from that of Schoeberl et al. [154], and informed our model with the HER3 phosphorylation rate constants calculated from the experiments reported by Shi et al. [163]. EGFR, HER2 and HER3 species were included in the model (HER4 was omitted as in the model by Schoeberl et al., to reflect its undetectable levels in most cancer cell lines) and allowed to form EGFR-HER3, HER2-HER3, and HER3-HER3 dimers upon binding to the ligand NRG-1β (as NRG-1β is a HER3-specific ligand, it is not able to induce EGFR or HER2 homodimers[146]). To represent HER3 catalytic activity, we implemented HER3 homodimerization and *trans*-autophosphorylation in our model, resulting in induction of PI3K-AKT signaling (Figure 2.2C). Based on the previous experimental results indicating a 1000-fold weaker phosphoryl transfer mechanism operating in HER3 [163], we assumed that the rate of HER3 phosphorylation is 1000-fold slower than that of EGFR. The addition of the HER3 homodimer species does not significantly alter the signaling dynamics of the model; levels of phosphorylated HER3 (pHER3), pHER2 and pAKT peak within minutes of NRG-1β stimulation (Figure 2.8), whereas pEGFR exhibits a weak signal due to poor induction of EGFR dimers by NRG-1β. Thus the phosphorylated HER2/3 heterodimer remains a predominant species in the model, although levels of

pHER3 are slightly higher due to activation and *trans*-autophosphorylation of HER3 homodimers.

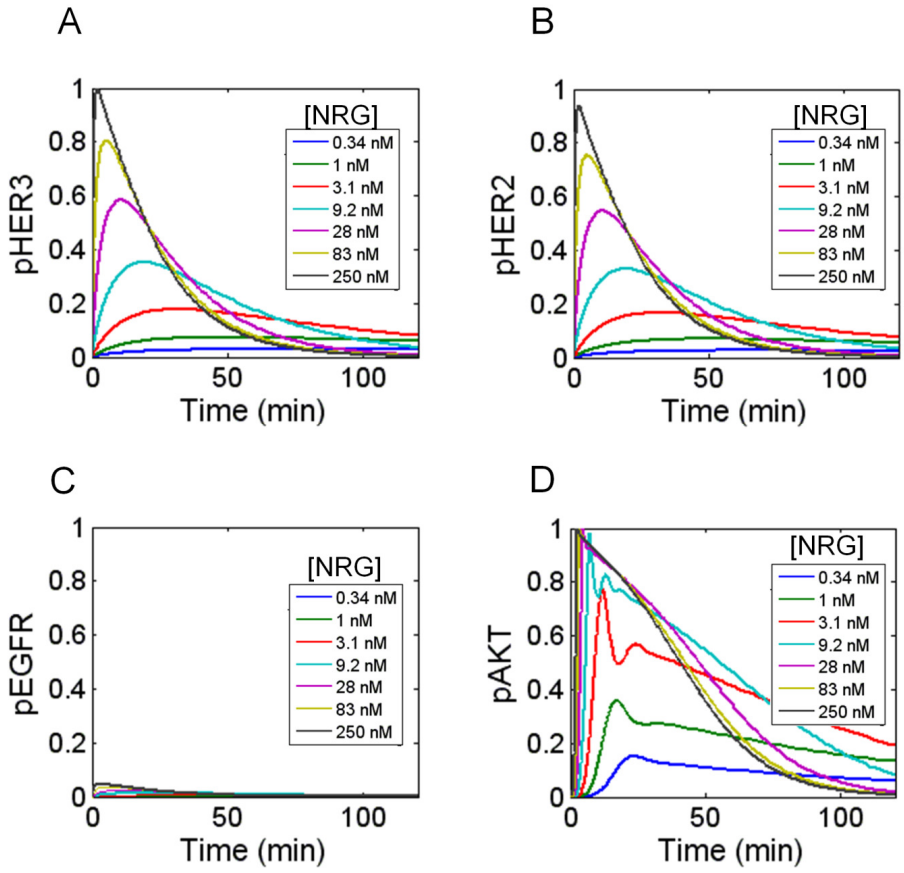

Figure 2.8. Time course plots for (A) pHER3, (B) pHER2, (C) pEGFR and (D) pAKT.
For each phosphorylated ErbB species, data was normalized to the maximum pHER3 signal observed, to facilitate comparison of the RTK activation levels. For pAKT, data was normalized to the maximum pAKT signal observed.

To identify the key proteins that direct signaling in our model of the ErbB network, which includes HER3 catalytic activity, parameter sensitivity analysis was performed with respect to pAKT. Figure 2.9A displays the normalized sensitivity of pAKT to various species in the model: it is clear that HER3 and NRG-1β represent the most sensitive species in the signaling network, followed by HER2 concentration. EGFR is not

a strong determinant of the extent of AKT phosphorylation, as expected from the weak

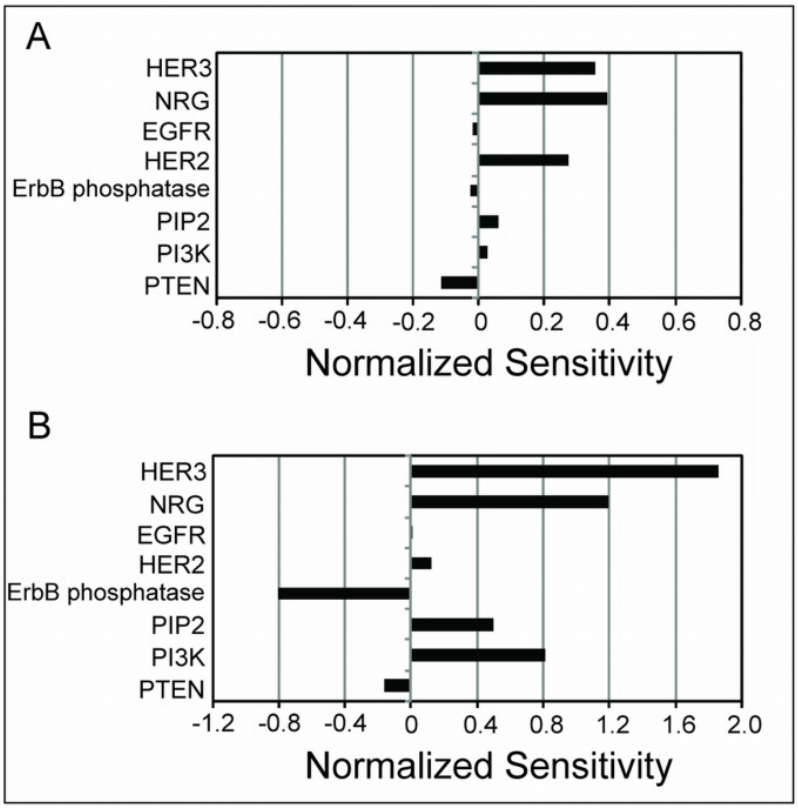

Figure 2.9. Parameter sensitivity analysis of the HER3 signaling model.
The normalized, time-integrated sensitivity of pAKT to key molecular species was computed in response to NRG-1β=25 nM by making a 0.1% change in each species concentration for (A) the model representing weak HER3 activity, and (B) the model representing weak HER3 activity in the presence of the EGFR/HER2 inhibitor lapatinib. Sensitivity to certain species, including HER3, HER2 and the ErbB phosphatase, changed upon addition of lapatinib.

ability of NRG-1β to elicit EGFR dimers. EGFR displayed a negative sensitivity because an increase in EGFR-HER3 dimerization results in fewer HER2-HER3 dimers, which produce the most pAKT. PTEN (the PIP_3 phosphatase) and the ErbB phosphatase (labeled as PTP in Fig. 2.2C) also exhibited a negative sensitivity in the analysis, as these phosphatases negatively regulate the signaling network through dephosphorylation of key molecular species.

As previous experimental studies have demonstrated that HER3 is a key mediator

of resistance to various tyrosine kinase inhibitors (TKIs) currently in use [8, 44, 76, 115, 157], we next tested whether weak HER3 catalytic activity could explain potential resistance mechanisms. Previously postulated mechanisms include leaky HER2-catalyzed phosphorylation of HER3 (*i.e.*, incomplete inhibition of HER2 catalytic activity by the TKI) [69, 76, 157], which remains a plausible hypothesis. However, these models of resistance do not consider the potential for HER3 catalytic activity, due to HER3's formerly assumed inactive pseudokinase status. Thus we incorporated the TKI lapatinib, which inhibits EGFR and HER2 catalytic activity, into our model of HER3 signaling using rate constants obtained from Schoeberl and colleagues [154], and simulated the effects of the TKI on pAKT induction upon NRG-1β stimulation. Figure 2.10, which displays normalized levels of pEGFR, pHER2 and pHER3 for increasing concentrations of lapatinib, illustrates that phosphorylated EGFR and HER2 are effectively inhibited at sufficiently high TKI concentrations, as expected. The pHER3 and pAKT signals are slightly more persistent; at the maximum simulated lapatinib concentration (2 μM) and 25 nM NRG-1β, pHER3 and pAKT levels persist at 3-4% of their respective no-inhibitor control levels (Fig. 2.10C-D). Although Fig. 2.10 demonstrates that the implementation of HER3 catalytic activity allows HER3 to escape TKI inhibition and sustain a weak level of AKT signaling, this weak activity is not expected to be phenotypically observable. Indeed, we computed the IC_{50} for lapatinib/pHER3 to be 94 nM (for 25 nM NRG-1β), which agrees with published *in vitro* and simulated values ranging from 100-152 nM [154].

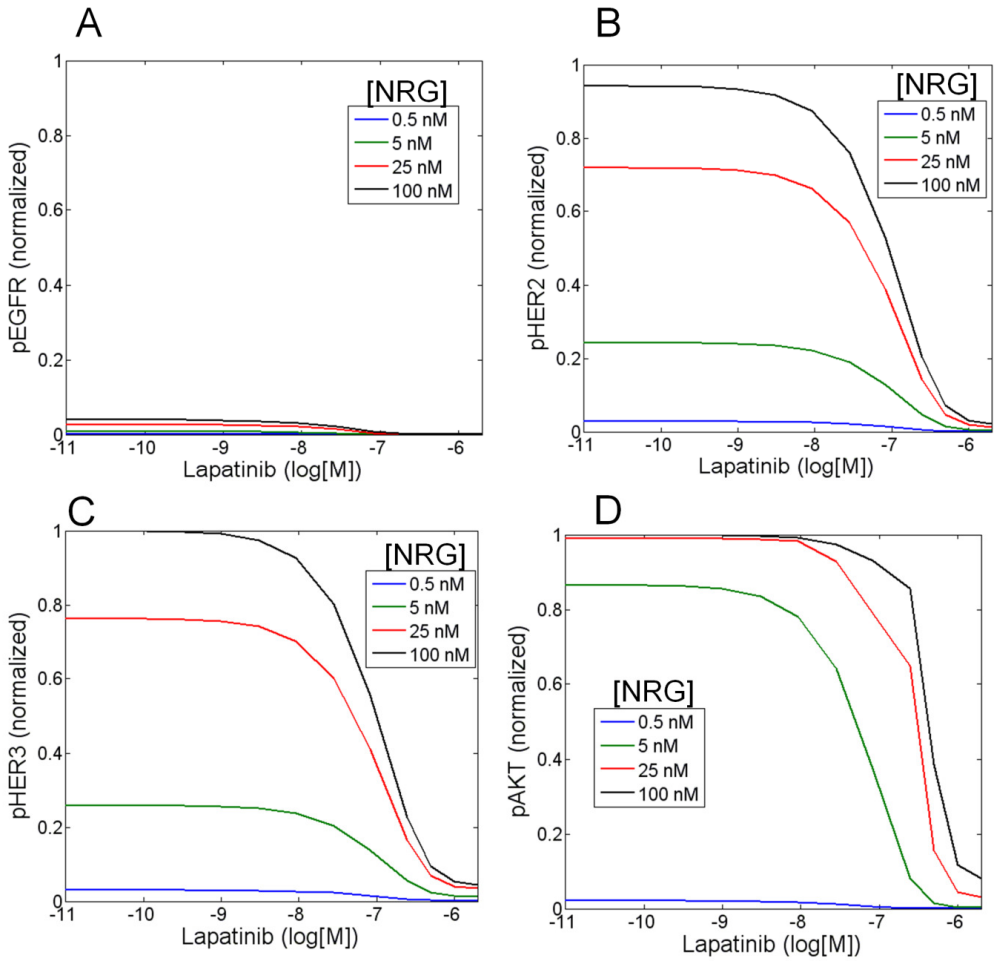

Figure 2.10. Dose-response curves of lapatinib treatment in the HER3 signaling model. The response to the TKI was computed following a 30 minute pre-incubation with lapatinib and 10 min stimulation with increasing concentrations of NRG1-β. Results for pEGFR, pHER2 and pHER3 were normalized to the no-inhibitor control value for 100 nM pHER3 to facilitate comparison of the profiles for the three ErbB kinases. Results for pAKT were normalized to the no-inhibitor control value for 100 nM pAKT.

Figure 2.9B compares the results of pAKT sensitivity analysis of the lapatinib-treated model to those of the inhibitor-free model in Fig. 2.9A. It is apparent that sensitivity to HER3 and NRG-1β increases, whereas sensitivity to EGFR and HER2 decreases, as lapatinib sequesters EGFR and HER2 molecules. The negative normalized sensitivity to the ErbB phosphatase also increases (Fig. 2.9B), as the pool of ErbB dimers has diminished due to sequestration of EGFR and HER2 by lapatinib. Thus a single alteration

to the signaling model (in this case, addition of lapatinib) significantly redefines the most perturbation-sensitive nodes in the network.

Although the pAKT signal induced by HER3 phosphorylation in our *in silico* lapatinib-treated cell is relatively weak, in an actual physiological context, a tumor cell may employ several resistance mechanisms at once [50, 76, 120, 157]. Indeed, previous experimental studies have demonstrated that leaky HER2 phosphorylation of HER3 in TKI-bound HER2/3 heterodimers is amplified by additional resistance mechanisms, such as inhibition of cellular phosphatases by TKI-mediated production of reactive oxygen species (ROS), and increased expression of HER3 at the plasma membrane [66, 127, 157]. To test whether weak HER3 activity during lapatinib treatment could be augmented by other processes, we simulated the lapatinib-treated model for decreased initial concentrations of the ErbB phosphatase, which represents a sensitive node in our pAKT sensitivity analysis (Fig. 2.9B). The pAKT signal was recorded in response to 2 μM lapatinib, as this concentration lies within the range of TKI concentrations shown to induce drug resistance in various tumor cell lines [50, 76, 157]. At the lowest concentration of phosphatase tested (1% of its nominal value represents a limiting phosphatase concentration with respect to HER3), even low levels of pAKT induced by HER3 signaling were amplified for varying concentrations of NRG-1β (Figure 2.11A). As an alternative resistance mechanism, we simulated the lapatinib-treated model for augmented levels of plasma membrane-bound HER3, which represents another highly sensitive node in our sensitivity analysis of the lapatinib-treated model (Fig. 2.9B). It has been demonstrated that in certain cases of TKI resistance, the tumor cell responds to the reduction in pAKT levels by upregulating vesicular transport of HER3 from the cytoplasm to the plasma membrane [127, 157]. Figure

2.11B illustrates that for a 2-fold increase in surface HER3 levels, the pAKT signal is amplified for varying concentrations of NRG-1β, in similarity to our results for decreased phosphatase concentration. The extent of amplification is not as pronounced for 5 nM NRG-1β as for 25-100 nM NRG-1β, as the K_d for HER3/NRG binding is 10 nM.

Figure 2.11C displays the results of decreased phosphatase levels combined with increased HER3 concentration, a phenomenon that, in effect, shifts the HER3 phosphorylation/dephosphorylation equilibrium and has been observed in various tumor cell lines [157]. For 25 nM NRG-1β, the pAKT signal is restored to nearly 60% of its no-inhibitor control level, and pAKT levels are nearly 100% regained for 100 nM NRG-1β, effectively recreating drug resistance *in silico*. Our data parallels the experimental studies performed by Sergina and colleagues [157], which describe HER3-mediated resistance and pAKT signaling in various TKI-treated tumor cell lines as well as *in vivo*. Thus our model demonstrates that even a weak level of HER3 signaling may be physiologically relevant in the context of an ErbB-driven tumor cell, and illustrates several routes through which HER3 signaling may be compounded by other previously postulated resistance mechanisms to generate TKI resistance.

Based on our model results[182], we propose two potential therapeutic strategies for overcoming TKI resistance in certain tumor cell lines: (a) treatment with an anti-HER3 monoclonal antibody (mAb) to prevent HER3 dimerization, or (b) treatment with a TKI specific for HER3 in addition to dampening the amplifying mechanisms using inhibitors of vesicular trafficking (for blocking upregulation of surface HER3) or antioxidants (for inhibition of TKI-produced ROS), which have been shown to overcome TKI resistance *in vitro* [157]. Our data suggest that application of a TKI targeting HER3 would be ineffective

as a standalone strategy and would require supplementation with an additional therapy, as even a weak level of HER3 activity resulting from incomplete TKI inhibition is sufficient to induce drug resistance.

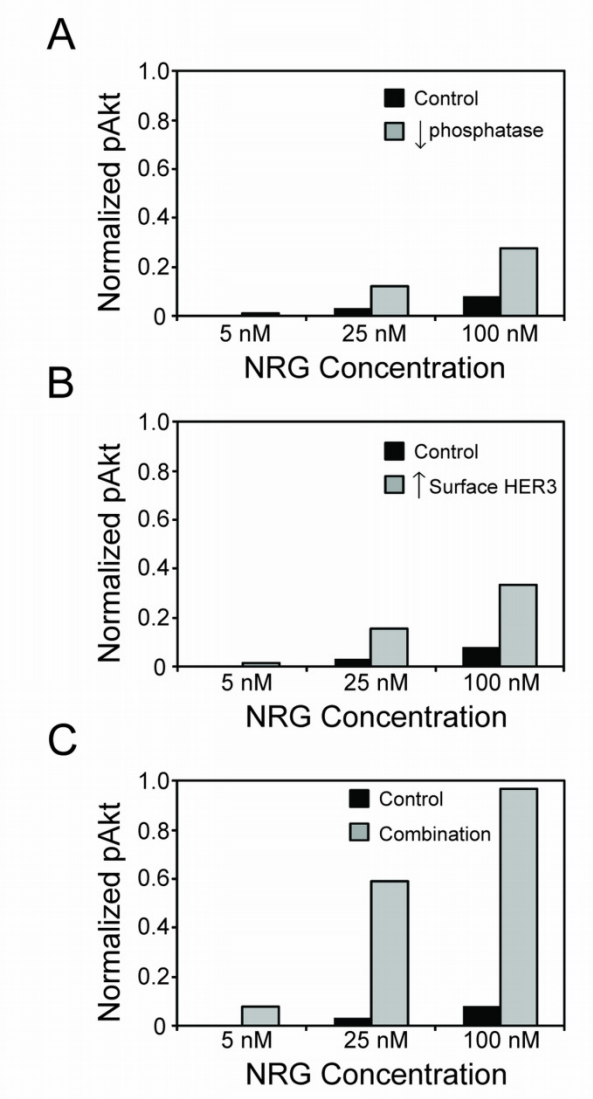

Figure 2.11. Levels of pAKT expressed following a 30 minute pre-incubation with 2 μM lapatinib and 10 min stimulation with increasing concentrations of NRG-1β.
(A) Decreased expression of the ErbB phosphatase (1% of nominal value), (B) increased expression (2-fold) of HER3 at the cell surface, and (C) combination of decreased ErbB phosphatase and increased surface HER3. The control represents the pAKT signal in the presence of inhibitor only (no amplification mechanism).

2.4 Discussion

In this work we applied a multiscale modeling approach to investigate the non-canonical catalytic mechanism employed by the ErbB kinase HER3, a RTK formerly categorized as an inactive pseudokinase, and the physiological relevance of this activity to mechanisms of drug resistance in an ErbB-driven tumor cell *in silico*. The results of our molecular-scale simulations support the characterization of HER3 as a weakly active kinase that, in contrast to its fully-active ErbB family members, depends upon a unique hydrophobic interface to coordinate the alignment of specific catalytic residues required for its activity[182]. Translating our molecular simulation results of the uniquely active behavior of the HER3 kinase into a physiologically relevant environment, our HER3 signaling model demonstrated that even a weak level of HER3 activity may be sufficient to induce AKT signaling and TKI resistance in the context of an ErbB signaling–dependent tumor cell, and therefore therapeutic targeting of HER3 may represent a superior treatment strategy for specific ErbB-driven cancers.

Our homology modeling analysis of the HER3 kinase, using the EGFR and HER4 structures as templates, provided a framework for highlighting the key structural differences between the catalytic site of HER3 and the other ErbB kinases, and for identifying the molecular features that distinguish HER3 from the other nodes (EGFR and HER2) in our multiscale model of HER3 activity. The homology modeling results reflected close evolutionary ties between HER3 and HER4, suggesting that despite employing different catalytic mechanisms, these two kinases may share other molecular regulatory mechanisms, such as autoinhibition through N-lobe dimer formation [85]. We also determined that the application of multiple templates in the initial sequence alignment improved the stereochemical quality of our HER3 homology model. This

methodological result is relevant as our homology modeling platform may be extended to other pseudokinases that have not yet been crystallized. Indeed, several putative pseudokinases share structural features with HER3, among them JAK2, TYK2, guanylate cyclase D (CYGF) and protein serine kinase H (PSKH2), all of which lack the catalytic aspartate in the HRD motif [18]. Application of our modeling and analysis procedure to these other kinases will facilitate identification of residues that may compensate for the lack of key catalytic motifs in the pseudokinase domains and provide insight into how unusual modes of phosphoryl transfer may have evolved in eukaryotic pseudokinases.

At the atomic level, our molecular dynamics simulations and PCA analysis of the HER3 kinase crystal structure revealed that the HER3 system exhibits tightly coordinated fluctuations of its catalytic sub-domains, which may facilitate the alignment of key catalytic residues involved in phosphoryl transfer. Despite sharing a pattern of concerted sub-domain motion with its fully-active ErbB family members, the HER3 system appeared to achieve this state through a unique mechanism, namely, a tightly coordinated set of hydrophobic contacts. By contrast, an extensive network of hydrophilic interactions is employed by EGFR, HER2 and HER4. The hydrogen bonding analysis and hydrophobicity calculations demonstrated that the conserved hydrophilic contacts present in the active sites of EGFR, HER2 and HER4 are replaced in HER3 by a unique hydrophobic interface formed by the extended β3-αC loop, the molten αC helix and the 3_{10} helix in the A-loop [180]. Thus our atomic-level simulations suggest that the ErbB kinases may have evolved two different modes of achieving the shared goal of prearranging the catalytic machinery for phosphoryl transfer. Moreover, the MD results highlight the most significant molecular interactions that contribute to potential

mechanisms of HER3 activity, which will be clinically relevant in terms of informing the design of small-molecule inhibitors targeted to the HER3 kinase.

Employing a systems model, we extended our results to the cellular level to rationalize the clinical implications of HER3 catalytic activity for elucidating mechanisms of resistance to TKIs in ErbB-driven tumor cells. To this end we implemented weak HER3 catalytic activity in a signaling model of the ErbB kinase network and tested the effects of the TKI lapatinib on phosphorylation of AKT in our *in silico* system. The simulation results revealed that, through dimerization and *trans*-autophosphorylation, HER3 is able to escape TKI inhibition and sustain a weak level of AKT signaling, which may be amplified by additional resistance mechanisms, as we illustrated for decreased expression of the ErbB phosphatase and increased expression of HER3 at the plasma membrane, to recreate drug resistance *in silico.* Our results parallel the experimental studies performed by Sergina et al. [157], which demonstrate HER3-mediated resistance and pAKT signaling in various TKI-treated tumor cell lines as well as *in vivo*. Furthermore, our model is sufficiently versatile to rationalize additional experimental results obtained by Sergina and colleagues: in their study it was observed that siRNA-induced knockdown of HER2 abrogates drug resistance, which suggests that HER2 is critically involved in the resistance mechanism, for example, by providing an activation stimulus for HER3 in the context of a HER2/3 dimer. Our model could be altered to reflect this result by implementing HER3 phosphorylation in the context of a HER2/3 heterodimer, rather than a HER3 homodimer. In fact, HER3 catalytic activity may be explicitly represented by several different model topologies, but the implication for the cell phenotype is the same: namely, that a weak level of HER3 activity may

induce drug resistance through amplification by additional resistance mechanisms, and therefore HER3, rather than the routinely-targeted EGFR and HER2 kinases, may represent a more viable candidate for inhibition in the context of an ErbB signaling–dependent tumor.

Indeed, TKI therapy will remain limited in its efficacy until drugs that target the HER3 kinase become routinely available. Recently, Schoeberl and colleagues evaluated the *in vitro* and *in vivo* efficacy of a novel therapeutic anti-HER3 antibody, MM-121, and reported that MM-121 reduced ligand-dependent phosphorylation of HER3 and abrogated resistance to the TKI gefitinib by preventing reactivation of HER3 [155]. Although HER3 activity is weak compared with that of other RTKs, HER3 may remain catalytically active in situations where the activities of its ErbB family members have been attenuated by kinase inhibitors, such as lapatinib or gefitinib, which do not target HER3. Since the catalytic site of HER3 differs structurally from the other ErbB kinases, it may be possible to design TK inhibitors that demonstrate improved selectivity for HER3 as the target kinase. A clinically valuable feature of our multiscale modeling approach is that the flow of information between models occurs in a bidirectional manner, so that it is possible to apply the results of our ErbB signaling model, which indicate that targeting of HER3 may be a superior therapeutic strategy for certain ErbB-driven cancers, to the design of molecular inhibitors of HER3 activity at the atomic scale. In this way, our atomic and subatomic simulations of HER3 activity inform our systems model, and the systems model, in turn, informs the atomic/subatomic-level simulations (*i.e.*, design of molecular inhibitors or mAbs against HER3). Thus multiscale modeling provides a powerful and quantitative platform for investigating the complexity inherent in

intracellular signaling pathways and rationalizing the effects of molecular perturbations on downstream signaling events and ultimately, on the cell phenotype.

Chapter 3

Molecular Simulation of Structure, Dynamics, and Function in the ErbB Receptor Tyrosine Kinases

Adapted from: Telesco SE and Radhakrishnan R. Atomistic Insights into Regulatory Mechanisms of the HER2 Tyrosine Kinase Domain: a Molecular Dynamics Study. (2009) *Biophys J.* 96(6):2321-2334.

Shih A, Telesco SE, Choi S, Lemmon M, Radhakrishnan R. Conserved Hydrophobic and Hydrophilic Bond Interaction Networks in ErbB Family Kinases. (2011) *Biochemical Journal* 436:241-51.

In Chapter 2, we focused on the unique activity of the HER3 tyrosine kinase in order to highlight the implications for elucidating mechanisms of drug resistance in specific human cancers. In Chapter 3, we extend our molecular dynamics study to the remaining ErbB family members (EGFR, HER2, HER4) to compare and contrast mechanisms of regulation and activation in the ErbB RTK family. We bias certain sections of this chapter toward the HER2 kinase, as HER2 exhibits several structural and functional features which are unique among the ErbB RTKs.

3.1 Introduction

Considering the high degree of sequence similarity and structural homology across the ErbB family members, we sought to elucidate the degree to which molecular mechanisms of activation are conserved across the ErbB RTKs, and to identify differences in overall function that arise from variability in primary structure. Recently, we and others have hypothesized the existence of distinct networks of intramolecular non-covalent bonds that characterize the active and inactive conformations of kinases (for Lyn [129, 130], Abl [38],

EGFR [38, 134, 165] and HER2 [181]), with transitions between the activation states necessitating a shift in these bond networks. In this Chapter, we present bioinformatics and fluctuation analyses of molecular dynamics trajectories of the ErbB kinase domains and relate sequence similarities to correspondence of specific bond-interaction networks and collective dynamical modes. We investigate how the various stimuli/perturbations, such as dimerization, phosphorylation, and mutation, impact both the active and inactive conformations of the ErbB kinase domains.

Although the ErbB kinases share certain structural and functional features, one family member, HER2 kinase, stands out in several ways. HER2, the only ErbB family member for which a crystal structure of the kinase domain has not yet been resolved, is unique in that it does not require ligand-binding in order to dimerize with other ErbB RTKs, and hence its extracellular domain is constitutively poised for dimerization. In order to counterbalance this easily perturbed activation state, a number of regulatory mechanisms exist for HER2, such as auto-inhibition of HER2 kinase domain dimerization through binding of chaperone proteins[30, 203, 202]. In this Chapter, we explore several of these regulatory mechanisms, which are crucial for preventing constitutive activation and uncontrolled proliferative signaling, and are disrupted in many HER2-driven cancers.

HER2 is further distinguished from its family members in regard to the role played by several of its key catalytic domains, including 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 [73]. 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 [186]. Contrastingly, the role of A-loop phosphorylation in HER2 is controversial, as several studies have highlighted the importance of Y877 phosphorylation for kinase activity [17, 207]. Xu et al. [203] have reported 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. Therefore, it is possible that phosphorylation of Y877 augments HER2 kinase activity.

In this Chapter, we perform molecular dynamics (MD) simulations of inactive and active EGFR, HER2, and HER4 kinase structures to elucidate details of the mechanism by which the ErbB RTKs are regulated and activated. We investigate the dimer-mediated allosteric activation of the ErbB kinases through dynamics simulations of a HER2/EGFR heterodimeric system as well as EGFR and HER4 homodimers, and delineate the role of phosphorylation of the A-loop tyrosine residue, Y877, through the free energy perturbation (FEP) method.

3.2 Materials and Methods

3.2.1 Homology modeling of the HER2 kinase domain

The HER2 kinase domain was modeled in homology to EGFR, with which it shares 83% sequence identity, using the program MODELLER [51, 150]. The coordinates of the inactive EGFR kinase domain were downloaded from the Protein Data Bank (PDB code 2GS7) [208] and a sequence alignment between EGFR and HER2 was performed in MODELLER. Missing residues were built using the loop modeling algorithm in MODELLER and

hydrogen atoms were added by employing the hbuild routine in CHARMM27 [112]. The structure of the active HER2 kinase domain was based on the model generated by Bagossi et al. [6]. The Bagossi model, which was constructed in homology to active EGFR (PDB code 1M14) [172], was refined by applying the loop-modeling routine to segments of missing residues and by performing additional energy minimization steps. The final models for the inactive and active HER2 kinase domains were assessed for stereochemical quality using PROCHECK [99].

3.2.2 Molecular dynamics simulations

Models for ErbB1 (EGFR) kinase were derived from the 1M14 (active) and 2GS7 (inactive) structures [172, 208]. Models for HER4 were derived from the structures of Qiu et al., PDB ID: 3BCE and 3BBW [141]. Models for kinase dimers were constructed based on the asymmetric dimer interface described in [208]. Each system was simulated as a fully atomistic, explicitly solvated-system in NAMD [135], using the CHARMM 27 forcefield [112]. The MD simulations and analyses were performed as described in Section 2.2.

3.2.3 Construction of the ErbB dimer systems

The ErbB dimer systems were modeled on the structure of the EGFR homodimer published by Zhang et al. [208]. The structures were minimized to remove unfavorable contacts and hydrogen atoms were added using the hbuild routine in CHARMM. The dimers were explicitly solvated using the TIP3P water potential with the buffering distance set to 15 Å for a total system size of approximately 120,000 atoms. Sodium (Na^+) and chloride (Cl^-) ions were added to achieve net electroneutrality of the system and an ionic strength of 75 mM. All Na^+ and Cl^- ions were placed at least 8 Å away from

any protein atoms and from each other. Minimization and MD steps were performed as for the monomeric systems, except that the production simulations were run for 20 ns.

3.2.4 Free energy perturbation (FEP) simulations

The free energy perturbation (FEP) method [212] was employed to compute the Helmholtz free energy difference between the Y877-unphosphorylated and phosphorylated HER2 structures. Calculations were performed using the alchemical free energy perturbation feature in NAMD [11, 136]. The Helmholtz free energy difference between two thermodynamic states connected by M intermediate, nonphysical substrates in the NVT ensemble is expressed as:

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

$$\beta = \frac{1}{k_B T}$$

where k_B is the Boltzmann constant, T is the temperature, and $v(x;\lambda_i)$ is the potential energy function that depends upon the Cartesian coordinates of the system $\{x\}$, and the coupling parameter, λ_i, that connects the initial and the final states of the transformation. The dual-topology paradigm [55] was utilized, in which the initial and the final state are defined in terms of distinct, noninteracting topologies. Four different simulations were performed, including the transformation of Y877 to pY877 for the solvated unphosphorylated HER2 structures (inactive and active), and the transformation of pY877 to Y877 for the solvated phosphorylated HER2 systems. Each perturbation was divided into 48 windows with 20 ps of equilibration and 80 ps of data collection per window, producing a total of 4.8 ns of simulation time for each FEP transformation.

Harmonic positional restraints were imposed on the sodium and chloride ions, and system electroneutrality was maintained by transforming counterions during the course of the alchemical perturbation. Estimation of error was based on two different sets of coordinates, *ie*, performing the alchemical transformations on configurations obtained after 5 and 10 ns of MD trajectory, and errors are reported in terms of the standard deviation from the mean ΔF. Another source of error involves the end points of the transformation, which correspond to the creation or elimination of a group of atoms and are subject to van der Waals clashes that result in end-point catastrophes [10, 137]. To obtain an accurate estimate of the free energy at the diverging end points, we increased the number of windows at the beginning and end of the FEP simulations, collecting data at several points with λ values close to 0 or 1.

3.3 Results

3.3.1 Molecular dynamics simulations of the ErbB kinases

Following molecular dynamics simulation of each active or inactive monomeric kinase system for at least 10 ns, the time evolution of the RMSD was used to monitor equilibration and to track any reorganization of the A-loop and αC-helix conformations; no conformational switching towards active or inactive states was observed (Figure 3.1 illustrates the RMSD plots for the HER2 systems). Whereas the majority of the protein backbone, including the C-loop and the N-loop, aligns closely between the inactive and active states, the A-loop and αC-helix conformations differ considerably. In transitioning from the inactive to the active conformation, the αC-helix rotates toward the C-lobe, with the rotating end shifting by ~9 Å toward the base of the cleft between the N- and C-lobes.

The helix is also extended by two turns in the active conformation compared with the inactive conformation. In the inactive kinase, the A-loop maintains a 'closed' conformation (mainly through inter-region hydrogen bonds) and partially blocks the catalytic site. By contrast, the A-loop appears 'unfurled' in the active kinase, and lies against the C-lobe.

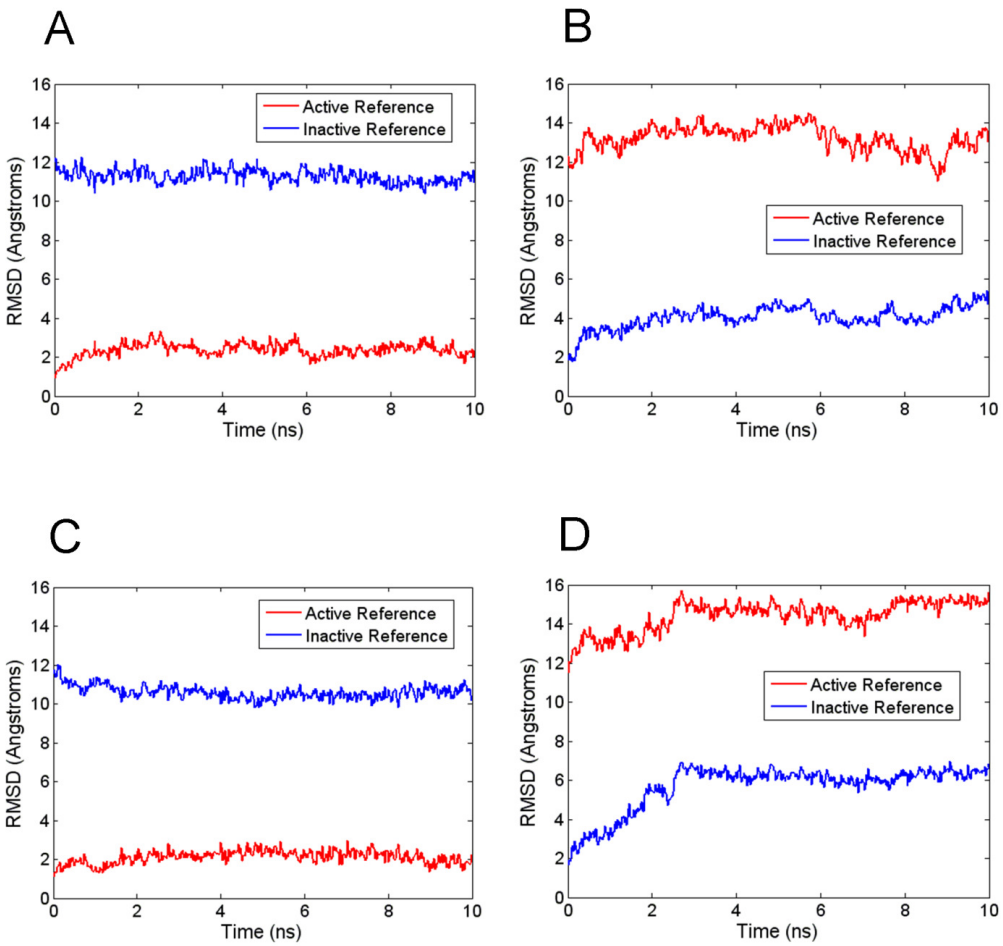

Figure 3.1. RMSD for all backbone atoms in the A-loop of the following HER2 systems. (*A*) the Y877-unphosphorylated active and (*B*) the Y877-unphosphorylated inactive trajectories, and (*C*) the Y877-phosphorylated active and (*D*) the Y877-phosphorylated inactive trajectories. The RMSD is plotted in reference to the initial active structure (*red*) and in reference to the initial inactive structure (*blue*).

3.3.2 PCA reveals a tightly coordinated motion in all active ErbB members

A principal component analysis (PCA) was applied to each MD trajectory in order to characterize the global motions of the ErbB kinases. An objective of the PCA is to delineate the differences in atomistic fluctuations among the inactive and active structures, as conformational rearrangement of the kinase domain is expected to correlate with dramatic changes in the dynamical behavior of the protein. PCA of the ErbB structures was performed on an active site region which comprises all domains critical for catalysis, including the A-, C- and N-loops and the αC helix. Specifically, the C_α atoms of the C-loop/A-loop region (residues 843-888 in HER2) and the αC helix/N-loop region (residues 725-778 in HER2) were chosen as the active site, and the analysis was applied to the NVT-equilibrated trajectory for each monomeric structure (Fig. 3.2). The PCA revealed that motions occurring within the active sites of the inactive and active ErbB monomers differ significantly, particularly in the A-loop. The inactive EGFR monomer exhibits large-amplitude motion in both the αC-helix and the A-loop (4.95 and 7.34 Å, respectively), with smaller fluctuations in the N-loop and C-loop (4.17 and 3.65 Å, respectively). By contrast, the active EGFR monomer demonstrates a uniform level of motion across all four subdomains of the active site with low-amplitude fluctuations (2-3Å), and shows no significant local deformations. The HER2 and HER4 kinase monomers demonstrate similar motions, as shown in Figure 3.2. The inactive HER4 kinase exhibits a dominant motion in the A-loop (6.46 Å), while the N-loop, C-loop, and the αC-helix undergo smaller lateral motions (4.45 Å, 2.64 Å, and 2.30 Å, respectively). In the active conformation, the HER4 kinase presents a fluctuation profile similar to EGFR: the subdomains all have similar small-amplitude motions (2-3 Å) with no large local deformations. Thus, for the three homologous members of the ErbB family, which

have a high degree of sequence similarity, not only are the principal motions conserved across the systems, but the characteristic differences between the inactive and active kinase conformations are maintained in character. We reason that the interactions among the A-, C-, and N-loops and the αC helix in the active ErbB systems are crucial for alignment of the key domains for catalysis. The correlation between the A-loop and C-loop is especially pronounced, ensuring appropriate positioning of the catalytic aspartate, D845, and the coordinating aspartate, D863, for the phosphoryl transfer reaction. This finding suggests large (and possibly similar) differences in the internal network of bonds between the two activity states of each kinase, which will be discussed in the following section.

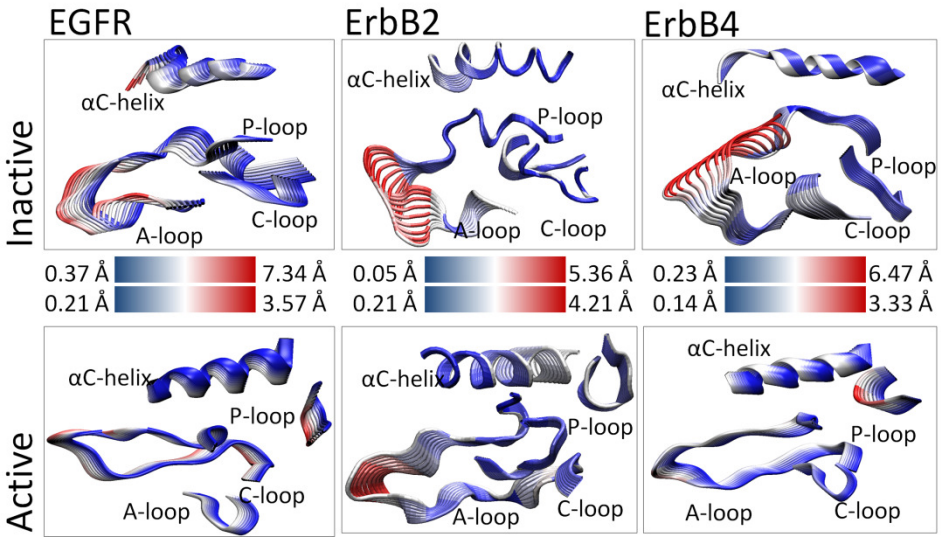

Figure 3.2. PCA of the key subdomains in the inactive and active ErbB kinases.
The motions are overlaid sequentially where the large-amplitude motion in each frame is highlighted in red and the low-amplitude motion is highlighted in blue. The inactive conformations exhibit large, localized motion while the active conformations demonstrate smaller, coordinated motions. Note that Fig. 2.6 also compares the PCA results for HER3.

3.3.3 Hydrogen bonding analysis reveals a conserved network of interactions in the active ErbB kinases

In order to identify specific interactions which could be contributing to the differences in global motions of the inactive and active systems and to investigate the mechanistic basis for the coupling of the A- and C-loops, individual salt bridges and hydrogen bonds were tabulated for each ErbB system through a hydrogen bonding analysis of the 10 ns trajectories. The hydrogen bonds for HER2 are illustrated in Figure 3.3, and are labeled according to the numbered residue pairs listed in Table 3.1. Our analysis reveals that six hydrogen bonds occur in the C-loop of the active HER2 structure, the majority of which link the C-loop and the A-loop (Table 3.1). Specifically, R844-L866, N850-T862, V842-R868, and L852-K860 connect the C-loop to the N-terminal end of the A-loop, while the L846-W888 bond couples the C-loop and the C-terminal end of the A-loop. Contrastingly, only two hydrogen bonds bridge the C-loop and the A-loop in the inactive HER2 system (Table 3.1). The extensive hydrogen bonding that preferentially links the A- and C-loops in the active conformations of the kinase provides a rationale for the cooperative fluctuations between these regions as revealed by the results of the PCA.

The EGFR and HER4 systems revealed a similar bonding pattern in the inactive and active states (see Reference [166] for a complete list of bonds present in the EGFR and HER4 kinases). Hence the similarity in the pattern of specific interactions that preferentially bridge the A- and C-loops in the active conformations of the EGFR, HER2, and HER4 kinases suggests that the tight coupling of the A- and C-loops may be a general feature of the architecture of ErbB receptor kinases which likely aids in the assembly of a catalytically competent active site.

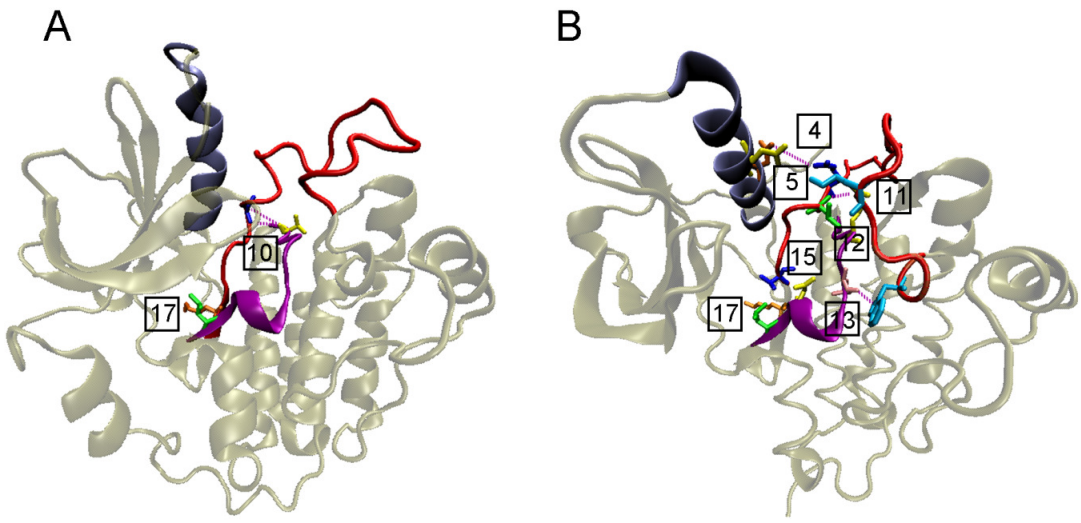

Figure 3.3. Hydrogen bonding analysis for HER2 (A) inactive and (B) active.
Labels correspond to the numbered residue pairs listed in Table 3.1. Only those bonds which couple the key subdomains of the kinase (A-loop, C-loop, and αC helix) are shown to highlight the preferential bridging of the subdomains in the active conformation. The A-loop is shown in *red*, C-loop in *purple*, and αC helix in *gray*.

We extended the evaluation of hydrogen bonds to the A-loop and αC helix, as these regions undergo the most pronounced structural shifts upon transition from the inactive to the active state of the ErbB kinases. In similarity to the hydrogen bonding pattern in the C-loop, we observe a considerable difference in the number of interactions in the A-loop and αC helix of the active and inactive HER2 systems (Table 3.1 and Fig. 3.3). Specifically, there are four hydrogen bonds and two salt bridges in the A-loop of the inactive HER2 structure compared to nine hydrogen bonds and two salt bridges in the active A-loop. The αC helix reflects a more marked distinction between the two systems, as two salt bridges occur in the active structure while salt bridges are entirely lacking in the inactive system. The majority of bonds in the active system connect the key domains of the kinase; for instance, the E766-K883 and D769-R868 bonds link the A-loop and the αC helix (Fig. 3.3B). Furthermore, the R844-L866, N850-T862, V842-R868, L852-K860

and L846-W888 bonds couple the A-loop and the C-loop. The EGFR and HER4 kinases reveal a similar bonding pattern[166]. Hence, the network of specific hydrogen bonds that bridge the key subregions of the kinase (A-loop, C-loop, and αC helix) provides a rationale for the concerted motion of these subdomains observed in the PCA.

Label*	Subregion	Inactive HER2	Active HER2
1	αC helix	–	A763 HN, S760 OG
2	αC helix	N764 HN, S760 O	N764 HN, S760 O
3	αC helix	–	E766 OE1/2, R756 HE
4	αC helix/A-loop	–	**E766, K883**†
5	αC helix/A-loop	–	D769 OD1/2, R868 HH12
6	αC helix	–	**E770, K753**
7	αC helix	Y772 O, G776 HN	–
8	αC helix	–	V773 O, V777 HN
9	αC helix	M774 O, L785 HN	–
10	C-loop/A-loop	V842 O, G865 HN	–
11	C-loop/A-loop	–	V842 O/HN, R868 HN/O
12	C-loop/A-loop	–	R844 HE, L866 O
13	C-loop/A-loop	–	L846 O, W888 HE1
14	C-loop	A848 O, L807 HN	–
15	C-loop/A-loop	–	N850 O, T862 HN
16	C-loop	V851 O, L806 HN	V851 O, L806 HN
17	C-loop/A-loop	L852 HN, K860 O	L852 HN, K860 O
18	A-loop	**D863, K753**	–
19	A-loop	–	L870 HN, R840 O
20	A-loop	–	D871 O, R840 HE/HH12
21	A-loop	D873 OD1/2, R897 HE	–
22	A-loop	–	E876 OE1/2, R898 HE
23	A-loop	**D880, R897**	**D880, R897**
24	A-loop	K883 HZ1/2, E757 OE1	**–**

* Labels correspond to the numbered hydrogen bonds illustrated in Figure 3.3.
† Salt bridges are highlighted in bold.

Table 3.1. Hydrogen bonding analysis for the inactive and active HER2 systems.

3.3.4 Variation among HER2, EGFR, and HER4 in the αC-β4 loop region of the kinase

A recent study by Fan et al. [45] reported that HER2 is strongly autoinhibited relative to EGFR and HER4, 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. Specifically, the polar residues in the αC-β4 loop of EGFR are replaced by nonpolar residues in HER2, which form a hydrophobic patch that contacts another segment of hydrophobic residues located in the A-loop. Residues comprising the hydrophobic patch in HER2 include V773, M774, G776, V777, G778 and V782 in the αC-β4 loop, and I861, T862, F864, L866 and L869 in the A-loop. This hydrophobic motif has been investigated by several groups in relation to its association with various molecules, such as the molecular chaperone Hsp90 [30, 202]. Fan et al. postulate 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 HER4.

In light of these experimental findings, we extended our hydrogen bonding analysis to the αC-β4 loop for HER2, EGFR, and HER4. In tabulating the bonds for each system, we observe that both the inactive and active HER2 structures contain only two hydrogen bonds in the αC-β4 region (Table 3.2). One of the bonds, S783-I861, is shared by both systems and couples the αC-β4 loop to the A-loop. Contrastingly, EGFR and HER4 contain a significantly greater number of hydrogen bonds in the αC-β4 loop. The active EGFR and HER4 structures form ten and eight hydrogen bonds, respectively (Table 3.2). Several bonds in active EGFR, including H749-I829, C751-I829, and H749-

V827, link residues in the αC-β4 loop and the A-loop, while other bonds, including S744-Y740, V745-V741, and L753-M742, connect the αC-β4 loop and the αC helix. The active HER2 system lacks many of these hydrogen bonds because hydrophobic interactions, rather than hydrophilic contacts, predominate in the αC-β4 region. In particular, the H749-I829 and H749-V827 bonds are absent in HER2 because the kinase contains a relatively nonpolar tyrosine residue in the position analogous to the positively-charged H749 in EGFR. Figure 3.4 contrasts the network of hydrogen bonds (in *stick* representation) in the αC-β4 loop of the active EGFR system with the hydrophobic residues (in *van der Waals* representation) in HER2. The individual bonds are labeled according to the numbered residue pairs listed in Table 3.2.

Label*	HER2 Active	EGFR Active	HER4 Active	HER2 Inactive	EGFR Inactive	HER4 Inactive
1	–	A743, L679	A748, Q684	–	–	–
2	–	S744, N676	–	–	–	–
3	–	S744, Y740	S749, L745	G776, Y772	S744, Y740	S749, L745
4	–	V745, V741	M750, I746	–	–	–
5	–	–	–	–	–	M750, M747
6	G778, Y835	D746, Y803	–	–	–	–
7	–	–	H752, Y808	–	–	H752, Y808
8	–	H749, V827	–	–	–	–
9	–	H749, I829	H754, I834	–	H749, I829	H754, I834
10	S783, I861	C751, I829	V756, I834	S783, I861	C751, I829	V756, I834
11	–	R752, Q767	R757, Q772	–	–	–
12	–	L753, M742	L758, M747	–	L753, M742	L758, M747

* Labels correspond to the numbered hydrogen bonds illustrated in Figure 3.4B for EGFR kinase in the active conformation.

Table 3.2. Comparison of the Hydrogen Bonding Network in the αC-β4 loop.

The bonding analysis of the inactive ErbB systems reveals a trend similar to that in the active systems. The inactive EGFR and HER4 conformations contain four and six hydrogen bonds, respectively, in contrast to two bonds in the αC-β4 loop of inactive HER2 (Table 3.2). The lack of hydrogen bonds in the αC-β4 loop of inactive HER2 is consistent with the prominence of hydrophobic interactions in this region as proposed by Fan et al. Despite sharing similar bonding patterns with EGFR in the A-loop and αC helix, HER2 differs markedly from EGFR in the αC-β4 region due to the presence of the hydrophobic patch of residues. The dearth of a hydrogen bonding network and the dominance of hydrophobic interactions surrounding the αC-β4 loop in HER2 is thought to contribute to its interaction with the molecular chaperone Hsp90 [202]. As a mature protein, only HER2 among the members of the ErbB family associates with Hsp90, as EGFR and HER4 lack the segment of hydrophobic residues in the αC-β4 loop. It has been proposed that binding of Hsp90 to the αC-β4 region in HER2 provides an inhibitory mechanism for regulation of HER2 activity by preventing dimerization and subsequent activation of the HER2 kinase [30].

The hydrophobic association between HER2 and Hsp90 is relevant to the effects of clinical mutations in the HER2 kinase domain. HER2 gene mutations have been identified in a cohort of non-small cell lung cancers (NSCLCs) which involve in-frame duplications/insertions within exon 20, a region that corresponds to the αC-β4 loop in the kinase domain. The most frequently occurring abnormality is the in-frame YVMA insertion at residue G776 ($G776^{YVMA}$), which has been shown to undergo markedly higher tyrosine phosphorylation than wild-type HER2, resulting in increased

tumorigenicity [195]. Another prevalent activating mutation is G776S, which has been found

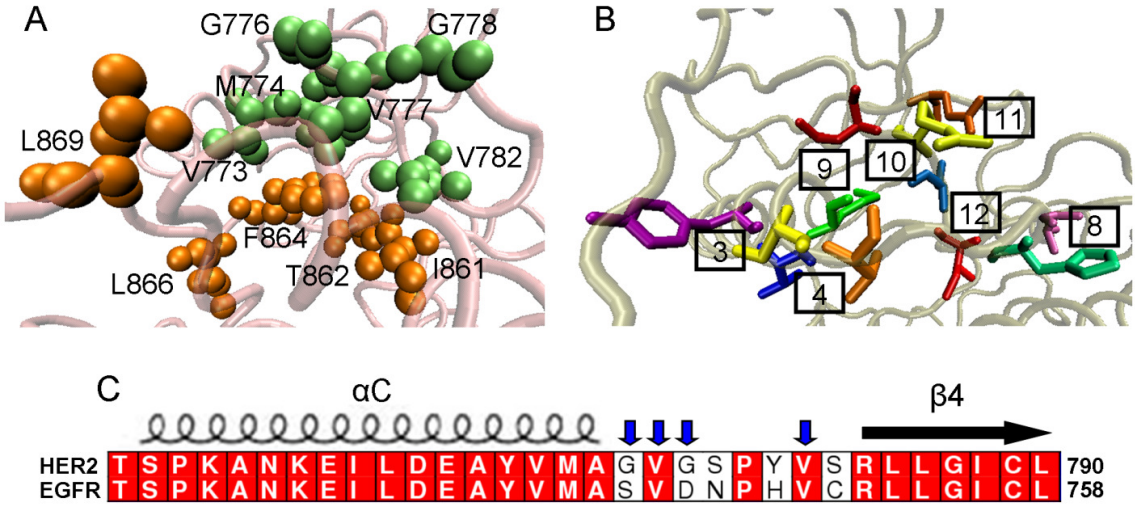

Figure 3.4. Hydrophobic interactions in HER2 in the αC-β4 loop.
(A) The patch of hydrophobic residues in the αC-β4 loop of the active HER2 structure. Residues in the αC-β4 loop are colored *green* and residues in the A-loop are colored *orange*. (B) The hydrogen bonding network in the αC-β4 loop of the active EGFR structure. The bonds are labeled according to the numbering scheme in Table 3.2. (C) Sequence alignment between HER2 and EGFR in the αC-β4 region of the kinase. *Blue arrows* indicate residues contributing to the hydrophobic patch in HER2.

in gastric tumors [175]. Such mutations weaken the hydrophobic interactions surrounding the αC-β4 loop and likely promote a hydrogen bonding network similar to those we have identified in EGFR and HER4, which can disrupt the inhibitory stimulus provided by HER2-Hsp90 association.

3.3.5 Activation in the ErbB dimer systems occurs through disruption of the inactivating bonding network

As postulated by Zhang et al. [208], the ErbB kinases undergo activation via an asymmetric dimer mechanism. In order to investigate potential mechanisms of dimer-induced activation, we constructed the following ErbB dimer systems: HER2-EGFR heterodimer, EGFR homodimer, and HER4 homodimer. The following results will be described for the HER2-EGFR heterodimer, but the results for the ErbB homodimers displayed similar trends[166]. In the context of a HER2-EGFR heterodimer, the residues comprising the dimeric interface for HER2 (kinase undergoing activation) include P707, Q711, M712, I714, L768, L790, and V794, and for EGFR (activating kinase) include I917, Y920, M921, V924, M928, I929, and V956 (Fig. 3.5). A second dimer was constructed in which the Y877-phosphorylated inactive HER2 monomer from the 10 ns equilibrated system was input as the activated kinase. Both heterodimeric systems were solvated and subjected to 20 ns of molecular dynamics simulation (see Methods). The objective was to determine whether dimerization promotes the active state and if so, whether phosphorylation of Y877 facilitates the activation mechanism.

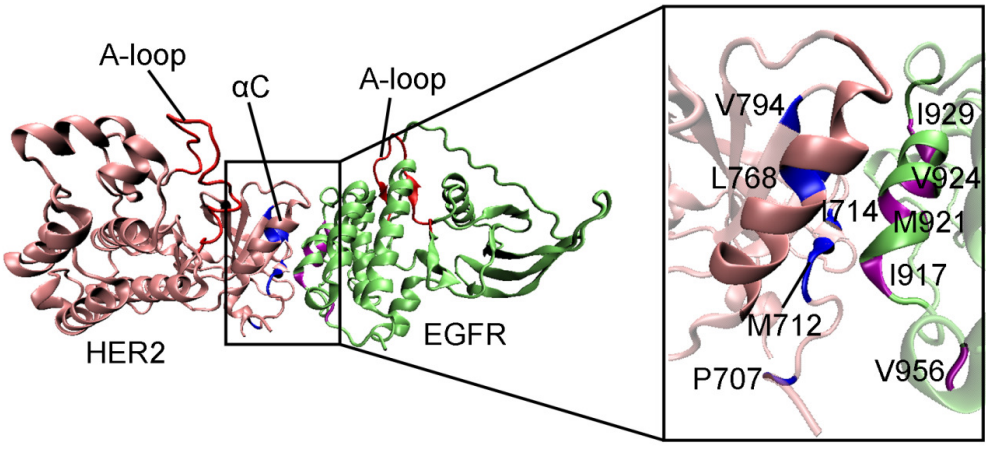

Figure 3.5. Snapshot of the modeled HER2-EGFR heterodimer.

Shown are the residues comprising the dimeric interface for the kinase undergoing activation (HER2, *pink*) and for the activator kinase (EGFR, *green*). The residues that constitute the interface for HER2 are P707, Q711, M712, I714, L768, L790, and V794 (highlighted in *blue* in the inset). For EGFR, the interface residues are I917, M921, V924, I929, and V956 (highlighted in *purple* in the inset).

PCA was performed for the 20 ns MD trajectories in order to characterize the extent of correlation of atomistic fluctuations. As for the monomeric systems, the active site region was chosen to include the C_α atoms of the A-, C-, and N-loops and the αC helix of HER2, the kinase that is undergoing activation in the context of the heterodimer. In similarity to the inactive monomeric simulations, the principal eigenmodes are dominated by A-loop movement (Fig. 3.6). However, the phosphorylated heterodimer exhibits notable fluctuations in the αC helix of HER2 (Fig. 3.6B), in similarity to the active monomeric structures. We also observe a repositioning of the αC helix toward the active conformation during the 20 ns MD simulation, shifting from an RMSD of 2 to 6 Angstroms relative to the inactive state. The HER2 structure in the unphosphorylated dimer also exhibits a slight conformational rearrangement of its αC helix toward the active state in response to the dimerization interface (Fig. 3.6A).

As in our analysis of the monomeric systems, we aimed to rationalize the pattern of global motions in the dimers in terms of specific interactions. In order to identify stabilizing bonds that are perturbed upon dimerization, we performed a hydrogen bonding analysis for the 20 ns MD trajectories. An examination of the bonding network in the heterodimeric systems reveals destabilization and severance of several interactions present in the monomeric systems (Table 3.3). In the dimer featuring Y877-unphosphorylated HER2 as the activated kinase, several bonds responsible for maintaining the inactive state in the monomer are broken, including N764-S760 and Y772-G776 in the αC helix. Indeed, disruption of these bonds is expected as a result of

their position within the dimeric interface. Additional bonds that are disturbed include G865-V842 and D873-R897 in the A-loop. In the dimer involving Y877-phosphorylated

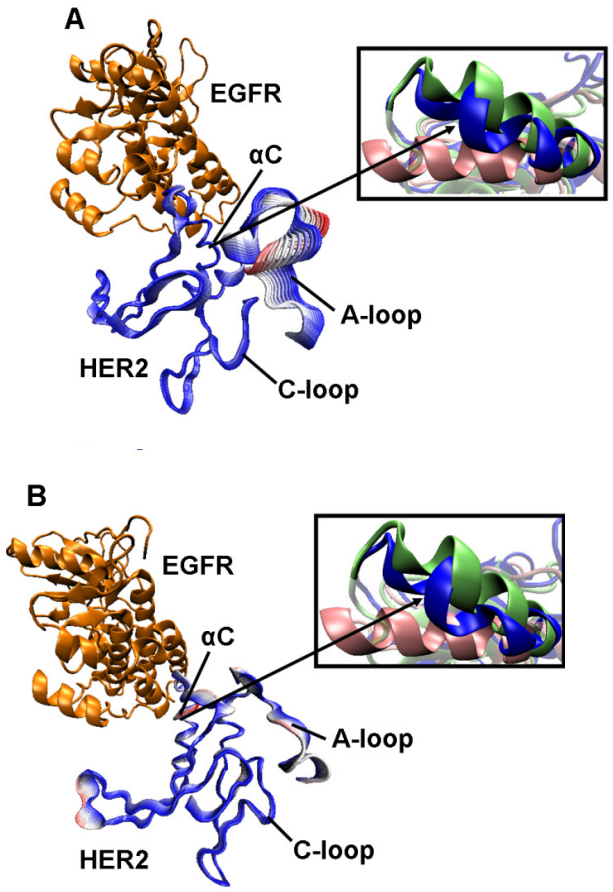

Figure 3.6. Global motions of the HER2-EGFR dimers as determined by PCA.
The structures represent displacements along the first eigenmode for (A) the Y877-unphosphorylated system and (B) the Y877-phosphorylated system. The structures are color-coded according to the RMSD (*red* and *white* indicate large fluctuations, *blue* indicates smaller fluctuations) to illustrate the motions in individual residues. Insets reveal shifting of the αC helix in the dimer (*blue*) away from the inactive conformation (*green*) and toward the active conformation (*pink*).

HER2 as the activated kinase, several bonds expected to break are indeed severed, including T759-E874, L785-M774 and Y772-G776 in the αC helix. Key interactions in the A-loop, such as G865-H843, R868-R840, and V884-K887, have also been perturbed. Our findings indicate that the dimerization interface directly alters the pattern of

stabilizing hydrogen bonds in the inactive system, in agreement with the allosteric activation mechanism proposed by Zhang et al. [208]. Disruption of specific interactions in the αC helix suggests that the pattern of hydrogen bonds in the dimer is shifting away from the inactive state and that the kinase domain is more susceptible to perturbations that would enable a conformational rearrangement toward the active state. Moreover, the dimeric interface directly induces the repositioning of the αC helix toward the active conformation even in the relatively short timescale of 20 ns.

	Y877-Unphosphorylated Dimer	Y877-Phosphorylated Dimer
Proximal, Broken*	–	T759-E874
	N764-S760	–
	Y772-G776	Y772-G776
	–	L785-M774
Proximal, Unbroken†	K753-D863	K753-D863
	E757-K883	E757-K883
	L785-M774	–
Not Proximal, Broken‡	G865-V842	–
	–	G865-H843
	–	R868-R840
	D873-R897	–
	–	V884-K887

* Refers to bonds containing residues that are proximal to the dimer interface (within 3 Å of at least one of the dimeric interface residues for 75% of the production trajectory) and have broken in the dimer trajectory.
† Refers to bonds containing residues that are proximal to the dimer interface, but have not broken in the dimer trajectory.
‡ Refers to bonds that are not proximal to the dimer interface, but have broken in the dimer trajectory.

Table 3.3. Summary of broken and unbroken bonds for the HER2-EGFR dimer.

Several bonds 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 is one of the salt bridges involved in the dual autoinhibitory mechanism in the monomeric simulations, persists throughout the dimeric trajectories. We also note the importance of the K883-E757 bond, which is conserved among all three inactive ErbB kinases, in governing stabilization of the inactive structure, as it must break in order for the K883-E766 salt bridge to assemble in the active state.

3.3.6 Free energy perturbation analysis of the role of Y877 phosphorylation in HER2

Several experimental studies have highlighted the importance of A-loop phosphorylation in HER2 activation. In the Y877-phosphorylated active system, we identify a network of hydrogen bonds that maintain the A-loop in the open conformation; as the A-loop extends, it uncovers the catalytic loop and promotes access of peptides to the active site. The hydrogen bonds fasten the A-loop to a segment of the αF helix (residues 896 to 901) and to the region between the αE helix and the C-loop (residues 840 to 844), ensuring that the A-loop remains in the active state (Fig. 3.7). Three hydrogen bonds, L866-R844, V842-R868, and R840-L870, secure the A-loop at its N-terminal end (Fig. 3.7B). Likewise, three bonds, Y877-F899, A879-R897, and E876-R898, fasten the A-loop at its C-terminal end (Fig. 3.7C). It is noteworthy that several of the bonds, such as L866-R844 and Y877-R844, link residues in the A-loop and C-loop, underscoring the tight coupling of these regions in the active state. However, the key residues required for kinase activity, D863 (the coordinating aspartate) and D845 (the catalytic aspartate), do not participate in

the hydrogen bonding network and hence remain poised for catalysis. Based on these results, we propose an activation model in which residues neighboring D863 and D845 compose hydrogen bonds that stabilize the A-loop in the active state, while ensuring availability of the catalytic aspartate residues for kinase activity.

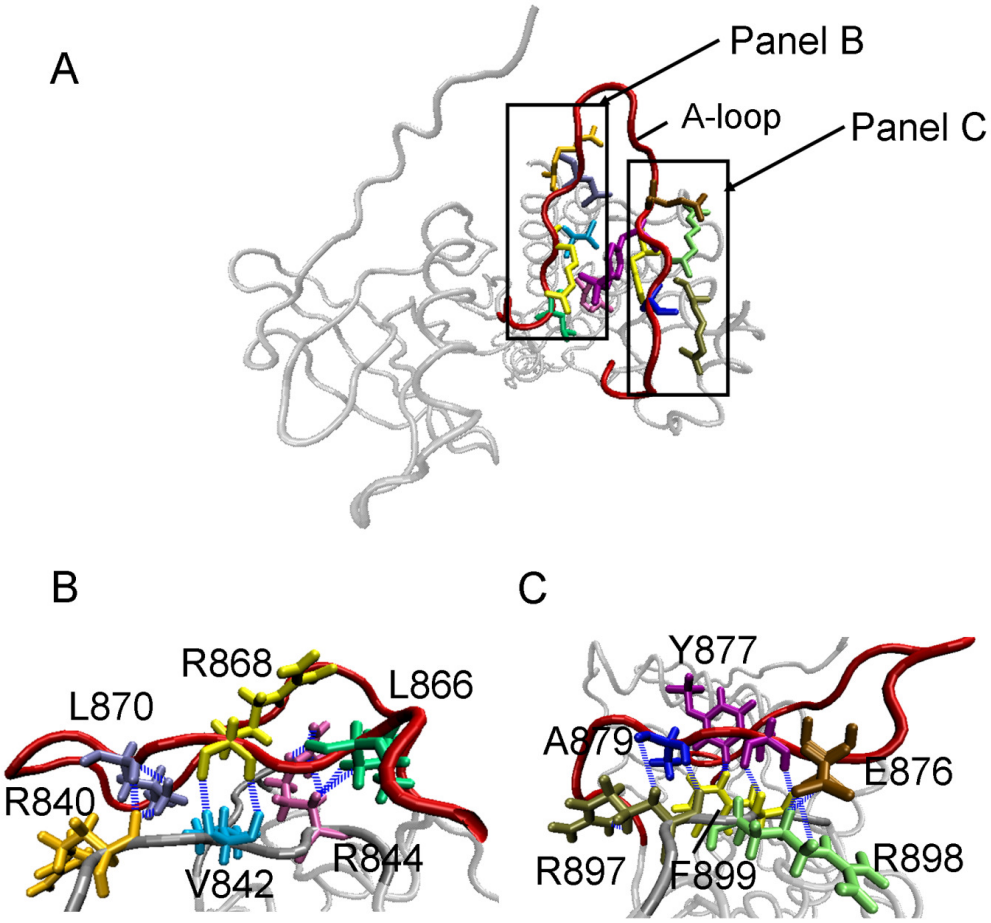

Figure 3.7. The stabilizing hydrogen bonds in the A-loop of the pY877 system.
The structures highlight the hydrogen bonds present at (*B*) the N-terminal end of the A-loop and (*C*) the C-terminal end of the A-loop. The bonds maintain the A-loop in the active state while ensuring availability of D863 and D845 for catalysis. Hydrogen bonds are depicted as *blue dashed lines*.

Intriguingly, we observe that phosphorylated Y877 performs a key function in linking the network of hydrogen bonds that maintain the A-loop in its extended form. The

phosphoryl group forms bonds with R844, K883, and R868, thereby bridging the C- and N-terminal ends of the A-loop (Fig. 3.8A). Furthermore, the main-chain oxygen of Y877 hydrogen-bonds with F899, contributing to the pattern of fastening interactions in the A-loop. Contrastingly, fewer hydrogen bonds occur in the A-loop of the Y877-unphosphorylated active system, indicating that the absence of the phosphoryl group results in a decreased number of intra-A-loop interactions (3.8B). We define a similar role for phosphorylated Y845 in the active EGFR system. Y845-phosphorylated EGFR shares eight of nine hydrogen bonds present in the A-loop of Y877-phosphorylated active HER2, including bonds between the phosphoryl group and the A-loop[181]. Our results for HER2 and EGFR suggest that the role of the phosphoryl group in Y877 (or Y845 in EGFR) is to bridge the stabilizing bonds on either side of the A-loop in the active system. In further support of this bridging mechanism of phosphorylated Y877 in HER2 and Y845 in EGFR, we note an analogous function for the phosphorylated tyrosine residue in insulin receptor tyrosine kinase (IRK), for which there exists a crystal structure of the phosphorylated active form of the protein [72].

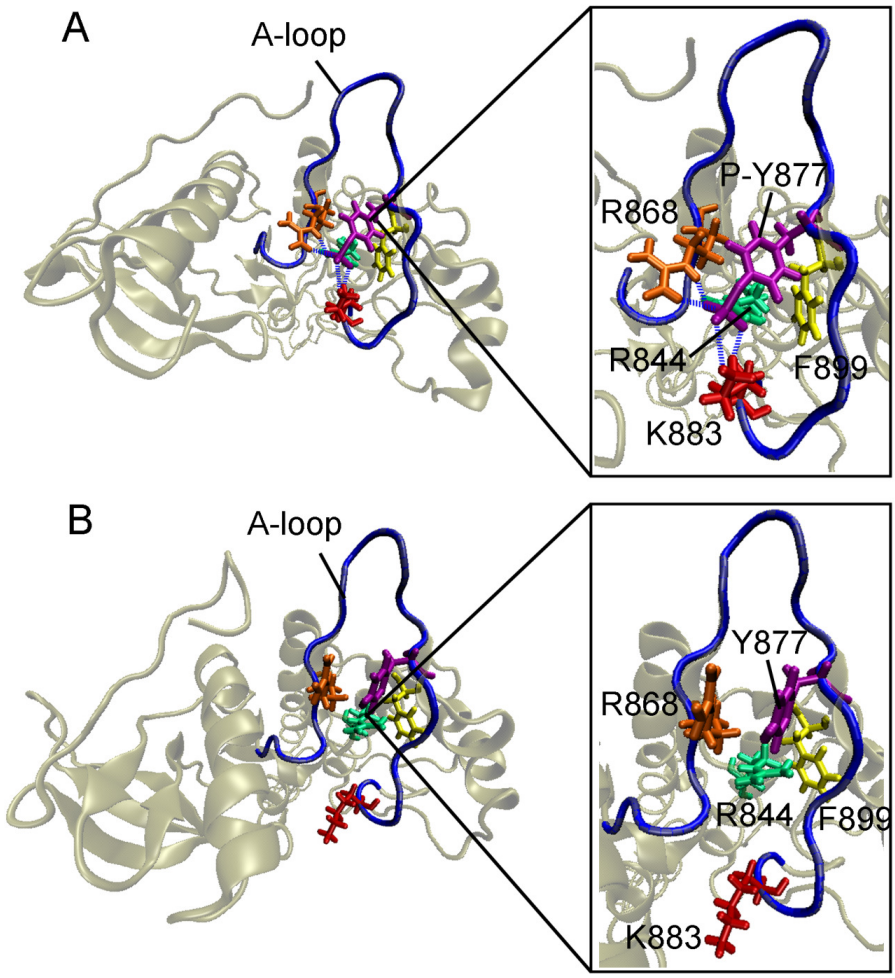

Figure 3.8. Bridging role of pY877.
(A) Snapshot of the Y877-phosphorylated active HER2 system, illustrating the role of phosphorylated Y877 in bridging the network of hydrogen bonds on either side of the A-loop. The phosphoryl group forms bonds with R844, K883, R868 and F899, thereby connecting the N- and C-terminal ends of the A-loop. Hydrogen bonds are depicted as *blue dashed lines*. (B) Snapshot of the unphosphorylated active HER2 system, showing the lack of stabilizing hydrogen bonds due to the absence of the phosphoryl group.

Analysis of the MD trajectories for the inactive systems reveals that phosphorylation of Y877 promotes formation of additional hydrogen bonds and salt bridges in the A-loop (see Table S4 in [181]). The phosphoryl group hydrogen-bonds with R844 and K883, spanning the A-loop as it does in the active system. Additional pairs of residues include G865-H843 and R868-R840, which parallel L866-R844 and R868-V842 in the active

structure. Although phosphorylation of Y877 in the inactive system alters the hydrogen bonding pattern so that it more closely resembles the network in the active system, phosphorylation is insufficient for promotion of conformational shifting to the active state within the short timescale of our 10 ns simulations.

To further investigate the effect of Y877-phosphorylation on kinase activity, we employed the free energy perturbation (FEP) method to calculate the Helmholtz free energy difference between the Y877-unphosphorylated and Y877-phosphorylated states in the NVT ensemble. The alchemical transformations were performed using the dual-topology paradigm [55], in which the initial and the final states are defined in terms of distinct, noninteracting topologies, and the interactions of the transformed atoms with their environment are scaled in terms of a linear parameter, λ (see Methods for details). Four different simulations were performed, 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. Replacement of Y877 with phosphorylated Y877 resulted in a free energy change of -385.1 $\pm$ 1.2 kcal/mol for the active structure and -384.0 $\pm$ 0.8 kcal/mol for the inactive structure, yielding a $\Delta\Delta F$ value of -1.1 $\pm$ 1.4 kcal/mol (Fig. 3.9A). The ΔF values for the reverse transformation, pY877 to Y877, were calculated in a similar manner, and were found to be 405.8 $\pm$ 1.1 kcal/mol for the phosphorylated active structure and 404.6 $\pm$ 1.1 kcal/mol for the phosphorylated inactive system, resulting in a $\Delta\Delta F$ value of 1.2 $\pm$ 1.5 kcal/mol (Fig. 3.9B). The individual ΔF values for the forward and reverse transformations differ slightly, as the structures contain different numbers of water molecules and ions, yet the error estimates are within the ranges that have been computed for other solvated systems [40, 67, 210]. The

ΔΔF values are significantly close and indicate 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. Additional perturbations are required for full catalytic competency of the kinase, suggesting that phosphorylation of Y877 is unlikely to be the primary stimulus for shifting to the active state.

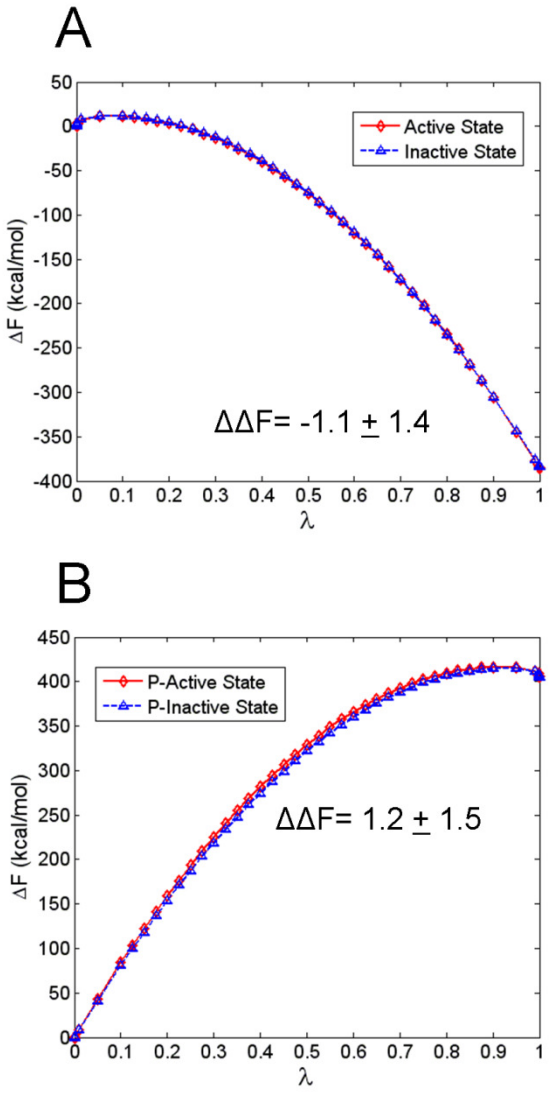

Figure 3.9. Evolution of the Helmholtz free energy as a function of the coupling parameter λ for the HER2 systems.

FEP simulations were performed for (A) the transformation of Y877 to pY877 and for (B) the transformation of pY877 to Y877 in both the inactive and active HER2 systems. Phosphorylation of Y877 stabilizes the inactive and active structures, while removal of the phosphoryl group results in an increase in the free energy.

It is noteworthy that differences in specific interactions between phosphorylated and unphosphorylated systems identified in the FEP trajectories also appear in our hydrogen bonding analysis. In the active structure, replacement of Y877 with pY877 produces several new hydrogen bonds, including R844-Y877, K883-Y877, and R868-Y877, all of which bridge the C- and N-terminal ends of the A-loop. Toward the end of the transformation, the main-chain oxygen of Y877 pairs with F899, which is consistent with the pattern of hydrogen bonds highlighted in the previous section. The perturbation of Y877 into pY877 in the inactive system effects formation of additional hydrogen bonds, such as R844-Y877 and K883-Y877, which link the ends of the A-loop as they do in the active system. Thus, the FEP results are consistent with our hydrogen bonding analysis, validating the robustness of the identified interactions despite the differences in simulation protocols.

3.4 Discussion

Given the involvement of the ErbB receptor tyrosine kinases in a wide range of human diseases, including schizophrenia and various types of cancer [106], it is imperative to understand their mechanism of activation at the molecular level. In this Chapter, we have investigated the mechanisms important in ErbB kinase domain regulation, as well as the molecular basis for HER2's unique mode of activation. Our MD simulations provide us with a framework for studying the atomistic behavior of the ErbB kinase domains in both monomeric and dimeric contexts. We have characterized the global motions of the ErbB

systems using principal component analysis and rationalized the differences in fluctuations in terms of specific interactions, namely, salt bridges and hydrogen bonds. Specifically, the existence of an extensive bonding network among the four key domains of the kinase in its active state correlates with the concerted motions of the four main loops as revealed by the PCA. The tight coupling of the A-loop and C-loop in the active system is especially pronounced, leading us to postulate the significance of the interaction in alignment of the catalytic residues. We also identify several commonalities in hydrogen bonding patterns among the three ErbB kinases in the A-loop, C-loop, and αC helix. Despite these shared trends in hydrogen bonds in the key activation loops, the bonding pattern in HER2 differs markedly in the αC-β4 loop. HER2 lacks many of the hydrogen bonds that occur in the αC-β4 loop in EGFR and HER4 due to the presence of the hydrophobic patch of residues in the αC-β4 region. The unique hydrogen bonding network in the αC-β4 loop in HER2 is relevant to the putative hydrophobic interaction between HER2 and the molecular chaperone Hsp90. The association between HER2 and Hsp90 serves a regulatory role in preventing HER2 dimerization and subsequent activation [30]. Unlike the other ErbB family members, the extracellular domain of HER2 is poised for dimerization in the absence of ligand binding [27], rationalizing the requirement for such tight regulatory mechanisms. The decrease in hydrophobic character 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 HER4. Such mutations are expected to be activating by disrupting the HER2-Hsp90 association.

We extended our analysis of global kinase motions and hydrogen bonding patterns to several different ErbB dimer systems, including a HER2-EGFR heterodimer, as well as EGFR and HER4 homodimers. Molecular dynamics simulations of the dimeric systems result in destabilization of several bonds present in the inactive monomeric structures, including interactions in the αC helix, as the αC helix comprises much of the dimerization interface for the activated kinase. The disruption of the hydrogen bonding pattern in the dimers provides a rationale for the observed conformational rearrangement of the αC helix toward the active state. Furthermore, we identify several interactions that persist throughout the dimeric trajectories of the inactive states and are hence candidates for further investigation through free energy methods such as umbrella sampling simulations, as the molecular environment surrounding such residues may define the pathway for conformational change and the associated barriers to activation. Owing to the fact that the residues involved in the N- and C-lobe faces of the dimer are essentially invariant among the ErbB family members, our results can in principle be extended to predict the behavior of other ErbB dimer combinations.

We have also investigated the effect of phosphorylation of the regulatory tyrosine residue, Y877 in HER2, on kinase activation. The results of our FEP simulations support the hypothesis that phosphorylation of Y877 in HER2 is unlikely to provide a dominant stimulus for activation. However, our structural and 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 stabilizing residues at the C- and N-terminal ends of the loop, thereby considerably altering the conformational environment surrounding the A-loop. Consistent with these findings, although phosphorylation of the

analogous A-loop tyrosine residue Y845 is unnecessary for catalytic activity of EGFR, it is known to significantly alter downstream signaling events, including activation of STAT5b and EGF-induced DNA synthesis [15, 90]. Likewise, Ishizawar et al. have postulated a role for phosphorylated Y877 in enhancing HER2-HER3 heterodimer formation by potentially changing the conformation of the kinase and engaging other molecules [78]. Hence the involvement of phosphorylated Y877 in the network of fastening interactions in the A-loop and the resultant alteration in the conformational environment may influence recruitment of signaling mediators involved in mitogenesis and other downstream processes. We also note that the residues involved in fastening the A-loop, which include R897, R898, and F899 in the αF helix, warrant further investigation, as mutation of these residues may destabilize the fastening bonding network and this effect may further elucidate the role of Y877-phosphorylation in HER2. These predictions can be experimentally tested and validated using HER2 mutagenesis assays.

The protein kinase genes are among the most frequently mutated in human cancers, and several ErbB kinase domain mutants have been determined. Our simulations provide insight into the effect of these mutations in the αC-β4 loop region of the ErbB kinases through assessment of structural dynamics and hydrogen bonding patterns in the kinase domains. Our results are consistent with a large body of experimental data [45, 72, 141, 202, 208] and provide a framework for highlighting the most crucial bonding interactions in the monomeric and dimeric ErbB systems. Elucidation of the molecular regulatory mechanisms will help establish structure-function relationships in the wildtype ErbB kinases as well as predict mutations with propensity for constitutive activation. Such

molecular variants in the EGFR, HER2, and HER4 RTKs are known to profoundly impact specific therapies targeting ErbB-mediated cancers.

Chapter 4

Multiscale Modeling of the Anti-cancer Role of the HER4 Tyrosine Kinase: Molecular Scale

In Chapters 4 and 5, we present a multiscale model of another member of the ErbB family, HER4/ErbB4. Chapter 4 focuses on molecular scale simulations of dimer-mediated activation in HER4, in order to elucidate the mechanisms by which the WT kinase is regulated and activated, and to rationalize the effects of several HER4 somatic mutants which have recently been discovered in a subset of cancer patients. Chapter 5 translates our molecular scale results of HER4 kinase activity into a cellular pathway model of WT versus mutant HER4 signaling, in which the WT and mutant cells trigger divergent signaling networks. Chapter 5 also discusses several experimental studies which we have performed in a HER4-stimulated mammary epithelial cell line, in support of our computational predictions.

4.1 Introduction

Whereas deregulation of the ErbB kinases is associated with many types of human cancer, HER4 kinase has recently been shown play an anti-carcinogenic role in certain tumors[48]. One mechanism by which HER4 is thought to impede tumor progression in mammary cells is through the activation of genes that promote cellular differentiation and inhibit proliferation, in effect, steering the cell away from a program of uncontrolled growth and instead toward a program of differentiation [80, 187]. HER4 is unique from the

other ErbB receptors in that binding of the ligands neuregulin (NRG) or heparin-binding epidermal growth factor (HB-EGF) induces proteolytic cleavage of the 80 kDa kinase domain, termed the s80 or soluble cleavage product, and binding of s80 to the transcription factor (TF) STAT5a. The s80-STAT5a complex then translocates to the nucleus to regulate expression of genes involved in mammary cell differentiation pathways, including the milk protein genes β-casein and whey acidic protein (WAP)[117, 197]. Hence studies are underway to determine the molecular pathways that are stimulated by the soluble HER4 protein, particularly the network of transcriptional regulatory elements that are activated upon nuclear translocation of HER4 and STAT5a. Delineation of the transcriptional regulatory network associated with HER4/STAT5a activity would enable the exploitation of the pathway for targeting of malignant cells. Specifically, activation of HER4 signaling in aggressive breast tumors would present a novel therapeutic approach to suppress growth of these malignancies.

The HER4-mediated proliferation-to-differentiation switch in mammary cells may be encoded in terms of differential spatial and temporal regulation of HER4 activity. The phenotypic response of the mammary cell to HER4 stimulation is associated with a specific subcellular context, and the shuttling of HER4 among various compartments may determine the cellular decision to proliferate or differentiate. For instance, the membrane-bound HER4 kinase may dimerize with HER2 to produce a mitogenic response in the mammary cell, whereas the soluble HER4 domain (s80) translocates to the nucleus to effect cellular differentiation. Hence a comparison of HER4 activity at each relevant spatio-temporal scale would aid in delineation of the mechanisms by which the proliferation-to-differentiation switch is regulated.

Despite the unique role played by wild-type (WT) HER4 in certain tumor types, several somatic mutations in the kinase domain of HER4 were recently discovered in patients with breast, gastric, colorectal, and non-small cell lung cancer[171]. A study by Tvorogov et al.[191] analyzed the basal and ligand-induced phosphorylation levels of the 9 HER4 kinase domain mutants in various cell backgrounds, and found that two of the mutations, G802dup and D861Y, disrupt the catalytic activity of the HER4 kinase, whereas the remaining mutations do not appear to have an effect on HER4 phosphorylation. The group discovered that, despite loss of kinase activity, the two mutant HER4 receptors were able to heterodimerize with HER2 and signal through the ERK and PI3K/AKT pathways. However, kinase activity of HER4 was required for ligand-induced activation of the STAT5a differentiation pathway, and as a result, the HER4 mutants were unable to activate STAT5a signaling or cell differentiation. Not only were the two mutants unable to induce cell differentiation, but when over-expressed in a breast cancer cell line, they actively suppressed the formation of differentiated acinar structures, in contrast to wild-type HER4, which induced differentiation. Thus the HER4 mutants exhibit a selective loss-of-function phenotype which is biased toward proliferative signaling pathways and against the cell differentiation pathway.

In this Chapter, we begin by simulating and analyzing a molecular model of the WT HER4 homodimer, in order to elucidate molecular mechanisms of activation in the WT kinase. We then apply the results of our WT simulations to help rationalize the effects of the clinically-identified HER4 somatic mutants on the cell phenotype.

4.2 Materials and Methods

4.2.1 Construction of the HER4 homodimer

The HER4 homodimer system was modeled on the structure of the HER4 dimer published by Qiu et al.[141]. We modeled the receiver kinase as a HER4 monomer in the inactive conformation, and the activator kinase as a HER4 monomer in the active conformation. The structure was minimized to remove unfavorable contacts and hydrogen atoms were added using the hbuild routine in CHARMM[112]. The dimer was explicitly solvated with the buffering distance set to 15 Å for a total system size of approximately 120,000 atoms. Sodium (Na^+) and chloride (Cl^-) ions were added to achieve net electroneutrality of the system and an ionic strength of 75 mM. All Na^+ and Cl^- ions were placed at least 8 Å away from any protein atoms and from each other. Minimization and MD steps were performed for a total of 10 ns. A second dimer was constructed in which the receiver kinase is in the active form, to use as the target structure for the TMD simulations. This dimer was also simulated (MD) for 10ns.

4.2.2 Targeted molecular dynamics (TMD) simulations

TMD simulations were implemented in NAMD[135]. In TMD simulations, a chosen subset of atoms is steered toward a target structure, while the other atoms may also move in response to this structural change. Hence, this approach can simulate large scale conformational transitions in biomolecules which are otherwise difficult or impossible to realize using conventional MD simulations. The RMSD between the current structure and the target structure was used to calculate an additional steering force in TMD simulations, which was applied on every atom from the chosen molecular subset, and is calculated as follows:

$$U_{TMD} = 0.5\frac{k}{N}[RMSD(t) - RMSD_0(t)]^2$$

RMSD(t) represents the RMSD, for a chosen molecular subset, between the current structure and the reference target structure at simulation time t. $RMSD_0(t)$ is the target RMSD value at simulation time t. The force constant, k, was tested for values ranging from 0.5-20 kcal/mol/Å^2, and ultimately, a value of 20 was chosen. N represents the number of atoms included in the target subset. During optimization of our TMD simulations, a variety of molecular subsets were tested, including all backbone atoms of the receiver kinase or of the catalytic site only. Ultimately, a subset which included the heavy atoms of the catalytic site was chosen. TMD simulations were run for 25 ns with a time step of 2 fs. The conformational transition from the initial to the target structure was determined by decreasing the value of $RMSD_0(t)$ as a function of the simulation time. The value of $RMSD_0(t)$ was linearly decreased to 0 Angstroms during the 25 ns TMD simulation. Upon reaching the target structure, MD simulations were run to allow the system to fully relax into the target conformation.

4.2.3 Analysis of the TMD trajectories

Several types of analyses were performed on the TMD trajectories. Hydrogen bonding analysis was performed in order to define those bonds which broke or formed during the course of the TMD simulation. Hydrogen bonds were defined by a bond length cutoff of 3.4 Å and an angle cutoff of 150º. Bonds that fulfilled these criteria and were present in at least 60% of the trajectory were tabulated in CHARMM. Salt bridges were defined as hydrogen bonds occurring between an acidic and a basic residue and satisfying a bond

length cutoff of 1.6 Å. All hydrogen bonds and salt bridges were also visualized in VMD[74] for the duration of the 10 ns simulation.

Ramachandran plots were constructed for key residues, in order to track any flips or changes in the dihedral angles phi and psi. The phi and psi angles were computed for the specified residues for each frame of the TMD trajectory, using VMD. The residues chosen included R841, L843, S749, and F872, as these residues undergo bonding changes during transition to the active conformation.

Root mean square deviation (RMSD) values were measured for the dimer undergoing TMD, with respect to the active (target) structure and the inactive (starting) structure. The RMSD was measured with respect to the backbone atoms of the catalytic site, and also with respect to the entire receiver kinase, to gauge the progression of the TMD system along the activation pathway.

4.3 Results

4.3.1 Simulation of the HER4 homodimer

Based on the asymmetric dimer interface observed in both the EGFR and HER4 crystal structures, we constructed a HER4 homodimer system in which an inactive HER4 monomer serves as the receiver kinase and an active HER4 monomer serves as the activator kinase (see Methods). We note that the activator kinase may be in either the active or inactive conformation, as the dimer interface is the same for both conformations. The dimer interface residues comprise portions of the C-lobe in the activator kinase and segments of the N-lobe (including the juxtamembrane segment and the αC helix) in the receiver kinase (Fig. 4.1). The dimer system was solvated and

energy-minimized prior to TMD simulation (see Methods), to remove any unfavorable contacts.

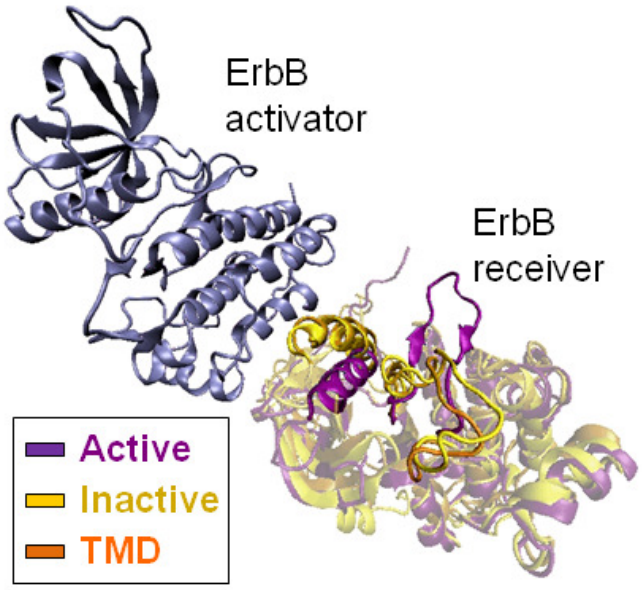

Figure 4.1. The ErbB dimer system.
The system consists of an activator kinase (in either the inactive or active conformation) and a receiver kinase (in the inactive conformation). The inactive state is displayed in yellow, the starting state for the TMD simulation is shown in orange, and the fully activated state is in purple. During activation, the most significant motions occur in the A-loop and αC helix. The A-loop extends to its open, active state, and the αC helix shifts into the catalytic site.

Recently, the 37-residue segment connecting the kinase domain to the ErbB transmembrane domain, or the juxtamembrane (JM) domain, has been shown to play an important role in the complete activation of the ErbB RTKs[144, 185]. Red Brewer et al.[144] demonstrated that mutations in the C-terminal 19 residues of the EGFR JM domain abolish EGFR activation. Furthermore, Thiel and Carpenter[185] showed that the deletion of the JM region results in significant loss of EGFR kinase phosphorylation. Although the molecular mechanism underlying the activating role of the JM region is unknown, a recent crystal structure of the EGFR kinase dimer[144], which included the complete JM domain, revealed that the JM region in the receiver kinase makes extensive contacts with

the C-lobe of the activator kinase, essentially forming a latch that 'cradles' the activator kinase to stabilize the asymmetric dimer interface. Although there have been several reported TMD studies of the EGFR dimer[38, 134], none have included the JM domain; rather, they have utilized EGFR structures in which the receiver kinase is truncated at the N-terminus of the kinase domain. Thus, not only do we report the first (which we are aware of) TMD study of the HER4 dimer, our study is also the first to include the crucial JM region.

In addition to the TMD simulations, we also constructed and simulated two control systems: 10ns molecular dynamics simulations of a HER4 dimer in which the receiver kinase is in the activated state, as well as in the inactive state. The purpose of the control simulations is to provide a frame of reference for analysis of any significant motions or changes in bond patterns observed in the TMD simulations. No significant motions were revealed for either control simulation, although the active dimer is more stable than is the inactive dimer. A principal component analysis of the control system trajectories revealed that the inactive dimer does move slightly toward the active conformation, in that there is a dominant shifting motion of its αC helix into the active site, with relatively large fluctuations (Fig. 4.2A). As in our previous simulations of the EGFR dimer (Fig. 4.2B), the dimer interface provides a stimulus to shift the inactive receiver kinase (mainly, the αC helix) away from the inactive conformation and toward the active state. The A-loop, which is not positioned directly in the dimer interface, does not globally shift during the control simulations, and most likely requires other perturbations in order to transition to its open, extended state. Within the short time scale of our MD simulations, the presence of the dimer interface is insufficient to fully activate

the receiver kinase, an event which, in actuality, would occur on the order of milliseconds to seconds.

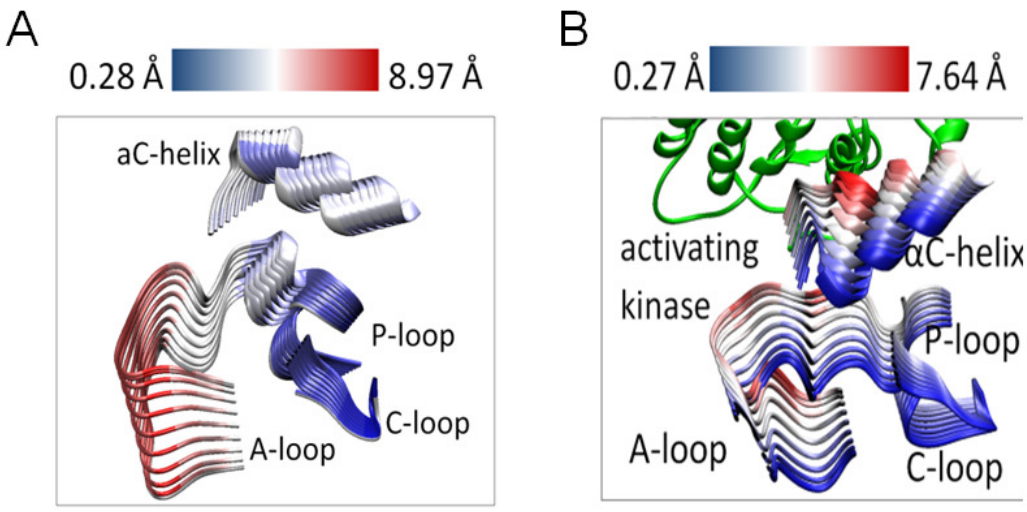

Figure 4.2. PCA of the HER4 dimer.
PCA of the 10ns control MD trajectories for (A) the HER4 homodimer system and (B) the EGFR homodimer system. The PCA was performed on the catalytic site (A-loop, P-loop, C-loop, αC helix) of the receiver kinase. Motions along the first eigenmode are displayed by superimposing several frames from each PCA trajectory. The structures are colored according to the RMSD, where blue regions indicate small fluctuations and red regions indicate larger fluctuations. The fluctuations are quantified according to the scale bar at top.

4.3.2 Optimization of TMD simulation restraints

As the catalytic sub-domains undergoing the most pronounced conformational shifts during the transition from the inactive to the active state are the A-loop and αC helix, we included these regions in our set of TMD restraints. Previous TMD studies of the EGFR kinase[38, 134] were biased solely toward these regions; however, other regions of the kinase may also be important in the activation mechanism. To incorporate a more global view of the allosteric activation mechanism, we extended our list of TMD restraints to include the entire catalytic site. However, it is important to not over-constrain the system, and so we allowed the remainder of the kinase, including the JM domain, to move freely. Additional

degrees of freedom are allowed in that the activator kinase is also completely free to move, as it provides the activation stimulus, but does not undergo any well-defined conformational changes itself.

We experimented with several different sets of restraints for the TMD simulations. The spring constant used to apply force to the dimer system (see Methods) was varied from k=0.5-20 kcal/mol/Å^2. In addition to optimizing the force constant, we varied the set of targeted atoms. Specifically, we tested the following sets of target atoms: all backbone atoms of the receiver kinase, heavy atoms of the A-loop and αC helix, alpha carbons of the catalytic site, and backbone atoms of the catalytic site. Restraining the heavy atoms of the catalytic site produced the most successful result in terms of achieving the target conformation. We also tested the application of restraints to the JM domain, but as it is a terminus, this resulted in overconstraint of the system. The TMD simulation time was varied from 2-25ns. The longer the simulation and the smaller the force constant, the more accurate the results, and so we tested several different combinations of simulation time and force constant. We found that the minimum force constant required for our large dimer system (approximately 120,000 atoms) was k=5 kcal/mol/Å^2, which, following a 10ns TMD simulation, shifted the dimer by 1 Angstrom toward the target conformation. Application of a slightly larger force constant, k=10 kcal/mol/Å^2, resulted in a small improvement, shifting the dimer by about 2 Angstroms toward the active conformation. However, in running these simulations longer than 10-12ns, we did not see an improvement in the RMSD, suggesting that, using a force constant of 5-10 kcal/mol/Å^2, a prohibitively long simulation is required for such a large dimer system.

A force constant of k=20 kcal/mol/Å^2 proved to be the minimum value required to achieve the target conformation in a reasonable simulation time, 25ns. In order to gauge whether the dimer had achieved its active state, we calculated the RMSD with respect to the active dimer (target structure) and also analyzed the pattern of hydrogen bond changes during the TMD simulation. For the k=20, 25ns simulation, the final RMSD of the backbone atoms of the receiver kinase with respect to the target structure was approximately 1.2 Angstroms, whereas the receiver had moved approximately 3.6 Angstroms away from the starting (inactive) conformation (Fig. 4.3). Due to the flexible nature of the A-loop and several other loop regions in the kinase, we do not aim to attain an RMSD of 0 Angstroms. Rather, we focus on the key changes in the hydrogen bonding pattern and the catalytic sub-domain conformations. We also ensure that the A-loop is in its open, extended form, allowing for the binding of peptide and ATP.

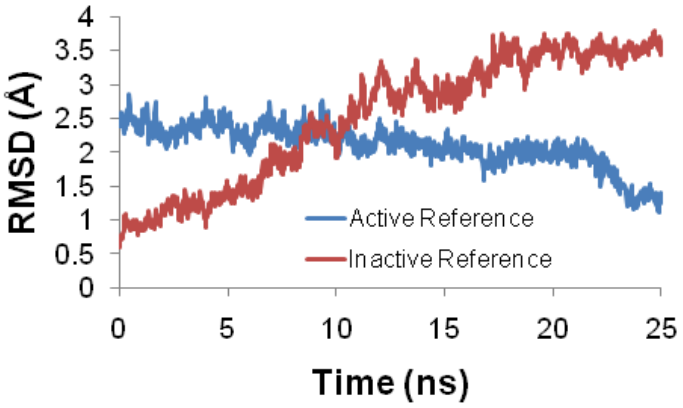

Figure 4.3. Transition of the HER4 homodimer system during the TMD simulation.
The RMSD is plotted for all backbone atoms of the receiver kinase, with respect to the active (target) structure (*blue*) and with respect to the inactive (starting) structure (*red*). Results are shown for the simulation in which t=25ns and k=20 kcal/mol/Å^2.

4.3.3 Analysis of global changes during the TMD simulation

Upon optimization of the parameters and restraints for the TMD simulation, the TMD trajectory was analyzed using several methods. First, global changes in the HER4 system were assessed, to gain insight into the sequence of molecular events required to achieve the target (active) conformation. Figure 4.4 tracks the RMSD of the backbone atoms of the A-loop and αC helix, two of the major catalytic sub-domains in the HER4 active site, during the progression of the 25ns TMD simulation. Relative to the active (target) structure, the A-loop moves from an RMSD of approximately 14 to 4 Angstroms, and relative to the inactive (starting) structure, it moves from an RMSD of 1 to 12 Angstroms (Figure 4.4). The αC helix moves from 4 to 2 Angstroms away from the active structure, and from 1 to 4 Angstroms away from the inactive structure. Interestingly, the αC helix becomes closer to the active conformation (relative to the inactive conformation) at approximately 18ns, whereas the A-loop becomes closer to the active conformation a bit later, at 21-22ns (Figure 4.4). This order of events (the αC helix shifting before the A-loop) can partially be rationalized by the location of the αC helix within the dimer interface, such that it is expected to be directly impacted by this stimulus (dimer interface). We have also observed this sequence of molecular events for the EGFR dimer. This mechanism may be unique to the ErbB kinases in that other kinase families, such as Src, are thought to undergo the reverse order of events (A-loop moves before αC helix[205]).

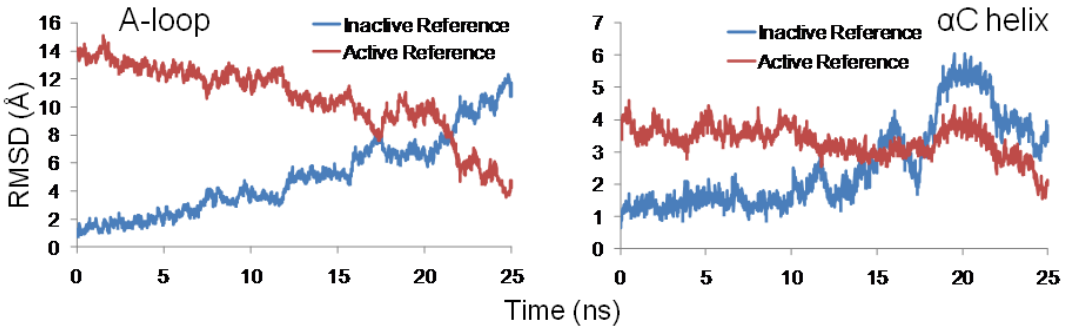

Figure 4.4. Transition of the A-loop and the αC helix toward the active (target) conformation during TMD.
The RMSD is plotted for the backbone atoms of the A-loop and αC helix, in reference to the active structure (*red*) and the inactive structure (*blue*).

4.3.4 Analysis of local interactions during the TMD simulation

In order to translate the global pattern of sub-domain motion that we observe for the TMD system into changes in specific, or local, interactions, we next analyzed the TMD trajectory for hydrogen bonds and salt bridges which may have broken or formed during the conformational transition to the active state. As the pattern of hydrogen bonds maintaining the inactive state differs significantly from the bond pattern in the active state, we expect to observe these changes during the course of the TMD. Table 4.1 summarizes all bonds which have either broken or formed during the transition from the inactive to the active conformation. The E743-K726 salt bridge, which is conserved among the active ErbB kinases and is required for the coordination of the α and β phosphates of ATP, represents one of the key bonds which must form during kinase activation. In the inactive state, there are two residues which sequester E743 and K726, thus autoinhibiting kinase activation by preventing formation of the key salt bridge. These two 'sequestering' bonds are D836-K726 and R841-E743, which must break during the activation process. Indeed, in our TMD simulation, these two bonds have

broken, freeing the E743 and K726 residues to form the crucial salt bridge (Table 4.1 and Fig. 4.5). Figure 4.5 tracks the change in bond length for the R841-E743 sequestering bond during the course of the TMD simulation. In comparison to the 10ns MD control simulation (Fig. 4.5, top panel), in which the R841-E743 bond spontaneously forms and breaks during the course of the simulation, the bond successfully breaks (at *t*=18ns) during the 25ns TMD simulation (Fig. 4.5, bottom panel). The Ramachandran plot in Figure 4.6A tracks the change in the dihedral angles phi and psi for the R841 residue (see Methods), revealing that the angle phi flips from a small negative value to a large positive value during the conformational transition, facilitating breakage of the R841-E743 bond. Other interactions which are severed during the TMD simulation include E730-K856, which allows the E739-K856 bond (between the A-loop and αC helix) to form, and the E847-R870 bond in the A-loop.

Bond	TMD result
W712-E692	Broken
R695-Q768	Formed
E715-K689	Broken
P722-L773	Formed
K726-D836	Broken
L728-T766	Broken
E730-K856	Broken
E739-K856	Formed
E743-R841	Broken
E743-K726	Formed
S749-N681	Weakened
R817-D846	Weakened
E847-R870	Broken

Table 4.1. Bonds formed or broken during the HER4 TMD simulation.

Another important interaction is L843-R813, which does not appear in the inactive dimer, but forms during the TMD simulation. This bond couples the A-loop and C-loop in the active state, and helps to maintain the A-loop in its open, extended form. In fact, several bonds bridging the A-loop and C-loop occur in the active kinase, and can be considered 'fastening bonds', in that they stabilize the A-loop in its activated conformation. Several of these interactions were observed to form during our TMD simulation. Figure 4.6B depicts the Ramachandran plot for the L843 residue, in order to track any changes in the dihedral angles phi and psi. It is apparent from the plot that the angle phi flips from a small negative value to a large positive value during the course of the TMD simulation, facilitating the formation of the L843-R813 fastening bond.

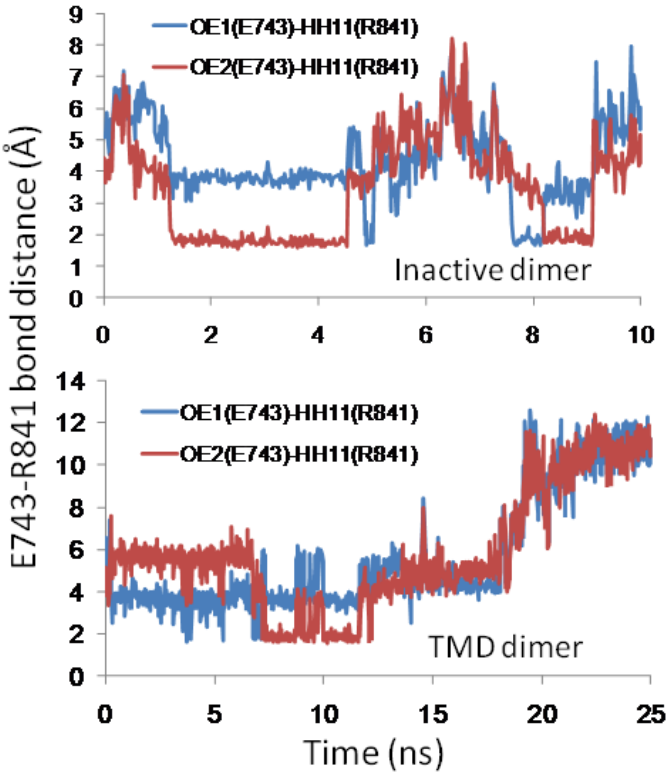

Figure 4.5. Bond distance for the E743-R841 salt bridge during the TMD simulation. The top panel displays the bond distance in the inactive dimer (10ns MD control simulation), in which the bond spontaneously breaks and forms during the course of the simulation. The bottom panel displays the bond distance in the 25ns TMD simulation; at *t*=18ns, the bond breaks.

In addition to the bonding pattern in the receiver kinase changing in response to the TMD restraints, several inter-dimer bonds (*i.e.,* bonds between the receiver and activator kinases) also broke during the 25ns TMD. Specifically, the D742-K917 bond, which occurs between the αC helix of the receiver kinase and the C-lobe of the activator kinase, is significantly weakened. The S764-E896 bond fully breaks during the TMD simulation. Such inter-dimer bonds have not been explored in previous TMD simulations of the EGFR kinase, yet may be worthwhile to investigate further, as interactions between previously-overlooked regions of the receiver and activator monomers may play a role in kinase activation.

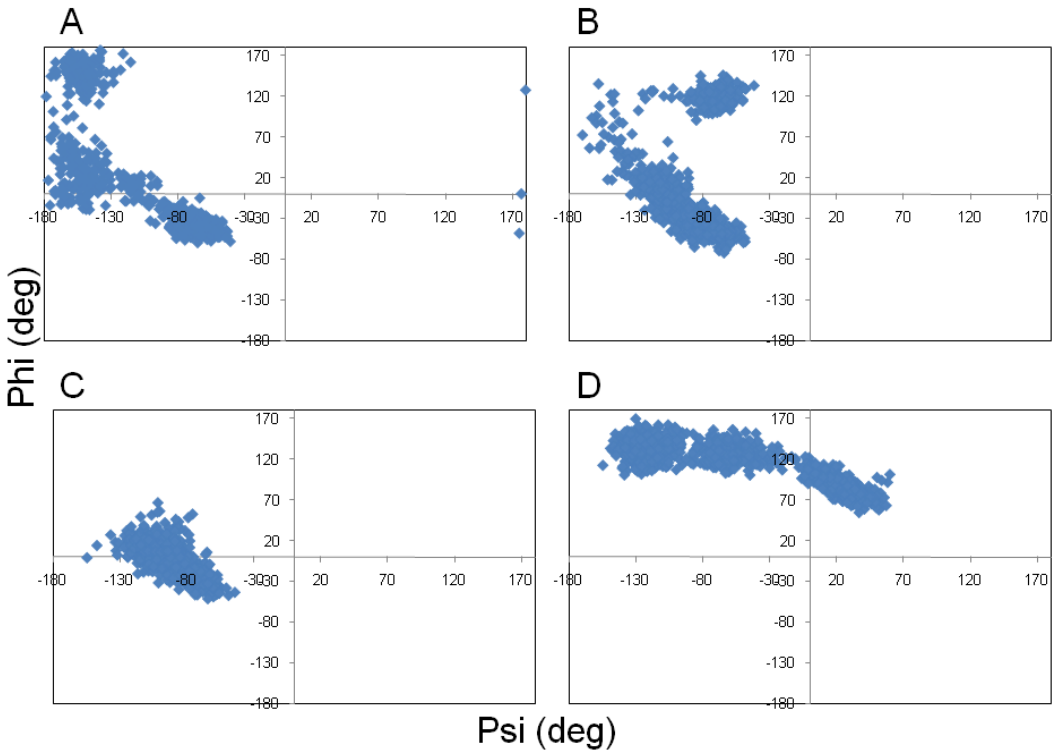

Figure 4.6. Ramachandran plots for key residues during the 25ns TMD simulation.
(A) R841, (B) L843, (C) S749, and (D) F872. Each data point marks a time point in the TMD simulation; in this way, any flipping of the dihedral angles can be detected.

The bonding pattern in the JM region of the receiver kinase also shifts during the course of the TMD simulation. The following bonds change: N681-S749 is weakened, E690-S764 forms, and K689-E715 and E692-W712 break. In addition, one bond forms between the JM domain of the receiver kinase and the C-lobe of the activator kinase: L685-Q903. Figure 4.6C displays the Ramachandran plot for S749, revealing that the phi angle flips from a negative to a postive value. Cross-referencing these JM residues with the list of amino acids shown to abolish EGFR phosphorylation in the scanning alanine mutagenesis assay performed by Red Brewer et al.[144], we discovered significant overlap. In particular, N681 (N676 in EGFR) and L685 (L680 in EGFR) both participate in JM interactions that are altered during our HER4 TMD simulation, and mutation of the analogous residues in EGFR was shown to abrogate kinase activity[144]. Thus our results for the HER4 dimer are in agreement with the experimental results, which emphasize the importance of the JM region as an important component of the activation mechanism in the ErbB kinases (at least, EGFR and HER4).

4.3.5 Comparison to TMD of the EGFR dimer

In order to assess whether the TMD pathway is conserved across other members of the ErbB family, we performed TMD simulations of an EGFR homodimer, based on the EGFR crystal structure[208]. The full study is detailed in our previous work, but we will highlight the most relevant points here, for comparison with our HER4 TMD results. The EGFR TMD simulations were performed by applying the RMSD of the backbone atoms of the αC-helix and the A-loop as the reaction coordinate (χ_i). The activation pathway was divided into 20 smaller windows and each segment of the window was simulated individually to ensure adequate sampling. To enhance the sampling of low probability

events, a harmonic restraint of constant k=20 kcal/mol/Å^2 was applied along the reaction coordinate, χ_i.

The TMD study of the EGFR dimer captured the transition of the kinase from the inactive to the active conformation. The PCA of the TMD trajectory is depicted in Figure 4.7A. The TMD study did not completely sample the activation pathway, though we observed several significant molecular events, including the extension of the A-loop and the rotation of the αC helix into the active site (Fig. 4.7A). Several specific interactions were conserved across the EGFR and HER4 dimer TMD simulations, including breakage of the E738-K836 salt bridge (EGFR numbering), the formation of the crucial E738-K721 salt bridge, and the formation of the fastening bonds between the A-loop and C-loop. In similarity to the HER4 TMD results, the data for the EGFR TMD strongly suggested that the αC helix shifts before the A-loop, contrary to the pathway mapped out for the Src kinase Hck[205]. Despite several differences in bonding events, the major molecular events occurring in the HER4 TMD simulation are conserved in EGFR.

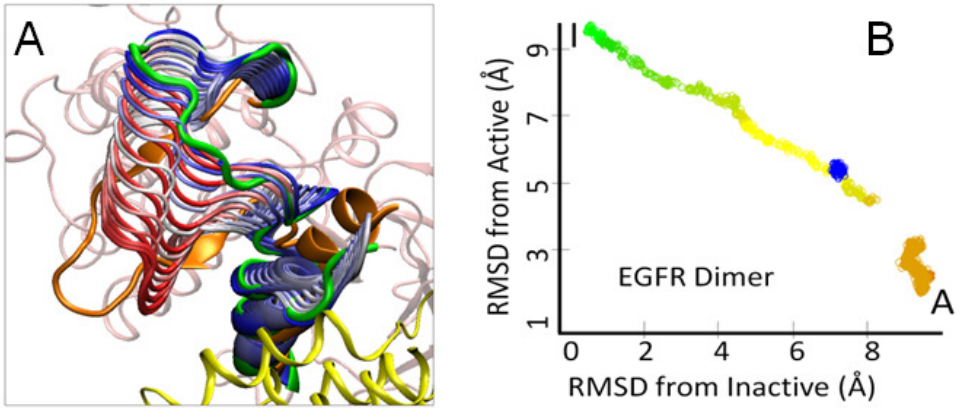

Figure 4.7. TMD of the EGFR dimer.
(A) PCA of the EGFR dimer TMD simulation. The active (target) structure is shown in orange, the inactive (starting) structure is in green, and the dimer undergoing TMD is colored according to RMSD value (in red/white/blue). The PCA motions are overlaid sequentially, where red regions indicate large motions and blue regions indicate smaller motions. The activator kinase is shown in yellow. (B) The TMD pathway for the EGFR dimer, showing the transition from the inactive state toward the active state.

4.3.6 Rationalizing the effects of the HER4 somatic mutants

As discussed in Section 4.1, two of the nine clinically-identified HER4 somatic mutants exhibit a selective loss-of-function phenotype in terms of the downstream signaling pathways which are preferentially activated. Specifically, the mutants signal through the proliferative ERK and PI3K/AKT pathways but not through the STAT5a differentiation pathway. To provide insight into the molecular mechanisms by which the HER4 mutants exert their preferential signaling effects, we assessed the mutants, D861Y (biochemical numbering is D836 and will be used henceforth) and G802dup (biochemical numbering is G777 and will be used henceforth), in the framework of our WT HER4 simulations. We first constructed the mutant structures in MODELLER, using our equilibrated WT HER4 structure as the template. Energy minimization was performed in MODELLER to remove any unfavorable contacts, and nearest neighbor amino acids were allowed to optimize their positions. The D836Y mutation is located in the DFG motif, a highly conserved amino acid segment which is positioned at the N-terminal end of the A-loop in many eukaryotic kinases, and is critical for efficient phosphotransfer. Based on our WT HER4 simulations and our mutant D836Y structure, it is apparent that in the WT kinase, D836 forms a crucial set of interactions with N823 and D818 (C-loop), the catalytic aspartic acid which deprotonates the hydroxyl group of the substrate tyrosine residue. The bulky tyrosine side chain in the D836Y mutant appears to disrupt these interactions, which are key in orienting the peptide substrate for phosphoryl transfer; thus the catalytic activity of the kinase is predicted to be abrogated. Furthermore, in our TMD simulation of the WT HER4 dimer, D836 sequesters the key activating salt bridge by bonding with K726, an interaction which is released in the active kinase. Thus D836 serves important roles in

both the inactive and active kinase conformations.

G777 forms part of the ATP-binding pocket in the HER4 kinase. Our structural mutant model revealed that the G777dup mutation perturbed the relative orientation of the energy-minimized neighboring residues, including H776, C778, and L779. Furthermore, other neighboring residues which were perturbed played key roles in our TMD simulation of the WT kinase. In particular, L773, which contributes to the ATP-binding pocket and is displaced in the G777dup mutant, bonds with P722 during the TMD simulation of the WT HER4 kinase. Although ATP is not explicitly included in our models, the disruption of the ATP-binding pocket introduced by the G777dup mutation is expected to significantly impair the proper docking of ATP into the catalytic site. Thus our simulations help to predict, at molecular resolution, the specific structural perturbations which are likely to be introduced by the D836Y and G777dup mutants. Due to the disrupted kinase activity of the HER4 mutants, they are unable to activate STAT5a (which requires HER4 kinase activity). Thus there is a bias toward the cellular proliferative signaling pathways which are induced by HER4 and ErbB heterodimers.

We also examined the two mutations in the context of the HER4 dimer interface. Our WT simulations confirmed that the mutations are unlikely to disrupt dimerization of HER4 directly, due to their position relative to our simulated dimer interface. The message supported by our simulations is that the mutants abolish kinase activity but not dimerization of HER4, a hypothesis which is underscored by the experiments of Tvorogov et al.[191]. Although the effects of these mutants may be predicted through analysis of the crystal structure of the HER4 homodimer, this gives only a static picture of the impact of such perturbations. Thus our HER4 simulations provide a dynamic

picture of the crucial roles played by D836 and G777 in the WT kinase activation mechanism.

4.4 Discussion

In this chapter, we analyzed the WT activation pathway for the HER4 kinase through simulations of a HER4 homodimer system. The simulations highlighted key molecular interactions which we predict are important in the HER4 activation mechanism. Specifically, we first assessed global patterns which emerged from simulation of the HER4 dimer system. We discovered that the order of events, in terms of motions of the individual catalytic sub-domains, involved the conformational shifting of the αC helix before the A-loop, which can partially be rationalized by the location of the αC helix in the dimer interface, such that it is expected to be directly impacted by this stimulus (dimer interface). We also observed this sequence of molecular events for the EGFR dimer. This mechanism may be unique to the ErbB kinases in that other kinase families, such as Src, are thought to undergo the reverse order of events (A-loop moves before αC helix). This difference may partially be due to the fact that phosphorylation of the A-loop tyrosine residue is required for the activation of Src kinase, whereas it does not seem to be necessary for ErbB kinase activation. Hence phosphorylation of the A-loop tyrosine residue in Src kinase may induce the A-loop to move earlier in the activation process.

The global motions of the HER4 dimer system were then transduced into local interactions through a hydrogen-bonding analysis of the dimer TMD trajectory. The bonding analysis, which identified key interactions which either broke or formed during the TMD simulation, highlighted a conserved auto-inhibitory mechanism. The E743-K726 salt bridge, which is conserved among the active ErbB kinases and is required for

coordination of the α and β phosphates of ATP, represents one of the key bonds which must form during kinase activation. In the inactive state, there are two residues which sequester E743 and K726, thus auto-inhibiting kinase activation by preventing formation of the key salt bridge. These two 'sequestering' bonds are D836-K726 and R841-E743, which must break during the activation process. Indeed, in our TMD simulation, these two bonds have severed, freeing the E743 and K726 residues to coordinate the crucial salt bridge. Additional local interactions which were highlighted include the formation of specific bonds which couple the A-loop and C-loop, such as L843-R813, helping to 'fasten' the A-loop in its open, extended conformation.

In addition to the bonding pattern in the receiver kinase changing in response to the TMD restraints, we also identified several inter-dimer bonds (*i.e.,* bonds between the receiver and activator kinases) which broke during the TMD simulation, including the D742-K917 bond, which occurs between the αC helix of the receiver kinase and the C-lobe of the activator kinase, and the S764-E896 bond. Moreover, the bonding pattern in the JM region of the receiver kinase shifts during the course of the TMD simulation, a result which is in agreement with previous experimental studies demonstrating that the JM domain is a crucial contributor to the kinase allosteric activation mechanism[144, 185]. These inter-dimer and JM domain interactions have not been explored in previous simulations of the EGFR kinase, yet our results indicate that they warrant further investigation, in order to piece together the complete activation pathway of the ErbB kinases.

We then applied the results of our WT analysis to rationalize the effects of two clinically-reported HER4 somatic mutations, D836Y and G777dup. We first constructed

the mutant structures in MODELLER, using our equilibrated WT HER4 structure as the template. Based on our WT HER4 simulations and our mutant D836Y structure, we found that the D836Y mutation disrupted interactions with the catalytic base in the C-loop, which is crucial for effective phospho-transfer. Similarly, the G777dup mutant perturbed the relative orientation of key residues in the ATP-binding pocket of HER4. We were able to cross-reference specific perturbed residues with bonds that were broken or formed during our TMD simulation of the WT HER4 kinase, underscoring the importance of D836 and G777 in the activation of the WT kinase. Our results confirmed the effects of the mutants on disruption of the catalytic site, and predicted resultant kinase impairment. However, as the mutants retain the ability to dimerize with other ErbB kinases, there is a bias toward cellular proliferative signaling and away from the STAT5a-mediated differentiation pathway.

The HER4 mutants are unique among the ErbB kinases in that similar mutations have been reported in EGFR and HER2, yet these mutations confer a gain-of-function phenotype for EGFR and HER2 [111, 131]. Specifically, tumors expressing these mutant EGFR and HER2 RTKs demonstrate increased ErbB phosphorylation and kinase activity. In addition, these gain-of-function mutations in EGFR and HER2 confer sensitivity of tumors to specific TKIs, suggesting that these tumors have become addicted to signaling through the mutant ErbB kinases. The mechanism by which these EGFR and HER2 mutants enhance ErbB signaling is thought to be through disruption of specific auto-inhibitory interactions in the hydrophobic core of the kinase. By contrast, similar mutations in HER4, which we have analyzed in this chapter, result in a selective loss-of-function phenotype, by abolishing HER4 kinase activity but not dimerization.

Furthermore, these HER4 mutants do not demonstrate sensitivity to specific TKIs[191], as do the analogous EGFR and HER2 mutants.

Based on our hypotheses regarding the HER4 mutants, parallels can be drawn to themes which emerged from our multiscale study of the HER3 kinase (see Chapter 2). In particular, we have shown, through our multiscale studies of the HER3 and HER4 kinases, that ErbB kinases devoid of intrinsic kinase activity are not necessarily benign in terms of their role in cell signaling and induction of cancer. Indeed, although the kinase activity of the ErbB RTKs plays a crucial function in cell signaling, there are other, equally important roles performed by the ErbB kinases, which allow for participation in signaling despite lack of catalytic activity. In the case of HER3 kinase, we found that, although HER3 exhibits relatively weak catalytic activity in comparison to the other ErbB family members, it plays a key role in the development of drug resistance to certain TKIs, due to its ability to synergize with other signaling processes in the cell (such as upregulation of HER3 expression) to amplify its weak signal. In the case of HER4, the kinase-dead HER4 somatic mutants retain the ability to induce proliferative signaling, due to their ability to dimerize with other ErbB RTKs. Indeed, the various activities of the ErbB kinases complement each other to produce a holistically functioning signaling unit, thus compensating for any deficiencies in the individual ErbB proteins. In certain tumor cells predisposed to uncontrolled signaling, this otherwise advantageous evolutionary characteristic of the ErbB system can have deleterious consequences, as we have illustrated in this Chapter focused on the HER4 somatic mutants.

Chapter 5

Multiscale Modeling of the Anti-cancer Role of the HER4 Tyrosine Kinase: Cellular Scale

In Chapter 4, we focused on the molecular mechanisms underlying the activation of the HER4 kinase dimer, and rationalized the effects of two clinically-identified HER4 somatic mutants from the perspective of our WT HER4 simulation. In Chapter 5, we translate our results into a cell signaling model representing the two divergent ('branching') networks induced by the WT and mutant HER4 kinase: the JAK2/STAT5a and PI3K/AKT pathways, respectively, in order to elucidate the parameters governing the selection of one branch versus the other. As the PI3K/AKT signaling branch has been extensively modeled, yet the HER4-mediated JAK/STAT branch has not, we undertake experimental assays of HER4 signaling in a murine mammary epithelial cell line, HC11, in order to provide parameters and more fully characterize the JAK/STAT branch of the HER4 signaling pathway.

5.1 Introduction

As discussed in Chapter 4, several studies have indicated an anti-cancer role for the HER4 kinase in various types of cancer, especially breast cancer, through steering of the mammary tumor cells toward a program of cell differentiation and away from a program of uncontrolled cell proliferation[48, 122, 124, 177]. Several somatic mutants recently discovered in HER4 abrogate this ability of HER4 to stimulate cell differentiation, and

instead bias signaling toward various proliferative pathways, such as PI3K/AKT[191]. Thus a better understanding of the molecular parameters governing the cell differentiation pathway induced by HER4 will allow for potential modulation of this pathway in the treatment of certain ErbB-driven cancers.

To quantitatively investigate this 'branched' signaling pathway (*i.e.,* PI3K/AKT/proliferation versus JAK2/STAT5a/differentiation) in HER4-stimulated mammary cells, we present a signaling model of WT versus mutant HER4 induction. As the PI3K/AKT signaling branch has been extensively studied and modeled, yet the HER4-induced JAK2/STAT5a branch has not, we undertake experiments in a HER4-stimulated mammary epithelial cell line, HC11, in order to more fully characterize the JAK/STAT branch of the HER4 signaling pathway. HC11 is a murine mammary epithelial cell line derived from the mammary gland of a female Balb/C mouse at mid-pregnancy, and has been extensively employed as an *ex vivo* model of mammary cell differentiation in response to various hormones and ligands. Although the mammary differentiation pathway has been investigated in response to various hormones and ligands, it has not been well-studied in response to HER4. Indeed, to date, only a few studies which examine the role of HER4 in STAT5-mediated mammary cell differentiation have been performed [122, 197]. In these studies, it was demonstrated that ligand-induced activation and phosphorylation of HER4 results in phosphorylation of STAT5a, an event which is mediated (either directly or indirectly) by JAK2. Thus HER4 performs several different functions in terms of activating the JAK2/STAT5 pathway: it is required for STAT5a phosphorylation (either directly or indirectly), and also to chaperone the activated STAT5a dimer into the nucleus[197]. Indeed, these studies revealed

that blocking either the kinase activity of HER4 or the ability of HER4 to cleave from the cell membrane abrogated STAT5a nuclear translocation and stimulation of differentiation genes[122].

These same studies highlighted a role for crosstalk between the HER4 pathway and other signaling cascades in the promotion of mammary cell differentiation. Specifically, the prolactin (PRL) pathway was shown to synergize with the HER4 pathway to produce a greater level of STAT5a activation and gene expression than that observed in either pathway independently[123]. Interestingly, knock-down of HER4 expression (using siRNA) or kinase activity resulted in a significant reduction of PRL-mediated STAT5a phosphorylation and gene expression. Indeed, it was shown that the PRL pathway critically depends upon HER4 for complete induction of STAT5a and mammary differentiation.

Given the crucial role of HER4 expression and kinase activity in full induction of the JAK2/STAT5a mammary differentiation pathway, we combine the results of our HER4 systems model and experiments to more fully characterize this signaling network, with the ultimate aim of motivating the modulation of this pathway for treatment of specific ErbB-driven cancers.

5.2 Materials and Methods

5.2.1 Systems model of WT versus mutant HER4 signaling pathways

The computational signaling model was based on modules from two published signaling models: the PI3K/AKT signaling branch was based on the model by Schoeberl et al. [154], and the JAK2/STAT5a signaling branch was based on the model by Yamada et al.[204]. Mass-action reactions describing ligand-induced ErbB receptor homo- and

heterodimerization, receptor internalization and degradation, constitutive dimerization, and activation were included. The PI3K/AKT model was modified to include HER4 dimers (the original model excludes HER4, as HER4 expression levels are low in the cancer cell lines on which the model is based). Hence all four ErbB RTKs are included in the model, and allowed to dimerize in response to stimulation with either the neuregulin-1β (NRG-1β) or HB-EGF ligands. The WT HER4 kinase was assumed to preferentially induce the JAK/STAT5 pathway via HER4 homodimers, whereas the mutant HER4 kinase was assumed to preferentially induce the PI3K/AKT pathway via HER4 heterodimers. Levels of phosphorylated ppAKT and pSTAT5a nuclear dimers were considered as read-outs of pathway activation for mutant and WT HER4 dimers, respectively. All simulations were performed in MATLAB 7.10 (MathWorks, Natick, MA).

5.2.2 HC11 cell differentiation assay

HC11 mammary epithelial cells, a kind gift from the Chodosh lab at UPenn, were grown and maintained at 5% CO2 in complete medium (RPMI 1640, 10% FBS, 2 mM L-glutamine, 100 U/mL penicillin, 100 μg/mL streptomycin, 1 μg/mL hydrocortisone, 10 ng/mL murine EGF and 5 μg/mL insulin). To induce differentiation, cells were seeded in 6-well plates at a density of 2 x 10^4 cells/cm^2 and allowed to grow to 100% confluence. The cells were maintained at confluence for 2 days in serum-free/EGF-free medium to induce competence. The competent cells were then stimulated for the indicated time points (1-3 d) with ovine prolactin (PRL, 5 μg/mL; Sigma, St. Louis, MO), EGF (10 ng/mL; BD Biosciences, San Jose, CA), NRG-β1 (20 ng/mL; Peprotech, Rocky Hill, NJ), HB-EGF (20 ng/mL; Peprotech, Rocky Hill, NJ), or a combination of PRL and HB-EGF

to induce differentiation. Differentiation was monitored by following the formation of blister-like structures or 'domes' that appear in confluent cultures and are thought to result from fluid secretion. Throughout this Chapter, we refer to 'Stage 1, Stage 2, and Stage 3' of the differentiation process: Stage 1 refers to growing the cells to confluence, Stage 2 refers to maintaining the cells at confluence for 48h to induce a state of competence, and Stage 3 refers to cell stimulation with HER4 ligand.

5.2.3 Reverse transcription-PCR

Total RNA from differentiated HC11 cells was prepared using TRIzol reagent (Invitrogen, Carlsbad, CA) following the manufacturer's instructions. Reverse transcription was performed using Superscript II reverse transcriptase and oligo(dT) primers (Invitrogen). PCR was performed using the following gene-specific primers: mouse β-casein (forward 5' ACT GTA TCC TCT GAG ACT G and reverse 5' TCT AGG TAC TGC AGA AGG TC, producing a 578-bp amplicon), mouse whey acidic protein (WAP, forward 5' TGC CTC ATC AGC CTT GTT CTT G and reverse 5' CAG CTT TCG GAA CAC CAA TGT TG, producing a 235-bp amplicon), and mouse keratin 18 as a housekeeping gene (forward 5' CAT CGT CTT GCA GAT CGA CA and reverse 5' GCT GAG ACC AGT ACT TGT CCA G, producing a 376-bp amplicon). RT-PCR products were resolved on 2% agarose gels.

5.2.4 Transcription factor activation assays

Nuclear extracts from differentiated HC11 cells were isolated using the Nuclear Extraction Kit from Active Motif (Carlsbad, CA). The nuclear extract was then used to profile the activation of the STAT5a TF using the TransAM STAT family ELISA-based

system from Active Motif (Carlsbad, CA). The TransAM system contains a 96-well plate on which has been immobilized oligonucleotide containing the STAT consensus binding site, 5'-TTCCCGGAA-3'. The active form of STAT contained in the nuclear extract specifically binds to the oligonucleotide sequence, and primary antibodies are used to detect specific subtypes of STAT, in our case, STAT5a. Subsequent incubation with an HRP-conjugated secondary antibody provides a colorimetric readout of the relative amount of activated STAT5a in the sample.

The activation of additional transcription factors, including glucocorticoid receptor (GR) and C/EBP, were also tested. The consensus binding site for GR is 5'-GGTACAnnnTGTTCT-3', and for C/EBP, 5'-GCAAT-3'.

5.2.5 Quantitative real-time RT-PCR

The cDNA templates were synthesized from total RNA using SuperScript II reverse transcriptase and oligo(dT) primer (Invitrogen, Carlsbad, CA). Quantitative PCR was performed in a 96-well micro titer plate format on an Applied Biosystems 7300 real-time PCR system (Applied Biosystems, Foster City, CA) using TaqMan Universal PCR master mix and TaqMan Gene Expression Assays (Applied Biosystems) specific for the genes of interest. Each experiment was repeated using three independent sources of RNA and fold change relative to control was calculated and normalized to hPRT (endogenous control) RNA levels. The PCR program used was: 2 min at 50 °C, 10 min at 95 °C, and 40 cycles of 15 sec at 95 °C, 1 min at 60 °C.

5.3 Results

5.3.1 Branched signaling model of WT versus mutant HER4 pathways

The HER4 signaling model comprises two branches: PI3K/AKT, which is stimulated by ErbB heterodimers and mutant HER4, and JAK2/STAT5a, which is induced by WT HER4 homodimers. The PI3K/AKT pathway has been previously well-characterized and modeled, whereas the HER4-JAK-STAT pathway has not (to our knowledge) been modeled, and thus we delineate the major events here. Upon binding to NRG, HER4 (which can basally associate with JAK2, as has been shown experimentally[123]) homodimerizes and becomes phosphorylated. The HER4-JAK2 complex then binds and activates STAT5a, which becomes phosphorylated and dimerizes. We incorporate JAK2 into the model in this way to reflect that it is required for the HER4-mediated activation of STAT5a, though the mechanism is currently unknown. The HER4-STAT5a dimer then cleaves from the membrane and translocates to the nucleus; this event is built into the model to reflect that STAT5a is unable to accomplish nuclear translocation without HER4 as a chaperone[197]. STAT5a and HER4 dissociate in the nucleus, where STAT5a induces expression of various genes. The only gene expression event which is explicitly represented in our model is the expression of SOCS, which, upon translation into SOCS protein, inhibits the JAK-STAT phosphorylation event in the cytoplasm. We note that there are a total of four negative regulators present in the model, including SOCS, SHP-2, which dephosphorylates the HER4-JAK2 complex, PPX, which represents the cytoplasmic STAT5a phosphatase, and PPN, which represents the nuclear STAT5 phosphatase[204].

Figure 5.1 displays time course plots for several model species, including phosphorylated HER4-JAK2 dimer ('pHER4-JAK2-dimer'), cytoplasmic STAT5

('STATc'), phosphorylated nuclear STAT5a dimers ('pStatn-pStatn'), and cytoplasmic SOCS mRNA, in response to stimulation with 5 nM NRG. It is apparent that levels of pHER4-JAK2 peak at approximately 30 minutes post-ligand stimulation, and levels of phosphorylated nuclear STAT5a dimer peak at approximately 1 hour post-stimulation. Due to the time lag during which STAT5a translocates to the nucleus and induces expression of SOCS mRNA, levels of cytoplasmic mRNA do not peak until approximately 2 hours post-induction.

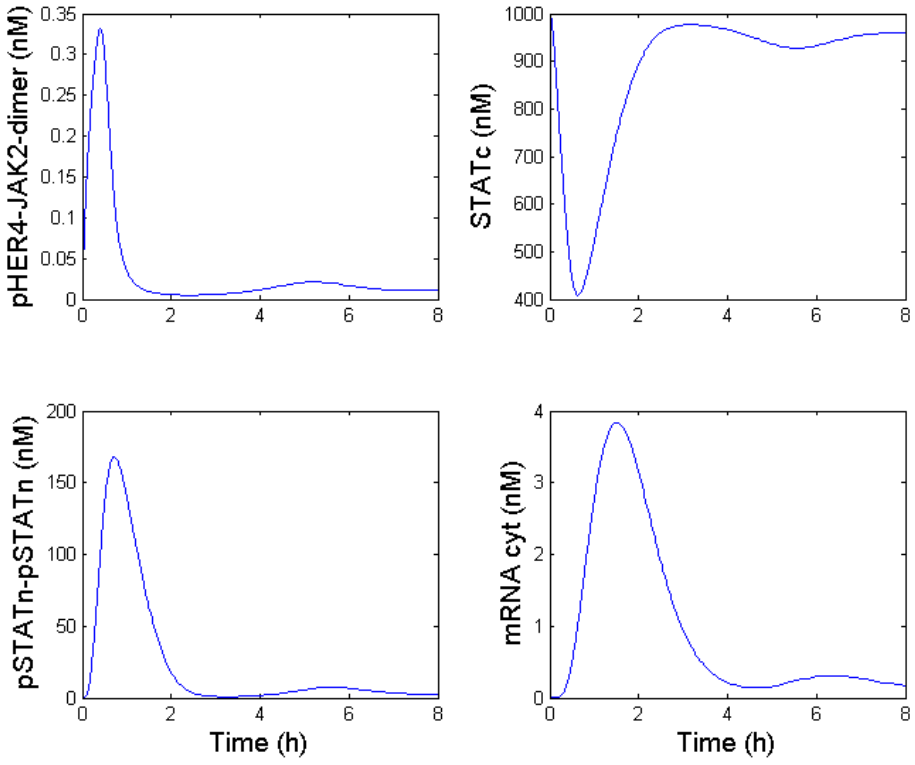

Figure 5.1. Time courses for key species in the HER4-JAK-STAT model, 5 nM NRG.

In response to an *in silico* cell expressing WT HER4, the activation of HER4 homodimers and the JAK2/STAT5a signaling branch is increased, whereas the response of the other ErbB dimers (and subsequent PI3K-AKT signaling) is significantly dampened. Thus the JAK2/STAT5a cascade becomes the predominant signaling

pathway. In response to 'expression' of mutant HER4, we tune the model parameters such that signaling through HER4 homodimers is essentially turned off, and PI3K/AKT is the predominant pathway induced by activated ErbB dimers (Fig. 5.2).

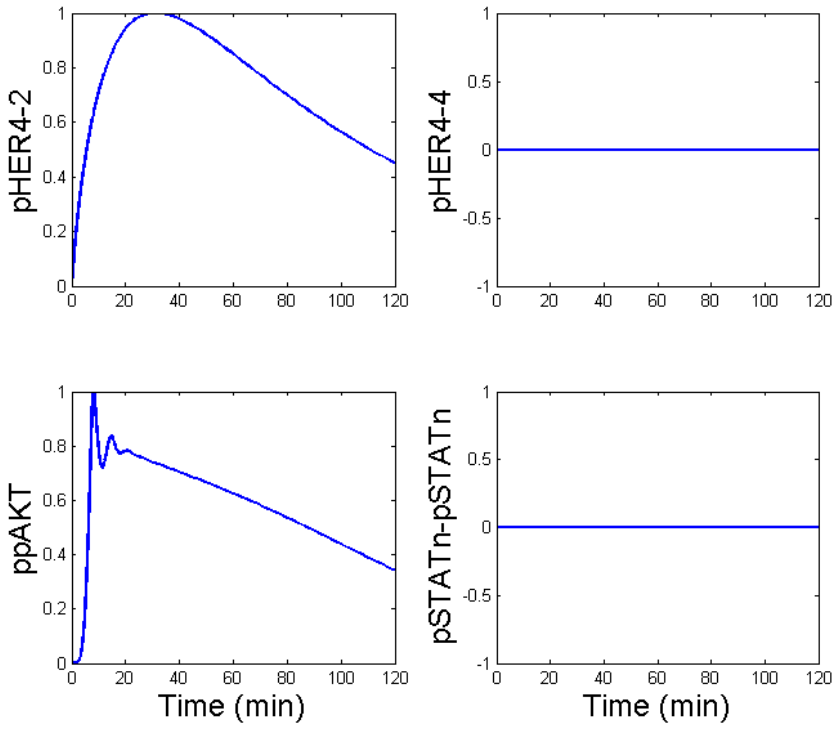

Figure 5.2. Time course plots for key species in the branched model for mutant HER4.

We find that our model recapitulates the major features of PI3K/AKT and JAK/STAT signaling dynamics, including the importance of the STAT5 nuclear phosphatase in governing both peak and steady state levels of nuclear pSTAT dimer. However, our model is also the first (to our knowledge) to explicitly incorporate the HER4-JAK-STAT signaling branch, which includes cleavage of the HER4 kinase domain from the membrane and translocation of the HER4-STAT5a complex to the nucleus. We note that additional rate constants, such as the rate of protease-induced cleavage of HER4, are not

required for the current version of our model, but will be explicitly incorporated as they become experimentally available.

5.3.2 Optimization of HER4-mediated HC11 cell differentiation assay

To more fully characterize the signaling dynamics in the HER4-mediated JAK/STAT pathway, we performed RT-PCR in differentiating HC11 cells, a murine mammary epithelial cell line (see Methods). Specifically, we assayed for the following read-outs of HC11 mammary differentiation: levels of activated STAT5a and glucocorticoid receptor (GR) as well as expression of the β-casein milk gene. STAT5a and GR are two key TFs which are both activated during HC11 cell differentiation, and synergize on the β-casein gene promoter.

Although transcription of the differentiation marker genes has not been extensively investigated in HER4-stimulated HC11 cells, several studies have reported expression of β-casein 24 to 72 hours after ligand stimulation, yet no detectable levels of β-casein mRNA at earlier time points [123, 122]. Indeed, we could not detect expression of β-casein at early time points (Figure 5.3) in response to the ligand NRG or HB-EGF. Interestingly, we discovered that β-casein expression could be induced at earlier time points only if FBS was kept in the cell medium during the differentiation process. Whereas we had been serum-starving the cells initially, we increased the concentration of FBS to 10% and observed β-casein production at several hours post-stimulation (Fig. 5.3). However, this gene expression is most likely due to other factors contained within the FBS, and thus we omitted FBS from our differentiation protocol. In addition, we noted that, although NRG and HB-EGF were unable to effect β-casein expression at early

time points, prolactin (PRL) induced robust gene expression at all time points tested, regardless of FBS concentration. Thus, despite sharing certain elements of signaling in mammary cells, the PRL and HER4 pathways utilize distinct mechanisms to a certain extent. We also tested a range of ligand concentrations and time points, which are discussed in the following sections, in order to optimize the differentiation assay.

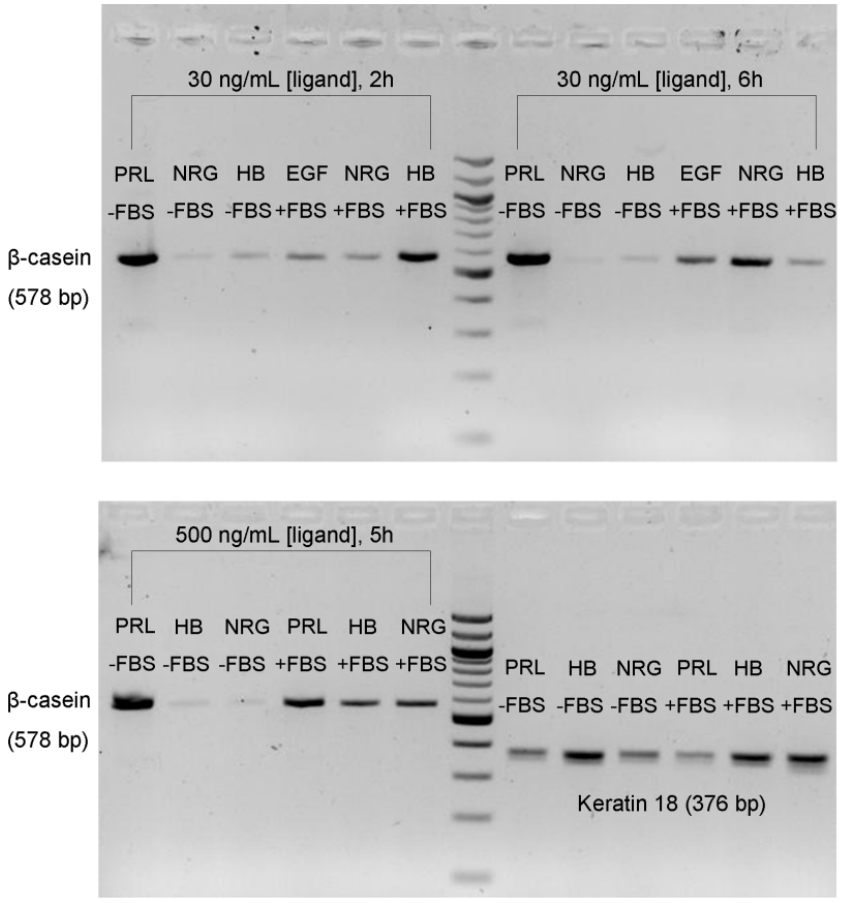

Figure 5.3. RT-PCR of β-casein gene expression in HC11 cells stimulated with NRG, HB-EGF, or PRL, in the presence or absence of FBS.
Several time points were assayed (2h, 5h and 6h are shown here) as well as several different ligand concentrations. Keratin 18 was employed as a housekeeping gene.

5.3.3 Glucocorticoid receptor is required for HER4-mediated signaling via JAK2/STAT5a

HER4 stimulation of the JAK2/STAT5a pathway has only recently been discovered; as a result, much remains to be elucidated regarding the dynamics of this pathway and interactions with other signaling cascades. Recently, Muraoka-Cook et al.[123] highlighted the synergy between the HER4 and prolactin (PRL) pathways. As summarized in Section 5.1, PRL-induced differentiation of mammary epithelial cells, in particular, HC11 cells, has been studied extensively[87, 88]. It is postulated that, *in vivo*, interactions among the PRL, insulin receptor, and HER4 pathways occur in a defined sequence during differentiation of the mammary epithelium throughout pregnancy and lactation. In a recent study in HC11 cells, Muraoka-Cook et al.[123] revealed that the PRL and HER4 signaling cascades converge upon JAK2 to effect mammary cell differentiation. Indeed, the group determined that in cells over-expressing kinase-dead HER4 or siRNA targeted to HER4, PRL-induced differentiation was markedly diminished. They also found that PRL induced phosphorylation of HER4 via JAK2, and that HER4 and PRL formed a protein complex in stimulated cells. In addition, it was observed that stimulation of both PRL and HER4 resulted in greater levels of activated STAT5a and β-casein expression than either ligand alone, highlighting the synergy between the two pathways. Besides PRL, it is currently unknown which other signaling cascades may be important in HER4-induced mammary cell differentiation.

In our experimental studies, we included the steroid hydrocortisone (HC), which signals through the glucocorticoid receptor (GR), in our HC11 differentiation assay, as the few previous studies of HER4-mediated differentiation have included HC in the cell growth medium (yet have not commented on its importance). However, we wondered to what extent HC plays a role in the HER4 differentiation process. In cells stimulated with

the HER4 ligand NRG-1β but not HC, no β-casein was expressed (Fig. 5.5). We also tested the β-casein response upon stimulation with another HER4 ligand, HB-EGF, and discovered the same result. However, upon addition of HC to the differentiation medium, HER4 induced robust expression of β-casein (Fig. 5.4). HC was required in Stage 2 of the differentiation process (see Methods) as well as Stage 3. Interestingly, HC was not needed for PRL-mediated expression of β-casein, although it did enhance PRL signaling. We therefore conclude that HC (and hence, GR), is required for HER4-mediated differentiation of HC11 mammary cells: in particular, it is necessary for induction of cell competence (Stage 2) as well as the early steps of differentiation (Stage 3).

Our results agree with previous ChIP studies which demonstrated the synergy between PRL and GR[87, 88]. In these studies, HC11 cells were stimulated with PRL, and various nuclear events were examined. In particular, the dynamics of assembly of various transcription factors (STAT5, GR) on the β-casein promoter were analyzed. It was discovered that stimulation with both HC and PRL was required for recruitment of the RNA polymerase II to the β-casein promoter and for assembly of the complete transcriptional activation complex, which included STAT5a (induced by PRL) and GR (induced by HC). Our results suggest that HER4 operates via a mechanism similar to PRL, in terms of synergizing with GR to effect maximal expression of β-casein. Indeed, we find that HER4 cannot induce differentiation of HC11 cells on its own.

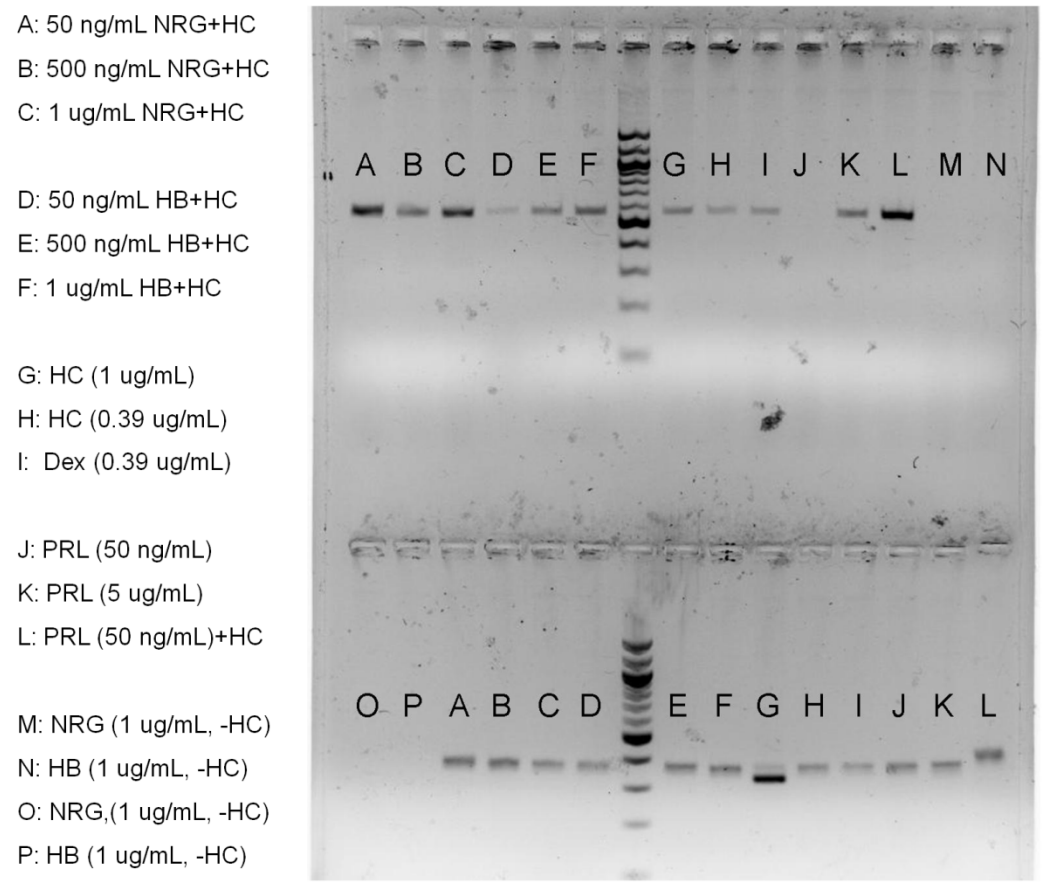

Figure 5.4. RT-PCR of β-casein gene expression in HC11 cells at 48h.
NRG and HB-EGF require HC in order to stimulate gene expression, whereas only low concentrations of PRL require HC. HC alone can stimulate β-casein gene expression at sufficiently high concentrations. NRG stimulates a higher level of β-casein gene expression than does HB-EGF. The bottom row (samples A-L) represent the housekeeping gene, keratin 18.

5.3.4 Different ligands signaling through the same receptor trigger divergent cellular outcomes

In the few previous studies of HER4-induced differentiation of mammary cells, two different HER4 ligands, NRG and HB-EGF, were employed. Although HB-EGF can also bind EGFR, the previous studies had shown that in HC11 mammary cells, very little EGFR is phosphorylated in response to HB-EGF; indeed, the predominant phosphorylated ErbB receptor is HER4 [122]. With this in mind, we tested the effects of

HB-EGF upon induction of STAT5a and β-casein expression, in order to compare to the results obtained for NRG. Interestingly, we found that addition of HB-EGF resulted in the opposite effect as stimulation with NRG: rather than stimulate β-casein expression, HB-EGF inhibited β-casein (Figure 5.5). At 48h post-ligand stimulation, concentrations as low as 10 ng/mL HB-EGF inhibited levels of β-casein. Higher concentrations of HB-EGF, ranging from 50-500 ng/mL, resulted in a fold change in β-casein expression of 0.3-0.5. This result suggests that there is a basal level of β-casein expression in response to HC alone (which is in the medium at Stage 2 of the differentiation process), and HB-EGF suppresses the differentiation process, perhaps by biasing the cell toward an alternative signaling pathway.

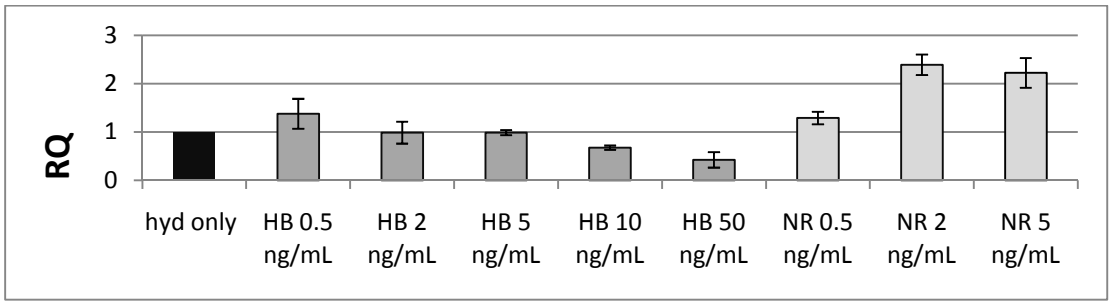

Figure 5.5. qRT-PCR of β-casein expression levels for increasing ligand levels.
Error bars are reported as the result of 3 replicate samples, each run in duplicate.

By contrast, at 48h post-ligand stimulation, increasing concentrations of NRG induce greater β-casein expression levels (Fig. 5.6). NRG concentrations ranging from 0.5-50 ng/mL result in a 1.3-2.5-fold increase in β-casein transcripts. Our result, which demonstrates that NRG stimulates β-casein expression whereas HB-EGF inhibits it, is in agreement with earlier work by Amin et al.[3, 4]. Amin et al. stimulated mammary

carcinoma cells with a variety of ErbB ligands, and showed that two different ligands signaling through the same ErbB receptor could trigger different signaling cascades and patterns of gene expression, and hence, different cellular decisions. In our study, we cannot be certain that HB-EGF and NRG are solely stimulating HER4, as HB-EGF also binds EGFR and NRG also binds HER3; however, based on earlier studies which showed that HER4 is the primary ErbB receptor that is phosphorylated in response to the HC11 differentiation program[123, 122], we can be sure that HER4 plays a dominant role in response to these ligands in the context of HC11 cell differentiation. Thus, two mechanisms by which NRG and HB-EGF may exert their opposing effects are: stimulation of different ErbB heterodimer combinations (unlikely to be the primary mechanism, based on earlier work), and recruitment of different signaling mediators via the same receptor (HER4), potentially through specific allosteric changes upon receptor-ligand binding. Thus it is possible that the same receptor, through binding to different ligands, can trigger divergent cellular outcomes (in this case, differentiation versus inhibition of differentiation).

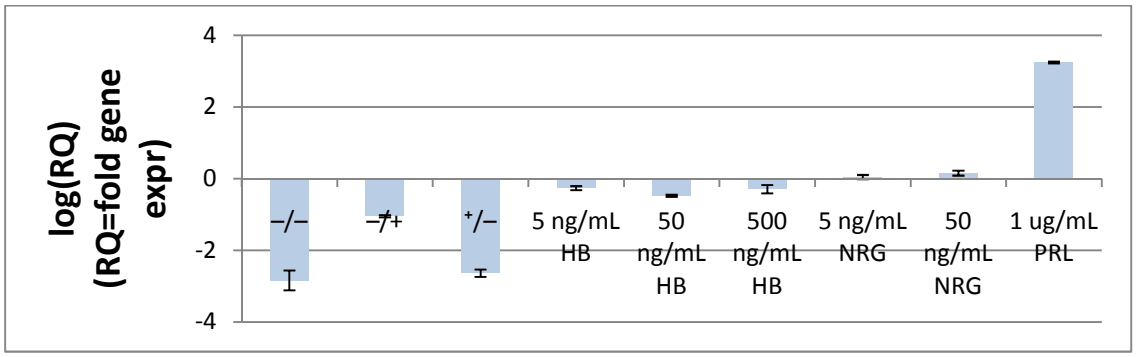

Figure 5.6. qRT-PCR results for β-casein in response to HB-EGF, NRG, PRL.
The first three samples (labeled -/-, -/+, and +/-, in reference to differentiation Stage 2/Stage 3) refer to the presence (+) or absence (-) of HC in Stage 2 and Stage 3 differentiation medium.

5.3.5 Interaction between PRL and HER4 signaling pathways

Based on the previous experimental studies, which suggested a synergistic interaction between the PRL and HER4 signaling pathways in the differentiation of mammary epithelial cells, we experimented with different ligand combinations. In response to stimulation with both PRL and HB-EGF, we found that levels of β-casein decreased (Figure 5.7), and this effect was amplified at lower PRL concentrations. Specifically, at a constant PRL concentration of 10 ng/mL, increasing levels of HB-EGF decreased the β-casein signal; at 10 ng/mL PRL/10 ng/mL HB-EGF, the expression level was 10-fold lower than that induced by 10 ng/mL PRL alone (Fig. 5.7). At 50 ng/mL PRL/10 ng/mL HB-EGF, the expression level was 3.4-fold lower than that of PRL alone, and at 500 ng/mL PRL/10 ng/mL HB-EGF, the fold β-casein expression was halved, relative to 500 ng/mL PRL alone. Thus, the addition of HB-EGF has an inhibitory effect on PRL-induced cell differentiation, although this effect is dampened at higher PRL concentrations.

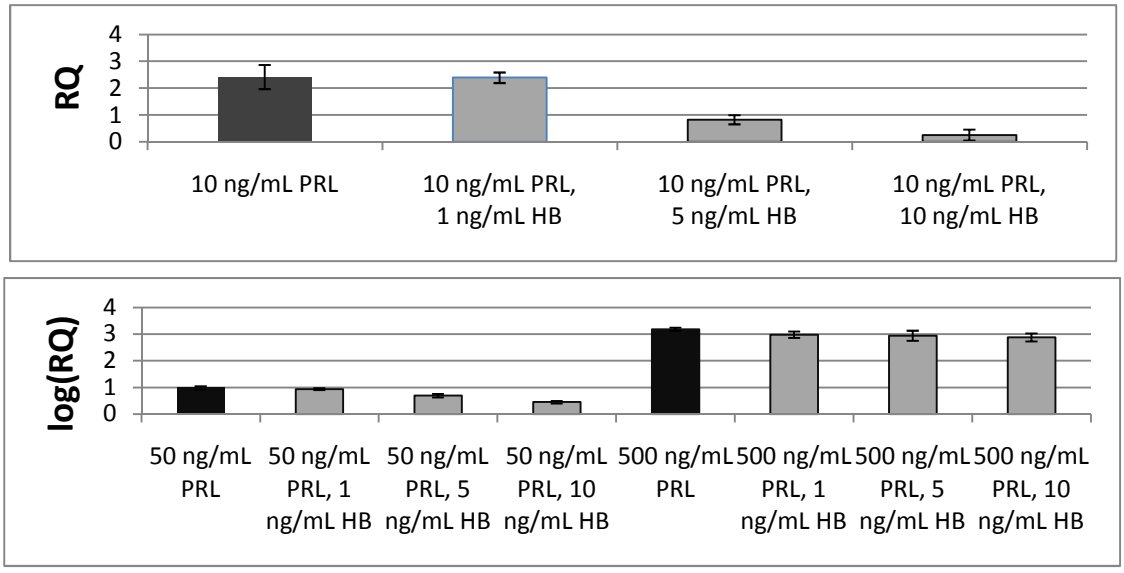

Figure 5.7. qRT-PCR results for β-casein expression in response to PRL and HB-EGF.
Top panel shows a constant PRL concentration of 10 ng/mL; bottom panel shows constant PRL concentrations of 50 and 500 ng/mL.

Interestingly, we obtained a similar result for stimulation with both PRL and NRG (Figure 5.8). Addition of 50 ng/mL PRL/10 ng/mL NRG resulted in a 3.2-fold decrease in β-casein levels, relative to 50 ng/mL PRL alone, and a combination of 500 ng/mL PRL/10 ng/mL NRG resulted in a 1.4-fold decrease in β-casein expression. Although the inhibitory effect of NRG on PRL-induced β-casein expression was not as pronounced as the effect of HB-EGF, the inhibition was statistically significant. Thus our results suggest that HER4 synergizes with the GR pathway yet antagonizes the PRL pathway, perhaps due to a specific subpopulation of HC11 cells. Our data helps to parse the contributions of these intersecting signaling cascades (HER4, PRL, GR), with the ultimate goal of incorporating these interactions into a more complete model of HER4-induced mammary cell differentiation.

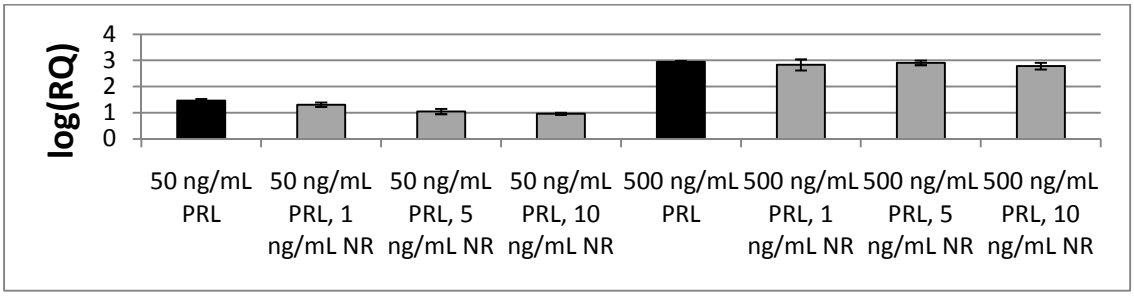

Figure 5.8. qRT-PCR for β-casein in response to a combination of PRL, NRG.
PRL is kept at a constant concentration of either 50 or 500 ng/mL. The addition of NRG inhibits PRL-induced β-casein expression levels, although this effect is less pronounced than for HB-EGF/PRL.

5.3.6 Temporal switch between two cellular decisions

In tracking the response of the HC11 cells to HER4 stimulation at various time points, we came upon an intriguing observation. At early time points post-stimulation, NRG does not exhibit any effect on β-casein expression, as we previously discussed. However, at 12h post-stimulation, we found that NRG actually decreased levels of β-casein, relative to

the HC control (Fig. 5.9). At 24-48h post-stimulation, NRG began to increase β-casein expression levels at sufficiently high ligand concentrations (Fig. 5.9). We reason that this result has not been discovered previously, as the few studies examining the HER4-STAT5a pathway focused on gene expression 24-48h post-stimulation. The mechanism by which this 'switch' occurs is an area of future investigation, and may be due to competition between different signaling pathways during the differentiation process. Currently, our HER4 signaling model truncates at the point of activation and nuclear translocation of STAT5a (or AKT). However, future iterations of the model will incorporate this gene expression switch, which occurs on much longer time scales than are currently represented in our model.

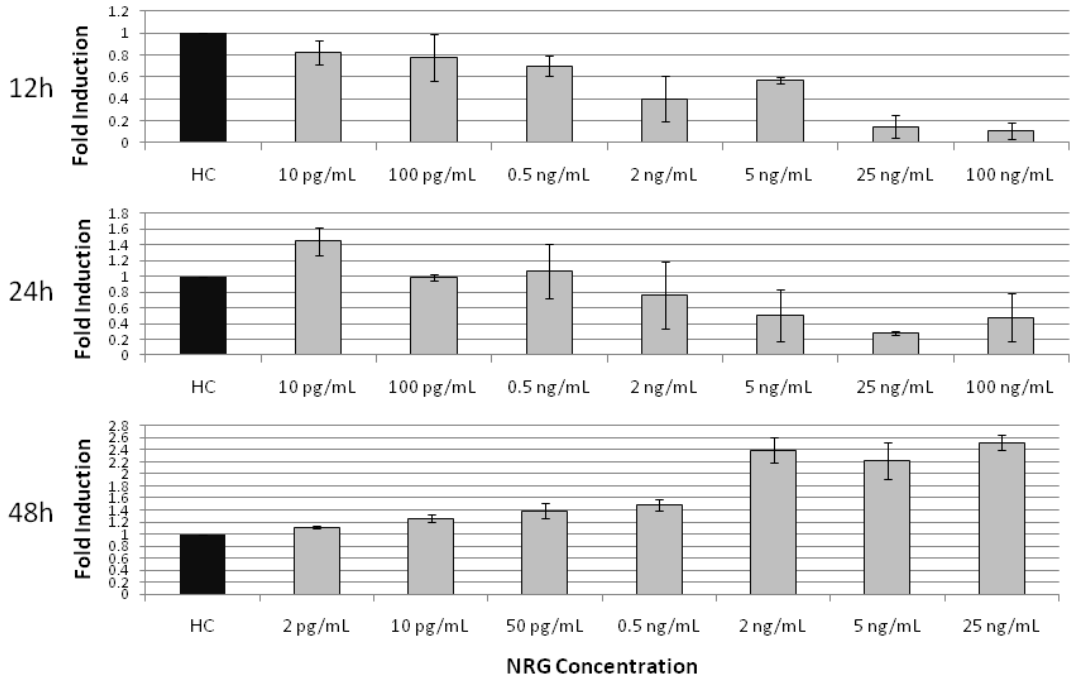

Figure 5.9. qRT-PCR time-course for NRG stimulation of HC11 cells.

The time course of β-casein gene expression correlates well with transcription factor activation dynamics for STAT5a and GR. In our ELISA-based TF activation assay (see Methods), we detected activity of both STAT5a and GR at various time points post-stimulation (Fig. 5.10). Figure 5.10A shows that, although levels of activated nuclear STAT5a are highest 15-30 minutes post-stimulation, STAT5a activity persists until 24h post-ligand stimulation. Similarly, the TF activation profile for GR (Fig. 5.10B) shows that levels of activated nuclear GR are highest at early time points, but persist until 24h post-stimulation. Our results agree with those of Kabotyanski et al. [87, 88], who assayed for binding of various TFs (including STAT5a and GR) to the β-casein gene promoter at various time points following stimulation with PRL, and found that, although several of the TFs assemble on the promoter at early time points, the RNA polymerase does not bind and commence transcription until 24h post-stimulation. Thus, additional nuclear events are required for assembly of the complete transcriptional apparatus and initiation of β-casein expression.

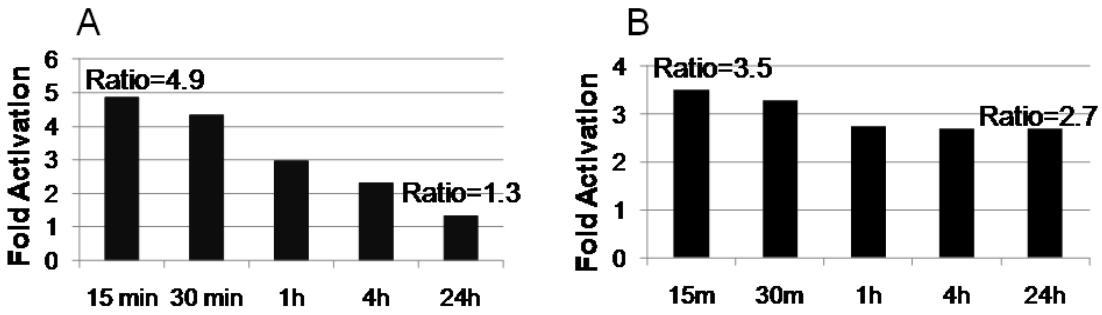

Figure 5.10. TF activation assays for (A) STAT5a and (B) GR.
The assays measure relative levels of activated TF in the nucleus, upon ligand stimulation. The ratio of levels of activated STAT5a in a ligand-stimulated cell versus an unstimulated cell is recorded. For example, in panel (A), at *t*=15 min, there is 4.9 times as much activated STAT5a in cells stimulated with ligand relative to unstimulated cells. Both the STAT5a and the GR signals persist until 24h post stimulation.

5.4 Discussion

As HER4 may represent the only member of the ErbB family of RTKs to play an anti-cancer role in certain tumor contexts, it is crucial to quantitatively evaluate the signaling dynamics of the HER4 pathway in the mammary cell environment. Here we have translated the results from our structural studies of the WT HER4 kinase, which corroborated the hypothesis that the HER4 somatic mutants abrogate the ability of HER4 to activate but not dimerize, into a cell signaling model of HER4 activity, which branches into two different pathways: the JAK2/STAT5a network (induced by WT HER4) and the PI3K/AKT cascade (induced by mutant HER4). We find that HER4 homodimers predominantly induce the JAK/STAT5 pathway, although in an *in vivo* context, it is likely that a combination of hormones and ligands (including PRL and HC) contributes to the preferential stimulation of HER4 homodimers rather than other ErbB dimer combinations, in response to ligands such as NRG (which can also bind to HER3). Our model is able to recapitulate the major features of the PI3K/AKT and JAK/STAT pathways, yet is the first, to our knowledge, to explicitly incorporate the HER4-JAK-STAT axis as well as the unique events occurring in the HER4-STAT5a pathway, including cleavage of the activated HER4 kinase from the membrane and translocation into the nucleus to effect gene transcription.

As the HER4-JAK-STAT pathway has not been well-characterized, we undertook experimental analysis of a model mammary epithelial cell line, HC11, in response to HER4 stimulation, to quantitatively assess several key aspects of the HER4-JAK-STAT cascade. We found that several other signaling cascades are either necessary for or synergize with the HER4 pathway; these cascades include the glucocorticoid receptor and PRL pathways. Although the HC11 cells mimic many aspects of mammary gland

differentiation (which occurs during the later stages of pregnancy), in an *in vivo* context, it is likely that there is a defined sequence of events, in terms of the interactions among HER4, PRL, and GR, which culminates in the differentiated mammary cell phenotype. Our finding that the NRG and HB-EGF ligands decreased PRL-induced gene expression could potentially be attributed to the presence of different HC11 sub-populations, in comparison to previous studies[123]. We also discovered a temporal 'switch' whereby the NRG ligand initially inhibits the expression of the β-casein milk gene (at 12h post-stimulation), followed by a switch to activation of the β-casein gene at 24-48h post-ligand stimulation. This event has not previously been noted, as there are few studies of the HER4 differentiation pathway to begin with, and additionally, the previous studies concentrate on gene expression at 24-48h. Although the molecular mechanism underlying this switch has yet to be determined, we reason that it may constitute a cellular shift from a program of proliferation to a program of differentiation: precisely the same shift that can potentially be modulated to bias tumor cells toward the HER4 differentiation pathway for therapeutic purposes. Future studies will examine the activation kinetics of other molecular species, such as the TF C/EBP-β, which is known to be involved in mammary cell differentiation, to determine whether they might play a role in mediating this cellular switch.

Future iterations of our branched HER4 signaling model will incorporate these additional layers of complexity revealed by our experimental analysis, including the crosstalk among the HER4, GR, and PRL pathways, as well as the cellular 'switch' from inhibition to activation of gene expression involved in cell differentiation. However, we currently note that it is crucial to further investigate the interactions among the HER4,

PRL, and GR pathways if HER4-mediated cancer therapies are eventually to be considered in the clinic. We also emphasize the importance of considering the complementary interactions among the ErbB kinases, especially in terms of rationalizing ErbB mutation mechanisms. In the case of the HER4 somatic mutants, the mutations effectively alter the cell fate by rendering HER4 dependent upon the other ErbB kinases (and no longer able to signal as a HER4 homodimer). Indeed, the various activities of the ErbB kinases complement each other to produce a holistically functioning signaling unit, thus compensating for any deficiencies in the individual ErbB proteins. In certain tumor cells predisposed to uncontrolled signaling, this otherwise advantageous evolutionary characteristic of the ErbB system can have deleterious consequences, and must be considered when designing therapeutics targeted to the ErbB family.

Chapter 6

Conclusions and Future Directions

6.1 Summary of results

In this thesis, we have applied a multiscale modeling framework to investigate the molecular regulatory mechanisms governing the activation of the ErbB RTKs (with a focus on the HER3 and HER4 kinases), which play a key role in crucial celluar processes as well as in various types of human cancer. Chapter 2 presented a multiscale model of activity in the HER3 RTK, a kinase which, until recently, has been considered kinase-dead. We applied our modeling scheme to investigate the non-canonical catalytic mechanism employed by HER3 and the physiological relevance of this activity to mechanisms of drug resistance in an ErbB-driven tumor cell *in silico*. The results of our molecular-scale simulations supported the characterization of HER3 as a weakly active kinase that, in contrast to its fully-active ErbB family members, depends upon a unique hydrophobic interface to coordinate the alignment of specific catalytic residues required for its activity. Translating our molecular simulation results of the uniquely active behavior of the HER3 kinase into a physiologically relevant environment, our HER3 signaling model demonstrated that even a weak level of HER3 activity may be sufficient to induce AKT signaling and TKI resistance in the context of an ErbB signaling–dependent tumor cell, and therefore therapeutic targeting of HER3 may represent a superior treatment strategy for specific ErbB-driven cancers.

Chapter 3 extended our analysis of HER3 to the other members of the ErbB family (EGFR, HER2, HER4) to facilitate a comparison of the activation and regulatory mechanisms across the ErbB family members. We biased certain sections of Chapter 3 toward the HER2 kinase, as HER2 exhibits several structural and functional features which are unique among the ErbB RTKs. Our analysis identified several commonalities in hydrogen bonding patterns among the three ErbB kinases (EGFR, HER2, HER4) in the A-loop, C-loop, and αC helix. We extended our analysis of global kinase motions and hydrogen bonding patterns to several different ErbB dimer systems, including a HER2-EGFR heterodimer, as well as EGFR and HER4 homodimers, and found that disruption of the hydrogen bonding pattern in the dimers shifted the systems toward the active state. Furthermore, we identified several interactions that persist throughout the dimeric trajectories of the inactive states and are hence candidates for further investigation through free energy methods such as umbrella sampling simulations, as the molecular environment surrounding such residues may define the pathway for conformational change and the associated barriers to activation.

Chapters 4 and 5 presented a multiscale model of activity in the HER4 RTK, which has been postulated to play an anti-cancer role in certain tumor types. Our model begins with molecular simulations of the HER4 dimer activation mechanism and extends to the cellular scale through a signaling model of the HER4 differentiation pathway in mammary cells. Chapter 4 focused on molecular scale simulations of dimer-mediated activation in HER4, in order to elucidate the mechanisms by which the WT kinase is regulated and activated, and to rationalize the effects of several HER4 somatic mutants which have recently been discovered in a subset of cancer patients. The simulations

highlighted key molecular interactions which we predict are important in the HER4 activation mechanism. We discovered that the order of events, in terms of motions of the individual catalytic sub-domains, differed from those of other kinase families, such as Src, a distinction which may partially be attributed to the fact that phosphorylation of the A-loop tyrosine residue does not seem to be required for the activation of the ErbB kinases. In addition to identifying global motions and hydrogen bonding events taking place within the activated kinase, we also identified several inter-dimer bonds (*i.e.,* bonds between the receiver and activator kinases) which broke during the TMD simulation. Moreover, the bonding pattern in the JM region of the receiver kinase shifted during the course of the TMD simulation, a result which is in agreement with previous experimental studies demonstrating that the JM domain is a crucial contributor to the kinase allosteric activation mechanism. These inter-dimer and JM domain bonds have not been explored in previous simulations of the EGFR kinase, yet our results indicate that they warrant further investigation, in order to piece together the complete activation pathway of the ErbB kinases.

Chapter 5 translated our molecular scale results of HER4 kinase activity into a cellular pathway model of WT versus mutant HER4 signaling, in which the WT and mutant cells trigger divergent signaling networks. In particular, induction of the JAK2/STAT5a signaling pathway by WT HER4 in mammary epithelial cells culminates in differentiation of the cells, whereas the HER4 somatic mutants preferentially signal through PI3K/AKT, resulting in cell proliferation. The mechanism by which this cell fate decision is made involves the ability of the HER4 mutants to dimerize with other ErbB kinases (hence stimulating PI3K/AKT) but not activate (hence abrogating signaling

through STAT5a). Thus the modulation of this cell fate decision may represent a potential therapeutic strategy for specific ErbB-driven cancers. To provide insight into the HER4-induced JAK2/STAT5a differentiation pathway in mammary cells, we performed experiments in HER4-stimulated mammary epithelial cells, and discovered several system properties which may be incorporated into our multiscale model of HER4 activity. Specifically, we highlighted the role of crosstalk between the HER4 pathway and other signaling cascades, including the glucocorticoid receptor and prolactin pathways, in promoting the HER4-induced differentiation signal, emphasizing that these synergistic networks must be taken into account when considering potential therapeutic modulation of the HER4 pathway. We also identified a temporal 'switch' in the expression of cell differentiation genes, in which HER4 ligand initially inhibited expression of these mammary cell differentiation genes but subsequently increased their expression levels. This result may represent a particular stage in the proliferation-to-differentiation switch of mammary epithelial cells, and may be exploited in future studies investigating HER4-mediated cancer therapeutics.

The protein kinase genes are among the most frequently mutated in human cancers, and several ErbB kinase domain mutants have been determined. Our simulations provide insight into the effect of these mutations in the ErbB kinases through assessment of structural dynamics and hydrogen bonding patterns in the kinase domains. Furthermore, we predict the importance of previously overlooked regions in regulation of kinase activation. Elucidation of the molecular regulatory mechanisms will help establish structure-function relationships in the wildtype ErbB kinases as well as predict mutations with propensity for constitutive activation. Such molecular variants in the EGFR, HER2,

and HER4 RTKs are known to profoundly impact specific therapies targeting ErbB-mediated cancers. It is worthwhile to note that mutation predictions made in mouse models of cancer can be extended to human studies insofar as the receptor crystal structures are homologous in mouse and human (which, for the ErbB kinases, they are).

Our multiscale models of ErbB kinase activity, which investigate the molecular regulatory mechanisms in the ErbB kinases, will help to elucidate structure-function relationships in drug-resistant cell lines and motivate the development of more efficacious TKI therapies targeting ErbB-mediated cancers. A clinically valuable feature of our multiscale modeling approach is that the flow of information between models occurs in a bidirectional manner, so that it is possible to apply the results obtained from one scale to the generation of new hypotheses at another scale, and vice versa. We have demonstrated that modeling a biological system at multiple levels of resolution constitutes another type of model 'validation', which should be appreciated in conjunction with experimental validation. Multiscale modeling provides a powerful and quantitative platform for investigating the complexity inherent in intracellular signaling pathways and rationalizing the effects of molecular perturbations on downstream signaling events and ultimately, on the cell phenotype.

6.2 Extensions and future work

6.2.1 Future work: Construction and simulation of macromolecular complexes

The work presented in this thesis focuses on the structure and dynamics of the monomeric (or dimeric) forms of the ErbB tyrosine kinases. In order to extend our

models to downstream signaling events, the formation of multi-protein complexes must be considered. For example, upon ligand-induced activation and phosphorylation of the HER4 kinase, various adaptor proteins, such as STAT5a, are recruited to the kinase domain. Precisely how these proteins interact with HER4 remains an open question. However, structural modeling can be utilized to investigate and predict likely conformations for these macromolecular complexes. This will be an important step in extending our multiscale model of ErbB activity, as it will promote the investigation of the effects of molecular perturbations not only on the ErbB kinases, but on the multi-protein complexes formed by the ErbB kinases and their downstream signaling mediators.

We applied our modeling method to construct a protein complex consisting of the HER4 kinase and the STAT5a transcription factor. As discussed in Chapters 4 and 5, one of the first events to occur following HER4 activation is recruitment of STAT5a. Several studies have revealed a physical association between HER4 and STAT5a [32, 197], and have also demonstrated that HER4 serves as a chaperone to shuttle STAT5a into the nucleus. Indeed, STAT5a is incapable of nuclear translocation in the absence of HER4[197]. However, the physical nature of the HER4-STAT5a association is unknown. Both STAT5a and HER4 form dimers when activated, and it is uncertain how these dimers associate with each other. Elucidating the mechanism of HER4-STAT5a binding will help to predict the effects of molecular perturbations, such as mutations, on the downstream signaling events that are induced by the HER4-STAT5a complex.

To construct such a complex, we first required structures for both HER4 and STAT5a. As no active STAT5a crystal structure currently exists, we homology modeled

the STAT5a protein based on the highly homologous STAT1 family member, for which a crystal structure of the active conformation does exist[25]. Here we note a few points regarding the active STAT structure. The unphosphorylated and phosphorylated STAT monomers essentially have the same structure; no conformational change occurs within the STAT core fragment upon phosphorylation. Rather, the role of phosphorylation is to alter the STAT dimerization mode: unphosphorylated STAT proteins dimerize in one of two modes (antiparallel or parallel), while phosphorylated STAT proteins dimerize as a 'pair of pliers'. This 'pliers' interface is necessary for subsequent DNA-binding in the nucleus. In this active conformation, the SH2 domains and phosphotyrosine residues of two STAT monomers form a reciprocal interaction such that the dimer is shaped like a pair of pliers. This mode of dimerization allows for high-affinity DNA-binding.

We generated the active STAT5a structure based on the STAT1 crystal structure, and constructed the active HER4 dimer as described in Chapter 4. We then applied various alignment algorithms to construct several preliminary HER4-STAT5a complexes, based on the structure of STAT bound to phosphotyrosine residues located in its SH2 domain. We also based our structures on the experimental evidence indicating that STAT5a interacts with HER4 via the STAT SH2 domain.

Once our structures have been ranked and further refined, we aim to answer questions such as: is it possible for HER4 to phosphorylate STAT5a directly, or is a protein mediator required? This is an important question, as there has been controversy regarding the role of HER4 in phosphorylation of STAT5a; indeed, one view is that HER4 may serve as a scaffold protein to indirectly promote the phosphorylation of STAT5a by JAK2[123]. Thus our structure will be employed to sort through possibilities

regarding the nature of the HER4-STAT5a interaction, and also help to delineate the sequence of signaling events downstream of HER4 activation (*e.g.*, does STAT5a dimerize and then bind to HER4, or do STAT5a monomers bind to HER4 and subsequently dimerize?). We also plan to extend our analysis protocol to the construction of other macromolecular complexes involved in the ErbB signaling network, including the EGFR-STAT5b interaction.

6.2.2 Future work: Transcriptional regulatory network analysis of the anti-cancer HER4 signaling pathway in mammary cells

As studies of gene expression analysis have become increasingly high-dimensional and multiplexed, so have studies of gene regulation by transcriptional regulatory networks. A computational program called Promoter Analysis and Interaction Network Toolset (PAINT) was recently developed by Vadigepalli et al.[192], and combines bioinformatics with statistical significance testing to predict candidate regulatory interactions between TFs and the transcriptional regulatory elements (TREs) of differentially expressed genes. PAINT retrieves upstream TREs from the TRANSFAC database[114] of promoter sequences of known or predicted genes in several eukaryotic species, including human, mouse, and rat, and has been successfully applied to the generation of candidate transcriptional regulatory networks in a wide range of biological systems[86, 89, 139, 148, 149]. Specifically, given a gene expression data set, PAINT will: (1) fetch potential promoter sequences for the genes in the list, (2) find TF-binding sites on the sequences, (3) analyze the TF-binding site occurrences for over/under-representation compared with a reference, and (4) generate multiple visualizations for these analyses. PAINT is currently available as an online, web-based service located at: http://www.dbi.tju.edu/dbi/tools/paint.

As a sample dataset, we applied the PAINT method to microarray data that had been collected to investigate gene expression downstream of HER4 activation in several mammary cell lines[3, 4]. In this study, the molecular mechanisms contributing to the functional differences in HER2 versus HER4 overexpression in breast cancer were investigated using microarray analysis. Specifically, agonistic antibodies were employed to activate HER2 and HER4 in isolation from the other ErbBs in breast cancer cells, and gene expression was performed to identify transcriptional targets of receptor activation. It was demonstrated that there are receptor-specific targets that are preferentially regulated by each receptor, and moreover, that two ligands acting via the same receptor dimer may induce expression of different subsets of genes.

To extend the microarray study to the transcriptional regulatory networks operating in HER2 versus HER4-stimulated cells, we applied the PAINT analysis to the gene expression dataset. We first clustered the differentially expressed genes based on the corresponding stimulatory ligand and identified TREs positioned 5000 base pairs upstream of the transcriptional start site of each gene. A Feasnet was then produced, which is a matrix consisting of each gene and its putative TREs/TFs. The Feasnet was then subjected to further PAINT analysis in order to obtain only the most statistically important TREs. In this refined matrix, each gene is assigned to its respective cluster (for example, genes that are upregulated by the ligand HB-EGF versus genes that are upregulated by the HER4 Ab). To determine which TREs are over-represented in each cluster, PAINT requires some type of reference file. We designated the entire list of genes included in the microarray as our reference set, as this selection removes any bias inherent in the microarray relative to the entire genome. The enrichment analysis results

were filtered using p-value ranges of $0<p<0.01$ and $0.01<p<0.05$ to derive the regulatory network hypotheses (Figure 6.1).

We observed over-representation of several regulatory elements in the interaction matrix, which suggests their potential role in regulation of specific gene clusters. The Ingenuity Pathways Analysis software (Ingenuity® Systems, www.ingenuity.com), which aids in gene ontology annotation and construction of putative signaling pathways, was applied to assign functional properties to the genes and TFs included in our analysis (Fig. 6.2). Specifically, we identified several TFs associated with cellular endocytosis, including LRF and its target gene SCAMP3, which is a component of the endosomal sorting machinery, and YIF1A, which is required for vesicle transport from the endoplasmic reticulum to the Golgi stack. It is plausible that these TFs are upregulated to facilitate internalization of the ErbB receptors in response to ligand binding. Intriguingly, stimulation of the mammary cell lines with the HER4 ligands NRG and HB-EGF resulted in different sets of activated genes and PAINT TF predictions, although the putative TFs perform similar functional roles in endocytosis. Hence different ligands signaling through the same receptor, in this case, the HER4 homodimer, generate differential TF regulatory patterns that nonetheless culminate in equivalent cellular outcomes.

In addition to endocytosis, we observed enrichment of several regulatory elements involved in the estrogen receptor (ER) pathway, including the TFs ARNT and COUPTF. This finding is in agreement with previous studies, which have demonstrated the formation of a HER4-ER complex and its translocation to the nucleus, resulting in upregulation of ER target genes. As both HER4 and ER are nuclear receptors with transcriptional properties, it is plausible that the crosstalk between the HER4 and ER

signaling pathways produces a synergistic effect at the level of TFs to direct mammary gland development. Hence our preliminary PAINT analysis of the ErbB dataset generated

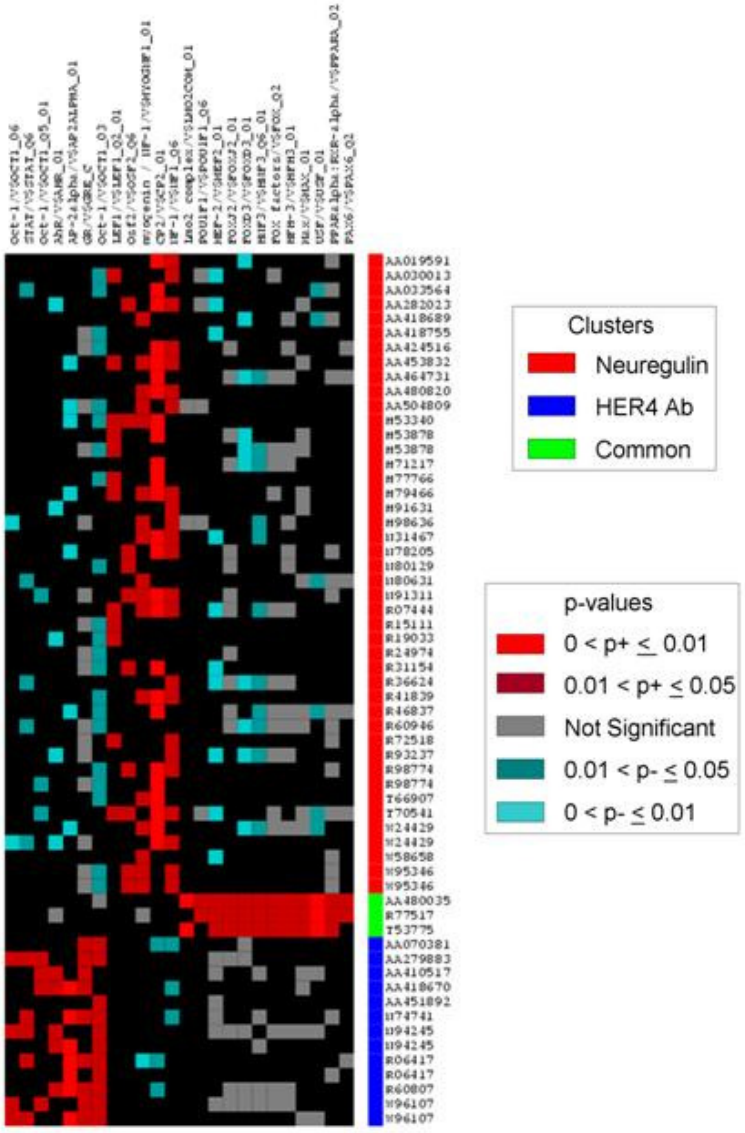

Figure 6.1. PAINT interaction matrix describing interactions in HER4-stimulated cells.
The list of differentially expressed genes is displayed on the y-axis, and the putative TREs are displayed on the x-axis at top. TREs which are over-represented in a specific gene subset are highlighted in red; TREs which are under-represented in a specific gene subset are in cyan.

testable hypotheses regarding the role of specific regulatory elements in the ErbB pathway in mammary carcinoma cells. Additionally, we optimized several PAINT parameters for implementation in our analysis of the forthcoming HC11 cell datasets.

Future studies will focus on microarray analysis and the application of PAINT to datasets obtained from HER4-stimulated HC11 cells during mammary differentiation, using the ligands NRG and HB-EGF. A subset of the genes predicted by PAINT to be differentially regulated by the activated TFs will be validated by qRT-PCR and clustered according to time point or type of ligand stimulation. Promoter regions of genes in these clusters will then be examined for shared TF binding sites by comparing the over-representation of each binding site relative to a reference set of genes. Appropriate statistical thresholds will be implemented in PAINT to rank the likelihood of involvement of individual TFs in each cluster and to refine the list of co-regulated genes. Cluster analysis of the genes is an important step because it links each TF binding event with either activation or suppression of specific genes involved in the HER4-mediated signaling pathway. The hypothesis motivating the PAINT analysis is that co-expressed genes share binding elements in their promoters, leading to co-regulation of the genes by specific TFs. Hence genes sharing identical TREs can be associated in a molecular network that may represent key pharmacological targets for modulating or exploiting the HER4 pathway in malignant cells.

To elucidate the linkage between differential gene expression and cell decision-making during HC11 cell differentiation, genes will be categorized according to their functional annotations (gene ontology). We predict increased expression of genes associated with cell differentiation concomitant with decreased expression of genes

involved in proliferation, depending on the ligand used for cell stimulation, as we have shown that NRG and HB-EGF produce slightly different functional outcomes. In particular, NRG has been demonstrated to produce more pronounced morphological signs of differentiation, whereas HB-EGF may stimulate a slightly weaker response due to co-activation of proliferation pathways.

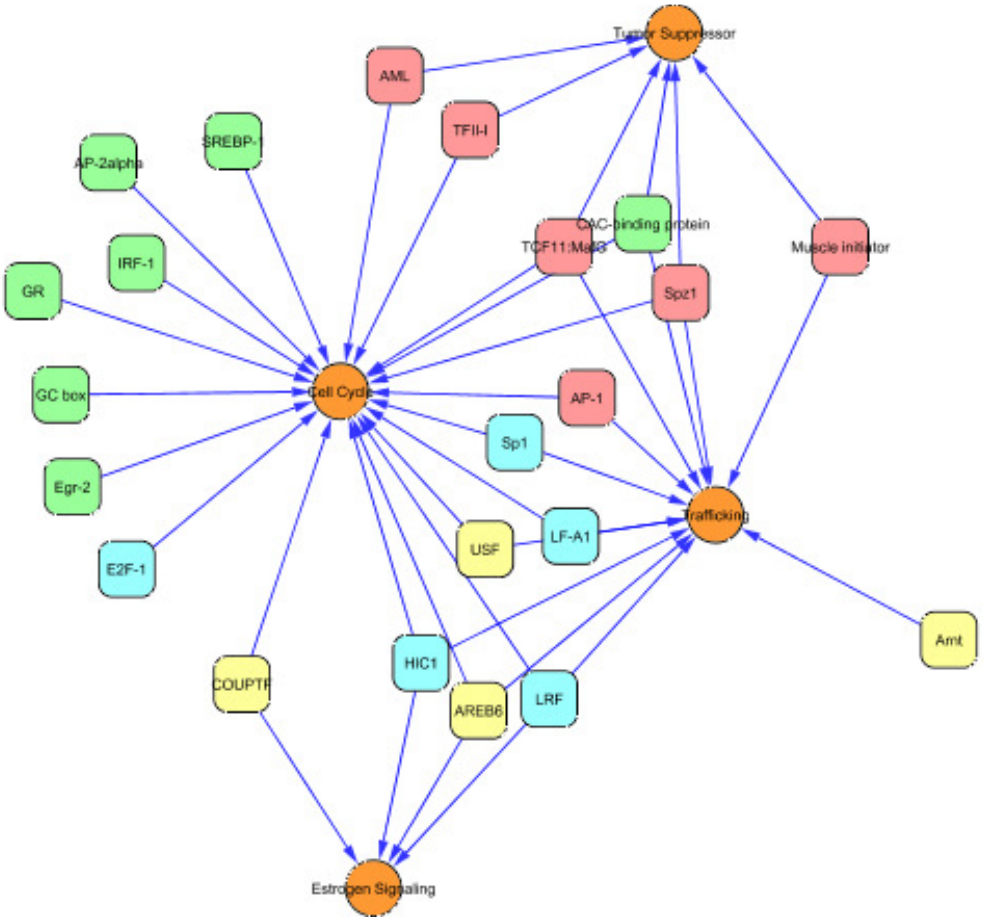

Figure 6.2. Transcription factor interaction matrix generated in Cytoscape[159].
The TFs postulated to play a role in the biological network of interest are associated with specific biological functions (cell cycle, trafficking, tumor suppressor) using gene ontology.

Bibliography

1. Acharya, H., S. Vembanur, S. N. Jamadagni, and S. Garde. (2010). Mapping hydrophobicity at the nanoscale: applications to heterogeneous surfaces and proteins. *Faraday Discuss* 146:353-365; discussion 367-393, 395-401.
2. Alimandi, M., A. Romano, M. C. Curia, R. Muraro, P. Fedi, S. A. Aaronson, P. P. Di Fiore, and M. H. Kraus. (1995). Cooperative signaling of ErbB3 and ErbB2 in neoplastic transformation and human mammary carcinomas. *Oncogene* 10:1813-1821.
3. Amin, D. N., A. S. Perkins, and D. F. Stern. (2004). Gene expression profiling of ErbB receptor and ligand-dependent transcription. *Oncogene* 23:1428-1438.
4. Amin, D. N., D. Tuck, and D. F. Stern. (2005). Neuregulin-regulated gene expression in mammary carcinoma cells. *Exp Cell Res* 309:12-23.
5. Antonescu, C. R., A. Yoshida, T. Guo, N.-E. Chang, L. Zhang, N. P. Agaram, L.-X. Qin, M. F. Brennan, S. Singer, and R. G. Maki. (2009). KDR Activating Mutations in Human Angiosarcomas Are Sensitive to Specific Kinase Inhibitors. *Cancer Research* 69:7175-7179.
6. Bagossi, P., G. Horvath, G. Vereb, J. Szollosi, and J. Tozser. (2005). Molecular modeling of nearly full-length ErbB2 receptor. *Biophys J* 88:1354-1363.
7. Balak, M. N., Y. Gong, G. J. Riely, R. Somwar, A. R. Li, M. F. Zakowski, A. Chiang, G. Yang, O. Ouerfelli, M. G. Kris, M. Ladanyi, V. A. Miller, and W. Pao. (2006). Novel D761Y and Common Secondary T790M Mutations in Epidermal Growth Factor Receptorâ€“Mutant Lung Adenocarcinomas with Acquired Resistance to Kinase Inhibitors. *Clinical Cancer Research* 12:6494-6501.
8. Baselga, J., and S. M. Swain. (2009). Novel anticancer targets: revisiting ERBB2 and discovering ERBB3. *Nat Rev Cancer* 9:463-475.
9. Beadling, C., E. Jacobson-Dunlop, F. S. Hodi, C. Le, A. Warrick, J. Patterson, A. Town, A. Harlow, F. Cruz, S. Azar, B. P. Rubin, S. Muller, R. West, M. C. Heinrich, and C. L. Corless. (2008). KIT Gene Mutations and Copy Number in Melanoma Subtypes. *Clinical Cancer Research* 14:6821-6828.
10. 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.
11. Bhandarkar, M., R. Brunner, C. Chipot, A. Dalke, A., S. Dixit, P. Grayson, J. Gullingsrud, A. Gursoy, D. Hardy, J. Hénin, W. Humphrey, D. Hurwitz, N. Krawetz, S.

Kumar, M. Nelson, J. Phillips, A. Shinozaki, G. Zheng, and F. Zhu. NAMD user's guide, version 2.6b1. Theoretical biophysics group, University of Illinois and Beckman Institute, 405 North Mathews, Urbana, Illinois 61801, July 2005.

12. Bill, A., A. Schmitz, B. Albertoni, J.-N. Song, L. C. Heukamp, D. Walrafen, F. Thorwirth, P. J. Verveer, S. Zimmer, L. Meffert, A. Schreiber, S. Chatterjee, R. K. Thomas, R. T. Ullrich, T. Lang, and M. Famulok. (2010). Cytohesins Are Cytoplasmic ErbB Receptor Activators. *Cell* 143:201-211.
13. Birtwistle, M. R., M. Hatakeyama, N. Yumoto, B. A. Ogunnaike, J. B. Hoek, and B. N. Kholodenko. (2007). Ligand-dependent responses of the ErbB signaling network: experimental and modeling analyses. *Mol Syst Biol* 3:144.
14. Blencke, S., A. Ullrich, and H. Daub. (2003). Mutation of Threonine 766 in the Epidermal Growth Factor Receptor Reveals a Hotspot for Resistance Formation against Selective Tyrosine Kinase Inhibitors. *Journal of Biological Chemistry* 278:15435-15440.
15. Boerner, J. L., J. S. Biscardi, C. M. Silva, and S. J. Parsons. (2005). Transactivating agonists of the EGF receptor require Tyr 845 phosphorylation for induction of DNA synthesis. *Mol Carcinog* 44:262-273.
16. Boldrini, L., S. Ursino, S. Gisfredi, P. Faviana, V. Donati, T. Camacci, M. Lucchi, A. Mussi, F. Basolo, R. Pingitore, and G. Fontanini. (2004). Expression and Mutational Status of c-kit in Small-Cell Lung Cancer. *Clinical Cancer Research* 10:4101-4108.
17. 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.
18. Boudeau, J., D. Miranda-Saavedra, G. J. Barton, and D. R. Alessi. (2006). Emerging roles of pseudokinases. *Trends Cell Biol* 16:443-452.
19. 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.
20. Burgess, A. W., H. S. Cho, C. Eigenbrot, K. M. Ferguson, T. P. J. Garrett, D. J. Leahy, M. A. Lemmon, M. X. Sliwkowski, C. W. Ward, and S. Yokoyama. (2003). An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. *Mol. Cell* 12:541-552.
21. Buttitta, F., F. Barassi, G. Fresu, L. Felicioni, A. Chella, D. Paolizzi, G. Lattanzio, S. Salvatore, P. P. Camplese, S. Rosini, T. Iarussi, F. Mucilli, R. Sacco, A. Mezzetti, and A. Marchetti. (2006). Mutational analysis of the HER2 gene in lung tumors from Caucasian

patients: Mutations are mainly present in adenocarcinomas with bronchioloalveolar features. *International Journal of Cancer* 119:2586-2591.

22. Carraway, K. L., 3rd, M. X. Sliwkowski, R. Akita, J. V. Platko, P. M. Guy, A. Nuijens, A. J. Diamonti, R. L. Vandlen, L. C. Cantley, and R. A. Cerione. (1994). The erbB3 gene product is a receptor for heregulin. *J Biol Chem* 269:14303-14306.
23. Carraway, K. L., 3rd, S. P. Soltoff, A. J. Diamonti, and L. C. Cantley. (1995). Heregulin stimulates mitogenesis and phosphatidylinositol 3-kinase in mouse fibroblasts transfected with erbB2/neu and erbB3. *J Biol Chem* 270:7111-7116.
24. Chen, W. W., B. Schoeberl, P. J. Jasper, M. Niepel, U. B. Nielsen, D. A. Lauffenburger, and P. K. Sorger. (2009). Input-output behavior of ErbB signaling pathways as revealed by a mass action model trained against dynamic data. *Mol Syst Biol* 5:239.
25. Chen, X., U. Vinkemeier, Y. Zhao, D. Jeruzalmi, J. E. Darnell, Jr., and J. Kuriyan. (1998). Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to DNA. *Cell* 93:827-839.
26. Cho, H. S., and D. J. Leahy. (2002). Structure of the extracellular region of HER3 reveals an interdomain tether. *Science* 297:1330-1333.
27. 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.
28. Cho, K. S., J. S. Lee, N. H. Cho, K. Park, W. S. Ham, and Y. D. Choi. (2008). Gene amplification and mutation analysis of epidermal growth factor receptor in hormone refractory prostate cancer. *The Prostate* 68:803-808.
29. 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.
30. 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.
31. Citri, A., and Y. Yarden. (2006). EGF-ERBB signalling: towards the systems level. *Nat Rev Mol Cell Biol* 7:505-516.
32. Clark, D. E., C. C. Williams, T. T. Duplessis, K. L. Moring, A. R. Notwick, W. Long, W. S. Lane, I. Beuvink, N. E. Hynes, and F. E. Jones. (2005). ERBB4/HER4 potentiates STAT5A transcriptional activity by regulating novel STAT5A serine phosphorylation events. *J Biol Chem* 280:24175-24180.

33. Coker, K. J., J. V. Staros, and C. A. Guyer. (1994). A kinase-negative epidermal growth factor receptor that retains the capacity to stimulate DNA synthesis. *Proc. Natl. Acad. Sci. U. S. A.* 91:6967-6971.

34. Curtin, J. A., K. Busam, D. Pinkel, and B. C. Bastian. (2006). Somatic Activation of KIT in Distinct Subtypes of Melanoma. *Journal of Clinical Oncology* 24:4340-4346.

35. Da Silva, L., P. Simpson, C. Smart, S. Cocciardi, N. Waddell, A. Lane, B. Morrison, A. Vargas, S. Healey, J. Beesley, P. Pakkiri, S. Parry, N. Kurniawan, L. Reid, P. Keith, P. Faria, E. Pereira, A. Skalova, M. Bilous, R. Balleine, H. Do, A. Dobrovic, S. Fox, M. Franco, B. Reynolds, K. Khanna, M. Cummings, G. Chenevix-Trench, and S. Lakhani. (2010). HER3 and downstream pathways are involved in colonization of brain metastases from breast cancer. *Breast Cancer Research* 12:R46.

36. Davies, H., C. Hunter, R. Smith, P. Stephens, C. Greenman, G. Bignell, J. Teague, A. Butler, S. Edkins, C. Stevens, A. Parker, S. O'Meara, T. Avis, S. Barthorpe, L. Brackenbury, G. Buck, J. Clements, J. Cole, E. Dicks, K. Edwards, S. Forbes, M. Gorton, K. Gray, K. Halliday, R. Harrison, K. Hills, J. Hinton, D. Jones, V. Kosmidou, R. Laman, R. Lugg, A. Menzies, J. Perry, R. Petty, K. Raine, R. Shepherd, A. Small, H. Solomon, Y. Stephens, C. Tofts, J. Varian, A. Webb, S. West, S. Widaa, A. Yates, F. Brasseur, C. S. Cooper, A. M. Flanagan, A. Green, M. Knowles, S. Y. Leung, L. H. J. Looijenga, B. Malkowicz, M. A. Pierotti, B. T. Teh, S. T. Yuen, S. R. Lakhani, D. F. Easton, B. L. Weber, P. Goldstraw, A. G. Nicholson, R. Wooster, M. R. Stratton, and P. A. Futreal. (2005). Somatic Mutations of the Protein Kinase Gene Family in Human Lung Cancer. *Cancer Research* 65:7591-7595.

37. Ding, L., G. Getz, D. A. Wheeler, E. R. Mardis, M. D. McLellan, K. Cibulskis, C. Sougnez, H. Greulich, D. M. Muzny, M. B. Morgan, L. Fulton, R. S. Fulton, Q. Zhang, M. C. Wendl, M. S. Lawrence, D. E. Larson, K. Chen, D. J. Dooling, A. Sabo, A. C. Hawes, H. Shen, S. N. Jhangiani, L. R. Lewis, O. Hall, Y. Zhu, T. Mathew, Y. Ren, J. Yao, S. E. Scherer, K. Clerc, G. A. Metcalf, B. Ng, A. Milosavljevic, M. L. Gonzalez-Garay, J. R. Osborne, R. Meyer, X. Shi, Y. Tang, D. C. Koboldt, L. Lin, R. Abbott, T. L. Miner, C. Pohl, G. Fewell, C. Haipek, H. Schmidt, B. H. Dunford-Shore, A. Kraja, S. D. Crosby, C. S. Sawyer, T. Vickery, S. Sander, J. Robinson, W. Winckler, J. Baldwin, L. R. Chirieac, A. Dutt, T. Fennell, M. Hanna, B. E. Johnson, R. C. Onofrio, R. K. Thomas, G. Tonon, B. A. Weir, X. Zhao, L. Ziaugra, M. C. Zody, T. Giordano, M. B. Orringer, J. A. Roth, M. R. Spitz, I. I. Wistuba, B. Ozenberger, P. J. Good, A. C. Chang, D. G. Beer, M. A. Watson, M. Ladanyi, S. Broderick, A. Yoshizawa, W. D. Travis, W. Pao, M. A.

Province, G. M. Weinstock, H. E. Varmus, S. B. Gabriel, E. S. Lander, R. A. Gibbs, M. Meyerson, and R. K. Wilson. (2008). Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* 455:1069-1075.

38. Dixit, A., and G. M. Verkhivker. (2009). Hierarchical Modeling of Activation Mechanisms in the ABL and EGFR Kinase Domains: Thermodynamic and Mechanistic Catalysts of Kinase Activation by Cancer Mutations. *PLoS Comput Biol* 5:e1000487.
39. Doherty, G. J., and H. T. McMahon. (2009). Mechanisms of Endocytosis. *Annual Review of Biochemistry* 78:857-902.
40. Donnini, S., A. E. Mark, A. H. Juffer, and A. Villa. (2005). Incorporating the effect of ionic strength in free energy calculations using explicit ions. *J Comput Chem* 26:115-122.
41. Douillard, J.-Y., E. Kim, V. Hirsh, T. Mok, M. Socinski, R. Gervais, Y.-L. Wu, L. Li, M. Sellers, and E. Lowe. (2007). Gefitinib (IRESSA) versus docetaxel in patients with locally advanced or metastatic non-small-cell lung cancer pre-treated with platinum-based chemotherapy: a randomized, open-label Phase III study (INTEREST): PRS-02. *Journal of Thoracic Oncology* 2:S305-S306 310.1097/1001.JTO.0000283087.0000271346.0000283019.
42. Downward, J., Y. Yarden, E. Mayes, G. Scrace, N. Totty, P. Stockwell, A. Ullrich, J. Schlessinger, and M. D. Waterfield. (1984). Close similarity of epidermal growth factor receptor and v-erb-B oncogene protein sequences. *Nature* 307:521-527.
43. Engelman, J. A., T. Mukohara, K. Zejnullahu, E. Lifshits, A. M. BorrÃ¡s, C.-M. Gale, G. N. Naumov, B. Y. Yeap, E. Jarrell, J. Sun, S. Tracy, X. Zhao, J. V. Heymach, B. E. Johnson, L. C. Cantley, and P. A. JÃ¤nne. (2006). Allelic dilution obscures detection of a biologically significant resistance mutation in EGFR-amplified lung cancer. *The Journal of Clinical Investigation* 116:2695-2706.
44. Engelman, J. A., K. Zejnullahu, T. Mitsudomi, Y. Song, C. Hyland, J. O. Park, N. Lindeman, C. M. Gale, X. Zhao, J. Christensen, T. Kosaka, A. J. Holmes, A. M. Rogers, F. Cappuzzo, T. Mok, C. Lee, B. E. Johnson, L. C. Cantley, and P. A. Janne. (2007). MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* 316:1039-1043.
45. 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.
46. Fedosov, D. A., B. Caswell, and G. E. Karniadakis. (2010). A multiscale red blood cell model with accurate mechanics, rheology, and dynamics. *Biophys J* 98:2215-2225.

47. Feller, S. E., Y. Zhang, R.W. Pastor, and B.R. Brooks. (1995). Constant pressure molecular dynamics simulation: The Langevin piston method. *J Chem Phys* 103:4613-4621.

48. Feng, S. M., C. I. Sartor, D. Hunter, H. Zhou, X. Yang, L. S. Caskey, R. Dy, R. S. Muraoka-Cook, and H. S. Earp, 3rd. (2007). The HER4 cytoplasmic domain, but not its C terminus, inhibits mammary cell proliferation. *Mol Endocrinol* 21:1861-1876.

49. 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.

50. Ferrer-Soler, L., A. Vazquez-Martin, J. Brunet, J. A. Menendez, R. De Llorens, and R. Colomer. (2007). An update of the mechanisms of resistance to EGFR-tyrosine kinase inhibitors in breast cancer: Gefitinib (Iressa) -induced changes in the expression and nucleo-cytoplasmic trafficking of HER-ligands (Review). *Int J Mol Med* 20:3-10.

51. Fiser, A., R. K. Do, and A. Sali. (2000). Modeling of loops in protein structures. *Protein Sci* 9:1753-1773.

52. Forbes, S. A., G. Tang, N. Bindal, S. Bamford, E. Dawson, C. Cole, C. Y. Kok, M. Jia, R. Ewing, A. Menzies, J. W. Teague, M. R. Stratton, and P. A. Futreal. (2010). COSMIC (the Catalogue of Somatic Mutations in Cancer): a resource to investigate acquired mutations in human cancer. *Nucleic Acids Research* 38:D652-D657.

53. Frey, K. M., I. Georgiev, B. R. Donald, and A. C. Anderson. (2010). Predicting resistance mutations using protein design algorithms. *Proceedings of the National Academy of Sciences* 107:13707-13712.

54. Fukuoka, M., S. Yano, G. Giaccone, T. Tamura, K. Nakagawa, J.-Y. Douillard, Y. Nishiwaki, J. Vansteenkiste, S. Kudoh, D. Rischin, R. Eek, T. Horai, K. Noda, I. Takata, E. Smit, S. Averbuch, A. Macleod, A. Feyereislova, R.-P. Dong, and J. Baselga. (2003). Multi-Institutional Randomized Phase II Trial of Gefitinib for Previously Treated Patients With Advanced Non-Small-Cell Lung Cancer. *Journal of Clinical Oncology* 21:2237-2246.

55. Gao, J., K. Kuczera, B. Tidor, and M. Karplus. (1989). Hidden thermodynamics of mutant proteins: a molecular dynamics analysis. *Science* 244:1069-1072.

56. 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.

57. Gartside, M. G., H. Chen, O. A. Ibrahimi, S. A. Byron, A. V. Curtis, C. L. Wellens, A. Bengston, L. M. Yudt, A. V. Eliseenkova, J. Ma, J. A. Curtin, P. Hyder, U. L. Harper, E. Riedesel, G. J. Mann, J. M. Trent, B. C. Bastian, P. S. Meltzer, M. Mohammadi, and P. M. Pollock. (2009). Loss-of-Function Fibroblast Growth Factor Receptor-2 Mutations in Melanoma. *Molecular Cancer Research* 7:41-54.

58. Gassmann, M. (1995). Aberrant neural and cardiac development in mice lacking the ErbB4 neuregulin receptor. *Nature* 378:390-394.

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

60. Godawat, R., S. N. Jamadagni, and S. Garde. (2009). Characterizing hydrophobicity of interfaces by using cavity formation, solute binding, and water correlations. *Proc Natl Acad Sci U S A* 106:15119-15124.

61. Griffith, J., J. Black, C. Faerman, L. Swenson, M. Wynn, F. Lu, J. Lippke, and K. Saxena. (2004). The structural basis for autoinhibition of FLT3 by the juxtamembrane domain. *Molecular Cell* 13:169-178.

62. Grubmuller, H., B. Heymann, and P. Tavan. (1996). Ligand binding: molecular mechanics calculation of the streptavidin-biotin rupture force. *Science* 271:997-999.

63. Hafner, C., J. M. M. van Oers, A. Hartmann, M. Landthaler, R. Stoehr, H. Blaszyk, F. Hofstaedter, E. C. Zwarthoff, and T. Vogt. (2006). High Frequency of FGFR3 Mutations in Adenoid Seborrheic Keratoses. *Journal of Investigative Dermatology* 126:2404-2407.

64. Hafner, C., A. Hartmann, F. X. Real, F. Hofstaedter, M. Landthaler, and T. Vogt. (2007). Spectrum of FGFR3 Mutations in Multiple Intraindividual Seborrheic Keratoses. *Journal of Investigative Dermatology* 127:1883-1885.

65. Hahn, C. G., H. Y. Wang, D. S. Cho, K. Talbot, R. E. Gur, W. H. Berrettini, K. Bakshi, J. Kamins, K. E. Borgmann-Winter, S. J. Siegel, R. J. Gallop, and S. E. Arnold. (2006). Altered neuregulin 1-erbB4 signaling contributes to NMDA receptor hypofunction in schizophrenia. *Nature Medicine* 12:824-828.

66. Hamburger, A. W. (2008). The role of ErbB3 and its binding partners in breast cancer progression and resistance to hormone and tyrosine kinase directed therapies. *J Mammary Gland Biol Neoplasia* 13:225-233.

67. Henin, J., B. Maigret, M. Tarek, C. Escrieut, D. Fourmy, and C. Chipot. (2006). Probing a model of a GPCR/ligand complex in an explicit membrane environment: the human cholecystokinin-1 receptor. *Biophys J* 90:1232-1240.

68. Holbro, T., R. R. Beerli, F. Maurer, M. Koziczak, C. F. Barbas, 3rd, and N. E. Hynes. (2003). The ErbB2/ErbB3 heterodimer functions as an oncogenic unit: ErbB2 requires ErbB3 to drive breast tumor cell proliferation. *Proc Natl Acad Sci U S A* 100:8933-8938.

69. Hsieh, M. H., Y. F. Fang, W. C. Chang, H. P. Kuo, S. Y. Lin, H. P. Liu, C. L. Liu, H. C. Chen, Y. C. Ku, Y. T. Chen, Y. H. Chang, Y. T. Chen, B. L. Hsi, S. F. Tsai, and S. F. Huang. (2006). Complex mutation patterns of epidermal growth factor receptor gene associated with variable responses to gefitinib treatment in patients with non-small cell lung cancer. *Lung Cancer* 53:311-322.

70. Huang, S.-F., H.-P. Liu, L.-H. Li, Y.-C. Ku, Y.-N. Fu, H.-Y. Tsai, Y.-T. Chen, Y.-F. Lin, W.-C. Chang, H.-P. Kuo, Y.-C. Wu, Y.-R. Chen, and S.-F. Tsai. (2004). High Frequency of Epidermal Growth Factor Receptor Mutations with Complex Patterns in Non-Small Cell Lung Cancers Related to Gefitinib Responsiveness in Taiwan. *Clinical Cancer Research* 10:8195-8203.

71. Hubbard, S. R., L. Wei, L. Ellis, and W. A. Hendrickson. (1994). Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature* 372:746-754.

72. Hubbard, S. R. (1997). Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *Embo J* 16:5572-5581.

73. Hubbard, S. R., and J. H. Till. (2000). Protein Tyrosine Kinase Structure and Function. *Annual Review of Biochemistry* 69:373-398.

74. Humphrey, W., A. Dalke, and K. Schulten. (1996). VMD: visual molecular dynamics. *J Mol Graph* 14:33-38, 27-38.

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

76. Hutcheson, I. R., J. M. Knowlden, S. E. Hiscox, D. Barrow, J. M. Gee, J. F. Robertson, I. O. Ellis, and R. I. Nicholson. (2007). Heregulin beta1 drives gefitinib-resistant growth and invasion in tamoxifen-resistant MCF-7 breast cancer cells. *Breast Cancer Res* 9:R50.

77. Inoue, A., K. Setoguchi, Y. Matsubara, K. Okada, N. Sato, Y. Iwakura, O. Higuchi, and Y. Yamanashi. (2009). Dok-7 Activates the Muscle Receptor Kinase MuSK and Shapes Synapse Formation. *Sci. Signal.* 2:ra7,1-8.

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

79. 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.

80. Jones, F. E., T. Welte, X. Y. Fu, and D. F. Stern. (1999). ErbB4 signaling in the mammary gland is required for lobuloalveolar development and Stat5 activation during lactation. *J Cell Biol* 147:77-88.

81. Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. (1983). Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics* 79:926--935.

82. Joslin, E. J., L. K. Opresko, A. Wells, H. S. Wiley, and D. A. Lauffenburger. (2007). EGF-receptor-mediated mammary epithelial cell migration is driven by sustained ERK signaling from autocrine stimulation. *J Cell Sci* 120:3688-3699.

83. 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.

84. 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.

85. 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.

86. Juskeviciute, E., R. Vadigepalli, and J. B. Hoek. (2008). Temporal and functional profile of the transcriptional regulatory network in the early regenerative response to partial hepatectomy in the rat. *BMC Genomics* 9:527.

87. Kabotyanski, E. B., M. Huetter, W. Xian, M. Rijnkels, and J. M. Rosen. (2006). Integration of prolactin and glucocorticoid signaling at the beta-casein promoter and enhancer by ordered recruitment of specific transcription factors and chromatin modifiers. *Mol Endocrinol* 20:2355-2368.

88. Kabotyanski, E. B., M. Rijnkels, C. Freeman-Zadrowski, A. C. Buser, D. P. Edwards, and J. M. Rosen. (2009). Lactogenic hormonal induction of long distance interactions between beta-casein gene regulatory elements. *J Biol Chem* 284:22815-22824.

89. Khan, R. L., R. Vadigepalli, M. K. McDonald, R. F. Rogers, G. R. Gao, and J. S. Schwaber. (2008). Dynamic transcriptomic response to acute hypertension in the nucleus tractus solitarius. *Am J Physiol Regul Integr Comp Physiol* 295:R15-27.

90. Kloth, M. T., K. K. Laughlin, J. S. Biscardi, J. L. Boerner, S. J. Parsons, and C. M. Silva. (2003). STAT5b, a Mediator of Synergism between c-Src and the Epidermal Growth Factor Receptor. *J Biol Chem* 278:1671-1679.

91. Knowles, P. P., J. Murray-Rust, S. Kjaer, R. P. Scott, S. Hanrahan, M. Santoro, C. F. Ibanez, and N. Q. McDonald. (2006). Structure and Chemical Inhibition of the RET Tyrosine Kinase Domain. *Journal of Biological Chemistry* 281:33577-33587.

92. Kong-Beltran, M., S. Seshagiri, J. Zha, W. Zhu, K. Bhawe, N. Mendoza, T. Holcomb, K. Pujara, J. Stinson, L. Fu, C. Severin, L. Rangell, R. Schwall, L. Amler, D. Wickramasinghe, and R. Yauch. (2006). Somatic Mutations Lead to an Oncogenic Deletion of Met in Lung Cancer. *Cancer Research* 66:283-289.

93. 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.

94. Kornev, A. P., S. S. Taylor, and L. F. Ten Eyck. (2008). A helix scaffold for the assembly of active protein kinases. *Proc Natl Acad Sci U S A* 105:14377-14382.

95. Kosaka, T., Y. Yatabe, H. Endoh, H. Kuwano, T. Takahashi, and T. Mitsudomi. (2004). Mutations of the Epidermal Growth Factor Receptor Gene in Lung Cancer: Biological and Clinical Implications. *Cancer Research* 64:8919-8923.

96. Kosaka, T., Y. Yatabe, H. Endoh, K. Yoshida, T. Hida, M. Tsuboi, H. Tada, H. Kuwano, and T. Mitsudomi. (2006). Analysis of Epidermal Growth Factor Receptor Gene Mutation in Patients with Non-Small Cell Lung Cancer and Acquired Resistance to Gefitinib. *Clinical Cancer Research* 12:5764-5769.

97. Kris, M. G., R. B. Natale, R. S. Herbst, T. J. Lynch, D. Prager, C. P. Belani, J. H. Schiller, K. Kelly, H. Spiridonidis, A. Sandler, K. S. Albain, D. Cella, M. K. Wolf, S. D. Averbuch, J. J. Ochs, and A. C. Kay. (2003). Efficacy of Gefitinib, an Inhibitor of the Epidermal Growth Factor Receptor Tyrosine Kinase, in Symptomatic Patients With Non-Small Cell Lung Cancer. *JAMA* 290:2149-2158.

98. Lahiri, S. D., G. Zhang, D. Dunaway-Mariano, and K. N. Allen. (2003). The pentacovalent phosphorus intermediate of a phosphoryl transfer reaction. *Science* 299:2067-2071

99. 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.

100. Le Novere, N., B. Bornstein, A. Broicher, M. Courtot, M. Donizelli, H. Dharuri, L. Li, H. Sauro, M. Schilstra, B. Shapiro, J. L. Snoep, and M. Hucka. (2006). BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Res* 34:D689-691.

101. Le Novere, N., M. Hucka, H. Mi, S. Moodie, F. Schreiber, A. Sorokin, E. Demir, K. Wegner, M. I. Aladjem, S. M. Wimalaratne, F. T. Bergman, R. Gauges, P. Ghazal, H. Kawaji, L. Li, Y. Matsuoka, A. Villeger, S. E. Boyd, L. Calzone, M. Courtot, U. Dogrusoz, T. C. Freeman, A. Funahashi, S. Ghosh, A. Jouraku, S. Kim, F. Kolpakov, A. Luna, S. Sahle, E. Schmidt, S. Watterson, G. Wu, I. Goryanin, D. B. Kell, C. Sander, H. Sauro, J. L. Snoep, K. Kohn, and H. Kitano. (2009). The Systems Biology Graphical Notation. *Nat Biotechnol* 27:735-741.

102. 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.

103. 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. *Clin. Cancer Res.* 12:57-61.

104. Lemmon, M. A., and J. Schlessinger. (2010). Cell Signaling by Receptor Tyrosine Kinases. *Cell* 141:1117-1134.

105. Li, C., M. Donizelli, N. Rodriguez, H. Dharuri, L. Endler, V. Chelliah, L. Li, E. He, A. Henry, M. I. Stefan, J. L. Snoep, M. Hucka, N. Le Novere, and C. Laibe. (2010). BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Syst Biol* 4:92.

106. Linggi, B., and G. Carpenter. (2006). ErbB receptors: new insights on mechanisms and biology. *Trends Cell Biol* 16:649-656.

107. Liu, B., D. Ordonez-Ercan, Z. Fan, S. M. Edgerton, X. Yang, and A. D. Thor. (2007). Downregulation of erbB3 abrogates erbB2-mediated tamoxifen resistance in breast cancer cells. *Int J Cancer* 120:1874-1882.

108. Liu, Y., J. Purvis, A. Shih, J. Weinstein, N. Agrawal, and R. Radhakrishnan. (2007). A multiscale computational approach to dissect early events in the Erb family receptor

mediated activation, differential signaling, and relevance to oncogenic transformations. *Ann Biomed Eng* 35:1012-1025.

109. Logie, A., C. Dunois-Larde, C. Rosty, O. Levrel, M. Blanche, A. Ribeiro, J.-M. Gasc, J. Jorcano, S. Werner, X. Sastre-Garau, J. P. Thiery, and F. Radvanyi. (2005). Activating mutations of the tyrosine kinase receptor FGFR3 are associated with benign skin tumors in mice and humans. *Human Molecular Genetics* 14:1153-1160.

110. Lorenzato, A., M. Olivero, S. PatanÃ¨, E. Rosso, A. Oliaro, P. M. Comoglio, and M. F. Di Renzo. (2002). Novel Somatic Mutations of the MET Oncogene in Human Carcinoma Metastases Activating Cell Motility and Invasion. *Cancer Research* 62:7025-7030.

111. 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. *New England Journal of Medicine* 350:2129-2139.

112. MacKerell, A. D., 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, 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. *Journal of Physical Chemistry B* 102:3586-3616.

113. Manning, G., D. B. Whyte, R. Martinez, T. Hunter, and S. Sudarsanam. (2002). The protein kinase complement of the human genome. *Science* 298:1912-1934.

114. Matys, V., E. Fricke, R. Geffers, E. Gossling, M. Haubrock, R. Hehl, K. Hornischer, D. Karas, A. E. Kel, O. V. Kel-Margoulis, D. U. Kloos, S. Land, B. Lewicki-Potapov, H. Michael, R. Munch, I. Reuter, S. Rotert, H. Saxel, M. Scheer, S. Thiele, and E. Wingender. (2003). TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res* 31:374-378.

115. Menendez, J. A., and R. Lupu. (2007). Transphosphorylation of kinase-dead HER3 and breast cancer progression: a new standpoint or an old concept revisited? *Breast Cancer Res* 9:111.

116. Min, X., B. H. Lee, M. H. Cobb, and E. J. Goldsmith. (2004). Crystal structure of the kinase domain of WNK1, a kinase that causes a hereditary form of hypertension. *Structure* 12:1303-1311.

117. Minami, Y., T. Shimamura, K. Shah, T. LaFramboise, K. A. Glatt, E. Liniker, C. L. Borgman, H. J. Haringsma, W. Feng, B. A. Weir, A. M. Lowell, J. C. Lee, J. Wolf, G. I.

Shapiro, K. K. Wong, M. Meyerson, and R. K. Thomas. (2007). The major lung cancer-derived mutants of ERBB2 are oncogenic and are associated with sensitivity to the irreversible EGFR/ERBB2 inhibitor HKI-272. *Oncogene* 26:5023-5027.

118. Mohammadi, M., J. Schlessinger, and S. R. Hubbard. (1996). Structure of the FGF receptor tyrosine kinase domain reveals a novel autoinhibitory mechanism. *Cell* 86:577-587.

119. Mol, C. D., D. R. Dougan, T. R. Schneider, R. J. Skene, M. L. Kraus, D. N. Scheibe, G. P. Snell, H. Zou, B. C. Sang, and K. P. Wilson. (2004). Structural basis for the autoinhibition and STI-571 inhibition of c-Kit tyrosine kinase. *Journal of Biological Chemistry* 279:31655-31663.

120. Motoyama, A. B., and N. E. Hynes. (2003). {BAD}: a good therapeutic target? *Breast Cancer Research* 5:27--30.

121. Mukherjee, K., M. Sharma, H. Urlaub, G. P. Bourenkov, R. Jahn, T. C. Sudhof, and M. C. Wahl. (2008). CASK Functions as a Mg2+-independent neurexin kinase. *Cell* 133:328-339.

122. Muraoka-Cook, R. S., M. Sandahl, C. Husted, D. Hunter, L. Miraglia, S. M. Feng, K. Elenius, and H. S. Earp, 3rd. (2006). The intracellular domain of ErbB4 induces differentiation of mammary epithelial cells. *Mol Biol Cell* 17:4118-4129.

123. Muraoka-Cook, R. S., M. Sandahl, D. Hunter, L. Miraglia, and H. S. Earp, 3rd. (2008). Prolactin and ErbB4/HER4 signaling interact via Janus kinase 2 to induce mammary epithelial cell gene expression differentiation. *Mol Endocrinol* 22:2307-2321.

124. Muraoka-Cook, R. S., M. A. Sandahl, K. E. Strunk, L. C. Miraglia, C. Husted, D. M. Hunter, K. Elenius, L. A. Chodosh, and H. S. Earp, 3rd. (2009). ErbB4 splice variants Cyt1 and Cyt2 differ by 16 amino acids and exert opposing effects on the mammary epithelium in vivo. *Mol Cell Biol* 29:4935-4948.

125. Nagashima, T., R. Ushikoshi-Nakayama, A. Suenaga, K. Ide, N. Yumoto, Y. Naruo, K. Takahashi, Y. Saeki, M. Taiji, H. Tanaka, S. F. Tsai, and M. Hatakeyama. (2009). Mutation of epidermal growth factor receptor is associated with MIG6 expression. *Febs J* 276:5239-5251.

126. Niho, S., Y. Ichinose, T. Tamura, N. Yamamoto, M. Tsuboi, K. Nakagawa, T. Shinkai, H. Jiang, Y. Nishiwaki, and M. Fukuoka. (2007). Results of a randomized phase III study to compare the overall survival of gefitinib (IRESSA) versus docetaxel in Japanese patients with non-small cell lung cancer who failed one or two chemotherapy regimens. *J. Clin. Oncol. (Meeting Abstracts)* 25:LBA7509-.

127. Offterdinger, M., C. Schofer, K. Weipoltshammer, and T. W. Grunt. (2002). c-erbB-3: a nuclear protein in mammary epithelial cells. *J Cell Biol* 157:929-939.

128. Ogiso, H., R. Ishitani, O. Nureki, S. Fukai, M. Yamanaka, J.-H. Kim, K. Saito, A. Sakamoto, M. Inoue, M. Shirouzu, and S. Yokoyama. (2002). Crystal Structure of the Complex of Human Epidermal Growth Factor and Receptor Extracellular Domains. *Cell* 110:775-787.

129. Ozkirimli, E., and C. B. Post. (2006). Src kinase activation: A switched electrostatic network. *Protein Science* 15:1051-1062.

130. Ozkirimli, E., S. S. Yadav, W. T. Miller, and C. B. Post. (2008). An electrostatic network and long-range regulation of Src kinases. *Protein Science* 17:1871-1880.

131. 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.

132. 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.

133. Pao, W., V. A. Miller, K. A. Politi, G. J. Riely, R. Somwar, M. F. Zakowski, M. G. Kris, and H. Varmus. (2005). Acquired Resistance of Lung Adenocarcinomas to Gefitinib or Erlotinib Is Associated with a Second Mutation in the EGFR Kinase Domain. *PLoS Med* 2:e73.

134. Papakyriakou, A., D. Vourloumis, F. Tzortzatou-Stathopoulou, and M. Karpusas. (2009). Conformational dynamics of the EGFR kinase domain reveals structural features involved in activation. *Proteins: Structure, Function, and Bioinformatics* 76:375-386.

135. 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.

136. 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. *J Comput Chem* 26:1781-1802.

137. 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.

138. Plowman, G. D., G. S. Whitney, M. G. Neubauer, J. M. Green, V. L. McDonald, G. J. Todaro, and M. Shoyab. (1990). Molecular cloning and expression of an additional epidermal growth factor receptor-related gene. *Proc Natl Acad Sci U S A* 87:4905-4909.

139. Pratt, C. H., R. Vadigepalli, P. Chakravarthula, G. E. Gonye, N. J. Philp, and G. B. Grunwald. (2008). Transcriptional regulatory network analysis during epithelial-mesenchymal transformation of retinal pigment epithelium. *Mol Vis* 14:1414-1428.

140. Prickett, T. D., N. S. Agrawal, X. Wei, K. E. Yates, J. C. Lin, J. R. Wunderlich, J. C. Cronin, P. Cruz, S. A. Rosenberg, and Y. Samuels. (2009). Analysis of the tyrosine kinome in melanoma reveals recurrent mutations in ERBB4. *Nature Genetics* 41:1127-1132.

141. 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.

142. Ramanan, V., N. J. Agrawal, J. Liu, S. Engles, R. Toy, and R. Radhakrishnan. Systems Biology and Physical Biology of Clathrin-Mediated Endocytosis: An Integrative Experimental and Theoretical Perspective. *Submitted to Integr. Biol.*

143. Rausenberger, J., A. Hussong, S. Kircher, D. Kirchenbauer, J. Timmer, F. Nagy, E. Schafer, and C. Fleck. (2010). An integrative model for phytochrome B mediated photomorphogenesis: from protein dynamics to physiology. *PLoS One* 5:e10721.

144. 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.

145. Reynolds, A. R., C. Tischer, P. J. Verveer, O. Rocks, and P. I. H. Bastiaens. (2003). EGFR activation coupled to inhibition of tyrosine phosphatases causes lateral signal propagation. *Nat Cell Biol* 5:447-453.

146. Riese, D. J., 2nd, T. M. van Raaij, G. D. Plowman, G. C. Andrews, and D. F. Stern. (1995). The cellular response to neuregulins is governed by complex interactions of the erbB receptor family. *Mol Cell Biol* 15:5770-5776.

147. Rosta, E., H. L. Woodcock, B. R. Brooks, and G. Hummer. (2009). Artificial reaction coordinate "tunneling" in free-energy calculations: the catalytic reaction of RNase H. *J. Comput. Chem.* 30:1634-1641.

148. Saban, M. R., H. L. Hellmich, M. Turner, N. B. Nguyen, R. Vadigepalli, D. W. Dyer, R. E. Hurst, M. Centola, and R. Saban. (2006). The inflammatory and normal transcriptome of mouse bladder detrusor and mucosa. *BMC Physiol* 6:1.

149. Saban, R., C. Simpson, R. Vadigepalli, S. Memet, I. Dozmorov, and M. R. Saban. (2007). Bladder inflammatory transcriptome in response to tachykinins: neurokinin 1 receptor-dependent genes and transcription regulatory elements. *BMC Urol* 7:7.

150. Sali, A., and T. L. Blundell. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234:779-815.

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

152. Schlessinger, J. (2000). Cell Signaling by Receptor Tyrosine Kinases. *Cell* 103:211-225.

153. Schmidt, H. (2007). SBaddon: high performance simulation for the Systems Biology Toolbox for MATLAB. *Bioinformatics* 23:646-647.

154. Schoeberl, B., E. A. Pace, J. B. Fitzgerald, B. D. Harms, L. Xu, L. Nie, B. Linggi, A. Kalra, V. Paragas, R. Bukhalid, V. Grantcharova, N. Kohli, K. A. West, M. Leszczyniecka, M. J. Feldhaus, A. J. Kudla, and U. B. Nielsen. (2009). Therapeutically targeting ErbB3: a key node in ligand-induced activation of the ErbB receptor-PI3K axis. *Sci Signal* 2:ra31.

155. Schoeberl, B., A. C. Faber, D. Li, M. C. Liang, K. Crosby, M. Onsum, O. Burenkova, E. Pace, Z. Walton, L. Nie, A. Fulgham, Y. Song, U. B. Nielsen, J. A. Engelman, and K. K. Wong. (2010). An ErbB3 antibody, MM-121, is active in cancers with ligand-dependent activation. *Cancer Res* 70:2485-2494.

156. Schulze, W. X., L. Deng, and M. Mann. (2005). Phosphotyrosine interactome of the ErbB-receptor kinase family. *Mol Syst Biol* 1:2005 0008.

157. Sergina, N. V., M. Rausch, D. Wang, J. Blair, B. Hann, K. M. Shokat, and M. M. Moasser. (2007). Escape from HER-family tyrosine kinase inhibitor therapy by the kinase-inactive HER3. *Nature* 445:437-441.

158. Shah, S. P., R. D. Morin, J. Khattra, L. Prentice, T. Pugh, A. Burleigh, A. Delaney, K. Gelmon, R. Guliany, J. Senz, C. Steidl, R. A. Holt, S. Jones, M. Sun, G. Leung, R. Moore, T. Severson, G. A. Taylor, A. E. Teschendorff, K. Tse, G. Turashvili, R. Varhol, R. L. Warren, P. Watson, Y. Zhao, C. Caldas, D. Huntsman, M. Hirst, M. A. Marra, and S. Aparicio. (2009). Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* 461:809-813.

159. Shannon, P., A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498-2504.

160. Sharma, S. V., D. W. Bell, J. Settleman, and D. A. Haber. (2007). Epidermal growth factor receptor mutations in lung cancer. *Nat Rev Cancer* 7:169-181.

161. Shen, M. Y., and A. Sali. (2006). Statistical potential for assessment and prediction of protein structures. *Protein Sci* 15:2507-2524.

162. Shewchuk, L. M., A. M. Hassell, B. Ellis, W. D. Holmes, R. Davis, E. L. Horne, S. H. Kadwell, D. D. McKee, and J. T. Moore. (2000). Structure of the Tie2 RTK Domain: Self-Inhibition by the Nucleotide Binding Loop, Activation Loop, and C-Terminal Tail. *Structure* 8:1105-1113.

163. 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.

164. 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.

165. 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.

166. Shih, A. J., S.E. Telesco, S.H. Choi, M.A. Lemmon, and R. Radhakrishnan. (2011). Molecular dynamics analysis of conserved hydrophobic and hydrophilic bond interaction networks in ErbB family kinases. *Biochemical Journal* 436:241-251.

167. Shih, A. J., Telesco, S.E., Radhakrishnan R. (2011). Analysis of somatic mutations in cancer: molecular mechanisms of activation in the ErbB family of receptor tyrosine kinases. *Cancers* 3:1195-1231.

168. 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 HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science* 244:707-712.

169. Sordella, R., D. W. Bell, D. A. Haber, and J. Settleman. (2004). Gefitinib-Sensitizing EGFR Mutations in Lung Cancer Activate Anti-Apoptotic Pathways. *Science* 305:1163-1167.

170. Sorkin, A., and M. von Zastrow. (2009). Endocytosis and signalling: intertwining molecular networks. *Nat Rev Mol Cell Biol* 10:609-622.

171. Soung, Y. H., J. W. Lee, S. Y. Kim, Y. P. Wang, K. H. Jo, S. W. Moon, W. S. Park, S. W. Nam, J. Y. Lee, N. J. Yoo, and S. H. Lee. (2006). Somatic mutations of the ERBB4 kinase domain in human cancers. *Int. J. Cancer* 118:1426-1429.

172. 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.

173. Stein, R. A., and J. V. Staros. (2006). Insights into the evolution of the ErbB receptor family and their ligands from sequence analysis. *BMC Evol Biol* 6:79.

174. Steitz, T. A. (1993). DNA- and RNA-dependent DNA polymerases. *Curr. Opin. Struct. Biol.* 3:31-38.

175. 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.

176. Stern, D. F. (2008). ERBB3/HER3 and ERBB2/HER2 duet in mammary development and breast cancer. *J Mammary Gland Biol Neoplasia* 13:215-223.

177. Strunk, K. E., C. Husted, L. C. Miraglia, M. Sandahl, W. A. Rearick, D. M. Hunter, H. S. Earp, 3rd, and R. S. Muraoka-Cook. (2007). HER4 D-box sequences regulate mitotic progression and degradation of the nuclear HER4 cleavage product s80HER4. *Cancer Res* 67:6582-6590.

178. Sweeney, C., D. Fambrough, C. Huard, A. J. Diamonti, E. S. Lander, L. C. Cantley, and K. L. Carraway, 3rd. (2001). Growth factor-specific signaling pathway stimulation and gene expression mediated by ErbB receptors. *J Biol Chem* 276:22685-22698.

179. Tarcic, G., S. K. Boguslavsky, J. Wakim, T. Kiuchi, A. Liu, F. Reinitz, D. Nathanson, T. Takahashi, P. S. Mischel, T. Ng, and Y. Yarden. (2009). An Unbiased Screen Identifies DEP-1 Tumor Suppressor as a Phosphatase Controlling EGFR Endocytosis. *Current Biology* 19:1788-1798.

180. Taylor, S. S., and A. P. Kornev. (2010). Yet another "active" pseudokinase, Erb3. *Proc Natl Acad Sci U S A* 107:8047-8048.

181. 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.

182. Telesco, S. E., A. J. Shih, F. Jia, and R. Radhakrishnan. (2011). A multiscale modeling approach to investigate molecular mechanisms of pseudokinase activation and drug resistance in the HER3/ErbB3 receptor tyrosine kinase signaling network. *Molecular Biosystems* 7:2066-2080.

183. Telesco, S. E., Shih, A.J., Liu, Y., Radhakrishnan, R. (2011). Molecular Simulation of Structure, Dynamics, and Function in the HER2 Receptor Tyrosine Kinase and Relevance to Cancer Mutations. *Cancer Research Journal* 4:1-35.

184. Thatcher, N., A. Chang, P. Parikh, J. Rodrigues Pereira, T. Ciuleanu, J. von Pawel, S. Thongprasert, E. H. Tan, K. Pemberton, V. Archer, and K. Carroll. (2005). Gefitinib plus best supportive care in previously treated patients with refractory advanced non-small-cell lung cancer: results from a randomised, placebo-controlled, multicentre study (Iressa Survival Evaluation in Lung Cancer). *The Lancet* 366:1527-1537.

185. Thiel, K. W., and G. Carpenter. (2007). Epidermal growth factor receptor juxtamembrane region regulates allosteric tyrosine kinase activation. *Proceedings of the National Academy of Sciences* 104:19238-19243.

186. 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. *Proc Natl Acad Sci U S A* 96:1415-1420.

187. Tidcombe, H., A. Jackson-Fisher, K. Mathers, D. F. Stern, M. Gassmann, and J. P. Golding. (2003). Neural and mammary gland defects in ErbB4 knockout mice genetically rescued from embryonic lethality. *Proceedings of the National Academy of Sciences* 100:8281-8286.

188. Till, J. H., M. Becerra, A. Watty, Y. Lu, Y. Ma, T. A. Neubert, S. J. Burden, and S. R. Hubbard. (2002). Crystal Structure of the MuSK Tyrosine Kinase: Insights into Receptor Autoregulation. *Structure* 10:1187-1196.

189. Tonks, N. K. (2006). Protein tyrosine phosphatases: from genes, to function, to disease. *Nat Rev Mol Cell Biol* 7:833-846.

190. Torres-Cabala, C. A., W.-L. Wang, J. Trent, D. Yang, S. Chen, J. Galbincea, K. B. Kim, S. Woodman, M. Davies, J. A. Plaza, J. W. Nash, V. G. Prieto, A. J. Lazar, and D. Ivan. (2009). Correlation between KIT expression and KIT mutation in melanoma: a study of

173 cases with emphasis on the acral-lentiginous/mucosal type. *Modern Pathology* 22:1446-1456.

191. Tvorogov, D., M. Sundvall, K. Kurppa, M. Hollmen, S. Repo, M. S. Johnson, and K. Elenius. (2009). Somatic Mutations of ErbB4. *Journal of Biological Chemistry* 284:5582-5591.

192. Vadigepalli, R., P. Chakravarthula, D. E. Zak, J. S. Schwaber, and G. E. Gonye. (2003). PAINT: a promoter analysis and interaction network generation tool for gene regulatory network identification. *Omics* 7:235-252.

193. Vasalou, C., and M. A. Henson. (2010). A multiscale model to investigate circadian rhythmicity of pacemaker neurons in the suprachiasmatic nucleus. *PLoS Comput Biol* 6:e1000706.

194. Vecchi, M., J. Baulida, and G. Carpenter. (1996). Selective Cleavage of the Heregulin Receptor ErbB-4 by Protein Kinase C Activation. *Journal of Biological Chemistry* 271:18989-18995.

195. 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.

196. Weinstein, I. B. (2002). Cancer. Addiction to oncogenes--the Achilles heal of cancer. *Science* 297:63-64.

197. Williams, C. C., J. G. Allison, G. A. Vidal, M. E. Burow, B. S. Beckman, L. Marrero, and F. E. Jones. (2004). The ERBB4/HER4 receptor tyrosine kinase regulates gene expression by functioning as a STAT5A nuclear chaperone. *J Cell Biol* 167:469-478.

198. Willmore-Payne, C., J. A. Holden, S. Tripp, and L. J. Layfield. (2005). Human malignant melanoma: detection of BRAF- and c-kit-activating mutations by high-resolution amplicon melting analysis. *Human Pathology* 36:486-493.

199. 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.

200. 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.

201. Wybenga-Groot, L. E., B. Baskin, S. H. Ong, J. F. Tong, T. Pawson, and F. Sicheri. (2001). Structural basis for autoinhibition of the EphB2 receptor tyrosine kinase by the unphosphorylated juxtamembrane region. *Cell* 106:745-757.

202. 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.

203. 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.

204. Yamada, S., S. Shiono, A. Joo, and A. Yoshimura. (2003). Control mechanism of JAK/STAT signal transduction pathway. *FEBS Lett* 534:190-196.

205. Yang, S., N. K. Banavali, and B. t. Roux. (2009). Mapping the conformational transition in Src activation by cumulating the information from multiple molecular dynamics trajectories. *Proceedings of the National Academy of Sciences* 106:3776-3781.

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

207. 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.

208. 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.

209. Zheng, J., E. A. Trafny, D. R. Knighton, N. H. Xuong, S. S. Taylor, L. F. Ten Eyck, and J. M. Sowadski. (1993). 2.2 Å refined crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MnATP and a peptide inhibitor. *Acta Crystallogr. D Biol. Crystallogr.* 49:362-365.

210. 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.*

211. Zi, Z., Y. Zheng, A. E. Rundell, and E. Klipp. (2008). SBML-SAT: a systems biology markup language (SBML) based sensitivity analysis tool. *BMC Bioinformatics* 9:342.

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