UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Studies of the Allosteric Activation of the Epidermal Growth Factor Receptor

Permalink
https://escholarship.org/uc/item/6g31k0nr

Author
Engel, Katherine Anne

Publication Date
2013

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Studies of the Allosteric Activation of the Epidermal Growth Factor Receptor

By

Katherine Anne Engel

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular and Cell Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor John Kuriyan, Chair
Professor Susan Marqusee
Professor David Savage
Professor David Wemmer

Spring 2013

Abstract

Studies of the Allosteric Activation of the Epidermal Growth Factor Receptor
by
Katherine Anne Engel
Doctor of Philosophy in Molecular and Cell Biology
University of California, Berkeley
Professor John Kuriyan, Chair

Ligand binding to the extracellular domain of the epidermal growth factor receptor (EGFR) results in receptor dimerization and allosteric activation of the kinase domain through formation of a specific asymmetric dimer. This asymmetric dimer is required for EGFR to undergo autophosphorylation and propagate the signal. Structurally, formation of the asymmetric dimer stabilizes the active conformation of the kinase domain, wherein the structural rearrangements create a hydrophobic pocket on the amino-terminal lobe (N-lobe) of one kinase at the dimer interface. However, the affinity for formation of the dimer by the kinase domain alone in solution is very weak. As a result the kinase domain is predominantly monomeric and bears very low catalytic activity. Studies of mechanisms that promote formation of the asymmetric dimer or otherwise stabilize the active conformation may provide additional insights into the allosteric activation of the kinase domain.

We hypothesized that a small molecule could allosterically activate the kinase domain by binding at the N-lobe hydrophobic pocket. Furthermore, a compound that binds at the N-lobe may sterically block formation of the asymmetric dimer and inhibit dimerization-mediated autophosphorylation. A small molecule library was screened for activation of the EGFR kinase domain, and five compounds were discovered to increase kinase activity. Characterization of the compounds shows that each compound has unique properties of binding to EGFR, dependence on reductant, reversibility, and activity with EGFR mutants, which suggests distinct mechanisms for each compound in activation of EGFR. The molecular mechanisms by which the compounds activate EGFR are unknown due to the difficulty of obtaining co-crystal structures. Nonetheless, it appears that at least one compound may block formation of the asymmetric dimer of kinase domains. This work demonstrates that the EGFR kinase domain may be activated by small molecules, and the small molecules that were discovered may help elucidate the mechanisms that control allosteric activation of the EGFR kinase domain.

The EGFR kinase domain phosphorylates most peptide substrates with a relatively low catalytic efficiency, likely due to the poor affinity for kinase domain dimerization in solution. Peptide C is a synthetic peptide substrate of EGFR developed by others that is phosphorylated with a significantly higher catalytic efficiency, and we sought to understand the basis for this. Peptide C was found to increase EGFR kinase activity by promoting formation of the EGFR kinase domain asymmetric dimer. Activation of the kinase domain by Peptide C also enhances phosphorylation of other substrates. Aggregation of the EGFR kinase domain by Peptide C likely underlies activation, and Peptide C precipitates several other proteins. Peptide C was found to

form fibrils independent of the presence of EGFR, and these fibrils may facilitate aggregation and activation of the kinase domain. These results establish that a peptide substrate of EGFR may increase catalytic activity by promoting kinase domain dimerization and allosteric activation by an aggregation-mediated mechanism.

Table of Contents

Abstract 1

Table of Contents i
List of Figures iii
List of Tables iv
Acknowledgements v

Chapter 1: Introduction 1
1.1 Historical overview of kinase research 2
1.2 Protein kinases 2
1.3 Receptor tyrosine kinase structure and function 5
1.4 Overview of the EGFR family proteins 5
1.5 Signaling by EGFR family proteins and its implications in cancer 9
1.6 Structure and regulation of the EGFR kinase domain 9
1.7 Expanding our understanding of the mechanisms controlling EGFR kinase activation 13

Chapter 2: Discovery of small molecule activators of the EGFR kinase domain 15
2.1 Overview 16
2.2 Results and Discussion 20
2.2.1 Selection of the ADP Hunter screen to identify activators of the EGFR kinase domain 20
2.2.2 Optimal parameters for screening EGFR activity 22
2.2.3 Identification of small molecule activators of the EGFR kinase domain by a high throughput screen 26
2.2.4 Initial validation of the hit compounds 28
2.2.5 Five of the hit compounds are true activators of the EGFR kinase domain 29
2.2.6 The activating effects of the compounds are largely reversible and sensitive to reductant 35
2.2.7 The compounds do not require the asymmetric dimer interface to activate the EGFR kinase domain 36
2.2.8 Structure activity relationship studies examine the roles of the compound functional groups 39
2.2.9 The effects of the activating compounds on formation of the EGFR kinase domain asymmetric dimer 40
2.2.10 Attempts to obtain co-crystal structures of compounds bound to the EGFR kinase domain 45
2.3 Conclusions 46
2.4 Materials and Methods 48
2.4.1 EGFR kinase domain expression and purification 48
2.4.2 ADP Hunter assay 48
2.4.3 High throughput screen 49
2.4.4 Surface plasmon resonance binding assay 49

2.4.5 Data fitting 50
2.4.6 Radiometric kinase assay 50
2.4.7 Precipitation analysis 50

Chapter 3: A highly efficient peptide substrate for EGFR activates the kinase by inducing aggregation 51
3.1 Overview 52
3.2 Results and Discussion 53
3.2.1 Analysis of catalytic parameters and cooperativity for EGFR peptide substrates 53
3.2.2 Peptide C stimulates phosphorylation of other substrates 56
3.2.3 Increased EGFR activity with Peptide C requires formation of the asymmetric dimer 58
3.2.4 The dimerization mutants exhibit similar cooperativity with Peptide C to wild type EGFR 61
3.2.5 Peptide C precipitates the EGFR kinase domain 61
3.2.6 Peptide C precipitates proteins and forms aggregates on its own 62
3.2.7 Electron microscopy shows that Peptide C forms fibrils 63
3.3 Conclusions 66
3.4 Materials and Methods 67
3.4.1 Protein expression and purification 67
3.4.2 Peptide substrates 68
3.4.3 Enzyme-coupled kinase assays 68
3.4.4 Hill analysis 68
3.4.5 Immunoblot analysis 69
3.4.6 Radiometric kinase assays 69
3.4.7 Src-mediated phosphorylation of EGFR 69
3.4.8 Precipitation analysis 69
3.4.9 Dynamic light scattering measurements 70
3.4.10 Transmission electron microscopy 70

References 71

List of Figures

Figure 1.1 Features of the inactive and active states of kinase domains and receptor tyrosine kinases 4
Figure 1.2 The domain organization and activation mechanism of EGFR 7
Figure 1.3 Structural features of the inactive and active EGFR kinase domains and the asymmetric dimer formed during kinase activation 11

Figure 2.1 The expected properties of a small molecule activator of the EGFR kinase domain *in vitro* and in cells 19
Figure 2.2 Detection of EGFR activity using the ADP Hunter assay 21
Figure 2.3 The workflow for the EGFR kinase reaction and ADP Hunter assay and results of optimized assay conditions 25
Figure 2.4 Scatter plot of the results of the screen for small molecule activators of the EGFR kinase domain 28
Figure 2.5 Validation of hit compounds as activators of the EGFR kinase domain 31
Figure 2.6 EGFR kinase domain activation and binding data for each of the five hit compounds 34
Figure 2.7 The reversibility and DTT dependence of compound activation of EGFR. . 36
Figure 2.8 The activation of dimerization mutant EGFR kinase domains by the compounds 38
Figure 2.9 Structure activity relationship analysis of the hit compounds 40
Figure 2.10 The effects of the hit compounds on EGFR dimerization-dependent activity 43

Figure 3.1 Michaelis-Menten fits for Tail Peptide A and Tyrsub 54
Figure 3.2 Hill coefficient measurements for EGFR kinase domain catalysis with EGFR peptide substrates 55
Figure 3.3 The effects of Peptide C on EGFR kinase domain phosphorylation of other substrates 57
Figure 3.4 Effect of activation loop Tyr-845 phosphorylation on EGFR kinase domain catalytic activity 58
Figure 3.5 Catalytic activities of EGFR kinase domain dimerization mutants with Peptide C 60
Figure 3.6 Precipitation of proteins by Peptide C 62
Figure 3.7 Dynamic light scattering autocorrelation curves for Peptide C samples 63
Figure 3.8 TEM images of Peptide C with and without the EGFR kinase domain 65

List of Tables

Table 3.1 Kinetic parameters for EGFR kinase domain activity with EGFR peptide substrates..... 54

Acknowledgements

I am extremely grateful to John Kuriyan for serving as my advisor during my graduate studies. I am very fortunate to have had the opportunity to learn how to approach both my research and science in general from a scientist who is so brilliant and inspirational. John provided me with exceptional support through my most difficult times in my graduate education, and I will always be thankful for his dedication in seeing me successfully complete my doctoral work.

I thank Natalia Jura, who mentored me during my early years in the lab and served as a constant source of knowledge about the intricacies of EGFR signaling. I also thank the other lab members who have studied EGFR with me, including Nick Endres, Rahul Das, Erika Kovacs, Aaron Cantor, and Yongjian Huang. All of our work was preceded by the very significant contributions to our understanding of the EGFR kinase domain made by the mighty Xuewu Zhang.

I am thankful for all of the exceptional people who I have worked with in the Kuriyan Lab over my graduate career. I thank the talented young scientists who I had the opportunity to mentor, Charlene Bashore and Tomoaki Sasaki, for their contributions to my work and teaching how to be an effective mentor. I also thank Julie Zorn and Qi Wang for helping my Peptide C project reach fruition. Steven Jacques continuously does an amazing job managing the Kuriyan Lab, and Xiaoxian Cao provided priceless support in performing all of the insect cell protein expression. My fellow graduate students, especially Patrick Visperas and Aaron Cantor, receive special acknowledgements for making this journey with me and providing laughter (and sometimes distractions) during our many days in the lab.

I am grateful to Komen for the Cure for providing financial support for the small molecule screen work. I thank everyone who has had a role in this project. Jim Wells, Michelle Arkin, and Mark Moasser helped to conceive this work and provided additional guidance, and Kenny Ang, Steven Chen, and Preeti Chugha assisted in the discovery process. I also thank David King and Jack Sadowsky for peptide synthesis.

Finally, I am very grateful to everyone else who has helped me during my graduate studies. My mom has been an endless source of support and love, despite the thousands of miles between us, and I am incredibly fortunate to have her in my life. I thank the rest of my big, crazy family, too. Additionally, I am extremely thankful to my friends, who comprise my family of the West Coast and with whom I have had some of my greatest life experiences.

Chapter 1

Introduction

1.1 Historical overview of kinase research

In 1906 the first observation of protein phosphorylation was published, yet it would be nearly 50 years until the discovery of the enzymes that are responsible for this modification, the protein kinases (Levene and Alsberg, 1906). The work of George Burnett and Eugene Kennedy in 1954 described the phosphorylation of casein by "protein phosphokinase," now known to be a mixture of casein kinases 1 and 2, in rat liver mitochondrial extracts. The phosphorylation of serine that they observed was found to be dependent on Mg^{2+} and adenosine triphosphate (ATP) (Burnett and Kennedy, 1954). The following year Edmond Fischer and Edwin Krebs described the effects of serine phosphorylation of glycogen phosphorylase by phosphorylase kinase in converting glycogen phosphorylase from its "b" form to the more active "a" form (Fischer and Krebs, 1955). Today we recognize this regulation of the activity of another protein by phosphorylation as one of the key roles of kinase activity.

1.2 Protein kinases

Protein kinases catalyze the process of protein phosphorylation, a post-translational modification in which the γ-phosphoryl group of ATP is covalently transferred to the hydroxyl group of serine, threonine, or tyrosine side chains or to the imidazole nitrogens of histidine side chains of proteins. Phosphorylation may alter the activity of a protein by causing a conformational change, affecting cellular localization, or by providing docking sites for other molecules. Protein phosphorylation serves many roles within the cell by regulating the proteins involved in signal transduction, metabolism, cell cycle progression, and transcription, among others, and approximately 30% of human cellular proteins undergo phosphorylation (Cohen, 2000; Pawson and Scott, 2005).

Protein kinases are ubiquitous in all life forms. Histidine kinases are present in bacteria, lower eukaryotes, and plants as a part of two-component signal transduction systems that catalyze phosphorylation of a histidine residue within the kinase itself as part of a response to extracellular cues. These histidine kinases are structurally and functionally divergent from the eukaryotic protein kinases described herein (Khorchid and Ikura, 2006; Santos and Shiozaki, 2001). Approximately 500 protein kinases have been identified within the human genome based on conservation of kinase-specific sequences. Of these, 90 specifically phosphorylate tyrosines while the wide majority of the others phosphorylate serine and threonine and a third type, termed dual-specificity kinases, phosphorylate serine, threonine, and tyrosine (Manning et al., 2002). While kinases within the superfamily are divergent in terms of the architecture and regulation of the full-length proteins, all kinase domains share defining structural and functional features. The kinase domain is roughly 30 kDa in size and consists of two distinct lobes, the amino-terminal lobe (N-lobe) which is made of a β-sheet and the catalytically important αC helix, and the carboxy-terminal lobe (C-lobe), which is mainly α-helical. The active site is formed at the cleft where these two lobes join (Figure 1.1A) (Huse and Kuriyan, 2002).

Susan Taylor and coworkers solved the atomic structure of the serine-threonine kinase protein kinase A in 1991, and the work of her group spearheaded the extensive biochemical and structural studies that have subsequently elucidated the general features of kinase activation (Knighton et al., 1991). The conformation of the active site and N-lobe predominantly define the structural differences between the inactive and active conformations of kinases. Protein substrates bind to the active kinase at the periphery of the active site near a dynamic loop of the kinase domain known as the activation loop. ATP binds near the hinge region that links the N- and C-lobes and is optimally positioned for the γ-phosphoryl group to undergo nucleophilic attack by the hydroxyl group of the protein substrate. In inactive conformations the activation loop of the kinase often packs within the active site and occludes the binding of substrate proteins as well as causes other rearrangements of critical active site residues. In many kinases phosphorylation of the activation loop stabilizes the activation loop in an extended conformation that opens the active site to be accessible to substrates, thereby activating the kinase. Additionally, the position of helix-αC relative to the rest of the N-lobe is critical in determining the activation state of the kinase domain. In the active kinase conformation, helix-αC is positioned closer to the core of the kinase domain. A glutamate in the helix forms a salt bridge with a lysine within the N-lobe, and this salt bridge is necessary for lysine-mediated coordination of the phosphates of ATP in the active state. This Lys-Glu salt bridge along with other highly-conserved active site residues are properly oriented in the active conformation to coordinate ATP, Mg^{2+}, and catalyze phosphorylation (Figure 1.1A) (Huse and Kuriyan, 2002; Jura et al., 2011).

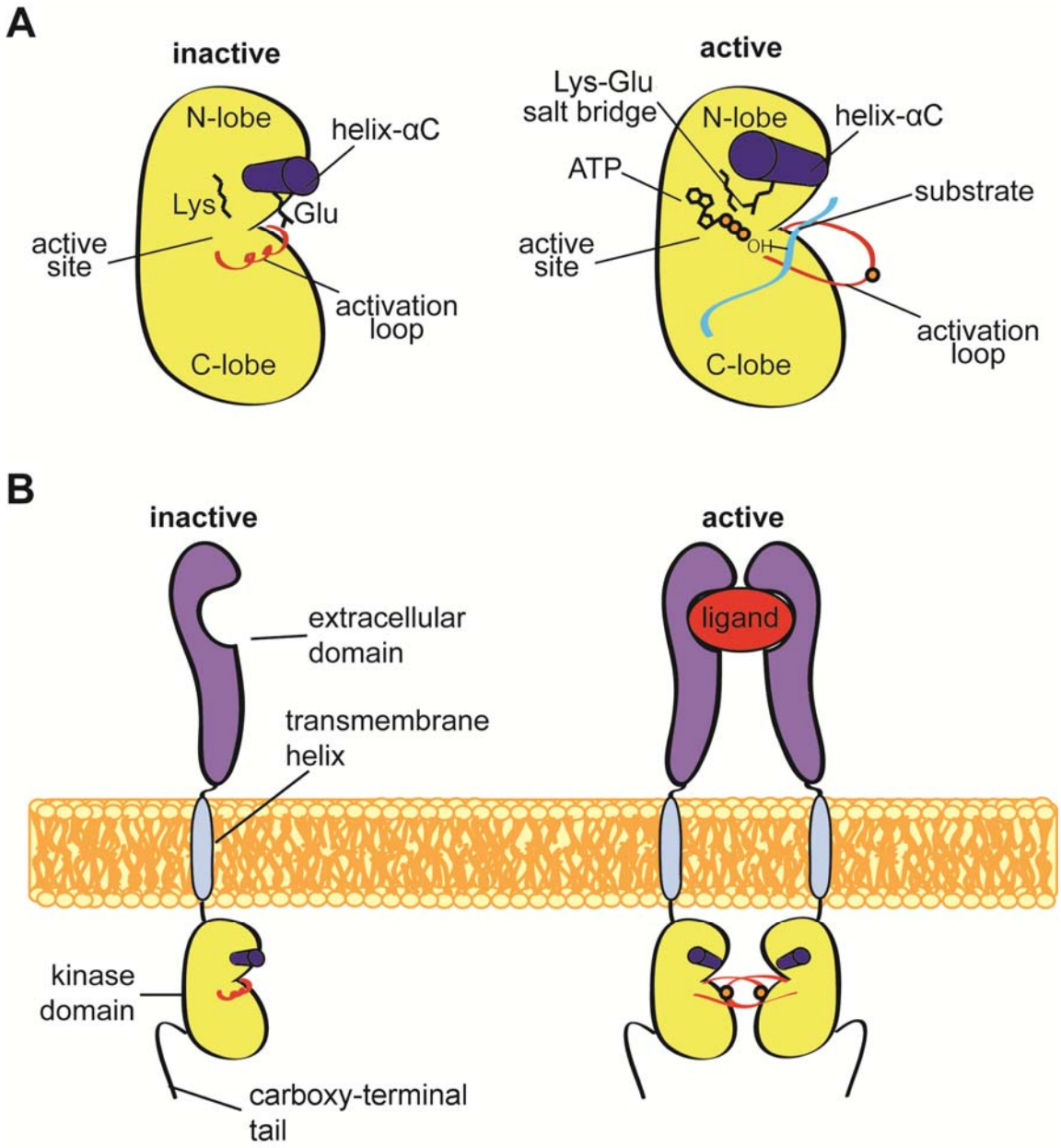


Figure 1.1 Features of the inactive and active states of kinase domains and receptor tyrosine kinases. (A) The general features of kinase domains in the inactive and active states are juxtaposed. Changes between the inactive and the active states occur predominantly within the N-lobe and active site, including rotation of helix-αC to form a critical Lys-Glu salt bridge and the activation loop adopting an extended conformation away from the active site upon kinase activation (Huse and Kuriyan, 2002). (B) The general architecture of receptor tyrosine kinases includes a ligand-binding extracellular domain, transmembrane helix, and cytoplasmic kinase domain. Binding of ligand to the extracellular domains typically induces receptor dimerization. The kinase domains undergo autophosphorylation (represented by orange circles) on the activation loops in *trans*, and the activated kinases proceed to phosphorylate additional protein substrates that serve as platforms for signal transduction (Schlessinger, 2000).

1.3 Receptor tyrosine kinase structure and function

Studies of tyrosine phosphorylation began 25 years after those examining serine-threonine phosphorylation, likely due to the relatively low abundance of tyrosine-phosphorylated proteins in the cell. Tony Hunter first observed tyrosine phosphorylation in 1979 and subsequently discovered the transforming protein of Rous sarcoma virus pp60^{src} as the first known tyrosine kinase one year later (Eckhart et al., 1979; Hunter and Sefton, 1980). Soon after, Hiroshi Ushiro and Stanley Cohen identified the epidermal growth factor receptor (EGFR) as the first receptor tyrosine kinase (RTK) (Ushiro and Cohen, 1980). This work collectively recognized the two forms of tyrosine kinases: the soluble form, such as Src, and the membrane-bound RTK, such as EGFR.

RTKs conserve a general structure and mechanism. Each protein consists of a ligand-binding extracellular domain, a single pass transmembrane helix, and an intracellular kinase domain (Figure 1.1B). The ligand for the receptor, which may be a growth factor, hormone, cytokine, or a transmembrane receptor on another cell, binds the extracellular domain and induces or alters receptor oligomerization (Figure 1.1B). For most RTKs, the ligand directly mediates dimerization. The close proximity of the intracellular kinase domains that occurs as a result of ligand binding stimulates kinase autophosphorylation in *trans* on the activation loop and increases catalytic activity of the protein for protein substrates. The substrates of RTKs include sites on the RTK itself and other proteins, and the phosphorylated tyrosines serve as docking sites for signal transduction proteins that contain Src homology 2 (SH2) or phosphotyrosine binding (PTB) domains (Schlessinger, 2000).

RTKs transduce an extracellular signal to produce a cellular response, which often involves relay of the signal to the nucleus to alter gene transcription. The signaling pathways that are commonly stimulated as a result of RTK activation include Janus kinase/signal transducer and activator of transcription (JAK/STAT), phospholipase Cγ (PLCγ), phosphatidylinositol-3 kinase (PI-3K), mitogen associated protein kinase (MAPK), and c-Jun N-terminal kinase (JNK). The variety of cellular outputs that may occur as a result of RTK activity include promotion of cell survival, differentiation, migration, and proliferation and are dependent on both the unique combinations of signaling proteins that each RTK recruits and the overall properties of the cell itself (Schlessinger, 2000).

1.4 Overview of the EGFR family proteins

Research on EGFR preceded the discovery of RTKs, beginning with the discovery of the EGFR ligand epidermal growth factor (EGF) by Stanley Cohen in 1962 as the component responsible for precocious tooth eruption and eyelid opening in newborn mice that were injected with a mouse salivary gland extract (Cohen, 1962). Subsequent research found the existence of a plasma membrane-bound receptor for EGF and characterized the internalization and degradation dynamics of this receptor, the increase in protein phosphorylation accompanying EGF stimulation, and cell growth as a physiological output of EGF stimulation (Carpenter and Cohen, 1979). Following

the discovery of EGFR and other RTKs, these proteins have become the focus of extensive research for their roles in cell signaling, and EGFR is the best-studied.

EGFR consists of a glycosylated extracellular ligand-binding domain, a single pass transmembrane α-helix that is flanked by extra- and intracellular juxtamembrane segments, the kinase domain, and a ~200 residue C-terminal tail, which contains the tyrosines that serve as autophosphorylation sites (Figure 1.2A). The receptor rests in an autoinhibited state. Binding of ligand such as EGF, amphiregulin, or transforming growth factor α, among others, releases a tether in the extracellular domain and permits formation of an active dimer that is competent to undergo autophosphorylation (Figure 1.2B) (Endres et al., 2011; Ferguson, 2008; Jorissen et al., 2003). The oligomeric state of EGFR prior to and following ligand stimulation is an ongoing area of research, and recent evidence suggests that the inactive protein may exist as a dimer and the active signaling complex may be a tetramer or higher order oligomer (Chung et al., 2010; Clayton et al., 2008; Jura et al., 2009a).

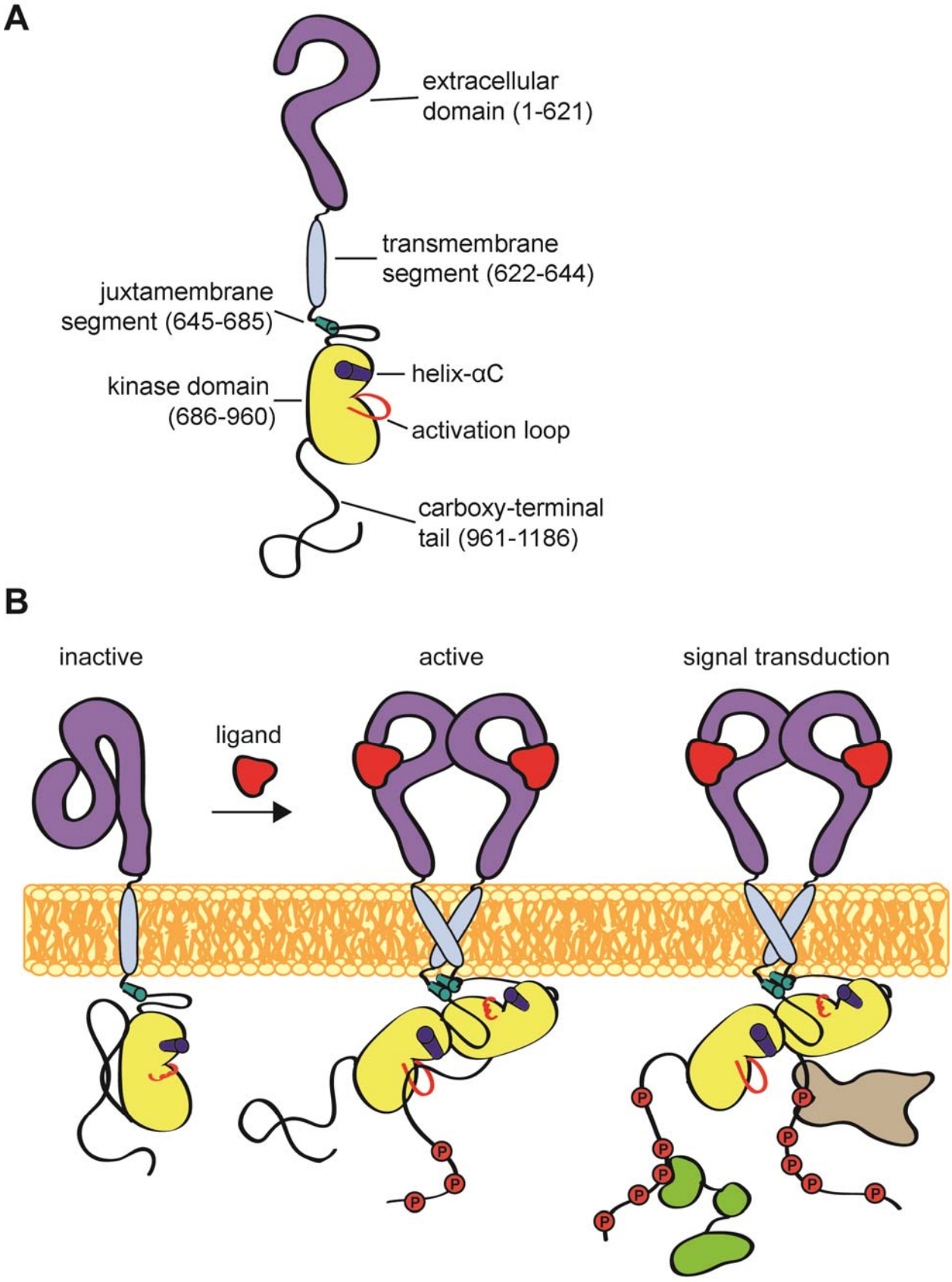

A
extracellular
domain (1-621)
transmembrane
segment (622-644)
juxtamembrane
segment (645-685)
kinase domain
(686-960)
helix-αC
activation loop
carboxy-terminal
tail (961-1186)
B
inactive
active
signal transduction
ligand
P

Figure 1.2 The domain organization and activation mechanism of EGFR. (A) The domain organization of EGFR includes a ligand-binding extracellular domain, transmembrane helix that is flanked by extracellular and cytoplasmic juxtamembrane segments, cytoplasmic kinase domain, and long carboxy-terminal tail that contains the sites of autophosphorylation. The residue range for each of these elements is indicated in parentheses and uses the EGFR numbering system that excludes the 24-residue N-terminal signal sequence. (B) Prior to ligand stimulation, EGFR adopts an autoinhibited state. Binding of ligand releases a tether in the extracellular domain and induces EGFR dimerization. Unlike other RTKS, the dimer interface is mediated solely by EGFR and not the bound ligand, and the active kinase domains form an asymmetric dimer. Activated EGFR undergoes autophosphorylation on tyrosines within the C-terminal tail, which serve as docking sites for signaling proteins involved in differentiation, migration, and growth. All members of the EGFR family of proteins function generally by this mechanism, with limited exceptions for Her2 and Her3 (Endres et al., 2011; Jorissen et al., 2003).

While EGFR shares the same general activation mechanism as other RTKs, three notable features distinguish it: first, the dimer formed by the extracellular domain is mediated solely by the protein and not the bound ligand, suggesting that the extracellular domain must adopt an autoinhibited conformation to prevent unstimulated dimerization and activation (Ferguson et al., 2003; Garrett et al., 2002; Ogiso et al., 2002). Secondly, phosphorylation of the kinase domain activation loop is not necessary for activation of EGFR kinase activity, an event that is required for activation of most kinases (Gotoh et al., 1992). Finally, the dimer formed by the EGFR kinase domains is not symmetric. The asymmetric dimer that is formed by the EGFR kinase domains is essential for EGFR activation and is discussed later in greater detail (Figure 1.2B) (Zhang et al., 2006).

EGFR belongs to the EGFR family of proteins, which includes three additional members in humans, Her2, Her3, and Her4. (The proteins are also known as ErbB1, ErbB2, ErbB3, and ErbB4.) The EGFR family of proteins share the same overall architecture and regulation and as a result are generally capable of pairing in different combinations to form homo- and heterodimers (Tzahar et al., 1996). While Her4 is largely similar to EGFR, Her2 and Her3 each have unique structural and functional features. Her2 lacks a known ligand and forms obligate heterodimers with other family members (Cho et al., 2003). Her3 has crippled kinase activity because residues that are essential for ATP binding and coordination are mutated (Guy et al., 1994; Sierke et al., 1997). As a result, Her3 must heterodimerize to form dimers that are capable of phosphorylation and signaling (Berger et al., 2004). Furthermore, the C-terminal tail sequences among the family members are relatively divergent, and each receptor recruits a variety of signaling proteins in different stoichiometries upon phosphorylation. Altogether, the capacity of the EGFR family members to homo- and heterodimerize expands the diversity of signaling that occurs following activation (Olayioye et al., 2000).

1.5 Signaling by EGFR family proteins and its implications in cancer

The EGFR family has the capacity to produce a variety of signaling outputs as the result of integrating a number of factors. The identity of the bound ligand, EGFR family member composition of the homo- or heterodimer, and presence of positive and negative effector proteins are all determinants in the outcome of EGFR family signaling (Avraham and Yarden, 2011). Of the four family members, EGFR couples to the largest number of unique signaling pathways, stimulating the protein kinase C (PKC), MAPK, PI3-K, and JNK pathways to ultimately increase transcription of target genes. The specific dynamics of EGFR signaling are not well understood, due to the many factors that affect the signaling output. Nonetheless, EGFR signaling is associated with a variety of outcomes, such as proliferation, migration, adhesion, growth, differentiation, and inhibition of apoptosis (Olayioye et al., 2000; Yarden and Sliwkowski, 2001). Collectively, fidelity of EGFR function at the cellular level bears great implications for the entire organism, as EGFR null mutations are either embryonic or newborn lethal in mice (Sibilia and Wagner, 1995; Threadgill et al., 1995).

In agreement with roles in regulating cell growth, proliferation, and inhibition of apoptosis, deregulation of the EGFR family members is frequently observed in numerous cancers. In the case of EGFR, mutations that bias the receptor to adopt the active conformation or deletions that remove autoinhibitory interactions result in increased kinase activity, and these genetic alterations have been observed in all regions of the full-length receptor (Pines et al., 2010). Alternatively, receptor overexpression as a result of gene amplification circumvents the requirement of ligand for activation and is also oncogenic (Hynes and MacDonald, 2009). The best-studied alterations to EGFR family proteins in cancers include activating mutations of EGFR in non-small cell lung cancer, overexpression or truncations of EGFR in glioblastoma, and overexpression of Her2 and Her3 as a signaling unit in breast cancer (Emde et al., 2012; Huang et al., 2009; Sharma et al., 2007). Therefore, understanding the molecular mechanisms of alterations that deregulate the EGFR family proteins in cancer and development of methods to inhibit the aberrant signaling constitute major goals of cancer research. Currently approved cancer treatments that are antagonists of EGFR family proteins include small molecule inhibitors that target the kinase active site and humanized antibodies that bind the extracellular domain and block binding of ligand or receptor dimerization. However, cancer cells frequently incur mutations or other changes to evade inhibition by these molecules, and the occurrence of drug resistance necessitates the discovery of novel inhibitors of EGFR family protein signaling (Emde et al., 2012; Pines et al., 2010).

1.6 Structure and regulation of the EGFR kinase domain

Kinases must assume one of several observed stable inactive conformations to prevent spurious activation and signaling (Jura et al., 2011). The EGFR kinase adopts the Src/CDK-like inactive conformation, named for its similarity to the inactive conformations observed in crystal structures of those kinases. In this inactive

conformation, helix-αC is swung away from the N-lobe in an orientation that is stabilized by hydrophobic packing interactions (Figure 1.3A) (Wood et al., 2004; Zhang et al., 2006). Upon activation, helix-αC rotates inward, causing exposure of a hydrophobic pocket on the N-lobe (Figure 1.3B). Furthermore, the Lys-721-Glu-738 salt bridge (the residue numbering system used here excludes the 24-residue N-terminal signal sequence) forms between helix-αC and the N-lobe, and other active site residues become positioned to be catalytically competent (Stamos et al., 2002; Zhang et al., 2006). The monomeric EGFR kinase domain is energetically more stable in the inactive conformation than the active conformation, and as a result the kinase domain is primarily inactive in its basal state (Shan et al., 2012).

Activation of the EGFR kinase domain is primarily achieved through stabilization of the active conformation by interactions between two kinase domains that form an asymmetric dimer. In this arrangement the C-lobe of one kinase domain, termed the activator, binds to the exposed hydrophobic pocket of the N-lobe of the second kinase, termed the receiver (Figures 1.3C). The entire surface area buried at the dimer interface is approximately 2000 $Å^2$. The asymmetric dimer was identified through crystal contacts and validated as biologically relevant in cell-based assays, wherein mutation of the dimer interface prevents autophosphorylation (Zhang et al., 2006). The allosteric activation of the EGFR kinase domain by an interaction at an N-lobe allosteric site is similar to activation mechanisms utilized by other kinases, most notably the activation of cyclin dependent kinases by cyclins (Jeffrey et al., 1995; Jura et al., 2011). The kinase domains of all members of the EGFR family dimerize by this mechanism, with Her3 restricted to serving as an activator due to its impaired catalytic activity (Jura et al., 2009b; Monsey et al., 2010).

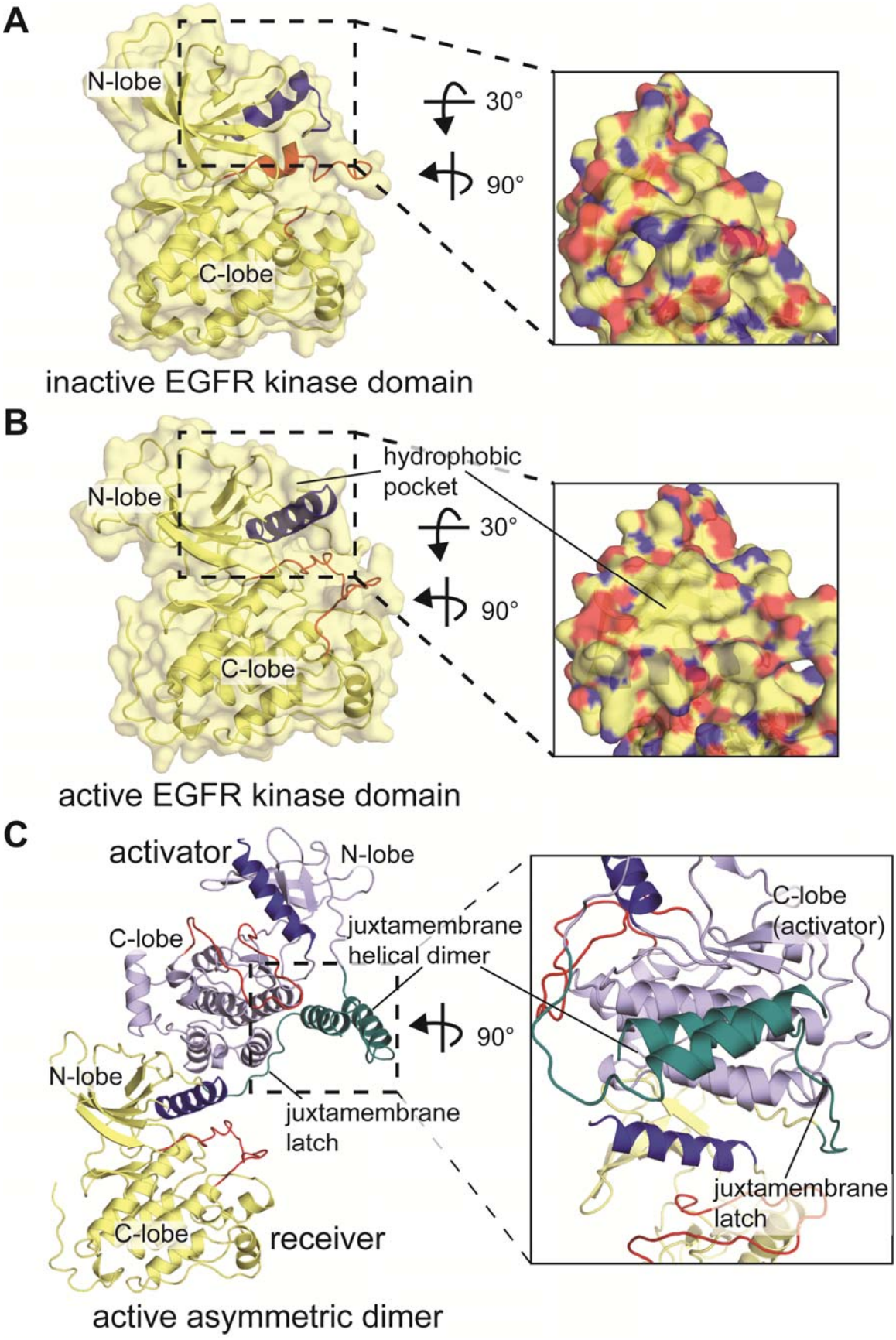

A
N-lobe
C-lobe
30°
90°
inactive EGFR kinase domain
B
hydrophobic
pocket
N-lobe
C-lobe
30°
90°
active EGFR kinase domain
C
activator
N-lobe
C-lobe
juxtamembrane
helical dimer
90°
C-lobe
(activator)
N-lobe
juxtamembrane
latch
juxtamembrane
latch
C-lobe
receiver
active asymmetric dimer

Figure 1.3 Structural features of the inactive and active EGFR kinase domains and the asymmetric dimer formed during kinase activation. Conformational changes between the (A) inactive EGFR kinase domain and (B) active EGFR kinase domain occur primarily in the N-lobe. In the inactive state, helix-αC is rotated away from the kinase domain, and the kinase domain adopts the Src/CDK-like inactive conformation. The repositioning of helix-αC towards the core of the N-lobe upon activation exposes a hydrophobic pocket on the surface of the N-lobe, shown in the surface representation on the right, which is not observed in the inactive conformation (Stamos et al., 2002; Wood et al., 2004; Zhang et al., 2006). (C) The conformational changes in the active structure are stabilized by formation of an asymmetric dimer of kinase domains. The C-lobe of one subunit, the activator (light blue), interacts with the exposed hydrophobic surface on the N-lobe of the second subunit, the receiver (yellow), and the receiver undergoes allosteric activation (Zhang et al., 2006). The juxtamembrane segment (teal), which is N-terminal to the kinase domain, stabilizes the formation of the asymmetric dimer through two sets of interactions. The N-terminal portion of the juxtamembrane segments from the activator and receiver kinases are modeled to form a helical dimer, while the C-terminal portion of the juxtamembrane segment from the receiver kinase docks onto the C-lobe of the activator kinase in an interaction known as the juxtamembrane latch (Jura et al., 2009a; Red Brewer et al., 2009). PDB codes for the structures shown here are: Inactive – 2GS7, Active – 2GS6, Asymmetric Dimer – 2GS6 (kinase domain) and 3GOP (juxtamembrane latch) (Red Brewer et al., 2009; Zhang et al., 2006). The juxtamembrane helical dimer has not been observed crystallographically and was modeled. Helix-αC is colored deep blue, and the activation loop is colored red in all structures to assist with orientation.

The tendency for the isolated kinase domain to form the asymmetric dimer is very low, and EGFR is monomeric *in vitro* at concentrations of at least 50 μM, as detected by light scattering. However, tethering of the kinase domain to the surface of small unilamellar vesicles restricts the kinase domain to two dimensions and thus increases the local concentration, promoting dimerization and activation (Zhang et al., 2006). Additionally, inclusion of the juxtamembrane residues N-terminal to the kinase domain within the recombinant kinase construct dramatically increases the affinity of the asymmetric dimer in solution by two mechanisms. First, the C-terminal half of the juxtamembrane region (residues 664-685) of the receiver kinase docks onto the C-lobe of the activator kinase, forming the asymmetric dimer-stabilizing interaction known as the juxtamembrane latch (Jura et al., 2009a; Red Brewer et al., 2009). Second, the N-terminal halves of the juxtamembrane segment (residues 645-663) from both proteins in the dimer form a helical coil (Figure 1.3C). The juxtamembrane latch interaction and helical dimer increase the asymmetric dimer affinity to values of the dissociation constant, K_D, of approximately 8 μM and 200 nM, respectively. These values were estimated by measuring changes in kinase specific activity that track the transition from the relatively inactive monomer to the active dimer (Jura et al., 2009a). In the cell the full-length receptor includes the juxtamembrane segment and is also concentrated on the plasma membrane, necessitating EGFR to employ autoinhibitory mechanisms to

prevent aberrant dimerization, activation, and signaling. These autoinhibitory mechanisms are not as well-characterized for the kinase domain.

1.7 Expanding our understanding of the mechanisms controlling EGFR kinase activation

While the structure and function studies of the EGFR kinase domain over the last decade have illuminated the mechanisms that regulate kinase activation, the functional origins of several changes to the primary sequence of EGFR that alter kinase activity remain unexplained and indicate that our understanding of EGFR regulation is incomplete. Alterations to the EGFR sequence that are observed in cancers in which residues 958-1030 or 958-1043 of the EGFR C-terminal tail are missing activate the kinase, while experiments in which the C-terminal tail sequence is modified result in either an increase or reduction in kinase activity (Bublil et al., 2010; Jura et al., 2009a; Pines et al., 2010; Zhang et al., 2006). This suggests a function for the C-terminal tail in regulating kinase activity that is likely to be exerted through direct interactions with the kinase domain. One role for the C-terminal tail in inhibition of kinase activity is suggested by the multiple crystal structures of a symmetric inactive kinase domain dimer, in which the C-terminal tail forms extensive interactions with the kinase domain both in *cis* and *trans* at the dimer interface (Jura et al., 2009a). However, the biological relevance of the interactions observed in the inactive dimer structure has not been resolved, nor do the interactions fully explain why changes to the C-terminal tail alter kinase activity.

As mentioned previously, the conformation of the N-lobe largely determines the activation state of the EGFR kinase domain. Recent molecular dynamics simulations have depicted the kinase N-lobe as a highly dynamic structure that accesses many conformations, of which the vast majority are inactive, as opposed to the two rigid states observed in the inactive and active crystal structures (Shan et al., 2012). Therefore, mechanisms to energetically bias the N-lobe conformation toward an active state should be sufficient to increase the basal activity of the kinase. Indeed, the L834R mutation that is observed in cancers activates EGFR by increasing the sampling of the active conformation (Shan et al., 2012; Yun et al., 2007). Likewise, it is then plausible that there remain uncharacterized sites on the kinase domain that could serve as points to regulate the conformation of the N-lobe and thus kinase activation.

One strategy to probe unknown sites of kinase regulation is to identify molecules that modulate EGFR activity and study their mechanisms. An additional benefit of this work is the therapeutic potential of such molecules, as novel methods to control the activity of the EGFR family proteins are desirable for the cancers that are driven in part by the aberrant activation of these kinases. The second chapter of this dissertation presents the identification of small molecule activators of the EGFR kinase domain and the characterization of the effects of these compounds on EGFR kinase domain activation and dimerization. In the third chapter, I present a mechanistic study of a peptide that is phosphrylated by the EGFR kinase domain with very high catalytic

efficiency and show that this peptide increases the catalytic activity of EGFR by inducing kinase domain dimerization through aggregation.

Chapter 2

Discovery of small molecule activators of the EGFR kinase domain

2.1 Overview

The EGFR kinase domain undergoes allosteric activation upon formation of an asymmetric dimer, wherein the C-lobe of one subunit, the activator, binds the N-lobe of the second subunit, the receiver, to induce and stabilize the active conformation of the N-lobe of the receiver (Zhang et al., 2006). The mechanism of allosteric activation via binding of a protein effector at the N-lobe is similarly employed by multiple other kinases (Jura et al., 2011). In the case of the AGC family of kinases, intra- or intermolecular binding events at an N-lobe allosteric site on the kinase, which is known as the PIF-pocket, regulate kinase activity by either activating or inhibiting the kinase. For example, the sequences of substrates of the AGC kinase PDK1 contain a PIF-pocket binding motif, and the binding of substrates stimulates phosphorylation (Biondi et al., 2000; 2002).

Biondi and coworkers hypothesized that a small molecule could bind the PIF-pocket to allosterically activate PDK1, similar to the mechanism employed by substrates of the kinase. Screening of a virtual compound library successfully identified small molecules that bind the PDK1 PIF-pocket and activate PDK1 (Engel et al., 2006). Multiple subsequent studies have yielded more potent activating compounds as well as compounds that bind the PIF-pocket but inhibit PDK1 activity (Hindie et al., 2009; Sadowsky et al., 2011; Wei et al., 2010). In cells these compounds selectively inhibit phosphorylation of substrate proteins that dock to the PIF-pocket, indicating the compounds disrupt a protein-protein interaction (Busschots et al., 2012).

Similar discoveries of compounds that bind the allosteric N-lobe site have been made for other kinases. A small molecule was identified to bind near helix-αC of the cyclin dependent kinase CDK2. The bound compound repositions helix-αC into a conformation that is incompatible with binding of the allosteric activator of CDK2, cyclin A, and thus the compound serves as an allosteric inhibitor of cyclin-mediated activation of CDK2 (Betzi et al., 2011). In another example, intramolecular binding of a regulatory domain to the PIF-pocket of the protein kinase C isoform PKCζ inhibits kinase activity. A compound was discovered that binds the PKCζ PIF-pocket and inhibits activity by substituting for the effects of its regulatory domain (Lopez-Garcia et al., 2011).

These examples demonstrate the different capacities for compounds that bind the allosteric N-lobe site of kinases. The compounds may substitute for the natural effects of the proteins that bind to the site to either activate the kinase, in the case of PDK1, or inhibit the kinase, in the case of PKCζ (Engel et al., 2006; Hindie et al., 2009; Lopez-Garcia et al., 2011; Sadowsky et al., 2011; Wei et al., 2010). Alternatively, the compounds may bind to the site and prevent activation by the natural protein activators of the kinases, as was found with the allosteric inhibitors of PDK1 and CDK2 (Betzi et al., 2011; Sadowsky et al., 2011). Overall, these studies suggest the N-lobe allosteric site as a relatively unique and potentially druggable target in kinases that are allosterically regulated by interactions there. Traditionally, kinase inhibitors have targeted the active site, which is highly conserved and makes specificity relatively more difficult to achieve (Balzano et al., 2011). Discovery of compounds that bind the

allosteric sites of kinases may circumvent this problem by regulating kinase activity at an alternative location to the active site.

This section of the dissertation focuses on the identification and characterization of small molecules that activate the EGFR kinase domain with the goal that such compounds might bind the allosteric site on the N-lobe, akin to the work on PDK1 and other kinases. The EGFR kinase domain is predominantly monomeric *in vitro* and consequently bears very low catalytic activity, which facilitates a screen for compounds that increase activity (Zhang et al., 2006). We discovered five compounds that activate the EGFR kinase domain in a small molecule screen and characterized these compounds in subsequent studies. We anticipate that these small molecules might stabilize the active conformation of the kinase by binding to the hydrophobic pocket of the N-lobe that is exposed upon activation, substituting for the role of the activator kinase subunit in the asymmetric dimer (Figure 2.1A).

Allosteric activators may significantly impact function of the full-length EGFR because compounds that bind the N-lobe should also block kinase domain dimerization within the full-length protein in cells and consequently inhibit autophosphorylation (Figure 2.1B). The EGFR feedback inhibitor mitogen-inducible gene 6 (MIG6) has been shown previously to block EGFR signaling through a similar mechanism, in which binding of MIG6 to the EGFR kinase domain C-lobe inhibits formation of the asymmetric dimer and signaling (Zhang et al., 2007). The concept of a small molecule that allosterically activates the kinase but inhibits autophosphorylation is counterintuitive. We expect that small molecules that activate the EGFR kinase domain and block dimerization would mimic the effects of the combination of the EGFR mutations V924R and L834R. Val-924 is on the C-lobe at the asymmetric dimer interface, and mutation to arginine blocks kinase domain dimerization (Zhang et al., 2006). The L834R mutation, which naturally arises in some cancers, is localized within the kinase activation loop and increases basal kinase activity (Yun et al., 2007). Full-length EGFR in cells bearing both mutations is unable to undergo autophosphorylation despite the activating effects of the L834R mutation, underscoring the necessity of the kinase domain asymmetric dimer in order for autophosphorylation to occur (Zhang et al., 2007). We hope that these compounds are able to block EGFR dimerization and autophosphorylation and could therefore serve as a novel class of EGFR inhibitors.

The mechanisms by which the five compounds discovered in these studies activate EGFR and the capacities of these compounds to inhibit EGFR dimerization are not fully resolved, despite extensive work towards biochemical and structural characterization. However, preliminary evidence suggests one of the discovered compounds blocks formation of the asymmetric dimer. Furthermore, compounds that allosterically activate the EGFR kinase need not bind exclusively at the N-lobe but may bind at any site that might bias the kinase domain towards the active conformation. Regardless of the specific mechanisms of the discovered EGFR activators, all of the compounds should have some capacity to directly or indirectly affect kinase domain dimerization since activation and formation of the asymmetric dimer are coupled. For example the EGFR kinase domain mutation L834R increases kinase activity by

destabilizing the inactive conformation, which causes dimerization of the kinase domain (Shan et al., 2012). Ultimately, the compounds discovered and discussed here may serve as tools for studying regulation of EGFR kinase activity and are potentially effectors of dimerization.

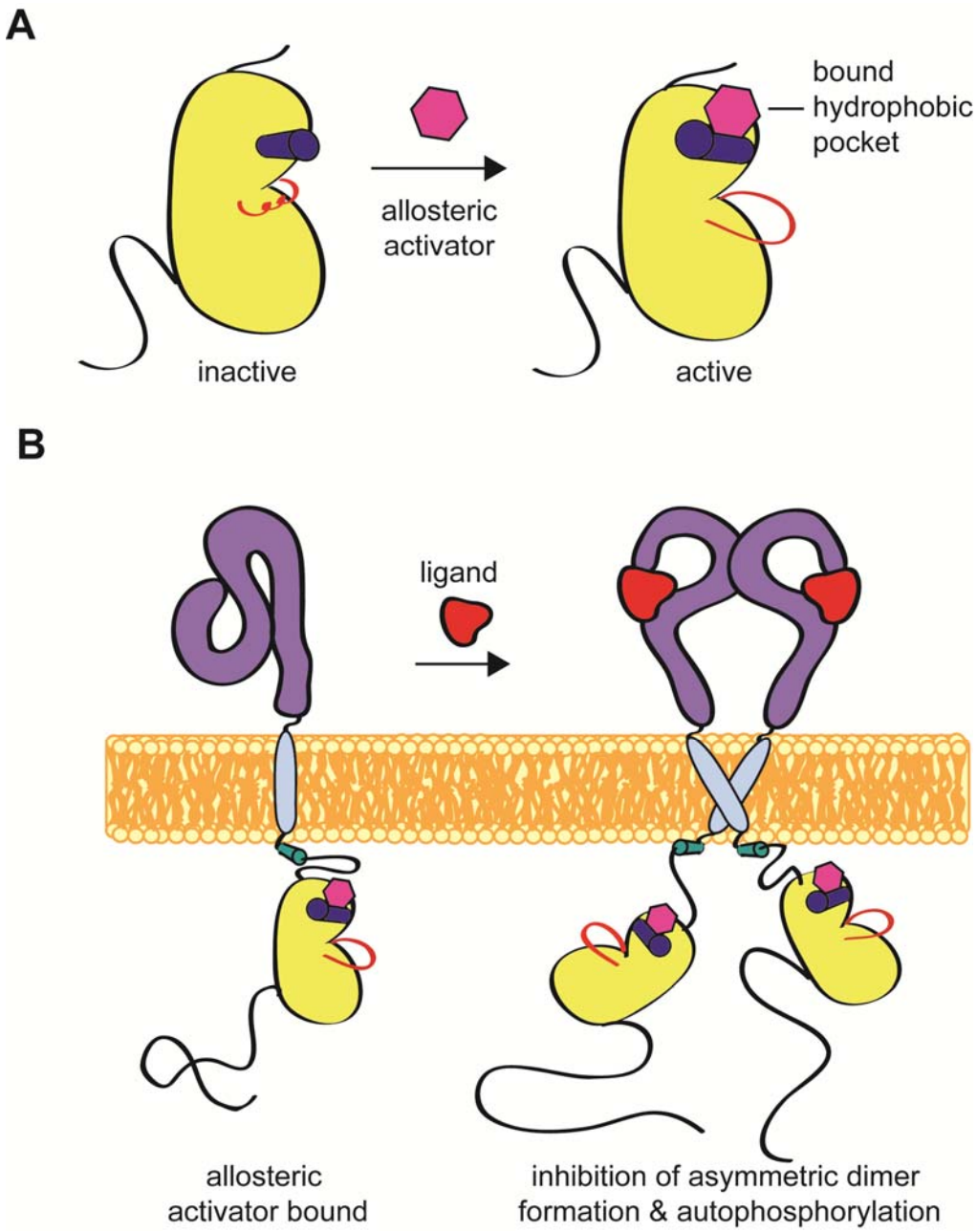


Figure 2.1 The expected properties of a small molecule activator of the EGFR kinase domain *in vitro* and in cells. (A) The kinase domain is predominantly monomeric *in vitro* and does not readily undergo allosteric activation through formation of the asymmetric dimer of kinase domains (Zhang et al., 2006). A small molecule could bind to the hydrophobic pocket on the N-lobe to allosterically activate the kinase, assuming the role of the activator subunit within the asymmetric dimer. (B) Full-length EGFR in cells that is bound by a compound that occupies the N-lobe allosteric site should not be able to form the asymmetric dimer and thus should not undergo autophosphorylation. Mutagenesis data suggest that the ability to form the asymmetric dimer is required for autophosphorylation of EGFR, regardless of the activation state of the receptor (Zhang et al., 2007).

2.2 Results and Discussion

2.2.1 Selection of the ADP Hunter screen to identify activators of the EGFR kinase domain

High throughput screens (HTS) are subject to stringent requirements, and numerous qualifications must be addressed when developing a HTS assay. The following questions are representative of the considerations that must be made during selection and optimization of the assay: Is the assay consistently reproducible? Can the assay be miniaturized in the 384-well plate format? What is the potential for compound interference with the readout of the assay? Are the steps of the assay compatible with automation? Is the assay dependent on kinetic measurements or on quantification of an end-point concentration? Does the assay have a robust signal window, as measured by the statistical variation described by the *Z'*-factor, which is described below (Zhang et al., 1999)?

Following evaluation of multiple assay formats, the commercial end-point assay known as ADP Hunter from the company DiscoveRx was selected to screen for changes in EGFR activity. The ADP Hunter screen employs a series of linked enzymatic reactions to couple ADP production, a byproduct of the EGFR kinase reaction, to production of the Resorufin fluorophore. A schematization of the assay is presented in Figure 2.2A, and a standard curve for the fluorescent assay output generated by titration of ADP is shown in Figure 2.2B.

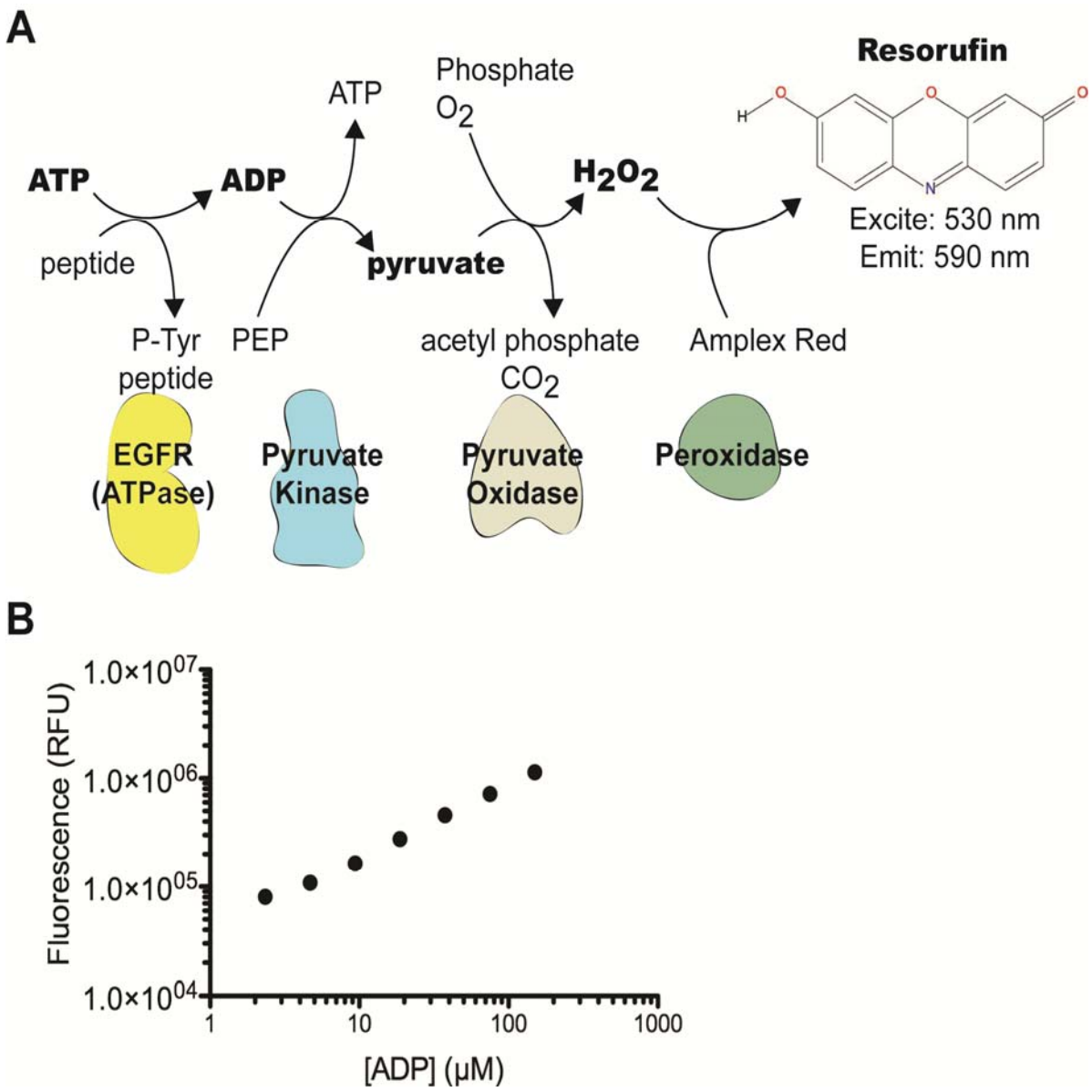


Figure 2.2 Detection of EGFR activity using the ADP Hunter assay. (A) A schematization of the ADP Hunter assay depicts how ADP production as a byproduct of the kinase reaction is linked to the fluorescent product Resorufin. Relevant reaction substrates and products are in bold. The phosphorylation activity of EGFR creates ADP, which then serves as a substrate in the pyruvate kinase-catalyzed reaction. The pyruvate produced by pyruvate kinase becomes a substrate in the pyruvate oxidase-catalyzed reaction, which produces hydrogen peroxide. A perixodase catalyzes the production of Resorufin from the precursor Amplex Red and hydrogen peroxide. (B) A standard curve generated for the ADP Hunter assay indicates the Resorufin fluorescence output as a function of ADP concentration in solution. The curve was generated through titration of ADP in the ADP Hunter reaction buffer followed by the ADP Hunter detection reaction. Data points are the means of four replicates, and error bars plotted as ± standard deviations of mean are too small to be visible on this graph.

One major advantage of the ADP Hunter assay was the flexibility in EGFR peptide substrate selection and concentration of the peptide substrate in the assay. Because the screen sought to identify activators of kinase activity, substrate concentrations should be saturating to prevent limitations on the signal output from substrate depletion or partial occupancy of the EGFR active site. Peptide substrates of EGFR (discussed extensively in Chapter 3) typically exhibit low millilmolar values for the Michalis constant K_M (Fan et al., 2005; Zhang et al., 2006). The ADP Hunter assay does not place restrictions on the peptide substrate and allows use of saturating concentrations of a generic tyrosine kinase peptide substrate. In contrast other assay formats require relatively low concentrations of substrates that are often proprietary, which would cause the catalytic rate of EGFR to be limited by substrate concentration and make activation of EGFR more difficult to detect.

Furthermore, the readout process for the ADP Hunter assay suggested there should be a relatively low prevalence of interference and false positive hits. The red-shifted fluorescence properties of the assay output Resorufin, which has an excitation maximum at 530 nm and emission maximum at 590 nm, decreases the likelihood of spectral interference from the compounds in the screening library, which typically have fluorescence profiles at shorter wavelengths (Simeonov et al., 2008). Furthermore, we believed that the coupled-enzyme basis of the assay itself should limit the occurrence of false positives, as production of ADP should be required initially for the series of enzymatic reactions that ultimately convert Amplex Red to Resorufin. One drawback of a coupled-enzyme assay is the potential of the compounds to inhibit any of the enzymes that are required for the readout, which would reduce the fluorescent signal. However, since the assay was screening for EGFR activation and not inhibition, we reasoned that the likelihood of a false negative (i.e. a compound that activates EGFR but inhibits the coupled-enzyme readout) was very low.

2.2.2 Optimal parameters for screening EGFR activity

The ultimate goal in development of a screen for EGFR kinase domain activation was to determine assay conditions that create a large and robust signal window for EGFR activity from which hit compounds could be identified readily. The calculation usually used to determine the suitability of an assay for HTS is the *Z'*-factor measurement, as given by:

$$Z' = 1 - \frac{(3\sigma_{c+} + 3\sigma_{c-})}{|\mu_{c+} - \mu_{c-}|}$$

Here, μ_c and σ_c represent the mean and standard deviation, respectively, for the positive (+ subscript) and negative (- subscript) controls of the assay. This measurement accounts for the dynamic signal window and the inherent variation of the assay. An assay with a *Z'*-factor value of at least 0.5 exhibits acceptable performance for HTS (Zhang et al., 1999).

Optimal screening conditions should support reliable identification of allosteric activators of the EGFR kinase domain. We used peptide substrates that are phosphorylated by EGFR with varying efficiencies as negative and positive controls for EGFR activity in the assay. Measurement of basal EGFR activity with a generic tyrosine kinase peptide substrate served as the negative control. A signal for basal activity of EGFR that was low but significantly above the background signal of the assay was desirable and indicated the assay functioned within the dynamic range of EGFR activity. We measured a loose upper bound for EGFR activation by addition of a particularly efficient EGFR substrate, known as Peptide C, as a positive control for EGFR activity because catalysis with Peptide C occurs at a rate that is approximately ten-fold faster than conventional peptide substrates. The work in Chapter 3 characterizes the mechanism by which the high catalytic efficiency of EGFR with Peptide C is achieved, which was unknown at the time of these HTS studies.

The workflow for the ADP Hunter assay used to quantify EGFR activity is depicted in Figure 2.3A. The entire assay is executed in a 384-well low-volume plate. Following performance of the EGFR kinase reaction, proprietary detection Reagents A and B of the ADP Hunter kit are added individually, which terminates the kinase reaction. The detection reaction is developed for one hour, and the Stop solution is added, which stabilizes the fluorescent signal. While the EGFR reaction parameters required extensive optimization, the performance of the ADP Hunter components within the guidelines provided by the manufacturer was satisfactory and did not require additional development. The entire assay performed well with the reaction buffer supplied with the kit (15 mM HEPES, pH 7.4, 10 mM $MgCl_2$, 20 mM NaCl, 0.1 mg/ml Bovine γ Globulin, 1 mM EGTA and 0.02% v/v Polysorbate-20), and so buffer conditions were not optimized further.

Substrate titration assays were performed to determine the concentrations of peptide substrates and ATP to use in the screen. The generic tyrosine kinase substrate Poly 4Glu:Tyr, a random co-polyamino acid with a molecular weight range of 5,000-20,000 Da that consists of a ratio of four glutamates per one tyrosine, was selected as the peptide substrate for EGFR for negative-control and experimental wells. While typically a more natural and homogeneous peptide substrate would be desirable, the cost to supply sufficient amounts of such peptides would be prohibitive, and so Poly 4Glu:Tyr was the best substrate option due to its commercial availability. The Poly 4Glu:Tyr concentration selected for screening was 2 mg/ml, which is saturating of EGFR activity and ensured that substrate depletion would not be a limiting factor for kinase activity. Similarly, the concentration of Peptide C, the positive control, was selected to be 500 μM, which is greather than 16-fold higher than the measured K_M value of 30 μM. However, the supply of Peptide C was limited as it required custom synthesis, and we could not assay kinase activity with the peptide in greater excess. The ATP concentration was selected to be 200 μM, which was the maximum limit of the ADP Hunter assay and should prevent substrate depletion.

Given conditions in which substrate concentrations are saturating, the main factors influencing the signal window for EGFR activity are EGFR kinase domain

concentration and the length of the kinase reaction. Recombinant EGFR kinase domain (residues 672-998) is expressed in *Spodoptera frugiperda* Sf9 cells, and purification yields approximately 1 mg/L of cells. Due to our production limits, the highest EGFR concentration that could be feasibly used was 100 nM. EGFR activity was measured over time, and at one hour, the signal from the positive-control wells containing Peptide C became maximized, likely due to substrate depletion, while the signal from the negative-control wells containing Poly 4Glu:Tyr continued to increase. This time point was determined to have the largest signal window between the positive and negative controls, and the quality of the signal window was confirmed by its *Z'*-factor value of 0.63. Additionally, the basal activity of EGFR in the negative-control samples was low but significantly above the assay background, indicating that the assay functioned within the dynamic range for EGFR activity. The collective results of these optimized reaction parameters are shown in Figure 2.3B.

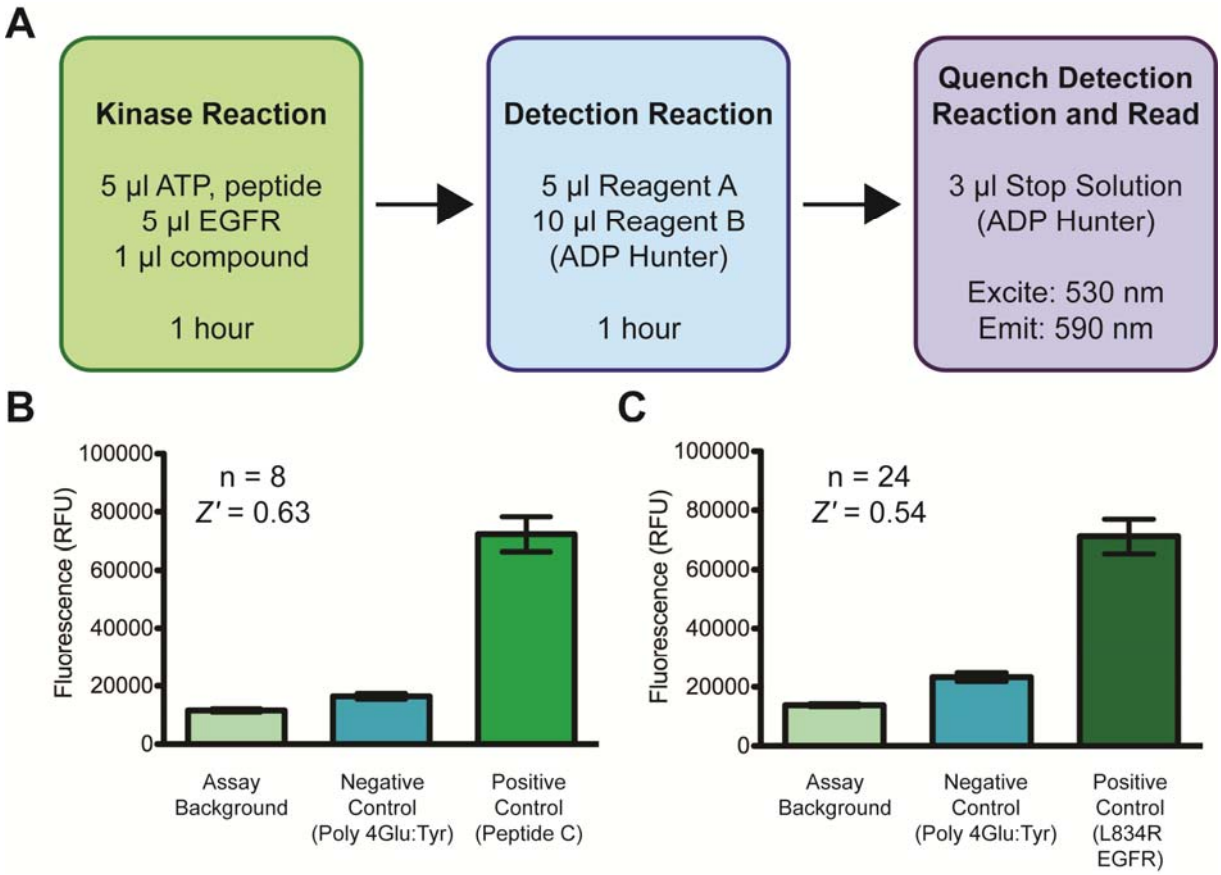


Figure 2.3 The workflow for the EGFR kinase reaction and ADP Hunter assay and results of optimized assay conditions. (A) The EGFR kinase reaction and ADP Hunter reaction to detect kinase activity are executed in three steps. Five µl EGFR kinase domain, 5 µl ATP and peptide substrate mix, and 1 µl compound or 11% v/v DMSO are mixed and incubated for one hour at room temperature. The proprietary detection reagents of the ADP Hunter kit Reagent A (5 µl) and Reagent B (10 µl) are added sequentially and incubated for one hour at room temperature. The ADP Hunter Stop solution (3 µl) is added to stabilize the fluorescent signal, and fluorescence is detected by excitation at 530 nm and measuring the emission at 590 nm. (B) The optimized reaction conditions create a robust assay with a large signal window between unactivated (negative control) and activated (positive control) EGFR samples. This assay was smaller scale, with eight wells per condition, and set up manually. (C) The optimized reaction conditions are executed on a larger scale, with 24 wells per condition, and set up using the automation equipment at the Small Molecule Discovery Center. The Z' values for both assays indicate that there is a statistically significant signal window that should allow identification of compounds that activate the EGFR kinase domain. 100 nM EGFR and 200 µM ATP were incubated with either 2 mg/ml Poly 4Glu:Tyr (negative control) or 500 µM Peptide C (positive control) in the ADP Hunter assay buffer for one hour at room temperature. In the case of (C), the positive control was the high activity EGFR mutant L834R with 2 mg/ml Poly 4Glu:Tyr, and 1% v/v DMSO was included. The activity of EGFR was detected using the ADP Hunter assay kit. The values plotted are means ± standard deviations of mean from either eight or 24 replicates.

The final step in HTS assay development is determining the DMSO tolerance of the assay, which is the highest concentration of DMSO that minimally impacts assay performance. The DMSO tolerance ultimately determines the concentration at which compounds can be delivered, making the highest DMSO concentration possible desirable. The DMSO tolerance of the EGFR activation assay was measured to be 1% v/v, which corresponds to a compound final assay concentration of 10 μM.

To guarantee that the selected assay parameters were indeed optimal, measurements of the assay background, negative control, and positive controls were performed in 24 replicates each using the liquid handling robotics at the Small Molecule Discovery Center screening facility at the University of California, San Francisco (Figure 2.3C). In contrast the previous assays had been performed manually using standard laboratory equipment. Notably, the positive control in this assay was not EGFR with Peptide C as a substrate but instead the L834R variant of the EGFR kinase domain with the peptide substrate Poly 4Glu:Tyr because supplies of Peptide C were limited (L834R is a mutation that increases the basal activity of EGFR (Yun et al., 2007).) Because Peptide C and L834R increase EGFR activity to a similar level, the results with the L834R variant positive control were expected to be transferable to assays employing Peptide C as the positive control. This larger-scale assay verified that the HTS assay development goals were attained. First, the signal for basal EGFR activity was ~1.7-fold higher than the background assay signal, indicating that a low but significant level of EGFR activity was detected. Second, the *Z'*-factor of the assay was calculated to be 0.54, which confirmed that the span and variance of the assay signal window should be robust to detect compounds that activate the EGFR kinase domain. Finally, execution of the assay using the liquid handling robotics indicated that the assay performed satisfactorily when automated.

2.2.3 Identification of small molecule activators of the EGFR kinase domain by a high throughput screen

A library of over 103,000 compounds at the Small Molecule Discovery Center was screened for the ability to activate the EGFR kinase domain using the optimized protocol described earlier. The compound library consisted of commercially available compounds that represent a diversity of chemical space and had generally drug-like properties, as predicted by Lipinski's Rule of 5 (Lipinski et al., 1997). A 384-well plate consists of 16 rows labeled A-P and 24 columns labeled 1-24. The wells in columns 3-22 of the 384-well plate were the experimental wells, and each contained a unique compound. The 1, 2, 23 and 24 columns all contained negative-control wells with the exception of the A, B, O, and P wells of columns 1 and 24, which were positive-control wells. In typical HTS campaigns, more positive-control wells would be included on each plate, but the supply of the positive control Peptide C was limited. The inclusion of both negative- and positive-control wells on each plate allowed the performance of each plate to be monitored through calculation of *Z'*-factors.

A set of sixteen plates was screened at a time typically, and selection of hit compounds was performed for each set of plates to circumvent issues with day-to-day

assay variation. For every plate, the value of fold-activation for each compound was calculated by dividing the fluorescence value for the well containing the compound by the mean fluorescence value of the negative-control wells. The mean and standard deviation of all of the fold-activation values for the compound-containing experimental wells were calculated for each set of plates. Experimental wells for which the fold-activation values were greater than three standard deviations above the mean were scored as hits. On this basis, a total of 172 hits were identified from the approximately 103,000 compounds screened, which corresponds to a hit rate of about 0.17%. The screening data is grossly represented by scatter plot in Figure 2.4. The day-to-day variation in measurements is evident in the deviation of the scatter of points progressing across the X-axis.

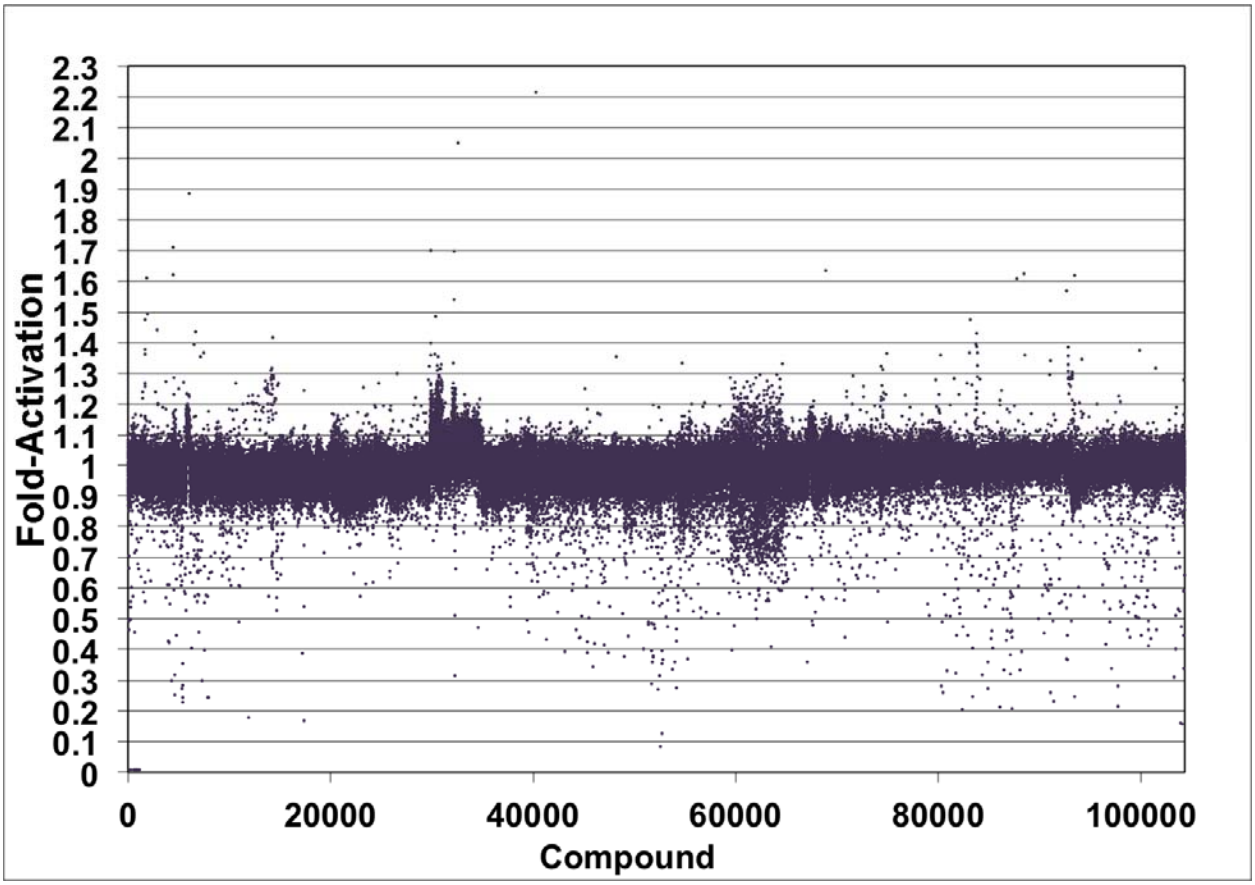


Figure 2.4 Scatter plot of the results of the screen for small molecule activators of the EGFR kinase domain. Over 103,000 compounds of the Small Molecule Discovery Center compound library were assayed for effects on EGFR activity by performance of the kinase assay under the optimized reaction conditions and detection of EGFR activity using the ADP Hunter assay. Typically, plates were screened in sets of 16, and fold-activation values were calculated by dividing the fluorescence value for each compound by the mean fluorescence value of the negative-control wells for each plate. Compounds with fold-activation values that were greater than three standard deviations above the mean fold-activation of each set of plates were scored as hits, and hit selection was performed for each set of plates. Based on this hit selection process, 172 hits were identified.

2.2.4 Initial validation of the hit compounds

The capacities of the 172 hit compounds to increase catalytic activity of the EGFR kinase domain required additional validation in order to select compounds to pursue further. Samples of the hit compounds were obtained from the stocks of the Small Molecule Discovery Center compound library, and dose response curves for each compound were generated using the ADP Hunter assay. Fluorescence increased in a dose-dependent manner for 42 of the 172 hit compounds, indicating that the assay

format and hit selection criteria successfully identified compounds that increased the assay signal. The remaining 130 compounds were false positives, likely due to the natural signal variation of the assay.

The 42 compounds with measurable EC_{50} values (compound concentrations at which 50% of maximal fold-activation by the compound is achieved) were evaluated for binding to the EGFR kinase domain in a surface plasmon resonance (SPR) assay, which was developed and performed by our collaborator Preeti Chugha of the Small Molecule Discovery Center. In this assay, the EGFR kinase domain was biotinylated and immobilized on a neutravidin-coated SPR chip. The kinase inhibitor AG1478, which is expected to bind the active conformation of EGFR, was bound to EGFR to prevent other compounds from binding to the active site and to increase the accessibility of the active conformation (Levitzki and Gazit, 1995; Liu and Gray, 2006). A titration of each of the compounds was flowed over the chip, and binding was quantified by fitting of the response data to hyperbolic binding curves. The binding curves were rough approximations, though, as the maximum concentration of the compounds in the assay was 125 μM and therefore below the levels required for saturation. Seventeen of these compounds were shown to have varying affinities for EGFR, with two estimated to have values of the dissociation constant, K_D, lower than 10 μM. However, five of the apparent K_D values were greater than 500 μM, and 25 of the compounds failed to bind, suggesting that the vast majority of hit compounds do not bind EGFR.

The quality of the data and the performances of the 42 compounds in the activity dose response assay and SPR binding assay were evaluated to select compounds for more-detailed analysis. The curves generated from the activation and binding data were considered to be higher quality if they reached saturation with low apparent cooperativity (as judged by the Hill slope), and attention was also paid to compounds with similar values for K_D and EC_{50}. Compounds with superstoichiometric binding were considered to bind non-specifically or have a propensity to aggregate. Based on these criteria, twenty compounds were selected for additional studies.

2.2.5 Five of the hit compounds are true activators of the EGFR kinase domain

Compounds that truly activate the EGFR kinase should consistently increase catalytic activity in orthogonal kinase assays. Analysis of the twenty compounds that were selected for further evaluation in a coupled-enzyme ATPase assay indicated that at least some of the compounds were false positives for EGFR activation. This assay couples ADP production to NADH oxidation, which produces a change in absorbance at 340 nm that can be read kinetically (Barker et al., 1995). One hit compound, which is a quinone, was found to catalyze NADH oxidation directly and demonstrated that compounds may produce false positives with coupled-enzyme assays in unprecedented ways (Scherbak et al., 2005).

To test the requirement for EGFR in generation of the assay signal and thereby eliminate false positives, the signal generated by the compounds with and without EGFR present was measured in the ADP Hunter assay. Twelve of the twenty selected

compounds increase the assay signal in a dose-dependent manner independent of the presence of EGFR. The other eight compounds required the EGFR kinase domain and increased EGFR activity with varying efficacy (Figure 2.5A-B). False positives are inherent to every HTS campaign. In the case of the ADP Hunter assay, the false positives were likely generated either by compounds with similar fluorescent properties to Resorufin or high reduction potentials, such as the quinone, that can catalyze hydrogen peroxide production (which is coupled to Resorufin production) from oxygen.

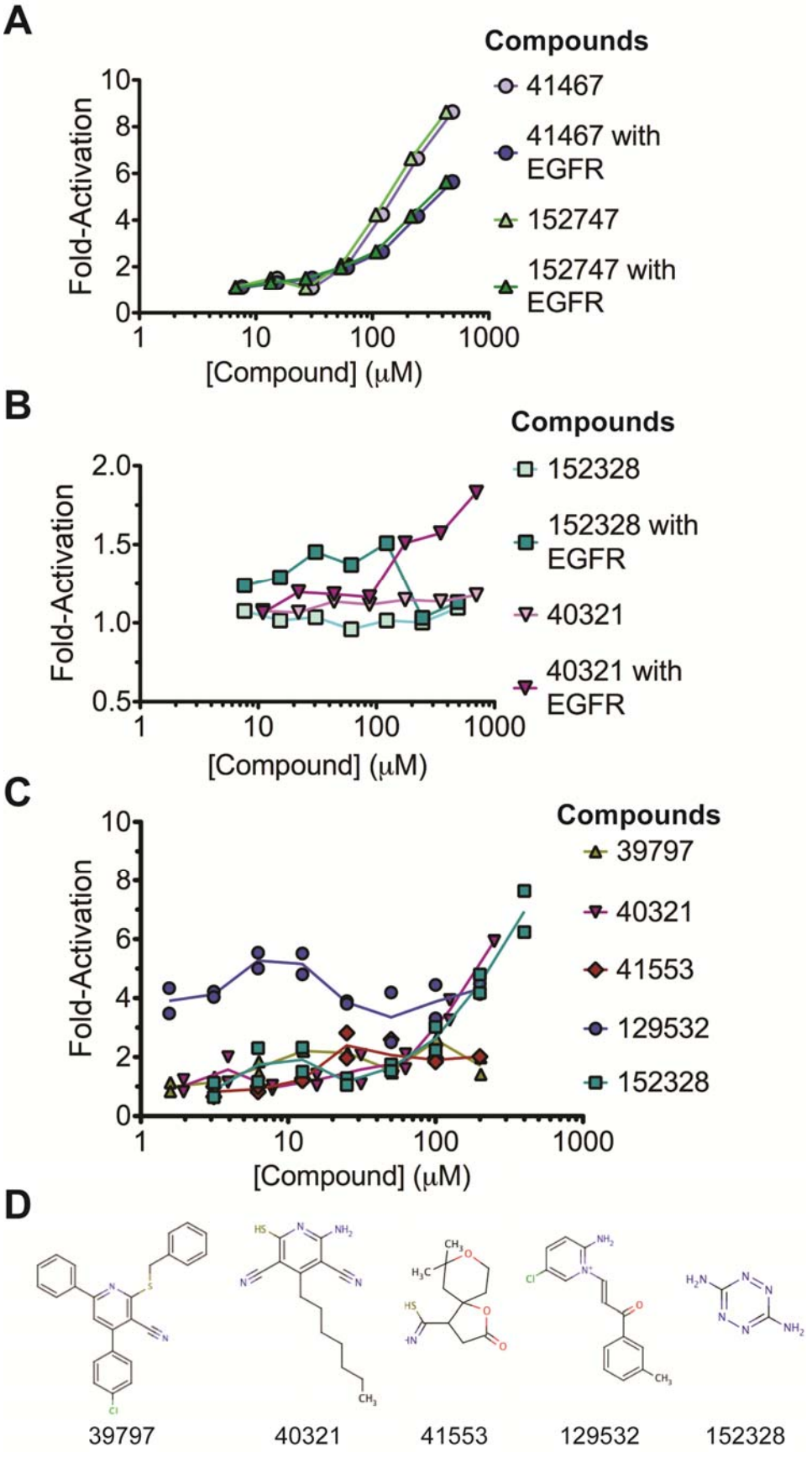
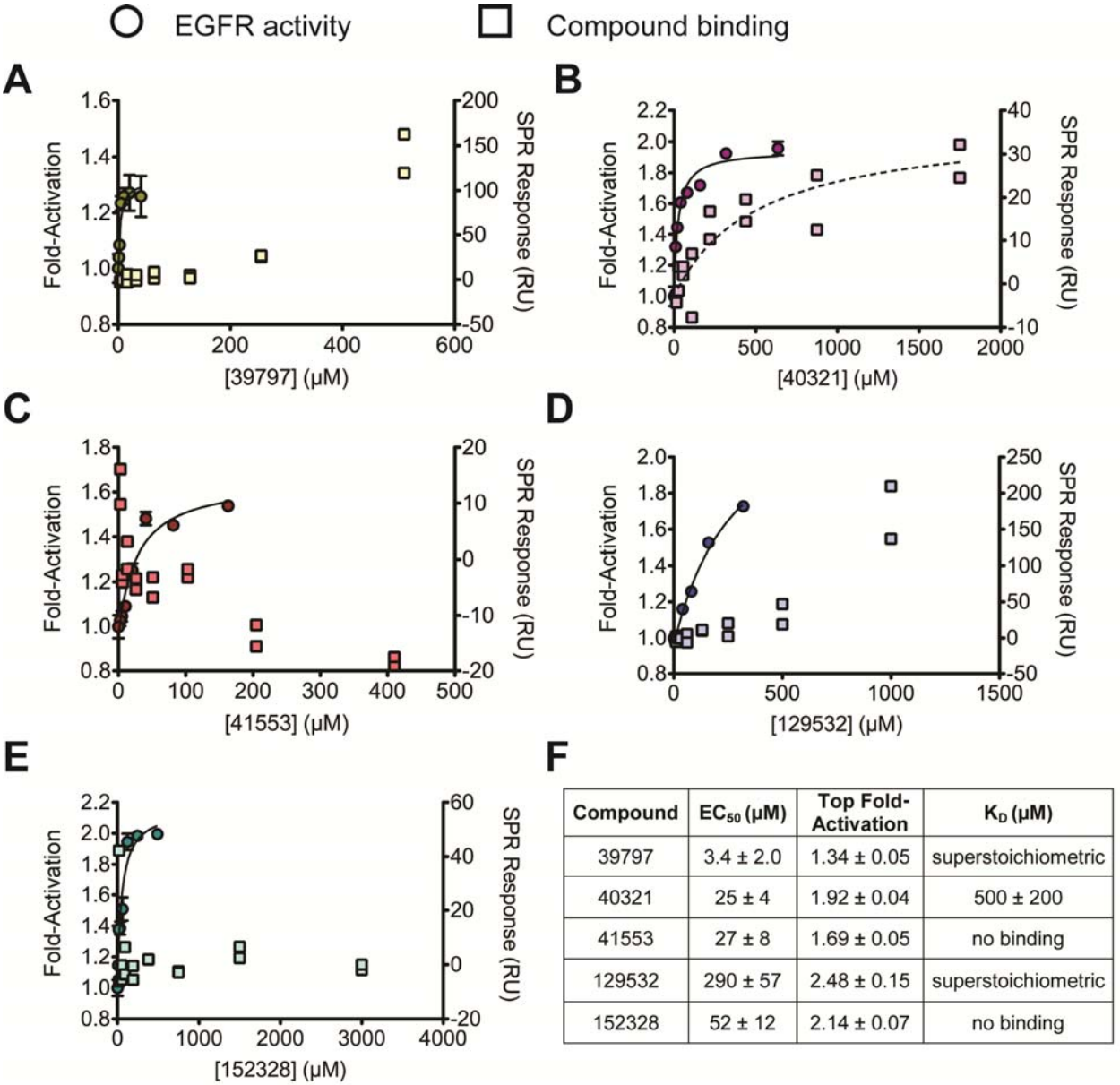

A
Compounds
41467
41467 with EGFR
152747
152747 with EGFR
Fold-Activation
[Compound] (μM)
B
Compounds
152328
152328 with EGFR
40321
40321 with EGFR
Fold-Activation
[Compound] (μM)
C
Compounds
39797
40321
41553
129532
152328
Fold-Activation
[Compound] (μM)
D
39797
40321
41553
129532
152328

Figure 2.5 Validation of hit compounds as activators of the EGFR kinase domain. Dose response curves of hit compounds in the absence and presence of EGFR indicate whether the increase in signal from the ADP Hunter assay (A) is independent of EGFR and constitutes a false positive compound (compounds 41467 and 152747 shown as examples) or (B) requires EGFR and thus indicates the compound stimulates EGFR activity (compounds 40321 and 152328 shown as examples). The data points were generated using a titration of compounds and the optimized ADP Hunter reaction conditions with no EGFR or 100 nM EGFR and a one hour reaction time. (C) Five compounds (39797, 40321, 41553, 129532, and 152328) are confirmed to activate EGFR in a dose-dependent manner in a radiometric kinase assay. Reactions were performed for ten minutes with 500 nM EGFR kinase domain, 500 μM Tail Peptide A, and 50 μM ATP labeled at 6.67 μCi/ml [γ-^{32}P]ATP and terminated by addition of 0.5% v/v phosphoric acid. Reactions were spotted onto filters, and the radioactivity was quantified by liquid scintillation counting. Fold-activation values were calculated by dividing either the fluorescence value (ADP Hunter) or counts per minute values (radiometric assay) by the mean value of samples that do not contain compound. (D) The chemical structures of the five compounds that activate the EGFR kinase domain: 39797, 40321, 41553, 129532, and 152328.

The most direct method to assay kinase activity is through radioactivity-based experiments, in which the incorporation of ^{32}P from [γ-^{32}P]ATP into the substrate peptide is measured by scintillation counting. The simplicity of radiometric assays minimizes the potential for compounds to interfere with the assay readout. A radiometric assay for EGFR kinase activity was developed to measure the phosphorylation of Tail Peptide A, a peptide that is derived from the autophosphorylation site at Tyr-1173 of the EGFR C-terminal tail. Each of the eight compounds that required EGFR for activation in the ADP Hunter Assay was tested in the radiometric assay, and five of these compounds were shown to increase EGFR activity in a dose-dependent manner (Figure 2.5C). These five compounds are considered true activators of EGFR kinase because the compounds consistently stimulate EGFR activity regardless of the assay. The data tend to be noisy, making quantitation difficult, but they do confirm that these compounds increase EGFR activity with varying efficacies. These five compounds are structurally diverse (Figure 2.5D). Conversely, stimulation of EGFR activity by the other three compounds was limited to the ADP Hunter assay, and since these compounds exhibit inconsistent activation of EGFR, they were not pursued further. Taken together, five compounds were confirmed to activate EGFR kinase, and the goal of the high throughput screen to identify compounds that simulate the activity of the EGFR kinase domain was successfully attained.

The new compound stocks, which were dissolved at higher concentrations than those used previously and thus make saturation easier to attain, permitted better measurements of activation and binding by the ADP Hunter and SPR assays, respectively (Figure 2.6). The values of EC_{50} range from approximately 3-300 μM. Compounds 40321 and 152328 exhibit the strongest stimulation of EGFR activity. Notably, the fold-activation values that are plotted do not take into account the large contribution of the assay background to the signal, as shown in Figure 2.3B,C, and so the true fold-activation is actually several fold greater. Out of the five compounds shown to increase EGFR activity, only compound 40321 exhibits hyperbolic and stoichiometric binding to the kinase domain, as quantified by the SPR assay, with a K_D value of 500±200 μM. Compounds 39797 and 129532 bind superstoichiometrically in the SPR assay, and compounds 41553 and 152328 do not bind. Though we do expect allosteric activators of EGFR to bind to the kinase domain, it is possible the compounds do not interact with sufficient stability to be detected in the SPR binding assay. Alternatively, the compounds that do not bind according to the SPR data may stabilize the active conformation of the kinase domain by a mechanism that does not require binding to EGFR.



Compound	EC_{50} (μM)	Top Fold-Activation	K_D (μM)
39797	3.4 ± 2.0	1.34 ± 0.05	superstoichiometric
40321	25 ± 4	1.92 ± 0.04	500 ± 200
41553	27 ± 8	1.69 ± 0.05	no binding
129532	290 ± 57	2.48 ± 0.15	superstoichiometric
152328	52 ± 12	2.14 ± 0.07	no binding

Figure 2.6 EGFR kinase domain activation and binding data for each of the five hit compounds. Each of the activating compounds (A) 39797, (B) 40321, (C) 41553, (D) 129532, and (E) 152328 were titrated in activation (circles, left Y axis) and binding assays (squares, right Y axis), and (F) the data were quantified by fits to hyperbolic binding equations (activation data as solid lines, binding data for 40321 as a dashed line) to derive the activation parameters EC_{50} and top fold-activation from the activation data or dissociation constant K_D from the SPR data. Activation data were generated in the ADP Hunter assay using the optimized reaction conditions with 500 nM EGFR kinase domain for a reaction time of ten minutes, and fold-activation values were calculated by dividing the fluorescence value by the mean value of samples that do not contain compound. Data points are means ± standard deviations of mean from four replicates. Binding data were generated by SPR. Biotinylated EGFR kinase domain was immobilized on a neutravidin-coated chip, and the binding response of different

concentrations of each compound to the EGFR kinase domain was measured. Binding measurements were performed in duplicate. Data were fit to a hyperbolic hyperbolic binding equation using Prism (Graphpad) to derive values of EC_{50}, top fold-activation, and K_D.

2.2.6 The activating effects of the compounds are largely reversible and sensitive to reductant

Compounds that are robust to assay conditions and reversible are desirable from the perspective of medicinal chemistry as the desired properties of the compounds may be more likely to be transferable if eventually introduced to cells. The ability of each activating compound to function reversibly was assessed to ensure that the mechanism of activation did not involve permanent modification of the EGFR kinase domain. Each compound was incubated with EGFR at a compound concentration at which EGFR activity approaches maximal stimulation for that compound. The compound and EGFR mix was then diluted to a compound concentration at which the stimulation of EGFR activity should be reduced either partially or fully, and the activity of EGFR was measured in the radiometric kinase assay. If the effects of the compound are not reversible, the activity of EGFR with the diluted compound should resemble the activity observed when the compound is assayed at a high concentration. All of the compounds with the exception of 39797 exhibit reduced activation of the EGFR kinase domain after dilution (Figure 2.7A). This suggests that the effects of compounds 40321, 41553, 129532, and 152328 on EGFR activity are reversible. Compound 39797 appears irreversible, but the large range of the data for the diluted compound may mask any reversibility.

The effects of the reductant dithiothreitol (DTT) on the abilities of the compounds to stimulate EGFR activity were evaluated. Compounds that produce different effects in the presence and absence of reductant are less desirable as this suggests specific redox requirements for the function of the compound. The activity of the EGFR kinase domain is insensitive to the presence of DTT in the radiometric kinase assay (Figure 2.7B, inset). The activation of EGFR by compounds 129532 and 152328 is dramatically reduced in the presence of DTT, with values of fold-activation that are decreased by factors of two and three, respectively (Figure 2.7B). DTT slightly decreases the activation of EGFR by compounds 39797 and 41553. Conversely, the addition of DTT increases the fold-activation of EGFR by compound 40321 by about 25%. Both compounds 40321 and 41553 contain thiol groups. These two compounds are not expected to covalently modify the EGFR kinase domain since the effects of the compounds are reversible. However, the presence of DTT may affect modification of the thiols of the compounds in a way that either positively, in the case of compound 40321, or negatively, in the case of 41553, affects activation of EGFR. It is currently unknown why the majority of compounds, including those with no thiols, appear to be redox state sensitive as demonstrated by the loss of efficacy in activating EGFR in the presence of DTT. It is possible the compounds themselves are modified in the presence of DTT or the sites on EGFR at which the compounds interact are affected by addition of DTT.

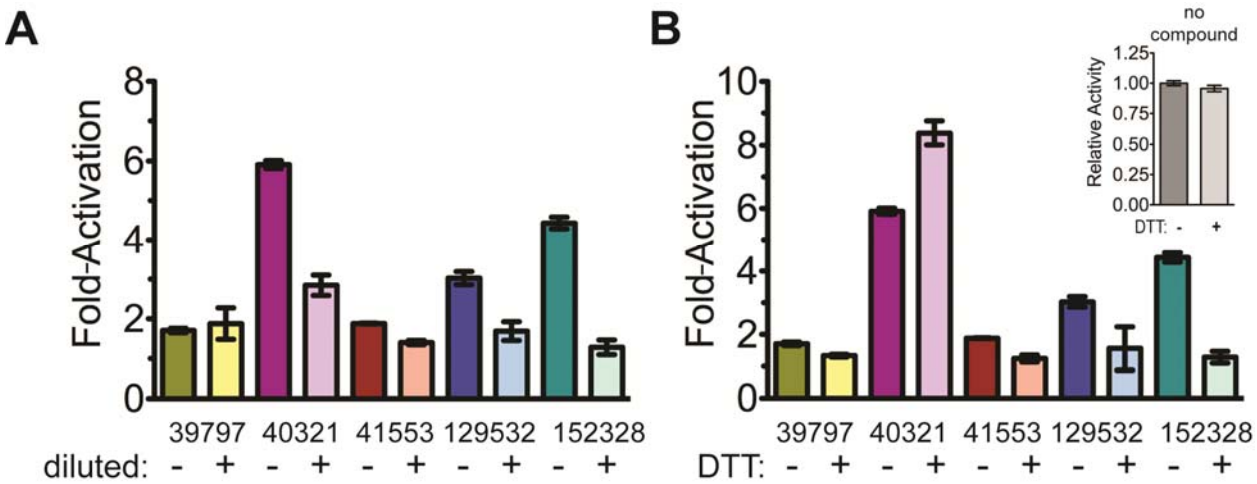


Figure 2.7 The reversibility and DTT dependence of compound activation of EGFR. (A) Incubation of the compounds at a high concentration with EGFR followed by dilution indicates that the activating effects of all compounds except 39797 are reversible. The compound concentrations prior to and after (+) dilution are: 39797 – 18 and 3 μM, 40321 - 240 and 40 μM, 41553 – 60 and 10 μM, 129532 – 1.0 and 0.17 μM, and 152328 – 300 and 50 μM. Notably, the pre-dilution compound concentrations do not exactly match the compound concentrations (listed below) of the non-diluted (-) samples that are plotted, but the results are expected to be similar. (B) The activation of EGFR by the majority of the compounds is reduced by the addition of 1 mM DTT, with the exception of compound 40321. The addition of DTT does not affect the activity of the EGFR kinase domain itself (inset). All reactions were performed using the radiometric kinase assay. Reactions were performed for ten minutes with 500 nM EGFR kinase domain, 500 μM Tail Peptide A, and 50 μM ATP labeled at 6.67 μCi/ml [γ-^{32}P]ATP and terminated by addition of 0.5% v/v phosphoric acid. Reactions were spotted onto filters, and the radioactivity was quantified by liquid scintillation counting. The concentrations of compounds in the non-diluted samples and DTT dependence samples were 39797 – 100 μM, 40321 – 250 μM, 41553 – 200 μM, 129532 – 6.25 μM, and 152328 – 400 μM. Fold-activation values were calculated by dividing either the counts per minute values by the mean value of samples that do not contain compound. The plotted values are means, and error bars represent the range of two replicates.

2.2.7 The compounds do not require the asymmetric dimer interface to activate the EGFR kinase domain

One mechanism by which the compounds could increase activity of the EGFR kinase domain is through promoting the formation of the activating asymmetric dimer. The abilities of the compounds to activate mutants of the EGFR kinase domain that are unable to form the asymmetric dimer were tested to establish if the compounds require EGFR dimerization and/or the site of the mutation to be intact for activation to occur. First, the L736R mutant, which is at the asymmetric dimer interface on the N-lobe, was assayed for activation by the compounds at a single concentration in the radiometric

kinase assay (Figure 2.8A). Notably, this mutation is located within the hydrophobic pocket that was targeted in the small molecule screen. All of the compounds activate the L736R kinase to similar extents as observed with the wild type kinase, with the exception of compound 152328, which activates this kinase variant only weakly (Figure 2.8B). This suggests that compound 152328 either activates EGFR by interacting at this site or requires EGFR kinase dimerization for activation to occur. Since the mutation does not affect the activities of the other compounds, either the compounds do not bind at the hydrophobic pocket or do bind at this site but Leu-736 is not required. Additionally, these four compounds are unlikely to activate EGFR by a dimerization-dependent mechanism.

The compounds were also assayed for activation of the EGFR kinase domain mutant V924R, which is located on the C-lobe at the asymmetric dimer interface, to confirm that the intact asymmetric dimer is not required for compound-mediated activation (Figure 2.8A). A titration of each compound was assayed with the V924R kinase domain, and EGFR activity was measured in the ADP Hunter assay. All of the compounds activate the V924R kinase, confirming that the asymmetric dimer interface is not required for activation by any of the compounds (Figure 2.8C). As discussed in Chapter 3, the EGFR kinase domain may be activated by an aggregation-mediated mechanism that promotes formation of the asymmetric dimer. Because the compounds are able to activate the L736R and V924R kinase domains (with the exception of compound 152328, which strongly activates only the V924R kinase domain), the compounds are unlikely to activate EGFR by aggregation. To confirm this, a titration of each compound was incubated with the EGFR kinase domain. The mixture was centrifuged, and the resuspended pellet was resolved by SDS-PAGE. Incubation of the compounds with the EGFR kinase domain does not increase band intensity for insoluble EGFR on the gels, which confirms the compounds do not aggregate EGFR (data not shown).

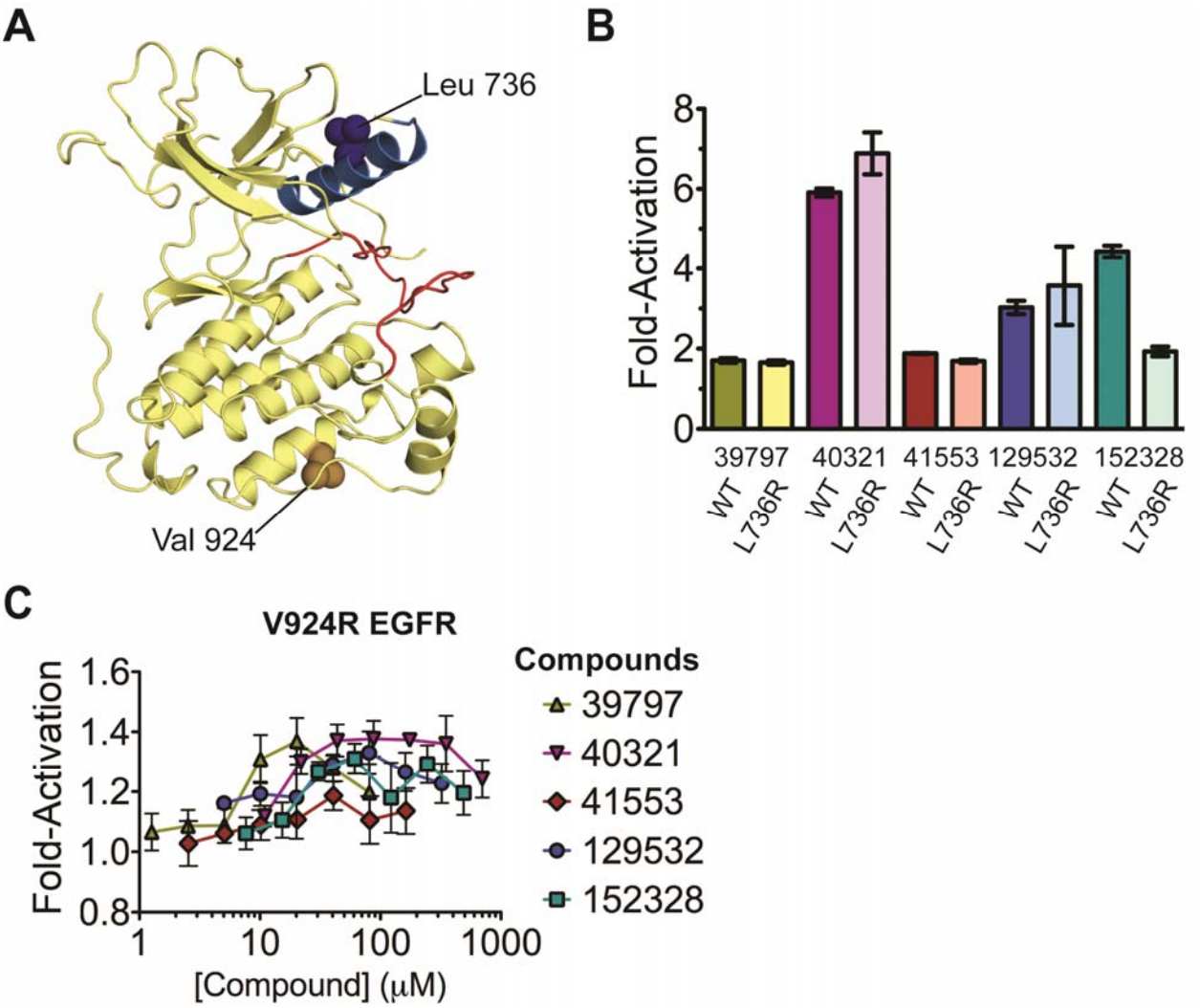


Figure 2.8 The activation of dimerization mutant EGFR kinase domains by the compounds. (A) The locations of the two residues at the asymmetric dimer interface, Leu-736 on the N-lobe and Val-924 on the C-lobe, that were mutated are depicted as spheres (PDB: 2GS6) (Zhang et al., 2006). Additionally, Leu-736 is located within the hydrophobic pocket that was targeted in the HTS campaign. (B) The activation of wild type (WT) and L736R EGFR kinase domain by each of the compounds indicates that the activation by only compound, 152328, is reduced by the L736R mutation. All reactions were performed using the radiometric kinase assay in a ten-minute reaction time using 500 nM WT or L736R EGFR kinase domain, 500 μM Tail Peptide A, and 50 μM ATP labeled at 6.67 μCi/ml [γ-^{32}P]ATP and terminated by addition of 0.5% v/v phosphoric acid. Reactions were spotted onto filters, and the radioactivity was quantified by liquid scintillation counting. The concentrations of compounds were 39797 – 100 μM, 40321 – 250 μM, 41553 – 200 μM, 129532 – 6.25 μM, and 152328 – 400 μM. (C) The activation of V924R kinase domain by a titration of each of the compounds indicates that all compounds are able to activate this EGFR dimerization mutant. Activation data were generated in the ADP Hunter assay using the optimized reaction conditions and with 500 nM V924R EGFR in a reaction time of ten minutes. Fold-

activation values were calculated by dividing either the counts per minute values (radiometric assay) or fluorescence values (ADP Hunter) by the mean value for samples that do not contain compound. The values plotted are means, and error bars represent the range of two replicates (radiometric assay) or the standard deviations of mean of four replicates (ADP Hunter).

2.2.8 Structure activity relationship studies examine the roles of the compound functional groups

Structure activity relationship (SAR) studies of hit compounds are useful to identify the roles of the functional groups of compounds in achieving the desired effect and to guide development of more potent compounds. Limited SAR studies for the activating compounds were performed using structurally similar compounds that were available within the Small Molecule Discovery Center compound library or that were synthesized by Clifford Bryant, a medicinal chemist at the Small Molecule Discovery Center. The majority of compounds that were tested, compounds 41225, 87195, 530545, and 530546, are analogues of the compound 40321, which was considered to be the most promising hit compound, while the fifth compound, compound 530547, is an analogue of compound 39797 (Figure 2.9A). The effects of each compound on EGFR kinase domain activity were measured in the ADP Hunter assay (Figure 2.9B). Of the 40321 analogue compounds, only compound 530545 activated the EGFR kinase domain, albeit weakly. The results for compounds 41225 and 87195 indicate that the aliphatic segment of compound 40321 is required for activation of EGFR. Furthermore, activity of EGFR decreases with these two compounds, but this may be a result of limited compound solubility upon dilution in the assay buffer and not specific inhibition of EGFR. Likewise, the results for compounds 530545 and 530546 indicate that compounds in which the thiol of compound 40321 is substituted with either methylsulfanyl or methylsulfinyl are no longer able to strongly activate the EGFR kinase domain. Compound 530547, the analogue of compound 39797 in which the chlorine is substituted with bromine, does not significantly activate EGFR.

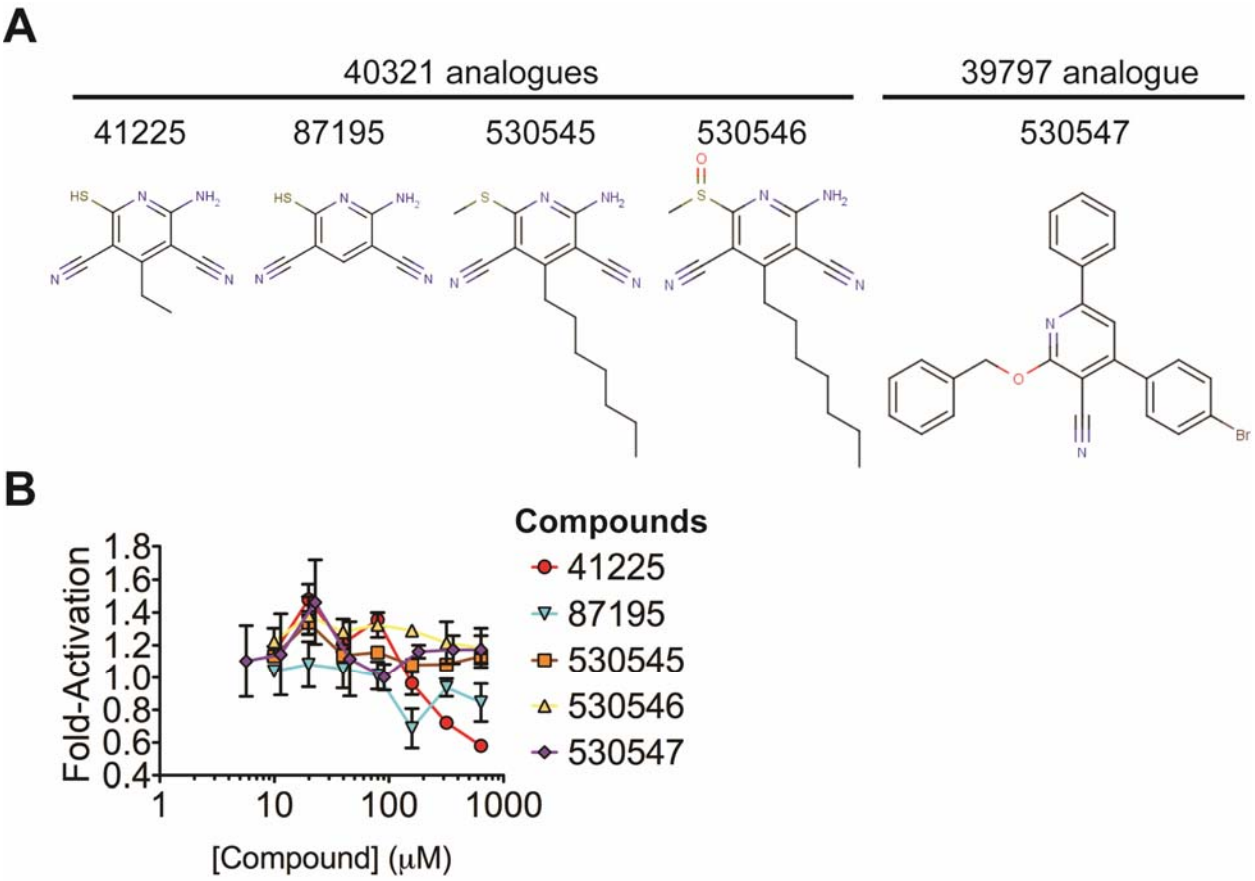


Figure 2.9 Structure activity relationship analysis of the hit compounds. (A) Compounds 41225, 87195, 530545, and 530546 are all analogues of hit compound 40321, while compound 530547 is an analogue of hit compound 39797. (B) Activation of EGFR in response to titration of each of the compounds in the ADP Hunter kinase assay shows that none of the compounds increase EGFR activity significantly, relative to what was observed previously with the hit compounds. Data were generated in the ADP Hunter assay using the optimized reaction conditions and with 500 nM EGFR and a reaction time of ten minutes. Fold-activation values were calculated by dividing the fluorescence values by the mean value of samples that do not contain compound. Data points are means ± standard deviations of mean from four replicates

2.2.9 The effects of the activating compounds on formation of the EGFR kinase domain asymmetric dimer

A primary goal in the discovery of compounds that activate the EGFR kinase domain was to evaluate if the compounds can inhibit formation of the activating asymmetric dimer of EGFR kinase domains. One mechanism by which the compounds may increase EGFR kinase domain activity is by binding at the hydrophobic pocket on the kinase domain N-lobe and substituting for the effects of the activator subunit in the asymmetric dimer, as outlined in Figure 2.1. Compounds that bind here should also

block formation of the asymmetric dimer. However, the data presented in section 2.2.7, in which all of the compounds except 152328 are shown to activate the L736R mutant kinase domain, suggest that most of these compounds may not bind at the hydrophobic pocket that was targeted.

As mentioned earlier, the EGFR kinase domain is predominantly monomeric *in vitro*, which complicates studying the formation of the asymmetric dimer in biochemical assays. However, concurrent with the small molecule screen for activators of the EGFR kinase domain, the role of the cytoplasmic juxtamembrane segment in stabilizing formation of the asymmetric dimer was revealed (Jura et al., 2009a; Red Brewer et al., 2009). The N-terminal portion of the cytoplasmic juxtamembrane segment (residues 645-663) from each subunit of the EGFR dimer forms a helical dimer, and the dimerization affinity of a kinase domain construct that includes the full cytoplasmic juxtamembrane segment, which was measured by examining specific activity at different kinase concentrations, is estimated to be 200 nM. The C-terminal portion of the cytoplasmic juxtamembrane segment (residues 664-685) of the receiver subunit binds the C-lobe of the activator subunit in an interaction termed the juxtamembrane latch, and the dimerization affinity of the construct consisting of the C-terminal half of the juxtamembrane segment and kinase domain is estimated to be 8 μM (Jura et al., 2009a). These interactions are depicted in Figure 1.3C. Therefore, a recombinantly-expressed kinase construct that includes the juxtamembrane segment can be used to study formation of the kinase domain asymmetric dimer in solution. Previously, the only mechanism available to dimerize the kinase domain was to tether the kinase to the surface of small unilamellar vesicles (Zhang et al., 2006). Additionally, assaying pairs of kinase domain variants that each bear one of two asymmetric dimer interface mutations, either the I682Q mutation on the N-lobe or the V924R mutation on the C-lobe, limits the kinase domains to forming dimers and prevents formation of oligomeric chains of kinases through the asymmetric dimer interface, like those observed in the crystal packing interactions of the active conformation structure (Zhang et al., 2006).

The protein MIG6, a feedback inhibitor of EGFR, inhibits EGFR activity in part by binding to the C-lobe of the kinase domain and preventing formation of the asymmetric dimer. Peptides derived from MIG6 bind to the kinase domain and inhibit dimerization-dependent activation of the EGFR kinase domain *in vitro* (Zhang et al., 2007). These MIG6-derived peptides can be used as a positive control for inhibition of the EGFR asymmetric dimer and an indicator that the assay specifically examines kinase domain dimerization.

Multiple biophysical assays were developed in an attempt to examine dimerization of the EGFR kinase domain. The increase in kinase activity due to dimerization mediated by the C-terminal portion of the juxtamembrane segment is similar in magnitude to the increase in activity from addition of the activating compounds, making it difficult to deduce the individual effects of dimerization and the compounds in activity-based assays. In the majority of the assays that were developed, the kinase domain constructs included the C-terminal portion of the juxtamembrane segment, herein called JM-B, to support dimerization of EGFR. In one assay JM-B

kinase domain constructs with either the I682Q or V924R mutations were combined in solution with a MIG6 peptide that was labeled with fluorescein. If dimerization of the kinase domains is disrupted, the fluorescence polarization (FP) signal for the labeled MIG6 peptide should increase, as the MIG6 peptide can now bind to the C-lobe of the I682Q kinase domain. Previous work suggested the MIG6 peptide should bind exclusively to the I682Q kinase domain and not the V924R mutant because the V924R mutation interferes with the kinase domain epitope to which MIG6 binds (Zhang et al., 2007). However, the results indicate that the MIG6 peptide binds both I682Q and V924R kinase domains, albeit the latter with lower affinity, which causes the FP data to be too convoluted to analyze any dimerization. In another assay chimeric JM-B kinases with fluorescent proteins fused to the C-terminal tails, either I682Q JM-B kinase domain with cerulean fluorescent protein (CFP) or V924R JM-B kinase domain with the yellow fluorescent protein variant Venus, were combined, and dimerization was examined by Förster resonance energy transfer (FRET) between the fluorescent proteins. Control assays utilizing the MIG6 peptide or the fluorescent proteins alone indicate the FRET values obtained were not specific to EGFR dimerization, suggesting that either the kinases do not stably dimerize or the distance between the fluorescent proteins within the dimer is too great to allow detection by FRET. Other simpler assays examining dimerization-dependent autophosphorylation, complex size by analytical size exclusion chromatography, or co-immunoprecipitation of the JM-B kinase domains also failed to demonstrate kinase domain dimerization. From these attempts, the conclusion was reached that the JM-B segment may only weakly support kinase dimerization, which results in short-lived dimers. Therefore the JM-B segment kinase construct was ultimately not useful in developing a dimerization assay.

Use of the kinase domain construct that includes the full juxtamembrane segment (JM-AB) in biophysical assays is problematic because the kinase construct may dimerize through the helical dimer independent of forming the kinase domain asymmetric dimer (Natalia Jura, personal communication). Thus, inhibition of the asymmetric dimer may not be detected due to the helical dimer remaining intact. However, since the JM-AB kinase domain is approximately 70-fold more active in solution than the kinase domain and the activating compounds are not as efficacious in increasing kinase activity, kinase domain dimerization can be examined using an activity-based assay (Jura et al., 2009a). The effects of a small concentration range of the compounds on the activity of the JM-AB kinase domain were measured using the ADP Hunter assay (Figure 2.10A). Compounds 40321 and 129532 both slightly but significantly decrease the kinase activity, suggesting that these compounds may inhibit dimerization of the kinase domain. The remaining three activating compounds have no significant effect on the activity of the JM-AB kinase domain.

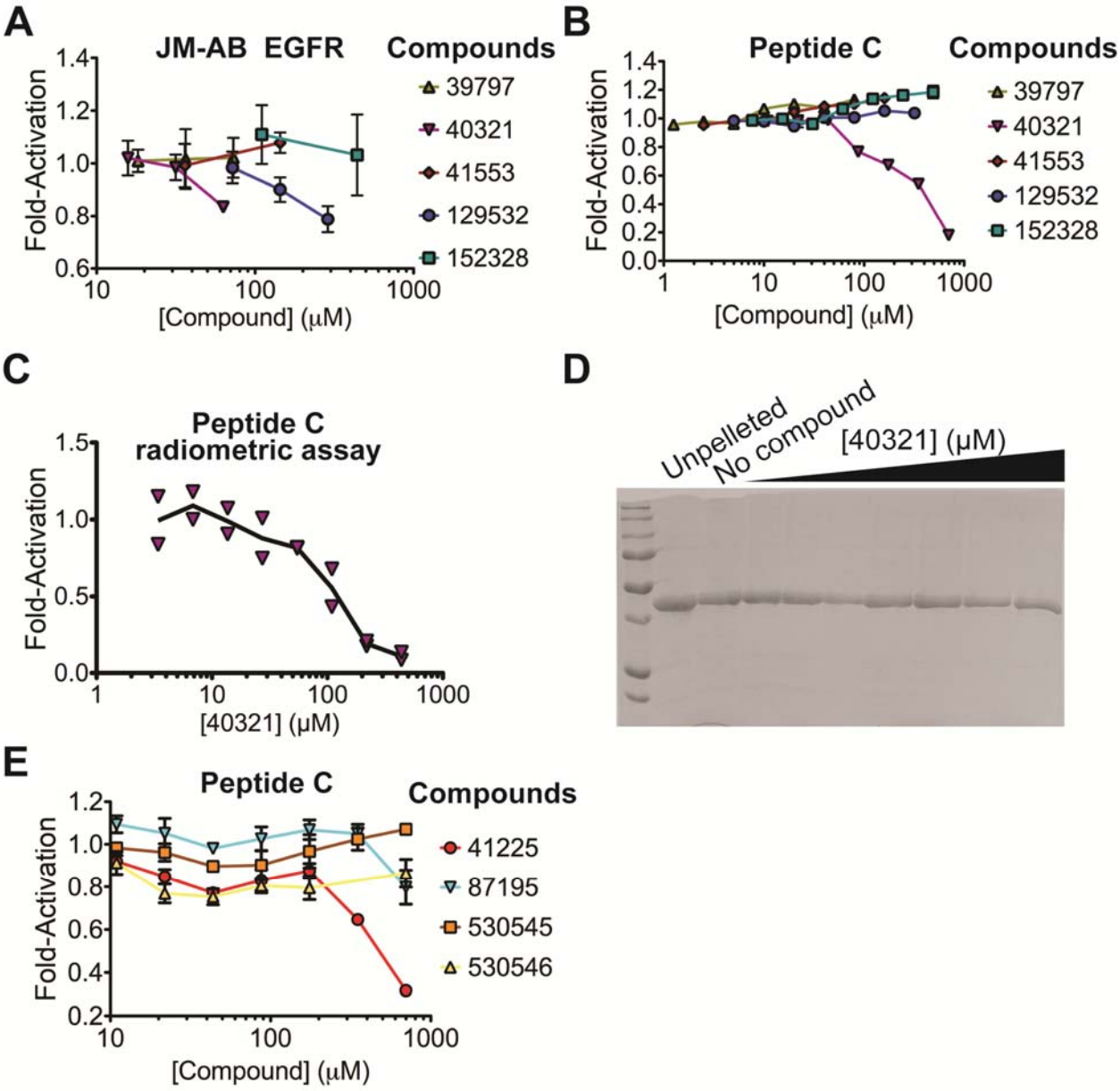


Figure 2.10 The effects of the hit compounds on EGFR dimerization-dependent activity. (A) The EGFR kinase domain construct that includes the juxtamembrane segment (JM-AB) assayed with a titration of the hit compounds indicates compounds 40321 and 129532 decrease the activity of this kinase domain construct while the remaining compounds do not. Activation data with JM-AB EGFR kinase domain were generated in the ADP Hunter assay using the optimized reaction conditions and with 200 nM JM-AB EGFR, 2 mg/ml Poly 4Glu:Tyr, and a reaction time of ten minutes. (B) EGFR kinase domain catalysis with Peptide C as a peptide substrate, which increases EGFR activity by promoting formation of the asymmetric dimer, assayed with a titration of the compounds shows that compound 40321 decreases EGFR activity, while the other compounds have no effect or increase EGFR activity. Activation data with Peptide C were generated in the ADP Hunter assay using 500 nM EGFR kinase domain, 100 μM Peptide C, and a twenty-minute reaction time. (C) The inhibitory effect of compound

40321 on EGFR catalysis with Peptide C is reproducible in the radiometric kinase assay. Reactions were performed for ten minutes with 500 nM WT EGFR kinase domain, 100 μM Peptide C, and 25 μM ATP labeled at 8.325 μCi/ml [γ-^{32}P]ATP and terminated by addition of 0.5% v/v phosphoric acid. Reactions were spotted onto filters, and the radioactivity was quantified by liquid scintillation counting. The data points are for two individual filters. (D) The EGFR kinase domain precipitates in the presence of Peptide C and compound 40321 over a compound concentration range of 11-700 μM, which was generated by a series of two-fold dilutions. Two μM EGFR kinase domain, 100 μM Peptide C, and compound 40321 in 5% v/v DMSO were incubated for 20 minutes, centrifuged, and the pelleted materials were resolved by SDS-PAGE. (E) The analogues of compound 40321 do not inhibit EGFR catalytic activity with Peptide C. Though the activity with compound 41225 decreases, this is postulated to not be specific to dimerization. Data were generated as described in (B). Fold-activation values were calculated by dividing either the counts per minute values (radiometric assay) or fluorescence values (ADP Hunter) by the mean value for samples that do not contain compound. Data points in (A), (B), and (E) are means ± standard deviations of mean from four replicates. The error bars in (B) are too small to be visible on this graph.

The EGFR peptide substrate Peptide C, which is phosphorylated with much higher efficiency than other peptide substrates of EGFR, was discovered to increase the catalytic efficiency of the EGFR kinase domain through an aggregation-based mechanism that promotes formation of the asymmetric dimer. (This work that describes the properties of Peptide C is the subject of Chapter 3 of this dissertation.) Each of the EGFR dimerization mutants I682Q and V924R are inactive when assayed with Peptide C as a substrate, indicating the absolute requirement for formation of the asymmetric dimer in EGFR catalysis with Peptide C. Additionally, a MIG6 peptide inhibits the activity of EGFR with Peptide C in a dose-dependent manner, supporting that catalysis with Peptide C requires formation of the asymmetric dimer (data not shown). Similar to the JM-AB kinase domain activity assays, a decrease in EGFR activity when Peptide C is used as the substrate may indicate that the compounds disrupt dimerization. The activity of the EGFR kinase domain with Peptide C as a substrate was assayed in the presence of each of the activating compounds using the ADP Hunter assay (Figure 2.10B). Compound 40321 decreases the activity of EGFR with Peptide C in a dose-dependent manner, while the other compounds either increase EGFR activity or have no effect. Similarly, compound 40321 decreases the phosphorylation of Peptide C in the radiometric kinase assay (Figure 2.10C). These data complement the data obtained for the JM-AB kinase that suggest compound 40321 might block kinase domain dimerization.

Compound 40321 could inhibit the activity of EGFR with Peptide C by disrupting Peptide C-mediated aggregation of the kinase domain to prevent formation of the asymmetric dimer as opposed to inhibiting formation of the dimer directly. To rule the former scenario out as the mechanism of inhibition, samples of the EGFR kinase domain were incubated with Peptide C and a titration of compound 40321, and insoluble materials were resolved by SDS-PAGE following centrifugation (2.10D). The addition of

compound 40321 does not affect the capacity of Peptide C to aggregate EGFR, suggesting that the observed inhibition occurs because the asymmetric dimer is directly inhibited from forming.

The analogue compounds of compound 40321 were tested for the ability to inhibit Peptide C-mediated activation of the EGFR kinase domain. Though these compounds were not effective in stimulating EGFR kinase domain activity, there may still be a capacity to block dimerization. The effects of the 40321 analogue compounds on Peptide C activation of EGFR were examined using the ADP Hunter assay (Figure 2.10E). Compound 41225 decreases EGFR activity, though a similar result was observed with the kinase domain and the substrate Poly 4Glu:Tyr, suggesting that the compound inhibits EGFR activity by an unknown mechanism that may be related to limited solubility of the compound (Figure 2.9B). Compound 530546 slightly decreases the activity of EGFR with Peptide C, while the other analogue compounds have no significant effect. Therefore, compound 40321 is both the most potent activator of the EGFR kinase domain and inhibitor of dimerization from this group of compounds.

2.2.10 Attempts to obtain co-crystal structures of compounds bound to the EGFR kinase domain

The binding sites of the activating compounds on the EGFR kinase domain are unknown, and extensive efforts to obtain co-crystal structures were made to identify the sites. Not only might identification of the binding sites suggest the mechanisms by which the compounds increase EGFR activity, it would also permit structure-aided design to guide development of molecules with increased efficacy and improved binding energies. Numerous attempts were made to co-crystallize the compounds with either the wild type kinase domain or the V924R kinase domain that cannot form the asymmetric dimer. Despite the identification of tens of crystallization conditions that produced proteinaceous crystals, all of the structures that were solved lack electron density for the compounds. This suggests the accessible surface of the kinase domain in the crystal structures that were solved does not provide a binding site for the compounds.

One apparent issue is that the wild type kinase domain crystallizes consistently in the asymmetric dimer crystal form, which forms an infinite chain of kinases interacting through the asymmetric dimer interface, while the V924R kinase domain always adopts the inactive conformation (Zhang et al., 2006). In both cases the hydrophobic pocket of the N-lobe that was targeted in the HTS campaign is inaccessible, either due to crystal packing interactions or the conformation of the kinase. Therefore, in the case that the compounds bind the hydrophobic pocket, the compounds cannot be successfully co-crystallized with these kinase domains unless significant structural changes are sufficiently stabilized. Furthermore, crystals of the wild type or V924R kinase domains likely cannot serve as vehicles for compound soaks due to the same issues with accessibility of the hydrophobic pocket. Alternatively, the compounds may not co-crystallize because the majority of compounds may not bind to the kinase domain, as is suggested by the SPR data.

To circumvent this problem, two strategies to crystallize an active kinase domain in the monomeric state were implemented, which would result in a kinase domain with the hydrophobic pocket exposed. There are currently no known structures in which the kinase domain is a monomer and in the active conformation. Such a crystal form of the EGFR kinase domain could either be co-crystallized with the compounds or serve as a vehicle for compound soaks. To promote the active conformation, either the L834R mutation was introduced or the covalent active site inhibitor canertinib (also called Ci-1033) was bound. Canertinib was expected to bind the active conformation of the kinase, based on what was previously observed for the structurally similar inhibitors gefitinib and erlotinib (Kotra et al., 2008; Stamos et al., 2002; Yun et al., 2007). To prevent crystallization in the asymmetric dimer crystal form, the V924R mutation on the kinase domain C-lobe was included. Crystallization attempts were done either in the presence or absence of the activating compounds. The canertinib-bound V924R EGFR kinase domain structure was solved, but the kinase domain crystallized as a symmetric dimer in the inactive conformation (Jura et al., 2009a). Electron density for the activating compounds was never observed. Crystals of the L834R/V924R kinase domain were not obtained, but another group solved the structure of this construct subsequently. This kinase adopts the inactive conformation, similar to the canertinib-bound kinase structure (Gajiwala et al., 2012). Altogether, these structures highlight the relative instability of the active conformation of the EGFR kinase domain in the absence of the stabilizing interactions of the asymmetric dimer. The N-lobe of the kinase domain was recently shown in molecular dynamics simulations to be intrinsically disordered. Introduction of the L834R activating mutation suppresses this disorder but also increases the propensity of the kinase domain to dimerize (Shan et al., 2012). This suggests that without the ability to form the asymmetric dimer, the kinase domain is energetically more stable in the inactive conformation and explains the difficulty in obtaining monomeric EGFR kinase domains that are in the active conformation. Though the compounds promote the active conformation of the kinase domain, the interactions may not be sufficiently stable for crystallization.

2.3 Conclusions

In this work, five compounds have been discovered to increase the catalytic activity of the EGFR kinase domain. The biochemical work provides a framework for how these compounds increase the catalytic activity of EGFR, but the lack of co-crystal structures precludes specific mechanisms. Most of the compounds are unlikely to covalently modify the kinase domain as the activating effects of all but compound 39797 are reversible. However, the degree to which EGFR is activated by all compounds except compound 40321 is reduced in the presence of DTT, which may reflect chemical modifications of the compounds that prevent activation of EGFR. For example, compound 41553 has a thiol that may be redox active, while the 1,2,4,5-tetrazine scaffold of compound 152328 is expected to be reduced readily at the 1 and 4 positions of the compound (Kaim, 2002). Because the compounds are all able to activate the V924R kinase domain that cannot form the asymmetric dimer, the compounds are not expected to increase EGFR activity through promoting formation of the asymmetric dimer like the mechanism of EGFR activation in the presence of Peptide C.

The inability to obtain co-crystal structures of the compounds bound to the EGFR kinase domain and lack of demonstrated binding for all compounds except compound 40321 by the SPR assay suggests two scenarios. In one case these compounds may not bind EGFR but instead exert the activating effects by altering the bulk solution conditions in a way that stabilizes the active conformation of the kinase domain or otherwise increases the efficiency of phosphorylation (Savelli et al., 2000). Alternatively, the intrinsic disorder of the N-lobe and consequent instability of the active conformation may result in transient binding of the compounds to the active conformation that is not detectable by SPR and co-crystallization (Shan et al., 2012). The latter problem is highlighted by the fact that the active conformation of the EGFR kinase domain is always resolved as forming the asymmetric dimer, and a crystal structure of a monomeric, active kinase domain has never been solved.

The ultimate goal of discovering activators of the EGFR kinase domain was to identify compounds that bind at the N-lobe hydrophobic pocket and could block the formation of the asymmetric dimer. Development of a biophysical assay that measures dimerization has been problematic due to issues with stably forming asymmetric dimers, and there are currently no published biophysical methods that examine dimerization of the EGFR kinase domain directly *in vitro*. However, the high catalytic activity of the JM-AB EGFR kinase domain and of the EGFR kinase domain with Peptide C are both dependent on formation of the asymmetric dimer, and the effects of the activating compounds in these two scenarios were examined. Compound 40321 reduces the dimerization-dependent EGFR activity consistently in these assays, suggesting that this compound indeed inhibits kinase domain dimerization. This compound was also shown to bind stoichiometrically by SPR and functioned in the presence and absence of DTT. However, from a medicinal chemistry standpoint, compound 40321 is not favorable for additional development due to the results obtained in the limited SAR studies. The long alykl moiety may produce problems with solubility in cells, and truncation or removal of this group reduces the efficacy of the compound. Likewise, the thiol, which was shown to be necessary for activation of EGFR, could produce issues with unwanted reactivity in cells. Without a co-crystal structure to visualize the binding mode of compound 40321, it is extremely difficult to perform directed SAR studies to develop a molecule that blocks dimerization and has more tolerable chemical properties.

Nonetheless, the primary goal of the small molecule screening work was achieved, and five compounds that activate the EGFR kinase domain were successfully identified. This work confirms that the EGFR kinase domain can be activated by a small molecule, as was hypothesized based on similar studies of the kinase PDK1 (Engel et al., 2006). The outstanding question that remains is to understand the molecular mechanisms by which the effects of these compounds are exerted, which could involve a variety of possibilities. The compounds could directly bind at the N-lobe hydrophobic pocket, as originally predicted. Alternatively, the compounds may bind at an uncharacterized site on the kinase domain that either releases autoinhibition or stabilizes the active conformation. An additional mechanism may involve alteration to the solution conditions that generally affect the conformational stability and activity of the kinase domain. Taken together, these compounds offer new tools for studying the

activation of the EGFR kinase domain and when understood better will provide additional insights into the mechanisms that regulate the activity of EGFR.

2.4 Materials and Methods

2.4.1 EGFR kinase domain expression and purification

DNA from the human EGFR gene was cloned into pFAST Bac HT (Invitrogen) using the NcoI and HindIII restriction sites. The different constructs included the EGFR residues 672-998 (kinase domain), 645-998 (JM-AB kinase domain), or 658-998 (JM-B kinase domain). A N-terminal 6-His tag, linker, and TEV protease cleavage site (MSHHHHHHDYDIPTTENLYFQGAM) was included in each construct. Mutations to the kinase were generated using the Quikchange mutagenesis kit (Stratagene) and were confirmed by DNA sequencing. Recombinant bacmid (Bac-to-Bac expression system, Gibco BRL) were transfected into Sf9 cells for recombinant baculovirus production. Sf9 cells were grown in suspension and infected with the recombinant baculovirus. The cells were harvested after 2-3 days by centrifugation at 4000 x g and resuspended in lysis buffer of 50 mM Tris, pH 8.0, 5% v/v glycerol, 1 mM DTT, and protease inhibitor cocktail (Roche). The resuspended cells were lysed by French press, and the lysate was clarified by centrifugation at 40000 x g for one hour. The supernatant was loaded on a 60 ml Q-Sepharose Fastflow column (GE) that was equilibrated in buffer A (50 mM Tris, pH 8.0, 5% glycerol, and 15 mM β-mercaptoethanol). The protein-bound column was washed with buffer A, and protein was eluted with buffer B (50 mM Tris, pH 8.0, 5% v/v glycerol, 1 M NaCl, and 7.5 mM β-mercaptoethanol). The eluted protein was loaded onto a 1 ml HisTrap FF column (GE) that was equilibrated with buffer C (20 mM Tris, pH 8.0, 5% v/v glycerol, 500 mM NaCl, 20 mM imidazole and 7.5 mM β-mercaptoethanol). The protein-bound column was washed with buffer C, and protein was eluted over a gradient of buffer C and buffer D (20 mM Tris, pH 8.0, 5% v/v glycerol, 500 mM NaCl, 250 mM imidazole and 7.5 mM β-mercaptoethanol). The protein was stored overnight at 4°C. The protein was diluted 10-fold into buffer E (20 mM Tris, pH 8.0, 5% v/v glycerol, 20 mM NaCl, and 2 mM DTT) and loaded onto a 6 ml Uno-Q column (Bio-Rad) that was equilibrated with buffer E. The protein-bound column was washed with buffer E, and the protein was eluted over a gradient of buffer E and buffer F (20 mM Tris pH 8.0, 5% v/v glycerol, 1 M NaCl, and 2 mM DTT). The eluted protein was analyzed by SDS-PAGE, and fractions containing pure EGFR kinase domain were pooled, concentrated, and buffer exchanged into storage buffer (20 mM Tris pH 8.0, 10% v/v glycerol, 50 mM NaCl, 2 mM TCEP and 2 mM DTT). Protein concentration was measured by amino acid analysis. The protein was flash frozen with liquid nitrogen and stored at -80°C.

2.4.2 ADP Hunter assay

The optimized EGFR kinase reaction was performed with 100, 200, or 500 nM EGFR, 200 μM ATP, 1% v/v DMSO or compound dissolved in DMSO, and 2 mg/ml Poly 4Glu:Tyr (Sigma Aldrich) in the ADP Hunter reaction buffer (15 mM HEPES, pH 7.4, 10 mM $MgCl_2$, 20 mM NaCl, 0.1 mg/ml Bovine γ Globulin, 1 mM EGTA and 0.02% v/v Polysorbate-20) in a volume of 11 μl for either one hour, 10 minutes, or 20 minutes at

room temperature. When Peptide C was used as a substrate, the concentration was either 500 µM, when used as a positive control for EGFR activity, or 100 µM, when used in dimerization assays. Five µl of Reagent A and 10 µl of Reagent B of the ADP Hunter assay kit (DiscoveRx) were added sequentially, and the detection reaction was incubated for an hour at room temperature. The detection reaction was terminated and stabilized by addition of 2.5 µl of the Stop Solution from the ADP Hunter kit. The fluorescent signal of Resorufin was detected using an excitation wavelength of 530 nm and emission wavelength of 590 nm with a Synergy HT plate reader (Biotek). All assays were performed in 384-well low-volume flat bottom black plates (Corning).

2.4.3 High throughput screen

The ADP Hunter assay in the high throughput screen was performed generally as described earlier. The reaction conditions were 100 nM EGFR, 200 µM ATP, 1% v/v DMSO or compound dissolved in DMSO, and either 2 mg/ml Poly 4Glu:Tyr (Sigma Aldrich) (negative-control and experimental wells) or 500 µM Peptide C (positive-control wells) in the ADP Hunter reaction buffer (15 mM HEPES, pH 7.4, 10 mM $MgCl_2$, 20 mM NaCl, 0.1 mg/ml Bovine γ Globulin, 1 mM EGTA and 0.02% v/v Polysorbate-20). All assays were performed in 384-well low-volume flat bottom black plates (Corning). Five µl of the substrate solutions (ATP and Poly 4Glu:Tyr or Peptide C) and 1 µl of compound were aliquoted to the plates using a Biomek Fx (Beckman Coulter) liquid handling robot. Peptide C was added to rows A, B, O, and P of columns 1, 2, 23, and 24, and all other wells contained Poly 4Glu:Tyr. Individual compounds dissolved in DMSO from the Small Molecule Discovery Center library were added to columns 3-22, and DMSO was added to columns 1, 2, 23, and 24. Five µl of EGFR was added using a Matrix Wellmate (Thermo) bulk liquid dispenser to initiate the reactions. The plates were briefly centrifuged, and reactions were incubated for one hour at room temperature. Five µl of Reagent A followed by 10 µl of Reagent B were added using the Matrix Wellmate. The plates were briefly centrifuged, and detection reactions were incubated for one hour at room temperature. Three µl of the Stop solution were added using the Matrix Wellmate. The plates were briefly centrifuged, and Resorufin fluorescence was detected using an excitation wavelength of 530 nm and emission wavelength of 590 nm with an Analyst HT plate reader (Molecular Devices).

2.4.4 Surface plasmon resonance binding assay

EGFR kinase domain was labeled with NHS-ester biotin at a ratio of 0.75:1, which resulted predominantly in labeling of one lysine. A CM5 sensor chip was functionalized with neutravidin (Thermo), and biotinylated EGFR kinase domain was bound to the surface to a signal of ~8800 response units. The inhibitor AG1478 was bound to the EGFR kinase domain. Compounds were flowed over the EGFR-coated sensor chip and a control sensor chip at varied concentrations in 2.5% v/v DMSO, and the differences in the binding response units between the two channels were recorded. The sensor chips were washed and compounds dissociated between subsequent measurements. Measurements were made using a Fujifilm AP3000.

2.4.5 Data fitting

Activation data for all compounds and binding data for compound 40321 were fit to the equation Y = X*Top/(Kd+X) – NS, where Y is the fold-activation or SPR response units, X is the compound concentration, Top is the maximum value for fold-activation or response units, Kd is the concentration of compound at which 50% activation (EC_{50}) or binding occurs (K_d), and NS is the non-specific background using Prism 5 (Graphpad).

2.4.6 Radiometric kinase assay

The reactions with Tail Peptide A consisted of 500 nM EGFR kinase domain, 500 µM Tail Peptide A, 50 µM ATP, 6.67 µCi/ml [γ-^{32}P]ATP, compound at indicated concentrations in 2.5% v/v DMSO, 15 mM HEPES, pH 7.5, 0.02% v/v Polysorbate-20, 1 mM EGTA, 20 mM NaCl, and 10 mM $MgCl_2$ in a final volume of 60 µl. DTT dependence assays were performed with 1 mM DTT included in the buffer. The Peptide C assay reaction consisted of 500 nM EGFR kinase domain, 100 µM Peptide C, 25 µM ATP, 8.325 µCi/ml [γ-^{32}P]ATP, compound 40321 at indicated concentrations in 2.5% v/v DMSO, 15 mM Tris, pH 7.5, 0.02% v/v Polysorbate-20, and 10 mM $MgCl_2$ in a final volume of 60 µl. Reactions were incubated for 10 minutes at room temperature. Reactions were terminated by addition of 60 µl of 0.5% v/v phosphoric acid, and 55 µl of each quenched reaction was spotted on each of two P81 phosphocellulose filter discs (Whatman). The discs were filtered and washed three times with 0.5% v/v phosphoric acid using a vacuum manifold (Hoefer Scientific Instruments). Five ml of UniverSol scintillation cocktail (MP Biomedicals) were added to the dried discs, and the radiation was measured by liquid scintillation counting (Beckman Coulter).

2.4.7 Precipitation analysis

Samples of compound 40321 at indicated concentrations in 5% v/v DMSO with 2 µM WT EGFR kinase domain and 100 µM Peptide C were incubated in 10 mM Tris, pH 7.5, 20 mM NaCl, and 0.02% v/v Polysorbate-20 in a final volume of 75 µl for about 20 minutes at room temperature. The samples were centrifuged at 15800 x g for 10 minutes at 4°C. Supernatants were removed, and pelleted materials were resuspended and resolved by SDS-PAGE.

Chapter 3

A highly efficient peptide substrate for EGFR activates the kinase by inducing aggregation

3.1 Overview

The kinase domain of EGFR is activated by an allosteric mechanism upon receptor dimerization. The C-lobe of one kinase subunit (the activator) interacts with the N-lobe of a second kinase subunit (the receiver) and results in the allosteric activation of the receiver through the formation of an asymmetric dimer of the kinase domains (Zhang et al., 2006). The recombinant kinase domain of EGFR is predominantly monomeric *in vitro* and thus does not readily undergo allosteric activation, resulting in very low catalytic activity. A construct of EGFR that includes the cytoplasmic juxtamembrane segment and the kinase domain forms the asymmetric dimer much more readily, and the catalytic efficiency is at least 70-fold higher than for the isolated kinase domain (Jura et al., 2009a; Red Brewer et al., 2009). The catalytic efficiency of the isolated kinase domain can also be increased by tethering it to small unilamellar vesicles, thereby increasing its effective concentration (Zhang et al., 2006).

The most important substrates of EGFR are tyrosine residues located in the long C-terminal tail of the receptor. When phosphorylated, these tyrosines serve as docking sites for effector proteins that relay the signal onward through the cell (Jorissen et al., 2003). Short peptides containing these tyrosine residues ("tail peptides") are poor substrates for the EGFR kinase domain, presumably because efficient phosphorylation of the tail occurs within an EGFR dimer or higher order oligomer (Clayton et al., 2005; Fan et al., 2005; Muthuswamy et al., 1999). For example, a 15 residue peptide containing Tyr-1173 from the tail of human EGFR, which we refer to as Tail Peptide A, has an estimated value of 1 mM for the Michaelis constant, K_M, when phosphorylated by the EGFR kinase domain.

Researchers at GlaxoSmithKline screened several peptide libraries for EGFR activity and identified a peptide substrate that is phosphorylated with a much higher efficiency than tail peptides (referred to as Peptide C in their work; we retain this name here) (Blackburn et al., 2003). The value of the catalytic efficiency, k_{cat}/K_M, for this peptide was reported to be ~26-fold higher than that for Tail Peptide A (Brignola et al., 2002). Additional studies with peptides derived from various proteins that are EGFR substrates as well as from synthetic peptide libraries have identified other peptide substrates for EGFR. Notably, the primary sequence of Peptide C is more hydrophobic than the other peptide substrates, which are predominantly acidic (Guyer et al., 1994; Songyang et al., 1995).

Owing to the increased catalytic efficiency of the EGFR kinase domain when Peptide C is used as a substrate, we sought to use Peptide C to develop robust biochemical assays for EGFR. The reported value of K_M of 128 μM for Peptide C permits EGFR activity to be measured at saturating peptide concentrations, which is not feasible using tail peptides for which the values of K_M are in the low millimolar range (Brignola et al., 2002; Fan et al., 2005). Several other studies have used Peptide C as an EGFR substrate peptide due to these favorable catalytic parameters (Wang et al., 2011; Wood et al., 2004; Yun et al., 2007).

Our studies confirm the remarkable activity of EGFR with Peptide C but also reveal unexpected properties that are likely to be undesirable in studies of EGFR enzymatic activity. We find that the value of k_{cat}/K_M for Peptide C in our assays is over 250-fold increased when directly compared to Tail Peptide A. We confirm that the high activity with Peptide C requires formation of the asymmetric dimer of EGFR kinase domains. Furthermore, the activity of EGFR with Peptide C as a substrate shows a steep dependence on the concentration of Peptide C, with a Hill coefficient of nearly 3. This behavior suggested to us initially that the peptide could be binding at an allosteric site on EGFR and thereby promoting kinase dimerization and increased activity. However, a detailed analysis of Peptide C found that it most likely stimulates EGFR activity by promoting aggregation of the kinase domain. Previous reports have described the ability of other molecules to aggregate and activate EGFR, and Peptide C represents the first substrate peptide of EGFR shown to have this effect (Hubler et al., 1992a; 1992b; Mohammadi et al., 1993).

3.2 Results and Discussion

3.2.1 Analysis of catalytic parameters and cooperativity for EGFR peptide substrates

We used a spectrophotometric enzyme-coupled kinase assay to characterize the kinetic parameters for Peptide C phosphorylation by EGFR in comparison to two other peptide substrates, Tail Peptide A and Tyrsub, a 13 residue peptide based on a sequence from the protein human erythrocyte Band 3 (Barker et al., 1995; Guyer et al., 1994). We were unable to reach saturating levels with Tail Peptide A and Tyrsub, and thus their catalytic parameters are estimates based on fits to the Michaelis-Menten equation (Figure 3.1). Peptide C does indeed stimulate EGFR kinase activity, with a catalytic rate constant, k_{cat}, of 0.160 s^{-1} that is roughly 8-fold higher than that for Tail Peptide A (0.021 s^{-1}) and Tyrsub (0.019 s^{-1}) (Table 3.1). More dramatically, the value of the Michaelis constant, K_M, for Peptide C (30 μM) is greatly decreased relative to the low millimolar values obtained for the peptides with natural sequences. Taken together, the catalytic efficiency (k_{cat}/K_M) for Peptide C is approximately 250- and 760-fold higher than those for Tail Peptide A and Tyrsub, respectively, confirming that Peptide C is a very efficient substrate for EGFR.

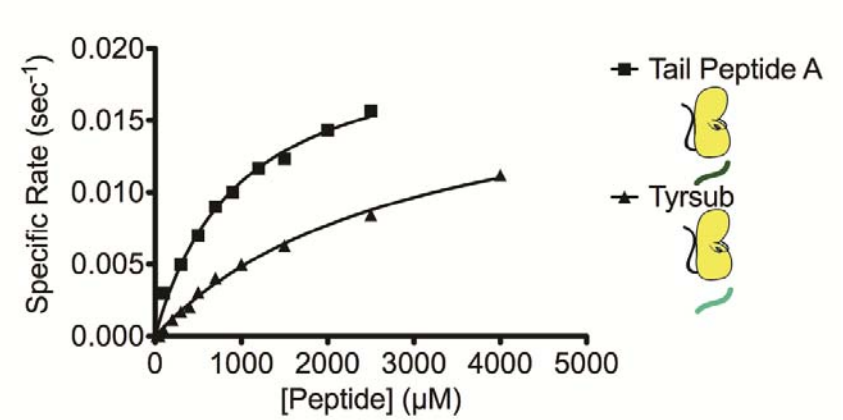


Figure 3.1 Michaelis-Menten fits for Tail Peptide A and Tyrsub. The specific rates of catalysis as a function of peptide concentration for the peptide substrates Tail Peptide A and Tyrsub are fit to the Michaelis-Menten equation to derive the catalytic parameters of k_{cat} and K_M (Table 3.1) using Prism software by GraphPad. Specific rates for each of the peptides at the indicated peptide concentrations were measured using the enzyme-coupled kinase assay with 8 μM EGFR kinase domain. Data points represent the means from a minimum of three replicates. Error bars plotting ± standard errors of mean are too small to be visible on this graph.

Peptide Substrate	k_{cat} (s^{-1})	K_M (μM)	$10^{-3}\times k_{cat}/K_M$ ($s^{-1}M^{-1}$)
Peptide C	0.160 ± 0.003	30 ±1	5.3
Tail Peptide A	0.021 ± 0.001	980 ± 80	0.021
Tyrsub	0.019 ± 0.001	2900 ± 300	0.007

Table 3.1 Kinetic parameters for EGFR kinase domain activity with EGFR peptide substrates. The specific rates of EGFR kinase domain catalysis as a function of peptide concentration were fit to the Michaelis-Menten equation to derive the catalytic rate constant k_{cat} and Michaelis constant K_M for Tail Peptide A and Tyrsub. Because the curve generated for specific rate as a function of peptide concentration for Peptide C is sigmoidal, the values of k_{cat} and K_M are instead derived from fitting the data to the Hill equation (Figure 3.2A). Specific rates for each of the peptides were measured in at least triplicate using the enzyme-coupled kinase assay with either 2 or 8 μM EGFR kinase domain for Peptide C or Tail Peptide A and Tyrsub, respectively. The values are ± standard errors of mean and generated from data points measured in at least triplicate.

The rate of EGFR-catalyzed phosphorylation rises steeply in response to increased Peptide C concentration, and fitting of the data to the Hill equation confirms positive cooperativity with a Hill coefficient of ~2.8 (Figure 3.2A). Conversely, Tail Peptide A and Tyrsub exhibit little or no cooperativity (Figure 3.2B,C). The observed cooperativity with Peptide C suggested that the peptide might achieve its high catalytic efficiency by a mechanism not employed by the other peptide substrates.

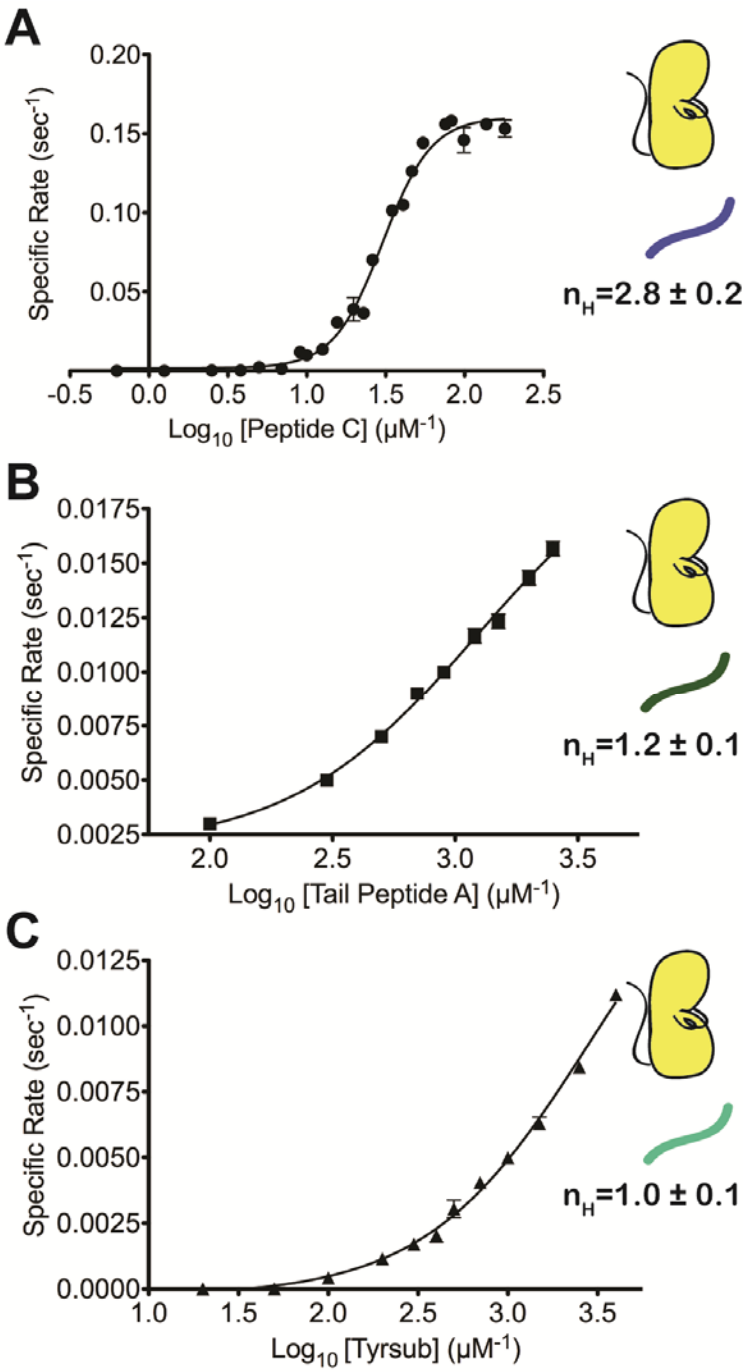


Figure 3.2 Hill coefficient measurements for EGFR kinase domain catalysis with EGFR peptide substrates. Hill coefficients for EGFR kinase domain catalysis with EGFR kinase peptide substrates (A) Peptide C, (B) Tail Peptide A, and (C) Tyrsub were derived by fitting the specific rates of catalysis at Log_{10} (peptide concentration) to the Hill equation using Prism software by GraphPad. Because the Tail Peptide A and Tyrsub measurements did not attain saturating concentrations, better fits for the Hill analysis were obtained when the Top (k_{cat}) values were constrained to those derived from the Michaelis-Menten fits (Table 3.1 and Figure 3.1). Specific rates for each of the peptides were measured using the enzyme-coupled kinase assay with either 2 or 8 μM

EGFR kinase domain for Peptide C or Tail Peptide A and Tyrsub, respectively. Data points represent the means and error bars are ± standard errors of mean from a minimum of three replicates. The error bars are too small to be visible for some data points. n_H values are ± standard errors of mean.

3.2.2 Peptide C stimulates phosphorylation of other substrates

Although the activation of the EGFR kinase domain relies principally on formation of the asymmetric dimer, phosphorylation of Tyr-845 in the activation loop of the kinase domain is also expected to increase activity (Khan et al., 2008). We wondered whether Peptide C could increase the catalytic activity of EGFR by increasing phosphorylation of Tyr-845. Immunoblotting with a phosphotyrosine-845 specific antibody confirmed that Tyr-845 phosphorylation increases in a dose-dependent manner with addition of Peptide C (Figure 3.3A,B). Surprisingly, the same dose-dependent increase in phosphorylation is observed for the tyrosines in the truncated C-terminal tail that is within the kinase domain construct, Tyr-974 and Tyr-992, indicating that Peptide C generally increases the rate of EGFR kinase domain autophosphorylation (Figure 3.3A,B). Phosphorylation of EGFR at Tyr-845 only results in a slight increase in *in vitro* catalytic activity, and so autophosphorylation at Tyr-845 does not account for the marked increase in activity observed with Peptide C (Figure 3.4).

The increase in autophosphorylation in the presence of Peptide C suggested that phosphorylation of other peptide substrates may also be promoted. To test this, we measured the phosphorylation of Tail Peptide A in a radiometric assay in the presence of a Peptide C variant in which the tyrosine was replaced with an alanine. This variant peptide cannot be phosphorylated by EGFR. Phosphorylation of Tail Peptide A by EGFR increases in response to the concentration of the Peptide C variant (Figure 3.3C). Interestingly, the activity profile with the Peptide C variant resembles that of Peptide C itself; the EC_{50} is 45 μM and a rough fit to the Hill equation yields a Hill coefficient of ~2.4 (Figure 3.3D). Thus, the ability of Peptide C to stimulate EGFR activity is apparently independent of its role as a substrate.

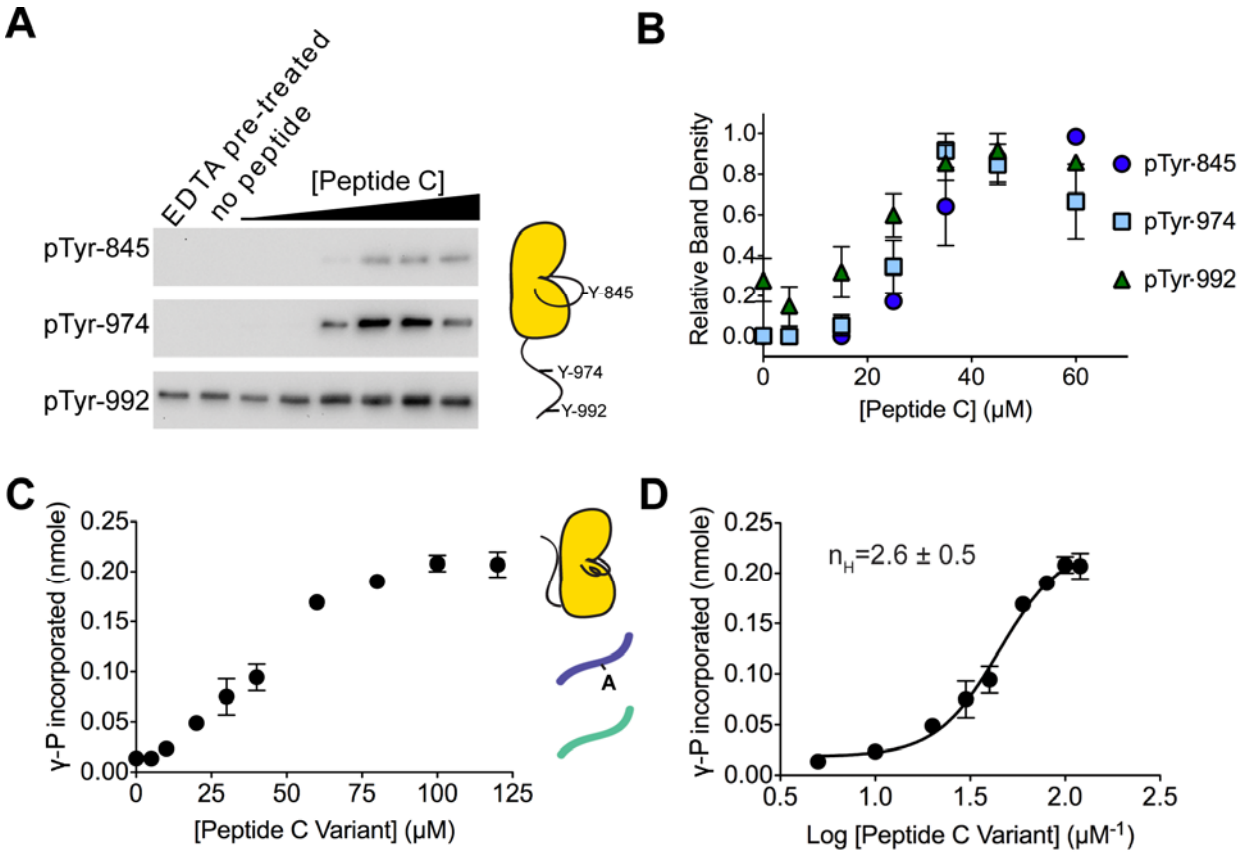


Figure 3.3 The effects of Peptide C on EGFR kinase domain phosphorylation of other substrates. Autophosphorylation of the EGFR kinase domain on the activation loop at Tyr-845 and the C-terminal tail at Tyr-974 and Tyr-992 in the presence of 0-60 μM Peptide C was measured by immunoblot and is shown by (A) representative immunoblots and (B) quantified band densities. Autophosphorylation reactions were performed with 800 nM EGFR kinase domain for one minute and terminated by addition of EDTA prior to immunoblotting with antibodies specific to EGFR and each phosphotyrosine. Band densities were quantified using ImageJ and normalized to EGFR kinase domain levels (Abramoff et al., 2004). Values are relative to the maximum levels of phosphorylation for each phosphotyrosine. Data points represent the averages ± standard errors of mean from a minimum of three replicates. (C) Phosphorylation of Tail Peptide A was measured in a radiometric kinase assay in the presence of a Peptide C variant in which the substrate tyrosine was replaced by alanine and could not undergo phosphorylation. (D) The phosphorylation of Tail Peptide A by EGFR kinase domain in the presence of a non-phosphorylatable Peptide C variant data are plotted as γ-phosphoryl (γ-P) incorporation as a function of Log_{10} (Peptide C variant concentration) and fit to the Hill equation. Reactions were performed for ten minutes with 500 nM EGFR kinase domain, 750 μM Tail Peptide A, Peptide C variant at indicated concentrations, and 25 μM ATP labeled at 8.325 μCi/ml [γ-^{32}P]ATP and terminated by addition of 0.5% v/v phosphoric acid. Reactions were spotted onto filters, and the radioactivity was quantified by liquid scintillation counting. Data points represent the means ± standard errors of mean from four replicate filters, and the value of n_H is ± one standard error of mean.

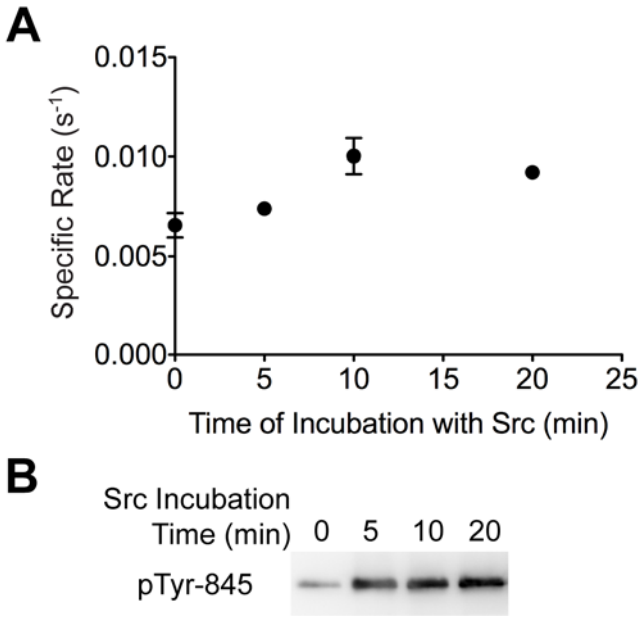


Figure 3.4 Effect of activation loop Tyr-845 phosphorylation on EGFR kinase domain catalytic activity. (A) EGFR kinase domain was incubated with Src kinase domain for the times indicated to increase phosphorylation on the EGFR kinase domain activation loop at Tyr-845, and the specific rates of catalysis of EGFR phosphorylation of Tail Peptide A by the pre-phosphorylated EGFR kinase domain were measured. (B) A representative immunoblot for phosphorylation of Tyr-845 indicates levels of EGFR kinase domain activation loop phosphorylation following incubation with Src and performance of the enzyme-coupled kinase reactions. Both Src-mediated phosphorylation and EGFR autophosphorylation contribute to phosphorylation of Tyr-845 *in vitro*. The Src-treated EGFR mixture was diluted into the enzyme-coupled kinase assay, and EGFR activity was measured with 500 μM Tail Peptide A and 8 μM EGFR at room temperature. The enzyme-coupled kinase reactions proceeded for approximately 20 minutes, and the reaction mixtures were subject to immunoblot analysis with an antibody specific to phosphorylation on Tyr-845. Data points represent the means ± standard errors of mean from three replicates.

3.2.3 Increased EGFR activity with Peptide C requires formation of the asymmetric dimer

We wondered whether Peptide C promotes the formation of the asymmetric dimer as a mechanism to increase the catalytic activity of EGFR. Such a mechanism could also underlie the observed positive cooperativity. To test whether asymmetric dimer formation is required, we utilized EGFR kinase domain constructs bearing mutations that block formation of the asymmetric dimer. One mutation, I682Q, is located on the N-lobe of the kinase domain and restricts the mutant kinase domain to serve only as an activator. The second mutation, V924R, is on the C-lobe and restricts the mutant kinase domain to being a receiver. When assayed independently, neither of these

mutants alone is sufficient to produce high activity. However, combining the two mutants at high effective concentrations permits reconstitution of the asymmetric dimer (Zhang et al., 2006).

The activating properties of Peptide C are entirely abrogated when assayed with either the I682Q or V924R mutants alone. However, when the two mutant EGFR kinase domains are combined, catalysis is regained to about 60% of that of the wild type kinase domain (Figure 3.5A). Therefore, the ability to form the asymmetric dimer is required for activation by Peptide C. In the EGFR asymmetric dimer only the receiver is activated, and this might explain why lower activity is observed with the combination of the V924R and I682Q kinase domains than for the wild type protein.

The requirements for catalytic activity of the activator and receiver kinase domains were tested by assaying kinase domains in which a catalytic residue, Asp-813, was mutated to asparagine (D813N), a mutation commonly made to inactivate kinases (Ge et al., 2002). The combination of the inactive mutant activator I682Q/D813N and competent receiver V924R exhibits nearly identical activity with Peptide C to the combination of I682Q and V924R (Figure 3.5B). Alternatively, inactive mutant receiver D813N/V924R paired with competent activator I682Q has no catalytic activity, as expected from the asymmetric dimer model. Taken together, these results indicate that the mechanism by which Peptide C stimulates EGFR catalysis is consistent with formation of the asymmetric dimer; both activator and receiver kinase domains are required for activation, but only the receiver undergoes activation (Zhang et al., 2006).

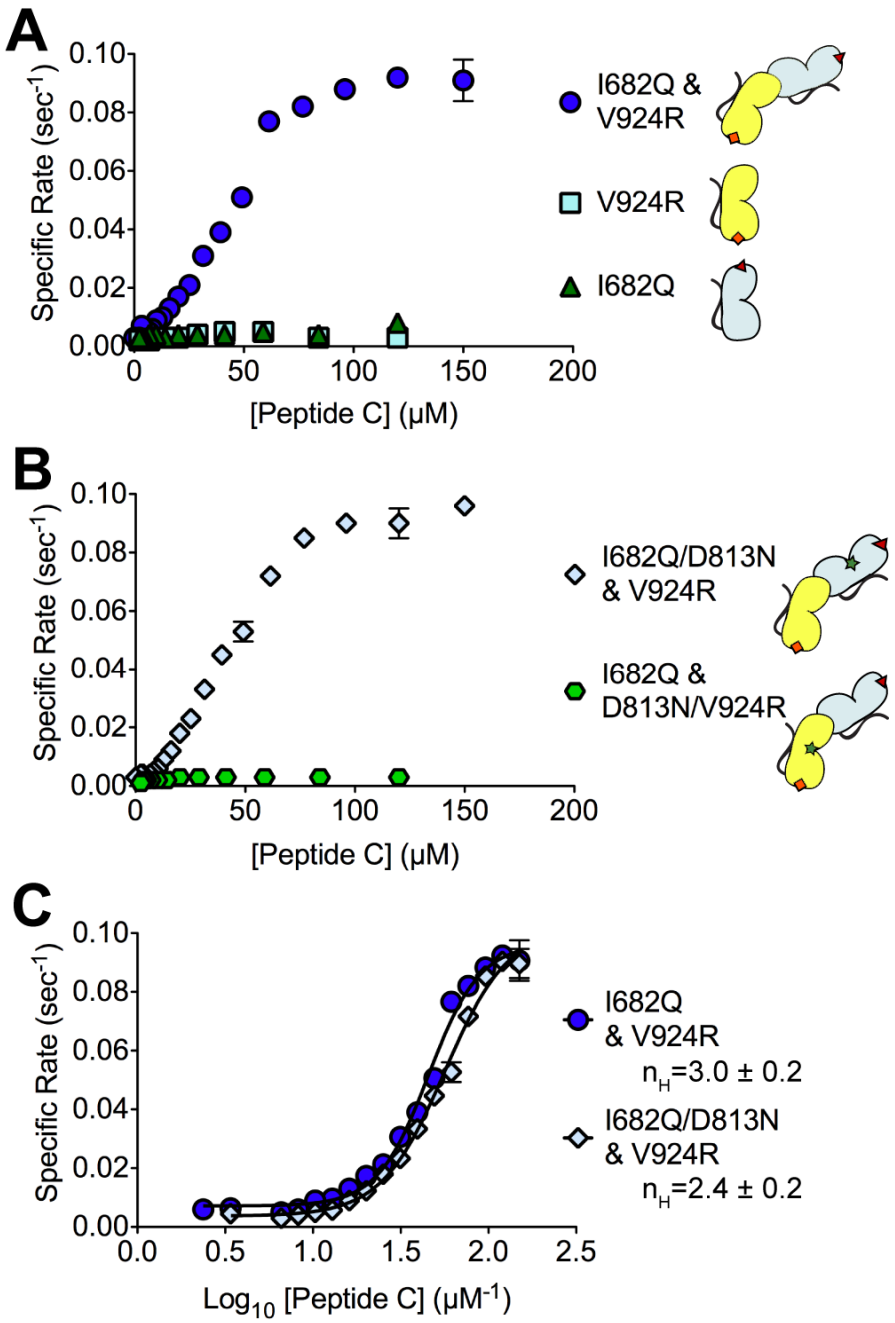


Figure 3.5 Catalytic activities of EGFR kinase domain dimerization mutants with Peptide C. (A) The specific rates of catalysis of EGFR kinase domain dimerization mutants I682Q, which is on the N-lobe, and V924R, which is on the C-lobe, either alone or combined were measured over a titration of Peptide C. (B) The kinase inactivating mutation D813N was introduced to each of the I682Q and V924R dimerization mutants, and specific rates for each of the kinase domain pairs I682Q/D813N with V924R or I682Q with D813N/V924R over a titration of Peptide C were measured. (C) Hill coefficients for the pairs of I682Q and V924R or I682Q/D813N and V924R were derived by fitting the specific catalytic rates from (A) and (B) at Log_{10} (peptide concentration) to the Hill equation using Prism software by GraphPad. Specific rates at indicated Peptide

C concentrations were measured using the enzyme-coupled kinase assay. All measurements were made using 2 μM of total kinase domain, and 1 μM of each of the mutant kinase domains when measured as a pair. Data points represent the means ± standard errors of mean from a minimum of three replicates, and N_H values are ± standard errors of mean.

3.2.4 The dimerization mutants exhibit similar cooperativity with Peptide C to wild type EGFR

The increases in specific rate in response to titration of Peptide C for the pairs of EGFR kinase domain dimerization mutants occur with high cooperativity, similar to what was observed with the wild type EGFR kinase domain. The Hill coefficients derived from fitting the data to the Hill equation are ~3.0 and ~2.5 for the V924R kinase domain assayed with the I682Q and I682Q/D813N kinase domains, respectively, and are similar to the Hill coefficient of ~2.8 for the wild type kinase domain (Figure 3C). This result is unexpected. We expected that the high cooperativity of the wild type kinase domain with Peptide C might result from the ability of the kinase domain to oligomerize as a chain through the asymmetric dimer interface as seen in crystal structures (Stamos et al., 2002; Zhang et al., 2006). The I682Q and V924R kinase domain mutants should not, however, form chains when combined, and so the high values of the Hill coefficients observed with these combinations could not be rationalized.

Aggregation may also give rise to steep dose response curves, and in this case high Hill coefficients are often transferable between assays and enzymes (Feng et al., 2007). Previous work has shown the abilities of several molecules to aggregate EGFR, and aggregation of EGFR correlates with increased kinase activity (Hubler et al., 1992a; 1992b; Mohammadi et al., 1993). We hypothesized that, likewise, Peptide C could increase catalytic efficiency by functioning as an aggregator of the EGFR kinase domain.

3.2.5 Peptide C precipitates the EGFR kinase domain

To test the capacity of Peptide C to aggregate EGFR, we incubated the kinase domain with Peptide C, pelleted the insoluble materials, and resolved the samples by SDS-PAGE. The kinase domain precipitates as a function of Peptide C concentration, and the Peptide C concentration dependence of precipitation resembles that for kinase activity (Figure 3.6A,D). The correlation of aggregation and activation of EGFR in the presence of Peptide C suggests two possibilities. Peptide C could promote the active conformation and consequently the aggregation of the kinase domain. EGFR kinase domain constructs that are more catalytically active also have a higher propensity to aggregate (Shan et al., 2012). Alternatively, Peptide C may aggregate EGFR, and the asymmetric dimer is stabilized as a byproduct of aggregation.

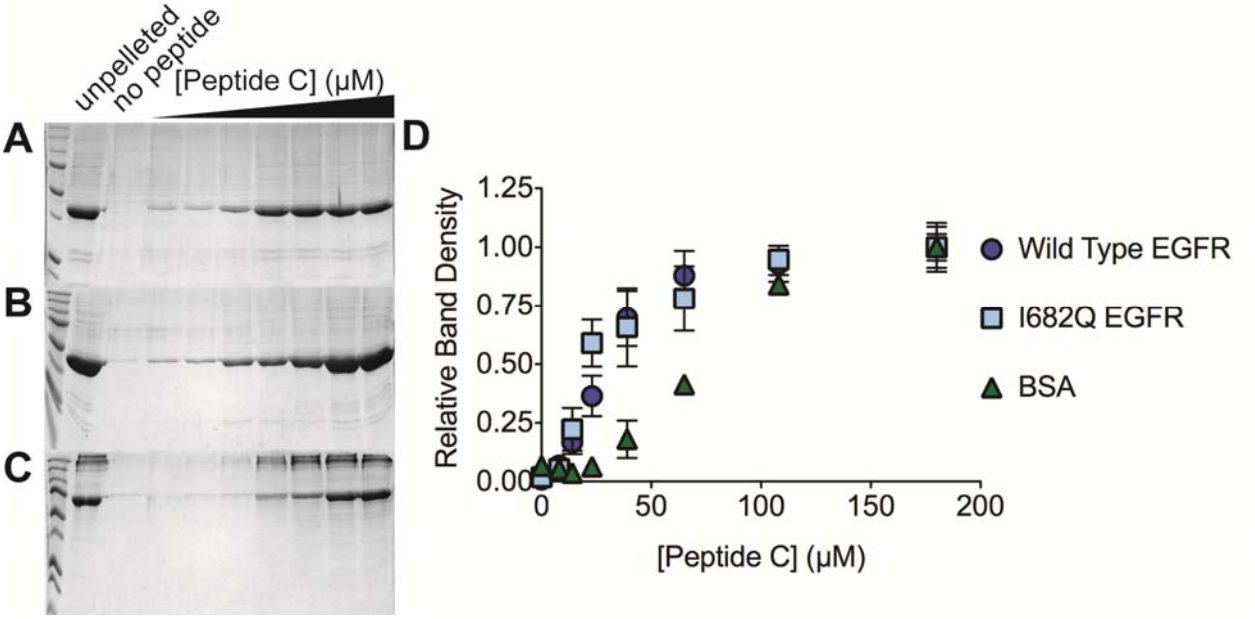


Figure 3.6 Precipitation of proteins by Peptide C. Co-sedimentation of (A) wild type EGFR kinase domain, (B) I682Q EGFR kinase domain, and (C) BSA in the presence of Peptide C is visualized by representative SDS-PAGE gels, and (D) band densities of the precipitated proteins are quantified. Each protein (2 μM) was incubated with 0-175 μM Peptide C for ten minutes and centrifuged, and the resuspended pellets were resolved by SDS-PAGE. Band densities were quantified using ImageJ (Abramoff et al., 2004). Data points represent the means ± standard errors of mean from a minimum of three replicates.

3.2.6 Peptide C precipitates proteins and forms aggregates on its own

To clarify the connection between activity and aggregation, we tested the ability of Peptide C to precipitate additional proteins. Peptide C precipitates the I682Q dimerization mutant EGFR kinase domain similarly to the wild type kinase domain (Figure 3.6B,D). This suggests that kinase activity is not required for precipitation and rather that activation occurs as result of aggregation. Furthermore, Peptide C appears to precipitate other proteins. BSA precipitates in the presence of Peptide C over the same concentration range of the peptide, although to a slightly lesser extent (Figure 3.6C,D). These data suggest that Peptide C precipitates proteins over a narrow concentration range and that this precipitation appears linked to the dimerization-mediated activation of the EGFR kinase domain.

We monitored the aggregation of Peptide C by itself in solution using dynamic light scattering (DLS) and found that the peptide alone is sufficient to form aggregates. A 1 mM solution of the peptide in water exhibited limited aggregation, and the hydrodynamic radius of the sample was almost entirely described by a polydisperse distribution of 1-10 nm. Preparation of 1 mM Peptide C in an assay buffer (15 mM Tris, pH 7.5, 0.02% v/v Polysorbate-20, and 20 mM sodium chloride) resulted in a highly

aggregated sample with very low mobility, which was indicated by a spike in laser intensity and minimal decay within the intensity autocorrelation decay function (Figure 3.7). A 1 mM Peptide C sample in 20 mM sodium chloride showed aggregation to an intermediate degree between the aqueous solution and the buffered samples. Thus Peptide C has a propensity to form large aggregates in solution, and aggregation of Peptide C appears to be increased by the addition of other solutes.

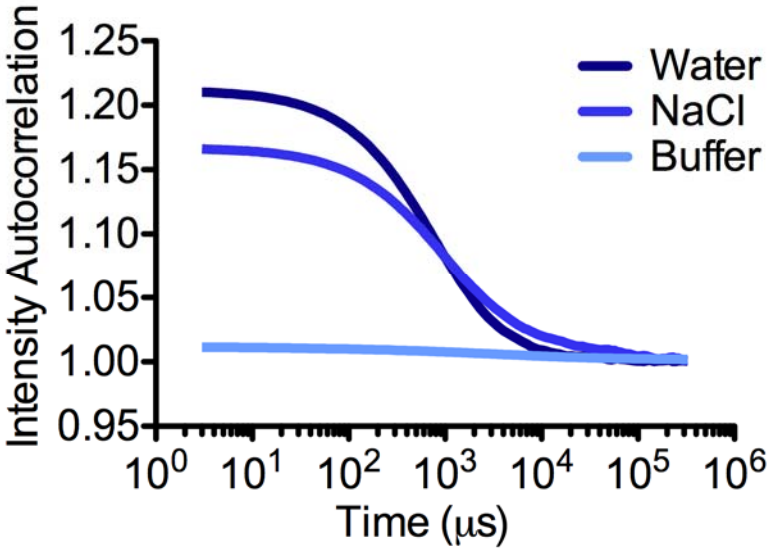


Figure 3.7 Dynamic light scattering autocorrelation curves for Peptide C samples. The autocorrelation curves for 1 mM samples of Peptide C in water, 20 mM sodium chloride, and buffer (15 mM Tris, pH 7.5, 0.02% v/v Polysorbate-20, and 20 mM sodium chloride) derived from dynamic light scattering measurements indicate Peptide C aggregation occurs in the presence of other solutes. Samples were measured at 25°C using 15% of the 60 mW laser power at a 90° angle on a DynaPro Titan instrument from Wyatt.

3.2.7 Electron microscopy shows that Peptide C forms fibrils

We examined samples of Peptide C in the absence and presence of EGFR kinase domain by transmission electron microscopy (TEM) to gain insight into the structures formed during aggregation. Samples of Peptide C at concentrations of 20, 40, and 80 μM all deposit as clumps of flexible fibrils that are several hundreds of nanometers in length (Figure 3.8A-C). The overall morphology of the fibrils appears to be independent of Peptide C concentration. Globular aggregates closely associated with the fibrils are also observed (Figure 3.8A-C, arrows). The TEM images show that Peptide C forms fibrils at concentrations relevant to activation of EGFR. Additionally, the globular aggregates formed by Peptide C and any monomeric, well-behaved Peptide C, which cannot be visualized by TEM, may also be important for activation of EGFR with Peptide C. The predominance and homogeneity of the fibrils suggests that if Peptide C

is present in a monomeric form, the transition of Peptide C to the fibrillar state is cooperative.

We also examined samples of Peptide C in the presence of the EGFR kinase domain to try to establish a link between Peptide C and EGFR kinase domain aggregation and activation. Fibrils similar to those formed by Peptide C alone as well as globular aggregates localized near the fibrils are observed in all samples containing Peptide C (Figure 3.8E,F). The addition of the EGFR kinase domain causes the edges of the fibrils to appear less crisp, suggesting that EGFR might associate with the surface of the fibrils. EGFR in samples without Peptide C forms numerous small aggregates on the grid, and these are also observed in the sample with 20 μM Peptide C (Figure 3.8D,E). These small, dispersed aggregates are not observed in samples of EGFR kinase domain with 80 μM Peptide C, which suggests that free EGFR becomes depleted from solution in the presence of increasing amounts of Peptide C (Figure 3.8F). This observation is consistent with our studies of EGFR precipitation as a function of Peptide C concentration.

Additionally, we examined samples of the I682Q EGFR kinase domain with Peptide C, and the samples appear indistinguishable from wild type (Figure 3.8G-I). This indicates the ability of EGFR to dimerize and activate does not grossly affect its interactions with Peptide C. The nature of the interaction between the EGFR kinase domain and Peptide C that underlies activation of the kinase domain remains unknown.

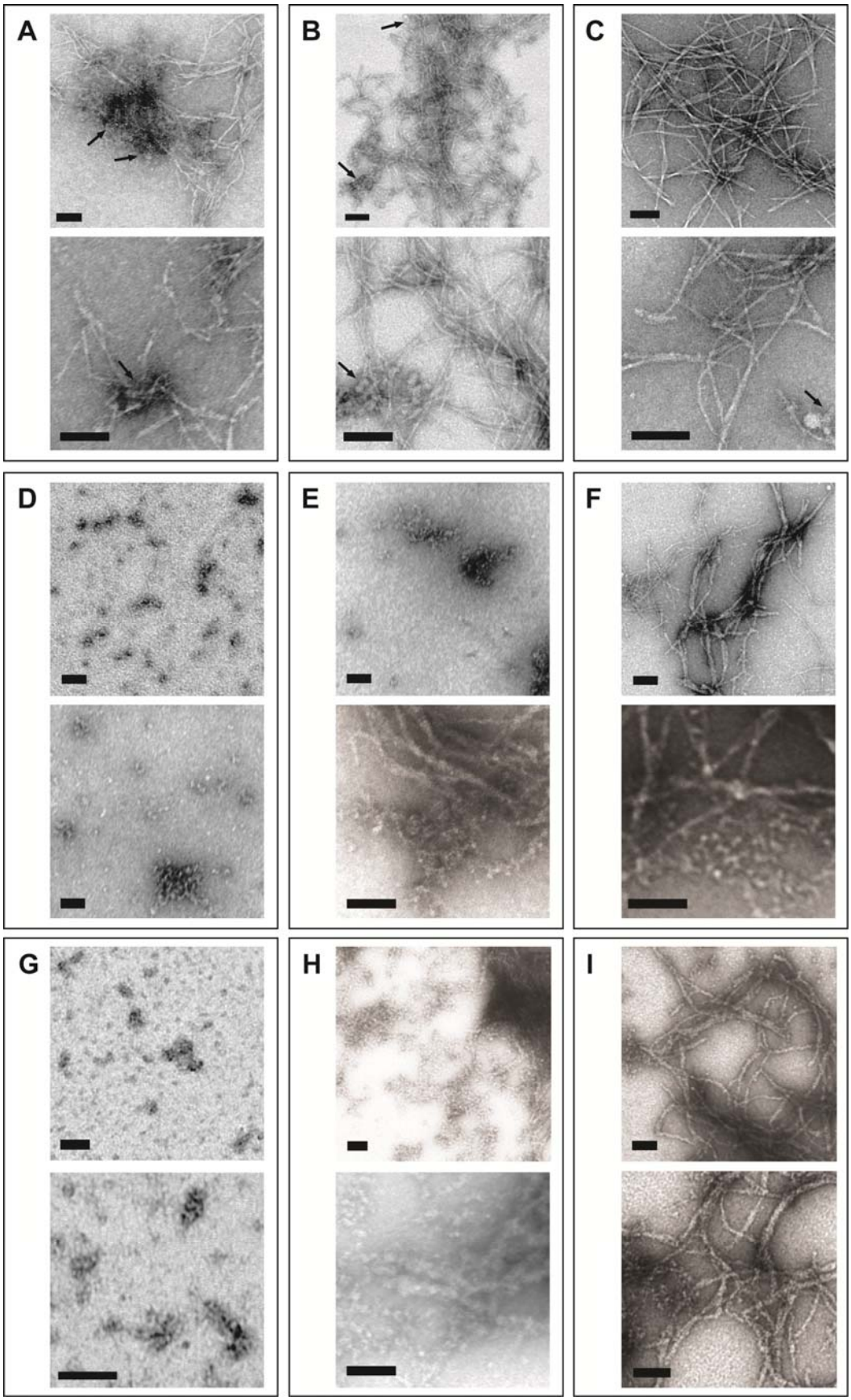

A
B
C
D
E
F
G
H
I

Figure 3.8 TEM images of Peptide C with and without EGFR kinase domain. Representative TEM images of (A) 20 μM, (B) 40 μM, and (C) 80 μM samples of Peptide C indicate Peptide C forms fibrils as well as globular aggregates, which are indicated by the arrows. Representative TEM images upon addition of 2 μM wild type EGFR kinase domain to (D) 0 μM, (E) 20 μM, and (F) 80 μM Peptide C indicate that Peptide C forms fibrils in the presence of EGFR. Representative TEM images upon addition of 2 μM I682Q EGFR kinase domain to (D) 0 μM, (E) 20 μM, and (F) 80 μM Peptide C indicate that Peptide C forms fibrils in the presence of an EGFR variant that cannot undergo dimerization. There appears to be lower levels of wild type and I682Q EGFR in solution at high Peptide C concentration, and the fibril edges are less defined. Samples were incubated for about 10 minutes prior to being deposited on Formvar carbon grids. The grids were negatively stained with 1% w/v uranyl acetate and viewed using a JEOL 1200 EX transmission electron microscope. Scale bars are 100 nm.

3.3 Conclusions

In this study, we have shown that Peptide C, a peptide selected for enhanced phosphorylation by EGFR, serves as a highly efficient substrate by aggregating EGFR. Aggregation appears to occur non-specifically, but because the activity of EGFR depends on the local concentration of the kinase domain, aggregation of EGFR by Peptide C results in formation of the asymmetric dimer and EGFR kinase domain activation. It has been shown previously that poly-lysine, poly-arginine, sphingosine, and protamine can aggregate and activate EGFR in an apparently cooperative manner, although it is not known if these molecules form macromolecular assemblies like those formed by Peptide C (Hubler et al., 1992b; Mohammadi et al., 1993). Peptide C appears to be the first such molecule, though, that serves as an EGFR substrate in addition to increasing kinase activity by an aggregation-mediated mechanism. All of these previously studied aggregators of EGFR are positively charged, and Peptide C contains an N-terminal arginine and three C-terminal lysines, suggesting that positive charge within the aggregating molecule may be required to aggregate EGFR.

The fibrils formed by Peptide C might provide a surface on which the EGFR can dimerize and activate. A similar mechanism has been observed for the allosteric activation of procaspase-3 upon binding to the surface of fibrils formed by a small molecule (Zorn et al., 2011). Our TEM data show that the edges of the fibrils are less clearly defined when the EGFR kinase domain is present, suggesting that the kinase domain may be associated with the surface of the fibrils.

Peptide C activates EGFR via formation of the asymmetric dimer. In one possible scenario, the fibrils formed by Peptide C might serve as a surface on which the EGFR kinase domain may bind to undergo aggregation and activation. This strategy resembles activation at the cell membrane, in which ligand binding to the full-length receptor causes dimerization and activation (Jorissen et al., 2003). This work studying Peptide C joins the growing body of evidence that recognizes the importance of EGFR

asymmetric dimer formation in a concentration dependent manner as the mechanism underlying kinase activation (Lu et al., 2012; Monsey et al., 2010; Zhang et al., 2006).

The phosphorylation of Peptide C by EGFR occurs with a desirable catalytic efficiency, and as a result other studies have utilized Peptide C as a substrate for EGFR (Wang et al., 2011; Wood et al., 2004; Yun et al., 2007). Despite the apparent excellent catalytic parameters of the peptide, Peptide C must be employed as a substrate with caution to avoid misinterpreting data as reflecting EGFR mechanism instead of the unusual properties of Peptide C.

3.4 Materials and methods

3.4.1 Protein expression and purification

DNA from the human EGFR gene spanning the kinase domain (672-998) was cloned into pFAST Bac HT (Invitrogen) using the NcoI and HindIII restriction sites. A N-terminal 6-His tag, linker, and TEV protease cleavage site (MSHHHHHHDYDIPTTENLYFQGAM) was included in the construct. Mutations to the kinase were generated using the Quikchange mutagenesis kit (Stratagene) and were confirmed by DNA sequencing. Recombinant bacmid (Bac-to-Bac expression system, Gibco BRL) were transfected into Sf9 cells for recombinant baculovirus production. Sf9 cells were grown in suspension and infected with the recombinant baculovirus. The cells were harvested after 2-3 days by centrifugation at 4000 x g and resuspended in lysis buffer of 50 mM Tris, pH 8.0, 5% v/v glycerol, 1 mM DTT, and protease inhibitor cocktail (Roche). The resuspended cells were lysed by French press, and the lysate was clarified by centrifugation at 40000 x g for one hour. The supernatant was loaded on a 60 ml Q-Sepharose Fastflow column (GE) that was equilibrated in buffer A (50 mM Tris, pH 8.0, 5% v/v glycerol, and 15 mM β-mercaptoethanol). The protein-bound column was washed with buffer A, and protein was eluted with buffer B (50 mM Tris, pH 8.0, 5% v/v glycerol, 1 M NaCl, and 7.5 mM β-mercaptoethanol). The eluted protein was loaded onto a 1 ml HisTrap FF column (GE) that was equilibrated with buffer C (20 mM Tris, pH 8.0, 5% v/v glycerol, 500 mM NaCl, 20 mM imidazole and 7.5 mM β-mercaptoethanol). The protein-bound column was washed with buffer C, and protein was eluted over a gradient of buffer C and buffer D (20 mM Tris, pH 8.0, 5% v/v glycerol, 500 mM NaCl, 250 mM imidazole and 7.5 mM β-mercaptoethanol). The protein was stored overnight at 4°C. The protein was diluted 10-fold into buffer E (20 mM Tris, pH 8.0, 5% v/v glycerol, 20 mM NaCl, and 2 mM DTT) and loaded onto a 6 ml Uno-Q column (Bio-Rad) that was equilibrated with buffer E. The protein-bound column was washed with buffer E, and the protein was eluted over a gradient of buffer E and buffer F (20 mM Tris, pH 8.0, 5% v/v glycerol, 1 M NaCl, and 2 mM DTT). The eluted protein was analyzed by SDS-PAGE, and fractions containing pure EGFR kinase domain were pooled, concentrated, and buffer exchanged into storage buffer (20 mM Tris, pH 8.0, 10% v/v glycerol, 50 mM NaCl, 2 mM TCEP and 2 mM DTT). Protein concentration was measured by amino acid analysis. The protein was flash frozen with liquid nitrogen and stored at -80°C.

3.4.2 Peptide substrates

The peptide substrates were produced by solid phase synthesis and purified by HPLC to a purity of at least 90% as analyzed by mass spectrometry. The sequence of Peptide C is RAHEEIYHFFFAKKK (Blackburn et al., 2003). The sequence of Tail Peptide A is RRKGSTAENAEYLRV and is derived from the Tyr-1173 autophosphorylation site of EGFR (Downward et al., 1985). The sequence of Tyrsub is EELEDDYEDDMEE and is derived from the protein human erythrocyte Band 3 (Guyer et al., 1994).

3.4.3 Enzyme-coupled kinase assays

The enzyme-coupled kinase assay was performed as described previously (Barker et al., 1995). EGFR kinase domain concentration was either 2 μM for assays with Peptide C or 8 μM for assays with Tail Peptide A or Tyrsub. In experiments in which pairs of different kinase domain dimerization mutants were assayed, each kinase domain was assayed at 1 μM for 2 μM total kinase domain concentration. Peptide concentrations were varied as indicated, and the concentration of ATP was 500 μM. The concentrations of coupled-kinase assay reagents were 1 mM phosphoenolpyruvate, 280 μM nicotinamide adenine dinucleotide, 89 units/ml pyruvate kinase, and 124 units/ml lactate dehydrogenase. Buffer conditions were 10 mM $MgCl_2$ and 20 mM Tris, pH 7.5. The final assay volume was 75 μl. Assays were performed in clear 96-well flat bottom plates, and the oxidation of NADH was measured by detecting decreases in absorbance at 340 nm. The rate of NADH oxidation was converted to the rate of ADP production using Beer's Law. Assays were performed at room temperature. Values for the kinetic parameters k_{cat} and K_M for Tail Peptide A and Tyrsub were derived from fits to the Michaelis-Menten equation using Prism, version 5.0c (Graphpad). Because the data for Peptide C fit better to a sigmoidal curve, the Top and EC_{50} values that were derived in the Hill analysis, which is described in the subsequent section, were used as the values for the k_{cat} and K_M, respectively, for EGFR catalysis with Peptide C.

3.4.4 Hill analysis

Hill coefficients were derived from fitting data to the Log(Agonist) vs. Response, Variable Slope equation defined as Y=Bottom + (Top-Bottom)/(1+10^((LogEC50-X)*HillSlope)) using Prism, version 5.0c (Graphpad), in which Y is the specific rate, Top and Bottom are maximum and minimum specific rates, respectively, EC50 is the concentration of peptide at which half of the maximum specific rate is attained, and X is the base-10 logarithm of the substrate concentration. Because the Tail Peptide A and Tyrsub analysis did not reach saturating concentrations, better fits for the Hill analysis were obtained when the Top (k_{cat}) values were constrained to those derived from the Michaelis-Menten fits.

3.4.5 Immunoblot analysis

Autophosphorylation reactions were performed with 800 nM EGFR kinase domain, 500 μM ATP, 10 mM $MgCl_2$, 20 mM Tris, pH 7.5, 400 μM NaF, 1 mM Na_3VO_4, and Peptide C at the indicated concentrations in 30 μl at room temperature. Reactions were terminated after one minute by addition of 20 mM EDTA. Samples were run on SDS-PAGE gels and subjected to immunoblot analysis. Levels of EGFR and autophosphorylation of EGFR were measured with primary antibodies specific to either EGFR (sc-03, Santa Cruz) or each of the phosphotyrosines at Tyr-845 (Cell Signaling), Tyr-974 (Cell Signaling), or Tyr-992 (Cell Signaling). The secondary antibody was goat anti-rabbit coupled to horseradish peroxidase (Cell Signaling). Band intensities were quantified using ImageJ and normalized to the level of EGFR for each sample (Abramoff et al., 2004).

3.4.6 Radiometric kinase assays

Reaction mixtures of 500 nM EGFR kinase domain, 750 μM Tail Peptide A, 25 μM ATP, 8.325 μCi/ml [γ-^{32}P]ATP, 15 mM Tris, pH 7.5, 0.02% v/v Polysorbate-20, and 10 mM $MgCl_2$ in a final volume of 60 μl were incubated for 10 minutes at room temperature. A variant of Peptide C with the sequence RAHEEIAHFFFAKKK was included at the indicated concentrations. Reactions were terminated by addition of 60 μl of 0.5% v/v phosphoric acid, and 55 μl of each quenched reaction was spotted on each of two P81 phosphocellulose filter discs (Whatman). The discs were filtered and washed three times with 0.5% v/v phosphoric acid using a vacuum manifold (Hoefer Scientific Instruments). Five ml of UniverSol scintillation cocktail (MP Biomedicals) were added to the dried discs in scintillation vials, and the radiation was measured by liquid scintillation counting (Beckman Coulter).

3.4.7 Src-mediated phosphorylation of EGFR

Src phosphorylation of EGFR was performed by incubating 200 nM Src kinase domain, 63.52 μM EGFR kinase domain, 500 μM ATP, 20 mM Tris, pH. 7.5, and 10 mM $MgCl_2$ for the indicated amounts of time at room temperature, and the reactions were inhibited by addition of 3.75 μM Dasatinib. This mix of EGFR, Src, and dasatinib was diluted into the enzyme-coupled kinase assay for a final EGFR concentration of 8 μM. Though the prior incubation of EGFR with Src results in addition of ~25 nM Src to the enzyme-coupled kinase assay, the 500 nM Dasatinib in the solution is sufficient to inhibit Src, and the enzyme-coupled kinase assay only measures EGFR catalysis.

3.4.8 Precipitation analysis

Samples of Peptide C at indicated concentrations and 2 μM WT EGFR kinase domain, I682Q EGFR kinase domain, or BSA (Sigma Aldrich) were incubated in 10 mM Tris, pH 7.5, 20 mM NaCl, and 0.02% v/v Polysorbate-20 in 75 μl total volume for about 20 minutes at room temperature. The samples were centrifuged at 15800 x g for ten minutes at 4° C. Supernatants were removed, and pelleted materials were resuspended

and resolved by SDS-PAGE. The bands were quantified by densitometry in ImageJ (Abramoff et al., 2004).

3.4.9 Dynamic light scattering measurements

One mM samples of Peptide C in water, 20 mM NaCl, or assay buffer (15 mM Tris, pH 7.5, 0.02% v/v Polysorbate-20, and 20 mM NaCl) were analyzed by DLS at 25° C using a DynaPro Titan (Wyatt). Measurements were made with the 60 mW laser at 15% power and at an angle of 90°. Buffer components were filtered prior to measurements. The filter retained Peptide C and as a result the peptide could not be filtered.

3.4.10 Transmission electron microscopy

Samples of Peptide C at indicated concentrations with and without 2 μM wild type or I682Q EGFR kinase domain were prepared in 10 mM Tris, pH 7.5, 20 mM NaCl, and 0.02% v/v Polysorbate-20 and incubated for approximately ten minutes at room temperature. Six μl of the samples were adsorbed on Formvar carbon-coated 400 mesh copper grids (Electron Microscopy Sciences), and the grids were negatively stained with three applications of 6 μl 1% w/v uranyl acetate. The grids were viewed with a JEOL 1200 EX transmission electron microscope.

References

Abramoff, M.D., Magalhães, P.J., and Ram, S.J. (2004). Image processing with ImageJ. Biophotonics Int. *11*, 36–42.

Avraham, R., and Yarden, Y. (2011). Feedback regulation of EGFR signalling: decision making by early and delayed loops. Nat. Rev. Mol. Cell Biol. *12*, 104–117.

Balzano, D., Santaguida, S., Musacchio, A., and Villa, F. (2011). A general framework for inhibitor resistance in protein kinases. Chem. Biol. *18*, 966–975.

Barker, S.C., Kassel, D.B., Weigl, D., Huang, X., Luther, M.A., and Knight, W.B. (1995). Characterization of $pp60^{c\text{-}src}$ tyrosine kinase activities using a continuous assay: autoactivation of the enzyme is an intermolecular autophosphorylation process. Biochemistry *34*, 14843–14851.

Berger, M.B., Mendrola, J.M., and Lemmon, M.A. (2004). ErbB3/HER3 does not homodimerize upon neuregulin binding at the cell surface. FEBS Lett. *569*, 332–336.

Betzi, S., Alam, R., Martin, M., Lubbers, D.J., Han, H., Jakkaraj, S.R., Georg, G.I., and Schönbrunn, E. (2011). Discovery of a potential allosteric ligand binding site in CDK2. ACS Chem. Biol. *6*, 492–501.

Biondi, R.M., Cheung, P.C., Casamayor, A., Deak, M., Currie, R.A., and Alessi, D.R. (2000). Identification of a pocket in the PDK1 kinase domain that interacts with PIF and the C-terminal residues of PKA. EMBO J. *19*, 979–988.

Biondi, R.M., Komander, D., Thomas, C.C., Lizcano, J.M., Deak, M., Alessi, D.R., and van Aalten, D.M.F. (2002). High resolution crystal structure of the human PDK1 catalytic domain defines the regulatory phosphopeptide docking site. EMBO J. *21*, 4219–4228.

Blackburn, R.K., Bramson, H.N., Moyer, M.B., and Stuart, J.D. (2003). Protein kinase peptide substrate determination using peptide libraries. US Patent Office, US Patent Application 2003.0113711.

Brignola, P.S., Lackey, K., Kadwell, S.H., Hoffman, C., Horne, E., Carter, H.L., Stuart, J.D., Blackburn, K., Moyer, M.B., Alligood, K.J., et al. (2002). Comparison of the biochemical and kinetic properties of the type 1 receptor tyrosine kinase intracellular domains. J. Biol. Chem. *277*, 1576–1585.

Bublil, E.M., Pines, G., Patel, G., Fruhwirth, G., Ng, T., and Yarden, Y. (2010). Kinase-mediated quasi-dimers of EGFR. FASEB J. *24*, 4744–4755.

Burnett, G., and Kennedy, E.P. (1954). The enzymatic phosphorylation of proteins. J. Biol. Chem. *211*, 969–980.

Busschots, K., Lopez-Garcia, L.A., Lammi, C., Stroba, A., Zeuzem, S., Piiper, A., Alzari, P.M., Neimanis, S., Arencibia, J.M., Engel, M., et al. (2012). Substrate-selective

inhibition of protein kinase PDK1 by small compounds that bind to the PIF-pocket allosteric docking site. Chem. Biol. *19*, 1152–1163.

Carpenter, G., and Cohen, S. (1979). Epidermal growth factor. Annu. Rev. Biochem. *48*, 193–216.

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

Chung, I., Akita, R., Vandlen, R., Toomre, D., Schlessinger, J., and Mellman, I. (2010). Spatial control of EGF receptor activation by reversible dimerization on living cells. Nature *464*, 783–787.

Clayton, A.H.A., Orchard, S.G., Nice, E.C., Posner, R.G., and Burgess, A.W. (2008). Predominance of activated EGFR higher-order oligomers on the cell surface. Growth Factors *26*, 316–324.

Clayton, A.H.A., Walker, F., Orchard, S.G., Henderson, C., Fuchs, D., Rothacker, J., Nice, E.C., and Burgess, A.W. (2005). Ligand-induced dimer-tetramer transition during the activation of the cell surface epidermal growth factor receptor-a multidimensional microscopy analysis. J. Biol. Chem. *280*, 30392–30399.

Cohen, P. (2000). The regulation of protein function by multisite phosphorylation-a 25 year update. Trends Biochem. Sci. *25*, 596–601.

Cohen, S. (1962). Isolation of a mouse submaxillary gland protein accelerating incisor eruption and eyelid opening in the new-born animal. J. Biol. Chem. *237*, 1555–1562.

Downward, J., Waterfield, M.D., and Parker, P.J. (1985). Autophosphorylation and protein kinase C phosphorylation of the epidermal growth factor receptor. Effect on tyrosine kinase activity and ligand binding affinity. J. Biol. Chem. *260*, 14538–14546.

Eckhart, W., Hutchinson, M.A., and Hunter, T. (1979). An activity phosphorylating tyrosine in polyoma T antigen immunoprecipitates. Cell *18*, 925–933.

Emde, A., Köstler, W.J., and Yarden, Y. (2012). Therapeutic strategies and mechanisms of tumorigenesis of HER2-overexpressing breast cancer. Crit. Rev. Oncol. Hematol. *84*, e49–e57.

Endres, N.F., Engel, K., Das, R., Kovacs, E., and Kuriyan, J. (2011). Regulation of the catalytic activity of the EGF receptor. Curr. Opin. Struct. Biol. *21*, 777–784.

Engel, M., Hindie, V., Lopez-Garcia, L.A., Stroba, A., Schaeffer, F., Adrian, I., Imig, J., Idrissova, L., Nastainczyk, W., Zeuzem, S., et al. (2006). Allosteric activation of the protein kinase PDK1 with low molecular weight compounds. EMBO J. *25*, 5469–5480.

Fan, Y.-X., Wong, L., and Johnson, G.R. (2005). EGFR kinase possesses a broad

specificity for ErbB phosphorylation sites, and ligand increases catalytic-centre activity without affecting substrate binding affinity. Biochem. J. *392*, 417–423.

Feng, B.Y., Simeonov, A., Jadhav, A., Babaoglu, K., Inglese, J., Shoichet, B.K., and Austin, C.P. (2007). A high-throughput screen for aggregation-based inhibition in a large compound library. J. Med. Chem. *50*, 2385–2390.

Ferguson, K.M. (2008). Structure-based view of epidermal growth factor receptor regulation. Annu. Rev. Biophys. *37*, 353–373.

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

Fischer, E.H., and Krebs, E.G. (1955). Conversion of phosphorylase b to phosphorylase a in muscle extracts. J. Biol. Chem. *216*, 121–132.

Gajiwala, K.S., Feng, J., Ferre, R., Ryan, K., Brodsky, O., Weinrich, S., Kath, J.C., and Stewart, Al (2012). Insights into the aberrant activity of mutant EGFR kinase domain and drug recognition. Structure *21*, 1–11.

Garrett, T.P.J., McKern, N.M., Lou, M., Elleman, T.C., Adams, T.E., Lovrecz, G.O., Zhu, H.-J., Walker, F., Frenkel, M.J., Hoyne, P.A., et al. (2002). Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor α. Cell *110*, 763–773.

Ge, G., Wu, J., Wang, Y., and Lin, Q. (2002). Activation mechanism of solubilized epidermal growth factor receptor tyrosine kinase. Biochem. Biophys. Res. Commun. *290*, 914–920.

Gotoh, N., Tojo, A., Hino, M., Yazaki, Y., and Shibuya, M. (1992). A highly conserved tyrosine residue at codon 845 within the kinase domain is not required for the transforming activity of human epidermal growth factor receptor. Biochem. Biophys. Res. Commun. *186*, 768–774.

Guy, P.M., Platko, J.V., Cantley, L.C., Cerione, R.A., and Carraway, K.L. (1994). Insect cell-expressed p180^{erbB3} possesses an impaired tyrosine kinase activity. P. Natl. Acad. Sci. U. S. A. *91*, 8132–8136.

Guyer, C.A., Woltjer, R.L., Coker, K.J., and Staros, J.V. (1994). Peptide substrate recognition by the epidermal growth factor receptor. Arch. Biochem. Biophys. *312*, 573–578.

Hindie, V., Stroba, A., Zhang, H., Lopez-Garcia, L.A., Idrissova, L., Zeuzem, S., Hirschberg, D., Schaeffer, F., Jørgensen, T.J.D., Engel, M., et al. (2009). Structure and allosteric effects of low-molecular-weight activators on the protein kinase PDK1. Nat. Chem. Biol. *5*, 758–764.

Huang, P.H., Xu, A.M., and White, F.M. (2009). Oncogenic EGFR signaling networks in glioma. Sci. Signal. *2*, re6.

Hubler, L., Kher, U., and Bertics, P.J. (1992a). Potentiation of epidermal growth factor receptor protein-tyrosine kinase activity by sulfate. Biochim. Biophys. Acta *1133*, 307–315.

Hubler, L., Leventhal, P.S., and Bertics, P.J. (1992b). Alteration of the kinetic properties of the epidermal growth factor receptor tyrosine kinase by basic proteins. Biochem. J. *281*, 107–114.

Hunter, T., and Sefton, B.M. (1980). Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. P. Natl. Acad. Sci. U. S. A. *77*, 1311–1315.

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

Hynes, N.E., and MacDonald, G. (2009). ErbB receptors and signaling pathways in cancer. Curr. Opin. Cell Biol. *21*, 177–184.

Jeffrey, P.D., Russo, A.A., Polyak, K., Gibbs, E., Hurwitz, J., Massagué, J., and Pavletich, N.P. (1995). Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. Nature *376*, 313–320.

Jorissen, R.N., Walker, F., Pouliot, N., Garrett, T.P.J., Ward, C.W., and Burgess, A.W. (2003). Epidermal growth factor receptor: mechanisms of activation and signalling. Exp. Cell Res. *284*, 31–53.

Jura, N., Endres, N.F., Engel, K., Deindl, S., Das, R., Lamers, M.H., Wemmer, D.E., Zhang, X., and Kuriyan, J. (2009a). Mechanism for activation of the EGF receptor catalytic domain by the juxtamembrane segment. Cell *137*, 1293–1307.

Jura, N., Shan, Y., Cao, X., Shaw, D.E., and Kuriyan, J. (2009b). Structural analysis of the catalytically inactive kinase domain of the human EGF receptor 3. P. Natl. Acad. Sci. U. S. A. *106*, 21608–21613.

Jura, N., Zhang, X., Endres, N.F., Seeliger, M.A., Schindler, T., and Kuriyan, J. (2011). Catalytic control in the EGF receptor and its connection to general kinase regulatory mechanisms. Mol. Cell *42*, 9–22.

Kaim, W. (2002). The coordination chemistry of 1,2,4,5-tetrazines. Coord. Chem. Rev. *230*, 127–139.

Khan, E.M., Lanir, R., Danielson, A.R., and Goldkorn, T. (2008). Epidermal growth factor receptor exposed to cigarette smoke is aberrantly activated and undergoes perinuclear trafficking. FASEB J. *22*, 910–917.

Khorchid, A., and Ikura, M. (2006). Bacterial histidine kinase as signal sensor and

transducer. Int. J. Biochem. Cell Biol. *38*, 307–312.

Knighton, D.R., Zheng, J.H., Eyck, Ten, L.F., Ashford, V.A., Xuong, N.H., Taylor, S.S., and Sowadski, J.M. (1991). Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. Science *253*, 407–414.

Kotra, S., Madala, K.K., and Jamil, K. (2008). Homology models of the mutated EGFR and their response towards quinazoline analogues. J. Mol. Graph. Model. *27*, 244–254.

Levene, P.A., and Alsberg, C.L. (1906). The cleavage products of vitellin. J. Biol. Chem. *2*, 127–133.

Levitzki, A., and Gazit, A. (1995). Tyrosine kinase inhibition: an approach to drug development. Science *267*, 1782–1788.

Lipinski, C.A., Lombardo, F., Dominy, B.W., and Feeney, P.J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Delivery Rev. *23*, 3-25.

Liu, Y., and Gray, N.S. (2006). Rational design of inhibitors that bind to inactive kinase conformations. Nat. Chem. Biol. *2*, 358–364.

Lopez-Garcia, L.A., Schulze, J.O., Fröhner, W., Zhang, H., Süss, E., Weber, N., Navratil, J., Amon, S., Hindie, V., Zeuzem, S., et al. (2011). Allosteric regulation of protein kinase PKCζ by the N-terminal C1 domain and small compounds to the PIF-pocket. Chem. Biol. *18*, 1463–1473.

Lu, C., Mi, L.-Z., Schürpf, T., Walz, T., and Springer, T.A. (2012). Mechanisms for kinase-mediated dimerization of the epidermal growth factor receptor. J. Biol. Chem. *287*, 38244–38253.

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

Mohammadi, M., Honegger, A., Sorokin, A., Ullrich, A., Schlessinger, J., and Hurwitz, D.R. (1993). Aggregation-induced activation of the epidermal growth factor receptor protein tyrosine kinase. Biochemistry *32*, 8742–8748.

Monsey, J., Shen, W., Schlesinger, P., and Bose, R. (2010). Her4 and Her2/neu tyrosine kinase domains dimerize and activate in a reconstituted in vitro system. J. Biol. Chem. *285*, 7035–7044.

Muthuswamy, S.K., Gilman, M., and Brugge, J.S. (1999). Controlled dimerization of ErbB receptors provides evidence for differential signaling by homo- and heterodimers. Mol. Cell. Biol. *19*, 6845–6857.

Ogiso, H., Ishitani, R., Nureki, O., Fukai, S., Yamanaka, M., Kim, J.-H., Saito, K., Sakamoto, A., Inoue, M., Shirouzu, M., et al. (2002). Crystal structure of the complex of

human epidermal growth factor and receptor extracellular domains. Cell *110*, 775–787.

Olayioye, M.A., Neve, R.M., Lane, H.A., and Hynes, N.E. (2000). The ErbB signaling network: receptor heterodimerization in development and cancer. EMBO J. *19*, 3159–3167.

Pawson, T., and Scott, J.D. (2005). Protein phosphorylation in signaling-50 years and counting. Trends Biochem. Sci. *30*, 286–290.

Pines, G., Köstler, W.J., and Yarden, Y. (2010). Oncogenic mutant forms of EGFR: lessons in signal transduction and targets for cancer therapy. FEBS Lett. *584*, 2699–2706.

Red Brewer, M., Choi, S.H., Alvarado, D., Moravcevic, K., Pozzi, A., Lemmon, M.A., and Carpenter, G. (2009). The juxtamembrane region of the EGF receptor functions as an activation domain. Mol. Cell *34*, 641–651.

Sadowsky, J.D., Burlingame, M.A., Wolan, D.W., McClendon, C.L., Jacobson, M.P., and Wells, J.A. (2011). Turning a protein kinase on or off from a single allosteric site via disulfide trapping. P. Natl. Acad. Sci. U. S. A. *108*, 6056–6061.

Santos, J.L., and Shiozaki, K. (2001). Fungal histidine kinases. Sci. STKE *2001*, re1.

Savelli, G., Spreti, N., and Di Profio, P. (2000). Enzyme activity and stability control by amphiphilic self-organizing systems in aqueous solutions. Curr. Opin. Colloid Interface Sci. *5*, 111–117.

Scherbak, N., Strid, Å., and Eriksson, L.A. (2005). Non-enzymatic oxidation of NADH by quinones. Chem. Phys. Lett. *414*, 243–247.

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

Shan, Y., Eastwood, M.P., Zhang, X., Kim, E.T., Arkhipov, A., Dror, R.O., Jumper, J., Kuriyan, J., and Shaw, D.E. (2012). Oncogenic mutations counteract intrinsic disorder in the EGFR kinase and promote receptor dimerization. Cell *149*, 860–870.

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

Sibilia, M., and Wagner, E. (1995). Strain-dependent epithelial defects in mice lacking the EGF receptor. Science *269*, 234–238.

Sierke, S.L., Cheng, K., Kim, H.H., and Koland, J.G. (1997). Biochemical characterization of the protein tyrosine kinase homology domain of the ErbB3 (HER3) receptor protein. Biochem. J. *322*, 757–763.

Simeonov, A., Jadhav, A., Thomas, C.J., Wang, Y., Huang, R., Southall, N.T., Shinn, P., Smith, J., Austin, C.P., Auld, D.S., et al. (2008). Fluorescence spectroscopic profiling of

compound libraries. J. Med. Chem. *51*, 2363–2371.

Songyang, Z., Carraway, K.L., Eck, M.J., Harrison, S.C., Feldman, R.A., Mohammadi, M., Schlessinger, J., Hubbard, S.R., Smith, D.P., Eng, C., et al. (1995). Catalytic specificity of protein-tyrosine kinases is critical for selective signalling. Nature *373*, 536–539.

Stamos, J., Sliwkowski, M.X., and Eigenbrot, C. (2002). Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. J. Biol. Chem. *277*, 46265–46272.

Threadgill, D.W., Dlugosz, A.A., Hansen, L.A., Tennenbaum, T., Lichti, U., Yee, D., LaMantia, C., Mourton, T., Herrup, K., Harris, R.C., et al. (1995). Targeted disruption of mouse EGF receptor: effect of genetic background on mutant phenotype. Science *269*, 230–234.

Tzahar, E., Waterman, H., Chen, X., Levkowitz, G., Lavi, S., Ratzkin, B.J., and Yarden, Y. (1996). A hierarchical network of interreceptor interactions determines signal transduction by Neu differentiation factor/neuregulin and epidermal growth factor. Mol. Cell. Biol. *16*, 5276.

Ushiro, H., and Cohen, S. (1980). Identification of phosphotyrosine as a product of epidermal growth factor-activated protein kinase in A-431 cell membranes. J. Biol. Chem. *255*, 8363–8365.

Wang, Z., Longo, P.A., Tarrant, M.K., Kim, K., Head, S., Leahy, D.J., and Cole, P.A. (2011). Mechanistic insights into the activation of oncogenic forms of EGF receptor. Nat. Struct. Mol. Biol. *18*, 1388–1393.

Wei, L., Gao, X., Warne, R., Hao, X., Bussiere, D., Gu, X.-J., Uno, T., and Liu, Y. (2010). Design and synthesis of benzoazepin-2-one analogs as allosteric binders targeting the PIF pocket of PDK1. Bioorg. Med. Chem. Lett. *20*, 3897–3902.

Wood, E.R., Truesdale, A.T., McDonald, O.B., Yuan, D., Hassell, A., Dickerson, S.H., Ellis, B., Pennisi, C., Horne, E., Lackey, K., et al. (2004). A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells. Cancer Res. *64*, 6652–6659.

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

Yun, C.-H., Boggon, T.J., Li, Y., Woo, M.S., Greulich, H., Meyerson, M., and Eck, M.J. (2007). Structures of lung cancer-derived EGFR mutants and inhibitor complexes: mechanism of activation and insights into differential inhibitor sensitivity. Cancer Cell *11*, 217–227.

Zhang, J., Chung, T., and Oldenburg, K. (1999). A simple statistical parameter for use in

evaluation and validation of high throughput screening assays. J. Biomol. Screen. *4*, 67–73.

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

Zhang, X., Pickin, K.A., Bose, R., Jura, N., Cole, P.A., and Kuriyan, J. (2007). Inhibition of the EGF receptor by binding of MIG6 to an activating kinase domain interface. Nature *450*, 741–744.

Zorn, J.A., Wille, H., Wolan, D.W., and Wells, J.A. (2011). Self-assembling small molecules form nanofibrils that bind procaspase-3 to promote activation. J. Am. Chem. Soc. *133*, 19630–19633.