4-28-2017

Quantitative Analysis of EGFR Phosphorylation and SH2 Domain Binding in vivo

JOSHUA A. JADWIN
University of Connecticut, joshuajadwin@gmail.com

Quantitative Analysis of EGFR Phosphorylation and SH2 Domain Binding *in vivo*

Joshua Adam Jadwin, M.D, PhD

University of Connecticut, 2017

Abstract

The work presented in the following thesis dissertation examines the regulation of phosphotyrosine (pY) signaling, an essential cellular process that relies on the activity of three major protein classes: Tyrosine kinases (TKs) which induce pY signaling by phosphorylating tyrosine residues on substrate proteins, Protein tyrosine phosphatases (PTPs) which suppress pY signaling by removing phosphate moieties from tyrosine phosphorylated proteins and Src-Homology 2 (SH2) containing proteins which bind to tyrosine phosphorylated proteins and connect them to downstream signaling pathways. The effects of kinase localization, temporal changes in kinase activation, SH2 protein concentration, and negative feedback from downstream signaling pathways are all examined by the research presented here. This is accomplished by exploiting the Epidermal Growth Factor Receptor (EGFR), a clinically important transmembrane TK, and its SH2 protein mediated downstream pathways. Using EGFR signaling as a tool, this dissertation research attempts to define innate properties of pY signaling systems which are broadly applicable and advance our understanding of the field.

Quantitative Analysis of EGFR Phosphorylation and SH2 Domain Binding *in vivo*

Joshua Adam Jadwin

B.S., State University of New York at Geneseo, 2005

A Dissertation
Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy
at the
University of Connecticut

2017

APPROVAL PAGE

Doctor of Philosophy Dissertation

Quantitative Analysis of EGFR Phosphorylation and SH2 Domain Binding *in vivo*

Presented by
Joshua A. Jadwin, B.S.

Major Advisor

_____

Bruce J. Mayer

Associate Advisor

_____

Arthur Gunzl

Associate Advisor

_____

Sandra Weller

Associate Advisor

_____

Kevin Claffey

Associate Advisor

_____

Daniel Rosenberg

University of Connecticut
2017

Acknowledgments:

This thesis dissertation represents not only my own academic efforts, but those of the family members, co-workers, faculty, and mentors who have inspired, encouraged, taught, challenged and believed in me. Its very existence is a testament to the quality of their characters and this work is dedicated to them.

Firstly I would like to thank Dr. Bruce Mayer, a mentor, colleague and friend, and one of the few individuals whom I look to as an example on which to model my own career. He is one of the most understated, kind, thoughtful and intelligent individuals I have ever had to the chance to work with. It has been an absolute honor to work with him in his lab. I would like to thank him for his never ending patience, curiosity, and optimism, and for allowing me pursue my own scientific endeavors and supporting me in these pursuits. Knowing him and working with him have made me both a better scientist and a better person.

I would next like to express my appreciation to the members of the Mayer Lab. First, to Dr. Kazuya Machida for his mentorship, for late nights spent writing and editing, and for graciously providing me with opportunities for academic collaboration. His work ethic and passion for science are an inspiration. I also thank Dr. Silas Ng, for many late night discussions, both scientific and otherwise, and for his friendship. Adam Lafontaine, for introducing me to computational biology, for his off topic and interesting discussions, and his camaraderie. Dr. Jonathan Ditlev, for his scientific insights, and modeling the balance of family and a successful scientific career. Dr. Mari Ikeda for her humor and her scientific rigor. Dr. Lin Jia for her patience and understated scientific acumen. Dr. Sofya Borinskaya for her insightful comments during lab meetings and for her help with computational biology. I am also grateful for my many students, who inspired my passion for teaching and through whom I learned to understand and critique my own work more effectively. I also thank my collaborators, Dr Ji Yu and Dr. Forest

White for their scientific insight and always taking the time to review, edit and often re-edit my manuscripts. I am also especially indebted to their students, Dr. Dongmyung Oh for his work on single molecule imaging and Dr. Timothy Curran for his work on quantitative mass spectrometry. Without their efforts, much of the work presented here would not be possible.

I thank my thesis committee members, Dr. Sandra Weller, Dr. Daniel Rosenberg, Dr. Arthur Gunzl and Dr.Kevin Claffey for their patience with my seemingly endless administrative issues and for their scientific insights. I chose them for their scientific rigor and then did not disappoint. I would also like to thank faculty, students and staff of the Department of Genetics and Developmental Biology, especially Dr. James Li, Dr. Blanka Rogina, Isolde Bates, and Gail Diamico, for their continued support, both academic and administrative, over the course of my tenure in the lab. I would also like to thank Andy for his conversation and friendship. I am also very grateful to Dr. Barbara Kream, Dr. Carol Pilbeam, Tracy Diely, and the entire University of Connecticut MD/PhD program for their patience and support over the entirety of my nine years.

I would also like to express my appreciation to the mentors and teachers who fostered my passion for science before my time at UConn. My many science teachers at Wellsville middle and high school, as well as Mrs. Figura, for fostering my interest in science. Drs Wendy Pogozelski and Harold Hoops at SUNY Geneseo for their love of science and their passion for teaching. Dr. Fadila Bouamr at the NIH NAID for taking a chance on me and giving me the intellectual space I needed to develop as a scientist.

I want to express my deepest gratitude to my parents, Bradley and Marguerite Jadwin, for always supporting me, for believing in me, for staying up all night with me to work on middle school science fair projects, for inspiring my love of science through trips to museums and aquariums, and for always challenging me to do my best. They are my role models in life, and

though they would like to believe otherwise, I owe everything I am to them. I want to thank my brother Bradely Jadwin Jr. for being an example of perseverance in the face of adversity, my brother Luke Jadwin for showing me how to stand up for myself without putting others down, and my sister Meghan Jadwin for being the glue that keeps our family together,

Finally and most importantly, I want to thank my beautiful and brilliant wife Naima Joseph. Thank you so much for your patience, for always supporting me and loving me across five years and two states. Your unyielding will and grace in the face of adversity never cease to amaze and inspire me. Thank you for being my partner and friend throughout this journey.

Table of Contents: **Page #**

Introduction

Assessing Extracellular Signals and Phosphotyrosine Dynamics

Multicellular organisms have evolved a number elegant mechanisms capable of interpreting the extracellular environment and translating into the language of cellular signaling (*1*). Typically extracellular ligands bind to cellular receptors, inducing allosteric changes within proteins which either directly induce enzymatic activity, or promote interactions which lead to receptor activation and/ or the recruitment of downstream enzyme effectors. These signal traducers include, but are not limited to G-protein coupled receptors, steroid hormone receptors, immune cell receptors, and receptor tyrosine kinases (RTKs). Each of these transducers classes encompasses a large number proteins which activate numerous molecular pathways, a complete discussion of which is well beyond the scope of any individual thesis project.

Protein signaling systems involved in inducing and regulating post-translational modification (PTMs) such as phosphorylation, acetylation, methylation and to potentially sumoylation and ubiquinylation can be distilled into three major components; Writers, Readers and Erasers. Writers function to conjugate PTMs to their target proteins, typically inducing signal transduction. Readers contain PTM-binding modular domains which translate PTMs into downstream signaling. Erasers remove PTMs, dampening or even preventing signaling from specific PTM mediated pathways(*2*). Put simply, total signal output is a function of the balance between Writer and Eraser activity, and Reader binding. The work presented here focuses phosphotyrosine (pY) mediated signaling in which signal output is defined by the interplay between tyrosine kinases (Writers), protein tyrosine phosphatases (Erasers) and Src-Homology 2 (SH2) domain containing proteins (Readers) (Figure i 1). Specifically we focus on how the intracellular concentration, localization and activities of each component can modify signal output in unexpected ways.

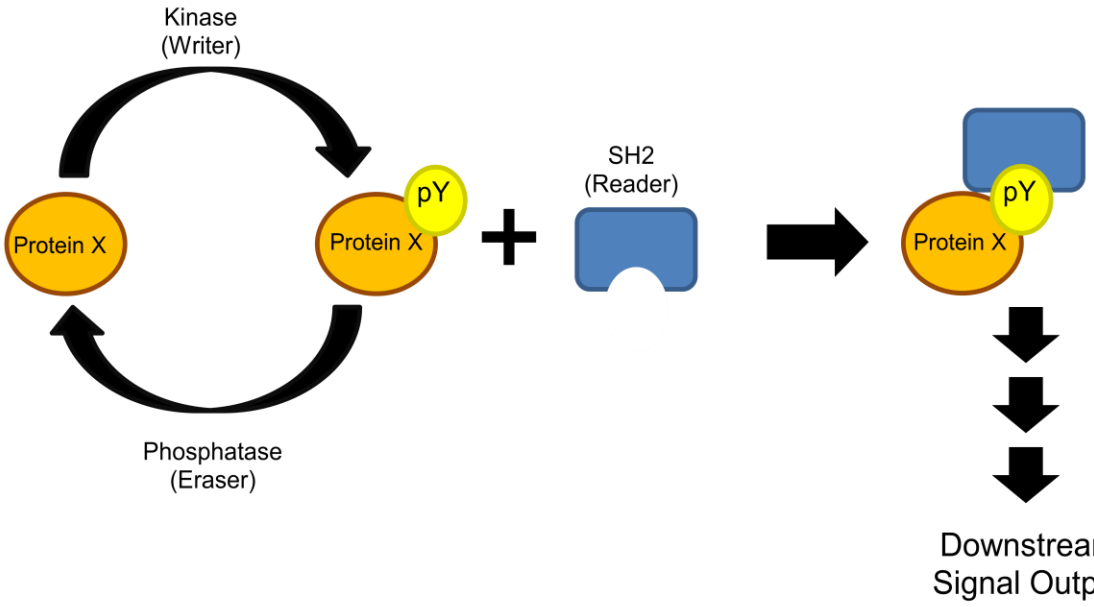

Figure i 1. Writer, Reader, Eraser model of phosphotyrosine signaling. Phosphotyrosine signaling is a classic example of the "writer, reader, eraser" signaling paradigm originally devised for chromatin modification signaling. Tyrosine kinase writers catalyze the transfer of the γ-PO_4 from ATP to a tyrosine residue on substrate proteins. These phosphorylated tyrosines function as docking sites for SH2 domain containing reader proteins which translate the phosphorylation of specific tyrosines into specific downstream signaling pathways. Phosphotyrosine signals are tempered by protein tyrosine phosphates (PTP) which catalyze the removal of phosphotyrosine residues from substrate proteins not bound to SH2 domains. The balance of kinase and PTP activities define the level protein tyrosine phosphorylation. Figure adapted B. J. Mayer, Perspective: Dynamics of receptor tyrosine kinase signaling complexes. *FEBS Lett.* **586**, 2575–9 (2012).

Tyrosine Kinases

Receptor tyrosine kinases are transmembrane proteins belonging to the larger family of approximately 90 tyrosine kinases (TKs) and 43 tyrosine kinase-like proteins (*3*–*5*). These proteins can be transmembrane, cytoplasmic and/or membrane associated. On the cellular level TK activation typically promotes growth, differentiation and proliferation. Specific TK driven processes include immune cell activation, vasculogenesis, tissue/growth and repair, homeostatic signaling and transmission of mechanical stresses from the cytoskeleton to the cellular signaling machinery, among many others (*6*–*10*). TK signaling is also arguably the most important driver of cellular transformation and tumor growth, with most tumors displaying excess TK activation or activation of a TK driven pathway(*11*–*14*). In fact, the non-receptor tyrosine kinase, v-SRC was the first oncogene discovered and shortly thereafter v-SRC activity was linked to tyrosine phosphorylation by studies which connected SRC expression with tyrosine phosphorylation in polyoma virus infected cells and polyoma virus middle-T protein tyrosine phosphorylation (*15*, *16*).

Tyrosine kinases function by catalyzing transfer of the γ-phosphate (-PO3) of ATP to the hydroxyl group on the amino acid tyrosine (*17*). Generally, TKs display relatively low basal activity when compared to their activated state which is induced by two major mechanisms, activation loop phosphorylation and in the case of RTKs, ligand binding (*3*, *18*–*20*). Most TKs contain a short unstructured loop within their kinase domain containing one or more tyrosines. Phosphorylation of these tyrosine residues induces structural changes within the kinase domain that allow for substrate binding and phosphorylation. RTK activation and signaling requires a second layer of regulation induced by the binding of a small protein ligand (also known as a mitogen or growth factor) to the extracellular domain. Ligand binding promotes allosteric changes in the ligand binding domain of the protein which promote receptor dimerization. Dimerization then significantly increases the chance activation loop phosphorylation.

Dimerized, activated RTKs then transphosphorylate substrate tyrosines on their binding partners. In either case, phosphorylated sites serve as docking sites for downstream signaling proteins, which are themselves often phosphorylated (see below). Isolated RTKs generally display only moderate substrate specificity, defined by the amino acid sequence surrounding the substrate tyrosine (*21*). Additional binding domains within the kinases provide further specificity by facilitating kinase-substrate interactions and subcellular localization (*22*).

Protein Tyrosine Phosphatases

The activity of tyrosine kinases is opposed by the action of protein tyrosine phosphatases (PTPs). PTPs are a diverse and rather understudied family of proteins which include both classical tyrosine specific phosphatases and dual specificity phosphates capable of dephosphorylating phosphotyrosine, phosphoserine and phosphothreonine residues. Classical PTPs which include 38 cytoplasmic and membrane-associated members, have been implicated as the major negative regulators of tyrosine phosphorylation. For the purposes of this work the acronym “PTP” will refer only to the classical, tyrosine specific PTPs (*23*, *24*).

All PTP family members contain the highly conserved active site motif [I/V]HCSXGXGR[S/T]G, which is essential for their function. A major mechanism of PTP regulation is the reversible oxidation of the active site cysteine. When oxidized, the cysteine thiolate group is converted to sulfenic acid (SOH), sulfenylamide (-SN) or forms of intramolecular disulfide bonds. All of these modifications prevent the sulfur atom on the cysteine from functioning as a nucleophile in the dephosphorylation reaction. Irreversible inhibition of PTPs by further oxidation to sulfinic (SO_2H) or sulfonic (SO_3H) acid can also occur within the cell (*25*, *26*). It should also be noted that serine, threonine, and tyrosine phosphorylation have also been described as mechanisms for controlling the activity of PTPs (*27*, *28*).

Like TKs, PTPs are thought to display only moderate to low specificity. Studies with substrate trapping mutants, inactivated PTP domains isolated from their parent proteins and used as probes for their interaction partners, support this assertion (*29*, *30*). However, more recent data mass spectrometry (MS) approaches using PTP knockdowns suggest that some PTPs can function as highly specific positive and negative kinase mediated pathways (*31*). Nevertheless, data for PTP specificity is limited, and specific substrate motifs have been difficult to pin down.

Src-Homology-2 and Phosphotyrosine Binding Domains

Originally identified as sequences homologous to a region of the v-SRC oncogene, SH2 domains are small modular phosphotyrosine binding domains, approximately 100 amino acids in length (*32*). In total 120 SH2 domains have been identified in 110 proteins, many of which are involved in oncogenic signaling (*33*, *34*). The binding affinities of SH2 domains are defined by specific amino acid motifs sequences up and downstream from the phosphorylated tyrosine residue, specifically, amino acids at positions -1 through +4 (-1:pY: +1:+2:+3:+4) (*35*, *36*). Binding motifs for specific SH2 domains can loosely be divided into four categories, pYE/D, pYXN, pXXM and pYXXP/L (X= any amino acid) (*36*) (Table i 1). However, there is significant overlap between these groups and the amino acid sequence surrounding these motifs can greatly influence SH2 affinity (*37*). It is also worth noting that many SH2s appear to display more complex binding affinities which cannot be defined by a specific motif. Also, a significant portion of these domains (~30) have not been studied, as they are either poorly soluble or are non-functional (unable to recognize phosphotyrosine controls) as recombinant proteins (*36*, *38*).

SH2 Domains	Motif
GRB2, GRB7, GRAP, GADS	pY-X-N
CRK, CRK-L, NCK1, NCK2, SHP1, SYK, SH2B, RASA1	pY-X-X-P/L/V
SLP76, MIST, FER, FES, ITK, GRB14, BRK	pY-D/E
PI3K1, PI3K3	pY-M-X-M
STAT, JAK, SOCS, RIN Families	Poorly Defined*

Table i 1. SH2 Binding Specificities. Table of SH2 domain binding specificities, with SH2 domain containing proteins on the left and binding motif on the right. The table grossly oversimplifies SH2 binding specificities, but is a useful starting place from which to start examining SH2-pY binding. The exact binding motifs of each SH2 domain are significantly more nuanced, with specific combinations of amino acids effecting binding in a both a positive and negative manner (*37*). As a result, defining an exact consensus motif for any one SH2 domain is extremely challenging.

SH2 domains are present in proteins with a wide variety of functions including: almost all cytoplasmic tyrosine kinases (e.g. SRC, ABL-Abelson Tyrosine Kinase, JAK-Janus Kinase and BTK-Bruton's Tyrosine Kinase families, among others), the PI3K (Phosphoinositoyl-3-kinase) family of lipid kinases, the SHIP family of lipid phosphatases, the SHP family of tyrosine phosphatases, the CBL family of ubiquitin ligases and a large number of adaptor proteins which connect tyrosine phosphorylation events with cellular signaling pathways, particularly through small GTPases (e.g. GRB2, GRB7 and GRAB with RAS, RIN family proteins with RAB7 and CRK family proteins with RAS/RAC1/RAP1) (*34*, *39*). For each of these protein types, SH2 domains function to localize enzyme activity to the site of their specific pY binding sites, often located within pY-signaling hubs such at RTKs and focal adhesion proteins. Many SH2 domain containing proteins also contain other modular binding domains such as phospholipid binding PH (Plekstrin Homology) and poly-proline binding SH3 (SRC-Homology 3) domains that further promote cellular localization, substrate recruitment and protein complex formation. Most SH2-domain containing proteins function within the cytoplasm where the bulk of tyrosine phosphorylated proteins reside. However a small number of SH2 domain containing proteins have been shown to function in the nucleus. For instance, the STAT family of proteins are translocated into the nucleus following phosphorylation and dimerization where they function as transcription factors (*40*). Likewise the SRC family kinase LYN has been shown to localize to the nucleus where it functions as in the DNA damage response and as a regulator of the cell cycle (*41*).

Quantifying Phosphotyrosine Dynamics: *Isolating TK and PTP activity in vivo*

Tyrosine phosphorylation is a highly dynamic process, with a continuous flux of phosphotyrosines passing through any pY signaling pathway (*18*). This flux, defined by

phosphorylation and desphosphorylation rates, when these two forces are at equilibrium these rates are the same and the absolute level of phosphorylation is constant (*18*, *42*). Quantification of these processes requires isolation of each process. This can be done *in vitro* using purified proteins and substrates, or *in vivo* using targeted inhibitors. A number of *in vitro* assays have been performed to quantify reaction and binding constants for kinases, phosphatases and SH2s. *In vitro* kinase and phosphatase activity assays typically require immunoprecipitated or recombinant proteins, and as a result do not take into account cellular factors which influence activity (*42–45*).

TK Activity

In vivo kinase activity likely depends on a number of factors including, kinase expression, spatial organization, activation loop phosphorylation, substrate concentration and the presence of kinase allosteric activators or inhibitors (*46–49*). Defining each of these factors can be difficult. PTP inhibition by active site oxidation has been exploited in multiple studies to measure TK activity *in vivo*. This method typically utilizes the irreversible phosphatase inhibitor pervanadate (PV), a peroxide of sodium orthovanadate, a phosphate-like competitive reversible inhibitor of PTP activity that is thought to be more specific for PTPs than peroxide treatment alone. PV is used to isolate kinase activity for *in vitro* and *in vivo* assays of tyrosine kinase activity (*44*, *50*). When live cells are treated with this reagent, cellular phosphorylation continues to rise at a rate defined by solely by the kinase activity, which can be quantified using phosphospecific antibodies. This method has been exploited in the work presented in this dissertation, and is further examined in Chapters 3 and 4.

PTP Activity

Assessment of *in vivo* phosphatase activity is equally complicated. Oxidation, expression and localization all affect PTP activity (*27*, *51*, *52*). Our limited understanding of PTP specificity further complicates the application of *in vitro* PTP activity measurements to in vivo system. To

overcome these issues, *in vivo* protein specific dephosphorylation rates can be measured by exploiting tyrosine kinase inhibitors (TKIs). By blocking the action of kinases, TKIs allow for quantification of the protein dephosphorylation without directly knowing which PTPs are mediating the effect (*42*). This method is useful when studying substrates for which the kinase responsible is well defined. Due to the important role they play in tumor growth, a number of cell permeable tyrosine kinase inhibitors have been developed over the past 20 years (*13*). Global tyrosine kinase activity can also be suppressed using the phytoestrogen Genestein. However, specific TKIs are preferred as this molecule has multiple non-TK effects on cellular physiology (*53*, *54*).

Epidermal Growth Factor Receptor Activation and Signaling

In 1986 Stanley Cohen was awarded the Nobel Prize for the discovery of EGFR and its ligand EGF (Epidermal Growth Factor) (*55*, *56*). Since the time of Cohen's seminal work EGFR has become one of, if not the most studied TK. The receptor belongs to the ErbB (erythroblastoma) family of RTKs, which includes ErbB1/EGFR, ErbB2/Her2, ErbB3 and Erb4. The intracellular domains all family members share sequence homology with the viral ErbB oncoprotein, a mutant form of the chicken EGFR homolog found in avian erythroblastosis virus (*57*). EGFR, Her2 and to a much lesser extent ErbB3,4 are known to function as oncogenes, with EGFR playing a role in nearly 30% of all andenocarcinomas. Cellular signaling proteins activated downstream of these receptors, including KRAS, BRAF, PI3K, and STAT family proteins, are also key drivers in the growth of many tumor types (*58*).

Upon binding of EGF, EGFR monomers undergo a conformational change that allows for homodimerization or heterodimerization with other members of the ErbB family (*48*). Receptor dimerization brings the cytoplasmic domains to interact (*59*, *60*). However, unlike most TKs,

EGFR kinase activity is not dependent on phosphorylation of a kinase domain activation loop (*61*). Instead the kinase domains of each EGFR are thought form an asymmetric dimer in which the C-terminal lobe of one receptor binds to and the activates the kinase domain of its partner in a manner resembling cyclin mediated activation of Cyclin Dependent Kinases (*59*).

Following activation, dimerized receptors autophosphorylate one another. The cytoplasmic tail of EGFR contains at least nine unique tyrosine phosphorylation motifs which have been shown to serve as docking sites for SH2 and PTB (phosphotyrosine binding) domain containing proteins(*58*, *62*). While a number of SH2 containing proteins have been implicated as phosphotyrosine specific EGFR binding proteins, five, including GRB2, PLCγ1, SHC1, SHP2 and CBL, have been thoroughly vetted in the literature (*43*, *58*) (Figure i 2). Each of these proteins promotes the activation of multiple downstream pathways including the RAS/RAF/MEK/ERK, PI3K/AKT, DAG/PKC and clathrin mediated receptor endocytosis and recycling. EGFR activation in is also known to induce disassembly of focal adhesions, leading to dephosphorylation of their constituent proteins, including the p130-CRK Associated Substrate (p130CAS), and a reduction in signaling through CRK family SH2 domain containing proteins (*63*–*66*).

EGFR hyperactivation as a result of kinase domain mutation and overexpression is a powerful mechanism for cellular transformation (*66*). EGFR mutations are a major driver of 30-60% of all non-small cell lung cancers, the leading cause of cancer death in the US (*66*–*68*). EGFR signaling has also been shown to be important in ERBB2 overexpressing breast cancers and in a large percentage of colorectal cancers (*69*, *70*). In these tumors there is constitutive activation of the mitogenic EGFR-dependent pathways listed above. As a result, a number of anti-EGFR agents have been developed for clinical use. Cetuximab, a chimeric (mouse/human) monoclonal antibody has been developed for use in colon cancers. This antibody prevents

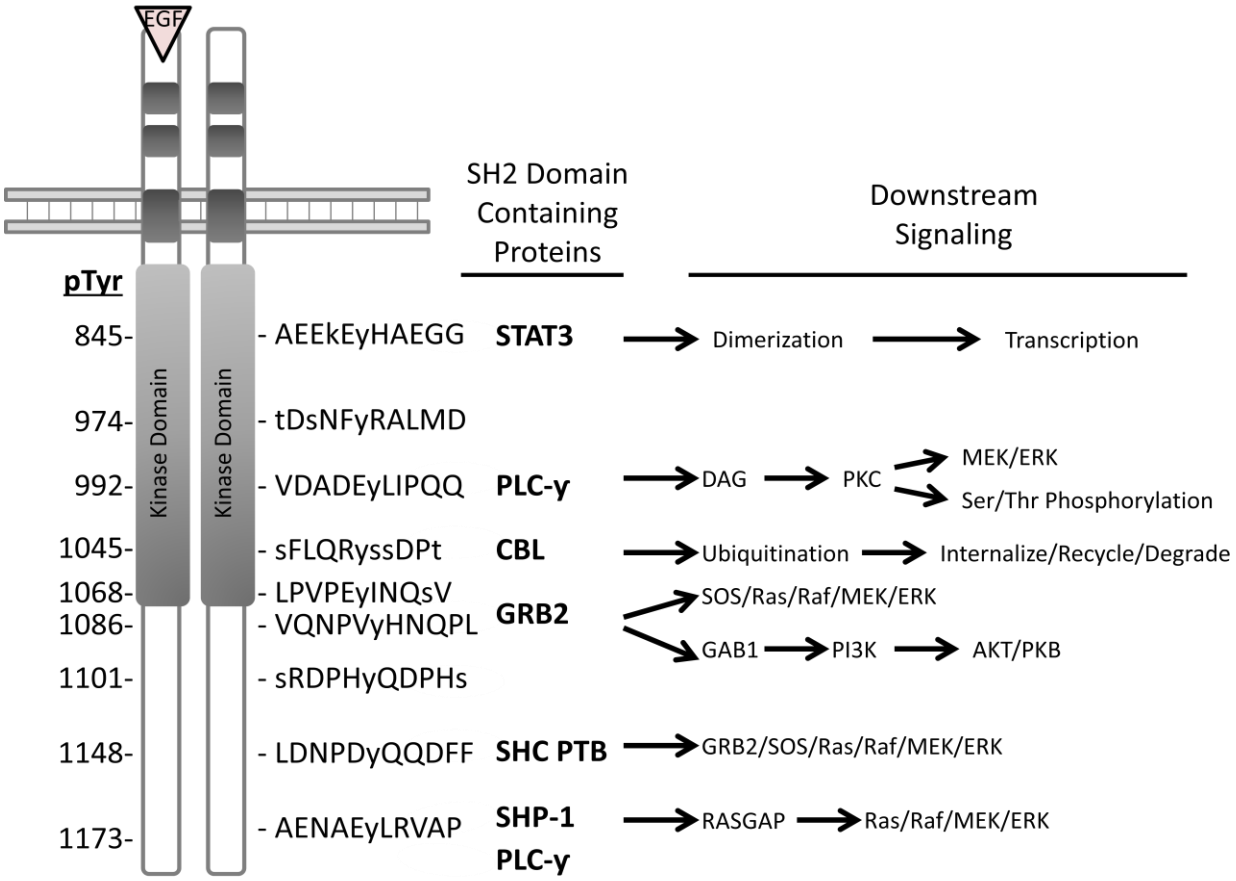

Figure i.2. Epidermal Growth Factor Receptor Phosphotyrosine Sites and Downstream Signaling. EGFR contains nine well studied tyrosines which are known to be phosphorylated and function as SH2-binding sites. The phosphosites are listed on the left and are placed in their approximate position along a linear model of the receptor. Specific motifs are listed to the right of enzyme, with phosphorylated tyrosines in lower case. The most well established SH2 domain containing proteins for each site are listed. These SH2 proteins represent a literature consensus and are not the only possible protein partners. Pathways activated by each protein are listed on the far right.

receptor activation and leads to receptor down regulation, but is only effective in cancers in which the EGFR downstream protein, KRAS has not been mutationally activated. ATP-analogues are another class of well studied EGFR inhibitors used both clinically and experimentally (see above). These include Erlotinib and Gefitinib which are used mainly in the treatment of NSCLCs with EGFR activating mutations. While both these classes of inhibitors prolong survival, resistance develops almost universally, either via activation of proteins downstream from EGFR or as a result of so-called "Gate Keeper" mutations in the receptor which allow for ATP hydrolysis and tyrosine phosphorylation but prevent the inhibitor from accessing the kinase active site (*71*, *72*).

KRAS Activation and Signaling

Nex to EGFR, the small GTPase RAS is one of the most clinically important pro-oncogenes gene families. This family includes K-, R- and N-RAS. RAS activation occurs following recruitment of SH2 domain containing proteins such as SHC1 and GRB2 to the phosphorylated proteins including EGFR. Together these proteins recruit the RAS activator SOS (Son of Sevenless) to the site of phosphorylation. SOS functions as a guanine nucleotide exchange which promotes the release of GDP from the RAS allowing it to bind GTP. GTP-binding activates RAS allowing it to bind and activate proteins futher downstream (*73*). In RAS driven tumors, RAS is typically activated as a result of missense mutations in either codon 12 or 13 (the common mutant KRASV12 being used in our proposed work) which prevents conversion to the inactive form of RAS. Activated RAS is known to binding to and activate a number proteins including RAF. RAF in turn actives the pro-proliferative mitogen-activated protein kinase pathway (MAPK pathway) and the PI3K pathway which promotes anti-apoptotic AKT pathway signaling (*14*, *74*).

Methodologies Overview:

Quantitative Mass Spectrometry

A complete understanding and analysis of phosphotyrosine dynamics and phosphoprotein-SH2 interactions requires at least some knowledge of the phosphosite sequences within a specific biological sample (tumor lysate, cell type, treatment condition). Currently phosphospecific mass spectrometry studies have identified over 40,000 non-redundant phosphotyrosine sites (*62*). However, to truly incorporate this data into dynamic models of cellular signaling we must have some measure of the abundance of specific phosphosites across multiple time points and/or experimental conditions.

Over the past few decades two major isotope-based tagging methods have been developed to tackle this issue; SILAC (Stable Isotope Labeling of Amino Acids in Cell Culture) and iTRAQ (Isobaric Tagging for Relative and Absolute Quantification). SILAC relies on the incorporation of a particular amino acid containing some variation of heavy or light isotopes (e.g. C12 and C13) into cell cultures assigned to different experimental conditions. Cells are combined, lysed and protease treated before MS analysis. In contrast, for iTRAQ (the method used in the studies presented in here) cells are lysed and protease treated first. Then phosphopeptides from each experimental treatment are isolated, covalently tagged with equivalent molecular weight tags, combined, separated by mass to charge ratio by an initial round of mass spectrometry and fragmented by ionization. Fragmentation causes the cleavage of the molecular weight tags in such a way that each experimental tag has a unique molecular weight. The peptide fragments are then analyzed again by mass spectrometry allowing for peptide sequence identification. The relative abundance of each individual phosphopeptide is determined by the relative amount of each tagged identified (Figure i.3). While both methods have their merits the major advantage of iTRAQ is that it can be applied to both cell cultures and isolated tissues.

In recent years, quantitative phospho-mass spectrometry has provided significant information about phosphotyrosine signaling networks, including that of EGFR. Temporal analyses of phosphotyrosine changes, as well as those in phosphoserine and phosphothreonine, have been used to reconstruct phosphorylation dependent signaling pathways and link temporal changes in phosphorylation at specific sites with changes in signaling output over time (*75*). MS has also been combined protein-protein interaction data obtained from SH2-domain based binding assays, SH2-pY interaction tools and protein interaction databases, such as those listed below, to reconstruct complex TK driven cellular signaling networks (*76*, *77*).

The Domainomics Approach: *Domainomics*, *Global profiling of in vivo SH2-pY interactions and Single Molecule Imaging of SH2-Phosphotyrosine Interactions.*

Domainomics

Many proteins contain independently folding globular domains capable of binding short peptide motifs even when both domain and motif are removed from the context of their full-length protein (*78*, *79*). Modular protein interacting domains facilitate protein-protein interactions required for a diverse set of cellular processes including signal transduction and subcellular localization. Domains are categorized based on structural and sequence homology, with each domain family recognizing motifs with similar characteristics, such as phosphorylated tyrosine (pTyr) or proline rich sequences. The combination of modular domains within a protein contributes to its biological function by defining its protein interaction network. Post-translational modification (PTM) of amino acid side chains within specific peptide motifs can modulate domain-motif binding. And it is the variation of in domain specificity and motif modification that is

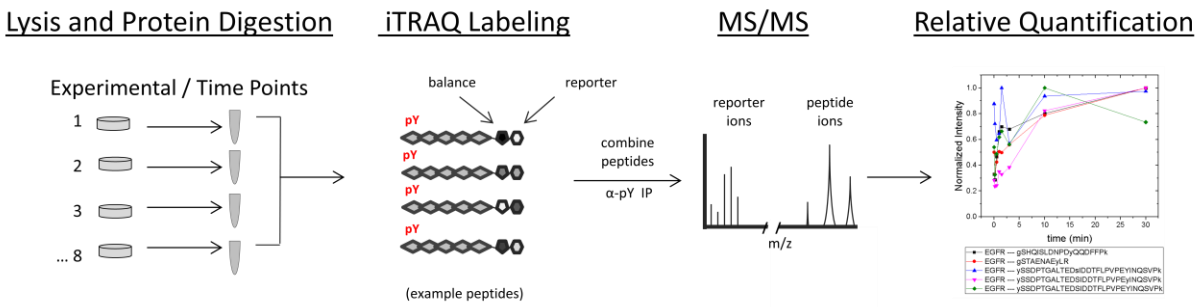

Figure i.3. Schematic of iTRAQ mass spectrometry. Isobaric tagging for relative and absolute quantification (iTRAQ) allows for comparison of the relative abundance of specific phospho-sites across multiple experimental samples. Cells from each experimental point are lysed, proteins are cleaved using trypsin and covalently labeled using an isobaric tags containing a reporter and balance. Balance and reporter used for different time points contain varying amounts of heavy and light isotopes, with each complete tag having the same molecular mass (therefore isobaric), but each reporter having a different mass. The peptides are then combined, and enriched for phosphorylation sites using TiO_2 and anti phosphotyrosine antibody columns. Peptides are then run through tandem MS, often using a triple quadrapole setup. Peptides are separated by mass and ionized which includes balance-reporter cleavage. Protein and reporter ions are separated again by MS. The molecular weight and specific pattern of peptide ions is used to determine peptide sequence. The relative abundance of each of the much lighter (lower mass to charge, m/z, ratio) reporter ions equates the relative abundance of that peptide in each experimental sample.

the basis for much of the elegant and complicated protein signaling networks required for life (*80*). Over the past decade, exploitation of a number of high-throughput proteomic technologies has led to the dissection of vast protein interaction networks and helped to decode the role of PTMs in altering network topology(*81*).

A significant portion of proteomics and signaling systems biology studies have taken advantage of modular domains as a means to assess protein-protein interactions. We refer to this modular domain-based proteomics as "domainomics" (*32*). While this may be a somewhat artificial segmentation, it is meant to draw attention to the potential of domains and their motifs as tools in contemporary proteomics. The work presented here focuses on SH2-domain based domainomics approaches.

Independent folding of SH2 domains, which preserves binding capabilities, allows for their use as affinity and detection reagents in a manner similar to that of phospho-specific antibodies. Unlike antibodies, SH2 domains naturally have wide-ranging specificity (see above). However, promiscuity in ligand selection is a physiological propensity rather than experimental noise, as it would be for antibodies. Interpretation of SH2 binding data is therefore more complex, requiring information about SH2-domain containing protein expression levels, phosphosite availability, and SH2-pY affinities and specificities.

Multiple technologies have been developed to assess SH2-pY interaction using the domainomics approach including: support dependent forward phase and reverse phase assays, in-solution assays and *in vivo* binding assays (figure i.4). In forward phase arrays, domains are immobilized on a solid support and probed with a tagged phosphopeptides or protein. Pioneered by Espejo et al 2002 and improved for uses with SH2 domains by the Macbeath group, this methodology has been used to define SH2-pY motif specificity and affinity (*43*, *82*). However, a major concern for the forward phase format is whether immobilization of

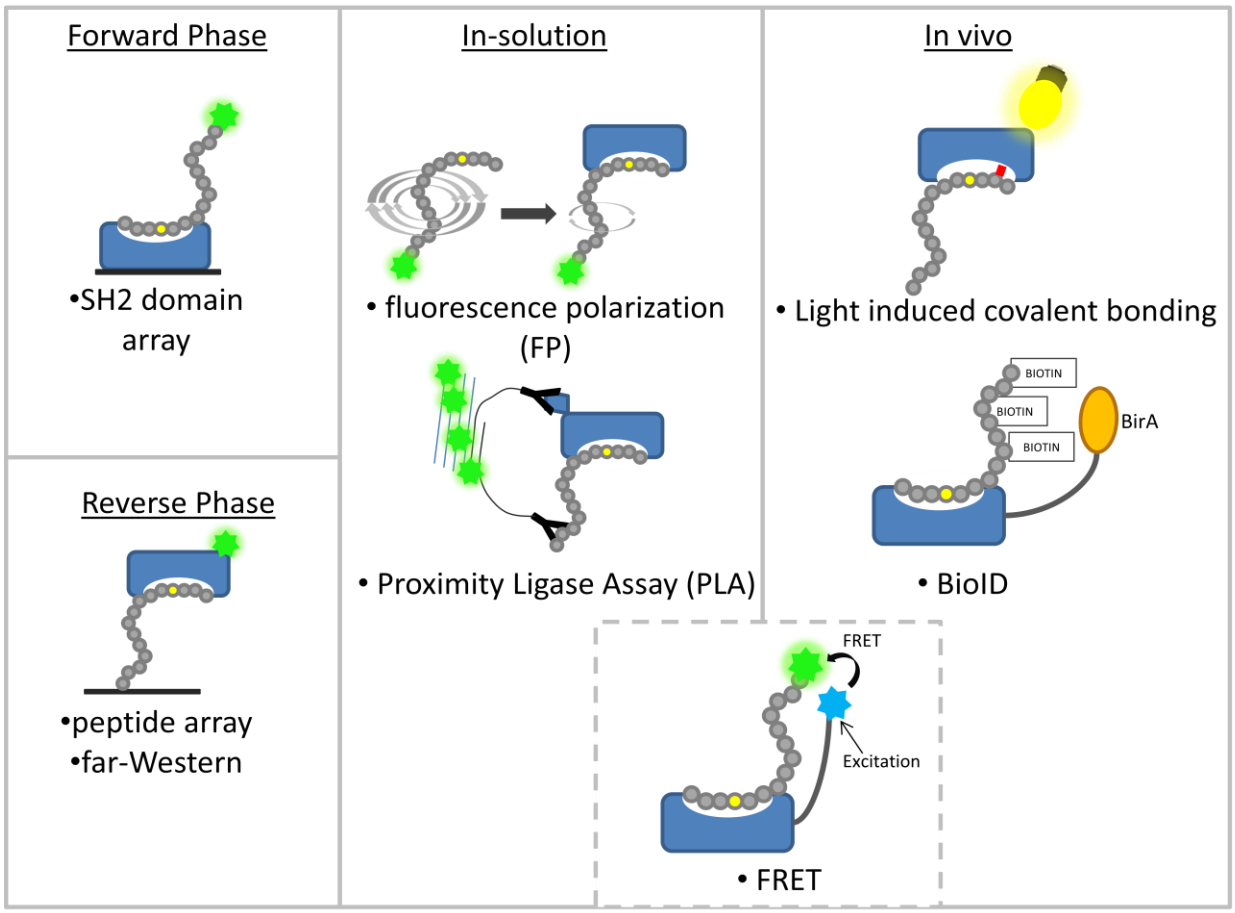

Figure i.4. Domainomic approaches to quantifying SH2-pY interaction. A number of domain-based approaches have been developed to identify and quantify SH2-pY interactions. Forward phase arrays utilize arrays of SH2 domains and can be useful for identifying the binding partners of a specific, usually small set of proteins. Reverse phase assays are useful for screening large peptide libraries for their ability to bind specific SH2 domains. In-solution assays remove the effect of non-specific substrate (e.g. nitrocellulose membrane) binding. These methods include: fluorescence polarization, proximity ligation and Förster Resonance Energy Transfer (FRET) which can also be used *in vivo*. A number of other *in vivo* SH2 domain based interaction assays have also been developed or are in development including light inducible pY-SH2 covalent bonding and BioID.

recombinant protein domains might disrupt the conformation and orientation required for ligand binding (*83*, *84*). The method also hampers high throughput screening as a unique SH2 domain array is required for every phosphopeptide tested.

In reverse phase arrays, cell lysates containing phosphorylated lysates or phosphorylated synthetic peptides are immobilized and probed with SH2 domains. This method is particularly suitable to screen large-scale libraries of peptide motifs to determine consensus sequences for domains, as thousands of phosphorylated peptides or even whole cell lysates can be probed with a single SH2 domain at once(*38*). Reverse phase assays have also have been extensively utilized in the Mayer Lab in the form of far-Western/SH2 profiling (see below) and by a recent large scale study of SH2 specificity that used recombinant SH2 domains to probe microarray of thousands of tyrosine phosphorylated peptides (*36*, *67*).

For in-solution methods, both SH2 domains and phosphosite are solubilized, which minimizes membrane associated binding artifacts seen in reverse and forward phase assays. In-solution assays include; FRET (Fluorescence Resonance Energy Transfer), fluorescence polarization (FP) and proximity ligation-based assays (PLA). FRET dependent methods measure the transfer of energy from excited electrons between nearby fluorophores on SH2-containing proteins and their phosphorylated targets. Unlike most other in-solution methods, they typically directly measure interactions and therefore can used for recombinant proteins, cell lysates and *in vivo* (*85*–*88*). However, these assays can have low sensitivity due to low signal to noise. FP-based assays assess the rotation of fluorescently tagged peptides bound to recombinant SH2 domains and are particularly useful for defining binding constant (Kd) values, though they too have issues with signal to noise ratios (*89*). Proximity ligation assays attempt to overcome low signal to noise issues using amplification. These assays use SH2 domain and protein specific probes tagged with oligonucleotides which are ligated when in close proximity and serve as

templates for PCR or rolling circle based amplification which is quantified using real-time PCR (*50*).

It is worth noting that all of the aforementioned assays require that proteins be removed from their native intracellular localization. Over the past few years *in vivo* binding assays have been developed, capable of detecting SH2-pY interactions occurring in live cells. These methods are typically significantly lower throughput than those mentioned above and include photo-crosslinking and BioID, as well as the phosphatase protection assay and super resolution imaging based methods that are discussed in Chapters 2 and 3, respectively. Photo-crosslinking methods utilize exogenous t-RNAs to incorporate synthetic amino acids into SH2 domains. The amino acids form covalent crosslinks with nearby proteins upon exposure to light of specific wavelength, allowing the experimenter to essentially freeze *in vivo* SH2-pY interactions in time, which can then be identified by MS after cell lysis (*90*). BioID utilizes proteins tagged with the bacterial biotin ligase (BirA) which biotinylate proteins within close proximity, which are then analyzed by Western or MS after cell lysis (*91*). This method is currently under investigation for analyzing SH2-pY interactions. Phosphatase protection assays utilize SH2 mediated protection to identify pY-SH2 interactions *in vivo* and will be discussed further in Chapter 3. Super resolution imaging can be used to monitor and quantify the movement and localization of individual ectopically expressed fluorescently tagged SH2 domain containing proteins in live cells. One such method, single particle tracking photo-activated light microscopy (sptPALM), is described in detail below and utilized in Chapter 2.

Profiling the global phosphoproteome using SH2 domains

To develop an unsupervised method for identifying SH2 domain binding sites and quantifying their dynamics, our laboratory has developed SH2 profiling (*38*). SH2 profiling utilizes far-Western blotting, in which non-antibody proteins are used as affinity reagents for proteins or

post-translational modifications of interest (*92*). In SH2 profiling recombinant SH2 domains are used as probes for phosphotyrosine sites within cells lysates which are either spotted onto membranes or run on PAGE (polyacrylamide gel electrophoresis) gels and transferred on nitrocellulose membranes. Binding of each SH2 domain is then detected using a non-specific high affinity probe conjugated with a detection agent (e.g. SH2-GST and anti-GST-HRP) and quantified using chemiluminesnce or fluorescence. In the case of lysates run on PAGE, signal quantification can be further divided by protein band or molecular weight. Clustering methods can then be applied to the data in order to detect significant subsets of lysates, SH2 domains, or phosphoproteins (Figure 1.5). Multiple experiments performed in the Mayer lab, as well as others, have shown that this method is capable of recapitulating much of the SH2-phosphotyrosine binding observed by more time consuming, lower throughput methods such as mutational analysis and immunoprecipiation (*12*). By quantifying and comparing the binding patterns and kinetics from multiple SH2 domain probes one can reconstruct the SH2 domain-phosphosite interaction landscape occurring within in the cell. SH2 profiling been shown to be useful tool for clustering cells and tumor tissues by functional characteristics such drug sensitivity and is currently under investigation for use as a prognostic marker (*12*, *50*, *67*). The method is also well suited for analysis of SH2 binding site kinetics, as it allows for the quantification and comparison SH2 binding across multiple phosphoproteins and time points (See Chapter 2).

In vivo single SH2 domain imaging

One drawback of SH2 profiling based analyses is that like other post-lysis interaction assays, the spatial orientation of phosphoproteins is lost prior to interaction profiling. *In situ* methods for identifying SH2–phosphoprotein interactions such as antibody-based immunofluoresence using fixed cells can provide some insight, but can be difficult to control across multiple time points and can be associated with a number of method dependent artifacts.

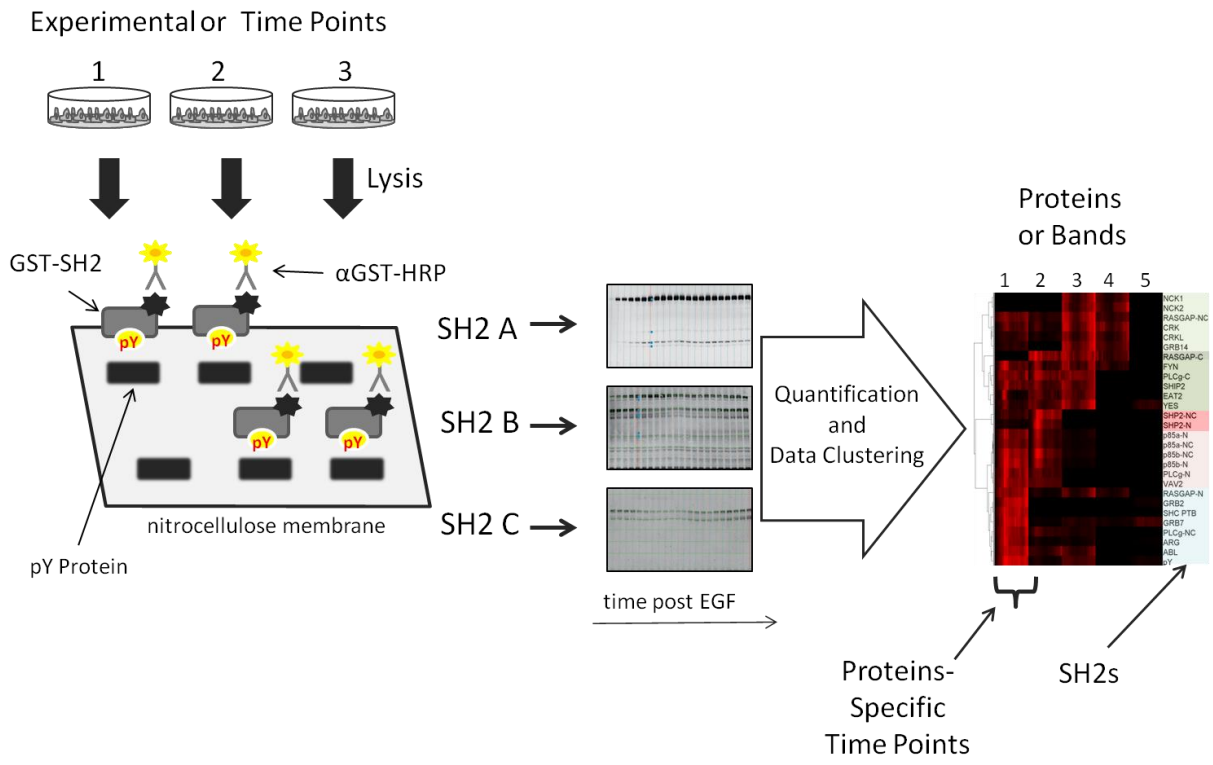

Figure i.5. SH2 Profiling. SH2 profiling is a reverse phase assay which uses SH2 domains as affinity reagents to assess the global phosphorylation state of lysates from different cell lines or experimental conditions. Cells from different experimental points are lysed, separated by gel electrophoresis, transferred to nitrocellulose membranes and probed with the large set of recombinant SH2 or PTB domain probes. Probe binding is detected using HRP- or fluorophore-tagged secondary antibodies by digital scanners, quantified and compared using hierarchical clustering.

To develop a method to monitor SH2 binding across both space and time in live cells we have collaborated with the Ji Yu laboratory, and specifically Dr. Dongmyung Oh, to develop a method for tracking single SH2 molecules as they are recruited to the surface of cells treated following RTK activation. The method incorporates and builds on a sptPALM method developed and used by Manley et al. to monitor tracking viral Gag and Vesicular Stomatitis Virus – G proteins on the plasma membrane of living cells (*93*) (Figure i.6). The protocol uses exogenously expressed photo-convertible fluorescently tagged SH2 domains to monitor EGF stimulated SH2 domain recruitment to the basal membrane of adherent cells on a glass slide. Total internal reflection (TIR) optics are used to excite only molecules at or near the plasma membrane (Z-axis resolution = 200nm) (*94*). To follow individual molecules, UV light is used to convert the excitation and emission spectrums (ex:506nm/em:561nm to ex:571nm/em:581nm) of a fraction of exogenously expressed SH2 domains tagged with the phophoactivatable fluorophore tdEOS, a fluorescent protein originally derived from the stony coral *Lobophyllia hemprichii* (*95*, *96*). Individual molecules can then be localized by fitting their point-spread functions to a Gaussian distribution, allowing for lateral resolution of approximately 30nm. The individual SH2 domains are then tracked by acquiring TIR images using short exposures and a rapid frame rate over a set period of time. This method provides a powerful tool for monitoring changes in the spatial orientation and mobility of individual SH2 molecules on the plasma membrane (*93*, *97*). Using this method we have developed an SH2 membrane hopping model in which SH2 domain interaction with the plasma membrane is sustained by multiple rebindings prior to returning to the cytosol (*94*) (Figure i.7). One major issue with the method is that it does not allow the user to easily identify the specific membrane phosphoprotein or set of proteins that each SH2 domain is binding. As a result, single molecule and TIR experiments must with combined with more

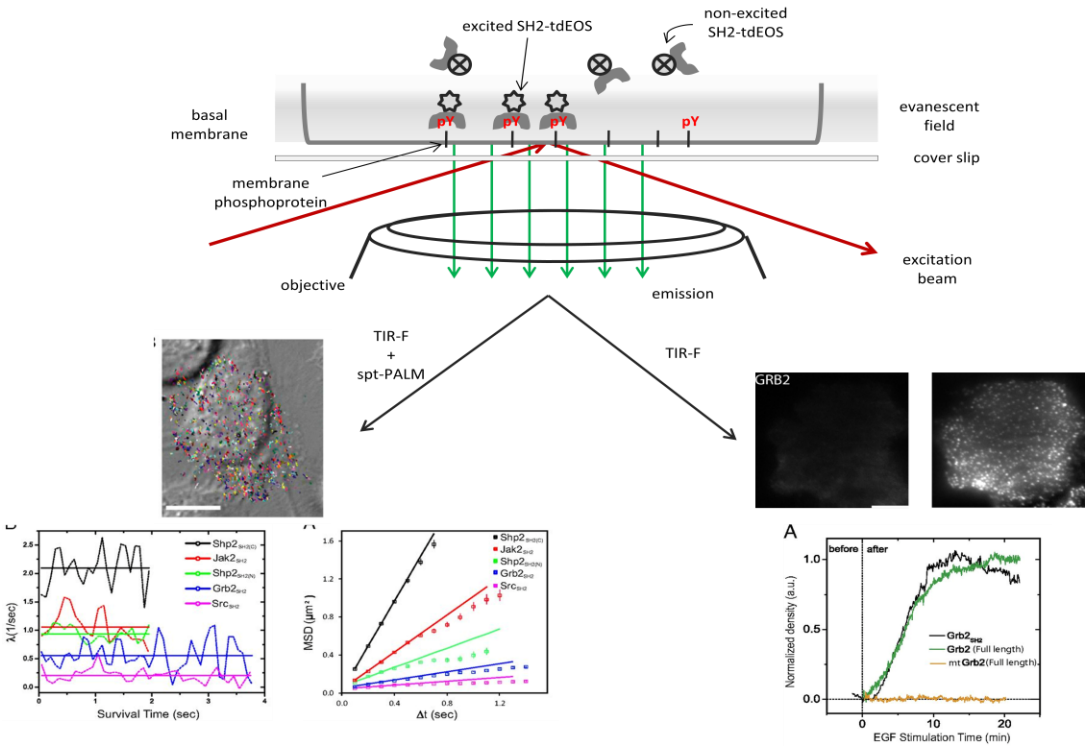

Figure i.6. Total Internal Fluorescence Microscopy and Single Particle Tracking Photoactivated Light Microscopy. An excitation beam focused at a cell membrane-glass slide interface at an angle sufficient to produce total reflection will produce an evanescent field along the interface with the same wavelength. This field extends into the cell with a depth of approximately 1/3 the wavelength of the excitation beam. This results in an excitation field extending about 200-300 nm into the cell and allows users to visualize only those fluorescently tagged molecules within that region in real time. This can be used to measure the recruitment of SH2 domains to phosphorylated membrane proteins in response to growth factor stimulation. The resolution of specific molecules on the surface can be greatly improved by using a photactivatable fluorophore, such as tdEOS. When exposed to higher energy blue or UV light, a small fraction of tdEOS molecules are “activated”, causing them to fluoresce at different excitation/emission spectrums. This allows users to track a small number of molecules as they are recruited to, and diffuse across or away from the excitation field. This method can be used to determine membrane on-rates, off-rates, diffusion constants and diffusion constraints.

traditional biochemical analyses, including Western, FW and MS, which can provide more information about SH2 domain binding partner phosphorylation dynamics.

Computational Modeling and EGFR Signaling

The mechanism by which dynamic or complex cellular processes function or produce a specific output are often difficult to determine or experimentally test using traditional biochemical and imaging methods. In these cases, computational reaction modeling can be a useful tool. In computational modeling interactions between biological species or components, such as structural proteins, enzymes, RNA, DNA etc., are represented by a set of interdependent reaction rates, which in turn are a function of component rate or binding constants and concentrations. By varying species parameters within this system of interconnected equations one can gain insight in those inputs that have the greatest effect on the output in question. The resulting hypothesis can then be tested by further wet lab experimentation. Data from these experiments can then be used to further refine the model.

Parameters which are not known can be approximated by determining which values allow the model to most closely recapitulate experimental results. However, in using a model containing too many approximated values one runs the risk of receiving erroneous results (*98*). The accuracy of computational models depends largely on the accuracy of the inputs and their relationship to one another. A significant effort has been made over the past twenty-five years to quantify EGFR-EGF and SH2 domain-pY EGFR interaction affinities and specificities (*21*, *35*, *43*, *89*, *99*, *100*). As a result, SH2 domain containing protein mediated EGFR signaling has been modeled by multiple groups in order to gain insight in the process. Many of these utilize complex modeling methods which are beyond the scope of the work presented here (*44*, *49*, *101*, *102*). However, only a few of these models utilize experimentally determined values for

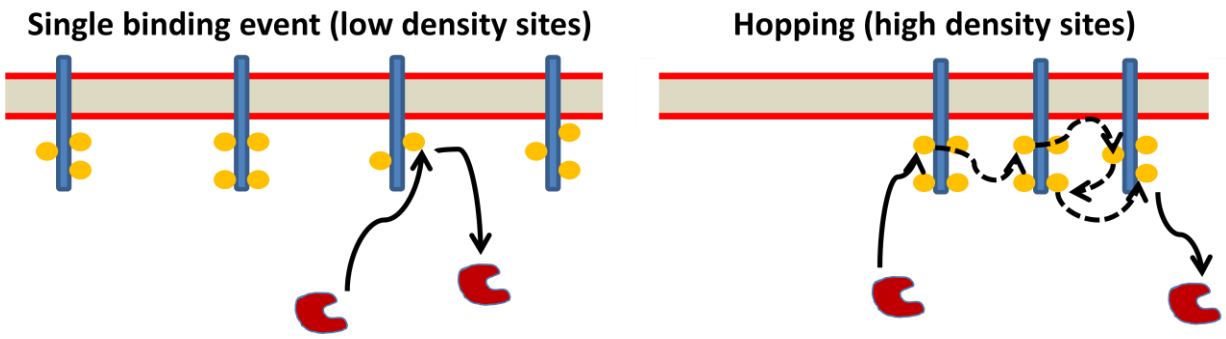

Figure i.7. Rebinding model of SH2-pY EGFR interaction. Phosphosite clustering increases the probability that an SH2 bound at the membrane will rebind a nearby membrane-associated phosphosite before exiting into the cytosol. As a result, SH2 domain containing proteins dwell at the membrane for a much greater time than would be expected from *in vitro* measured off-rates.

EGFR phosphorylation and dephosphorylation or SH2 protein concentrations. As a result, they are more susceptible to the parameter fitting hazard described above.

Traditional biologists and biochemists often do not have the necessary training or mathematical background to create a mathematical model from scratch based on their own data (*103*–*105*). To rectify this issue, the Center for Cell Analysis and Modeling at UConn developed the Virtual Cell (VCell), a modeling software with a reasonably friendly user interface whose base functions can be utilized with minimal training (*106*). In the work presented here, the VCell software is exploited to model our findings using experimentally determined parameters in an attempt to validate a proposed explanation for experimental findings.

Chapter 1

Time-resolved multimodal analysis of SH2 domain binding in EGF-stimulated cells

Joshua A. Jadwin[*1], Dongmyung Oh[*2], Timothy G. Curran[3], Mari-Ogiue Ikeda[1], Lin Jia[1], Forest M. White[3], Kazuya Machida[1], Ji Yu[2], Bruce J. Mayer[1,2]

Affiliations

1. Raymond and Beverly Sackler Laboratory of Molecular Medicine, Department of Genetics and Genome Sciences, University of Connecticut School of Medicine, 263 Farmington Avenue, Farmington, CT 06030

2. Richard D. Berlin Center for Cell Analysis and Modeling, University of Connecticut School of Medicine, 263 Farmington Avenue, Farmington, CT 06030

3. Department of Biological Engineering and Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139

Abstract

While the affinities and specificities of SH2 domain-phosphotyrosine interactions have been well characterized, the spatio-temporal changes in phosphosite availability in response to signals, and their affect on recruitment of SH2-containing proteins *in vivo*, are not well understood. To address this issue, we used three complementary experimental approaches to monitor phosphorylation and SH2 binding in A431 cells: 1) phospho-specific mass spectrometry; 2) far-Western blotting; and 3) live cell single-molecule imaging of SH2 membrane recruitment. Far-Western and MS analyses of EGF-stimulated cells identified both well-established and previously undocumented EGF-dependent tyrosine phosphorylation and binding events, as well as dynamic changes in binding patterns over time. In comparing SH2 binding site phosphorylation with SH2 domain membrane recruitment in living cells, we found *in vivo* binding to be much slower. Delayed *in vivo* SH2 domain recruitment correlated with clustering of SH2 domain binding sites on the membrane, consistent with membrane retention via SH2 rebinding.

Introduction

Receptor tyrosine kinases (RTK) and tyrosine kinase-associated receptors play an essential role in transducing extracellular signals into the cell. These proteins function as central signaling nodes for a diverse set of normal biological processes including proliferation, differentiation, immune cell activation, neuronal development, angiogenesis, and cell migration. Dysregulated tyrosine kinase activity is also a primary driver of human cancer. Upon activation, RTKs phosphorylate tyrosine residues on themselves and associated proteins, which then serve as binding sites for Src homology 2 (SH2) and phosphotyrosine binding (PTB) domains found within downstream signaling proteins (*3*, *107*, *108*).

120 SH2 domains have been identified in 110 different proteins (*33*, *36*). The central role these proteins play in cellular signaling has made them popular targets for study. The affinities and specificities of many phosphotyrosine (pY)-SH2 interactions have been quantified, and much of the complex web of downstream pathways initiated by SH2 domain binding has been unraveled (*36*, *43*, *89*, *109*, *110*). By contrast, much less is known about the role that spatial and temporal changes in protein phosphorylation play in signal transduction. Current experimental approaches, however, have the potential to address system dynamics directly.

Methods such as quantitative mass spectrometry (MS) and live cell fluorescence microscopy are now are capable of tracking temporal changes in cellular physiology across remarkably short time steps. For instance, isotopic protein-tagging MS methods, including SILAC and iTRAQ, allow us to quantify the relative and absolute abundance of pY sites found within hundreds of proteins across multiple time points (*111*, *112*). Data from these experiments has been used to map temporal changes in molecular signaling, providing us with a more comprehensive understanding of pathway dynamics (*75*, *112*). At the same time, high resolution microscopy, coupled with more traditional imaging methods and biochemical studies, has

allowed us to begin to dissect the spatial reorganization of RTKs and their effectors following receptor activation (*49*, *60*, *113–118*).

We previously used single particle tracking photoactivated localization microscopy (sptPALM) (*93*) to visualize and quantify individual SH2 domain binding events at the plasma membrane in response to receptor activation (Oh et al., 2012). From these studies we developed a model in which the high density of tyrosine-phosphorylated sites on the membrane results in the repeated rebinding of SH2 domain-containing proteins before they can escape into the cytosol. This local rebinding suppresses the apparent off-rate and prolongs the membrane dwell time of SH2 domain-containing proteins. Furthermore, we showed that clustering of SH2 binding sites, a well-known consequence of RTK activation (*60*, *115*, *119*, *120*), further suppressed the apparent off-rate of the GRB2 SH2 domain, suggesting that signal output could be influenced by changes in physical properties such as phosphosite distribution.

Traditional two-dimensional signaling diagrams and phosphopeptide-based affinity studies fail to fully capture the complexities and importance of spatial and temporal changes in receptor activation, protein phosphorylation, and SH2 domain-containing protein recruitment. Indeed, existing biochemical and proteomic methods each have strengths and technical limitations, such that no single approach provides a clear and unambiguous perspective. Thus to better understand RTK signaling dynamics, we have employed three orthogonal experimental techniques: SH2 domain-based reverse-phase binding assay (far-Western blotting) to assess global changes in SH2 binding sites, iTRAQ (isobaric tagging for relative and absolute quantification)-based phosphotyrosine-specific MS to identify changes in the abundance of specific phosphopeptides, and live cell single particle tracking using total internal reflection fluorescence (TIRF) microscopy to assay recruitment of SH2 domains to the membrane *in vivo (*Figure 1.1). Combining these complementary approaches to analyze the EGF response in the same experimental system revealed previously uncharacterized properties of the EGFR

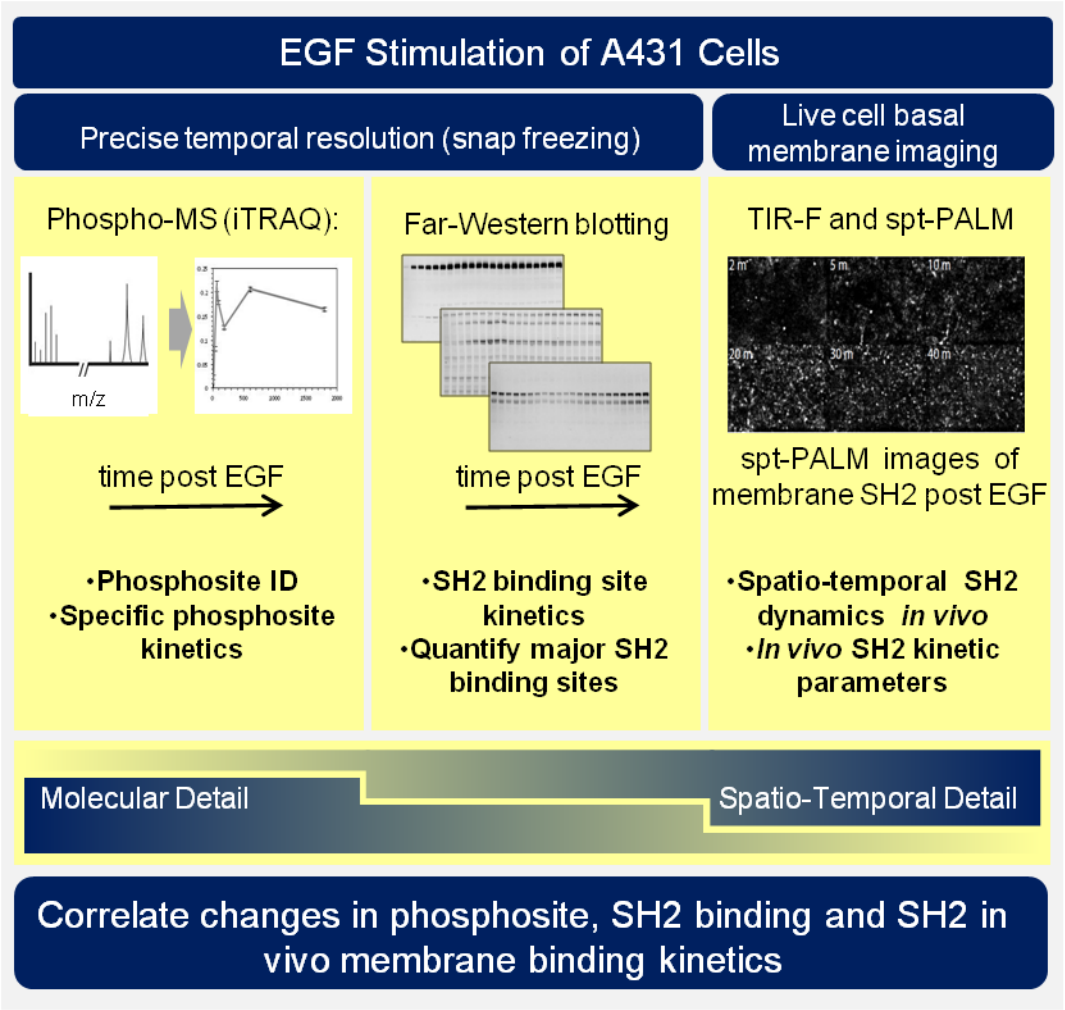

Figure 1.1. Outline of multimodal experimental approach to analysis of EGFR phosphorylation and SH2 binding in A431 cells. FW blotting provides a link between the phosphorylation level changes quantified by MS and the *in vivo* SH2 binding measurements made using TIR imaging.

signaling network. We quantified and compared binding site creation kinetics for a large set of SH2 domains with high temporal resolution, and were able to define a previously unappreciated role for phosphosite clustering in sculpting the dynamic profile of signal output downstream of tyrosine kinase activation.

Results

Dynamic EGF-dependent changes in tyrosine phosphorylation

We first performed far-Western blotting, a reverse-phase SH2 binding assay (*38*) to quantify changes in binding sites for multiple SH2 domains across the set of phosphoproteins in EGF-stimulated A431 cells, an EGFR-overexpressing squamous-cell carcinoma cell line (*121*, *122*). Cells were starved overnight, stimulated with 25 ng/ml EGF, flash frozen at multiple time points post-stimulation, and lysates run on LDS-PAGE in duplicate. All blots included positive and negative pY controls to assess non-specific binding. Blots were probed with recombinant GST-tagged SH2 and PTB probes and an anti-pY antibody (Table 1.1) (*38*). Of the probes tested, 27 were selected for further analysis based on significant, reproducible, and dynamic binding to EGF-stimulated A431 cell lysates (Figure 1.2). Anti-pY blots of the same lysates showed a rapid increase in total phosphorylation upon EGF treatment, dominated by a major band corresponding to EGFR (Figure 1.3). Blots also contained a number of minor bands whose phosphorylation varied over time. SH2 binding patterns varied across the set of SH2 domains tested, however the molecular weights of major bands were consistent, suggesting that most SH2 probes bound predominantly to a relatively small set of highly phosphorylated proteins (Figures 1.2 and 1.3). Most prominent were five highly dynamic bands, which were identified by immunodepletion (Figure 1.4) and exhibited three well-defined kinetic patterns. These major bands were (1) EGFR, a ~195kDa band whose phosphorylation increased rapidly following EGF treatment, dipped slightly and then remained relatively constant; (2 and 3) the focal adhesion protein p130CAS, a doublet at 150 and 115kDa whose phosphorylation was high in unstimulated cells, decreased rapidly upon EGF treatment, and then rebounded at later time points; (4) the scaffolding protein GAB1, a single band at 130kDa that was phosphorylated with kinetics similar to EGFR, then rapidly returned to near basal levels; and (5) the scaffold/adaptor SHCA, a relatively weak band at 71kDa which displayed kinetics similar to that of EGFR (Figures 1.2, 1.3, 1.4 and 1.5 A).

See attached Excel files

Table 1.1: SH2 domain constructs. Cloning information and amino acid sequences for all cDNA constructs used in this analysis are listed.

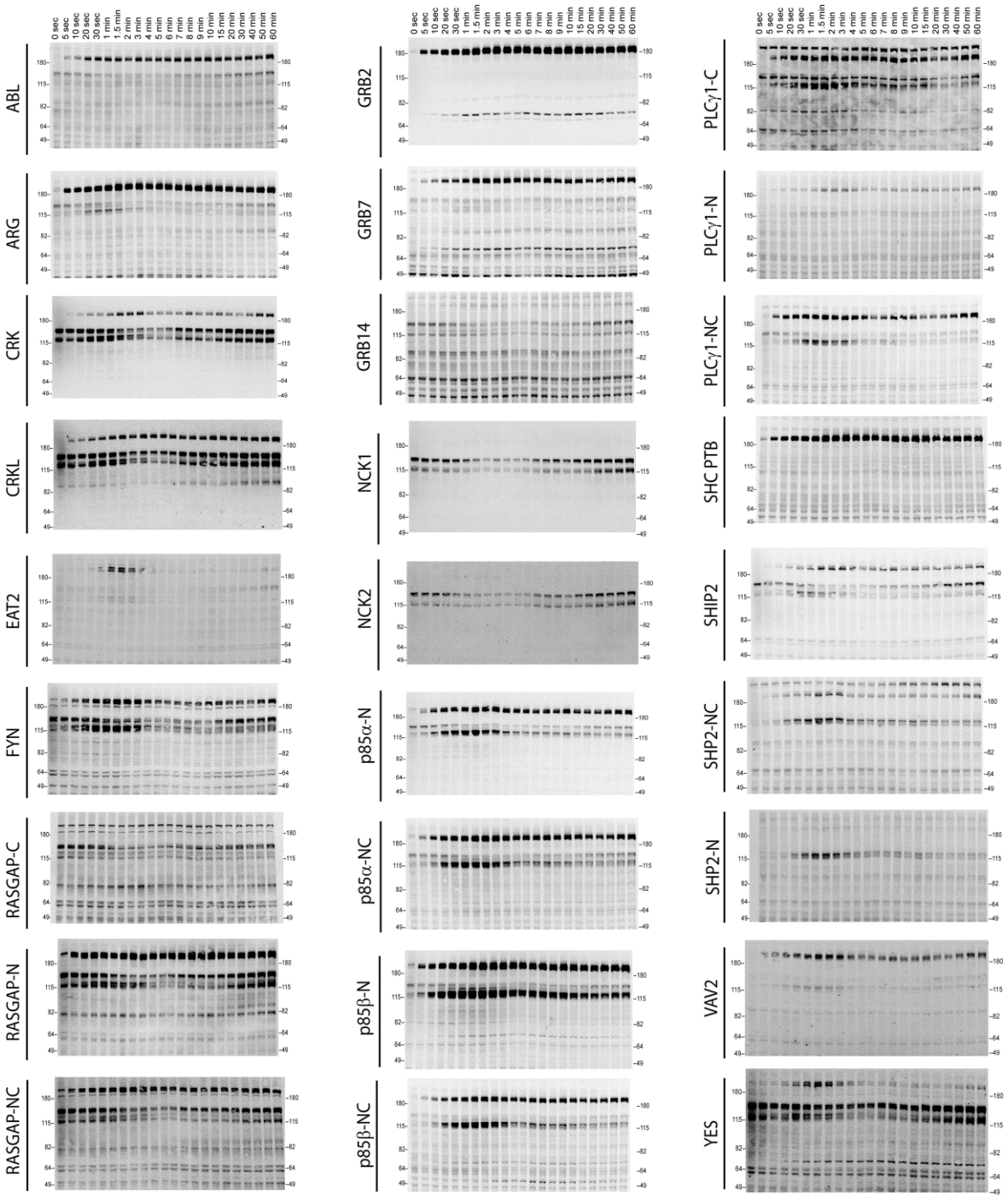

Figure 1.2: Representative FW blots using 1 PTB and 26 SH2 domain probes for EGF-stimulated A431 cells. Protein names are labeled on the left and stimulation times are on top. For probe names, "-N, -C, or -NC" indicates the use of N-terminal, C-terminal, or natural tandem SH2 domain probes. See Table 1.1 for construct information.

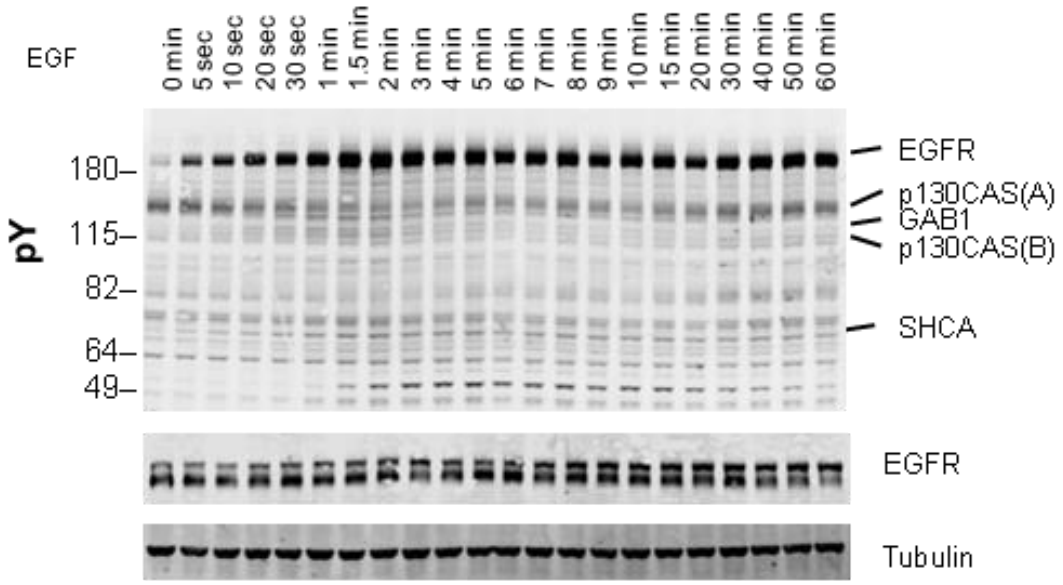

Figure 1.3. Representative anti-phosphotyrosine blot of EGF stimulation time course in A431 cells. The major SH2 binding proteins (identified by immunoprecipitation-see 1.4) are marked on the left. EGFR and tubulin loading controls shown in lower panels.

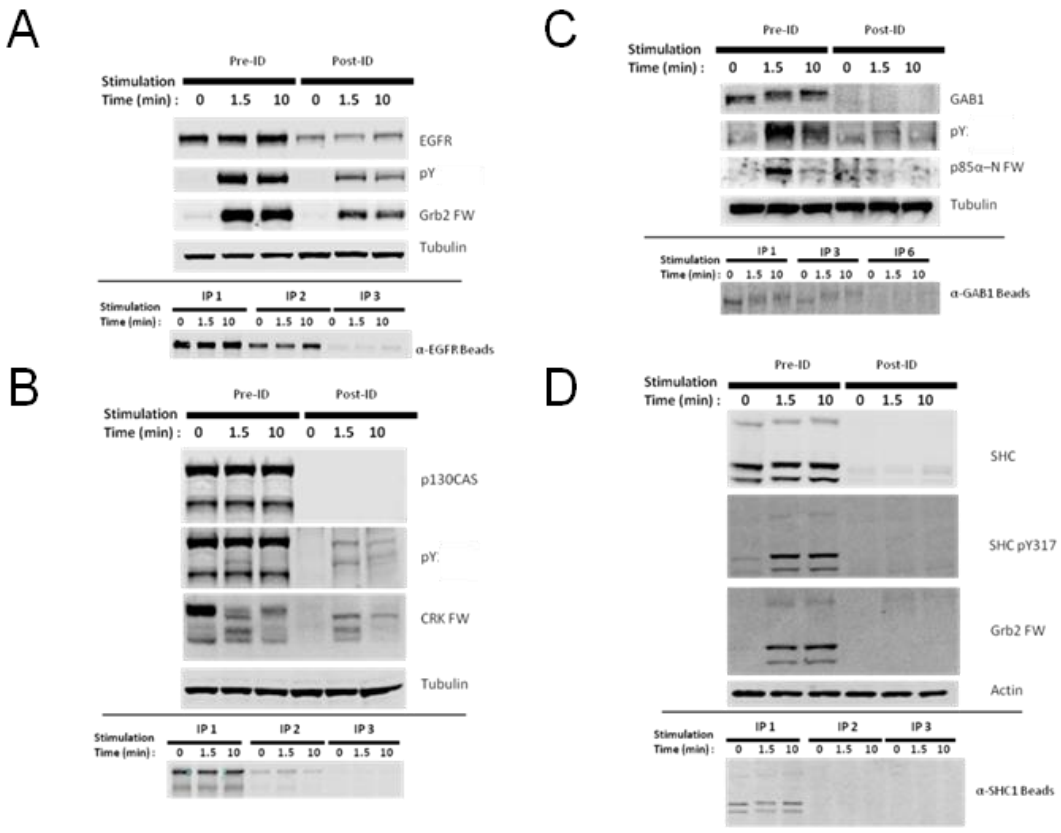

Figure 1.4 Identification of major SH2 domain binding bands by immunodepletion: A) EGFR, B) p130CAS, C) GAB1, and D) SHC1. Blots show expression, phosphorylation and representative far-Western for each protein before (pre-ID) and after (post-ID) immunodepletion. Immunodepletions were performed on lysates before (0 min) and after (1.5 and 10 min) EGF stimulation.

Quantification of SH2 domain binding sites

Far-Western provides a means to characterize the global binding patterns of multiple SH2 domains across multiple cellular states (*38*, *67*). To quantitatively compare changes in binding patterns of different SH2 domains upon EGF treatment, we performed unsupervised hierarchical clustering on SH2 domain binding data over the 60-minute EGF stimulation time-course (Figure 1.5 B). Six distinct SH2 domain-binding clusters (R^2 >0.85) were identified, which were broadly consistent with literature-reported interactions. For example, GRB2 SH2 and SHCA PTB domains most strongly interacted with EGFR (*123*, *124*), SHP2 and p85α/β SH2 domains interacted most strongly with GAB1 (*125*), and focal adhesion-associated proteins such as CRK, CRKL, NCK1 and NCK2 bound most strongly to p130CAS (*126*, *127*). All but three SH2 domains (RASGAP C-terminal, NCK1 and NCK2) showed at least some binding to EGFR. As expected, probes from homologous proteins clustered together (i.e. p85 α/β). However, SH2 probes isolated from proteins with two SH2 domains showed more variation (e.g. RASGAP-N, -C and -NC) (Table 1.2).

Tyrosine phosphorylation in EGF-stimulated A431 cells is dominated by phospho-EGFR (see Figure 1.3), and as a result most SH2 domains displayed strong EGFR binding. To better understand the relative affinity of each SH2 for various phosphoproteins, we normalized the quantified SH2 binding signal to pY binding (from anti-pY immunoblot) and reclustered the data (Figure 1.6). After normalization, EGFR no longer dominated binding. Only three domains (GRB2, SHCA PTB and to a lesser extent ARG) displayed a greater relative affinity for EGFR than would be expected from motif-independent pY binding (i.e. anti-pY blotting). Instead, most SH2 domains displayed greater than expected affinity for GAB1 and p130CAS. This analysis is designed to highlight SH2 binding specificity for particular protein targets; whereas the non-normalized FW results reflect a combination of binding specificity and the abundance of binding sites, and thus are more likely to predict the level of binding to different proteins *in vivo.*

See attached Excel files

Table 1.2:. Normalized FW, MS and *in vivo* imaging kinetics data. *Normalized Data-Interactive* tab provides a graph which allows for the comparison of data from FW, MS, imaging and pY EGFR immunoblotting. To use, select the desired data from drop down menu. To remove data select “Blank” located at the top of the menu. *Normalized Data* provides source data for the interactive chart (top) and all normalized data (below). Errors for FW and MS are shown as standard error of the mean and standard deviation, respectively. *FW Data* displays averaged binding quantifications for each SH2 probe. Data for each probe was normalized to the highest intensity band on each blot (Data used for Figure 1B). Error used is SEM. The number of replicate blots used is listed. *MS Data* tab provides sequence, Uniprot protein abbreviation and protein description for each peptide identified; indication of EGF dependence (two time points with Student’s t-test $p \leq .05$ and one time point with at least a two-fold increase compared to untreated samples); indication of sites not associated with EGF stimulation in PhosphoSitePlus database; and the number of biological replicates in which the peptides was detected. Phosphosite abundance data is normalized to sum of signal for all eight time points. Error is represented as standard or average deviation.

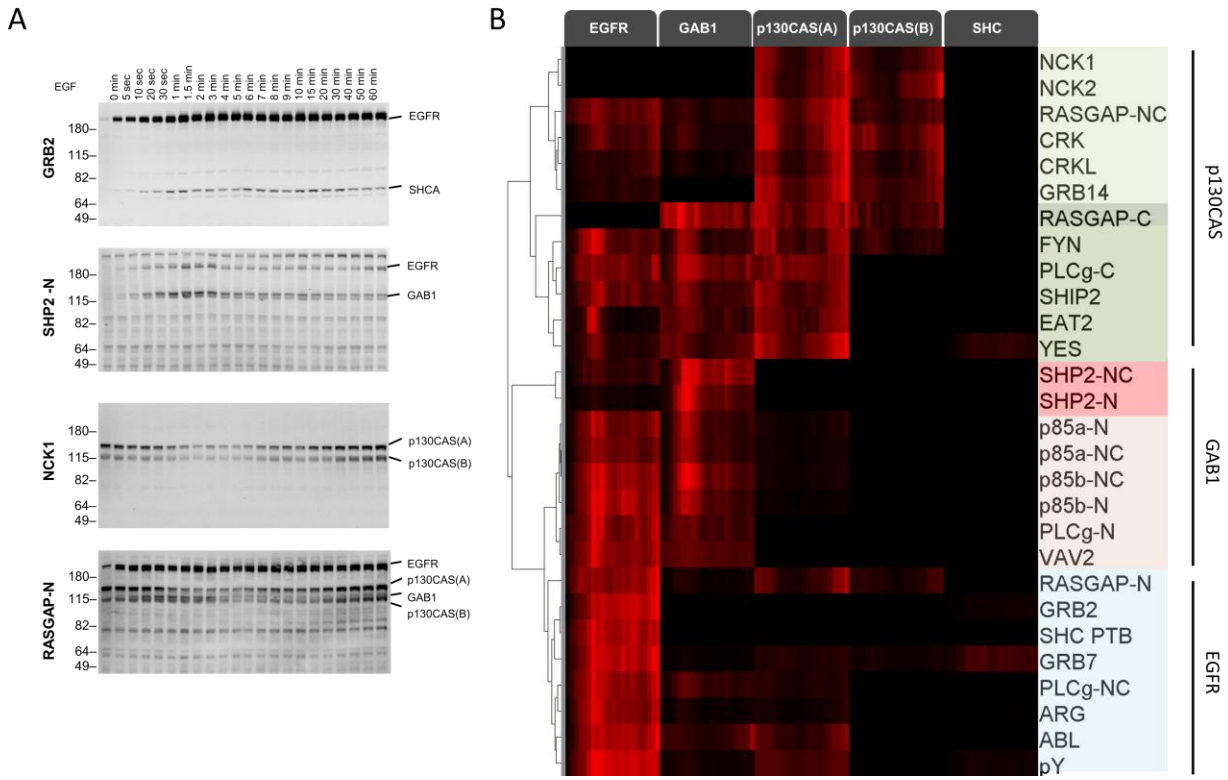

Figure 1.5: Dynamic EGF-dependent changes in tyrosine phosphorylation patterns revealed by SH2 domain far-Western analysis. A) Representative far-Western blots of 60-minute EGF stimulation time-course. Far-Westerns using GRB2, SHP2-N, NCK1 and RASGAP-N are shown to illustrate major binding patterns identified (see B) Immunoprepitation identified bands are shown to the left. B) Hierarchical clustering of SH2 domains on the basis of binding to four major phosphoproteins (EGFR, GAB1, p130CAS, and SHCA). Signal was normalized to maximum band intensity across all time points and all bands for each probe (red represents greater percentage of total signal, max = 1, min = 0). Names of SH2/PTB domain probes are indicated on the right. Colored boxes represent SH2 clusters defined by Pearson correlation coefficient > 0.85.

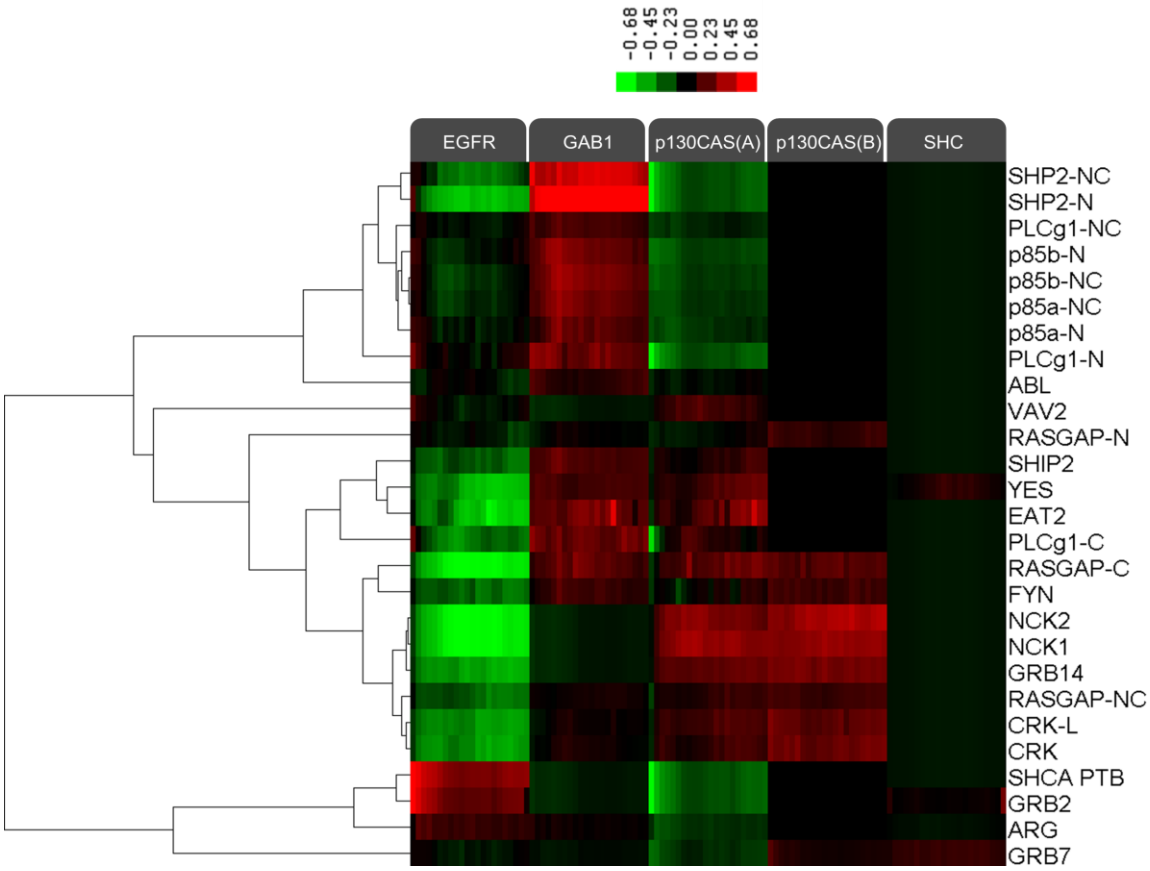

Figure 1.6: Hierarchical clustering of relative SH2 specificity. Values were obtained by subtracting percent bound by pY from percent bound by each SH2 for each time point. Red represents binding greater than pY and green represents binding less than pY. Relative SH2 specificity value = %Band intensity$_{\text{SH2 at time x}}$ - %Band intensity$_{\text{anti-pY at time x}}$.

pY-EGFR specific SH2 domain binding and downstream signaling

In principle, if different sites on a particular protein were phosphorylated with very different kinetics, different SH2 domains might bind to that protein with different kinetics. For the most part, however, we observed only modest differences in the temporal pattern of the binding of different SH2 domains to particular phosphoproteins. To better visualize the temporal variation in EGFR binding, SH2 domains were clustered based on changes in phospho-EGFR binding over time (Figure 1.7A). The majority of probes, including GRB2, ARG, p85 and the SHCA PTB, displayed dynamics similar to that of total EGFR phosphorylation, with a rapid and sustained increase in binding. However, a set of SH2 domains including CRK, CRKL, RASGAP and GRB14 showed a more gradual increase in phosphorylation over the 60-minute time-course. And a set of four SH2 domains including EAT2, SHP2-N and the only two Src-family kinases (SFK) tested, YES and FYN, displayed a rapid increase in binding followed by a rapid and sustained decrease (Table 1.2).

To test whether this variation was due to binding of SH2 domains to specific EGFR sites phosphorylated with different kinetics, we performed EGFR phosphosite-specific Western blotting and compared site-specific dynamics to those obtained by FW. Qualitatively, the phosphorylation dynamics of specific sites mirrored those of specific SH2 domains by far-Western (Figure 1.7 A and B). In particular, the kinetics of EGFR pY992, which is located within a canonical CRK binding motif (pYXXP), strongly correlated with CRKL probe EGFR binding kinetics (R^2=0.88). Both displayed an initial rapid rise to approximately half maximum, followed by a slow increase to maximum over the 60-minute time-course (Figure 1.7 C). Strong correlation was also observed between the dynamics of GRB2 SH2 domain binding and the GRB2 binding site EGFR pY1068 (R^2=0.91) (Figure 1.7 D) (*124*). SHCA PTB domain binding correlated well with total EGFR pY (R^2=0.89), consistent with a recent report suggesting that the SHCA PTB binding site on EGFR, pY1148, dominates receptor phosphorylation (*111*) (Figure

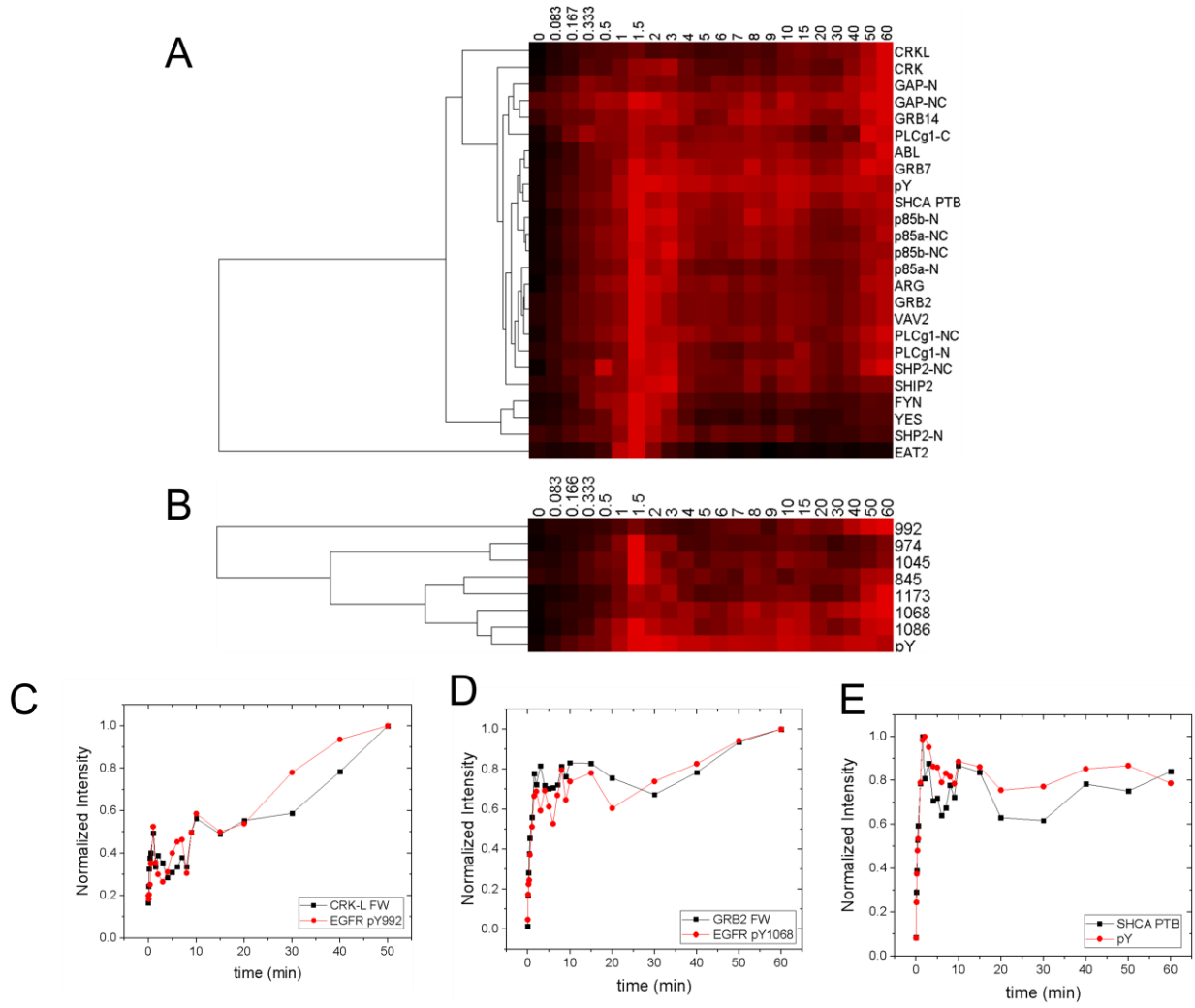

Figure 1.7: Comparison of pYEGFR-SH2 binding and EGFR site specific phosphorylation dynamics following EGF stimulation. A) Hierarchical clustering of SH2 binding to EGFR band (normalized to maximum, max=1, min=0) for 60-min time-course. B) Hierarchical clustering of EGFR phosphosite-specific Western blotting quantifications (normalized to maximum, max=1, min=0) for 60-min time-course. C and D) Plots comparing the EGFR phosphosite kinetics and SH2 binding kinetics for SH2 domains and their canonical binding motifs (C, CRKL and pY992, R^2=0.81; D, GRB2 and pY1168, R^2=0.91). E) Plot comparing total EGFR phosphorylation (pY EGFR band) with SHCA PTB binding (R^2=0.89). Correlation coefficients were calculated by plotting normalized pY Western intensity against normalized SH2 binding FW intensity.

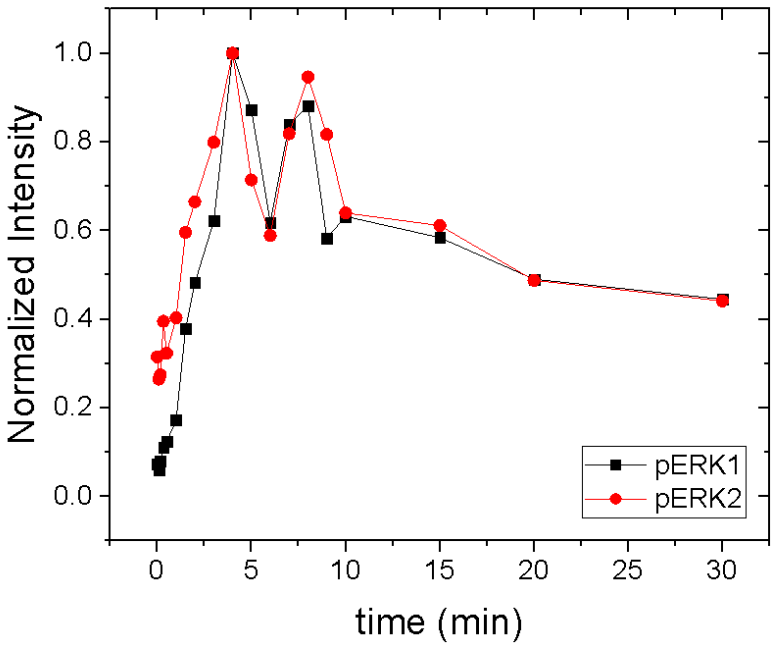

Figure 1.8. pERK1 and pERK2 pY/pT activation site phosphorylation kinetics following stimulation with EGF. Lines represent average of two technical replicates. Data is normalized to to maximum =1.

1.7 E). However, not all binding data could be predicted from known phosphosite specificities. For example, the far-Western binding kinetics of YES, FYN and EAT2 qualitatively resembled the rapid and transient phosphorylation EGFR pY974 and pY1045, providing evidence for potentially unappreciated interactions (Figure 1.7 A and B, Table 1.2).

Finally in order to understand the relationship between SH2 binding site creation and downstream signaling we assessed the activation time-course of ERK1 and ERK2, which are known to be activated after EGF stimulation via recruitment of proteins including SHP2, GRB2 and SHCA to EGFR (*58*). Activating phosphorylation of ERK1/2 reached a maximum approximately 4 minutes after EGF stimulation, compared with 1.5-2 minutes for tyrosine phosphorylation of EGFR (Figure 1.8).

Quantitative phospho-specific mass spectrometry

Although SH2-based far-Western blotting provides a unique insight into overall patterns of binding for different SH2 domains, it is highly dependent on phosphosite abundance, and does not identify specific phosphorylated sites. To obtain a complementary and less concentration-dependent view of EGF-induced changes in tyrosine phosphorylation, iTRAQ MS was used to quantify the phosphorylation kinetics of individual phosphosites. A431 cells were flash frozen and lysed at eight representative time points (0, 10s, 30s, 1 min, 1.5 min, 3 min, 10 min, 30 min) following EGF treatment. Prior to MS analysis, anti-pY immunoblotting was performed to ensure that phosphorylation kinetics and band patterns of MS samples were similar to those used for FW analysis (Figure 1.9). Lysates were then digested with protease, enriched for tyrosine-phosphorylated peptides by anti-pY immunoprecipitation, and purified using affinity chromatography prior to analysis by reverse phase liquid chromatography tandem mass spectrometry (LC-MS/MS). Relative phosphorylation levels of specific peptides were quantified

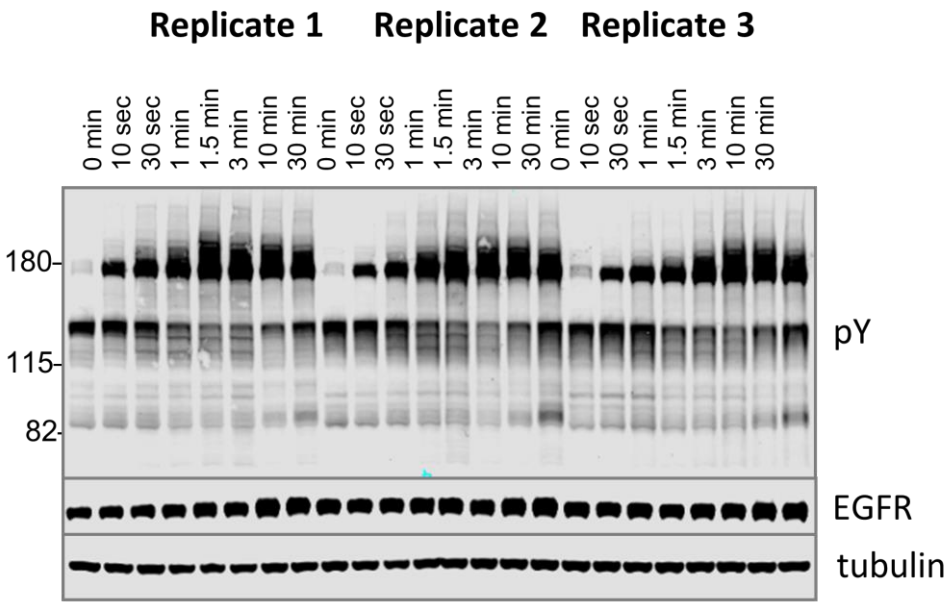

Figure 1.9 Anti-pY blot of three biological replicates analyzed by iTRAQ phospho-specific MS. Eight time points; 0, 0.167, 0.5, 1, 1.5, 3, 10 and 30 min post-EGF were analyzed. The kinetics and phosphorylation pattern observed is very similar to that observed in Figure 1.3.

by iTRAQ analysis using three biological replicates; only those phosphopeptides identified in all eight time points, in at least two replicates at each time point, were included in our analysis.

In total, 132 unique pY sites were identified from 93 proteins (Table 1.2). Of the phosphosites identified, 88 (67 proteins) displayed an EGF-dependent increase in tyrosine phosphorylation (defined by at least one time point with a two-fold or greater increase in abundance and a statistically significant increase in at least two time points following EGF treatment), including those from EGFR, GAB1 and the SH2-containing proteins SHCA (SHC1), PLCγ1, SHP2 and CRKL (Table 1.2). The percent of EGF-dependent phosphosites indentified in this study (65%) was higher than previous experiments using cell lines with more moderate EGFR expression (*77*, *112*, *128*). Among EGF-dependent sites identified in this study, 64 (57 proteins) were not found within a curated database of EGF-dependent phosphosites (*62*) (Table 1.2).

Gene ontology (GO)-based functional analysis of EGF-responsive phosphoproteins identified enrichment for terms associated with positive regulation of signaling and peptidyl tyrosine modification (*129*). EGF-nonresponsive sites, on the other hand, tended to be associated with regulation of cell adhesion and locomotion. However, a large number of terms were shared by the two protein sets, indicating significant functional overlap (Figure 1.10). By comparing the amino acid frequencies surrounding phosphotyrosines for each group, we found that EGF-nonresponsive site were more likely to contain the CRK SH2 binding motif (pYXDP/L), consistent with the functional association of these proteins with adhesion and locomotion. EGF-responsive peptides however, displayed little sequence consensus (Figure 1.11).

Phosphopeptide analysis of FW-identified proteins

iTRAQ analysis identified multiple sites from the four major phosphoproteins detected by FW (Figure 1.12 and Table 1.2). Comparison of MS and FW data for these proteins

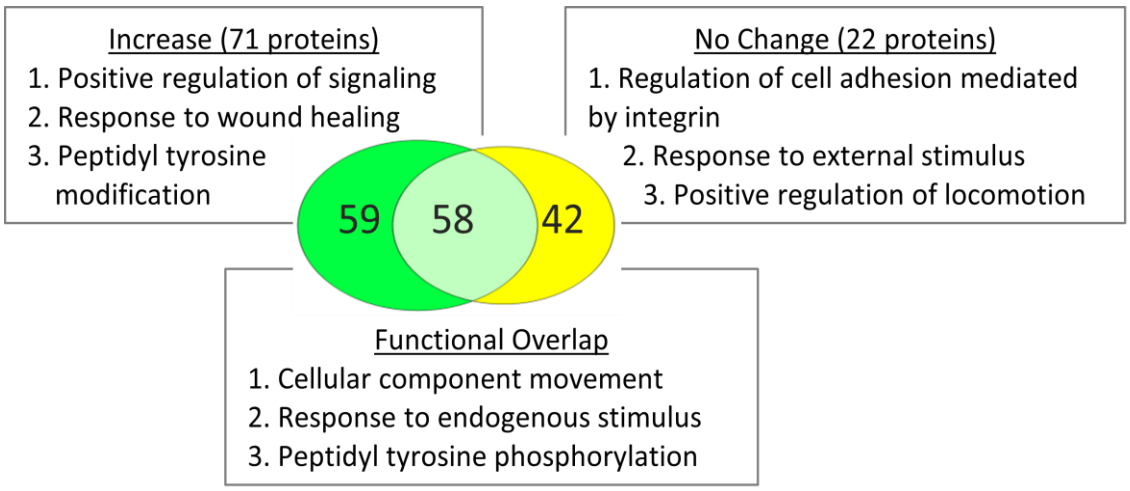

Figure 1.10 Venn diagram showing overlap of significant gene ontologies for proteins containing peptides whose phosphorylation was enhanced or unchanged by EGF ($p \leq 0.05$, Bonferroni corrected). The number of unique or overlapping ontologies observed for each protein set is indicated within the diagram. GO terms listed represent the three largest GO parent terms returned by REVIGO (*129*).

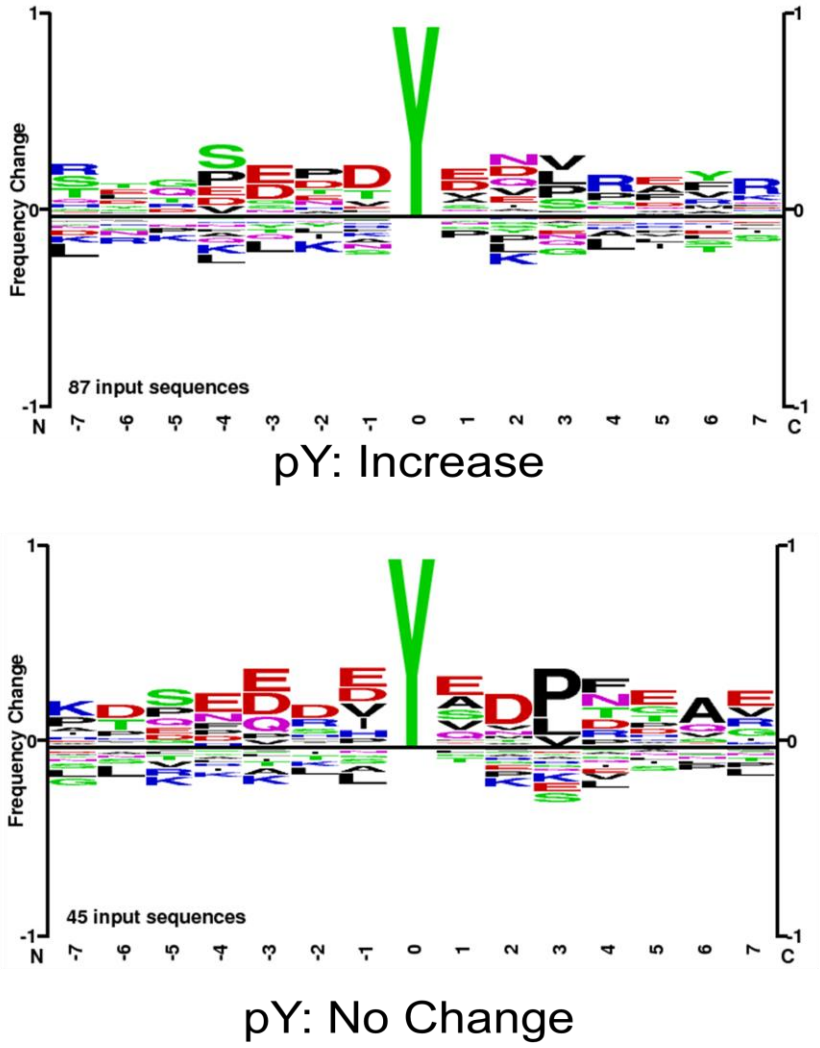

Figure 1.11 Amino acid frequency logos for sites whose phosphorylation was enhanced (upper) and unchanged (lower) by EGF stimulation. Logos were created with PhosphositePlus logo generator. Amino acid abundance data is realitve to PhosphoSitePlus pY database background.

revealed both similarities and differences. Four unique EGFR pY peptides and one tyrosine/serine dually phosphorylated peptide were identified (Figure 1.12 A). The tyrosine-only sites tended to display a continual increase in abundance over the first 10 minutes, compared to the more rapid plateau in signal intensity seen in FW and anti-pY (Figures 1.3, 1.9, Table 1.2). The dually phosphorylated peptide varied significantly over the time-course and was not scored as EGF-dependent. Five unique phosphopeptides were identified for GAB1, all of which increased rapidly, but did not show the rapid dephosphorylation seen in anti-pY blots and FW (Figure 1.12 B). Like the EGFR pY peptides, the single SHCA phosphopeptide, a known GRB2 binding site (*130*), displayed slower phosphorylation kinetics than those obtained from anti-pY and FW blots (Figure 1.12 C). MS phosphopeptide data also did not recapitulate the p130CAS dephosphorylation and rephosphorylation pattern seen on anti-pY immunoblots and SH2 far-Western. Instead, the six p130CAS phosphopeptides identified by iTRAQ MS trended toward an increase in phosphorylation following EGF stimulation, though only 2 of the 6 displayed a statistically significant change in abundance (Figure 1.12 D, Table 1.2). Overall, FW blotting and MS-based quantification of individual sites provided rather different perspectives on tyrosine phosphorylation dynamics, highlighting the importance of using complementary experimental approaches. These results suggest that changes in the binding of an SH2 domain to a multiply phosphorylated protein can be difficult to predict precisely from changes in the phosphorylation of individual phosphosites on that protein.

Finally, most phosphopeptides showed a transient decrease in phosphorylation at the 3 min time point, similar to the dip seen on far-Westerns and anti-pY Westerns, especially for EGFR and SHCA, at approximately 4 min (Figure 1.2, 1.7 C-E and Table 1.2). The fact that a similar pattern is seen using two very different experimental approaches increases confidence that it accurately reflects the actual behavior of the system.

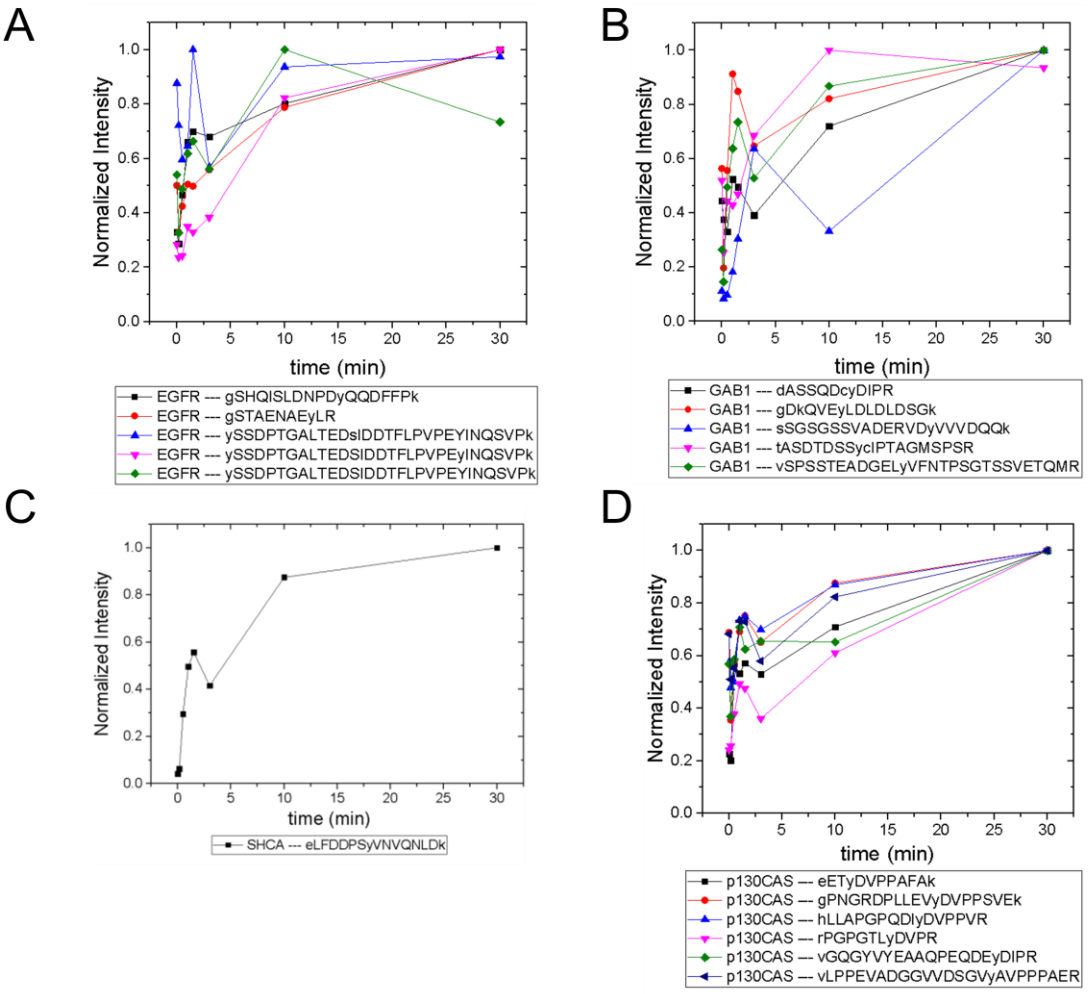

Figure 1.12. Relative MS quantified abundance of phosphosites from major phosphoproteins identified by FW. A) EGFR, B) GAB1 C), SHCA (F) and D) p130CAS (BCAR1). Specific phosphopeptide sequences are listed. Abundances are all normalized to maximum = 1.

In vivo membrane recruitment of SH2 domains

FW and MS data revealed dynamic changes in potential SH2 binding sites in response to EGF, but provide little insight into the actual binding kinetics of SH2 domain-containing proteins in living cells. To address this issue, we quantified SH2 basal membrane recruitment kinetics following EGF stimulation in A431 cells for a similar set of 25 SH2 domains, each fused with the photoactivatable fluorescent protein tdEOS. TIRF microscopy and sptPALM were used to identify SH2 localization patterns and quantify membrane binding, diffusion constants (D) and apparent membrane dissociation rates (λ) (Tables 1.2 and 1.3). It is important to note that all SH2-containing proteins contain additional interaction domains that may affect the recruitment kinetics of the full length proteins *in vivo*; however in order to directly compare results with SH2 binding site data from the FW and MS studies above, we assayed *in vivo* binding using isolated, ectopically expressed SH2 domains. In the case of GRB2, we previously showed that full-length GRB2 and the SH2 domain alone were very similar in their *in vivo* membrane binding behavior (*94*), indicating that for this protein the kinetics were dominated by SH2-pY interactions.

Live-cell imaging of EGF-induced SH2 domain membrane binding revealed a wide variety of dynamic behaviors among the different SH2 domains tested. However, a number of trends were apparent in light of the binding specificities identified by FW hierarchical clustering. SH2 domains that predominantly bound GAB1 (e.g. SHP2-C/-NC, SHIP2) tended to reach maximum binding rapidly, diffuse rapidly upon binding, and display a relatively diffuse spatial distribution before stimulation, and a mix of clustered and diffuse localization after stimulation. EGFR-binding SH2 domains (e.g. GRB2 and GRB7) displayed relatively slow recruitment kinetics and diffusion rates, and their binding sites tended to cluster together into discrete foci in response to EGF. p130CAS-binding domains (e.g. CRK, NCK, PLCγ-C) generally had either slower recruitment kinetics or exhibited small to no net recruitment following stimulation; they also

	in vivo (single molecule and TIRF imaging)			*in vitro* (FW
SH2 Domain	**recruitment time τ (min)**	**λ (s^{-1})**	**D ($\mu m^2/s$)**	**phosphorylation time τ (min)**
SHIP2	0.55 (0.15)	decay	0.04	1.13 (0.81)*
SHP2-N	0.91 (0.02)	0.96	0.187	0.60 (0.01)
GRB14	0.99 (0.47)	decay	-	-1.50 (0.08)
SHP2-NC	1.57 (0.18)	decay	0.081	1.57 (0.18)
SHP2-C	1.85 (0.49)	2.16	0.759	ND
EAT2	2.26 (1.10)	1.48	0.097	0.64 (0.51)*
PLCγ1-NC	2.44 (0.18)	1.01	0.022	0.62 (0.19)
p85α-NC	3.06 (0.32)	0.31	0.004	0.51 (0.09)
SHC PTB	3.25 (0.08)	0.36	0.007	0.69 (0.08)
GRB7	3.45 (0.67)	0.46	0.009	0.77 (0.05)
VAV2	4.27 (0.52)	0.64	0.034	0.80 (0.17)
SHC SH2	4.57 (0.20)	0.64	0.026	ND
GRB2	4.58 (0.41)	0.59	0.021	.059 (0.03)
PLCγ1-N	5.15 (2.28)	0.08	0.022	0.75 (0.33)
CRK	5.54 (0.64)	0.2	0.016	-2.80 (0.29)
RASGAP-NC	5.92 (0.80)	decay	0.076	-1.27 (0.54)*
PLCγ1-C	6.48 (1.27)	0.029	0.043	0.30 (0.03)
ARG	6.55 (0.90)	decay	0.01	0.54 (0.20)
p85α-N	6.88 (0.78)	decay	0.016	0.67 (0.10)
RASGAP-N	6.99 (2.53)	0.44	0.011	-2.51 (0.98)
FYN	constant	ND	0.008	0.43 (0.04)
NCK1	decrease	ND	0.009	-0.83 (0.03)
RASGAP-C	constant	ND	ND	1.27 (0.54)*
ABL	constant	ND	ND	0.70 (0.40)
YES	ND	ND	ND	-3.62*(1.48)

Table 1.3: Quantification of *in vivo* SH2 binding dynamics and binding site kinetics. To directly compare FW and *in vivo* imaging data, both SH2 binding site phosphorylation and *in vivo* SH2 membrane recruitment curves were fit to the first order exponential recovery function ($1-e^{-t/\tau}$), where τ is the binding time constant and t is time after EGF. D is the average diffusion rate of SH2 molecules on the membrane. λ is the apparent membrane dissociation rate constant. *Decay* identifies those SH2 domains with decreasing apparent dissociation rate constants (i.e. a decaying hazard rate function). * denotes data which fit poorly to the recovery function (R-square < 0.5). ND denotes data which was not determined. SEM reported in parentheses.

primarily localized to structures resembling focal adhesions (Figures 1.13A, 1.14 A-C, Tables 1.2 and 1.3).

To compare SH2 recruitment kinetics of different SH2 domains, time constants were calculated for the recruitment curves obtained for each SH2 domain. Membrane fluorescence measurements following EGF treatment were plotted with respect to time and fitted to the first order exponential recovery function, $1\text{-}e^{(-t/\tau)}$, where t equals the EGF stimulation time. The time constant (τ) is proportional to the time required to reach maximal binding. SH2 domain membrane recruitment time constants negatively correlated with both disassociation and diffusion rates (Table 1.3, Figure 3, Figure 1.14 D and E). In other words, those SH2 domains that were recruited most quickly tended to have the highest apparent diffusion and dissociation rates after binding.

One surprising result was that for most SH2 domains, *in vivo* binding kinetics lagged significantly when compared to the kinetics of SH2 binding site creation, as measured by far-Western (Figure 1.13 B, Table 1.2). On average, time constants calculated from membrane fluorescence intensity were ~6 times greater than time constants for the same domains calculated from far-Western blotting, and *in vivo* time constants had a much greater range when compared with those obtained by FW (Table 1.3). This lag was unexpected, as simple calculations using best-estimate SH2 concentrations and rate constants (k_{on} and k_{off}) suggested that binding and unbinding should equilibrate rapidly, within seconds. Using the GRB2 SH2 as an example, we examined two potential mechanisms that could give rise to this discrepancy, both which of turned out to be unsubstantiated. First, unexpectedly low concentrations of GRB2 SH2-tdEOS and phosphorylated EGFR *in vivo* could prolong the time to EGF-induced equilibrium. However, direct biochemical measurements of GRB2 SH2-tdEOS and pY-EGFR in these cells (6.5 μM and 1.5 μM, respectively) did not support this theory (Figure 1.15).

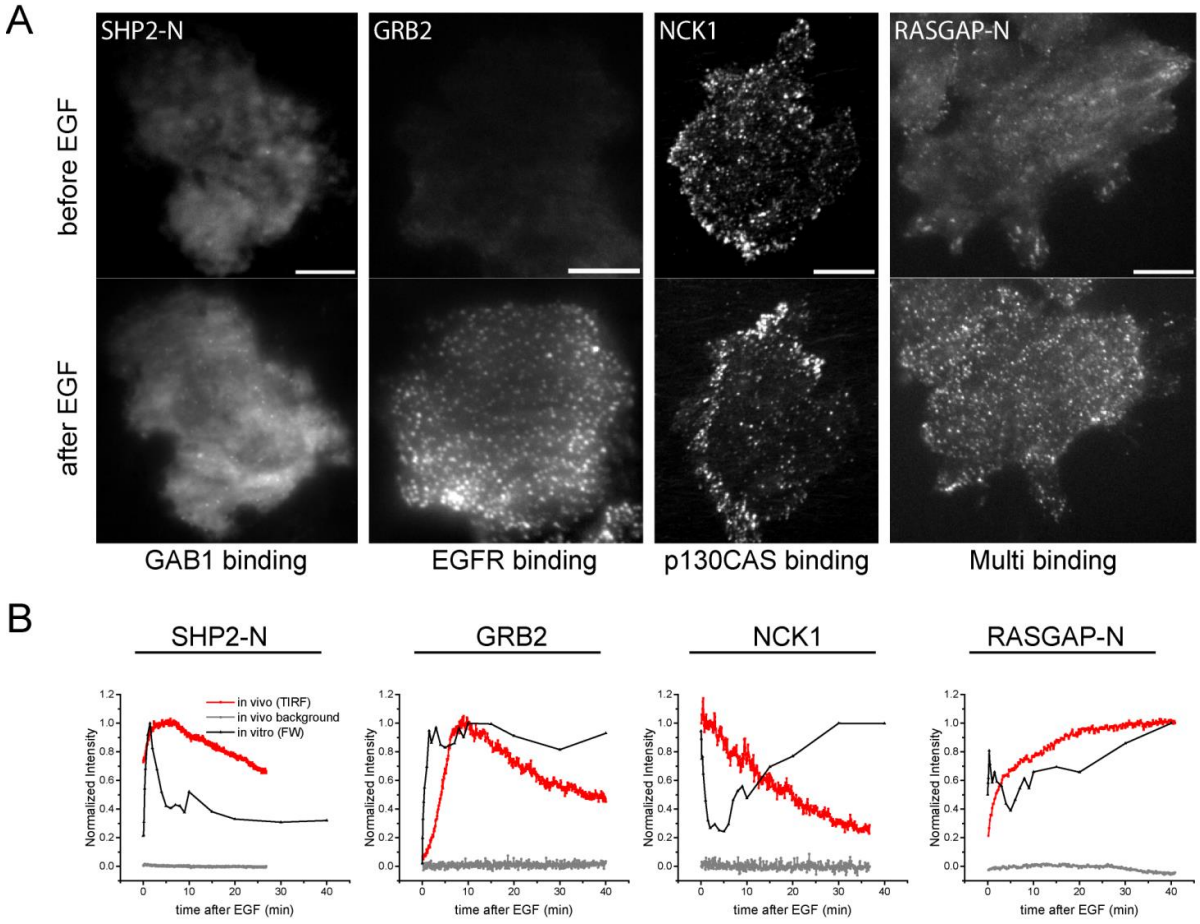

Figure 1.13: *In vivo* localization and recruitment kinetics of SH2 domains. A) Representative total internal reflection fluorescence (TIRF) microscopy images of fluorescently tagged SH2 domains before and 40 min after EGF stimulation, for SHP2-N, GRB2, NCK1 and RASGAP-N SH2 domains. B) Comparison of change in total membrane SH2 fluorescence from imaging live cells (*red*) and change in FW-based SH2 binding (*black*) following EGF stimulation. *Gray* lines indicate TIRF background signal. Data is normalized to maximum.

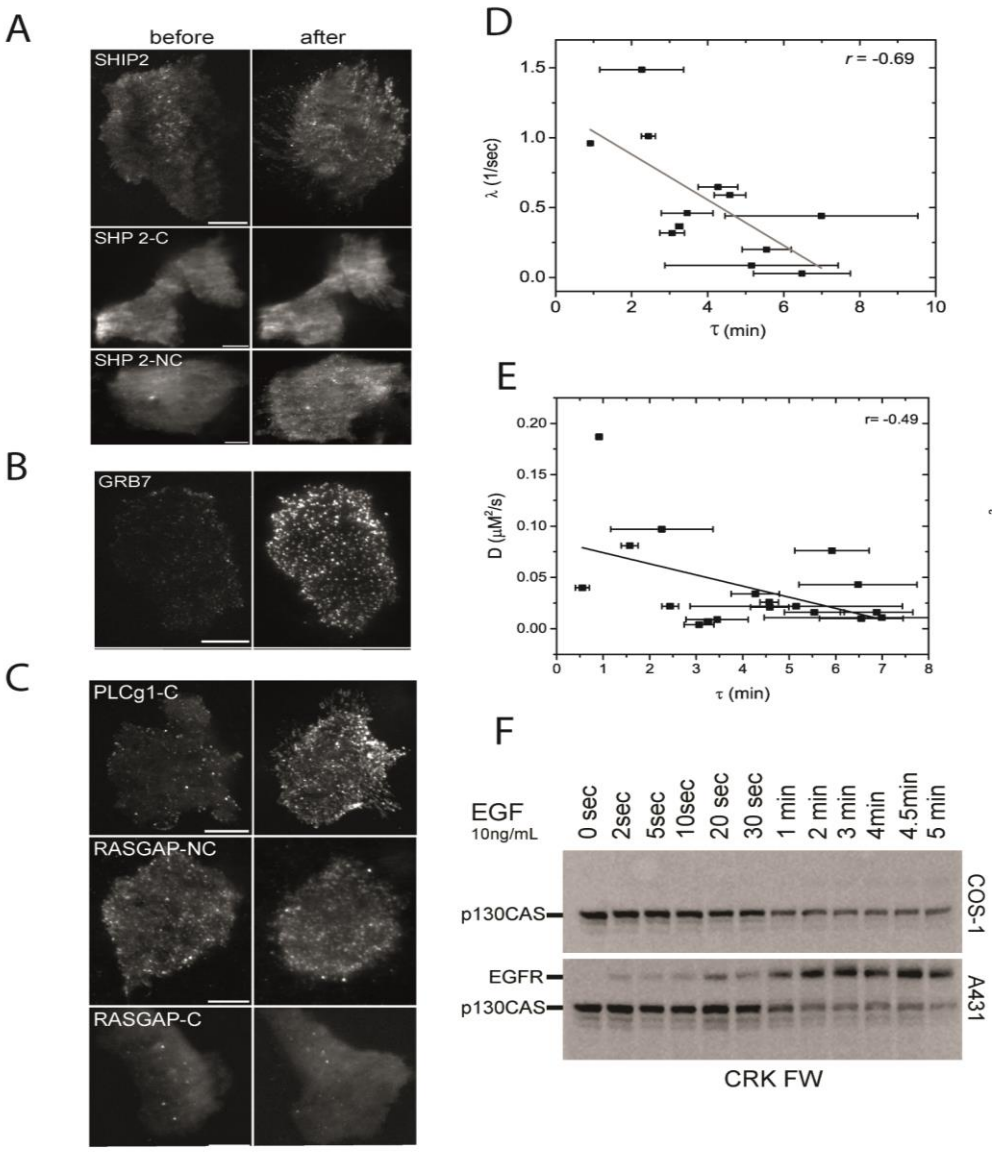

Figure 1.14 Analysis of *in vivo* SH2 domain localization and membrane binding. A-C) TIRF images of fluorescently tagged SH2 domains before and after EGF stimulation. A) GAB1 binding domains B) EGFR binding domains, and C) p130CAS binding domains. Domains were grouped according to clustering results from Figure 1B. Post-EGF images were taken ~40min after stimulation. D) Correlation of SH2 domain apparent membrane off-rate (λ, y-axis) and SH2 domain membrane recruitment time constant (τ, x-axis). E) Correlation plot of SH2 domain probe diffusion rate (D, y-axis) and recruitment time constant (τ, x-axis). Data for SHP2-C was

an outlier and was removed from the plot for clarity. F) CRK FW time-course blots from A431 and COS1 cells simulated in parallel with 10 ng/mL EGF. CRK SH2 shows little or no binding to the EGFR band in COS1 cells.

Then, we ruled out the possibility that SH2 binding site creation on the basal membrane (where *in vivo* binding is measured by TIRF) is much slower than on the apical membrane by quantifying confocal images of EGF-stimulated cells immunostained with anti-pY antibody (Figure 1.16).

It is also worth noting that most p130CAS-binding domains showed increased membrane binding in response to EGF, even as total binding site availability (as measured by FW) declined over the first 5-10 minutes (Figure 1.2, Table 1.2). This is likely the result of the relatively high concentration of EGFR phosphosites in the EGF-stimulated A431cells (Figure 1.14 F). One exception however was NCK1 SH2, which did in fact display an initial decrease in membrane binding upon EGF treatment, though it did not exhibit the late increase (10-60 min) or occur at the same rate seen in FW (Figure 1.13 B, *panel 3*, Table 1.2). Nevertheless, multiple parameters obtained from *in vivo* imaging were consistent with p130CAS being the major NCK1 SH2 binding partner. Diffusion and motility measurements of fluorescent p130CAS and NCK1 SH2 were remarkably similar (Figure 1.17 A and B). In addition, composite images created from sptPALM movies showed both the NCK1 SH2 domain and p130CAS localizing to focal adhesion-like structures on the cell periphery (Figure 1.17 C). Taken together, these results suggest that the SH2 domain binding *in vivo* is a phosphosite-dependent, but not fully equilibrated, process.

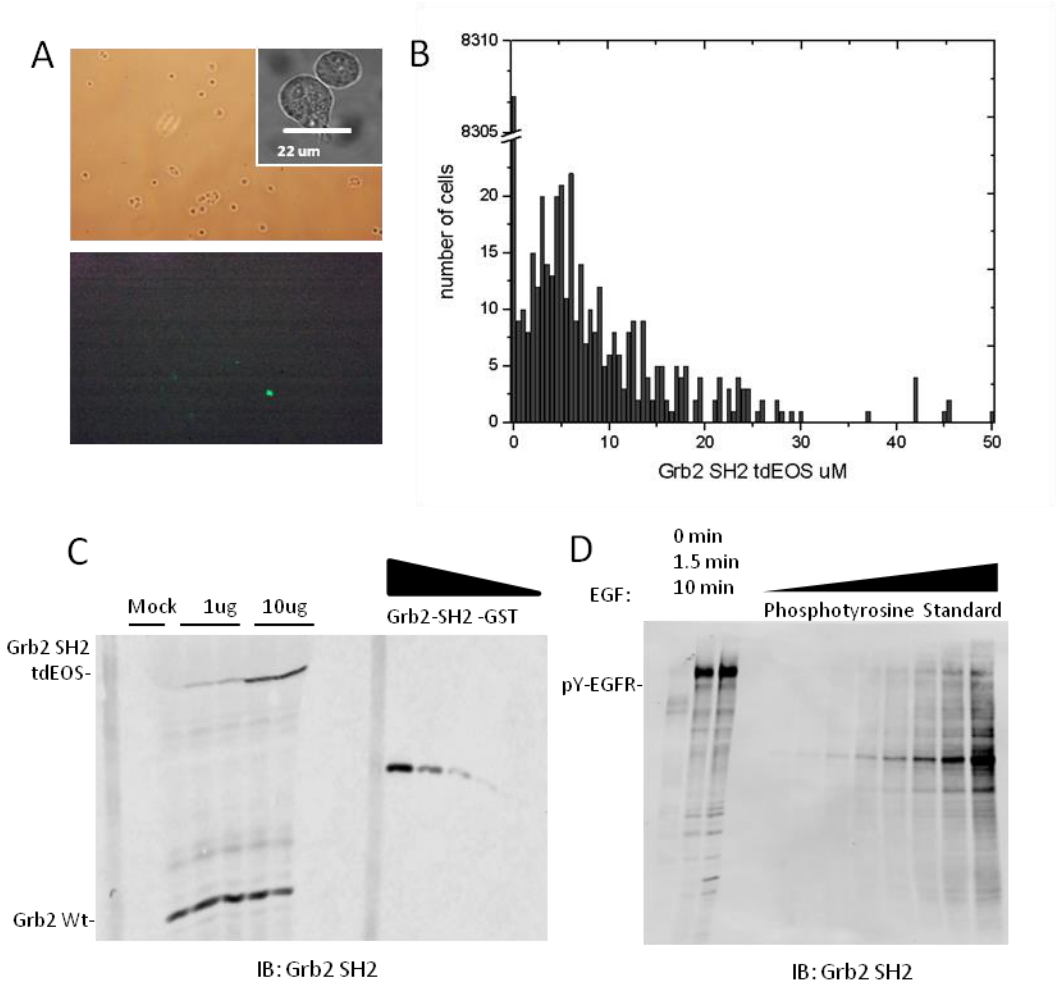

Figure 1.15 . Quantification of GRB2 SH2-tdEOS in A431 cells. A) Representative white light *(top panel*) and epifluorescence images *(lower panel)* of GRB2 SH2-tdEOS transfected A431 cells that were used to determine total cell number, transfection efficiency and relative expression level. Insert *(upper panel)* shows representative DIC image of nonadherent cells used to determine cell volume. B) Histogram of individual cell GRB2 SH2-tdEOS expression levels. Left skew in expression was compensated for in the final calculation. C) Anti-GRB2 SH2 blot used to calculate average GRB2 SH2-tdEOS concentration D) Anti-pY blot showing EGF-induced EGFR phosphorylation and phosphorylation standard titration used to calculate the cellular concentration of phosphorylated EGFR sites.

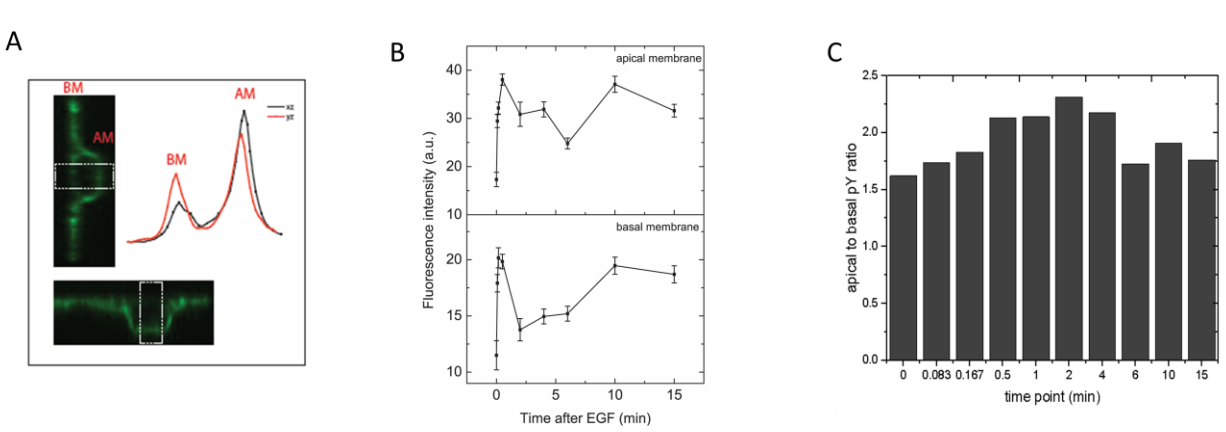

Figure 1.16 Quantification of pY on the basal membrane of A431 cells following EGF. A) Representative z-axis cross-sections of fixed A431 cells immunostained with anti-pY. The images and traces were obtained from the same cell along the x- and y-axes. White block indicates the quantified area. Curves represent an average of multiple line scan quantifications across an individual cell membrane. B) Apical and basal pY levels following EGF stimulation as measured by immunofluorescence. Intensity measurements from at least 10 cells were averaged for each time point. C) Ratio of apical to basal phosphorylation following stimulation with EGF.

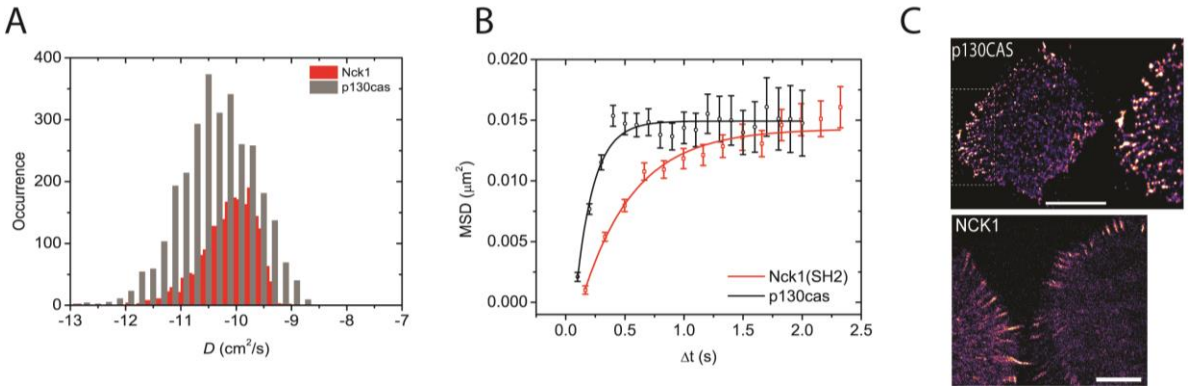

Figure 1.17. *In vivo* NCK SH2 and p130CAS display similar membrane interaction dynamics. A) Histogram of individual NCK1 and p130CAS molecule diffusion rates post-EGF. B) Mean square displacement (MSD) plots of p130CAS (*black*) and NCK1 SH2 domain (*red*) post-EGF. C) sptPALM images of tdEOS-labeled p130CAS and NCK1 SH2-tdEOS in A431 cells before EGF stimulation.

Changes in total cellular phosphorylation correlate with the apparent SH2 on-rate

If SH2 membrane binding *in vivo* is indeed far from equilibrium, then maximum phosphorylation and maximum binding need not be coincident. Instead, maximum phosphorylation may be more closely correlated with the maximum *rate* of recruitment, i.e. γ_{on} (t) = k_{on}(t)[pY][SH2], where γ_{on} (t) is the apparent on-rate at time t. To test this we directly measured the *rate* of binding by utilizing sptPALM to count the number of new SH2 molecules appearing at the membrane during a small time window (*97*). For these experiments, we utilized the GRB2 SH2 domain to minimize effects from binding to non-EGFR phosphoproteins (Figure 1.2 and 1.5). Using this method we found that the binding rate of GRB2 SH2 increased much more rapidly (τ=2.08 min) than total binding (Figure 1.18 A, *black line*). To corroborate this rate measurement, we attempted to recapitulate the experimentally obtained changes in GRB2 SH2 abundance at the membrane by combining apparent on-rate measurements, with apparent off-rate measurements (λ_{off}) calculated using a previously published method (Figure 1.18 A, *red line*) (*94*). This relationship can be represented as follows:

$$\frac{\mathrm{d[mem:SH2]}}{\mathrm{dt}} = \gamma_{on}(t) - \lambda_{off}(t) \cdot \ [mem{:}\,SH2](t) \quad (1)$$

The solution of this nonhomogeneous first order differential equation being:

$$[\mathrm{mem{:}\,SH2}](\mathrm{t}) = \mathrm{ce}^{-\int \lambda_{off}(\mathrm{t})\mathrm{dt}} + \mathrm{e}^{-\int \lambda_{off}(\mathrm{t})\mathrm{dt}} \int \gamma_{on}\,(\mathrm{t}) \cdot \mathrm{e}^{\int \lambda_{off}(\mathrm{t})\mathrm{dt}}\mathrm{dt} \quad (2)$$

Plugging in the measured values for $\gamma_{on}(t)$ and $\lambda_{off}(t)$ for GRB2 SH2 returns a SH2 membrane binding [Mem:SH2] curve with a relatively slow rate of recruitment. This rate was similar to that obtained experimentally (Figure 1.18 B), indicating that SH2 recruitment is likely kinetically controlled and not an equilibrated process.

GRB2 SH2 domain binds rapidly to non-clustered sites

We previously reported that clustering of SH2 binding sites was associated with a decrease in the apparent membrane dissociation rate (increase in dwell time) (*94*). We proposed that this was due to increased SH2 rebinding to phosphosites that were more closely packed upon clustering. It was therefore plausible that clustering might play a role in the apparent delay in maximal recruitment of SH2 domains to membrane binding sites. Consistent with this idea, analysis of GRB2 SH2 cluster size and cluster number in EGF-treated cells showed that cluster formation reaches a maximum at 10-15 min (Figure 1.18 C), a time-scale coincident with that of the recruitment of GRB2 SH2 to the cell membrane. Furthermore, as mentioned above, we found that the recruitment time constants for individual SH2 domains negatively correlated with their apparent dissociation rates, which we previously reported to depend on the extent of phosphosite clustering (Figure 1.14 D, Table 1.3) (*94*). These results suggested that maximal membrane recruitment of SH2 domains may lag behind the generation of phosphorylated SH2 binding sites due to the relatively slow clustering of those sites.

To assess more directly the role of clustering in SH2 recruitment kinetics, we counted the number of molecules detected within cluster regions on sptPALM images. We found that the binding rate within clusters increased at a relatively a slow pace, reaching maximum at approximately 10.5 min, with a time constant of $\tau_{cluster}$ = 4.05 min (Figure 1.18 D, *red*). In contrast, the binding rate in non-cluster regions, obtained by subtracting cluster-associated binding events from the total binding events detected, reached a maximum at approximately 2.5 min ($\tau_{\text{non-cluster}}$ = 1.08 min) (Figure 1.18 D, *blue*). These results strongly suggest that binding to non-clustered phosphosites is relatively fast and tracks with the level of phosphosites available, while increases in *in vivo* binding seen at later points (after 1-2 minutes) can be attributed to increased binding to clustered sites as they accumulate.

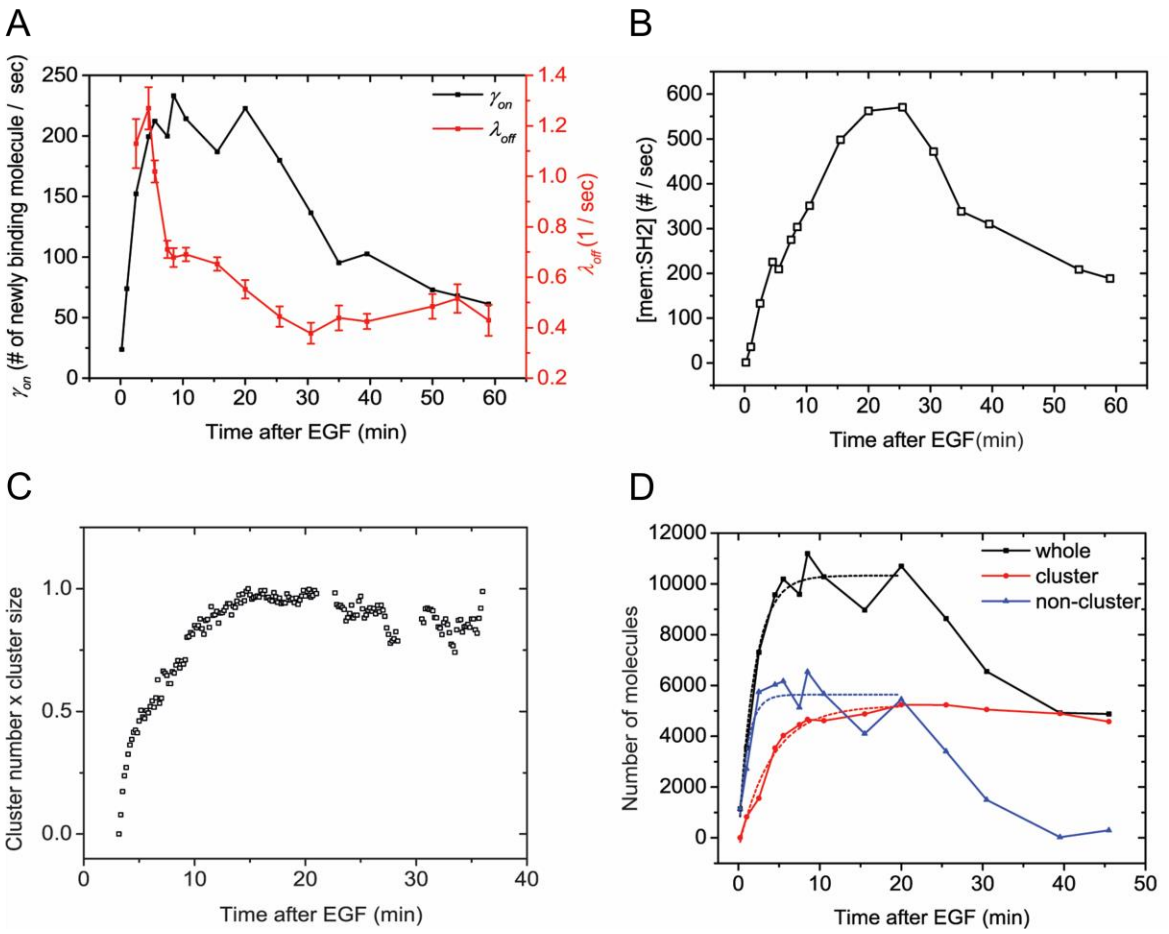

Figure 1.18. New molecule membrane recruitment and cluster-associated recruitment *in vivo*. A) Apparent membrane binding rate (g_{on}, *black line*) and apparent membrane dissociation rate (l_{off}, *red line*) following stimulation with EGF. B) SH2 membrane binding curve calculated using experimentally determined g_{on} and l_{off} values. C) Kinetics of GRB2 SH2 binding site clustering (cluster size x cluster number) after EGF treatment. D) Number of newly recruited GRB2 SH2 molecules (*black, whole*), and those within clusters (*red, cluster*) and outside of clusters (*blue, non-cluster*) after EGF stimulation. Dotted lines show fit with exponential recovery function.

Pervanadate-treated cells display reduced GRB2 binding site clustering and rapid recruitment

To further validate the relationship between SH2 recruitment and EGFR clustering, we sought a way to measure GRB2 SH2 domain recruitment in the absence of clustering. Binding of EGF to EGFR has been shown to induce dimerization, as well as higher order clustering of the receptor in a kinase-activity-dependent manner (*60*, *115*, *119*, *120*). We attempted to bypass EGF-induced receptor multimerization by treating cells with the competitive irreversible tyrosine phosphatase inhibitor pervanadate (PV), which increases cellular pY levels by blocking dephosphorylation. PV treatment led to a rapid increase in phosphotyrosine on a large number of proteins as shown by anti-pY immunoblot, but FW blotting with the GRB2 SH2 demonstrated that the EGF receptor was by far the major GRB2 SH2 binding protein in PV-treated cells (Figure 1.19 A). We confirmed the lack of GRB2 SH2 binding site clustering in PV-treated cells using a variety of measures. Firstly, GRB2 SH2-tdEOS showed a diffuse spatial distribution in PV-treated cells (Figure 1.19 B). Secondly, individual membrane-associated GRB2 SH2 molecules were much more mobile (i.e. had a higher diffusion rate) in PV-treated cells than in cells treated with EGF (Figure 1.19 C). Finally, the apparent dissociation rate of GRB2 SH2 was significantly greater in PV-treated cells (Figure 1.19 D). These results were all consistent with a lack of GRB2 SH2 binding site clustering in PV-treated cells.

In the absence of apparent binding site clustering, GRB2 SH2 membrane recruitment in PV-treated cells occurred rapidly (τ=0.81 + 0.15 min) (Figure 1.19 E). The rate of membrane binding in these cells was comparable to the rate of GRB2 binding site creation in PV-treated cells, as measured by GRB2 FW (Figure 1.19 A, E, G). New molecule recruitment was only marginally faster (τ=0.58 ±0.17) than total recruitment, indicating that little or no GRB2 SH2 rebinding was occurring at the plasma membrane (Figure 1.19 F). These results are consistent

with the interpretation that delayed SH2 domain binding kinetics in EGF-treated cells are the result of increased binding site clustering over time.

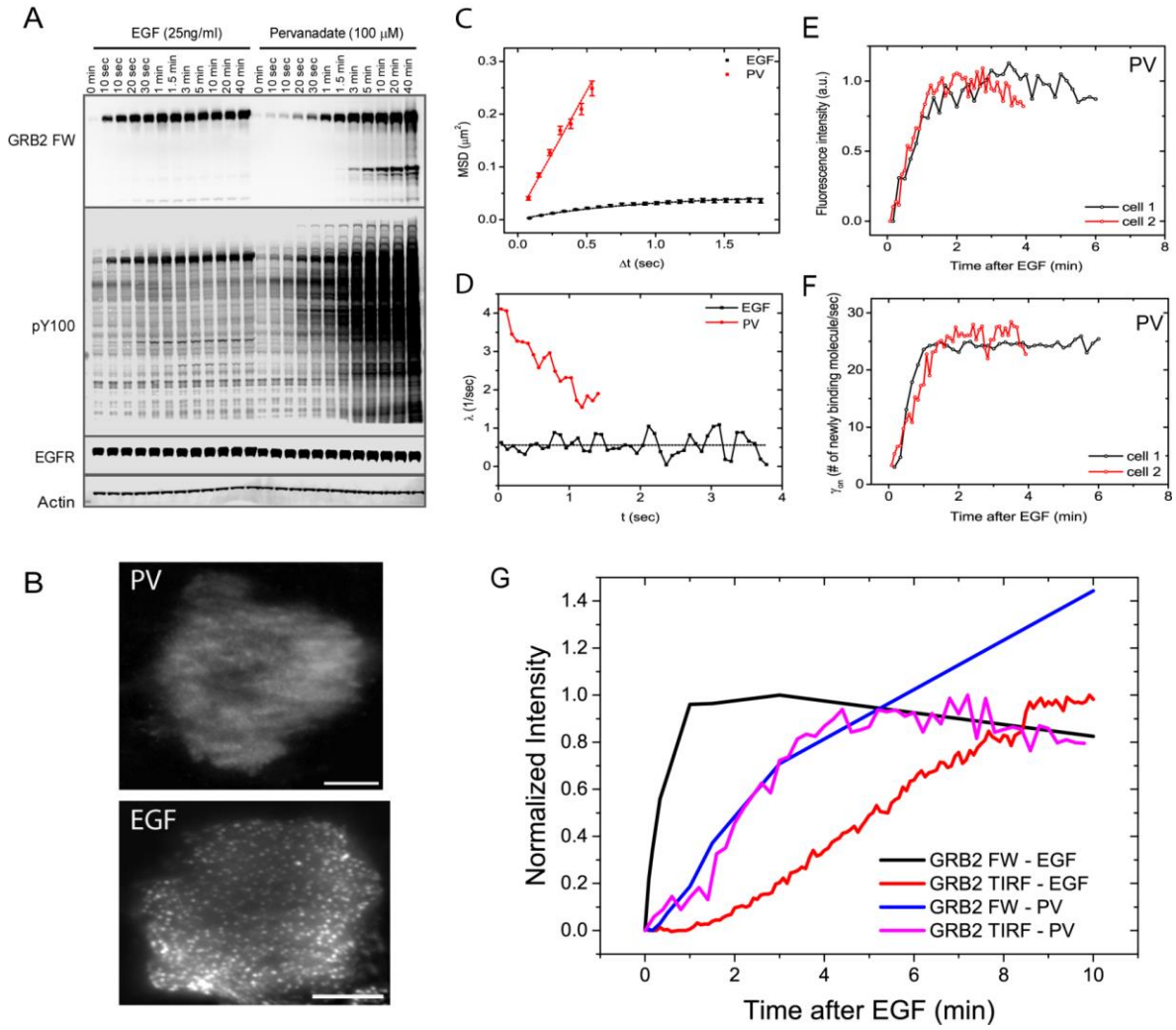

Figure 1.19. GRB2 SH2 recruitment in pervanadate (PV) treated cells. A) Representative GRB2 SH2 far-Western and anti-pY immunoblot for EGF and PV treated cells. B) TIRF microscopy images of fluorescently tagged GRB2 SH2 in pervanadate-treated (PV)- and EGF-stimulated cells 40 min post-stimulation. C) Mean square displacement (MSD) of GRB2 SH2 in EGF- (*black*) and pervanadate-treated cells (*red*). D) Change in mean apparent membrane dissociation (i.e. hazard rate, λ) of GRB2 SH2 following EGF and PV treatment. E) TIRF imaging of total GRB2 SH2-tdEOS fluorescence following PV treatment for two separate cells. F) New molecule recruitment rate (apparent on-rate) curves for two separate PV-treated cells. G) Plot of GRB2 membrane fluorescence kinetics (TIRF imaging, see E) and GRB2 FW (n=2)

total binding in EGF-treated cells (*red-in vivo, black-FW*) and PV-treated cells (*magenta-in vivo, blue-FW*). FW data is normalized so that the maximum signal in EGF treated cells equals 1.

Discussion

In this study we employed three methods to monitor EGF-induced tyrosine phosphorylation and SH2 domain binding in the EGFR overexpressing A431 cell line: far-Western blotting, iTRAQ MS, and *in vivo* imaging of plasma membrane SH2 domain recruitment. These methods proved to be complementary and far from interchangeable, as response kinetics showed surprising differences. Nevertheless, by combining and comparing data from each technique we were able gain unique insights into the relationship between protein tyrosine phosphorylation, creation of SH2 domain binding sites, and membrane recruitment of SH2 domains in response to EGF, as well as into the capabilities of each method.

FW data provided an unprecedented global overview of changes in binding sites for 27 representative pY binding domains in response to EGF. A unique aspect of the FW approach is that it provides insight into the overall binding preferences of particular SH2 domains among all the tyrosine phosphorylated proteins in a cell lysate. Our results are generally consistent with previously reported SH2 domain-phosphoprotein interactions. On the whole, however, EGFR binding dominated SH2 probe binding patterns in our experiments. For example, probes such as CRK and CRKL, which are generally considered specific for focal adhesion proteins, displayed significant EGFR binding. This is likely due in part to the high EGFR expression level in A431 cells. When the CRK SH2 domain was used to probe far-Western blots of EGF-stimulated COS1 cells, which express more typical amounts of EGFR, strong p130Cas binding was seen in the absence of significant binding to EGFR (Figure 1.14 F). This suggests that the concentration of SH2 domain binding sites can be as important as SH2 domain binding specificity for determining the output of tyrosine kinase signaling pathways, and is consistent with the idea that signaling pathways can be “re-wired” by dysregulated RTK activity (*131*).

By temporally resolving both short and long stimulation time points, we were also able to show that the time course of phosphorylation can differ when specific EGFR sites are compared. This

was demonstrated by EGFR phosphosite-specific immunoblotting, and was reflected in the differences observed in SH2 binding site kinetics by far-Western blotting. These results suggest that the stoichiometry of SH2 and PTB domain-containing effectors bound to EGFR is temporally regulated by differential phosphorylation/dephosphorylation of their specific binding sites following stimulation with EGF. However, these differences were generally modest, and as such, evaluation of their importance for EGFR signaling will require a more in-depth analysis. Nevertheless, for multiply phosphorylated proteins, such as EGFR, far-Western likely provides a more complete picture of SH2-mediated signaling than phosphosite-specific Western blotting alone, because many SH2 domains display at least some affinity for multiple phosphosites on these proteins, and the relative stoichiometry of phosphorylation of different sites is variable and is challenging to quantify (*43*, *111*).

Using iTRAQ MS we were able to show that the phosphorylation state of over 65% of observed pY peptides was significantly altered following EGF stimulation, a much higher percentage than detected in previous studies using similar significance thresholds in cell lines with more moderate EGFR expression. Many of these sites are not listed as EGF-dependent within the PhosphoSitePlus database. Again, these data suggest that EGFR overexpression is associated with a significant expansion of its classical downstream signal transduction pathways. However, a more systematic analysis is needed to validate and quantify the effects of receptor expression level on EGFR signalosome plasticity. iTRAQ MS was also able to identify functional differences in protein tyrosine phosphorylation. GO analysis revealed significant differences between EGF-dependent and -independent phosphoproteins. As expected, proteins with EGF-dependent phosphosites were associated with ontologies related to activation of pY-dependent growth factor signaling. On the other hand, proteins containing non-affected sites tended to be associated with ontologies related to integrin signaling and cellular adhesion. Phosphosites in

these unaffected proteins were also more likely to contain CRK or NCK binding sites, consistent with their role in cell adhesion signaling.

The relatively fine temporal resolution of our analyses allowed us to capture small dynamic changes in protein tyrosine phosphorylation that were likely missed by previous studies using broader time intervals. Of particular note is the oscillation in the tyrosine phosphorylation levels soon after stimulation with EGF. The phenomenon was reproducibly captured for EGFR and to a lesser extent GAB1 by anti-pY and far-Western blotting. The observation of a similar dip in most MS-detected phosphosites suggests that it is a general feature of EGF-induced tyrosine phosphorylation and not a methodological artifact. While we can only speculate about the specific mechanism, it is likely due to feedback mechanisms that function rapidly and are directly linked to EGFR kinase activity (e.g. recruitment of SH2-containing phosphatases (Ferrell and Ha, 2014)). Moreover, such phosphorylation-dependent phosphatase-mediated negative feedback could also explain why the phosphorylation of SH2 domain-containing proteins, such as SHCA, reached equilibrium long before their recruitment to the membrane reached a maximum. EGF-dependent recruitment or activation of phosphatases might serve to counter the effect of the continued increases in binding between SH2-containing proteins and EGFR, as detected by TIRF imaging.

We also observed a number of inconsistencies between MS and far-Western in phosphorylation kinetics. This was particularly true for the focal adhesion scaffold p130CAS, where individual phosphosites did not show the decline in abundance after EGF treatment seen in FW (and anti-pY) blots. Similarly, we also observed noticeable differences in the rate of phosphorylation of sites on EGFR when compared to matched anti-pY Westerns. How can different analytic methods give such different results for the same samples? It is important to realize that both FW and anti-pY blots integrate the total signal for all phosphosites on a given protein, which can vary widely both in their affinity for different SH2 domains and in their stoichiometry of

phosphorylation. Discrepancies can arise when the limited subset of phosphosites detected by MS is not phosphorylated and/or dephosphorylated with the same kinetics or to the same level as the totality of sites that contribute to Western or far-Western signals. This is more likely to be an issue for proteins with many potential phosphorylation sites, such as EGFR and p130CAS (Table 1.2, Figure 1.17 A and B). More directed analyses of specific multiply phosphorylated proteins will be needed to resolve this question definitively. Nevertheless, these findings highlight the need to use multiple orthogonal methods to gain a full and accurate picture of the molecular events downstream of RTK activation.

Perhaps the most notable and surprising finding from this study was the difference between the kinetics of SH2 binding site creation, as measured by FW, and *in vivo* membrane recruitment, as measured by live cell imaging. The average time required for membrane binding to reach a maximum following EGF stimulation was nearly six times longer than expected based on the abundance of binding sites, indicating that SH2 membrane recruitment is not in rapid equilibrium with tyrosine phosphorylation. The extent of the apparent lag in *in vivo* binding correlated to some extent with SH2 binding specificity and the diffusion rate of bound SH2 domains—in general, the lag was shorter for SH2 domains that bound mostly to GAB1. Furthermore, the extent of the lag was also correlated with the apparent membrane off-rate, in that the SH2 domains with the longest membrane dwell times were also typically those that took the longest to reach maximal binding (Figure 1.14 D). We recently reported that the apparent membrane off-rate is related to the clustering of SH2 binding sites, as clustering favors phosphosite hopping or rebinding at the membrane (*94*). The results presented here suggest that clustering plays a role in the apparent delay in reaching maximal SH2 binding *in vivo* and that this delay may be a natural feature of the interaction between SH2 domain-containing proteins and the membrane. Consistent with this hypothesis, the time-course of GRB2 SH2 binding site clustering in response to EGF was remarkably similar to the time-course of membrane

recruitment (Figures 1.13 B and 1.18 C). Moreover, when GRB2 binding sites were created without inducing clustering (by PV treatment), the apparent delay was no longer seen, and the rates of GRB2 binding site formation and membrane recruitment were nearly equal (Figure 1.19 G).

We were able to directly measure the effect of rebinding by using sptPALM to quantify the recruitment of new domains to the membrane (equivalent to the effective membrane on-rate). The new molecule membrane binding rate (i.e. the apparent on-rate) reached a maximum much more quickly than did total membrane binding. In particular, the kinetics of new molecule binding to unclustered sites were very similar to the kinetics of binding site generation measured by far-Western (Figure 1.18 D, Table 1.2). This strongly suggests that the apparent delay in total membrane binding observed under TIRF microscopy is the result of changes in the effective off-rate, and not an issue of binding site availability.

This novel observation of a lag in reaching maximal recruitment to phosphorylated sites *in vivo* raises the question of how such a delay might impact signaling. Signal outputs are subject to multiple positive and negative feedback loops, making it difficult to assess the specific role of such a phase delay. For example, our results show that ERK activation (assayed by ERK1/2 phosphorylation) reaches a maximum at ~4 min, between the time of maximal GRB2 binding sites (1-2 min) and maximal recruitment of GRB2 to the membrane (~10 min). It is possible that the sustained, clustering-dependent increase in SH2 binding might be involved in regulating the duration of ERK activation. Alternatively, it may be more important in the response to lower (more physiological) levels of receptor stimulation than those used for our experiments. Another interesting possibility is that clustering-dependent recruitment is a mechanism to diversify the response kinetics upon receptor stimulation. For example, recruitment of effectors to GAB1, a relatively freely diffusing scaffold protein, reaches a maximum much more rapidly than recruitment of effectors to EGFR itself. Phosphorylation of scaffolding and adaptor molecules

such as GAB1, SHCA, and others may play a general role in increasing the spatiotemporal diversity of receptor output (*75*, *133*).

Taken together, the data presented here provide a strong rationale for a multimodal approach to the analysis of PTM-mediated signaling. Such an approach allows the identification of areas of discrepancy, which can then be explored by more directed experimentation. This is analogous to the interplay between *in vivo* experimentation and computational modeling, where new insight is driven by discrepancies between model predictions and experimental results (*134*, *135*). In the process of reconciling seemingly discordant results obtained by interrogating the same system with multiple methodologies, one can gain novel insights into the system not possible from studies dependent on a single methodology.

Src Homology 2 domains enhance tyrosine phosphorylation in vivo by protecting their binding sites from dephosphorylation

Joshua A. Jadwin[1], Timothy G. Curran[2], Forest M. White[2], Bruce J. Mayer[1]

Affiliations

1. Raymond and Beverly Sackler Laboratory of Molecular Medicine, Department of Genetics and Genome Sciences, University of Connecticut School of Medicine, 263 Farmington Avenue, Farmington, CT 06030

2. Department of Biological Engineering and Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139

Abstract:

Phosphotyrosine (pY) dependent signaling is a highly dynamic process. Signal output depends not only on rates of phosphorylation and dephosphorylation, but also on rates of binding and dissociation of effectors containing phosphotyrosine-dependent binding modules such as Src Homology 2 (SH2) and phosphotyrosine binding (PTB) domains. Previous *in vitro* studies suggested binding of SH2 and PTB domains can enhance protein phosphorylation by protecting bound sites from phosphatase-mediated dephosphorylation. To test whether this occurs *in vivo*, we quantified the effects of SH2 domain overexpression on protein tyrosine phosphorylation by quantitative Western and far-Western blotting, mass spectrometry (MS), and computational modeling, using the binding of GRB2 (growth-factor receptor bound 2) to phosphorylated Epidermal Growth Factor Receptor (EGFR) as a model system. We found that SH2 overexpression resulted in a significant, dose-dependent increase in EGFR tyrosine phosphorylation, and the phosphosites enhanced corresponded with the binding specificity of the overexpressed SH2 domain. Computational modeling using experimentally determined EGFR phosphorylation rates, and pY-EGFR and GRB2 concentrations recapitulated experimental findings. Surprisingly, both modeling and biochemical analysis suggested that SH2 domain overexpression did not result in a major decrease in the number of unbound phosphorylated SH2 domain binding sites. Our results suggest that signaling via SH2 domain binding is buffered over a relatively wide range of effector concentrations, and that SH2 domain proteins with overlapping binding specificities are unlikely to compete with one another for phosphosites *in vivo*.

Introduction:

SH2 (Src-Homology 2) domains are small modular protein domains that bind specifically to tyrosine phosphorylated sites on proteins (*32*, *34*). In cell signaling, proteins that contain SH2 domains function to “read” post-translational marks that are “written” by activated tyrosine kinases and are “erased” by protein tyrosine phosphatases (PTPs) (*136*). Protein complexes mediated by SH2-phosphotyrosine (pY) interactions are critical for downstream signaling from a number important tyrosine phosphorylated proteins, including activated receptor tyrosine kinases (RTKs) such as the Epidermal Growth Factor Receptor (EGFR) and the proteins phosphorylated by RTKs as well as by non-receptor cytoplasmic tyrosine kinases, such as Abl and Src; (*107*, *108*).

In total, 120 unique SH2 domains have been identified in 110 proteins with varying functions including adaptors, tyrosine and serine/threonine kinases, tyrosine phosphatases, and lipid phosphatases (*107*). SH2 binding specificity and affinity is defined by the amino acid sequence flanking their pY binding site. In particular, amino acids at positions -1 through +3 relative to the phosphotyrosine have been shown to have the greatest influence. While the specificity of each SH2 domain is unique, most SH2s bind to one of a few general motifs; +1 D/E, +2 N, +3 P/L/V, and +1M +3M (*35*, *36*, *110*). SH2 domain phosphosite motif specificities and affinities have been elaborated using *in vitro* methods such as surface plasmon resonance, solution assays and protein microarray, and using pull-down based approaches (*43*, *89*, *137*, *138*). In addition to SH2 domains, there are a relatively small number of other phosphotyrosine-specific binding modules which include several PTB (phosphotyrosine binding) domains (*139*, *140*).

SH2 binding *in vivo* is highly dynamic (*49*, *141*). Not only do SH2 domains bind and unbind from phosphosites rapidly at rates determined by their binding constants, but phosphosites themselves turn over rapidly, with halftimes in the range of seconds (*49*). The rate of

phosphosite turnover is dependent on both kinase and phosphatase activity (*142*). Furthermore, multiple SH2 domains are expressed and can compete for binding to phosphosites. Thus understanding SH2-mediated signal output requires consideration of phosphotyrosine flux and local concentrations of SH2-containing proteins, in addition to binding site specificity.

To study the interplay between SH2 domain binding and phosphosite dynamics we have exploited EGFR, a major docking site for multiple SH2 domain containing proteins. EGFR kinase activity increases when its ligand, EGF (Epidermal Growth Factor), binds to the extracellular domain of EGFR inducing structural changes that promote receptor dimerization and kinase activation (*142*). As a result, cellular levels of EGFR kinase activity can easily be manipulated by varying ligand concentration. Moreover, unlike most kinases, EGFR kinase does not depend on tyrosine phosphorylation of the so called activation loop. This is important because the effects of SH2 expression on receptor phosphorylation can be assessed independent from phosphorylation-associated receptor activation (*48*, *59*, *61*, *143*, *144*). Activated dimerized receptors phosphorylate the C-terminal tyrosine residues that serve as binding sites for a set of SH2 domain-containing proteins including GRB2, SHCA, PLCγ1 and SHP2 (*43*, *66*, *145*). Each SH2 domain is thought to bind preferentially to a specific individual phosphosite or subset of phosphosites based on its individual binding specificity. For example, GRB2 has been shown bind to pYXN motifs (where X can be any amino acid) at EGFR pY1068, pY1086 and pY1114, while the SHCA PTB domain has been shown to bind strongly to pY1148, a NPDpY motif (*35*, *124*, *146*, *147*).

Previous studies suggested that SH2 domains could specifically prevent dephosphorylation of their phosphosite binding partners *in vitro* (*148–151*). However little is known about the impact in living cells, where phosphosite turnover is high and overall occupancy may be low. Here we use EGFR, as well as constructs containing the GRB2 and CRK (v-Crk avian sarcoma virus CT10 oncogene homolog) SH2 domains, to investigate the interplay between SH2 domain

binding and phosphosite dynamics *in vivo*, focusing specifically on a quantitative analysis of phosphosite protection. By employing biochemical analyses, phosphotyrosine-specific mass spectrometry (MS) and computation modeling we demonstrate that GRB2 can enhance the steady state tyrosine phosphorylation level of its binding sites *in vivo* through SH2-dependent protection from PTPs. Our results also suggest that SH2 protection has important implications for our understanding of binding site competition between SH2 domains with similar specificities. Furthermore, SH2-mediated pY protection might serve as the basis for a novel method for identifying SH2-pY interactions as they occur *in vivo*.

Results:

GRB2 SH2 domain overexpression enhances EGFR phosphorylation

To assess the effect of SH2 protein-phosphosite interaction on tyrosine phosphorylation we transiently overexpressed full-length wild type (wt) GRB2 in COS1 cells and monitored cellular tyrosine phosphorylation before and after stimulation with EGF (2.5ng/mL for 5 min) by anti-pY Western (Figure 2.1 A and 2.1 B *lanes 1-4*). GRB2 overexpression enhanced total EGFR phosphorylation before and after EGF treatment. Relative increases in pY-EGFR ranged between 1.5 and 4 fold and were more pronounced following EGF treatment.

GRB2 mediates signaling through a complex series of downstream pathways and it is possible that these pathways could lead indirectly to enhanced EGFR phosphorylation, for example by increasing cytoplasmic kinase activity or suppressing phosphatase activity (*152–155*). To rule out downstream signaling as a driver of GRB2-mediated EGFR phospho-enhancement, we compared protein tyrosine phosphorylation after expression of wt GRB2 with four GRB2 derived constructs: a fluorescently tagged GRB2 SH2 domain (tdEOS-GRB2 SH2), previously shown to be recruited to the plasma membrane of EGF stimulated cells; full length GRB2 and tdEOS-GRB2 SH2 constructs containing a mutation in the SH2 domain (R86K), previously shown to

abrogate phospho-dependent interaction; and a chimeric protein in which the SH2 domain of GRB2 is replaced by that from CRK, referred to here as GCG (Figure 2.1 A and 2.1 B *lanes 5-12*) (*94*, *151*, *156*). Unlike the GRB2 SH2, which binds predominantly to the activated EGFR, the CRK SH2 domain binds predominantly to phosphorylated p130CAS, and only weakly to phosphorylated EGFR (*157*).

As expected, SH2 constructs carrying the R86K mutant failed to increase EGFR phosphorylation. The tdEOS-GRB2 SH2 fusion, like full length GRB2, increased EGFR phosphorylation before and after EGF stimulation. Compared to the full length protein however, transient expression of this construct resulted in a more pronounced increase in EGFR phosphorylation, especially prior to EGF treatment. By contrast, overexpression of the CRK-GRB2 chimera resulted in a significant increase in p130CAS phosphorylation with only a minor increase (less than two fold) in the phosphorylation of EGFR, consistent with its binding specificity for p130CAS phosphotyrosine sites (*38*, *65*, *127*, *157*).

These results showed that overexpession of SH2 domains could specifically increase the tyrosine phosphorylation of their known cellular binding partners. To more directly address the SH2 binding specificity of the enhanced phosphosites, we performed far-Western blotting on lysates from COS1 cells expressing the various GRB2 constructs and probed with recombinant GRB2 and CRK SH2 domains (Figure 2.1 C). Expression of constructs containing the GRB2 SH2 domain specifically increased GRB2 SH2 probe binding, particularly of a band corresponding to EGFR. By contrast, the CRK SH2 probe bound predominantly to the p130CAS band in GCG-expressing cells.

GRB2 mediated pY-EGFR enhancement depends on GRB2 concentration and EGFR kinase activity

Next we assessed the dose dependence of the enhancement of EGFR phosphorylation by GRB2 expression. Increasing amounts of cDNA encoding full length GRB2 were transfected in COS1 cells and the average GRB2 concentration per cell was then calculated for each dose using a recombinant GST-GRB2 SH2 standard (Figure 2.1 D). As expected, EGFR phosphorylation increased with GRB2 concentration. As in Figure 2.1 B and C, the effect of GRB2 overexpression was not as dramatic in unstimulated serum starved cells (an approximately 2-fold increase)

To explore the relationship between phosphosite flux and SH2-mediated phosphosite protection, we modulated EGFR kinase activity in SH2-overexpressing COS1 cells and monitored EGFR phosphorylation. Cells were transfected with tdEOS-GRB2 SH2 or the tdEOS-GRB2 SH2 R86K mutant and then treated with increasing concentrations of EGF (Figure 2.1 E). Relative EGF-induced increases in phosphorylation were fairly constant between 0 and 2.5 ng/mL EGF (5-7 fold), but fell off at higher EGF concentrations and eventually plateaued. The absolute increase in phosphorylation (difference between mutant and wild type GRB2 pY-EGFR signal at a particular EGF concentration) peaked at 2.5 ng/mL EGF, and dropped off at both higher and lower concentrations. Quantification using a phosphospecific antibody for EGFR pY1068, an established GRB2 binding site, revealed similar results (Figure 2.1 E).

Specificity of SH2-dependent phosphotyrosine enhancement

If SH2 domains enhance phosphorylation by protecting their binding sites from dephosphorylation, then protected sites should be enriched for high-affinity sites containing

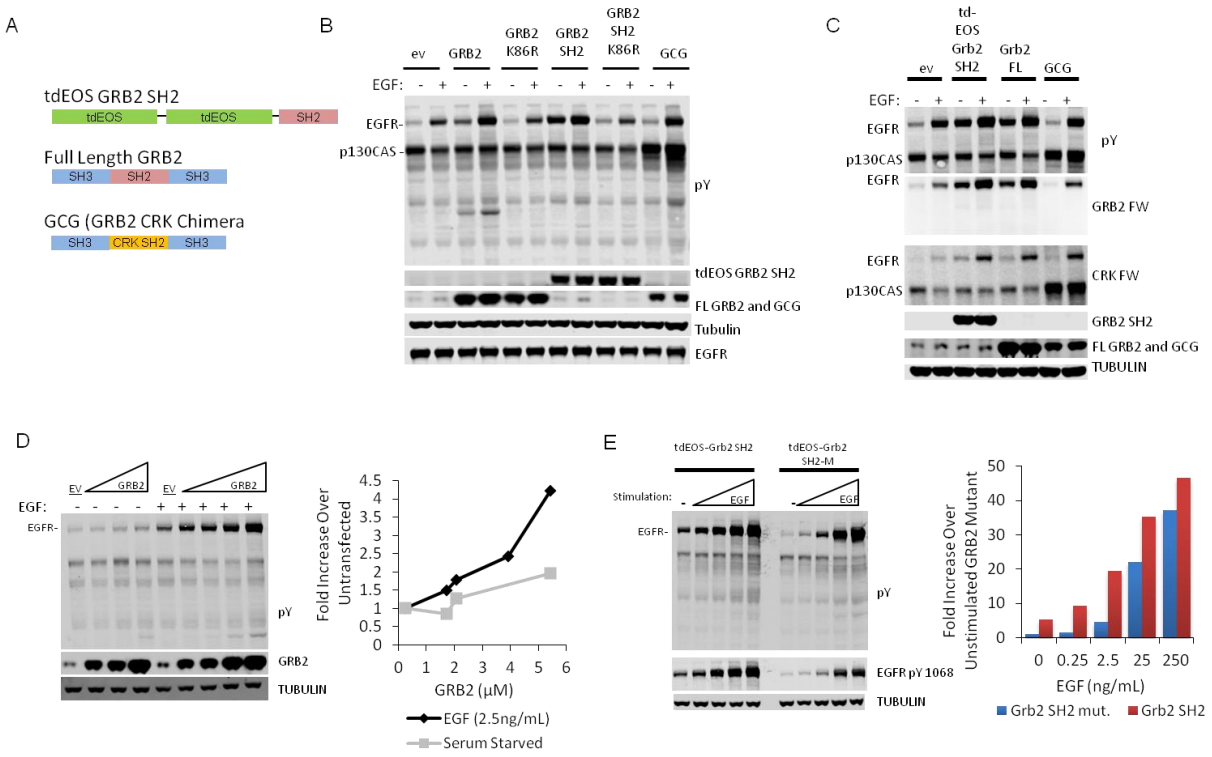

Figure 2.1: Effect of GRB2 expression in EGF-stimulated COS1 cells. A) Diagram of major constructs used for this study: tdEOS tagged GRB2 SH2, full length wt GRB2 and a chimera of GRB2 SH3 domains and the CRK SH2 domain (GCG). B) GRB2-mediated enhancement of EGFR phosphorylation is SH2-dependent. Western blotting of lysates from COS1 cells overexpressing GRB2, tdEOS-GRB2 SH2 and GCG constructs before and after stimulation with 2.5 ng/ml EGF. R86K constructs are mutants that are unable to bind phosphotyrosine sites. C) Far-Western blotting and immunoblotting of lysates from COS1 cells transfected with GRB2 constructs. In labels on right, “CRK FW” and “GRB2 FW” indicate far-Western blotting with CRK and GRB2 SH2 domains. D) Western blot showing the effect of GRB2 expression level on EGFR tyrosine phosphorylation. GRB2 expression levels were determined using a GRB2-SH2 standard and a GRB2-SH2-specific antibody. Densitometric quantification of this data is plotted on the right. E) Western blot showing the effect of EGF stimulation on GRB2-mediated

enhancement of total EGFR pY and EGFR pY1068 (a GRB2 SH2 binding site). “GRB2 SH2M” is the R86K mutant. Densitometric quantification of this data is shown to the right.

phosphorylation for seven EGFR phosphosites. In control (empty vector) cells, pY1148 dominated EGFR phosphorylation, a finding consistent with recent quantitative MS analyses (*111*). Cells expressing each of the three constructs displayed increased EGFR phosphorylation at all of the measured sites to varying degrees. Relative to the overall increase in phosphorylation of EGFR, as measured by anti-pY immunoblot, both wt GRB2 and tdEOS-GRB2 SH2 expression resulted in increased phosphorylation of the canonical GRB2 binding site pY1068 (EpYINQ), but had an equal or even greater effect on pY974 (FpYRAL), pY992 (EpYLIP), and pY1173 (EpYLRV). Nevertheless, GRB2 overexpression did result in a significant change in the absolute phosphorylation landscape of EGFR in transfected cells, with pY1068 phosphorylation nearly equaling that of pY1148 (Figure 2.2 A, *adjacent bar graphs*). As in previous experiments, GCG increased total EGFR total phosphorylation by less than two fold. In these cells, phosphorylation was significantly shifted in favor of pY992 (EpYLIP), a canonical CRK binding site; which was increased more than 10 fold when compared to EGF-treated empty vector controls (Figure 2.2 A, *adjacent bar graphs*).

canonical motifs for the overexpressed SH2 domains (e.g. pYXN motifs for GRB2). To test this, we transfected three constructs (tdEOS-GRB2 SH2, full length GRB2, and GCG) in COS1 cells and monitored the absolute level of phosphorylation at each EGFR site (with and without treatment with 2.5 ng/ml EGF) using phosphosite-specific antibodies (Figure 2.2 A). To compare signal levels across multiple antibodies we created a maximally phosphorylated pY-EGFR standard by treating COS1 cells with 200 ng/mL EGF and the phosphatase inhibitor pervanadate for 40 min (Figure 2.2 A, *far right lane*). Using this method we quantified the percent

Quantitative phosphotyrosine MS analysis of COS1 cells expressing the same four constructs revealed similar results (Figure 2.2 B and C, Table 2.1). Overexpression of GRB2 SH2 and full length GRB2, and to a much lesser extent the GCG construct, resulted in a generalized increase in the relative abundance of the four EGFR phosphosites identified (pY974, 1086, 1148 and 1173) both before and after EGF stimulation (Figure 2.2 B). Unlike pY-EGFR immunoblotting, MS data suggested that tyrosine phosphorylation of EGFR pY974 was only mildly increased by exogenous GRB2 expression. All five of the p130CAS phosphopeptides identified contained the canonical CRK SH2 binding motif, pYDXP. Of these, phosphorylation of four was enhanced by GCG chimera expression (Figure 2.2 C).

The specificity of GRB2 SH2-mediated phosphosite enhancement is concentration dependent

While Western, FW and MS data were broadly consistent with the hypothesis that SH2 domains protect their preferred binding sites from dephosphorylation by phosphatases *in vivo*, the specificity was rather modest compared with that seen using peptide-based *in vitro* interaction assays. However, most of our experiments were performed using high intracellular

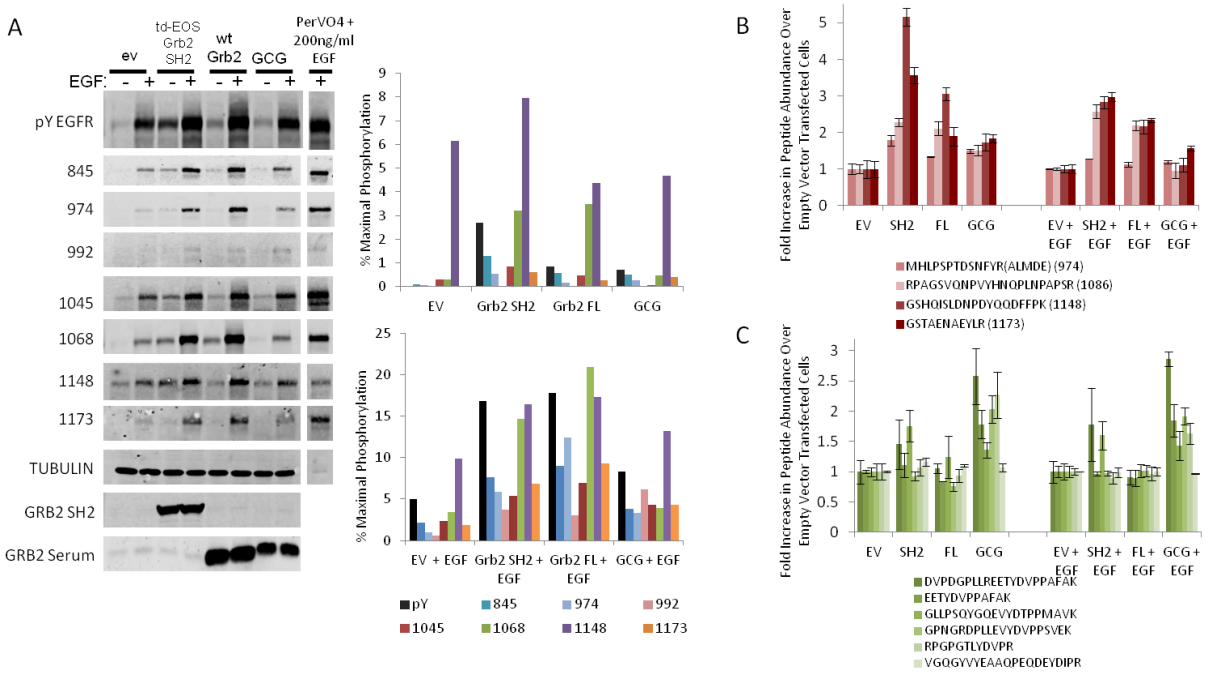

Figure 2.2: Site-specific increases in EGFR phosphotyrosine in SH2-overexpressing cells. A) Anti-pY and phosphosite-specific anti-pY-EGFR Western blots from COS1 cells transfected with empty vector, tdEOS-GRB2 SH2, FL GRB2 SH2, or GCG. EGF+Pervanadate (200 ng/mL EGF, 100 μM pervanadate, 40 min) was used as a maximally phosphorylated standard and run with the site-specific blots at a 1:10 dilution on the same membranes. Antibodies are indicated to the left; for phosphospecific antibodies, numbers indicate phosphosite recognized. The percent maximal phosphorylation for each site and total pY-EGFR without (top) and with (bottom) EGF stimulation is shown in the bar graphs to the right. B) The relative increase in abundance of EGFR phosphopeptides detected by quantitative phosphotyrosine specific mass spectrometry in COS1 lysates from cells expressing the SH2 constructs. C) The relative increase in abundance of p130CAS phosphopeptides detected as in B. Error bars for B and C represent standard error of the mean (SEM).

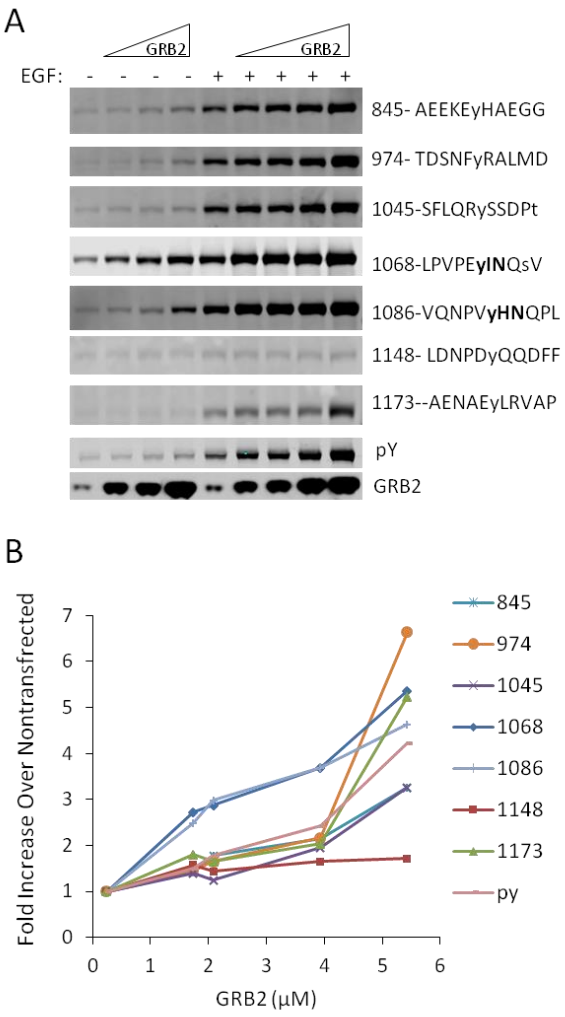

Figure 2.3: GRB2 specifically enhances its canonical binding motifs in a concentration-dependent manner. A) EGFR pY-specific Westerns from COS1 cells expressing an increasing amount of GRB2. EGFR pY amino acid position and pY motifs are indicated on the right. B) Quantification of EGFR pY site-specific phosphorylation following EGF stimulation plotted against overexpressed GRB2 concentration.

concentrations (approximately 5.5 μM) of SH2-containing constructs in order to maximize the effect of phosphosite protection. At such high concentrations, SH2 domains likely interact with and protect relatively low affinity sites. To better understand the effect of concentration on specificity in our system, we analyzed lysates of cells expressing a range of GRB2 concentrations by probing with EGFR phosphosite-specific antibodies (Figure 2.3). We found that the known high-affinity GRB2 binding sites, pY1068 and pY1086, were selectively enhanced at relatively low levels of GRB2 overexpression. By contrast, sites predicted to bind the GRB2 SH2 with lower affinity, such as pY974, pY1173 and pY1045, were enhanced only at the highest GRB2 concentrations (Figure 2.3 B).

SH2 domain overexpression enhances phosphosites within canonical binding motifs across the phosphoproteome

To delve more deeply into the specificity of phosphosite protection associated with SH2 domain expression, we used our quantitative MS data to examine the sequence specificity of SH2-mediated phosphosite enhancement (*128*, *157*). In total, we identified 118 tyrosine-phosphorylated peptides from 79 different proteins (Table 2.1). MS analysis indicated that for almost half of these phosphopeptides, their abundance was significantly increased or decreased in cells overexpressing SH2 domains relative to controls (Figure 2.4 A). In particular, the phosphosites enhanced by expression of GRB2 or the GRB2 SH2 were enriched for the GRB2 SH2 binding motif (pYXN), while GCG-enhanced sites were enriched for the CRK SH2 binding site (pYXXP) (Figure 2.4 4B and C). Collectively, the enhanced motifs closely resembled those previously identified using *in vitro* binding assays (*see insets in* 4B *and* C) (*36*). This suggests that GRB2 and CRK (i.e. GCG) SH2-mediated phosphosite

See attached Excel files

Table 2.1: iTRAQ MS data and SH2-enhanced phosphosites. *MS Data* tab lists all phosphopeptides identified by MS. The protein name, accession number, peptide sequence, relative abundance, and standard deviation of all peptides identified are shown. Peptides phosphorylated at serine and/or threonine are listed, but were not included in our analysis. Mean abundance values are such that the sum for each peptide equals eight. *GCG ENHANCED* tab lists the proteins, sequence and fold change in abundance for peptides that displayed increased abundance in GCG-transfected cells (over empty vector-transfected) for both starved and EGF treatment conditions. *GRB2 ENHANCED* tab lists the protein, sequence and fold change in abundance for peptides enhanced in tdEOS-GRB2 SH2 and full length GRB2-transfected cells (over empty vector transfected) for both starved and EGF treatment conditions. Data from GCG *ENHANCED* and *GRB2 ENHANCED* were used to create the LOGOs in Figure 2.4 B and C.

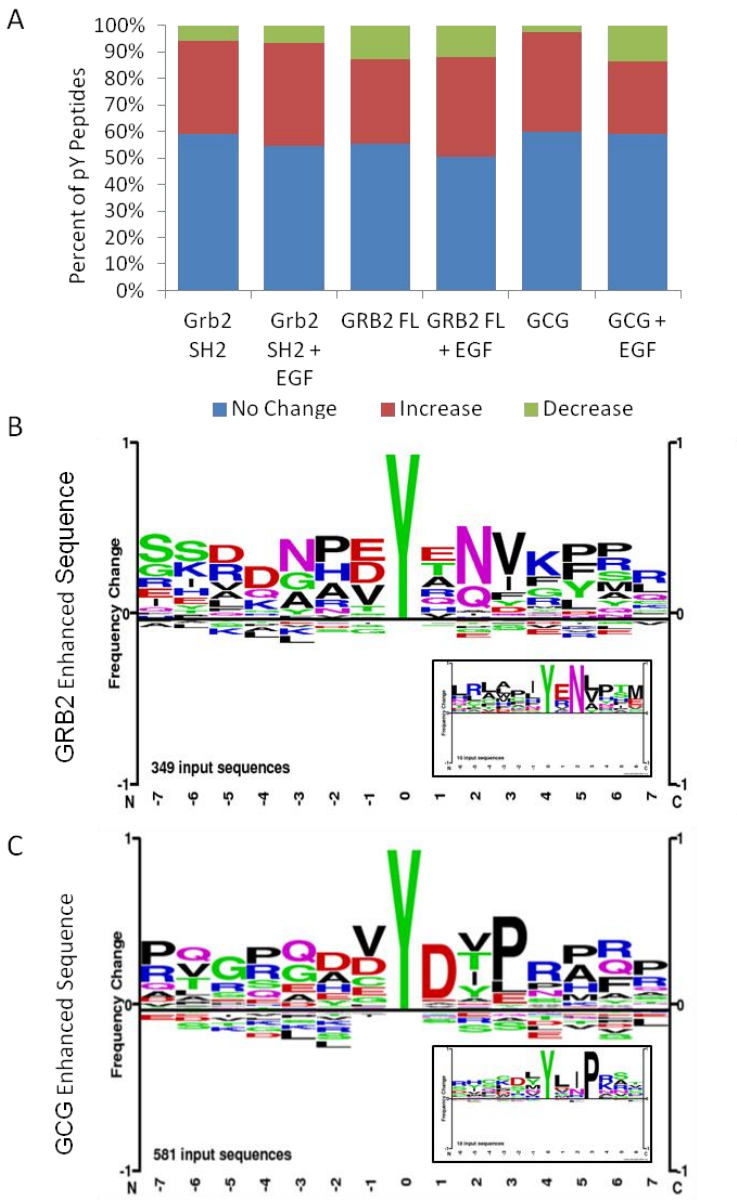

Figure 2.4: Quantitative MS-based analysis of tyrosine phosphoproteome in cells overexpressing SH2-containing proteins. A) Bar graph showing the percentage of phosphopeptides in each experimental treatment whose abundance was increased, decreased or unchanged. B) Peptide LOGO of phosphosites whose abundance was enhanced by GRB2 construct expression before and after EGF stimulation. Inset shows peptide LOGO for *in vitro* GRB2 binding phosphopeptides from by Tinti et al. C) Peptide LOGO of phosphosites whose

abundance was enhanced by GCG expression both before and after EGF stimulation. Inset shows peptide LOGO for *in vitro* CRK binding phosphopeptides from by Tinti et al.

enhancement was largely binding site specific when assessed across the entire phosphotyrosine proteome (*106*).

Computational model of SH2 phosphosite protection

These results strongly suggested that SH2 domains specifically prevent the dephosphorylation of their phosphorylated binding sites *in vivo* by shielding those sites from being accessed by phosphatases. To better understand the behavior of such a system we generated a quantitative computational reaction model of SH2-pY interactions in cells before and after EGF stimulation and compared model predictions with experimental data (Figure 2.5) (*106*).

To generate an accurate and realistic model, we experimentally determined as many parameters as possible in our COS1 cell system (Table 2.2). In particular, we quantified EGF-dependent and -independent phosphorylation and dephosphorylation rates at steady state in the absence or presence of EGF. Absolute rates were quantified by quantifying anti-pY immunoblots of lysates from cells treated with the tyrosine phosphatase inhibitor pervanadate before and after EGF stimulation, and from EGF-stimulated cells treated with the EGFR kinase inhibitor Erlotinib, using an absolute standard for phosphotyrosine developed in our lab (Figure 2.5 B-D) (*50*, *142*). The independently obtained measures of EGFR phosphorylation and dephosphorylation in EGF-treated cells were nearly identical as would be expected of a system at steady state. EGFR phosphorylation was modeled as a first order substrate-limited reaction, as we were unable to obtain reasonable V_{max} values by fitting to the Michaelis-Menton equation. EGFR dephosphorylation was modeled by fitting dephosphorylation data (i.e. Erolotinib treatment) to the Michaelis-Menton function. The phosphorylation rate constant (k_f) and dephosphorylation V_{max} and K_m values were calculated using the measured rates, and the percent pY-EGFR values obtained in Figure 2.2. Phosphorylation in unstimulated cells was

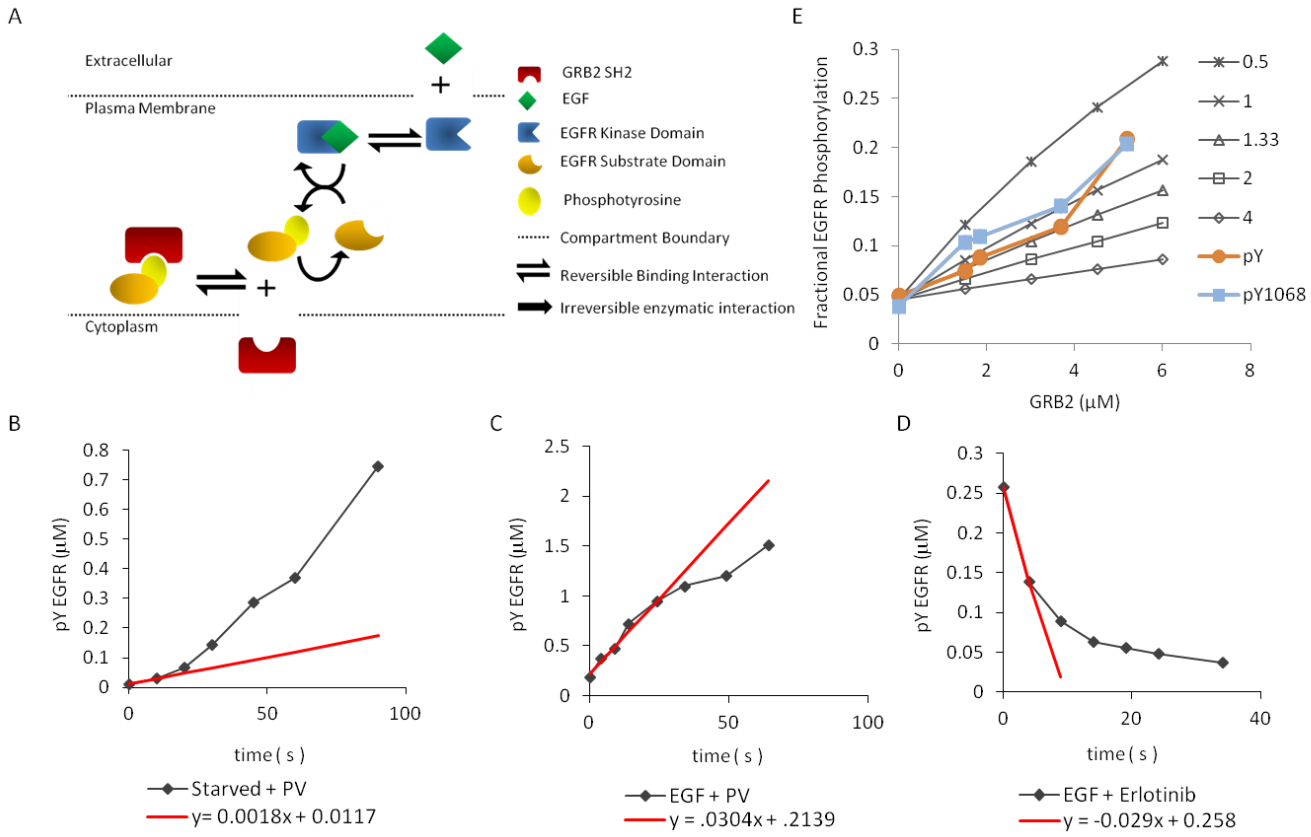

Figure 2.5: Computational model and parameter determination. A) Diagram of the computational model used to quantify the effect of SH2 expression on EGFR phosphorylation. The effect of SH2 domain expression in unstimulated cells utilized the same scheme, but used an EGFR k_f value obtained in unstimulated cells. B-D) Measurement of *in vivo* phosphorylation and dephosphorylation rates in COS1 cells. B) Plot of absolute phosphorylation rate in serum-starved cells treated with pervanadate. C) Plot of absolute phosphorylation rate in starved cells treated with pervanadate five minutes after stimulation with 2.5 ng/mL EGF. D) Plot of absolute dephosphorylation rate in cells treated with Erlotinib five minutes after stimulation with 2.5 ng/mL EGF. Black lines show amount of phosphorylation quantified from experimental data. Red lines show initial rate used to calculate model parameters. pY-EGFR concentrations were obtained by comparing anti pY Westerns of COS1 lysates with a recombinant v-Abl-derived phosphotyrosine standard. E) Plot of experimental data for fraction of total pY-EGFR (*orange*

circle) and EGFR pY1068 (*blue square*) overlaid on model predictions generated at varying k_{on} values. Plots were created by holding the k_{off} constant at 1 s^{-1} and varying k_{on} values (*black lines*).

Model Parameters		
Parameter	Description	Values
V_{cyto}	Cytoplasmic Volume	$2.59x10^{3}\ \mu m^{3}$
V_{mem}	Approximated Membrane Volume	$2.05\ x10^{2}\ \mu m^{3}$
$V_{extracell}$	Extracellular Volume	$4.0x10^{12}\ \mu m^{3}$
[EGF]	EGF concentration	$4.0x10^{-5}$ - $4\ x10^{-2}\ \mu M$
$k_{on\ EGF}$ *ψ	EGF-EGFR binding, forward rate constant	$63.0\ \mu M^{-1} * s^{-1}$
$k_{off\ EGF}$ *ψ	EGF-EGFR binding, reverse rate constant	$0.16\ s^{-1}$
$k_{f\ EGFR-US}$	Phosphorylation rate constant of unstimulated EGFR	$0.0036\ s^{-1} * \mu M^{-1}$
$k_{f\ EGF-EGFR}$	Phosphorylation rate constant of EGF bound EGFR	$0.54\ s^{-1} * \mu M^{-1}$
[EGFR] ψ	EGFR concentration	$0.81\ \mu M$
$V_{max\ PTP}$	Maximum EGFR dephosphorylation rate in unstimulated cells	$4.6\ \mu M \cdot s^{-1}$
$K_{m\ PTP}$	Km of EGFR dephosphorylation	$3.3\ \mu M$
$[GRB2]_{endo}$	Endogenous COS1 cell GRB2 concentration	$0.24\ \mu M$
$[GRB2]_{exo}$	Total GRB2 construct concentration in transfected cells	$1.7 – 5.4\ \mu M$
$k_{offGRB2-EGFR}$ Ω	GRB2-EGFR binding, off-rate constant	$1.0\ s^{-1}$
$k_{on\ GRB2-EGFR}$	GRB2-EGFR binding, on-rate constant	Experimental Fit
$K_{d\ GRB2-EGFR}$	GRB2-EGFR dissociation constant	Experimental Fit

Table 2.2: Model Parameters. All parameters were defined using experimental data unless otherwise noted below. Cell volumes were approximated from measurements of trypsinized cells. EGF binding constants were taken from the literature. COS1 cell phosphorylation and dephosphorylation rates were measured by comparing changes in phosphorylation under three conditions with a v-Abl-derived phosphotyrosine standard (see Figure 2.5 B-D). Measurements of *in vivo* EGFR phosphorylation rates and approximations of EGFR expression in COS1 cells were used to calculate the EGFR k_f values. Phosphatase V_{max} and K_m values were obtained by fitting the quantitative EGFR dephosphorylation data to the Michaelis-Menton function. GRB2 concentrations were calculated via Western blotting using a purified GRB2 standard run on the same membrane. GRB2 binding constants were determined using data from previously published work and by fitting experimental data to the model (See Figure 2.5 E). *= Berkers 1991, ψ=French 1995, Ω= Oh 2012.

modeled using the same reaction kinetics scheme, and the assumption that all EGFR kinases present in the cell had the same basal activity and contributed to substrate domain phosphorylation. EGFR and EGFR substrate concentrations were estimated using published estimates of EGFR expression in A431 cells and anti-EGFR Westerns comparing EGFR expression in COS1 and A431 cells. EGF concentrations are those used experimentally. EGF-EGFR binding parameters were previously published and used to define the percent of active EGFR kinases (*99*, *100*, *158*).

The GRB2 SH2-pY-EGFR dissociation constant (k_{off}) was obtained from recently published *in vivo* measurements (*94*). Binding (k_{on}) and dissociation (K_d) constant values were then approximated by fitting our experimental GRB2 titration data (see Figure 2.1 D) for both total pY-EGFR and pY1068 to modeling data generated using different k_{on} values (Figure 2.5 E). The resulting K_d values, approximately 1.0 and 0.66 μM respectively, are near the average of measurements for GRB2 SH2-pY interaction generated from solid phase assay (~0.2-0.7 μM) and solution methods (2.6 μM) (*43*, *89*, *159*, *160*).

Comparison of modeling and experimental data

Using the GRB2-pY-EGFR K_d value of 1 μM, we then modeled the effect of increasing concentrations of EGF on phosphosite abundance (Figure 2.6 A). The size of the effect predicted by the model closely matched experimental results, with increasing EGFR kinase activity significantly suppressing the relative effect of SH2 domain-mediated protection.

We then tested the effect of increasing GRB2 concentrations on EGFR phosphorylation using our computational model. Maximal percent phosphorylation was similar to that found experimentally; EV+EGF at ~5.0% and GRB2 SH2+EGF at ~20% (Figure 2.6 B and 2.2 A). Closer examination of the data revealed that GRB2 SH2-mediated increases in total

phosphorylation did not greatly affect the number of unbound sites within the system. Even at GRB2 concentrations twenty times that of the endogenous protein, resulting in a greater than four-fold increase in total phosphorylation, the number of unbound sites decreased only by approximately 25% (Figure 2.6 B). This minimal reduction in free phosphosites suggested that protection might minimize the effect of competition between SH2 domains with similar substrate specificities.

To more directly assess this hypothesis we added a GRB2 competitor SH2 to our model, referred to here as SH2competitor. We set the binding affinity of SH2competitor for phosphorylated EGFR equal to that of GRB2, varied its concentration, and monitored the amount of GRB2 bound to EGFR (Figure 2.6 C). The model predicted that a 60-fold increase in the amount of SH2competitor resulted in only a 22% decrease in the amount of GRB2 bound to EGFR. We surmised that the surprisingly modest effect of excess competitor depended on the turnover of phosphosites, which in turn drove protection of SH2-bound sites from dephosphorylation. For comparison, we modeled GRB2-EGFR binding in a system that contained the same initial amount of EGFR pY, but did not allow for pY turnover mediated by kinase and PTP activity (Figure 2.6 C). In this system, the addition of a similar amount of SH2competitor resulted in a nearly 84% decrease in the amount of GRB2 bound to EGFR.

To test effect of SH2 competition on downstream signaling *in vivo*, we assessed the phosphorylation of ERK1 and ERK2, well established effectors of EGFR activation, in cells transfected with tdEOS-GRB2 SH2 (*66*). Overexpression of tdEOS-GRB2 SH2, which lacks the GRB2 SH3 domains and cannot itself contribute to downstream signaling, resulted in an approximately 25% reduction in the phosphorylation of ERK1 and ERK2 (Figure 2.6 D). These data are consistent with our model predictions, as well as with a recent report showing that overexpression of a GFP-tagged GRB2 construct with an artificially enhanced phosphosite

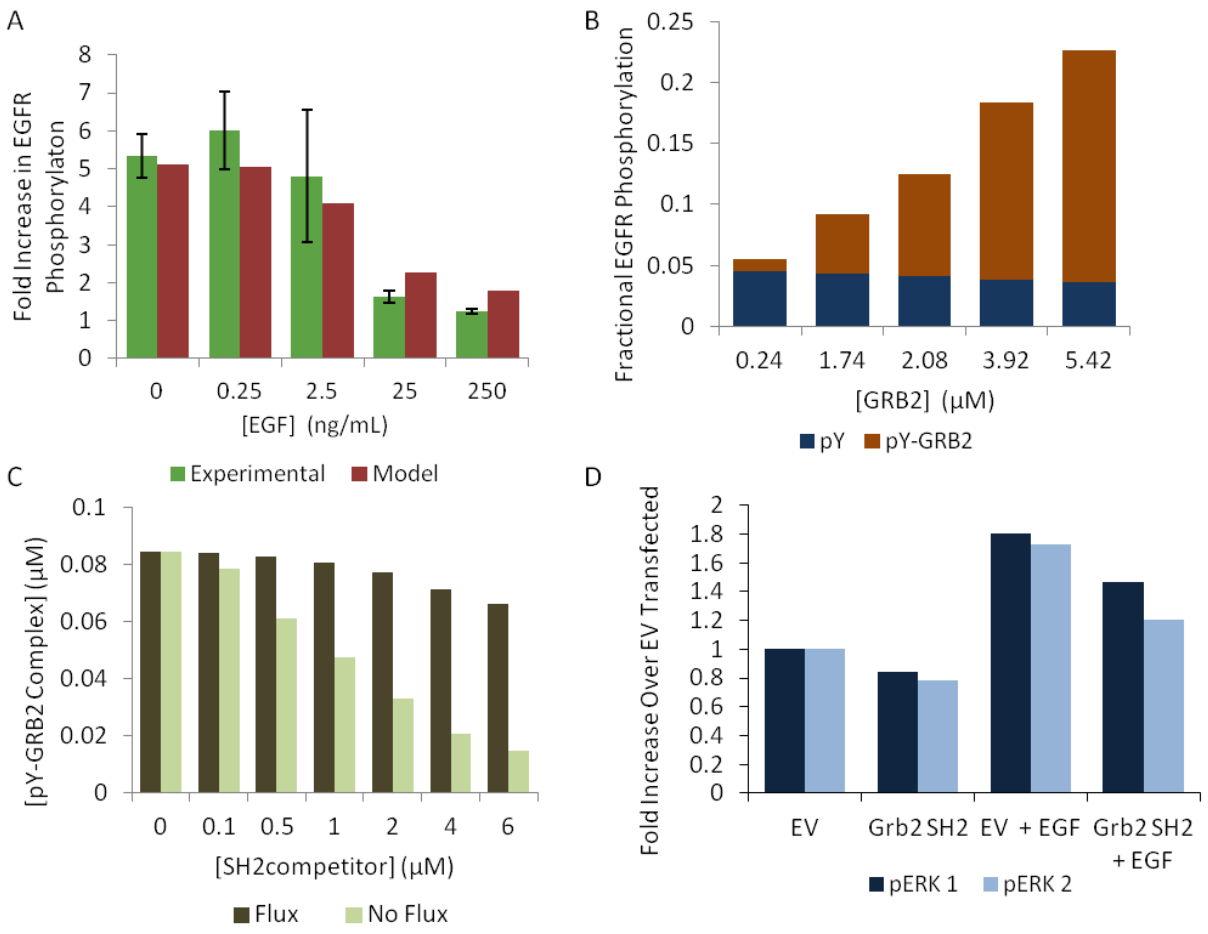

Figure 2.6: Computational modeling recapitulates experimental data. A) Comparison of model predictions (*red bars)* and experimental data (*green bars)* assessing the relative effect of EGF concentration on GRB2 SH2-mediated EGFR phosphosite enhancement. See Figure 2.1 E. Error bars represent the standard error of two technical replicates. For modeling data, the GRB2 concentration was set at 5.4 μM (maximum expression by transfection) B) Model predictions of the relationship between SH2-bound phosphosites (pY-SH2, *orange*) and unbound phosphosites (pY, *blue*). C) Model predictions showing the effect of an increasing concentration of a GRB2 binding site competitor (SH2competitor) on the amount of pY-bound GRB2. Dark green bars represent pY-GRB2 binding in a system with pY flux (i.e. phosphorylation and dephosphorylation; Figure 2.5 A). Light green bars represent pY-GRB2 binding data in a system

containing a constant 3.6% pY EGFR (equal to the amount of pY present in the Flux model without any SH2 present). D) Quantification of phosphorylated ERK1 (*dark blue*) and ERK2 (*light blue)* in COS-1 cells expressing tdEOS GRB2 SH2 before and after EGF stimulation. All values are normalized to empty vector-transfected unstimulated cells and total ERK expression.

affinity, but not GFP-tagged wt GRB2 SH2, suppressed the growth of EGFR-expressing cell lines in an anchorage independent growth assay (*161*).

Discussion

Phosphosite protection by SH2 domain containing proteins is a familiar concept within the field of phosphotyrosine signaling, but has not been rigorously investigated to date. Seminal experiments exploited this phenomenon to identify specific SH2 binding sites *in vitro*, but its occurrence and importance *in vivo* are less understood. (*146*, *147*, *151*). Here we present a focused study of this phenomenon using the EGFR-GRB2 interaction as a model.

The results of our work suggest that GRB2 overexpression can lead to enhanced phosphorylation of EGFR in an SH2- and dose-dependent manner. Replacement of the GRB2 SH2 domain with the CRK SH2 domain resulted in a shift in protection from EGFR to the focal adhesion protein p130CAS, an established CRK SH2 binding protein, suggesting that enhancement depended on SH2 binding specificity (*157*). The observed enhancement of specific binding sites was concentration-specific. Expression of GRB2 at concentrations close to the GRB2-EGFR dissociation constant resulted in high specificity for enhancement of canonical GRB2 binding motifs on EGFR; by contrast, very high concentrations of GRB2 (between 4 and 6 μM) enhanced the phosphorylation of both canonical and non-canonical binding sites. This loss of specificity is likely due to increased binding to lower affinity sites as the SH2 concentration increases. Alternatively, SH2 binding might prevent dephosphorylation of unbound phosphorylated sites by sterically preventing phosphatase access. It been suggested that no more than three SH2-containing proteins can bind to EGFR at once (*49*). In this scenario, as GRB2 SH2 concentration increases and binding nears saturation, phosphorylation of unbound sites increases because they are less accessible to cellular phosphatases even as they remain unbound.

Analysis of the tyrosine phosphoproteome by MS revealed enhancement in the expected canonical SH2 binding motifs for the GRB2 and CRK SH2 domains (Fig. 4B and C), consistent with a mechanism dependent on SH2-mediated phosphosite protection. While our experiments mostly identified known interaction partners for the SH2 domains tested, this approach may be useful to identify interaction partners for SH2 domains with more poorly defined specificities, particularly those which are difficult to purify for use in *in vitro* studies (*36*, *38*, *162*). Furthermore, this method would allow for the identification of interactions that occur *in vivo*, in the cell type of choice. For this purpose it would be significantly simpler than alternative methods utilizing inducible covalent binding or biotinylation, which require mutation and optimization of SH2 domains (*90*, *91*).

To better understand the basis for SH2-mediated phosphosite protection, we created a compartmental deterministic ordinary differential equation (ODE) model using the Virtual Cell reaction modeling software (*106*). An important feature not previously incorporated into models was experimentally determined steady state phosphorylation and dephosphorylation rates. We determined these values using a method we recently developed to quantify the absolute amount of phosphotyrosine in a sample (*50*). While not done in this study, this approach could be combined with percent phosphorylation and receptor concentration data to determine to absolute rates of phosphorylation and dephosphorylation for individual phosphosites.

Model predictions recapitulated experimental data quite accurately, indicating that phosphosite protection is sufficient to explain increased receptor phosphorylation upon SH2 domain overexpression in our system. Arguably, a more complex model incorporating rule-based modeling and multiple phosphorylated sites might provide more detailed insight into the system (*163*). However, the simplicity of the model used in here makes it much more flexible for application to other systems. For instance, our lab and others have previously shown that the

SH2 domain of CRK appears to prevent p130CAS tyrosine dephosphorylation (*164*, *165*). This could easily be modeled, using similar methods to quantify protein amounts and phosphorylation levels.

One rather counterintuitive prediction from the model is that phosphosite protection reduces competition between SH2 domains with overlapping specificities. Due to SH2-mediated phosphosite protection, output from one SH2-containing effector can increase or decrease depending on the local concentration of that SH2, without greatly affecting output from other SH2-containing effectors with similar binding specificity. In this way SH2 signaling through a specific pY motif is additive, if the following three conditions are met. First, SH2 domains in the system must bind similar pY sites. As mentioned above, most SH2 binding motifs fall into several broad classes, suggesting that significant specificity overlap exists (*37*, *166*). Second, phosphatase activity must be high and as a result most potential phosphosites must be unphosphorylated. Consistent with this, previous studies showed that the stoichiometry of tyrosine phosphorylation was low both in starved and stimulated cells (*167*). This is also demonstrated by the dramatic increase in tyrosine phosphorylation seen after treating cells with the phosphatase inhibitor pervanadate (Figure 2.5 B and C).

The final requirement is that SH2 domains have moderate affinities for their phosphotyrosine binding sites, so that binding is not saturated at intracellular concentrations. Recent work that exploited protein structure data to develop high-affinity SH2 domains suggests there is selective pressure to maintain moderate affinities (*161*). *In vitro* affinity measurements showed that K_d values for SH2-EGFR interactions largely fall between 0.8 and 4 μM, with each phosphosite interacting with multiple SH2 domains (*89*, *168*). Of course there are a few examples of SH2-pY interactions with high affinities, which may have evolved to drive specific signaling pathways and eliminate significant SH2-dependent signaling cross-talk (*169*, *170*).

Phosphosite protection may also play an important biological role in oncogenic signaling. SH2 proteins that are largely unbound at normal expression levels may behave very differently when highly overexpressed. Viral CRK (v-CRK) and its human homolog CRKI are able to transform fibroblasts, despite consisting of only an SH2 and SH3 domain. Presumably these CRK proteins exploit phosphotyrosine flux on their binding proteins to create their own binding sites through protection, allowing them to induce excess signaling through pro-growth pathways (*64*, *151*). Likewise, the SH2 containing adaptor GRB7 is often overexpressed in breast cancers overexpressing the orphan RTK and EGFR family member HER2. GRB7 overexpression is associated increased HER2 phosphorylation, which results in activation of pro-oncogenic downstream pathways and increases in cell growth and migration (*171*, *172*). The relative lack of SH2 competition associated with phosphosite protection may also allow for cancer cells to maintain homeostatic signaling processes in the setting of increased expression of oncogenic SH2-containing proteins such as SRC, BCR-ABL, JAK, or STAT.

Finally, it is important that phosphosite protection be considered both in experimental design and data interpretation. For example, recently published work has shown that SHCD overexpression results in a significant increase in phosphorylation of EGFR in a PTB-dependent manner. The largest increases occurred at the SHCD PTB binding site, pY1148, and to a lesser extent at pY1068 and pY1173 (*154*). Although this study does not consider protection from phosphatases as a potential explanation for the findings, the data are highly consistent with this mechanism. More generally, fluorescently tagged SH2 domains and SH2-domain containing proteins are sometimes used as probes *in vivo* to monitor the availability of binding sites for their endogenous counterparts (*94*, *157*). The data presented here suggest that SH2-only probes would only marginally affect binding of endogenous proteins, and therefore can be used as tracers of SH2-pY interactions *in vivo* without significantly disrupting signaling. On the other hand, isolated SH2 domains have also been used as dominant negative reagents, in principle

blocking signaling from phosphosites by preventing binding of endogenous, wt proteins that normally bind that site (*173–177*). Our work shows that such dominant negative approaches will typically not be effective unless expression levels are extremely high, as at more moderate expression levels significant unbound phosphosites remain.

Taken together, the work presented here suggests that phosphosite protection is an inherent characteristic of the phosphotyrosine signaling system *in vivo*. This feature may allow a specific pY site to provide both homeostatic signaling and a diverse set of cell-type specific SH2-mediated signaling outputs at the same time. The process is also likely an important feature of signaling through other dynamic post-translational modifications (PTMs) including serine/threonine phosphorylation and lysine acetylation, both of which can signal through modular reader domains in manner closely resembling SH2-pY interactions. Thus, a complete understanding of signaling from these and other writer-eraser-reader systems requires consideration of the innate ability of reader proteins to participate actively in the signaling process.

Chapter 3

KRASV12 Associated EGFR Dephosphorylation

Joshua A. Jadwin[1] and Bruce J. Mayer[1]

Affiliations

1. Raymond and Beverly Sackler Laboratory of Molecular Medicine, Department of Genetics and Genome Sciences, University of Connecticut School of Medicine, 263 Farmington Avenue, Farmington, CT 06030

Abstract

Non-small cell lung cancer (NSCLC) is a leading cause of cancer death in the United States and around the world. Approximately 70% of all NSCLC tumors contain an activating mutation in either the receptor tyrosine kinase (RTK) EGFR or the small GTPase RAS. Furthermore, dysregulation of signaling by RTKs, the RAS family of proteins and their downstream effectors has been shown to be essential for the growth of a number of tumor types. Specifically, elevated levels of EGFR phosphorylation due to receptor mutation or overexpression are known to induce cellular transformation. Recent data from our laboratory and others have shown that expression of the constitutively active RAS family mutant, KRASV12, results in reduced EGFR phosphorylation in NSCLC cells (*178–180*). Our preliminary data suggest that these changes to EGFR phosphorylation are mediated by protein-tyrosine phosphatases (PTPs), a relatively understudied family of proteins which function in opposition to tyrosine kinases by dephosphorylating tyrosine residues which prevents the binding of downstream effector proteins. We hypothesize that KRAS drives a previously undescribed feedback pathway which enhances the activity EGFR targeting PTPs.

Introduction:

Over the past twenty years there has been a concerted effort by the scientific and clinical communities to apply our understanding of the basic molecular biology of cancer to develop therapies that specifically target selected molecular pathways that drive tumor growth, with a particular emphasis placed on inhibitors of tyrosine kinases and their downstream effectors (*181*, *182*). Effective application of these new cancer treatments requires an in depth understanding of how each target protein fits into the web of interconnected molecular signaling pathways (*183*). Of particular importance is understanding the role potential drug targets play in regulating important upstream signaling hubs through negative feedback.

The Epidermal Growth Factor Receptor is an excellent example of such one such signaling hub, as it has multiple downstream effectors, many of which are potential targets for drug development. As outlined in the Introduction to this thesis, EGFR activation by EGF leads to reciprocal phosphorylation of a set of tyrosine residues on the C-terminal tail of the paired receptor. Each phosphotyrosine residue is then capable of recruiting a unique set of SH2 domain containing adapter proteins and enzymes to EGFR at the membrane where they propagate the signal to their downstream effectors. These effectors include KRAS, as well as a number of other proto-oncogenes and important signaling proteins that promote cell proliferation, motility survival and differentiation (*66*, *184*).

KRAS, which belongs to the RAS family of small GTPases, is constitutively activated in nearly 30% of all cancers and 30% of NSCLCs, as result of an activating missense mutation in either codon 12 or 13 (the common mutant KRASV12 being used here). This mutation prevents conversion to the inactive form of KRAS resulting in continuous mitogenic signaling through the RAS-RAF-MEK-ERK pathway (*185*). Interestingly, activation of these two genes is almost always mutually exclusive indicating an antagonistic relationship between EGFR and KRAS.

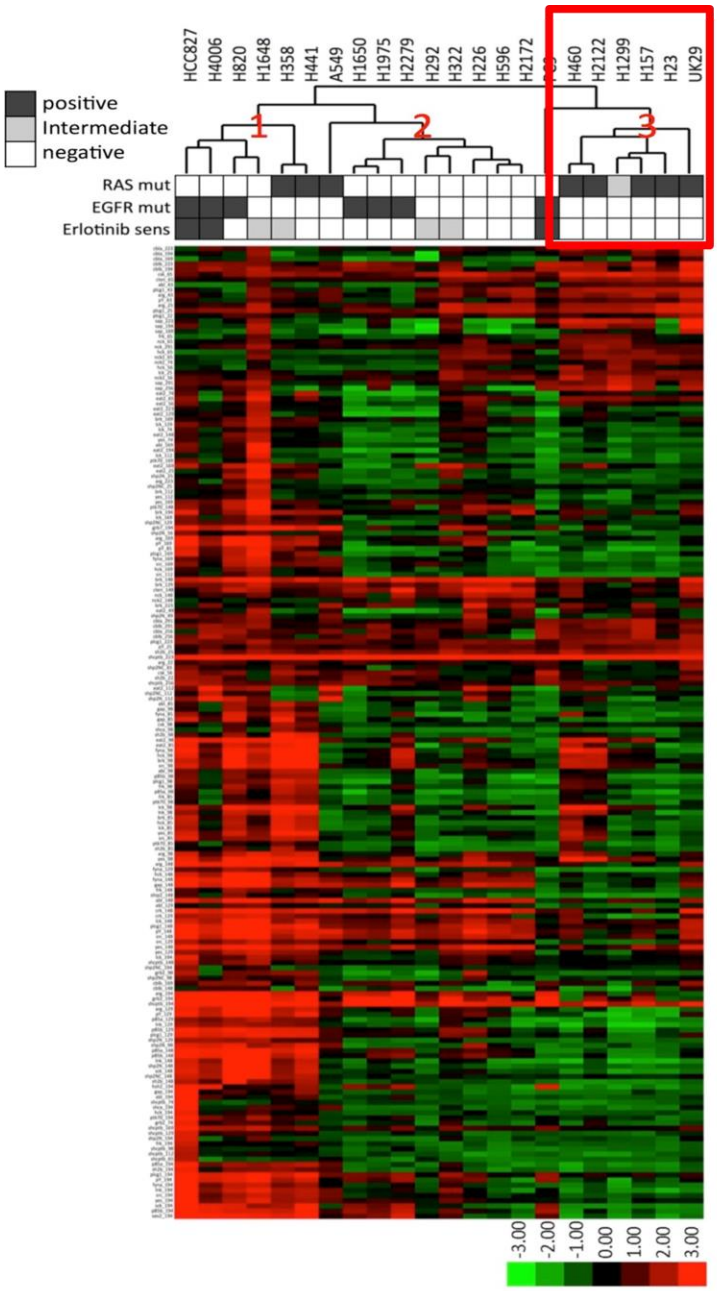

Figure 3.1: SH2 profiling of NSCLC cell lines. A subset of KRASV12 expression cell lines display a unique SH2 binding signature and a relatively low overall tyrosine phosphorylation. Red box indicates the KRASV12 cell subset, labeled with "3". The cell line with the light gray box was initially identified as KRAS mutant negative, but was later shown to express the constitutively active form of the NRAS. Figure adapted from Machida et al 2010.

A recent study published by our lab using SH2 domains to profile the global phosphorylation state of large set of NSCLC cells lines suggested that NSCLC lung cancer cells lines expressing KRASV12 displayed a unique global phosphotyrosine signature, including suppressed EGFR phosphorylation (Figure 3.1)(*67*). Previously published quantitative MS and KRAS knockdown experiments performed using NSCLC cell lines also suggested that KRAS activation was associated with the suppression of EGFR phosphorylation (*178*, *186*). Here we present preliminary data suggesting that KRASV12 expression suppresses EGFR phosphorylation by enhancing EGFR dephosphorylation through a PTP mediated mechansim.

Results

Exogenous KRASV12 expression suppresses EGFR phosphorylation

To explore the relationship between KRASV12 expression and EGFR phosphorylation we created KRASV12 expressing NSCLC cell lines (H322, H2172). KRASV12 cell lines displayed both increased KRAS expression and KRAS activity as measured by activated RAS pulldown. Both KRASV12 cells and their mock controls had very low levels of phosphorylation in serum starvation conditions. However, when these cells were stimulated with 2.5ng/mL of the EGF for 5 min, we observed a significant decrease in the phosphorylation of EGFR in the KRASV12 cells, as measured by anti-pY antibody (Figure 3.2).

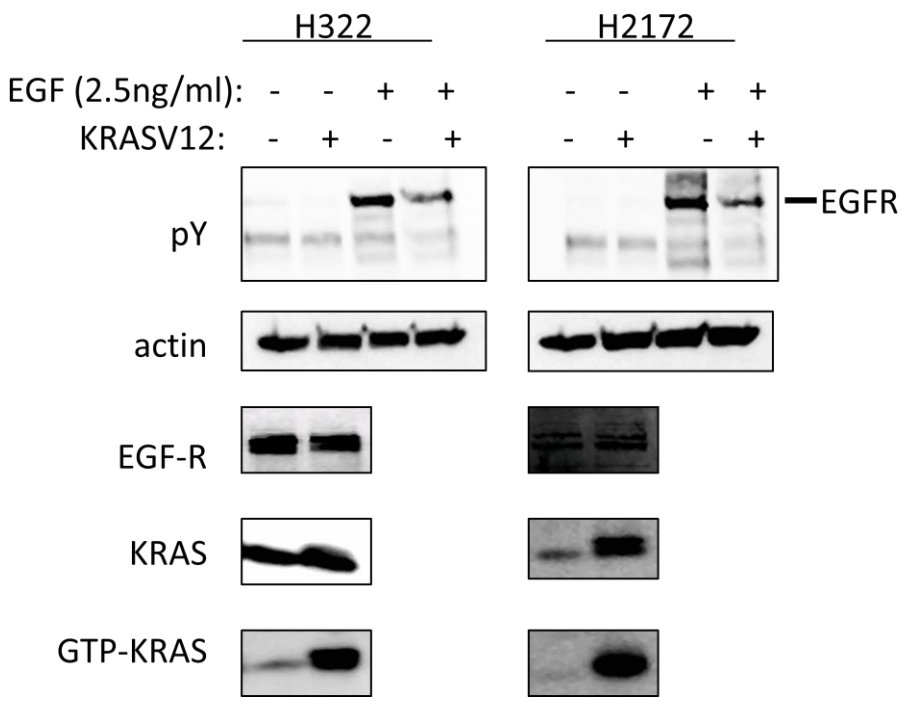

Figure 3.2 Exogenous expression of KRASV12 in NSCLC cell lines suppresses EGFR phosphorylation. Stable KRASV12 or empty vector expressing H322 and H2172 cell lines were created by retroviral infection and drug selection. Cells were serum starved overnight and stimulated with EGF. KRASV12 expression resulted in a significant reduction in phospho-EGFR. KRAS activity in each line was assessed by Ras-Binding-Protein pulldown and are shown in the bottom panel.

KRASV12 mediated suppression of EGFR phosphorylation is not the result of receptor internalization

One of the major mechanisms by which EGFR is regulated is via receptor internalization (*19*, *187*). Activation and autophosphorylation of the receptor leads to the recruitment of the ubiquitin E3 ligase Cbl, an SH2 domain containing protein (see Figure 1.2). Receptors can be internalized through both clathrin-dependent and clathrin-independent processes, which have been shown to occur between 10 and 60 minutes after receptor activation. Internalized receptors are sorted through the endosomal sorting complex and either recycled and sent back to the membrane or trafficked to the lysosome for degradation.

There is some evidence suggesting that KRAS activation stimulates receptor internalization (*188*). In an effort to rule out this possibility we compared the amount of EGFR at the plasma membrane in KRASV12 and empty vector expressing cells. Cells were treated with Sulfo-NHS-SS-Biotin, a membrane impermeable biotin conjugate which forms disulfide bonds with accessible lysine residues on surface proteins. The cells were then lysed and biotinylated proteins were pulled down using avidin beads. Both the mock and KRASV12 H2172 cell lines had similar levels of membrane associated EGFR (Figure 3.3). This suggested that the observed decrease in EGFR stimulated EGFR phosphorylation was not the result of KRAS driven receptor internalization and that other mechanisms should be explored.

KRASV12 enhances the rate of EGFR dephosphorylation

The level of tyrosine phosphorylation on any protein is a function of the balance between kinase and phosphatase activity and potentially the concentration of pY binding proteins (see Chapter 2). To examine the role of kinase activity in the KRASV12 associated decrease in EGFR

phosphorylation, we measured the basal rate of EGFR phosphorylation in cells with and without the constitutive KRAS by experimentally eliminating phosphatase activity using pervanadate. When EGFR kinase activity was isolated in serum starved mock and KRASV12 cells using pervanadate no difference in total receptor phosphorylation or the phosphorylation rate was observed, suggesting that KRASV12 was not somehow suppressing the EGFR kinase (Figure 3.4).

We then examined the effect of KRASV12 overexpression on the rate of EGFR dephosphorylation. EGFR targeted PTP activity was isolated by suppressing EGFR kinase activity using the TKI Erlotinib (*42*). Interestingly, we found a nearly 4 fold increase in the rate of dephosphorylation in H2172 cells expressing the KRAS mutant (Figure 3.5). These results suggest that KRASV12 expression enhances the activity or expression of PTPs acting on phosphorylated EGFR.

KRASV12 may suppress the growth NSCLC cells driven by constitutive EGFR activation

If KRAS activation does in fact suppress EGFR phosphorylation, this might explain why activating mutations in these proteins are rarely ever observed together in clinical samples. By suppressing EGFR phosphorylation and the resultant downstream signaling, KRASV12 might reduce cellular fitness compared to cells expressing the EGFR mutant alone. To examine this interesting correlation, we created a KRASV12 expressing cell line using HCC827 cells, an NSCLC cell line which expresses a constitutively active form of EGFR. Unlike non-mutant EGFR cell lines, KRASV12 expressing HCC827 cells did not display a noticeable difference in receptor phosphorylation (Figure 3.6 A). However, when we assessed the rate of EGFR dephosphorylation in these cells, we observed a noticeable increase in cells expressing KRASV12, similar to that seen in wild-type EGFR cells. (Figure 3.6 B).

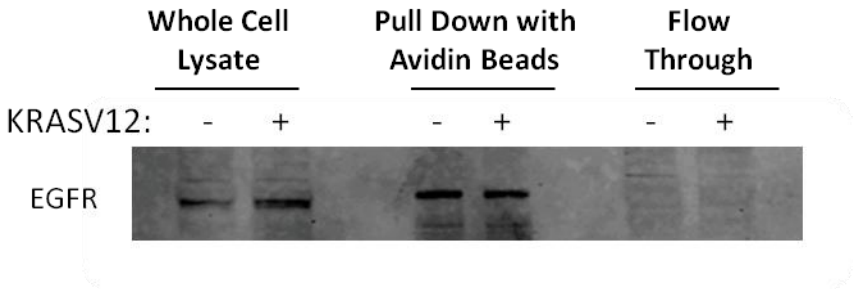

Figure 3.3: Membrane localization of EGFR in KRASV12 cells. Extracellular and transmembrane proteins on unpermeablized H2172 KRASV12 and empty vector cells were biotinylated using NHS-S-S-Biotin at 4^{o}C prior to lysis. Biotinylated proteins were isolated with avidin labeled beads, run on SDS-PAGE and immunoblotted for EGFR. No difference in total EGFR or membrane bound EGFR was observed suggesting that decreased EGFR phosphorylation in KRASV12 cells is not a result of increased receptor internalization.

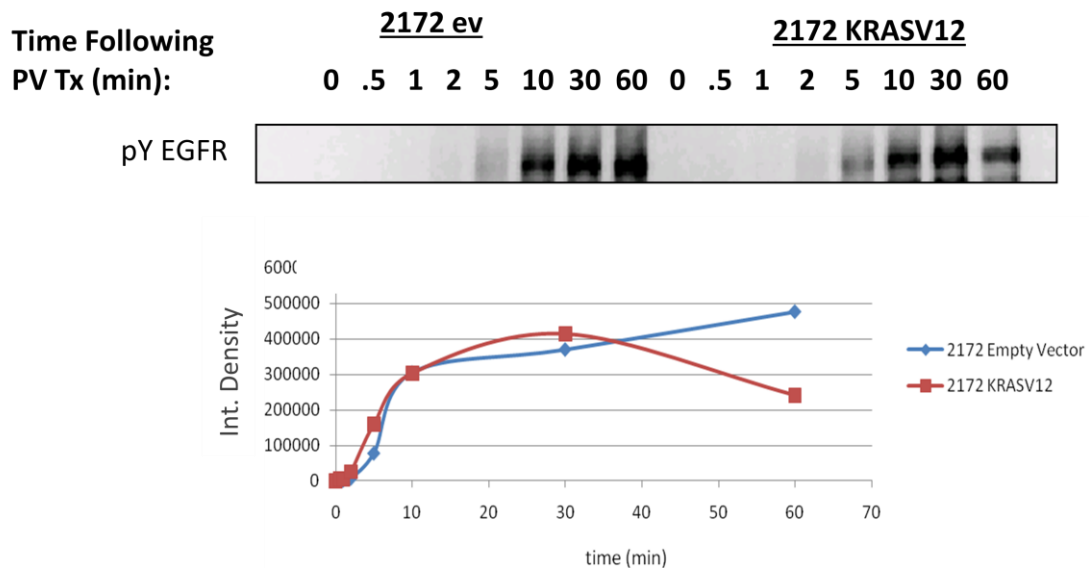

Figure 3.4 Quantification of basal EGFR kinase activity in KRASV12 cells. H2172 cells infected with empty vector (ev) or expressing KRASV12 were starved overnight, treated with 100nM pervanadate, lysed at the times shown and run on SDS-PAGE. No difference in basal EGFR phosphorylation rate was observed indicating that the EGFR phosphorylation rate was unchanged by KRASV12 expression. Y-axis values are non-normalized densitometry measurements.

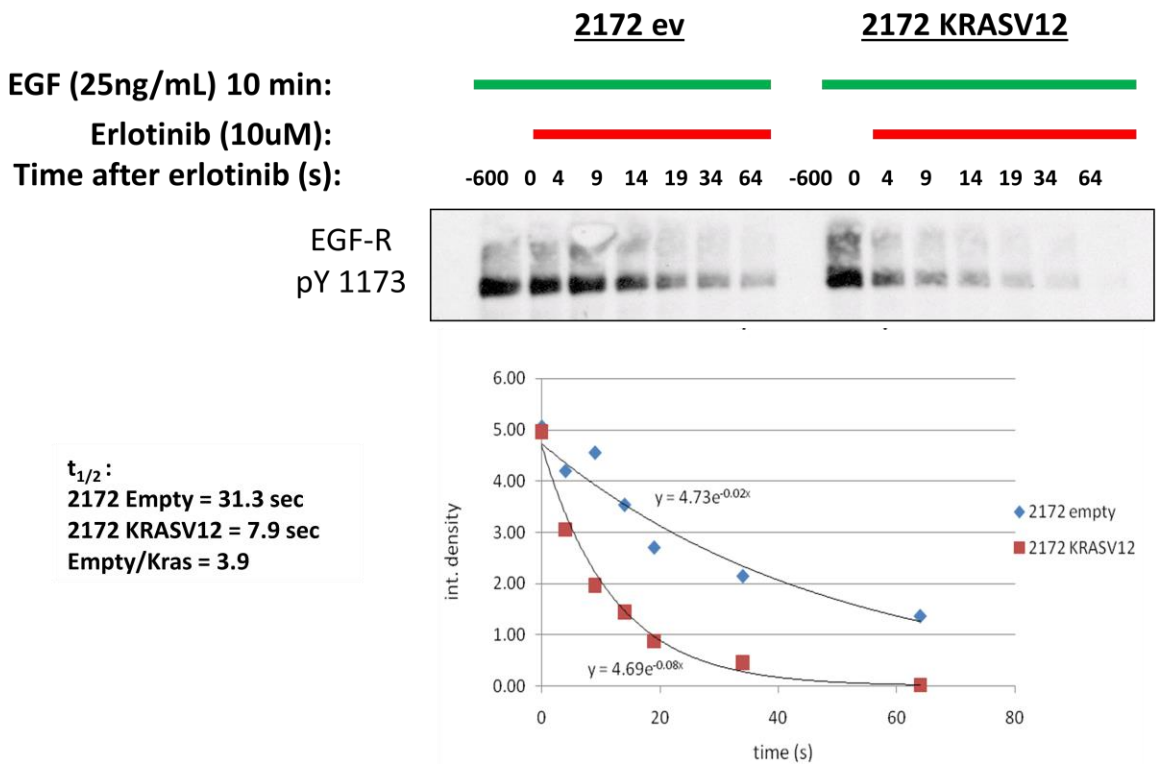

Figure 3.5: Quantification of EGFR dephosphorylation rate in KRASV12 expressing H2172 cells. Serum starved H2172 KRAV12 and empty vector cells were stimulated with 25ng/mL EGF for 10 min (-600 min). At the 10 minute point cells were treated with 10uM of the selective EGFR TKI Erlotinib, flash frozen at the times shown, lysed and immunoblotted using an antibody which specifically recognizes EGFR pY1173 (top panel). The EGFR dephosphorylation half-time was nearly 4 times shorter in KRASV12 expressing cells. Phosphorylation half-time values were determined by fitting each data set to an exponential decay function (bottom panel). Calculated half-times for each cell line are shown to the left of the graph. KRASV12 cells-red dots. EV cells-blue dots.

To assess the functional effect of KRASV12 expression on EGFR driven cell growth we compared the monolayer culture growth rate of the two cell lines using MTT assay. Over a period of 144 hrs no difference in growth was observed (Figure 3.7). However, most cells lines, both untransformed and transformed, grow rapidly in monolayer culture and as a result subtle differences in cellular growth may not be apparent. To more thoroughly assess the differences in cell transformation associated with the expression of both mutations, we the performed an anchorage independent growth assay. This method, also known as a soft agar assay, is a classic measure of cell transformation, in which cells are grown suspended in Bacto-agar prepared using cell culture media. Interestingly, using this method we observed, a nearly 10% decrease in the total number of colonies ($p=0.21$) and a 25% decrease in colony size associated with KRASV12 ($p=0.11$) (Figure 3.8 A-C). While these differences were not statistically significant they do suggest that KRAS signaling might suppress EGFR driven tumor growth.

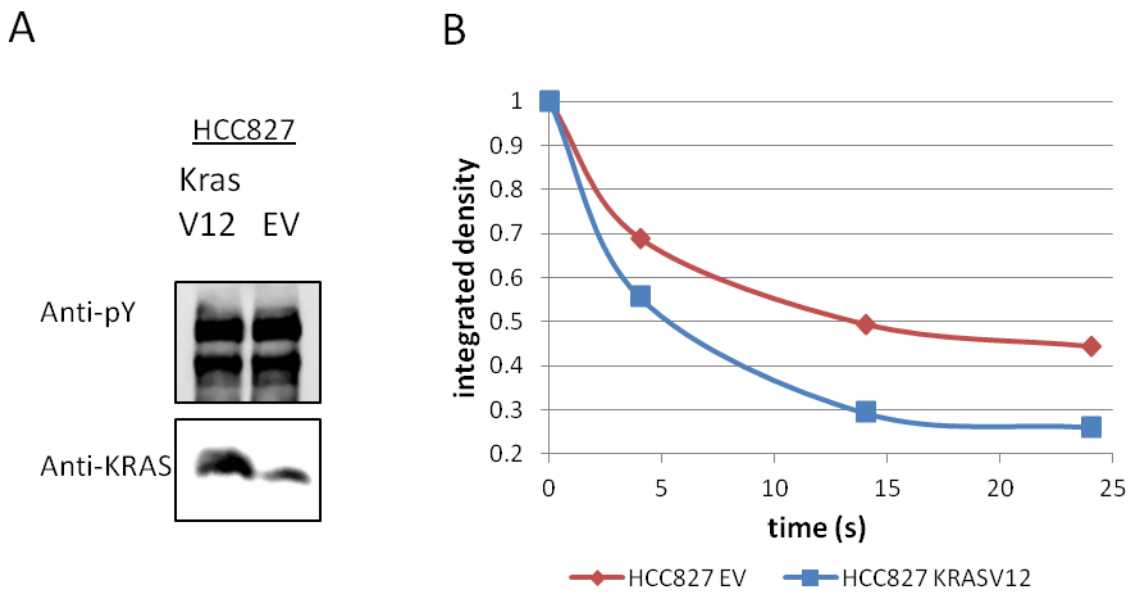

Figure 3.6: Effect of KRASV12 on EGFR dephosphorylation in HCC827 cells driven by a constitutively active form of EGFR. A) Western-blots showing EGFR phosphorylation and KRASV12 expression in KRASV12 and empty vector HCC827. B) Dephosphorylation curves obtained by treating starved HCC827 cells with 10uM Erlotinib. Data did not fit the exponential decay function. However, direct measurements of curves suggest that the half-time of EGFR pY1173 is approximately 3 times lower in KRASV12 expressing cells. KRASV12 expressing cells were also dephosphorylated to a lower level (last time points). Empty vector cells are shown in red. KRASV12 cells are shown in blue.

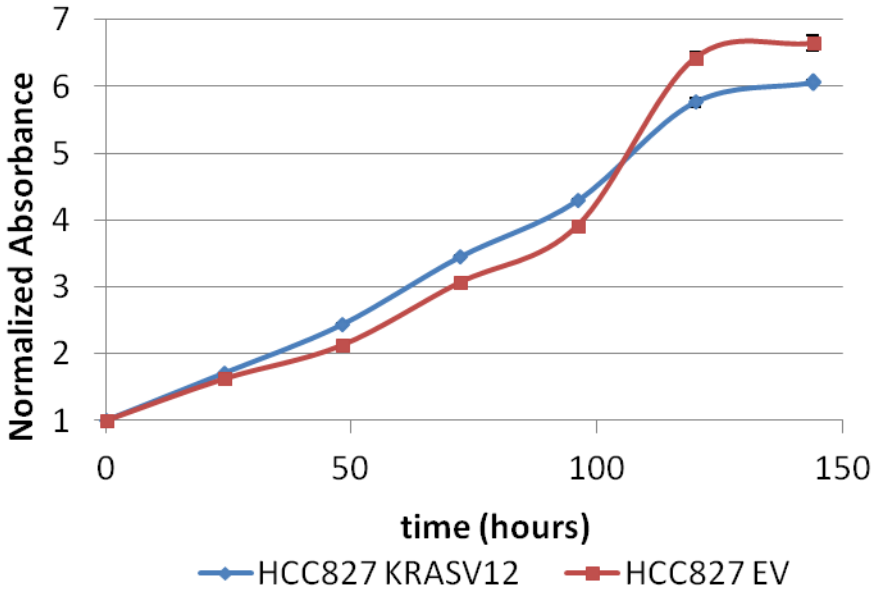

Figure 3.7: Monolayer cell growth of KRASV12 and empty vector expressing HCC827 cells. Plot showing the relative change in KRASV12 and empty vector cell number over 144 hrs in monolayer growth. Reduction of the tetrazolium dye MTT by cellular oxidoreductases was used as a proxy for total cell number and measured by absorbance. First time point values for both cell lines were are normalized to 1. Results represent average of three technical replicates.

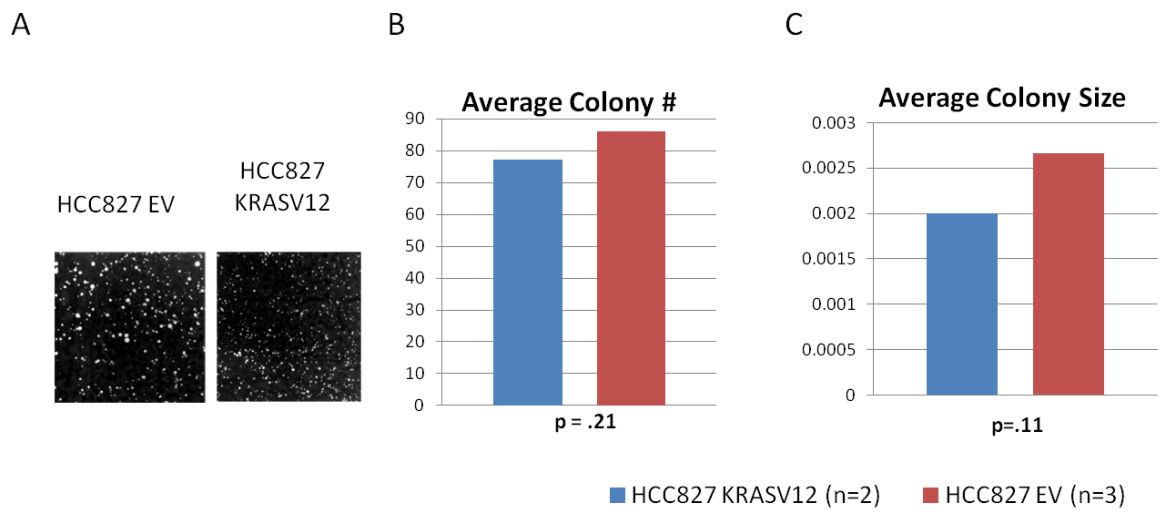

Figure 3.8: Anchorage-independent growth of KRASV12 and empty vector expressing HCC827 cells. A) Representative image of colonies from KRASV12 and empty vector cells. B) Quantification of average colony number for each cell line. C) Quantification of average colony size for each cell line (inches^2). Statistical analysis - Student's t-test. n= number of biological replicates.

Discussion:

As evident from the work presented so far in this thesis, regulation of receptor tyrosine kinase phosphorylation can occur at multiple levels. The work presented here suggest that elevated KRAS signaling leads to increased phosphatase activity which in turn suppresses EGFR phosphorylation. Expression of the constitutively active KRASV12 resulted in a decrease in EGFR phosphorylation, which was most prevalent following EGF stimulation. This effect appears to be a PTP mediated process, as receptor dephosphorylation occurred much more rapidly in KRASV12 expressing cells.

From a basic biological perspective KRAS mediated negative feedback is rational. Excessive EGFR signaling can lead to aberrant cell growth and eventually cellular transformation. By initiating downregulation of receptor phosphorylation, KRAS may prevent unopposed pro-growth, anti-apoptotic signaling. Interestingly, this is not the first example of KRAS induced phosphatase mediated regulation. KRAS activation has been shown to stimulate phosphatase mediated negative feedback of activating serine and threonine phosphorylation on mitogen activated protein kinases (MAPKs) activation. This process is mediated by dual-specificity phosphatases (DUSPs), capable of acting on pS and pT, as well as pY(*189*). While no DUSP has been shown to dephosphorylate EGFR, further inquiry into DUSPs as a potential mediator of the observed findings may be warranted.

While invoking a new role for DUSPs in EGFR signaling regulation is an interesting hypothesis, classical PTPs are the most likely drivers of our findings. A number of PTPs, including PTPN9, RPTPκ and RPTPη have been shown to dephosphorylate EGFR (*190*–*192*). However, identification of the specific PTPs involved in KRAS mediated EGFR dephosphorylation remains a difficult task due, at least in part, to the general lack of PTP substrate specifity. A recently developed method for identifying and quantifying a cell's specific PTP expression (PTPome) and PTP activation state (oxizided PTPome) may provide some insight in the specific PTPs

involved (*193*). While a preliminary assessment of the PTPome suggested that no large difference exists in PTP expression between these KRASV12 and empty vector cell lines (data not shown), a more in depth analysis of PTP expression and activity using both western and mass spectrometry could provide more insight.

Although the current results are still equivocal, our studies of KRASV12 expression in mutant EGFR driven NSCLC cells suggest that a PTP mediated process may contribute the clinical mutual exclusivity observed between KRASV12 and activated EGFR. These finding are particularly interesting because they suggest that oncogene induced transformation is a measured process, where more is not always better. RAS activation or expression might be tempered in order to maintain optimal signaling through all EGFR mediated pathways. However, further analysis of the interaction between the two processes is needed. It is also worth noting that the absolute level of phosphorylation in the presense or receptor specific inhibitors was lower in KRASV12 cells. This may be the result EGFR phosphorylation by non-EGFR TKs which is also suppressed in KRASV12 expressing cells.

Finally, the work presented here only examined total EGFR phosphorylation and EGFR pY1173. KRASV12 expression may lead to the dephosphorylation of specific phosphosites on EGFR. Understanding the effect on each site will be essential to understanding how different downstream pathways are affected. Moreover, it behooves us to examine the effect of KRASV12 expression on the phosphorylation status of other proteins, including other RTKs. The anti-pY blots in figure 4.2 suggest that this process may not be EGFR specific, but instead may be the result of a generalized increase in cellular PTP activity.

Summary

The work presented here examines the regulation of phosphotyrosine signaling on three different levels using the Epidermal Growth Factor Receptor as a tool to understand universal properties of pY signaling systems. These include temporal changes in receptor localization and phosphorylation, the effect of SH2 protein expression on maintenance of tyrosine phosphorylation and the role of negative feedback from downstream effectors. This work only touches on the plethora of known regulatory mechanisms known control signaling from this important oncogene. Nevertheless, our hope is that the mechanisms presented herein will have a much wider impact on how researchers in the field understand and approach phosphotyrosine signaling on the whole.

We begin with a general overview of pY signaling, the protein families involved, the specific proteins assayed and the different methodological approaches used. The introduction highlights the importance of approaching phosphotyrosine signaling as a dynamic process in which the flux of phosphotyrosines is dependent on kinase and phosphatase activity, as well the specificities and concentrations of specific SH2 domains. We focus on the utility of SH2 domains as a Domainomics tool to assess phosphotyrosine dependent signaling. Information about the orthogonal methods and techniques utilized in subsequent chapters are also discussed.

Chapter 1 follows with a large scale investigation into the spatial and temporal regulation of EGFR activation and SH2 binding. Here we compare phosphotyrosine and SH2 binding dynamics on three different levels; tyrosine phosphosite dynamics assayed by quantitative MS, SH2 binding site dynamics assayed by FW and SH2 domain membrane recruitment and binding *in vivo* assayed by live cell membrane imaging. The work, which was performed on a global scale includes: a temporal analysis 100 unique pY sites by MS, SH2 binding dynamics following EGF stimulation for 27 domains and 22 time points, and an *in vivo* analysis of membrane recruitment and diffusion values for 25 different SH2 domains. The most important finding from

this work was that SH2 recruitment to the membrane is delayed in comparison to SH2 site availability, which occurs rapidly in response to EGF. These findings indicated that SH2 membrane recruitment is not in equilibrium with pY site creation. Deconvolution of *in vivo* binding data suggested that this finding may be due to binding site clustering at the membrane. As clustering progresses in response to EGF stimulation, the likelihood of SH2 membrane rebinding increases. This means the rate at which SH2 domains leave the membrane decreases. As a result, SH2 recruitment continues to increase until SH2 binding sites have clustered. In all, this research both adds to a growing body of work the supporting role of spatial organization in regulating protein signaling and highlights the importance of challenging one's own results using orthogonal methodologies.

Chapter 2 focuses on the role SH2 domain containing proteins play in maintaining their own binding sites. This work revisits the concept of SH2 mediated phosphosite protection, a hypothesis derived from early the results of early *in vitro* studies, verifying and quantifying its effects *in vivo*. We exploit the well studied interaction between pY-EGFR and GRB2 to quantify role of SH2 concentration and kinase activity on binding site protection. The specificity of GRB2-SH2 mediated enhancement is assessed on two levels. First in relationship to the specific tyrosine sites on EGFR using traditional biochemical methods and then in relation to the entire proteome using quantitative MS. Finally, we recapitulate our protection results using computational reaction modeling, incorporating concentrations and reaction rates obtained by *in vivo* measurements (using live cells as opposed to recombinant proteins). Finally, using both computational and biochemical assays we propose a model in which competition between SH2 domains from different proteins with shared specificities is relieved by the ability of SH2 domains to maintain their own binding sites via protection for phosphatases. If this theory holds true it suggests that increased binding of effectors does not necessarily significantly decrease the

amount of unbound sites available. This would mean that signaling output from multiply phosphorylated proteins such as EGFR is an additive process and not a zero sum game.

Chapter 3 concludes with preliminary data from an investigation into the role KRAS signaling in inducing EGFR dephosphorylation. This paper builds on data published by our lab suggesting that KRAS signaling drives changes in the global phosphotyrosine landscape. Using NSCLC cell lines stably expressing the constitutively active KRASV12 mutant we demonstrated that elevated KRAS activity is associated with a significant reduction in EGFR phosphorylation. We then showed that this finding is associated with an increase the rate of EGFR dephosphorylation. We hypothesized that this negative interaction might contribute to the mutual exclusivity observed between EGFR and KRAS in almost all tumor types. To begin to assess this hypothesis, we stably expressed KRASV12 in a NSCLC cell line containing an activating EGFR mutation. EGFR phosphorylation was not suppressed in these cells, however we did were observe an increase in the EGFR dephosphorylation rate. Anchorage independent growth in these cells also trended lower, but the difference was not statically significant. Together these results support a mechanism by which KRAS negatively feeds back on EGFR by altering the kinase-phosphatase balance within the cell.

Future Directions

Chapter 1

Assess the variations in EGFR phosphosite phosphorylation dynamics.

EGFR far-Western and phosphosite specific Western blotting suggested that small but significant differences exist in the relative rate of phosphorylation at different EGFR phosphosites. These differences may represent an innate mechanism for regulating signal output from the receptor following stimulation. In order to understand this process more fully we propose quantifying the phosphorylation of specific EGFR phosphosites in response to variations in EGF concentration, using different EGFR ligands (e.g. EGF, TNF-α and amphiregulin), and in cells expressing different EGFR concentrations (*160*, *194*). The latter experiments would be best performed in a single cell line, such as A431 cells, with variable levels of EGFR knockdown. It may also be possible to quantify the relative phosphorylation of specific sites under specific conditions using a pY-EGFR standard as described in Chapter 2. By integrating this data we may be able to create topological maps of EGFR signaling output in response varying conditions.

Compare phosphorylation dynamics and SH2 binding for p130CAS using optimized MS protocol

The orthogonal analyses of SH2 binding sites (i. e. phosphosites) performed in Chapter 1 identified a number of inconsistencies between the phosphorylation dynamics quantified by FW and MS. Of particular note, was the phosphorylation of p130CAS, a highly phosphorylated focal adhesion protein which contains number pYXXP, CRK binding sites (*65*). Blot based analyses of p130CAS phosphorylation following EGF stimulation have always suggested that p130CAS is rapidly dephosphorylated in response to EGFR activation (*195*). MS based experiments have been less clear, often returning few peptides and may be confounded by the repetitive CRK

binding sequences found within the p130CAS substrate domain. (*3*). The iTRAQ MS analysis performed in Chapter 1 identified five p130CAS phosphosites, but failed to demonstrate the dephosphorylation observed by FW and pY immunoblot. One potential explanation for this discrepancy is that phosphosites on p130CAS are differentially phosphorylated in response to EGF, with some sites dephosphorylated and others remaining flat or even increasing in response to EGF. This hypothesis is consistent with the fact that FW analysis of phospho-p130CAS identified some SH2s with level or even increased binding following EGF treatment. Our MS analysis may have been skewed toward identifying these sites. In an attempt gain a more complete understanding of p130CAS phosphosite dynamics we propose performing a p130CAS-specific analysis using both FW and MS. Assay sensitivity may be improved by precipitating and concentrating p130CAS from EGF treated lysates prior to protease treatment. We also propose running the MS analysis using nontraditional proteolytic agents such as elastase or chymotrypsin along with trypsin, which was employed in our initial experiments in order to increase the likelihood of detecting sites that were missed in our analysis of tryptic peptides.

Quantify SH2 membrane recruitment and rebinding with induced receptor clustering.

The *in vivo* imaging and pervanadate treatment experiments presented in Chapter 1 provide strong evidence for delayed recruitment as a cluster dependent process. However, these results do leave room for alternate explanations. Validation of our proposed mechanism would require that we monitor SH2 membrane recruitment under specific controlled levels of receptor clustering. Light inducible clustering using Cry2-RTK chimeras provide a mechanism to do just this. Cry2 is a photoreceptor protein isolated from *Arabidopsis* which rapidly and reversibly multimerizes in response to blue light. Cry2 chimeras have been used to study the clustering in

a number of RTKs including TRK, FGFR, and PDGFR. Using one of these established proteins or a new Cry2-EGFR chimera, along with TIR and sptPALM imaging, it should be possible to directly assess the effect of clustering on SH2 membrane recruitment and binding.

Model SH2 membrane recruitment to clustered receptors using Langevin modeling.

As mentioned above, the role clustering plays in the recruitment of SH2 domains to binding sites on the membrane should be further validated using more controlled methods. One way to do this is by modeling the process computationally. Typical reaction modeling solvers such as those used in the Virtual Cell (see Chapter 2) are do not treat molecules as space filling objects and are therefore may not be appropriate for modeling the effects of receptor clustering. On the other hand molecular modeling, which does take molecular volumes into account, is highly computationally intensive, making it difficult to use for dynamic processes involving multiple copies of different proteins such as SH2 domains binding to clustered, phosphorylated receptors. In an effort to find some modeling middle ground, Dr. Paul Michalski, a former member of the CCAM modeling team, designed a general purpose Langevin dynamics simulator that models proteins as a set of sites connected by stiff links. The sites are modeled as impenetrable spheres, which captures the effects of excluded volume and steric hindrance. Using this simulator software, it should be possible to model the recruitment of SH2 domains to phosphorylated receptors at the membrane with variable levels of clustering. This would allow us demonstrate the role of membrane phosphosite rebinding and its effects on recruitment within a tunable system containing a defined set of species (i. e. SH2 and EGFR with binding sites).

Chapter 2

Evaluation SH2 phosphosite protection as a method for identifying SH2 binding motifs/binding partners for SH2 domains which are insoluble as recombinant proteins.

In our attempt to verify that the GRB2 SH2 domain was enhancing EGFR phosphorylation by protecting its binding sites, we assessed GRB2 SH2 associated changes across the entire phosphoproteome. To do this we compared the relative abundance of individual phosphosites in cells with and without exogenous SH2 expression. Using this method we were able to show that the SH2 domains used (GRB2 and CRK) were enhancing sites containing their canonical binding motifs. While these motifs have been validated by a number of previous *in vitro* binding experiments, the binding specificities of a number of SH2 domain families has not been assessed due to the poor solubility and/or pY-binding performance of their recombinant proteins. These include RIN, STAT and JAK family SH2 among others. However, we have expressed some of these SH2 domains in eukaryotic cells as GST-fusion proteins including those of JAK2, STAT2 and SOCS2 (data not shown). We propose using *in vivo* pY protection coupled with MS quantification to define pY binding motifs for such domains. The method would also likely identify new binding partners for these domains, which could then be verified by more traditional methods such as immunoprecipitation.

Utilize the newly developed Vcell hybrid modeling solver create a more complete model of EGFR activation and phosphosite protection.

Though the model used in Chapter 1 was able to recapitulate our *in vivo* data and using quantified parameters, it is still a fairly simplistic model of EGFR signaling. It contains only a single EGFR phosphosite, treats EGFR as a monomer and treats EGFR kinase and substrates domains as individual species. These assumptions were made largely to simplify the model for

our purposes. For instance, adding new phosphosites to a substrate domain can rapidly increase the total number of reactions required to complete a model. When an additional phosphosite is added to a partial differential equation model, reactions for the substrate domain in each of its four possible phosphorylation states must be defined for any interaction involving the receptor. Rule-based modeling can reduce this combinatorial explosion, by defining the interactions that can occur and the sequence in which they occur (*163*, *196*). Until recently this process required that these rules be defined using a specific nonintuitive coding language, which itself is somewhat opaque to individuals without a background in modeling or computer coding. However, a more biologically intuitive interface is currently being developed using the VCell platform which combines both rule-based and partial differential equation modeling. Most importantly though, this software attempts to treat molecules, PTMs and interactions in a graphical manner that is more intuitive to the typical biologist. This software should allow for the development of an EGFR-SH2 interaction model capable of incorporating multiple phosphosites and multiple SH2 domain containing proteins. Moreover, by applying the methods developed in Chapter 2 for absolute phosphotyrosine quantification, site specific percent phosphorylation and quantification of kinase/phosphatase rates, it should be possible create an accurate model of EGFR phosphorylation dynamics for a specific cell type. This might help us understand why, for instance, the GRB2 SH2 domain appears to protect noncanonical pY motifs on EGFR. It could also allow us to quantify the apparent *in vivo* dissociation constants of specific SH2 domains, such as GRB2 SH2 for noncanonical pY motifs.

Test the protection model using a non-pTyr system

The protection model proposed in Chapter 2 was only validated using the reader, writer and eraser proteins involved in pTyr signaling. However, there is no reason that the mechanism of

reader dependent post-translational modification protection should be limited to this pTyr signaling. Serine and threonine phosphorylation, lysine acetylation and arginine/lysine methylation, each have their own set of reader proteins. For instance, bromodomain (BRD) containing proteins bind acetylated lysine residues on histone tails. 61 BRD containing proteins have been identified and a significant body of research suggests that their affinity for any particular histone tail is highly dependent on the post-translational modification (methylation, serine/threonine phosphorylation) of nearby residues (*197*). This suggests that, as with SH2 domains, BRD specificity is both diverse and overlapping. It seems reasonable to suggest that BRDs protect their binding site acetyl groups from the activity of histone deacetylatases. Maintenance of specific histone marks could profoundly affect local transcriptional activity. Interestingly, this mechanism has been invoked in the literature to explain the ability of BDF1 BRDs to maintain euchromatin in yeast (*198*). It therefore seems reason to expand our study of protection into other modular domain containing proteins and in particular BRDs. By using methods such as chromatin immunoprecipitation, DNase protection and transcriptome sequencing to assess the affect of isolated BRDs and full length protein expression on local and global changes in histone acetylation, secondary histone PTMs and transcription patterns, we might gain insight into the innate properties of these domains and the role they play in regulating transcription.

Chapter 3

Assess EGF induced EGFR phosphorylation in KRAS (or a generalized RAS) knockdown cell line

The preliminary data presented in Chapter 3 suggests that constitutive KRAS activity suppresses EGFR phosphorylation by upregulating PTP activity. It is difficult to know if this

feedback is part of the normal signaling pathway, or if it is specific to KRASV12 signaling. As such it would be useful to assess the phosphorylation of EGFR in cells expressing varying levels of wild type KRAS. For example, EGFR phosphorylation in shRNA mock knockdown cells should be lower than in KRAS (or pan RAS) knockdown cells. These experiments could also provide some information about the temporality of this KRAS mediated response. If the feedback occurs through cytoplasmic signaling alone, one might expect EGFR phosphorylation in mock cells to display pY signal attenuation immediately following stimulation (minutes to hours). However, if no difference is observed, it is more likely that the mechanism requires transcription level changes.

Assess the EGFR phosphorylation/dephosphorylation in response to other RAS family proteins.

The RAS family of small GTPase, (ie. KRAS, NRAS, HRAS) share a high level of sequence homology. However, a body of research suggest that they are not entirely interchangeable, with each performing a unique function in cellular transformation (*199*, *200*). In order to better understand the connection between RAS signaling and phosphatase activity we propose repeating our KRAS experiments using the constitutively active mutants of NRAS and HRAS. While it is difficult to quantitatively compare the activity and effects of the different mutants, it would be useful to know if phosphatase activation and EGFR dephosphorylation is a general feature of RAS signaling.

Identify KRAS effectors which mediate phosphatase regulation

Our current data strongly point to a connection between KRAS and phosphatase activity. However, the signaling pathways linking these two remains a black box. To establish a molecular connection between KRASV12 and EGFR dephosphorylation, we propose assessing EGFR phosphorylation in cells in which the activity of well established KRAS effectors (i. e. PI3K, AKT, MEK, and PKC δ) have been manipulated. A combination of commercially available inhibitors, RNAi and constitutively active mutant overexpression could be used to identify proteins directly involved in our findings. It may also be useful to look for any direct KRASV12-PTP interactions by performing KRAS immunoprecipitation followed by PTPome analysis (anti-oxidized PTPome western on pervanadate treated lysates).

Assess the specific phosphorylation sites affected by KRASV12 induced EGFR dephosphorylation.

One of the major questions yet to be resolved in this work is the specificity of KRASV12 mediated phosphatase activity. Substrate specificity has been observed for some PTPs and therefore it is reasonable to believe that KRASV12 activated phosphatases may target specific sites on EGFR. The relative changes in phosphorylation at each EGFR pY site should be assessed using the phosphospecific antibodies utilized in Chapters 1 and 2. It would also be useful to quantify these site specific effects in a number of cell lines demonstrating KRAS associated EGFR dephosphorylation, as differential phosphorylation in a single cell line could have many explanations, including differential SH2 protein expression and EGFR phosphosite protection.

Determine if the activation of any other receptors (or other phosphoproteins) is suppressed by KRAS activation

In addition to EGFR phosphosite specificity we are also interested in the protein specificity of KRASV12 driven phosphatase activity. Is the effect specific to EGFR family proteins, RTKs or a generalized increase in activity? A quantitative phosphospecific MS study of phosphosite abundance in KRASV12 cell identified very few proteins, aside from EGFR, which are dephosphorylation in KRASV12 expressing cells (*186*). However, this work only examined EGF associated phosphorylation. By expressing our KRASV12 construct in cell lines expressing other RTKs (i. e. FGFR, HER2, PDGFR andMET) or other proteins with high levels of phosphorylation (i. e. p130CAS, FAK, paxillin, ABL, SRC) we should be able to assess the extent to which KRAS regulates tyrosine phosphorylation. In addition, it would also be useful to perform FW blotting on lysates from these cell lines using a few representative SH2 probes (e.g. GRB2, CRK, PI3K). This would allow us to detect changes in specific phosphosites on multiply phosphorylated proteins without having to acquire phosphosite specific antibodies for each protein. Moreover, by integrating protein and SH2 specific data we may be able to gain greater insight into the phosphatases involved, the mechanism by which they are being regulated and the signaling pathways which are being targeted.

Quantify PTP activity in KRASV12 expressing cells using PTPome MS-based analyses

Potential mechanisms by which KRASV12 drives EGFR dephosphorylation include altering PTP localization, activity or expression. In order to address the latter two possibilities we propose assessing PTP expression (PTPome) and oxidation-associated PTP inactivation (oxPTPome) in our KRASV12 and mock cell lines using a recently published protocol which exploits an antibody that recognizes the oxidized active site of all PTPs (*193*). Differentially regulated PTPs can be detected by western blotting and specifically identified by MS. If KRASV12 regulated

PTPs are identified, PTP-specific RNAi can then be employed to assess the role of these PTPs in EGFR dephosphorylation.

Contributions

Introduction

The Domainomics section of this Introduction section is adapted from the published review J. Jadwin, M. Ogiue-Ikeda, K. Machida, The application of modular protein domains in proteomics. *FEBS Lett.* **586**, 2586–96 (2012), on which I am the first author. All other text and figures in this section is my own original work.

Chapter 1

The work presented in Chapter 1 is an unpublished manuscript prepared by the authors listed below the title. The manuscript text was prepared by me, B. Mayer, J. Yu and T. Curran, with editing assistance and scientific input from K. Machida, D. Oh and F. White. Construct preparation was performed by M. Ikeda and preliminary experiments were performed by L. Jia and K. Machida. MS phosphopeptides identification was performed by T. Curran. All imaging, image analyses and image quantification (excluding immunofluorescence) were performed by D. Oh. FW blot quantification was performed by K, Machida and I. All other experimental work, analyses including and manuscript preparation including; western, FW, immunodepletion, cell culture, peptide analysis, GO analysis, immunofluorence preparation, manuscript figure layouts, and interactive chart preparation were performed by me. The research was supported by the National Cancer Institute Grant U01CA154966 (to B.J.M. and F.W.) and partly supported by a Quest for CURES (QFC) grant from the Leukemia and Lymphoma Society (to K.M.). We would like to thank Michael Blinov for assistance with the SH2:Mem equations and Ahmed Elmokadem for assistance with immunofluorescence imaging.

Chapter 2

Chapter 2 is an unpublished manuscript prepared by me and edited by B. Mayer. MS peptide identification and quantification was performed by T. Curran. B. Mayer cloned the GCG

construct. All other work presented here was performed by me. I would like to thank J. Yu for productive discussions related to reaction kinetics and modeling.

Chapter 3

Chapter 3 is preliminary data and text for an unpublished manuscript. Text preparation and experiments were all performed by me with editing assistance and scientific input from B. Mayer.

Summary and Future Directions

Summary and Future directions sections were prepared by me with editing and input from B. Mayer. I would also like to thank Adam Lafontaine for useful discussions on modeling and receptor clustering experiments.

All other text, figures and experiments not specifically attributed in the this Contributions sections can be were prepared and performed by me with editing assistance and suggestions provided by by B. Mayer.

Materials and Methods

Chapter 1:

SH2 domain expression constructs: Characteristics of SH2 domain constructs used in this study are provided in Table 1.1. GST-SH2 probes for far-Western were cloned in pGEX backbones as previously described (*38*) and are available through Addgene at https://www.addgene.org/Bruce_Mayer/. SH2-tdEOS clones were generated using Gateway cloning as previously described (*94*).

Antibodies and reagents: Anti-pY immunoblots and immunofluorescence experiments were performed using mouse monoclonal pY100 (CST, #9411). EGFR was detected using rabbit anti-EGFR (SCBT, #sc-31157) and monoclonal rabbit anti-EGFR (CST, #4267) and immunoprecipitated using rabbit anti-EGFR sepharose beads (CST, #5735). p130CAS was detected using mouse anti-p130CAS (BD, #610274) and immunoprecipitated using agarose beads conjugated with mouse anti-p130CAS (SCBT, #sc-20029). GAB1 was detected using mouse anti-GAB1 (Milipore, #06-579) and immunoprecipitated using the same antibody bound to Protein A sepharose CL-4B beads (GE Healthcare, #17-0780). SHCA was detected using rabbit anti-SHC (SCBT, #288) and immunoprecipitated using mouse anti-SHC (Thermo Scientific, #47F4) bound to Protein G agarose beads (Life Technologies, #20398). Phospho-SHCA was detected using rabbit anti-phospho-SHC pY317 (CST, #2431). EGFR phosphosite-specific Westerns were performed using the following antibodies: pY845 (SCBT, #sc-575442), pY974(CST, #2641S), pY992 (CST, #2235P), pY1045 (CST, #2237P), pY1068 (CST, #3777P), pY1086 (CST, #2220S), and pY1173 (CST, #4407S). Antibodies against pY1148 (CST, #4404S and Thermo Scientific, #44-792G) and pY1101 (Abcam, #ab76195) were tested but returned low signal to noise. GRB2 SH2-tdEOS and GST-GRB2-SH2 were detecting using mouse anti-GRB2 SH2 (R&D Systems, #669604). pERK1 and pERK2 were detected using rabbit anti p44/42 pT202/pY204 (CST, #9101S). Pervanadate (PV) was prepared fresh for each

experiment by mixing 1 vol 100 mM $NaVO_4$ with 0.32 vol 30% (w/w) hydrogen peroxide at room temperature (RT) for 30-90 min.

Sample preparation: All experiments were performed using the human squamous-cell carcinoma line A431. For FW and MS experiments, cells were grown to ~70% confluence in standard growth media, starved overnight and stimulated with 25 ng/mL EGF for the appropriate duration. Media was then removed and cells were snap frozen via submersion in liquid N_2 within 3 to 5 s. For imaging experiments, cells were plated onto acid-washed glass bottom dishes (< 50% confluent) (MatTek) and allowed to grow overnight. Cells were transfected with 100-200 ng of DNA using 1-3 µl of Lipofectamine 2000 (Life Technologies) in antibiotic-free Opti-MEM (Life Technologies). After 4 h, transfection media was aspirated, replaced with complete media and cells were allowed to grow overnight. Prior to imaging, DMEM culture media was removed, cells were washed with PBS and kept in phenol red minus media (BrainBits). For pervanadate experiments, pervanadate was prepared fresh, diluted in culture media and added to cell culture dishes at a final concentration of 100-200 µM.

Far-Western and Western blotting: Far-Western blotting was performed as previously described (*38*). *Far-Western blotting:* Briefly, snap frozen cells were thawed on ice at 4°C, lysed and scraped in Kinase Lysis Buffer (150 mM NaCl, 25 mM Tris-HCl pH 7.4, 5 mM ethylene diamine tetraacetic acid (EDTA), 1 mM phenyl methyl sulfonyl fluoride (PMSF), 1% Triton X-100, 10% glycerol, 0.1% sodium pyrophosphate, 10 mM β-glycerophosphate, 10 mM NaF, 5 µg/ml of Aprotinin (Sigma A6279), 50 µM pervanadate), cleared by centrifugation, run on Lithium Dodecyl Sulfate (LDS) PAGE using NuPAGE NOVEX 4-12% gradient gels (Life Technologies, #WG1403A) and transferred overnight onto nitrocellulose membranes. 30 duplicate membranes were created, with all membranes containing positive and negative pY controls. Blots were frozen at -20°C, thawed, blocked with 5% non-fat dry milk in Tris pH 8.0 buffered saline (TBST) for 1 h and blotted for 2 h with 1-5 ng/ml recombinant GST-SH2 labeled with GSH-HRP and

anti-GST-HRP diluted in 5% milk-TBST. Blots were then rinsed with TBST, washed for 20 min with two buffer changes and imaged using ECL (PerkinElmer, #NEL104001EA) on the Kodak Image Station 4000 MM for 1 h. Reprobing was performed using the procedure outlined above, after stripping by washing with 100 mM glycine, pH 2.0 twice for 15 min followed by thorough washing with TBST. Blotting was performed at least twice for all quantified SH2 probes. Quantified far-Westerns met the following conditions: 1) pY-dependent binding (pervanadate-treated positive control sample has greater signal than PTP-treated negative control); 2) reproducibility for replicate membranes; 3) EGF-dependent changes in band intensity; 4) high signal-to-noise (minimal non-specific bands) and 5) high blot quality (minimal background and or signal distortion). 67 probes were initially screened for this study. Of these, the SH2 domains of BLK, BLNK, BRDG1, BRK, CBLB, CIS1, CSK, CTEN, EMT/ITK, FER, FES, FGR, FRK, GRB10, HCK, JAK3, LNK, LYN, P55G(NC), PLCG1(SH2+3), SH2-B, SH3BP2, SHB, SHIP1, SOCS2, SOCS3, SOCS4, SRC, STAT1, STAT3, STAT5A, SYK(NC), and ZAP70(NC) and PTB domains of CTEN, DOK1, FRS2, IRS1, SHCD, TENC1, and SCK, were excluded due to data quality. It should be noted that many of these probes have shown good activity in other systems (*38*, *67*).

Western blotting: Nitrocellulose membranes were prepared in a manner similar to that outlined above for far-Western. Western blotting was performed using standard procedures (blocking in 5% milk, washes in TBST, overnight primary antibody incubation at 4^{o}C, 1 h secondary antibody incubation at RT). Both HRP/ECL and IRDye (680 and 800 nm) labeled secondary antibodies were used. Blots were imaged using the Kodak Image Station 4000 MM or Licor Odyssey imager as appropriate.

Protein identification by immunodepletion: SH2 binding proteins were tentatively identified via their SH2 binding affinities and molecular weights. Band identities were confirmed by repeated immunoprecipitation (3 or 6 times) of EGF-treated lysates using antibody-conjugated beads as

listed above, followed by comparison of pre- and post-depletion lysates using the listed antibodies for immunoblot.

SH2 binding band quantification and hierarchical clustering: Briefly, major protein bands of each blot were auto-detected and quantified using the Carestream MI software (Carestream) after subtracting background. Apparent non-specific bands with low reproducibility in multiple experiments were manually excluded. The molecular weight of each band was estimated using protein ladder lanes on the both sides of the blot. Hierarchical clustering was performed using average linkage and Pearson correlation using Cluster 3.0 software. Heat maps were created using Java Treeview software. For kinetics clustering (Figure 1B), all bands were normalized to the band with the maximum signal and averaged. For relative specificity clustering, bands from each time point (i.e. lane) were normalized to the total signal per lane for all SH2s and averaged. The normalized values obtained for pY blots were then subtracted from these normalized values on a band-by-band basis.

Determination of GRB2 SH2 and pY-EGFR concentration: The approximate cellular GRB2 SH2-tdEOS concentration was determined by comparing GRB2 SH2-tdEOS expression levels in transfected A431 cells with a GST-GRB2-SH2 standard of known concentration via immunoblot. Briefly, A431 cells (30 mm plate) were transfected with 1µg GRB2 SH2-tdEOS using Lipofectamine 2000 (Life Technologies), incubated for 18 h, lysed in KLB, run by LDS-PAGE along with a serial dilution of the GRB2 standard, transferred to nitrocellulose and immunoblotted with anti-GRB2 SH2. GRB2 SH2-tdEOS transfection efficiency and expression level distribution were determined by quantifying epifluorescence images, acquired with a 1.5 min exposure using a CCD camera, using ImageJ. Average cell volume was calculated from three DIC images of non-adhered A431 cells. Cellular pY-EGFR concentrations were calculated by comparing pY immunoblots of EGF-stimulated cells with a tyrosine-phosphorylated protein standard created in our lab and quantified using a malachite green free phosphate quantification

assay (*50*). A431 cells were plated and EGF-stimulated using the protocol performed for FW above. Lysates were run on LDS-PAGE along with a serial dilution of the pY standard and immunoblotted using anti-pY. Signal intensity of the 195 kDa EGFR band and the total pY standard signal were quantified using Image Studio ver. 4.0 (Licor). The number of phosphotyrosines in the standard lanes were calculated and used to create standard curve, which was then used to calculate the concentration of EGFR pY residues after correcting for the number of cells run per lane and the cell volume.

Mass spectrometry sample preparation: Samples were lysed with 8 M urea + 1 mM sodium orthovanadate and protein yield was quantified by BCA assay (Pierce). Samples were reduced with 10 µl 10 mM DTT in 100 mM ammonium acetate pH 8.9 (1 h at 56° C). Samples were alkylated with 75 µl of 55 mM iodoacetamide in 100 mM ammonium acetate pH 8.9 (1 h at RT). 1 ml of 100 mM ammonium acetate and 10 µg of sequencing grade trypsin (Promega, #V5111) were added and digestion proceeded for 16 h at RT. Samples were acidified with 125 µl of trifluoroacetic acid (TFA) and desalted with C18 spin columns (ProteaBio, #SP-150). Samples were lyophilized and subsequently labeled with iTRAQ 8plex (AbSciex) per manufacturer's directions. iTRAQ Channels were designated as follows: 0s-113, 10s-114, 30s-115, 1m-116, 1.5m-117, 3m-118, 10m-119, 30m-121.

Immunoprecipitation: 70 µl protein-G agarose beads (Calbiochem, #IP08) were rinsed in 400 µl IP Buffer (100 mM Tris, 0.3% NP-40, pH 7.4) and charged for 8 h with three pY-specific antibodies: (12 µg 4G10 (Millipore), 12 µg PT66 (Sigma), and 12 µg PY100 (CST)) in 200 µl IP Buffer. Beads were rinsed with 400 µl of IP Buffer. Labeled samples were resuspended in150 µl iTRAQ IP Buffer (100 mM Tris, 1% NP-40, pH 7.4) + 300 µl milliQ water and pH was adjusted to 7.4 (with 0.5 M Tris-HCl pH 8.5). Sample was added to charged beads for overnight incubation. Supernatant was removed and beads were rinsed 3 times with 400 µl Rinse Buffer (100mM

Tris-HCl, pH 7.4). Peptides were eluted in 70 µl of Elution Buffer (100 mM glycine, pH 2) for 30 min at RT.

Immobilized metal affinity chromatography (IMAC) purification: A fused silica capillary (FSC) column (200 µm ID x 10 cm length) was packed with POROS 20MC beads (Applied Biosystems, #1-5429-06). IMAC column was prepared by rinsing, at approximately 10 µl/min, with solutions in the following order: 100 mM EDTA pH 8.9 (10 min), H_2O (10 min), 100 mM $FeCl_3$ (20 min), 0.1% acetic acid (10 min). IP elution was loaded at a flow rate of 2 µl/min. The column was rinsed with 25% acetonitrile, 1% acetic acid, and 100 mM NaCl (10 min) and 0.1% acetic acid (10 min), both at 10 µl/min. Peptides were eluted with 50 µl 250 mM NaH_2PO_4 at 2 µl/min and collected in an autosampler vial. Eluent was acidified with 2 µl of 10% TFA prior to loading.

Liquid chromatography mass spectrometry: Acidified IMAC eluent was loaded onto an Acclaim PepMap 100 precolumn (Thermo Scientific, #164705) using an EASY-nLC 1000 (Thermo Scientific). Peptides were analyzed on a 1 h gradient from 100% A (0.1% formic acid) to 100% B (0.1% formic acid, 80% acetonitrile) spraying through a 50 cm analytical column (New Objective, #PF360-50-10-N-5) packed with 3 µm beads (YMC America, #AQ12S03).

MS data analysis: Thermo .RAW files were searched with MASCOT v2.4 using Proteome Discoverer (v1.4). Peptides that appeared in at least two biological replicates were included if their MASCOT scores exceeded 15 and they were designated as medium or high confidence by Proteome Discoverer. Normalized iTRAQ values for each biological replicate were averaged to produce the final dataset (Table 1.2, *MS Data*). Error is represented as the standard deviation (observed in all three biological replicates) or the average deviation (observed in two biological replicates).

Gene ontology, sequence motif and EGF dependence analysis: Gene ontology (GO) analyses were performed using STRING 9.1. Biological Processes GO terms were queried and significant terms were identified as those with *p*-values $\leq$ 0.05 after Bonferroni correction for multiple comparisons. General gene ontology used in the Venn diagram were obtained by GO semantic clustering using REVIGO (*129*). Sequence motif analysis was performed using the PhosphoSitePlus sequence logo generator using the frequency change algorithm and pY background (*62*, *201*). Previous documentation of EGF dependence was performed by comparing our data set with the PhosphositeSitePlus EGF-associated phosphorylation data set as of May 2015.

Apical and basal membrane pY quantification: A431 cells were grown to ~30% confluence in glass bottom 30 mm dishes and starved overnight in 0.1% FBS, 1% pen/strep DMEM. Cells were stimulated with 25 ng/mL EGF, quickly washed once in 4^{o}C PBS and fixed with 4% paraformaldehyde on ice for 20 min (~15 s media-to-fix time). Cells were then washed with PBS three times for 5 min each, permeabilized in 1.5% Triton-X PBS for 10 min and washed with PBS three times for 10 min each. Dishes were then blocked in 1.5% BSA PBS for 60 min at 4^{o}C, incubated with anti-pY (1:500 in blocking buffer) for 2 h while rocking at 4^{o}C, washed three times for 5 min in PBS, incubated with anti-mouse IgG conjugated with Alexfluor 594 at RT for 2 h, washed three times for 5 min with PBS, covered with Fluoromount-G and a glass coverslip. Confocal microscopy was used to capture z-stacks through cells with significant apical-basal separation. Time points (min) and number of cells quantified are as follows: 0 (n=7), 0.08 (n=20), 0.17 (n=20), 0.5 (n=29), 1 (n=8), 2 (n=21), 4 (n=20), 6 (n=20), 10 (n=20) and 15 (n=20). For each cell, multiple line scans through the z-stack were quantified and averaged in both the x and y planes and used to plot the change in apical and basal phosphorylation in response to EGF.

TIRF microscopy and tdEOS photactivation: A manually controllable home-built total internal reflection (TIR) microscopy set-up using an Olympus IX81 microscope was used for all single molecule and time-lapse imaging. For cytoplasmic or top membrane illumination, TIR angle was switched to epi illumination mode using a step motor (Thor Labs, Inc). A co-aligned 488 nm Argon ion laser (Melles Griot) was used to excite the population of non-photoactivated tdEOS proteins with a green channel filter set. tdEOS photoactivation was performed using a 405 nm diode laser (Cube laser system, Coherent). A 532 nm DPSS crystal diode laser (Crystal Laser Inc) was used to excite photoactivated tdEOS proteins using a red channel filter set. Fluorescence emission was collected using a thermoelectric cooled EM CCD camera (PhotonMax, Roper Scientific). Binding and dissociation rate measurements were performed by imaging the photoactivated population under red channel illumination (532 nm). The green channel (488 nm) was alternatively used to confirm both longer time recruitment kinetics and clustering processes as a function of EGF stimulation time.

Single particle tracking: Time-lapse single molecule videos were used to track individual SH2 domains and create individual molecule trajectories. Individual diffraction-limited fluorescence dots displaying single step fluorescence photobleaching were tracked using the standard cross-correlation calculation between raw data and Gaussian PSF kerner. The centroid position of cross-correlated particles was determined and used to define particle x-y coordinates. Tracking over subsequent images was performed based on the nearest neighbor method. The scan area used during position tracking of individual molecules from frame to frame was chosen based on the estimated diffusion rate of the molecule. To avoid tracking of non-specific fluorescence dots appearing within the scan area, we reduced the density of fluorescence particles to 0.1 – 0.05 /μm^2 using red channel illumination of tdEOS proteins. Dimers or small clusters were filtered out by size thresholding of correlated particles. Diffusive cytoplasmic-like fluorescence dots

occurring in single frames were filtered out by intensity-based thresholding of cross correlation images.

Trajectory analysis: Mean square displacement (MSD) was calculated as a function of time lag ($n \cdot \Delta t$, n= integer, Δt=frame interval) using all trajectories over 4 frames long. Individual MSD-Δt curves were fitted to the normal diffusion equation at a time scale of n = 2 to 3 (*202*). Diffusion coefficients of individual trajectories were used to create distribution histograms from which mean diffusion coefficients were determined.

Fluorescence localization images: SH2-specific fluorescence images were prepared from multiple images (50 - 200 frames) using time-lapse videos recorded at 0.1 Hz. Image sequences were z-stack projected using the standard deviation projection method in ImageJ (NIH) to create an image depicting the spatial distribution of single molecules. Although the resolution of SH2 domain clusters generated by this method is relatively lower when compared to that of reconstituted sptPALM images, it allowed us to quickly determine the localization patterns of SH2 molecules on the membrane.

SH2 apparent on- and off-rate determination: The binding rate of GRB2 SH2 was determined using multiple sptPALM image sets of newly recruited GRB2 SH2 obtained at 0.1 Hz for 0.8min throughout the EGF stimulation time-course. Counting of only newly emerged molecules in each video image sequence was performed based on tracking molecular ID. The apparent on-rate (γ_{on}) of GRB2 SH2 was calculated by dividing the number of newly binding molecules by the total acquisition time (0.8 min). The apparent membrane dissociation rate constant (λ_{off}) was determined using the same video dataset. The λ_{off} of individual SH2 molecules was measured using their trajectory lengths and is defined as the reciprocal of the SH2 membrane dwell time. As the membrane dwell time of all SH2 modules tested here is shorter than the photobleaching time tdEOS, λ_{off} values were calculated after compensating tdEOS photobleaching (*94*). All

time measures represent time after EGF stimulation. λ_{off} values presented in Table 1 were measured at 40-60 min after EGF stimulation.

Determination of clustered and non-clustered apparent on-rates: GRB2 SH2 clusters were identified by auto-thresholding sptPALM images (ImageJ). Thresholded binary and raw sptPALM images were used to characterize the number of clusters, cluster size and intensity (ImageJ). Maximum intensity measurements were used to determine the number of molecules in a cluster. The apparent binding rate of GRB2 SH2 (γ_{on}) was determined by dividing the total number of molecules by the cluster size and acquisition time. The average γ_{on} was derived from the distribution of individual γ_{on} for each movie. To obtain the number of newly binding GRB2 SH2 molecules in non-clustered regions, we subtracted the number of molecules in clusters from the number of molecules identified in the region of interest.

Chapter 2

Plasmids

The tdEOS GRB2 SH2 construct was cloned as previously described in Chapter 1 (*94*). The full length GRB2 and GCG constructs were cloned into pEBB (*173*). The GCG construct was creating by replacing the GRB2 SH2 domain (AA 58-159) with the chicken CRK SH2 domain (AA 1-124).

Antibodies and reagents

Anti-pY immunoblot experiments used mouse monoclonal pY100 (Cell Signaling Technology [CST], #9411). EGFR was detected using rabbit anti-EGFR (Santa Cruz Biotechnology [SCBT], #sc-31157). EGFR phosphosite-specific Westerns were performed using the following antibodies: pY845 (SCBT, #sc-575442), pY974 (CST, #2641S), pY992 (CST, #2235P), pY1045 (CST, #2237P), pY1148 (CST, #4404S), pY1068 (CST, #3777P), pY1086 (CST, #2220S), and

pY1173 (CST, #4407S). pERK1 and pERK2 were detected using rabbit anti p44/42 pT202/pY204 (CST, #9101S). Full length GRB2 and GRB2-SH2 were blotted using the GRB2-SH2 specific antibody mouse anti-GRB2 SH2 (R&D Systems, #669604). The GCG and full length GRB2 were detected using a rabbit antiserum raised against GST-GRB2 produced for our laboratory. Pervanadate (PV) was prepared fresh for each experiment by mixing 1 vol 100 mM Na_3VO_4 with 0.32 vol 30% (w/w) hydrogen peroxide at room temperature (RT) for 30–90 min.

.

.

Cell culture

COS1 cells, a fibroblast-like an African Green Monkey kidney cell line, were cultured in complete DMEM media (10% fetal bovine serum, 1% penicillin/streptomycin). Overnight serum starvation was performed by aspirating complete DMEM and replacing it with starvation DMEM (0% fetal bovine serum and 1% penicillin/streptomycin) and incubating for 6-14 hrs.

Plasmid expression

The tdEOS-GRB2 SH2, full length GRB2, GCG and empty vector constructs were exogenously expressed in COS1 cells using Lipofectamine 2000. 10µg of each construct was used for all Western experiments except for the GRB2 dose dependence experiments. For MS experiments 30 µg of each construct was used per 10-cm dish.

GRB2 SH2 concentration quantification

The approximate cellular tdEOS-GRB2 SH2 concentration was determined by comparing tdEOS-GRB2 SH2 expression levels in transfected COS1 cells with a GST-GRB2 SH2 standard of known concentration via immunoblot. Briefly, COS1 cells (30 mm plate) were transfected with 10 µg tdEOS-GRB2 SH2 using Lipofectamine 2000 (Life Technologies), incubated for 18 hr,

lysed in KLB, run by LDS-PAGE along with a serial dilution of the GRB2 standard, transferred to nitrocellulose and immunoblotted with anti-GRB2 SH2. The average cell volume was calculated from three DIC images of non-adhered cells.

Determination of site-specific relative EGFR phosphorylation

To determine to relative phosphorylation of each EGFR pY site, a maximally phosphorylated pY-EGFR control was created by treating COS1 cells with pervanadate (1:1000 dilution) and 200 ng/ml EGF for 40 min. Lysates from this control were then diluted 1:10, run alongside experimental lysates on LDS-PAGE, transferred to nitrocellulose, Western blotted using the site-specific anti-pY-EGFR antibodies, and quantified by using the LiCOR system. The percent of maximum phosphorylation for each site was then determined by comparing its specific western signal to that of the maximally phosphorylated standard.

EGFR phosphorylation/dephosphorylation rate quantification

To determine the basal EGFR phosphorylation rate, COS1 cells were starved overnight. The starvation media was then replaced with starvation media containing 100 μM pervanadate and cells were flash frozen in liquid N_2 after rapid media aspiration at the time points shown. To determine the rate of EGF-induced phosphorylation, the cells were treated with 2.5 ng/mL EGF. After 5 min, the EGF media was then replaced with starvation media containing 2.5 ng/mL EGF and 100 μM pervanadate, and the cells were flash frozen in liquid N_2 as above. Cells from each time course were then lysed in KLB and run on LDS-PAGE along with an increasing concentration of the malachite green-quantified pY-ABL standard (*50*), transferred to nitrocellulose and blotted using anti-pY100. The absolute rate of EGFR phosphorylation was then determined using a pY-ABL standard curve and determining the initial slope following pervanadate treatment (Fig. 5B, C).

To determine the rate of dephosphorylation, the cells were treated with 2.5 ng/mL EGF. After 5 min, the EGF media was then replaced with starvation media containing 2.5 ng/mL EGF and 10 μM Erlotinib and the cells were flash frozen in liquid N_2 at the time points shown. The rate of EGFR dephosphorylation was then determined using as described above for the phosphorylation rate (Fig. 5D).

Mass spectrometry

Cells were transfected with 10μg of the constructs listed, stimulated with 2.5ng/mL for 10 min and flash froze in liquid N2. Cells were lysed with 8 M urea + 1 mM sodium orthovanadate and protein yield was quantified by BCA assay (Pierce). Samples were reduced with 10 μl 10 mM DTT in 100 mM ammonium acetate pH 8.9 (1 hr at 56°C). Samples were alkylated with 75 μl of 55 mM iodoacetamide in 100 mM ammonium acetate pH 8.9 (1 hr at RT). 1 ml of 100 mM ammonium acetate and 10 μg of sequencing grade trypsin (Promega, #V5111) were added and digestion proceeded for 16 hr at RT. Samples were acidified with 125 μl of trifluoroacetic acid (TFA) and desalted with C18 spin columns (ProteaBio, #SP-150). Samples were lyophilized and subsequently labeled with iTRAQ 8plex (AbSciex) per manufacturer's directions.

Protected consensus sequence determination

The consensus sequence of protected sites was performed using MS data listed in Table 2.1. The abundance of each site in empty vector and SH2-containing protein expressing cells was compared. Those sites whose abundance was greater than the combined standard deviation of the abundance in each experimental point were determined to be significantly enhanced by SH2 protein expression. Significantly enhanced pY-sites were then compiled for each experimental treatment and the abundance of each significantly enhance peptide was weighted using its fold increase over its empty vector control. This weighted list was then used create an amino acid sequence LOGO plot using the PhosphositePlus LOGO generator by applying the Frequency

Change algorithm and the phospho Tyr background settings (*62*, *201*). The consensus LOGO for GRB2 was determined using only those pY peptides whose abundance was significantly enhanced under all four treatments (i.e. GRB-SH2, GRB FL, GRB2-SH2+EGF, GRB2 FL+EGF). For the CRK consensus sequence LOGO, pY-sites enhanced by both treatments, with and without EGF, were used. Phosphosite LOGOs for the data from Tinti et al were created using sequences from peptides whose binding signal exceeded the average signal by more than two SDs (Z score > 2) (*36*). The PhosphoSitePlus LOGO generator was also used and no weighting was applied.

Computational Reaction Modeling

All reaction modeling was performed using the Virtual Cell version 5.3.17 using the Combined Stiff Solver (IDA/CVCODE). All reaction parameters were approximated as listed in Table 2.2. Briefly, total cell volume of COS1 cells was calculated from an average diameter of tyrpsinized cells (~20μm). The plasma membrane reaction area was approximated as a hollow sphere 200nm deep, the nucleus was approximated as an internal sphere with a diameter of 10μm and the cytoplasmic volume consisted of the remaining volume. The extracellular volume was set at 4 ml as in our experiments. EGF concentrations were those used in our experiments. EGF-EGFR forward and reverse bindings constant values were obtained from Berkers et al 1991 and French et al 1995 (*99*, *100*). EGFR receptor numbers were estimated using previous estimates of EGFR expression in A431 cells and anti-EGFR Western blots of COS1 and A431 cell lysates. The EGFR K_f values was determined using the initial phosphorylation rate of EGFR in EGF stimulated or non-stimulated COS1 cells treated with pervanadate. The PTP V_{max} and K_m were obtained by fitting data to the Michaelis-Menton function using OriginPro 2017. The GRB2-EGFR k_{off} value was taken from Oh et al 2012 (*94*). The GRB2-EGF K_d and k_{on} values were determined by fitting experimental data to modeling data using variable k_{on} values.

Chapter 3

Antibodies

pY, pY EGFR and EGFR were detected as described in Chapter 1 methods. KRAS was detected using the mouse anti-KRAS (Abnova, #H00003845-M01).

Cell Culture

H2172, H322 and HCC827 lung adenocarcinoma (NSCLC) cell lines were all obtained from Dr. Eric Haura of the H. Lee Moffit Cancer as part for a previously published work, who also approved their use in these experiments (*67*). Cells were cultured in complete RPMI-1640 (10% fetal bovine serum, 1% penicillin/streptomycin). For EGF or pervanadate stimulation experiments cells were starved overnight (8-16 hrs) in starvation RPMI-1680 (0.1% fetal bovine serum, 1% penicillin/streptomycin).

Retrovirus production and retroviral transduction

VSV-G psudeotyped retrovirus was used create KRASV12 expressing NSCLC cell lines. 293T cells were grown to 70-80% confluency in 10cm cell culture dishes and transfected with retrovirus packaging plasmids (pMD.gag-pol –12ug and HCMV - 4ug) and pBABE KRASV1-24ug using Lipofectamine 2000. Media from these cells was collected beginning 24 hrs post – transfection every 4 hrs for 36hrs. Media was then filtered through a 0.45um syringe driven filter. To create the KRASV12 cell lines, cells were incubated in the virus containing for two consecutive 24hr periods. Infected cells were then grown for an additional day and passed as necessary. KRASV12 expressing cells were then selected for by passing them into complete RPMI-1640 containing 10ug/mL puromycin and grown for a period of 7day or until non-infected controls were had completely detached. Cells were then transferred to non-selection media, grown for an additional week before KRASV12 expression was verified by western-blot.

Surface protein biotinylation and pulldown

Cell surface EGFR was isolated using the Pierce Cell Surface Biotinylation Kit (#89881). Briefly, cells were grown in 10cm culture dishes to 85-95% confluence. Cells were then transferred to the cold room, culture media was removied and cells were washed twice with 10mL of cold PBS. Wash PBS was removed and 10mL of 1mg/mL Sulfo-NHS-SS-Biotin (prepared fresh) was added to each cell culture dish. Cells were set to rock gently at 4^{o}C for 30 min. Next the biotinylation reaction was quenched by washing the cells three times with Quenching Solution (0.5% BSA, 50mM Glycine), gently scraped into a 15 mL conical tube (Falcon), spun down at 500xg, washed with TBS and spun again 500xg. Cells were then lysed in 500uL RIPA lysis buffer (150 mM NaCl, 1.0% IGEPAL® CA-630, 0.5% sodium deoxycholate, 0.1% SDS, and 50 mM Tris, pH 8.0) by rotation at 4^{o}C for 1 hr. 20-50 ul of lysate were then taken to be used as prebinding lysate for comparison by Western. Protein concentrations were then determined via Bradford Assay and lysate concentration were normalized using RIPA. Lysates were then mixed with Neutravidin Beads (Pierce) in a 3:1 lysate to bead ratio and allowed to rotate for 2 hrs at 4^{o}C. Finally beads were spun down at 2000 rpm in a microcentrifuge, washed 2x's with high salt RIPA (500mM NaCl) and once with normal RIPA (150mM) and finally boiled in 6x laemelli buffer prior to analysis by SDS PAGE.

Phosphorylation rate quantification

The basal phosphorylation rate of cells with and without KRASV12 expression was performed as described in the Chapter 2 method *"EGFR phosphorylation/dephosphorylation rate quantification".*

Active-RAS pulldown

Active-RAS pulldown was preformed as previously describe by Knaus et al (*203*). Briefly, RAS binding beads were prepared by incubating 60 ug of recombinant GST-RAF1 RBD (AA 1-149, Ras Binding Domain, Addgene Plasmid 13338) with GSH-sepharose beads (GE Healthcare) overnight at 4^{o}C. NSCLC cell lines were lysed with GTPase lysis buffer (50mM Tris-HCL, 10mM $MgCL_2$, 200mM NaCl, 1% (v/v) Triton-X100, 5% glycerol (v/v), 1mM $NaVO_3$, 1mM PMSF, 1mg/mL aprotinin, pH 7.5) for 10 min on ice and cleared by centrifugation at 13,200 rpm for 10 min. Lysate protein concentration was determined by Bradford Assay and 500mg of lysate was incubated with 40ul of 1:1 with GST-RAF1 RBD GSH beads and incubated at 4^{o}C for 1hr. Beads were then washed 3x's with GTPase Washing Buffer (25mM Tris-HCL, 30mM $MgCl_2$, 40mM NaCl, 1% Triton-X100 (v/v) and once with GTPase Washing Buffer without Triton. Beads were then spun down and boiled in 6x laemelli buffer prior to analysis by SDS PAGE.

Adherent cell growth (MTT) assay

HCC827 cell lines were plated at 1000 cells/well in a cell culture coated 96 well plate and allowed to attach for 24 hrs. Growth was assay by MTT Assay every day for five days. For each time point 10uL of MTT reagent was added to the wells to be measured and allowed to incubate in the cell culture incubator for 3 hrs. 100mL of DMSO was then added to each well, to soluabalize the reduced MTT reagent crystals and absorbance of each well was measured at 570 nm. Background absorbance was subtracted using empty wells for each day. Absorbance values were then normalized in proportion to the Day 1 value for each cell line and plotted in Excel. Three technical replicates were measured for each cell line.

Anchorage-independent growth assay

Non-cell culture coated 6cm dishes were coated with 2 mL of a base layer agar (10% FBS , 0.54% Bacto-agar, 1x RPMI-1640). $1x10^4$ cells were then trypsinized and resuspended in a cell growth agar layer (10% FBS, 0.396% Bacto-agar, 1x RPMI-1640). 1mL of complete RPMI-1640

was then added on above the cell growth agar. Cells were allowed to grow at 37^{o}C in 5% CO_2 for approx. 3 weeks with 1mL of new complete media was added every 4 days. Cells were then stained with Crystal Violet (0.005%), imaged using the Kodak Image Station (see Chapter 1 methods) and quantified using ImageJ.

Works Cited

1. W. Lim, B. J. Mayer, T. Pawson, *Cell Signaling: principles and mechanisms* (Garland Science, ed. 1st, 2014).
2. W. a Lim, T. Pawson, Phosphotyrosine signaling: evolving a new cellular communication system. *Cell*. **142**, 661–7 (2010).
3. M. a. Lemmon, J. Schlessinger, Cell signaling by receptor tyrosine kinases. *Cell*. **141**, 1117–1134 (2010).
4. P. Blume-Jensen, T. Hunter, Oncogenic kinase signalling. *Nature*. **411**, 355–365 (2001).
5. S. R. Hubbard, J. H. Till, P t k s f (2000).
6. S. R. Hubbard, The insulin receptor: Both a prototypical and atypical receptor tyrosine kinase. *Cold Spring Harb. Perspect. Biol.* **5** (2013), doi:10.1101/cshperspect.a008946.
7. Y. Huang *et al.*, The c-Abl tyrosine kinase regulates actin remodeling at the immune synapse. *Blood*. **112**, 111–119 (2008).
8. K. Chen *et al.*, Mechanotransduction in Response to Shear Stress. *J. Biol. Chem.* **274**, 18393–18400 (1999).
9. K. G. Peters, C. De Vries, L. T. Williams, Vascular endothelial growth factor receptor expression during embryogenesis and tissue repair suggests a role in endothelial differentiation and blood vessel growth. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 8915–8919 (1993).
10. J. Toubiana *et al.*, Src-family-tyrosine kinase Lyn is critical for TLR2-mediated NF- B activation through the PI 3-kinase signaling pathway. *Innate Immun.*, 1–13 (2015).
11. J. Ursini-Siegel *et al.*, Receptor tyrosine kinase signaling favors a protumorigenic state in breast cancer cells by inhibiting the adaptive immune response. *Cancer Res.* **70**, 7776–7787 (2010).
12. K. Machida, M. Khenkhar, P. Nollau, Deciphering Phosphotyrosine-Dependent Signaling Networks in Cancer by SH2 Profiling. *Genes Cancer*, 353–361 (2012).
13. A. Levitzki, Tyrosine Kinase Inhibitors: Views of Selectivity, Sensitivity, and Clinical Performance. *Annu. Rev. Pharmacol. Toxicol.* **53**, 121008173501002 (2012).
14. A. De Luca, M. R. Maiello, A. D'Alessio, M. Pergameno, N. Normanno, The RAS/RAF/MEK/ERK and the PI3K/AKT signalling pathways: role in cancer pathogenesis and implications for therapeutic approaches. *Expert Opin. Ther. Targets*. **16**, S17–S27 (2012).
15. G. S. Martin, The road to Src. *Oncogene*. **23**, 7910–7917 (2004).
16. T. Hunter, B. M. Sefton, Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. *Proc. Natl. Acad. Sci. U. S. A.* **77**, 1311–1315 (1980).
17. S. R. Hubbard, Structural analysis of receptor tyrosine kinases. *Prog. Biophys. Mol. Biol.* **71**, 343–358 (1999).
18. B. J. Mayer, Perspective: Dynamics of receptor tyrosine kinase signaling complexes. *FEBS Lett.* **586**, 2575–9 (2012).
19. H. Waterman *et al.*, A mutant EGF-receptor defective in ubiquitylation and endocytosis unveils a role for Grb2 in negative signaling. *EMBO J.* **21**, 303–13 (2002).
20. D. Wheeler, Y. Yarden, Eds., *Receptor Tyrosine Kinases: Structure, Functions and Role in Human Disease* (Humana Press, New York, 201AD).

21. Z. Songyang *et al.*, Catalytic specificity of protein-tyrosine kinases is critical for selective signalling. *Nature*. **373** (1995), pp. 536–539.

22. J. Y. Gouzi, C. Moog-Lutz, M. Vigny, N. Brunet-de Carvalho, Role of the subcellular localization of ALK tyrosine kinase domain in neuronal differentiation of PC12 cells. *J. Cell Sci.* **118**, 5811–5823 (2005).

23. N. K. Tonks, Protein tyrosine phosphatases: from genes, to function, to disease. *Nat. Rev. Mol. Cell Biol.* **7**, 833–46 (2006).

24. F. Böhmer, S. Szedlacsek, L. Tabernero, A. Östman, J. Den Hertog, Protein tyrosine phosphatase structure-function relationships in regulation and pathogenesis. *FEBS J.* **280**, 413–431 (2013).

25. R. Karisch, B. G. Neel, Methods to monitor classical protein-tyrosine phosphatase oxidation. *FEBS J.* **280**, 459–475 (2013).

26. T.-C. Meng, T. Fukada, N. K. Tonks, Reversible Oxidation and Inactivation of Protein Tyrosine Phosphatases In Vivo. *Mol. Cell*. **9**, 387–399 (2002).

27. A. J. Garton, M. R. Burnham, A. H. Bouton, N. K. Tonks, Association of PTP-PEST with the SH3 domain of p130 cas ; a novel mechanism of protein tyrosine phosphatase substrate recognition (1997).

28. S. Dadke, A. Kusari, J. Kusari, Phosphorylation and activation of protein tyrosine phosphatase (PTP) 1B by insulin receptor. *Mol. Cell. Biochem.* **221**, 147–154 (2001).

29. S. Wälchli *et al.*, Probing protein-tyrosine phosphatase substrate specificity using a phosphotyrosine-containing phage library. *J. Biol. Chem.* **279**, 311–8 (2004).

30. a J. Flint, T. Tiganis, D. Barford, N. K. Tonks, Development of “substrate-trapping” mutants to identify physiological substrates of protein tyrosine phosphatases. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 1680–1685 (1997).

31. G. Fan, S. Aleem, M. Yang, W. T. Miller, N. K. Tonks, *J. Biol. Chem.*, in press, doi:10.1074/jbc.M115.651703.

32. J. A. Jadwin, M. Ogiue-Ikeda, K. Machida, The application of modular protein domains in proteomics. *FEBS Lett.* **586**, 2586–96 (2012).

33. B. a Liu *et al.*, The human and mouse complement of SH2 domain proteins-establishing the boundaries of phosphotyrosine signaling. *Mol. Cell*. **22**, 851–68 (2006).

34. T. Pawson, G. D. Gish, P. Nash, SH2 domains, interaction modules and cellular wiring. *Trends Cell Biol.* **11**, 504–511 (2001).

35. Z. Songyang *et al.*, Specific motifs recognized by the SH2 domains of Csk, 3BP2, fps/fes, GRB-2, HCP, SHC, Syk, and Vav. *Mol. Cell. Biol.* **14**, 2777–2785 (1994).

36. M. Tinti *et al.*, The SH2 Domain Interaction Landscape. *Cell Rep.* **3**, 1293–1305 (2013).

37. B. a Liu, B. W. Engelmann, P. D. Nash, The language of SH2 domain interactions defines phosphotyrosine-mediated signal transduction. *FEBS Lett.* **586**, 2597–605 (2012).

38. K. Machida *et al.*, High-throughput phosphotyrosine profiling using SH2 domains. *Mol. Cell*. **26**, 899–915 (2007).

39. M. B. Yaffe, Phosphotyrosine-binding domains in signal transduction. *Nat. Rev. Mol. Cell Biol.* **3**, 177–186 (2002).

40. T. Kisseleva, S. Bhattacharya, J. Braunstein, C. W. Schindler, Signaling through the JAK / STAT pathway , recent advances and future challenges. **285**, 1–24 (2002).

41. K. Ikeda *et al.*, Nuclear localization of Lyn tyrosine kinase mediated by inhibition of its kinase activity. *Exp. Cell Res.* **314**, 3392–3404 (2008).

42. L. B. Kleiman, T. Maiwald, H. Conzelmann, D. A. Lauffenburger, P. K. Sorger, Rapid phospho-turnover by receptor tyrosine kinases impacts downstream signaling and drug binding. *Mol. Cell*. **43**, 723–37 (2011).

43. R. B. Jones, A. Gordus, J. a Krall, G. MacBeath, A quantitative protein interaction network for the ErbB receptors using protein microarrays. *Nature*. **439**, 168–74 (2006).

44. C. S. Monast, C. M. Furcht, M. J. Lazzara, Computational analysis of the regulation of EGFR by protein tyrosine phosphatases. *Biophys. J.* **102**, 2012–2021 (2012).

45. N. Nair, R. J. Davis, H. L. Robinson, Protein tyrosine kinase activities of the epidermal growth factor receptor and ErbB proteins: correlation of oncogenic activation with altered kinetics. *Mol. Cell. Biol.* **12**, 2010–2016 (1992).

46. X. Zhang *et al.*, Inhibition of the EGF receptor by binding of MIG6 to an activating kinase domain interface. *Nature*. **450**, 741–744 (2007).

47. F. De Smet, A. Christopoulos, P. Carmeliet, Allosteric targeting of receptor tyrosine kinases. *Nat. Biotechnol.* **32**, 1113–1120 (2014).

48. M. a Lemmon, J. Schlessinger, K. M. Ferguson, The EGFR Family : Not So Prototypical Receptor Tyrosine Kinases. *Cold Spring Harb Perspect Biol*. **6**, 1–18 (2014).

49. M. Hsieh, S. Yang, M. A. Raymond-Stinz, J. S. Edwards, B. S. Wilson, Spatio-temporal modeling of signaling protein recruitment to EGFR. *BMC Syst. Biol.* **4**, 57 (2010).

50. C. M. Thompson, L. R. Bloom, M. Ogiue-Ikeda, K. Machida, SH2-PLA: a sensitive in-solution approach for quantification of modular domain binding by proximity ligation and real-time PCR. *BMC Biotechnol.* **15**, 60 (2015).

51. I. a Yudushkin *et al.*, Live-cell imaging of enzyme-substrate interaction reveals spatial regulation of PTP1B. *Science*. **315**, 115–119 (2007).

52. F. G. Haj, P. J. Verveer, A. Squire, B. G. Neel, P. I. H. Bastiaens, Imaging sites of receptor dephosphorylation by PTP1B on the surface of the endoplasmic reticulum. *Science*. **295**, 1708–1711 (2002).

53. T. Akiyama *et al.*, Genistein, a specific inhibitor of tyrosine-specific protein kinases. *J. Biol. Chem.* **262**, 5592–5595 (1987).

54. G. R. Yan *et al.*, Global phosphoproteomic effects of natural tyrosine kinase inhibitor, genistein, on signaling pathways. *Proteomics*. **10**, 976–986 (2010).

55. G. Carpenter, S. Cohen, Epidermal growth factor. *J. Biol. Chem.* **265**, 7709–12 (1990).

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

57. a Wells, J. M. Bishop, Genetic determinants of neoplastic transformation by the retroviral oncogene v-erbB. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 7597–7601 (1988).

58. R. Roskoski, The ErbB/HER family of protein-tyrosine kinases and cancer. *Pharmacol. Res.* **79**, 34–74 (2014).

59. X. Zhang, J. Gureasko, K. Shen, P. a Cole, J. Kuriyan, An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell*. **125**, 1137–49 (2006).

60. N. F. Endres *et al.*, Conformational coupling across the plasma membrane in activation of

the EGF receptor. *Cell*. **152**, 543–56 (2013).

61. N. Gotoh, A. Tojo, M. Hino, Y. Yazaki, M. Shibuya, WITHIN THE KJNASE DOMAIN IS NOT REQUIRED FOR. *Biochem. Biophys. Res. Commun.* **186**, 768–774 (1992).

62. P. V Hornbeck *et al.*, PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* **40**, D261-70 (2012).

63. Z. Lu, G. Jiang, P. Blume-jensen, T. Hunter, Epidermal Growth Factor-Induced Tumor Cell Invasion and Metastasis Initiated by Dephosphorylation and Downregulation of Focal Adhesion Kinase Epidermal Growth Factor-Induced Tumor Cell Invasion and Metastasis Initiated by Dephosphorylation and Downregulat (2001), doi:10.1128/MCB.21.12.4016.

64. R. B. Birge, C. Kalodimos, F. Inagaki, S. Tanaka, Crk and CrkL adaptor proteins: networks for physiological and pathological signaling. *Cell Commun. Signal.* **7**, 13 (2009).

65. P. Defilippi, P. Di Stefano, S. Cabodi, p130Cas: a versatile scaffold in signaling networks. *Trends Cell Biol.* **16**, 257–63 (2006).

66. N. Normanno *et al.*, Epidermal growth factor receptor (EGFR) signaling in cancer. *Gene*. **366**, 2–16 (2006).

67. K. Machida *et al.*, Characterizing tyrosine phosphorylation signaling in lung cancer using SH2 profiling. *PLoS One*. **5**, e13470 (2010).

68. G. da Cunha Santos, F. a Shepherd, M. S. Tsao, EGFR mutations and lung cancer. *Annu. Rev. Pathol.* **6**, 49–69 (2011).

69. H. Masuda *et al.*, Role of epidermal growth factor receptor in breast cancer. *Breast Cancer Res. Treat.* **136**, 331–345 (2012).

70. A. M. Krasinskas, EGFR Signaling in Colorectal Carcinoma. *Patholog. Res. Int.* **2011**, 932932 (2011).

71. P. Workman, P. a. Clarke, Resisting Targeted Therapy: Fifty Ways to Leave Your EGFR. *Cancer Cell*. **19**, 437–440 (2011).

72. A. Bardelli, P. a Jänne, The road to resistance: EGFR mutation and cetuximab. *Nat. Med.* **18**, 199–200 (2012).

73. J. M. Rojas, J. L. Oliva, E. Santos, Mammalian son of sevenless Guanine nucleotide exchange factors: old concepts and new perspectives. *Genes Cancer*. **2**, 298–305 (2011).

74. E. Castellano, J. Downward, RAS Interaction with PI3K : More Than Just Another Effector Pathway, 261–274 (2015).

75. Y. Zheng *et al.*, Temporal regulation of EGF signalling networks by the scaffold protein Shc1. *Nature*. **499**, 166–71 (2013).

76. A. M. Del Rosario, F. M. White, Quantifying oncogenic phosphotyrosine signaling networks through systems biology. *Curr. Opin. Genet. Dev.* **20**, 23–30 (2010).

77. B. Blagoev, S.-E. Ong, I. Kratchmarova, M. Mann, Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat. Biotechnol.* **22**, 1139–45 (2004).

78. P. I. Domains, Assembly of Cell Regulatory Systems Through. *Science (80-.).* **300**, 445–452 (2003).

79. M. A. Stiffler *et al.*, PDZ domain binding selectivity is optimized across the mouse

proteome. *Science*. **317**, 364–369 (2007).

80. B. T. Seet, I. Dikic, M.-M. Zhou, T. Pawson, Reading protein modifications with interaction domains. *Nat. Rev. Mol. Cell Biol.* **7**, 473–483 (2006).

81. A. Wolf-Yadlin, M. Sevecka, G. MacBeath, Dissecting protein function and signaling using protein microarrays. *Curr. Opin. Chem. Biol.* **13**, 398–405 (2009).

82. A. Espejo, J. Côté, A. Bednarek, S. Richard, M. T. Bedford, A protein-domain microarray identifies novel protein-protein interactions. *Biochem. J.* **367**, 697–702 (2002).

83. B. Asbach, M. Kolb, M. Liss, R. Wagner, M. Schäferling, Protein microarray assay for the screening of SH3 domain interactions. *Anal. Bioanal. Chem.* **398**, 1937–1946 (2010).

84. S. Hu, Z. Xie, J. Qian, S. Blackshaw, H. Zhu, Functional protein microarray technology. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **3**, 255–268 (2011).

85. a Y. Ting, K. H. Kain, R. L. Klemke, R. Y. Tsien, Genetically encoded fluorescent reporters of protein tyrosine kinase activities in living cells. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 15003–15008 (2001).

86. J. Seong *et al.*, Detection of focal adhesion kinase activation at membrane microdomains by fluorescence resonance energy transfer. *Nat. Commun.* **2**, 406 (2011).

87. A. Buntru, T. Zimmermann, C. R. Hauck, Fluorescence resonance energy transfer (FRET)-based subcellular visualization of pathogen-induced host receptor signaling. *BMC Biol.* **7**, 81 (2009).

88. J. Sun *et al.*, Antagonism between binding site affinity and conformational dynamics tunes alternative cis-interactions within Shp2. *Nat. Commun.* **4**, 1–11 (2013).

89. R. J. Hause *et al.*, Comprehensive Binary Interaction Mapping of SH2 Domains via Fluorescence Polarization Reveals Novel Functional Diversification of ErbB Receptors. *PLoS One*. **7** (2012), doi:10.1371/journal.pone.0044471.

90. N. Hino *et al.*, Genetic incorporation of a photo-crosslinkable amino acid reveals novel protein complexes with GRB2 in mammalian cells. *J. Mol. Biol.* **406**, 343–53 (2011).

91. K. J. Roux, D. I. Kim, M. Raida, B. Burke, A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J. Cell Biol.* **196**, 801–810 (2012).

92. Y. Wu, Q. Li, X.-Z. Chen, Detecting protein-protein interactions by Far western blotting. *Nat. Protoc.* **2**, 3278–3284 (2007).

93. S. Manley *et al.*, High-density mapping of single-molecule trajectories with photoactivated localization microscopy. *Nat. Methods*. **5**, 155–7 (2008).

94. D. Oh *et al.*, Fast rebinding increases dwell time of Src homology 2 (SH2)-containing proteins near the plasma membrane. *Proc. Natl. Acad. Sci.* **109** (2012), pp. 14024–14029.

95. J. Wiedenmann *et al.*, EosFP, a fluorescent marker protein with UV-inducible green-to-red fluorescence conversion. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 15905–15910 (2004).

96. K. Nienhaus, G. U. Nienhaus, J. Wiedenmann, H. Nar, Structural basis for photo-induced protein cleavage and green-to-red conversion of fluorescent protein EosFP. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 9156–9159 (2005).

97. S. Das *et al.*, Single-molecule tracking of small GTPase Rac1 uncovers spatial regulation of membrane translocation and mechanism for polarized signaling (2015), doi:10.1073/pnas.1409667112.

98. J. a Ditlev, B. J. Mayer, L. M. Loew, There is more than one way to model an elephant. Experiment-driven modeling of the actin cytoskeleton. *Biophys. J.* **104**, 520–32 (2013).

99. a. R. French, D. K. Tadaki, S. K. Niyogi, D. a. Lauffenburger, Intracellular trafficking of epidermal growth factor family ligands is directly influenced by the pH sensitivity of the receptor/ligand interaction. *J. Biol. Chem.* **270** (1995), pp. 4334–4340.

100. J. Berkers, P. Henegouwen, J. Boonstra, Three Classes of Epidermal Growth Factor Receptors in HeLa Cells. *J. Biol. Chem.* **266**, 922–927 (1991).

101. H. Resat, J. a Ewald, D. a Dixon, H. S. Wiley, An integrated model of epidermal growth factor receptor trafficking and signal transduction. *Biophys. J.* **85**, 730–743 (2003).

102. F. Capuani *et al.*, Quantitative analysis reveals how EGFR activation and downregulation are coupled in normal but not in cancer cells. *Nat. Commun.* **6**, 7999 (2015).

103. D. C. Hess *et al.*, Computationally driven, quantitative experiments discover genes required for mitochondrial biogenesis. *PLoS Genet.* **5** (2009), doi:10.1371/journal.pgen.1000407.

104. a. Finkelstein *et al.*, Computational challenges of systems biology. *Computer (Long. Beach. Calif).* **37** (2004), doi:10.1109/MC.2004.1297236.

105. L. M. Loew, J. C. Schaff, The Virtual Cell: A software environment for computational cell biology. *Trends Biotechnol.* **19**, 401–406 (2001).

106. B. M. Slepchenko, J. C. Schaff, I. Macara, L. M. Loew, Quantitative cell biology with the Virtual Cell☆. *Trends Cell Biol.* **13**, 570–576 (2003).

107. T. Pawson, Specificity in Signal Transduction: From Phosphotyrosine-SH2 Domain Interactions to Complex Cellular Systems. *Cell*. **116**, 191–203 (2004).

108. M. J. Wagner, M. M. Stacey, B. a. Liu, T. Pawson, Molecular mechanisms of SH2- and PTB-Domain-containing proteins in receptor tyrosine kinase signaling. *Cold Spring Harb. Perspect. Biol.* **5**, 1–19 (2013).

109. B. a Liu *et al.*, SH2 domains recognize contextual peptide sequence information to determine selectivity. *Mol. Cell. Proteomics*. **9**, 2391–404 (2010).

110. H. Huang *et al.*, Defining the specificity space of the human SRC homology 2 domain. *Mol. Cell. Proteomics*. **7**, 768–784 (2008).

111. T. Curran, Y. Zhang, D. Ma, J. Sarkaria, F. White, MARQUIS: A multiplex method for absolute quantification of peptides and posttranslational modifications. *Nat. Commun.* **6**, 1–11 (2015).

112. J. V Olsen *et al.*, Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell*. **127**, 635–48 (2006).

113. M. Grusch *et al.*, Spatio-temporally precise activation of engineered receptor tyrosine kinases by light. **33**, 1713–1727 (2014).

114. M. Morimatsu *et al.*, Multiple-state reactions between the epidermal growth factor receptor and Grb2 as observed by using single-molecule analysis. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 18013–8 (2007).

115. A. Abulrob *et al.*, Nanoscale imaging of epidermal growth factor receptor clustering: effects of inhibitors. *J. Biol. Chem.* **285**, 3145–56 (2010).

116. I. Chung *et al.*, Spatial control of EGF receptor activation by reversible dimerization on living cells. *Nature*. **464**, 783–7 (2010).

117. Y. Sako, S. Minoguchi, T. Yanagida, Single-molecule imaging of EGFR signalling on the surface of living cells. **2** (2000).

118. S. R. Needham *et al.*, Measuring EGFR separations on cells with ~10 nm resolution via fluorophore localization imaging with photobleaching. *PLoS One*. **8**, e62331 (2013).

119. M. Sergeev, J. L. Swift, A. G. Godin, P. W. Wiseman, Ligand-induced clustering of EGF receptors: A quantitative study by fluorescence image moment analysis. *Biophys. Chem.* **161**, 50–53 (2012).

120. J. Ichinose, M. Murata, T. Yanagida, Y. Sako, EGF signalling amplification induced by dynamic clustering of EGFR. *Biochem. Biophys. Res. Commun.* **324**, 1143–1149 (2004).

121. P. Stanton *et al.*, Epidermal growth factor receptor expression by human squamous cell carcinomas of the head and neck, cell lines and xenografts. *Br. J. Cancer*. **70**, 427–433 (1994).

122. M. M. Wrann, C. F. Fox, Identification of Epidermal Growth Factor Receptors in a Hyperproducing Human Epidermoid Carcinoma Cell Line. *J. Biol. Chem.* **254**, 8083–8086 (1979).

123. B. Blagoev *et al.*, A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat. Biotechnol.* **21**, 315–8 (2003).

124. T. Okutani *et al.*, Grb2/Ash binds directly to tyrosines 1068 and 1086 and indirectly to tyrosine 1148 of activated human epidermal growth factor receptors in intact cells. *J. Biol. Chem.* **269**, 31310–31314 (1994).

125. M. Holgado-Madruga, D. R. Emlet, D. K. Moscatello, a K. Godwin, a J. Wong, A Grb2-associated docking protein in EGF- and insulin-receptor signalling. *Nature*. **379** (1996), pp. 560–564.

126. G. M. Rivera *et al.*, Requirement of Nck adaptors for actin dynamics and cell migration stimulated by platelet-derived growth factor B. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 9536–9541 (2006).

127. R. Sakai *et al.*, A novel signaling molecule, p130, forms stable complexes in vivo with v-Crk and v-Src in a tyrosine phosphorylation-dependent manner. *EMBO J.* **13**, 3748–3756 (1994).

128. Y. Zhang *et al.*, Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. *Mol. Cell. Proteomics*. **4**, 1240–50 (2005).

129. F. Supek, M. Bošnjak, N. Škunca, T. Šmuc, REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One*. **6**, e21800 (2011).

130. S. L. Harmer, a L. DeFranco, Shc contains two Grb2 binding sites needed for efficient formation of complexes with SOS in B lymphocytes. *Mol. Cell. Biol.* **17**, 4087–4095 (1997).

131. T. Pawson, N. Warner, Oncogenic re-wiring of cellular signaling pathways. *Oncogene*. **26**, 1268–1275 (2007).

132. J. E. Ferrell, S. H. Ha, Ultrasensitivity part III: cascades, bistable switches, and oscillators. *Trends Biochem. Sci.* **39**, 612–618 (2014).

133. B. N. Kholodenko, J. F. Hancock, W. Kolch, Signalling ballet in space and time. *Nat. Rev. Mol. Cell Biol.* **11**, 414–426 (2010).

134. H. Kitano, Computational systems biology. *Nature*. **420**, 206–210 (2002).

135. P. J. Verveer, P. I. H. Bastiaens, Quantitative microscopy and systems biology: Seeing the whole picture. *Histochem. Cell Biol.* **130**, 833–843 (2008).

136. W. a Lim, Designing customized cell signalling circuits. *Nat. Rev. Mol. Cell Biol.* **11**, 393–403 (2010).

137. W. X. Schulze, M. Mann, A novel proteomic screen for peptide-protein interactions. *J. Biol. Chem.* **279**, 10756–64 (2004).

138. K. D. Beebe, P. Wang, G. Arabaci, D. Pei, Determination of the binding specificity of the SH2 domains of protein tyrosine phosphatase SHP-1 through the screening of a combinatorial phosphotyrosyl peptide library. *Biochemistry*. **39**, 13251–13260 (2000).

139. J. Schlessinger, M. A. Lemmon, *Sci. STKE*, in press, doi:10.1126/stke.2003.191.re12.

140. W. M. Kavanaugh, C. W. Turck, L. T. Williams, {PTB} domain binding to signaling proteins through a sequence motif containing phosphotyrosine. *Science (80-.).* **268**, 1177–1179 (1995).

141. R. B. Jones, Do low-affinity ErbB receptor protein interactions represent the base of a cell signaling iceberg? *Expert Rev Proteomics*. **10**, 115–118 (2013).

142. L. B. Kleiman, T. Maiwald, H. Conzelmann, D. A. Lauffenburger, P. K. Sorger, Rapid Phospho-Turnover by Receptor Tyrosine Kinases Impacts Downstream Signaling and Drug Binding. *Mol. Cell*. **43**, 723–737 (2011).

143. D. L. Morris, K. W. Cho, Y. Zhou, L. Rui, SH2B1 enhances insulin sensitivity by both stimulating the insulin receptor and inhibiting tyrosine dephosphorylation of insulin receptor substrate proteins. *Diabetes*. **58**, 2039–2047 (2009).

144. N. Welsh, N. Makeeva, M. Welsh, Overexpression of the Shb SH2 domain-protein in insulin-producing cells leads to altered signaling through the IRS-1 and IRS-2 proteins. *Mol. Med.* **8**, 695–704 (2002).

145. W. X. Schulze, L. Deng, M. Mann, *Mol. Syst. Biol.*, in press, doi:10.1038/msb4100012.

146. C. W. Ward *et al.*, Systematic Mapping of Potential Binding Sites for Shc and Grb2 SH2 Domains on Insulin Receptor Substrate-1 and the Receptors for Insulin, Epidermal Growth Factor, Platelet-derived Growth Factor, and Fibroblast Growth Factor. *J. Biol. Chem.* **271**, 5603–5609 (1996).

147. A. G. Batzer, P. Blaikie, K. Nelson, J. Schlessinger, B. Margolis, The phosphotyrosine interaction domain of Shc binds an LXNPXY motif on the epidermal growth factor receptor . These include : The Phosphotyrosine Interaction Domain of Shc Binds an LXNPXY Motif on the Epidermal Growth Factor Receptor. **15** (1995).

148. A. G. Batzer, D. Rotin, E. Y. Skolnik, J. Schlessinger, the epidermal growth factor receptor . Hierarchy of Binding Sites for Grb2 and Shc on the Epidermal Growth Factor Receptor. **14** (1994), doi:10.1128/MCB.14.8.5192.Updated.

149. R. J. Daly *et al.*, SH2 domains prevent tyrosine dephosphorylation of the EGF receptor : identification of Tyr992 as the high-affinity binding site for SH2 domains of phospholipase CT. **11**, 559–567 (1992).

150. F. G. Gervais, L. M. Chow, J. M. Lee, P. E. Branton, A. Veillette, The SH2 domain is required for stable phosphorylation of p56lck at tyrosine 505, the negative regulatory site. *Mol. Cell. Biol.* **13**, 7112–7121 (1993).

151. R. B. Birge, J. E. Fajardo, B. J. Mayer, H. Hanafusa, Tyrosine-phosphorylated epidermal growth factor receptor and cellular p130 provide high affinity binding substrates to

analyze Crk-phosphotyrosine-dependent interactions in vitro. *J. Biol. Chem.* **267**, 10588–10595 (1992).

152. Y. Xie, A. M. Pendergast, M. C. Hung, Dominant-negative mutants of Grb2 induced reversal of the transformed phenotypes caused by the point mutation-activated rat HER-2/neu. *J. Biol. Chem.* **270**, 30717–30724 (1995).

153. A. Giubellino, T. R. Burke, D. P. Bottaro, Grb2 signaling in cell motility and cancer. *Expert Opin. Ther. Targets*. **12**, 1021–1033 (2008).

154. M. K. B. Wills, J. Tong, S. L. Tremblay, M. F. Moran, N. Jones, The ShcD signaling adaptor facilitates ligand-independent phosphorylation of the EGF receptor. *Mol. Biol. Cell*. **25**, 739–52 (2014).

155. J. Biscardi, M. Maa, D. Tice, M. Cox, c-Src-mediated Phosphorylation of the Epidermal Growth Factor Receptor on Tyr 845 and Tyr 1101 Is Associated with Modulation of Receptor Function. *J. Biol. …*. **274**, 8335–8343 (1999).

156. D. G. Motto *et al.*, In vivo association of Grb2 with pp116, a substrate of the T cell antigen receptor-activated protein tyrosine kinase. *J. Biol. Chem.* **269**, 21608–21613 (1994).

157. J. A. Jadwin *et al.*, Time-resolved multimodal analysis of Src Homology 2 (SH2) domain binding in signaling by receptor tyrosine kinases. *Elife*. **5**, 1–26 (2016).

158. B. N. Kholodenko, Cell signaling dynamics in time and space. *Nat. Rev. Mol. Cell Biol.* **7**, 165–176 (2006).

159. Y. M. Chook *et al.*, Protein Chemistry and Structure : The Grb2-mSos1 Complex Binds Phosphopeptides with Higher Affinity than The Grb2-mSos1 Complex Binds Phosphopeptides with Higher Affinity than Grb2 *. **271**, 30472–30478 (1996).

160. M. a. Lemmon, J. E. Ladbury, V. Mandiyan, M. Zhou, J. Schlessinger, Independent binding of peptide ligands to the SH2 and SH3 domains of Grb2. *J. Biol. Chem.* **269**, 31653–31658 (1994).

161. T. Kaneko *et al.*, *Sci. Signal.*, in press, doi:10.1126/scisignal.2003021.

162. J. A. Jadwin, in *SH2 Domains: Methods and Protocols*, B. A. Machida, Kazuya, Liu, Ed. (Springer New York, 2017), pp. 477–492.

163. M. S. Creamer *et al.*, Specification, annotation, visualization and simulation of a large rule-based model for ERBB receptor signaling. *BMC Syst. Biol.* **6**, 107 (2012).

164. T. Takino *et al.*, Tyrosine phosphorylation of the CrkII adaptor protein modulates cell migration. *J. Cell Sci.* **116**, 3145–3155 (2003).

165. S. Antoku, B. J. Mayer, Distinct roles for Crk adaptor isoforms in actin reorganization induced by extracellular signals. *J. Cell Sci.* **122**, 4228–4238 (2009).

166. B. a Liu *et al.*, SH2 domains recognize contextual peptide sequence information to determine selectivity. *Mol. Cell. Proteomics*. **9**, 2391–2404 (2010).

167. K. Sharma *et al.*, Ultradeep Human Phosphoproteome Reveals a Distinct Regulatory Nature of Tyr and Ser/Thr-Based Signaling. *Cell Rep.*, 1583–1594 (2014).

168. A. Kaushansky, A. Gordus, B. Chang, J. Rush, G. MacBeath, A quantitative study of the recruitment potential of all intracellular tyrosine residues on EGFR, FGFR1 and IGF1R. *Mol. Biosyst.* **4**, 643–653 (2008).

169. T. J. Wilson *et al.*, Fine specificity and molecular competition in SLAM family receptor signalling. *PLoS One*. **9** (2014), doi:10.1371/journal.pone.0092184.

170. S. Deindl, T. Kadlecek, X. Cao, J. Kuriyan, A. Weiss, Stability of an autoinhibitory interface in the structure of the tyrosine kinase ZAP-70 impacts T cell receptor response. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 20699–20704 (2009).

171. T. Bai, S. W. Luoh, GRB-7 facilitates HER-2/Neu-mediated signal transduction and tumor formation. *Carcinogenesis*. **29**, 473–479 (2008).

172. D. Pradip, M. Bouzyk, N. Dey, B. Leyland-Jones, Dissecting GRB7-mediated signals for proliferation and migration in HER2 overexpressing breast tumor cells: GTP-ase rules. *Am. J. Cancer Res.* **3**, 173–95 (2013).

173. M. Tanaka, R. Gupta, B. J. Mayer, Differential inhibition of signaling pathways by dominant-negative SH2/SH3 adapter proteins. *Mol Cell Biol*. **15**, 6829–6837 (1995).

174. M. I. Kontaridis, K. D. Swanson, F. S. David, D. Barford, B. G. Neel, PTPN11 (Shp2) mutations in LEOPARD syndrome have dominant negative, not activating, effects. *J. Biol. Chem.* **281**, 6785–6792 (2006).

175. D. Qian, M. Mollenauer, A. Weiss, Dominant-negative zeta-associated protein 70 inhibits T cell antigen receptor signaling. *J. Exp. …*. **183** (1996) (available at http://jem.rupress.org/content/183/2/611.abstract).

176. J. P. Northrop, M. J. Pustelnik, A. T. Lu, J. R. Grove, Characterization of the roles of SH2 domain-containing proteins in T-lymphocyte activation by using dominant negative SH2 domains. *Mol. Cell. Biol.* **16**, 2255–63 (1996).

177. D. J. Carroll, D. T. Albay, M. Terasaki, L. A. Jaffe, K. R. Foltz, Identification of PLC gamma-dependent and -independent events during fertilization of sea urchin eggs. *Dev. Biol.* **206**, 232–247 (1999).

178. N. Sunaga *et al.*, Knockdown of oncogenic KRAS in non-small cell lung cancers suppresses tumor growth and sensitizes tumor cells to targeted therapy. *Mol. Cancer Ther.* **10**, 336–46 (2011).

179. W. Pao *et al.*, KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. *PLoS Med.* **2**, e17 (2005).

180. W. De Roock *et al.*, Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: A retrospective consortium analysis. *Lancet Oncol.* **11**, 753–762 (2010).

181. A. Gschwind, O. M. Fischer, A. Ullrich, The discovery of receptor tyrosine kinases: targets for cancer therapy. *Nat. Rev. Cancer*. **4**, 361–370 (2004).

182. J. a McCubrey *et al.*, Ras/Raf/MEK/ERK and PI3K/PTEN/Akt/mTOR cascade inhibitors: how mutations can result in therapy resistance and how to overcome resistance. *Oncotarget*. **3**, 1068–111 (2012).

183. I. I. Wistuba, J. G. Gelovani, J. J. Jacoby, S. E. Davis, R. S. Herbst, Methodological and practical challenges for personalized cancer therapies. *Nat. Rev. Clin. Oncol.* **8**, 135–141 (2011).

184. Z. Zhang, A. L. Stiegler, T. J. Boggon, S. Kobayashi, B. Halmos, EGFR-mutated lung cancer: a paradigm of molecular oncology. *Oncotarget*. **1**, 497–514 (2010).

185. J. L. Bos, ras Oncogenes in human cancer: A review. *Cancer Res.* **49** (1989), pp. 4682–4689.

186. U. Guha *et al.*, Comparisons of tyrosine phosphorylated proteins in cells expressing lung

cancer-specific alleles of EGFR and KRAS. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14112–7 (2008).

187. L. K. Goh, A. Sorkin, Endocytosis of receptor tyrosine kinases. *Cold Spring Harb. Perspect. Biol.* **5**, a017459 (2013).

188. J. C. Tomshine *et al.*, Cell proliferation and epidermal growth factor signaling in non-small cell lung adenocarcinoma cell lines are dependent on Rin1. *J. Biol. Chem.* **284**, 26331–9 (2009).

189. J. S. Britson, F. Barton, J. M. Balko, E. P. Black, Deregulation of DUSP activity in EGFR-mutant lung cancer cell lines contributes to sustained ERK1/2 signaling. *Biochem. Biophys. Res. Commun.* **390**, 849–54 (2009).

190. F. G. Haj, B. Markova, L. D. Klaman, F. D. Bohmer, B. G. Neel, Regulation of receptor tyrosine kinase signaling by protein tyrosine phosphatase-1B. *J Biol Chem*. **278**, 739–744 (2003).

191. Y. Xu, L.-J. Tan, V. Grachtchouk, J. J. Voorhees, G. J. Fisher, Receptor-type protein-tyrosine phosphatase-kappa regulates epidermal growth factor receptor function. *J. Biol. Chem.* **280**, 42694–700 (2005).

192. T. A. Berset, E. F. Hoier, A. Hajnal, The C. elegans homolog of the mammalian tumor suppressor Dep-1/Scc1 inhibits EGFR signaling to regulate binary cell fate decisions. *Genes Dev.* **19**, 1328–40 (2005).

193. R. Karisch *et al.*, Global proteomic assessment of the classical protein-tyrosine phosphatome and “redoxome.” *Cell*. **146**, 826–840 (2011).

194. M. R. Schneider, E. Wolf, The epidermal growth factor receptor ligands at a glance. *J. Cell. Physiol.* **218**, 460–466 (2009).

195. M. Ojaniemi, K. Vuori, Epidermal Growth Factor Modulates Tyrosine Phosphorylation of p130Cas: INVOLVEMENT OF PHOSPHATIDYLINOSITOL 3’-KINASE AND ACTIN CYTOSKELETON. *J. Biol. Chem.* **272**, 25993–25998 (1997).

196. V. Danos, J. Feret, W. Fontana, R. Harmer, J. Krivine, Rule-Based Modelling of Cellular Signalling. *Web*, 1–25 (2008).

197. P. Filippakopoulos *et al.*, Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell*. **149**, 214–231 (2012).

198. A. G. Ladurner, C. Inouye, R. Jain, R. Tjian, Bromodomains mediate an acetyl-histone encoded antisilencing function at heterochromatin boundaries. *Mol. Cell*. **11**, 365–376 (2003).

199. G. Mulligan *et al.*, Mutation of NRAS but not KRAS significantly reduces myeloma sensitivity to single-agent bortezomib therapy. *Blood*. **123**, 632–639 (2014).

200. P. P. Fotiadou, C. Takahashi, H. N. Rajabi, M. E. Ewen, Wild-Type NRas and KRas Perform Distinct Functions during Transformation. *Mol. Cell. Biol.* **27**, 6742–6755 (2007).

201. V. Vacic, L. M. Iakoucheva, P. Radivojac, Two Sample Logo: A graphical representation of the differences between two sets of sequence alignments. *Bioinformatics*. **22**, 1536–1537 (2006).

202. a Kusumi, Y. Sako, M. Yamamoto, Confined lateral diffusion of membrane receptors as studied by single particle tracking (nanovid microscopy). Effects of calcium-induced differentiation in cultured epithelial cells. *Biophys. J.* **65**, 2021–2040 (1993).

203. G. Knaus, Ulla, Bamberg, Alison, Bokoch, in *Methods in Microbiology, vo 412, Neutrophil*

Methods and Protcols, G. Quinn, Mark, DeLeo, Frank, Bokoch, Ed. (Humana Press, 2007), pp. 59–67.