SNAP-tag Based Affinity Reagents for Proteomic Profiling of Cell Signaling

Matthew E. K. Chang

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2017

Reading Committee:

Dustin Maly, Chair

Shao-En Ong

Champak Chatterjee

Program authorized to offer degree:

Department of Chemistry

University of Washington

SNAP-tag Based Affinity Reagents for Proteomic Profiling of Cell Signaling

Matthew E. K. Chang

Chair of the Supervisory Committee:
Professor Dustin J. Maly
Department of Chemistry

Abstract

Rapid and reliable propagation of signals through the cell is critical for all facets of cellular biology. Disrupting the fidelity of these responses leads to a number of diseases, including cancer, diabetes, and neurodegeneration. Thus, methods that allow the general characterization of signaling protein properties and function are of great practical use. The formation of multi-protein complexes is a general trait of signaling cascades and necessary for the fidelity of information transfer. The composition, spatial arrangement, and regulatory state of these signaling complexes is dynamic, and can change depending on the type and duration of the stimulus. These dynamic properties makes probing the identities and interactions of these proteins difficult. The first chapter of this thesis describes a unique chemical genetic method that can be utilized with quantitative proteomics to study different groups of signaling proteins involved in signaling. The second chapter details the profiling of a group of phosphotyrosine-recognizing domains that are important determinants of tyrosine kinase substrate selection in dynamic signaling pathways. Together, these efforts describe new proteomic tools for

uncovering signaling protein function and provide insight into fidelity of tyrosine phosphorylation networks.

TABLE OF CONTENTS

List of Abbreviations

AAK1	AP2-associated protein kinase 1
Abl	Abelson protein tyrosine kinase
Arg	Arginine
ATP	Adenosine triphosphate
BAR	Bivalent affinity reagent
BG	O^6-benzylguanine
Blk	B lymphoid tyrosine kinase
Brk	Breast tumor kinase (PTK6)
Btk	Bruton tyrosine kinase
BSA	Bovine serum albumin
C-terminus	Carboxy terminus
CAM	Chloroacetamide
Cbl	Casitas B-lineage lymphoma proto-oncogene
Cbl-b	Casitas B-lineage lymphoma proto-oncogene b
CLP	O^4-benzyl-2-chloro-6-aminopyrimidine
Csk	c-Src kinase
DARPin	Designed Ankyrin repeat protein
DMSO	Dimethylsulfoxide
Dok1	Docking protein 1
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
ERK	Extracellular signal-regulated protein kinase
ESI	Electrospray ionization
Fes	Feline sarcoma/Fujinami avian sarcoma oncogene homolog
Grb2	Growth factor receptor-bound protein 2
GST	Glutathione S-transferase
Hck	Hematopoietic cell kinase

HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
HeLa	Henrietta Lacks epithelial cells
HPLC	High performance liquid chromatography
IC^{50}	Half maximal inhibitory concentration
IPTG	Isopropyl β-D-1thiogalactopyranoside
JNK2	*c*-Jun N-terminal kinase 2
K_d	Dissociation constant
LC-MS	Liquid chromatography mass spectrometry
LFQ	Label free quantification
Lck	Leukocyte c-terminal Src kinase
Lyn	Lck/Yes-related novel protein tyrosine kinase
Lys	Lysine
m/z	Mass-to-charge ratio
MAPK	Mitogen-activated protein kinase
MAPK1	ERK2
N-terminus	Amino terminus
NEK9	Never in mitosis A-related kinase 9
OD	Optical density
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SH2	Src homology 2
SH3	Src homology 3
Shc1	SHC-transforming protein 1
SILAC	Stable isotope labeling with amino acids in cell culture
Src	Avian sarcoma viral oncogene tyrosine kinase
TFA	Trifluoroacetic acid
Tris	Tris(hydroxymethyl)aminomethane)
TUBE	Tandem ubiquitin binding entity
TYK2	Non-receptor tyrosine kinase 2
UBA	Ubiquitin-associated domain
Yes	Yamaguchi sarcoma virus oncogene tyrosine kinase

Units

°	degree
Å	angstrom
C	celcius
g	gram
hr	hour
kDa	kilodalton
L	liter
m	meter; milli
M	molar
min	minute
mol	mole
n	nano
p	pico
s	second
µ	micro

NMR

d	Doublet
Hz	Hertz
J	Coupling constant in Hz
m	Multiplet
MHz	Megahertz
ppm	Parts per million
s	Singlet
t	Triplet
δ	Chemical shift in parts per million

List of Figures

Chapter 1

Chapter 2

List of Tables

Chapter 1

Chapter 2

Acknowledgements

The work presented in this thesis is a result of the support of many people who have guided me along the way. I would like to thank my advisor, Dusty Maly, for his patience and pushing me to become a better scientist. Under his guidance, I have learned a wide variety of skills, but most invaluably, he has helped me to learn to be a more critical and logical thinker. For that, I am most grateful. The multiple members who have passed through the Maly lab have been wonderful and supportive of my time here.

I would also like to thank my collaborator, Shao-En Ong, who has become a second adviser to me during my graduate school career. Under his tutelage, I've been able to expand my knowledge and set of scientific skills beyond what I would have been able to do at only one lab. Shao-En, Ho-Tak, Martin, Emily, and Danny have been kind enough to incorporate me into their lab and bestow trust in me that they do not readily offer to others. The results of my work could not have been made possible without the guidance and help of Shao-En, Ho-Tak, and Martin in particular.

Finally, I thank my friends, family, and loved ones for all of their support that has sustained me throughout this journey. All of my friends that I've made have helped make my time here truly great. Although they are often far away, Mari and Chubs have supported me, showered me with love, and always cheered me up whenever we get a chance for a visit. My parents have sacrificed to enable me to pursue any path that I chose. Their love and support have truly driven me to become the best version of myself.

Dedication

To my Mom, Dad, Marsipan, and Mr. Chub

Chapter 1: Bivalent Affinity Reagents for Proteomic Profiling

I. Introduction

Signaling is a critical component of cellular functioning as it allows the cell to couple extracellular stimuli to a direct intracellular response. Cells have developed complex molecular systems to interpret a variety of analogue signals and transmit them as digital ones. Because their deregulation is involved in several disease states, most commonly cancer, understanding signaling pathway networks is of great utility (1). One of the principle means cells have developed to manage the complexity of signals and responses is through the dynamic arrangement of signaling molecules including kinases, scaffolding proteins, and regulatory proteins, into large signaling complexes. By organizing signaling components into complexes, cells can improve the efficiency, control, and robustness of cellular signaling as well as creating high functioning hubs of crosstalk between important pathways.

The traditional molecular biology paradigm of one gene-one protein-one function has been a driving force in research aimed at understanding signaling processes. However, it has become increasingly clear that this approach has certain limitations because it can neglect important contextual relationships (2). The dynamic nature of signaling complexes is a particularly illustrative example of this difficulty as the type and duration of the extracellular stimulus can alter complex composition, leading to significantly different phenotypic outcomes even within the same pathway (3, 4). In order to consider the signaling network as a whole, new approaches establish the particular relationships, or edges, between molecules of interest (nodes) in the context of the entire pathway. This requires an integrative view of these biological processes and presents an opportunity for new methods to visualize and probe the changes in these signaling networks.

The prevailing method for determining the nodes and edges of protein interaction and protein signaling networks has been the use of affinity purification-mass spectrometry (AP-MS) (5). In this technique, a “bait” protein or molecule captures “prey” proteins as well as other proteins interacting with the prey. These isolated proteins are subsequently identified by mass spectrometry. Protein enrichment from complex lysate has commonly utilized antibody-based capture of endogenous or tagged proteins. The number and quality of available immunoprecipitation-suitable antibodies and the need to express recombinant or tagged proteins, which can disrupt native interactions in the context of the signaling pathway, can limit the utility of these reagents. Other approaches have used small molecule inhibitors immobilized to resin as bait, but these generally lack the ability to distinguish between activation states of signaling

molecules and thus are limited in uncoupling the dynamic between a signaling protein and its complex. A general concern in the use of AP-MS for mapping networks is whether the proteins identified are truly specific protein-protein interactions as opposed to non-specifically copurified proteins (6).

In order to address these limitations, we sought to develop quantitative affinity reagents that can specifically interrogate cellular signaling cascades through the relationships and composition of post-translational modifications and signaling enzymes, most notably protein kinases, in macro-molecular signaling complexes. By using affinity reagents that display two separate binding ligands, true interactions can be analyzed and mapped with quantitative mass spectrometry. The development of this bivalent method will allow for quantitative and global profiling of signaling cascades and provide distinct signaling fingerprints of different cell types and states.

In this strategy, a bivalent affinity reagent (BAR), built from previous methodology used in the lab, will enrich signaling complexes containing kinases and signaling proteins containing phosphotyrosine motifs. Protein kinases are central cogs found nearly ubiquitously in all signaling pathways and catalyze the transfer of the γ-phosphate of ATP to alcohol-containing amino acids of specific substrates (7). This phosphorylation event is not only a key regulator in numerous biological functions such as cellular growth, proliferation, and differentiation, but also is a critical recognition motif for phosphorylation dependent interactions. The first discovered example of such a phosphorylation-dependent interaction was between phosphotyrosine and Src Homology 2 (SH2) domains (8). These protein interaction domains–along with others such as SH3, UBD, and PTB domains–enable modular organization of signaling proteins and allow them to be tethered in a signal dependent fashion. As a result, phosphorylation by protein kinases plays an important role in the assembly and composition of signaling complexes in response to cellular stimuli, shaping the network at large. In theory, probing the abundance and composition of kinase-containing signaling complexes can be used as reporters for cellular state.

SNAP-tag-Based Bivalent Methodology

Our approach tethers two binding ligands to target separate and unique members that are part of the same kinase-containing signaling complex. This methodology is adapted from our SNAP-tag-based platform of generating bivalent inhibitors for specific and potent kinase inhibition. Because probing kinase function among the greater than 500 members is difficult due to highly

conserved sequence homology in the ATP-binding active site, tools that mitigate these hurdles are desirable (9). In order to study the function of specific kinases, an inhibitor must not only show potency against its target, but also be selective over other members of the kinome. One method to successfully achieve greater potency and selectivity is by combining an ATP-competitive small molecule with a secondary binding ligand that binds to another site on the kinase involved in known protein-protein interactions to form a bivalent or bisubstrate inhibitor (10, 11).

We previously developed a chemical genetic method to generate bivalent kinase inhibitors that implements a self-labeling protein, engineered to display a selective secondary binding ligand of a protein kinase (12-14). The small molecule is tethered to the secondary binding ligand by

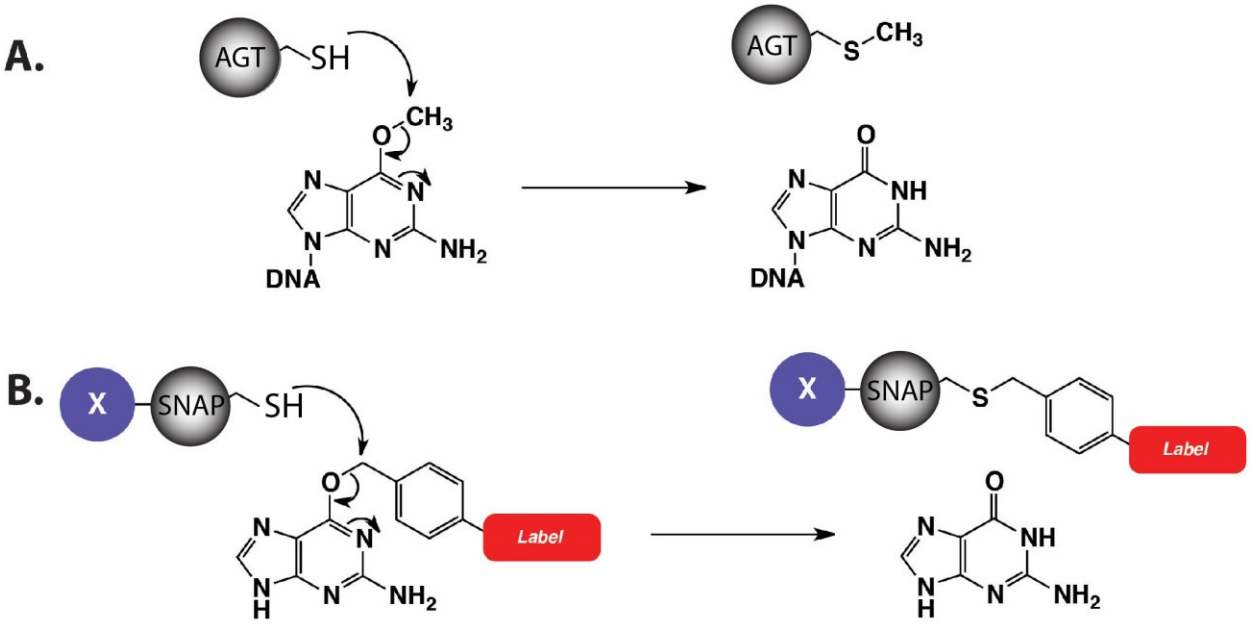


Figure 1-1. SNAP-tag labeling mechanism. (A) The normal function of AGT is to repair alkylated DNA lesions through an irreversible S_N2 reaction with its active site cysteine. (B) Fusion proteins with AGT can be labeled with small molecules with specific chemical tags.

SNAP-tag, an engineered form of O^6-alkylguanine-DNA alkyltransferase (AGT). The normal function of AGT in the cell is to repair O^6-alkylated guanine bases of DNA and acts by irreversibly transferring the alkyl lesion from the damaged base in an S_N2 reaction to its active site cysteine (Figure 1-1) (15). SNAP-tag was developed by Johnsson and coworkers to selectively label itself with small molecules linked to O^6-benzylguanine (BG) or O^4-benzyl-2-chloro-6-aminopyrimidine (CLP) (16, 17). This approach in generating bivalent reagents with SNAP-tag is unique in that an entire, folded protein is used as a scaffold for ligand display. The active site cysteine is located in a relatively shallow hydrophobic pocket, which can allow favorable access to the ATP-competitive small molecule, while the secondary binding ligand can be displayed from either the N- or C-

terminus of the protein scaffold through genetic fusion of these peptides with SNAP-tag. Conveniently, the N- and C-terminus both sit on the same face of the active site of SNAP-tag, which allows for favorable binding interactions with the target (18). Our lab has applied this methodology to develop potent and selective bivalent inhibitors of Src, Abl, EGFR, p38α and Pim1 (12, 13).

Bivalent Inhibition of Phosphorylation State-Specific ERK

We also began taking advantage of antibody mimetic proteins including monobodies and designed ankyrin repeat proteins (DARPins) as secondary binding ligands owing to their high stability and flexibility in generating specific binders. In 2012, Plückthun and coworkers exploited the conformational difference between ERK phosphorylation states and developed phosphorylation-specific ERK DARPins using ribosome display (19). We believed these proteins showed great promise as activation specific ligands for use as a bivalent reagent (Figure 1-2). The DARPin target, extracellular signal-regulated kinase 2 (ERK2), is a serine/threonine-specific mitogen-activated protein kinase (MAPK) part of the Ras-Raf-MEK-ERK pathway. This cascade is activated by extracellular stimuli and conveys this signal to the nucleus to regulate gene expression. Depending on the cell surface stimulus, the pathway can relay a signal that causes the prevention or induction of apoptosis or cell cycle progression (20). ERK2 is activated by phosphorylation of a threonine and tyrosine residue on its activation loop, causing a conformational change between the active and inactive forms of the kinase (21).

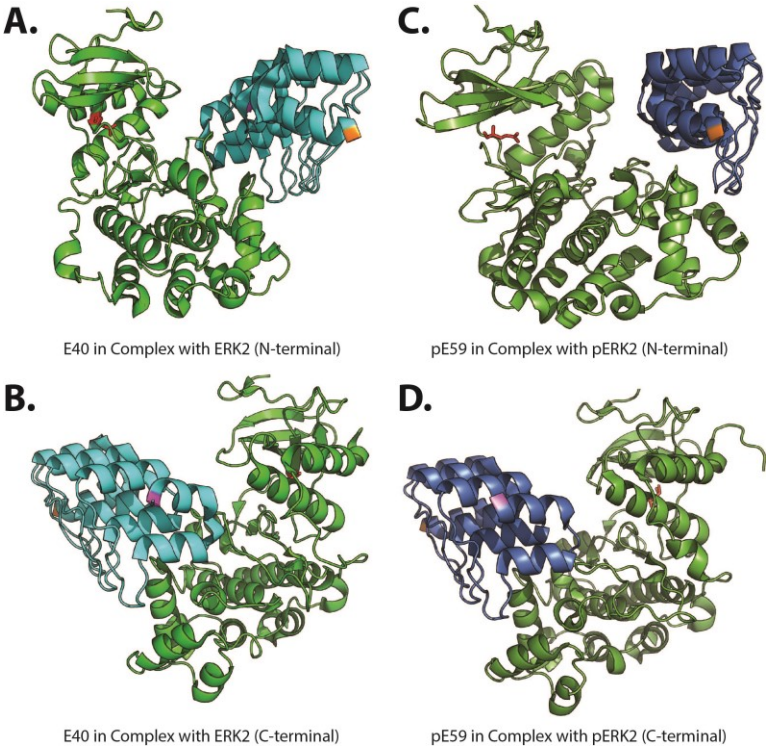


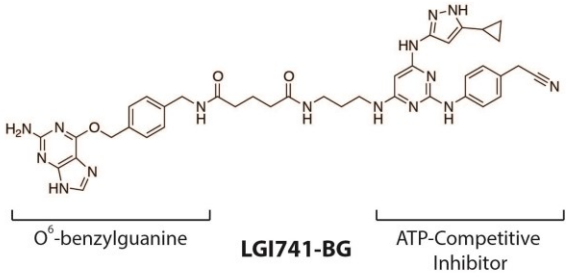
Figure 1-2. Crystallographic structures showing ERK DARPins bound to ERK2 with the ATP-binding site of ERK2 highlighted in red. (A) and (B) show E40 bound to ERK2 with the N- and C-terminus of the DARPin highlighted in orange and magenta, respectively. (C) and (D) show the bound pE59 DARPin.

The phospho-specific ERK DARPin, pE59, showed >74-fold selectivity for the active form of ERK2 (K_d = 117 nM) versus the inactive kinase (K_d > 8.7 μM). The non-phospho-specific ERK DARPin, E40, had a K_d = 6.6 nM for the inactive form of ERK2, which was 182-fold more selective over the phosphorylated form of ERK2 (K_d = 1.2 μM). With these new DARPins that are specific for a post-translational modification, the SNAP-tag based bivalent inhibitor platform was applied to generate phosphorylation-specific ERK inhibitors.

A. ERK DARPin SNAP-tag Constructs

SNAP-pE57 = SNAP-GSGTGSGS-pE57
SNAP-pE59 = SNAP-GSGTGSGS-pE59
SNAP-E40 = SNAP-GSGTGSGS-E40

B.



C.

ppERK2 IC_{50} (nM)

	No Inhibitor	LGI741
No Protein	N/A	171 ± 12
SNAP(wt)	NT	> 10000
SNAP-pE57	52* ± 21	< 0.32
SNAP-pE59	37* ± 4	< 0.32
SNAP-E40	> 10000	> 10000

Figure 1-3. Bivalent inhibitors of active and inactive ERK. (A) SNAP-tag fusions that target ppERK (SNAP-pE57, SNAP-pE59) and ERK (SNAP-E40). (B) BG-derivative of LGI741. (C) In vitro activities of the separate binding moieties and bivalent constructs. *SNAP-pE57 and SNAP-pE59 did not show maximal ppERK2 inhibition at the highest inhibitor concentration tested.

Two different phospho-specific ERK DARPins, pE57 and pE59, and one non-phospho ERK DARPin, E40, were selected as secondary binding ligands for bivalent constructs. C-terminal fusion proteins of SNAP-tag connected via a flexible, eight amino-acid linker to each selected DARPin were constructed using overlap-extension PCR and then inserted into the pMCSG7 LIC vector (Figure 1-3A).

Bivalent inhibitors with the three different selected DARPins were assembled in a labeling reaction with SNAP-tag and the small molecule, LGI741, chosen as the ATP-competitive small molecule (Figure 1-3B). The monovalent pieces and bivalent inhibitors were subsequently tested in an *in vitro* kinase activity assay against activated ERK2. LGI741-CLP showed potent inhibition against activated ERK2 with an IC_{50} = 171 ± 12 nM, but upon conjugation to SNAP-tag, saw knockback and was not active at all (IC_{50} > 10 μM). Although the two unconjugated ppERK–specific DARPin constructs had relatively low IC_{50} values, they only displayed partial inhibition against the active kinase as even the highest concentrations tested (~15 μM) never reached maximal inhibition. This may be due to the DARPin binding to the kinase on the activation loop and sterically interfering with the phosphorylation of substrate in certain conformations of the fusion protein. The bivalent inhibitors with pE57 and pE59, specific for phospho-ERK, displayed significant inhibition that

approached the enzyme concentration tested (IC_{50} < 0.32 nM). The inactive-specific bivalent inhibitor, SNAP-E40-LGI, showed no activity against ppERK (IC_{50} > 10 μM), validating its preference against phosphorylated kinase.

These inhibitors were also assayed against the related MAPKs p38α and JNK2 to offer an initial selectivity profile. IC_{50} values for each activated kinase were compared with their ppERK2 counterpart as displayed in Figure 1-4. The free small molecule showed potent inhibition against JNK2 (IC_{50} = 750 ± 30 nM), but the bivalent phospho-specific ERK inhibitors showed much lower potency with low micromolar activity. The bivalent inhibitors also showed no activity against p38α (IC_{50} > 10 μM). In addition to this initial selectivity test, the bivalent pE59 construct has undergone selectivity testing by Novartis in their kinase affinity matrix (KAM) screen, where it has selectively enriched spiked ppERK2 from a complex lysate mixture (22).

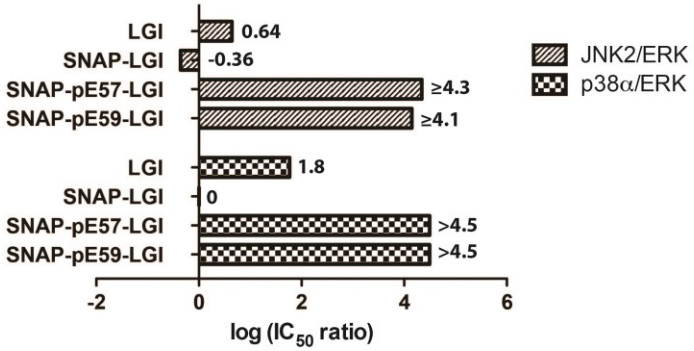


Figure 1-4. Kinase selectivity profile of the phosphorylation-specific ERK bivalent inhibitors and LGI against the MAPKs JNK2 and p38α.

Bivalent Affinity Reagents in Probing Signaling Networks

Given the encouraging results using the bivalent inhibitor of phospho-ERK, we looked to expand the utility of this methodology to probe signaling networks on a larger scale. One of the current limitations of AP-MS, the need for tight binders with high specificity to elucidate true interactions, lined up well with the hurdles faced in generating tools to probe kinase function. Furthermore, the SNAP-tag method for generating bivalent reagents has shown to be highly modular because each binding ligand can be rapidly swapped out to tune the binding specificity towards targets of interest.

For these bivalent affinity reagents or BARs, a pan kinase inhibitor displayed from SNAP-tag will target kinases in signaling complexes, while a secondary binding ligand, the Grb2 SH2 domain, will enrich phosphotyrosine-containing proteins. The reagent is immobilized on resin and used to affinity purify proteins from complex lysate (Figure 1-5). Enrichment of signaling complex proteins are analyzed by stable isotope labeling with amino acids in cell culture (SILAC), a quantitative proteomic technique that allows for a direct comparison between two or more conditions (23). In this technique, cells are subcultured into populations grown in media containing "heavy" amino acids, specifically arginine and lysine, with ^{13}C and ^{15}N stable isotopes or "light" amino acids (^{12}C, ^{14}N). The respective labeled amino acids are incorporated into the cell's proteins with full incorporation within five cell doublings. These separately labeled cell populations can be harvested following respective treatment and mixed equally into a single sample for MS analysis, where peptide signals from each population are distinguishable by mass shift. By comparing the integration of the light and heavy signal for a given peptide, relative quantitation for thousands of peptides in a sample can be achieved. SILAC is well suited for two or more condition comparisons and is a useful technique in elucidating true binders from non-specific ones.

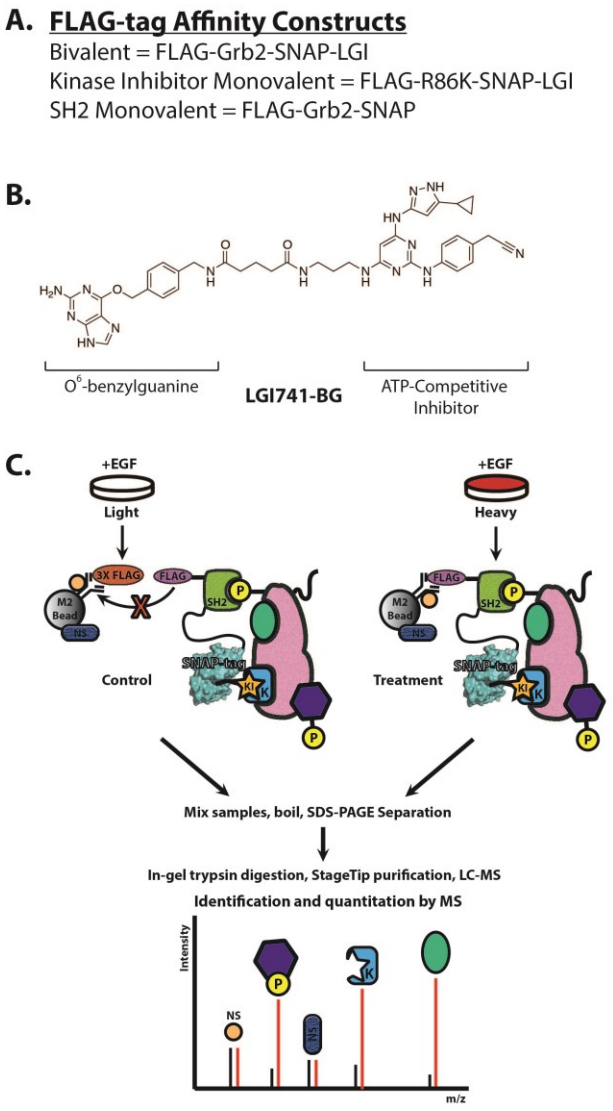


Figure 1-5. SNAP-tag-based affinity reagents for the enrichment of signaling complexes. (A) SNAP-tag fusions of bivalent and monovalent affinity reagents. (B) BG-derivative of LGI741. (C) Pulldown schematic with EGF stimulated, SILAC-labeled lysate to compare true protein binders in a treatment sample from the non-specific interactions in the control.

In 2003, Blagoev *et al.* used a GST-tagged fusion protein with the SH2 domain of the adapter protein, Grb2, to enrich SILAC-labeled signaling proteins from EGF stimulated and

unstimulated HeLa cell lysate (24). Using the resin-immobilized phosphotyrosine-binding SH2 domain to affinity purify proteins from lysate, the authors were able to identify 228 proteins, 28 of which were selectively enriched based on EGF stimulation. Signaling proteins important in the EGFR pathway were purified including Shc1, Sos1, Vav2, and Vav3. Furthermore, proteins involved in the recycling of EGFR such as ubiquitin, Cbl-b, and the AP-2 subunit proteins were identified as part of protein complexes.

With a bivalent affinity reagent featuring the same Grb2 SH2 domain used above as a binding ligand, the bait can enrich this subset of signaling proteins, but the kinase inhibitor can provide additional specificity towards kinase containing complexes. This reagent should be able to preferentially engage signaling complexes over monovalent signaling proteins because it can simultaneously bind to two separate sites in a complex. Enrichment conditions can be tuned to fully maximize the advantages provided by this bivalent interaction to specifically target these complexes. In addition, the modularity of using SNAP-tag display will allow for diverse targeting of different subsets of signaling proteins based on the general binding preferences for either small molecule or secondary binding ligand.

II. Results and Discussion

FLAG-Tagged Bivalent Affinity Reagents

In order to generate bivalent affinity reagents for the enrichment of diverse signaling complexes, the kinase inhibitor LGI741, which has a broad kinase profile, was chosen to complement the phosphotyrosine binding of Grb2's SH2 domain. The molecule was linked to the chemical recognition moiety benzylguanine, for rapid conjugation to SNAP-tag fusions (Figure 1-5B). The Grb2 SH2 domain was fused to the N-terminus of SNAP-tag through a flexible 24 amino acid spacer using overlap extension PCR. N-terminal to this was a FLAG-tag recognition sequence for anti-FLAG M2 resin capture. The final FLAG-BAR was assembled via a SNAP-tag labeling reaction with the kinase small molecule, and monitored by intact mass MS to compare the deconvoluted masses of the labeled and unlabeled peaks. Monovalent forms of these reagents were also generated for use as controls as well as competition reagents to titrate away monovalent signaling components. A phosphotyrosine "dead" SH2 mutant that does not bind to phosphorylated tyrosine residues was produced by mutating a critical arginine at position 86 to lysine using QuikChange. As part of a SNAP-tag fusion, the protein was labeled with the kinase small molecule to form the FLAG-R86K-SNAP-LGI monovalent affinity reagent. A FLAG-Grb2

SH2 monovalent reagent was generated by simply leaving the SNAP-tag active site unconjugated to the inhibitor.

In an initial profile, respective bivalent and monovalent affinity reagents were immobilized to anti-FLAG resin and incubated with EGF-stimulated (5 min), SILAC-labeled HeLa cell lysate. Excess 3X FLAG peptide was added in the control samples to compete away FLAG-tagged bait from resin, thus establishing background binders from the specific interactions in the treatment incubations with no competitor. Control samples were incubated with light-labeled SILAC lysate and treatment samples with heavy-labeled lysate for the forward conditions, and vice versa for the reverse in order to confirm changes in protein capture are not an artifact of the lysate. Following incubation and washes, light and heavy samples from the same respective conditions were mixed, eluted by boiling in SDS, and subject to in-gel trypsin digestion for processing by nano-flow LC-MS.

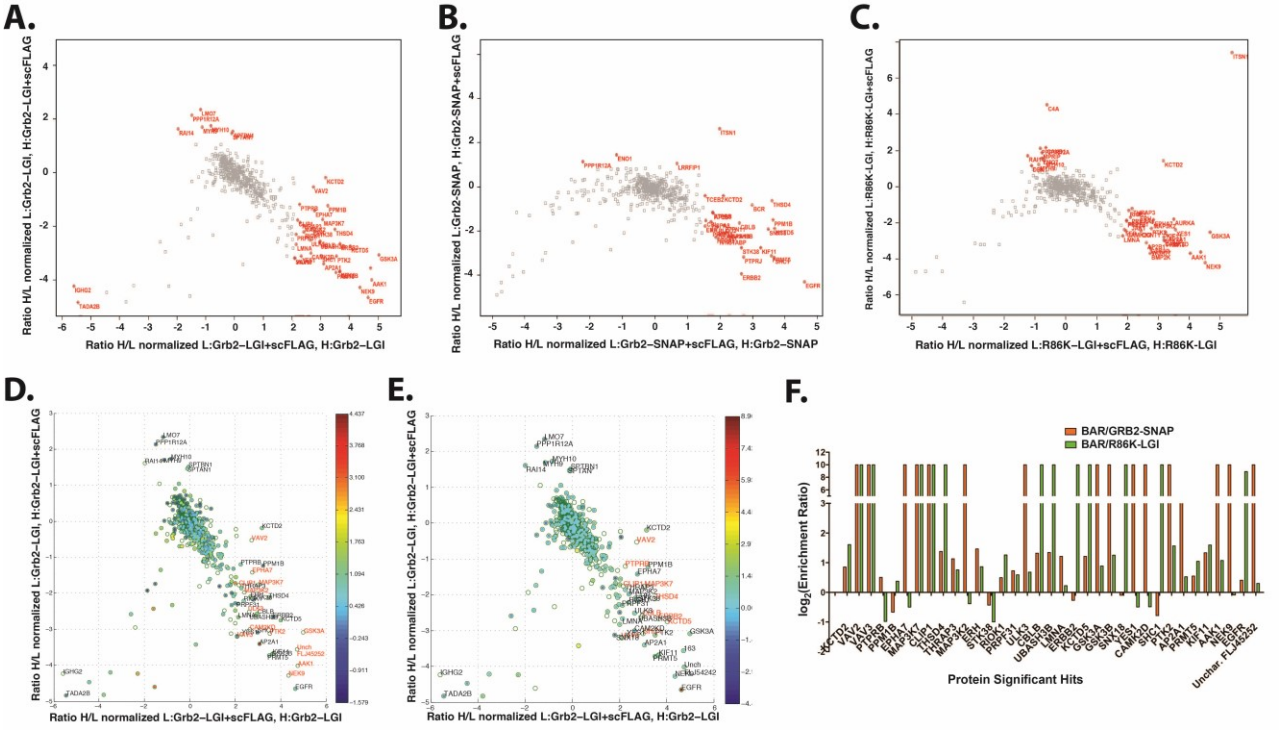


Figure 1-6. Scatter plots for initial profiling experiments with the bivalent and monovalent affinity reagents. (A) Bivalent Grb2-SNAP-LGI. Significant hits highlighted in red. (B) Monovalent Grb2-SNAP (C) Monovalent R86K-SNAP-LGI. (D,E) Scatter plots of the BAR with heat maps comparing the intensity ratios of the bivalent reagent to the respective monovalent signal (Grb2-SNAP=D, R86K-SNAP-LGI=E). Proteins highlighted in red indicate that the particular protein is unique to BAR enrichment in comparison to the respective monovalent reagent. (F) Relative enrichment between the BAR significant hits and its monovalent counterpart (Grb2-SNAP=orange, R86K-SNAP-LGI=green) is shown as a $\log_2$ comparison between respective intensity ratios. Proteins with values of 10 indicate that it was uniquely enriched by the BAR in comparison to the monovalent reagent.

Scatter plots of the monovalent affinity reagents showed bias towards certain subsets of proteins. For the Grb2 monovalent reagent, many of the significant hits were the same proteins

observed in previous SH2 domain pulldown experiments and are known interactors with Grb2 such as Shc1, EGFR, and Cbl-b, along with phosphatases and proteins involved in cytoskeletal rearrangement (Figure 1-6B). The R86K-SNAP-LGI construct enriched a large number of kinases, including Nek9, GSK3A, and Aurora kinase, which were shown to be targets of the LGI small molecule in profiling experiments by Novartis (Figure 6C). With the BAR, the scatter plot in Figure 6A shows many of the same significant hits greatly enriched by the monovalent reagents (e.g. EGFR, Nek9, GSK3A). However, a detailed comparison of the signal intensity ratio for respective peptides shows that the BAR enriches many of these shared hits to a greater extent than its monovalent counterparts (Figure 1-6D-F). In addition, the BAR also uniquely enriched proteins in comparison to the monovalent reagents such as the guanine nucleotide exchange factors Vav2 and Vav3, MAP3K7 kinase, and CLIP1, which is believed to be involved in the regulation of the microtubule cytoskeleton (25).

In order to determine the bivalent interactions made by our BAR and proteins in a signaling complex, competition experiments with the monovalent pieces were explored. By titrating in a specific monovalent piece during the capture step, bivalent interactions can be selected over monovalent ones. This competition also validates interaction strength by the monovalent affinity reagents. For the first competition experiment, five-fold excess of the untagged, free small molecule LGI was added during lysate-bait incubation in both the control and treatment samples for all three respective affinity reagents. SILAC-labeled HeLa cells were EGF stimulated for 5 min. With capture using the BAR, interactions between the SNAP-tag conjugated LGI small molecule and protein targets should be disrupted except for strong bivalent contacts. The Grb2 monovalent reagent should see very little change to its enrichment profile while the R86K-SNAP-LGI reagent should see a significant reduction in protein capture. These expectations were observed upon analysis of these samples as shown in Figure 1-7. Overall protein enrichment is reduced for the BAR and kinase capture in particular is drastically ameliorated based on significant hits. Interestingly, a number of proteins still show up as uniquely enriched by the BAR relative to both monovalent counterparts, such as Vav2 and Vav3 among others, suggesting these are bivalent interactions that still maintain high affinity even in the presence of the free small molecule.

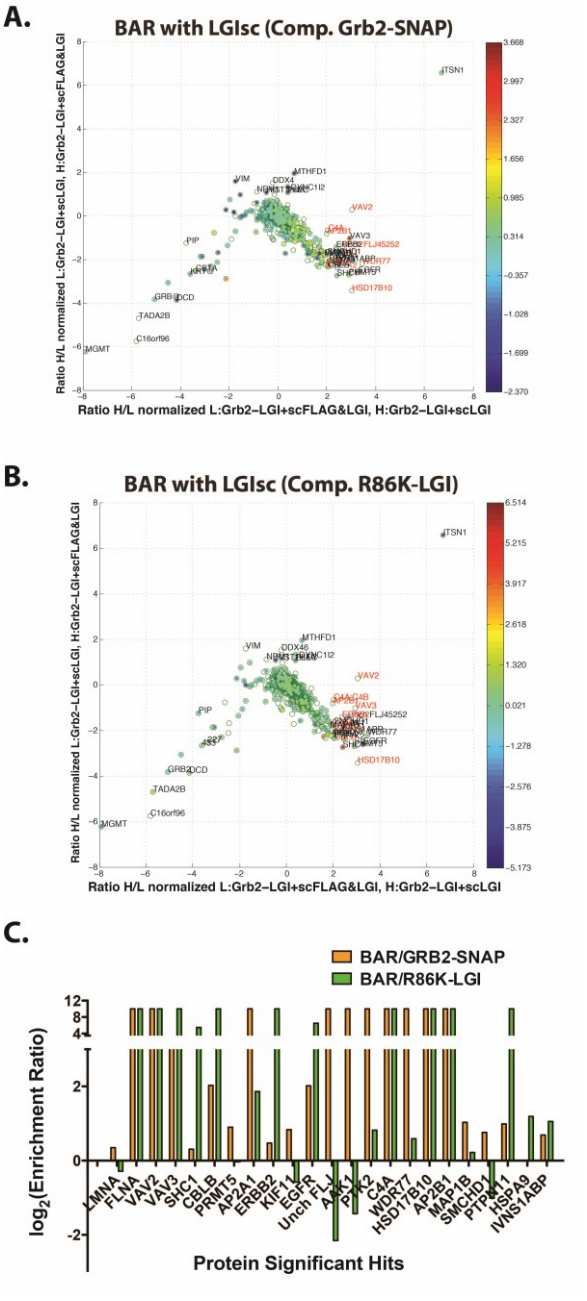


Figure 1-7. Scatter plots for competition with free inhibitor. (A) Comparison between BAR and monovalent Grb2-SNAP. (B) Comparison between BAR and monovalent R86K-SNAP-LGI. Heat maps compare the intensity ratio between the BAR and the respective monovalent signal. Proteins highlighted in red indicate that the particular protein is unique to BAR enrichment in comparison to the respective monovalent reagent. (C) Relative enrichment between the BAR significant hits and its monovalent counterpart (Grb2-SNAP=orange, R86K-SNAP-LGI=green) is shown as a $\log_2$ comparison between respective intensity ratios. Proteins with values of 10 indicate unique enrichment by the BAR in comparison to the monovalent reagent.

Given these encouraging results, the next step was to repeat the competition experiments with untagged versions of the monovalent affinity reagents used thus far. Each untagged monovalent affinity reagent would act as a competitor against the BAR with respect to that particular monovalent interaction. These competition profiles were compared to that of the BAR without competitor in a third series of conditions. In the control sample, the BAR was again incubated in the presence of excess 3X FLAG peptide, but ten-fold molar excess of untagged monovalent reagent was added to both the control and treatment capture samples. Overall, monovalent competition drastically reduced the number of proteins enriched by the BAR relative to no competition, which had 31 significant hits. With Grb2-SNAP competitor, the BAR only enriched 8 proteins characterized as significant, while 10 were identified when R86K-SNAP-LGI was the competitor (Figure 1-8A-B). In terms of relative level of enrichment based on signal intensity ratios of particular proteins, monovalent competition generally decreased the ability of the BAR to capture prey proteins from lysate (Figure 1-8C-E). Although the extent of this reduction is not as drastic as anticipated based on the previous data with free LGI as the competitor, each monovalent component was effective at reducing the number of proteins enriched by the BAR, which should allow it to determine monovalent interactions from bivalent ones as part of proteins on the same signaling complex.

Although these results were promising, a major limitation with the use of these FLAG-tagged constructs was high experiment turnaround times as a result of sample-processing. During SDS-PAGE separation of eluted samples, each gel lane, representing one replicate of an experimental condition (either the forward or reverse), is divided into 5 gel slices for in-gel trypsin digestion. This is necessary because during the elution step, the bait as well as the light and heavy chain of the FLAG antibody from resin is released into the sample. These three proteins on their own represent a large portion of the total protein sampling space and as such must be separated as their presence can mask other proteins during MS analysis. Thus, a forward and reverse experiment for just one condition expands from two samples to ten and an instrument run time of five hours to twenty-five. Given this bottleneck in experiment processing and the need to screen multiple conditions for tuning the BAR, a new experimental approach seemed warranted.

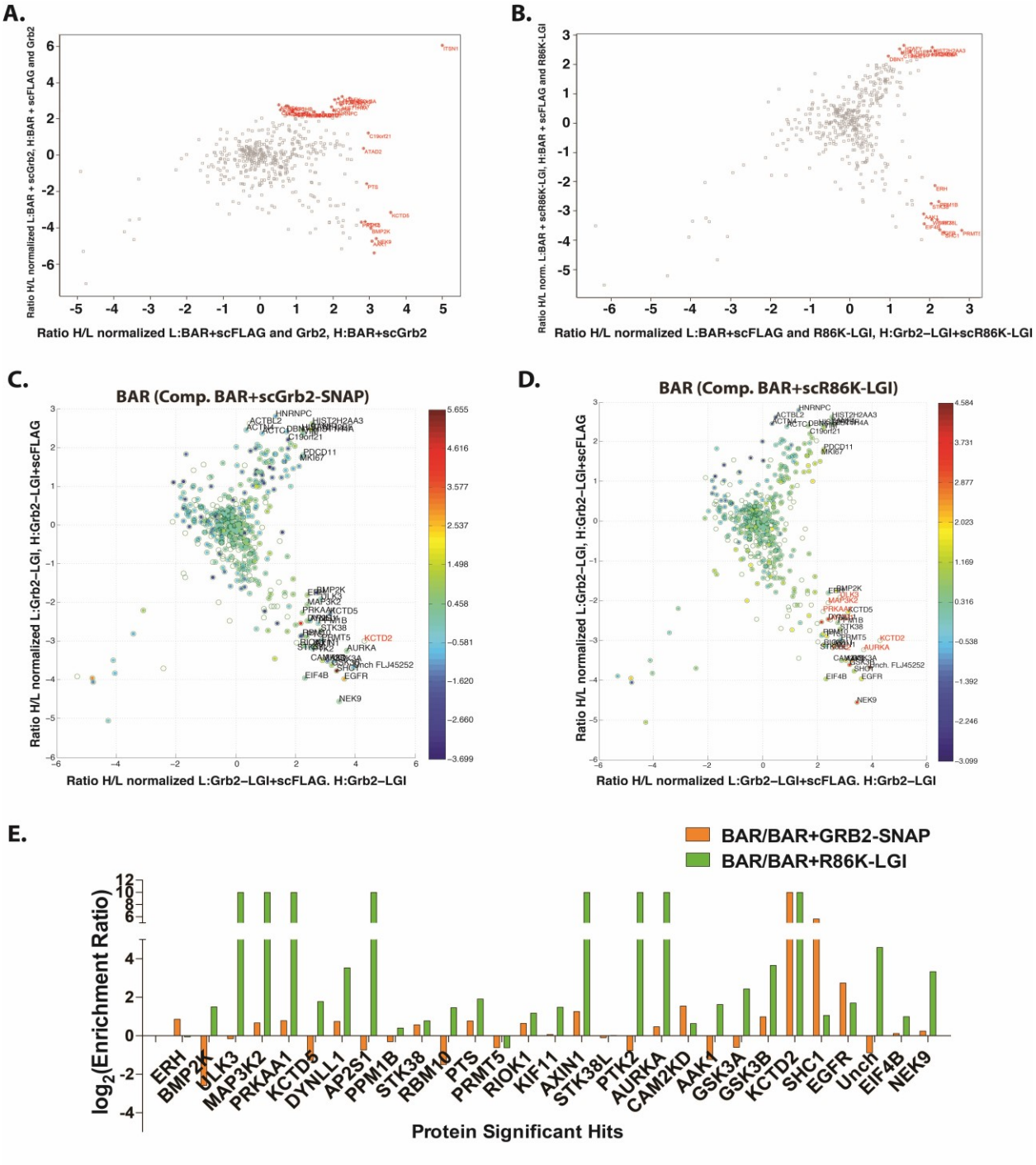


Figure 1-8. Scatter plots for competition with untagged monovalent affinity reagents. (A) BAR with Grb2-SNAP competition. Significant hits are highlighted in red. (B) BAR with R86K-SNAP-LGI competition. (C-D) Heat maps comparing the intensity ratio between the BAR and the respective BAR plus monovalent competition signal. Proteins highlighted in red indicate that the particular protein is unique to BAR enrichment in comparison to the competition experiment (E) Relative enrichment between the BAR significant hits and its monovalent competition counterpart (Grb2-SNAP=orange, R86K-SNAP-LGI=green) is shown as a $\log_2$ comparison between respective intensity ratios. proteins with values of 10 indicate unique enrichment by the BAR.

ASH*-Based Affinity Reagents

The first attempt at generating new affinity reagents to circumvent these processing issues used a hexylchloride 'catch-and-release' system developed by our lab (26). This ASH* (AGT-SUMO*-HALO) system features HaloTag, which is a self-labeling protein that reacts with alkylchloride-labeled molecules, AGT or SNAP-tag as discussed previously, and a cleavage site, SUMO*, a SUMO variant (R64T and R71E) recognized by an engineered form of the protease Ulp1, Ulp1*. For our affinity reagents, the Grb2-SH2 domain was connected to the N-terminus of ASH*. This construct could be covalently labeled with the kinase inhibitor via SNAP-tag as before, but directly conjugated to resin via HaloTag. Furthermore, the SUMO* cleavage site would allow for highly selective release of the bait and its prey with Ulp1* retaining non-specific binders onto the resin. Unfortunately, these new affinity reagents ended up being highly insoluble during bacterial expression and protein purification. Thus, they were never tested.

AviTag Affinity Reagents

For the third and final generation of affinity reagents, the SNAP-tag based fusion proteins are tagged with AviTag in place of FLAG-tag. AviTag is a fifteen amino acid sequence (GLNDIFEAQKIEWHE) recognized by the enzyme biotin ligase (BirA), which can conjugate biotin onto the lysine residue of the tag *in vitro* (27, 28). Streptavidin or avidin resin can then capture these biotinylated constructs for subsequent use in pulldown experiments. As with SNAP-tag labeling, biotinylation of the affinity reagent can be monitored by intact mass MS. This system is convenient because the biotin-Streptavidin interaction of the bait and resin is highly resistant to a variety of conditions and wash steps, which was a particular limitation of using FLAG-beads. Because of this versatility, selective elution steps may allow for the preservation of bait-resin interactions while prey proteins are separated. Reducing the amount of affinity reagent in the protein sample space should increase the identification of prey proteins as high amounts of bait will not drown out the rest of the signal. In addition, these types of elution steps can allow for in-solution trypsin digestion, which will overcome the need for SDS-PAGE separation and subsequent sample division.

Condition Optimization with AviTag Constructs

To facilitate this type of immobilization and elution system, we focused on optimizing and testing a variety of elution conditions with these new affinity reagents to determine bait retention, elution efficiency of captured protein, and downstream effects of the particular elution with regards to trypsin digestion and ESI-MS signal suppression. Elution with low pH glycine, acetic acid, and phosphoric acid showed poor disruption of bait and prey interaction as analyzed by immunoblot analysis of pY1058 EGFR during BAR pulldowns of EGF-stimulated HeLa cell lysate. Elution with LDS was promising as it was most effective at eluting EGFR from the bait and resin, while only leaching about 2% of the affinity reagent from resin even with 4X LDS (Figure 1-9). A single detergent removal column appeared sufficient in removing enough LDS for efficient trypsin digestion of captured pY1058-EGFR as immunoblot signal disappeared entirely following overnight incubation. However, SILAC analysis of untreated lysate comparing LDS elution with detergent removal to in-solution digest with urea denaturation showed variable signal loss (5-90% depending on the peptide) when using the removal columns.

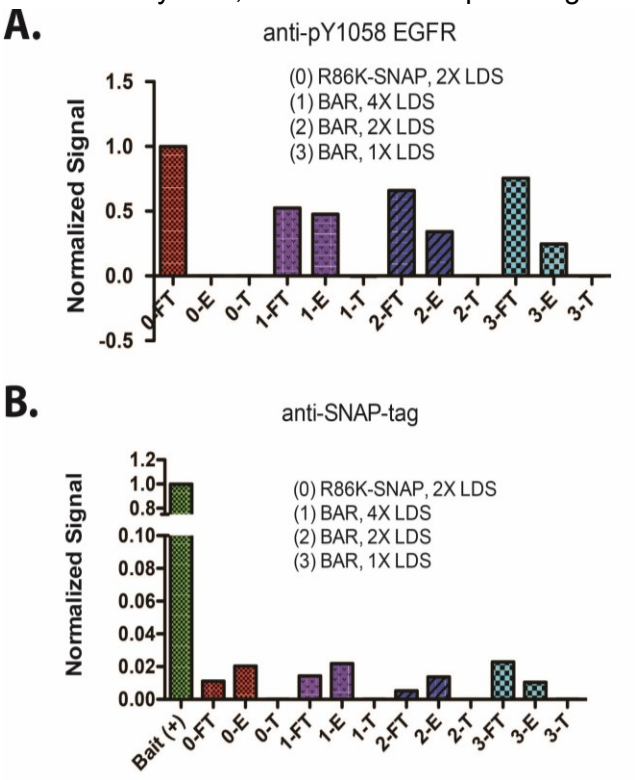


Figure 1-9. Immunoblot analysis of prey elution and bait retention on resin as well as downstream effects on trypsinolysis (Flowthrough=FT, Elution=E, After trypsin=T). (A) Bar graph showing normalized phosphotyrosine1058-EGFR signal for each pulldown step. Condition 0 is the negative control as the R86K-SNAP reagent should not bind EGFR. (B) Bar graph showing normalized SNAP-tag signal to evaluate bait retention at each step of the experiment.

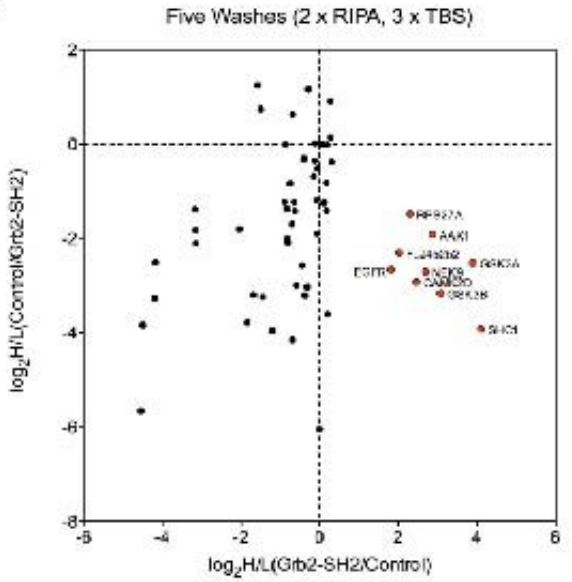
Establishing Baseline Proteomic Profiles of AviTag Affinity Reagents

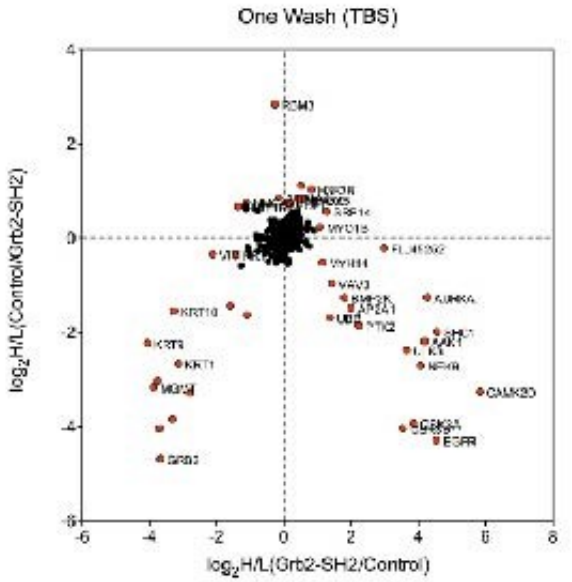
Improving the pulldown conditions with these AviTag reagents offers a more rapid profiling method to build upon the results with the FLAG-tagged affinity constructs. Initial test pulldowns with the AviTag BAR were initially very clean so optimization with regards to washing and addition of bait to lysate was required. We also experimented with different lysis conditions including

reducing the amount of detergent and using a Dounce tissue grinder versus cell scraping. The Grb2 BAR was profiled against the SH2 phosphotyrosine dead control in duplex SILAC experiments in EGF-stimulated HeLa cell lysate in various optimization studies. We reduced the washing steps from five washes all the way down to one, achieved in a single washing and SILAC channel mixing step. With regards to lysis, the amount of detergent was reduced from 1% to 0.5% and a Dounce tissue grinder was used in one set of forward-reverse SILAC experiments. Finally, instead of adding the lysate to Streptavidin-immobilized BAR, we included the BAR in the lysis buffer for immediate incubation of the affinity reagent with potential substrates. The lysate was then clarified by centrifugation and added to Streptavidin resin for incubation.

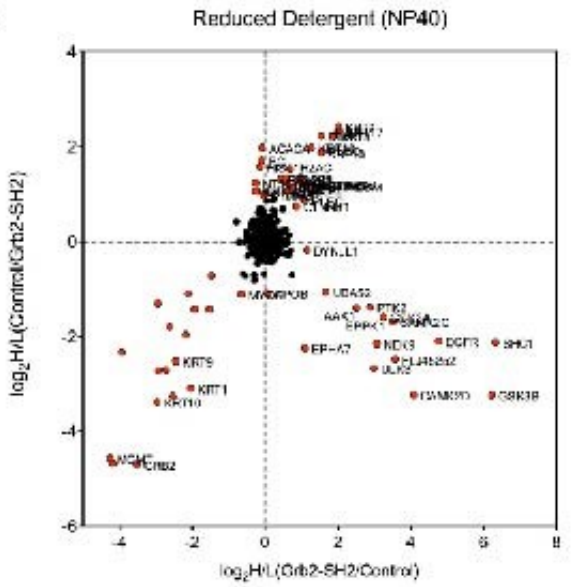
As expected, reducing the stringency of the wash step lead to an increase in the number of protein hits, most notably in the background. As seen in Figure 1-10A, five washes using a mixture of detergent-containing lysis buffer and TBS yielded a very clean and empty proteomic result. Reducing this to a single wash step, essentially merging it with the step to mix the SILAC channels, yielded an increase in background and significantly enriched proteins by the Grb2 BAR over the control. Reducing the amount of detergent in the lysis buffer and lysing with a Dounce homogenizer did not result in an appreciable difference in enrichment (Figure 1-10B). Changing the addition time of the bait lead to a slight increase in captured targets, which lead us to finalizing our pulldown conditions to direct addition of the affinity reagent at the time of lysis and only one wash step.

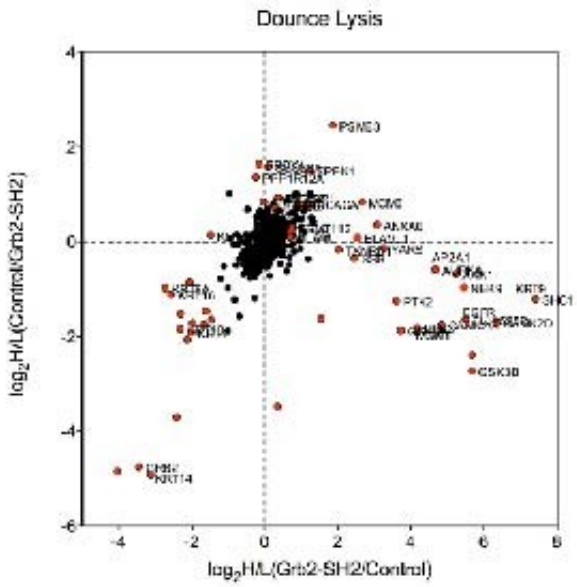
A


Five Washes (2 x RIPA, 3 x TBS)
log2H/L(Control/Grb2-SH2)
log2H/L(Grb2-SH2/Control)


One Wash (TBS)
log2H/L(Control/Grb2-SH2)
log2H/L(Grb2-SH2/Control)

B


Reduced Detergent (NP40)
log2H/L(Control/Grb2-SH2)
log2H/L(Grb2-SH2/Control)


Dounce Lysis
log2H/L(Control/Grb2-SH2)
log2H/L(Grb2-SH2/Control)

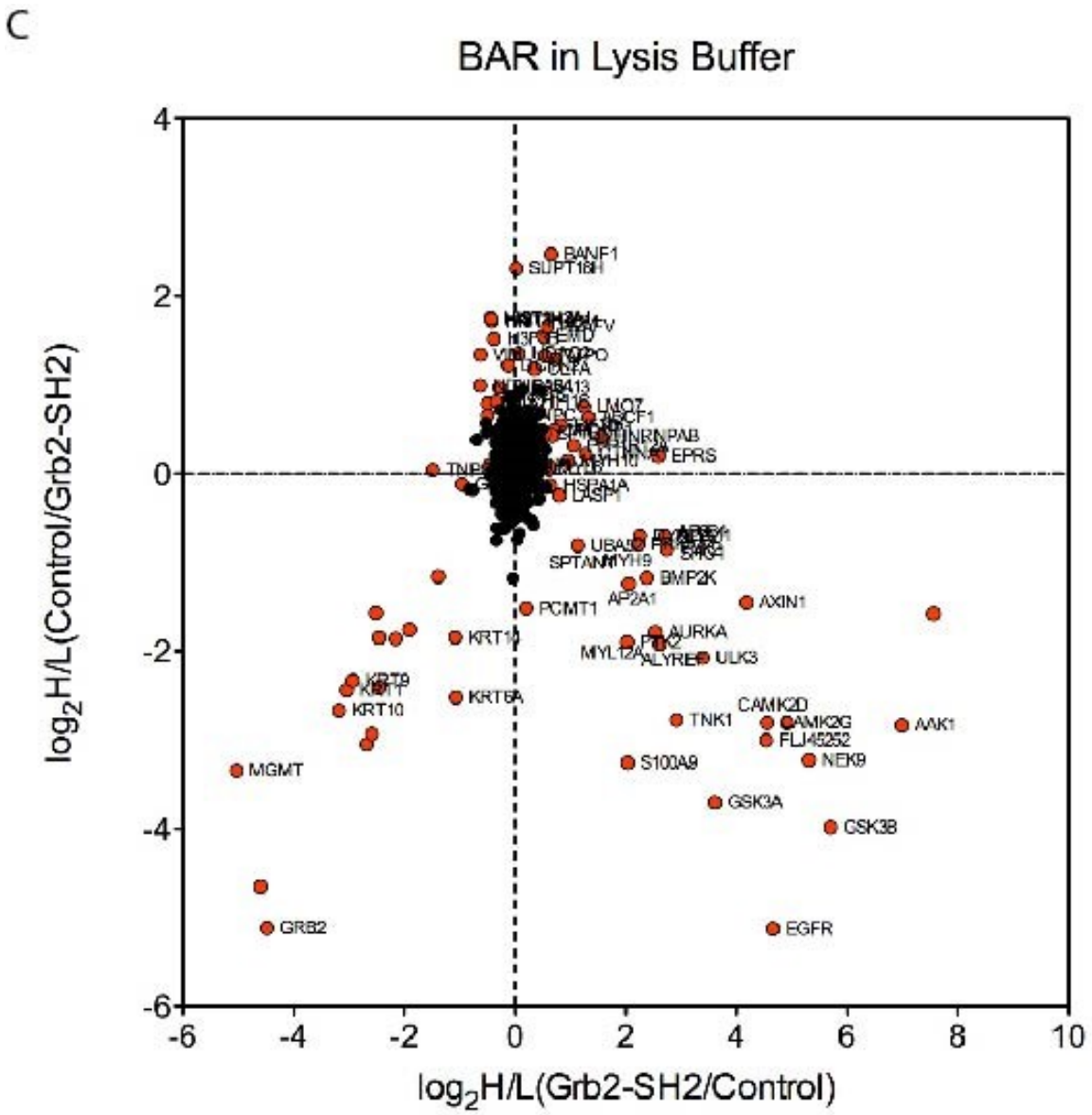


Figure 1-10. Scatter plots summarizing the AviTag pulldown optimization for washing steps (A), lysis (B), and timing of bait addition to lysate (C). Proteins that were significant hits in at least one of two replicates as determined by the Significance B calculation in Perseus are highlighted in red.

With these set of conditions in hand, we looked to replicate some of the results from the FLAG-tagged constructs and to establish baseline profiles for both bivalent and monovalent affinity reagents in EGF-stimulated HeLa cell lysate. At this point, only a Grb2 BAR had been used, so expanding the suite of reagents for both the secondary binding ligand and the kinase inhibitor was also prioritized.

Utilizing the Modularity of SNAP-tag to Expand Kinome Coverage

As we have demonstrated with SNAP-tag-based kinase inhibitors, a strength of this bivalent approach is the modularity of the binding domains, which can alter the specificity and subspace of targets. Shown below in Figure 1-11 is a CLP-functionalized kinase inhibitor that could be used as a ligand to change the profile of the affinity reagents. The compound is derived from a broad spectrum kinase inhibitor that enriched 235 protein kinases from MDA-MB-231 cells in a proteomics profile, including AKT kinase members that had not been previously detected by other pan-kinase inhibitors (29). Adding this inhibitor should offer differential and complementary coverage of the kinome in relationship to signaling complex composition than that of LGI alone. The kinase inhibitor was synthesized in our lab and conjugated to the CLP leaving group using click chemistry, hereafter referred to as the MG kinase probe.

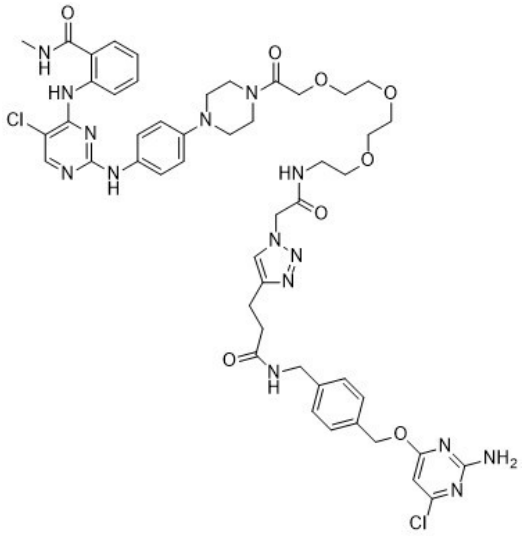


Figure 1-11. General kinase inhibitor-CLP conjugate to expand the reagent suite of BARs.

Expanding the Suite of Protein Secondary Binding Ligands

With the new elution conditions in place, we sought to generate a panel of AviTagged bivalent affinity reagents from a diverse set of proteins (Figure 1-12). In addition to targeting phosphotyrosine with SH2 domains, protein domains targeting ubiquitin and an EVH2 domain that interacts with Pro-rich areas, specifically with actin, were included to expand the scope of the secondary binding ligands. The ubiquitin-association domain (UBA) from the E3 ubiquitin-protein ligase Cbl-b, the VHS domain of signal transducing molecule 1 (STAM), which is believed to direct Stam1 towards longer ubiquitin chains and direct cleavage (30), and tandem ubiquitin-binding domains (TUBEs) from HR23 and ubiquilin comprised the ubiquitin interacting ligand type (31). The EVH2 domain was chosen from Ena/VASP-like protein, EVL. The SH2 domain ligands were from the tyrosine kinases Abl and Src, and the adapter proteins Grb2 and Crk. In addition, the R86K Grb2 mutant used in the previous FLAG system, that destroys its phosphotyrosine binding and a double mutant of Cbl-b (G943A, Y944Q) that abolishes ubiquitin binding were generated for use as control baits to establish background binding (31). Fusion proteins of this panel of

domains and SNAP-tag were generated, expressed in *E. coli,* and purified. As was done previously, both the biotinylation step of the AviTag with BirA and SNAP-tag labeling with the pan kinase inhibitor of choice were monitored by intact MS.

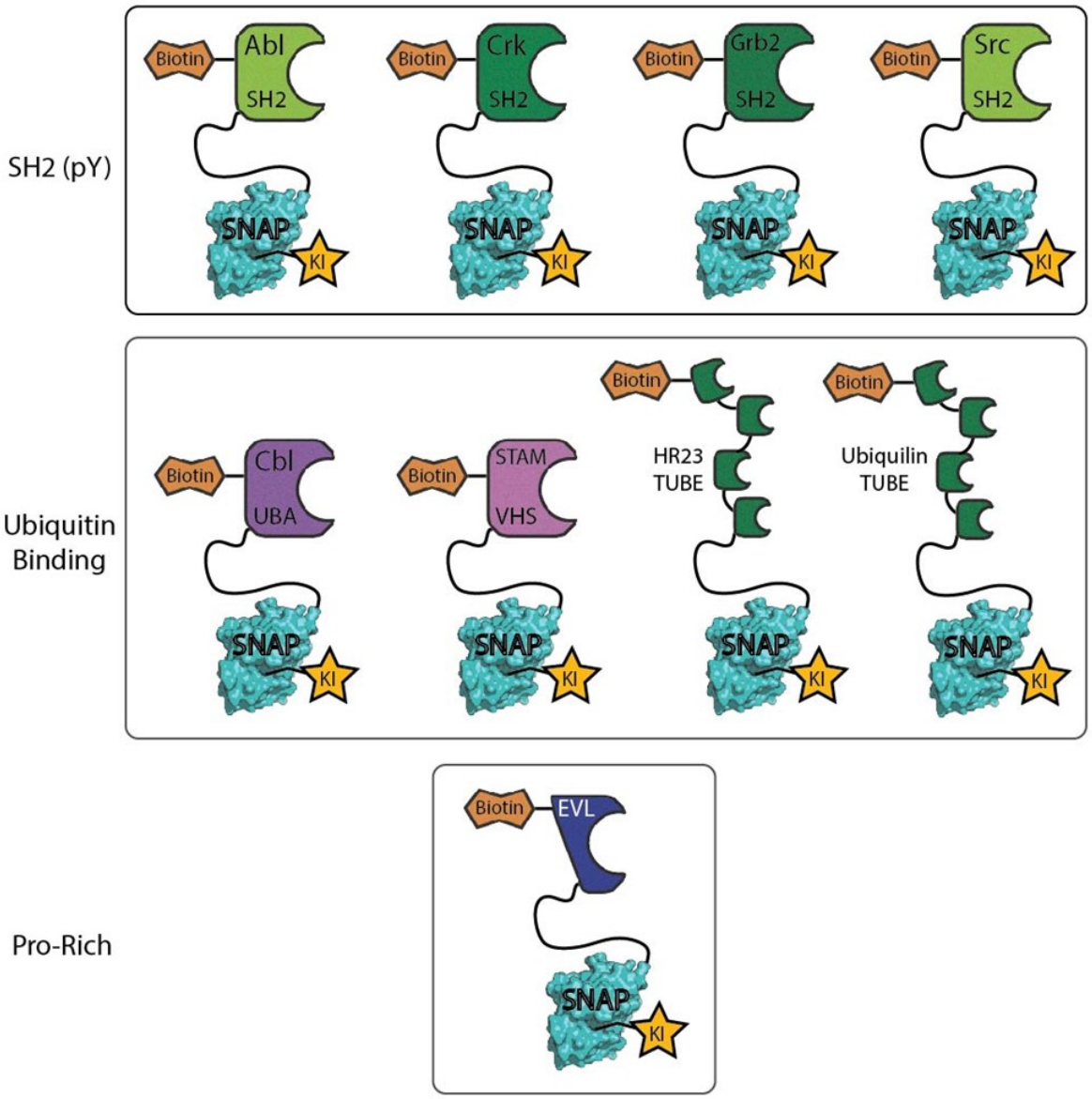


Figure 1-12. Suite of bivalent affinity reagents divided into three groups based on the domain type and the modified proteins they should specifically target. SH2 domains from four different proteins, Abl, Crk, Grb2, and Src, target phosphotyrosine. Domains that directly bind to ubiquitinated proteins, Cbl-b UBA, the TUBES comprised of ubiquitin interaction motifs from HR23 and ubiquilin, or direct other domains to ubiquitin, VHS domain of Stam1, comprise the second group. The third solitary member was the EVH2 domain of Evl that targets proline rich regions, especially actin.

Initial Profiling of Bivalent Affinity Reagent Panel in EGF-Stimulated HeLa Cell Lysate

The first step with the new affinity reagent suite was to test several members for their ability to enrich proteins versus the control bait. Dual labeled, EGF-stimulated HeLa cell lysate was probed with a select group of bivalent affinity reagents against a control bait on a separate corresponding SILAC channel to evaluate their enrichment ability and profile. Each bait was added directly to lysate and followed by addition of Streptavidin resin for subsequent binding of the bait plus captured prey protein. The panel included secondary binding ligands from the SH2 domains Grb2, Abl, and Src in addition to the UBA of Cbl-b. The two different pan kinase inhibitors, LGI and MG were also used to evaluate their effectiveness as one prong of the bivalent interaction. The number of proteins significantly enriched by these baits and the number of those that are kinases are summarized in Table 1-1 and plotted in Figure 1-13. Significant hits were determined by the significance B calculation in Perseus in one replicate in addition to the log2 intensities flipping in the label swap replicate.

Bait	Enriched Proteins	Enriched Kinases
Abl-SNAP-LGI	18	12
Cbl-SNAP-LGI	17	13
Cbl-SNAP-MG	45	34
Grb2-SNAP-LGI	28	19
Grb2-SNAP-MG	47	38
Src-SNAP-LGI	30	17

Table 1-1. Summary of the number of significantly enriched proteins and kinases for each of the BARs profiled in EGF-stimulated HeLa cell lysate. Significance was determined by the significance B calculation in at least one of the two replicates with ratio flipping occurring in the second label swap replicate.

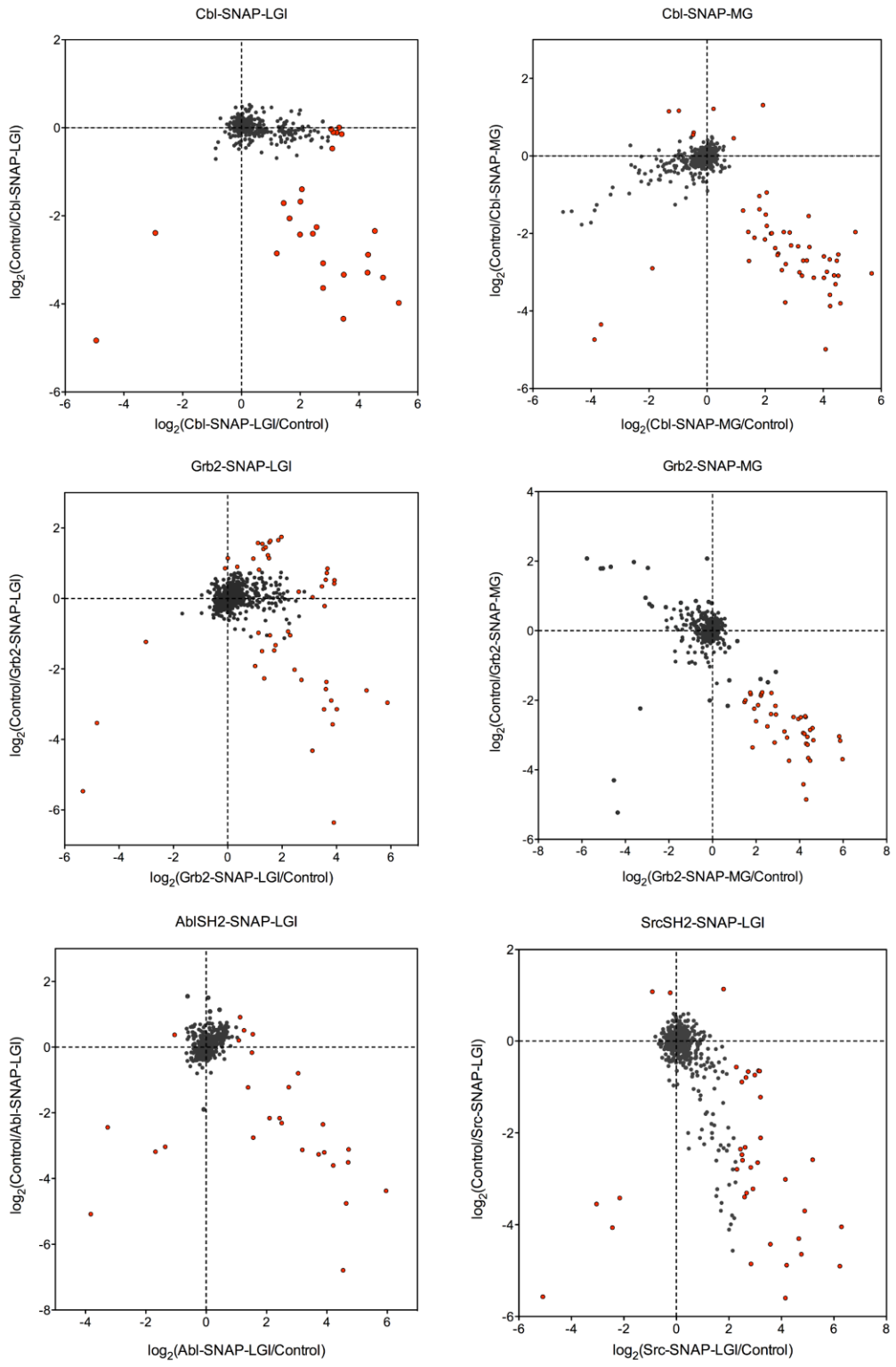


Figure 1-13. Scatter plots of the initial proteomic screen with four AviTag BAR constructs against EGF-stimulated HeLa cell lysate to compare secondary protein ligand capture and coverage of the two different kinase inhibitors.

Most of the enriched proteins were kinases, suggesting that the kinase inhibitor portion of the bivalent affinity reagents dominates. Given the higher binding affinity of the kinase inhibitor in comparison to these protein domains, which generally have low micromolar K_ds, this result was not surprising. The number of enriched hits also revealed that the bivalent affinity reagents bearing the MG kinase inhibitor had higher numbers of enriched hits. This was a result of not only expanding the number of kinase targets from 13 to 34, but increasing non-kinase hits from 4 to 11 for the Cbl construct. Of the three different SH2 domains profiled, Src showed the highest level of enrichment and Abl the lowest.

EGF-Stimulated HeLa Time Course

Signaling complexes are dynamic in nature, with the composition of the proteins part of the signaling hub as well as their phosphorylation state changing as the signaling event proceeds. The dynamic ability of proteins to form these signaling complexes is controlled in part by phosphorylation and the changes that this post-translational modification can impart, which can include altered activity, changes in conformation, and binding through topology changes (33, 34). As such, the level of phosphorylation changes during a signaling event. Because phosphorylation plays such a prominent role, especially in EGF signaling, it seemed natural to further evaluate the panel of bivalent affinity reagents in their ability to enrich phosphosites as signaling propagates from start to later time points.

Based on our initial proteomic results, the kinase inhibitor MG was chosen as the small molecule of choice moving forward as it consistently displayed broader kinome coverage than LGI. BARs of all nine secondary binding ligands were assembled with MG and then profiled in a brief time course of EGF stimulation. Triple SILAC-labeled HeLa cells were treated with EGF for 0, 15, and 30 min as shown in Figure 1-14, lysed, and then exposed to the given BAR. The resulting elutions from all three channels were mixed, trypsinized and then either processed directly on StageTip for background protein analysis, or further enriched for phosphopeptides with immobilized metal affinity chromatography (IMAC). Each time point was run in duplicate using this triplex SILAC experimental setup.

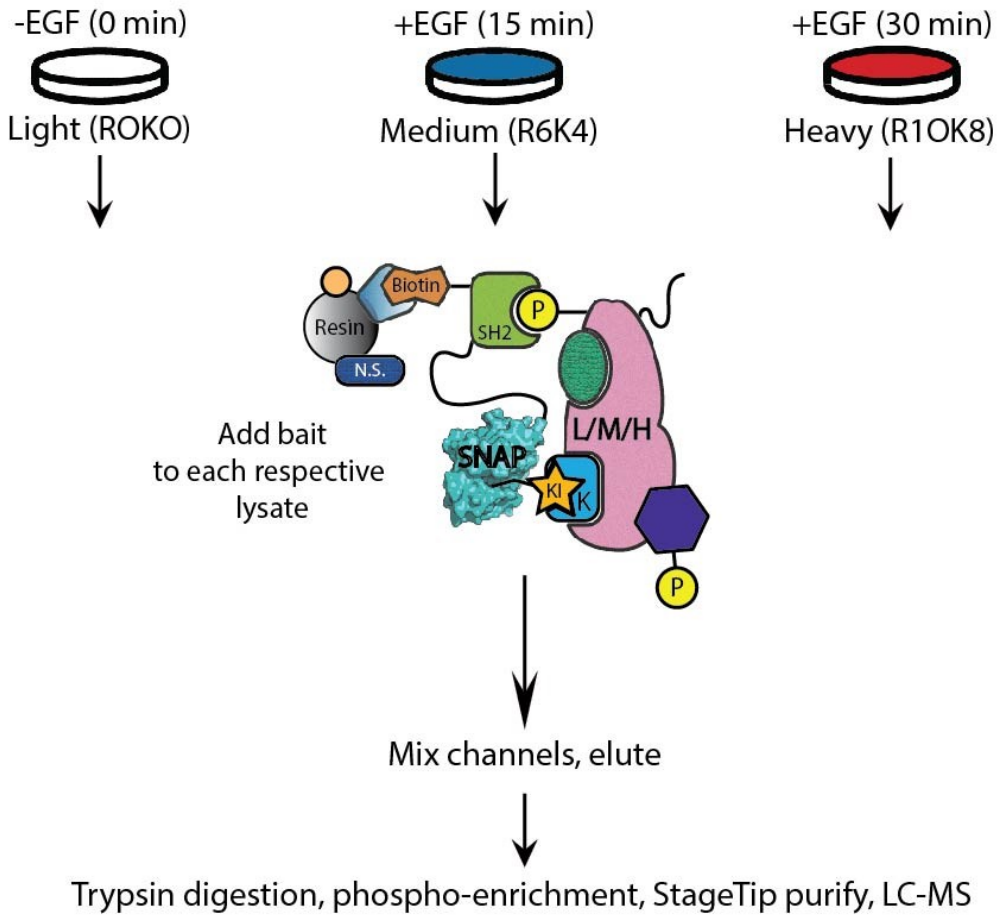


Figure 1-14. Schematic of a triple-label SILAC experiment probing the interactions of a bivalent affinity reagent at different time points throughout the course of EGF stimulation in HeLa cells. Each member of the nine BAR panel was exposed to EGF-stimulated HeLa cell lysate from three different time points. These three separate pulldowns were then mixed and eluted for subsequent processing and MS analysis.

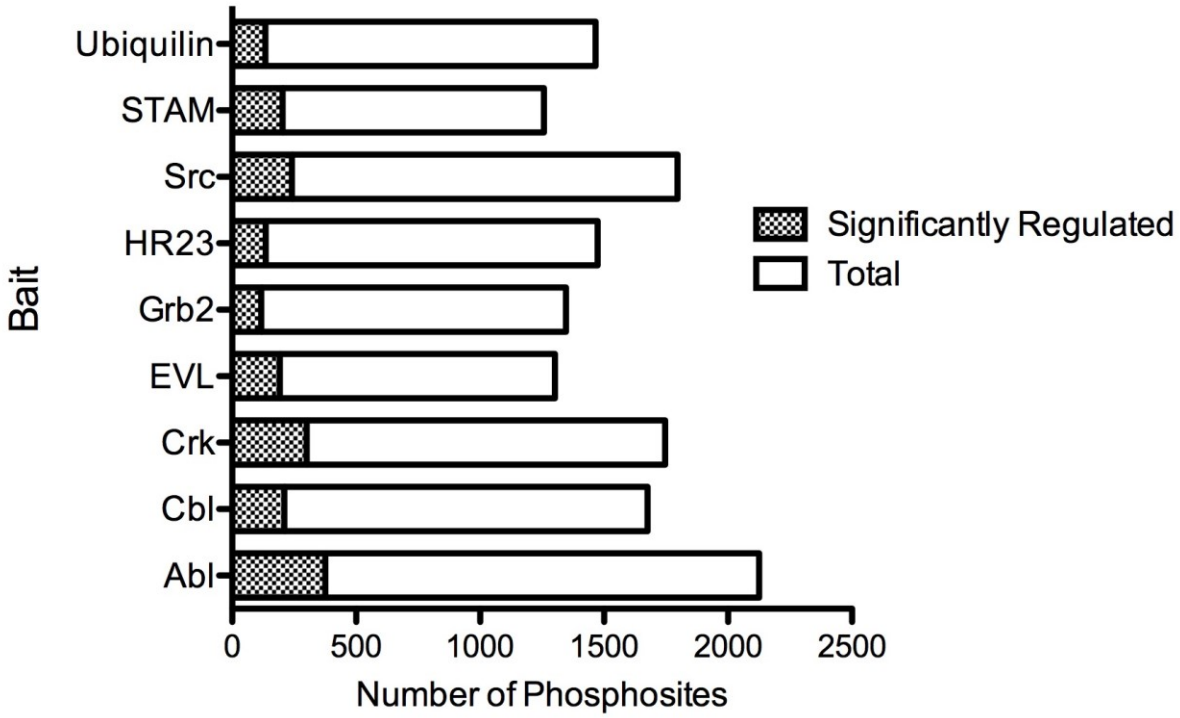


Figure 1-15. Graph showing the number of phosphosites quantified in total for each type of bivalent affinity reagent as well as the fraction of those sites that are significantly regulated for at least one time point in EGF-stimulated HeLa cell lysate.

For each BAR, the phosphosites were considered significantly regulated if the magnitude of the log2 ratio from one time point to another was >0.5 (approximately 50% change) in both the forward and reverse replicates. Changes from 0 to 15 min and 0 to 30 min were considered. The results from this analysis are summarized in Figure 1-15, which lists the total number of phosphosites quantified for each BAR as well as the subset of those sites that were significantly regulated by the criteria above. Every member of the BAR panel uses the kinase inhibitor MG for kinome coverage, so differences in enrichment are a consequence of the binding ability of the BAR's protein domain. These interactions would include direct protein-BAR domain interactions in addition to any proteins captured through a bivalent mechanism. Interestingly, the Abl SH2 BAR had the highest number of significantly regulated phosphosites even though it showed the most limited capture ability based on the initial profiling results (See Table 1-1). The SH2 domain from Grb2 and the TUBEs from HR23 and Ubiquilin gave the most limited coverage of significantly regulated phosphopeptides when used as secondary binding ligands.

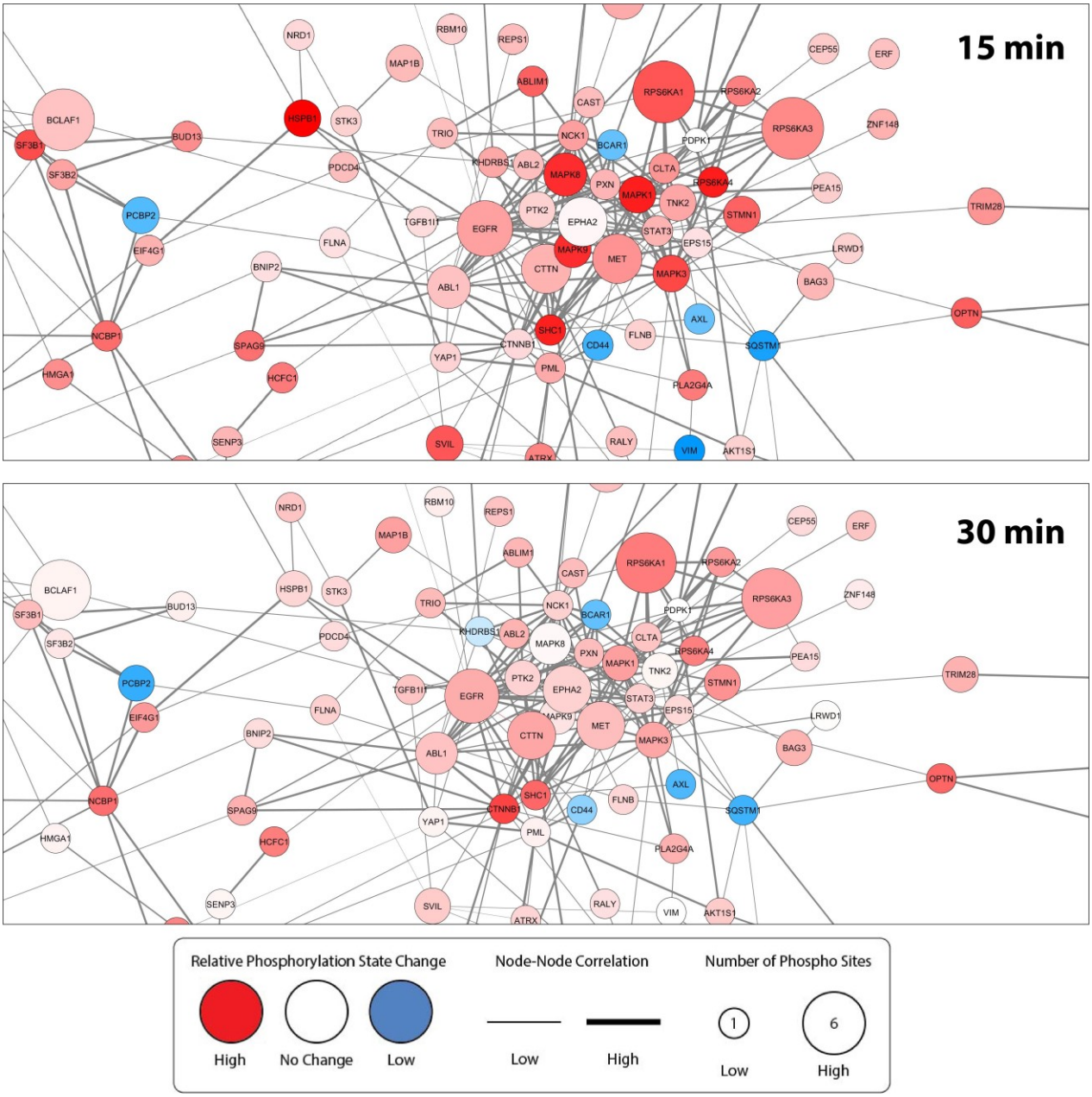


Figure 1-16. Network of Abl-SNAP-MG enriched proteins that showed temporal changes at 15 and 30 min compared to no EGF stimulation. Each node is color coded according to the change in enrichment compared to no stimulation (0 min). The number of significantly regulated phosphosites is illustrated by the node size. Edge width corresponds to the strength of the node-node interaction as determined by the combined score calculated by STRING.

The significantly regulated phosphosites from the bivalent Abl construct were analyzed for their connectivity using the STRING Database and then plotted using Cytoscape (Figure 1-16). Nodes were color coded according to the relative difference in enrichment at 15 or 30 min

compared to no stimulation. Node size corresponds to the number of phosphosites for that protein that were significantly regulated at these times. The width of the edge between nodes represents the confidence in that interaction as determined by the combined score calculation in STRING. At 15 min, there are a large number of phosphosites that have been significantly upregulated as would be expected of a signaling event such as EGFR activation. By comparison, many of these sites are still upregulated at 30 min, but the change has decreased, representing an overall downregulation of these signaling sites from 15 min post EGF addition to 30 min.

Secondary Binding Domain Specificity Profiling Through Inhibitor Competition

Both the initial profiling and the time course studies confirmed that the ATP-competitive inhibitor was the dominating binding interaction of the two ligands on the BAR. This result was not surprising considering the difference in average strength for these two types of interactions. Because of this bias, we noticed it was difficult to assess the contributions of the protein domain towards each BAR binding profile. A simple way to delineate these differences would be to profile each of the nine secondary binding ligands as a monovalent bait, one without the kinase inhibitor conjugated to the SNAP-tag active site. However, we also wanted to address the main hurdle, which was identifying any de facto bivalent targets. This monovalent profiling methodology would eliminate any possible unique bivalent interactions to each BAR. To circumvent this, we chose to profile some of the BARs plus their monovalent pieces in EGF-stimulated HeLa cell lysate, but added 20-fold excess of free kinase inhibitor MG to nullify the binding contributions of the inhibitor. Because bivalent interactions with a multiprotein signaling complex should be much tighter than any of the monovalent ones, we rationalized that these would remain despite the excess of competitive inhibitor. For the monovalent halves, the secondary binding domains alone would mostly be unaffected by inhibitor competition. The monovalent kinase inhibitor SNAP-tag conjugate with a dead protein domain should have nearly all of its binding muted. With these parameters, we profiled BARs and their monovalent counterparts for Abl SH2 and Cbl-b UBA in a triple-label SILAC experiment.

To elucidate the contributions of the kinase domains, we compared the phosphosites enriched with a log2 value greater than 0.5 by the BAR versus the monovalent version of the kinase inhibitor ligand, which essentially would work like a dead bait. Based on this, we were able to generate enrichment lists for the two domains profiled. Interaction networks for Abl SH2 and Cbl-b UBA are shown in Figures 1-17 and 1-18, respectively.

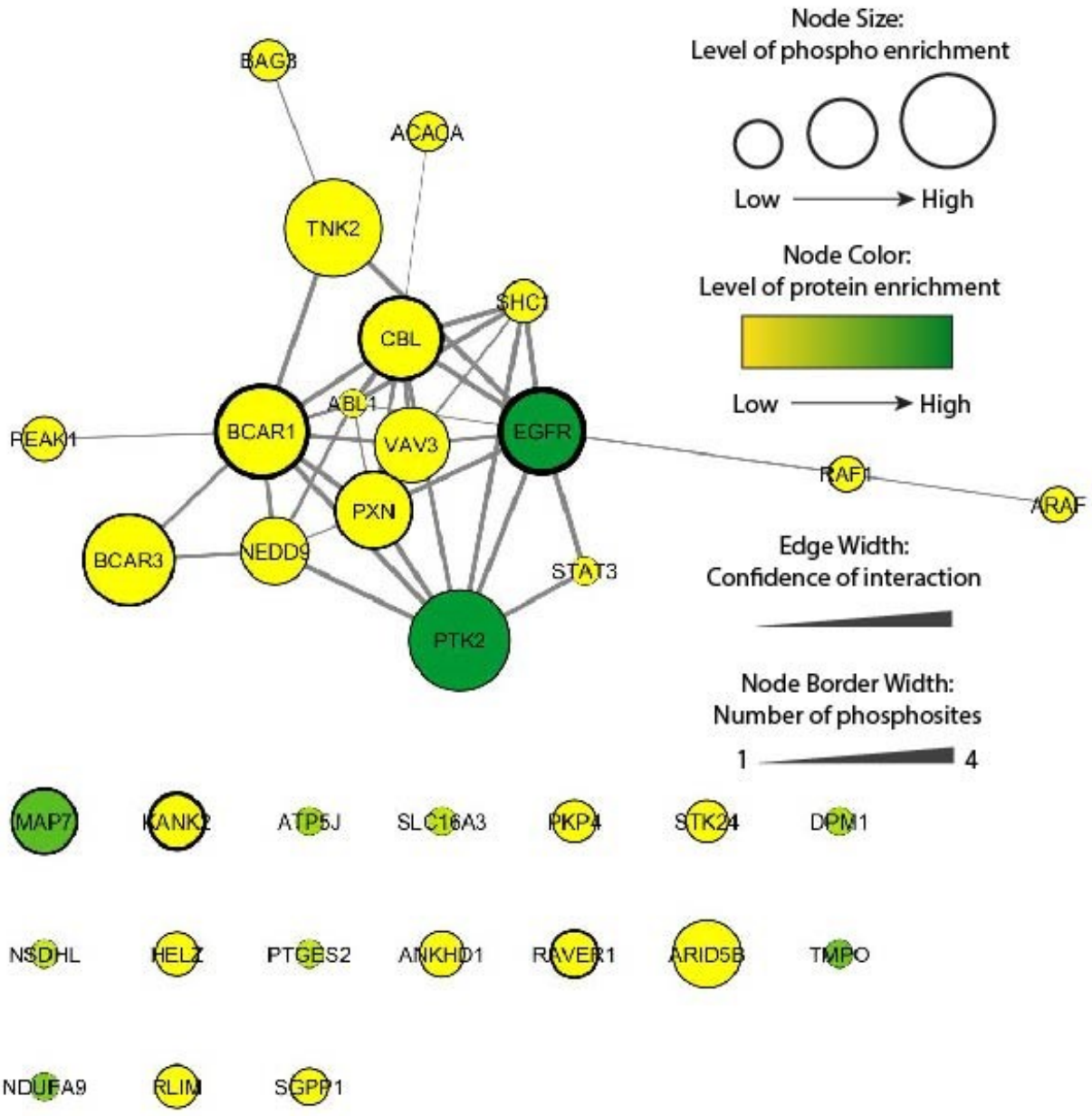


Figure 1-17. Network profile of enriched hits of bivalent Abl SH2 affinity reagent with competition against 20X kinase inhibitor. Node color corresponds to the fold enrichment for the protein. If the protein was not quantified, it is assigned an enrichment of zero. The size of the node indicates the relative level of phospho enrichment for a given protein. The border thickness of the node correlates to the number of phosphosites quantified for the protein and the edge width corresponds to the node interaction correlation.

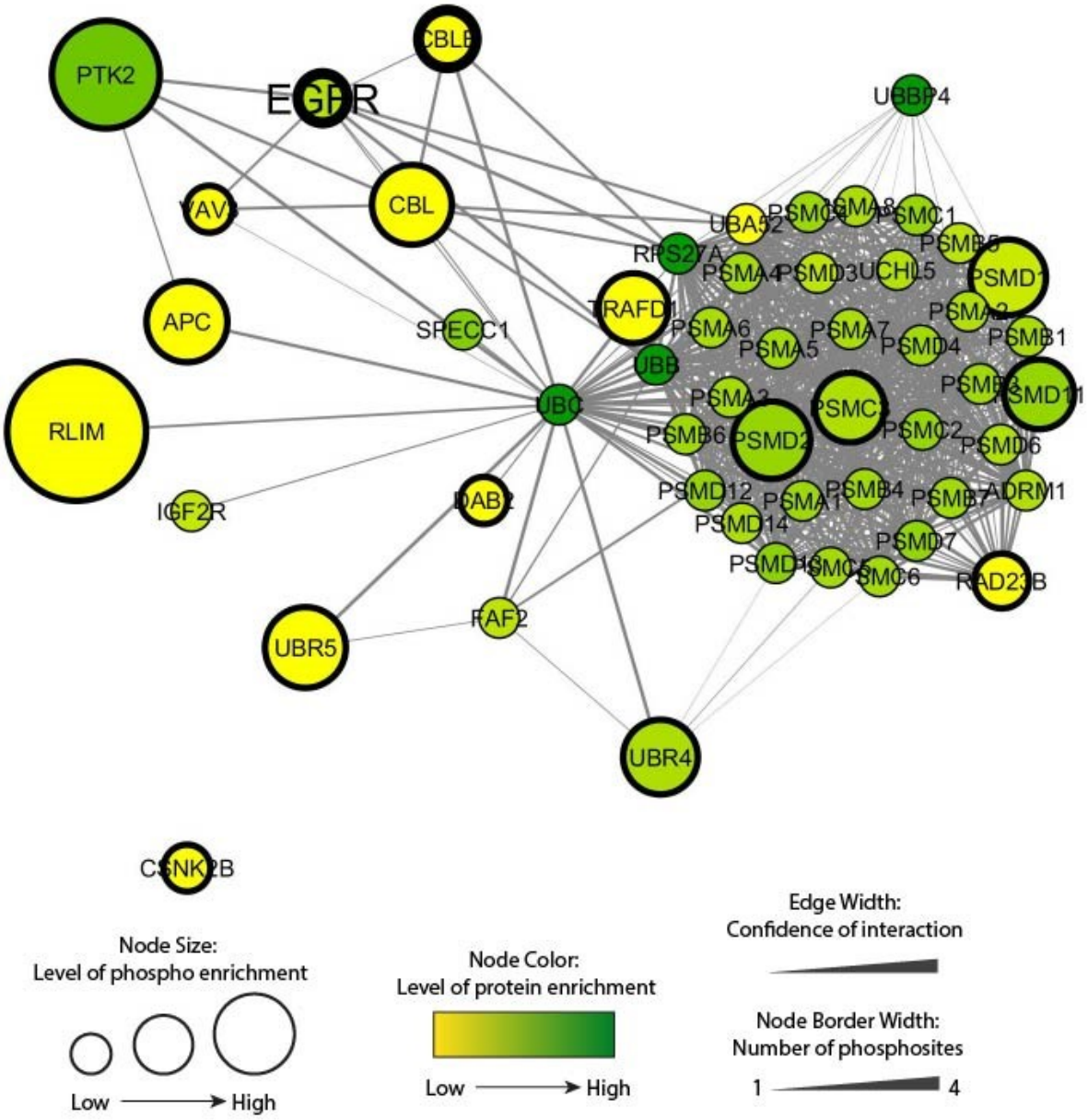


Figure 1-18. Network profile of enriched hits of bivalent Cbl UBA affinity reagent with competition against 20X kinase inhibitor. Node color corresponds to the fold enrichment for the protein. If the protein was not quantified, it is assigned an enrichment of zero. The size of the node indicates the relative level of phospho enrichment for a given protein. The border thickness of the node correlates to the number of phosphosites quantified for the protein and the edge width corresponds to the node interaction correlation.

The interaction list for the Abl and Cblb BAR are different in this competition experiment, validating that each secondary ligand has its own unique contribution. Both Abl and Cbl-b BAR enriched phosphosites on EGFR, Vav3 and PTK2, but their enrichment differs besides that. The Abl BAR enriched phosphosites and proteins from paxillin, which binds to PTK2 or FAK, BCAR1, and Nedd9, all involved in cell adhesion and focal adhesions. It also enriched EGFR adapter proteins such as Shc1 and Vav3. The Cbl-b BAR enriched proteins and phosphosites were primarily from the proteasomal complex in addition to several E3 ligases, UBR4, UBR5, and RLIM.

Bivalency Profiling

Although we were able to determine differences in enriched targets between the protein domains of several BARs, these studies did not yield any evidence of bivalent interactions. The enrichment of the BAR was compared to each monovalent piece, but there were not any consistent proteins or phosphopeptides that appeared with higher fold enrichment for both sides of the comparison. Given that bivalent enrichment was the driving point behind this approach, we sought a variety of profiling experiments to facilitate bivalent capture of signaling complexes. To this end, bivalency comparison pulldowns in addition to other experiments with increased free inhibitor competition with 50 and 100-fold excess relative to the bait, increased washes to remove weaker binders, increased bait amount, and probing at different stimulation time points were done with several different BARs. Each of these experiments utilized triplex SILAC to compare each piece, bivalent plus two monovalents, on a separate SILAC channel. The results of the most simple bivalency experiment, comparing the bivalent reagents to each monovalent portion on separate, triplex SILAC lysate is summarized in Figure 1-19.

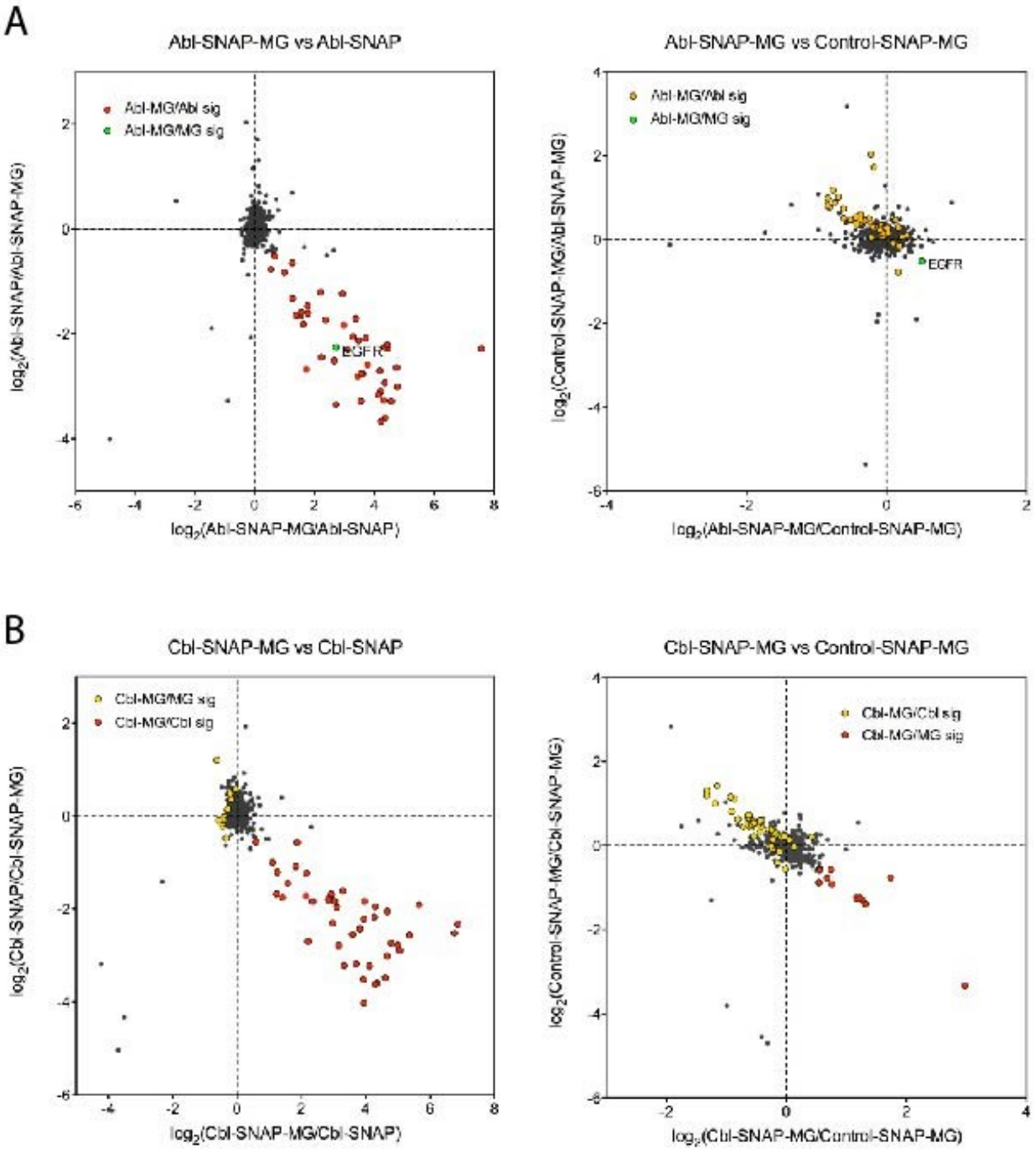


Figure 1-19. Scatter plots of bivalency experiments comparing the BAR to its respective monovalent pieces, the protein domain or the kinase inhibitor, for two different reagents, (A) Abl SH2 and (B) Cbl-b UBA. Proteins with a log2 > 0.5 enrichment for the BAR over the corresponding monovalent are colored in red while those that showed significant enrichment versus the other monovalent piece are in yellow. Proteins that had higher BAR enrichment versus both monovalent forms are in green.

The scatter plots compare the BAR to either monovalent secondary protein domain (Abl/Cbl-SNAP) or kinase inhibitor (Control-SNAP-MG). On each of these comparison plots, the proteins that the BAR enriched with a log2 > 0.5 than the monovalent form are highlighted in red. In addition, those proteins that had this higher BAR capture for the other monovalent piece are colored in yellow. When these two events merge, where the BAR enriches a protein to a greater extent than both monovalent portions, a bivalent interaction should be occurring and is highlighted in green. As evident by this figure, there is only a single such point, EGFR, in the Abl BAR profiling. Furthermore, these plots reveal that each prong of the BAR seems to act independently, enriching its own subset of proteins. The hits from the BAR vs Abl-SNAP in the first plot, primarily kinases, migrate towards the centroid in the second plot of the BAR vs Control-SNAP-MG, and vice versa. The same holds true for the comparison with Cbl BAR and its monovalent forms. All of the various conditions mentioned previously yielded similar results. Unfortunately, these reagents behaved like dual affinity reagent ligands that could enrich different subsets of proteins rather than a cooperative bivalent binding reagent.

III. Conclusion

By building upon the bivalent kinase inhibitor approach, we have generated a set of unique affinity reagents for proteomic profiling. While the original vision for these reagents was to bivalently enrich multiprotein signaling complexes with each ligand binding to a separate member of the complex, numerous experiments did not show this ability. The profiling suggested that each of the two ligands acts independently and captures a subset of proteins unique to its own binding profile. Although the kinase inhibitor portion of these reagents is the dominant ligand in this setup, competition experiments illustrated that each protein ligand contributes unique interactors to the total enrichment profile of the reagent. Rather than bivalent affinity reagents, these could be used as dual capture reagents for enrichment of two different protein or phosphosite subsets. Because the domains and small molecule portion of the reagent can be swapped out as desired, the type of proteins targeted can be changed as necessary.

IV. Methods

Protein Design, Expression, and Purification

Plasmids of pE57, pE59, and E40 were generated by Genscript and amplified along with SNAP-tag (pss26b) using PCR. The DNA for the affinity reagents were amplified from gBlocks containing the protein domain of interest (Abl SH2, Src SH2, Grb2 SH2, Crk SH2, Cbl-b UBA, VHS of STAM, HR23 TUBE, Ubiquilin TUBE, EVH2 of EVL) and a 24 amino acid GST linker. Fusion proteins between SNAP-tag and the DARPins were generated using overlap extension PCR, and then the resulting fusions were cloned into the pMCSG7 vector using ligation independent cloning (LIC). Overlap extension PCR and Gibson assembly were used for the affinity reagent constructs. The resulting vectors were transformed into *E. coli* XL1-Blue cells for subsequent isolation of plasmid DNA using Quiagen's QIAprep Spin Miniprep Kit. Sequence-verified plasmids were transformed into *E. coli* BL21(DE3) expression hosts. BL21 cells were inoculated in LB Miller broth with 100 µg/mL of ampicillin. Cells were grown at 37°C until an OD_{600} of 0.6-0.8 and then induced with 1 mM IPTG overnight at 18°C. Cells were isolated by centrifugation and then stored as a pellet at -80°C. In order to purify the protein, cells were thawed and resuspended in 10 mL of wash buffer (50 mM HEPES, pH 7.5, 10 mM imidazole) with 100 µg/mL of PMSF. Cells were lysed by sonication and the resulting lysate cleared by centrifugation at 4°C. The cleared lysate was incubated and rotated with 125 µL of 5 PRIME PerfectPro Ni-NTA Agarose (Thermo Fisher Scientific) at 4°C for 60 min. The resin was washed 10x with 1 mL of wash buffer and protein eluted with His6 elution buffer (50 mM HEPES, pH 7.5, 300 mM imidazole). Collected fractions were tested with Coomassie reagent and analyzed with SDS-PAGE. Pure fractions were desalted into storage buffer (50 mM Tris, pH 7.5, 100 mM NaCl, 5 mM DTT), aliquoted, snap-frozen, and stored at -80°C.

Preparation and Purification of Bivalent Inhibitor SNAP-tag Conjugates

SNAP-tag(wt), SNAP-pE57, SNAP-pE59, and SNAP-E40 were labeled with LGI741-CLP or LGI741-BG using the following conditions. Purified protein was incubated with 1.5-fold excess of the small molecule in labeling buffer (50 mM Tris, pH 7.5, 100 mM NaCl, 0.1% Tween 20, and 1 mM DTT) at 25°C for 2 hr. Reactions were applied to two sequential Thermo Scientific Zeba Spin Desalting columns (Thermo Fisher Scientific) equilibrated with Tris buffer in order to remove any remaining unlabeled small molecule from the resulting protein-small molecule conjugate. The concentration of the eluted proteins was quantified using Coomassie Plus, The Better

Bradford™Assay Kit (Thermo Fisher Scientific) and used in subsequent assays. Labeling efficiency was determined using ESI-MS intact mass analysis.

In Vitro Activity Assays (ERK2)

Inhibitors (initial concentration of 5-10 μM with three-fold serial dilutions for 10 data points) were assayed in triplicate against ppERK2 (final concentration = 0.32 nM) in KRB2 assay buffer (30 mM HEPES, pH 7.5, 0.6 mM EGTA, 10 mM $MgCl_2$, 0.2 mg/mL BSA, 3.75 mM DTT, 2.5 mM β-glycerophosphate), 0.2 mg/mL myelin basic protein, and 0.2 μCi/well γ^{32}P-ATP for a final reaction volume of 30 μL. The phosphorylation reaction was run for 3 hr and then terminated by spotting 4.6 μL of the reaction mixture onto phosphocellulose membrane. The membranes were washed 3 x 10 min with 0.5% phosphoric acid and dried with acetone. Radioactivity was measured by phosphorimaging with a GE Typhoon FLA 9000. ImageQuant was used to quantify the scanned images with the resulting data converted to percent inhibition. This data was then analyzed with GraphPad Prism and IC_{50} values calculated with nonlinear regression analysis.

In Vitro Activity Assays (JNK2 and p38α)

JNK2 and p38α were activated for use in activity assays as follows. The kinase was incubated in kinase phosphorylation buffer (50 mM MOPS pH 7.4, 10 mM $MgCl_2$, 1 mM DTT, 0.001% v/v Tween 20), 0.1 mg/mL BSA, 390 μM ATP, and upstream kinase (pMKK4, pMKK7 for JNK2 and MKK6 for p38α). Inhibitors (initial concentration of 5-10 μM with three-fold serial dilutions for 10 data points) were assayed in duplicate against JNK2 (final concentration = 2 nM) and p38α (final concentration 2 nM) in KRB2 assay buffer (30 mM HEPES, pH 7.5, 0.6 mM EGTA, 10 mM $MgCl_2$, 0.2 mg/mL BSA, 3.75 mM DTT, 2.5 mM β-glycerophosphate), 0.2 mg/mL myelin basic protein, and 0.2 μCi/well γ^{32}P-ATP for a final reaction volume of 30 μL. The phosphorylation reaction was run for 3 hr (p38α) or 4 hr (JNK2) and then terminated by spotting 4.6 μL of the reaction mixture onto phosphocellulose membrane. Subsequent wash steps, imaging, and analysis were performed as detailed above for ERK2.

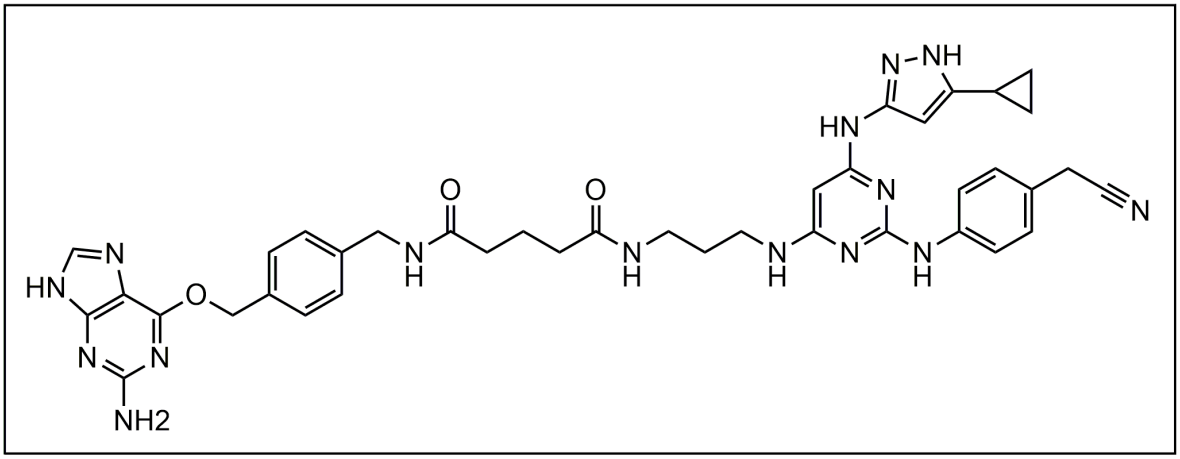
Kinase Phosphorylation Inhibition Assays

Kinase (40 nM) was incubated in phosphorylation buffer (50 mM MOPS pH 7.4, 10 mM $MgCl_2$, 1 mM DTT, 0.001% v/v Tween 20) and BSA (0.1 mg/mL) with inhibitor (initial concentration = 5-10 μM, two-fold serial dilutions for 8 data points) for 2 hr at room temperature. Upstream kinase (MEK2) was added with 5 μM ATP and the reaction was incubated for 2 hr at room temperature. The enzymatic reaction was quenched with 1X SDS and boiled for 30 min. Samples were separated by SDS-PAGE, transferred to nitrocellulose membrane, and blocked with 5% w/v non-fat milk in TBS. The membranes were probed with antibodies specific for phosphorylated ERK2 (Cell Signaling Phospho-p44/42 MAPK) and then with LI-COR IR Dye secondary antibody. Blots were fluorescently imaged with a LI-COR Odyssey Imager and quantified with Odyssey's accompanying software. EC_{50} values were calculated with non-linear regression analysis with GraphPad Prism.

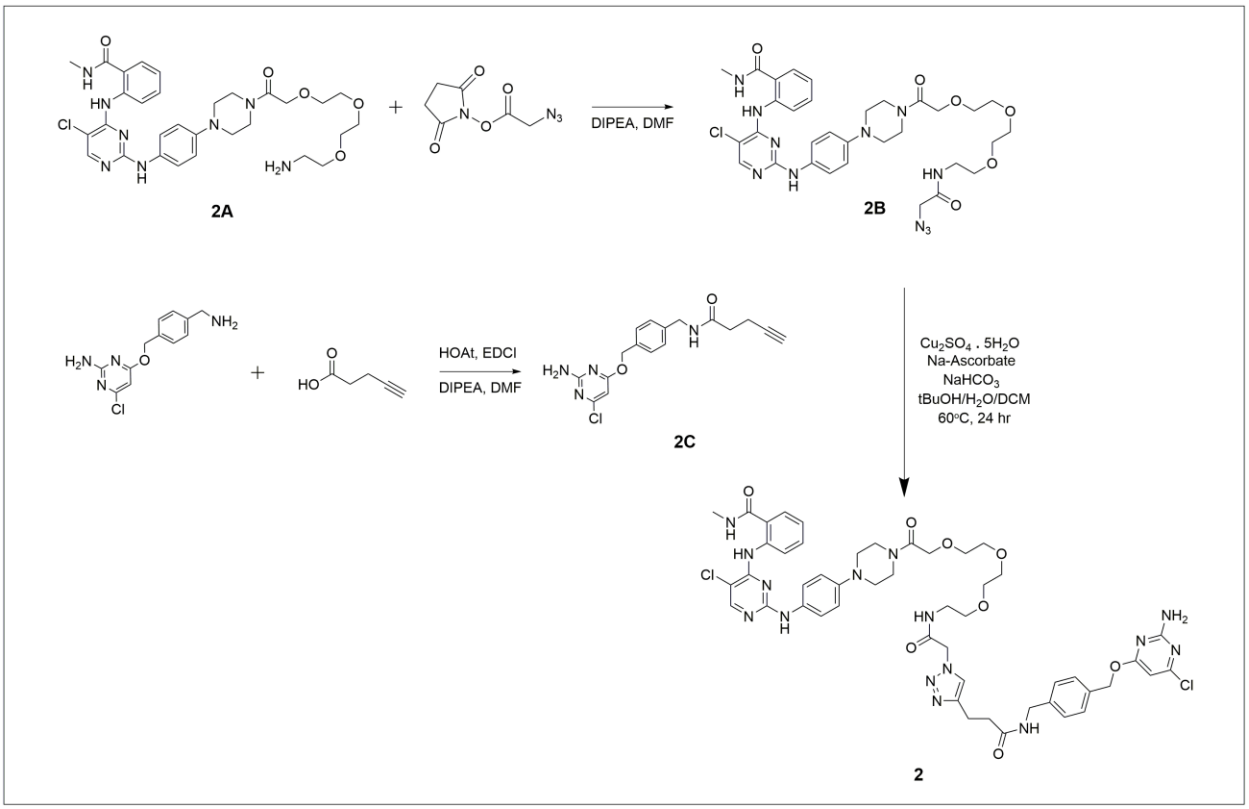
Synthesis of Kinase Inhibitor CLP and BG Conjugates

General Methods: Unless otherwise noted, all reagents were from commercial suppliers and used without further purification. Products were analyzed by ^{1}H-NMR obtained with a Bruker AV-300 or AV-301 instrument at room temperature. Mass spectrometry was done on a Bruker Esquire Ion Trap Mass Spectrometer.

General HPLC Purification Conditions: A 70 min gradient of 1:99 to 100:0 acetonitrile/H_2O, 0.1%TFA was used at 8 mL/min with 254 nm detection for 75 min.



[1]. Synthesis was perforemd by a collaborator at Novartis, Edmund Harrington, using a previously published protocol (22).



Synthetic scheme for synthesis of compound 2, MG-CLP.

[2A]. Synthesis was done by another lab member using a previously published protocol (35).

[2B]. To a stirred solution of 2A in DMF was added 1.5 molar excess of 2,5-dioxopyrrolidin-1-yl 2-azidoacetate (Click Chemistry Tools) and 3-fold molar excess of diisopropylethylamine. The reaction was stirred overnight at room temperature. The product was concentrated and purified by reverse-phase C18 HPLC.

[2C]. Synthesis was done by another lab member using method outline in the synthetic scheme above.

[2] To a stirred solution of **2B**, Cu_2SO_4 pentahydrate, sodium ascorbate and sodium bicarbonate in 1:1:1 of tBuOH/H_2O/DCM, 2C was added in 1.5 molar excess. The reaction was stirred at 60°C for 24 hr, concentrated and purified by reverse-phase C18 HPLC.

^{1}H-NMR (300 MHz, DMSO-d_6) ppm 2.59 – 2.64 (t, 2 H), 2.95 (s, 2 H), 3.00 – 3.05 (t, 2 H), 3.39 – 3.43 (m, 3 H), 3.54 – 3.58 (t, 2 H), 3.62 – 3.67 (m, 4 H), 3.72 (br, 6 H), 4.34 (m, 3 H), 5.08 (s, 2 H), 5.31 (s, 2 H), 6.09 (s, 1 H), 7.09 (br, 2 H), 7.19 – 7.24 (m, 2 H), 7.25 – 7.37 (m, 4 H), 7.42 – 7.48 (m, 1 H), 7.68 (s, 1 H), 7.74 – 7.76 (dd, 1 H), 8.04 (br, 1 H), 8.38 (br, 1 H), 8.59 – 8.62 (br, 1H).

Preparation of AviTag Bivalent Affinity Reagents

Biotinylation of AviTag SNAP-tag fusion proteins was done by incubating 50 uM of the protein in 50 mM bicine buffer, pH 8.3 with 10 mM ATP, 10 mM MgOAc, 500 uM d-biotin and 1 uM BirA biotin-protein ligase overnight at 4°C. This mixture was then incubated with 75 uM of the kinase inhibitor LGI-BG or MG-CLP with 1 mM DTT overnight at 4°C. To this was added 20 mM imidazole and 100 uL of HisPur Ni-NTA resin (Thermo Fisher Scientific). The slurry was incubated for 1 hr at 4°C and then washed five times with TBS on a column. The labeled protein was eluted off resin with elution buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 200 mM imidazole) and exchanged into 50 mM Tris pH 7.5, 150 mM NaCl using three sequential buffer exchanges in Slide-a-Lyzer dialysis cassettes (Thermo Fisher Scientific) or with a GE Healthcare Sephadex G-25 PD-10 desalting column (Thermo Fisher Scientific). Labeled proteins were analyzed for biotinylation of the AviTag and SNAP-tag labeling via intact ESI-MS. For this analysis, proteins were desalted into 50 mM ammonium bicarbonate, 0.1% formic acid with two sequential Zeba Spin 7K desalting columns (Thermo Fisher Scientific). Proteins were injected directly onto a Bruker Esquire LC-Ion Trap and labeling estimated by analysis of the deconvoluted spectra.

Cell Culture and Harvest

HeLa cells were grown in custom-made media with -Lys/-Arg Dulbecco's modified Eagle's medium, DMEM (Caisson Labs) supplemented with 200 ug/mL of L-proline, SILAC amino acids (0.2 mM Arg0/Lys0 for the light label, 0.2 mM Arg 6/Lys4 for the medium label, and 0.2 mM Arg10/Lys8 for the heavy label (Cambrdige Isotope Labs), Penicillin-Streptomycin (100X, 10,000 U/mL, Gibco), and 10% dFBS (Caisson Labs). Cells were cultured in SILAC media for at least

five cell doublings for optimal labeling. HeLa cells were grown to greater than 90 percent confluency and then serum-starved overnight. The following day, cells were stimulated with 150 ng/mL of EGF in 20% FBS supplemented media. Cells were washed twice with dPBS (Gibco) and lysed with a cell scraper in modified RIPA buffer (50 mM Tris pH 7.8, 150 mM NaCl, 1 mM EDTA, 1% NP-40) supplemented with Pierce Protease Inhibitor Cocktail (10X, Thermo Fisher Scientific) and Pierce Phosphatase Inhibitor Cocktail (10X, Thermo Fisher Scientific). For direct bait-lysate addition, the 0.5-2 nmol of bait (depending on the study) was added to 500 uL of lysis buffer for lysis of a 15 cm plate of cells. The lysate was vortexed intermittently on ice five times and then clarified by centrifugation at 10,000xg for 15 min at 4°C. Protein concentration was determined using the Pierce 660 nm Assay Reagent (Thermo Fisher Scientific) and then adjusted to 3 mg/mL with lysis buffer.

Affinity Purification and Sample Digestion

For bait preparation, an appropriate amount of affinity reagent (0.5 nmol -2 nmol, depending on the type of study) was added to Pierce Streptavidin Resin (20 uL per pulldown, Thermo Fisher Scientific) and incubated overnight at 4°C. Confirmation of complete bait capture was assessed by NanoDrop of the flowthrough. The resin was washed three times with TBS pH 7.5 and then added to lysate. For the direct bait to lysate addition, 20 uL of Streptavidin resin were added to the lysates already containing the appropriate bait. The slurry was incubated for 3 hr at 4°C and then washed once with TBS pH 7.5 while mixing the SILAC channels from each respective pulldown. Proteins were eluted from resin with 100 uL of 1X LDS for 30 min at 60°C. The supernatant was reduced with 10 mM DTT for 30 min at 37°C in a ThermoMixer and then alkylated with 20 mM CAM for 20 min at 37°C. LDS was removed from the sample with a Thermo Scientific Pierce Detergent Removal Spin Column (Thermo Fisher Scientific) pre-equilibrated in 100 mM tetraethylammonium bicarbonate (TEAB, Sigma). The detergent-free sample was diluted with 35 µL of fresh 8 M urea and digested with 1 µg of Lys-C (Wako) for 2 hr at 37°C. This solution was further diluted with 115 µL of TEAB and digested with 1 µg of Pierce Trypsin Protease, MS Grade (Thermo Fisher Scientific) overnight at 37°C. The samples were acidified with trifluoroacetic acid (TFA) and then diluted with 250 uL of 5% acetonitrile, 94.9% water, 0.1% TFA. Peptides were extracted with StageTips or used for additional phophopeptide enrichment (35).

Phosphopeptide Enrichment with IMAC

Peptides were desalted with Oasis HLB C18, 10 mg cartridges (Waters), dried down on a speed vac, and resuspended in 80% ACN, 19.9% H_2O, 0.1% TFA (StageTip Buffer B). A slurry containing a 1:1:1 mixture of Fe^{3+}-NTA, Ga^{2+}-NTA, and PHOS-Select Iron Affinity Gel (Sigma) was prepared and 20 µL of this mixture was added to each tube of peptide. The slurry was shaken in a ThermoMixer at 25°C for 1 hour and then added directly to a StageTip equilibrated first in methanol and then in Buffer B. The IMAC resin was washed twice with 50 µL of Buffer B and the StageTip equilibrated further with 2 x 100 µL of 1% formic acid. The phosphopetides were eluted from the resin and extracted to the StageTip by addition of 100 µL of 500 mM dibasic sodium phosphate buffer, pH 7.0. This elution was repeated and the StageTip washed twice with 100 µL of 1% formic acid. Phosphopeptide results were analyzed in a single nano-LC-MS/MS run.

LC-MS/MS and Data Analysis

A Thermo-Dionex RSLCnano UHPLC connected to a 20 cm long, 360 µm OD x 100 µm ID fused silica capillary column with a 7 µm tip, self-pulled with a laser puller (Sutter), was used to separate extracted peptides. Columns were packed with 3 µm Reprosil C18 beads (Dr. Maisch). Peptides were run on a 120 min, 5% to 30% acetonitrile gradient in 0.1% acetic acid at 300 nL/min. The LC solvents used were 0.1% acetic acid (A) and 0.1% acetic acid, 99.9% acetonitrile (B). FTMS spectral scans (*R* = 30 000 at 400 *m/z*; *m/z* 350-1600; 3*e*6 target; max 500 ms ion injection time) were done using a Thermo Orbitrap Elite mass spectrometer with data-dependent analysis using Top15 selection and CID fragmentation (1*e*4 target; max 100 ms injection time). Collection was done with dynamic exclusion for 30 s and exclusion list size of 50. CID was used with normalized collision energy of 35% for 10 ms. The raw files were analyzed with MaxQuant/Andromeda(58) version version 1.5.2.8. Protein, peptide, and site FDRs of 0.01 and a minimum score of 40 for modified peptides, 0 for unmodified peptides, minimum delta score of 17 for modified peptides, and 0 for unmodified peptides was selected. MS/MS spectra were searched against the UniProt human database from July 22, 2015. MaxQuant search parameters were conducted as follows: Variable modifications included Oxidation (M), Acetyl (Protein N-term), Phospho (STY). Carbamidomethyl (C) was a fixed modification. Maximum labeled amino acids was 3, max. missed cleavages was 2, enzyme was Trypsin/P, max charge of 7, multiplicity was either 1, 2, or 3. SILAC labels were Arg0/Lys0 for light, Arg6/Lys4 for medium, and Arg10/Lys8 for heavy. Match between runs was selected and requantification enabled. FTMS scans had an initial search

tolerance of 20 ppm and 0.5 Da for ITMS MS/MS scans. Data analysis was conducted using the Perseus software package version 1.5.8.5, Microsoft Excel, and Graphpad Prism. Significant hits were determined by the Significance B calculation in Perseus or using a log2 cutoff. Interaction networks were generated using data curated from STRING database and visualized using Ctyoscape (version 3.3.1).

V. References

1. Kim, E. K., and Choi, E. J. (2010) Pathological roles of MAPK signaling pathways in human diseases. *Biochim Biophys Acta*. *1802*, 396-405.

2. Cazier, J. B., and Tomlinson, I. (2010) General lessons from large-scale studies to identify human cancer predisposition genes. *J. Pathol*. *220,* 255-262.

3. Good, M. C., Zalatan, J. G., and Lim, W. A. (2011) Scaffold proteins: hubs for controlling the flow of cellular information. *Science*. *332,* 680-686.

4. Scott, J. D., and Pawson, T. (2009) Cell signaling in space and time: where proteins come together and when they're apart. *Science*. *326,* 1220-1224.

5. Gingras, A. C., Gstaiger, M., Raught, B., and Aebersold, R. (2007) Analysis of protein complexes using mass spectrometry. *Nat. Rev. Mol. Cell Biol. 8,* 645-654.

6. Bensimon, A., Heck, A. J. R., and Aebersold, R. (2012) Mass spectrometry-based proteomics and network biology. *Annu. Rev. Biochem*. *81,* 379-405.

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

8. Mayer B. J., and Baltimore, D. (1993) Signaling through SH2 and SH3 domains. *Trends Cell Biol*. *3,* 8-13.

9. Dar, A. C., and Shokat, K. M. (2011) The evolution of protein kinase inhibitors from antagonists to agonists of cellular signaling. *Annu. Rev. Biochem*. *80,* 769–795.

10. Lambda, V., and Ghosh, I. (2012) New directions in targeting protein kinases: focusing upon the true allosteric and bivalent inhibitors. *Curr. Pharm. Des*. *18,* 2936-2945.

11. Jencks, W. P. (1981) On the attribution and additivity of binding energies. *Proc. Natl. Acad. Sci. U. S. A. 78,* 4046-4050.

12. Hill, Z. B., Perera. G. K., and Maly, D. J. (2009) A chemical genetic method for generating bivalent inhibitors of protein kinases. *J. Am. Chem. Soc*. *131,* 6686-6688.

13. Hill, Z. B., Perera. G. K., Andrews, S. S., and Maly, D. J. (2012) Targeting diverse signaling interaction sites allows the rapid generation of bivalent kinase inhibitors. *ACS Chem. Biol*. *7,* 487-495.

14. Hill, Z. B., Perera, B. G., and Maly, D. J. (2011) Bivalent inhibitors of the tyrosine kinases ABL and SRC: determinants of potency and selectivity. *Mol. BioSyst. 7,* 447-456.

15. Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H., and Johnsson, K. (2003) A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol*. *21,* 86-89.

16. Juillerat, A., Gronemeyer, T., Keppler, A., Gendreizig, S., Pick, H., Vogel, H., and Johnsson, K. (2003) Directed evolution of O6-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules in vivo. *Chem. Biol*. *10,* 313-317.

17. Keppler, A., Pick, H., Arrivoli, C., Vogel, H., and Johnsson, K. (2004) Labeling of fusion proteins with synthetic fluorophores in live cells. *Proc. Natl. Acad. Sci. U. S. A. 101,* 9955-9959.

18. Binz, H. K., Amstutz, P., and Plückthun, A. (2005) Engineering novel binding proteins from nonimmunoglobulin domains. *Nat. Biotechnol*. *23,* 1257-1268.

19. Kummer, L., Parizek, P., Rube, P., Millgramm, B., Prinz. A., Mittl, P. R., Kaufholz, M., Zimmermann, B., Herberg, F. W., and Plückthun, A. (2012) Structural and functional analysis of phosphorylation-specific binders of the kinase ERK from designed ankyrin repeat protein libraries. *Proc. Natl. Acad. Sci. U. S. A. 109,* E2248-E2257.

20. Dhillon, A. S., Hagan, S., Rath, O., and Kolch, W. (2007) MAP kinasae signaling pathways in cancer. *Oncogene 26,* 3279-3290.

21. Canagarajah, B. J., Khokhlatchev, A., Cobb, M. H., and Goldsmith, E. J. (1997) Activation mechanism of the MAP kinae ERK2 by dual phosphorylation. *Cell. 90,* 859–869.

22. Gower, C. M., Thomas, J. R., Harrington, E., Murphy, J., Chang, M. E. K., Cornella-Taracido, I., Jain, R. K., Schirle, M., and Maly, D. J. (2016) Conversion of a single polypharmacological agent into selective bivalent inhibitors of intracellular kinase activity. *ACS Chem. Biol. 11*, 121-131.

23. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics*. *1,* 376-386.

24. Blagoev, B., Kratchmarova, I., Ong, S. E., Nielsen, M., Foster, L. J., and Mann, M. (2003) A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nature Biotech*. *21,* 315-318.

25. Perez, F., Diamantopoulos, G. S., Stalder, R., and Kreis, T. E. (1999) CLIP-170 highlights growing microtubule ends in vivo. *Cell. 96,* 517-527.

26. Brigham, J. L., Perera, G. K., and Maly, D. J. (2013) A hexylchloride-based catch-and-release system for chemical proteomic applications. *ACS Chem. Biol*. *8,* 691-699.

27. Cull, M. G., and Schatz, P. J. (2000) Biotinylation of proteins in vivo and in vitro using small peptide tags. *Methods Enzymol*. *326,* 430-440.

28. Wu, S. C., and Wong, S. L. (2004) Development of an enzymatic method for site-specific incorporation of desthiobiotin to recombinant proteins in vitro. *Anal. Biochem*. *331,* 430-438.

29. Zhang, L., Holmes, I. P., Hochgräfe, F., Walker, S.R., Ali, N. A.; Humphrey, E. S., Wu, J., de Silva, M., Kersten, W. J., Connor, T., Falk, H., Allan, L., Street, I. P., Bentley, J. D., Pilling, P. A., Monahan, B. J., Peat, T. S., and Daly, R. J. (2013) Characterization of the novel broad-spectrum kinase inhibitor CTx-0294885 as an affinity reagent for mass spectrometry-based kinome profiling. *J. Proteome Res*. *12,* 3104,-3116.

30. Baiady, N.. Padala, P., Mashahreh, B., Cohen-Kfir, E., Todd, E. A., Du Pont, K. E., Berndsen, C. E., and Wiener, R. (2016) The Vps27/Hrs/STAM (VHS) domain of the signal-transducing adaptor molecule (STAM) directs associated molecule with SH3 domain of STAM (AMSH) specificity to longer ubiquitin chains and dictates the position of cleavage. *J. Biol. Chem. 291*, 2033-2042.

31. Hjerpe, R., Aillet, F., Lopitz-Otsoa, F., Lang, V., England, P., and Rodriguez, M. S. (2009) Efficient protection and isolation of ubiquitylated proteins using tandem ubiquitin-binding entities. *EMBO Rep. 10*, 1250-1258.

32. Davies, G. C., Ettenberg, S. A., Coats, A. O., Mussante, M., Ravichandran, S., Collins, J., Nau, M. M., and Lipkowitz, S. (2004) Cbl-b interacts with ubiquitinated proteins; differential functions of the UBA domains of c-Cbl and Cbl-b. *Oncogene. 23*, 7104-7115.

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

34. Pawson, T., and Nash, P. (2003) Assembly of cell regulatory systems through protein interaction domains. *Science. 300*, 445-452.

35. Golkowski, M., Vidadala, R. S. R., Lombard, C. K., Suh, H. W., Maly, D. J., and Ong, S. E. (2017) Kinobead and single-shot LC-MS profiling identifies selective PKD inhibitors. *J. Proteome Res.* In press.

36. Rappsilber, J., Ishihama, Y., and Mann, M. (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem. 75*, 663-670.

Chapter 2: Tyrosine Kinase SH2 Domain Profiling

I. Introduction

The core ability enabling the vast array of interesting and critical biology within a cell lies in their ability to propagate and appropriately respond to cellular stimuli. The simplest concept of signaling is that one molecule acts on another, which subsequently acts on another down the line, and so forth as the signal travels through space and time. Signal transduction has evolved in a such a way to allow for modulation, regulation, and tuning at each step of the pathway. Tyrosine phosphorylation, which has evolved to be a critical component of signal transduction in all eukaryotic cells, is an important regulator of growth, cell cycle progression, differentiation, metabolic homeostasis, neural transmission, and aging (1).

Tony Pawson and Wendell Lim eloquently noted that signaling is a system of writers, erasers, and readers (2). In the case of phosphotyrosine, tyrosine kinases are the writers that phosphorylate specific tyrosine residues. Tyrosine phosphatases act as erasers by selectively removing the phosphorylation form of tyrosine. Finally, domains that selectively recognize phosphotyrosine over the unphosphorylated form–like Src homology 2 (SH2) domains–are the readers and that this transient tyrosine modification. This system of writers, erasers, and readers work in tandem to create a complex and diverse set of signaling outcomes that is information rich and robust.

Under basal conditions, tyrosine kinases are tightly autoinhibited and have low activity, but in response to certain stimuli, they can be rapidly activated. Dysregulation of tyrosine kinase autoinhibition can lead to disease states, most notably in many forms of cancer (3). Because tyrosine kinase activity facilitates the formation of multi-member signaling complexes, it is crucial that they are not only highly regulated in activity but that they also only interact with specific protein binding partners and substrates. The greater than ninety tyrosine kinases in the human kinome can be broken down into two different types: receptor tyrosine kinases and non-receptor tyrosine kinases whose 32 members can be placed into 10 subfamilies (4). Most non-receptor cytoplasmic tyrosine kinase family members contain an SH2 domain that can regulate catalytic activity and mediate cellular localization and substrate targeting. For example, the SH2 domain has a multi-functional role in the Src-family kinases; stabilizing an autoinhibited state when a C-terminal regulatory site is phosphorylated and directing substrate and scaffolding interactions of the activated enzyme (5).

Besides providing a regulatory role in tyrosine kinases, SH2 domains are important in selective phosphotyrosine recognition to specifically target kinases towards substrates and to localize in active signaling complexes. Various studies have revealed that phosphotyrosine binding by the SH2 domain is a two-pronged approach, with the SH2 binding pocket providing affinity towards the phosphotyrosine moiety and a second pocket or subsite that confers specificity by selectively recognizing residues–especially at positions -2 to +4 relative to phosphotyrosine–next to the site of phosphorylation (6-8). Given their importance in dictating substrate targeting for kinases, several groups have profiled the phosphotyrosine recognition motifs of SH2 domains with a variety of biochemical methods.

In 1993, Songyang and coworkers probed the SH2 domain interactions of 11 proteins, including the tyrosine kinases Fgr, Lck, Src, Fyn, and Abl against a synthetic phosphopeptide library (9). A library of phosphotyrosine-containing peptides was incubated with each immobilized, GST-tagged SH2 domain and bound peptides eluted and sequenced to provide the recognition motif around the phosphotyrosine. These studies revealed two distinct phosphotyrosine interaction motifs, with the five tyrosine kinases preferring a general motif of pY-hydrophilic (+1)-hydrophilic (+2)-Ile/Pro (+3). Notably, the SH2 domains of Src family kinases preferred the sequence motif pY-Glu (+1)-Glu (+2)-Ile (+3).

Similar approaches utilizing peptide arrays were done by Liu et al. in 2010 and Tinti et al. in 2013. Both groups employed the SPOT method to profile a large number of different phopshopeptides against an SH2 domain panel. Liu and coworkers measured the interactions of 50 SH2 domains with solid-phase peptide arrays using fluorescent polarization (10). The Liu et al. peptide library consisted of 192 peptides of 11 amino acids in length that corresponded to cytoplasmic tyrosine-containing regions of InsR, IRS-1/2, FGFR1/2/3, FRS-2, PLC-γ1, Crk, p130Cas, and p62Dok1. Members of the SH2 panel showed a high degree of selectivity with interactions ranging from a dissociation constant (K_d) of 5.8 μM to 0.18 nM. Clustering analysis highlighted several families based on peptide binding. Interestingly, many of the tyrosine kinases–including Src, Brk, Hck, Yes, Fgr, Fyn, Lyn, Lck, Blk, and Csk–clustered together.

The study by Tinti and coworkers expanded the peptide library to include 6,202 phosphopeptides on a chip that were profiled against 70 GST-tagged SH2 domains (11, 12). For each SH2 domain, phosphopeptides that provided a signal higher than the average signal by more than two standard deviations were aligned and used for sequence logo analysis. Based upon this analysis, recognition motifs were generated for each of the SH2 domains profiled as well as specificity relationships between each. This analysis revealed a different recognition motif

relationship in that Src family kinases Fgr, Yes, Src, and Lyn group in one specificity class while the tyrosine kinases Fes, Fyn, Blk, Hck Brk, Btk, Abl1, and Abl2 are in another.

Because ongoing proteomic efforts continually increase the number of discovered phosphorylated peptides, the authors attempted to use the results from their chip analysis to predict SH2 domain interactions with phosphotyrosine-containing peptides. To validate these predictions, 57 synthetic phosphopeptides were immobilized and used as affinity reagents in EGF-stimulated HeLa cells. Interactors were then analyzed by LC-MS. While a majority of the identified proteins were SH2 domains, only 33% of the SH2 domain-phosphopeptide interactions determined experimentally were highly ranked based on predictions from chip array analysis.

In 2007, Machida et. al. used a different approach–a binding assay based on a reverse-phase protein array–to profile SH2 domain-binding specificity (13). The authors performed comprehensive profiling of GST fusions of all known human SH2 domains. Of these, 74 of the GST-SH2 fusions had sufficient solubility for use. Their profiling procedure was two-fold. First, Far-Western blotting of each SH2 domain against a mixture of pervanadate-treated lysate. Second, a high-throughput reverse-phase binding assay where phosphotyrosine-containing lysate is spotted onto nitrocellulose and then probed with the SH2 domain of interest. These assays revealed distinct binding patterns for most SH2 domains to the mixed pool of phosphotyrosine. Notably, the Src family kinases could be subclassified into two distinct families as a consequence of their Far-Western binding. Blk, Lyn, Hck, and Lck clustered in one group, while Fyn, Yes, Fgr, and Src grouped into another.

The work described above has provided a great deal of insight into the diversity and specificity of SH2 domain interactions with phosphotyrosine-containing peptides. However, a comparison of these results also yields conflicting data, especially regarding the relationships in recognition specificity between different SH2 domains. The common theme in all previous studies is that SH2 domains are profiled against unfolded interactors, isolated from their native biological context. Phosphotyrosine-containing peptides are only short stretches of amino acids that are excised from larger folded domains and far-western analysis is performed on denatured proteins separated by gel electrophoresis. Furthermore, in both cases, only primary interactions are probed, with higher order structure and any secondary interactions lost. To attempt to address these limitations, we used tyrosine kinase SH2 domains as affinity purification reagents to pull down interactors at the protein level from a pool of lysate. By investigating the interactions of these domains against intact proteins in the context of a signaling pathway, we hoped to elucidate

additional insights into SH2 domain specificity relationships and SH2 domain-containing signaling complex identity.

II. Results and Discussion

Initial Characterization of SH2 Domain Panel

A panel of seven protein tyrosine kinase SH2 domains (Abl, Brk, Btk, Csk, Fes, Lyn, and Src) were chosen based on diverse evolutionary relationship and good solubility as determined by available crystal structures. An 'phosphotyrosine dead' mutant was generated from the SH2 domain of Grb2 (R86K) to act as a control protein bait that would not enrich any phosphotyrosine-containing proteins. Fusion proteins of SH2 domains and the labeling protein SNAP-tag were generated, expressed in *E. coli*, and purified. The purified SH2-SNAP-tag fusions were then covalently conjugated via the SNAP-tag labeling reaction to functionalized resin displaying O^4-benzyl-2-chloro-6-aminophosphotyrosinerimidine (CLP) as illustrated in Figure 2-1.

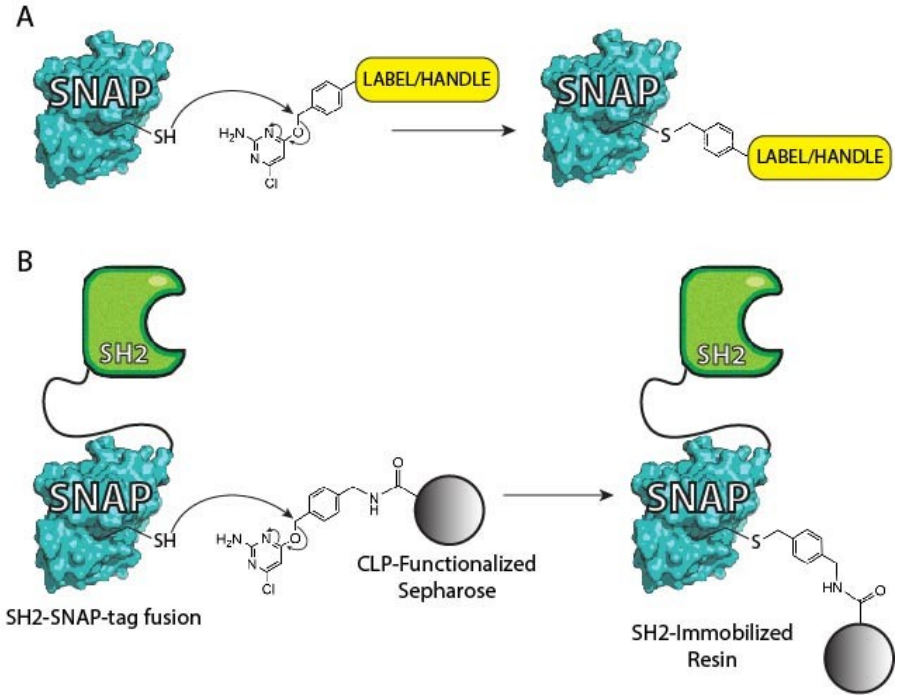


Figure 2-1. SNAP-tag labeling reaction for use in covalent conjugation of molecules to protein. (A) SNAP-tag can label itself with molecules bearing a CLP handle in an SN2 reaction via its active site cysteine. (B) SH2-domain fusion proteins are conjugated to functionalized resin through SNAP-tag mediated conjugation.

In order to determine the optimal window for probing phosphotyrosine-based interactions for a given signaling timeframe, two model signaling pathways were analyzed. First, timecourses for EGF stimulation of EGFR in HeLa cells and CD3/CD28 costimulation of the T-cell receptor pathway in Jurkats were performed to determine which timepoints provided optimal signaling output (Figure 2-2). HeLa cells were stimulated with EGF for 0, 1, 5, 15, 30, and 60 min and then subsequently analyzed for overall phosphotyrosine amount by Western blotting with the anti-phosphotyrosine antibody 4G10. A similar timecourse and analysis was performed for CD3CD28 stimulation of Jurkat cells. 5 min for EGF stimulation of HeLa cells and 15 min for CD3/CD28 stimulation in Jurkat cells showed maximum phosphotyrosine levels and lysates from HeLa and Jurkat cells stimulated for these durations were generated for proteomic profiling of our SH2 domain panel.

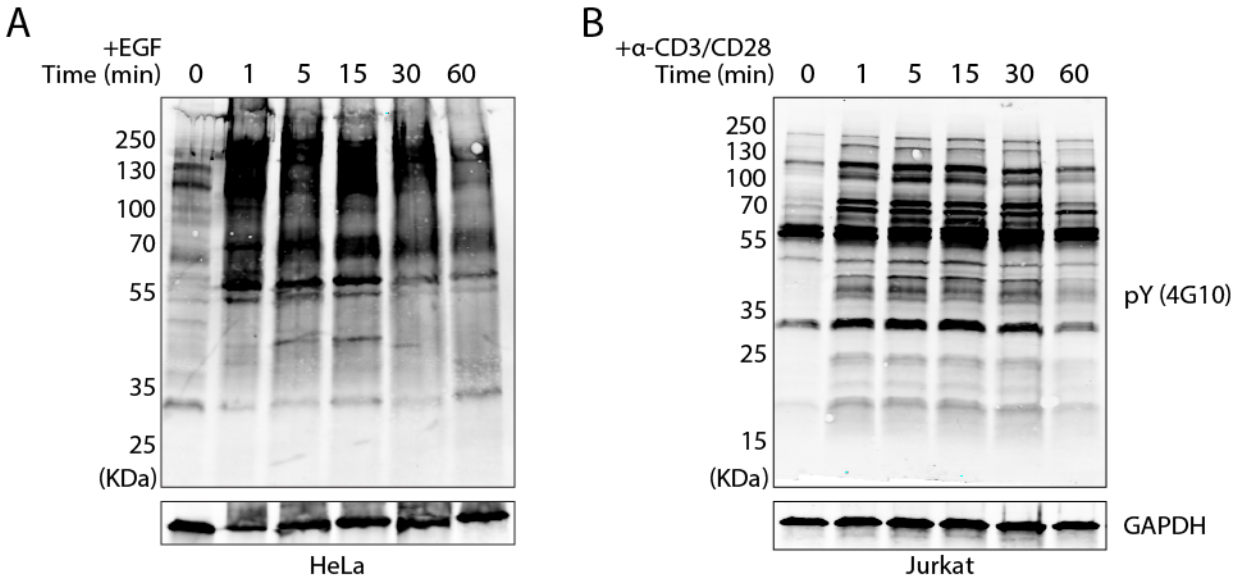


Figure 2-2. Time course probing for phosphotyrosine in EGF-stimulated HeLa cell lysate (A) and T cell activated Jurkat cell lysate(B).

The panel of SH2 domains were first analyzed by their overall ability to enrich phosphotyrosine from stimulated lysate. The lysates from EGF-stimulated HeLa and CD3/CD28-activated Jurkat cells were exposed to 4 nmol of each SH2 domain conjugated to resin. The resulting elutions from each SH2 were subsequently probed with anti-phosphotyrosine antibodies. Peptide microarray experiments described by Tinti et. al. in 2012 suggested that phosphotyrosine antibodies 4G10, pY20, and PY100 have different specificities (14). As such, we used several different phosphotyrosine antibodies–4G10, APY03, and PY100–to highlight the full range of enrichment differences between the SH2 domains.

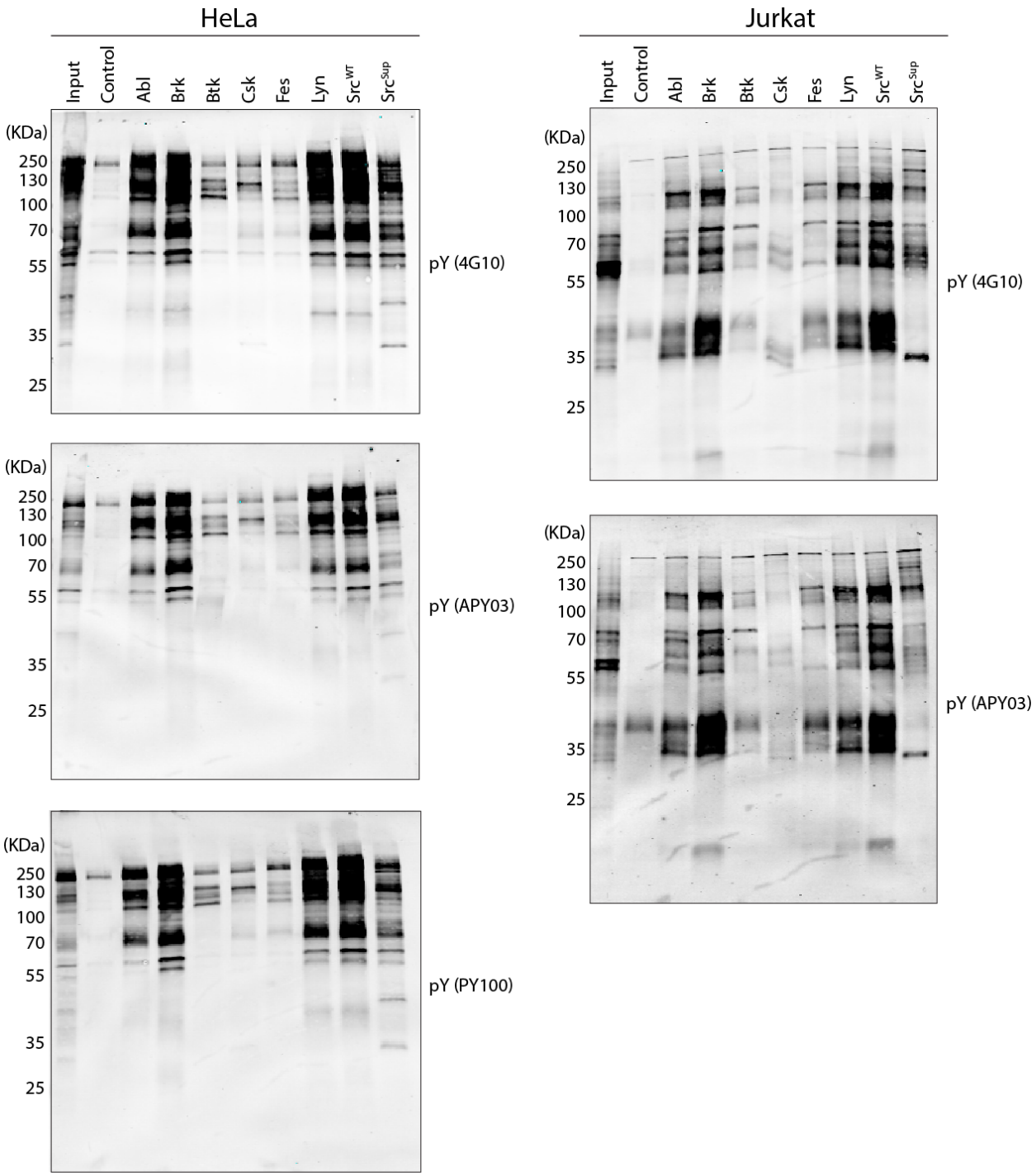


Figure 2-3. Initial profiling of SH2 domain panel in EGF stimulated HeLa and CD3/CD28-costimulated Jurkat lysates. Enriched proteins were immunoblotted with the anti-phosphotyrosine antibodies 4G10, APY03, and PY100.

As shown in Figure 2-3, the SH2 domains of Brk, Src, and Lyn pull down the largest number of phosphotyrosine-modified proteins, whereas Btk, Csk, and Fes show much more

selective and limited enrichment of phosphotyrosine-containing proteins in EGF-stimulated HeLa cells. The trend in Jurkat stimulated lysate was similar, with Src and Brk displaying the highest level of overall phosphotyrosine enrichment. Btk, Csk, and Fes still appear to be quite specific and show the lowest enrichment out of the panel. The results for Csk in Jurkat cells was surprising because Csk kinase is known to play a role in suppressing T cell receptor activation through phosphorylating several positive effectors (15, 16). In Jurkat lysate, Abl and Lyn have lower levels of enrichment relative to Src and Brk when compared to affinity purification from HeLa lysate.

The three different phosphotyrosine antibodies used in this study did not show appreciable differences in phosphotyrosine species detected. Overall, the 4G10 antibody gave the best results in terms of coverage, but all of the antibodies tested yielded the same trends when comparing the enrichment levels of the SH2 domain panel.

MS Characterization of SH2 Domain Interactions in EGF-Stimulated HeLa

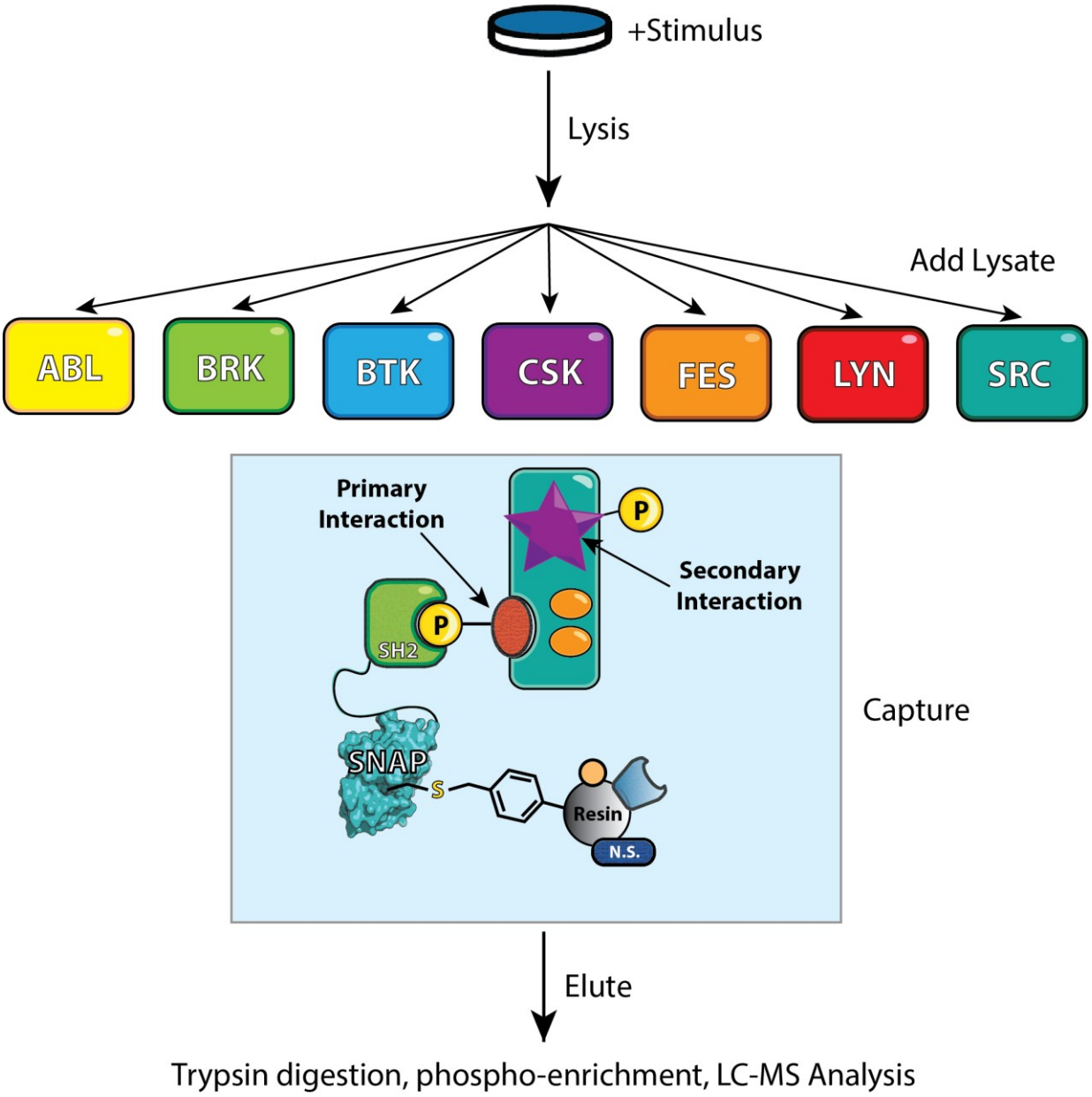
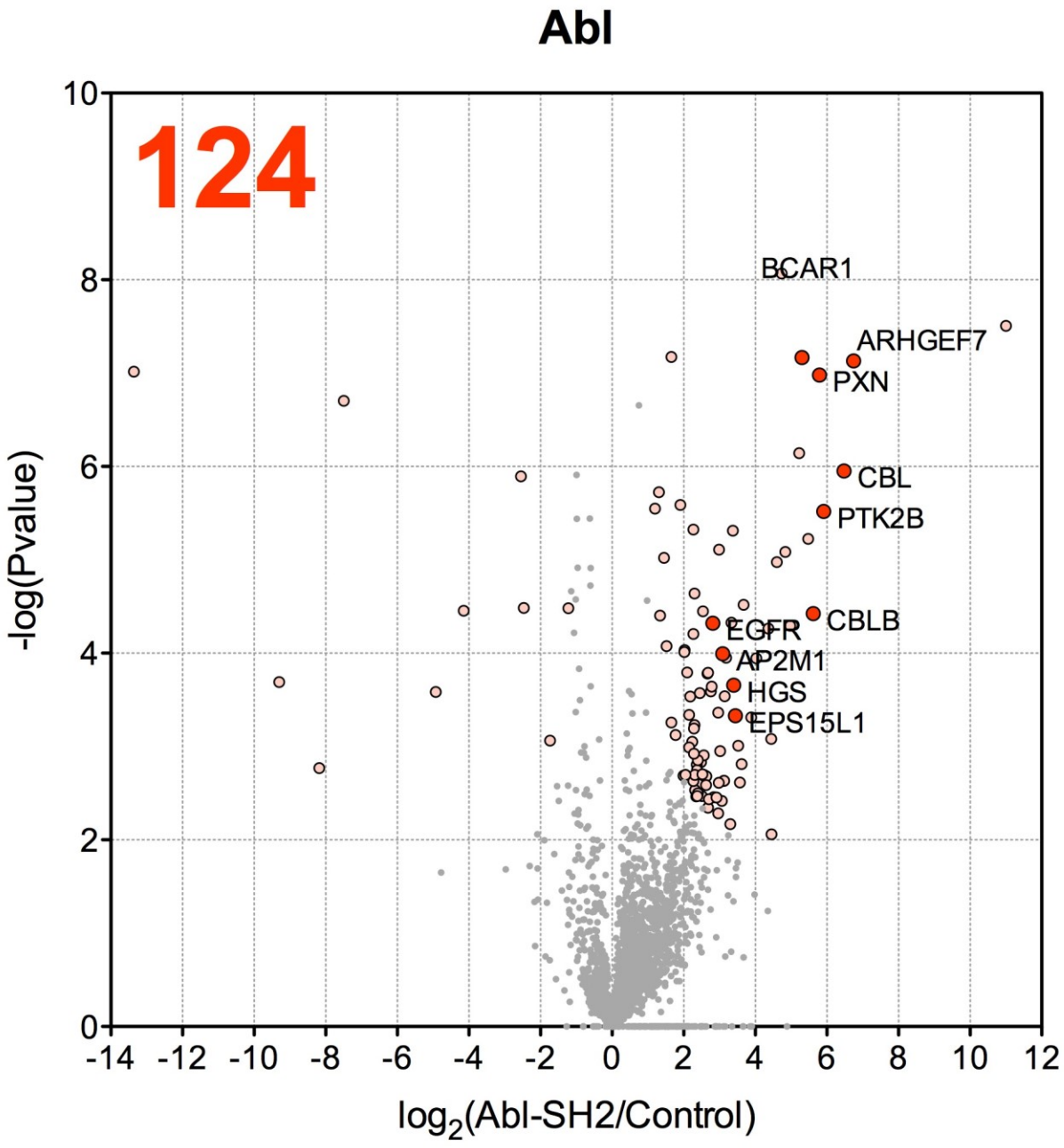
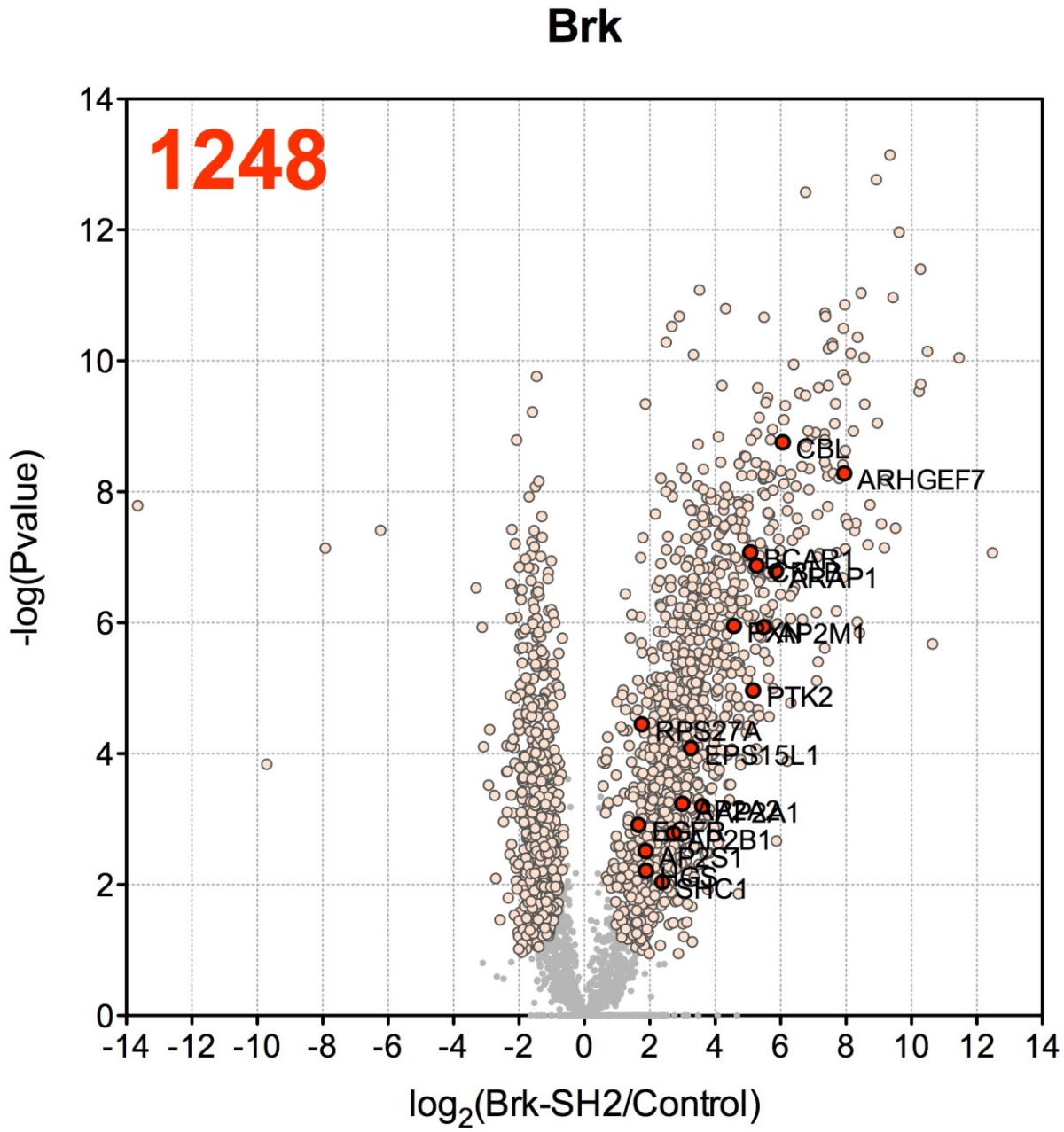
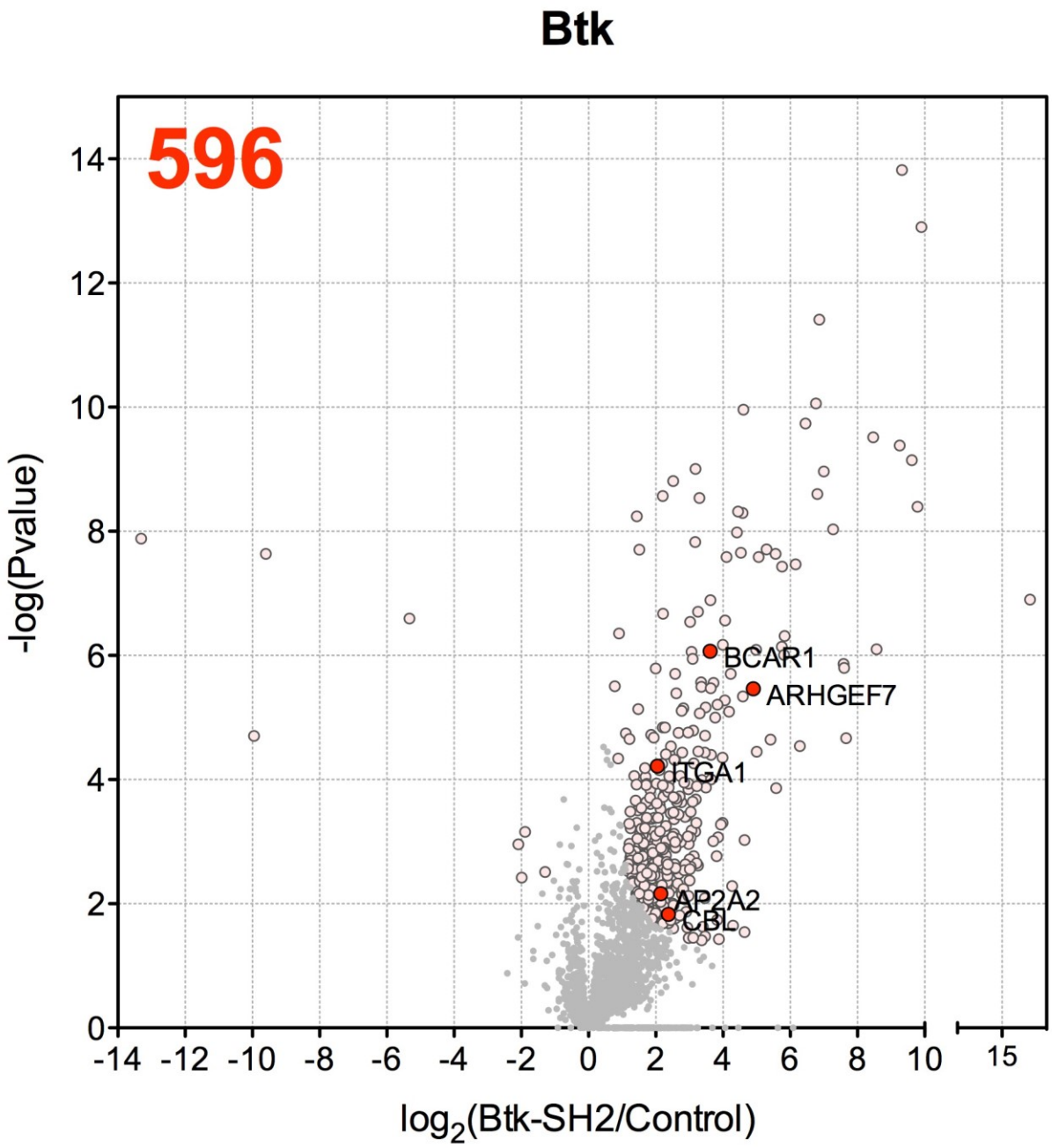
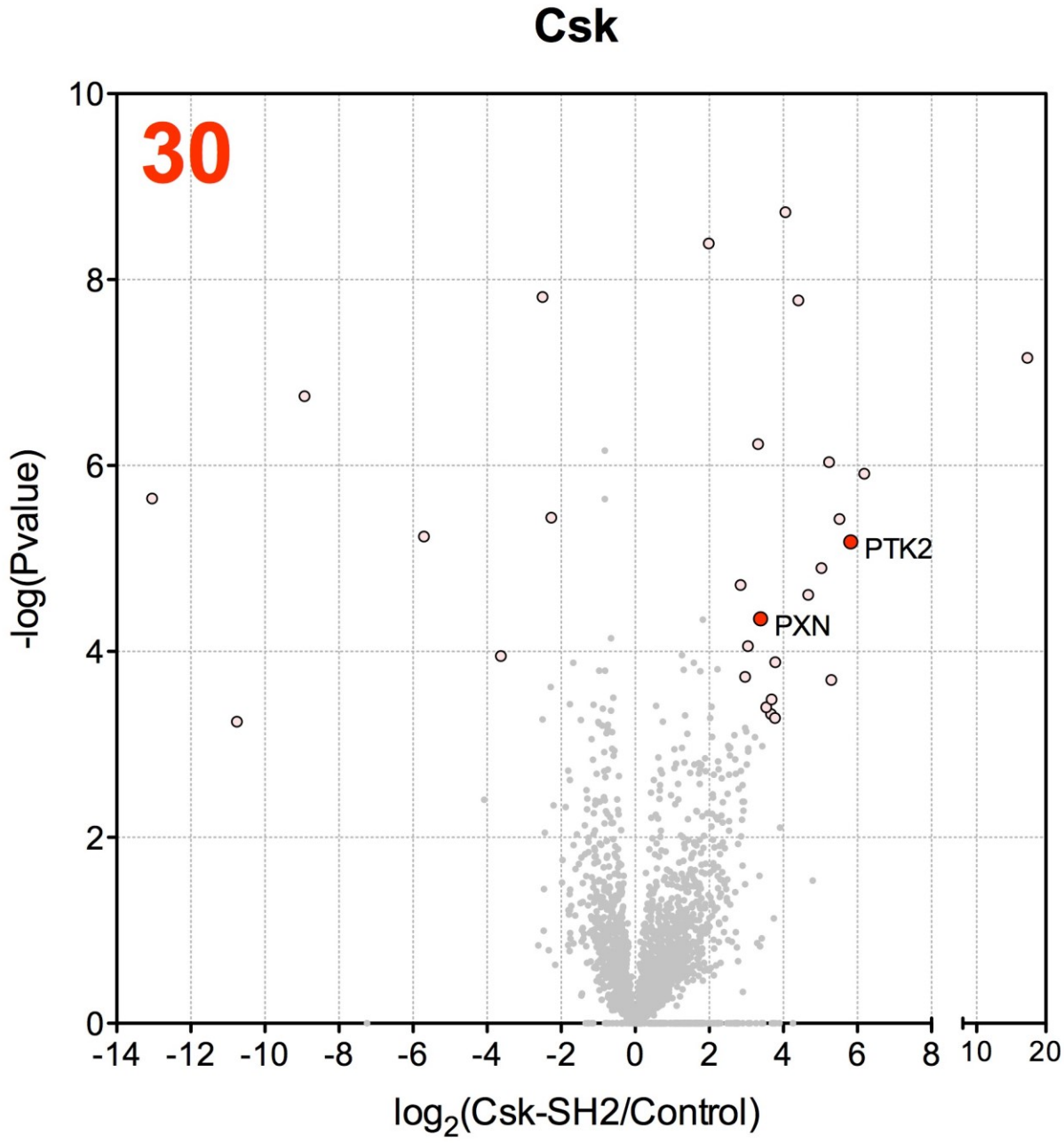


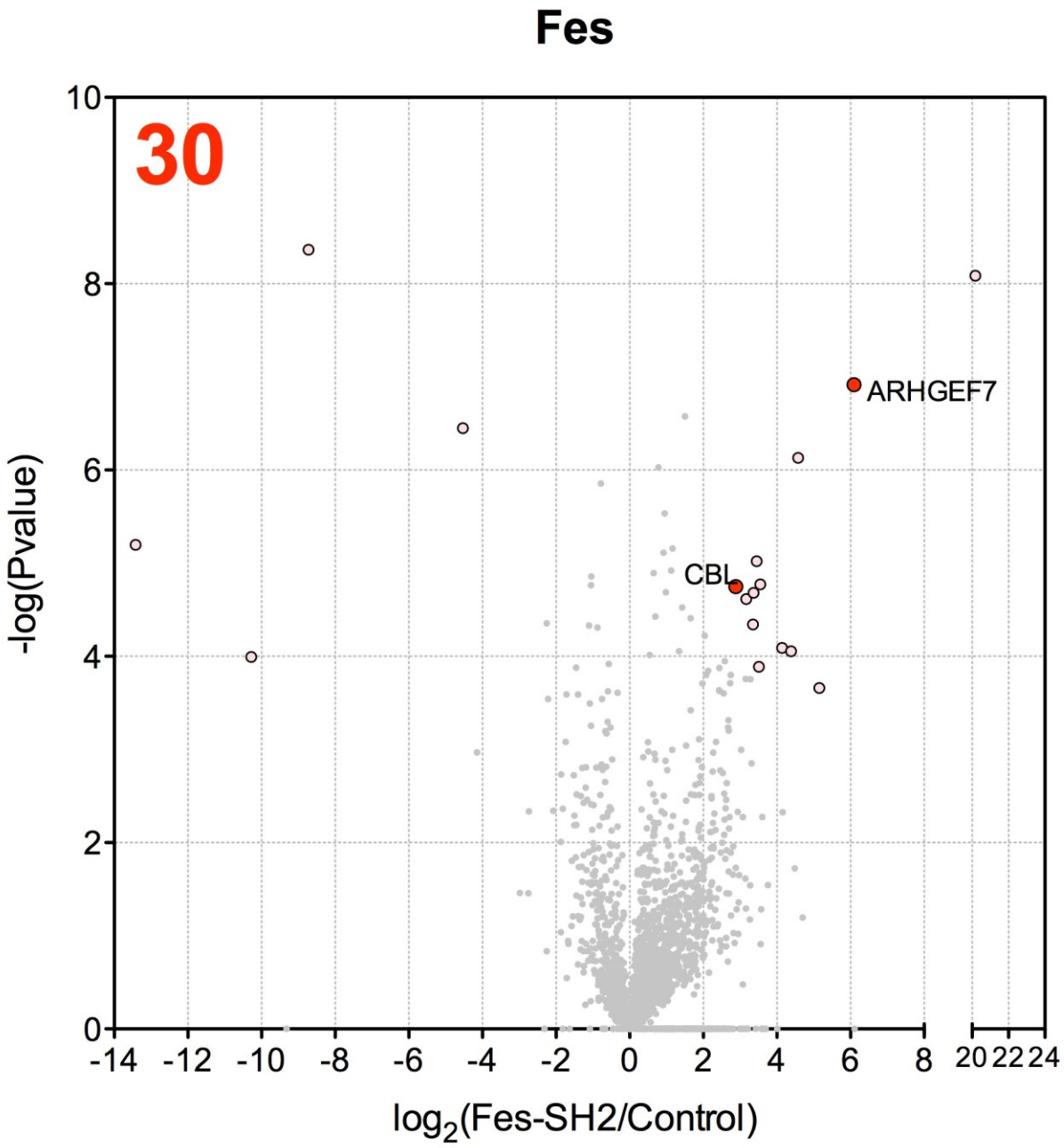
Figure 2-4. Schematic of proteomic profiling of the SH2 domain panel. The signaling pathway is activated in the cell (+Stimulus) and then cells are lysed. Lysate is then equally distributed in a pulldown with each of the seven SH2 domains immobilized via SNAP-tag to resin during the 'Capture' step. The SH2 domains should bind to interacting proteins with phosphotyrosine (P), but also potentially capture secondary interactors that are complexed with the primary interaction. This mix of captured proteins are eluted from the resin and then analyzed by LC-MS.

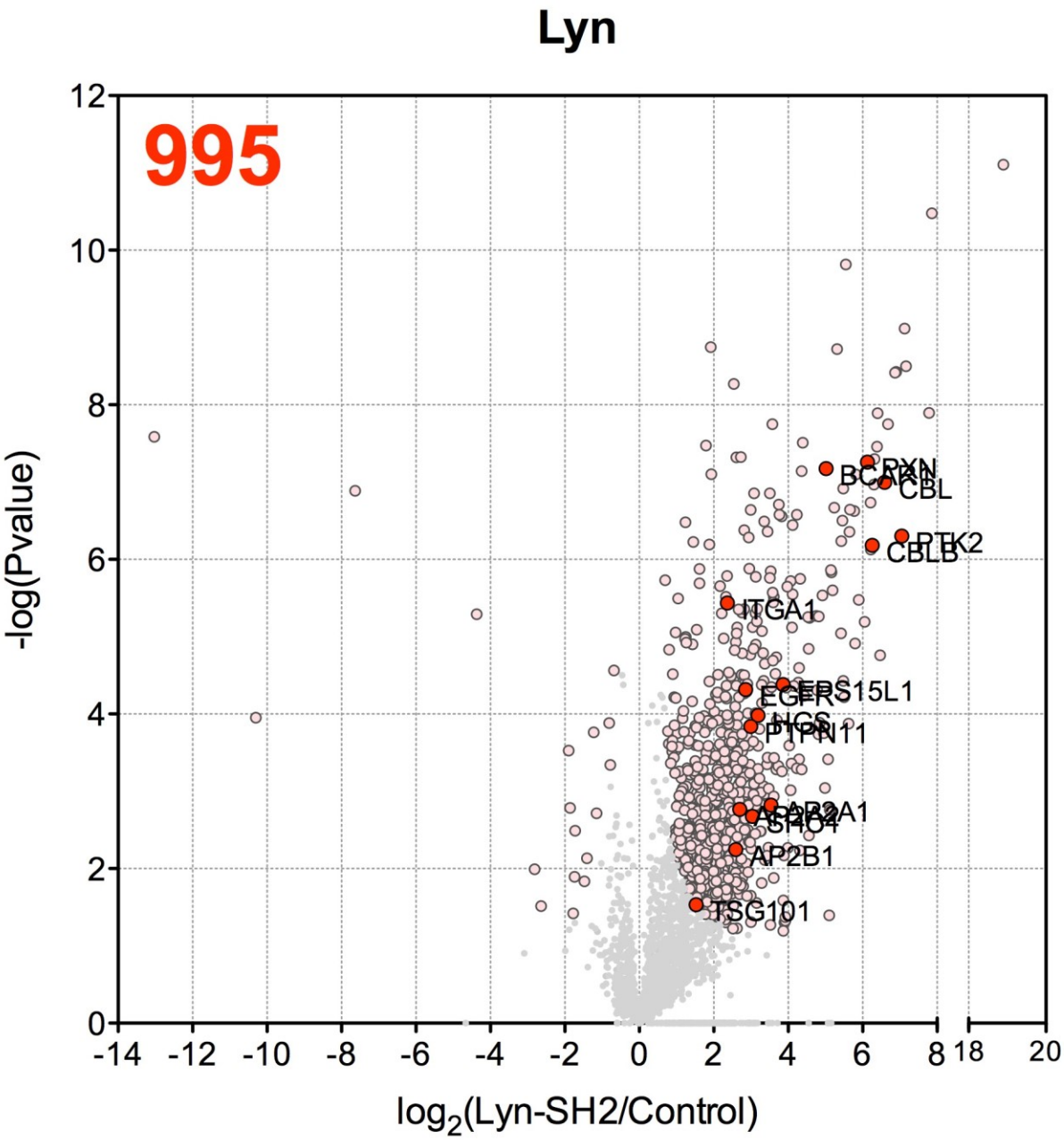

Abl
124
BCAR1
ARHGEF7
PXN
CBL
PTK2B
CBLB
EGFR
AP2M1
HGS
EPS15L1
-log(Pvalue)
log2(Abl-SH2/Control)
10
8
6
4
2
0
-14
-12
-10
-8
-6
-4
-2
0
2
4
6
8
10
12


Brk
1248
-log(Pvalue)
log2(Brk-SH2/Control)
CBL
ARHGEF7
PTK2
RPS27A
EPS15L1
SHC1
14
12
10
8
6
4
2
0
-14
-12
-10
-8
-6
-4
-2
0
2
4
6
8
10
12
14


Btk
596
BCAR1
ARHGEF7
ITGA1
AP2A2
CBL
-log(Pvalue)
log2(Btk-SH2/Control)
14
12
10
8
6
4
2
0
-14
-12
-10
-8
-6
-4
-2
0
2
4
6
8
10
15


Csk
30
PTK2
PXN
-log(Pvalue)
log2(Csk-SH2/Control)
10
8
6
4
2
0
-14
-12
-10
-8
-6
-4
-2
0
2
4
6
8
10
20


Fes
30
ARHGEF7
CBL
-log(Pvalue)
log2(Fes-SH2/Control)
10
8
6
4
2
0
-14
-12
-10
-8
-6
-4
-2
0
2
4
6
8
20
22
24


Lyn
995
-log(Pvalue)
log2(Lyn-SH2/Control)
12
10
8
6
4
2
0
-14
-12
-10
-8
-6
-4
-2
0
2
4
6
8
18
20
PXN
CBL
PTK2
CBLB
ITGA1
EPS15L1
EGFR
HGS
PTPN11
AP2A1
SHC1
AP2B1
TSG101

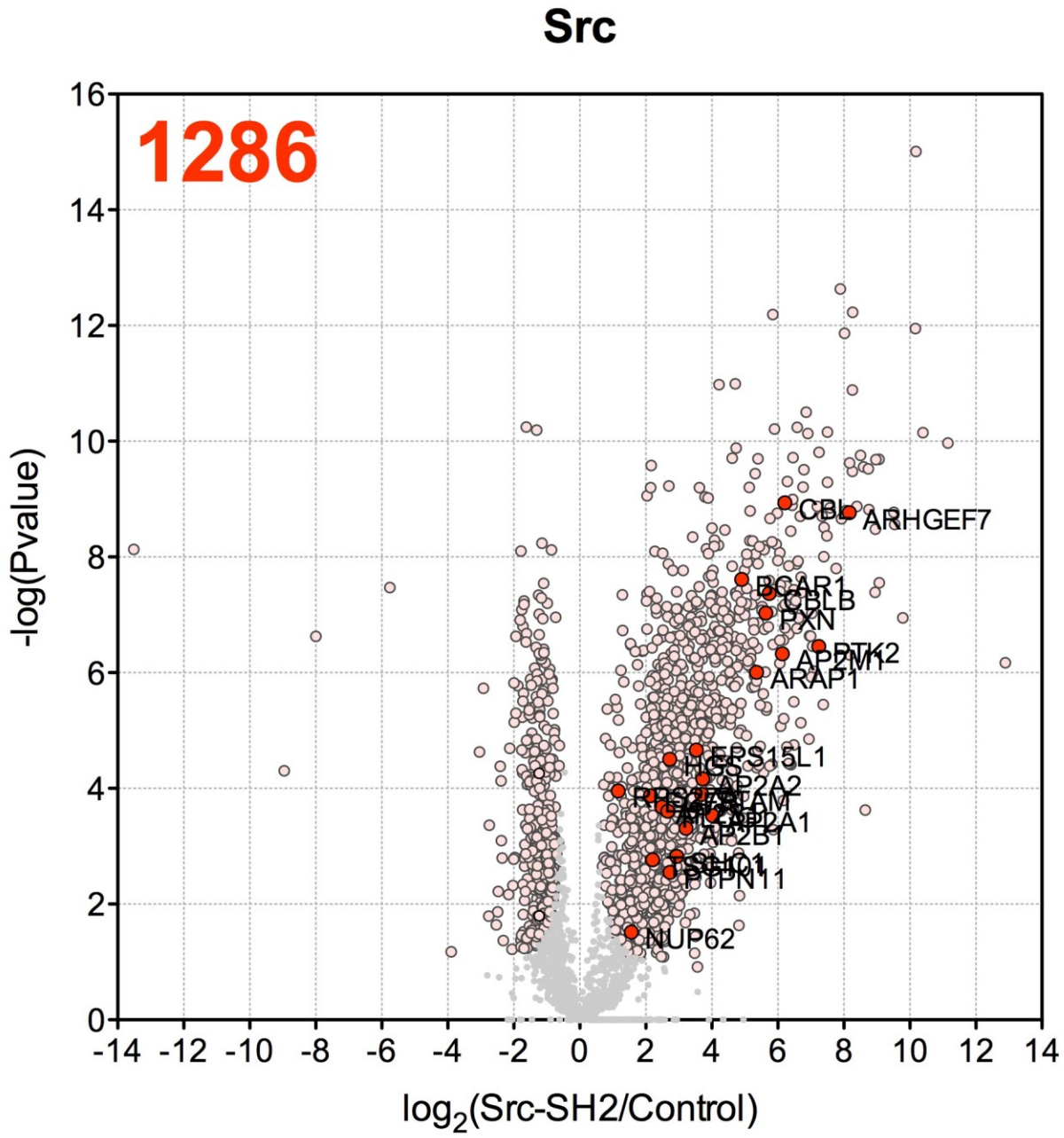


Figure 2-5. Volcano plots of the proteins enriched from EGF-stimulated HeLa cell lysate for each of the seven SH2 domains in the panel. Proteins that are significantly enriched by the listed SH2 domain compared to the Grb2 (R86K) control are listed in pink and the subset of those that are characterized with the EGFR signaling pathway by GO

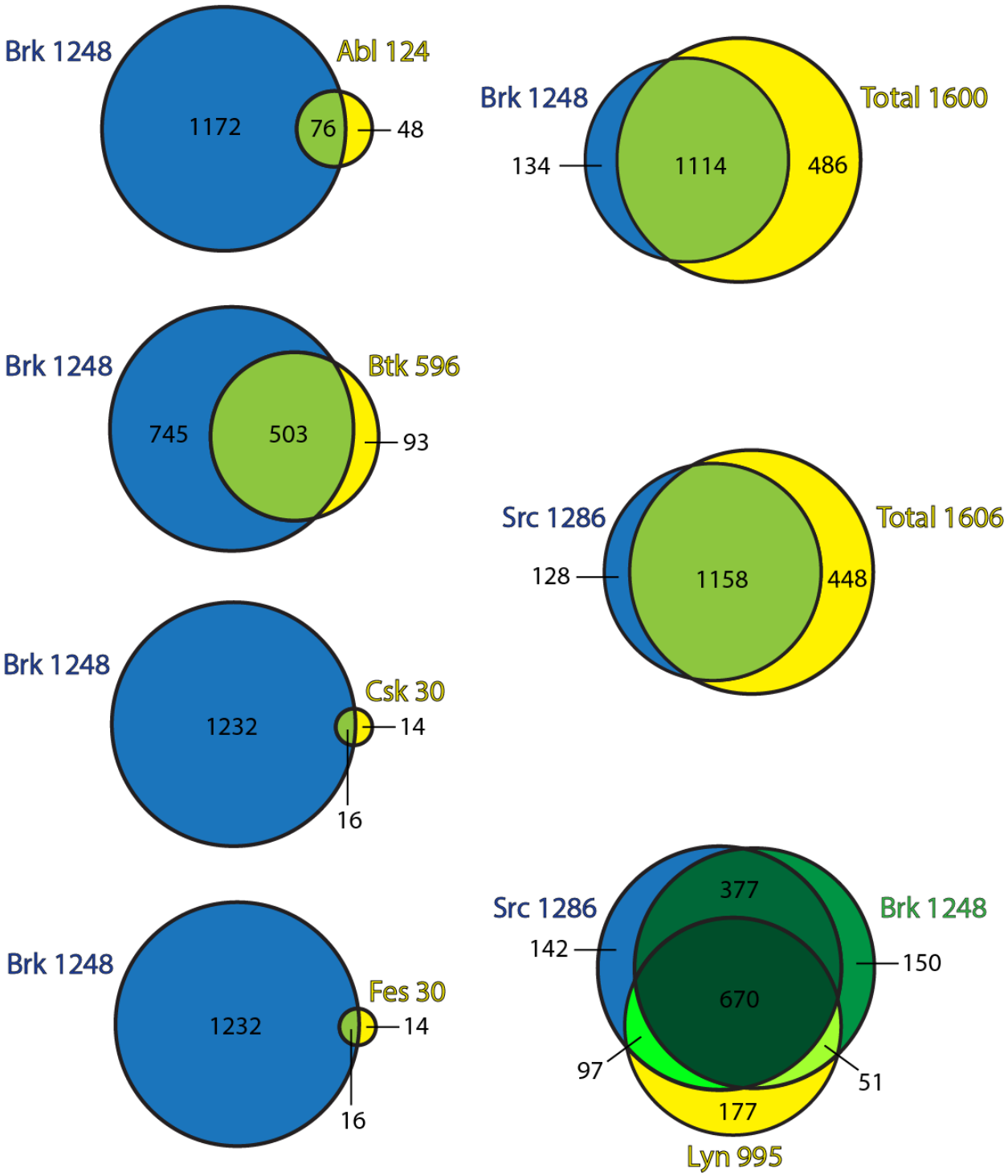


Figure 2-6. Diagrams comparing the enriched proteins from the panel of SH2 domains in EGF-stimulated HeLa cell lysate.

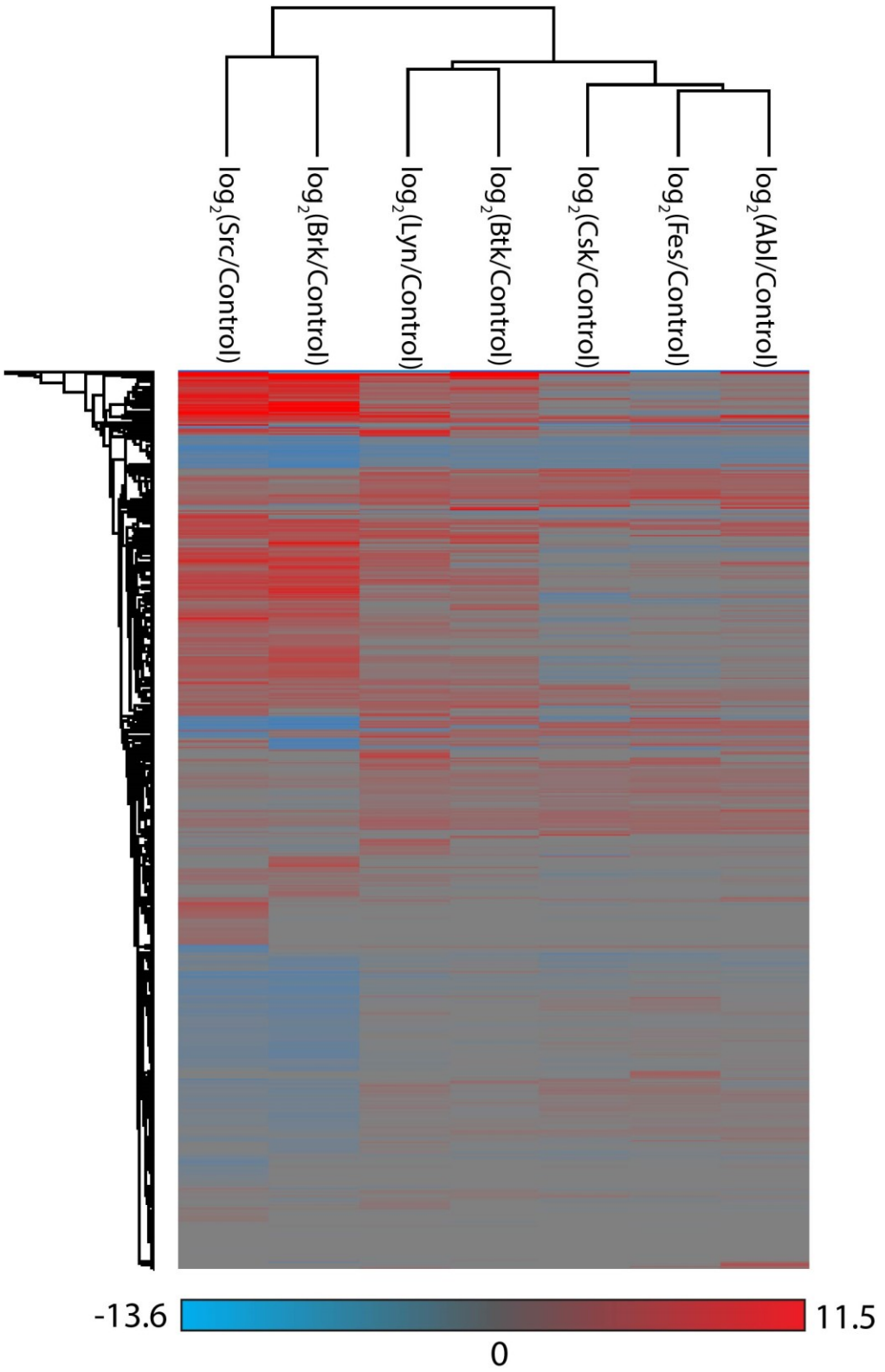


Figure 2-7. Hierarchical clustering analysis of the log2 ratio in enrichment for each SH2 domain and the control in EGF-stimulated HeLa cell lysate.

TERM	SRC	BRK	LYN	BTK	CSK	FES	ABL
GO:0007173~epidermal growth factor receptor signaling pathway	11	9	9	-	3	3	8
GO:0042059~negative regulation of epidermal growth factor receptor signaling pathway	16	15	12	7	-	-	6
GO:0023051~regulation of signaling	3	-	3	3	-	-	3
GO:0048013~ephrin receptor signaling pathway	14	11	19	10	-	3	6
GO:0016477~cell migration	-	-	18	-	3	-	7
GO:0006364~rRNA processing	138	145	101	85	-	-	-
GO:0000398~mRNA splicing, via spliceosome	138	142	89	70	-	-	5
GO:0006413~translational initiation	104	110	89	75	-	-	5
GO:0000184~nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	87	91	74	61	-	-	-
GO:0006614~SRP-dependent cotranslational protein targeting to membrane	75	78	66	61	-	-	-
GO:0019083~viral transcription	83	78	68	61	-	-	-
GO:0000082~G1/S transition of mitotic cell cycle	16	6	15	-	4	-	-
GO:0018105~peptidyl-serine phosphorylation	-	-	-	-	4	-	-
GO:0046777~protein autophosphorylation	-	-	-	-	4	-	5
GO:0006977~DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest	-	-	-	-	3	-	-
GO:0043547~positive regulation of GTPase activity	-	-	-	-	-	6	10
GO:0000086~G2/M transition of mitotic cell cycle	5	-	14	9	3	3	-
GO:0008283~cell proliferation	43	42	30	-	-	4	-
GO:0006281~DNA repair	35	40	27	19	3	-	5
GO:0098609~cell-cell adhesion	62	64	55	32	-	-	6
GO:0007165~signal transduction	-	-	-	-	-	-	11
GO:0051439~regulation of ubiquitin-protein ligase activity involved in mitotic cell cycle	20	12	9	-	2	-	-
GO:0006457~protein folding	26	-	25	-	-	-	-
GO:0072583~clathrin-mediated endocytosis	5	5	5	4	-	-	-

Table 2-1. GO enrichment analysis of enriched proteins and the difference between the profiled SH2 domains. The number of proteins that were significantly enriched by an SH2 domain that were identified for a given GO-BP term are listed. Each SH2 domain entry is color coded to show the percentage those proteins represented from the total number of enriched hits for that SH2.

Following the initial characterization of our panel of seven SH2 domains, we next sought to profile them in more depth with quantitative mass spectrometric characterization at the protein and phosphosite level. EGF-stimulated HeLa cell lysate was prepared as described above, quantified, and 1.5 mg of extract was added to 4 nmol of each SH2 domain. For each SH2 domain, three technical replicate pulldowns were performed in two separate biological replicates. Phosphopeptide enrichment was performed on 90% of the trypsinized peptide from each elution, while the remaining ten percent was used for MS characterization of the protein targets. Significantly enriched hits for a given SH2 domain were determined by *t*-test in comparison to the phosphotyrosine dead control. The list of enriched protein targets for each domain in EGF-stimulated HeLa lysate was populated if a given protein was significantly pulled down over control in one or more biological replicates.

As shown in Figure 2-5, the number of enriched protein targets for the panel mirrors the overall density of phosphotyrosine-containing proteins observed in the initial profiling by Western blot. Interestingly, the SH2 domains of Src and Brk had large numbers of enriched proteins and

in fact cover nearly 80 percent of the 1600 proteins enriched by all of the SH2 domains in the remaining panel (Figure 2-6). Src and Lyn SH2 domains, which are both derived fomr Src family kinases, share 767 enriched protein targets, but Src SH2 has 40% unique enriched proteins and Lyn has 23%. All of the SH2 domains except for Csk and Fes showed significant enrichment of EGFR by *t*-test comparison to the control. In addition, most of the SH2 domains significantly enriched the same proteins that are involved in the EGFR signaling pathway including c-Cbl, Cbl-b, PTK2, Vav3, BCAR1, and paxillin (17-23).

Hierarchical clustering analysis of the label-free quantification (LFQ) intensity for each replicate (Supplementary Figure) and the *t*-test difference between each SH2 domain versus the control both (Figure 6) show a distinct grouping of Src and Brk, separated from the other five kinase SH2 domains profiled. This grouping of SH2 domain enrichment differs from the one established by Machida et al. with their Far Western-based approach on pervanadate-treated lysate, which showed a grouping of Lyn – Abl – Src – Fes – Brk – Csk – Btk. We also compared the phosphotyrosine-spot density from the Western blotted elutions (Figure 2-3) and MS results (Figure 2-5) to the Far-Western results of Machida et al. Strikingly, the range of Far-Western signal variability for each SH2 domain is much lower than the variability observed in our pulldowns. The SH2 domains from Btk, Csk, and Fes appear to be more selective in both analyses, but we observe a much bigger difference in relative enrichment levels in comparison to the SH2 domains of Brk and Src. In addition, the SH2 domain of Abl showed a more robust enrichment in the Far-Western screen. The observed variations are likely a result of the differences in the methodology between our proteomic method and the profiling of Machida. In the Far-Western profiling method, the lysate proteins are denatured and separated by SDS-PAGE, resulting in a screen that determines direct SH2-protein interactions that specifically reports on the primary sequence specificity of the SH2 domain. In our affinity purification method, we are pulling down at the proteins in a more physiologically relevant signaling environment to maintain native protein complexes. Our method probes direct SH2 domain interactions at the protein level and should enrich secondary interactors as well.

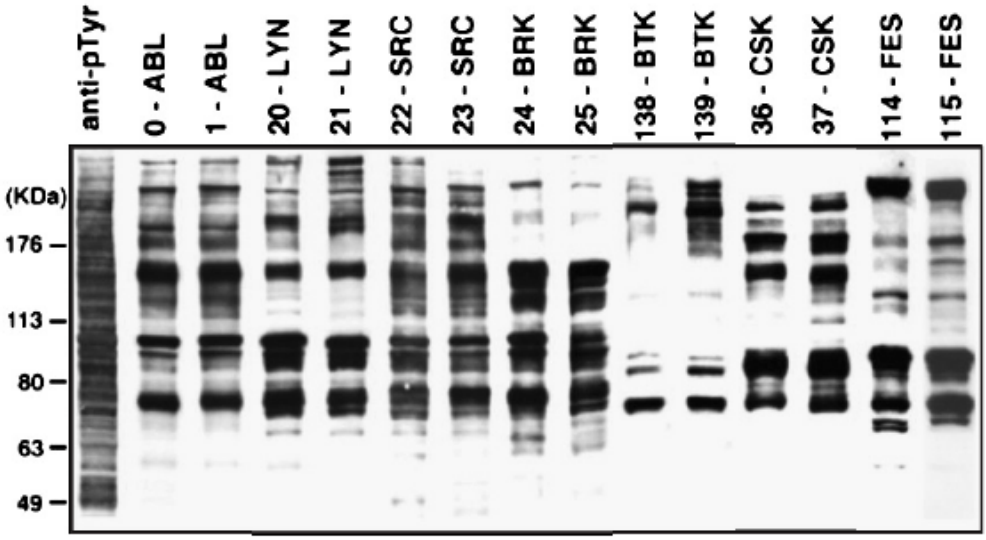
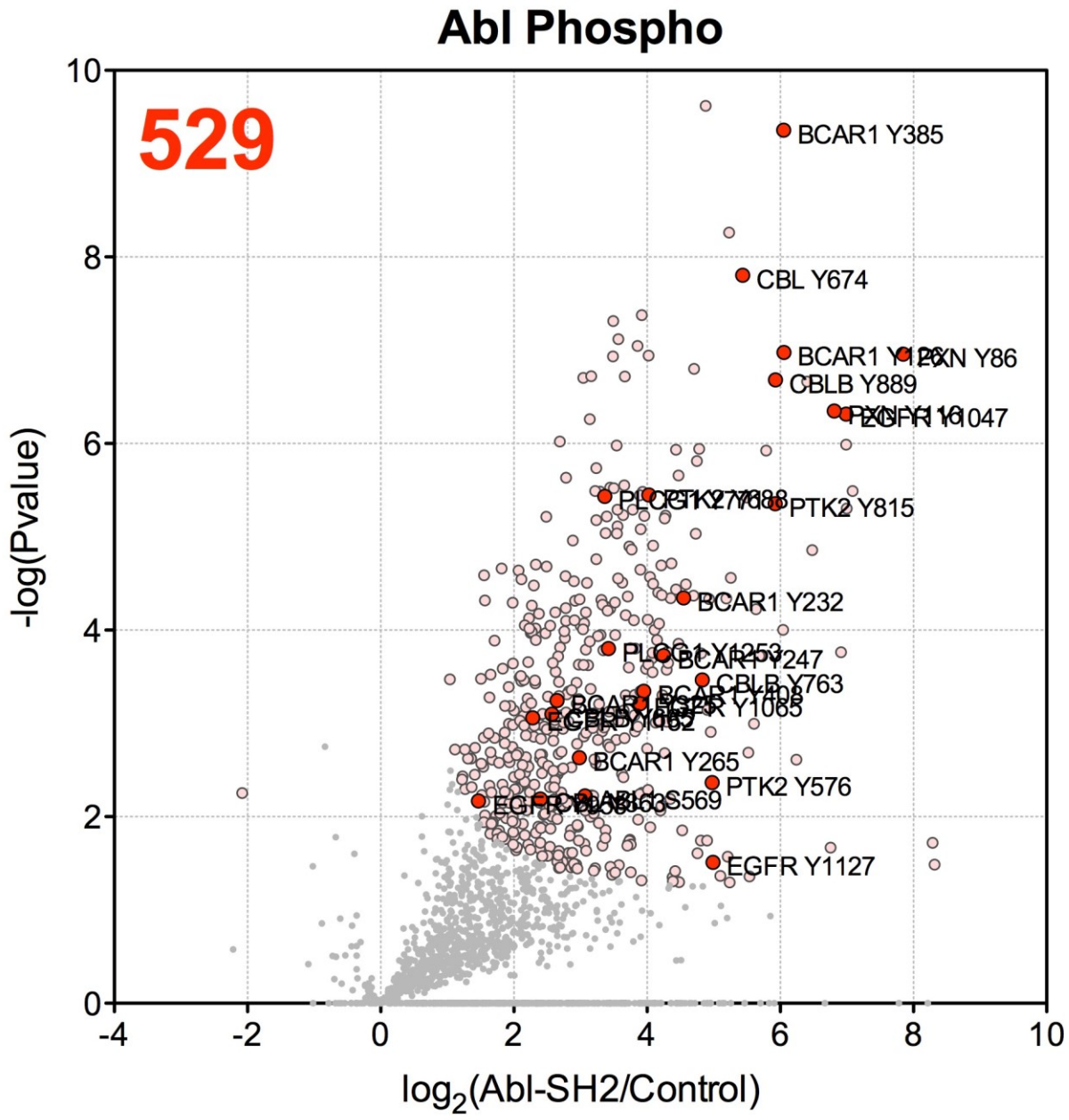
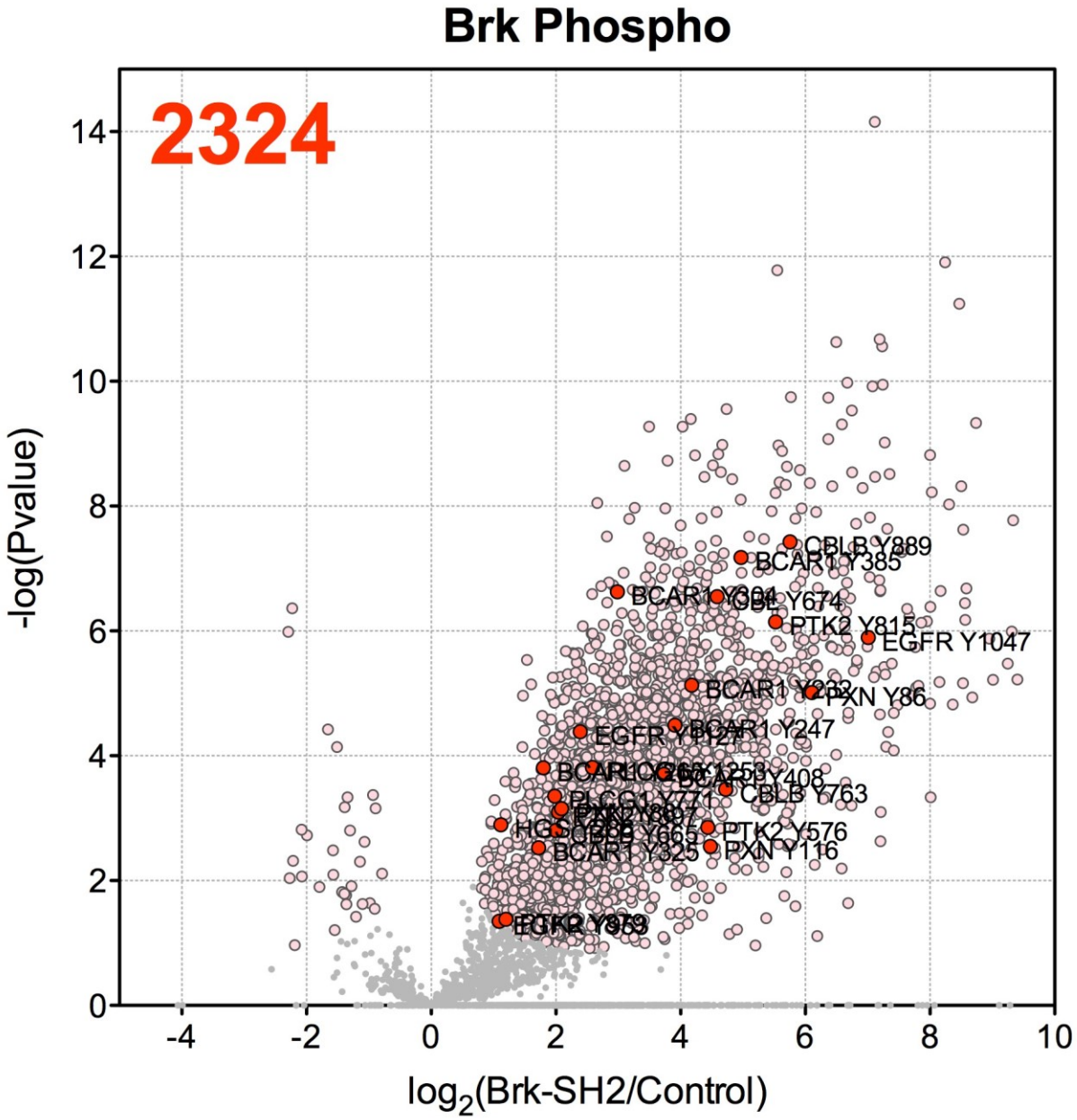


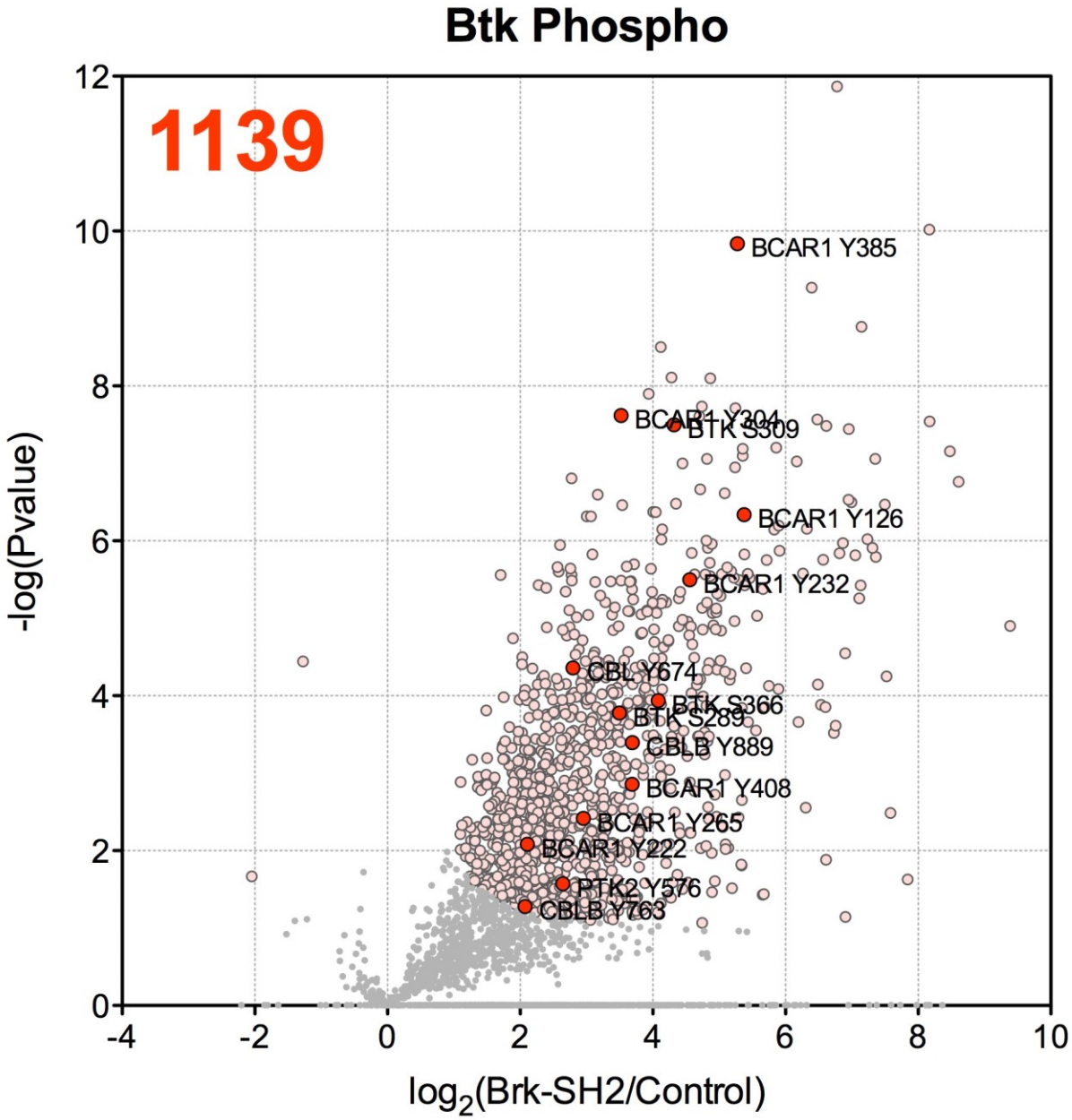
Figure 2-8. Binding preferences of the SH2 panel in duplicate as obtained by Machida et. al. against pervanadate-treated NIH 3T3, A431, HepG2-PDGFR, and MR20 cell lysate in a Far-Western assay. Reprinted with permission from the authors.

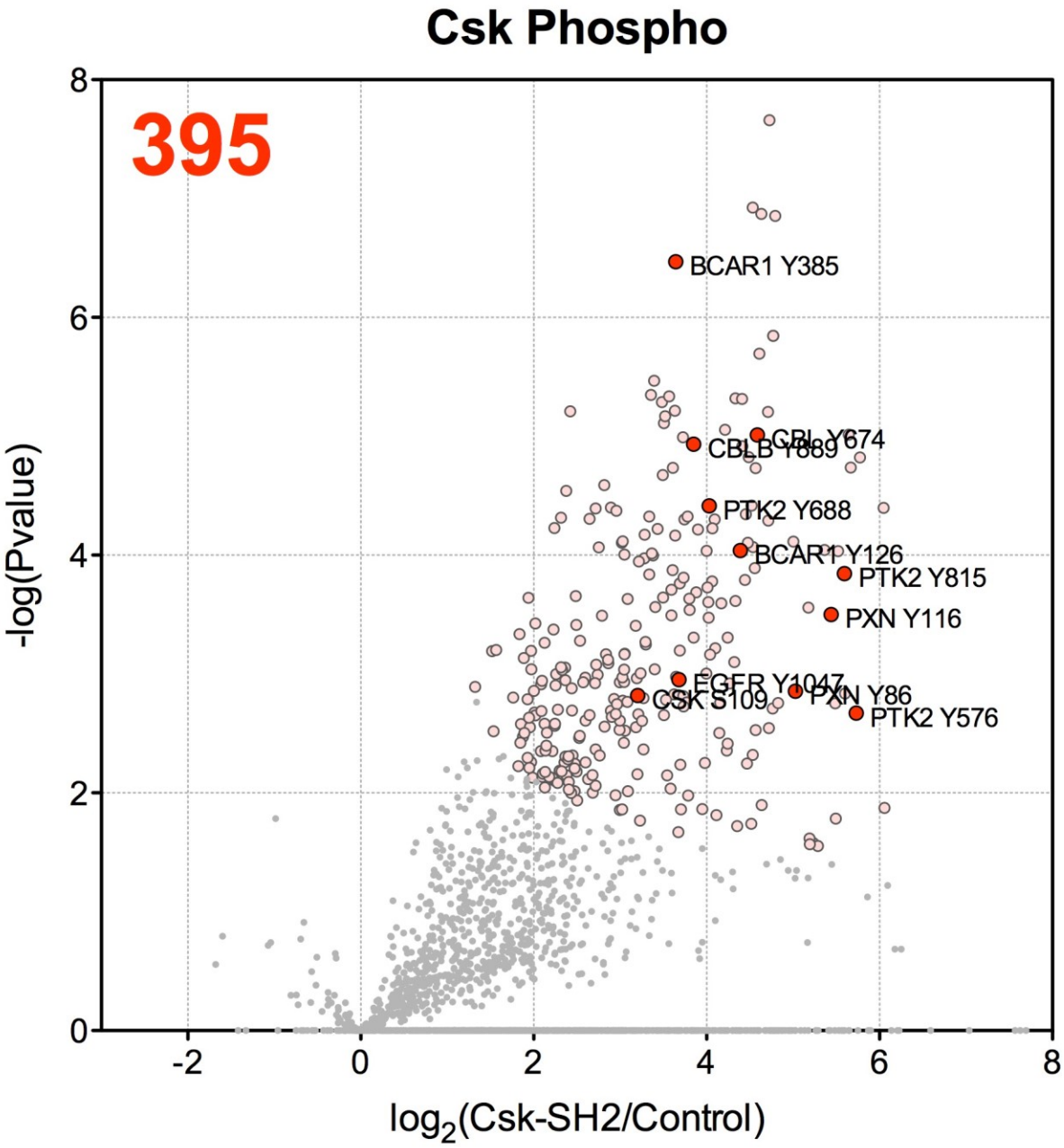
The enriched hits from each SH2 domain was further analyzed by gene ontology (GO) enrichment via the DAVID platform (24, 25) and then compared among the profiled SH2 domains. A summary of this analysis is shown in Table 2-1, which compares several shared, as well as unique GO biological process enrichment terms between the panel of SH2 domains. The overall number of enriched proteins within each GO term is listed for a given domain. In order to show the relative proportion that each GO term represents from the total protein hits for an SH2 domain, the terms in Table 2-1 are color coded according to the percentage of proteins within that term from the total number of proteins in the set. The SH2 domains from Src, Brk, Lyn, and Btk show strong enrichment for GO terms involved in RNA processing and splicing, whereas these terms are absent from Abl, Csk, and Fes. All SH2 domains except Btk's have the EGFR signaling pathway as an enriched term.

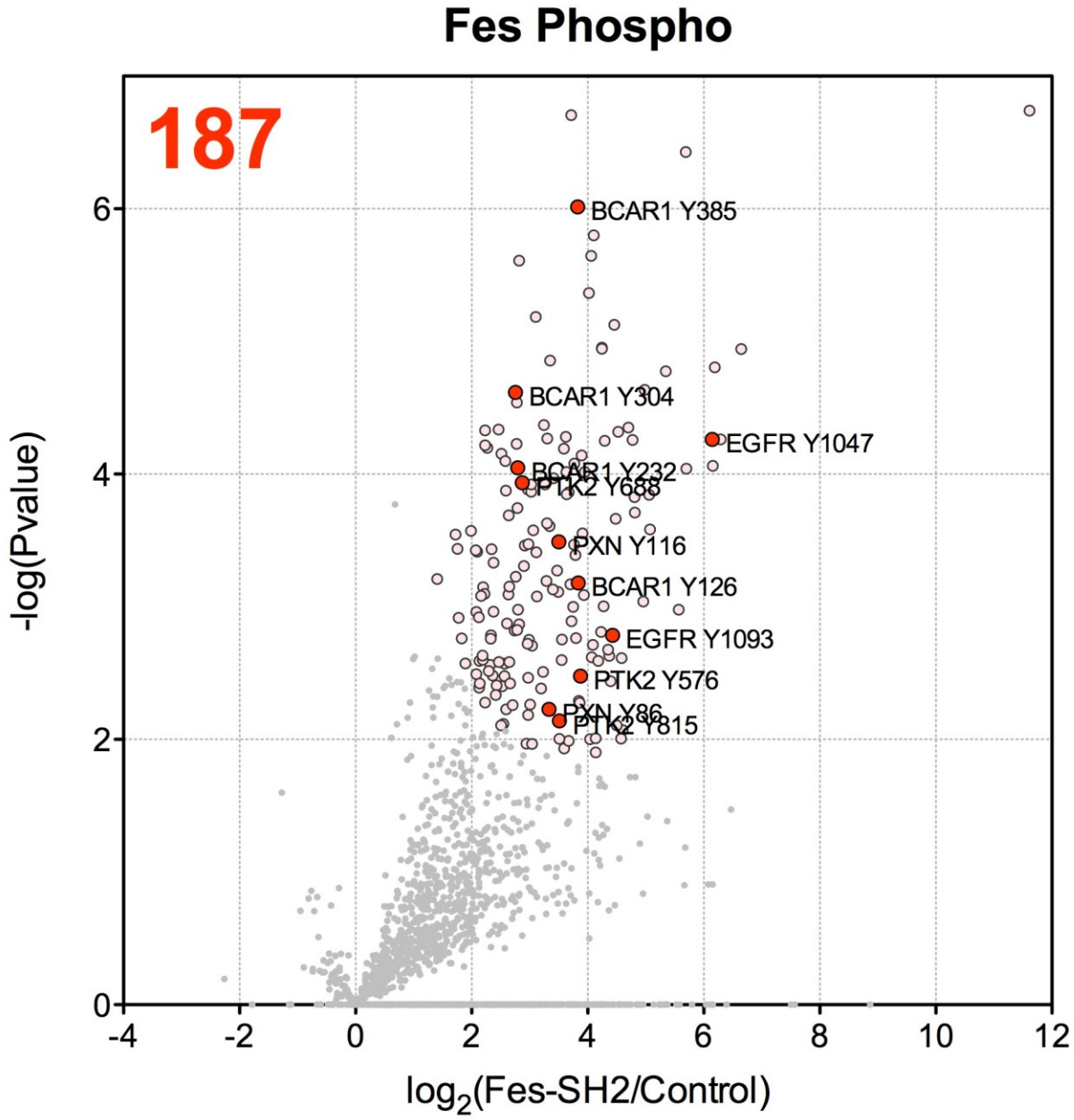
Phosphopeptide Anaysis of EGF-Stimulated HeLa Interactors




Brk Phospho
2324
-log(Pvalue)
log2(Brk-SH2/Control)
CBLB Y889
BCAR1 Y385
PTK2 Y815
EGFR Y1047
PXN Y86
BCAR1 Y247
CBLB Y763
PTK2 Y576
PXN Y116
BCAR1 Y325


Btk Phospho
1139
BCAR1 Y385
BCAR1 Y304
BTK S309
BCAR1 Y126
BCAR1 Y232
CBL Y674
BTK S366
BTK S289
CBLB Y889
BCAR1 Y408
BCAR1 Y265
BCAR1 Y222
PTK2 Y576
CBLB Y763
-log(Pvalue)
log2(Brk-SH2/Control)
12
10
8
6
4
2
0
-4
-2
0
2
4
6
8
10


Csk Phospho
395
BCAR1 Y385
CBL Y674
CBLB Y689
PTK2 Y688
BCAR1 Y126
PTK2 Y815
PXN Y116
EGFR Y1047
CSK S109
PXN Y86
PTK2 Y576
-log(Pvalue)
log2(Csk-SH2/Control)
8
6
4
2
0
-2
0
2
4
6
8


Fes Phospho
187
BCAR1 Y385
BCAR1 Y304
EGFR Y1047
BCAR1 Y232
PTK2 Y685
PXN Y116
BCAR1 Y126
EGFR Y1093
PTK2 Y576
PXN Y86
PTK2 Y815
-log(Pvalue)
log2(Fes-SH2/Control)
0
2
4
6
-4
-2
0
2
4
6
8
10
12

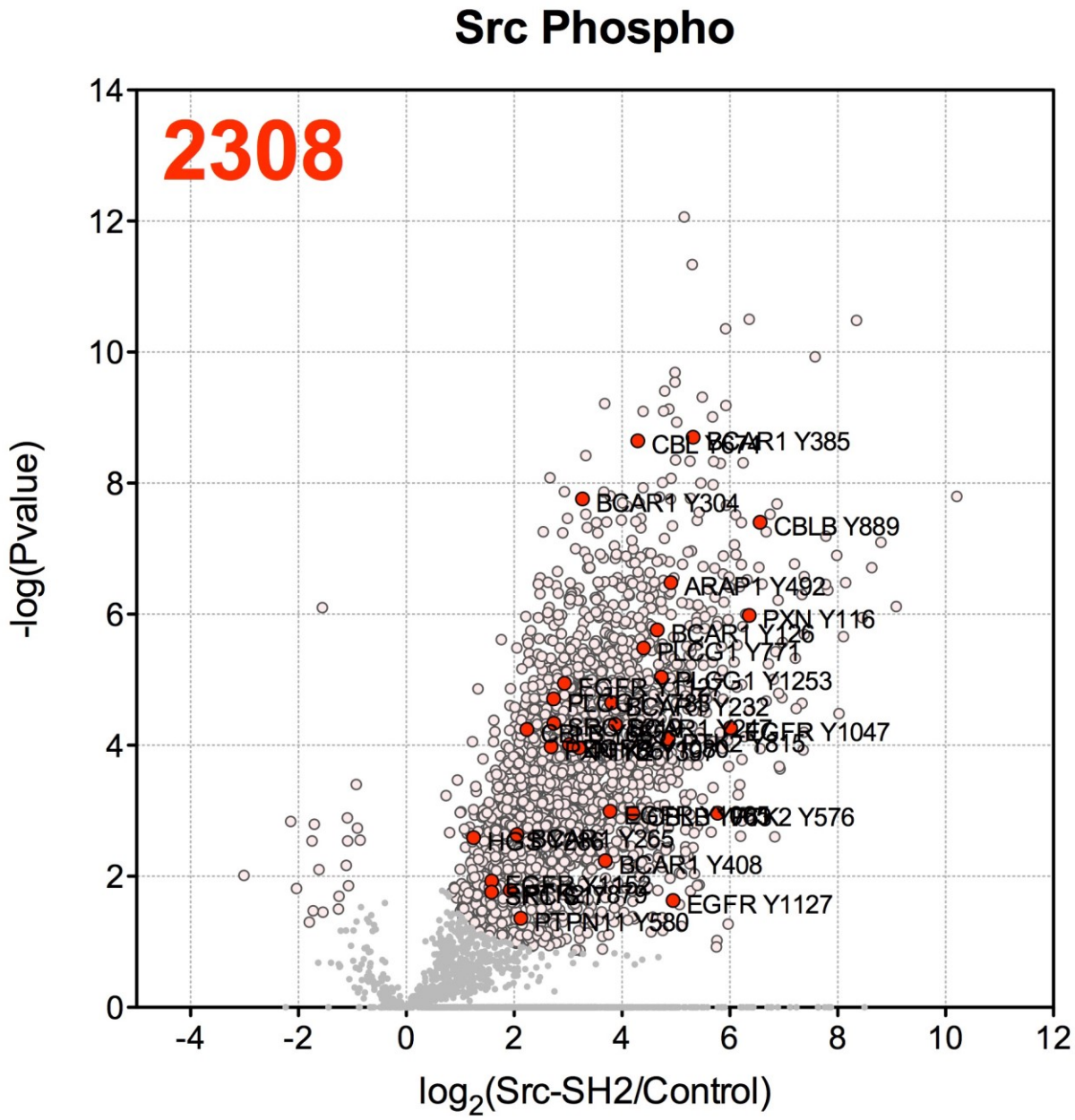


Figure 2-9. Volcano plots of the phosphosites enriched from EGF-stimulated HeLa cell lysate by the seven members of the SH2 domain panel. Significantly enriched targets are denoted in pink. Proteins that are designated in the EGFR signaling pathway by GO analysis have their phosphotyrosine sites in red.

Phosphopeptide enrichment using immobilized metal affinity chromatography (IMAC) was performed using a mixture of Sigma PHOSSelect and Fe- and Ga-immobilized resin to enrich

from the remaining 90% of trypsinized peptides. The number of enriched phosphosites, determined in the same way as protein interactors, for each SH2 domain mirrored what was observed at the protein level (Figure 2-9). The SH2 domains for Brk and Src again showed the largest amount of enriched phosphosites as shown in Table 2-2. In fact, Brk and Src themselves can account for over eighty percent of the total enriched phosphosites (Figure 2-10). Again, the SH2 domains for Src and Lyn show modest overlap in shared phosphosites, but Src SH2 has a significant fraction, ~65 percent, that is unique to its SH2 domain. The number of phosphotyrosine sites enriched also varies extensively by domain. Although the number of phosphotyrosine-containing peptides generally correlated with the overall level of phosphosite enrichment (see Brk, Src, and Lyn), there were some exceptions to this trend. For example, Btk, which enriched 1139 total phosphosites, only enriched a total of 43 phosphotyrosine-containing peptides. In contrast, Abl enriched 44 more phosphotyrosine-containing peptides than Btk, but only had a total of 529 significantly enriched phosphosites.

	ABL	BRK	BTK	CSK	FES	LYN	SRC
Enriched Sites	529	2324	1139	395	187	1001	2308
pY	87	121	43	51	37	118	139
Shared Sites	486	1926	1060	357	181	936	1960
Unique Sites	43	398	79	38	6	65	348

Table 2-2. Summary of enriched phosphosites for the panel of SH2 domains against EGF-stimulated HeLa cell lysate. The total number sites enriched by an SH2 domain in one or more biological replicates and the number of those that are tyrosine sites are listed. The number of sites that are shared by the other SH2 domains in the panel are listed as well as those that are unique for that SH2 domain only.

Phosphosite	log_2(Abl/ Control) Phospho	log_2(Brk/ Control) Phospho	log_2(Btk/ Control) Phospho	log_2(Csk/ Control) Phospho	log_2(Fes/ Control) Phospho	log_2(Lyn/ Control) Phospho	log_2(Src/ Control) Phospho
ABCF3 Y591___1	0.00	5.91	4.07	2.24	0.68	0.00	0.00
ANAPC1 Y571___1	0.00	4.01	2.45	0.00	0.00	0.00	0.00
ANKS1A Y455___1	3.67	3.06	0.00	0.00	3.95	3.98	4.27
ARAP1 Y492___1	0.00	4.57	0.00	0.00	0.00	4.97	4.91
ARHGAP42 Y342___1	0.00	2.78	0.00	0.00	0.00	0.00	3.72
ARHGEF40 Y242___1	0.00	4.26	2.80	0.00	0.00	4.97	4.45
ARHGEF5 Y656___1	1.58	1.90	0.00	1.96	0.00	0.00	1.61
ATP6V0D2 Y33___3	0.00	2.53	0.00	0.00	0.00	0.00	2.66
AXL Y598___1	0.00	0.00	0.00	0.00	0.00	1.78	1.79
BCAR1 Y126___1	6.06	4.90	5.38	4.39	3.84	5.40	4.65
BCAR1 Y232___1	4.55	4.18	4.56	1.85	2.80	2.82	3.81
BCAR1 Y247___1	4.25	3.91	2.57	0.00	0.00	4.64	3.87
BCAR1 Y265___1	2.98	1.80	2.95	0.00	0.00	3.22	2.05
BCAR1 Y325___1	2.65	1.72	2.46	0.00	0.00	0.00	1.83
BCAR1 Y385___1	6.05	4.97	5.27	3.64	3.83	5.69	5.32
BCAR1 Y408___1	3.95	3.73	3.69	3.20	3.04	4.22	3.69
BCAR1 Y304___1	3.44	2.99	3.52	2.04	2.76	3.71	3.27
BCAR1 Y222___1	0.00	2.73	2.11	0.00	0.00	1.60	0.00
BCAR3 Y42___1	3.68	1.96	2.45	2.54	2.53	3.00	2.46
BCAR3 Y117___1	0.00	4.81	0.00	0.00	0.00	0.00	4.68
BCLAF1 Y282___1	0.00	4.91	4.96	0.00	0.00	5.26	5.48
BCLAF1 Y282___2	0.00	3.18	2.41	0.00	0.00	0.00	2.81
BCLAF1 Y381___1	0.00	6.70	6.90	0.00	0.00	2.97	6.37
BRD2 Y637___1	0.00	4.47	0.00	0.00	0.00	0.00	0.00
CAV1 Y14___1	1.93	1.00	0.00	0.00	0.00	2.20	0.00
CBL Y674___1	5.44	4.59	2.79	4.59	2.85	4.99	4.29
CBL Y700___1	1.49	0.00	0.34	NaN	NaN	1.95	0.89
CBL;CBLB Y363___1	2.40	2.03	NaN	NaN	NaN	2.22	1.50
CBLB;DKFZp779A0729 Y665___1	2.57	2.00	0.00	0.00	0.00	2.31	2.24
CBLB;DKFZp779A0729 Y763___1	4.83	4.72	2.07	2.36	1.91	4.69	4.20
CBLB;DKFZp779A0729 Y889___1	5.93	5.75	3.69	3.85	0.00	5.62	6.56
CBLB;DKFZp779A0729 Y665___2	0.00	3.94	0.00	0.00	0.00	0.00	3.57
CCDC50 Y279___1	2.11	2.60	0.00	0.00	0.00	3.90	4.31
CCDC50 Y144___1	0.00	0.00	0.00	0.00	0.00	2.12	2.19
CCDC50 Y145___1	0.00	0.00	0.00	0.00	0.00	3.09	2.01
CDCP1 Y619___1	4.43	2.67	1.32	3.65	1.82	3.83	3.95
CDK1;CDC2;CDK2;CDK3 Y15___1	1.03	0.23	0.33	4.77	0.03	-0.38	0.47
CDK1;CDC2;CDK2;CDK3 Y15___2	2.25	1.91	1.45	3.73	1.42	1.54	1.17
CDK16;CDK17 Y136___1	0.00	0.00	0.00	4.83	0.00	0.00	0.00
CDK2;CDK3 Y15___1	1.03	0.23	0.33	4.77	0.03	-0.38	0.47
CDK2;CDK3 Y15___2	2.29	1.95	1.49	3.77	1.46	1.58	1.21
CDK5 Y15___1	0.00	0.83	0.00	3.51	0.00	0.00	0.85

CDKL5 Y171___1	0.00	2.04	0.00	0.00	0.00	0.00	1.32
CHCHD3 Y53___1	0.00	0.00	3.17	0.00	0.00	0.00	0.00
CTNND1 Y228___1	1.36	0.81	0.00	1.97	0.00	2.33	2.44
CTNND1 Y904___1	2.66	1.63	1.04	3.28	0.00	3.54	2.25
CTNND1 Y96___1	2.63	2.27	0.00	3.64	0.00	3.63	3.64
CTTN Y384___1	4.02	4.25	0.00	0.00	2.26	4.57	4.35
CTTN Y409___1	2.16	1.14	0.00	0.00	0.00	2.06	2.81
CYSRT1 Y52___1	0.91	0.00	0.00	0.00	0.00	1.71	0.00
DCBLD1 Y621___1	1.73	1.55	1.38	0.72	0.00	3.05	1.91
DCBLD1 Y578___1	0.00	1.18	0.00	0.00	0.00	0.00	1.47
DSP Y846___3	3.20	0.00	0.00	0.00	0.00	3.82	3.26
DSP Y849___3	2.87	0.00	0.00	0.00	0.00	3.49	2.93
DYRK1A;DYRK1B Y321___1	0.03	0.90	0.27	-0.52	-0.26	0.58	1.70
EFNB2 Y304___1	3.04	1.93	-0.34	0.00	2.09	2.69	2.21
EGFR Y1047___1	6.99	7.01	0.00	3.68	6.15	6.18	6.01
EGFR Y1065___1	3.89	3.32	0.69	1.84	1.04	3.13	3.78
EGFR Y1127___1	4.99	6.15	2.54	2.81	3.28	4.68	4.95
EGFR Y1152___1	2.28	0.99	-1.40	0.58	0.91	2.41	1.58
EGFR Y953___1	1.47	1.09	1.05	0.80	0.55	2.18	0.93
EGFR Y1127___2	0.00	2.39	0.00	0.00	0.00	0.00	2.93
EGFR Y1093___1	0.00	0.00	0.00	0.00	4.43	5.93	6.70
EGFR Y1080___2	0.00	0.00	0.00	0.00	0.00	0.00	3.03
EMD Y161___1	2.35	2.96	3.09	4.72	3.67	3.15	3.19
EMD Y94___1	2.39	2.25	1.96	2.10	1.39	2.51	1.98
EPHA2 Y575___1	3.57	0.73	0.00	0.00	0.00	3.24	1.86
EPHA2 Y588___2	3.90	2.94	0.00	0.00	0.00	2.63	3.28
EPHA2 Y594___1	4.23	1.95	0.00	0.00	0.00	4.09	0.00
EPHA2 Y772___1	4.28	1.76	1.17	2.25	2.08	3.91	2.81
EPHA2 Y588___1	2.41	0.02	0.00	1.24	0.00	2.85	1.18
EPHA2 Y594___2	3.72	2.91	0.00	0.00	0.00	2.08	2.51
EPHA5;EPHA3;EPHA4 Y670___1	1.98	-0.09	1.47	1.14	0.00	1.64	1.01
EPHA7 Y592___1	2.60	2.51	0.00	0.00	1.92	3.29	3.10
EPHA7 Y786___1	2.77	2.52	0.72	2.79	1.15	2.90	2.57
EPHB4 Y581___1	2.49	1.60	0.00	0.00	0.00	2.80	2.27
EPHB4 Y774___1	2.32	1.38	NaN	NaN	NaN	2.86	2.09
EPS15 Y715___1	0.00	0.00	0.00	0.00	0.00	0.00	2.97
EPS15L1 Y564___1	0.00	3.30	0.00	0.00	0.00	0.00	3.18
ERBB2 Y972___1	3.89	3.14	1.98	NaN	4.60	3.97	3.10
ERBB2 Y601___1	3.11	3.04	0.00	0.00	2.46	3.19	2.58
ERRFI1 Y394___1	0.00	2.07	0.00	0.00	0.00	0.00	2.77
FAF2 Y79___1	2.07	2.01	NaN	4.49	-0.26	3.19	3.08
FAM175B Y342___1	2.23	2.46	0.00	0.00	0.00	2.76	0.00
FAM63A Y3___1	0.00	0.00	0.00	0.00	0.00	1.99	0.59
FCHSD2 Y523___1	0.00	2.06	0.00	0.00	0.00	3.49	0.00
GAB1 Y156___1	3.29	2.37	0.00	2.48	0.00	3.31	2.83
GAB1 Y556___1	4.09	4.27	2.55	3.56	2.00	5.66	5.76
GAB2 Y576___2	0.00	0.00	0.00	0.00	0.00	2.18	2.76

GAB2 Y605___1	0.00	0.00	0.00	0.00	0.00	2.79	1.41
GIT1 Y545___1	6.99	6.38	5.65	4.10	6.19	6.92	6.90
GPRC5A Y300___1	4.93	0.00	4.68	0.00	4.78	0.00	0.00
GPRC5A Y317___1	5.26	3.28	4.72	4.88	4.98	4.83	3.91
HBS1L Y36___1	NaN	2.26	NaN	NaN	NaN	2.88	2.33
HGS Y286___1	1.24	1.12	0.00	1.30	0.00	2.14	1.25
HIPK1;HIPK2 Y352___1	0.14	NaN	-0.55	0.47	0.49	2.15	1.19
HNRNPA3 Y101___1	1.21	3.84	1.34	-0.04	0.81	1.86	3.57
HYDIN Y54___3	0.00	0.00	1.82	4.07	0.00	0.00	0.00
IFITM3 Y20___1	0.00	3.26	0.00	0.00	0.00	2.92	4.50
IMMP2L Y11___1	NaN	NaN	2.46	4.11	3.20	2.06	NaN
INPPL1 Y1069___1	4.47	3.44	2.10	4.00	2.74	4.96	5.40
INPPL1 Y920___1	3.76	3.39	3.57	3.59	3.30	5.75	4.66
INPPL1 Y820___1	0.00	1.15	0.00	0.00	0.00	4.20	3.01
ITSN2 Y951___1	3.54	2.24	0.00	4.41	0.00	4.14	3.40
KIF1B;KIF1Bbeta Y111___3	0.83	0.85	0.40	1.48	1.35	NaN	2.05
KIRREL Y535___1	2.22	1.88	0.00	1.97	3.10	2.60	1.85
KIRREL Y538___1	2.25	1.56	0.00	2.23	1.41	1.84	1.39
LDLR Y930___1	1.46	1.66	0.00	1.43	0.00	1.35	0.70
LRRD1 Y842___2	2.05	2.62	3.15	NaN	NaN	2.19	2.41
MAPK1 Y187___1	2.39	1.95	3.06	1.47	2.86	4.23	3.76
MAPK1 Y187___2	0.00	3.24	2.60	0.00	0.00	0.00	2.67
MAPK14 Y105___1	2.71	2.84	1.75	2.21	2.84	2.79	2.47
MAPK3 Y90___1	2.20	2.30	1.61	NaN	1.56	3.70	2.62
MLLT1 Y158___1	0.00	4.13	0.00	0.00	0.00	0.00	0.00
MPP5 Y209___1	0.00	0.00	0.00	0.00	0.00	2.91	1.92
MPZL1 Y117___1	4.72	2.00	0.00	0.00	4.04	0.00	4.83
NCOR1 Y89___1	0.00	1.97	0.00	0.00	0.00	0.57	1.55
NEDD9 Y166___1	3.16	2.28	3.17	1.89	3.71	2.92	2.34
NEDD9 Y345___1	3.05	1.97	3.87	0.00	3.41	3.27	2.18
NEDD9 Y92___1	2.27	1.34	2.03	1.65	2.27	2.15	1.60
NEDD9 Y214___1	3.39	2.41	4.14	0.00	4.02	3.72	2.93
NEDD9 Y317___1	3.52	1.79	3.83	0.00	3.62	3.46	2.24
NUP160 Y1116___1	0.00	4.58	5.02	4.71	4.46	4.85	5.18
NUP160 Y1116	0.00	4.58	5.02	4.71	4.46	4.85	5.18
PADI6 Y444___3	2.43	1.65	3.07	0.11	2.81	1.53	1.05
PAG1 Y227___1	2.67	1.40	1.08	1.57	2.97	1.58	1.82
PAG1 Y359___1	2.40	1.21	0.85	1.03	3.05	1.74	1.59
PALMD Y140___1	0.00	0.00	0.00	0.00	0.00	0.00	2.39
PARD3 Y1020___1	1.92	1.64	0.00	1.93	0.00	3.00	2.80
PEAK1 Y635___1	3.89	3.69	0.00	3.59	2.18	4.31	4.29
PIK3C2B Y244___1	2.60	0.56	1.20	1.33	2.51	2.76	1.62
PIK3C2B Y228___1	0.00	2.99	0.00	0.00	0.00	0.00	3.01
PKP4 Y1166___1	3.14	2.39	0.79	3.39	2.23	3.19	2.78
PKP4 Y1137___1	0.00	1.22	0.00	0.00	0.00	0.00	1.39
PLA2G4A Y535___1	1.59	1.54	0.00	0.00	0.00	0.00	0.00
PLCG1 Y1253___1	3.42	2.59	0.00	0.00	0.00	4.94	4.73

PLCG1 Y771___1	3.36	1.98	0.00	0.00	0.00	4.21	4.40
PLCG1 Y783___1	0.00	0.00	0.00	0.00	0.00	2.67	2.73
PLEKHA6 Y512___1	0.00	0.00	0.00	3.21	0.00	2.30	0.00
POLR2A Y1853___1	0.00	1.48	0.00	0.00	0.00	0.00	1.87
PRKCD Y313___1	0.00	1.32	0.00	1.82	0.00	1.74	2.28
PRPF4B Y849___1	2.11	5.54	3.80	NaN	NaN	3.22	4.51
PRR14L Y1691___3	NaN	2.85	0.10	NaN	1.54	NaN	1.44
PTK2 Y576___1	4.98	4.44	2.64	5.73	3.88	7.71	5.77
PTK2 Y688___1	4.03	3.56	0.00	4.03	2.87	4.02	0.00
PTK2 Y815___1	5.92	5.52	1.72	5.59	3.51	4.74	4.86
PTK2 Y397___1	2.63	2.04	0.00	0.00	0.00	3.81	3.21
PTK2 Y879___1	0.00	1.20	0.00	1.96	0.00	2.26	1.93
PTK2 Y720___1	0.00	0.00	0.00	0.00	0.00	2.38	2.51
PTK2B Y579___1	0.71	-1.19	0.00	0.79	0.00	2.26	0.80
PTK2B Y580___1	1.11	0.53	0.00	1.67	0.00	2.71	1.77
PTK2B Y792___1	2.12	0.00	0.00	0.00	0.00	1.95	0.47
PTPN11 Y580___1	1.28	0.45	NaN	-0.26	-0.45	2.86	2.12
PTPN18 Y194___1	1.68	-0.02	0.00	0.69	0.00	0.96	0.47
PXN Y116___1	6.81	4.48	1.50	5.44	3.50	4.22	6.36
PXN Y86___1	7.85	6.11	0.61	5.03	3.33	5.97	5.40
PXN Y86___2	2.98	2.09	0.00	0.00	0.00	1.55	2.69
RAB7A Y136___1	2.27	2.52	0.00	0.00	0.00	3.68	3.45
RBCK1 Y288___1	0.00	3.94	0.00	0.00	0.00	0.00	3.46
RIN1 Y36___1	3.08	3.33	1.53	0.00	0.00	3.33	0.00
RPL38 Y43___1	0.00	0.00	0.00	2.82	0.00	0.00	0.00
RRAGC Y16___1	0.00	3.05	0.00	0.00	0.00	3.77	4.59
SCN2A Y548___3	0.00	2.87	4.57	0.00	3.59	0.00	3.14
SDCBP Y46___1	3.86	0.00	0.00	0.00	0.00	3.34	3.19
SEC16A Y1390___1	0.00	0.00	3.09	0.00	0.00	0.00	3.11
SGK223 Y413___1	3.55	1.90	0.84	4.63	1.15	3.83	3.14
SHB Y246___1	4.27	2.85	2.53	2.57	0.00	3.48	3.72
SHB Y268___1	0.00	-0.39	0.00	0.00	0.00	1.65	0.00
SHC1 Y317___1	1.31	1.55	0.12	0.26	0.03	2.06	0.82
SLITRK5 Y704___1	0.00	3.36	0.00	0.00	0.00	4.29	3.72
SPACA5 Y24___2	2.76	1.89	0.00	0.00	2.59	2.11	2.58
STAM Y384___1	0.00	0.00	0.00	0.00	0.00	1.28	0.00
STAM2 Y192___1	3.96	3.05	0.00	2.27	0.00	3.55	3.15
STAM2 Y374___1	2.26	0.00	0.00	0.00	0.00	2.17	0.54
STAM2 Y371___1	2.88	1.27	0.00	2.90	0.00	2.33	0.00
STAT3 Y607___1	2.09	0.82	1.13	1.02	1.16	2.01	0.78
STAT5B;STAT5A Y699___1	3.92	3.53	0.00	3.52	0.00	3.05	3.89
STX4 Y167___1	0.00	0.00	3.01	2.54	3.86	0.00	0.00
TGFB1I1 Y43___1	5.23	4.61	0.11	4.79	0.57	4.74	4.87
THOC7;NIF3L1BP1 Y19___1	5.85	5.20	6.90	5.17	5.67	6.47	5.75
THRAP3 Y293___1	0.00	1.42	5.09	0.00	0.00	3.33	4.32
TJP2 Y1095___1	2.02	1.47	0.00	2.96	0.00	0.00	3.04
TNK2 Y859___1	4.75	2.97	1.43	3.69	1.42	3.92	3.56

TNK2 Y891___1	3.57	2.72	0.00	2.75	0.00	4.84	3.52
TNK2 Y891___2	4.09	3.29	0.00	0.00	0.00	4.17	4.14
TNK2 Y892___2	4.17	3.36	0.00	0.00	0.00	4.24	4.21
TOM1L2 Y354___1	4.17	4.62	NaN	NaN	NaN	4.31	3.04
TRA2B;TRA2A Y264___2	0.00	2.18	0.00	0.00	0.00	0.00	2.59
TRIM41 Y139___2	NaN	NaN	NaN	NaN	NaN	NaN	2.25
TTC6 Y850___1	0.00	0.00	0.00	6.05	0.00	0.00	0.00
TYK2 Y71___1	1.75	2.48	1.86	1.85	2.13	2.86	1.85
TYRO3 Y804___1	2.53	2.10	0.00	0.00	0.00	4.41	3.60
TYRO3 Y783___1	0.00	0.00	0.00	0.00	0.00	2.17	0.81
TYRO3;MERTK Y636___1	1.71	0.61	0.00	0.00	1.02	2.98	1.56
VAV3 Y171___1	2.87	2.80	0.00	0.00	0.00	0.00	0.00
WDR3 Y10___2	0.00	1.83	0.20	0.00	0.00	0.00	2.89
WDR3 Y7___2	0.00	2.25	0.63	0.00	0.00	0.00	3.31
Y1602___2	3.21	1.11	3.37	2.66	4.35	1.58	0.40
Q5TH74 Y16___3	0.00	2.71	0.00	0.00	0.00	0.00	2.59
Q5T5M9 Y186___2	0.00	5.70	2.74	0.00	0.00	3.75	5.53

Table 2-3. Phosphotyrosine sites enriched by at least one member of the SH2 domain panel in EGF-stimulated HeLa cell lysate. Relative enrichment for the site is heatmapped as the log2 ratio for the enrichment of each SH2 domain over the phosphotyrosine control.

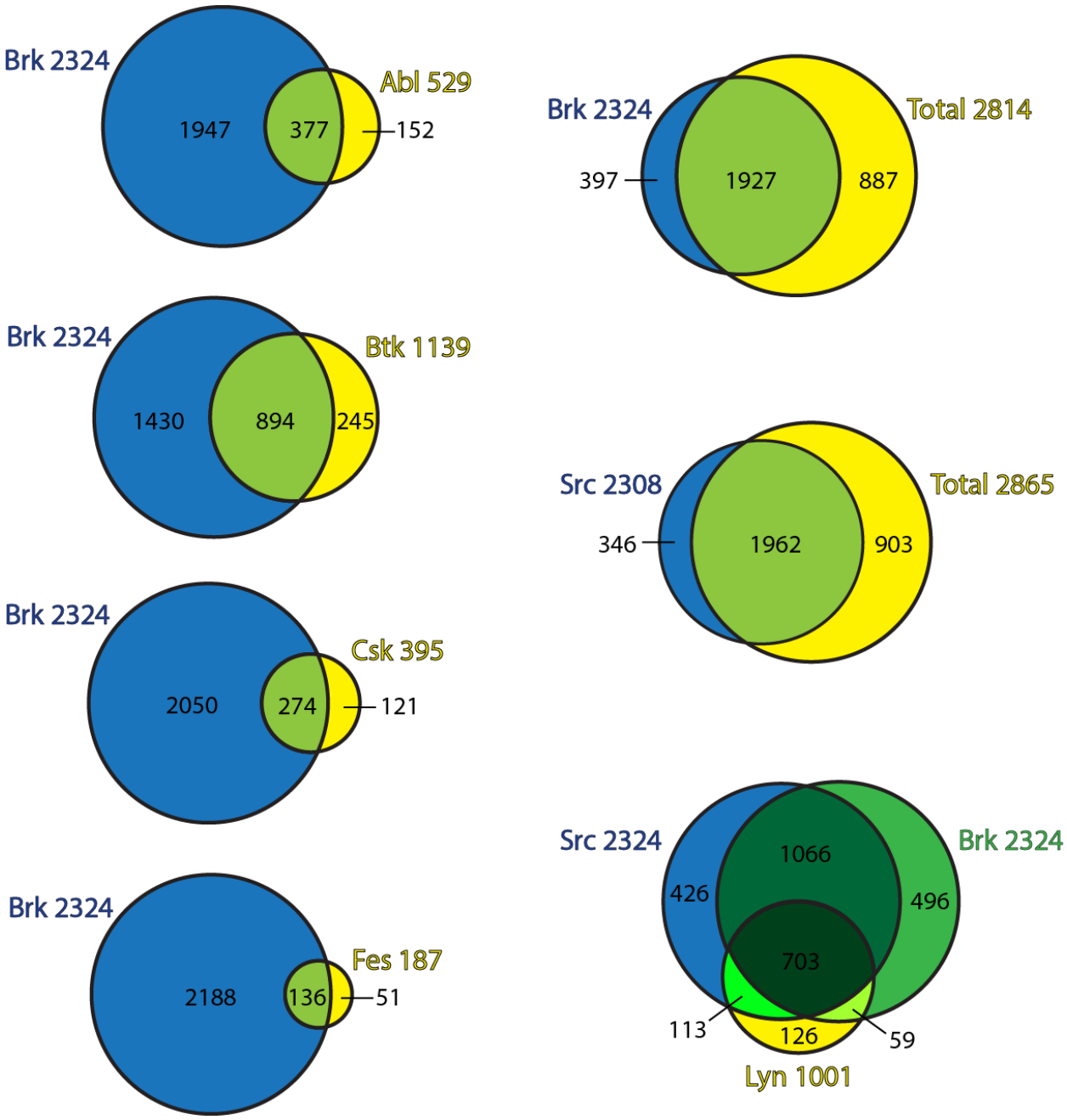


Figure 2-10. Venn diagrams comparing the phosphosite enrichment for each SH2 domain in EGF-stimulated HeLa cell lysate. The relative proportion that Src and Brk, the two SH2 domains that had the highest enrichment, with the rest of the panel are compared most extensively.

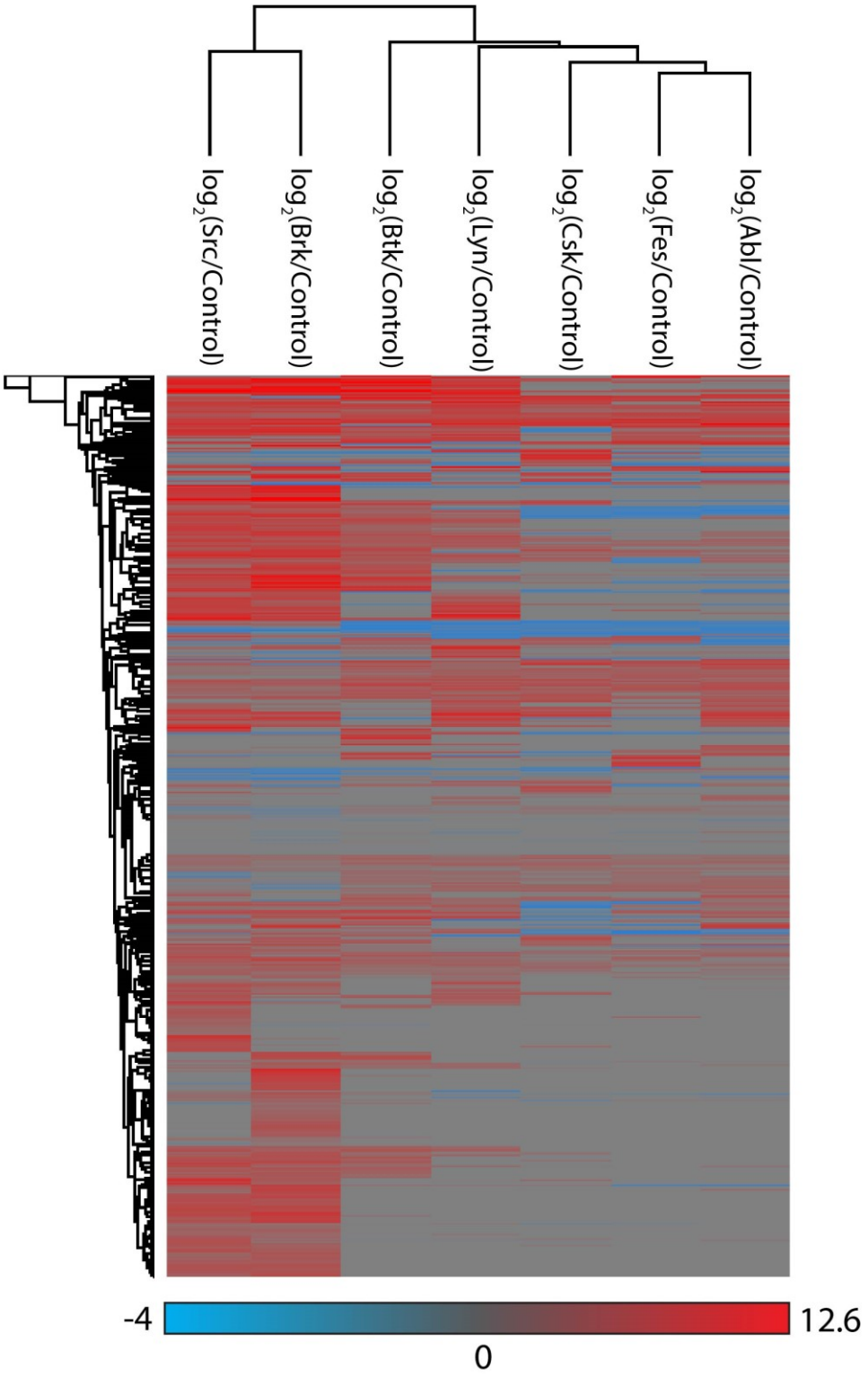


Figure 2-11. Hierarchical clustering of the log2 difference in phosphopeptide enrichment between SH2 domain over the control in EGF-stimulated HeLa cell lysate.

Hierarchical clustering of the log2 ration of enrichment between each SH2 domain over the phosphotyrosine control revealed the same enrichment relationship between SH2 domains as that observed previously (compare Figures 2-7 and 2-11). Based on phosphopeptide enrichment, Src and Brk cluster together and branch off separately from Lyn, Btk, Fes, and Abl.

Given that SH2 domains bind to phosphotyrosine, we chose to focus further analysis on these sites which are listed in detail in Table 2-3. Because previous microarray studies have established differential binding preferences around the phosphotyrosine for different SH2 domains, we expected to observe very distinct lists of significantly enriched phosphotyrosine sites for this panel. The enriched phosphotyrosine sites are actually relatively similar with a majority of their shared tyrosine sites part of core EGFR signaling machinery. The differences in tyrosine between the SH2 domains appears more around the periphery of these signaling components.

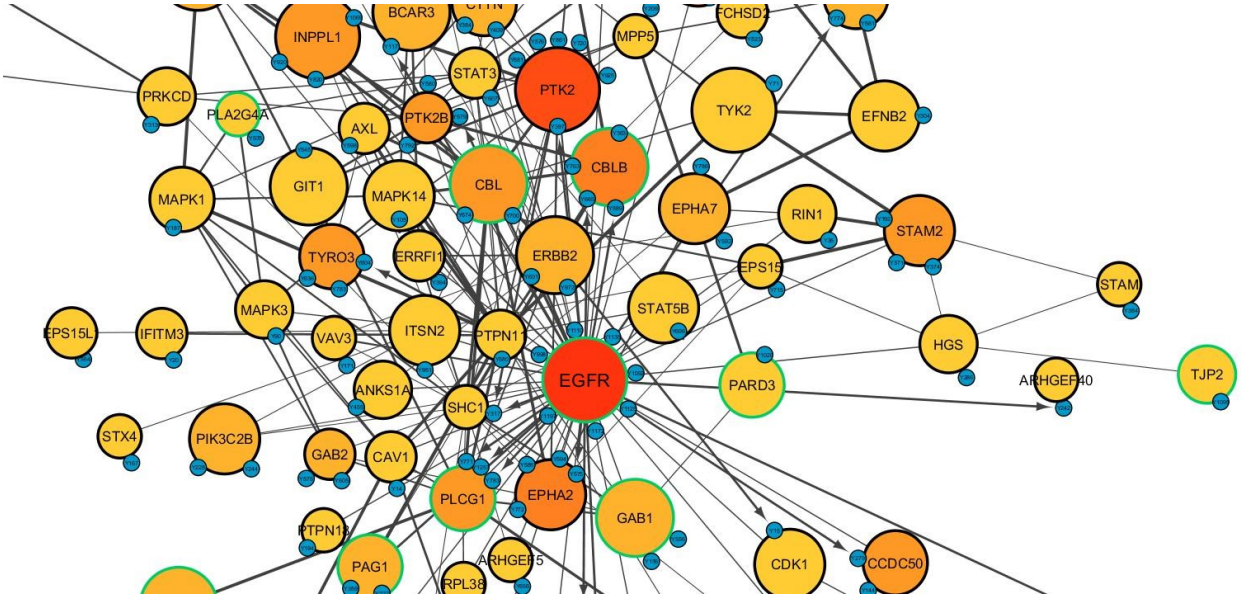


Figure 2-12. Interaction map generated with PhosphoPath and Cytoscape illustrating the density of phosphotyrosine at a given protein (node color) and the number of SH2 domains in the panel that enriched at least one tyrosine site on that protein target (node size) for EGF-stimulated HeLa. Phosphotyrosine sites are blue nodes. Proteins with phosphotyrosine sites that showed differential enrichment with Lyn and Src against +-EGF lysate pools are outlined in green.

To visualize these differences among the phosphotyrosine network, we used the database from STRING and PhosphoPath, which is an application in Cytoscape (Figure 2-12) to generate an interactome of the proteins bearing tyrosine sites enriched by at least one member of the SH2 domain panel. The proteins containing phosphotyrosine are plotted as nodes with any interactions between proteins and sites curated from the STRING database in addition to Biogrid and

PhosphositePlus. The node size represents the number of SH2 domains in the panel that enriched at least one phosphotyrosine on the given protein. The color of the node is heat mapped to show the density of phosphotyrosine enriched for that protein. Additionally, proteins with phosphotyrosine sites that had stimulation-specific enrichments with Lyn and Src SH2 are outlined in green. Larger nodes, which are phosphotyrosine-containing peptides enriched by multiple members of the SH2 panel, are clustered around EGFR and downstream signaling members such as Cbl, PTK2, STAT5B, BCAR1 in addition to other kinases such as TYK2, EPHA2, EPHA7 and p38a. The smaller nodes, which are phosphotyrosine-containing peptides enriched by fewer SH2 domains, tend to lie on the periphery of the EGFR-MAPK signaling pathway.

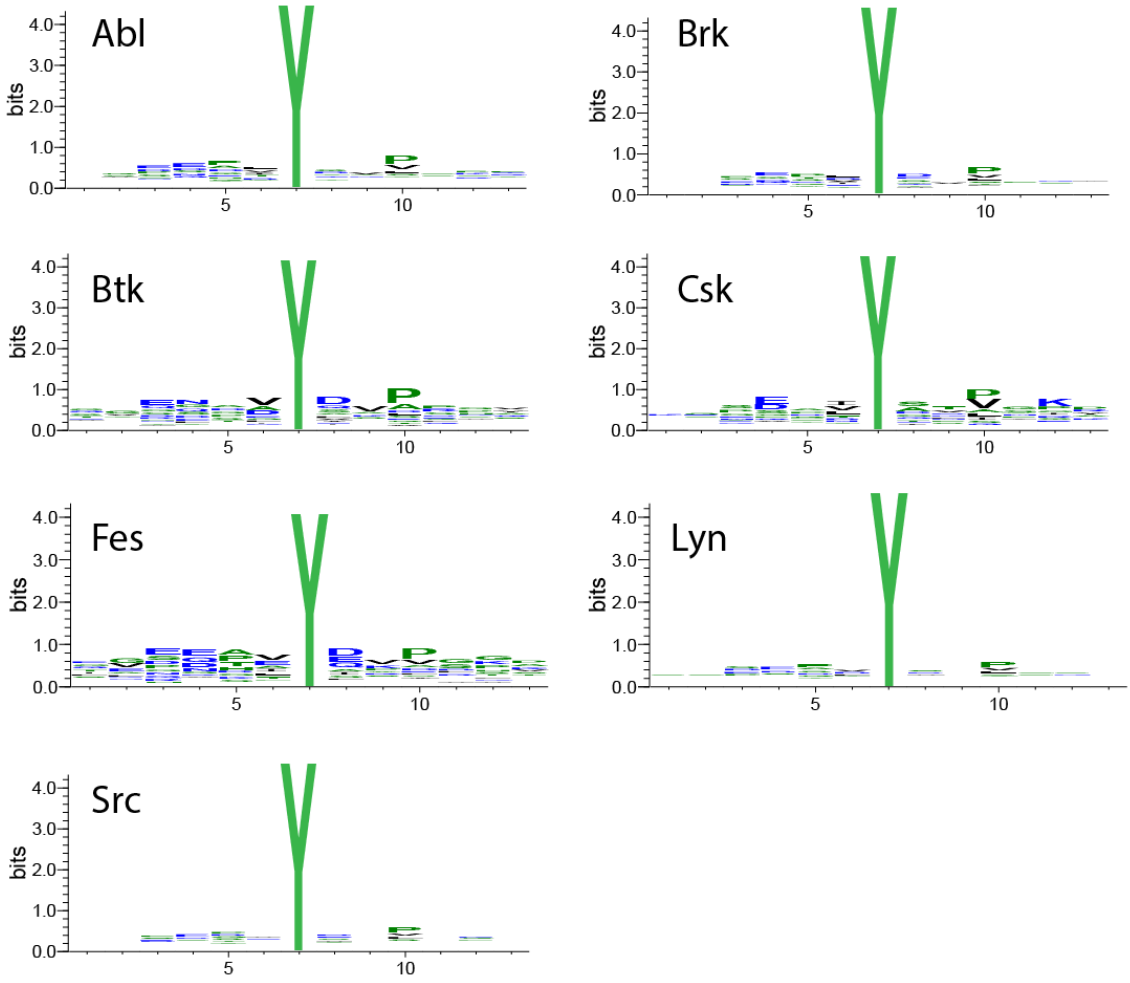


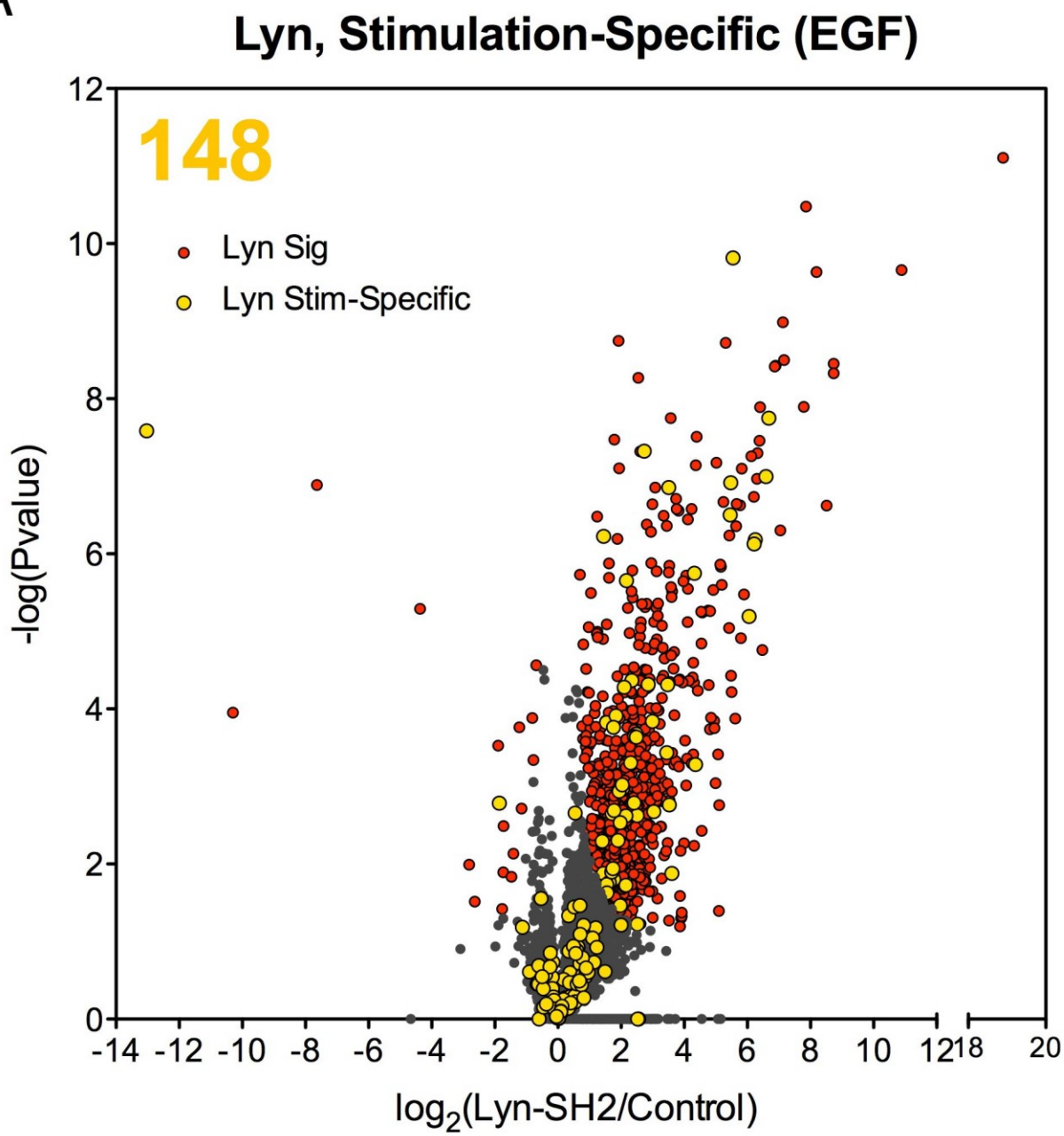
Figure 2-13. Sequence logos of the SH2 domain panel based on the phosphotyrosine sites enriched from EGF-stimulated HeLa cell lysate. Logos were generated using WebLogo3.

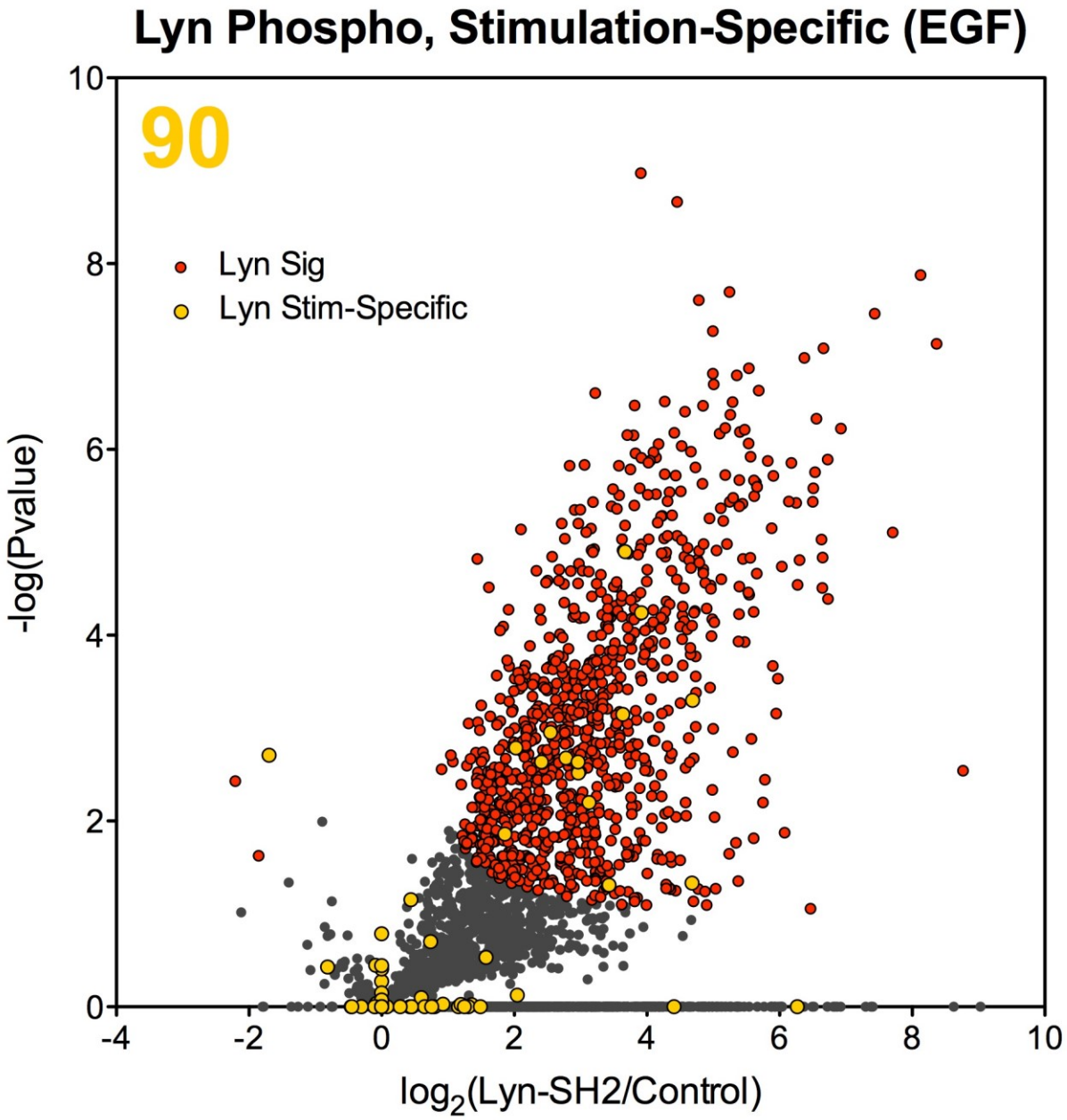
For each SH2 domain, the sequence specificity was analyzed based on the enriched phosphotyrosine peptides. WebLogo3 was used to generate sequence logos of the phosphotyrosine sites plus six amino acids N- and C-terminal to the phosphotyrosine (Figure 2-13). The amino acid sequence motifs appear muted for all SH2 domains, which is likely a combination of having multiple different tyrosine sites on the same protein and some number of sites that are actually secondary interactors. However, we do observe a relationship between the relative ambiguity around the phosphotyrosine to the overall number of phosphosites enriched. For example, Src and Brk have relatively undistinguishable sequence features on the sequence logo analysis and showed fairly promiscuous enrichment of proteins and phosphosites. Conversely, Csk and Fes had much more limited enrichment profiles and thus display more specific recognition motifs around the tyrosine.

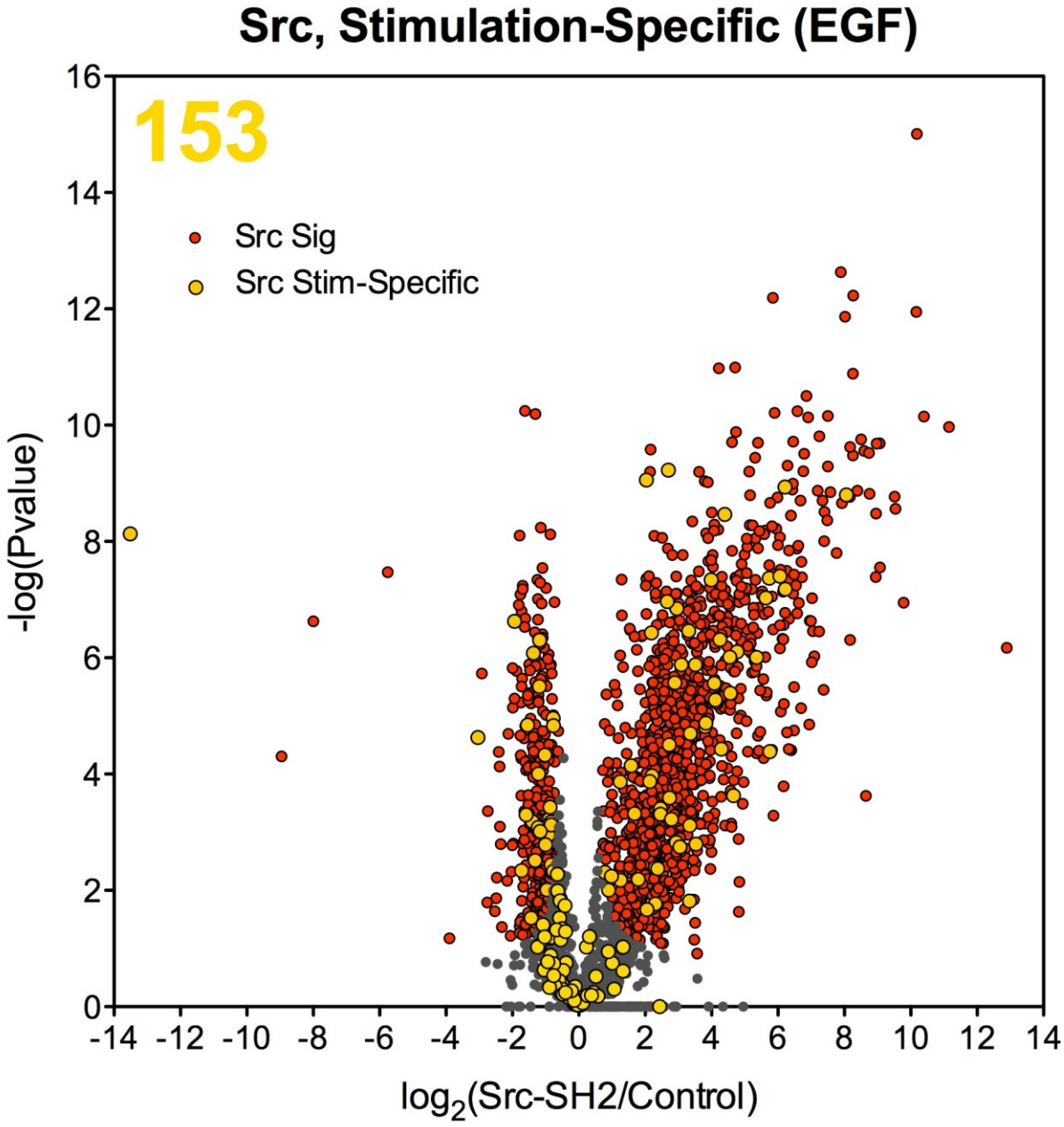
Comparison of Lyn and Src SH2 Interactions in Stimulated and Non-Stimulated HeLa

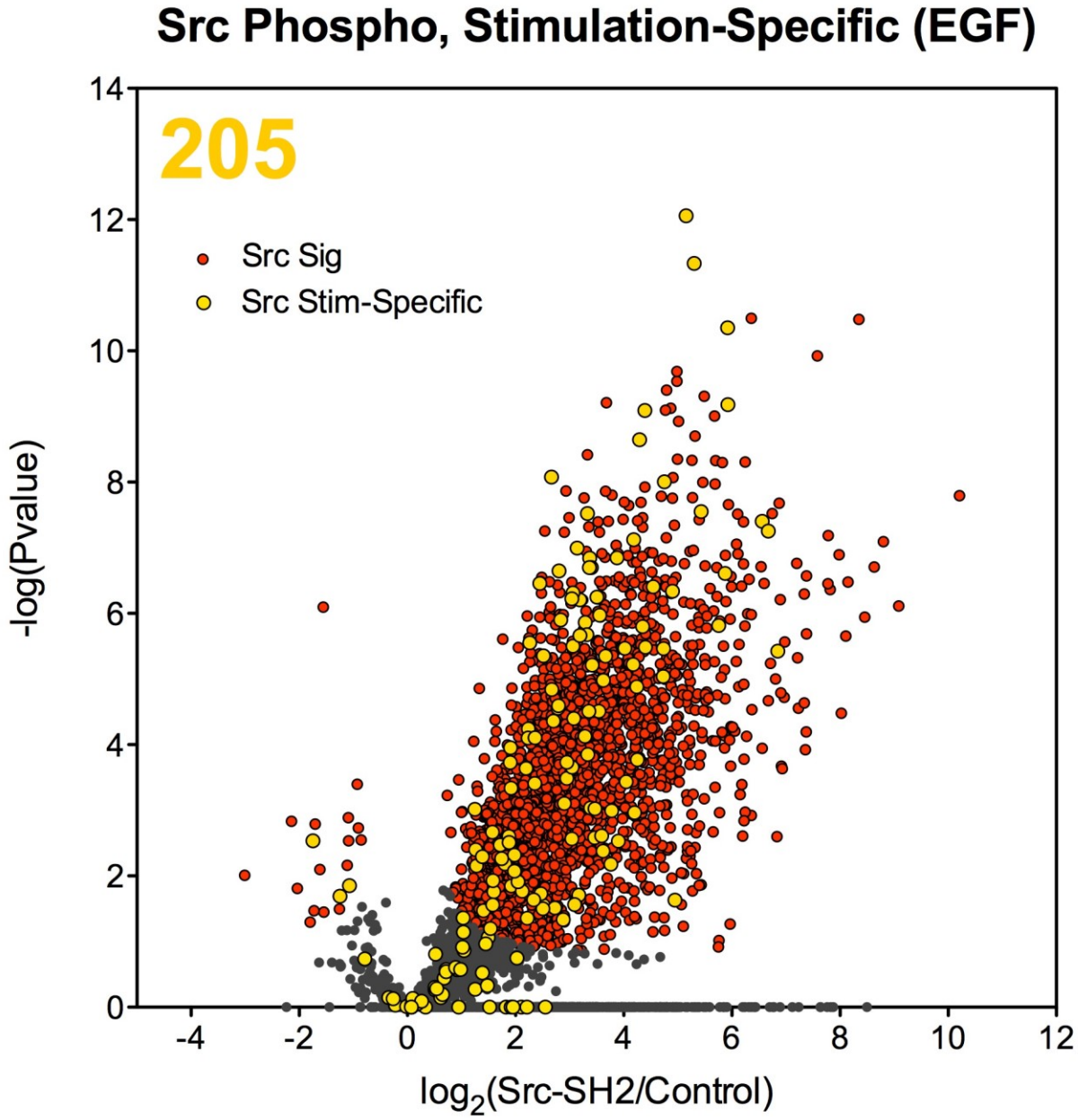
The enriched proteins and phosphosites for the SH2 domain panel are composed of a mixture of targets that arose because of EGF stimulation and those that are independent of that event. In order identify the subset of those proteins and sites that are EGF stimulation-dependent, non-stimulated HeLa cell lysate was prepared and incubated with immobilized SH2 domains from Lyn and Src. We then compared differences in the proteins and phosphopeptides that were enriched by the SH2 domains of Src and Lyn between unstimulated and EGF-stimulated HeLa cell lysate. This comparison was done using a two-sample *t*-test comparing SH2 domain enrichment in the stimulated pool versus non-stimulated. The results of these stimulation-dependent hits are summarized in Figure 2-14.

A


Lyn, Stimulation-Specific (EGF)
148
Lyn Sig
Lyn Stim-Specific
-log(Pvalue)
log2(Lyn-SH2/Control)
0
2
4
6
8
10
12
-14
-12
-10
-8
-6
-4
-2
0
2
4
6
8
10
12
18
20


Lyn Phospho, Stimulation-Specific (EGF)
90
Lyn Sig
Lyn Stim-Specific
-log(Pvalue)
log2(Lyn-SH2/Control)
10
8
6
4
2
0
-4
-2
0
2
4
6
8
10


Src, Stimulation-Specific (EGF)
153
Src Sig
Src Stim-Specific
-log(Pvalue)
log2(Src-SH2/Control)
16
14
12
10
8
6
4
2
0
-14 -12 -10 -8 -6 -4 -2 0 2 4 6 8 10 12 14


Src Phospho, Stimulation-Specific (EGF)
205
Src Sig
Src Stim-Specific
-log(Pvalue)
log2(Src-SH2/Control)
14
12
10
8
6
4
2
0
-4
-2
0
2
4
6
8
10
12

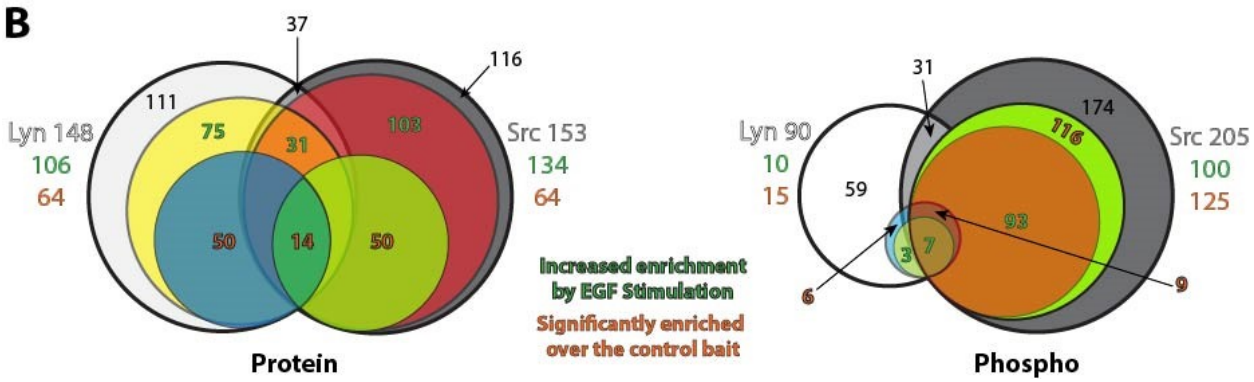
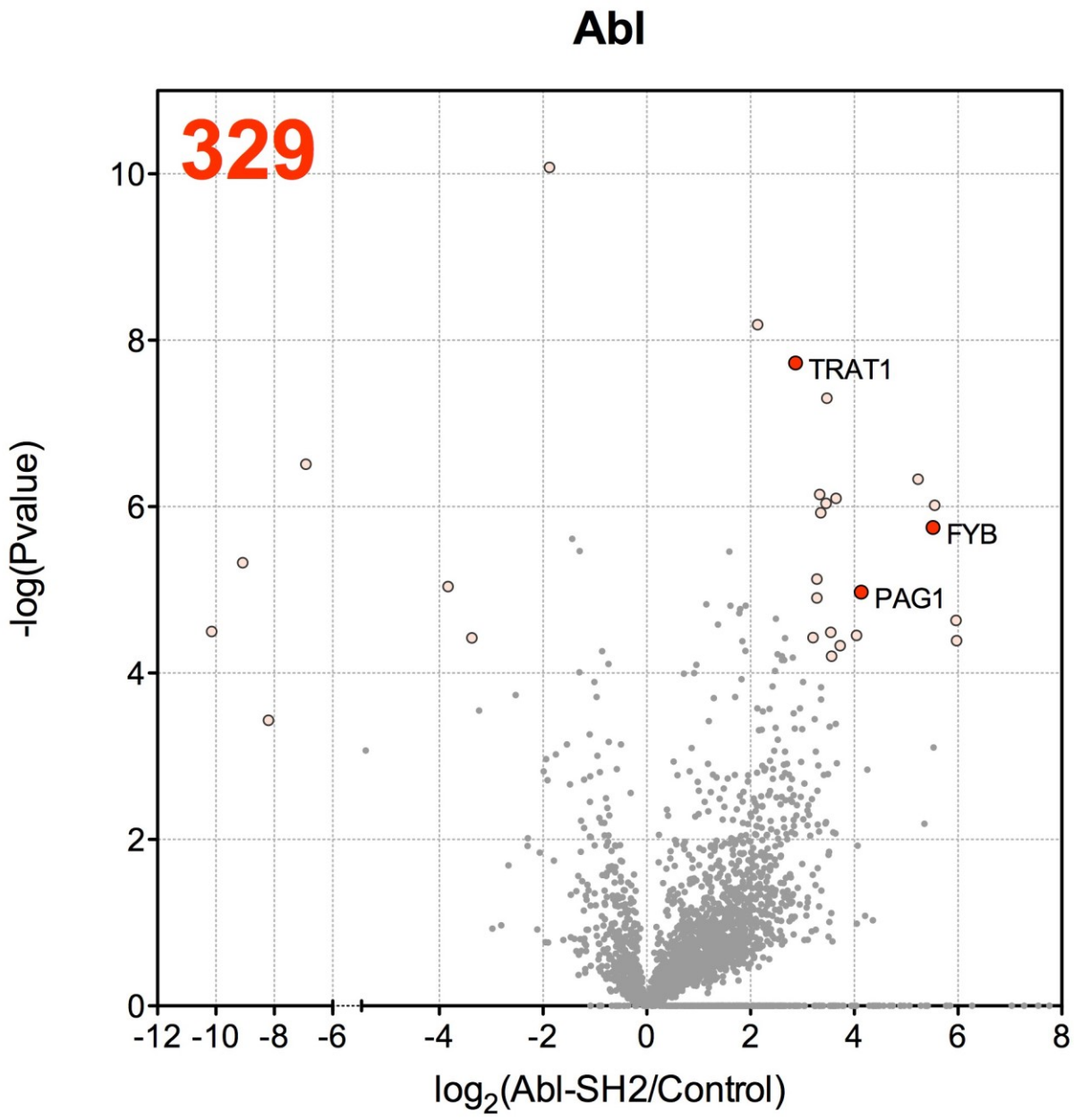


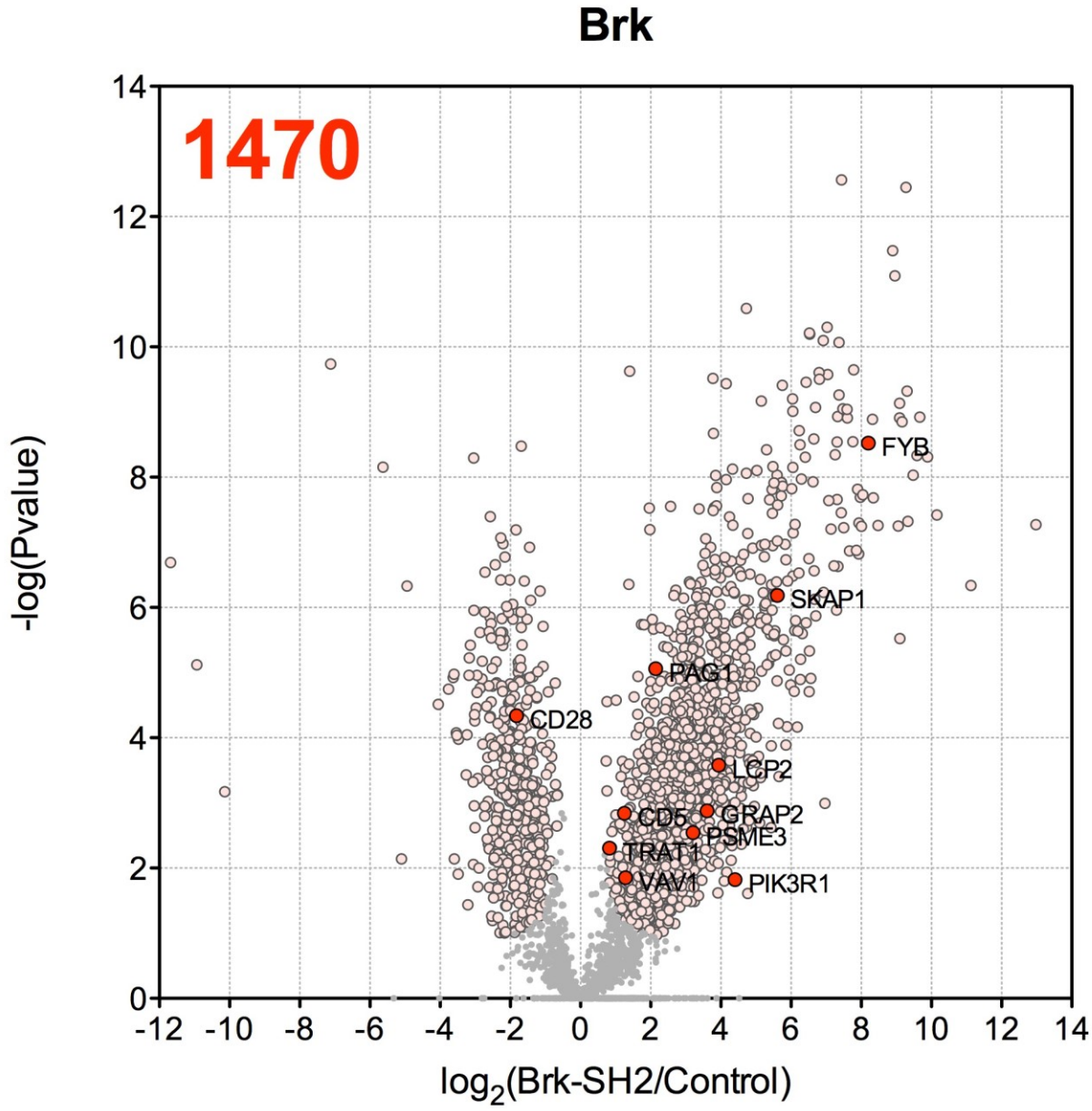
Figure 2-14. Summary of the EGF-stimulation-dependent proteins and phosphosites enriched by Lyn and Src SH2 domain. (A) Volcano plots overlayed with stimulation-specific enrichment by Lyn and Src highlighted in yellow. (B) Venn diagrams comparing the total number of proteins or phosphosites that showed significant differential enrichment between +-EGF treatment. The subset of those proteins/phosphosites that were enriched more with EGF treatment are in green text while those that were significantly enriched by the respective SH2 domain over the control bait in EGF stimulated HeLa are in red text.

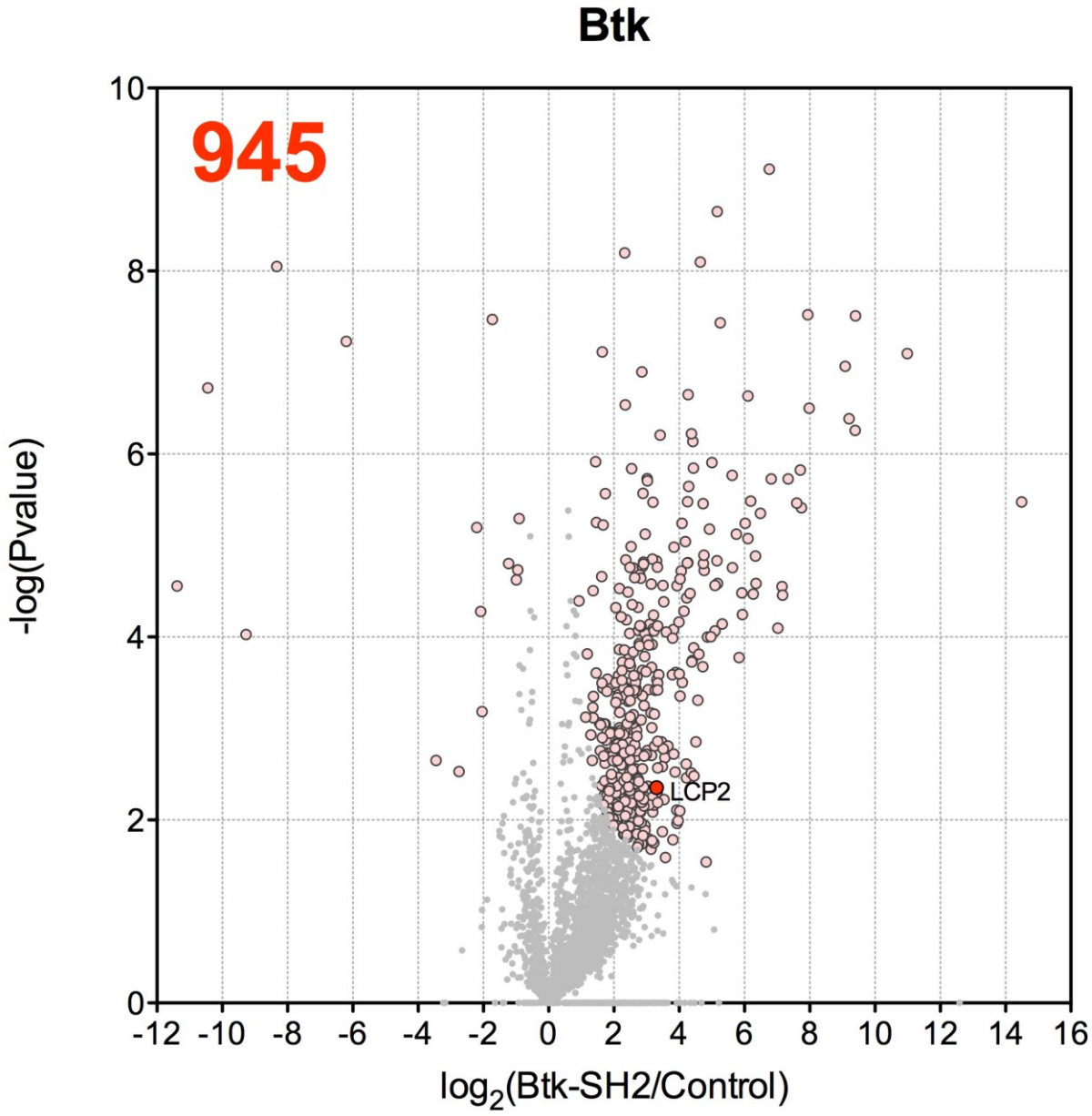
The proteins and phosphosites that are differentially captured depending on EGF activation are from both interactions specific to Lyn or Src SH2 domain and from those proteins or phosphopeptides that are enriched equally well by the phosphotyrosine dead control. In fact, only about half of these are SH2 domain-specific interactions as shown in Figure 2-14B. The other half are proteins or phosphopeptides that the SH2 domain and the phosphotyrosine dead control were able to enrich. Most of the activation-dependent hits are pulled down upon EGF stimulation, but there is a significant fraction that actually decreased upon activation. Lyn SH2 enriched 90 stimulation-dependent phosphopeptides from HeLa cell lysate, but only 10 of these phosphosites are enriched more in the stimulated cells than unstimulated (Figure 2-14B). Lyn SH2 also had only 90 stimulation-dependent phosphosites compared to 148 proteins. The overlap between Src and Lyn's significant hits when only comparing stimulation-dependent hits is decreased from their comparison of total significant hits from EGF-stimulated HeLa lysate (Compare Figures 2-6, 2-10, and 2-14). Closer examination of the activation-dependent phosphotyrosine sites and where they lie with regards to the EGFR signaling cascade show that EGFR, Cbl, Cblb, GAB1, PLCG1, and PAG1 had phosphotyrosine sites that were activation-specific, but a majority the other proteins did not show differential capture based on stimulation (Figure 2-12).

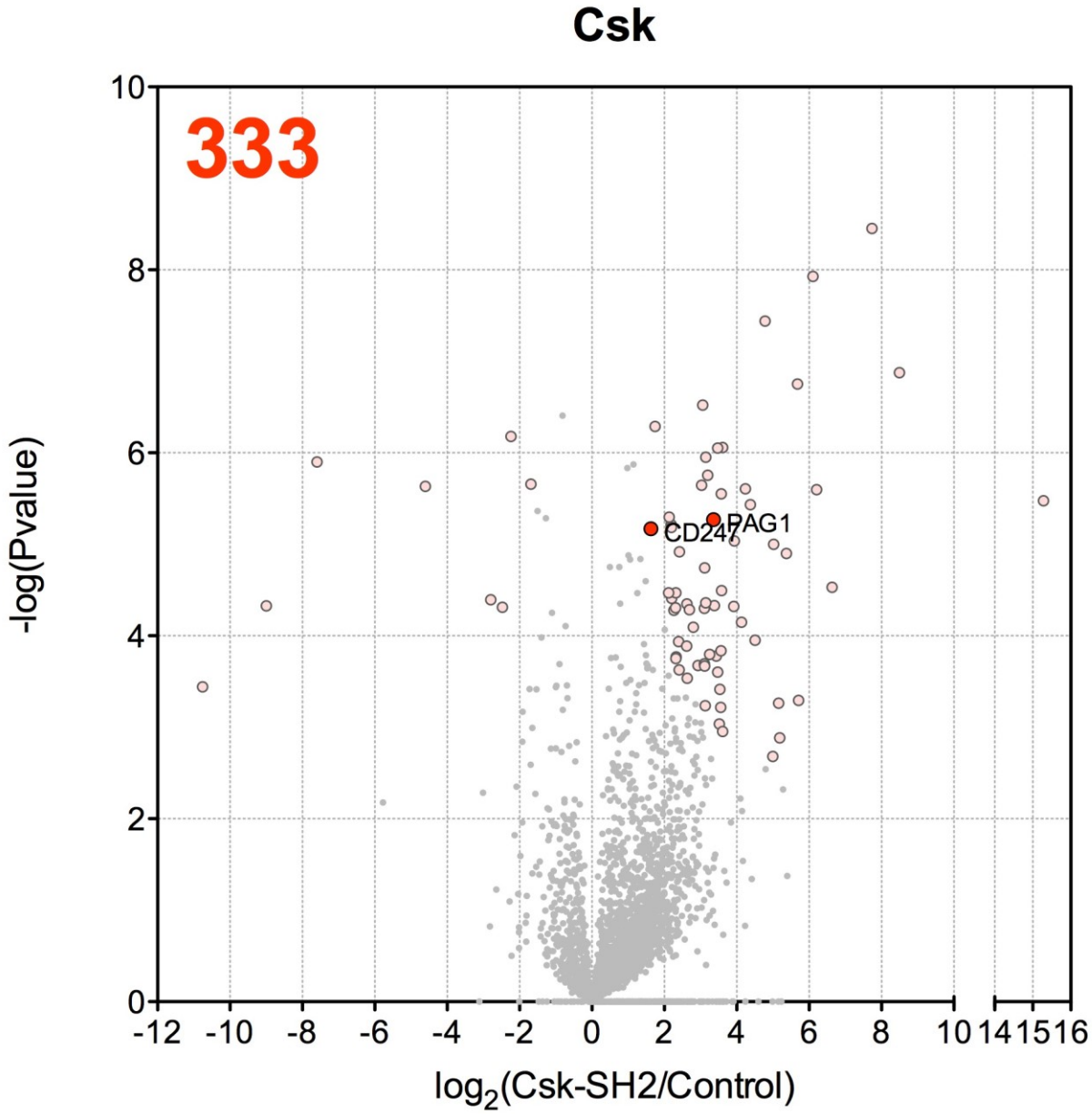
MS Characterization of SH2 Domain Interactions in T Cell Activated Jurkat

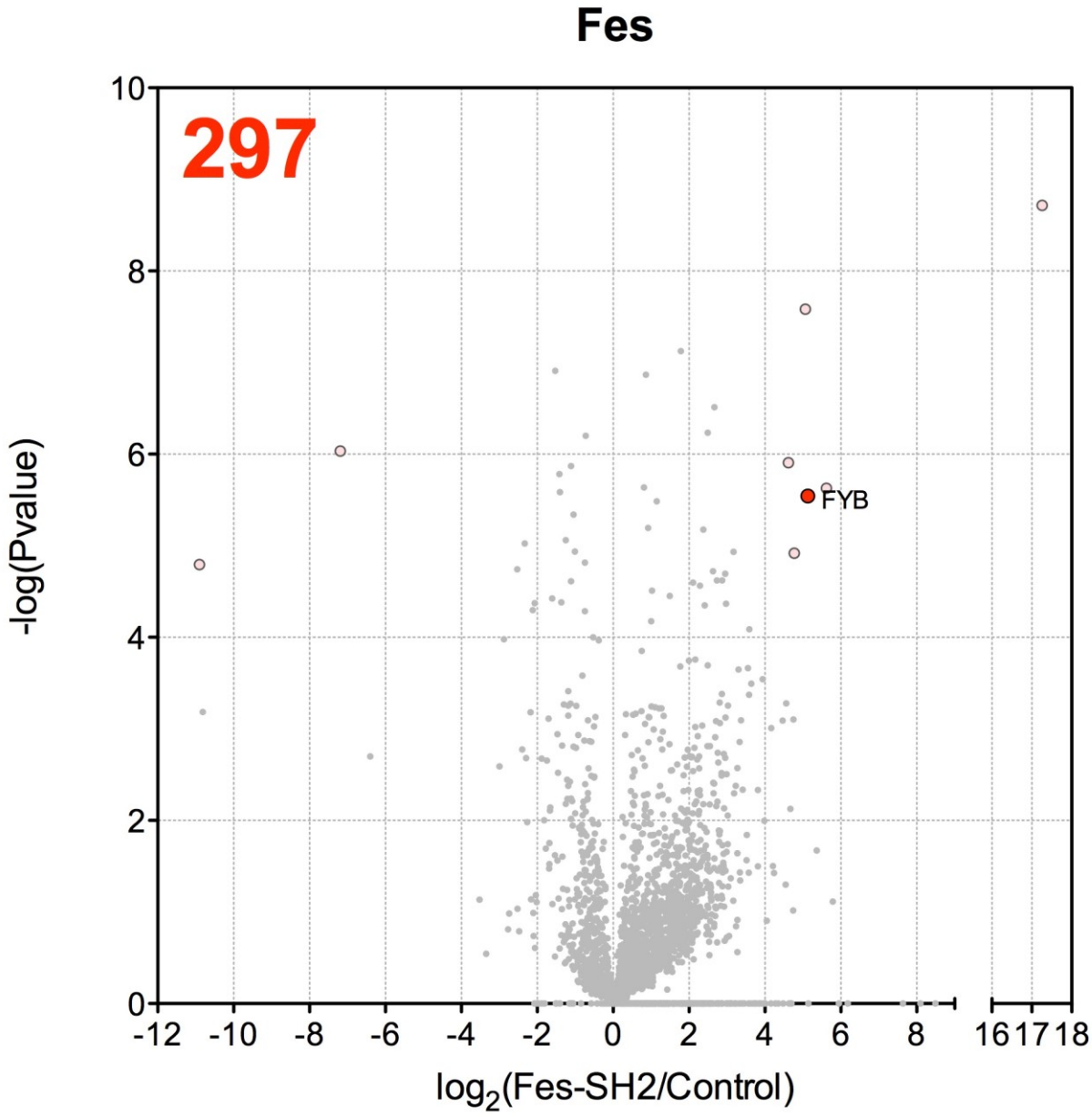
The SH2 domain panel was chosen to represent a diverse set of tyrosine kinase members that are involved in several different biological processes. While Abl and Src are known to be involved in EGF signaling, Lyn and Csk are more commonly associated with immune signaling, more specifically with T cell receptor activation. As such, we wanted to profile the domains in more than one biological context. To this aim, the panel of seven SH2 domains were used as affinity purification reagents in 1.5 mg of 15 min CD3/CD28 costimulated Jurkat cell lysate. For each SH2 domain, three technical replicate pulldowns were performed in two separate biological replicates. The experimental workflow for the pulldown, elution, and proteomic analysis remained the same as it was in HeLa. Significantly enriched proteins were similarly determined by *t*-test for the given SH2 domain against the phosphotyrosine dead control. Each list was populated with those proteins that made the *t*-test cutoff for at least one set of biological replicates.

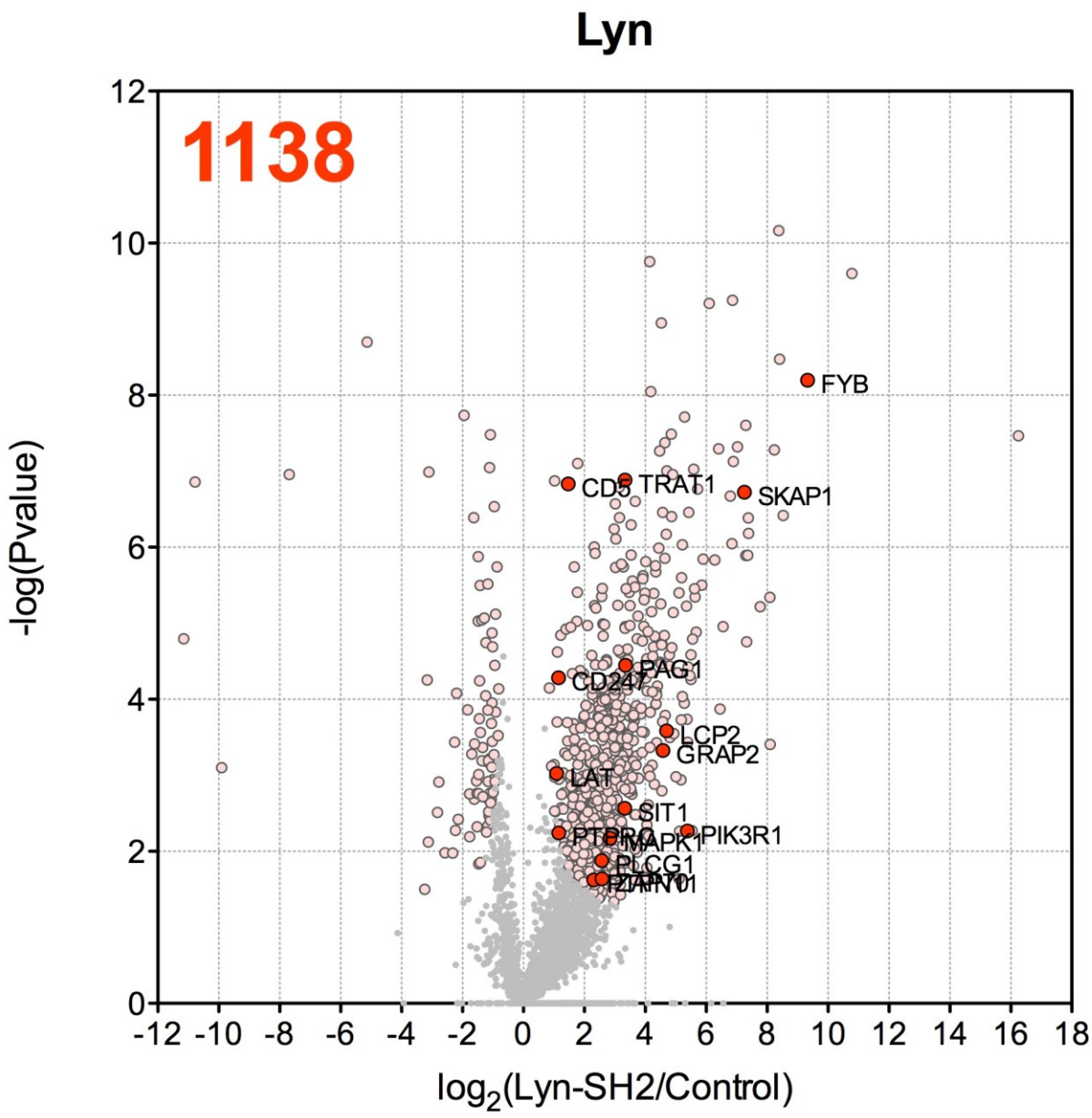

Abl
329
TRAT1
FYB
PAG1
-log(Pvalue)
log2(Abl-SH2/Control)
10
8
6
4
2
0
-12
-10
-8
-6
-4
-2
0
2
4
6
8


Brk
1470
-log(Pvalue)
log2(Brk-SH2/Control)
FYB
SKAP1
PAG1
CD28
LCP2
CD5
GRAP2
PSME3
TRAT1
VAV1
PIK3R1


Btk
945
LCP2
-log(Pvalue)
log2(Btk-SH2/Control)
10
8
6
4
2
0
-12
-10
-8
-6
-4
-2
0
2
4
6
8
10
12
14
16


Csk
333
CD247
PAG1
-log(Pvalue)
log2(Csk-SH2/Control)
10
8
6
4
2
0
-12
-10
-8
-6
-4
-2
0
2
4
6
8
10
14
15
16


Fes
297
FYB
-log(Pvalue)
log2(Fes-SH2/Control)
10
8
6
4
2
0
-12
-10
-8
-6
-4
-2
0
2
4
6
8
16
17
18


Lyn
1138
FYB
CD5
TRAT1
SKAP1
PAG1
CD247
LCP2
GRAP2
LAT
SIT1
PIK3R1
MAPK1
PLCG1
-log(Pvalue)
log2(Lyn-SH2/Control)

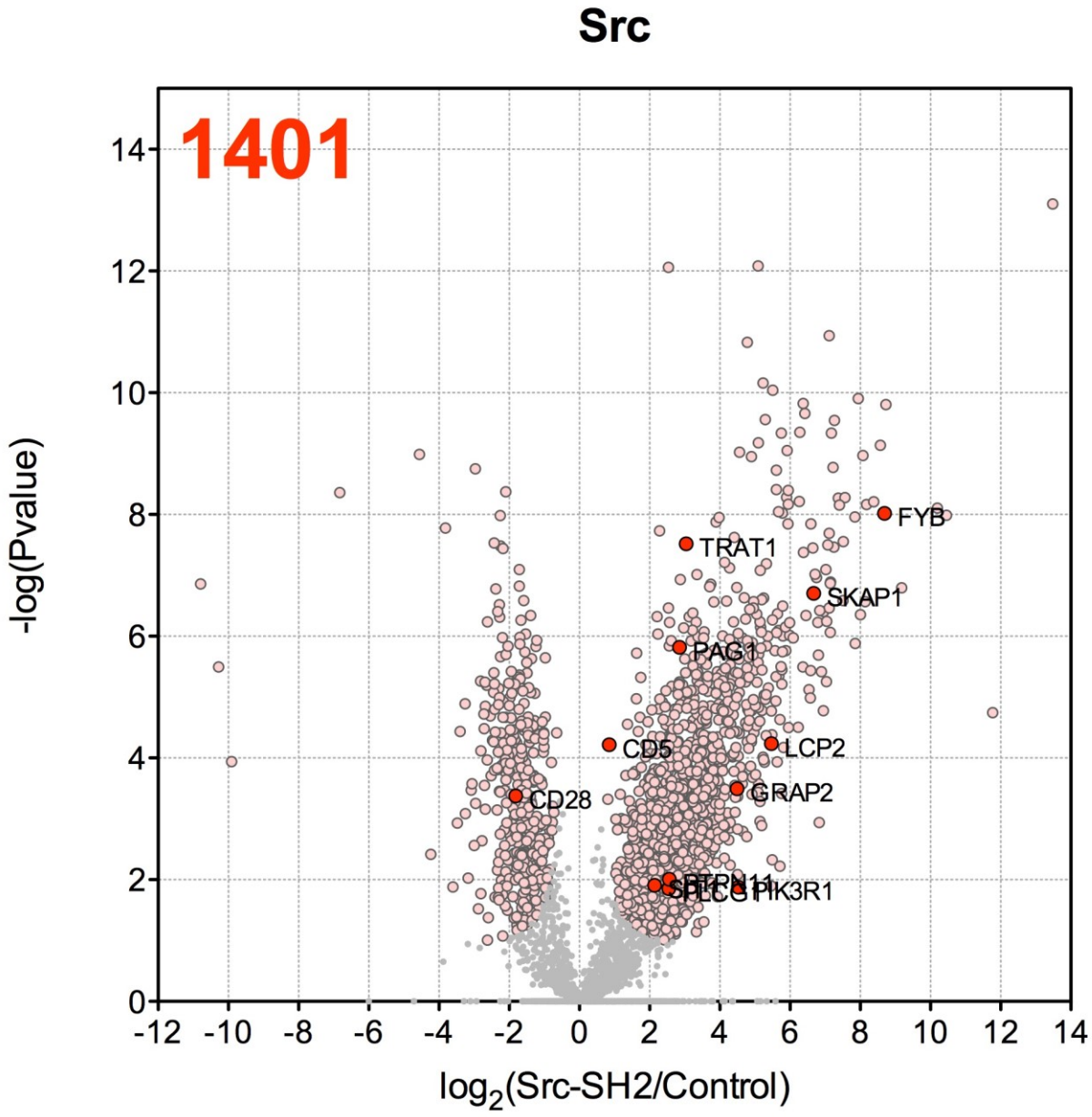


Figure 2-15. Volcano plots of proteins enriched from T cell activated Jurkat cell lysate for each of the seven SH2 domains in the panel. Significantly enriched targets are listed in pink and those that are designated as part of the GO-BP T cell receptor signaling pathway are noted in red.

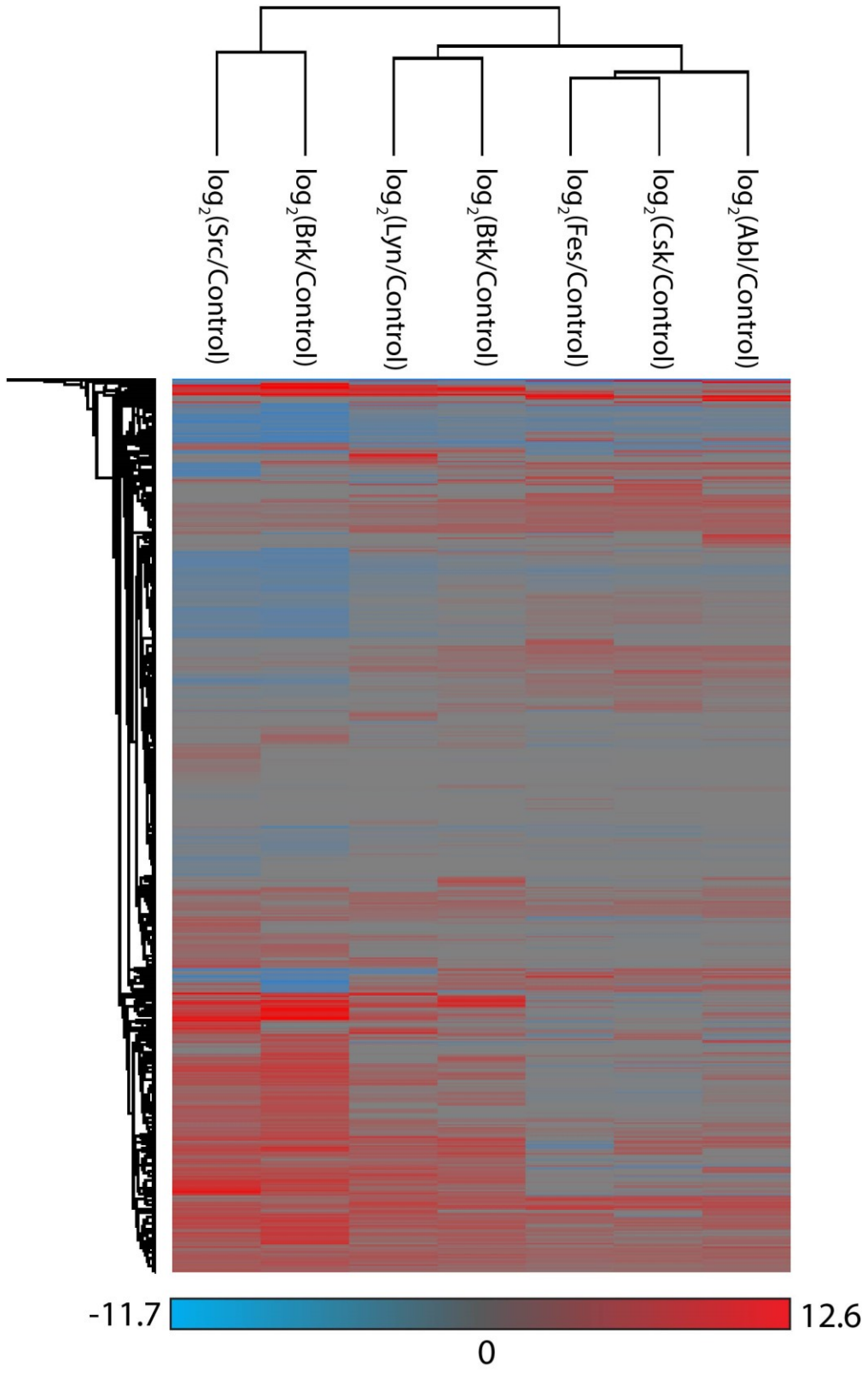


Figure 2-16. Hierarchical clustering analysis of the log2 ratio between enrichment for each SH2 domain over the control in T cell activated Jurkat cell lysate.

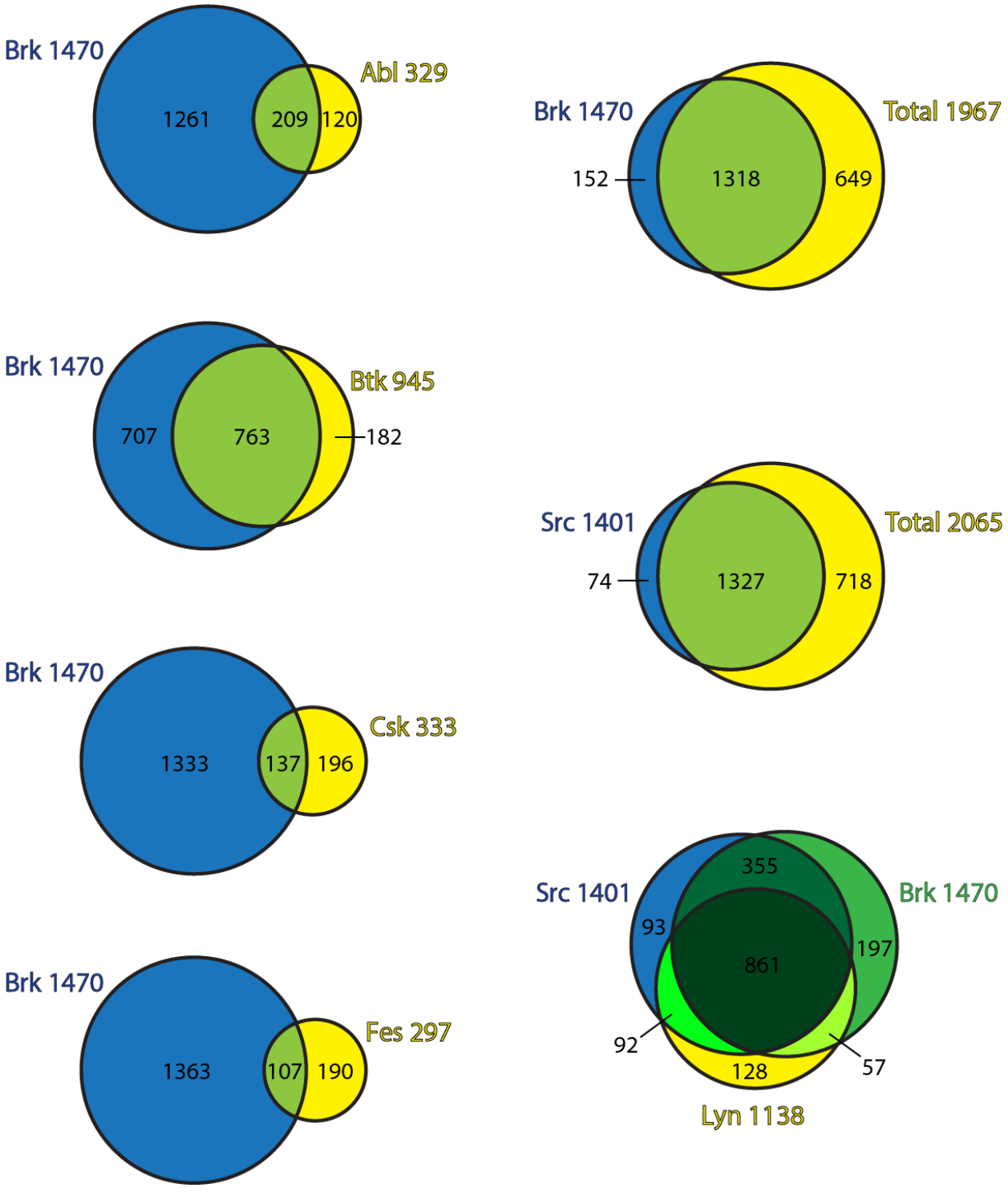


Figure 2-17. Venn diagram analysis displaying the enriched hits for Src and Brk, which had the highest enrichment, with the rest of the panel in T cell activated Jurkat cell lysate.

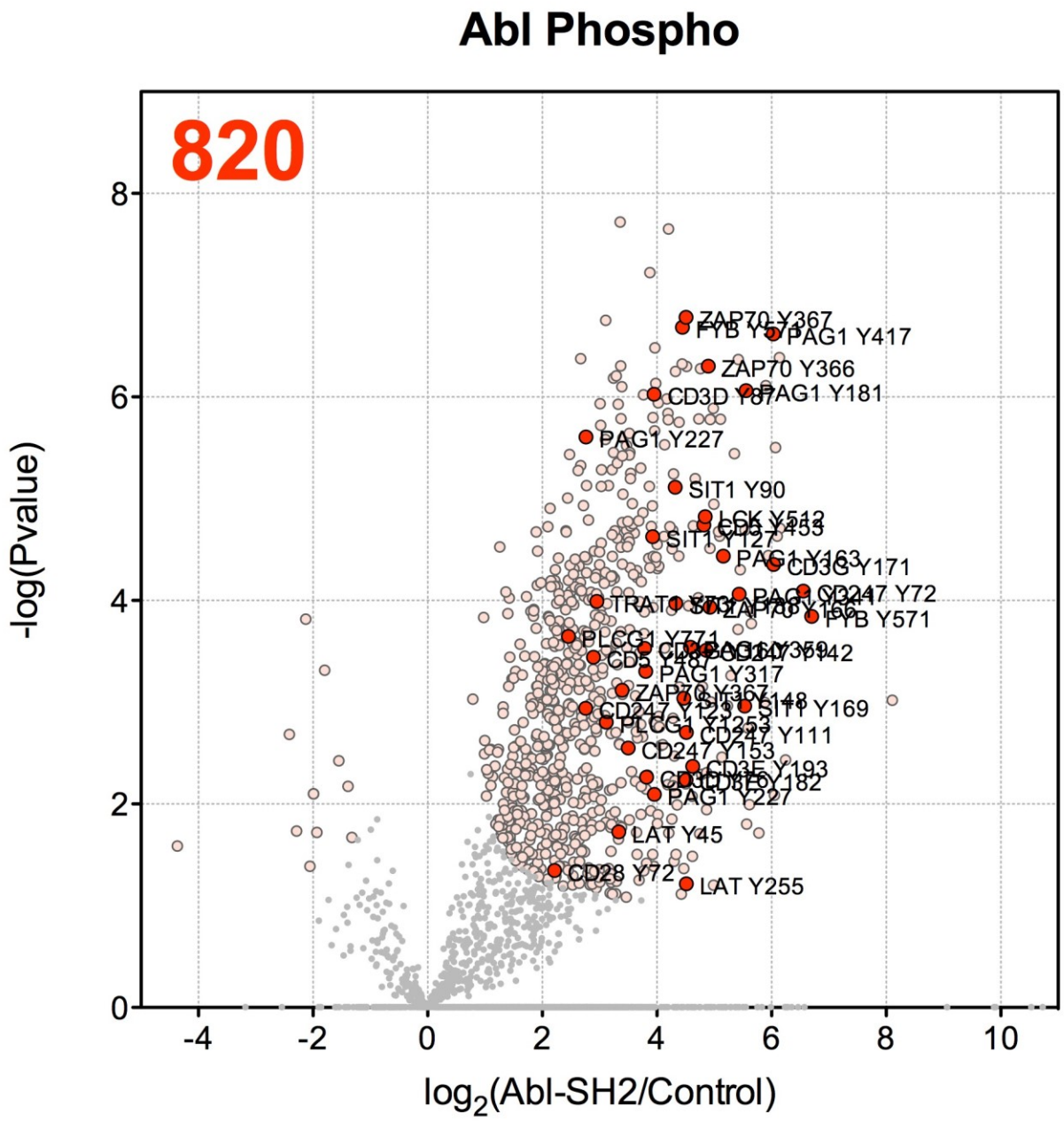
The relative enrichment profiles for the SH2 domains in the T cell activated Jurkat cell lysate were quite similar to the trends observed in EGF-stimulated HeLa cells. Once again, the

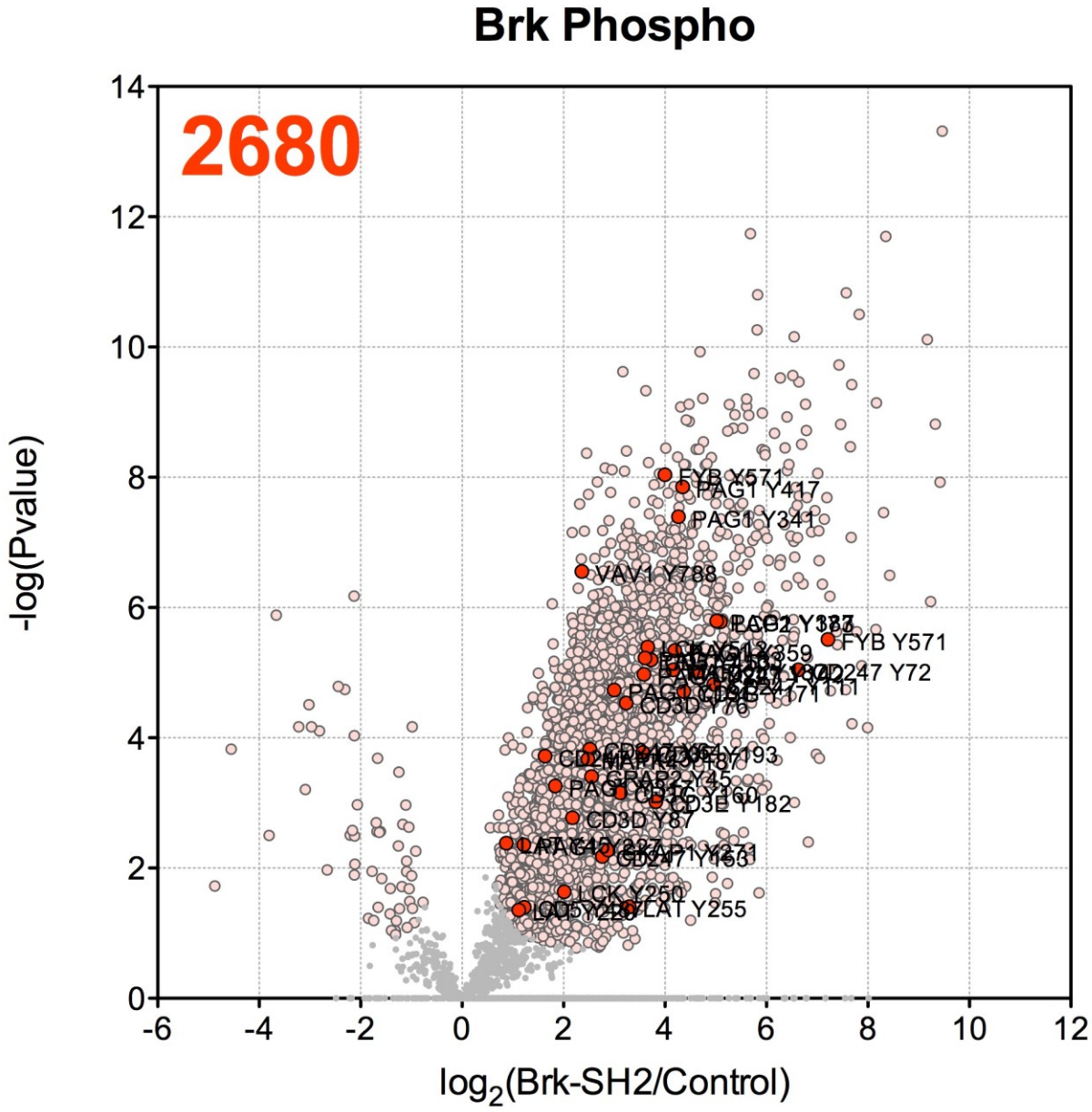
SH2 domains of Brk and Src had the most robust list of identified proteins over the control while Csk, Fes, and Abl SH2 displayed the most restrictive enrichment (Figure 2-15 and 2-16). The SH2 domain of Abl had a much more limited enrichment than expected when comparing its results from the initial Western blot studies (compare Figure 2-1 and Figure 2-15). Additionally, the SH2 domain of Btk exhibited very specific and low amounts of enriched phosphotyrosine-containing species in the Western blot studies, but had the fourth largest number of significantly identified protein hits of the panel. Csk SH2 increased the number of proteins pulled down in T cell activated Jurkat cell lysate, 333, when compared to 30 proteins in EGF-stimulated HeLa cell lysate. This was a positive and welcome result as Csk kinase has a role in T cell receptor regulation, and thus we expected to see a larger number of enriched proteins in this system.

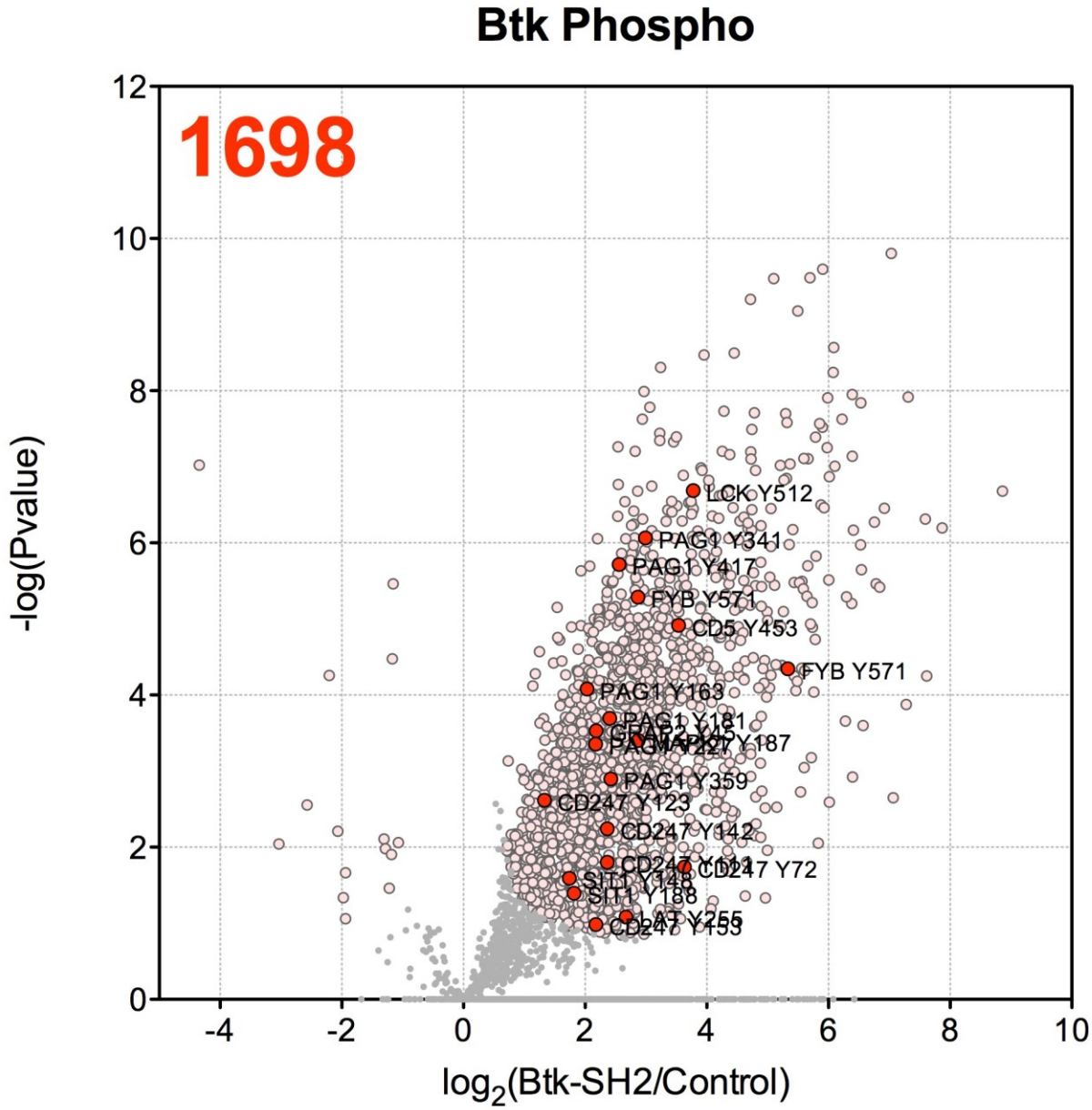
Hierarchical clustering of the log2 ratio in enrichment between SH2 domain versus control (Figure 2-16) revealed the same relationship between SH2 domains as that observed for HeLa (compare to Figure 2-7 and 2-11). This adds support to our prior proteomic results in EGF-stimulated HeLa cell lysate, which displayed a different SH2 domain relationship than observed by Machida and coworkers in the Far-Western method.

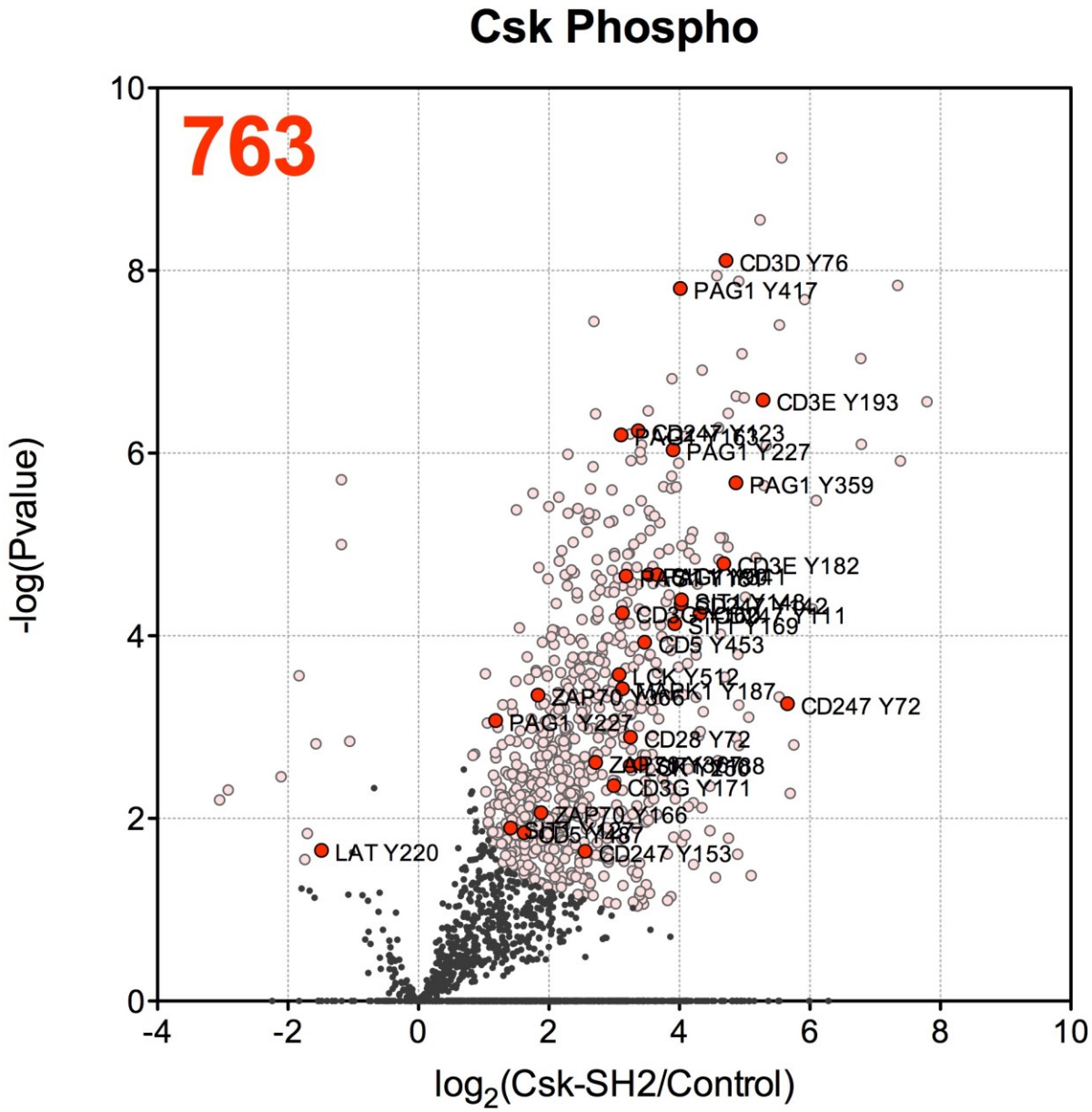
Phosphopeptide Analysis of T Cell Activated Jurkat Interactors

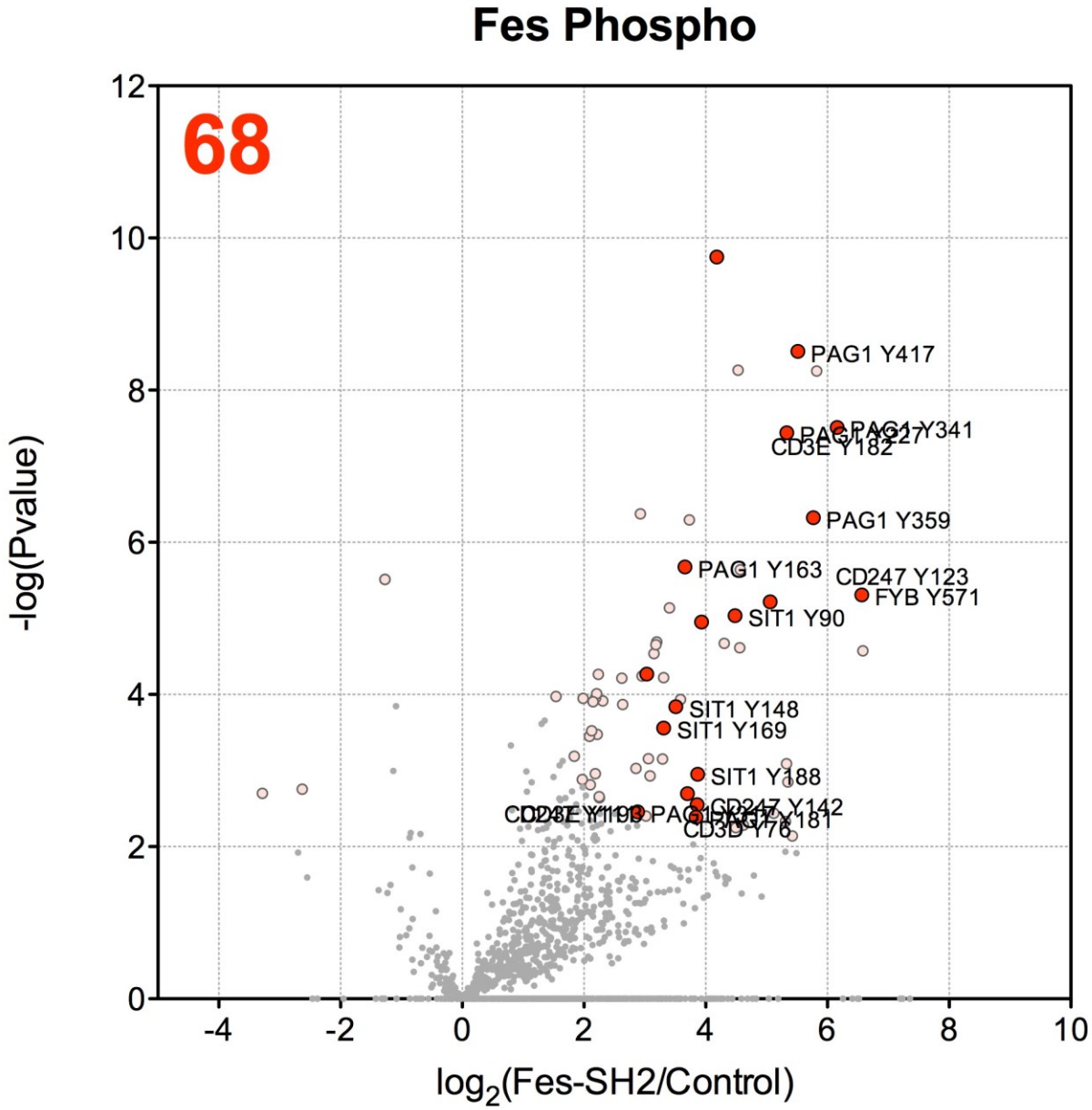
The remaining 90% of trypsinized peptide from each pulldown were enriched for phosphopeptides using IMAC and then analyzed with LC/MS. Significantly enriched peptides were determined as usual. Phosphosite results mirrored those of protein in terms of level of enrichment (compare Figure 2-15 and 2-18). For all members of the SH2 panel, the number of phosphotyrosine-containing peptides was less than observed when pulling down from EGF-stimulated HeLa cell lysate (Table 2-3). Again, this confirms the initial Western blot profiling that showed there is less enriched phosphotyrosine-containing proteins from a Jurkat lysate pulldown (Figure 2-3). However, the number of phosphotyrosine pulled down also shows a limitation in the Western blot profiling. For example, the SH2 domain of Btk had a low density of phosphotyrosine-containing proteins in the Western blot profiling, but relatively high enrichment of 945 proteins and 1698 phosphopeptides as determined by MS. Btk SH2 simply pulls down a larger number of proteins and phosphosites from a smaller pool of phosphotyrosine-containing sites.

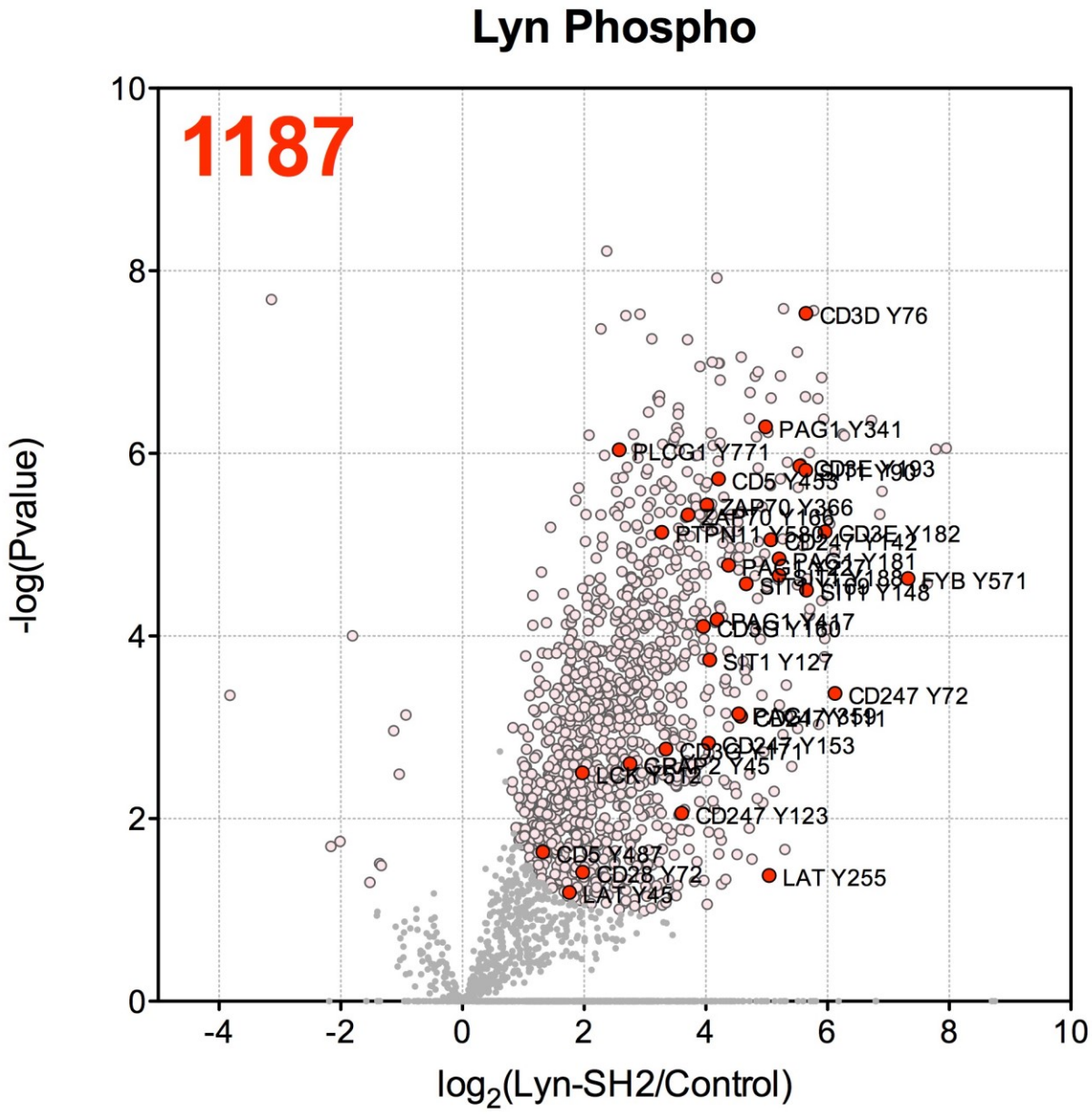

Abl Phospho
820
-log(Pvalue)
log2(Abl-SH2/Control)
ZAP70 Y367
PAG1 Y417
ZAP70 Y366
PAG1 Y181
CD3D Y87
PAG1 Y227
SIT1 Y90
SIT1 Y127
PAG1 Y163
CD3G Y171
FYB Y571
PAG1 Y317
SIT1 Y169
CD247 Y111
CD247 Y153
PAG1 Y227
LAT Y45
CD28 Y72
LAT Y255


Brk Phospho
2680
-log(Pvalue)
log2(Brk-SH2/Control)
FYB Y571
PAG1 Y417
PAG1 Y341
VAV1 Y788
FYB Y571
CD3D Y87
LCK Y250
LAT Y255


Btk Phospho
1698
-log(Pvalue)
log2(Btk-SH2/Control)
LCK Y512
PAG1 Y341
PAG1 Y417
FYB Y571
CD5 Y453
FYB Y571
PAG1 Y163
PAG1 Y181
PAG1 Y359
CD247 Y123
CD247 Y142
CD247 Y72
SIT1 Y188
CD247 Y153


Csk Phospho
763
CD3D Y76
PAG1 Y417
CD3E Y193
PAG1 Y227
PAG1 Y359
CD3E Y182
CD5 Y453
CD247 Y72
PAG1 Y227
CD28 Y72
CD3G Y171
ZAP70 Y166
LAT Y220
CD247 Y153
-log(Pvalue)
log2(Csk-SH2/Control)
10
8
6
4
2
0
-4
-2
0
2
4
6
8
10


Fes Phospho
68
PAG1 Y417
PAG1 Y341
CD3E Y182
PAG1 Y359
PAG1 Y163
CD247 Y123
FYB Y571
SIT1 Y90
SIT1 Y148
SIT1 Y169
SIT1 Y188
CD247 Y142
CD3D Y76
-log(Pvalue)
log2(Fes-SH2/Control)


Lyn Phospho
1187
-log(Pvalue)
log2(Lyn-SH2/Control)
CD3D Y76
PAG1 Y341
PLCG1 Y771
CD5 Y455
ZAP70 Y366
PTPN11 Y580
CD3E Y182
PAG1 Y181
FYB Y571
SIT1 Y148
PAG1 Y417
SIT1 Y127
CD247 Y72
CD247 Y153
CD247 Y123
CD5 Y487
CD28 Y72
LAT Y255
LAT Y45

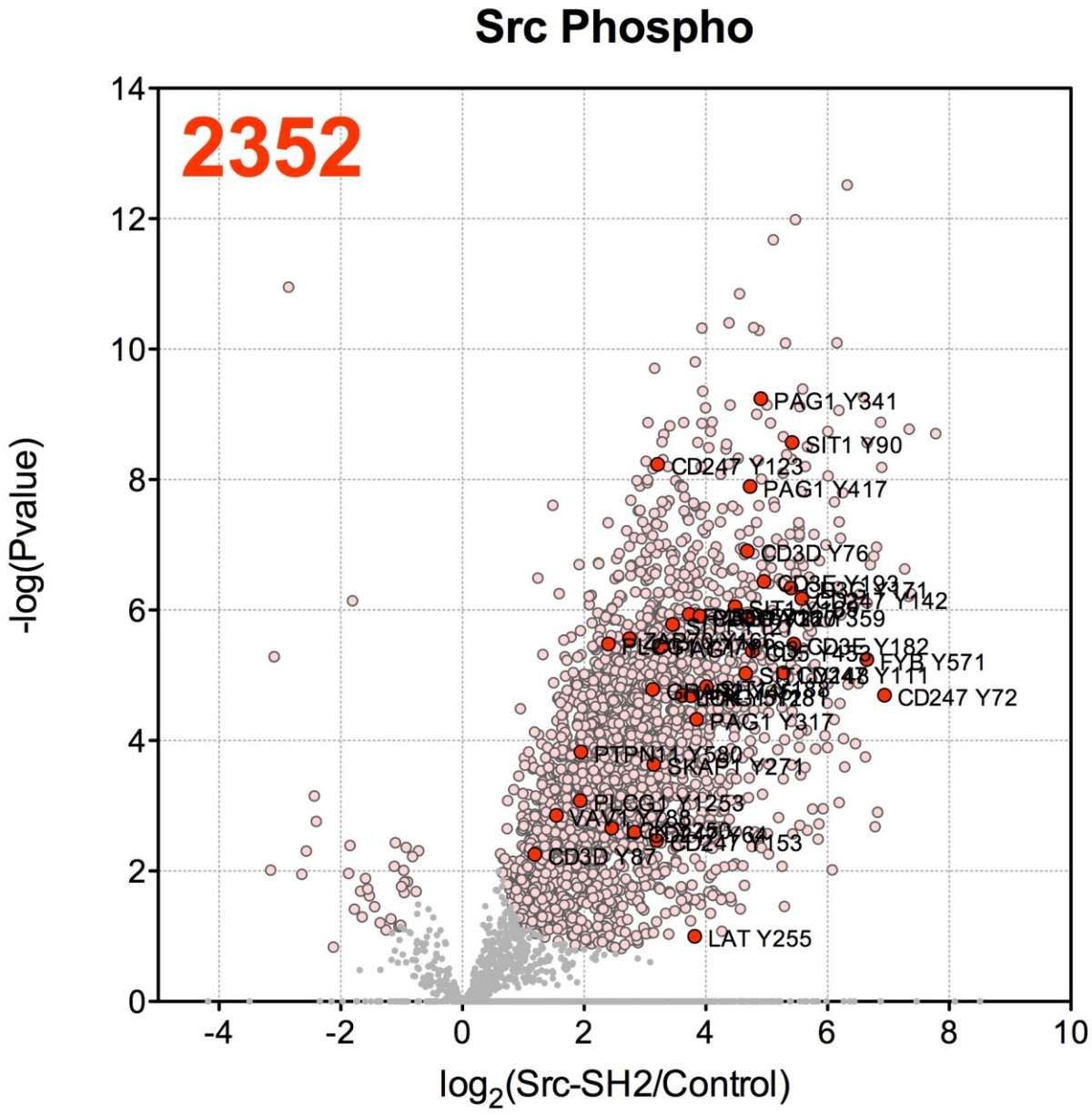


Figure 2-18. Volcano plots displaying the significantly enriched phosphosites in T cell activated Jurkat cell lysate for each SH2 domain. Phosphosites that were significantly enriched are denoted in pink and those phosphopeptides involved in T cell receptor signaling by GO analysis are in red.

	ABL	BRK	BTK	CSK	FES	LYN	SRC
Enriched Sites	820	2680	1698	763	68	1187	2352
pY	65	86	43	49	21	64	78
Shared Sites	773	2086	1572	699	66	1095	2034
Unique Sites	47	594	126	64	2	92	318

Table 2-4. Summary of enriched phosphosites for the panel of SH2 domains against T cell activated Jurkat cell lysate. The total number of sites enriched by an SH2 domain in one or more biological replicates and the number of those that are phosphotyrosine are listed. The number of sites that are shared by the other SH2 domains in the panel and those unique for a given SH2 domain are also provided.

Phosphosite	log_2(Abl/ Control) Phospho	log_2(Brk/ Control) Phospho	log_2(Btk/ Control) Phospho	log_2(Csk/ Control) Phospho	log_2(Fes/ Control) Phospho	log_2(Lyn/ Control) Phospho	log_2(Src/ Control) Phospho
ALDH7A1 Y452___3	0.00	2.91	1.90	0.00	0.00	2.14	2.80
ALDH7A1 Y462___3	0.00	2.83	1.82	0.00	0.00	2.07	2.73
ARAP2 Y77___1	0.00	0.00	0.00	4.28	0.00	3.47	2.40
ARHGAP27 Y28___1	2.24	1.77	0.00	2.72	0.00	4.09	3.23
BCLAF1 Y381___2	0.00	1.83	0.00	0.00	0.00	0.00	0.00
BCLAF1 Y282___2	NaN	3.51	3.55	-0.04	NaN	NaN	2.94
BCLAF1 Y282___1	0.00	2.90	3.43	0.00	0.00	0.00	4.59
BCLAF1 Y381___1	0.00	3.32	0.00	0.00	0.00	0.00	2.70
BUD13 Y320___1	0.00	0.00	0.00	0.00	0.00	4.23	0.00
C9orf78 Y147___1	0.00	2.20	0.00	0.00	0.00	0.00	1.25
CBL Y674___1	2.55	1.03	0.00	0.00	0.00	0.00	1.71
CBL Y700___1	3.23	1.56	0.00	0.79	0.00	2.20	1.36
CBL;CBLB Y371___1	4.12	2.54	0.00	0.00	0.00	2.80	3.22
CCDC132 Y344___1	3.65	4.71	4.02	3.13	0.00	0.00	1.65
CCDC132 Y344___2	0.00	4.43	1.85	0.00	0.00	0.00	0.00
CD226 Y167___1	0.00	0.00	0.00	3.98	0.00	0.00	0.00
CD247 Y123___1	2.76	1.63	1.33	3.37	3.03	3.61	3.21
CD247 Y153___1	3.50	2.76	2.18	2.55	2.70	4.05	3.19
CD247 Y111___1	4.51	4.96	2.36	4.32	3.70	4.57	5.27
CD247 Y142___1	4.86	4.64	2.37	4.03	3.86	5.07	5.58
CD247 Y72___1	6.55	6.62	3.63	5.65	3.71	6.12	6.94
CD247 Y64___1	0.00	2.52	0.00	0.00	0.00	2.28	2.83
CD28 Y72___1	2.21	NaN	0.93	3.25	1.90	1.98	1.24
CD3D Y76___1	3.82	3.23	0.86	4.72	4.18	5.65	4.68
CD3D Y87___1	3.95	2.17	0.00	2.49	0.00	0.00	1.19
CD3E Y182___1	4.51	3.81	0.87	4.68	5.06	5.96	5.45
CD3E Y193___1	4.62	3.56	0.62	5.28	3.93	5.55	4.96
CD3G Y160___1	3.79	3.12	0.00	3.13	2.26	3.96	3.88
CD3G Y171___1	6.04	4.37	1.20	2.99	2.38	3.35	5.40
CD5 Y487___1	2.89	1.22	0.84	1.62	0.60	1.32	0.67
CD5 Y453___1	4.82	3.72	3.54	3.47	1.91	4.21	4.76
CD84 Y185___1	2.02	0.00	0.00	0.00	0.00	0.98	1.52
CD84 Y165___1	3.22	0.94	0.00	0.00	0.00	3.24	2.53
CDK1;CDK2;CDK3;CDC2 Y15___1	0.68	0.88	0.91	5.53	0.26	0.48	0.81
CDK1;CDK2;CDK3;CDC2 Y15___2	3.71	2.83	2.69	6.10	1.34	0.60	1.06
CDK5 Y15___1	1.14	1.07	0.78	3.88	0.00	0.52	1.08
CRACR2A Y464___3	0.00	3.73	3.51	0.00	0.00	0.00	0.00
DAPP1 Y139___1	0.00	1.30	0.00	0.00	0.00	0.00	1.59
DDX42 Y183___1	0.00	3.42	0.00	0.00	0.00	0.00	3.55
DOK1 Y157___1	3.61	1.25	2.83	2.69	0.00	2.96	4.32
DOK1 Y176___1	5.45	3.43	5.67	5.33	3.58	5.67	3.94
DOK1 Y310___1	0.48	0.82	3.02	1.71	0.39	1.59	2.29

DOK1 Y270___1	0.00	0.00	4.56	0.00	0.00	0.00	3.85
DOK2 Y299___1	4.02	3.71	3.75	3.42	2.31	4.29	4.01
DYRK1A;DYRK1B Y312___1	-0.67	1.37	1.09	0.37	-0.15	1.04	1.96
ETS1 Y283___2	0.00	1.31	0.00	0.00	0.00	0.00	0.00
FYB Y571___2	4.44	3.99	2.87	0.00	0.00	1.97	3.73
FYB Y571___1	6.70	7.21	5.33	NaN	6.56	7.33	6.65
GCSAM Y133___1	3.37	0.84	0.00	0.00	0.00	0.00	1.19
GRAP2 Y45___1	0.00	2.55	2.18	0.00	0.00	2.76	3.13
GSK3B;GSK3A Y216___1	2.11	3.25	3.61	4.59	3.97	3.48	2.55
HIPK1;HIPK2 Y352___1	1.12	0.14	1.41	0.26	0.08	1.52	1.43
HIPK3 Y359___1	1.37	0.00	1.31	0.00	0.00	2.08	1.70
HNRNPA3 Y364___1	0.00	5.69	0.00	3.63	0.00	2.74	4.62
HNRNPA3 Y360___1	NaN	NaN	4.91	NaN	2.39	NaN	NaN
ITK Y512___1	3.35	3.16	2.51	3.91	1.24	2.55	3.45
ITSN2 Y951___1	4.29	1.85	0.00	5.92	0.00	4.21	3.30
KIF1B;KIF1Bbeta Y111___3	2.86	2.43	2.81	2.12	1.50	2.61	2.44
KIRREL Y535___1	2.04	1.10	0.00	0.89	1.84	0.00	0.39
KIRREL Y538___1	3.29	0.93	0.00	0.00	1.79	1.39	1.25
LAT Y220___1	1.50	1.11	-0.08	-1.48	0.89	1.08	0.96
LAT Y45___1	3.33	1.27	0.70	NaN	2.00	1.76	0.97
LAT Y255___1	4.51	3.29	2.67	NaN	4.92	5.04	3.82
LAT Y45___2	0.00	0.87	0.00	0.00	0.00	0.00	0.95
LAX1 Y218___1	1.79	0.25	0.00	0.00	0.00	1.86	0.83
LCK Y512___1	4.84	3.65	3.78	3.08	0.00	1.97	3.62
LCK Y250___1	4.36	2.01	4.41	3.25	0.00	0.00	2.46
LCK;YES1;FYN;SRC Y419___1	-1.35	-0.32	0.74	1.06	0.24	-1.06	0.21
LCP2 Y173___1	0.00	5.09	4.23	0.00	0.00	4.90	4.17
LMO7 Y133___1	5.42	1.64	3.12	0.00	0.00	3.11	3.02
MAPK1 Y187___1	3.04	4.16	2.87	3.13	0.00	0.65	2.92
MAPK1 Y187___2	0.00	2.48	0.00	0.00	0.00	0.00	0.00
MAPK14 Y105___1	-0.58	-0.57	-0.41	0.00	0.00	1.30	-0.74
MAPK3 Y90___1	0.00	2.30	0.00	0.00	0.00	0.00	0.00
MAX Y87___1	0.00	2.87	0.00	0.00	0.00	2.10	1.40
MDGA1 Y58___3	0.00	4.49	2.55	0.00	3.35	0.00	0.00
MDGA1 Y69___3	0.00	4.95	3.01	0.00	3.81	0.00	0.00
NEK1 Y1081___1	0.00	0.46	1.84	0.00	0.00	0.00	0.49
PAG1 Y227___2	2.76	1.21	1.16	1.18	0.00	0.00	0.00
PAG1 Y317___1	3.81	1.83	0.00	3.10	2.88	4.34	3.85
PAG1 Y227___1	3.96	3.58	2.17	3.90	5.33	4.37	3.90
PAG1 Y359___1	4.59	4.19	2.42	4.87	5.77	4.54	4.69
PAG1 Y163___1	5.16	3.60	2.03	3.11	3.66	0.00	3.27
PAG1 Y341___1	5.43	4.26	2.99	3.53	6.16	4.98	4.91
PAG1 Y181___1	5.56	2.99	2.40	3.18	3.84	5.21	3.76
PAG1 Y417___1	6.03	4.34	2.56	4.01	5.51	4.19	4.73
PAG1 Y387___1	4.36	5.01	0.00	0.00	7.20	0.00	5.75
PECAM1 Y713___1	1.40	0.00	0.00	1.50	0.00	0.19	-0.76
PGRMC2 Y210___2	0.00	2.85	4.47	0.00	0.00	0.00	0.00

PGRMC2 Y210___1	9.06	-0.34	0.14	2.81	2.40	1.93	4.11
PKP4 Y372___1	2.05	-0.33	NaN	1.70	NaN	NaN	0.89
PLCG1 Y771___1	2.45	0.00	0.00	0.00	0.00	2.58	2.40
PLCG1 Y1253___1	3.12	0.00	0.00	0.00	0.00	0.00	1.94
PRPF4B Y849___1	1.40	5.21	3.59	1.89	NaN	NaN	4.36
PTK2B Y326___1	0.00	0.00	0.00	0.72	0.00	2.27	0.78
PTK2B Y325___1	-0.21	0.00	0.00	0.23	0.00	1.77	0.38
PTPN11 Y580___1	1.02	0.00	0.00	0.00	0.00	3.28	1.95
PTPRA Y789___1	2.31	1.21	1.38	1.33	0.00	3.11	3.55
RBM15 Y68___1	0.00	0.00	0.00	0.00	0.00	0.00	1.36
RBM8A Y53___1	0.00	1.40	0.00	0.00	0.00	0.00	0.00
RFC1 Y67___2	0.00	2.51	2.94	0.00	0.00	0.00	2.02
SH2D2A Y188___1	0.00	1.55	0.00	0.00	0.00	0.00	1.63
SHISA9 Y196___3	0.00	3.00	2.67	0.00	0.00	0.00	0.00
SIT1 Y127___1	3.92	2.60	0.00	1.41	0.00	4.06	3.45
SIT1 Y90___1	4.32	3.36	2.85	3.66	4.48	5.64	5.42
SIT1 Y188___1	4.33	2.88	1.82	3.41	3.87	5.20	4.01
SIT1 Y148___1	4.47	3.64	1.74	4.03	3.51	5.65	4.66
SIT1 Y169___1	5.53	3.36	1.23	3.93	3.31	4.67	4.48
SKAP1 Y271___1	0.00	2.87	0.00	0.00	0.00	3.43	3.15
SLAMF6 Y198___1	0.00	0.00	0.00	2.18	0.00	2.83	3.13
SMARCAD1 Y217___2	0.00	1.38	0.00	0.00	0.00	0.00	0.00
SMARCAD1 Y91___2	0.00	3.55	0.00	0.00	0.00	0.00	0.00
THRAP3 Y293___1	0.00	5.25	0.00	0.00	0.00	4.96	0.00
TMEM192 Y209___1	0.00	0.00	0.00	0.00	0.00	2.69	0.00
TNK2 Y394___1	2.51	0.00	0.00	1.40	0.00	1.13	-0.15
TRAT1 Y73___1	2.95	0.00	0.00	0.00	0.00	0.00	1.64
TRIM4 Y355___2	4.46	0.87	3.47	2.47	2.85	NaN	3.92
TYK2 Y292___1	3.77	2.48	3.36	3.64	2.85	2.91	2.36
UBASH3A Y9___1	3.99	3.15	0.48	4.91	-0.18	4.86	4.34
UQCRB Y7___1	3.27	0.00	0.00	0.00	0.00	0.00	0.00
VAV1 Y788___1	0.00	2.36	0.00	0.00	0.00	0.00	1.54
ZAP70 Y367___1	3.39	0.00	0.00	2.72	0.00	0.00	0.00
ZAP70 Y367___2	4.51	0.22	1.05	1.76	0.00	0.00	0.00
ZAP70 Y366___2	4.90	1.13	1.66	1.83	0.00	0.00	0.00
ZAP70 Y166___1	4.93	2.20	0.74	1.88	0.00	3.71	2.75
ZAP70 Y366___1	0.00	0.07	0.00	1.19	0.68	4.02	0.00

Table 2-5. Phosphotyrosine sites enriched by at least one member of the SH2 domain panel in T cell activated Jurkat cell lysate. Relative enrichment for the site is heatmapped as the log2 ratio of enrichment for each SH2 domain over the phosphotyrosine dead control.

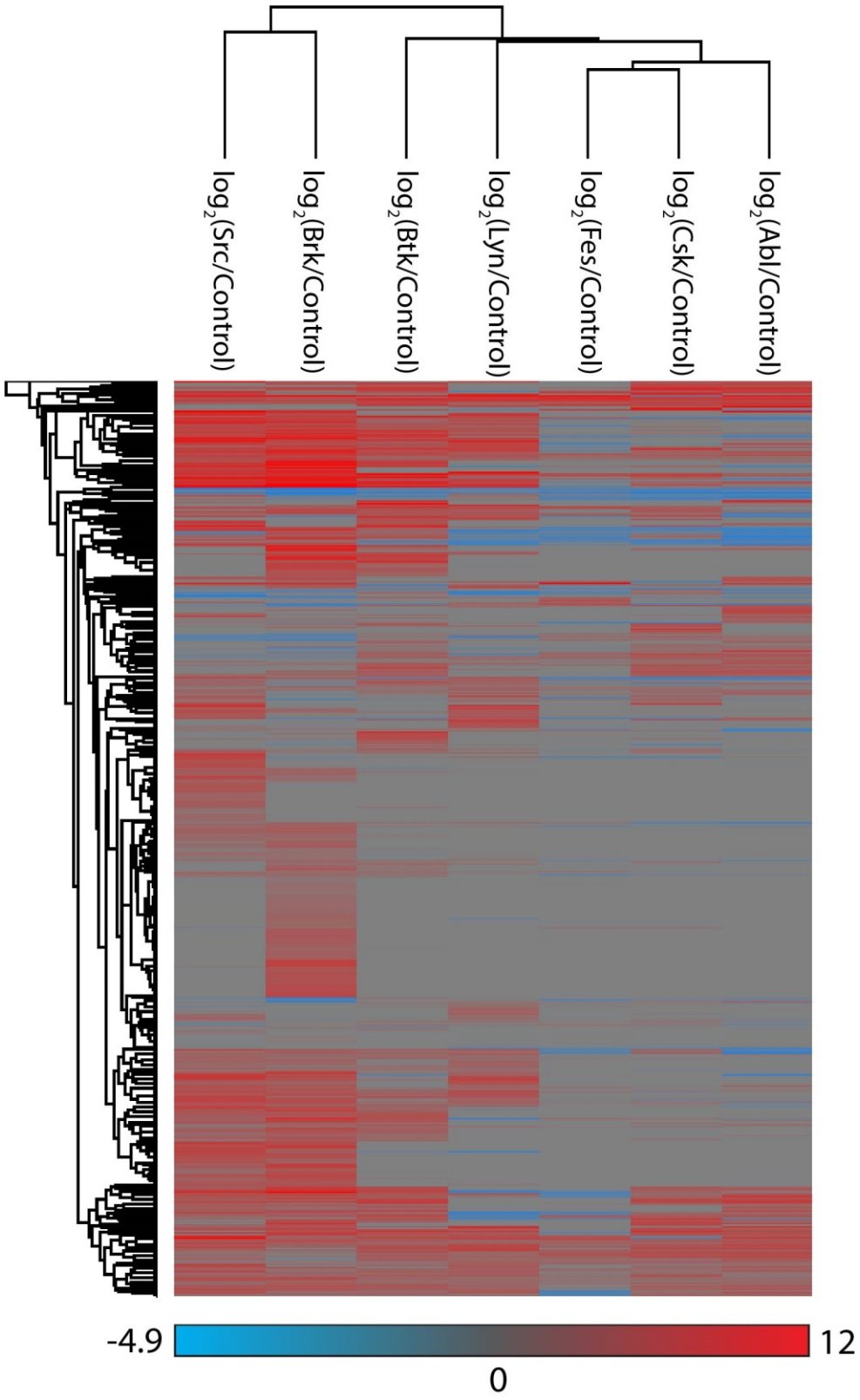


Figure 2-19. Hierarchical clustering analysis on the log2 ratio in phosphopeptide enrichment for each SH2 domain over the control in T cell activated Jurkat cell lysate.

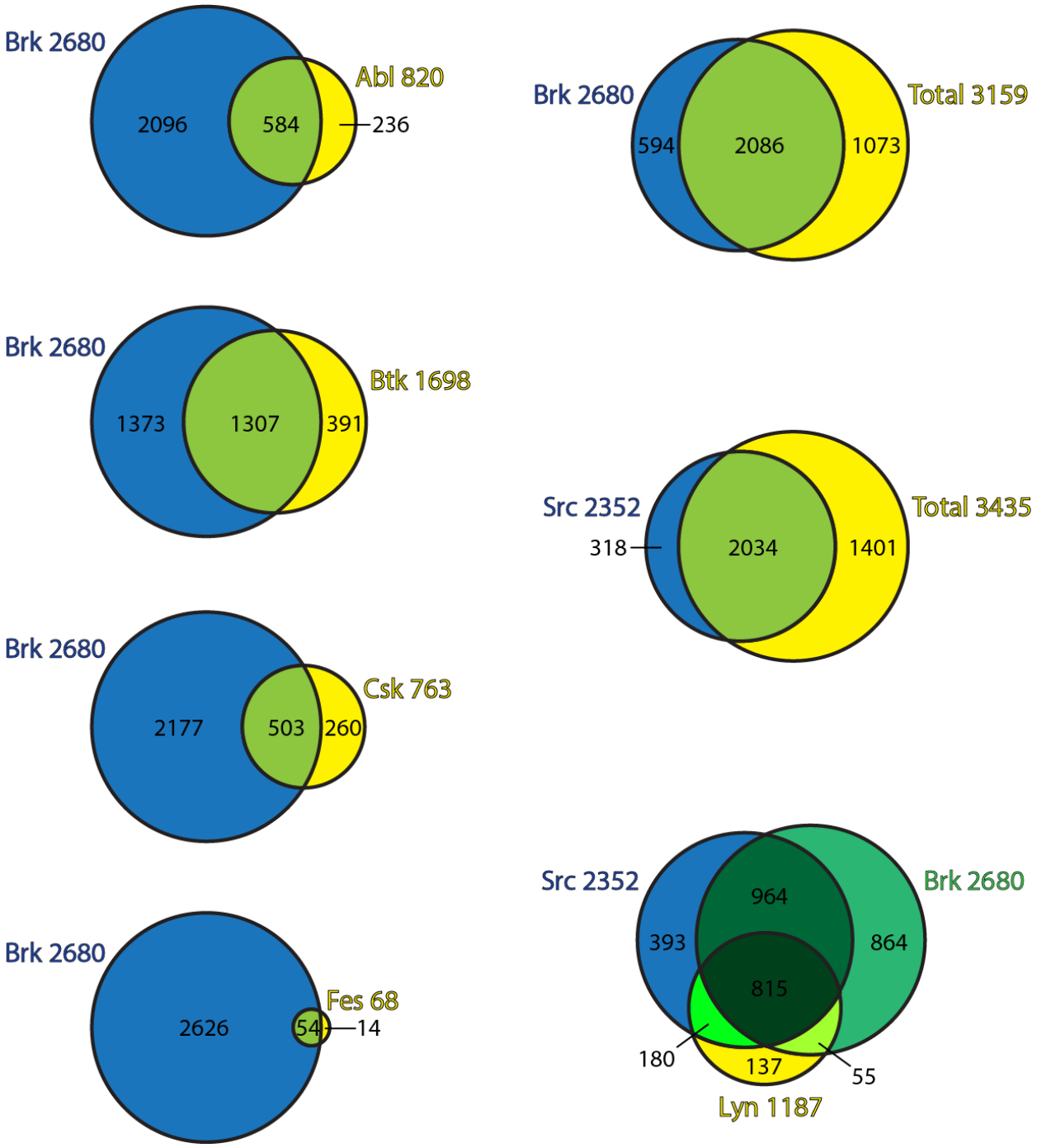


Figure 2-20. Venn diagram analysis comparing the phosphosites enriched from the SH2 domain panel in T cell activated Jurkat cell lysate. Brk and Src are highlighted in particular because they showed the highest number of significantly identified phosphopeptides.

The SH2 domains of Brk and Src had large numbers of identified phosphosites, 2680 and 2352 phosphopeptides, respectively, and by themselves captured nearly two-thirds of the remaining pool of phosphopeptides from the other SH2 domains (Figure 2-20, Table 2-4). Btk SH2 enriched 1698 phosphoeptides and Lyn SH2 enriched 1187, which were similar levels of overall phosphosite enrichment in T cell activated Jurkat cell lysate as compared to EGF-stimulated HeLa cell lysate. The SH2 domains of Abl and Csk had increased numbers of significant hits when moving to T cell activated Jurkat cell lysate. Abl enriched 820 phosphopeptides and Csk SH2 enriched 763 phosphopeptides compared to 529 and 395 phosphopeptides in EGF-stimulated HeLa cell lysate, respectively. The enrichment of the SH2 domain of Fes decreases slightly, going from 68 phosphosites in T cell activated Jurkat cell lysate as compared to 187 phosphopeptides in EGF-stimulated HeLa cell lysate. The relationship between domains at the phosphosite level remained the same when analyzed by clustering (Figure 2-19).

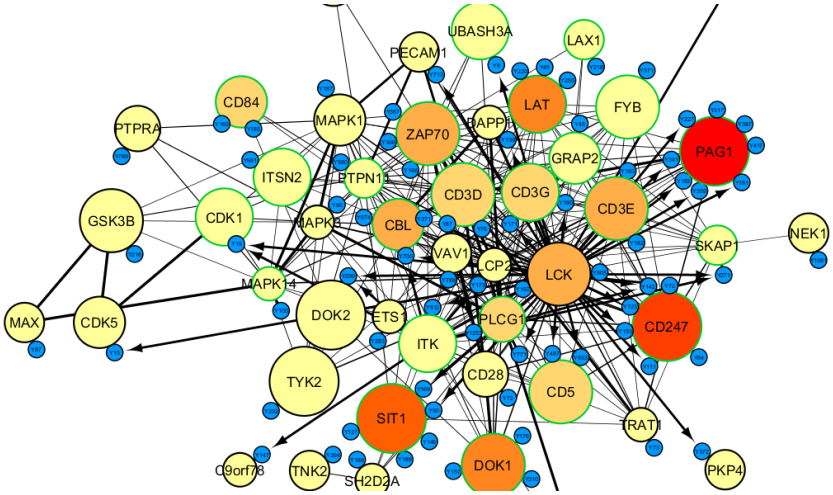


Figure 2-21. PhosphoPath-generated interaction map of proteins with at least one significantly enriched phosphotyrosine in T cell activated Jurkat cell lysate. Node color is the density of phosphotyrosine at a given protein and node size is represented by the number of SH2 domains in the panel that enriched at least one phosphotyrosine site on that protein target. Phosphotyrosine residues are colored as blue nodes. Proteins with phosphotyrosine sites that were enriched differentially, specific to the activation, by the SH2 domains from Lyn and Src are outlined in green.

Plotting the interaction network of proteins with significantly enriched phosphotyrosine sites revealed a that many of the important canonical T cell receptor signaling proteins are enriched by multiple members of the SH2 panel (Figure 2-21). The T cell surface glycoproteins

CD3E, CD3D, CD5, and CD247 as well as Zap70, Lck, Dok1, and SIT1, which are all important in T cell signaling, had at least one phosphotyrosine site enriched by six or more members of the SH2 domain panel (26-31). The proteins who were only enriched by a few of the SH2 domains are diverse in identity: the kinases p38a, ERK1, PTK2B, TNK2, and NEK1; Rho activity regulators plakophilin 4, and GCSAM; and E3-ligase TRIM4. Amazingly, the panel of seven SH2 domains was able to enrich a majority of proteins at the site of T cell receptor signaling, all of which had enriched phosphotyrosine sites, such as the CD3 chains, Zap70, Lck, Lat, GRAP2, ITK, Dok1, Dok2, Csk, Pag1, and Fyn (31).

The SH2 domain of Fes was the most specific and limited out of the SH2 domain panel at both the protein and phosphosite level in both signaling systems. Fes kinase plays a role in regulating the actin cytoskeleton and microtubule assembly, and even requires a functional SH2 domain for its microtubule localization (32). The Fes SH2 domain additionally has been shown to be involved in mast cell signaling through interaction with phosphorylated FεRI beta chain and HS1, a protein that regulates actin (33). Fes kinase has not explicitly been implicated in the EGFR or T cell receptor pathway thus far and as such, its limited enrichment profile is perhaps not surprising. A protein microarray study using 61 different peptides from physiological sites of phosphotyrosine from four ErbB receptors reported an EGFR-Fes SH2 interaction, but we only observed roughly a 0.5-fold enrichment over control on the log2 scale (34).

For both biological systems, Brk and Src SH2 exhibited the highest levels of enrichment relative to other members of the panel. In the EGFR signaling pathway, Src is known to phosphorylate EGFR in its kinase domain at Y867 and there is evidence that Src plays some role as a signaling pathway mediator of other EGFR tyrosine phosphorylation sites, including Y1110 and Y1016 (35-37). The role of Src in the T cell receptor pathway is less clear as the Src-family kinases Fyn and Lck are known to be the key players in these signaling events.

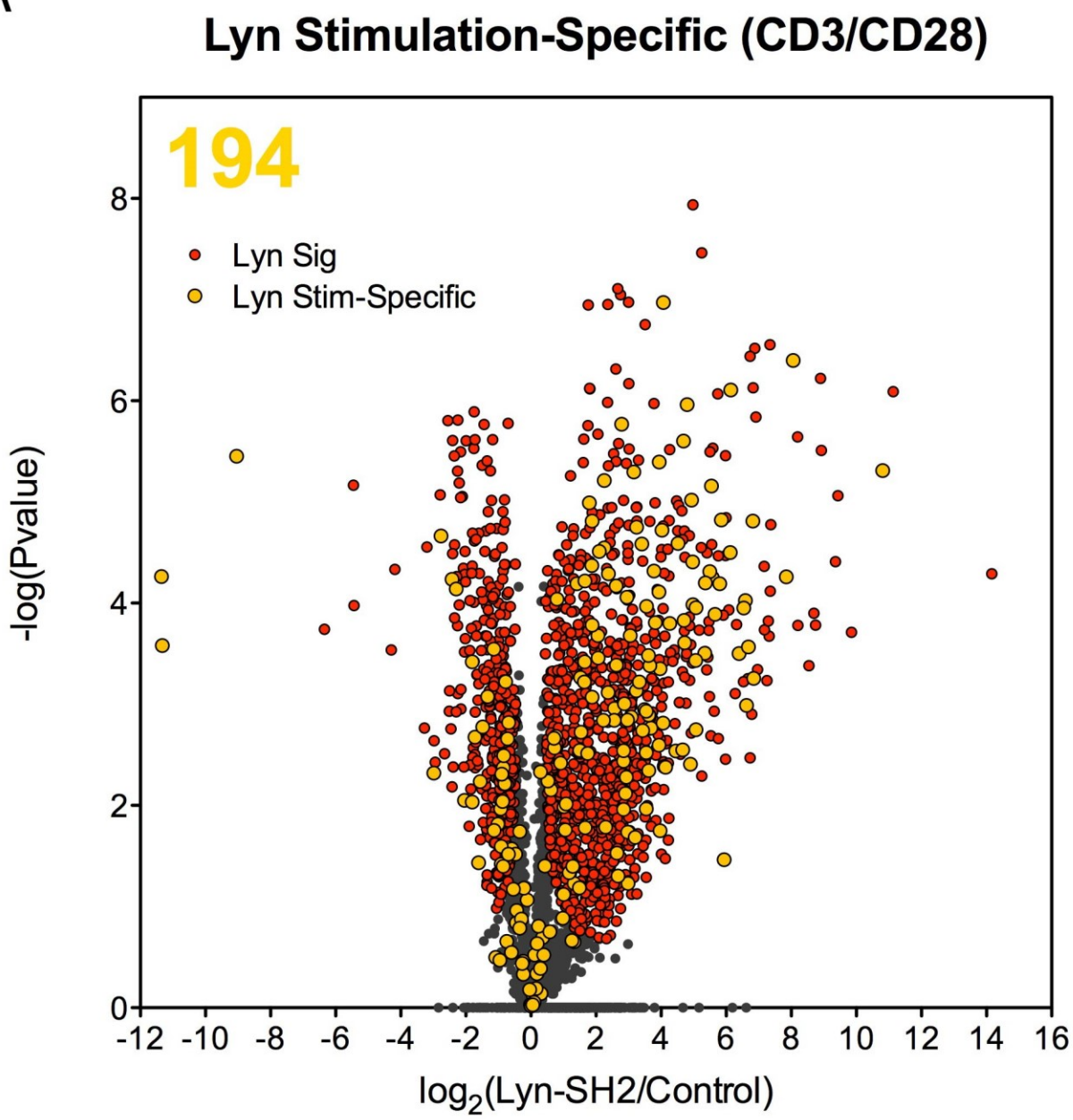
Brk is a non-receptor tyrosine kinase that has been implicated with a wide variety of signaling molecules including EGFR and paxillin (38, 39). Interestingly, Brk was found to be activated by EGFR and subsequently phosphorylates Y867 in the EGFR catalytic domain to sustain activated EGFR signaling (40). With regards to the T cell receptor signaling pathway, Brk expression was found to be induced upon activation of T cell lymphomas and that it may play a role in lymphomagenesis as si-RNA-mediated knockdown of the kinase decreased T cell growth and survival (41). However, there has not been any study showing evidence for a direct interaction with T cell receptors. In our study, we did not observe the T cell receptor proteins in the proteomic scan, but we did observe multiple phosphopeptides from these receptors. It seems more

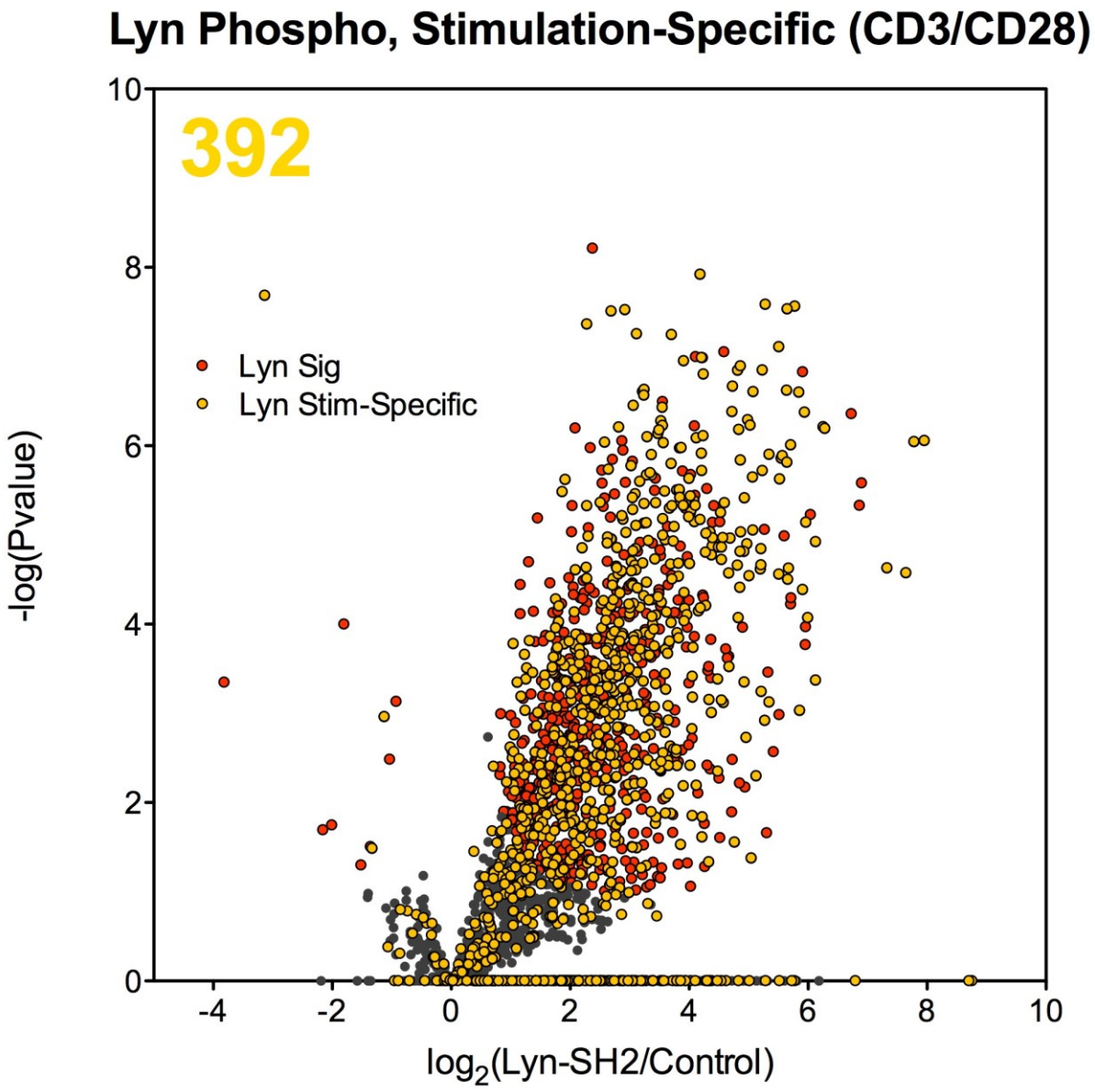
reasonable that these interactions may be secondary and mediated by Brk's interaction with Cbl and possibly Dok1. We observed that Brk pulled down a large number of proteins involved in mRNA transcription, translation, and splicing, which was evidenced from our GO enrichment analysis. This result corroborates studies that indicate Brk regulates proteins involved in RNA processing, specifically by binding to and phosphorylating nuclear RNA-binding proteins such as the SAM68 family member KHDRBS1 to negatively regulate its RNA binding capacity (41). Furthermore, KHDRBS1 seems to function as an adapter protein downstream of several signaling proteins including EGFR (43) and the T cell receptor (44). In our proteomic profiling, we observed greater than 2^5 enrichment for KHDRBS1 in both Brk and Src, which may point to their ability to pull down multiple nuclear and RNA/DNA binding proteins.

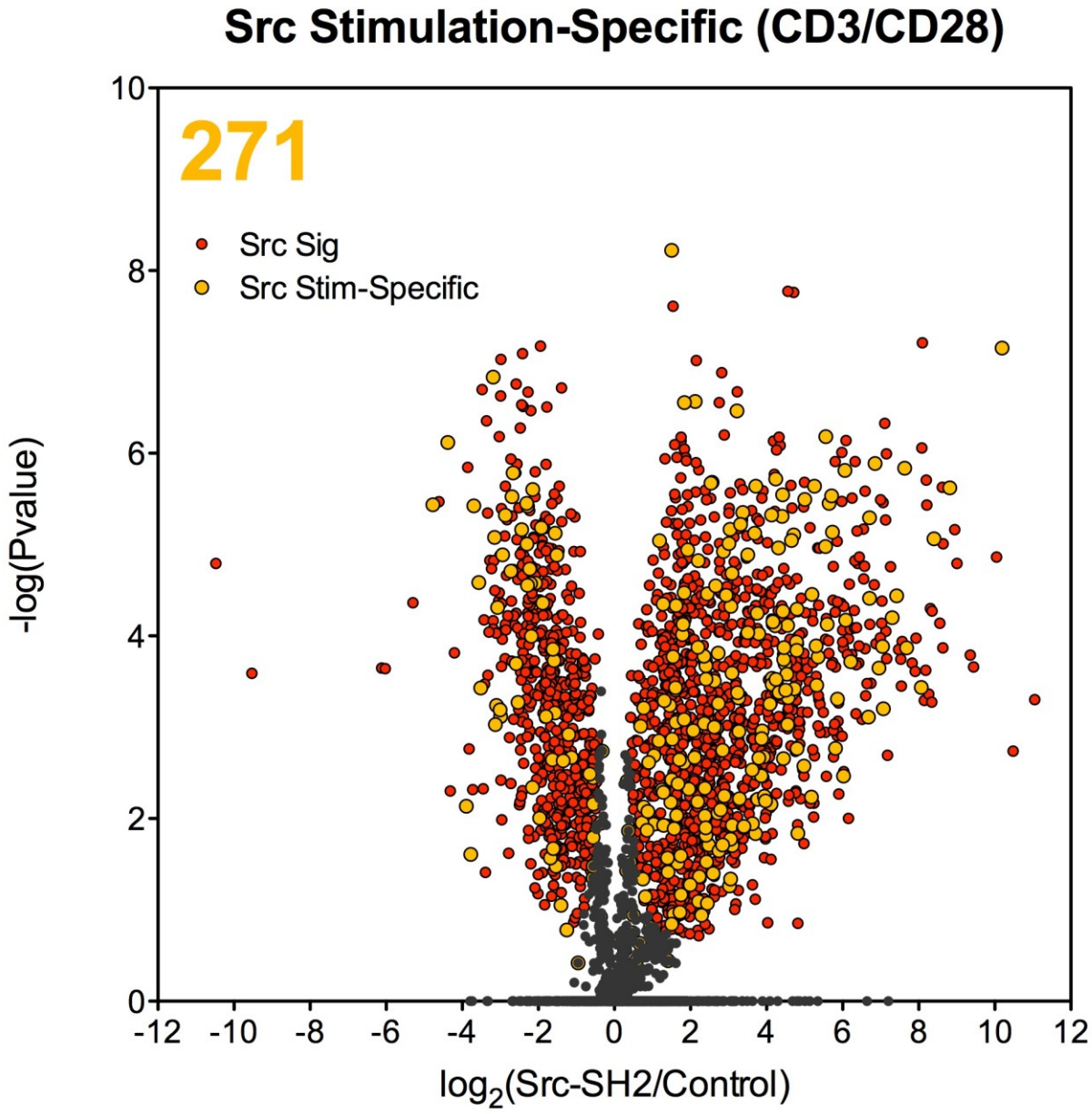
Comparing Interactors from Stimulated and Non-Stimulated Jurkat for Lyn and Src SH2

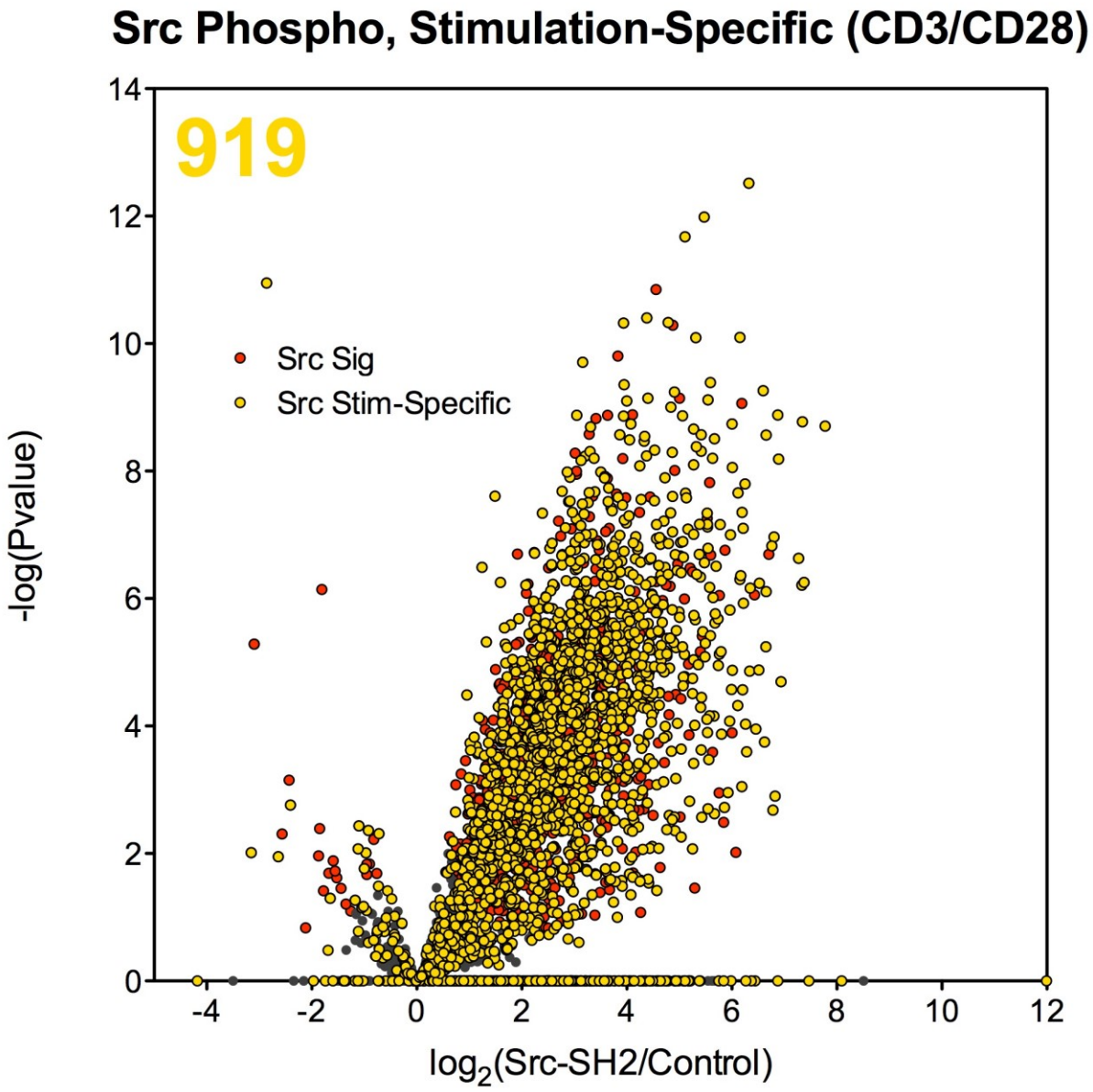
Similar for EGF-stimulated HeLa cell lysate, we next profiled the SH2 domains of Lyn and Src against non-stimulated Jurkat lysate to identify the interacting hits that were a result of T cell activation. The preparation of samples and data analysis remained the same as was done for the prior study.

A


Lyn Stimulation-Specific (CD3/CD28)
194
Lyn Sig
Lyn Stim-Specific
-log(Pvalue)
log2(Lyn-SH2/Control)
0
2
4
6
8
-12
-10
-8
-6
-4
-2
0
2
4
6
8
10
12
14
16


Lyn Phospho, Stimulation-Specific (CD3/CD28)
392
Lyn Sig
Lyn Stim-Specific
-log(Pvalue)
log2(Lyn-SH2/Control)
10
8
6
4
2
0
-4
-2
0
2
4
6
8
10


Src Stimulation-Specific (CD3/CD28)
271
Src Sig
Src Stim-Specific
-log(Pvalue)
log2(Src-SH2/Control)
10
8
6
4
2
0
-12
-10
-8
-6
-4
-2
0
2
4
6
8
10
12


Src Phospho, Stimulation-Specific (CD3/CD28)
919
Src Sig
Src Stim-Specific
-log(Pvalue)
log2(Src-SH2/Control)
14
12
10
8
6
4
2
0
-4
-2
0
2
4
6
8
10
12

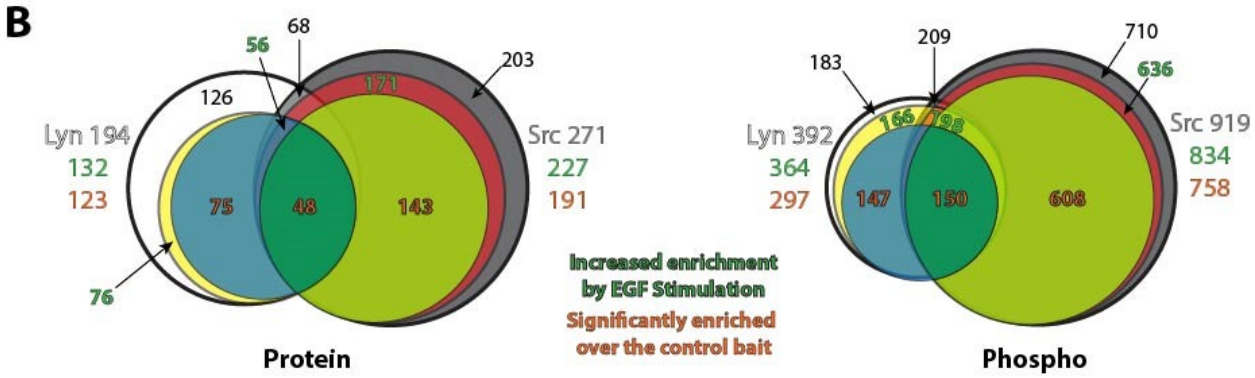


Figure 2-22. Analysis of the enrichment with non-stimulated lysate for Src and Lyn, showing proteins and phosphosites that showed differential identification as a result of T cell activation in Jurkat cells. (A) Volcano plots comparing each SH2 domain versus the control in stimulated lysate with the stimulation specific hits overlayed in yellow. (B) Venn diagrams comparing the total number of activation-specific proteins and phosphosites captured by Lyn- and Src-SH2. The subset of those proteins and sites that were enriched more in the stimulated lysate versus non-stimulated are highlighted with green text. The stimulation-dependent hits that were specifically captured by the respective SH2 domain over the control bait are detailed in orange text.

The most immediate and visible result from this comparative profiling is that there is a significantly higher amount of proteins and phosphosites that arose from T cell activation than was observed for EGF stimulation, especially for Src (Figure 2-22A). The number of activation-dependent phosphosites for Lyn SH2 in this signaling pathway also dramatically increased compared to HeLa (392 versus 90), suggesting a larger role in T cell activation than EGF signaling. While the proportion of these proteins and phosphosites that are preferentially enriched in the stimulated lysate remains high, there is a much larger subset of these that are also specific to the SH2 domain interaction than observed in HeLa (Figure 2-22B). We also observed that a larger number of the activation-dependent phosphotyrosine sites for these two domains lie near the stimulated T cell surface glycoprotein receptors (Figure 2-21). The concentration of these phosphotyrosine sites suggests that Lyn and Src are more intimately involved near the sites of receptor activation and regulation. Lyn and Src have less overlap in shared activation-specific sites, which is similar to the observation with EGF signaling HeLa. This may indicate that there is greater divergence in recognition specificity at the respective signaling machinery level between these two kinase SH2 domains and that their enrichment similarities lie in other functional classes.

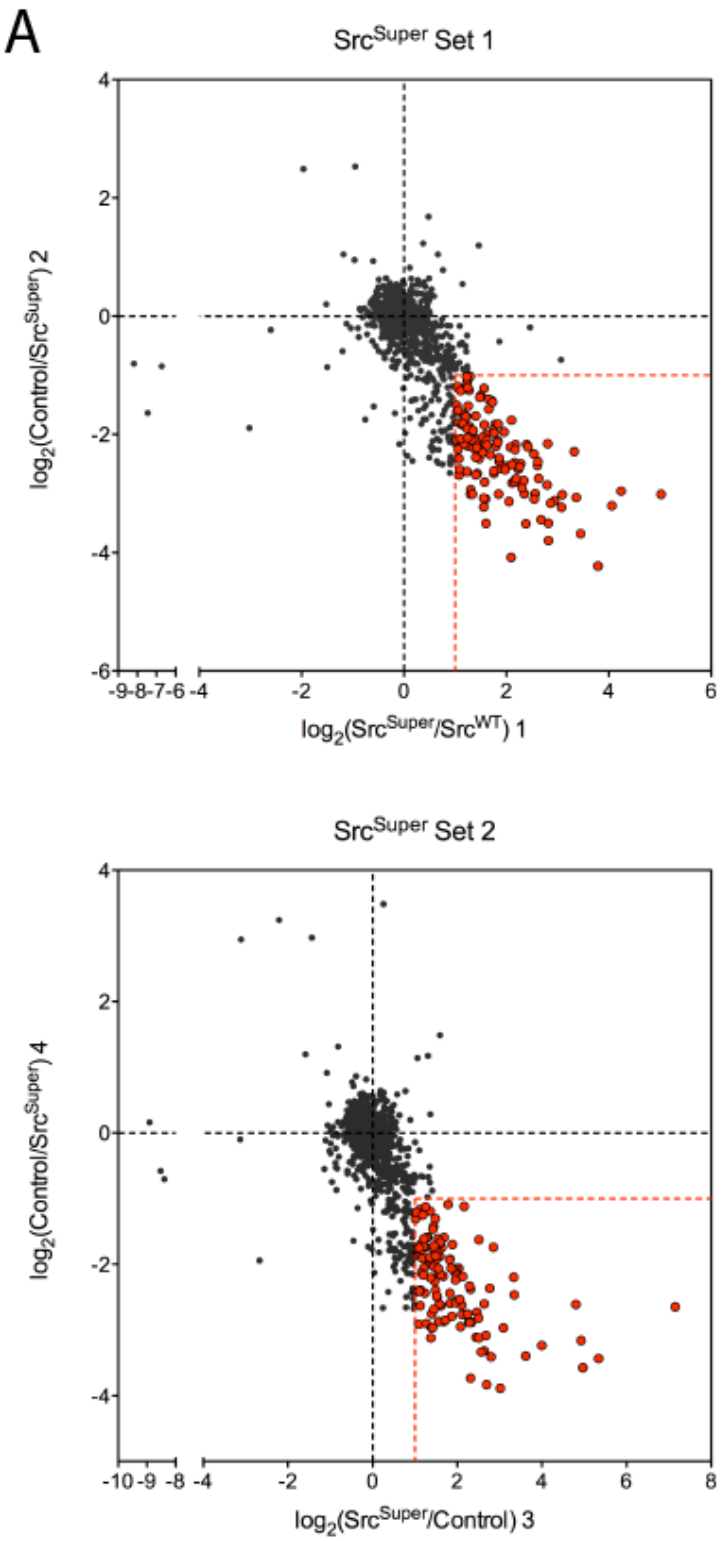
Src SH2 Superbinder Profiling

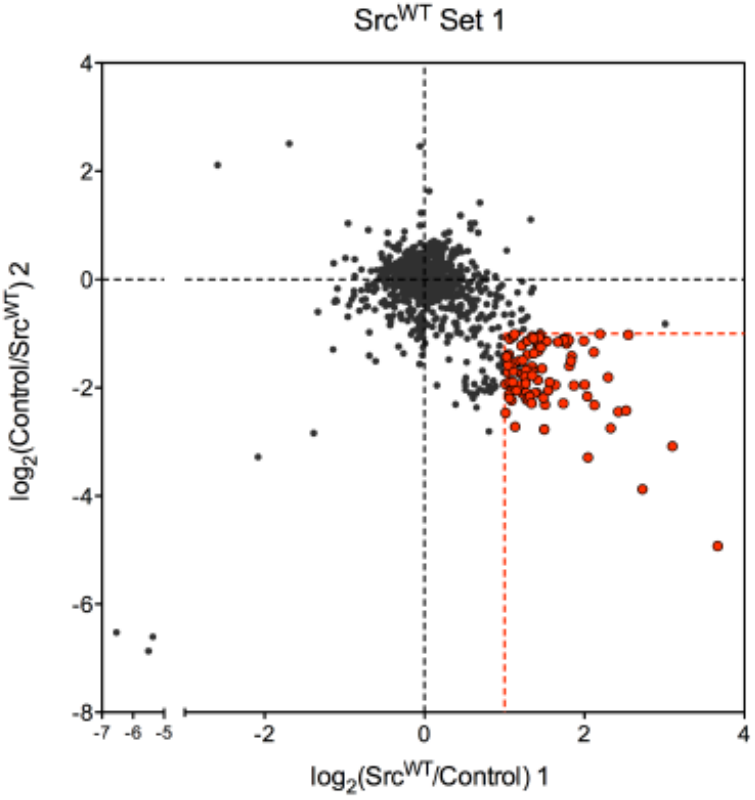
The basic phosphotyrosine binding ability of SH2 domains is a result of a conserved binding pocket. A second pocket confers enhanced specificity and binding to a recognition sequence that is C-terminal to the phosphotyrosine of the ligand, resulting in a dual binding mode that is moderate in strength. In 2012, Kaneko and coworkers used phage display to optimize fifteen amino acid residues in the phosphotyrosine binding pocket of the SH2 domains from Fyn, Src, and Lyn to enhance binding strength (45). The resulting optimized triple-mutant SH2 domains, dubbed "superbinders," yielded submicromolar K_d values against a GGpTyrGG peptide, indicating that the binding is a result of affinity for the phosphotyrosine alone. In a fluorescent polarization assay, the Src SH2 triple mutant T183V/C188A/K206L, or Src^{Super}, showed a K_d of 3.8 nM as compared to a K_d = 0.70 μM for the wild type. This increase in binding affinity remained true for several other phosphotyrosine peptides from physiological phosphotyrosine sources, including EGFR, ShcA, and VEGFR.

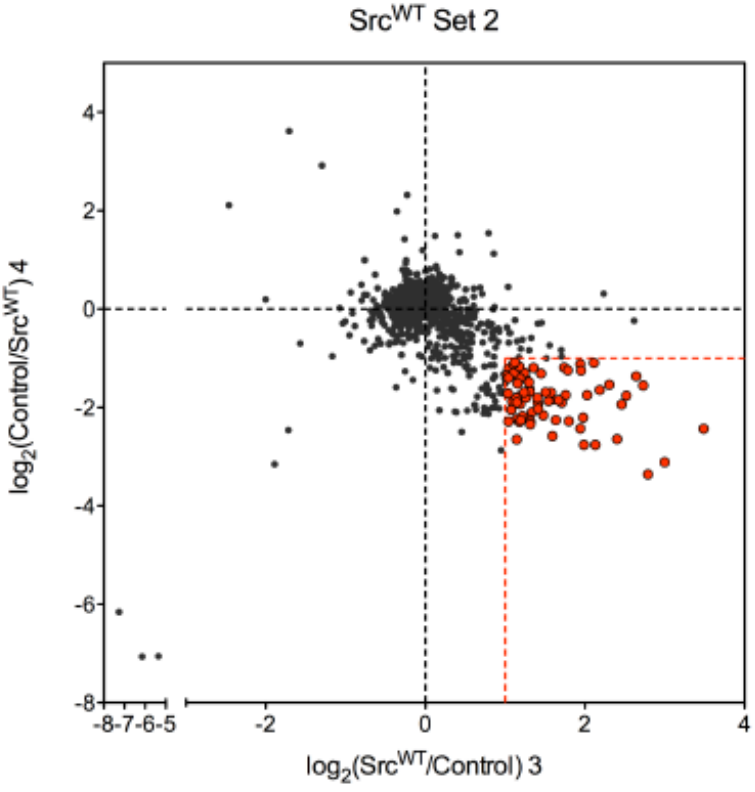
Src^{Super} was subsequently used as a phosphoproteomics affinity reagent in 2016 to identify greater than 10,000 phosphotyrosine sites, 36% of which were novel (46). The authors immobilized Src^{Super} onto Ni-NTA resin and then pulled down from trypsinized lysate from pervanadate-treated HeLa or Jurkat cells. In this approach, the superbinder enriched more phosphotyrosine sites from HeLa lysate than the two phosphotyrosine-specific antibodies 4G10 and P-Tyr-100. In fact, when used in sequential enrichment after first enriching with 4G10 antibody, 2,783 new phosphotyrosine sites were identified with the median intensity of phosphotyrosine peptides from Src^{Super} enrichment, 20-fold higher. Src^{Super} was able to capture 82% of the phosphotyrosine peptides pulled down by the antibody. When assayed at low amounts, 0.375 nmol, the authors observed that Src^{Super} tended to select for phosphopeptides with a phosphotyrosine-E-X-L/I-K motif, but at larger amounts including up to 11.25 nmol, the Grb2 and Src Superbinder forms converged in specificity due to high enrichment of the available phosphotyrosine proteome overall.

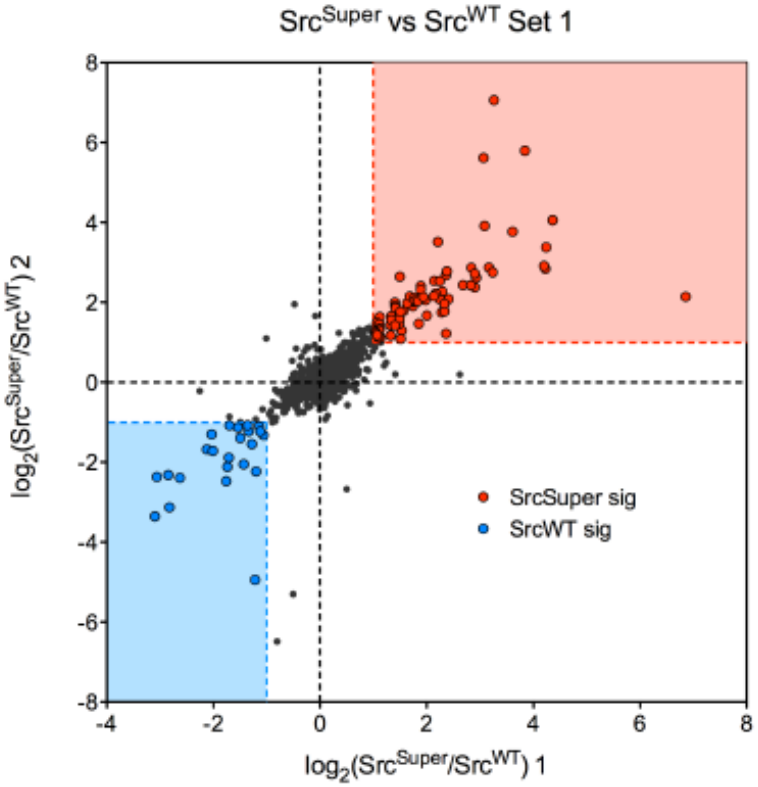
Given these interesting properties, we generated a Src^{Super}-SNAP-tag fusion to test in our AP-MS approach. Interestingly, we observed from our initial Western blot profiling that this superbinder showed a distinct enrichment motif from its wild-type counterpart (Figure 2-3). Src^{Super} did not have a denser region of phosphotyrosine enrichment as expected, but simply appeared to be a unique SH2 domain among the panel. To follow up on this result, triple SILAC-labeled HeLa cells were EGF-stimulated for 5 minutes, lysed, and then exposed to 1.875 nmol of either resin-immobilized phosphotyrosine dead control, Src^{WT} SH2, or Src^{Super} SH2. The resulting

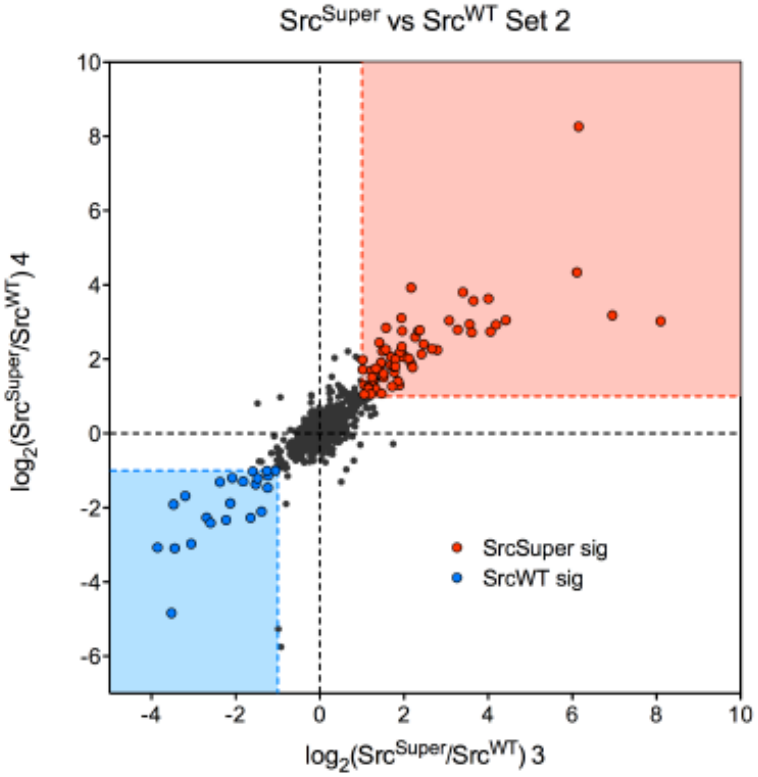
elutions were subsequently mixed, trypsinized, and submitted for proteomic and phosphoproteomic analysis.



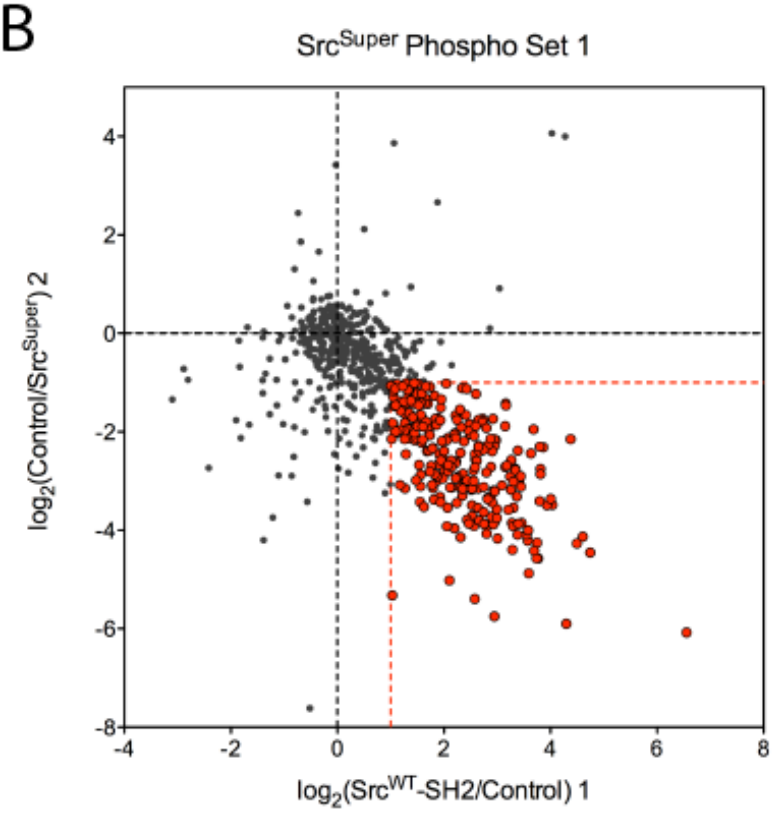

Src^WT Set 1
log2(Control/Src^WT) 2
log2(Src^WT/Control) 1
4
2
0
-2
-4
-6
-8
-7 -6 -5
-2
0
2
4

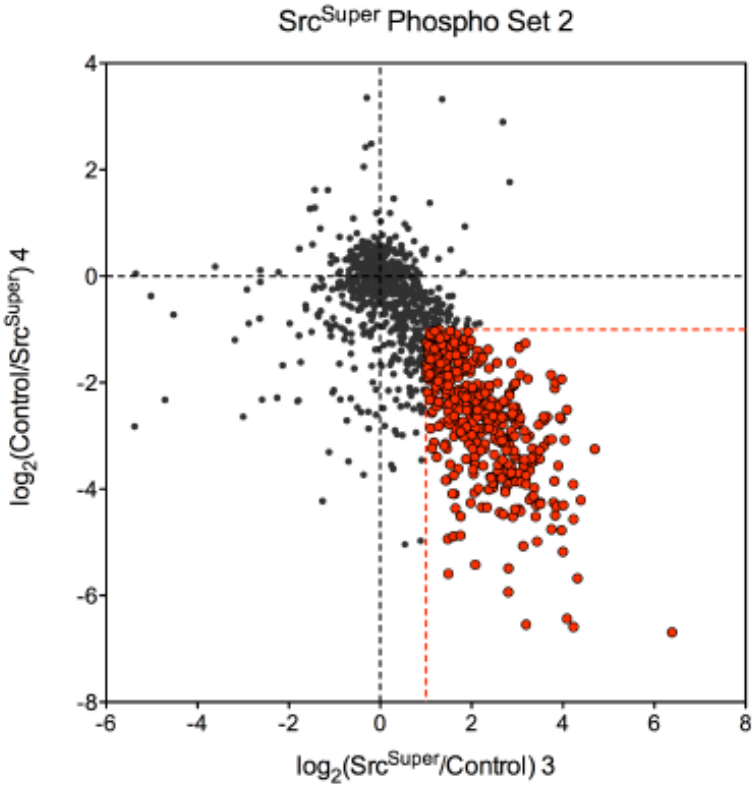

Src^WT Set 2
log2(Control/Src^WT) 4
log2(Src^WT/Control) 3
4
2
0
-2
-4
-6
-8
-8-7-6-5
-2
0
2
4

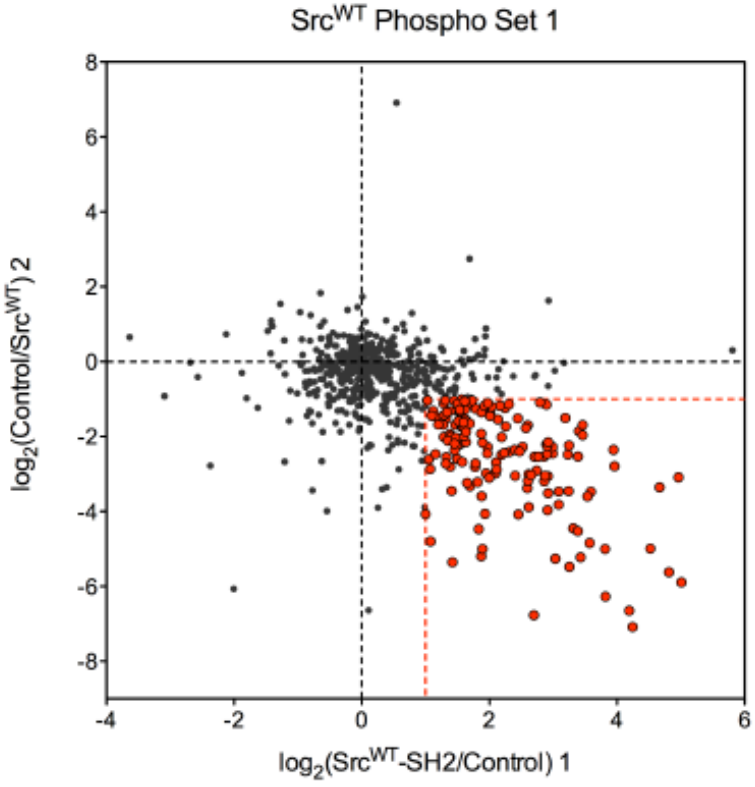

Src^Super vs Src^WT Set 1
log2(Src^Super/Src^WT) 2
log2(Src^Super/Src^WT) 1
SrcSuper sig
SrcWT sig
8
6
4
2
0
-2
-4
-6
-8
-4
-2
0
2
4
6
8

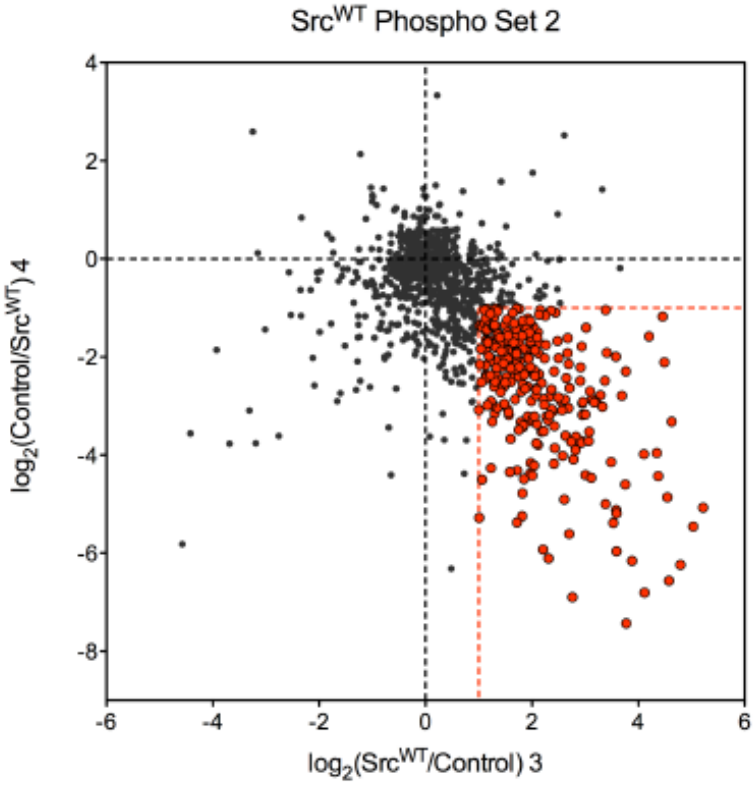

Src^Super vs Src^WT Set 2
log2(Src^Super/Src^WT) 4
log2(Src^Super/Src^WT) 3
SrcSuper sig
SrcWT sig
10
8
6
4
2
0
-2
-4
-6
-4
-2
0
2
4
6
8
10

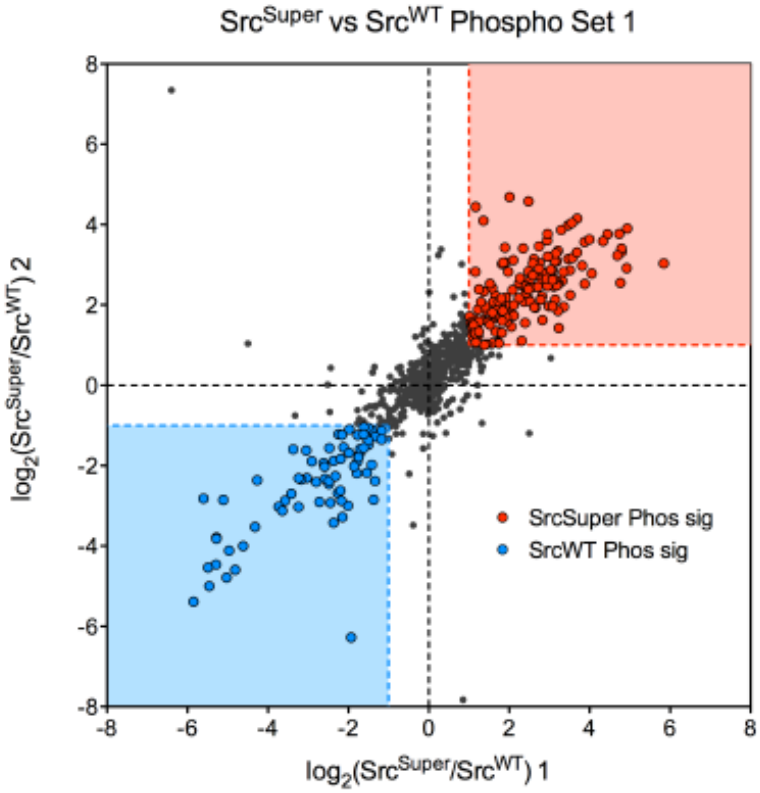
B

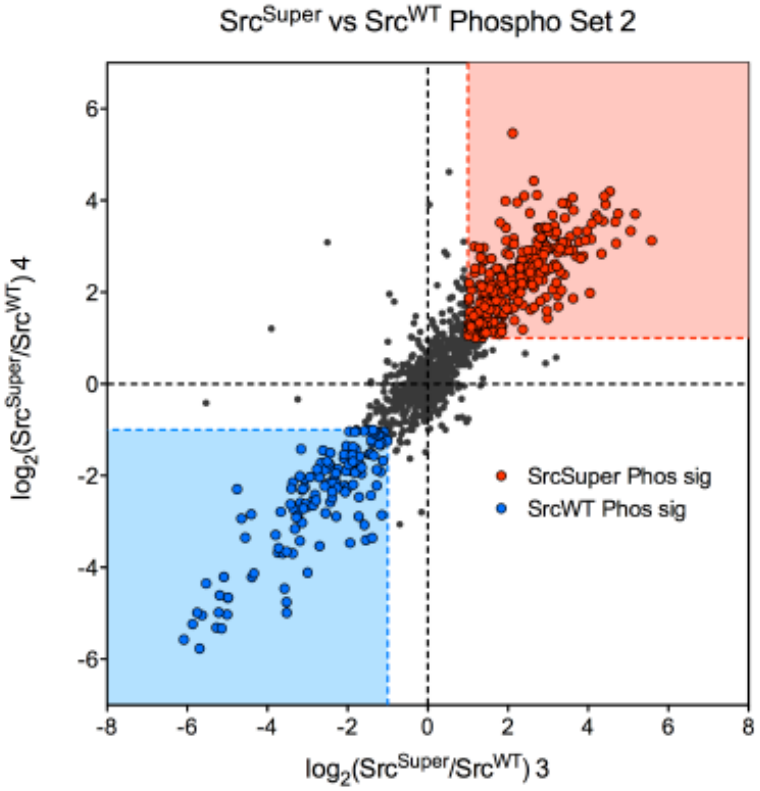

Src^Super Phospho Set 1
log2(Control/Src^Super) 2
log2(Src^WT-SH2/Control) 1
4
2
0
-2
-4
-6
-8
-4
-2
0
2
4
6
8


Src^Super Phospho Set 2
log2(Control/Src^Super) 4
log2(Src^Super/Control) 3
4
2
0
-2
-4
-6
-8
-6
-4
-2
0
2
4
6
8

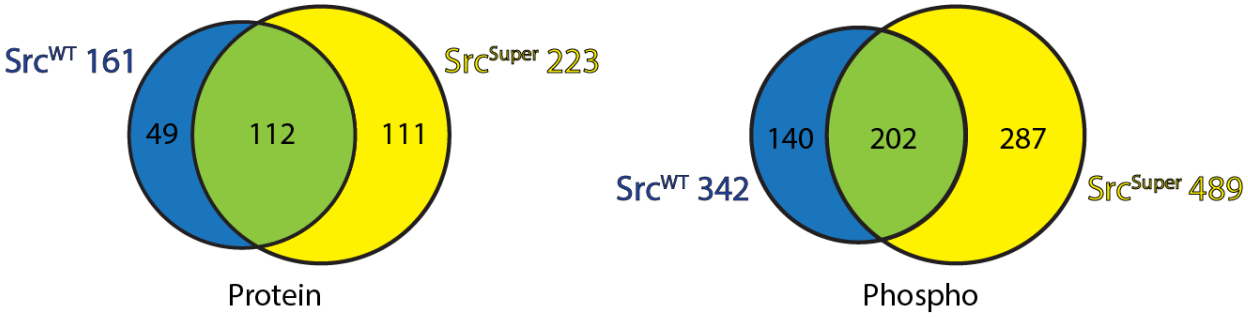

Src^WT Phospho Set 1
log2(Control/Src^WT) 2
log2(Src^WT-SH2/Control) 1
8
6
4
2
0
-2
-4
-6
-8
-4
-2
0
2
4
6


Src^WT Phospho Set 2
log2(Control/Src^WT) 4
log2(Src^WT/Control) 3
4
2
0
-2
-4
-6
-8
-6
-4
-2
0
2
4
6


Src^Super vs Src^WT Phospho Set 1
log2(Src^Super/Src^WT) 2
log2(Src^Super/Src^WT) 1
SrcSuper Phos sig
SrcWT Phos sig

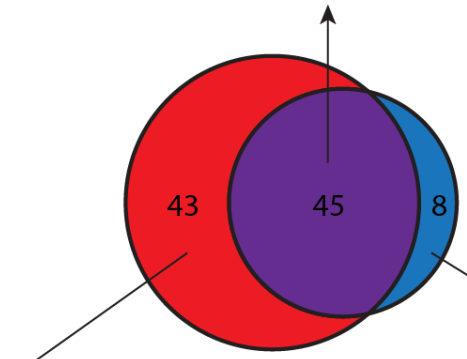

Src^Super vs Src^WT Phospho Set 2
log2(Src^Super/Src^WT) 4
log2(Src^Super/Src^WT) 3
SrcSuper Phos sig
SrcWT Phos sig

C



D

TERM	PROTEIN COUNT
cytoplasmic translation	13
ribosomal small subunit biogenesis	8
ribosomal small subunit assembly	7
ribosomal large subunit assembly	6
CRD-mediated mRNA stabilization	4
epidermal growth factor receptor signaling pathway	5



TERM	PROTEIN COUNT
antigen processing and presentation of exogenous peptide antigen via MHC class II	10
G2/M transition of mitotic cell cycle	11
microtubule-based process	5
ER to Golgi vesicle-mediated transport	7
activation of MAPK activity	6
COPII vesicle coating	5

TERM	PROTEIN COUNT
RNA splicing, via transesterification reactions	3
regulation of cell communication by electrical coupling involved in cardiac conduction	2
nucleosome assembly	3
regulation of release of sequestered calcium ion into cytosol by sarcoplasmic reticulum	2
regulation of ryanodine-sensitive calcium-release channel activity	2

Figure 2-23. Summary of the proteomic analysis of Src^{Super} and Src^{WT} SH2 in EGF-stimulated HeLa cell lysate. Scatter plots displaying the enrichment results of protein (A) and phosphopeptide (B) targets for Src^{Super} and Src^{WT} against one another or versus the SH2-dead control. (C) Venn diagrams comparing the enriched protein and phosphosites for Src^{Super} and Src^{WT}. (D) Comparison of the GO-BP enrichment terms for the significantly identified proteins for each domain.

Proteins or phosphopeptides with enrichment ratios greater than 2-fold over the control in at least three out of the four SILAC replicates, or 2-fold in one replicate with at least an enrichment

of 0.5 on the log2 scale in 3 of 4 replicates were considered significantly enriched. Scatter plots displaying the enrichment with 2-fold cutoff of the SH2 domains for Src^{WT} and Src^{Super} versus the control bait, and also versus one another are shown in Figure 2-23A-B. When plotted against one another, Src^{WT} and Src^{Super} show distinct sets of proteins and phosphopeptides identified for each. This was a surprising result as we had expected to see the superbinder encompass the majority of the identified proteins and phosphopeptides for Src^{WT} in addition to its own large subset of enriched species. Based on these results, the three point mutations in the SH2-binding pocket appear to not only greatly increase phosphotyrosine binding affinity, but suggests that these also affect binding specificity from the wild-type. As shown in Figure 2-23C, the two SH2 domains share a subset of proteins and phosphopeptides, but also differentially enrich their own groups with the superbinder displaying higher numbers of enriched hits overall at both the protein and phosphopeptide level.

If the identified protein hits unique to each domain and those shared are analyzed by GO enrichment, a larger amount of enriched terms is observed for the superbinder, which is expected given that it has a larger set of protein hits (Figure 2-23D). Only GO-BP terms with P-values of 0.05 or less were considered for analysis and comparison. One of the shared terms between the two sets of protein lists, the EGFR signaling pathway is reasonable given the enriched sample space is from EGF-stimulated HeLa cell lysate. Interestingly, a dominating shared GO term deals with ribosome biogenesis and assembly, which is a result of enrichment of several ribosomal proteins for both superbinder and WT SH2 domain. The top unique terms by P-value for Src^{Super} appear to be related to processes involving transport and signal processing via vesicles and microtubule movement. In contrast, the few unique terms for the wild-type SH2 domain are populated by a diverse set of processes that have very small numbers of proteins populating these search terms. The GO enrichment analysis points towards a more compelling case that the superbinder simply enhances the phosphotyrosine binding of the wild type SH2 domain as the set of proteins unique to Src^{WT} appear to disjointed biologically. However, this is still difficult to point to given the numbers of non-overlapping protein and phosphosite hits by MS profiling.

Confirmation of MS-Determined SH2 Domain Targets via Western Blot

A surprising result of the MS analysis of the SH2 domain panel from affinity purification at the protein level was the widespread enrichment profile of Src and Brk. In addition, most of the SH2 domains were able to enrich some of the core members of the signaling pathway being

activated in each system. Considering the experimental procedures in place for MS analysis, the only reasonable hypothesis was that each SH2 domain was enriching from different subsets of signaling complexes that were comprised of some of the same, main signaling pathway members. Some of these SH2 domains were very specific in sets they bound to while others such as Src and Brk were much more promiscuous. Thus, the secondary interactions for these protein domains led to some of the differences observed in protein and phosphosite enrichment. We sought to confirm some of the protein and phosphosite targets that were either core signaling pathway members or appeared to be secondary interactors based on analysis of the STRING network between baits and enriched phosphotyrosine containing proteins (Figure 2-12, 2-21).

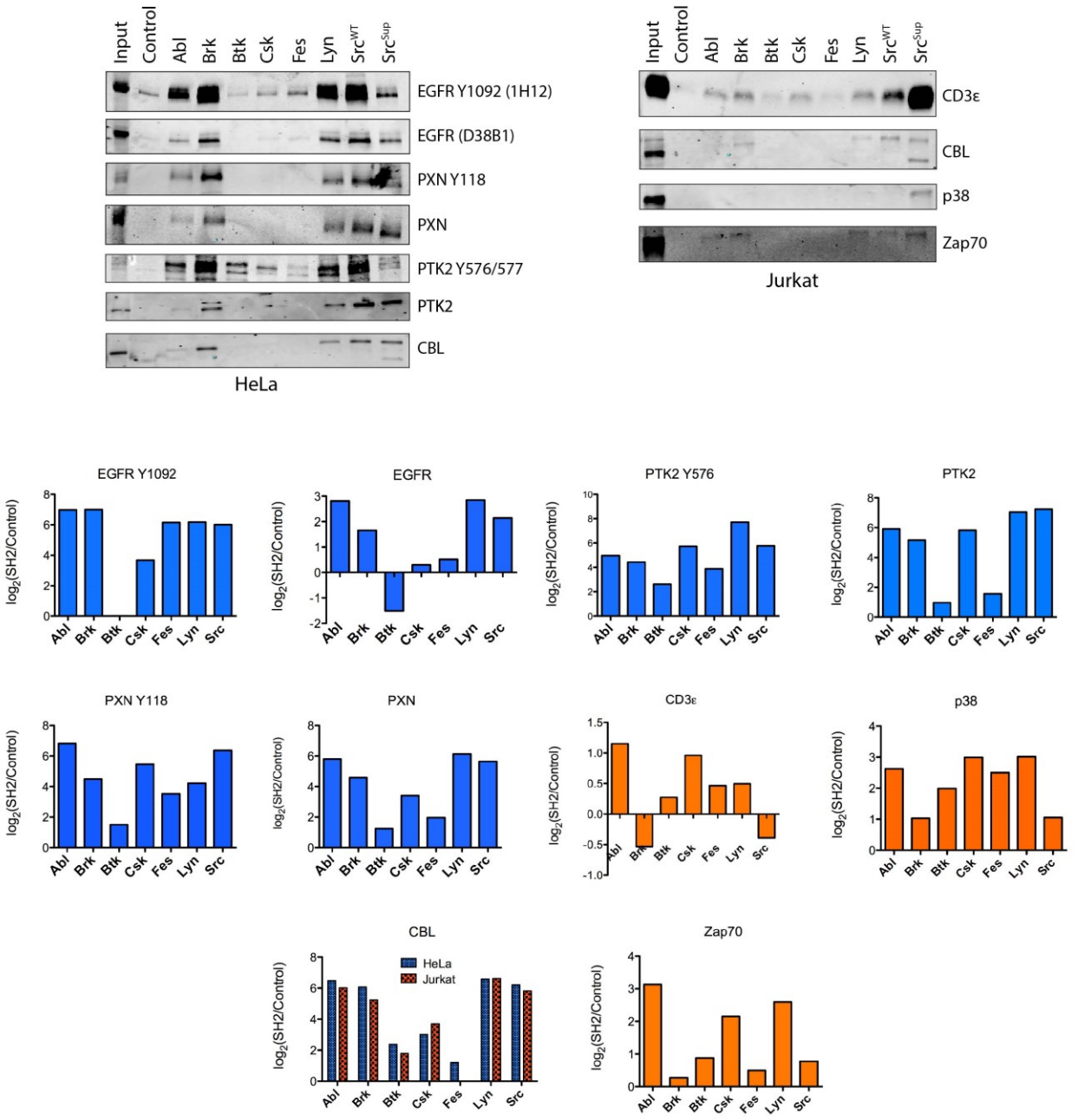


Figure 2-24. Affinity purification and Western blot analysis on a selected panel of enriched targets from EGF-stimulated HeLa cell lysate and T cell activated Jurkat cell lysate. The corresponding MS quantification for a each SH2 domain over the control is plotted for each protein or phosphotyrosine.

Immunoblot confirmation of a variety of protein and phosphotyrosine-containing targets was performed on the eluted fraction of each SH2 domain following enrichment from EGF-stimulated HeLa or T cell activated Jurkat cell lysate. As shown in Figure 2-24, EGFR and EGFR Y1092, an important regulatory site implicated in a variey of processes such as altered enzymatic activity, intracellular localization, and molecular association, matches closely in the Western blot

with the enrichment observed in the proteomic studies for each SH2 domain in EGF-stimulated HeLa lysate. For paxillin and its phosphotyrsoine site Y118, enrichment is observed for Lyn and Src, but not for Csk and Fes, which had been detected in the MS method. Abl SH2 also has a much lower signal relative to the control by immunoblot than MS quantification. The blot for PTK2 and PTK Y576 yielded mixed results as we observed immunoprepicipation for the protein only for Brk, Lyn, Src^{WT}, and Src^{Super}, but some amount of enrichment for the phosphotyrosine site for all eight members of the panel. The SH2 domain of Csk displayed significant enrichment of PTK2 at the protein level in the proteomic study, but we did not detect this difference over the control in the Western blot in EGF-stimulated HeLa cell lysate. The Y576 site did show enrichment for Csk, but did not match the overall level from phosphoproteomics, where it was close in magnitude to that of Brk and Src. Cbl, which is an important handle that potentially links several of these SH2 domains to the signaling receptors studied, did not corrobate the proteomic scan. In both the T cell activated Jurkat and EGF-stimulated HeLa cell lsyates, Abl, Brk, Csk, Lyn, and Src showed high levels of enrichment, but only Brk, Src^{WT}, and Src^{Super} showed appreciable amounts in EGF-stimulated HeLa cell lysate and only Src^{WT} and Src^{Super} in T cell activated Jurkat cell lysate.

In T cell activated Jurkat cell lysate, we also probed for the enrichment of CD3ε, p38α, and Zap70 by each of the seven SH2 domains in the panel plus Src^{Super}. The MS enrichment of the T cell surface glycoprotein CD3ε was validated for the SH2 domains of Abl, Csk, and Lyn. However, we observed conflicting results for the SH2 domains of Brk and Src, which did not enrich CD3ε in the proteomic scan, but had strong capture as detected by Western blot. Both the SH2 domains of Src and Brk enriched phosphopeptides from CD3ε, which provides additional evidence in favor of the Western blot results. Interestingly, Src^{Super} showed by far the strongest enrichment of CD3ε in the Western blot although we do not have proteomic results for comparison. All seven members of the SH2 domain panel significantly enriched p38α from the MS results, but none showed any enrichment by Western blot. The SH2 domain of Src^{Super} was the only one to show enrichment against p38α. The results for Zap70 were also mixed. The SH2 domains of Brk and Src showed stronger enrichment of Zap70 by Western blot in T cell activated Jurkat cell lysate than in the MS method. Zap70 enrichment by the SH2 domains of Abl and Lyn was confirmed, but the magnitude of the enrichment relative to Brk and Src was larger in the proteomic results than via Western blot. Additionally, we expected to see Csk SH2 enrich Zap70, but this was not corroborated by Western blot.

Src and Lyn SH2 Chimeras for Differential Kinase-Phosphotyrosine Targeting

The proteomic profiling in both HeLa and Jurkat cell lysate showed varied enrichment sets for the SH2 domains of Lyn and Src. These two related Src family tyrosine kinases enriched similar proteins and phosphosites in both signaling pathways, but we noticed that Src had a large fraction of interactors not covered by Lyn's SH2. Inspired by chimeric approaches such as the work of Miller and coworkers, who replaced the SH2 and SH3 domains of Src family kinase Hck with the syntrophin PDZ domain, we sought to generate chimeric, full-length Lyn and Src that have swapped SH2 domains (47). We also chose to include a third choice, the Src^{Super} triple mutant SH2 domain, for inclusion into full-length Src kinase. Additionally, the tails of both kinases are mutated (Y527F in Src, Y508F in Lyn) to constituetively activate the kinase. These activated, chimeric kinases should have altered substrate targeting because of the replaced SH2 domains and will ideally mirror our MS profiling results.

III. Conclusion

Profiling tyrosine kinase SH2 domains in stimulated cell lysate offers information that complements previous peptide microarray studies, but this approach can also elucidate secondary protein-protein interactions that are lost in isolated SH2 domain-peptide interactions or gel-based assays. In addition, there are differences between SH2 recognition in denatured proteins or small peptides when compared to the dynamic structure present within the cell. The cellular interactome of each SH2 domain in HeLa and Jurkat signaling illustrates that while there are distinct enrichment profiles for each, they also share common signaling hubs. The differing levels of enrichment for each member of the SH2 domain panel, most notably for the large numbers of protein and phosphosites for the SH2 domains of Brk and Src versus the much more limited and discriminating pools of Fes and Csk, corroborates some of the biology known to these kinases and their potential involvement in the two signaling pathways studied.

Phosphotyrosine analysis of the panel of tyrosine kinase SH2 domains reveals that for enrichment at the protein level, there is not the significant difference between enriched phosphotyrosine that was expected based on peptide array studies and previously determined preferred phosphotyrosine recognition motifs. We do not believe this indicates that these SH2 domains are non-discriminating between phosphotyrosine motifs, but rather that in the context of a signaling cascade, multiple phosphotyrosine handles are created on a core, interacting set of

proteins and these SH2 domains interact with varied complexes differing in their secondary interactions.

IV. Methods

SH2 Domain-SNAP-tag Fusion Design

The SNAP-tag gene fragment was amplified from the pSS26b plasmid (Covalys) and then fused to a gBlock containing the SH2-domain plus a variable GST linker (12-24 amino acids) and part of the SNAP-tag gene into the pMCSG7 vector using Gibson Assembly.

Protein Expression and Purification

Plasmids containing the SH2 domains were transformed into BL21(DE3) *E. coli* cells and a single colony from the transformation was inoculated into 10 mL of LB Broth with 100 ug/mL of ampicillin overnight at 37°C. The inoculation was added the next day to 1 L of LB Broth with 100 ug/mL of ampicillin and grown at 37°C until an optical density at 600 nm of (O.D.) of approximately 0.8-0.9. The culture was cooled to 18°C, induced with 1 mM IPTG, and then grown overnight at 18°C. Cells were pelleted by centrifugation and stored at -80°C. For protein purification, the cell pellets were thawed and resuspended in ice-cold lysis buffer containing 50 mM HEPES pH 7.5, 100 mM NaCl, 20 mM imidazole and supplemented with 1 mM PMSF and Protease Inhibitor Cocktail (Pierce). The cells were lysed with sonication (3 x 30 s pulses) and then clarified by centrifugation at 10,000xg at 4°C for 10 minutes. The clarified lysate was incubated with 1 mL of HisPur Ni-NTA (Thermo Scientific) for 1 hour at 4°C. The resin was spun down, supernatant removed, and washed twice with fresh lysis buffer for 30 minutes at 4°C. Protein was eluted from the resin by adding 5 mL of elution buffer (50 mM HEPES pH 7.5, 100 mM NaCl, 200 mM imidazole) and incubating for 30 minutes at 4°C. The resulting elution was collected using a column. Proteins were desalted into 50 mM Tris pH 7.5, 150 mM NaCl using Slide-a-Lyzer dialysis cassettes, 10K MWCO (Thermo Scientific) for three sequential buffer exchange steps. The purity of each protein was determined by SDS-PAGE and then aliquoted for storage at -80°C.

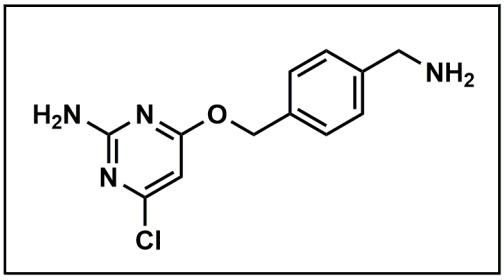
Cell Culture

HeLa and Flp-In Jurkat cells were cultured and maintained according to ATCC standards at 37°C and 5% CO_2 in ATCC recommended media supplemented with Penicillin-Streptomycin (100X, 10,000 U/mL, Gibco) and 10% FBS (One-shot, US origin, Gibco). For triplex-SILAC studies, HeLa cells were SILAC labeled in -Lys/-Arg Dulbecco's modified Eagle's medium, DMEM (Caisson Labs) supplemented with 200 ug/mL of L-proline, SILAC amino acids (0.2 mM Arg0/Lys0 for the light label, 0.2 mM Arg 6/Lys4 for the medium label, and 0.2 mM Arg10/Lys8 for the heavy label, Cambrdige Isotope Labs), Penicillin-Streptomycin (100X, 10,000 U/mL, Gibco), and 10% dFBS (Sigma). Cells were grown in the respective SILAC media for at least five cell doublings.

Cell Stimulation and Harvest

HeLa cells were grown to greater than 90 percent confluency and then serum-starved overnight. The cells were treated with 150 ng/mL of human EGF (Cell Signaling Technology) in media with 20% FBS for 5 minutes and then washed twice with ice-cold dPBS (Gibco). Cells were lysed on ice with 500 uL per 15-cm plate of modified RIPA buffer (50 mM Tris pH 7.8, 150 mM NaCl, 1 mM EDTA, 5% glycerol, 1% NP-40) containing Protease Inhibitor Cocktail (10X, Pierce) and Phosphatase Inhibitor Cocktail 2 and 3 (100X, Sigma). The resultant lysate was vortexed briefly on ice and then centrifuged at 10,000xg for 15 min at 4°C. The clarified lysate was quantified for protein concentration using the Pierce 660 nm Assay Reagent (Thermo Fisher Scientific) and adjusted with modified RIPA buffer to a final concentration of 3 mg/mL. The final sample was snap-frozen in liquid nitrogen and stored at -80°C. Jurkat cells grown to a concentration of approximately 5 x 10^5 cells/mL and then serum starved overnight. The following day, the cells were spun down and concentrated to a final concentration of 500 x 10^6 cells in 10 mL of serum-free media and then added to media containing 5 ug/mL and 2.5 ug/mL final concentration respectively, of anti-CD3 and CD28 antibodies (10 ug/mL, Miltenyi Biotechnology). The cells were incubated for 15 minutes and then spun down for media removal. The stimulated cells were lysed in 10 mL of modified RIPA as described previously and then processed similarly as in HeLa.

Chemical Synthesis



[1] was prepared by another lab member using a previously published procedure (48).

CLP-Functionalized Sepharose Preparation

NHS-Activated Sepharose 4 Fast Flow (GE Healthcare Life Sciences) was washed twice with 2:1:1 of H_2O:EtOH:DMF and resuspended in 1:1 DMF/EtOH. The pH was adjusted to 8 with diisoprophosphotyrosinelethlamine and 1.5-fold molar excess of **[1]** added. The slurry was incubated on a rotator overnight at room temp. The slurry was drained, 0.5 M methoxyethanolamine, 0.5 M NaCl pH 8.5 added, and incubated overnight with rotation at room temp to cap any unreacted NHS. The resin was drained and washed with 0.1 M Tris pH 8.5 and then with 0.1 M sodium acetate pH 5, 0.5 M NaCl. This series of wash steps was repeated twice and the resin stored in 20% EtOH.

SH2 Domain-Resin Immobilization

Each SH2 domain was conjugated to resin via a SNAP-tag labeling reaction by incubating with CLP-functionalized sepharose in 50 mM Tris pH 7.5, 150 mM NaCl, 4mM DTT overnight at 4°C. The SH2 domain was added in slight excess such that approximately 4 nmol of bait would be conjugated on 20 uL of resin. Resin loading was estimated by measuring the protein concentration in the flowthrough by NanoDrop. The resulting SH2-domain immobilized resin was washed three times with 1 mL of cold TBS pH 7.5 and used the same day.

SH2 Domain Affinity Enrichment and Sample Digestion

For enrichment of SH2 domain interactors, an appropriate amount of SH2-immobilized resin such that 4 nmol of the SH2 domain was added to 1.5 mg of cell extract (500 uL, 3 mg/mL protein

concentration) in a 1.5 mL microtube. The tube was agitated on a rotator for 3 hours at 4°C. Following incubation, the supernatant was aspirated and the resin washed three times with cold wash buffer containing 50 mM Tris pH 7.8, 150 mM NaCl. For the triplex SILAC studies, each pulldown containing lysate from the resultant channel was aspirated and then washed once separately. The samples from each channel (light, medium, heavy) were then mixed, and washed twice with 1 mL of the wash buffer. Proteins were eluted using an adapted protocol from Masuda et. al. that utilizes a surfactant mixture of 100 mM Tris pH 9, 12 mM sodium deoxycholate, 12 mM sodium lauryl sarcosinate for 5 minutes at room temperature (49). The elution was reduced with 10 mM DTT for 30 min. and then subsequently alkylated with 20 mM CAM for 20 min. at 37°C and 1400 rpm in a ThermoMixer. The sample was diluted with 400 µL of 50 mM ammonium bicarbonate and 1 µg of LysC (Wako) was added. The elution was shaken for 2 hours at 37°C and 1400 rpm. Following this initial digestion, 1 µg of Pierce Trypsin Protease, MS Grade (Thermo Fisher Scientific) was added and the samples were digested overnight on a Thermomixer at 1400 rpm and 37°C. The surfactant was removed by adding TFA to a final concentration of 5% and then extracted with 500 µL of ethyl acetate. The samples were dried down using a speed vac and then resuspended in 5% aq. acetonitrile (ACN) and 0.1% TFA (StageTip Buffer A). From this pool of trypsinized peptide, ten percent was used for protein analysis and the remaining fraction went towards phosphopeptide enrichment. For the protein-based samples, peptides were extracted using StageTips (50) and analyzed in a single nano-LC-MS/MS run.

Phosphopeptide Enrichment with IMAC

Peptides were desalted on Oasis HLB C18, 10 mg cartridges (Waters), dried down using a speed vac, and resuspended in 80% ACN, 19.9% H_2O, 0.1% TFA (StageTip Buffer B). A slurry containing a 1:1:1 mixture of Fe^{3+}-NTA, Ga^{2+}-NTA, and PHOS-Select Iron Affinity Gel (Sigma) was prepared and 20 µL of this mixture was added to each tube of peptide. The slurry was shaken in a ThermoMixer at 25°C for 1 hour and then added directly to a StageTip equilibrated first in methanol and then in Buffer B. The IMAC resin was washed twice with 50 µL of Buffer B and the StageTip equilibrated further with 2 x 100 µL of 1% formic acid. The phosphopetides were eluted from the resin and extracted to the StageTip by addition of 100 µL of 500 mM dibasic sodium phosphate buffer, pH 7.0. This elution was repeated and the StageTip washed twice with 100 µL of 1% formic acid. Phosphopeptide results were analyzed in a single nano-LC-MS/MS run.

LC-MS/MS and Data Analysis

A Thermo-Dionex RSLCnano UHPLC connected to a 20 cm long, 360 µm OD x 100 µm ID fused silica capillary column with a 7 µm tip, self-pulled with a laser puller (Sutter), was used to separate extracted peptides. Columns were packed with 3 µm Reprosil C18 beads (Dr. Maishch). Peptides were run on a 120 min, 5% to 30% acetonitrile gradient in 0.1% acetic acid at 300 nL/min. The LC solvents used were 0.1% acetic acid (A) and 0.1% acetic acid, 99.9% acetonitrile (B). FTMS spectral scans (R = 30 000 at 400 m/z; m/z 350-1600; 3e6 target; max 500 ms ion injection time) were done using a Thermo Orbitrap Elite mass spectrometer with data-dependent analysis using Top15 selection and CID fragmentation (1e4 target; max 100 ms injection time). Collection was done with dynamic exclusion for 30 s and exclusion list size of 50. CID was used with normalized collision energy of 35% for 10 ms. The raw files were analyzed with MaxQuant/Andromeda(58) version version 1.5.7.4 for the label-free SH2 domain profiling and version 1.5.2.8 for the triplex-SILAC studies. Protein, peptide, and site FDRs of 0.01 and a minimum score of 40 for modified peptides, 0 for unmodified peptides, minimum delta score of 17 for modified peptides, and 0 for unmodified peptides was selected. MS/MS spectra were searched against the UniProt human database from July 6, 2016 for the label-free SH2 domain profiling and July 22, 2015 for the triplex SILAC. MaxQuant search parameters were conducted as follows: Variable modifications included Oxidation (M), Acetyl (Protein N-term), Phospho (STY). Carbamidomethyl (C) was a fixed modification. Maximum labeled amino acids was 3, max. missed cleavages was 2, enzyme was Trypsin/P, max charge of 7, multiplicity was either 1, 2, or 3. For the triplex SILAC studies, SILAC labels were Arg0/Lys0 for light, Arg6/Lys4 for medium, and Arg10/Lys8 for heavy. Match between runs was selected and requantification enabled for the SILAC profiling. FTMS scans had an initial search tolerance of 20 ppm and 0.5 Da for ITMS MS/MS scans. Data analysis was conducted using the Perseus software package version 1.5.8.5, Microsoft Excel, Venny (version 2.1) and Graphpad Prism. Significant hits were determined by a two-sample t-test in Perseus using using a permutation-based FDR of 0.05 and S0 of 0.3. LFQ and phosphopeptide intensities were normalized to the median within the population for each respective replicate. The LFQ and phosphopeptide intensities for the control were imputed as its pool of quantified proteins was limited with respective to the functional SH2 domains. For the unstimulated versus stimulated comparisons, LFQ and phosphopeptide intensities were normalized after log2 transformation to the median within each respective replicate. Missing values in both stimulated and non-stimulated were imputed. Interaction networks were generated using data curated from STRING database and visualized using Ctyoscape (version 3.5.1) and PhosphoPath (version 3.1). Sequence logos

were generated using WebLogo 3 (version 2.8.2). Gene ontology enrichment was analyzed using The Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.8 (24, 25).

Western Blot Confirmation

Samples for immunoblot studies were prepared in the same manner as for the MS profiling, boiled in SDS, and then separated by SDS-PAGE. Transfer to nitrocellulose was done using a Trans-Blot Turbo Transfer System (Bio-Rad) for 30 minutes. The blots were blocked with Odyssey Blocking Buffer (Li-Cor) for 1 hr at room temp and incubated with the primary antibody of choice overnight at 4°C. Antibody dilutions were in Blocking Buffer and wash buffer, TBS with 0.1% Tween-20 (v/v). Secondary antibody incubations were performed for 1 hr at room temperature followed by three washes with washing buffer and one wash with TBS. Blots were detected with near-infrared-dye-conjugated secondary antibodies and imaged using an Odyssey IR Imager (Li-Cor). Primary antibodies used and their dilutions are as follows: phosphotyrosine (4G10, 1:2000) (EMD Millipore); phosphotyrosine (APY03, 1:500) (Cytoskeleton); phosphotyrosine (PY100, 1:2000), total EGFR (#4267, 1:2000), phospho-EGFR (Y1068, #2236, 1:2000), FAK (#3285, 1:1000), phospho-FAK (Y576/577, #3281, 1:1000), Paxillin (#12065, 1:1000), phospho-Paxillin (Y118, #2541, 1:1000), Cbl (#2747, 1:1000), CD3ε (#85061, 1:1000), Zap70 (#3165, 1:1000), p38α (#9212, 1:1000) (all from Cell Signaling Technologies).

V. References

1. Hunter, T. (2009) Tyrosine phosphorylation: thirty years and counting. *Curr. Opin. Cell. Biol. 21*, 140-146.

2. Lim, W. A., and Pawson, T. (2010) Phosphotyrosine signaling: evolving a new cellular communication system. *Cell. 142*, 661-667.

3. Kolibaba, K. S., and Druker, B. J. (1997) Protein tyrosine kinases and cancer. *Biochim. Biophys. Acta. 1333*, F217-F248.

4. Robinson, D. R., Wu, Y. M., and Lin, S. F. (2000) The protein tyrosine kinase family of the human genome. *Oncogene. 19*, 5548-5557.

5. Filippakopoulos, P., Muller, S., and Knapp, S. (2009) SH2 domains: modulators of nonreceptor tyrosine kinase activity. *Curr. Opin. Struct. Biol. 19*, 643-649.

6. Songyang, Z., and Cantley, L. C. (1995) Recognition and specificity in protein tyrosine kinase-mediated signaling. *Trends Biochem. Sci. 20*, 470-475.

7. Wavreille, A. -S., Garaud, M., Zhang, Y., and Pei, D. (2007) Defining SH2 domain and PTP specificity by screening combinatorial peptide libraries. *Methods. 42*, 207-219.

8. Kaneko, T., Huang, H., Zhao, B., Li, L., Liu, H., Voss, C. K., Wu, C., Schiller, M. R., and Li, S. S. C. (2010) Loops govern SH2 domain specificity by controlling access to binding pockets. *Sci. Signal. 3*, Ra34.

9. Songyang, Z., Shoelson, S. E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnofksy, S., Lechleider, R. J., Neel, B. G., Birge, R. B., Fajardo, J. E., Chou, M. M., Hanafusa, H., Schaffhausen, B., and Cantley, L. C. (1993) SH2 Domains Recognize Specific Phosphopeptide Sequences. *Cell. 72*, 767-778.

10. Liu, B. A., Jablonowski, K., Shah, E. E., Engelmann, B. W., Jones, R. B., and Nash, P. D. (2010) SH2 domains recognize contextual peptide sequence information to determine selectivity. *Mol. Cell. Proteomics. 9*, 2391-2404.

11. Tinti, M., Kiemer, L., Costa, S., Miller, M. L., Sacco, F., Olsen, J. V., Carducci, M., Paoluzi, S., Langone, F., Workman, C. T., Blom, N., Machida, K., Thompson, C. M., Schutkowski, M., Brunak, S., Mann, M., Mayer, B. J., Castagnoli, L., and Cesareni, G. (2013) The SH2 domain interaction landscape. *Cell Rep. 3*, 1293-1305.

12. Frank, R. (1992) Spot-synthesis: an easy technique for the positionally addressable, parallel chemical synthesis on a membrane support. *Tetrahedron. 48*, 9217-9232.

13. Machida, K. Thompson, C. M., Dierck, K., Jablonowski, K., Kärkkäinen, S., Liu, B., Zhang, H., Nash, P. D., Newman, D. K., Nollau, P., Pawson, T., Renkema, G. H., Saksela, K., Schiller, M. R., Shin, D. -G., and Mayer, B. J. (2007) High-Throughput Phosphotyrosine Profiling Using SH2 Domains. *Mol. Cell. 26*, 899-915.

14. Tinti, M., Nardozza, A. P., Ferrari, E., Sacco, F., Corallino, S., Castagnoli, L., and Cesareni, G. (2012) The 4G10, phosphotyrosine20, and p-TYR-100 antibody specificity: profiling by peptide microarrays. *N. Biotechnol. 29,* 571-577.

15. Bergman, M., Mustelin, T., Oetken, C., Partanen, J., Flint, N. A., Amrein, K. E., Autero, M., Burn, P., and Alitalo, K. (1992) The human p50csk tyrosine kinase phosphorylates p56lck at Tyr-505 and down regulates its catalytic activity. *EMBO J. 11*, 2919-2924.

16. Davidson, D., Bakinowski, M., Thomas, M. L., Horejsi, V., and Veilette, A. (2003) Phosphorylation-dependent regulation of T-cell activation by PAG/Cbp, a lipid raft-associated transmembrane adaptor. *Mol. Cell. Biol. 23*, 2017-2028.

17. Visser Smit, G. D., Place, T. L., Cole, S. L., Clausen, K. A., Vemuganti, S., Zhang, G., Koland, J. G., and Lill, N. L. (2009) Cbl controls EGFR fate by regulating early endosome fusion. *Sci. Signal. 2*, ra86.

18. Grovdal, L. M., Stang, E., Sorkin, A., and Madshus, I. H. (2004) Direct interaction of Cbl with pTyr 1045 of the EGF receptor (EGFR) is required to sort the EGFR to lysosomes for degradation. *Exp. Cell. Res. 300*, 388-395.

19. Ettenberg, S. A., Keane, M. M., Nau, M. M., Frankel, M. F., Wang, L. -M., Pierce, J. H., and Lipkowitz, S. (1999) *cbl*-b inhibits epidermal growth factor receptor signaling. *Oncogene. 18*, 1855-1866.

20. Long, W., Yi, P., Amazit, L., LaMarca, H. L., Ashcroft, F., Kumar, R., Mancini, M. A., Tsai, S. Y., Tsai, M. J., and O'Malley, B. W. (2010) SRC-3Delta4 mediates the interaction of EGFR with FAK to promote cell migration. *Mol. Cell. 37*, 321-332.

21. Zeng, L., Sachdev, P., Yan, L., Chan, J. L., Trenkle, T., McClelland, M., Welsh, J., and Wang, L. (2000) Vav3 mediates receptor protein tyrosine kinase signaling, regulates GTPase activity, modulates cell morphology, and induces transformation. *Mol. Cell. Biol. 20*, 9212-9224.

22. Kang, Y. S., Kim, W., Huh, Y. H., Bae, J., Kim, J. S., and Song, W. K. (2011) P130Cas attenuates epidermal growth factor (EGF) receptor internalization by modulating EGF-triggered dynamin phosphorylation.

23. Wu, D. W., Chen, C. Y., Chu, C. L., and Lee, H. (2016) Paxillin confers resistance to tyrosine kinase inhibitors in EGFR-mutant lung cancers via modulating BIM and Mcl-1 protein stability. *Oncogene. 35*, 621-630.

24. Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009) Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. *Nature Protoc. 4*, 44-57.

25. Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res. 37*, 1-13.

26. Fischer, A., Picard, C., Chemin, K., Dogniaux, S., le Deist, F., and Hivroz, C. (2010) ZAP70: a master regulator of adaptive immunity. *Semin. Immunopathol. 32*, 107-116.

27. Gelkop, S., Gish, G. D., Babichev, Y., Pawson, T., and Isakov, N. (2005) T cell activation-induced CrkII binding to the Zap70 protein tyrosine kinase is mediated by Lck-dependent phosphorylation of Zap70 tyrosine 315. *J. Immunol. 175*, 8123-8132.

28. Wiede, F., Shields, B. J., Chew, S. H., Kyparissoudis, K., van Vliet, C., Galic, S., Tremblay, M. L., Russell, S. M., Godfrey, D. I., and Tiganis, T. (2011) T cell protein tyrosine phosphatase attenuates T cell signaling to maintain tolerance in mice. *J. Clin. Invest. 121*, 4758-4774.

29. Yasuda, T., Bundo, K., Hino, A., Honda, K., Inoue, A., Shirakata, M., Osawa, M., Tamura, T., Nariuchi, H., Oda, H., Yamamoto, T., and Yamanahi, Y. (2007) Dok-1 and Dok-2 are negative regulators of T cell receptor signaling. *Int. Immunol. 19*, 487-495.

30. Marie-Cardine, A., Kirchgessner, H., Bruyns, E., Shevchenko, A., Mann, M., Autschbach, F., Ratnofsky, S., Meuer, S., and Schraven, B. (1999) SHP2-interacting transmembrane adaptor protein (SIT), a novel disulfide-linked dimer regulating human T cell activation. *J. Exp. Med. 189*, 1181-1194.

31. Brownlie, R. J., and Zamoyska, R. (2013) T cell receptor signaling networks: branched, diversified and bounded. *Nat. Rev. Immunol. 13*, 257-269.

32. Laurent, C. E., Delfino, F. J., Cheng, H. J., and Smithgall, T. E. (2004) The human c-Fes tyrosine kinase binds tubulin and microtubules through separate domains and promotes microtubule assembly. *Mol. Cell. Biol. 24*, 9351-9358.

33. McPherson, V. A., Everingham, S., Karisch, R., Smith, J. A., Udell, C. M., Zheng. J., Jia, Z., and Craig, A.W. (2009) Contributions of F-BAR and SH2 domains of Fes protein tyrosine kinase for coupling to the FcepsilonRI pathway in mast cells. *Mol. Cell. Biol. 2*, 389-401.

34. Jones, R. B.; Gordus, A.; Krall, J. A.; and MacBeath, G. (2006) A quantitative protein interaction network for the ErbB receptors using protein microarrays. *Nature*. *439*, 168-174.

35. Chung, B. M., Dimri, M., George, M., Reddi, A. L., Chen, G., Band, v., and Band, H. (2009) The role of cooperativity with Src in oncogenic transformation mediated by non-small cell lung cancer-associated EGF receptor mutants. *Oncogene. 28*, 1821-1832.

36. Wu, W., Graves, L. M., Gill, G. N., Parsons, S. J., and Samet, J. M. (2002) Src-dependent phosphorylation of the epidermal growth factor receptor on tyrosine 845 is required for zinc-induced Ras activation. *J. Biol. Chem. 277*, 24252-24257.

37. Mueller, K. L., Hunter, L. A., Ethier, S. P., and Boerner, J. L. (2008) Met and c-Src cooperate to compensate for loss of epidermal growth factor receptor kinase activity in breast cancer cells. *Cancer Res. 68*, 3314-3322.

38. Kamalati, T., Jolin, H. E., Mitchell, P. J., Barker, K. T., Jackson, L. E., Dean, C. J., Page, M. J. Gusterson, B. A., and Crompton, M. R. (1996) Brk, a breat tumor-derived non-receptor protein-tyrosine kinase, sensitizes mammary epithelial cells to epidermal growth factor. *J. Biol. Chem. 271*, 30956-30963.

39. Chen, H. Y., Shen, C. H., Tsai, Y. T., Lin, F. C., Huang, Y. P., and Chen, R. H. (2004) Brk activates rac1 and promotes cell migration and invasion by phosphorylating paxillin. *Mol. Cell. Biol. 24*, 10558-10572.

40. Li, X., Lu, Y., Liang, K., Hsu, J. M., Albarracin, C., Mills, G. B., Hung, M. C., and Fan, Z. (2012) Brk/PTK6 sustains activated EGFR signaling through inhibiting EGFR degradation and transactivating EGFR. *Oncogene. 31*, 4372-4383.

41. Kasprzycka, M., Majewski, M., Wang, Z., Ptasznik, A., Wysocka, M., Zhang, Q., Marzec, M., Gimotty, P., Crompton, M. R., and Wasik, M. A. (2006) Expression and Oncogenic Role of Brk (PTK6/Sik) Protein Tyrosine Kinase in Lymphocytes. *Am. J. Pathol. 168*, 1631-1641.

42. Derry, J. J., Richard, S., Valderamma Caravajal, H., Ye, X., Vasioukhin, V., Cochrane, A. W., Chen, T., and Tyner, A. L. (2000) Sik (BRK) phosphorylates Sam68 in the nucleus and negatively regulates its RNA binding ability. *Mol. Cell. Biol. 20*, 6114-6126.

43. Huot, M. E., Vogel, G., and Richard, S. (2009) Identification of a Sam68 ribonucleoprotein complex regulated by epidermal growth factor. *J. Biol. Chem. 284*, 31903-31913.

44. Fusaki, N., Iwamatsu, A., Iwashima, M., and Fujisawa, J. (1997) Interaction between Sam68 and Src family kinases, Fyn and Lck, in T cell receptor signaling. *J. Biol. Chem. 272*, 6214-6219.

45. Kaneko, T., Huang, H., Xuan, C., Xing, L., Li, X., Chengjun, L., Voss, C., Sachdev., S. S., and Li, S. S. (2012) Superbinder SH2 domains act as agonists of cell signaling. *Sci. Signal. 5*, ra68.

46. Bian, Y., Li, L., Dong, M., Liu, X., Kaneko, T., Cheng, K., Liu, H., Voss, C., Cao, X., Wang, Y., Litchfield, D., Ye, M., Li, S. S., and Zou, H. (2016) Ultra-deep tyrosine phosphoproteomics enabled by a phosphotyrosine superbinder. *Nat. Chem. Biol. 12*, 959-966.

47. Yadav, S. S., Yeh, B. J., Craddock, B. P., Lim, W. A., and Miller, W. T. (2009) Reengineering the Signaling Properties of a Src Family Kinase. *Biochemistry, 48*, 10956-10962.

48. Hill, Z. B., Perera, B. G. K., Andrews, S. S., and Maly, D. J. (2012) Targeting diverse signaling interaction sites allows the rapid generation of bivalent kinase inhibitors. *ACS Chem. Biol. 7,* 487-495.

49. Masuda, T., Tomita, M., and Ishihama, Y. (2008) Phase transfer surfactant-aided trypsin digestion for membrane proteome analysis. *J. Proteome Res. 7*, 731-740.

50. Rappsilber, J., Ishihama, Y., and Mann, M. (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem. 75*, 663-670.