



Published in final edited form as:

J Proteome Res. 2008 September ; 7(9): 3900–3910. doi:10.1021/pr800198w.

Identification of c-Src tyrosine kinase substrates using mass spectrometry and peptide microarrays

Ramars Amanchy¹, Jun Zhong¹, Henrik Molina^{1,2}, Chaerkady Raghothama¹, Akiko Iwahori¹, Dario Eluan Kalume^{1,5}, Mads Grønborg¹, Jos Joore⁵, Leslie Cope⁴, and Akhilesh Pandey^{1,6}

¹McKusick-Nathans Institute of Genetic Medicine and the Departments of Biological Chemistry, Pathology and Oncology, Johns Hopkins University, Baltimore, Maryland 21205, USA ²Department of Biochemistry and Molecular Biology, University of Southern Denmark, Center for Experimental Bioinformatics, Odense 5230, Denmark ³Pepsican Systems, Edelhertweg 15, 8219 PH Lelystad, The Netherlands ⁴Sidney Kimmel Comprehensive Cancer Center and the Department of Biostatistics, Bloomberg School of Public Health, and Johns Hopkins University, Baltimore, Maryland, USA ⁵Current address: Department of Tropical Medicine, Oswaldo Cruz Foundation - FIOCRUZ, Rio de Janeiro, RJ, 21040-900, Brazil

Summary

c-Src tyrosine kinase plays a critical role in signal transduction downstream of growth factor receptors, integrins and G protein-coupled receptors. We used stable isotope labeling with amino acids in cell culture (SILAC) approach to identify additional substrates of c-Src tyrosine kinase in human embryonic kidney 293T cells. We have identified 10 known substrates and interactors of c-Src and Src family kinases along with 26 novel substrates. We have experimentally validated 4 of the novel proteins (NICE-4, RNA binding motif 10, FUSE-binding protein 1 and TRK-fused gene) as direct substrates of c-Src using *in vitro* kinase assays and cotransfection experiments. Significantly, using a c-Src specific inhibitor, we were also able to implicate 3 novel substrates (RNA binding motif 10, EWS1 and Bcl-2 associated transcription factor) in PDGF signaling. Finally, to identify the exact tyrosine residues that are phosphorylated by c-Src on the novel c-Src substrates, we designed custom peptide microarrays containing all possible tyrosine-containing peptides (312 unique peptides) and their mutant counterparts containing a Tyr→Phe substitution from 14 of the identified substrates. Using this platform, we identified 34 peptides that are phosphorylated by c-Src. We have demonstrated that SILAC based quantitative proteomics approach is suitable for identification of substrates of non-receptor tyrosine kinases and can be coupled to peptide microarrays for high-throughput identification of substrate phosphopeptides.

Introduction

Most signaling pathways include protein kinases and their substrates that serve as means to amplify signals from extracellular signals and other stimuli. However, the precise connectivity between protein kinases and their downstream substrates has not been fully elucidated for most protein kinases. c-Src is a classic non-receptor tyrosine kinase that has been implicated in regulation of cytoskeletal rearrangement and cell adhesion networks that control cell migration, cell proliferation and cell survival (1). One vital step in understanding the role of c-Src kinase in cellular transformation and signaling is systematic identification of all of its potential cellular substrates involved in these processes.

⁶Corresponding author: Tel: 410-502-6662, FAX: 410-502-7544, E-mail: pandey@jhmi.edu..

Recent studies based on advances in mass spectrometry-based proteomics have provided large-scale catalogs of phosphorylation sites (2-5). However, determination of kinases responsible for these phosphorylation events is not an easy task owing to the transient interaction between kinases and their substrates. Chemical and genetic approaches have been previously used to identify c-Src substrates. Such studies include use of an ATP analog that is a specific substrate for an analog-specific allele of v-Src (6), screening of cDNA expression libraries with anti-phosphotyrosine antibodies (7) and use of mutant inducible forms of c-Src (8). To date, several c-Src substrates as well as interactors have been reported. Human Protein Reference Database (HPRD) (9) provides a list of 132 c-Src mediated phosphorylation sites in 64 known substrates along with 204 proteins that interact with c-Src.

We have used the SILAC approach which enables identification of tyrosine kinase substrates based on a unique signature in mass spectrometry experiments (10-12). The main objective of this work was to identify novel c-Src substrates by overexpression of a constitutively active form of c-Src followed by enrichment of tyrosine phosphorylated proteins. We have identified 26 novel c-Src tyrosine kinase substrates in addition to 10 others, which were either known Src family kinase substrates or proteins known to associate with Src family kinases. We have experimentally confirmed 4 novel substrates, NICE-4, RNA binding motif 10, FUSE-binding protein 1 and TRK-fused gene, to be direct substrates of c-Src using *in vitro* kinase assays. We were also able to implicate EWS1, RNA binding motif 10 and Bcl-2 associated transcription factor in PDGF signaling using a chemical inhibitor of c-Src. Our peptide microarray approach led to identification of a number of peptides that are phosphorylated by c-Src. To our knowledge, this is the first reported integrated proteomics strategy that couples cell culture, mass spectrometry and peptide microarrays to identify tyrosine kinase substrates.

Experimental procedures

Chemicals and Antibodies

Stable isotope containing amino acids, $^{12}\text{C}_6$ -arginine, $^{13}\text{C}_6$ -arginine and $^{13}\text{C}_6$ - $^{15}\text{N}_4$ -arginine, were purchased from Cambridge Isotope Labs (Andover, MA). Complete protease inhibitor cocktail tablets were purchased from Roche (Indianapolis, IN), sodium orthovanadate and anti-Flag M2 monoclonal antibody from Sigma-Aldrich Co (St. Louis, MO), SU6656 from EMD Biosciences, Inc. (San Diego, CA), anti-phosphotyrosine antibodies (4G10) agarose-conjugate and streptavidin-agarose beads from Upstate Biotechnology (Lake Placid, NY), antiphosphotyrosine-RC20 biotin conjugate from BD Biosciences (San Jose, CA) and PDGF-BB from Invitrogen (Carlsbad, CA). Sequencing grade trypsin was purchased from Promega (Madison, WI). Antibodies against cortactin were purchased from Upstate USA, Inc. (Chicago, IL), p130CAS and EWS1 from Santa Cruz Biotechnology, Inc. (Santa Cruz, CA), BTF from Bethyl, Inc. (Montgomery, TX), and RBM10 was from Abcam Inc. (Cambridge, MA). Phospho-Src (Tyr416) antibody and PhosphoScan Kit (P-Tyr-100) was purchased from cell signaling technology (Boston, MA).

Cell culture and Stable isotope labeling with amino acid in cell culture (SILAC)

Human embryonic kidney 293T cells were grown in Dulbecco's modified Eagle's medium containing 'light', 'medium' or 'heavy' arginine supplemented with 10% dialyzed fetal bovine serum plus antibiotics. The 293T cells were adapted to growing in isotope rich-medium supplemented with dialyzed serum prior to initiating these experiments. In each experiment, twenty 10 cm dishes were used per condition and the cells were transfected with 15 μg of DNA using the standard calcium phosphate method (Invitrogen, Carlsbad, CA). 6 h after transfection, the cells were serum starved for 10 h or 20 h. After starvation, the cells were lysed in modified RIPA buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1mM EDTA, 1% Nonidet P-40, 0.25% sodium deoxycholate, and 1 mM sodium orthovanadate in the presence of protease inhibitors).

Upon cell lysis proteins lysates were either subjected to affinity purification of tyrosine phosphorylated proteins (13) or peptides containing phosphotyrosine were enriched directly from trypsin-digested cell lysates (14) using specific antibodies against phosphotyrosine and are identified by tandem mass spectrometry.

Immunoprecipitation and Western blotting

Light, medium and heavy cell lysates were precleared with protein A-agarose, mixed and incubated with 400 µg of 4G10 monoclonal antibodies coupled to agarose beads, 75 µg of biotin-conjugated RC20 antibody and streptavidinagarose beads overnight at 4°C. Precipitated immune complexes were then washed three times with lysis buffer. Agarose beads were boiled and resolved by 10% SDS-PAGE. The gel was silver stained for visualizing protein bands. Western blotting experiments were performed using anti-phosphotyrosine antibody (4G10) and reprobing was carried out using anti-Flag antibody.

Cloning and Transfection

NICE-4 protein (NP_055662), RNA Binding Motif protein 10 isoform 1 (NP_05667) and Far upstream element-binding protein (NP_003893), TRK-Fused gene (NP_006061) were subcloned into a Flag epitope-tagged mammalian expression vector, pCMVtag4A. 293T cells were grown in 10 cm dishes. One dish was co-transfected wild c-Src and pCMVtag4A vector as control; one was co-transfected wild type c-Src and Flag-tagged cDNAs. The expressed proteins were immunoprecipitated using anti-Flag antibody, followed by SDS-PAGE and Western blotting. The blots were probed with anti-phosphotyrosine antibody followed by stripping and re-probing with anti-Flag antibodies.

In vitro kinase assays using GST-fusion proteins

Fusion proteins were made by using TNT-coupled rabbit reticulocyte lysate system (Promega, Madison, WI) with the cDNAs cloned in GST expression vector, PGEX4T1. The *in vitro* translated GST-tagged proteins were purified with 10 µg of GST beads for 12 h at 4 °C. After incubation, the beads were washed two times in lysis buffer and two times in kinase buffer (20 mM Hepes, pH 7.4, 5 mM MgCl₂, 2 mM MnCl₂, 50 µM sodium vanadate, 50 µM DTT). Immune complexes were incubated for 30 min at 30 °C in 5 µl of ATP mixture (10 µM cold ATP and 10 µCi of [γ -³²P] ATP) and c-Src Kinase. Protein samples were then eluted by boiling in sample buffer and resolved by SDS-PAGE. The gel was dried and exposed to x-ray film to visualize the ³²P-labeled protein bands.

PDGF stimulation and inhibition of Src kinase

NIH3T3 cells were grown in DMEM containing 10% FBS supplemented with antibiotics. For all PDGF stimulation experiments, cells were stimulated with 100 ng / ml of PDGF-BB for 5 min. Cells were treated with 2 µM c-Src kinase inhibitor, SU6656 for 1 hour prior to stimulation of cells with PDGF-BB for 5 min.

In-gel trypsin digestion and in solution trypsin digestion

The silver stained protein bands were excised and in-gel trypsin digestion was performed as described previously (15). Briefly, the gel slices were excised, and incubated with trypsin overnight at 37°C to allow digestion of proteins after a reduction and alkylation step. After in-gel digestion, the tryptic peptides were extracted. The supernatants from the in-gel and in solution trypsin digestion containing the peptide mixture were partially dried down in a vacufuge to approximately 10 µL. For in solution digestion and enrichment of phosphopeptides phosphoscan kit was used according to manufacturer's prescribed conditions.

Liquid chromatography-mass spectrometry

The extracted peptide mixture was centrifuged for 2 min, $12\,000 \times g$ at $4\text{ }^{\circ}\text{C}$ and resolved by reversed phase liquid chromatography on Agilent 1100 Series LC system (Agilent Technologies, Palo Alto, CA) equipped with a well plate sampler, a vacuum degasser, and a capillary pump. Each fraction from the digested peptide mixture was analyzed by automated nanoflow LC-MS/MS. An Agilent Technologies 1100 series system was used to deliver a flow of $1.5\text{ }\mu\text{l}/\text{min}$ during desalting of the sample and $250\text{ nl}/\text{min}$ during elution of the peptides into the mass spectrometer as described before (10). Each sample was loaded onto an on-line analytical fused silica needle column (Proxion Biosystems, Odense, Denmark) packed with $5\text{-}\mu\text{m}$ Vydac C18 resin. Washing and desalting was done with 95% mobile phase A (H_2O with 0.4% acetic acid and 0.005% heptafluorobutyric (v/v)) and 5% mobile phase B (90% acetonitrile, 0.4% acetic acid, 0.005% heptafluorobutyric acid in water). Samples were eluted from the analytical column by a linear gradient of 90% mobile phase A to 60% mobile phase A. A 34-min gradient was used for elution. A potential of 2.8 kV was applied to the emitter (Proxion Biosystems). The spectra were acquired on a quadrupole time-of-flight mass spectrometer (Q-TOF US-API, Micromass, Manchester, UK) equipped with an ion source sample introduction system designed by Proxeon Biosystems (Odense, Denmark). Data were obtained in positive ion mode. Data-dependent acquisition was performed with an ion mass window to 2.5 Da. MS to MS/MS switch was set to a threshold of 10 counts/s, and MS/MS to MS was set to an intensity below a threshold of 2 counts/s. Charge state recognition was used to estimate the collision energy for the fragmented precursor. Scan time was set to 0.9 s, and interscan time was set to 0.1 s. The number of components (i.e. number of MS/MS per MS scan) was set to three resulting in a total cycle time (one MS and three MS/MS spectra) of 10 s. The acquisition of data was performed using MassLynx (version 4.0). The parameters used for generating peak lists from the raw data were the following: smooth window, 4.00; number of smooth, 2; smooth mode, Savitzky Golay; and percentage of peak height to calculate centroid spectra, 80% with no base-line subtraction. The generated peak lists (pkl-file) were searched against the RefSeq human protein database (build 33, 29572 sequences) (www.ncbi.nlm.nih.gov/RefSeq/) using Mascot version 2.0, with a mass accuracy of 1.1 Da for the parent ion (MS) and 0.2 for the fragment ions (MS/MS), allowing a maximum of two missed cleavages. Carbamidomethylation of cysteines was considered as fixed modification, and oxidations of methionine residues, “medium” arginine (+6 Da), “heavy” arginine (+10 Da), and phosphorylation of tyrosine residue were considered as variable modifications. An initial protein list was generated using the following criteria. Only proteins containing at least one unique peptide (if the sequence has not been assigned to a different protein) with a Mascot score over 30, were considered in the dataset. The sequence of higher scoring peptides was manually verified. Quantitation was performed on three to four peptides (wherever available) by comparing the extracted ion chromatogram of the corresponding light and heavy peptides using MS-Quant (16). Reproducibility of measurements was performed by using two analysis of variance models as described earlier (17). The tandem mass spectra were manually verified to assign the sequence and phosphorylation sites for all peptides identified in this study (Supplementary Table 1). The phosphorylation sites were manually verified and assigned after confirmation of a mass difference of 243 Da corresponding to phosphotyrosine residue.

Peptide microarrays and data analysis

The WT (peptides containing a tyrosine residue in the center) and MUT (peptides where the central tyrosine residue was replaced by phenylalanine) peptides (Supplementary Table 2) were each spotted as triplicates on glass slides (Pepscan Systems, Lelystad, Netherlands) as described earlier (18). c-Src kinase assays were carried out using these custom peptide microarrays by incubating 50 ng of recombinant c-Src Kinase (Invitrogen, Carlsbad, CA) in kinase reaction buffer and $300\text{ }\mu\text{Ci}/\text{ml}$ γ33P -labeled ATP (AH9968; GE Healthcare Biosciences Corp., Piscataway, NJ) at $25\text{ }^{\circ}\text{C}$ for 1 h in a $120\text{ }\mu\text{l}$ reaction volume supplemented with

200 μM ATP. The reaction was stopped and the following washing steps were performed; 2 washes in 2M sodium chloride containing 1% Triton X-100 followed by 3 washes in phosphate buffered saline containing Triton X-100 and 1 wash in distilled water. The glass slides were then air dried and exposed to the phosphorimager screen for 12 hours and scanned using Biorad Molecular Imager FX (Bio-Rad Laboratories, Inc, Hercules, CA). The image was processed using GenePix Pro 6.0 software (Molecular Devices Corporation, Sunnyvale, CA). Autoradiographs were obtained by using a phosphorimager screen. The assay was performed in triplicate. The intensity values obtained were transformed to the log base 2 scale. Effects on intensity due to the position of the spot on the slide were estimated by performing a local regression analysis (loess) with respect to chip coordinates, and subtracted out (19). Normalized log₂ intensities for triplicate spots were averaged, and the mean log₂ MUT intensity was subtracted from the corresponding mean log₂ WT intensity for each peptide. The resulting background adjusted values were averaged over replicate arrays. The mean background adjusted log intensity for each peptide on the Y-axis, and the average of WT and MUT log intensities on the X-axis were plotted to estimate the distribution of the intensity values arising from the phosphorylated peptides. A key assumption for selection of positive (phosphorylated) peptides is that WT peptide intensity values are greater than MUT peptide intensities. It is also assumed that higher intensity peptides are more likely to be positive. We note that consistent with this assumption, the WT intensity is consistently higher than MUT intensity for the high intensity peptides on the right side of the plot. Likewise, there is greater symmetry on the left side quadrant of the MvA plot, where we expect non-phosphorylated peptides to have WT intensities that are as likely to be lower than MUT values as higher. The classical False Positive Rate (expected error rate for a set of points) is derived by evaluating symmetry in the Y-axis. For each value of A, the classical FPR blue curve gives: (number of points below the X-axis to the right of A) / (number of points above the X-axis to the right of A), and describes the expected number of false positives in the upper right quadrant of the plot. Sometimes it is desirable to estimate the probability that a single given point is a false positive, allowing us to move the threshold to the left until the price of adding one more peptide is too high. This is described by the local false positive rate curve. The local FPR curve gives the probability of a positive peptide located at A being a false positive. We selected all peptides where the local false positive rate was lower than 0.15 and WT intensities were two fold greater compared to MUT. The resulting set has an overall FPR of 0.085.

Results and discussion

Stable isotope labeling of cellular proteins for the identification of c-Src kinase substrates

Stable isotope labeling with amino acids in cell culture (SILAC) involves metabolic labeling of cellular proteomes by growing the cells in media containing amino acids labeled with stable isotopes. SILAC enables identification of peptides labeled *in vivo* and relative quantitation of abundance of the peptides arising out of a mixture of labeled and unlabeled protein samples (15). This method also allows one to distinguish contaminating proteins in immunoprecipitates that arise due to non-specific binding.

Because of the complexity of cell lysates, and because kinase substrates generally exhibit low stoichiometry of tyrosine phosphorylation, specific identification of tyrosine kinase substrates can be facilitated by prior enrichment of tyrosine-phosphorylated proteins with anti-phosphotyrosine antibodies. We applied SILAC for the identification of the c-Src kinase substrates in human embryonic kidney cells by overexpression of a constitutively active form of c-Src followed by affinity purification of tyrosine phosphorylated proteins. We transiently overexpressed either a kinase inactive c-Src (K298M) as a control, or, a constitutively active c-Src (Y527F) kinase in 293T cells. Phosphorylation of the C-terminal tyrosine by C-terminal Src Kinase (CSK) allows inactivation of c-Src (20-22). Hence mutation of this tyrosine residue to phenylalanine allows c-Src kinase to be constitutively active by preventing its folding and

by allowing the kinase domain access to its substrates (23,24). Inhibition of c-Src activity is often achieved by co-expression of the c-Src-inactivating C-terminal Src Kinase (CSK), the kinase-inactive Src mutant Src K298M (25), or by treatment of the cells with c-Src inhibitors.

Three populations of human embryonic kidney cells were grown in Dulbecco's modified eagle's media containing, $^{12}\text{C}_6$ arginine (light), $^{13}\text{C}_6$ -arginine (medium) and $^{13}\text{C}_6$ - $^{15}\text{N}_4$ -arginine (heavy), respectively (Figure 1). The cells grown in light medium were transfected with a kinase inactive c-Src as a negative control (26). Cells grown in medium and heavy isotope containing media were transfected with a constitutively active form of c-Src (Y527F) and harvested at 12 h or 24 h post-transfection, respectively. We found an increased tyrosine phosphorylation upon transfection of constitutively active form of c-Src for 24h (Figure 2). We have also observed that the Y416 in c-Src gets phosphorylated more in this state which points to the increased c-Src tyrosine kinase activity (Figure 2). In addition, the trend of increasing tyrosine phosphorylation could also serve as one more surrogate signature of substrates as one would expect the phosphorylation level to increase during this time course.

Mixing of light, medium and heavy isotope labeled cell lysates allowed us to compare the profile of proteins in a single MS experiment. In MS/MS spectra, fragmentation patterns generated by light, medium and heavy peptide pairs are identical except for the expected mass shift of the fragment ions. The ratio of the intensity of the heavier versus the light peptides provides information about the degree of phosphorylation of a protein and hence its enrichment upon expression of an active c-Src kinase. Thus, the greater the extent of phosphorylation of a protein by c-Src kinase, the higher should be its abundance in anti-phosphotyrosine antibody immunoprecipitates. Peptide sets with a little or no increase in intensity indicate that the protein is not different in abundance in the different states being compared. Such proteins were not investigated further as they are likely nonspecifically bound proteins. An increase in heavy/light intensity ratio, indicating an increase in total phosphotyrosine content upon Src kinase expression and activity, was found in peptides derived from 36 proteins (Table 1 and 2). Of these, 10 proteins were either known Src family kinase substrates or proteins known to interact with Src family kinases (Table 1) whereas the remaining 26 proteins have not previously been described as substrates of c-Src or Src family members (Table 2) in higher eukaryotes. The known substrates identified in this screen included EWS1 (Ewing sarcoma breakpoint region 1) (27), cortactin (28), calponin-3 (29), hnRNP-K (Heterogeneous nuclear ribonucleoprotein K) (30,31), G3BP (RasGAP SH3-domain binding protein) (32-34) and c-Src itself (35). The protein with maximum increase in tyrosine phosphorylation upon c-Src overexpression was c-Src itself. Other known and novel Src family substrates displayed >2 fold increase in intensity of phosphorylation. Apart from signaling and cytoskeletal proteins we also identified DNA and RNA binding proteins in our analysis. Some of the reasons why it is not possible in a kinase-substrate identification screen of this type to possibly identify every known substrate are: i) Previously described substrates might not be expressed in the cell line that we have used, ii) Although the substrates might be expressed, they might not be abundant enough to be enriched and detected in our experiments, iii) The phosphorylation might not occur or occur at a lower level in the cells that we have used; and iv) The time course and kinetics of phosphorylation in the system that we have employed might be different from the systems previously used in the literature to describe substrates. Thus, although it is not possible to identify all of the known substrates of Src, identification of 6 known substrates of Src along with validation of some novel ones is indicative of the success of this type of phosphoproteomic screen.

We performed relative quantitation of tyrosine phosphorylation for each protein as a measure of increase in intensity ratios from light isotope to heavier isotope containing peptides (Supplementary Table 1). Figure 3A-F shows representative MS spectra for six of the proteins identified from our screen. In addition to identification of proteins with increased

phosphorylation and quantitation, we also mapped 6 phosphorylation sites Y334 in cortactin (10) (Figure 4A) and Y72 in hnRNP-K (30) (Figure 4B) identified in this study were reported earlier. We also identified 4 novel tyrosine phosphorylation sites on each on EWS1 (Fig 4C), FUSE-binding protein 1 (Fig 4D), calponin-3 (Fig 4E) and FIP1-like1 (Fig 4F). We note that although EWS1 was a known c-Src substrate, no tyrosine phosphorylation sites were previously localized.

One of the drawbacks of our experiments is that lysine was not available as 3 different isotopic forms at the time we initiated our experiments. By using 3 isotopes of lysine along with arginine, we would have obtained a better peptide coverage for each protein and likely identified additional proteins as substrates of c-Src. Nevertheless, we performed a lysine and arginine labeled SILAC experiment in a similar manner but followed by enrichment of phosphopeptides using antibodies against phosphotyrosine to see if we could identify phosphorylated peptides, but unfortunately, we could only identify 8 tyrosine phosphorylation sites (data not shown). We do not know the reasons for such a low yield. However, a protein IP serves our purpose of identifying the proteins that are substrates of Src even though the site is still not identified. This is why we coupled our approach with peptide microarrays to aid in identification of phosphopeptides.

Validation of a subset of novel c-Src substrates

Further validation of the proteins identified by SILAC to prove that they are *bona fide* substrates usually involves the use of antibodies against these proteins. The validation of all of the protein candidates is not always possible, especially for novel proteins, as it depends on the availability of good antibodies. As commercial antibodies were not available for many of the proteins identified, we chose to investigate a subset of proteins, if they were direct substrates of c-Src using *in vitro* kinase assays. We selected NICE-4, RBM10, FBP1 and TRK-fused gene for this purpose, as their cDNAs were readily available. We used rabbit reticulocytes to perform *in vitro* transcription and translation reactions. These proteins were purified and incubated along with c-Src kinase to investigate if it could phosphorylate these proteins. After incubation with c-Src kinase, the proteins were resolved by SDS-PAGE and autoradiographs were obtained. We found that all of the tested proteins were tyrosine phosphorylated upon incubation with c-Src kinase (Figure 5A). TRK-fused gene from *Xenopus laevis* has been shown to interact with SH3 domains of various proteins including v-Src but did not bind to neuronal specific Src *in vitro* (36).

We verified if the above c-Src substrates were also substrates of c-Src *in vivo* by cotransfecting these proteins with wild type c-Src in 293T cells. We also subcloned these cDNAs into a Flag epitope-tagged vector, pCMVtag4A, and cotransfected 293T cells with wild type c-Src kinase or with an empty vector. The proteins were immunoprecipitated using anti-Flag antibodies, resolved by SDS-PAGE, immunoblotted with anti-phosphotyrosine antibodies. Upon cotransfection with c-Src kinase, we again observed increased tyrosine phosphorylation of all the tested proteins suggesting that these proteins were also *in vivo* substrates of c-Src (Figure 5B). We have previously identified NICE-4 as a tyrosine phosphorylated protein in a global phosphoproteomic study of HeLa cells (10). RBM10, FBP1 and TRK-fused gene are novel tyrosine phosphorylated proteins and further investigations need to be carried out to determine how they transduce signals downstream of c-Src kinase and if tyrosine phosphorylation regulates this process.

Involvement of a subset of novel substrates in platelet-derived growth factor signaling

Since kinase activity of c-Src is required for modulating cellular responses to PDGF receptor stimulation (37), we chose to study the role of a subset of novel substrates in PDGF signaling. The role of c-Src has been well studied in PDGF signaling (38). We investigated the

involvement of EWS1, BTF and RBM10 in PDGF receptor signaling. Cortactin and p130CAS were used as positive controls. For this experiment, NIH3T3 cells, which express endogenous PDGF receptors, were treated with PDGF-BB in the presence or absence of a c-Src kinase inhibitor, SU6656 (2-oxo-3-(4,5,6,7-tetrahydro-1 H-indol-2-ylmethylene)-2,3-dihydro-1H-indole-5-sulfonic acid dimethylamide) (38). By activating PDGF signaling in NIH3T3 cells, we investigated the ability of these proteins to get tyrosine phosphorylated upon ligand-induced stimulation of the PDGF receptor. We found that all of the novel proteins were tyrosine phosphorylated upon stimulation of PDGF receptor (Figure 5C), as was the case with the two known substrates. Using SU6656, a potent inhibitor of c-Src kinase, we have shown the involvement of three novel proteins RBM10, EWS1 and BTF as c-Src substrates in PDGF signaling (Figure 5C). c-Src has been shown to have been involved in the regulation of nuclear proteins and transcription factors downstream of PDGF signaling and plays an important role in controlling DNA synthesis (39). We have also examined the involvement of RasGAP SH3-domain binding protein and Thyroid hormone receptor associated protein 3 in PDGF signaling using SU6656 but could not detect any tyrosine phosphorylation of these proteins in any state (data not shown). Further experiments need to be done in order to examine the precise role of these proteins and the importance of their phosphorylation in PDGF signaling. One important caveat of these experiments using SU6656 for inhibition of Src kinase is that it could bind and inhibits other related kinases. When better inhibitors become available this would be a great approach to identify the kinase-substrate relationships.

Ewing sarcoma breakpoint region 1 (EWS1) is an RNA binding protein and has been shown to be involved in gene translocations and often appear as fusion of EWS with ETS family transcription factor genes and known to cause Ewing sarcoma tumors (40). One of the fusion proteins EWS-Fli1 has been implicated in insulin-like growth factor 1 (IGF-1) (41) as well as PDGF-BB (42) induced proliferation of Ewing sarcoma cells. The implication of EWS1 as a downstream substrate of c-Src in PDGF-signaling might shed light into the mechanism of induction of proliferation by these genes in signaling and tumors. Bcl-2-associated transcription factor 1 (BTF) is an apoptotic transcriptional repressor (43) localized to the nucleus and whose function is still under investigation. RNA binding motif 10 (RBM10) is an RNA binding protein with a zinc finger domain and associated with the expression of Bax family members in breast cancers and VEGF (44). This was the first time BTF and RBM10 were identified as a tyrosine phosphorylated proteins and the finding that they are components of PDGF signaling downstream of c-Src might help to understand the role of BTF and RBM10 in growth factor receptor signaling. Although we have observed tyrosine phosphorylation of a subset of proteins in PDGF signaling, it is not possible to validate all candidates in any proteomics experiment because of the following reasons: i) Availability of good immunoprecipitating antibodies against these proteins is limited and hence we have tagged a subset of these proteins to show that these are indeed substrates; and, ii) To validate these proteins in a specific signaling pathway, it is not easy to predict in which signaling pathway (s) these proteins are involved downstream of c-Src. However, studies are currently in progress on a subset of proteins to show their physiological relevance and also to identify tyrosine phosphorylated proteins in growth factor signaling pathways downstream of Src kinase.

Development of peptide microarrays for high throughput validation of c-Src substrates

Although we have established above that a number of novel proteins are potential substrates of c-Src, the exact residues that undergo phosphorylation have not been identified in most of these instances. Hence, we developed a custom peptide microarray as a platform to rapidly identify the phosphopeptides on these proteins which are phosphorylated by c-Src. We systematically designed 312 peptides encompassing all tyrosines from 14 selected proteins (Table 3). These were synthesized in such a fashion that they contained the tyrosine residue

being tested in the center. In parallel an equal number of peptides that have the centric tyrosine residues mutated to phenylalanine were designed.

In all, 624 WT or mutant (312WT and 312MUT) peptides from 14 proteins were spotted with each sequence being represented in triplicate, on to the glass slides as described earlier (18). The design of these peptide microarrays is analogous to DNA microarrays manufactured by Affymetrix for mRNA expression studies. c-Src kinase assays were performed on the peptide microarrays and the arrays subsequently exposed to phosphorimager screen (Figure 6A). Intensity values were obtained using GenePix software as described under 'materials and methods.' We normalized the intensity values and compared the log₂ intensities of the WT peptides against their corresponding MUT peptides. The intensity values from 3 different experiments were averaged individually for WT and MUT peptides and plotted (Figure 5B). The spots in the upper right quadrant were taken as true positives (Figure 6B). The false positive rates (FPR) were calculated by assuming that those peptides for which MUT intensity exceeded WT intensity were not phosphorylated, and that the variation of WT-MUT intensities for those points was representative of non-phosphorylated peptides (Figure 6C). Based on the calculated FPRs, a line was drawn which separates the true positives from the remainder of the peptides.

From this analysis, we have identified phosphorylation sites on 12 out of 14 proteins that were spotted on peptide microarrays. Peptides containing multiple tyrosines were mutated systematically and all the tyrosine containing peptides were looked at along with their corresponding mutated counterparts to deduce the correct phosphorylated peptide. A total of 34 peptides from 12 proteins (Table 3) out of 312 peptides from 14 proteins spotted (Supplementary Table 2) were phosphorylated by c-Src in our analysis. This included 6 phosphopeptides from Bcl2-associated transcription factor, 5 from Thyroid hormone receptor associated protein 3 and 4 phosphopeptides from ARS2. We did not detect phosphorylation on any of the peptides derived from two proteins, Zinc finger, CCHC domain containing 8 and Splicing factor proline/glutamine-rich. Thus peptide arrays allowed assignment of phosphorylation sites in a high-throughput fashion and also served as an additional validation step for potential substrates identified from our SILAC experiments. Because of the tremendous potential and promise that peptide microarrays hold for the identification kinase-substrate identification, studies are ongoing to spot the peptides from other proteins identified in this and other mass spectrometry based studies to identify the kinase specific phosphorylation sites and in building phosphorylation motifs.

Conclusions

This study describes identification of substrates of a non-receptor tyrosine kinase using a combination of proteomic approaches. Use of SILAC methodology allowed us to identify a number of known and novel substrates of c-Src in human embryonic kidney 293T cells. We identified 4 new phosphorylation sites and also validated a subset of the novel substrates as direct c-Src substrates using *in vitro* kinase assays. We corroborated our results for four proteins as direct substrates of c-Src. We also implicated three of the novel c-Src substrates as tyrosine phosphorylated proteins in PDGF receptor signaling. Since identification of phosphopeptides is still a challenge in proteomics, we designed a custom peptide microarray platform for high-throughput identification of peptides phosphorylated by specific kinase, c-Src in this case.

Using peptide microarrays, we identified 34 phosphopeptides phosphorylated by c-Src that are derived from 12 novel candidate substrate proteins that were identified by SILAC as potential c-Src substrates. Identification of the phosphopeptides from these 12 new substrates also provides validation of this approach. It is worth noting here that most tyrosine containing peptides were not phosphorylated by c-Src. Peptide microarray technology has its own limitations. Although it can help identify *bona fide* substrate peptides in many instances, it is

possible that some sequences that are phosphorylated *in vivo* do not get phosphorylated because of lack of secondary and tertiary structure of the immobilized peptides. Further, although the peptide microarray analysis has shed light on the phosphopeptides preferentially phosphorylated by c-Src, these sites still remain to be investigated *in vivo* using other methodologies.

The present study offers many encouraging leads, such as the identification of tyrosine phosphorylation of a subset of new c-Src substrates that are components in the PDGF signaling downstream of c-Src. Clearly, further characterization of these novel sites and proteins will result in a significant expansion of our knowledge of the c-Src kinase signaling network. The significance of tyrosine phosphorylation on each of the newly discovered sites remains to be determined. In any case, these results demonstrate that c-Src-mediated tyrosine phosphorylation is extensive and implicates a number of hitherto unrecognized proteins as c-Src kinase substrates.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENT

AP is supported by grants from the National Institutes of Health (CA106424 and U54 RR020839), Department of Defense Era of Hope Scholar award (W81XWH-06-1-0428) and by the Beckman Young Investigator award. LC is supported by grants from the National Cancer Institute (P30 CA06973-44). We thank members of Pandey lab for fruitful discussions. We also thank Hopkins Expressionists Working Group for advice and comments on peptide microarray analysis. Jos Joore is VP of Array Technology at PepScan Systems. We also thank Jurriaan Tuynman and Maikel Peppelenbosch, Academic Medical Center, Amsterdam, The Netherlands for their helpful suggestions on peptide microarrays.

References

1. Thomas SM, Brugge JS. Cellular functions regulated by Src family kinases. *Annu Rev Cell Dev Biol* 1997;13:513–609. [PubMed: 9442882]
2. Olsen JV, Blagoev B, Gnäd F, Macek B, Kumar C, Mortensen P, Mann M. Global, *in vivo*, and site-specific phosphorylation dynamics in signaling networks. *Cell* 2006;127:635–648. [PubMed: 17081983]
3. Ficarro SB, McClelland ML, Stukenberg PT, Burke DJ, Ross MM, Shabanowitz J, Hunt DF, White FM. Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat Biotechnol* 2002;20:301–305. [PubMed: 11875433]
4. Beausoleil SA, Villen J, Gerber SA, Rush J, Gygi SP. A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat Biotechnol* 2006;24:1285–1292. [PubMed: 16964243]
5. Molina H, Horn DM, Tang N, Mathivanan S, Pandey A. Global Proteomic Profiling of Phosphopeptides using Electron Transfer Dissociation Tandem Mass Spectrometry. *Proc Natl Acad Sci U S A* 2007;104:2199–2204. [PubMed: 17287340]
6. Shah K, Shokat KM. A chemical genetic screen for direct v-Src substrates reveals ordered assembly of a retrograde signaling pathway. *Chem Biol* 2002;9:35–47. [PubMed: 11841937]
7. Lock P, Abram CL, Gibson T, Courtneidge SA. A new method for isolating tyrosine kinase substrates used to identify fish, an SH3 and PX domain-containing protein, and Src substrate. *Embo J* 1998;17:4346–4357. [PubMed: 9687503]
8. Qiao Y, Molina H, Pandey A, Zhang J, Cole PA. Chemical rescue of a mutant enzyme in living cells. *Science* 2006;311:1293–1297. [PubMed: 16513984]
9. Peri S, Navarro JD, Amanchy R, Kristiansen TZ, Jonnalagadda CK, Surendranath V, Niranjan V, Muthusamy B, Gandhi TK, Gronborg M, Ibarrola N, Deshpande N, Shanker K, Shivashankar HN, Rashmi BP, Ramya MA, Zhao Z, Chandrika KN, Padma N, Harsha HC, Yatish AJ, Kavitha MP,

- Menezes M, Choudhury DR, Suresh S, Ghosh N, Saravana R, Chandran S, Krishna S, Joy M, Anand SK, Madavan V, Joseph A, Wong GW, Schiemann WP, Constantinescu SN, Huang L, Khosravi-Far R, Steen H, Tewari M, Ghaffari S, Blobel GC, Dang CV, Garcia JG, Pevsner J, Jensen ON, Roepstorff P, Deshpande KS, Chinnaiyan AM, Hamosh A, Chakravarti A, Pandey A. Development of human protein reference database as an initial platform for approaching systems biology in humans. *Genome Res* 2003;13:2363–2371. [PubMed: 14525934]
10. Amanchy R, Kalume DE, Iwahori A, Zhong J, Pandey A. Phosphoproteome analysis of HeLa cells using stable isotope labeling with amino acids in cell culture (SILAC). *J Proteome Res* 2005;4:1661–1671. [PubMed: 16212419]
 11. Ibarrola N, Molina H, Iwahori A, Pandey A. A novel proteomic approach for specific identification of tyrosine kinase substrates using [¹³C]tyrosine. *J Biol Chem* 2004;279:15805–15813. [PubMed: 14739304]
 12. Kratchmarova I, Blagoev B, Haack-Sorensen M, Kassem M, Mann M. Mechanism of divergent growth factor effects in mesenchymal stem cell differentiation. *Science* 2005;308:1472–1477. [PubMed: 15933201]
 13. Pandey A, Podtelejnikov AV, Blagoev B, Bustelo XR, Mann M, Lodish HF. Analysis of receptor signaling pathways by mass spectrometry: identification of vav-2 as a substrate of the epidermal and platelet-derived growth factor receptors. *Proc Natl Acad Sci U S A* 2000;97:179–184. [PubMed: 10618391]
 14. Rush J, Moritz A, Lee KA, Guo A, Goss VL, Spek EJ, Zhang H, Zha XM, Polakiewicz RD, Comb MJ. Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat Biotechnol* 2005;23:94–101. [PubMed: 15592455]
 15. Amanchy R, Kalume DE, Pandey A. Stable isotope labeling with amino acids in cell culture (SILAC) for studying dynamics of protein abundance and posttranslational modifications. *Sci STKE* 2005;2005:12.
 16. Schulze WX, Mann M. A novel proteomic screen for peptide-protein interactions. *J Biol Chem* 2004;279:10756–10764. [PubMed: 14679214]
 17. Molina H, Parmigiani G, Pandey A. Assessing reproducibility of a protein dynamics study using in vivo labeling and liquid chromatography tandem mass spectrometry. *Anal Chem* 2005;77:2739–2744. [PubMed: 15859588]
 18. Diks SH, Kok K, O'Toole T, Hommes DW, van Dijken P, Joore J, Peppelenbosch MP. Kinome profiling for studying lipopolysaccharide signal transduction in human peripheral blood mononuclear cells. *J Biol Chem* 2004;279:49206–49213. [PubMed: 15355981]
 19. Cleveland, WS.; Loader, CR. Technical Report. AT&T Bell Laboratories; Murray Hill, NY: 1995. Smoothing by local regression: Principles and methods.
 20. Cooper JA, Gould KL, Cartwright CA, Hunter T. Tyr527 is phosphorylated in pp60c-src: implications for regulation. *Science* 1986;231:1431–1434. [PubMed: 2420005]
 21. Bagrodia S, Chackalaparampil I, Kmiecik TE, Shalloway D. Altered tyrosine 527 phosphorylation and mitotic activation of p60c-src. *Nature* 1991;349:172–175. [PubMed: 1702522]
 22. Kaplan KB, Bibbins KB, Swedlow JR, Arnaud M, Morgan DO, Varmus HE. Association of the amino-terminal half of c-Src with focal adhesions alters their properties and is regulated by phosphorylation of tyrosine 527. *Embo J* 1994;13:4745–4756. [PubMed: 7525268]
 23. Thomas JE, Soriano P, Brugge JS. Phosphorylation of c-Src on tyrosine 527 by another protein tyrosine kinase. *Science* 1991;254:568–571. [PubMed: 1719633]
 24. Schuh SM, Brugge JS. Investigation of factors that influence phosphorylation of pp60c-src on tyrosine 527. *Mol Cell Biol* 1988;8:2465–2471. [PubMed: 2457151]
 25. Luttrell LM, Hawes BE, van Biesen T, Luttrell DK, Lansing TJ, Lefkowitz RJ. Role of c-Src tyrosine kinase in G protein-coupled receptor- and Gbetagamma subunit-mediated activation of mitogen-activated protein kinases. *J Biol Chem* 1996;271:19443–19450. [PubMed: 8702633]
 26. Chang JH, Wilson LK, Moyers JS, Zhang K, Parsons SJ. Increased levels of p21ras-GTP and enhanced DNA synthesis accompany elevated tyrosyl phosphorylation of GAP-associated proteins, p190 and p62, in c-src overexpressors. *Oncogene* 1993;8:959–967. [PubMed: 7681161]
 27. Kim J, Lee JM, Branton PE, Pelletier J. Modulation of EWS/WT1 activity by the v-Src protein tyrosine kinase. *FEBS Lett* 2000;474:121–128. [PubMed: 10838070]

28. Huang C, Liu J, Haudenschild CC, Zhan X. The role of tyrosine phosphorylation of cortactin in the locomotion of endothelial cells. *J Biol Chem* 1998;273:25770–25776. [PubMed: 9748248]
29. Abouzaglou J, Benistant C, Gimona M, Roustan C, Kassab R, Fattoum A. Tyrosine phosphorylation of calponins. Inhibition of the interaction with F-actin. *Eur J Biochem* 2004;271:2615–2623. [PubMed: 15206927]
30. Ostareck-Lederer A, Ostareck DH, Cans C, Neubauer G, Bomsztyk K, Superti-Furga G, Hentze MW. c-Src-mediated phosphorylation of hnRNP K drives translational activation of specifically silenced mRNAs. *Mol Cell Biol* 2002;22:4535–4543. [PubMed: 12052863]
31. Yang JP, Reddy TR, Truong KT, Suhasini M, Wong-Staal F. Functional interaction of Sam68 and heterogeneous nuclear ribonucleoprotein K. *Oncogene* 2002;21:7187–7194. [PubMed: 12370808]
32. Rahmouni S, Vang T, Alonso A, Williams S, van Stipdonk M, Soncini C, Moutschen M, Schoenberger SP, Mustelin T. Removal of C-terminal SRC kinase from the immune synapse by a new binding protein. *Mol Cell Biol* 2005;25:2227–2241. [PubMed: 15743820]
33. Parker F, Maurier F, Delumeau I, Duchesne M, Faucher D, Debussche L, Dugue A, Schweighoffer F, Tocque B. A Ras-GTPase-activating protein SH3-domain-binding protein. *Mol Cell Biol* 1996;16:2561–2569. [PubMed: 8649363]
34. Cichowski K, McCormick F, Brugge JS. p21rasGAP association with Fyn, Lyn, and Yes in thrombin-activated platelets. *J Biol Chem* 1992;267:5025–5028. [PubMed: 1544885]
35. Kmiecik TE, Shalloway D. Activation and suppression of pp60c-src transforming ability by mutation of its primary sites of tyrosine phosphorylation. *Cell* 1987;49:65–73. [PubMed: 3103925]
36. Ohan N, Sabourin D, Booth RA, Liu XJ. *Xenopus laevis* TRK-fused gene (TFG) is an SH3 domain binding protein highly expressed in the cement gland. *Mol Reprod Dev* 2000;56:336–344. [PubMed: 10861999]
37. Broome MA, Hunter T. Requirement for c-Src catalytic activity and the SH3 domain in platelet-derived growth factor BB and epidermal growth factor mitogenic signaling. *J Biol Chem* 1996;271:16798–16806. [PubMed: 8663329]
38. Blake RA, Broome MA, Liu X, Wu J, Gishizky M, Sun L, Courtneidge SA. SU6656, a selective src family kinase inhibitor, used to probe growth factor signaling. *Mol Cell Biol* 2000;20:9018–9027. [PubMed: 11074000]
39. Shah K, Vincent F. Divergent roles of c-Src in controlling platelet-derived growth factor-dependent signaling in fibroblasts. *Mol Biol Cell* 2005;16:5418–5432. [PubMed: 16135530]
40. Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, Peter M, Kovar H, Joubert I, de Jong P, Rouleau G, et al. Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* 1992;359:162–165. [PubMed: 1522903]
41. Prieur A, Tirode F, Cohen P, Delattre O. EWS/FLI-1 silencing and gene profiling of Ewing cells reveal downstream oncogenic pathways and a crucial role for repression of insulin-like growth factor binding protein 3. *Mol Cell Biol* 2004;24:7275–7283. [PubMed: 15282325]
42. Nozawa S, Ohno T, Banno Y, Dohjima T, Wakahara K, Fan DG, Shimizu K. Inhibition of platelet-derived growth factor-induced cell growth signaling by a short interfering RNA for EWS-Flil1 via down-regulation of phospholipase D2 in Ewing sarcoma cells. *J Biol Chem* 2005;280:27544–27551. [PubMed: 15919668]
43. Kasof GM, Goyal L, White E. Btf, a novel death-promoting transcriptional repressor that interacts with Bcl-2-related proteins. *Mol Cell Biol* 1999;19:4390–4404. [PubMed: 10330179]
44. Martinez-Arribas F, Agudo D, Pollan M, Gomez-Esquer F, Diaz-Gil G, Lucas R, Schneider J. Positive correlation between the expression of X-chromosome RBM genes (RBMX, RBM3, RBM10) and the proapoptotic Bax gene in human breast cancer. *J Cell Biochem* 2006;97:1275–1282. [PubMed: 16552754]
45. Bernhard OK, Cunningham AL, Sheil MM. Analysis of proteins copurifying with the CD4/lck complex using one-dimensional polyacrylamide gel electrophoresis and mass spectrometry: comparison with affinity-tag based protein detection and evaluation of different solubilization methods. *J Am Soc Mass Spectrom* 2004;15:558–567. [PubMed: 15047060]
46. Scott MP, Zappacosta F, Kim EY, Annan RS, Miller WT. Identification of novel SH3 domain ligands for the Src family kinase Hck. Wiskott-Aldrich syndrome protein (WASP), WASP-interacting protein (WIP), and ELMO1. *J Biol Chem* 2002;277:28238–28246. [PubMed: 12029088]

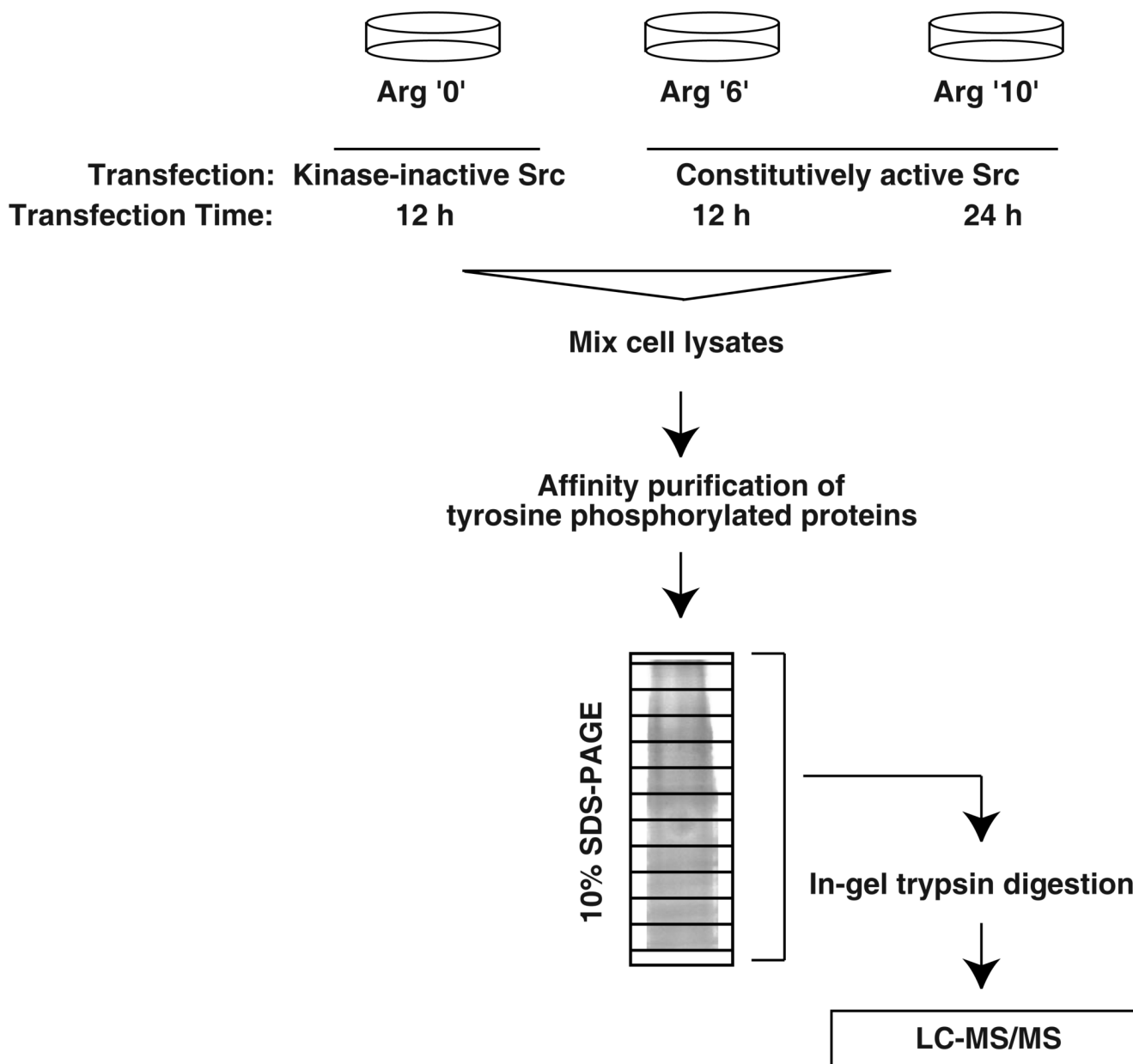


Figure 1.

(A) Schematic for the integrated proteomic approach for the identification of c-Src kinase substrates. Human embryonic kidney (HEK) 293T cells growing in Arg `0' containing medium were transiently transfected with a kinase-dead Src (K298M) and 293T cells growing in Arg `6' and Arg `10' were transiently transfected with constitutively active Src kinase (Y527F). Arg `0' refers to $^{12}\text{C}_6$ -arginine, Arg `6' refers to $^{13}\text{C}_6$ -arginine and Arg `10' refers to $^{13}\text{C}_6$ - $^{15}\text{N}_4$ -arginine, isotopic labeled forms of arginine used to differentially label 293T cells for identification of Src substrates.

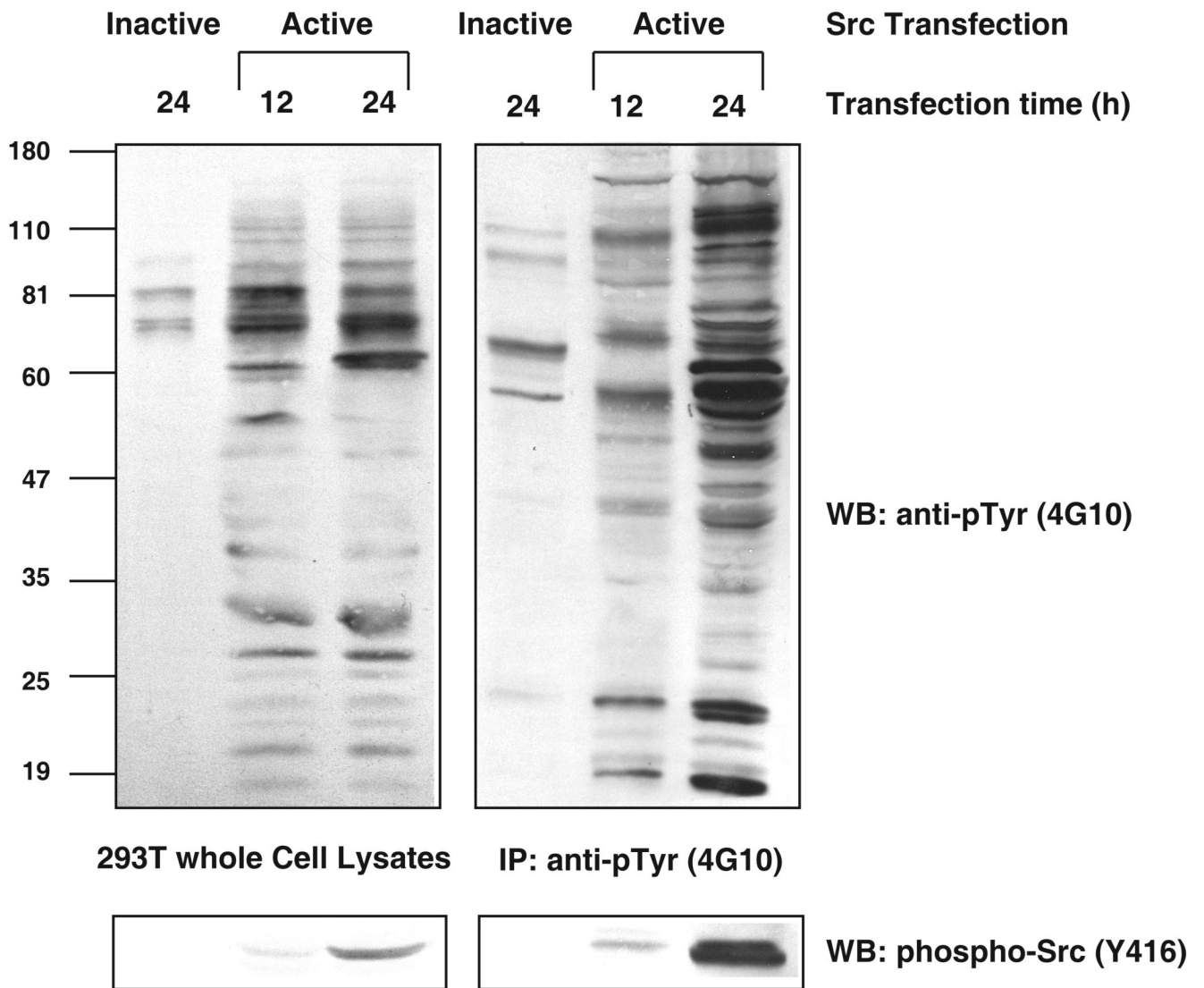


Figure 2.

Tyrosine phosphorylation profile of proteins on transfection with inactive and active forms of c-Src. 293T cells were transfected with inactive and active forms of c-Src and cells were lysed and tyrosine phosphorylated proteins were immunoprecipitated from the cell lysates as described in materials and methods. Cell lysates and immunoprecipitates were then run on a 10% SDS-PAGE and transferred to nitrocellulose membranes. The membranes were probed with anti-phosphotyrosine antibodies against and re-probed with phospho (Y416)-Src antibody.

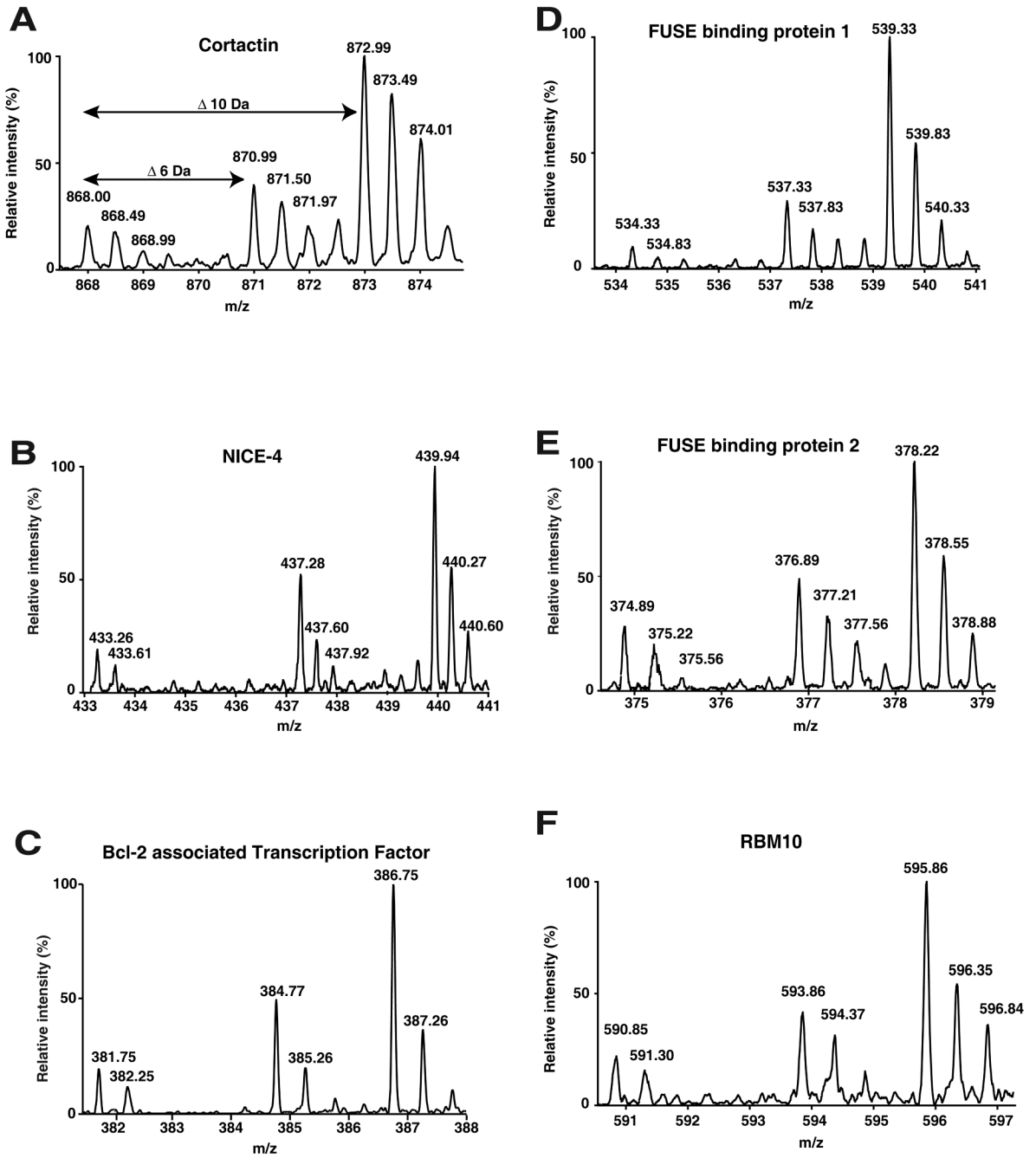


Figure 3. MS spectra of 6 proteins identified as Src substrates by SILAC. The 3 spectral peaks in each figure represent the mass shift of the same peptide. The relative increase in intensity ratios between light to heavy are represented below in parentheses (A) A doubly charged peptide from cortactin (1:7) (B) A triply charged peptide from NICE-4 (1:7) (C) A doubly charged peptide from Bcl2-associated transcription factor (1:7.5) (D) A doubly charged peptide from FUSE-binding protein 1 (1:8) (E) A doubly charged peptide from FUSE-binding protein 2 (1:4) (F) A doubly charged peptide from RNA binding motif 10 (1:7.5).

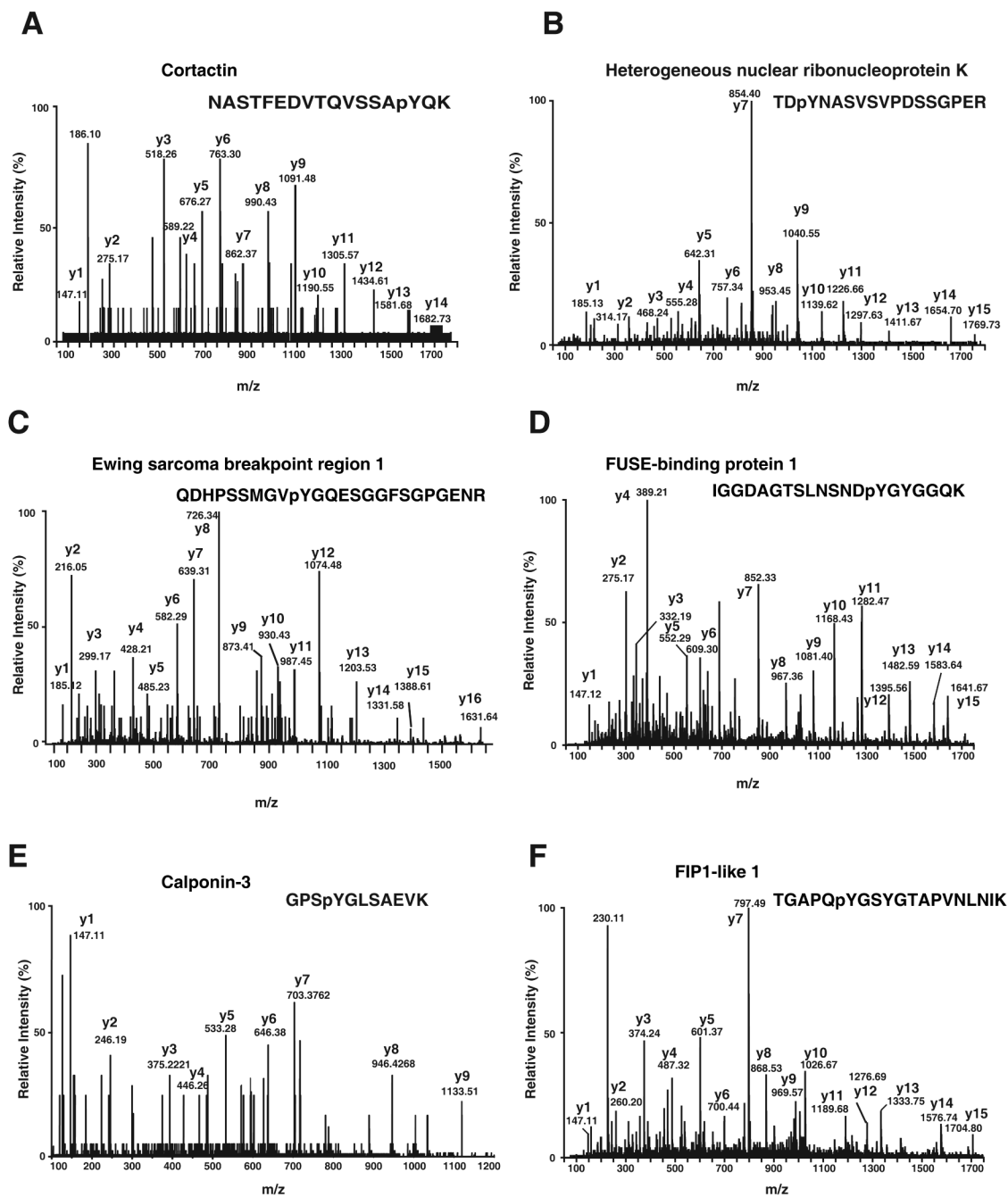


Figure 4. MS/MS spectra of novel phosphorylation sites identified in this study (A) Phosphopeptide NASTFEDVTQVSSApYQK derived from Cortactin (B) Phosphopeptide TDpYNASVSPDSSGPER derived from hnRNP K (C) Phosphopeptide QDHPSSMGVpYQGESGGFSGPGENR derived from Ewing sarcoma breakpoint region 1 (D) Phosphopeptide IGGDAGTSLNSNDpYGYGGQK derived from FUSE-binding protein 1 (E) Phosphopeptide GPSpYGLSAEVK derived from Calponin-3 (F) Phosphopeptide TGAPQpYGSYGTAPVNLNIK derived from FIP1-like 1.

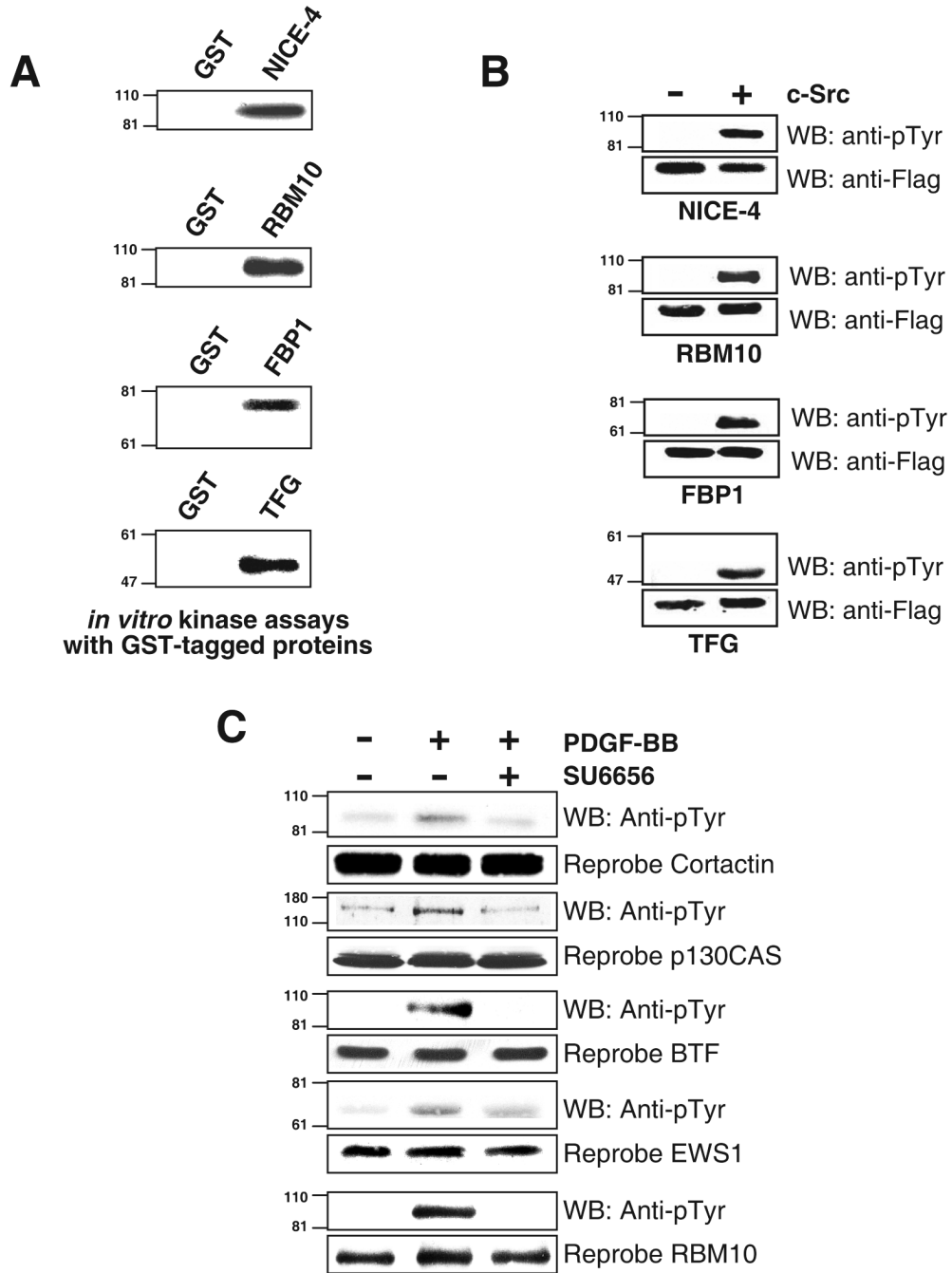


Figure 5. Experimental validation of tyrosine phosphorylation of proteins obtained from c-Src kinase overexpression in 293T cells (A) *In vitro* Kinase assays using GST tagged proteins and c-Src using a rabbit reticulocyte *in vitro* transcription and translation system (B) 293T cells were co-transfected with genes of interest along with either empty vector PCMVtag4A or with c-Src. Culture media was changed 12 hours after transfection and cells were serum starved after serum starved for 12 hours and lysed 48 hours after transfection. Proteins were immunoprecipitated using anti-Flag antibodies and Western blotting was performed using phosphotyrosine antibodies and re-probed (C) Validation of a subset of proteins in PDGF signaling. NIH3T3 cells have been grown to confluence and serum starved for 12 hours followed by stimulation

with PDGF-BB (100 ng/ml for 5 min) and PDGF stimulation after treatment with SU6656 (2 μ M for 1 hour prior to lysis or stimulation) and cell lysates were subjected to immunoprecipitation using anti-phosphotyrosine antibodies and probed with respective antibodies, and re-probed in whole cell lysates.

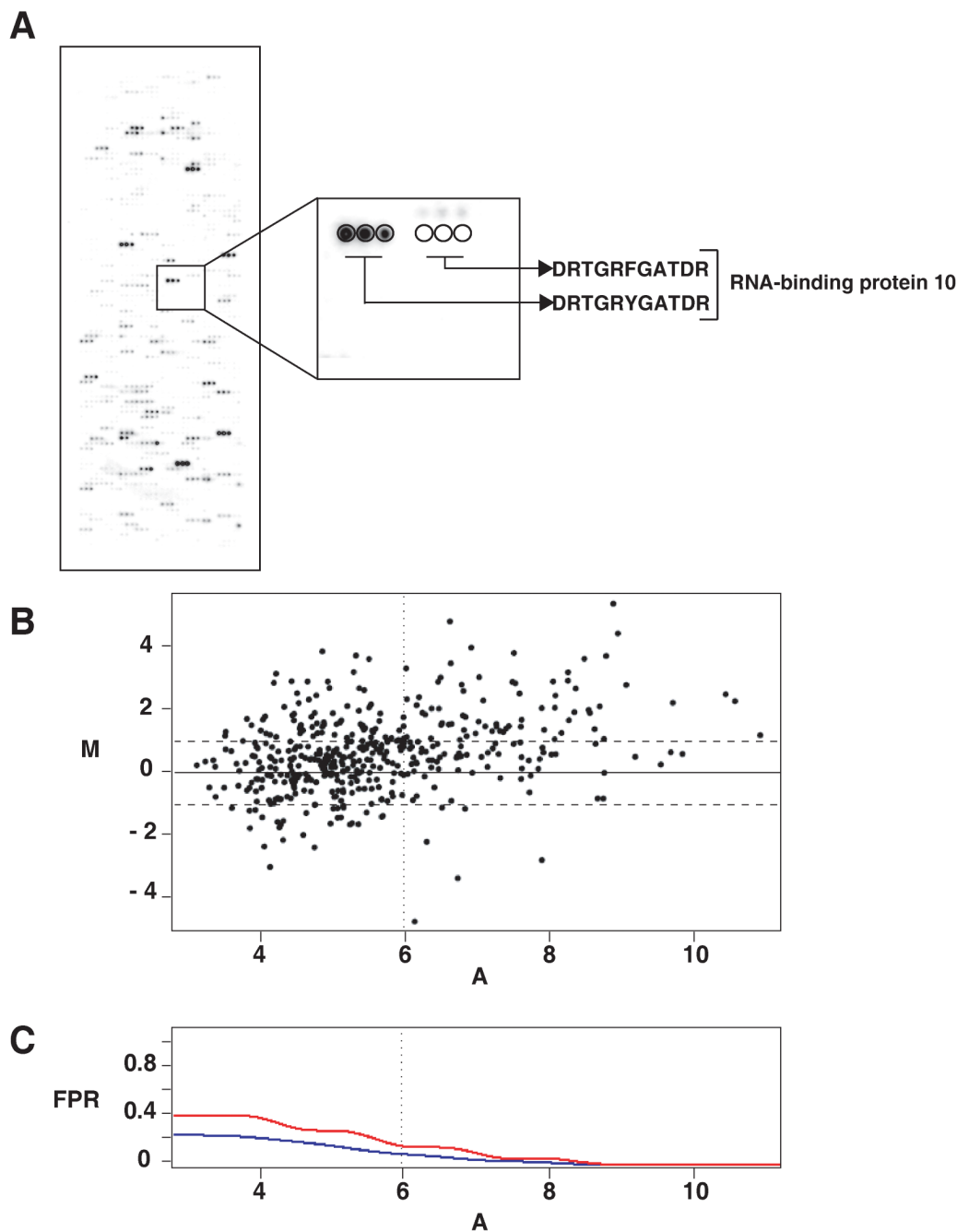


Figure 6. (A) Peptide microarrays: *in vitro* kinase assays performed on peptide microarrays, where all tyrosine containing peptides and their corresponding Y→F mutant counterparts were spotted on glass slides. A representative section of the peptide microarray is magnified to show the signal corresponding to a peptide and its Y→F mutant. (B) A classical MvA plot displaying data pertaining phosphorylation intensities on peptide microarrays. The horizontal dotted lines indicate 2-fold difference between WT and MUT intensity values. The vertical dotted line corresponds to a local false positive rate of 0.15. M on the Y- axis represents differential of (log₂ WT - log₂ MUT intensity values for each peptide and A on the X-axis represents average intensities ([log₂ WT + log₂ MUT]/2) (C) A plot displaying classical and local false positive

rates (FPR). The red line represents local false positive rate curve and the blue line represents the classical false positive rate. The vertical dotted line shows where the local FPR=0.15. All peptides to the right of the vertical dotted line and above two fold (horizontal line) were selected as true positives.

Table 1
Known Src Family kinase substrates and interactors identified in this study by overexpression of c-Src Kinase in 293T cells followed by SILAC

	NCBI Accession No.	Protein	Fold increase (Heavy: light) [*] ± S.D.	Reference
Known c-Src substrates				
1	NP_005408	c-Src	30 ± 9.1	(35)
2	NP_005222	Cortactin	4.9 ± 2.6	(28)
3	NP_005234	Ewing sarcoma breakpoint region 1	4.6	(27)
4	NP_112552	Heterogeneous nuclear ribonucleoprotein K	3.7 ± 0.2	(30,31)
5	NP_001830	Calponin 3	3.4	(29)
6	NP_005745	RasGAP SH3-domain binding protein	2.2	(32-34)
Known Src family kinase interactors				
7	NP_112533	Heterogeneous nuclear ribonucleoprotein A2/B1	10.2 ± 5.1	(45)
8	NP_002129	Heterogeneous nuclear ribonucleoprotein D	4.3	(45)
9	NP_006833	Splicing factor 3B subunit 2	5.6 ± 0.9	(46)
10	NP_005511	Heterogeneous nuclear ribonucleoprotein H1	3.0	(45)

* Heavy refers to Arg¹⁰ and light refers to Arg⁰ containing peptides

Table 2

List of novel potential Src substrates identified using SILAC by overexpression of Src Kinase in 293T cells

Accession #	Protein	Fold increase (Heavy:light)* ± S.D.
NP_005454	Heterogeneous nuclear ribonucleoprotein D-like	10.0 ± 5.2
NP_113680	RNA binding motif protein 4B	9.4 ± 3.7
NP_003760	Splicing factor, arginine/serine-rich 9	9.3 ± 3.6
NP_004490	Heterogeneous nuclear ribonucleoprotein AB	8.0
NP_005667	RNA binding motif protein 10	7.2 ± 0.6
NP_003893	FUSE binding protein	6.9 ± 0.8
NP_055662	NICE-4	6.6 ± 2.1
NP_919223	Heterogeneous nuclear ribonucleoprotein A3	6.1 ± 2.5
NP_003676	FUSE binding protein 2	5.7 ± 2.2
NP_004951	FUS/TLS oncogene	5.0 ± 0.9
NP_005849	A-kinase anchor protein 8	5.5
NP_055554	Bcl-2-associated transcription factor 1	5.5
NP_877952	Arsenate resistance protein ARS2	5.4
NP_473357	FUS interacting protein (serine/arginine-rich) 1	5.1
NP_005110	Thyroid hormone receptor associated protein 3	5.1
NP_008937	Cleavage and polyadenylation specific factor 5	4.8 ± 3.1
XP_028253	Similar to Zinc finger CCCH-type domain-containing protein 6	3.8 ± 1.9
NP_079222	NEFA-interacting nuclear protein	3.8 ± 0.8
NP_060082	Zinc finger, CCHC domain containing 8	3.7
NP_005057	Splicing factor proline/glutamine-rich	3.1 ± 1.0
NP_005780	Proteasome activator subunit 3	2.6 ± 0.4
NP_006061	TRK-fused gene	2.4
NP_003746	Eukaryotic translation initiation factor 3, subunit 4	2.2
NP_663760	Ataxin 2 related protein	2.2
NP_067038	Chromosome 20 open reading frame 77	2.0 ± 0.3
NP_057123	Homeobox prox 1	2.0 ± 0.3

* Heavy refers to Arg `10' and light refers to Arg `0' containing peptides

Table 3

Peptides from the newly identified Src substrates that were phosphorylated by c-Src on peptide microarrays

	Accession No.	Protein	Phosphopeptide
1	NP_003893	FUSE binding protein 1	1 EVRNEYGSRIG 2 RQQAAYYAQTS
2	NP_003676	FUSE binding protein 2	1 GDRNEYGSRIG 2 AYYSHYYQQPP 3 RQQAAYYGQTP
3	NP_055554	BCL2-associated transcription factor 1	1 YHRGGYRPVWN 2 EETEDYRQFRK 3 GRGRGYYQGGG 4 NGSSRYSPSQN 5 RGRGYYQGGGG 6 RSSSPYKSPV
4	NP_005667	RNA binding motif protein 10	1 ARGSSYGVTST 2 EPPVDYSYYQQ 3 DRTGRYGATDR
5	NP_006061	TRK-fused Gene	1 NEDITYDELVL 2 QMYQQYQQQAG
6	NP_005745	RasGAP SH3-domain-binding protein	1. NDIFRYQDEVF
7	NP_663760	Ataxin 2 related protein	1. GQQGKYRGAKG
8	NP_005849	A-Kinase anchor protein 8	1. RPSYSYDYEFD
9	NP_005110	Thyroid hormone receptor associated protein 3	1 NYRQAYSPRRG 2 NHPRVYQNRDF 3 SGGAAATKRYL 4 GGYGNYSNWQ 5 GTPAGYGRGRE
10	NP_877952	Arsenate resistance protein ARS2	1 AGRGNYDAFRG 2 QGLMPYQPRP 3 HSSDPYHSGYE 4 KRYNDYKLDLR
11	NP_006833	Splicing factor 3B subunit 2	1 HGDLYYEGKEF 2 EEPEIYEPNFI 3 QREESYSRMGY
12	NP_055983	hypothetical protein LOC23211	1 GRGRGYRGRGS 2 DRYNSYNRPRP 3 RKYREYSPPYA