Resource

Cell

Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks

Jesper V. Olsen,[1,2,3] Blagoy Blagoev,[1,3,*] Florian Gnad,[2,3] Boris Macek,[1,2] Chanchal Kumar,[2] Peter Mortensen,[1] and Matthias Mann[1,2,*]

[1] Center for Experimental BioInformatics, Department of Biochemistry and Molecular Biology, University of Southern Denmark, DK-5230 Odense, Denmark

[2] Department of Proteomics and Signal Transduction, Max-Planck-Institute for Biochemistry, D-82152 Martinsried, Germany

[3] These authors contributed equally to this work.

*Contact: bab@bmb.sdu.dk (B.B.), mmann@biochem.mpg.de (M.M.)

DOI 10.1016/j.cell.2006.09.026

SUMMARY

Cell signaling mechanisms often transmit information via posttranslational protein modifications, most importantly reversible protein phosphorylation. Here we develop and apply a general mass spectrometric technology for identification and quantitation of phosphorylation sites as a function of stimulus, time, and subcellular location. We have detected 6,600 phosphorylation sites on 2,244 proteins and have determined their temporal dynamics after stimulating HeLa cells with epidermal growth factor (EGF) and recorded them in the Phosida database. Fourteen percent of phosphorylation sites are modulated at least 2-fold by EGF, and these were classified by their temporal profiles. Surprisingly, a majority of proteins contain multiple phosphorylation sites showing different kinetics, suggesting that they serve as platforms for integrating signals. In addition to protein kinase cascades, the targets of reversible phosphorylation include ubiquitin ligases, guanine nucleotide exchange factors, and at least 46 different transcriptional regulators. The dynamic phosphoproteome provides a missing link in a global, integrative view of cellular regulation.

INTRODUCTION

The mammalian cell constantly receives signals from its surroundings to which it has to respond appropriately. For example, growth-factor signals are integrated with internal-state information and lead to decisions on cell growth, differentiation, or proliferation (Hunter, 2000; Pawson and Nash, 2003; Schlessinger, 2000). Many human diseases, including multiple forms of cancer, arise through deregulation of this information processing capability. In recent decades, our knowledge of the players in signal transduction mechanisms has been painstakingly accumulated, mainly through the study of individual molecules in specific pathways. More recently, the emergence of technologies allowing high-throughput, system-wide experiments—such as microarray analysis—has provided a detailed and objective view of downstream transcriptional changes following various stimuli. However, many critical events involved in cellular responses are mediated by changes in posttranslational protein modifications rather than transcriptional changes. Thus, protein modification can influence and control enzymatic activity, protein conformation, protein-protein interactions, and cellular localization. Even for protein phosphorylation, which affects an estimated one-third of all proteins and is the most widely studied posttranslational modification (Cohen, 2001), only a small subset of total in vivo sites has been discovered so far. Development of global and quantitative methods for elucidating dynamic phosphorylation events is therefore essential for a systematic understanding of cellular behavior.

Phosphorylation has traditionally been studied largely by in vitro assays, a method which has recently been extended using protein chip arrays (Ptacek et al., 2005). Likewise, synthetic peptides have served as kinase substrates and allowed extraction of consensus motifs (Songyang et al., 1994), which were then incorporated into in silico prediction programs (Yaffe et al., 2001). Unfortunately, kinases are often less specific in vitro than they are in vivo, necessitating additional experimental approaches. Mass spectrometry (MS) has become a powerful technology for proteomics and a method of choice for unbiased (i.e., hypothesis-free) analysis of in vivo phosphorylation (Aebersold and Mann, 2003; Chen and White, 2004; Ficarro et al., 2002; Mumby and Brekken, 2005; Rush et al., 2005; Salomon et al., 2003; Stover et al., 2004).

Intracellular signal transduction mediated by receptor tyrosine kinases, such as the epidermal growth factor (EGF) receptor, is governed by phosphorylation of

downstream protein kinases and their substrates (Hunter, 2000; Pawson and Nash, 2003), and several MS studies have focused on tyrosine-phosphorylation "early events" after growth-factor treatment (Blagoev et al., 2003, 2004; Kratchmarova et al., 2005; Zhang et al., 2005). These approaches employed anti-phosphotyrosine antibodies, which exclude detection of downstream serine/threonine kinase signaling. Other studies have identified large numbers of phosphorylation sites, but without a functional context (Beausoleil et al., 2004). We have recently described an integrated phosphoproteomic technology combining phosphopeptide enrichment, high-accuracy identification, and stable isotope labeling by amino acids in cell culture (SILAC) (Ong et al., 2002) to quantify changes in phosphopeptide levels. Initially we applied these methods to the yeast pheromone response pathway (Gruhler et al., 2005). We have now improved and extended this approach and combined it with the time-course method (Andersen et al., 2005; Blagoev et al., 2004) to study, for the first time, the global in vivo phosphoproteome and its temporal dynamics upon growth-factor stimulation.

RESULTS

Phosphopeptide Sequencing and Determination of Temporal Profiles

Our quantitative, phosphopeptide-specific approach combines SILAC for quantitation, strong-cation exchange chromatography (SCX) and titanium dioxide (TiO_2) chromatography for phosphopeptide enrichment, and high-accuracy multistage MS (Figure 1). Three populations of HeLa cells are SILAC encoded with both arginine and lysine using three distinct isotope forms ("double-triple labeling") and stimulated by EGF (150 ng/ml) for different times. Cells are mixed and separated into cytosolic and nuclear fractions, and the proteins are enzymatically digested. The resulting peptide mixtures are then separated into 13 fractions by SCX, and phosphopeptides are enriched on TiO_2 beads in the presence of 2,5-dihydrobenzoic acid (Larsen et al., 2005) (Figure 1A). Each fraction is analyzed by online liquid chromatography (LC) MS on a hybrid linear ion trap/Fourier transform mass spectrometer (LTQ-FT) with two consecutive stages of fragmentation for unambiguous phosphopeptide identification (Beausoleil et al., 2004; Olsen and Mann, 2004) (Figure 1B). In a second, independent experiment, we used the novel linear ion trap/orbitrap mass spectrometer (LTQ-Orbitrap) (Makarov et al., 2006). In this case, phosphopeptides were fragmented by multistage activation ("pseudo-MS^3") (Schroeder et al., 2004) (Figure 1C). All labeled peptides elute as characteristic triplets, with their intensity reflecting the relative amounts at the three time points. Two time-course experiments are combined using the common time point (5 min of EGF stimulation), providing a five-time-point profile (Figure 1D). Quantitation of phosphorylation sites was done by MSQuant software (Schulze and Mann, 2004) and checked manually for all peptides. Extracted ion current (XIC) values indicate that our measurements spanned a phosphopeptide abundance range of between 10^3 and 10^4.

More than 10,000 phosphopeptides were detected in 116 LC MS analyses. We achieved greater than 99% certainty of phosphopeptide identification through the sub-ppm accuracy of peptide mass measurements, double fragmentation of peptides losing a phospho-group (MS^2 and MS^3), and the isotope-state information present in the SILAC peptide triplet (see Table S1 and Figure S1 in the Supplemental Data available with this article online). Furthermore, peptides were typically sequenced several times in different forms—for instance, those containing oxidized methionine or missed tryptic cleavage sites (see Supplemental Experimental Procedures and Table S2 for a listing of all peptides and their ratios). Our analysis identified 6,600 phosphorylation sites on a total of 2,244 proteins (Figure 2).

The above analysis establishes the sequence and number of phosphorylation sites for each phosphopeptide. In some cases, the fragmentation spectra do not contain sufficient information to determine the site of phosphorylation within a peptide with single amino acid accuracy. We therefore developed a phospho-site analysis pipeline using a posttranslational modification (PTM) score for localization, which assigns probabilities to each of the possible sites based on their distinguishing fragment ions (Figures S2–S4). The PTM score is based on an algorithm that makes use of the four most intense fragment ions per 100 m/z units in an MS^2 or MS^3 spectrum (Olsen and Mann, 2004). All possible combinations of serine, threonine, and tyrosine phosphorylation are tested, and the combinations with the highest scores are reported (for details, see Experimental Procedures and Supplemental Data). Furthermore, we extracted from the literature 22 consensus motifs from 15 of the most common kinases and matched all candidate phospho-sites against these motifs (Table S3). Given the peptide sequence and number of phosphorylation sites for each phosphopeptide, we grouped potential phosphorylation sites into four categories depending on their PTM localization score and motifs (Figure 2B). In the category with highest confidence in localization (class I), the site in question had a localization probability for the phospho-group of at least 0.75. That is, the added probability of all other potential sites is less than 0.25. In class II, the localization probability is between 0.75 and 0.25, but these sites also had to match one of the 22 kinase motives. Class III sites had the same localization probabilities as class II but did not match any of the motifs. Finally, class IV sites (not shown in the figure) had localization probabilities below 0.25. Note, however, that the probability that the peptide is phosphorylated is still larger than 99%, even for class IV peptides. Many of the kinase motifs have very limited specificity, which is reflected in the fact that a third of the phospho-sites matched with two or more motifs (Table S4).

To determine the distribution of quantitation errors, we measured the SILAC ratios of all nonphosphorylated peptides. Protein expression is not expected to change

appreciably within the time frame of our experiment, so the SILAC triplets of nonphosphorylated peptides should be 1:1:1. The standard deviation of the measured ratios was below 20% for all four preparations (Figure S5). For the analysis of dynamic changes as a result of growth-factor treatment, a cutoff value of a 2-fold change in either direction is therefore very conservative. A total of 1,046 phosphopeptides were in this category, which we term here "regulated phosphopeptides."

The In Vivo Phosphoproteome

In their classic study using phosphoamino acid analysis, Hunter and coworkers found relative abundances of 0.05%, 10%, and 90% for phosphotyrosine (pY), phosphothreonine (pT), and phosphoserine (pS) in normally growing cells (Hunter and Sefton, 1980). Here we determined the distribution between individually identified sites to be 103 pY (including 53 induced by EGF), 670 pT, and 4,901 pS sites (class I sites, Figure 2C). Thus, our new data set, based on more than 2,000 phosphoproteins, suggests that the distribution of pY, pT, and pS sites is 1.8%, 11.8%, and 86.4%—very close to the original estimate for serine and threonine, but an order of magnitude higher for tyrosine. This apparent discrepancy is not due simply to tyrosine phosphorylation induced by growth-factor treatment because many EGF-regulated sites were also detected in the basal state. Instead, the likely reason is that tyrosine phosphorylation tends to occur on less abundant proteins compared to serine and threonine phosphorylation. On a per-protein basis, pY sites would therefore be underrepresented in the phosphoamino acid analysis. Another contributing factor may be the fact that pY is less stable in phosphoamino acid analysis than pS/pT, whereas we have found no evidence that the phosphopeptide enrichment method discriminates between these modification types.

We next averaged the time profiles of the upregulated phosphorylation sites and normalized them to their maximal fold change. Figure 2D shows that, on average, dynamic changes in tyrosine phosphorylation occur much faster and from a lower basal level compared to serine/threonine phosphorylation.

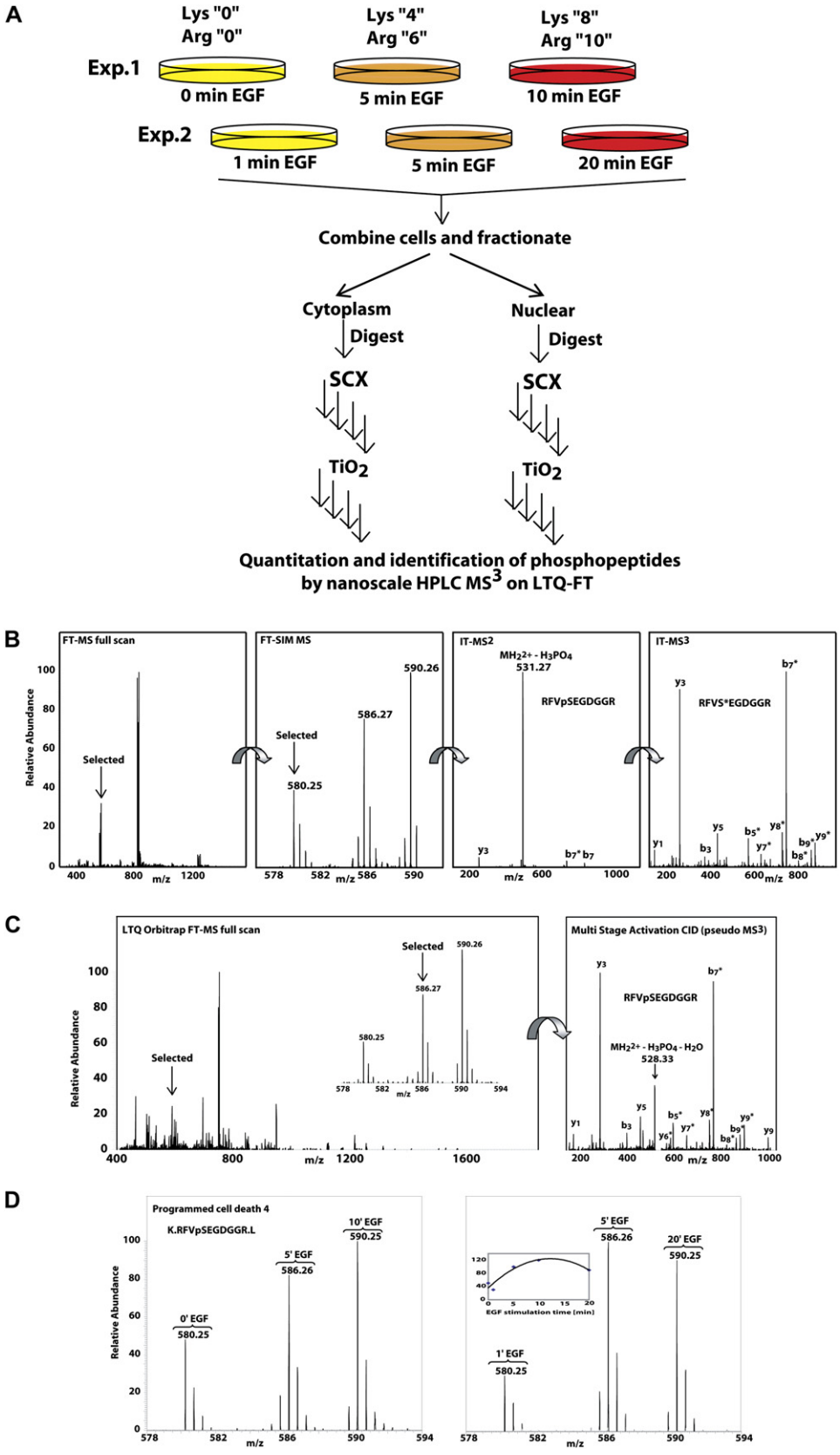
We found phosphorylation sites on a wide variety of proteins, and about half the proteins contained more than one site (Table S2). Figure 2E shows a Gene Ontology (GO) (Zeeberg et al., 2003) analysis of the in vivo phosphoproteome. GO localization information existed for 1,589 of the 2,244 phosphoproteins. Almost half of the phosphorylation events occurred on nuclear proteins, whereas only one-third of all proteins in the database were assigned as nuclear by GO, indicating that phosphorylation preferentially targets nuclear proteins. The proportion of regulated proteins closely paralleled the total number of phosphorylations of nuclear proteins. As expected, proteins annotated as extracellular were significantly underrepresented in the phosphoproteome. Although there is evidence of a mitochondrial phosphoproteome (Pagliarini and Dixon, 2006), proteins annotated as mitochondrial by GO were underrepresented, as were plasma membrane proteins. As membrane proteins are often of low abundance and we did not specifically enrich the plasma membrane in this analysis, we are less likely to identify phosphorylation events associated with them. Conversely, proteins associated with the cytoskeleton, which are generally of higher abundance, were almost twice as likely to be detected in the phosphoproteome as compared with their proportion in the database, and they accounted for an even higher share of the regulated sites. This illustrates that, at the current state of technology, our phosphopeptide screen is still not comprehensive. We may also miss certain phosphopeptides that are either too small or too large in mass to be readily measurable in our current MS setup. Therefore, absence of a phosphorylation site from our data set does not necessarily mean that it does not exist. Given that we have observed 2,244 phosphoproteins in a single cell line and taking into account the sensitivity and dynamic range limitations of current MS technology (de Godoy et al., 2006), it is clear that a large proportion of cellular proteins are phosphorylated.

We compared our data set with all annotated human phospho-sites in the SwissProt database from experimental data or inferred from homologous sequences (3,262 sites in version 48.0) and also included four previous phosphoproteomics data sets in our analysis (Amanchy et al., 2005; Beausoleil et al., 2004; Stover et al., 2004; Thelemann et al., 2005). We found that more than 90% of our sites were novel with respect to SwissProt. From the four previous data sets, we were able to map 1,890 sites onto IPI version 3.13. Of these, 691 (37%) were also found in our data set (Table S5). Given that the previous studies generally used lower-resolution MS and different experimental conditions, this is a relatively high value. As in the case of SwissProt, close to 90% of our data are novel compared with these previous studies. Taken together, our data suggest that, despite several decades of research into phosphorylation, most in vivo phosphorylation sites have still not been detected.

Phosphoproteome Dynamics

EGF signaling begins with activation of the EGF receptor and extends through a cascade of downstream kinases to mediate the increased phosphorylation of a large number of substrate proteins. The overall levels of phosphorylated proteins within a specific cellular compartment are also affected by other processes, including dephosphorylation by protein phosphatases, protein degradation via ubiquitination, and translocation of the protein to another subcellular location. Our experiment measures the net effect of all these diverse processes that collectively regulate the dynamic phosphoproteome.

To search for patterns in the time profiles of the regulated phosphopeptides, we explored several clustering techniques. Clustering algorithms have been applied extensively to the analysis of gene-expression data, including time-series data sets. We found fuzzy c-means (FCM) clustering to be better suited to our analysis than


A
Lys "0"
Arg "0"
Lys "4"
Arg "6"
Lys "8"
Arg "10"
Exp.1
0 min EGF
5 min EGF
10 min EGF
Exp.2
1 min EGF
5 min EGF
20 min EGF
Combine cells and fractionate
Cytoplasm
Digest
Nuclear
Digest
SCX
SCX
TiO2
TiO2
Quantitation and identification of phosphopeptides
by nanoscale HPLC MS3 on LTQ-FT
B
FT-MS full scan
FT-SIM MS
IT-MS2
IT-MS3
Relative Abundance
Selected
590.26
586.27
580.25
MH2 2+ - H3PO4
531.27
RFVpSEGDGGR
RFVS*EGDGGR
m/z
C
LTQ Orbitrap FT-MS full scan
Multi Stage Activation CID (pseudo MS3)
Selected
590.26
586.27
580.25
RFVpSEGDGGR
MH2 2+ - H3PO4 - H2O
528.33
D
Programmed cell death 4
K.RFVpSEGDGGR.L
0' EGF
580.25
5' EGF
586.26
10' EGF
590.25
1' EGF
580.25
5' EGF
586.26
20' EGF
590.25
EGF stimulation time (min)

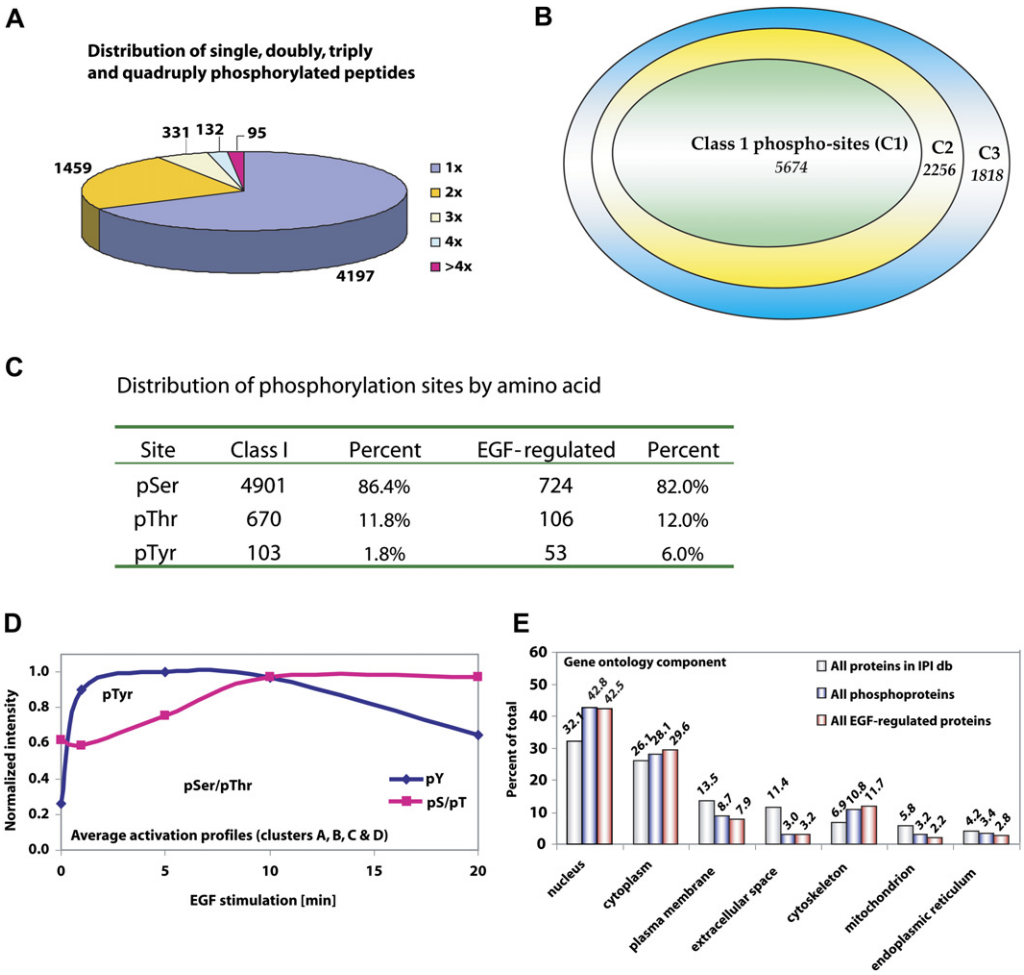


Figure 2. Results of Large-Scale Phosphopeptide Sequencing

(A) Distribution between singly and multiply phosphorylated peptides (some peptides were both singly and multiply phosphorylated and occur in more than one category).

(B) Phosphorylation-site analysis. See text for definition of classes I, II and III.

(C) Distribution of the pY/pS/pT phosphoproteome.

(D) Averaged and normalized time profile of all regulated tyrosine and serine/threonine phosphorylation sites.

(E) Gene Ontology (GO) analysis of the in vivo phosphoproteome.

hard partitioning algorithms, such as k-means, hierarchical clustering, or self-organizing maps (Figure 3 and Experimental Procedures). In FCM clustering, each profile is assigned a grade of membership for a set of clusters (represented by color in Figure 3). FCM offers robust clustering with regards to noise by variation of a fuzzification parameter m, which limits the contribution of ill-behaved profiles to the clustering process. We iteratively explored combinations of cluster sizes and fuzzification parameters and found optimal partitioning with c = 6 and m = 2. There

Figure 1. Quantitative and Time-Resolved Phosphoproteomics Using SILAC

(A) Three cell populations are SILAC encoded with normal and stable isotope-substituted arginine and lysine amino acids, creating three states distinguished by mass. Each population is stimulated for a different length of time with EGF, and the experiment is repeated to yield five time points. Cells are combined, lysed, and enzymatically digested, and phosphopeptides are enriched and analyzed by mass spectrometry.

(B) Mass spectra of eluting peptides reveal SILAC triplets (same peptide from the three cell populations), and these triplets are remeasured in selected ion monitoring (SIM) scans for accurate mass determination. Phosphopeptides are identified by loss of the phospho-group in a first fragmentation step followed by sequence-related information from a second fragmentation step.

(C) Same peptide as in (B) but measured on the LTQ-Orbitrap. Inset shows a magnification of the SILAC peptide selected for fragmenting. Right-hand panel shows the result of multistage activation of the peptide.

(D) Raw data of a phosphopeptide from the protein programmed cell death 4. The three peptide intensities in the two experiments are combined using the 5 min time point, resulting in the quantitative profile shown in the inset.

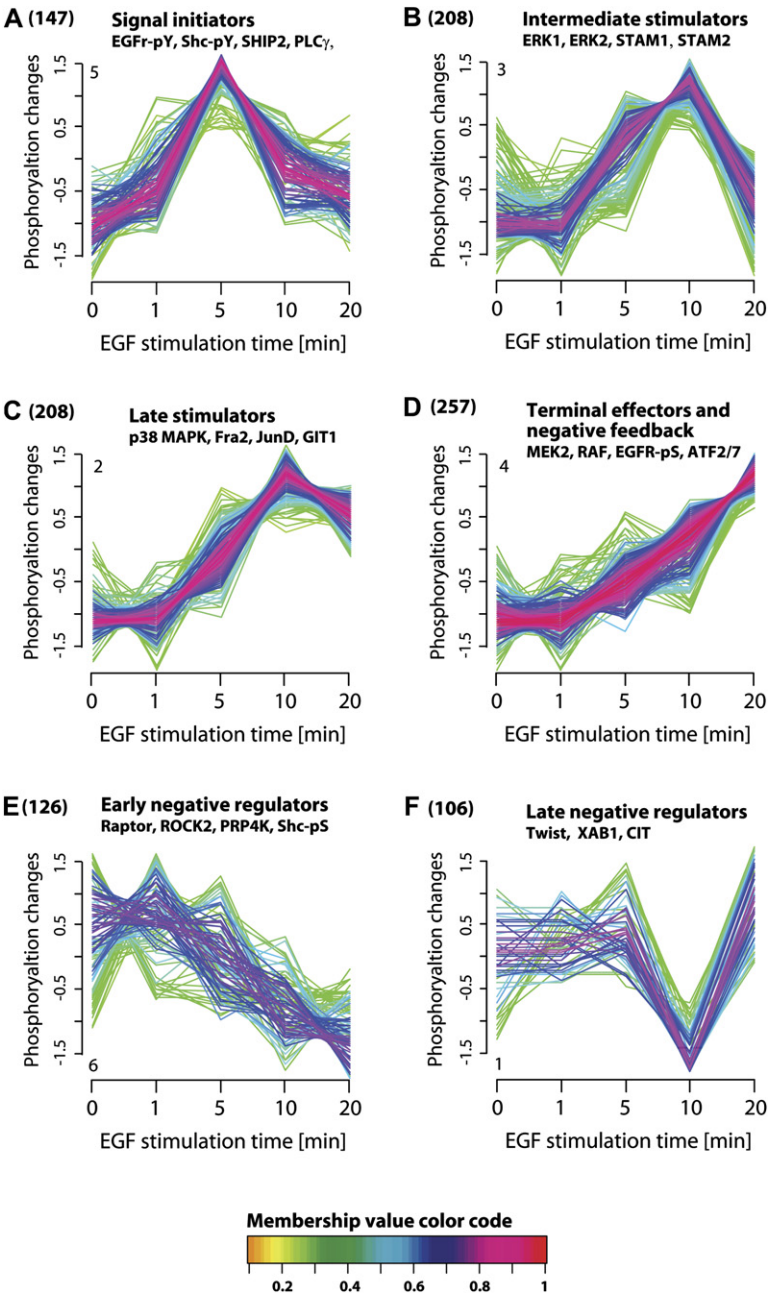


Figure 3. Clustering of Dynamic Phosphorylation Profiles

Temporal profiles were assigned to six clusters by fuzzy c-means clustering. The y axis is $\log_{10}$ transformed and normalized, and the number of phosphopeptides is given in parentheses. Each trace is color coded according to its membership value for the respective cluster (see color bar). Each cluster is designated by the function of prominent members. Examples of such members are given for each cluster.

are four clusters with upregulated phosphopeptides and two with downregulated ones.

Clusters turned out to contain functionally related members and were named to reflect their specific components. For example, cluster A was enriched in tyrosine-phosphorylated peptides involved in membrane-proximal signal events and was therefore termed "signal initiators." Similarly, cluster D was termed "negative feedback" because it contained phospho-sites from RAF, MEK2, and EGFR, which are each involved in downregulation of the EGF pathway (Hunter, 1998; Schlessinger, 2000).

We next investigated whether different phospho-sites within the same protein were regulated in the same manner. Interestingly, most proteins (77%) that had a regulated phosphopeptide also had at least one additional phosphopeptide that behaved differently—i.e., that was either unchanging or else in a different cluster. For instance, Figure 3 contains a tyrosine phosphorylation site of the adaptor protein Shc with the signal initiator time profile and a serine phosphorylation site in the "late negative regulators" cluster (Pelicci et al., 1992). As another example, we detected two serine phosphorylated peptides from

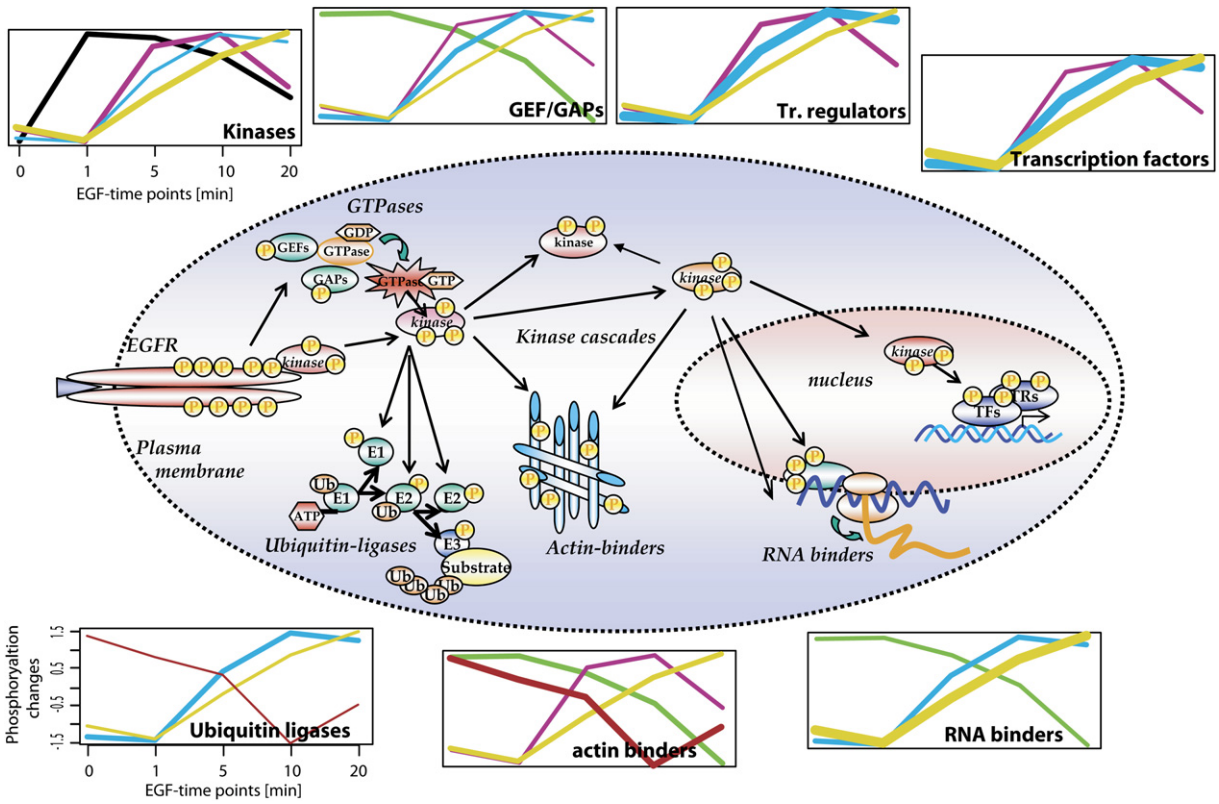


Figure 4. Functional Categories of Regulated Phosphoproteins

Regulated phosphoproteins fall into seven major functional categories. For each category, the consensus kinetic curves of its members (from Figure 3) are drawn in proportion to the number of phosphopeptides in that cluster. For additional data, see Table S6.

tuberin (TSC2) (Roux et al., 2004). One of these was unchanging, while the other had a "late stimulator" profile following EGF stimulation. Our data set shows that, in general, phosphorylation is regulated differently on different sites within the same protein. This finding suggests that the degree of phosphorylation should always be measured site specifically rather than for the protein as a whole in order to obtain accurate and functionally relevant understanding of activation kinetics. It also suggests that protein phosphorylation typically serves different functions on different sites of the protein, a fact that is already appreciated for a number of well-studied signaling proteins.

We grouped regulated phosphopeptides into functional protein classes in Table S6 and Figure 4. One class of regulated proteins comprises kinases and phosphatases (represented by 23 phosphopeptides), spanning the entire EGF cascade from the receptor through RAF and MEK to the MAP kinases (ERK1, ERK2, and p38). The main kinetic clusters in this functional class are initiators, activators, and terminators of EGF signaling, all showing increasing levels of phosphorylation during the 20 min time course of EGF treatment. Interestingly, transcriptional regulators made up a large class of regulated proteins. In addition to 26 different transcription factors with 34 regulated phosphopeptides, there were 20 transcriptional coregulators with EGF-dependent phosphorylation sites (Table 1). Traditionally, cellular signaling research has focused either on the early events in the cytoplasm or else directly on the resulting changes in transcriptional levels; thus, information about the transcription factors responsible for the observed gene-expression differences is lacking. Our study now reveals that the EGF signal spreads to at least 46 transcriptional regulators within the relatively short time frame of 20 min. Only a subset of these transcription factors have previously been known to be involved in growth-factor signaling, and some were not even known to be regulated by phosphorylation. WBRS9 and DAFT-1 are such examples, from which we identified seven phosphopeptides and two phosphopeptides, respectively (Table S2). However, only pS-189 in WBRS9 and pS-1456 in DAFT-1 increased upon EGF stimulation, implicating them as the regulatory sites (Table 1). Moreover, analysis of the regulated phosphorylation sites using Phosida (see below) indicated 33 novel phosphorylation sites from 19 different transcription factors. Both transcriptional regulators and coregulators display mainly late stimulator and "terminal effector" kinetic classes.

Apart from the canonical signaling axis leading to gene-expression changes, EGF also affects the cytoskeleton. Table S6 contains 27 regulated phosphopeptides of

Table 1. EGF-Regulated Transcription-Factor Phospho-Sites

IPI Number	Protein Name	Phosphopeptide Sequence	Cluster
IPI00234446	ATF-2	KMPLDLpSPLATPIIR	D
IPI00234446	ATF-2	NDSVIVADQpTPpTPTR	D
IPI00009975	ATF-7	ESSEPTGpSPAPVIQHSSATAPSNGLSVR	C
IPI00009975	ATF-7	TDSVIIADQpTPpTPTR	D
IPI00006079	Bclaf1	STFREEpSPLR	D
IPI00029795	CTF	KTEMDKpSPFNSPSPQDSPR	A
IPI00026673	CUTL1	RRHpSpSVSDSQPCEPPpSVGTEYSQGASPQPQHQLK	D
IPI00220159	DAFT-1	RNpSVERPAEPVAGAATPSLVEQQK	C
IPI00032936	ERF	RVpSSDLQHATAQLSLEHR	C
IPI00032936	ERF	TPADTGFAFPDWAYKPESpSPGSR	C
IPI00011593	Fra2	RSpSSSGDQSSDSLNpSPTLLAL	C
IPI00011593	Fra2	SHPYpSPLPGLASVPGHMALPRPGVIK	D
IPI00289547	Jun D	DEPQTVPDVPSFGEpSPPLpSPIDMDTQER	C
IPI00395737	LUZP1	RSpSSEGLSK	D
IPI00163729	MRTF-A	FGSTGSpTPPVpSPTPSER	B
IPI00163729	MRTF-A	pSPAAFHEQR	B
IPI00329152	MRTF-B	pSPAAFHEQIK	B
IPI00102820	MTSG1	NSGSFPpSPSIpSPR	F
IPI00033016	Myc	KFELLPTPPLpSPSRR	B
IPI00418606	NFIB	MYpSPICLTQDEFHPFIEALLPHVR	B
IPI00166491	PHF2	ALRPPTpSPGVFGALQNFK	C
IPI00166491	PHF2	RKGpSDDAPYpSPTAR	C
IPI00291638	RB1CC1	pSTELVLpSPDMPR	D
IPI00297694	RNF4	RLPQDHADSCVVpSpSDDEELSR	D
IPI00549473	Similar to ILF-1	pSAPApSPTHPGLMSPR	E
IPI00164672	SMIF	HAPTYTIPLpSPVLpSPTLPAEAPTAQVPPSLPR	D
IPI00063647	SPAG9	ERPISLGIFPLPAGDGLLpTPDAQK	C
IPI00030783	STAT5 (cytoplasm)	AVDGpYVKPQIK	A
IPI00030783	STAT5 (nuclear)	AVDGpYVKPQIK	C
IPI00414482	TF3C-alpha	NSSTDQGpSDEEGSLQK	C
IPI00414482	TF3C-alpha	RRApSWASENGETDAEGTQMTPAK	C
IPI00217957	TMBS62	ETRIpSFVEEDVHPK	C
IPI00069817	WBRS9	EDEGRREpSINDR	D
IPI00008137	Zfp295	EHAPLApSPVENK	D

actin-binding proteins, with diverse kinetic profiles. There were also 21 proteins interacting with small GTPases, many of them regulating cytoskeleton-associated functions, with late stimulator or late negative regulator time profiles (Bos, 2005; Soderling and Scott, 2006). Ubiquitination, like phosphorylation, is a diverse and widespread cellular control mechanism (Devoy et al., 2005; Dikic and Giordano, 2003; Waterman and Yarden, 2001). While ubiquitination was not measured directly in this experiment, a group of 12 ubiquitin ligases, including E1-, E2-, and E3-type enzymes, was dynamically phosphorylated following EGF stimulation and provide an interesting link between these two regulatory systems. Ubiquitin ligases had late stimulator or "early negative regulator" time profiles. Another intriguing class of proteins with very different time profiles was a diverse set of RNA-binding proteins

and RNA-processing factors (represented by 26 phosphopeptides). Figure 4 summarizes the variety of cellular functions directly influenced by EGF signaling.

Cellular Control by Dynamic Phosphorylation

A major advantage of following the temporal dynamics of phospho-sites instead of whole proteins is that signaling outcomes can be more directly connected to responsible upstream or downstream events. The phosphorylation of receptor molecules not only activates various signaling cascades but also deactivates the processes in later stages. These events are separated in time and were easily resolved in our analysis. The autophosphorylation of the EGF receptor on multiple tyrosine residues is the critical step initiating the propagation of the signal inside the cell (Hunter, 2000; Pawson and Nash, 2003; Schlessinger, 2000). In contrast, Ser/Thr phosphorylation often attenuates the signal by negative feedback (Countaway et al., 1989; Hunter, 1998; Schlessinger, 2000). Accordingly, we detected signal initiator profiles for all of the EGFR peptides containing Tyr1069, Tyr1092, Tyr1110, Tyr1138, Tyr1172, and Tyr1197, while phosphorylation of EGFR on Thr693, Ser991, and Thr993 instead showed delayed kinetics, peaking after 10 min (Figure 5A). Moreover, it was recently reported that, in addition to ERK1/2, p38 MAPK is also capable of phosphorylating EGFR at Thr693, which could account for its considerably higher phosphorylation at later time points (Winograd-Katz and Levitzki, 2006) (Figures 5A and 5B).

Fine-tuning mechanisms are not limited to receptors and may affect proteins at all levels of signaling pathways. Supporting this notion, we identified two novel in vivo phosphorylation sites in kinases within the canonical RAF-MEK-ERK cascade (Marshall, 1994), namely Ser186 on A-RAF and Ser23 on MEK2. The phosphorylation motifs—and more importantly, the kinetic profiles of these sites—clearly suggest negative feedback regulation by the downstream kinases ERK1/2 (Table S6 and Figure 3D). In a similar fashion, direct activation of protein kinases can be precisely monitored by the phosphorylation in their activation loops (Karin and Hunter, 1995; Marshall, 1994). For example, we observed that EGF stimulation of HeLa cells leads to more rapid and transient activation of both ERK1 and ERK2 as compared with the delayed and prolonged activation of p38 MAPK (Figure 5B).

Another general principle commonly used in cellular signaling is the phosphorylation-dependent translocation of proteins such as kinases and transcriptional regulators into the nucleus. Signal transducer and activator of transcription 5 (STAT5) is a latent transcription factor retained in its inactive state in the cytosol. Upon growth-factor stimulation, STAT5 is activated by phosphorylation on Tyr694, leading to its dimerization and nuclear translocation (Bromberg and Darnell, 2000). This translocation was readily apparent in our analysis through the dynamic phospho-profiles of Tyr694 of STAT5 in the cytosol and the nucleus (Figure 5C). We observed rapid phosphorylation of STAT5 in the cytosolic fractions, peaking as early as the first minute of EGF stimulation. In the nuclear fraction, however, an increasing profile was detected for the same phosphopeptide, with an accumulation rate mirroring its decreasing amounts in the cytosol (Figure 5C).

Our data identify a unique combination of in vivo signaling components and their relevant regulatory sites. This could be particularly useful for deciphering the action of complex transcription factors, as illustrated below for activator protein 1 (AP-1), a dimeric complex composed of combinations of members of the Jun, Fos, and ATF families of transcription factors. The dimer partner composition of AP-1 determines differential DNA binding-site specificity and as a result regulates the expression of a distinct subset of genes (Shaulian and Karin, 2002; van Dam and Castellazzi, 2001). Moreover, MAPK-dependent phosphorylation of members from all three families has been shown to regulate AP-1 activity (Karin et al., 1997; Murphy et al., 2002; Ventura et al., 2003). Here we provide evidence that EGF-dependent gene regulation in HeLa cells involves at least four members of the AP-1 complex, namely JunD, Fra2, ATF2, and ATF7 (Figures 5D–5F). Although we identified five phosphorylation sites on Fra2, only the peptide containing pSer308 and pSer320 changed significantly, associating these two sites with relevant regulatory activity (Figure 5D). Likewise, we have identified activating phosphorylations in JunD, while the corresponding phospho-sites in JunB did not change during the course of stimulation (Figure 5E). These results indicate that JunD plays a more prominent role downstream of EGFR than its close relative JunB. For the remaining family of AP-1-related transcription factors, we have observed three regulatory sites for ATF2 and three for ATF7. Only the ATF2 phosphorylation on Thr69 and Thr71 by p38 MAPK, which leads to enhanced AP-1 activity (Karin et al., 1997; Raingeaud et al., 1996), has previously been described. The phosphorylation of the corresponding sequence in ATF7 exhibited similar delayed kinetics strictly following the activation profile of p38 (Figure 5F). Thus, our results not only position ATF7 downstream in EGF-stimulated signaling cascades and provide its regulatory sites but also point toward p38 MAPK as the kinase responsible for their phosphorylation.

Phosida Database

The current data set comprises more phosphorylation sites than all previous studies combined. To enable efficient use of these data by the scientific community, we have created a database called Phosida, for *pho*sphorylation *si*te *da*tabase (http://www.phosida.com). Phosida lists phosphorylation sites associated with particular projects and proteomes or, alternatively, displays phosphorylation sites found for any protein or protein group of interest (Figure 6). The data are crossreferenced to information in SwissProt and the International Protein Index (IPI) database. Importantly, Phosida links extensive mass spectrometric peptide information to the phosphorylation sites, such as identical sites on several peptides with the same temporal profiles in response to stimulus (EGF stimulation,

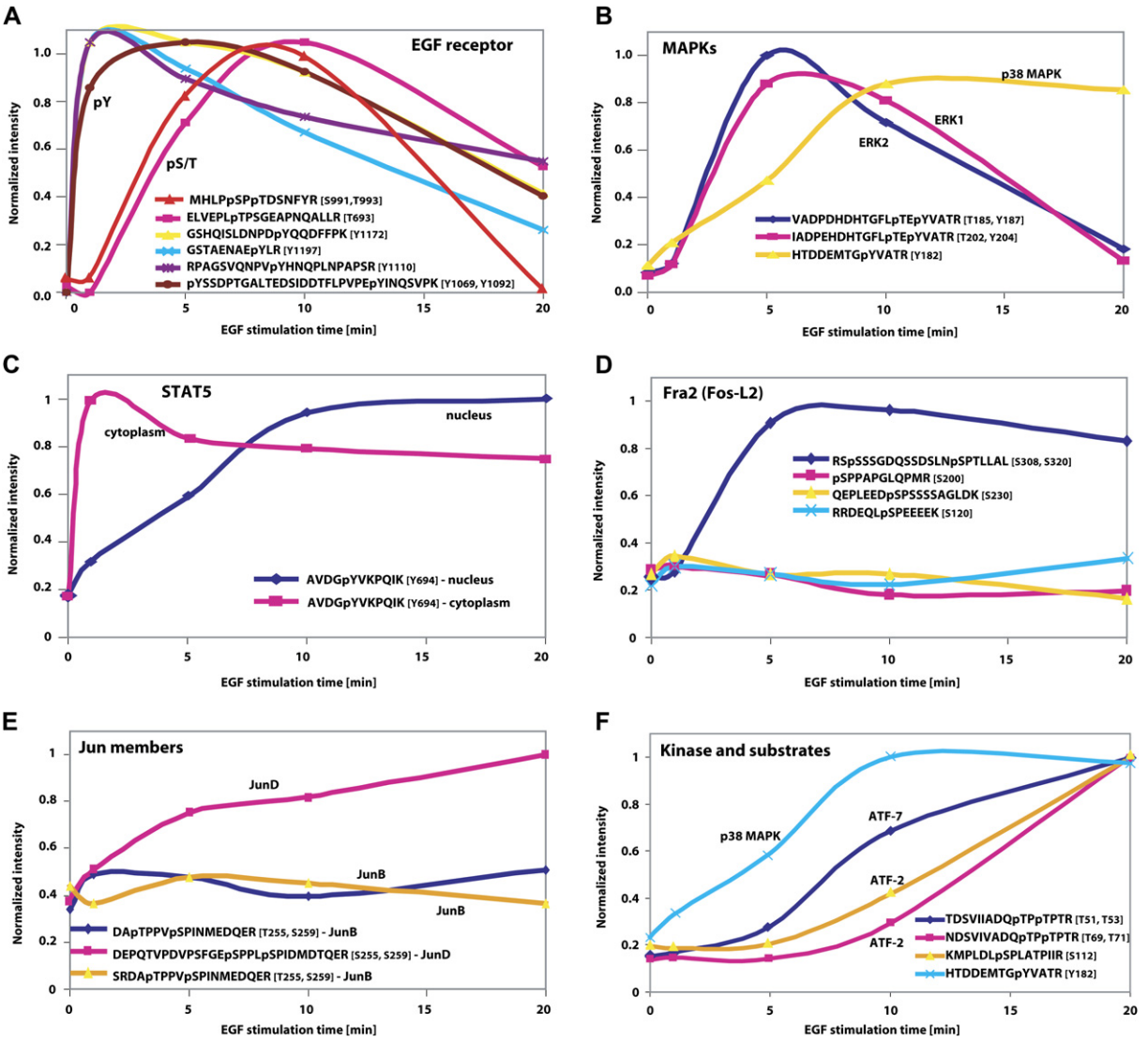


Figure 5. Regulatory Information from Specific Phosphorylation Sites

(A) EGF receptor and negative feedback. Tyrosine phosphorylation sites (pY) have fast kinetics, whereas serine/threonine phosphorylation (pS, pT) occurs with a time delay.

(B) Activation profiles from the mitogen-activated protein kinase (MAPK) family.

(C) Translocation of activated STAT5 from the cytosol to the nucleus.

(D) Multiply phosphorylated transcription factor Fra2. Only one of the phosphopeptides is regulated by EGF.

(E) Related but differentially regulated transcription factors. Both JunD and JunB are phosphorylated at the paralogous sites, but only in JunD are these sites regulated (shown by overlapping peptides for the Jun B site, yellow and violet traces).

(F) Profile for a kinase and its known substrates. p38 starts to phosphorylate its substrates shortly after its own phosphorylation and activation.

in this case). Furthermore, we have submitted our phosphorylation data to the Phospho.ELM database (Diella et al., 2004) and the Human Proteome Reference Database (Peri et al., 2004) and have uploaded previous phosphoproteome data sets to Phosida.

DISCUSSION

We have developed and applied a strategy combining triple-encoding SILAC for monitoring activation profiles, SCX and TiO_2 chromatography for phosphopeptide enrichment, and high-accuracy mass spectrometric characterization. Identification of numerous phosphorylation sites of transcription factors and other low-abundance regulatory proteins demonstrates that the technology can probe the phosphoproteome in considerable depth. The approach is completely generic for identification of key phosphorylation events in signaling pathways and is applicable to any cell culture system that can be SILAC labeled. It can also be used to study crosstalk between

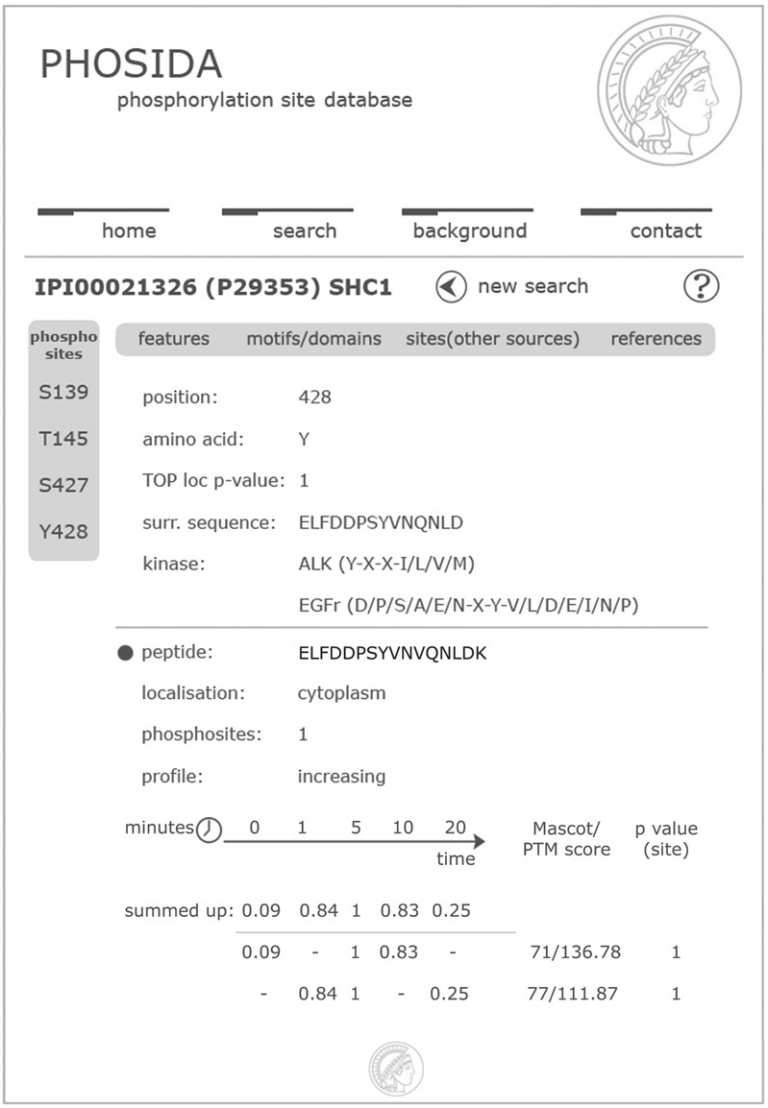


Figure 6. The Phosida Database

Phosida information for the phosphopeptides of Shc, including matching kinase motifs, localization p value, cellular compartment where the peptide was found, and data for its phosphorylation dynamics. Information for the tyrosine-phosphorylated peptide ELFDDP-SpYVNVQNLDK is displayed. The database is freely accessible at http://www.phosida.com.

signals, perturbations such as drug treatment, and knockouts or siRNA-induced knockdowns of signaling molecules. While acquisition and especially analysis of the data is currently very time consuming, advances in automation should make quantitative phosphoproteomics easier to perform in the future.

A global, unbiased view of the in vivo phosphoproteome reveals that a large proportion of cellular proteins are phosphorylated and that only a small subset of total phosphorylation sites are regulated in response to a stimulus. The observation that individual phospho-sites on a protein are typically regulated differently suggests that proteins generally serve as integrating platforms for a variety of incoming stimuli. This integration of signals could be independent, with phosphorylation of each site occurring separately from the others, or it could be dependent, for example, on when a "priming site" is necessary for subsequent phosphorylation events.

While phosphorylation is not the only regulatory system in the cell, it is intimately tied to many other systems, as demonstrated here by identification of regulatory sites on ubiquitin ligases, GEFs, actin-binding proteins, RNA-interacting proteins, and other important protein classes. Our data also pinpoint a large number of transcriptional regulators or coregulators to which the growth-factor stimulus is distributed.

Future studies could connect these data directly to downstream gene-expression changes as measured by microarray technology and/or to protein-expression changes as measured by quantitative proteomics. Integrated with other large-scale data sets, quantitative phosphoproteomics should provide an interesting foundation for a system-wide modeling of cell signaling events.

Perhaps of more immediate use, it contains a treasure trove of detailed and time-resolved information about numerous individual signaling events controlled by

phosphorylation. For example, debate has raged for years as to whether Grb2, a key adaptor protein, undergoes tyrosine phosphorylation (Li et al., 2001). Here we identified two phosphorylation sites on Grb2 with certainty: Tyr169 and Tyr209, both located within the carboxy-terminal SH3 domain (see Table S2). The different kinetic profiles of the two sites now support the notion that efficient Sos dissociation from the activated EGFR-Grb2-Sos complex requires combinatorial Grb2 phosphorylation (Blagoev et al., 2004; Li et al., 2001). In a similar way, such data sets may greatly accelerate cell signaling research by providing scientists with critical information about regulatory sites and their dynamics on various kinases, adaptors, transcriptional regulators, and other key signaling molecules. Only a limited number of kinase mutations have been found in cancers (Stephens et al., 2005), indicating that changes in activity are more often the cause of disease progression and suggesting that large-scale and quantitative phosphorylation screens could be used to monitor cancer development.

EXPERIMENTAL PROCEDURES

Cell Culture, Fractionation, and Peptide Preparation

Serum-starved HeLa cells (human cervix epithelial adenocarcinoma cells) labeled with L-arginine and L-lysine, L-arginine-U-$^{13}C_6{}^{14}N_4$ and L-lysine-$^{2}H_4$, or L-arginine-U-$^{13}C_6$-$^{15}N_4$ and L-lysine-U-$^{13}C_6$-$^{15}N_2$ (6 × 15 cm dishes per condition; ~95% confluent cells) were treated with 150 ng/ml of EGF for 0 min, 5 min, and 10 min. A second, identically labeled set of HeLa cells was treated with EGF for 1 min, 5 min, and 20 min. Otherwise, conditions were similar to those described in Blagoev et al. (2004). Mixed cells were centrifuged and cell membranes were disrupted using a Dounce glass homogenizer. Nuclear pellets were separated from cytoplasm by centrifugation.

Protein digestion was performed essentially as in Foster et al. (2003), strong-cation exchange chromatography of digests as in Gruhler et al. (2005), and TiO_2 enrichment of phosphopeptides essentially as in Larsen et al. (2005). Details are given in Supplemental Experimental Procedures.

Mass Spectrometric Analysis

Titansphere eluates were analyzed by online C_{18} reversed-phase nanoscale liquid chromatography-tandem mass spectrometry essentially as described in Gruhler et al. (2005) with a few modifications. Briefly, experiments were performed on an Agilent 1100 nanoflow system (Agilent Technologies) connected to a 7 tesla Finnigan LTQ-FT (Thermo Electron, Bremen, Germany) equipped with a nanoelectrospray ion source (Proxeon Biosystems, Odense, Denmark). The mass spectrometer was operated in the data-dependent mode to automatically switch between MS, MS^2, and neutral loss-dependent MS^3 acquisition.

Data-dependent settings were chosen to trigger an MS^3 scan when a neutral loss of 97.97, 48.99, or 32.66 Da was detected among the ten most intense fragment ions. Former target ions selected for MS^2 were dynamically excluded for 60 s. Total cycle time was approximately 3 s. In a second set of experiments, a hybrid linear ion trap/orbitrap instrument was used (LTQ-Orbitrap, Thermo Electron). The instrument was operated with the "lock mass" option as recently described (Olsen et al., 2005). Survey spectra were acquired with a resolution of 60,000 in the orbitrap while acquiring up to five tandem mass spectra in the LTQ part of the instrument. During fragmentation, the neutral loss species at 97.97, 48.99, or 32.66 m/z below the precursor ion were activated in turn for 30 ms each (pseudo-MS^3; see Schroeder et al., 2004).

Assigning Peptide Sequences Using MASCOT and MSQuant

Raw MS^2 and MS^3 spectra were centroided and merged into a single peak-list file and searched against the human IPI protein database versions 3.04 and 3.07, the former of which contains 48,968 protein sequences including known nonhuman contaminants such as porcine trypsin and *Achromobacter lyticus* endoproteinase Lys-C.

Statistical Analysis of Assigned Peptide Sequences and Quantitation

To establish a cutoff score threshold for a false-positive rate of less than one percent ($p < 0.01$), we performed a MASCOT search against a concatenated target/decoy database (Elias et al., 2005) consisting of a combined forward and reverse version of the IPI human database. All spectra and all sequence assignments made by MASCOT were imported into MSQuant (http://msquant.sourceforge.net). Here, the assignment of individual phosphorylation sites in MS^2 and MS^3 spectra were automatically scored using an algorithm similar to the one we have published previously for MS^3 spectra scoring (Olsen and Mann, 2004). The scoring is probability based and makes use of the four most intense fragment ions per 100 m/z units. For the localization PTM score from MS^2 spectra, fragment ions are expected to retain the phospho-group (+80 Da), whereas in neutral loss-triggered MS^3 spectra, the fragments are matched as $-H_2O$ ($+HPO_3 - H_3PO_4$; −18 Da). The algorithm calculates the putative b and y ions in the observed mass range for all possible combinations of phosphorylation sites within a peptide sequence and determines the number of matches, k.

The PTM localization probability score is $-10 \times \log_{10}(p)$, where the probability p is calculated as

$$p = (k!/[n!(n-k)!] \cdot [p^k] \cdot [(1-p)^{(n-k)}])$$
$$= (k!/[n!(n-k)!] \cdot [0.04^k] \cdot [0.96^{(n-k)}]),$$

where n is the total number of possible b and y ions (see Figures S3 and S6 for examples).

Phosphopeptide quantitation was likewise performed with the help of MSQuant. For each identified SILAC triplet, MSQuant calculated the three XIC values, and the assignments used for quantitation were visually displayed and validated. XICs for the first and last member of the triplet were normalized with respect to the common 5 min point of stimulation, which was scaled to one.

Clustering of Regulated Phosphopeptides

The raw ratios for the time profiles of regulated phosphopeptides were $\log_{10}$ transformed and then normalized so that, for each profile, the mean was zero and standard deviation was one. The normalization of data ensures that phosphopeptides with similar temporal patterns are close in Euclidean space. The transformed profiles were then clustered using the Mfuzz toolbox (Futschik and Carlisle, 2005), which is based on the open-source statistical language R (RDC Team, 2006). We used the fuzzy c-means (FCM) clustering algorithm, which is part of the toolbox. FCM clustering is a soft partitioning clustering method that requires two main parameters (c = number of clusters, m = fuzzification parameter) and uses Euclidean distance as the distance metric. FCM assigns to each profile a membership value in the range [0,1] for each of the c cluster. The algorithm iteratively assigns the profile to the cluster with the nearest cluster center while minimizing an objective function. Parameter m plays an important role in deriving robust clusters that are not greatly influenced by noise and random artifacts in data. For our analysis, the optimal values of c and m were derived by the iterative refinement procedure as described in Futschik and Carlisle (2005). The final clustering was done with the parameters c = 6 and m = 2.

Supplemental Data

Supplemental Data include Supplemental Experimental Procedures, Supplemental References, six figures, and six tables and can be found with this article online at http://www.cell.com/cgi/content/full/127/3/635/DC1/.

ACKNOWLEDGMENTS

We thank other members of the Center for Experimental BioInformatics (CEBI) and the Department of Proteomics and Signal Transduction at the Max-Planck Institute for Biochemistry for help and fruitful discussions, especially Sonja Krüger, Marcus Krüger, Gaby Sowa, and Cuiping Pan for experimental assistance. Angus Lamond, Leonard Foster, and Shao-En Ong provided important comments on the manuscript. TiO_2 spheres were a kind gift from GL Sciences (Tokyo). Alexandre Podtelejnikov and Søren Gade of Proxeon Biosystems A/S are acknowledged for their help with bioinformatic analysis using ProteinCenter. Work at CEBI is supported by a generous grant from the Danish National Research Foundation.

Received: June 16, 2006
Revised: August 1, 2006
Accepted: September 19, 2006
Published: November 2, 2006

REFERENCES

Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. Nature *422*, 198–207.

Amanchy, R., Kalume, D.E., Iwahori, A., Zhong, J., and Pandey, A. (2005). Phosphoproteome analysis of HeLa cells using stable isotope labeling with amino acids in cell culture (SILAC). J. Proteome Res. *4*, 1661–1671.

Andersen, J.S., Lam, Y.W., Leung, A.K., Ong, S.E., Lyon, C.E., Lamond, A.I., and Mann, M. (2005). Nucleolar proteome dynamics. Nature *433*, 77–83.

Beausoleil, S.A., Jedrychowski, M., Schwartz, D., Elias, J.E., Villen, J., Li, J., Cohn, M.A., Cantley, L.C., and Gygi, S.P. (2004). Large-scale characterization of HeLa cell nuclear phosphoproteins. Proc. Natl. Acad. Sci. USA *101*, 12130–12135.

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. Nat. Biotechnol. *21*, 315–318.

Blagoev, B., Ong, S.E., Kratchmarova, I., and Mann, M. (2004). Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. Nat. Biotechnol. *22*, 1139–1145.

Bos, J.L. (2005). Linking Rap to cell adhesion. Curr. Opin. Cell Biol. *17*, 123–128.

Bromberg, J., and Darnell, J.E., Jr. (2000). The role of STATs in transcriptional control and their impact on cellular function. Oncogene *19*, 2468–2473.

Chen, W.G., and White, F.M. (2004). Proteomic analysis of cellular signaling. Expert Rev. Proteomics *1*, 343–354.

Cohen, P. (2001). The role of protein phosphorylation in human health and disease. The Sir Hans Krebs Medal Lecture. Eur. J. Biochem. *268*, 5001–5010.

Countaway, J.L., Northwood, I.C., and Davis, R.J. (1989). Mechanism of phosphorylation of the epidermal growth factor receptor at threonine 669. J. Biol. Chem. *264*, 10828–10835.

de Godoy, L.M., Olsen, J.V., de Souza, G.A., Li, G., Mortensen, P., and Mann, M. (2006). Status of complete proteome analysis by mass spectrometry: SILAC labeled yeast as a model system. Genome Biol. *7*, R50.

Devoy, A., Soane, T., Welchman, R., and Mayer, R.J. (2005). The ubiquitin-proteasome system and cancer. Essays Biochem. *41*, 187–203.

Diella, F., Cameron, S., Gemund, C., Linding, R., Via, A., Kuster, B., Sicheritz-Ponten, T., Blom, N., and Gibson, T.J. (2004). Phospho.ELM: a database of experimentally verified phosphorylation sites in eukaryotic proteins. BMC Bioinformatics *5*, 79.

Dikic, I., and Giordano, S. (2003). Negative receptor signalling. Curr. Opin. Cell Biol. *15*, 128–135.

Elias, J.E., Haas, W., Faherty, B.K., and Gygi, S.P. (2005). Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. Nat. Methods *2*, 667–675.

Ficarro, S.B., McCleland, M.L., Stukenberg, P.T., Burke, D.J., Ross, M.M., Shabanowitz, J., Hunt, D.F., and White, F.M. (2002). Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. Nat. Biotechnol. *20*, 301–305.

Foster, L.J., De Hoog, C.L., and Mann, M. (2003). Unbiased quantitative proteomics of lipid rafts reveals high specificity for signaling factors. Proc. Natl. Acad. Sci. USA *100*, 5813–5818.

Futschik, M.E., and Carlisle, B. (2005). Noise-robust soft clustering of gene expression time-course data. J. Bioinform. Comput. Biol. *3*, 965–988.

Gruhler, A., Olsen, J.V., Mohammed, S., Mortensen, P., Faergeman, N.J., Mann, M., and Jensen, O.N. (2005). Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway. Mol. Cell. Proteomics *4*, 310–327.

Hunter, T. (1998). The Croonian Lecture 1997. The phosphorylation of proteins on tyrosine: its role in cell growth and disease. Philos. Trans. R. Soc. Lond. B Biol. Sci. *353*, 583–605.

Hunter, T. (2000). Signaling–2000 and beyond. Cell *100*, 113–127.

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

Karin, M., and Hunter, T. (1995). Transcriptional control by protein phosphorylation: signal transmission from the cell surface to the nucleus. Curr. Biol. *5*, 747–757.

Karin, M., Liu, Z., and Zandi, E. (1997). AP-1 function and regulation. Curr. Opin. Cell Biol. *9*, 240–246.

Kratchmarova, I., Blagoev, B., Haack-Sorensen, M., Kassem, M., and Mann, M. (2005). Mechanism of divergent growth factor effects in mesenchymal stem cell differentiation. Science *308*, 1472–1477.

Larsen, M.R., Thingholm, T.E., Jensen, O.N., Roepstorff, P., and Jorgensen, T.J. (2005). Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. Mol. Cell. Proteomics *4*, 873–886.

Li, S., Couvillon, A.D., Brasher, B.B., and Van Etten, R.A. (2001). Tyrosine phosphorylation of Grb2 by Bcr/Abl and epidermal growth factor receptor: a novel regulatory mechanism for tyrosine kinase signaling. EMBO J. *20*, 6793–6804.

Makarov, A., Denisov, E., Kholomeev, A., Balschun, W., Lange, O., Strupat, K., and Horning, S. (2006). Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. Anal. Chem. *78*, 2113–2120.

Marshall, C.J. (1994). MAP kinase kinase kinase, MAP kinase kinase and MAP kinase. Curr. Opin. Genet. Dev. *4*, 82–89.

Mumby, M., and Brekken, D. (2005). Phosphoproteomics: new insights into cellular signaling. Genome Biol. *6*, 230.

Murphy, L.O., Smith, S., Chen, R.H., Fingar, D.C., and Blenis, J. (2002). Molecular interpretation of ERK signal duration by immediate early gene products. Nat. Cell Biol. *4*, 556–564.

Olsen, J.V., and Mann, M. (2004). Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation. Proc. Natl. Acad. Sci. USA *101*, 13417–13422.

Olsen, J.V., de Godoy, L.M., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S., and Mann, M. (2005). Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. Mol. Cell. Proteomics *4*, 2010–2021.

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.

Pagliarini, D.J., and Dixon, J.E. (2006). Mitochondrial modulation: reversible phosphorylation takes center stage? Trends Biochem. Sci. *31*, 26–34.

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

Pelicci, G., Lanfrancone, L., Grignani, F., McGlade, J., Cavallo, F., Forni, G., Nicoletti, I., Pawson, T., and Pelicci, P.G. (1992). A novel transforming protein (SHC) with an SH2 domain is implicated in mitogenic signal transduction. Cell *70*, 93–104.

Peri, S., Navarro, J.D., Kristiansen, T.Z., Amanchy, R., Surendranath, V., Muthusamy, B., Gandhi, T.K., Chandrika, K.N., Deshpande, N., Suresh, S., et al. (2004). Human protein reference database as a discovery resource for proteomics. Nucleic Acids Res. *32*, D497–D501.

Ptacek, J., Devgan, G., Michaud, G., Zhu, H., Zhu, X., Fasolo, J., Guo, H., Jona, G., Breitkreutz, A., Sopko, R., et al. (2005). Global analysis of protein phosphorylation in yeast. Nature *438*, 679–684.

Raingeaud, J., Whitmarsh, A.J., Barrett, T., Derijard, B., and Davis, R.J. (1996). MKK3- and MKK6-regulated gene expression is mediated by the p38 mitogen-activated protein kinase signal transduction pathway. Mol. Cell. Biol. *16*, 1247–1255.

RDC Team (R Development Core Team). (2006). R: A Language and Environment for Statistical Computing (Vienna: R Foundation for Statistical Computing).

Roux, P.P., Ballif, B.A., Anjum, R., Gygi, S.P., and Blenis, J. (2004). Tumor-promoting phorbol esters and activated Ras inactivate the tuberous sclerosis tumor suppressor complex via p90 ribosomal S6 kinase. Proc. Natl. Acad. Sci. USA *101*, 13489–13494.

Rush, J., Moritz, A., Lee, K.A., Guo, A., Goss, V.L., Spek, E.J., Zhang, H., Zha, X.M., Polakiewicz, R.D., and Comb, M.J. (2005). Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. Nat. Biotechnol. *23*, 94–101.

Salomon, A.R., Ficarro, S.B., Brill, L.M., Brinker, A., Phung, Q.T., Ericson, C., Sauer, K., Brock, A., Horn, D.M., Schultz, P.G., and Peters, E.C. (2003). Profiling of tyrosine phosphorylation pathways in human cells using mass spectrometry. Proc. Natl. Acad. Sci. USA *100*, 443–448.

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

Schroeder, M.J., Shabanowitz, J., Schwartz, J.C., Hunt, D.F., and Coon, J.J. (2004). A neutral loss activation method for improved phosphopeptide sequence analysis by quadrupole ion trap mass spectrometry. Anal. Chem. *76*, 3590–3598.

Schulze, W.X., and Mann, M. (2004). A novel proteomic screen for peptide-protein interactions. J. Biol. Chem. *279*, 10756–10764.

Shaulian, E., and Karin, M. (2002). AP-1 as a regulator of cell life and death. Nat. Cell Biol. *4*, E131–E136.

Soderling, S.H., and Scott, J.D. (2006). WAVE signalling: from biochemistry to biology. Biochem. Soc. Trans. *34*, 73–76.

Songyang, Z., Blechner, S., Hoagland, N., Hoekstra, M.F., Piwnica-Worms, H., and Cantley, L.C. (1994). Use of an oriented peptide library to determine the optimal substrates of protein kinases. Curr. Biol. *4*, 973–982.

Stephens, P., Edkins, S., Davies, H., Greenman, C., Cox, C., Hunter, C., Bignell, G., Teague, J., Smith, R., Stevens, C., et al. (2005). A screen of the complete protein kinase gene family identifies diverse patterns of somatic mutations in human breast cancer. Nat. Genet. *37*, 590–592.

Stover, D.R., Cadwell, J., Marto, J.A., Root, K., Mestan, J., Stumm, M., Ornatsky, O., Orsi, C., Radosevic, N., Liao, L., et al. (2004). Differential phosphoprofiles of EGF and EGFR kinase inhibitor-treated human tumor cells and mouse xenografts. Clin. Proteomics *1*, 69.

Thelemann, A., Petti, F., Griffin, G., Iwata, K., Hunt, T., Settinari, T., Fenyo, D., Gibson, N., and Haley, J.D. (2005). Phosphotyrosine signaling networks in epidermal growth factor receptor overexpressing squamous carcinoma cells. Mol. Cell. Proteomics *4*, 356–376.

van Dam, H., and Castellazzi, M. (2001). Distinct roles of Jun: Fos and Jun: ATF dimers in oncogenesis. Oncogene *20*, 2453–2464.

Ventura, J.J., Kennedy, N.J., Lamb, J.A., Flavell, R.A., and Davis, R.J. (2003). c-Jun NH(2)-terminal kinase is essential for the regulation of AP-1 by tumor necrosis factor. Mol. Cell. Biol. *23*, 2871–2882.

Waterman, H., and Yarden, Y. (2001). Molecular mechanisms underlying endocytosis and sorting of ErbB receptor tyrosine kinases. FEBS Lett. *490*, 142–152.

Winograd-Katz, S.E., and Levitzki, A. (2006). Cisplatin induces PKB/Akt activation and p38(MAPK) phosphorylation of the EGF receptor. Oncogene. Published online June 19, 2006. 10.1038/sj.onc.1209737.

Yaffe, M.B., Leparc, G.G., Lai, J., Obata, T., Volinia, S., and Cantley, L.C. (2001). A motif-based profile scanning approach for genome-wide prediction of signaling pathways. Nat. Biotechnol. *19*, 348–353.

Zeeberg, B.R., Feng, W., Wang, G., Wang, M.D., Fojo, A.T., Sunshine, M., Narasimhan, S., Kane, D.W., Reinhold, W.C., Lababidi, S., et al. (2003). GoMiner: a resource for biological interpretation of genomic and proteomic data. Genome Biol. *4*, R28.

Zhang, Y., Wolf-Yadlin, A., Ross, P.L., Pappin, D.J., Rush, J., Lauffenburger, D.A., and White, F.M. (2005). Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. Mol. Cell. Proteomics *4*, 1240–1250.