

Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics

Blagoy Blagoev^{1,2}, Shao-En Ong^{1,2}, Irina Kratchmarova¹ & Matthias Mann¹

To study the global dynamics of phosphotyrosine-based signaling events in early growth factor stimulation, we developed a mass spectrometric method that converts temporal changes to differences in peptide isotopic abundance. The proteomes of three cell populations were metabolically encoded with different stable isotopic forms of arginine. Each population was stimulated by epidermal growth factor for a different length of time, and tyrosine-phosphorylated proteins and closely associated binders were affinity purified. Arginine-containing peptides occurred in three forms, which were quantified; we then combined two experiments to generate five-point dynamic profiles. We identified 81 signaling proteins, including virtually all known epidermal growth factor receptor substrates, 31 novel effectors and the time course of their activation upon epidermal growth factor stimulation. Global activation profiles provide an informative perspective on cell signaling and will be crucial to modeling signaling networks in a systems biology approach.

In growth factor signaling, stimulation of cell-surface receptors first triggers activation of the receptor itself and then of a large number of intracellular effector molecules^{1–3}. The stimulus is integrated with a host of other cellular processes, leading to cytoskeletal changes, activating transcriptional programs in the nucleus and ultimately resulting in cell proliferation, differentiation or motility. Classical signaling pathways and networks depict potential protein-protein interactions only in a static form. In the cell, these interactions are dynamic and occur in an ordered fashion. So far, there have been no experimental methods for learning about the dynamics of protein activation and interaction at the proteomic scale. Dynamic profiling using phosphospecific antibodies or fluorescence-based methods has had to be done one protein at a time, in different cell lines and different conditions, making it virtually impossible to obtain a global kinetic profile of interaction or activation. Large-scale mapping of phosphorylation sites has shown great promise^{4,5}, but has only given a partial picture of signaling dynamics. Here we address this important practical and conceptual problem by studying the temporal dimension of the epidermal growth factor receptor (EGFR) signaling network as a model system.

The critical step in early growth factor signaling is tyrosine phosphorylation of the receptor and subsequently of key substrates.

This leads to activation of these proteins and recruitment of signaling partners. The tyrosine phosphorylation status of any protein over time thus indicates its temporal involvement in the signaling process. We therefore devised a proteomics strategy to enable global, time-dependent analysis of the EGFR phosphotyrosine proteome (Fig. 1).

Quantitative proteomics is useful not only for the measurement of protein abundance levels^{6–8} but also for differential labeling of protein populations in functional experiments^{9,10}. Previously, our laboratory described a mass spectrometry-based approach termed stable isotope labeling by amino acids in cell culture (SILAC)^{11,12}. One cell population is grown in medium containing normal arginine and another population is grown in medium with ¹³C-substituted arginine until the arginine is completely incorporated into each protein in the cell. The only difference between the two proteomes is the form of arginine in the proteins, making their peptides distinguishable by mass spectrometry. Because the proteome is completely labeled at the start of the experiment, protein synthesis and degradation have no effect on measured signal intensities, unlike in the case of classical metabolic labeling with radioisotopes. Therefore, relative peak heights of the peptides are a direct measure of the relative amounts of the proteins.

So far, quantitative proteomic methods have been used to compare only two states, such as 'on-off' or disease versus normal. Here we introduce a second isotopic label, allowing simultaneous quantification of three cellular states, for example, different time points of growth factor treatment (Fig. 1a). Three populations of HeLa cells were grown in media containing different versions of arginine, leading to mass separation of the arginine-containing peptides by 6 or 10 Da (see Fig. 1). The cells were stimulated for different lengths of time with epidermal growth factor (EGF; 150 ng/ml), lysed and their combined cell extracts immunoprecipitated with anti-phosphotyrosine antibodies (4G10 and P-Tyr-100). Arginine-containing peptides generated by proteolytic cleavage of eluted proteins occur as triplets in the mass spectra (Fig. 1b). The intensity of each peak directly indicates the relative amount of the tyrosine-phosphorylated protein or its binding partners at the corresponding time point of growth factor treatment.

To obtain a higher-resolution time profile of the phosphotyrosine phosphoproteome, we combined two three-state experiments. The unstimulated cells were divided to serve as a common zero point for both three-state experiments, resulting in a combined time profile for 0, 1, 5, 10 and 20 min. All changes are expressed as fold-change relative

¹Center for Experimental Bioinformatics (CEBI), Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark. ²These authors contributed equally to this work. Correspondence should be addressed to M.M. (mann@bmb.sdu.dk).

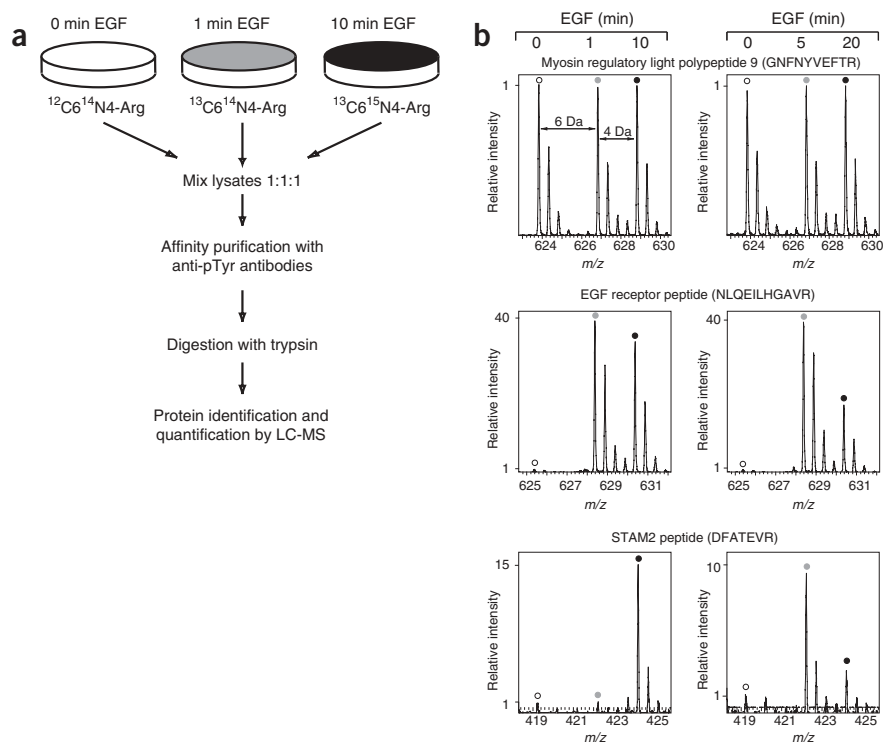


Figure 1 Determining the activation profile of the EGFR phosphotyrosine-dependent proteome. **(a)** Cells are grown in medium with one of three different forms of arginine such that this amino acid is completely incorporated into the proteome. Each population is stimulated by EGF for the indicated time interval and lysed. Lysates are mixed and tyrosine-phosphorylated proteins affinity purified, followed by tryptic digestion and mass spectrometric analysis. The experiment is then repeated for another triplet of time points to obtain a higher resolution profile. **(b)** Peptide mass spectra of three different proteins displaying distinct activation profiles. The peptide GNFNYVEFTR of myosin light chain in the upper left panel shows three isotope distributions with the same intensity (stimulation for 0, 1 and 10 min). The upper right panel shows the stimulation for 0, 5 and 20 min. This phosphoprotein was therefore basally phosphorylated and did not change at the five time points of EGF stimulation. The peptide NLQEILHGAVR of EGFR in the middle panels displays a rapid increase from the 0 time point (intensity set equal to one) through the 1- and 5-min points, followed by a decrease at the 10- and 20-min points. The lower panels show the peptide DFATEVR of STAM2. The increase is slower with a maximum between 10 and 15 min.

to the common zero point. This can lead to relatively large standard deviations for some proteins with low basal phosphorylation, but the curve shapes are still highly constrained, as discussed in Methods and the **Supplementary Figure 1** online.

We determined the activation profile of 202 proteins, precipitated by anti-phosphotyrosine antibodies, upon EGF stimulation. Of these, 81 showed a 1.5 fold or higher change in at least one time point and are listed in **Table 1**. Although the SILAC technique can quantify changes smaller than 10% (ref. 12), we chose 1.5 as a conservative cutoff for biological significance.

Signaling factors are frequently of very low abundance, making their detection and quantification a challenge to mass spectrometry. To estimate the coverage of the EGF tyrosine phosphoproteome achieved in our study, we first compared the identified proteins against the results of our previous proteomic screens in the EGF system^{13–15} and found that indeed all proteins were present. Second, we surveyed the literature for known effector molecules of EGF in early signaling events. We found 50 such proteins, and all appeared in our screen (marked in **Table 1**). These effectors have been determined by a wide variety of biochemical and genetic methods, many of which have no sensitivity bias. Third, in sensitivity-limited analysis of complex mixtures, many proteins are identified by only one peptide. However, we sequenced at least two peptides for all proteins listed in **Table 1**. Furthermore, peptides from these proteins were observed in both of the experiments used to construct the dynamic profile. The above evidence strongly suggests that we have covered an extensive part of the EGF-induced phosphoproteome.

Our approach retrieved proteins that were tyrosine phosphorylated by the activated EGF receptor as well as by downstream kinases. Tyrosine phosphorylation also triggers interactions of signaling factors that are not necessarily phosphorylated themselves. For example, the adaptor protein Grb2 was probably identified in our screen through its association with activated EGFR. As such proteins bind to phosphorylated partners in a stimulation-dependent manner and

play roles in the signaling cascade, we use the term ‘activation’ for these molecules as well.

We compared the profiles obtained by quantitative proteomics with those from standard western blot analyses. For this purpose, we grew and stimulated HeLa cells under the same experimental conditions used for the proteomic experiments up to the point of immunoprecipitation. Lysates derived from the different time points of stimulation were then immunoprecipitated separately with phosphotyrosine antibodies. We strove to obtain antibodies against proteins representing different classes of dynamic profiles for immunoblotting. As can be seen in **Figure 2** and **Table 1**, the western blots were consistent with the results from quantitative mass spectrometry for all the proteins tested, including for a novel protein we generated antibodies against.

Activation profiles for some of the key proteins in EGF signaling are presented in **Figure 3a,b**. The receptor itself is activated about 40-fold within 1 min, remains fully activated for 5 min and then slowly decreases to less than half the peak value at the last measured time point at 20 min. Eps15 and the E3 ubiquitin ligase c-Cbl show the same kinetics as the receptor, but with lower fold activation. These two proteins are known to be recruited early to the activated EGFR and to play an important role in early steps of receptor internalization^{16,17}. Because Cbl transfers the ubiquitin to EGFR, it is not surprising that the profile of ubiquitinated receptor (quantified from ubiquitin peptides derived from the same gel band as the receptor) is delayed compared with EGFR and Cbl. The ubiquitinated receptor is recognized by HRS and STAM/STAM2 (ref. 18), which are involved in the subsequent events of EGFR endosomal trafficking. Consistent with this role, their activation profile is delayed with respect to that of the ubiquitinated receptor (**Fig. 3a**).

In canonical growth factor signaling, the activated receptor recruits adaptor proteins that bring to the plasma membrane guanine exchange factors for small GTPases such as Ras, Rac and Rho, which then relay the signal to downstream effectors. For example, Ras signals to Raf and

Table 1 Proteins activated upon EGF stimulation

Protein name	1 min	5 min	10 min	20 min	Protein name	1 min	5 min	10 min	20 min
Upregulated proteins									
Known effectors of EGFR signaling									
EGFR	39	40	31.5	17.8	PLC-gamma	29.7	11	3.2	2.8
ErbB 2	35	30	18.5	9.7	PI 3-kinase, p85-alpha	1.3	1.8	1.6	1.2
Eps15	34	28.6	17.9	2.8	STAT 5	19.7	19.1	4.6	2.1
c-Cbl	23.6	20.5	8.8	3.7	Annexin I	9.3	10	2.7	1.7
Cbl-B	3.5	3.6	2.7	1.7	Vav-2	2.6	1.7	1.6	1.2
Ubiquitin	10.7	30.1	18.7	5.9	Vav-3	4.0	3.8	2.5	2.2
Hrs	1.6	25.4	33.2	7.6	Tubulin alpha-1	1.6	1.62	1.6	1.3
STAM	2.8	24.0	57.1	8.06	Tubulin beta-2	1.5	1.5	1.4	1.2
STAM 2	1.1	7.1	15.4	2.1	ARF GAP, GIT1	1.5	2.0	1.5	1.06
Shc	5.6	6.0	3.7	2.9	Srcasm	2.7	6.8	2.6	1.5
Grb 2	1.6	1.4	1.4	1.3	Ras GAP-like protein IQGAP1	1.7	2.0	1.06	1.1
Sos 1	1.6	1.6	0.78	0.76	AP-2, beta 1 subunit	1.5	2.1	1.9	1.9
Sos 2	1.7	1.8	1.2	1.4	Clathrin light chain A	0.9	1.2	1.6	1.4
Erk 2	1.8	8.6	6.1	1.7	ACK1	3.0	3.2	3.9	0.9
p38 MAP kinase	1.0	2.1	4.0	1.2	Acid phosphatase 1	2.1	2.84	2.6	1.94
p62Dok	18.1	10.2	4.7	1.0	Vimentin	2.1	0.92	1.9	0.7
DOC 2	13.3	16.0	4.2	3.2	Desmoglein 2	13.5	9.71	6.2	4.2
SHIP 2	34.0	19.1	5.3	1.9	Gamma-catenin	4.8	4.0	2.1	1.4
Odin	8.3	6.3	3.0	1.8	Proteins not previously reported in EGFR signaling				
hnRNP R	8.1	10	14.5	10	Nuclear domain 10 protein	3.1	7.8	4.8	2.1
hnRNP A2/B1	6.0	1.5	1.45	1.3	ADP, ATP carrier protein	1.7	1.77	1.5	1.32
hnRNP Q2	2.2	3.9	2.9	2.22	Sorcin	7.2	4.7	3.3	1.5
hnRNP JKTBP	1.1	1.5	1.6	1.2	N-cadherin	4.6	3.3	1.3	1.1
DEAD-box protein 3	22.8	32.7	21.1	15.58	WW domain binding protein 2	2.9	2.5	2.2	1.3
Major vault protein	1.7	2.5	1.1	0.97	HECT domain protein LASU1	2.6	9.8	5.1	1.95
Axl	1.2	1.4	1.6	1.5	PDCD6-interacting protein	2.0	8.8	37	2.8
Anthrax toxin receptor	1.2	1.7	5.3	6.6	Syntenin 1	1.4	2.1	5.1	3.3
Transferrin receptor 1	1.1	2.2	9.8	6.9	TOM1-like protein 2 ^a	3.9	14.6	8.5	1.7
Protein kinase PKX1	111	89	3.2	3.2	Ymer ^a	42	60	18.8	7.1
CYLD	23.9	22.2	17.1	2.9	LIM domains containing 1 ^a	2.4	2.0	1.3	1.1
ARF GAP, GIT2	1.19	2.1	1.4	1.1	FLJ21610 ^a	14.3	5.7	3.0	1.7
Centaurin delta 2	29.6	14.4	7.4	1.8	FLJ00269 ^a	3.2	1.7	1.22	0.7
Rho GEF 7	1.03	1.8	1.09	1.1	Downregulated proteins				
Known effectors of EGFR signaling									
ARP 2/3, subunit 1A	0.58	1.07	0.44	0.72	ARP2/3, 34 kDa subunit	0.72	1.04	0.51	0.76
ARP2/3, 20 kDa subunit	0.60	1.00	0.43	0.72	ARP2/3, 16 kDa subunit	0.80	1.02	0.54	0.71
ARP2/3, 21 kDa subunit	0.63	1.06	0.47	0.69	Gelsolin	0.71	0.99	0.66	0.80
Similar to ARP2/3, sub 5	0.60	1.06	0.49	0.78	p130Cas/BCAR 1	0.44	0.57	0.66	0.76
ARP2/3, 41 kDa subunit	0.79	1.05	0.49	0.74	Annexin A2	0.68	0.64	0.65	0.76
ARP3	0.61	1.06	0.46	0.70	Cytochrome c	0.59	0.22	0.15	0.11
ARP2	0.62	1.07	0.53	0.81	Proteins not previously reported in EGFR signaling				
AND-34/BCAR 3	0.67	0.70	0.82	0.81	Endothelial lectin HL-2	0.23	0.16	0.17	0.30
NDP kinase B	0.85	0.66	0.78	0.66	Swiprosin 1 ^a	0.59	1.03	0.61	0.94

The upper part of the table shows proteins with at least 1.5-fold activation and the lower part proteins with at least 1.5-fold downregulation. The table is further subdivided into known EGFR effectors and proteins not previously implicated. Fold activation relative to the zero time point (unstimulated cells) are given for the 1-, 5-, 10- and 20-min time points. See Methods for experimental details and **Supplementary Table 1** for standard deviations. ^aNovel molecules.

activates a mitogen-activated protein (MAP) kinase cascade that transmits the signal into the nucleus. Consistent with this sequence, a chronological ordering of dynamic profiles from EGFR via adaptors (Grb2, Shc) and guanine exchange factors (GEFs; Sos1/2) to MAP kinases was observed (**Fig. 3b**). Interestingly, Sos activation decays faster than Grb2 activation (**Fig. 3b**), which may correlate with negative feedback mechanisms dissociating Sos from the activated EGFR-Grb2 complex¹⁹. We also note that two MAP kinases, Erk and p38,

showed different activation profiles, although both have the same position in the signaling cascade and translocate into the nucleus (**Fig. 3b**).

One of the major effects of growth factor stimulation is the rearrangement of the actin cytoskeleton through small GTPases. **Figure 3c** shows the dynamics of Arp2/3, a complex responsible for the branching of actin microfilaments. Arp2/3 has a maximum at 5 min, followed by another increase at the 20-min time point. We quantified nine Arp2/3 members; strikingly, all display the same

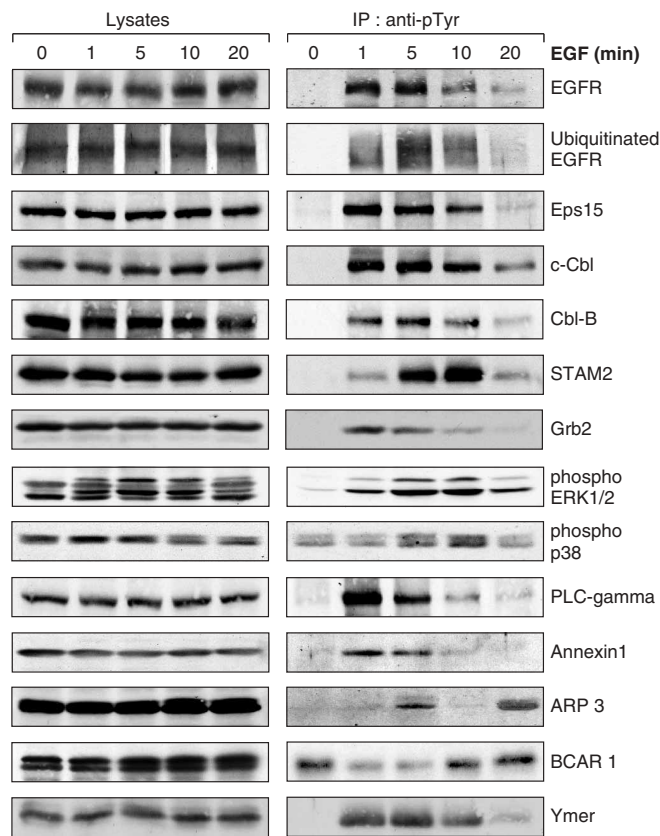


Figure 2 Western blot analysis of selected EGFR effectors. HeLa cells were stimulated with EGF for the indicated time intervals, matching the proteomics experiments. After immunoprecipitation with anti-phosphotyrosine antibodies and SDS-PAGE, blots were probed with antibodies against the specified proteins. For ERK1/2 and p38, cellular lysates were probed directly with phosphospecific antibodies. On the left hand side the cellular lysates were probed directly with the corresponding antibodies as a loading control, except for ubiquitin, where anti-EGFR antibody was used for this purpose. IP, immunoprecipitation.

quantitative behavior. Although the data do not indicate whether Arp2/3 subunits are directly tyrosine phosphorylated, it is clear from **Figure 3c** as well as from western blots (**Fig. 2**) that Arp2/3's involvement goes through at least two peaks in the time scale studied here. This is further supported by the very small standard deviation of quantification (**Supplementary Table 1** and **Supplementary Fig. 1c** online). Interestingly, Gelsolin, which is known to be tyrosine phosphorylated and to actively participate in actin dynamics, has the same profile as the Arp2/3 complex (**Fig. 3c**)²⁰. A novel protein, Swiprosin 1, exhibits the same dynamics (**Fig. 3c**), suggesting a potential role in growth factor-dependent actin remodeling.

Figure 3c also shows p130CAS and AND-34, two proteins with a unique profile of rapid decrease upon stimulation and slow recovery of activity over the 20-min period. Although AND-34 has not previously been tied to EGF signaling, both proteins are known to be tyrosine phosphorylated and can interact with each other²¹. Another shared feature is that they were both found in screens for antiestrogen resistance genes in breast cancer cells, where they were termed BCAR1 and BCAR3, respectively^{22,23}. Thus, common function can be reflected in common activation profiles.

Almost all of the known proteins in **Table 1** are readily associated with growth factor signaling. One exception is a group of six RNA-

binding proteins, with different levels and kinetics of upregulation (**Table 1**). Although the role of RNA-associated proteins has not usually been studied in this context, connections to the regulation of mRNA stability have been observed²⁴. Four of the six RNA binding proteins in **Table 1** are heterogeneous nuclear ribonucleoproteins, one is an RNA helicase, and the final one, major vault protein, is a component of a large cytosolic RNP complex.

Six of the proteins in **Table 1** are novel, with no associated functional information. We already connected Swiprosin 1 with actin rearrangement through its time profile (see above). The activation profiles of the other five are shown in **Figure 3d**. One of these is predicted to be an atypical protein kinase, and its activation profile associates it with early events at the plasma membrane. Another novel protein contains three LIM domains. LIM domain proteins have been found in proteins that associate with actin filaments²⁵, and, in the example of the LIM domain-containing protein Enigma, shown to associate with receptor tyrosine kinases²⁶. Because this protein also showed the same early activation profile, it may be involved in early actin remodeling. A third protein with an early activation profile contains an ATPase domain but no other recognizable domain structure or obvious homology. TOM1-like protein 2 is a novel protein named for its homology to TOM1-like protein 1/Srcasm, which interacts with several adapter proteins in a phosphotyrosine-dependent manner²⁷. Like Srcasm, the novel protein has a VHS domain, which implicates it in endosomal trafficking²⁸. Further support for this functional context comes from the fact that the time profile for Hrs is very similar to that of the novel protein (**Fig. 3a,d**). The final novel protein did not contain any recognizable domains or homology to known proteins. We raised an antibody against it, and the dynamic profile observed by western blotting (**Fig. 2**) fully correlates with the time profile obtained by quantitative proteomics (**Fig. 3d**). Continuing our tradition of naming growth-factor substrates after figures from Nordic mythology, we termed this molecule Ymer.

This single unbiased study reveals virtually the entire known set of tyrosine-phosphorylated effectors of EGFR and the time profile of their involvement. We classified the 81 effectors into major categories, recognizing that signaling molecules may have several functional roles (see **Supplementary Table 1** online). Only about a third of the proteins appear to have direct roles in transducing the signal to the cell nucleus. At least as many participate in cytoskeletal remodeling and cell adhesion. Of the remaining third, half are involved in receptor downregulation and signal termination. We found ten proteins that directly regulate the activities of the small GTPases, a finding consistent with their key role in signal transduction (**Supplementary Table 1** online).

Our data provide insight into early EGFR signaling. Many proteins not previously known to be linked with EGFR signaling were identified. The data also contain information about a rich variety of growth factor-related phenomena. For example, we observed the transactivation of several transmembrane receptors, including tyrosine kinase receptors (Axl) as well as G-protein coupled receptors (Anthrax toxin receptor; **Table 1**). As another example, STAT5 is known to be activated by EGFR but it is unclear if this activation is direct or indirect³. Because STAT5 is immediately activated, with the same kinetics as the receptor, our data suggest that the activation is direct. The tumor suppressor CYLD has recently been shown to be a critical regulator of NF kappa B-dependent apoptosis, and our data now link this protein to EGFR signaling²⁹⁻³¹. Likewise, the RNA binding proteins found in this study may help open up another area of growth factor signaling.

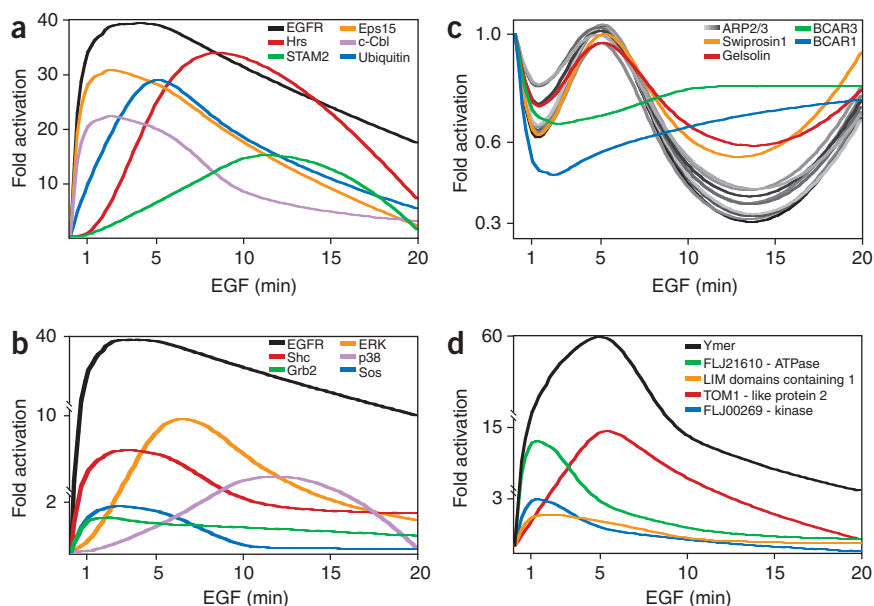


Figure 3 Activation profiles of different categories of EGFR signaling effectors, derived from mass spectrometric data as shown in **Figure 1b** and documented in **Table 1**. Continuous activation curves were obtained from the data in **Table 1** as described in Methods. **(a)** Key proteins involved in receptor internalization and endosomal trafficking. **(b)** Proteins from the Ras-MAP kinase pathways. The y-axis is discontinuous in **b** and **d** for better visualization. **(c)** Proteins involved in actin remodeling. **(d)** Novel proteins found in this study.

EGFR has been a case study for the modeling of signaling for several decades³². One difficulty in modeling has been the complexity of the system, caused by its many interdependencies and absence of *in vivo* rate constants. Our global view of EGF stimulation should help modeling efforts in three ways: by providing a much more complete ‘cast of characters’ as well as the magnitude of their activation; by allowing temporal ordering of the participants in the signaling network; and by providing a strong constraint on possible models, which will have to reproduce the measured activation curves for each protein.

Our temporal analysis can be extended to many other cellular situations. Other systems that involve tyrosine phosphorylation cascades can be studied in the same way. Different affinity strategies, such as the use of recombinant domains as bait¹⁰, would allow other signaling questions to be addressed. Furthermore, when enriching phosphopeptides instead of phosphoproteins^{4,5}, the time course of any individual phosphorylation site can, in principle, be determined. The strategy of triple labeling is likely to find use beyond the study of temporal variations. For example, it would be interesting to examine response to pharmacological agents by a stepwise increase of the dose. Finally, this global approach to growth factor signaling will also help in understanding and modeling growth deregulation in cancer and metabolic disorders.

METHODS

Materials and reagents. L-arginine U-¹³C₆ (¹³C₆¹⁴N₄-Arg) and L-arginine ¹³C₆,¹⁵N₄ (¹³C₆¹⁵N₄-Arg) were from Cambridge Isotope Laboratories. The cell culture medium, Dulbecco’s Modified Eagle Medium (DMEM) deficient in all amino acids, was a custom medium preparation from Gibco-Invitrogen. All other L-amino acids were obtained from Sigma Chemicals. Dialyzed serum was obtained from Gibco-Invitrogen. Trypsin was from Promega and HeLa cells were from the American Type Culture Collection (ATCC). EGF was obtained from PeproTech.

Antibodies against the following proteins were used: Grb2, c-Cbl, p130Cas/BCAR1 and Annexin I from Transduction Laboratories; p38, PLC γ -1 and agarose-conjugated phosphotyrosine antibody (4G10) from Upstate Biotechnology; phospho-p38 MAP kinase, p44/42 MAP kinase, phospho-p44/42 MAP kinase and immobilized phosphotyrosine antibody (P-Tyr-100) from Cell Signaling Technologies; EGF Receptor, Ubiquitin, Eps15, Cbl-B and Arp3 were from Santa Cruz Biotechnology. Secondary horseradish peroxidase (HRP)-conjugated anti-goat antibody was from Sigma and anti-mouse and anti-rabbit antibodies were from Amersham Biosciences. KLH-conjugated peptides corresponding to amino acids AHPVAQQHT-NYHQPLL and FSKSESSHKGFFHYKH, synthesized by Boston Biomolecules, were used to generate rabbit polyclonal antibodies against human STAM2 and Ymer, respectively.

Cell culture media preparation and cell culture conditions.

The appropriate amount of amino acids was added to the amino acid-deficient medium as concentrated stock solutions in PBS, according to manufacturer’s specifications for standard DMEM with the exception of arginine. The arginine-free medium was then divided into three lots and 21 mg/l of L-arginine, L-arginine-U-¹³C₆¹⁴N₄, L-arginine-U-¹³C₆-¹⁵N₄ were added separately to make up the Arg-0, Arg-6, Arg-10 media, respectively. Each medium with the full complement of amino acids and the specific form of arginine was sterile-filtered through a 0.22 μ m filter (Milipore) and divided into two lots—for dialyzed serum-containing and serum-free media.

HeLa cells were grown in DMEM-labeling medium, prepared as described above, supplemented with 2 mM L-glutamine, and 10% dialyzed fetal bovine serum plus antibiotics, in a humidified atmosphere with 5% CO₂ in air. Cells were grown for at least six cell divisions in labeling medium and were placed in serum-free medium 16 h before EGF stimulation.

Immunoprecipitation and western blotting. HeLa cells labeled with either L-arginine, L-arginine-U-¹³C₆¹⁴N₄ or L-arginine-U-¹³C₆-¹⁵N₄ (5×10^7 cells per condition) were treated with 150 ng/ml of EGF for 0 min, 1 min and 10 min. A second similarly labeled set of HeLa cells were treated with EGF for 0 min, 5 min and 20 min, respectively. All treatments were carried out at 37 °C. The cells were then lysed in a buffer containing 1% NP-40, 0.1% sodium deoxycholate, 150 mM NaCl, 1 mM EDTA, 50 mM Tris, pH 7.5, 1 mM sodium orthovanadate and protease inhibitors (Complete tablets, Roche Diagnostics). The lysates were centrifuged at 14,000g to pellet cellular debris. A Bradford assay was performed to equalize protein concentrations for subsequent 1:1:1 (L-arginine, L-arginine-U-¹³C₆¹⁴N₄, L-arginine-U-¹³C₆-¹⁵N₄) mixing of the lysates before the immunoprecipitation step. The unstimulated Arg-0 cells served as a common reference point in the two experimental sets, and lysates from these cells were pooled and divided into two equal lots for mixing with treated samples.

Mixed lysates were pre-cleaned on protein A beads for 1 h to reduce nonspecific binding of abundant proteins. Protein A beads were separated from lysates by passing the mixture through a Econo-Pac column (Bio-Rad). Agarose-conjugated anti-phosphotyrosine antibody 4G10 was incubated with each of the two sets of pre-cleaned lysates for 2 h followed by addition of agarose-conjugated anti-phosphotyrosine P-Tyr-100 for an additional 4 h. Precipitated complexes were then washed with 30 column volumes of lysis buffer and proteins were subsequently eluted with five column volumes of glycine pH 2.4. The eluate was neutralized with a neutralizing buffer containing 2 M Tris, pH 8.0, 1.5 M NaCl, 1 mM EDTA and concentrated on Centricon Spin columns (10 kDa MW cutoff; Milipore) for separation on a NuPAGE 4%–12% Bis-Tris gel (Invitrogen).

The gel was then stained with the Colloidal Blue staining Kit (Invitrogen) to visualize the gel lane. The two gel lanes (one for 0, 1 and 10 min and one for 0, 5 and 20 min of stimulation) were divided into ten slices each for gel-enhanced liquid chromatography–mass spectrometry (GelLC–MS) and excised from the gel and subjected to in-gel reduction and alkylation, followed by trypsin digestion as previously described³³. Peptides were extracted twice by adding an equal volume of 30% acetonitrile/0.3% trifluoroacetic acid (TFA) in water to the digest mixture followed by a final extraction with 100% acetonitrile. Extracts were dried in a Speedvac to remove organic solvent and subsequently acidified to 0.5% TFA. Samples were desalted and concentrated with STAGE tips³⁴ and resuspended in 5 μ l of 0.5% acetic acid/0.02% heptafluorobutyric acid for sample injection.

For the western blotting experiments, 1×10^7 HeLa cells per condition were used. The cells were grown, serum-starved and EGF-treated essentially the same way as described above for the large scale experiment. Briefly, the cells were starved for 16 h and then left untreated or treated with the growth factor for 1, 5, 10 and 20 min, respectively. Subsequently, the cells were scraped in a lysis buffer and the cleared lysates incubated with agarose-conjugated 4G10 antibody for 6 h. After the initial 2 h of incubation, anti-phosphotyrosine P-Tyr-100 was added to the immunoprecipitations. After three washes with lysis buffer, immunoprecipitated proteins were resolved on SDS-PAGE. After transfer, the nitrocellulose membrane was blocked with 2% BSA and then incubated with the corresponding antibodies, followed by HRP-conjugated secondary antibodies (Fig. 2). The membranes were subjected to chemiluminescent detection according to manufacturer's instructions (ECL, Amersham). In the case of p44/42 and p38 MAP kinases, the western blotting was performed directly on the lysates using phospho-specific antibodies.

Mass spectrometry and data analysis. Mass spectrometric analyses were done with nanoscale LC-mass spectrometry (LC-MS) and LC-tandem mass spectrometry and a quadrupole time-of-flight instrument (QSTAR-Pulsar, ABIMDS-SCIEX) with sample introduction with a 96-well autosampler (Agilent HP1100). The column was a self-packed, 10-cm long 8- μ m tip opening/100 μ m ID capillary needle (PicoTip, New Objective) packed with Vydac 218MSB3 bulk material (3 μ m prototype reversed phase material, a gift from Grace-Vydac). The chromatographic separation was done with a linear gradient elution from 95% buffer A (H₂O/acetic acid, 100:0.5 vol/vol) to 50% buffer B (H₂O/acetonitrile/acetic acid, 20:80:0.5 vol/vol) in 80 min.

Data files of completed LC-MS runs were converted to Mascot generic format flat files by the Mascot.dll script supplied with Analyst QS software (MDS-Sciex). The Mascot software package (Matrix Science London) was used for database search and protein identification with IPI human protein database. Both peptide mass tolerance and fragment mass tolerances were set at 0.2 Da for an initial search. An iterative calibration algorithm based on Mascot protein identifications was applied to the raw Analyst data file and gave mass accuracies within 20 p.p.m. All protein identifications were based on at least two matching peptides and the tandem mass spectra were additionally validated manually.

Arginine-containing peptides were extracted from the search results. Their retention times were used to locate the peptides in the raw mass spectral files CL with the help of an in-house software program capable of calling Analyst functions. (The program is available as 'open-source' at <http://www.sourceforge.net/> under the name MSQuant.) The program then identified the eluting peptide peak and determined the peak areas of peptide triplets for each single scan mass spectrum. This data were collated and verified manually for each arginine peptide cluster for each protein.

Each arginine peptide yielded two ratios expressed as the peak area of Arg-6 divided by the peak area of Arg-0, and Arg-10 divided by Arg-0. Averaging these peptide ratios for the different arginine peptides sequenced yielded the ratios shown in Table 1, for two of the four ratios. The other two ratios were obtained in the same way from the second set of experiments, using a common zero point and two additional durations of stimulation as described above.

The typical quantification error arising from ¹³C-substituted SILAC quantification methods were better than 10% relative standard deviation (RSD) for an individual peptide¹². We separately assessed the quantification variability

between different peptides of a protein, which also resulted in a variability well within 20% RSD. However, we observe that activation ratios >25 show higher variability (as high as 30% RSD) because of small peak areas of the Arg-0 peak¹². In these cases, the ratio of signal-to-noise is much greater for the Arg-0 peak as a percentage of the total peak area and the observed ratio is generally an underestimation of the true ratio. Because all points are normalized to the common zero time point, this error would contribute to all peptides by the same factor while leaving the curve shape unaffected. This is graphically illustrated in Supplementary Figure 1 online.

The typical quantification errors from mass spectrometric quantification of 20% RSD were smaller than the selected, conservative threshold of 1.5-fold chosen for biological significance. This cutoff yielded virtually the entire known EGF tyrosine phosphoproteome. The entire experiment was performed in an independent, three-time-point experiment on a smaller scale, which yielded reproducible activation profiles for all the proteins quantified.

Activation curves were generated from protein ratios in Table 1 by plotting interpolated curves with KaleidaGraph (Synergy Software) using the Stineman function. The peak in activation profiles as described in the text is the maximum of this plotted curve.

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS

We thank other members of our laboratory for help and fruitful discussions. Peter Mortensen is acknowledged for help with the bioinformatic analysis and Carmen de Hoog and Leonard Foster for critical reading of the manuscript. Work in the Center for Experimental BioInformatics (CEBI) is supported by a generous grant by the Danish National Research foundation and by the 6th Framework Program of the European Union (Interaction Proteome).

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 21 January; accepted 8 June 2004

Published online at <http://www.nature.com/naturebiotechnology/>

- Pawson, T. & Nash, P. Assembly of cell regulatory systems through protein interaction domains. *Science* **300**, 445–452 (2003).
- Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* **103**, 211–225 (2000).
- Blume-Jensen, P. & Hunter, T. Oncogenic kinase signalling. *Nature* **411**, 355–365 (2001).
- Ficarro, S.B. *et al.* Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat. Biotechnol.* **20**, 301–305 (2002).
- Salomon, A.R. *et al.* Profiling of tyrosine phosphorylation pathways in human cells using mass spectrometry. *Proc. Natl. Acad. Sci. USA* **100**, 443–448 (2003).
- Gygi, S.P. *et al.* Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17**, 994–999 (1999).
- Lill, J. Proteomic tools for quantitation by mass spectrometry. *Mass Spectrom. Rev.* **22**, 182–194 (2003).
- Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422**, 198–207 (2003).
- Ranish, J.A. *et al.* The study of macromolecular complexes by quantitative proteomics. *Nat. Genet.* **33**, 349–355 (2003).
- Blagoev, B. *et al.* A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat. Biotechnol.* **21**, 315–318 (2003).
- Ong, S.E. *et al.* 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 (2002).
- Ong, S.E., Kratchmarova, I. & Mann, M. Properties of ¹³C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). *J. Proteome Res.* **2**, 173–181 (2003).
- Pandey, A. *et al.* Identification of a novel immunoreceptor tyrosine-based activation motif-containing molecule, STAM2, by mass spectrometry and its involvement in growth factor and cytokine receptor signaling pathways. *J. Biol. Chem.* **275**, 38633–38639 (2000).
- Pandey, A. *et al.* 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. USA* **97**, 179–184 (2000).
- Pandey, A. *et al.* Cloning of a novel phosphotyrosine binding domain containing molecule, Odin, involved in signaling by receptor tyrosine kinases. *Oncogene* **21**, 8029–8036 (2002).
- Waterman, H. & Yarden, Y. Molecular mechanisms underlying endocytosis and sorting of ErbB receptor tyrosine kinases. *FEBS Lett.* **490**, 142–152 (2001).
- Dikic, I. & Giordano, S. Negative receptor signalling. *Curr. Opin. Cell Biol.* **15**, 128–135 (2003).

18. Bache, K.G., Raiborg, C., Mehlum, A. & Stenmark, H. STAM and Hrs are subunits of a multivalent ubiquitin-binding complex on early endosomes. *J. Biol. Chem.* **278**, 12513–12521 (2003).
19. Dong, C., Waters, S.B., Holt, K.H. & Pessin, J.E. SOS phosphorylation and disassociation of the Grb2-SOS complex by the ERK and JNK signaling pathways. *J. Biol. Chem.* **271**, 6328–6332 (1996).
20. De Corte, V. *et al.* Gelsolin-induced epithelial cell invasion is dependent on Ras-Rac signaling. *EMBO J.* **21**, 6781–6790 (2002).
21. Riggins, R.B., Quilliam, L.A. & Bouton, A.H. Synergistic promotion of c-Src activation and cell migration by Cas and AND-34/BCAR3. *J. Biol. Chem.* **278**, 28264–28273 (2003).
22. Brinkman, A., van der Flier, S., Kok, E.M. & Dorsers, L.C. BCAR1 a human homologue of the adapter protein p130Cas, and antiestrogen resistance in breast cancer cells. *J. Natl. Cancer Inst.* **92**, 112–120 (2000).
23. van Agthoven, T. *et al.* Identification of BCAR3 by a random search for genes involved in antiestrogen resistance of human breast cancer cells. *EMBO J.* **17**, 2799–2808 (1998).
24. Ostareck-Lederer, A. *et al.* c-Src-mediated phosphorylation of hnRNP K drives translational activation of specifically silenced mRNAs. *Mol. Cell. Biol.* **22**, 4535–4543 (2002).
25. Bach, I. The LIM domain: regulation by association. *Mech. Dev.* **91**, 5–17 (2000).
26. Wu, R. *et al.* Specificity of LIM domain interactions with receptor tyrosine kinases. *J. Biol. Chem.* **271**, 15934–15941 (1996).
27. Seykora, J.T., Mei, L., Dotto, G.P. & Stein, P.L. 'Srcasm: a novel Src activating and signaling molecule. *J. Biol. Chem.* **277**, 2812–2822 (2002).
28. Lohi, O., Poussu, A., Mao, Y., Quijcho, F. & Lehto, V.P. VHS domain—a longshoreman of vesicle lines. *FEBS Lett.* **513**, 19–23 (2002).
29. Trompouki, E. *et al.* CYLD is a deubiquitinating enzyme that negatively regulates NF-kappaB activation by TNFR family members. *Nature* **424**, 793–796 (2003).
30. Kovalenko, A. *et al.* The tumour suppressor CYLD negatively regulates NF-kappaB signalling by deubiquitination. *Nature* **424**, 801–805 (2003).
31. Brummelkamp, T.R., Nijman, S.M., Dirac, A.M. & Bernards, R. Loss of the cylindromatosis tumour suppressor inhibits apoptosis by activating NF-kappaB. *Nature* **424**, 797–801 (2003).
32. Wiley, H.S., Shvartsman, S.Y. & Lauffenburger, D.A. Computational modeling of the EGF-receptor system: a paradigm for systems biology. *Trends Cell Biol.* **13**, 43–50 (2003).
33. Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. Mass spectrometric sequencing of proteins from silver stained polyacrylamide gels. *Anal. Chem.* **68**, 850–858 (1996).
34. Rappsilber, J., Ishihama, Y. & Mann, M. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nano-electrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **75**, 663–670 (2003).