Ligand binding to cell surface receptors initiates a cascade of signaling events regulated by dynamic phosphorylation events on a multitude of pathway proteins. Quantitative features, including intensity, timing, and duration of phosphorylation of particular residues, may play a role in determining cellular response, but experimental data required for analysis of these features have not previously been available. To understand the dynamic operation of signaling cascades, we have developed a method enabling the simultaneous quantification of tyrosine phosphorylation of specific residues on dozens of key proteins in a time-resolved manner, downstream of epidermal growth factor receptor (EGFR) activation. Tryptic peptides from four different EGFR stimulation time points were labeled with four isoforms of the iTRAQ reagent to enable downstream quantification. After mixing of the labeled samples, tyrosine-phosphorylated peptides were immunoprecipitated with an anti-phosphotyrosine antibody and further enriched by IMAC before LC/MS/MS analysis. Database searching and manual confirmation of peptide phosphorylation site assignments led to the identification of 78 tyrosine phosphorylation sites on 58 proteins from a single analysis. Replicate analyses of a separate biological sample provided both validation of this first data set and identification of 26 additional tyrosine phosphorylation sites and 18 additional proteins. iTRAQ fragment ion ratios provided time course phosphorylation profiles for each site. The data set of quantitative temporal phosphorylation profiles was further characterized by self-organizing maps, which resulted in identification of several cohorts of tyrosine residues exhibiting self-similar temporal phosphorylation profiles, operationally defining dynamic modules in the EGFR signaling network consistent with particular cellular processes. The presence of novel proteins and associated tyrosine phosphorylation sites within these modules indicates additional components of this network and potentially localizes the topological action of these proteins. Additional analysis and modeling of the data generated in this study are likely to yield more sophisticated models of receptor tyrosine kinase-initiated signal transduction, trafficking, and regulation. Molecular & Cellular Proteomics 4:1240–1250, 2005.

Cell signaling downstream of receptor tyrosine kinases, such as epidermal growth factor receptor (EGFR),\(^1\) comprises an interconnected network of pathways associated with various regulatory processes. These pathways involve a number of components whose interactions are required to carry out regulatory processes and that are themselves regulated by covalent modifications, including phosphorylation on tyrosine as well as serine and threonine residues. In the particular case of EGFR, ligand binding activates the receptor and dimerization results in autophosphorylation of selected tyrosine sites in the C-terminal region (1). Different tyrosine phosphorylation sites in EGFR (i.e. Tyr-1092 and Tyr-1110 for Grb2 (2), Tyr-1172 and Tyr-1197 for SHC (3)) act as docking sites for a variety of proteins upstream of several signaling cascades, most prominently mitogen-activated protein kinase, phospholipase C-\(\gamma\), and phosphoinositide 3-kinase pathways, leading to proliferation, differentiation, migration, and antiapoptotic effects (4–6). Because EGFR is at the origin of pathways governing diverse cell behavioral responses such as cell survival, proliferation, differentiation, and motility, ascertaining quantitative and dynamic features of the various regulatory pathways will be imperative for determining which features are related to particular responses. Indeed, because dysregulation of EGFR-activated pathways, often a consequence of receptor overexpression or mutation, has been shown to be correlated with many types of cancer, one promising step toward identifying mechanisms underlying tumorigenesis associated with aberrant EGFR signaling would be to generate a quantitative comparison of a broad variety of specific cellular signaling events downstream of this RTK under multiple bio-

\(^{1}\) The abbreviations used are: EGFR, epidermal growth factor receptor; SOM, self-organizing map; IP, immunoprecipitation.
logical cell states. These data could then be used to implement models of cellular signaling pathways, from which predictions could be made as to the most beneficial intervention strategies. Toward this end, we describe here a quantitative mass spectrometric method for time-resolved analysis of dynamic tyrosine phosphorylation at specific sites on multiple proteins simultaneously, using the EGFR signaling cascade as a model.

All phosphorylation-mediated cellular signaling cascades are bound by the same principles. A dynamic relationship between component proteins in the pathway generates site- and time-specific phosphorylation/dephosphorylation events that propagate down the cascade until the desired response is elicited. Most proteins have more than one potential phosphorylation site, and phosphorylation/dephosphorylation of different sites in the same protein may lead to different responses; i.e. activation events may involve a particular protein, but specific phosphorylation sites regulate the cellular activity (7, 8). So far, most of the literature deals with either phosphorylation site, and phosphorylation/dephosphorylation of sites in receptor tyrosine kinase pathways (15–18) and to mass spectrometry been used to identify many tyrosine phosphorylation predictions could be made as to the most beneficial intervention strategies. Toward this end, we describe here a quantitative mass spectrometric method for time-resolved analysis of dy-

stimulation, cells were lysed on ice with 3 ml of 8 M urea supple-
mented with 1 M Na2O4. A 10-μl aliquot was taken from each sample to perform binchoninic acid protein concentration assay (Pierce) according to the manufacturer’s protocol. Cell lysates were reduced with 10 mM DTT for 1 h at 56 °C, alkylated with 55 mM iodoacetamide for 45 min at room temperature, and diluted to 12 ml with 100 mM ammonium acetate, pH 8.9. 40 μg of trypsin (Promega) was added to each sample (~ 100:1 substrate/trypsin ratio), and the lysates were digested overnight at room temperature. The whole-cell digest solutions were acidified to pH 3 with acetic acid (HOAc) and loaded onto C18 Sep-Pak Plus Cartridges (Waters). The peptides were desalted (10 ml of 0.1% HOAc) and eluted with 10 ml of a solution composed of 25% acetonitrile (MeCN) with 0.1% HOAc. Each sample was divided into 10 aliquots and lyophilized overnight to dryness for storage at ~80 °C.

**iTRAQ Labeling and Peptide IP—** Peptide labeling with iTRAQ reagent (Applied Biosystems) was performed according to the manufacturer’s protocol. In brief, each aliquot (3 × 10⁶ cell equivalent) was reacted with one tube of iTRAQ reagent; i.e. acquisition of this data set from the first biological sample required 8 tubes of iTRAQ reagent (2 × iTRAQ-114 (0 min), 2 × iTRAQ-115 (5 min), 2 × iTRAQ-116 (10 min), and 2 × iTRAQ-117 (30 min)) after the sample was dissolved in 30 μl of 0.5 M triethylammonium bicarbonate (NEt3HCO3), pH 8.5, and the iTRAQ reagent was dissolved in 70 μl of ethanol. The mixture was incubated at room temperature for 1 h and concentrated to ~20 μl. Samples labeled with four different isotopic iTRAQ reagents were combined and concentrated to 10 μl and then dissolved in 200 μl of IP buffer (100 mM Tris, 100 mM NaCl, and 1% Nonidet P-40, pH 7.4) and 200 μl of water, and pH was adjusted to 7.4. The mixed sample was incubated with 4 μg of immobilized anti-phosphotyrosine antibody (Santa Cruz Biotechnology) overnight at 4 °C. The antibody beads were spun down for 5 min at 7000 rpm, and the supernatant was separated and saved. The antibody-bound beads were washed twice with 200 μl of IP buffer for 10 min and twice with rinse buffer (100 mM Tris, 100 mM NaCl, pH 7.4) for 5 min at room temperature. The phosphotyrosine-containing peptides were eluted from antibody with 50 μl of 100 mM glycine pH 2.5 for 1 h at room temperature.

**IMAC and Mass Spectrometry—** Phosphopeptide enrichment on IMAC was performed as described previously (14), except that peptides were not converted to methyl esters. After antibody elution, peptides were loaded onto a 10-cm self-packed IMAC (20M C, Applied Biosystems) capillary column (inner diameter, 200 μm; outer diameter, 360 μm) and rinsed with organic rinse solution (25% MeCN, 1% HOAc, 100 mM NaCl) for 10 min at 10 μl/min. The column was equilibrated with 0.1% HOAc for 10 min at 10 μl/min and eluted onto a 10-cm self-packed C18 capillary precolumn (inner diameter, 100 μm; outer diameter, 360 μm) with 50 μl of 250 mM NaHPO4, pH 8.0. After a 10-min rinse (0.1% HOAc), the precolumn was connected to a 10-cm self-packed C18 precolumn (5 μm ODS-AQ) analytical capillary column (inner diameter, 50 μm; outer diameter, 360 μm) with an integrated electrospray tip (~1-μm orifice). Peptides were eluted using a 100-min gradient with solvent A (H2O/HOAc, 99:1 (v/v)) and B (H2O/MeCN/HOAc, 29:70:1 (v/v)) at 0% from 0% to 15% B, 75 min from 15% to 40% B, and 15 min from 40% to 70% B. Eluted peptides were directly electrosprayed into a Q-TOF mass spectrometer (QSTAR XL Pro; Applied Biosystems). MS/MS spectra of the five most intense peaks with two to five charges in the MS scan were automatically acquired in information-dependent acquisition with previously selected peaks excluded for 40 s.

**Western Blot Analysis—** 40 μg of protein from each sample was mixed with 4× sample buffer (250 mM Tris-HCl, pH 6.8, 8% SDS, 40% glycerol, 0.04% bromphenol blue, and 400 mM dithiothreitol) and boiled for 5 min. Proteins were separated by SDS-PAGE and transferred onto PVDF membranes. After blocking for 1 h at room temper-
ature membranes were incubated overnight at 4 °C in primary antibody, washed three times for 5 min in TBS-Tween 20 (20 mM Tris-HCl, pH 7.5, 137 mM NaCl, and 0.1% Tween 20), incubated for 1 h at room temperature in secondary antibody (dilution 1:2500 horseradish peroxidase conjugated donkey anti-rabbit in TBS-Tween 20, 5% non-fat milk powder) (Amersham Biosciences), and finally washed three times for 5 min with TBS-Tween 20. Blots were developed with ECL Advance Western blotting detection kit (Amersham Biosciences) and scanned on a Kodak Image Station 1000. Primary antibodies used were anti-EGFR, for loading control; anti-EGFR pY1172; anti-Fak pY576; anti-STAT-3 pY705 (all from Cell Signaling Technologies), and anti-Gsk-3-β pY216 (Upstate Biotechnology).

Phosphopeptide Sequencing, Data Clustering, and Analysis—MS/MS spectra were extracted and searched against human protein database (NCBI) using ProQuant (Applied Biosystems) and MASCOT (Matrix Science). For ProQuant, an interrogator database was generated by predigesting the human protein database with trypsin and allowing one miscleavage and up to six modifications on a single peptide (phosphotyrosine ≤ 2, phosphoserine ≤ 1, phosphothreonine ≤ 1, iTRAQ-lysine ≤ 4, and iTRAQ-tyrosine ≤ 4). Mass tolerance was set to 0.15 atomic mass units for precursor ions and 0.1 atomic mass units for fragment ions. For MASCOT, data were searched against the human non-redundant protein database with trypsin specificity, two missed cleavages, precursor mass tolerance set to 1.5 Da, and fragment ion tolerance set to 0.2 Da. Phosphotyrosine-containing peptides were manually validated and quantified. Peak areas for each of the four signature peaks (m/z: 114, 115, 116, 117) were obtained and corrected according to the manufacturer’s instructions to account for isotopic overlap. Data were further corrected with values generated from the peak areas of non-phosphorylated peptides to account for possible variations in the starting amounts of sample for each time point. Finally, all data were normalized by the anti-phosphotyrosine peptide immunoprecipitation was analyzed, and iTRAQ marker ions from non-phosphorylated peptides were averaged and used to correct the data (Fig. 1c; data available in Supplemental Table III).

Application of this method to investigate EGFR signaling in human mammary epithelial cells resulted in the identification of 78 tyrosine phosphorylation sites on 58 proteins in a single analysis. For each phosphorylation site, a quantitative temporal phosphorylation profile was generated by comparing the relative ratios of peak areas for the iTRAQ marker ions (m/z 114–117) in the MS/MS spectrum. The mean values and standard deviations for a subset of the tyrosine phosphorylation sites are provided in Table I along with an estimation of the statistical significance of each value relative to the 5-min time point (a full list of phosphorylation sites, time course phosphorylation profiles, and statistical analysis can be found in Supplemental Table I). For instance, the EGFR pY1172 autophosphorylation site increased 20-fold from 0 to 5 min, decreased by 10% from 5 to 10 min, and decreased by another 40% to 30 min. Based on the number of replicate measurements (as is typical for data-dependent acquisition of the peptides were selected for MS/MS more than once during the chromatographic elution profile, resulting in multiple independent measurements of the relative ratios for the iTRAQ marker ion peak areas) in the LC/MS/MS analysis, the difference between the 0- and 5-min time points and the 30- and 5-min time points is statistically significant (p < 0.05), whereas the difference between the 10- and 5-min time points does not meet this significance criterion for this particular phosphorylation site.

Of the 58 proteins identified in this analysis, 52 have been associated with the EGFR signaling network, whereas the
other six proteins (listed in Table II) have not been previously identified in either proteomic or biochemical analyses of EGFR signaling. Representative spectra for two of the six peptides listed in Table II are presented in Fig. 2; the mass range spanning the iTRAQ marker ion region is highlighted for each peptide. Included in this group of six proteins is KIAA0606, a protein also known as suprachiasmatic nucleus circadian oscillatory protein (SCOP). Regulation of the circadian clock in the suprachiasmatic nucleus drives daily rhythms of behavior and has been associated with EGF receptor signaling in the hypothalamus (25). In addition, EGF activated expression of clock genes Per1 and Per2, and modulation of the circadian rhythm in cell culture has also been documented (26). The mechanism by which signal is transduced from the EGF receptor to the clock gene promoter region has not been elucidated. In this study, we detected rapid and transient tyrosine phosphorylation at pY2060 in SCOP in response to EGF stimulation. This protein and phosphorylation site may be directly involved in signal transduction from activated EGF receptor to downstream up-regulation of clock gene expression, but this relationship needs to be biologically validated. Also included in this group of six proteins is KIAA1217, a protein whose function has not yet been characterized. The specific peptide and phosphorylation site that we have identified on this protein display surprising homology (12/15 identical residues) to a region of p130Cas interacting protein (p140Cas), a protein known to be tyrosine-phosphorylated in response to stimulation with exogenous growth factors (27).

In addition to the six proteins that had not been previously characterized in the EGFR signaling network, we have also identified several novel phosphorylation sites on proteins known to be in the network (a subset of these sites is presented in Table III). Contained in this group are phosphorylation sites on hypothetical protein FLJ00269, hypothetical protein FLJ21610, target of myb1-like 2 protein, and chromosome 3 open reading frame 6 (previously named Ymer by Blagoev et al. (19)). Identification and quantification of temporal phosphorylation profiles at specific tyrosine sites in these proteins further confirms the involvement of these proteins in the EGFR signaling cascade and should facilitate biological manipulation of these proteins to help characterize
their functional role in the EGFR network.

To validate the method, we chose to compare our results with those generated by Western blot with phospho-specific antibodies. Four of the phosphorylation sites identified by mass spectrometry were selected; total cell lysate was loaded on the gel and blotted, and selected anti-phosphotyrosine antibodies were used to develop the Western blots. As demonstrated in Fig. 3, anti-phosphotyrosine Western blots recapitulated the general trend measured by mass spectrometry.

As further validation of this data set, a second set of cells were cultured, stimulated with EGF for the appropriate times, lysed, processed, and analyzed to quantify temporal tyrosine phosphorylation profiles. Two replicate analyses were performed on this set of samples to estimate analytical reproducibility of the method, and the iTRAQ labeling scheme was reversed within these replicates to ensure that the quantification was not biased by a particular isofrom of the iTRAQ reagent. A subset of the results from these replicate analyses are presented in Table I, and the full data set is contained in Supplemental Table II. Of the 78 tyrosine phosphorylation sites identified from analysis of the first set of cells, 65 were detected and quantified in the analyses of the second set of stimulated cells. The absence of the remaining 13 phospho-

### Table I

| Protein Site Ratio | Sample 1, average of analysis 1 and 2 | Sample 2, average of analysis 1 and 2 |
|--------------------|--------------------------------------|--------------------------------------|
| **Ratio**          | 0 min:5 min                          | 10 min:5 min                         |
| c-Cbl              | 0.000 ± 0.001***                     | 0.160 ± 0.015                        |
| Caveolin 1          | 0.391 ± 0.074                        | 1.043 ± 0.009                        |
| EGFR               | 0.505 ± 0.031*                       | 0.505 ± 0.012*                       |
| EGFR               | 0.059                                | 0.429                                |
| EphB1              | 0.656 ± 0.090                        | 1.164 ± 0.049                        |
| EphrinB2           | 0.837                                | 1.156                                |
| EPS15              | 0.080                                | 0.652                                |
| FAK                | 0.637 ± 0.051**                      | 1.629 ± 0.036**                      |
| FRK                | 0.614                                | 1.340                                |
| GIT 1              | 0.575 ± 0.057                        | 1.090 ± 0.040                        |
| GSK3-β             | 0.846 ± 0.014**                      | 0.908 ± 0.015                        |
| HER2               | 0.141                                | 0.532                                |
| Paxillin           | 0.528 ± 0.131                        | 0.792 ± 0.164                        |
| Plakophilin 3      | 0.194                                | 0.778                                |
| PLC-γ              | 0.043                                | 0.335                                |
| PTP NR 11          | 0.945 ± 0.032                        | 1.122 ± 0.023                        |
| PTPR A             | 0.893 ± 0.005                        | 0.724 ± 0.030                        |
| Rho GEF 5          | 0.201 ± 0.037**                      | 0.632 ± 0.047                        |
| SHC                | 0.147 ± 0.015**                      | 0.902 ± 0.059                        |
| SHC                | 0.028 ± 0.046*                       | 0.296 ± 0.392                        |
| SHIP-2             | 0.043 ± 0.034**                      | 0.363 ± 0.028**                      |
| STAM-1             | 0.309                                | 0.926                                |
| STAM-2             | 0.218                                | 0.908                                |
| STAT-3 Isoform 1   | 0.209 ± 0.026**                      | 0.244 ± 0.046**                      |
| STAT-3 Isoform 2   | 0.195                                | 0.275                                |

*p < 0.05; **p < 0.01; ***p < 0.001.
FIG. 2. Representative MS/MS spectra for two phosphorylated peptides from proteins not previously reported in EGFR signaling. For each spectrum, y- and b-type fragment ions present in full scan mass spectra enable peptide identification and phosphorylation site assignment, whereas peak areas for each of the iTRAQ marker ions (inset for top and bottom) enable quantification of the temporal phosphorylation profiles.

TABLE II

Tyrosine-phosphorylated proteins not previously reported in EGFR signaling

Of the 58 proteins identified in the analysis of the first biological sample, six have not been previously associated with the EGFR signaling network. Protein identification, tyrosine phosphorylation site, and time course profile data are provided for these six proteins. *, p < 0.05; **, p < 0.01; ***, p < 0.001.

| Protein                                    | Site     | 0 min:5 min | 10 min:5 min | 30 min:5 min |
|--------------------------------------------|----------|-------------|--------------|--------------|
| Chromosome 20 ORF 18 (HOIL-1)              | pY288    | 0.026 ± 0.034*** | 1.950 ± 0.156** | 0.441 ± 0.012*** |
| Hypothetical protein FLJ14564              | pY413    | 0.66 ± 0.04  | 0.956 ± 0.063  | 0.719 ± 0.023  |
| Hypothetical protein FLJ30532 (MARVELD2)  | pY23     | 0.000       | 0.753         | 0.534         |
| Retinoic acid induced 3                    | pY320    | 0.725 ± 0.117 | 1.177 ± 0.027  | 0.891 ± 0.029  |
| KIAA0608 protein                          | pY2060   | 0.140       | 0.671         | 0.413         |
| KIAA1217 protein                          | pY393    | 0.456       | 1.102         | 0.952         |

Fig. 2. Representative MS/MS spectra for two phosphorylated peptides from proteins not previously reported in EGFR signaling. For each spectrum, y- and b-type fragment ions present in full scan mass spectra enable peptide identification and phosphorylation site assignment, whereas peak areas for each of the iTRAQ marker ions (inset for top and bottom) enable quantification of the temporal phosphorylation profiles.
Time-resolved Analysis of EGFR Signaling

TABLE III
Tyrosine phosphorylation sites not previously reported in the literature

In addition to the sites listed in Table II, many additional tyrosine phosphorylation sites that have not been previously reported in the literature have been identified in this work. A subset of these sites is presented along with the protein identification and the temporal phosphorylation profile quantification.

| Protein | Site | Ratio 0 min:5 min | Ratio 10 min:5 min | Ratio 30 min:5 min |
|---------|------|-------------------|-------------------|-------------------|
| Hypothetical protein FLJ00269 | pY411 | 0.808 | 0.841 | 1.026 |
| Hypothetical protein FLJ21610 | pY453 | 0.261 | 1.123 | 0.888 |
| Chromosome 3 ORF 6 (Ymer) | pY145 | 0.00 ± 0.003*** | 1.359 ± 0.024* | 0.444 ± 0.005** |
| Target of myb1-like 2 | Tyr-160 & Tyr-168 | 0.059 ± 0.023* | 0.640 ± 0.152 | 0.249 ± 0.058* |

*, p < 0.05; **, p < 0.01; ***, p < 0.001.

Fig. 3. Comparison of tyrosine phosphorylation dynamics as measured by mass spectrometry and Western blot analysis. Tyrosine phosphorylation profiles determined by mass spectrometry were recapitulated by Western blot analysis, further validating the results generated by this approach. Gel loading levels were normalized by probing with an anti-EGFR antibody (bottom).

min) and the shape of the temporal profile, perhaps reflecting signaling pathways that are potentiated differently between the two sets of cell cultures. Many possible sources of variation between the cultures exist, including slight changes in the confluence of the cells (different cell-cell contacts would prime signaling pathways differently) and changes in the passage number of the cells (the second set of cells were cultured and stimulated 1 month after the first set of cells). Given the sources of variation, it is not surprising to see slight alteration of the temporal profiles for some of the phosphorylation sites. However, most of the temporal phosphorylation profiles are very similar between the cell cultures, which is indicative of the reproducibility of the method and helps to further validate the results of this study.
Bioinformatic Analysis Reveals Dynamic Modules within the EGFR Signaling Network—To ascribe potential functionality to proteins and phosphorylation sites not previously associated with the EGFR signaling network, we attempted to find similar features in temporal phosphorylation profiles between poorly and well characterized protein phosphorylation sites. Sorting through the data manually tends to introduce bias and proved to be time intensive and non-productive. Therefore, with the goal of clustering self-similar phosphorylation profiles (modules) within the EGFR signaling network, we resorted to bioinformatic analysis and generated a self-organizing map (SOM) with the data set from the first set of cells. SOM is a mathematical technique designed to identify underlying patterns in complex data sets; SOMs have been used to analyze gene expression patterns in hematopoietic differentiation, creating biologically relevant clusters and enabling the generation of interesting hypotheses (28). Using the quantitative tyrosine phosphorylation profiles we tested several different options for SOM matrix size before settling on a $3 \times 3$ matrix (Fig. 4a).

Fig. 4. SOMs facilitate the identification of EGFR signaling modules. a, after loading the mass spectrometry data into Spotfire, a $3 \times 3$ self-organizing map was generated to facilitate interpretation of the data and to uncover groups of phosphorylation sites that displayed similar patterns. b, a single cluster within the SOM (blue box) contained many phosphorylation sites involved in the immediate-early response. c, exception for the doubly phosphorylated peptide from ACK1, HOIL-1, and Ymer, phosphorylation sites contained in two adjacent clusters (red box) are from proteins implicated in regulation of receptor endocytosis and trafficking. Each of these sites was maximally phosphorylated at 10 min of EGF stimulation. The level of phosphorylation at 10 and 30 min relative to 5 min separated the phosphorylation sites into different compartments.

**Molecular & Cellular Proteomics 4.9**

1247
tyrosine phosphorylation sites on SHC, a protein whose PTB (phosphotyrosine-binding) domain binds to EGFR pY1172 (29). In fact, almost all of the phosphorylation sites in this cluster are located on proteins known to interact with EGFR or other receptor tyrosine kinases. For instance, phosphoinositide 3-kinase p85α has been shown to interact (both directly and indirectly) with both EGFR and PDGFR (30). Phosphoinositide 3-kinase p85α pY580 has been ascribed to insulin receptor tyrosine kinase activity (31) but is most probably the result of phosphorylation by EGFR under these stimulation conditions. Other proteins in this group include c-Cbl, Rho-GEF 5, ACK1, BDP1, Erk1, and hypothetical protein FLJ30532, one of the six proteins not previously associated with the EGFR network. c-Cbl is tyrosine-phosphorylated after EGF stimulation and has recently been shown to interact with pY1045 of EGFR after EGF stimulation and before endocytosis of the receptor (32). Tyrosine phosphorylation of activated CDC42 kinase 1 (ACK1) in response to EGF stimulation has been established (33). This protein is most likely to be localized to the receptor through an interaction with CDC42 and Rho-GEF 5; Rho-GEF proteins have been shown to interact directly with tyrosine kinase receptors (34). BDP1 is a phosphatase involved in regulation of Gab1, mitogen-activated protein kinase, and HER2 signaling after EGF stimulation (35); it is likely that phosphorylation of this protein stimulates phosphatase activity in a negative feedback mechanism.

Erk1 and hypothetical protein FLJ30532, the remaining two proteins in this cluster, are not known to associate directly with EGFR or other receptor tyrosine kinases. From their profiles and grouping in this cluster, the tyrosine phosphorylation sites on these proteins are involved in the immediate early response to EGF stimulation. Hypothetical protein FLJ30532 is also known as MARVELD2, named for MARVEL, a recently discovered tetra-spanning transmembrane domain (36). It is interesting that another MARVEL-domain containing protein has recently been implicated in mediating signal transduction from the B-cell receptor to Erk2 (37). Identification of the tyrosine phosphorylation site on MARVELD2 provides a handle by which to investigate the role of this protein in EGFR-mediated signal transduction, whereas localization to this cluster potentially localizes the topological action of this protein to the immediate-early response to EGF stimulation. It should be noted that although the Erk1 tyrosine site identified in this study is located in the activation loop, the doubly phosphorylated, fully active form of the protein was not detected; this peptide was probably present at a lower level and was below the detection limit of the current approach.

In addition to the immediate-early response module, SOM analysis of the data also uncovered an endocytosis and trafficking module in a metacluster of self-similar time-course phosphorylation profiles found in two adjacent compartments of the SOM (Fig. 4c). These phosphorylation sites are all maximally phosphorylated at 10 min after EGF stimulation and have separated into two SOM compartments based on phosphorylation level at 10 min relative to 5 min of stimulation. Proteins within this cluster known to be involved in endocytosis and trafficking include Eps15, STAM1, STAM2, and Annexin A2. EPS15 is necessary for receptor-mediated endocytosis of EGF, and tyrosine phosphorylation of Eps15 after EGF stimulation has been shown (38). Monoubiquitination of Eps15 occurs after EGF-induced internalization of the EGFR receptor (39) and may be related to c-Cbl activity, but a phosphorylation site from an alternate E3 ubiquitin-protein ligase, chromosome 20 open reading frame 18 (also known as HOIL-1 (40)) is localized to this cluster and may be responsible for further ubiquitination within the endosomes. Involvement of HOIL-1 in the EGFR pathway has not been characterized, but its biological function fits in well with endocytosis and trafficking. STAM1 and STAM2 phosphorylation sites are also located in these clusters, these proteins bind to ubiquitinated proteins via their Vps27p, Hrs, and STAM domains and ubiquitin-interacting motifs in early endosomes and are involved in endosomal trafficking (41). The N terminus of annexin A2 is tyrosine-phosphorylated after growth factor stimulation and may play a role in receptor trafficking; recently it has been shown that phosphorylation is blocked when receptor internalization is inhibited (42). Based on the clustering of these phosphorylation sites, it is possible that phosphorylation of these sites occurs after EGFR ubiquitination by Cbl and endocytosis via clathrin-coated early endosomes. If so, it may be that Ymer, a protein recently found to be modulated by EGF stimulation (19) and located in these clusters, may be localized to early endosomes. If Ymer follows the role of the other proteins within this cluster, tyrosine phosphorylation of this protein may regulate endosomal trafficking of the receptor. Identification of this regulated site will enable further investigation to determine the functional role of phosphorylation of this protein in EGFR signaling.

DISCUSSION

Quantitative, time-resolved analysis of phosphorylated-tyrosine mediated signaling cascades downstream of activated receptor tyrosine kinases will be critical in developing a better understanding of the molecular mechanisms underlying a variety of disease states. To this end, the method that we have developed has enabled the quantification of the temporal phosphorylation profile of specific tyrosine phosphorylation sites after EGF stimulation. This method is not limited to EGF signaling but is applicable to the investigation of a broad variety of other tyrosine-mediated signal transduction pathways. Stable-isotope labeling of peptides with the iTRAQ reagent occurs after cell lysis and tryptic digestion but before immunoprecipitation. Quantification is therefore not limited to cell culture-derived samples, and variations in immunoprecipitation do not affect quantification. In fact, because only several hundred micrograms of protein is required per sample per analysis, this method should be readily amenable to the
investigation of signal transduction in samples with limited quantities of protein, such as clinical tumor tissues.

It is worth noting that Blagoev et al. (19) have recently published an alternative method combining stable isotope labeling with amino acids in cell culture and immunoprecipitation of tyrosine-phosphorylated proteins, thereby enabling the quantification of temporal involvement in the EGFR signaling network at the protein level. Although the two methods seem quite similar, there are important advantages and disadvantages of each that should be noted. The method that we describe here provides identification of specific protein phosphorylation sites and quantitative temporal phosphorylation profiles for each of these sites; the method of Blagoev et al. (19) monitors tyrosine phosphorylation (or association with tyrosine-phosphorylated proteins) at the protein level without site specification. Our site-specific monitoring of protein phosphorylation provides more explicit detail regarding the regulation of proteins within the network but precludes analysis of potential other, non-phosphorylated proteins associated (and therefore co-immunoprecipitated) with tyrosine-phosphorylated proteins. In both methods, only those peptides (and therefore proteins) amenable to LC/MS/MS analysis will be identified; this limitation is more significant for peptide immunoprecipitation in that there are fewer peptides with which to identify each protein. Our current method uses only a single anti-phosphotyrosine antibody such that the results may be constrained by the bias associated with this particular antibody. Future work will investigate the use of multiple different antibodies to increase coverage of the network. Despite differences between the two methods, there is significant consistency at the protein level between the Blagoev et al. (19) data set and our data set, with ours then providing additional site-specific information. Differences in the temporal phosphorylation profiles between the two studies are probably caused by multiple tyrosine phosphorylation sites on a given protein but may also be caused by the different quantification methods (stable isotope labeling with amino acids in cell culture versus iTRAQ).

Another study using anti-phosphotyrosine protein immunoprecipitation, stable isotope (ICAT) labeling, and mass spectrometry to study EGFR signaling was recently published (43). Similar to Blagoev et al. (19), Thelemann et al. (43) monitored tyrosine phosphorylation (or association with tyrosine-phosphorylated proteins) at the protein level without site specification, although many phosphorylation sites were identified without stable-isotope quantification in a separate section of the manuscript. It is difficult to directly compare our data with those reported by Thelemann et al. (43) because the total number of proteins quantified in their report was significantly fewer than contained in either our data set or the data set reported by Blagoev et al. (19).

In this article, we have demonstrated an analytical method enabling the quantification of time-resolved tyrosine phosphorylation profiles with site-specific resolution and estimated the statistical significance of the relative phosphorylation levels at each time point after stimulation. Data generated by this method self-organize into clusters of phosphorylation sites that recapitulate and extend biological findings in the literature. Generation of an SOM has enabled the identification of modules in the EGFR signaling network. Based on the presence of particular proteins within these modules, we can generate additional hypotheses aimed at defining the potential function of tyrosine phosphorylation on several proteins not previously implicated in EGFR signaling. Further analysis and modeling of the data generated in this study are likely to result in the development of more sophisticated models of receptor-initiated signal transduction, endocytosis, trafficking, and regulation. Future application of this method to interrogate aberrant signaling downstream of constitutively active tyrosine kinases may reveal mechanisms of pathogenesis in these systems, and may provide additional targets for novel therapeutics.

Acknowledgments—We thank Rebecca Fry for assistance with the SpotFire program and other members of the White and Lauffenburger laboratories for helpful discussions.

* This work was supported by NCI, National Institutes of Health, Bioengineering Research Partnership Grant CA96504 (to D. A. L) and National Institutes of Health Grants GM68762 (to D. A. L.), DK070172 (to F. M. W.), and DK42816 (to F. M. W.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ The on-line version of this article (available at http://www.mcponline.org) contains Supplemental Tables I–III.

** To whom correspondence should be addressed: Biological Engineering Division, 56–787, MIT, Cambridge, MA 02139. Tel.: 617-258-8949; E-mail: fwhite@mit.edu.

REFERENCES

1. Holbro, T., and Hynes, N. E. (2004) ErbB receptors: directing key signaling networks throughout life. Annu. Rev. Pharmacol. Toxicol. 44, 195–217
2. Batzer, A. G., Rotin, D., Urena, J. M., Skolnik, E. Y., and Schlessinger, J. (1994) Hierarchy of binding sites for Grb2 and Shc on the epidermal growth factor receptor. Mol. Cell. Biol. 14, 5192–5201
3. Okabayashi, Y., Kido, Y., Okutani, T., Sugimoto, Y., Sakaguchi, K., and Kasuga, M. (1994) Tyrosines 1148 and 1173 of activated human epidermal growth factor receptors are binding sites of Shc in intact cells. J. Biol. Chem. 269, 18674–18678
4. Navolanic, P. M., Steelman, L. S., and McCubrey, J. A. (2003) EGFR family signaling and its association with breast cancer development and resistance to chemotherapy. Int. J. Oncol. 22, 237–252
5. Schoebert, B., Eichler-Jonsson, C., Gilles, E. D., and Muller, G. (2002) Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGFR receptors. Nat. Biotechnol. 20, 370–375
6. Wells, A., Ware, M. F., Allen, F. D., and Lauffenburger, D. A. (1999) Shaping up for shipping out: PLCgamma signaling of morphology changes in EGFR-stimulated fibroblast migration. Cell Motil. Cytochem. 44, 227–233
7. Moro, L., Dolce, L., Cabodi, S., Bergatto, E., Erba, E. B., Smeriglio, M., Turco, E., Retta, S. F., Giarfida, M. G., Venturino, M., Godovac-Zimmermann, J., Conti, A., Schaefer, E., Beguinot, L., Tacchetti, C., Gaggin, P., Silengo, L., Tarone, G., and Defilippi, P. (2002) Integrin-induced epidermal growth factor (EGF) receptor activation requires c-Src and p130Cas and leads to phosphorylation of specific EGF receptor tyrosines. J. Biol.
by hypothalamic EGFR receptor signalling. Novartis Found. Symp. 253, 250–262; discussion 102–259, 263–256, 281–254

Balsalobre, A., Maracci, L., and Schibler, U. (2000) Multiple signaling pathways elicit circadian gene expression in cultured Rat-1 fibroblasts. Curr. Biol. 10, 1291–1294

Di Stefano, P., Cabodi, S., Boeri Erba, E., Margaria, V., Bergatto, E., Giuffrida, M. G., Silengo, L., Tarone, G., Turco, E., and Defilippi, P. (2004) P130Cas-associated protein (p140Cap) as a new tyrosine-phosphorylated protein involved in cell spreading. Mol. Biol. Cell 15, 787–800

Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., Lande, R. S., and Golub, T. R. (1999) Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. Proc. Natl. Acad. Sci. U. S. A. 96, 2907–2912

Sakaguchi, K., Okabayashi, Y., Kido, Y., Kimura, S., Matsumura, Y., In-ussin, K., and Kasaq, M. (1998) Shc phosphorylation-binding domain dominantly interacts with epidermal growth factor receptors and mediate Ras activation in intact cells. Mol. Endocrinol. 12, 536–543

Hu, P., Margolis, B., Skolnik, E. Y., Lammers, R., Ulrich, A., and Schlessinger, J. (1992) Interaction of phosphatidylinositol 3-kinase-associated p85 with epidermal growth factor and platelet-derived growth factor receptors. Mol. Cell. Biol. 12, 981–991

Hayashi, N., Nishioka, Y., Kondaira, S., Kanai, I., Ishi, K., Fuku, Y., Shibasaki, F., Takenawa, T., Kido, H., and Katsunuma, N. (1993) The alpha-type 85-kDa subunit of phosphatidylinositol 3-kinase is phosphorylated at tyrosines 368, 580, and 607 by the insulin receptor. J. Biol. Chem. 268, 7107–7117

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

Satoh, T., Kato, J., Nishida, K., and Kaziro, Y. (1996) Tyrosine phosphorylation of ACK in response to temperature shift-down, hyperosmotic shock, and epidermal growth factor stimulation. FEBS Lett. 386, 230–234

Taya, S., Inagaki, N., Senguku, H., Makino, H., Iwamatsu, A., Urakawa, I., Nagao, K., Kataoka, S., and Kabuchi, K. (2001) Direct interaction of insulin-like growth factor/insulin receptor with leukemia-associated RhoGEF. J. Cell Biol. 155, 889–902

Gensler, M., Buschbeck, M., and Ulrich, A. (2004) Negative regulation of HER2 signaling by the PEST-type protein-tyrosine phosphatase BDP1. J. Biol. Chem. 279, 12110–12116

Sanchez-Pulido, L., Martin-Belmonte, F., Valencia, A., and Alonso, M. A. (2002) MARVEL: a conserved domain involved in membrane apposition events. Trends Biochem. Sci. 27, 599–601

Imamura, Y., Katahira, T., and Kitamura, D. (2004) Identification and characterization of a novel BASH N terminus-associated protein, BNAS2. J. Biol. Chem. 279, 26425–26432

Salcini, A. E., Chen, H., Iannolo, G., De Camilli, P., and Di Fiore, P. P. (1999) Epidermal growth factor pathway substrate 15, Epis15. Int. J. Biochem. Cell Biol. 31, 805–809

van Delft, S., Goverts, R., Strous, G. J., Verkleij, A. J., and van Bergen en Henegouwen, P. M. (1997) Epidermal growth factor induces ubiquitination of Eps15. J. Biol. Chem. 272, 14013–14016

Yamanaka, K., Ishikawa, H., Megumi, Y., Tokunaga, F., Kanie, M., Rouault, T. A., Morishima, I., Minato, N., Ishimori, K., and Iwai, K. (2003) Identification of the ubiquitin-protein ligase that recognizes oxidized IRP2. Nat. Cell Biol. 5, 336–340

Mizuno, E., Kawahata, K., Kato, M., Kitamura, N., and Komada, M. (2003) STAM proteins bind ubiquitinated proteins on the early endosome via the VHS domain and ubiquitin-interacting motif. Mol. Biol. Cell 14, 3675–3689

Biener, Y., Feinstein, R., Mayak, M., Kaburagi, Y., Kadowaki, T., and Zick, Y. (1996) Annexin II is a novel player in insulin signal transduction. Possible association between annexin II phosphorylation and insulin receptor internalization. J. Biol. Chem. 271, 29489–29496

Thelemann, A., Petti, F., Griffin, G., Iwata, K., Hunt, T., Settnari, T., Feny, 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