Single Molecule Pull-down for Quantification of Membrane Receptor Multisite Phosphorylation

Emanuel Salazar-Cavazos[1,2], Keith A Lidke[2,3], Diane S. Lidke[1,2]

[1]Department of Pathology, [2]Comprehensive Cancer Center and [3]Department of Physics & Astronomy, University of New Mexico, Albuquerque, NM 87131

The single molecule pull-down (SiMPull) assay allows for capture of individual proteins or macromolecular complexes that can then be interrogated via single molecule imaging. We describe several technical improvements over previous protocols that make it possible to directly detect and quantify the phosphorylation state of thousands of individual membrane receptors, and thereby estimate both the fraction of receptors phosphorylated at specific tyrosine residues and the frequency of multisite phosphorylation. These improvements include 1) the reduction of autofluorescence in the green spectral channel; 2) a simplified imaging chamber that accommodates higher sample number with lower sample volume; 3) corrections for membrane receptor surface expression; 4) three-color multiplex imaging; and 5) corrections for steric hindrance of dual antibody binding. These improvements enabled the first direct detection of multisite phosphorylation on full-length Epidermal Growth Factor Receptor (EGFR) and revealed that phosphorylation fraction varied by tyrosine residue. These SiMPull measurements provide a new level of detail in the status of receptor phosphorylation that was previously inaccessible with traditional biochemical techniques.

INTRODUCTION

The ability of a cell to respond rapidly and specifically to changes in the surrounding environment is controlled by protein-protein interactions at the plasma membrane and along the signaling cascade. While much is known about the biochemical events that govern signaling pathways, this information has mostly been derived from population-based measurements that typically average over millions of cells and/or proteins. However, there is growing evidence that the heterogeneity of the system contributes to how cellular information is processed[1–4]. To better understand the role of protein phosphorylation heterogeneity in directing signaling outcomes, we have adapted the single molecule pull down (SiMPull) assay to identifying the phosphorylation state of individual receptors.

SiMPull is a powerful technique that allows for interrogation of macromolecular complexes at the individual protein level. Jain *et al.* first demonstrated the ability of this technique to capture macromolecular complexes[5]. SiMPull samples are prepared in a manner similar to IP/Western Blot protocols, but the sample is interrogated using single molecule microscopy. Briefly, cells are lysed and the protein of interest is captured by antibodies bound to the coverglass. If the proteins are fluorescently tagged, either by fluorescent proteins or subsequent antibody labeling, their presence will be quantified by single molecule imaging (Fig. 1A).

Here, we describe a modification of SiMPull for the study of phosphorylation patterns of transmembrane receptors. Traditionally, protein phosphorylation has been measured using ensemble techniques, such as Western Blot analysis or flow cytometry, which provide information on the relative changes of a protein phosphorylation amount. However, these techniques cannot

determine the fractions of proteins in a specific phosphorylation state, much less identify when an individual protein contains multiple sites of phosphorylation. While mass spectrometry has the potential to detect multisite phosphorylation, the residues of interest must be found in the same small peptide that is generated by enzymatic digestion (typically 7-35 amino acids) or the protein of interest must be small[6–8]. Therefore, new techniques are needed to better understand the phosphorylation status of individual proteins. Recently, Kim *et al.* used a modified SiMPull approach, termed SiMBlot, to pull-down surface biotinylated proteins and identify phosphorylation using denaturing conditions and phosphorylation-specific antibody labeling[9]. Our approach differs in several significant ways from SiMBlot and provides important improvements over previous protocols[9,10], including the reduction of autofluorescence in the green spectral channel and a simplified imaging chamber that accommodates higher sample number with lower sample volume. We also demonstrate the importance of optimizing antibody labeling and fixation conditions. To quantify receptor phosphorylation, we have used two- and three-color imaging to identify individual proteins and their corresponding phosphorylation status. Corrections to account for membrane receptor surface expression and steric hindrance in the case of dual antibody labeling are described.

We apply this method to the study of the classical Epidermal Growth Factor Receptor (EGFR). EGFR has 20 tyrosines in its cytoplasmic tail, at least 12 of which are known to recruit specific adaptor proteins[11]. The potential for multisite phosphorylation provides a mechanism through which the cell might differentially respond to extracellular cues, depending on the extent and combination of receptor phosphorylation[3,12–14]. We found that only a subpopulation of EGFR become phosphorylated under what is considered maximal activation conditions and that the extent of phosphorylation varies by tyrosine residue. Multiplex imaging of the GFP-tagged receptors and antibodies directed to two distinct phosphotyrosines revealed that multisite phosphorylation frequently occurs. The extent of phosphorylation at individual tyrosines along with the existence of multisite phosphorylation has implications for how EGFR translates extracellular cues into downstream signaling outcomes.

RESULTS

Visualization of individual protein phosphorylation status by SiMPull

Phosphorylation of EGFR was assayed at the single molecule level using the SiMPull concept depicted in Figure 1A, where the GFP-tagged receptors from cell lysates are immunoprecipitated by antibodies bound to the coverglass and subsequently labeled with fluorescently-tagged antibodies detecting phosphorylated tyrosines (anti-PY). Figure 1B shows the capture of single EGFR-GFP from cell lysate on the coverglass surface and the corresponding labeling of phosphorylated EGFR using a pan-phosphotyrosine antibody conjugated to Alexa Fluor 647 (anti-PY AF647; PY99 antibody). Individual molecules are identified in each image (Fig. 1B). Images are then overlaid to identify phosphorylated receptors. Colocalization of receptor and PY localizations provide an initial estimate of the fraction of phosphorylated receptors (Fig. 1B). Optimization of the experimental process and image analysis, including reduction in autofluorescence, corrections for the level of receptor surface expression and the appropriate antibody labeling conditions, are described in the following sections.

Simplified sample chamber increases throughput and reduces sample volume

Jain *et al*. originally described the use of a fluidic chamber for imaging that consisted of 4-6 channels generated between a coverglass and microscope slide using epoxy[10]. While these

types of flow chambers are straightforward to produce, the protocol is time consuming (30-60 min) and larger volumes (~70 µL) are required to fill each channel. We developed a sample holder using a hydrophobic barrier pen to create isolated sample regions on a coverglass (Fig. 1C). Rectangular (24x60mm, #1.5) coverglasses are treated as described and an array of up to 20 squares can be drawn with the hydrophobic ink pen in a matter of minutes. As little as 10 µL of sample is needed to fill each region, which is seven times less than for the original flow channels. This is particularly useful considering the high cost of fluorescently labeled antibodies and that for some applications sample availability may be limited. Time for sample preparation is also reduced as washing and labeling steps are simplified without the need for flowing of buffers through channels.

Quenching with $NaBH_4$ reduces background autofluorescence

It has been previously noted that autofluorescent background is detected in the spectral region corresponding to green emitting fluorophores[10]. We also observed autofluorescent puncta in our green spectral channel (503-548 nm) that was identified as single GFP molecules in the absence of cell lysate. Since our experimental approach relies on GFP to identify the location of EGFR on the coverglass, it was important to reduce this background to avoid over-counting of receptors. We found that incubating the PEG-coated coverglass with Sodium Borohydride (10 mg/mL $NaBH_4$ for 4 min) significantly reduced the number of background fluorescent molecules (Fig. 2). Despite the improvement in background signal, we routinely acquired background measurements for each coverglass preparation to enable background correction for each experiment (see Methods for details).

Antibody optimization is required for accurate phosphotyrosine detection

Since antibodies are used to quantify protein phosphorylation, it is critical to optimize the antibody labeling protocol. Figure 3 presents the results from optimization of anti-EGFR-pY1068. In these experiments, EGF stimulated cells were co-treated with a phosphatase inhibitor (pervanadate, PV) to increase the amount of receptor phosphorylation. We found that 60 min incubation of EGFR-pY1068 on ice was needed to ensure maximal labeling (Fig. 3A). Importantly, we observed that over time the antibody would dissociate from EGFR, with ~37% reduction after 1 hr at room temperature (Fig. 3B, no fixative). Complete imaging of the sample array can take up to 1 hr, therefore, loss of antibody over this period would lead to an underestimate of receptor phosphorylation for samples imaged later in time. We tested multiple fixation protocols to minimize antibody unbinding and found that fixation with 4% Paraformaldehyde/0.1% Glutaraldehyde (PFA/GA) for 10 min stabilized the antibody levels for at least 1 hr (Fig. 3B, PFA/GA). To ensure that saturating levels of antibody are used concentration curves for each antibody were generated, using the PFA/GA fixation for optimal results. The example in Figure 3C shows the titration curve for anti-EGFR-pY1068, which saturates at ~20 µg/mL. Consequently, we used 20 µg/mL for all experiments. Binding affinity will vary for each antibody and fluorescent-conjugation may also alter antibody affinity. Therefore, it is necessary to perform a binding curve for each antibody and we routinely made this measurement after each new antibody conjugation. Another important consideration is the specificity of the antibody for its binding site. Kim et al demonstrated an elegant way to determine specificity by using purified proteins with individual tyrosines mutated to alanine. We use here the same EGFR-pY1068 and EGFR-pY1173 antibodies that they found to have high specificity from their *in vitro* measurements[9].

Correction is required to account for non-surface localized receptors

At any point in time, a fraction of membrane receptors is trafficking through intracellular compartments. These internal receptors are not accessible during addition of extracellular ligand, but will be captured by the antibody during SiMPull sample preparation and result in an underestimate of receptor phosphorylation. In our CHO-EGFR-GFP cells, we observed intracellular EGFR-GFP (Fig. 4A, left). To determine the fraction of EGFR accessible to ligand, we labeled all surface proteins on the CHO-EGFR-GFP cells with membrane-impermeable AF647-NHS Ester (Fig. 4A, right) and used SiMPull to visualize the amount of EGFR-GFP colocalized with AF647. By increasing the concentration of AF647-NHS until saturation is achieved, we estimated that ~65% of the receptors are located at the plasma membrane (Fig. 4B). With this information, we corrected our measurements to account for only those receptors available to bind ligand. After correction, we find that ~14% of the receptors are phosphorylated at Y1068 after 1 min stimulation with 50 nM EGF (Fig. 4C). We note that while surface labeling of receptors with AF647-NHS ester allows for identification of surface proteins, we found that pre-labeling of EGFR in this way reduced EGF binding (data not shown). Therefore, pre-labeling receptors in this way should not be used for the study of EGFR activation. To validate our correction method, we analyzed the phosphorylation levels of receptors from CHO cells expressing ACP-tagged EGFR. We directly labeled the plasma membrane localized EGFR using membrane-impermeable CoA-Atto488 as describe previously[15,16]. Cells were then exposed to EGF and probed for EGFR phosphorylation with SiMPull, this time using Atto488 as the marker for plasma membrane EGFR. The percent of phosphorylated EGFR was similar when comparing the membrane-localized ACP-EGFR and the membrane-corrected EGFR-GFP samples (Fig. 4C). Therefore, the effects of EGF binding to EGFR on the plasma membrane can be accurately determined from whole cell lysates and we apply this correction for the remainder of the results.

Extent of phosphorylation varies by tyrosine residue

We used SiMPull to characterize the kinetics and dose response of EGFR activation. The multi-well hydrophobic array (Fig. 1C) made it possible to efficiently examine a full dose response or time course of activation in a single imaging session. We quantified total EGFR tyrosine phosphorylation (PY) and the phosphorylation patterns for two specific tyrosine sites (Y1068 and Y1173). Cells simulated for 5 min (Fig. 5A) with increasing concentrations of EGF showed the expected increase in total phosphorylation with ligand dose (Fig 5A, PY, blue bars). This fraction reached 64% with 50 nM EGF, a dose that is considered saturating. While both specific tyrosines show less phosphorylation than total PY, the fraction of EGFR with phosphorylation at Y1173 was consistently higher than at Y1068 (Fig. 5A). The kinetics of phosphorylation between PY, pY1068 and pY1173 are similar (Fig. 5B). These results indicate several important outcomes. First, phosphorylation detection by SiMPull is sensitive, capable of detecting receptor phosphorylation at low ligand dose and early time points. Second, even under saturating ligand conditions, only a fraction of receptors is phosphorylated, reaching a maximum of 64% with 5 min stimulation. Third, the extent of phosphorylation varies by tyrosine residue. The detected phosphorylation levels are not due to limitations in antibody labeling, since cells stimulated in the presence of phosphatase inhibitors showed increased receptor phosphorylation (Supplementary Fig. 1A).The use of high salt (500 mM NaCl) concentration during cell lysis did not change the detected phosphorylation, indicating that adaptor proteins are not interfering with antibody recognition (Supplementary Fig. 1B).

Three-color SiMPull reveals multisite phosphorylation

The observation that Y1068 and Y1173 have different phosphorylation levels suggests that there are subpopulations of receptors with differing phosphorylation patterns. However, examining a single tyrosine site at a time cannot address the coincidence of phosphotyrosines. To assess the potential of multisite protein phosphorylation, we developed simultaneous three-color SiMPull imaging. We began by testing whether receptors phosphorylated at Y1068 were also phosphorylated at other tyrosine residues by co-labeling receptors with anti-pY1068 and anti-PY antibodies. When labeling a single protein with two or more antibodies, the effects of steric hindrance must be considered. In this case, we found that labeling first with anti-pY1068 followed by anti-PY did not alter PY levels (Supplementary Fig. 2). However, labeling with anti-PY first did cause a loss of pY1068 signal. Therefore, we performed the experiments with sequential labeling by anti-pY1068 first, followed by anti-PY. As before, the addition of EGF resulted in increased phosphorylation, with the PY antibody showing more labeling than the site-specific antibody (Fig. 6A,B) and the presence of multi-phosphorylated receptors was observed in the three-color images (Fig. 6A, white circles). Quantification of the three-color colocalization revealed that ~12% of EGFR were labeled by both antibodies (Fig. 6B, pY1068+PY, orange bar). Importantly, we find that nearly 76% of the receptors phosphorylated at Y1068 were co-labeled with PY. Therefore, multisite phosphorylation is a prevalent outcome in EGFR activation.

To ensure receptors detected as multi-phosphorylated were individual receptors labeled with both antibodies rather than two nearby labeled receptors detected as one in a diffraction-limited spot, we performed step-photobleaching analysis (Fig. 6C). Our analysis showed that the majority of doubly labeled receptors (~98%) were associated with a single EGFR-GFP molecule (Fig. 6C, right). It is important to note that the number of GFP spots demonstrating two-step photobleaching increased as the sample density increased (data not shown). Therefore, we recommend a pulldown protein density in the range of 0.04-0.08/μm^2. Alternatively, photobleaching traces can be performed in each measurement to exclude those spots showing more than one-step photobleaching.

With the knowledge that the majority of pY1068+ receptors are also phosphorylated in at least one other tyrosine residue, we examined the pairwise phosphorylation of Y1068 and Y1173. In contrast to dual labeling with pY1068 and PY, the close proximity of these two tyrosines did result in steric hindrance of antibody binding. This is seen as a reduction in labeling when the antibodies are applied second as compared to first (Supplementary Fig. 2). We used the reduced labeling as measured in sequential labeling controls to correct for steric blocking (see Supplementary Fig. 2 and Methods). We found that approximately 50% of the pY1068+ receptors are co-phosphorylated at Y1173. This is an enrichment of approximately two-fold as compared to pY1173+ in the total EGFR population (Fig. 6D). Stimulation of cells with lower doses of EGF also resulted in multisite EGFR phosphorylation (Fig. 6E). Notably, at 1 nM EGF, multi-phosphorylation is already considerable within the site-specific subpopulations, with 58 +/- 14% of pY1068+ receptors and 29 +/- 8% of pY1173+ receptors being co-labeled with PY. Therefore, multisite phosphorylation is not merely a consequence of saturating ligand conditions. The use of a three-color imaging scheme to correlate phospho-antibody labeling directly with GFP-tagged receptors was critical, due to the relatively high non-specific binding of the antibodies (Supplementary Fig. 3). In the absence of the GFP channel to remove the non-specific binding, the values for dual labeling are underestimated (Supplementary Fig. 3)**.** These results show that SiMPull, when performed using the improvements described here, can be used to quantify the extent and coincidence of phosphorylation at multiple tyrosines.

DISCUSSION

We have described a series of modifications to SiMPull that allow this technique to obtain quantitative information about multiple post-translational modifications (PTMs) at the single-protein level. Using SiMPull we have monitored EGFR phosphorylation patterns, quantified subpopulations of phosphorylated receptors, and directly observed the existence multisite phosphorylation. This approach holds distinct advantages over other techniques. Detailed information on protein PTMs is not accessible by traditional biochemical methods that can only determine relative changes from an average of the population. While mass spectrometry has the potential to detect multisite phosphorylation, the residues of interest must be found in close proximity[6]. In SiMBlot, which is also a single molecule approach to detecting PTMs, surface proteins are first biotinylated and then pulled-down via streptavidin-coated surface, rather than by a specific antibody[9]. The SiMPull method is not restricted to surface proteins and is therefore compatible with the interrogation of intracellular proteins. PTMs other than phosphorylation can also be studied as long as a suitable antibody is available. Thus, detection of PTMs by SiMPull enables measurements that were previously difficult to perform.

In addition to its advantages, SiMPull has a number of caveats that must be considered to ensure rigorous quantification. As with any antibody-dependent technique, the affinity and specificity of the antibody must be determined. We have shown that it is important to establish proper concentrations and labeling times for each antibody used, as well as the importance of post-fixation to prevent antibody dissociation during imaging. In addition, directly labeling the primary antibody with the fluorophore eliminates the need for secondary antibodies, which may add additional labeling efficiency artifacts and restrict options due to the limited availability of species used to generate primary antibodies. The phosphotyrosines probed in EGFR are located in an intrinsically disordered region of the C-terminal tail, therefore these sites are likely to be more accessible to antibodies than if they were located in structured regions. If the PTMs of interest are found in structured regions, a protein denaturation step may be used[9]. Steric hindrance of two or more antibodies binding to the same protein is another possible complication. We did find steric hindrance in the case of co-labeling pY1068 and pY1173. However, a simple control experiment and mathematical correction are described to avoid undercounting of dual-phosphorylated receptors. For future studies, the use of fluorescently-conjugated Fab fragments may reduce the impact of steric hindrance. It is worth noting that while detailed information on the phosphorylation status of individual proteins status is obtained, the cell-to-cell variability is lost with SiMPull.

Using SiMPull, we have quantified EGFR phosphorylation patterns. The new level of detail afforded by SiMPull has provided several important insights. First, we directly determined that only a fraction of EGFR was found to be phosphorylated, even under saturating ligand conditions. Second, the phospho-EGFR is further divided into subpopulations that vary in the extent of phosphorylation at individual tyrosine residues. Third, the use of three-color imaging allowed us to probe for multisite phosphorylation. Comparisons of pY1068 with a pan-phosphotyrosine antibody revealed that many receptors are indeed phosphorylated at more than one tyrosine simultaneously. Strikingly, the majority of pY1068+ receptors are co-labeled with PY and ~50% of pY1068+ are also positive for pY1173. These results are in contrast to recent SiMBlot studies of EGFR concluding that multisite phosphorylation was not a common occurrence. These differences may be explained by optimization of our labeling protocol that provided the sensitivity needed to detect multisite phosphorylation, including the use of fluorescently-conjugated primary antibodies, labeling under saturating antibody conditions and post-fixation to prevent antibody dissociation. Notably, our results are consistent with previous work indicating that multisite phosphorylation is important in the efficient recruitment of certain adaptor proteins to activated EGFR[17,18]. The existence of multisite phosphorylation holds significant functional implications. By modulating protein phosphorylation patterns, both single- and multisite combinations,

downstream signaling pathways may be differentially activated and lead to biased signaling. Consistent with this idea, it has been shown that biased signaling arises with different ligand types and doses, as well as with the relative abundance of receptors and their signaling partners[19–22].

Interestingly, we found that dual Y1068/Y1173 phosphorylation occurs about two-times more frequently than expected if these sites were independent of each other, suggesting positive correlation between the sites. Mechanistically, this enrichment could be a result of either long-lived receptor interactions or repeated dimerization events. If dimer lifetimes are sufficiently long, then phosphorylation of multiple sites could happen in a single dimerization event, suggesting that phosphorylation occurs in a semi-processive manner. Alternatively, if a receptor undergoes many dimerization and dissociation events, then these repeated interactions could result in the phosphorylation of a unique tyrosine in each encounter. This would be similar to quasi-processive phosphorylation as described for ERK[23]. Experimental evidence exists to support each of these mechanisms. Both long-lived and transient EGFR dimerization has been observed on living cells, with dimer lifetimes dependent on ligand occupancy[24]. Recent work from the Lemmon group has shown that high and low affinity ligands induce distinct dimer structures, where low affinity ligands lead to less stable dimers and differential signaling outcomes[21]. Oncogenic signaling by EGFR mutants has been shown to be driven by enhanced dimerization and increased catalytic activity that could amplify multi-phosphorylation[15,25–27]. Interplay between receptors and the membrane environment has also been shown to affect the efficiency of EGFR encounters[24,28]. Therefore, the frequency of dimerization and the duration of dimer lifetimes may serve as a kinetic proofreading mechanism, regulating the EGFR phosphorylation patterns and dictating cellular outcome. Additionally, adaptor protein binding and phosphatase activity likely play roles in phosphorylation extent. For example, Capuani et al have shown that Grb2 and Cbl can protect Y1045/Y1068 from dephosphorylation[17]. These mechanisms are not mutually exclusive and may be more or less relevant depending on the cellular contexts. We would, therefore, predict that phospho-EGFR patterns will be modulated by differences in ligand dose, ligand-dependent dimer lifetimes, membrane architecture and adaptor protein abundance. The contributions that these mechanisms have in EGFR and other signaling pathways, remain unclear. The unique datasets provided by SiMPull, combined with other experimental and computational modeling tools, should prove useful in addressing these types of questions.

Acknowledgements

This work was supported by the New Mexico Spatiotemporal Modeling Center (NIH P50GM085273), Mexican Secretariat of Public Education SEP (E.S.-C.), National Science Foundation CAREER MCB-0845062 (D.S.L.), and the UNM Comprehensive Cancer Center (NIH P30CA118100). We thank Dr. Bridget Wilson, Dr. William Hlavacek for useful discussions, and Shayna Lucero for assistance with cell culture. We thank Dr. Donna Arndt-Jovin for the CHO-ACP-EGFR cell line and suggesting the ACP-EGFR comparison experiment. We gratefully acknowledge use of the University of New Mexico Comprehensive Cancer Center fluorescence microscopy and flow cytometry facilities, as well as the NIH P30CA118100 support for these cores. The authors declare no competing financial interests.

Author contributions

E.S.-C. performed all experiments and data analysis; D.S.L and E.S.-C. designed experiments and wrote the manuscript; E.S.-C. and K.A.L. developed analysis methods and algorithms. All authors discussed the results and commented on the manuscript.

Competing financial interests

The authors declare no competing financial interests.

METHODS

Cell lines and reagents

CHO cells expressing GFP-tagged[29,30] or ACP-tagged EGFR (provided by Dr. Donna Arndt-Jovin) were cultured in DMEM supplemented with 10% FBS, penicillin–streptomycin and 2 mM L-glutamine (Thermo Fisher Scientific). ACP-tagged EGFR was as described in [15,16] with the exception that a shortened 16 aa sequence was introduced at the EGFR N-terminus[31]. EGF, Protease and Phosphatase Inhibitor Cocktail, Alexa Fluor 647 NHS Ester, and NeutrAvidin were purchased from Thermo Fisher Scientific. CoA 488 and ACP Synthase were purchased from New England Biolabs. N-(2-aminoethyl)-3-aminopropyltrimethoxysilane was purchased from United Chemical Technologies (#A0700). Sodium bicarbonate and sodium borohydride were purchased from EMD Millipore (#SX0320-1, #SX0380-3). mPEG-Succinimidyl Valerate (MPEG-SVA-5000-5g) and biotin-PEG-Succinimidyl Valerate (Biotin-PEG-SVA-5000-500mg) were from Laysan Bio. Biotinylated anti-EGFR antibody (E101) was obtained from Leinco Technologies. Antibodies in carrier-free buffer were purchased from Cell Signaling Technologies: EGFR pY1068 (clone 1H12, 2236BF) and EGFR pY1173 (clone 53A5, 4407BF). Monoclonal antibody pre-labeled with AF647 to detect pan-tyrosine phosphorylation (PY99 antibody, sc-7020 AF647) was purchased from Santa Cruz Biotechnology. Mix-n-Stain CF555 and CF640R antibody labeling kits were purchased from Biotium Inc. Paraformaldehyde and glutaraldehyde were purchased from Electron Microscopy Sciences.

Labeling of antibodies

Carrier-free antibodies (50 µg at 0.5-1 mg/mL per reaction) were labeled using Mix-n-Stain antibody labeling kits following the manufacturer's instructions. Briefly, the labeling reaction was carried out for 30 min at room temperature and antibodies were centrifuged using the ultrafiltration vial provided to remove the unconjugated dye. Antibodies were resuspended in PBS and stored at 4 °C. The labeling efficiency achieved was between 2.7-4.4 dyes/antibody.

Cell treatment and lysate preparation

CHO-EGFR-GFP cells were plated overnight in 60 mm tissue culture dishes at 800,000 cells/dish and CHO-ACP-EGFR cells in 24-well plates at 50,000 cells/well. For ACP labeling, CHO-ACP-EGFR cells were washed with serum-free DMEM medium (SFM), incubated with ACP labeling solution (SFM, 10 mM $MgCl_2$, 4 µM CoA 488 and 1 µM ACP) for 20 minutes at 37°C and washed three times with SFM previous to stimulation. Cells were washed in Tyrode's solution (135 mM NaCl, 10 mM KCl, 0.4 mM $MgCl_2$, 1 mM $CaCl_2$, 10 mM HEPES, 20 mM glucose, 0.1% BSA, pH 7.2) and treated with 25 nM EGF or Tyrode's solution alone (resting cells) at 37°C. At the indicated time points, cells were placed on ice, washed one time with cold PBS followed by addition of lysis buffer (1% IGEPAL CA-630, 150 mM NaCl, 50 mM Tris pH 7.2) containing Protease and Phosphatase Inhibitors. Cell lysates were collected using cell scrapers (Greiner Bio-One North

America, #541070), transferred to fresh tubes on ice and vortexed every 5 min for a total of 20 min. Lysates were centrifuged at 16,000× g for 20 min at 4 °C and the supernatant was transferred to a new tube and stored at -80 °C. For experiments involving treatment of cells with phosphatase inhibitors, cells were pre-treated for 15 min with a Tyrode's solution containing 1 mM pervanadate (PV) followed by incubation for 5 min in a solution with 50 nM EGF and 1 mM PV. A stock solution of 30 mM PV was prepared before each experiment by mixing equimolar concentrations of hydrogen peroxide and activated sodium orthovanadate that was incubated in the dark for at least 15 min before use.

Fabrication of hydrophobic arrays and surface functionalization

Coverglasses (24x60mm, #1.5; Electron Microscopy Sciences, #63793) were Piranha-cleaned[32] and placed in a coverglass holder (Fisher Scientific, #08-817). Coverglasses were sequentially sonicated in Methanol and Acetone for 10 min each, and in 1M KOH for 20 min using a bath sonicator (Branson Ultrasonics, B1200R-1). These solutions were stored in polypropylene 50 mL tubes (VWR, #89401-564) and reused up to five times. Coverglasses were rinsed with Milli-Q water two times, dried by quickly passing them multiple times over the flame of a Bunsen burner using metal tweezers and placed in a dry coverglass holder. A solution containing 76 mL of methanol, 4 mL of acetic acid and 0.8 mL of aminosilane (N-(2-aminoethyl)-3-aminopropyltrimethoxysilane) was prepared in an Erlenmeyer flask, immediately poured into the coverglass holder and incubated at room temperature for 10 min in the dark, followed by 2 min sonication and another 10 min incubation in the dark. Coverglasses were next washed with methanol for 2 min, rinsed and washed for 2 min with water, and dried in the dark. Treated coverglasses were placed on top of a parafilm-covered coverglass containing a guide pattern, which was used as reference to draw the Sample Array with a hydrophobic barrier pen (Vector Laboratories, #H-4000). Ink was allowed to dry for at least 5 min before coverglasses were placed in a humidified chamber (empty tip rack with 50 mL of water; USA Scientific #1111-2820). For surface functionalization, 50 mg of mPEG-Succinimidyl Valerate, 1.3 mg of biotin-PEG-Succinimidyl Valerate and 200 µL of freshly prepared 10 mM sodium bicarbonate were mixed thoroughly by pipetting up and down, centrifuged for 1 min at 10,000 g at room temperature and immediately applied to the SiMPull array (10-13 uL per region). After incubating for 3-4 hours in the dark inside the humidified boxes, arrays were washed by sequential 30 sec submersions into three water-filled 250 mL glass beakers. Coverglasses were dried with nitrogen gas, stored in pairs (back to back) inside 50 mL tubes, which were filled with nitrogen gas before closing and sealing with Parafilm. Coverglasses were stored in the dark at -20°C for up to a week before use.

Labeling and quantification of surface receptors

CHO-EGFR-GFP cells grown in 24-well plates were placed on ice and washed 3 times with cold PBS. AF647-NHS Ester was dissolved at the indicated concentrations in PBS. Cells were incubated with this solution for 30 min at 4°C with gentle agitation, washed 3 times with cold PBS and subjected to cell lysis. The percent of receptors labeled with AF647 across different dye concentrations was assessed with SiMPull. To estimate the percent of receptors at the cell surface the AF647-labeling curve was fitted to a biexponential decay curve in its increasing form using the 'fit' function in MATLAB: $y = C1\ (1 - e^{-ax}) + C2\ (1 - e^{-bx})$, where y is the % of AF647-labeled receptors, x is the concentration of reactive AF647-NHS ester used, and $a>0$, $b>0$, $C1$ and $C2$ are coefficients to be fitted. The sum of the coefficients $C1$ and $C2$ represent the asymptote of the curve and an approximation of the fraction of receptors at the cell surface.

Single-Molecule Pulldown and phospho-site labeling

T50 (10 mM Tris pH 8.0, 50 mM NaCl) and T50-BSA (T50 with 0.1 mg/mL BSA) solutions were prepared and stored for up to a month at room temperature. SiMPull arrays were equilibrated at room temperature and placed on a TC100 plate lined with Parafilm. Each region of the SiMPull array was treated with 10-15 µL of a 10 mg/mL sodium borohydride ($NaBH_4$)/PBS solution for 4 min at room temperature and washed 3 times with PBS. SiMPull regions were then incubated with a 0.2 mg/mL NeutrAvidin/T50 solution for 5 min and washed three times with T50, followed by incubation with a 2 µg/mL biotinylated anti-EGFR/T50-BSA solution for 10 min and washed three times with T50-BSA.

The plate containing the SiMPull array(s) was kept on ice during sample preparation. Lysates were diluted in cold T50-BSA with Protease and Phosphatase Inhibitors (T50-BSA/PPI), vortexed at medium speed, and added to the SiMPull array. After 10 min incubation, the lysates were removed and the SiMPull regions washed 4 times with cold T50-BSA/PPI. To determine appropriate dilution factor, the density of pulldown receptors as a function of lysate concentration was first assessed to achieve a pulldown density 0.04-0.08/μm^2. Antibodies were diluted in cold T50-BSA/PPI, incubated for 1 hr, washed 6 times with cold T50-BSA for a total of 6-8 minutes, and washed twice with cold PBS. Immediately after, antibodies were fixed for 10 min with a 4% PFA/0.1% GA solution (paraformaldehyde/glutaraldehyde) and washed 2 times with 10 mM Tris (pH 7.4)/PBS for a total of 10 min to inactivate fixatives. For 3-color SiMPull experiments the same antibody incubation and fixation procedure was performed for the second antibody. Tris solution was replaced by T50-BSA and the SiMPull array was equilibrated to room temperature before proceeding to imaging.

SiMPull imaging

Imaging of SiMPull samples was performed using an inverted microscope (Olympus America, model IX71) equipped with a 150×/1.45 NA oil-immersion objective for Total Internal Reflection Fluorescence Microscopy (Olympus America, UAPON 150XOTIRF) and a three-dimensional piezostage (Mad City Labs, Nano-LPS100). Excitation of CF640R- or AF647-labeled antibodies was done using a 642-nm laser (Thorlabs, HL63133DG), CF555-labeled antibodies using a 561-nm laser (Coherent Inc, Sapphire 561-100 CW CDRH), and of GFP- and CoA 488-tagged receptors using a 488-nm laser (Spectra Physics, Cyan 100mW). All lasers were set in total internal reflection configuration, and laser powers were adjusted to prevent photobleaching of the sample at the timescale of the image exposure time (300 msec). Sample illumination and emission were filtered using a quad-band dichroic and emission filter set (Semrock, LF405/488/561/635-A-000). Emission light was separated into four channels using a quad-view multichannel imaging system (Photometrics, model QV2) equipped with the appropriate dichroics (Chroma, 495 DCLP, 565 DCLP, 660 DCLP) and emission filters (Semrock, 685/40 nm, 600/37 nm, 525/45 nm). Emission light was collected with an electron-multiplying charge-coupled device (EMCCD) camera (Andor Technology, DU-897E-C50-#BV) with EM gain set to 200. Each channel was 256 x 256 pixels, with a pixel size of 106.7 nm. Photobleaching and bleed through were prevented by controlling the laser shutters and microscope stage through a MATLAB script to sequentially excite and acquire the different fluorophores (642-nm laser first, 488-nm laser last). A minimum of 20 regions of interest were acquired per condition. For quantification of step photobleaching of EGFR-GFP molecules, a 100 frame time series (300 msec exposure time) was acquired after imaging of the other two channels.

Quantification of Receptor Phosphorylation

All image processing was performed using MATLAB together with the MATLAB toolbox for image-processing DIPImage (Delft University of Technology)[33] and all software is available upon request. The location of emitters in each channel was calculated using graphics processor unit (GPU) computing as previously described[34]. Fits in the GFP channel were filtered based on the quality of the fit to the point spread function to reduce the chances of detecting multiple receptors in close proximity as a single molecule. Image registration was performed as previously described [35]. In our case, the root mean square error for image registration was <10 nm. For visualization purposes, Gaussian blob representations of the fluorophore localizations were generated. A receptor was considered to be phosphorylated when the localization centers of the receptor and labeled antibody were at a distance <106.7 nm (within 1 pixel).

Phosphorylation percentages were calculated as $100*(N_{Phos})/(N_{GFP}-N_{BG})$ where N_{Phos} is the number of receptors identified as phosphorylated, N_{GFP} is the number of observed single molecules in the GFP channel and N_{BG} is the non-specific background rate in the GFP channel.

The number of GFP localizations was calculated by subtracting background spots and accounting only for surface receptors as follows: $N_{GFP} = (N_{LOC}-N_{BG})*SR$, where N_{LOC} is the total number of emitters localized, N_{BG} is the expected number of background emitters in the area imaged, and *SR* (surface ratio) is the fraction of receptors located at the cell surface. The density of background emitters was quantified for each SiMPull array and used for background correction of samples in that array. For 3-color SiMPull experiments where steric hindrance between sequentially incubated antibodies was observed (i.e. pY1068-pY1173 detection), estimations of dual phosphorylation were corrected to account for this hindrance as explained in Supplementary Fig. 2.

Statistical Analysis

Based on the consideration that the phosphorylation state of each receptor analyzed has the properties of a Bernoulli trial, standard errors (SE) of phosphorylation measurements were calculated as for sample proportions in a binomial distribution: $SE= p(1-p)/n$, where p is the fraction of receptors phosphorylated and n is the total number of receptors. The condition np>10 (with the exception of Figure 6E, np>5) and np(1-p)>10 was ensured to be met to allow this approximation to be adequate. Two-sample Z-test (two-tailed) was used to estimate p-values[36].

Step-photobleaching Analysis

For step-photobleaching analysis of multi-phosphorylated receptors the average fluorescence intensity of the area (200x200 nm) surrounding each of these EGFR-GFP molecules was quantified and plotted for the duration of the time series. Intensity plots were manually analyzed and the number of photobleaching steps was quantified. For a small fraction of the emitters the number of molecules could not be reliably counted because either they photobleached too quickly (<2 frames) or did not photobleach during the duration of the movie, and therefore were excluded from the analysis.

REFERENCES

1. Lahav, G. *et al.* Dynamics of the p53-Mdm2 feedback loop in individual cells. *Nat. Genet.* **36,** 147–150 (2004).

2. Feinerman, O., Veiga, J., Dorfman, J. R., Germain, R. N. & Altan-Bonnet, G. Variability and Robustness in T Cell Activation from Regulated Heterogeneity in Protein Levels. *Science.* **321,** 1081–1084 (2008).

3. Coba, M. P. *et al.* Neurotransmitters Drive Combinatorial Multistate Postsynaptic Density Networks. *Sci. Signal.* **2,** ra19 (2009).

4. Spencer, S. L., Gaudet, S., Albeck, J. G., Burke, J. M. & Sorger, P. K. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature* **459,** 428–432 (2009).

5. Jain, A. *et al.* Probing cellular protein complexes using single-molecule pull-down. *Nature* **473,** 484–488 (2011).

6. Swaney, D. L., Wenger, C. D. & Coon, J. J. Value of using multiple proteases for large-scale mass spectrometry-based proteomics. *J. Proteome Res.* **9,** 1323–1329 (2010).

7. Curran, T. G., Zhang, Y., Ma, D. J., Sarkaria, J. N. & White, F. M. MARQUIS: A multiplex method for absolute quantification of peptides and posttranslational modifications. *Nat. Commun.* **6,** 5924 (2015).

8. Brunner, A. M. *et al.* Benchmarking multiple fragmentation methods on an orbitrap fusion for top-down phospho-proteoform characterization. *Anal. Chem.* **87,** 4152–4158 (2015).

9. Kim, K. L. *et al.* Pairwise detection of site-specific receptor phosphorylations using single-molecule blotting. *Nat. Commun.* **7,** 11107 (2016).

10. Jain, A., Liu, R., Xiang, Y. K. & Ha, T. Single-molecule pull-down for studying protein interactions. *Nat. Protoc.* **7,** 445–452 (2012).

11. Schulze, W. X., Deng, L. & Mann, M. Phosphotyrosine interactome of the ErbB-receptor kinase family. *Mol. Syst. Biol.* **1,** 2005.0008 (2005).

12. Gibson, S. K., Parkes, J. H. & Liebman, P. A. Phosphorylation modulates the affinity of light-activated rhodopsin for G protein and arrestin. *Biochemistry* **39,** 5738–5749 (2000).

13. Salazar, C. & Höfer, T. Multisite protein phosphorylation - From molecular mechanisms to kinetic models. *FEBS J.* **276,** 3177–3198 (2009).

14. Lau, E. K. *et al.* Quantitative Encoding of the Effect of a Partial Agonist on Individual Opioid Receptors by Multisite Phosphorylation and Threshold Detection. *Sci. Signal.* **4,** ra52 (2011).

15. Valley, C. C. *et al.* Enhanced dimerization drives ligand-independent activity of mutant epidermal growth factor receptor in lung cancer. *Mol. Biol. Cell* **26,** 4087–4099 (2015).

16. Ziomkiewicz, I. *et al.* Dynamic conformational transitions of the EGF receptor in living mammalian cells determined by FRET and fluorescence lifetime imaging microscopy. *Cytom. Part A* **83,** 794–805 (2013).

17. Sigismund, S. *et al.* Threshold-controlled ubiquitination of the EGFR directs receptor fate.

EMBO J. **32,** 2140–2157 (2013).

18. Fortian, A. & Sorkin, A. Live-cell fluorescence imaging reveals high stoichiometry of Grb2 binding to the EGF receptor sustained during endocytosis. *J. Cell Sci.* **127,** 432–444 (2014).

19. Chen, W. W. *et al.* Input–output behavior of ErbB signaling pathways as revealed by a mass action model trained against dynamic data. *Mol. Syst. Biol.* **5,** 239 (2009).

20. Wilson, K. J. *et al.* EGFR ligands exhibit functional differences in models of paracrine and autocrine signaling. *Growth Factors* **30,** 107–16 (2012).

21. Freed, D. M. *et al.* EGFR Ligands Differentially Stabilize Receptor Dimers to Specify Signaling Kinetics. *Cell* **171,** 683–695 (2017).

22. Wolf-Yadlin, A. *et al.* Effects of HER2 overexpression on cell signaling networks governing proliferation and migration. *Mol. Syst. Biol.* **2,** 54 (2006).

23. Aoki, K., Yamada, M., Kunida, K., Yasuda, S. & Matsuda, M. Processive phosphorylation of ERK MAP kinase in mammalian cells. *Proc. Natl. Acad. Sci.* **108,** 12675–12680 (2011).

24. Low-Nam, S. T. *et al.* ErbB1 dimerization is promoted by domain co-confinement and stabilized by ligand binding. *Nat. Struct. Mol. Biol.* **18,** 1244–1249 (2011).

25. Kim, Y. *et al.* Temporal resolution of autophosphorylation for normal and oncogenic forms of EGFR and differential effects of gefitinib. *Biochemistry* **51,** 5212–5222 (2012).

26. Shan, Y. *et al.* Oncogenic mutations counteract intrinsic disorder in the EGFR kinase and promote receptor dimerization. *Cell* **149,** 860–870 (2012).

27. Zhang, X., Gureasko, J., Shen, K., Cole, P. A. & Kuriyan, J. An Allosteric Mechanism for Activation of the Kinase Domain of Epidermal Growth Factor Receptor. *Cell* **125,** 1137–1149 (2006).

28. Chung, I. *et al.* Spatial control of EGF receptor activation by reversible dimerization on living cells. *Nature* **464,** 783–787 (2010).

29. Brock, R., Hamelers, I. H. L. & Jovin, T. M. Comparison of fixation protocols for adherent cultured cells applied to a GFP fusion protein of the epidermal growth factor receptor. *Cytometry* **35,** 353–362 (1999).

30. Lidke, D. S. *et al.* Quantum dot ligands provide new insights into erbB/HER receptor–mediated signal transduction. *Nat. Biotechnol.* **22,** 198–203 (2004).

31. George, N. A new method for protein labeling with small molecules based on acyl carrier protein. (EPFL, 2006). doi:10.5075/epfl-thesis-3675

32. Labit, H. *et al.* A simple and optimized method of producing silanized surfaces for FISH and replication mapping on combed DNA fibers. *Biotechniques* **45,** 649–658 (2008).

33. Hendriks, C. L. L., van Vliet, L. J., Rieger, B., van Kempen, G. M. P. & van Ginkel, M. DIPimage: a scientific image processing toolbox for MATLAB. *Quant. Imaging Group, Fac. Appl. Sci. Delft Univ. Technol. Delft, Netherlands* (1999).

34. Smith, C. S., Joseph, N., Rieger, B. & Lidke, K. A. Fast, single-molecule localization that achieves theoretically minimum uncertainty. *Nat. Methods* **7,** 373–375 (2010).

35. Schwartz, S. L. *et al.* Differential Mast Cell Outcomes Are Sensitive to FcεRI-Syk Binding Kinetics. *Mol. Biol. Cell* **28,** 3397–3414 (2017).

36. LeBlanc, D. C. *Statistics: concepts and applications for science*. **Vol. 2,** (Jones & Bartlett Learning, Burlington, 2004).

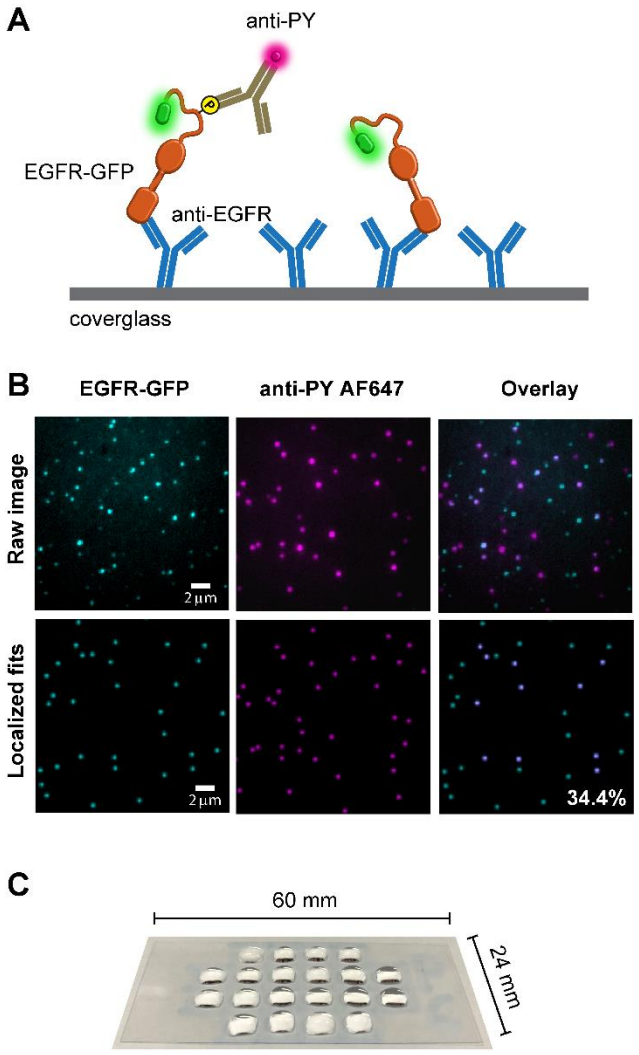

Figure 1. **SiMPull to quantify protein phosphorylation.** **(A)** Illustration depicting overall principle for assessing phosphorylation at the single molecule level using GFP-tagged EGFR (EGFR-GFP) as an example. **(B)** Representative images showing raw data (top) and blob-reconstructed localized molecules (bottom). CHO-EGFR-GFP cells were stimulated for 5 min with 25 nM EGF at 37°C before lysis for SiMPull. Raw images are brightness and contrast enhanced for visualization. The EGFR-GFP fits were filtered based on their fit to the microscope point spread function and the GFP-channel used as a mask to create the overlay. The number in the bottom right image represents the phosphorylation percentage estimated for this field of view. **(C)** Hydrophobic array for preparation of SiMPull samples.

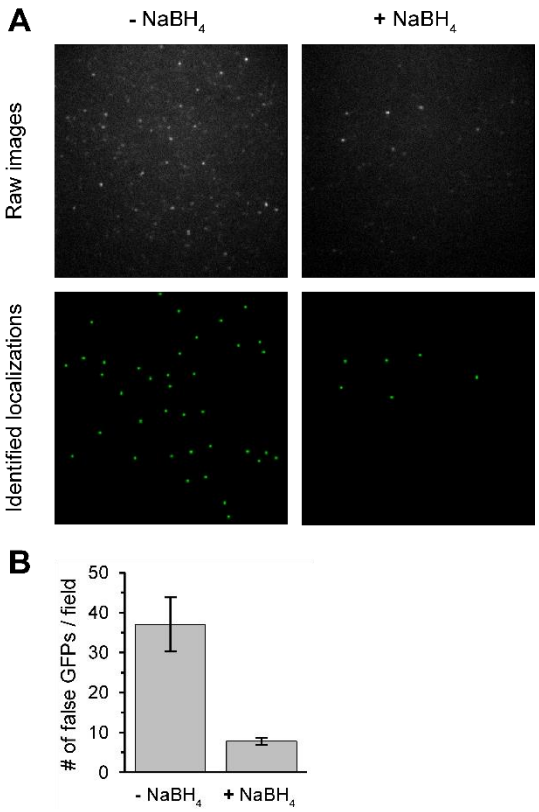

Figure 2. **Reduction of autofluorescence with Sodium Borohydride ($NaBH_4$). (A)** Raw images and blob-reconstructions from a typical field of view of a PEG/PEG-biotin functionalized surface without (left) and with (right) $NaBH_4$-treatment. **(B)** Quantification of the average number of false-positive localizations per field of view in surfaces with or without treatment with $NaBH_4$. For each condition N > 12 fields of view were analyzed. Error bars represent mean +/- S.E.M.

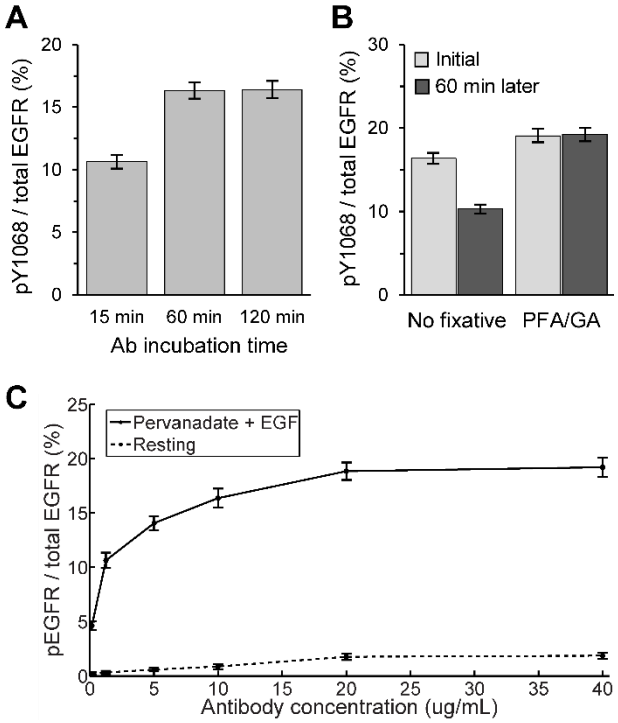

Figure 3. **Optimization of antibody labeling for accurate quantification of receptor phosphorylation.** CHO-EGFR-GFP cells were pre-treated with 1 mM PV for 15 min and stimulated with 50 nM EGF+1mM PV for 5 min at 37°C to enhance receptor phosphorylation and interrogated for anti-EGFR-pY1068-CF555 labeling. **(A)** Antibody labeling with anti-pY1068 requires 60 min to reach maximal labeling. A 20 μg/mL antibody concentration was used. Number of receptors analyzed per condition, N>3400. **(B)** Addition of PFA/GA post-fixation prevents loss of antibody over time. N>2700 per condition. **(C)** Increase in labeling as a function of antibody dose. EGFR-pY1068-CF555 saturates at ~20 μg/mL. Antibody was incubated for 1 hour on ice and post-fixed with PFA/GA. Resting cells were used as a control for non-specific labeling. N>1700 per data point. Error bars are standard error of measured phosphorylation percentages.

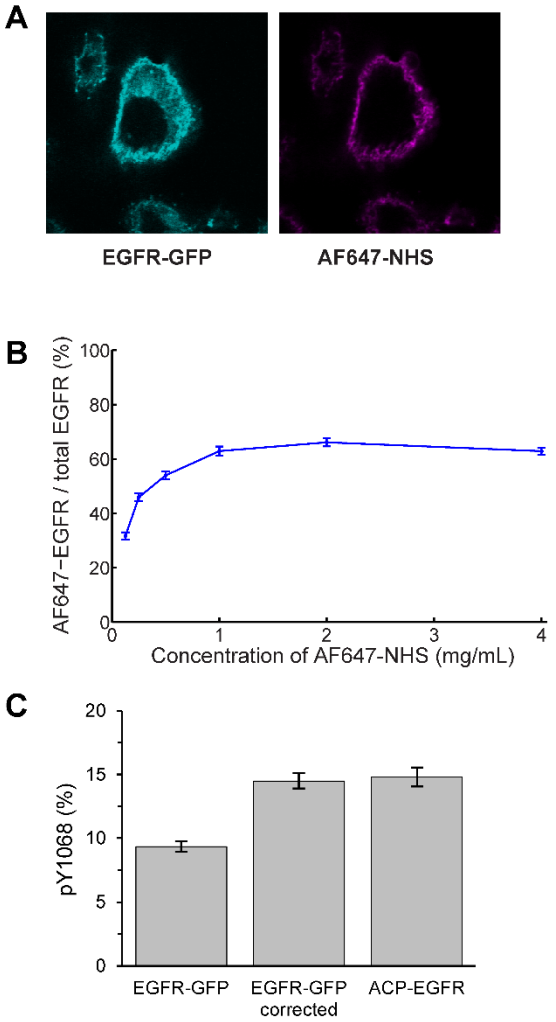

Figure 4. **Correction for cellular distribution of receptors. (A)** Confocal images showing typical distribution of EGFR-GFP in CHO cells (left) and the labeling of surface proteins achieved with the AF647-NHS ester (right). **(B)** Cells were incubated with increasing concentrations of AF647-NHS and assayed by SiMPull to determine the percentage of EGFR-GFP molecules labeled with AF647. Number of receptors analyzed per data point, 850 < N < 1550. **(C)** Percentage of pY1068+ receptors estimated for EGFR-GFP before and after correcting for surface expression. The corrected phosphorylation percentage for EGFR-GFP corresponds to the value measured for ACP-EGFR, which only includes plasma membrane localized receptors. N>2400 for each EGFR type. Error bars are standard error of measured phosphorylation percentages.

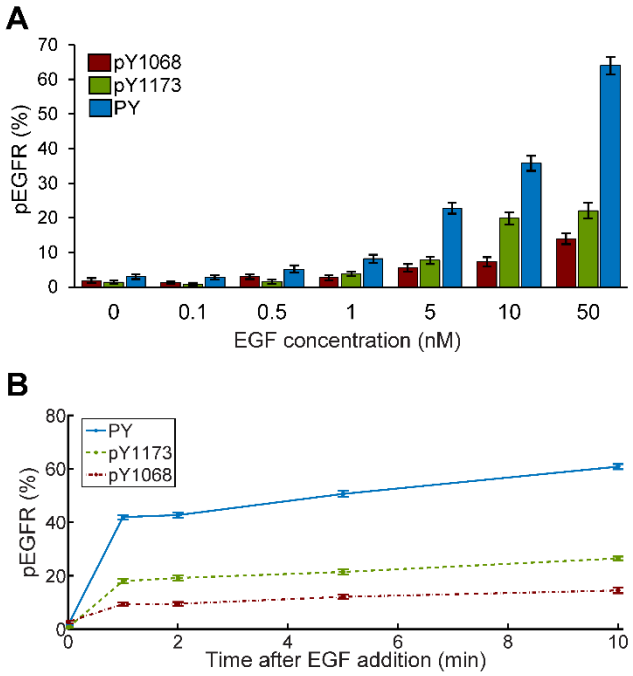

Figure 5. The extent of phosphorylation varies by tyrosine residue. (A) Dose response curve for CHO-EGFR-GFP cells after 5 min of EGF addition at 37°C. Number of receptors analyzed per condition, 800 < N < 1800. **(B)** Site-specific EGFR phosphorylation kinetics. Phosphorylation time course for CHO-EGFR-GFP cells stimulated with 25 nM EGF at 37°C. N>1800. Error bars are standard error of measured phosphorylation percentages.

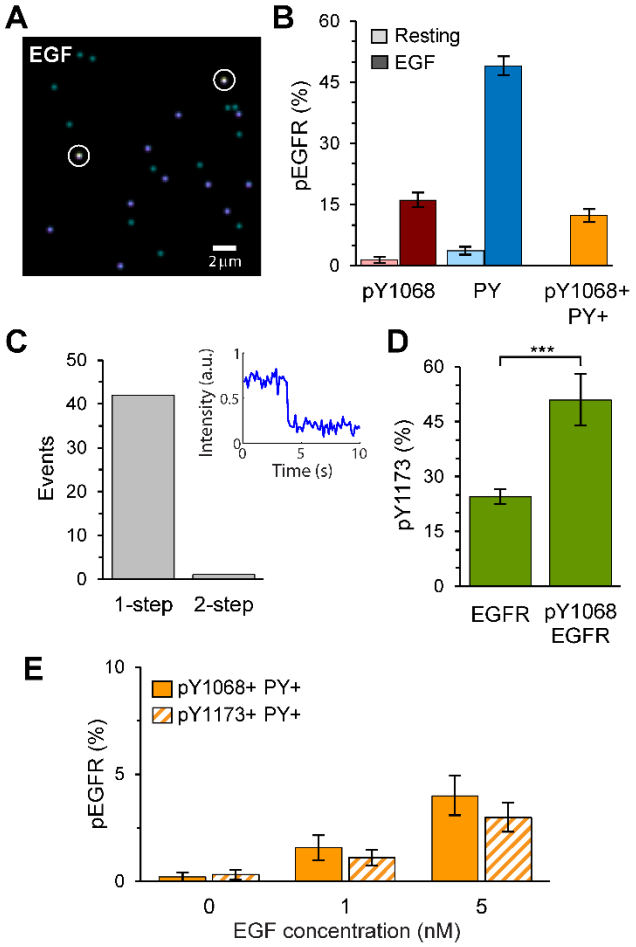

Figure 6. SiMPull reveals EGFR multisite phosphorylation. (A) Representative 3-color SiMPull image showing detection of EGFR-GFP (cyan), where receptors positive for PY labeling appear purple and white circles mark receptors labeled for both PY and pY1068. This image does not contain receptors labeled with pY1068 alone. Cells were treated with 25 nM EGF for 5 min. **(B)** Quantification of single and multi-phosphorylation in EGFR. Number of receptors analyzed per condition, N>500 for resting condition and N>840 for EGF condition. **(C)** Step-photobleaching analysis of multi-phosphorylated EGFR-GFP from (B). The majority (98%) of diffraction limited GFP spots show single-step bleaching, consistent with the pull-down of receptors as monomers. Inset shows example GFP-intensity trace of a multi-phosphorylated EGFR-GFP. **(D)** Percentage of Y1173 phosphorylation in overall population of surface receptors compared to that in pY1068+ receptors. N>780 for EGFR and N=51 for pY1068 EGFR. *** $P < 0.001$. **(E)** Multisite phosphorylation is also observed at lower EGF dose. Cells stimulated for 5 min with indicated EGF dose. $970 < N < 1700$ per condition. Error bars are standard error of measured phosphorylation percentages.