Supporting Information

Yen et al. 10.1073/pnas.1507329112

SI Materials and Methods

Cell Culture. Lung cancer cell lines were cultivated in RPMI medium 1640 (Life Technologies) supplemented with 10% (vol/vol) FBS, 100 U/mL penicillin, 100 μg/mL streptomycin, 1 mM L-glutamine, and 1 mM sodium pyruvate at 37 °C in 5% (vol/vol) CO_2. To identify EGFR phosphorylation, CL1-0 and CL1-5 cells were cultured overnight in a Hyperflask (Corning) with a total number of 5×10^7 cells and then were starved in serum-free RPMI medium for 24 h.

Purification of FLAG-Tagged sEGFR. By using anti-FLAG (M2) agarose (Sigma-Aldrich), recombinant sEGFR was affinity purified from 50 mL of the concentrated culture medium (~10 fold) from 293F cells transiently overexpressing FLAG-tagged sEGFR. The concentrated medium first was treated with sialidase (from *Clostridium perfringens*, 40 mU/mL; Roche) at 37 °C overnight in the presence of EDTA-free protease inhibitor mixture (Roche) before affinity purification. The anti-FLAG affinity column containing 500 μL of agarose beads was washed three times with 10 mL Tris-buffered saline [TBS, 25 mM Tris (pH 7.4), 137 mM NaCl, 2.7 mM KCl] before incubation with sEGFR-containing medium, and sEGFR was eluted with 0.1 M glycine-HCl (pH 2.5) followed by neutralization with 1:15 volume of 1 M Tris·HCl (pH 9). The eluted FLAG-tagged sEGFR was buffer-exchanged to 50 mM sodium phosphate and SPR running buffer [10 mM Hepes (pH 7.4), 150 mM NaCl, 3 mM EDTA, 0.005% Tween-20] for MALLS and SPR assay, respectively.

Desialylation and Detection of Sialylation of Desialylated Samples. Sialidase (α2,3/6/8-sialidase purchased from Roche; 100 mU/mL) was used to treat cell lysate (1 mg/mL) by incubation at 4 °C for overnight before affinity purification. For treatment of cancer cells, 10^6 cells were seeded in a 10-cm dish and incubated in RPMI medium 1640 supplemented with 0.1% BSA and sialidase (20 mU/mL; Roche) at 37 °C overnight. For treatment with STI, 50 μM of STI was added to the culture medium, and the cells (10^5/mL) were treated for 3 d. To detect the level of sialylation, the glycoforms of sialidase-treated sEGFR were analyzed by mass spectrometry as described previously (1). For monitoring the removal of sialic acids on flEGFR used for in vitro phosphorylation assays, 200 μg of sialidase-treated lysates was immunoprecipitated with 2 μg of biotinylated SNA lectin (Vector Laboratories) and 50 μL of NeutrAvidin beads (Thermo Scientific) in 0.4 mL lysis buffer, followed by anti-EGFR immunoblotting, and the relative sialylation of EGFR was quantified by the intensity of SNA-precipitated EGFR/total EGFR in cell lysate. To detect the levels of sialylation in desialylated cells, 1 μg of biotinylated SNA or *Maackia amurensis* lectin II (MALII) was used to stain 10^5 cells at 4 °C for 30 min in 50 μL FACS staining/washing buffer [2% (vol/vol) FCS, 0.1% NaN3 in PBS], followed by staining with Alexa Fluor 488-conjugated streptavidin (BD Biosciences) at 4 °C for 30 min. Then cells were analyzed with FACSCanto (Becton Dickinson) using FACSDiva (Becton Dickinson) and FlowJo (Tree Star) software.

SPR Study. The binding of sEGFR to EGF was analyzed by using a BIAcore T200 (GE Healthcare). All experiments were performed at 25 °C in degassed 25 mM Hepes buffer (pH 7.4) containing 150 mM NaCl, 3 mM EDTA, and 0.005% Tween 20. EGF was immobilized on a BIAcore CM5 biosensor chip via amine coupling. Briefly, the BIAcore CM5 chip was activated with N-hydroxysuccinimide and 1-ethyl-3(3-diethylaminopropyl)-carbodiimide hydrochloride. Then EGF (200 μg/mL; Millipore) in 10 mM sodium acetate (pH 4) was flowed over the activated surface at 5 μL/min for 10 min. Free ligands were washed away, and the remaining reactive functional groups were blocked with 1 M ethanolamine-HCl. Immobilized EGF contributed a signal of 327 response units (RU). A series of concentrations of purified sEGFR, as the analyte, were flowed over the EGF-immobilized surface (and over the control surface without ligand) at 10 μL/min for 10 min using a multicycle method. The RU signal corresponding to the height of the plateau was corrected by the RU signal generated with the control surface. The kinetic parameters, including the K_d, K_{on}, and K_{off}, were calculated by BIAcore T200 evaluation software. Each experiment was performed at least three times, and SD values were calculated and shown.

Tip-Based pH/Acid Controlled IMAC. EGFR tryptic peptides were desalted by self-made reversed-phase StageTips with poly(styrene-divinylbenzene) copolymer (SDB-XC) (2) as the reversed-phase sorbent. Phosphopeptides were purified through tip-based IMAC procedures constructed in house as previously described (3). All purification steps for buffer exchange and sample loading involved manipulation via centrifugation. Briefly, Ni^{2+} ions were removed with 50 mM EDTA in 1 M NaCl. Then the tip was activated with 100 mM $FeCl_3$ and equilibrated with loading buffer [6% (vol/vol) acetic acid at pH 3.0] before sample loading. Tryptic peptides were reconstituted in loading buffer and loaded onto the IMAC tip. After successive washes with 6% (vol/vol) acetic acid, 25% (vol/vol) acetonitrile, and 6% (vol/vol) acetic acid, the bound peptides were eluted with 200 mM $NH_4H_2PO_4$. The eluted phosphopeptides and unbound peptides in flow-through were desalted using reversed-phase StageTips.

LC-MS/MS Analysis. The TripleTOF 5600 system (AB SCIEX) was equipped with a nanoACQUITY UPLC (Waters). Three micrometers of ReproSil-Pur C18-AQ particles (Dr. Maisch) were packed into a 15-cm self-pulled column with a 100-μm inner diameter and 7-μm opening to prepare an analytical column using "stone-arch" frit (4). The LC system consisted of water with 0.1% formic acid (buffer A) and acetonitrile with 0.1% formic acid (buffer B). Peptides were separated through a gradient of up to 80% (vol/vol) buffer B over 120 min at the flow rate of 500 nL/min. Data were acquired using an ion spray voltage of 2.5 kV, curtain gas at 20 psi, nebulizer gas at 15 psi, and an interface heater temperature of 150 °C. For information-dependent acquisition, the MS survey scan range was m/z 300–1,500, and data were acquired for 250 ms. The top 10 precursor ions were selected based on exceeding a threshold of 100 cps in each MS survey scan, and 10 MS/MS scans were performed for 200 ms each. The collision energy was adjusted automatically by the rolling CID function of Analyst TF 1.5. To minimize repeated scans, dynamic exclusion was set at 6 s, and the precursor then was removed from the exclusion list.

Data Processing and Protein Identification. The raw MS/MS data were processed using the AB_SCIEX MS Data Converter and analyzed using Mascot (Matrix Science; version 2.3) against the UniProt database (version 57.8, *Homo sapiens*, 20,329 sequences) with the following constraints: allowing for tryptic peptides with up to two missed cleavage sites, a fragment ion mass tolerance of 0.1 Da, and a parent ion tolerance of 20 ppm. For unlabeled phosphopeptides, phosphorylation (S, T, Y) and oxidation (M) were

selected as variable modifications. The identification false-discovery rate was evaluated by search against a randomized decoy database created by Mascot at the peptide spectrum match (PSM) level. In addition, only PSMs with P value less than 0.05 were accepted.

Quantitative Analysis by IDEAL-Q. Quantitative analysis of phosphopeptide was performed by IDEAL-Q software as previously described (5, 6). Briefly, the raw data files were converted into files of mzML format by using the AB_SCIEX MS Data Converter. IDEAL-Q performs quantitation analysis using spectral data in mzXML or mzML format and Mascot search results in XML format. IDEAL-Q was used to process sequentially all peptides in each LCMS/MS run, both identified and unidentified, to quantify as many peptides as possible. IDEAL-Q first predicted the retention time of identified peptides in its current run and then determined the peak cluster based on the predicted retention time. Therefore, the unidentified peptides could be detected and aligned according to these peak clusters with a similar peptide m/z (<20 ppm) and elution time (<3 min). The identified and assigned unidentified peptide peaks were validated further by signal-to-noise ratio, charge state, and isotopic distribution criteria as follows: (*i*) signal-to-noise ratio >3, (*ii*) correct charge state, and (*iii*) correct isotope pattern to ensure accuracy. To quantify the peptide abundance of these identified and assigned peptides, extracted ion chromatography areas of each peptide were calculated. Fold change of each peptide was determined further between different samples.

SWATH-MS Measurement. SWATH-MS measurements were acquired by the TripleTOF 5600 system. The instrument was specifically tuned to allow a quadrupole resolution of 25 Da per mass selection. Using an isolation width of 26 Da (25 Da of optimal ion transmission efficiency + 1 Da for the window overlap), a set of 26 overlapping windows was constructed covering the mass range from 350 to 1,000 Da. The collision energy for each window was determined based on the appropriate collision energy for a 2+ ion centered upon the window with a spread of ± 15 eV. An accumulation time of 100 ms was used for each fragment ion scan, and total cycle time was about 2.7 s (2.7 s total for stepping through the 26 isolation windows + 0.05 s for the optional survey scan). The targeting quantitation analysis of SWATH was implemented by Skyline (7). The isolation scheme in Skyline for MS1 and MS/MS filtering was set up as "TOF mass analyzer." The resolution setting for MS1 and MS/MS was 20,000 and 10,000, respectively. For retention filtering, only scans within 5 min of MS/MS IDs were used. To quantify the targeting peptide abundance, the three most abundant transition ions were used for quantification.

Western Blot. Cells were lysed by lysis buffer [1% (vol/vol) Nonidet P-40, 10% (vol/vol) glycerol, 150 mM NaCl, 100 mM sodium phosphate (pH 7.2), 1× EDTA-free protease inhibitor cocktail from Roche] and then were incubated on ice for 15 min. After incubation, the samples were centrifuged at 13,000 × g for 15 min. The supernatants were collected, and protein concentrations were determined by bicinchoninic acid protein assay kit (Thermo Scientific). The samples were separated on 4–12% NuPAGE (Life Technologies) and transferred to PVDF membranes (Millipore). The membranes were blocked with blocking buffer [5% (wt/vol) BSA in TBS] for 1 h and then were incubated with anti-EGFR or EGFR phosphosite-specific antibodies (1:1,000 diluted in blocking buffer). After washing with 0.05% Tween 20 in TBS, the membranes were incubated with peroxidase-conjugated second antibodies, and the signal was developed by adding chemiluminescent substrates.

Proliferation Assay. Cell proliferation was measured with the INCUCYTE kinetics imaging system (Essen BioScience). Growth tests were conducted with 1×10^3 cells per well in 200 μL of culture medium containing the indicated concentrations of gefitinib for 4 d. For the groups pretreated with STI, 50 mM STI was used to treat the cells for 3 d and maintained throughout the additional 4-d gefitinib treatment. The degree of cell confluence was calculated based on the images taken (three distinct fields for each well) every 4 h. Each condition was triplicated in the experiments.

Statistical Analysis. Values are expressed as mean ± SD of at least three experiments. Paired t tests were used to analyze the statistical significance of the differences, with a value of $P < 0.05$ considered statistically significant.

1. Liu YC, et al. (2011) Sialylation and fucosylation of epidermal growth factor receptor suppress its dimerization and activation in lung cancer cells. *Proc Natl Acad Sci USA* 108(28):11332–11337.
2. Rappsilber J, Mann M, Ishihama Y (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* 2(8):1896–1906.
3. Tsai CF, et al. (2014) Sequential phosphoproteomic enrichment through complementary metal-directed immobilized metal ion affinity chromatography. *Anal Chem* 86(1):685–693.
4. Ishihama Y, Rappsilber J, Andersen JS, Mann M (2002) Microcolumns with self-assembled particle frits for proteomics. *J Chromatogr A* 979(1-2):233–239.
5. Tsou CC, et al. (2010) IDEAL-Q, an automated tool for label-free quantitation analysis using an efficient peptide alignment approach and spectral data validation. *Mol Cell Proteomics* 9(1):131–144.
6. Wang YT, et al. (2010) An informatics-assisted label-free quantitation strategy that depicts phosphoproteomic profiles in lung cancer cell invasion. *J Proteome Res* 9(11): 5582–5597.
7. MacLean B, et al. (2010) Skyline: An open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 26(7):966–968.

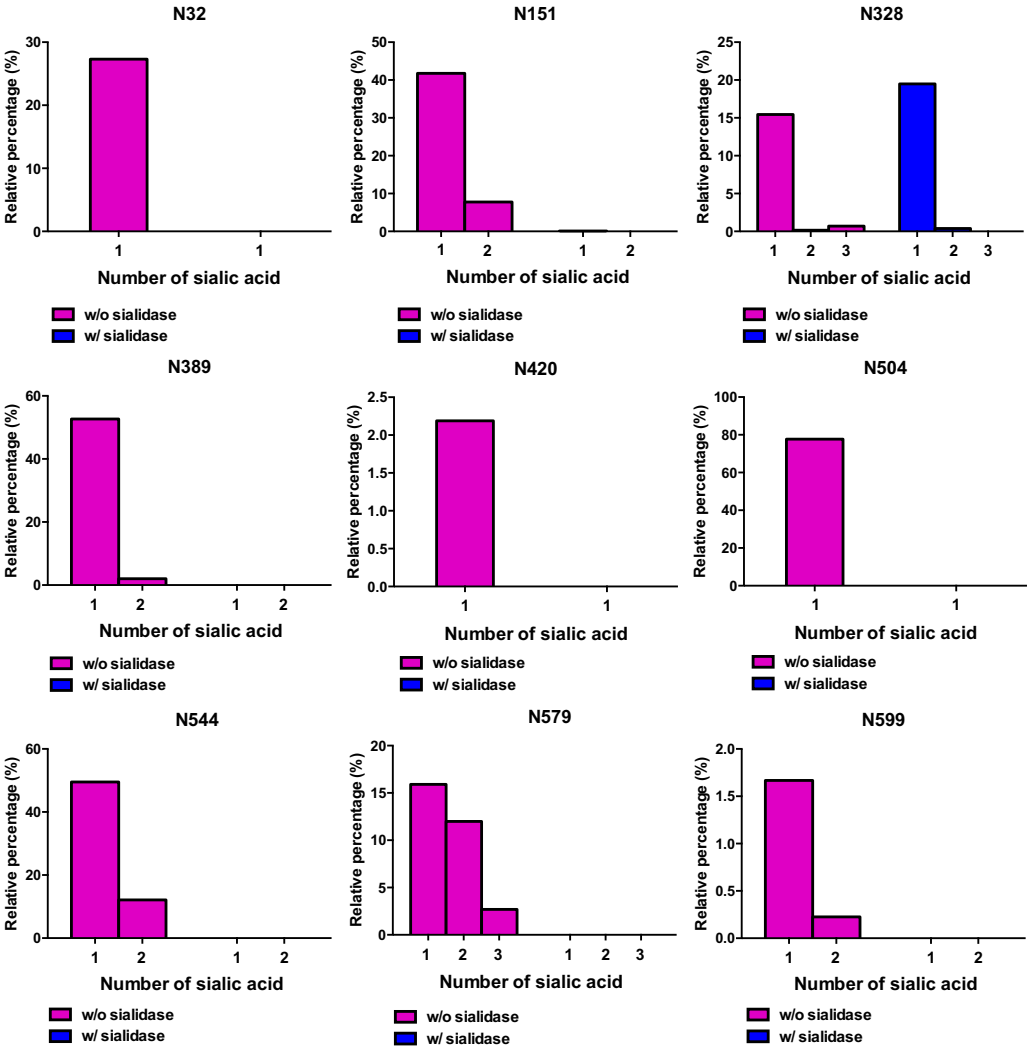

Fig. S1. Sialylation analysis of sEGFR. The glycoforms of sEGFR, with or without sialidase treatment, were analyzed by LC-MS/MS. The peak areas of identified glycopeptides with various numbers of sialic acid were quantified, and the relative percentage of each sialic acid-containing glycopeptide was calculated against the sum of peak area of all signals with the same peptide backbone (with or without glycosylation).

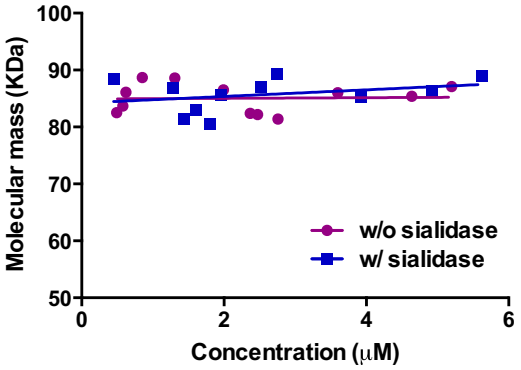

Fig. S2. Determination of the molecular mass of sEGFR. The molecular mass of sEGFR and desialylated sEGFR at various concentrations was measured by MALLS.

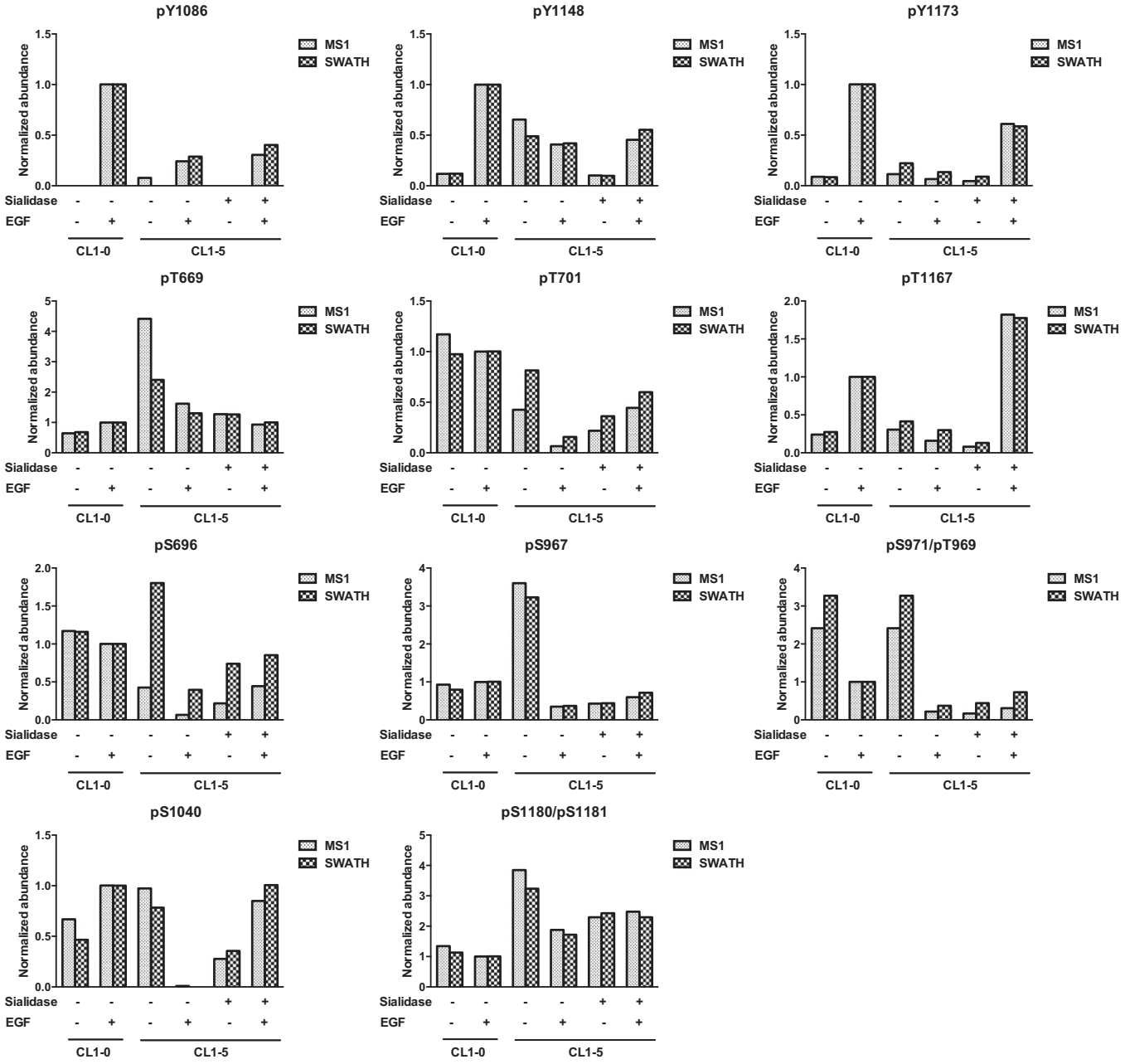

Fig. S3. Quantification of phosphopeptides by sequential window acquisition of all theoretical spectra mass spectrometry (SWATH-MS).

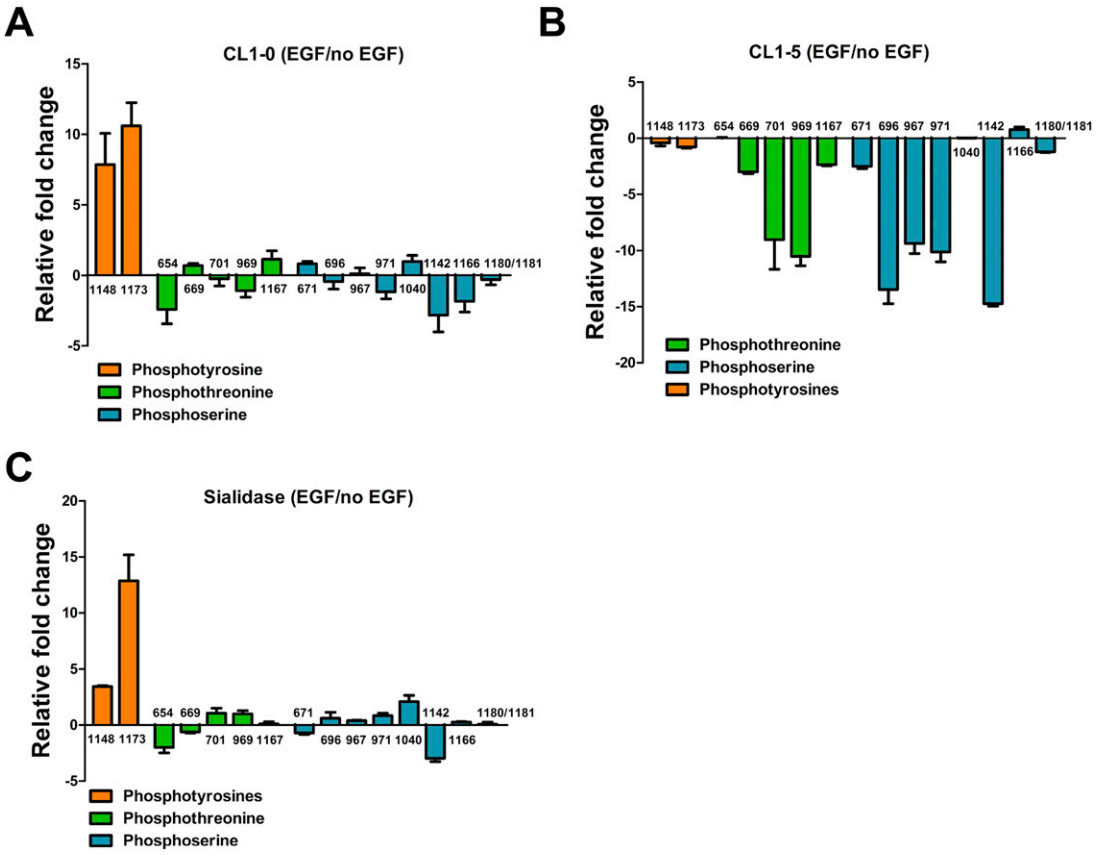

Fig. S4. Responsiveness of EGFR phosphorylation to EGF in lung cancer cells. The phosphorylation of immunoprecipitated EGFR from lung cancer cells CL1-0 (*A*), CL1-5 (*B*), and sialidase-treated CL1-5 (*C*) was identified by MS and quantified with the label-free method. The relative fold change of each phosphosites is shown. Error bars represent SD values.

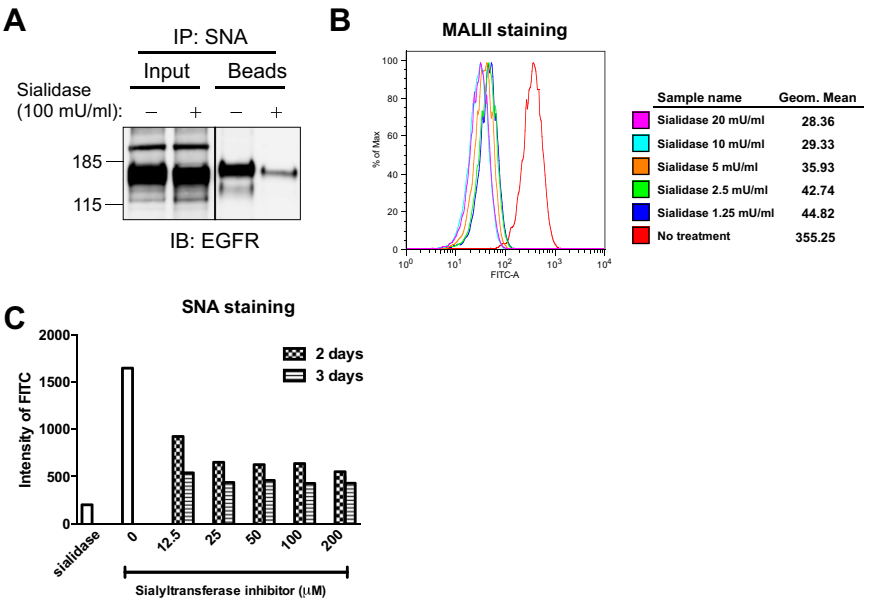

Fig. S5. Detection of sialylation in desialylated samples. (*A*) Lectin pull-down experiment with SNA for detection of sialylation on flEGFR in cell lysates. (*B* and *C*) Lectin staining followed by flow cytometric analysis for detecting the sialylation of CL1-5 (MALII staining, *B*) and H1975 (SNA staining, *C*) cell lines after desialylation. The levels of sialylation are shown as geometric means of fluorescence intensity.

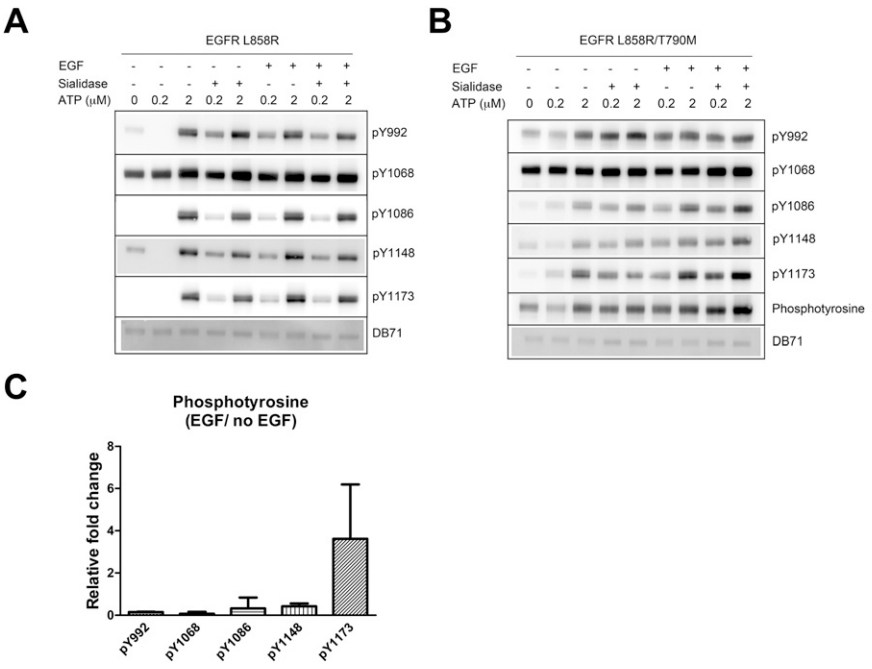

Fig. S6. In vitro phosphorylation profile of EGFR mutants. (*A* and *B*) Purified flEGFR and desialylated flEGFR with the L858R (*A*) or L858R/T790M (*B*) mutation were incubated with or without EGF at 0.02 or 0.2 μM of ATP. The level of phosphorylation was analyzed by site-specific anti-EGFR phosphotyrosine antibodies. (*C*) The relative fold change of phosphorylation on five tyrosine residues upon EGF stimulation. The signal intensity of tyrosine phosphorylation in 0.2 μM ATP was quantified. Error bars represent SD values.

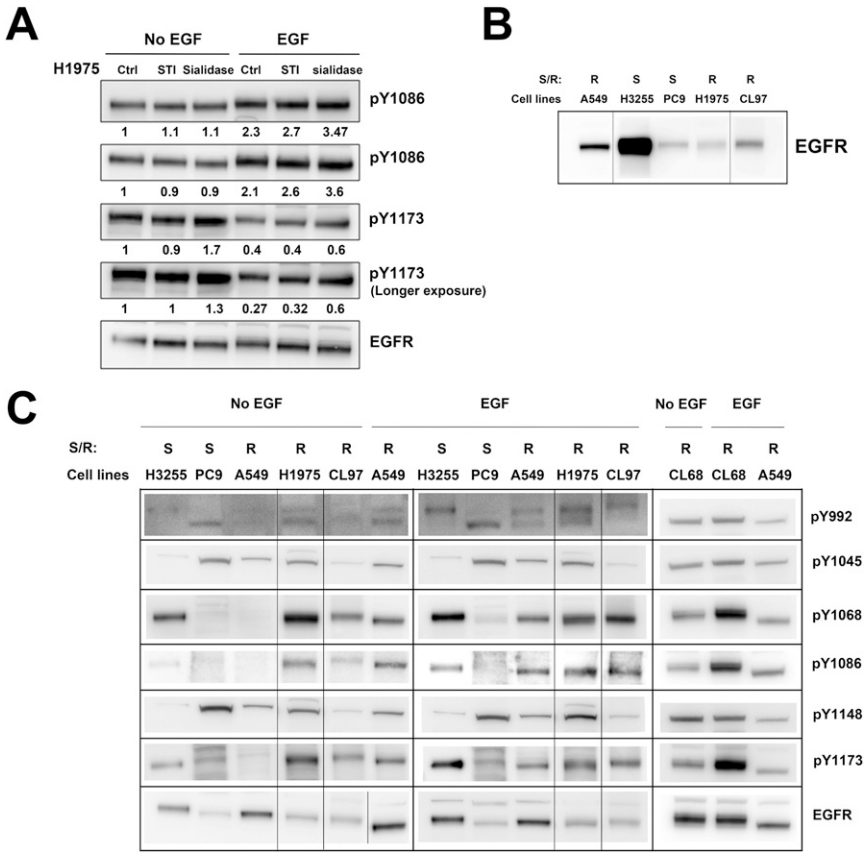

Fig. S7. Profile of EGFR tyrosine phosphorylation in lung cancer cell lines with EGFR mutations. (*A*) Tyrosine phosphorylation of H1975 cells treated with STI or sialidase. (*B*) The expression levels of EGFR in different lung cancer cell lines. (*C*) Phosphorylation profiles of EGFR from seven lung cancer cell lines (H3255, PC9, A549, H1975, CL97, CL68) with or without EGF treatment. The intensity of A549-derived EGFR in immunoblots was used to normalize between different blots. Site-specific EGFR phosphorylation was analyzed by specific antibodies. S, TKI sensitive; R, TKI resistant.

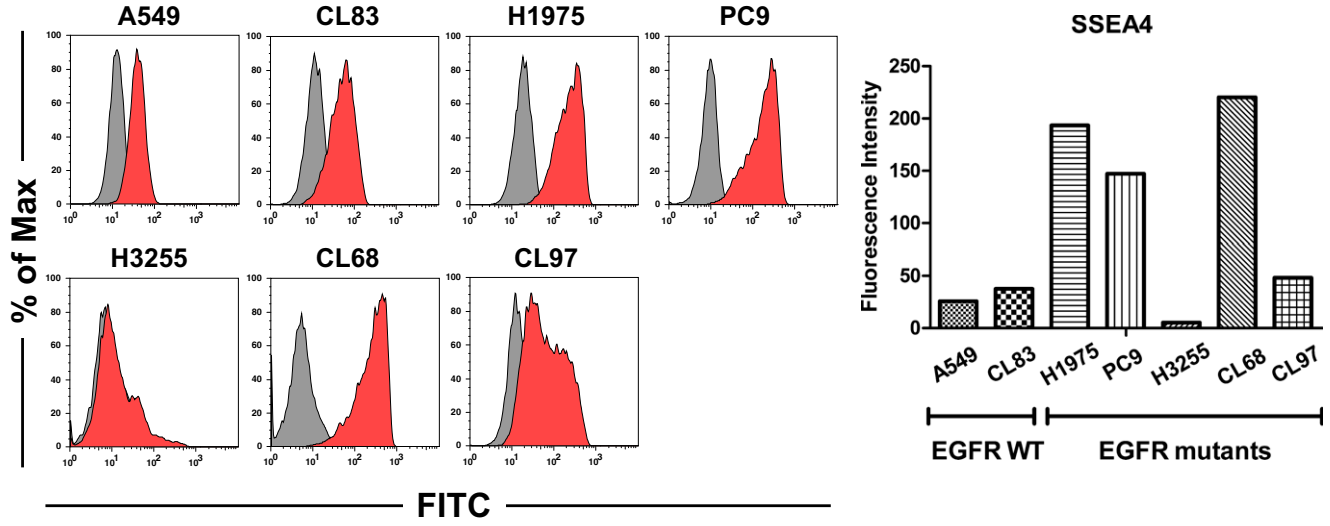

Fig. S8. Expression of SSEA4 on the surface of lung cancer cells. Cell-surface expression of SSEA4 was analyzed by flow cytometry. Lung cancer cell lines with EGFR wild-type (CL83 and A549) or EGFR mutants (H1975, PC9, H3255, CL68, and CL97) were stained with FITC-conjugated anti-SSEA4 mAb. The histograms of the cells stained with anti-SSEA4 mAb and isotype control are shown in red and gray, respectively. In the right panel, the expression level of SSEA4 is presented as geometric mean of fluorescence intensity.

Table S1. EGFR phosphopeptides identified by LC-MS/MS analysis

Peptide sequence	Phosphosites	Miss cleavage	Mascot score	z	Observed *m/z*
1. pTLRRLLQER	T654	0	30.37	3	422.233
2. ELVEPLpTPSGEAPNQALLR	T669	0	120.72	2	1,057.539
3. ELVEPLpTPSGEAPNQALLR	T669	0	98.43	3	705.365
4. ELVEPLpTPSGEAPNQALLR	T669	0	80.12	4	529.269
5. ELVEPLTPpSGEAPNQALLR	S671	0	94.17	2	1,057.529
6. ELVEPLTPpSGEAPNQALLR	S671	0	79.83	3	705.358
7. VLGSGAFGTVYK	T701	0	108.54	2	639.809
8. VLGSGAFGTVYK	S696	0	86.44	2	639.813
9. MHLPSPpTDSNFYR	T969	0	65.14	3	548.901
10. MHLPpSPTDSNFYR	S967	0	80.02	2	822.845
11. MHLPpSPTDSNFYR	S967	0	85.13	3	548.902
12. MHLPSPTDpSNFYR	S971	0	66.18	3	548.9
13. EDSFLQR	S1040	0	54.52	2	487.701
14. RPAGSVQNPVYHNQPLNPAPSR	Y1086	0	84.23	3	827.069
15. RPAGSVQNPVYHNQPLNPAPSR	Y1086	0	95.3	4	620.555
16. GSHQISLDNPDpYQQDFFPK	Y1148	0	110.43	2	1,158.524
17. GSHQISLDNPDpYQQDFFPK	Y1148	0	98.42	3	772.671
18. GSHQISLDNPDpYQQDFFPK	Y1148	0	73.83	4	579.755
19. GSHQIpSLDNPDYQQDFFPK	S1142	0	113.09	2	1,158.495
20. GSHQIpSLDNPDYQQDFFPK	S1142	0	99.82	3	772.681
21. GSHQIpSLDNPDYQQDFFPK	S1142	0	71.77	4	579.754
22. EAKPNGIFKGSTAENAEpYLR	Y1173	2	87.56	3	759.039
23. EAKPNGIFKGSTAENAEpYLR	Y1173	2	45.46	4	569.526
24. EAKPNGIFKGSpTAENAEYLR	T1167	2	68.21	3	759.038
25. EAKPNGIFKGSpTAENAEYLR	T1167	2	36.83	4	569.524
26. EAKPNGIFKGpSTAENAEYLR	S1166	2	97.95	3	759.031
27. EAKPNGIFKGpSTAENAEYLR	S1166	2	50.67	4	569.526
28. GSTAENAEYLR	Y1173	0	80.09	2	645.769
29. GSTAENAEYLR	T1167	0	77.24	2	645.771
30. VAPQSSEFIGA	S1180/S1181	0	33.55	2	593.262