Absolute Quantitation of Isoforms of Post-translationally Modified Proteins in Transgenic Organism*

Yaojun Li, Yiwei Shu, Changchao Peng, Lin Zhu, Guangyu Guo, and Ning Li‡

Post-translational modification isoforms of a protein are known to play versatile biological functions in diverse cellular processes. To measure the molar amount of each post-translational modification isoform \( (P_{\text{isf}}) \) of a target protein present in the total protein extract using mass spectrometry, a quantitative proteomic protocol, absolute quantitation of isoforms of post-translationally modified proteins (AQUIP), was developed. A recombinant ERF110 gene over-expression transgenic Arabidopsis plant was used as the model organism for demonstration of the proof of concept. Both Ser-62-independent \(^{14}\text{N}\)-coded synthetic peptide standards and \(^{15}\text{N}\)-coded ERF110 protein standard isolated from the heavy nitrogen-labeled transgenic plants were employed simultaneously to determine the concentration of all isoforms \( (T_{\text{rel}}) \) of ERF110 in the whole plant cell lysate, whereas a pair of Ser-62-dependent synthetic peptide standards were used to quantify the Ser-62 phosphorysite occupancy \( (R_{\text{aqv}}) \). The \( P_{\text{isf}} \) was finally determined by integrating the two empirically measured variables using the following equation: \( P_{\text{isf}} = T_{\text{rel}} \cdot R_{\text{aqv}} \). The absolute amount of Ser-62-phosphorylated isoform of ERF110 determined using AQUIP was substantiated with a stable isotope labeling in Arabidopsis-based relative and accurate quantitative proteomic approach. The biological role of the Ser-62-phosphorylated isoform was demonstrated in transgenic plants. Molecular & Cellular Proteomics 11: 10.1074/mcp.M111.016568, 272–285, 2012.

Protein post-translational modification (PTM)\(^1\) plays versatile regulatory roles in numerous cellular processes ranging from cell cycle control, to signaling, to differentiation (1, 2).

Hundreds of post-translational modifications of proteins are currently known to occur physiologically (3); among them, phosphorylation of protein is one of the most important. It has been estimated that a third of all proteins are phosphorylated at any given stages of a eukaryotic cell (4). Different phosphorylation isoforms of a protein may play unique functional roles in cellular processes (5–7). Therefore, quantitative measurement of the site-specific protein phosphorylation in a temporal and spatial fashion upon a specific induction either by an environmental or developmental factor is able to provide useful information on cascades of phosphorylation events and to map out phosphor relay networks during various cellular processes (8–14).

Mass spectrometry, on the other hand, has emerged as a powerful and indispensable tool for the quantitative analysis of changes in these protein PTMs (13, 15–18). However, difficulties lie in the exhaustive and complete quantitation of a eukaryotic plant cell proteome because of the intrinsic limitation of mass spectrometry measurement. Thus, enrichment of phosphorylated protein has been a common practice to compensate the inadequate sensitivity of mass spectrometry and to identify and quantify the functional phosphorylation events beyond the detection limits of mass spectrometers.

The most frequently used approach to enrich the in vivo expressed proteins of research interest is to fuse a polypeptide tag to a target protein in a transgenic eukaryotic cell, followed by tandem affinity purification (TAP) of the recombinant protein under both native (19, 20) and fully denaturing condition (21). Especially, the expression of recombinant proteins in transgenic cell lines and organisms is a common practice in addressing the molecular mechanisms underlying regulation of gene expression, cell signaling, as well as protein structure and function relationship (22–27). Thus, versatile PTMs of target proteins may be investigated through the affinity-purified recombinant proteins from transgenic organisms. This in vivo overexpression of fusion proteins is particularly attractive when it has been found to be of the same post-translational modification profile as that of endogenous proteins (28–31).

The quantitative proteomic methods developed thus far for measuring phosphorylation can be classified into two subgroups in general: (i) the isotope labeling approach, such as ICAT (32), isobaric tags for relative and absolute quantitation...
(33), stable isotope labeling with amino acids in cell culture (34), stable isotope labeling in planta (35), stable isotope labeling in Arabidopsis (SILIA) (36), stable isotope labeling in mammals (13), and enzymatic labeling with $^{18}$O water (8) and (ii) the isotope labeling-free approach (10, 12, 37). These novel and relative measurement techniques have been successfully applied to explore the dynamic changes at the cellular phosphoproteome (10, 11, 13, 14) and to provide relative abundances in phosphopeptides among a set of treated and untreated samples (8, 9, 12, 38). In particular, when the phosphorylation stoichiometry of a phosphosite is under consideration, isotope-free quantitation of the has become a way to monitor the ratio of ion counts of a phosphopeptide and its unmodified cognate (10, 14). However, to absolutely quantify a change of phosphorylation level of a protein, the isotope-labeled peptide standards (AQUA peptides) approach is often the method of choice, in which the stable isotope-labeled synthetic phosphopeptide standards are spiked into the protease-digested peptide samples in known quantities to measure the absolute levels of post-translational modifications (7, 9, 39). In addition, an artificial concatamer of stable isotope-labeled standard peptides (QconCAT) can be spiked into the targeted protein samples before the proteolysis step to measure the targeted protein quantities (40–43). Although AQUA and QconCAT approaches are very easily multiplexed and straightforward to perform for quantitative analysis of protein phosphorylation, both of these two quantitative approaches become limited either if a multiple or partial isolation of targeted protein phosphorylation, both of these two quantitative approaches become limited either if a multiple or partial isolation of proteins of interest is needed or in cases where protease-mediated peptide digestion is incomplete (44).

To further improve the absolute protein quantitation methodology, a protein standard absolute quantification (PSAQ) strategy was developed (44, 45). According to this strategy, a full-length stable isotope-labeled protein is introduced as a quantification standard into a crude protein extract containing the targeted protein. The labeled protein standard is believed to go through an identical in vitro manipulation procedure with the targeted protein during the entire process of protein isolation and peptide production. The PSAQ has emerged as a popular approach especially in the absolute quantitation of nonphosphorylated proteins. However, in the case of absolute measurement of phosphorylated protein isoform, it is quite difficult to obtain a specific isoform of phosphoprotein standard either from an in vitro kinase reaction, because of the possible incomplete phosphorylation, or from an overexpression of phosphoprotein in a heterologous system where no specific kinase may be present to catalyze the site-specific phosphorylation.

It is therefore that the measurement of the concentration of an isoform of phosphorylated proteins is an important issue because of its significant biological implication. For example, a 4-fold increase in site-specific phosphorylation from 25 to 100% of the total protein may result in a complete different cellular and phenotypic response than that from 1 to 4% (14, 39). Because the absolute amount of the phosphorylated isoform is calculated either in moles or grams of cell and tissue, the concentration of a phosphoprotein isoform produced under a number of treatments can be compared directly among various protein extracts (46). The comprehensive statistical analysis of the absolute amount of a protein isoform (e.g. a PTM isoform of a tissue- and stage-specific biomarker) can also provide correlations between levels of specific biomarkers and severities of symptoms of those individuals afflicted by diseases (47, 48).

To absolutely quantify a phosphoprotein isoform from the total cellular proteins, an alternative method, absolute quantitation of isoforms of post-translationally modified recombinant proteins (AQUIP), is hereby proposed. This AQUIP strategy is as follows: (i) two $^{14}$N-coded (light nitrogen) synthetic peptides are made according to PTM site-independent regions of a target protein and then used as the AQUA peptide standards to measure the absolute concentrations of the targeted peptides deriving from a corresponding stable isotope-labeled protein; at the same time, these two PTM site-independent peptide standards are employed to evaluate not only the peptide yield but also the absolute amount of trypsin-digested and $^{15}$N-coded peptide samples of a target protein; (ii) a tandem affinity purification tag, namely His$_{10}$-Bio-His$_{10}$ (HBH) tag, is employed that allows the stable isotope-labeled targeted phosphoproteins to be purified and enriched under a fully denaturing condition, leading to the preservation of the phosphorylation status of a recombinant phosphoprotein attained in a transgenic organism throughout the protein purification; (iii) both $^{14}$N-coded phosphopeptide standard and its unmodified cognate are synthesized and used to absolutely quantify the $^{15}$N-coded phosphorylated and nonphosphorylated peptides derived directly from trypsin-digested and $^{15}$N-coded recombinant proteins that are produced via SILIA(36); and (iv) the absolute amount of a site-specific phosphorylation isoform of a protein ($P_{\text{ist}}$) is determined based on the following equations,

$$T_{\text{ist}} = NP_{\text{ist}} + P_{\text{ist}}$$  \hspace{1cm} (Eq. 1)

$$R_{\text{ist}} = P_{\text{ist}}/(P_{\text{ist}} + NP_{\text{ist}})$$  \hspace{1cm} (Eq. 2)

$$P_{\text{ist}} = T_{\text{ist}} \cdot R_{\text{ist}}$$  \hspace{1cm} (Eq. 3)

where $T_{\text{ist}}$ is the molar amount of all isoforms of a protein, including both the site-specifically phosphorylated and its nonphosphorylated cognates. $R_{\text{ist}}$ is defined as the phosphorylation occupancy at a specific phosphosite of a protein and is a ratio of site-specifically phosphorylated isoform over all isoforms of the protein. $P_{\text{ist}}$ is the molar amount of a site-specifically phosphorylated isoform, whereas $NP_{\text{ist}}$ stands for all of its nonphosphorylated cognates. To attain a proof-of-concept for AQUIP in a model organism, such as Arabidopsis, a transcription factor named ERF110 (ethylene response factor 110) is therefore chosen as the targeted protein to be measured for its absolute phos-
phylation level at Ser-62 position. The ethylene response factors (ERF) belong to a subfamily of ethylene-responsive element-binding transcription factors, or AP2/EREBP (49, 50). These transcription factors are well known to participate in regulation of diverse plant growth and development processes, as well as plant adaptation to environment. Ethylene regulates these transcription factors via the ethylene receptors (such as ETR1), mitogen-activated protein kinase cascades (such as CTR1, MKK9, and MAPK3/6) and a putative metal ion transporter (EIN2) and finally down to ethylene related transcription (5, 12). Quantitative proteomic study of PTMs of this class of important plant transcription factors will contribute to the understanding of the PTM networks governing cellular processes in an eukaryotic model organism.

**EXPERIMENTAL PROCEDURES**

**Plant and Materials**—Murashige and Skoog basal salt mixture was purchased from Amersham Biosciences Bioscience. Sucrose and HPLC grade methanol, formic acid, and ACN were from Thermo Fisher Scientific. Ni²⁺-nitrilotriacetic acid-agarose beads were purchased from Qiagen. Dynabead M-280 streptavidin was from Invitrogen. ZipTIP-C18 was purchased from Millipore (Billerica, MA). Potassium nitrate coded with ¹⁵N stable isotope (99% purity) was purchased from Cambridge Isotopes Laboratories (Andover, MA). Trypsin modified by L-1-tosylamido-2-phenylethyl chloromethyl ketone, aminooxy-acetic acids, phosphatase inhibitor mixture 2, and other standard-grade chemicals were obtained from Sigma-Aldrich. The seeds of wild type Arabidopsis thaliana, Col-0 ecotype, and the ethylene-insensitive mutant etr1-1 were bought from the Arabidopsis Biological Resource Center (Columbus, OH). The ethylene-insensitive mutant ein2-5 Arabidopsis was a gift from Dr. Joseph Ecker.

**Construction of ERF110-HBH Fusion Gene and Binary Vector**—An intron-free 5'-UTR of ribulose-1,5-bisphosphate carboxylase oxygenase small subunit 1A (TAIR accession number AT1G67090) was constructed by self-annealing of two single-stranded oligonucleotides, Ru-F (5'-GAT CCT CAG TCA CAC AAA GAG TAA AGA AGA ACA ATG G-3') and Ru-R (5'-TGG ACC ATT GTT CCT TAC TCT TTG TGT GAC TGA G-3'). His₅-N-Biotin carboxyl carrier protein (TAIR accession number AT5G16390)-His₅ (HBH) peptide tag was generated with one-step PCR using primers, HBH-F (5'-AAA GTC GAC GGC GGC CCT CAT CAT CAT CAC CAT CCA GAC AAA TCG TCA-3') and HBH-R (5'-AAA GAG CTC CTA CTT AAT TAA GAC ACC ATG ATG ATG GTG GTG ATG ATG CGG TTG AAC CAC AA-3'). The purified HBH PCR product was then digested with Sall and Snacl. Purification was carried out between Sall/Snacl-cut His₅-BIO-His₅, sticky end UTR oligonucleotide, and BamHII/Sacl-cut pBl121 vector at a ratio of 15:15:1, respectively. Both 0.9-kb ERF110 (TAIR accession number At5g50080; GenBank™ accession number JN819205) genomic DNA fragment and the ERF genomic DNA coding region with point mutations were amplified by PCR using the following primers: gERF-F, 5'-CAT AGT CGA CTC CAT GAT CCT GTC GGC C-3' (Sall, underlined); gERF-R, 5'-TCA TGG CCC GCC TCT GTT AGG TAG AGA AGG-3' (Ascl, underlined); gERFS-A-F, 5'-CTG GTA GAC GCC TCA CAT CAA-3' and gERFS-R-A, 5'-GAT TAT GTG AAC GGT CTA CAC G-3'. PCR products were digested with Ascl and Sall and inserted into a modified binary vector pBl121 derivative between a double CaMV 35S promoter and HBH tag. Both of these constructs, genomic DNA sequence encoding ERF110 and Ser-62 to Ala-62 (A) mutant genomic ERF110 gene, were transformed into ein2-5 Arabidopsis via a floral dip protocol (51). Transgenic plants were selected from Murashige and Skoog medium supplemented with 25 mg liter⁻¹ hygromycin (Invitrogen) on Petri dishes.

**Growth Conditions and Hormone Treatment**—The second generation (T2) seeds of ERF110-ox/ein2-5 Arabidopsis were sown in rows on agar-based SILIA medium containers (either with K¹⁵NO₃ or K¹⁵NO₃) and grown in the plant growth chambers as described previously (36). To reduce the endogenous ethylene effects, 150 µM aminoxy-acetic acid was added into the medium. Two batches of 3-week-old plants were placed separately in two independent air flow chambers for 5 h to reduce endogenous ethylene level. One of them was treated with 10 ppm ethylene for 12 h, whereas the other was retained in air for the same period of time. Both groups of plants were harvested, frozen by liquid nitrogen, and stored in −140 °C freezer.

**Tandem Affinity Purification of ¹⁵N-coded Recombinant ERF110 Protein**—Heavy nitrogen-labeled total cellular proteins were produced via the previously published SILIA protocol (36). Both ¹⁴N- and ¹⁵N-coded proteins were extracted using a urea extraction buffer and by a previously described protocol (36). After several steps of high speed centrifugation, precipitation, and rinse, protein pellet was air-dried and resuspended in a buffer A containing 8 mM urea, 300 mM NaCl, 0.5% Nonidet P-40, 50 mM sodium phosphate, pH 8.0, 50 mM Tris, pH 8.0, 10 mM NaF, 1 mM glycero-2-phosphate disodium, and 1× phosphatase inhibitor mixture 2. Then protein sample was incubated with Ni²⁺-nitrilotriacetic acid beads (500 µl of slurry, pre-equilibrated with buffer A) at room temperature for 1 h, washed three times with 3 ml of buffer A followed by 3 ml of loading buffer A containing 20 mM imidazole for another three times. The proteins were eluted three times with 1 ml of buffer B (8 mM urea, 200 mM NaCl, 0.2% SDS, 10 mM sodium phosphate, pH 8.0, 100 mM Tris, pH 8.0, 250 mM imidazole). The elutes were then loaded onto immobilized streptavidin magnetic beads (200 µl of slurry pre-equilibrated by buffer C containing 8 mM urea, 200 mM NaCl, 0.2% SDS, 100 mM Tris, pH 8.0) and incubated at room temperature for at least 6 h or overnight. Streptavidin beads were washed three times with 1 ml of buffer C, three times with 1 ml of buffer C plus 2% SDS, and three times with 1 ml of 1× phosphate-buffered saline, respectively. Biotin-labeled protein was eluted using 1× SDS loading buffer containing 30 mM d-biotin at 96 °C for 15 min. The supernatant was ice-chilled quickly and collected for SDS-PAGE fractionation.

To perform immunoblotting analysis, the proteins resolved on SDS-PAGE gel were transferred to a Hybond-P PVDF membrane (GE Healthcare) and analyzed by immunoblotting. The anti-His₅ monoclonal antibody (Clontech) was diluted at a ratio of 1:5000.

**Preparation of Heavy Nitrogen (¹⁵N)-coded Peptides for Absolute Quantitation**—To determine the absolute amount of recombinant ERF110 protein in the whole plant cell lysate, a total of 0.7 mg of ¹⁴N-coded total cellular proteins were prepared from a whole cell lysate of ERF110-ox/ein2-5 transgenic Arabidopsis labeled via SILIA (36) and subjected to a 6 × 9-cm SDS-PAGE gel. After the protein gel was stained with Coomassie Brilliant Blue, the ¹⁵N-coded recombinant ERF110 protein-containing gel slice (~40 kDa) was excised out. The in-gel tryptic digestion and peptide extraction were performed as previously described (52). The ¹⁵N-coded peptide sample was then divided into seven parts in an equal amount, and each part was mixed with the two ¹⁴N-coded peptide standards (peptide P₁, and P₂), which were chemically synthesized for the absolute quantitative analysis of an isoform of a target protein. A series of ¹⁴N-coded peptide standards (from 100 to 1000 fmol) were spiked into aliquots of ¹⁵N-coded peptide samples, followed by ZipTipC₁₈ desalting and analysis on LC-Q-TOF-mass spectrometry.

To measure the peptide yield of the peptide Pₐ from trypsin digestion of a ¹⁵N-coded recombinant protein (Pₐ), 294 fmol (12 ng) and 882 fmol (36 ng) of ¹⁵N-coded recombinant ERF110 proteins, which were highly purified from the ¹⁴N stable isotope-labeled transgenic
Arabidopsis tissue, were resolved on SDS-PAGE gel and then subjected to protease digestion. Two 15N-coded synthetic peptide standards (P1 and P2), both of which serve as an internal control and had a molar amount (either 294 fmol or 882 fmol) equal to that of 14N-coded ERF110 protein to be measured, were spiked into two batches of 15N-coded peptide samples, respectively. After MS analysis, the two sets of MS raw data for each peptide were integrated into the above seven data points to establish a standard curve for peptide standard P1 and P2 according to a correlation between the molar amounts and the ion intensities of 14N-coded peptide standards, respectively.

Establishment of Standard Curves for Absolute Measurement of Ser-62 Phosphorylation—To determine ERF110 phosphosite occupancy calculated by the AQUA approach (R NP), a known amount (0.5 μg = 12 pmol) of recombinant ERF110 proteins isolated from both air- or ethylene-treated ERF110-ox/ein2-5 Arabidopsis were divided into six aliquots, followed by both SDS-PAGE fractionation and in-gel trypsin digestion. The 14N-coded synthetic Ser-62-phosphopeptide and its nonphosphorylated cognate were spiked into six aliquots of 15N-coded peptide samples with a concentration ranging from 5 to 160 fmol and from 62.5 to 2000 fmol, respectively. Then sodium periodate with a final concentration of 2 mM was used to chemically oxidize the methionine residues on the peptides in these peptide samples as described by Kettenbach et al. (53). After LC-Q-TOF mass spectrometry analysis, two standard curves were then built according to the relationship between ion counts of whole isotopic envelope and molar ratios of 14N-coded synthetic peptide pairs to determine molar amounts of 15N-coded Ser-62-phosphopeptide (P NP) and its nonphosphorylated cognates (NP NP).

Nano LC-Q-TOF Mass Spectrometry Analysis—LC-MS/MS was performed with a nanoflow LC (nano AcquityTM; Waters) coupled to an ESI hybrid Q-TOF Premier tandem mass spectrometer (Waters). The program MassLynx (version 4.1; Waters) was used for data acquisition and instrument control. A 180-μm × 20-mm Symmetry C18 trap column and 75-μm × 250-mm BEH130 C18 analytical column were used. The following parameters were used for accurate mass screening (54) on the Q-TOF mass spectrometer. A single survey scan (MS1 mode) from m/z 300 to 2000 was followed by up to three MS/MS scans from 50 to 2000. The top three most abundant precursors from the inclusion list (if present) were targeted for MS/MS spectrum acquisition over the course of a 120-min experiment. The information in the inclusion list, including mass to charge (m/z), charge state (q), and retention time, were imported into the mass spectrometer operating software (MassLynx version 4.1; Waters) for selection of precursors. The m/z tolerance around targeted precursors was 50 mDa. Dynamic exclusion was set with exclusion duration of 30 s. Again the m/z tolerance for dynamic exclusion was 50 mDa. The intensity threshold for triggering of a detected peak was set to 50, and the collision energy is variable during MS/MS scan according to m/z and m/z 2, and the exact values are from factory instructions.

The raw data were processed using both ProteinLynx 2.2.5 and commercially available Mascot search engine (http://www. matrixscience.com) as described previously (36). The observed isotopic envelopes of precursor ions of each peptide mentioned in this study were examined manually and performed noise elimination. An average of 14N (99.64%) and 15N (97.44%) incorporation rate was used, respectively, to make calculations, as described in our previous study using both Isopro3.1 (https://sites.google.com/site/isoproms/) and Siliamass 1.0 scripts (36).

RESULTS

Absolute Quantitation of a Recombinant Protein in Arabidopsis Cell Lysate—To develop an absolute quantitative proteomic method, AQUIP, for quantitation of a phosphorylated (one of many types of PTMs) isoform of a recombinant protein from the total cellular protein of a transgenic organism, Arabidopsis overexpressing a HBH-tagged recombinant ERF110 was utilized in this experiment. The recombinant ERF110 protein was first isolated from the whole plant cell lysate using TAP under the fully denaturing condition (see “Experimental Procedures”). Because the concentration of an in vivo site-specifically phosphorylated isoform of a protein (P NP) can be determined according to Equations S1-S3 (supplemental Fig. 1), in which the P NP is equal to that of the total amount of recombinant ERF110 protein (T NP) multiplied by Ser-62 site phosphorylation occupancy (R NP) of the recombinant ERF110 protein, the first experiment is to empirically determine the total molar amount (T NP) of all isoforms of the recombinant ERF110 present in the whole plant cell lysate. Thus, T NP can be determined according to the following equation,

\[ T_{NP} = \frac{NP_{NP} + P_{NP}}{NP_{NP}} \times \frac{1}{m} \sum \frac{1}{k_i/k_{NP}} + \ldots + \frac{1}{k_m/k_{NP}} \]

(Eq. 4)

in which T NP stands for the absolute molar amount of a standard peptide P NP derived from a trypsin digestion of 15N-coded recombinant ERF110 protein, and k NP stands for the peptide yield of the peptide P NP from a trypsin digestion of 14N-coded recombinant ERF110 protein (Equations S4–S6). Because one or two peptides are usually used as standards to absolutely quantitate protein in both QconCAT(40, 42) and AQUA strategies (9, 16), two PTM site-independent and 14N-coded synthetic standard peptides (m = 2), P 1 and P 2, were then selected and made according to primary sequence of ERF110 protein. The amino acid sequences of these two peptide standards are SPAPGEPPFIK 227 (P 1) and GWLGIDSAPIPSS-FAP 50 (P 2). The rationales for choosing them as standards were that (i) they are unique in sequence; (ii) they should be free of PTMs because the post-translational modification may complicate the subsequent mass spectrometry analysis; and (iii) the standard peptide ion counts (or called MS signal) should have a good reproducibility and a high resolution on mass spectrometry analysis.

To determine the absolute amount of recombinant ERF110 protein in the whole plant cell lysate, a known amount of 15N-coded total cellular proteins were prepared from the whole cell lysate of ERF110-ox/ein2-5 transgenic Arabidopsis labeled via SILIA (36). The crude 15N-coded ERF110 peptide mixture was generated by in-gel trypsin digestion of the in planta expressed ERF110 protein, which was resolved on a SDS-PAGE gel and located at a region near 40 kDa (Fig. 1, left panel). The crude 15N-coded peptide sample was then divided into seven aliquots, each of which was mixed with the two 14N-coded peptide standards (P 1 and P 2; supplemental Table 1). A standard curve was then established according to the correlation between the molar amount and the ion intensity of the whole isotopic envelope of each 15N-coded peptide
FIG. 1. An AQUIP strategy for measurement of the molar amount of recombinant proteins from the whole cell lysate. Transgenic Arabidopsis overexpressing a recombinant gene of interest, such as ERF110, was labeled with the stable isotope $^{15}$N via SILIA protocol (top panel). The $^{15}$N-coded crude total cellular proteins were divided into two parts. One part was first subjected to TAP, whereas the other was resolved on SDS-PAGE gel together with the TAP-purified recombinant protein (right middle panel). The highly purified recombinant ERF110 protein serves as an internal control to measure the peptide yield from tryptic digestion and enrichment of ERF110 present in the total cellular protein as suggested by PSAQ strategy (right middle panel). After determination of the mass amount of purified ERF110, 12 ng (294 fmol) and 36 ng (882 fmol) of $^{15}$N-coded recombinant ERF110 proteins were subjected to SDS-PAGE gel, followed by in-gel trypsin digestion and peptide extraction (right middle panel). To absolute quantitate (AQUA) the $^{15}$N-coded tryptic peptides derived from a target protein using the $^{14}$N-coded synthetic peptide standards, two oligopeptides, SPAPGEPPFIK$_{227}$ ($P_1$) and GWLGIDSAPIPSSFAR$_{59}$ ($P_2$) were synthesized and used to

$$T_{isf} = \left( \frac{T_1}{k_1} + \frac{T_2}{k_2} \right) / 2$$
Absolute Quantitation of Isoforms of PTM Protein

**Fig. 2. Absolute quantitation of recombinant ERF110 protein in cell lysate.**

(a) the molar amount of a standard peptide is correlated to its ion intensity: $T_1 = 0.1307T_{m,ion} - 140.64$ ($r^2 = 0.9934$) for peptide $P_1$, and $T_2 = 0.3703T_{m,ion} - 101.28$ ($r^2 = 0.9866$) for peptide $P_2$. The open red stars indicate the $T_m$ value calculated according to the ion intensity $T_{m,ion}$ from whole isotopic envelope of precursor ion of the peptide $P_m$, b, the average peptide yield ($k_m$) of peptide $P_m$ is presented in the percentage of the loading amount of the pure recombinant protein standard. Here, the $m$ value is 1 or 2. The amino acid sequences of peptide $P_1$ and $P_2$ are SPAPGEPPFK$^2_{277}$ and GWLIDSAPIPSSSFAR$^{590}$, respectively.

(b) Absolute quantitation of isoforms of recombinant ERF110 protein in the whole plant cell lysate—To absolutely quantitate Ser-62-phosphorylated ERF110 isoform, $P_{ist}$, the phosphorylation occupancy ($R_{ist}$) in Equation 2 needs to be determined. To do that, a general workflow was designed as shown in Fig. 3. First, two $^{14}$N-coded peptide standards were synthesized. One was a phosphopeptide (or a PTM peptide), and the other was its unmodified cognate. Because both peptides share the same amino acid sequence and the same trypsin digestion site, the peptide yield (supplemental Fig. 1) was therefore considered to be identical. The Ser-62 phosphorylation occupancy, $R_{ist}$, can be determined according to the following equation.

---

establish standard curves (bottom panel). The $^{15}$N-coded total cellular proteins (0.7 mg) from the whole cell lysate of ERF110-ox/ein2-5 Arabidopsis (air-treated for 12 h) were loaded into seven lanes of SDS-PAGE gel, each with an equal amount. The $^{15}$N-coded peptide sample mixture was generated from in-gel trypsin digestion of a fraction of total cellular proteins with a molecular size of ~40 kDa (left middle panel). A series of peptide standards (100, 200, 400, 500, 600, 800, and 1000 fmol) were then spiked into the seven $^{15}$N-coded peptide samples, followed by LC-MS/MS analysis (bottom panel). A standard curve was then established according to the correlation between the molar amount and the ion intensity of each $^{15}$N-coded peptide standard on MS.

---

Molecular & Cellular Proteomics 11.8
Absolute Quantitation of Isoforms of PTM Protein

Table I

| Parameter | Treatment | Peptide P<sub>a</sub> | Peptide P<sub>b</sub> | PTM site |
|-----------|-----------|----------------------|----------------------|-----------|
| T<sub>m</sub> (fmol·mg⁻¹) | Air/ethylene | 4,362.8 ± 422.9 | 6,795.8 ± 561.6 | 1.27 ± 0.07% |
| k<sub>m</sub> (fmol·mg⁻¹) | Air/ethylene | 55.7 ± 5.7% | 86.4 ± 5.46% | 0.56 ± 0.02% |
| R<sub>isf</sub> | Ethylene | 7,849.1 ± 693.8 | 7,849.1 ± 693.8 | 99.6 ± 5.3 |
| P<sub>isf</sub> (fmol·mg⁻¹) | Ethylene | 99.6 ± 5.3 |

Thus, P<sub>isf</sub> of recombinant ERF110 protein was determined to be 99.6 ± 5.3 fmol·mg⁻¹ and 43.6 ± 1.6 fmol·mg⁻¹ in the total cellular protein sample of air- and ethylene-treated ERF110-ox/ein2-5 Arabidopsis, respectively (Table I). It is clear that ethylene treatment reduced the Ser-62 phosphorylation level to 44.0 ± 3.1% of that of the air control in the transgenic plant (Fig. 4c).

Substantiation of the AQUIP Results Using the SILIA-based Relative and Accurate Measurement Approach—A SILIA-based relative and quantitative phosphoproteomic method has recently been developed to measure fold changes in a target phosphoprotein abundance in Arabidopsis (36), which essentially measures the mass ratio of a phosphoprotein produced in two different plant tissue samples. Given that the molar ratio of a pair of phosphoproteins present in a pair of differentially treated plant tissues is equivalent to the mass ratio of the same pair of phosphoproteins, the molar ratio of the Ser-62-phosphorylated isoform of recombinant ERF110 protein determined using the above AQUIP method should be substantiated with the relative and accurate quantitation phosphoproteomic protocol as illustrated in the relative quantitation workflow (supplemental Fig. 3). A standard curve was then established by correlating the ratios of the observed ion intensities of the whole isotopic envelopes (abbreviated as RI) of ¹⁴N/¹⁵N-coded recombinant ERF110 peptide precursors (all three standard peptides, P<sub>1</sub>, P<sub>2</sub>, and Ser-62 nonphosphorylated peptide NP<sub>a</sub>, were used) and the mass ratios of the

| Parameter | Treatment | Peptide P<sub>a</sub> | Peptide P<sub>b</sub> | PTM site |
|-----------|-----------|----------------------|----------------------|-----------|
| T<sub>m</sub> (fmol·mg⁻¹) | Air/ethylene | 4,362.8 ± 422.9 | 6,795.8 ± 561.6 | 1.27 ± 0.07% |
| k<sub>m</sub> (fmol·mg⁻¹) | Air/ethylene | 55.7 ± 5.7% | 86.4 ± 5.46% | 0.56 ± 0.02% |
| R<sub>isf</sub> | Ethylene | 7,849.1 ± 693.8 | 7,849.1 ± 693.8 | 99.6 ± 5.3 |
| P<sub>isf</sub> (fmol·mg⁻¹) | Ethylene | 99.6 ± 5.3 | 43.6 ± 1.6 |

The molar amount of Ser-62-phosphorylated isoform of recombinant ERF110 protein, P<sub>isf</sub>, in the whole plant cell lysate was then determined according to the following equation:

\[ P_{isf} = T_{isf} \cdot R_{isf} = T_{isf} \cdot R_{aqq} \]  

(Eq. 6)

in which R<sub>aq</sub> stands for the phosphosite occupancy calculated by AQUA approach, and P<sub>aq</sub> and NP<sub>aq</sub> stand for the molar amount of phosphopeptide P and its nonphosphorylated cognate NP, respectively, derived from ¹⁵N-coded and highly purified recombinant ERF110 protein (Equation S8). To determine the molar ratio, R<sub>aq</sub>, a known quantity (12 pmol) of highly purified and ¹⁵N-coded recombinant ERF110 protein from ERF110-ox/ein2-5 Arabidopsis was divided into six aliquots and subjected to SDS-PAGE fractionation and trypsin digestion. Among them, both ¹⁵N-coded synthetic Ser-62 phosphopeptide (P) and its nonphosphorylated cognate (NP) were spiked into six mixtures of ¹⁴N-coded peptide samples with a concentration range across from 5 to 160 fmol (P) and 62.5 to 2000 fmol (NP), respectively (Supplementary Tables 2 and 3 and Fig. 4, a and b). The mass spectrometric signals of these ¹⁴N-coded synthetic Ser-62 phosphopeptide and its nonphosphorylated cognate were used to establish two individual standard curves against their molar amounts (Fig. 4, a and b, and Equations S9–S12 in supplemental Fig. 1b). The molar amounts of P<sub>aq</sub> and NP<sub>aq</sub> were therefore determined to be 16.6 ± 0.9 and 1,290.5 ± 20.3 fmol, respectively, in the air-treated sample, whereas those of P<sub>aq</sub> and NP<sub>aq</sub> were determined to be 7.5 ± 0.3 and 1,334.5 ± 27.1 fmol, respectively, in the ethylene-treated sample (Fig. 4, a and b, and Equations S9 and S10). The Ser-62 phosphorylation occupancy R<sub>isf</sub> (or R<sub>aq</sub>, supplemental Fig. 1b and Equations S13 and S15) were therefore determined to be 1.27 ± 0.07 and 0.56 ± 0.02%, respectively, for air- and ethylene-treated ERF110-ox/ein2-5 Arabidopsis (Table I).

Given that T<sub>isf</sub> and R<sub>isf</sub> values have been determined empirically according to Equations 4 and 5, the absolute molar amount of Ser-62-phosphorylated isoform of recombinant ERF110 protein, P<sub>isf</sub>, in the whole plant cell lysate was then determined according to the following equation:

\[ P_{isf} = T_{isf} \cdot R_{isf} = T_{isf} \cdot R_{aqq} \]  

(Eq. 6)

Thus, P<sub>isf</sub> of recombinant ERF110 protein was determined to be 99.6 ± 5.3 fmol·mg⁻¹ and 43.6 ± 1.6 fmol·mg⁻¹ in the total cellular protein sample of air- and ethylene-treated ERF110-ox/ein2-5 Arabidopsis, respectively (Table I). It is clear that ethylene treatment reduced the Ser-62 phosphorylation level to 44.0 ± 3.1% of that of the air control in the transgenic plant (Fig. 4c).

Substantiation of the AQUIP Results Using the SILIA-based Relative and Accurate Measurement Approach—A SILIA-based relative and quantitative phosphoproteomic method has recently been developed to measure fold changes in a target phosphoprotein abundance in Arabidopsis (36), which essentially measures the mass ratio of a phosphoprotein produced in two different plant tissue samples. Given that the molar ratio of a pair of phosphoproteins present in a pair of differentially treated plant tissues is equivalent to the mass ratio of the same pair of phosphoproteins, the molar ratio of the Ser-62-phosphorylated isoform of recombinant ERF110 protein determined using the above AQUIP method should be substantiated with the relative and accurate quantitation phosphoproteomic protocol as illustrated in the relative quantitation workflow (supplemental Fig. 3). A standard curve was then established by correlating the ratios of the observed ion intensities of the whole isotopic envelopes (abbreviated as RI) of ¹⁴N/¹⁵N-coded recombinant ERF110 peptide precursors (all three standard peptides, P<sub>1</sub>, P<sub>2</sub>, and Ser-62 nonphosphorylated peptide NP<sub>a</sub>, were used) and the mass ratios of the
An AQUA strategy for relative quantitation of Ser-62-phosphorylated isoform over all isoforms of recombinant ERF110 proteins. A known quantity (12 pmol) of $^{15}$N-coded recombinant ERF110 proteins, including both Ser-62-phosphorylated isoform (circled P) and all of its cognates of different types of PTMs (circled P or triangle), were highly purified from ERF110-ox/ein2-5 Arabidopsis (either air- or ethylene-treated for 12 h), divided into six aliquots, and resolved on SDS-PAGE gel, followed by in-gel trypsin digestion. The $^{14}$N-coded synthetic Ser-62-phosphopeptide and its nonphosphorylated cognate were mixed together with concentrations 5, 10, 20, 40, 80, or 160 fmol and 62.5, 125, 250, 500, 1000, or 2000 fmol, respectively. Then the $^{14}$N-coded peptide standard pairs were spiked into six aliquots of $^{15}$N-coded peptide samples. After the oxidization of methionine residues on the peptides and LC-MS/MS analysis, two standard curves were built by this AQUA method between the ion intensity of whole isotopic envelopes of standard peptide $P$ and NP. The amino acid sequence of Ser-62 phosphosite-containing oligopeptide is VD$^S$SHNPIEESMSK. The molar amounts of Ser-62-phosphorylated peptide ($P_{aqu}$) and its nonphosphorylated cognate ($NP_{aqu}$), both of which were derived from the highly purified and $^{15}$N-coded ERF110 protein isoforms, were determined according to the two standard curves, respectively. The percentage of Ser-62-phosphorylated ERF110 isoform among all of ERF110 isoforms in the total cellular proteins is defined as the site-specific phosphorylation occupancy $R_{isf}$, which is equivalent to Ser-62 phosphosite occupancy $R_{aqu}$ in the peptide sample and determined using AQUA method. The error values indicate standard deviations.
**Absolute Quantitation of Isoforms of PTM Protein**

**Fig. 4.** Determination and substantiation of Ser-62-phosphorylated ERF110 isoform concentration via AQUIP and relative quantitation. 

- **a** and **b**, standard curves constructed for the absolute quantitation of the molar amount of Ser-62-phosphopeptide (P<sub>aqu</sub>) and its nonphosphorylated cognate (NP<sub>aqu</sub>) from the peptide sample derived from the highly purified recombinant ERF110 proteins. The standard curves for measuring the Ser-62 phosphosite peptides, P<sub>aqu</sub> and NP<sub>aqu</sub> value, from air-treated ERF110-ox/ein2-5 Arabidopsis are P<sub>aqu</sub> = 0.0427P<sub>ion</sub> + 2.1545 (r² = 0.9984), and NP<sub>aqu</sub> = 0.0516P<sub>ion</sub> + 12.224 (r² = 0.9996), respectively, whereas the standard curves for measuring Ser-62 phosphosite peptides from the ethylene-treated protein sample are P<sub>aqu</sub> = 0.0437P<sub>ion</sub> + 2.3165 (r² = 0.9977) and NP<sub>aqu</sub> = 0.0526P<sub>ion</sub> + 11.564 (r² = 0.9976), respectively. P<sub>aqu</sub> and NP<sub>aqu</sub> values of ERF110 proteins in both air- and ethylene-treated total cellular protein samples are labeled with open and solid stars on the standard curves, respectively. 

- **c**, relative measurement of both differentially labeled total cellular proteins (equivalent to the relative mass ratio of Ser-62-phosphorylated ERF110 isoform, abbreviated as RM). The fold change in 14N/15N-coded total cellular protein or the mass ratios (RM) of recombinant ERF110 varied from 16:1 to 1:16 (supplemental Fig. 4, a–d, and supplemental Table 4), and the formula for this standard curve was determined to be log<sub>RM</sub> = 0.8226log<sub>RI</sub> – 0.0355, r² = 0.995 (supplemental Fig. 4d). Both 14N/15N-labeled protein mixtures used to build the standard curve were from air-grown ERF110-ox/ein2-5 Arabidopsis because it has been determined that ethylene treatment did not affect the total molar amount of recombinant ERF110 protein in Arabidopsis (supplemental Figs. 5, a and b, and 6). In contrast, 12 h of ethylene treatment decreased Ser-62 phosphorylation to 34.8 ± 3.1% of that of air control when the calculation was made based on the ratio (RI = P<sub>ion-ethylene/P<sub>ion-air</sub></sub>) of ion count intensities of a differently labeled Ser-62-phosphorylated peptide pair (Fig. 4c and supplemental Fig. 5, a and b). Interestingly, if the mass ratio (RM) is taken into consideration, the Ser-62-phosphorylated isoform of recombinant ERF110 decreased to 43.2 ± 3.4% (Fig. 4c). By Student’s t test, the ethylene-altered Ser-62 phosphorylation stoichiometry (Fig. 4c) determined solely by ion count (RI) was significantly different from those determined using either the mass ratio (RM, p = 5.2 × 10⁻³) of recombinant ERF110 protein or the molar ratio of phosphorylated isoforms of recombinant ERF110 proteins from air- and ethylene-treated plant tissues (RP, p = 2.4 × 10⁻¹⁰). In contrast, the difference between molar ratio (RM) and mass ratio (RP) of Ser-62-phosphorylated ERF110 isoform measured from both air- and ethylene-treated tissues was insignificant (p = 0.48). The MS/MS spectra and Mascot results for identification of the Ser-62 phosphorylation of ERF110 were shown in supplemental Fig. 7 and supplemental Table 5, respectively.

**The Functional Role of Ser-62 Phosphorylated ERF110 Isoform in Regulation of Transgenic Arabidopsis**—While we were measuring the absolute molar amount of the Ser-62-phosphorylated ERF110 isoform in transgenic Arabidopsis, we also investigated the ethylene regulation of gene expression of the Ser-62-phosphorylated ERF110 isoform and Ser-62-phosphopeptide from both air- and ethylene-treated (12 h) ERF110-ox/ein2-5 Arabidopsis. The ratio of ion intensity (RI) of the whole isotopic envelopes of 1⁴N-coded over 1⁵N-coded phosphopeptides derived from both air- and ethylene-treated (12 h) ERF110-ox/ein2-5 Arabidopsis (RI = P<sub>ion-ethylene/P<sub>ion-air</sub></sub> where P<sub>ion</sub> = (P<sub>ion</sub> + NP<sub>ion</sub>) × R<sub>ion</sub>, Equation S14 in supplemental Fig. 1b). The mass ratio (RM) of the 1⁴N/1⁵N-coded ERF110 phosphoprotein in between air- and ethylene-treated total cellular protein samples is determined using a standard curve log<sub>RM</sub> = 0.8226log<sub>RI</sub> – 0.0355, r² = 0.995 (supplemental Fig. 4d). The molar ratio (RP) of Ser-62-phosphorylated ERF110 isoform (P<sub>aq</sub>) between air- and ethylene-treated protein samples is determined according to a ratio of P<sub>lat-ethylene/P<sub>lat-air</sub></sub> where P<sub>lat</sub> = T<sub>lat</sub>R<sub>lat</sub> (or R<sub>aq</sub>) and R<sub>aq</sub> = P<sub>aq</sub>/P<sub>aq</sub> + NP<sub>aq</sub>p (t test): 5.2 × 10⁻³ (RI versus RM), 2.4 × 10⁻¹⁰ (RI versus RP). ***; t test significance p < 0.001.
native ERF110 gene and the biological effect of Ser-62 phosphorylated isoform of recombinant ERF110 protein (a PTM isoform) on Arabidopsis growth and development. It was found that ethylene indeed up-regulated the gene expression of the putative ethylene response transcription factor 110 (ERF110) under the wild type background but not in the ethylene-sensing defective mutants, ein2-5 and eptr1-1 (supplemental Fig. 8). These molecular biological results strongly suggest that the ERF110 gene is among one of many ethylene-regulated genes and it plays a role in ethylene response in Arabidopsis. Overexpression of ERF110 in ein2-5 (ERF110-ox/ein2-5) leads to the bypassing of EIN2 signaling component-mediated transcriptional regulation of ERF110 gene expression. Consequently, ERF110-ox/ein2-5 plant surprisingly developed a bushy phenotype with the increased number of rosette leaf and inflorescence (Fig. 5). Both rosette leaf and inflorescence number increased from 34.5 ± 1.0 to 65.8 ± 4.4 (p = 2.4 × 10^-5) and from 4.8 ± 0.2 to 7.6 ± 0.4 (p = 2.9 × 10^-6), respectively. To determine whether it was the high level of Ser-62-phosphorylated ERF110 isoform that induced such an interesting phenotype, a site-directed mutagenesis was made at Ser-62 position on ERF110, and Ser-62 was changed into Ala-62. The resulting S62A ERF110 mutant was transformed into ein2-5 Arabidopsis background. The point mutation S62A in ERF110 protein, which mimics the Ser-62 dephosphorylation on ERF110, clearly reduced its ability in promoting the bushy phenotype (Fig. 5). In fact, both the rosette leaf and the inflorescence number decreased from 65.8 ± 4.4 to 40.2 ± 2.1 (p = 1.4 × 10^-5) and from 7.6 ± 0.4 to 5.9 ± 0.4 to (p = 0.003), respectively, suggesting a functional role for the Ser-62-phosphorylated isoform of ERF110 in Arabidopsis growth and development.

DISCUSSION

A population of proteins expressed from a single gene normally consists of numerous substoichiometric post-translational modification isoforms, the composition of which is dynamic and highly regulated in a temporal and spatial fashion and in response to environmental cues. The PTMs of a protein, which occur at distinct amino acid side chains or peptide linkages and are typically catalyzed by substrate-specific enzymes, play a critical role in regulation of gene expression and versatility of gene function (55, 56). There have been successful examples where the amount of each protein PTM isoform present in various tissue samples was characterized, quantitated, and compared (7, 9, 39), and a comprehensive statistical analysis of a protein isoform (e.g. a PTM isoform of a tissue- and stage-specific biomarker) provided predictive information on the prognosis of diseases (47, 48). Thus, in this study, a plant hormone response transcription factor ERF110 (supplemental Fig. 8) was selected for quantitative PTM analysis for reasons that (i) Ser-62 phosphosite of ERF110 originates from a computational prediction of phosphorylation motifs, and its prediction was based upon an ethylene-regulated phosphopeptide from the ethylene-insensitive mutant ein2-5 (12); (ii) Ser-62 phosphosite has only been found to be phosphorylated during the in vitro plant kinase assays, the level of which is down-regulated by ethylene in ethylene-treated ein2-5 mutant (12). Thus, in this experiment, overexpression of the ethylene-inducible gene, ERF110, in ein2-5 should result in a constitutive expression of this gene under an EIN2 loss-of-function genetic back-
ground (ein2-5) and bypass the suppression of ein2-5 mutation on ERF110 gene expression. More importantly, the ERF110 overexpression provides, for the first time, an in vivo validation on regulation by ethylene of Ser-62 phosphorylation of ERF110.

The phosphorylation/dephosphorylation events occurred on Ser-62 phosphosite are independent to the function of EIN2 (Table I). To our surprise, the overexpression of ERF110 gene generates a bushy phenotype in the ethylene-insensitive ein2-5 mutant line (Fig. 5). This phenotype is similar to that of Arabidopsis bud1 mutant, the genetic defect of which has been found to result from an overexpression of a mitogen-activated protein kinase kinase 7 (At-MKK7) gene (57). Site-directed mutagenesis of ERF110 protein at Ser-62 phosphosite, S62A, confirms a functional role for the Ser-62 phosphorylation in promoting bushy phenotype (Fig. 5). Thus, our in vivo evidence suggests a link between AtMKK7 and AtERF110 gene in regulation of the bushy phenotype. It is possible that an EIN2-independent ethylene signaling pathway(s) suppresses the bushy phenotype by removing Ser-62 phosphorylation, whereas both EIN2- and MKK7-dependent cascade signaling enhances the ERF110 gene expression, as well as produces more Ser-62-phosphorylated ERF110 isoform to elicit the bushy phenotype. In fact, the dual and opposing effects of ethylene have been documented before on numerous ethylene-regulated gene expression and responses (5, 58–60). Although the actual role for ERF110 in regulation of reproductive growth and development in Arabidopsis is still under investigation, the molecular genetics data on the Ser-62 phosphosite of ERF110 nonetheless justifies a biological relevance for the MS-based proteomic quantitation of Ser-62-phosphorylation of ERF110 using AQUIP.

An ideal scenario of application of MS in protein PTMs researches would be to quantitatively measure the abundance, processing, and PTMs of all proteins in a single run. Not surprisingly, the reality of realizing such a goal is still far-fetched at this moment (17, 61). One argument is that the total number of tryptic peptides in a eukaryotic cell probably enormously exceeds the peptide identification capacity of every mass spectrometry device available to date, and it becomes even more complicated when considering missed cleavages, peptide modifications, redundant sequencing, and so on (17). Another practical problem associated with MS application is the challenge that PTMs often present at lower levels in biological samples, such as in the case of ERF110 gene products. Throughout our studies, the native ERF110 protein has never been detected using MS except using Western blot analysis with polyclonal anti-ERF110 antibodies (data not shown). Thus, an efficient affinity-based enrichment becomes an inevitable approach in study of PTMs of ERF110 that are extremely low in quantity and heterogeneous in population (7, 61, 62). One of the affinity-based enrichment strategies is the TAP, which has been commonly applied in rapid and highly specific purification of protein complexes under native conditions from diverse organisms ranging from yeast (19, 20, 63) and Escherichia coli (64), to Drosophila (65), Arabidopsis (66–68), and human cells (69–71). However, an intrinsic limitation commonly associated with the TAP strategy is the requirement of native conditions in at least one of purification phases, which therefore makes it susceptible to a loss of PTMs during cell lysis and purification because of the presence of active modifying as well as demodifying enzymes, such as kinase and phosphatase, under the native conditions. To overcome these limitations, a modified version of TAP method using the histidine-biotin fusion polypeptides as a tandem affinity tag was developed, which allows a two-step purification of ubiquitinated proteins from yeast under a fully denaturing condition (21). To increase the binding capacity of histidine-biotin-tagged recombinant protein to affinity beads (Ni²⁺-nitrilotriacetic acid-agarose) under a highly denaturing condition (8 M urea), we have modified the histidine-biotin tag into a new HBH tag (supplemental Fig. 2a), in which a polyhistidine (His₆) oligopeptide was genetically engineered and fused to both the N and C termini of biotin carboxyl carrier protein. Furthermore, recombinant protein has been shown to share a similar PTM profile with that of its native gene product (28–31). Thus, an ERF110-HBH hybrid protein was therefore made and overexpressed in ein2-5 Arabidopsis to study the effect of ethylene on Ser-62 phosphorylation on ERF110 using AQUIP.

Metabolic labeling of organisms with heavy nitrogen (¹⁵N) has been successfully introduced into differential proteome analysis of both plant and mammalian cells and tissues (13, 35, 36). The ¹⁵N stable isotope labeling peptide pairs used in AQUIP demonstrates a much bigger mass difference (usually >10 Da), making it easier to distinguish between the peptide pairs for multiply charged species when using lower resolution mass spectrometers (supplemental Table 1). More importantly, AQUIP integrates advantages from both AQUA and QconCAT strategies. For examples, the commercially available ¹⁴N-coded synthetic peptides were used like the AQUA approach in measuring the absolute levels of peptides (72) and protein phosphorylations with phosphopeptide standards (7, 9). In contrast, ¹⁵N-coded synthetic peptides were not employed because of their relative higher cost. In the case of QconCAT strategy, it is excellent for multiplex absolute quantitation of isoforms of proteins by introducing the isotope-labeled artificial polypeptide concatemer into the sample (40–43), yet the molecular biological manipulation involved in preparation of a recombinant QconCAT gene may demand extra effort (44). Thus, an approach of choice may be to combine the use of ¹⁵N-coded synthetic AQUA peptide together with the less costly in planta production of ¹⁴N-coded target proteins via SILIA.

Another aspect of AQUIP is that it integrates the widely accepted concept of PSAQ strategy (73) with the in planta expression of a ¹⁵N-coded recombinant target protein and measures the actual efficiency of peptide yield of a target
protein from a total cellular protein sample (Figs. 1 and 2). Moreover, it avoids performing kinase reaction in vitro to get a phosphorylated PSAQ protein as the PTM isoform standard that is required in absolute quantitation of a target protein. In fact, the upstream kinase activities for a specific phosphorylation site are not readily measurable. The AQUIP strategy simply employs four stable isotope-free peptides in total to absolutely quantitate a site-specific phosphorylation stoichiometry of a target protein like ERF110. As shown in Fig. 2b and Table I, the peptide yield was peptide-dependent and consistent with what Havlis and Shevchenko (46) once reported. Our results also suggest that if AQUA peptide standards were used, the peptide yield for each individual peptide should be evaluated to avoid system errors in absolute quantitation (Table I). In our study, the molar amount of highly purified and 15N-coded recombinant target protein (Figs. 1 and 3), which was used for determining the peptide yield, should be close to the molar amount of the same target protein in the whole Arabidopsis cell lysate. Otherwise, a unique standard curve should be established individually for determining all isoforms (T_isf) of the recombinant target protein from either the whole Arabidopsis cell lysate or the highly purified recombinant target protein preparation. In addition, a pair containing a phosphopeptide and its nonphosphorylated cognate should be mixed together at a series of molar ratios and then spiked into aliquots of trypsin-digested peptide samples derived from highly purified recombinant 15N-coded ERF110 protein because they share the identical amino acid sequence and the same trypsin digest site, and it is expected that they have the same peptide yield from in-gel trypsin digestion so that the phosphorylation occupancy (R_isf), i.e. the Ser-62-phosphorylated isoform over all isoforms of ERF110, is equal to the phosphosite occupancy (R_aq) of Ser-62-containing phosphopeptide over the total Ser-62-containing tryptic peptides of ERF110 (supplemental Fig. 1). If the molar amount of NP_aq peptide is far bigger than that of P_aq, an approximate equation, R_isf = R_aq = P_aq/NP_aq, can be used for determination of phosphorylation occupancy at Ser-62 site.

Moreover, throughout the AQUIP quantitative proteomic analysis, R_isf of the recombinant ERF110 protein is determined under a fully denaturing condition because it is able to ensure the population of a specific PTM isoform unchanged during protein extraction. Especially, the direct in-gel tryptic digestion of unpurified target protein following urea-based protein extraction (Fig. 1, left panel) is a core step in AQUIP strategy. Although the efficiency of in-solution digestion normally exceeds 80–90% and is much higher than the 10–40% resulting from in-gel digestion (46), the in-gel digestion was still chosen because the target protein, such as ERF110, can be enriched from the total cellular proteins by one-dimen-
sional SDS-PAGE separation under SDS-denaturing condition, and it most likely excludes the possibility of refolding of modifying enzymes like protease, kinase, and phosphatase during fractionation and in-gel digestion. According to the same rationale, the highly purified 15N-coded recombinant ERF110 protein was also isolated through an identical procedure (Fig. 1 and 3, right panel). For exactly the same reason, we suspect that the application of AQUIP strategy on the quantitation of PTMs of a target protein purified through antibodies and co-immunoprecipitation from plant total cellular protein may introduce artifacts because immunoprecipitation and co-immunoprecipitation requires non-denaturing buffer to perform. Even direct performance of in-solution digestion on 8 M urea-denatured protein may introduce a change on PTMs because protease digestion solution requires a decrease in urea concentration from 8 to 2 M for the basic trypsin digestion reaction to work (74). The lower molar concentration of urea may result in renaturation of modifying and demodifying enzymes.

Because MS equipped with electrospray dissociation or electrospray dissociation detector becomes more popular in characterizing certain PTMs that are labile under most MS fragmentation conditions (75), middle- and top-down proteomic analysis strategies are frequently adopted in quantitative proteomics. Because a recombinant protein highly purified through a fully denaturing procedure like what is applied here is likely to retain most, if not all, PTM isoforms produced in vivo, the combinatorial PTM occupancy (R_iso) of a protein will be measured directly using top-down MS analysis of the whole isoform, and at the same time the absolute total amount of all isoforms (T_iso) will be determined using PTM-free and 14N-coded synthetic AQUA peptides against 15N-coded recombinant target protein (Fig. 1 and Table I). It is therefore believed that AQUIP may be applicable in the absolute quantitation of a protein isoform (Piso) of combinatorial post-translational modifications in both prokaryotes and eukaryotes.

* This work was supported by Grants CAS10SC01, CAS10SC01M1, GMGS11SC02, HKUST10/CRF/10, NMES11SC01, W0298, 661207, and 661408 from the Hong Kong Research Grant Council, Croucher Foundation, The Hong Kong University of Science and Technology (to N. L.).

† To whom correspondence should be addressed. Tel.: 852-23587335; Fax: 852-23581559; E-mail: boningli@ust.hk.

REFERENCES
1. Mann, M., and Jensen, O. N. (2003) Proteomic analysis of post-translational modifications. Nat. Biotechnol. 21, 255–261
2. Witze, E. S., Old, W. M., Resing, K. A., and Ahn, N. G. (2007) Mapping protein post-translational modifications with mass spectrometry. Nat. Methods 4, 798–806
3. Krishna, R. G., and Wold, F. (1993) Post-translational modification of proteins. Adv. Enzymol. Relat. Areas Mol. Biol. 67, 265–298
4. Cohen, P. (2002) The origins of protein phosphorylation. Nat. Cell Biol. 4, E127–E130
5. Yoo, S. D., Cho, Y. H., Tena, G., Xiong, Y., and Sheen, J. (2008) Dual control of nuclear EIN3 by bifurcate MAPK cascades in C2H4 signalling. Nature 451, 789–795
6. Suzuma, K., Takahara, N., Suzuma, I., Ishikii, K., Ueki, K., Leitges, M., Aiello, L. P., and King, G. L. (2002) Characterization of protein kinase Cζ isoform’s action on retinoblastoma protein phosphorylation, vascular
endothelial growth factor-induced endothelial cell proliferation, and retinal neovascularization. *Proc. Natl. Acad. Sci. U.S.A.* 99, 721–726

7. Mayya, V., Rezual, K., Wu, L., Fong, M. B., and Han, D. K. (2006) Absolute quantification of multisite phosphorylation by selective reaction monitoring mass spectrometry: Determination of inhibitory phosphorylation status of cyclin-dependent kinases. *Mol. Cell. Proteomics* 5, 1148–1157

8. Bonenfant, D., Schmelzle, T., Jacinto, E., Crespo, J. L., Mini, T., Hall, M. N., and Jenow, P. (2003) Quantitation of changes in protein phosphorylation: A simple method based on stable isotope labeling and mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* 100, 880–885

9. Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., and Gygi, S. P. (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. U.S.A.* 100, 8940–8945

10. Steen, H., Jobanathan, A. J., Springer, M., and Kirschner, M. W. (2005) Quantitative proteomics: relative and absolute quantitation of protein phosphorylation stoichiometry by MS. *Proc. Natl. Acad. Sci. U.S.A.* 102, 3948–3953

11. Olsen, J. V., Blagoev, B., Nead, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 127, 635–648

12. Li, H., Weng, W. S., Zhu, L., Guo, H. W., Ecker, J., and Li, N. (2009) PhosphoTAP: proteomic analysis of ethylene-regulated protein phosphorylation in etiolated seedlings of Arabidopsis mutant et2 using two-dimensional separations coupled with a hybrid quadrupole time-of-flight mass spectrometer. *Proteomics* 9, 1646–1661

13. McClatchy, D. B., Liao, P., Park, S. K., Xu, T., Li, Y., Zhang, H., and An, L. (2010) The tandem affinity purification (TAP) method: A general procedure of protein complex purification and protein interaction identification. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6069–6074

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

15. Kirkpatrick, D. S., Gerber, S. A., and Gygi, S. P. (2005) The absolute quantitative proteomics strategy: A general procedure for the quantification of proteins and post-translational modifications. *Methods* 35, 265–273

16. Macek, B., Mann, M., and Olsen, J. V. (2009) Global and site-specific quantitative phosphoproteomics: Principles and applications. *Annu. Rev. Pharmacol. Toxicol.* 49, 199–221

17. Domon, B., and Aebersold, R. (2010) Options and considerations when selecting a quantitative proteomics strategy. *Nat. Biotechnol.* 28, 747–751

18. Xu, X., Song, Y., Li, Y., Chang, J., Zhang, H., and An, L. (2010) The tandem affinity purification method: An efficient system for protein complex purification and protein interaction identification. *Protein Expr. Purif.* 72, 149–156

19. Puig, O., Caspary, F., Rigbolt, K. T., Nielsen, M. M., and Blagoev, B. (2011) Characterization of ubiquitination dependent dynamics in growth factor receptor signaling by quantitative proteomics. *Mol. Biosyst.* 7, 3223–3233

20. Ramus, C., Gonzalez de Peredo, A., Dahout, C., Gallagher, P., and Jan, J. (2006) An optimized strategy for ICAT quantification of membrane proteins. *Mol. Cell. Proteomics* 5, 88–78

21. Robertson, M., Helliwell, C. A., and Dennis, E. S. (2008) Post-translational modifications of the endogenous and transgenic FLC protein in Arabidopsis thaliana. *Plant Cell Physiol.* 49, 1859–1866

22. Rebecchi, K. R., Go, E. P., Xu, L., Woodin, C. L., Mure, M., and Desaire, H. (2011) A general protease digestion procedure for optimal protein sequence coverage and PTM analysis of recombinant glycoproteins: Application to the characterization of hLOXL2 glycosylation. *Anal. Chem.* 83, 4848–4849

23. Kwon, J., Jeong, J., Kim, Y. M., Hwang, N., Paek, E., and Lee, K. J. (2008) Strategy for comprehensive identification of post-translational modifications in cellular proteins, including low abundant modifications: Application to glyceraldehyde-3-phosphate dehydrogenase. *J. Proteome Res.* 7, 587–602

24. Kwon, J., Jeong, J., Kim, Y. M., Hwang, N., Paek, E., and Lee, K. J. (2008) Strategy for comprehensive identification of post-translational modifications in cellular proteins, including low abundant modifications: Application to glyceraldehyde-3-phosphate dehydrogenase. *J. Proteome Res.* 7, 587–602

25. Lebert, D., Dupuis, A., Detry, M. K., Doherty, M. K., and Barnidge, D. R. (2011) Absolute quantification of the model biomarker prostate-specific antigen in serum by LC-MS/MS using protein cleavage and isotope dilution mass spectrometry. *J. Proteome Res.* 10, 1416–1427

26. Brun, V., and Doherty, M. K. (2006) Absolute quantification of the model biomarker prostate-specific antigen in serum by LC-MS/MS using protein cleavage and isotope dilution mass spectrometry. *J. Proteome Res.* 10, 1416–1427

27. Lebert, D., Doherty, M. K., and Freudenstein, J. V. (2008) Absolute quantification of proteins in plasma. *Anal. Chem.* 80, 4026–4036

28. Schiff, J. A., and Mann, M. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 127, 635–648

29. Mayya, V., Rezual, K., Wu, L., Fong, M. B., and Han, D. K. (2006) Absolute quantification of multisite phosphorylation by selective reaction monitoring mass spectrometry: Determination of inhibitory phosphorylation status of cyclin-dependent kinases. *Proc. Natl. Acad. Sci. U.S.A.* 100, 880–885

30. Li, H., Weng, W. S., Zhu, L., Guo, H. W., Ecker, J., and Li, N. (2009) PhosphoTAP: proteomic analysis of ethylene-regulated protein phosphorylation in etiolated seedlings of Arabidopsis mutant et2 using two-dimensional separations coupled with a hybrid quadrupole time-of-flight mass spectrometer. *Proteomics* 9, 1646–1661

31. McClatchy, D. B., Liao, P., Park, S. K., Xu, T., Li, Y., Zhang, H., and An, L. (2010) The tandem affinity purification (TAP) method: A general procedure of protein complex purification and protein interaction identification. *Protein Expr. Purif.* 72, 149–156

32. Puig, O., Caspary, F., Rigbolt, K. T., Nielsen, M. M., and Blagoev, B. (2011) Characterization of ubiquitination dependent dynamics in growth factor receptor signaling by quantitative proteomics. *Mol. Biosyst.* 7, 3223–3233

33. Ramus, C., Gonzalez de Peredo, A., Dahout, C., Gallagher, P., and Jan, J. (2006) An optimized strategy for ICAT quantification of membrane proteins. *Mol. Cell. Proteomics* 5, 88–78

34. Robertson, M., Helliwell, C. A., and Dennis, E. S. (2008) Post-translational modifications of the endogenous and transgenic FLC protein in Arabidopsis thaliana. *Plant Cell Physiol.* 49, 1859–1866

35. Rebecchi, K. R., Go, E. P., Xu, L., Woodin, C. L., Mure, M., and Desaire, H. (2011) A general protease digestion procedure for optimal protein sequence coverage and PTM analysis of recombinant glycoproteins: Application to the characterization of hLOXL2 glycosylation. *Anal. Chem.* 83, 4848–4849

36. Kwon, J., Jeong, J., Kim, Y. M., Hwang, N., Paek, E., and Lee, K. J. (2008) Strategy for comprehensive identification of post-translational modifications in cellular proteins, including low abundant modifications: Application to glyceraldehyde-3-phosphate dehydrogenase. *J. Proteome Res.* 7, 587–602

37. Lebert, D., Doherty, M. K., and Freudenstein, J. V. (2008) Absolute quantification of the model biomarker prostate-specific antigen in serum by LC-MS/MS using protein cleavage and isotope dilution mass spectrometry. *J. Proteome Res.* 10, 1416–1427

38. Brun, V., and Doherty, M. K. (2006) Absolute quantification of the model biomarker prostate-specific antigen in serum by LC-MS/MS using protein cleavage and isotope dilution mass spectrometry. *J. Proteome Res.* 10, 1416–1427

39. Lebert, D., Doherty, M. K., and Freudenstein, J. V. (2008) Absolute quantification of the model biomarker prostate-specific antigen in serum by LC-MS/MS using protein cleavage and isotope dilution mass spectrometry. *J. Proteome Res.* 10, 1416–1427

40. Lebert, D., Doherty, M. K., and Freudenstein, J. V. (2008) Absolute quantification of the model biomarker prostate-specific antigen in serum by LC-MS/MS using protein cleavage and isotope dilution mass spectrometry. *J. Proteome Res.* 10, 1416–1427
59. Li, N. (2008) The dual-and-opposing effect of ethylene on the negative gravicurvature of Arabidopsis thaliana. Plant J. 16, 735–743.

50. Nakano, T., Suzuki, K., Fujimura, T., and Shinshi, H. (2006) Genome-wide analysis of the ERF gene family in Arabidopsis and rice. Plant Physiol. 140, 411–432.

51. Clough, S. J., and Bent, A. F. (1998) Floral dip: A simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J. 16, 735–743.

52. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V., and Mann, M. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat. Protoc. 1, 2856–2860.

53. Kettenbach, A. N., Rush, J., and Gerber, S. A. (2011) Absolute quantification of protein and post-translational modification abundance with stable isotope-labeled synthetic peptides. Nat. Protoc. 6, 175–186.

54. Jaffe, J. D., Keshishian, H., Chang, B., Addona, T. A., Gillette, M. A., and Jensen, O. N. (2004) Modification-specific proteomics: Characterization of post-translational modifications by mass spectrometry. Curr. Opin. Chem. Biol. 8, 33–41.

55. Jensen, O. N. (2004) Modification-specific proteomics: Characterization of post-translational modifications by mass spectrometry. Curr. Opin. Chem. Biol. 8, 33–41.

56. Walsh, C., ed. (2006) Posttranslational Modification of Proteins: Expanding the proteome. Roberts and Company Publishers, Greenwood, CO.

57. Dai, Y., Wang, H., Li, B., Huang, J., Liu, X., Zhou, Y., Mou, Z., and Li, J. (2008) Accurate inclusion mass screening: A bridge from unbiased discovery to targeted assay development for biomarker verification. Mol. Cell. Proteomics 7, 1952–1962.

58. Lu, B. W., Pei, L. K., Chan, W. K., Zhang, H., Zhu, G., Li, J. Y., and Li, N. (2006) Genome-wide proteomic characterization of combinatorial post-translational modification patterns. Exp. Rev. Proteomics 3, 517–520.