A Self-validating Quantitative Mass Spectrometry Method for Assessing the Accuracy of High-content Phosphoproteomic Experiments*§

Pedro Casado‡ and Pedro R. Cutillas‡§

Protein kinase pathways play pivotal roles in cell signaling and biology. The phosphoproteome is a reflection of protein kinase pathway activation and therefore there is considerable interest in its quantification as a means to assess the wiring of signaling networks. Although different approaches for quantitative phosphoproteomics have been described, there is no data on how accurate these are for each quantified phosphorylated site. We report a liquid chromatography-MS approach to objectively assess data quality in high-content comparison of phosphoproteomes in which samples to be compared are mixed at different proportions. The experimental data is then used to derive a linear regression function that allows calculating correlation values, linearity, and accuracy. We applied the technique to investigate phosphorylation in P31/Fuj and Kasumi-1, two leukemia cell lines showing strikingly different sensitivities to scr and PI3K inhibitors. We found that phosphopeptides quantified with accuracy were not always quantified with precision because of low ion statistics contributing to variability. Thus our approach was complementary to standard methods for calculating the precision of replicate measurements based on the coefficient of variation and provided additional information on data quality for each quantified phosphopeptide. We quantified > 2250 phosphorylation sites across cell lines with different levels of sensitivity to kinase inhibitors, of which 1847 showed an accuracy variation of < 30% (with an overall mean of 22%). Hundreds of phosphorylation sites on proteins with diverse function (including kinases, transcription, and translation factors) showed significantly distinct intensities across sensitive and resistant cells lines, indicating that kinase pathways are differentially regulated in cancer cells of distinct sensitivity to signaling inhibitors. Molecular & Cellular Proteomics 10.1074/mcp.M110.003079, 1–11, 2011.

Protein kinases have diverse roles in cell signaling processes that in turn control normal and disease physiology.

The human genome contains more than 500 kinase genes and these phosphorylate thousands of amino acid residues on proteins. Quantification of phosphorylation sites therefore provides a means to assess kinase pathway activation, as phosphorylation site abundance also reflects phosphatase activity and the gene expression of their phosphoprotein substrates. There is therefore a great interest in the development and use of phosphoproteomics approaches as a means to quantify kinase pathway activation in many areas of biomedical research.

Examples of successful application of phosphoproteomics to understand cell biology have been reported. Most of these studies used metabolic or chemical labeling, such as SILAC or iTRAQ, prior to liquid chromatography-mass spectrometry (LC-MS) for relative quantification of phosphopeptides derived from the proteolytic digestion of whole cell lysates (1–4). Studies in which relative phosphoprotein quantification is performed using label-free approaches have also been reported (5–8). Label-free quantitative phosphoproteomics circumvents shortcomings inherent to labeling strategies, which include the difficulty of comparing large sample numbers, and their cumbersome and costly nature. However, although in principle attractive, it is at present not known how accurate label-free phosphoproteomics data are (it may also be argued that, because of the difficulty in obtaining this information, there is not data on how accurate label-based strategies are for each quantified phosphorylation site in large-scale experiments). The purpose of the present study was to develop a method that could be used to assess the accuracy of quantification for each of the thousands of phosphorylation sites that can be quantified by LC-MS in high-content phosphoproteomics experiments.

Label-free quantitative phosphoproteomics based on LC-MS involves comparing MS or MS/MS intensities of ionized peptides in different samples and assumes that peptide ion intensities are a measure of peptide abundance across the samples. However, this assumption does not always hold true; for example it has been shown that, when quantifying...
proteins, the ionic intensities of only a subset of peptides reflect protein abundance accurately (9, 10), whereas other peptides, derived from the same protein, poorly reflect protein amounts. The reasons for this phenomenon are not known but potential causes include presence of unsuspected modifications, chemical instability of some peptides, differences in ionization efficiencies across distinct experimental conditions, and/or localization of these poorly behaved peptides in protein regions that are not well digested during proteolysis. Analysis of variation of replicate experiments may not find these nonproteotypic peptides because this approach cannot detect systematic biases in sample processing and analysis. This is not a major problem for protein quantification as several peptides derived from the same protein add confidence to the quantitative data (11); analytically badly behaved peptides, which for the purpose of this article we term nonproteotypic, may then be considered as outliers and excluded from the analysis. However, the existence of nonproteotypic peptides is a major problem for the quantification of phosphopeptides because only one phosphopeptide ion is normally detected per phosphopeptide molecule in data dependent acquisition experiments of complex peptide mixtures; thus limited redundancy of data means that at present it is not possible to assess which of the thousands of phosphopeptides detectable in label-free LC-MS experiments (2, 3, 12, 13) are well behaved analytically (i.e. proteotypic) and can thus be quantified with accuracy.

The importance of measuring linearity and accuracy of quantification for each analyte is illustrated by well-developed and validated MS-based clinical and forensic assays. These procedures use calibration curves for each analyte to be quantified to determine the linear dynamic range of quantification and typical values of accuracy and precision of the method. These calibration curves are needed for accurate quantification even when the analysis involves the use of isotope labeled internal standards (14). This approach is being adopted by the proteomics community for targeted quantification of proteins (15, 16), most commonly accomplished by selected reaction monitoring. However, calibration curves can only be constructed for a limited number of analytes because this entails synthesizing standards for each of the compounds one wishes to quantify. Therefore this approach is not suitable for high-content or large-scale experiments, such as in phosphoproteomics, in which thousands of phosphorylation sites are quantified per experiment.

The approach described here for determining linearity and accuracy of quantitative phosphoproteomic data involves mixing the cell lysates of samples to be quantified at different proportions prior to digestion, phosphopeptide enrichment and LC-MS analysis (Fig. 1A). The linear regression function is then used to derive objective measures of linearity and accuracy for each phosphopeptide. This accuracy parameter is determined by measuring how much experimental intensities deviate from the values calculated using the linear regression function (Fig. 1B).

As a proof of principle, we used this procedure to determine phosphorylation sites that are differentially regulated in two acute myeloid leukemia (AML) cell lines that showed different sensitivities to kinase inhibitors. This biological system is part of our ongoing research program aimed at understanding the
molecular mechanisms that underlie heterogeneous responses of cancer cells to different inhibitors. In this regard, Kasumi-1 and P31/Fuj present an interesting model as they show markedly different sensitivities to several kinase inhibitors. An understanding of the differences in kinase signaling in these cells may shed light into the mechanisms that make cancer cells resistant to kinase inhibitors. We demonstrate that the technique can be used to assess linearity and accuracy of quantification for each of the hundreds/thousands of phosphorylation sites that can be quantified by LC-MS. Using this technique, we identified several phosphorylated peptides that were differentially regulated in AML cell lines with different sensitivities to kinase inhibitors. In addition to its use as a tool for relative quantification of phosphoproteomes in its own right, this method could also be used to compare the accuracy of different techniques and to validate proteomics workflows during the developmental stage.

**MATERIALS AND METHODS**

**Cell Culture**—P31/Fuj and Kasumi-1 acute myeloid leukemia (AML) cell lines were maintained in RPMI 1640 medium supplemented with 10% FBS, 100 units/ml of penicillin/streptomycin, and 50 μM β-mercaptoethanol at 37 °C in a humidified atmosphere at 5% CO₂. Cells were grown at about 0.5 to 2 × 10⁵ cells/ml.

**Sensitivity of AML Cell Lines to Inhibitor Treatment**—P31/Fuj and Kasumi-1 AML cell lines were seeded in 96-well plates at 1 × 10⁴ cell/ml. Following a recovery period of 24 h, cells were treated with 1 μM PI-103, 10 μM PP2, or vehicle (dimethylsulfoxide). Following 72 h treatment, cell viability was assessed by MTS assay (CellTiter 96® AQueous One Solution Cell Proliferation assay, Promega Corporation, Madison, WI). Each condition was assayed five times.

**Cell Lysis**—Twenty four hours prior to lysis, 50 × 10⁴ cells of each AML cell line were seeded at a density of 0.5 × 10⁴ cells/ml in fresh medium. Cells were harvested by centrifugation at 300 × g for 5 min, washed twice with ice cold PBS supplemented with phosphatase inhibitors (1 mM Na₂VO₃ and 1 mM NaF) and lysed with a denaturing buffer (20 mM HEPES pH 8.0, 8.3 mM urea, 1 mM Na₂VO₃, 1 mM NaF, 2.5 mM Na₃P₂O₇, 1 mM β-glycerol-phosphate) at a concentration of 10 × 10⁶ cells/ml. Following sonication, lysate debris was cleared by centrifugation at 20,000 × g for 10 min and the protein concentration in the supernatants was determined by Bradford analysis.

**Digestion and Solid-Phase Extraction**—P31/Fuj and Kasumi-1 cell lysates were mixed to a final protein concentration of 0.5 mg/ml in a volume of 1 ml. The proportions used were 0%, 30%, 50%, 70%, and 100% of Kasumi-1 extracts mixed with 100%, 50%, 30%, and 0% of P31/Fuj extracts, respectively. Protein mixtures were reduced and alkylated by sequential incubation with 4.1 mM dithiothreitol and 8.3 mM iodoacetamide. For digestion, samples were diluted to 20 mM HEPES pH 8.0 and incubated with immobilized TLCK-trypsin (20 TAME units/mg) for 16 h at 37 °C. Digestion was stopped by adding trifluoroacetic acid (TFA) at a final concentration of 1%. The resultant peptide solutions were desalted using Sep-Pak C₁₈ columns (Waters UK Ltd, Manchester, UK) according to manufacturer instructions.

**Immunobilized Metal Ion Affinity Chromatography (IMAC)**—Phosphopeptide enrichment was performed using an adapted IMAC enrichment protocol as previously reported (12). In short, Sep-Pack eluted samples were incubated for 1 h at room temperature with 300 μl of Fe(III)-coated Sepharose high-performance beads used as a 50% slurry in 50% acetonitrile/0.1% TFA. Unbound peptides were discarded and beads were sequentially washed with 300 μl of 50% acetonitrile/0.1% TFA and 300 μl of 50% acetonitrile/1% TFA. For phosphopeptide elution, beads were incubated twice with 300 μl 50% acetonitrile/1.5% ammonia water pH 11 for 1 min at room temperature. Recovered peptides were acidified by the addition of 10% formic acid, dried in a SpeedVac and stored at −80 °C.

**Nanoflow-Liquid Chromatography Tandem Mass Spectrometry**—For LC-MS/MS analysis, dried phosphopeptides were dissolved in 10 μl of 0.1% TFA and analyzed in a LTQ-Orbitrap XL mass spectrometer (Thermo Fisher Scientific, Hemel Hempstead, UK) connected online to a nanoflow ultra-high pressure liquid chromatography (nanoAcquity, Waters). This ultra-high pressure liquid chromatography delivered a flow rate of 5 μl/min (loading) and 400 nL/min (gradient elution) with an operating back pressure of about 3000 psi. Separations were performed in a BEH 100 μm × 100 mm column (Waters). The mobile phases were solution A: 0.1% formic acid in LC-MS grade water; and solution B: 0.1% formic acid in LC-MS grade acetonitrile. Gradient runs were from 1% B to 35% B in 100 min followed by a 5 min wash at 85% B and a 7 min equilibration step at 1% B. Full scan survey spectra (m/z 350–1600) were acquired in the Orbitrap with a resolution of 60,000 at m/z 400. A data dependent analysis was employed in which the five most abundant multiply charged ions present in the survey spectrum were automatically mass-selected, fragmented by collision-induced dissociation (normalized collision energy 35%), and analyzed in the LTQ. Thus, Full-MS scans were followed by a maximum of five MS/MS scans (m/z 50–2000) resulting in a maximum duty cycle of 2.5 s. Because chromatographic peaks were about 30 s at the base, these settings ensured that there were at least 10 data points per extracted ion chromatogram (XIC). Dynamic exclusion was enabled with the exclusion list restricted to 500 entries, exclusion duration of 40 s and mass window of 10 ppm.

**MS Data Analysis**—MS/MS data were smoothed and centroided using Mascot Distiller 2.3.2. The processed files were searched against the human sequence library in the international protein index (Human.v3.56, 76539 sequences) using the Mascot search engine (17). Searches were automated with Mascot Daemon (v2.2.2; Matrix Science, London, UK). The parameters included, choosing trypsin as digestion enzyme with two missed cleavage allowed, carbamidomethyl (C) was set as fixed modification, and Pyro-glu (N-term), Oxidation (M) and Phospho (STY) were variable modifications. Data sets were searched with a mass tolerance of ±7 ppm and a fragment mass tolerance of ±800 mnu. Hits were considered significant when they had an Expectation value <0.05 (as returned by Mascot). False discovery rates were ~2% as determined by decoy database searches. Results from Mascot searchers were uploaded in Pride (Accession numbers: 11380–11381, inclusive). No attempt was made to identify sites of modification as this was not the aim of the present study. Phosphopeptides are thus reported as the start-end residues within the protein sequence.

Because of undersampling and the stochastic nature of peak selection for fragmentation in data dependent acquisition experiments, MS/MS data was not obtained from all the phosphopeptides in all the runs. To overcome this issue, phosphopeptides identified by Mascot with a statistical significant threshold were placed in a database of peptides quantifiable by LC-MS. PESCAL (11) was used to quantify the intensities of the peptides present in the database across all the samples. PESCAL uses the m/z and retention time of the selected peptides to construct extracted ion chromatograms (XICs) for the first three isotopes of each ion. This applies restrictions on the molecular mass, retention time, charge, and isotope distribution, which permits the identification of the LC-MS elution profiles corresponding to the studied phosphopeptides with high confidence (false positive discovery rate < 5%). Windows for XIC construction were 7 ppm and 5 min for m/z and retention.
The intensity values could then be calculated by determining the peak height and areas of each individual XIC. The resulting quantitative data were parsed into Excel files for further normalization and statistical analysis. Peptide intensities were normalized to the total chromatogram intensity and further expressed as a percentage relative to the largest intensity value across samples.

Proteins containing the annotated phosphopeptides were classified by function using Ingenuity Pathways Analysis software (Ingenuity IPA 8.7-3203).

Statistical Analysis — Cell viability data were expressed as mean ± standard deviation (n = 5). Microsoft Excel was used to calculate Pearson’s correlation coefficient (R²), slopes and intercepts of linear regression functions between phosphopeptide ion intensities and percentage of cell extract in the protein mixture. Because linear regression functions were constructed using five data points, correlations were considered statistically significant at R² ≥ 0.878 (p < 0.05). The deviation from total accuracy for each phosphopeptide was calculated by subtracting the calculated relative intensity of the phosphopeptide to the theoretical intensity obtained using the linear regression function, dividing this value by the relative intensity and expressing as a percentage. This deviation from total accuracy was defined as %Acc deviation.

Cell sensitivities of P31/Fuj and Kasumi-1 cell lines to PI3K and src inhibitors.

P31/Fuj and Kasumi-1 cells were treated with vehicle (dimethylsulfoxide), 1 μM PI103, or 10 μM PP2 for 72 h and the relative number of cells was measured using an MTS assay.

Example of assessment of data quality for one phosphopeptide.

P31/Fuj and Kasumi-1 cells were processed and run in the LC-MS/MS system in three technical replicates. The phosphopeptide (SASH3 p126–143) at m/z 722.67002 was selected as an illustrative example. A, Relative intensities of XICs for this phosphopeptide in each of the individual three replicates. B, The fold change in signal intensities of m/z 722.67002 was calculated for each replicate. Mean, S.D. and %CV of fold change were also calculated. C, Cells were mixed at the ratios shown and the signal of m/z 722.67002 used to construct a linear regression function and to calculate values of linearity and accuracy as shown in Fig. 1B. D, This function was then used to recalculate intensities (see Fig 1B). Dilution refers to the percentage of P31/Fuj cell extracts in the mixture (balance was with Kasumi-1 cell extracts). Accuracy was defined as the difference between experimental and calculated intensities relative to the calculated intensity. %Acc deviation is as defined in Fig. 1B. CV of relative intensity was also calculated for each dilution.
multiplying it by 100 (formulas are shown in Fig. 1). The mean percentage accuracy (%Acc) was then calculated by averaging the accuracies of the data points from each dilution. Coefficient of variation (CV) was calculated as S.D./mean × 100.

RESULTS AND DISCUSSION

Throughout this paper we use the definition of “precision” as how reproducible the measurements can be obtained whereas the term “accuracy” is used to denote how close measured values are to the actual ones (this is related to the linearity and dynamic range of measurement). These definitions are standard in other fields of analytical chemistry (18). The analysis of variation in replicate experiments allows calculating the precision of the measurement, but replicates do not allow identifying systematic biases, assessing whether measurements are within a linear dynamic range or determining how calculated fold changes deviate from the actual differences between samples. Therefore full assessment of data quality requires methods complementary to the analysis of variance of replicate experiments (which measure the precision of the analysis) (14, 18).

Principle of the Method—The approach described here for determining linearity and accuracy of quantitative phosphoproteomic data involves mixing the cell lysates of samples to be quantified at different proportions prior to digestion, phosphopeptide enrichment and LC-MS analysis (Fig. 1A). This allows the calculation of a linear regression function, which is then used to derive objective measures of accuracy for each phosphopeptide, including Pearson’s correlation coefficients ($R^2$) and percentage of accuracy, which gives a measure of how much the measured intensities deviate from the values calculated using the linear regression function (Fig. 1B).

In order to identify differences between the samples being compared, in addition to fold change, we also report a fold coefficient (Fc), which we defined as the change in normalized phosphopeptide intensity (RI in Fig. 1B) as a function of change in the composition of the cell mixture (PE in Fig. 1B). Fc is taken as the slope of the linear regression function and ranges from 0 (no differences in phosphopeptide abundance between the samples being compared) to 1. As expected, values of Fc and fold change correlated (see below). Fc is a useful measure of differences between samples because it avoids divisions by zero, which would occur when a phosphopeptide is not detected in one of the samples, or by negative numbers, which would in some instances occur if we calculated fold changes using the intercept of the linear regression equation as an intensity value.

Biological Model—The application of the technique is best explained by an example of its use in a real study, in which we aimed to compare phosphorylation in cancer cells of distinct sensitivity to kinase inhibitors. We measured phosphorylation in Kasumi-1 and P31/Fuj, two acute myeloid leukemia (AML)
cell lines exhibiting strikingly different sensitivities to PP2 and PI-103 (Fig. 2); these two kinase inhibitors have src and phophoinositide 3-kinase (PI3K) as main targets, respectively.

Cells were lysed and mixed at the ratios shown in Fig. 1B. Following enrichment, phosphopeptides were identified by LC-MS/MS and quantified by LC-MS. The experiment was performed in two independent occasions (two biological replicates). Technical replicates (three injections of same samples) were used to evaluate technical variation.

Comparison of Precision of Measurements with their Accuracy and Linearity—We first evaluated precision and accuracy of phosphopeptides quantified from P31/Fuj and Kasumi-1 cells (Fig. 3). The data of the three technical replicates were normalized to total intensity, averaged, and precision (estimated by calculating the coefficient of variation (CV) of the measurement; see Methods section) was calculated for the fold change of each phosphopeptide in P31/Fuj relative to Kasumi-1. In order to illustrate the approach, Fig. 3 shows the analysis of one phosphopeptide at m/z 722.6700 ± 5 ppm. The intensities derived from the XIC of this peptide ion in three replicates are shown in Fig. 3A. Calculation of fold changes was then performed for each replicate (Fig. 3B) and the CV of these measurements calculated to be ~117%. Fig. 3A clearly shows that this relatively large CV is because of the large variation of the low abundant peptide in the Kasumi-1 sample. These results are not surprising as it is well known that low intensity peaks often show large variability because of poor ion statistics (19, 20).

We then calculated Fold, Fc, R2, and %Acc as described in Fig. 1B for the same phosphopeptide. These results (Fig. 3C and 3D) show that this peptide could be quantified with excellent linearity (R² = 0.996) and acceptable accuracy (Acc = 18%). We also noticed that the precision of the measurement correlated with the proportion of cells in the mixture (Table I). There also were 73 and 45 phosphopeptides increased in P31/Fuj or Kasumi-1 cells, respectively, quantified with good linearity but whose precision was less than 50% (Fig. 4A). This analysis is therefore consistent with the data shown in Fig. 3 suggesting that there was little correlation between linearity of LC-MS signals and how precise these measurements can be made. This conclusion is further supported with the plot of precision and mean %Acc shown in Fig. 4B. Although these data indicate that there is no correlation between precision and %Acc, the ability of measuring LC-MS signals with precision probably has an impact in the accuracy and linearity of quantification because data points measured with precision should contribute to a better quality of the regression function.

We sought to investigate in more detail the reasons why linearity/accuracy and precision did not correlate. As mentioned above, low precision of measurement can be attributed to poor ion statistics of low intensity peaks. This is illustrated in Fig. 3 for one phosphopeptide. We therefore hypothesized that phosphopeptides of low intensity were responsible for the low correlation between precision and linearity and accuracy of quantification (Fig. 4A and B). Fig. 4C shows the analysis of phosphopeptides with greater signals in Kasumi-1 over P31/Fuj cells that were measured with linearity/accuracy (45 phosphopeptides) but measured with low precision. As expected, there was a positive correlation between the percent of Kasumi-1 extracts in the sample and the intensity of the peptides. More interestingly, there was a negative correlation between the percent of Kasumi-1 extracts and the mean %CV values of the intensity of all these peptides (Fig. 4C), and same negative correlation was observed for percent of Kasumi-1 extracts and %Acc values of these phosphopeptides; in other words Fig. 4C shows that as phosphopeptide signal intensities decrease, the measurements become less precise and less accurate (their CV and increase). Similar results were found for the analysis of phosphopeptides increased in P31/Fuj relative to Kasumi-1 that were measured with linearity and accuracy but not precision (73 phosphopeptides) (Fig. 4D). These results confirmed that signal strength

| Dilution | Normalized Intensity | Accuracy | %Acc deviation |
|----------|----------------------|----------|----------------|
|          | Experimental      | Calculated|               |
| 100      | 100.0             | 102.3    | 102.3          |
| 70       | 75.6              | 73.3     | 97.0           |
| 50       | 56.2              | 53.9     | 95.9           |
| 30       | 33.2              | 34.6     | 104.4          |
| 0        | 4.8               | 5.6      | 117.0          |
| Mean     | 103.3             | 5.7      | 14.6           |

A Self-validating Method for Accurate Phosphoproteomics

Table I
Calculation of accuracy values for RASAL3 p-840–849
Experimental normalized phosphopeptide intensities were used to derive a linear regression function as outlined in Fig. 1B which was then used to recalculate intensities. Dilution refers to the percentage of Kasumi-1 cell extracts in the mixture (balance was with P31/Fuj cell extracts). Accuracy was defined as the difference between experimental and calculated intensities relative to the calculated intensity. %Acc deviation is as defined in Fig. 1B.
has a dramatic effect on the precision of quantification and also indicated a small effect on linearity and accuracy. However, values of accuracy and linearity, as calculated here, were less affected by signal intensity because these values were calculated by taking all the dilutions into consideration (Fig. 1) contrary to the CV of the fold changes, which were calculated by using the values at the extremes only.

As for the correlation of Fc and fold with accuracy of quantification, there was no indication that large fold differences could be quantified with greater accuracy than smaller ones, and the correlation between fold and %Acc was not statistically significant (data not shown).

### Table II: Calculation of accuracy values for Nucleolin p-63-71

| Dilution | Normalized Intensity | Accuracy |
|----------|----------------------|----------|
|          | Experimental         | Calculated | %Acc deviation |
| 100      | 56.6                 | 52.9      | 93.6         | 6.9 |
| 70       | 54.6                 | 63.5      | 116.3        | 14.0 |
| 50       | 63.8                 | 70.6      | 110.6        | 9.5 |
| 30       | 97.6                 | 77.6      | 79.6         | 25.6 |
| 0        | 80.4                 | 88.2      | 109.8        | 8.9 |
| Mean     |                      |           | 102.0        | 13.0 |

**Application of the Approach to Biological Replicates**—Our approach can be used to robustly assess the quality of phosphopeptide quantifications. However, this does not exclude the necessity of running biological replicates to assess biological significance. Therefore, to demonstrate the applicability of our approach to compare phosphorylation in a biological model, we repeated the experiment so that we could analyze the data from biological replicates. Thus, two independent sets of samples were analyzed as indicated in Fig. 1. Following normalization, we calculated R², Fc, fold changes, and accuracy values for each phosphopeptide quantified in these experiments. Tables I and II show examples of how these values were derived for the phosphopeptides (840–849) on RASAL3 (RAS protein activator like-3) and (63–71) on Nucleolin, respectively. We considered the mean deviation of accuracy (mean %Acc, which measures how far the values are from 100% accuracy) to be a more useful estimate of data quality than mean accuracy (in which a value of 100% would mean perfect accuracy) as averaging values above and below 100% often results in a value close to 100% even though individual values may be far from 100% accuracy. This is illustrated in Table II for the analysis of a

---

**Fig. 5. Examples of phosphorylated peptides measured in P31/Fuj and Kasumi-1 cell lines.** Phosphorylated peptides were measured in these two cell lines following mixing them at the percentages shown. Figures display quantitative results of two independent experiments separately (black triangles and circles, respectively) and the mean of both values (clear squares). A, Examples of measured peptides that were differentially phosphorylated in both cell lines. B, Examples of measured phosphorylation levels at positions showing no or (C) subtle differences between cells.
A Self-validating Method for Accurate Phosphoproteomics

phosphopeptide derived from Nucleolin, in which the mean accuracy was close to 100%, yet the mean %Acc was relatively high at 13%.

**Examples of Differentially Regulated Phosphorylation in Cells of Distinct Sensitivity to Inhibitors**—Our analysis measured phosphorylations at positions that were clearly differentially regulated between cell lines (see Fig. 5A for illustrative examples). Phosphorylations that did not present any variation were also identified (Fig. 5B). More importantly and interestingly, this technique also measured, with high accuracy, phosphorylations showing subtle differences in abundance between samples (examples are given in Fig. 5C). The high

---

**Table III**

Summary of phosphopeptide quantification

|                  | Total number of phosphopeptides | Acc < 30% | R^2 > 0.878 | Acc < 30% and R^2 > 0.878 |
|------------------|---------------------------------|-----------|-------------|--------------------------|
| Total            | 2250                            | 1847      | 528         | 477                      |
| Elevated in P31/Fuj versus Kasumi-1 | 535                              | 412       | 265         | 232                      |
| Elevated in Kasumi-1 versus P31/Fuj  | 338                              | 231       | 130         | 112                      |

---

**Fig. 6.** Correlation between Fold Coefficient (Fc) and fold of change. The relationship between Fold and Fc is shown for all the phosphopeptides identified by LC-MS/MS and those with R^2 > 0.879 only in left and right panels, respectively.

**Fig. 7.** Relationship between accuracy of quantification and ion intensity. The distribution of the percentages of phosphopeptides with different accuracy deviations (%Acc) are shown as a function of cutoff intensity values. Accuracy was increased when the data was filtered to only include phosphopeptides with large intensity values.
correlation values and accuracy of these analyses indicate that these subtle differences (less than 1.5-fold) in phosphorylation between samples are not because of technical variability and are thus real biological differences. Therefore, a strength of the technique is that it can be used to confidently measure subtle differences in phosphorylation extent by providing the researcher with an objective means to assess data quality for each identified phosphopeptide and to assure these small changes are accurately quantified.

**Correlation Between Parameters Describing Data Quality**—Our results confirmed a correlation between fold coefficient and fold change (Fig. 6) indicating that either value may be used to estimate the magnitude of phosphorylation differences between samples. Out of the 2250 phosphopeptides identified in these experiments (with a 2% false discovery rate), 1847 were quantified with an accuracy of 30% or less. We chose this cutoff for accuracy somewhat arbitrarily and other accuracy thresholds can be chosen to match the purpose of the experiment (this is in analogy to choosing an appropriate p value in other statistical analyses).

The proportion of phosphopeptides that were measured with high accuracy (low mean %Acc’s) could be increased by restricting the analysis to phosphopeptides that produced relatively large intensity values. As anticipated, these data show that accuracy of quantification is greater for phosphopeptides showing large intensity values (Fig. 7). But interestingly, even in unfiltered data 80% phosphopeptides were measured with accuracy deviations 30% with a mean of 22%. By restricting the analysis to phosphopeptides with a signal greater than 200,000 the accuracy was increased to 16.9% on average and 87% of these were quantified with accuracy deviations 30%. However, this increase in accuracy was at the expense of quantifying 733 phosphopeptides only.

**Overview of Differentially Regulated Phosphopeptides Between P31/Fuj and Kasumi-1 Cells**—As summarized in Table III, we accurately (with R² > 0.878 and % Acc < 30%) quantified 477 phosphopeptides, of which 232 showed greater intensity in P31/Fuj cells than in Kasumi-1, whereas 112 phosphopeptides were increased in Kasumi-1 cells relative to P31/Fuj (see Supplemental Table 2 for a list of all the identifications). In both cases, we chose twofold as the criteria to consider significance. Fig. 8 shows several examples of phosphorylation sites on kinases and transcription factors.
that were differentially regulated in these cells and thus correlated with the sensitivity of these AML cells to signaling inhibitors. P31/Fuj cells carry an inactivating mutation in the phosphatase PTEN (21), which opposes PI3K signaling. This PTEN mutation may explain the increased phosphorylation of kinases downstream of PI3K (22) such as PDK1 (that is coded by PDPK1) and PKC-β (coded by PRKCB) in P31/Fuj cells relative to Kasumi-1 (Fig. 8). However, other differentially regulated phosphorylations identified in our experiments have not been linked to the PI3K signaling network. It is also noteworthy to observe that the resistant cells (P31/Fuj) showed greater phosphorylation on kinases than the sensitive cells (Kasumi-1) (Fig. 8 and Fig. 9); these observations are consistent with the notion that cells resistant to kinase inhibitors activate several kinase pathways parallel to those being targeted by the compounds (23).

CONCLUSIONS

The technique presented here allows distinguishing phosphorylations that can be quantified by LC-MS in a linear dynamic range with accuracy. The comparison of different methods used to assess data quality indicated that these provided complementary information. About 50% of the phosphopeptides quantified with precision (CV < 50% and fold > 2) were classified as nonproteotypic with our approach. Conversely, several phosphopeptides quantified with linearity and accuracy with our approach showed poor precision of quantification because of the low intensity of their signals in one of the samples. Thus, our approach, rather than substituting for methods to assess data quality based on measuring reproducibility (needed to determine technical and biological variation), it provides additional information on quality data for each identified phosphorylation site.

An important feature of the approach described here is that relative phosphopeptide abundance across samples is calculated concurrently with parameters that quantify accuracy and linearity, thus allowing the researcher to make decisions regarding data quality for each of the thousands of phosphorylation sites that can be quantified by LC-MS. In our experiments, about 8% of the phosphopeptides detected in unfiltered data had accuracy deviations > 50% and these were therefore not quantifiable by our LC-MS technique with acceptable levels of accuracy. These data indicate that many of the detectable phosphopeptides may not always be quantifiable in each experimental setting and illustrate the importance of identifying proteotypic phosphopeptides for accurate quantification.

This is the first time, to the best of our knowledge, that rigorous measures of accuracy and linearity are provided for each phosphorylation site quantified in a global phosphoproteomics experiment and that self-validating parameters are included as part of the technique. In addition to being a powerful tool for accurate phosphoproteomics, the technique may also be used to identify proteotypic phosphopeptides amenable to be quantified using other methods and to quantify proteins and other types of post-translational modifications. Finally, although as a proof-of-principle we have here exemplified the use of the technique by employing label-free LC-MS, the described approach can also be used in combination with stable isotope labeling methods, and can be par-
particularly valuable during the developmental stage in order to validate experimental workflows before embarking in the analysis of experimental samples.

Acknowledgments—We thank Bart Vanhaesebroeck for encouragement and support, Juan Carlos Rodriguez-Prados for discussions, Alex Montoya for technical assistance and other members of the Centre for Cell Signaling for feedback on the manuscript.

* This work was supported by grants from the NIH, BBSRC, Bart’s and the London Charity and the Medical Research Council. § Correspondence should be addressed to: Analytical Signalling Group, Centre for Cell Signalling, Institute of Cancer, Bart’s and the London Medical School, Queen Mary University of London, John Vane Science Centre, Charterhouse Square, United Kingdom. Tel: +44 20 7882 8264; E-mail: p.cutillas@qmul.ac.uk.

In order to cite this article properly, please include all of the following information: Casado, P. and Cutillas, P. R. (2011) A Self-validating Method for Accurate Phosphoproteomics. Mol. Cell. Proteomics 10(1):M110.003079. DOI: 10.1074/mcp.M110.003079.