TITLE:
Motif-centric phosphoproteomics to target kinase-mediated signaling pathways

AUTHOR(S):
Tsai, Chia-Feng; Ogata, Kosuke; Sugiyama, Naoyuki; Ishihama, Yasushi

CITATION:

ISSUE DATE:
2022-01

URL:
http://hdl.handle.net/2433/267848

RIGHT:
© 2021 The Author(s); This is an open access article under the Creative Commons Attribution 4.0 International license.
Motif-centric phosphoproteomics to target kinase-mediated signaling pathways

Successful generation of motif-centric carrier peptides by \textit{in vitro} kinase reactions

Subsequent tandem mass tag quantification targeted to motif-centric phosphopeptides

Successful quantitation of more than 7,000 pY sites from 25 µg of peptides achieved

Multiple kinase-mediated pathways can be targeted for TMT quantitation

Chia-Feng Tsai, Kosuke Ogata, Naoyuki Sugiyama, Yasushi Ishihama

yishiham@pharm.kyoto-u.ac.jp

Tsai et al., 2022, Cell Reports Methods 2, 100138
January 24, 2022 © 2021 The Author(s).
https://doi.org/10.1016/j.crmeth.2021.100138
Motif-centric phosphoproteomics to target kinase-mediated signaling pathways

Chia-Feng Tsai,1,3,4 Kosuke Ogata,1,3 Naoyuki Sugiyama,1 and Yasushi Ishihama1,2,5,*
1Graduate School of Pharmaceutical Sciences, Kyoto University, Kyoto 606-8501, Japan
2Laboratory of Clinical and Analytical Chemistry, National Institute of Biomedical Innovation, Health and Nutrition, Ibaraki, Osaka 567-0085, Japan
3These authors contributed equally
4Present address: Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA 99354, USA
5Lead contact
*Correspondence: yishiham@pharm.kyoto-u.ac.jp
https://doi.org/10.1016/j.crmeth.2021.100138

SUMMARY
Identifying cellular phosphorylation pathways based on kinase-substrate relationships is a critical step to understanding the regulation of physiological functions in cells. Mass spectrometry-based phosphoproteomics workflows have made it possible to comprehensively collect information on individual phosphorylation sites in a variety of samples. However, there is still no generic approach to uncover phosphorylation networks based on kinase-substrate relationships in rare cell populations. Here, we describe a motif-centric phosphoproteomics approach combined with multiplexed isobaric labeling, in which in vitro kinase reactions are used to generate targeted phosphopeptides, which are spiked into one of the isobaric channels to increase detectability. Proof-of-concept experiments demonstrate selective and comprehensive quantification of targeted phosphopeptides by using multiple kinases for motif-centric channels. More than 7,000 tyrosine phosphorylation sites were quantified from several tens of micrograms of starting materials. This approach enables the quantification of multiple phosphorylation pathways under physiological or pathological regulation in a motif-centric manner.

INTRODUCTION
Protein kinase-mediated phosphorylation on serine, threonine, and tyrosine residues is one of the most ubiquitous post-translational modifications (PTMs). Signaling cascades via protein phosphorylation play key roles in multiple cellular processes in mammals, including intra- and intercellular signaling, protein synthesis, gene expression, cell survival, and apoptosis (Cohen, 2002; Hunter, 2000; Needham et al., 2019). The relative abundances of phosphoserine (pS), phosphothreonine (pT), and phosphotyrosine (pY) sites in the human proteome have been estimated to be 90:10:0.05 based on the traditional method of $^{32}$P labeling (Hunter and Sefton, 1980). There are many possible reasons for the extreme paucity of pY sites compared with pS and pT sites in mammals, including the fact that tyrosine kinases are activated only under certain conditions, and that the high activity of protein tyrosine phosphatase leads to a short half-life of pY sites (Hunter, 2009).

Great advances in the analytical workflows of shotgun phosphoproteomics, in which metal affinity chromatography is integrated with liquid chromatography-tandem mass spectrometry (LC-MS/MS), have made it possible to identify more than 30,000 phosphorylation sites (Hogrebe et al., 2018; Humphrey et al., 2015; Mertins et al., 2018). In general, LC-MS/MS has an identification bias toward the more abundant phosphopeptides in a sample, whereas in kinase substrates, sequence features such as Pro-directed, basophilic, acidophilic, and tyrosine-containing motifs have an important influence (Villen et al., 2007).
Also, since biological importance does not necessarily correlate with protein expression levels, it is possible that important signals are transduced via specific kinases that are expressed at extremely low levels. Therefore, advanced pre-fractionation or enrichment methods before LC-MS/MS are needed to identify a wide range of kinase substrates (Tsai et al., 2014). Especially for low-abundance pY peptides, MS detectability is affected by the ionization suppression caused by the presence of more abundant pS and pT peptides in the complex phosphoproteomes. However, the combination of metal affinity chromatography with immunoaffinity purification using a pY antibody (Abe et al., 2017) or a recently developed SH2 domain-derived pY-superbinder (Bian et al., 2016; Dong et al., 2017) has been reported to increase the identification of pY peptides. In addition, immunoaffinity-based methods using multiple antibodies have been developed for the identification and quantitation of phosphopeptides derived from proteins in various pathways or pY peptides derived from tyrosine kinases (Stokes et al., 2012). Nevertheless, a large amount of starting material (1–10 mg) is generally necessary for deep tyrosine phosphoproteome analysis.

One of the major advantages of multiplexed isobaric tandem mass tag (TMT)-based methods for relative quantitation is that the differentially labeled peptides appear as a single peak at the MS1 level (Thompson et al., 2003), enhancing the detectability of low-abundance peptides. TMT strategies using a large amount of relevant “boosting” (or “carrier”) peptides labeled with one or several TMT channels have been successfully used for single-cell proteomics analysis (Budnik et al., 2018; Dou et al., 2019; Tsai et al., 2020). For instance, Yi et al. (2019) developed a Boosting to Amplify Signal with Isobaric Labeling (BASIL) strategy to quantify more than 20,000 phosphorylation sites in human pancreatic islets. However, the identification number of pY peptides was less than 1%. Recently, Chua et al. developed a Broad-spectrum Optimization Of Selective Triggering (BOOST) method, in which pervanadate (a tyrosine phosphatase inhibitor)-treated cells were used as a boosting channel to increase the detectability of pY peptides (Chua et al., 2020). The BOOST method coupled with antibody-based pY enrichment could quantify more than 2,300 unique pY peptides. However, the required amount of starting material was in the milligram range, making it difficult to apply to small samples such as clinical specimens, which often contain less than 100 μg of extractable material.

We previously developed an LC-MS/MS-based in vitro kinase assay using dephosphorylated lysate proteins as the substrate source for in vitro kinase reactions to profile human protein kinomes (Imamura et al., 2014). A total of 175,574 direct kinase substrates were identified from 354 wild-type protein kinases, 21 mutant protein kinases, and 10 lipid kinases (Sugiyama et al., 2019). In addition, we used the in vitro kinase reactions with CK2, MAPK, and EGFR to generate phosphopeptides with targeted motifs to measure the phosphorylation stoichiometry of more than 1,000 phosphorylation sites, including 366 low-abundance tyrosine phosphorylation sites (Tsai et al., 2015).

In the present study, we aimed to develop a motif-centric TMT approach in which phosphopeptides having targeted sequence motifs are generated by in vitro kinase reactions for the boosting TMT channel to increase the detectability of kinase substrates, including tyrosine kinase substrates, without immunoaffinity enrichment. To demonstrate the feasibility of this strategy, phosphopeptides with targeted motifs of CK2, PKA, CDK1, ERK2, JNK1, p38α, SRC, and EGFR were used for the boosting TMT channel to monitor the perturbation of kinase-mediated signaling pathways by tyrosine kinase inhibitor treatment.

RESULTS

Workflow for isobaric motif-centric phosphoproteome analysis
We previously developed a motif-centric approach (Tsai et al., 2015) in which dephosphorylation and isotope tagging are integrated with in vitro kinase reactions to improve the sensitivity and reproducibility for determining the absolute phosphorylation stoichiometry of targeted kinase substrates. However, the number of phosphosites commonly identified in endogenous and motif-centric phosphopeptides was not as large as expected, because some endogenous signals with specific kinase motifs are below the detection limit. In this study, we developed a motif-centric approach in which the same peptides from different samples were labeled with multiplexed TMT reagents and assembled as a single peak at the MS1 level to increase the sensitivity. In addition, we set one of the TMT channels for signal boosting, using phosphopeptides having targeted sequence motifs generated by in vitro kinase reactions to increase the detectability of targeted kinase substrates. The entire workflow is shown in Figure 1. The motif-centric peptides are generated by in vitro kinase reactions using Pro-directed, acidophilic, basophilic, or tyrosine kinase (Figure 1A). The same biological resource (same cell type or tissue) can be used to accomplish the back-phosphorylation (Li et al., 2016; Mundina-Wellenmann et al., 1991) without any pre-dephosphorylation process(Tsai et al., 2015). In vitro kinase reactions both at the protein and tryptic peptide levels can be used in most cases, except for basophilic kinase reactions, where tryptic peptides cannot be used as substrates owing to the lack of K or R at the N-terminal side of the phospho accepting sites. In such a case, the kinase reactions at the protein level, followed by trypic digestion, is used to generate the basophilic motif-centric phosphopeptides. After TMT labeling, phosphopeptides are enriched by immobilized metal ion affinity chromatography (IMAC) and analyzed by nanoLC-MS/MS. We recently reported that TMT-labeled phosphopeptides tend to pass through TiO2 columns (Ogata et al., 2021), and the same phenomenon was also observed in IMAC system (Figure S1). Therefore, the IMAC protocol has been modified to increase the recovery of TMT phosphopeptides by decreasing the concentration of acetic acid in the loading buffer from 6% to 0.5%. We used this modified TMT-IMAC protocol for all experiments unless otherwise noted.

In tandem MS, the TMT-labeled precursor ions from the endogenous and back-phosphorylated peptides are fragmented and the assembled signals in b or y ions are helpful for peptide identification, especially for endogenous phosphopeptides with low abundance (Figure 1B). The relative quantitation of endogenous phosphopeptides across different samples is done by the use of reporter ions at the MS3 level, by the implementation of
Preparation of motif-centric phosphopeptides

(A) The motif-centric peptides were generated by in vitro kinase reactions. Tryptic peptides from study samples and motif-centric peptides were labeled with different isobaric tags (TMT in this study). After mixing, the TMT-labeled phosphopeptides were enriched and analyzed by LC-MS/MS.

(B) TMT-labeled ions from endogenous and motif-centric phosphopeptides are assembled as a single peak at the MS1 level and then fragmented at the MS2 level to identify the peptide sequence. Relative quantification of endogenous phosphopeptides between different samples is achieved by reporter ions at the MS3 level.

Acidophilic motif-centric phosphoproteome analysis

We first used the motif-centric TMT approach to quantify the kinase-perturbed phosphorylation changes in HeLa cells using CK2 kinase inhibitor (CKi, Silmitasertib, CX-4945). The trypptic peptides from CKi-treated HeLa cells were phosphorylated by CK2 in vitro and labeled with both TMT128 and TMT131 for CK2 motif-centric boosting channels. The phosphopeptides from DMSO-treated cells were labeled with both TMT128 and TMT130, and phosphopeptides from CKi-treated cells were labeled with both TMT127 and TMT130. An MS2 spectrum and an MS3 spectrum of a known CK2 substrate, the pS66 site in LIG1, are shown in Figure 2. After fragmention by collision-induced dissociation, the peptide sequence and phosphosite localization information can be annotated at the MS3 level (Figure 2A). Then, the MS3 spectrum was obtained, demonstrating that the TMT signals of endogenous phosphopeptides were decreased after the CKi treatment (TMT127 and TMT130 in Figure 2B) and that the TMT signals of CK2-motif targeting phosphopeptides were increased after in vitro kinase reactions (TMT128 and TMT131 in Figure 2B). These results indicate that the isobaric motif-centric approach can be used to monitor CK2 phosphorylation signaling in terms of the TMT ratios.

In total, this isobaric motif-centric approach quantified 4,559 unique phosphopeptides (91% specificity in phosphopeptide enrichment) from 25 μg of peptides per channel. The logarithm of the median ratio of CKi-treated to untreated peptides was negative, while the logarithm of the median ratio of motif-centric (CK2-motif) to CKi-treated peptides was positive and this increase was larger than the decrease caused by CKi treatment (Figure 2C), indicating that the in vitro kinase reactions effectively increased the TMT signals of peptides directly phosphorylated by CK2. Silmitasertib has been used to exclusively inhibit CK2 activity in previous studies (Chon et al., 2015; Wang et al., 2017). Therefore, we can discriminate direct CK2 substrates from others by comparison of TMT ratios such as CKi/DMSO and CK2-motif/CKi using the Student t test. Based on the CKi/DMSO ratios, 66 (1.6%) up-regulated and 750 (18%) down-regulated phosphorylation sites were identified (Figure 2D). Among the down-regulated phosphorylation sites, up to 35% (n = 266) were significantly increased after in vitro kinase reactions (group B in Figure 2D). After sequence motif analysis (O’Shea et al., 2013), we found that the sequence motif of the phosphopeptides in group B agreed with the known CK2 substrate motif (acidic motif), whereas the motif logo from group A contained both Pro and Asp at the +1 position (Figure 2D). Among the phosphorylation sites (n = 266) in group B, 222 sites were registered in a public phosphorylation sites database (Hornbeck et al., 2015), including six known CK2 substrates: PTGES3 (S113), LIG1 (S66), SLC3A2 (S375), CDK1 (S39), TOP2A (S1377), and ABCF1 (S110). For group B proteins, we also examined the overlap with CK2 interactors in STRING and found 20 known CK2 interactors (Figure 2E). We further performed Gene Ontology and Reactome Pathway enrichment analysis for group B proteins (Szklarczyk et al., 2019). As a result, 137 proteins in group B were annotated as nuclear proteins, and a majority of them possessed functions related to ATP binding, DNA binding, nucleotide binding and so on (Table S1). Among the annotated pathways, the top hit was the cell cycle pathway, in which 21 group B proteins were down-regulated, including CDK1, a known CK2 substrate that affects cell cycle regulation upon S39 phosphorylation (Bloom and Cross, 2007; Russo et al., 1992) (Figure 2F). All these results indicate that this motif-centric approach is an effective tool for monitoring specific kinase-mediated signaling pathways.

Basophilic motif-centric phosphoproteome analysis

The isobaric motif-centric approach was further examined with PKA as a basophilic kinase, using forskolin as an activator. Unlike acidophilic CK2, PKA cannot be used to generate
motif-centric peptides from tryptic peptides because it requires basic amino acid residues at the N-terminal side of the phospho-acceptor. Therefore, PKA motif-centric peptides were generated by in vitro kinase reactions at the protein level, followed by tryptic digestion (Figure 1A). After the isobaric labeling, the TMT ratios of three channels, forskolin-treated, DMSO-treated, and motif-centric samples, were used to discriminate the peptides phosphorylated directly by PKA from others, as in the case of CK2. The results are illustrated in Figure S2. In total, we quantified 7,855 phosphorylation sites from biological replicate experiments. Among them (Figure S2A), 620 phosphorylation sites were up-regulated (group 1) and 322 phosphorylation sites were inhibited by forskolin (group 2). We previously established a computational model (primary sequence preference [PSP score]) to characterize the kinase sequence specificity toward the substrate target site based on known kinase-substrate relationships, to exclude potential indirect targets of PKA (Imamura et al., 2017). We applied PSP scoring to the phosphorylation sites in groups 1 and 2, and found that the PSP scores of group 1 were significantly higher than those of group 2, although the sequence motifs of phosphorylation sites in groups 1 and 2 both belonged to the basophilic category (Figure S2B). Based on the above results, the motif-centric approach was able to discriminate peptides phosphorylated directly by PKA from indirectly phosphorylated ones.

**Tyrosine motif-centric phosphoproteome analysis**

Although the detectability of pY sites has been restricted by their extremely low phosphorylation stoichiometry compared with pS and pT sites (Sharma et al., 2014; Tsai et al., 2015), this lower stoichiometry results in a larger amount of unphosphorylated counterparts, which can be used for back-phosphorylation. To increase the identification number of pY sites, we firstly tried to use pervanadate (PV)-treated HeLa cells as a pY-centric sample (Figure S3A). PV is well-known as a tyrosine phosphatase inhibitor causing an increase in the stoichiometry of endogenous pY sites (Sharma et al., 2014). As expected, the pY content was increased from 0.8% to 11.7% (Figure S3B) by PV treatment. Then we used the PV-treated peptides as pY-centric peptides with 25 µg of untreated HeLa peptides. Indeed, the number of quantifiable pY sites increased from 74 to 595 as the spiking amount of the pY-centric peptides was increased from 0 to 150 µg, without any antibody-based enrichment (Figure S3B).

The amplified signals caused by the spiked pY peptides resulted in an even higher detection efficiency of pY peptide sites compared to pS and pT peptide sites (Figure S3C). With this approach, we were able to detect 309 additional pY peptide sites with a 10-fold increase in the amount of pY-centric peptides. This indicates that the motif-centric phosphoproteome analysis provides an efficient strategy for the comprehensive detection of pY peptide sites.
in a much larger number of identified pY sites than in usual phosphoproteome analysis. However, the number of identified pY sites was still lower than that obtained using the recently published antibody-based boosting strategy (Chua et al., 2020), which also used PV-treated cells as boosting samples to quantify more than 2,300 unique pY peptides from 1 mg of starting materials.

To extend the tyrosine phosphoproteome coverage, the back-phosphorylation sites of the untreated HeLa peptides were phosphorylated via in vitro kinase reactions by using tyrosine kinases such as EGFR and SRC. Note that EGFR and SRC have different phosphorylation motifs (Sugiyama et al., 2019). Then, the kinase-treated peptides were labeled with one of the 6-plexed TMT reagents (each TMT channel contained 25 μg tryptic peptides), to detect the endogenous pY sites in EGF or EGF/afatinib (EGFR inhibitor)-treated HeLa cells (Figure S3A). After TMT labeling followed by IMAC enrichment, the TMT-labeled endogenous phosphopeptides mixed with PV treated, EGFR-centric or SRC-centric phosphopeptides were analyzed on a 2-m-long monolithic silica column system with the SPS-MS3 technique (McAlister et al., 2014). Compared with the result from PV-treated HeLa as pY motif-centric peptides, the content of pY increased from 15% to 87% (Figure 3A), which resulted in an approximately 10-fold increase in the numbers of identified and quantified class 1 pY sites (localization probability >0.75), as shown in Figure 3B. However, Cheung et al. (2021) reported that the isobaric labeling-based quantitative approaches have technical limitations that potentially affect data quality and biological interpretation, owing to the large amounts of spiked carrier samples. In addition, it is difficult to control the TMT ratio within the quantifiable range because the phosphorylation stoichiometry in cells depends on each pY site. Furthermore, we should reject TMT peptides without reporter ion signals in the sample channels. Therefore, we examined the distribution of the TMT reporter ion intensity of each sample channel and found a notch to discriminate the signal from the noise (Figure S3C); this was also mentioned in the previous study (Hughes et al., 2017). Based on this observation, we set the acceptance criterion for the minimum TMT intensity in the sample channels to be greater than 40 on a log2 scale for the total TMT intensity of the sample channels. We also set another criterion—that the maximum TMT ratio of the motif-centric channel to the sample channel should be less than 100 (Figure 3C)—to minimize the ion sampling effect (Cheung et al., 2021; Tsai et al., 2020). By
assuming these two criteria, we found that 668, 5,425, and 3,688 pY sites could be quantified in the PV-treated, EGFR-centric, and SRC-centric samples, respectively. Among these pY sites, the sequence motifs of unique pY sites in each dataset were different (Figure 3D), which indicated that complementary pY sites can be identified by using different motif-centric peptides. For quantitation based on the TMT reporter ion intensities, the EGF and EGF-afatinib treated cells were separated in the principal component analysis (Figure 3E). The above results demonstrate that this motif-centric approach through spiking isobaric pY peptides with high purity was able to increase the detectability of pY sites without the need for pY-specific affinity purification, such as using pY antibodies.

We further evaluated the spiking effect of the carrier/boost amount of motif-centric peptides on the number of quantifiable phosphosites. Different amounts of TMT-labeled peptides (5, 10, 25, and 60 μg) from untreated HeLa cells were mixed with different amounts (5, 25, and 125 μg) of pY motif (EGFR)-centric peptides (Figure 4A). In these experiments, we applied a nanoscale solid phase TMT labeling protocol to phosphopeptides enriched by TiO₂ chromatography to accommodate smaller amounts of peptides (Ogata et al., 2021). The distribution of the quantifiable pY site fraction for different boost ratios is shown in Figure 4B. When we applied the acceptance criterion that the boost ratio should be less than 100, a higher fraction of pY sites was obtained as we increased the amount of endogenous peptides with the constant amount of motif-centric peptides. The same tendency was observed for the number of quantifiable pY sites (Figure 4C). As the amount of the endogenous peptides increased, the number of quantifiable pY sites increased linearly, regardless of the amount of motif-centric peptides. In contrast, when the amount of the endogenous peptides was limited, the number of quantifiable pY sites was saturated or even decreased with increasing amounts of motif-centric peptides. Based on these results, in this study, the peptide amount per TMT channel was set to 25 μg for both endogenous and motif-centric channels.

Multiple motifs-centric approach to depict perturbed phosphoproteome

We further used multiple kinases to generate a wide variety of motif-centric peptides to monitor the phosphorylation signals generated by different kinases. For proof-of-concept, we analyzed the phosphoproteome of EGF-treated and EGF/afatinib co-treated HeLa cells. We selected four Pro-directed kinases (ERK1 [MAPK3], JNK1 [MAPK8], p38α [MAPK14] and CDK1), two tyrosine kinases (SRC and EGFR), and one acidophilic kinase (CK2) to detect the different motif-centric phosphorylation sites (Figure S4A). After applying the following acceptance criteria—the boost ratio should be less than 100, the total TMT intensity of the sample channels must be greater than 40 on a log₂ scale, and two valid TMT values must be obtained in at least one of duplicate channels—up to 11,895 class 1 phosphorylation sites were quantified, including 5,045 pS, 1,756 pT, and 5,094 pY sites (Figure 5A). Of the 11,895 quantifiable class 1 phosphorylation sites, up to 8,482 were quantified as motif-centric phosphopeptides, indicating that endogenous peptides that are not motif-centric do not interfere with the quantification of the target phosphopeptides under the current conditions. According to the Student t test (EGF vs. EGF/afatinib), the ratio of commonly regulated phosphorylation sites was consistent between experiments using different kinases, regardless of the kinase used (Figure 5B), indicating that the boosting channel did not significantly affect the quantitative results.

In addition to identifying which phosphorylation sites were regulated after afatinib treatment, we were also able to estimate which kinases phosphorylated which sites by using this motif-centric approach. For example, up to 18 class 1 phosphorylation sites, including 7 pY sites, on EGFR kinase were quantified without any immunoprecipitation step from 25 μg of starting materials per TMT channel (Figure 5C). Among these sites, pY1144 is a known autophosphorylation site which was down-regulated by afatinib and detected in the EGFR motif-centric experiment. This result indicates that the activity of this phosphorylation site is controlled by EGFR and also inhibited by afatinib. In contrast, pS1051 and pS992 were down-regulated by afatinib, but there is no information about the corresponding kinase in the public database. Through this motif-centric approach, we could determine that pS1051 and pS992 are likely phosphorylated by CK2 and ERK2, respectively.

Another advantage of the motif-centric approach is the specificity of the in vitro kinase reactions. The Student t test (EGF/afatinib vs. motif-centric) revealed a low overlap of motif-centric peptides among the kinases (Figure 5D). Although ERK2, JNK1, p38α, and CDK1 are all Pro-directed kinases, complementary profiles for the quantified motif-centric phosphorylation sites were observed (Figure 5D). The motif logos of these sites showed slight differences among the four Pro-directed kinases (Figure S4B). For example, the proportion of pT motifs was higher with p38α kinase. In addition, more acidic amino acids were located at the C-terminal side of the phospho-acceptor site in the case of ERK2 kinase compared with CDK1 kinase. In addition to S/T sites, the pY motifs also differed between EGFR- and SRC-centric phosphopeptides (Figure S4C).

Furthermore, from the two ratios (EGF vs. motif-centric and EGF/afatinib vs. motif-centric), we can estimate the stoichiometry of motif-centric phosphorylation sites (Figure 5E), assuming that the efficiency of in vitro kinase reactions is 100%. The phosphorylation sites containing acidophilic kinase substrates targeted by CK2 generally exhibited higher phosphorylation stoichiometry than sites targeted by Pro-directed kinase and tyrosine kinases. The phosphorylation stoichiometry distribution that we observed here is consistent with our previous findings (Tsai et al., 2015). To further validate that the endogenous signals boosted by motif-centric peptides are lower than other peptide signals, the peak areas in the XICs of MS1 signals of quantified phosphopeptides were calculated based on the proportion of TMT intensity (Figures S4D and S4E). The XICs of endogenous phosphopeptides boosted by the motif-centric peptides were lower than those of other phosphopeptides, which indicates the motif-centric approach is effective to identify these low-abundance kinase substrates.

In this motif-centric approach, a specific kinase can be chosen to target an endogenous peptide that is phosphorylated by that kinase. To confirm this, 8,482 motif-centric phosphopeptides prepared with these seven kinases were subjected to KEGG

Multiple motifs-centric approach to depict perturbed phosphoproteome

We further used multiple kinases to generate a wide variety of motif-centric peptides to monitor the phosphorylation signals generated by different kinases. For proof-of-concept, we analyzed the phosphoproteome of EGF-treated and EGF/afatinib co-treated HeLa cells. We selected four Pro-directed kinases (ERK1 [MAPK3], JNK1 [MAPK8], p38α [MAPK14] and CDK1), two tyrosine kinases (SRC and EGFR), and one acidophilic kinase (CK2) to detect the different motif-centric phosphorylation sites (Figure S4A). After applying the following acceptance criteria—the boost ratio should be less than 100, the total TMT intensity of the sample channels must be greater than 40 on a log₂ scale, and two valid TMT values must be obtained in at least one of duplicate channels—up to 11,895 class 1 phosphorylation sites were quantified, including 5,045 pS, 1,756 pT, and 5,094 pY sites (Figure 5A). Of the 11,895 quantifiable class 1 phosphorylation sites, up to 8,482 were quantified as motif-centric phosphopeptides, indicating that endogenous peptides that are not motif-centric do not interfere with the quantification of the target phosphopeptides under the current conditions. According to the Student t test (EGF vs. EGF/afatinib), the ratio of commonly regulated phosphorylation sites was consistent between experiments using different kinases, regardless of the kinase used (Figure 5B), indicating that the boosting channel did not significantly affect the quantitative results.

In addition to identifying which phosphorylation sites were regulated after afatinib treatment, we were also able to estimate which kinases phosphorylated which sites by using this motif-centric approach. For example, up to 18 class 1 phosphorylation sites, including 7 pY sites, on EGFR kinase were quantified without any immunoprecipitation step from 25 μg of starting materials per TMT channel (Figure 5C). Among these sites, pY1144 is a known autophosphorylation site which was down-regulated by afatinib and detected in the EGFR motif-centric experiment. This result indicates that the activity of this phosphorylation site is controlled by EGFR and also inhibited by afatinib. In contrast, pS1051 and pS992 were down-regulated by afatinib, but there is no information about the corresponding kinase in the public database. Through this motif-centric approach, we could determine that pS1051 and pS992 are likely phosphorylated by CK2 and ERK2, respectively.

Another advantage of the motif-centric approach is the specificity of the in vitro kinase reactions. The Student t test (EGF/afatinib vs. motif-centric) revealed a low overlap of motif-centric peptides among the kinases (Figure 5D). Although ERK2, JNK1, p38α, and CDK1 are all Pro-directed kinases, complementary profiles for the quantified motif-centric phosphorylation sites were observed (Figure 5D). The motif logos of these sites showed slight differences among the four Pro-directed kinases (Figure S4B). For example, the proportion of pT motifs was higher with p38α kinase. In addition, more acidic amino acids were located at the C-terminal side of the phospho-acceptor site in the case of ERK2 kinase compared with CDK1 kinase. In addition to S/T sites, the pY motifs also differed between EGFR- and SRC-centric phosphopeptides (Figure S4C).

Furthermore, from the two ratios (EGF vs. motif-centric and EGF/afatinib vs. motif-centric), we can estimate the stoichiometry of motif-centric phosphorylation sites (Figure 5E), assuming that the efficiency of in vitro kinase reactions is 100%. The phosphorylation sites containing acidophilic kinase substrates targeted by CK2 generally exhibited higher phosphorylation stoichiometry than sites targeted by Pro-directed kinase and tyrosine kinases. The phosphorylation stoichiometry distribution that we observed here is consistent with our previous findings (Tsai et al., 2015). To further validate that the endogenous signals boosted by motif-centric peptides are lower than other peptide signals, the peak areas in the XICs of MS1 signals of quantified phosphopeptides were calculated based on the proportion of TMT intensity (Figures S4D and S4E). The XICs of endogenous phosphopeptides boosted by the motif-centric peptides were lower than those of other phosphopeptides, which indicates the motif-centric approach is effective to identify these low-abundance kinase substrates.

In this motif-centric approach, a specific kinase can be chosen to target an endogenous peptide that is phosphorylated by that kinase. To confirm this, 8,482 motif-centric phosphopeptides prepared with these seven kinases were subjected to KEGG
Figure 4. Effects of the boost ratio of spiking motif-centric peptides on quantifiable pY sites
(A) Different amount of TMT labeled peptides from un-treated HeLa cell (5, 10, 25, and 60 μg) were mixed with different amount (5, 25, and 125 μg) of pY motif (EGFR)-centric peptides. After phosphopeptides enrichment, the TMT-labeled phosphopeptides were analyzed by LC-MS/MS.
(B) The distribution of the quantifiable pY site fraction for different boosting ratios.
(C) The number of detected pY sites under different amount of motif-centric peptides (boost ratio <100x). Triplicate sample preparation of 4 different amount of HeLa peptides with three different amounts of pY motif-centric peptides was performed and triplicate LC-MS/MS analyses were performed for each TMT set as shown in Table S3.
pathway enrichment analysis using DAVID (Huang da et al., 2008). As a result, known pathways involving the seven kinases, such as ErbB and insulin signaling, were enriched as expected (Figure 6A and Table S2). We then mapped the responsible kinases for the in vitro phosphorylation sites on the 33 phosphoproteins comprising the ErbB pathway identified in this enrichment analysis (Figure 6B), showing how the kinases used for motif-centric peptides covered the targeted phosphoproteins within the targeted pathway. In other words, we can manipulate the targeted pathway by choosing the appropriate kinases, using our large-scale library of in vitro kinase-substrate relationships (Sugiyama et al., 2019). Overall, our findings indicate that the motif-centric approach can provide system-wide customizable maps consisting of targeted pathways under physiological or pathological regulation.

**DISCUSSION**

Previously, the kinase assay-linked phosphoproteomics approach (Xue et al., 2012, 2013) has been developed to find potential kinase substrates which were found in the overlap of the in vitro phosphopeptides generated by kinase reactions with dephosphorylated peptides and in vivo kinase-dependent phosphorylation events in different LC-MS/MS runs. However, the dephosphorylation is not complete, and this may result in the false-positive identification of putative kinase substrates. In addition, the analytical throughput is still limited by the need for multiple LC-MS/MS analysis, and the background effect owing to incompletely dephosphorylated peptides decreases the detection sensitivity for putative kinase substrates. To tackle these challenges, the isobaric motif-centric approach developed in this study can link in vitro substrates and physiological phosphorylation events by monitoring both endogenous and back (specific kinase motif-targeting) phosphorylated signals in a single LC-MS/MS run without the need for a further dephosphorylation step, thereby enabling the high-throughput analysis of putative kinase substrates.

Unlike metal affinity chromatography for enrichment of phosphopeptides, motif-specific immunoaffinity precipitation (IAP)-based LC-MS/MS makes it possible to recognize a characteristic sequence motif from a broad range of peptides by using different motif antibodies. Because each antibody binds phosphopeptides followed by a specific peptide motif, the overlap of identified phosphopeptides among different antibodies is low. The results of this approach are similar to those obtained with our motif targeting approach (Figure 4D), in which recognition between kinase and peptides is based on the specific sequence motif (Stokes et al., 2015; Sugiyama et al., 2019). The overlap between a given antibody and metal affinity chromatography ranged from roughly 16% with the all Ser/Thr antibody mix, to a low of only 3.6% with pY-1000 (Stokes et al., 2015).
Anthony et al. aimed to enlarge the phosphoproteome coverage by using both TiO₂ followed by basic pH reversed-phase fractionation and motif-specific IAP with four different phosphorylation motif-specific antibodies (Possemato et al., 2017). In total, 8,947 nonredundant peptides were identified in the TiO₂ dataset, of which only 852 (9.5%) were in common with the peptides identified in the IAPs. These results suggest that the range of phosphorylation within a given system is so broad that no single approach is likely to provide comprehensive coverage.

In the case of IAPs, the specificity of the antibodies is not high enough to distinguish phosphopeptides with similar motifs; for example, MAPK phosphorylates substrates with the consensus sequence PX(S/T)P and CDKs phosphorylate substrates containing the consensus sequence (S/T)PXR/K (Songyang et al., 1996). However, the complementary regulated motif-centric phosphorylation sites (Figures 5D and S4B) were differentiated in this study, even though the kinases all belong to the same Pro-directed kinase group. The motif-centric approach was also able to distinguish the motif difference (PX(S/T)P and (S/T)PXR/K) for ERK2 and CDK1 (Figure 5D). In addition, the purification specificity in IAP (<50%) is much lower than that in the metal affinity-based approach (Possemato et al., 2017), which means that milligram amounts of starting materials are necessary for the IAPs method. For our isobaric motif-centric approach, we obtained a purification specificity of more than 90% (Figure S4F). In addition, the IAPs approach uses selected antibodies to isolate targeting peptides, and this may cause unnecessary sample loss during the purification step. In contrast, the motif-centric approach uses kinase to recognize specific substrates and transfer the phosphate group to back-phosphorylated peptides. Endogenous phosphopeptides, especially low-abundance tyrosine phosphopeptides, are not removed before IMAC enrichment. Therefore, only a few tens of micrograms of material was required for our isobaric motif-centric approach.

Recently, a novel antibody-based method, called PTMScan Direct, was developed for the identification and quantitation of peptides derived from proteins that are critical signaling nodes of various pathways (Stokes et al., 2012). However, the coverage of the IAPs approach is still limited by the availability and quality of antibodies. In contrast, recombinant active kinases are much more readily available than antibodies. In our previous study (Sugiyama et al., 2019), we were able to identify a total of 175,574 potential direct kinase substrates by using 385 active kinases (354 wild-type protein kinases, 21 mutants, and 10 lipid kinases). Based on this kinase substrate library, it is easily possible to select multiple kinases for targeting pathway analysis.

The multiplexing nature of isobaric labeling is particularly useful to achieve greater sensitivity with limited individual sample amounts, as in the case of tyrosine phosphopeptides. Chua et al. (2020) and Fang et al. (2020) used samples with and without PV (tyrosine phosphatase inhibitor)-treated cells as a boosting channel to increase the relative abundance of tyrosine phosphorylation sites. Although approximately 2,300 (Chua et al., 2020) and 835 pY phosphopeptides (Fang et al., 2020) were detected, the required amount of starting material is at the milligram level. In addition, the use of antibodies for pY phosphopeptides enrichment is also necessary for their approach. Here, in contrast, we generate higher purity tyrosine phosphopeptides via in vitro kinase reactions (motif-centric). By using motif-centric peptides, we could detect up to 7,129 (SRC targeting) and 7,280 (EGFR targeting) tyrosine phosphopeptides without the need for immunoprecipitation, using only a few tens of µg of starting materials.
Conclusions
The isobaric motif-centric strategy presented here can be used to enhance the sensitivity of specific kinase downstream signaling analysis, especially for tyrosine phosphopeptides. It provides a simple yet highly effective quantitative phosphoproteomic workflow suitable for multiplexed analysis of relatively small biological or clinical samples (less than milligram size), including cells or tissues. This approach enables the quantitation of both fold-change and stoichiometry among thousands of phosphopeptides generated by specific kinases in signaling pathways. The use of multiple kinases for motif targeting analysis effectively increases the phosphoproteome coverage. Overall, we anticipate this strategy should find broad biomedical applications for targeting kinase/pathway analysis where limited amounts of starting cells or tissues are available.

Limitations of the study
Although the performance of isobaric labeling-based quantitative approaches is affected by co-selected precursor ions, the use of SPS-MS\(^3\) can minimize this effect and improve the detection sensitivity by using co-fragmented multiple (≤10) MS\(^2\) fragment ions with higher intensity (McAlister et al., 2014). Moreover, the newly available Real Time Search-MS\(^3\) method (RTS-MS\(^3\)) (Erickson et al., 2019) or ion mobility technique (Bekker-Jensen et al., 2020; Hebert et al., 2018; Ogata and Ishihama, 2020) provides a solution for precise and accurate quantitation without sacrificing proteome coverage. Recently, the use of spiked TMT labeling peptides with boosting or carrier samples has decreased the TMT reporter ions dynamic range and decreased the quantitation accuracy (Tsai et al., 2020). Cheung et al. (2021) and Stopfer et al. (2021a) also demonstrated that an increase in carrier proteome level requires a concomitant increase in the number of ions sampled to maintain quantitative accuracy. Therefore, it will be important to optimize the spiking amount of carrier/boosting and the TMT channel design for the isobaric motif-centric strategy. As shown in Figure S7, the spiking amount of carrier/boosting affected the number of quantifiable pY sites (boost ratio <100). In addition, optimization of the MS parameters to improve the ion sampling by adjusting the ion injection time and AGC will also be beneficial to improve the quantitation performance (Cheung et al., 2021; Tsai et al., 2020). To overcome the challenge presented by the large quantitation dynamic range, it may be useful to integrate the motif-centric approach with isotope-based SureQuant (Stopfer et al., 2021b) or Internal Standard Triggered-Parallel Reaction Monitoring (Gallien et al., 2015) quantitation.

For the phosphorylation stoichiometry, it was assumed that the kinase reaction efficiency was 100%. Similar to our previous motif-targeting approach for measuring phosphorylation stoichiometry (Tsai et al., 2015), we used a high concentration of ATP (1 mM) and a long incubation time (overnight) to maximize the efficiency of the reaction. However, it is not always possible to achieve 100% kinase reaction efficiency for all sites. If the reaction efficiency is not 100%, then the measured stoichiometry will be overestimated.

Although we used a single kinase to produce the motif-centric peptides for each TMT set, it is possible to pool multiple kinase products to increase the throughput of this analysis. However, before applying this method to real samples, it is necessary to optimize how many and what kinases can be combined to avoid excessive increase in sample complexity and dynamic range. Otherwise, the quantification of peptides with low amounts would be problematic.

STAR METHODS
Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
- **METHOD DETAILS**
  - Tryptic peptides from HeLa cell lysate
  - In vitro kinase reactions
  - TMT labeling for digested peptides
  - IMAC
  - LC-MS/MS analysis
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

SUPPLEMENTAL INFORMATION
Supplemental information can be found online at https://doi.org/10.1016/j.crmeth.2021.100138.

ACKNOWLEDGMENTS
We would like to thank all lab members for fruitful discussions. C.-F.T. was supported by a JSPS Grant-in-Aid for a postdoctoral fellowship for overseas researchers (15F15343). This work was supported by the JST Strategic Basic Research Program, CREST (grant No. 18070870) to Y.I., and by a JSPS Grants-in-Aid for Scientific Research (No. 17H03605, No. 21H02459 to Y.I., No. 20K21478 to Y.I. and K.O., and No. 20H04845, No. 21H02466 to N.S.)

AUTHOR CONTRIBUTIONS

DECLARATION OF INTERESTS
The authors declare no competing interests.

Received: July 1, 2021
Revised: October 8, 2021
Accepted: December 13, 2021
Published: January 14, 2022

REFERENCES


STAR METHODS

KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chemicals, peptides, and recombinant proteins</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Triethylammonium bicarbonate</td>
<td>Sigma</td>
<td>Catalog: T7408</td>
</tr>
<tr>
<td>Phosphatase Inhibitor Cocktail 2</td>
<td>Sigma</td>
<td>Catalog: P0044</td>
</tr>
<tr>
<td>Phosphatase Inhibitor Cocktail 3</td>
<td>Sigma</td>
<td>Catalog: P5726</td>
</tr>
<tr>
<td>BCA Protein Assay Kit</td>
<td>Thermo Scientific Pierce</td>
<td>Catalog: 23225</td>
</tr>
<tr>
<td>Sodium deoxycholate</td>
<td>FUJIFILM Wako</td>
<td>Catalog: 190-08313</td>
</tr>
<tr>
<td>Sodium lauroyl sarcosinate</td>
<td>FUJIFILM Wako</td>
<td>Catalog: 198-14745</td>
</tr>
<tr>
<td>Iron-(III) chloride</td>
<td>FUJIFILM Wako</td>
<td>Catalog: 091-00872</td>
</tr>
<tr>
<td>Dithiothreitol</td>
<td>Thermo Scientific</td>
<td>Catalog: 20291</td>
</tr>
<tr>
<td>Iodoacetamide</td>
<td>Thermo Scientific</td>
<td>Catalog: A3221</td>
</tr>
<tr>
<td>Lysyl endopeptidase</td>
<td>FUJIFILM Wako</td>
<td>Catalog: 129-02541</td>
</tr>
<tr>
<td>Sequencing-grade modified trypsin</td>
<td>Promega</td>
<td>Catalog: V517</td>
</tr>
<tr>
<td>Ni-NTA silica resins</td>
<td>QIAGEN</td>
<td>Catalog: 31314</td>
</tr>
<tr>
<td>Empore SDB-XC membrane disks</td>
<td>CDS</td>
<td>Catalog: 13-110-020</td>
</tr>
<tr>
<td>Titanium dioxide (10 μm)</td>
<td>GL Sciences</td>
<td>Catalog: 5020-75010</td>
</tr>
<tr>
<td>CK2α2/b (CSNK2A2/B)</td>
<td>Carna Biosciences</td>
<td>Catalog: 05-185</td>
</tr>
<tr>
<td>PKAc (PRKACA)</td>
<td>Carna Biosciences</td>
<td>Catalog: 01-127</td>
</tr>
<tr>
<td>ERK2 (MAPK1)</td>
<td>Carna Biosciences</td>
<td>Catalog: 04-143</td>
</tr>
<tr>
<td>EGFR (ERBB1)</td>
<td>Carna Biosciences</td>
<td>Catalog: 08-115</td>
</tr>
<tr>
<td>SRC</td>
<td>Carna Biosciences</td>
<td>Catalog: 08-173</td>
</tr>
<tr>
<td>JNK1 (MAPK8)</td>
<td>Carna Biosciences</td>
<td>Catalog: 04-163</td>
</tr>
<tr>
<td>CDK1 (CDC2/CycB1)</td>
<td>Carna Biosciences</td>
<td>Catalog: 04-102</td>
</tr>
<tr>
<td>p38α (MAPK14)</td>
<td>Carna Biosciences</td>
<td>Catalog: 04-152</td>
</tr>
<tr>
<td>TMTsixplex™</td>
<td>Thermo Scientific</td>
<td>Catalog: 90061</td>
</tr>
<tr>
<td><strong>Deposited data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Experimental models: Cell lines</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HeLa S3</td>
<td>ATCC</td>
<td>Catalog: CCL-2.2</td>
</tr>
<tr>
<td><strong>Software and algorithms</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MaxQuant</td>
<td>PMID: 27809316</td>
<td><a href="https://www.maxquant.org/">https://www.maxquant.org/</a></td>
</tr>
<tr>
<td>Perseus</td>
<td>PMID: 27348712</td>
<td><a href="https://maxquant.net/perseus/">https://maxquant.net/perseus/</a></td>
</tr>
</tbody>
</table>

RESOURCE AVAILABILITY

**Lead contact**
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Yasushi Ishihama (yishiham@pharm.kyoto-u.ac.jp).

**Materials availability**
This study did not generate new unique reagents.

**Data and code availability**
- Data described in this paper have been deposited at https://zenodo.org and are publicly available as of the date of publication. DOIs are listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.
EXPERIMENTAL MODEL AND SUBJECT DETAILS

HeLa S3 cells were cultured in DMEM containing 10% fetal bovine serum and 100 µg/mL kanamycin. For isobaric acidophilic motif-centric phosphoproteomes, cells were not stimulated (mock) or were stimulated with 10 mM CK2 inhibitor (CX-494S) for 30 min. For isobaric basophilic motif-centric phosphoproteomes, cells were not stimulated (mock) or were stimulated with 10 µM PKA activator (forskolin) for 30 min. For isobaric tyrosine and multiple motif-centric phosphoproteomes, cells were treated with 10 µM EGF, 10 µM EGFr/10 µM afatinib, and 500 µM PV (pH 10 with 0.14% H2O2), respectively, for 30 min before harvesting. Two biological replicates were performed.

METHOD DETAILS

Tryptic peptides from HeLa cell lysate

Cells were washed three times with ice-cold phosphate-buffered saline (phosphate-buffered saline, 0.01 M sodium phosphate, 0.14 M NaCl, pH 7.4) and then lysed in lysis buffer containing 12 mM sodium deoxycholate, 12 mM sodium lauroyl sarcosinate in 100 mM triethanolammonium bicarbonate. Protein concentration was determined by means of BCA protein assay. The lysates were digested based on the reported phase-transfer surfactants protocol (Masuda et al., 2008). The digested peptides were desalted on SDB-XC StageTips (Rappsilber et al., 2007).

In vitro kinase reactions

For acidophilic, Pro-directed and tyrosine kinase reactions, the trypic peptides were dissolved in 40 mM Tris-HCl (pH 7.5) and incubated with each kinase (0.2 µg CK2, ERK2, JNK1, p38α, CDK1 or SRC) at 37 °C overnight for in vitro kinase reactions in the presence of 1 mM ATP and 20 mM MgCl2. For the EGFR kinase reactions, tryptic peptides were firstly passed through SCX StageTips (Rappsilber et al., 2007) to remove afatinib. Eluted peptides were further desalted on SDB-XC StageTips. Then, the desalted peptides were dissolved in 40 mM Tris-HCl (pH 7.5) and incubated with EGFR (0.2 µg) for in vitro kinase reactions in the presence of 1 mM ATP and 4 mM MnCl2 at 37 °C overnight. For basophilic kinases such as PKA, the lysates were loaded onto a 10-kDa ultrafiltration device (Amicon Ultra, Millipore). The device was centrifuged at 14,000 g to remove the detergents. Subsequently, the original lysis buffer was replaced with 40 mM Tris-HCl (pH 7.5) followed by centrifugation. Then, the proteins were incubated with 0.2 µg PKA for in vitro kinase reactions in the presence of 1 mM ATP and 20 mM MgCl2 at 37 °C overnight. After the kinase reactions, the proteins were reduced with 10 mM DTT for 30 min at 37 °C and alkylated with 50 mM iodoacetamide in the dark for 30 min at 37 °C. The resulting samples were digested by Lys-C (1:100, w/w) at 37 °C for 3 h followed by trypsin (1:50, w/w) overnight at 37 °C. All the peptides were desalted on SDB-XC StageTips.

TMT labeling for digested peptides

The desalted peptides were dissolved in 200 mM HEPES (pH 8.5). Then, the resuspended digested peptides were mixed with TMT reagent dissolved in 100% ACN for 1 h. The labeling reaction was stopped by adding 5% hydroxyamine for 15 min, followed by acidification with TFA. All the peptides labeled with each multiplexed TMT reagent were mixed into the same tube and the mixture was diluted to decrease the concentration of ACN to less than 5%. The TMT-labeled peptides were desalted on SDB-XC StageTips. The information on the peptide amount in each TMT channel for all experiments is shown in Table S3. Note that the peptide amount for each TMT channel was quantified by means of nanoLC-UV at 210 nm using a Thermo Ultimate 3000 RSLCnano system (Germering), an MU701 UV detector (GL Sciences), and a C18 analytical column (150 mm length × 100 µm ID) packed with Reprosil-Pur 120 C18-AQ material (3 µm, Dr. Maisch).

IMAC

The procedure for phosphopeptides purification with an Fe3+-IMAC tip was as described previously (Tsai et al., 2014, 2015) with minor modifications. In brief, a buffer consisting of 50 mM EDTA in 1 M NaCl was used for removing Ni2+ ions. Then, the metal-free NTA was activated by loading 100 mM FeCl3 into the IMAC tip. The Fe3+-IMAC tip was equilibrated with 0.5% (v/v) acetic acid at pH 3.0 before sample loading. Tryptic peptides from HeLa lysates were reconstituted in 0.5% (v/v) acetic acid and loaded onto the IMAC tip. For acidophilic, Pro-directed and tyrosine kinase reactions in the presence of 1 mM ATP and 20 mM MgCl2, the IMAC tip was coupled to an activated SDB-XC StageTip and the bound phosphopeptides were eluted onto the SDB-XC StageTip with 200 mM NH4H2PO4 buffer. Then, the eluted phosphopeptides were desalted with SDB-XC StageTip.

LC-MS/MS analysis

NanoLC-MS/MS analyses were performed on an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Scientific), which was connected to the Thermo Ultimate 3000 RSLCnano system and an HTC-PAL autosampler (CTC Analytics). Peptide mixtures were loaded onto and separated on self-pulled needle columns (150 mm length × 100 µm inner diameter) packed with Reprosil-Pur 120 C18-AQ material (3 µm) or a 2-m-long C18 monolithic silica capillary column (Iwasaki et al., 2010). The mobile phases consisted of (A) 0.5% acetic acid and (B) 0.5% acetic acid and 80% acetonitrile. Peptides were separated through a gradient from 17.5% to 45% buffer B at a flow rate of 500 nL/min. Full-scan spectra were acquired at a target value of 4 × 106 with a resolution of 60,000. Data
were acquired in a data-dependent acquisition mode using the top-speed method (3 s). The peptides were isolated using a quadrupole system (the isolation window was 0.7). The MS² analysis was performed in the ion trap using collision-induced dissociation fragmentation with a collision energy of 35 at a target value of $1 \times 10^4$ with 100 ms maximum injection time. The MS³ analysis was performed for each MS² scan acquired by using multiple MS² fragment ions isolated by an ion trap as precursors for the MS³ analysis with a multinotch isolation waveform (McAlister et al., 2014). HCD fragmentation was used for MS³ scan with an NCE of 65%, and the fragment ions were detected by the Orbitrap (resolution 15,000). The AGC target was $5 \times 10^4$ with a maximum ion injection time of 22 ms. The raw data sets have been deposited at the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the jPOST partner repository (https://jpostdb.org) (Moriya et al., 2019) with the dataset identifier JPST001027 (PXD026996).

**Data analyses**

**Database search.** The raw MS/MS data were processed with MaxQuant (Cox and Mann, 2008; Tyanova et al., 2016a). Peptide search with full tryptic digestion and a maximum of two missed cleavages was performed against the SwissProt human database (20,102 entries). The mass tolerance for precursor and MS³ ions was 4.5 ppm, whereas the tolerance for MS² ions was 0.5 Th. Acetylation (protein N-terminal), oxidation (M) and phospho (STY) were set as variable modifications and carbamidomethyl (C) was set as a fixed modification. The quantitation function of reporter ion MS³ (6-plexed TMT) was turned on. The false discovery rate was set to 1% at the level of PSMs and proteins. A score cut-off of 40 was used for identified modified peptides.

**QUANTIFICATION AND STATISTICAL ANALYSIS**

The abundances of TMT were log₂-transformed and further analyzed by Perseus (Tyanova et al., 2016b) for statistical evaluation such as principal component analyses and t tests. The PSP logo generator (Hornbeck et al., 2015) was used for sequence motif analysis. DAVID (Huang da et al., 2009) was used for gene ontology and pathway enrichment analysis. STRING v11 (Szklarczyk et al., 2019) was used for protein-protein interaction analysis. SigmaPlot (Systat Software), was used for preparing box plots.