analytical chemistry

# Gaining Confidence in the Elusive Histidine Phosphoproteome

Clement M. Potel,[†,‡,§] Miao-Hsia Lin,[†,‡] Nadine Prust,[†,‡] Henk W. P. van den Toorn,[†,‡] Albert J. R. Heck,[†,‡] and Simone Lemeer*[†,‡]

†Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, University of Utrecht, Padualaan 8, 3584 CH Utrecht, The Netherlands

‡Netherlands Proteomics Center, Padualaan 8, 3584 CH Utrecht, The Netherlands

S *Supporting Information*
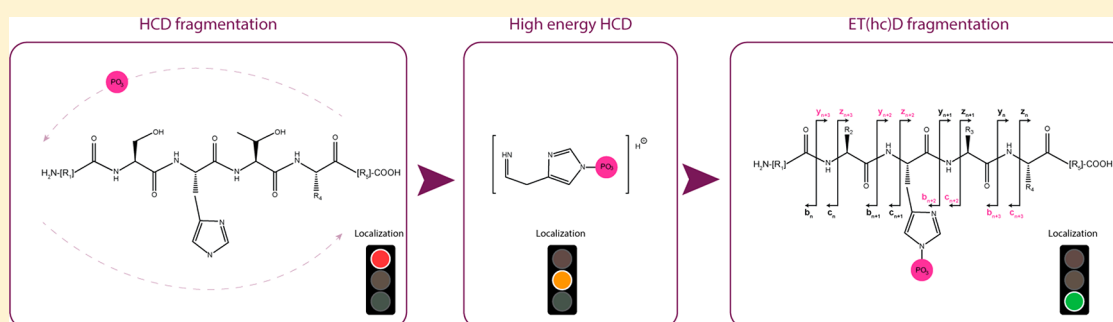
ABSTRACT: Recent technological advances have made it possible to investigate the hitherto rather elusive protein histidine phosphorylation. However, confident site-specific localization of protein histidine phosphorylation remains challenging. Here, we address this problem, presenting a mass-spectrometry-based approach that outperforms classical HCD fragmentation without compromising sensitivity. We use the phosphohistidine immonium ion as a diagnostic tool as well as ETD-based fragmentation techniques to achieve unambiguous identification and localization of histidine-phosphorylation sites. The work presented here will allow more confident investigation of the phosphohistidine proteome to reveal the roles of histidine phosphorylation in cellular signaling events.

Through chemical alteration of the nature of protein surfaces, reversible protein phosphorylation constitutes the primary means of regulating protein activity and is thus regarded as one of the most important protein post-translational modifications (PTMs).[1−4] Compared with the well-studied phosphohydroxyamino acids (serine, threonine, and tyrosine phosphorylation or pSTY), histidine phosphorylation (pHis) has received much less attention and was for a long-time considered an archaic type of modification. Over recent years, it has become apparent that protein histidine phosphorylation also occurs in higher organisms, and its biological functions have started to become unraveled,[5,6] underlining the importance of overcoming technical challenges hampering the study of the histidine phosphoproteome. Fortunately, some novel analytical tools allowing the enrichment and detection of the labile phosphohistidine have been developed,[7−9] opening the way to in-depth investigations. Antibody-based approaches pointed to a potential role of histidine phosphorylation in cell-cycle regulation,[7] as well as to potential oncogenicity of pHis deregulation,[10] but these antibodies are so far not well-suited for the enrichment of pHis phosphopeptides. We recently demonstrated that under optimized conditions, immobilized-metal affinity chromatography (Fe³⁺-IMAC) can be used to enrich histidine phosphorylated peptides along with pSTY peptides.[9] Despite

the fact that the first obstacle concerning the enrichment of pHis peptides seems to be largely tackled, the validation of identified phosphohistidine sites remains challenging. Because of a relatively low activation barrier, neutral loss of the phosphate group is a dominant dissociation pathway upon collision-induced dissociation (CID), complicating phospho-site localization.[11] This issue, inherent to the localization of most phosphorylation events, is even more prominent in the case of histidine phosphorylation, as the chemical nature of the phosphorus−nitrogen bond results in extensive neutral losses upon fragmentation as well as in potential gas-phase rearrangement.[12] Moreover, it has been demonstrated that confident localization of phosphosites after CID fragmentation becomes more challenging when the number of residues that may harbor a phosphate moiety increases.[13,14] In the human proteome, His has a relative low frequency of occurrence (∼3%), as compared with Ser (∼8%) and Thr (∼6%). It is then not surprising that the so-far reported pHis peptides often additionally harbor one or more Ser and Thr residues, enhancing the risk of false-positive identification and hampering correct phosphosite localization. The low con-

fidence in the pHis events identified constitutes to this day one of the biggest hurdles in the investigation of the biological functions of histidine phosphorylation, and thus tools validating identified pHis phosphosites are seriously needed.

EXPERIMENTAL SECTION

All peptides were synthesized on the basis of the Fmoc-based solid-phase synthesis strategy. Potassium phosphoramidate (PPA) was synthesized according to Wei and Matthews[15] (see the Supporting Information for more details). *Escherichia coli* strain BL21DE3 was grown in M9 minimal medium with 0.5% additional glucose at 37 °C. Cells were collected and lysed in lysis buffer with subsequent sonication. Next, methanol–chloroform protein precipitation was performed, and the protein precipitate was then digested overnight at room temperature. Fe^{3+}-IMAC enrichment was performed as described previously[16] (see the Supporting Information for more details). Nanoflow LC-MS/MS was performed by coupling an Agilent 1290 (Agilent Technologies, Middelburg, Netherlands) to an Orbitrap Fusion Lumos or an Orbitrap Q-Exactive HF (Thermo Scientific, Bremen, Germany). The mass spectrometer was operated in data-dependent acquisition mode. Several parameters were optimized for the immonium-triggering method using synthetic pHis peptides and can be found in Table S4 (see the Supporting Information for more details). Raw data files were processed using the MaxQuant[17] software (ver. 1.5.3.30) and searched against the Swissprot *E. coli* K12 database (4434 entries) in the case of synthetic pHis peptides and the Swissprot + TrEMBL *E. coli* BL21DE3 database (4294 entries) in the case of the endogenous pHis peptides. The false-discovery rate was set to 1%, a score cutoff of 40 in the case of modified peptides was used, and the minimum peptide length was set to seven residues (see the Supporting Information for more details).

RESULTS AND DISCUSSION

In comparison with the commonly studied phosphohydroxyamino acids (pSTY), it has earlier been reported that phosphohistidine peptides are more susceptible to exhibiting triplets of neutral losses (i.e., HPO_3, H_2PO_4, and H_3PO_5) from the precursor ion upon low-energy CID.[18] Here, we found that a similar prevalence of neutral-loss triplets is observed when endogenous pHis peptides are subjected to higher-energy-collisional-dissociation (HCD) fragmentation (Figure 1A and Table S1). However, despite the fact that pHis peptides are more prone to exhibiting triplets of neutral losses upon fragmentation, it does not constitute an indisputable diagnostic tool, as only 30% of the PSMs matching phosphohistidine peptides exhibit neutral-loss triplets. More importantly, neutral-loss triplets were also observed in approximately 10% of the spectra deriving from pST phosphopeptides. We also report here that reversed-phase-liquid-chromatography (RP-LC) retention times can be used to differentiate histidine phosphorylated peptides from STY phosphorylated isoforms, as synthetic pHis phosphopeptides elute predominantly later than their pSTY counterparts (Figure 1B and Table S2). This is in agreement with the fact that nonmodified histidine is the amino acid residue that contributes the least to peptide retention on C18 material,[13] and histidine phosphorylation is thus likely to result in an increase in retention time, possibly through charge neutralization.[19] However, a validation strategy based on synthetic peptides requires the prior correct
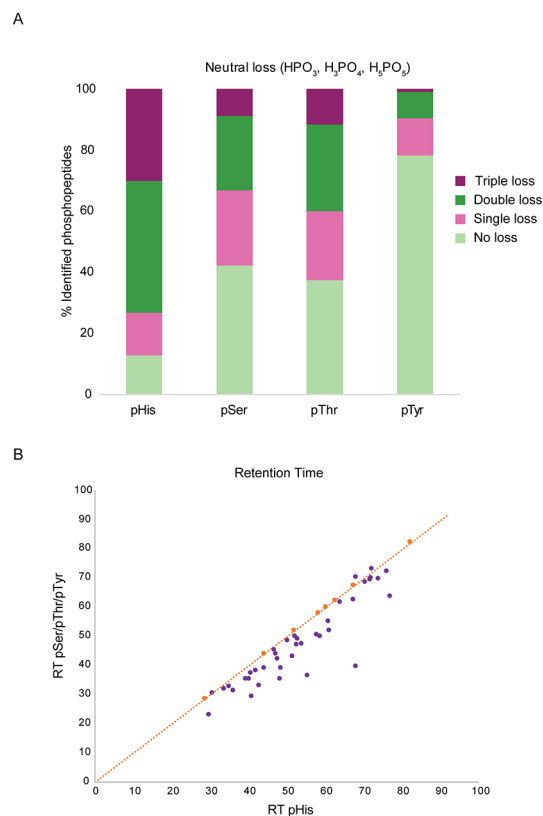

Figure 1. (A) Phosphate neutral-loss triplets (i.e., 80, 96, and 116 Da, respectively) prevalently, although not exclusively, observed in spectra derived from HCD fragmentation of pHis peptides. Here, the percentages of peptide-spectrum matches (PSMs) corresponding to endogenous class I phosphopeptides (MaxQuant localization probability ≥0.75) identified in *E. coli* samples that exhibit neutral losses are displayed for pHis and pSTY peptides. (B) Synthetic phosphohistidine peptides (purple dots) generally eluting later than their pSTY counterparts during RP-LC. Orange dots indicate RT standards.

identification of phosphohistidine peptides and, evidently, the synthesis of all pHis peptide candidates as well as all their pSTY isoforms, which is costly and time-consuming.

Phosphohistidine Immonium Ion as a True Diagnostic-Marker Ion. We demonstrated earlier that pHis immonium fragment ions are sometimes detected in the MS/MS spectra of pHis peptides.[9] We then observed that the chance of detecting the phosphohistidine immonium ion is strongly dependent on the pHis-site position within the peptide and on the phosphopeptide abundance.[9] In these earlier studies, only one-third of the identified histidine phosphorylated endogenous bacterial peptides exhibited the pHis-diagnostic immonium ion (m/z 190.0376). Still, this finding spurred us to investigate ways to increase the probability of detecting the pHis immonium ion, as well as the presence of any other potential diagnostic ions in mass spectra originating from synthetic and endogenous pHis peptides. To do so, 60 peptides harboring previously identified pHis sites were chemically phosphorylated on histidine residues following the protocol of Wei and Matthews[15] (Table S3). Binning of the low-mass region of the resulting MS2 spectra revealed another unique fragment ion at m/z 218.032, corresponding to sequential b-type and y-type cleavages of pHis peptides. Importantly, low-mass binning of

MS2 spectra derived from the analysis of full proteome samples and histidine-phosphorylation-depleted endogenous human phosphoenriched samples validated the specificity of these two marker ions (Figure S1). Because of the overall lower probability of observing the m/z 218.032 ion, we next focused exclusively on the m/z 190.0376 ion to develop a more robust diagnostic tool.

Optimized Settings for the Detection of the pHis Immonium Ion. Fragmentation pathways and fragment abundances heavily depend on collision-energy settings. For instance, the phosphotyrosine-immonium-ion intensity increases with increasing collision energy.[20] Here we found that the probability of detecting the two pHis marker ions is dependent on the HCD normalized-collision-energy (NCE) value and that incremental increases of the NCE value lead to an increase in the number of observed occurrences of the pHis immonium ion (Figure S2A). As expected, such high HCD collision energies led to peptide overfragmentation, impairing the ability to sequence the peptides by standard database search engines. We consequently developed an acquisition method relying on an initial high-energy-HCD-fragmentation step, followed by high-quality triggered MS2 scans with ETD-based fragmentation only when the pHis immonium ion is detected (Figure 2).

Immonium-Ion-Triggered EThcD Fragmentation Allowing Unambiguous pHis-Site Localization. It has been well-documented that electron-transfer-dissociation (ETD) techniques allow better preservation of labile phosphorylations, but the fragmentation efficiency highly depends on the precursor's charge density.[11,21] For this reason, supplemental activation is beneficial for low-charge peptides. In our case, ETD-only fragmentation was used for precursors of charges superior or equal to 4, and ETD with supplemental HCD fragmentation[22] (EThcD) was used for precursors of charges 2 and 3. This triggering strategy enable circumventing a drawback of ETD-based fragmentation (i.e., the low sampling rate limiting the depth of ETD experiments).

In addition to the collision energy, other settings were also optimized for the initial high-energy-HCD-fragmentation step, including the scan range, the Orbitrap mass resolution, the automatic-gain-control target (AGC, the setting allowing the determination of the number of charges to be accumulated before fragmentation), and the maximum injection time. A narrow scan range centered on the immonium-ion m/z as well as an Orbitrap resolution of 30k appeared to be optimal for maximizing the number of detected occurrences of the pHis diagnostic ions (Figure S2B,C). No significant differences in the numbers of pHis sites (Figure S3) or pHis peptide-spectrum matches (PSMs, Figure 3A) were observed between NCE values of 50 and 60 at the tested AGC target values, and thus an NCE value of 55 was chosen. The dependence of our approach on the number of precursor ions accumulated was more critical. Logically, low values of AGC targets resulted in significant drops in pHis-identification numbers. It however appeared that no significant difference was observed when the AGC target value was doubled from 5×10^4 to 1×10^5 charges (Figure 3A). These AGC target settings are similar to the values used in standard (phospho)proteomics high-speed acquisition methods with Orbitrap readout. The immonium-triggering method thus enables the use of ETD-based fragmentation techniques without sacrificing too much of the acquisition speed and thus the sensitivity. The final optimal parameters in the current method are summarized in Table S4.
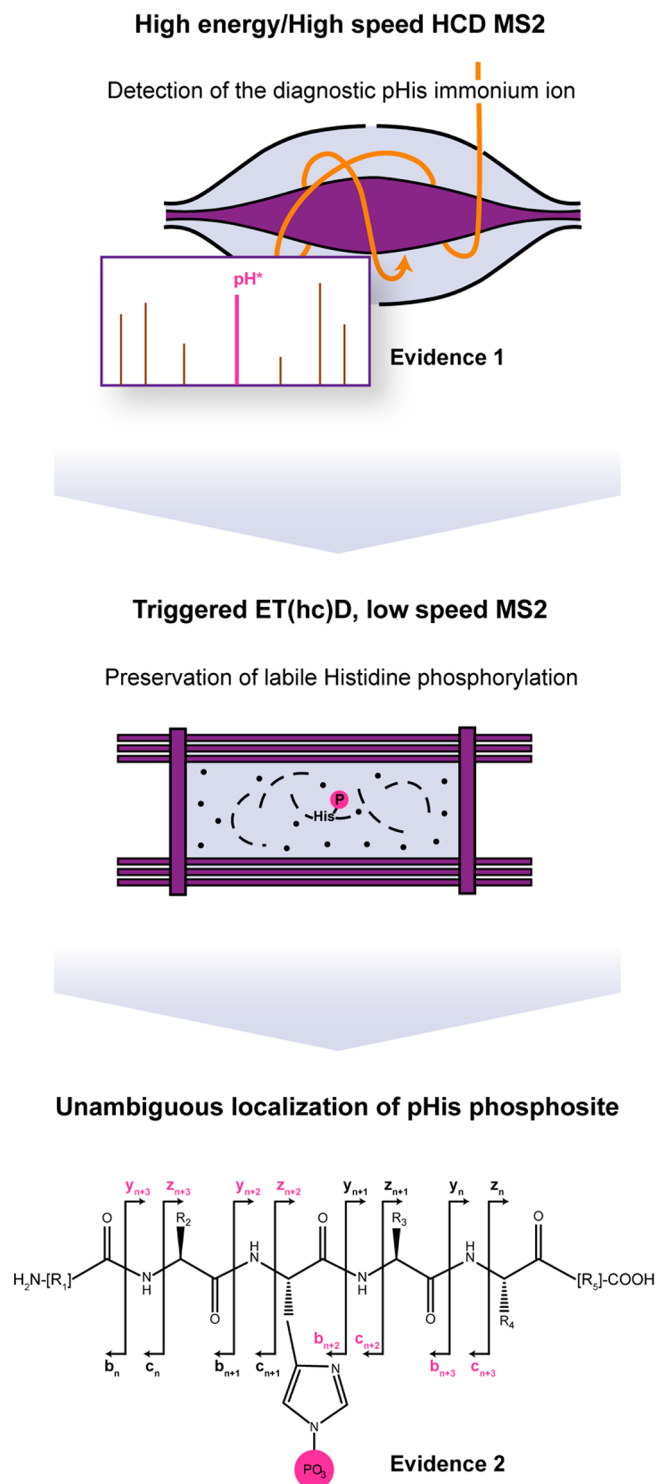

Figure 2. Phosphohistidine-immonium-ion-triggering method. The initial step consists of recording a high-energy-HCD-fragmentation spectrum, which is used to detect the pHis immonium marker ion. Upon detection of this diagnostic ion, the same precursor is subjected to ETD-based fragmentation, achieving improved peptide-sequence coverage while preserving the histidine phosphorylation.

Our optimized method led to the correct sequencing of ca. 90% of the 60 synthetic pHis peptides in a single LC-MS/MS run (Table S5). Notably, the immonium-ion-triggering method outperformed HCD fragmentation for the analysis of the synthetic pHis peptides, in terms of identification numbers and
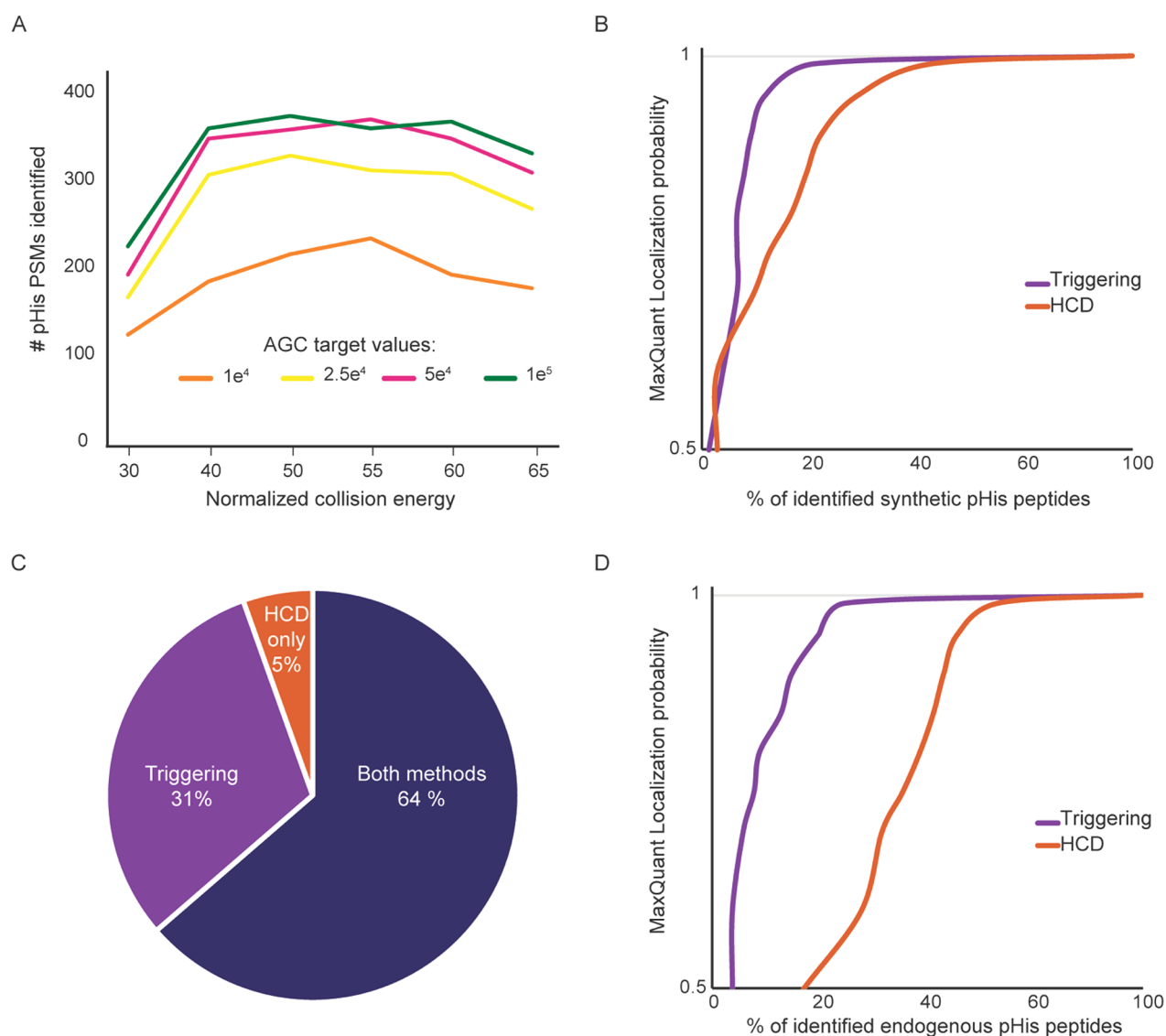
Figure 3. (A) Number of pHis peptide-spectrum matches (PSMs) identified by the immonium-triggering method as a function of the automatic-gain-control (AGC) target and the HCD normalized collision energy of the first high energy HCD MS2 step. (B) Comparison of MaxQuant localization probabilities of synthetic pHis phosphosites after HCD fragmentation vs those from the pHis-immonium-ion-triggering method. (C) Number of synthetic phosphohistidine sites identified after HCD fragmentation vs those from the immonium-ion-based-triggering method, using a MaxQuant-localization-probability cutoff of 0.9. (D) Comparison of localization probabilities of endogenous pHis phosphosites after HCD fragmentation vs those from the pHis-immonium-ion-triggering method.

identification scores (Figure S4). The primary purpose of this study is to improve localization confidence of histidine phosphosites. Besides using the immonium ion as a signature for the presence of histidine phosphorylation, this method uses ETD-based fragmentation techniques, which are well-known to result in higher phosphosite-localization confidence.[22,23] This is here reiterated by the fact that our method also outperforms HCD in terms of localization probabilities (Figure 3B). Consequently, the immonium-triggering method yielded 37% more pHis-site identifications than HCD at a MaxQuant-localization-probability cutoff of 0.9 (Figure 3C and Table S6).

Immonium-Ion Triggering Allowing Correct Site Localization in Complex Samples. Following this optimization, we next applied our method to biological samples. We chose to study *Escherichia coli*, as histidine phosphorylation plays known important roles in this system, and we were able to confidently identify 80 pHis sites (Table S7) by combining

the Fe^{3+}-IMAC enrichment protocol with the here presented immonium-ion-triggering method. For comparison, a previously reported strategy using a combination of pHis antibodies and neutral-loss triggering enabled the identification of 15 pHis sites in *E. coli*.[18] More than the number of pHis sites identified, the high confidence of the histidine-phosphosite identification and localization constitutes the real strength and novelty of the method presented here. This is exemplified by comparing our results to our previously reported *E. coli* data set, in which we identified a similar number of pHis sites per growth condition.[9] However, when comparing the distributions of the localization scoring for these identified endogenous peptides, it appears that the triggering method significantly enhanced the confident localization of pHis sites (Figure 3D). HCD fragmentation performed significantly worse for endogenous pHis peptides compared with for synthetic pHis peptides, which we attribute to the
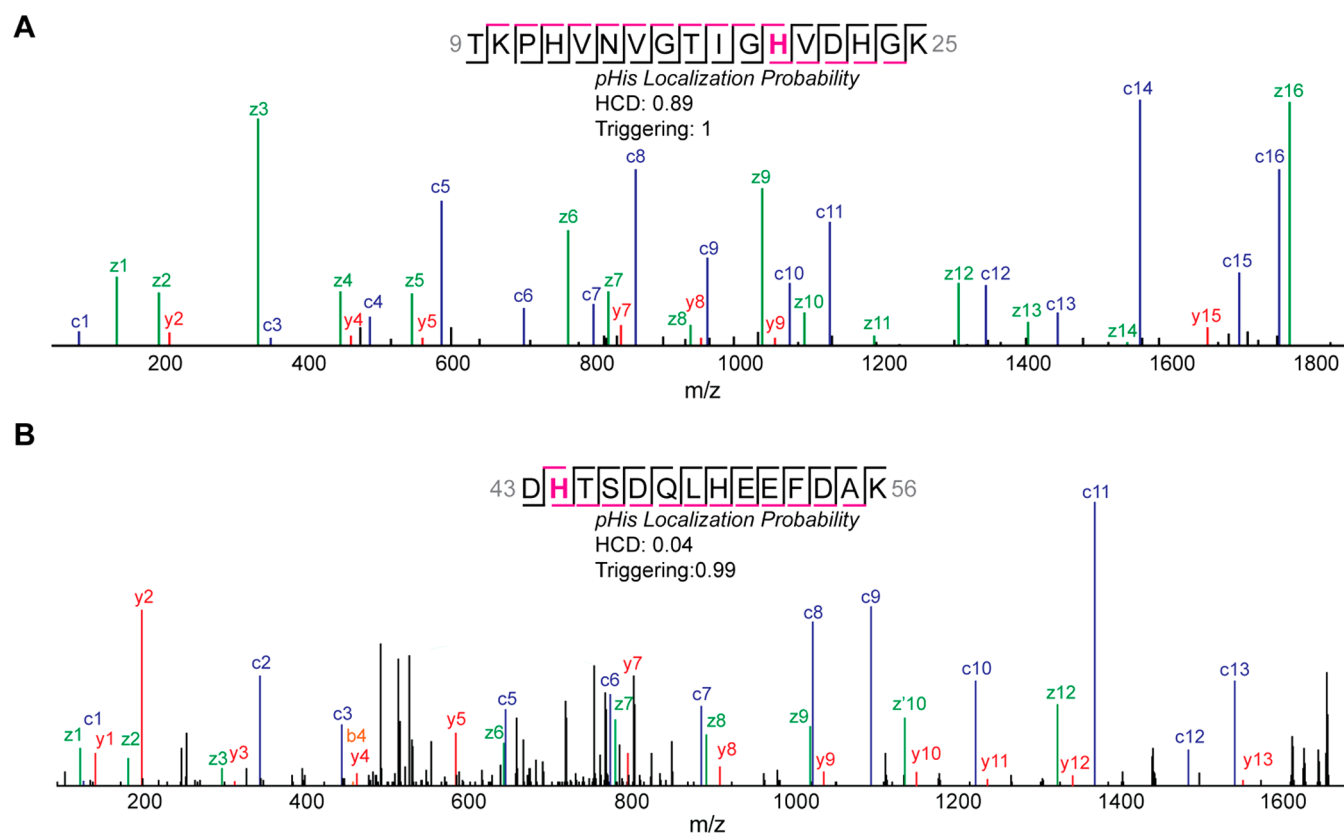
Figure 4. (A) Triggered ETD fragmentation of a phosphorylated peptide of the TufA−TufB proteins enabling the unambiguous localization of the phosphorylation site on the His20 residue because of the full sequence coverage and absence of neutral losses. (B) Triggered EThcD fragmentation of a phosphorylated peptide belonging to the LysU protein clearly pointing toward phosphorylation of His44. HCD fragmentation localized it to the adjacent threonine.

higher dynamic range and complexity of the endogenous sample, resulting in the generation of lower-quality MS2 spectra (Figure 3B,D). Our method enabled improved sequence coverage while preserving the labile histidine phosphorylation, as illustrated, for instance, by the identification of the His20 phosphosite on TufA−TufB and the His44 phosphosite on LysU (Figure 4A), which both harbor numerous amino acids that could accept a phosphate moiety. Examples of confident localization of histidine phosphorylation within peptides presenting adjacent putative phosphosites further emphasize the strength of the immonium-triggering method. Attributing the phosphosite on the DHTSDQLHEE-FDAK phosphopeptide, belonging to the lysine−tRNA ligase (LysU) protein, is very challenging by HCD, because of the close proximity of the histidine, serine, and threonine residues. Using HCD, the phosphosite was previously assigned to the Thr45 residue.[9] Here, we enhanced the MaxQuant-localization-probability score from 0.04 to 0.99 by utilizing EThcD instead of HCD, validating that the phosphorylation is undoubtedly localized at the His44 residue (Figure 4B). Similarly, the phosphoenolpyruvate-dependent-phosphotransferase-system protein (PtsP) is known to be phosphorylated on its active site His356. In our earlier report,[9] we were unable to confidently localize this histidine phosphosite, as the algorithm either reported high-confidence phosphorylation on the adjacent Ser355 or was unable to assign the phosphosite. As shown in Figure S5, the immonium-ion-triggering method in combination with EThcD enabled the confident localization of the phosphorylation site on His356, as both the y- and z-ion

series are complete, without neutral losses. As a final example, the phosphorylation of the His83 of the GatY protein, a residue responsible for zinc chelation, was not confidently identified by HCD because of the presence of adjacent potential phosphorylation-acceptor sites, whereas the immonium-ion-triggering method permitted, once again, unambiguous localization (Figure S6).

■ CONCLUSION

To summarize, we presented evidence that an optimized novel phosphohistidine-immonium-ion-triggering method can be used to extend the coverage of the still largely elusive phosphohistidine proteome. By combining two layers of evidence (i.e., the presence of the immonium diagnostic ion and the use of ETD-based fragmentation techniques), this method enhances confident identification of protein histidine phosphosites. Current knowledge of the biological role of histidine phosphorylation, in prokaryotes but also in eukaryotes, represents only the tip of the iceberg. In our view, our method should help the community gain in-depth insight in the biological function and abundance of histidine phosphorylation in organisms across the whole tree of life.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.anal-chem.9b00734.

Figure S1. Number of MS2 spectra presenting m/z 190.0376 and 218.0320 ions. Figure S2. Number of occurrences of the ions as functions of the HCD normalized collision energy, scan range, and Orbitrap resolution. Figure S3. Number of pHis sites identified by the immonium-triggering method. Figure S4. Average Andromeda score for peptide identification. Figures S5 and S6. Identification of the His356 phosphosite of the PtsP protein and the His83 phosphosite of the GatY protein by the immonium-ion method (PDF)

Materials and methods concerning peptide synthesis and histidine phosphorylation, *E. coli* growth and lysis conditions, Fe^{3+}-IMAC phosphoenrichment, LC-MS/MS, and data analysis (PDF)

Table S1. Detected neutral losses (XLSX)

Table S2. Retention times of pHis phosphopeptides and their pSTY counterparts (XLSX)

Table S3. List of synthetic peptides chemically phosphorylated on histidine residues and their pSTY counterparts (XLSX)

Table S4. Optimized parameters of the triggering mass-spectrometry acquisition method (XLSX)

Table S5. List of pHis synthetic peptides identified in a single LC-MS/MS experiment using the pHis-immonium-triggering method followed by database searches using MaxQuant software (XLSX)

Table S6. List of pHis synthetic peptides identified using either the pHis-immonium-triggering method or classical HCD fragmentation (XLSX)

Table S7. List of pHis endogenous peptides identified using the pHis-immonium-triggering method followed by database searches using the MaxQuant software (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: s.m.lemeer@uu.nl. Tel.: +31-2539974.

ORCID ●

Albert J. R. Heck: 0000-0002-2405-4404

Simone Lemeer: 0000-0002-6230-9471

Present Address

§C.M.P.: Genome Biology Unit, European Molecular Biology Laboratory, 69117 Heidelberg, Germany

Author Contributions

C.M.P., M.H.L., and S.L. designed the study. C.M.P., M.H.L., and N.P. performed the experiments. C.M.P., M.H.L., H.T., and S.L. performed data analysis. C.M.P, A.J.R.H., and S.L. wrote the manuscript.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Cohen, P. *Eur. J. Biochem.* **2001**, *268* (19), 5001−5010.

(2) Deribe, Y. L.; Pawson, T.; Dikic, I. *Nat. Struct. Mol. Biol.* **2010**, *17* (6), 666−672.

(3) Brognard, J.; Hunter, T. *Curr. Opin. Genet. Dev.* **2011**, *21* (1), 4−11.

(4) Pawson, T.; Scott, J. D. *Trends Biochem. Sci.* **2005**, *30* (6), 286−290.

(5) Fuhs, S. R.; Hunter, T. *Curr. Opin. Cell Biol.* **2017**, *45*, 8−16.

(6) Adam, K.; Hunter, T. *Lab. Invest.* **2018**, *98*, 233−247.

(7) Fuhs, S. R.; Meisenhelder, J.; Aslanian, A.; Ma, L.; Zagorska, A.; Stankova, M.; Binnie, A.; Al-Obeidi, F.; Mauger, J.; Lemke, G.; et al. *Cell* **2015**, *162* (1), 198−210.

(8) Kee, J.-M.; Oslund, R. C.; Perlman, D. H.; Muir, T. W. *Nat. Chem. Biol.* **2013**, *9* (7), 416−421.

(9) Potel, C. M.; Lin, M.-H.; Heck, A. J. R.; Lemeer, S. *Nat. Methods* **2018**, *15*, 187.

(10) Hindupur, S. K.; Colombi, M.; Fuhs, S. R.; Matter, M. S.; Guri, Y.; Adam, K.; Cornu, M.; Piscuoglio, S.; Ng, C. K. Y.; Betz, C.; et al. *Nature* **2018**, *555* (7698), 678−682.

(11) Potel, C. M.; Lemeer, S.; Heck, A. J. R. *Anal. Chem.* **2019**, *91* (1), 126−141.

(12) Gonzalez-Sanchez, M.-B.; Lanucara, F.; Hardman, G. E.; Eyers, C. E. *Int. J. Mass Spectrom.* **2014**, *367*, 28−34.

(13) Marx, H.; Lemeer, S.; Schliep, J. E.; Matheron, L.; Mohammed, S.; Cox, J.; Mann, M.; Heck, A. J. R.; Kuster, B. *Nat. Biotechnol.* **2013**, *31* (6), 557−564.

(14) Wiese, H.; Kuhlmann, K.; Wiese, S.; Stoepel, N. S.; Pawlas, M.; Meyer, H. E.; Stephan, C.; Eisenacher, M.; Drepper, F.; Warscheid, B. *J. Proteome Res.* **2014**, *13* (2), 1128−1137.

(15) Wei, Y.-F.; Matthews, H. R. Identification of Phosphohistidine in Proteins and Purification of Protein-Histidine Kinases. In *Methods in Enzymology*; Tony Hunter, B. M. S., Ed.; Academic Press, 1991; Vol. *200*, pp 388−414.

(16) Potel, C. M.; Lin, M.-H.; Heck, A. J. R.; Lemeer, S. *Mol. Cell. Proteomics* **2018**, *17* (5), 1028−1034.

(17) Cox, J.; Mann, M. *Nat. Biotechnol.* **2008**, *26* (12), 1367−1372.

(18) Oslund, R. C.; Kee, J.-M.; Couvillon, A. D.; Bhatia, V. N.; Perlman, D. H.; Muir, T. W. *J. Am. Chem. Soc.* **2014**, *136* (37), 12899−12911.

(19) Steen, H.; Jebanathirajah, J. A.; Rush, J.; Morrice, N.; Kirschner, M. W. *Mol. Cell. Proteomics* **2006**, *5* (1), 172−181.

(20) Diedrich, J. K.; Pinto, A. F. M.; Yates, J. R. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (11), 1690−1699.

(21) Riley, N. M.; Coon, J. J. *Anal. Chem.* **2018**, *90* (1), 40−64.

(22) Frese, C. K.; Zhou, H.; Taus, T.; Altelaar, A. F. M.; Mechtler, K.; Heck, A. J. R.; Mohammed, S. *J. Proteome Res.* **2013**, *12* (3), 1520−1525.

(23) Riley, N. M.; Coon, J. J. *Anal. Chem.* **2016**, *88* (1), 74−94.