Widespread protein histidine phosphorylation in bacteria

Clement Potel[1,2,#], Miao Hsia Lin[1,2,#], Albert J.R. Heck[1,2,*] and Simone Lemeer[1,2,*]

1. 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

2. Netherlands Proteomics Center, Padualaan 8, 3584 CH Utrecht, The Netherlands

[#] contributed equally

* Corresponding authors

For decades, the extreme difficulties in analyzing histidine phosphorylation have limited the study of phosphohistidine signaling. Here we report a method revealing widespread and abundant protein histidine phosphorylation in *E. coli*. Our new approach enhanced sensitivity and specificity of phosphopeptide enrichment in *E. coli* and more particularly of histidine phosphopeptides by a factor ~10. Thus, we generated the largest *E. coli* phosphoproteome dataset to date, with a total number of 2129 phosphosites, of which a remarkable high percentage (~10%) are phosphohistidine sites. This resource enables a major step forward towards a better understanding of the biological function of histidine phosphorylation, as well as of bacterial signaling pathways in general.

In bacteria, as in eukaryotes, reversible protein phosphorylation plays a key role in the regulation of virtually every cellular process[1]. It has been widely documented that protein phosphorylation primarily takes place at Ser, Thr and Tyr residues. Although protein phosphorylation can also occur at His residues, this was for a long time considered as an archaic type of phosphorylation. It has been well established that histidine phosphorylation plays crucial roles in bacterial signaling pathways, linking external stimuli to gene expression (the so-called two-component systems)[2] or bacterial cell metabolism (phosphotransferase system)[3]. In recent years, evidences that protein histidine phosphorylation also occurs in more complex organisms has emerged[4], but its role or abundance remain largely unknown, primarily due to the lack of adequate analytical tools. Indeed, since the discovery of histidine phosphorylation in 1962[5], only limited advances have been made in the detection of protein histidine phosphorylation and in the elucidation of its biological function[6-8]. Although excellent methods exist to analyze protein phosphorylation at Ser, Thr and Tyr residues[9,10], such approaches are considered not to be suited for the study of histidine protein phosphorylation. Several fundamental issues hinders the large-scale analysis of protein histidine phosphorylation. The high ΔG of hydrolysis of the phosphoramidate bond makes the phosphohistidine relatively unstable in comparison with commonly studied O-phosphorylations and protonation of the second nitrogen atom of the imidazole ring of histidine leads to the hydrolysis of the phosphate group under acidic conditions[11]. Consequently, commonly used methods for the enrichment of substoichiometric phosphopeptides, such as IMAC (Immobilized Metal Affinity Chromatography)[12] and MOAC (Metal Oxide Affinity Chromatography)[10], which involve strong acidic conditions, were so far incompatible with the study of protein phosphorylation at His residues. Here we challenge this dogma and present a new method allowing

the very efficient and sensitive enrichment of histidine phosphorylated peptides using Fe^{3+}-IMAC-columns, as well as an optimized sample preparation workflow.

The acid lability of histidine phosphorylation was considered to be an unsurmountable obstacle, as such acidic conditions are essential for selective phosphopeptides enrichment, good chromatographic separation or efficient ionization in positive ion mode. We assessed the validity of this statement by generating chemical histidine phosphorylation standards[13,14] that were subjected to different treatments (**Online methods**). As expected, strong acidic conditions led to hydrolysis of the histidine phosphorylation, but interestingly, most of the histidine phosphorylation could be preserved under relatively mild acidic conditions (pH=2.3) (**Figure S1**). This unexpected observation prompted us to attempt phosphohistidine enrichment using Fe^{3+}-IMAC columns, as it has been recently reported that this workflow does not necessitate strong acidic conditions to achieve high enrichment specificity[15]. As a first proof of principle, we performed enrichment of chemically histidine phosphorylated peptides and observed that pHis peptides were efficiently retained and preserved during enrichment (**Figure S2, Table S1**).

Encouraged by these result, we set out to study protein histidine phosphorylation in *Escherichia coli*, grown under four different biological conditions, *i.e.* exponential and stationary growth phase with either glucose or glycerol as carbon source. In previous studies on bacterial phosphoproteomes only a limited number of phosphosites (~100) were reported[16]. Here we present several optimization steps to expand the coverage of the bacterial phosphoproteome while preserving histidine phosphorylation. These critical steps included: (i) improved cell lysis using a stronger lysis buffer, (ii) removal of molecular interferences by protein precipitation, (iii) enzymatic depletion of nucleic acids, (iv) no use of temperatures above room temperature, (v) sample desalting at low temperature and (vi) fast (total time of 14 minutes) Fe^{3+}-IMAC enrichment

under relatively mild acidic conditions (**Figure 1A**). The remarkable change of the UV trace at the outlet of the Fe^{3+}-IMAC column attests of the high efficiency of the elimination of interfering components inherent to bacterial cell structure (**Figure S3**). As a result, the number of identified phosphorylations drastically increased, with a cumulative number of 2555 unique phosphopeptides (**Figure 1B, Figure S4**), representing a ~10-fold improvement when compared to earlier published bacterial phosphoproteome datasets (**Figure S5**).

With a total number of 246 pHis sites on 173 histidine phosphoproteins, we for the first time show the widespread occurrence of protein histidine phosphorylation in *E. coli* (**Figure 1C**). When comparing to previously published immunoprecipitation data[6,8], our results confirm 85% and 73%, of the 14 previously identified histidine phosphorylated proteins and 15 phosphosites respectively (**Figure 1D**). While the ratio between Ser/Thr/Tyr phosphorylations is in accordance with previous studies on the *E. coli* phosphoproteome[17], the major surprise comes from the fact that histidine phosphorylation represents around 10% of the *E. coli* phosphoproteome. Site localization was assessed by the database search algorithm MaxQuant[18] and of the 246 pHis sites, 135 phosphosites could be retained as high-confidence Class I pHis sites (localization confidence ≥ 75%) (**Table S2**). However, amongst the remaining 111 non-Class I pHis sites, 9 phosphohistidine sites were reported as genuine in the literature, possibly indicating the difficulty of localizing histidine phosphorylation sites due to prominent neutral losses of the phosphate group **(Figure 2A**). Even when considering only the class I phosphosites, we increased the number of pHis sites identified by a factor ~10 when compared to the recently reported strategies using pan-pHis antibody-based immunoprecipitation[8]. Notably, the detected class I pHis peptides were on average two times more abundant than the O-phosphorylated peptides, revealing the high abundance of protein histidine phosphorylation and good preservation of labile pHis peptides during the enrichment (**Figure S6**).

To further validate the pHis site assignments, the fragmentation patterns of several pHis peptides were compared to their corresponding synthetic analogues, for which possible STY-phosphorylated isoforms were also synthesized. The HCD fragmentation patterns of the synthetic pHis peptides presented the highest correlations with the endogenous *E. coli* pHis peptides, excluding the possibility of a miss assignment of the phosphorylation site (**Figure S7 & S8**).

Markedly, histidine exhibits an intense immonium ion in the low mass region of beam-type collision induced dissociation spectra. We observed an ion signal at 190.037 *m/z* in the HCD spectra of histidine phosphorylated peptides, corresponding to the phosphohistidine immonium ion (**Figure 2B**). This unique marker ion signal was nearly not observed in the *E. coli* proteome or strong acid treated phosphoproteome (**Figure S9**). The existence of the histidine immonium ion has previously been hypothesized[19,20], but to our knowledge, this constitutes the first report of the use of this diagnostic pHis immonium ion as a signature for the presence of phosphohistidine in large-scale phosphoproteomics datasets. In comparison with the well-known phosphotyrosine diagnostic ion, the presence of the phosphohistidine immonium ion was observed to be strongly dependent on the intensity of the precursor pHis peptide ion and the position of the phosphohistidine residue within the peptide (**Figure S9**).

No clear histidine phosphorylation motif could be statistically enriched, which is consistent with the fact that known histidine kinases all phosphorylate distinct substrates (**Figure S10**)[21]. We observed that histidine phosphorylation is widespread in several important bacterial metabolic pathways, such as glycolysis **(Figure 2C**) and the phosphotransferase system, thus explaining the relatively high abundance of detected histidine phosphorylations. In addition to detecting pHis on highly abundant metabolic enzymes (~ 15,000 copies per cell), we also detected histidine phosphorylation on the histidine kinase sensors arcB and dcuS, which are present at very low levels

(~10-100 copies per cell)[22], demonstrating that our strategy also enables the preservation of low abundant histidine phosphorylation, and covers a broad dynamic range. Our method also proved to be highly reproducible in terms of quantification of endogenous pHis peptides, which is a prerequisite for a full understanding of signaling networks (**Figure S11**). We observed significant regulation of pHis sites when comparing exponential and stationary phases (**Figure 2D, Table S3**), as for example the histidine phosphorylation on the stationary-phase-induced ribosome-associated (Sra) protein being up-regulated in the stationary phase. However, the biological relevance of these phosphorylations regulations needs to be further investigated.

When compared to previously antibody-based approaches[6-8], in which histidine phosphorylated proteins are enriched, the approach presented here: (i) allows for direct localization of the histidine phosphorylation site on the peptide (ii) allows for comprehensive identification and quantification of all phosphorylation events simultaneously (iii) requires less input material and (iv) is less expensive. In conclusion, this work paves the way for a better understanding of the biological role of bacterial phosphorylation and in particular protein histidine phosphorylation. Better understanding the biological role of protein histidine phosphorylation will be of crucial importance, as many pathways in which histidine phosphorylation is involved are key to the survival and physiology of *E. coli* and many pathogenic microorganisms. In addition, the methodological advances presented here could also answer a long standing question: is histidine phosphorylation also playing a major role in higher organisms?

Acknowledgments

This work was supported by the Roadmap Initiative *Proteins@Work* (project number 184.032.201), funded by the Netherlands Organization for Scientific Research (NWO). Additionally, we acknowledge support from the MSMed program, funded by the European Union's Horizon 2020 Framework Programme (grant agreement number 686547). SL acknowledges support from the Netherlands Organization for Scientific Research (NWO) through a VIDI grant (project 723.013.008). We thank Runsheng Zheng and Bernhard Kuster of the Technical University Munich for providing us with synthetic peptides used in this study.

References

1 Hunter, T. Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell* **80**, 225-236 (1995).

2 Laub, M. T. & Goulian, M. Specificity in two-component signal transduction pathways. *Annu Rev Genet* **41**, 121-145, doi:10.1146/annurev.genet.41.042007.170548 (2007).

3 Deutscher, J., Francke, C. & Postma, P. W. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol Mol Biol Rev* **70**, 939-1031, doi:10.1128/MMBR.00024-06 (2006).

4 Fuhs, S. R. & Hunter, T. pHisphorylation: the emergence of histidine phosphorylation as a reversible regulatory modification. *Curr Opin Cell Biol* **45**, 8-16, doi:10.1016/j.ceb.2016.12.010 (2017).

5 Boyer, P. D., Deluca, M., Ebner, K. E., Hultquist, D. E. & Peter, J. B. Identification of phosphohistidine in digests from a probable intermediate of oxidative phosphorylation. *J Biol Chem* **237**, PC3306-PC3308 (1962).

6 Kee, J. M., Oslund, R. C., Perlman, D. H. & Muir, T. W. A pan-specific antibody for direct detection of protein histidine phosphorylation. *Nat Chem Biol* **9**, 416-421, doi:10.1038/nchembio.1259 (2013).

7 Fuhs, S. R. *et al.* Monoclonal 1- and 3-Phosphohistidine Antibodies: New Tools to Study Histidine Phosphorylation. *Cell* **162**, 198-210, doi:10.1016/j.cell.2015.05.046 (2015).

8 Oslund, R. C. *et al.* A phosphohistidine proteomics strategy based on elucidation of a unique gas-phase phosphopeptide fragmentation mechanism. *J Am Chem Soc* **136**, 12899-12911, doi:10.1021/ja507614f (2014).

9 Rush, J. *et al.* Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat Biotechnol* **23**, 94-101, doi:10.1038/nbt1046 (2005).

10 Sugiyama, N. *et al.* Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications. *Mol Cell Proteomics* **6**, 1103-1109, doi:10.1074/mcp.T600060-MCP200 (2007).

11 Kee, J. M. & Muir, T. W. Chasing phosphohistidine, an elusive sibling in the phosphoamino acid family. *ACS Chem Biol* **7**, 44-51, doi:10.1021/cb200445w (2012).

12 Villen, J. & Gygi, S. P. The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. *Nat Protoc* **3**, 1630-1638, doi:10.1038/nprot.2008.150 (2008).

13 Hohenester, U. M., Ludwig, K., Krieglstein, J. & Konig, S. Stepchild phosphohistidine: acid-labile phosphorylation becomes accessible by functional proteomics. *Anal Bioanal Chem* **397**, 3209-3212, doi:10.1007/s00216-009-3372-x (2010).

14 Wei, Y. F. & Matthews, H. R. Identification of phosphohistidine in proteins and purification of protein-histidine kinases. *Methods Enzymol* **200**, 388-414 (1991).

15 Ruprecht, B. *et al.* Comprehensive and reproducible phosphopeptide enrichment using iron immobilized metal ion affinity chromatography (Fe-IMAC) columns. *Mol Cell Proteomics* **14**, 205-215, doi:10.1074/mcp.M114.043109 (2015).

16 Mijakovic, I. & Macek, B. Impact of phosphoproteomics on studies of bacterial physiology. *FEMS Microbiol Rev* **36**, 877-892, doi:10.1111/j.1574-6976.2011.00314.x (2012).

17 Lin, M. H., Sugiyama, N. & Ishihama, Y. Systematic profiling of the bacterial phosphoproteome reveals bacterium-specific features of phosphorylation. *Sci Signal* **8**, rs10, doi:10.1126/scisignal.aaa3117 (2015).

18 Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* **11**, 2301-2319, doi:10.1038/nprot.2016.136 (2016).

19 Zu, X. L., Besant, P. G., Imhof, A. & Attwood, P. V. Mass spectrometric analysis of protein histidine phosphorylation. *Amino Acids* **32**, 347-357, doi:10.1007/s00726-007-0493-4 (2007).

20 Medzihradszky, K. F., Phillipps, N. J., Senderowicz, L., Wang, P. & Turck, C. W. Synthesis and characterization of histidine-phosphorylated peptides. *Protein Sci* **6**, 1405-1411, doi:10.1002/pro.5560060704 (1997).
21 Capra, E. J. & Laub, M. T. Evolution of two-component signal transduction systems. *Annu Rev Microbiol* **66**, 325-347, doi:10.1146/annurev-micro-092611-150039 (2012).
22 Li, G. W., Burkhardt, D., Gross, C. & Weissman, J. S. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* **157**, 624-635, doi:10.1016/j.cell.2014.02.033 (2014).

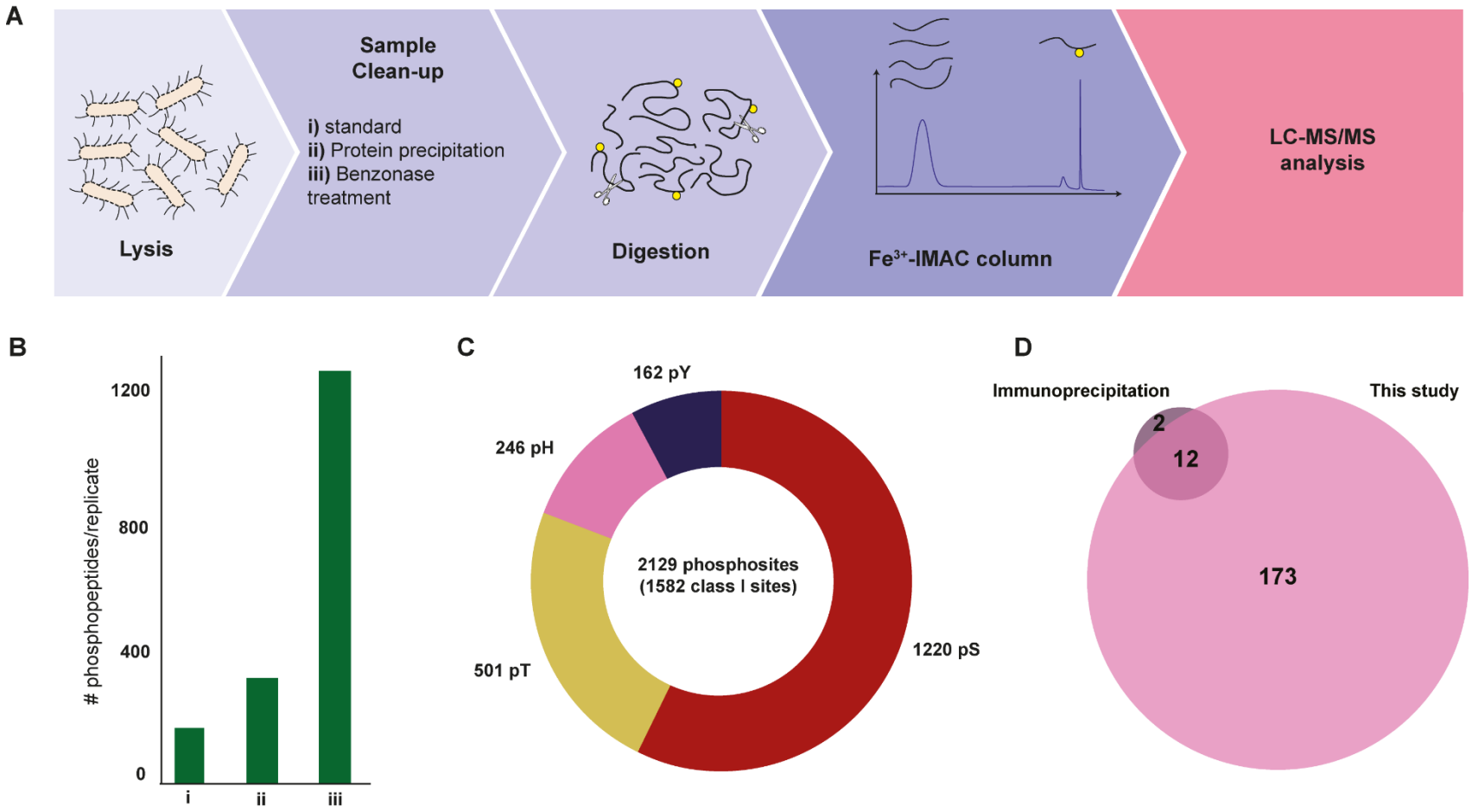
A
Lysis
Sample
Clean-up
i) standard
ii) Protein precipitation
iii) Benzonase
treatment
Digestion
Fe3+-IMAC column
LC-MS/MS
analysis
B
phosphopeptides/replicate
1200
800
400
0
i
ii
iii
C
162 pY
246 pH
501 pT
1220 pS
2129 phosphosites
(1582 class I sites)
D
Immunoprecipitation
This study
2
12
173

Figure 1: Experimental workflow and summary of results. (A) After cell lysis, efficient sample clean-up and protein digestion, the phosphopeptides are enriched using a Fe^{3+}-IMAC column prior to LC-MS/MS analysis. Several steps were optimized to ensure the preservation of the acid-labile histidine phosphorylation. (B) Total number of phosphopeptides identified per LC-MS/MS run. Through sequential optimizations, the number of phosphopeptides identified drastically increased. Amongst the different optimizations, there are two key steps: protein precipitation (ii) and nuclease treatment (iii). (C) Distribution of Ser/Thr/Tyr/His phosphosites identified. Of the 2129 identified phosphosites (the deepest and most comprehensive overview of a bacterial phosphoproteome to date), 246 are pHis making pHis more frequent than pTyr and only half as frequent as pThr. (D) Number of identified histidine phosphorylated proteins in comparison with the recently reported antibody-based affinity strategy[8]. When compared to the pan-pHis immunoprecipitation method, the number of identified phosphohistidine sites increased by more than 10 fold.

Figure 2

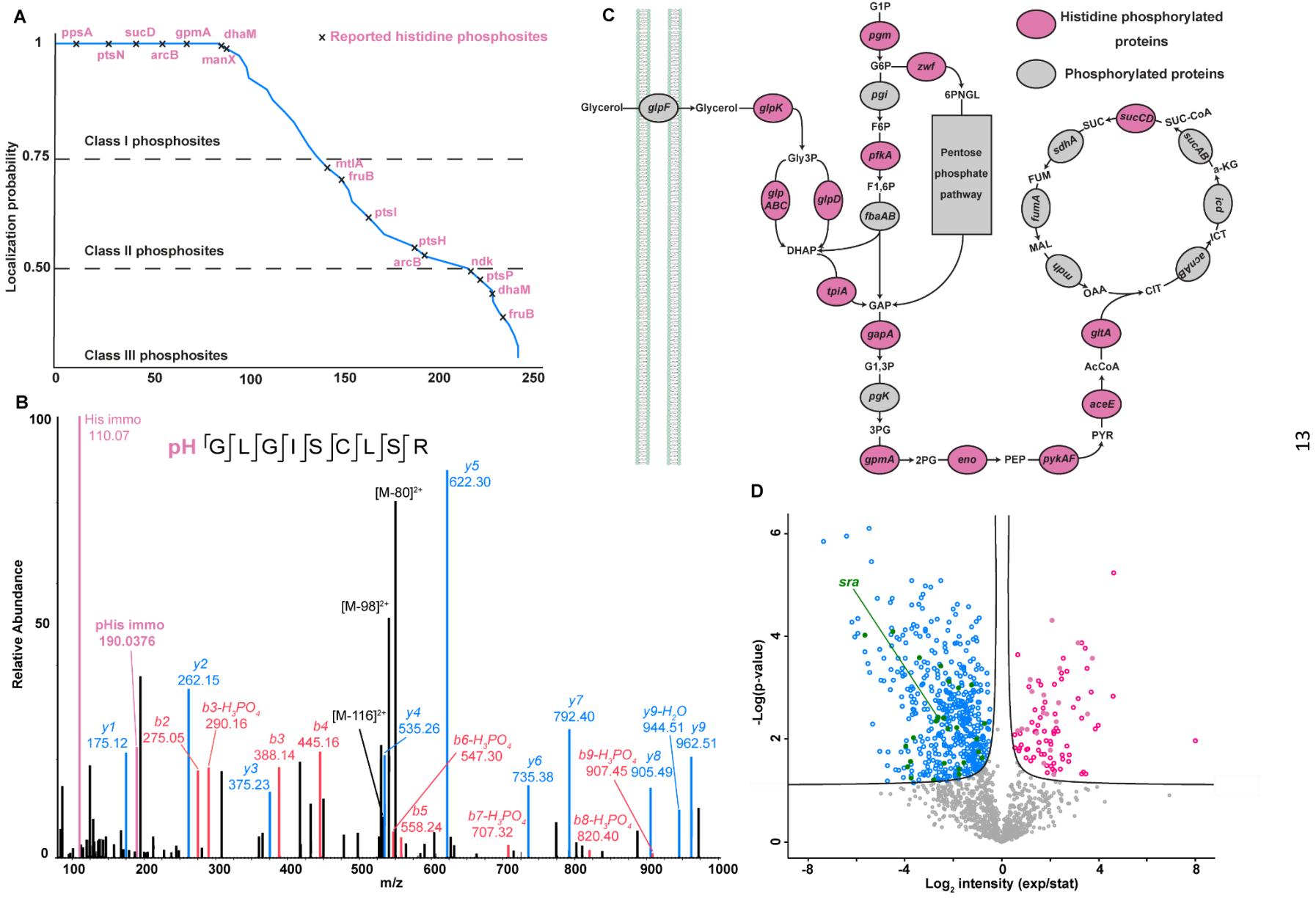
A
Localization probability
1
0.75
0.50
Class I phosphosites
Class II phosphosites
Class III phosphosites
ppsA
ptsN
sucD
arcB
gpmA
manX
dhaM
mtlA
fruB
ptsI
ptsH
arcB
ndk
ptsP
dhaM
fruB
× Reported histidine phosphosites
0
50
100
150
200
250
B
Relative Abundance
100
50
0
His immo
110.07
pH ⌈G⌋L⌋G⌋I⌋S⌋C⌋L⌋S⌋R
pHis immo
190.0376
y5
622.30
[M-80]2+
[M-98]2+
[M-116]2+
y2
262.15
y1
175.12
b2
275.05
b3-H3PO4
290.16
b3
388.14
y3
375.23
b4
445.16
y4
535.26
b5
558.24
b6-H3PO4
547.30
b7-H3PO4
707.32
y6
735.38
y7
792.40
b8-H3PO4
820.40
b9-H3PO4
907.45
y8
905.49
y9-H2O
944.51
y9
962.51
100
200
300
400
500
600
700
800
900
1000
m/z
C
Glycerol
glpF
Glycerol
glpK
Gly3P
glp ABC
glpD
DHAP
tpiA
GAP
gapA
G1,3P
pgK
3PG
gpmA
2PG
eno
PEP
pykAF
PYR
aceE
AcCoA
gltA
CIT
acnAB
ICT
icd
a-KG
oxoG
OAA
mdh
MAL
fumA
FUM
sdhA
SUC
sucCD
SUC-CoA
sucAB
G1P
pgm
G6P
zwf
6PNGL
pgi
F6P
pfkA
F1,6P
fbaAB
Pentose phosphate pathway
Histidine phosphorylated proteins
Phosphorylated proteins
D
-Log(p-value)
0
2
4
6
sra
-8
-4
0
4
8
Log2 intensity (exp/stat)

Figure 2. Site localization and biological regulation of phosphohistidine (A) Site localization of histidine phosphosites using the search algorithm (MaxQuant)[18], optimized for pSer and pThr, is less accurate. A significant number of histidine phosphosites reported in the literature and identified in this study exhibited a localization probability below the commonly accepted cut-off of 75%. This can be explained by the fact that this arbitrary cut-off might not be suitable for the localization of more than 3 possible phosphorylated residues, or by extensive neutral losses inherent to collision induced dissociation of histidine phosphorylated peptides. (B) HCD spectra of an endogenous *E. coli* phosphohistidine peptide corresponding to the His15 phosphosite of the Uncharacterized HTH-type transcriptional regulator YeiE protein. The phosphohistidine diagnostic immonium ion can be observed in the low mass region, as well as the characteristic neutral loss triplet[8]. (C) Histidine phosphorylation was detected on the majority of metabolic enzymes involved in the glycolysis. (D) Distribution of the 1469 quantified phosphosites in glycerol-fed *E. coli* cells in the exponential phase compared to the stationary phase, according to the t-test and fold change. The black curve indicates the significance level (false discovery rate (FDR) = 0.05, S0 = 0.1). Amongst the significantly regulated phosphosites, 19 pHis sites were up-regulated in the exponential phase (right, pink filled circle), while 26 were up-regulated in the stationary phase (left, green filled circle). Amidst the significantly up-regulated pHis sites is the stationary-phase-induced ribosome-associated protein (Sra) in the stationary phase.