ARTICLE

Global Histidine Phosphoproteomics in Human Prostate Cancer Cells

Yan Gao[1], Doeun Kim[1], Eunji Sung[1], Minjia Tan[2], Tae Gyun Kwon[3,4], Jun Nyung Lee[3], and Sangkyu Lee[1]*

[1]*BK21 Plus KNU Multi-Omics Based Creative Drug Research Team, College of Pharmacy, Research Institute of Pharmaceutical Sciences, Kyungpook National University, Daegu, Republic of Korea*
[2]*Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, China*
[3]*Department of Urology, School of Medicine, Kyungpook National University, Daegu, Republic of Korea*
[4]*Joint Institute for Regenerative Medicine, Kyungpook National University, Daegu, Republic of Korea*

Abstract : Histidine phosphorylation (pHis) is increasingly recognized as an important post translational modification (PTM) in regulating cellular functions in eukaryotes. In order to clarify the role of pHis in mammalian cell signaling system, a global phosphorylation study was performed in human prostate cancer cells, PC-3M, using a TiO_2 affinity chromatography. A total number of 307 pHis sites were identified on the 268 proteins among total identified 9,924 phosphorylation sites on 3,316 proteins. In addition, 22 pHis proteins were classified in enzyme category. This report provides the first database for the study of pHis in prostate cancer cells.

Keywords : Histidine phosphorylation, TiO_2-affinity chromatography, Mammalian cells, LC-MS

Introduction

Although protein histidine phosphorylation (pHis) was discovered in 1962, it was a non-popular protein post-translational modification (PTM) that has not been noticed for 50 years.[1,2] In contrast, phosphorylation studies at serine (Ser), threonine (Thr) or tyrosine (Tyr) residues have been actively conducted following various enrichment technology have been developed. Development of the enrichment method for pHis was relatively poor because the dephosphorylation proceeds faster because the pHis is labile than the phosphorylation of other amino acids.[3] To overcome the limitations, various methods for enrichment of pHis proteins have been developed in recent decades. Typically, a method for producing pan-antibody (mAb) from stable phosphohistidine analogues has been proposed, and studies have been conducted to identify the pHis phosphopeptides.[4-6] Immobilized metal affinity chromatography

(IMAC) and titanium dioxide (TiO_2)-affinity chromatography are also used to enrich the pHis phosphopeptides.[7,8] In addition, phosphopeptide enrichment methods using strong anion exchange (SAX) is reported in 2020.[9] Comparing these phosphorylation enrichment methods, IMAC affinity the phosphopeptides through the attraction of metal cations and negatively charged phosphate groups, TiO_2 is a versatile chemo-affinity chromatographic sorbent used in biomolecular research, SAX can enrich acid-labile phosphopeptides at near neutral pH.

In previous studies, pHis is crucial for prokaryotic signal transduction and as an intermediate for several metabolic enzymes. For example, NME family members catalyze transfer of phosphate from ATP onto Nucleoside diphosphates (NDPs) through a 1-pHis enzyme intermediate, 3-pHis is used to initiate phosphotransfer cascades and played as metabolic enzymes such as succinyl-CoA synthetase (SCS) and ATP-citrate lyase (ACLY).[5,10] However, its role in mammalian cells remains still largely uncharted. In order to characterize the biomechanism of PTM, substrate proteins should be identified as much as possible. In the case of pHis, limited information on the substrate proteins having pHis is a cause of the unclear function in mammalian cells. It was reported that, 246 pHis sites were identified in *Escherichia Coli* (*E. Coli*) using Fe3$^+$-IMAC affinity technology.[7] For the purpose of identifying higher abundance of pHis sites, our research group conduct global phosphoproteome research coupled slightly modified TiO_2-enrichement to identify the substrate proteins of pHis.

Experimental

Protein extraction

For global pHis phosphosproteome analysis, PC-3M cells purchased from Korea Cell Line Bank (KCLB, Seoul, Korea) were cultured in RPMI 1640 medium (Thermos fisher scientific, Waltham, MA, USA) containing 10% fetal bovine serum (HyClone laboratories, Inc. Logan, UT, USA) and 1% penicillin-streptomycin (Thermos fisher scientific, Waltham, MA, USA) in incubator at 5% CO_2, 37°C. Cells were subculture at 70%-80% confluent and were lysed by four types of lysis buffer comprised 1% protease inhibitor and 1% phosphatase inhibitor (Thermo Fisher Scientific, Waltham, MA, USA); ice-cold RIPA buffer (Thermos Fisher Scientific, Waltham, MA, USA), 4% SDS buffer (100 mM Tris-HCl, pH 8.0), Tris-buffer (25 mM Tris-HCl, pH 8.5) and 8 Urea buffer (100 mM Tris-HCl, pH 8.0). Protein amount was determined by BCA protein assay kit (Thermo Fisher Scientific, Waltham, MA, USA). Protein lysate was kept at -80°C until use.

Western blotting

PC-3M cell lysate was denatured and loaded 10 μg per/well to a 10% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) gel. The SDS-PAGE gel was running at 120 V in RT. The protein was then electro-transferred from the gel to polyvinylidene difluoride (PVDF) membrane (Roche, Basel, Switzerland) by wet transfer. Transferring was performed at a voltage of 90 V for 120 min in an ice-cold environment. After transferring, the membrane was blocking in 5% (*v/v*) bovine serum albumin (BSA) (GenDEPOT, Barker, TX, USA) in 1X TBST (137 mM sodium chloride and 20 mM Tris with 0.1% Tween-20) at room temperature for 2 h. The membrane was then incubated with primary antibodies: anti-phosphoserine and anti-phosphohistidine were purchased from abcam, (Abcam, Cambridge, MA, USA), anti-phosphothreonine and anti-phosphotyrosine (Cell Signaling Technology, Beverly, MA, USA). Antibodies were dilution as 1:1000 (*v/v*) in 5% (*v/v*) BSA, incubated membrane with gently shaking in 4°C for overnight. Membranes were then washed in the 1X Tris-Buffered Saline, 0.1% Tween 20 (TBST) buffer for 10 min, 3 times. The membrane was then incubated with (HRP)-linked anti-mouse or anti-rabbit antibody (Cell Signaling Technology, Danvers, MA, USA) at room temperature for 90 min. The membrane was then washed in TBST for 10 min, 3 times. For western blotting visualization, the membranes were dampened by 1 mL ECL reagent (GE Healthcare, Chicago, IL, USA), and exposure was performed with the Image Quant LAS 4000 mini (GE Healthcare, Chicago, IL, USA).

Proteins in-solution digestion and enrichment of phosphopeptide

Finally, two milligram protein lysates were utilized for in-solution digestion. Protein lysate was reduced in 5 mM dithiothreitol at 56°C for 30 min and alkylated in 15 mM iodoacetamide at room temperature for 45 min. To purify the protein, 4 folds volume of cold acetone was added and precipitated in -20°C for 4 h. After centrifugation at 16,000 g, 4°C for 15 min and pellet was re-dissolved in 100 mM ammonium bicarbonate (ABC) buffer. For the trypsin digestion, sequencing grade modified trypsin (Promega Corporation, Madison, WI, USA) was added as 1:50 (*w/w*) and incubated at 37°C for 12 h. Trypsin digestion was quenched by adding 10% (*v/v*) trifluoroacetic acid (TFA) at a final concentration of 0.4% (*v/v*) TFA. Peptide concentration was measured by quantitative colorimetric peptide assay kit (Thermo Fisher Scientific, Waltham, MA, USA). Phosphorylated peptides were enriched by TiO_2 Phosphopeptide Enrichment Kit (Thermo Fisher Scientific, Waltham, MA, USA) and Titansphere Phos-TiO MP kit (GL Science, Tokyo, Japan) refer to user's manual. Eluted phosphor-peptides were fractionated using a high-pH reversed-phase peptide fractionation kit (Thermo Fisher Scientific, Waltham, MA, USA) and desalted through SDB and GC tips (GL Science, Tokyo, Japan). Desalted peptides elution was then dried in a speed vacuum without heating. The dried peptides were kept at -80°C before use.

LC-MS analysis

Desalted peptide sample was dissolved in 2.5 μL of solvent A consisted of 2% acetonitrile, 0.1% acetic acid in water. Peptide sample was separated by nano-liquid chromatograph (Eksigent, Dublin, CA, USA) equipped with auto-sampler system. The C12 column (Proteo C12, 4 μm beads, 90-Å pore size, Phenomenex, Torrance, CA, USA) was packed with capillary tubing (length: 10 cm, OD: 75 μm, ID: 150 μm, Polymicro Technologies, Molex, Lisle, IL, USA). For the nano-LC, a 75 min gradient was employed for phosphorylated peptide analysis, and the flow rate was set to 300 nL/min. The procedure started with 6% of solvent B consisted of 0.1% acetic acid and 0.005% TFA in acetonitrile, and end with 21% solvent B. For the peptide analysis, nano-LC was coupled with the LTQ Orbitrap Velos mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) in the positive ion mode at Mass Spectrometry Convergence Research Center. Mass data was acquired in the data-dependent mode. The electrospray voltage was set to 1.8 kV. The precursor ion scans ranging from *m/z* 300–1,800 were acquired at a resolution of 60,000, and the automatic gain control (AGC) target value was 1.0×10^6 for the MS scan. Tandem MS was obtained using the collision-induced dissociation (CID) mode. The data-dependent mode that produced 20 of the most abundant ions from the full scans were fragmented in HCD mode with 35% normalized collision energy (NCE). The maximum IT was 100 ms for a full scan and 60 ms for MS^2. The mass data are available via ProteomeXchange with the identifier PXD016954.

Yan Gao, Doeun Kim, Eunji Sung, Minjia Tan, Tae Gyun Kwon, Jun Nyung Lee, and Sangkyu Lee

Database search and bioinformatics analysis

MS raw files were searched in MaxQuant (version 1.5) against a *Homo sapiens* database. For global proteomic peptide search, parameter setting as following: oxidation and acetyl (protein N-term) were chosen as variable modifications, Trypsin/P was chosen as digest enzymes, protein false discovery ratio (FDR) was set at 0.01. For phosphorylated peptide analysis data search, phosphorylation (S/Y/T/H) was added as an option in variable modifications. *Homo sapiens* proteome database was downloaded from UniProt (https://www.uniprot.org/proteomes/UP000005640). For result filtration, "contamination +" and "reverse +" were removed from the result file, and a protein "score" below 40 and phosphorylation "localization prob" below 75% were filtered from the results.

Phosphorylated protein subcellular localization was predicated by WoLF PSORT program (https://www.genscript.com/wolf-psort.html). The protein sequences that were downloaded from UniProt were inputted as a query. The subcellular location can be predicated by converting the amino acid sequences into numerical vectors. The detailed prediction method has been reported. Gene functional annotation and Kyoto Encyclopedia of Genes and Genomes (KEEG) pathway mapping was analysis by DAVID (https://david.ncifcrf.gov/). –log 10 of fisher exact test P value was used as an enrichment factor, the top five enriched terms were selected to visualized in bioinformatics figure. Proteomics date was analysis by Cytoscape software (version 3.7.2). The protein–protein interaction (PPI) network was retrieved from the web resource STRING (Version 10.5 https://string-db.org/). During STRING analysis, the UniProt accession IDs were inputted as a query. To characterize the histidine phosphorylation in the PC-3M cells, a sequence logo was generated using an open access, PhosphoSitePlus (https://plogo.uconn.edu/), against the *Homo sapiens* database. The pre-aligned sequences had a length of 15 amino acids, which were used for the foreground format. Histidine was arranged as a central character respectively.

Result and Discussion

Firstly, we performed Western blotting analysis of cell lysate using S/T/T/H-phosphorylation antibodies, to evaluate the abundance and cellular distribution of pHis in PC-3M cells (Figure 2A). The experimental results showed that pHis are widely spread like other phosphorylation of S/T/Y in mammalian cells. The global phosphoproteome analysis was based on LC-MS/MS analysis of phosphopeptides enriched by TiO$_2$ affinity chromatography as described above. To maximize the possibility of pHis peptide enrichment, experiments were conducted using products from two different brands. Using this procedure, we detected 9,924 unique phosphopeptides in 3,616 phosphoproteins from 1,463,120 MS/MS spectra with a Mascot score >40,
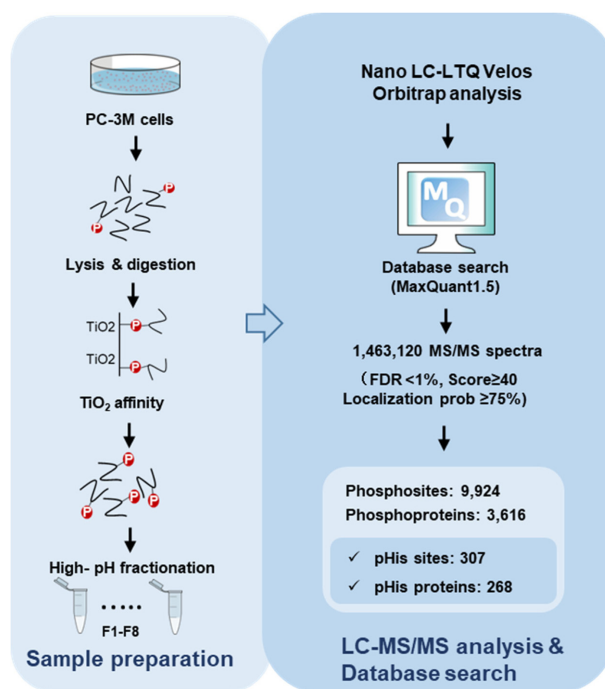

Figure 1. Experimental work-flow for global proteomics analysis of phosphorylation in prostate cancer cell. PC-3M cells were lysed and tryptic-digested, phosphorylated peptides were enriched by TiO$_2$ tips. Eluted peptides were analysis by a nano-LC-MS/MS system. Generated raw data was searched in database by MaxQuant (version 1.5). Totally, a number of 307 pHis sites on 268 proteins were identified among 9,924 phosphorylated sites on 3,616 proteins in PC-3M cells.

localization probability of >75% and a false discovery rate (FDR) <1% (Figure 1). The sample preparation was conducted by technically duplicated experiments. Distribution of S/T/Y/H-phosphorylation in this study was 80.4%, 14.9%, and 1.56%, 3.10% respectively (Figure 2B).

To further validate of pHis phosphorylated peptides, we manually confirmed the localization of phosphorylated sites to exclusive phosphorylation at other adjacent amino acid residues.[11] In total, we identified 307 pHis sites on 268 histidine phosphoproteins corresponding to 3.1% of the total identified phosphorylated sites (Figure 2B). The prevalence of pHis was smaller than our reported value, 6.3% in zebrafish embryos, it was higher than phosphorylation in Tyr, that 1% of eukaryotic cellular phosphoproteins.[12] As a result of subcellular localization of pHis, 45.0% was predicated to localize in the nucleus, and followed by 26.8% in cytoplasm, 8.1% secreted, 7.4 % in membrane, 7.4% in cell membrane, 3.4% mitochondrion, 2.0% in peroxisome, individually (Figure 2C). As a result of analyzing the sequence of pHis using the sequence log, S and P were enriched at positions 1 and 2, respectively, but other characteristic motifs could not be identified (Figure 2D).
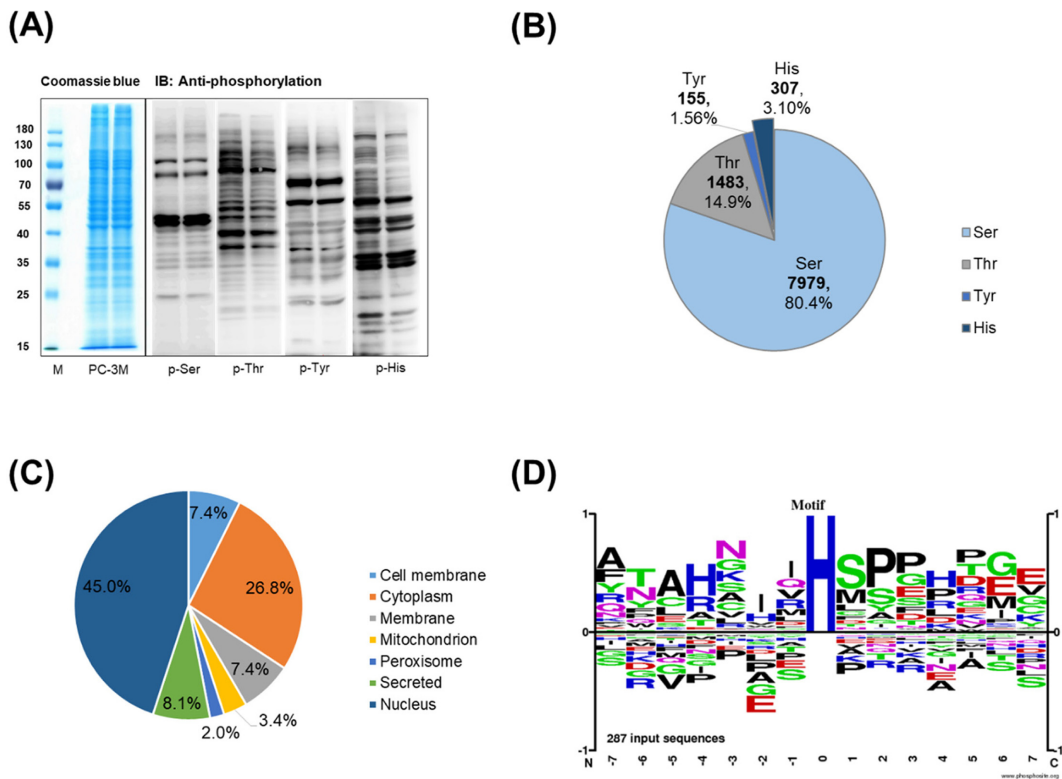
Figure 2. Proteomics profiling of phosphorylation (S/T/Y/H) in prostate cancer cell. (A) Western blotting scanning of phosphorylation in PC-3M cell that occurred on serine, threonine, tyrosine, as well as histidine. (B) Distribution of phosphorylation on serine (80.4%), threonine (14.9%), tyrosine (1.6%) and histidine (3.1%). (C) Subcellular distribution of phosphorylated protein. (D) Sequence logo of pHis sequences.
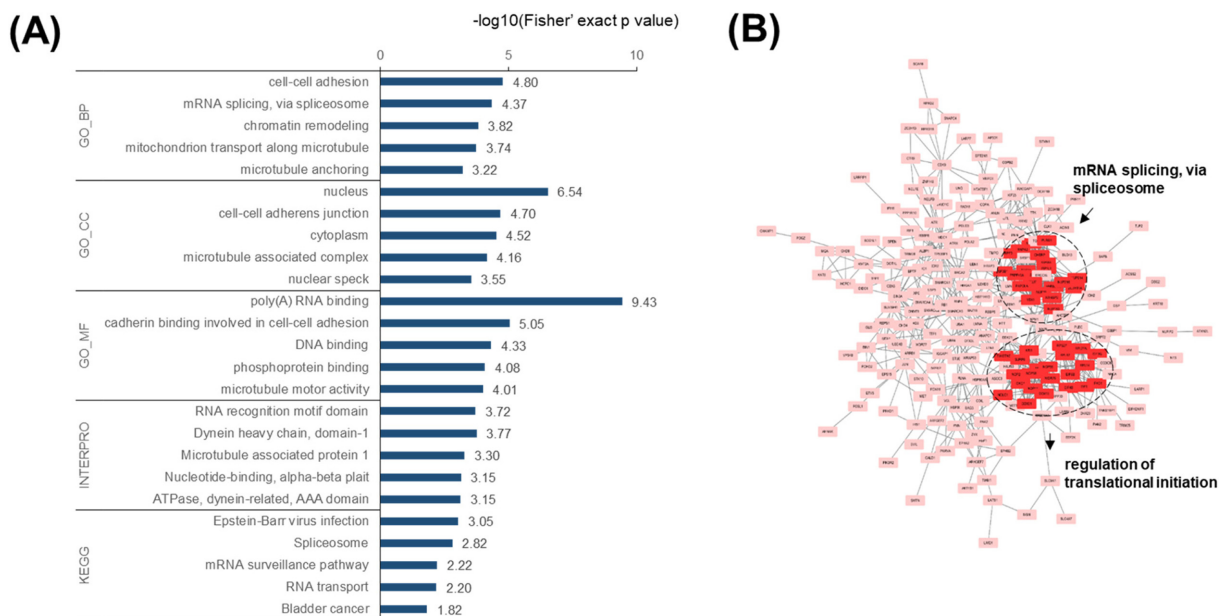

Figure 3. Function annotation of histidine phosphorylation in mammalian cells. (A) Bioinformatics analysis of Gene Ontology (GO) term, INTERPRO and KEGG pathway. (B) Protein-protein interaction of histidine phosphorylated protein, highly interacted in the mRNA splicing and regulation of translational initiation pathway.

Table 1. Selected histidine phosphorylated proteins in enzyme category.

Protein ID	Gene name	Protein name	Site	Sequence
Q63HN8	RNF213	E3 ubiquitin-protein ligase RNF213	H_1682	QME**H** (ph)FLDSWKR
Q9HAT8	PELI2	E3 ubiquitin-protein ligase pellino homolog 2	H_263	TADGLF**H** (ph)TPTQK
Q9C0C9	UBE2O	Ubiquitin-conjugating enzyme E2 O	H_83	LI**H** (ph)GEDSDSEGEEEGR
Q8IYL2	TRMT44	Probable tRNA (uracil-O(2')-)-methyltransferase	H_240	VKMSNVYQIQLS**H** (ph)SK
Q03164	KMT2A	Histone-lysine N-methyltransferase 2A	H_1842	GPGEPDSPTPL**H** (ph)PPTPPILSTDR
Q03164	KMT2A	Histone-lysine N-methyltransferase 2A	H_2097	TIA**H** (ph)SPTSFTESSSK
Q9H2M3	BHMT2	S-methyltransferase BHMT2	H_60	QL**H** (ph)MEFLR
P13051	UNG	Uracil-DNA glycosylase	H_20	**H** (ph)APSPEPAVQGTGVAGVPEESGDAAAIPAK
Q8NEJ0	DUSP18	Dual specificity protein phosphatase 18	H_165	NTV**H** (ph)MVSSPVGMIPDIYEK
Q92769	HDAC2	Histone deacetylase 2	H_391	MLPHAPGVQMQAIPEDAV**H** (ph)EDSGDEDGEDPDKR
P49753	ACOT2	Acyl-coenzyme A thioesterase 2, mitochondrial	H_9	SNKLLSP**H** (ph)PHSVVLR
Q8TBZ6	TRMT10A	tRNA methyltransferase 10 homolog A	H_320	NELDSP**H** (ph)EEK
P43403	ZAP70	Tyrosine-protein kinase ZAP-70	H_25	RSLEPAENV**H** (ph)GAGGGAFPASQTPSKPASADGHR
P12931	SRC	Proto-oncogene tyrosine-protein kinase Src	H_459	YLEEKNFV**H** (ph)R
Q15139	PRKD1	Serine/threonine-protein kinase D	H_271	VP**H** (ph)TFVIHSYTR
Q15139	PRKD1	Serine/threonine-protein kinase D	H_276	VPHTFVI**H** (ph)SYTR
Q9Y5S2	CDC42BPB	Serine/threonine-protein kinase MRCK beta	H_1695	HSTPSNSSNPSGPPSPNSP**H** (ph)R
Q14566	MCM6	DNA replication licensing factor MCM6	H_15	MDLAAAAEPGAGSQ**H** (ph)LEVR
Q9NVF9	ETNK2	Ethanolamine kinase 2	H_15	MAVPPSAPQPRASF**H** (ph)LR
Q06278	AOX1	Aldehyde oxidase	H_335	LPEEKTQMY**H** (ph)ALLK
O00116	AGPS	Alkyldihydroxyacetonephosphate synthase, peroxisomal	H_335	NIYGNIEDLVV**H** (ph)IKMVTPR
Q9BTU6	PI4K2A	Phosphatidylinositol 4-kinase type 2-alpha	H_54	VAAAAGSGPSPPGSPG**H** (ph)DR
P11678	EPX	Eosinophil peroxidase	H_388	STETPKLAAM**H** (ph)TLFMR
Q13093	PLA2G7	Platelet-activating factor acetylhydrolase	H_367	GSVHQNFADFTFATGKIIG**H** (ph)MLK

To characterize the potential roles of pHis in mammalian cells, we analyzed gene ontology (GO) functional classification analyses based on biological process (BP), molecular function (MF), and cellular component (CC) of the identified histidine-phosphorylated proteins and others (Figure 3A). In our previous study, the GO results of pHis and pS/T/Y were different in zebrafish embryos.[8] In this study, we identified the same and different points between pHis and pSTY. First, both pHis and pSTY were enriched in cell-cell adhesion and mRNA splicing, via spliseosome in COBP and enriched in poly (A) RNA binding in GOMF. Secondary, in GOCC, pHis showed enrichment during nucleus, whereas pS/T/Y were enriched nucleoplasm. Looking at the overall GO results, it was found that pHis is related to protein having cell-cell adhesion, which these was ranked higher in three GO categories. Protein-protein interaction (PPI) analysis of pHis proteins depicted that, proteins involved in mRNA splicing, via spliseosome, regulation of translational initiations were highly interacted (Figure 3B).

In IntroPro protein sequence analysis and KEGG pathway analyses showed a clear difference between pHis and pS/T/Y in mammalian cells (Figure 3A). pS/T/Y were enriched to kinase sections in IntroPro protein sequence analysis and signaling pathway in KEGG pathway, which was similar to our previous Zebrafish data.[8] However, in the case of pHis, it was enriched in a different category from those traditionally known as the mechanism of phosphorylation. pHis were shown to be involved in RNA recognition motif domain and cell structure proteins, which dynein heavy chain, microtubule associated protein 1, ATPase, dynein-related, AAA domain in IntroPro protein sequence analysis. In KEGG pathway analyses, pHis were enriched in Spliceosome, mRNA surveillance pathway and RNA transport category.

In prokaryotes, it is well-known that pHis is widespread and important in several crucial metabolic pathways, and pathogenesis-related or essential processes for survival in the host.[7,13] Additionally, pHis on Phosphoenolpyruvate synthase (PpsA) in *E. coli* using pan-His antibody was identified in 2013, which was found the amount of pHis on PpsA is sensitive to nitrogen availability in vivo.[5] In eukaryotes, 786 pHis proteins were enriched in the 293T cells lysates, which proteins were selectively classified to determine whether phosphorylation was 1-pHis or 3-pHis based on monoclonal antibodies with SILAC technology.[6] Additionally, we reported the 63 pHis proteins in Zebrafish larvae based on TiO$_2$-enrichment technology.[8] When comparing the results of this study with previous papers studied in eukaryote, it was not duplicated with the protein detected in zebrafish. Comparing the results with 293T cells, 39 proteins were found to overlap, but there was no pHis site information in the previous study, so only proteins could be identified.

So far, information on the pHis substrate protein is limited, and the need to identify more pHis site information and the substrate protein has been raised to characterize the role in biological systems. In order to closely approach the role of pHis in cancer cells, proteins in enzyme categories among the sites detected in this study was summarized in Table 1. There are 24 p-His sites in 22 proteins. It has been suggested that it can be a pHis for various proteins, from kinase to enzymes involved in histone modification. Although the role of pHis in individual proteins has not been identified in this study, further studies should continue to elucidate its function.

Conclusions

In conclusion, in order to approach the role of pHis in mammalian cells, we investigated the substrate and site information of pHis using prostate cancer cells. As a result of global phosphoproteomics analysis based on TiO$_2$ affinity chromatography, it was possible to provide 307 sites and 268 proteins information. Among the results, a list that classified the sites belonging to the enzyme protein in detail was also presented. The usefulness of the TiO$_2$ affinity assay for the identification of pHis and the wide presence of pHis in mammalian cells are presented.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

References

1. Boyer, P. D.; Deluca, M.; Ebner, K. E.; Hultquist, D. E.; Peter, J. B. *J. Biol. Chem.* **1962,** 237, PC3306.

2. Hunter, T. *Cell* **1995,** 80, 225, DOI: 10.1016/0092-8674(95)90405-0.

3. Kee, J. M.; Muir, T. W. *ACS Chem. Biol.* **2012,** 7, 44, DOI: 10.1021/cb200445w.

4. Kee, J. M.; Villani, B.; Carpenter, L. R.; Muir, T. W. *J. Am. Chem. Soc.* **2010,** 132, 14327, DOI: 10.1021/ja104393t.

5. Kee, J. M.; Oslund, R. C.; Perlman, D. H.; Muir, T. W. *Nat. Chem. Biol.* **2013,** 9, 416, DOI: 10.1038/nchembio.1259.

6. Fuhs, S. R.; Meisenhelder, J.; Aslanian, A.; Ma, L.;

Yan Gao, Doeun Kim, Eunji Sung, Minjia Tan, Tae Gyun Kwon, Jun Nyung Lee, and Sangkyu Lee

Zagorska, A.; Stankova, M.; Binnie, A.; Al-Obeidi, F.; Mauger, J.; Lemke, G.; Yates, J. R.; Hunter, T. *Cell* **2015,** 162, 198, DOI: 10.1016/j.cell.2015.05.046.

7. Potel, C. M.; Lin, M. H.; Heck, A. J. R.; Lemeer, S. *Nat. Methods* **2018,** 15, 187, DOI: 10.1038/nmeth.4580.

8. Gao, Y.; Lee, H.; Kwon, O. K.; Cheng, Z.; Tan, M.; Kim, K. T.; Lee, S. *Proteomics* **2019,** 19, e1800471, DOI: 10.1002/pmic.201800471.

9. Hardman, G.; Eyers, C. E. *Methods Mol. Biol.* **2020,** 2077, 225, DOI: 10.1007/978-1-4939-9884-5_15.

10. Fuhs, S. R.; Hunter, T. *Curr. Opin. Cell Biol.* **2017,** 45, 8, DOI: 10.1016/j.ceb.2016.12.010.

11. Chen, Y.; Kwon, S. W.; Kim, S. C.; Zhao, Y. *J. Proteome Res.* **2005,** 4, 998, DOI: 10.1021/pr049754t .

12. Beausoleil, S. A.; Jedrychowski, M.; Schwartz, D.; Elias, J. E.; Villen, J.; Li, J.; Cohn, M. A.; Cantley, L. C.; Gygi, S. P. *Proc. Natl. Acad. Sci. U.S.A.* **2004,** 101, 12130, DOI: 10.1073/pnas.0404720101.

13. Lai, S. J.; Tu, I. F.; Wu, W. L.; Yang, J. T.; Luk, L. Y. P.; Lai, M. C.; Tsai, Y. H.; Wu, S. H. *BMC Microbiol.* **2017,** 17, 123, DOI: 10.1186/s12866-017-1034-2.