Expanding and applying reactivity-based protein profiling (RBPP) in *Arabidopsis*

Brian C. Mooney, [1], Farnusch Kaschani[2], Joy Lyu[1] , Markus Kaiser[2] Renier A. L. van der Hoorn [1*]

[1] The Plant Chemetics Laboratory, Department of Biology, University of Oxford, South Parks Road, OX1 3RB Oxford, UK

2 ZMB Chemical Biology, Faculty of Biology, University of Duisburg-Essen; Essen, Germany.

*, for correspondence: renier.vanderhoorn@biology.ox.ac.uk

Abstract

Reactivity-based protein profiling (RBPP) is a powerful chemical proteomics technique that uses chemical probes to gain new insights into the surface exposure of amino acid residues within native proteomes. While RBPP has primarily been employed to identify proteins that could be potential targets for novel small molecule drugs, its applications in biological research are limited, especially in plant science. In this study, we introduce RBPP in plants by identifying over 15,000 labelling sites in native protein extracts from Arabidopsis cell cultures using an amine-reactive N-hydroxysuccinimide (NHS) ester probe linked to biotin. We demonstrate the utility of this approach by examining differential labelling in the Arabidopsis proteome in response to ATP. Pre-treatment with ATP results in decreased biotin-NHS labelling at residues near known ATP-binding sites and also revealed several novel ATP-responsive proteins. Lastly, we established RBPP with phosphonic acid-based NHS probes and immobilised metal affinity chromatography (IMAC) enrichment, revealing over 5,700 additional labelling sites. Collectively, these findings highlight that RBPP is a powerful and adaptable tool that can address key questions in plant biology.

Keywords

Chemical proteomics, reactivity-based protein profiling, RBPP, Arabidopsis, ATP

Introduction

Reactivity-based protein profiling (RBPP) is a powerful chemical proteomics approach that can provide insights into the reactivity and surface exposure of amino acid residues across whole proteomes (Hacker *et al.*, 2017; Ward *et al*. 2017). These sites are monitored using small molecule probes that selectively label specific amino acid side chains. For example, *N*-hydroxysuccinimide (NHS) and sulfotetrafluorophenyl (STP) esters display strong proteomic reactivity towards the primary amino group in the ε-amino side chain of lysines (Ward *et al*., 2017). These reactive groups (sometimes called 'warheads') are typically conjugated to an affinity tag to facilitate enrichment and identification of the labelled sites by liquid chromatography tandem mass spectrometry (LC-MS/MS) (Morimoto & van der Hoorn, 2016). Biotin is most widely used tag as the resulting peptides can be

enriched using avidin- or streptavidin-coated beads. The extremely strong natural affinity between biotin and avidin enables efficient capture of labelled peptides but requires stringent elution conditions for their release. More recently, phosphonate-tagged probes have emerged as practical alternatives for protein labelling as the resulting phosphopeptides can be enriched with high sensitivity using immobilized metal affinity chromatography (IMAC) resins (Paulo *et al*., 2018; Kleinpenning *et al*., 2020; van Bergen *et al*., 2023).

The application of RBPP has primarily been limited to biomedical science to identify 'hotspots' within proteins that could be targeted by novel small molecule therapeutic drugs (Hacker *et al*., 2017; Ward *et al*. 2017). Its potential to address wider questions in biology remains relatively overlooked. We hypothesize that RBPP could be deployed as a discovery tool to characterise plant responses to various biological stimuli by monitoring proteome-wide changes in surface-exposure or reactivity at labelled sites. Dynamic induced responses including conformational changes, protein-protein interactions, ligand binding and post-translational modifications among others would all be expected to result in labelling differentials on affected proteins (**Figure 1A**). RBPP could be used in this manner to identify novel proteins of interest involved in the response to virtually any experimental treatment, including biotic or abiotic stressors.

Here, we establish RBPP in plant science by identifying over 15,000 labeling sites in the native proteome of the model plant *Arabidopsis thaliana* using a biotin-NHS (BioNHS) probe and enrichment based on monomeric avidin agarose beads. We demonstrate the utility of RBPP to address biological questions by investigating differential labelling across the proteome induced by ATP, recovering known and putatively novel ATP binding sites. Finally, we describe an alternative RBPP strategy based on an IMAC-enrichable phosphonic acid-NHS (PhoNHS) probe and compare this with the BioNHS probe.

Results

RBPP monitors over 15,000 reactive residues in Arabidopsis

Proteins were extracted from Arabidopsis cells and incubated with a concentration gradient of an amine-reactive biotin-NHS probe (**Figure 1B**). Robust biotinylation of the Arabidopsis proteome was confirmed by immunoblot using streptavidin-HRP (**Figure 1C**).

Following tryptic digestion of the labelled proteome, biotinylated peptides were enriched using monomeric avidin beads and prepared for analysis by LC-MS/MS. NHS esters predominantly react with the primary amino group in the ε-amino side chain of lysines (K) but may also react with the hydroxyl groups of serine (S), threonine (T) and tyrosine (Y) (Ward *et al*., 2017). We searched for modified K, S, T or Y residues and identified 15,117 unique sites labelled by BioNHS (**Figure 1D**). Lysines accounted for approximately 86.95 % of all labelled sites identified (13,235 of 15,221), with 791 serines (5.20 %), 698 threonines (4.59 %) and 497 tyrosines (3.26 %) (**Figure 1D**).

ATP disrupts biotin-NHS labelling at known and novel ATP-binding sites

To investigate whether RBPP could be used to characterise proteome-wide changes induced by treatment with a small molecule, we treated an Arabidopsis protein extract with 1 mM ATP before labelling with BioNHS (**Figure 2A**). Although ATP did not lead to substantial changes in the overall number of labelled sites detected (**Figure 2B-C**), ATP pre-treatment did induce a significant decrease in labelling at 60 sites from 52 different proteins (**Figure 2D**). This set of 52 proteins showed a significant enrichment for GO terms relating to ATP, including 'ATP-dependent activity' and 'ATP hydrolysis activity' (**Figure 2E**).

To examine the position of labelling sites within the three-dimensional structure of selected ATP-binding proteins, we took advantage of AlphaFold predictions and experimentally determined structures. For example, labelling was significantly reduced at three lysine residues in the oxalyl-CoA synthetase AAE3 (ACYL-ACTIVATING ENZYME, Uniprot = Q9SMT7) in the presence of ATP; K178, K411 and K482. These residues surround known ATP-binding pockets at His 214 and Ser 291 (**Figure 3A)** (Fan *et al*., 2016). The previously resolved crystal structure of AAE3 in complex with ATP further reveals that K178 is directly involved in ATP binding via hydrogen bonding (Figure 3B). Taken together, these findings demonstrate that ligand binding sites can be identified using RBPP.

Phosphonic acid NHS labels a distinct subset of Arabidopsis proteins

Next, we investigated whether we could improve upon the biotin-avidin RBPP approach using the alternative phosphonate-IMAC enrichment strategy. We custom-synthesized a phosphonic acid-NHS (PhoNHS) probe (**Figure 4A**) and tested its reactivity with the Arabidopsis proteome. Robust labelling of the Arabidopsis proteome occurs at concentrations above 100 µM PhoNHS (**Figure 4B**). We then directly compared labelling of an Arabidopsis proteome labelled with either PhoNHS or BioNHS by LC-MS/MS with 4 replicates. In this experiment, we identified over 12,000 unique KSTY sites labelled with PhoNHS compared to 9,000 with BioNHS (**Figure 4C**). Of these sites, 3,566 were labelled with both probes while 5,734 and 3,117 were labelled only with either PhoNHS or BioNHS respectively (**Figure 4D**). Of the 3,566 sites commonly labelled by both probes, most were detected with a higher intensity upon labelling by PhoNHS (**Figure 4E**).

Taken together, these data demonstrate that RBPP is a highly flexible approach that could be used to investigate plant responses to stimuli by simultaneously monitoring thousands of amino acid sites across plant proteomes.

Discussion

We identified over 15,000 unique labelling sites in the native Arabidopsis proteome using a BioNHS probe providing broad coverage of surface-exposed residues that compares favourably with the number of labelling sites identified in other studies using similar approaches (Hacker *et al*., 2017; Morimoto *et al*., 2022). Our study focused on probes based on NHS chemistry because of their strong

reactivity with lysine which is one of the most exposed amino acids on protein surfaces. However, many other residues could be targeted by RBPP probes. For example, iodoacetamide probes have been used to monitor cysteine reactivity in mice (Xiao *et al*., 2020) and the plant pathogen *Pseudomonas syringae* (Morimoto *et al*., 2022). Considering BioNHS alone labels over 15,000 unique KSTY sites in Arabidopsis, using multiple probes to simultaneously monitor different amino acids could provide unrivalled proteome coverage amounting to many tens of thousands of sites.

We demonstrated that RBPP could be used to identify ATP binding sites, as inferred by reduced labeling in the presence of the ligand. This indirect approach contrasts with the strategy of using acyl-ATP probes to directly identify proteins bound by ATP (Villamor *et al*., 2013). We do not foresee RBPP as a substitute for such direct labelling strategies, but as a highly flexible discovery tool that can be used to investigate the response to virtually any type of stimulus using a single probe. Although we used ATP treatment here, RBPP could be applied to study any treatment that induces changes in reactivity or surface-exposure. Similarly, although we labelled a total protein extract, RBPP could be performed *in vivo* using cell permeable probes including BioNHS.

Biotin-conjugated chemical probes such as BioNHS are widely used for protein labelling because of the ability to enrich biotinylated proteins or peptides using streptavidin and related molecules. However, purification of biotinylated peptides has several limitations: i) many proteins stick non-specifically to streptavidin beads; ii) biotin is a relatively large modification that shifts many labelled peptides out of the MS detection window; and (iii) because of the high affinity of the biotin-streptavidin interaction, the elution of biotinylated peptides requires harsh conditions, resulting in the modification of 95-99% of the biotinylated peptides, such that these peptides can no longer be recognized by MS analysis (Kleinpenning *et al*., 2020). Attempts to improve the detection of labelled peptides include the use of desthiobiotin and/or neutravidin to reduce the affinity, and the use of cleavable linkers in probes (Morimoto & van der Hoorn, 2016), but these methods are also often inefficient and still cause peptide modifications. Phosphonate-conjugated labelling reagents serve as alternatives and generate phosphopeptides that can be enriched efficiently using immobilized metal-ion affinity chromatography (IMAC) resins.

Here, we show here that a phosphonate-IMAC approach results in the identification of over 12,000 labelled amino acids across the Arabidopsis proteome (**Figure 4C)**. Interestingly, these sites do not completely overlap with BioNHS labelled sites, indicating that PhoNHS labels a distinct set of residues (**Figure 4D).** As phosphonate is negatively charged at physiological pH, we investigated whether these differences could be attributed to differences in atomic charge at labelled sites. However, we did not identify a correlation between charge at labelled sites and preferential labelling by either BioNHS or PhoNHS. These differences could also be attributable to the slightly longer linker in the PhoNHS probe used here and the drastic differences in the avidin and IMAC enrichment procedures.

Experimental procedures

Arabidopsis cell culture - Cells were obtained from a suspension culture of wild-type *Arabidopsis thaliana* (L.) Heynh. ecotype Landsberg erecta, derived from an original stock established by May and Leaver (1993). The culture was maintained in the light at 22 °C on a rotary shaker in Murashige and Skoog (MS) medium supplemented with 160 mM glucose, 0.5 mg l-1 naphthelene acetic acid and 0.05 mg l-1 kinetin. Cells were harvested for experiments 5-7 days after subculturing.

Protein extraction - Arabidopsis cell samples were collected by centrifugation of 1 mL aliquots of the cell suspension. Cell pellets were flash frozen in liquid nitrogen and stored at -80 °C. Cells were lysed via addition of 800 µL of extraction buffer (50 mM HEPES pH 7, 1% PVPP, 5 mM MgCl2, 10% glycerol, 0.3% Triton-X). Lysates were incubated with the Benzonase nuclease (25 units per mL) for 30 minutes on ice before centrifugation at 15,000 x g to remove nucleic acids and cellular debris. Soluble proteins in the supernatant were retained and combined into a single tube. Protein yields were measured using the DC protein assay (Bio-Rad).

ATP treatment - The Arabidopsis protein extract was divided into 8 identical aliquots containing 700 µg protein. 1 mM ATP or a mock control was added to 4 replicates each and samples were incubated at room temperature for 30 minutes before labelling with NHS reagents.

Labelling with NHS reagents - The biotin-NHS probe used in this study is commercially available from Thermo Scientific™ as EZ-Link™ NHS-Biotin (20217). The phosphonic-acid NHS probe was synthesized by Haoyuan CehmExpress Co., Ltd. Stock solutions of NHS reagents were prepared in DMSO. Protein extracts were incubated with 1 mM biotin-NHS or phosphonic acid-NHS or an equivalent volume of DMSO (mock) on a rotary wheel for 1 hour at room temperature. Biotin-NHS labelling was assessed by immunoblotting using Streptavidin-HRP (Sigma; S2438). Phosphonic acid-NHS labelling was determined using the Pro-Q™ Diamond Phosphoprotein Gel Stain (Invitrogen) following SDS-PAGE. Following labelling proteins were precipitated by incubation with 4 x volumes of acetone and 20 mM NaCl at -20 °C. Excess NHS reagent was removed in the supernatant following centrifugation of the acetone precipitation.

Tryptic digest - Protein pellets were resuspended in denaturing buffer (6 M urea, 2 M thiourea in 50 mM ammonium bicarbonate) and incubated for 10 minutes at room temperature. Proteins were reduced with 10 mM TCEP and incubated at RT in the dark for 30 minutes. 50 mM IAA was then added and samples were incubated in the dark dark for a further 20 minutes. The urea concentration was reduced to approximately 1 M by the addition of 50 mM ammonium bicarbonate. Trypsin/Lys-C (Promega) was added to the samples at a 1:100 ratio of protease to protein. Digestion reactions were incubated overnight with shaking at 37 °C.

Peptide desalting - Peptides were desalted using Sep-Pak C18 cartridges (Waters) and stored at -20 °C. Briefly, tryptic peptides were acidified by the addition of trifluoroacetic acid (TFA) to 0.5%. The peptide samples were then diluted in solution A (2% acetonitrile, 0.1% TFA) before being passed through the C18 cartridge using a needleless syringe. After washing with solution A, desalted peptides

were eluted from the C18 membrane with solution B (65% acetonitrile, 0.1% TFA). Purified peptides were dried in a vacuum concentrator.

Enrichment of biotin-NHS labelled peptides - Desalted peptides were resuspended in 25 mM ammonium bicarbonate. Peptide concentration was measured using a BCA assay (Pierce Quantitative Colorimetric Peptide Assay). Biotinylated peptides were enriched using monomeric avidin agarose beads (Pierce). Beads were washed 3 times with H_2O. 500 µg input peptides were added to the washed beads and incubated on a rotating wheel for 1 hour at room temperature. Beads were washed 3 times with 300 uL MS Grade H_2O to remove nonspecific peptides. Biotinylated peptides were eluted from the beads with 500 uL elution buffer (80% acetonitrile; 1 % TFA) for 30 minutes at room temperature followed by 2 minutes at 96 °C. Beads were pelleted by centrifugation and eluted peptides in the supernatant were moved to a clean tube. Peptides were dried down in a vacuum concentrator.

Enrichment of phosphonic acid-NHS labelled peptides -Labelled phosphopeptides were enriched using MagReSyn Zr-IMAC and Ti-IMAC HP magnetic microparticles (ReSyn Biosciences). Dried, desalted peptides were resuspended in 200 uL IMAC loading buffer (0.1 M glycolic acid in 80% acetonitrile, 5% TFA). Samples were centrifuged to pellet and remove insoluble material. Zr and Ti-IMAC beads were combined at a 1:1 ratio and equilibrated in loading buffer before the addition of the peptides. Peptides were incubated with the beads for 30 minutes at room temperature on a rotating wheel. Flow-through supernatant was removed using a magnetic separation rack. Beads were washed with with loading buffer, wash buffer 1 (80% acetonitrile; 1% TFA) and wash buffer 2 (10% acetonitrile; 0.2% TFA) for 2 minutes each with gentle agitation. Bound peptides were eluted by incubation with 1% ammonium hydroxide for 20 minutes. 2.5% formic acid was added to the eluate. Peptides were dried down using a vacuum concentrator.

Analysis of protein structures - AlphaFold models were predicted using ChimeraX and ColabFold and visualised using PyMOL.

Mass spectrometry - Dried peptide samples were sent to the Analytics Core Facility Essen (ACE) for MS analysis.

ACKNOWLEDGEMENTS

We thank Urzsula Pyzio for excellent plant care and Sarah Rodgers, Caroline O'Brien and Patricia Bowman for technical support.

Funding - This project was financially supported by ERC AdG project 101019324 'ExtraImmune' (BM, RH), and Interdisciplinary DTC project DDT00230 (JL).

Author contributions - RH conceived the project; BM performed the experiments; JL performed bioinformatic analysis; FK and MK performed proteomic analysis; BM and RH wrote the manuscript with input from all authors.

Competing interests - none declared.

References

Fan, M., Xiao, Y., Li, M., & Chang, W. (2016). Crystal structures of *Arabidopsis thaliana* oxalyl-CoA synthetase essential for oxalate degradation. *Molecular Plant* **9:** 1349–1352.

Hacker, S., Backus, K., Lazear, M. *et al.* (2017) Global profiling of lysine reactivity and ligandability in the human proteome. *Nature Chem* **9:** 1181–1190.

Kleinpenning, F., Steigenberger, B., Wu, W. et al. (2020) Fishing for newly synthesized proteins with phosphonate-handles. *Nat Commun.* **11:** 3244.

May, M. J., & Leaver, C. J. (1993). Oxidative Stimulation of Glutathione Synthesis in Arabidopsis thaliana Suspension Cultures. *Plant physiology* **103:** 621–627.

Morimoto, K., & van der Hoorn, R. A. L. (2016). The increasing impact of activity-based protein profiling in plant science. *Plant & Cell Physiology*, **57:** 446–461.

Morimoto, K., Stegmann, M., Kaschani, F., Mohammed, S., & van der Hoorn, R. A. L. (2022). Cysteine reactivity profiling to unveil redox regulation in phytopathogens. *Methods in Mol. Biol.* **2447:** 105–117.

Paulo, J. A., Navarrete-Perea, J., Erickson, A. R., Knott, J., & Gygi, S. P. (2018). An Internal standard for assessing phosphopeptide recovery from metal ion/oxide Enrichment strategies. *J. Am. Soc. Mass Spec.* **29:** 1505–1511.

van Bergen, W., Hevler, J. F., Wu, W., Baggelaar, M. P., & Heck, A. J. R. (2023). Site-Specific activity-based protein profiling using phosphonate handles. *Mol. Cell. Proteomics* **22:** 100455.

Villamor, J. G., Kaschani, F., Colby, T., Oeljeklaus, J., Zhao, D., Kaiser, M., Patricelli, M. P., & van der Hoorn, R. A. L. (2013). Profiling protein kinases and other ATP binding proteins in Arabidopsis using Acyl-ATP probes. *Mol. & Cell. Proteomics* **12:** 2481–2496.

Ward, C. C., Kleinman, J. I., & Nomura, D. K. (2017). NHS-esters as versatile reactivity-based probes for mapping proteome-wide ligandable hotspots. *ACS Chem. Biol.* **12:** 1478–1483.

Xiao, H., Jedrychowski, M. P., Schweppe, D. K., Huttlin, E. L., Yu, Q., Heppner, D. E., Li, J., Long, J., Mills, E. L., Szpyt, J., He, Z., Du, G., Garrity, R., Reddy, A., Vaites, L. P., Paulo, J. A., Zhang, T., Gray, N. S., Gygi, S. P., & Chouchani, E. T. (2020). A quantitative tissue-specific landscape of protein redox regulation during aging. *Cell* **180:** 968–983.

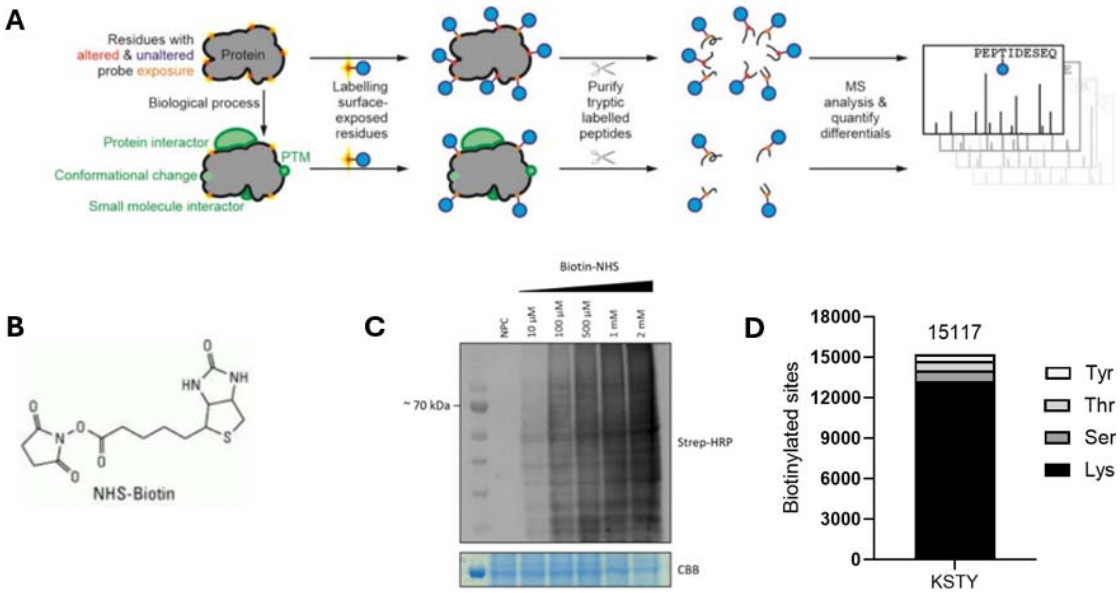

Figure 1. Establishing RBPP in *Arabidopsis thaliana*.

A. Schematic of RBPP approach. **B.** Structure of NHS-biotin probe. **C.** Biotin-NHS labelling leads to biotinylation of the *Arabidopsis* proteome in a concentration-dependent manner. Biotinylation was detected by streptavidin-HRP immunoblot. **D.** The number and identity of unique biotin-NHS labelled sites in the Arabidopsis proteome.

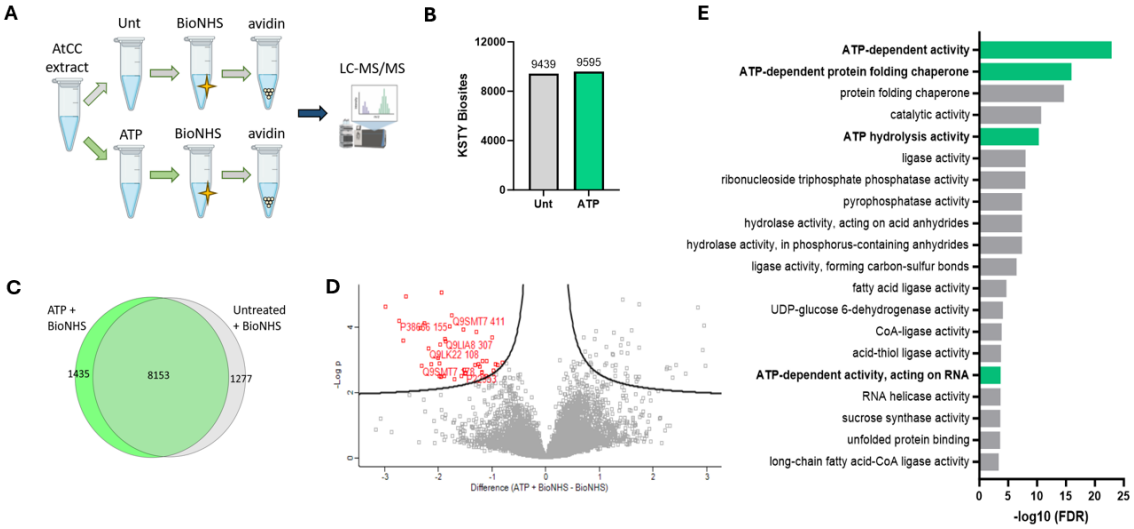

Figure 2. ATP alters labelling by bioNHS.

A. Schematic showing the experimental design for ATP pre-treatments. A total protein extract from Arabidopsis cells was pre-treated with 1 mM ATP or a mock solution before labelling for 1 h with 1 mM biotin-NHS. Biotinylated peptides by enriched using monomeric avidin agarose beads and identified by LC-MS/MS. **B.** Number of uniquely labelled KSTY sites following mock or ATP pre-treatment. **C.** Venn diagram showing overlap of labelled sites detected in the ATP treated and untreated groups. **D.** Volcano plot showing differentially labelled peptides +/- ATP. Left (red) indicates peptides with reduced labelling in the presence of ATP. **E.** Residues with reduced labelling after ATP treatment derive from proteins enriched in ATP-related GO terms.

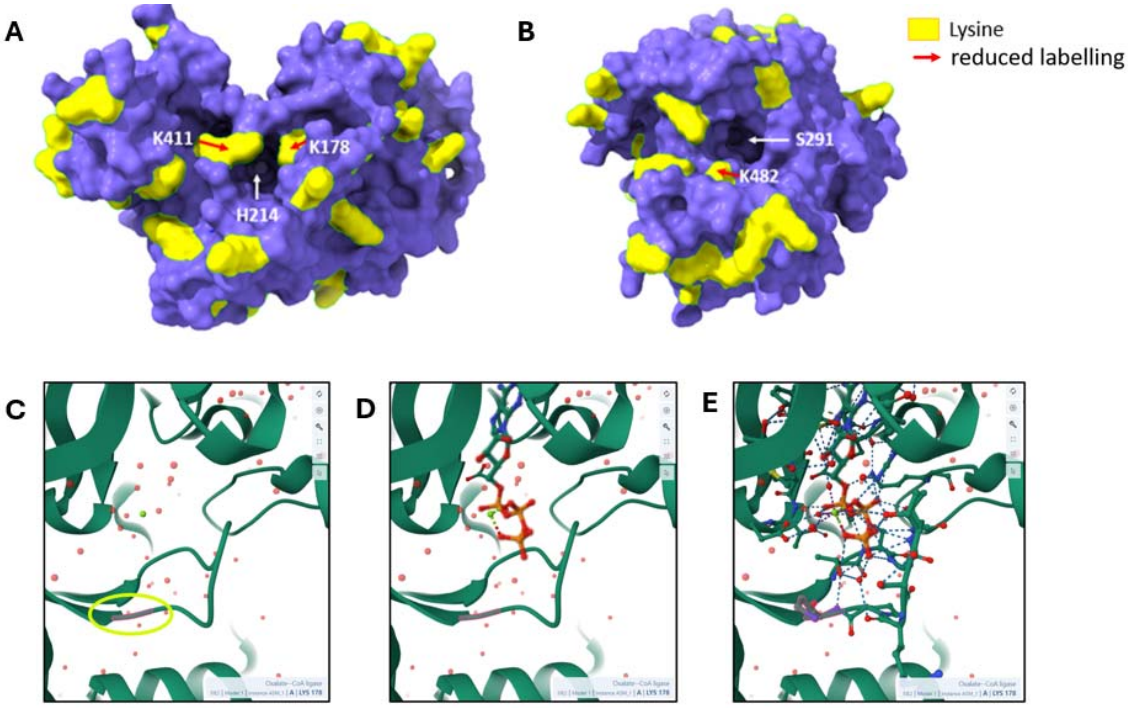

Figure 3. ATP pre-treatment induces loss of RBPP labelling close to ATP-binding sites.

A & B. AlphaFold-predicted structure of AAE3 showing ATP-binding sites H214 and S291 (white arrows). RBPP labelling was significantly reduced in the presence of ATP at nearby residues K411, K178 and K482 (red arrows). **C.** Crystal structure of AAE3 in the absence with K178 highlighted (yellow box). **D.** K178 in the presence of ATP **E.** Hydrogen bonds between K178 and ATP.

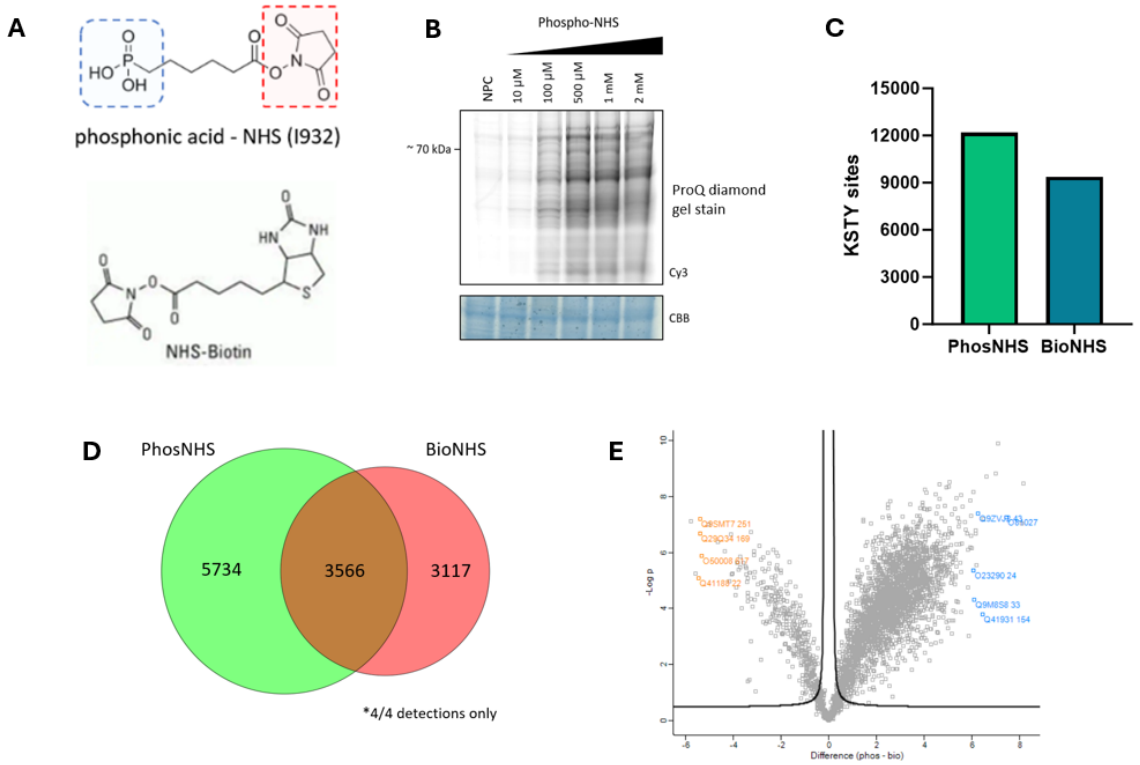

Figure 4. RBPP with phospho-NHS probe.

A. Chemical structures of the phospho-NHS and biotin-NHS probes. **B.** Labelling of total extract from Arabidopsis cells with a concentration gradient of phosphor-NHS after staining SDS-PAGE gel with ProQ diamond gel stain. **C.** Number of unique lysine sites labelled with either Phos-NHS or Bio-NHS probe. **D.** Overlap of lysine sites labelled by the biotin NHS and phospho NHS probes. **E.** Volcano plot comparing the intensity of sites labelled by both biotin-NHS and phospho-NHS. Left = stronger labelling following the biotin-avidin approach. Right =stronger labelling following the phosphonate-IMAC approach.