A Journal of the Gesellschaft Deutscher Chemiker

Angewandte International Edition GDCh Chemie

www.angewandte.org

## Accepted Article

**Title:** Chemical proteomic profiling of protein 4'-phosphopantetheinylation in mammalian cells

**Authors:** Nan Chen, Yuan Liu, Yuanpei Li, and Chu Wang



**To be cited as:** *Angew. Chem. Int. Ed.* 10.1002/anie.202004105

**Link to VoR:** https://doi.org/10.1002/anie.202004105

WILEY-VCH

WILEY-VCH

RESEARCH ARTICLE

# Chemical proteomic profiling of protein 4'-phosphopantetheinylation in mammalian cells

Nan Chen[a], Yuan Liu[a], Yuanpei Li[a] and Chu Wang[a,b] *

Dedicated to the 100th Birthday of Professor Youqi Tang

[a] Dr. N. Chen, Dr. Y. Liu, Y. P. Li and Prof. Dr. C. Wang
College of Chemistry and Molecular Engineering
Synthetic and Functional Biomolecules Center
Beijing National Laboratory for Molecular Sciences
Key Laboratory of Bioorganic Chemistry and Molecular Engineering of Ministry of Education, Peking University
Beijing 100871 (China)
E-mail: chuwang@pku.edu.cn
[b] Prof. Dr. C. Wang
Peking-Tsinghua Center for Life Sciences, Peking University

Supporting information for this article is given via a link at the end of the document.

**Abstract:** Protein 4'-phosphopantetheinylation is an essential post-translational modification (PTM) in prokaryotes and eukaryotes. So far, only five protein substrates of this specific PTM have been discovered in mammalian cells. These proteins are known to perform important functions including fatty acid biosynthesis and folate metabolism, as well as β-alanine activation. In order to explore existing and new substrates of 4'-phosphopantetheinylation in mammalian proteomes, we designed and synthesized a series of new pantetheine analogue probes enabling effective metabolic labelling of 4'-phosphopantetheinylated proteins in HepG2 cells. In combination with a quantitative chemical proteomic platform, we enriched and identified all the currently known 4'-phosphopantetheinylated proteins with high confidence, and unambiguously determined their exact sites of modification. More encouragingly, we discovered, via targeted proteomics, a potential 4'-phosphopantetheinylated site in the protein of mitochondrial dehydrogenase/reductase SDR family member 2 (DHRS2).

## Introduction

Co-enzyme A (CoA) participates in a large array of biochemical processes as a fundamental cofactor in essentially all living organisms.[1] In addition to serving as an acyl carrier, CoA can also modulate protein structures and functions via post-translational modifications (PTMs), one of which is protein 4'-phosphopantetheinylation[2]. The modification was first discovered in 1965[3] and has been known to be mediated by a phosphopantetheinyl transferase (PPTase) with CoA as the substrate[2]. In prokaryotic species, protein 4'-phosphopantetheinylation has been extensively documented as an "acyl carrier prosthetic arm" for natural product chain elongation reactions in fatty acid synthase (FAS), polyketide synthase (PKS) and non-ribosomal peptide synthetase (NRPS)-mediated biosynthesis.[4] Mechanistically, all intermediates are tethered on the terminal thiol through a reactive thioester linkage and elongated by a series of orderly enzymes to yield the final natural products with distinct structures and bioactivities.

In mammalian cells, only five protein substrates with 4'-phosphopantetheinylation have been reported (Table S1). They play important functions including fatty acid synthesis[3b, 5], folate metabolism[6] and β-alanine activation[7]. Whether there are other novel 4'-phosphopantetheinylated substrates in mammalian proteomes remains unexplored.

Chemical biology tools have been developed to aid functional study of protein 4'-phosphopantetheinylation. Burkart and colleagues pioneered in synthesizing unnatural pantetheine analogue probes to enable chemical labelling of acyl carrier proteins in bacteria.[8] They further demonstrated that the strategy could be exploited to capture transient protein-protein interactions with acyl carrier proteins for structural studies[9] and to discover novel carrier proteins from unsequenced prokaryotic organisms[10]. Unfortunately, when they tested the performance of these probes in human cells, they could only label one substrate protein, FAS, with a fluorescent analogue probe, likely due to the difference in pantetheine importing mechanisms between eukaryotic and prokaryotic organisms.[10] Therefore, despite its functional essentiality in mammalian biology, there is still lack of efficient tools for comprehensive profiling of substrates of 4'-phosphopantetheinylation in mammalian cells to date.

Herein, we designed and synthesized a series of new pantetheine analogue probes that showed improved metabolic labelling in mammalian cells. By integrating these probes with quantitative chemical proteomics, we performed a global profiling of 4'-phosphopantetheinylated proteins in mammalian proteomes for the first time. Implementation of a targeted chemical proteomic strategy unambiguously identified the exact modification sites, which include all five of the known sites of 4'-phosphopantetheinylation and a potential site from a new substrate. Our study complements the toolbox of unnatural pantetheine probes and opens new opportunities for functional study of protein 4'-phosphopantetheinylation in mammalian cells.

## Results and Discussion

### Establishment of a quantitative MS-based platform for chemoproteomic profiling

In a previous study[10], Burkart and colleagues tested the labelling of two pantetheine analogue probes in the human breast cancer SKBR3 cell line. They observed by in-gel fluorescence that only the fluorophore-functionalized probe could result in modest labelling of human FAS while the azide-functionalized probe showed no detectable labelling at all. Notably, although another alkyne-functionalized analogue was shown to harness bacterial pantothenate-uptake and sequential enzymatic processes to yield a functional CoA mimic with a similar kinetic performance as the natural pantothenate[8b], it was never reported to be tested in mammalian cells. We therefore set out to test if such an alkynyl pantetheine probe, herein referred to as **PP-1** (Figure 1A), can be metabolically incorporated into mammalian cells to label 4'-phosphopantetheinylated proteins efficiently. If so, the labelled proteins can be profiled by mass spectrometry-based quantitative chemoproteomic analysis.

We first synthesized **PP-1** according to the reference[8b]. To control for false positive identifications resulting from non-specific probe incorporation or unexpected reactivity, we also designed and synthesized an inactive probe, **PP-1-C** (Figure 1A). **PP-1-C** does not have the head hydroxyl group, and theoretically, should not be capable of being phosphorylated to form the CoA analogue[11]. Both probes showed no obvious cytotoxicity at concentrations of up to 1 mM in HepG2 cells (Figure S1). Probe incorporation was examined by incubating HepG2 cells with 1 mM of **PP-1** or **PP-1-C** for 24 h, followed by cell lysis and conjugation with biotin-azide. Labelled cell lysates were resolved by immunoblotting and the results showed that the proteome labelling by **PP-1** was much stronger than that by the control probe (Figure 1B).

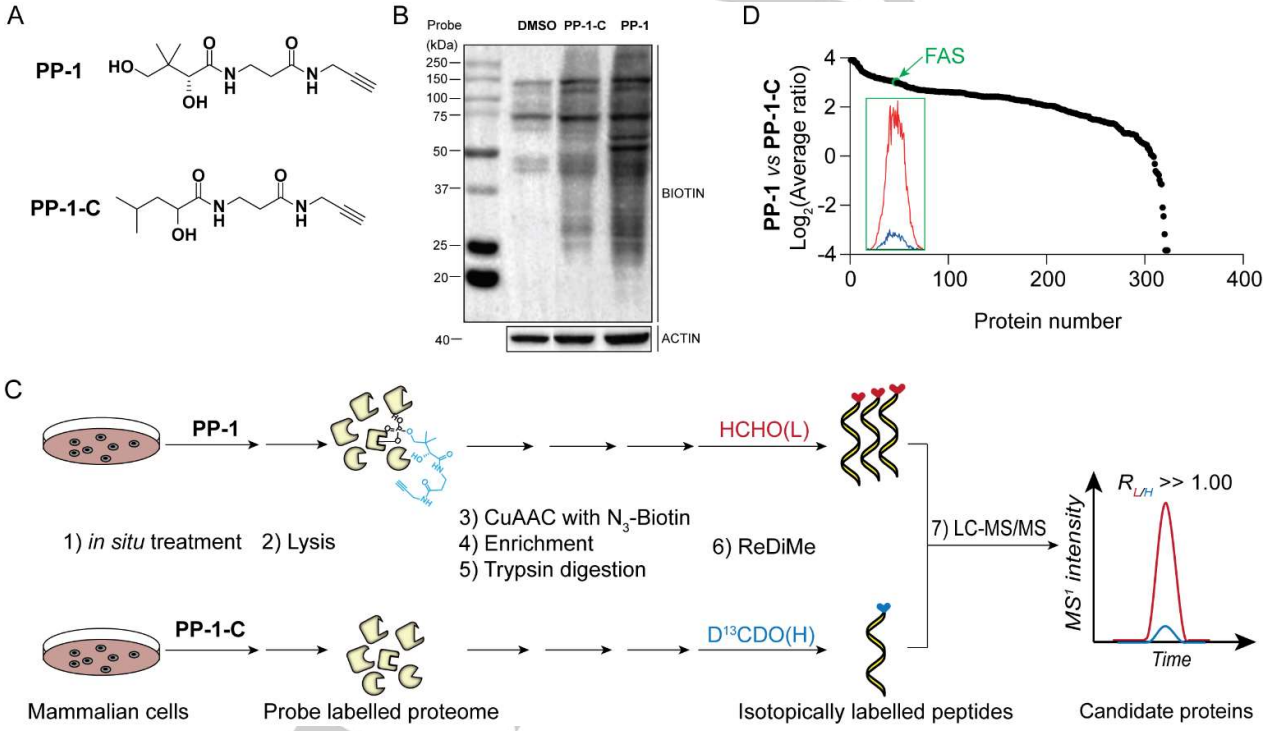



**Figure 1.** Quantitative MS-based identification of protein targets labelled by the pantetheine analogue probe, **PP-1**. (A) Chemical structures of pantetheine analogue probe, **PP-1** and the control probe, **PP-1-C**. (B) Evaluation of metabolic labelling by **PP-1** and **PP-1-C** in HepG2 cells. Labelling signals were detected by immunoblotting against biotin. The signals from DMSO-treated samples result from endogenous biotinylated proteins and they could also be observed in both **PP-1-C** and **PP-1**-treated samples. The other signals from **PP-1-C**-treated samples are likely caused by unintended metabolic labelling. (C) Schematic workflow of the quantitative chemical proteomic platform to profile protein targets labelled by the pantetheine analogue probe. CuAAC, copper catalyzed alkyne-azide cycloaddition; $N_3$-Biotin, biotin-azide; ReDiMe, reductive demethylation; $R_{L/H}$, exacted $MS^1$ ion intensities ratio of the peptide labelled by 'Light' HCHO *vs* by 'Heavy' $D^{13}CDO$. (D) Average ratios of the identified proteins from HepG2 cells treated with **PP-1** *vs* **PP-1-C** (n = 3). A representative peptide (R.SLLVNPEGPTLMR.L) $MS^1$ chromatographic peak for FAS is shown as insets ($R_{L/H}$ = 9.75). Green point denotes the rank of FAS in the ratio distribution curve.

We next established a quantitative chemical proteomic strategy to identify protein targets labelled by **PP-1**. As shown in Figure 1C, HepG2 cells treated with **PP-1** and **PP-1-C**, respectively, were subjected to cell lysis, conjugation with biotin-azide through copper catalyzed alkyne-azide cycloaddition (CuAAC)[12], affinity purification with streptavidin-beads and trypsin digestion. The released peptides were isotopically labelled by reductive dimethylation (ReDiMe)[13] prior to combination for nano liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis. Proteins were quantified based on their extracted $MS^1$ ion intensities and the **PP-1**-labelled proteins were expected to show high enrichment ratios *vs* those labeled by **PP-**

**1-C**. In total, 323 proteins were profiled if they are commonly quantified twice from the three replicates that were obtained by two independent batches of the **PP-1** probe (Figure S2, Table S2). Among these proteins, we found only one positive control substrate protein with known 4'-phosphopantetheinylation, FAS, which has an average enrichment ratio of 8.10 for **PP-1** *vs* **PP-1-C** (Figure 1D and S3). The result suggested that the alkynyl probe could be metabolically incorporated into mammalian proteomes, however, its labeling efficiency might not be very high since the other four positive substrate proteins were not detected. Moreover, we identified a number of proteins that have enrichment ratios similar to or even higher than that of FAS. As many of these proteins have been functionally well characterized before but without obvious connections with 4'-phosphopantetheinylation and the enrichment ratios showed some noticeable variations across different batches (Table S2), we therefore suspected that the **PP-1** probe might undergo certain conversions in cells, resulting in unintended metabolic labelling.

## Development of new pantetheine analogue probes with improved labelling and profiling performance

In addition to its use as bioorthogonal reporters for protein 4'-phosphopantetheinylation, pantetheine analogues have also been implemented as antimalarial agents previously[14]. One study reported that pantothenamides, one class of the amide analogues of pantetheine, exhibited increased potency towards human malaria parasites when the serum pantetheniase was inactivated.[15] Pantetheinase is an ectoenzyme widely expressed in liver, kidney and gut, which accounts for the hydrolysis of pantetheine to pantothenic acid (Vitamin B5 or VB5) and cysteamine (Figure S4A).[16] Since the enzyme has a high promiscuous activity for the cysteamine group, it can break down the amide bond of pantothenamides as well.[17] Subtle structural modifications of pantothenamides by either shortening or growing one carbon atom of the middle β-alanine moiety have proved effective to counteract the degradation by pantetheinase (Figure S4B).[18]

Inspired by these findings, we hypothesized that **PP-1** might be susceptible to similar enzymatic degradation by pantetheinase, releasing a propargylamine fragment that can be potentially incorporated into proteins and cause nonspecific labelling[19] (Figure S4C). To minimize the possibility of such an undesired labelling, we designed and synthesized two new pantetheine analogue probes, namely **PP-2** and **PP-3**, by replacing the β-alanine moiety with glycine and γ-aminobutyric acid, respectively (Figure 2A). Both probes showed similar cytotoxicity to that of **PP-1** in HepG2 cells (Figure S5A). When we evaluated their ability of metabolic labelling by immunoblotting, we found that the proteome labelling signals of **PP-2** and **PP-3** increased significantly compared to that of **PP-1**. One notable evidence was exemplified by the top band around 250 kDa that probably resulted from the dramatically improved labelling of FAS (the protein size is ~270 kDa) (Figure 2B).

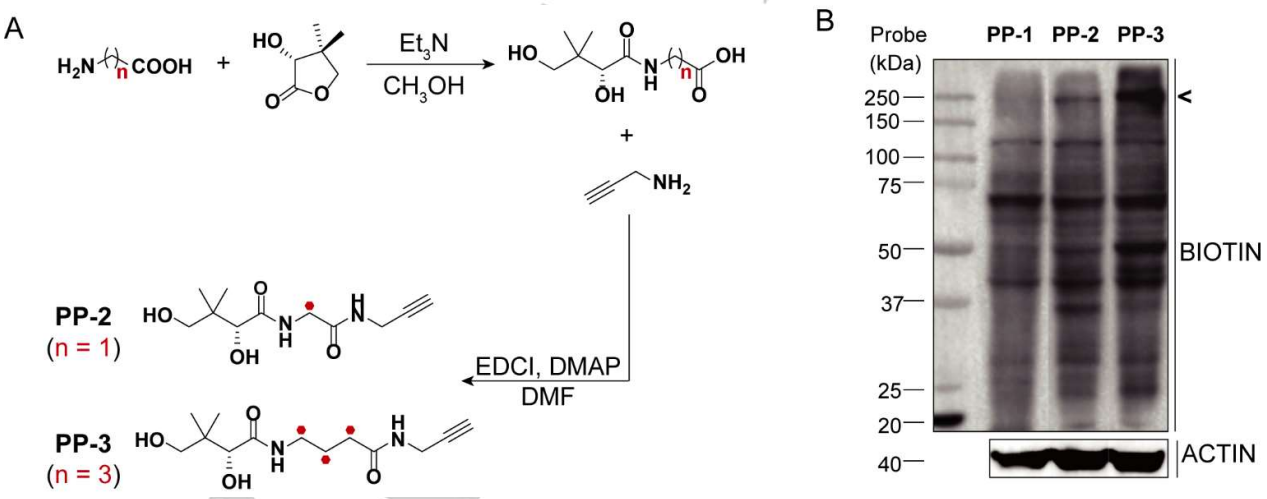


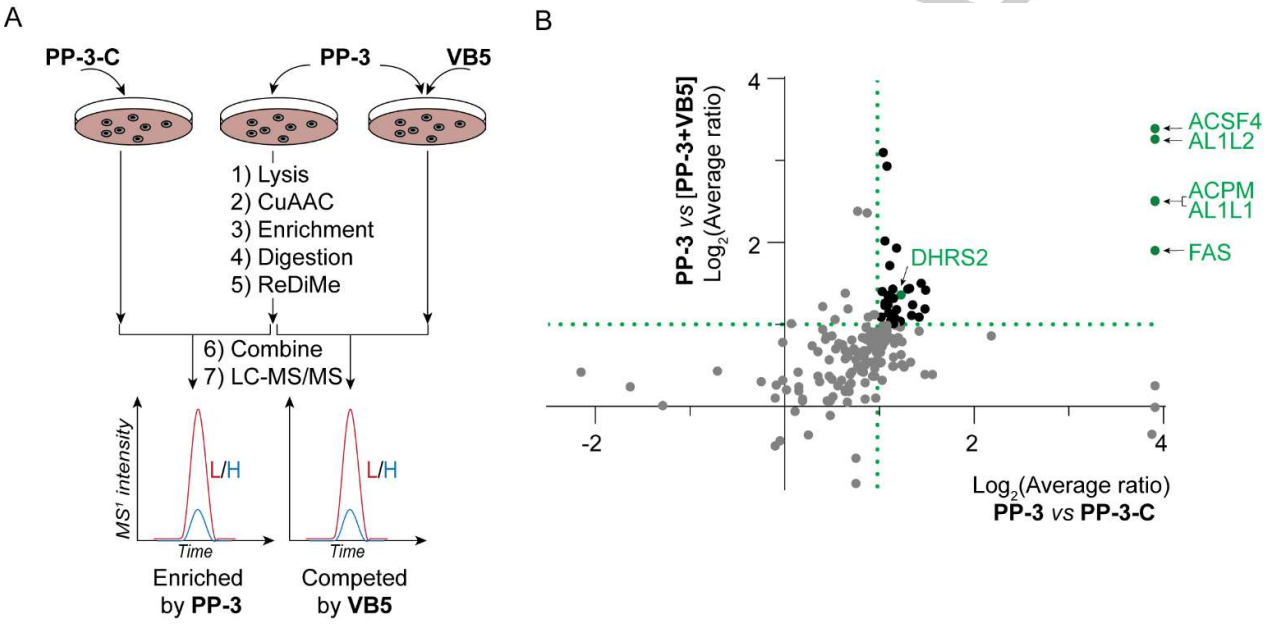

**Figure 2.** Development of new pantetheine analogue probes with improved metabolic labelling efficiency. (A) Synthetic routes and chemical structures of the newly developed pantetheine analogue probes. The β-alanine moiety of **PP-1** was modified by deducting or adding one carbon atom to yield **PP-2** and **PP-3**, respectively. Methanol solution containing glycine (n = 1) or γ-aminobutyric acid (n = 3) was allowed to refluxed with *D*-(-)-pantolactone in the presence of trimethylamine ($Et_3N$). The generated pantothenate analogues were conjugated with propargylamine to afford **PP-2** (n = 1) and **PP-3** (n = 3), respectively. EDCI, *N*-(3-Dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride; DMAP, 4-dimethylaminopyridine; DMF, dimethylformamide. (B) Immunoblotting showed significantly improved labelling by **PP-2** and **PP-3** in HepG2 cells. A band around 250 KDa that is more likely to correspond to the labelling of FAS is indicated.

Encouraged by these results, we performed the MS-based chemical proteomic experiment to identify the proteins labelled by **PP-3**, which showed slightly stronger proteome labelling signals by immunoblotting. With the ReDiMe-based quantitative chemical proteomic platform (Figure 3A), we enriched and quantified a total of 427 proteins from **PP-3**-treated HepG2 cells in at least two biological replicates (Figure S6A and Table S3) with **PP-3-C** (Figure S5B) as the inactive control probe. In contrast to the profiling results using **PP-1**, all five of the 4'-phosphopantetheinylated positive-control proteins were satisfyingly quantified with the 'singleton' enrichment ratio of 15.0 (Figure 3B-C and S7A), which made them well distinguished from

10.1002/anie.202004105




the rest of population. This result fully supported our hypothesis that **PP-3** could label true 4'-phosphopantetheinylated targets with enhanced efficiency by resisting pantetheinase degradation and minimizing "off-target" labeling. In order to further confirm that **PP-3** is able to mimic natural 4'-phosphopantetheinylation, we next performed a competitive chemical proteomic experiment with the natural VB5 as the competitor (Figure 3A, S6B and Table S3). As expected, all the positive control proteins showed strong competition ratios (Figure 3B-C and S7B). Collectively, we developed a new pantetheine analogue probe with significantly improved efficiency of metabolic labelling in mammalian cells, which enables enrichment and identification of all the known 4'-phosphopantetheinylated proteins in their endogenous forms directly from mammalian proteomes.




C

| Protein name | FAS | ACPM | AL1L1 | AL1L2 | ACSF4 |
|---|---|---|---|---|---|
| $R_{L/H}$ | 15.00 | 15.00 | 15.00 | 15.00 | 15.00 |
| **PP-3** *vs* **PP-3-C** | | | | | |
| $R_{L/H}$ | 3.57 | 5.28 | 6.99 | 10.43 | 6.12 |
| **PP-3** *vs* **[PP-3+VB5]** | | | | | |

**Figure 3.** MS-based quantitiatve profiling of 4'-phosphopantetheinylated protein targets by **PP-3** in HepG2 cells. (A) Scheme for the probe enrichment and competition experiments to profile 4'-phosphopantetheinylated protein targets using **PP-3**. (B) Scatter plot of the enrichment (X-axis) and competition (Y-axis) ratios quantified from the profiling experiments. The ratios are averaged for each protein quantified in the direct profiling experiments (**PP-3** *vs* **PP-3-C**, n = 3) and the competitive profiling experiments (**PP-3** *vs* [**PP-3 + VB5**], n = 2), respectively. The proteins within the cut-off lines (green lines) show both the enrichment and competition ratios above 2.00. All five of the known 4'-phosphopantetheinylated proteins (FAS, ACPM, AL1L1, AL1L2 and ACSF4) and a potential candidate (DHRS2) are denoted. (C) Representative peptide $MS^1$ chromatograms for all the five known 4'-phosphopantetheinylated proteins which showed strong enrichement and competition ratios in the profiling experiments .

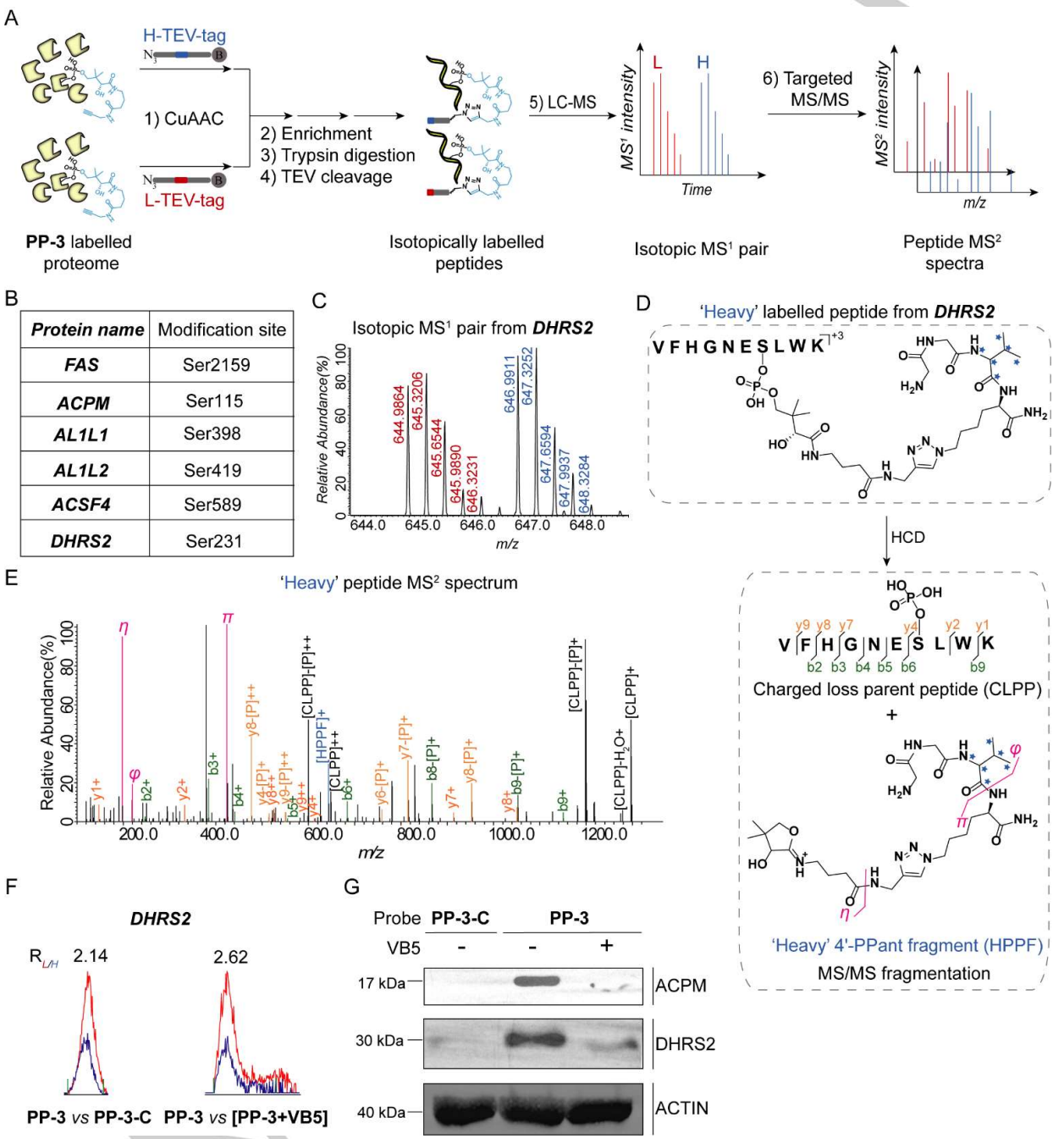
## Identification of 4'-phosphopantetheinylation sites by targeted chemical proteomics

One notable feature in Figure 3B is that only the five currently known 4'-phosphopantetheinylated proteins showed outstanding ratios in both the probe enrichment (**PP-3** *vs* **PP-3-C**) and competition experiments (**PP-3** *vs* **[PP-3 + VB5]**), while the rest of quantified targets failed to meet either of the cutoffs. Though this might be used as an empirical criterion to distinguish true 4'-phosphopantetheinylated protein candidates from false positives, the ultimate evidence to support the existence of 4'-phosphopantetheinylation on a substrate protein is to identify the

10.1002/anie.202004105




exact site of modification by MS/MS. We therefore proceeded to employ a tandem orthogonal proteolysis strategy[20] to perform site-specific identification of 4'-phosphopantetheinylation using **PP-3** (Figure 4A). Lysates from the **PP-3**-treated cells were divided into two parts and were conjugated by CuAAC with equal amounts of isotopically differentiated TEV (tobacco etch virus protease)-cleavable biotin-azide tags ("H-TEV-tag" and "L-TEV-tag"), respectively. After subsequent combination, enrichment




| *Protein name* | Modification site |
|---|---|
| ***FAS*** | Ser2159 |
| ***ACPM*** | Ser115 |
| ***AL1L1*** | Ser398 |
| ***AL1L2*** | Ser419 |
| ***ACSF4*** | Ser589 |
| ***DHRS2*** | Ser231 |

**Figure 4.** Identification of sites of 4'-phosphopantetheinylation via targeted chemical proteomics. (A) Scheme for identification of sites of 4'-phosphopantetheinylation by targeted chemical proteomics in combination with an isotopic tandem orthogonal proteolysis strategy. (B) List of six sites of 4'-phosphopantetheinylation identified in the experiment. Ser, serine residue. (C) Representative MS¹ spectrum of the 4'-phosphopantetheinylated peptide (VFHGNES*LWK, S* is the modified serine residue) in DHRS2. The MS¹ spectrum displays an isotopic pair with mass *vs* charge (*m/z*) in red for 'Light' labelled peptide and those in blue for 'Heavy' labelled peptide. (D) Chemical structures of the 'Heavy' labelled peptide from DHRS2 and its fragments generated by HCD (higher-energy collisional dissociation). All detected *b*, *y* ions from the charged loss parent peptide (CLPP) and characteristic fragment ions (φ, π and η) from the 'Heavy' 4'-PPant fragment (HPPF) are indicated. $^{13}C$ and $^{15}N$ atoms are denoted by blue star symbols. (E) The representative MS² spectrum corresponding to the 'Heavy' labelled peptide from DHRS2 shows the fragmentation ions indicated in Figure 3D. [P], [phosphate]. (F) Representative peptide MS¹ chromatograms of DHRS2 from the probe enrichment and competition profiling experiments shown in Figure 3A. (G) Immunoblots showing selective metabolic labelling of ACPM and DHRS2 by **PP-3** which can be strongly competed by VB5. (n = 2)

and tandem trypsin/TEV proteolytical digestions on beads, the peptides harboring the **PP-3**-derived modifications on serine residues that mimic 4'-phosphopantetheinylation were obtained. More importantly, only the modified peptides should appear in a light/heavy pair as the isotopic labels were introduced into TEV tags. We reasoned that isotopic pairs could facilitate peptide identification by targeted chemical proteomics, a strategy that has been applied to improve the detection of glycosylated peptides[21], selenopeptides[22] as well as drug-binding sites[23]. We analyzed one aliquot of the sample first by the regular shotgun proteomic method and processed the raw data by an in-house software (as available in https://github.com/wangchulab/CIMAGE_Lite) to detect the isotopic patterns with equal light/heavy pair in $MS^1$ spectra. We then generated an inclusion list of all the detected isotopic peptides pairs (Table S4) and directed them to the targeted LC-MS/MS analysis on the next replicated sample.

The data were searched with 'Light' and 'Heavy' variable modifications on serine residues with MSFragger[24] (Figure S8). However, the scores of MS/MS spectra were low due to poorly matched b/y fragment ions. Previous studies of 4'-phosphopantetheinylation on carrier proteins have observed characteristic "neutral loss" of the 4'-phosphopantetheine group ('4'-PPant fragment') during MS/MS fragmentation, which further suggested the use of its $MS^3$ fragmentation pattern (ion "*a-l*" as originally designated) as a diagnostic tool to detect 4'-phosphopantetheinylation (Figure S9A).[25] In light of such observations, we manually checked the MS/MS spectra by pLabel[26] using a phosphorylation modification on serine residues instead. Satisfyingly, we identified a total of 6 unique modification sites with high confidence scores, which include all five of the known 4'-phosphopantetheinylation sites (Figure 4B-C and S10). To our best knowledge, these results highlighted the first successful example of site-specific identification of 4'-phosphopantetheinylation by chemoproteomic methods, which largely benefit from the significantly improved metabolic labelling efficiency of the **PP-3** probe.

Interestingly, the site-specific profiling using the **PP-3** probe identified a previously unreported modification site on Ser231 of mitochondrial dehydrogenase/reductase SDR family member 2 (DHRS2) (Figure 4B). The $MS^1$ spectra of the modified peptide, VFHGNES*LWK (S* is the probe-modified serine residue), showed a characteristic light/heavy pair as expected (Figure 4C). In addition to the corresponding b/y fragment ions, we also observed a few common ion peaks within the low *m/z* range (*m/z* 100-700) from the $MS^2$ spectra which can be consistently mapped to the isotopically labelled "4'-PPant fragment" and their associated fragments (Figure 4D-E). For example, in one representative $MS^2$ spectrum generated by higher-energy collisional dissociation (HCD) fragmentation, we could confidently assign the peaks at *m/z* 643.3911 (Cal'd. 643.3918) as the 'Heavy' 4'-PPant fragment (HPPF). Additionally, fragment ions (η, π and φ, corresponding chemical structures shown in Figure S9B) from the HPPF could be observed as well.

DHRS2 belongs to the short-chain dehydrogenase/reductase (SDR) superfamily that usually catalyze NAD(P)(H) dependent reactions with a large array of substrates including polyols, steroids, fatty acid derivatives and xenobiotics,[27] however, its detailed molecular function remains elusive. Based on the structural homology to DHRS4, Ser231 of DHRS2 is located in the short helices (αFG2) right above the substrate-binding site[28]. Although DHRS2 showed relatively lower enrichment (2.34) and competition (2.56) ratios in the protein-based ReDiMe profiling experiments, respectively (Figure 3B and 4F), these numbers are reliably quantified (as averaged from the quantified peptides across all the replicates) and statistically significant (*p*-value = 1.01E-11 and 1.19E-6 with $R_{L/H}$ = 1.0 as the reference) (Table S3 and Figure S7), suggesting the protein was indeed enriched by the **PP-3** probe and the labelling could be competed by excessive VB5 treatment. Consistently, when we tested the **PP-3** probe labelling of DHRS2 and its competition by VB5 using affinity purification and immunoblotting, we observed that the protein behaved similarly as the positive control ACPM (Figure 4G and S11). Collectively, these data suggest that DHRS2 might be a potential 4'-phosphopantetheinylated protein with the modification on Ser231. However, it remains to be explored whether this site has a native 'prosthetic arm' and if so, what is the functional relevance of this modification.

Taken together, a total of six 4'-phosphopantetheinylation sites, including a potential one in DHRS2, were identified by targeted chemical proteomics. The number is still much less than that of potential candidates identified from the ReDiMe experiments with both the enrichment and competition ratios above 2.00 (Table S3). Multiple reasons may account for this discrepancy. Firstly, multiple peptides per protein from the trypsin digestion are available for identification while there is normally only one 4'-phosphopantetheinylated peptide from each candidate protein after the TEV cleavage. Secondly, as peptides with length of 6-25 amino acids are mostly suitable for standard LC-MS/MS identification, the "4'-PPant arm" might create additional challenge when it is attached to peptides that are too short or long, resulting in reduced 'detectability'. In fact, the modified peptides of FAS and ACPM after trypsin digestion have relatively long lengths (25 amino acids for FAS and 31 for ACPM, respectively) and were fortunately eluted at the end of gradient (>35% acetonitrile) for MS/MS analysis. Thirdly, as exemplified by DHRS2, substrates with low protein turnover rate or low modification stoichiometry will be metabolically labelled with reduced efficiency, resulting in the seemingly low enrichment and competition ratios, as well as insufficient amount of enriched peptides for site identification. Lastly, the possibility cannot be excluded that some of the candidates might still result from nonspecific or off-target labelling of **PP-3** even after it is optimized, and biochemical validation of native modifications on these potential substrates should be rigorously performed before their functional studies are taken.

## Conclusion

In summary, we have developed a new generation of unnatural pantetheine analogues for effective metabolic labelling of protein 4'-phosphopantetheinylation in mammalian cells. Since FAS was first determined as a 4'-phosphopantetheinylated protein in mammalian cells, it took almost half a century to discover the rest of four protein substrates of 4'-

phosphopantetheinylation that are known to date. In combination with advanced chemical proteomic technologies, our work here provides an all-in-one snapshot of these known 4'-phosphopantetheinylated proteins with the resolution of exact modification sites and suggests potential candidates such as DHRS2. Methodologically, sample preparations need further optimization (e.g. auxiliary proteases for peptide digestion, new cell lines and longer LC gradients) to increase the probability of identifying new 4'-phosphopantetheinylation sites. Biologically, our work opens new opportunities to study the functional role of 4'-phosphopantetheinylation on these substrate enzymes and the biological pathway they are involved in. Chemically, it is foreseeable that our probe can aid development of new tool compounds that aims to perturb the state of this functional modification. Collectively, we envision that the analytical and functional study of protein 4'-phosphopantetheinylation will be greatly accelerated with this powerful chemical biology tool.

## Acknowledgements

We thank the Computing Platform of the Center for Life Science for supporting the LC-MS/MS data analysis. This work was supported by Ministry of Science and Technology of China (2016YFA0501500), National Natural Science Foundation of China (21925701, 91953109 and 21778004). There are no conflicts of interest.

**Keywords:** 4′-phosphopantetheinylation • chemoproteomics • pantetheine analogue probe • site identification • mammalian cell

[1] F. L. Theodoulou, O. C. Sibon, S. Jackowski, I. Gout, *Biochem. Soc. Trans.* **2014**, *42*, 1025-1032.

[2] J. Beld, E. C. Sonnenschein, C. R. Vickery, J. P. Noel, M. D. Burkart, *Nat. Prod. Rep.* **2014**, *31*, 61-108.

[3] a) P. W. Majerus, A. W. Alberts, P. R. Vagelos, *Proc. Natl. Acad. Sci. U S A* **1965**, *53*, 410-417; b) A. R. Larrabee, E. G. McDaniel, H. A. Bakerman, P. R. Vagelos, *Proc. Natl. Acad. Sci. U S A* **1965**, *54*, 267-273.

[4] J. L. Meier, M. D. Burkart, *Chem. Soc. Rev.* **2009**, *38*, 2012-2045.

[5] J. E. Cronan, I. M. Fearnley, J. E. Walker, *FEBS Lett* **2005**, *579*, 4892-4896.

[6] a) H. Donato, N. I. Krupenko, Y. Tsybovsky, S. A. Krupenko, *J. Biol. Chem.* **2007**, *282*, 34159-34166; b) N. I. Krupenko, M. E. Dubard, K. C. Strickland, K. M. Moxley, N. V. Oleinik, S. A. Krupenko, *J. Biol. Chem.* **2010**, *285*, 23056-23063.

[7] J. Drozak, M. Veiga-da-Cunha, B. Kadziolka, E. Van Schaftingen, *Febs j.* **2014**, *281*, 1585-1597.

[8] a) K. M. Clarke, A. C. Mercer, J. J. La Clair, M. D. Burkart, *J. Am. Chem. Soc.* **2005**, *127*, 11234-11235; b) J. L. Meier, A. C. Mercer, H. Rivera, Jr., M. D. Burkart, *J. Am. Chem. Soc.* **2006**, *128*, 12174-12184.

[9] C. Nguyen, R. W. Haushalter, D. J. Lee, P. R. L. Markwick, J. Bruegger, G. Caldara-Festin, K. Finzel, D. R. Jackson, F. Ishikawa, B. O'Dowd, J. A. McCammon, S. J. Opella, S.-C. Tsai, M. D. Burkart, *Nature* **2014**, *505*, 427-431.

[10] A. C. Mercer, J. L. Meier, J. W. Torpey, M. D. Burkart, *Chembiochem* **2009**, *10*, 1091-1100.

[11] K. G. Virga, Y. M. Zhang, R. Leonardi, R. A. Ivey, K. Hevener, H. W. Park, S. Jackowski, C. O. Rock, R. E. Lee, *Bioorg. Med. Chem.* **2006**, *14*, 1007-1020.

[12] H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed. Engl.* **2001**, *40*, 2004-2021.

[13] P. J. Boersema, R. Raijmakers, S. Lemeer, S. Mohammed, A. J. Heck, *Nat. Protoc.* **2009**, *4*, 484-494.

[14] C. Spry, K. Kirk, K. J. Saliba, *FEMS Microbiol. Rev.* **2008**, *32*, 56-106.

[15] C. Spry, C. Macuamule, Z. Lin, K. G. Virga, R. E. Lee, E. Strauss, K. J. Saliba, *PLoS One* **2013**, *8*, e54974.

[16] Y. L. Boersma, J. Newman, T. E. Adams, N. Cowieson, G. Krippner, K. Bozaoglu, T. S. Peat, *Acta Crystallogr. D Biol. Crystallogr.* **2014**, *70*, 3320-3329.

[17] C. Wittwer, B. Wyse, R. G. Hansen, *Anal. Biochem.* **1982**, *122*, 213-222.

[18] M. de Villiers, C. Macuamule, C. Spry, Y. M. Hyun, E. Strauss, K. J. Saliba, *ACS Med. Chem. Lett.* **2013**, *4*, 784-789.

[19] a) L. Lorand, R. M. Graham, *Nat. Rev. Mol. Cell Biol.* **2003**, *4*, 140-156; b) K. Ikura, K. Kita, I. Fujita, H. Hashimoto, N. Kawabata, *Arch Biochem. Biophys.* **1998**, *356*, 280-286; c) J. E. Folk, M. H. Park, S. I. Chung, J. Schrode, E. P. Lester, H. L. Cooper, *J. Biol. Chem.* **1980**, *255*, 3695-3700.

[20] E. Weerapana, A. E. Speers, B. F. Cravatt, *Nat. Protoc.* **2007**, *2*, 1414-1425.

[21] C. M. Woo, A. T. Iavarone, D. R. Spiciarich, K. K. Palaniappan, C. R. Bertozzi, *Nat. Methods* **2015**, *12*, 561-567.

[22] J. Gao, F. Yang, J. Che, Y. Han, Y. Wang, N. Chen, D. W. Bak, S. Lai, X. Xie, E. Weerapana, C. Wang, *ACS Cent. Sci.* **2018**, *4*, 960-970.

[23] J. Gao, A. Mfuh, Y. Amako, C. M. Woo, **2018**, *140*, 4259-4268.

[24] A. T. Kong, F. V. Leprevost, D. M. Avtonomov, D. Mellacheruvu, A. I. Nesvizhskii, *Nat. Methods* **2017**, *14*, 513-520.

[25] a) D. Meluzzi, W. H. Zheng, M. Hensler, V. Nizet, P. C. Dorrestein, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 3107-3111; b) J. L. Meier, A. D. Patel, S. Niessen, M. Meehan, R. Kersten, J. Y. Yang, M. Rothmann, B. F. Cravatt, P. C. Dorrestein, M. D. Burkart, V. Bafna, *J. Proteome Res.* **2011**, *10*, 320-329.

[26] L. H. Wang, D. Q. Li, Y. Fu, H. P. Wang, J. F. Zhang, Z. F. Yuan, R. X. Sun, R. Zeng, S. M. He, W. Gao, *Rapid Commun. Mass Spectrom.* **2007**, *21*, 2985-2991.

[27] J. E. Bray, B. D. Marsden, U. Oppermann, *Chem. Biol. Interact.* **2009**, *178*, 99-109.

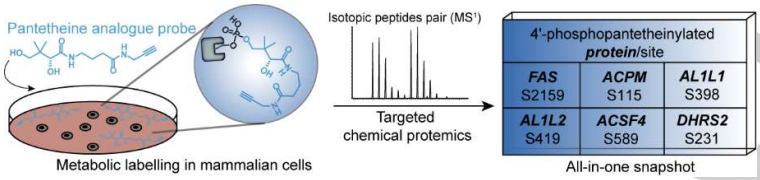
[28] F. Gabrielli, S. Tofanelli, *Gene* **2012**, *511*, 461-469.

10.1002/anie.202004105

WILEY-VCH

RESEARCH ARTICLE

**Entry for the Table of Contents**




By developing a new generation of pantetheine analogue probes, we achieved the metabolic labelling and chemoproteomic profiling of protein 4'-phosphopantetheinylation in mammalian cells for the first time. Targeted cheimcal proteomics facilitated identification of exact sites of 4'-phosphopantetheinylation that included all five of the currently known ones and a potential one in the protein of mitochondrial dehydrogenase/reductase SDR family member 2 (DHRS2).

*Nan Chen, Yuan Liu, Yuanpei Li and Chu Wang**

**Chemical proteomic profiling of protein 4'-phosphopantetheinylation in mammalian cells**