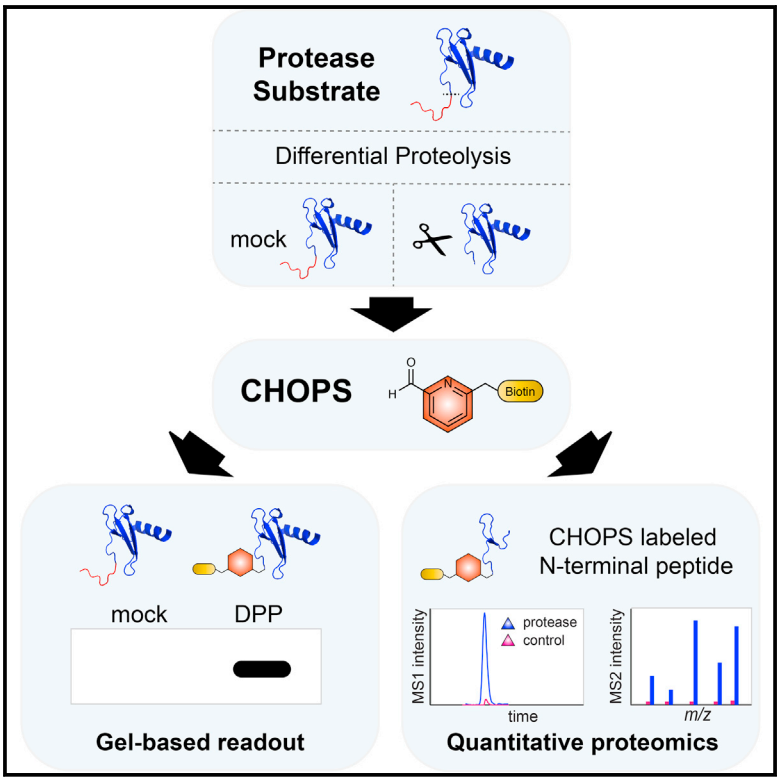
Brief Communication

Cell Chemical Biology

A Chemical Strategy for Protease Substrate Profiling

Graphical Abstract



Authors

Andrew R. Griswold, Paolo Cifani, Sahana D. Rao, ..., Ronald C. Hendrickson, Alex Kentsis, Daniel A. Bachovchin

Correspondence

bachovcd@mskcc.org

In Brief

Proteases regulate countless (patho) physiological processes, but the identification of protease substrates is challenging. Here, Griswold et al. introduce a simple chemoproteomic strategy, termed CHOPS, for profiling protease substrates. Using CHOPS, the authors identify the cleavage specificities of proteases in cellular lysates and show that DPP9 preferentially processes short peptides.

Highlights

- A 2PCA-biotin probe enables chemical enrichment of protease substrates (CHOPS)
- CHOPS rapidly identifies DPP substrates with a gel-based readout
- DPP9 preferentially cleaves short peptides, not whole proteins (e.g., Nlrp1b)
- CHOPS enables unbiased protease substrate profiling with quantitative proteomics

Griswold et al., 2019, Cell Chemical Biology *26*, 1–7
June 20, 2019
https://doi.org/10.1016/j.chembiol.2019.03.007

CellPress

Please cite this article in press as: Griswold et al., A Chemical Strategy for Protease Substrate Profiling, Cell Chemical Biology (2019), https://doi.org/10.1016/j.chembiol.2019.03.007

Cell Chemical Biology

Brief Communication

CellPress

A Chemical Strategy for Protease Substrate Profiling

Andrew R. Griswold,[1,2] Paolo Cifani,[3] Sahana D. Rao,[4] Abram J. Axelrod,[5] Matthew M. Miele,[6] Ronald C. Hendrickson,[6] Alex Kentsis,[2,3,4,7] and Daniel A. Bachovchin[2,4,5,8,*]

[1]Weill Cornell/Rockefeller/Sloan Kettering Tri-Institutional MD-PhD Program, New York, NY 10065, USA
[2]Pharmacology Program, Weill Cornell Graduate School of Medical Sciences, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
[3]Molecular Pharmacology Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
[4]Tri-Institutional PhD Program in Chemical Biology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
[5]Chemical Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
[6]Proteomics Core Laboratory, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
[7]Department of Pediatrics, Memorial Sloan Kettering Cancer Center and Weill Medical College of Cornell University, New York, NY 10065, USA
[8]Lead Contact
*Correspondence: bachovcd@mskcc.org
https://doi.org/10.1016/j.chembiol.2019.03.007

SUMMARY

The dipeptidyl peptidases (DPPs) regulate hormones, cytokines, and neuropeptides by cleaving dipeptides after proline from their amino termini. Due to technical challenges, many DPP substrates remain unknown. Here, we introduce a simple method, termed CHOPS (chemical enrichment of protease substrates), for the discovery of protease substrates. CHOPS exploits a 2-pyridinecarboxaldehyde (2PCA)-biotin probe, which selectively biotinylates protein N-termini except those with proline in the second position. CHOPS can, in theory, discover substrates for any protease, but is particularly well suited to discover canonical DPP substrates, as cleaved but not intact DPP substrates can be identified by gel electrophoresis or mass spectrometry. Using CHOPS, we show that DPP8 and DPP9, enzymes that control the Nlrp1 inflammasome through an unknown mechanism, do not directly cleave Nlrp1. We further show that DPP9 robustly cleaves short peptides but not full-length proteins. More generally, this work delineates a practical technology for identifying protease substrates, which we anticipate will complement available "N-terminomic" approaches.

INTRODUCTION

The DPP4 activity and/or structure homolog (DASH) sub-family of serine proteases, which include DPP4, DPP7, DPP8, DPP9, and FAP, have attracted significant attention as potential therapeutic targets (Adams et al., 2004; Busek et al., 2004; Lankas et al., 2005; Rosenblum and Kozarich, 2003). DASH enzymes share the rare ability to cleave after proline residues in the second position of polypeptide substrates. DPP4, the best characterized DASH enzyme, cleaves and regulates the activity of dozens of biologically important peptides, including neuropeptides, chemokines, and incretins (Mulvihill and Drucker, 2014), and DPP4 inhibitors are approved anti-diabetic drugs (Deacon and Lebovitz, 2016). However, many critical substrates of DASH enzymes, including substrates of DPP4, are unknown (Mulvihill and Drucker, 2014; Tagore et al., 2009; Waumans et al., 2015). For example, DPP8 and DPP9 act as an intracellular checkpoint to restrain the Nlrp1 inflammasome (Okondo et al., 2017, 2018), but the key substrate that controls inflammasome activation has not been identified.

DPPs remain poorly characterized in large part due to technical challenges in identifying endogenous substrates (Mulvihill and Drucker, 2014; Tagore et al., 2009; Tinoco et al., 2010; Wilson et al., 2016; Yates et al., 2007). Intact and cleaved DPP substrates are similar in size and typically inseparable by gel electrophoresis, and thus gel-based platforms that exploit size differences cannot be used for DPP characterization (Dix et al., 2008; Shao et al., 2007). Moreover, DPPs recognize the free N-terminal amines of their substrates (Green et al., 2004; Rasmussen et al., 2003; Ross et al., 2018), limiting the utility of approaches that involve N-terminal substrate modification before protease digestion (Tonge et al., 2001; Zhang et al., 2015). Mass spectrometry (MS)-based global peptide profiling (Jost et al., 2009; Tagore et al., 2009; Tammen et al., 2008; Tinoco et al., 2010, 2011; Yates et al., 2007) and N-terminomics (Kleifeld et al., 2010; Wilson et al., 2013) methodologies have been used to measure changes in intact and/or cleaved peptides in response to DPP modulation. However, these methods do not enrich substrates, and, as such, low abundance peptides are often difficult to detect (Tagore et al., 2009; Yates et al., 2007).

The Francis group recently published an amine-reactive chemical group, 2-pyridinecarboxyaldehyde (2PCA, Figure 1A), which we hypothesized could be used to label and enrich DPP substrates from complex lysates (MacDonald et al., 2015). 2PCA condenses with N-terminal amino groups to form imines,

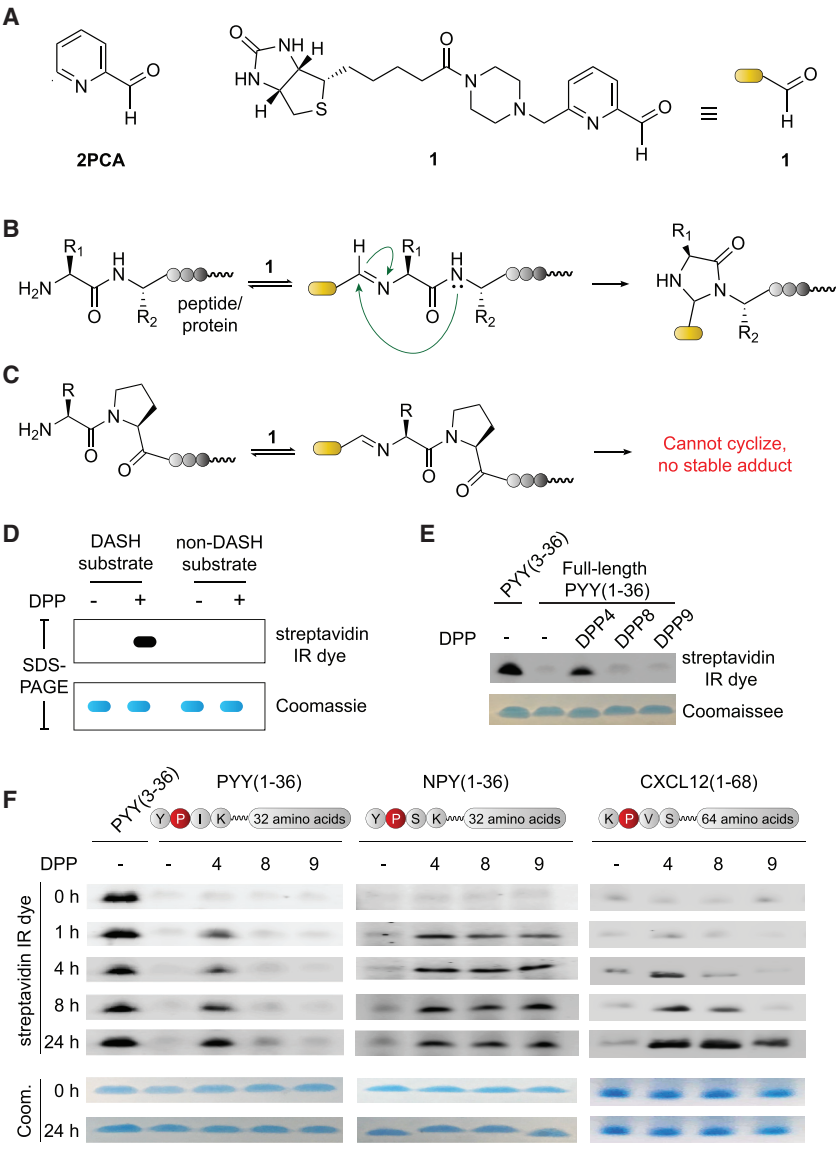


Figure 1. A Chemical Strategy for DPP Substrate Profiling

(A) The structure of 2PCA (left), and the biotin-coupled 2PCA probe **1** (right).

(B and C) **1** forms stable adducts with N-termini (B) unless proline is in position 2 (C). R_1 can be any side chain, R_2 can be any side chain except proline. The gray circles represent the remaining peptide chain.

(D) Schematic of the expected results for a DASH and a non-DASH substrate containing a proline in position 2 incubated with or without enzyme and labeled with **1**.

(E) Full-length PYY(1-36) was incubated with the indicated enzyme (10 nM) for 4 h before addition of **1** (10 mM, 16 h). Cleaved PYY(3-36) was labeled with **1** as a control. See also Figure S1.

(F) PYY, NPY, and CXCL12 (10 μM) were incubated with PBS, DPP4, 8, or 9 (10 nM) for the indicated time, before aliquots were removed, boiled, and labeled with **1** (10 mM). Blots depict probe labeling (streptavidin IR dye) and total protein (Coom.).

which then cyclize with the next nitrogen in the peptide chain to form stable 5-membered rings (Figure 1B). All N-termini are stably labeled, except for those with penultimate amino-terminal prolines, which are unable to cyclize (Figure 1C). Importantly, lysine amino groups are not stably modified by 2PCA as they lack nearby amide groups for cyclization. Here, we leverage this unique reactivity of 2PCA to label cleaved, but not intact DPP substrates (Figure 1D). We show that a biotin-coupled 2PCA-based probe (**1**, Figure 1A) can readily identify cleaved DPP substrates *in vitro* and selectively enrich cleaved DPP substrates from cell lysates. Using this platform, termed chemical enrichment of protease substrates (CHOPS), we show that DPP8 and DPP9, enzymes that restrain the Nlrp1 inflammasome, do not directly cleave Nlrp1b. We furthermore demonstrate that DPP9 has little activity against full-length proteins but robustly cleaves short peptides. Although CHOPS is particularly useful for DPP profiling, we show that it is, in theory, compatible with any protease of interest, and therefore this work lays the foundation for a technology to unbiasedly identify protease substrates in whole proteomes.

RESULTS

Visualization of DPP Proteolysis by SDS-PAGE

We initially tested whether **1** could detect the processing of DPP4 substrate peptide YY (PYY), a 36 amino acid peptide that regulates postprandial satiety. PYY exists in two forms, PYY(1-36) and PYY(3-36) (Batterham et al., 2003); the latter is generated by DPP4-mediated removal of its N-terminal NH_2-Tyr-Pro dipeptide. We observed that **1** biotinylated PYY(3-36), but not PYY(1-36), by SDS-PAGE (Figure 1E) and MALDI-TOF (Figure S1). Moreover, upon incubating PYY(1-36) with DPP4, we observed an increase in biotinylation (Figures 1E and S1). Notably, the two forms of PYY are indistinguishable by electrophoretic migration, but **1** enables PYY cleavage to be visualized by SDS-PAGE without the need for MS.

To further explore the capability of CHOPS to profile DPP substrates *in vitro*, we examined PYY(1-36) processing by DPP4, DPP8, and DPP9. After 1 h, DPP4 treatment resulted in strong labeling (Figure 1F), consistent with published kinetic data (Bjelke et al., 2006). DPP8 and DPP9 cleave PYY(1-36) much less efficiently than DPP4 (Bjelke et al., 2006), and accordingly only a slight increase in biotinylation was observed even after 24 h in DPP8- and DPP9-treated samples (Figure 1F). We next asked whether other DPP substrates were compatible with this method. Neuropeptide Y (NPY) is critical for nervous system function (Reichmann and Holzer, 2016), and the two N-terminal amino acids, NH_2-Tyr-Pro, are known to be efficiently cleaved by DPP4, DPP8, and DPP9 (Bjelke et al., 2006). Indeed, we observed that all three enzymes rapidly induced biotinylation of NPY (Figure 1F). Chemokine ligand 12 (CXCL12) is a known,

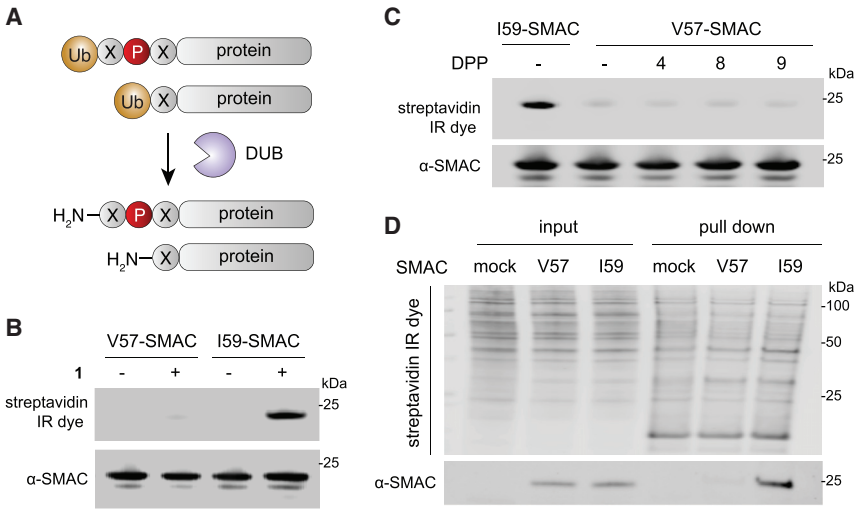


Figure 2. Characterization of Probe 1 Reactivity in Lysates

(A) Graphical depiction of ubiquitin fusion strategy. A deubiquitinase (DUB), either endogenous in HEK293T cells or recombinantly added to bacterial lysates, generates the desired N-terminal residue.

(B and C) **1** labeling (10 mM, 16 h) of purified SMAC proteins (25 μM) alone (B) or treated with the indicated DPP (10 nM, 16 h) (C).

(D) Lysates from HEK293T cells expressing the indicated SMAC proteins were harvested, labeled with **1** (10 mM, 16 h), and enriched on streptavidin agarose beads. Immunoblots depict probe labeling (streptavidin IR dye) and total SMAC protein.

albeit slowly cleaved, substrate of DPP4 and DPP8 (Ajami et al., 2008). Our SDS-PAGE assay confirmed the cleavage kinetics for DPP4 and DPP8, and further showed that DPP9, which had not been previously tested for its ability to cleave CXCL12, also slowly cleaves this substrate.

Detection of DPP Substrates in Complex Lysates

We next wanted to determine if CHOPS could identify loss of an N-terminal NH_2-Xaa-Pro dipeptide from a protein in a cellular lysate. We initially chose to evaluate the pro-apoptotic protein SMAC (Du et al., 2000; Wu et al., 2000), which is proteolytically processed in the mitochondria to reveal an N-terminal IAP-binding motif (IBM), NH_2-Ala-Val-Pro-Ile, at residue 56. We purified SMAC proteins starting with V57 (V57-SMAC) or I59 (I59-SMAC), which differ only by an N-terminal NH_2-Val-Pro dipeptide, using a ubiquitin fusion strategy (Figure 2A) (Bachmair et al., 1986). As expected, **1** strongly biotinylated purified I59-SMAC but not V57-SMAC (Figure 2B). DPP4, DPP8, and DPP9 did not increase probe labeling of V57-SMAC (Figure 2C), indicating that V57-SMAC is not a DPP substrate. Next, we expressed both SMAC variants in HEK293T cells, labeled the resulting lysates with **1**, and enriched the probe-labeled proteins using streptavidin-coupled beads. Consistent with the purified protein results, only I59-SMAC was enriched (Figure 2D). Importantly, these results demonstrate that this strategy is compatible with whole cellular proteomes.

DPP8/9 Do Not Cleave Nlrp1b

Inhibitors of DPP8/9 activate the Nlrp1 inflammasome (Okondo et al., 2017, 2018), which in turn activates caspase-1 and induces a lytic form of cell death called pyroptosis. However, the substrates of DPP8/9 that control the Nlrp1 inflammasome remain unknown. Intriguingly, Nlrp1 has a "function-to-find" (FIIND) domain, which undergoes post-translational autoproteolysis and generates N- and C-terminal polypeptide fragments that remain associated in an auto-inhibited state (D'Osualdo et al., 2011; Finger et al., 2012; Frew et al., 2012) (Figure 3A). The new N-terminus of the C-terminal fragment is NH_2-Ser-Pro, a dipeptide sequence potentially removed by DPP8/9. DPP9 was also recently reported to directly associate with Nlrp1, and the authors of this study similarly speculated, but did not directly assess, whether DPP9 cleaves the NH_2-Ser-Pro dipeptide (Zhong et al., 2018). To determine if Nlrp1 is a substrate of DPP8/9, we overexpressed two C-terminal fragments of mouse Nlrp1b (starting at S984 or M986), treated lysates from these cells with **1**, and enriched biotinylated proteins. As expected, only M986-Nlrp1b, which does not have a proline in the second position, was enriched (Figure 3B). We then treated lysates containing S984-Nlrp1b with DPP8 or DPP9, before performing a similar pull-down experiment with **1** (Figure 3C). We did not observe an increase in Nlrp1b, demonstrating that DPP8/9 do not efficiently remove the N-terminal dipeptide.

It remained possible that DPP8/9 could only cleave this peptide bond in the context of the full-length protein. To evaluate if this was the case, we harvested lysates from HEK293T cells expressing full-length Nlrp1b. Importantly, these cells were treated with Val-boroPro, a potent inhibitor of DPP4, DPP8, and DPP9 (Okondo et al., 2017), to ensure that any Nlrp1b cleavage by endogenous DPP8/9, if possible, was blocked (Figure S2). We then treated these lysates with recombinant DPP8 or DPP9, boiled the samples to separate the N- and C-terminal fragments of Nlrp1b, and performed a pull-down experiment with **1**. We observed no increase in the amount of captured C-terminal fragment (Figure 3D), further indicating that DPP8/9 do not cleave Nlrp1b. Consistent with this result, CARD8, a homolog of Nlrp1 that mediates DPP8/9 inhibitor-induced pyroptosis in human AML cells (Johnson et al., 2018), has an NH_2-Ser-Leu sequence following autoproteolysis, and therefore is unlikely to be directly cleaved by DPP8/9.

An MS-Based Platform for Proteome-wide Evaluation of Protease Substrates

We next developed an MS-coupled CHOPS assay to enable the unbiased identification of protease substrates across the entire proteome (Figure 4A). In this format, isotopically labeled cells or cell lysates are subjected to differential proteolysis, mixed, labeled with **1**, enriched on streptavidin beads, and digested with trypsin. The probe-labeled peptides are then eluted from the beads and identified and quantified by high-resolution tandem mass spectrometry. Importantly, this format, unlike the SDS-PAGE format, should enable the interrogation of all proteases, regardless of cleavage specificity (i.e., an MS readout

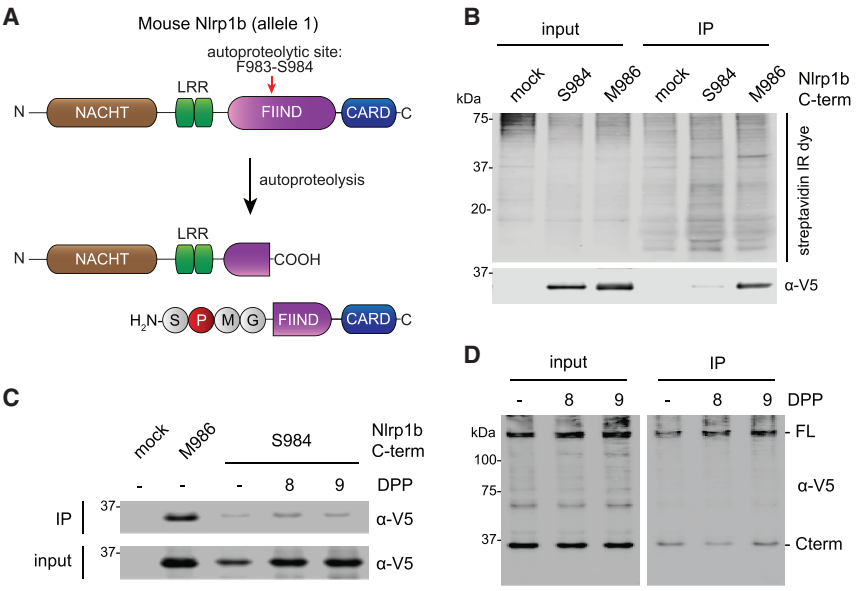


Figure 3. DPP8/9 Do Not Cleave the C-Terminal Fragment of Nlrp1b

(A) Graphical depiction of mouse Nlrp1b. The nucleotide-binding (NACHT), leucine-rich repeat (LRR), function-to-find (FIIND), and caspase activation and recruitment (CARD) domains are shown. The FIIND domain undergoes post-translational autoproteolysis generating N- and C-terminal polypeptide fragments.

(B and C) **1** labeling (10 mM, 16 h) and enrichment of HEK293T lysates ectopically expressing the indicated Nlrp1b fragment (both with C-terminal V5 tags) alone (B), or treated with the indicated DPP (10 nM, 24 h) (C). Immunoblots depict probe labeling (streptavidin IR dye) and total Nlrp1b C-terminus (V5).

(D) HEK293T cells expressing full-length Nlrp1b with a C-terminal V5 tag were grown in the presence of Val-boroPro to ensure that any Nlrp1b cleavage by DPP8/9 cleavage, if possible, was blocked. See also Figure S2. Lysates were harvested, treated with the indicated DPP (10 nM, 24 h), labeled with **1** (10 mM, 16 h), and enriched on streptavidin agarose beads. The immunoblot depicts total Nlrp1b protein (V5).

does not require exclusive labeling of just the cleaved substrate). As such, we predicted that this platform could serve as an alternative to the subtiligase-based platform (Mahrus et al., 2008), which uses an engineered enzyme (subtiligase) to conjugate a biotin-containing glycolate ester chemical probe to protein N-termini.

As an initial validation of this method, we analyzed the endogenous N-termini of HEK293T cells and THP-1 cells by mixing untreated heavy and light SILAC-labeled lysates before enriching N-termini. We identified 1902 **1**-labeled-N-termini peptides with reproducible relative abundance (i.e., similar heavy:light ratios) across two replicates (Figures 4B, S3A, and S3B; Tables S1 and S2). The frequencies of the inferred P1 residues preceding these N-termini closely matched those observed in Jurkat cells using subtiligase (Mahrus et al., 2008) (Figure 4C). Interestingly, we observed higher frequencies of several P1′ residues (Figure 4C), including Pro, Glu, Asp, and Leu, which are disfavored by subtiligase (Weeks and Wells, 2018) but efficiently labeled by 2PCA. We next treated heavy-labeled lysates with trypsin before mixing with untreated light lysates and performing CHOPS. We observed a dramatic increase in the number of heavy-enriched N-terminal-labeled peptides (Figure 4B, 3,210 in two tryptic experiments versus 24 in two untreated experiments), with nearly all following an inferred P1 Arg or Lys residue (Figures 4E and 4F). As expected, the P1′ residues largely matched the untreated sample. Together, these data demonstrate that CHOPS can profile protease substrates in lysates.

We next wanted to identify DPP9 substrates in lysates. Thus, we treated lysates from heavy-labeled *DPP8/9* KO THP-1 cells (Figure S3C) (Okondo et al., 2017) with DPP9 before mixing with untreated, light-labeled *DPP8/9* KO THP-1 lysates and performing CHOPS. We identified only nine enriched **1**-labeled N-terminal peptides (of 429) (Figures 4G and 4H). Moreover, none of these nine peptides followed an inferred P1 Pro, suggesting that DPP9 likely has little activity against full-length proteins. Consistent with these data, we observed no change in protein biotinylation as observed by SDS-PAGE (Figure S3D). Similarly, a previous study using TAILS (terminal amine isotopic labeling of substrates) found very few peptides corresponding to DPP8/9 substrates (Wilson et al., 2013). DPP8/9 can efficiently cleave short peptides *in vitro* (Geiss-Friedlander et al., 2009; Justa-Schuch et al., 2016; Wilson et al., 2013; Zhang et al., 2015), and we speculated that DPP8/9 might also cleave short peptides, but not globular proteins, in cells. We therefore trypsinized *DPP8/9* KO THP-1 lysates to generate peptides before treating the heavy-labeled proteome with DPP9 and performing CHOPS. In this experiment, we observed 80 heavy-enriched N-termini across two replicates (Figures 4G and 4H), 30 of which followed inferred P1 prolines (Figures 4H–4J). Overall, these data show that CHOPS can indeed identify DPP8/9 substrates in complex mixtures and suggest that DPP8/9 likely act as peptidases to digest small peptide substrates.

DISCUSSION

Here, we have introduced a chemical strategy to detect and/or enrich protease substrates called CHOPS. CHOPS can be used with a simple gel-based readout to evaluate N-terminal NH_2-Xaa-Pro loss from peptides or proteins, corresponding to canonical DPP cleavage. Using this gel-based method, we showed that DPP8/9 do not cleave several proteins starting with NH_2-Xaa-Pro, including SMAC and Nlrp1b. Using an MS-based readout, we extended these findings to show that DPP9 does not readily cleave full-length proteins, but instead robustly cleaves short peptides. Although DPP8/9 have been speculated to cleave several full-length proteins, including SYK and AK2 (Justa-Schuch et al., 2016; Wilson et al., 2013), all confirmatory assays have been conducted exclusively on peptides corresponding to the N-terminus of the protein rather than the

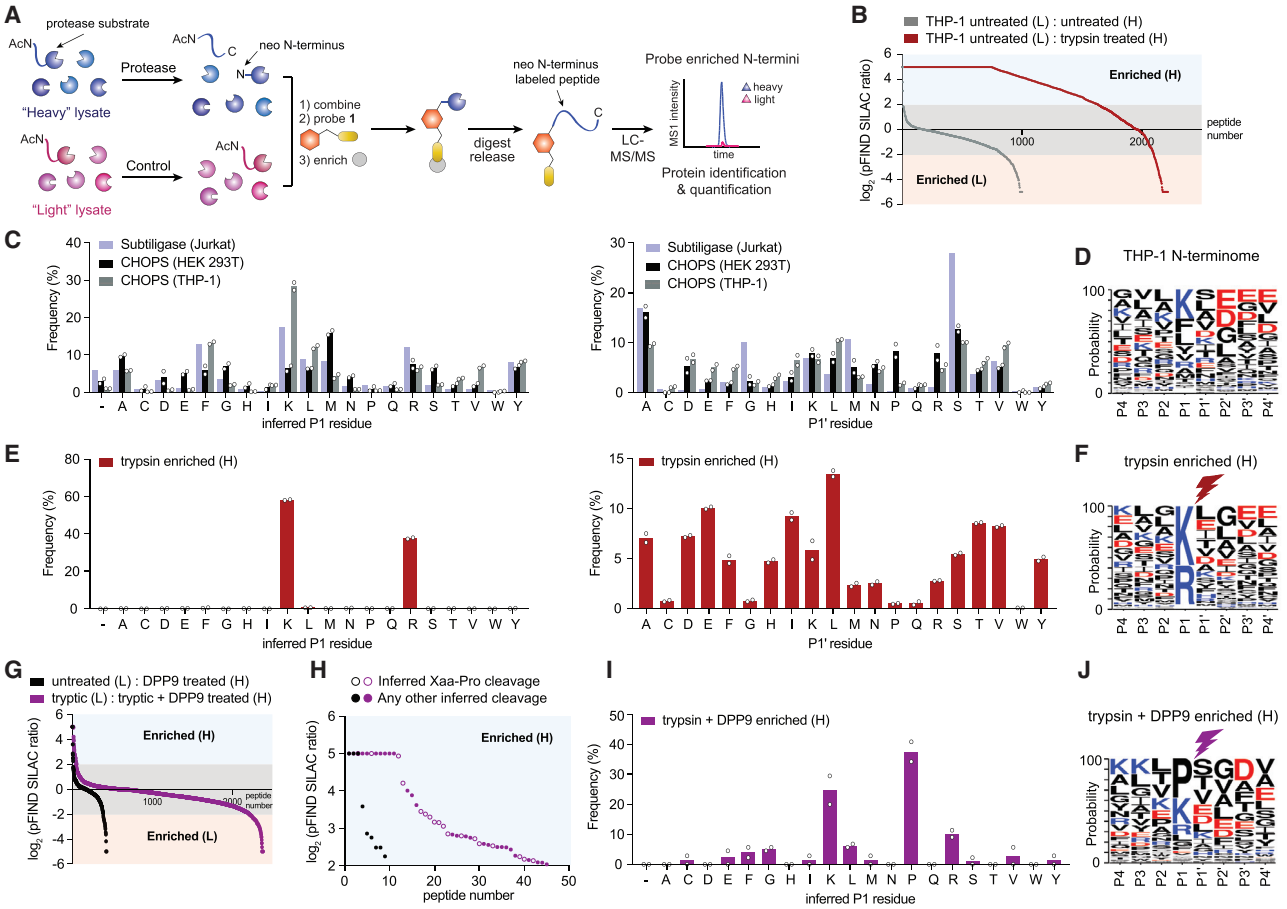


Figure 4. An MS-Based Platform for Proteome-wide Evaluation of Protease Substrates

(A) Graphical depiction of MS-CHOPS. "AcN" indicates an acetylated N-terminus, which will not be labeled by **1**.

(B–F) CHOPS identified **1**-labeled peptides from untreated heavy- and light-labeled THP-1 lysates (n = 2 biological replicates; replicate 1 = 863 peptides, replicate 2 = 1,003 peptides), trypsin-treated heavy- and light-labeled THP-1 lysates (n = 2 biological replicates; replicate 1 = 2,219 peptides, replicate 2 = 2,048 peptides), and untreated heavy- and light-labeled lysates HEK293T lysates (n = 2 biological replicates; replicate 1 = 194 peptides, replicate 2 = 353 peptides). Isotopic enrichments from one replicate of the indicated experiments are shown (B). The mean frequency of inferred P1 residues and P1′ residues of probe **1**-labeled peptides from the untreated experiments (n = 2) at similar heavy:light ratios (gray box in B) were compared with a previous analysis of Jurkat cells using subtiligase (Mahrus et al., 2008) (C). Sequence logo representation of residues P4–P4′ for the THP-1 "N-terminome" determined by CHOPS is shown (D). Frequency of inferred P1 residues and P1′ residues (E) and sequence logo representation (F) for heavy-enriched probe-**1**-labeled peptides (blue box in B) from the trypsin-treated experiment are shown.

(G–J) CHOPS identified **1**-labeled peptides from heavy DPP9-treated and light untreated *DPP8/9* KO THP-1 lysates (n = 1 biological replicate; 429 peptides) and heavy DPP9-treated and light untreated tryptic *DPP8/9* KO THP-1 proteomes (n = 2 biological replicates; replicate 1 = 2,390 peptides, replicate 2 = 2,198 peptides). All peptides (G) and heavy-enriched peptides (H, blue box in G) from a single replicate of the indicated experiments are shown. Open circles indicate inferred Xaa-Pro cleavage, filled circles indicate non-P1 Pro sequences (H). Frequency of inferred P1 residues (I) and sequence logo representation (J) or heavy-enriched probe-**1**-labeled peptides from the tryptic DPP9-treated experiment.

full-length protein. Although we cannot rule out that CHOPS may have failed to identify a key full-length protein substrate of DPP8/9, we speculate that DPP8/9 are enzymes primarily involved in the cleavage of cellular peptides.

More generally, we have shown that CHOPS can be coupled with quantitative MS to identify substrates for proteases outside the DPP family. In this regard, we predict that CHOPS will form the basis of a complementary N-terminomic strategy to subtiligase and TAILS. As CHOPS has more general reactivity with N-termini than subtiligase (Figure 4C), it may offer compatibility with a broader substrate repertoire; as CHOPS involves positive enrichment, unlike TAILS, it may detect lower abundance substrates. However, CHOPS does have a clear limitation in that cannot detect substrates containing a P2′ Pro. Future work will be needed to fully characterize the advantages and disadvantages of CHOPS as a general N-terminomic strategy. Regardless, the work here delineates a simple strategy for identifying protease substrates in complex lysates and encourages its further development and application.

SIGNIFICANCE

The dipeptidyl peptidases (DPPs) are a family of serine proteases that cleave substrates after penultimate

amino-terminal proline residues. The DPPs cleave and regulate a wide range of peptide hormones, cytokines, and neuropeptides, and several have attracted significant interest as therapeutic targets. However, the DPPs remain poorly characterized due to technical challenges in discovering their physiologically relevant substrates, which are often in low abundance and indistinguishable in size by gel electrophoresis after cleavage. Here, we show that 2-pyridinecarboxaldehyde (2PCA)-based chemical probes can be used to selectively label and enrich canonical DPP substrates. Importantly, this chemical-based strategy, called CHOPS, not only readily identifies DPP substrates *in vitro* but also enables the enrichment of cleaved DPP substrates from cell lysates. Using this approach, we show that DPP8 and DPP9, enzymes that control the Nlrp1b inflammasome through an unknown mechanism, do not cleave Nlrp1b or other full-length proteins. More generally, these findings define a useful strategy for the rapid evaluation of DPP substrates and provide the bases for future development of a chemical reactivity-based protease profiling platform.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENTS AND RESOURCES SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Biological Methods
  - Mass Spectrometry-based CHOPS
  - LC-MS/MS Method
  - Proteomic Analysis
  - FP-Biotin Enrichment of Serine Proteases
  - Chemical Synthesis
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.chembiol.2019.03.007.

ACKNOWLEDGMENTS

This work was supported by the Josie Robertson Foundation (D.A.B. and A.K.), a Stand Up to Cancer-Innovative Research grant (grant number SU2C-AACR-IRG11-17 to D.A.B.; Stand Up to Cancer is a program of the Entertainment Industry Foundation. Research grants are administered by the American Association for Cancer Research, the scientific partner of SU2C), the Pew Charitable Trusts (D.A.B. is a Pew-Stewart Scholar in Cancer Research), the Pershing Square Sohn Cancer Research Alliance (D.A.B.), the NIH (R01 AI137168 to D.A.B.; T32 GM007739-Andersen to A.R.G; R21 CA188881 and R01 CA204396 to A.K; and the MSKCC Core Grant P30 CA008748), an Alfred P. Sloan Foundation Research Fellowship (D.A.B.), Burroughs Wellcome Fund (A.K), Alex's Lemonade Stand Foundation (A.K.), Gabrielle's Angel Foundation (D.A.B. and A.K.), and the Damon Runyon Cancer Research Foundation (A.K. is the Damon Runyon-Richard Lumsden Foundation Clinical Investigator). We thank E. Weerapana, J. Peeler, C. Wang, Y. Liu, and K. Boehm for helpful comments and discussion.

AUTHOR CONTRIBUTIONS

A.R.G. and D.A.B. designed and performed experiments, analyzed data, and wrote the paper. A.R.G. with assistance from A.J.A. synthesized compounds. P.C., M.M.M., A.K., and R.C.H. assisted with LC-MS/MS experiments and MS data analysis. S.D.R. assisted with cloning and transfections.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: June 21, 2018
Revised: February 14, 2019
Accepted: March 12, 2019
Published: April 18, 2019

REFERENCES

Adams, S., Miller, G.T., Jesson, M.I., Watanabe, T., Jones, B., and Wallner, B.P. (2004). PT-100, a small molecule dipeptidyl peptidase inhibitor, has potent antitumor effects and augments antibody-mediated cytotoxicity via a novel immune mechanism. Cancer Res. *64*, 5471–5480.

Ajami, K., Pitman, M.R., Wilson, C.H., Park, J., Menz, R.I., Starr, A.E., Cox, J.H., Abbott, C.A., Overall, C.M., and Gorrell, M.D. (2008). Stromal cell-derived factors 1alpha and 1beta, inflammatory protein-10 and interferon-inducible T cell chemo-attractant are novel substrates of dipeptidyl peptidase 8. FEBS Lett. *582*, 819–825.

Bachmair, A., Finley, D., and Varshavsky, A. (1986). In vivo half-life of a protein is a function of its amino-terminal residue. Science *234*, 179–186.

Batterham, R.L., Cohen, M.A., Ellis, S.M., Le Roux, C.W., Withers, D.J., Frost, G.S., Ghatei, M.A., and Bloom, S.R. (2003). Inhibition of food intake in obese subjects by peptide YY3-36. N. Engl. J. Med. *349*, 941–948.

Bjelke, J.R., Christensen, J., Nielsen, P.F., Branner, S., Kanstrup, A.B., Wagtmann, N., and Rasmussen, H.B. (2006). Dipeptidyl peptidases 8 and 9: specificity and molecular characterization compared with dipeptidyl peptidase IV. Biochem. J. *396*, 391–399.

Busek, P., Malik, R., and Sedo, A. (2004). Dipeptidyl peptidase IV activity and/or structure homologues (DASH) and their substrates in cancer. Int. J. Biochem. Cell Biol. *36*, 408–421.

Chi, H., Liu, C., Yang, H., Zeng, W.F., Wu, L., Zhou, W.J., Wang, R.M., Niu, X.N., Ding, Y.H., Zhang, Y., et al. (2018). Comprehensive identification of peptides in tandem mass spectra using an efficient open search engine. Nat. Biotechnol. https://doi.org/10.1038/nbt.4236.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. Genome Res. *14*, 1188–1190.

D'Osualdo, A., Weichenberger, C.X., Wagner, R.N., Godzik, A., Wooley, J., and Reed, J.C. (2011). CARD8 and NLRP1 undergo autoproteolytic processing through a ZU5-like domain. PLoS One *6*, e27396.

Deacon, C.F., and Lebovitz, H.E. (2016). Comparative review of dipeptidyl peptidase-4 inhibitors and sulphonylureas. Diabetes Obes. Metab. *18*, 333–347.

Dix, M.M., Simon, G.M., and Cravatt, B.F. (2008). Global mapping of the topography and magnitude of proteolytic events in apoptosis. Cell *134*, 679–691.

Du, C., Fang, M., Li, Y., Li, L., and Wang, X. (2000). Smac, a mitochondrial protein that promotes cytochrome c-dependent caspase activation by eliminating IAP inhibition. Cell *102*, 33–42.

Finger, J.N., Lich, J.D., Dare, L.C., Cook, M.N., Brown, K.K., Duraiswami, C., Bertin, J., and Gough, P.J. (2012). Autolytic proteolysis within the function to find domain (FIIND) is required for NLRP1 inflammasome activity. J. Biol. Chem. *287*, 25030–25037.

Frew, B.C., Joag, V.R., and Mogridge, J. (2012). Proteolytic processing of Nlrp1b is required for inflammasome activity. PLoS Pathog. *8*, e1002659.

Geiss-Friedlander, R., Parmentier, N., Moller, U., Urlaub, H., Van den Eynde, B.J., and Melchior, F. (2009). The cytoplasmic peptidase DPP9 is rate-limiting

for degradation of proline-containing peptides. J. Biol. Chem. *284*, 27211–27219.

Green, B.D., Mooney, M.H., Gault, V.A., Irwin, N., Bailey, C.J., Harriott, P., Greer, B., O'Harte, F.P., and Flatt, P.R. (2004). N-terminal His(7)-modification of glucagon-like peptide-1(7-36) amide generates dipeptidyl peptidase IV-stable analogues with potent antihyperglycaemic activity. J. Endocrinol. *180*, 379–388.

Johnson, D.C., Taabazuing, C.Y., Okondo, M.C., Chui, A.J., Rao, S.D., Brown, F.C., Reed, C., Peguero, E., de Stanchina, E., Kentsis, A., et al. (2018). DPP8/DPP9 inhibitor-induced pyroptosis for treatment of acute myeloid leukemia. Nat. Med. *24*, 1151–1156.

Jost, M.M., Lamerz, J., Tammen, H., Menzel, C., De Meester, I., Lambeir, A.M., Augustyns, K., Scharpe, S., Zucht, H.D., Rose, H., et al. (2009). In vivo profiling of DPP4 inhibitors reveals alterations in collagen metabolism and accumulation of an amyloid peptide in rat plasma. Biochem. Pharmacol. *77*, 228–237.

Justa-Schuch, D., Silva-Garcia, M., Pilla, E., Engelke, M., Kilisch, M., Lenz, C., Moller, U., Nakamura, F., Urlaub, H., and Geiss-Friedlander, R. (2016). DPP9 is a novel component of the N-end rule pathway targeting the tyrosine kinase Syk. Elife *5*, https://doi.org/10.7554/eLife.16370.

Kleifeld, O., Doucet, A., auf dem Keller, U., Prudova, A., Schilling, O., Kainthan, R.K., Starr, A.E., Foster, L.J., Kizhakkedathu, J.N., and Overall, C.M. (2010). Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products. Nat. Biotechnol. *28*, 281–288.

Lankas, G.R., Leiting, B., Roy, R.S., Eiermann, G.J., Beconi, M.G., Biftu, T., Chan, C.C., Edmondson, S., Feeney, W.P., He, H., et al. (2005). Dipeptidyl peptidase IV inhibition for the treatment of type 2 diabetes: potential importance of selectivity over dipeptidyl peptidases 8 and 9. Diabetes *54*, 2988–2994.

MacDonald, J.I., Munch, H.K., Moore, T., and Francis, M.B. (2015). One-step site-specific modification of native proteins with 2-pyridinecarboxyaldehydes. Nat. Chem. Biol. *11*, 326–331.

Mahrus, S., Trinidad, J.C., Barkan, D.T., Sali, A., Burlingame, A.L., and Wells, J.A. (2008). Global sequencing of proteolytic cleavage sites in apoptosis by specific labeling of protein N termini. Cell *134*, 866–876.

Mulvihill, E.E., and Drucker, D.J. (2014). Pharmacology, physiology, and mechanisms of action of dipeptidyl peptidase-4 inhibitors. Endocr. Rev. *35*, 992–1019.

Okondo, M.C., Johnson, D.C., Sridharan, R., Go, E.B., Chui, A.J., Wang, M.S., Poplawski, S.E., Wu, W., Liu, Y., Lai, J.H., et al. (2017). DPP8 and DPP9 inhibition induces pro-caspase-1-dependent monocyte and macrophage pyroptosis. Nat. Chem. Biol. *13*, 46–53.

Okondo, M.C., Rao, S.D., Taabazuing, C.Y., Chui, A.J., Poplawski, S.E., Johnson, D.C., and Bachovchin, D.A. (2018). Inhibition of Dpp8/9 activates the Nlrp1b inflammasome. Cell Chem Biol *25*, 262–267.e5.

Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D.J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., et al. (2019). The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. *47*, D442–D450.

Rasmussen, H.B., Branner, S., Wiberg, F.C., and Wagtmann, N. (2003). Crystal structure of human dipeptidyl peptidase IV/CD26 in complex with a substrate analog. Nat. Struct. Biol. *10*, 19–25.

Reichmann, F., and Holzer, P. (2016). Neuropeptide Y: a stressful review. Neuropeptides *55*, 99–109.

Rosenblum, J.S., and Kozarich, J.W. (2003). Prolyl peptidases: a serine protease subfamily with high potential for drug discovery. Curr. Opin. Chem. Biol. *7*, 496–504.

Ross, B., Krapp, S., Augustin, M., Kierfersauer, R., Arciniega, M., Geiss-Friedlander, R., and Huber, R. (2018). Structures and mechanism of dipeptidyl peptidases 8 and 9, important players in cellular homeostasis and cancer. Proc. Natl. Acad. Sci. U S A *115*, E1437–E1445.

Shao, W., Yeretssian, G., Doiron, K., Hussain, S.N., and Saleh, M. (2007). The caspase-1 digestome identifies the glycolysis pathway as a target during infection and septic shock. J. Biol. Chem. *282*, 36321–36329.

Tagore, D.M., Nolte, W.M., Neveu, J.M., Rangel, R., Guzman-Rojas, L., Pasqualini, R., Arap, W., Lane, W.S., and Saghatelian, A. (2009). Peptidase substrates via global peptide profiling. Nat. Chem. Biol. *5*, 23–25.

Tammen, H., Hess, R., Rose, H., Wienen, W., and Jost, M. (2008). Peptidomic analysis of blood plasma after in vivo treatment with protease inhibitors–a proof of concept study. Peptides *29*, 2188–2195.

Tinoco, A.D., Kim, Y.G., Tagore, D.M., Wiwczar, J., Lane, W.S., Danial, N.N., and Saghatelian, A. (2011). A peptidomics strategy to elucidate the proteolytic pathways that inactivate peptide hormones. Biochemistry *50*, 2213–2222.

Tinoco, A.D., Tagore, D.M., and Saghatelian, A. (2010). Expanding the dipeptidyl peptidase 4-regulated peptidome via an optimized peptidomics platform. J. Am. Chem. Soc. *132*, 3819–3830.

Tonge, R., Shaw, J., Middleton, B., Rowlinson, R., Rayner, S., Young, J., Pognan, F., Hawkins, E., Currie, I., and Davison, M. (2001). Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. Proteomics *1*, 377–396.

Waumans, Y., Baerts, L., Kehoe, K., Lambeir, A.M., and De Meester, I. (2015). The dipeptidyl peptidase family, prolyl oligopeptidase, and prolyl carboxypeptidase in the immune system and inflammatory disease, including atherosclerosis. Front. Immunol. *6*, 387.

Weeks, A.M., and Wells, J.A. (2018). Engineering peptide ligase specificity by proteomic identification of ligation sites. Nat. Chem. Biol. *14*, 50–57.

Wilson, C.H., Indarto, D., Doucet, A., Pogson, L.D., Pitman, M.R., McNicholas, K., Menz, R.I., Overall, C.M., and Abbott, C.A. (2013). Identifying natural substrates for dipeptidyl peptidases 8 and 9 using terminal amine isotopic labeling of substrates (TAILS) reveals in vivo roles in cellular homeostasis and energy metabolism. J. Biol. Chem. *288*, 13936–13949.

Wilson, C.H., Zhang, H.E., Gorrell, M.D., and Abbott, C.A. (2016). Dipeptidyl peptidase 9 substrates and their discovery: current progress and the application of mass spectrometry-based approaches. Biol. Chem. *397*, 837–856.

Wu, G., Chai, J., Suber, T.L., Wu, J.W., Du, C., Wang, X., and Shi, Y. (2000). Structural basis of IAP recognition by Smac/DIABLO. Nature *408*, 1008–1012.

Yang, X., Boehm, J.S., Yang, X., Salehi-Ashtiani, K., Hao, T., Shen, Y., Lubonja, R., Thomas, S.R., Alkan, O., Bhimdi, T., et al. (2011). A public genome-scale lentiviral expression library of human ORFs. Nat. Methods *8*, 659–661.

Yates, N.A., Deyanova, E.G., Geissler, W., Wiener, M.C., Sachs, J.R., Wong, K.K., Thornberry, N.A., Roy, R.S., Settlage, R.E., and Hendrickson, R.C. (2007). Identification of peptidase substrates in human plasma by FTMS based differential mass spectrometry. Int. J. Mass Spectrom. *259*, 174–183.

Zhang, H., Maqsudi, S., Rainczuk, A., Duffield, N., Lawrence, J., Keane, F.M., Justa-Schuch, D., Geiss-Friedlander, R., Gorrell, M.D., and Stephens, A.N. (2015). Identification of novel dipeptidyl peptidase 9 substrates by two-dimensional differential in-gel electrophoresis. FEBS J. *282*, 3737–3757.

Zhong, F.L., Robinson, K., Teo, D.E.T., Tan, K.Y., Lim, C., Harapas, C.R., Yu, C.H., Xie, W.H., Sobota, R.M., Au, V.B., et al. (2018). Human DPP9 represses NLRP1 inflammasome and protects against autoinflammatory diseases via both peptidase activity and FIIND domain binding. J. Biol. Chem. *293*, 18864–18878.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Smac/Diablo Mouse monoclonal Ab	Cell Signaling Tech	2954S; RRID: AB_2131196
V5 Rabbit polyclonal Ab	Abcam	Ab9116; RRID: AB_307024
DPP9 Rabbit polyclonal Ab	Abcam	Ab42080; RRID: AB_731947
DPP8 Rabbit polyclonal Ab	Abcam	Ab42076; RRID: AB_731944
GAPDH Rabbit monoclonal Ab	Cell Signaling Tech	14C10; RRID: AB_10693448
IRDye 800CW anti-rabbit	LICOR	925-32211; RRID: AB_2651127
IRDye 800CW anti-Mouse	LICOR	925-32210; RRID: AB_2687825
IRDye 680 RD Streptavidin	LICOR	926-68079
Chemicals, Peptides, and Recombinant Proteins		
Biotin-coupled 2PCA probe, **1**	This paper	N/A
Gly-Pro-AMC	VWR	100042-646
Val-boroPro	Okondo et al., 2017	N/A
L-lysine:2HCl	Cambridge Isotope Laboratories	ULM-8766-0.05, lot: I-17807
L-lysine:2HCl 13C6, 99%; 15N2, 99%	Cambridge Isotope Laboratories	CNLM-291-H-0.05, lot: PR-26193
L-arginine:HCl	Cambridge Isotope Laboratories	ULM-8347-0.1, lot: I1-12740A
L-arginine:HCl 13C6, 99%; 15N4, 99%	Cambridge Isotope Laboratories	CNLM-539-0.1, lot: PR-27681
human peptide YY(1-36)	Anaspec	AS-24401
human peptide YY(3-36)	Anaspec	AS-24405
human neuropeptide Y	Anaspec	AS-22464
human chemokine CXCL12/SDF-1 alpha	Kerafast	EMW001
DPP4	R&D	9168-SE-010, Lot: DFQM0417091
DPP8	Enzo Life	BML-SE527-0010, Lot: 05311606
DPP9	Enzo Life	BML-SE528-0010, Lot: 07221432
sequencing grade trypsin	Promega	V5113
Deposited Data		
MS-CHOPS proteomics data	This paper	ProteomeXchange PXD013019
Experimental Models: Cell Lines		
Human HEK 293T	ATCC	CRL-3216
Human THP-1	ATCC	TIB-202
Human *DPP8/9* KO THP-1	Okondo et al., 2017	N/A
Oligonucleotides		
Primers for cloning, please see Table S3.	This paper	N/A
Recombinant DNA		
pLEX_307	Addgene	Addgene Plasmid Cat #41392
Ubiquitin WT	Addgene	Addgene Plasmid Cat#12647
pET DEST42	ThermoFisher	Cat#12276010
pLEX_307_Ubiquitin_V57-SMAC	This paper	N/A
pLEX_307_Ubiquitin_I59-SMAC	This paper	N/A
pET-DEST42_V57-SMAC	This paper	N/A
pET-DEST42_I59-SMAC	This paper	N/A
pLEX_307_Ubiquitin_S984-Nlrp1b	This paper	N/A
pLEX_307_Ubiquitin_M986-Nlrp1b	This paper	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
GraphPad Prism version 7	GraphPad Software	www.graphpad.com
pFIND version 3.1.5	Chi et al., 2018	www.pfind.ict.ac.cn
WebLogo 3.0	Crooks et al., 2004	www.weblogo.threeplusone.com
Mnova version 11.0.4	Mestrelab Research	www.mestrelab.com
FlexControl v3.4	Bruker	www.bruker.com
FlexAnalysis v3.4	Bruker	www.bruker.com
ChemDraw Professional 16.0	Perkin Elmer	www.perkinelmer.com

CONTACT FOR REAGENTS AND RESOURCES SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Daniel A. Bachovchin (bachovcd@mskcc.org).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines and cell culture information is listed below. HEK 293T (ATCC, fetal) cells were grown in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum (FBS). THP-1 (ATCC, male) and *DPP8/9* KO THP-1 ((Okondo et al., 2017), male) cells were grown in Roswell Park Memorial Institute medium (RPMI) with 10% FBS. All cells were grown at 37°C in a 5% CO_2 incubator. Cell lines were regularly tested for mycoplasma using the MycoAlert™ Mycoplasma Detection Kit (Lonza). For Stable Isotope Labeling with Amino acids in Cell culture (SILAC), 50 mL of dialyzed FBS (Sigma-Aldrich) was added to 500 mL RPMI 1640 Media for SILAC (Thermo) followed by either 50 mg unlabeled L-Lysine:2HCl, 100 mg unlabeled L-Arginine:HCl or 50 mg labeled L-Lysine:2HCl (13C6, 99%; 15N2, 99%), 100 mg labeled L-Arginine:HCl (13C6, 99%; 15N4, 99%). Media was sterile-filtered with a 0.22 μm filter, aliquoted into sterile 50 mL bottles, labeled light/heavy and stored at 4°C protected from light. All SILAC cells underwent at least 20 cell doublings in respective SILAC media (light/heavy) prior to MS-based experiments.

METHOD DETAILS

Biological Methods

Cloning

cDNA encoding the full-length mouse Nlrp1b gene was cloned from RAW 264.7 macrophages. The full-length gene was shuttled into a pLEX_307 vector (Addgene) with a C-terminal V5 tag using Gateway technology (Thermo Fisher Scientific). Two C-terminal constructs of Nlrp1b starting at S984 and M986 were generated by PCR amplification using specific primers, which contained a 5' sequence overlap with ubiquitin. The ubiquitin (Ub) sequence was PCR amplified (Addgene 12647) with primers containing a 3' overlap with corresponding Nlrp1b C-termini. The partially overlapping Ub and Nlrp1b C-termini products were mixed and assembled by PCR to yield the ubiquitin fused constructs. These products were then shuttled into a pLEX_307 vector (Addgene) with a C-terminal V5 tag using Gateway technology (Thermo Fisher Scientific). cDNA encoding SMAC was cloned from The Broad Institute's ORFeome (Yang et al., 2011), and similarly amplified to generate the ubiquitin fused constructs. These products were then shuttled into a modified pLEX_307 vector (Addgene) with a C-terminal FLAG tag and a pET-DEST42 vector (Addgene) using Gateway technology (Thermo Fisher Scientific).

GP-AMC Assay

DPP4, DPP8, and DPP9 were diluted to 100 nM in PBS with 10% glycerol. A solution of substrate (1 mM Gly-Pro-AMC) was prepared in DMSO. 16 μL of PBS or the indicated cell lysate was added to a 394-well black clear-bottom plate (Corning), followed by 2 μL of 100 nM enzyme and 2 μL of 1 mM substrate. A blank well with 2 μL of 1 mM substrate and 18 μL of PBS was used as a negative control. Fluorescence was recorded at ambient temperature every minute for 30 min at 380-nm excitation and 460-nm emission wavelengths.

Gel-Based In Vitro *DPP Substrate Assays*

Lyophilized peptide YY(1-36), peptide YY(3-36), neuropeptide Y, and Human Chemokine CXCL12 were resuspended in water to 100 μM and stored at −20°C. Enzymatic activity was confirmed by GP-AMC assay. Reactions were prepared with 8 μL of 100-μM peptide stock, 8 μL of 100 nM enzyme (DPP4, DPP8 and DPP9) or PBS (for control), and 64 μL of PBS (Corning). Reactions incubated at 37°C for 24 h. 9 μL aliquots were taken at 0, 1, 4, 8 and 24 h and immediately boiled to deactivate the enzyme. 1 μL of 100 mM **1** was added to the aliquots. After 16 h of incubating at 37°C, an equal volume of 2× loading dye was added, samples were boiled and run on 16% tricine gels (Invitrogen) at 125 V for 90 min. A semi-dry transfer at 25 V, 2.5 A for 15 min was used to transfer

proteins to the membrane, which was then blocked and blotted for biotin with streptavidin IR dye. Additional 9 μL aliquots were also taken at 0, 24 h for Coomassie staining (Invitrogen) according to the manufacturer's specifications.

Detection of DPP Substrates in Lysates

HEK 293T cells were seeded at 5 × 10^5 cells/well in 6-well plates. The next day cells were transfected by manufacturer's instructions (Fugene, Promega) with the ubiquitin-fused construct (2 μg/well) or red fluorescent protein (mock). Cells were harvested 48 h later, pelleted by centrifugation (400 × *g*, 4°C), washed twice with cold PBS (Corning), resuspended in 250 μL PBS, sonicated (3 × 10 s), and clarified by centrifuging 15 min at 15,000 × *g*. The resulting soluble proteins were transferred to a new microcentrifuge tube; the concentration was measured by DC protein assay kit (Bio-Rad), and diluted to 1.5 mg/mL.

To test probe labeling of these constructs, 5 μL of 100 mM **1** was added to a 50 μL aliquot of 1.5 mg/mL lysate. After incubating for 16 h at 37°C unbound probe was removed by buffer exchanging into PBS using a 3-kDa cut-off ultra-centrifugal filter unit (Amicon) according to the manufacturer's instructions. The resulting concentrated solution was then boiled in 100 μL of 1% SDS for 10 min and diluted to 500 μL with the addition of PBS. 100 μL of high capacity neutroavidin agarose resin slurry (Pierce) was added and the mixture was rotated end-over-end for 2 h at ambient temperature. Agarose resin was pelleted by centrifugation (500 × *g* for 1 min), washed 3 times with 0.2% SDS in PBS (1 mL), and 3 times with PBS (1 mL). Boiling the resin with an equivalent volume of 2× SDS loading dye for 10 min eluted the immunoprecipitated peptides, which were separated by SDS-PAGE gel, immunoblotted, and visualized using the Odyssey Imaging System (LI-COR).

For the DPP8/9 cleavage assay of ubiquitin-fusion generated fragments, 15 μL of 100 nM DPP8, DPP9, or PBS (control) were added to a 135 μL aliquot of 1.5 mg/mL lysate. The mixture incubated 24 h at 37°C. The samples were boiled to deactivate enzymes, 15 μL of 100 mM **1** was added and the reaction incubated at 37°C for 16 h. Unbound probe was removed by buffer exchanging into PBS using a 3-kDa cut-off ultra-centrifugal filter unit (Amicon) according to the manufacturer's instructions. The samples then processed as described above.

For evaluating DPP8/9 cleavage of full-length Nlrp1b, HEK 293T cells were seeded at 5 × 10^5 cells/well in 6-well plates. The next day, cells were transiently transfected with the indicated constructs encoding full-length Nlrp1b protein. The cells were then cultured in the presence of Val-boroPro (10 μM, with fresh drug added each day) for 3 days before the soluble proteome was harvested as described above. These cells were cultured with Val-boroPro to ensure that any potential DPP cleavage activity on the neo N-terminus generated by Nlrp1b autoproteolysis was blocked (i.e., it would not be possible to see an increase in probe **1** labeling after recombinant DPP9 treatment if endogenous DPP9 had already cleaved this substrate completely). Treatment with Val-boroPro was unnecessary for the ubiquitin-fusion generated C-terminal fragment constructs because a control fragment lacking the N-terminal NH_2-Ser-Pro dipeptide could be generated for comparison. Inhibition of endogenous DPP activity was confirmed by GP-AMC assay (Figure S2A). 20 μL of 100 nM DPP8, DPP9, or PBS (control) were then added to a 200 μL aliquot of 1.5 mg/mL lysate, and activity of recombinant DPP9 in the lysate was confirmed by GP-AMC assay (Figure S2B). The mixtures were incubated 24 h at 37°C before being boiled for 10 min to separate the N- and C- terminal fragments. 20 μL of 100 mM **1** was added and the reaction incubated at 37°C for 16 h. Unbound probe was removed by buffer exchanging into PBS using a 3-kDa cut-off ultra-centrifugal filter unit (Amicon) according to the manufacturer's instructions. The samples were then processed as described above.

Mass Spectrometry-based CHOPS

SILAC labeled (light/heavy) cell pellets were suspended in 25 mM ammonium bicarbonate (ABC) and sonicated (3 × 10 s). The resulting proteomes were transferred to a new microcentrifuge tubes; the concentration was measured by DC protein assay kit (Bio-Rad).

To evaluate the native N-terminome 100 μL of heavy and light lysates (HEK 293T 1 mg/mL, THP-1 2 mg/mL) were mixed together and labeled with **1** (5 mM) for 16 h at 37°C. Proteins were precipitated by adding seven volumes (1.4 ml) of acetone, vortexing briefly, incubating for 1 h at −20°C, centrifuging 10 min at 15,000 × *g* (4°C) and decanting the supernatant. The residual protein pellet was suspended with sonication in 1 mL of 50 mM Tris HCl, 150 mM NaCl with 1% NP-40. 200 μl of high capacity neutroavidin agarose resin slurry (Pierce) was added and the mixture was rotated end-over-end for 2 h at ambient temperature. Agarose resin was pelleted by centrifugation (500 × *g* for 1 min), washed 2 times with 50 mM Tris HCl, 150 mM NaCl with 1% NP-40 (1 ml), 2 times with PBS (1 mL), 2 times with water (1 mL). Proteins were reduced on-bead with 10 mM DTT (42°C for 30 min), and alkylated with 20 mM iodoacetamide (37°C for 30 min). Agarose resin was pelleted, resuspended in 2 M urea, 1 mM $CaCl_2$ in 25 mM ABC and trypsinized overnight (2 ug, Pierce) while shaking at 37°C. The resulting tryptic peptides were washed once with 50 mM Tris HCl, 150 mM NaCl with 1% NP-40 (1 mL), 2 times with PBS (1 mL), and 2 times with water (1 mL). Bound peptides were eluted twice with 200 μL of 80% acetonitrile in water containing 0.1% formic acid (first at RT for 10 min, second at 70°C for 10 min). The combined elution fractions were dried using a Genevac EZ-2 evaporator, resuspended in 100 μL water containing 0.1% formic acid, desalted with a C18 Silica MicroSpin Column (The Nest Group) according to the manufacturer's instructions and re-dried using a Genevac EZ-2 evaporator to give the final peptide pellet for LC-MS/MS analysis.

To evaluate tryptic protease substrates, 200 μL of 3 mg/mL SILAC labeled THP-1 proteomes were boiled for 15 min and sonicated to resuspend precipitated proteins. Sequencing grade trypsin (2 μg) was added along with 1 mM $CaCl_2$ to the heavy proteome, an equal volume of PBS was added to the light proteome. The samples shook overnight at 37°C, at which time the samples were boiled to inactivate trypsin, sonicated and combined pairwise. Samples were labeled with **1** and processed as above.

To evaluate DPP9 whole proteome substrates, 200 μL of 3 mg/mL SILAC labeled *DPP8/9 KO* THP-1 proteome was incubated with either PBS (light) or DPP9 (200 U). The samples shook overnight at 37°C, at which time the samples were boiled to inactivate DPP9,

sonicated and combined pairwise. Samples were labeled with **1** and processed as above. Rather than trypsinizing on bead, labeled peptides/proteins were eluted with acetonitrile, dried, trypsinized in solution and then desalted for LC-MS/MS analysis.

To evaluate DPP9 tryptic proteome substrates, 200 μL of 3 mg/mL SILAC labeled *DPP8/9 KO* THP-1 proteome were trypsinized for 8 h at 37°C. Samples were boiled to inactivate trypsin, sonicated and incubated with either PBS (light) or DPP9 (200 U). The samples shook overnight at 37°C, at which time the samples were boiled to inactivate DPP9, sonicated and combined pairwise. Samples were labeled with **1** and processed as above.

LC-MS/MS Method

The desalted peptides were resuspended in 50 ul of 0.1% formic acid in water. 1 ul sample was injected for microcapillary liquid chromatography with tandem mass spectrometry using the NanoAcquity (Waters) with a 100- μm-inner-diameter × 10-cm-length C18 column [1.7 (μm) BEH130, Waters] configured with a 180-μm × 2-cm trap column coupled to a Q Exactive Plus mass spectrometer (Thermo Fisher Scientific). Trapping was performed at 15 uL/min (0.1% formic acid) for 1 min. Peptides were eluted with a linear gradient of 0-50% acetonitrile (0.1% formic acid) in water (0.1% formic acid) over 90 min with a flow rate of 300 (nL/min). Full scan MS1 spectra were acquired over 400-1600 (m/z) at 70,000 resolution with max IT of 50 ms and automatic gain control (AGC) at 1×10^6 ions. MS data was collected in data dependent acquisition (DDA) mode scanning the top 10 most intense precursor ions for HCD fragmentation performed at normalized collision energy (NCE) 27% with AGC at 5×10^4 ions, isolation window 1.5 m/z, and dynamic exclusion of 15s. MS/MS spectra were collected with resolution of 17,500.

Proteomic Analysis

MS/MS data (.raw files) were analyzed with pFIND version 3.1.5 (Chi et al., 2018). Probe **1** was added as a variable peptide N-terminal modification ($C_{21}N_5O_2H_{27}S_1$, +413.1885) and searches were conducted with the following parameters: open search against semi-tryptic peptides, allowing precursor and fragment tolerances of ± 20 ppm, using a human SwissProt FASTA database with reverse sequence decoys and common contaminants, quantified with SILAC labeling (Light label = none, Heavy label = SILAC-Arg10Lys8), and allowing an FDR of less than 1% at the peptides level. The median SILAC ratio and spectra score was calculated for unique peptide sequences identified with the correct N-terminal modification and at least 2 spectral counts in a single run. Log_2 (median SILAC ratio) were calculated setting the upper limit to 5 and lower limit to -5. The P1-P6 residues were determined for the resulting peptide sequences by searching against the human SwissProt FASTA database, given the identified protein in pFIND. Peptides were classified as 'heavy enriched' given a log_2 (median SILAC ratio) of greater than 2 and 'light enriched' given a log_2 (median SILAC ratio) of less than -2. Peptides were considered 'at similar level' given a log_2 (median SILAC ratio) between -2 and 2. Sequence logo were generated using WebLogo (Crooks et al., 2004).

FP-Biotin Enrichment of Serine Proteases

THP-1 cells were pelleted at 400 g, resuspended in 1 mL of PBS and lysed by sonication. Protein concentrations were determined using the DC Protein Assay kit (Bio-Rad) and adjusted to 2 mg/mL. One milliliter of 1 mg/mL lysate was reacted with the FP-biotin probe (5 μM) for 1 h at 25°C. Triton-X100 was then added to a final concentration of 1%, and samples were incubated for 1 h while rotating at 4°C to solubilize membranes. Lysates were then separated from free probe by passage over a Sephadex G-25M column (GE Healthcare). SDS was added to a final concentration of 0.5% (in 3.5 mL total volume), and samples were boiled for 10 min at 95°C to denature proteins. Samples were then diluted in PBS to a final volume of 11 mL before the addition of 200 μL of NeutrAvidin-Agarose resin (Pierce) and incubation for 1 h while rotating at 25°C. Samples were then washed with 4 × 10 mL PBS and eluted by resuspension in 150 mL 2× sample loading dye and boiling at 95°C for 10 min.

Purification of SMAC Proteins

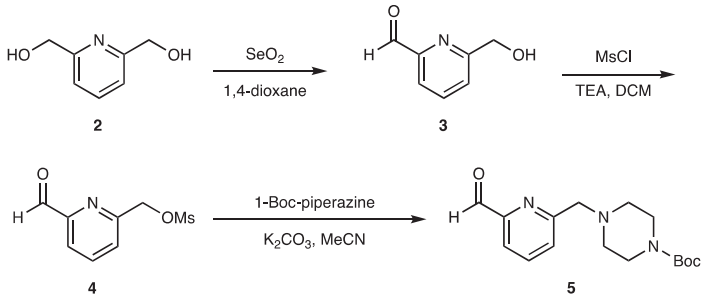
Rosetta 2 (DE3) cells (Novagen) containing SMAC constructs in pET-DEST42 vectors were grown in TB media containing 75 mg/L carbenicillin with shaking at 37°C to an OD_{600} of 0.5. The cells were then induced with 1 mM IPTG and harvested 4 h later by centrifugation. Cells were lysed by stirring for 20 min at 4°C in DUB enzyme buffer (50 mM HEPES, pH 7.5, 0.5 mM EDTA, 1 mM DTT) supplemented with 1 mg/mL lysozyme. The lysate was then sonicated and centrifuged at 10,000 × *g* for 10 min. 15 μL of USP2 (60 nM final concentration) was added to the lysates, which were incubated at room temperature for 5 h. Lysates were then buffer exchanged using Sephadex G-25M columns (GE Healthcare). Talon cobalt affinity resin (Clontech; 400 μL of slurry/g of cell paste) was added to the supernatant, and the mixture was rotated at 4°C overnight. Beads were collected by centrifugation at 700 × *g* for 3 min, washed three times with PBS, and applied to a 1 cm column. The column was washed twice with PBS buffer (10 mL/400 μL of resin slurry). The bound protein was eluted by the addition of 200 mM imidazole (2 mL/400 μL of resin). Imidazole was removed by passage over a Sephadex G-25M column (GE Healthcare). Protein concentrations were determined using the Bio-Rad DC Protein Assay kit. Glycerol was added to a final concentration of 10%, and proteins were aliquoted stored at −80°C until use.

MALDI-TOF

DHBA matrix was prepared freshly by saturating 2,5-Dihydroxybenzoic acid (Sigma) in 40% acetonitrile in water with 0.1% TFA. 10 μM peptide samples in water were diluted 1:10 with 5% acetonitrile in water with 0.1% TFA. 0.5 μL of the resulting solution was spotted on the MALDI plate followed by 0.5 μL of saturated DHBA matrix. MALDI-TOF was conducted on a Bruker Autoflex instument. Data was collected and analyzed using "Compass for flexSeries" v1.4, "flexControl" v3.4 and "flexAnalysis" v3.4 software.

Chemical Synthesis

All reagents used for chemical synthesis were purchased from Sigma-Aldrich or Alfa Aesar unless otherwise specified and used without further purification. All anhydrous reactions were performed under nitrogen atmosphere. Analytical thin-layer chromatography (TLC) was conducted on EMD Silica Gel 60 F254 plates with detection under UV (254 nm) or by staining with potassium permanganate ($KMnO_4$). NMR spectra were recorded on a Bruker UltraShield Plus 500 MHz Avance III NMR. Chemical shifts are recorded in ppm (δ) relative to solvent. Coupling constants (*J*) are reported in Hz. NMR spectra were processed using Mnova (www.mestrelab.com/software/mnova-nmr) software. High-resolution mass spectra were obtained on a Waters Acuity Premiere XE TOF LC-MS by electrospray ionization.



Scheme S1. Synthetic route for **5** according to literature procedure (MacDonald et al., 2015)

Compound 3

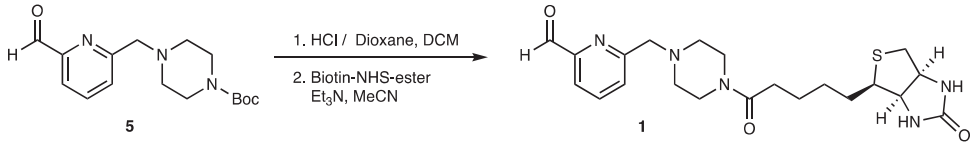
6-(hydroxymethyl)-2-pyridinecarboxaldehyde (MacDonald et al., 2015). To a stirred solution of 2,6-pyridinedimethanol **2** (3.0 g, 21.6 mmol) in 1,4-dioxane (60 ml) was added selenium dioxide (1.2 g, 10.8 mmol). The mixture was sonicated for 2 min and then stirred at 65°C for 24 h. The reaction was cooled to ambient temperature, diluted with dichloromethane (100 ml), filtered through a plug of Celite, and concentrated under reduced pressure. This crude material was then purified by flash chromatography using 2.5% methanol in dichloromethane to afford 2.63 g (89%) of the title compound as an off-white solid. ^{1}H NMR (500 MHz, Chloroform-d): δ 10.08 (s, 1H), 7.92–7.85 (m, 2H), 7.57–7.51 (m, 1H), 4.86 (s, 2H), 3.34 (br s, 1H). ^{13}C NMR (500 MHz, Chloroform-d): δ 193.2, 160.1, 151.7, 137.7, 125.1, 120.6, 64.1. HRMS (ESI) calculated for $C_7H_8N_1O_2$ ([M+H]$^+$) 138.0550, found 138.0553.

Compound 4

6-(formylpyridin-2-yl)methyl methanesulfonate (MacDonald et al., 2015). Methanesulfonyl chloride (1.8 ml, 22.8 mmol) was added dropwise at 0°C to a stirred solution of 6-(hydroxymethyl)-2-pyridinecarboxaldehyde **3** (2.63 g, 19 mmol) and triethylamine (5.4 ml, 57 mmol) in dichloromethane (65 ml). After 1 h, the reaction was quenched with saturated aqueous sodium bicarbonate. The aqueous layer was separated and extracted three times with dichloromethane. The organic layers were combined, dried over sodium sulfate, filtered and concentrated under reduced pressure to afford 3.25 g (78%) of the title compound as a brown oil. ^{1}H NMR (400 MHz, Chloroform-d): δ 10.03 (s, 1H), 7.96–7.92 (m, 2H), 7.75–7.72 (m, 1H), 5.40 (s, 2H), 3.13 (s, 3H). ^{13}C NMR (100 MHz, $CDCl_3$): δ, 192.6, 154.5, 152.6, 138.6, 126.4, 121.9, 70.6, 38.1. HRMS (ESI) calculated for $C_8H_{10}N_1O_4S_1$ ([M+H]$^+$) 216.0325, found 216.0321.

Compound 5

Tert-butyl 4-((6-formylpyridin-2-yl)methyl)piperazine-1-carboxylate (MacDonald et al., 2015). To a stirred solution of 6-(formylpyridin-2-yl)methyl methanesulfonate **4** (3.2 g, 15 mmol) and 1-Boc-piperazine (3.25 g, 15 mmol) in acetonitrile (65 ml) was added potassium carbonate (4.23 g, 30 mmol). The mixture was stirred at 60°C for 16 h at which time the solvent was removed under reduced pressure. The resulting residue was then partitioned between dichloromethane and saturated aqueous sodium bicarbonate. The aqueous layer was separated and extracted three times with dichloromethane. The combined organic layers were dried over sodium sulfate, filtered and concentrated under reduced pressure. The crude material was purified by flash chromatography with 33% ethyl acetate in hexanes to afford 2.18 g (48%) of the title compound as a tan solid. ^{1}H NMR (500 MHz, Chloroform-d) ∂ 10.0 (s, 1H), 7.81 – 7.75 (m, 2H), 7.63 (p, *J* = 4.0 Hz, 1H), 3.71 (s, 2H), 3.40 (t, *J* = 5.1 Hz, 4H), 2.42 (t, *J* = 5.1 Hz, 4H), 1.39 (s, 9H). ^{13}C NMR (500 MHz, Chloroform-d) ∂ 193.59, 159.43, 154.75, 152.38, 137.45, 127.40, 120.31, 79.71, 64.10, 53.12, 43.65, 28.42. HRMS (m/z) calculated for $C_{16}H_{23}N_3O_3$ [M+H$^+$] 306.1812, found 306.1809.



Scheme S2. Synthetic route for **1**

Compound 1

6-((4-(5-((3a*S*,4*S*,6a*R*)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl) pentanoyl) piperazin-1-yl) methyl)picolinaldehyde. Compound **5** (153 mg, 0.5 mmol) was dissolved in dichloromethane (5 ml), to which 4M HCl in 1,4 dioxane (1.4 ml) was added and the reaction was allowed to stir for 2 h at room temperature. The resulting white precipitate was then concentrated under vacuum and the solids were triturated with dichloromethane and dried under reduced pressure. The resulting hydrochloride salt was suspended in acetonitrile (4 ml) with biotin-NHS-ester (204 mg, 0.6 mmol). Triethylamine (0.21 ml, 1.5 mmol) was added dropwise and the solution was left stirring at room temperature overnight. After 16 h, the solvent was removed under reduced pressure and the resulting residue was partitioned into dichloromethane and saturated aqueous sodium bicarbonate. The aqueous layer was separated and extracted three times with dichloromethane. The combined organic layers were dried over sodium sulfate, filtered, and concentrated under reduced pressure. The crude material was purified by flash chromatography on a Teledyne ISCO CombiFlash with a gradient of methanol (5% to 15%) in dichloromethane to afford 56.3 mg (26%) of the title compound as an off-white solid. ^{1}H NMR (500 MHz, Chloroform-d) ∂ 10.03 (s, 1H), 7.84 (d, *J* = 5.7 Hz, 2H), 7.71 – 7.63 (m, 1H), 6.25 (s, 1H), 5.78 (s, 1H), 4.47 (dd, *J* = 7.8, 4.9 Hz, 1H), 4.27 (dd, *J* = 8.0, 4.5 Hz, 1H), 3.75 (s, 2H), 3.61 (d, *J* = 5.3 Hz, 2H), 3.48 (t, *J* = 5.0 Hz, 2H), 3.13 (q, *J* = 2.9 Hz, 1H), 2.92 – 2.80 (m, 1H), 2.70 (d, *J* = 12.7 Hz, 1H), 2.49 (dt, *J* = 19.0, 5.1 Hz, 4H), 2.33 (t, *J* = 7.5 Hz, 2H), 1.73 – 1.61 (m, 4H), 1.44 – 1.36 (m, 2H). ^{13}C NMR (500 MHz, Chloroform-d) ∂ 192.53, 170.50, 162.93, 158.16, 151.38, 136.55, 126.45, 119.45, 62.88, 60.87, 59.19, 54.50, 52.46, 52.01, 44.56, 40.48, 39.55, 31.74, 27.45, 27.30, 24.13. HRMS (m/z) calculated for $C_{21}H_{29}N_5O_3S$ [M+H$^+$] 432.2069, found 432.2062.

QUANTIFICATION AND STATISTICAL ANALYSIS

Proteomic statistical analysis was performed with pFIND. GraphPad Prism 7 software was used for SEM calculations. WebLogo 3.0 was used to generate sequence logos. For *in vitro* cell assays, n refers to number of replicate wells analyzed for each treatment group. Error bars for *in vitro* assay figures represent SEM from mean.

DATA AND SOFTWARE AVAILABILITY

MS-CHOPS proteomic data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (Perez-Riverol et al., 2019). The accession number for the data reported is: PXD013019.