

Small Molecule Interactome Mapping by Photoaffinity Labeling **Reveals Binding Site Hotspots for the NSAIDs**

Jinxu Gao, Adelphe Mfuh, Yuka Amako, and Christina M. Woo*

Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford St., Cambridge, Massachusetts 02138, United States

Supporting Information

ABSTRACT: Many therapeutics elicit cell-type specific polypharmacology that is executed by a network of molecular recognition events between a small molecule and the whole proteome. However, measurement of the structures that underpin the molecular associations between the proteome and even common therapeutics, such as the nonsteroidal antiinflammatory drugs (NSAIDs), is limited by the inability to map the small molecule interactome. To address this gap, we developed a platform termed small molecule interactome mapping by photoaffinity labeling (SIM-PAL) and applied it to the in cellulo direct characterization of specific NSAID binding



sites. SIM-PAL uses (1) photochemical conjugation of NSAID derivatives in the whole proteome and (2) enrichment and isotope-recoding of the conjugated peptides for (3) targeted mass spectrometry-based assignment. Using SIM-PAL, we identified the NSAID interactome consisting of over 1000 significantly enriched proteins and directly characterized nearly 200 conjugated peptides representing direct binding sites of the photo-NSAIDs with proteins from Jurkat and K562 cells. The enriched proteins were often identified as parts of complexes, including known targets of NSAID activity (e.g., NF- κ B) and novel interactions (e.g., AP-2, proteasome). The conjugated peptides revealed direct NSAID binding sites from the cell surface to the nucleus and a specific binding site hotspot for the three photo-NSAIDs on histones H2A and H2B. NSAID binding stabilized COX-2 and histone H2A by cellular thermal shift assay. Since small molecule stabilization of protein complexes is a gain of function regulatory mechanism, it is conceivable that NSAIDs affect biological processes through these broader proteomic interactions. SIM-PAL enabled characterization of NSAID binding site hotspots and is amenable to map global binding sites for virtually any molecule of interest.

INTRODUCTION

Polypharmacology, wherein one drug interacts with multiple protein targets, is an outcome of molecular recognition events between small molecules and the whole proteome. Polypharmacology manifests in increased efficacy when properly exploited or tragic unanticipated off-target effects when not fully understood. Many pharmaceuticals in diverse therapeutic areas possess either known or uncharacterized polypharmacology,¹ such as the nonsteroidal anti-inflammatory drugs (NSAIDs),² the immunomodulatory drugs,³ or the opioids.⁴ Differences in protein networks across cell types form the basis for molecular interactions that culminate in an observed phenotype,⁵ suggesting that a map of the protein-ligand interaction network may eventually predict these polypharmacology outcomes.^{6,7}

A method to directly map the noncovalent small molecule interactome has the potential to accelerate drug discovery by providing structural insight and instant validation of the binding interaction, yet such a characterization is rarely performed. Common methods to structurally reveal small molecule binding sites, such as X-ray crystallography or NMR spectroscopy, are constrained to the measurement of stable interactions between a single compatible protein and small molecule pair in vitro.

Global proteomic profiles from small molecule affinity purification strategies are commonly obtained using mass spectrometry (MS);⁸ however, the vast majority of proteomics studies stop short of obtaining direct structural evidence for the molecular interaction. Inherent challenges have prevented the mapping of small molecule binding sites to the whole proteome. Small molecule interactions occur over a range of concentrations that require a general mechanism for capture and enrichment prior to MS analysis. The conjugation chemistry to capture the binding event must be rapid and general for unbiased covalent bond formation at the small molecule binding site. Binding sites of small molecules to defined protein isolates can be determined by application of photoaffinity labeling (PAL) to conjugate the small molecule to the protein prior to MS analysis.⁹ Yet, the demand for a general chemical strategy to covalently conjugate a small molecule locally to the protein interaction site poses great challenges to spectral assignment by database searching. Database searching methods are not adapted to the computational complexity

Received: November 2, 2017

Journal of the American Chemical Society

yielded by amino acid residue-agnostic conjugation chemistry to the whole proteome.

Translation of MS-based binding site identification by PAL from a single protein to the whole proteome thus requires (1) a selective chemical workflow to isolate the conjugated peptide and (2) a targeted MS technique for confident characterization. PAL covalently conjugates small molecules to the proteome for stringent enrichment of interacting proteins.¹⁰ Application of cleavable enrichment handles enables recovery of the small molecule-conjugated peptide.⁹ Critically, a targeted MS strategy, wherein unique isotopic markers are installed specifically to the small molecule conjugated peptide, provides an orthogonal handle for detection and validation that is transformative during database assignment of peptides carrying heterogeneous modifications by MS.¹¹ Recent strides have enabled the identification of fragment-based small molecule ligands to the whole proteome.¹²

The knowledge gap caused by the lack of a small molecule interactome map extends to common pharmaceuticals like the NSAIDs. The NSAIDs potently suppress inflammation, pain, and fever, and have been further explored as potential treatments for cancer^{2,13} and Alzheimer's disease.¹⁴ NSAID mechanisms have been primarily characterized through inhibition of the enzymes cyclooxygenase-1 and -2 (COX-1, COX-2, respectively).¹⁵ Inhibition of COX-2, the primary cyclooxygenase involved in inflammation, prevents the production of prostaglandins, thereby reducing inflammation. However, a wealth of biomedical evidence points to broader COX-2-independent mechanisms of the NSAIDs for which a molecular basis remains poorly defined.^{16–18} Prior studies suggest that specific NSAIDs inhibit the nuclear factor- κ B (NF- κ B) pathway¹⁹ and caspases.²⁰ A detailed understanding of NSAID—protein interactions and their underlying structures.

Herein, we report the development of a platform termed small molecule interactome mapping by photoaffinity labeling (SIM-PAL) and its application to the NSAIDs. SIM-PAL is designed to characterize the protein interactions and direct binding site hotspots of a small molecule in a whole cell proteome using a PAL-based enrichment strategy coupled to isotope-targeted MS (Figure 1A). Our platform involves: (1) photoconjugation of NSAID derivatives in cells, (2) enrichment and isotopic recoding of NSAID-labeled peptides, and (3) isotope-targeted assignment of the conjugated peptides. We show that photo-NSAID derivatives are effective reporters of NSAID binding sites with recombinant COX-2 and the global whole cell proteome from Jurkat and K562 cell lines. By virtue of direct characterization of the conjugated peptide, we localized the photo-NSAIDs to nearly 200 binding sites from over 150 proteins, including histones H2A and H2B. Histone H2A was stabilized by interacting with the NSAIDs by cellular thermal shift assay. SIM-PAL revealed the precise binding sites for the photo-NSAIDs via an approach that is readily translated to broad classes of small molecules.

RESULTS

Development of Photo-NSAIDs as Reporters of NSAID Binding Sites. Three NSAIDs, naproxen (1), celecoxib (2), and indomethacin (3), were conjugated to diazirine-based photoaffinity labels ("photo-NSAIDs") to serve as reporters for NSAID binding sites (Figure 1B). These NSAIDs were selected for their different structure-activity relationship (SAR) between COX-1 and COX-2.^{21,22} Naproxen (1) is a non-



Figure 1. Strategy to profile NSAID interactome by SIM-PAL. (A) Photo-NSAIDs are applied in cellulo and photoconjugated to protein binding partners. Conjugated proteins are tagged by probe **10** using CuAAC and enriched to separately obtain the protein interactome and conjugated peptides representing binding site hotspots. Conjugated peptides are analyzed by isotope-targeted MS. (B) Structures of NSAIDs naproxen (1), celecoxib (2), and indomethacin (3), their photo-NSAID analogs **4–6**, and the negative controls tag 7, the orthogonal compound photoglutarimide **8**, and the celecoxib analog **9**. COX-2 IC₅₀ by ELISA is shown below. (C) Structure of the cleavable biotin azide probe **10**. Probe **10** is prepared as a 1:3 mixture of stable ¹²C:¹³C isotopes (highlighted in red).

Journal of the American Chemical Society

selective COX-1 and COX-2 inhibitor, but is commonly employed for chronic use due to low rates of gastrointestinal side effects.²³ Celecoxib (2) is a selective COX-2 inhibitor developed by Pfizer, yet possesses off-target cardiovascular and gastrointestinal complications.²⁴ Indomethacin (3) is a member of the indole class of NSAIDs and possesses known COX-2independent anti-inflammatory mechanisms.¹⁶ The design of photo-NSAIDs was based upon previous SAR studies^{25,26} and the crystal structure between mouse COX-2 and indomethacin (3).²⁷ During our studies, a crystal structure of a NSAID with human COX-2 was reported.²⁸ In addition to photo-NSAIDs **4–6**, we developed the tag 7,²⁹ a structurally orthogonal photoglutarimide **8**, and a celecoxib analog **9** to assess selectivity of the binding site identification assay (Scheme S1).

All three photo-NSAIDs maintained COX-2 inhibition by ELISA (Figure 1B, Figure S1),³⁰ although some variation in activity was observed. Photocelecoxib (5) was the most potent $(IC_{50} = 36.6 \text{ nM})$, and photonaproxen (4) was the least potent (IC₅₀ = 36.0 μ M). All photo-NSAIDs possessed antiproliferative properties within 1.3-1.4-fold of their parent compound in Jurkat cells (Figure S2). Elimination of the sulfonamide from celecoxib (2) to give the analog 9 is known to reduce COX-2 inhibition,²⁶ and we found additionally attenuated antiproliferative activity of the analog 9 in Jurkat cells (Figure S2). Furthermore, all photo-NSAIDs were competitively displaced from recombinant COX-2 by the native NSAID (Figure 2A). COX-2 was separately incubated with each of the photo-NSAIDs with or without a 100-fold excess of the parent compound as a competitor.¹⁰ The samples were photoirradiated, tagged with the fluorophore TAMRA-azide by copper-mediated azide-alkyne cycloaddition (CuAAC) and fluorescently visualized to reveal selective and reversible binding of COX-2 to the photo-NSAIDs. The tag 7 did not produce observable conjugation to COX-2 by fluorescence.

Photo-NSAIDs Possess Known and Transient Binding Sites with COX-2. We used recombinant COX-2 to validate fragmentation patterns of conjugated NSAIDs to a protein and to determine binding site selectivity for each of the photo-NSAIDs. Small molecule modification on a peptide may perturb MS fragmentation pathways in unexpected ways, rendering the spectra unassignable by database searching. To evaluate this possibility, photo-NSAIDs (10 μ M) were incubated with recombinant COX-2 $(1 \mu g)$ for 30 min and photo-irradiated. The irradiated samples were appended to the cleavable biotin azide probe 10 to simulate the conjugated species ultimately observed after enrichment (Figure 1C). The probe 10 is a multifunctional probe developed to possess a biotin affinity enrichment handle, an acid-labile diphenylsilane, and a stable isotope-coded azidoacetate for CuAAC and isotope-targeted MS. We previously established the compatibility of a similar cleavable probe scaffold in targeted MS experiments.¹¹ Following CuAAC, the samples were trypsindigested and the biotin probe was cleaved in situ (2% formic acid). The resulting peptides were analyzed by LC-MS/MS. MS data was searched by SEQUEST HT against recombinant COX-2 with the photo-NSAID as a modification on any amino acid (Table S1). Due to the nature of photochemical conjugation, a binding site may be represented by multiple conjugation events to several amino acid residues on one or more peptides. All peptide spectral matches (PSMs) assigned to a conjugated peptide were manually validated.

Photo-NSAIDs were readily assigned by database searching. Manual inspection of these PSMs indicated that in the case of



Figure 2. Comparative analysis of NSAIDs and photo-NSAIDs with recombinant COX-2. A. Fluorescence image of photo-NSAID binding to COX-2 and competitive displacement by the respective parent compound. COX-2 (125 ng) was incubated with photo-NSAIDS with or without the parent molecule, photo-conjugated, and tagged with TAMRA-azide. B. Docking structure of photonaproxen (4, red) and naproxen (1, blue) with COX-2. C. Docking structure of photocelecoxib (5, green) and celecoxib (2, blue) with COX-2. D. Docking structure of photoindomethacin (6, yellow) and indomethacin (3, blue). E. Docking structure of the tag 7 (purple) with COX-2. The blue box indicates the region of COX-2 that is enlarged in sections B-D. COX-2 $(1 \mu g)$ was separately conjugated to each of the photo-NSAIDs (10 μ M) or the tag 7 (10 μ M), tryptically digested, and analyzed by MS on an Orbitrap Elite. Conjugated peptides observed by MS are highlighted for each photo-NSAID. Docked structures were either the lowest desolvation energy or highest interface area size created by Patchdock (October, 2017). Structure of COX-2 from PDB: 5KIR.

the photo-NSAIDs 4–6, no irregular fragmentation pathways were observed. As expected, at a consistent dose of 10 μ M across all photo-NSAIDs, photocelecoxib (5) possessed the highest number of observed conjugated peptides (seven conjugated peptides from 14 PSMs), including within the active site of COX-2. Analysis of COX-2 treated by photonaproxen (4) and photoindomethacin (6) revealed six

conjugated peptides across 19 PSMs and eight PSMs, respectively (Table S1). By contrast, the tag 7 was found conjugated to one peptide, which was not marked by the photo-NSAIDs. Within each peptide, the specific conjugation site localized to a range of 2-4 amino acid residues that reflect the specificity of the photochemical conjugation event.

The photo-NSAIDs, the parent compounds, and the tag 7 were structurally minimized [Gaussian 16, basis set: HF 3-21g(d)] and individually docked in the crystal structure of human COX-2 (PDB: 5KIR).²⁸ Docking was performed using Patchdock, a molecular docking algorithm based on shape complementarity using rigid structures.³¹ Structures with the lowest desolvation energy or highest interface area size docked the photo-NSAID or parent compound to the same binding site, although the orientation between each pair of compounds differed (Figure 2B-D). The docked structure of photonaproxen (4, red), photocelecoxib (5, green), and photoindomethacin (6, yellow) overlaid with their respective parent compound (blue). Conjugated amino acid residues within 5 Å of the docked photo-NSAID are highlighted in Figure 2B-D. The orientation of each of the photo-NSAIDs positioned the diazirine in close proximity to one or two specific amino acids on a conjugated peptide observed by MS. Other marked residues were located on solvent exposed areas of the protein. As photochemistry captures dynamic processes and the photochemical tag is structurally flexible, the additional conjugation events on recombinant COX-2 may represent transient interactions with COX-2 in vitro, which may not be observed in cellulo. The tag 7 was additionally docked to COX-2 and, in combination with the observed MS data, was shown to transiently bind to an orthogonal region of COX-2 (Figure 2E). No conjugated peptides from amino acids near to the canonical binding site were observed with the tag 7.

Characterization of the NSAID interactome in Jurkat and K562 cells. Confident that photo-NSAIDs were producing defined linkages with COX-2 that recapitulated NSAID activity, we next sought to characterize global NSAID interactions within the whole cell proteome. Photo-NSAIDs, the tag 7, and the photoglutarimide 8 were added to activated Jurkat T cells as a model system for inflammation. Small molecule labeling was photo-irradiation-dependent and dosedependent up to 250 μ M (Figure S3A,B). Competitive displacement of the tightest COX-2 binder photocelecoxib (5) by the parent compound occurred at a 1:10 molar ratio (Figure S3C). Jurkat cells were stimulated with phorbol myristyl acetate and ionomycin for 18 h prior to photo-NSAID exposure.³² The stimulated Jurkat cells were exposed to each compound (250 μ M, 1 h) and photo-irradiated in situ to conjugate the small molecule to the proteome. The resulting NSAID-conjugated proteins were enriched using the probe 10 in a biotin-dependent manner (Figure S4). COX-2 from Jurkat cells was enriched by photo-NSAIDs, and this enrichment was abrogated by competition with the parent compound, as indicated by Western blot (Figure S5A). Cellular thermal shift assay, a mechanism to measure stabilization of protein-ligand interactions,³³ showed that COX-2 was stabilized by celecoxib (2) and photocelecoxib (5) at an apparent aggregation temperature (T_{agg}) over 10 °C relative to the tag 7 (Figure S5B,C).

The enriched proteomes were digested with trypsin, and the released peptides were analyzed by LC-MS/MS on an Orbitrap Fusion Tribrid with collision-induced dissociation (CID) and higher energy CID (HCD) fragmentation modes. MS data

were assigned by SEQUEST HT. Two biological replicates were collected for each of the photo-NSAIDs that displayed high reproducibility across the enriched proteome (>60%) and protein abundance (Figure S6A). Using the normalized spectral abundance factor (NSAF) for PSM-based label free quantification, $^{34-36}$ proteins that were statistically significant (*t* test, pvalue <0.05) and the PSM ratio was \geq 2-fold enriched were considered significantly enriched by the photo-NSAID relative to the control tag 7. Following application of the NSAF, a standard distribution of values was obtained to test for statistical significance. The majority of proteins that passed statistical significance were also enriched by the photo-NSAIDs as compared to the tag 7 (Figure 3A). This analysis yielded approximately 700 proteins significantly enriched by at least one of the photo-NSAIDs in Jurkat cells (Table S2). High proteomic overlap among photo-NSAIDs, but not negative controls 7 and 8, was observed. Across each of the three compounds, at least 40% of the identified photo-NSAIDbinding proteins were enriched by all three photo-NSAIDs and 53% of the proteins were enriched by at least two photo-NSAIDs (Figure 3B). By comparison, 140 proteins that were significantly enriched by any photo-NSAID were also enriched by the photoglutarimide 8 (24%).

Jurkat cells were initially dosed with a concentration of 250 μ M of each photo-NSAID to maximize downstream observation of conjugated peptides by MS. Although naproxen (1) enters blood plasma at concentrations that surpass 250 μ M, our cell viability data against Jurkat cells showed an IC₅₀ range of 23–216 μ M across the NSAIDs and their derivatives (Figure S2). We thus examined the NSAID interactome at 50 μ M and found 260 proteins enriched only by the three photo-NSAIDs and not by the tag 7 (Table S3). Enrichment was determined by inclusion of proteins only observed in the photo-NSAID treated samples. At 50 μ M, the proteomic overlap between the photo-NSAIDs was lower, indicative of higher binding selectivity between the ligands (Figure S6B). A majority of proteins identified at 50 μ M were likewise identified at 250 μ M for each photo-NSAID (86–92%). The celecoxib analog 9 (50 μ M) was additionally tested in activated Jurkat cells and displayed moderate proteomic overlap with photocelecoxib (5)(Figure S6C).

To determine the generality of these observations, photo-NSAID interactions were measured against K562 cells, a human chronic myeloid leukemia cell line. Several lines of evidence point to NSAID-dependent inhibition and apoptosis of K562 cells.³⁸ A total of 513 proteins were significantly enriched across two biological replicates from K562 cells, of which 42% of the proteins were enriched by at least two of the three photo-NSAIDs (Figures S7 and S8). Significantly enriched proteins from K562 cells possessed a moderate overlap with proteins from Jurkat cells (206 proteins), indicating a high degree of specificity across cell lines (Figure 3C). In sum, a total of 1033 proteins were considered significantly enriched from Jurkat and K562 cells by photo-NSAIDs (Table S2).

Photo-NSAIDs interacted with proteins distributed throughout the cell (Figure 3D). Approximately two-thirds of the NSAID interactome localized to the nucleus and cytoplasm. Photo-NSAIDs additionally captured proteins annotated as localized to the mitochondria (10%), endoplasmic reticulum and Golgi (11%), and membrane or secreted proteins (10%). These data are a close reflection of the natural distribution of proteins throughout the cell.³⁹ Only 30% of these proteins were previously annotated as interacting with a small molecule, let



- AP2A2 celecoxib (2) 10.4 AP2B1 photo-celecoxib (5) 12.1 AP2 adaptor complex DMSO 20.3 Figure 3. Photo-NSAID protein interactome at 250 μ M. (A) Fold change enrichment of the photo-NSAIDs against the tag 7 based on PSMs. Proteins are ordered in descending fold change ratio. Proteins that were not statistically significant across the two biological replicates are not displayed. (B) Overlap across the Jurkat photo-NSAID significantly enriched proteome. (C) Proteomic overlap between significantly enriched proteins from Jurkat and K562 cells. (D) Subcellular localization of the 1033 significantly enriched proteins. (E) Selected protein interaction networks captured by at least one photo-NSAID. Color scheme: red = significantly enriched; pink = \geq 2-fold

enriched; gray = identified in data; white = member of the protein complex not identified in data. Dashed border indicates direct observation of at least one conjugated peptide. Gray lines indicate the protein interaction network, built using CORUM as a reference (July, 2017). Half maximal effective concentration (EC₅₀) values for inhibition of the NF- κ B pathway as determined by a NF- κ B luciferase reporter assay shown below.

alone one of the NSAIDs (BindingDB, ChEMBL, DrugBank), pointing to the broader range of molecular interactions that remain to be revealed by the reported binding site mapping approach.

In line with evidence for capture of a broader range of associated proteins, a number of protein complexes were selectively enriched. Comparison of enriched proteins to CORUM⁴⁰ revealed that NF- κ B subunits (NFKB1, NFKB2) are directly interacting with the photo-NSAIDs, confirming their known inhibition of the NF- κ B pathway,² which was revalidated for celecoxib (2) and photocelecoxib (5) via a NF- κ B luciferase reporter assay (Figure 3E, Figure S9). Additional

protein complexes that were enriched include members of the proteasome and the adaptor protein complex 2 (Figure 3E). These proteins may have existed as complexes leading to photo-NSAID conjugation by virtue of proximity or were enriched due to associative protein interactions. While protein complexes can be enriched through associative protein—protein interactions, our use of strong dissociative detergents to prepare cell lysates (1% Rapigest, sonication) does not typically lead to observation of protein complexes following enrichment.¹¹

Direct Photo-NSAID Interaction Mapping Reveals Binding Site Hotspots. Following tryptic digestion of the photo-NSAID proteome, the probe 10 was acid cleaved to release and isolate the conjugated peptides from the enrichment media. The conjugated peptides were analyzed by application of isotope-targeted MS. A unique isotopic signature was embedded to the probe 10 using a ¹³C-derived stable isotope ratio of 1:3 over [M:M + 2] spacing to perform isotope-targeted MS. During click chemistry, the unique isotopic signature was exclusively transferred to the photo-NSAIDs conjugated to the proteome. The isotopic signature is therefore only found on small molecule-conjugated peptides and is used during isotope-targeted MS to overcome barriers in detection of low abundance species and validation of modified peptides against a background of unmodified peptides.

During data collection, the isotopically recoded species is immediately recognizable by full scan MS and may be selected for fragmentation by use of an inclusion list.^{41,42} This selection process increases the fraction of isotopically recoded, small molecule-conjugated peptides selected for fragmentation.¹ More critically, the isotope signature played a crucial role in manual validation of database search assignments due to the ambiguity of the amino acid modification site. Database searching was performed against the tryptic protein database with each of the photo-NSAIDs as a modification on any amino acid residue. Small molecule conjugation to any amino acid drastically increases the size of the protein database. For example, a single modification increases the size of the fully tryptic human protein database by 60-fold, while two modifications increase the database size by 1000-fold. This exponential increase in the tryptic peptide database leads to a breakdown in false discovery rate (FDR).43 We thus used a two-tier validation process for confident assignment of conjugated peptides. First, modified peptides at 5% FDR were filtered based on visual inspection of the MS2 spectral assignment. Second, each of the precursor spectra were individually validated for the isotopic signature in the MS1. Based on this analysis, 575 PSMs, corresponding to 194 individual conjugated peptides, were characterized across the photo-NSAIDs, tag 7, and photoglutarimide 8 (Table S4). Selected frequently observed conjugated peptides are displayed in Table 1.

The individual peptides derived from approximately 150 proteins, of which over 90% of the proteins were \geq 2-fold enriched. Forty-three of these proteins were considered significantly enriched by at least one of the photo-NSAIDs. A number of isotopically coded species that were either not selected or not confidently assigned by SEQUEST were also observed. Detection of enriched conjugated species using a pattern searching algorithm⁴² revealed nearly 1000 isotopically coded precursor ions in the MS1 spectra across our photo-NSAID conjugated peptide data.

Each ligand produced a unique profile of conjugated peptides. Photocelecoxib (5) and photonaproxen (4) repre-

			count of conjugated peptide PSMs				
	conjugated peptide	protein (GENE)	4	5	6	7	8
1	AMGIMNSFVNDIFER	histone H2B type 1-K (HIST1H2BK)	41	12	1		
2	VGAGAPVYLAAVLEYLTAEILELAGNAAR	histone H2A type 2-B (HIST2H2AB)	8	28	9		
3	VAPEEHPVLLTEAPLNPK	actin, cytoplasmic 1 (ACTB)	3	14	3		
4	VGAGAPVYMAAVLEYLTAEILELAGNAAR	histone H2A type 2-A (HIST2H2AA3)	2	12	5		
5	MSVQPTVSLGGFEITPPVVLR	Nucleophosmin (NPM1)	7	7	1		
6	NLEALALDLMEPEQAVDLTLPK	X-ray repair cross-complementing 6 (XRCC6)	7	3		1	2
7	IHFPLATYAPVISAEK	tubulin α -1A chain (TUBA1A)		13			
8	VGAGAPVYLAAVLEYLTAEILELAGNAARDNKK	histone H2A type 2-B (HIST1H2AB)	5	8			
9	ISGLIYEETR	histone H4 (HIST1H4A)	2	10	1		
10	VETGVLKPGMVVTFAPVNVTTEVK	elongation factor 1- α 1 (EEF1A1)	5	7			
11	AIGAVPLIQGEYMIPCEK	Cathepsin D (CTSD)	4	2	3		
12	AFADAMEVIPSTLAENAGLNPISTVTELR	T-complex protein 1 subunit delta (CCT4)	5		1		
13	NASDMPETITSR	minor histocompatibility antigen H13 (HM13)		2	2		
14	AFLASPEYVNLPINGNGK	glutathione S-transferase P (GSTP1)	4				
15	LKNDQANYSLNTDDPLIFK	adenosine deaminase (ADA)		4			
16	FGPYYTEPVIAGLDPK	proteasome subunit β type-3 (PSMB3)	3				
17	EGNDLYHEMIESGVINLK	ATP synthase β , mitochondrial (ATP5B)		2			
^{<i>a</i>} Presente	d data are in aggregate across Jurkat and K562 cells	. For the full dataset, see Table S4.					

Table 1. Selected Conjugated Peptides Observed with High PSM Frequency Across Photo-NSAIDs, the Tag 7, and the Photo-Glutarimide 8^a

sented the bulk of the identified interactions and were found conjugated to 93 and 85 peptides, respectively. Photoindomethacin (6) was conjugated to 34 peptides in total. A degree of overlap between the photo-NSAID-conjugated peptides was observed, where 30 peptides were identified by at least two of the three photo-NSAIDs. Of these, a single conjugated peptide on Ku70, a member of the DNA repair pathway, was observed by two photo-NSAIDs, the tag 7 and photoglutarimide **8** (entry 6, Table 1). The remaining 14 conjugated peptides detected by the tag 7 and nine conjugated peptides from the photoglutarimide **8** were detected exclusively by that compound.

Conjugated peptides were mapped to proteins found throughout the cell. Peptides from nuclear proteins (entries 1, 2, 4, 5, 6, 8, 9, 14) were most frequently observed (Table 1). Mitochondrial proteins (entry 17) and cell surface or secreted proteins (entries 11, 12, 13) were also observed. NSAID interactions were mapped to diverse protein types, including enzymes (entries 6, 11, 14, 15, 17), parts of the proteasome (entries 11, 16), and structural proteins (entries 3, 7), which demonstrates the unbiased discovery power of MS. In many cases, these interactions were highly specific and observed with only one of the photo-NSAIDs (entries 7, 14, 15, 16, 17).

We found a number of PSMs assigned to NSAID-conjugated histone peptides (Figure 4A, Figure S10). All directly observed histone interactions were highly interconnected by clustering and imply significantly upregulated conjugation of histone complexes by the photo-NSAIDs (Figure 4B). These proteins were enriched in photo-NSAID proteomic data as compared to the tag 7, but were not considered statistically significant. In particular, two peptides from histone H2A and histone H2B were primarily detected. The histone H2B peptide was detected in a total of 54 PSMs across our data sets, predominantly by photonaproxen (4, entry 1, Table 1). Histone H2A type 2-A and type 2-B were detected in a total of 62 and 19 PSMs, respectively (entries 2, 4, and 8, Table 1), predominantly in Jurkat cells. These PSMs related to peptides conjugated to photocelecoxib (5), followed by photonaproxen (4), and to a lesser degree photoindomethacin (6). By virtue of directly

observing the conjugated peptide, photo-NSAIDs were mapped to a specific binding site hotspot around these two peptides, which are in close proximity in structures of the nucleosome (Figure 4C).⁴⁴

Validation of SIM-PAL data by competition assay revealed that binding of all three photo-NSAIDs with histone H2A and H2B was competitively exchanged with the parent compound by Western blot (Figure 5A). The interaction of photo-NSAIDs with NF- κ B p65, cathepsin D, and nucleophosmin was additionally displaced by the parent compound. Selective photoconjugation and partial competition was observed when probing for Ku70/XRCC6, in agreement with the conjugated peptide PSM frequency (entry 6, Table 1). The tag 7 and photo-NSAIDs labeled elongation factor $1-\alpha$ 1 and rho GDPdissociation inhibitor 2 similarly, although partial competition was still observed in some cases. Cellular thermal shift assay in Jurkat cells showed stabilization of histone H2A by all three NSAIDs and the celecoxib analog 9 relative to the tag 7 by 10-15 °C (Figure 5B). The T_{agg} was estimated by densitometry and fitted to a Boltzmann sigmoid equation. Photonaproxen (4) and photocelecoxib (5) also stabilized histone H2A, while photoindomethacin (6) possessed a destabilizing effect. These data are in line with the PSM frequency in Table 1. Further characterization of these binding sites for biological function will be pursued.

DISCUSSION

The development and application of SIM-PAL to enable characterization of the small molecule interactome and direct binding interactions is described (Figure 1A). SIM-PAL provides several key advantages, including: (1) instant validation of the interaction by virtue of direct observation, (2) measurement of a concentration-dependent range of interactions, including transient and highly specific binding events, and (3) structural information about the binding site and the small molecule structure. These advantages are highlighted by our analysis of three NSAIDs for their binding sites with a single protein (COX-2) and against the whole proteome from Jurkat and K562 cells.



Figure 4. Spectral matches of photo-NSAID binding site hotspots to histone proteins. (A) Precursor pattern distribution (MS1) and database assignment (MS2) for a histone H2A peptide conjugated to each of the photo-NSAIDs. (B) Cluster diagram of histone protein

Figure 4. continued

complexes with at least one observed binding site to a photo-NSAID. High confidence interactions between two clusters made using STRING. (C) Structure of the nucleosome (PDB: 2CV5). Peptides from histone H2A (red) and histone H2B (blue) that were conjugated by all three photo-NSAIDs are highlighted.

We applied SIM-PAL to three NSAIDs, naproxen (1), celecoxib (2), and indomethacin (3), due to their different structures yet similar anti-inflammatory effects. Variation in their off-target toxicity and the affected COX-2-independent anti-inflammatory pathways has been reported with a limited molecular basis for these effects. For example, celecoxib was the result of a medicinal chemistry optimization for COX-2 selectivity relative to COX-1,²² yet additional COX-2independent mechanisms have been subsequently reported.¹⁸ To illuminate the molecular driving force behind NSAID biology, we developed a set of photo-NSAIDs (4-6, Figure 1B) that recapitulated the activity of the parent compounds and demonstrated their direct binding interaction with COX-2 (Figure 2). Due to the dynamic nature of PAL, we identified additional transient interactions with COX-2 in vitro and the broader proteome in cellulo. These data report on access and molecular recognition of the compounds to the proteome and afford a structural basis for further functional analysis or protein degradation strategies.⁴⁵ A systematic study of the kinetics of PAL will provide concrete measurement of the range of transient interactions observable by SIM-PAL.

Application of the photo-NSAIDs 4-6 to stimulated Jurkat and K562 cells at 250 µM revealed over 1000 significantly enriched protein interactions that were more homologous across the three photo-NSAIDs than with the tag 7 or the photoglutarimide 8. Analysis of the photo-NSAID protein interactome revealed entire protein complexes of both known and novel interactions. For example, specific proteins in the NF- κ B pathway are now identified as direct interactors (Figure 3E and Figure S9). Enrichment of several protein complexes also indicates that the photoconjugation event is dynamic and highly specific to the local environment of the small molecule at the time of activation. Novel interactions with the proteasome and the adaptor protein complex 2 were also notable. Observation of these proteins may be in part due to the broader range of interactions captured by photoconjugation as compared to existing target identification strategies. Comparison of photo-NSAID interactions with proteomics profiles derived from fragment-based small molecules¹² revealed a 60% overlap. Thus, photo-NSAIDs possess a protein interaction profile that is specific to the molecular structure and cell type.

The photoconjugation step is highly concentration dependent. Analysis of the photo-NSAIDs at 50 μ M yielded 260 enriched proteins across the three compounds. The high proteomic overlap observed at 50 μ M and 250 μ M implies that the protein interactions enriched at 50 μ M are a more stringent subset of the photo-NSAID interactome. Carrying out the described photoconjugation experiments at functionally relevant concentrations is important for enrichment of binding partners that would be physiologically observed. These results provide insight into the molecular interactions within the proteome generated by the NSAIDs, which may or may not possess direct biological function.

Subsequent evaluation of the conjugated peptides using isotope-targeted MS revealed nearly 200 conjugated peptides,



Figure 5. Validation of photo-NSAID binding interactions identified by SIM-PAL. (A) Competitive displacement of photo-NSAIDs by the parent compound. Jurkat cell lysates conjugated to the indicated compound with or without the parent compound were clicked with the biotin probe **10**, captured on streptavidin–agarose, and probed for the indicated protein. (B) Cellular thermal shift assay performed on Jurkat cells in the presence of the indicated compound, probed for histone H2A. Densitometry measurement of the Western blot data was used to estimate T_{agg} (°C). Data are representative of two independent biological replicates.

representing captured photo-NSAID binding sites. These conjugated peptides mapped to proteins that were enriched in the proteomic data, but not always classified as statistically significant by the label-free quantification method.³⁴ The direct observation of these conjugated peptides provides novel insight into real molecular interactions that are occurring in the whole cell proteome. Although the majority of conjugated peptides were identified by only one of the photo-NSAIDs, 30 peptides were observed conjugated to two or more of the photo-NSAIDs (Table 1, Table S4). Selection of proteins for follow up functional studies is typically based on existing biological relevance, relative abundance, or dose dependence. SIM-PAL reports on both the enriched proteome and directly on the conjugated peptide, which provides a higher resolution picture to select proteins for further study.

We elected to evaluate a specific conjugation event that was observed on an interface between histone H2A and H2B in the nucleosome due to the high frequency of PSMs in our MS data. The two peptides are within 6 Å of each other in a crystal structure of the nucleosome, indicative of a binding site hotspot recognized by the photo-NSAIDs.44 While all three photo-NSAIDs were able to conjugate these peptides, photonaproxen (4) conjugation to histone H2B was observed the most frequently, while photocelecoxib (5) was most frequently observed conjugated to histone H2A. Comparison of these data to fragment-based small molecule interaction sites revealed that a diazirine-tagged coumarin, most structurally similar to photonaproxen (4), likewise conjugated a similar peptide from histone H2AZ.¹² The interaction of photo-NSAIDs with histones was competitively displaced by the parent compound by Western blot (Figure 5A). Although a functional attribution cannot be inferred simply by observation of an interaction, celecoxib (2) and photocelecoxib (5) stabilized histone H2A by cellular thermal shift assay (Figure 5B). Stabilization of protein complexes leading to downstream signaling changes is reminiscent of other immunomodulators, including cyclosporin, FK506, and rapamycin,⁴⁶ and the immunomodulatory drug lenalidomide.^{47,48} The interaction may also represent a more promiscuous binding site hotspot of histones with small molecules that has never previously been observed, analogous to the influence of serum albumin on drug pharmacology.⁴ The photo-NSAIDs reported here are validated probes that may be applied to additional biological studies. Further studies may reveal functional relevance and the basis for some of the poorly understood biological effects of NSAIDs.

While photo-NSAIDs interacted with and stabilized COX-2 in cellulo, COX-2 was only observed in PSMs at a FDR > 1% that were thus filtered out from the proteomics data set (Table S2). Furthermore, conjugated peptides from only a fraction of the enriched proteins were identified. These detection differences reflect the increased challenge in seeking to perform site-specific identification of a single conjugated peptide as opposed to protein identification that may derive from multiple peptides from the same protein. Deeper analysis of the NSAID interactome may be obtained by increasing protein inputs, increasing chromatographic separation, or cellular fractionation methods to deplete abundant proteins. Alternatively, application of additional MS fragmentation methods, use of a second protease (e.g., chymotrypsin), or alternative spectral assignment method would provide additional conjugated peptide assignments.

SIM-PAL represents the culmination of advances in chemical enrichment strategies coupled to MS technology and a computational pattern searching algorithm to lay the groundwork for rapid progress in direct structural characterization of small molecule binding sites within the whole proteome. Recent work in profiling small molecule modification sites has begun to drastically expand the number of interactions measured from the complex proteome and enable a deeper understanding of the molecular underpinnings of polypharmacology. SIM-PAL revealed the global interaction map for the three NSAIDs profiled and is readily translated to other clinically relevant agents. For example, the immunomodulatory drugs have widely established pluripotent activity, and the mechanism of action is only partly understood.³ Alternatively, metformin is widely used to treat diabetes with little understanding of the underlying mechanism.⁵⁰ SIM-PAL is poised for broad application to bioactive small molecules for identification of proteomic interactions and binding site hotspots using an unbiased whole cell assay.

ASSOCIATED CONTENT

S Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.7b11639.

Supplementary figures and experimental details. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the data set identifier PXD007094 (PDF)

Supplementary tables (XLSX)

AUTHOR INFORMATION

Corresponding Author

*cwoo@chemistry.harvard.edu

ORCID [©]

Christina M. Woo: 0000-0001-8687-9105

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Dr. Bogdan Budnik (Harvard Proteomics Resource Laboratory) for acquiring MS data. Financial support from the Burroughs Wellcome Fund (C.M.W.), the Uehara Memorial Foundation Postdoctoral Fellowship (Y.A.), Murata Overseas Scholarship Foundation (Y.A.), and Harvard University is gratefully acknowledged.

REFERENCES

(1) Lounkine, E.; Keiser, M. J.; Whitebread, S.; Mikhailov, D.; Hamon, J.; Jenkins, J. L.; Lavan, P.; Weber, E.; Doak, A. K.; Cote, S.; Shoichet, B. K.; Urban, L. *Nature* **2012**, *486*, 361.

(2) Jana, N. R. Cell. Mol. Life Sci. 2008, 65, 1295.

(3) Zhu, Y. X.; Kortuem, K. M.; Stewart, A. K. Leuk. Lymphoma 2013, 54, 683.

- (4) Pasternak, G. W. Neuropharmacology 2014, 76, 198.
- (5) Hopkins, A. L. Nat. Chem. Biol. 2008, 4, 682.

(6) Keiser, M. J.; Setola, V.; Irwin, J. J.; Laggner, C.; Abbas, A. I.;

Hufeisen, S. J.; Jensen, N. H.; Kuijer, M. B.; Matos, R. C.; Tran, T. B.; Whaley, R.; Glennon, R. A.; Hert, J.; Thomas, K. L. H.; Edwards, D.

D.; Shoichet, B. K.; Roth, B. L. Nature 2009, 462, 175.

(7) Paolini, G. V.; Shapland, R. H. B.; van Hoorn, W. P.; Mason, J. S.; Hopkins, A. L. *Nat. Biotechnol.* **2006**, *24*, 805.

(8) Schenone, M.; Dancik, V.; Wagner, B. K.; Clemons, P. A. Nat. Chem. Biol. 2013, 9, 232.

(9) Gertsik, N.; am Ende, C. W.; Geoghegan, K. F.; Nguyen, C.; Mukherjee, P.; Mente, S.; Seneviratne, U.; Johnson, D. S.; Li, Y. M. *Cell Chem. Biol.* **2017**, *24*, 3. (10) Mackinnon, A. L.; Taunton, J. Current Protocols in Chemical Biology; Wiley: Hoboken, NJ, 2009; Vol. 1, pp55-73.10.1002/9780470559277.ch090167

(11) Woo, C. M.; Iavarone, A. T.; Spiciarich, D. R.; Palaniappan, K. K.; Bertozzi, C. R. *Nat. Methods* **2015**, *12*, 561.

(12) Parker, C. G.; Galmozzi, A.; Wang, Y.; Correia, B. E.; Sasaki, K.; Joslyn, C. M.; Kim, A. S.; Cavallaro, C. L.; Lawrence, R. M.; Johnson, S. R.; Narvaiza, I.; Saez, E.; Cravatt, B. F. *Cell* **2017**, *168*, 527.

(13) Zha, S.; Yegnasubramanian, V.; Nelson, W. G.; Isaacs, W. B.; De Marzo, A. M. *Cancer Lett.* **2004**, 215, 1.

(14) Stewart, W. F.; Kawas, C.; Corrada, M.; Metter, E. J. *Neurology* **1997**, 48, 626.

(15) Vane, J. R. Nat. New Biol. 1971, 231, 232.

(16) Tegeder, I.; Pfeilschifter, J.; Geisslinger, G. FASEB J. 2001, 15, 2057.

(17) Hanif, R.; Pittas, A.; Feng, Y.; Koutsos, M. I.; Qiao, L.; Staiano-Coico, L.; Shiff, S. I.; Rigas, B. Biochem. Pharmacol. **1996**, 52, 237.

(18) Jones, M. K.; Wang, H. T.; Peskar, B. M.; Levin, E.; Itani, R. M.; Sarfeh, I. J.; Tarnawski, A. S. *Nat. Med.* **1999**, *5*, 1418.

(19) Yin, M.-J.; Yamamoto, Y.; Gaynor, R. B. Nature 1998, 396, 77.
(20) Smith, C. E.; Soti, S.; Jones, T. A.; Nakagawa, A.; Xue, D.; Yin, H. Cell Chem. Biol. 2017, 24, 281.

(21) Mitchell, J. A.; Akarasereenont, P.; Thiemermann, C.; Flower, R. J.; Vane, J. R. Proc. Natl. Acad. Sci. U. S. A. **1993**, 90, 11693.

(22) Flower, R. J. Nat. Rev. Drug Discovery 2003, 2, 179.

(23) Schnitzer, T. J.; Burmester, G. R.; Mysler, E.; Hochberg, M. C.; Doherty, M.; Ehrsam, E.; Gitton, X.; Krammer, G.; Mellein, B.; Matchaba, P.; Gimona, A.; Hawkey, C. J.; Grp, T. S. *Lancet* **2004**, *364*, 665.

(24) McGettigan, P.; Henry, D. JAMA-J. Am. Med. Assoc 2006, 296, 1633.

(25) Kalgutkar, A. S.; Crews, B. C.; Rowlinson, S. W.; Marnett, A. B.; Kozak, K. R.; Remmel, R. P.; Marnett, L. J. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 925.

(26) Chandna, N.; Kumar, S.; Kaushik, P.; Kaushik, D.; Roy, S. K.; Gupta, G. K.; Jachak, S. M.; Kapoor, J. K.; Sharma, P. K. *Bioorg. Med. Chem.* **2013**, *21*, 4581.

(27) Kurumbail, R. G.; Stevens, A. M.; Gierse, J. K.; McDonald, J. J.; Stegeman, R. A.; Pak, J. Y.; Gildehaus, D.; iyashiro, J. M.; Penning, T.

D.; Seibert, K.; Isakson, P. C.; Stallings, W. C. Nature **1996**, 384, 644. (28) Orlando, B. J.; Malkowski, M. G. Acta Crystallogr., Sect. F: Struct. Biol. Commun. **2016**, 72, 772.

(29) Li, Z.; Hao, P.; Li, L.; Tan, C. Y. J.; Cheng, X.; Chen, G. Y. J.; Sze, S. K.; Shen, H.-M.; Yao, S. Q. Angew. Chem., Int. Ed. 2013, 52, 8551.

(30) Seibert, K.; Zhang, Y.; Leahy, K.; Hauser, S.; Masferrer, J.; Perkins, W.; Lee, L.; Isakson, P. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 12013.

(31) Schneidman-Duhovny, D.; Inbar, Y.; Nussinov, R.; Wolfson, H. J. Nucleic Acids Res. **2005**, 33, W363.

(32) Iñiguez, M. A.; Punzón, C.; Fresno, M. J. Immunol 1999, 163, 111.

(33) Jafari, R.; Almqvist, H.; Axelsson, H.; Ignatushchenko, M.; Lundbäck, T.; Nordlund, P.; Molina, D. M. *Nat. Protoc.* **2014**, *9*, 2100.

(34) Zybailov, B.; Mosley, A. L.; Sardiu, M. E.; Coleman, M. K.; Florens, L.; Washburn, M. P. J. Proteome Res. 2006, 5, 2339.

(35) Choi, H.; Fermin, D.; Nesvizhskii, A. I. Mol. Cell. Proteomics 2008, 7, 2373.

(36) Zhang, Y.; Wen, Z.; Washburn, M. P.; Florens, L. Anal. Chem. 2015, 87, 4749.

(37) Zhou, D.; Zhang, Q.; Lu, W.; Xia, Q.; Wei, S. J. Clin. Pharmacol. 1998, 38, 625.

(38) Dharmapuri, G.; Doneti, R.; Philip, G. H.; Kalle, A. M. Leuk. Res. 2015, 39, 696.

(39) Thul, P. J.; Åkesson, L.; Wiking, M.; Mahdessian, D.; Geladaki, A.; Ait Blal, H.; Alm, T.; Asplund, A.; Björk, L.; Breckels, L. M.; Bäckström, A.; Danielsson, F.; Fagerberg, L.; Fall, J.; Gatto, L.; Gnann, C.; Hober, S.; Hjelmare, M.; Johansson, F.; Lee, S.; Lindskog, C.; Mulder, J.; Mulvey, C. M.; Nilsson, P.; Oksvold, P.; Rockberg, J.; Schutten, R.; Schwenk, J. M.; Sivertsson, Å.; Sjöstedt, E.; Skogs, M.; Stadler, C.; Sullivan, D. P.; Tegel, H.; Winsnes, C.; Zhang, C.; Zwahlen, M.; Mardinoglu, A.; Pontén, F.; von Feilitzen, K.; Lilley, K. S.; Uhlén, M.; Lundberg, E. *Science* **2017**, *356*, eaal3321.

(40) Ruepp, A.; Waegele, B.; Lechner, M.; Brauner, B.; Dunger-Kaltenbach, I.; Fobo, G.; Frishman, G.; Montrone, C.; Mewes, H. W. *Nucleic Acids Res.* **2010**, *38*, D497.

(41) Palaniappan, K. K.; Pitcher, A. A.; Smart, B. P.; Spiciarich, D. R.; Iavarone, A. T.; Bertozzi, C. R. *ACS Chem. Biol.* **2011**, *6*, 829.

(42) Woo, C. M.; Felix, A.; Byrd, W. E.; Zuegel, D. K.; Ishihara, M.; Azadi, P.; Iavarone, A. T.; Pitteri, S. J.; Bertozzi, C. R. J. Proteome Res. 2017, 16, 1706.

(43) Bern, M.; Kil, Y. J.; Becker, C. Current Protocols in Bioinformatics; Wiley: Hoboken, NJ, 2012; Chpt 13, pp 13.20.1–13.20.14

(44) Tsunaka, Y.; Kajimura, N.; Tate, S.-i.; Morikawa, K. *Nucleic Acids Res.* **2005**, 33, 3424.

(45) Lu, J.; Qian, Y.; Altieri, M.; Dong, H.; Wang, J.; Raina, K.; Hines, J.; Winkler, J. D.; Crew, A. P.; Coleman, K.; Crews, C. M. *Chem. Biol.* **2015**, *22*, 755.

(46) Schreiber, S. Science 1991, 251, 283.

(47) Krönke, J.; Udeshi, N. D.; Narla, A.; Grauman, P.; Hurst, S. N.; McConkey, M.; Svinkina, T.; Heckl, D.; Comer, E.; Li, X.; Ciarlo, C.; Hartman, E.; Munshi, N.; Schenone, M.; Schreiber, S. L.; Carr, S. A.; Ebert, B. L. *Science* **2014**, *343*, 301.

(48) Lu, G.; Middleton, R. E.; Sun, H.; Naniong, M.; Ott, C. J.; Mitsiades, C. S.; Wong, K.-K.; Bradner, J. E.; Kaelin, W. G. *Science* **2014**, 343, 305.

(49) Kragh-Hansen, U. Pharmacol Rev. 1981, 33, 17.

(50) Pernicova, I.; Korbonits, M. Nat. Rev. Endocrinol. 2014, 10, 143.