

Donor-Acceptor Pyridinium Salts for Photo-Induced Electron Transfer Driven Modification of Tryptophan in Peptides, Proteins, and Proteomes using Visible Light

Caleb R. Hoopes^{†a}, Francisco J. Garcia^{†b}, Akash M. Sarkar^a, Nicholas J. Kuehl^a, David T. Barkan^b, Nicole L. Collins^a, Chien-Hsiang Hsu^b, Michael D. Jones^b, Markus Schirle^b and Michael T. Taylor^{a*}

^aUniversity of Wyoming, Department of Chemistry, Laramie, WY 82071

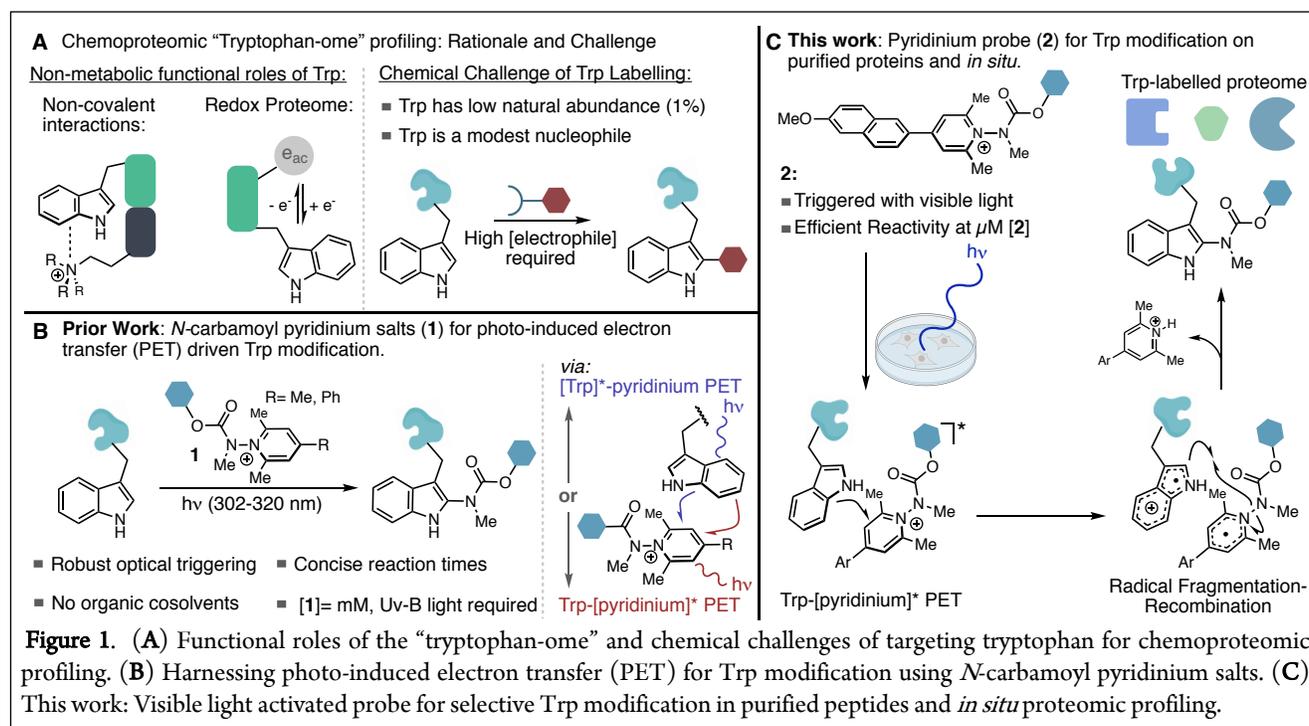
^bNovartis Institutes for Biomedical Research, 181 Massachusetts Ave, Cambridge, MA, 02138

Supporting Information Placeholder

ABSTRACT: Tryptophan (Trp) plays a variety of critical functional roles in protein biochemistry however, owing to its low natural frequency and poor nucleophilicity, the design of effective methods for both single protein bioconjugation at Trp as well as for *in situ* chemoproteomic profiling remains a challenge. Here, we report a method for covalent Trp modification that is suitable for both scenarios by invoking photo-induced electron transfer (PET) as a means of driving efficient reactivity. We have engineered biaryl *N*-carbamoyl pyridinium salts that possess a donor-acceptor relationship enabling optical triggering with visible light whilst simultaneously attenuating the probe's photo-oxidation potential in order to prevent photodegradation. This probe was assayed against a small bank of eight peptides and proteins, where it was found that micromolar concentrations of probe and short irradiation times (10-60 min) with violet light enabled efficient reactivity towards surface exposed Trp residues. The carbamate transferring group can be used to transfer useful functional groups to proteins including affinity tags and click handles. DFT calculations and other mechanistic analyses reveal correlations between excit-

ed state lifetimes, relative fluorescent quantum yields, and chemical reactivity. Biotinylated and azide-functionalized pyridinium salts were used for Trp profiling in HEK293T lysates and *in situ* in HEK293T cells using 450 nm LED irradiation. Peptide level enrichment from live cell labelling experiments identified 290 Trp modifications, with an 82% selectivity for Trp modification over other π -amino acids; demonstrating the ability of this method to identify and quantify reactive Trp residues from live cells.

The union of covalent protein modification chemistry with proteomic analysis has enhanced the molecular precision of drug discovery by enabling the target identification of small molecule therapeutics through activity-based protein profiling, the precision design of covalent inhibitors, and through the characterization of transient and dynamic aspects of a proteome¹. Recently, site-selective protein modification technologies have been adapted for use in broad spectrum chemoproteomic interrogation of a given reactive moiety



(typically, an amino acid side-chain)². This approach enables the characterization of reactive hotspots surrounding the moiety of interest. This information can then lead to the discovery of new sites of biological significance and, ultimately, the discovery of new druggable sites. Many chemoproteomic profiling strategies make use of either direct or *in-situ* generated electrophiles which target either strongly nucleophilic species, such as cysteine^{2a-c}, or higher abundance residues of moderate to weak nucleophilicity such as lysine^{2d-f}, serine/threonine^{2f}, tyrosine^{2g,h}, and aspartate/ glutamate^{2i,h}. An elegant example from Chang and Toste targeted weakly nucleophilic, low abundance methionine residues in lysates^{2k}.

However, electrophilic approaches directed towards other functionally important residues, such as tryptophan (Trp), remain a persistent challenge. The challenge of chemoproteomic profiling of Trp stems from a combination of Trp's very low natural abundance and the modest nucleophilicity of its indole sidechain³. This combination necessitates that any Trp-selective electrophile would have to possess both sufficient kinetic competency and selectivity to label Trp (ideally *in situ*) in a complex proteome at low concentrations of reagent. As such, Trp has been the subject of many innovative labelling approaches, including through the use of sulfonyl chlorides^{4a,b}, metalcarbenoids^{4c,d}, nitroxides^{4e}, transition metal catalyzed processes^{4f}, oxidative strategies^{4g} and photoredox catalysis^{4h}. Whilst comprising an elegant body of work, these methods would require significant adaptation for use as chemical probes for Trp residues *in situ*. In 2003, Nishimura used a 2-nitrobenzenesulfonyl chloride isotopologue for chemoselective Trp modification in rat serum; resulting in the identification of three tryptic peptides harboring a Trp modification^{4b}. In 2007, Leitner reported an approach that enabled enrichment of Trp-containing peptides through reversible condensation of malondialdehyde onto the Trp-indolic N-H followed by enrichment on hydrazide-derivatized beads; an approach that enabled the identification of nine proteins from yeast lysates⁴ⁱ. Whilst demonstrating the premise, these strategies required either pre-denaturation of proteins and labelling performance under acidic conditions or the use of unselective chemistries; precluding the use of these methods in more complex environments.

Despite these challenges, chemoproteomic profiling of Trp presents a unique chemical challenge and opportunity (Figure 1A). Trp is widely dispersed throughout the proteome and is also relatively evenly distributed between surface exposed and buried positions⁵. Trp possesses the most electron-rich π -system of the naturally occurring amino acids; a property that enables Trp to engage in electrostatically-driven non-covalent interactions such as π - π , cation- π , and X-H- π interactions, as well as through H-bonding to neighboring functionality *via* the indolic N-H bond⁶. As a result, Trp can be found enriched at centers of biochemical signifi-

cance such as protein-protein interfaces, protein-lipid interfaces (2.9% abundance in membrane proteins^{7d}), as well as serving to maintain protein structural integrity through intraprotein non-covalent interactions⁷. Owing to Trp's electron-rich π -system, further enrichment of Trp can be observed in the redox proteome⁸.

We recently reported a method for Trp modification that circumvents Trp's modest nucleophilicity by instead utilizing Trp's propensity to engage in photo-induced electron transfer (PET)⁹. Specifically, we utilized an *N*-carbamoylpyridinium salt (**1**) that, upon irradiation with 302 nm light, couples PET with selective bond formation at Trp through a radical fragmentation/recombination process (Figure 1B). The reaction displayed good selectivity for Trp, as well as short reaction times, with the pyridinium reagent displaying high degrees of both thermal- and photo-stability. Following our report, Melchiorre^{10a}, Beier^{10b}, Paixão^{10c}, and Chiang^{10d} reported photochemical Trp modifications on small peptides using either Uv-A or visible light and organic solvents. Davis and Gouveurner also reported a three-step sequence to access functionalized Trp constructs using a difluoromethylation process featuring a ruthenium photocatalyst in a strongly reducing environment as part of the 24 hour long process^{10e}. Given the observed photo- and thermal stability of the **1** scaffold, coupled with its compatibility with pure aqueous environments and biological antioxidants such as glutathione (GSH), we considered that this scaffold could be readily adapted for use in chemoproteomic profiling. Additionally, the spatiotemporal control afforded by optical triggering enables control over aspects such as probe incubation time, allowing for sufficient time for cellular penetration and compartmentalization. Thus, chemoproteomic profiling using pyridinium salts to invoke the PET mechanistic manifold presents the opportunity to enrich an underexplored sub-section of the proteome that compliments current electrophilic strategies through enabling the discovery of new functionally critical Trp residues as well as sites of non-covalent interactions and biomolecular interfaces; in turn enabling their consideration as ligandable sites for drug discovery¹¹.

Results and Discussion

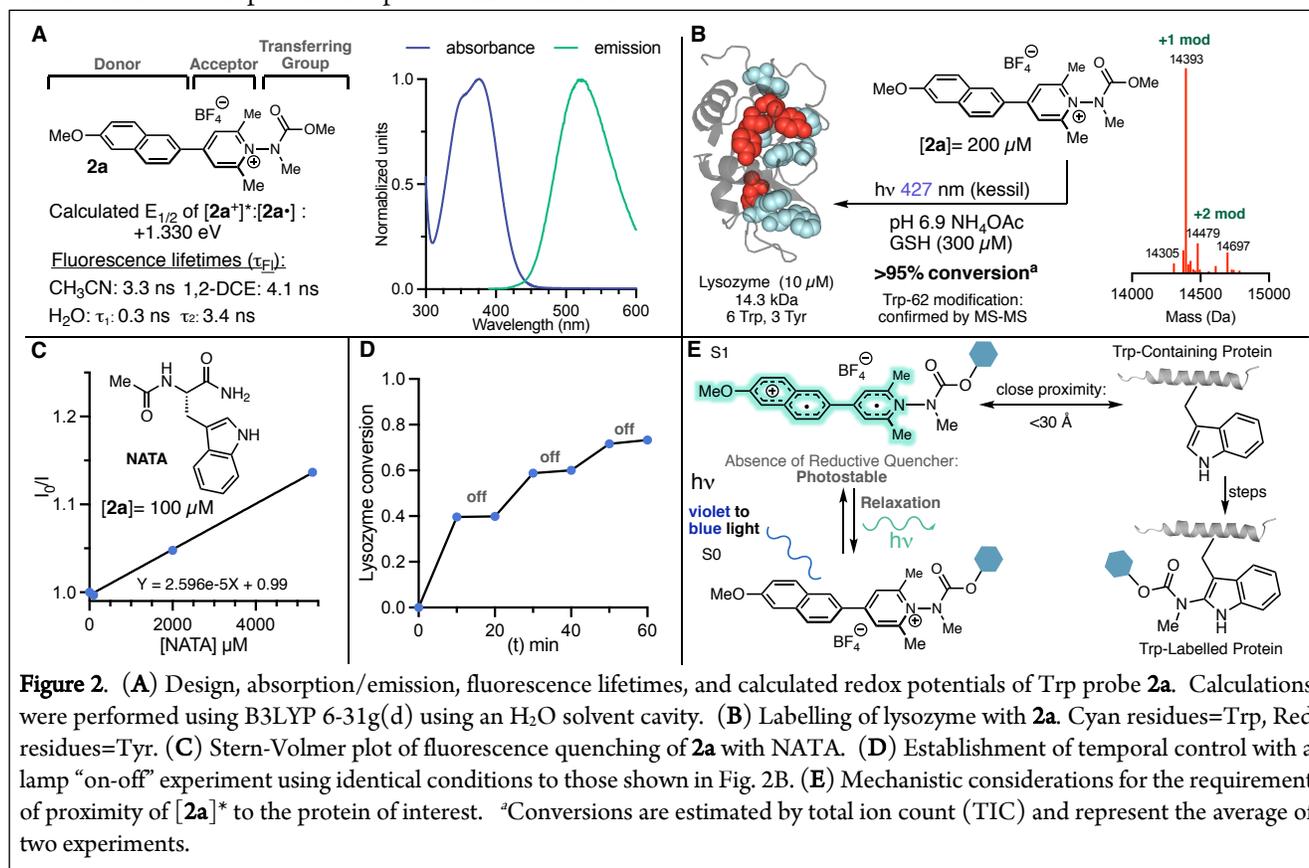
In seeking to develop a robust chemoproteomic profiling workflow using **1**, we noted the practical challenges of using filtered (to avoid wavelengths of <300 nm) Uv-B light as an irradiation source, as well as the potential for Uv-B light to induce photodegradation of labile proteins and unintended cellular stress¹². We have previously demonstrated that the mechanism of Trp labelling with **1** can be controlled in a reagent- and wavelength-dependent fashion by either invoking direct Trp photoexcitation ($[\text{Trp}]^* \rightarrow \mathbf{1}$, Trp acts as a photoreductant) or by tuning the structure of **1** to enable exclusive photoexcitation of **1** ($\text{Trp} \rightarrow [\mathbf{1}]^*$, Trp acts as a reductive quencher); with the latter approach enabling modification of photolabile proteins. Using this observation, we rationalized

that it should be possible to extend the π -conjugation of **1** to enable direct photoexcitation of **1** using the visible spectrum. Concurrently, we recognized that structurally analogous cationic-aromatic photosensitizers, such as triaryl pyrylium salts, acridinium salts, and flavins are potent photo-oxidants that possess excited state redox potentials ($S^*/S \cdot E_{1/2}$) in excess of +1.8 eV¹³; which could possibly lead to a loss of chemoselectivity or photo-induced degradation of sensitive biological structure.

These considerations have led to the development of pyridinium scaffold **2**, which possesses an electron-rich 6-methoxynaphthyl ring appended at the 4-position to the cationic pyridinium moiety. This structural manifold creates a donor-acceptor relationship between the two aryl rings that both tempers the photo-oxidation potential of **2** and bathochromically shifts the absorbance of model probe **2a** to beyond 450 nm (Figure 2A); in turn enabling optical activation with simple violet and blue LEDs. Irradiation of a degassed solution of 10 μ M lysozyme in pH 6.9 NH_4OAc buffer, 300 μ M glutathione (GSH) and 200 μ M of **2a** with a 427 nm LED (Kessil PR-160, 50W) in a Hepatochem Photoredox-box photoreactor for 60 minutes resulted in the efficient labelling of lysozyme in >95% conversion and only low levels of glutathionylation observed (Figure 2B). MS-MS analysis confirms selectivity for Trp-62, analogous to our prior report^{9a}. Compared to our first generation probe designs (**1**), **2a** demonstrates a marked enhancement in performance; efficiently labeling proteins at an order of magnitude lower concentration of both protein and probe whilst simultane-

ously affording optical triggering with visible light.

To better understand this enhanced kinetic profile, we studied the photophysical properties of **2a**. Temporal control experiments verify the requirement of light to achieve labelling, and suppress chain-transfer driven mechanisms (Figure 2D). Stern-Volmer plot of fluorescence quenching of **2a** with small molecule Trp analog *N*-acetyltryptophamide (NATA) shows a measurable linear correlation, suggesting dynamic quenching of $[\mathbf{2a}]^*$ with Trp (Figure 2C). **2a** displays solvatochromatic properties that are associated with significant charge transfer in the photoexcited state, with TDDFT calculations supporting this excitation mode through a naphthyl- \rightarrow pyridinium charge transfer upon excitation from S_0 to S_1 , with the biaryl motif adopting a coplanar orientation in S_1 (Figure 2E). Owing to the irreversible peak reduction potential typical of the SET/fragmentation reactivity manifold of **2a** and related scaffolds¹⁴, we computationally estimated the redox potential of **2a** using the method of Nicewicz¹⁵. This approach estimates the $[\mathbf{2a}^+]^*/\mathbf{2a}^{\cdot}$ excited state redox potential to be +1.330 eV; nearly 1 eV less oxidizing than tri-arylpopyrylium catalysts^{13a}. Fluorescence lifetime measurements in 1,2-dichloroethane and acetonitrile both show one-phase exponential decays that correlate to $\tau_1 = 4.1$ ns and $\tau_1 = 3.3$ ns respectively; suggesting a general decrease in lifetime as solvent polarity increases. However, two lifetimes are observed in water ($\tau_1 = 0.3$ ns and $\tau_2 = 3.4$ ns); with the second lifetime being of very weak intensity. Concurrently, the fluorescence intensity of **2a** decreases with



increasing solvent polarity; with the Φ_{rel} = 24.8 between acetonitrile and water (see supplementary figure S14). These data suggest that (1) the excited state of **2a** is generally suppressed in aqueous systems compared to organic solvents and (2) **2a*** cannot diffuse beyond 30 Å. When considered together, these observations lead us to propose that the observed selectivity for Trp modification by **2a** is likely driven by a synergistic combination of a short-lived excited state that lacks the persistence to enable kinetically slower processes and protein-**2a** pre-complexation *via* hydrophobic effects, as well as a kinetic preference for Trp over other redox-active residues (Figure 2E).

Next, we sought to explore the scope of the labelling process. We therefore assembled a small battery of Trp-containing peptides and proteins in which the Trp residues are located in an array of differing environments (Figure 3A).

Smaller peptides, such as Octreotide and Leuprolide are each labelled efficiently under violet LED irradiation conditions. Exenatide•OAc, a 4.2 kDa polypeptide, possesses a single turn of the peptide backbone displays slightly diminished, but still efficient conversion (70%). Thioredoxin (Trx), an 11 kDa protein essential to disulfide formation, possesses two Trp residues, two Tyr residues and, in its reduced form, two Cys residues. Irradiation of Trx proceeded smoothly to afford Trp labelling at Trp-31. During the course of the labelling process, we noted disulfide bond formation. We next explored the labelling of β 2-microglobulin (B2M), an 11.8 kDa structural component of the class I major histocompatibility complex, which features two Trp residues, one that is fully solvent-exposed and one that is largely buried, as well as 6 surface exposed Tyr residues. Irradiation of B2M with **2a** for 30 minutes led to almost exclusive mono-labelling at sol-

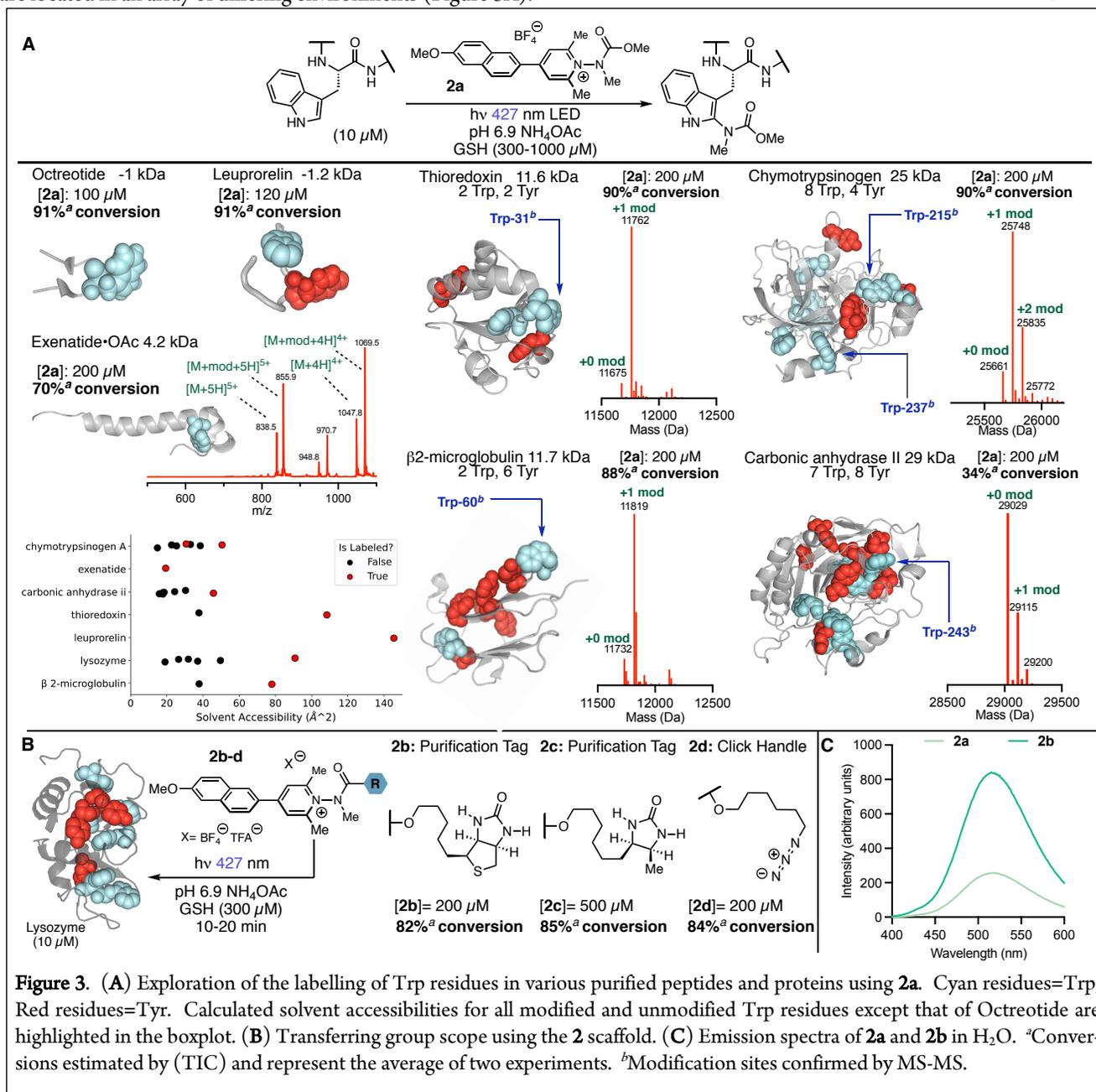


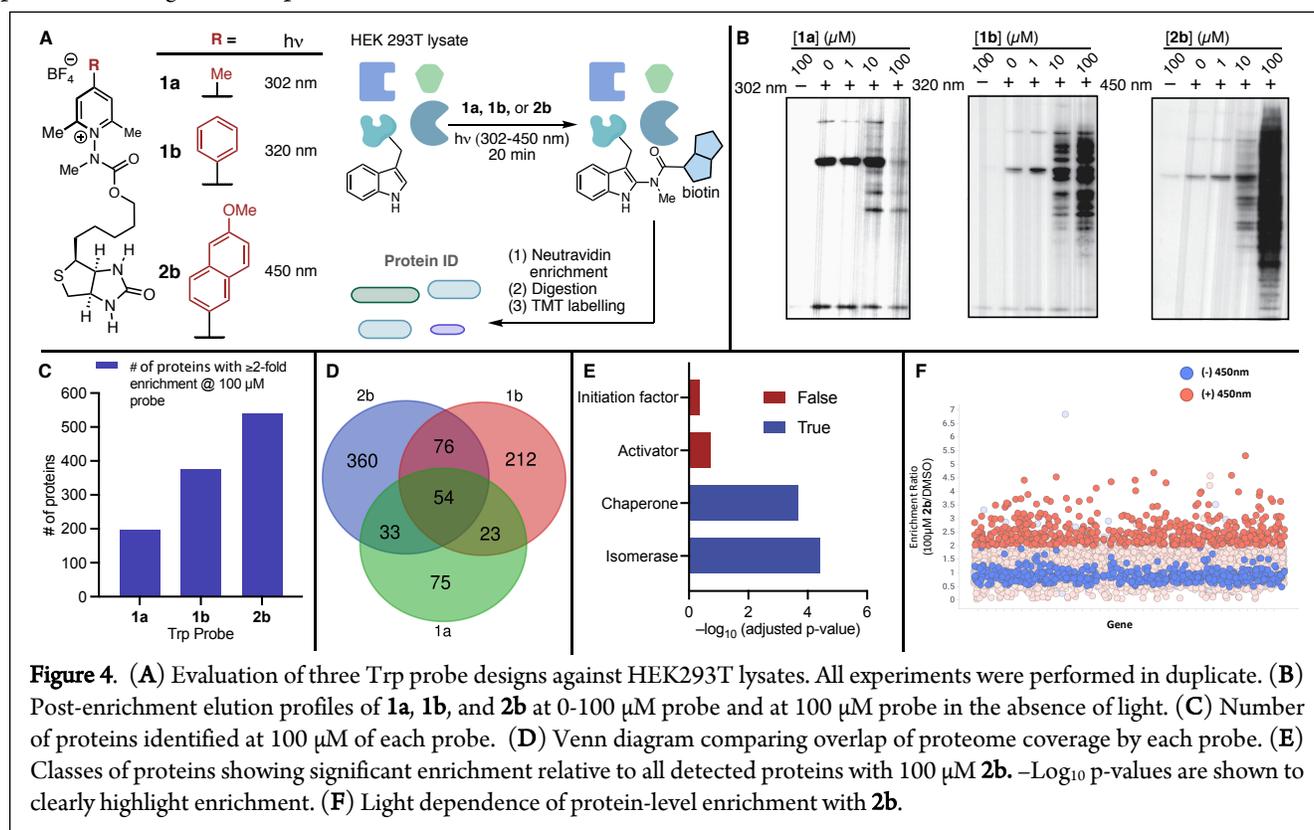
Figure 3. (A) Exploration of the labelling of Trp residues in various purified peptides and proteins using **2a**. Cyan residues=Trp, Red residues=Tyr. Calculated solvent accessibilities for all modified and unmodified Trp residues except that of Octreotide are highlighted in the boxplot. (B) Transferring group scope using the **2** scaffold. (C) Emission spectra of **2a** and **2b** in H_2O . ^aConversions estimated by (TIC) and represent the average of two experiments. ^bModification sites confirmed by MS-MS.

vent exposed Trp-60 in 88% conversion. Chymotrypsinogen A, a 25 kDa protein, possesses 8 Trp residues and 4 Tyr residues, efficiently labels in 90% conversion, with a 21:8:1 ratio of mono:di:tri labelling. Digestion/MS-MS elucidated Trp-215 and Trp-237 to be the detectable locations of labelling in this protein. Finally, we investigated the labelling of bovine carbonic anhydrase II (CAII). CAII possesses 7 Trp residues, most of which are either completely buried or have very limited solvent exposure. As a result, CAII displayed much lower levels of conversion (34%), with Trp-243 being the identifiable site of modification. Taken together, these results establish the selectivity of **2a** for Trp over an array of environments. The solvent accessibility of each Trp residue, excluding Octreotide, was calculated and indicates that **2a** selects for the most solvent accessible residue when multiple choices are present. In examples where there is a single Trp residue that is poorly exposed, such as Exenatide, modification can be still be observed, albeit with reduced conversion. It also suggests that **2a** is sensitive to local steric environments, not just with respect to surface exposure, but also to adequate surface exposure of the indolic C2 position as this position is the preferred site of carbamylation. We also note that these labelling reactions generally demonstrate clean reaction profiles, indicating that **2a** does not cause significant levels of photodegradation.

We then turned our attention to the use of **2** to transfer useful functionality to Trp residues (Figure 3B). In that regard, we found that affinity tags such as biotin (**2b**) are efficiently transferred to lysozyme in 82% conversion. Notably, protein labelling with **2b** proceeded at a much faster rate

than with **2a**; with complete reactivity observed in less than 10 minutes, as opposed to 60 minutes with **2a**. Fluorescence lifetime experiments show that **2b** has a shorter τ_2 ($\tau_1=0.5$ ns and $\tau_2=2.8$ ns) than **2a** whilst comparison of emission spectra reveal that **2b** has a higher fluorescence quantum yield (Φ_{Fl}) than **2a** (relative Φ_{Fl} **2b**:**2a**= 3.27) (Figure 3C); suggesting a correlation between Φ_{Fl} and kinetics. Desthiobiotinylation with **2c** also proceeded with an enhanced rate, but required a larger excess of probe ($[2c]=500$ μ M vs. $[2b]=200$ μ M) in order to observe comparable labelling efficiency. The difference in reactivity between these two structurally similar probes can, in part, be attributed to subtle differences in water solubility, but the origins of this difference require further investigation. Azide probe **2d** also showed enhanced rates compared to **2a**, and was smoothly transferred to lysozyme in 84% conversion with just 20 minutes of irradiation time.

Having demonstrated efficient reactivity and selectivity towards Trp modification in purified peptide/protein systems, we next sought to assess the capabilities of **2** in chemoproteomic profiling. We assembled three biotinylated Trp probes, **2b**, as well as probes **1a** and **1b**, which featured our previously reported pyridinium designs against proteome profiling of HEK293T lysates. Lysates (5 mg/ml) were incubated with the desired probe at varying concentrations (0-100 μ M) for 60 minutes and subsequently irradiated for 20 minutes with the following light sources: unfiltered 302 nm light for **1a**, filtered 320 nm light for **1b**, and 450 nm light for **2b** (Figure 4A). Biotinylated proteins were then enriched *via*



neutravidin and washed to remove nonspecifically bound proteins. The enriched proteins were eluted and subsequently digested in solution followed by Tandem Mass Tag (TMT) labelling, and protein identification and quantitation by LC/MS-MS. Western blot analysis using streptavidin-HRP of elution profiles for probes **1a**, **1b** and **2b** before digestion indicated that each probe displays concentration and optically dependent proteome labelling profiles. Notably, we observed a loss of signal with trimethyl probe **1a** at the highest assayed conditions, 100 μM (Figure 4B). It is possible that this loss of signal correlates to photodegradation associated with the generation of an excess of high energy photo-excited states of **1a**, which could originate from the lower wavelength light emitted by the unfiltered 302 nm light source. Both probes **1b** and **2b** showed expected dose-response profiles. Following enrichment and digestion, LC/MS-MS analysis led to the identification (from ≥ 2 unique peptides) of 185 proteins with **1a**, 365 proteins with **1b**, and 523 proteins with **2b** that showed ≥ 2 -fold enrichment at 100 μM concentration vs. DMSO (Figure 4C). A comparison of proteome coverage of each probe reveals **2b** has the highest percentage (69%) of unique enriched proteins (Figure 4D). In order to understand what types of protein functions are enriched with **2b**, an enrichment analysis of proteins displaying ≥ 2 -fold enrichment with **2b** versus all proteins detected in the experiment was performed (Figure 4E). This analysis indicated that chaperone and isomerase proteins an especially high degree of enrichment against the background. A complete enrichment analysis and list of enriched proteins organized by class are available in supplementary information figures S22 and S23). Analysis of enrichment profiles for identified proteins with probe **2b** in both the presence and absence of photoirradiation clearly illustrates both photo-induced reactivity of this probe as well as its thermal inertness in the absence of light (Figure 4F). During the completion of our studies, a pre-print from Hacker used an alkynylated derivative of **1a** as part of an elegant study of profiling *S. aureus* and MDA-MB-231 cell lysates with multiple bioconjugation probes¹⁶.

Next, we sought to label and identify Trp modifications through intracellular PET labelling in cultured HEK293T cells. Owing to the superior performance of clickable azide probe **2d** over desthiobiotin probe **2c** in labelling purified proteins, we opted for use of **2d** in these studies. Thus, live HEK293T cell cultures were incubated with 1-100 μM concentrations of **2d** for 60 minutes at 37°C, followed by photoirradiation with 450 nm light for 20 minutes at 4°C. Lysates were then generated, followed by CuAAC to install an acid-cleavable alkyne-DADPS-biotin (DADPS=dialkoxydiphenyl silane) affinity tag¹⁷, digestion, TMT labelling, and recombination of TMT-labelled samples. Enrichment of this mixture with neutravidin, followed by release of captured tryptic peptides with 10% formic acid yielded peptides harboring covalent modifications that were then identified and quantified by LC/MS-MS (Figure 5A). As evidenced in

Figure 5B, **2d** enables light and dose-dependent enrichment of Trp-modified peptides, with measurable 2-fold enrichment at low (10 μM) concentrations of probe. We also noted that **2d** displays robust thermal stability under whole cell culture conditions and the lack of peptide enrichment without 450 nm photoirradiation demonstrates spatiotemporal control. At 100 μM of **2d**, the above workflow enabled the detection of 290 Trp-modified peptide fragments showing dose-response behavior that correspond to 209 unique proteins (blue dots). Selectivity for Trp modification is maintained *in situ* and demonstrates the ability to enrich the tryptophan-ome directly from whole cell culture. Analysis of the chemoselectivity profile of **2d** reveals 221 unique peptides harboring a Trp modification, with minor levels of modifications observed at His, Tyr, and Cys (Figure 5C). When the relative amino acid frequency of these residues in human proteins is taken into account, **2d** displays 82% selectivity for Trp (Figure 5D).

Proteins with detected Trp modifications were analyzed against the UniProt database for subcellular localization profiles (Figure 5E). A majority of identified proteins were found to have at least partial localization within the nucleus and the cytoplasm, followed by the mitochondria and endoplasmic reticulum. In addition to evidencing that **2d** readily crosses the plasma membrane, this analysis also indicates that the net positive charge of **2d** does not restrict Trp modification to mitochondrial proteins. Rather, **2d** is able to penetrate a wide variety of important subcellular regions and can enrich the tryptophan-ome in a wide array of areas. Cross-referencing identified proteins against the Drugbank database revealed that a majority of identified proteins (75%) do not have Drugbank entries (Figure S24). A search of modified Trp residues for residue/region specific annotation in the UniProt database indicated that a large majority (92%) of modified residues have no functional annotation assigned to them (Figure S24). The most common functional annotations found involved Trp residues located in trans-membrane regions of the parent protein (6%), whilst 2% of Trp residues had established function based upon mutagenesis experiments. This very low percentage of residue annotation, as well as low percentage of DrugBank proteins, contrasts to other chemoproteomic profiling workflows that target more highly studied residues, such as lysine^{2d-f} and cysteine^{2a-c}, and suggests potential for this method to enable the discovery of functionally important Trp residues.

Whilst most modified Trp residues have no known annotation, chemoproteomic profiling with **2d** did enable *in situ* modification of functionally critical Trp residues (Figure 5H). For example, chemoproteomic Trp profiling with **2d** enabled *in situ* labelling of Trp290 of the nuclear protein NPM1. NPM1 is a protein involved in a wide array of cellular functions revolving around genomic homeostasis¹⁸. Trp290 is located in the C terminal domain of NPM1, which plays a critical role with respect to colocalization of NPM1 into the

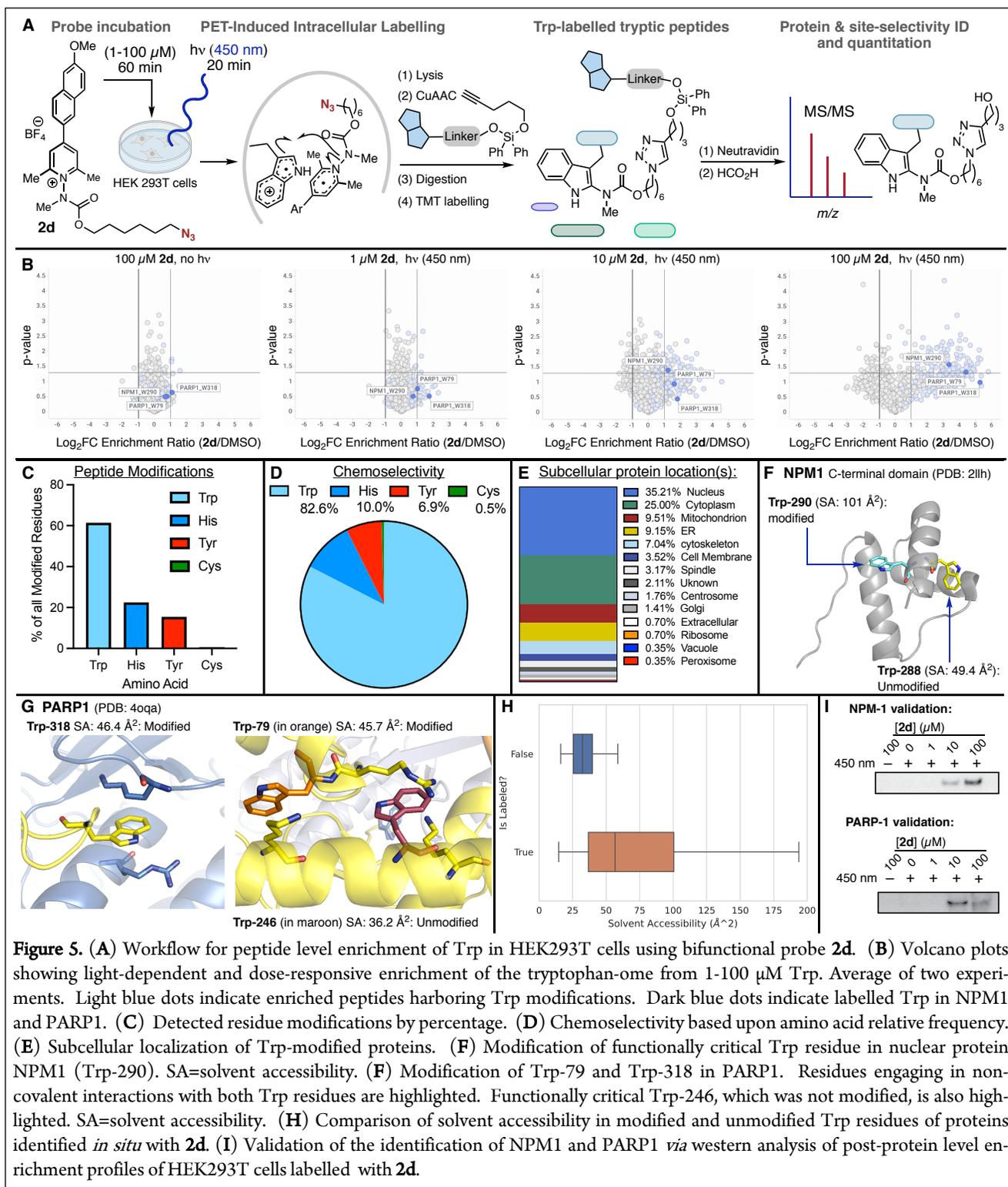


Figure 5. (A) Workflow for peptide level enrichment of Trp in HEK293T cells using bifunctional probe **2d**. (B) Volcano plots showing light-dependent and dose-responsive enrichment of the tryptophan-ome from 1-100 μ M Trp. Average of two experiments. Light blue dots indicate enriched peptides harboring Trp modifications. Dark blue dots indicate labelled Trp in NPM1 and PARP1. (C) Detected residue modifications by percentage. (D) Chemoselectivity based upon amino acid relative frequency. (E) Subcellular localization of Trp-modified proteins. (F) Modification of functionally critical Trp residue in nuclear protein NPM1 (Trp-290). SA=solvent accessibility. (G) Modification of Trp-79 and Trp-318 in PARP1. Residues engaging in non-covalent interactions with both Trp residues are highlighted. Functionally critical Trp-246, which was not modified, is also highlighted. SA=solvent accessibility. (H) Comparison of solvent accessibility in modified and unmodified Trp residues of proteins identified *in situ* with **2d**. (I) Validation of the identification of NPM1 and PARP1 *via* western analysis of post-protein level enrichment profiles of HEK293T cells labelled with **2d**.

nucleolus. Mutation of Trp290 and Trp288, which has a lower calculated solvent accessibility than Trp290 and was not labelled, result in protein misfolding of the C-terminal domain. This in turn prevents NPM1 from both localizing to the nucleus and binding to DNA^{18b}. Mutations to Trp290, as well as Trp288 are commonly found in cases of acute myeloid leukemia^{18c}.

We also observed the *in situ* modification of Trp79 and Trp318 in PARP1; an essential protein in DNA damage detection and repair, gene transcription regulation, and cell death signaling¹⁹. PARP1 is a single chain, multidomain protein consisting of six DNA-damage recognition and catalytic cofactors. Trp79 and Trp318 participate in non-covalent interactions at the interface between these cofactors; with Trp79 participating in a cation- π interaction with a lysine

residue at the Zn1-Zn3 interface whilst Trp318 participates in non-covalent interactions with an arginine of the catalytic domain and a lysine of the WGR domain. Mutation of these residues did not hinder DNA-damage recognition^{19d} but instead resulted in either partial (Trp79) or complete (Trp318) loss of catalytic activity^{19c,d}. PARP1 possesses a third functionally critical Trp residue, Trp264, which was not modified by **2d**. Solvent accessibility calculations reveal that Trp79 and Trp318 have similar solvent accessibilities and are both more solvent accessible than Trp264. This Trp modification selectivity trend correlates well with the selectivity trends observed in our studies from Figure 3 as well as NPM1; namely that **2** selects for Trp residues with the greatest solvent accessibility when multiple Trp residues are present.

We then expanded upon these observations by conducting a survey of all Trp residues found on the proteins identified from our *in situ* studies (for which structural data is available) for solvent accessibility (Figure 5H). Given a maximum theoretical surface exposure of 264 Å², modified Trp residues were found to be 21% accessible (average accessible area = 56 Å²) whilst unmodified Trp residues were only 12% accessible (average accessible area = 32 Å²); a statistically significant difference²⁰. This data further reinforces a preference for **2** to label Trp residues with enhanced surface exposure.

Finally, we validated the identification of NPM1 and PARP1 as enriched targets with **2d** (Figure 5I). HEK293T cells were treated with varying concentrations of **2d** both in the presence and absence of blue light, followed by protein level enrichment. The resultant enriched protein mixtures were then analyzed *via* western analysis and found to contain both proteins when cultures were irradiated in the presence of **2d**. No enrichment of these proteins was found in either the absence of probe or the absence of light.

Conclusions

In summary, we have developed a donor-acceptor pyridinium salt scaffold (**2**) that enables PET-driven Trp modification in proteins using visible light. We demonstrate that **2** labels purified Trp-containing proteins at an order of magnitude more dilute conditions over our previous design whilst maintaining good conversions and selectivity. The carbamate transferring group can be used to install functional handles useful for chemoproteomics and general bioconjugation strategies on to proteins. Photophysical analyses and DFT calculations provide insight into the electronic structure and photochemical behaviors of **2** in the context of observed chemical outcomes; providing mechanistic insight for future probe designs. We also demonstrated that **2** is suited for chemoproteomic profiling; enabling enrichment of the tryptophan-ome from both lysates and live cell culture. We anticipate that **2** can be readily incorporated into chemoproteomic workflows that both probe for ligandable sites that do

not possess traditional nucleophilic residues as well as to explore the basic *in situ* chemical behavior of Trp.

ASSOCIATED CONTENT

Supporting information includes experimental procedures for probe synthesis, protein modification experiments, photophysical studies, chemoproteomic profiling and computational analysis. A second spreadsheet contains chemoproteomic profiling data tables.

The Supporting Information is available free of charge on the ACS Publications website.

AUTHOR INFORMATION

Corresponding Author

*mtaylo45@uwyo.edu

ORCID

Michael T. Taylor: 0000-0002-7655-7222

Nicholas J. Kuehl: 0000-0003-3640-8596

Akash M. Sarkar: 0000-0002-0277-6916

Author Contributions

‡These authors contributed equally.

Funding Sources

No competing financial interests have been declared.

ACKNOWLEDGMENT

We acknowledge the SBC COBRE program (SP20GM121310-02) for support. Additionally, this material is based upon work supported by the Chemistry of Life Processes Program in the National Science Foundation Division of Chemistry under Grant CHE-2048201. CRH acknowledges the Wyoming Research Scholars Program and the Wyoming NASA space grant consortium (NNX15AI08H) for student fellowships. Partial instrumentation support was provided by the NSF grant CHE-1429615. We thank Prof. Caleb Hill for assistance with acquiring cyclic voltammetry data for **2a**.

REFERENCES

- (1)(a) Niphakis, M.J.; Cravatt, B.F. Enzyme Inhibitor Discovery by Activity-Based Protein Profiling. *Annu. Rev. Biochem.* **2014**, *83*, 341–77. (b) Cravatt, B.F.; Wright, A.T.; Kozarich, J.W. Activity-Based Protein Profiling: From Enzyme Chemistry to Proteomic Chemistry. *Annu. Rev. Biochem.* **2008**, *77*, 383–414. (c) Heal, W.P.; Dang, T.H.T.; Tate, E.W. Activity-based probes: discovering new biology and new drug targets. *Chem Soc Rev.* **2011**, *40*, 246–57. (d) Chan, W.C.; Sharifzadeh, S.; Buhlage, S.J.; Marto, J.A. Chemoproteomic methods for covalent drug discovery. *Chem. Soc. Rev.* **2021**, *50*, 8361–8381. (e) Spradlin, J.N.; Zhang, E.; Nomura, D.K. Reimagining Druggability Using Chemoproteomic Platforms. *Acc. Chem. Res.* **2021**, *54*, 1801–1813.
- (2) For reviews summarizing cysteine profiling: (a) Maurais, A.J.; Weerapana, E. Reactive-cysteine profiling for drug discovery. *Curr. Op. Chem. Biol.* **2019**, *50*, 29–36. (b) Backus, K.M.; Applications of Reactive Cysteine Profiling. *Curr. Top. Microbiol. Immunol.* **2019**, *420*, 375–417. (c) Guan,

- I.; Williams, K.; Pan, J.; Liu, X. New Cysteine Covalent Modification Strategies Enable Advancement of Proteome-wide Selectivity of Kinase Modulators. *Asian J. Org. Chem.* **2021**, *10*, 949-963. For examples of lysine profiling: (d) Hacker, S.M.; Backus, K.M.; Lazear, M.R.; Forli, S.; Correia, B.E.; Cravatt, B.F. Global profiling of lysine reactivity and ligandability in the human proteome. *Nature Chemistry* **2017**, *9*, 1181-1190. (e) Shannon, D.A.; Banerjee, R.; Webster, E.R.; Bak, D.W.; Wang, C.; Weerapana, E. Investigating the Proteome Reactivity and Selectivity of Aryl Halides. *J. Am. Chem. Soc.* **2014**, *136*, 3330-3333. For examples targeting *N*, *O*- and *S*-nucleophiles: (f) Ward, C.C.; Kleinman, J.I.; Nomura, D.K. NHS-Esters As Versatile Reactivity-Based Probes for Mapping Proteome-Wide Ligandable Hotspots. *ACS Chem. Biol.* **2017**, *12*, 1478-1483. For examples of Tyrosine profiling: (g) Brulet, J.W.; Borne, A.L.; Yuan, K.; Libby, A.H.; Hsu, K-L. Liganding Functional Tyrosine Sites on Proteins Using Sulfur-Triazole Exchange Chemistry. *J. Am. Chem. Soc.* **2020**, *142*, 18, 8270-8280. (h) Sun, F.; Suttapitugsakul, S.; Wu, R. An Azo Coupling-Based Chemoproteomic Approach to Systematically Profile the Tyrosine Reactivity in the Human Proteome. *Anal. Chem.* **2021**, *93*, 10334-10342.
- For Carboxylic acid profiling: (i) Bach, K.; Beerkens, B.L.H.; Zanon, P.R.A.; Hacker, S.M. Light-Activatable, 2,5-Disubstituted Tetrazoles for the Proteome-wide Profiling of Aspartates and Glutamates in Living Bacteria. *ACS Cent. Sci.* **2020**, *6*, 546-554. (j) Ma, N.; Hu, J.; Zhang, Z-M.; Liu, W.; Huang, M.; Fan, Y.; Yin, X.; Wang, J.; Ding, K.; Ye, W.; Li, Z. 2*H*-Azirine-Based Reagents for Chemoselective Bioconjugation Carboxyl Residues Inside Live Cells. *J. Am. Chem. Soc.* **2020**, *142*, 6051-6059. For Methionine profiling: (k) Lin, S.; Yang, X.; Jia, S.; Weeks, A.M.; Hornsby, M.; Lee, P.S.; Nichiporuk, R.V.; Iavarone, A.T.; Wells, J.A.; Toste, F.D.; Chang, C.J. Redox-based reagents for chemoselective methionine bioconjugation. *Science* **2017**, *355*, 597-602.
- (3) Lakhdar, S.; Westermaier, M.; Terrier, R.; Boubaker, T.; Ofial, A.R.; Mayr, H. Nucleophilic Reactivities of Indoles. *J. Org. Chem.* **2006**, *71*, 24, 9088-9095.
- (4) (a) Scoffone, E.; Fontana, A.; Rocchi, R. Sulfenyl halides as Modifying Reagents for Polypeptides and Protein I. Modification of Tryptophan Residues. *Biochemistry* **1968**, *7*, 971-979. (b) Kuyama, H.; Watanabe, M.; Toda, C.; Ando, E.; Tanaka, K.; Nishimura, O. An approach to quantitative proteome analysis by labeling tryptophan residues. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 1642-1650. (c) Antos, J.M.; McFarland, J.M.; Ivarone, A.T.; Francis, M.B. Chemoselective Tryptophan Labeling with Rhodium Carbenoids at Mild pH. *J. Am. Chem. Soc.* **2009**, *131*, 6301-6308. (d) Ball, Z.T. Designing Enzyme-like Catalysts: A Rhodium(II) Metallopeptide Case Study. *Acc. Chem. Res.* **2013**, *46*, 560-570. (e) Seki, Y.; Ishiyama, T.; Sasaki, D.; Abe, J.; Sohna, Y.; Oisaka, K.; Kanai, M. Transition Metal-Free Tryptophan-Selective Bioconjugation of Proteins. *J. Am. Chem. Soc.* **2016**, *138*, 10798-10801. (f) Wang, W.; Lorion, M.M.; Shah, J.; Kapdi, A.R.; Ackermann, L. Late-Stage Peptide Diversification by Position-Selective C-H activation. *Angew. Chem. Int. Ed.* **2018**, *57*, 2-20. (g) Petersen, J.; Christensen, K.E.; Nielsen, M.T.; Mortensen, K.T.; Komatnyy, V.V.; Nielsen, T.E.; Qvortrup, K. Oxidative Modification of Tryptophan-Containing Peptides. *ACS Comb. Sci.* **2018**, *20*, 344-349. (h) Foettinger, A.; Leitner, A.; Lindner, W. Selective Enrichment of Tryptophan-Containing Peptides from Protein Digests Employing a Reversible Derivatization with Malondialdehyde and Solid-Phase Capture on Hydrazide Beads. *J. Proteome Res.* **2007**, *6*, 3827-3834. (i) Yu, Y.; Zhang, L.K.; Buevich, A.V.; Li, G.; Tang, H.; Vachal, P.; Colletti, S.L.; Shi, Z.C. Chemoselective peptide Modification via Photocatalytic Tryptophan β -position Conjugation. *J. Am. Chem. Soc.* **2018**, *140*, 6797-6800.
- (5) (a) Tsai, C.-J.; Lin, S.L.; Wolfson, H.J.; Nussinov, R. Studies of protein-protein interfaces: A statistical analysis of the hydrophobic effect. *Protein Sci.* **1997**, *6*, 53-64. (b) Ma, B.; Elkayam, T.; Wolfson, H.; Nussinov, R. Protein-protein interactions: Structurally conserved residues distinguish between binding sites and exposed protein surfaces. *Proc. Natl. Acad. Sci.* **2003**, *100*, 5772-5777.
- (6) (a) Ma, J.C.; Dougherty, D.A. The Cation- π Interaction. *Chem. Rev.* **1997**, *97*, 1303-1324. (b) Mecozi, S.; West Jr., A.P.; Dougherty, D.A. Cation- π interactions in aromatics of biological and medicinal interest: electrostatic potential surfaces as a useful qualitative guide. *Proc. Natl. Acad. Sci. USA.* **1996**, *93*, 10566-10571. (c) Gallivan, J.P.; Dougherty, D.A. Cation- π interactions in structural biology. *Procl. Natl. Acad. Sci.* **1999**, *17*, 9459-9464. (d) Steiner, T.; Koellner, G. Hydrogen Bonds with π -Acceptors in Proteins: Frequencies and Role in Stabilizing Local 3D Structures. *J. Mol. Bio.* **2001**, *305*, 535-557. (e) Bissantz, C.; Kuhn, B.; Stahl, M. A Medicinal Chemist's Guide to Molecular Interactions. *J. Med. Chem.* **2010**, *53*, 5061-5084.
- (7) (a) Ma, B.; Nussinov, R. Trp/Met/Phe Hot Spots in Protein-Protein Interactions: Potential Targets in Drug Discovery. *Curr. Top. Med. Chem.* **2007**, *7*, 999-1005. (b) de Jesus, A.J.; Allen, T.W. The role of tryptophan side chains in membrane protein anchoring and hydrophobic mismatch. *BBA-Biomembranes.* **2013**, *1828*, 864-876. (c) Liu, W.; Caffrey, M. Interactions of Tryptophan, Tryptophan Peptides, and Tryptophan Alkyl Ester at Curved membrane Interfaces. *Biochemistry.* **2006**, *45*, 11713-11726. (d) Barik, S. The Uniqueness of Tryptophan in Biology: Properties, Metabolism, Interactions and Localization in Proteins. *Int. J. Mol. Sci.* **2020**, *21*, 8776-8797.
- (8) Winkler, J.R.; Gray, H.B. Could tyrosine and tryptophan serve multiple roles in biological redox processes? *Phil. Trans. R. Soc. A.* **2015**, *373*, 20140178.
- (9) (a) Tower, S.J.; Hetcher, W.J.; Myers, T.E.; Kuehl, N.J.; Taylor, M.T. Selective Modification of Tryptophan Residues in Peptides and Proteins Using a Biomimetic Electron Transfer Process. *J. Am. Chem. Soc.* **2020**, *142*, 9112-9118. (b) Orellana, N.V.; Taylor, M.T. Targeting Trp for Tagging through Photoinduced Electron Transfer. *Synlett.* **2021**, DOI: 10.1055/a-1479-6366
- (10) (a) Laroche, B.; Tang, X.; Archer, G.; Di Sanza, R.D.; Melchiorre, P. Photochemical Chemoselective Alkylation of Tryptophan-Containing Peptides. *Org. Lett.* **2021**, *23*, 285-289. (b) Rahimidashghoul, K.; Klimánková, I.; Hubálek, M.; Matoušek, V.; Filgas, J.; Slaviček, P.; Slanina, T.; Beier, P. Visible-Light-Driven Fluoroalkylation of Tryptophan Residues in Peptides. *ChemPhotoChem.* **2021**, *5*, 43-50. (c) Lima, R.N.; Delgado, J.A.; Bernardi, D.I.; Berlinck, R.G.S.; Kaplaneris, N.; Ackermann, L.; Paixão, M.W. Post-synthetic functionalization of tryptophan protected peptide sequences through indole (C-2) photocatalytic alkylation. *Chem. Commun.* **2021**, *57*, 5758-5761. (d) Weng, Y.; Ding, B.; Liu, Y.; Song, C.; Chan, L-Y.; Chiang, C-W. Late-Stage Photoredox C-H Amidation of *N*-Unprotected Indole Derivatives: Access to *N*-(Indol-2-yl)amides. *Org. Lett.* **2021**, *23*, 2710-2714. (e) Imiolek, M.; Isenegger, P.G.; Ng, W-L.; Khan, A.; Davis, B.G. Residue-Selective Protein C-Formylation via Sequential Difluoroalkylation-Hydrolysis. *ACS Cent. Sci.* **2021**, *7*, 145-155.
- (11) Scott, D.E.; Bayley, A.R.; Abell, C.; Skidmore, J. Small molecules, big targets: drug discovery faces the protein-protein interaction challenge. *Nat. Rev. Drug Discov.* **2016**, *15*, 533-550.
- (12) (a) Pattison, D.I.; Davis, M.J. Actions of ultraviolet light on cellular structure. *Cancer: Cell Structures, Carcinogens and Genomic Instability.* **2006**, 131-157. (b) Pattison, D.I.; Rahmanto, A.S.; Davies, M.J. Photo-oxidation of proteins. *Photochem. Photobiol. Sci.* **2012**, *11*, 38-53.
- (13) (a) Romero, N.A.; Nicewicz, D.A. Organic Photoredox Catalysis. *Chem. Rev.* **2016**, *116*, 10075-10166. (b) Dongare, P.; Mackenzie, I.; Wang, D.; Nicewicz, D.A.; Meyer, T.J. *Proc. Natl. Acad. Sci.* **2017**, *114*, 9279-9283.
- (14) He, F-S.; Ye, S.; Wu, J. Recent Advances in Pyridinium Salts as radical Reservoirs in organic Synthesis. *ACS Catal.* **2019**, *9*, 8943-8960.
- (15) Roth, H.G.; Romero, N.A.; Nicewicz, D.A. Experimental and Calculated Electrochemical Potentials of Common Organic Molecules for Applications to Single-Electron Redox Chemistry. *Synlett.* **2016**, *27*, 714-723.
- (16) Zanon, P.R.A.; Yu, F.; Musacchio, P.; Lewald, L.; Zollo, M.; Krauskopf, Mrdović, D.; Raunft, P.; Maher, T.E.; Cigler, M.; Chang, C.; Lang, K.; Toste, F.D.; Nesvizhskii, A.I.; Hacker, S.M. Profiling the proteome-wide selectivity of diverse electrophiles. *ChemRxiv.* **2021**. This content is a preprint and has not been peer-reviewed. DOI: 10.33774/chemrxiv-2021-w7rsv-v2.
- (17) Szychowski, J.; Mahdavi, A.; Hodas, J. J. L.; Bagert, J. D.; Ngo, J. T.; Landgraf, P.; Dieterich, D. C.; Schuman, E. M.; Tirrell, D. A. Cleavable Biotin Probes for Labeling of Biomolecules via Azide-Alkyne Cycloaddition. *J. Am. Chem. Soc.* **2010**, *132*, 18351-18360.

(18)(a) Box, J.K.; Paquet, N.; Adams, M.N.; Boucher, D.; Bolderson, E.; O'Byrne, K.J.; Richard, D.J. Nucleophosmin: from structure and function to disease development. *BMC Molecular Biol.* **2016**, 17, 19. (b) Grummitt, C.G.; Townsley, F.M.; Johnson, C.M.; Warren, A.J.; Bycroft, M. Structural Consequences of Nucleophosmin Mutations in Acute Myeloid Leukemia. *J. Biol. Chem.* **2008**, 283, 23326-23332. (c) Döhner, K.; Schlenk, R.F.; Habdank, M.; Scholl, C.; Rücker, F.G.; Corbacioglu, A.; Bullinger, L.; Fröhling, S.; Döhner, H. Mutant nucleophosmin (NPM1) predicts favorable prognosis in younger adults with acute myeloid leukemia and normal cytogenetics: interaction with other gene mutations. *Blood* **2005**, 106, 3740-3746.

(19) (a) Chaudhuri, A.R.; Nussenzweig, A. The multifaceted roles of PARP1 in DNA repair and chromatin remodeling. *Nat. Rev. Mol. Cell Biol.*

2017, 18, 610-621. (b) Langelier, M-F.; Pascal, J.M. PARP-1 mechanism for coupling DNA damage detection to poly(ADP-ribose) synthesis. *Curr. Opin. Struct. Biol.* **2014**, 23, 134-143. (c) Langelier, M-F.; Plank, J.L.; Pascal, J.M. Structural Basis for DNA Damage-Dependent Poly(ADP-ribose) synthesis by Human PARP-1. *Science* **2012**, 336, 728-732. (d) Steffen, J.D.; Tholey, R.M.; Langelier, M-F.; Planck, J.L.; Schiewer, M.J.; Lal, S.; Bildzukevicz, N.A.; Yeo, C.J.; Knudsen, K.E.; Brody, J.R.; Pascal, J.M. Targeting PARP-1 Allosteric Regulation Offers Therapeutic Potential against Cancer. *Cancer Res.* **2014**, 74, 31-37.

(20) Tien, M.Z.; Meyer, A.G.; Sydykova, D.K.; Spielman, S.J.; Wilke, C.O. Maximum Allowed Solvent Accessibilities of Residues in Proteins. *PLoS One* / **2013**, 8, e80635.

Insert Table of Contents artwork here

