HHS Public Access

Author manuscript

Chem Commun (Camb). Author manuscript; available in PMC 2023 May 11.

Published in final edited form as:

Chem Commun (Camb). ; 58(65): 9116–9119. doi:10.1039/d2cc02263j.

Harnessing the Intrinsic Photochemistry of Isoxazoles for the Development of Chemoproteomic Crosslinking Methods

Marshall G. Lougee[a], **Vinayak Vishnu Pagar**[a,b], **Hee Jong Kim**[c], **Samantha X. Pancoe**[a], **W. Kit Chia**[b], **Robert H. Mach**[b], **Benjamin A. Garcia**[c], **E. James Petersson**[a]

[a.]Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA.

[b.]Department of Radiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA.

[c.]Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104.

Abstract

The intrinsic photochemistry of the isoxazole, a common heterocycle in medicinal chemistry, can be applied to offer an alternative to existing strategies using more perturbing, extrinsic photo-crosslinkers. The utility of isoxazole photo-crosslinking is demonstrated in a wide range of biologically relevant experiments, including common proteomics workflows.

Photo-crosslinking or photoaffinity labeling (PAL) connects small molecule drug discovery to proteomics, providing detailed information like ligand binding sites or global information like off-target profiling.[1] Despite wide usage of photo-crosslinking in chemical biology, few advancements have been made regarding the photochemistry employed. Emphasis has been on time-tested photoreactive groups such as aryl azides, diazirines, and benzophenones.[2–4] However, the synthesis and attachment of these motifs is not always straightforward and may alter binding of the parent molecule. Recently, the Woo and Parker laboratories have investigated how the selectivity of these existing methods function and how it can be optimized.[5–7] A study from GlaxoSmithKline also enforced the utility of photo-crosslinkers in drug discovery, using diazirines in fragment-based screening to identify binders of protein targets.[8] There have been some recent expansions of the photo-crosslinking toolbox, but the collection is still limited.[9–11] Considering all of this, we sought to develop new photo-crosslinkers that could potentially be embedded in bioactive molecules.

Isoxazoles represent a common motif in medicinal chemistry.[12] Intriguingly, several studies have highlighted the labile nature of the isoxazole N-O bond. The Khlebnikov group has achieved transformations such as: rearrangement of vinyl-isoxazoles to pyrroles[13], isoxazole-azirine and isoxazole-oxazole isomerization[14], and azirine-oxazole conversion followed by reductive isoxazole ring-opening.[15] The majority of these reactions utilized

ejpetersson@sas.upenn.edu; Tel: 1-215-746-2221.

Electronic Supplementary Information (ESI) available: Synthesis and character-rization of all probes, protein expression and purification, all molecular biology protocols, proteomics methods and analysis. See DOI: 10.1039/x0xx00000x

a catalyst. However, there is significant precedent for light-activated ring opening of isoxazoles.[16] The isoxazole produce intermediates reminiscent of photo-crosslinking motifs during these transformations, including nitrenes and azirines (Scheme 1).[17] Indeed, muscimol and thiomuscimol, isoxazole-containing substrates of the gamma amino butyric acid receptor, have been shown to crosslink, presumably due to the photoreactivity of the isoxazole.[18–20] Isoxazoles represent an untapped functional group in terms of their innate ability to act as a minimalist photo-crosslinker and their presence in a wide variety of molecules in medicinal chemistry. We therefore sought to establish proof-of-principle for isoxazoles as photo-crosslinkers/PAL probes by demonstrating their utility in binding site validation and off-target analysis with implications for drug repurposing.

α-Synuclein (αS) is a 14.5 kDa protein that is implicated in several neurodegenerative diseases. αS insoluble aggregates (fibrils) are highly concentrated in Lewy Bodies, the pathological hallmark of Parkinson's disease.[21] In our efforts to develop positron emission tomography (PET) imaging probes for Parkinson's disease, isoxazole-containing molecules have been shown to bind to *in vitro* generated αS pre-formed fibrils with nM affinity.[22] We therefore used derivatives of these molecules as a first test of isoxazole crosslinking. We were excited to find that our PET leads could photo-crosslink to αS (ESI, Fig. S1)

We then synthesized a series of compounds to systematically assess electronic effects on isoxazole crosslinking by varying the ring at the 3-position of the isoxazole. Inclusion of a propargyl ether enables subsequent copper-catalyzed 3+2 cycloaddition (click) chemistry functionalization.[23, 24] The series **AS1–7** was validated against fibrils via matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry (MS) and fluorescent SDS-PAGE analysis. Upon irradiation with 254 nm light, all probes formed covalent adducts with αS that were labeled using an Alexa Fluor 488 (AF488) azide probe. (Fig. 1B) Thus, isoxazole crosslinking is not significantly affected by the electronic nature of the pendant ring system. Alternative wavelengths (365 nm, 427 nm) were examined as well, but resulted in decreased labeling efficiency. (ESI, Fig. S2–S3) A mass adduct corresponding to isobaric labeling was noted via whole protein MALDI-TOF MS, as well as an additional peak, possibly due to in-source decay (Fig. 1D).

To demonstrate that our molecules enable target enrichment, cross-linked αS was labeled with desthiobiotin azide (DTBA), disaggregated by boiling, bound to streptavidin beads, and eluted with excess biotin. A product with the additional mass of DTBA was observed in the elution fraction, indicating successful enrichment. (ESI, Fig. S5) Probe **AS7** was examined in more detail for the ability to label αS in the presence of human embryonic kidney (HEK) cell lysate. 100 μM **AS7** in the presence of 2 mg/ml HEK cell lysate showed little decrease in labeling efficiency, with a detection limit of 1–5 μM αS, demonstrating utility of isoxazoles as photo-labeling motifs even in the complex milieu of cell lysates. (Fig. 1C)

Given the efficiency of *in vitro* photo-crosslinking, we wished to determine whether we could obtain binding site information for our probe molecules by performing bottom-up proteomics. αS contains a variety of small-molecule binding sites,[25] and it is not obvious which of these are targeted by a given ligand. Upon irradiation of **AS3**, a substantial adduct

was noted by whole-protein MALDI-TOF MS (Fig. 1D). Upon tryptic digest, Orbitrap liquid chromatograph MS/MS (LC-MS/MS) was used to identify a match for residues 24–43 of αS (Fig. 1E). Analysis of the fragment MS/MS spectrum allowed us to confirm the identity of this crosslinked peptide and localize the crosslink to Gly_{31} (Fig. 1F). The majority of the other identified crosslinked peptides were also localized to this region (ESI, Fig. S6), consistent with previous mapping of the binding site of a similar molecule to this site using a diazirine.[22] Amino acid selectivity of the isoxazole can be inferred via detailed analysis of the identified crosslinks. Lysine and glycine showed the most reactivity, with minor crosslinks also identified to threonine, serine, alanine, and glutamine. (ESI, Fig. S6–S8) These results demonstrate the utility of isoxazole as a minimal photo-crosslinking motif in protein binding site analysis with a suitably promiscuous reactivity for proteomic applications.

To validate the use of isoxazole photo-crosslinkers more broadly, we shifted focus to FDA-approved drugs, where at least 15 compounds are currently or have previously been approved. Amongst these molecules, sulfisoxazole (SFX, Fig. 2A) is an antibacterial agent that functions as a competitive inhibitor of dihydropteroate synthase, which is necessary to bacterial folate biosynthesis.[26] In addition to its role as an antibiotic, **SFX** has been identified as a hit in anticancer drug repurposing screens, where it is proposed to inhibit extracellular vesicle (exosome) release.[27] There has been some debate as to the mechanism of reduction in tumor burden,[28, 29] motivating **SFX** PAL studies. Regardless of previously identified targets, novel binders of **SFX** would be useful for drug-repurposing as well as benchmarking of isoxazole photochemistry.

A sulfisoxazole probe (SFX-1, Fig. 2A) was synthesized to contain an alkyne for subsequent labeling and enrichment studies. **SFX-1** was first validated by photo-crosslinking in the presence of 2 mg/mL HEK cell lysate with subsequent conjugation to tetramethyl rhodamine (TAMRA) azide. From this, the optimal irradiation time was determined to be ~10 min, as a time-dependent decrease in protein banding upon prolonged UV exposure was observed. (ESI, Fig. S9) The application of this probe in a chemoproteomic workflow was then examined quantitatively in cytosolic MDA-MB-231 cell lysate via stable isotopic labeling using amino acids in cell culture (SILAC).[30] Freshly prepared lysate was chosen as a model, as the current isoxazole crosslinkers require 254 nm light, which may be damaging to cells and lead to altered protein profiles. The SILAC lysates were treated with either 250 μM probe (heavy “H” samples, **SFX-1**) or 250 μM probe and 1 mM competitor (light “L” samples, **SFX-1** + **SFX**). The resulting samples were then crosslinked, labeled with a photo-cleavable biotin azide, enriched on streptavidin beads, eluted by 365 nm irradiation, and analyzed via LC-MS/MS. (Fig. 2B) It should be noted that the SILAC experiments focused on cytosolic targets (Triton-soluble fraction) to enable optimal enrichment on streptavidin beads without denaturing the proteome.

A robust list of proteins was identified after filtering of contaminants (1555 total) that was further enriched (785 total) via thresholding (H/L ≥ 1.5, Andromeda score > 15, p-value < 0.05). (Fig. 2C) Given the remaining complexity of the target list, a bioinformatic similarity-based approach was adopted to select potential high-value targets of **SFX**. Proteins were analyzed via domain annotation in Uniprot (60% of initial hits annotated), and the proteins

pertaining to the top-10 enriched domains (122 total) were one-hot encoded and grouped using k-means clustering. (Fig. 2D; ESI, Table S1)

The enriched domains all represent interesting targets with potential ties to exosome biogenesis, the proposed mechanism of **SFX** as an anti-cancer agent. The top enriched domain was the RNA recognition motif (RRM), followed by thioredoxin and EF-hand. The RRM domain protein list was mainly composed of heterogenous nuclear ribonucleoproteins (hnRNPs), which have been shown to be enriched and shuttled in extracellular vesicles.[31] Thioredoxins, which maintain the proper oxidation of cysteine residues, are primarily intracellular proteins, but have been shown to be secreted, and thioredoxin-associated proteins are found in exosomes.[32, 33] The EF-hand domain is found in a variety of calcium-binding proteins, typified by calmodulin (CaM), which are implied in exosome biogenesis.[34] Aside from the EF-hand domain proteins, a number of other calcium-modulating proteins were enriched that have been tied to exosome formation including: synaptotagmin,[35] calpain,[36] and annexin family proteins.[37, 38] Additional gene-ontology (GO) analysis[39] revealed prominent enrichment of exosome cellular compartmentalization in the identified proteins. (ESI, Fig. S10)

To validate the observed targets, we attempted *in vitro* photo-crosslinking with RRM protein hnRNP A1, thioredoxin, and CaM. CaM and thioredoxin both showed photo-crosslinking with **SFX-1** and competition with various dosages of **SFX**. (ESI, Fig. S10). We did not see crosslinking to hnRNP A1, but this may be because the protein is not in its bioactive conformation in the absence of its cognate RNA. All of the targets will require further study to understand the basis for their interactions with **SFX** and the implications for its anti-cancer potential. More broadly, the discovery of novel targets for **SFX** reinforces the ability of isoxazole photo-crosslinking methods to spearhead drug-repurposing screens.

In this initial report, we have demonstrated that isoxazole compounds can be activated with UV light for photo-crosslinking to proteins and that this is compatible with several chemoproteomic workflows, including photo-cleavable biotin tags, with the addition of a bio-orthogonal reaction handle elsewhere on the molecule. We have validated isoxazole crosslinking in αS fibril experiments and showed that we identify the same binding site previously observed with a conventional diazirine tag. Isoxazole crosslinking has also provided insight into the interactome of **SFX**, with several novel targets identified. Certainly, there are some limitations to isoxazoles, such as the requirement for UV light, which may limit their application in live cells and tissues. As we broaden our understanding of the crosslinking mechanism, we may be able to address this, as well as further determine amino acid crosslinking preferences and the effects of pendant functional groups. We note that alkyl substituted isoxazoles have also shown promise, and preliminary proof of principle has been demonstrated with crosslinking of Valdecoxib (**VCOX-1**) in whole-cell lysates. (ESI, Fig. S31)

We have focused on the use of the isoxazole as an intrinsic crosslinker, but optimized versions could also be used as novel extrinsic crosslinkers to complement diazirines, aryl azides, and benzophenones. The immediate usefulness of isoxazoles as a photochemical motif emphasizes the need for the identification and development of additional novel photo-

crosslinkers, further enabling chemists to utilize functional, minimalist, and easily accessible small molecule probes. Indeed, studies of the photo-reactivity of other common medicinal chemistry heterocycles is already underway in our laboratory.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This research was supported by the National Institutes of Health (NIH R01-NS103873 to E.J.P., NIH U19-NS110456 to R.H.M. and E.J.P., and P01-CA196539 and R01-AI118891 to B.A.G.). Instruments supported by the NIH and National Science Foundation include NMR (NSF CHE-1827457) and mass spectrometers (NIH RR-023444 and NSF MRI-0820996) M.G.L. was supported by an Age Related Neurodegenerative Disease Training Grant fellowship (NIH T32-AG000255). H.J.K. was supported by an NIH Predoctoral Fellowship (NIH F31-AG069390). S.X.P. thanks the Vagelos Molecular Life Science Scholars program for support.

Notes and references

1. Smith E and Collins I, Future Med. Chem, 2015, 7, 159–183. [PubMed: 25686004]
2. Dorman G and Prestwich GD, Biochemistry, 1994, 33, 5661–5673. [PubMed: 8180191]
3. Dubinsky L, Krom BP and Meijler MM, Bioorg. Med. Chem, 2012, 20, 554–570. [PubMed: 21778062]
4. Murale DP, Hong SC, Haque MM and Lee J-S, Proteome Sci, 2017, 15, 14. [PubMed: 28652856]
5. Gao J, Mfuh A, Amako Y and Woo CM, J. Am. Chem. Soc, 2018, 140, 4259–4268. [PubMed: 29543447]
6. West AV, Muncipinto G, Wu H-Y, Huang AC, Labenski MT, Jones LH and Woo CM, J. Am. Chem. Soc, 2021, 143, 6691–6700. [PubMed: 33876925]
7. Conway LP, Jadhav AM, Homan RA, Li W, Rubiano JS, Hawkins R, Lawrence RM and Parker CG, Chem. Sci, 2021, 12, 7839–7847. [PubMed: 34168837]
8. Grant EK, Fallon DJ, Hann MM, Fantom KGM, Quinn C, Zappacosta F, Annan RS, Chung CW, Bamborough P, Dixon DP, Stacey P, House D, Patel VK, Tomkinson NCO and Bush JT, Angew. Chem. Int. Ed, 2020, 59, 21096–21105.
9. Herner A, Marjanovic J, Lewandowski TM, Marin V, Patterson M, Miesbauer L, Ready D, Williams J, Vasudevan A and Lin Q, J. Am. Chem. Soc, 2016, 138, 14609–14615. [PubMed: 27740749]
10. Page ACS, Scholz SO, Keenan KN, Spradlin JN, Belcher BP, Brittain SM, Tallarico JA, McKenna JM, Schirle M, Nomura DK and Toste FD, Chem. Sci, 2022, 13, 3851–3856. [PubMed: 35432890]
11. Cauwel M, Guillou C, Renault K, Schapman D, Bénard M, Galas L, Cosette P, Renard P-Y and Sabot C, Chem. Commun, 2021, 57, 3893–3896.
12. Zhu J, Mo J, Lin H.-z., Chen Y and Sun H.-p., Bioorg. Med. Chem, 2018, 26, 3065–3075. [PubMed: 29853341]
13. Galenko EE, Bodunov VA, Galenko AV, Novikov MS and Khlebnikov AF, J. Org. Chem, 2017, 82, 8568–8579. [PubMed: 28726412]
14. Serebryannikova AV, Galenko EE, Novikov MS and Khlebnikov AF, J. Org. Chem, 2019, 84, 15567–15577. [PubMed: 31658810]
15. Galenko EE, Bodunov VA, Kryukova MA, Novikov MS and Khlebnikov AF, J. Org. Chem, 2021, 86, 4098–4111. [PubMed: 33571409]
16. Galenko EE, Khlebnikov AF and Novikov MS, Chem. Heterocylcic Compounds, 2016, 52, 637–650.
17. Pavlik JW, Martin HST, Lambert KA, Lowell JA, Tsefrikas VM, Eddins CK and Kebede N, J. Heterocycl. Chem, 2005, 42, 273–281.
18. Smith GB and Olsen RW, J. Biol. Chem, 1994, 269, 20380–20387. [PubMed: 8051133]

19. Nielsen M, Witt M-R, Ebert B and Krogsgaard-Larsen P, Eur. J. Pharmacol, 1995, 289, 109–112. [PubMed: 7781704]
20. Huh KH, Delorey TM, Endo S and Olsen RW, Mol. Pharmacol, 1995, 48, 666. [PubMed: 7476892]
21. Spillantini MG, Schmidt ML, Lee VMY, Trojanowski JQ, Jakes R and Goedert M, Nature, 1997, 388, 839–840. [PubMed: 9278044]
22. Ferrie JJ, Lengyel-Zhand Z, Janssen B, Lougee MG, Giannakoulias S, Hsieh C-J, Pagar VV, Weng C-C, Xu H, Graham TJA, Lee VMY, Mach RH and Petersson EJ, Chem. Sci, 2020, 11, 12746–12754. [PubMed: 33889379]
23. Parker CG and Pratt MR, Cell, 2020, 180, 605–632. [PubMed: 32059777]
24. Rostovtsev VV, Green LG, Fokin VV and Sharpless KB, Angew. Chem. Int. Ed, 2002, 41, 2596–2599.
25. Hsieh C-J, Ferrie JJ, Xu K, Lee I, Graham TJA, Tu Z, Yu J, Dhavale D, Kotzbauer P, Petersson EJ and Mach RH, ACS Chem. Neurosci, 2018, 9, 2521–2527. [PubMed: 29750499]
26. Hong YL, Hossler PA, Calhoun DH and Meshnick SR, Antimicrob. Agents Chemother, 1995, 39, 1756–1763. [PubMed: 7486915]
27. Im EJ, Lee CH, Moon PG, Rangaswamy GG, Lee B, Lee JM, Lee JC, Jee JG, Bae JS, Kwon TK, Kang KW, Jeong MS, Lee JE, Jung HS, Ro HJ, Jun S, Kang W, Seo SY, Cho YE, Song BJ and Baek MC, Nat. Commun, 2019, 10.
28. Fonseka P, Chitti SV, Sanwlani R and Mathivanan S, Nat. Commun, 2021, 12, 977. [PubMed: 33579909]
29. Lee C-H, Bae J-H, Kim J-I, Park J-M, Choi E-J and Baek M-C, Nat. Commun, 2021, 12, 976. [PubMed: 33579906]
30. Mann M, Nat. Rev. Mol. Cell Biol, 2006, 7, 952–958. [PubMed: 17139335]
31. Fabbiano F, Corsi J, Gurrieri E, Trevisan C, Notarangelo M and D'Agostino VG, J. Extracell. Vesicles, 2020, 10, e12043. [PubMed: 33391635]
32. Plugis Nicholas M, Weng N, Zhao Q, Palanski Brad A, Maecker Holden T, Habtezion A and Khosla C, Proc. Natl. Acad. Sci. USA, 2018, 115, 8781–8786. [PubMed: 30104382]
33. Kowal J, Arras G, Colombo M, Jouve M, Morath Jakob P, Primdal-Bengtson B, Dingli F, Loew D, Tkach M and Théry C, Proc. Natl. Acad. Sci. USA, 2016, 113, E968–E977. [PubMed: 26858453]
34. Chazin WJ, Acc. Chem. Res, 2011, 44, 171–179. [PubMed: 21314091]
35. Tucker WC and Chapman ER, Biochem. J, 2002, 366, 1–13. [PubMed: 12047220]
36. Taylor J, Jaiswal R and Bebawy M, Curr. Cancer Drug Targets, 2017, 17, 486–494. [PubMed: 27799031]
37. Théry C, Boussac M, Véron P, Ricciardi-Castagnoli P, Raposo G, Garin J and Amigorena S, J. Immunol, 2001, 166, 7309. [PubMed: 11390481]
38. Koumangoye RB, Sakwe AM, Goodwin JS, Patel T and Ochieng J, PLOS ONE, 2011, 6, e24234. [PubMed: 21915303]
39. Schölz C, Lyon D, Refsgaard JC, Jensen LJ, Choudhary C and Weinert BT, Nat. Methods, 2015, 12, 1003–1004. [PubMed: 26513550]

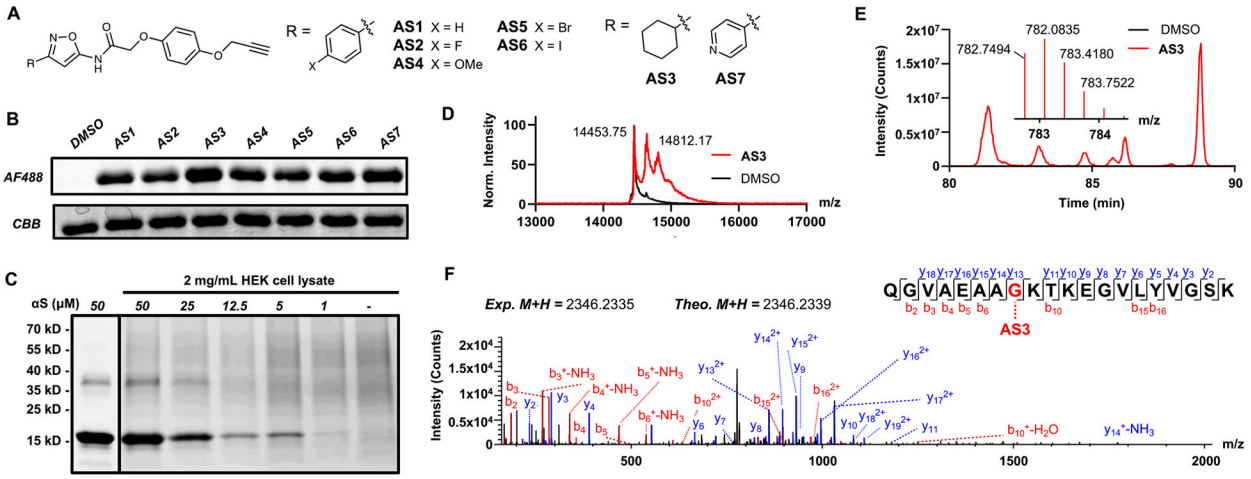


Fig. 1.

A. **AS1–7** isoxazole probes. B. Fluorescent SDS-PAGE of isoxazole-containing photo-crosslinkers irradiated at 254 nm in the presence of fibrillar αS. AF488: Fluorescent detection of αS photoproducts click labeled with AF488. CBB: Coomassie staining showing protein present in each lane. C. Crosslinking of fibrillar αS to **AS7** in the presence and absence of HEK cell cytosolic lysate. D. Whole protein MALDI-TOF MS of **AS3** crosslinking to fibrillar αS. E. αS fragment crosslinked to **AS3** identified via Orbitrap LC-MS/MS (extracted and parent ion chromatograms.) DMSO controls show spectra for un-crosslinked αS. F. Fragment (MS/MS) spectrum of **AS3** crosslinked peptide with annotated b and y ion series.

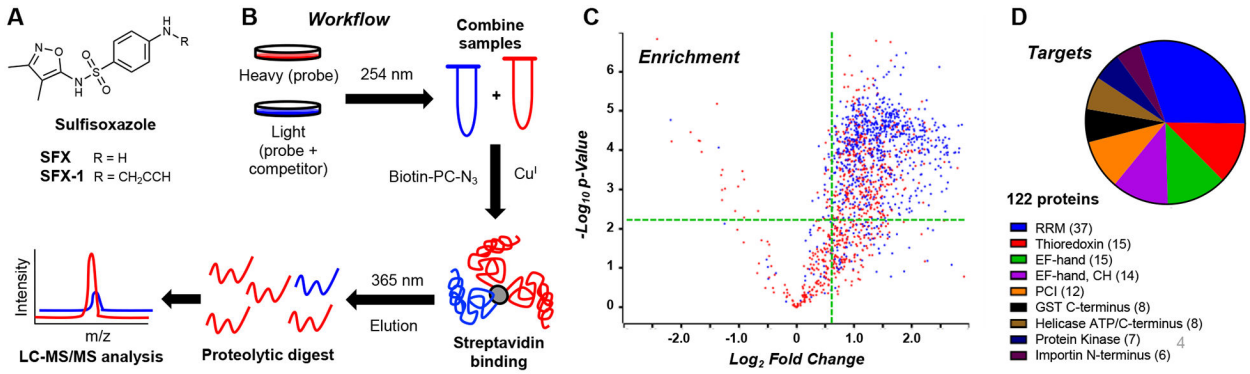
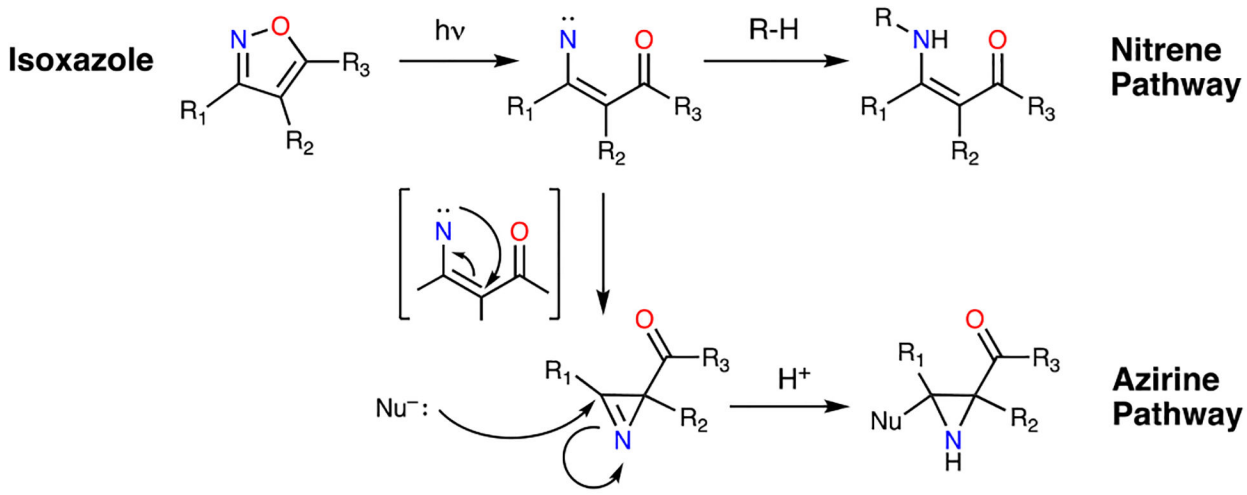


Fig. 2.

A. Structures of sulfisoxazole probes. B. Schematic representation of the chemoproteomic workflow utilized in this study. C. Volcano plot representing **SFX-1** enrichment of proteins with an andromeda score greater than 15 colored in blue and the enrichment region for the top domains highlighted in green. D. Pie chart representing proteins targeted by **SFX-1**, organized by k-means clustering of specific protein domains, with sections sized depending on number of proteins. See ESI for details.



Scheme 1.
Isoxazole photo-crosslinking mechanisms