Enhancing Cysteine Chemoproteomic Coverage Through Systematic Assessment of Click Chemistry Product Fragmentation

Tianyang Yan[†,‡], Andrew B. Palmer[†,‡], Daniel J. Geiszler[§], Daniel A. Polasky[∥], Ernest Armenta[†‡], Alexey I. Nesvizhskii[§,∥], Keriann M. Backus[†,‡,₵,Ŧ,Ł,ʍ,*]

† Biological Chemistry Department, David Geffen School of Medicine, UCLA, Los Angeles, CA, 90095, USA.

‡ Department of Chemistry and Biochemistry, UCLA, Los Angeles, CA, 90095, USA.

§ Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, 48109, USA.

∥ Department of Pathology, University of Michigan, Ann Arbor, MI, 48109, USA.

₵ Molecular Biology Institute, UCLA, Los Angeles, CA, 90095, USA.

Ŧ DOE Institute for Genomics and Proteomics, UCLA, Los Angeles, CA, 90095, USA.

Ł Jonsson Comprehensive Cancer Center, UCLA, Los Angeles, CA, 90095, USA.

ʍ Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, UCLA, Los Angeles, CA, 90095, USA.

ABSTRACT: Mass spectrometry-based chemoproteomics has enabled functional analysis and small molecule screening at thousands of cysteine residues in parallel. Widely adopted chemoproteomic sample preparation workflows rely on the use of pan-cysteine reactive probes such as iodoacetamide alkyne combined with biotinylation via copper-catalyzed azide–alkyne cycloaddition (CuAAC) or 'click chemistry' for cysteine capture. Despite considerable advances in both sample preparation and analytical platforms, current techniques only sample a small fraction of all cysteines encoded in the human proteome. Extending the recently introduced labile mode of the MSFragger search engine, here we report an in-depth analysis of cysteine biotinylation via click chemistry (CBCC) reagent gas-phase fragmentation during MS/MS analysis. We find that CBCC conjugates produce both known and novel diagnostic fragments and peptide remainder ions. Among these species, we identified a candidate signature ion for CBCC peptides, the oxonium-biotin fragment ion that is generated upon fragmentation of the N(triazole)–C(alkyl) bond together with cyclization. Guided by our empirical comparison of the fragmentation patterns of five CBCC reagent combinations, we achieved enhanced coverage of cysteine labeled peptides. For larger, fragmentation-prone biotinylation reagents, implementation of labile search afforded unique PSMs and provides a roadmap for the utility of such searches in enhancing chemoproteomic peptide coverage.

Introduction

Chemoproteomics has become a workhorse technology for functional biology and drug discovery efforts, enabling target deconvolution for a number of bioactive molecules, clinical candidates, and even drugs.[1-4] Most sample preparation workflows rely on the same general strategy. Cells or cellular lysates are first labeled with chemical probes that incorporate an electrophilic moiety or a photoactivatable group for irreversible modification of the probe-binding proteins. Probe-labeled proteomes are then conjugated via bioorthogonal copper-catalyzed azide–alkyne cycloaddition (CuAAC) or 'click chemistry' to biotin or desthiobiotin capture reagents (typically azide-or alkyne-modified).[5-7] The use of bioorthogonal chemistry for protein and peptide capture, in contrast with probes derivatized directly with biotin, offers the advantage of decreasing probe size and increasing cell permeability.[8,9] After enrichment on avidin resin and proteolytic digest, labeled peptides are then identified through liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis.

Cysteine-reactive probes are particularly favored for chemical probe development campaigns, given the unique chemistry of the cysteine thiol[10] and its important roles in functional and therapeutically relevant targets.[11,12] Chemoproteomics studies using cysteine-reactive probes such as iodoacetamide alkyne have demonstrated that the proteome harbors thousands of potential ligandable cysteines.[13,14] However, a central challenge of these studies is that they still only sample a small fraction (2.4%) of all cysteines in the human proteome (estimated to comprise >260,000 residues).[15] Closing this gap requires the development of optimized chemoproteomic sample preparation and data analysis workflows. Supporting this premise, our recent work revealed that using a SP3 cleanup method combined

with on-line sample fractionation could achieve 5.5-fold increased coverage compared to prior methods, with decreased instrument time.[15]

One aspect of global cysteine chemoproteomics workflows that remains largely unexplored is the impact of reagent gas-phase fragmentation during MS/MS analysis and whether such fragmentation might impact data analysis or coverage of labeled peptides. The inertness of the triazole formed during 'click chemistry' to gas-phase reactions has been widely documented.[16–18] In one notable exception to this paradigm, collision-induced dissociation (CID) of triazole-modified peptides has been found to afford nucleophilic displacement of the N3 nitrogen in the 1,2,3-triazole ring.[19] This fragmentation and the resulting stable quaternary ammonium ion were previously leveraged to generate custom isobaric tags.[20] Whether and to what extent 1,2,3-triazoles show additional gas-phase fragments remains unexplored.

In contrast with the seeming gas-phase inertness of the triazole moiety, the MS/MS fragmentation of biotin conjugates have been widely documented, including the application of signature reporter ions for identification of biotinylated peptides. These studies have extended to the fragmentation of biotin-lysine conjugates, biotin-phenol modified tyrosine residues (e.g. conjugates formed during APEX proximity labeling[21]) and biotin conjugates formed during covalent modification of cysteine,[22] serine,[23] and tyrosine residues.[24,25] Most notably, incorporation of signature biotin fragments into MS/MS analysis workflows has proven instrumental for identification of the specific sites of labeling for serine hydrolase reactive fluorophosphonate-biotin reagents (FP-biotin).[23] Dehydrobiotin (m/z 227.0845) is a diagnostic ion shared across most biotin fragmentation studies.

The variable modification approach employed by search algorithms (e.g. SEQUEST,[26] PROLUCID[27]) widely used for chemoproteomics data analysis looks for fragment ions that contain intact biotin conjugates, thereby under sampling peptides with labile modifications. Some algorithms, (e.g. COMET[28]), allow for neutral losses to be specified. Open search is an alternative approach in which the peptide mass is determined by matching fragment ions without knowledge of the precursor mass.[29,30] For labile modifications, open searches, together with the more focused mass offset[31] or multinotch[32] searches enable spectral matching for a larger proportion of observed fragmentation ions. Showcasing the utility of these algorithms for achieving increased coverage of peptides with labile modifications, offset searches of glycoproteomics datasets were found to significantly increase the number of successfully annotated spectra.[33]

Here we report the discovery that gas phase fragmentation of triazole-and biotin-modified peptides affords several characteristic fragment ions, including previously reported and newly identified species. Using open and mass offset search strategies, we conduct an in-depth analysis of the relative intensities and specificity of these signature fragments. By varying the nature of the labeling reagent, including linker length and position of the triazole, we achieved improved coverage of labeled peptides. The combination of labile and non-labile ion searches further enhanced coverage of chemoproteomic detected cysteines identified using larger biotinylation reagents. Collectively our study demonstrates the utility of labile ion and mass offset search strategies in the analysis of chemoproteomics datasets and reveals the ubiquity of triazole fragmentation in MS/MS analysis of chemically modified peptides.

Experimental section

Proteomic sample preparation. Samples were prepared as reported.[15] Jurkat proteome (200 µL of 1 mg/mL) was first labeled with IAA **1** or other reagents (4 µL of 100 mM stock solution in DMSO, final concentration = 2 mM) for 1h at ambient temperature. CuAAC was performed with biotin-azide **2** or other reagents (4 µL of 200 mM stock in DMSO, final concentration = 4 mM), TCEP (4 µL of fresh 50 mM stock in water, final concentration = 1 mM), TBTA (12 µL of 1.7 mM stock in DMSO/t-butanol 1:4, final concentration = 100 µM), and CuSO4 (4 µL of 50 mM stock in water, final concentration = 1 mM) for 1h at ambient temperature. After CuAAC labeling, each sample was treated with 20 µL 10% SDS/PBS and then 0.5 µL benzonase (Fisher Scientific, 70-664-3) for 30 min at 37 °C. For each 200 µL sample (1 mg/mL protein concentration), 20 µL Sera-Mag Speed-Beads Carboxyl Magnetic Beads, hydrophobic (GE Healthcare, 65152105050250) and 20 µL Sera-Mag Speed-Beads Carboxyl Magnetic Beads, hydrophilic (GE Healthcare, 45152105050250) were mixed and washed with water for three times. The bead slurries were then transferred to the CuAAC samples, incubated for 5 min at RT with shaking (1000 rpm). Absolute ethanol (400 µL) was added to each sample, and the samples were incubated for 5 min at RT with shaking (1000 rpm). Samples were then placed on a magnetic rack, washed three times with 80% ethanol in water (400 µL). After washing, beads were resuspended in 200 µL 2 M urea in 0.5% SDS/PBS. DTT (10 µL of 200 mM stock in water, final concentration = 10 mM) was added into each sample and the sample was incubated at 65 °C for 15 min. Then, iodoacetamide (10 µL of 400 mM stock in water, final concentration = 20 mM) was added and the solution was incubated for 30 min at 37 °C with shaking in the dark. After that, absolute ethanol (400 µL) was added to each sample, and the samples were incubated for 5 min at RT with shaking (1000 rpm). Beads were then again washed three times with 80% ethanol in water (400 µL). Next, beads were resuspended in 200 µL 2 M urea in PBS and 2 µL trypsin solution was added. Digest was overnight at 37 ºC with shaking. After digestion, ∼ 4 mL acetonitrile (> 95% of the final volume) was added to each sample and the mixtures were incubated for 10 min at RT with shaking (1000 rpm). The beads were then washed (3 × 1 mL acetonitrile) with a magnetic rack. Peptides were eluted from SP3 beads with 100 µL of 2% DMSO in MB water for 30 min at 37 °C with shaking (1000 rpm). The elution was repeated again with 100 µL of 2% DMSO in MB water. For each sample, 50 µL of NeutrAvidin Agarose resin slurry (Pierce, 29200) was washed three times in 10 mL IAP buffer (50 mM MOPS pH 7.2, 10 mM sodium phosphate, and 50 mM NaCl buffer) and

then resuspended in 800 μL IAP buffer. Peptide solutions eluted from SP3 beads were then transferred to the NeutrAvidin Agarose resin suspension, and the samples were rotated for 2h at RT. After incubation, the beads were pelleted by centrifugation (21,000 g, 1 min) and washed (3 × 1 mL PBS, 6 × 1 mL water). Bound peptides were eluted twice with 60 μL of 80% acetonitrile in MB water containing 0.1% FA. The first 10 min incubation at RT and the second one at 72 °C. The combined eluants were dried (SpeedVac), then reconstituted with 5% acetonitrile and 1% FA in MB water and analyzed by LC-MS/MS.

Liquid-chromatography tandem mass-spectrometry (LC-MS/MS) analysis. The samples were analyzed by liquid chromatography tandem mass spectrometry using a Thermo Scientific™ Orbitrap Eclipse™ Tribrid™ mass spectrometer or coupled with a High Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS) Interface. Peptides were fractionated online using a 18cm long, 100 μM inner diameter (ID) fused silica capillary packed in-house with bulk C18 reversed phase resin (particle size, 1.9 μm; pore size, 100 Å; Dr. Maisch GmbH). The 70-minute water-acetonitrile gradient was delivered using aThermo Scientific™ EASY-nLC™ 1200 system at different flow rates (Buffer A: water with 3% DMSO and 0.1% formic acid and Buffer B: 80% acetonitrile with 3% DMSO and 0.1% formic acid). The detailed gradient includes 0 – 5 min from 3 % to 10 % at 300 nL/min, 5 – 64 min from 10 % to 50 % at 220 nL/min, and 64 – 70 min from 50 % to 95 % at 250 nL/min buffer B in buffer A **(Table S7)**. Data was collected with charge exclusion (1, 8,>8). Data was acquired using a Data-Dependent Acquisition (DDA) method consisting of a full MS1 scan (Resolution = 120,000) followed by sequential MS2 scans (Resolution = 15,000) to utilize the remainder of the 1 second cycle time. Precursor isolation window and normalized collision energy were set as described in the study. The MS data have been deposited to the ProteomeXchange Consortium[34] via the PRIDE partner repository[35] with the dataset identifier PXD028853. File details were listed in **Table S6**.

Protein and peptide identification. Raw data collected by LC-MS/MS were searched with MSFragger (v3.3) and FragPipe (v16.0). For closed search, the proteomic workflow and its collection of tools was set as default. Precursor and fragment mass tolerance was set as 20 ppm. Missed cleavages were allowed up to 1. Peptide length was set 7 - 50 and peptide mass range was set 500 - 5000. Cysteine residues were searched with differential modifications. For labile search, mass offsets were set restricted to cysteines. Y ion masses and diagnostic fragment masses were set as in **Scheme 1**, **Scheme 2** and **Figure S8** for different proteomic samples. PTM-Shepherd was enabled for localization.[36] A sample workflow can be found in SI. Diagnostic fragment ion and peptide remainder masses were identified using a labile offset search described above without diagnostic and Y ions specified. Calibrated and deisotoped spectrum files produced by FragPipe were retained and reused for this analysis. The MS search results have been deposited to the ProteomeXchange Consortium[34] via the PRIDE partner repository[35] with the dataset identifier PXD028853. File details were listed in **Table S6**.

Results

Identification and verification of oxonium-biotin as the major fragmentation product of cysteine biotinylation via click chemistry (CBCC).

During our prior chemoproteomics analysis of the cysteinome[15] following the cysteine biotinylation via 'click' chemistry (CBCC) sample preparation workflow shown in **Figure 1A**, we observed that nearly all MS/MS spectra displayed signals at m/z 227.085 and m/z 284.143 (**Figure 1B**). The m/z 227.085 fragment ion matches closely with the calculated exact mass of dehydrobiotin **F1**, which has been previously reported as signature ion for biotin-modified peptides.[22,24,25,37] While the m/z 284.143 fragment ion has not been previously reported as a signature ion associated with biotin, close inspection of potential labile bonds in the biotin-triazole product **M1**, suggested that this signature ion likely results from fragmentation at the N(triazole)–C(alkyl) bond (N6-C5) together with cyclization to afford cyclic oxonium species **F2**, termed here oxonium-biotin (**Figure 1C**).

As this fragmentation of biotin-triazole conjugate is, to our knowledge, unprecedented, we next sought to further vet the identity of the m/z 284.143 fragment ion. MS3 analysis of biotinylated peptides revealed that the dehydrobiotin and biotin-oxonium ions were the only major fragment ions produced from CBCC, and both ions were observed with high relative intensity (**Figure 1D**). Consistent with prior MS3 studies, for example those that rely on isobaric tags,[38] these data support these fragment ions are likely derived from the cysteine biotin modification.

To further pinpoint the source of m/z 284.143 fragment ion, we next synthesized and applied a deuterated isotopologue **3** of our biotin-azide capture reagent (**Scheme S1** and **Figure 1E**) to chemoproteomic analysis of cysteine-containing peptides. IAA-labeled samples were subjected to capture, enrichment, and MS/MS analysis (**Figure 1A**) using biotin-D-azide **3**. As predicted, results revealed the production of a new characteristic fragment ion with m/z 290.180. Compared to oxonium-biotin, the mass shift is consistent with the -d6 modification of the reagent **S1–S3 in Scheme S1** compared to biotin-azide **2**. A comparison of samples labeled with the 'light' **1** and 'heavy' **3** biotin reagents, revealed comparable PSMs and unique peptides were detected (**Figure 1F**), consistent with equal performance of the light and heavy reagents. For samples labeled with 1:1 heavy and light biotin-azide tags, quantification of the relative intensities of labeled peptides, using FragPipe's IonQuant[39] revealed a median log2(H:L) ratio close to zero with minimal variance across peptides quantified, which indicated equal labeling efficiency for both reagents (**Figure S2A**). Consistent with prior reports,[40] we do observe a minor (~3 sec) average retention time shift for peptides labeled with the deuterated reagent **3** (**Figure S2B**).
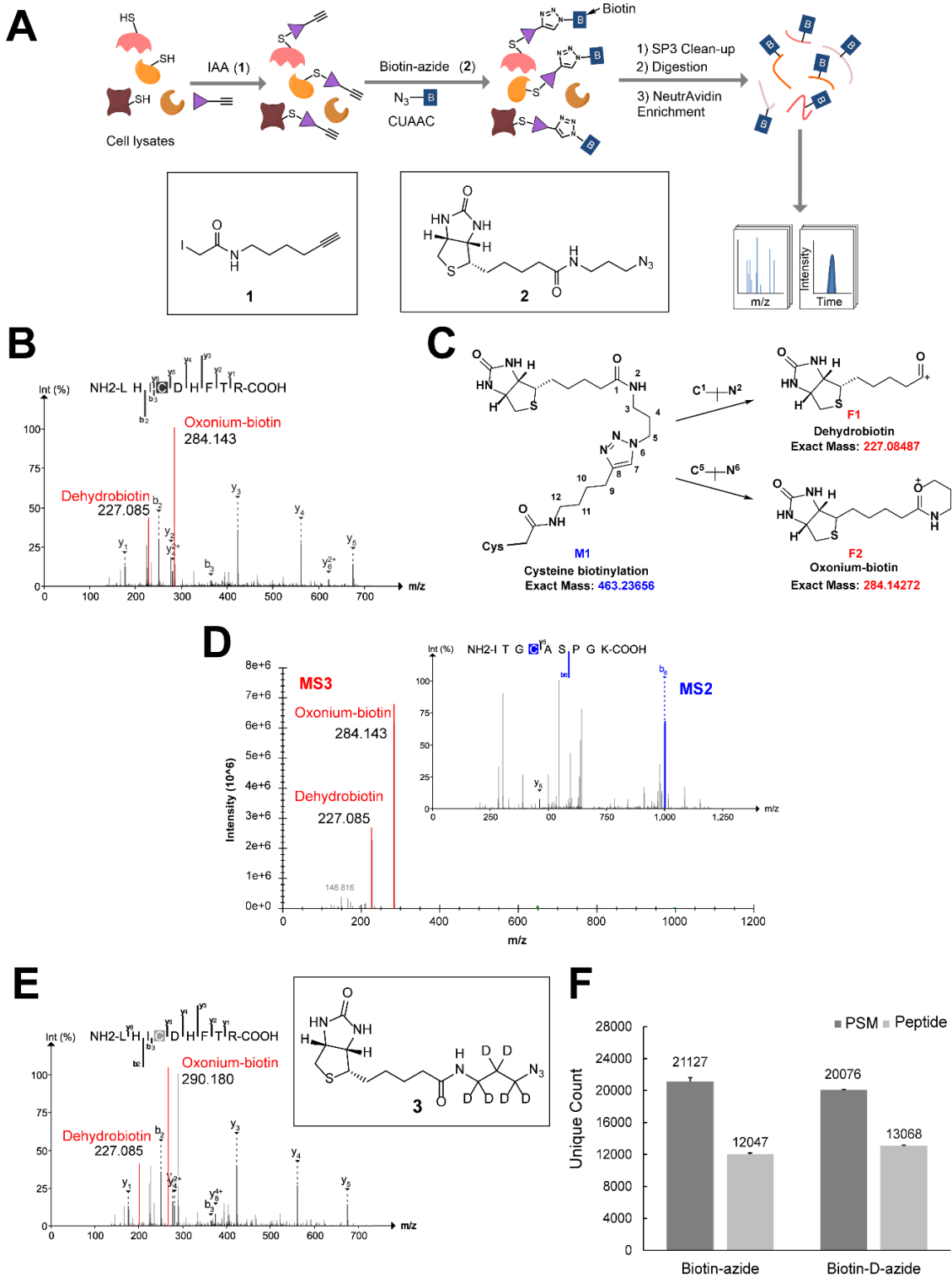
Figure 1. (A) Workflow for cysteine chemoproteomic identification. Cysteines are capped with the pan cysteine reactive probe iodoacetamide alkyne (IAA or **1**), followed by copper-catalyzed azide-alkyne cycloaddition (CuAAC or 'click' chemistry) conjugation to biotin-azide **2**, SP3 sample cleanup, tryptic digest, neutravidin enrichment and LC-MS/MS analysis. (B) Representative MS/MS spectrum of biotinylated cysteine peptide. (C) Scheme of potential fragmentations of the click product and the resulting fragment ions. Blue labeled remainder modification of the biotinylated cysteines and red labeled fragment ions. (D) Representative MS3 spectrum of biotinylated cysteine b/y ions. (E) Representative MS/MS spectrum of cysteine peptide labeled with biotin-D-azide **3**. (F) Mean unique counts of PSMs and peptides identified with biotin-azide **2** and biotin-D-azide **3**. Experiments were performed in duplicate. Data can be found in **Table S1**.

Scheme 1. Fragment ions and remainder modification on cysteines from fragmentation of cysteine biotinylation with IAA 1 and biotin-azide 2 via click chemistry. Blue labeled remainder modification of the biotinylated cysteines and red labeled fragment ions.
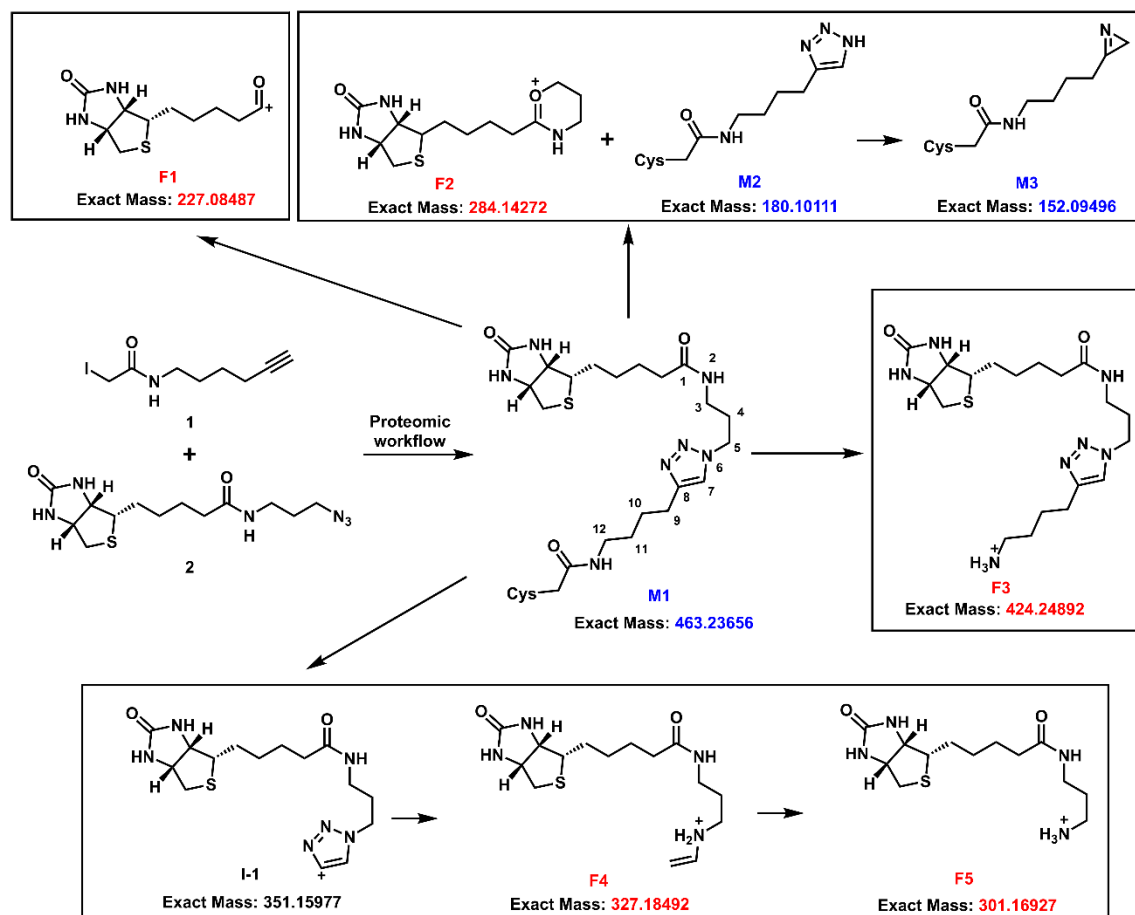
In depth analysis of the fragmentation products of CBCC.

Having pinpointed the likely nature of the *m/z* 284.143 fragment ion, we next expanded our analysis of the CBCC fragmentation products with the overarching objective of more fully deciphering how chemoproteomic-labeled peptides behave in the gas phase. Inspired by recent advances in the identification of peptides with labile modifications, we opted to test whether the mass offset search of the MSFragger search engine[33,41] could be extended to analysis of CBCC fragmentation productions.

We first extended PTM-Shepherd to identify biotinylation-specific spectral features corresponding to fragments of the biotinylation modification (diagnostic ions) or partially fragmented biotinylation remaining on the peptide (diagnostic peptide remainder masses) (**Figure 2A**). For biotinylated PSMs, spectra had all possible diagnostic features calculated and aggregated into a common histogram. For diagnostic ions, ions were inserted into the histogram at their *m/z* observed in the experimental spectrum. For peptide remainder masses, the distance between every ion in the experimental spectrum and the theoretical, unmodified peptide mass was calculated and inserted into the common histogram. Recurring features were identified based on bin height. The recurring features were then quantified across modified and unmodified spectra, and their specificity for biotinylation was assessed via comparisons to the unmodified PSMs.

As a result, several diagnostic fragment ions and peptide remainder masses were identified, including *m/z* 227.085 (dehydrobiotin), *m/z* 284.143 (oxonium-biotin), *m/z* 424.249, *m/z* 327.185, *m/z* 301.169 as fragment ions and +180.101, +152.095, +463.237 (intact modification) as peptide mass shifts. Based on the diagnostic masses and structure of the modification, several fragmentation pathways were proposed (**Scheme 1**). We expected that N6–C5 bond cleavage and cyclization with C1 carbonyl oxygen of the **M1** ion would afford biotin oxonium ion **F2** and peptide remainder ions **M2** and **M3**, with **M3** produced by an additional loss of nitrogen **M2**.[42] **M1** would also fragment to produce both the aforementioned dehydrobiotin **F1** and diagnostic fragmentation **F3**, which is generated by cleavage of the C12-N13 bond. While still somewhat speculative, we anticipate that **F4** and **F5** form as a result of cleavage of the C9-C8 bond followed by loss of nitrogen and acetylene gas, respectively. Manual annotation of MS2 spectra of representative peptides using an integrative proteomics data viewer (PDV)[43] (**Figure 2B** and **S3**) identified spectra
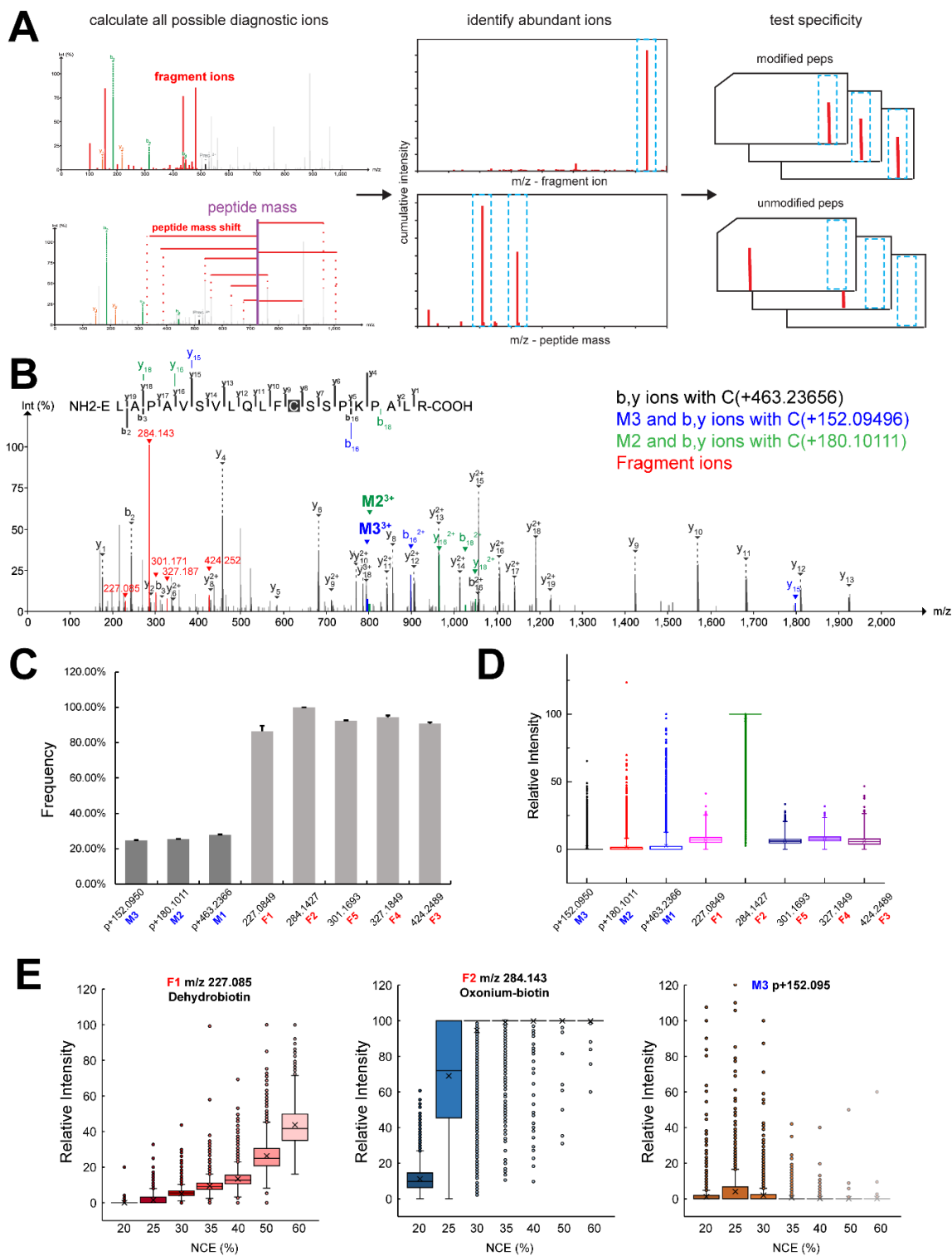
Figure 2. (A) Workflow for the identification of diagnostic ions and peptide remainder masses. All possible diagnostic ions (top) and peptide remainder masses (bottom) were calculated using an extended version of PTM-Shepherd (left panel), common features (light blue) across spectra were identified (middle panel) and specificity were tested against unmodified peptides (right panel). (B) Representative MS/MS spectra of biotinylated cysteine peptides with annotations of fragment ions (**F1-F5**), peptide remainder ions (**M2**, **M3**), as well as b/y ions with different cysteine modifications. (C) Frequency distribution and (D) Relative intensities of signature fragment ions and peptide remainder ions of cysteine biotinylation. (E) Relative intensities of representative ions (**F1**, **F2**, **M3**) in HCD fragmentation mode with varying normalized collision energies (NCE). Experiments were performed in duplicate. Bar plots display mean values across replicates and box plots display minimum, first quartile (Q1), median, third quartile (Q3), and maximum values of the sample (similarly hereinafter). All data can be found in **Table S1**.

Table 1. Number of biotinylated PSMs and peptides identified in HCD fragmentation mode with varying NCE.

NCE (%)	20	25	30	35	40	50	60
PSMs	16570	16664	17214	15412	14105	5618	953
Peptides	11882	11757	12037	11155	10366	4602	874

corresponding to fragments **F1-F5** (red ions), peptide remainder ions **M2** (green), **M3** (blue) as peptide ions with a diagnostic mass shift and b/y ions with corresponding cysteine modifications.

Looking across our entire labile ion search dataset (**Table S1**), we next analyzed the frequency distribution of all the identified fragment ions. Unsurprisingly, frequency analysis showed that **F2** oxonium-biotin ion is the most common fragment ion generated from LC-MS/MS analysis of CBCC samples, with nearly 100% occurrence in all MS2 scans. The **F1** and **F3** fragment ions were also ubiquitous, appearing in >80% of MS2 scans (**Figure 2C**). In contrast, **M1** (p+463.237)-, **M2** (p+180.101)-, and **M3** (p+152.095)-modified peptide remainder ions are only found in ~25% of all MS2 scans. The relatively modest frequency distribution of modified peptide precursor ions can be rationalized in part by only partial fragmentation of the biotin modification, natural low abundance of unfragmented peptide backbone on MS2 level, together with the inherent ambiguity and challenges associated with accurate annotation of tandem mass spectra for modified peptides, particularly those featuring more labile modifications.[44]

As demonstrated by our representative annotated spectra (**Figure 2B** and **S3**), inspection of MS2 scans for representative peptides revealed that the oxonium-biotin **F2** fragment ion was almost always the dominant ion with 100% relative intensity. To more rigorously quantify the relative intensities of each characteristic fragmentation ion, we established a customized PTM-Shepherd workflow in FragPipe that reports the relative intensities for all the fragment and peptide remainder ions of biotinylated cysteine peptides (See methods). Consistent with our observations from manual inspection of spectra, dataset-wide quantitation revealed a 100% median intensity for the oxonium-biotin ion (**Figure 2D** and **Table S1**). All other fragment and peptide remainder ions were found to have substantially reduced median of relative intensities spanning 0% - 7.6% (**Figure 2D** and **Table S1**), which is consistent with their low frequency of detection. The dominance of the biotin-oxonium **F2** ion in nearly all spectra makes it an intriguing candidate signature ion for CBCC-labeled peptides.

To further understand how collision energy would impact the signal intensity of the observed fragment and peptide remainder ions, we next subjected CBCC peptide samples to a collision energy ramping experiment comparing LC-MS-HCD-MS methods varying normalized collision energy (NCE) from NCE = 20 to NCE = 60 (**Figure 2E** and **Table 1**). We found that the greatest number of biotinylated PSMs and unique peptides were identified at NCE=30 and the lowest number of peptides at NCE=60 (**Table 1**). Analysis of the relative intensities of the fragment and peptide remainder

ions by labile search, revealed a dramatic increase in the intensity of diagnostic ions including *m/z* 227.085 and *m/z* 284.143 with ramping of collision energy in HCD modes (**Figure 2E**), while other fragment ions (**F3**, **F4**, **F5**) remained relatively low intensity (**Figure S4**). In contrast, the intensities of signature peptide precursor ions decrease as collision energy is increased, consistent with over-fragmentation of precursor ions (**Figure 2E** and **S3**). On the other hand, when comparing biotinylated PSMs and peptides identified with CID or HCD, we observed ~30% fewer IDs using CID at 30% NCE (**Table S2**). No apparent change of relative intensities of fragment ions and peptide remainder ions was observed with varying NCE, stating the fragmentation of CBCC was favorable with only HCD (**Figure S5**).

Assessment of the specificity of the fragmentation products

High specificity is essential for a fragment ion to serve as a signature or diagnostic ion for a specific chemical modification. To test the specificity of each observed fragment ion for precursor ion modification state, we next quantified the frequency distribution of each ion for the fragment and peptide remainder ions for PSMs harboring or lacking cysteine modifications (with or without (w/o) mass shift, respectively; **Figure 3A, B**). This comparison of biotin-labeled and unlabeled PSMs revealed high and statistically significant specificity for all three signature peptide precursor ions (**M1**, **M2**, and **M3**; **Figure 3A** – Gray), with frequency of detection of these ions near 0% for unlabeled PSMs—for samples that had been enriched on neutravidin resin (**Figure 3A**), the unlabeled peptide subset constitutes nonspecific, unmodified peptides carried through the neutravidin enrichment. We observe more moderate, albeit still statistically significant, specificity for all fragment ions (**F1**, **F3**, **F4,** and **F5**; **Figure 3A** – Gray), with detection frequencies of ~50% for PSMs lacking the biotin mass shift compared with almost 100% for biotin-modified PSMs. In contrast, the **F2** ion was identified close to 100% both modified and unmodified PSMs. The general modest to low specificity observed for these fragment ions can likely be rationalized by co-isolation of the modified and unmodified cysteine peptides as precursor ions.[45]

We expected that increased specificity of fragment ions for CBCC modified peptides could be achieved by decreasing the proportion of CBCC-labeled peptides in the MS sample, thereby reducing co-isolation of labeled and unlabeled precursors. Therefore, we next extended fragment ion frequency analysis to unenriched (pre-neutravidin enrichment) trypsin digested CBCC-labeled lysates (**Figure 3B**). An increase in specificity for all fragment ions was observed
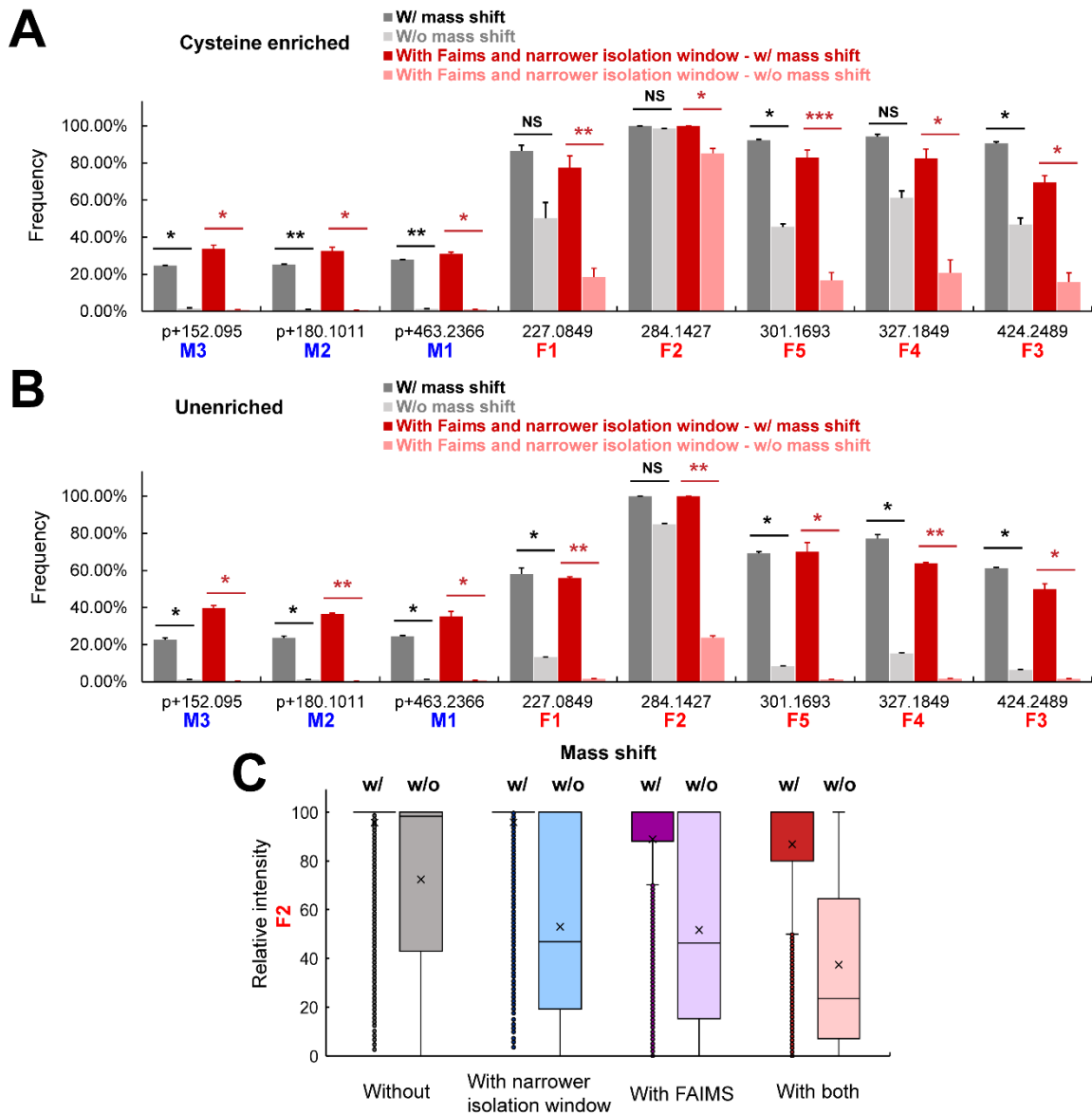
Figure 3. Frequency distribution of fragment and peptide remainder ions in PSMs with or without cysteine modifications in (A) neutravidin enriched samples or (B) unenriched samples with narrower isolation window and FAIMS or without. (C) Relative intensity of oxonium-biotin in PSMs with or without cysteine modifications with narrower isolation window, with FAIMS, with both or without. Samples were enriched on neutravidin resin prior to analysis. Statistical significance was calculated with unpaired Student's t-tests with equal variance, * $p<0.05$, ** $p<0.01$, *** $p<0.001$, NS $p>0.05$. Experiments were performed in duplicate. All data can be found in **Table S3**.

for the unenriched samples, as indicated by the decrease in detection frequency compared with the neutravidin enriched samples shown in **Figure 3A**. Surprisingly, even in these unenriched samples, the specificity of the **F2** ion remained modest, with a frequency of detection in unlabeled peptide PSMs of >80%.

Given the observed high intensity of the oxonium-biotin **F2** fragment ion for PSMs lacking a biotin modification, both for neutravidin-enriched and unenriched samples, we next sought to assess whether specificity of fragment ions, particularly the **F2** fragment ion, could be improved by modifi-

cations to our data acquisition method. On-line sample fractionation using a FAIMS ion mobility device has been found to increase sample quantitation at the MS2 level, affording decreased ratio compression for TMT samples.[46] Therefore, we next applied FAIMS ion mobility to CBCC labeled samples and, as described above, compared the frequency distribution of each ion for biotinylated vs non-biotinylated precursor ions. We find that the use of FAIMS improves the specificity of all fragment ions (**Figure S6**). Similarly, the use of a narrower precursor isolation window (0.5 Da compared with our standard isolation window 1.6 Da) also afforded improved specificity with no decrease in frequency distribution of PSMs.

Scheme 2. Major fragmentation anticipated in different reagent combinations. (A) IAA + Biotin-azide. (B) IAA + Biotin-D-azide. (C) PIAA + Biotin-azide. (D) IAAz + Biotin-alkyne. (E) IAA + Biotin-C4-Azide. Blue labeled peptide remainder modification of the biotinylated cysteines and red labeled fragment ions.
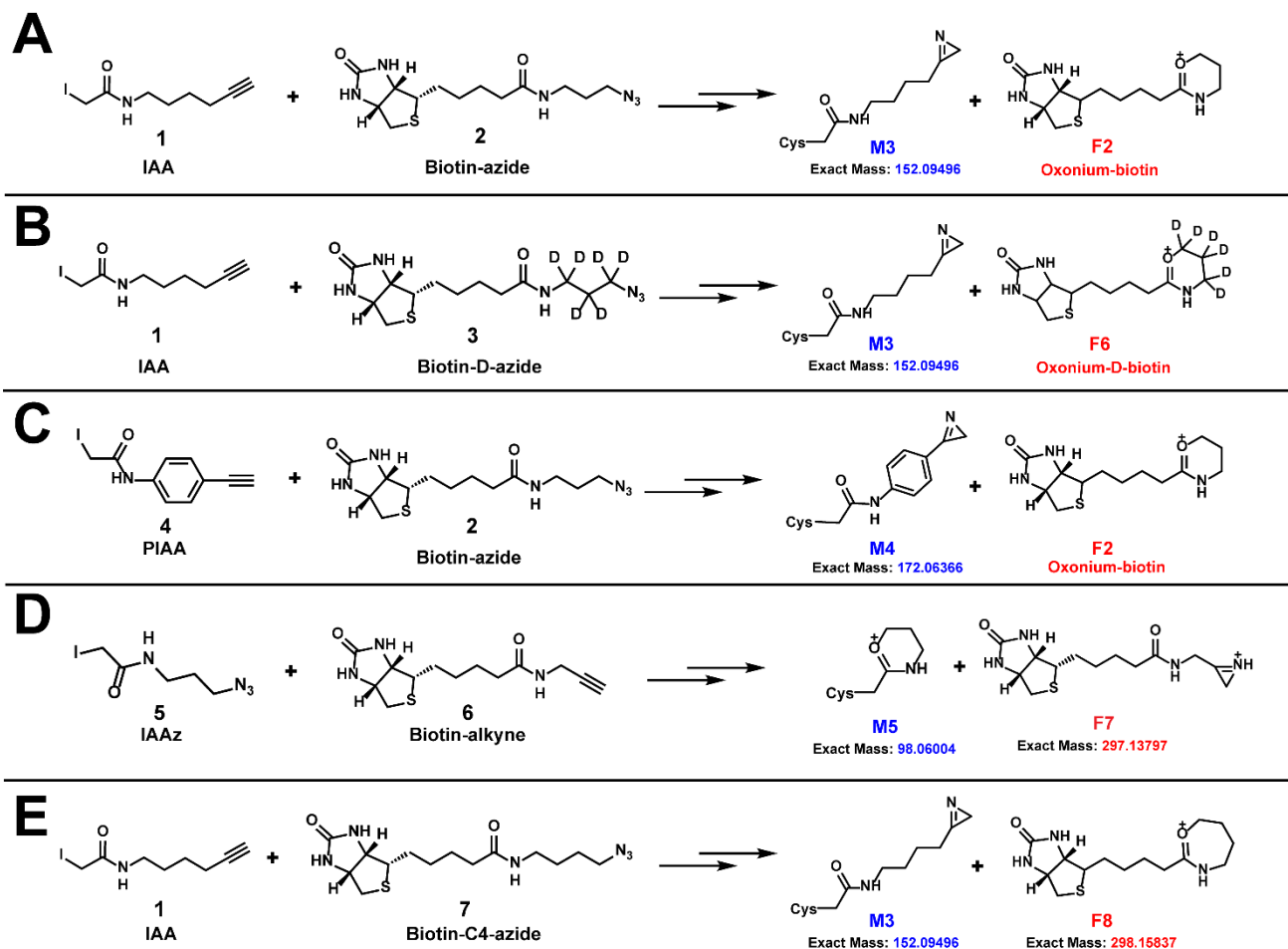
By combining both FAIMS and the narrower precursor isolation window the frequency of detection of fragment ions **F1**, **F3**, **F4**, and **F5** dropped to <20% for neutravidin enriched samples (**Figure 3A**– Red) and to close to 0% for unenriched samples (**Figure 3B**). For the **F2** oxonium-biotin, while the frequency of detection for unlabeled peptides remained relatively high (~80% for neutravidin enriched samples and ~20% for unenriched samples, **Figure 3A, B**, respectively), median ion intensity analysis (**Figure 3C**) revealed that addition of FAIMS, the narrow precursor isolation window, and, most significantly, the combination of both FAIMS and the narrower precursor isolation window reduced the median intensity of the F1 ion to 20%. Collectively, this analysis points to the potential utility of the F1 ion as a characteristic ion for CBCC peptide identification, with particular relevance in the analysis of unenriched samples using both a FAIMS device and narrow precursor isolation window. With implementation of an intensity threshold cutoff of ~40%, we expect that this ion will also prove diagnostic of biotinylation state for neutravidin enriched samples.

Unmatched fragment ions in MS/MS spectra can sometimes be rationalized as stemming from contaminating ions. To rule out such ions, we expanded our analysis to include additional labeling reagent combinations (**Scheme 2 and Table 2**), with the expectation that each reagent pair would afford a unique fragment profile. CBCC samples were prepared and analyzed using three reagent combinations **A** (IAA **1** + Biotin-azide **2**), **B** (IAA **1** + Biotin-D-azide **3**) and **C** (phenyl-iodoacetamide alkyne (PIAA) **4** + biotin-azide **2**)[47] (**Scheme 2**). Labile ion search revealed that the observed frequency distribution of peptide remainder and fragment ions for each reagent combination matched with the expected fragmentation pattern (**Table 2**, **Scheme 2**). The **F1** ion was detected for both reagent combinations prepared using biotin-azide reagent **2**. For the reagent combination **B**, the **F2** ion was only observed in 1.22% of all PSMs and instead the **F2** + 6.038 ion (m/z 290.180, Oxonium-D-biotin) was detected in 100% of PSMs. Similarly, both reagent combinations **A** and **B** afforded the expected p+152.095 peptide remainder ion with a detection frequency of 25% and 29%, respectively. For combination **C**, the p+152.095 ion was only detected in 1% of all scans, and instead peptide remainder ion p+172.064 was detected in 42% of all PSMs,

Table 2. Frequency distribution of signature fragment ions and peptide remainder ions using different labeling reagent combinations to generate CBCC modified PSMs.

	Reagent Combination	m/z							
		p+152.095 (IAA)	p+172.064 (PIAA)	p+98.060 (IAAz)	284.143 (Biotin-azide)	290.180 (Biotin-D-azide)	298.158 (Biotin-C4-azide)	297.138 (Biotin-alkyne)	227.085
A	IAA + Biotin-azide	25.44	1.42	1.70	100.00	0.02	0.60	0.06	89.60
B	IAA + Biotin-D-azide	30.75	1.85	2.36	1.22	100.00	0.67	0.09	76.34
C	PIAA + Biotin-azide	1.03	42.52	1.33	100.00	0.03	0.44	0.15	97.48
D	IAAz + Biotin-alkyne	4.84	1.78	22.05	1.50	0.12	0.60	12.89	88.03
E	IAA + Biotin-C4-azide	23.20	1.83	1.75	2.11	0.00	99.66	0.04	88.47

which matches the phenyl azirine **M4** resulting from fragmentation of the N(triazole)–C(alkyl) bond of the click product.

Investigation of fragmentation of CBCC peptides labeled with different reagent combinations

Given the seeming exquisite selectivity of the identified fragment ions for each reagent combination tested as shown in **Table 2**, we next expanded our panel of reagents to include two additional reagent pairs with the goal of further probing the nature of the fragment ions and determining whether structural modifications to the reagents would impact the frequency, intensity, or identify of observed fragments. We were optimistic that a thorough investigation of reagent fragmentation might point to avenues for achieving increased coverage of CBCC peptides.

We hypothesized that the length of the C3 alkyl chain and corresponding stability of the six-member ring in the **F2** oxonium-biotin cyclized product was likely contributing to the observed high intensity and low specificity of the fragment ion. We synthesized two additional reagents, Iodoacetamide-azide (IAAz) **5** and Biotin-C4-azide **7** that together would allow us to further probe the impact of changes to linker length and triazole orientation on fragmentation. We then subjected these reagents combination **D** (IAAz **5** + Biotin-alkyne **6**) and **E** (IAA **1** + Biotin-C4-azide **7**) to chemoproteomic analysis, with labile ion search for predicted fragmentation products (**Scheme 2**). As noted for the first three reagent combinations, we observed high selectivity for the ions produced by each reagent combination with <2.2% detection of ions not predicted to be generated by each reagent combination (**Table 2**). We were pleased to observe the formation of p+98.060 peptide remainder ion and m/z 297.138 fragment ion in reagent combination **D**, which are consistent with our predicted reagent fragmentation to form 1,3-oxazonium **M5** peptide remainder ion and biotin-azirinium **F7** fragment ion, detected in 22% and 13% of all labeled PSMs, respectively. Similarly, extension of the alkyl linker by one carbon in reagent combination **E** afforded production of m/z 297.138 fragment ion, which is

consistent with the structure of **F8** oxonium-C4-biotin. Given the general instability of primary carbocations, we expect that this ion also cyclizes as proposed for the **F2** fragment ion, although we cannot exclude that the decreased favorability of seven-member ring formation might preclude cyclization. Together these ions provide further evidence of the gas phase lability of triazole modifications, including that this fragmentation is generalizable to multiple combinations of reagents of varying linker lengths and triazole configurations.

We were intrigued by the striking decreased detection rate of oxonium species formation with the reagent combination **D** when compared with our standard combination **A** (compare 100% for **F2** species with 22% for **M5**). We find that these reagent differences extend to the frequency of ion detection and relative intensities of observed ions (**Figure 4A, B**). For CBCC PSMs enriched with **D**, we observe a marked decrease in the frequency and intensity of both peptide remainder ion **M5** (p+98.060) and fragment ion **F7** (m/z 297.138), when compared with the intensities observed for peptides labeled with **A** (**Figure 3A, B**). We suspect that the swapped azide-alkyne configuration results in less favorable fragmentation of the N(triazole)–C(alkyl) bond and/or decreased gas phase stability of the afforded product ions.

In stark contrast with the analysis of **D**, we observed for **C** that the frequency of detection of the peptide remainder ion **M4** (p+172.064) increased to 40% (**Figure 4A**), compared with 20% for the corresponding p+152.095 observed for reagent pair **A** (**Figure 3A**). We hypothesized that the pi conjugation of the aromatic PIAA probe likely stabilizes the proposed azirine species, thereby increasing the favorability of fragmentation. The frequency of detection of the dehydrobiotin and oxonium-biotin species was observed to be comparable for both aryl (PIAA **4**, **Figure 4A**) and alkyl (IAA **1**, **Figure 3A**) probes. Similar to our standard reagent combination **A**, low specificity was observed for the **F1** and **F2** ions, with detection in most PSMs that lacked the biotin modification. The biotin-C4-azide **7** labeling experiment afforded improved selectivity when compared to the biotin-azide **2** labeled samples, which was expected and consistent
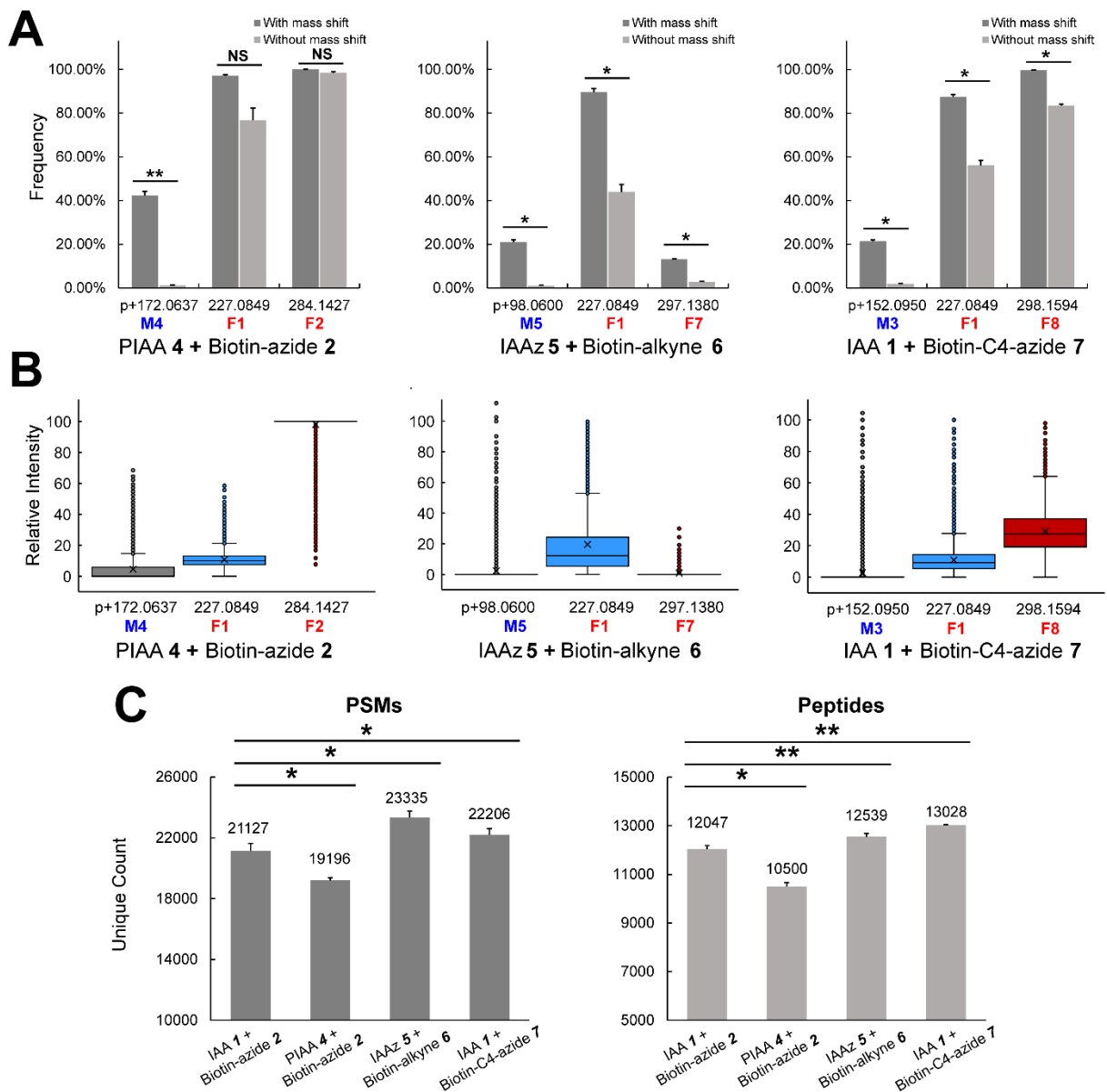
Figure 4. (A) Frequency distribution of signature fragment ions and peptide remainder ions, (B) Relative intensity of signature ions, and (C) Peptides and PSMs identified in CBCC with different labeling reagents. Statistical significance was calculated with unpaired Student's t-tests with equal variance, * $p<0.05$, ** $p<0.01$, NS $P>0.05$. Experiments were performed in duplicate. All data can be found in **Table S4**.

with the less favorable fragmentation of the longer alkyl linker. The improved selectivity was indicated by the decreased frequency of detection of signature ions **F1**, **F8**, and **M3** in PSMs without biotin modifications. Notably, the specificity of **F8** for peptide biotinylation status increased dramatically with the use of FAIMS and a narrower precursor isolation window (**Figure S7**), consistent with our aforementioned observation that these acquisition conditions dramatically impact the frequency of detection of signature fragment ions in PSMs lacking biotin.

Quantification of the median relative ion intensities (**Figure 4B**) supports the likely different favorability of frag-

mentation and/or differences in stability of cognate fragment ions derived from CBCC peptides labeled by reagent pairs **C**, **D** and **E**. While the intensities of the dehydrobiotin fragment ion **F1** were comparable for all three reagent combinations, the **F2, F7,** and **F8** ions differed dramatically with near zero relative intensity for **F7**, approximately 30% relative intensity for **F8** and close to 100% for **F2**. Further supporting increased fragmentation of labeling combination **C**, the **M4** (p+172.064) median ion intensity was 4.46% compared with 0% for both the **M3** and **M5** peptide remainder ions.

Curious as to how fragmentation of CBCC affects the cysteine chemoproteomic coverage during the MS/MS analysis, we quantified the mean PSMs and unique CBCC peptides for
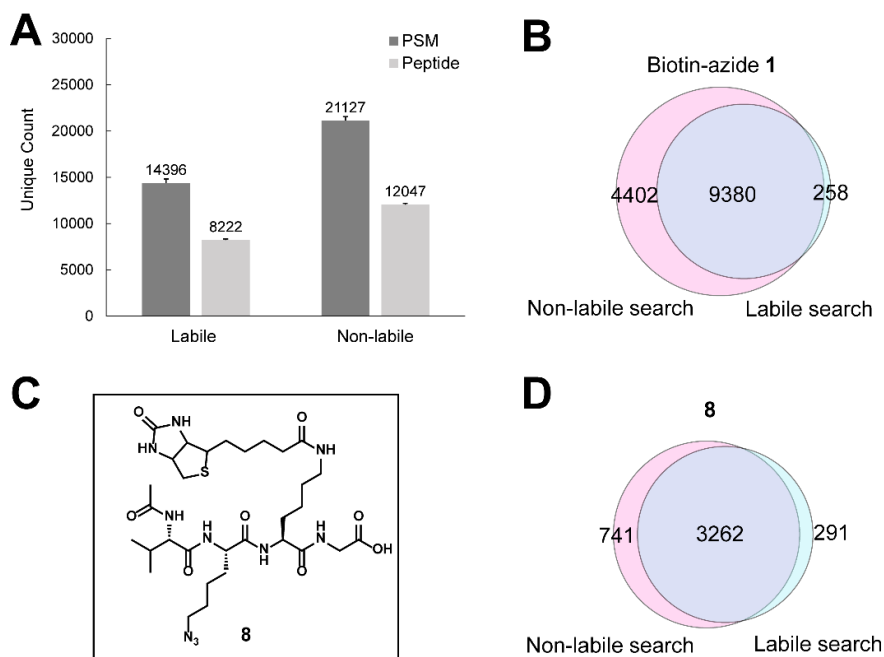
Figure 5. (A) Cysteine peptides and PSMs identified with labile and non-labile search in samples labeled with IAA **1** and Biotin-azide **2**. (B) Venn diagram of cysteine peptides identified with labile and non-labile search in samples prepared with IAA **1** clicked with biotin-azide **2** and (D) N-Ac-Val-Azido-Lys-Lys(biotin)-Gly-COOH tetrapeptide **8**. (C) Structure of N-Ac-Val-Azido-Lys-Lys(biotin)-Gly-COOH tetrapeptide reagent **8**. Experiments were performed in duplicate. All data can be found in **Table S5**.

each reagent combination. Consistent with a model where increased formation of these diagnostic fragment ions can decrease coverage, we observed that both reagent combinations D and E afforded a ~9% increase in PSMs and a smaller, albeit still significant increase in unique peptides (**Figure 4C**) when compared with the standard reagent pair **A**. In contrast, fragmentation-prone combination **C** decreased the PSMs by ~8% and the unique peptides by a more substantial ~13%. We also extended this analysis to include the fraction of peptides that were uniquely identified by one reagent combination. We were surprised to observe relatively modest overlap between the combinations tested (**Figure S8**). While we acknowledge the semi-stochastic nature of data-dependent acquisition (DDA) proteomic data acquisition and resulting problem of "missing" data may in part rationalize some of these differences, collectively we expect that these data point to the possible utility of varying CBCC labeling reagents as a strategy to improve detection of certain low abundance or otherwise tough-to-detect peptides.

Application of labile search options in MSFragger to enhance cysteine chemoproteomic coverage

With our successful identification of a panel of labile ions derived from the CBCC modifications using the labile MSFragger search options, we next asked whether such la-

bile ion searches would also increase coverage of CBCC peptides. Due to the high abundance of the oxonium-biotin species in all MS2 scans, we initially expected that, as has been reported for analysis of glycopeptide datasets[33], such search would similarly boost coverage for CBCC. Curiously, we did not observe such an increase (**Figure 5A**), and instead found that labile search consistently underperformed, with only ~66% the number of PSMs and ~75% the number of peptides compared with our established closed search method. We suspect that this decrease may stem from the relative gas-phase stability of many of the ions detected, despite the observation of intense diagnostic ions. Partial fragmentation of the modification on multiple peptide fragment peaks can be concentrated into a single diagnostic peak, resulting in high relative intensity despite much of the modification remaining intact. Comparison of closed vs labile ion searches for two aggregate replicates revealed that nearly all cysteines identified in the labile search are share with the closed search, with only 2.7% unique cysteines identified for reagent combination **A** (**Figure 5B**).

Motivated by the high performance of labile search for glycoproteomics, we wanted to further investigate whether such searches could prove beneficial for chemoproteomics studies using more fragmentation prone reagents. We synthesized N-Ac-Val-Azido-Lys-Lys(biotin)-Gly-COOH tetrapeptide reagent **8** (**Figure 5C**). We expected that the increased size (861.453 Da modification) and peptidic back-

bone of this reagent would afford increased gas-phase fragmentation compared with the more minimalist IAA **1** + Biotin-azide **2** (463.237 Da) CBCC labeling protocol. We subjected proteome to labeling with IAA **1 + 8**. LC-MS/MS and non-labile search identified 10710 PSMs and 4003 unique CBCC peptides across two replicate experiments **(Table S5)**, which represents substantially reduced coverage when compared to the aforementioned labeling with IAA **1** + Biotin-azide **2**.

We observed a number of characteristic fragment ions produced by the reagent **8** (proposed fragmentation is shown in **Figure S9**), including known and novel species. In addition to formation of dehydrobiotin, we also detected a m/z 310.158 (**F9**) fragment ion, which we assigned as Im-Kbiotin-NH3,[37] as characterized previously for biotinylated peptides. Additional we detected a fragment ion of m/z 430.21187 (**F10**), that we assigned as likely corresponding to the Lys(biotin)-Gly dipeptide and fragment ion m/z 682.35926 (**F11**), corresponding to the teatrapeptide formed from C-N triazole bond fragmentation. Together with these fragment ions we also identified peptide remainder ions p+449.275 (**M7**), p+433.256 (**M8**), p+180.101 (**M2**), p+152.095 (**M3**), which match with the proposed fragmentation scheme. As a demonstration of the complexity of the fragmentation pattern observed for this larger biotin reagent, we also identified several ions that appeared diagnostic of peptide labeling state as indicated by presences in labeled PSMs but did not readily match with any fragmentation pattern tested, including m/z 243.08, m/z 412.20, m/z 464.26, and m/z 491.14. The ubiquity and intensity of these ions together with the reduced coverage of CBCC peptides supports the general dogma of the field that smaller labeling reagents are likely preferable when compared to bulkier tags.

To determine whether labile search would increase coverage for this bulkier reagent, we subjected the **8**-labeled samples to labile search analysis, following a similar workflow to that shown in **Figure 2A**. Using all fragment and peptide remainder ions shown in **Figure S9**, labile search of these samples afforded increased coverage, identifying ~8.2% unique cysteines (**Figure 5D**). While these gains remain modest, this finding points to the possible utility of labile search in analysis of more complex chemoproteomic labeling reagents, particularly for those with substituents prone to fragmentation.

Discussion

In this study, we performed a detailed characterization of the gas-phase fragmentation products generated from cysteine biotinylation via click chemistry (CBCC) chemoproteomic analysis. Manual inspection of CBCC MS/MS spectra first revealed two fragment ions, the well characterized dehydrobiotin (m/z 227.085) and a new species m/z 284.143. Using a biotin isotopologue alongside MS3 analysis, we then interrogated the nature of the m/z 284.143 fragment ion and propose it to be the oxonium-biotin species **F2**. We propose that the oxonium biotin forms as a result of fragmentation at the N(triazole)–C(alkyl) bond together with cyclization to afford cyclic oxonium species. This finding goes against the generally held assumption that triazoles are relatively inert to gas phase chemistry. There is some uncertainty surrounding the formal nature of the m/z 284.143 fragment ion as either cyclic oxazonium or linear carbocation. However, the general instability of primary carbocations together with the favorability of the six-member ring formation supports the oxonium-biotin structure as the most likely nature of the fragment ion.

We next leveraged the labile search features of the MSFragger algorithm together, and extended the PTM-Shepherd module, to identify additional fragment and peptide remainder ions associated with the CBCC labeled peptides. Consistent with our proposed fragmentation of the N6–C5 bond in the clicked conjugate, we also identified peptide remainder ions that correspond to peptides retaining the triazole **M2** and together with the corresponding azirine species **M3**, afforded by loss of nitrogen. Analysis of the specificity of these ions for the CBCC peptide modification state revealed a striking difference between the fragment ions and the peptide remainder ions, with all peptide remainder ions only identified in PSMs for biotinylated peptides. Overall, peptide remainder ions were identified at a lower frequency and with lower relative intensity compared to the fragment ions.

A narrower isolation window minimizes the co-isolation of modified and unmodified peptides, which together with the use of FAIMS, successfully improved the specificity of all fragment ions, including oxonium-biotin and dehydrobiotin. The improvement in specificity was observed particularly for unenriched samples, which contain lower abundance of the biotinylated species. When compared to the oxonium-biotin species, several of the other peptide remainder (**M2**, **M3**) and fragment ions (**F1**, **F3**, **F4**, **F5**) showed considerably greater specificity for the peptide biotinylation state. However, given the moderate to low detection frequency of such ions, they are likely suboptimal as diagnostic species. Distinguished by its high intensity and ubiquity, we expect that with an appropriate intensity threshold, the oxonium-biotin species could serve as a diagnostic ion for CBCC peptides.

Extension of these analyses to four additional reagent pairs revealed several striking results. First, we found that the production of the oxonium-biotin species was not restricted our IAA probe and instead was produced by both alkyl and aryl iodoacetamide probes, with apparent increased favorability of fragmentation for the aryl probe likely due to the pi conjugation stabilizing the peptide remainder ion. The ion's generalizability points to its utility for a wide variety of clickable probes, including more advanced drug-like scaffolds. Our reagent panel also revealed that a ~10% increase in PSMs could be achieved by small and easily implementable modifications to the reagents themselves. For example, the use of iodoacetamide azide reagent IAAz **5** and biotin-alkyne **6**, dramatically decreased the intensity of detected fragments and simultaneously afforded an increase in PSMs. Similarly, we also observed that extension of the alkyl linker by one carbon in the biotin-C4-azide reagent also afforded an increase in PSMs together

with a decrease in the intensity of the corresponding fragment ion that is likely due to the decreased favorability of cyclization of the seven-member ring.

Taken together, our comprehensive analysis of CBCC reagents fragmentation not only sheds light on the fundamental behavior of these labels in the gas phase, but also allows us to put forth several general principles for the choice of biotinylation reagents. For traditional closed searches, reagents should be chosen judiciously to disfavour fragmentation thereby improving coverage, for example, with the biotin-C4-azide or iodoacetamide azide reagent IAAz. For larger and/or more labile reagent combinations, labile search option can match 10% additional spectra compared to the closed search, as demonstrated by reagent **8.** As shown here, the labile options in FragPipe are well equipped for case-by-case analysis of peptide biotinylation reagent combinations. Therefore, we recommend that the chemoproteomics community consider adoption of such labile ion analyses both as a tool to improve the confidence in labeling site identification and to expand coverage of labeled peptides, particularly in those cases where labeled peptides are lowly abundant.

More broadly, fragment ions, such as those identified here, have potential use in quantification of labeled sites at the MS2 and MS3 levels, as showcased by established isobaric labeling methods (e.g. TMT and ITRAQ). The design of new isobaric tags remains challenging, given the lack of generalizable rules for patterns of fragmentation and generation of reporter ions in the gas phase. Looking ahead, as we piloted here with biotin-azide and biotin-D-azide, we imagine quantification will easily be achieved based on the intensity of oxonium biotin m/z 284.143 and its matched isotopologue m/z 290.180 with a 6-dalton mass shift. The limitation of retention time shifts caused by deuterium could be addressed through the strategic incorporation of deuterated balancer moieties.

Alongside providing an improved fundamental understanding of how chemoproteomics samples behave in the gas phase, our study offers several added benefits. First, we present the low cost synthesis of an isotopically labeled pair of biotin-azide reagents, which compares favorably to the cost and complexity of established isotopically labeled reagents, both azide-containing and those that feature cysteine-reactive electrophiles.[11,13,48,49] Our demonstration of the labile ion search features built into FragPipe should also provide a generalizable computational platform for others interested in leveraging fragmentation of chemoproteomics samples. Exemplifying the utility of such studies, gas phase fragmentation of cysteines modified by covalent drugs, such as ibrutinib, have been leveraged to improve the identification of labeled cysteine residues.[50,51] In addition, fragmentation of sulfonyl-triazole probes has been harnessed for site-of-labeling studies.[52] When combined with custom isobaric data analysis algorithms, these modifications should also provide an avenue to improve MS2-level quantification of peptide labeling. We anticipate that realizing the full potential of labile search algorithms in chemoproteomic applications may depend on advances in these algorithms in order to take full advantage of partial fragmentation, such as that observed for the CBCC reagents described here.

ASSOCIATED CONTENT

Supporting Information

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

Detailed methods of chemical synthesis and chemoproteomic sample preparation; Supplementary figures and tables (PDF)

Aggregated datasets for each result section (xlsx)

Sample workflow for MSFragger search (workflow)

AUTHOR INFORMATION

Corresponding Author

* Keriann M. Backus, Biological Chemistry Department, David Geffen School of Medicine, UCLA, Los Angeles, CA, 90095, USA, E-mail: kbackus@mednet.ucla.edu.

Note

The authors declare no competing financial interest.

ACKNOWLEDGMENT

REFERENCES

(1)   Cheng, K.; Lee, J. S.; Hao, P.; Yao, S. Q.; Ding, K.; Li, Z. Tetrazole-Based Probes for Integrated Phenotypic Screening, Affinity-Based Proteome Profiling, and Sensitive Detection of a Cancer Biomarker. *Angew. Chemie - Int. Ed.* **2017**, *56* (47), 15044–15048. https://doi.org/10.1002/anie.201709584.

(2)   Drewes, G.; Knapp, S. Chemoproteomics and Chemical Probes for Target Discovery. *Trends Biotechnol.* **2018**, *36* (12), 1275–1286. https://doi.org/10.1016/j.tibtech.2018.06.008.

(3)   Friedman Ohana, R.; Kirkland, T. A.; Woodroofe, C. C.; Levin, S.; Uyeda, H. T.; Otto, P.; Hurst, R.; Robers, M. B.; Zimmerman, K.; Encell, L. P.; Wood, K. V. Deciphering the Cellular Targets of Bioactive Compounds Using a Chloroalkane Capture Tag. *ACS Chem. Biol.* **2015**, *10* (10), 2316–2324. https://doi.org/10.1021/acschembio.5b00351.

(4)   Moellering, R. E.; Cravatt, B. F. How Chemoproteomics Can Enable Drug Discovery and Development. *Chem. Biol.* **2012**, *19* (1), 11–22. https://doi.org/10.1016/j.chembiol.2012.01.001.

(5)   Martell, J.; Weerapana, E. Applications of Copper-Catalyzed Click Chemistry in Activity-Based Protein Profiling. *Molecules* **2014**, *19* (2), 1378–1393. https://doi.org/10.3390/molecules19021378.

(6)   Parker, C. G.; Pratt, M. R. Click Chemistry in Proteomic Investigations. *Cell* **2020**, *180* (4), 605–632. https://doi.org/10.1016/j.cell.2020.01.025.

(7)   Speers, A. E.; Cravatt, B. F. Activity-Based Protein Profiling (ABPP) and Click Chemistry (CC)–ABPP by MudPIT Mass

Spectrometry. *Curr. Protoc. Chem. Biol.* **2009**, *1* (1), 29–41. https://doi.org/10.1002/9780470559277.ch090138.

(8) Chan, W. C.; Sharifzadeh, S.; Buhrlage, S. J.; Marto, J. A. Chemoproteomic Methods for Covalent Drug Discovery. *Chem. Soc. Rev.* **2021**, *50* (15), 8361–8381. https://doi.org/10.1039/D1CS00231G.

(9) Kuljanin, M.; Mitchell, D. C.; Schweppe, D. K.; Gikandi, A. S.; Nusinow, D. P.; Bulloch, N. J.; Vinogradova, E. V.; Wilson, D. L.; Kool, E. T.; Mancias, J. D.; Cravatt, B. F.; Gygi, S. P. Reimagining High-Throughput Profiling of Reactive Cysteines for Cell-Based Screening of Large Electrophile Libraries. *Nat. Biotechnol.* **2021**, *39* (5), 630–641. https://doi.org/10.1038/s41587-020-00778-3.

(10) Go, Y.-M.; Chandler, J. D.; Jones, D. P. The Cysteine Proteome. *Free Radic. Biol. Med.* **2015**, *84*, 227–245. https://doi.org/10.1016/j.freeradbiomed.2015.03.022.

(11) Backus, K. M.; Correia, B. E.; Lum, K. M.; Forli, S.; Horning, B. D.; González-Páez, G. E.; Chatterjee, S.; Lanning, B. R.; Teijaro, J. R.; Olson, A. J.; Wolan, D. W.; Cravatt, B. F. Proteome-Wide Covalent Ligand Discovery in Native Biological Systems. *Nature* **2016**, *534* (7608), 570–574. https://doi.org/10.1038/nature18002.

(12) Bar-Peled, L.; Kemper, E. K.; Suciu, R. M.; Vinogradova, E. V.; Backus, K. M.; Horning, B. D.; Paul, T. A.; Ichu, T. A.; Svensson, R. U.; Olucha, J.; Chang, M. W.; Kok, B. P.; Zhu, Z.; Ihle, N. T.; Dix, M. M.; Jiang, P.; Hayward, M. M.; Saez, E.; Shaw, R. J.; Cravatt, B. F. Chemical Proteomics Identifies Druggable Vulnerabilities in a Genetically Defined Cancer. *Cell* **2017**, *171* (3), 696-709.e23. https://doi.org/10.1016/j.cell.2017.08.051.

(13) Abo, M.; Li, C.; Weerapana, E. Isotopically-Labeled Iodoacetamide-Alkyne Probes for Quantitative Cysteine-Reactivity Profiling. *Mol. Pharm.* **2018**, *15* (3), 743–749. https://doi.org/10.1021/acs.molpharmaceut.7b00832.

(14) Weerapana, E.; Wang, C.; Simon, G. M.; Richter, F.; Khare, S.; Dillon, M. B. D.; Bachovchin, D. A.; Mowen, K.; Baker, D.; Cravatt, B. F. Quantitative Reactivity Profiling Predicts Functional Cysteines in Proteomes. *Nature* **2010**, *468* (7325), 790–797. https://doi.org/10.1038/nature09472.

(15) Yan, T.; Desai, H. S.; Boatner, L. M.; Yen, S. L.; Cao, J.; Palafox, M. F.; Jami-Alahmadi, Y.; Backus, K. M. SP3-FAIMS Chemoproteomics for High-Coverage Profiling of the Human Cysteinome**. *ChemBioChem* **2021**, *22* (10), 1841–1851. https://doi.org/10.1002/cbic.202000870.

(16) Chowdhury, S. M.; Du, X.; Tolić, N.; Wu, S.; Moore, R. J.; Mayer, M. U.; Smith, R. D.; Adkins, J. N. Identification of Cross-Linked Peptides after Click-Based Enrichment Using Sequential Collision-Induced Dissociation and Electron Transfer Dissociation Tandem Mass Spectrometry. *Anal. Chem.* **2009**, *81* (13), 5524–5532. https://doi.org/10.1021/ac900853k.

(17) Lee, S. S.; Lim, J.; Tan, S.; Cha, J.; Yeo, S. Y.; Agnew, H. D.; Heath, J. R. Accurate MALDI-TOF/TOF Sequencing of One-Bead−One-Compound Peptide Libraries with Application to the Identification of Multiligand Protein Affinity Agents Using in Situ Click Chemistry Screening. *Anal. Chem.* **2010**, *82* (2), 672–679. https://doi.org/10.1021/ac902195y.

(18) Wang, Z.; Udeshi, N. D.; O'Malley, M.; Shabanowitz, J.; Hunt, D. F.; Hart, G. W. Enrichment and Site Mapping of O-Linked N-Acetylglucosamine by a Combination of Chemical/Enzymatic Tagging, Photochemical Cleavage, and Electron Transfer Dissociation Mass Spectrometry. *Mol. Cell. Proteomics* **2010**, *9* (1), 153–160. https://doi.org/10.1074/mcp.M900268-MCP200.

(19) Sohn, C. H.; Agnew, H. D.; Lee, J. E.; Sweredoski, M. J.; Graham, R. L. J.; Smith, G. T.; Hess, S.; Czerwieniec, G.; Loo, J. A.; Heath, J. R.; Deshaies, R. J.; Beauchamp, J. L. Designer Reagents for Mass Spectrometry-Based Proteomics: Clickable Cross-Linkers for Elucidation of Protein Structures and Interactions. *Anal. Chem.* **2012**, *84* (6), 2662–2669. https://doi.org/10.1021/ac202637n.

(20) Sohn, C. H.; Lee, J. E.; Sweredoski, M. J.; Graham, R. L. J.; Smith, G. T.; Hess, S.; Czerwieniec, G.; Loo, J. A.; Deshaies, R. J.;

Beauchamp, J. L. Click Chemistry Facilitates Formation of Reporter Ions and Simplified Synthesis of Amine-Reactive Multiplexed Isobaric Tags for Protein Quantification. *J. Am. Chem. Soc.* **2012**, *134* (5), 2672–2680. https://doi.org/10.1021/ja2099003.

(21) Udeshi, N. D.; Pedram, K.; Svinkina, T.; Fereshetian, S.; Myers, S. A.; Aygun, O.; Krug, K.; Clauser, K.; Ryan, D.; Ast, T.; Mootha, V. K.; Ting, A. Y.; Carr, S. A. Antibodies to Biotin Enable Large-Scale Detection of Biotinylation Sites on Proteins. *Nat. Methods* **2017**, *14* (12), 1167–1170. https://doi.org/10.1038/nmeth.4465.

(22) Borisov, O. V.; Goshe, M. B.; Conrads, T. P.; Rakov, V. S.; Veenstra, T. D.; Smith, R. D. Low-Energy Collision-Induced Dissociation Fragmentation Analysis of Cysteinyl-Modified Peptides. *Anal. Chem.* **2002**, *74* (10), 2284–2292. https://doi.org/10.1021/ac010974p.

(23) Liu, Y.; Patricelli, M. P.; Cravatt, B. F. Activity-Based Protein Profiling: The Serine Hydrolases. *Proc. Natl. Acad. Sci.* **1999**, *96* (26), 14694–14699. https://doi.org/10.1073/pnas.96.26.14694.

(24) Kim, D. I.; Cutler, J. A.; Na, C. H.; Reckel, S.; Renuse, S.; Madugundu, A. K.; Tahir, R.; Goldschmidt, H. L.; Reddy, K. L.; Huganir, R. L.; Wu, X.; Zachara, N. E.; Hantschel, O.; Pandey, A. BioSITe: A Method for Direct Detection and Quantitation of Site-Specific Biotinylation. *J. Proteome Res.* **2018**, *17* (2), 759–769. https://doi.org/10.1021/acs.jproteome.7b00775.

(25) Schopfer, L. M.; Champion, M. M.; Tamblyn, N.; Thompson, C. M.; Lockridge, O. Characteristic Mass Spectral Fragments of the Organophosphorus Agent FP-Biotin and FP-Biotinylated Peptides from Trypsin and Bovine Albumin (Tyr410). *Anal. Biochem.* **2005**, *345* (1), 122–132. https://doi.org/10.1016/j.ab.2005.07.016.

(26) Eng, J. K.; McCormack, A. L.; Yates, J. R. An Approach to Correlate Tandem Mass Spectral Data of Peptides with Amino Acid Sequences in a Protein Database. *J. Am. Soc. Mass Spectrom.* **1994**, *5* (11), 976–989. https://doi.org/10.1016/1044-0305(94)80016-2.

(27) Xu, T.; Park, S. K.; Venable, J. D.; Wohlschlegel, J. A.; Diedrich, J. K.; Cociorva, D.; Lu, B.; Liao, L.; Hewel, J.; Han, X.; Wong, C. C. L.; Fonslow, B.; Delahunty, C.; Gao, Y.; Shah, H.; Yates, J. R. ProLuCID: An Improved SEQUEST-like Algorithm with Enhanced Sensitivity and Specificity. *J. Proteomics* **2015**, *129*, 16–24. https://doi.org/10.1016/j.jprot.2015.07.001.

(28) Eng, J. K.; Jahan, T. A.; Hoopmann, M. R. Comet: An Open-Source MS/MS Sequence Database Search Tool. *Proteomics* **2013**, *13* (1), 22–24. https://doi.org/10.1002/pmic.201200439.

(29) Chick, J. M.; Kolippakkam, D.; Nusinow, D. P.; Zhai, B.; Rad, R.; Huttlin, E. L.; Gygi, S. P. A Mass-Tolerant Database Search Identifies a Large Proportion of Unassigned Spectra in Shotgun Proteomics as Modified Peptides. *Nat. Biotechnol.* **2015**, *33* (7), 743–749. https://doi.org/10.1038/nbt.3267.

(30) Yu, F.; Teo, G. C.; Kong, A. T.; Haynes, S. E.; Avtonomov, D. M.; Geiszler, D. J.; Nesvizhskii, A. I. Identification of Modified Peptides Using Localization-Aware Open Search. *Nat. Commun.* **2020**, *11* (1), 4065. https://doi.org/10.1038/s41467-020-17921-y.

(31) Swearingen, K. E.; Eng, J. K.; Shteynberg, D.; Vigdorovich, V.; Springer, T. A.; Mendoza, L.; Sather, D. N.; Deutsch, E. W.; Kappe, S. H. I.; Moritz, R. L. A Tandem Mass Spectrometry Sequence Database Search Method for Identification of O-Fucosylated Proteins by Mass Spectrometry. *J. Proteome Res.* **2019**, *18* (2), 652–663. https://doi.org/10.1021/acs.jproteome.8b00638.

(32) Solntsev, S. K.; Shortreed, M. R.; Frey, B. L.; Smith, L. M. Enhanced Global Post-Translational Modification Discovery with MetaMorpheus. *J. Proteome Res.* **2018**, *17* (5), 1844–1851. https://doi.org/10.1021/acs.jproteome.7b00873.

(33) Polasky, D. A.; Yu, F.; Teo, G. C.; Nesvizhskii, A. I. Fast and Comprehensive N- and O-Glycoproteomics Analysis with MSFragger-Glyco. *Nat. Methods* **2020**, *17* (11), 1125–1132. https://doi.org/10.1038/s41592-020-0967-9.

(34)   Deutsch, E. W.; Csordas, A.; Sun, Z.; Jarnuczak, A.; Perez-Riverol, Y.; Ternent, T.; Campbell, D. S.; Bernal-Llinares, M.; Okuda, S.; Kawano, S.; Moritz, R. L.; Carver, J. J.; Wang, M.; Ishihama, Y.; Bandeira, N.; Hermjakob, H.; Vizcaíno, J. A. The ProteomeXchange Consortium in 2017: Supporting the Cultural Change in Proteomics Public Data Deposition. *Nucleic Acids Res.* **2017**, *45* (D1), D1100–D1106. https://doi.org/10.1093/nar/gkw936.

(35)   Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D. J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M.; Pérez, E.; Uszkoreit, J.; Pfeuffer, J.; Sachsenberg, T.; Yilmaz, Ş.; Tiwary, S.; Cox, J.; Audain, E.; Walzer, M.; Jarnuczak, A. F.; Ternent, T.; Brazma, A.; Vizcaíno, J. A. The PRIDE Database and Related Tools and Resources in 2019: Improving Support for Quantification Data. *Nucleic Acids Res.* **2019**, *47* (D1), D442–D450. https://doi.org/10.1093/nar/gky1106.

(36)   Geiszler, D. J.; Kong, A. T.; Avtonomov, D. M.; Yu, F.; Leprevost, F. da V.; Nesvizhskii, A. I. PTM-Shepherd: Analysis and Summarization of Post-Translational and Chemical Modifications From Open Search Results. *Mol. Cell. Proteomics* **2021**, *20*, 100018. https://doi.org/10.1074/mcp.TIR120.002216.

(37)   Renuse, S.; Madugundu, A. K.; Jung, J. H.; Byeon, S. K.; Goldschmidt, H. L.; Tahir, R.; Meyers, D.; Kim, D. I.; Cutler, J.; Kim, K. P.; Wu, X.; Huganir, R. L.; Pandey, A. Signature Fragment Ions of Biotinylated Peptides. *J. Am. Soc. Mass Spectrom.* **2020**, *31* (2), 394–404. https://doi.org/10.1021/jasms.9b00024.

(38)   McAlister, G. C.; Nusinow, D. P.; Jedrychowski, M. P.; Wühr, M.; Huttlin, E. L.; Erickson, B. K.; Rad, R.; Haas, W.; Gygi, S. P. MultiNotch MS3 Enables Accurate, Sensitive, and Multiplexed Detection of Differential Expression across Cancer Cell Line Proteomes. *Anal. Chem.* **2014**, *86* (14), 7150–7158. https://doi.org/10.1021/ac502040v.

(39)   Yu, F.; Haynes, S. E.; Nesvizhskii, A. I. IonQuant Enables Accurate and Sensitive Label-Free Quantification With FDR-Controlled Match-Between-Runs. *Mol. Cell. Proteomics* **2021**, *20*, 100077. https://doi.org/10.1016/j.mcpro.2021.100077.

(40)   Boutilier, J. M.; Warden, H.; Doucette, A. A.; Wentzell, P. D. Chromatographic Behaviour of Peptides Following Dimethylation with H2/D2-Formaldehyde: Implications for Comparative Proteomics. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.* **2012**, *908*, 59–66. https://doi.org/10.1016/j.jchromb.2012.09.035.

(41)   Kong, A. T.; Leprevost, F. V.; Avtonomov, D. M.; Mellacheruvu, D.; Nesvizhskii, A. I. MSFragger: Ultrafast and Comprehensive Peptide Identification in Mass Spectrometry-Based Proteomics. *Nat. Methods* **2017**, *14* (5), 513–520. https://doi.org/10.1038/nmeth.4256.

(42)   Palacios, F.; de Retana, A. M. O.; de Marigorta, E. M.; de los Santos, J. M. 2H-Azirines as Synthetic Tools in Organic Chemistry. *European J. Org. Chem.* **2001**, *2001* (13), 2401–2414. https://doi.org/10.1002/1099-0690(200107)2001:13<2401::AID-EJOC2401>3.0.CO;2-U.

(43)   Li, K.; Vaudel, M.; Zhang, B.; Ren, Y.; Wen, B. PDV: An Integrative Proteomics Data Viewer. *Bioinformatics* **2019**, *35* (7), 1249–1251. https://doi.org/10.1093/bioinformatics/bty770.

(44)   Shao, W.; Lam, H. Tandem Mass Spectral Libraries of Peptides and Their Roles in Proteomics Research. *Mass Spectrom. Rev.* **2017**, *36* (5), 634–648. https://doi.org/10.1002/mas.21512.

(45)   Sandberg, A.; Branca, R. M. M.; Lehtiö, J.; Forshed, J. Quantitative Accuracy in Mass Spectrometry Based Proteomics of Complex Samples: The Impact of Labeling and Precursor Interference. *J. Proteomics* **2014**, *96*, 133–144. https://doi.org/10.1016/j.jprot.2013.10.035.

(46)   Schweppe, D. K.; Prasad, S.; Belford, M. W.; Navarrete-Perea, J.; Bailey, D. J.; Huguet, R.; Jedrychowski, M. P.; Rad, R.; McAlister, G.; Abbatiello, S. E.; Woulters, E. R.; Zabrouskov, V.; Dunyach, J.-J.; Paulo, J. A.; Gygi, S. P. Characterization and Optimization of Multiplexed Quantitative Analyses Using High-Field Asymmetric-Waveform Ion Mobility Mass Spectrometry. *Anal. Chem.* **2019**, *91* (6), 4010–4016. https://doi.org/10.1021/acs.analchem.8b05399.

(47)   Cao, J.; Boatner, L. M.; Desai, H. S.; Burton, N. R.; Armenta, E.; Chan, N. J.; Castellón, J. O.; Backus, K. M. Multiplexed CuAAC Suzuki–Miyaura Labeling for Tandem Activity-Based Chemoproteomic Profiling. *Anal. Chem.* **2021**, *93* (4), 2610–2618. https://doi.org/10.1021/acs.analchem.0c04726.

(48)   Wdowiak, A. P.; Duong, M. N.; Joyce, R. D.; Boyatzis, A. E.; Walkey, M. C.; Nealon, G. L.; Arthur, P. G.; Piggott, M. J. Isotope-Coded Maleimide Affinity Tags for Proteomics Applications. *Bioconjug. Chem.* **2021**, *32* (8), 1652–1666. https://doi.org/10.1021/acs.bioconjchem.1c00206.

(49)   Zanon, P. R. A.; Lewald, L.; Hacker, S. M. Isotopically Labeled Desthiobiotin Azide (IsoDTB) Tags Enable Global Profiling of the Bacterial Cysteinome. *Angew. Chemie Int. Ed.* **2020**, *59* (7), 2829–2836. https://doi.org/10.1002/anie.201912075.

(50)   Ficarro, S. B.; Browne, C. M.; Card, J. D.; Alexander, W. M.; Zhang, T.; Park, E.; McNally, R.; Dhe-Paganon, S.; Seo, H.-S.; Lamberto, I.; Eck, M. J.; Buhrlage, S. J.; Gray, N. S.; Marto, J. A. Leveraging Gas-Phase Fragmentation Pathways for Improved Identification and Selective Detection of Targets Modified by Covalent Probes. *Anal. Chem.* **2016**, *88* (24), 12248–12254. https://doi.org/10.1021/acs.analchem.6b03394.

(51)   Browne, C. M.; Jiang, B.; Ficarro, S. B.; Doctor, Z. M.; Johnson, J. L.; Card, J. D.; Sivakumaren, S. C.; Alexander, W. M.; Yaron, T. M.; Murphy, C. J.; Kwiatkowski, N. P.; Zhang, T.; Cantley, L. C.; Gray, N. S.; Marto, J. A. A Chemoproteomic Strategy for Direct and Proteome-Wide Covalent Inhibitor Target-Site Identification. *J. Am. Chem. Soc.* **2019**, *141* (1), 191–203. https://doi.org/10.1021/jacs.8b07911.

(52)   McCloud, R. L.; Yuan, K.; Mahoney, K. E.; Bai, D. L.; Shabanowitz, J.; Ross, M. M.; Hunt, D. F.; Hsu, K.-L. Direct Target Site Identification of a Sulfonyl–Triazole Covalent Kinase Probe by LC-MS Chemical Proteomics. *Anal. Chem.* **2021**, acs.analchem.1c01591.