analytical chemistry

Technical Note

Cite This: Anal. Chem. 2017, 89, 11208-11213

pubs.acs.org/ac

Differential Enzymatic $^{16}O/^{18}O$ Labeling for the Detection of Cross-Linked Nucleic Acid–Protein Heteroconjugates

Fiona J. Flett,[†] Timo Sachsenberg,[‡] Oliver Kohlbacher,[‡,∥] C. Logan Mackay,[§] and Heidrun Interthal*[,†]

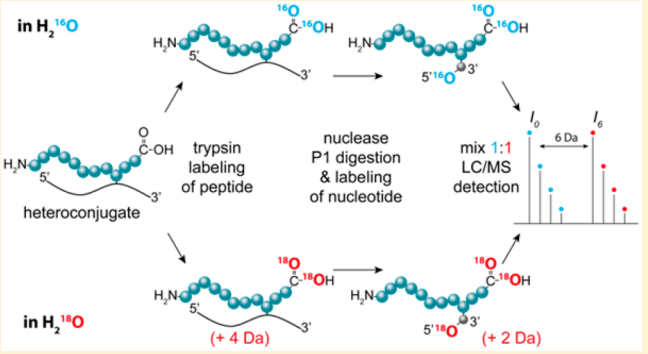
[†]Institute of Cell Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3FF, Scotland, U.K.

[‡]Center for Bioinformatics, Quantitative Biology Center and Department of Computer Science, University of Tübingen, Sand 14, 72076 Tübingen, Germany

[§]School of Chemistry, University of Edinburgh, Edinburgh EH9 3FJ, Scotland, U.K.

[∥]Max Planck Institute for Developmental Biology, Spemannstr. 35, 72076 Tübingen, Germany

ABSTRACT: Cross-linking of nucleic acids to proteins in combination with mass spectrometry permits the precise identification of interacting residues between nucleic acid–protein complexes. However, the mass spectrometric identification and characterization of cross-linked nucleic acid–protein heteroconjugates within a complex sample is challenging. Here we establish a novel enzymatic differential $^{16}O/^{18}O$-labeling approach, which uniquely labels heteroconjugates. We have developed an automated data analysis workflow based on OpenMS for the identification of differentially isotopically labeled heteroconjugates against a complex background. We validated our method using synthetic model DNA oligonucleotide–peptide heteroconjugates, which were subjected to the labeling reaction and analyzed by high-resolution FTICR mass spectrometry.



Mass spectrometry in combination with UV cross-linking is a powerful technique which can be applied to nucleic acid–protein complexes in order to identify proteins, peptides, and the amino acids involved in intermolecular interactions within nucleic acid–protein complexes.[1,2] Noncovalent nucleic acid–protein interactions are first stabilized by UV cross-linking to form covalent heteroconjugates. Such heteroconjugates can then be characterized by mass spectrometry to identify the cross-linked peptides and amino acids, following protease and nuclease digestion and an enrichment step. However, the mass spectrometric identification of cross-linked nucleic acid–protein heteroconjugates within a complex mixture is still challenging and cannot be performed by conventional MS search engines. Here we introduce a novel sequential differential enzymatic $^{16}O/^{18}O$ isotope-labeling strategy which has been designed to facilitate the mass spectrometric identification of oligonucleotide–peptide heteroconjugates, allowing them to be readily distinguished from non-cross-linked peptides, and their detection can be easily automated.

Normally following UV cross-linking, heteroconjugates are digested with trypsin resulting in a mixture of oligonucleotide–peptide heteroconjugates and non-cross-linked peptides and oligonucleotides. To establish this method, we have used synthetic model DNA oligonucleotide–peptide heteroconjugates,[3] labeled them using our differential-labeling approach described in Figure 1, and then analyzed them by high-resolution Fourier transform ion cyclotron resonance (FTICR) mass spectrometry. First, heteroconjugates are labeled by postdigestion trypsin- catalyzed labeling in the presence of either $H_2{}^{16}O$ or $H_2{}^{18}O$. During this step, heteroconjugates are differentially labeled at the C-terminus of the peptide moiety with $^{16}O_2$/ $^{18}O_2$. Non-cross-linked peptides are also labeled; therefore, to clearly identify heteroconjugates, a second labeling step is required. Nuclease P1 digestion in the presence of either $H_2{}^{16}O$ or $H_2{}^{18}O$ allows heteroconjugates to be uniquely labeled with either $^{16}O_1/^{18}O_1$ at the 5′ phosphate of the remaining DNA moiety. Following the two-step labeling approach, differentially labeled $^{16}O_3$ (I_0 isotopomers) and $^{18}O_3$ (I_6 isotopomers) samples are mixed 1:1 and analyzed by LC/MS. The differentially labeled heteroconjugates are unambiguously identified by the presence of a characteristic doublet, in which the monoisotopic peaks are clearly separated by 6 Da.

In addition, we have developed an OpenMS data analysis pipeline in which the detection of 6 Da doublets is automated, therefore allowing the easy identification of differentially labeled nucleic acid-peptide heteroconjugates within an LC/MS data set. Thus, a complete workflow for the labeling and detection of nucleic acid–peptide heteroconjugates is illustrated, which is applicable to any nucleic acid–protein cross-linking study.

Received: May 2, 2017
Accepted: September 8, 2017
Published: September 8, 2017

DOI: 10.1021/acs.analchem.7b01625
Anal. Chem. 2017, 89, 11208–11213

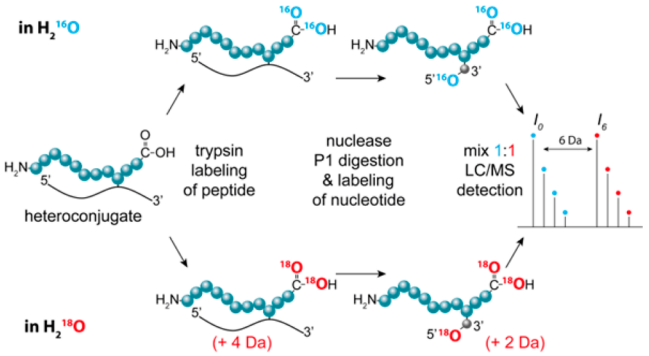


Figure 1. Sequential differential enzymatic $^{16}O/^{18}O$ labeling of nucleic acid–peptide heteroconjugates. Oligonucleotide–peptide heteroconjugates are first labeled by trypsin in $H_2{}^{16}O$ or $H_2{}^{18}O$ to label the C-terminus of the peptide moiety with $^{16}O_2$ or $^{18}O_2$. Heteroconjugates are then digested and labeled using nuclease P1 to label the 5′ monophosphate of the DNA moiety with ^{16}O or ^{18}O. Samples are then combined 1:1 and analyzed by LC/MS. The resulting heteroconjugates are labeled with $^{16}O_3$ or $^{18}O_3$, which appear as a characteristic doublet of I_0 and I_6 isotopomers, separated by 6 Da. Amino acids are depicted as green circles, an oligonucleotide as a black line, and a single nucleoside 5′ monophosphate as a single small gray circle.

MATERIALS AND METHODS

Chemicals, Solvents, and Reagents. All chemicals, solvents, and reagents were purchased from Sigma-Aldrich or Fisher Scientific. For mass spectrometry, solvents were of LC/MS grade or higher. For all stable isotope-labeling experiments, $H_2{}^{18}O$ (97% Cambridge Stable Isotope Laboratories, Inc., U.S.A.) was used to prepare all buffers and labeling solutions (except where stated). The synthetic heteroconjugates were synthesized as described previously.[3] For all heteroconjugates, the 20-mer oligonucleotide (5′GTAGAGGATCTAAAAGAC-XT-Biotin-TEG3′), where X was 5-ethynyl-2′-deoxyuridine (5EdU), was purchased from BaseClick, Germany. For heteroconjugate HC20-A, the peptide (LDIAFGTF*ATK), where F* was 4-azidophenylalanine, was purchased from Eurogentec, U.K. For heteroconjugates HC20-B, HC20-C, and HC20-D, peptides B (LDNAHF*GDATK), C (LDFAH-F*GDATK), and D (LDNSHF*GDATK), were synthesized in-house. For click chemistry, the Oligo-Click Kit was used as per the manual (BaseClick, Germany) and as described previously.

Trypsin-Catalyzed Labeling of HC20-A. Two reactions were set up in parallel to label HC20-A with either ^{16}O or ^{18}O. Prior to labeling, 500 pmol of HC20-A was washed in 3 × 10 μL of $H_2{}^{16}O$ or $H_2{}^{18}O$ and then reduced to dryness using a SpeedVac. For labeling, HC20-A was reconstituted in 50 μL (10 μM) containing 50 mM NH_4HCO_3, 10 mM $CaCl_2$, and 2 μg of trypsin (sequencing grade modified trypsin, Promega, 1 μg/μL stock dissolved in 50 mM NH_4HCO_3). Reactions were incubated at 37 °C for 5 h, and then, 10 μL of each reaction was analyzed by RP-HPLC/ESI-FTICR-MS in negative mode as described previously for HC20-A[3].

Nuclease P1 Labeling of HC20-A. Two reactions were set up in parallel to digest and label HC20-A with either ^{16}O or ^{18}O. Prior to labeling, 200 pmol of HC20-A was washed as described above. For labeling, HC20-A was reconstituted in 20 μL (10 μM) containing 50 mM $NH_4C_2H_3O_2$ (pH 5.2), 100 nM $ZnCl_2$, and 0.1 units nuclease P1 (N8630, Sigma-Aldrich, 250 units dissolved in 1 mL of $H_2{}^{16}O$). Reactions were incubated at 50 °C for 30 min, and then, 10 μL of each reaction was analyzed individually by monolith-HPLC/ESI-FTICR-MS in positive mode as described below.

Differential Sequential Labeling of Heteroconjugates. All labeling reactions were set up in parallel for $H_2{}^{16}O$ and $H_2{}^{18}O$: either a pure sample of 500 pmol of HC20-A in H_2O or samples containing the indicated mixtures of HC20-A, HC20-B, HC20-C, and HC20-D in molar ratios of 1:1, 1:2, or 1:5 with BSA tryptic peptides (10 μM BSA predigested) in 50 mM NH_4HCO_3, in either ^{16}O or ^{18}O. Trypsin-catalyzed labeling of all samples was then performed, as described above, in 50 μL (10 μM of each heteroconjugate and 10, 20, or 50 μM predigested BSA). Following trypsin labeling, samples were incubated at 100 °C for 10 min and phenylmethylsulfonyl fluoride was added to a final concentration of 1 mM. Samples were then reduced to dryness using a SpeedVac and digested with nuclease P1 as described above but in 50 μL (10 μM). For titanium dioxide heteroconjugate enrichment, samples were acidified by 2% formic acid and then loaded onto GELoader tip microcolumns packed in-house with titanium dioxide (TiO_2 Sachtopore NP 5 μm/300A) and washed three times with 30 μL of 2% formic acid (diluted in $H_2{}^{16}O/H_2{}^{18}O$). Samples were eluted with 20 μL of 1% ammonia into 10% formic acid and then reduced to dryness and reconstituted in 50 μL of $H_2{}^{16}O/H_2{}^{18}O$. To minimize back exchange, 5 μL of the differentially labeled ^{16}O and ^{18}O samples, containing 5 μM of each heteroconjugate, were combined immediately prior to analysis by monolith-HPLC/ESI-FTICR-MS in positive mode.

FTICR Mass Spectrometry. Analysis was performed using reverse-phase high-performance liquid chromatography (RP-HPLC) using an U3000 HPLC system (Dionex, UK) coupled to a standard electrospray source (Bruker Daltonics) and a SolariX FTICR mass spectrometer equipped with a 12 T superconducting magnet (BrukerDaltonics). Acquisition of LC/MS data was controlled by HyStar, version 3.4, build 8 (Bruker Daltonics).

Analysis of labeled HC20-A in negative mode was performed as described previously.[3]

For analysis of nuclease P1 digested and labelled heteroconjugates, RP-HPLC was performed using a monolith column (500 μm × 50 mm pepswift poly(styrene/divinylbenzene) (PSDVB) column, Thermo Finnigan, U.S.A.). For analysis of either ^{16}O- or ^{18}O- individually labeled heteroconjugate samples, 10 μL of a 10 μM heteroconjugate sample was injected. For analysis of differentially $^{16}O/^{18}O$-labeled heteroconjugates, 5 μL of a 10 μM ^{16}O-labeled heteroconjugate sample was mixed with 5 μL of a 10 μM ^{18}O-labeled heteroconjugate sample, and 10 μL was injected. Heteroconjugates were analyzed in positive mode using buffer A (0.1% formic acid in H_2O) and buffer B, (0.1% formic acid in acetonitrile). With use of a flow rate of 20 μL/min at 60 °C, the column was first run at 0% B for 2 min, followed by a linear ramp from 0 to 100% B from 2 to 20 min. The column was then washed with 100% B from 20 to 25 min and then with 0% B from 25 to 30 min.

For electrospray ionization, the gas pressure was typically ~2.2 psi and the spray voltage was 4.5 kV. For mass spectrometry, ion accumulation times were typically 0.3 s. Ions were trapped using a 6 × 10 cm infinity cell. Each individual LC/MS spectrum was the sum of two acquisitions. The transient data size was typically 1 or 2 Mword for each acquisition, and the sine-bell multiplication apodization was applied to each transient during FTMS postprocessing. All

DOI: 10.1021/acs.analchem.7b01625
Anal. Chem. 2017, 89, 11208–11213

mass spectra were analyzed using DataAnalysis software version 4.1 SR1 build 362.7 (Bruker Daltonics).

OpenMS Data Analysis Pipeline. Data was exported to mzML files using CompassXport 3.0 (Bruker Daltonics). The mzML files formed the input for the data analysis pipeline outlined in Figure 2. To reduce the impact of low intensity signals on doublet detection, a signal processing step was performed that retains the highest intensity mass peak in a sliding window of size 0.2 Thomson using the OpenMS tool SpectraFilterWindowMower.[4,5] In the second step of the pipeline, we configure the tool FeatureFinderMultiplex[6] to detect eluting species exhibiting the characteristic 6 Da shift with an m/z tolerance set to 5 ppm. The FileFilter tools then discarded all detected singlets, and only doublets were retained. The list of 6 Da doublets is then exported to a tabular text file using the TextExporter tool.

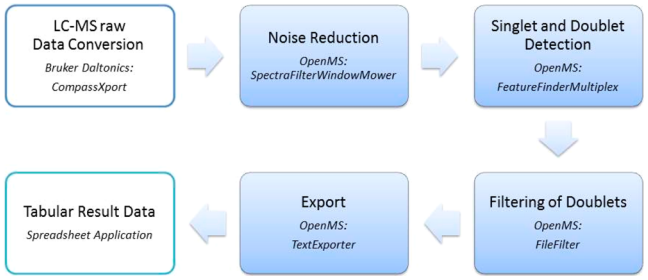


Figure 2. Overview of the data processing pipeline. Raw data is exported and input into the automated OpenMS data processing workflow (shaded boxes). Each node corresponds to an OpenMS pipeline tool (name in italic letters) that performs a distinct data processing step. Detected doublets are then written as a tabular file and can be inspected using a spreadsheet application.

RESULTS

Trypsin-Catalyzed Labeling of Heteroconjugates. Serine proteases including trypsin, LysC, and GluC can catalyze the incorporation of two ^{18}O atoms at the carboxyl terminus of a proteolytic peptide in the presence of the heavy isotopic form of water $H_2{}^{18}O$.[7] For ^{18}O labeling, proteolytic digestion can be decoupled from enzyme-catalyzed oxygen exchange,[8] which offers the advantage that both reactions can be performed separately. Therefore, postdigestion trypsin-catalyzed oxygen exchange can be optimized to promote the incorporation of two ^{18}O atoms and to achieve a high labeling efficiency, which is required for the differential-labeling strategy.

To demonstrate efficient postdigestion trypsin-catalyzed ^{18}O labeling of DNA-peptide heteroconjugates, two labeling reactions were set up in parallel to label a synthetic DNA oligonucleotide–peptide heteroconjugate HC20-A with either ^{16}O or ^{18}O and then were analyzed individually by RP-HPLC/ESI-FTICR mass spectrometry in negative mode (Figure 3A).

In the presence of $H_2{}^{18}O$, the monoisotopic peak of HC20-A was increased in mass by 4 Da following the incorporation of two ^{18}O atoms into the C-terminus of the peptide moiety to form the I_4 isotopomer. The high mass-resolving power of FTICR mass spectrometry allowed the clear distinction between the I_0 and I_4 isotopomers. The associated change in mass did not affect the HPLC retention time of the HC20-A.

Nuclease P1-Catalyzed Labeling of Heteroconjugates. Nuclease P1 digests the DNA moiety of HC20-A into a single nucleoside 5′ monophosphate covalently bound to the peptide via the triazole linkage.[3] In the presence of $H_2{}^{18}O$, we expected nuclease P1 to label heteroconjugates by incorporating a single ^{18}O atom on the remaining nucleotide.

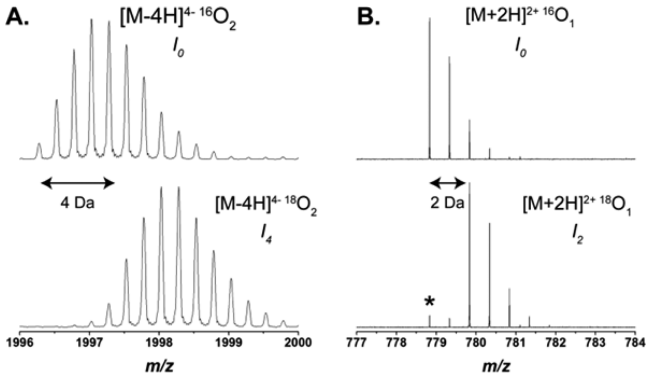


Figure 3. Trypsin- and nuclease P1-catalyzed $^{16}O/^{18}O$ labeling of HC20-A. (A) RP-HPLC/ESI-FTICR-MS negative-mode mass spectra of HC20-A following trypsin-catalyzed labeling. The $[M - 4H]^{4-}$ ions of I_0 and I_4 isotopomers are present at m/z 1996.2232 and 1997.2254 in $H_2{}^{16}O$ and $H_2{}^{18}O$, respectively; (B) Monolith-HPLC/ESI-FTICR positive-mode mass spectra of HC20-A following nuclease P1 digestion. The $[M + 2H]^{2+}$ ions of I_0 and I_2 isotopomers are present at m/z 778.8336 and 779.8358 in $H_2{}^{16}O$ and $H_2{}^{18}O$, respectively. An asterisk indicates remaining I_0 species present in the $H_2{}^{18}O$-labeled sample.

To demonstrate nuclease P1-catalyzed labeling, two reactions were performed in parallel to digest and label HC20-A with either ^{16}O or ^{18}O and then were analyzed individually by RP-HPLC/ESI-FTICR mass spectrometry in positive mode (Figure 3B). In the presence of 95% $H_2{}^{18}O$, the monoisotopic peak of the digested HC20-A had increased in mass by 2 Da following the incorporation of a single ^{18}O label into the remaining 5′ monophosphate to form the I_2 isotopomer. Although very efficiently labeled, ^{18}O labeling to form the I_2 isotopomer was not 100%, with some I_0 isotopomer also present (Figure 3B, annotated with an asterisk), likely due to the occurrence of minimal back exchange with residual ^{16}O. Again, the associated change in mass did not affect the retention time of the digested HC20-A. In conclusion, HC20-A can be $^{16}O/^{18}O$-labeled by either trypsin introducing a 4 Da shift or nuclease P1 introducing a 2 Da shift.

Differential Sequential Labeling of Heteroconjugates. To uniquely label heteroconjugates with $^{18}O_3$, we next combined both the trypsin and nuclease P1-labeling steps. To demonstrate this, two reactions were performed in parallel to label heteroconjugate HC20-A with either $^{16}O_3$ or $^{18}O_3$. Trypsin-catalyzed ^{18}O labeling was performed first, and following this, to prevent digestion of nuclease P1 in the next step, trypsin was chemically and heat inactivated. Next, nuclease P1-catalyzed ^{18}O labeling was performed. Samples were then were analyzed individually, mixed 1:1, and analyzed by RP-HPLC/ESI-FTICR mass spectrometry in positive mode. In the presence of $H_2{}^{18}O$, the monoisotopic peak of HC20-A was increased in mass by 6 Da following the dual-labeling approach (Figure 4A,B) to form the I_6 isotopomer. When the labeled samples were mixed 1:1 prior to analysis, the heteroconjugate was detected as a doublet with the coeluting I_0 and the I_6 isotopomers, 6 Da apart, within a single LC/MS run (Figure 4, lower spectrum).

In a UV cross-linking experiment, heteroconjugates are present within a complex sample containing a background of

DOI: 10.1021/acs.analchem.7b01625
Anal. Chem. 2017, 89, 11208–11213

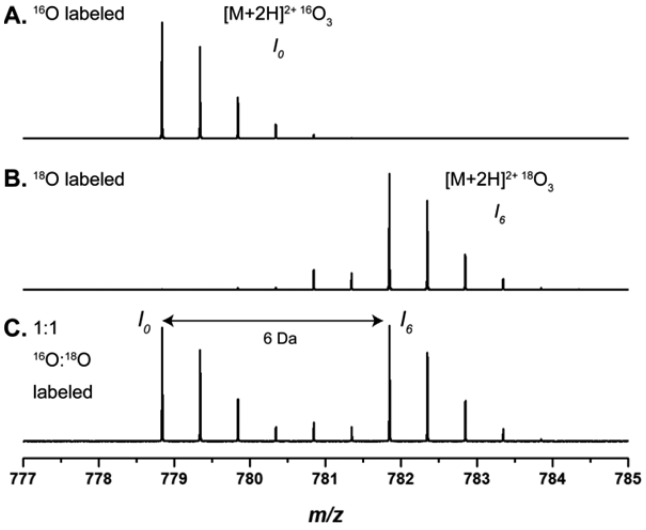


Figure 4. Differential $^{16}O/^{18}O$ labeling of the heteroconjugate. Monolith-HPLC/ESI-FTICR positive-mode mass spectra of differentially labeled HC20-A. Differentially ^{16}O- and ^{18}O-labeled samples were either (A and B) run individually, or (C) mixed 1:1 prior to MS analysis. The $[M + 2H]^{2+}$ ions of I_0 and I_6 isotopomers are present at m/z 778.8535 and 781.8569 in $H_2{}^{16}O$ and $H_2{}^{18}O$, respectively.

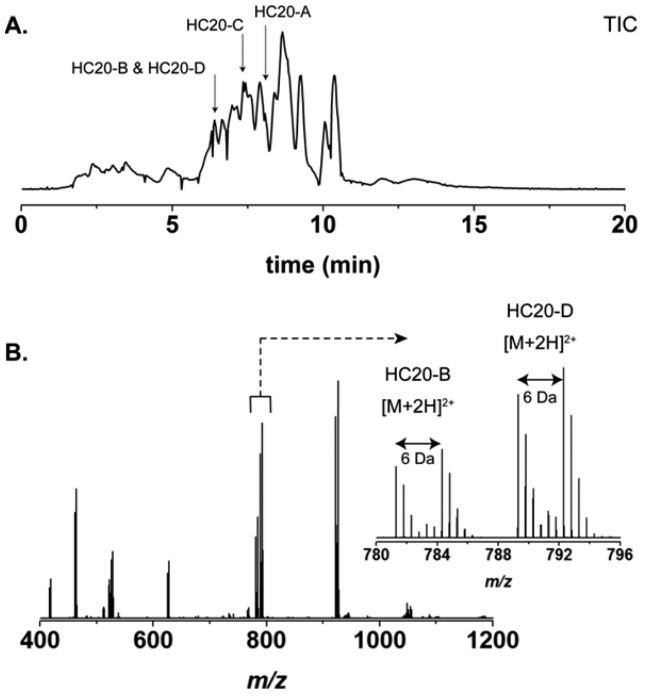


Figure 5. Differential $^{16}O/^{18}O$-labeling of the heteroconjugates. (A) Total ion chromatogram of a differentially labeled BSA tryptic digest containing HC20-A, HC20-B, HC20-C, and HC20-D, analyzed by Monolith-HPLC/ESI-FTICR positive-mode mass spectrometry. The elution time of each heteroconjugate is annotated. (B) The mass spectrum at time 6.4 min, and inset, zoomed in to shown differentially labeled HC20-B and HC20-D.

non-cross-linked species. Therefore, to ensure that the differential labeling of heteroconjugates was done efficiently within a more complex sample, we next performed differential labeling of HC20-A within a tryptic digest of BSA. In addition to HC20-A, we employed a range of heteroconjugates which varied in their overall sequence and charge to ensure labeling and detection was not restricted to HC20-A. Differential labeling was performed as previously.

Heteroconjugates HC20-A, -B, -C, and -D were observed to elute at 8.0, 6.4, 7.4, and 6.4 min, respectively, as indicated on the total ion chromatogram shown in Figure 5. All four heteroconjugates were efficiently labeled and observed as unique doublets within this complex sample. In Figure 5B, the mass spectrum at 6.4 min is shown, in which both differentially labeled HC20-B and differentially labeled HC20-D were present as unique 6 Da doublets within this spectrum. In conclusion, a range of heteroconjugates can be uniquely and efficiently labeled within a more complex sample using a two-step differential-labeling approach.

Detection of Differentially Labeled Heteroconjugates. In the experiments described above, differentially labeled heteroconjugate doublets were easily identified as we employed defined synthetic model heteroconjugates. However, finding unknown doublets in a real cross-linking experiment is a challenging task. To automate the detection of differentially labeled doublets and allow the efficient detection of potential heteroconjugate species, we developed an OpenMS data analysis pipeline. Doublets detected by the pipeline represent potential heteroconjugate precursor ions that could then be further characterized and confirmed by MS/MS analysis. If the detected doublet is a true heteroconjugate, fragmentation of both the light- and heavy-labeled species would give complementary labeled fragment spectra. The ^{18}O labeling may also aid in sequencing of the nucleic acid−peptide heteroconjugates, similar to its application for peptide and cross-linked peptide de novo sequencing.[9,10]

To test the data analysis pipeline, data was analyzed from samples in which equimolar mixtures of BSA tryptic peptides and HC20-B, HC20-C, and HC20-D were differentially labeled in triplicate and analyzed by LC/MS. The pipeline detected four 6 Da doublet species that were present in all three replicates, whereas no doublets were detected in control samples containing only labeled BSA peptides.

The detected doublets correspond to the $[M + 2H]^{2+}$ and $[M + 3H]^{3+}$ ions of HC20-B and HC20-D (Table 1). HC20-C was not detected, but manual inspection of the data revealed that a highly abundant BSA peptide with a similar mass coeluted with HC20-C, which resulted in an overlapping isotopic distribution and interfered with the detection of HC20-C by the pipeline. This highlights the importance of enrichment strategies that are required in all cross-linking workflows to remove the majority of excess non-cross-linked species and, thus, enable the identification of cross-linked heteroconjugates.[1,2] To test if our labeling strategy is compatible with enrichment methods and if enrichment could improve the detection of heteroconjugates, the samples were enriched by using titanium dioxide columns.[11] Importantly, the 6 Da label of heteroconjugates was maintained following enrichment, and in addition, HC20-C could now be detected by the pipeline (Table 1).

In a true cross-linking experiment, owing to low cross-linking efficiency, heteroconjugates are present in substochiometric amounts compared with non-cross-linked peptides. Therefore, the pipeline was tested using different molar ratios of BSA peptides/heteroconjugates. At a ratio of 2:1, the pipeline detected HC20-B and HC20-D (Table 1). However, with a larger excess of BSA peptides at 5:1, only a single HC20-B $[M + 2H]^{2+}$ ion was detected. Following titanium dioxide enrichment, a further three heteroconjugates ions were

DOI: 10.1021/acs.analchem.7b01625
Anal. Chem. 2017, 89, 11208−11213

Table 1. Heteroconjugates Detected by the OpenMS Pipeline, with and without Titanium Dioxide Enrichment[a]

molar ratio of BSA peptides/heteroconjugates			1:1		2:1		5:1	
titanium dioxide enrichment			-	+	-	+	-	+
HC20-B	$[M+2H]^{2+}$	m/z = 781.3070	Y	Y	Y	Y	Y	Y
	$[M+3H]^{3+}$	m/z = 521.2071	Y	Y	Y	Y	–	Y
HC20-C	$[M+2H]^{2+}$	m/z = 797.8197	–	Y	–	Y	–	–
HC20-D	$[M+2H]^{2+}$	m/z = 789.3045	Y	Y	Y	Y	–	Y
	$[M+3H]^{3+}$	m/z = 526.5387	Y	Y	Y	Y	–	Y

[a]Y = Yes, ion detected and – = No, ion not detected.

detected (Table 1). Titanium dioxide enrichment increased the sensitivity of the detection in every scenario tested.

In conclusion, the OpenMS-based analysis pipeline efficiently detects differentially labeled heteroconjugates within an LC/MS data set. In cross-linking experiments, candidate heteroconjugate species could be confirmed with targeted MS/MS approaches. Lastly, our labeling strategy is compatible with standard heteroconjugate enrichment protocols, a key step in all cross-linking workflows.

DISCUSSION

UV cross-linking and mass spectrometry is a useful method to characterize interactions in different nucleic acid–protein complexes.[1,2] However, the identification of heteroconjugates is a challenging task. We have developed a novel isotopic differential-labeling strategy to uniquely label heteroconjugates and enable their detection by an OpenMS data analysis pipeline by the presence of isotopic doublets. This labeling approach could also be used to complement and increase confidence in heteroconjugates identified by the recently developed software program RNPxl. RNPxl detects heteroconjugates by calculating and searching for a variety of nucleic acid modifications.[12] Length and composition of nucleic acid moieties remaining after nuclease digestion are variable, and any amino acid could potentially be cross-linked. Therefore, a large number of nucleic acid modifications must be considered in an RNPxl database search, and conventional search engines are not suitable for such analysis.

Nuclease P1 digestion alone can uniquely label heteroconjugates. However, automated detection of 2 Da differentially labeled heteroconjugate doublets is challenging due to the overlapping isotope distributions of the I_0 and I_2 isotopomers. We therefore combined both the trypsin- and nuclease P1-labeling steps to uniquely label heteroconjugates (with 6 Da) and provide optimal resolution between the isotope distributions of the I_0 and I_6 isotopomers. In addition, the 6 Da shift is small enough to allow accurate detection of the mass difference using high-resolution mass spectrometry. Labeling heteroconjugates with $^{18}O_3$ (6 Da) allows them to be easily distinguished from contaminants such as non-cross-linked (but labeled) peptides (4 Da) and nucleic acids (2 Da), which are commonly present in cross-linking samples, even following heteroconjugate enrichment.

This labeling approach is applicable to existing methods published thus far for generating nucleic acid-peptide heteroconjugates, as it exploits the protease and nuclease digestion steps necessary for the sample preparation of heteroconjugates. This method can detect heteroconjugates with a variety of nucleic acid modifications that have been UV cross-linked with native DNA or photoactivatable DNA, or cross-linked chemically, and does not require that the cross-linking mechanism and any associated losses are known in advance. It also does not require specially synthesized, isotopically labeled nucleic acids,[13] and therefore it would be suitable for cross-linking experiments using any synthetic nucleic acid as well as nucleic acids in or isolated from cells. Synthetic heteroconjugates as described here may be included to serve as internal controls for efficient labeling and enrichment. This labeling approach is relevant to both the study of DNA–protein and RNA–protein heteroconjugates and would be compatible with other serine proteases and nucleases that act by a hydrolysis mechanism. Lastly, the steps of protease and nuclease labeling are flexible and could be reversed or incorporated at any stage within the sample preparation and enrichment protocol as appropriate for the user. In conclusion, the differential-labeling approach introduced and developed here facilitates the identification of cross-linked nucleic acid-peptide heteroconjugates by mass spectrometry.

AUTHOR INFORMATION

Corresponding Author

*E-mail: heidrun.interthal@ed.ac.uk.

ORCID

Heidrun Interthal: 0000-0002-3827-9460

Notes

The authors declare no competing financial interest.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[14] partner repository with the data set identifier PXD003673. The OpenMS workflow can be found on our Web site at http://www.OpenMS.de/workflows.

ACKNOWLEDGMENTS

We would like to thank David Clarke and Jeffrey G. A. Walton for helpful discussion. This work was supported by Medical Research Council Career Development Award G0700257 and a Scottish Universities Life Sciences Alliance grant (to H.I.). F.J.F. was supported by a BBSRC postgraduate studentship. T.S. and O.K. are supported by the BMBF grant de.NBI—German Network for Bioinformatics Infrastructure (FKZ 031 A 535A).

REFERENCES

(1) Schmidt, C.; Kramer, K.; Urlaub, H. *J. Proteomics* **2012**, *75*, 3478–3494.

(2) Steen, H.; Jensen, O. N. *Mass Spectrom. Rev.* **2002**, *21*, 163–182.

(3) Flett, F. J.; Walton, J. G.; Mackay, C. L.; Interthal, H. *Anal. Chem.* **2015**, *87*, 9595–9599.

(4) Kohlbacher, O.; Reinert, K.; Gröpl, C.; Lange, E.; Pfeifer, N.; Schulz-Trieglaff, O.; Sturm, M. *Bioinformatics* **2007**, *23*, e191–e197.

DOI: 10.1021/acs.analchem.7b01625
Anal. Chem. 2017, 89, 11208–11213

(5) Sturm, M.; Bertsch, A.; Gröpl, C.; Hildebrandt, A.; Hussong, R.; Lange, E.; Pfeifer, N.; Schulz-Trieglaff, O.; Zerck, A.; Reinert, K.; et al. *BMC Bioinf.* **2008**, *9*, 163.
(6) Nilse, L.; Sigloch, F. C.; Biniossek, M. L.; Schilling, O. *Proteomics: Clin. Appl.* **2015**, *9*, 706−714.
(7) Ye, X.; Luke, B.; Andresson, T.; Blonder, J. *Briefings Funct. Genomics Proteomics* **2009**, *8*, 136−144.
(8) Yao, X.; Afonso, C.; Fenselau, C. *J. Proteome Res.* **2003**, *2*, 147−152.
(9) Back, J. W.; Notenboom, V.; de Koning, L. J.; Muijsers, A. O.; Sixma, T. K.; de Koster, C. G.; de Jong, L. *Anal. Chem.* **2002**, *74*, 4417−4422.
(10) Shevchenko, A.; Chernushevich, I.; Ens, W.; Standing, K. G.; Thomson, B.; Wilm, M.; Mann, M. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 1015−1024.
(11) Kramer, K.; Hummel, P.; Hsiao, H.-H.; Luo, X.; Wahl, M.; Urlaub, H. *Int. J. Mass Spectrom.* **2011**, *304*, 184−194.
(12) Kramer, K.; Sachsenberg, T.; Beckmann, B. M.; Qamar, S.; Boon, K.-L.; Hentze, M. W.; Kohlbacher, O.; Urlaub, H. *Nat. Methods* **2014**, *11*, 1064−1070.
(13) Lelyveld, V. S.; Björkbom, A.; Ransey, E. M.; Sliz, P.; Szostak, J. W. *J. Am. Chem. Soc.* **2015**, *137*, 15378−15381.
(14) Vizcaíno, J. A.; Csordas, A.; del-Toro, N.; Dianes, J. A.; Griss, J.; Lavidas, I.; Mayer, G.; Perez-Riverol, Y.; Reisinger, F.; Ternent, T.; et al. *Nucleic Acids Res.* **2016**, *44*, D447−D456.

DOI: 10.1021/acs.analchem.7b01625
Anal. Chem. 2017, 89, 11208−11213