Journal of proteome research

pubs.acs.org/jpr

Article

Reversible Click Chemistry Tag for Universal Proteome Sample Preparation for Top-Down and Bottom-Up Analysis

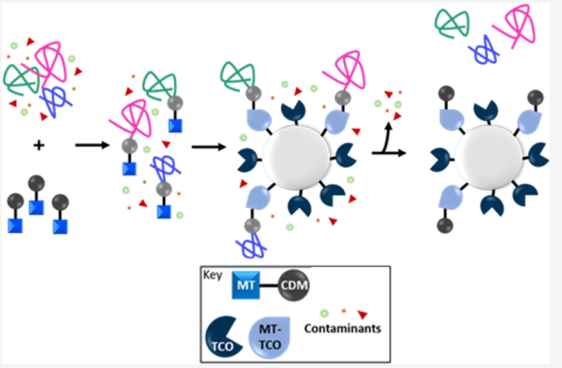
Stephanie Biedka, Brigitte F. Schmidt, Nolan M. Frey, Sarah M. Boothman, Jonathan S. Minden, and Amber Lee Wilson*

Cite This: *J. Proteome Res.* 2021, 20, 4787−4800

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Successful proteome analysis requires reliable sample preparation beginning with protein solubilization and ending with a sample free of contaminants, ready for downstream analysis. Most proteome sample preparation technologies utilize precipitation or filter-based separation, both of which have significant disadvantages. None of the current technologies are able to prepare both intact proteins or digested peptides. Here, we introduce a reversible protein tag, ProMTag, that enables whole proteome capture, cleanup, and release of intact proteins for top-down analysis. Alternatively, the addition of a novel Trypsin derivative to the workflow generates peptides for bottom-up analysis. We show that the ProMTag workflow yields >90% for intact proteins and >85% for proteome digests. For top-down analysis, ProMTag cleanup improves resolution on 2D gels; for bottom-up exploration, this methodology produced reproducible mass spectrometry results, demonstrating that the ProMTag method is a truly universal approach that produces high-quality proteome samples compatible with multiple downstream analytical techniques. Data are available via ProteomeXchange with identifier PXD027799.



KEYWORDS: *sample preparation, click chemistry, proteomics, protein, protein modification, protein chemistry, sample cleanup, two-dimensional gel electrophoresis, protein mass spectrometry, reversible chemistry*

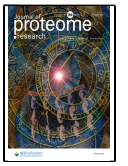
INTRODUCTION

The first, and arguably most important, step in any proteomics workflow is sample preparation. High-quality protein extraction and cleanup during sample preparation are critical for achieving the yields, reproducibility, coverage, and sensitivity required for successful and reliable proteome analysis. Studies comparing the currently available technology for protein sample preparation have revealed that each option has unique advantages and disadvantages.[1−6] Current technologies for proteomics sample preparation generally rely on precipitation or filtration-based technologies, both of which can suffer from sample loss, long processing times, incompatibility with certain buffers that aid in protein solubilization, bias for certain proteins, and limitations for automation.[7−13] Currently, no single sample preparation technique can be applied universally to yield either intact protein or digested peptide proteome samples. Thus, sample preparation optimization typically requires considerable effort, resources, and sample material to develop reliable extraction and cleanup protocols.

Nearly all proteins contain primary amines in the form of lysine residues and their amino termini. Reliance on this fact is the basis behind several successful proteomics methods such as two-dimensional difference gel electrophoresis (2D-DIGE), isobaric tag for relative and absolute quantitation, and tandem mass tags where N-hydroxysuccinimide (NHS) ester-based tags are covalently linked to lysine residues and amino termini.[14−19] These amide linkages are irreversible under typical conditions used in proteomics research. In an effort to develop a truly universal proteome sample preparation approach, we developed a bifunctional protein tag where one end of the tag forms a reversible covalent linkage with any protein that contains a primary amine, and the other end forms an irreversible covalent linkage with a bead-based matrix.

More than half a century ago, it was shown that dicarboxylic acid anhydride derivatives such as succinic anhydride, maleic anhydride, and citraconic anhydride could be used to reversibly modify amine residues on proteins.[20−23] These reversible modifications are pH-dependent; at slightly basic pH (~8), dicarboxylic acid anhydrides form amides with protein amine

Received: May 24, 2021
Published: September 15, 2021



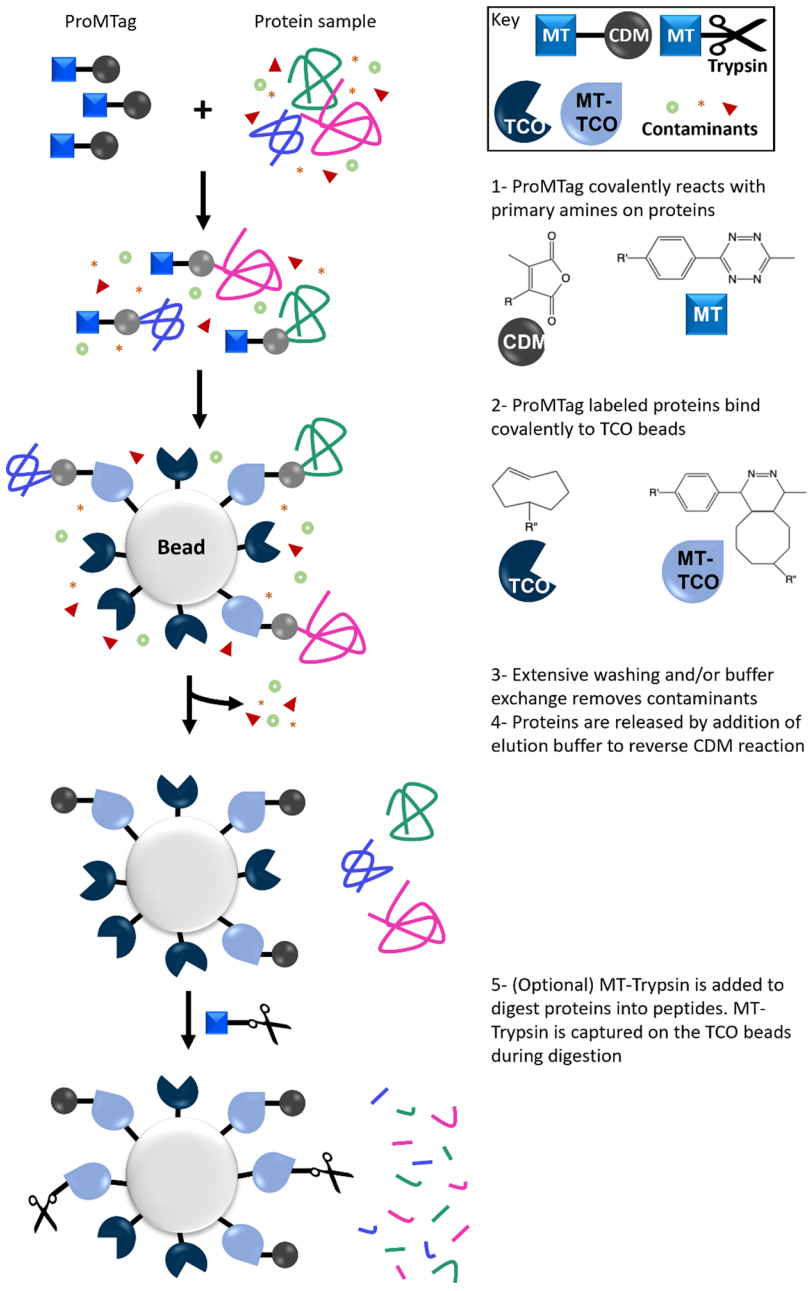


Figure 1. ProMTag cleanup workflow.

residues. Lowering the pH below 5 reverses this amide linkage to recyclize the dicarboxylic acid anhydride and returns the proteins to their original, unmodified state.[21] While the mechanism for amide reversal has not been completely investigated, mechanisms have been proposed (Figure 2C).[24]

Ganesan et al. capitalized on a derivative of the dicarboxylic acid anhydride, carboxyl dimethyl maleic anhydride (CDM) for the separation of antigens and antibodies for immunoproteomic analysis.[25] They showed that by joining CDM to Biotin via a short PEG linker, antigens could be reversibly tagged with Biotin-CDM prior to immunoprecipitation. After immunoprecipitation and release from Protein A beads, tagged antigen proteins were captured on Streptavidin beads and separated from the untagged immunoglobulins. Purified antigens were isolated by reversing the CDM linkage.

While this specific study demonstrated the utility of CDM as a reversible tag, the Biotin moiety was not ideal for whole proteome capture and release. To tag an entire proteome, a large excess of tag was required, thus requiring an equally large number of Avidin monomers to bind all of the added Biotin. Because Avidin tetramers are relatively large and cannot pack as tightly on the surface of a bead as smaller molecules, whole proteome capture and release would require a prohibitively large amount of Avidin beads or a dialysis step that would result in protein loss/dilution. In addition, Avidin itself is a protein so use of harsh denaturants or digestive enzymes would negatively impact downstream proteome analysis. Recently, studies have been performed to attempt to overcome these deficiencies by sequestering free interfering Biotin, by rendering Streptavidin resistant to Trypsin/Lys-C digestion, and by accounting for Streptavidin shedding contamina-

https://doi.org/10.1021/acs.jproteome.1c00443
J. Proteome Res. 2021, 20, 4787−4800

tion.[26−28] However, this still did not solve the problem of the large amount of Avidin that would be required for whole proteome cleanup and the inability to use Avidin-based cleanup with samples in harsh denaturants.

Rather than using the Biotin−Avidin interaction to bind the CDM tag to beads, we took advantage of a fast, bio-orthogonal click chemistry reaction, using the covalent reaction between methyl-tetrazine (MT) and trans-cyclooctene (TCO). MT and TCO make an ideal reaction pair for bio-orthogonal chemistry for several reasons. First, the reaction does not require a catalyst, so there are no copper-associated side reactions and the reaction has faster kinetics than both copper-catalyzed and strain-promoted azide−alkyne reactions, with rate constants from 800 to 30,000 M^{-1} s^{-1}, depending on the tetrazine derivative used.[29,30] Additionally, this reaction proceeds best in an aqueous environment at physiological pH and room temperature.[31] It also has high selectivity and very little cross reactivity, unlike the previously described pairs. To date, this click chemistry pair has been used for fluorescent probing of cellular structures, fluorescent labeling of specific cell types, and ^{18}F labeling of drugs.[30,32−36] The ability to conjugate other functional groups to the tetrazine and TCO groups makes them ideal for tagging and labeling cellular structures. Here, we describe the synthesis of MT-CDM, referred to as **Pro**tein **M**ethyl-tetrazine-CDM **Tag** (ProMTag) that includes a CDM moiety that reversibly reacts with protein primary amines and a MT moiety that reacts quickly and irreversibly with TCO-agarose beads.

In this study, we show that ProMTag paired with TCO-agarose beads can be used to tag, capture, cleanup, and release protein samples in their original, unmodified state for fast, high-yield, high-coverage protein sample preparation (Figure 1). We also introduce a new form of modified Trypsin that is MT-tagged, making it resistant to autodigestion and allowing sequestration and removal from solution by binding to TCO-beads. Thus, we demonstrate that this workflow can be used for either intact protein applications or for digested peptide applications by adding a digestion step with MT-Trypsin, showing proteome analysis of samples derived from a range of cell types using both two-dimensional electrophoresis (2-DE) of intact proteins and mass spectrometry (MS) of digested peptides. We compared a conventional precipitation/in-solution digestion workflow to the ProMTag workflow which revealed that the ProMTag workflow leads to a higher number of protein identifications and higher reproducibility. This novel protein sample preparation technique utilizing reversible click chemistry represents an alternative to filter and precipitation-based sample preparation methods. It is relatively fast, unbiased, of high yield, and highly reproducible.

METHODS AND MATERIALS

Synthesis of ProMTag and Confirmation via MS and Nuclear Magnetic Resonance (NMR)

2-(2-Carboxyethyl)-3-methyl-maleic anhydride (TCI Chemicals, 1 mmol, 184.15 mg) was placed in a 25 mL round bottom flask and dissolved in 10 mL of anhydrous dichloromethane (DCM). Oxalyl chloride (TCI Chemicals, 1.1 mmol, 130 mg) was added followed by one drop of dimethylformamide (DMF). The reaction mixture was stirred for about 1 h at 40 °C until gas evolution ceased. The solvent was removed by drying with a rotary evaporator with a final vacuum 200 of mm/Hg. The resulting residue was dissolved in 10 mL of anhydrous acetonitrile (ACN).

Methyltetrazine-PEG4-amine (MT-PEG-NH_2-HCl) (Click Chemistry Tools, 0.25 mmol, 100 mg) was dissolved in 5 mL anhydrous ACN. N,N-Diisopropylethylamine (Beantown Chemicals, 1 mmol, 175 μL) (DIPEA) was added, and the resulting solution was added to the 2-carboxyethyl-3-methyl-maleic anhydride solution over a 5 min time period while stirring. The reaction mixture was stirred at room temperature for 2 h.

The reaction was quenched by adding 10 mL of 1 N HCl and stirring for 1 h. The solvent was evaporated using a rotary evaporator with a water bath temperature of 40 °C and setting 200 mm/Hg until the product started to precipitate. This mixture was then purified by medium pressure liquid chromatography (MPLC) on a RP-18 column (15 mm/96 mm) on a Buchi Sepacore System. The product was purified using water/ 10% ACN, 0.1% trifluoroacetic acid. Single fractions were analyzed by ultraperformance liquid chromatography (UPLC) (Waters Acquity). The product eluted in two peaks, which were both isolated and concentrated to dryness. The residue was dissolved in ACN, transferred to a vial, and again dried. 1H-NMR analysis showed that both peaks were the ProMTag product, with different elution times because of two isoforms of the product. Both peaks were shown to react identically with primary amines, so the fractions were combined and used as a single product for all future experiments.

The product was confirmed using MS by Carnegie Mellon University's Center for Molecular Analysis using the following parameters: the ProMTag product was diluted to 20 μM in 100% ACN and acquired in flow injection analysis mode with background solvent 100% ACN. The sample was run in positive ion mode on a Finnigan LCQ Mass Spectrometer. The source voltage was 3.54 kV, the source current was 0.24 μA, the sheath gas flow rate was 18.85, the auxiliary gas flow rate was 0.86, the capillary voltage was 16.10 V, and the capillary temperature was 200.10 °C. For the syringe pump, the flow rate was 8 μL/min, the infused volume was 207 μL, and the syringe diameter was 3.26 mm.

Protein Sample Sources

Single-protein experiments were carried out with carbonic anhydrase, lysozyme, and alcohol dehydrogenase (ADH) (all from Sigma) dissolved in 100 mM HEPES pH 8.0, 1% SDS, and 200 mM NaCl.

Yeast (*Saccharomyces cerevisiae*, strain W303 or JWY6147) was grown at 30 °C in YEPD (2% dextrose, 2% peptone, and 1% yeast extract) to the mid-log phase. Cells were washed with water and stored at −80 °C. Cells were lysed by boiling in 200 mM HEPES pH 8.0, 2% SDS, and 400 mM NaCl for 10 min. Cell debris was pelleted by centrifugation at 3300 x g for 10 min.

Tissue culture (TC) cell lysis was performed by first washing 30 μL of frozen, packed TC cells (a mixture of mammary tumor cell lines: 10A, NeoT, T1, and CA1) in ice-cold PBS. A surfactant cocktail (130 μL of 1% SDS, 1% IGEPAL, 0.5% Na deoxycholate, and 100 mM HEPES pH 8.0) was added to lyse cells.[37] The cells were sonicated with 10 pulses at 30% power and 30% duty cycle. The sample was then boiled for 10 min and spun at 15,000 rpm for 20 min in a desktop microcentrifuge at room temperature. The supernatant was removed and stored at −20 °C.

https://doi.org/10.1021/acs.jproteome.1c00443
J. Proteome Res. 2021, 20, 4787−4800

Bacterial culture (BC) was prepared from a midlog culture of *Berkholderia cenocepacia* that was pelleted and washed twice with water and stored at −80 °C. Cells were lysed by boiling in 100 mM HEPES pH 8.0 and 2% SDS for 10 min. Cell debris was pelleted by centrifugation at 15,000 ×g for 10 min.

Labeling of Proteins with ProMTag

Proteins were labeled in buffer containing 100 mM HEPES pH 8.0, 1% SDS, and 200 mM NaCl at a protein concentration of 1−2 mg/mL. ProMTag (30 mg/mL) was added at a 1- to 10-fold mass excess of ProMTag over protein. The reaction mixture was incubated on the benchtop at room temperature for 30 min.

A fluorescamine assay was utilized to determine the extent of labeling of primary amines. A standard curve was constructed using 0−20 μg of the corresponding unlabeled protein or proteome sample. Each protein standard and ProMTag-labeled sample were diluted to a final volume of 200 μL with 100 mM HEPES pH 8.0. Then, 50 μL of a fluorescamine solution (3 mg/mL fluorescamine in acetone) was added to each sample. Samples were vortexed briefly, and 200 μL of each sample was transferred to a black polystyrene flat bottom Greiner CELLSTAR 96-well plate (Millipore Sigma). Fluorescence measurements were carried out on a Tecan Spark microplate reader (Tecan Group Ltd.) with excitation and emission wavelengths of 400 and 460 nm, respectively.

Binding of ProMTagged Proteins to TCO-Beads

To determine the ratio of TCO-beads (Click Chemistry Tools) to protein necessary to achieve complete protein capture, 20 μg of lysozyme was labeled with a 2.5-fold excess of ProMTag over protein as described above. The labeled lysozyme was added to varying amounts of TCO-beads in slit-bottom (SB) tubes (Impact Proteomics). These tubes are designed with a small slit in the bottom that allows for the passing of liquid with very little dead volume, but retains the solid resin. Binding to TCO-beads was carried out at room temperature with gentle rotation for 15 min. The flowthrough was collected by nesting the SB-tube in a 1.5 mL collection tube and then centrifugation briefly (<5 s) in a benchtop centrifuge. The TCO-beads and any proteins bound to the beads were retained in the SB-tube. A fraction of each flowthrough was run on a 4−20% SDS-PAGE gel (BioRad) for 1 h at 120 V followed by silver staining.

To determine how long labeled proteins must be incubated with TCO-beads to allow for complete protein capture, proteins were first labeled with a DIGE CyDye Cy3-NHS minimal dye (Cytiva, 10−20 μM) in buffer containing 100 mM HEPES pH 8.0, 1% SDS, and 200 mM NaCl at a protein concentration of 1−2 mg/mL for 15 min at room temperature. Proteins were then labeled with ProMTag as described above at a ratio of 2.5 μg of ProMTag per microgram of protein. After ProMTag labeling, the proteins were brought to a final concentration of 1 mg/mL with 100 mM HEPES pH 8.0, 1% SDS, and 200 mM NaCl.

Cy3-ProMTag-proteins were added to 10 μL of packed TCO-agarose beads that had been equilibrated with 100 mM HEPES pH 8.0 and 10% ACN in SB-tubes. Binding of labeled proteins to TCO-beads was carried out at room temperature with gentle rotation for up to 60 min. After the appropriate incubation time, any unbound protein was collected by brief centrifugation.

Binding of labeled proteins to TCO-beads was assessed by running the effluents and a load control on a 4−20% SDS-PAGE gel (BioRad) for 1 h at 120 V. Fluorescence images were acquired using a custom-built imager.[38] The amount of unbound protein remaining in the effluent was quantified according to the pixel intensity of the fluorescence images using ImageJ with the protein amount being calculated as a percentage of the load.[38]

SDS Quantification after Wash Steps

Yeast lysate (50 μg) was labeled with a 2.5-fold excess of ProMTag over protein as described above. The labeled lysate was brought to a final volume of 50 μL with 100 mM HEPES pH 8.0, 2% SDS before being added to 25 μL of packed TCO-beads in a SB-tube. Binding to TCO-beads was carried out at room temperature with gentle rotation for 15 min. The flowthrough and washes were collected, and the beads were then extensively washed, with each wash being collected separately, as follows: one wash with 200 μL 100 mM HEPES pH 8.0, 250 mM NaCl, and 10% ACN, two washes with 200 μL 100 mM HEPES pH 8.0, and 75% ACN, one wash with 200 μL 100 mM HEPES pH 8.0, 10% ACN, and two washes with 200 μL ultrapure water. Proteins were released from the TCO-beads by addition of 40 μL of 100 mM formic acid (FA) followed by a 15 min incubation at room temperature with gentle rotation. The released protein was treated with MT-Trypsin, as will be described below, and the peptide eluate was assayed for SDS contamination.

The SDS concentration in each control, flowthrough, wash, and peptide eluate sample was quantified with a Stains-All assay.[39] Briefly, 10 μL of each sample was transferred to a clear polystyrene flat bottom Corning 96-well plate (Sigma-Aldrich). Next, a Stains-All working solution was prepared by diluting 50 μL of a stock Stains-All solution in 900 μL ultrapure water and 50 μL formamide. The Stains-All stock solution was composed of a 1 mg/mL Stains-All solution in 50% isopropanol. Next, 200 μL Stains-All working solution was added to each sample, and absorption at 438 nm was measured on a Tecan Spark microplate reader.

Relative Drop-Size Assay for Detergent Content

To assess commonly used detergents such as CHAPS and IGEPAL CA-360 (also known as NP-40, herein referred to as IgePAL), we adapted a drop-geometry assay based on placing drops on a hydrophobic surface and measuring the wettability of the drop.[40] Detergent-free drops round up on hydrophobic surfaces, while the addition of detergent or surfactants allows the drop to spread on the surface. Here, we placed a 20 μL drop on a fresh piece of parafilm mounted on a black acrylic sheet. An image of the drop was captured with a video camera. This image was captured within 10 s of placing the drop to avoid any evaporation artifacts. Three input solutions were prepared: 1% SDS, 0.1% IgePAL, and 1% CHAPS. Because the TCO-bead elution solution contains 0.1 M FA, standard curves of 10-fold dilution series in elution solution were created for each detergent. Mock cleanup experiments were performed by exposing 25 μL of packed TCO-beads to 50 μL of the input detergent solutions. The flowthrough from these experiments was collected. The TCO-beads were washed with the same protocol as described above for SDS removal. The beads were then eluted with 50 μL of 100 mM FA. The drop sizes of flowthrough and elution fractions were measured in duplicate. The drop sizes were reported relative to 100 mM FA.

https://doi.org/10.1021/acs.jproteome.1c00443
J. Proteome Res. 2021, 20, 4787−4800

Elution of Proteins from TCO-Beads

Proteins were labeled and bound to TCO-beads, and the TCO-beads were washed extensively as described above. Proteins were eluted from the beads by addition of 20 μL of elution buffer (100 mM FA, 1.5% SDS) followed by incubation with gentle rotation at room temperature for 60 min. The eluate was collected by brief centrifugation, and the TCO-beads were washed three additional times by the addition of 20 μL elution buffer followed by a brief vortex and centrifugation with no further incubation. Elution of proteins from TCO-beads was assessed by sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS-PAGE) as described above.

MS of Lysozyme after ProMTag Reversal

To show that reversal of the ProMTag leaves proteins in their original, unmodified state, purified lysozyme that had been dialyzed against water were either made up as a 1 mg/mL solution in 100 mM FA (the control solution) or as a 1 mg/mL solution in 100 mM HEPES pH 8.0. The latter solution (200 μg) was ProMTagged, bound to TCO-beads, washed, and eluted as described above with 100 mM FA. The control solution and eluate were analyzed as intact proteins by MS (LTQ-XL) at Carnegie Mellon University's Center for Molecular Analysis. The control solution, ProMTagged lysozyme, and eluate were also compared using SDS-PAGE. Because the pH during SDS-PAGE is maintained around a pH of 8.0, mass changes after labeling and after elution can be observed.

Whole Proteome Sample Cleanup Using the ProMTag Workflow

The TC cell lysate (200 μg) was diluted with 100 mM HEPES pH 8.0 and 0.1% SDS to a final volume of 200 μL. ProMTag (16.7 μL of 30 mg/mL, 945 nmol) was added and incubated at room temperature for 30 min. The ProMTagged lysate was added to an SB-tube containing 75 μL of packed TCO-beads that had been previously equilibrated with 100 mM HEPES pH 8.0, 10% ACN and then incubated for 30 min at room temperature with end-over-end mixing. The unbound material was removed by brief centrifugation. The TCO-beads were resuspended in 75 μL of wash buffer (WB) (7 M urea, 2 M thiourea, 4% CHAPS, and 10 mM DTT), and 1 μL of 1 mM DIGE CyDye Cy3-NHS in dry DMF was added and incubated for 15 min with gentle mixing. The fluorescent labeling reaction was quenched by adding 2 μL of 5 M Methylamine-HCl and 10 mM HEPES pH 8.0 followed by a 30-minute incubation. The TCO-beads were washed three times with 300 μL WB containing 100 mM HEPES pH 8.0. The TCO-beads were washed four times with 300 μL of a solution containing WB containing 100 mM HEPES pH 8.0 combined with ACN at a 1:1 ratio. The TCO-beads were then washed three times with 300 μL WB. The clean proteome sample was eluted in two 50 μL aliquots of WB containing 100 mM FA, where each elution aliquot was incubated for 15 min prior to centrifugation. 2-DE was performed as described previously.[41]

To demonstrate that the ProMTag does not interfere with top-down proteomes, 2-DE of 100 μg of BC lysate protein was cleaned up and electrophoresed as described in the preceding paragraph. Protein spots were excised from the gel using a custom-built fluorescence imager with a robotic gel cutting tool.[38] The gel plugs were equilibrated in 0.1% acetic acid and sent to Stanford University Mass Spectrometry (SUMS) for analysis. The proteins were prepared for MS by the standard in-gel digestion protocol.

Synthesis of MT-Trypsin

To prepare MT-Trypsin, 3 mg of TPCK-Trypsin (Thermo Fisher) was dissolved in binding buffer (BB) (50 mM HEPES pH 8.0, 500 mM NaCl) to a final concentration of 2 mg/mL. To remove unwanted Trypsin autolysis fragments and inactive Trypsin, functionally intact Trypsin was bound to Benzamidine Sepharose 4 Fast Flow (high sub) beads (GE Healthcare), which had been equilibrated with BB. The loading ratio was 60 μL of Benzamidine Sepharose beads per 1 mg of Trypsin, and the binding was performed at 4 °C with end-over-end rotation for 30 min. The beads were pelleted, and the supernatant was removed. The beads were washed four times with BB. After each addition of BB, the beads were incubated for 5 min at room temperature with end-over-end rotation.

To elute cleaned-up TPCK-Trypsin, elution buffer (EB) (50 mM HEPES pH 8.0, 500 mM NaCl, and 20 mM benzamidine HCl) was added to the beads followed by a 30-min incubation at room temperature with end-over-end rotation. The eluate was collected by brief centrifugation. The elution step was repeated once, and the eluates were pooled. The concentration of the recovered TPCK-Trypsin was determined by a BCA assay.

The TPCK-Trypsin was labeled with methyltetrazine-PEG12-NHS (BroadPharm) by adding two 5 μL aliquots of 100 mM mTet-PEG12-NHS per 1 mg of TPCK-Trypsin, each followed by incubation on ice for 30 min. The labeled MT-Trypsin was then brought to 1.4 mg/mL with BB and dialyzed against 0.5 M ammonium bicarbonate overnight in a 10 K molecular weight cutoff Slide-A-Lyzer dialysis cassette (Thermo Fisher). MT-Trypsin stock was stored at −80 °C, with smaller aliquots stored at 4 °C. MT-Trypsin retained >90% activity over a 6-month period at 4 °C, as tested with a p-toluene-sulfonyl-L-arginine methyl ester assay (data not shown).[42]

Testing Capture and Activity of MT-Trypsin

To assess the binding of MT-Trypsin to TCO-beads, 20 μL of MT-Trypsin was added to 10 μL of packed TCO-beads that had been equilibrated with 100 mM HEPES pH 8.0, 10% ACN in an SB-tube and incubated at room temperature with gentle rotation for 15 min. The unbound MT-Trypsin was collected into a 1.5 mL tube by brief centrifugation. Binding to TCO-beads was assessed by SDS-PAGE as described above.

The digestion efficiency of MT-Trypsin was examined by first labeling 10 μg of yeast lysate with Cy3-NHS as described above with one modification; the labeling reaction was incubated for 45 min instead of 15 min to ensure that the all the Cy3-NHS was hydrolyzed to avoid labeling the MT-Trypsin. Next, 10 μg of MT-Trypsin was added to the labeled lysate, and the digestion reaction was incubated for 120 min at 37 °C. Digestion was assessed at various time points by SDS-PAGE. Images were acquired and processed as described above.

Putting the two previous steps together, 20 μg of yeast lysate was reduced with 10 mM DTT for 30 min at 56 °C and then alkylated with 20 mM IAA for 30 min at room temperature in the dark. ProMTag labeling of the yeast lysate, binding to TCO-beads, and washing of the TCO-beads were carried out as described above. The proteins were released from the TCO-beads by addition of 20 μL of 100 mM FA followed by incubation at room temperature with gentle rotation for 15 min. MT-Trypsin was then added to the proteins at a ratio of one-to-two mass ratio of MT-Trypsin to ProMTag-lysate.

https://doi.org/10.1021/acs.jproteome.1c00443
J. Proteome Res. 2021, 20, 4787−4800

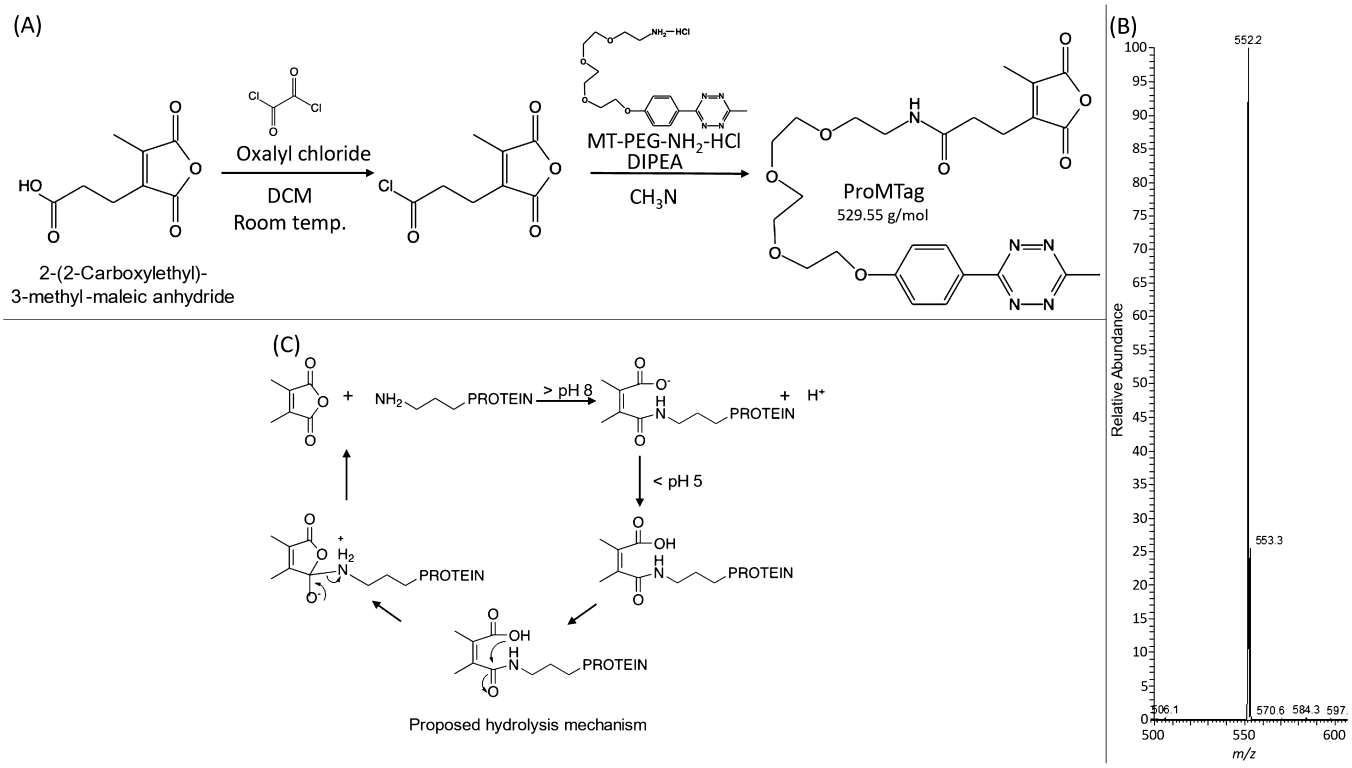


Figure 2. Synthesis of ProMTag and confirmation by MS. (A) ProMTag was synthesized from starting materials and methyltetrazine-PEG4-HCl. (B) Synthesis was confirmed using MS, which showed a major peak at 552.2 *m/z*, representing ProMTag with a sodium adduct and a minor peak at 553.3 *m/z* representing protonated ProMTag with a sodium adduct. (C) Proposed mechanism for the reversible coupling of dialkyl maleic anhydride to protein.

Digestion was carried out at 37 °C for the indicated amount of time, after which the supernatant was collected by brief centrifugation. Any unbound protein/peptides remaining in the supernatant were assessed by SDS-PAGE and silver staining, as described previously.

MS Analysis of Peptides after Precipitation/in-Solution Digestion or ProMTag Cleanup

For in-solution digestion cleanup, eight volumes of ice-cold acetone were added to yeast lysate followed by addition of trichloroacetic acid to a final concentration of 10%. Protein precipitation was carried out overnight at −20 °C. Samples were centrifuged at 15,100 ×*g* for 10 min at 4 °C, washed once with cold acetone, and centrifuged again. The pellet was allowed to air-dry for 5 min before being suspended in 50 mM ammonium bicarbonate. Proteins were reduced and alkylated as described above. Trypsin Gold (Promega) was added at an enzyme to protein ratio of 1:50, and samples were digested overnight at 37 °C. FA was added to a final concentration of 0.5% FA, and samples were dried in a speed vac.

For ProMTag cleanup to produce peptide fragments, yeast lysate was reduced, alkylated, and labeled with ProMTag as described above. The labeled proteins were added to 20 μL of packed TCO-beads in an SB-tube and incubated at room temperature with gentle rotation for 30 min. The TCO-beads were then washed as follows: one wash with 200 μL 100 mM HEPES pH 8.0, 250 mM NaCl, 10% ACN, two washes with 200 μL each of 100 mM HEPES pH 8.0, 75% ACN, one wash with 200 μL 100 mM HEPES pH 8.0, 10% ACN, and two washes with 200 μL each of ultrapure water. The proteins were released from the TCO-beads by the addition of 40 μL 100 mM FA followed by a 15-minute incubation at room temperature with gentle rotation. MT-Trypsin was then added at a 1:1 protein to enzyme ratio and digested at 37 °C for 1 h. The resulting peptides were recovered by briefly centrifuging the SB-tube. To recover any remaining peptides, 40 μL of 100 mM FA was added to the beads followed by a 15-minute incubation at room temperature with gentle rotation and recovery of the supernatant by centrifugation. Yields were determined using Pierce Quantitative Fluorometric Peptide Assay (ThermoFisher Scientific). Peptides were dried in a vacuum concentrator and sent to SUMS for analysis.

In a typical MS experiment, peptides were reconstituted in 2% aqueous ACN prior to analysis. MS experiments were performed using liquid chromatography (LC) using either an Acquity M-Class UPLC (Waters) followed by MS using an Orbitrap Fusion Tribrid MS (Thermo Scientific) or a nanoAcquity UPLC (Waters) connected to an Orbitrap Q Exactive HF-X MS (Thermo Scientific). A flow rate of 300 nL/min was used, where mobile phase A was 0.2% (v/v) FA in water and mobile phase B was 0.2% (v/v) FA in ACN. Analytical columns were either prepared in-house, with an internal diameter of 100 microns packed with NanoLCMS solutions 1.9 μm C18 stationary phase to a length of approximately 20 or 50 cm uPAC columns (Pharmafluidics), for the Fusion and HF-X, respectively. Peptides were directly injected into the analytical column using a gradient (2−45% B followed by a high-B wash) of 90 min. MS was operated in a data-dependent fashion using collision-induced dissociation (CID) fragmentation for MS/MS spectra generation collected in the ion trap on the fusion or higher energy collision dissociation (HCD) fragmentation on the HF-X.

https://doi.org/10.1021/acs.jproteome.1c00443
J. Proteome Res. 2021, 20, 4787−4800

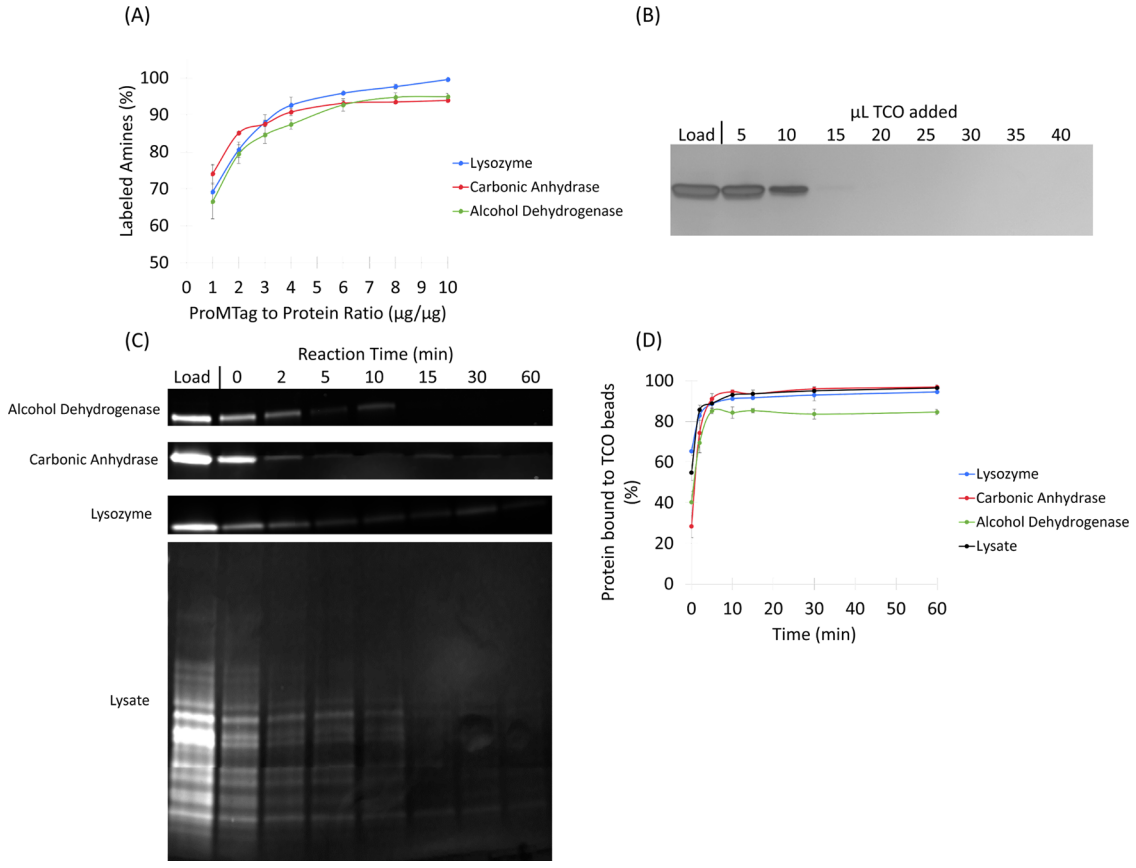


Figure 3. Protein Labeling with ProMTag and Binding of ProMTag-proteins to TCO-beads. (A) Lysozyme, carbonic anhydrase, and ADH were incubated with increasing amounts of ProMTag. The extent of labeling of primary amines was assessed using a fluorescamine assay. (B) Titration of TCO-beads to optimize ProMTagged lysozyme binding. (C) Composite image of the time course of ProMTagged lysozyme, carbonic anhydrase, ADH, and yeast lysate binding to TCO-beads. (D) Quantification of ProMTagged protein shown in (C).

For data analysis, the RAW data files were checked using Preview (Protein Metrics) to verify calibration and quality prior to further analysis. Data were then processed using Byonic v3.7.13 (Protein Metrics) to identify peptides and infer proteins using the 2019−12 release of the UniProt Knowledgebase downloaded on 02/24/2020 containing 6078 yeast proteins as the *S. cerevisiae* database, including isoforms and concatenated with common contaminant proteins. Proteolysis with Trypsin was assumed to be semispecific with up to two missed cleavage sites and allowing for common modifications. Precursor mass accuracies were held within 12 ppm with fragment ions held within 0.4 Da (for CID) and 12 ppm (for HCD). Proteins were held to a false discovery rate of 1%, using standard approaches described previously.[43] The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the data set identifier PXD027799.[44]

RESULTS AND DISCUSSION

Synthesis of ProMTag

ProMTag was synthesized by covalently linking 2-(2-carboxyethyl)-3-methyl-maleic anhydride with methyltetrazine-PEG4-amine (MT-PEG-NH_2-HCl) (Figure 2A). The final ProMTag product was purified by MPLC on RP-18 with single fractions analyzed by UPLC. The structure was confirmed by MS and NMR.

There were two elution peaks containing a potential ProMTag product after UPLC. 1H-NMR analysis showed that both peaks were the ProMTag product, with different elution times because of two isoforms (possibly because of the anhydride or di-acid states of the CDM; or a mixed anhydride joining two ProMTags, CDM-to-CDM) of the product. 1H NMR (500 MHz, Chloroform-d) δ 11.64 (s, 1H, COOH), 8.57−8.49 (m, 2H), 7.13−7.05 (m, 2H), 6.85 (s, 1H,NH), 4.30−4.22 (m, 2H), 3.98−3.90 (m, 2H), 3.84−3.74 (m, 2H), 3.79−3.70 (m, 2H), 3.70−3.59 (m, 4H), 3.56 (t, J = 5.0 Hz, 2H), 3.42 (q, J = 5.1 Hz, 2H), 3.08 (d, J = 1.3 Hz, 3H), 2.79 (t, J = 7.2 Hz, 2H), 2.61 (t, J = 7.2 Hz, 2H), 2.09 (s, 3H). 13C NMR (126 MHz, CDCl3) δ 172.41, 166.59, 165.87, 165.77, 163.73, 162.32, 142.62, 142.01, 129.78, 124.43, 115.21, 77.32, 77.06, 76.81, 70.76, 70.47, 70.32, 70.07, 69.59, 69.12, 67.52, 39.80, 32.83, 21.00, 20.55, 9.56. The MS data showed two peaks: a major peak at 552.2 *m/z* and a minor peak at 553.3 *m/z* corresponding to the ProMTag and protonated ProMTag final product, respectively (Figure 2B). The difference between the ProMTag molecular weight and actual weight shown via MS is due to a sodium adduct formed with both products. Both fractions were tested for their ability to react with primary amines and were shown to be identical in their reaction, so the two fractions were combined and used as one product for the remainder of this study (data not shown).

https://doi.org/10.1021/acs.jproteome.1c00443
J. Proteome Res. 2021, 20, 4787−4800

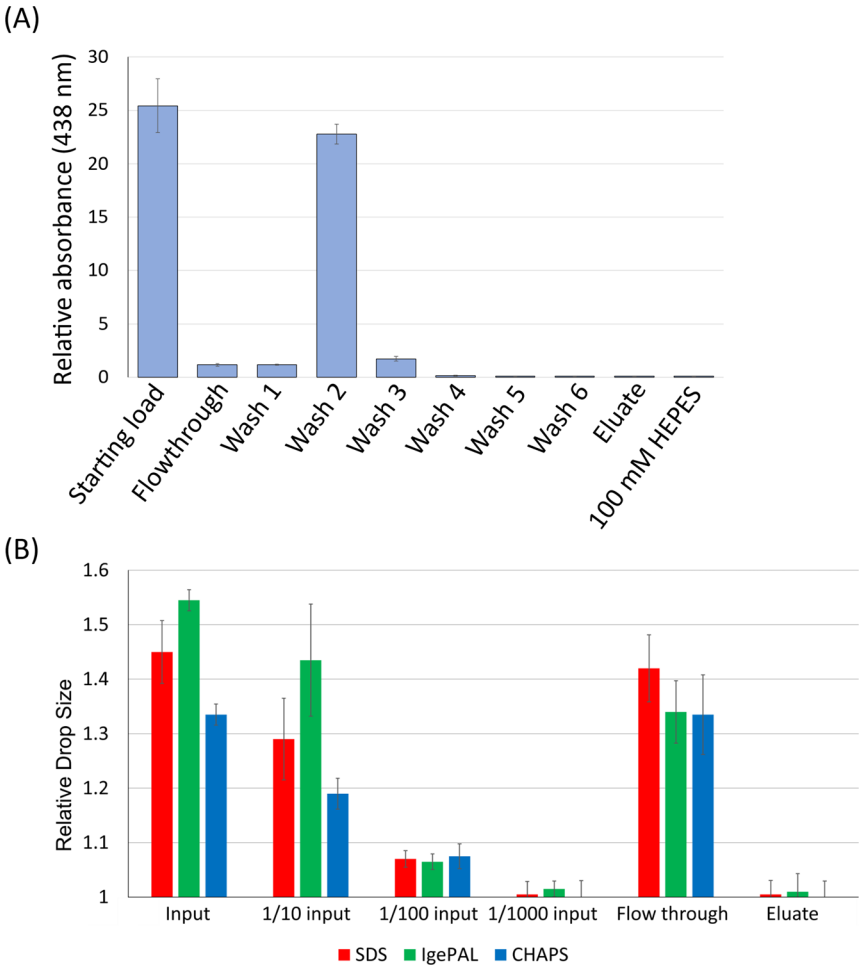


Figure 4. Detergent removal during wash steps. (A) Assessment of SDS removal using the Stains-all assay. (B) Assessment of SDS, IgePAL, and CHAPS removal using a drop-size assay.

ProMTag Covalently Reacts with Primary Amines on the Surface of Proteins

To assess the ability of ProMTag to label proteins, a fluorescamine assay was used to determine the number of free primary amines after reacting proteins with increasing amounts of ProMTag (Figure 3A). Three model proteins, lysozyme (14 kDa), carbonic anhydrase (30 kDa), and ADH (150 kDa), were used to determine the protein labeling efficiency of ProMTag. All three proteins reached >90% amine labeling with a mass ratio of 6 mg ProMTag per mg of input protein. Even at a 1:1 ProMTag to protein ratio (wt/wt), the majority of the primary amines on the surface of the proteins were labeled. Theoretically, each protein needs only a single ProMTag to be captured, so the labeling extent in these experiments demonstrated the feasibility of ProMTag coupling to proteins for efficient capture and cleanup.

ProMTag-Labeled Proteins Rapidly Bind to TCO-Beads

The ratio of TCO-beads to protein necessary to achieve complete protein capture was assessed by adding 20 μg of ProMTagged lysozyme to increasing amounts of TCO-beads and incubating for 1 h. The eluate was then analyzed using SDS-PAGE followed by silver staining to determine the amount of lysozyme that remained in solution (Figure 3B). Complete binding was observed after incubation with 20 μL of TCO-beads. Therefore, a ratio of 1 μL of TCO-beads per 1 μg of labeled protein was used for most subsequent experiments.

The rate of ProMTag-protein capture by TCO-beads was assessed by measuring the amount of protein left in solution after various times of exposure to TCO-beads. To quantify protein abundance, the model proteins were first minimally labeled with Cy3-NHS DIGE-Cy dye and then labeled with ProMTag at 2.5 μg/μg of protein. While Cy3-NHS DIGE-Cy dye also reacts with primary amines, this minimally labeling dye was designed to react with less than 2% of available primary amines, leaving many available amines for the ProMTag to react with. After labeling, the protein capture reaction was initiated by the addition of TCO-beads. At various times thereafter, the beads were rapidly pelleted, and the supernatant was withdrawn, run on an SDS-PAGE gel, and analyzed by fluorescence imaging (Figure 3C). On average, 88% of each model protein was bound after just a 10-minute incubation with TCO-beads (Figure 3D). The proteins had small differences in binding, with carbonic anhydrase showing the best binding at 94% after 10 min and ADH showing the lowest binding after 10 min at 88% (Figure 3D). Binding only increased 2−4% from 10 to 60 min. The binding curves demonstrate that ProMTag-labeled proteins of various sizes can be rapidly and efficiently bound to TCO-beads.

To determine that ProMTag:TCO-bead capture was similarly effective for the entire proteome, the previous labeling and capture workflow was repeated using a whole protein yeast-cell lysate (Figure 3C). We observed that 93% of

https://doi.org/10.1021/acs.jproteome.1c00443
J. Proteome Res. 2021, 20, 4787−4800

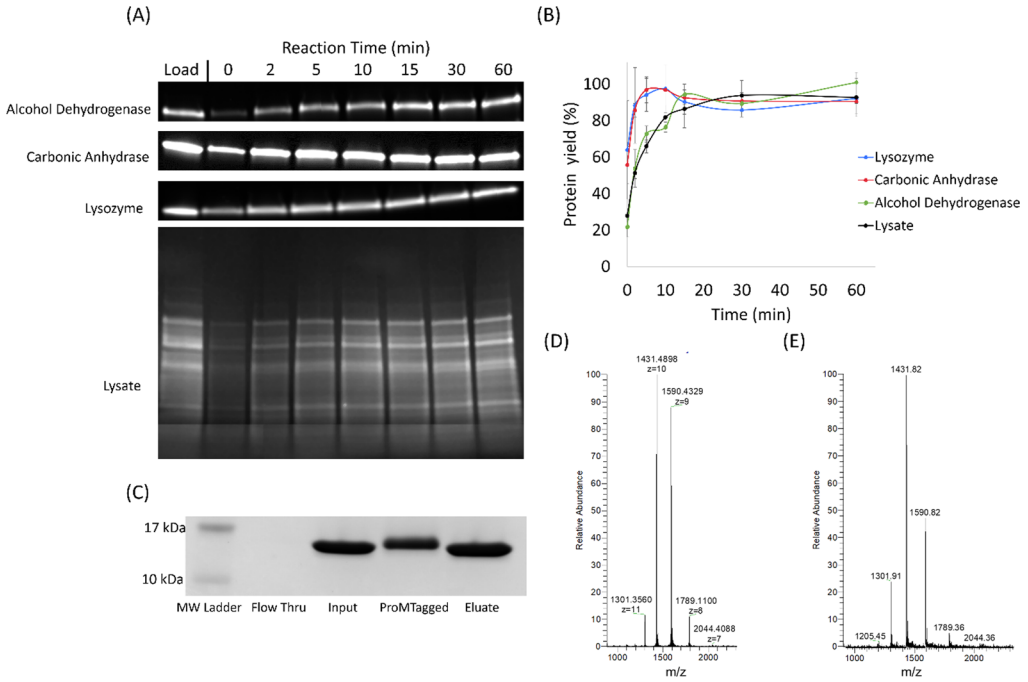


Figure 5. Elution of individual proteins and whole proteomes from TCO-beads. ProMTag-labeled proteins were captured on TCO-beads, washed to remove SDS, and eluted from the beads by the addition of a FA solution. (A) Shows a composite image of lysozyme, carbonic anhydrase, ADH, and yeast lysate elution from TCO-beads at various times. (B) Shows the fluorescence quantification of these elution images normalized relative to the input protein sample. (C) Shows a silver-stained SDS-PAGE of lysozyme with and without being ProMTagged. (D) MS analysis of intact lysozyme without being ProMTagged. (E) MS analysis of ProMTagged lysozyme after ProMTag reversal.

the proteome was able to bind to the TCO-beads in as little as 10 min, reaching a maximum of 96% bound by the end of the one-hour incubation (Figure 3C,D). This showed that an entire proteome could be successfully labeled with ProMTag and bound to TCO-beads quickly and efficiently.

Common Detergents Are Completely Removed during Washing of TCO-Beads

SDS is a ubiquitous protein denaturant that is the preferred reagent for the first step of proteome sample solubilization. Unfortunately, SDS greatly interferes with MS and 2DE, so it must be removed from proteome samples after lysis. The covalent linkage of ProMTagged protein and TCO-beads allows for extensive and varied washing of the bead-bound proteins to ensure the complete removal of SDS.

Yeast-cell lysate was labeled with ProMTag in the presence of 2% SDS and incubated with the TCO-beads for 15 min. The flowthrough was collected, and the beads were washed with a series of various WBs. The proteins were then released from the TCO-beads and digested with MT-Trypsin, and the eluate was collected. The presence of SDS in each sample was assessed via a Stains-All assay (Figure 4A).[41] The flowthrough contained little SDS, as did the first wash, which contained high salt to wash away nucleic acids. The second and third washes contained high concentrations of ACN to enable SDS removal. The second wash successfully removed the majority of the input SDS, with trace amounts of SDS in the third wash. Three subsequent washes and the eluate contained no detectable SDS. Thus, one high salt wash and two high ACN washes are sufficient to remove all SDS in a series of <10 second centrifugation steps.

IgePAL and CHAPS are common detergents used in proteomics research. To assess their removal (as well as SDS), we adapted a drop-size surface tension assay to measure the amount of detergent remaining in the final 100 mM FA eluate from TCO-beads.[40] Twenty microliter drops were placed on a parafilm-covered surface, and the diameter of the droplet was measured from above. A 10-fold dilution series shows that a 100-fold dilution of the original detergent solution was measurably different from the elution solution, 100 mM FA. A 1000-fold-dilution was indistinguishable from the elution solution (Figure 4B). The majority of the input detergent was found in the TCO-bead flowthrough. All three detergents were undetectable after using the same washing scheme as described for SDS.

Rapid Elution of ProMTagged Proteins from TCO-Beads

Following thorough washing and removal of contaminants, the next step in the ProMTag workflow is release of proteins from the TCO-beads by reversal of the CDM linkage. This reversal step is carried out in 100 mM FA. To assess the rate and efficiency of protein release, the three model proteins were labeled with Cy3-NHS and ProMTag, bound to TCO-beads, and washed as described previously. They were then eluted in 100 mM FA at various time points over the course of an hour (Figure 5A,B). For each of these time points, the total amount of eluted protein was analyzed by SDS-PAGE and fluorescence imaging. Proteins were quantified relative to a loading control. More than 90% of each of the three model proteins was released from the beads. The same experimental workflow was repeated with an intact protein yeast-cell lysate. While elution for the intact proteome was slightly slower than the individual proteins, after 15 min 86% of the proteome had been recovered, reaching a maximum of 93% after 30 min. This demonstrated that both individual proteins of various sizes and intact proteomes can be successfully labeled with ProMTag, captured with TCO-beads, washed, and released with very high yields.

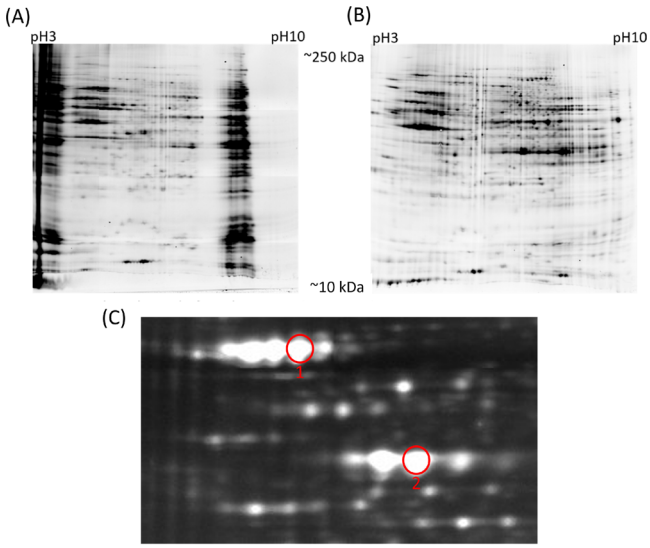
https://doi.org/10.1021/acs.jproteome.1c00443
J. Proteome Res. 2021, 20, 4787−4800

One concern in using a reversible tag would be if the reversal was incomplete or left residual modifications on the original protein that would interfere with downstream analysis. Methods like MS and 2-DE are very sensitive to mass and/or charge changes, and any modifications left on the proteins after cleanup could negatively affect analysis. To show that the ProMTag is fully reversible, lysozyme was ProMTagged, bound to TCO-beads, washed, and released as previously stated and analyzed using SDS-PAGE and MS. The complete capture of the ProMTagged lysozyme was demonstrated by its disappearance in the TCO-bead flowthrough (Figure 5C, lane 1). Lysozymes before ProMTagging, after ProMTagging, and after reversal of the ProMTag were run on an SDS-PAGE gel to analyze mass changes that occur (Figure 5C, lanes 2−4, respectively). After tagging with ProMTag, a slight mass shift for lysozyme was observed as expected, but the mass shift was not observed after reversal of the ProMTag. MS analysis of intact protein was performed on lysozymes before (Figure 5D) and after ProMTagging (Figure 5E). If ProMTag was not completely removed from the surface of lysozymes, we would expect to observe mass differences from the original lysozyme in increments of 552 Da (the mass of ProMTag). MS of lysozymes showed that the mass spectra before and after ProMTagging were identical, and no mass shifts different from the native protein were observed. Deconvolution of the spectra showed a mass of 14,308.5 Da for untreated and cleaned up lysozymes. No molecular species with a mass of 14,860 were observed, indicating that no single ProMTag adducts were present.

Cleanup of an Intact-Protein Proteome Sample Using the ProMTag Workflow

While MS of intact lysozyme showed that top-down proteomics is compatible with the ProMTag workflow in theory, we wanted to elaborate on this idea to include top-down proteomics that utilizes gel-based separation of proteoforms within a whole proteome. To understand how the ProMTag cleanup workflow could improve the resolution of a top-down proteome analytical technique, intact proteins from TC cell lysates with and without ProMTag sample cleanup were analyzed by 2DE. TC cells were lysed using an SDS-containing lysis buffer, and the sample was split in half: one half was cleaned up using the ProMTag workflow and the other half was not cleaned up. The cleaned-up sample was eluted from the TCO-beads with a urea-based buffer containing FA and applied directly to the isoelectric focusing strips. The two samples were then analyzed using 2DE (Figure 6A,B). Sample cleanup using the ProMTag workflow prevented precipitation zones on both the acidic and basic portions of the gel that are typically caused by high salt, nucleic acids, and SDS, allowing for the resolution of individual protein spots in these areas. Additionally, the cleaned-up sample showed resolution of more individual proteins throughout the pH range. Together this showed that the ProMTag cleanup workflow worked well for removal of interfering contaminants and improved resolution of an intact proteome using a top-down analytical technique.

Top-down proteomics either involves the MS analysis of intact proteins (Figure 5E) or high-resolution separation and isolation of intact proteins followed by protease digestion and MS analysis of the resulting peptides. To further demonstrate the utility of the ProMTag-TCO cleanup workflow for top-down proteomics, a cell lysate of a BC was separated by 2DE,



Spot #	Description	|Log Prob|	Total Intensity	# of spectra	# of unique peptides	Coverage %
1	UPI000052E2E4 Gro EL	133.77	31358813.6	101	53	66.85
2	UPI00002E1B9C EF-Tu	267.18	1555891389.8	461	92	74.24

Figure 6. 2DE of intact lysate before and after cleanup with ProMTag and MS to identify proteins from 2DE gel. (A) Shows a 2DE gel of TC lysate that was not cleaned up. (B) Shows a 2DE gel of a cleaned-up TC sample. Both samples were fluorescently tagged with Cy3-NHS. These images are shown in reverse contrast. (C) Shows a segment of a 2DE gel of Cy3-tagged bacterial cell culture where two prominent protein spots were excised, in-gel digested with Trypsin, and analyzed by LC−MS/MS. The table below shows the MS analysis of the excised spots.

and selected proteins were excised, digested, and analyzed by LC−MS/MS (Figure 6C). The 2DE gel showed well-resolved protein spots, and the two indicated spots in Figure 6C were able to be identified via MS. The top MS hit for spot 1 revealed this protein to be Gro EL, and the top MS hit for spot 2 was EF-Tu. The MS analysis showed over 66% coverage with 53 unique peptides identified for Gro EL and 74% coverage with 92 unique peptides identified for EF-Tu. Taken together, the whole protein MS both in-solution and after separation on a 2DE gel shows that the ProMTag workflow can be used for top-down protein analysis and is very versatile in its compatibility with multiple technologies.

Modification of Trypsin for Protein Digestion in the ProMTag Workflow

The majority of proteome analyses are performed at the level of digested peptides. To make the ProMTag cleanup workflow amenable to generating cleaned-up, protease-digested samples, we developed a new, modified version of Trypsin, referred to as MT-Trypsin. This was created by modifying the solvent available primary amines of Trypsin with NHS-PEG-methyltetrazine, which made the Trypsin resistant to autolysis, more stable, and capturable by TCO-beads. This allows for shorter digestion times at higher Trypsin concentrations without the risk of overwhelming the final sample with Trypsin autolysis peptides. Any autolysis products that may arise will already be modified for capture on the TCO-beads and will be prevented from being eluted in the final sample.

We first assessed the efficacy of MT-Trypsin capture by TCO-beads. MT-Trypsin was incubated with TCO-beads for 15 min, and then the unbound MT-Trypsin fraction was assessed by SDS-PAGE. Silver staining revealed that there was

https://doi.org/10.1021/acs.jproteome.1c00443
J. Proteome Res. 2021, 20, 4787−4800

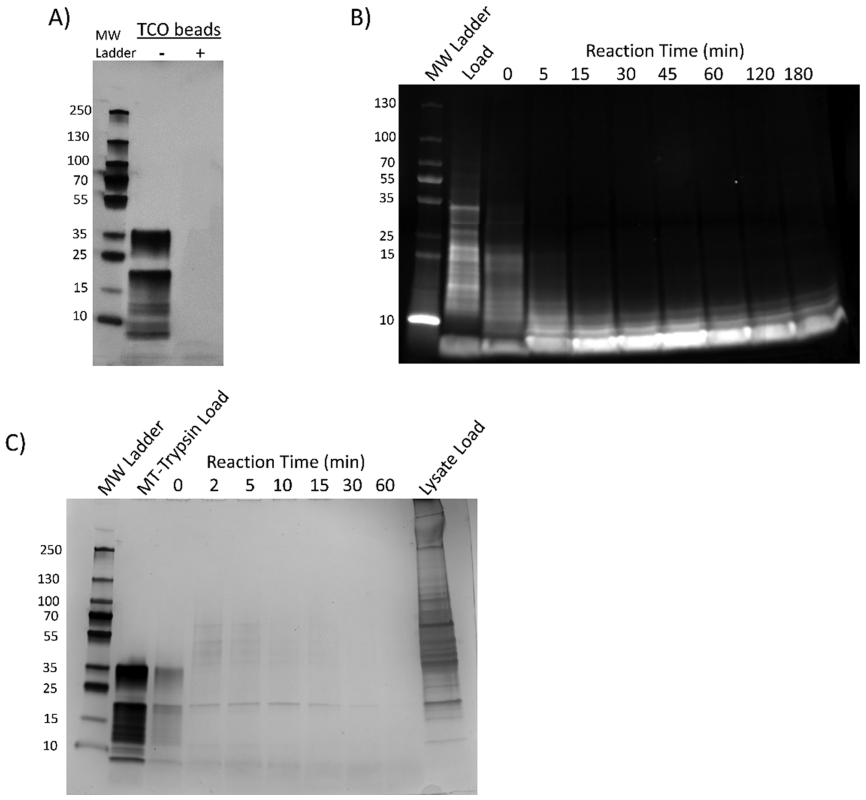


Figure 7. MT-Trypsin characterization. (A) MT-Trypsin was incubated with TCO-beads for 15 min and then the supernatant was analyzed by SDS-PAGE and silver staining. (B) Fluorescently tagged yeast lysate was incubated with MT-Trypsin at a 1:1 mass ratio over a 180-minute period. (C) MT-Trypsin digestion time course of ProMTagged-yeast lysate captured on TCO-beads. MT-Trypsin input load is shown on the left and lysate input load on the right.

no detectable MT-Trypsin in the unbound fraction (Figure 7A). Next, MT-Trypsin was combined with a Cy3-labeled intact yeast lysate at a 1:1 ratio (wt/wt), and digestion of proteins into peptides was assessed by SDS-PAGE over a 3-hour time course. Almost all protein bands had collapsed into less than ∼11 kDa peptide bands on the gel after the 15-minute time point, with little further degradation thereafter (Figure 7B). This demonstrated that modified MT-Trypsin retained its ability to digest protein. Finally, MT-Trypsin was incorporated into the ProMTag cleanup workflow by adding it to the cleaned-up sample after the ProMTag linkage had been reversed. At this step, MT-Trypsin was added to the sample at a 1:1 protein to enzyme ratio. The supernatant was assessed for MT-Trypsin capture on the TCO-beads over the course of an hour (Figure 7C). The zero-time point was taken immediately after the addition of MT-Trypsin and mixing, and much of the MT-Trypsin was already bound to the TCO-beads. At this time point, most of the intact protein from the lysate remained bound to the beads because of the lack of solubilizing agents and the hydrophobic nature of intact proteins. By 2 minutes, very little MT-Trypsin remained in solution and a small amount of digested lysate proteins appeared in the supernatant. After 60 minutes, no protein was detectable by silver stain. Thus, cell lysates prepared by SDS lysis were efficiently cleaned-up and digested into peptides using the ProMTag workflow with the addition of a digestion step using MT-Trypsin.

MS Analysis of Peptides after the ProMTag Cleanup Workflow

Finally, all of the aforementioned steps were put together to cleanup and digest a yeast-cell lysate for MS analysis. Yeast proteins were solubilized by boiling for 10 min in a lysis buffer containing 2% SDS and subsequently reduced and alkylated using DTT and iodoacetamide, respectively. The lysate was labeled with ProMTag, ProMTagged proteins were bound to TCO-beads, and the protein-ProMTag-TCO-beads were washed to eliminate nonprotein components, including SDS. The proteins were released from the ProMTag-TCO-beads by addition of FA. After a 15-minute incubation to allow reversal of the protein-CDM linkage, MT-Trypsin was added at a 1:1 protein to MT-Trypsin ratio. Digestion was performed for 1 h at 37 °C, and the resulting peptides were separated from the TCO-beads by brief centrifugation.

Sample cleanup and peptide generation were also performed using conventional precipitation and in-solution digestion for comparison. Each of these peptide preparations was performed in triplicate starting with the same initial yeast lysate. The yeast proteome peptide samples were vacuum-dried and analyzed by LC−MS/MS.

Peptide quantification of three independent sample prep replicates showed an average yield of 62.65 ± 16.7 for the conventional cleanup and 88% ± 11 for the ProMTag cleanup. On average, the conventional workflow led to the identification of 856 ± 26 proteins from 6004 ± 251 peptides and 15,196 ± 331 spectra (Figure 8A). The ProMTag workflow nearly doubled the number of identified proteins and more than doubled the number of peptide identifications, with an average

https://doi.org/10.1021/acs.jproteome.1c00443
J. Proteome Res. 2021, 20, 4787−4800

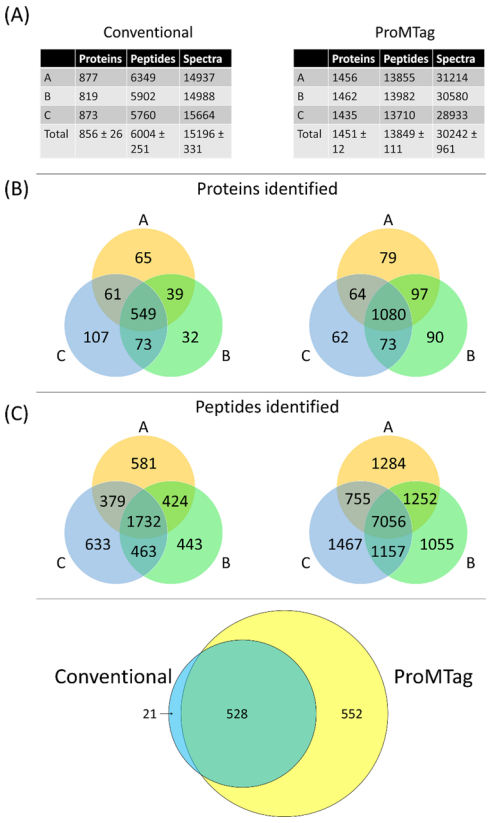


Figure 8. MS analysis of a yeast proteome processed using ProMTag cleanup. (A) Protein, peptide, and spectra counts for three independent yeast sample preparations that were processed using either the conventional method or the ProMTag cleanup. (B) Venn diagram of protein overlap for the three lysates based on the *p*-value score. (C) Venn diagrams of peptide overlap for the three lysates based on log (F) Comparison of proteins identified in all three replicates of the two different methods.

of 1451 ± 12 protein identifications from 13,849 ± 111 peptides and 30,242 ± 961 spectra (Figure 8A). Protein and peptide identification was also analyzed for reproducibility across samples. Proteins were scored and included in the comparison if they had a *p*-value ≤0.001 (Figure 8B) or at least two unique peptides (data not shown). Peptides were scored and included in the comparison if they had a log probability of ≥3 (the log *p*-value of the peptide-spectrum match) (Figure 8C) or a "good" Byonic score of 300 (the primary indicator of peptide-spectra match correctness, data not shown). Using the conventional workflow, roughly 59% of identified proteins were scored in all three replicates. This value increased to 78% for proteins found in at least two of the three replicates. The ProMTag workflow showed a significant increase in protein identification reproducibility with 71% of proteins being scored in all three replicates and 85% of proteins being scored in at least two of the three replicates. For peptides, using the conventional workflow resulted in 37% of all peptides scored in all three replicates, and an average of 64% of peptides scored in two of the three samples. Using the ProMTag workflow, peptide identification reproducibility increased significantly with 50% of all peptides being scored in all three replicates and 73% of all peptides being scored in at least two out of three replicates. Finally, each data set was compared for proteins identified in all three replicates (Figure 8F). Of the 549 proteins identified in all three conventional prep replicates, 528 were also found in all three ProMTag replicate samples. Only 21 proteins were exclusively identified in the conventional replicates. Conversely, 552 additional proteins were identified in all three ProMTag replicates. Taken together, these data show that the ProMTag cleanup workflow is compatible with MS, shows high reproducibility and sensitivity, and can identify many more proteins than a conventional precipitation/in-solution digestion workflow.

CONCLUSIONS

In this study, a new universal protein cleanup workflow was developed which utilizes the reversible click chemistry tag ProMTag for rapid and unbiased protein capture, cleanup, and elution. Using this workflow, individual proteins, protein mixtures, or whole protein lysates can be cleaned up with minimal protein loss. This workflow allows for removal of contaminants, such as SDS, as well as buffer exchange and sample desalting. The universality of this technique was further demonstrated by showing that with only one additional step, the ProMTag cleanup workflow can also be applied to the preparation of peptide samples for MS analysis. Utilizing the ProMTag cleanup workflow also led to higher rates of protein identification and reproducibility than a conventional cleanup workflow. This is the first proteome sample cleanup workflow that can be used to yield either intact proteins or tryptic peptides.

The ProMTag's dual covalent coupling moieties, one reversible and one irreversible, enable a wide range of potential uses from rapid buffer exchange to protein concentration to targeted protein capture, allowing this technology to be applied to a wide range of applications involving any protein-based samples. One key advantage of the ProMTag system is that it is nondestructive, leaving proteins, peptides, and other biomolecules intact after reversal of the CDM linkage. Thus, the proteins can be treated as gently or as harshly as the experiment requires. The click coupling of MT to TCO is rapid and extremely robust, enabling a wide range of coupling conditions.

Here, we also introduce MT-Trypsin. This stabilized, capturable Trypsin enables the use of high concentrations of MT-Trypsin for rapid digestion. The MT moiety allows for rapid and complete separation of peptide products from the enzyme. Coupling MT-Trypsin to TCO-beads obviates the need for membrane filtration, which can be slow and lossy. Finally, this single-pot proteome cleanup approach is amenable to automation. Thus, it affords a rapid, high-yield, unbiased, and high-throughput method for either top-down or bottom-up proteome sample preparation.

AUTHOR INFORMATION

Corresponding Author

Amber Lee Wilson − *Impact Proteomics, LLC., Pittsburgh, Pennsylvania 15206, United States;* orcid.org/0000-0002-4434-7802; Email: amberleewilson20@gmail.com

Authors

Stephanie Biedka − *Impact Proteomics, LLC., Pittsburgh, Pennsylvania 15206, United States*

Brigitte F. Schmidt − *JGS Research Co., Pittsburgh, Pennsylvania 15212, United States; Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States*

https://doi.org/10.1021/acs.jproteome.1c00443
J. Proteome Res. 2021, 20, 4787−4800

Nolan M. Frey − *Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States*
Sarah M. Boothman − *Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States*
Jonathan S. Minden − *Impact Proteomics, LLC., Pittsburgh, Pennsylvania 15206, United States; Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.1c00443

Author Contributions

B.F.S. contributed design and execution of the ProMTag synthesis, purification, and analysis. N.M.F. and S.M.B. contributed the 2DE analysis of tissue culture and bacterial samples. J.S.M. contributed the ProMTag cleanup of tissue culture samples and experimental design, troubleshooting, experimental interpretation, and manuscript editing. A.L.W. contributed experimental design, troubleshooting, experimental interpretation, manuscript drafting, and manuscript editing. S.B. contributed all other experiments not described above and assisted in experimental design, troubleshooting, experimental interpretation, and manuscript editing.

Notes

The authors declare the following competing financial interest(s): A patent application related to this work was filed by Impact Proteomics, LLC. Patent Application No. PCT/US19/36035 with assignors Jonathan S. Minden and Amber Lee Wilson filed 06/07/2019 and titled "Protein and Peptide Purification Methods".

■ ACKNOWLEDGMENTS

This work was supported in part by NSF PFI AIR TT Award 1700833 and NSF SBIR Phase I Award 1843332. This work was also supported in part by NIH P30 CA124435 utilizing the Stanford Cancer Institute Proteomics/Mass Spectrometry Shared Resource at the Vincent Coates Foundation Mass Spectrometry Laboratory, Stanford University Mass Spectrometry (RRID:SCR_017801). The authors thank Ryan Leib and the Stanford University Mass Spectrometry Facility for their help with mass spectrometry analysis, data analysis, and statistical analysis. Thanks to Mark Bier and Liam Dugan and Carnegie Mellon University's Center for Molecular Analysis. Additional thanks to members of the John Woolford lab at Carnegie Mellon University for yeast cells and to Camille Jacqueline at University of Pittsburgh Medical Center for tissue culture cells used for two-dimensional gel electrophoresis. Thanks to Marcel Bruchez for use of his lab's equipment and insightful discussions.

■ REFERENCES

(1) Chandramouli, K.; Qian, P.-Y. Proteomics: Challenges, Techniques and Possibilities to Overcome Biological Sample Complexity. *Hum. Genomics Proteomics* **2009**, *1*, 23904.
(2) Jiang, L.; He, L.; Fountoulakis, M. Comparison of Protein Precipitation Methods for Sample Preparation Prior to Proteomic Analysis. *J. Chromatogr. A* **2004**, *1023*, 317−320.
(3) Sielaff, M.; Kuharev, J.; Bohn, T.; Hahlbrock, J.; Bopp, T.; Tenzer, S.; Distler, U. Evaluation of FASP, SP3, and IST Protocols for Proteomic Sample Preparation in the Low Microgram Range. *J. Proteome Res.* **2017**, *16*, 4060−4072.
(4) Kuljanin, M.; Dieters-Castator, D.; Hess, D.; Postovit, L.; Lajoie, G. Comparison of sample preparation techniques for large-scale proteomics. *Proteomics* **2017**, *17*, No. 1600337.
(5) Yang, Y.; Anderson, E.; Zhang, S. Evaluation of Six Sample Preparation Procedures for Qualitative and Quantitative Proteomics Analysis of Milk Fat Globule Membrane. *Electrophoresis* **2018**, *39*, 2332−2339.
(6) Ludwig, K. R.; Schroll, M. M.; Hummon, A. B. Comparison of In-Solution, FASP, and S-Trap Based Digestion Methods for Bottom-Up Proteomic Studies. *J. Proteome Res.* **2018**, *17*, 2480−2490.
(7) Wang, W.-Q.; Jensen, O. N.; Møller, I. M.; Hebelstrup, K. H.; Rogowska-Wrzesinska, A. Evaluation of Sample Preparation Methods for Mass Spectrometry-Based Proteomic Analysis of Barley Leaves. *Plant Methods* **2018**, *14*, 72.
(8) Thomas, S.; Hao, L.; Ricke, W. A.; Li, L. Biomarker Discovery in Mass Spectrometry-Based Urinary Proteomics. *Proteomics Clin. Appl.* **2016**, *10*, 358−370.
(9) Hernandez-Valladares, M.; Aasebø, E.; Selheim, F.; Berven, F. S.; Bruserud, Ø. Selecting Sample Preparation Workflows for Mass Spectrometry-Based Proteomic and Phosphoproteomic Analysis of Patient Samples with Acute Myeloid Leukemia. *Proteomes* **2016**, *4*, 24.
(10) Lipecka, J.; Chhuon, C.; Bourderioux, M.; Bessard, M.-A.; van Endert, P.; Edelman, A.; Guerrera, I. C. Sensitivity of Mass Spectrometry Analysis Depends on the Shape of the Filtration Unit Used for Filter Aided Sample Preparation (FASP). *Proteomics* **2016**, *16*, 1852−1857.
(11) Lever, J. E. Quantitative Assay of the Binding of Small Molecules to Protein: Comparison of Dialysis and Membrane Filter Assays. *Anal. Biochem.* **1972**, *50*, 73−83.
(12) Hajduk, J.; Matysiak, J.; Kokot, Z. J. Challenges in Biomarker Discovery with MALDI-TOF MS. *Clin. Chim. Acta* **2016**, *458*, 84−98.
(13) Manza, L. L.; Stamer, S. L.; Ham, A.-J. L.; Codreanu, S. G.; Liebler, D. C. Sample Preparation and Digestion for Proteomic Analyses Using Spin Filters. *Proteomics* **2005**, *5*, 1742−1745.
(14) Mujumdar, R. B.; Ernst, L. A.; Mujumdar, S. R.; Lewis, C. J.; Waggoner, A. S. Cyanine Dye Labeling Reagents: Sulfoindocyanine Succinimidyl Esters. *Bioconjugate Chem.* **1993**, *4*, 105−111.
(15) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. Quantitative Analysis of Complex Protein Mixtures Using Isotope-Coded Affinity Tags. *Nat. Biotechnol.* **1999**, *17*, 994−999.
(16) Thompson, A.; Schäfer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Johnstone, R.; Mohammed, A. K.; Hamon, C. Tandem Mass Tags: A Novel Quantification Strategy for Comparative Analysis of Complex Protein Mixtures by MS/MS. *Anal. Chem.* **2003**, *75*, 1895−1904.
(17) Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J. Multiplexed Protein Quantitation in Saccharomyces Cerevisiae Using Amine-Reactive Isobaric Tagging Reagents. *Mol. Cell. Proteomics* **2004**, *3*, 1154−1169.
(18) Baslé, E.; Joubert, N.; Pucheault, M. Protein Chemical Modification on Endogenous Amino Acids. *Chem. Biol.* **2010**, *17*, 213−227.
(19) Hermanson, G. T. *Bioconjugate Techniques*, 3rd edition; Elsevier Science & Technology: San Diego, United States, 2013.
(20) Habeeb, A. F. S. A.; Cassidy, H. G.; Singer, S. J. Molecular Structural Effects Produced in Proteins by Reaction with Succinic Anhydride. *Biochim. Biophys. Acta* **1958**, *29*, 587−593.
(21) Dixon, H. B.; Perham, R. N. Reversible Blocking of Amino Groups with Citraconic Anhydride. *Biochem. J.* **1968**, *109*, 312−314.
(22) Butler, P. J. G.; Hartley, B. S. [14] Maleylation of Amino Groups. In *Methods in Enzymology; Enzyme Structure, Part B*; Academic Press, 1972; Vol. *25*, pp. 191−199.
(23) Atassi, M. Z.; Habeeb, A. F. S. A. [49] Reaction of Proteins with Citraconic Anhydride. In *Methods in Enzymology*; Elsevier, 1972; Vol. *25*, pp. 546−553.
(24) Su, S.; Du, F.-S.; Li, Z.-C. Synthesis and PH-Dependent Hydrolysis Profiles of Mono- and Dialkyl Substituted Maleamic Acids. *Org. Biomol. Chem.* **2017**, *15*, 8384−8392.

(25) Ganesan, V.; Ascherman, D. P.; Minden, J. S. Immunoproteomics Technologies in the Discovery of Autoantigens in Autoimmune Diseases. *Biomol. Concepts* **2016**, *7*, 133−143.
(26) Lucas, A.; Kaupbayeva, B.; Murata, H.; Cummings, C.; Russell, A. J.; Minden, J. S. Utilization of the Polymer Sieving Effect for the Removal of the Small Molecule Biotin-CDM. *ACS Appl. Polym. Mater.* **2019**, *1*, 2897−2906.
(27) Rafiee, M. R.; Sigismondo, G.; Kalxdorf, M.; Förster, L.; Brügger, B.; Béthune, J.; Krijgsveld, J. Protease-resistant streptavidin for interaction proteomics. *Mol. Syst. Biol.* **2020**, *16*, No. e9370.
(28) St-Germain, J. R.; Samavarchi Tehrani, P.; Wong, C.; Larsen, B.; Gingras, A.-C.; Raught, B. Variability in Streptavidin−Sepharose Matrix Quality Can Significantly Affect Proximity-Dependent Biotinylation (BioID) Data. *J. Proteome Res.* **2020**, *19*, 3554−3561.
(29) Karver, M. R.; Weissleder, R.; Hilderbrand, S. A. Synthesis and Evaluation of a Series of 1,2,4,5-Tetrazines for Bioorthogonal Conjugation. *Bioconjugate Chem.* **2011**, *22*, 2263−2270.
(30) Devaraj, N. K.; Upadhyay, R.; Haun, J. B.; Hilderbrand, S. A.; Weissleder, R. Fast and Sensitive Pretargeted Labeling of Cancer Cells via Tetrazine/Trans-Cyclooctene Cycloaddition. *Angew. Chem. Int. Ed.* **2009**, *48*, 7013−7016.
(31) Meijer, A.; Otto, S.; Engberts, J. B. F. N. Effects of the Hydrophobicity of the Reactants on Diels−Alder Reactions in Water. *J. Org. Chem.* **1998**, *63*, 8989−8994.
(32) Devaraj, N. K.; Hilderbrand, S.; Upadhyay, R.; Mazitschek, R.; Weissleder, R. Bioorthogonal Turn-On Probes for Imaging Small Molecules inside Living Cells. *Angew. Chem. Int. Ed.* **2010**, *122*, 2931−2934.
(33) Rossin, R.; Renart Verkerk, P.; van den Bosch, S. M.; Vulders, R. C. M.; Verel, I.; Lub, J.; Robillard, M. S. In Vivo Chemistry for Pretargeted Tumor Imaging in Live Mice. *Angew. Chem. Int. Ed.* **2010**, *122*, 3447−3450.
(34) Keliher, E. J.; Reiner, T.; Turetsky, A.; Hilderbrand, S. A.; Weissleder, R. High-Yielding, Two-Step 18 F Labeling Strategy for 18F-PARP1 Inhibitors. *ChemMedChem* **2011**, *6*, 424−427.
(35) Li, Z.; Cai, H.; Hassink, M.; Blackman, M. L.; Brown, R. C. D.; Conti, P. S.; Fox, J. M. Tetrazine−Trans-Cyclooctene Ligation for the Rapid Construction of 18F Labeled Probes. *Chem. Commun.* **2010**, *46*, 8043−8045.
(36) Haun, J. B.; Castro, C. M.; Wang, R.; Peterson, V. M.; Marinelli, B. S.; Lee, H.; Weissleder, R. Micro-NMR for Rapid Molecular Analysis of Human Tumor Samples. *Sci. Transl. Med.* **2011**, *3*, No. 71ra16.
(37) Shen, S.; An, B.; Wang, X.; Hilchey, S. P.; Li, J.; Cao, J.; Tian, Y.; Hu, C.; Jin, L.; Ng, A.; Tu, C.; Qu, M.; Zand, M. S.; Qu, J. A Surfactant Cocktail-Aided Extraction/Precipitation/On-Pellet Digestion (SEPOD) Strategy Enables Rapid, Efficient and Reproducible Sample Preparation for Large-Scale Quantitative Proteomics. *Anal. Chem.* **2018**, *90*, 10350−10359.
(38) Van, P. T.; Bass, V.; Shiwarski, D.; Lanni, F.; Minden, J. High Dynamic Range Proteome Imaging with the Structured Illumination Gel Imager. *Electrophoresis* **2014**, *35*, 2642−2655.
(39) Rusconi, F.; Valton, É.; Nguyen, R.; Dufourc, E. Quantification of Sodium Dodecyl Sulfate in Microliter-Volume Biochemical Samples by Visible Light Spectroscopy. *Anal. Biochem.* **2001**, *295*, 31−37.
(40) Kaufmann, T. C.; Engel, A.; Rémigy, H.-W. A Novel Method for Detergent Concentration Determination. *Biophys. J.* **2006**, *90*, 310−317.
(41) Van, P. T.; Ganesan, V.; Bass, V.; Parthasarathy, A.; Schlesinger, D.; Minden, J. S. In-Gel Equilibration for Improved Protein Retention in 2DE-Based Proteomic Workflows. *Electrophoresis* **2014**, *35*, 3012−3017.
(42) Ford, J. R.; Chambers, R. P.; Cohen, W. An Active-Site Titration Method for Immobilized Trypsin. *Biochim. Biophys. Acta* **1973**, *309*, 175−180.
(43) Elias, J. E.; Gygi, S. P. Target-Decoy Search Strategy for Increased Confidence in Large-Scale Protein Identifications by Mass Spectrometry. *Nat. Methods* **2007**, *4*, 207−214.
(44) 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.; Yılmaz, 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*, D442−D450.

https://doi.org/10.1021/acs.jproteome.1c00443
J. Proteome Res. 2021, 20, 4787−4800