

✂ Author's Choice

# Multiplex N-terminome Analysis of MMP-2 and MMP-9 Substrate Degradomes by iTRAQ-TAILS Quantitative Proteomics\*<sup>§</sup>

Anna Prudova<sup>‡§</sup>, Ulrich auf dem Keller<sup>‡¶</sup>, Georgina S. Butler, and Christopher M. Overall<sup>||</sup>

Proteolysis is a major protein posttranslational modification that, by altering protein structure, affects protein function and, by truncating the protein sequence, alters peptide signatures of proteins analyzed by proteomics. To identify such modified and shortened protease-generated neo-N-termini on a proteome-wide basis, we developed a whole protein isobaric tag for relative and absolute quantitation (iTRAQ) labeling method that simultaneously labels and blocks all primary amines including protein N-termini and lysine side chains. Blocking lysines limits trypsin cleavage to arginine, which effectively elongates the proteolytically truncated peptides for improved MS/MS analysis and peptide identification. Incorporating iTRAQ whole protein labeling with terminal amine isotopic labeling of substrates (iTRAQ-TAILS) to enrich the N-terminome by negative selection of the blocked mature original N-termini and neo-N-termini has many advantages. It enables simultaneous characterization of the natural N-termini of proteins, their N-terminal modifications, and proteolysis product and cleavage site identification. Furthermore, iTRAQ-TAILS also enables multiplex N-terminomics analysis of up to eight samples and allows for quantification in MS2 mode, thus preventing an increase in spectral complexity and extending proteome coverage by signal amplification of low abundance proteins. We compared the substrate degradomes of two closely related matrix metalloproteinases, MMP-2 (gelatinase A) and MMP-9 (gelatinase B), in fibroblast secreted proteins. Among 3,152 unique N-terminal peptides identified corresponding to 1,054 proteins, we detected 201 cleavage products for MMP-2 and unexpectedly only 19 for the homologous MMP-9 under identical conditions. Novel substrates identified and biochemically validated include insulin-like growth factor binding protein-4, complement C1r component A, galectin-1, dickkopf-related protein-3, and thrombospondin-2. Hence, N-terminomics analyses using iTRAQ-TAILS links gelatinases with new mecha-

nisms of action in angiogenesis and reveals unpredicted restrictions in substrate repertoires for these two very similar proteases. *Molecular & Cellular Proteomics* 9: 894–911, 2010.

From maturation to degradation, proteolysis is a ubiquitous posttranslational modification that irreversibly modifies the structure and function of every protein in the cell (1). The N-terminal sequence of a protein can determine protein structure, function, localization, interacting partners, and turnover rates and hence is important information needed to functionally annotate the proteome. Selective proteolytic processing generates new protein N-termini, also known as neo-N-termini, sometimes with only one or a few amino acids trimmed off. Nonetheless, even such subtle changes to protein sequences often have a dramatic effect on protein function (2–4) and can serve as initiating or key steps in the proteolytic control of signaling cascades, such as cytokine activation and removal of inhibitory binding proteins (5–7).

Proteases form 5–10% of all drug targets because proteolysis is an important causal or progression factor in many diseases including chronic inflammation, neurodegeneration, heart disease, and cancer (8–11). Successful anti-proteolytic therapies include those targeting angiotensin convertase in heart disease, dipeptidyl-peptidase IV in diabetes, and human immunodeficiency virus protease-1 in AIDS (12), whereas some, such as matrix metalloproteinase (MMP)<sup>1</sup> inhibitors in cancer, have failed (13, 14). Such drug failures have been attributed to deficiencies in knowledge, specifically limited information on substrate repertoires (also known as substrate degradomes), contributing to the poor understanding of complex protease function in health and disease (7, 15). Indeed,

From the Department of Biochemistry and Molecular Biology, Department of Oral Biological and Medical Sciences, and Centre for Blood Research, University of British Columbia, 4.401 Life Sciences Institute, 2350 Health Sciences Mall, Vancouver, British Columbia V6T 1Z3, Canada

✂ Author's Choice—Final version full access.

Received, February 1, 2010, and in revised form, March 18, 2010

Published, MCP Papers in Press, March 20, 2010, DOI 10.1074/mcp.M000050-MCP201

<sup>1</sup> The abbreviations used are: MMP, matrix metalloproteinase; APMA, 4-aminophenylmercuric acetate; MCP3, monocyte chemoattractant protein 3; CCL7, CC motif chemokine 7; iTRAQ, isobaric tag for relative and absolute quantification; Dkk-3, dickkopf-related protein-3; IGFBP, insulin-like growth factor-binding protein; TAILS, terminal amine isotopic labeling of substrates; COFRADIC, combined fractional diagonal chromatography; SILAC, stable isotope labeling with amino acids in cell culture; ROC, receiver operating characteristic; Tricine, *N*-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine; IGF, insulin-like growth factor.

for half of the 569 human proteases, no substrate is known, whereas processing of known targets of the other half is not well characterized (16). Hence, complete annotation of substrates and their cleavage sites is warranted, and this is best done in an unbiased manner on a global scale.

In addition to constitutive proteolysis during protein synthesis and maturation, the processing of a mature protein often irreversibly changes its activity. Hence, within each substrate, it is important to determine the cleavage site because the biological activity of the cleavage products is commonly determined by the precise fragmentation pattern. Although the changes in protein structure induced by proteolytic processing often translate into effects on biological function, such alterations in substrate sequence are inherently difficult to detect. Thus, it has not been possible to assess all the proteolytic cleavages in a biological sample, especially for substrates of low abundance. Toward the goal of complete degradome annotation of complex biological samples, several proteomics methods have been developed. The first was ICAT or isobaric tag for relative and absolute quantitation (iTRAQ) labeling followed by shotgun proteomics where protease substrates can be identified after calculating peptide and hence protein relative ratios in cellular samples derived from protease-transfected or control cells (17–19). Although proving very successful in identifying many novel protease substrates in the cellular context, this represents an indirect method requiring biochemical determination of the cleavage site. In contrast, detection and analysis of the cleavage product prime side neo-N-termini allows for direct monitoring of proteolysis and is achieved by methods that enrich for protein N-termini.

The most comprehensive of the N-terminomics approaches, COFRADIC, uses several chemical derivatization steps and relies on multiple rounds of chromatographic separation followed by MS/MS analyses of ~150 fractions per sample (20). The original COFRADIC method was modified to incorporate isotopic labeling with  $H_2^{16}O/H_2^{18}O$  and was successfully used for the detection of proteolytic events in several *in vitro* systems (21, 22). Although labeling with light/heavy water is a simple way to introduce isotopes during trypsin digestion, this modification is labile and thus is a source of error in quantification. More recently reported, SILAC-COFRADIC (23) eliminates this problem but requires expensive reagents, making it impractical for the analysis of tissues from animal experiments and inapplicable for analysis of clinically relevant patient samples because proteins need to be SILAC-labeled during translation. Nonetheless, it is the most expansive and successful technique to date. Alternative methods use positive selection approaches where protein N-termini are enriched through avidin affinity chromatography following their chemical (24) or enzymatic modification using engineered subtiligase (25) with biotinylated tags. A serious drawback of these methods is the lack of reliable quantification that is key for distinguishing specific proteolytic events from basal proteolysis present in every sample. Hence, it can

be difficult to definitively ascribe specific cleavage products to a specific protease. This is an especially critical issue when analyzing proteases with no known consensus cleavage site or having broad specificity that prevents manual parsing of the data for cleavage site motifs. Subtiligase also shows some amino acid bias in labeling (26) that can preclude analyses of certain proteases or skew substrate analyses of others. Moreover, all current positive selection approaches suffer in neopeptide identification of the prime side cleavage products that are present as shortened semitryptic peptides after trypsinization in proteomics workflows. In contrast, gel-based approaches such as isotopic labeling and one-dimensional gel analysis (27), protein topography and migration analysis platform (PROTOMAP) (28), and/or two-dimensional DIGE (29, 30) rely on detection of cleaved proteins by the shift in their migration on SDS-PAGE. Although easily adapted for many proteases, SDS-PAGE limits sensitivity, the resolution of cleaved *versus* intact proteins is sometimes difficult, and gel-based MS approaches do not allow for automatic/simultaneous detection of substrate cleavage sites. Indeed, the absence of N-terminal information dramatically reduces substrate identification (31).

We recently introduced a new proteomics technique for protein N-terminal characterization and substrate discovery we call terminal amine isotopic labeling of substrates (TAILS) (32). TAILS utilizes negative selection to enrich original mature protein N-termini and proteolysis-derived neo-N-termini from a proteome labeled with stable isotope tags for quantification. Two samples containing active or inactive protease are reduced, alkylated, and labeled by dimethylation with primary amine-reactive formaldehyde (heavy or light) to block protein N-termini and side chains of lysine residues. The samples are mixed at a 1:1 ratio and digested with trypsin. Primary amino groups of trypsin-generated internal peptides are covalently coupled to a high molecular weight polyaldehyde polyglycerol polymer developed for proteomics and are removed by filtration. The filtrate contains the labeled original mature protein N-termini and proteolysis-derived neo-N-termini that are identified and quantified by mass spectrometry in MS1 mode where the areas under the ion peaks of the corresponding peptide variants doublets are compared (32). Quantitative and statistical analysis of peptide isotopic ratios set proteolysis-derived neo-N-termini apart from original N-termini of the proteins.

Although dimethylation serves as an inexpensive option providing wide N-terminome coverage (32), it is commonly used as two isotopic variants and therefore limits experiments to pairwise sample comparisons (33). In addition, labeling with heavy and light formaldehyde, which differ in mass by 6 Da (33), like SILAC labeling (34), doubles sample complexity, thus reducing the chances for identification of rare proteins. To address these issues, we developed whole protein iTRAQ labeling to incorporate multiplex analysis and MS2 quantification into the TAILS procedure.

Here we describe and report validation of iTRAQ-TAILS using a mixture of chemokines treated with the endoprotease Glu-C as a model protease of known specificity. Despite their importance in maintaining homeostasis in health and disease, precise mechanisms of the actions of closely related gelatinases MMP-2 and MMP-9 are not well understood. Thus, we used iTRAQ-TAILS to simultaneously characterize and compare their substrate degradomes from the N-terminome. Novel substrates were identified, and many were biochemically validated including insulin-like growth factor binding protein-4, galectin-1, dickkopf-related protein-3, and thrombospondin-2.

#### EXPERIMENTAL PROCEDURES

**Cell Culture and Secretome Preparation—Mmp2<sup>-/-</sup>** mouse embryonic fibroblasts used in this study were described previously (19). To collect secreted proteins, ~70% confluent cells in T175 flasks (Corning) were washed four times with 20 ml of PBS each and incubated in Dulbecco's modified essential medium lacking phenol red (Invitrogen) and serum at 37 °C in a 5% CO<sub>2</sub> atmosphere. After 48 h, the supernatant was collected, the cells were discarded, and the medium was clarified for 5 min at 400 × g to remove any dead cells. Next, PMSF (0.5 mM final) and EDTA (1 mM final) were added followed by centrifugation at 4 °C for 30 min at 2,250 × g to remove smaller debris. Supernatants were sterile filtered and frozen at -80 °C. Secreted proteins were concentrated at 4 °C by ultrafiltration using 5-kDa-cutoff membranes (Amicon) followed by several volume washes with 100 mM HEPES buffer (pH 8). Protein concentration was measured by Bradford assay (Bio-Rad), and the sample was adjusted with 1.0 M HEPES (pH 8) to 2 mg/ml protein in 250 mM HEPES. The secretome preparations were aliquoted in 0.5-mg amounts and stored at -80 °C.

**In Vitro Protease Digestion of Fibroblast Secretomes—Glu-C** was from New England Biolabs, and human MMP-2 and MMP-9 were expressed and purified (35) in their proform to single band homogeneity (see supplemental Fig. 2). MMPs were activated with 1 mM APMA for 1 h at 37 °C in the dark. APMA was removed by two buffer exchanges using desalting spin columns (Pierce) and 100 mM HEPES (pH 8). Any carryover of APMA was <1 μM. Activation of MMP-2 and MMP-9 as well as the activity state of endogenous pro-MMP-9 in the secretomes was monitored by gelatin zymography using 9% polyacrylamide gels co-polymerized with 1 mg/ml gelatin and by a quenched fluorescent substrate assay using a synthetic MMP peptidic substrate (7-methoxycoumarin-4-yl)acetyl-Pro-Leu-Gly-Leu-[3-(2,4-dinitrophenyl)-L-2,3-diaminopropionyl]-Ala-Arg-NH<sub>2</sub> (35). For MMP secretome digests (1:100 protease/secretome), 10 mM CaCl<sub>2</sub> and 100 mM NaCl final were added, and the samples were then incubated for 18 h at 37 °C. As a positive control, CCL7/MCP-3, a known MMP-2 substrate (3), was added (0.5–1 μg/100 μg of secretome). Control samples received an equivalent volume of buffer only. For each individual experiment, identical secretome aliquots from the same batch were used for all conditions and digestions. Repeat experiments performed with secretomes from the same batch were defined as technical replicates. Experiments with secretomes collected on different dates from different passage cells were considered biological replicates.

**Preparation of Digested Secretomes—**A typical iTRAQ-TAILS analysis utilized 0.5 mg of secretome/condition for each iTRAQ channel. Prior to labeling, protease-treated secretomes were first denatured, and the cysteine residues were reduced and alkylated as follows: 8.0 M guanidinium chloride and 1.0 M HEPES (pH 8) were used to achieve final concentrations of 2.5 M and 250 mM guanidinium chloride and

HEPES, respectively. After 15-min incubation at 65 °C, the samples were reduced by tris(2-carboxyethyl)phosphine (1 mM final) for 45 min. Next, alkylation was performed using 5 mM iodoacetamide for 30 min at 65 °C in the dark.

**Whole Protein iTRAQ Labeling—**We developed and highly optimized a novel whole protein iTRAQ labeling method. Labeling was performed using 1:5 protein/iTRAQ (w/w) ratio and 50% DMSO final as a solvent. Thus, for 0.5-mg secretome reactions, 2.5 mg of each individual iTRAQ isotopic tag were dissolved in 100% DMSO using an amount equal to the volume of each alkylated sample. After mixing with the sample and 30-min incubation at room temperature, the iTRAQ reagent was quenched with 100 mM ammonium bicarbonate for 15 min (to prevent label carryover to tryptic digestions). After labeling, the reactions were combined at a 1:1 ratio. The samples were cleaned up using ice-cold acetone/methanol (8:1 volume ratio) to precipitate 1 volume of sample at -80 °C for 2–3 h. The pellet was sedimented by centrifugation at 4 °C for 30 min at 10,000 × g. After discarding the supernatant, the sample was resuspended in ice-cold methanol to remove any precipitated guanidinium chloride and centrifuged again (two washes in total). The final pellet was briefly air-dried, first resuspended in 50 μl of 100 mM NaOH, and then adjusted to 1 mg/ml protein, 100 mM HEPES (pH 8). We use sequencing grade TrypsinGold (Promega) at a 1:100–1:200 ratio (enzyme/proteome weight ratio) for 18 h at 37 °C with trypsinization assessed by 10% SDS-PAGE and silver staining.

**N-terminal Enrichment—**Trypsinized samples were enriched for N-terminal peptides using a dendritic polyglycerol aldehyde polymer to deplete the internal and C-terminal tryptic peptides (32). The pH of the reactions was first adjusted with HCl to 6–7 followed by addition of sodium cyanoborohydride (30 mM final) and the polymer. To ensure complete capture of tryptic peptides bearing a free α-amine, the polymer was used in 3–5-fold excess (w/w) for 18 h at 37 °C. After coupling, the unbound peptides were separated from the polymer-bound peptides by filtration using ultrafiltration tubes with 3-kDa-cutoff membranes (Amicon). When concentrated to ~100–200 μl, the polymer was washed with 200 μl of 100 mM ammonium bicarbonate. Polymer-bound internal peptides on the filter were discarded, and the flow-through containing N-terminal peptides was frozen at -80 °C until further analysis.

**Off-line High Performance Liquid Chromatography—**N-terminal peptides were fractionated by strong cation exchange HPLC on an Agilent Technologies 1200 series HPLC system (Agilent Technologies) using a PolySULFOETHYL A 100 × 4.6-mm, 5-μm, 300-Å column (PolyLC Inc.). Solvent A consisted of 10 mM potassium phosphate and 25% acetonitrile (pH 2.7). Solvent B included 10 mM potassium phosphate, 25% acetonitrile, and 1.0 M sodium chloride (pH 2.7). Separation was monitored by absorbance at 214 and 280 nm. Peptides were bound to the column and washed for 15 min with 100% solvent A. Peptides were then eluted at 1 ml/min where solvent B was gradually increased from 0 to 30% from 15 to 37 min followed by a sharp increase to 40% at 43 min. Solvent B increased to 100% at 45 min and was kept at 100% for 8 min. The mobile phase was switched back to solvent A from 53 to 55 min, and the column was equilibrated with 100% solvent A for 10 min. Peptide-containing fractions were collected every 1.5 min, concentrated to 0.1 ml under vacuum, and desalted using C<sub>18</sub> OMIX tips.

**In-line Liquid Chromatography and Mass Spectrometry Analysis—**Peptide samples (2–10 μl) were analyzed by nanospray LC-MS/MS using a C<sub>18</sub> column (150-mm × 100-μm column at a flow rate of 100–200 nl/min) interfaced with a QStar XL Hybrid ESI mass spectrometer (Applied Biosystems, MDS-Sciex, Concord, Canada). After loading the samples onto a trapping column, the column was washed with 5% acetonitrile and 0.1% formic acid (v/v). Elution and separation of peptides were achieved with a 40–100-min linear 5–40%



acetonitrile gradient (containing 0.1% formic acid) at a flow rate of 150–200 nl/min. MS data were acquired automatically with Analyst QS version 1.1 software (Applied Biosystems, MDS-Sciex). An information-dependent acquisition method consisting of a 1-s TOF MS survey scan of mass range 400–1500 amu and three 3-s product ion scans of mass range 75–1500 amu was used. The three most intense peaks over 20 counts with a charge state of 2+ to 4+ were selected for fragmentation.

**MS2 Peptide Assignments and iTRAQ Quantification**—Acquired MS2 scans were searched against a mouse International Protein Index (IPI) protein database (v.3.24) supplemented with the sequences for human CCL7 and MMP-2 (total 52,415 protein entries) by Mascot version 2.2 (Matrix Science) and X! Tandem 2007.07.01 release. Searches were performed with the following parameters: semi-Arg-C cleavage specificity with up to two missed cleavages; cysteine carbamidomethylation and peptide lysine iTRAQ as fixed modifications; and N-terminal iTRAQ, N-terminal acetylation, N-terminal amino acid cyclization (chemokine dataset only), and methionine oxidation as variable modifications. Peptide tolerance and MS/MS tolerance were set at 0.4 Da, and the scoring scheme was ESI-QUAD-TOF. Search results were evaluated on the Trans Proteomic Pipeline (TPPv.4.2, rev 0, Build 200811181145) (36, 37) using PeptideProphet (38) and iProphet (39) for peptide/protein identification and Libra for quantification of iTRAQ reporter ion intensities. Final data sets included only the peptides with an iProphet probability error rate  $\leq 5\%$  as was statistically modeled by the corresponding Trans Proteomic Pipeline software for each individual data set.

**Statistical Data Analysis and Peptide Annotation**—The development and validation of our novel statistical analysis for N-terminomics is fully detailed in the accompanying paper (40). In brief, for the generation of MA plots, histograms, probability densities, distributions, and non-linear curve fitting, the R statistical environment was used (version 2.8.0). Receiver operating characteristic (ROC) curve analysis was performed using the ROC package (41). To calculate standard deviations by a sliding window approach, for spectrum merging and weighted averaging, and for peptide isoform assignments and positional annotation, we used in-house Perl scripts using appropriate BioPerl packages (40). Peptide positions within protein sequences represent amino acid numbering in the mature fully processed form of each protein. For identifier mapping, the Protein Information and Property Explorer system (42) was used. For active site mapping, amino acid occurrences were calculated as described previously (43), and heat maps were generated using TM4:MeV. Protein sequence logos were generated using the ice-Logo software package with random sampling of the reference database (44).

**Substrate Cleavage Assays**—Human C1r component A protein and rabbit pyruvate kinase M1/M2 were from Sigma. Human thrombospondin-2 and antibodies were a kind gift from Dr. Paul Bornstein (University of Washington, Seattle, WA). Human dickkopf-related protein 3 (Dkk-3) was from R&D Systems. Human galectin-1 was from Research Diagnostics Inc. Human peptidyl-prolyl cis-trans isomerase A and antibodies were from BIOMOL International. Human insulin-like growth factor-binding protein (IGFBP)-4 and cystatin C were from Cell Sciences Inc.

MMP-2 and MMP-9 were incubated with the candidate substrates in 50 mM Tris-HCl, 200 mM NaCl, 5 mM CaCl<sub>2</sub>, 1 mM APMA, and 0.05% Brij 35 for 16 h at 37 °C. To minimize proteolysis of the MMPs by the serine protease C1r, 0.5 mM PMSF was included in the assay that was for 4 h only. Reaction products were analyzed by Tris-Tricine SDS-PAGE or Tris-glycine SDS-PAGE and silver staining, Western blotting, and Edman sequencing as appropriate.

## RESULTS

**iTRAQ Whole Protein Labeling**—Here we modified our recently published TAILS protocol (32) with the incorporation of iTRAQ labeling for the multiplexing of four to eight samples per analysis that allows for more accurate quantification and versatile experimental designs (Fig. 1). Unlike conventional iTRAQ where labeling is performed after trypsin digestion on peptides (45), we aimed to quantify cleavage events in native proteins and so label and quantify these events at the protein level rather than at the peptide level. To do so, we devised a novel protocol for efficient, rapid, and amino group-specific protein iTRAQ labeling performed before trypsinization. This represents a key step of the new TAILS workflow as it (i) blocks protein N-termini, allowing for their subsequent enrichment by negative selection along with all naturally blocked N-termini, such as those with acetylated or cyclized N-termini (32); (ii) simultaneously introduces a stable isotope tag to  $\alpha$ -amino groups of protein N-termini and  $\epsilon$ -amino groups of lysine residues, thus allowing for accurate quantification; and (iii) because trypsin cannot cleave at iTRAQ-blocked lysine residues, trypsinization yields peptides with Arg-C-like specificity, lengthening most protease-truncated semitryptic neo-N-peptides for improved MS/MS identification and substrate coverage. Our highly optimized conditions, using 50% DMSO as a solvent, yielded  $97 \pm 3\%$  labeling efficiency within 30 min for protein N-terminal  $\alpha$ -amines and lysine  $\epsilon$ -amines with minimal side reactions on Tyr, Ser, or Thr (when the data sets were searched using iTRAQ-Tyr, -Ser, or -Thr as variable modifications). Hence, this whole protein iTRAQ labeling method has applicability for other applications aimed at protein characterization.

**iTRAQ-TAILS Validation by Glu-C Cleavage of Native Proteins**—For iTRAQ-TAILS validation, we used a mixture of 13 chemokines of which one was a synthetic analog of MMP-cleaved CXC motif chemokine 8 (from position 5 to 72 (5–72)) of the full-length (1–72) chemokine (supplemental Table 1). To check the accuracy of neo-N-terminal peptide identification, the chemokine mixture was incubated with the protease Glu-C (which cleaves only after Glu or Asp, so allowing for manual data parsing) (43) or buffer alone (control) under conditions to retain the native protein structure. The Glu-C-incubated and control samples were labeled with iTRAQ-116 and iTRAQ-115, respectively. The principal of the quantitative aspects of TAILS (32) is that original mature protein N-terminal peptides of proteins that are unmodified by proteolysis should be present in both samples in equal amounts and therefore have a protease/control iTRAQ ratio of 1.0 (Fig. 1). Neo-N-termini generated by proteolysis will have an iTRAQ ratio  $\gg 1$  (singletons; indicated here by a ratio of 35) because they will not be present in the undigested sample. Blocked mature protein N-terminal peptides cleaved within their sequence will decrease to give iTRAQ ratios  $< 1$ ; in this case, the ratio is proportional to the extent of proteolysis.

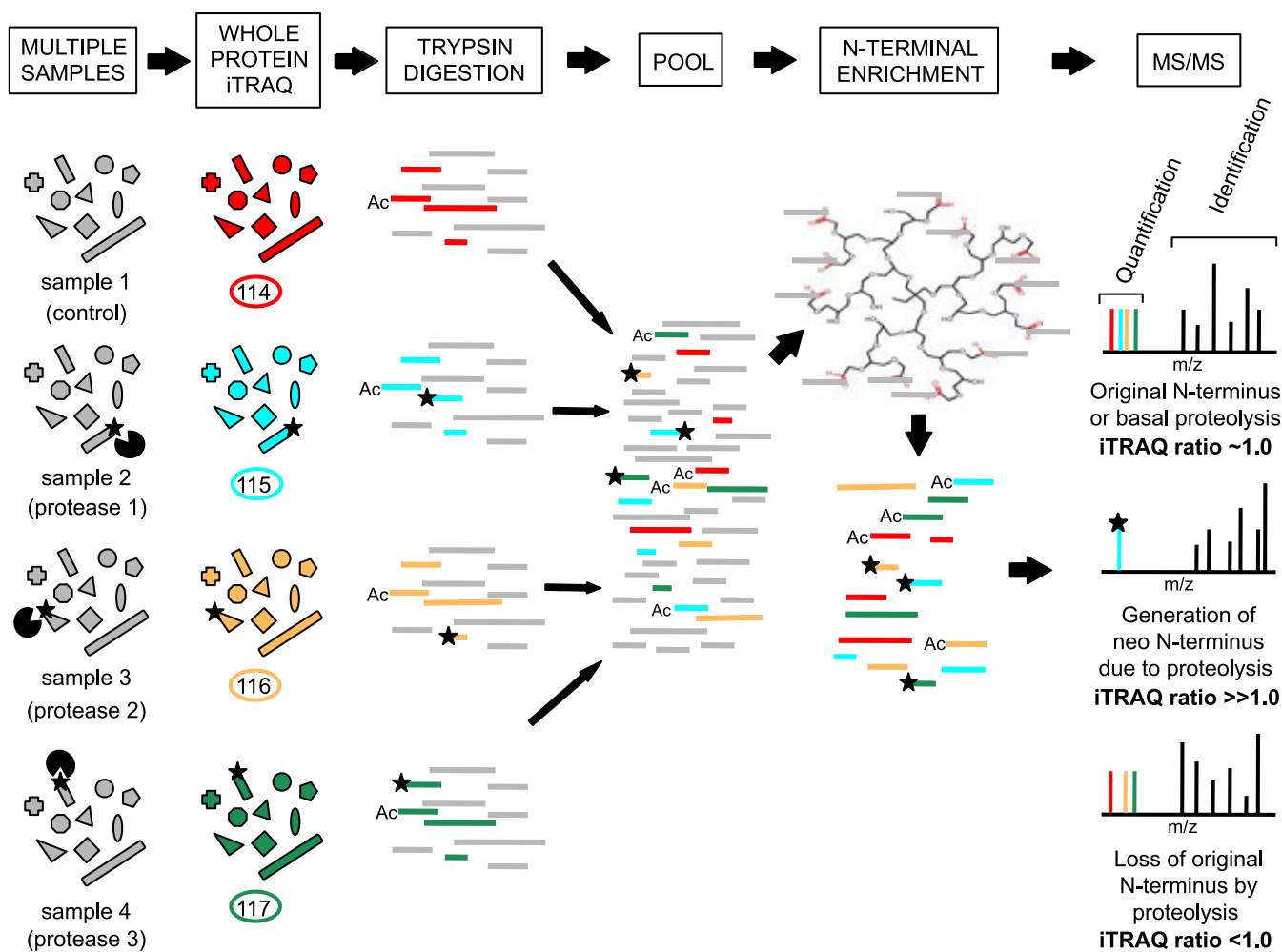
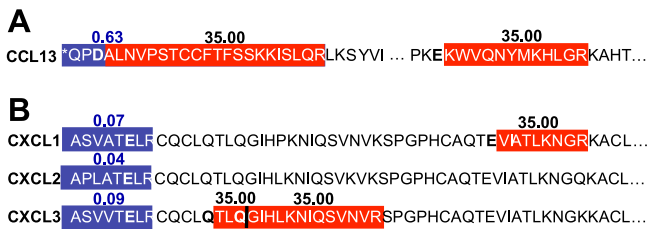


FIG. 1. **iTRAQ-TAILS workflow.** Multiple samples (two up to eight) can be co-analyzed and compared using iTRAQ-TAILS. In the schematic, a proteolyzed protein is indicated by a *black star* in *samples 2, 3, and 4* generated by different proteases (or under different conditions). The proteins in each complex sample are labeled with distinct iTRAQ labels (that give 114, 115, 116, or 117 reporter ions in MS2 mode) using whole protein iTRAQ labeling on their free N-termini and lysine residues (represented by *colored* peptides). This not only distinguishes proteins derived from each sample (indicated by *different colors*) but can also be used for conventional iTRAQ quantification of the proteins present after trypsinization. Pooled samples are subjected to trypsin digestion to generate the following peptides: N-terminal peptides that are modified by iTRAQ reagents or pre-existing modifications such as acetylation (Ac) (indicated by *color*) or internal and C-terminal peptides having free N-termini (indicated in *gray*). The peptides are pooled, and the N-terminal peptides from the original proteins are isolated by negative selection: a polyaldehyde dendritic polymer binds all of the peptides with free (trypsin-generated) N-termini but not the peptides with blocked N-termini (iTRAQ-labeled or naturally acetylated or cyclized). The polymer and bound peptides are removed by ultrafiltration, and the N-terminal peptides are subjected to MS/MS. In MS2, following peptide fragmentation, the iTRAQ labels are quantified, giving the relative amount of peptide derived from each sample, whereas the rest of the b and y ions permit peptide identification using the search engines Mascot and X! Tandem. Original N-termini present in all samples give an iTRAQ ratio centered on 1.0. The appearance of a singleton reporter ion represents the neo-N-terminus (indicated by the *black star*), which has an iTRAQ ratio significantly >1. Alternatively, indirect evidence for proteolysis is revealed by a disappearance of the original N-terminus (which will have an iTRAQ ratio <1).

Supplemental Table 2 lists the 342 spectra used to identify 23 unique peptides (supplemental Table 3) representing 12 of the 13 chemokines encompassing seven original protein N-termini and 14 Glu-C neo-N-termini. Several predicted peptides are too short for MS/MS identification or have redundant sequences (see below). Notably, searches were performed without specifying an enzyme, and because all identified peptides were either Arg-C/Arg-C original protein N-termini (the original protein N-terminus with Arg at the C-terminus), Glu-C/

Arg-C neopeptides (due to Glu-C cleavage after Glu and Asp or deamidated Gln and Asn), or Glu-C/Glu-C neopeptides ( $n = 5$ ), this unbiased search provides additional validation for iTRAQ-TAILS and our bioinformatics pipeline (40).

As expected, all Glu-C/control iTRAQ ratios for N-termini with internal Glu or Asp residues were <1.0, and all neo-N-peptides were present as high iTRAQ ratio ( $\geq 35$ ) singletons (see note for peptide 8 in supplemental Table 3). For example, for CC motif chemokine 13, both the original N-terminus



**FIG. 2. iTRAQ-TAILS validation using simple chemokine mixture: examples of identified peptides and their iTRAQ ratios.** A mixture of 13 chemokines (see supplemental Table 1 for the complete list) was divided in two and incubated with Glu-C or buffer alone. After TAILS and analysis by tandem mass spectrometry, the data were analyzed using Mascot and X! Tandem followed by validation with iProphet. The sequences of the chemokines CCL13 (A) and the first 52 residues of CXCL1, -2, and -3 (B) are shown with the peptides identified by TAILS and their corresponding Glu-C/control iTRAQ ratios indicated above. Peptides shown in red correspond to cleavage fragments generated by Glu-C processing after Asp, Glu, and deamidated Gln and Asn (indicated in bold). Peptides in blue represent original protein N-termini. \* indicates pyro-Gln as identified by TAILS. The sequence <sup>40</sup>VIATLKNGR<sup>48</sup> in CXCL1 was identified as <sup>40</sup>VIATLK-DGR<sup>48</sup>, indicative of Asn<sup>46</sup> deamidation.

pyro<sup>1</sup>QPDA2LNVPSTCC3CFSSK4KISLQR<sup>24</sup> (supplemental Table 3), with an iTRAQ ratio of 0.63 (Fig. 2A), and the neo-N-terminus (underlined) generated by Glu-C cleavage after Asp<sup>3</sup> with a Glu-C/control iTRAQ ratio  $\geq 35$  (supplemental Table 3) were identified. Such results highlight three key features of TAILS. First, cyclization of N-terminal Gln to pyro-Gln prevents N-terminal iTRAQ labeling. However, because TAILS quantification does not rely just on labeling peptide N-termini, the two lysine residues in this peptide were both labeled, and so the peptide was quantifiable (Fig. 2A and supplemental Tables 2 and 3). Second, proteolytic cleavage can result in semitryptic peptides too short for MS detection or reliable identification due to sequence redundancy. Blocking of lysine residues mitigates this problem in TAILS. However, cleavage at a nearby arginine residue can still be problematic. For example, Glu-C cleavage of CXCL1, -2, and -3 chemokines after Glu<sup>6</sup> generates the redundant two-residue fragment <sup>7</sup>LR (Fig. 2B). However, the third advantage of TAILS is the ability to also indirectly identify substrates by quantifying a decrease in original N-termini when proteolytically processed as was the case for these three chemokines (Fig. 2B). For CC motif chemokine 13, we identified a decrease in the intact N-terminal peptides that corresponded with the generation of its internal neopeptide only in the Glu-C sample. Such reciprocal examples provide very strong confirmation of cleavage. Thus, iTRAQ-TAILS allows both qualitative (*i.e.* identification of true N-termini, both original and neo) and quantitative (*i.e.* determination of iTRAQ ratios) analysis of proteolysis products, both directly and indirectly, providing a quantitative approach with extended substrate coverage.

iTRAQ quantification of peptides is carried out at the MS2 level following MS/MS fragmentation. Because some peptides are sampled multiple times during their elution, some

reporter ion intensities are quantified several times under varying conditions with more intense ions providing higher accuracy ratios than lower intensity ions. To assure accurate quantification, we calculated intensity-dependent weighting factors for spectrum merging for such peptides as described in full in the accompanying paper (40). In this experiment, we divided a secretome sample (from *Mmp2*<sup>-/-</sup> murine embryonic fibroblasts) and subjected each half to iTRAQ-TAILS (iTRAQ-115 versus iTRAQ-116 labels) without any protease treatment, expecting iTRAQ 115/116 ratios of 1.0 for all peptides. iTRAQ-TAILS analysis yielded 1,593 spectra corresponding to 987 quantifiable peptides identified with high confidence (supplemental Table 4). These spectra were used to perform an intensity-dependent ratio MA analysis where *M* values ( $\log_2(115/116)$ ) are plotted against corresponding *A* values ( $0.5 \times \log_2(115 \times 116)$ ) (40). As expected, experimental deviation from the predicted ratio of 1.0 was lowest for high *A* values (peptides with strong ion intensities and therefore providing a more reliable ratio) and highest for low intensity reporter ion quantifications (supplemental Fig. 1A). Next, we calculated *A* value-dependent standard deviations followed by non-linear curve fitting to derive relative intensity-dependent weighting factors for reporter ion intensity ratios. These were applied in quantitative analysis of all data sets, thus assuring highly accurate peptide quantification (40).

For proteases with broad or unknown specificity, N-terminome data cannot be manually parsed to identify cleavage sites, and so quantitative techniques are necessary for these proteases. iTRAQ-TAILS quantification relies on reporter ion intensity ratios to distinguish between original N-termini with expected ratios of 1.0 and proteolysis-derived neo-N-termini with ratios  $\gg 1.0$ . However, setting arbitrary ratio cutoffs (*e.g.* 2-fold (17)) can cause either undersampling of cleavage events (if set too high) or result in high rates of false positives (if set too low). Hence, we established a statistically defined reporter ion ratio cutoff using Glu-C because its specificity is known, and therefore the cutoff ratio can be validated. To do this, equal aliquots of MDA-MB-231 human breast cancer cell secretome were digested with Glu-C under native conditions or incubated with buffer alone (control) followed by TAILS using iTRAQ-117 and iTRAQ-114 tags, respectively. The resulting 1,793 spectra corresponding to 1,272 high confidence peptides are listed in supplemental Table 5. From these, a subset of 157 spectra was selected that (i) are unambiguously assigned to semitryptic peptides cleaved C-terminal to Glu or Asp in position P1 and (ii) have no internal Glu or Asp residue (as cleavages here will alter their ratio). Reflecting the large number of high ratio cleaved peptides, these spectra had the maximum of the probability density function of iTRAQ ratios of 2<sup>5.11</sup> corresponding to a dynamic range of  $\sim 35$  (supplemental Fig. 1B). Using this value as the upper quantification threshold, we performed ROC curve analysis (supplemental Fig. 1C) to determine the optimal singleton ratio cutoff for cleavage event spectra. Thereby, a 117(Glu-

C)/114(control) reporter ion intensity ratio of 10 has a true positive rate of 93% and a false positive rate of 15% (40). Notably, ROC analysis performed for the same data set at the peptide level rather than spectral level (using intensity-weighted means of reporter ion ratios after spectrum merging) produced similar results (supplemental Fig. 1C, gray dashed line), thus validating our approach for multiple spectrum peptide quantification (40).

Finally, we tested our substrate ratio cutoff model with an independent data set where again Glu-C was used but on a different collection of secretome from MDA-MB-231 human breast cancer cells. TAILS identified 1,695 peptides matching our high confidence peptide identification criteria from 2,288 spectra (supplemental Table 6). 43% (639 of 1,499) of all quantifiable peptides (supplemental Table 7) had an unambiguously assigned Glu or Asp in the P1 position, and 84% (536 of 639) of these had a ratio above the substrate cutoff of 10, which is in good agreement with our statistical model. Moreover, this increased to 86% (147 of 172) when peptides with internal Glu and/or Asp residues that are affected by internal Glu-C cleavage were excluded, thus validating our approach. Hence, at a ratio of 10, this provides a conservative number of true substrates, but at the same time, the cleavage sites are identified with very high confidence.

*Analysis of MMP-2 and MMP-9 Substrate Repertoires by 4-plex iTRAQ-TAILS*—Recent proteomics analyses from our laboratory and others reveal a diverse MMP-2 substrate degradome that includes a wide range of bioactive molecules (19, 46), but less is known about the substrates for the closely related and pathologically important MMP-9 (7, 47–49). Moreover, few studies directly compare these two proteases, and confounding effects of different experimental models complicate interpretation of their substrates and biological roles. Therefore, we directly compared the substrate repertoires of MMP-2 and MMP-9 on the same secretome sample assayed and analyzed under identical conditions using 4-plex iTRAQ-TAILS.

Secretomes from *Mmp2*<sup>-/-</sup> murine embryonic fibroblasts (which lack MMP-2 and express only the inactive proform of MMP-9 (supplemental Fig. 2)) were incubated with active human MMP-2 or MMP-9, buffer alone (undigested control), or 100  $\mu$ M APMA (an activator of MMPs (50)) as a control for any endogenous MMPs present in the secretome. As a positive control, human chemokine CCL7/MCP3, a known MMP substrate with a characterized cleavage site (3), was spiked into each sample. Samples from two biological replicates were processed by TAILS using 114, 115, 116, and 117 iTRAQ reagents for labeling the control, MMP-2-, MMP-9-, and APMA-treated samples, respectively. Aliquots from each replicate, before and after TAILS N-terminal enrichment, were analyzed by MS/MS, searched with Mascot and X! Tandem, and validated with PeptideProphet, and the results from both search engines were then analyzed by iProphet using identical search and secondary validation parameters. Finally, both

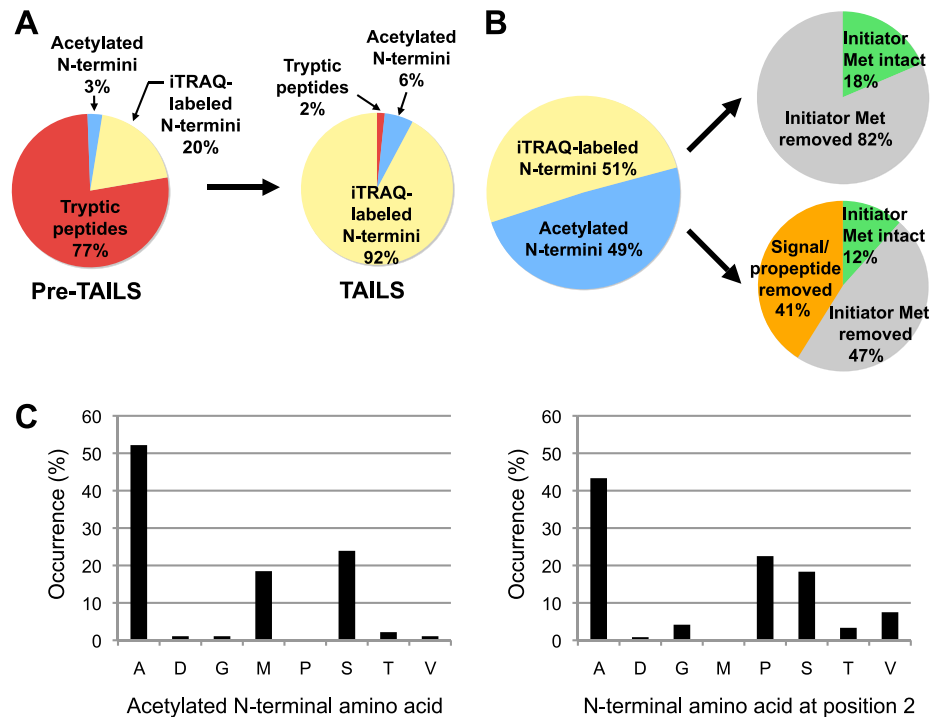
iProphet files were combined in a single list for subsequent hierarchical substrate winnowing as described (32). The pre-TAILS analysis yielded 1,197 unique peptides (1,915 spectra) (supplemental Table 8). After TAILS, 3,152 unique peptides were identified from 5,972 spectra (supplemental Tables 9 and 10). As expected, before N-terminal enrichment, the sample is dominated by internal tryptic peptides (77%) with a smaller fraction of iTRAQ-labeled (20%) and naturally acetylated (3%) N-terminal peptides (Fig. 3A). In contrast, the TAILS sample was highly enriched for protein N-termini (98% total) with only a 2% bleed-through of tryptic peptides. 92% of the peptides were iTRAQ-labeled mature protein and neo-N-termini, and 6% were naturally acetylated N-terminal peptides.

Among the 1,054 proteins identified by TAILS, 259 were represented by their original mature N-termini (supplemental Table 11), *i.e.* protein N-termini as translated or generated during maturation with and without removal of the initiator methionine as well as after signal or propeptide removal. These were approximately equally split between acetylated and non-blocked N-termini, now labeled by iTRAQ (Fig. 3B), as found before (32). Thereby, the preferred acetylated residues in position 2 after removal of Met<sup>1</sup> were alanine, serine, and methionine, and alanine, proline, serine, and valine were the most abundant amino acids in position 2 of non-blocked proteins with initiator methionine removed (Fig. 3C). Again, this was in very good agreement with our previous findings (32). As expected, iTRAQ ratios of original N-termini are closely distributed around 1.0 and thus were used to derive factors for intra- and interexperimental quantification normalization (supplemental Fig. 3). A smaller subpopulation of peptides with ratios <1.0 are likely to be the peptides that harbor an MMP cleavage site where processing results in reduced amounts of the original N-termini in the control sample, similar to what was seen with the CXC chemokines described above.

In MMP-2-treated samples, we identified 201 unique peptides from 162 proteins (supplemental Table 12) with an MMP-2/control iTRAQ ratio  $\geq 10$ , and so, according to our statistical model, these are high confidence substrates. In contrast, MMP-9 digestion under identical conditions yielded only 19 unique peptides from 19 proteins (supplemental Table 13). Notably, these were also identified as MMP-2 substrates; that is, MMP-9 cleaved 19 of the same substrates of MMP-2 in this secretome sample. As expected, the positive control human MCP3/CCL7 yielded iTRAQ ratios of 29 and 14 for processing by MMP-2 and MMP-9, respectively. MMP-2 and MMP-9 process some proteins annotated as intracellular, which is consistent with the previous findings of “mislocalized” intracellular proteins among secreted proteins (7, 47, 51). Although this may represent an important clearance mechanism for intracellular proteins released upon cell death and lysis, some intracellular proteins are now recognized as having *bona fide* extracellular roles independent of their intracellular functions,



**FIG. 3. N-terminome analysis of *Mmp2*<sup>-/-</sup> fibroblast secretomes by iTRAQ-TAILS.** Proteins secreted by *Mmp2*<sup>-/-</sup> murine fibroblasts were incubated with APMA-activated human MMP-2, MMP-9, 100  $\mu$ M APMA, or buffer alone. The samples were processed by TAILS and analyzed by MS/MS. **A**, distribution of nonredundant peptides in the aliquot of the sample analyzed by mass spectrometry before (pre-TAILS;  $n = 1,197$  peptides) and after N-terminal enrichment (TAILS;  $n = 3,152$  peptides). In the pre-TAILS and TAILS analyses, 95 and 97% of all identified naturally acetylated peptides, respectively, were found to contain  $\geq 1$  lysine residue(s), which was iTRAQ-labeled and therefore quantifiable. **B**, distribution of N-terminal modifications among original N-termini identified by TAILS analysis ( $n = 190$  proteins). **C**, frequency distributions of acetylated amino acids at the protein N-termini and of residues found at non-blocked protein N-termini after initiator methionine removal.



and so their cleavage has the potential to modify their role outside of the cell (7, 51).

**Biochemical Validation of Novel MMP-2 and MMP-9 Substrates**—Based on our Glu-C studies, we have determined stringent conditions for substrate identification: an iTRAQ ratio cutoff of 10 reflects an optimal 93% true positive rate of cleavage event identification with a 15% false positive rate. As determined from ROC analysis, peptides with a ratio of 30 have only a 5% false positive rate (supplemental Fig. 1C). Thus, we selected and successfully validated processing of several high confidence substrates with ratios  $\geq 10$  in follow-up biochemical experiments. Substrate candidates with iTRAQ ratios  $< 10$  are more likely to include false positive substrates, and so these must be validated biochemically. Therefore, we applied additional hierarchical substrate winnowing criteria (32), choosing proteins that are known substrates for other MMPs or proteins that have other family members processed by MMP-2, MMP-9, or other MMPs (18). An important qualification is that although human and murine MMP-2 are remarkably homologous, showing 97% identity with only eight conservative substitutions found in the catalytic domain, it is possible that subtle differences might occur between the substrate specificities determined by TAILS for human MMP-9 cleaving murine proteins *versus* human MMP-9 cleaving human substrates in the biochemical validation experiments where murine proteins could not be obtained.

The therapeutic cancer target galectin-1 (52) is a known MMP-2 (19) and MMP-11 (32) substrate, and another family member, galectin-3, is processed by MMP-2 and MMP-9 (53).

Galectin-1 was identified by iTRAQ-TAILS as a high confidence substrate of MMP-2 with a ratio of  $\geq 35$  for the neo-N-terminus produced by cleavage at  $ACG \downarrow^4LVA$  (supplemental Table 12) and as a candidate MMP-9 substrate with a ratio of 3.63 (Fig. 4, A and B). Galectin-1 was also identified by the original acetylated (Ac) N-terminus,  $Ac\text{-}^2ACGLVASNLNLK\text{-}PGECLKVR$  (supplemental Table 11, peptide 124), which was also labeled at two lysine residues (K) and exhibited an iTRAQ ratio of 0.35 and 0.44 for MMP-2- and MMP-9-treated samples, respectively, consistent with cleavage near the protein N-terminus to deplete this peptide. Identification of the same substrate by two complementary peptides with mirrored iTRAQ ratios is a hallmark of the TAILS approach and lends higher confidence to such candidate substrates.

Incubation of recombinant human galectin-1 with MMP-2 and MMP-9 resulted in lower molecular weight laddering by both MMPs (Fig. 4C). Thus, we identified galectin-1 as a novel MMP-9 substrate by TAILS and biochemically confirmed the result. In addition, we detected two novel high confidence MMP-2 cleavage sites at positions  $ACG \downarrow^5LVA$  (ratio,  $\geq 35$ ) (supplemental Table 12, peptide 31) and  $ASN \downarrow^9LNL$  (ratio, 9.57) (supplemental Table 10, peptide 1682). It should be noted that previously reported MMP-2 cleavage sites of human galectin-1 determined by Edman degradation (19) were not detected because processing at these cleavage sites would result in peptides too short for MS analysis: cleavage at  $PGE \downarrow^{16}CLR$  in human galectin-1 corresponds to  $PGE \downarrow^{16}CLKVR$  in murine galectin-1, which generates a peptide only five amino acids long. The second site,  $KCV \downarrow^{132}AFD$ , near the C terminus corresponds to



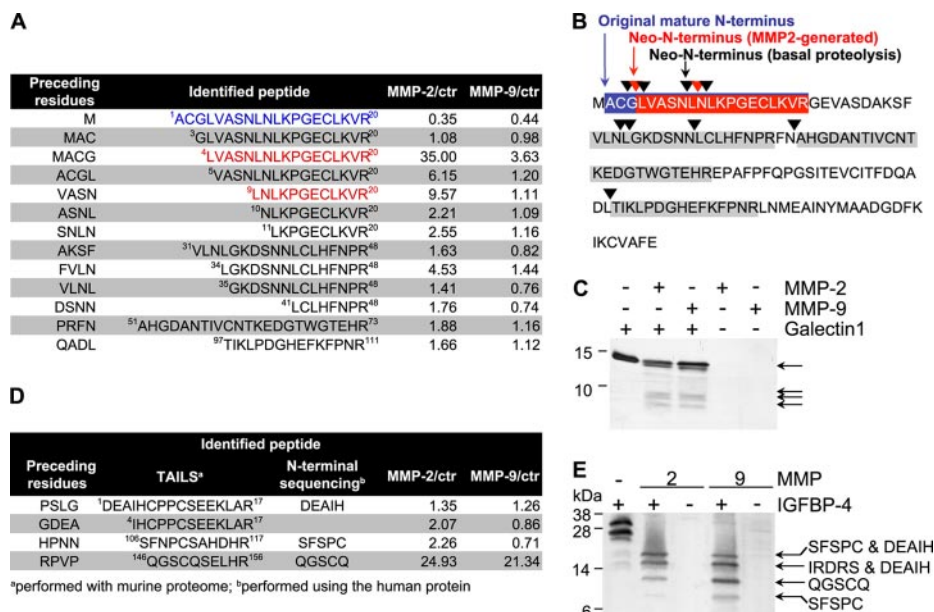


FIG. 4. Galectin-1 and IGFBP-4 are novel substrates of MMP-9. **A**, summary of galectin-1 peptides identified by iTRAQ-TAILS analysis in fibroblast secretomes treated with MMP-2 or MMP-9. **B**, the peptides identified by iTRAQ-TAILS are highlighted within the sequence of human galectin-1. Peptides with iTRAQ ratios  $\geq 10$  were identified as high confidence cleavage fragments and are shown in red. The peptide representing the mature original N-terminus is shown in blue. The neo-N-termini due to proteolysis by MMP-2 are indicated by red arrowheads, and those generated by proteolysis by other proteases in the sample, as suggested by the lower ratios, are indicated by black arrowheads. **C**, recombinant human galectin-1 was incubated for 18 h at 37 °C at a 10:1 molar ratio with MMP-2, MMP-9, or buffer alone, or MMP-2 and MMP-9 were incubated alone. The digestion products were separated on a 15% Tris-Tricine gel and visualized by silver staining. Arrows indicate cleavage fragments of galectin-1. **D**, summary of IGFBP-4 peptides identified by Edman sequencing in **E** and their corresponding iTRAQ ratios identified by iTRAQ-TAILS analysis of fibroblast secretomes treated with MMP-2 or MMP-9. **E**, recombinant human IGFBP-4 was incubated for 18 h at 37 °C at a 10:1 molar ratio with MMP-2, MMP-9, or buffer alone, or MMP-2 and MMP-9 were incubated alone as indicated. The digestion products were separated on a 15% Tris-Tricine gel followed by transfer to the PVDF membrane, R-250 Coomassie Blue staining, and Edman sequencing of the visible cleavage fragments. Cleavage products with their corresponding sequences identified by Edman degradation are indicated. ctr, control.

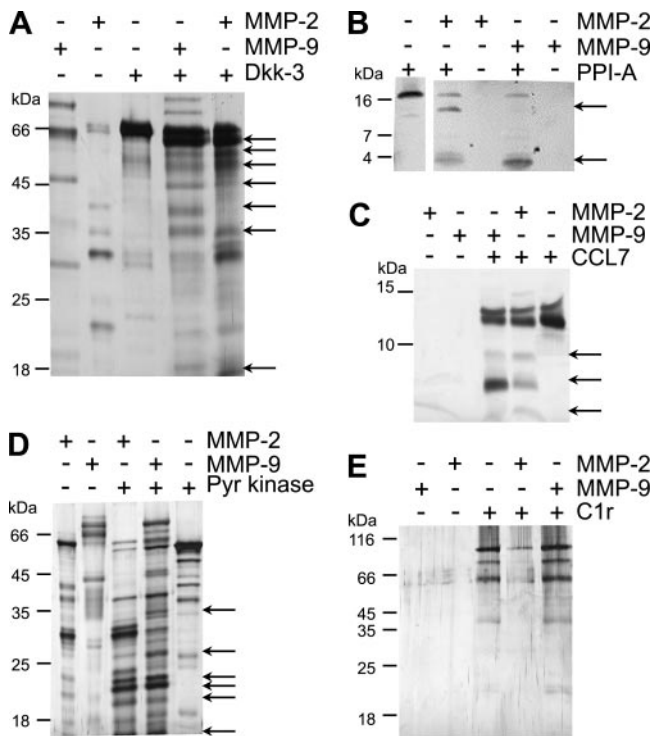
KCV ↓ <sup>132</sup>AFE in murine galectin-1 with cleavage at this site producing a three-residue peptide that is not identifiable by MS.

IGFBP-4 and IGFBP-6 are known MMP-2 substrates (19), but their cleavage sites are unknown. IGFBP-4 is a high confidence substrate of both MMP-2 and MMP-9 with the identification of four N-terminal peptides (Fig. 4D), one of which, <sup>146</sup>QGSC-QSELHR (supplemental Table 12, peptide 77), had iTRAQ ratios of 25 and 21, respectively. *In vitro* incubation of recombinant human IGFBP-4 resulted in similar processing by MMP-2 and MMP-9 to four distinct lower molecular weight fragments. Edman sequencing of digested human IGFBP-4 revealed the mature protein N-terminal sequence <sup>1</sup>DEAIH and the neo-N-terminal sequences SFSPC and QGSCQ. Notably, these are the exact cleavage sites identified by iTRAQ-TAILS in the murine homolog (Fig. 4E). The fourth cleavage site identified by Edman sequencing, IRDRS, was not detected by TAILS because it would be represented by a two-residue-long neo-N-terminal peptide (underlined) that cannot be identified by MS. Another TAILS peptide, <sup>4</sup>IHCPCSEEKLAR (supplemental Table 10, peptide 1080), where cleavage occurred close to the N-terminus <sup>1</sup>DEAIHCPCSEEKLAR, was not detected by Edman se-

quencing, likely due to the limited resolution of gel-based methods. Notably, the identification of the same cleavages in human proteins by Edman degradation as those detected in murine homologs by TAILS demonstrates the validity of our approach of using human protease in a murine system.

Pyruvate kinase M1/M2 was identified as a high confidence MMP-2 substrate and as a candidate MMP-9 substrate as evident from three<sup>2</sup> distinct peptides, <sup>17</sup>LHAAMADTFLEH-MCR (supplemental Table 12, peptide 148), <sup>21</sup>MADTFLEH-MCR (supplemental Table 12, peptide 141), and <sup>109</sup>VALDT-KGPEIR (supplemental Table 12, peptide 173), with MMP-2 iTRAQ ratios of 15.27, 16.08, and 13.11 and MMP-9 iTRAQ ratios of 2.15, 1.41, and 1.62, respectively. Incubation of

<sup>2</sup> An alternative analysis of the TAILS data set (data not shown) using PeptideProphet (as described in Ref. 40) yielded an additional high confidence peptide of pyruvate kinase M1/M2, <sup>422</sup>CCSGAIV-LTKSGR, that was not identified in iProphet analysis probably due to the differences in the algorithm between the two platforms. Observed iTRAQ ratios for this peptide of 1.08 and 9.21 for MMP-2 and MMP-9, respectively, provide further evidence of differential processing of pyruvate kinase by these proteases and are consistent with *in vitro* cleavage patterns obtained with purified proteins.



**FIG. 5. Biochemical validation of MMP-2 and MMP-9 substrates.** Recombinant proteins identified by iTRAQ-TAILS as high confidence or potential substrates were incubated for 18 h at 37 °C with MMP-2, MMP-9, or buffer alone, or MMP-2 and MMP-9 were incubated alone. **A**, digestion products of recombinant human Dkk-3 were separated by 10% SDS-PAGE. Cleavage fragments of recombinant human peptidyl-prolyl cis-trans isomerase A (*PPI-A*) (**B**) and CCL7 (**C**) were separated on 15% Tris-Tricine gels. Digestion products of recombinant pyruvate kinase M1/M2 (**D**) and human C1r subcomponent A (**E**) were separated by 10% SDS-PAGE. Digestion products were visualized by silver staining except for peptidyl-prolyl cis-trans isomerase A, which was detected by Western blotting using the corresponding rabbit polyclonal antibody. Arrows indicate major cleavage products. C1r protein was degraded by MMP-2 as shown by loss of the intact protein band. Autolysis of the MMP-2 and MMP-9 proteases during the 18-h assay often occurred and is typical, although the proteases used were purified to single band homogeneity as shown in supplemental Fig. 2. *Pyr*, pyruvate.

pyruvate kinase M1/M2 with MMP-2 and MMP-9 resulted in loss of intensity of the full-length protein (more so with MMP-2 than with MMP-9) and appearance of multiple lower molecular weight fragments with distinct pattern differences between the two MMPs (Fig. 5).

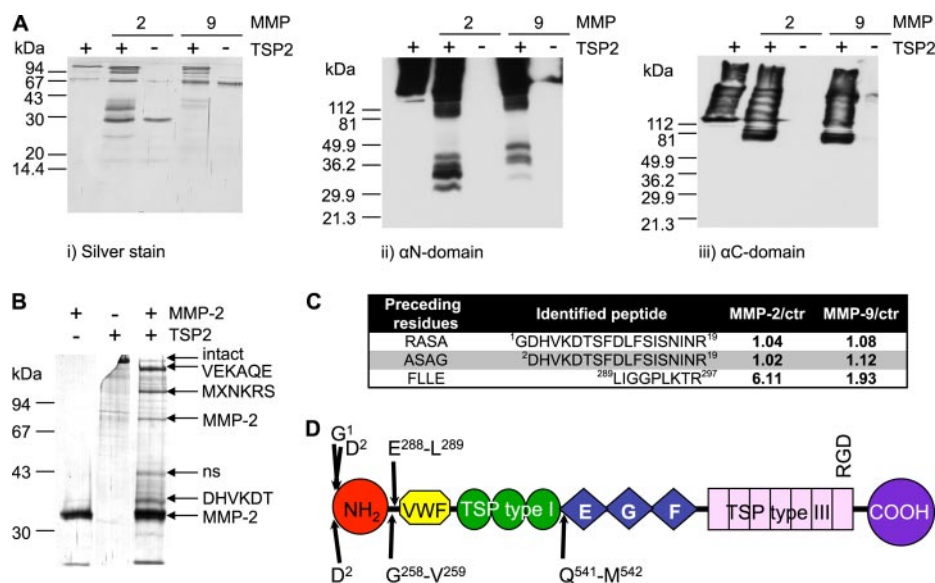
Complement C1q is cleaved by MMP-1, -2, -3, and -9 (54) as is the related protein mannose binding lectin (55). We identified complement C1r-A subcomponent as a high confidence MMP-2 substrate with an iTRAQ ratio of  $\geq 35$  for the cleavage site VSS↓<sup>191</sup>LEY (supplemental Table 12, peptide 13). For MMP-9, C1r-A was only a candidate substrate because it had an iTRAQ ratio of 4.36, considerably less than the statistically defined cutoff ratio of 10 needed for high confidence substrate identification. Indeed, full-length C1r-A was degraded by MMP-2, which caused disappearance of the

protein, but MMP-9 incubation had no effect on C1r-A (Fig. 5). This illustrates the higher false positive rate at lower iTRAQ ratios and underscores the importance of stringent statistically derived cutoffs for high confidence substrate identification rather than using arbitrary values such as 2-fold. Nonetheless, other lower iTRAQ ratio candidates that were successfully validated as novel MMP-9 substrates included dickkopf-related protein-3 (Fig. 5), previously shown to be processed by MMP-2 at an unknown site (19) but identified here, whereas the related dickkopf-1 protein was processed by MMP-14 (18). We also biochemically verified a known MMP-2 cleavage of peptidyl-prolyl cis-trans isomerase A (32) (Fig. 5) and confirmed a novel MMP-9 cleavage prompted by previous studies that detected processing of peptidyl-prolyl cis-trans isomerase A by MMP-14 (18). Human chemokine CC7 was used as a positive control for MMP-2 (Fig. 5), and we found for the first time that it too was cleaved by MMP-9. Finally, thrombospondin-1 and thrombospondin-2 are MMP-2 (19) and MMP-14 (18) substrates, respectively. Several neo-N-termini in thrombospondin-2 were identified (supplemental Table 12), and its cleavage by MMP-2 and MMP-9 was confirmed (Fig. 6 and supplemental Fig. 4).

**Analysis of MMP-2 and MMP-9 Substrate Specificity—**MMP-2 and MMP-9 are 72% identical and also possess a remarkably high degree of similarity in their structures with the main differences being a longer glycosylated hinge region in MMP-9 as well as in the S1' selectivity pocket of the active site (56, 57). Therefore, we tested whether the structural differences between MMP-2 and MMP-9 are discernible in their substrate specificity profiles. We used the substrate cleavage sites that were identified at high confidence in native proteins to create a consensus sequence spanning four amino acids upstream and downstream of the cleavage site (positions P4 to P4'). Similar to that previously reported for MMP-2 using denatured proteome-derived peptide libraries in the proteomic identification of protease cleavage site (PICS) technique (43), our consensus sequence resulting from digestion of native proteins showed a strong preference for Leu in position P1' and Pro in P3. On the other hand, MMP-9 exhibited a more relaxed preference for Leu in P1' and still a very strong preference for Pro in P3, reflecting the differences in the S1' subsites of these gelatinases (Fig. 7).

#### DISCUSSION

TAILS was recently demonstrated to be an efficient yet simple technology for N-terminome characterization and substrate discovery (32). TAILS relies on negative selection using a novel polyaldehyde polymer to enrich protein N-termini and isotopic dimethylation to distinguish specific protease cleavage products from basal proteolysis products generated by other proteases that are present in every sample. Without quantification, it can be extremely difficult to accurately discern protease-specific cleavage events from cleavage signatures already present in the sample. Nonetheless, similar to



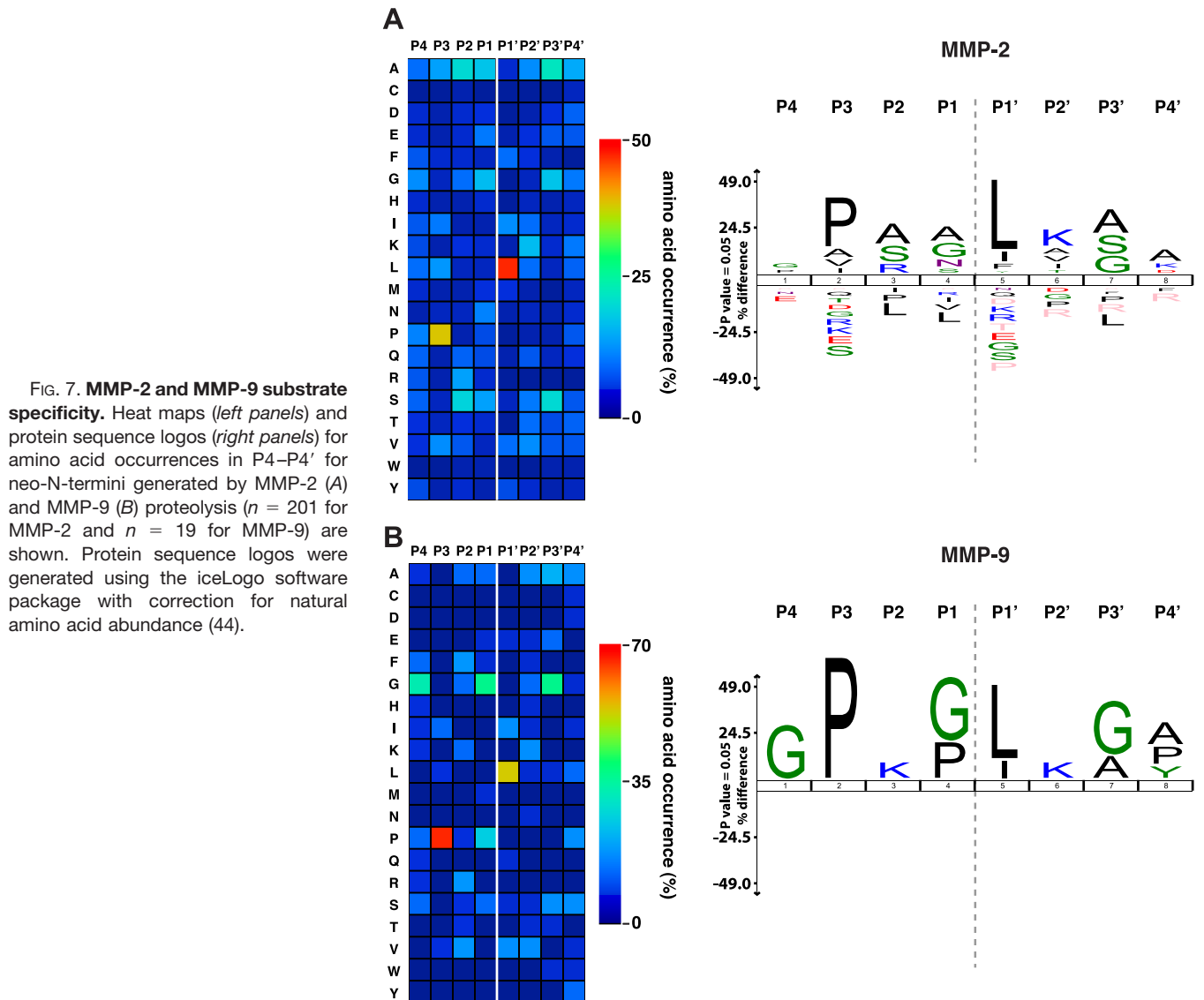
**FIG. 6. Thrombospondin-2 is a novel MMP-2 and MMP-9 substrate.** *A*, recombinant human thrombospondin-2 (*TSP2*) was incubated for 18 h at 37 °C at a 10:1 molar ratio with MMP-2, MMP-9, or buffer alone. The digestion products were separated on by 10% SDS-PAGE and visualized by silver staining (*left panel*) or detected by Western blotting using an antibody raised against the N-terminal (*middle panel*) or C-terminal domain of the protein (*right panel*). *B*, fragments of thrombospondin-2 after incubation with MMP-2 were transferred to a PVDF membrane, and N-terminal sequences were derived by Edman degradation as indicated by arrows. *ns* indicates that no sequence was obtained. All lanes were from the same gel. *C*, summary of thrombospondin-2 peptides and corresponding iTRAQ ratios identified by iTRAQ-TAILS in fibroblast secretomes treated with MMP-2 or MMP-9. *D*, schematic diagram of thrombospondin-2. N-terminal, von Willebrand factor (*VWF*), thrombospondin (*TSP*) type I, epidermal growth factor (*EGF*), thrombospondin (*TSP*) type III, and C-terminal domains are shown. Positions of cleavage sites or the mature N-terminus of the murine protein identified by iTRAQ-TAILS are shown above, and those identified by Edman degradation of the human protein after digestion by human MMP-2 are shown below the protein. *ctr*, control.

other N-terminal enrichment methods using MS1 quantification with SILAC (23, 34) and <sup>18</sup>O labeling (21, 22), TAILS labeling by dimethylation results in spectral duplication and an increase in sample complexity, thus limiting proteome coverage. Furthermore, increased complexity precludes sequencing of every peptide pair, thus potentially increasing false positive peptide assignments and quantification.

Although formaldehyde is widely available and inexpensive for dimethylation labeling, there are just two isotopic variants, limiting each analysis to two samples. By using isotopic cyanoborohydride, a third label can be introduced (58). More expensive SILAC reagents allow for comparison of up to five conditions. However, the use of more isotopic variants exacerbates the problem of multiplied peptide spectra and requires multiple rounds of chromatography and MS analyses, up to 150 per data set, for excellent proteome coverage by COFRADIC, which is expansive but expensive (23). In contrast, labeling with isobaric mass tags, e.g. iTRAQ reagents, allows for quantification in MS2 mode and simultaneous comparison of up to four and recently eight samples (using 8-plex iTRAQ reagents; Applied Biosystems) without spectral multiplication. Also, iTRAQ labeling increases the coverage of lower abundance proteins as the amount of precursor ion is amplified in MS1 mode by summation of identical peptides labeled with different isotopic iTRAQ tags from contributing samples. Following fragmentation in MS2 mode, peptides can be identified and quantified from the

same MS/MS spectra, thus minimizing the possibility of erroneous quantification based on misidentifications that can occur with MS1 quantification (59, 60).

Our novel protocol for efficient, rapid, and amino group-specific whole protein iTRAQ labeling addresses these issues. High fidelity labeling and virtually complete (97 ± 3%) blocking of lysine side chains and N-terminal primary amino groups with minimal side reactions on Tyr, Ser, and Thr side chains minimize false positive assignments of cleavage events for peptides commencing with any of these residues that can be seen in other N-terminal characterization methods (61). High labeling efficacy is essential for accurate quantification and assures minimal losses of original and neo-N-termini (by coupling to the polymer if it were incomplete), thereby promoting maximal proteome coverage. Notably, iTRAQ labeling is rapid compared with dimethylation and reduces the workflow time of TAILS from 3 to 2 days. The polymer is also key to the success of TAILS. With virtually undetectable nonspecific peptide binding, this step enhances sample recovery and is extremely fast (32). Finally, due to the ready cyclization of glutamine, cysteine (cyclized in proteomics workflows), and to a lesser extent glutamate, when screening for substrates of proteases cleaving at these residues, only a negative peptide selection procedure reliably detects such cleavage products as these are unreactive in positive selection workflows.



Accurate quantification of iTRAQ ratios is a key feature of TAILS as it enables differentiation of proteolytic activity of the test protease from basal proteolysis that occurs in every biological sample. This is especially crucial when analyzing proteases of broad or unknown specificity where manual data parsing cannot be performed to search for known cleavage sites such as for proteases like Glu-C, granzymes, and caspases. To assure accurate identification of proteins from single peptides, stringent criteria were developed before (32). Now, for accurate quantification of the single peptides, we have developed and applied a rigorous bioinformatics protocol for data analysis to the TAILS platform (described in detail in the accompanying paper (40)). Specifically, we built peptide ion intensity-dependent error models to account and correct for peptide ion intensity-dependent variation in iTRAQ ratios. Furthermore, we used Glu-C, a protease of canonical specificity, to digest a com-

plex proteome to statistically derive an iTRAQ ratio cutoff for high confidence substrate identification. The use of stringent statistical validation assures low rates of false positive substrate identification and renders iTRAQ-TAILS a highly robust assay system so elevating it from a substrate identification “screen.” Thus, our iTRAQ-TAILS data analysis yields two broad categories of substrates: (i) high confidence identifications with iTRAQ ratios  $\geq 10$  that require virtually no further validation and (ii) candidate substrates with iTRAQ ratios  $< 10$  that must be confirmed in follow-up biochemical experiments to differentiate between true and false positive hits. To select the most promising candidates for validation in this study, we also applied additional hierarchical substrate winnowing criteria developed before (32) where priority was given to proteins with family members that are known substrates of MMP-2, MMP-9, or other MMPs or that are proteins cleaved by other MMPs.



We used iTRAQ-TAILS to analyze the substrate degradome of the closely related MMP-2 and MMP-9 proteases that are important in cancer and chronic inflammation (62, 63). Using an MMP-2/MMP-9-naïve proteome from murine fibroblasts incubated with recombinant human MMP-2 or MMP-9, TAILS identified 201 MMP-2 cleavage sites in 162 proteins, whereas MMP-9 processed a much smaller subset of 19 proteins under identical conditions. The control APMA channel showed no evidence of endogenous MMP activity. Among the MMP-2 and MMP-9 substrates identified were several known MMP targets, validating the approach. For instance, for the known MMP-2 substrate IGFBP-4 (19), we report a novel cleavage site by MMP-2 (PVP ↓<sup>147</sup>QGS) and previously undescribed processing by MMP-9 at the same location that were biochemically validated *in vitro*. TAILS also identified cystatin C as a high confidence substrate of both MMP-2 and MMP-9 with a cleavage site homologous to that in human cystatin C (19) that was confirmed by MALDI-TOF MS for MMP-9 for the first time (supplemental Fig. 5). In total, we biochemically confirmed the cleavage of seven new MMP-9 targets (three high confidence hits (cystatin C, galectin-1, and IGFBP-4) and four candidate substrates (thrombospondin-2, pyruvate kinase isozymes M1/M2, Dkk-3, and peptidyl-prolyl cis-trans isomerase A)) and four new MMP-2 substrates (two high confidence hits (pyruvate kinase M1/M2 isozymes and complement C1r-A subcomponent) and two candidate substrates having lower ratios (peptidyl-prolyl cis-trans isomerase A and thrombospondin-2)). Interestingly, although complement C1r-A subcomponent was identified by TAILS as a potential MMP-9 substrate (iTRAQ ratio of 4.36), no cleavage could be detected in follow-up *in vitro* biochemical analyses. This is contrary to processing of the same protein by MMP-2 that was detected by TAILS with high confidence (iTRAQ ratio of  $\geq 35$ ) and successfully confirmed *in vitro*. This example provides an illustration of the high rate of true positive substrate identification at iTRAQ ratios  $\geq 10$  and a higher rate of false positive substrates at lower iTRAQ ratios that are predicted by our statistical model for substrate ratio cutoff. This also emphasizes the need for statistically defined ratio cutoffs rather than selecting arbitrary values.

It should be noted that iTRAQ-TAILS substrate discovery is limited to detection of stable unique N-terminal cleavage fragments that are sufficiently long for MS analysis upon trypsin digest after arginine residues. For instance, if the protease under study processes a protein at the same site as other proteases this would result in a decreased iTRAQ ratio (with the extent of decrease dependent on the relative efficiency of cleavage by “basal” versus “experimental” proteases) and thus lower the chance of cleavage fragment detection. This scenario is illustrated by examples of thrombospondin-2 and IGFBP-4 where several low ratio peptides were confirmed by Edman sequencing to be the products of true cleavage events (Figs. 4 and 6). The second limiting factor, *i.e.* the size of the fragment, is embedded in the

proteome sequences but can be easily overcome by blocking lysine residues as performed in TAILS and COFRADIC and using proteases other than trypsin for digestion of the proteome. *In silico* analysis of the mouse proteome indicates that tryptic digest alone would result in 50% unambiguous coverage of N-terminal peptides, whereas using both trypsin and Glu-C in parallel would increase the coverage to 80% with a further increase to 90% if higher accuracy mass spectrometers are used (64).

A unique advantage of TAILS over other N-terminal enrichment methods is the retention and quantification of original N-termini. This is facilitated by two key features of the TAILS protocol: negative selection, which allows for retention of all original N-termini including those that are cyclized or blocked by acetylation as seen in up to 85% of eukaryotic proteins (65), and labeling of lysine side chains, which allows quantification of even N-terminally blocked peptides. Interestingly, iTRAQ-TAILS analysis of *Mmp2*<sup>-/-</sup> secretomes indicates that >95% of all acetylated N-termini harbor lysine residues N-terminal of the first arginine and so after tryptic digestion are labeled and quantified by TAILS. In contrast, only 40–60% of all acetylated peptides identified by TAILS using dimethylation to block lysines are quantifiable,<sup>3</sup> likely due to better ionization of iTRAQ-labeled peptides compared with dimethylated counterparts, a further advantage of iTRAQ-TAILS.

The retention of N-terminally naturally blocked peptides allowed for the identification of several substrates by both their original N-termini and overlapping neo-N-termini with inverse ratios. For digestion of galectin-1 by MMP-2 and MMP-9, the iTRAQ ratios of the acetylated original N-terminus Ac-<sup>1</sup>ACGLVASNLNLKPGECKLVR were 0.35 and 0.44, respectively, showing depletion compared with the control sample. In comparison, the neo-N-terminus <sup>4</sup>LVASNLNLKPGECKLVR had ratios of  $\geq 35$  and 3.63 for MMP-2 and MMP-9, respectively. Notably, the galectin-1 original N-terminus was acetylated after removal of the initiator methionine, thus providing evidence of a posttranslational modification that was previously inferred by similarity but not experimentally demonstrated. Cofilin-1, a known MMP-2 substrate, was identified by multiple peptides that included the original mature acetylated N-terminus Ac-<sup>1</sup>ASGVAVSDGVIKVFNDMKVR with iTRAQ ratios of 0.51 and 0.53 for MMP-2 and MMP-9, respectively, and the complementary neo-N-terminus <sup>4</sup>VAVSDGVIKVFNDMKVR with corresponding iTRAQ ratios of 20.79 and 1.28, respectively. Although a high MMP-2 iTRAQ ratio for the second peptide defines this fragment as an MMP-2 cleavage product and hence cofilin-1 as a high confidence substrate, MMP-9 cleavage can be inferred indirectly from the decreased (0.53) iTRAQ ratio for the original N-terminus, but this needs validation.

In other cases, additional cleavage events can also be inferred from the decreased iTRAQ ratio of the original N-

<sup>3</sup> O. Kleifeld, A. Doucet, and C. M. Overall, unpublished data.

termini even if the corresponding neo-N-terminus is unsuitable for MS analysis and goes undetected. For example, a membrane protein, ATP synthase-coupling factor 6, was identified only by its original mature N-terminus, <sup>1</sup>NKELD-PVQKLFVDKIR, with iTRAQ ratios of 0.53 and 0.90 for MMP-2 and MMP-9, respectively. In the absence of identified neo-N-termini, these results suggest proteolytic processing of the protein by MMP-2. Hence, retention of original mature protein N-termini by TAILS and their quantification through lysine labeling expand substrate coverage, proteome annotation, and characterization of posttranslational N-terminal processing such as methionine or propeptide removal, acetylation, cyclization etc. as well as substrate identification in a single analysis. Thus, iTRAQ-TAILS has a unique advantage over other global methods for characterization of proteolytic events where proteolytic processing of proteins with modified N-termini would be overlooked: methods utilizing positive selection (24, 25); those omitting naturally acetylated original N-termini from analysis or cleaved peptides containing P1' cyclized Asn or Glu or Cys (which cyclizes due to the chemicals used in the proteomics workflow) (20); those using non-quantitative negative selection (64) or enzymatic biotinylation, which does not recognize Pro and has lower specificity for Glu and Asp (24, 25); and gel-based approaches (27–30), which lack resolution.

The highly efficient processing of one or two specific bonds in many bioactive signaling molecules profoundly alters their activity and cell function in a variety of processes including chemotaxis, mitogenesis, and angiogenesis. Processing of the novel substrates provides insights into new important roles and potential mechanisms of MMP-2 and MMP-9 in regulating cells, the pericellular environment, and immune responses. The new MMP-9 substrate Dkk-3 is a tumor suppressor protein that is an antagonist of the Wnt signaling pathway (66, 67). Overexpression of Dkk-3 arrests growth and induces apoptosis in cancer cell lines, ultimately resulting in their lower tumorigenicity in mice (66, 68). In contrast, down-regulation of Dkk-3 levels observed in many tumors correlates with faster progression and shorter survival times (69) and was previously attributed to transcriptional suppression via promoter hypermethylation (70). Our results demonstrating Dkk-3 processing by MMP-2 and MMP-9 provide an alternative mechanism for Dkk-3 down-regulation at the posttranslational level and are consistent with the proangiogenic role of MMP-2 and MMP-9.

Pyruvate kinase M1/M2, a multitasking intracellular/extracellular (7, 51) glycolytic enzyme (71) that had been implicated as a cancer marker and a therapeutic target (45), was identified as a novel MMP-2 and MMP-9 substrate by multiple cleavage fragments that were biochemically validated. In tumors, pyruvate kinase acts as a metabolic switch activated by oncoproteins: the dimeric form favored by binding of oncoproteins controls consumption of glucose carbons into biosynthetic processes as compared with the tetrameric form, which

results in glycolytic ATP production (72). In response to apoptotic stimuli, pyruvate kinase is translocated into the nucleus where it induces cell death in a manner independent of caspases or its own enzymatic activity (73–75). Pyruvate kinase M1/M2 is secreted by tumors into plasma, although a specific function here remains to be determined (76). Given the many facets of pyruvate kinase activity, functional modulation by MMP-2 and MMP-9 processing is of great interest and is currently under investigation in our laboratory.

We have also demonstrated that galectin-1, a cancer drug target (52), is a novel MMP-9 substrate. MMP-2, MMP-9, and MMP-14 were previously shown to cleave galectin-3 (18, 53), and galectin-1 has been validated *in vitro* as an MMP-2 (19) and MMP-11 (32) substrate. Although the earlier study successfully used MALDI MS to identify two cleavage sites in galectin-1 (19), we were able to detect additional cleavage sites at ACG ↓ <sup>4</sup>LVA and ASN ↓ <sup>9</sup>LNL.

Growth factor-binding proteins from the IGFBP superfamily are known substrates of many MMPs (18, 19, 77–80). For instance, MMP-2 cleaves connective growth tissue factor (77), IGFBP-3, -4, and -6, albeit at unknown sites (19, 79), whereas MMP-9 processes IGFBP-1 and -3 (81, 82). We confirmed processing of IGFBP-4 by MMP-2 and identified IGFBP-4 as a novel MMP-9 substrate. TAILS analysis identified four novel cleavage sites, three of which were confirmed by Edman sequencing. IGFBPs have been previously shown to modulate stability and/or biological functions of insulin-like growth factors (IGFs) and also affect cellular responses in an IGF-independent manner (83). For instance, release of IGF-1 through proteolytic clearance of IGFBP-4 by pregnancy-associated plasma protein-A protease promotes tumor angiogenesis and inhibits apoptosis in a breast cancer mouse model (84). We predict that processing of IGFBP-4 by MMP-2 and MMP-9 will increase bioavailability of IGF and lead to enhanced cell migration and tumorigenesis, thus providing a new MMP-2 and MMP-9 proangiogenic mechanism.

We identified and confirmed the novel processing of thrombospondin-2 by MMP-2 and MMP-9. Intact thrombospondin-2 prevents tumor angiogenesis and induces apoptosis of vascular endothelial cells *in vivo* (85, 86). Thus, proteolytic clearance is likely to be proangiogenic, consistent with the role of gelatinases in cancer progression. In addition, release of any of the functional domains of the multimodular thrombospondin-2 (87) by proteolysis might also result in different biological outcomes, but this requires further investigation.

The similar substrate specificity of MMP-2 and MMP-9 observed in this study and illustrated by overlapping substrate degradomes and similar substrate consensus sequences can be explained by the high degree of structural similarity between the two gelatinases that share the same modular organization (57, 88). Indeed, some studies report a lack of phenotype in MMP-2 and MMP-9 knock-out mice in certain models (89–92), suggesting redundant activity and mutual compensation for each other. On the contrary, other animal

and cancer patient studies demonstrate measured differences in the roles of these two MMPs (93, 94). Thus, although ablation of both MMP-2 and MMP-9 resulted in decreased tumor number and growth (95, 96), only MMP-9 triggered the angiogenic switch during carcinogenesis albeit via an unknown substrate (95). Furthermore, MMP-2 and MMP-9 have been shown to act as antagonists or to cooperate with each other, depending on the experimental model (97–101). Although MMP-2 is widely expressed and is a highly abundant protease in serum, MMP-9 is produced by fewer types of cells and can be sequestered in neutrophil granules where its expression and release are tightly controlled (102–105). Potentially other important MMP-9 substrates will be found by screening other proteomes, but nonetheless, it was unexpected that MMP-9 cleaved so few of the proteins present in the fibroblast secretome under identical conditions as MMP-2. Thus, the apparent redundancy in activity between the two enzymes might be compensated for, at least partially, by their separation in time and space.

In conclusion, iTRAQ-TAILS is a high throughput and multiplex quantitative proteomics technique for N-terminome characterization and global substrate discovery. Overall, we have identified multiple known and, most importantly, many novel substrates of MMP-2 and MMP-9 with their corresponding cleavage sites in a multiplex manner for the first time. A big advantage of multiplex analyses is the higher confidence in comparisons made between different samples, such as the comparison of two closely related proteases. Here, under identical experimental conditions for digestion and simultaneous sample handling, TAILS, and MS analysis, a vastly more expansive substrate repertoire was identified for MMP-2 over MMP-9. This suggests that the controlled expression and activation of these two proteases at different sites and times provides some of the explanation for the differences in phenotypes of their corresponding gene knock-out mice. Processing of the new protein substrates by MMP-2 and MMP-9 identified here provides new mechanisms for gelatinase involvement in angiogenesis and carcinogenesis rather than extracellular matrix degradation (7, 51) and suggests new avenues for therapeutic intervention.

**Acknowledgments**—We thank Dr. Wei Chen from the University of British Columbia Centre for Blood Research Mass Spectrometry Suite for excellent mass spectrometry analyses. We are grateful to Dr. Paul Bornstein (University of Washington, Seattle, WA) for providing human thrombospondin-2 and the thrombospondin-2 antibodies used in this study.

\* This work was supported in part by a grant from the Canadian Institutes of Health Research, a program project grant in Breast Cancer Metastases from the Canadian Breast Cancer Research Alliance with funds from the Canadian Breast Cancer Foundation and the Cancer Research Society, and an infrastructure grant from the Michael Smith Foundation for Health Research.

□ This article contains supplemental Figs. 1–5 and Tables 1–13.

‡ Both authors contributed equally to this work.

§ Supported by the University of British Columbia Centre for Blood Research Strategic Training Program in Transfusion Science.

¶ Supported by a German Research Foundation (Deutsche Forschungsgemeinschaft) research fellowship. Present address: ETH Zurich, Inst. of Cell Biology, Schafmattstrasse 18, CH-8093 Zurich, Switzerland.

|| Supported by a Canada Research Chair in Metalloproteinase Proteomics and Systems Biology. To whom correspondence should be addressed. Tel.: 604-822-2958; Fax: 604-822-7742; E-mail: chris.overall@ubc.ca.

## REFERENCES

- Doucet, A., Butler, G. S., Rodríguez, D., Prudova, A., and Overall, C. M. (2008) Metadegradomics: toward in vivo quantitative degradomics of proteolytic post-translational modifications of the cancer proteome. *Mol. Cell. Proteomics* **7**, 1925–1951
- Ajami, K., Pitman, M. R., Wilson, C. H., Park, J., Menz, R. I., Starr, A. E., Cox, J. H., Abbott, C. A., Overall, C. M., and Gorrell, M. D. (2008) Stromal cell-derived factors 1alpha and 1beta, inflammatory protein-10 and interferon-inducible T cell chemo-attractant are novel substrates of dipeptidyl peptidase 8. *FEBS Lett.* **582**, 819–825
- McQuibban, G. A., Gong, J. H., Wong, J. P., Wallace, J. L., Clark-Lewis, I., and Overall, C. M. (2002) Matrix metalloproteinase processing of monocyte chemoattractant proteins generates CC chemokine receptor antagonists with anti-inflammatory properties in vivo. *Blood* **100**, 1160–1167
- Cox, J. H., and Overall, C. M. (2008) Cytokine substrates: MMP regulation of inflammatory mediator signalling, in *The Cancer Degradome: Proteases and Cancer Biology* (Edwards, D., Hoyer-Hansen, G., Blasi, F., and Sloane, B. F., eds) pp. 519–539, Springer, New York
- LeMosy, E. K., Hong, C. C., and Hashimoto, C. (1999) Signal transduction by a protease cascade. *Trends Cell Biol.* **9**, 102–107
- Affara, N. I., and Coussens, L. M. (2008) Proteolytic pathways: intersecting cascades in cancer development, in *The Cancer Degradome: Proteases and Cancer Biology* (Edwards, D., Hoyer-Hansen, G., Blasi, F., and Sloane, B. F., eds) pp. 157–182, Springer, New York
- Butler, G. S., and Overall, C. M. (2009) Updated biological roles for matrix metalloproteinases and new “intracellular” substrates revealed by degradomics. *Biochemistry* **48**, 10830–10845
- Barrett, A. J. (2004) Bioinformatics of proteases in the MEROPS database. *Curr. Opin. Drug Discov. Devel.* **7**, 334–341
- Tremblay, G. M., Janelle, M. F., and Bourbonnais, Y. (2003) Anti-inflammatory activity of neutrophil elastase inhibitors. *Curr. Opin. Investig. Drugs* **4**, 556–565
- Overall, C. M., and Kleifeld, O. (2006) Towards third generation matrix metalloproteinase inhibitors for cancer therapy. *Br. J. Cancer* **94**, 941–946
- Stoch, S. A., and Wagner, J. A. (2008) Cathepsin K inhibitors: a novel target for osteoporosis therapy. *Clin. Pharmacol. Ther.* **83**, 172–176
- Turk, B. (2006) Targeting proteases: successes, failures and future prospects. *Nat. Rev. Drug Discov.* **5**, 785–799
- Coussens, L. M., Fingleton, B., and Matrisian, L. M. (2002) Matrix metalloproteinase inhibitors and cancer: trials and tribulations. *Science* **295**, 2387–2392
- Overall, C. M., and López-Otín, C. (2002) Strategies for MMP inhibition in cancer: innovations for the post-trial era. *Nat. Rev. Cancer* **2**, 657–672
- Overall, C. M., and Kleifeld, O. (2006) Tumour microenvironment—opinion: validating matrix metalloproteinases as drug targets and anti-targets for cancer therapy. *Nat. Rev. Cancer* **6**, 227–239
- López-Otín, C., and Overall, C. M. (2002) Protease degradomics: a new challenge for proteomics. *Nat. Rev. Mol. Cell Biol.* **3**, 509–519
- Tam, E. M., Morrison, C. J., Wu, Y. I., Stack, M. S., and Overall, C. M. (2004) Membrane protease proteomics: Isotope-coded affinity tag MS identification of undescribed MT1-matrix metalloproteinase substrates. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 6917–6922
- Butler, G. S., Dean, R. A., Tam, E. M., and Overall, C. M. (2008) Pharmacoproteomics of a metalloproteinase hydroxamate inhibitor in breast cancer cells: dynamics of membrane type 1 matrix metalloproteinase-mediated membrane protein shedding. *Mol. Cell. Biol.* **28**, 4896–4914



19. Dean, R. A., and Overall, C. M. (2007) Proteomics discovery of metalloproteinase substrates in the cellular context by iTRAQ labeling reveals a diverse MMP-2 substrate degradome. *Mol. Cell. Proteomics* **6**, 611–623
20. Gevaert, K., Goethals, M., Martens, L., Van Damme, J., Staes, A., Thomas, G. R., and Vandekerckhove, J. (2003) Exploring proteomes and analyzing protein processing by mass spectrometric identification of sorted N-terminal peptides. *Nat. Biotechnol.* **21**, 566–569
21. Vande Walle, L., Van Damme, P., Lamkanfi, M., Saelens, X., Vandekerckhove, J., Gevaert, K., and Vandennebeele, P. (2007) Proteome-wide identification of Htra2/Omi substrates. *J. Proteome Res.* **6**, 1006–1015
22. Van Damme, P., Martens, L., Van Damme, J., Hugelier, K., Staes, A., Vandekerckhove, J., and Gevaert, K. (2005) Caspase-specific and non-specific in vivo protein processing during Fas-induced apoptosis. *Nat. Methods* **2**, 771–777
23. Van Damme, P., Maurer-Stroh, S., Plasman, K., Van Durme, J., Colaert, N., Timmerman, E., De Bock, P. J., Goethals, M., Rousseau, F., Schymkowitz, J., Vandekerckhove, J., and Gevaert, K. (2009) Analysis of protein processing by N-terminal proteomics reveals novel species-specific substrate determinants of granzyme B orthologs. *Mol. Cell. Proteomics* **8**, 258–272
24. Timmer, J. C., Enoksson, M., Wildfang, E., Zhu, W., Igarashi, Y., Denault, J. B., Ma, Y., Dummitt, B., Chang, Y. H., Mast, A. E., Eroshkin, A., Smith, J. W., Tao, W. A., and Salvesen, G. S. (2007) Profiling constitutive proteolytic events in vivo. *Biochem. J.* **407**, 41–48
25. Mahrus, S., Trinidad, J. C., Barkan, D. T., Sali, A., Burlingame, A. L., and Wells, J. A. (2008) Global sequencing of proteolytic cleavage sites in apoptosis by specific labeling of protein N termini. *Cell* **134**, 866–876
26. Chang, T. K., Jackson, D. Y., Burnier, J. P., and Wells, J. A. (1994) Subtiligase: a tool for semisynthesis of proteins. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 12544–12548
27. Guo, L., Eisenman, J. R., Mahimkar, R. M., Peschon, J. J., Paxton, R. J., Black, R. A., and Johnson, R. S. (2002) A proteomic approach for the identification of cell-surface proteins shed by metalloproteases. *Mol. Cell. Proteomics* **1**, 30–36
28. Dix, M. M., Simon, G. M., and Cravatt, B. F. (2008) Global mapping of the topography and magnitude of proteolytic events in apoptosis. *Cell* **134**, 679–691
29. Bredemeyer, A. J., Lewis, R. M., Malone, J. P., Davis, A. E., Gross, J., Townsend, R. R., and Ley, T. J. (2004) A proteomic approach for the discovery of protease substrates. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 11785–11790
30. Bech-Serra, J. J., Santiago-Josefat, B., Esselens, C., Saftig, P., Baselga, J., Arribas, J., and Canals, F. (2006) Proteomic identification of desmoglein 2 and activated leukocyte cell adhesion molecule as substrates of ADAM17 and ADAM10 by difference gel electrophoresis. *Mol. Cell. Biol.* **26**, 5086–5095
31. Simon, G. M., Dix, M. M., and Cravatt, B. F. (2009) Comparative assessment of large-scale proteomic studies of apoptotic proteolysis. *ACS Chem. Biol.* **4**, 401–408
32. Kleifeld, O., Doucet, A., auf dem Keller, U., Prudova, A., Schilling, O., Kainthan, R. K., Starr, A. E., Foster, L. J., Kizhakkedathu, J. N., and Overall, C. M. (2010) Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products. *Nat. Biotechnol.* **28**, 281–288
33. Guo, K., Ji, C., and Li, L. (2007) Stable-isotope dimethylation labeling combined with LC-ESI MS for quantification of amine-containing metabolites in biological samples. *Anal. Chem.* **79**, 8631–8638
34. Gioia, M., Foster, L. J., and Overall, C. M. (2009) Cell-based identification of natural substrates and cleavage sites for extracellular proteases by SILAC proteomics. *Methods Mol. Biol.* **539**, 131–153
35. Butler, G. S., Tam, E. M., and Overall, C. M. (2004) The canonical methionine 392 of matrix metalloproteinase 2 (gelatinase A) is not required for catalytic efficiency or structural integrity: probing the role of the methionine-turn in the metzincin metalloprotease superfamily. *J. Biol. Chem.* **279**, 15615–15620
36. Pedrioli, P. G. (2010) Trans-proteomic pipeline: a pipeline for proteomic analysis. *Methods Mol. Biol.* **604**, 213–238
37. Deutsch, E. W., Mendoza, L., Shteynberg, D., Farrah, T., Lam, H., Tasman, N., Sun, Z., Nilsson, E., Pratt, B., Prazen, B., Eng, J. K., Martin, D. B., Nesvizhskii, A. I., and Aebersold, R. (2010) A guided tour of the Trans-Proteomic Pipeline. *Proteomics* **10**, 1150–1159
38. Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74**, 5383–5392
39. Shteynberg, D., Deutsch, E. W., Lam, H., Aebersold, R., and Nesvizhskii, A. I. (2008) iProphet: improved validation of peptide identification in shotgun proteomics, in *HUPO 7th Annual World Congress, Amsterdam, August 16–20, 2008*, Abstr. P-TUE-181, Human Proteome Organisation, Montreal
40. auf dem Keller, U., Prudova, A., Gioia, M., Butler, G., and Overall, C. (2010) A statistics based platform for quantitative N-terminome analysis and identification of protease cleavage products. *Mol. Cell. Proteomics* **9**, 912–927
41. Sing, T., Sander, O., Beerwinkel, N., and Lengauer, T. (2005) ROCr: visualizing classifier performance in R. *Bioinformatics* **21**, 3940–3941
42. Ramos, H., Shannon, P., and Aebersold, R. (2008) The protein information and property explorer: an easy-to-use, rich-client web application for the management and functional analysis of proteomic data. *Bioinformatics* **24**, 2110–2111
43. Schilling, O., and Overall, C. M. (2008) Proteome-derived, database-searchable peptide libraries for identifying protease cleavage sites. *Nat. Biotechnol.* **26**, 685–694
44. Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J., and Gevaert, K. (2009) Improved visualization of protein consensus sequences by ice-Logo. *Nat. Methods* **6**, 786–787
45. DeSouza, L. V., Taylor, A. M., Li, W., Minkoff, M. S., Romaschin, A. D., Colgan, T. J., and Siu, K. W. (2008) Multiple reaction monitoring of mTRAQ-labeled peptides enables absolute quantification of endogenous levels of a potential cancer marker in cancerous and normal endometrial tissues. *J. Proteome Res.* **7**, 3525–3534
46. Morrison, C. J., Butler, G. S., Rodriguez, D., and Overall, C. M. (2009) Matrix metalloproteinase proteomics: substrates, targets, and therapy. *Curr. Opin. Cell Biol.* **21**, 645–653
47. Cauwe, B., Martens, E., Proost, P., and Opendakker, G. (2009) Multidimensional degradomics identifies systemic autoantigens and intracellular matrix proteins as novel gelatinase B/MMP-9 substrates. *Integr. Biol.* **1**, 404–426
48. Masson, V., de la Ballina, L. R., Munaut, C., Wielockx, B., Jost, M., Maillard, C., Blacher, S., Bajou, K., Itoh, T., Itoharu, S., Werb, Z., Libert, C., Foidart, J. M., and Noël, A. (2005) Contribution of host MMP-2 and MMP-9 to promote tumor vascularization and invasion of malignant keratinocytes. *FASEB J.* **19**, 234–236
49. Xu, D., Suenaga, N., Edelmann, M. J., Fridman, R., Muschel, R. J., and Kessler, B. M. (2008) Novel MMP-9 substrates in cancer cells revealed by a label-free quantitative proteomics approach. *Mol. Cell. Proteomics* **7**, 2215–2228
50. Springman, E. B., Angleton, E. L., Birkedal-Hansen, H., and Van Wart, H. E. (1990) Multiple modes of activation of latent human fibroblast collagenase: evidence for the role of a Cys73 active-site zinc complex in latency and a “cysteine switch” mechanism for activation. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 364–368
51. Butler, G. S., and Overall, C. M. (2009) Proteomic identification of multitasking proteins in unexpected locations complicates drug targeting. *Nat. Rev. Drug Discov.* **8**, 935–948
52. Salatino, M., Croci, D. O., Bianco, G. A., Ilarregui, J. M., Toscano, M. A., and Rabinovich, G. A. (2008) Galectin-1 as a potential therapeutic target in autoimmune disorders and cancer. *Expert Opin. Biol. Ther.* **8**, 45–57
53. Ochieng, J., Fridman, R., Nangia-Makker, P., Kleiner, D. E., Liotta, L. A., Stetler-Stevenson, W. G., and Raz, A. (1994) Galectin-3 is a novel substrate for human matrix metalloproteinases-2 and -9. *Biochemistry* **33**, 14109–14114
54. Ruiz, S., Henschen-Edman, A. H., Nagase, H., and Tenner, A. J. (1999) Digestion of C1q collagen-like domain with MMPs-1,-2,-3, and -9 further defines the sequence involved in the stimulation of neutrophil superoxide production. *J. Leukoc. Biol.* **66**, 416–422
55. Butler, G. S., Sim, D., Tam, E., Devine, D., and Overall, C. M. (2002) Mannose-binding lectin (MBL) mutants are susceptible to matrix metalloproteinase proteolysis: potential role in human MBL deficiency. *J. Biol. Chem.* **277**, 17511–17519
56. Van den Steen, P. E., Van Aelst, I., Hvidberg, V., Piccard, H., Fiten, P., Jacobsen, C., Moestrup, S. K., Fry, S., Royle, L., Wormald, M. R., Wallis, R., Rudd, P. M., Dwek, R. A., and Opendakker, G. (2006) The he-



- mopexin and O-glycosylated domains tune gelatinase B/MMP-9 bio-availability via inhibition and binding to cargo receptors. *J. Biol. Chem.* **281**, 18626–18637
57. Rowsell, S., Hawtin, P., Minshull, C. A., Jepson, H., Brockbank, S. M., Barratt, D. G., Slater, A. M., McPheat, W. L., Waterson, D., Henney, A. M., and Pauptit, R. A. (2002) Crystal structure of human MMP9 in complex with a reverse hydroxamate inhibitor. *J. Mol. Biol.* **319**, 173–181
  58. Hsu, J. L., Huang, S. Y., and Chen, S. H. (2006) Dimethyl multiplexed labeling combined with microcolumn separation and MS analysis for time course study in proteomics. *Electrophoresis* **27**, 3652–3660
  59. Mueller, L. N., Brusniak, M. Y., Mani, D. R., and Aebersold, R. (2008) An assessment of software solutions for the analysis of mass spectrometry based quantitative proteomics data. *J. Proteome Res.* **7**, 51–61
  60. Nesvizhskii, A. I., Vitek, O., and Aebersold, R. (2007) Analysis and validation of proteomic data generated by tandem mass spectrometry. *Nat. Methods* **4**, 787–797
  61. Doucet, A., and Overall, C. M. (2008) Protease proteomics: revealing protease in vivo functions using systems biology approaches. *Mol. Aspects Med.* **29**, 339–358
  62. Egeblad, M., and Werb, Z. (2002) New functions for the matrix metalloproteinases in cancer progression. *Nat. Rev. Cancer* **2**, 161–174
  63. Parks, W. C., Wilson, C. L., and López-Boado, Y. S. (2004) Matrix metalloproteinases as modulators of inflammation and innate immunity. *Nat. Rev. Immunol* **4**, 617–629
  64. McDonald, L., Robertson, D. H., Hurst, J. L., and Beynon, R. J. (2005) Positional proteomics: selective recovery and analysis of N-terminal proteolytic peptides. *Nat. Methods* **2**, 955–957
  65. Plevoda, B., and Sherman, F. (2000) N-alpha-terminal acetylation of eukaryotic proteins. *J. Biol. Chem.* **275**, 36479–36482
  66. Tsuji, T., Nozaki, I., Miyazaki, M., Sakaguchi, M., Pu, H., Hamazaki, Y., Iijima, O., and Namba, M. (2001) Antiproliferative activity of REIC/Dkk-3 and its significant down-regulation in non-small-cell lung carcinomas. *Biochem. Biophys. Res. Commun.* **289**, 257–263
  67. Krupnik, V. E., Sharp, J. D., Jiang, C., Robison, K., Chickering, T. W., Amaravadi, L., Brown, D. E., Guyot, D., Mays, G., Leiby, K., Chang, B., Duong, T., Goodearl, A. D., Gearing, D. P., Sokol, S. Y., and McCarthy, S. A. (1999) Functional and structural diversity of the human Dickkopf gene family. *Gene* **238**, 301–313
  68. Hsieh, S. Y., Hsieh, P. S., Chiu, C. T., and Chen, W. Y. (2004) Dickkopf-3/REIC functions as a suppressor gene of tumor growth. *Oncogene* **23**, 9183–9189
  69. Mizobuchi, Y., Matsuzaki, K., Kuwayama, K., Kitazato, K., Mure, H., Kageji, T., and Nagahiro, S. (2008) REIC/Dkk-3 induces cell death in human malignant glioma. *Neuro-Oncol.* **10**, 244–253
  70. Kobayashi, K., Ouchida, M., Tsuji, T., Hanafusa, H., Miyazaki, M., Namba, M., Shimizu, N., and Shimizu, K. (2002) Reduced expression of the REIC/Dkk-3 gene by promoter-hypermethylation in human tumor cells. *Gene* **282**, 151–158
  71. Mazurek, S. (2007) Pyruvate kinase type M2: a key regulator within the tumour metabolome and a tool for metabolic profiling of tumours. *Ernst Schering Found. Symp. Proc.* **4**, 99–124
  72. Mazurek, S., Drexler, H. C., Troppmair, J., Eigenbrodt, E., and Rapp, U. R. (2007) Regulation of pyruvate kinase type M2 by A-Raf: a possible glycolytic stop or go mechanism. *Anticancer Res.* **27**, 3963–3971
  73. Steták, A., Veress, R., Ovádi, J., Csermely, P., Kéri, G., and Ullrich, A. (2007) Nuclear translocation of the tumor marker pyruvate kinase M2 induces programmed cell death. *Cancer Res.* **67**, 1602–1608
  74. Lee, J., Kim, H. K., Han, Y. M., and Kim, J. (2008) Pyruvate kinase isozyme type M2 (PKM2) interacts and cooperates with Oct-4 in regulating transcription. *Int. J. Biochem. Cell Biol.* **40**, 1043–1054
  75. Shimada, N., Shinagawa, T., and Ishii, S. (2008) Modulation of M2-type pyruvate kinase activity by the cytoplasmic PML tumor suppressor protein. *Genes Cells* **13**, 245–254
  76. Kumar, Y., Tapuria, N., Kirmani, N., and Davidson, B. R. (2007) Tumour M2-pyruvate kinase: a gastrointestinal cancer marker. *Eur. J. Gastroenterol. Hepatol.* **19**, 265–276
  77. Dean, R. A., Butler, G. S., Hamma-Kourbali, Y., Delbé, J., Brigstock, D. R., Courty, J., and Overall, C. M. (2007) Identification of candidate angiogenic inhibitors processed by matrix metalloproteinase 2 (MMP-2) in cell-based proteomic screens: disruption of vascular endothelial growth factor (VEGF)/heparin affinity regulatory peptide (pleiotrophin) and VEGF/Connective tissue growth factor angiogenic inhibitory complexes by MMP-2 proteolysis. *Mol. Cell. Biol.* **27**, 8454–8465
  78. Mañes, S., Mira, E., Barbacid, M. M., Ciprés, A., Fernández-Resca, P., Buesa, J. M., Mérida, I., Aracil, M., Márquez, G., and Martínez-A, C. (1997) Identification of insulin-like growth factor-binding protein-1 as a potential physiological substrate for human stromelysin-3. *J. Biol. Chem.* **272**, 25706–25712
  79. Fowlkes, J. L., Enghild, J. J., Suzuki, K., and Nagase, H. (1994) Matrix metalloproteinases degrade insulin-like growth factor-binding protein-3 in dermal fibroblast cultures. *J. Biol. Chem.* **269**, 25742–25746
  80. Sadowski, T., Dietrich, S., Koschinsky, F., and Sedlacek, R. (2003) Matrix metalloproteinase 19 regulates insulin-like growth factor-mediated proliferation, migration, and adhesion in human keratinocytes through proteolysis of insulin-like growth factor binding protein-3. *Mol. Biol. Cell* **14**, 4569–4580
  81. Coppock, H. A., White, A., Aplin, J. D., and Westwood, M. (2004) Matrix metalloproteinase-3 and -9 proteolyse insulin-like growth factor-binding protein-1. *Biol. Reprod.* **71**, 438–443
  82. Mañes, S., Llorente, M., Lacalle, R. A., Gómez-Moutón, C., Kremer, L., Mira, E., and Martínez-A, C. (1999) The matrix metalloproteinase-9 regulates the insulin-like growth factor-triggered autocrine response in DU-145 carcinoma cells. *J. Biol. Chem.* **274**, 6935–6945
  83. Kim, H. S., Nagalla, S. R., Oh, Y., Wilson, E., Roberts, C. T., Jr., and Rosenfeld, R. G. (1997) Identification of a family of low-affinity insulin-like growth factor binding proteins (IGFBPs): characterization of connective tissue growth factor as a member of the IGFBP superfamily. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12981–12986
  84. Ryan, A. J., Napoletano, S., Fitzpatrick, P. A., Currid, C. A., O'Sullivan, N. C., and Harmey, J. H. (2009) Expression of a protease-resistant insulin-like growth factor-binding protein-4 inhibits tumour growth in a murine model of breast cancer. *Br. J. Cancer* **101**, 278–286
  85. Noh, Y. H., Matsuda, K., Hong, Y. K., Kunstfeld, R., Riccardi, L., Koch, M., Oura, H., Dadras, S. S., Streit, M., and Detmar, M. (2003) An N-terminal 80 kDa recombinant fragment of human thrombospondin-2 inhibits vascular endothelial growth factor induced endothelial cell migration in vitro and tumor growth and angiogenesis in vivo. *J. Invest. Dermatol.* **121**, 1536–1543
  86. Streit, M., Riccardi, L., Velasco, P., Brown, L. F., Hawighorst, T., Bornstein, P., and Detmar, M. (1999) Thrombospondin-2: a potent endogenous inhibitor of tumor growth and angiogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14888–14893
  87. LaBell, T. L., and Byers, P. H. (1993) Sequence and characterization of the complete human thrombospondin 2 cDNA: potential regulatory role for the 3' untranslated region. *Genomics* **17**, 225–229
  88. Xu, X., Chen, Z., Wang, Y., Yamada, Y., and Steffensen, B. (2005) Functional basis for the overlap in ligand interactions and substrate specificities of matrix metalloproteinases-9 and -2. *Biochem. J.* **392**, 127–134
  89. Itoh, T., Ikeda, T., Gomi, H., Nakao, S., Suzuki, T., and Itohara, S. (1997) Unaltered secretion of beta-amyloid precursor protein in gelatinase A (matrix metalloproteinase 2)-deficient mice. *J. Biol. Chem.* **272**, 22389–22392
  90. Martin, M. D., Carter, K. J., Jean-Philippe, S. R., Chang, M., Mobashery, S., Thiollou, S., Lynch, C. C., Matrisian, L. M., and Fingleton, B. (2008) Effect of ablation or inhibition of stromal matrix metalloproteinase-9 on lung metastasis in a breast cancer model is dependent on genetic background. *Cancer Res.* **68**, 6251–6259
  91. Robinson, S. N., Pisarev, V. M., Chavez, J. M., Singh, R. K., and Talmadge, J. E. (2003) Use of matrix metalloproteinase (MMP)-9 knockout mice demonstrates that MMP-9 activity is not absolutely required for G-CSF or Flt-3 ligand-induced hematopoietic progenitor cell mobilization or engraftment. *Stem Cells* **21**, 417–427
  92. Asahi, M., Sumii, T., Fini, M. E., Itohara, S., and Lo, E. H. (2001) Matrix metalloproteinase 2 gene knockout has no effect on acute brain injury after focal ischemia. *Neuroreport* **12**, 3003–3007
  93. Sinnamon, M. J., Carter, K. J., Fingleton, B., and Matrisian, L. M. (2008) Matrix metalloproteinase-9 contributes to intestinal tumorigenesis in the adenomatous polyposis coli multiple intestinal neoplasia mouse. *Int. J. Exp. Pathol.* **89**, 466–475
  94. Kuittinen, O., Soini, Y., and Turpeenniemi-Hujanen, T. (2002) Diverse role of MMP-2 and MMP-9 in the clinicopathological behavior of Hodgkin's

- lymphoma. *Eur. J. Haematol.* **69**, 205–212
95. Bergers, G., Brekken, R., McMahon, G., Vu, T. H., Itoh, T., Tamaki, K., Tanzawa, K., Thorpe, P., Itohara, S., Werb, Z., and Hanahan, D. (2000) Matrix metalloproteinase-9 triggers the angiogenic switch during carcinogenesis. *Nat. Cell Biol.* **2**, 737–744
96. Itoh, T., Tanioka, M., Yoshida, H., Yoshioka, T., Nishimoto, H., and Itohara, S. (1998) Reduced angiogenesis and tumor progression in gelatinase A-deficient mice. *Cancer Res.* **58**, 1048–1051
97. Garg, P., Rojas, M., Ravi, A., Bockbrader, K., Epstein, S., Vijay-Kumar, M., Gewirtz, A. T., Merlin, D., and Sitaraman, S. V. (2006) Selective ablation of matrix metalloproteinase-2 exacerbates experimental colitis: contrasting role of gelatinases in the pathogenesis of colitis. *J. Immunol.* **177**, 4103–4112
98. Longo, G. M., Xiong, W., Greiner, T. C., Zhao, Y., Fiotti, N., and Baxter, B. T. (2002) Matrix metalloproteinases 2 and 9 work in concert to produce aortic aneurysms. *J. Clin. Investig.* **110**, 625–632
99. Itoh, T., Matsuda, H., Tanioka, M., Kuwabara, K., Itohara, S., and Suzuki, R. (2002) The role of matrix metalloproteinase-2 and matrix metalloproteinase-9 in antibody-induced arthritis. *J. Immunol.* **169**, 2643–2647
100. Garg, P., Vijay-Kumar, M., Wang, L., Gewirtz, A. T., Merlin, D., and Sitaraman, S. V. (2009) Matrix metalloproteinase-9-mediated tissue injury overrides the protective effect of matrix metalloproteinase-2 during colitis. *Am. J. Physiol. Gastrointest. Liver Physiol.* **296**, G175–G184
101. Thomas, A. C., and Newby, A. C. (2009) Effect of matrix metalloproteinase-9 knockout on vein graft remodelling in mice. *J. Vasc. Res.* **47**, 299–308
102. Mook, O. R., Frederiks, W. M., and Van Noorden, C. J. (2004) The role of gelatinases in colorectal cancer progression and metastasis. *Biochim. Biophys. Acta* **1705**, 69–89
103. Hahn-Dantona, E., Ramos-DeSimone, N., Siple, J., Nagase, H., French, D. L., and Quigley, J. P. (1999) Activation of proMMP-9 by a plasmin/MMP-3 cascade in a tumor cell model. Regulation by tissue inhibitors of metalloproteinases. *Ann. N.Y. Acad. Sci.* **878**, 372–387
104. Ramos-DeSimone, N., Hahn-Dantona, E., Siple, J., Nagase, H., French, D. L., and Quigley, J. P. (1999) Activation of matrix metalloproteinase-9 (MMP-9) via a converging plasmin/stromelysin-1 cascade enhances tumor cell invasion. *J. Biol. Chem.* **274**, 13066–13076
105. Strongin, A. Y., Collier, I., Bannikov, G., Marmer, B. L., Grant, G. A., and Goldberg, G. I. (1995) Mechanism of cell surface activation of 72-kDa type IV collagenase. Isolation of the activated form of the membrane metalloprotease. *J. Biol. Chem.* **270**, 5331–5338
106. Radisky, D. C., Staffings-Mann, M., Hizai, G., and Bisseff, M. G. (2009) Single proteins might have dual but related functions in intracellular and extracellular microenvironments. *Nat Rev Mol Cell Biol.* **10**, 228–234