Combinatorial degradomics: Precision tools to unveil proteolytic processes in biological systems

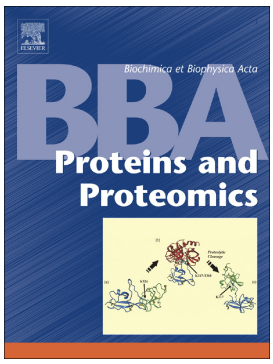
Savickas, Simonas; Kastl, Philipp; auf dem Keller, Ulrich

Published in:
Biochimica et Biophysica Acta - Proteins and Proteomics

Link to article, DOI:
10.1016/j.bbapap.2020.140392

Publication date:
2020

Document Version
Peer reviewed version



Combinatorial degradomics: Precision tools to unveil proteolytic processes in biological systems

Simonas Savickas, Philipp Kastl, Ulrich auf dem Keller

PII: S1570-9639(20)30033-9

DOI: https://doi.org/10.1016/j.bbapap.2020.140392

Reference: BBAPAP 140392

To appear in: *BBA - Proteins and Proteomics*

Received date: 22 November 2019

Revised date: 13 February 2020

Accepted date: 14 February 2020

Please cite this article as: S. Savickas, P. Kastl and U.a.d. Keller, Combinatorial degradomics: Precision tools to unveil proteolytic processes in biological systems, *BBA - Proteins and Proteomics*(2020), https://doi.org/10.1016/j.bbapap.2020.140392

Combinatorial degradomics: precision tools to unveil proteolytic processes in biological systems

Simonas Savickas[*], Philipp Kastl[*], and Ulrich auf dem Keller[#]

Department of Biotechnology and Biomedicine, Technical University of Denmark,
DK-2800 Kongens Lyngby, Denmark

#Address for correspondence:

Prof. Ulrich auf dem Keller
Technical University of Denmark
Department of Biotechnology and Bioengineering
Søltofts Plads, Building 224, Room 114
DK-2800 Kongens Lyngby
Denmark

e-mail: uadk@dtu.dk

Tel.: +45 45 25 27 91

*These authors contributed equally to this work.

Running title: Protease degradomics

Abstract

The biological activity of a protein is regulated at many levels ranging from control of transcription and translation to post-translational modifications (PTM). Proteolytic processing is an irreversible PTM generating novel isoforms of a mature protein termed proteoforms. Proteoform dynamics is a major focus of current proteome research, since it has been associated with many pathological conditions. Mass-spectrometry (MS)-based proteomics and PTM-specific enrichment workflows have become the methods of choice to study proteoforms *in vitro* and *in vivo*. Here, we give an overview of currently available MS-based degradomics methods and outline how they can be optimally applied to study protease cleavage events. We discuss the advantages and disadvantages of selected approaches and describe state-of-the-art improvements in degradomics technologies. By introducing the concept of combinatorial degradomics, a combination of global discovery degradomics and highly sensitive targeted degradomics, we demonstrate how MS-based degradomics further evolves as a powerful tool in biomedical protease research.

Keywords: protease, mass spectrometry, proteomics, degradomics, targeted degradomics, combinatorial degradomics

1. Introduction

Over the past decades, the dogma of molecular biology has dramatically evolved from the simple concept of single gene-transcript-protein relations. It is now clear that one gene can encode many transcriptional products, which are translated into even more amino acid chains making up proteins with multiple biological activities under different conditions. Thus, to understand the function of a protein in a specific biological setting it is important to elucidate its identity, abundance and how generation of different proteoforms is dynamically regulated during its lifetime [1,2]. At the transcriptional level, proteoform dynamics is mediated by transcription factors and alternative splicing, leading to generation of protein isoforms, while at the protein level generation of new proteoforms is conferred by post-translational modifications (PTMs). In recent years the analysis of regulation at the protein level has become a major focus, which resulted in establishment of large initiatives, such as the Human Proteome Organization (HUPO, www.hupo.org) and its associated human protein project (HPP). Regulation of protein activity at the protein level is dynamic, depends on the physiological context and ultimately determines the terminal function of a protein. This knowledge could therefore be used to identify potential novel disease markers and allow for the development of specialized therapeutics that target only one specific proteoform [3].

State-of-the-art system-wide analytical methods like mass spectrometry (MS)-based proteomics have become the preferred approaches to analyze protein functions *in vitro* and *in vivo*. Although proteomics is based on the simple physical concept that by applying an electric field in a high vacuum charged peptides and/or proteins can be separated and identified by their masses, it has become a highly specialized and complex field of research. As an example, MS is used to study different protein proteoforms resulting from PTMs, like phosphorylation, ubiquitination or acetylation [4]. Limited proteolysis of substrates is a very important but often neglected irreversible PTM, which terminally alters protein fate and creates novel proteoforms. Substrates undergoing proteolysis by proteases are often signaling proteins like cytokines, chemokines, other proteases or extracellular matrix components. Those molecules are all involved in regulating development, immune and inflammatory responses, as well as many pathological conditions ranging from cardiovascular to chronic inflammatory diseases or cancer. In order to better understand those processes and identify new biomarkers for the diagnosis of their associated diseases, substrate proteolysis is now the focus of many research groups around the globe. Based on the interest in proteolysis the field of degradomics has been established, aiming at analyzing proteolysis both at the substrate and the protease level by MS-based methods. Over the past

years, multiple specialized degradomics methods have been developed and it is getting more and more challenging to keep an overview of these approaches and to understand how they can be optimally applied to study a specific problem. In the following, we will provide an overview of major degradomics approaches, how currently available MS-based methods can be used in protease research and discuss the future potential of the field.

2. Proteomics and degradomics

As a first step to characterize a proteome and its dynamics, usually a whole proteome analysis of a biological sample of interest is conducted, termed discovery-based proteomics. Most common discovery-based proteomics analyses follow the same, classic bottom-up-proteomics workflow. After isolating the proteome from a biological sample, proteins are digested into peptides using a site-specific endopeptidase (e.g. Trypsin, LysC, ArgC, Lysarginase or GluC). Peptides are then fractionated by reverse-phase chromatography, positively ionized, and analyzed in the mass spectrometer according to their mass to charge (m/z) value. The detected peptides and their abundances can then be related to the original protein by a database search against the whole proteome sequence of the target organism. The constant development of new, highly sensitive, high-resolution mass spectrometers in conjunction with optimized chemical labeling reagents to analyze multiple samples in a single experiment (multiplexing) allows for unprecedented peptide identification and quantification sensitivity below femtomolar ranges. Indeed, as whole proteome analysis of biological samples showed, the abundance of proteins can span a range of more than ten orders of magnitude [5,6]. This is a major problem, because due to the principles of MS analysis peptides derived from high abundance proteins have the tendency to mask low abundance peptides derived from low abundance proteins (Fig. 2A). Moreover, proteolytically generated proteoforms and their peptides are even less abundant, since substrate proteolysis is often incomplete. This can potentially be overcome by prior depletion of high-abundance proteins (e.g. using combinatorial peptide ligand libraries (CPLLs) [7]), as well as affinity purification (e.g. upon biotinylation [8]) and additional sample fractionation (e.g. by UHPLC or affinity chromatography). In the case of *in vitro* experiments, this might be a feasible approach, since access to sample is usually not a problem, but amounts of *in vivo* material is limited and any additional purification step increases the risk of loss of proteins and in particular the lowly abundant proteoforms. In order to solve this challenge and allow for the nearly exclusive identification of protease-generated peptides, specialized positional proteomics methods have been invented. These methods combine multiplexing of samples by isotopic labeling with a specific peptide enrichment step during sample preparation, depleting the sample of internal tryptic

endopeptidase-generated peptides and enriching for only natural as well as protease-generated N-terminal peptides. Over the past years multiple different positional degradomics methods have been presented, with some being project-oriented and rarely been used besides for their original studies. Others have become established discovery degradomics methods used to identify a multitude of protease cleavage events (Fig. 1). Here, we will focus on the two majorly applied methods in detail, but also give a brief overview of some more specialized alternatives.

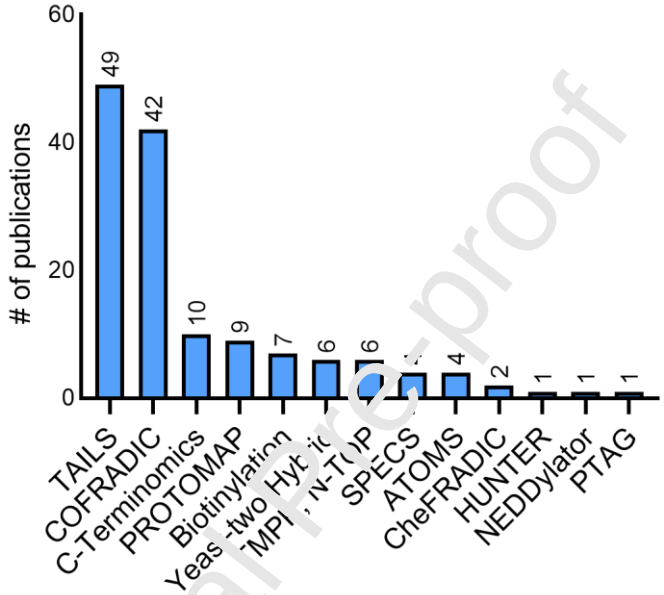


Figure 1: Number of published studies using specified degradomics techniques according to PubMed searches. Data collected in November 2019.

3. Discovery degradomics as tool to identify protease substrates

The first method developed for specific enrichment of protein N-termini from whole proteomes was combined fractional diagonal chromatography (COFRADIC) [9]. In COFRADIC, proteins are reduced and alkylated to block cysteine side chains followed by trideutero-acetylation of primary N-termini (natural and protease-generated) (Fig. 2B, right panel). Afterwards, proteins are trypsin digested, resulting in two species of N-terminal peptides: (i) acetylated original peptides and (ii) tryptic peptides with free α-amines. After the tryptic digest, samples are purified by two orthogonal reverse-phase high performance liquid chromatography (HPLC) steps. The first HPLC step fractionates the sample and reduces its complexity, whereas prior to the second step the C-terminal and tryptic peptides are hydrophobically labeled with 2,4,6-trinitrobenzenesulfonic acid (TNBS). Primary N-termini are then isolated from the

sample due to their altered retention time on the column. A drawback of COFRADIC are limited multiplexing capabilities due to the acetylation of free N-termini and restriction of sample labeling to stable isotope labeling with amino acids in cell culture (SILAC), since other commercial chemical labeling reagents bind to the primary amines of a peptide. A modified version of COFRADIC termed charge-based fractional diagonal chromatography (ChaFRADIC) was developed, which improves its efficiency by employing strong cation exchange chromatography (SCX) prior to the orthogonal HPLC step [10]. This allows for multiplex labeling of native and protease-generated N-termini with amine-reactive chemical reagents, instead of blocking by acetylation. ChaFRADIC has recently been miniaturized, so that the separation can be performed on a self-made column in a pipet tip, reducing sample input to below five µg and making the workflow less dependable on a core facility with expensive, reliable HPLC systems [11].

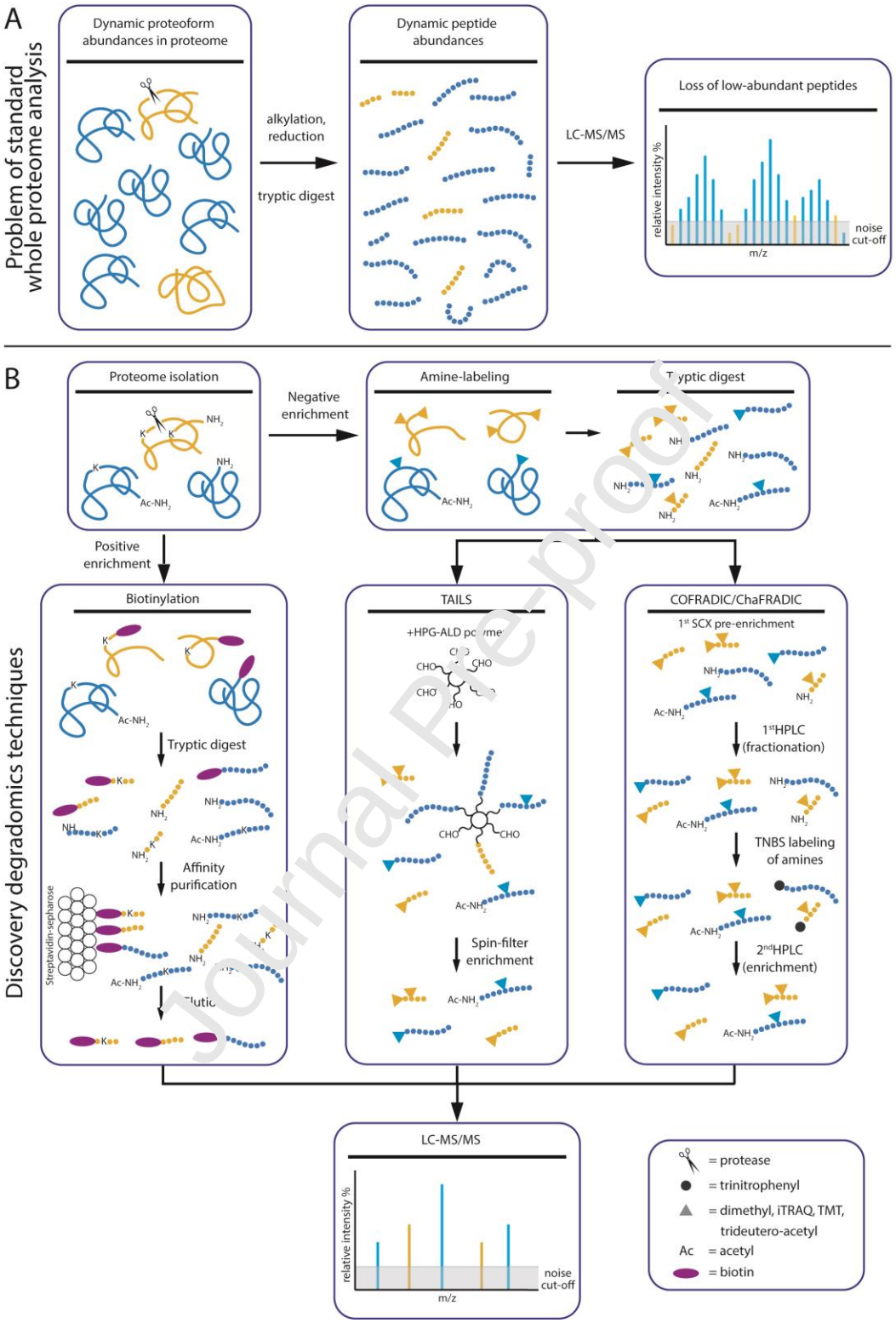


Figure 2: Discovery degradomics allow for identification of low abundance proteoforms in biological samples. (A) Traditional whole proteome shotgun proteomics fails to detect many low abundance proteins due to low signal-to-noise ratios. **(B)** Peptide enrichment methods in discovery degradomics allow for detection of low abundance protease-generated peptides in complex biological samples. Identification of N-terminal peptides can be improved by either positive (left panel) or negative (middle and right panel) enrichment strategies.

Another powerful positional proteomics method widely applied is terminal amine isotopic labeling of substrates (TAILS) ([12]; for a recent protocol see [13]). Like COFRADIC and ChaFRADIC, TAILS is based on negative enrichment of protein N-termini and protease-generated neo-N-terminal peptides (Fig. 2B, middle panel). Here, an isotopic label introduced by whole protein labeling chemically blocks all primary α- and ε-amines. After the tryptic digest, tryptic peptides are removed from the sample by binding to an amine-reactive highly-branched polyglycerol aldehyde polymer (HPG-ALD polymer) via their unblocked primary N-terminal α-amines. Since chemically or natively blocked (e.g. acetylated) N-terminal peptides cannot bind to this polymer, they can be collected in the flow-through of size-exclusion spin-columns. Similar to ChaFRADIC, TAILS allows for multiplexing of samples by isotopic labeling at the protein level prior to the tryptic digest. However, for TAILS this can be extended up to 16-plex by use of amine-reactive tandem mass tags (TMT). Since less chromatography steps are involved, there is a lower risk of peptide losses during sample preparation, but the polymer pull-out depends on prior N-terminal blocking and efficiency of proteome labeling [14]. Multiplexing provides an additional benefit, since increasing the number of samples analyzed within a single TMT-based TAILS experiment results in more total protein, which on the other hand allows to reduce the amount of total protein needed from individual samples.

An alternative to COFRADIC, ChaFRADIC or TAILS is based on removing internal tryptic peptides via phospho-tagging (PTAG) followed by titanium dioxide (TiO_2) affinity chromatography [15]. Additionally, natural und protease-generated N-termini can be positively enriched by N-terminalomics by Chemical Labeling of the α-Amine of Proteins (N-CLAP) [16] or subtiligase enzymatic labeling [8], which both specifically chemically label α-amines allowing for their affinity purification (Fig. 2B, left panel). Secretome Protein Enrichment with Click Sugars (SPECS) metabolically labels glycosylated proteins, followed by click chemistry-mediated biotinylation for avidin affinity enrichment and enables in particular identification of protease-mediated surface shedding events but without providing information on protease cleavage sites [17] [18].

Even though identifying the N-terminal peptide of a cleaved substrate is often sufficient for protease substrate discovery, C-terminal cleavage products are of as much interest. However, enrichment of C-terminal peptides is not as commonly used, since carboxyl groups are less reactive than primary amines. Nevertheless, TAILS and COFRADIC have been adjusted to allow for enrichment of natural and protease-generated protein C-termini. For TAILS, additionally to N-terminal blocking, ethanolamine is coupled to the C-terminus of peptides allowing for the negative enrichment of protein C-termini using a high-

molecular-weight poly-allylamine polymer [19]. C-terminal COFRADIC employs an extra labeling step of primary amines of C-terminal peptides, followed by an additional round of HPLC fractionation [20].

Positive and negative enrichment methods both come with their own advantages and disadvantages and are thus suitable to address different biological questions. Negative enrichment methods allow for in-depth analysis of both the proteome (before enrichment) and the degradome (after enrichment) with the possibility to multiplex experiments by chemical labeling of α- and ε-amines, which increases quantification and identification confidence between multiple samples. Importantly, they also enrich for naturally modified N-/C-terminal peptides (e.g. acetylated N-termini), since those will not bind to the matrix used for negative enrichment. While using readily available chemicals for labeling, negative enrichment methods require either extensive fractionation (COFRADIC, ChaFRADIC) or specialized polymers (TAILS) for depletion of internal tryptic peptides, which, however, are now publicly available. As a major technical challenge, positive enrichment of N-terminal peptides requires selective labeling of N-terminal α-amines. This precludes identification of naturally blocked protein N-termini but might be advantageous for identification of protease cleavage events in cytoplasmic samples with high numbers of naturally acetylated N-terminal peptides. Furthermore, multiplexing and peptide quantification is mostly restricted to SILAC or label-free quantification, since isotopic labels would have to be introduced with N-terminal affinity tags. An exception is subtiligase enzymatic labeling, which allows isotopic labeling of ε-amines after positive enrichment but is restricted to multiplexed quantification of lysine-containing N-terminal peptides and requires access to the subtiligase enzyme.

Since their first publication discovery degradomics methods have been extensively used to identify proteases and their specific cleavage products in various biological backgrounds and from all kinds of different species ranging from *E. Coli* [9,21], *Arabidopsis thaliana* [10], model animals like rodents [22] and pigs [23] to humans [24]. At time of writing, a simple PubMed search with the keywords “Terminal Amine Isotopic Labeling of Substrates” or “COFRADIC” resulted in 49 (for TAILS) and 42 (for COFRADIC) research articles, respectively (Fig. 1). Additional manual curation revealed that those search terms did not even provide a full list of all articles published using these methods. Since discussing all studies that employed a discovery degradomics technique would be beyond the scope of this review, we provide a compilation of research articles published up till November 2019 and based on PubMed searches for the associated method names (Table 1). Reference numbers in this table refer to PubMed accession numbers of corresponding publications listed in Table S1.

Table 1: Compilation of published discovery degradomics methods. Ref. # refers to numbers in Table S1.

#	Technique	Preparation time	Organism	Sample type	Ref. #
49	TAILS	48h	Popular technique: from recombinant proteins to clinical patient testing		1-49
42	COFRADIC	48h	Popular technique: from recombinant proteins to clinical patient testing		50-91
1	HUNTER	48h	Rat	Brain	135
			Plant	*Arabidopsis thaliana*	135
			Human	HeLa, B-ALL, 697, blood plasma, primary B-ALL, primary AML, B-cells, monocytes, natural killer cells, NKT	135
2	CheFRADIC	48h	Plant	*Arabidopsis thaliana*	136
			Human	Platelets	137
7	Enzymatic Biotinylation	60h	Bacteria	*Escherichia coli*	123, 125, 126
			Yeast	*Saccharomyces cerevisiae*	123
			Recombinant Protein	Aprotinin	123,126
			Human	HEK-293A, Serum, Jurkat cells, blood plasma	123, 124, 126, 127
6	Yeast-two-Hybrid in Proteolysis		Human	HEK293T, HCT116, Daudi, THP-1, OCI-AML2, dendritic cells,	129-132
			Recombinant protein	PLXAN1, Fibrinogen	129, 133
			Mouse	Megakaryocytes, platelets	134
4	SPECS	96h	Mouse	Cortical neurons (secretome), MEF (secretome), primary neurons (secretome)	92,94,95
			Human	HEK293T (secretome)	93,94
9	PROTOMAP	24h	Human	Blood, Jurkat T Cells, erythrocytes, MHCC97L, HCCLM6, A549, PC-3	97-104
6	TMPP, N-TOP, dN-TOP	24h	Recombinant protein	Recombinant glycoproteins, Hz6F4-2	105, 106, 107, 110
			Bacterium	*Myobacterium smegmatis, Herminiimonas arsenicoxydans*	108, 109
1	NEDDylator		Human	Jurkat	128
4	ATOMS	48h	Human	Primary macrophages, BJ cells, HUVEC, skin	111, 112
			Recombinant protein	Fibronectin-1, LM-111	113
10	C-Terminomics	48h	Recombinant protein	Bovine serum albumin, β-casein	114
			Human	293T	115
			Bacteria/Yeast	*Saccharomyces cerevisiae, Escherichia coli, Thermoanaerobacter tengcongensis*	116-121
1	PTAG		Bacteria/Yeast	*Neisseria meningitides, Saccharomyces cerevisiae*	122

4. Targeted degradomics to validate and extend results from discovery degradomics approaches

Discovery proteomics and degradomics approaches are inherently challenged by the complexity of the sample proteome. Only applying differential depletion steps or immunoprecipitation methods using specific antibodies might enable identification and quantification of specific low abundance proteins and protease cleavage events. Moreover, inherent undersampling in shotgun proteomics can prevent comprehensive positional peptide mapping of identified substrate proteins. Thus, it might not be clear, if a unique proteoform was generated by limited proteolysis or if it resulted from rather unspecific degradation of the mature protein. By harnessing the sensitivity, precision and ability to overcome interference of the whole proteome, targeted degradomics ensures that proteoforms and cleavage events of interest can be analyzed with high sensitivity and comprehensive coverage of substrate proteins.

The vast amount of information accessible today provides ample resources to identify a proteoform cleavage as well as a candidate protease of interest. Data obtained by discovery degradomics approaches is stored in databases like MEROPS (https://www.ebi.ac.uk/merops/), CutDB (https://omictools.com/cutdb-tool), TopFIND (http://clipserve.clip.ubc.ca/topfind), CaspDB (https://omictools.com/caspdb-tool), Degrabase (https://wellslab.ucsf.edu/degrabase/) and others [25–29]. Those databases may also contain multiple isoforms of a cleaved proteoform associated with different biological contexts [29,30]. Additionally, extracted datasets allow for prediction of candidate proteases based on known substrate cleavage motif preferences of endopeptidases. Validation of direct protease-substrate relations revealed by discovery degradomics require experimental evidence that identified proteoforms are indeed the result of a “true” cleavage event performed by a specific protease. This validation can be achieved by a targeted degradomics experiment. In contrast to discovery degradomics, targeted degradomics aims at monitoring specific proteoforms and their cleavage products rather than measuring as many events as possible. As an example, a substrate of interest is co-incubated with a candidate protease and the mass spectrometer set to specifically scan either the original or the proteolytically generated peptide. Thereby, both proteoforms can be identified and their relative abundances determined [31]. As long as the peptide target is known, targeted degradomics allows identifying the same target peptide e.g. in a simple cell-free protein system, in cell culture and in *in vivo* systems [32]. For instance, targeted proteomics was successfully applied to compare the same peptide target across different cancer cell lines, ranging from pancreatic cancer to colorectal adenosarcoma, breast cancer, and metastatic adenosarcoma [33]. Similarly, it is possible to

monitor the same cleaved peptide in human plasma, urine, liver, kidney, heart, skin, etc. Furthermore, targeted proteomics and degradomics is specifically suited for cross-tissue and cross-laboratory studies and basically only limited by robustness of chromatography and instrument availability [34,35].

In general, trypsin is used as a standard working protease for bottom-up proteomics and degradomics, but with the emergence of alternative endopeptidases, such as LysC, AspN, GluC, Pepsin, and LysargiNase the coverage of the full proteome as well as the N-terminome has significantly increased [31,36]. Like in discovery approaches, different endopeptidases can be applied to targeted proteomics and degradomics to cover the naturally occurring proteome and degradome with maximum depth. Therefore, using the appropriate endopeptidase it is now possible to develop targeted degradomics assays for most protease cleavage events. As long as the peptide of interest has a length between 6 and 26 amino acids, the mass-spectrometer will identify the fragment [36]. Still, due to a limited variety of endopeptidases identification of specific fragments might be prohibited. As an alternative, a top-down proteomics approach might be applied, in which the intact, undigested proteome is directly injected into the LC-MS system. However, top-down-proteomics faces other challenges, such as limited solubility, insignificant ionization and complex fragmentation of proteins [37].

Through PTM crosstalk other PTMs of the parent protein can affect generation of proteolytic fragments. Using targeted degradomics, PTMs such as phosphorylation, glycosylation and ubiquitination that may interfere with proteolytic cleavage can be concomitantly monitored, often without additional enrichment steps [13]. Monitoring the target peptide with and without a modification only requires correction of the target mass and determination of respective elution times. Concomitant analysis of multiple PTMs in an endogenous biological environment allows for a more in-depth analysis of the sample and provides additional insight into complex biological systems [38,39].

In addition to validating peptide cleavage events, targeted degradomics data can be quantitatively analyzed with high accuracy, e.g. to proportionally assess cleaved and non-cleaved proteoforms. This can be achieved by measuring (i) the intensity of the ion chromatographic peak, (ii) counting peptide spectrum matches or (iii) comparing absolute peptide concentration to the synthetic reference spectrum [40]. The most common quantification technique determines the area under the curve of each transition recorded at the peptide fragment level [41,42], which is a flexible and convenient method to select the right values for the identified peptide. These values can then be analyzed using standard data processing software, such as Microsoft Excel, GraphPad Prism, SPSS, or statistical programming

languages like R, Matlab or Python [43,44]. All of these provide convenient packages for targeted degradomics data analysis and statistical methods to assist in data interpretation.

Targeted degradomics is becoming more and more popular as a technique to verify and reliably quantify a cleavage event of choice [31]. This is supported by the advances in sensitivity, precision, speed, and robustness of mass spectrometers that have reached and in some cases already surpassed traditional antibody-based methods [45]. The lack of reliable antibodies, the inability to raise an antibody to a specific cleavage product, or the time it takes to produce an antibody for a specific target protein is a limitation that might be overcome by mass spectrometry-based targeted degradomics.

With better access, affordable prices and availability of proteomics core facilities, mass spectrometry-based proteomics is readily becoming available to a wider community of protease researchers. Initiatives like the Clinical Proteomics Tumor Analysis Consortium (CPTAC, https://www.hupo.org/Clinical-Proteome-Tumor-Analysis-Consortium-(CPTAC)) and the European Proteomics Infrastructure Consortium providing access (EPIC-XS, https://epic-xs.eu/) now also offer degradomics technologies and have opened up many services to biomedical scientists and clinical researchers.

5. Future perspective for combinatorial degradomics technologies

If combined in an approach we term combinatorial degradomics (Fig. 3), discovery and targeted degradomics become a powerful tool to identify, validate and characterize proteolytic events in every type of biological sample.

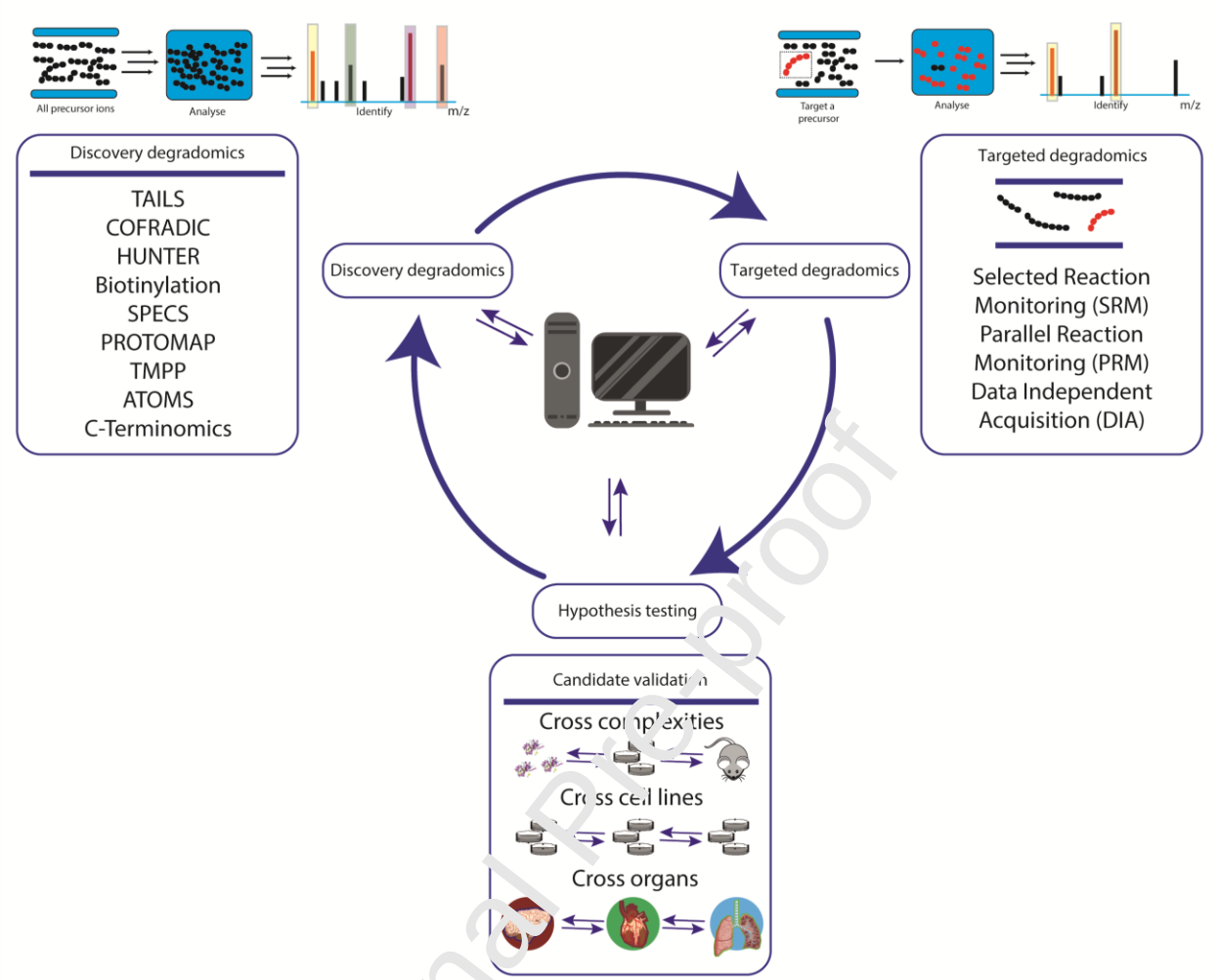


Figure 3: Overview of combinatorial degradomics. Combination of discovery and targeted degradomics methods allows monitoring of protease cleavage events in all kinds of biological samples. Candidate cleavages identified from combinatorial degradomics have to be ultimately verified by *in vitro*/*in vivo* testing across different systems. All three approaches integrate with database profiling and next-generation machine learning algorithms.

Combinatorial degradomics has already been applied to better understand known and to decipher novel protease signaling pathways. Uncovering new modes of complement activation, providing unprecedented insight into inflammation and cancer signaling and unraveling specificity profiles for entire protease families are only a few examples of how degradomics technologies have revolutionized protease research towards a comprehensive understanding of the protease web [46–49]. Although data recorded in system-wide analyses is compiled in data repositories, designing experiments or software algorithms to orthogonally but comprehensively validate the insights gained by degradomics experiments is still a major challenge. Furthermore, more streamlined techniques are needed to confirm biological relevance of phenomena observed in such hypothesis-generating studies.

One of the techniques, which has been applied to discovery degradomics is data independent acquisition (DIA) mass spectrometry, allowing discovery and quantification of thousands of peptides and newly generated proteoforms [50,51]. DIA shifts the paradigm of identifying the most abundant ions in a peptide mixture to measuring specified windows of masses and thus turns the stochastic nature of data dependent acquisition (DDA) shotgun proteomics into a more unbiased approach [52]. Given the fundamental differences in modes of data acquisition, DIA can also serve as orthogonal technology for high-throughput validation of DDA data. In general, DIA relies on extensive spectral libraries that need to be recorded in separate DDA measurements, significantly increasing mass spectrometry time. However, use of pre-generated libraries can now be completely avoided, e.g. by combining DIA with Prosit, a deep learning algorithm for peptide tandem mass spectra prediction [53].

Quantitative discovery degradomics has highly benefitted from extended multiplexing capabilities, now allowing concomitant analysis of up to sixteen samples with the newly released 16plex Tandem Mass Tag (TMT) reagents. By combining sixteen samples into a single injection, the possibility to identify a target cleavage can be increased, the time taken for mass spectrometry analysis reduced and a significantly higher sample throughput achieved. Dependent on the chemistry of isobaric labeling reagents, even higher levels of multiplexing can be realized [54], which are expected to soon be commercially available for proteomics and degradomics applications.

Together with the development of novel amine-reactive labeling reagents, extensive off-line peptide fractionation has the potential to significantly increase throughput and N-terminome coverage in discovery degradomics. By establishing automated set-ups using reverse phase high pH chromatography a sample can be fractionated into more than 48 fractions, dramatically increasing the coverage of the proteome [51,55]. Injecting each fraction independently using a classical DDA mode for analysis and combining all fractions into a single output file, it is possible to identify over 10.000 N-terminal peptides from a single sample. This includes multiple proteoforms that were previously undiscovered or required complex steps of targeted precipitation [45]. Combining fractionation and DIA, the groups of Huesgen and Lange have taken another approach to N-terminal enrichment. Instead of using the HPG-ALD polymer regularly applied in TAILS, they modified the HYTANE method for negative enrichment of N-termini [56]. With this approach, they have managed to identify more than 8000 N-termini out of 1 Million HeLa cells starting material [51].

High-throughput validation of discovery degradomics results does not only depend on the sensitivity and throughput of the validation methods, but also on data processing strategies. As an example, Triggered

by Offset, Multiplexed, Accurate mass, High resolution, and Absolute Quantitation (TOMAHAQ) is a new software tool for targeted proteomics that combines TMT-based multiplexing with targeted analyses. In conjunction with the new TMT 16plex kit, TMT0 and TMT *super heavy*, TOMAHAQ and TOMAHAQ companion allow rapid and reliable target quantification of thousands of targets at a time. Compared to regular targeted proteomics workflows this is a major boost in throughput and may also be used for validation of proteolytic cleavage events identified in discovery degradomics experiments [57–59].

Furthermore, by combining already existing techniques like TAILS and TOMAHAQ it will be possible to use N-terminally enriched fractions as template libraries to monitor and quantify the N-termini of samples without enrichment. This technique will reduce sample loss due to long enrichment workflows, reducing sample-to-sample variation from handling errors and will increase the speed of target verification.

The combination of these new analysis tools pushes combinatorial proteomics to new depths. With the advent of single cell proteomics and new updates on TOMAHAQ it may even become possible to monitor proteolysis events in single cells [60]. Due to an inherent capacity of TMT tags and the TOMAHAQ targeted method to distinguish between a carrier channel with over 500 cells and a single cell carrying a TMT tag this might enable analyzing the heterogeneity of post-translational modifications at the single cell level. This will allow us to draw a comprehensive map of cellular dynamics according to the size of a single cell, its protein content, cell type, etc. The opportunities opened up by single cell proteomics can hardly be grasped, but they will move the field even further towards clinical diagnostics.

Declaration of Conflict of Interests

None.

Acknowledgments

Ulrich auf dem Keller acknowledges support by a Novo Nordisk Foundation Young Investigator Award [NNF16OC0020670]. Philipp Kastl is supported by a research fellowship from the German Research Foundation [project no. 415888450].

References

[1] Y. Liu, A. Beyer, R. Aebersold, On the Dependency of Cellular Protein Levels on mRNA Abundance, Cell. 165 (2016) 535–550. https://doi.org/https://doi.org/10.1016/j.cell.2016.03.014.

[2] L.M. Smith, N.L. Kelleher, T.C. for T.D. Proteomics, M. Linial, D. Goodlett, P. Langridge-Smith, Y. Ah Goo, G. Safford, L. Bonilla*, G. Kruppa, R. Zubarev, J. Rontree, J. Chamot-Rooke, J. Garavelli, A. Heck, J. Loo, D. Penque, M. Hornshaw, C. Hendrickson, L. Paša-Tolić, C. Borchers, D. Chan, N. Young*, J. Agar, C. Masselon, M. Gross*, F. McLafferty, Y. Tsybin, Y. Ge, I. Sanders*, J. Langridge, J. Whitelegge*, A. Marshall, Proteoform: a single term describing protein complexity, Nat. Methods. 10 (2013) 186. https://doi.org/10.1038/nmeth.2369.

[3] R. Aebersold, G.D. Bader, A.M. Edwards, J.E. van Eyk, M. Kussmann, J. Qin, G.S. Omenn, The Biology/Disease-driven Human Proteome Project (B/D-HPP): Enabling Protein Research for the Life Sciences Community, J. Proteome Res. 12 (2013) 23–27. https://doi.org/10.1021/pr301151m.

[4] E.S. Witze, W.M. Old, K.A. Resing, N.G. Ahn, Mapping protein post-translational modifications with mass spectrometry, Nat. Methods. 4 (2007) 798–806. https://doi.org/10.1038/nmeth1100.

[5] A.F. Hühmer, R.G. Biringer, H. Amato, A.N. Fonteh, M.G. Harrington, Protein analysis in human cerebrospinal fluid: Physiological aspects, current progress and future challenges, Dis. Markers. 22 (2006) 3–26. https://doi.org/10.1155/2006/158797.

[6] M. Wang, C.J. Herrmann, M. Simonovic, D. Szklarczyk, C. von Mering, Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines, Proteomics. 15 (2015) 3163–3168. https://doi.org/10.1002/pmic.201400441.

[7] E. Boschetti, P.G. Righetti, The ProteoMiner in the proteomic arena: A non-depleting tool for discovering low-abundance species, J. Proteomics. 71 (2008) 255–264. https://doi.org/https://doi.org/10.1016/j.jprot.2008.05.002.

[8] H.A.I. Yoshihara, S. Mahrus, J.A. Wells, Tags for labeling protein N-termini with subtiligase for proteomics, Bioorg. Med. Chem. Lett. 18 (2008) 6000–6003. https://doi.org/10.1016/j.bmcl.2008.08.044.

[9] K. Gevaert, M. Goethals, L. Martens, J. Van Damme, A. Staes, G.R. Thomas, J. Vandekerckhove, Exploring proteomes and analyzing protein processing by mass spectrometric identification of sorted N-terminal peptides, Nat. Biotechnol. 21 (2003) 566–569. https://doi.org/10.1038/nbt810.

[10] A.S. Venne, F.A. Solari, F. Faden, T. Paretti, N. Dissmeyer, R.P. Zahedi, An improved workflow for quantitative N-terminal charge-based fractional diagonal chromatography (ChaFRADIC) to study proteolytic events in Arabidopsis thaliana, Proteomics. 15 (2015) 2458–2469.

https://doi.org/10.1002/pmic.201500014.

[11] G. Shema, M.T.N. Nguyen, F.A. Solari, S. Loroch, A.S. Venne, L. Kollipara, A. Sickmann, S.H.L. Verhelst, R.P. Zahedi, Simple, scalable, and ultrasensitive tip-based identification of protease substrates, Mol. &Amp; Cell. Proteomics. 17 (2018) 826 LP – 834. https://doi.org/10.1074/mcp.TIR117.000302.

[12] O. Kleifeld, A. Doucet, A. Prudova, U. auf dem Keller, M. Gioia, J.N. Kizhakkedathu, C.M. Overall, Identifying and quantifying proteolytic events and the natural N terminome by terminal amine isotopic labeling of substrates, Nat. Protoc. 6 (2011) 1578–1611. https://doi.org/10.1038/nprot.2011.382.

[13] E. Madzharova, F. Sabino, U. auf dem Keller, Exploring Extracellular Matrix Degradomes by TMT-TAILS N-Terminomics BT - Collagen: Methods and Protocols, in: I. Sagi, N.A. Afratis (Eds.), Springer New York, New York, NY, 2019: pp. 115–126. https://doi.org/10.1007/978-1-4939-9095-5_8.

[14] V. Guryča, J. Lamerz, A. Ducret, P. Cutler, Qualitative improvement and quantitative assessment of N-terminomics, Proteomics. 12 (2012) 1207–1216. https://doi.org/10.1002/pmic.201100430.

[15] G.P.M. Mommen, B. van de Waterbeemd, H.D. Meiring, C. Kersten, A.J.R. Heck, A.P.J.M. de Jong, Unbiased Selective Isolation of Protein N-terminal Peptides from Complex Proteome Samples Using Phospho Tagging (PTAG) and TiO₂-based Depletion, Mol. &Amp; Cell. Proteomics. 11 (2012) 832 LP - 842. https://doi.org/10.1074/mcp.O112.018283.

[16] G. Xu, S.B.Y. Shin, S.R. Jaffrey, Global profiling of protease cleavage sites by chemoselective labeling of protein N-termini, Proc. Natl. Acad. Sci. 106 (2009) 19310 LP – 19315. https://doi.org/10.1073/pnas.0908358106.

[17] P.-H. Kuhn, K. Koroniak, S. Hogl, A. Colombo, U. Zeitschel, M. Willem, C. Volbracht, U. Schepers, A. Imhof, A. Hoffmeister, C. Haass, S. Roßner, S. Bräse, S.F. Lichtenthaler, Secretome protein enrichment identifies physiological BACE1 protease substrates in neurons, EMBO J. 31 (2012) 3157–3168. https://doi.org/10.1038/emboj.2012.173.

[18] A. Serdaroglu, S.A. Müller, U. Schepers, S. Bräse, W. Weichert, S.F. Lichtenthaler, P.-H. Kuhn, An optimised version of the secretome protein enrichment with click sugars (SPECS) method leads to enhanced coverage of the secretome, Proteomics. 17 (2017) 1600423. https://doi.org/10.1002/pmic.201600423.

[19] O. Schilling, P.F. Huesgen, O. Barré, C.M. Overall, Identification and Relative Quantification of Native and Proteolytically Generated Protein C-Termini from Complex Proteomes: C-Terminome Analysis BT - Network Biology: Methods and Applications, in: G. Cagney, A. Emili (Eds.), Humana Press, Totowa, NJ, 2011: pp. 59–69. https://doi.org/10.1007/978-1-61779-276-2_4.

[20] P. Van Damme, A. Staes, S. Bronsoms, K. Helsens, N. Colaert, E. Timmerman, F.X. Aviles, J. Vandekerckhove, K. Gevaert, Complementary positional proteomics for screening substrates of endo- and exoproteases, Nat. Methods. 7 (2010) 512–515. https://doi.org/10.1038/nmeth.1469.

[21] K. Gevaert, J. Van Damme, M. Goethals, G.R. Thomas, B. Hoorelbeke, H. Demol, L. Martens, M. Puype, A. Staes, J. Vandekerckhove, Chromatographic Isolation of Methionine-containing Peptides for Gel-free Proteome Analysis, Mol. &Amp; Cell. Proteomics. 1 (2002) 896 LP –

903. https://doi.org/10.1074/mcp.M200061-MCP200.

[22] P. Schlage, T. Kockmann, F. Sabino, J.N. Kizhakkedathu, U. auf dem Keller, Matrix Metalloproteinase 10 Degradomics in Keratinocytes and Epidermal Tissue Identifies Bioactive Substrates With Pleiotropic Functions, Mol. Cell. Proteomics . 14 (2015) 3234–3246. https://doi.org/10.1074/mcp.M115.053520.

[23] F. Sabino, O. Hermes, F.E. Egli, T. Kockmann, P. Schlage, P. Croizat, J.N. Kizhakkedathu, H. Smola, U. auf dem Keller, In Vivo Assessment of Protease Dynamics in Cutaneous Wound Healing by Degradomics Analysis of Porcine Wound Exudates, Mol. Cell. Proteomics . 14 (2015) 354–370. https://doi.org/10.1074/mcp.M114.043414.

[24] T. Klein, S.-Y. Fung, F. Renner, M.A. Blank, A. Dufour, S. Kang, M. Bolger-Munro, J.M. Scurll, J.J. Priatel, P. Schweigler, S. Melkko, M.R. Gold, R.I. Viner, C.H. Régnier, S.E. Turvey, C.M. Overall, The paracaspase MALT1 cleaves HOIL1 reducing linear ubiquitination by LUBAC to dampen lymphocyte NF-κB signalling, Nat. Commun. 6 (2015) 8777. https://doi.org/10.1038/ncomms9777.

[25] E.D. Crawford, J.E. Seaman, N. Agard, G.W. Hsu, O. Julien, S. Mahrus, H. Nguyen, K. Shimbo, H.A.I. Yoshihara, M. Zhuang, R.J. Chalkley, J.A. Wells, The DegraBase: A Database of Proteolysis in Healthy and Apoptotic Human Cells, Mol. &Amp; Cell. Proteomics. 12 (2013) 813 LP – 824. https://doi.org/10.1074/mcp.O112.024372.

[26] N. Fortelny, S. Yang, P. Pavlidis, P.F. Lange, C.M. Overall, Proteome TopFIND 3.0 with TopFINDer and PathFINDer: database and analysis tools for the association of protein termini to pre- and post-translational events, Nucleic Acids Res. 43 (2014) D290–D297. https://doi.org/10.1093/nar/gku1012.

[27] Y. Igarashi, A. Eroshkin, S. Gramatikova, K. Gramatikoff, Y. Zhang, J.W. Smith, A.L. Osterman, A. Godzik, CutDB: a proteolytic event database, Nucleic Acids Res. 35 (2006) D546–D549. https://doi.org/10.1093/nar/gkl813.

[28] S. Kumar, B.J. van Raam, G.S. Salvesen, P. Cieplak, Caspase Cleavage Sites in the Human Proteome: CaspDB, a Database of Predicted Substrates, PLoS One. 9 (2014) e110539. https://doi.org/10.1371/journal.pone.0110539.

[29] N.D. Rawlings, A.J. Barrett, P.D. Thomas, X. Huang, A. Bateman, R.D. Finn, The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database, Nucleic Acids Res. (2018). https://doi.org/10.1093/nar/gkx1134.

[30] N. Fortelny, G.S. Butler, C.M. Overall, P. Pavlidis, Protease-inhibitor interaction predictions: Lessons on the complexity of protein-Protein interactions, Mol. Cell. Proteomics. (2017). https://doi.org/10.1074/mcp.M116.065706.

[31] S. Savickas, U. Auf Dem Keller, Targeted degradomics in protein terminomics and protease substrate discovery, Biol. Chem. (2017). https://doi.org/10.1515/hsz-2017-0187.

[32] H.A. Ebhardt, A. Root, C. Sander, R. Aebersold, Applications of targeted proteomics in systems biology and translational medicine, Proteomics. (2015). https://doi.org/10.1002/pmic.201500004.

[33] M. Tachikawa, Y. Kaneko, S. Ohtsuki, Y. Uchida, M. Watanabe, H. Ohtsuka, T. Terasaki, Targeted proteomics-based quantitative protein atlas of pannexin and connexin subtypes in mouse and

human tissues and cancer cell lines, J. Pharm. Sci. (2019). https://doi.org/10.1016/j.xphs.2019.09.024.

[34] T.A. Addona, S.E. Abbatiello, B. Schilling, S.J. Skates, D.R. Mani, D.M. Bunk, C.H. Spiegelman, L.J. Zimmerman, A.J.L. Ham, H. Keshishian, S.C. Hall, S. Allen, R.K. Blackman, C. H. Borchers, C. Buck, H.L. Cardasis, M.P. Cusack, N.G. Dodder, B.W. Gibson, J.M. Held, T. Hiltke, A. Jackson, E.B. Johansen, C.R. Kinsinger, J. Li, M. Mesri, T.A. Neubert, R.K. Niles, T.C. Pulsipher, D. Ransohoff, H. Rodriguez, P.A. Rudnick, D. Smith, D.L. Tabb, T.J. Tegeler, A.M. Variyath, L.J. Vega-Montoto, Å. Wahlander, S. Waldemarson, M. Wang, J.R. Whiteaker, L. Zhao, N.L. Anderson, S.J. Fisher, D.C. Liebler, A.G. Paulovich, F.E. Regnier, P. Tempst, S.A. Carr, Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma, Nat. Biotechnol. (2009). https://doi.org/10.1038/nbt.1546.

[35] H.D. Cox, F. Lopes, G.A. Woldemariam, J.O. Becker, M.C. Parkin, A. Thomas, A.W. Butch, D.A. Cowan, M. Thevis, L.D. Bowers, A.N. Hoofnagle, Interlaboratory agreement of insulin-like growth factor 1 concentrations measured by mass spectrometry, Clin. Chem. (2014). https://doi.org/10.1373/clinchem.2013.208538.

[36] P. Giansanti, L. Tsiatsiani, T.Y. Low, A.J.R. Heck, Six alternative proteases for mass spectrometry-based proteomics beyond trypsin, Nat. Protoc. (2016). https://doi.org/10.1038/nprot.2016.057.

[37] A.D. Catherman, O.S. Skinner, N.L. Kelleher, Top Down proteomics: Facts and perspectives, Biochem. Biophys. Res. Commun. 445 (2014) 683–693. https://doi.org/https://doi.org/10.1016/j.bbrc.2014.02.041.

[38] S.L. King, C.K. Goth, U. Eckhard, H.J. Joshi, A.D. Haue, S.Y. Vakhrushev, K.T. Schjoldager, C.M. Overall, H.H. Wandall, TAILS N-terminomics and proteomics reveal complex regulation of proteolytic cleavage by O-glycosylation, J. Biol. Chem. (2018). https://doi.org/10.1074/jbc.RA118.001978.

[39] D.B. Trentini, M.J. Suskiewicz, A. Heuck, R. Kurzbauer, L. Deszcz, K. Mechtler, T. Clausen, Arginine phosphorylation marks proteins for degradation by a Clp protease, Nature. (2016). https://doi.org/10.1038/nature20122.

[40] J.W.H. Wong, G. Cagney, An overview of label-free quantitation methods in proteomics by mass spectrometry., Methods Mol. Biol. (2010). https://doi.org/10.1007/978-1-60761-444-9_18.

[41] A. Bourmaud, S. Gallien, B. Domon, Parallel reaction monitoring using quadrupole-Orbitrap mass spectrometer: Principle and applications, Proteomics. (2016). https://doi.org/10.1002/pmic.201500543.

[42] B. MacLean, D.M. Tomazela, N. Shulman, M. Chambers, G.L. Finney, B. Frewen, R. Kern, D.L. Tabb, D.C. Liebler, M.J. MacCoss, Skyline: An open source document editor for creating and analyzing targeted proteomics experiments, Bioinformatics. (2010). https://doi.org/10.1093/bioinformatics/btq054.

[43] M. Choi, C.Y. Chang, T. Clough, D. Broudy, T. Killeen, B. MacLean, O. Vitek, MSstats: An R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments, Bioinformatics. (2014). https://doi.org/10.1093/bioinformatics/btu305.

[44] C. Galitzine, J.D. Egertson, S. Abbatiello, C.M. Henderson, L.K. Pino, M. MacCoss, A.N. Hoofnagle, O. Vitek, Nonlinear regression improves accuracy of characterization of multiplexed mass

spectrometric assays, Mol. Cell. Proteomics. (2018). https://doi.org/10.1074/mcp.RA117.000322.

[45] S.M. Henry, E. Sutlief, O. Salas-Solano, J. Valliere-Douglass, ELISA reagent coverage evaluation by affinity purification tandem mass spectrometry, MAbs. (2017). https://doi.org/10.1080/19420862.2017.1349586.

[46] U. Auf Dem Keller, A. Prudova, U. Eckhard, B. Fingleton, C.M. Overall, Systems-level analysis of proteolytic events in increased vascular permeability and complement activation in skin inflammation, Sci. Signal. (2013). https://doi.org/10.1126/scisignal.2003512.

[47] U. Eckhard, P.F. Huesgen, O. Schilling, C.L. Bellac, G.S. Butler, J.H. Cox, A. Dufour, V. Goebeler, R. Kappelhoff, U. auf dem Keller, T. Klein, P.F. Lange, G. Marino, C.J. Morrison, A. Prudova, D. Rodriguez, A.E. Starr, Y. Wang, C.M. Overall, Active site specificity profiling of the matrix metalloproteinase family: Proteomic identification of 4300 cleavage sites by nine MMPs explored with structural and synthetic peptide cleavage analyses, Matrix Biol. 49 (2016) 37–60. https://doi.org/https://doi.org/10.1016/j.matbio.2015.09.003.

[48] N. Fortelny, J.H. Cox, R. Kappelhoff, A.E. Starr, P.F. Lange, I. Pavlidis, C.M. Overall, Network Analyses Reveal Pervasive Functional Regulation Between Proteases in the Human Protease Web, PLOS Biol. 12 (2014) e1001869. https://doi.org/10.1371/journal.pbio.1001869.

[49] T. Klein, U. Eckhard, A. Dufour, N. Solis, C.M. Overall, Proteolytic Cleavage - Mechanisms, Function, and "omic" Approaches for a Near-Ubiquitous Posttranslational Modification, Chem. Rev. (2018). https://doi.org/10.1021/acs.chemrev.7b00120.

[50] M. Grozdanić, R. Vidmar, M. Vizovišek, M. Fonović, Degradomics in Biomarker Discovery, PROTEOMICS – Clin. Appl. (2019). https://doi.org/10.1002/prca.201800138.

[51] S.S.H. Weng, F. Demir, E.K. Ergin, S. Dirnberger, A. Uzozie, D. Tuscher, L. Nierves, J. Tsui, P.F. Huesgen, P.F. Lange, Sensitive determination of proteolytic proteoforms in limited microscale proteome samples, Mol. Cell. Proteomics. (2019). https://doi.org/10.1074/mcp.tir119.001560.

[52] R. Bruderer, O.M. Bernhardt, T. Gandhi, Y. Xuan, J. Sondermann, M. Schmidt, D. Gomez-Varela, L. Reiter, Optimization of experimental parameters in data-independent mass spectrometry significantly increases depth and reproducibility of results, Mol. Cell. Proteomics. (2017). https://doi.org/10.1074/mcp.RA117.000314.

[53] S. Gessulat, T. Schmidt, D.P. Zolg, P. Samaras, K. Schnatbaum, J. Zerweck, T. Knaute, J. Rechenberger, B. Delanghe, A. Huhmer, U. Reimer, H.C. Ehrlich, S. Aiche, B. Kuster, M. Wilhelm, Prosit: proteome-wide prediction of peptide tandem mass spectra by deep learning, Nat. Methods. (2019). https://doi.org/10.1038/s41592-019-0426-7.

[54] R. Bąchor, M. Waliczek, P. Stefanowicz, Z. Szewczuk, Trends in the Design of New Isobaric Labeling Reagents for Quantitative Proteomics, Mol. . 24 (2019). https://doi.org/10.3390/molecules24040701.

[55] T.S. Batth, J. V Olsen, Offline high pH reversed-phase peptide fractionation for deep phosphoproteome coverage, in: Methods Mol. Biol., 2016. https://doi.org/10.1007/978-1-4939-3049-4_12.

[56] L. Chen, Y. Shan, Y. Weng, Z. Sui, X. Zhang, Z. Liang, L. Zhang, Y. Zhang, Hydrophobic Tagging-Assisted N-Termini Enrichment for In-Depth N-Terminome Analysis, Anal. Chem. 88 (2016) 8390–8395. https://doi.org/10.1021/acs.analchem.6b02453.

[57] B.K. Erickson, C.M. Rose, C.R. Braun, A.R. Erickson, J. Knott, G.C. McAlister, M. Wühr, J.A. Paulo, R.A. Everley, S.P. Gygi, A Strategy to Combine Sample Multiplexing with Targeted Proteomics Assays for High-Throughput Protein Signature Characterization, Mol. Cell. (2017). https://doi.org/10.1016/j.molcel.2016.12.005.

[58] C.M. Rose, B.K. Erickson, D.K. Schweppe, R. Viner, J. Choi, J. Rogers, R. Bomgarden, S.P. Gygi, D.S. Kirkpatrick, TomahaqCompanion: A Tool for the Creation and Analysis of Isobaric Label Based Multiplexed Targeted Assays, J. Proteome Res. (2019). https://doi.org/10.1021/acs.jproteome.8b00767.

[59] X. Zhong, Q. Yu, F. Ma, D.C. Frost, L. Lu, Z. Chen, H. Zetterberg, C. Carlsson, O. Okonkwo, L. Li, HOTMAQ: A Multiplexed Absolute Quantification Method for Targeted Proteomics, Anal. Chem. (2019). https://doi.org/10.1021/acs.analchem.8b04580.

[60] B. Budnik, E. Levy, G. Harmange, N. Slavov, SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation, Genome Biol. (2018). https://doi.org/10.1186/s13059-018-1547-5.

Table S1: Publication number and corresponding pubmedID as referred to from table 1 and figure 1.

TAILS	PMID
1	31629960
2	31463921
3	31601107
4	31337208
5	31462935
6	31393699
7	31606086
8	30926607
9	30705123
10	30267877
11	30840238
12	30257879
13	30371095
14	30284925
15	30377279
16	29593093
17	29437971
18	29358340
19	29035686
20	29318539
21	28730820
22	28082348
23	28315245
24	28299737
25	28109429
26	2747726[illegible]2
27	2708954
28	2674052
29	26587561
30	26475864
31	26371235
32	26061547
33	26258467
34	25871442
35	25516628
36	25331112
37	25310974
38	23811845
39	24281761
40	23519473
41	22940918
42	23667905
43	22367194
44	22577022
45	21972973
46	21959240
47	21693781
48	21646356
49	21604129
COFRADIC	PMID
50	29358340
51	2831524[illegible]
52	2831524[illegible]
53	2[illegible]933[illegible]8
54	2[illegible]153[illegible]30
55	25877158
56	25381060
57	25383893
58	24816145
59	24536027
60	24505135
61	24974297
62	24567107
63	23620545
64	23705892
65	23429522
66	22825847
67	22559236
68	22313033
69	22065552
70	21790200
71	21619078
72	21799483
73	21750686
74	21953040
75	21070068
76	19911852
77	19660099
78	19420222
79	19065607
80	19544022
81	18402783
82	18318009
83	18021241
84	17918875
85	17444671
86	17461103
87	17266347
88	17891751
89	16179924
90	16038019
91	12665801
SPECS	PMID
92	29716987
93	27991726
94	25827571
95	22728825
PROTOMAP	PMID
96	28315249
97	27708557
98	27451424
99	27378148
100	24967658
101	22623320
102	20943600
103	21584940
104	18724940
TMPP	PMID
105	26481993
106	22749915
107	22695497
108	18955433
109	23641718
110	25769014
ATOMS	PMID
111	31324718
112	29946113
113	20876890
114	22078539

C-Terminomics	PMID
115	20526347
116	31634432
117	31195796
118	27071760
119	26361894
120	21877277
121	26356223
122	21812462
123	27363206
124	28315248
PTAG	PMID
125	22729381
Biotinylation	PMID
126	16299481
127	17650073
128	21604127
129	18722006
130	24974296
131	29155430
132	19767749
NEDDylator	PMID
133	23201124
Yeast-2-Hybrid	PMID
134	29362422
135	27471094
136	22898924
137	2054690[illegible]
138	162[illegible]
139	1531596[illegible]
HUNTER	PMID
140	31471496
CheFRADIC	PMID
141	26010716
142	27535140

CRediT author statement

Simonas Savickas: Writing – Original Draft **Philipp Kastl:** Writing – Original Draft **Ulrich auf dem Keller:** Writing – Review & Editing

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Highlights

- Whole proteome mass spectrometry fails to detect low-abundant proteolytic peptides
- Degradomics techniques enrich and detect low-abundant proteoforms
- Global discovery degradomics allows for protease-generated biomarker identification
- Highly sensitive targeted degradomics validates identified candidates
- Combinatorial degradomics further evolves MS towards clinical application