Targeted degradomics in protein terminomics and protease substrate discovery

Savickas, Simonas; auf dem Keller, Ulrich

Published in:
Biological Chemistry

Link to article, DOI:
10.1515/hsz-2017-0187

Publication date:
2017

Document Version
Peer reviewed version

Biological Chemistry ‘Just Accepted’ Papers

Biological Chemistry ‘Just Accepted’ Papers are papers published online, in advance of appearing in the print journal. They have been peer-reviewed, accepted and are online published in manuscript form, but have not been copy edited, typeset, or proofread. Copy editing may lead to small differences between the Just Accepted version and the final version. There may also be differences in the quality of the graphics. When papers do appear in print, they will be removed from this feature and grouped with other papers in an issue.

Biol Chem ‘Just Accepted’ Papers are citable; the online publication date is indicated on the Table of Contents page, and the article’s Digital Object Identifier (DOI), a unique identifier for intellectual property in the digital environment (e.g., 10.1515/hsz-2011-xxxx), is shown at the top margin of the title page. Once an article is published as **Biol Chem ‘Just Accepted’ Paper** (and before it is published in its final form), it should be cited in other articles by indicating author list, title and DOI.

After a paper is published in **Biol Chem ‘Just Accepted’ Paper** form, it proceeds through the normal production process, which includes copy editing, typesetting and proofreading. The edited paper is then published in its final form in a regular print and online issue of **Biol Chem.** At this time, the **Biol Chem ‘Just Accepted’ Paper** version is replaced on the journal Web site by the final version of the paper with the same DOI as the **Biol Chem ‘Just Accepted’ Paper version.**

Disclaimer

Biol Chem ‘Just Accepted’ Papers have undergone the complete peer-review process. However, none of the additional editorial preparation, which includes copy editing, typesetting and proofreading, has been performed. Therefore, there may be errors in articles published as **Biol Chem ‘Just Accepted’ Papers** that will be corrected in the final print and online version of the Journal. Any use of these articles is subject to the explicit understanding that the papers have not yet gone through the full quality control process prior to advanced publication.

Biological Chemistry **'Just Accepted' paper**
ISSN (online) 1437-4315
DOI: 10.1515/hsz-2017-0187

Minireview

Targeted degradomics in protein terminomics and protease substrate discovery

Simonas Savickas and Ulrich auf dem Keller*

Institute of Molecular Health Sciences, ETH Zurich, Otto-Stern-Weg 7, CH-8093 Zurich, Switzerland

*Corresponding author

e-mail: ulrich.aufdemkeller@biol.ethz.ch

Abstract

Targeted degradomics integrates positional information into mass spectrometry-based targeted proteomics workflows and thereby enables analysis of proteolytic cleavage events with unprecedented specificity and sensitivity. Rapid progress in establishment of protease-substrate relations provides extensive degradomics target lists that now can be tested with help of selected and parallel reaction monitoring (S/PRM) in complex biological systems, where proteases act in physiological environments. In this minireview, we describe the general principles of targeted degradomics, outline the generic experimental workflow of the methodology and highlight recent and future applications in protease research.

Keywords: degradomics; parallel reaction monitoring; protease substrates; selected reaction monitoring, targeted proteomics.

Introduction

Proteolysis is a protease-induced irreversible post translational modification, resulting in hydrolysis of peptide or isopeptide bonds with distinct consequences for target proteins. Thereby, substrates might be either eliminated by proteolytic degradation or functionally modified through limited proteolysis by specific removal of a defined subsequence of amino acids. Cellular homeostasis depends on tightly regulated proteolytic activity, which is spatially and temporally maintained by cofactors, inhibitors and other surrounding elements. Dysregulated proteolysis may lead to fatal pathologies, such as impaired tissue repair, inflammation, neurodegenerative disorders or cancer (Drag and Salvesen, 2010; Turk *et al.*, 2012). Latest databases contain information about 2700 proteases of various model organisms with more than 1000 in queue still to be sequenced and characterized (Rawlings *et al.*, 2016). Major protease groups can be clustered on the basis of their catalytic mechanisms into aspartic, glutamic, cysteine, metallo, serine and threonine endopeptidases (Turk *et al.*, 2012). Multiple strategies, such as substrate- and activity-based probes and natural substrate turnover characterized the mechanisms of these proteases, and more than 17.000 physiological substrates were assigned to their related cleaving enzymes (auf dem Keller and Schilling, 2010; Rawlings *et al.*, 2016).

Rapid advancements in mass spectrometry (MS)-based proteomics has helped to detect protease substrates and active proteases in the context of their biological roles and significantly increased the number of known protein targets (Marino *et al.*, 2015; Schlage and auf dem Keller, 2015). However, although these global approaches accumulated also an array of potential biomarkers, their application in the clinics is lagging behind (Anderson *et al.*, 2013). Techniques complimentary to shotgun proteomics, such as selected reaction monitoring (SRM), selected ion monitoring (SIM) and parallel reaction monitoring (PRM) may bridge target discovery, context dependent activity and clinical diagnosis. Methodologies first described by Agard *et al.* (2012) and Fahlman *et al.* (2014) have utilized previously collected biochemical data of proteolytic cleavages and applied SRM and targeted analysis to monitor substrate turnover in complex biological samples with attomolar sensitivity and high specificity. They have indicated that targeted degradomics, i.e. the specific proteomics-based analysis of selected proteolytic cleavage events, has the quantitative strength, unbiased reproducibility, and systematic simplicity known from general targeted proteomics, which has been widely applied to monitor relative protein abundances. Thus, this pioneering work can

have a wide spectrum of applications from clinical assay improvement to elucidation of complex proteolytic networks. Prominent examples for proteolytic fragments, which have been already tested in the clinics as biomarkers are peptides released from the amyloid beta 4 protein (APP) in Alzheimer's disease or fibrinogen degradation products in breast and colon cancer (Huesgen *et al.*, 2014).

In this minireview, we introduce targeted degradomics as a methodology, which is complementary to shotgun MS-based degradomics approaches. Particularly, we focus on design and implementation of targeted degradomics experiments and recent applications of PRM and SRM in the context of proteolytic processing. Moreover, we outline the power of these technologies as independent functional discovery tools in addition to providing enhanced sensitivity in validation of degradomics screens. Finally, we discuss the potential of targeted degradomics methods in elucidating complex interplays between proteolysis and other post-translational modifications (PTMs) exploiting their high specificity and sensitivity.

Concept of targeted degradomics

By specifically monitoring preselected peptides that have been generated from proteins in a bottom-up approach in most cases by tryptic digest, targeted proteomics methods allow assessing and quantifying the same set of related proteins reproducibly in complex proteomes (Rost *et al.*, 2015). For technical details of these powerful approaches, the reader is referred to excellent recent review articles published by pioneers in the field (Bourmaud *et al.*, 2016; Ebhardt *et al.*, 2015; Picotti and Aebersold, 2012). Targeted degradomics further extends this concept by selectively including semi-tryptic peptides, which are released from proteins that had been proteolytically processed. Concomitant analysis of the corresponding fully tryptic peptide (cleavage site spanning peptide) from the non-processed form of the target protein allows drawing conclusions with regard to degree of processing, while recording of 'classical' tryptic peptides from other regions of the protein relates differential processing to general changes in protein abundances in proteomes across multiple conditions. In particular, this enables quantifying degrees of maturation of proteins whose function is dependent on proteolytic removal of modulatory propeptides. Classical examples are protease zymogens, e.g. members of the blood coagulation cascade, which are activated by proteolytic processing. Hence, by integrating positional information into selection of target peptides, targeted degradomics quantitatively monitors specific proteolytic cleavage events and their dynamics

in complex proteomes with high throughput and unprecedented precision. Ultimately, this can be exploited to test for dynamic activities of individual proteases or net outcomes of interconnected protease networks through targeted analysis of related indicative proteolytic events.

Generation of degradomics target lists

Key to targeted degradomics is the availability of information on proteolytic cleavage events, which are translated into libraries of unique proteolytically processed peptides. Due to extensive biomedical research in the last decades and increasing numbers of high-throughput studies, a wealth of data has been deposited into custom protease-centered databases, such as MEROPS, TopFIND, CutDB, CaspDB and Degrabase (Crawford *et al.*, 2013; Fortelny *et al.*, 2015; Igarashi *et al.*, 2007; Kumar *et al.*, 2014; Rawlings *et al.*, 2016). For specific target proteins of interest or to assess activities of particular proteases, mining of these databases can be complemented by extensive literature searches to include more recently described putative cleavage events, which have not yet been included in common repositories. Even if no published data are available, indicative cleavage events for many proteases might be obtained with help of emerging protease predictor algorithms, which achieve high levels of accuracy, particularly for proteases whose cleavage specificity is dominated by selective amino acid residues (Song *et al.*, 2011; Soste *et al.*, 2014). Such substrate and cleavage site predictors are available for several groups of proteases (PROSPER: https://prosper.erc.monash.edu.au/home.html, PeptideCutter: http://web.expasy.org/peptide_cutter/, Cascleave: http://www.structbioinfor.org/cascleave2/index.html, SitePrediction: http://www.dmbr.ugent.be/prx/bioit2-public/SitePrediction/, ProP: http://www.cbs.dtu.dk/services/ProP/, PoPS: http://pops.csse.monash.edu.au) and combine sequence information, structural properties and additional biophysical and biological features to assign predictive values to theoretical cleavages in target proteins. As an additional custom resource, cleavage data might be derived from own shotgun discovery degradomics experiments, applying methodologies that have been extensively reviewed elsewhere (Rogers and Overall, 2013; Schlage and auf dem Keller, 2015).

Since cleavage events collected from public repositories or generated by predictive algorithms are not necessarily derived from MS-based proteomics studies, they mostly need to be translated into suitable peptide lists for targeted degradomics analysis. This is achieved by

in silico digestion of related cleaved and non-cleaved substrate proteins using sequence specific endopeptidases (Fahlman *et al.*, 2014; Giansanti *et al.*, 2016). As mentioned above, trypsin is the most widely used endopeptidase in bottom-up proteomics to cleave amino acid sequences carboxy-terminal to arginine and lysine residues. However, this specificity might be problematic for targeted degradomics experiments, since it overlaps with many important trypsin-like endopeptidases, complicating distinction between cleavages of the test protease and the proteomics working protease used in bottom-up approaches. Moreover, the high frequency of lysine and arginine in protein sequences might often lead to very short (<6 amino acids) semi-tryptic peptides, which are not accessible to identification by MS-based proteomics and thus cannot be reliably monitored in targeted experiments. Therefore, alternative endopeptidases, such as Lys-C, Glu-C, Asp-N or pepsin should be considered that have been successfully employed in bottom-up proteomics (Giansanti *et al.*, 2016). With LysargiNase, a very interesting novel endopeptidase has been introduced that is particularly interesting for targeted degradomics, since it generates peptides with N-terminal lysine or arginine and thus alleviates monitoring of C-terminal halves of cleaved cleavage site spanning peptides (Huesgen *et al.*, 2015). The generated list of truncated semi-specific peptide fragments should be complemented with intact fully specific peptides per protein ideally N- and C-terminal to respective cleavage sites. These peptides benchmark the general abundance of protease target proteins in analyzed proteomes and allow evaluating proteolytic processing as differential event, e.g. in response to cellular stimuli.

In contrast to classical targeted proteomics, for which extensive repositories of proteotypic peptides with associated MS spectra and in-line liquid chromatography (LC) retention times for design of optimal peptide target lists are available (Kusebauch *et al.*, 2016), targeted degradomics generally requires physical validation of selected targets with help of synthetic peptides that are measured under standardized LC-MS conditions. Finally, introducing heavy isotopes in form of ^{13}C- or ^{15}N-amino acids provides the researcher with a mix of validated peptides that can be spiked into analytes for most specific monitoring of proteolytic events in complex biological matrices.

Targeted degradomics experimental workflow

A validated target list of unique peptides allows directly querying complex proteomes for neo-N-termini, neo-C-termini and cleavage site spanning peptides with high specificity and

sensitivity. In a generic targeted degradomics workflow (Figure 1), proteomes could be derived from any biological source material of different levels of complexity that had been exposed to differential protease activity. This might have happened through endogenous proteases whose basal activity is modulated by external stimuli or ablated in loss-of-function models, or upon incubation of the source proteome with selected exogenously added proteolytic modifiers. Moreover, substrate peptides of known cleavage events might be added to bioactive proteomes to specifically test for activities of endogenous proteases (Dutta *et al.*, 2016). The resulting substrate degradomes are further digested with respective working endopeptidases to generate peptides of same length and properties as in the provided target list.

In a next step, prepared protein digests are analyzed by LC-MS/MS in SRM or PRM mode, whereby the mass spectrometer is programmed to specifically monitor peptides provided in the unique degradomics target list. Targeted degradomics analyses use reverse phase (RP) chromatography as classical in-line LC component. Normalization of LC retention times is achieved with help of standard peptides, e.g. iRT peptides (Biognosys, Switzerland), which are measured together with target peptides when preparing the target list and spiked into the experimental protein digest prior to LC-MS/MS analysis. The combination of defined masses and retention times allows narrowing in on short time windows in so called time-scheduled runs, increasing the number of simultaneously monitored targets with high sensitivity. Depending on nature and complexity of analyzed proteomes and substrate degradomes, additional dimensions of LC might be added either off- or in-line with RP to enhance chances of precisely identifying and quantifying proteolytic end products of low abundance (Di Palma *et al.*, 2012).

Targeted degradomics assays have been mostly implemented on triple-quadrupole mass spectrometers (Agard *et al.*, 2012; Dutta *et al.*, 2016; Julien *et al.*, 2016; Sabino *et al.*, 2015; Shimbo *et al.*, 2012; Soste *et al.*, 2014). Triple-quadrupole instrument measurements rely on both precursor and fragment isolation for targeted analysis, also known as SRM. Small mass windows for quadrupole filters ensure reproducible and sensitive selection of target peptides and their fragment ions. However, the more recently introduced PRM methodology performed on quadrupole-Orbitrap instruments isolates only the precursor but measures all of its possible fragment ions (Bourmaud *et al.*, 2016). It has been observed that in some conditions quantification of peptides with PRM allows higher resolution and higher sensitivity than SRM (Kockmann *et al.*, 2016). The improved methodological accuracy enables a more stringent

quantification of proteolytically generated and modified peptides (neo-N and neo-C-terminus, cleavage site spanning peptide). Thus, lower limits of detection might be reached to help evaluating properties of low abundant proteases or proteases with low activity in a complex proteome.

For design and interpretation of targeted degradomics experiments, software packages, such as Analyst, Skyline, and SpectroDive are applied, which are commonly used in targeted proteomics (MacLean *et al.*, 2010; Poli *et al.*, 2015). The software recognizes, annotates and quantifies peptides specified in unique degradomics target lists. Recovered peptide intensity values can be further subjected to stringent statistical normalization as well as ad hoc and post hoc testing using statistical downstream processing packages like MSstats (Choi *et al.*, 2014).

To increase sensitivity and coverage, specific subspecies of cleavage events, e.g. neo-termini, might be selectively enriched with help of appropriate positional proteomics approaches like terminal amine isotopic labeling of substrates (TAILS), combined fractional diagonal chromatography (COFRADIC) or subtiligase treatment (Schlage and auf dem Keller, 2015). Thereby, proteomes are chemically labeled on the protein level after differential exposure to proteases but before digestion with working endopeptidases. Notably, since application of any of these techniques is associated with chemical modification of target peptides, degradomics target lists have to be adjusted accordingly and any synthetic template peptides subjected to the same treatment. Furthermore, selective enrichment of terminal peptides prohibits analysis of e.g. tryptic peptides of the monitored protease substrate of interest from the same peptide mixture and thus requires separate analyses of samples prior to the enrichment step. Finally, several additional strategies can be applied to further enhance coverage and improve quantification of proteins and protease substrates of interest (Eichelbaum *et al.*, 2012).

Recent applications of targeted degradomics

In recent years, the power of targeted degradomics has been exploited to tackle several important questions in protease research. The flexibility in study design and highly reproducible identification of the same peptide species in multiple samples by SRM is particularly suited to simultaneously monitor multiple proteolytic events over time and thereby determine kinetic parameters for proteases in complex systems. This was first demonstrated by Agard *et al.* (2012) who determined hundreds of catalytic efficiencies of

caspase-dependent cleavage events in cells both upon incubation of lysates with recombinant proteases and upon activation of endogenous caspases in cells by apoptosis. This study applied subtiligase-tagging to enrich for N-terminal peptides in combination with SRM for quantification, a strategy that was also successful in determining substrate preferences of caspase-2 and caspase-6 by quantitative kinetics (Julien *et al.*, 2016).

Biological fluids and specifically blood plasma present enormous challenges to stochastic sampling in shotgun proteomics due to their complexity and extreme dynamic range of protein concentrations (Hu *et al.*, 2006). Consequently, analysis of proteolytic end products in body fluids highly benefits from selectivity and specificity of targeted degradomics approaches. As a powerful alternative to immunoblotting, sensitive SRM assays have been established to monitor proteolytically generated forms of cardiac troponin T in patient serum that are associated with severity of cardiac damage and state of a patient prior and after developing acute myocardial infarction (Streng *et al.*, 2016). To test for activity of a specific protease – aparaginyl endopeptidase – in blood plasma, Dutta *et al.* (2016) developed a cleavable synthetic peptide whose substrate products they monitored by SRM upon incubation with human plasma. Since asparaginyl endopeptidase activity has been implicated in diseases, such as breast cancer, leukemia and dementia, this assay might serve as diagnostic biomarker test in future applications. A study by Wiita *et al.* (2014) monitored proteolytic signatures in serum from cancer patients in response to chemotherapy induced cell death. By utilizing the power of subtiligase assisted N-terminal enrichment and targeted degradomics, they identified and quantified processing events including many caspase cleavage products that significantly increased in amount in postchemotherapy plasma and might lead to a novel class of biomarkers to monitor response to chemotherapy. As another example, Sabino *et al.* (2015) used multiplexed iTRAQ-based TAILS to record proteolytic fragments released into wound fluids at multiple time points after wounding in a pig vacuum therapy model. Among them, a cleavage fragment of the integrin adapter protein kindlin-3 was detected, which could be assigned to caspase-3 activity and readily validated by SRM in samples from an independent experiment.

It should be noted that so far, most studies applied targeted degradomics to validate and specifically monitor protease cleavage products that had been identified after enrichment of terminal peptides. However, it can be expected that the importance of termini enrichment will decrease with increasing performance of mass spectrometers. As an example, the Kindlin-3

neo-N terminus monitored by SRM in pig wound exudates by Sabino *et al.* (2015) was identified both before and after TAILS enrichment of protein N-termini.

Future directions

The unique capabilities of targeted degradomics approaches are particularly suited to address current major challenges in understanding the complex functions of proteases in biological systems. For instance, targeted degradomics allows deciphering interconnected protease activation networks and their relations by monitoring zymogen propeptide removal as crucial regulatory event in protease activation. Starting with the famous waterfall cascade of blood coagulation, interdependent zymogen activation networks have been described for many protease groups including caspases, cathepsins, kallikreins and matrix metalloproteinases that all together with their substrates and inhibitors form the protease web (Fortelny *et al.*, 2014). With help of targeted analyses, it will be possible to develop quantitative models of protease activation networks by concomitant monitoring of decrease in abundance of spanning peptides and increase in abundance of corresponding neo-termini at sites of proteolytic removal of inhibitory propeptides (Fahlman *et al.*, 2014) (Figure 2). Respective abundance ratios will determine activation potentials of individual network nodes and ultimately delineate proteolytic potentials of modules and subnetworks and their disturbance by gain- or loss-of-function in model systems of increasing levels of complexity. Thereby, the high specificity and sensitivity of targeted degradomics will enable direct transfer of results from test tube and cell-based assays to analysis of biological material from animal models and patients suffering from diseases associated with detrimental aberrant protease activities. Ultimately, studies of this kind have the potential to convert current limited attempts of interfering with single protease-substrate relations into powerful strategies of attacking disturbances in protease networks as underlying causes of many pathologies.

Another area of research that is predestined to strongly benefit from targeted degradomics is the analysis of interplay between proteolysis and other post-translational modifications in form of PTM crosstalk. As a prime example of such complicated interactions, proteolytic processing of fibroblast growth factor 23 is regulated by reciprocal O-glycosylation and phosphorylation of a critical serine residue in P1’ position of a furin-type cleavage site (Tagliabracci *et al.*, 2014). Moreover, Dix *et al.* (2012) demonstrated the functional interplay between phosphorylation and proteolysis during apoptosis, and TAILS has been applied to

phosphorylated and non-phosphorylated proteomes to identify phosphorylation-dependent cleavages of target proteins by caspases (Turowec *et al.*, 2014). Similarly, glycosylation of either the cleaving protease or the substrate protein can strongly affect cleavage kinetics and modulate specificities (Goettig, 2016). Several powerful enrichment techniques have been developed and applied to study single PTMs, but stochastic sampling in shotgun approaches complicates analysis of the same peptide in different PTM-modified forms. This might be overcome by targeted proteomics and enable efficient characterization of PTM-dependent proteolytic cleavage events when combined with targeted degradomics methods.

In addition to a strong contribution of targeted degradomics to solving important issues in protease biology, it can be expected that rapid advancements in targeted proteomics technologies will steadily increase the power of the concept. Currently, targeted proteomics assays are typically limited to simultaneous analysis of a few hundred peptides per LC-MS/MS run with high resolution. Recently, Gallien *et al.* (2015) introduced internal standard triggered-parallel reaction monitoring (IS-PRM), which they showed to significantly increase the number of measurable peptides in a single LC-MS/MS run by using synthetic isotopically labeled peptides to trigger scheduled high resolution measurements. As a very powerful emerging technology, data-independent acquisition (DIA) combines advantages in coverage of shotgun proteomics for discovery-based applications with sensitivity and reproducible monitoring of selected peptides by SRM or PRM. This is achieved by selecting all instead of only the most intense precursor ions for fragmentation and further analysis on time-of-flight or Orbitrap analyzers. Here, a particular challenge is the deconvolution of multiplexed fragment spectra, a computationally intensive task that is more and more alleviated by introduction of novel software packages for data interpretation (Navarro *et al.*, 2016). Advanced sample preparation strategies and selective enrichment of protein termini together with DIA could be a key workflow to characterize proteolytic events with highest coverage and precision in increasingly complex biological matrices.

Conclusions

Targeted degradomics is a powerful concept to precisely monitor proteolytic events and their modulation under changing conditions in complex biological systems with high specificity and sensitivity. With the growing knowledge of specific protease activities and the improvements in high-throughput protease substrate discovery, more and more

comprehensive degradomics target lists can be compiled to analyze protease action and function with extensive coverage and at increasing levels of complexity. Numerous software packages are under active development that will further leverage data interpretation and exchange of results within the protease research community. Several studies have already exploited the power of targeted degradomics to define kinetic parameters of hundreds of proteolytic cleavages in parallel and to assess protease activities in blood and body fluids, which pose particular challenges to MS-based proteomics. Targeted proteomics opens up many new avenues of research, e.g. to explore interconnected protease activation networks under physiological conditions and to study the complex interplay between proteolysis and other major PTMs. Rapid developments in MS-based proteomics technologies will further push the boundaries of throughput, sensitivity and specificity to get closer to the ultimate aim of understanding the entire protease web and its perturbations in health and disease.

Acknowledgements

We thank S. Werner (ETH Zurich, Switzerland) for her continuous support of our work and the Swiss National Science Foundation for funding with help of a project grant (31003A_163216).

References

Agard, N.J., Mahrus, S., Trinidad, J.C., Lynn, A., Burlingame, A.L., and Wells, J.A. (2012). Global kinetic analysis of proteolysis via quantitative targeted proteomics. Proc. Natl. Acad. Sci. USA *109*, 1913-1918.

Anderson, N.L., Ptolemy, A.S., and Rifai, N. (2013). The riddle of protein diagnostics: future bleak or bright? Clin. Chem. *59*, 194-197.

auf dem Keller, U. and Schilling, O. (2010). Proteomic techniques and activity-based probes for the system-wide study of proteolysis. Biochimie *92*, 1705-1714.

Bourmaud, A., Gallien, S., and Domon, B. (2016). Parallel reaction monitoring using quadrupole-Orbitrap mass spectrometer: principle and applications. Proteomics *16*, 2146-2159.

Choi, M., Chang, C.Y., Clough, T., Broudy, D., Killeen, T., MacLean, B., and Vitek, O. (2014). MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. Bioinformatics *30*, 2524-2526.

Crawford, E.D., Seaman, J.E., Agard, N., Hsu, G.W., Julien, O., Mahrus, S., Nguyen, H., Shimbo, K., Yoshihara, H.A., Zhuang, M., *et al.* (2013). The DegraBase: a database of proteolysis in healthy and apoptotic human cells. Mol. Cell. Proteomics *12*, 813-824.

Di Palma, S., Hennrich, M.L., Heck, A.J., and Mohammed, S. (2012). Recent advances in peptide separation by multidimensional liquid chromatography for proteome analysis. J. Proteomics *75*, 3791-3813.

Dix, M.M., Simon, G.M., Wang, C., Okerberg, E., Patricelli, M.P., and Cravatt, B.F. (2012). Functional interplay between caspase cleavage and phosphorylation sculpts the apoptotic proteome. Cell *150*, 426-440.

Drag, M. and Salvesen, G.S. (2010). Emerging principles in protease-based drug discovery. Nat. Rev. Drug Discov. *9*, 690-701.

Dutta, A., Potier, D.N., Walker, M.J., Gray, O.J., Parker, C., Holland, M., Williamson, A.J., Pierce, A., Unwin, R.D., Krishnan, S., *et al.* (2016). Development of a selected reaction monitoring mass spectrometry-based assay to detect asparaginyl endopeptidase activity in biological fluids. Oncotarget *7*, 70822-70831.

Ebhardt, H.A., Root, A., Sander, C., and Aebersold, R. (2015). Applications of targeted proteomics in systems biology and translational medicine. Proteomics *15*, 3193-3208.

Eichelbaum, K., Winter, M., Berriel Diaz, M., Herzig, S., and Krijgsveld, J. (2012). Selective enrichment of newly synthesized proteins for quantitative secretome analysis. Nat. Biotechnol. *30*, 984-990.

Fahlman, R.P., Chen, W., and Overall, C.M. (2014). Absolute proteomic quantification of the activity state of proteases and proteolytic cleavages using proteolytic signature peptides and isobaric tags. J. Proteomics *100*, 79-91.

Fortelny, N., Cox, J.H., Kappelhoff, R., Starr, A.E., Lange, P.F., Pavlidis, P., and Overall, C.M. (2014). Network analyses reveal pervasive functional regulation between proteases in the human protease web. PLoS Biol. *12*, e1001869.

Fortelny, N., Yang, S., Pavlidis, P., Lange, P.F., and Overall, C.M. (2015). 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*, D290-297.

Gallien, S., Kim, S.Y., and Domon, B. (2015). Large-scale targeted proteomics using internal standard triggered-parallel reaction monitoring (IS-PRM). Mol. Cell. Proteomics *14*, 1630-1644.

Giansanti, P., Tsiatsiani, L., Low, T.Y., and Heck, A.J. (2016). Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. Nat. Protoc. *11*, 993-1006.

Goettig, P. (2016). Effects of glycosylation on the enzymatic activity and mechanisms of proteases. Int. J. Mol. Sci. *17*.

Hu, S., Loo, J.A., and Wong, D.T. (2006). Human body fluid proteome analysis. Proteomics *6*, 6326-6353.

Huesgen, P.F., Lange, P.F., and Overall, C.M. (2014). Ensembles of protein termini and specific proteolytic signatures as candidate biomarkers of disease. Proteomics Clin Appl *8*, 338-350.

Huesgen, P.F., Lange, P.F., Rogers, L.D., Solis, N., Eckhard, U., Kleifeld, O., Goulas, T., Gomis-Ruth, F.X., and Overall, C.M. (2015). LysargiNase mirrors trypsin for protein C-terminal and methylation-site identification. Nat. Methods *12*, 55-58.

Igarashi, Y., Eroshkin, A., Gramatikova, S., Gramatikoff, K., Zhang, Y., Smith, J.W., Osterman, A.L., and Godzik, A. (2007). CutDB: a proteolytic event database. Nucleic Acids Res. *35*, D546-549.

Julien, O., Zhuang, M., Wiita, A.P., O'Donoghue, A.J., Knudsen, G.M., Craik, C.S., and Wells, J.A. (2016). Quantitative MS-based enzymology of caspases reveals distinct protein substrate specificities, hierarchies, and cellular roles. Proc. Natl. Acad. Sci. USA *113*, E2001-2010.

Kockmann, T., Trachsel, C., Panse, C., Wahlander, A., Selevsek, N., Grossmann, J., Wolski, W.E., and Schlapbach, R. (2016). Targeted proteomics coming of age - SRM, PRM and DIA performance evaluated from a core facility perspective. Proteomics *16*, 2183-2192.

Kumar, S., van Raam, B.J., Salvesen, G.S., and Cieplak, P. (2014). Caspase cleavage sites in the human proteome: CaspDB, a database of predicted substrates. PLoS One *9*, e110539.

Kusebauch, U., Campbell, D.S., Deutsch, E.W., Chu, C.S., Spicer, D.A., Brusniak, M.Y., Slagel, J., Sun, Z., Stevens, J., Grimes, B., *et al.* (2016). Human SRMAtlas: a resource of targeted assays to quantify the complete human proteome. Cell *166*, 766-778.

MacLean, B., Tomazela, D.M., Shulman, N., Chambers, M., Finney, G.L., Frewen, B., Kern, R., Tabb, D.L., Liebler, D.C., and MacCoss, M.J. (2010). Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics *26*, 966-968.

Marino, G., Eckhard, U., and Overall, C.M. (2015). Protein termini and their modifications revealed by positional proteomics. ACS Chem. Biol. *10*, 1754-1764.

Navarro, P., Kuharev, J., Gillet, L.C., Bernhardt, O.M., MacLean, B., Rost, H.L., Tate, S.A., Tsou, C.C., Reiter, L., Distler, U., *et al.* (2016). A multicenter study benchmarks software tools for label-free proteome quantification. Nat. Biotechnol. *34*, 1130-1136.

Picotti, P. and Aebersold, R. (2012). Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. Nat. Methods *9*, 555-566.

Poli, M., Ori, A., Child, T., Jaroudi, S., Spath, K., Beck, M., and Wells, D. (2015). Characterization and quantification of proteins secreted by single human embryos prior to implantation. EMBO Mol. Med. *7*, 1465-1479.

Rawlings, N.D., Barrett, A.J., and Finn, R. (2016). Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res. *44*, D343-350.

Rogers, L.D. and Overall, C.M. (2013). Proteolytic post-translational modification of proteins: proteomic tools and methodology. Mol. Cell. Proteomics *12*, 3532-3542.

Rost, H.L., Malmstrom, L., and Aebersold, R. (2015). Reproducible quantitative proteotype data matrices for systems biology. Mol. Biol. Cell *26*, 3926-3931.

Sabino, F., Hermes, O., Egli, F.E., Kockmann, T., Schlage, P., Croizat, P., Kizhakkedathu, J.N., Smola, H., and auf dem Keller, U. (2015). *In vivo* assessment of protease dynamics in cutaneous wound healing by degradomics analysis of porcine wound exudates. Mol. Cell. Proteomics *14*, 354-370.

Schlage, P. and auf dem Keller, U. (2015). Proteomic approaches to uncover MMP function. Matrix Biol. *44-46C*, 232-238.

Shimbo, K., Hsu, G.W., Nguyen, H., Mahrus, S., Trinidad, J.C., Burlingame, A.L., and Wells, J.A. (2012). Quantitative profiling of caspase-cleaved substrates reveals different drug-induced and cell-type patterns in apoptosis. Proc. Natl. Acad. Sci. USA *109*, 12432-12437.

Song, J., Tan, H., Boyd, S.E., Shen, H., Mahmood, K., Webb, G.I., Akutsu, T., Whisstock, J.C., and Pike, R.N. (2011). Bioinformatic approaches for predicting substrates of proteases. J. Bioinform. Comput. Biol. *9*, 149-178.

Soste, M., Hrabakova, R., Wanka, S., Melnik, A., Boersema, P., Maiolica, A., Wernas, T., Tognetti, M., von Mering, C., and Picotti, P. (2014). A sentinel protein assay for simultaneously quantifying cellular processes. Nat. Methods *11*, 1045-1048.

Streng, A.S., de Boer, D., Bouwman, F.G., Mariman, E.C., Scholten, A., van Dieijen-Visser, M.P., and Wodzig, W.K. (2016). Development of a targeted selected ion monitoring assay for the elucidation of protease induced structural changes in cardiac troponin T. J. Proteomics *136*, 123-132.

Tagliabracci, V.S., Engel, J.L., Wiley, S.E., Xiao, J., Gonzalez, D.J., Nidumanda Appaiah, H., Koller, A., Nizet, V., White, K.E., and Dixon, J.E. (2014). Dynamic regulation of FGF23 by Fam20C phosphorylation, GalNAc-T3 glycosylation, and furin proteolysis. Proc. Natl. Acad. Sci. USA *111*, 5520-5525.

Turk, B., Turk, D.S.A., and Turk, V. (2012). Protease signalling: the cutting edge. EMBO J. *31*, 1630-1643.

Turowec, J.P., Zukowski, S.A., Knight, J.D., Smalley, D.M., Graves, L.M., Johnson, G.L., Li, S.S., Lajoie, G.A., and Litchfield, D.W. (2014). An unbiased proteomic screen reveals caspase cleavage is positively and negatively regulated by substrate phosphorylation. Mol. Cell. Proteomics *13*, 1184-1197.

Wiita, A.P., Hsu, G.W., Lu, C.M., Esensten, J.H., and Wells, J.A. (2014). Circulating proteolytic signatures of chemotherapy-induced cell death in humans discovered by N-terminal labeling. Proc. Natl. Acad. Sci. USA *111*, 7594-7599.

Figures

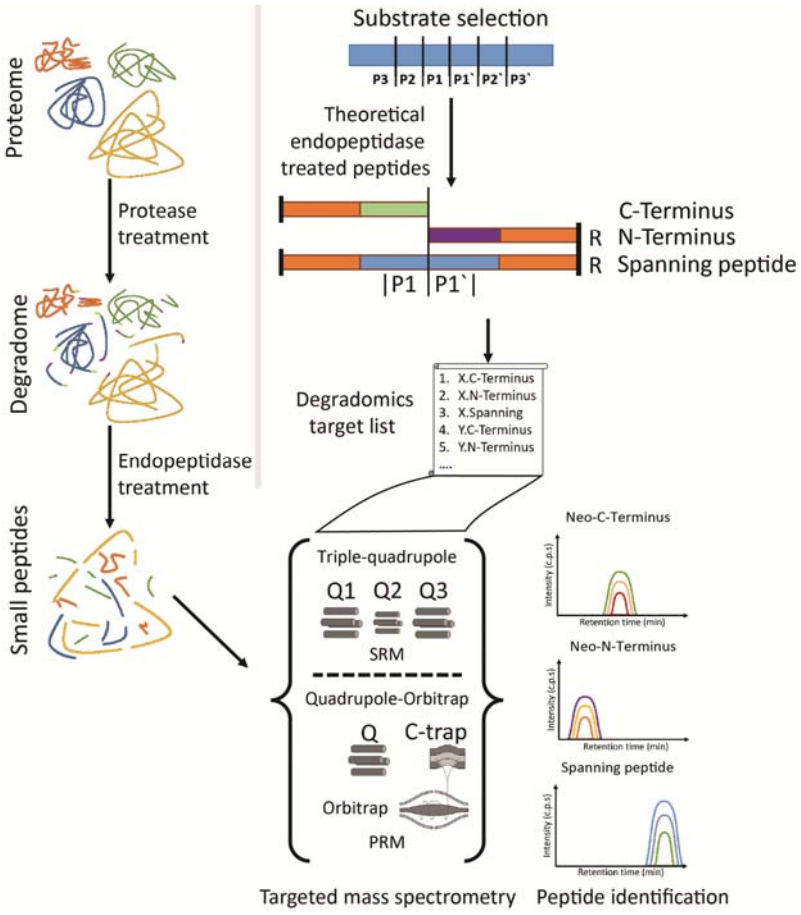

Figure 1 Generic targeted degradomics workflow.

Targeted degradomics workflows start with identifying the proteolytic cleavage of interest in protease cleavage databases like CutDB, MEROPS or UniProt. Numerous, high confidence proteolytic cleavages have been identified and annotated. Substrate proteins are then subjected to theoretical endopeptidase treatment with mass spectrometry working proteases like trypsin, ArgC or AspN to generate a unique peptide list of newly formed neo-N-termini, neo-C-termini and cleavage site spanning peptides. After selection and validation of degradomics target lists, sample proteomes are subjected to test protease treatment and digested by selected sequence specific working endopeptidases. Peptides are separated by in-line liquid chromatography and analyzed by targeted mass spectrometry for members of degradomics target lists. In SRM mode, triple-quadrupole instruments select both precursor ions (Q1) and predefined fragments (Q3), whereas quadrupole-Orbitrap instruments only select the precursor ion (Q) but follow all fragments over the chromatographic elution peak that is integrated for quantitation.

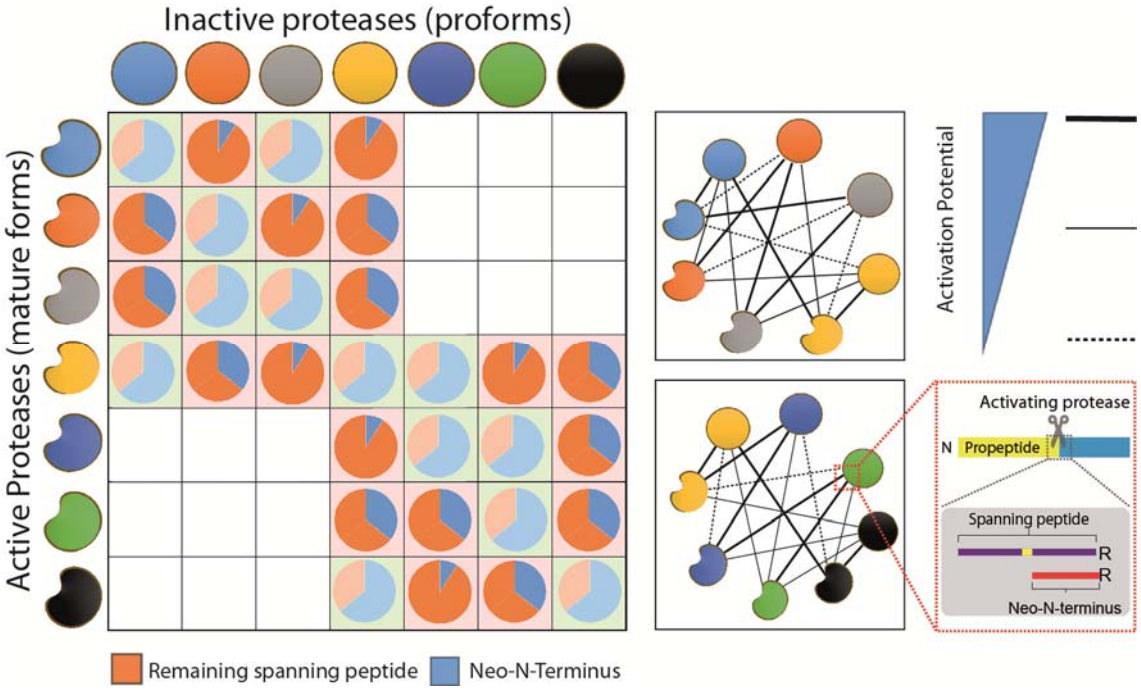

Figure 2 Mapping protease activation networks by targeted degradomics.

Activation of proteases is monitored by recording generation of neo-N termini and loss of cleavage site spanning peptides upon zymogen propeptide removal by targeted degradomics. Relative quantification of the spanning peptide of the inactive proform and the neo-N terminus of the active mature protease in relation to activities of activating upstream proteases determines degree of activation and thus activation potentials within an activation network. Integration of data and functional modulation of individual nodes allows mapping of complicated interconnected protease activation networks in complex biological matrices.