



Proteome-wide structural changes measured with limited proteolysis-mass spectrometry: an advanced protocol for high-throughput applications

Liliana Malinovska^{1,6}, Valentina Cappelletti^{1,6}, Devon Kohler^{2,6}, Ilaria Piazza^{1,3}, Tsung-Heng Tsai^{1,4}, Monika Pepelnjak¹, Patrick Stalder¹, Christian Dörig¹, Fabian Sesterhenn¹, Franziska Elsässer¹, Lucie Kralickova¹, Nigel Beaton⁵, Lukas Reiter^{1,5}, Natalie de Souza¹, Olga Vitek^{2,✉} and Paola Picotti^{1,✉}

Proteins regulate biological processes by changing their structure or abundance to accomplish a specific function. In response to a perturbation, protein structure may be altered by various molecular events, such as post-translational modifications, protein-protein interactions, aggregation, allostery or binding to other molecules. The ability to probe these structural changes in thousands of proteins simultaneously in cells or tissues can provide valuable information about the functional state of biological processes and pathways. Here, we present an updated protocol for LiP-MS, a proteomics technique combining limited proteolysis with mass spectrometry, to detect protein structural alterations in complex backgrounds and on a proteome-wide scale. In LiP-MS, proteins undergo a brief proteolysis in native conditions followed by complete digestion in denaturing conditions, to generate structurally informative proteolytic fragments that are analyzed by mass spectrometry. We describe advances in the throughput and robustness of the LiP-MS workflow and implementation of data-independent acquisition-based mass spectrometry, which together achieve high reproducibility and sensitivity, even on large sample sizes. We introduce MSstatsLiP, an R package dedicated to the analysis of LiP-MS data for the identification of structurally altered peptides and differentially abundant proteins. The experimental procedures take 3 d, mass spectrometric measurement time and data processing depend on sample number and statistical analysis typically requires ~1 d. These improvements expand the adaptability of LiP-MS and enable wide use in functional proteomics and translational applications.

This protocol is an update to *Nat. Protoc.* 12, 2391–2410 (2017): <https://doi.org/10.1038/nprot.2017.100>.

Introduction

The structure of proteins is tightly linked to their functional state. Multiple structural changes, such as post-translational modifications, binding of interaction partners, multimerization, aggregation and cleavage, can regulate protein activity. Therefore, the ability to simultaneously assess structural alterations of many proteins in their native environment is a valuable tool to evaluate the functional state of biological systems.

In 2014 we introduced limited proteolysis coupled with mass spectrometry (LiP-MS), which allows the proteome-wide assessment of protein structural alterations in complex biological samples. LiP-MS uses a nonspecific protease to digest the protein or proteome of interest under native conditions to yield a structure-specific proteolytic pattern, amenable to bottom-up mass spectrometry (MS) analysis. We first demonstrated the power of the method in a global assessment of altered protein conformations in *Saccharomyces cerevisiae* cultured in different nutrients¹. Since then, we have used it to study proteome thermostability in multiple organisms² and to map the metabolite-protein interactome in native *Escherichia coli* lysates (LiP-small molecule mapping, LiP-SMap)³. In a modified version (LiP-Quant), we applied the approach in proteome-wide drug target identification

¹Institute of Molecular Systems Biology, Department of Biology, ETH Zurich, Zurich, Switzerland. ²Khoury College of Computer Sciences, Northeastern University, Boston, MA, USA. ³Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC Berlin), Berlin, Germany.

⁴Department of Mathematical Sciences, Kent State University, Kent, OH, USA. ⁵Biognosys AG, Schlieren, Switzerland. ⁶These authors contributed equally: Liliana Malinovska, Valentina Cappelletti, Devon Kohler. [✉]e-mail: o.vitek@northeastern.edu; picotti@imsb.biol.ethz.ch

screens in human proteomes⁴. Most recently, we have shown that LiP-MS enables the analysis of protein functional alterations *in situ*⁵. The dynamic 3D proteome map generated by a LiP-MS experiment captures protein functional alterations due to enzymatic activity changes, allosteric regulation, phosphorylation, protein complex formation and protein aggregation, in acute stress responses in *S. cerevisiae* and nutrient adaptation in *E. coli*.

Since our first report of LiP-MS¹ and the associated protocol⁶, we have continuously optimized the method and have gained deeper insight into the various experimental parameters that affect a LiP-MS experiment. Here, we present an updated and improved version of the protocol that incorporates these optimizations, implements data-independent acquisition (DIA)-based MS and allows multiplexed, high-throughput sample processing. Together, these changes increase sensitivity and substantially reduce experimental variability, even with large sample numbers. This should improve performance in large-scale studies such as systematic functional proteomic studies and translational applications. Finally, we introduce MSstatsLiP, an analysis package for all-in-one quality control, data processing, statistical analysis and visualization of LiP-MS data, including visualization of structural barcodes as recently reported⁵.

Applications

LiP-MS is of interest to diverse research fields. Analogous to traditional proteomic approaches, it can be applied to the targeted or unbiased study of protein abundance changes induced by specific perturbations, with the important advantage of also detecting structural changes proteome wide. Such structural measurements are complementary to abundance measurements and substantially increase the information content of proteomic screens⁵. Thus, LiP-MS is a powerful new approach for functional proteomics. It can be applied to systematically study the response of cells and organisms to different environmental, chemical or genetic perturbations and characterize cellular processes or functions that are altered during the response.

Moreover, LiP-MS is a new tool within structural biology. It provides not only the possibility to monitor structural changes of thousands of proteins within the complex cellular milieu, but it can be used to pinpoint structural changes, such as binding of small molecules, with the resolution of single functional sites^{4,5}. Combined with other high-resolution structural methods, such as NMR spectroscopy or X-ray crystallography, it can be used to bridge protein structural information from *in vitro* and *in vivo* analyses. Specifically, with experimental adjustments (see Experimental design), the proteolytic pattern obtained from a protein *in vitro*, where the structure can be determined with high resolution, can be compared to the proteolytic pattern obtained from a protein in its native environment to make inferences about the structural state *in situ*^{1,5}.

Within systems biology, LiP-MS can be used for interactome analysis of proteins with small molecules^{3,4}. This can be extended to protein-protein interactions, as suggested by our studies on chaperone-client interactions⁵, and it could be applied to study global protein interactions with biopolymers such as DNA, RNA and lipids in the future. Lastly, the high-throughput capability of the optimized LiP-MS protocol described here enables translational approaches. LiP-MS can be used for drug target identification⁴, as well as in screens to identify a novel class of disease biomarkers that rely on structural alterations of proteins⁷.

Comparison with other structural methods

Other MS-based techniques could in principle be used to detect structural changes in complex proteomes. Cross-linking-MS has emerged as an appealing approach for the identification of protein-protein interactions and neighboring residues in complex samples^{8–10}. In addition, hydroxyl radical footprinting-MS^{11,12} could be applied to detect protein structural changes on a large scale. However, it is still challenging to apply both these approaches to complex proteomes or to perform comparative analysis of structural proteome profiles. Nevertheless, recent technical developments suggest that cross-linking-MS and surface footprinting-MS hold promise for obtaining large-scale protein structural information^{13–15}.

Among non-structural methods, the thermal proteome profiling (TPP) technique¹⁶, which measures protein precipitation due to heat-induced unfolding, has enabled unbiased identification of drug targets^{16,17}, identification of proteins binding to metabolites and proteome-wide exploration of phosphorylation events and their effect on protein thermal stability¹⁸. TPP probes changes in the biophysical properties of proteins and can be applied *in vitro* or *in vivo* without requiring a labeling step. Although TPP has the advantage of slightly deeper proteome coverage,

as recently shown by a direct comparison between the two methods⁴, it does not have structural resolution. LiP-MS data instead can pinpoint structural changes to regions of ~10 amino acids or ~6.4 Å in a 3D structure. Thus, for assessment of compound binding, LiP-MS can identify binding sites, whereas TPP identifies targets based on overall stability changes. More broadly, both LiP-MS⁵ and TPP¹⁹ can be used to gain mechanistic insights into protein functions and, in the case of LiP, to link the structural effects of perturbations to specific functional sites. Finally, while LiP-MS can in principle probe structural alterations for all detectable proteins, TPP requires the structural change to affect protein stability to produce a change in the melting profile.

The LiP-MS approach has some limitations. First, in order for a structural change to be detected by LiP-MS, the relevant protein region must be detected in the mass spectrometer. Second, in part because of the versatility of the method in detecting multiple types of molecular events, mechanistic interpretation of a LiP-MS change typically requires orthogonal data. Finally, the proteolysis step is carried out within the lysate, rather than in an intact cell or organism; in-cell LiP-MS requires further method development.

Overview of the protocol

A LiP-MS experiment typically tests the effects of a perturbation, whether environmental, chemical, genetic or other, on the system of interest. Briefly stated, LiP-MS uses a short, controlled proteolysis with a non-specific protease to probe proteome-wide differences in susceptibility to proteolysis between conditions. LiP-MS thus probes for structural changes in thousands of proteins simultaneously in the condition(s) under study. The LiP-MS method consists of five main steps: (i) sample preparation, in which each sample is split into two aliquots (LiP and trypsin-only control (TrP) samples); (ii) limited proteolysis under native conditions in the LiP sample; (iii) complete proteolytic digestion in denaturing conditions in both the LiP and TrP samples; (iv) data acquisition; and (v) data analysis and visualization (Fig. 1).

Sample preparation

LiP-MS is applicable to a broad variety of samples, ranging from purified proteins or complexes to whole-cell or tissue lysates and body fluids. The extraction method is thus dictated by the type of sample and should be selected for the highest protein yield while preserving non-denaturing and, ideally, physiological conditions, which are most likely to also preserve native protein structure. Recommended lysis methods for bacteria, yeast and mammalian cells have been described previously⁶ as well as methods to enrich for membrane proteins⁴. During the cell lysis step, endogenous proteases can lead to unwanted proteolysis. We therefore recommend performing the extraction step at low temperature (4 °C) and adding protease inhibitors to the extraction buffer (see also Influence of reaction condition on protease activity under Experimental design).

Each sample in a LiP-MS experiment is split into two aliquots and processed in two parallel pipelines, in each case in at least four replicates. The first aliquot (four replicates) is subjected to limited proteolysis with the nonspecific protease proteinase K (PK) in native conditions, which will

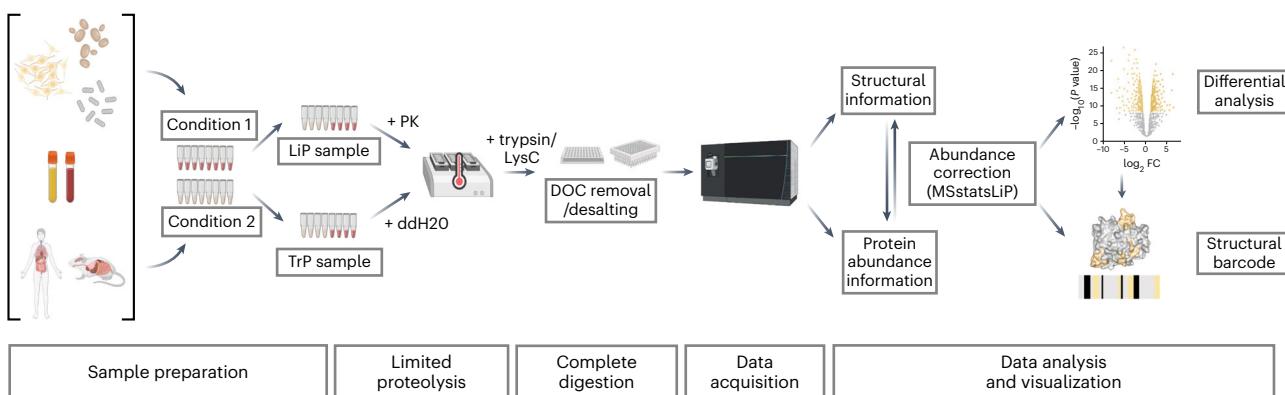


Fig. 1 | Schematic of LiP-MS protocol showing the five main steps. DOC, sodium deoxycholate; LysC, lysyl endopeptidase; PK, proteinase K. Created with BioRender.com.

generate a structure-specific proteolytic pattern, followed by complete digestion with the endoproteases trypsin and lysyl endopeptidase (LysC) under denaturing conditions (LiP sample, LiP). The other aliquot (four replicates) is directly denatured and subjected to complete digestion and is used for the assessment of protein abundance changes (TrP sample).

The LiP-SMap and LiP-Quant pipelines map the interaction of proteins with small molecules; in these approaches, small molecules are added to the lysate before performing limited proteolysis. In the LiP-SMap pipeline, small molecules are added at defined concentrations, whereas for LiP-Quant a concentration range is used. The samples are then processed by using the same workflow as LiP-MS.

Limited proteolysis

The controlled partial digestion of proteins under native conditions is the key step of the method. On the basis of their structural state, proteins are differentially susceptible to time-limited digestion by PK (for a discussion on alternative proteases, see ref. ⁶). Thus, flexible and accessible regions of a protein, like loops or unstructured domains, are preferentially targeted (Extended Data Fig. 1) while folding or aggregation reduces PK accessibility, and the extent of cleavage by PK can be used as a proxy for structural flexibility and accessibility. A molecular event that alters the structure of a protein can therefore alter susceptibility to PK and result in a distinct proteolytic pattern for each protein conformation.

To ensure high reproducibility in the cleavage pattern, the effect of several variables on this key proteolysis step must be considered in the experimental design (see Experimental design). In particular, the enzyme/substrate (E:S) ratio and the incubation time of the limited proteolysis step should be tightly controlled (see Effects of enzyme/substrate ratio and incubation time). PK is efficiently inactivated by heating (99 °C) and subsequent denaturation by using sodium deoxycholate (DOC) (for a discussion of suitable chaotropes and surfactants, see ref. ⁶). The complete inactivation of PK is key for the success of the experiment, because residual activity can lead to prolonged digestion and impede the generation of structure-specific digestion patterns.

Complete digestion

In this step, all samples are digested by trypsin and LysC under denaturing conditions to generate peptides amenable to bottom-up MS analysis. First, disulfide bonds are reduced and subsequently alkylated, to ensure full accessibility for digestion. Next, the samples are digested to completion with the proteases LysC and trypsin. For the LiP samples, the preceding digestion by protease PK will produce cleavage sites with non-tryptic ends. We refer to peptides harboring only one tryptic end as ‘half-tryptic’ (HT) peptides. For the control samples, which were digested solely with LysC and trypsin, peptides will have two tryptic ends; we refer to these as ‘fully tryptic’ peptides. The digestion reaction is stopped by the addition of formic acid, which leads to precipitation of DOC. Precipitated DOC is removed, and the peptides resulting from tryptic digestion are desalted by using a reverse phase C18 matrix.

Data acquisition

The samples are measured with bottom-up MS by using DIA, as described elsewhere²⁰. For a description of data-dependent acquisition (DDA) methods and targeted reaction monitoring, which can also be used in LiP-MS, see ref. ⁶. To deconvolute the fragment ion spectra, spectral libraries are used in the post-acquisition *in silico* data processing. In a classical proteomics experiment, public repositories can be used for this purpose; however, because LiP-MS data include HT peptides, a study-specific library should typically be generated. This can be done by using DDA methods on the same set of samples or by searching DIA data directly^{3–5,20}. The former ensures a more extensive spectral library and increases accuracy of identifications and quantifications but comes with the cost of longer acquisition times if all replicates of a sample are measured; this can be mitigated by pooling samples across replicates for the library-generation step.

Data analysis and visualization

The main goal of LiP-MS data analysis is the identification of proteins that are structurally altered in the condition of interest; by comparing the samples, it is possible to find out which protein regions are involved in the alteration (see ‘Data processing, quality control and statistical analysis’). Before the structural analysis, a LiP-MS dataset is assessed by using several quality control measures. Briefly, the peptide intensity distribution of each run, the coefficient of variation across biological replicates, the fraction of HT peptides in each replicate and the number of identified proteins and peptides are evaluated.

Box 1 | Correction of LiP-MS data for protein abundance changes

Structurally altered proteins are identified on the basis of a significant change in LiP peptide intensities between different conditions. However, if a peptide stems from a protein that is differentially abundant between the compared conditions, the change in LiP peptide intensity may be a consequence of abundance changes or may result from a combination of structural and abundance changes. Therefore, it is important to disentangle the contributions of abundance and structural changes to the intensity change of a LiP peptide.

One approach for this⁶ uses the fold change in the TrP samples to correct for protein abundance changes. Briefly, if a protein shows significant abundance changes, the fold change of all corresponding LiP peptides is corrected by the fold change of the protein. This approach has two main limitations: first, the definition of statistical significance for corrected LiP peptides is different from that for uncorrected LiP peptides; second, uncertainty in the protein abundance measurements is not incorporated into the statistical analysis of protein structural changes.

We describe here an alternative approach, based on a previously described method used to solve a similar problem in post-translational modification profiling experiments^{30,31}. This method applies linear mixed effect models to infer the relative abundance of LiP peptides (from the LiP dataset) and proteins (from the TrP dataset). Peptide and protein abundance estimates, and their standard error, are combined to adjust the LiP peptide abundance with respect to the underlying protein abundance. Structurally altered peptides are then detected on the basis of the adjusted peptide abundance. This approach is incorporated in the new LiP-tailored analysis pipeline MSstatsLiP.

We tested the performance of both correction approaches on a benchmark dataset. We used the disordered monomeric form (M) and the aggregated fibrillar form (F) of human α -synuclein, which can be distinguished by using LiP-MS¹⁵. We separately spiked the M and F conformations of α -synuclein, each at two concentrations (5 and 20 pmol/ μ g lysate), into a yeast lysate (Extended Data Fig. 9a). We then performed LiP-MS on all samples and compared the distribution of structurally altered α -synuclein peptides between samples (i) without adjustment for protein abundance changes (Fig. 8a, blue bar), (ii) using the classic approach (Fig. 8a, orange bar) and (iii) using MSstatsLiP (Fig. 8a, yellow bar). We first compared the same conformation spiked in at different concentrations (F1 versus F2 and M1 versus M2). Although the approach should detect no structural alterations in this setup, we observed many altered α -synuclein peptides ($|log_2FC| > 1$ and $q\text{-value} < 0.05$) without abundance correction due to the altered abundance of α -synuclein (41 for the F1 versus F2 comparison and 48 for the M1 versus M2 comparison; Fig. 8). This number is drastically reduced upon abundance correction using either method (Fig. 8a). The classic approach reduced the number of significant peptides ($|log_2FC| > 1$ and $q\text{-value} < 0.05$, from 41 to 2 for the F1 versus F2 comparison and from 48 to 13 for the M1 versus M2 comparison; Fig. 8a) solely based on the corrected fold change (Fig. 8b). MSstatsLiP outperformed the classic approach and reduced the number of peptides passing the significance cutoff ($|log_2(FC)| > 1$, $q\text{-value} < 0.05$) to 0 for the comparison of the fibrillar form (F1 versus F2) and to 4 for the comparison of the monomer (M1 versus M2). The peptides that survived the corrections are probably due to interactions of α -synuclein monomer with the yeast lysate (see also Supplementary Discussion, Benchmarking abundance correction approaches for LiP-MS).

When comparing the two structural states (F versus M), we analyzed all four possible combinations (F1 versus M1, F2 versus M2: same concentration; F1 versus M2, F2 versus M1: different concentration). Many significant peptides were identified after both correction approaches, as expected for structural comparisons of two protein conformations. Because the four combinations compare the same proteins, we expect to identify the same peptides in all comparisons. Here, MSstatsLiP also outperformed the classical approach, because it identified a larger fraction of peptides in all four comparisons (45% MSstatsLiP, 26% classical approach; Fig. 8c). In all comparisons, the significantly changing peptides mapped predominantly to the NAC region of the protein, which is known to form the amyloid core in the fibrillar form³² (Fig. 8d).

We further tested the performance of the abundance correction approaches on a previously characterized LiP-MS dataset from *S. cerevisiae* treated with a short osmotic stress⁵. As a result of the short perturbation, only a few proteins changed in abundance, while many more were structurally altered⁵, including several proteins of the MAPK-Hog1 signaling pathway, as expected on the basis of current knowledge of this perturbation. To investigate the impact of abundance correction when only a few proteins show significantly altered abundance, we compared the results of the differential analysis before and after correction for protein abundance changes by using MSstatsLiP. Around 20% of LiP peptides passing the significance cutoffs in the corrected data was not found in the non-corrected data, and vice versa (Fig. 9a), revealing the importance of correcting even for subtle protein abundance changes. Within the MAPK-Hog1 pathway, we identified significantly changing peptides ($|log_2(FC)| > 1$ and $q\text{-value} < 0.05$; Fig. 9b, orange labels) for most proteins irrespective of the correction method used. Only a few proteins of this pathway did not show any significant structural alteration (Fig. 9b, gray labels) or were not detected at all by MS (Fig. 9b, white labels). Notably, for most of the 21 structurally altered proteins, we detected one or more structurally altered peptides in the known functional site (using a distance cutoff of 6.4 Å as reported in ref.³ with both correction approaches (Fig. 9b). However, compared to the classical approach⁵, correction for protein abundance with MSstatsLiP identified a structural alteration on four additional proteins from this pathway (Pgi, Pbs2, Ypk1 and Eno1), and for two of them (Pgi and Pbs2), this structural alteration (Fig. 9b, Hog1 and Pgi boxes, orange peptide) mapped to a known functional site, a catalytic site and a phosphorylation site (Fig. 9b, colored green in the Pgi box and blue in the Hog1 box), respectively.

In conclusion, we strongly recommend correcting LiP-MS data for protein abundance changes. The correction method proposed by Schopper et al.⁶ and the new approach embedded in the MSstatsLiP package can both be used for this purpose. For datasets in which protein abundance changes are limited to a small number of proteins, we recommend using the MSstatsLiP package. When applied to the benchmark dataset, the MSstatsLiP approach indeed demonstrated better performance than the classic correction approach (Fig. 8). Nonetheless, the number of significant structural alterations might be underestimated after correction with MSstatsLiP for perturbations causing extensive protein abundance changes, because with this approach, the uncertainty associated with protein abundance measurement is incorporated into the statistical analysis of the LiP data. In these cases, the approach described previously⁶ can be used instead with the recommendation of using prior knowledge or orthogonal data (e.g., positive controls) to further increase the confidence of the result.

LiP-MS data analysis consists of differential analysis of peptide intensities between conditions, performed on the LiP samples. Both fully tryptic and HT peptides can be used. In this new version of the LiP protocol, the differential analysis is performed by using linear mixed models that integrate all quantitative information across all conditions, which increases the sensitivity²¹. Any contribution of protein abundance change is corrected by using the TrP samples (Box 1). Proteins whose structures are significantly different between conditions are identified in a differential analysis of the LiP samples; LiP peptides that change their amounts between conditions are identified on the basis of the statistical significance and the fold change of their intensity. Similarly, proteins whose abundances are significantly different in two conditions are identified in a differential analysis of the fully tryptic

peptides of the TrP samples. The cutoffs for statistical significance might be different depending on the experiment:

- For a global exploration of structural alterations generated upon a perturbation, such as the response to osmotic stress, we recommend significance cutoffs of $q\text{-value} < 0.05$ and $|\log_2(\text{fold change (FC)})| > 1$.
- For identifications of specific interactions with molecules, like small molecules in the LiP-SMap workflow or in vitro experiments, a more stringent $q\text{-value} < 0.01$ can be applied³.
- To extract binding parameters in the LiP-Quant pipeline that involves a dosage series, no hard significance cutoffs are set a priori, and structurally informative peptides are ranked mainly on the basis of the goodness of fit to a binding model over the ligand concentration range⁴.

In addition to selecting significant peptides by applying the above cutoffs, it is possible to prioritize true identifications by computing a frequency score, which considers how often a hit is found in unrelated LiP experiments⁴.

In all LiP-MS experiments, the distribution of altered peptides can be visualized in volcano plots and quantified by using bar diagrams. We have recently introduced a new visualization of structural alteration of an individual protein⁵, in which we map the intensity of the altered peptide(s) along the sequence of the protein and visualize protein coverage in addition, generating structural barcodes (Fig. 1).

Protocol optimization

Here, we describe several improvements to the LiP-MS protocol⁶, improving reproducibility and enabling analysis of large sample sizes.

Higher throughput

We have adapted the sample preparation procedure to substantially increase the number of samples that can be processed, from up to 30 samples in 2 d⁶ to up to 192 samples in 2 d or several hundred samples in 3 d. These adaptations include:

- 1 *Heat inactivation using a thermocycler instead of a water bath*^{3–5}. This allows parallel sample processing, reduces experimental variability (Extended Data Fig. 2a) and uses a commercially available instrument instead of custom equipment.
- 2 *An additional pausing point directly after PK inactivation*. This increases the flexibility of sample collection and allows parallel downstream sample processing, thus minimizing technical variability.
- 3 *The simultaneous digestion of samples with LysC and trypsin*. This reduces the experimental time, allowing for better control of batch effects⁵.
- 4 *The use of filters for removal of precipitated DOC*. This reduces experimental variability (Extended Data Fig. 2b) and enables sample processing in 96-well format after the LiP step, until and including peptide cleanup⁵.

Improved reproducibility and proteome coverage

We have implemented DIA of LiP-MS data. DIA-MS has emerged as a powerful alternative to common acquisition methods for quantitative proteomics²⁰. Combining the advantage of shotgun (DDA) and targeted (selected reaction monitoring and parallel reaction monitoring) proteomics, DIA provides better reproducibility, quantification and proteome coverage. Therefore, by implementing DIA of LiP-MS data instead of the previously described DDA method, we substantially increase experimental reproducibility and obtain a high proteome coverage^{3,5}.

R package for data analysis

We introduce MSstatsLiP, an R package specifically designed for the analysis, quality assessment and visualization of LiP-MS data. MSstatsLiP provides functions for data quality control, normalization to minimize systematic variations across MS runs, summarization of spectral features, quantification of LiP peptides and their corresponding proteins and detection of protein structural alterations and/or abundance changes across experimental conditions. MSstatsLiP offers the possibility to visualize LiP-MS data as structural barcodes, representing proteolytic fingerprints along the sequence of a protein and providing a concise visual summary of protein regions showing conformation changes between conditions⁵. Finally, MSstatsLiP provides a more sophisticated statistical method to correct LiP-MS data for abundance changes (see the next section), based on an approach described for post-translational modification data²².

As part of this optimized protocol, we also discuss several key considerations for experimental design (see Experimental design), reflecting our current understanding of the variables that can affect a LiP-MS experiment.

Experimental design

The results of a LiP-MS experiment can be affected by several variables. Typically, the impact of each of these variables on the outcome will depend on the biological question. In this section, we provide a summary of key variables that should be considered in a LiP-MS experiment as well as recommendations on how to incorporate them into experimental design (see also Supplementary Discussion and Extended Data Fig. 3).

Effects of E:S ratio and incubation time

The readout of a LiP-MS experiment is differential susceptibility to PK digestion, and factors that influence digestion will therefore affect the outcome. We explored the effects of different E:S ratios as well as different incubation times on the outcome of a LiP-MS experiment, such as variability and coverage. We varied either the incubation time (at a fixed 1:100 E:S ratio (w/w)) or the E:S ratio (with a fixed 5-min incubation time) (Fig. 2a). In general, we observed that 70–80% of the detected peptides are found across all conditions (Extended Data Fig. 4a). We further assessed the effect of each stepwise increase in incubation time or decrease in E:S ratio (i.e., 30 s versus 1 min, 1 min versus 2 min, etc.) on the number of identified peptides and on peptide intensity (Fig. 2b,c). The pairwise comparisons of incubation times showed that most (>90%) peptides were identified in both conditions compared (Fig. 2b, left panel) and that these peptides also showed very few significant changes in intensity, when we applied the standard cutoff ($|\log_2(\text{FC})| > 1$, $\text{q-value} < 0.05$; Fig. 2c, left panel). Comparisons of E:S ratios also resulted in most peptides (>84%) being identified in both conditions and in few significant changes in peptide intensities, although the changes were larger than in the incubation time comparison (Fig. 2b,c, right panels). Overall this suggests that, for the conditions examined, LiP-MS results are robust over a range of incubation times and E:S ratios but that a change in E:S ratio can be more influential.

Conditions that promote increased digestion (i.e., long digestion times) could lead to secondary cleavages, which could affect the structural information obtained. To assess this, we compared peptide identification (Extended Data Fig. 4b) and intensity (Extended Data Fig. 4c) at each incubation time (1–60 min) versus the shortest incubation time (30 s). We observed that most peptide identifications were shared between each pair of conditions in particular at shorter PK incubation times; for instance, >93% of peptides were shared between the 5-min and the 30-s PK incubation conditions (Extended Data Fig. 4b), and among these, >90% showed no significant change in intensity ($|\log_2(\text{FC})| > 1$, $\text{q-value} < 0.05$; Extended Data Fig. 4c). The fraction of shared peptides and of peptides with unaltered intensity was reduced at longer incubation times but was still relatively low (e.g., >86% of peptides were shared between the 60-min and 30-s incubation condition, and of these, >83% showed unchanged intensity). Interestingly, across the dataset, peptides that were lost upon increased digestion were on average longer than the shared peptides, as were also HT peptides that decreased in intensity compared to those that stayed the same (Extended Data Fig. 4d,e). Although these data suggest a low extent of secondary cleavage events, this is likely to be negligible at 5-min or shorter PK incubation times because most peptides do not change under these conditions.

We further assessed whether incubation time and E:S ratio influence the number of peptides and proteins identified in a LiP-MS experiment, the overall HT content and the coefficient of variation of the peptide intensities (Fig. 2d). With increased protein digestion (i.e., increasing incubation time or PK amount), the number of identified proteins in LiP samples decreases. In addition, the number of identified peptides is highest after 1 min of digestion or at a 1:500 E:S (w/w) ratio; it decreases with prolonged incubation time or higher PK amounts. The overall HT content increases at higher digestion conditions. The variability of the experiment (median coefficient of variation (CV)) is lowest for a 5-min incubation time or for a 1:100 E:S ratio (w/w), respectively. Moreover, PK digestion itself can induce structural alterations of the proteins being probed, which is expected to be more pronounced with increased digestion. A LiP-MS experiment should therefore use as short an incubation time and as low an E:S ratio as possible. Therefore, balancing the need for limited digestion with high robustness, we generally recommend using an E:S ratio of 1:100 (w/w) and an incubation time of 5 min to ensure the most robust and structurally informative results. If high identification rates are a priority, the incubation time or PK amount can be decreased further.

Note that the two datasets were acquired independently on different instruments, so that absolute numbers are not comparable between them. In general, the number of identifications in a LiP-MS

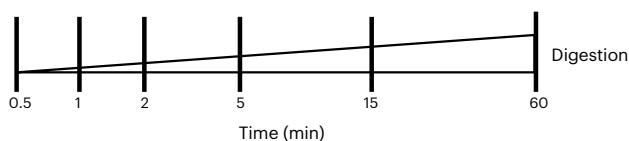
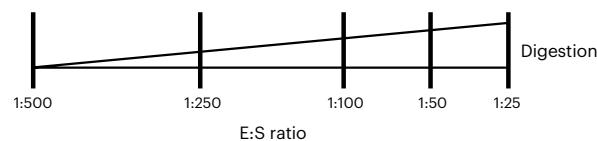
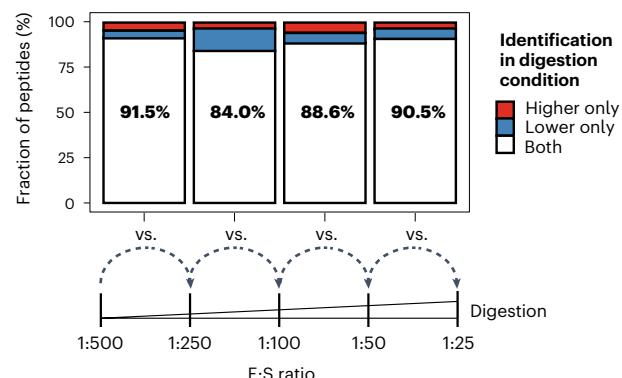
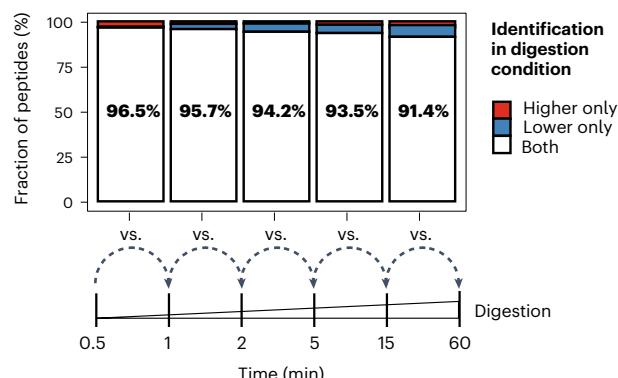
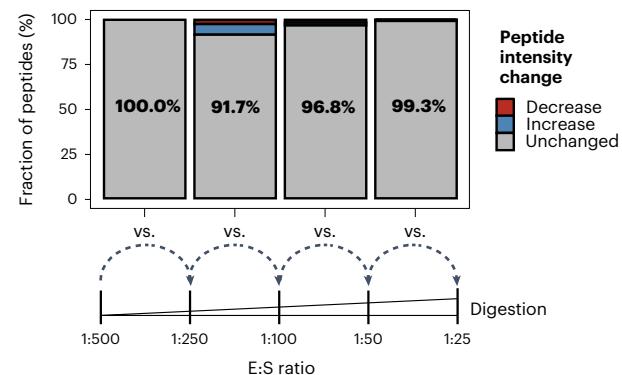
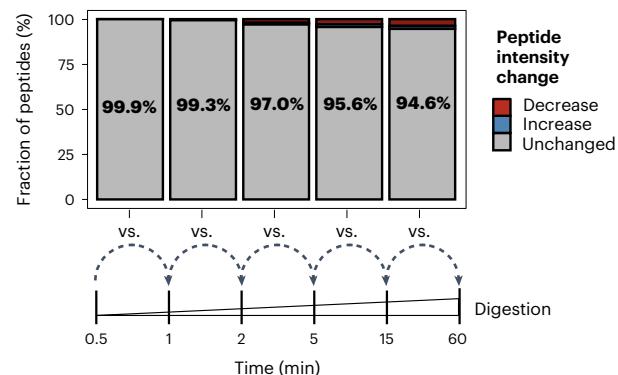
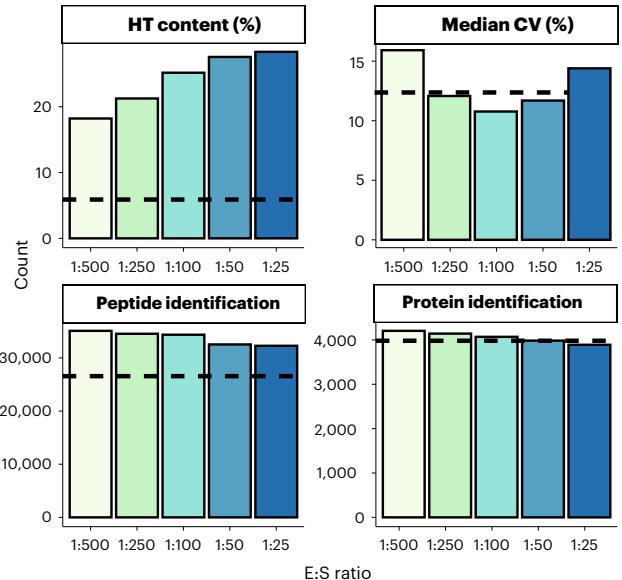
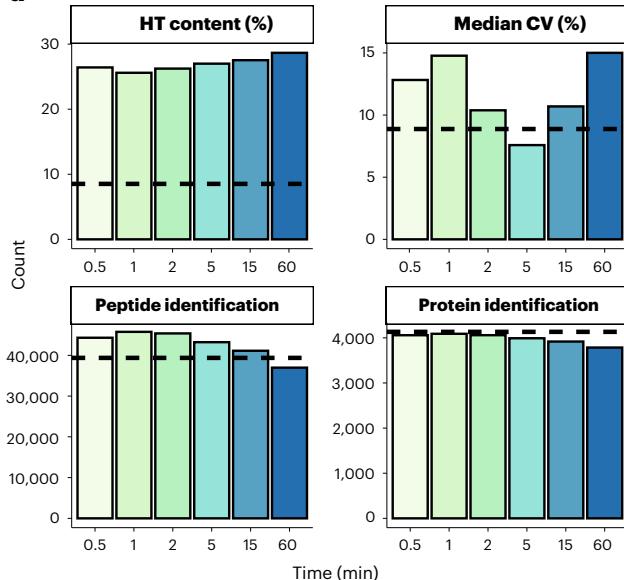
a**Incubation time****E:S ratio****b****c****d**

Fig. 2 | Influence of E:S ratio and incubation time. **a**, Experimental design. Left: LiP experiment at 1:100 E:S ratio (w/w) with increasing incubation time; the time points used are indicated. Right: LiP experiment with increasing amount of PK, incubated for 2.5 min. The E:S ratios used are indicated. **b** and **c**, Pairwise comparisons of peptide identifications and intensity changes in each stepwise increase of incubation time (left) or E:S ratio (right). **b**, Fraction of peptides identified only in the higher digestion condition (red), only in the lower digestion condition (blue) or shared between both conditions (white). **c**, Fraction of peptides with significantly changed intensities ($|\log_2(\text{fold change})| > 2$, $\text{q-value} < 0.05$; increase indicated in blue, decrease indicated in red and no change indicated in gray. **d**, Experiment characteristics under different conditions. The content of HT peptides (percentage), the median coefficient of variation (CV) (percentage), the number of identified proteins and the number of identified peptides are calculated in each condition. Conditions are colored by incubation time or E:S ratio, respectively. The black dotted line represents the value of each characteristic in the trypsin control sample.

experiment, as in any MS experiment, will depend on the instrument and settings used for data acquisition. These datasets were acquired on a mammalian cell extract. We also show a comparison of the experimental characteristics in different organisms at an E:S ratio of 1:100 (w/w) and an incubation time of 5 min (Extended Data Fig. 5).

Influence of reaction condition on protease activity

Because a LiP-MS experiment is frequently used to determine structural changes upon perturbations, it sometimes requires the addition of salts or small molecules, or a change of buffer, pH or reaction temperature or some other perturbation to the samples. In this case, it is necessary to assess the effect of the perturbation on the activity of PK itself.

To test PK activity, we suggest using the proteolytic resistance values of a well-defined set of reporter peptides. Proteolytic resistance is calculated as the ratio of the intensity of fully tryptic peptides in the LiP condition to the TrP condition and can be compared across different conditions by using the linear mixed effects model-based differential analysis implemented in the MSstatsLiP package (see Box 1; for step-by-step instructions, see the proteolytic resistance analysis notebook, LiP-MS_data_analysis_proteolytic_resistance.Rmd, <https://doi.org/10.5281/zenodo.5749994>). In general, a low proteolytic resistance value is indicative of a high extent of cleavage, whereas high proteolytic resistance values indicate a low cleavage extent.

An appropriate set of reporter peptides consists of peptides that, in the absence of any perturbation, are cleaved by PK to different extents, ranging from almost completely digested to not digested at all. Such a peptide set can be generated by digesting a whole-cell extract of *E. coli* (also other organisms can be used) with trypsin, followed by standard DDA^{6,23,24}. We suggest choosing ~100 fully tryptic peptides with the following criteria:

- Peptides with a CV of <5% across replicates
- Peptides with a proteolytic resistance value between 0 and 0.8

We recommend generating selected reaction monitoring assays for the reporter peptides and using these assays²⁵ for all subsequent PK activity tests. This increases the identification rate for the reporter peptide set. If this is not possible, the peptides can also be measured by using a shotgun MS approach. In this case, the analysis should be performed with the set of reporter peptides that is identified in all samples.

A LiP experiment using the peptides as a substrate should be performed with and without the perturbation. The effect on PK activity can then be visualized by using a proteolytic resistance plot, in which peptides are ranked by proteolytic resistance (Fig. 3). For instance, addition of trifluoroethanol (TFE) to the LiP buffer results in a global increase in proteolytic resistance for most reporter peptides (Fig. 3a). This indicates that PK is not active in TFE. In contrast, we illustrate a condition in which PK activity is not impaired (Fig. 3b). Here, we test a proteinase inhibitor cocktail that is typically added to the lysis buffer in a LiP-MS experiment to prevent degradation by endogenous proteases (see ‘Sample preparation’). Although this cocktail targets serine/asparagine proteases and would therefore inhibit PK at high concentrations, we show here two concentrations within the range used in a typical LiP experiment and that are not expected to inhibit PK in this context. We observe no reduction in PK activity in either condition.

It is important to note that different preparations of trypsin-digested lysates might lead to the identification and use of different sets of reporter peptides (as is the case in Fig. 2a versus 2b). Therefore, we highly recommend always including a positive control with the reporter peptides (i.e., where PK treatment is carried out in standard LiP buffer).

Structural comparison of pure proteins and proteins in complex backgrounds

A LiP-MS experiment may involve comparing the structure of a protein in its purified/in vitro and *in situ* states. Such studies can provide important insights into the functional state of the protein in its cellular context. For instance, by comparing the proteolytic pattern of an *in vitro* preparation of the

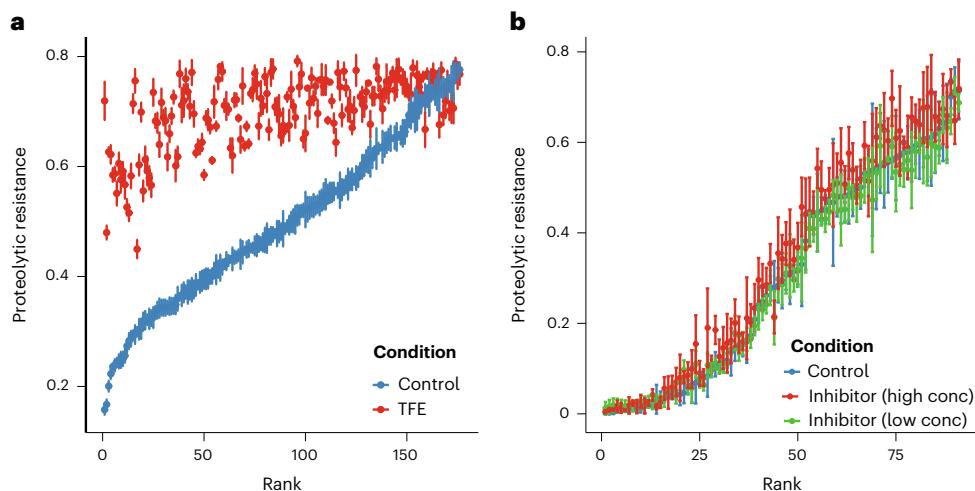


Fig. 3 | Proteolytic resistance plots of reporter peptides. **a**, Effect of trifluoroethanol (TFE) on PK activity. Proteolytic resistance values of reporter peptides are plotted against their ranks. The rank is determined by the proteolytic resistance value in the control condition. The comparison of LiP buffer (blue) and LiP buffer with TFE (red) shows a change in proteolytic resistance value. Error bars are standard deviation. **b**, Effect of proteinase inhibitor cocktail on PK activity, plotted as in **a**. conc, concentration.

substrate-bound form of an enzyme to its in situ proteolytic pattern, one could draw conclusions on the structural and functional state of that protein within the cell⁵. A prerequisite for this is that the proteolytic pattern of a given protein is comparable across different backgrounds. However, some proteins show a low correlation of peptide intensities depending on whether they are processed in buffer or spiked in to a complex cell lysate (Fig. 4a–d). Notably, this effect varies depending on the protein: for instance, myoglobin, a globular protein, shows a good correlation (0.92), whereas monomeric α -synuclein, an intrinsically disordered protein, shows a relatively poor correlation of peptide intensities (0.66) measured in different backgrounds.

Background-dependent low correlation in peptide intensities can be the result of a background-dependent differential extent of cleavage by PK or can be caused by interactions of the protein of interest with proteins or other factors in the background. Background-induced interactions are expected to affect certain regions of a protein; thus, in such a case, only a subset of peptides should show poor correlation. Moreover, if the peptide intensities in various complex (i.e., non-buffer) conditions show a high correlation (Extended Data Fig. 6), background-induced interactions are unlikely. The extent of background-dependent PK cleavage can be adjusted by titrating the E:S ratio and the PK incubation time for each protein of interest. We show an example of adjusting the E:S ratio and incubation time for monomeric α -synuclein in a complex background (Fig. 4e). The relatively poor correlation with monomeric α -synuclein in buffer when both samples are processed identically (E:S ratio of 1:100 (w/w), 5 min, yellow) can be improved if the sample in complex background is processed at a lower E:S ratio or a shorter incubation time. The best correlation is achieved at an E:S ratio of 1:250 (w/w) and an incubation time of 5 min (green).

Influence of confounding factors on reproducibility

In LiP-MS, as in any high-throughput experiment, non-biological confounding factors can cause systematic technical variations or batch effects. These effects can arise from multiple steps in a LiP-MS experiment, such as sample collection, sample preparation and liquid chromatography (LC)–tandem MS data acquisition (Fig. 5). Efforts should be made to reduce such effects at all steps; if they remain in the data, statistical approaches should be used to mitigate them.

Expertise

The steps described in this protocol do not need specialized expertise. However, several critical steps require precise timing. Especially when larger numbers of samples need to be processed, sample handling and timing can be challenging and can affect the quality of the data (see the Troubleshooting section).

The following experiment is a standard training step in our laboratory. In brief, the experiment aims to identify the target of rapamycin in a yeast lysate, as described previously⁴.

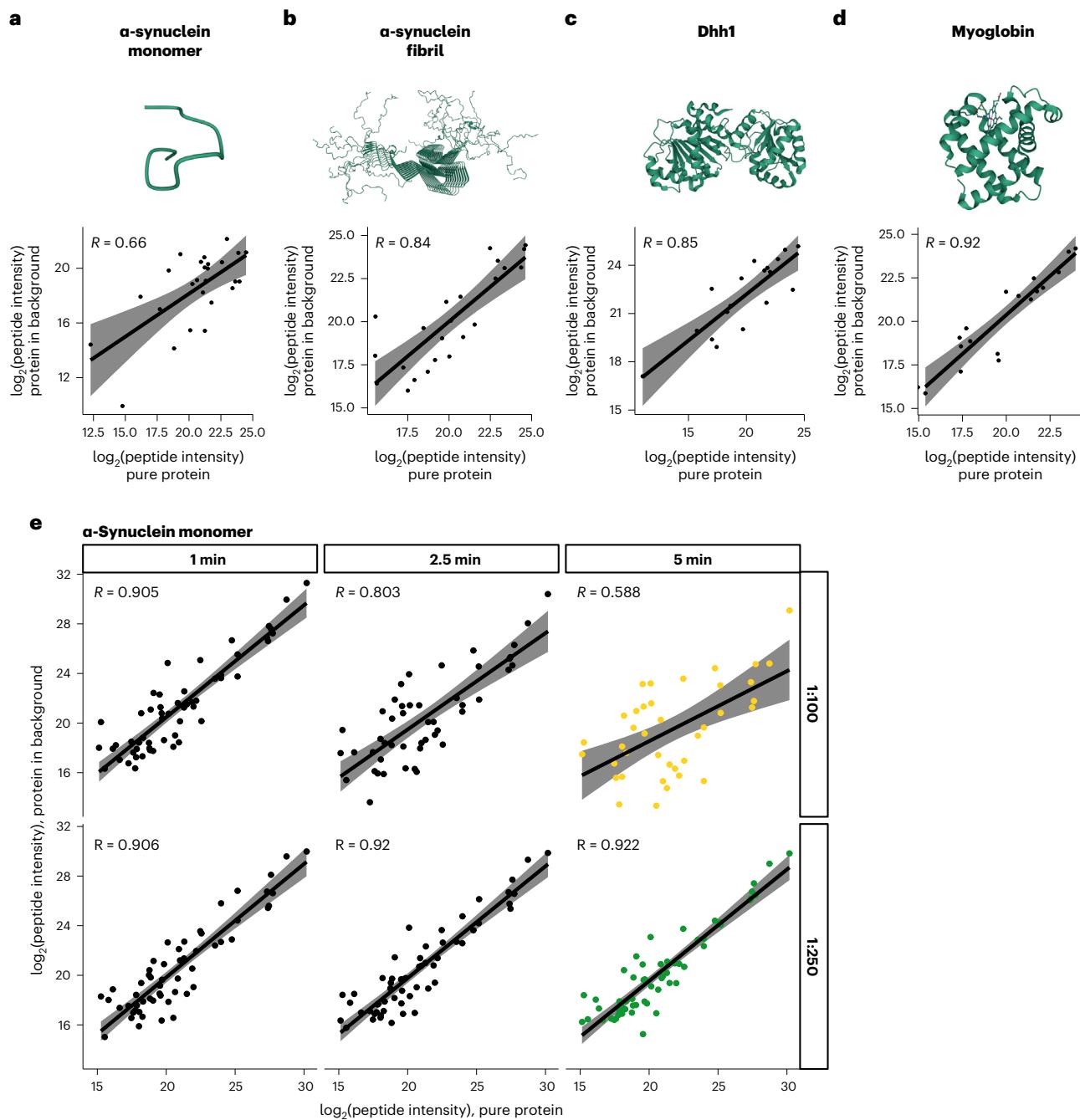


Fig. 4 | Comparison of LiP peptide intensities of proteins measured in buffer or in a proteinaceous background. **a-d**, Monomeric α -synuclein (**a**), fibrillar α -synuclein (**b**), Dhh1 (**c**) and myoglobin (**d**). Upper panel: protein structure (monomeric α -synuclein: schematic representation; fibrillar α -synuclein: PDB 2NDA; Dhh1: PDB 1S2M; myoglobin: PDB 1XCH). Lower panel: correlation analysis of \log_2 -transformed peptide intensities of protein in buffer versus in a complex background. The Pearson correlation coefficient is indicated. **e**, Adjusting the E:S ratio for monomeric α -synuclein. Correlation of peptide intensities of monomeric α -synuclein in buffer (X axis), in all cases incubated at an E:S ratio of 1:100 for 5 min, and in proteinaceous background (Y axis), incubated at the indicated E:S ratio and time points. The best correlation (green) is achieved at an E:S ratio of 1:250 (w/w) and 5-min incubation time, and the correlation improves substantially when both samples are processed by using the same condition (yellow).

- Prepare yeast cell lysates from four independent biological replicates at a final concentration of 2 mg/ml and divide into aliquots of 100 μ g of total protein.
- Divide the samples in two batches and incubate one at 25 °C for 10 min with 2 μ M rapamycin and the other without addition of the drug (see Step 5B).
- Add PK at an E:S ratio of 1:100 (w/w) to the treated and the untreated lysates and incubate at 25 °C (see Step 6).

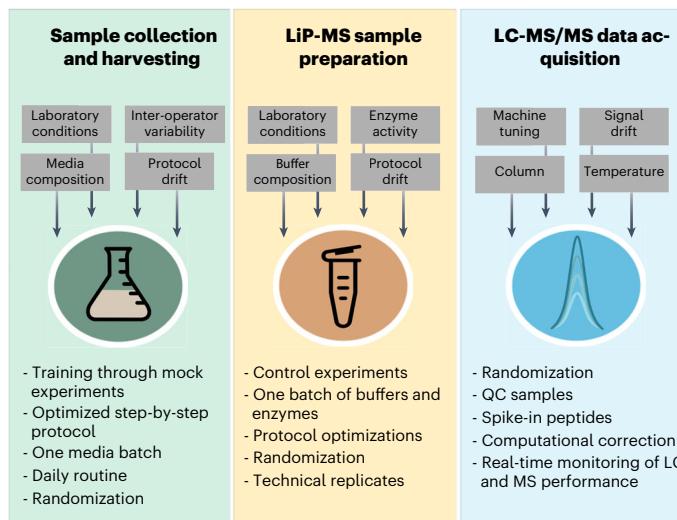


Fig. 5 | Potential sources of batch effects throughout the LiP-MS workflow and measures to minimize them. Batch effects can occur during sample collection and harvesting, LiP-MS sample preparation and LC-MS/MS data acquisition.

- Stop the reaction after precisely 5 min by heat inactivation, followed by denaturation with 5% (wt/vol) DOC (Steps 8–10).
- Process the samples as described below (Step 12 onward).

The CVs across replicates in each condition (treated and untreated) should be low (mean CV < 0.2) (Fig. 6a). If the effect of the applied perturbation is large enough to be the primary source of variance between samples, as in this example, replicates should cluster together, and the two conditions should be clearly separated (Fig. 6b,c). The peptide intensities in each run should display a good correlation (Pearson correlation coefficient > 0.9) within a condition and lower correlation between conditions (Fig. 6d). In case the data do not meet these conditions, refer to the Troubleshooting section for more details.

Upon differential analysis of protein structural changes (see ‘Data processing, quality control and statistical analysis’), peptides of the known rapamycin target FRP1 should be among the highest-scoring peptides. In the example in Fig. 6e, seven peptides (of 24,802 identified peptides, mapping to 2,417 proteins) passed the significance cutoffs, two of which map to FRP1.

We note that there may be cases in which the perturbation does not produce enough of a global change to yield good clustering and separation; in fact, even rapamycin treatment of mammalian cells (as opposed to yeast, as in this case) can prove problematic in this regard. When the magnitude of the expected change is not known, a lack of replicate clustering and condition separation on its own does not mean that the experiment is invalid. In such a case, a multiple-data-points experiment can help increase confidence in the results (see also Supplementary Discussion, Titration analysis with MSstatsLiP).

Materials

Reagents

- HEPES BioPerformance, certified 99.5% (Sigma-Aldrich, cat. no. H4034)
- Potassium chloride (Merck, cat. no. 104.936.1000)
- Magnesium chloride hexahydrate, puriss p.a. (Fluka, cat. no. 63072)
- cComplete, EDTA-free protease inhibitor cocktail (Merck, cat. no. 11873580001)
- Albumin standard (Thermo Fisher Scientific, cat. no. 23209)
- Pierce bicinchoninic acid assay, BCA (Thermo Fisher Scientific, cat. no. 23227) **! CAUTION** Pierce BCA protein assay contains a component classified as an aquatic environmental hazard (GHS09).
- PK from *Tritirachium album*, lyophilized powder, BioUltra (Sigma-Aldrich, cat. no. P2308) **! CAUTION** PK is classified as harmful (acute toxicity, GHS07) and a health hazard (GHS08).
- Ammonium bicarbonate (Ambic; Fluka Analytical, cat. no. 40867) **! CAUTION** Ammonium bicarbonate is classified as harmful (acute toxicity, GHS07).
- Tris(2-carboxyethyl) phosphine hydrochloride (Pierce TCEP-HCl; Thermo Fischer Scientific, cat. no. 20490) **! CAUTION** TCEP-HCl is classified as corrosive (GHS05).

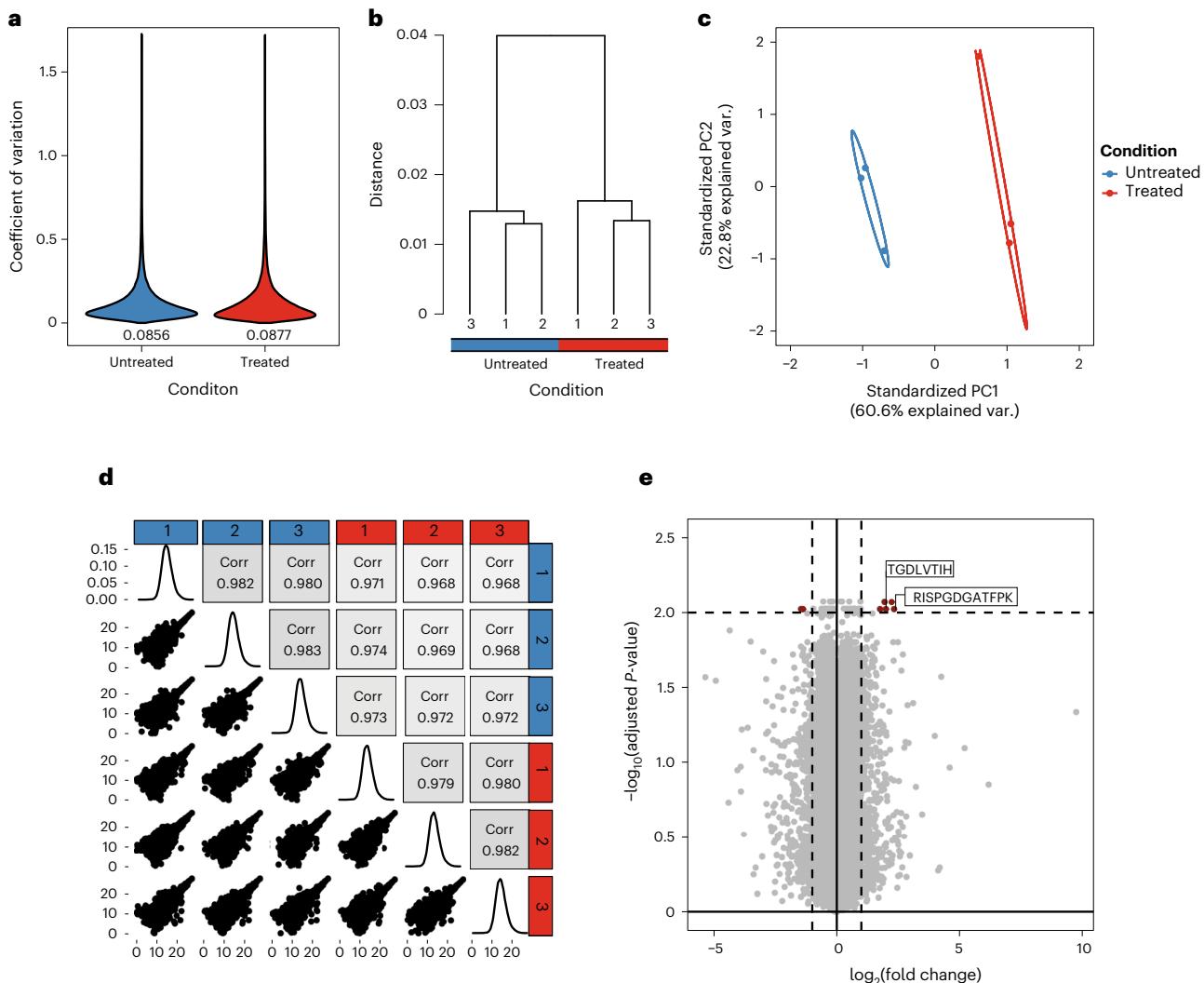


Fig. 6 | Anticipated results from a LiP-MS training experiment: rapamycin treatment of a yeast lysate. **a**, CV (calculated as the ratio of the standard deviation to the mean) of non-transformed fragment ion intensity for biological replicates in rapamycin-treated (red) and untreated (blue) conditions. Data distribution analysis by violin plots. The median CV is displayed below the violin plot. **b**, Dendrogram clustering: hierarchical clustering of Pearson correlation coefficient of \log_2 -transformed fragment ion intensities using Ward clustering ref. ²⁸. Replicates are colored in grayscale. Conditions are colored as in **a**. **c**, Principal component analysis (PCA) of \log_2 -transformed fragment ion intensities. Conditions are colored as in **a**. **d**, Run correlation. Pearson correlation coefficient of \log_2 -transformed fragment ion intensities visualized as a colored square (upper triangle) with values ranging from 0.968 (gray) to 0.982 (white). Scatter plot of \log_2 -transformed fragment ion intensity correlation (lower triangle) and data distribution density plot (diagonal). Conditions are colored as in **a**. **e**, Differential analysis of structurally altered peptides, visualized in a volcano plot. Each point represents a peptide. Upon application of the significance cutoffs ($|\log_2(\text{FC})| > 1$, $\text{q-value} < 0.01$), seven peptides are identified as significant (red). Two of them are of the known target of rapamycin, FPR1. Their sequence is shown. var., variance.

- Iodoacetamide BioUltra (IAA; Sigma-Aldrich, cat. no. I1149) !CAUTION IAA is classified as toxic (GHS06) and a health hazard (GHS08). ▲CRITICAL IAA is light-sensitive.
- LysC for biochemistry (Wako Pure Chemical Industries, cat. no. 129-02541)
- Sequencing-grade modified trypsin, frozen (Promega, cat. no. V5113)
- DOC, 97% (Sigma-Aldrich, cat. no. D6750) !CAUTION DOC is classified as harmful (acute toxicity, GHS07).
- Formic acid, ~98% (Fluka Analytical, cat. no. 94318). !CAUTION Formic acid is classified as flammable (GHS02), corrosive (GHS05) and toxic (GHS06).
- Methanol ROTISOLV, >99.95%, LC-MS grade (Roth, cat. no. AE71.2) !CAUTION Methanol is classified as flammable (GHS02), toxic (GHS06) and a health hazard (GHS08).
- Acetonitrile ROTISOLV, >99.98%, ultra LC-MS (Roth, cat. no. HN40.2) !CAUTION Acetonitrile is classified as flammable (GHS02), harmful (GHS07) and a health hazard (GHS08).

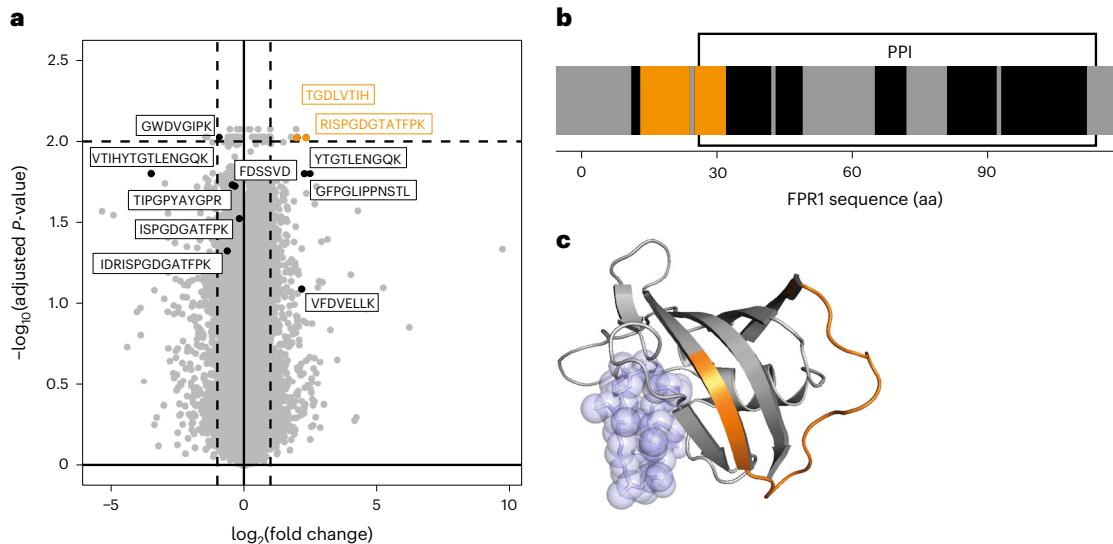


Fig. 7 | Structural alterations of FPR1 upon rapamycin treatment. **a**, Differential analysis of structurally altered peptides, visualized in a volcano plot. Peptides from FPR1 are highlighted (non-significant: black; significant: orange; significance cutoffs: $|\log_2(\text{FC})| > 1$, $\text{q-value} < 0.01$), and the sequence of the peptides is displayed. **b**, Structural barcode visualizing peptides mapped along the sequence of FPR1. Significant peptides are colored in orange, and non-significant peptides are colored in black. Gray regions represent regions with no identified matching peptide. Position of the peptidyl-proline isomerase (PPI) domain is indicated by a rectangle. **c**, Structural representation of FPR1; homology modeling based on human FK506-binding protein complexed with rapamycin (PDB 2DG3). Significant peptides are colored in orange, and bound target (rapamycin) is colored in purple.

Reagent setup

Native lysis buffer

Native lysis buffer consists of 20 mM HEPES pH 7.4, 150 mM KCl and 10 mM MgCl₂. It can be stored at 4 °C.

Protease inhibitor

Protease inhibitor is a 25× stock solution. Place one protease inhibitor tablet in 2 ml of H₂O. The solution can be stored at -20 °C for ≥12 weeks.

DOC solution

DOC solution is 10% (wt/vol) DOC in HPLC-grade water. The solution should be freshly prepared.

Ammonium bicarbonate solution

The solution is 100 mM ammonium bicarbonate. The solution should be freshly prepared.

TCEP solution

The solution is 300 mM TCEP-HCl in HPLC-grade water. It can be stored at -20 °C for ≥1 month.

IAA solution

The solution is 400 mM IAA in HPLC-grade water. IAA is light-sensitive; protect it from light with aluminium foil. The solution should be freshly prepared.

Buffer A

Buffer A consists of 5% (vol/vol) acetonitrile and 0.1% (vol/vol) formic acid in HPLC-grade water. It can be stored at room temperature (20–24 °C).

Buffer B

Buffer B consists of 50% (vol/vol) acetonitrile and 0.1% (vol/vol) formic acid in HPLC-grade water. It can be stored at room temperature.

LC buffer A

LC buffer A consists of 0.1% (vol/vol) formic acid in HPLC-grade water. It can be stored at room temperature.

LC buffer B

LC buffer B consists of 0.1% (vol/vol) formic acid in 99.9% (vol/vol) HPLC-grade acetonitrile. It can be stored at room temperature.

LiP input samples

Appropriate input for the LiP-MS workflow can be (i) a complex sample, such as a whole-cell extract or body fluid; (ii) partially fractionated samples; or (iii) pure protein preparations of single proteins or complexes. We have used whole-cell extracts of *E. coli*^{2,3,5}, *S. cerevisiae*^{1,2,5}, *Thermus thermophilus*² and HeLa cell lines⁴. ▲CRITICAL If possible, perform lysis or fractionation in LiP buffer. To avoid degradation through endogenous proteases, we recommend adding proteinase inhibitor. At the recommended concentrations, the inhibitor does not affect PK activity (see Experimental design, Influence of reaction condition on protease activity). ▲CRITICAL If the pure protein preparation is stable in LiP buffer, consider a buffer exchange to LiP buffer in advance. Alternatively, test the activity of PK in the new buffer (see Experimental design, Influence of reaction condition on protease activity).

Equipment

- Thin-walled PCR tubes (Thermo Scientific, cat. no. AB-1182)
- Multichannel pipette (Eppendorf, cat. nos. 3125000010 and 3125000052)
- Thermocycler (AnalytikJena, cat. no. 846-x-070-723)
- Plate reader (BMG LABTECH, CLARIOstar)
- Thermoblock (Eppendorf, cat. no. 5384000012) with 96-well setup (Eppendorf, cat. no. 5306000006) and cover (Eppendorf, cat. no. 5308000003)
- 96-well sample collection plates, 2 ml (Waters, cat. no. WAT058958) or 1 ml (Waters, cat. no. WAT058957)
- FiltrEX 96-well filter plates (Corning, cat. no. CLS3505)
- 96-well MACROSpin plate (The Nest Group)
- MilliporeSigma Supelco PlatePrep 96-well vacuum manifold (Fisher Scientific, cat. no. 11-100-3078)
- Sonicator (Sono Swiss, SW 12 H)
- Centrifuge (Eppendorf, cat. no. 5810000320)
- SpeedVac (Eppendorf, cat. no. 5305000304)
- MS instrument: Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Scientific) equipped with nanoelectrospray ion source and coupled to a 40 cm × 75 µm (i.d.) HPLC column (New Objective, cat. no. PF360-50-10-N-5) packed with 1.9-µm C18 beads (Dr. Maisch, cat. no. r119.aq)
- NanoACQUITY ultraperformance LC (UPLC) system (Waters)

Equipment setup

Liquid chromatography

Use a Waters ACQUITY M-Class UPLC system to separate peptides on a 40 cm × 0.75 mm i.d. column (New Objective, cat. no. PF360-75-10-N-5) packed with 1.9-µm C18 beads (Dr. Maisch ReproSil-Pur 120). Fractionation is achieved with a gradient of buffer A (0.1% (vol/vol) formic acid, Carl Roth GmbH) and buffer B (99.9% (vol/vol) acetonitrile, 0.1% (vol/vol) formic acid, Carl Roth GmbH): a linear gradient from 3% to 35% buffer B over 120 min, followed by 5 min with an isocratic constant concentration of 90% buffer B. If a UPLC system, such as the Thermo Scientific EASY-nLC 1200, is not compatible with 99.9% (vol/vol) acetonitrile, adjust it to the highest concentration possible. The flow rate is set to 300 nl/min throughout the gradient.

DDA MS method parameters

Acquire MS1 scans over a mass range of 350–1,150 mass-to-charge ratio (m/z) with a resolution of 120,000. Select the most intense precursors for 3 s with an intensity threshold of 5×10^4 for collision-induced dissociation. Acquire the corresponding MS2 spectra with a resolution of 30,000, for maximally 54 ms. Use all multiply charged ions to trigger MS-MS scans followed by a dynamic exclusion for 60 s. Exclude singly charged precursor ions and ions of undefinable charge states from fragmentation.

DIA MS method parameters

Acquire 41 variable-width DIA isolation windows with a 1- m/z overlap between windows as described in Supplementary Table 1. Acquire DIA-MS2 spectra at a resolution of 30,000 with a fixed

scan range of 200–1,800 m/z and an automatic gain control (AGC) target of 1×10^5 . To mimic DDA fragmentation, calculate the normalized collision energy on the basis of the doubly charged center m/z of the DIA window. Choose maximum injection times to maximize parallelization. Acquire a survey MS1 scan from 350 to 1,400 m/z at a resolution of 120,000, with an AGC target of 2×10^5 or 100-ms injection time in between the acquisitions of the full DIA isolation window sets.

Software

- Proteome discoverer, version 1.4 and newer (Thermo Fisher Scientific, <http://planetorbitrap.com/proteome-discoverer>)
- Spectronaut Pulsar (Biognosys, <https://biognosys.com/shop/spectronaut>)
- MSstatsLiP R-package (<https://www.bioconductor.org/packages/release/bioc/html/MSstats.html>)

Procedure

Sample preparation ● Timing 1 h

▲ **CRITICAL STEP** When analyzing large numbers of samples, we recommend precisely controlling as many aspects of the procedure as possible, to reduce the effect of confounding factors on the variability and reproducibility of the results (see Influence of confounding factors on reproducibility).

- 1 Measure the protein concentration by using commercially available BCA colorimetric assays and adjust the total protein concentration.
 - (i) Adjust the protein concentration to 2 mg/ml by using LiP buffer.
▲ **CRITICAL STEP** In case of low protein concentrations (<2 mg/ml), adjust the concentration of all samples to the lowest protein concentration in the experimental setup.
- 2 Divide the sample in two. Perform limited proteolysis with one half of the sample (LiP sample); the other half will be digested with trypsin only (trypsin control sample) (see Step 6). The trypsin control samples are used to correct the contribution of protein abundance (Box 1 and Figs. 8 and 9). Distribute aliquots of the sample in a total volume of 50 μ l in PCR tube strips and store on ice.
▲ **CRITICAL STEP** We recommend performing the experiment in quadruplicate.
▲ **CRITICAL STEP** We recommend preparing aliquots of all LiP samples and all TrP samples on different strips, because they will be processed and analyzed independently.

Limited proteolysis ● Timing 25 min per PCR tube strip

▲ **CRITICAL** LiP-MS can be used to assess structural alterations due to various causes. To investigate structural differences in two or more biological conditions (e.g., response to treatment, environmental change or comparison of pathological states), follow option A at Step 5. To assess structural changes as a consequence of binding events to a specific molecule of interest (e.g., metabolite, drug target or other protein) by using LiP-SMap, follow option B at Step 5. To extract binding parameters from interactions with a molecule of interest by using LiP-Quant, follow option C at Step 5.

▲ **CRITICAL STEP** If using a thermocycler with three blocks, two strips can be processed in parallel. Make sure that the interval between the starting times of both strips allows sufficient time to perform each step of the protocol for each strip. We recommend a starting interval of 1.5–2.0 min, which can be reduced if the experimenter is experienced with the workflow and handling.

▲ **CRITICAL STEP** The overall time needed can further be reduced by establishing a concatenated workflow. The next strip can be started for incubation at 25 °C in one block while the current strip is being incubated at 99 °C in the other block. The best starting time for the next strip needs to be determined depending on the heating/cooling time of the instrument and the experience of the experimenter with the workflow.

- 3 Pre-heat the blocks of the thermocycler to 25 °C and 99 °C.
▲ **CRITICAL STEP** Use a heated lid for the block set to 99 °C.
▲ **CRITICAL STEP** Keep the lid open for the block set to 25 °C.
- 4 Prepare proteinase solution and ddH₂O; store in ice.
▲ **CRITICAL STEP** Use PK at a 1:100 (wt/wt) E:S ratio.
▲ **CRITICAL STEP** We recommend distributing aliquots of PK solution and ddH₂O in PCR tube strips, which allows the subsequent use of multi-channel pipettes.

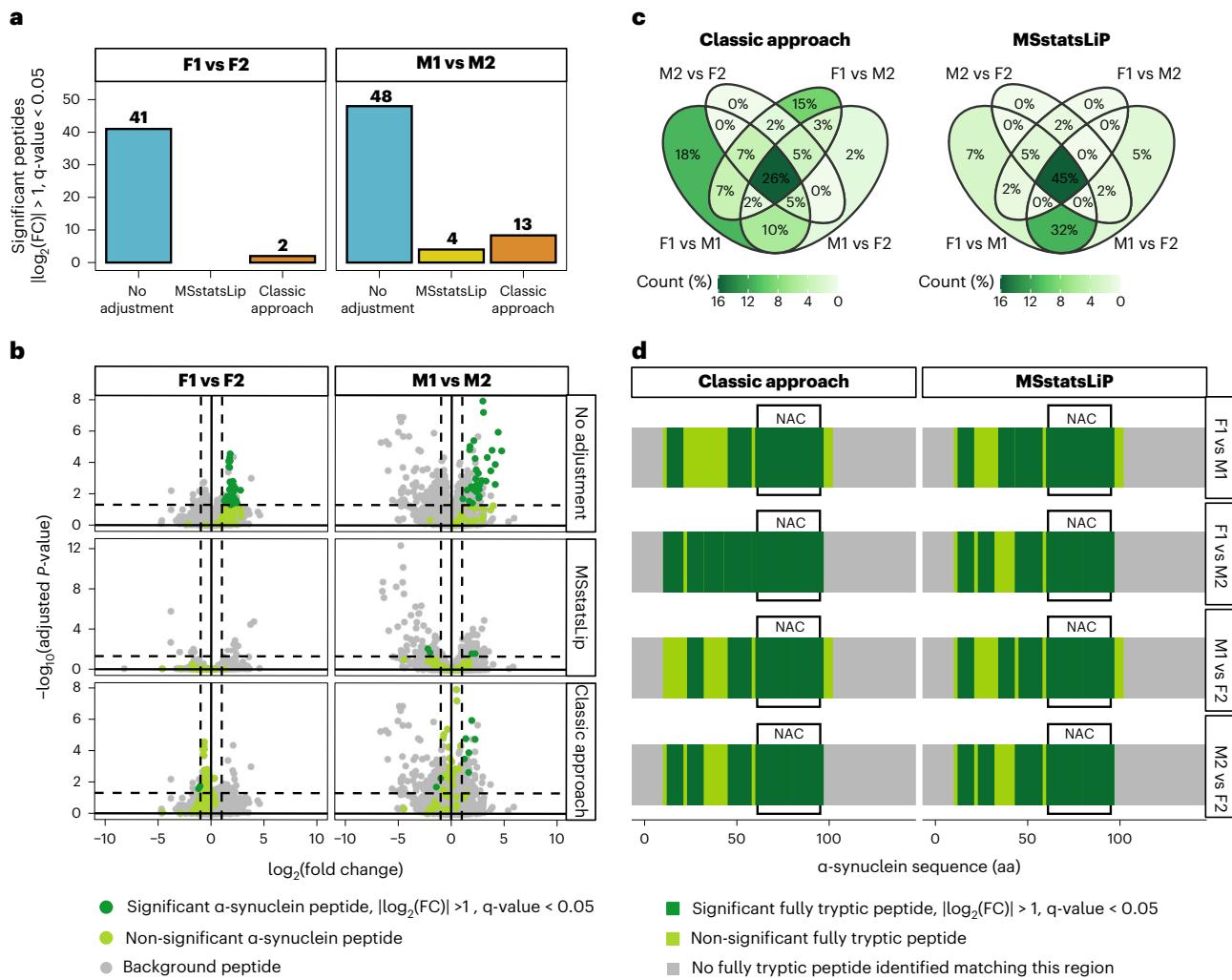


Fig. 8 | Benchmarking abundance correction approaches for LiP-MS. **a**, Quantification of significant peptides ($|\log_2(\text{FC})| > 2$, $\text{q-value} < 0.05$) of the indicated sample comparisons (α -synuclein spiked into yeast lysate). The plots compare the same structure at different concentrations (monomer (M) 1 versus M2 and fibril (F)1 versus F2), with no abundance correction (blue) or the indicated correction approach (yellow and orange). **b**, Differential analysis of LiP peptides in the comparisons described in **a**. Each dot represents an individual peptide; green indicates peptides of α -synuclein. Dark green, peptides passing the significance cutoff ($|\log_2(\text{FC})| > 2$, $\text{q-value} < 0.05$); light green, non-significant peptides. **c**, Venn Diagram visualizing the overlap of identified significant peptides in the four comparisons of F versus M by using the classical approach (left) and MSstatsLiP (right). The proportions of identified significant peptides in each set are visualized by color intensity. **d**, Structural barcodes visualizing fully tryptic peptides mapped along the sequence of α -synuclein in the comparisons described in **a** and corrected for protein abundance as indicated. Significant peptides are colored in dark green, and non-significant peptides are colored in light green. Gray regions indicate no identified fully tryptic peptide matching this region. The position of the aggregation region (NAC) is indicated as a line below the barcode. AA, amino acids.

5 Incubate the sample.

▲ CRITICAL STEP We recommend starting with the strips containing LiP samples.

(A) **Structural comparison across multiple conditions**

(i) Incubate the strip in the thermocycler block at 25 °C for 5 min.

(B) **LiP-SMap pipeline for single treatment**

(i) Incubate the treated sample with the molecule of interest. The incubation time, temperature and concentration of the target should be chosen according to the experimental setting. Incubate untreated sample with an adequate mock solution.

(ii) Incubate the strip in the thermocycler block at 25 °C for 5 min.

(C) **LiP-Quant pipeline for treatment series**

(i) Incubate the treated sample with the molecule of interest at a suitable concentration or time range. The incubation time, temperature and concentration range of the target should

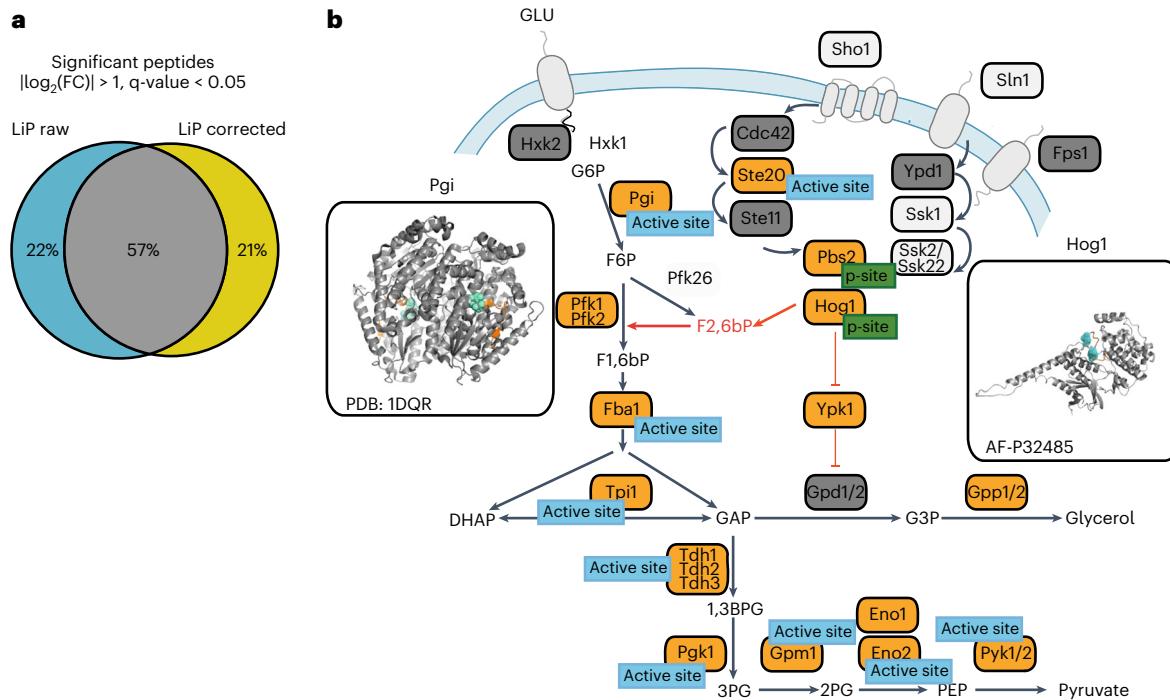


Fig. 9 | Performance of abundance correction approaches on proteins of the Hog1 signaling pathway upon yeast osmotic shock. **a**, Overlap of structurally altered peptides identified before (LiP raw) and after (LiP corrected) correction for protein abundance changes by using MSstatsLiP. Peptides are considered significantly altered when passing the cutoff of $|\log_2(\text{FC})| > 1$ and $\text{q-value} < 0.05$. **b**, Schematic representation of the HOG1-MAPK signaling pathway. Proteins with structurally altered peptides upon osmotic stress were identified by using MSstatsLiP. Proteins with structural changes are indicated with orange labels (significance cutoffs: $|\log_2(\text{FC})| > 1$ and $\text{q-value} < 0.05$), proteins with no detected structural changes are indicated with gray labels and proteins not detected by MS are indicated with white labels. If altered LiP peptides map to a distance $< 6.4 \text{ \AA}$ from a functional site, the type of site is indicated. The structural models show structurally altered LiP peptides (orange) mapped onto the 3D protein structures of available models (Hog1, AlphaFold prediction: AF-P32485 (ref. ²⁹); Pgi, PDB: 1DQR). For Hog1, the phosphorylation sites are indicated in blue, whereas for Pgi, the ligand (D-gluconate 6-phosphate) is indicated in green. Adapted with permission from ref. ⁵, Elsevier.

be chosen according to the experimental setting. Incubate untreated sample with an adequate mock solution.

(ii) Incubate the strip in the thermocycler block at 25 °C for 5 min.

- 6 Add 5 μl of PK to the LiP samples. Add 5 μl of ddH₂O to the control samples. Mix well by pipetting up and down ≥ 20 times.

▲ **CRITICAL STEP** Visually inspect the height of the volume in the pipette strips and ensure that the levels are consistent.

▲ **CRITICAL STEP** We do not recommend using volumes $> 5 \mu\text{l}$, because the addition of solution previously incubated on ice will change the temperature.

▲ **CRITICAL STEP** We do not recommend using volumes $< 5 \mu\text{l}$, to minimize pipetting error effects.

- 7 Incubate the strip at 25 °C for 5 min.

▲ **CRITICAL STEP** Exact timing of the incubation step is crucial. We strongly advise timing the proteolytic reaction precisely to reduce variability. In cases in which the experimental design requires a decrease in proteolytic activity, we recommend adjustment of the E:S ratio over adjustment of incubation time (see Effects of E:S ratio and incubation time and Structural comparison of pure proteins and proteins in complex backgrounds).

- 8 Transfer the strip to the block at 99 °C and incubate for 5 min.

▲ **CRITICAL STEP** Ensure that the lid of the block is tightly closed to avoid evaporation, and refrain from opening the lid during heat inactivation.

- 9 Change the temperature of the block to 4 °C and let the sample cool down.

▲ **CRITICAL STEP** Maintain a heated lid to avoid condensation.

- 10 Add DOC to a final concentration of 5% (wt/vol) by using the 10% (wt/vol) stock solution.

- 11 Transfer to a 96-deep-well plate.

■ **PAUSE POINT** Samples in DOC can be stored at -20°C . We recommend freezing the samples if not all samples from an experimental setup can be processed simultaneously. We also recommend

freezing the samples if their number (including control samples) exceeds 192, because the time required for Steps 3–11 and the following steps can exceed the capacity of a normal working day.

Sample processing for MS ● **Timing** 90–120 min of handling time, overnight incubation, 60–90 min of handling time

- 12 To reduce cysteine residues, add TCEP at a final concentration of 5 mM and incubate at 37 °C for 30 min.
- 13 To alkylate reduced cysteine residues, add IAA at a final concentration of 40 mM and incubate at room temperature for 30 min.
▲ **Critical Step** IAA is light sensitive; both the stock solution and the plate should be protected from light.
- 14 Dilute the samples to 1% (wt/vol) DOC by adding 0.1 M ammonium bicarbonate.
- 15 Add LysC and trypsin at a 1:100 (wt/wt) E:S ratio and incubate at 37 °C overnight with agitation at 800 rpm.
- 16 Stop the digestion by adding formic acid to a final concentration of 2% (vol/vol) and mix by pipetting up and down.
▲ **Critical Step** Verify that the pH of the sample is <3. If this is not the case, the sample needs to be discarded because this indicates an error in preparation.
- 17 The addition of acid releases CO₂, which can cause air bubbles to occur. Handle with care if the 96-well plates are very full.
▲ **Critical Step** We recommend mixing the solution by pipetting up and down several times before transferring the solution to the filter plate.
- 18 The addition of acid will cause precipitation of DOC. DOC can be removed by using FiltrEX 96-well filter plates. Place a filter plate on a fresh 96-deep-well plate to collect the flow-through. Transfer 200 µl of the sample into the plate by using multi-channel pipettes and centrifuge at 800g for 1 min. Repeat until all sample has been filtered.
▲ **Critical Step** Do not fill the wells of the plate with >200 µl. If more volume is added, it can result in cross-contamination with other wells during the centrifugation. During the filtering, precipitated DOC will be retained in the wells, and adding less volume might become necessary to avoid overflow.
▲ **Critical Step** Ensure that all solution has passed through the filter. After the last transfer and centrifugation step, centrifuge at 800g for an additional 1 min.
- 19 ■ **Pause Point** Samples can be stored at –20 °C for ≥6 months.

Peptide cleanup ● **Timing** 2–3 h

Use a 96-well C18 plate to desalt the samples. Choose the resin capacity by following the manufacturer's instruction and according to the total protein amount used in Steps 1 and 2.

- 20 Wash the C18 resin with 200 µl of methanol by using a vacuum manifold.
- 21 Wash the samples three times with desalting buffer A by using a vacuum manifold.
- 22 Load the samples by using a vacuum manifold. Repeat loading until the whole sample volume is loaded.
- 23 Wash the samples three times with desalting buffer A by using a vacuum manifold.
- 24 Elute the samples with 4 × 50 µl of buffer B by centrifugation at 800g for 1 min.
- 25 Evaporate the solvent from the eluted peptides by using a vacuum centrifuge at 45 °C.

■ **Pause Point** Samples can be stored at –20 °C for ≥6 months.

▲ **Critical Step** For samples with high complexity, it can be beneficial to perform fractionation²⁶.

Data acquisition and analysis ● **Timing** 2–3 h of measurement time per sample, 1–2 d for spectral library generation, 2–3 d for targeted data extraction

Data acquisition is performed in two modes. For generation of spectral libraries, the samples are acquired in DDA mode. For label-free quantification, the samples are acquired in DIA mode. Here, we describe label-free quantification using the Software Spectronaut (Biognosys).

▲ **Critical Step** We recommend using standardized peptides for retention time correction.

▲ **Critical Step** We recommend visually inspecting the total ion count during or immediately after acquisition to ensure stable acquisition.

Spectral library

- 24 Resuspend peptides in an appropriate volume of buffer A to a final concentration of 1 mg/ml.
- ▲ **Critical step** For purified proteins or complexes, we recommend adding a heterologous whole-cell digest to reduce the false discovery rate (FDR) and avoid mis-annotations during the database search.
- ▲ **Critical step** We recommend acquiring a spectrum for each replicate of the sample for highest library quality. For time constraints, the replicates of one sample can be pooled, and one spectrum per sample can be obtained instead.
- 25 Perform DDA measurements by using method parameters such as those described in Equipment setup. Inject 1–3 µl of peptide sample and separate with a 2-h water/acetonitrile gradient as described in Equipment setup.
- 26 Search the obtained spectra of the LiP samples and trypsin control samples individually. Perform a database search with a search engine tool by using the following general settings for all searches:
- Up to two missed cleavages
 - Cysteine carbamidomethylation (+57.0214 Da) as fixed modification
 - Methionine oxidation (+15.99492 Da) as variable modification
 - Monoisotopic peptide tolerance of 10 ppm
 - Fragment mass tolerance of 0.02 Da
 - 1% FDR filter on the peptide level
- For the LiP samples, apply a semi-specific tryptic digestion rule; for the TrP samples, apply a specific tryptic digestion rule.
- ▲ **Critical step** We recommend including a list of known contaminants in the search. The list is provided as Supplementary Data 1.
- ▲ **Critical step** For the analysis of proteolytic resistance (see Experimental design, Influence of reaction condition on protease activity), search the LiP and TrP samples together by applying a specific tryptic digestion rule.
- 27 Generate spectral libraries for the LiP samples and the control samples by using the software Spectronaut (Biognosys), with the default setting and matching the tryptic digestion rule of the search settings (Step 25).

Label-free quantification

- 28 Resuspend peptides in an appropriate volume of buffer A to a final concentration of 1 mg/ml.
- 29 Perform DIA measurements by using method parameters such as those described in Equipment setup. Inject 1–3 µl of peptide sample and separate with a 2-h water/acetonitrile gradient as described in Equipment setup.
- 30 Perform targeted data extraction for LiP and TrP samples individually by using the software Spectronaut (Biognosys), with the spectral libraries generated in Step 26 and with the default settings.
- ▲ **Critical step** For the analysis of proteolytic resistance (see Experimental design, Influence of reaction condition on protease activity), use the corresponding combined spectral library.

Data processing, quality control and statistical analysis ● **Timing 1 d**

▲ **Critical** Here, we describe the analysis pipeline for a case-control study, in which protein structural alterations and abundance changes are compared between two conditions. Data processing, quality control and statistical analysis are performed by using the R-base package MSstatsLiP.

We provide RMarkdown analysis notebooks for step-by-step operation instructions for the analysis of case-control studies (LiP-MS_data_analysis_case_control.Rmd, <https://doi.org/10.5281/zenodo.5749994>) and the analysis of multiple-dose studies (LiP_MS_data_analysis_multiple_dose.Rmd, <https://doi.org/10.5281/zenodo.5749994>; see also Supplementary Discussion and Extended Data Fig. 7). We also provide an RMarkdown analysis notebook for data visualization (LiP_MS_data_visualization.Rmd, <https://doi.org/10.5281/zenodo.5749994>).

Data processing

- 31 Export the analyzed data from Spectronaut by using the default MSstats report scheme.
- 32 In R, load the data of the LiP samples (LiP data) and the trypsin control samples (TrP data). Convert the data by using the function SpectronauttoMSstatsLiPFormat. The .fasta file(s) used for the spectral searches (Step 25) are also loaded.

▲ **CRITICAL STEP** See the notebook (LiP_MS_data_visualization.Rmd, 10.5281/zenodo.5749994) for details on proteotypicity filter and feature filtering settings.

▲ **CRITICAL STEP** Ensure that the condition nomenclature (Condition) is identical in both datasets and that the replicate nomenclature (BioReplicate) complies with the experimental setup recognized by MSstats. Adjust if necessary.

33 Summarize the data by using the function `dataSummarizationLiP()`.

▲ **CRITICAL STEP** The cross-run normalization method can be set at this point. The default option is using a median normalization approach.

▲ **CRITICAL STEP** Missing values can be imputed at this point. The default option is set to FALSE. Imputation is performed by using an accelerated failure time model.

Quality control

34 Check that the workflow has been implemented correctly:

- The intensity distribution of each run before and after normalization can be visualized in box plots; it should span the same range for all runs.
- Intensity traces of individual proteins or peptides can be visualized in profile plots.
- The CV in the biological replicates is calculated as the ratio of the standard deviation to the mean for the different fragment ions and visualized in a violin plot. The distribution should be uniform toward lower CVs (see Fig. 6a). High CVs ($CV > 0.25$) can be, for example, a result of inefficient mixing, imprecise timing or problems in handling large numbers of samples.
- The fraction of HT peptides is determined by using the function `calculateTrypticity()`. For estimates of HT contents in different sample types, see Supplementary Table 2 and Extended Data Fig. 5a.

? TROUBLESHOOTING

35 The data exploration step visualizes the distribution of the data and can be used to identify outliers and detect batch effects:

- The intensities of each run are correlated by using Pearson correlation (Fig. 6d). Runs within one group should exhibit a higher correlation value than runs from different conditions. For an overview, the correlation coefficient is visualized in a tile map. We recommend also plotting the individual intensities in a scatter plot to inspect the intensity distribution.
- Hierarchical clustering of the correlation coefficients can help to identify signal drifts or batch effects from other sources. It is visualized in a dendrogram (Fig. 6b).
- Principal component analysis (PCA) of the intensities can also be used to visualize the distribution of the data (Fig. 6c). As for hierarchical clustering, the replicates should cluster according to the condition.

36 Visualize the number of identified peptides and proteins in both the LiP samples and the TrP samples.

▲ **CRITICAL STEP** The number of identified proteins should be similar for both the LiP samples and the control samples across the different conditions. A significant decrease or increase in the number of identified proteins can be indicative of an experimental problem. Similarly, the number of identified peptides in both sample types should be of the same order of magnitude. However, a slight increase in the total number of identified peptides is often observed in the LiP samples as a consequence of the generation of HT peptides. For estimates of identification numbers in different species, see Extended Data Fig. 5.

Statistical analysis

37 Perform the group comparison by using the function `groupComparisonLiP()`. The default setting is a pairwise comparison of all conditions.

Data visualization

38 Visualize structurally altered peptides and differentially abundant proteins.

- A volcano plot can be used to illustrate the distribution of hits. The $-\log_{10}(\text{adjusted } P \text{ value})$ is plotted against the fold change value. In the volcano plot of structurally altered peptides, both fully tryptic and HT peptides are visualized (Fig. 6e).
- The quantitative assessment of the structurally altered peptides and differentially abundant proteins can be visualized in bar plots.
- The position of structurally altered peptides along the protein sequence, as well as all identified but non-altered peptides, can be displayed by using barcode plots.

? TROUBLESHOOTING

Troubleshooting

In Table 1, we describe commonly observed errors and their impact on the identification of structural alterations. We also provide examples of the errors in the supplemental material.

Table 1 | Troubleshooting

Step	Problem	Possible reason	Solution	Impact on analysis
34	High CV	Mixing problems in Step 6: in solutions with high viscosity, PK is not uniformly distributed in the sample (Extended Data Fig. 8a, upper panel), thus resulting in a differential extent of cleavage and consequently, high variation within the replicates ($CV > 0.25$) (Extended Data Fig. 8a, lower panel)	Upon addition of PK, perform the mixing with a higher volume (e.g., 30 μ l). We recommend switching to a 100- μ l pipette and using an empty tip set to 30 μ l. Note: Avoid vortexing	A high variability will impair the statistical analysis, and the high error will result in non-significant values for the identifications
		Evaporation in Step 8: during PK inactivation at 99 °C, the lids of the reaction tubes can snap open and lead to loss of material and/or evaporation. This leads to a global increase in variability, independent of, for example, the viscosity of the solutions (Extended Data Fig. 8b)	Ensure that the reaction tubes are properly closed before transferring them to the heat block. Avoid unnecessary opening and closing of the reaction tubes. Use a heated lid and ensure that the lid is closed tightly on the reaction tubes	
		Timing problems in Steps 5–9: choosing too short of intervals between processing different samples can lead to inconsistent timing between the samples and therefore a different digestion extent, if the replicates are distributed across multiple reaction tube strips. This in turn leads to different peptide intensities and higher CVs.	Increase the time intervals between processing different samples to ensure enough time for the precise inactivation of PK. Keeping all replicates of a sample on one strip can also help to reduce variability from imprecise inactivation timing	
34	Alterations in HT content	Insufficient inactivation of PK in Step 8 (Extended Data Fig. 8c)	Ensure precise timing and correct temperature (>99 °C) at the inactivation step	Identification of structural alterations might be hindered
34	Replicates do not cluster; experimental groups do not separate	High variability: poor distribution can be a consequence of high variability (see High CV, above)	See above (High CV)	See above (High CV)
		Subtle alterations: in perturbation experiments, as in Step 5, B and C, the effect caused by the perturbation might be in the range of the variability of some features and therefore too subtle to drive separation at the global level	Orthogonal measures for identification of positive hits can improve the outcome (e.g., perturbation range instead of single perturbation; see ref. ⁴ , Supplementary Discussion and Extended Data Fig. 7)	Structural alterations can still be identified, but additional (orthogonal) analyses might be needed to validate the structural alterations
		Spike-in: in a spike-in experiment, the separation of experimental groups is driven by the signal of the spike-in protein. Log ₂ -transformation of the peptide intensity values minimizes their range and diminishes the contribution of the spike-in signal (Extended Data Fig. 8d, left panel)	For visualization, use non-transformed values (Extended Data Fig. 8d, right panel)	Will not impair the identification of structural alterations of the spike-in protein
38	Unidirectional fold changes	Differences in PK activity across experimental conditions will result in a predominantly unidirectional change of the fully tryptic peptides in one direction and HT peptides in the opposite direction (Extended Data Fig. 8e)	Test experimental conditions for effects on PK activity (see Experimental design, Influence of reaction condition on protease activity)	Identification of structural alterations might be hindered

Anticipated results

Here, we describe the characterization of a structurally altered protein that was identified in the teaching data set (see Experimental design, Expertise). The protein is the target of rapamycin in *S. cerevisiae*, FK506-binding protein 1 (FPR1). 11 peptides from this protein were identified in the experiment, with two peptides exceeding the significance cutoffs (Fig. 7a). The two structurally altered peptides are HT peptides, which map to the regions 13–24 (peptide sequence RISPGDGATFPK) and 25–32 (peptide sequence TGDLVTIH); the non-tryptic ends of the peptides are at positions 13 and 32, respectively (Fig. 7b). The second cleavage site resides within the N-terminal region of the peptidyl-proline isomerase domain, and the significantly altered peptide TGDLVTIH is in close spatial proximity to the rapamycin binding side (6.6 Å) (Fig. 7c).

Data availability

The mass spectrometry proteomics data presented in Fig. 8 and Extended Data Fig. 7 and 9 have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository²⁷ with the dataset identifier [PXD031616](#) and [PXD031627](#) (<http://www.proteomexchange.org/>). The mass spectrometry proteomics data presented in Fig. 6, Fig. 7 and Extended Data Fig. 8 were generated as part of ref.⁴. The mass spectrometry data presented in Fig. 9 were generated as part of ref.⁵. The corresponding spectral libraries, Spectronaut Reports and statistical source data presented in Figs. 6–9 and Extended Data Figs. 7 and 9 are available at <https://doi.org/10.5281/zenodo.5749994>.

Code availability

The R Markdown notebooks for data analysis are available at <https://doi.org/10.5281/zenodo.5749994>. The MSstatsLiP R-package can be installed from Bioconductor (<https://www.bioconductor.org/packages/release/bioc/html/MSstats.html>).

References

1. Feng, Y. et al. Global analysis of protein structural changes in complex proteomes. *Nat. Biotechnol.* **32**, 1036–1044 (2014).
2. Leuenberger, P. et al. Cell-wide analysis of protein thermal unfolding reveals determinants of thermostability. *Science* **355**, eaai7825 (2017).
3. Piazza, I. et al. A map of protein-metabolite interactions reveals principles of chemical communication. *Cell* **172**, 358–372.e23 (2018).
4. Piazza, I. et al. A machine learning-based chemoproteomic approach to identify drug targets and binding sites in complex proteomes. *Nat. Commun.* **11**, 4200 (2020).
5. Cappelletti, V. et al. Dynamic 3D proteomes reveal protein functional alterations at high resolution in situ. *Cell* **184**, 545–559.e22 (2021).
6. Schopper, S. et al. Measuring protein structural changes on a proteome-wide scale using limited proteolysis-coupled mass spectrometry. *Nat. Protoc.* **12**, 2391–2410 (2017).
7. Mackmull, M.-T. et al. Global, in situ analysis of the structural proteome in individuals with Parkinson’s disease to identify a new class of biomarker. *Nat. Struct. Mol. Biol.* **29**, 978–989 (2022).
8. Leitner, A. et al. Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol. Cell. Proteom.* **9**, 1634–1649 (2010).
9. Rappaport, J. The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J. Struct. Biol.* **173**, 530–540 (2011).
10. Sinz, A. Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. *Mass Spectrom. Rev.* **25**, 663–682 (2006).
11. Kiselar, J. G. & Chance, M. R. Future directions of structural mass spectrometry using hydroxyl radical footprinting. *J. Mass Spectrom.* **45**, 1373–1382 (2010).
12. Maleknia, S. D. & Downard, K. M. Advances in radical probe mass spectrometry for protein footprinting in chemical biology applications. *Chem. Soc. Rev.* **43**, 3244–3258 (2014).
13. Espino, J. A. & Jones, L. M. Illuminating biological interactions with in vivo protein footprinting. *Anal. Chem.* **91**, 6577–6584 (2019).
14. Rinas, A., Espino, J. A. & Jones, L. M. An efficient quantitation strategy for hydroxyl radical-mediated protein footprinting using Proteome Discoverer. *Anal. Bioanal. Chem.* **408**, 3021–3031 (2016).
15. Liu, F., Rijkers, D. T. S., Post, H. & Heck, A. J. R. Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. *Nat. Methods* **12**, 1179–1184 (2015).
16. Savitski, M. M. et al. Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science* **346**, 1255784 (2014).
17. Tan, C. S. H. et al. Thermal proximity coaggregation for system-wide profiling of protein complex dynamics in cells. *Science* **359**, 1170–1177 (2018).
18. Potel, C. M. et al. Impact of phosphorylation on thermal stability of proteins. *Nat. Methods* **18**, 757–759 (2021).
19. Mateus, A. et al. Thermal proteome profiling for interrogating protein interactions. *Mol. Syst. Biol.* **16**, e9232 (2020).

20. Ludwig, C. et al. Data-independent acquisition-based SWATH- MS for quantitative proteomics: a tutorial. *Mol. Syst. Biol.* **14**, e8126 (2018).
21. Clough, T., Thaminy, S., Ragg, S., Aebersold, R. & Vitek, O. Statistical protein quantification and significance analysis in label-free LC-MS experiments with complex designs. *BMC Bioinforma.* **13**, 1–17 (2012).
22. Tsai, T. & Vitek, O. Introduction to MSstatsPTM. Available at <https://rdrr.io/bioc/MSstatsPTM/f/vignettes/MSstatsPTM.Rmd> (2020).
23. Moulder, R., Goo, Y. A. & Goodlett, D. R. Label-free quantitation for clinical proteomics. *Methods Mol. Biol.* **1410**, 65–76 (2016).
24. Richards, A. L. et al. One-hour proteome analysis in yeast. *Nat. Protoc.* **10**, 701–714 (2015).
25. Feng, Y. & Picotti, P. Selected reaction monitoring to measure proteins of interest in complex samples: a practical guide. *Methods Mol. Biol.* **1394**, 43–56 (2016).
26. Rappaport, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
27. Perez-Riverol, Y. et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450 (2019).
28. Murtagh, F. & Legendre, P. Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion?. *J. Classif.* **31**, 274–295 (2014).
29. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
30. Kohler, D. et al. Bioconductor - MSstatsPTM. Available at <https://www.bioconductor.org/packages/release/bioc/html/MSstatsPTM.html> (2021).
31. Choi, M. et al. MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics* **30**, 2524–2526 (2014).
32. Ueda, K. et al. Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease. *Proc. Natl Acad. Sci. USA* **90**, 11282–11286 (1993).

Acknowledgements

We thank the Picotti laboratory for continuous input on the protocol optimization and the R package. I.P. and L.M. were supported by long-term EMBO postdoctoral fellowships (ALTF 846-2014 and ALTF 538-2016). This project received funding from the European Research Council (grant agreement no. 866004) and through the EPIC-XS Consortium (grant agreement no. 823839) both under the European Union's Horizon 2020 research and innovation program. It was supported by a Personalized Health and Related Technologies (PHRT) grant (PHRT-506), a Sinergia grant from the Swiss National Science Foundation (SNSF grant CRSII5_177195) and the National Center of Competence in Research AntiResist funded by the SNSF (grant number 51NF40_180541).

Author contributions

L.M., V.C. and N.d.S. wrote the paper. L.M., V.C. and I.P. optimized the original version of the protocol. M.P., P.S., C.D., F.E., L.K., N.B. and L.R. contributed to protocol optimization. L.M., V.C., D.K., T.-H.T. and F.S. performed bioinformatic analysis of the data. D.K. and T.-H.T. optimized algorithms and tools for the R package MSstatsLiP. O.V. supervised the development of the R package. P.P. supervised the optimization of the protocol.

Competing interests

N.B. and L.R. are employees of Biognosys AG (Zurich, Switzerland). P.P. is a scientific advisor for the company Biognosys AG (Zurich, Switzerland) and an inventor of a patent licensed by Biognosys AG that covers the LiP-MS method used in this protocol. The remaining authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41596-022-00771-x>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41596-022-00771-x>.

Correspondence and requests for materials should be addressed to Olga Vitek or Paola Picotti.

Peer review information *Nature Protocols* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

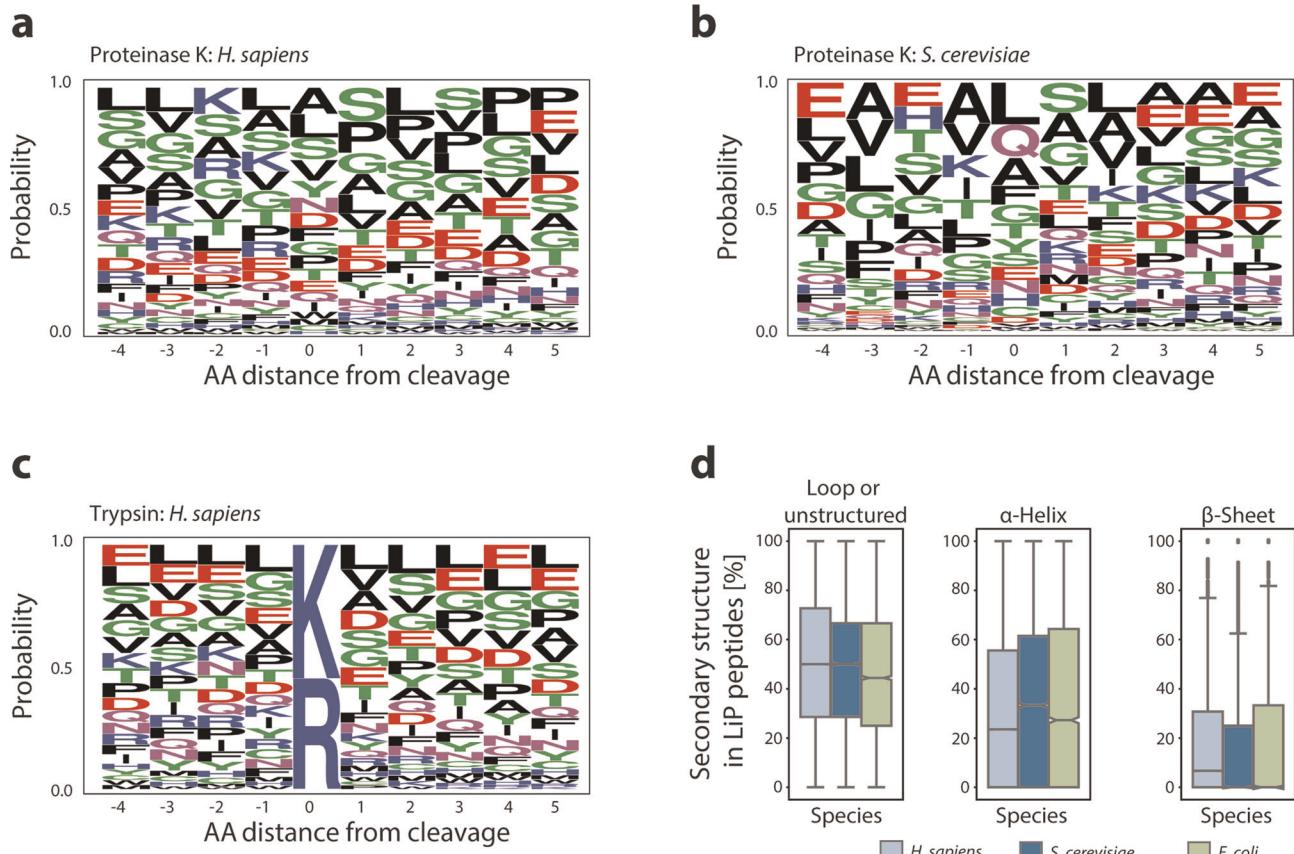
Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Received: 5 October 2021; Accepted: 8 July 2022;
Published online: 16 December 2022

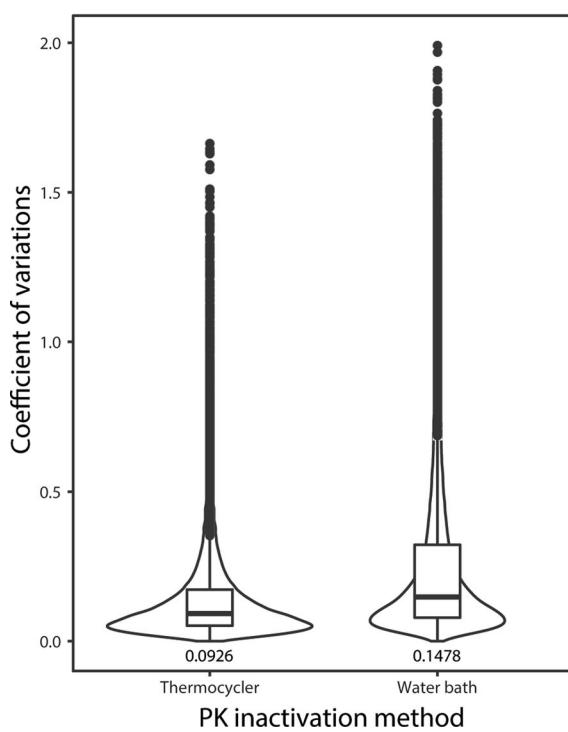
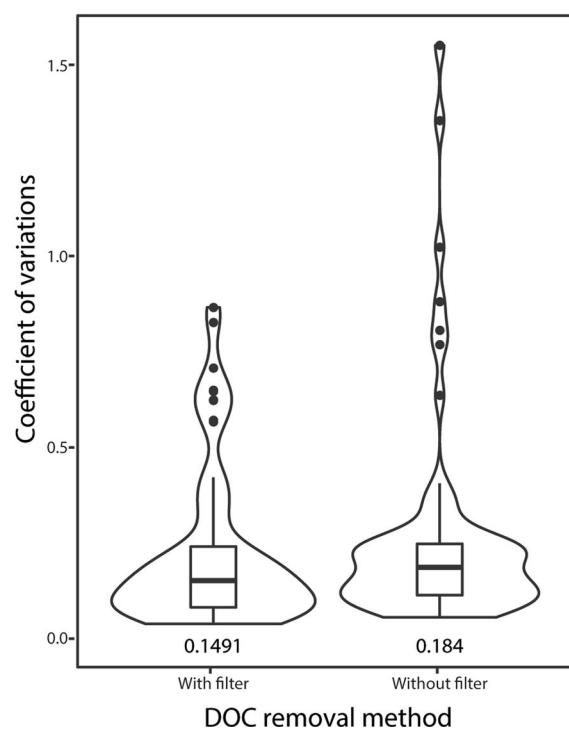
Related Links

Key references using this protocol

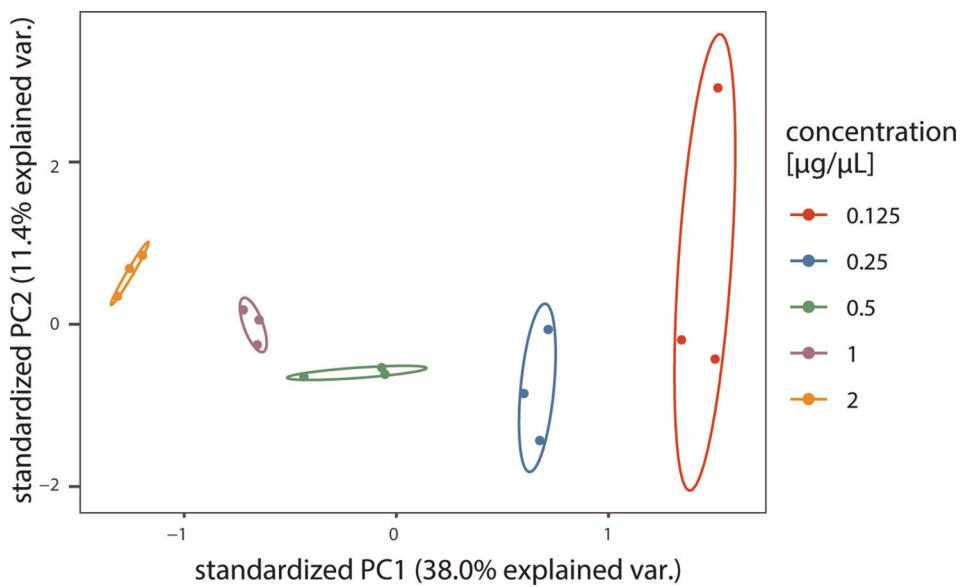
Piazza, I. et al. *Nat. Commun.* **11**, 4200 (2020): <https://doi.org/10.1038/s41467-020-18071-x>
Cappelletti, V. et al. *Cell* **184**, 545–559.e22 (2021): <https://doi.org/10.1016/j.cell.2020.12.021>



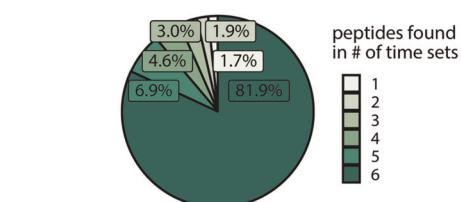
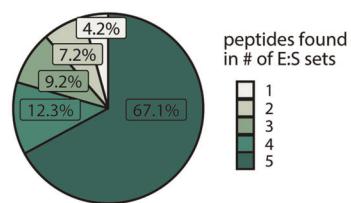
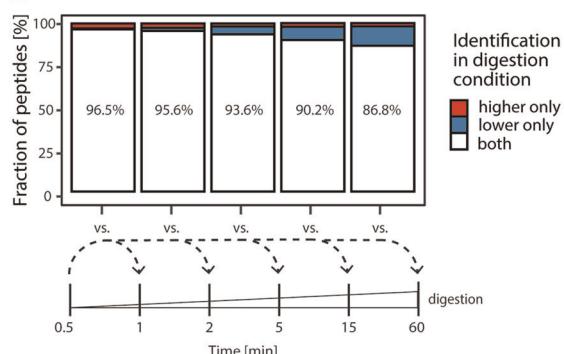
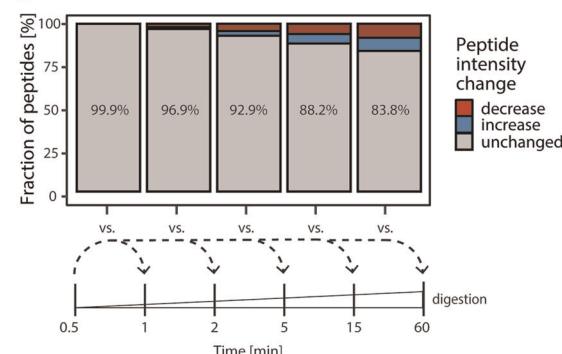
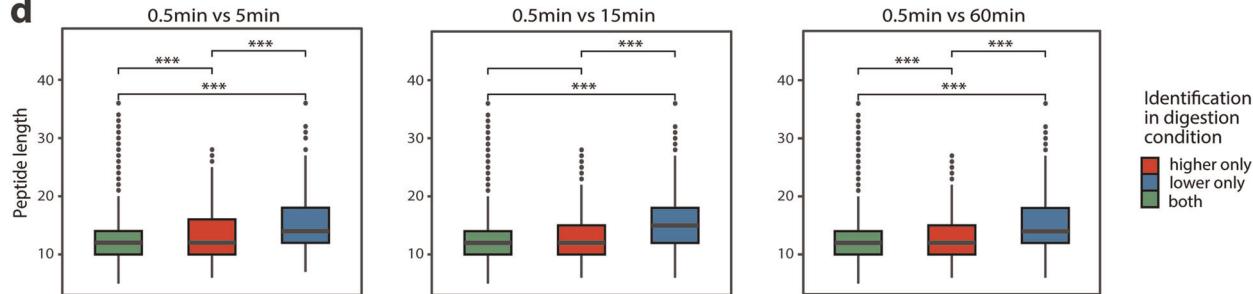
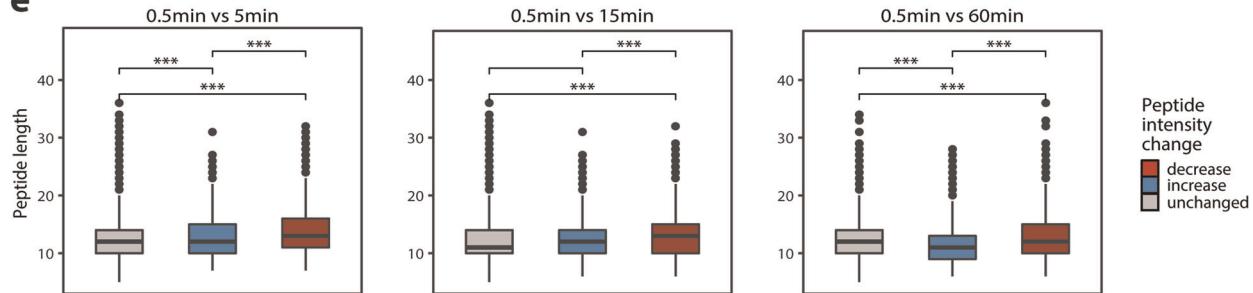
Extended Data Fig. 1 | Analysis of cleavage preference of PK. **a** and **b**, WebLogo diagram of the cleavage site of PK and the neighboring five amino acids in a mammalian cell lysate (**a**) and yeast cell lysate (**b**). Amino acids are plotted according to the frequency of occurrence at the cleavage side and colored on the basis of their physicochemical properties. **c**, WebLogo diagram of the cleavage site of a specific protease (trypsin) in a mammalian cell extract. **d**, Boxplots showing the distribution of secondary structure elements in half-trypic peptides in different organisms. Box shows the quartiles of the dataset, median values are represented by a vertical line in the center of the box, bars extend to the rest of the distribution and dots represent outliers. AA, amino acid.

a**b**

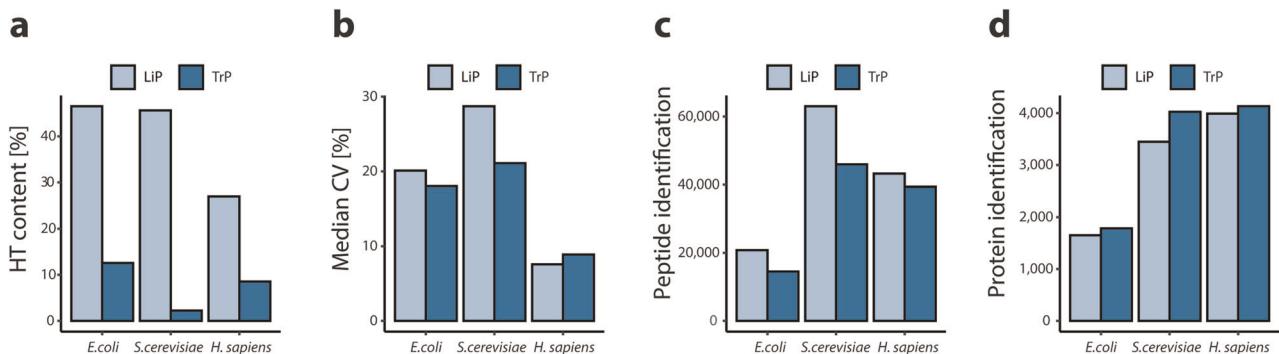
Extended Data Fig. 2 | Improvement of reproducibility and robustness of a LiP-MS experiment. **a**, Coefficient of variation of a LiP experiment performed in a thermocycler as described in this protocol (new) and performed in a boiling pot as described in Schopper et al.⁶ (old). The median value is displayed below the plot. **b**, Coefficient of variation of a LiP experiment with DOC removal through centrifugation in individual reaction tubes as described in Schopper et al.⁶ (old) and with DOC removal through filtration (new). The median value is displayed below the plot.



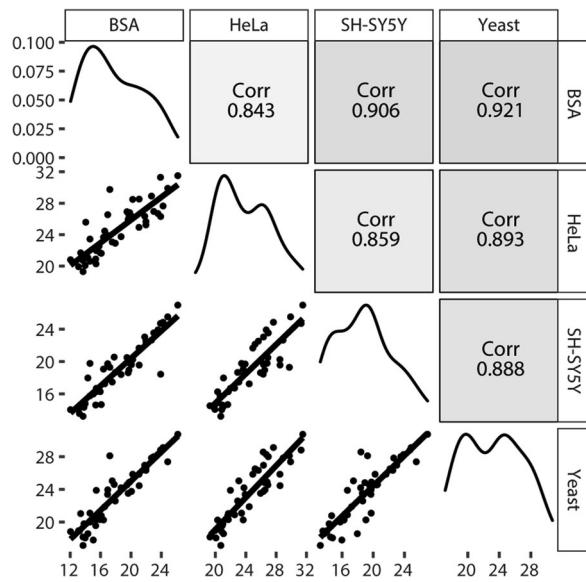
Extended Data Fig. 3 | Comparison of PK cleavage pattern at different sample concentrations. Principal component (PC) analysis of MS1 features. The colors indicate the different samples amounts in $50 \mu\text{l}$.

a Incubation time**Enzyme : substrate ratio****b****c****d****e**

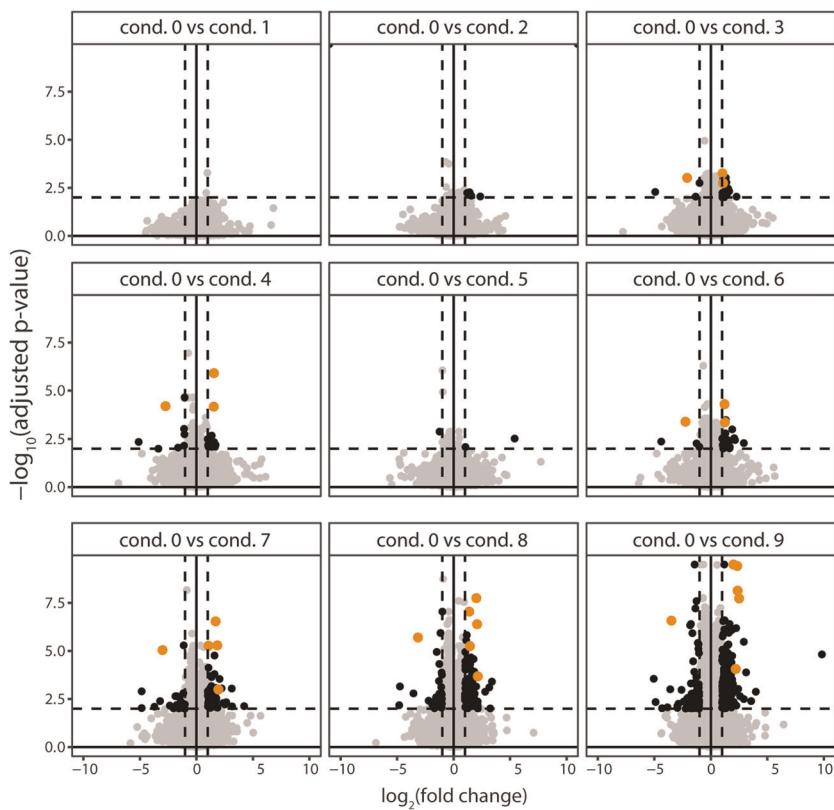
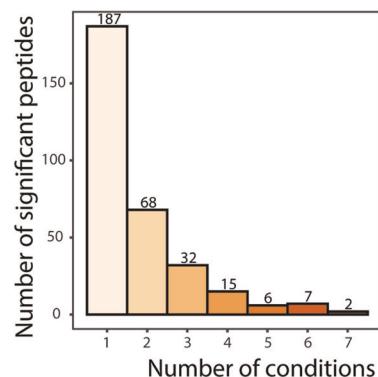
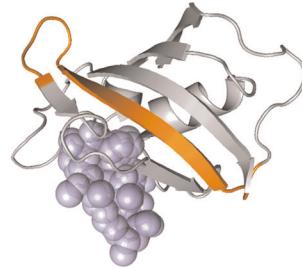
Extended Data Fig. 4 | Influence of E:S ratio and incubation time. **a**, Summary characterization of peptides in LiP-MS experiments performed as described in Fig. 2a. Left pie chart: Number of peptides identified across the different incubation times. Right pie chart: Number of peptides identified across the different E:S ratios. **b** and **c**, Pairwise comparisons of peptide identifications and intensity changes for each incubation time (1–60 min) versus the shortest incubation time (30 s). **b**, Fraction of peptides identified only in the higher digestion condition (red), only in the lower digestion condition (blue) or shared between both conditions (white). **c**, Fraction of peptides with significantly changed intensities ($|\log_2(\text{fold change})| > 2$, $\text{q-value} < 0.05$); increase indicated in blue, decrease indicated in red and no change indicated in gray. **d** and **e**, Distribution of peptide length (number of amino acids) in groups defined in **b** and **c**. Boxplot of peptide length in amino acids, colored by groups. **d**, Peptides identified only in the higher digestion condition (red), only in the lower digestion condition (blue) or shared between both conditions (green). **e**, Peptides with increased intensity (blue), decreased intensity (red) or no change (gray). A significant difference in mean peptide length is indicated by asterisks ($P < 0.05$, Wilcoxon test).



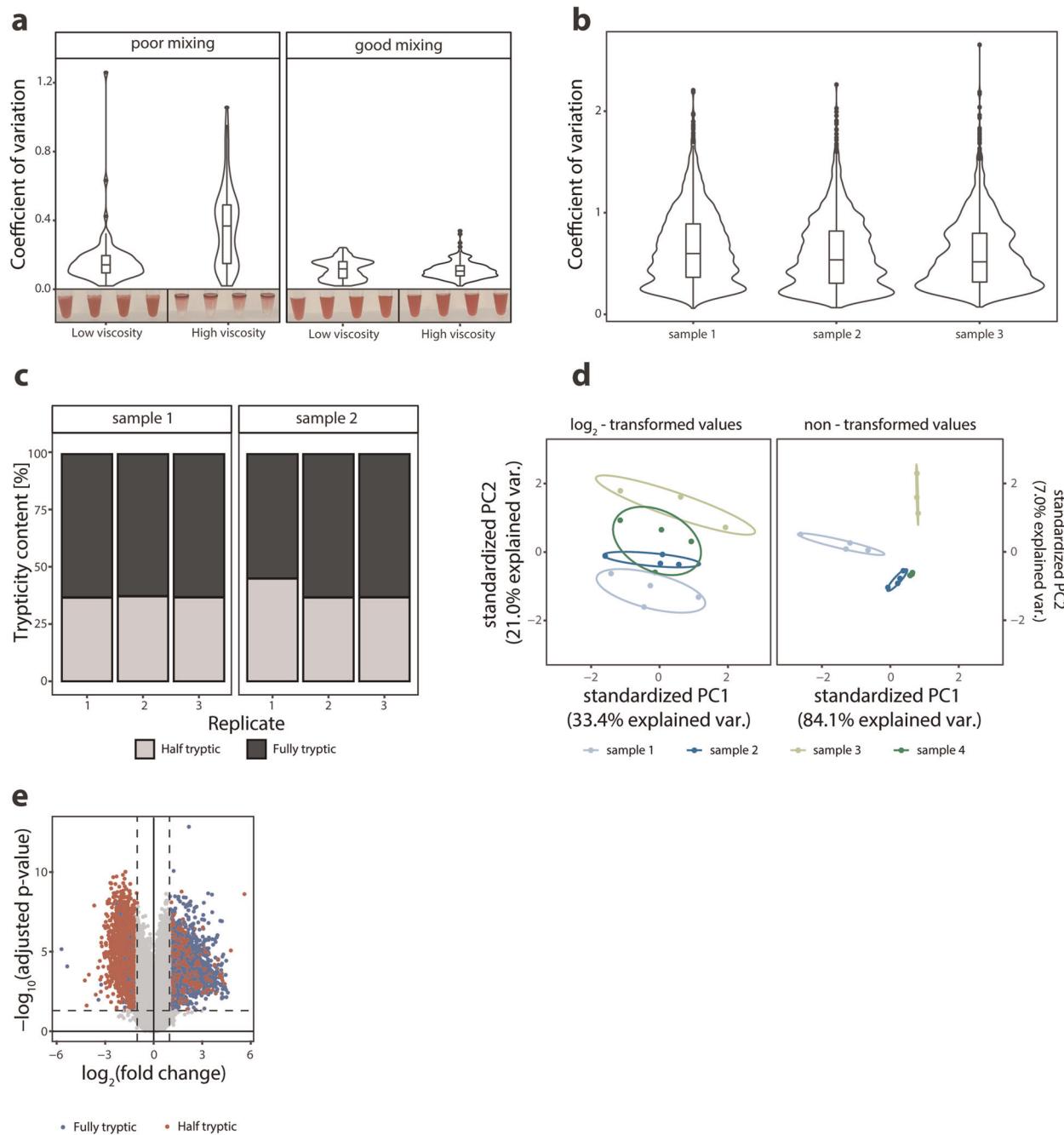
Extended Data Fig. 5 | Experimental characteristics in different organisms. Quantification in three different species: *E. coli*, *S. cerevisiae* and human (*Homo sapiens*) cell lines, both in LiP samples (light blue) and TrP samples (dark blue). **a**, HT content (percentage). **b**, Median CV (percentage). **c**, Number of peptide identifications. **d**, Number of protein identifications.



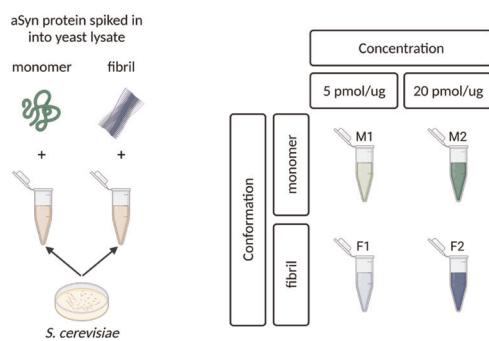
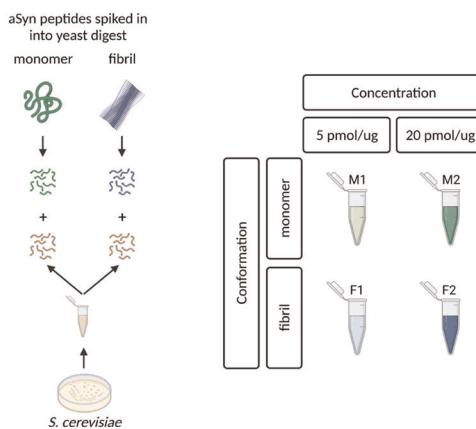
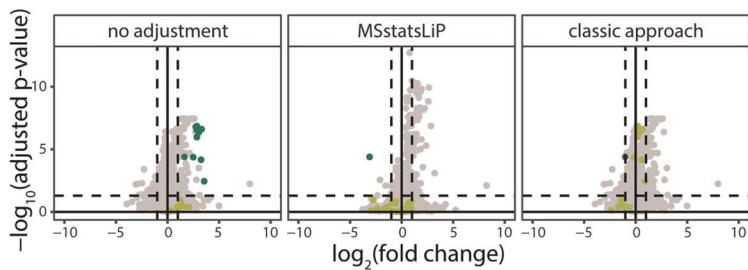
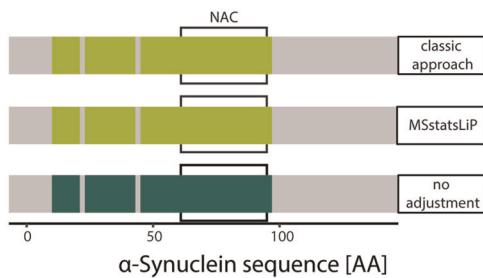
Extended Data Fig. 6 | Comparison of LiP peptide intensities in different proteinaceous backgrounds. Correlation of peptide intensities of monomeric α -synuclein spiked into 1.8 mg/ml BSA, whole-cell lysate of HeLa cells, SH-SY5Y cells and yeast *S. cerevisiae* cells. Pearson correlation coefficient of log₂-transformed fragment ion intensities visualized as colored squares (upper triangle) with values ranging from 0.983 (gray) to 0.968 (white). Scatter plot of log₂-transformed fragment ion intensity correlation (lower triangle) and data distribution density plot (diagonal).

a**b****c**

Extended Data Fig. 7 | Commonly observed errors or issues. **a**, CV (ratio of the standard deviation to the mean) of peptide intensity for biological replicates in samples with varying viscosity. Mixing efficiency is visualized by the addition of red color to the samples. In solutions with high viscosity, poor mixing can lead to non-uniform distribution of PK. Poor distribution of PK results in differential extents of cleavage and thus high variability (median CV > 0.25). Proper mixing can counteract this effect. **b**, High CV (median CV > 0.25) of peptide intensity for biological replicates in three samples. Evaporation can occur during PK inactivation as a result of opening lids. This leads to a global increase in variability in all samples of an experiment. **c**, Distribution of HT peptides (gray) and fully tryptic peptides (black) in two samples. Insufficient inactivation of PK can lead to variations in the proportions of HT peptides as reflected by the inconsistent trypticity content among replicates in sample 2. In this case, replicate 1 of sample 2 cannot be used for the statistical analysis. **d**, PCA of fragment ion intensities in four samples. Spike-in of single proteins at high concentrations into a complex background can display poor separation of experimental groups when using \log_2 -transformed data (left panel). Using non-transformed data emphasizes the separation effect of the spike-in protein (right panel). **e**, Differential analysis of structurally altered peptides, visualized in a volcano plot. Each point represents a peptide. Peptides passing the significance cutoff ($|\log_2(\text{FC})| > 1$, $\text{q-value} < 0.05$) are colored by trypticity (blue: fully tryptic; red: HT). Impaired PK activity in one of the conditions results in predominantly unidirectional fold changes of peptide classes (upregulation of almost all fully tryptic peptides and downregulation of mostly semi-tryptic peptides).



Extended Data Fig. 8 | Multiple dose analysis of yeast cells treated with rapamycin at nine different concentrations. **a**, Volcano plots of altered LiP peptide intensities generated with MSstatsLiP. Significantly altered peptides ($|\log_2\text{FC}| > 1$, $\text{q-value} < 0.01$) are indicated in black. Significantly changing peptides mapping to the direct target of rapamycin (FPR1) are colored in orange. **b**, Bar plot showing in how many conditions a significant FPR1 peptide was found. Only two LiP peptides were found in six conditions: LGLSNEDFFHK mapping to Sac1p and TGDLVTIHYTGTLENGQK mapping to Fpr1p, the known target of rapamycin. **c**, The Fpr1p LiP peptide found altered in six conditions is mapped onto the 3D protein structure complexed with rapamycin obtained by homology modeling (PDB: 2DG3). Rapamycin bound to target protein is indicated in violet. Orange indicates the TGDLVTIHYTGTLENGQK LiP peptide.

a**b****c****d**

Extended Data Fig. 9 | Benchmarking abundance correction approaches for LiP-MS. **a**, Schematic of the experimental setup. Spike-in of intact proteins. The M and F conformation of α -synuclein (α syn) are spiked into yeast lysate at two different concentrations (5 and 20 pmol/ μ g lysate). **b**, Schematic of the experimental setup. Spike-in of peptides. Peptides obtained from the digestion of M and F α syn are spiked into a yeast lysate digest at the same concentrations as in **a**. **c**, Differential analysis of LiP peptides in the comparison M1 versus M2 (i.e., two different concentrations of monomer) described in **b**. Each dot represents an individual peptide, and green indicates peptides of α syn. Dark green, peptides passing the significance cutoff ($|\log_2(\text{FC})| > 2$, $q\text{-value} < 0.05$); light green, non-significant peptides. **d**, Structural barcodes visualizing fully tryptic peptides mapped along the sequence of α syn in the M1 versus M2 comparison described in **c** and corrected for protein abundance as indicated. Significant peptides are colored in dark green, and non-significant peptides are colored in light green. Gray regions indicate no identified fully tryptic peptide matching this region. The position of the NAC region is indicated as a line below the barcode.