nature portfolio

https://doi.org/10.1038/s41596-022-00771-x

Supplementary information

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

In the format provided by the authors and unedited

Supplementary Discussion

Benchmarking abundance correction approaches for LiP-MS

To identify structurally altered peptides, LiP-MS data are corrected using the tryptic controls to distinguish between protein abundance and structural contributions to the detected change. We adapted the correction method described in Tsai and Vitek (2020) to estimate peptide and protein abundances in the LiP and TrP datasets and combine them to adjust LiP peptide changes with respect to protein abundance variations. To test the performance of this correction method over the one described in Schopper et al. (2017), we generated a combinatorial dataset where the same protein, human α-Synuclein, was spiked-in into a yeast background at two different concentrations (5 pmol/µg lysate and 20 pmol/µg lysate) and two different conformations (monomer, M, and fibril, F) (Extended Data Figure 9a). The resulting data were analyzed following the classic workflow and with the MSstatsLiP package and the results of the two correction approaches were compared. Overall, both methods performed well in term of protein abundance correction but when we compared the monomer form at two different concentrations, where no structural rearrangements are expected, both analyses identified structurally altered LiP peptides. To investigate the nature of the structural alterations detected when comparing monomeric α-Synuclein, we repeated the LiP experiment on pure preparations of α-Synuclein, mixing the resulting peptides (instead of the proteins) with a yeast digest right before the MS analysis (Extended Data Figure 9b). We then repeated the same comparative analysis and focused on the comparison between monomeric α-Synuclein at two different concentrations. Interestingly, all the significant peptides identified in the non-adjusted dataset disappeared after correction, except for one HT peptide identified with the classic approach and one with MSstatsLiP (Extended Data Figure 9c) indicating the structural alteration detected in the benchmark dataset was a consequence of the interaction of monomeric α-Synuclein with the yeast background. When restricting the analysis to only fully tryptic peptides, none of them were significantly affected in the M1 vs M2 comparison, in contrast to the unadjusted data where the structurally affected region covered a significant portion of the protein, including the NAC domain (Ueda et al., 1993) (Extended Data Figure 9d).

Titration analysis with MSstatsLiP

MSstatsLiP offers the possibility to analyze data sets with multiple data points by adapting the linear mixed effect model to account for between-subject heterogeneity over different data points. Here we test the potential benefit of employing such a multiple data points model for a repeated measures design (e.g. where a sample is treated with multiple drug doses), since multiple-dose analysis outperforms single-dose experiments for the identification of drug targets (Piazza et al., (2020). We applied MSstatsLiP to a drug dose titration experiment where Rapamycin was added to a yeast lysate at nine different concentrations (Piazza et al., 2020). When analysing every drug concentration as an independent experiment, we identified two structurally altered peptides in the known target of rapamycin (FPR1) but only at the highest concentration (significance cutoffs: abs(log2FC)>1, qvalue<0.01, Fig. 7a). On the contrary, when we performed the multiple dose analysis including all nine rapamycin concentrations (See LiP.MS data analysis multiple dose study notebook) we identified two peptides as significantly changing in the highest number of conditions (Extended Data Figure 7a,b), and one of the two mapped to the binding site of rapamycin in the target FPR1 (Extended Data Figure 7c).

Influence of total protein concentration on proteinase K cleavage patterns

Some sample types – for example, clinical samples – are limited in protein amount and it is thus not possible to adjust all samples to a standard concentration (step 1). In these cases, total protein concentration may vary between samples, which could affect PK cleavage patterns.

To probe the effect of varying total protein concentration, we analyzed a cell lysate with protein concentrations ranging from 0.125 to 2 µg/µl using an E:S ratio of 1:100. The total protein concentration of each sample has a pronounced impact on the observed cleavage pattern, which is apparent in a principal component analysis (Extended Data Figure 3). Since this effect reduces comparability of samples with different total protein concentrations, we recommend diluting all samples to the concentration of the most dilute sample (see step 1).

Supplementary Tables

Table 1: DIA isolation window settings

Lower limit	Center	Upper limit	Isolation window width
350	358	366	16
365	373	381	16
380	388	396	16
395	403	411	16
410	418	426	16
425	433	441	16
440	448	456	16
455	463	471	16
470	478	486	16
485	493	501	16
500	508	516	16
515	523	531	16
530	538	546	16
545	553	561	16
560	568	576	16
575	583	591	16
590	598	606	16
605	613	621	16
620	628	636	16
635	643	651	16
650	659	668	18
667	676	685	18
684	693	702	18
701	710	719	18
718	727	736	18
735	744	753	18
752	761	770	18
769	778	787	18
786	795	804	18
803	813	823	20

822	832	842	20
841	851	861	20
860	870	880	20
879	889	899	20
898	908	918	20
917	929.5	942	25
941	953.5	966	25
965	977.5	990	25
989	1006.5	1024	35
1023	1048	1073	50
1072	1111	1150	78

Table 2: Expected HT content in different sample types

Sample type	**Sample**	**Expected HT content**
complex, lysate	*S. cerevisiae*	~30%
complex, lysate	*E. coli* lysate	~40%
complex, lysate	Mammalian cell lysate	~30%
complex, lysate	Brain homogenate	~30%
complex, fluid	Red blood cell fraction	~45%