Efficient Tandem LysC/Trypsin Digestion in Detergent Conditions

Anna Hakobyan, Martin Bernd Schneider, Werner Liesack, and Timo Glatter*

All shotgun proteomics experiments rely on efficient proteolysis steps for sensitive peptide/protein identification and quantification. Previous reports suggest that the sequential tandem LysC/trypsin digest yields higher recovery of fully tryptic peptides than single-tryptic proteolysis. Based on the previous studies, it is assumed that the advantageous effect of tandem proteolysis requires a high sample denaturation state for the initial LysC digest. Therefore, to date, all systematic assessments of LysC/trypsin proteolysis are done in chaotropic environments such as urea. Here, sole trypsin is compared with LysC/trypsin and it is shown that tandem digestion can be carried with high efficiency in Mass Spectrometry-compatible detergents, thereby resulting in higher quantitative yields of fully cleaved peptides. It is further demonstrated that higher cleavage efficiency of tandem digests has a positive impact on absolute protein quantification using intensity-based absolute quantification (iBAQ) values. The results of the examination of divergent urea tandem conditions imply that beneficial effects of the initial LysC digest do not depend on the sample denaturation state, but, are mainly caused by different target specificities of LysC and trypsin. The observed detergent compatibility enables tandem digestion schemes to be implemented in efficient cellular solubilization proteomics procedures without the need for buffer exchange to chaotropic environments.

The endopeptidase trypsin is the most widely used protease in shotgun MS experiments.[1–4] This is due to the high efficiency and specificity of trypsin in cleaving the peptide-bonds on the C-terminal side of lysine and arginine residues, thereby generating peptides with suitable properties for MS analysis. Efficient tryptic cleavage is characterized by the generation of the maximum number of most intense fully cleaved peptides (FCPs), while keeping the miscleavage events minimal. One suggested strategy to improve digestion efficiency is to include a lysine-specific endopeptidase in a sequential or tandem digest together with trypsin to obtain a tandem proteolysis workflow (further referred to as LT).[1,5–7] Due to its higher stability than trypsin, LysC is utilized in LT strategies as the first protease in a high denaturation environment, followed by trypsin digest in diluted denaturant concentrations.[1,6,7] Urea is often used to create such differential denaturation state across the LT digestion procedure, which can result in superior cleavage efficiency compared to sole trypsinization.[1] However, it is well known that the use of urea may have negative side effects due to an increased protein carbamylation at higher temperatures.[8] Furthermore, urea creates biases when extracting proteins from intact cells, especially in the recovery of membrane proteins,[9] and when heat exposure is a critical parameter for protein extraction.[10] Alternatively, different detergents are utilized in in-solution digestion (ISD) schemes with less solubilization bias[9] and protein-modifying properties than urea. However, the assessment of LT digests in detergent buffers using state-of-the-art quantitative proteomics approaches is still missing.

The general advantage of detergents in one-step ISD routines prompted us to compare the efficiency of tandem LT and single-trypsin digest in MS-compatible detergent buffer systems. We

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/pmic.201900136

DOI: 10.1002/pmic.201900136
decided to focus our comparative study on the MS-compatible and budget-friendly products, like sodium deoxycholate (DOC) and sodium lauroyl sarcosinate (SLS). While DOC is broadly used for cellular proteomics workflows, SLS is not frequently applied. Recent studies, however, have shown that SLS can be efficiently implemented in the ISD strategies of comparative proteomics studies. 

Our research was strictly focused on the digestion effect, with the aim to minimize solubilization-mediated differences between the detergents. Therefore, we first extracted proteins from frozen *Escherichia coli* cell pellets using SLS, followed by acetonitrile protein precipitation and reconstitution of the protein pellets in the detergent buffers. Then protein digestion in LT and single-trypsin workflows (for more details see Supporting Information) was carried out in constant detergent concentration of 0.5% (Figure 1A); thereby anticipating that potential differences between LT and single-trypsin digest could not be the results of a differential denaturation state of the sample as assumed for urea buffers. In a first attempt to investigate the differences between LT and single-trypsin digest in detergent buffers, we extracted total spectrum counts from the data set and grouped the identified spectra into fully and miscleaved (Figure 1B and Table S1, Supporting Information). As expected, the number of fully cleaved peptide (FCP) spectra were the most prominently identified. They exceeded the number of miscleaved peptide (MCP) spectrum counts by at least 3.5-fold. Both LT workflows identified more FCP spectra than single-trypsin digestion. Thus, in both LT workflows, the MCP counts concurrently decreased with the increase in FCP spectra. Nonetheless, we did not observe a major influence of the digestion scheme on the protein identification rate in the moderately complex *E. coli* lysate (Figure S1, Supporting Information).

In order to compare the workflow efficiencies in greater detail, we performed label-free quantification (LFQ) of the different digestion strategies. Out of a total of 9618 FCPs quantified in DOC (Table S2, Supporting Information) and 9025 FCPs quantified in SLS (Table S3, Supporting Information), FCP levels were consistently higher in all LT strategies than in single-trypsin workflows (Figure 1C,D). In contrast to SLS-trypsin, 1298 FCPs were significantly enriched in SLS-LT (Table S3, Supporting Information). This particular enrichment was the most striking difference between LT and single-trypsin digests and corresponded to 14.4% of all quantified FCPs. On the contrary, the MCPs (n = 2462 for DOC and n = 3113 for SLS) showed reversed asymmetric distribution, with MCPs being more abundantly represented after single-trypsin digest (Figure 1D).

Next, we assessed how the digestion strategies affect quantification reproducibility and variation. The box plots in Figure 2A,B illustrate median FCP-coefficient of variation (CV) ranging from 8.9% to 13%, which are consistently lower than the CVs determined for MCPs (10.7–18%). In SLS-LT, CVs were lower than in the corresponding single-trypsin workflow. In DOC, however, we noted that LT exhibited a higher signal variation than single trypsin. This may be caused by the influence of DOC on LysC activity, for which critical digestion parameters are less studied than for trypsin.

To further monitor the cleavage efficiency of the proteolytic workflows, we identified all MCPs with a corresponding FCP counterpart and calculated the LFQ intensity ratio of the FCP/MCP pairs (Figure 2C,D). The distribution of the FCP/MCP ratios showed that the FCP/MCP LT median in both DOC and SLS is shifted to higher values than for single-trypsin digestion.

Based on our observations, we hypothesize that a high FCP rate has an effect on absolute protein quantification. Accordingly, absolute quantification strategies that exclusively incorporate FCP intensities to calculate protein quantities, such as the intensity-based absolute quantification index (iBAQ), may get compressed when a digestion strategy generates reduced FCP intensities. In order to test our assumption, we generated iBAQ values from LT and single-trypsin data sets and found that indeed LT generates higher iBAQ values than single-trypsin digest (Figure S2 and Tables S4 and S5, Supporting Information). This finding is further illustrated by the LT-directed shape of the iBAQ volcano plots and the associated sum of all log$_2$-fold change ratios of iBAQ values biased toward LT in both SLS and DOC (Figure 2E,F). Therefore, the more comprehensive cleavage makes LT to the potentially more reliable procedure in representing absolute protein abundances via indices like iBAQ than single-trypsin strategies.

Analysis of the sequence context of occurring cleavage and miscalavage events provides valuable information on workflow specificities. For this analysis, we first assessed the global MCP target residues and created iLogo plots for all detected MCPs. In agreement with other studies, we observed that the pool of totally detected MCPs is characterized by the frequent occurrence of acidic residues adjacent to scissile lysine and arginine (Figure S3, Supporting Information). In addition, proline in +1 position and basic residues at both sides of cleavage occur. This cleavage consensus was observed for both detergent buffers.

Interestingly, we found a highly similar cleavage context in FCPs originating from tandem digests (Figure 3A,B). Here LT-FCPs clearly showed a particular enrichment of surrounding acidic amino acids as well as proline in +1, but also leucine in –1 positions next to cleaved lysine residues. This enrichment pattern was observed in both DOC and SLS (Figure 3A). Compared to LT-FCPs, no obvious cleavage preference was observed for trypsin-FCPs, except for an expected overrepresentation of arginine and a decrease in lysine. This finding was anticipated due to the well-known target selectivity toward arginine residues in single-trypsin digests relative to tandem digests with an initial LysC treatment. Nonetheless, consistent with the enriched LT-FCP motifs, trypsin-MCPs (Figure 3B) exhibited more acidic and basic amino acids surrounding the target side. This corresponds well to results obtained previously for urea–LT and LysC cleavage.

Following the extraction of peptides matching to all major MCP classes, we found that such peptides are highly enriched in the total MCP pools of the DOC (92%) and SLS (86%) experiments (Figure 3C). The Venn diagram shows that the numbers for each class are homogenously distributed, but with partial overlaps. In particular, we found that acidic-MCPs were considerably enriched in all trypsin conditions. This is illustrated in Figure 3D showing enriched “acidic” MCPs after single-trypsin digest in SLS as the most affected MCP class (Figure S4A, Supporting Information). In addition, we observed an increased proteolytic cleavage after proline in LT digests (Figure S4, Supporting Information), which agrees with a previous report.
Figure 1. A) Experimental design to assess the efficiency of tandem LysC/trypsin (LT) versus single-trypsin digest in sodium lauroyl sarcosinate (SLS) and sodium deoxycholate (DOC). E. coli strain MG1655 cell pellets were lysed using SLS buffer, acetone/methanol precipitated and reconstituted in DOC or SLS. Proteins were then digested by LysC/trypsin (LT) or sole trypsin in 0.5% detergent. Following digestion, the peptide samples were analyzed using LC–MS with subsequent label-free quantification (LFQ). B) Bar chart representing the number of fully (FCP) and miscleaved peptide (MCP) spectra detected upon applying the different digestion workflows. C) Volcano plot comparing the relative enrichment of FCPs and MCPs as a result of the LT and single-trypsin workflows in DOC. D) The same as in (C) but in SLS buffer.

Our data demonstrate that LT-tandem digests perform efficiently in constant detergent concentrations, and do not require a high chaotropic state for the initial LysC digest. To further support our conclusion, we performed another LT digest experiment using 6 or 1.6 M urea for the LysC step and compared it to single-trypsin digest in 1.6 M urea. We observed that tandem proteolysis was consistently more efficient in increasing the quantity of FCPs (Figure 3E–G). The comparison of the two tandem digestion strategies, however, showed that the different urea concentrations (1.6 or 6 M) used in LysC digest had no major differential impact on the proteolytic result (Figure 3G; Figure S5 and Table S6, Supporting Information). We additionally confirmed that the
sequence cleavage context for FCPs and MCPs was preserved among all comparisons, including the LT urea (Figure S6, Supporting Information) and LT detergent workflows (Figure 3A,B).

Taking together, we aimed to investigate the digestion efficiency of tandem LysC/trypsin digestion in DOC and SLS. As concluded from the increased number of FCPs, increased cleavage efficiency and conserved cleavage motifs, we show that LT digests in these two MS-compatible detergents are highly efficient. Furthermore, the increased FCP yield has a positive impact on quantitative approaches building on them. Finally, we suggest
Figure 3. IceLogo plots showing enriched sequences at A) fully cleaved and B) mis-cleaved peptide stretches. Plots were generated using significantly enriched peptides from LT versus trypsin LFQ comparison in DOC and SLS. C) Venn diagrams showing the overlap between main classes of MCPs. D) Volcano plot illustrating the log2-fold change in all MCPs with acidic amino acids surrounding the mis-cleavage side, which resulted from LT versus single-trypsin digests in SLS. E and F) Volcano plots of all FCPs (E) or MCPs (F), resulting from a LT versus trypsin digestion with all tandem steps including LysC digest carried out in 1.6 M urea. G) Bar chart representing the number of significantly enriched FCPs and MCPs from tandem 1.6 M urea, tandem 6 M urea, and single-trypsin 1.6 M urea digestion comparisons.
that the main positive attributes of LT strategies are not an immediate effect of the sample denaturation state, but rather a result of the different target specificities between LysC and trypsin.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD013273.

Supporting Information
Supporting Information is available from the Wiley Online Library or from the author.

Conflict of Interest
The authors declare no conflict of interest.

Keywords
digestion efficiency, LysC, shotgun proteomics, trypsin

Received: April 8, 2019
Revised: August 22, 2019
Published online: October 1, 2019

[1] T. Glatter, C. Ludwig, E. Ahrne, R. Aebersold, A. J. Heck, A. Schmidt, J. Proteome Res. 2012, 11, 5145.
[2] E. I. Chen, D. McClatchy, S. K. Park, J. R. Yates, 3rd, Anal. Chem. 2008, 80, 8694.
[3] W. J. Hervey, M. B. Strader, G. B. Hurst, J. Proteome Res. 2007, 6, 3054.
[4] J. R. Wisniewski, A. Zougman, N. Nagaraj, M. Mann, Nat. Methods 2009, 6, 359.
[5] J. R. Wisniewski, M. Mann, Anal. Chem. 2012, 84, 2631.
[6] L. H. Betancourt, A. Sanchez, I. Pla, M. Kuras, Q. Zhou, R. Andersson, G. Marko-Varga, J. Proteome Res. 2018, 17, 2556.
[7] C. Chiva, M. Ortega, E. Sabido, J. Proteome Res. 2014, 13, 3979.
[8] J. W. Poulson, C. T. Madsen, C. Young, F. M. Poulson, M. L. Nielsen, J. Proteome Res. 2013, 12, 1020.
[9] T. Glatter, E. Ahrne, A. Schmidt, J. Proteome Res. 2015, 14, 4472.
[10] A. Hakobyan, W. Liesack, T. Glatter, J. Proteome Res. 2018, 17, 3086.
[11] A. Schmidt, K. Kochanowski, S. Vedelaar, E. Ahrne, B. Volkmer, L. Callipo, K. Knoops, M. Bauer, R. Aebersold, M. Heinemann, Nat. Biotechnol. 2016, 34, 104.
[12] N. A. Kulak, G. Pichler, I. Paron, N. Nagaraj, M. Mann, Nat. Methods 2014, 11, 319.
[13] T. Sander, N. Farke, C. Diehl, M. Kuntz, T. Glatter, H. Link, Cell Syst. 2019, 8, 66.
[14] B. Schwanhausser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, M. Selbach, Nature 2011, 473, 337.
[15] N. Colaert, K. Helsens, L. Martens, J. Vandekerckhove, K. Gevaert, Nat. Methods 2009, 6, 786.
[16] R. Rajmakers, P. Neerincx, S. Mohammed, A. J. Heck, Chem. Commun. 2010, 46, 8827.
[17] P. Giansanti, L. Tsiatsiani, T. Y. Low, A. J. R. Heck, Nat. Protoc. 2016, 11, 993.