Contribution ID: **58** Type: **Oral Presentation**

Properties, Origin, and Reproducibility of Truncated Proteoforms Across Top-Down Proteomic Studies

Wednesday, August 27, 2025 10:35 AM (5 minutes)

Top-down proteomics (TDP) offers a powerful approach to identifying intact proteoforms, inherently providing information about protein modifications. Truncated proteoforms are among the most frequently observed modifications in TDP studies. Here, we selected fifty datasets, including over 140,000 proteoforms, published over the past decade, spanning various organisms, sample preparation approaches, data acquisitions, and data analysis pipelines, and investigated the reported proteoforms. On average, across all studies, 70% of the proteoforms were truncated, and only 30% were identified as full-length proteoforms (including those with only N-terminal methionine excision). Only 16% of the exclusively N-terminally, 5% of the C-terminally, and 1% of the N- and C-terminally truncated proteoforms have been described in the UniProt database, highlighting that the biological function of most truncations is unknown. To understand cleavage patterns more clearly, we determined the amino acids N- and C-terminally located from the truncation sites. We found truncation sites to be very diverse, with specific datasets showing unique patterns. Several truncation sites indicated artifacts introduced during sample preparation (e.g., between the aspartate-proline bond due to sample heating), mass spectrometric analysis (e.g., N-terminal to proline residues due to in-source fragmentation), or data analysis (e.g., due to database and precursor tolerance settings). Moreover, substantial differences between samples of the same organism but different tissues or growth conditions were also observed. Importantly, our analysis revealed specific protein termini not linked to artifactual cleavages that were consistently reported across numerous studies, implicating these proteoforms to have biological significance. This includes previously unannotated mitochondrial signal peptide sites and cleavages with protein domain or structural specificity. In conclusion, this study provides a comprehensive overview of truncated proteoforms identified in recent TDP studies, highlighting both methodological and biological influences. We believe these results can also serve as a resource for other scientists to investigate non-canonical termini from proteins of interest.

User consent

yes

Author: Dr KAULICH, Philipp T. (Systematic Proteome Research & Bioanalytics, Institute for Experimental Medicine, Christian-Albrechts-Universität zu Kiel, 24105 Kiel, Germany)

Co-authors: Prof. THOLEY, Andreas (Systematic Proteome Research & Bioanalytics, Institute for Experimental Medicine, Christian-Albrechts-Universität zu Kiel, 24105 Kiel, Germany); Dr FULCHER, James M. (Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, USA)

Presenter: Dr KAULICH, Philipp T. (Systematic Proteome Research & Bioanalytics, Institute for Experimental Medicine, Christian-Albrechts-Universität zu Kiel, 24105 Kiel, Germany)

Session Classification: Lightning Talks