Enabling Technologies for N-terminal & 'Genome free' proteomics; *de novo* sequence analysis by a combination of LysN protein digestion and electron transfer dissociation

Albert J.R. Heck

Biomolecular Mass Spectrometry and Proteomics Group, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands and Netherlands Proteomics Centre

In this talk targeted novel targeted proteomics technologies will be discussed used to analyze I) protein N-termini and II) proteomes of species of uncharacterized genomes.

Although N-terminal processing of proteins is an essential process, not many large inventories are available, in particular not for human proteins. Using modern day mass spectrometry based proteomics techniques it is now possible to unravel N-terminal processing in a semi-comprehensive way. Strong cation exchange chromatography with improved separation of singly charged peptides was exploited for the targeted analysis of N-acetylated protein termini from human HEK293 cells. Taking advantage of the complementarity between Lys-N, Lys-C, and trypsin for protein digestion, a total of 1391 non-redundant acetylated protein N-termini could be identified in a multi-protease approach, representing the largest dataset of human acetylated protein N-termini to date. Sequence analysis and comparison of the dataset with related datasets from D. melanogaster, S. cerevisiae and H. salinarum provides new insights into N-terminal processing across these species.

For species with un-sequenced or poorly characterized genomes de novo sequencing of MS/MS fragmentation spectra is essential. However, de novo sequencing is challenging due to the complexity of common CID fragmentation spectra. Lys-N enzymatic cleavage in combination with ETD analysis results in fragmentation spectra almost exclusively containing N-terminal fragment ions. These, easy to interpret, ladder sequences open up a completely new window for de novo sequencing. As a proof of concept we analyzed the proteomes of ostrich muscle and hibernating bear heart. We performed a proteomics study of ostrich through Lys-N proteolytic cleavage followed by low-pH SCX fractionation, RP-nanoLC separation and ETD dissociation. The SCX fractionation is used for isolation of the 'de novo sequence-able' peptides. These peptides produce fragmentation spectra, after ETD, dominated by c-type ions, which are relatively easy to interpret. De novo analyses of the ETD spectra is performed by an in-house developed algorithm, called LysNDeNovo, which utilizes the presence of a single fragment ion series to assign the peptide sequence. Our de novo sequencing approach results in a significant higher number of peptides identified than searching the ETD Ostrich proteomics dataset using the Mascot search engine. Moreover, the de novo results allow the determination of point mutations as well as conserved regions between proteins of different species.

References

SCX-based fractionation of Lys-N generated peptides facilitates the targeted analysis of post-translational modifications
N. Taouatas , A.F. M. Altelaar, M. M. Drugan, A. O. Helbig, S. Mohammed and A. J.R. Heck.
Mol. Cell. Proteomics 8 (2009) 190-200.

Lys-N and Trypsin cover complementary parts of the phosphoproteome in a refined SCX-based approach
S. Gauci, A. O. Helbig, M. Slijper, J. Krijgsveld, A. J.R. Heck, S. Mohammed
Anal Chem 81 (2009) 4493-4501.