On the Advantages of Admixed Lys-C/Lys-N Digests for Proteome Depth and DeNovo Sequencing

Christopher M. Adams, Ryan D. Leib, Allis S. Chien
Vincent Coates Foundation Mass Spectrometry Laboratory, Stanford University, Stanford CA

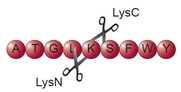
Overview

The preparative cornerstone of bottom-up or shotgun proteomics is enzymatic protein digestion to peptides. Digestion has many advantages over intact approaches, including increased fragmentation efficiency and availability of robust database search methods vs. in-silico peptide sequences to assign fragment ion spectra.

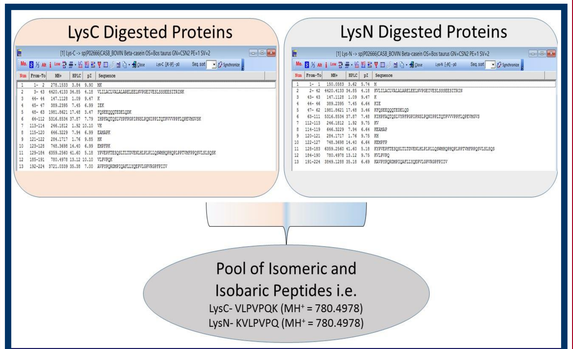
We examine the utility of combining LysC and LysN proteolytic digests, which doubles digest complexity but increases the probability of peptide identification for isobaric and isomeric peptides.

Introduction

Most enzymatic proteases specifically target a particular amino acid residue at a particular terminus.



Heck[1] pioneered the use of LysN, with advantageous b- and c- fragment ion series via HCD and ETD. We rationalized that combining LysC and LysN digests (LysCN) would provide isomeric *and* isobaric peptide pairs, potentially providing more complete b/y and/or c/z fragment ions for superior scoring, terminal characterization and de novo sequencing.



Scheme 1.
LysC and LysN (LysCN) in silico digest of Bovine beta casein demonstrating the isobaric nature of the digested peptides for all portions of the protein except the terminal sequences. Pooling the two digests doubles the complexity of the sample.

Methods

Sample preparation

A 10 protein mix, yeast lysate, and human K562 lysate were each solubilized in 8M urea, 50mM ammonium bicarbonate, then reduced and alkylated using DTT (5mM) and propionamide (10mM), respectively. Samples were adjusted to <1M urea with 50mM ammonium bicarbonate pH 8.0 and digested overnight at 37˚C using 1:25-50 protease to protein ratios of LysC and LysN in separate reactions. The resulting digests were quenched and cleaned up on stage tips. Terminal enrichment was done using a "positional proteomics" approach as previously reported[2].

Mass Spectrometry and Liquid Chromatography

NanoLC-MS/MS analysis was run on an Acquity M-Class using an in-house packed C18 reversed-phase column, with either a LTQ Orbitrap Velos or Fusion mass spectrometer using both High-High (HCD) and High-Low (CID) methodologies.

Database Searching and Statistical Evaluation

.RAW files were searched using Byonic to a 1% FDR and post-processed using Combyne, Matlab, and Excel. The processed data was visualized with the Byonic viewer and Excel.

Results

Retention Time Shift

Differences in pI for isomeric peptides (Scheme 1) resulted in significant retention time shifts. Universally, the peptides containing C-terminal K residues eluted earlier than the N-terminal K isomers. While some instances of co-elution were observed, they were not amenable to co-fragmentation.

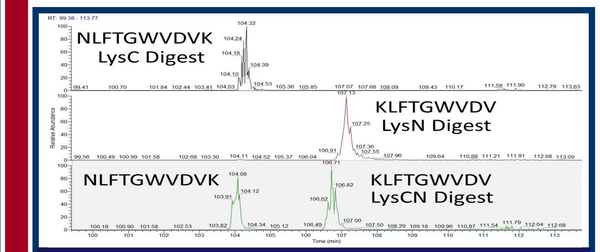


FIGURE 1. Typical peptide chromatography
N-terminal K containing peptides eluted an average of 2.5 minutes later than their C-terminal K isomers.

Proteome Depth

With complex mixtures, e.g. yeast and K562 lysates, the doubled complexity of LysCN admixed digests overwhelms the instrument duty cycle. This saturation results in minimal differences in proteome depth vs. independent digests at typical chromatographic timescales, but significant improvements in protein assignment scores. The result metrics for 90 min. and 120 min. chromatographic runs of LysC, LysN, and the combined LysCN in Table 1 below demonstrate the importance of understanding sample complexity when selecting analytical conditions for admixed digests.

Digest Type/ Acquisition (90min)	# Protein	# Spectra (PSM)	# Unique Peptides	Avg Log Probability	Avg % Coverage
LysC (.5ug)	1210	8840	5508	19.24	17.95
LysN (0.5ug)	1129	7060	4629	14.28	15.19
LysC+N (.25ug each)	1006	8583	5531	22.02	18.80

Digest Type/ Acquisition (120min)	# Protein	# Spectra (PSM)	# Unique Peptides	Avg Log Probability	Avg % Coverage
LysC (.5ug)	1437	10358	6444	20.27	17.42
LysN (0.5ug)	1280	8233	5247	14.86	15.26
LysC+N (.5ug each)	1150	10277	6577	19.43	19.10

Table 1. Comparison of whole cell yeast lysate results
Yeast lysates were run independently with LysC only, LysN only, and combined LysCN. The number of proteins or peptides does not improve in the combined approach in such complex mixtures, due to saturation of the MS duty cycle. The improvement is found in the the average protein log probability score as well as in the average % sequence coverage. Note: the LysCN numbers are underestimated as a result of using Combyne to integrate the data.

De Novo Sequencing and Probability Scores

The advantages of the LysCN approach are best applied in recovering low scoring fragment ion spectra or pseudo de novo analysis. Ideally, the isobaric/isomeric peptide pairs would co-elute. In most instances, the peptides did not co-elute in reversed phase chromatography. Nonetheless, spectral assignment confidence increased in many instances because of the complementary LysCN spectra. Figure 2 exemplifies the scenario on a low scoring 26mer peptide from the yeast HSP71 protein.

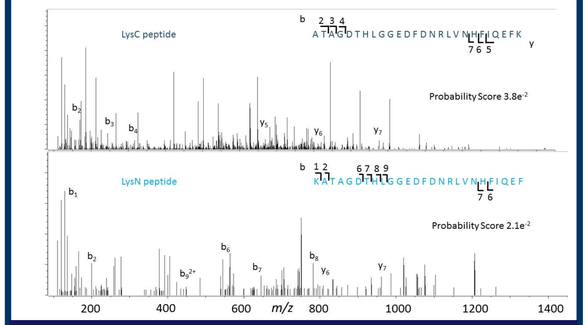


Figure 2. Complementarity of LysCN peptide fragment spectra Probability scoring of a isomeric peptide increased as a result of observing both LysC and LysN generated peptides.

Future

- Identify chromatographic conditions that enable co-isolation of LysC and LysN peptide pairs
- Apply strategy to areas which would benefit most:
  - De novo sequencing of antibodies
  - Characterization of lysine post-translational modifications
  - Identification of peptide N-termini

Conclusions

When LysC and LysN digests are combined, the LysC and LysN peptide pairs typically elute several minutes apart under reversed phase chromatographic conditions. This lack of co-elution doubles the sample complexity when LysC and LysN digests are combined, and highlights the importance of adjusting the chromatographic separation to suit the complexity of the sample.

The average log probabilities of identified proteins increased, and low scoring peptides can be rescued, when both LysC and LysN isomeric peptides are present and provide complementary information.

References

1. Taouatas N, *Nature Methods* **5** (5): 405–7 (2008)
2. McDonald L, *Nature Methods* **2**, 955 – 957 (2005)

Acknowledgements

The Vincent and Stella Coates Foundation, NIH NCRR Award Number S10RR027425.