

Integrated Positive and Negative Enrichment of N-terminal Peptides Reveals the Complexity of Proteolytic Processing in *Aspergillus niger*

Jong-Seo Kim, Ziyu Dai, Uma K. Aryal, Qibin Zhang, Joseph N. Brown, Samuel O. Purvine, Matthew E. Monroe, Ronald J. Moore, David G. Camp II, Scott E. Baker, Richard D. Smith, and Wei-Jun Qian
Biological Sciences Division and Energy Process and Materials Division, Pacific Northwest National Laboratory, Richland, WA



Pacific Northwest
NATIONAL LABORATORY

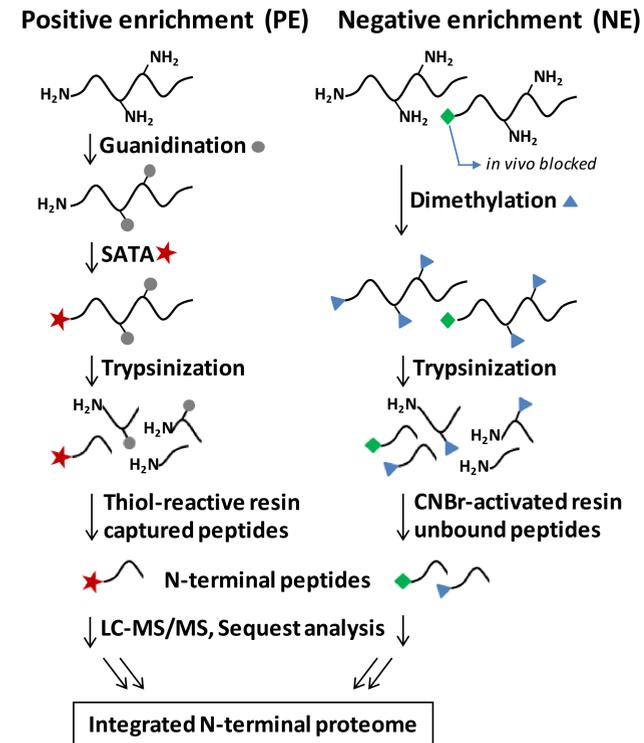
Overview

- Proteomic profiling of protein N-terminal peptides represents an effective way to characterize proteolytic processing, an important post-translational modification.
- Both positive and negative enrichment approaches were developed and optimized for effective isolation of N-terminal peptides with good specificity.
- The integration of these two approaches provided significantly increased coverage of the N-terminome.
- The resulting data from *Aspergillus niger* revealed the complexity of proteolytic processing.

Introduction

- Proteolytic processing of proteins, the irreversible hydrolysis by proteases, has been recognized as an important post-translational modification that regulates cellular protein functions.¹
- Due to the inherent complexity of the proteome, selective enrichment methods of protein N-terminal peptides are necessary for effective N-terminomics.²
- We developed a novel thiol reactive resin-assisted positive enrichment method with high specificity (>97%).
- To complement this positive enrichment, we optimized a negative enrichment method based on CNBr-activated resin³ and dimethylation, after evaluating several amine-reactive resins and N-blocking chemistries.
- The integrated results from the positive and negative selection strategies showed significantly enhanced coverage of proteolytic cleavage sites in cell lysates of *Aspergillus niger*.

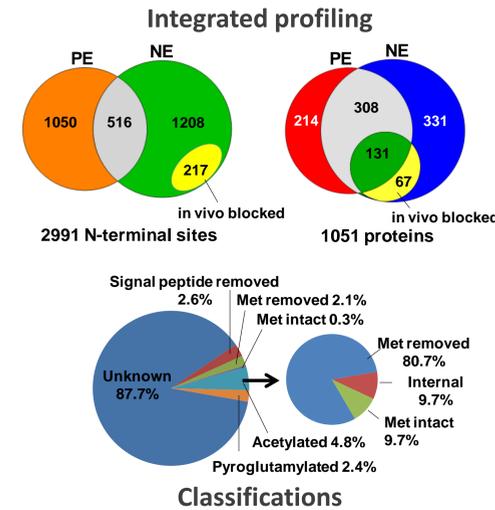
Methods



- Sample preparation.** *Aspergillus niger* (NRRL3122) was grown on potato dextrose agar (PDA) plates. Proteins were extracted in the presence of a protease inhibitor cocktail and subjected to reduction and alkylation of cysteine residues.
- Enrichment.** For positive enrichment, the protein lysine ϵ -amines are converted into homoarginine. A protected thiol group is then introduced onto the protein N-terminus using the SATA reagent. Following digestion and deprotection of thiols, the thiol-containing N-terminal peptides are enriched by a thiol-reactive resin and released by DTT. For negative enrichment, both protein α - and ϵ -amines are blocked via reductive dimethylation. After digestion, the internal peptides with newly generated primary amines were removed by amine-reactive scavenging CNBr-activated resin, resulting in enrichment of N-terminal peptides in the flow-through.
- LC-MS/MS.** Enriched N-terminal peptides were analyzed using LC-MS/MS (LTQ-Orbitrap). Spectra were searched by SEQUEST against the *A. niger* JGI database. Peptides were filtered based on the MS generating function⁴ value with <1% FDR for peptides and a 10 ppm mass tolerance for parent ions.

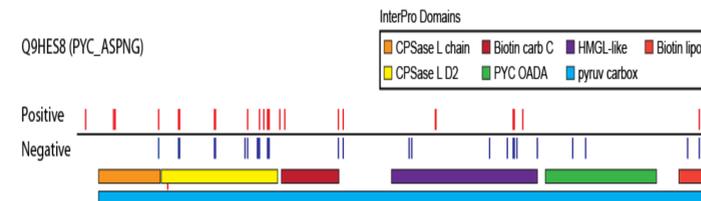
Results

Proteolytic processing sites (PPS)

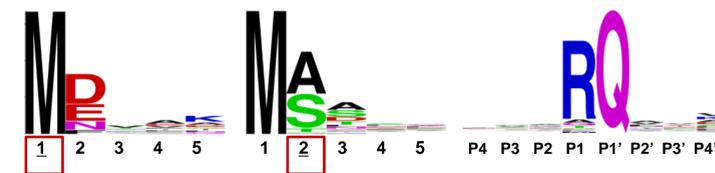


The *in vivo* blocked N-terminal sites were those identified with acetylation or pyroglutamylation modifications. Some proteins (or genes) containing *in vivo* N-terminal modifications were identified both PE and NE methods via different N-terminal peptides.

Complexity of proteolytic processing and sequence motif of *in vivo* blocked N-termini

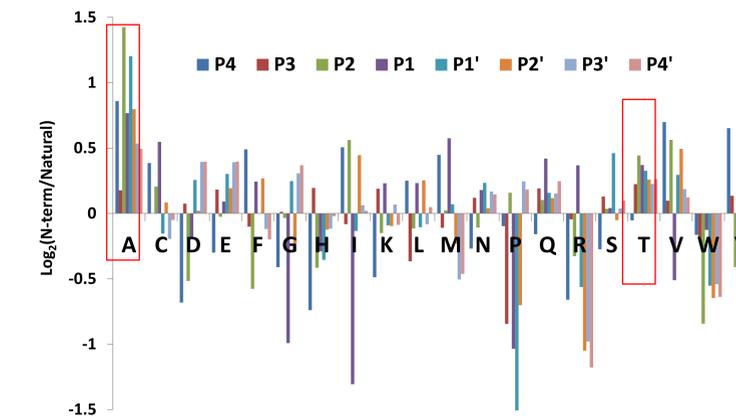


A representative protein, pyruvate carboxylase, illustrating the complexity of proteolytic processing by observing forty proteolytic sites identified from positive and negative enrichment methods through seven domains.

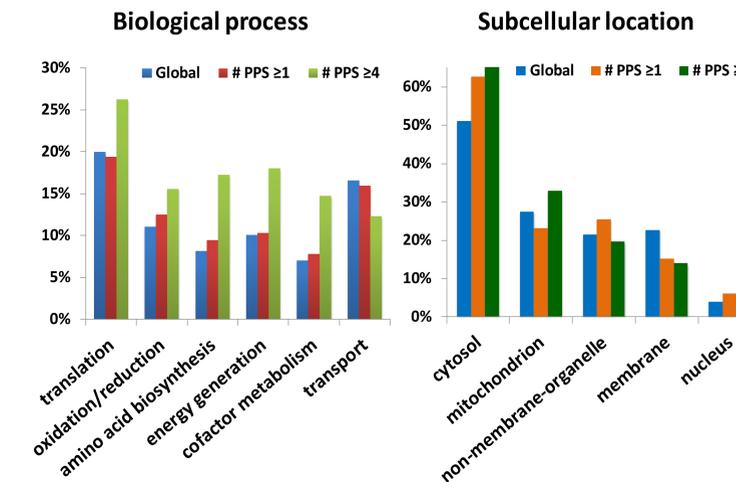


Sequence motifs of N-terminal acetylated peptides at initiator methionine or second residue highlighted with red square box, and sequence motif for pyroglutamylated peptides.

Implications for *in vivo* protease activity



Normalized frequencies of amino acid residues around cleavage site named P4, P3, P2, P1, P1', P2', P3', P4', where the peptide bond between P1 and P1' is cleaved. Each amino acid at individual positions from P4 to P4' was normalized by the natural occurrence of each amino acid in *A. niger* whole proteome and plotted in log₂ scale. The overall complex pattern indicates the potential of activities from many different proteases. The higher frequencies of A and T residues may indicate that a substantial number of the identified N-terminal sites originated from the removal of uncharacterized signal peptides.



The integrated proteins were sorted according to biological process or subcellular location. # PPS: number of proteolytic processing sites per protein.

Conclusions

- Both positive and negative enrichment have limitations to providing an unbiased coverage of the N-terminal proteome due to the different blocking and enrichment strategies being employed.
- The integrated results suggest that the positive and negative enrichment methods are complementary for achieving a broad coverage of the proteolytic processing events.
- The majority of the identified N-terminal sites were not previously annotated, suggesting the complexity of proteolytic processing in *A. niger*.
- The integrated method should be powerful for identifying potential new functional protein proteolytic forms.

Acknowledgements

This work was supported by a U. S. Department of Energy (DOE) Early Career Research Award and by grants from the National Center for Research Resources (5P41RR018522-10) and the National Institute of General Medical Sciences (8 P41 GM103493-10) from the National Institutes of Health. Experimental work was performed in the Environmental Molecular Science Laboratory, a DOE/BER national scientific user facility at Pacific Northwest National Laboratory (PNNL) in Richland, Washington. PNNL is operated by Battelle for the DOE under contract DE-AC05-76RLO-1830.

References

- C Lopez-Otin and CM Overall, *Nat Rev Mol Cell Biol* 3:509 (2002)
- O Kleifeld, A Doucet, UAD Keller, et al., *Nat Biotechnol* 28:281 (2010)
- XM Zhang, JY Ye, and P Hojrup, *J Proteomics* 73:240 (2009)
- S Kim, N Gupta, and PA Pevzner, *J Proteome Res* 7:3354 (2008)

CONTACT: Wei-Jun Qian, Ph.D.
Biological Sciences Division, K8-98
Pacific Northwest National Laboratory
P.O. Box 999, Richland, WA 99352
E-mail: weijun.qian@pnnl.gov