ENRICHMENT OF NATURAL AND PROTEOLYTICALLY GENERATED PROTEIN C–TERMINI ON A PROTEOME–WIDE SCALE

Schilling Oliver
Institute for Molecular Medicine and Cell Research, University of Freiburg, Germany

Barré Olivier
Centre for Blood Research and Department of Oral Biological and Medical Sciences, University of British Columbia, Vancouver, Canada

Huesgen Pitter F.
Centre for Blood Research and Department of Oral Biological and Medical Sciences, University of British Columbia, Vancouver, Canada

Overall Christopher M.
Centre for Blood Research and Department of Oral Biological and Medical Sciences and Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, Canada

Method Article

Keywords: Proteomics, Proteolysis, Degradomics, C-terminal

Posted Date: June 10th, 2010

DOI: https://doi.org/10.1038/nprot.2010.100

Abstract

Introduction

The sequence and nature of the amino \(N) and carboxyl \(C) termini of proteins provides crucial functional annotation of proteomes as their modification or truncation affects all proteins, often influencing protein fate and function. While protein N termini are relatively well studied \[1-3] comparatively little is known about protein C termini. Techniques for proteome–wide characterization of C–terminal sequences have long been considered a fundamentally missing tool in functional proteomics. In contrast to commonly used techniques for C–terminal sequencing, which are restricted to samples consisting of only few proteins, the present protocol constitutes one of the first strategies for the targeted analysis of protein C termini from complex cellular proteomes \[4]. With proteolytic processing being a key post-translational modification, interest in the focused analysis of protein N and C termini stems from research on proteolysis on a proteome-wide level \[5, 6]. System–wide approaches to elucidate proteases and their substrates are commonly referred to as degradomics \[7] and include more specialized techniques focusing on the N terminome, also called N terminomics \[8-11]. Examples of important N-terminal cleavages are many \[5, 6], but these have been less studied for C-terminal truncations generating neo C termini. In part this is because few techniques have been developed to target the carboxyl terminus of proteins. In turn, this is due to its lower chemical reactivity compared to primary amino groups. This is a serious deficit because proteolytic processing close to protein C termini and truncations by carboxypeptidases—increasingly recognized for their involvement in physiological and pathological processes \[12-14]—can only be investigated with degradomic approaches focusing on neo C termini. We present a strategy for the enrichment of C-terminal peptides from complex proteomes and their identification by liquid chromatography-tandem mass spectrometry \(LC–MS/MS). We call this focused analysis of the C termini of proteomes on a global scale C-terminomics. Stable isotope labeling can be incorporated to compare C–terminal processing in different samples. The protocol focuses on commercially available, cost–efficient reagents. The essence of the approach is a negative selection strategy that involves binding and removal of internal tryptic peptides to enrich for the C- terminal peptides. To do this thiol groups of full-length proteins are first reduced and alkylated followed by reductive methylation of N-terminal α-amines and lysine ε-amines. Now one key step is performed: carboxyl groups—including protein carboxy-termini together with aspartate and glutamate side-chains—are protected by carbodiimide-mediated condensation with ethanolamine. At a concentration of 1.0 M, ethanolamine is used in large excess over the protein components, similar to other carboxyl-labeling approaches \[15]. Proteins are then trypsin digested, yielding internal tryptic peptides with free, newly formed neo N and C termini and original protein C-terminal peptides that have protected carboxyl groups from the previous step. To prevent cross-reactivity at the following selection step, newly formed N termini are also chemically protected by reductive methylation. Internal and N-terminal peptides are then covalently coupled through their free carboxyl groups to the primary amine groups of a linear polymer poly-allylamine by carbodiimide-mediated condensation. Poly-allylamine is used at a 1.0 mM concentration, which represents a primary amine concentration of ~1.0 M. The protected protein C-

terminal peptides remain uncoupled and are separated from the high molecular weight polymer by ultrafiltration using a size cutoff of 10 kDa. Uncoupled protein C-terminal peptides remain in the flow-through and are analyzed by LC-MS/MS. Chemical labeling of protein C termini performed at the protein level provides an essential validation step for natural and proteolytically generated C termini since it clearly distinguishes in vivo cleavage events from proteolysis during sample preparation. Similarly, incorporation of stable isotope labeling is essential for quantitative comparison of C–terminal processing in different samples in order to distinguish altered proteolysis \(e.g. by protease over–expression or depletion) from ongoing background proteolysis. The present labeling strategy makes use of the amine protection steps that we call “carboxy-terminal amine incorporated labeling of substrates \(C-TAILS)”.

Reagents

Cellular proteome from an organism with a fully sequenced genome. 1-ethyl-3-\[3-dimethylaminopropyl]carbodiimide hydrochloride \(EDC) 2-\(N-morpholino)ethanesulfonic acid \(MES) 4-\(2-hydroxyethyl)piperazine-1-ethanesulfonic acid \(HEPES) Acetonitrile Dithiothreitol \(DTT) DNAse I Ethanolamine Formaldehyde \(light isotope d\(0)C12), for stable isotope labeling also heavy formaldehyde \(d\(2)C13, Cambridge Isotopes) Guanidine hydrochloride Iodoacetamide Methanol N-hydroxysuccinimide \(NHS) Ortho-phenanthroline Phenylmethanesulfonyl fluoride \(PMSF) Poly-allylamine \(average molecular weight ~56 kDa) RNAse A Trans-epoxysuccinyl-L-leucylamido\(4-guanidino)butane \(E-64) Trichloracetic acid \(TCA)

Equipment

10-kDa cut-off Microcon filter devices C18 solid phase extraction columns Standard laboratory equipment for cell culture, molecular biology, and protein chemistry Tabletop centrifuge Tandem mass spectrometer coupled to a liquid chromatography frontend Ultrasonication bath Water bath or thermo block

Procedure

Proteome Harvest The present protocol allows for the enrichment of C–terminal peptides from complex cellular proteomes. Stable isotope labeling is optionally incorporated to allow for the quantitative comparison of C-termini from different samples. 1. Lyse cells in 100 mM 4-\(2-hydroxyethyl)piperazine-1-ethanesulfonic acid \(HEPES), pH 7.5 in the presence of 1 µg/ml of DNAse I and RNAse A and the following protease inhibitors: 1 mM trans-epoxysuccinyl-L-leucylamido\(4-guanidino)butane \(E-64), 10 mM phenylmethanesulfonyl fluoride \(PMSF), and 1 mM ortho-phenanthroline. Lysis methods such as ultrasonication or freeze–thaw cycles can be employed. Keep samples separate in experiments where comparisons will be made. 2. Centrifuge at 12,000 x g for 10 min and proceed with supernatant. 3. Determine protein concentration and use 2 mg as starting material. 4. Add guanidine hydrochloride to a final concentration of 2.0 M. Protection of Sulfhydryls and Primary Amines 5. Reduce disulfide bonds by incubation with 20 mM dithiothreitol \(DTT, final concentration) for 30 – 60 min at 70 °C. 6. Protect free

sulfhydryls by incubation with 50 mM iodoacetamide \(final concentration) for 3 h at 25 °C. 7. Quench unreacted iodoacetamide by incubation with 5 – 10 mM DTT \(final concentration) for 30 min at 25 °C. 8. Protect primary amines by dimethylation using 20 mM formaldehyde \(final concentration) and 20 mM NaCNBH3 \(ALD coupling solution, final concentration) for 16 h at 25 °C. In case of a comparative experiment, label one sample with light formaldehyde \(d\(0)C12) and one sample with heavy formaldehyde \(d\(2)C13) \(Cambridge Isotopes). 9. Precipitate proteins by addition of 15 % \(v/v, final concentration) trichloracetic acid \(TCA). 10. Incubate on ice for 1 h. 11. Centrifuge at 12,000 x g for 10 min, keep pellet and discard supernatant. 12. Wash pellet three times with –20 °C cold methanol. If pellet loosens, repeat centrifugation. Protection of Carboxyls 13. Resolubilize proteins in 200 mM 2-\(N-morpholino)ethanesulfonic acid \(MES), pH 5.0, 2.0 M guanidine hydrochloride, 1.0 M ethanolamine. Incubate in an ultrasonication bath to enhance resolubilization. 14. Check that pH equals 5.0 and adjust if required. 15. Add 10 mM N-hydroxysuccinimide \(NHS) and 20 mM 1-ethyl-3-\[3-dimethylamino-propyl]carbodiimide hydrochloride \(EDC). 16. After 1 min incubation, check that pH equals 5.0 and adjust if required. 17. Incubate for 1 h at 25 °C. 18. Repeat EDC step – including pH control – twice, with the last incubation lasting 16 h. 19. Precipitate proteins by addition of 15 % \(v/v, final concentration) TCA. 20. Incubate on ice for 1 h. 21. Centrifuge at 12,000 x g for 10 min, keep pellet and discard supernatant. 22. Wash pellet three times with –20 °C cold methanol. If pellet loosens, repeat centrifugation. Trypsin Digestion and Chemical Protection of new N Termini 23. Resuspend proteins in a minimal volume of 20 mM HEPES pH 7.5, 2.0 M guanidine hydrochloride. Incubate in an ultrasonication bath to improve resolubilization. 24. Reduce guanidine hydrochloride concentration to 0.5 M by addition of 20 mM HEPES pH 7.5. 25. Determine protein concentration. 26. Add sequencing-grade trypsin at a ratio of 1:100 \(w/w, trypsin/total protein) and incubate for 16 h at 37 °C. 27. After digestion, protect neo-N-termini of internal peptides by reductive dimethylation with 20 mM \(final concentration) formaldehyde and 20 mM \(final concentration) NaCNBH3 and incubation for 6 – 16 h at 37 °C. In case of a comparative experiment, label one sample with light formaldehyde \(d\(0)C12) and one sample with heavy formaldehyde \(d\(2)C13) formaldehyde \(Cambridge Isotopes). Use the same labeling scheme as in step 8. After amine protection, mix samples at equal abundance. Depletion of Internal and N-terminal Peptides 28. Prepare a 2 mM stock solution of poly-allylamine \(average molecular weight ~56 kDa, Sigma) in 200 mM MES pH 5.0, 2 M guanidine hydrochloride, 20 % \(v/v) acetonitrile. 29. Check that pH of poly-allylamine stock equals 5.0 and adjust if required. 30. Add one volume of poly-allylamine stock to the tryptic digest to reach a final poly-allylamine concentration of 1 mM. 31. Add 10 mM NHS and 50 mM EDC \(final concentrations). 32. After 1 min incubation, check that pH equals 5.0 and adjust if required. 33. Incubate for 3 h at 25 °C. 34. Repeat EDC step – including pH control – twice, with the last incubation lasting 16 h. 35. Pass polymer / tryptic digestion mixture through a 10-kDa cut-off Microcon filter device \(Millipore). Due to the hygroscopicity of poly-allylamine, the ultrafiltration step can be lengthy. Keep the filtrate, which contains uncoupled, C-terminal peptides. 36. Repeat ultrafiltration step twice with 500 µl water. 37. Pool the three filtrates and desalt using C18 solid phase extraction columns. 38. Determine peptide concentration. Liquid Chromatography-Tandem Mass Spectrometry and Data Analysis 39. Analyze C-terminal peptides by liquid chromatography-tandem mass spectrometry \(LC-MS/MS). Standard LC-MS/MS settings for peptide/proteome analysis are employed. These settings are machine

specific and not within the scope of this protocol. 40. Spectrum-to-sequence assignment using an appropriate software such as Mascot \[16] or X\!Tandem \[17]. Since trypsin does not cleave at dimethylated lysines, set cleavage-specificity as C-terminal to arginine \(ArgC-like specificity). Apply semi-style cleavage searches with up to three missed cleavages. Set static modifications as carboxyamidomethylation of cysteine residues \(+57.02 Da), dimethylation of lysines and amino termini \(+28.03 Da for normal isotope abundance; 34.06 Da for d\(2)C13 formaldehyde), and ethanolamine-modification \(+43.04 Da) for aspartate and glutamate residues as well as peptide C-termini. For quantitative comparison with stable isotope labeling, use software tools like XPRESS \[18] to determine Heavy:Light \(H/L) ratios.

Timing

4 days

Critical Steps

EDC is unstable in water; always use fresh EDC solutions. Control pH frequently during EDC coupling.

Anticipated Results

10s – 100s of C-terminal peptides.

References

\[1] Auf dem Keller U., Prudova A., Gioia M., Butler G.S., Overall C.M., A statistics-based platform for quantitative N-terminome analysis and identification of protease cleavage products, Mol. Cell. Proteomics 9 \(2010) 912-927. \[2] Kleifeld O., Doucet A., Auf dem Keller U., Prudova A., Schilling O., Kainthan R.K., Starr A.E., Foster L.J., Kizhakkedathu J.N., Overall C.M., Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products, Nat. Biotechnol. 28 \(2010) 281-288. \[3] Prudova A., Auf dem Keller U., Butler G.S., Overall C.M., Multiplex N-terminome analysis of MMP-2 and MMP-9 substrate degradomes by iTRAQ-TAILS quantitative proteomics, Mol. Cell. Proteomics 9 \(2010) 894-911. \[4] Schilling O., Barre O., Huesgen P.F., Overall C.M., Proteome-wide analysis of protein carboxy termini: C terminomics, Nat. Metods in press \(2010). \[5] Doucet A., Butler G.S., Rodríguez D., Prudova A., Overall C.M., Metadegradomics: toward in vivo quantitative degradomics of proteolytic post-translational modifications of the cancer proteome, Mol. Cell. Proteomics 7 \(2008) 1925-1951. \[6] Schilling O., Overall C.M., Proteomic discovery of protease substrates, Curr. Opin. Chem. Biol. 11 \(2007) 36-45. \[7] López-Otín C., Overall C.M., Protease degradomics: a new challenge for proteomics, Nat. Rev. Mol. Cell. Biol. 3 \(2002) 509-519. \[8] Dix M.M., Simon G.M., Cravatt B.F., Global mapping of the topography and magnitude of proteolytic events in apoptosis, Cell 134 \(2008) 679-691. \[9] Gevaert K., Goethals M., Martens L., Van Damme J., Staes A., Thomas G.R., Vandekerckhove J., Exploring proteomes and analyzing protein processing by mass spectrometric identification of sorted N-terminal peptides., Nat.

Biotechnol. 21 \(2003) 566-569. \[10] Mahrus S., Trinidad J.C., Barkan D.T., Sali A., Burlingame A.L., Wells J.A., Global sequencing of proteolytic cleavage sites in apoptosis by specific labeling of protein N termini, Cell 134 \(2008) 866-876. \[11] Timmer J.C., Enoksson M., Wildfang E., Zhu W., Igarashi Y., Denault J.-B., Ma Y., Dummitt B., Chang Y.-H., Mast A.E., Eroshkin A., Smith J.W., Tao W.A., Salvesen G.S., Profiling constitutive proteolytic events in vivo, Biochem. J. 407 \(2007) 41-48. \[12] Gomis-Rüth F.X., Structure and mechanism of metallocarboxypeptidases, Crit. Rev. Biochem. Mol. Biol. 43 \(2008) 319-345. \[13] Reznik S.E., Fricker L.D., Carboxypeptidases from A to Z: implications in embryonic development and Wnt binding, Cell. Mol. Life Sci. 58 \(2001) 1790-1804. \[14] Vendrell J., Querol E., Avilés F.X., Metallocarboxypeptidases and their protein inhibitors. Structure, function and biomedical properties, Biochim. Biophys. Acta 1477 \(2000) 284-298. \[15] Panchaud A., Hansson J., Affolter M., Bel Rhlid R., Piu S., Moreillon P., Kussmann M., ANIBAL, stable isotope-based quantitative proteomics by aniline and benzoic acid labeling of amino and carboxylic groups, Mol. Cell. Proteomics 7 \(2008) 800-812. \[16] Perkins D.N., Pappin D.J., Creasy D.M., Cottrell J.S., Probability-based protein identification by searching sequence databases using mass spectrometry data, Electrophoresis 20 \(1999) 3551-3567. \[17] Craig R., Beavis R.C., TANDEM: matching proteins with tandem mass spectra, Bioinformatics 20 \(2004) 1466-1467. \[18] Han D.K., Eng J., Zhou H., Aebersold R., Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry, Nat. Biotechnol. 19 \(2001) 946-951.