Protease subsite profiling with proteome-derived peptide libraries (PICS)

Oliver Schilling

UBC Centre for Blood Research, Department of Oral Biological and Medical Sciences, 4.401 Life Sciences Institute, 2350 Health Sciences Mall, University of British Columbia, Vancouver, BC, V6T 1Z3, Canada

Christopher M. Overall

UBC Centre for Blood Research, Departments of Oral Biological and Medical Sciences, and Biochemistry and Molecular Biology, 4.401 Life Sciences Institute, 2350 Health Sciences Mall, University of British Columbia, Vancouver, BC, V6T 1Z3, Canada

Method Article

Keywords: proteases, subsite profiling, peptide libraries, proteomics

Posted Date: May 14th, 2008

DOI: https://doi.org/10.1038/nprot.2008.88

Abstract

Introduction

Specificity of numerous proteases is determined by enzyme-substrate interactions in an extended active site cleft. The Schechter and Berger nomenclature[1] designates residues carboxy-terminal to the scissile peptide bond as prime side \(P') and amino-terminal residues as non-prime side \(P). P and P' residues interact with complementary protease subsites called S and S'. While some proteases show very restrictive specificity profiles \(e.g. trypsin is specific for lysine and arginine in P1), other proteases display broad specificity profiles involving multiple subsite preferences. A bottleneck in protease characterization is the determination of consensus cleavage sites as all current techniques provide information on either the prime or nonprime side only–none provide both in the same experiment. Genetic library approaches such as phage display provide cleavable sequences only without identifying the actual scissile bond. Proteome-derived peptide libraries employ biological sequence diversity to robustly characterize protease subsite specificities by mass spectrometry[2]. Termed Proteomic Identification of protease Cleavage Sites \(PICS), peptide libraries are generated from cell lysates or secreted proteins. These libraries are used as substrate screens for test proteases. Prime-side cleavage products are isolated and identified by liquid chromatography/tandem mass spectrometry \(LC-MS/MS). The corresponding non-prime side sequences are derived by bioinformatic database searches. Hence, the entire sequence of the peptidic substrate is reconstructed. The cohort of cleaved peptides, often surpassing 200 unique cleavage sequences in single experiments, characterizes subsite preferences \(both positive and negative) together with subsite cooperativity. Like any other peptide-based approaches for the determination of active site specificity, PICS does not aim to identify physiological protease substrates. PICS is optimally suited to determine sequence specificity of endoproteases. Since the method isolates prime-side cleavage products, it is not suited to profile carboxypeptidases as these prime-side cleavage products are too short for LC-MS/MS identification. Many amino peptidases can not cleave at blocked N-termini, which PICS proteome wide libraries feature. PICS utilizes peptide libraries with modified lysine and cysteine residues and hence cannot profile specificity for these residues. PICS screens employing unmodified lysine or cysteine residues are currently under development[2]. Proteome-derived peptide libraries are generated by endoproteolytic digestion of proteomes such as cell culture lysates. The specificity of the applied digestion protease is important since it determines sequence features of the peptide library. PICS libraries have been generated with the following three digestion proteases: – Trypsin cleaves C-terminal to lysine and arginine, hence tryptic peptides lack internal basic residues. – GluC \(_Staphylococcus aureus_ protease V8) cleaves C-terminal to glutamate and, to a lesser extent, aspartate; hence GluC peptides lack internal glutamate residues. – Chymotrypsin cleaves C-terminal to large hydrophobic residues, hence chymotryptic peptides have a lower internal content of such residues. PICS profiling of matrix metalloprotease \(MMP)-2 PICS using chymotryptic libraries confirmed the MMP-2 preference for leucine in P1', thereby indicating that numerous potential chymotryptic cleavage sites remained intact[2]. Profiling proteases using different PICS libraries generated with different digestion proteases provides an unbiased analysis of cleavage site specificity. Our data reported in Schilling and Overall \(2008)[2] confirms that similar specificity profiles are generated using these different libraries. Proteome-derived peptide libraries were shown to be sufficiently rich in sequence diversity to profile sequence specificity of proteases[2]. Library generation and isolation of modified peptides can easily be modified to screen the sequence specificity of any posttranslational modification that is amenable to LC-MS/MS detection \(e.g phosphorylation or acetylation). Proteome-derived peptide libraries can also be used to characterize sequence specificity of ligand or receptor binding.

Reagents

Strains Eukaryotic cell line from organism with sequenced genome. **Reagents** 4-\(2-hydroxyethyl)piperazine-1-ethanesulfonic acid \(Sigma-Aldrich) Acetonitrile \(Sigma-Aldrich) Bicinchoninic acid assay \(Pierce) Chymotrypsin \(1-chloro-3-tosylamido-7-amino-2-heptanone treated, Worthington) Cyanoborohydride \(ALD coupling solution, Sterogene)

Dimethylsulfoxide \(Sigma-Aldrich) Dithiothreitol \(Sigma-Aldrich) DNase I \(Sigma-Aldrich) Ethylenediaminetetraacetic \(Sigma-Aldrich) Formaldehyde \(Sigma-Aldrich) Formic acid \(Sigma-Aldrich) GluC \(V8 protease, Worthington) Methanol \(Sigma-Aldrich) micro bicinchoninic acid assay \(Pierce) NaCl \(Sigma-Aldrich) Phenylmethylsulphonyl fluoride \(PMSF, Sigma-Aldrich) Phosphate buffered saline Potassium phosphate buffer Reversed-phase solid phase extraction cartridges \(e.g. Sep-Pak, Waters, or equivalent) RNase E \(Sigma-Aldrich) Sephadex G-10 resin \(GE Healthcare) Sodium dodecyl sulfate \(SDS, Sigma-Aldrich) Spin columns, ~10 µm pore size, top and bottom plug \(Pierce) Streptavidin sepharose \(GE Healthcare) Sulfopropyl sepharose fast flow resin \(GE Healthcare) Sulfosuccinimidyl 2-\(biotinamido)-ethyl-1,3-dithiopropionate \(Pierce) Test protease for PICS assay Trans-epoxysuccinyl-L-leucylamido\(4-guanidino)butan \(E-64,Sigma-Aldrich) Trichloracetic acid \(Sigma-Aldrich) Trifluoracetic acid \(Sigma-Aldrich) Trypsin \(1-chloro-3-tosylamido-4-phenyl-2-butanone treated, Sigma-Aldrich) **Abbreviations** BCA, bicinchoninic acid DMSO, dimethylsulfoxide DTT, dithiothreitol EDTA, ethylenediaminetetraacetic E-64, trans-epoxysuccinyl-L-leucylamido\(4-guanidino)butan HEPES, 4-\(2-hydroxyethyl)piperazine-1-ethanesulfonic acid LC-MS/MS, liquid chromatography-tandem mass spectrometry PBS, phosphate buffered saline PMSF, phenylmethylsulphonyl fluoride SDS, sodium dodecyl sulfate SPE, solid phase extraction Sulfo-NHS-SS-biotin, sulfosuccinimidyl 2-\(biotinamido)-ethyl-1,3-dithiopropionate PAGE, polyacrylamid gel electrophoresis TCA, trichloracetic acid TFA, trifluoracetic acid

Equipment

Standard laboratory equipment for cell culture, molecular biology, and protein chemistry Centrifuge for volumes up to 50 ml at up to 20000 g Chromatography columns to pack up to 5 ml resin Chromatography system \(e.g. GE Healthcare or equivalent) with UV/VIS and conductivity detector LC-MS/MS Personal computer with Perl interpreter \(see "www.clip.ubc.ca/resources/index.html":http://www.clip.ubc.ca/resources/index.html) and Microsoft EXCEL NanoDrop spectrophotometer pH test strips \(e.g. from Merck, pH range 6–9) Polystyrene-divinylbenzene resin column \(e.g. GE Healthcare or equivalent) Table top centrifuge for 2 ml reaction tubes Ultrasonication bath Vacuum evaporation system \(commonly referred to as speed-vac) XK 26/20 columns \(GE Healthcare)

Procedure

Protease profiling with proteome derived peptide libraries consists of – library preparation – PICS assay Since proteome-derived peptide libraries can be prepared in batches, library generation is uncoupled from protease profiling. The two sections on library preparation and protease profiling are therefore independent and do **not** represent one continuous workflow. **Preparation of proteome derived peptide libraries** _Cell culture_ Proteome-derived peptide libraries can be generated from any proteome source with known genome or proteome sequence, such as human or murine cell lines. Prokaryotic cell lines are less suited since they represent lower sequence diversity. We have used the human fibrosarcoma cell line HT1080 in monolayer culture and the 293 human embryonic kidney cell line in suspension culture using the chemically defined, protein-free CD 293 medium \(Invitrogen). In order to produce sufficient starting material for several subsite screening assays, we recommend to harvest cell material from at least 6 cell culture flasks of 175 cm^2 surface each. When using serum-supplemented medium, cells are grown serum-free when reaching 50 % confluency. Cells are washed three times in phosphate buffered saline \(PBS) \(138 mM NaCl, 2.7 mM KCl, 20 mM Na_2HPO_4, 1.5 mM KH_2PO_4, pH 7.4) and grown in serum-free medium for 12–24 h. _Proteome harvest_ Cell-conditioned medium, cell lysate, and detergent-solubilized lysis pellet can be used to generate proteome-derived peptide libraries. In our experience and using the cell strains described above, cell lysate and detergent-solubilized lysis pellet result in markedly higher library yields than cell conditioned medium. _Library preparation from cell-conditioned medium_ 1| Decant medium from cell culture flasks. See section “Library preparation from cell lysate” on how to proceed with harvesting of cells. 2| Add protease inhibitors: 1.0 mM phenylmethylsulphonyl fluoride \(PMSF, serine protease inhibitor, 100x stock in acetonitrile should be freshly prepared, Sigma-Aldrich) 10 µM trans-epoxysuccinyl-L-leucylamido\(4-guanidino)butan \(E-64, cysteine protease

inhibitor, 100x stock in dimethylsulfoxide \(DMSO), Sigma-Aldrich) 10 mM ethylenediaminetetraacetic acid \(EDTA, for inhibition of metalloproteases, EDTA is easier to remove than the more specific reagent 1,10-phenanthroline, Sigma-Aldrich). 3| Centrifuge at 2,200 g at 4 °C for 30 min, keep supernatant and discard pellet. 4| Precipitation of proteins with trichloracetic acid \(TCA, Sigma-Aldrich) in order to remove protease inhibitors. TCA is added to a final concentration of 10 % \(vol/vol), followed by 1 h incubation on ice and centrifugation at 8,500 g at 4 °C for 1 h. 5| Discard supernatant. 6| Wash pellet twice with -20 °C methanol. 7| Air-dry for 10 min and overlay with 1 mL of 100 mM 4-\(2-hydroxyethyl)piperazine-1-ethanesulfonic acid \(HEPES, Sigma-Aldrich), 0.1 % \(wt/vol) sodium dodecyl sulfate \(SDS, Sigma-Aldrich), pH 7.4. 8| TCA pellet resolubilization by mild ultrasonication for 30 min in a chilled ultrasonication bath, followed by incubation at -20 °C \(-80 °C if suspension does not freeze) for 1 h and a second incubation at 70 °C for 30 min. 9| Protein quantization by bicinchoninic acid \(BCA, Pierce) assay and/or absorbance at 280 nm. 10| Dilute 5-fold with 200 mM HEPES pH 7.4, to a final SDS concentration of 0.02 %. Then add 5 mM dithiothreitol \(DTT, Sigma-Aldrich). 11| Keep 20 µl aliquot for SDS-polyacrylamide gel electrophoresis \(PAGE) analysis. 12| Specific endoproteolytic digestion using one of the following three endoproteases: trypsin \(1-chloro-3-tosylamido-4-phenyl-2-butanone treated, Sigma-Aldrich) GluC \(V8 protease, Worthington) chymotrypsin \(1-chloro-3-tosylamido-7-amino-2-heptanone treated, Worthington). 13| Enzyme/proteome ratio for each of the digestion proteases is 1:50. 14| Add 5 µg/ml DNase I \(Sigma-Aldrich) and 5 µg/ml RNase E \(Sigma-Aldrich). 15| Incubate at 37 °C for 16 h. 16| Add 1 mM PMSF and centrifuge at 40,000 x g for 0.5 h. 17| Keep supernatant, discard pellet. 18| Take 20 µl aliquot to control digest efficiency by SDS-PAGE. 19| Continue at section sulfhydryl and amine protection. _Library preparation from cell lysate_ 1| Wash cells three times with PBS. 2| Detach cells in PBS supplemented with 0.2 % \(wt/vol) EDTA. Cell detachment scheme can be adjusted to meet specific needs of particular cell strains; however proteases such as trypsin should not be used. 5–10 mL detachment solution are sufficient for a 175 cm^2 cell culture flask. 3| Centrifuge cells at 400 g at 4 °C for 5 min; decant supernatant and keep pellet. 4| Resuspend cells in 10 ml hypotonic lysis buffer containing protease inhibitors: 10 mM HEPES, pH 7.4, 1.0 mM PMSF, 10 µM E-64, and 10 mM EDTA. 5| Detergent free cell lysis. We routinely employ nitrogen cavitation. Alternative lysis methods include cell membrane shearing by repeated aspiration of the cell suspension with 20- or 27-gauge needles. 6| Centrifuge cell lysate at 40,000 x g for 1 h. Both the soluble fraction \(supernatant) and the insoluble lysis pellet \(containing, among other components, cell organelles, nuclei and unlysed cells) can be used to generate peptide libraries. This section describes usage of the soluble cell lysate, usage of the cell lysis pellet is described in the following section. 7| TCA precipitation of the cell lysate in order to remove protease inhibitors. TCA is added to a final concentration of 10 % \(vol/vol) followed by 1 h incubation on ice and centrifugation at 8,500 g at 4 °C for 1 h. 8| Discard supernatant. 9| Wash pellet twice with -20 °C methanol 10| Air-dry for 10 min and overlay with 1 mL 100 mM HEPES pH 7.4, 0.1 % \(wt/vol) SDS. 11| Protein quantization by BCA assay and/or absorbance at 280nm. 12| Dilute 5-fold with 200 mM HEPES pH 7.4 to a final SDS concentration of 0.02 %, then add 5 mM DTT. 13| Keep 20 µl aliquot for SDS-PAGE analysis. 14| Specific endoproteolytic digestion using one of the following three endoproteases: trypsin \(1-chloro-3-tosylamido-4-phenyl-2-butanone treated) GluC chymotrypsin \(1-chloro-3-tosylamido-7-amino-2-heptanone treated). Enzyme/proteome ratio for each of the digestion proteases is 1:50. 15| Add 5 µg/ml DNase I and 5 µg/ml RNase E. 16| Incubate at 37 °C for 16 h. 17| Add 1 mM PMSF and centrifuge at 40,000 x g for 0.5 h. 18| Keep supernatant, discard pellet. 19| Take 20 µl aliquot to control digest efficiency by SDS-PAGE. 20| Continue at section sulfhydryl and amine protection. _Library preparation from detergent-solubilized lysis pellet_ 1| Carefully remove remaining supernatant from lysis pellet. 2| Resuspend lysis pellet in 5 ml of 100 mM HEPES pH 7.4, 0.02 % SDS \(wt/vol). 3| \(optional) Freeze-thaw cycle of lysis pellet to enhance lysis extent. 4| Incubate at 70 °C for 30 min. Let cool to ambient temperature \(approx. 22 °C). 5| Take 100 µl aliquot, centrifuge at 14,000 g for 10 min and determine protein concentration of supernatant \(BCA assay). 6| Add 5 mM DTT. 7| Add 5 µg/ml DNase I and 5 µg/ml RNase E and incubate at 37 °C for 2–3 h. 8| Keep 20 µl aliquot for SDS-PAGE analysis. 9| Specific endoproteolytic digestion using one of the following three endoproteases: trypsin \(1-chloro-3-tosylamido-4-phenyl-2-butanone treated) GluC chymotrypsin \(1-chloro-3-tosylamido-7-amino-2-heptanone treated). Enzyme/proteome ratio for each of the digestion proteases is 1:50. 10| Incubate at 37 °C for 16 h. 11| Add 1 mM PMSF and centrifuge at 40,000 x g for 0.5 h. 12| Keep supernatant, discard pellet. 13| Take 20 µl

aliquot to control digest efficiency by SDS-PAGE. 14| Continue at section sulfhydryl and amine protection. _Sulfhydryl and amine protection_ 1| Add 5 mM DTT, incubate at 70 °C for 30 min, let cool to 22 °C. 2| Add 40 mM iodoacetamide \(Sigma-Aldrich). 3| After 5 min incubation control that pH remains 7.0–7.5. If required, adjust pH to 7.5. 4| Incubate at 22 °C for 3 h. 5| Add 10 mM DTT to quench remaining iodoacetamide. 6| After 15 min incubation control that pH remains 7.0–7.5. If required, adjust pH to 7.5. 7| Add 10 mM formaldehyde \(Sigma-Aldrich) and 10 mM $NaCNBH_3$ \(cyanoborohydride, “ALD coupling solution”, Sterogene), incubate for 3–16 h at 22 °C. 8| Repeat formaldehyde labeling once. _Peptide library purification_ - The first purification step employs preparative size exclusion chromatography using Sephadex G-10 resin \(GE Healthcare). For samples volumes of 5–15 ml we employ two self-packed XK 26/20 columns \(GE Healthcare) in tandem. Other chromatography setups might be suitable as well, should however be tested beforehand. Samples are buffer exchanged to 10 mM potassium phosphate buffer, pH 2.7, 20 % \(vol/vol) acetonitrile. Peptide elution is monitored by absorption at 214 nm or 220 nm; elution of sample buffer and salt is monitored by conductivity. Sephadex G-10 has an exclusion limit of 700 Da. Salt and buffer are retained and elute after the PICS peptides. The 700 Da exclusion limit also retains peptides with less than six residues. Such peptides are not suited for PICS analysis since potential cleavage products are too short for LC-MS/MS analysis. Flow rate with an AKTA chromatography system is 5 ml/min. PICS peptides are collected for further purification. - The second purification step employs strong cation exchange chromatography using sulfopropyl sepharose fast flow resin \(GE Healthcare, ca 5 mL resin in self-packed columns); binding and washing in 10 mM potassium phosphate buffer, pH 2.7, 20 % \(vol/vol) acetonitrile. Peptide binding and elution is monitored by absorption at 214 nm or 220 nm. Extensive washing \(approximately 10 column volumes) is required. Elution is performed by a one-step gradient of 1 M NaCl in 10 mM potassium phosphate buffer, pH 2.7, 20 % \(vol/vol) acetonitrile. 1| Remove acetonitrile by vacuum-evaporation \(“speed-vac”). Reduce sample volume by 50 % and replenish with water. 2| The third purification step employs reversed phase chromatography using polystyrene-divinylbenzene resin. 0.6 % \(vol/vol) formic acid \(Sigma-Aldrich) is added to the sample for binding. Washing is performed with 0.3 % \(vol/vol) formic acid and elution with 80% \(vol/vol) acetonitrile. Other setups for reversed-phase solid phase extraction \(SPE) are suitable if they provide sufficient binding capacity. 3| Concentrate eluate to a minute amount of solvent in a vacuum concentration system. 4| Resuspend in 0.5 ml water. 5| Take 10 µl aliquot for peptide quantization by BCA assay. 6| Adjust concentration to 2–10 mg/ml. 7| Store at -70 °C. **Test protease assay and LC-MS/MS readout** 1| Thaw peptide library; 200 µg library per protease if no clean-up is required before biotinylation, 400 µg per protease if clean-up is required before biotinylation \(see below). 2| Adjust appropriate buffer conditions for the test protease \(pH, type of buffer, reducing agents, co-factors). Use buffer that is free of detergents and primary or secondary amines \(exception: enzyme co-factors). 3| Adjust peptide library concentration to 1 mg/ml. 4| Add activated protease; typical protease library ratios are 1:1000–1:50 \(wt/wt) with 1:100 \(wt/wt) typically applied for a first enzyme characterization. 5| Incubate for 1–16 h at a temperature suitable for the protease under investigation. 6| \(optional) Heat de-activate protease. - If PICS assay buffer and protease preparation neither contain primary amines nor reducing substances such as DTT, skip this step. If primary amines or reducing substances are present, perform reversed-phase SPE using C18 SPE cartridges with sufficient peptide binding capacity \(e.g. Sep-Pak, Waters, or equivalent products): acidify sample with 0.1–0.5 % trifluoracetic acid \(TFA, Sigma-Aldrich), bind to column, wash with 5 column volumes of 0.5 % \(vol/vol) formic acid, elute with 2 column volumes 80 % acetonitrile, vacuum-evaporate to near-dryness and reconstitute sample in 200 µl 200 mM HEPES pH 7.4. 7| Control that pH is 7–8, adjust if required. 8| Biotinylation of PICS assay using 0.5 mM sulfosuccinimidyl 2-\(biotinamido)-ethyl-1,3-dithiopropionate \(sulfo-NHS-SS-biotin, amine-reactive biotin with a cleavable disulfide linker, Pierce). Like all NHS-esters, sulfo-NHS-SS-biotin readily hydrolyzes in aqueous buffers, particularly at basic pH conditions. A 10 mM sulfo-NHS-SS-Biotin solution \(20x stock) in DMSO is prepared immediately prior to mixing with the PICS assay. 9| Incubate at 22 °C for 2 h. 10| Equilibrate high-capacity streptavidin sepharose \(GE Healthcare) in 50 mM HEPES, 150 mM NaCl, pH 7.4. Resin slurry volume is 1.5-times the volume of the PICS assay, providing sufficient binding capacity for bound and unbound biotin in PICS assay without further cleanup. Note that high-capacity streptavidin sepharose from GE Healthcare is supplied as a diluted slurry without explicit labeling. The commercially available unit of 5 ml slurry is provided as a 1:4 dilution

resulting in a total volume of 20 mL. Sepharose resin can be buffer equilibrated by repeated centrifugation and resuspension in the desired buffer. 11| Add buffer-equilibrated streptavidin sepharose to biotinylated PICS assay. Required volume of undiluted slurry is 1.5-times the volume of the PICS assay. 12| Incubate at 0.5 h at 22 °C. Apply mild agitation to keep slurry in suspension. 13| Pour slurry in a spin column of sufficient volume with a filter of approximately 10 µm pore size. In a typical PICS experiment, the total volume at this point is 550–600 µl \(200 µl library + buffer + protease + sulfo-NHS-SS-Biotin solution + 300 µl streptavidin sepharose). For this volume, we routinely use spin columns from Pierce. 14| Place column in a 2 ml reaction tube. 15| Centrifuge column. Centrifugation speed must be adjusted to pass the PICS assay solution through the column without letting the column run completely dry. After centrifugation, resin should still be wet but without supernatant. Since resin packing can vary, we recommend performing centrifugation first at 150 g for 1 min. If this centrifugation speed proves to be insufficient, elongate centrifugation time or increase centrifugation speed \(increments of 200 rpm). During every centrifugation, cover column top loosely with the supplied screw cap to minimize cross contamination. 16| Re-apply the first flow through to resin and centrifuge a second time. 17| Discard flow-through. 18| Wash resin: apply 500 µl of washing buffer \(50 mM HEPES, 150 mM NaCl, pH 7.4). 19| Centrifuge column and discard flow-through. 20| Repeat washing 8 times. 21| Prepare fresh elution buffer: 50 mM HEPES, 10 mM DTT, pH 7.4. 22| Press bottom plug in column. 23| Apply 500 µl of elution buffer. 24| Close column top with screw cap. 25| Incubate 2 h at 22 °C with mild agitation. 26| Place column in a clean 2 ml reaction tube. 27| Remove bottom plug and loosen screw cap. 28| Centrifuge and keep eluate. 29| Place column in a second clean 2 ml reaction tube. 30| Apply 500 µl of elution buffer. 31| Centrifuge and keep eluate. 32| Pool the two elution fractions. – Estimated eluate amount is 1–5 % of library amount \(2–10 µg for 200 µg library) depending on biotinylation efficiency as well as protease activity and specificity. 33| Reversed-phase SPE of eluate using C18 cartridges or tips with sufficient peptide binding capacity \(OMIX tips, Varian, or Sep-Pak cartridges, Waters). Acidify sample with 0.5 % \(vol/vol) formic acid, bind to column, wash with 5 column volumes of 0.5 % \(vol/vol) formic acid, elute with 2 column volumes 80 % acetonitrile, and vacuum-evaporate to near-dryness. 34| \(optional) Overlay dried sample with 5 µl of 1 mM DTT. Note that peptide concentration can only be determined in absence of DTT \(see below). 35| \(optional) Determine peptide concentration by reconstituting dried sample in water and use an aliquot for µBCA analysis \(Pierce). As stated above, the estimated eluate amount is 2–10 µg for 200 µg library. In our experience, this estimate has allowed for successful LC-MS/MS analysis. - Samples can be stored at -80 °C for several weeks. - Note that strong cation exchange fractionation cannot be applied to PICS eluate since thioacylated N-termini are not charged. - Description of LC-MS/MS setup is not within the scope of this protocol. PICS eluates have been successfully analyzed by electrospray quadrupole time-of-flight mass spectrometers, however other LC-MS/MS setups might be suitable as well. LC-MS/MS details can be found elsewhere[2]. _Data analysis_ 1| Several steps in data analysis depend on LC-MS/MS vendor specific data formats and access to commercial software. We provide examples for data originating from Applied Biosystems QSTAR instruments \(“.wiff” file format) and access to Mascot[3] for uninterpreted database searches. 2| Generate peak list from LC-MS/MS raw data. For wiff data, use the mascot.dll plug-in \(Matrix Science, "www.matrixscience.com/help/instruments_analyst.html":http://www.matrixscience.com/help/instruments_analyst.html) for Analyst software \(Applied Biosystems) with determination of charge state from survey scan, default precursor charge states of 2+ and 3+, MS/MS averaging of precursor scans \(precursor mass tolerance for grouping = 1, maximum number of cycles between groups = 10, and minimum number of cycles per group = 1), centroiding of MS/MS data, de-isotoping of MS/MS data, and rejection of MS/MS spectra with less than 10 peaks. – Peptide-to-spectrum assignment by uninterpreted database searches. We routinely employ Mascot. Peptides are identified from the IPI database[4]. In accordance with established protocols for QSTAR data analysis[5,6], mass tolerance is 0.4 Da for parent ions and 0.2 Da for fragment ions, allowing up to two missed cleavages. GluC is defined to cleave C-terminal to either \(i) glutamate or \(ii) glutamate and aspartate with results from both searches combined. Semi-style cleavage searches are applied with no constraints for the orientation of the specific terminus and the following static modifications: carboxyamidomethylation of cysteine residues \(+57.02 Da), dimethylation of lysine ε-amines \(+28.03 Da), and thioacylation of amino termini \(+88.00 Da). – We recommend post-analysis validation using PeptideProphet[7] at a 95% confidence level. – To assist with

the remaining steps of PICS analysis we supply a documented EXCEL spreadsheet and a documented Perl script at "www.clip.ubc.ca/resources/index.html":http://www.clip.ubc.ca/resources/index.html , section CLIP-PICS. - Obtain high-confidence peptide identifications from Mascot and/or PeptideProphet. Note that PICS is a peptide-centric method, hence only peptide, but not protein identifications are of interest. - Follow instructions in the PICS EXCEL spreadsheet and in the PICS Perl script to complete PICS analysis, including bioinformatic determination of non-prime sequences. - To display PICS subsite profiles in the form of heat maps, the EXCEL spreadsheet compiles position-specific amino acid frequencies from PICS sequence data \(“analysis tab”). Template files to generate heat maps with pro Fit software \(QuantumSoft) are available at "www.clip.ubc.ca/resources/index.html":http://www.clip.ubc.ca/resources/index.html . - To display PICS cleavage sites as sequence logos, sequences are aligned according to their subsites and a web-based utility is used \("www.cbs.dtu.dk/$_{\text{gorodkin/appl/plogo.html":http://www.cbs.dtu.dk/}}$gorodkin/appl/plogo.html). A pre-formatted alignment of PICS sequences is provided in the PICS EXCEL spreadsheet. – To investigate subsite cooperativity, the EXCEL spreadsheet \ (“analysis tab”) allows a focused analysis of substrates having a particular amino acid at a specific position. By comparing position-specific amino acid frequencies of such a focused analysis with position-specific amino acid frequencies of all substrates, a first characterization of subsite cooperativity can be performed. Additional software tools for the analysis of subsite cooperativity are currently being investigated and will be presented on "www.clip.ubc.ca/resources/index.html":http://www.clip.ubc.ca/resources/index.html .

Timing

Library preparation: 2–3 d PICS assay: 1–2 d

Critical Steps

Control pH during library generation. Sulfo-NHS-SS-biotin stock must be prepared freshly in DMSO. Equilibrate vial to room temperature before opening. Biotinylation can only be performed in absence of primary amines and reducing agents.

Troubleshooting

Very few \(< 10) cleavage products are identified. Use different PICS libraries made with complementary digestion proteases or made from different proteome fractions. Control library integrity by performing a PICS assay with a protease of known specificity. If possible, use a control test protease. PICS libraries have modified lysine and cysteine residues hence specificity for these residues cannot be profiled. Use PICS library from a source with little exposure to endogenous levels of the test protease. For example, use PICS libraries from cell lysate to profile extracellular proteases. Vary enzyme:library ratio.

Anticipated Results

10s–100s of cleavage site sequences representing specificity of the test protease.

References

1. Schechter, I. & Berger, A. On the size of the active site in proteases. I. Papain. _Biochem. Biophys. Res. Commun_. **27** \(1967). 2. Schilling, O. & Overall, C.M. Proteome-derived, database searchable peptide libraries for identifying protease cleavage sites. _Nat. Biotechnol_. **26** \(2008). 3. 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). 4. Kersey, P.J. et al. The International Protein Index: an integrated database for

proteomics experiments. _Proteomics_ **4** \(2004). 5. Elias, J.E., Haas, W., Faherty, B.K. & Gygi, S.P. Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. _Nat. Methods_ **2** \(2005). 6. Stephan, C. et al. Automated reprocessing pipeline for searching heterogeneous mass spectrometric data of the HUPO Brain Proteome Project pilot phase. _Proteomics_ **6** \(2006). 7. Keller, A., Nesvizhskii, A.I., Kolker, E. & Aebersold, R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. _Anal. Chem_. **74** \(2002).