Biochimie xxx (2015) 1–11

Contents lists available at ScienceDirect

Biochimie

journal homepage: www.elsevier.com/locate/biochi

Review

Current trends and challenges in proteomic identification of protease substrates

Matej Vizovišek [a, 1], Robert Vidmar [a, c, 1], Marko Fonović [a, b, *], Boris Turk [a, b, d, **]

[a] Department of Biochemistry and Molecular and Structural Biology, Jožef Stefan Institute, Jamova cesta 39, SI-1000 Ljubljana, Slovenia
[b] Centre of Excellence for Integrated Approaches in Chemistry and Biology of Proteins, Jamova cesta 39, SI-1000 Ljubljana, Slovenia
[c] International Postgraduate School Jožef Stefan, Jamova cesta 39, SI-1000 Ljubljana, Slovenia
[d] Faculty of Chemistry and Chemical Technology, University of Ljubljana, Aškerčeva cesta 5, SI-1000 Ljubljana, Slovenia

ARTICLE INFO

Article history:
Received 23 July 2015
Accepted 23 October 2015
Available online xxx

Keywords:
Proteomics
Proteases
Degradomics
Substrates
Terminomics

ABSTRACT

Proteolytic cleavage is a ubiquitous, irreversible, posttranslational modification that changes protein structure and function and plays an important role in numerous physiological and pathological processes. Over the last decade, proteases have become increasingly important clinical targets because many of their inhibitors are already used in the clinic or in various stages of clinical testing. Therefore, a better understanding of protease action and their repertoires of physiological substrates can not only provide an important insight into their mechanisms of action but also open a path toward novel drug design. Historically, proteases and their substrates were mainly studied on a case-by-case basis, but recent advancements in mass spectrometry-based proteomics have enabled proteolysis studies on a global scale. Because there are many different types of proteases that can operate in various cellular contexts, multiple experimental approaches for their degradomic characterization had to be developed. The present paper reviews the mass spectrometry-based approaches for determining the proteolytic events in complex biological samples. The methodologies for substrate identification and the determination of protease specificity are discussed, with a special focus on terminomic strategies, which combine peptide labeling and enrichment.

Contents

1. Introduction

The human genome encodes approximately 600 proteases that have important roles in vital physiological and pathological processes, such as proliferation, the immune response, physiological homeostasis, cell death, inflammation, cancer, cardiovascular and

* Corresponding author. Department of Biochemistry and Molecular and Structural Biology, Jožef Stefan Institute, Jamova cesta 39, SI-1000 Ljubljana, Slovenia.
** Corresponding author. Department of Biochemistry and Molecular and Structural Biology, Jožef Stefan Institute, Jamova cesta 39, SI-1000 Ljubljana, Slovenia.
E-mail addresses: marko.fonovic@ijs.si (M. Fonović), boris.turk@ijs.si (B. Turk).
[1] Both authors contributed equally to this manuscript.

http://dx.doi.org/10.1016/j.biochi.2015.10.017
0300-9084/

Abbreviations

2D-PAGE two-dimensional polyacrylamide electrophoresis
2D-DIGE two-dimensional difference gel electrophoresis
CLiPS cellular libraries of peptide substrates
ChaFRADIC charge-based fractional diagonal chromatography
COFRADIC combined fractional diagonal chromatography
(d)N-TOP (double) TMMP labeling approach
FPPS fast profiling of protease specificity
ICAT isotope-coded affinity tags
iTRAQ Isobaric tags for relative and absolute quantitation
LC-MS/MS liquid chromatography coupled with tandem mass spectrometry
MMP matrix metalloprotease
N-CLAP N-terminalomics by Chemical Labeling of the α-Amine of Proteins
NHS N-Hydroxysuccinimide
PICS Proteomic Identification of Protease Cleavage Sites
PITC phenyl isothiocyanate
PROTOMAP protein topography and migration analysis platform
PS-SCL positional scanning-substrate combinatorial assays
PTAG phospho tagging
PTM posttranslational modification
SAX strong anion exchanger
SCX strong cation exchanger
SDS-PAGE sodium dodecyl sulphate polyacrylamide electrophoresis
SILAC stable isotope labeling by amino acid in cell culture
SPECS secretome protein enrichment using click sugars
TAILS terminal amine isotopic labeling of substrates
TMMP trimethoxphenylphopshonium
TMMP-Ac-OSu (N-succinimidyloxycarbonylmethyl) tris (2,4,6-trimethoxyphenyl)
TopFIND Terminus Oriented Protein Function Inferred database
TOPPR the online protein processing resource
TRAIL TNF-related apoptosis-inducing ligand

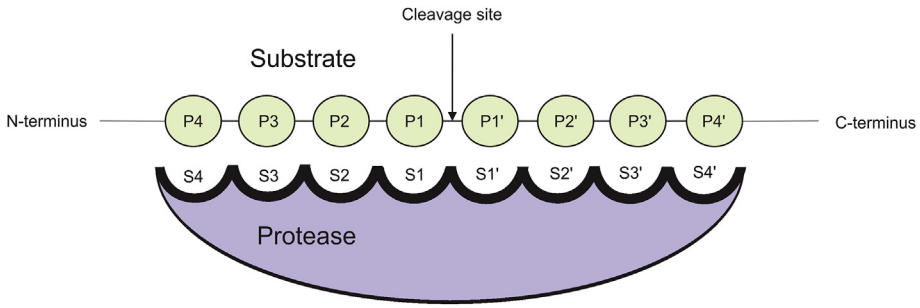

Fig. 1. *Nomenclature of protease-substrate interaction*. The substrate binding sites downstream of the cleavage site are numbered S1–Sn towards the N-terminus of the substrate (non-primed sites) and S1′–Sn′ towards the C-terminus (primed sites). The substrate residues are numbered P1–Pn, and P1′–Pn′ [5]. In either case, the numbering starts at the scissile bond.

neurodegenerative diseases, and infections [1,2]. Based on their catalytic mechanism, proteases are classified into serine, cysteine, metallo, aspartic, threonine and glutamyl proteases, and proteases of an unknown catalytic mechanism (MEROPS database, reviewed in Refs. [3,4]). Proteases can be further divided into endopeptidases, which cleave proteins inside the polypeptide chain, and exopeptidases, which cleave at the N- or C-terminus (aminopeptidases or carboxypeptidases). Accordingly, the cleavage results in the formation of two novel protein fragments or, in the case of exopeptidase, N- or C-terminally trimmed proteins. Thus proteolytic processing is an irreversible posttranslational modification (PTM) that changes the structure and function of their protein substrates. Protease-substrate interactions play a major role in the specificity of the proteolytic cleavage [1]. Schechter-Berger nomenclature (Fig. 1) is used to annotate the positions upstream or downstream of the cleavage site, with the substrate binding subsites on the surface of the protease numbered S1–Sn towards the N-terminus of the substrate (the so called non-primed sites) and S1′–Sn′ towards the C-terminus of the substrate (the so called primed sites), whereas the substrate residues they bind are numbered P1–Pn, and P1′–Pn′, respectively. In both cases, the numbering begins at the scissile bond [5].

Proteases with narrow specificity generally execute limited proteolysis (e.g., caspases during apoptosis), while proteases with broad specificity, such as cysteine cathepsins or the proteasome, often have major roles in general protein degradation and clearance, thereby governing the proteome composition of a cell [1]. In addition, the efficiency of the cleavage *in vivo* is determined by several other factors. First, the protease and the target substrate must be present in sufficient concentrations and must interact in the cellular environment under the favourable conditions required for protease activity (e.g., pH and redox state). Second, the presence of posttranslational modifications, endogenous inhibitors, allosteric effectors and other proteases can also significantly impact substrate processing *in vivo* [1,6–10]. Moreover, because even a small quantity of an active protease can trigger a physiological response, their *in vivo* activity is tightly regulated on several levels, including transcription (different expression levels of a protease), activation (synthesis as inactive zymogens), inhibition by endogenous inhibitors, compartmentalization [11]) and protease half-life [1,2,12].

It is crucial to identify a protease's physiological substrates to understand its action and position inside the proteolytic web [13]. However, although a substantial amount of data on proteases has been gathered over the past decade, we have still only identified a very limited subset of true physiological substrates. During the last 15 years, mass spectrometry has become an indispensable tool for identifying protease substrates in complex biological samples but also for determining protease specificities. However, a single experimental design is generally not sufficient for the study of complex proteolytic pathways, and various methodological approaches for proteomic studies of proteases had to be developed.

Please cite this article in press as: M. Vizovišek, et al., Current trends and challenges in proteomic identification of protease substrates, Biochimie (2015), http://dx.doi.org/10.1016/j.biochi.2015.10.017