Nicholas J. Kramer

Proteases

Bowers Lab

Proteolysis – Proteolytic cleavage or proteolysis is the enzymatic hydrolysis of a peptide bond in a peptide or protein substrate by a family of specialized enzymes termed proteases.

Protease – enzymes that aid in or carry out proteolysis; protein catabolism by hydrolysis of peptide bonds. Proteases can be categorized into seven broad groups, according to the specific residues and/or sites at which hydrolysis occurs. These include ***serine***, ***cysteine***, ***threonine***, ***aspartic***, ***glutamic***, ***metallo-***, and ***asparagine*** *peptide* proteases.

Useful Reviews:

- Klein, T.; Eckhard, O.; Dufour, A.; Solis, N.; Overall, C. M. Proteolytic Cleavage – Mechanisms, Function, and "Omic" Approaches for a Near-Ubiquitous Posttranslational Modification, *Chem. Rev.* **2018**, 118, 1137-1168
- López-Otín, C.; Bond, J. S. Proteases: Multifunctional Enzymes in Life and Death, *J. Bio. Chem.* **2008**, 283, 30433-30437
- Bordusa, F. Proteases in Organic Synthesis, *Chem. Rev.* **2002**, 102, 4817-4867
- Hedstrom, L. Serine Protease Mechanism and Specificity, *Chem. Rev.* **2002**, 102, 4501-4523
- Barrett, A. J.; Woessner, J. F.; Rawlings, N. D. Handbook of Proteolytic Enzymes; Elsevier: **2012**.

Common Commercial Proteases:

Cleavage Tag (Protease Name)	Sequences and Cleavage Site (= |)
3C 'PreScission'	LEVLFQ|GP
EKT (Enterokinase)	DDDDK|
Fxa (Factor Xa)	IEGR|
TEV (tobacco etch virus)	ENLYFQ|G
Thrombin	LVPR|GS
Trypsin	R| or K|
Chymotrypsin	F| or W| or Y|
Lys-C (Endoproteinase)	K|
Lys-N (Endoproteinase)	|K
Asp-N (Endoproteinase)	|D
Glu-C (Endoproteinase)	E| or D|
Furin	RXXR|
Proteinase K	see note below
Pepsin	(F/W/Y) |(F/W/Y)
Papain	(hydrophobic)RK|
Clostripain	R| or K|
Thermolysin	(L/F)|(F/V/M/A/I)
Cathepsin C	removes N-terminal dipeptide
Carboxypeptidase A	(nonspecific)|(aromatic or branched)
Carboxypeptidase B	specific for C-terminal R or K
TVMV	GTVAFG|S
WNV protease	(K/R)R|GS
Elastase	(A/G/V)|

3C 'PreScission' – Highly specific protease; cleaves between the Glu and Gly residues in the cleavage tag. Often produced with the tradename 'PreScission protease or PSP'.[1]

EKT (Enterokinase) – an intestinal enzyme normally involved in the protease cleavage of Trypsin. Cleaves after the Lysine (K) in its recognition sequence.[1]

Fxa (Factor Xa) – Cleaves after the Arg residue but can also cleave less frequently at secondary basic sites. Its most common secondary cleavage site is between the Gly and Arg residues in its own recognition site, although the frequency of these events is protein specific.[1]

TEV (tobacco etch virus) – Cleavage occurs between the Glu and Gly residues. TEV is often reported to have better specificity for its recognition site compared to EKT, Thrombin or Factor Xa.[1]

Thrombin – Cleaves preferentially between the Arg and Gly residues. Off target cleavage can occur at non-specific sites, normally from contaminating proteases. To ensure maximal protein integrity the enzyme reagent must be very pure.[1]

Trypsin – Cleaves peptides on the C-terminal side of lysine and arginine amino acid residues. If a proline is on the carboxyl side of the cleavage site, the cleavage will not occurs. If an acidic residue is on either side of the cleavage site, the rate of hydrolysis has been shown to be slower.[2]

Chymotrypsin – Cleaves at the C-terminus of the aromatic amino acids such as phenylalanine, tryptophan, and tyrosine. The specificity for aromatic side chains is because of the protein's hydrophobic pocket.[3]

Lys-C (Endoproteinase) –is a serine proteinase which cleaves peptide bonds at the carboxyl side of lysine. LysC is a sequencing grade enzyme and is subtle for proteomics and glycobiology applications.[4,5]

Lys-N (Endoproteinase) – Endoproteinase LysN cleaves at the amino-terminus of lysine residues (which includes methylated lysines).[4]

Asp-N (Endoproteinase) – Endoproteinase AspN (flavastacin) is a zinc metalloendopeptidase which selectively cleaves protein and peptide bonds N-terminal to aspartic acid residues.[4,5]

Glu-C (Endoproteinase) – Preferentially cleaves peptide bonds C-terminal to glutamic acid residues. Also cleaves at aspartic acid residues at a rate 100-300 times slower than at glutamic acid residues.[4,5]

Furin – Is a recombinant, ubiquitous subtilisin-like protein convertase with a minimal cleavage site of Arg-X-X-Arg/. However, the enzyme prefers the site Arg-X-Lys/Arg-Arg.[5]

Proteinase K – Hydrolyzes a variety of peptide bonds and is frequently used to cleanup enzymatic reactions or cell lysates. This cleaves at the carboxyl side of aliphatic, aromatic, and hydrophobic residues.[4,5]

Pepsin – Pepsin is most efficient in cleaving bonds between hydrophobic and preferably aromatic residues.[6]

Papain – cleaves after arginine or lysine that follow a hydrophobic residue; valine not tolerated on the other side of the site.[6]

Clostripan – Clostripan hydrolyses arginyl bonds and lysyl bonds at a slower rate.[6]

Thermolysin – preferentially cleaves at the N-terminus of hydrophobic residues.[6]

Cathepsin C – catalyzes excision of dipeptides from the N-terminus of protein and peptide substrates, except if (i) the amino group of the N-terminus is blocked, (ii) the site of cleavage is on either side of a proline residue, (iii) the N-terminal residue is lysine or arginine, or (iv) the structure of the peptide or protein prevents further digestion from the N-terminus.[6]

Carboxipeptidase A – hydrolyzes at the C-terminus of aromatic or aliphatic sidechains.[6]

Carboxipeptidase B – hydrolyzes at the C-terminus of arginine or lysine.[6]

TVMV – specifically cleaves between the Gly and Ser of the cleavage tag.[6]

WNV protease – specifically cleaves between Arg and Gly of the cleavage tag.[6]

Elastase – Porcine elastase I is specific for Ala-Ala and Ala-Gly bonds.[6]

Recent Protease Advancements:

- **CylA** – van der Donk et al. *J. Ind. Microbiol. Technol.*, **2018**
- **LahT150** – Nair and van der Donk et al. *eLife* **2019**;8:e42305

Tools for *In Silico* Prediction of Protease Cleavage Sites:

- **Sigma Aldrich Enzyme Explorer Protease Finder** – a tool for locating both endo and exoproteases for specific cleavage sites. By inputting relevant protein and peptide sequence surrounding a desired cleavage site, Protease Finder will select the proteolytic enzyme(s) best suited for the required hydrolysis. The Protease Finder utilizes cleavage specificity data from over 140 Sigma proteolytic enzymes, composing 73 different types of proteases. Link: https://www.sigmaaldrich.com/life-science/metabolomics/enzyme-explorer/learning-center/protease-finder.html
- **ExPASy Bioinformatic Resource Portal PeptideCutter** – a tool that predicts potential cleavage sites cleaved by proteases or chemicals in a given protein sequence. PeptideCutter return the query sequence with the possible cleavage sites mapped on it and/or a table of cleavage site positions. Link: https://web.expasy.org/peptide_cutter/
- **PROSPER Protease specificity prediction server** – an integrated feature-based webserver for *in silico* prediction of protease substrates and their cleavage sites for twenty-four different types, covering four major protease families- Aspartic (A), Cysteine ©, Metallo (M), and Serine (S). Link: https://prosper.erc.monash.edu.au/webserver.html
- **MEROPS** – an information resource for peptidases (also termed proteases, proteinases and proteolytic enzymes) and the proteins that inhabit them. The database uses an hierarchal, structure-based classification of the peptidases. In this, peptidase is assigned to a Family of the basis of statistically significant similarities in amino acid sequence, and families that are thought to be homologous are grouped together in a Clan. Link: https://www.ebi.ac.uk/merops/
- **TopFIND** – a public knowledgebase and analysis resource for protein termini and protease processing. This tool provides integration of protein termini & function with proteolytic processing, alternative transcription & translation. Link: http://clipserve.clip.ubc.ca/topfind
- **CutDB** – a tool that allows the used to browse cleavage sites in proteins of interest by disease, species, etc. Link: http://cutdb.burnham.org

Reference Websites and Publications for Common Commercially Available Proteases:

1. Oxford Genetics, *Cleavage Tag Guide* (www.oxfordgenetics.com)
2. Worthington biochemical Corporation, *Trypsin* (http://www.worthington-biochem.com/try/)
3. Chemistry LibreTexts™, Chymotrypsin (https://chem.libretexts.org)
4. ThermoFisher Scientific, *Proteases and Protein-Cleaving Reagents* (www.thermofisher.com)
5. New England BioLabs® Inc., Protease Selection Chart (https://www.neb.com/tools-and-resources/selection-charts/)
6. Mótyán, J. A.; Tóth, F.; Tőzsér, Research Applications of Proteolytic Enzymes in Molecular Biology, *Biomolecules* **2013**, 3, 923-942