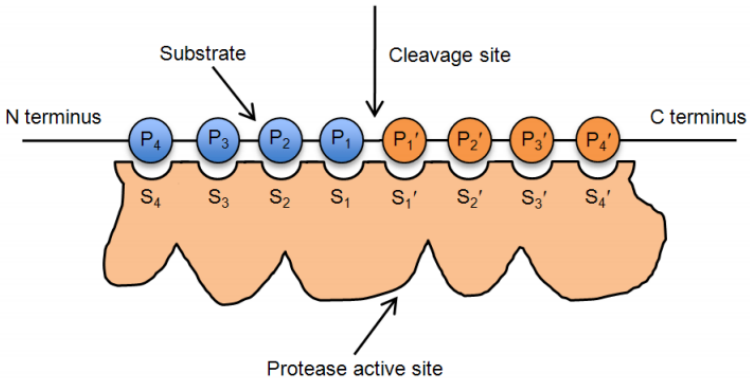
PROTEASES

INTRODUCTION

Many proteases are used as tools in protein biochemistry and mass spectrometry experiments. For example, the cleavage and removal of purification tags and trypsin digestion for peptide mapping. This document lists the target cleavage sequences for commonly used proteases along with useful information that is important for their efficient and effective use.

The system for protease cleavage site nomenclature was originally proposed by Schechter and Berger [1], [2] in which the cleavage site in the peptide substrate is between the amino acid residues designated P1 and P1'. The numbering of the amino acids then goes upwards away from the cleavage site. P2, P3, P4 etc in the N terminal direction and P2', P3', P4' etc in the C terminal direction. The amino acids that pair these from the protease are designated S1 – 4 and S'1 – 4 etc.



COMMONLY USED PROTEASES

Protease cleavage sequence for commonly used proteases		
TEV	ENLYFQG (P1' can vary but not completely [3])	cleaves after Q
PreScission	LEVLFQGP	cleaves after Q
Thrombin	LVPRGS	cleaves after R
Xa	IEGR	cleaves after R
Enterokinase	DDDDK	cleaves after K

Protease	Type	EC	Usual source and/or expression system	pH Optimum	Activators	Inhibitors	Specificity	Mr	References and Notes
Enterokinase (aka Enteropeptidase)	Serine-	EC 3.4.21.9	Duodenum E. coli S. cerevisiae	5.6		Reducing agents	DDDDK↓ P1' ≠ Pro, Trp	110kDa + 35kDa light chain is active by itself	[4]
Thrombin	Serine-	EC 3.4.21.5	Plasma CHO cells	6.8 – 8.5		Reducing agents	LVPR↓GS	32kDa + 4.5kDa	[5]
Factor Xa	Serine-	EC 3.4.21.6	Plasma HEK 293 cells	6.5		Reducing agents Chelating agents Phosphate ions	IEGR↓	42kDa + 17kDa	Very promiscuous [5], [6]
TEV Protease Tobacco Etch Virus -nuclear inclusion a endopeptidase	Cysteine-	EC 3.4.22.44	E. coli	8.0		Thiol alkylating agents	ENLYFQ↓G P1' can vary [3] P2' ≠ Pro	Catalytic domain 27kDa	Ac-TEV™ is a S219V mutant
Rhinovirus 3C Protease (aka PreScission)	Cysteine-	EC 3.4.22.28	E. coli	7.5		Thiol alkylating agents	LEVLFQ↓GP	27kDa	[7]
Carboxypeptidase A	Metallo-	EC 3.4.17.1	Bovine pancreas E. coli S. cerevisiae S. frugiperda (baculovirus)	7.0 – 8.0	None required	Reducing agents EDTA and other chelators of Zn, 2-Bz-mercaptopropionic acid Anions (citrate, P_i, oxalate) Cu^{2+}, Pb^{2+}, Fe^{3+}	H_2N-R_nY↓X-COOH X ≠ Arg, Lys, Pro Asp, Glu, Gly cleaved slowly	34.5kDa	Note 1
Carboxypeptidase B	Metallo-	EC 3.4.17.2	Porcine pancreas E. coli P. pastoris	7.0 – 9.0	None required	Reducing agents EDTA and other chelators of Zn. Heavy metals	H_2N-R_nY↓X-COOH X = basic amino acids only. e.g. Lys, Arg, Ornithine. Y = non-specific	34.3kDa	Note 2 Will cleave hydrophobic resides under certain conditions [8]
Trypsin	Serine-	EC 3.4.21.4	Bovine pancreas	7.5 – 8.5	(Stabilised by 20mM Ca^{2+})	TLCK, a_1-antitrypsin, DFP, PMSF, Aprotinin, Leupeptin, a_2-Macroglobulin	X↓Y X = Lys or Arg Y = non-specific	24kDa	Active in 1mg/mL SDS, can have chymotrypsin contamination. Notes 3 to 5

Protease	Type	EC	Usual source and/or expression system	pH Optimum	Activators	Inhibitors	Specificity	Mr	References and Notes
Chymotrypsin A_4	Serine-	EC 3.4.21.1	Porcine pancreas	7.5 – 8.5	Ca^{2+}	TPCK, DFP, PMSF, Aprotinin, a_1-antitrypsin, a_2-Macroglobulin	X↓Y X = Aromatic (Trp, Tyr, Phe), leu, Met, Ala Y = non-specific	25kDa	Active in 1mg/mL SDS Note 5
Thermolysin	Metallo-	EC 3.4.24.27	Bacillus thermoproteolyticus	7.0 – 9.0	(Stabilised by 20mM Ca^{2+})	EDTA (5mM, 40°C, 3min). Hg^{2+}, $AgNO^3$, Oxalate, Citrate, P_i, a_2-Macroglobulin	X↓YZ X = non-specific Y = Leu, Phe, Ile, Val, Met, Ala Z ≠ Pro	38kDa	Not inhibited by DFP or reagents which attach -SH groups. Retains 50% activity after 1hr at 80°C (in presence of Ca^{2+}) Note 6.
Dipeptidyl amino peptidase (DAPase) (aka Cathepsin C)	Cysteine-	EC 3.4.14.1	Kidney S. frugiperda (baculovirus)	5.5	Cl^-	Reducing agents Thiol alkylating agents	N terminal dipeptides P2 ≠ Pro, Lys, Arg P1 ≠ Pro	23kDa + 16kDa + 6Kda	Not inhibited by DFP, PCMB. Notes 7 to 9.
Endoproteinase Arg-C	Serine-	EC 3.4.21.40	Murine submaxillary glands	8.0 – 8.5	None required	DFP, TLCK, Hg^{2+}, Cu^{2+}, Zn^{2+}, a_2-Macroglobulin	Arg↓Y Y = non-specific	25kDa	Not inhibited by EDTA or hydroxyquinoline. Cleavage at other amino acids may occur at slow rate under certain conditions.
Endoproteinase Glu-C (aka V8)	Serine-	EC 3.4.21.19	Staphylococcus aureus V8	4.0 & 7.8	None required	DFP, a_2-Macroglobulin	Glu↓Y and Asp↓Y Y = non-specific Note 10	27kDa	Notes 10 and 11
Endoproteinase Lys-C	Serine-	EC 3.4.99.30	Lysobacter enzymogenes	8.5 – 8.8	None required	DFP, TLCK, Leupeptin, Aprotinin	Lys↓Y Y = non-specific	30kDa	Not inhibited by EDTA or PMSF or antitrypsin
Endoproteinase Asp-N	Metallo-	EC 3.4.24.33	Pseudomonas fragi or Flavobacterium menigosepticum.	6.0 – 8.5	Zn^{2+}	EDTA and other chelators of zinc.	Asp↓Y and Cys↓Y Y = non-specific	27kDa or 40.1kDa	Reduce and alkylate Cys to achieve Asp↓Y only cleavage

NOTES FOR TABLE ABOVE

1. Carboxypeptidase A is inactivated by freezing, lyophilisation, by pH<6.5 or by high concentrations of urea >4.5M.
2. Carboxypeptidase B is stable in SDS (1mg/mL) and in 1M urea.
3. Trypsin preparations usually contain trace amounts of chymotrypsin. If the specificity of trypsin is critical for a particular application, the trypsin should first be treated with TPCK to inactivate the chymotrypsin.
4. Trypsin is reversibly inactivated by high concentrations of urea, but is full active in urea <6.5M.
5. There are special rules for predicting the cleavage of trypsin and chymotrypsin based upon data presented in [9]. See [10] for a summary.
6. Low cleavage specificity. Relative rates of cleavage with thermolysin for P1 position. Leu > Phe > (Ile, Val, Met, Ala) > (Tyr, Gly, Thr, Ser).
7. DAPase (Cathepsin C) will not hydrolyse native proteins such as haemoglobin, fibrinogen, RNase and insulin.
8. Natural “road blocks” to DAPase (Cathepsin C) are Lys and Arg in the P2 position. This can be used to ensure only desired sequence is removed [8]. E.g. By constructing MGHHHHHHR/K will ensure only 6His tag is removed. It will of course leave an N terminal Lys or Arg.
9. DAPase (Cathepsin C) tends to be unstable during dialysis or purification procedures. Inclusion of NaCl can prevent some of the activity loss.
10. Endoproteinase Glu C exhibits differing specificity depending on the incubation buffer. In 50mM ammonium bicarbonate buffer pH 7.5 or 50mM ammonium acetate pH4.0 the protease cleaves only Glu-Y bonds. In 50mM phosphate buffer pH 7.8 the protease will cleave Asp-Y and Glu-Y bonds. If Y is a bulky hydrophobic residue cleavage is slow.
11. Endoproteinase Glu C retains full activity in 2mg/mL SDS and most of its activity in 5mg/mL SDS and 4 – 6M Urea and 5.5M Guanidine.

USEFUL LINKS

The Peptidase Database: http://merops.sanger.ac.uk/

Expasy peptide cutter tool: https://web.expasy.org/peptide_cutter/

Handbook or proteolytic enzymes:
https://www.sciencedirect.com/book/9780123822192/handbook-of-proteolytic-enzymes

PEAK PROTEINS

REFERENCES

[1] I. Schechter and A. Berger, “On the size of the active site in proteases. I. Papain,” *Biochem. Biophys. Res. Commun.*, vol. 27, no. 2, pp. 157–162, Apr. 1967.

[2] I. Schechter and A. Berger, “On the active site of proteases. III. Mapping the active site of papain; specific peptide inhibitors of papain,” *Biochem. Biophys. Res. Commun.*, vol. 32, no. 5, pp. 898–902, Sep. 1968.

[3] R. B. Kapust, J. Tözsér, T. D. Copeland, and D. S. Waugh, “The P1′ specificity of tobacco etch virus protease,” *Biochem. Biophys. Res. Commun.*, vol. 294, no. 5, pp. 949–955, Jun. 2002.

[4] Y. Kitamoto, X. Yuan, Q. Wu, D. W. McCourt, and J. E. Sadler, “Enterokinase, the initiator of intestinal digestion, is a mosaic protease composed of a distinctive assortment of domains.,” *Proc. Natl. Acad. Sci.*, vol. 91, no. 16, pp. 7588–7592, Aug. 1994.

[5] R. J. Jenny, K. G. Mann, and R. L. Lundblad, “A critical review of the methods for cleavage of fusion proteins with thrombin and factor Xa,” *Protein Expr. Purif.*, vol. 31, no. 1, pp. 1–11, Sep. 2003.

[6] J. P. Ludeman *et al.*, “Determination of the P1′, P2′ and P3′ subsite-specificity of factor Xa,” *Int. J. Biochem. Cell Biol.*, vol. 35, no. 2, pp. 221–225, Feb. 2003.

[7] R. Ullah *et al.*, “Activity of the Human Rhinovirus 3C Protease Studied in Various Buffers, Additives and Detergents Solutions for Recombinant Protein Production,” *PLoS One*, vol. 11, no. 4, p. e0153436, Apr. 2016.

[8] D. S. Waugh, “An overview of enzymatic reagents for the removal of affinity tags,” *Protein Expr. Purif.*, vol. 80, no. 2, pp. 283–293, Dec. 2011.

[9] B. Keil, *Specificity of Proteolysis*. 1992.

[10] Expasy, “Cleavage probabilities for trypsin and chymotrypsin.” .