

Review

Research Applications of Proteolytic Enzymes in Molecular Biology

János András Mótyán, Ferenc Tóth and József Tózsér *

Department of Biochemistry and Molecular Biology, Faculty of Medicine,
Medical and Health Science Center, University of Debrecen, POB 6, Debrecen H-4012, Hungary;
E-Mails: motyan.janos@med.unideb.hu (J.A.M.); tothfree@gmail.com (F.T.)

* Author to whom correspondence should be addressed; E-Mail: tozser@med.unideb.hu;
Tel./Fax: +36-52-416-432.

*Received: 15 October 2013; in revised form: 4 November 2013 / Accepted: 6 November 2013 /
Published: 8 November 2013*

Abstract: Proteolytic enzymes (also termed peptidases, proteases and proteinases) are capable of hydrolyzing peptide bonds in proteins. They can be found in all living organisms, from viruses to animals and humans. Proteolytic enzymes have great medical and pharmaceutical importance due to their key role in biological processes and in the life-cycle of many pathogens. Proteases are extensively applied enzymes in several sectors of industry and biotechnology, furthermore, numerous research applications require their use, including production of Klenow fragments, peptide synthesis, digestion of unwanted proteins during nucleic acid purification, cell culturing and tissue dissociation, preparation of recombinant antibody fragments for research, diagnostics and therapy, exploration of the structure-function relationships by structural studies, removal of affinity tags from fusion proteins in recombinant protein techniques, peptide sequencing and proteolytic digestion of proteins in proteomics. The aim of this paper is to review the molecular biological aspects of proteolytic enzymes and summarize their applications in the life sciences.

Keywords: proteolytic enzymes; proteases; molecular biology research applications

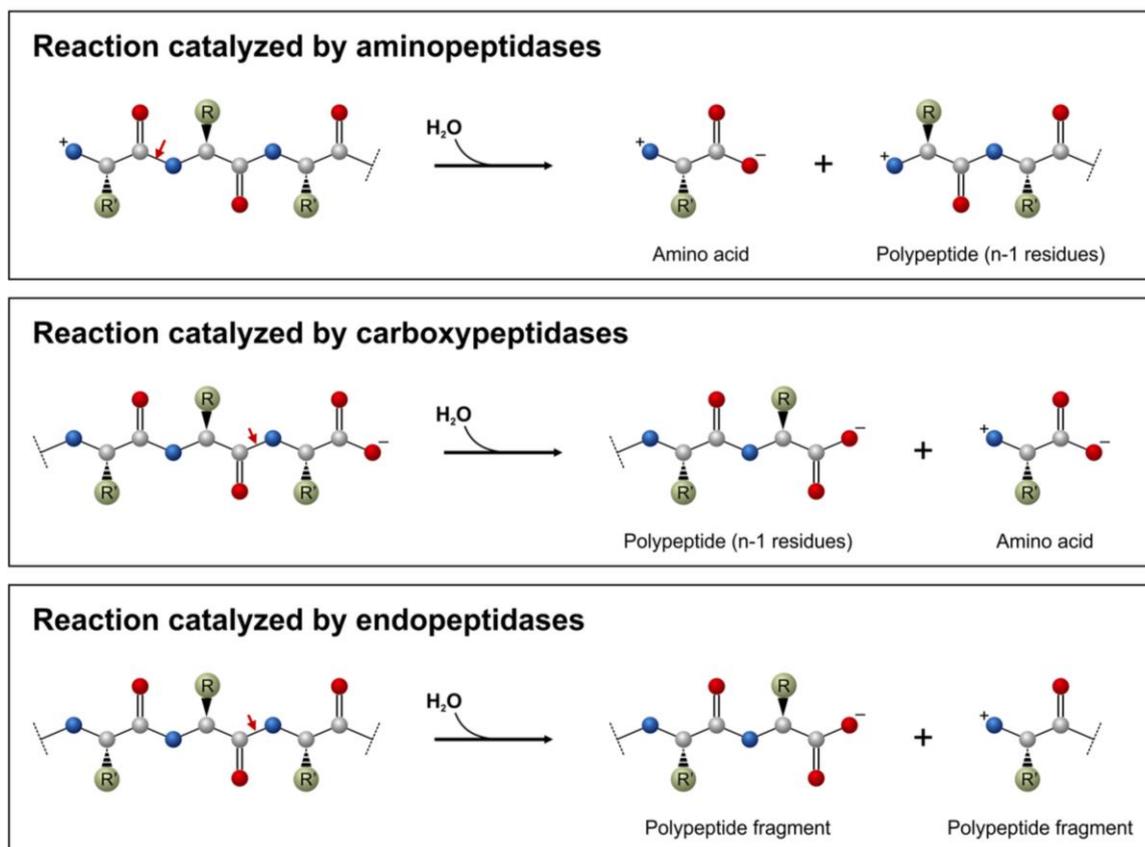
1. Scope of the Review

Proteolytic enzymes are capable of hydrolyzing peptide bonds and are also referred to as peptidases, proteases or proteinases [1].

The physiological function of proteases is necessary for all living organisms, from viruses to humans, and proteolytic enzymes can be classified based on their origin: microbial (bacterial, fungal and viral), plant, animal and human enzymes can be distinguished.

Proteolytic enzymes belong to the hydrolase class of enzymes (EC 3) and are grouped into the subclass of the peptide hydrolases or peptidases (EC 3.4). Depending on the site of enzyme action the proteases can also be subdivided into exopeptidases or endopeptidases. Exopeptidases catalyze the hydrolysis of the peptide bonds near the *N*- or *C*-terminal ends of the substrate. Aminopeptidases (Figure 1) can liberate single amino acids (EC 3.4.11), dipeptides (dipeptidyl peptidases, EC 3.4.14) or tripeptides (tripeptidyl peptidases EC 3.4.14) from the *N*-terminal end of their substrates. Single amino acids can be released from dipeptide substrates by dipeptidases (EC 3.4.13) or from polypeptides by carboxypeptidases (EC 3.4.16-3.4.18) (Figure 1), while peptidyl dipeptidases (EC 3.4.15) liberate dipeptides from the *C*-terminal end of a polypeptide chain. Endopeptidases (Figure 1) cleave peptide bonds within and distant from the ends of a polypeptide chain [2].

Figure 1. Action of aminopeptidases and carboxypeptidases removing the terminal amino acid residues as well as endopeptidases on a polypeptide substrate (having *n* residues). Red arrows show the peptide bonds to be cleaved.



Based on the catalytic mechanism and the presence of amino acid residue(s) at the active site the proteases can be grouped as aspartic proteases, cysteine proteases, glutamic proteases, metalloproteases, asparagine proteases, serine proteases, threonine proteases, and proteases with mixed or unknown catalytic mechanism [3].

The current classification system further classifies the proteases into families based on sequence similarities, furthermore, homologous families are grouped into clans using a structure-based classification [3,4]. Classification and nomenclature of proteolytic enzymes as well as a detailed description of individual proteases is available in the MEROPS database [3].

Action of the proteolytic enzymes is essential in several physiological processes, e.g., in digestion of food proteins, protein turnover, cell division, blood-clotting cascade, signal transduction, processing of polypeptide hormones, apoptosis and the life-cycle of several disease-causing organisms including the replication of retroviruses [5,6]. Due to their key role in the life-cycle of many hosts and pathogens they have great medical, pharmaceutical, and academic importance [7–9].

It was estimated previously that about 2% of the human genes encode proteolytic enzymes [8] and due to their necessity in many biological processes proteases have become important therapeutic targets [8]. They are intensively studied to explore their structure-function relationships, to investigate their interactions with the substrates and inhibitors, to develop therapeutic agents for antiviral therapies [9] or to improve their thermostability, efficiency and to change their specificity by protein engineering for industrial or therapeutic purposes [7]. Studying proteolytic enzymes is highly justified by their key role in several fields of industry [2,10–12], as well. The worldwide market of industrial enzymes was estimated to reach \$3.3 billion value in 2010 and the largest segment of this market is related to proteases [13].

Proteases are extensively applied enzymes in several sectors of industry and biotechnology, furthermore, numerous research applications require the use of them, including the production of Klenow fragments, peptide synthesis, digestion of unwanted proteins during nucleic acid purification, use of proteases in cell culture experiments and in tissue dissociation, preparation of recombinant antibody fragments for research, diagnostics and therapy, exploration of the structure-function relationships by structural studies, removal of affinity tags from fusion proteins in recombinant protein techniques, peptide sequencing, and proteolytic digestion of proteins in proteomics.

This review focuses on the application of proteolytic enzymes in life sciences, especially in the field of molecular biology. The summary table of proteases discussed in this review (Table 1) contains the substrate specificities of the enzymes which are grouped based on their catalytic mechanisms.

2. Molecular Biology Research Applications

2.1. Klenow Fragment Production

The Klenow fragment is the large fragment of the *E. coli* DNA polymerase I enzyme. While the holoenzyme has 5'→3' polymerase, 3'→5' and 5'→3' exonuclease activities, the Klenow fragment has only the polymerase and the 3'→5' exonuclease activities. The Klenow fragment has several applications in the recombinant DNA technology, including the labeling, sequencing, and site-specific mutagenesis of DNA.

Table 1. Substrate specificity of some proteolytic enzymes used in molecular biology research. Proteases are classified based on their catalytic mechanisms, furthermore, the main sources and enzyme specificities are indicated. The arrows indicate the sites of cleavages.

Enzyme	Main source	Cleavage site
Endopeptidases		
<i>Serine proteases</i>		
Trypsin	bovine	-Arg or Lys↓nonspecific-
Chymotrypsin	bovine	-Trp (or Phe, Leu, Tyr)↓nonspecific-
Enterokinase	bovine	Asp-Asp-Asp-Lys↓nonspecific-
Endoproteinase Arg-C	microbial	-Arg↓nonspecific-
Endoproteinase Glu-C	microbial	-Glu (or Asp)↓nonspecific-
Endoproteinase Lys-C	microbial	-Lys↓nonspecific-
Elastase	porcine	-Ala (or Gly or Val)↓nonspecific-
Subtilisin	microbial	-Trp (or Tyr, Phe, Leu)↓nonspecific-
Proteinase K	fungal	-aromatic, aliphatic or hydrophobic ↓nonspecific-
Thrombin	bovine	-Arg (or Lys)↓nonspecific- specific for -Leu-Val-Pro-Arg-↓Gly-Ser-
Factor Xa	bovine	-Arg (or Lys)↓nonspecific- specific for -Leu-Val-Pro-Arg-↓Gly-Ser-
WNV protease	<i>E. coli</i>	-Lys (or Arg)-Arg↓Gly-Ser-
<i>Cysteine proteases</i>		
Bromelain	plant	-nonspecific↓nonspecific-
Papain	plant	-Arg (or Lys)↓nonspecific-
Ficin (ficain)	plant	-nonspecific↓nonspecific-
Rhinovirus 3C	<i>E. coli</i>	Gly-Pro dipeptide after the scissile bond highly specific for -Leu-Glu-Val-Leu-Phe-Gln↓Gly-Pro-
TEV protease	<i>E. coli</i>	specific for -Gln-Asn-Leu-Tyr-Phe-Gln↓Gly-
TVMV protease	<i>E. coli</i>	specific for -Glu-Thr-Val-Arg-Phe-Gln↓Ser-
<i>Metalloproteases</i>		
Endoproteinase Asp-N	microbial	-nonspecific↓Asp-
Thermolysin	microbial	-Leu (or Phe)↓Leu (or Phe, Val, Met, Ala, Ile)-
Collagenase	microbial	-Pro-neutral↓Gly-Pro-
Dispase	microbial	-nonspecific↓non-polar-
<i>Aspartic proteases</i>		
Pepsin	porcine	-Phe (or Tyr, Leu, Trp)↓Trp (or Phe, Tyr, Leu)-
Cathepsin D	bovine	-Phe (or Leu)↓nonspecific (not Val, Ala)-
Exopeptidases		
<i>Serine proteases</i>		
Carboxypeptidase Y	yeast	-nonspecific↓nonspecific
<i>Cysteine proteases</i>		
Cathepsin C	bovine	removes N-terminal dipeptide
DAPase	porcine	removes N-terminal dipeptide
<i>Metalloproteases</i>		
Carboxypeptidase A	bovine	-nonspecific↓aromatic or branched preferred
Carboxypeptidase B	porcine	specific for C-terminal Arg or Lys

The enzymatic method to release the large protein fragment from the DNA polymerase I holoenzyme by proteolysis was published in 1970 [14]. Subtilisin-catalyzed proteolytic cleavage was used to produce Klenow fragment leading to the retention of the polymerase and the 3'→5' exonuclease activities and to the loss of 5'→3' exonuclease activity of the intact polymerase [6].

Nowadays, commercially available Klenow fragment is produced in recombinant ways in *E. coli* strains which carry the gene of large fragment of DNA polymerase I, therefore, the proteolytic production of Klenow fragment has mainly historical significance.

2.2. Enzymatic Peptide Synthesis

While enzymatic peptide synthesis has been frequently used to synthesize peptides for pharmaceutical and nutritional purposes (Table 2), this method is also essential for several research applications. The enzymatic method has several advantages compared to chemical methods, such as stereo specificity with side-chain protection, and the non-toxic nature of solvents coupled with the possibility of recycling the reagents used for synthesis. Enzymes have been selected considering their specificity for amino acid residues (Table 2), but this type of application is limited by the possibility of the hydrolysis of the peptide bond. The types of the enzymatic synthesis and its requirements have been reviewed [15–17]. Enzymatic peptide synthesis can be made by equilibrium- or kinetically-controlled methods.

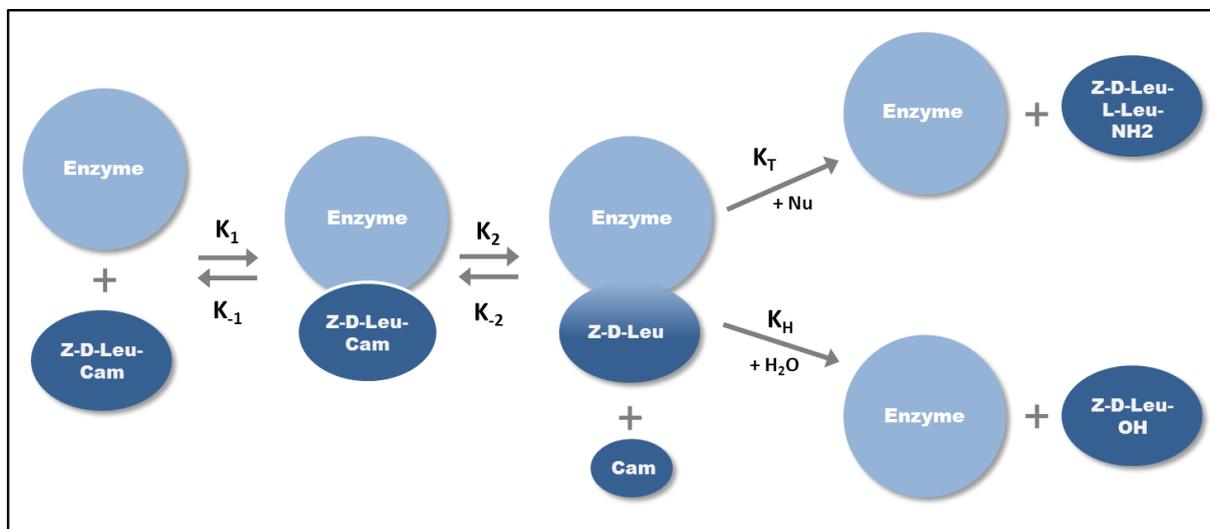
Table 2. Examples of peptides synthesized by proteases.

Peptide	Sequence	Enzyme(s)	Reference
Aspartame	Asp-Phe	Thermolysin	[18]
Nutritional peptide	Tyr-Trp-Val	α -Chymotrypsin, papain	[19]
Somatostatin	Ala-Gly-Cys-Lys-Phe-Phe-Trp- Lys-Thr-Phe-Thr-Ser-Cys	Thermolysin, chymotrypsin	[20]
Vasopressin	Tyr-Phe-Phe-Gln	Thermolysin, chymotrypsin	[21]
Oxytocin	Cys-Tyr Tyr-Ile Pro-Leu Leu-Gly	Papain, thermolysin, chymotrypsin	[21]
mouse EGF (21–31)	His-Ile-Glu-Ser-Leu-Asp- SerTyr-Thr-Cys	Papain, trypsin	[22]

2.2.1. Kinetically Controlled Peptide Synthesis

The scheme for chymotrypsin-catalyzed kinetically-controlled Z-D-Leu-L-Leu-NH₂ synthesis [23] is illustrated in Figure 2. The acyl donor Z-D-Leu that is activated by carbamoylmethyl (Cam) ester and chymotrypsin (E) form the enzyme-substrate complex first and after that the covalently linked Z-D-Leu-E intermediate with the loss of the carbamoylmethyl ester. If this intermediate is attacked by water, hydrolysis occurs, which results in the Z-D-Leu-OH fragment. However if a more powerful nucleophile (e.g., alcohol or thiol) is present in the media, the enzyme produces a peptide bond instead of the cleavage [15] and the Z-D-Leu-L-Leu-NH₂ dipeptide may be formed in the presence of H-L-Leu-NH₂ nucleophile (Figure 2). The product yield depends on the kinetics of the two nucleophilic reactions, however, the reaction is faster and requires lower substrate enzyme ratios compared to the equilibrium-controlled synthesis, due to the activated acyl donor.

Figure 2. Kinetically-controlled synthesis of Z-D-Leu-L-Leu-NH₂ dipeptide. After the formation of enzyme-substrate complex (K_1) a covalent enzyme-substrate intermediate is formed (K_2). The intermediate is subjected to the attack from H₂O or other nucleophiles (Nu). K_H is the equilibrium constant of hydrolysis, K_T is the equilibrium constant of the transferase reaction.



The acyl donor activating agent must not only be an ester, but an amide or a nitrile as well. Only serine or cysteine proteases can be used to perform the kinetically controlled peptide synthesis, as these enzymes can act as transferase and hence are able to catalyze the transfer of an acyl group from the acyl donor to the nucleophile through the formation of a covalent acyl enzyme intermediate. Papain, thermolysin, trypsin and α -chymotrypsin are mostly used for kinetically-controlled peptide synthesis [24]. The yield of peptide product will depend on the apparent ratio of transferase to hydrolase rate constants $(K_T/K_H)_{app}$ and the rate at which the peptide product is hydrolyzed. The $(K_T/K_H)_{app}$ values of proteases used for kinetically-controlled synthesis are in the range of 10^2 – 10^4 [17].

2.2.2. Equilibrium-Controlled Synthesis

In the case of the equilibrium-controlled synthesis the process is the reverse of hydrolysis. Important problems of this enzymatic method of peptide synthesis are the low reaction rates and the need of increased yield because proteases do not alter the equilibrium of the reaction. A high amount of the enzyme is often required together with the precise reaction conditions to drive the equilibrium towards synthesis [15]. A higher rate of peptide bond formation can be reached by the appropriate pH of the reaction medium (changing the equilibrium of ionization) but there are several other ways to increase the yield of the reaction:

(a) Precipitation of the product is the classical method. When certain soluble carboxyl amine components are used as reactants the products will precipitate and the reduced concentration of soluble products will drive the equilibrium towards synthesis. In this case the yield of the product depends on the concentration of the starting materials and can be determined by the solubility of the product [17].

(b) In a biphasic system the enzyme acts in an aqueous environment that is surrounded by a water immiscible medium; the water content of the system is around 2–5%. The reactants (dissolved in high

concentrations) can diffuse from the organic phase into the water until the equilibrium is reached; the enzyme-catalyzed synthesis is followed by the diffusion of the products back into the organic phase. The organic phase reduces the dielectric constant of the medium and thus the acidity of the carboxyl group of the acyl donor as well, which in turn promotes the synthesis of the peptide bond instead of hydrolysis [24]. This method is applicable only for the synthesis of water insoluble products.

(c) The dissolved state system can be used for the synthesis of water-soluble products (short peptides, high molecular weight peptides and proteins). In this environment forcing the reaction towards peptide synthesis requires the mass action, the addition of a water-miscible organic co-solvent in high concentration or the excess of one reactant. In serine protease-catalyzed reaction in water, the rate determining feature is the acylation of the enzyme while the product yield at equilibrium depends on the partition coefficient and the ratio of the aqueous and organic volumes [17]. The chymotrypsin- and subtilisin-catalyzed synthesis of N-Bz-L-Tyr-L-Leu-NH₂ is more efficient in hydrophobic organic solvents; adding water in sub-saturating concentration increases the yield of the chymotrypsin-catalyzed peptide synthesis [25].

2.2.3. Strategies Used in Enzymatic Synthesis

The use of enzymes in organic solvents have several advantages compared to aqueous solvents which have led to their widespread application: the thermodynamic equilibrium can be shifted towards synthesis, the undesirable side reactions can be reduced, the nonpolar substrates are more soluble in organic solvents, the separation process and enzyme recovery is more effective in a low water-containing environment [26–29]. Many proteases, such as thermolysin, subtilisin and α -chymotrypsin [26,29] can maintain their active conformation in organic solvents and show good functionality in the synthesis of aspartame and demorphin derivatives.

However, the use of enzymes in organic solvent has also disadvantages such as the unfavorable effects of the organic solvents on enzyme activity and stability. The modification of biocatalysts by protein engineering [30,31] and/or chemical modification or the use of naturally solvent-tolerant proteases [32] for peptide synthesis is a developing field. The driving force of this field is the aim of making biocatalysts with proper features to suit them for the reactions under specific synthesis conditions. Site-directed mutagenesis is a very effective tool and can be used by protein engineers to screen mutants with enhanced stability, activity or specificity, furthermore, this method can be used to explore structure-function relationships (rational design).

Subtilisin has been extensively studied and engineered via site-directed mutagenesis to make it more capable of peptide bond formation in aqueous solution [33]. Single and multiple mutations have been introduced into subtilisin to increase its stability and make it more resistant against oxidizing agents, thermal denaturations and inactivation effects of polar solvents [28].

Thermolysin has a higher synthesis rate compared to the solvent stable PST-01 protease from *Pseudomonas aeruginosa*. Considering the high structural similarity of these enzymes, the synthetic activity of PST-01 protease was increased by the Y114F mutation [31], moreover, the Y114R and Y114S mutations resulted in better activity enhancement.

Chemical modification is also an efficient method to modify the properties of enzymes used for peptide synthesis. Thiol-subtilisin, in which the serine residue has been chemically changed to cysteine

at the active site, shows an enhanced aminolysis to hydrolysis ratio in aqueous solution and in dimethyl sulphoxide. The stability of proteases can also be increased by chemical modification e.g., a hydrophilic carbohydrate-polyacrylate polymer coat can make the enzymes highly active and stable in polar solvents and more resistant against thermal inactivation [28].

Immobilization of proteases is the most frequently applied method for the recovery of products without great loss of the catalysts, which greatly decreases the cost of the synthesis. This approach also ensures better operational stability of biocatalysts and control of the reaction. Enzyme immobilization techniques can be divided into five groups: (a) covalent attachment to solid support; (b) absorption on solid support; (c) entrapment in polymeric gel; (d) crosslinking with bifunctional reagents and (e) encapsulation [17,34].

Substrate engineering means the manipulation of the leaving group and is a powerful tool to increase the specificity of the proper enzyme and/or increase the rate and the yield of the reaction [24].

Protease-catalyzed synthesis of stereochemically modified peptides is also a preferable application compared to chemical synthesis due to stereospecificity of the proteases.

Proteases can bind not only natural substrates, but also specifically designed substrate mimetics, which are also very useful tools to increase the yield of peptide synthesis. Substrate mimetics can bind to the active site of the enzyme and in this way proteases can be used for the synthesis of products containing non-specific amino acids. The undesired cleavage of the newly synthesized peptide bonds can be avoided using this method and it is not required to change the properties of the medium or the enzyme [17].

The production of peptides with amides at their C-termini is a great challenge for enzymatic peptide synthesis, but amidation may be required to retain biological activity.

2.3. Nucleic Acid Isolation

Generally, the first step of nucleic acid isolation protocols is the lysis of the biological material containing the DNA or RNA of interest. Before the purification and concentration of nucleic acids the contaminating proteins and other macromolecules have to be removed from the sample. Undamaged nucleic acids can be isolated when the degradation of the DNA and RNA present in the sample is avoided by the inhibition and removal of DNases and RNases. The nucleases can be inhibited by the addition of chelators (e.g., EDTA) which bind the ions essential for their action. Besides the inactivation of nucleases, proteolytic enzymes are applied during the nucleic acid isolation to remove total protein content of the sample.

The most widely used proteolytic enzyme in nucleic acid purification is the Proteinase K, which was described in 1974 [35]. Proteinase K is a non-specific serine endopeptidase which can catalyze the cleavage of peptide bonds at the carboxylic side of aromatic, aliphatic, or hydrophobic amino acid residues. Besides the digestion of unwanted proteins, Proteinase K also quickly inactivates the nucleases which might degrade the nucleic acids present in the sample [36]. This proteolytic digestion decreases the level of contaminants in the nucleic acid extract and prevents nucleic acids from degradation leading to a higher yield of the DNA or RNA to be isolated.

2.4. Cell Isolation and Tissue Dissociation

Cell biology studies frequently require the dissociation of primary tissues and the isolation of viable cells for tissue culturing. The most common method for cell isolation is the enzymatic digestion of the junctions connecting the cells and the components of the surrounding extracellular matrix, by which the cells can be released from a wide variety of tissues. Several enzymes are available in the market for the detachment of cultured cells, cell dissociation and cell component or membrane-associated protein isolation [37,38]. Besides the polysaccharidases, nucleases and lipases, the proteases are the most important enzymes used widely to dissociate cells from tissues.

The experimental conditions of cell isolation are functions of several parameters, including the type of the tissue and the source of its origin. Cells with high viability can be isolated in high yield using a suitable enzyme or the optimal combination of enzymes. As proteases differ in their specificities, different enzymes are recommended to be used for most effective tissue disruption, depending on the origin and type of the tissue. We describe below the enzymes most commonly used for cell isolation.

The matrix metalloproteinase collagenase was first isolated in 1953 [39]. This endopeptidase can digest the collagenous extracellular matrix in a zinc-dependent manner. Collagenase cleaves the peptide bonds within the triple helices of native collagen, between a neutral amino acid and Gly within the Pro-X-Gly-Pro sequence. This sequence can be found most frequently in the collagen; therefore collagenases digest other proteins less efficiently. A commercially available collagenase (clostridiopeptidase A) is produced by *Clostridium histolyticum*, and it is capable of digesting collagen fibers very effectively. Solutions supplied for tissue dissociation contain collagenase and other additional proteinases which can digest the components of the extracellular matrix [38,40]. The serine protease elastase is a unique enzyme which can cleave the peptide bonds in elastin, therefore, it is generally used to dissociate tissues containing a high amount of elastin connective fibers. Elastase cleaves peptide bonds next to smaller neutral amino acids and besides its protease activity it also has esterase and amidase activities. Papain is a cysteine peptidase of *Carica papaya* latex. Papain, similarly to elastase, also has amidase and esterase activities and has a broad specificity. Papain has less damaging effects on tissues and therefore it is typically applied for cell dissociation of neuronal tissues. Besides cell dissociation, papain is also widely used for integral membrane protein solubilization and for digestion of proteoglycans. The serine protease trypsin is a very specific proteinase cleaving the peptide bonds at the C-terminal end of positively charged Lys and Arg side chains. Due to the high specificity of trypsin the digestion of tissue proteins is less effective and it is generally used for tissue dissociation together with other proteolytic enzymes. Serine protease chymotrypsin cleaves peptide bonds preferentially at the carboxyl side of aromatic Tyr, Trp and Phe residues. Chymotrypsin is less widely used for tissue dissociation; the use of other additional proteases is required for efficient digestion. The Zn-metalloprotease dispase is also a neutral protease. This non-specific protease cleaves the peptide bonds of proteins at the amino side of non-polar amino acid residues.

2.5. Cell Culturing

Cells isolated from a tissue can be cultured separately from the organism in cell culture flasks using appropriate growth medium. Adherent cells grown in a cell culture flask are attached to the surface by

protein bridges which have to be disrupted during passaging. The cells can be released from the cell flask surface mechanically using a cell scraper or can be detached by a protease treatment using trypsin solution.

Trypsinization means the process used for the detachment of adherent cells using trypsin solution to digest the adhesion molecules by which the cells are attached to the surface of the culture flask. Trypsin solutions generally contain EDTA to reduce the concentration of metal ions that might inhibit trypsin.

Although trypsinization is the most commonly used method to detach adherent cells from cell culture flasks, the effects of this protease treatment were only recently studied in detail. It was found that trypsinization can affect the extracellular matrix surrounding the cells [41] and has physiological effects on cells grown in cell cultures [42]. Trypsin treatment can lead to cleavage of membrane proteins and receptors, which can cause significant changes in the expression level of different proteins: level of growth- and metabolism-related protein expressions were found to be down-regulated after trypsinization, while up-regulation of apoptosis-related protein expressions was seen after the protease treatment [42,43]. This effect should be taken into account when trypsinization is involved in experimental design.

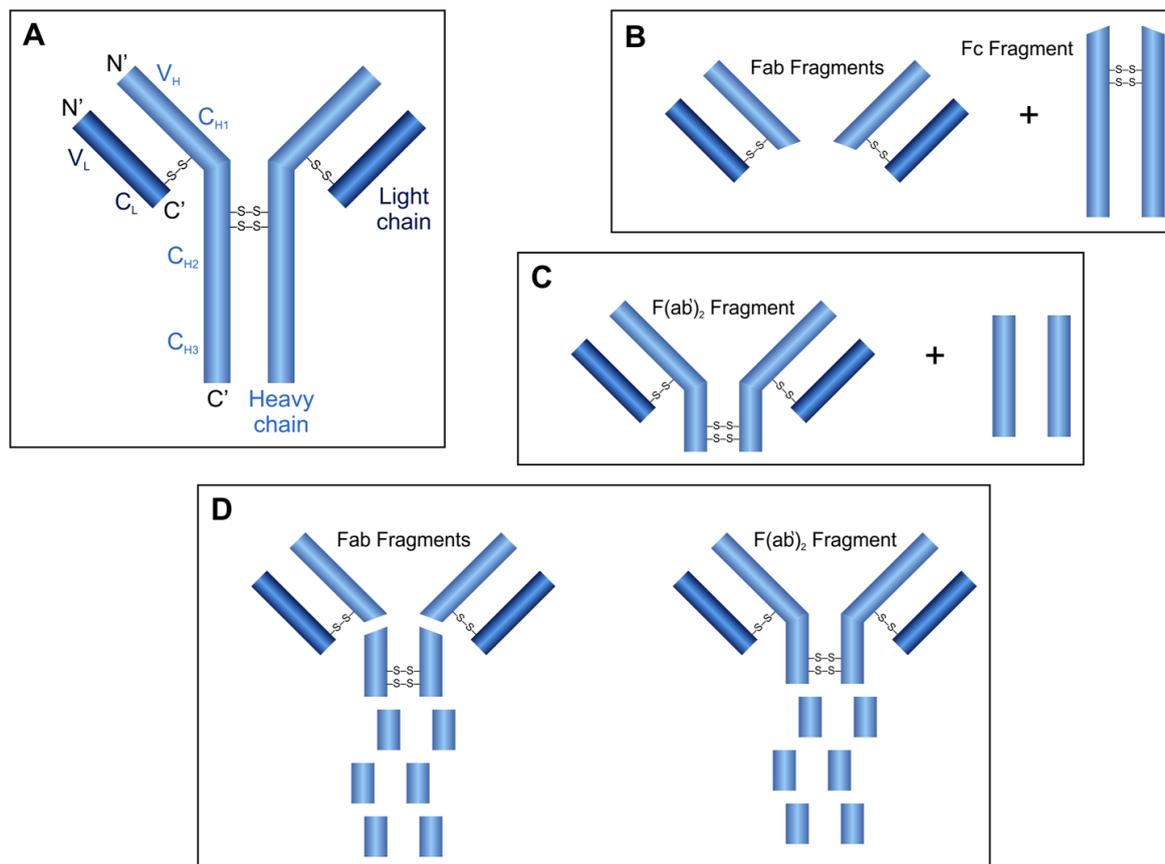
2.6. Antibody Fragment Production

Antibody molecules are produced by the immune system against foreign substances and are classified into the immunoglobulin superfamily of the proteins (Figure 3A). They consist of four polypeptide chains, two identical heavy chains (H) and two identical light chains (L) which are connected by disulfide bridges. Both the H and the L chains contain variable (V_H and V_L) and constant (C_{H1} , C_{H2} , C_{H3} and C_L) regions, respectively. The V_H and V_L chains, containing hypervariable regions, are responsible for the antigen-antibody interactions and determine the antigen specificity [6].

Fragments of the monoclonal antibodies are widely used in diagnostics, therapeutics and in biopharmaceutical research [44–46] having beneficial properties compared to the whole immunoglobulin molecules due to their smaller size and lower immunogenicity [44]. Fragments of whole immunoglobulin molecules can be produced using recombinant DNA technology or can be generated by enzymatic digestion. Here we discuss the proteolytic antibody fragmentation method.

Generally, the papain, pepsin and ficin proteases are used for the specific digestion of IgG molecules. Digestion of an antibody by the cysteine protease papain results in three fragments due to the cleavage of peptide bonds in the hinge region between C_{H1} and C_{H2} domains: one Fc (crystallizable) and two identical Fab (antigen binding) fragments are released (Figure 3B). While both released Fab fragments carry one antigen-binding site, the Fc fragment does not have antigen-binding ability. The aspartic acid protease pepsin cleaves the peptide bonds of the antibody near the disulfide bonds connecting the H chains (Figure 3C). This digestion results in the release of the peptides of the Fc region and one $F(ab')_2$ fragment containing both antigen binding sites. The cysteine protease ficin can release both $F(ab')_2$ or Fab fragments (Figure 3D), depending on the cysteine concentration [44].

Figure 3. Structure of IgG antibody molecules (A) and fragments released after proteolytic digestion using papain (B), pepsin (C) or ficin (D).



The *in vivo* applications of the large, pentameric IgM immunoglobulins could be restricted because of their large size. Proteolytic fragmentation of IgM molecules is also a useful method to produce smaller, active Fab, Fab', F(ab')₂ or “IgG-like” fragments which have beneficial properties compared to the whole IgM molecules. Trypsin and pepsin are useful enzymes for IgM fragmentation, pepsin can generate Fab and F(ab')₂ fragments, while trypsin can generate “IgG-like” fragments, as well. Papain cannot be used for IgM fragmentation because of the production of heterogeneous fragments.

Digestion of monoclonal antibodies by papain or pepsin is still used to produce Fab or F(ab')₂ fragments [44]. There are several products available on the market obtained by proteolytic cleavage of antibodies, furthermore, several companies supply fragmentation kits for antibody fragment preparation. In some cases it could be preferable to produce the Fab fragments in high quality by recombinant expression in cell lines and obtain the fragment in sufficient quantity. For example, the Fab fragments produced by recombinant expression were found to be more preferable for the crystallization experiments because of the higher quality of the protein sample [47].

Antibody fragments have several beneficial properties for *in vivo* applications compared to whole antibody molecules. Due to their smaller size they have higher mobility and can penetrate tissues or permeate cell membranes more easily. As antigen-binding fragments lack the Fc fragments, they have lower immunogenicity, as they contain only their antigen-binding domain sites (Fab, Fab' or F(ab')₂ fragments) and do not carry the regions responsible for antibody effector functions, furthermore, they do not form large immunocomplexes.

The antigen-binding fragments of antibodies have great importance from the viewpoint of clinical and therapeutic applications. Antibody fragments could be administered to prevent the development of a disease (e.g., restenosis), could be applied during the diagnosis (e.g., metastatic breast and colon cancer) or for the treatment of some diseases (e.g., macular degeneration) or to detect toxins or neutralize snake venoms [44,45]. The current number of antibody-based therapeutics approved by the FDA is 35, while several other antibodies are in clinical trials. The relevance of antibody fragments in structural studies is discussed in detail in the following paragraph. Further important biotechnological applications of antibody fragments, e.g., as immunodetection, immunopurification and detoxification, have been reviewed in the recent past [46].

2.7. Structural Studies

Crystallographers have great challenges in the structure determination of transmembrane proteins. It is difficult to crystallize these proteins due to their high molecular flexibility, hydrophobic surfaces and low solubility. Antibody fragments are very useful tools for solving these problems. Specific binding of antibody fragments can increase the overall hydrophilicity of proteins and the solubility of the transmembrane proteins, furthermore, they can decrease the flexibility and stabilize the conformation of the molecule [48]. The structures of the membrane proteins co-crystallized with antibody fragments can be determined at higher resolution because these crystals have a higher diffraction quality [49–51].

The proteases have great indirect significance in structure determination, because the antibody fragments produced by proteolytic cleavage of whole immunoglobulins can be used to improve the crystallization properties [52–54].

The proteases are not only tools for crystallographers but are also important target molecules for structural biologists and have great relevance in antiviral therapies, drug and therapeutics development from the viewpoint of structural biology. Besides interest in proteases with unknown structure, the results of structural studies can help researchers evaluate the structure-function relationships more efficiently; increasing knowledge on the structural organization of viral proteases can help to explore their action, perform comparative studies by which we can better understand the structure-function and evolutionary relationships and recognize general or specific features [55–57]. One of the main driving forces of structure determination of proteases is the need for the development of efficient drugs for antiviral therapies [7,9,58–60]. Both structural and enzymatic inhibition studies are required for the structure-based drug development of protease inhibitors [61]. Structural data can also help protein engineers to alter the specificity and to improve the enzymatic properties of proteases by structure-guided mutagenesis [62–64] for several purposes.

2.8. Fusion Tag Removal

The proteins produced by recombinant techniques are typically linked with a fusion partner termed a fusion tag. The introduction of a fusion tag means the fusion of an additional protein or peptide to the recombinant protein. These fusion tags are extensively used from basic research to high-throughput structural biology owing to the several advantages they provide in the expression of different recombinant proteins [65,66]. The tags largely aid the detection and purification of proteins; moreover they also could have a favorable effect on protein yield and/or solubility. Tags can prevent proteins

from proteolytic digestion, can protect antigenicity or facilitate the folding of the fusion protein. On the other hand, they can also negatively alter solubility, structural integrity and biological activity [47,67,68] or may cause a disadvantage for further application of the protein, so the removal of a tag can be crucial [66].

Commonly used affinity tags and fusion protein partners are the hexahistidine-tag (His₆), FLAG-tag, maltose binding protein (MBP), glutathione S-transferase (GST), thioredoxin (TRX), small ubiquitin-like modifier (SUMO), ubiquitin (Ub) and green fluorescent protein (GFP) [69].

In some cases the tag can be removed by a chemical treatment but those methods are rather unspecific compared to enzymatic cleavages and may lead to protein denaturation and/or side chain modifications of amino acids in the target protein. The specificity and detergent sensitivity of common proteases used for tag removal have been examined and reviewed previously [70–73]. Both endo- and exo-proteases could be suitable for fusion tag removal.

2.8.1. Endoproteases

Serine proteases such as enterokinase (also referred to as enteropeptidase), factor Xa and thrombin have been widely used for many years to remove N-terminal tags, but several cases have been reported in which they cleaved not only at the desired cleavage site but also in the protein of interest. These incidents led to the extensive application of viral proteases like human rhinovirus (HRV) 3C protease and tobacco etch virus (TEV) protease. While sequences recognized by a cellular serine protease and the viral proteases could be similar, the viral proteases cleave the protein substrates at the undesired sites less efficiently due to their high specificity and low catalytic rate, moreover, recombinant viral proteases can be produced in high quantities in *E. coli* [73]. These findings and the limited activity of the generally used serine proteases in some detergents, which are needed to study the membrane proteins, inspired the search for other viral proteases for tag removal, such as Tobacco Vein Mottling Virus (TVMV) protease [74], West Nile Virus (WNV) protease [75] and some alphaviral proteases: Venezuelan Equine Encephalitis Virus (VEEV) protease, Semliki Forest Virus (SFV) and Sindbis Virus (SIN) protease [76].

2.8.2. Exopeptidases

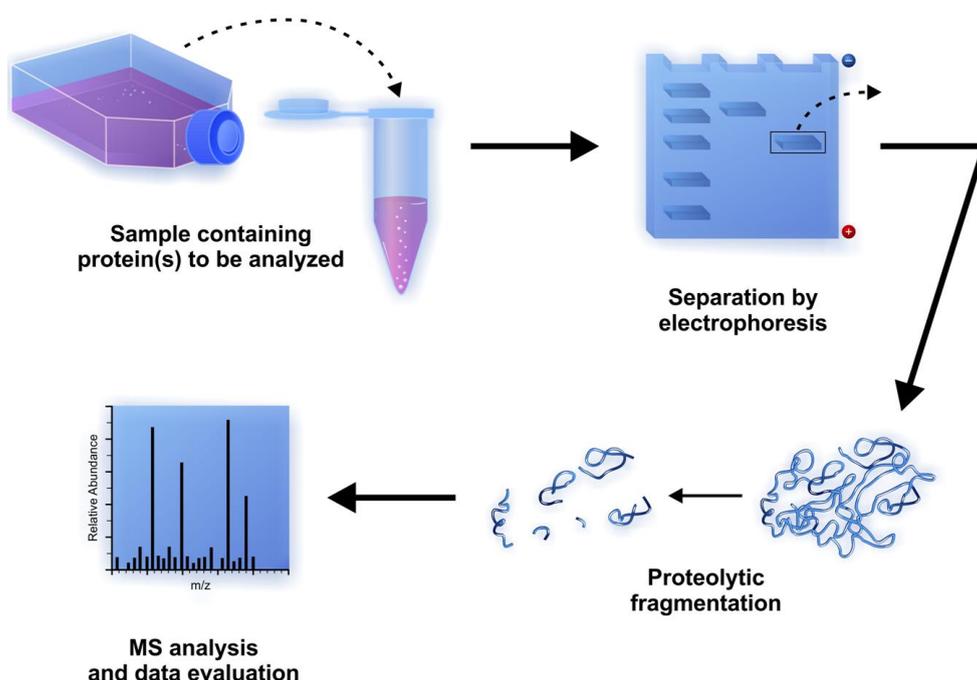
Amino peptidases and carboxypeptidases are not as widely used as endopeptidases, as they frequently leave amino acid residues on the target protein, while it is easier to design cleavage sites with endopeptidases not to leave extra residues after the cleavage. However, if it is still desired to add a tag onto the C-terminal of the target protein a carboxypeptidase may be used for its removal. Among metallo-carboxypeptidases, type A carboxypeptidases remove mostly aromatic or branched aliphatic side chain containing amino acids, while type B carboxypeptidases prefer basic amino acids. Carboxypeptidases can be utilized to remove a His₆ tag from the C-terminal end of a protein [77]. Dipeptidyl aminopeptidase (DAPase) is a useful enzyme for the removal of N-terminal dipeptides.

2.9. Proteomic Applications

Proteomic studies are made with the aim to identify, characterize, and quantify the required samples and typically involve mass spectrometric (MS) analysis. Besides determination of the composition of protein complexes, chemical properties, post-translational modifications and structural properties of proteins can also be revealed by MS.

MS is a powerful analytical method to measure the mass of proteins or peptides by the analysis of mass-to-charge ratio (m/z) ratio. Generally, the samples to be analyzed contain a mixture of various proteins and/or polypeptides that have to be separated and digested into smaller fragments before the MS analysis. Protein separation can be performed efficiently by polyacrylamide gel electrophoresis. The separated proteins can be digested after the electrophoresis by chemical cleavage or by enzyme-catalyzed digestion of peptide bonds. The process in which the bands or spots are cut out from the gel followed by the addition of protease(s) to the gel containing the protein(s) of interest is called in-gel digestion [78]. Proteolysis of whole proteins leads to the release of smaller peptides with different molecular masses which are suitable for MS analysis (Figure 4).

Figure 4. Steps of proteomic analysis using mass-spectrometry after separation and in-gel digestion of proteins of interest.



Trypsin is the most widely used proteolytic enzyme for protein digestion in MS analysis. Due to its high specificity it is easy to predict the cleavage sites and to compare the results of experimental enzymatic and theoretical *in-silico* digestion. The Asp-N, Lys-C, Arg-C and Glu-C enzymes are also highly sequence-specific endoproteases but they are less active [78]. Chymotrypsin, pepsin, papain, elastase, subtilisin, proteinase K, thrombin, factor Xa and some other proteases are also suitable enzymes for fragmentation in MS analysis [79].

Several proteases are available for this fragmentation of which specificities are well established [79,80]. The number and length of released peptide fragments depends on the protease(s) applied for selective

proteolysis of the targeted protein. Applications developed for *in silico* protein fragmentation are useful to predict the proteolytic fragments and to choose the most proper enzyme for the most effective digestion based on the enzyme specificities (<http://prospector.ucsf.edu>).

Generally, highly sequence-specific proteases are preferred for protein fragmentation instead of less specific enzymes, as the latter ones produce a very complex mixture of fragments. In the case of the efficient fragmentation the peptide, the fragments have a proper length and are released in high yield, and the complete sequence of the whole protein can be covered by the analysis of the proteolytic fragments.

3. Summary

Besides extended application for nutritional and pharmaceutical purposes, proteases from natural sources are also widely used tools in molecular biology practice. Their degradative properties make them useful for general protein digestion in tissue dissociation, cell isolation, and cell culturing. The specificity and the predictability of cleavages by proteases enables their use for more specific tasks such as antibody fragment production, the removal of affinity tags from recombinant proteins and specific protein digestion in the proteomics field mainly for protein sequencing. Moreover, the already mentioned specificity makes proteases—in a water restricted environment—able to synthesize the peptide bonds instead of hydrolyzing them. This property combined with their enantioselectivity has also promoted their use in peptide synthesis.

The expansion of knowledge has assisted the increase of applications of proteolytic enzymes for several purposes, and the application fields are widening with the help of protein engineering techniques and by chemical modification of the enzymes [7,62]. Studies made with the aim to better understand the structure and function of existing proteolytic enzymes and to obtain new, engineered proteases with altered properties for therapeutic, industrial or research fields require the use of the applications discussed in this paper.

Acknowledgments

This work was supported by the TÁMOP 4.2.2.A-11/1/KONV-2012-0023 “VÉD-ELEM” and by the Hungarian Science and Research Fund (OTKA 101591) implemented through the New Hungary Development Plan co-financed by the European Social Fund and the European Regional Development Fund.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Barrett, A.J.; McDonald, J.K. Nomenclature: Protease, proteinase and peptidase. *Biochem. J.* **1986**, *237*, 935.
2. Rao, M.B.; Tanksale, A.M.; Ghatge, M.S.; Deshpande, V.V. Molecular and biotechnological aspects of microbial proteases. *Microbiol. Mol. Biol. Rev.* **1998**, *62*, 597–635.

3. Rawlings, N.D.; Barrett, A.J.; Bateman, A. MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* **2012**, *40*, D343–D350.
4. Rawlings, N.D.; Barrett, A.J. Evolutionary families of peptidases. *Biochem. J.* **1993**, *290*, 205–218.
5. Neurath, H.; Walsh, K.A. Role of proteolytic enzymes in biological regulation (a review). *Proc. Natl. Acad. Sci. USA* **1976**, *73*, 3825–3832.
6. Devlin, T.M. *Textbook of Biochemistry with Clinical Correlations*, 5th ed.; Wiley & Sons: New York, NY, USA, 2002.
7. Li, Q.; Yi, L.; Marek, P.; Iverson, B.L. Commercial proteases: Present and future. *FEBS Lett.* **2013**, *587*, 1155–1163.
8. Craik, C.S.; Page, M.J.; Madison, E.L. Proteases as therapeutics. *Biochem. J.* **2011**, *435*, 1–16.
9. Antonelli, G.; Turriziani, O. Antiviral therapy: Old and current issues. *Int. J. Antimicrob. Agents* **2012**, *40*, 95–102.
10. Kirk, O.; Borchert, T.V.; Fuglsang, C.C. Industrial enzyme applications. *Curr. Opin. Biotechnol.* **2002**, *13*, 345–351.
11. Rani, K.; Rana, R.; Datt, S. Review on latest overview of proteases. *Int. J. Curr. Life Sci.* **2012**, *2*, 12–18.
12. Ray, A. Protease enzyme- potential industrial scope. *Int. J. Technol.* **2012**, *2*, 1–4.
13. Sarrouh, B.; Santos, T.M.; Miyoshi, A.; Dias, R.; Azevedo, V. Up-to-date insight on industrial enzymes applications and global market. *J. Bioprocess. Biotech.* **2012**, *S4*:002
14. Klenow, H.; Henningsen, I. Selective elimination of the exonuclease activity of the deoxyribonucleic acid polymerase from *Escherichia coli* B by limited proteolysis. *Proc. Natl. Acad. Sci. USA* **1970**, *65*, 168–175.
15. Morihara, K. Using proteases in peptide synthesis. *Trends Biotechnol.* **1987**, *5*, 164–170.
16. Bhalla, T.C.; Kumar, D.; Gajju, H.; Agrawal, H.O. Thermophilic bacterial proteases. *J. Punjab Acad. Sci.* **1999**, *1*, 77–91
17. Kumar, D.; Bhalla, T.C. Microbial proteases in peptide synthesis: Approaches and applications. *Appl. Microbiol. Biotechnol.* **2005**, *68*, 726–736.
18. Kühn, D.; Dürrschmidt, P.; Mansfeld, J.; Ulbrich-Hofmann, R. Biolysin and thermolysin in dipeptide synthesis: A comparative study. *Biotechnol. Appl. Biochem.* **2002**, *36*, 71–76.
19. Kimura, Y.; Muraya, K.; Araki, Y.; Matsuoka, H.; Nakanishi, K.; Matsuno, R. Synthesis peptides consisting of essential amino acids by a reactor system using three proteinases and an organic solvent. *Agric. Biol. Chem.* **1990**, *54*, 3331–3333.
20. Bille, V.; Ripak, C.; van Aasche, I.; Forni, I.; Degelaen, L.; Searso, A. A Semi-Enzymatic Synthesis of Somatostatin. In Proceedings of 21st European Peptide Symposium. ESCOM, Leiden, The Netherlands, 1991; Giralt, E., Andreu, D., Eds.; pp. 253–254.
21. Rizo, J.; Gierarch, L.M. Constrained peptides: Models of bioactive peptides and protein structures. *Ann. Rev. Biochem.* **1992**, *61*, 387–418.
22. Widmer, F.; Bayne, S.; Houen, G.; Rigby, R.B.; Whittaker, R.G.; Johansen, J.T. Use of Proteolytic Enzymes for the Synthesis of Fragments of Mouse Epidermal Growth Factor. In Proceedings of Peptides 1984, Almquist, Stockholm, 1985; Ragnursson, U., Ed.; pp. 193–196.

23. Salam, S.M.; Kagawa, K.; Kawashiro, K. Alpha-chymotrypsin-catalyzed peptide synthesis using N-protected D-amino acid carbamoylmethyl esters as acyl donors. *Biotechnol. Lett.* **2005**, *27*, 1199–1203.
24. Guzmán, F.; Barberis, S.; Illanes, A. Peptide synthesis: Chemical or enzymatic. *Electron. J. Biotechnol.* **2007**, *10*, 279–314.
25. Sergeeva, M.V.; Paradkar, V.M.; Dordick, J.S. Peptide synthesis using proteases dissolved in organic solvents. *Enzyme Microb. Technol.* **1997**, *20*, 623–628.
26. Dordick, J.S. Enzymatic catalysis in monophasic organic solvents. *Enzyme Microb. Technol.* **1989**, *11*, 194–211.
27. Khemlnitski, Y.L.; Levashov, A.V.; Klyachko, N.L.; Martinek, K. Engineering biocatalytic systems in organic media with low water content. *Enzyme Microb. Technol.* **1988**, *10*, 710–724.
28. Gill, I.; Fandino, R.L.; Jobra, X.; Vulfson, E.N. Biologically active peptides and enzymatic approaches to their production. *Enzyme Microb. Technol.* **1996**, *18*, 162–183.
29. Shen, H.Y.; Tian, G.L.; Ye, Y.H. Synthesis of demorphin (1–4) derivatives catalysed by proteases in organic solvents. *J. Pept. Res.* **2004**, *65*, 143–148.
30. Wells, J.A.; Estell, D.A. Subtilisin: An enzyme designed to be engineered. *Trends Biochem. Sci.* **1988**, *13*, 291–297.
31. Ogino, H.; Tsuchiyama, S.; Yasuda, M.; Doukyu, N. Enhancement of the aspartame precursor synthetic activity of an organic solvent-stable protease. *Protein Eng. Des. Sel.* **2010**, *23*, 147–152.
32. Xu, J.X.; Jiang, M.; Sun, H.L.; He, B.F. An organic solvent-stable protease from organic solvent-tolerant *Bacillus cereus* WQ9–2: Purification, biochemical properties, and potential application in peptide synthesis. *Bioresour. Technol.* **2010**, *101*, 7991–7994.
33. Abrahmsén, L.; Tom, J.; Burnier, J.; Butcher, K.A.; Kossiakoff, A.; Wells, J.A. Engineering subtilisin and its substrates for efficient ligation of peptide bonds in aqueous solution. *Biochemistry* **1991**, *30*, 4151–4159.
34. Adamczak, M.; Krishna, S.H. Enzyme for efficient biocatalysis. *Food Technol. Biotechnol.* **2004**, *42*, 251–264.
35. Ebeling, W.; Hennrich, N.; Klockow, M.; Metz, H.; Orth, H.D.; Lang, H. Proteinase K from *Tritirachium album* Limber. *Eur. J. Biochem.* **1974**, *47*, 91–97.
36. Carpi, F.M.; di Pietro, F.; Vincenzetti, S.; Mignini, F.; Napolioni, V. Human DNA extraction methods: Patents and applications. *Recent Pat. DNA Gene Seq.* **2011**, *5*, 1–7.
37. Sigma-Aldrich Corporation. Enzymes for cell dissociation and lysis. *Biofiles Life Sci. Res.* **2006**, Available online: http://www.sigmaaldrich.com/etc/medialib/docs/Sigma/General_Information/2/biofiles_issue2.Par.0001.File.tmp/biofiles_issue2.pdf (accessed on 20 July 2013).
38. Santangelo, C. Worthington biochemical online tissue dissociation guide. *Worthington Biochem. Corp.* **2011**. Available online: <http://www.worthington-biochem.com/tissuedissociation/default.html> (accessed on 20 July 2013).
39. Mandl, I.; Maclennan, J.D.; Howes, E.L. Isolation and characterization of proteinase and collagenase from *Cl. histolyticum*. *J. Clin. Investig.* **1953**, *32*, 1323–1329.
40. Kin, T.; Johnson, P.R.; Shapiro, A.M.; Lakey, J.R. Factors influencing the collagenase digestion phase of human islet isolation. *Transplantation* **2007**, *83*, 7–12.

41. Canavan, H.E.; Cheng, X.; Graham, D.J.; Ratner, B.D.; Castner, D.G. Cell sheet detachment affects the extracellular matrix: A surface science study comparing thermal liftoff, enzymatic, and mechanical methods. *J. Biomed. Mater. Res. A* **2005**, *75*, 1–13.
42. Huang, H.L.; Hsing, H.W.; Lai, T.C.; Chen, Y.W.; Lee, T.R.; Chan, H.T.; Lyu, P.C.; Wu, C.L.; Lu, Y.C.; Lin, S.T.; *et al.* Trypsin-induced proteome alteration during cell subculture in mammalian cells. *J. Biomed. Sci.* **2012**, *17*, doi:10.1186/1423-0127-17-36.
43. Danhier, P.; Copetti, T.; de Preter, G.; Leveque, P.; Feron, O.; Jordan, B.F.; Sonveaux, P.; Gallez, B. Influence of cell detachment on the respiration rate of tumor and endothelial cells. *PLoS One* **2013**, *8*, e53324.
44. Rader, C. Overview on concepts and applications of Fab antibody fragments. *Curr. Protoc. Protein Sci.* **2009**, *55*, 6.9.1–6.9.14.
45. Flanagan, R.J.; Jones, A.L. Fab antibody fragments: Some applications in clinical toxicology. *Drug Saf.* **2004**, *27*, 1115–1133.
46. De Marco, A. Biotechnological applications of recombinant single-domain antibody fragments. *Microb. Cell Fact.* **2011**, *10*, doi:10.1186/1475-2859-10-44.
47. Zhao, Y.; Gutshall, L.; Jiang, H.; Baker, A.; Beil, E.; Obmolova, G.; Carton, J.; Taudte, S.; Amegadzie, B. Two routes for production and purification of Fab fragments in biopharmaceutical discovery research: Papain digestion of mAb and transient expression in mammalian cells. *Protein Expr. Purif.* **2009**, *67*, 182–189.
48. Hunte, C.; Michel, H. Crystallisation of membrane proteins mediated by antibody fragments. *Curr. Opin. Struct. Biol.* **2002**, *12*, 503–508.
49. Kovari, L.C.; Momany, C.; Rossmann, M.G. The use of antibody fragments for crystallization and structure determinations. *Structure* **1995**, *3*, 1291–1293.
50. Caffrey, M. Membrane protein crystallization. *J. Struct. Biol.* **2003**, *42*, 108–132.
51. Bill, R.M.; Henderson, P.J.; Iwata, S.; Kunji, E.R.; Michel, H.; Neutze, R.; Newstead, S.; Poolman, B.; Tate, C.G.; Vogel, H. Overcoming barriers to membrane protein structure determination. *Nat. Biotechnol.* **2011**, *29*, 335–340.
52. Zhou, Y.; Morais-Cabral, J.H.; Kaufman, A.; MacKinnon, R. Chemistry of ion coordination and hydration revealed by a K⁺ channel-Fab complex at 2.0 Å resolution. *Nature* **2001**, *414*, 43–48.
53. Rasmussen, S.G.F.; Choi, H.J.; Rosenbaum, D.M.; Kobilka, T.S.; Thian, F.S.; Edwards, P.C.; Burghammer, M.; Ratnala, V.R.; Sanishvili, R.; Fischetti, R.F.; *et al.* Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature* **2007**, *450*, 383–387.
54. Kwong, P.D.; Wyatt, R.; Robinson, J.; Sweet, R.W.; Sodroski, J.; Hendrickson, W.A. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **1998**, *393*, 648–659.
55. Bagossi, P.; Sperka, T.; Fehér, A.; Kádas, J.; Zahuczky, G.; Miklóssy, G.; Boross, P.; Tózsér, J. Amino acid preferences for a critical substrate binding subsite of retroviral proteases in type 1 cleavage sites. *J. Virol.* **2005**, *79*, 4213–4218.
56. Eizert, H.; Bander, P.; Bagossi, P.; Sperka, T.; Miklóssy, G.; Boross, P.; Weber, I.T.; Tózsér, J. Amino acid preferences of retroviral proteases for amino-terminal positions in a type 1 cleavage site. *J. Virol.* **2008**, *82*, 10111–10117.

57. Tózsér, J. Comparative studies on retroviral proteases: Substrate specificity. *Viruses* **2010**, *2*, 147–165.
58. Patick, A.K.; Potts, K.E. Protease inhibitors as antiviral agents. *Clin. Microbiol. Rev.* **1998**, *11*, 614–627.
59. Wlodawer, A.; Gustchina, A. Structural and biochemical studies of retroviral proteases. *Biochim. Biophys. Acta* **2000**, *1477*, 16–34.
60. De Clercq, E. Strategies in the design of antiviral drugs. *Nat. Rev. Drug Discov.* **2000**, *1*, 13–25.
61. Wlodawer, A. Structure-based design of AIDS drugs and the development of resistance. *Vox Sang.* **2002**, *83* (Suppl. 1), 23–26.
62. Pogson, M.; Georgiou, G.; Iverson, B.L. Engineering next generation proteases. *Curr. Opin. Biotechnol.* **2009**, *20*, 390–397.
63. Lim, E.J.; Sampath, S.; Coll-Rodriguez, J.; Schmidt, J.; Ray, K.; Rodgers, D.W. Swapping the substrate specificities of the neuropeptidases neurolysin and thimet oligopeptidase. *J. Biol. Chem.* **2007**, *282*, 9722–9732.
64. Villa, J.P.; Bertenshaw, G.P.; Bond, J.S. Critical amino acids in the active site of meprin metalloproteinases for substrate and peptide bond specificity. *J. Biol. Chem.* **2003**, *278*, 42545–42550.
65. Derewenda, Z.S. The use of recombinant methods and molecular engineering in protein crystallization. *Methods* **2004**, *34*, 354–363.
66. Waugh, D.S. Making the most of affinity tags. *Trends Biotechnol.* **2005**, *23*, 316–320.
67. Renzi, F.; Panetta, G.; Vallone, B.; Brunori, M.; Arceci, M.; Bozzoni, I.; Laneve, P.; Caffarelli, E. Large-scale purification and crystallization of the endoribonuclease XendoU: Troubleshooting with His-tagged proteins. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2006**, *62*, 298–301.
68. Horchani, H.; Ouertani, S.; Gargouri, Y.; Sayari, A. The N-terminal His-tag and the recombination process affect the biochemical properties of *Staphylococcus aureus* lipase produced in *Escherichia coli*. *J. Mol. Catal. B Enzym.* **2009**, *61*, 194–201.
69. Terpe, K. Overview of tag protein fusions: From molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* **2003**, *60*, 523–533.
70. Jenny, R.J.; Mann, K.G.; Lundblad, R.L. A critical review of the methods for cleavage of fusion proteins with thrombin and factor Xa. *Protein Expr. Purif.* **2003**, *31*, 1–11.
71. Arnau, J.; Lauritzen, C.; Petersen, G.E.; Pedersen, J. Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expr. Purif.* **2005**, *48*, 1–13.
72. Vergis, J.M.; Wiener, M.C. The variable detergent sensitivity of proteases that are utilized for recombinant protein affinity tag removal. *Protein Expr. Purif.* **2011**, *78*, 139–142.
73. Waugh, D.S. An overview of enzymatic reagents for the removal of affinity tags. *Protein Expr. Purif.* **2011**, *80*, 283–293.
74. Nallamsetty, S.; Kapust, R.B.; Tózsér, J.; Cherry, S.; Tropea, J.E.; Copeland, T.D.; Waugh, D.S. Efficient site-specific processing of fusion proteins by tobacco vein mottling virus protease *in vivo* and *in vitro*. *Protein Expr. Purif.* **2004**, *38*, 108–115.
75. Huang, Q.; Li, Q.; Chen, A.S.; Kang, C. West Nile virus protease activity in detergent solutions and application for affinity tag removal. *Anal. Biochem.* **2013**, *435*, 44–46.

76. Zhang, D.; Tózsér, J.; Waugh, D.S. Molecular cloning, overproduction, purification and biochemical characterization of the p39 nsp2 protease domains encoded by three alphaviruses. *Protein Expr. Purif.* **2009**, *64*, 89–97.
77. Austin, B.P.; Tózsér, J.; Bagossi, P.; Tropea, J.E.; Waugh, D.S. The substrate specificity of *Metarhizium anisopliae* and *Bos taurus* carboxypeptidases A: Insights into their use as tools for the removal of affinity tags. *Protein Expr. Purif.* **2011**, *77*, 53–61.
78. Steen, H.; Mann, M. The ABC's (and XYZ's) of peptide sequencing. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 699–711.
79. Granvogl, B.; Plösch, M.; Eichacker, L.A. Sample preparation by in-gel digestion for mass spectrometry-based proteomics. *Anal. Bioanal. Chem.* **2007**, *389*, 991–1002.
80. Meyers, A.; Trauger, S.; Webb, W.; Reisdorph, N.; Wranik, C.; Peters, E.; Siuzdak, G. Protein identification and profiling with mass spectrometry. *Spectroscopy* **2003**, *17*, 1–15.

© 2013 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).