catalysts

MDPI

Review

Reversed Proteolysis—Proteases as Peptide Ligases

Peter Goettig

Structural Biology Group, Department of Biosciences, University of Salzburg, 5020 Salzburg, Austria; peter.goettig@sbg.ac.at; Tel.: +43-662-8044-7283

Abstract: Historically, ligase activity by proteases was theoretically derived due to their catalyst nature, and it was experimentally observed as early as around 1900. Initially, the digestive proteases, such as pepsin, chymotrypsin, and trypsin were employed to perform in vitro syntheses of small peptides. Protease-catalyzed ligation is more efficient than peptide bond hydrolysis in organic solvents, representing control of the thermodynamic equilibrium. Peptide esters readily form acyl intermediates with serine and cysteine proteases, followed by peptide bond synthesis at the N-terminus of another residue. This type of reaction is under kinetic control, favoring aminolysis over hydrolysis. Although only a few natural peptide ligases are known, such as ubiquitin ligases, sortases, and legumains, the principle of proteases as general catalysts could be adapted to engineer some proteases accordingly. In particular, the serine proteases subtilisin and trypsin were converted to efficient ligases, which are known as subtiligase and trypsiligase. Together with sortases and legumains, they turned out to be very useful in linking peptides and proteins with a great variety of molecules, including biomarkers, sugars or building blocks with non-natural amino acids. Thus, these engineered enzymes are a promising branch for academic research and for pharmaceutical progress.

Keywords: protease catalysis; peptide bond synthesis; natural ligases

Citation: Goettig, P. Reversed Proteolysis—Proteases as Peptide Ligases. *Catalysts* **2021**, *11*, 33. https://doi.org/10.3390/catal11010033

Received: 9 November 2020
Accepted: 27 December 2020
Published: 30 December 2020

1. Introduction

A large number of data on proteases cleaving all sorts of peptide bonds is available in the scientific publications and textbooks. About 100 years ago, the prototypic mammalian food digesting enzymes were already investigated, such as pepsin, an aspartic protease, as well as the serine proteases trypsin, chymotrypsin, and elastases, followed by carboxypeptidases, which belong to the metalloproteinases. Later on, it turned out that blood coagulation including fibrinolysis and the innate immune response of the complement system depend on trypsin-like serine proteases. In the following decades, more proteases were discovered in all living organisms and of new types, e.g., the eponymous cysteine protease papain from the plant *Carica papaya*. Although their major task is always the cleavage of one or more peptide bonds, their physiological roles range from the unspecific digestion of proteins in food to different protein substrates with varying specificity. Many diverse substrates are found in different tissues under healthy conditions or even in cancer cells, whereas some proteases exhibit exclusive specificity for a single cut in only one protein, such as in several viral proteases. All these proteases or peptidases are catalysts, which reduce the required activation energy of a given reaction, following in case of proteolysis always the thermodynamically favored direction of cleavage. However, as pure catalysts, they all are capable of catalyzing the reverse or backward reaction. This concept was already suggested in the year 1898 by van 't Hoff for trypsin as a potential catalyst in protein synthesis, linking its own cleavage products again, while the experimental phenomenon was reported by Sawjalow in 1901 as plastein formation [1,2]. Later studies confirmed that pepsin, papain, or chymotrypsin synthesized from protein hydrolysates new polypeptides with an average of 40 amino acids, the insoluble plastein [3].

Since all living cells possess a highly efficient protein synthesis machinery with the ribosomes, there is usually little need for enzymes as catalysts in peptide ligation. However, in order to incorporate *D*- or β-amino acids, methyl-, glyco- and phosphoryl-residues, as well as heterocycles and fatty acids into peptides, some bacteria, fungi and lower eukaryotes utilize peptide ligases, which are organized in multimodular "megaenzymes" as non-ribosomal peptide synthetases (NRPS) [4]. Typically, they contain an ATP-driven activation domain, as in the NRPS of *Bacillus subtilis*, which synthesizes the cyclic antibiotic bacitracin with four *D*- and five *L*-residues. Another remarkable example of an enzyme complex, which synthesizes the tripeptide δ-(*L*-α-aminoadipyl)-*L*-cysteinyl-*D*-valine (ACV), was discovered in *Penicillium chrysogenum*, comprising subunits for activation, condensation, thiolation, epimerization of *L*-valine, and thioesterase function [5].

It should be noted that the common ubiquitinylation and sumoylation of eukaryotes are post-translational modifications, which involve protein ligation; however, both processes result in isopeptide bond formation of εNH_2-groups of Lys side-chains with the C-terminal carboxylates of ubiquitin or small ubiquitin-like modifier (SUMO) molecules [6]. In addition, non-Lys ubiquitinylation on the Ser, Thr, and Cys side chains of target proteins was described [7]. Otherwise, only a few natural examples are known for true peptide ligases and some proteases, which work both ways, cleaving and/or synthesizing peptide bonds, although protease-catalyzed protein splicing was suggested as a new post-translational modification [8]. Such enzyme-dependent processing is distinct from intein splicing. Self-splicing natural proteins excise inteins from two exteins with neighboring group participation (anchimeric assistance). Inteins are present in all kingdoms of life, requiring the consensus sequence Nterm-Extein1-[Cys/Ser-His-Asn]-Cys/Ser/Thr-Extein-2-COOH, with His assisting the nucleophilic attacks at carbonyl C atoms [9,10]. Several hundred natural inteins excise themselves from exteins in this manner, but they are rare among eukaryotes [11]. By contrast, immunoproteasomes, in which the active subunits β1, β2, and β5 are replaced by β1i, β2i, and β5i subunits, are capable of religating peptide fragments [12,13]. They cut out up to 40 residues from a protein and ligate the fragments via acyl intermediates bound to the catalytic threonine residues, followed by transpeptidation, resulting in antigen presentation at major histocompatibility complex I (MHC I) class molecules (Figure 1A) [14]. The ligation can proceed in common or so-called reverse splicing as two remaining N-termini can react with bound acyl intermediates, which generates overall 25% of all MHC I presented antigens [15]. A recent study employed synthetic peptides derived from HIV antigens or human "self" peptides, which resulted in 27% spliced peptides, e.g., SSCLPCLPS***F***K could be generated from SSCLPCLPSEK***F***, whereas other combinations are possible [16]. An overview of proteasome-mediated peptide splicing in the immunological context was written by Vigneron and coworkers [17].

Otherwise, protein splicing has become a useful tool in biotechnological applications for the labeling, purification, and biosensoring of proteins [18]. Together with several engineered proteases with enhanced ligation capacity, they are an alternative to a variety of chemical ligation methods, employed to link chemically synthesized building blocks to larger proteins [19]. Native chemical ligation, which can be performed in solution or with a solid phase support is a powerful method for in vitro polypeptide synthesis, in particular when chemically modified or non-natural molecular building blocks are incorporated [20,21]. However, chemical ligation is limited in the size of proteins to roughly 200 residues, while the average size of polypeptides in corresponding studies is around 80 residues [22]. Nevertheless, combined approaches of chemical synthesis with ligation by proteases have been successfully performed [23,24].

Sortase A from *Staphylococcus aureus* is a transpeptidase with an active site Cys, requiring a LPXTGXX motif with GXX as a leaving group, in order to form a thioester intermediate and conjugate a LPXT-CO acyl moiety with the N-terminus of another peptide (Figure 1A). The only other prominent type of natural peptide ligases are several plant legumains, which are also known as asparaginyl endopeptidase (AEP) or occasionally termed butelase.

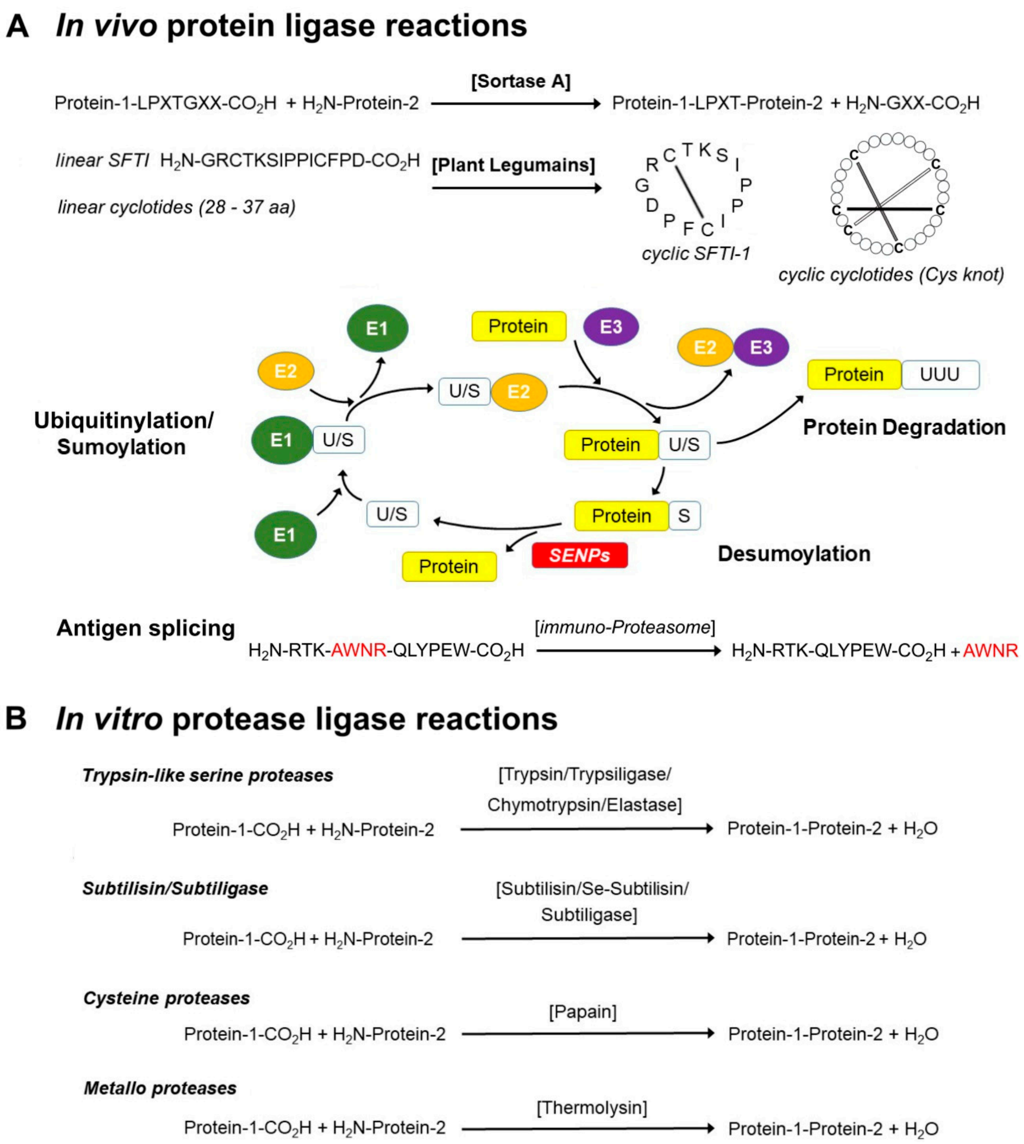

Figure 1. Overview of Proteases/Ligases in Biology and Organic Synthesis. (**A**) Natural processes involving protein ligases. Cyclotides are 28 to 37 amino acids (aa) long. Ubiquitinylation and sumoylation require the specialized ligases E1, E2, and E3, being various enzymes in both pathways. E1 utilizes ATP hydrolysis for the formation of thioesters with the C-terminus of ubiquitin (U) or SUMO (S) and transfers them to E2 as thioesters. Eventually, up to 1000 specialized E3 ligases, such as the numerous really interesting new gene (RING) finger domain ligases, form isopeptide bonds between U/S and Lys side-chains of target proteins. Whereas polyubquitinylated proteins are degraded by proteasomes, several sentrin-specific proteases (SENPs) cleave the SUMO tags in regulatory processes. (**B**) Synthetic approaches using various proteases and their engineered variants.

Legumains are lysosomal caspase-related cysteine proteases, which are widespread among eukaryotes, with a marked specificity for Asn in a wider range around neutral pH or Asp at acidic pH as P1-residue of the scissile bond [25]. In particular, plant legumains have gained additional interest in recent years, as some members are distinct ligases, which are important in the synthesis of cyclic plant peptides [26]. Another special field in which peptide ligation is thought to play a role is the interaction of a variety of proteases with inhibitors that are first cleaved and then religated in the active site. Eventually, common proteases and optimized, and engineered variants such as trypsiligase and subtiligase can be used as peptide synthetases under nearly physiological conditions or in organic solvents, which increases the tendency for peptide ligation (Figure 1B) [19].

2. Historical Outline of Developments in the Last Century

The very first experiments demonstrating that proteases can be used for the synthesis of peptide bonds were published by Bergmann and Fraenkel-Conrat [27,28]. A comprehensive overview of the field until the 1980s is given by Fruton [29]. To the early pioneers, it was obvious that thermodynamics disfavored peptide bond formation of charged N- and C-termini based on the zwitterionic nature of the amino acids themselves. However, as oligopeptides are often less soluble in aqueous buffers compared with the charged or more polar precursors, their removal from the solution, e.g., by precipitation, can shift the equilibrium toward the polypeptide product. Kinetic parameters of the peptidolytic hydrolysis reaction such as k_{cat} and K_M are useful indicators for the efficiency of the reverse reaction, the peptide ligation. In particular, K_M can be used to assess at least partially the efficiency of the synthesis, whereas calculation of the ligation k_{cat} is more complicated, due to the linking of two substrates. As catalysts, proteases accelerate the forward and backward reaction and the eventual reaching of the equilibrium. Since the chemical equilibrium constant for dipeptide bond formation of Xaa–Gly ranges from about 0.001 to 0.005 L/mol and the average free energy $\Delta G \approx 5.0$ kJ/mol, the hydrolysis reaction is favored in aqueous solution [3,30]. Bordusa gives a good overview of the two basic approaches in peptide synthesis catalyzed by proteases, namely thermodynamic or equilibrium control and kinetic control; see Figure 2 [31]. In addition, N-protected amino acids and peptides, esters of the C-terminal moiety and various other precursors of amino acids were used as substrates of protease–peptide ligases. As a general rule of thumb, it turned out that charged residues are better ligated in organic solvents and esters and N-terminally protected residues are suitable for aqueous solutions, when trypsin is used as catalyst. The esters of the carboxylate moiety form proper acyl intermediates as in the hydrolysis reaction. Similar observations were made for chymotrypsin in systematic analyses, which confirmed that an acetylated N-terminus is favorable for the ester component, such as Ac–NH_2–Phe–CO–OEt. However, it was noticed that serine proteases, such as bovine ß-trypsin, α-chymotrypsin, and porcine plasmin not only cleave the reactive bond Lys15-Ala16 of the bovine pancreatic trypsin inhibitor (BPTI) but also resynthesize it until the thermodynamic equilibrium is reached [32].

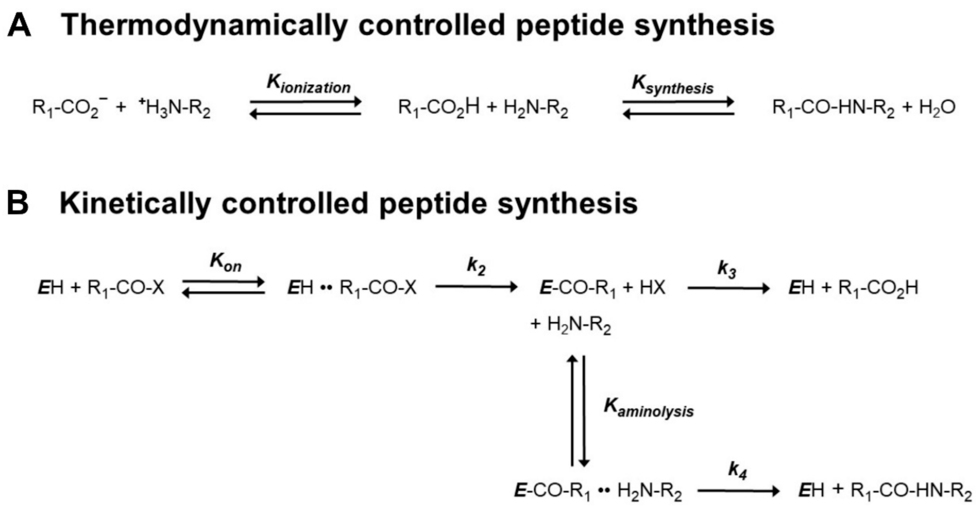

Figure 2. Thermodynamic and kinetic control of peptide bond synthesis with protease catalysts. (**A**) Thermodynamic or equilibrium control dominates in catalysis by natural ligases or in organic media with low water content that shifts the chemical equilibrium. (**B**) Kinetic control is mostly achieved by synthetic precursors of peptides or proteins, containing a derivative of the C-terminal carboxylate with a good leaving group, such as carboxy esters of the residue with methanol (MeOH) and various other alcohols, as well as corresponding thioesters and thiols as leaving moieties. The formation of an acyl intermediate with serine or cysteine proteases is required [33].

Among serine proteases, the bacterial subtilisin BPN′ was quite successful in various ligation reactions. Eventually, an engineered version, the so-called subtiligase, was widely used as peptide ligating enzyme. Cysteine proteases such as carboxypeptidase Y from yeast or papain from the papaya plant showed similar results according to their specificity. Essentially, metalloproteinases, e.g., thermolysin (see below) and aspartic proteases, such as pepsin, rennin, or cathepsin D, synthesized peptide bonds in a similar manner as the above-mentioned cysteine and serine proteases, according to their own specificity, but with different kinetics, which depend on their underlying basic reaction mechanisms [34]. Then, in the late 1970s, progress was made, with the first spectacular example of porcine derived desalanine-B30-insulin, which was generated by cleavage with Zn^{2+}-dependent carboxypeptidase A and coupled to Thr–OBut by trypsin catalysis to generate human insulin, whereby Thr–OBut was present in 50-fold excess in a dimethylformamide (DMF)/ethanol mixture [35]. Subsequently, this trypsin-catalyzed transpeptidation was commercially exploited to produce humanized insulin from the porcine variant by Novo A/S [36]. A similar approach succeeded in the papain and chymotrypsin catalyzed synthesis of opioid pentapeptides with N–tBOC amino acids as precursors or corresponding esters and phenylhydrazides as nucleophiles [37]. Another outstanding example is the industrial production of the food sweetener aspartame by ligation of amino acids with thermolysin, shifting the equilibrium by high molar excess of reactants Z–Asp–COOH and H_2N–Phe–OMe and precipitation of the insoluble product [38].

The basic requirements and limitations of protease catalyzed ligation were well known in the 1990s, as aqueous solutions in the pH range of 5 to 9 favor aminolysis, which was mostly due to thermodynamic barrier set by the charged terminal carboxylate and ammonium groups, resulting in $\Delta G \approx 2.1$ kJ/mol derived from $K_{synthesis}$ of a dipeptide with one N/C-protecting group on the amino acids [39]. The first calorimetric analyses in the 1950s had already found $\Delta G \approx 1.8$ kJ/mol for a corresponding chymotrypsin catalyzed peptide bond formation from Bz–Tyr–COOH and H_2N–Gly–$CONH_2$ [40]. However, the thermodynamic equilibrium can be largely shifted to peptide synthesis by using organic solvents, which reduce the ionization, as was demonstrated for a similar reaction to Z-Trp-Gly-$CONH_2$ in mixtures of water and 85% butanediol or other organic solvents, such as dimethyl sulfoxide (DMSO) [41]. Moreover, some organic solvents reduce the activity of proteolytic enzymes to a great extent, while others in particular polyalcohols such as glycerol stabilize them. Thermodynamic control of the equilibrium in aqueous solution may utilize the excess of one reactant or removal of the product, as demonstrated by the precipitation of the aspartame salt of Z–L–Asp–L–Phe–COOMe with D–Phe–COOMe, which forms by catalysis with thermolysin [42]. The special case of intramolecular peptide bond synthesis in the soybean trypsin inhibitor (SBTI) was already described in 1969, showing tryptic cleavage of the Arg–Ile64 bond, the subsequent cleavage of Arg63, and the ligation of a new Lys63 by carboxypeptidase B catalysis and peptide bond synthesis with trypsin to a SBTI containing Lys–Arg64 [43]. In general, biphasic aqueous–organic media are a way to circumvent the problem of inactivation of enzymes; e.g., thermolysin or pepsin are present in the water phase, whereby the ligation products accumulate in the organic phase [44].

Another route to peptide synthesis follows kinetic control in aqueous solution, involving peptidic carboxylate ester precursors with leaving groups X, resulting in acyl intermediates with serine or cysteine proteases, followed by aminolysis at higher basic pH, which ensures an excess of nucleophilic α-NH_2 groups [39,45]. These reactions of the type R1–COX + H_2N-R2 $\rightarrow$ R1–CO–NH–R2 + X are not so much influenced by the thermodynamic equilibrium and its $K_{synthesis}$ constant but rather by the velocity of the competing aminolytic and hydrolytic reaction (i.e., the ratio of vA/vH), which are both catalyzed by the protease. As the reaction starts with a burst, the maximum yield can be obtained by inactivation of the protease at acidic pH, while an excess of acyl donor precursor seems favorable [46]. Nevertheless, examples of mixed thermodynamic and kinetic control were reported, such as the synthesis of the tripeptide Z–Ala–Phe–Leu–NH_2

by chymotrypsin, which was present in aqueous reverse micelles as activity enhancing microreactors [47].

Ideally, the protease specificity ensures little side products without the need for protecting groups and, in particular, the proper stereochemistry of natural proteins, consisting exclusively of *L*-amino acids, in contrast to organic synthesis with the ever present risk of racemization and increasingly lower yields with each additional peptide residue [48]. Basically, any natural protease has the potential to synthesize peptide bonds according to its specificity for the amino acid residues around the scissile or to be ligated bond. An overview of the most frequently used proteases as ligases is displayed with respective catalytic and specificity determining components in Figure 3. However, some proteases turned out to be more suitable under kinetic control, in particular serine and cysteine proteases that form acyl intermediates. Interestingly, the S1′ specificity plays a crucial role in enhancing the nucleophilicity of carboxyamide precursors such as H_2N–Arg–CO–NH_2 with acyl intermediates formed by Mal–Phe–OMe esters with α-chymotrypsin, while the known S3′ specificity for Arg side chains might further exploited [49,50]. Experiments in the late 1980s and early 1990s focused on the synthesis of physiologically relevant peptides and small hormones, e.g., the cysteine protease papain and α-chymotrypsin and were used to generate the δ-sleep inducing peptide from three tripeptides [48,51]. Further examples are the ligation of two segments to gonadotropin-releasing hormone by chymotrypsin and human growth-hormone releasing factor (GRF) by trypsin, respectively [52,53].

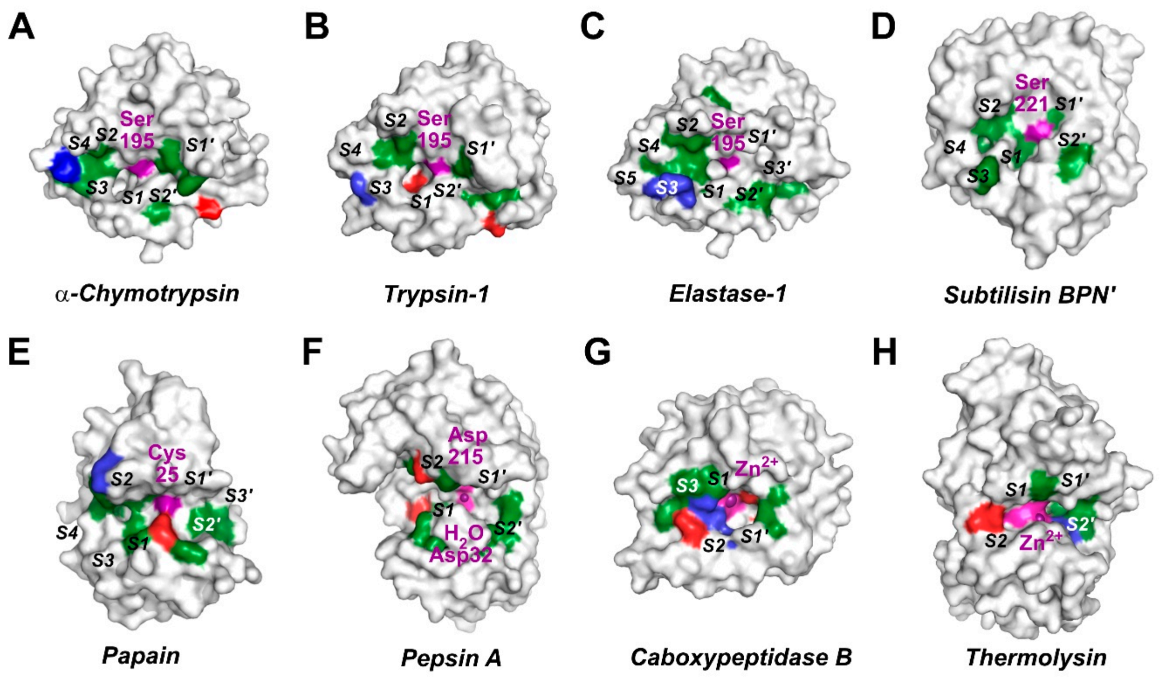

Figure 3. Proteases as ligases. Molecular surfaces are depicted white in overall polar and hydrophilic regions. Positively (blue) and negatively (red) charged residues, as well as hydrophobic patches (green) contributing to the specificity are colored. Crucial catalytic residues and ions (purple) are labeled, including catalytic water molecules (magenta), as well as the most relevant specificity pockets. All proteases are shown in the standard orientation, with the N-terminus of the substrate to the left and the C-terminus on the right. (**A**) Bovine α-chymotrypsin. (**B**) Bovine β-trypsin. (**C**) Porcine pancreatic elastase 1. (**D**) Bacterial subtilisin BPN′; from *Bacillus ameloliquefaciens*. (**E**) The plant cysteine protease papain. (**F**) The aspartic protease porcine pepsin. (**G**) Porcine carboxypeptidase B. (**H**) Thermolysin from the bacterium *Bacillus thermoproteolyticus*.

In addition, enzymatic modifications of polypeptides and proteins, which were obtained from recombinant expression, have been performed. Thus, it was possible to convert GRF(1–29)–COOH into the more active GRF(1–29)–CO–NH_2 with NH_4OAc/NH_3 in a 90% 1,4-butanediol solution by trypsin [54]. In addition, the zinc metalloprotease thermolysin from *Bacillus thermoproteolyticus* and the mammalian serine protease elastase were successful catalysts in coupling amino acid amides, such as Ty–CO–NH_2 to human neuropeptide Y (1–35) [55]. Similar modifying syntheses were performed with the serine proteases carboxypeptidase Y from yeast, V8/endoproteinase Glu-C (*Staphylococcus aureus*), and the Glu/Asp-specific endoprotease (GSE, *Bacillus licheniformes*) [56–58]. The previously mentioned enzymes were utilized in the more demanding semisyntheses of larger proteins. An outstanding example is the semisynthesis of the Hemoglobin A apo form, which was linked by V8 protease catalysis at positions 30–31 in 30% iso-propanol, resulting in the 141 residue long full length chain and subsequent reconstitution with a heme group [59]. Tryptic cleavage at Lys15–Ala16 and removal of the dipeptide Ala16-Arg17 by aminopeptidase K allowed incorporating the dipeptides Gly–Arg, Ala–Lys, and Leu–Arg in functional BPTI variants [60].

Further steps toward true ligases were the chemical conversion of the serine protease subtilisin from *Bacillus subtilis* thiolsubtilisin and selenolsubtilisin, in which the catalytic Ser221 was replaced by cysteine an selenocysteine, respectively (Figure 4A) [61,62]. As the Cys221 variant displayed a strong enhancement of the ligase character in 1:1 mixtures of water and dimethylformamide (DMF), the Sec221 variant appeared to be an even better acyl transferase, increasing aminolysis with respect to Ser and Cys subtilisin by factors of 1000 and 10, respectively [63]. However, the catalytic Sec221 is prone to oxidation and forms the seleninic acid R–SeO_2H, as observed in the crystal structure, which most likely impedes applications of selenosubtilisin as ligase [64]. Recombinant engineering of subtilisin led to the thiolsubtilisin variant Pro225Ala, with a 450-fold increase of the ratio aminolysis to hydrolysis, while additional mutations increased the stability in DMF and aqueous solution 50- to 100-fold [65,66]. Further studies of modified subtilisin variants resulted in the so-called subtiligases, which are discussed in the following Section 3.1.

In particular, the group around Jakubke demonstrated that low temperatures in aqueous solution, with an optimum below −10 °C in the frozen state, favored the formation of peptide bonds by α-chymotrypsin and β-trypsin in kinetically controlled reactions [67,68]. Similarly, porcine elastase synthesized peptide bonds according to its P1-specificity for Ser, Ile, and Val, while a bacterial Glu/Asp-specific endopeptidase synthesized Glu/Asp-Xaa bonds around −15 °C [69,70]. In addition, the group found significant ligase activity by the zymogens trypsinogen and chymotrypsinogen, e.g., with increased binding of nucleophiles in kinetically controlled reactions [71].

3. Developments in the 21st Century

Some routes employing trypsin and chymotrypsin as ligases were further followed, as in kinetically controlled reactions in the frozen state, resulting in the suppression of hydrolysis and increase of yields [68]. Established protease-ligases as subtilisin and V8 protease from *Staphylococcus aureus*, which is trypsin-like with P1–Glu specificity, were used to demonstrate that macromolecular crowding agents, such as polyethylene glycols (PEGs) and dextran enhanced the ligase reaction for the synthesis of triose-phosphate isomerase [72]. Similarly, subtilisin Carlsberg catalyzed the condensation of a 15-residue glycopeptide from a tripeptide as acyl acceptor and a 12-mer as acyl donor in a mixture of water and DMF (1:9) [73].

3.1. Subtiligase

Bacterial subtilisins were already structurally characterized in the early 1970s as subtilisin BPN′ from *Bacillus amyloliquefaciens*, which exhibits the α/β-hydrolase fold of serine proteases [74]. A good example demonstrating that subtilisin BPN′ can ligate peptide bonds is the resynthesis of the scissile bond in chymotrypsin inhibitor 2 (Figure 4B) [75]. The significantly reduced hydrolytic activity of subtilisin BPN′; in organic solvent mixtures, such as 50% DMF, was explained by corresponding crystal structures, with His64 rotated by 180° around the Cβ-Cγ bond, thereby disrupting the catalytic triad with Asp32 and Ser221 [76]. Engineered thermostable variants became increasingly important due to their capacity for protein degradation in laundry detergents [77]. As already mentioned, subtilisin variants gained interest as protein ligases from the mid-1990s onwards. Thus, the subtilisin BPN′; double mutant Ser221Cys/Pro225Ala allowed synthesizing full-length Ribonuclease A (124 residues) from six esterified peptides in aqueous solution, with a ligation efficiency of about 70% in each step and final milligram yields [78]. Early applications were the ligation of various proteins preferentially at their N-terminus with biotin or PEG-linked peptides, i.e., their esterified precursors [24]. Among the precursor peptides, glycolate phenylalanyl esters and benzyl thioesters could be linked with free amino termini of peptides by subtiligases [79]. Phage display resulted in additional mutations, such as Met124Leu/Ser125Ala, which increased the activity nearly 3-fold (Figure 4C) [80].

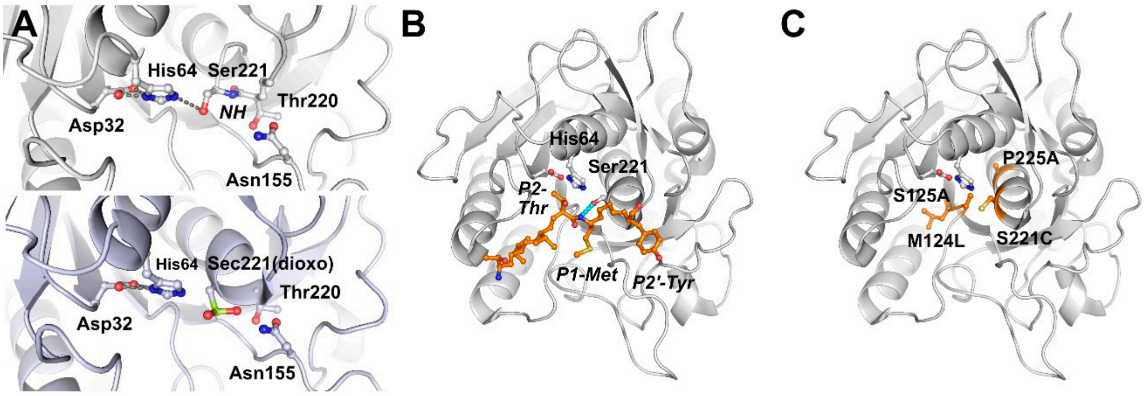

Figure 4. Subtilisin and derived ligases. (**A**) Active sites of subtilisin BPN′ and the chemically modified Ser221 variant selenolsubtilisin Carlsberg (Protein Data Bank (PDB) codes 1S02 and 1SEL). The oxyanion hole is formed by the Ser221 amide (NH) and the side chains of Asn155 and Thr220 (upper panel). Selenocysteine (Sec) enhances the acyl transferase activity and, consequently, ligase activity, although it is oxidized in crystals (lower panel). (**B**) The complex of subtilisin BPN′ and chymotrypsin inhibitor 2 (CI2), with relevant residues shown as ball-and-sticks (PDB 1LW6). CI2 binds in a substrate-like manner to the specificity pockets S4 to S2′. For clarity, most residues of CI2 and the P1′ side chain (Glu) were omitted. The scissile and religated bond near Ser221 is depicted with the carbonyl O (red) and amide N (blue), while the nucleophile–electrophile interaction between the Ser-Oγ and the carbonyl C is shown as cyan dots. (**C**) The two most critical mutations that turn subtilisin into subtiligase are the catalytic nucleophile Ser221Cys and Pro225Ala. In addition, Met124Leu and Ser125Ala enhance the ligase activity (model based on PDB 1SCJ).

It was pointed out that subtiligase favors thioester binding associated thiolysis with a broad substrate specificity over hydrolysis, in particular at low pH around 4.5 [81]. Subtiligase facilitated N-terminal tagging for proteomics either combined with click chemistry-based derivatization for nanomolar concentration quantification of products or with high affinity binding of biotin/avidin for the analysis of human blood proteins [82,83]. Moreover, subtiligase catalyzed labeling contributed to the investigation of the so-called α-aminome, whereby a database of proteolysis in healthy and apoptotic cells was established [84]. Mutants of subtiligase improved the efficiency in ligating N-terminally Cys-free peptides to protein thioesters, as well as a significant higher ligation rate for Glu at P1′ with the Tyr217Lys mutant (Figure 4C) [85]. Recently, an extensive engineering and selection process yielded more than 70 subtiligase mutants with differential specificity around the scissile/ligated bond, allowing for the N-terminal extension of most P1–P1′ combinations of amino acid residues [86,87]. Known as peptiligase or omniligase-1, a subtilisin variant ensured the cyclization of θ-defensins or retrocyclins, antimicrobial, and fungicidal peptides with 18 residues, which are expressed in old world monkeys, but not in the great apes or in humans [88,89]. Thymoligase is a further engineered and structurally determined peptiligase, which can synthesize the therapeutic polypeptide thymosin-α1 (28 aa long) with yields > 90% [90]. It deviates from subtiligase with several mutations such as an

Asn225, while Asn156 and Asp166 at S1 and Arg217 at S1′ shift the specificity toward basic P1 and acidic P1′ residues.

3.2. Sortases

Although several bacterial transpeptidases were known before, the characterization of sortase A from *Staphylococcus aureus* in 1999 was a hallmark for enzymes that can act as proteases and peptide ligases [91]. Sortase A cleaves polypeptides between threonine and the glycine of the Leu–Pro–Xaa–Thr–Gly (LPXTG) motif, i.e., the sorting signal, and transfers the acyl intermediates to the bacterial cell wall. Thus, the switch from protease to ligase activity is integrated in the overall transpeptidase function. The NMR structure revealed a β-barrel structure, apparently with a catalytic dyad, consisting of Cys184 and His120 and the activity stimulating a Ca^{2+}-binding site (Figure 5A) [92]. X-ray crystallography provided a corresponding ligand-free enzyme structure and a substrate complex (Figure 5B,C) [93]. It turned out that nearly all Gram-positive bacteria possess one or more sortases of different classes: while sortase A (SrtA) uses LPXTG, SrtB has a different sorting motif, e.g., NPQTN in *S. aureus*, as SrtC and SrtD appear to be more specialized, such as targeting LPNTA (Figure 5C) [94].

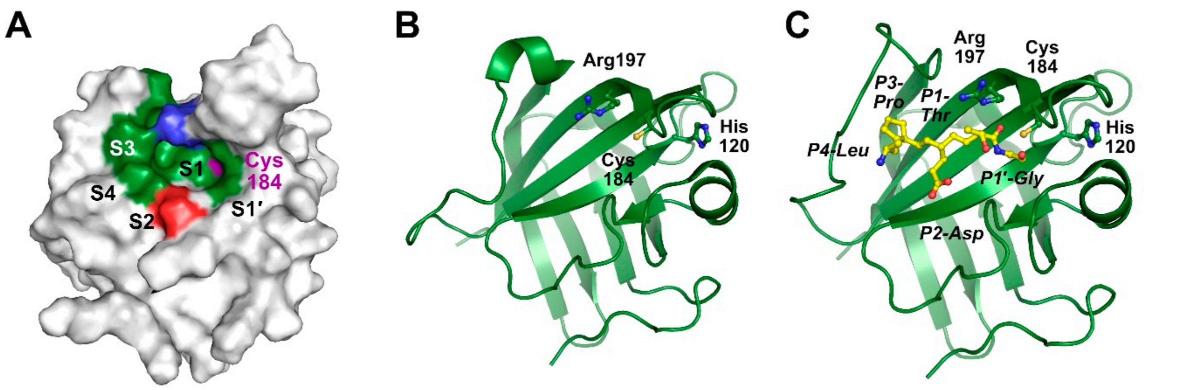

Figure 5. Sortase ligases. (**A**) Sortase A surface representation, derived from the first NMR structure (PDB 1IJA). The non-prime side of the active site is dominated by hydrophobic residues (green), less by basic (blue) and acidic (red) ones, whereby the catalytic Cys184 is shown in purple. (**B**) Ligand-free sortase A (PDB 1T2P). The most important catalytic residues are labeled: the nucleophile Cys184, its dyad partner His120, and Arg197. (**C**) Sortase A substrate complex with the peptide LPETG bound to the active site of a Cys184Ala mutant (PDB 1T2W). The role of Arg197 as an equivalent to an oxyanion hole is corroborated by binding the carbonyl O of P1–Thr.

Early studies demonstrated that SrtA could ligate polypeptides, e.g., enhanced green fluorescent protein (eGFP) was linked to dimers and higher oligomers, and it could attach PEGylated peptides to eGFP [95,96]. Moreover, peptide nucleic acids were ligated to other polypeptides and conjugate peptides to ethanol amine-linked glycosylphosphatidyl inositol, corresponding to natural GPI anchors [97,98]. Apart from the artificial sortase specific linker, various chimeric proteins were built from domains that kept their full functionality, such as GFP bound to immunoglobulin G (IgG) molecules [99]. It was demonstrated that GFP "sortagging" of various protein targets was possible in the cytosol of yeast and HEK293 cells [100]. The ligation reaction catalyzed by SrtA was exploited in various approaches, among them the generation of chimeric proteins that were linked to peptidic precursors for click reactions or of an immunotoxin consisting of a Fab fragment and gelonin, which was directed against αHer2 as tumor specific target, resulting in strong cell killing activity [101,102]. Using LPET or Gly-rich motifs, SrtA was capable of ligating cyclic SFTI-1 and cyclotide variants, with proper disulfide formation [103]. SrtA conjugated camelid nanobodies in high yields to labels for single-photon emission computed tomography (SPECT) with indium 111, positron emission tomography (PET) with gallium 68, and the fluorescent dye Cy5 for fluorescence reflectance imaging (FRI) [104].

Similarly, SrtA ligation and corresponding labeling facilitated NMR studies of the PSD-93 and-95 oligomers and their interactions with PDZ3/SH3–GK, which are important for the assembly of the mega-N-methyl-D-aspartate receptor synaptic signaling complex at glutamatergic synapses [105]. Furthermore, SrtA was part of an immobilization strategy for highly specific receptor proteins on sensorchips used in biolayer interferometers, as well as for immobilization on magnetic nanoparticles, employed in GFP label single molecule fluorescence spectroscopy [105,106]. Triple to hepta mutants of sortase showed enhanced performance for in vivo labeling and domain linking due to Ca^{2+} independent activity [107,108].

3.3. Legumains

Whereas the caspase-related legumains or asparaginyl endopeptidases (AEPs) of animals cleave various protein targets in lysosomes at acidic pH with a distinct specificity for P1-Asn and Asp residues, their plant counterparts diverged in proteases and enzymes that preferentially ligate peptide bonds (Figure 6A,B) [25,109]. The latter reaction is favored when cyclic polypeptides are formed, such as the prototypic sun flower trypsin inhibitor (SFTI-1), which comprises 14 residues and an internal disulfide bridge [110]. Similarly, several plant families produce cyclotides, comprising 28 to 37 residues and a so-called cysteine knot, which is formed by three internal disulfides [111]. Both cyclotides and cyclic SFTI molecules appear to be host defense molecules with antimicrobial and insecticidal properties, which gained interest as potential new pharmaceutical drugs [112]. The cyclization takes place after the cleavage of the C-terminal pro-peptide with a P1-Asn from the cyclotide precursor protein through a transpeptidation reaction, which is even catalyzed in alien foreign cyclotides by AEPs from *Nicotiana benthamiana* [113]. The transpeptidation reaction involves an acyl-transfer step from the acyl-AEP intermediate to the N-terminal residue of the cyclotide domain. In the context of plant legumains, occasionally, some publications use the term "butelase"; however, these enzymes are closely related, true AEP/ligases of the legumain family [114].

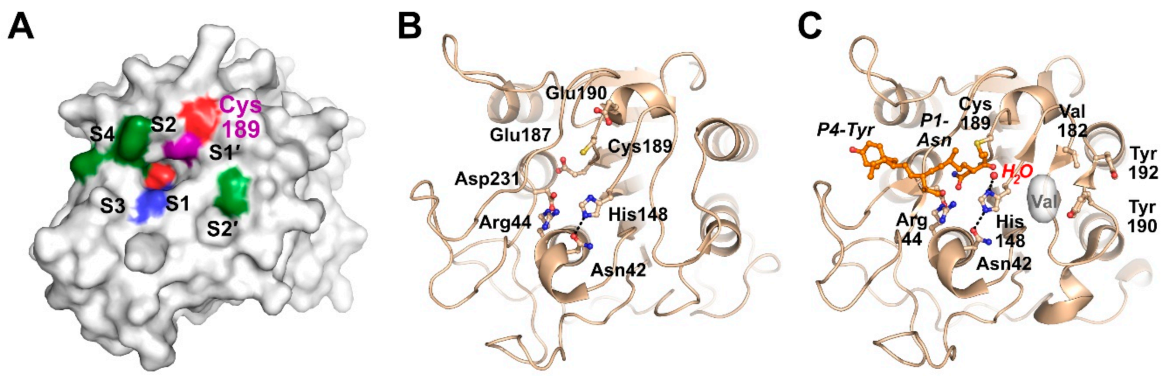

Figure 6. Legumain ligases. (**A**) Human legumain structure as surface representation. The catalytic Cys189 is labeled, as well as the specificity pockets from S4 to S2′ (PDB 4AW9). Residues that shape these pockets are colored as hydrophobic (green), basic (blue), and acidic (red). (**B**) Active site of plant γ-legumain from *A. thaliana*, shown without the so-called LSAM domain (PDB 5NIJ). The catalytic triad of Cys189, His148, and Asn42 is depicted as ball-and-stick models, as well as residues that shape the S1 subsite (Glu187, Asp231, Arg44, while His45 was omitted) and the S1′ pocket (Glu190). (**C**) Plant γ-legumain displayed as an acyl intermediate model based on the chloromethyl ketone complex (PDB 5OBT). The P1–Asn is well accommodated by the charged residues in the S1 pocket. The catalytic water is depicted as a red sphere, whereby its hydrolytic reactivity depends on the presence residues such as a Val in the prime side of human legumain. In plant legumains, it can be replaced by a Gly, resulting in a larger hydrophobic region, bordered by Val182, Tyr190, and Tyr192, favoring peptide ligase activity.

As already known for human legumain with an Asp near the active site converted to a succinimide in the backbone, a shift from acidic pH ≤ 5 to neutral pH ≈ 7 suffices to favor the cyclization activity as well in plant AEPs [115,116]. For both human and plant legumains, the major factor of the switch from protease activity to ligase activity is neutral pH; however, this is predominantly dependent on the respective tissue expression. Based on structures of the proform and mature γ-legumain from *Arabidopsis thaliana*, functional assays and molecular dynamics suggested a hydrophobic residue in position 184 (Gly in human legumain) nearby the catalytic water as a major mechanistic determinant of the ligase, enhanced by an interplay with the P2′-residue of the substrate (Figure 6C) [117,118]. A comparison of several plant AEPs identified a loop close to the active site as additional ligase enhancer, the so-called marker of ligase activity (MLA) [119]. Structural studies on AEPs from the *Violaceae* family identified the determinants of the ligase activity near the specificity subsite S1 as LAD1, including the central gatekeeper residue Gly(184) or Ile/Val/Cys/Ala in AEP-ligases, and near S2 and S1′ as LAD2, which is mostly a Gly–Pro pair, respectively [120]. Although first approaches of engineered AEP ligases for special peptide targets by plant and *E. coli* expression are made, their potential seems limited to the special case of cyclic peptides with certain sequence requirements [121]. Recently, it was shown that the predominantly proteolytic β-legumain from *Arabidopsis* can easily ligate linear precursors for SFTI variants, corroborating the general idea of proteases as catalyzing the forward and backward reaction [122]. However, a computational approach using quantum mechanics/molecular mechanics (QM/MM) calculations for human legumain indicated that two different backward reactions of the ligation are possible [123].

3.4. Trypsin and Trypsiligase

The interaction of trypsin with the above-mentioned inhibitors, e.g., SFTI, demonstrates the potential of trypsin as ligase in one of best studied examples. An NMR study confirmed the very rigid structure of SFTI bound to the active of trypsin, exhibiting only marginal conformational deviations with respect to the crystal structure (Figure 7A,B) [124]. It was also shown that trypsin resynthesizes the open P1–P1′ bond of the acyclic SFTI-1 [5,6] permutant, while it cleaves this scissile bond in cyclic SFTI-1, until in both cases, the equilibrium is reached with a ratio 9:1 (cyclic/open) [125]. Interestingly, polymer supported trypsin and chymotrypsin successfully linked the linear precursors of various chemically synthesized cyclotides, which in turn inhibited the free proteases in the picomolar range [126].

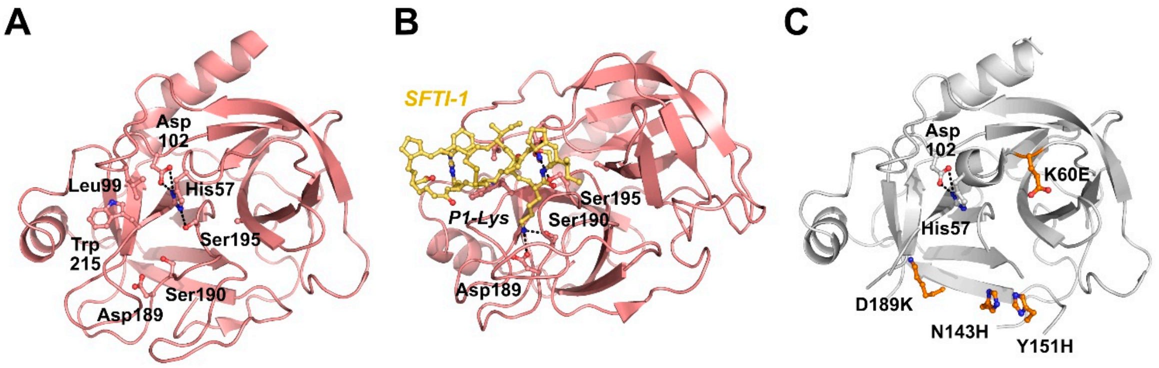

Figure 7. Trypsin ligase properties and trypsiligase. (**A**) The prototypic bovine β-trypsin structure (PDB 1TLP). The catalytic triad and specificity determining residues are shown as ball-and-stick models. (**B**) Trypsin bound sunflower trypsin inhibitor-1 (SFTI-1) in the active site with a scissile bond that is religated in the thermodynamic equilibrium (PDB 1SFI). The major specificity determinant is P1-Lys, while several prolines and an internal disulfide rigidify the cyclic inhibitor. (**C**) Trypsiligase derived from rat trypsin II with the four crucial mutations (orange) depicted as ball-and-stick side-chains (PDB 4NIV). The crystal structure shows a zymogen-like conformation, with large disordered regions around the catalytic Ser195, the S1 pocket and the 148-loop. Only the Lys60Glu mutation could be confirmed by electron density. The mutant residues Asn143His, Tyr151His, and Asp189Lys were modeled according to the positions in mature bovine trypsin.

A first step toward a trypsin-based ligase was the trypsin mutant Asp189Glu, with a significantly improved ligase activity, followed by the more efficient triple mutant Lys60Glu/Asp189Ser/Asp194Asn of trypsin [127,128]. In addition, changes at the oxyanion hole of trypsin reduced the hydrolytic activity and improved the ligase efficiency, in particular for the mutant Gln192Pro with protein substrates [129]. The anionic rat trypsin II quadruple mutant Lys60Glu/Asn143His/Glu151His/Asp189Lys, termed trypsiligase, adopts a zymogen-like state with a disordered activation domain in the absence of ligands, whereas an active and ordered conformation is induced by inhibitor binding (Figure 7C) [130]. Similar to sortases and bacterial transglutaminases, trypsiligase catalyzes N-terminal linking of antibody Fab fragments in high yields [131]. The proteolytic activity of this trypsin variant is strictly limited to the cleavage sequence Tyr–Arg–His, with the relatively rare occurrence of 0.5% in the human proteome, allowing for the successful N- and C-terminal ligation of Fab fragments with PEG and carboxyfluorescein [132]. A recent overview was published as a book chapter with detailed protocols for the usage of trypsiligase in polypeptide and protein synthesis [133].

4. Conclusions

Protease-derived peptide ligases cannot compete with recombinant expression in the area of protein synthesis. Nevertheless, they gain utility as increasingly valuable tools in combined approaches of organic polypeptide synthesis, which generates building blocks of up to 50 residues, and the covalent linkage by various ligases. An overview of the most relevant protease-ligases with *in vitro* reactions is shown in Table 1. In addition, ligases are most efficient in a variety of protein-modifying reactions, by attaching natural and synthetic peptides, molecular labels, glycans, membrane anchors, reactive organic molecular moieties, drugs, or even protein domains.

Table 1. Structural and functional parameters of ligating proteases. PDB codes, ligated bond (often involving modified amino acids or peptides), reaction conditions, and yields are shown. The listed enzymes were regularly used to catalyze peptide bond formation and for linking proteins with modifying molecules of various types. [1] HFIP, hexafluoroisopropylalcohol; [2] the yield is estimated from a chromatogram.; [3] Selenolsubtilisin C derived from subtilisin Carlsberg transfers a cinnamoyl acyl intermediate to H_2N–Gly; [4] yield estimated from a 4-times higher rate constant compared to thiolsubtilisin.

Protease/Ligase	PDB	P1–P1′ Bond	Reaction	Best Solvent	Yield	Citation
α-Chymotrypsin	2CHA	Pro–Arg	kinetic control	H_2O, pH 8.0, 10% MeOH	99%	[134]
α-Chymotrypsin		Tyr–Gly	mixed control	H_2O (−15 °C)	98%	[135]
β-Trypsin	1TPP	Arg–Lys	kinetic control	DMF/HFIP [1] 1:1, 4% H_2O	90%	[53]
β-Trypsin		Arg–Ala	mixed control	H_2O (−15 °C)	94%	[135]
Trypsiligase (rat)	4NIV	Arg–His	thermodynamic	H_2O, pH 7.8	40%	[131]
Elastase-1 (porcine)	3EST	Leu–Ala	kinetic control	H_2O, pH 9.0 (NaOH)	17%	[69]
Papain	1PPN	Phe–Leu	kinetic control	H_2O, 0.2 M K_2HPO_4, pH 7.8	56%	[136]
Pepsin	4PEP	Phe–Leu	thermodynamic	H_2O, pH 4.5, 10% DMF	100%	[137]
Carboxypeptidase B	1Z5R	Tyr–Lys	thermodynamic	H_2O, pH 6.7	10% [2]	[43]
Thermolysin	1LNF	Phe–Phe	thermodynamic	H_2O, pH 7.5, 20% DMSO	95%	[138]
Subtilisin BPN′	1S02	Phe–Leu	kinetic control	DMF/ H_2O 1:1	30%	[139]
Thiolsubtilisin C		Phe–Val	kinetic control	H_2O, pH 8.0, 40% DMF	90%	[63]
Selenosubtilisin C	1SEL	Cin–Gly [3]	kinetic control	H_2O, 0.1 M borate, pH 9.3	99% [4]	[62]
Subtiligase BPN′	-	Phe–Ala	thermodynamic	H_2O, pH 8.0	95%	[65]
Thymoligase	5OX2	Lys–Asp	kinetic control	H_2O, phosphate pH 8.3	94%	[90]
Sortase A	1T2P	Thr–Gly	thermodynamic	H_2O, phosphate pH 6.0	95%	[140]
γ-Legumain (plants)	5OBT	Asn–Gly	thermodynamic	H_2O, phosphate pH 6.0	95%	[140]

The incorporation of 4-fluorohistidine as a catalytic residue in RNase based on subtiligase ligation is one of the most stunning historical examples of the usefulness of protein ligases [78]. Nowadays, the rapidly evolving field of biological click chemistry is based on the apparently unlimited usage of non-natural amino acids (nnaa or nnca for non-canonical) with numerous applications from biophysical applications to analytical and pharmaceutical research. In combination with protease-derived ligases, the tedious recombinant expression of proteins for the incorporation of nnaas might be facilitated or circumvented. It remains to be seen how useful the more advanced, engineered protease-ligases can become outside labs performing basic research, such as in biotechnological and pharmaceutical applications. At least in theory, sortases, subtiligases, and trypsiligases should allow modifying any protein, which contains the required amino acid sequences for the specific ligation reaction. Thus, labels for monitoring and warheads for disabling disease-related enzymes could be linked to highly specific therapeutic proteins in an easy and elegant manner.

Funding: This study was supported by the Austrian Science Fund (FWF) with the D-A-CH grant I 3877-B21.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this review article.

Acknowledgments: I am grateful to Hans Brandstetter for critical assessment of the manuscript and his continuous support.

Conflicts of Interest: The author declares no conflict of interest.

References

1. van Hoff, J.H. Über die zunehmende Bedeutung der anorganischen Chemie. Vortrag, gehalten auf der 70. Versammlung der Gesellschaft deutscher Naturforscher und Ärzte zu Düsseldorf. *Z. Anorg. Allg. Chem.* **1898**, *18*, 1–13. [CrossRef]
2. Sawjalow, W.W. Zur Theorie der Eiweissverdauung. *Pflügers Arch. Ges. Physiol. Menschen Tiere* **1901**, *85*, 171–225. [CrossRef]
3. Borsook, H. Peptide Bond Formation. In *Advances in Protein Chemistry*; Anson, M.L., Bailey, K., Edsall, J.T., Eds.; Academic Press: New York, NY, USA, 1953; Volume 8, pp. 127–174.
4. Marahiel, M.A. Working outside the protein-synthesis rules: Insights into non-ribosomal peptide synthesis. *J. Pept. Sci.* **2009**, *15*, 799–807. [CrossRef] [PubMed]
5. Martin, J.F. α-Aminoadipyl-cysteinyl-valine Synthetases in β-Lactam Producing Organisms. *J. Antibiot.* **2000**, *53*, 1008–1021. [CrossRef] [PubMed]
6. Eifler, K.; Vertegaal, A.C.O. SUMOylation-Mediated Regulation of Cell Cycle Progression and Cancer. *Trends Biochem. Sci.* **2015**, *40*, 779–793. [CrossRef] [PubMed]
7. McClellan, A.J.; Laugesen, S.H.; Ellgaard, L. Cellular functions and molecular mechanisms of non-lysine ubiquitination. *Open Biol.* **2019**, *9*, 190147. [CrossRef] [PubMed]
8. Saska, I.; Craik, D.J. Protease-catalysed protein splicing: A new post-translational modification? *Trends Biochem. Sci.* **2008**, *33*, 363–368. [CrossRef]
9. Perler, F.B.; Olsen, G.J.; Adam, E. Compilation and analysis of intein sequences. *Nucleic Acids Res.* **1997**, *25*, 1087–1093. [CrossRef]
10. Novikova, O.; Topilina, N.; Belfort, M. Enigmatic Distribution, Evolution, and Function of Inteins. *J. Biol. Chem.* **2014**, *289*, 14490–14497. [CrossRef]
11. Aranko, A.S.; Wlodawer, A.; Iwaï, H. Nature's recipe for splitting inteins. *Prot. Eng. Des. Sel.* **2014**, *27*, 263–271. [CrossRef]
12. Ferrington, D.A.; Gregerson, D.S. Chapter 3—Immunoproteasomes: Structure, Function, and Antigen Presentation. In *Progress in Molecular Biology and Translational Science*; Grune, T., Ed.; Academic Press: San Diego, CA, USA, 2012; Volume 109, pp. 75–112.
13. Hanada, K.-i.; Yewdell, J.W.; Yang, J.C. Immune recognition of a human renal cancer antigen through post-translational protein splicing. *Nature* **2004**, *427*, 252–256. [CrossRef] [PubMed]
14. Vigneron, N.; Ferrari, V.; Stroobant, V.; Abi Habib, J.; Van den Eynde, B.J. Peptide splicing by the proteasome. *J. Biol. Chem.* **2017**, *292*, 21170–21179. [CrossRef] [PubMed]
15. Liepe, J.; Marino, F.; Sidney, J.; Jeko, A.; Bunting, D.E.; Sette, A.; Kloetzel, P.M.; Stumpf, M.P.H.; Heck, A.J.R.; Mishto, M. A large fraction of HLA class I ligands are proteasome-generated spliced peptides. *Science* **2016**, *354*, 354–358. [CrossRef] [PubMed]
16. Paes, W.; Leonov, G.; Partridge, T.; Nicastri, A.; Ternette, N.; Borrow, P. Elucidation of the Signatures of Proteasome-Catalyzed Peptide Splicing. *Front. Immunol.* **2020**, *11*, 11. [CrossRef]
17. Vigneron, N.; Stroobant, V.; Ferrari, V.; Abi Habib, J.; Van den Eynde, B.J. Production of spliced peptides by the proteasome. *Mol. Immunol.* **2019**, *113*, 93–102. [CrossRef]
18. Sarmiento, C.; Camarero, J.A. Biotechnological Applications of Protein Splicing. *Curr. Protein Pept. Sci.* **2019**, *20*, 408–424. [CrossRef]
19. Nuijens, T.; Toplak, A.; Schmidt, M.; Ricci, A.; Cabri, W. Natural Occurring and Engineered Enzymes for Peptide Ligation and Cyclization. *Front. Chem.* **2019**, *7*, 1–8. [CrossRef]
20. Kent, S. Chemical protein synthesis: Inventing synthetic methods to decipher how proteins work. *Bioorganic. Med. Chem.* **2017**, *25*, 4926–4937. [CrossRef]
21. Agouridas, V.; Diemer, V.; Melnyk, O. Strategies and open questions in solid-phase protein chemical synthesis. *Curr. Opin. Chem. Biol.* **2020**, *58*, 1–9. [CrossRef]

22. Agouridas, V.; El Mahdi, O.; Cargoët, M.; Melnyk, O. A statistical view of protein chemical synthesis using NCL and extended methodologies. *Bioorganic Med. Chem.* **2017**, *25*, 4938–4945. [CrossRef]
23. Schmohl, L.; Schwarzer, D. Chemo-enzymatic three-fragment assembly of semisynthetic proteins. *J. Pept. Sci.* **2014**, *20*, 145–151. [CrossRef] [PubMed]
24. Chang, T.K.; Jackson, D.Y.; Burnier, J.P.; Wells, J.A. Subtiligase: A tool for semisynthesis of proteins. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 12544–12548. [CrossRef] [PubMed]
25. Dall, E.; Brandstetter, H. Structure and function of legumain in health and disease. *Biochimie* **2016**, *122*, 126–150. [CrossRef] [PubMed]
26. James, A.M.; Haywood, J.; Mylne, J.S. Macrocyclization by asparaginyl endopeptidases. *New Phytol.* **2018**, *218*, 923–928. [CrossRef] [PubMed]
27. Bergmann, M.; Fraenkel-Conrat, H. The Role of Specificity in the Enzymatic Synthesis of Proteins: Syntheses with Intracellular Enzymes. *J. Biol. Chem.* **1937**, *119*, 707–720.
28. Bergmann, M.; Fraenkel-Conrat, H.; Doherty, D.G. The enzymatic synthesis of peptide bonds. *J. Biol. Chem.* **1938**, *124*, 1–6.
29. Fruton, J.S. Proteinase-catalyzed synthesis of peptide bonds. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1982**, *53*, 239–306. [CrossRef]
30. Martin, R.B. Free energies and equilibria of peptide bond hydrolysis and formation. *Biopolymers* **1998**, *45*, 351–353. [CrossRef]
31. Bordusa, F. Proteases in Organic Synthesis. *Chem. Rev.* **2002**, *102*, 4817–4868. [CrossRef]
32. Tschesche, H.; Kupfer, S. Hydrolysis-Resynthesis Equilibrium of the Lysine-15–Alanine-16 Peptide Bond in Bovine Trypsin Inhibitor (Kunitz). *Biol. Chem.* **1976**, *357*, 769–776. [CrossRef]
33. Guzman, F.; Barberis, S.; Illanes, A. Peptide synthesis: Chemical or enzymatic. *Electron. J. Biotechnol.* **2007**, *10*, 279–314. [CrossRef]
34. Morihara, K. Using proteases in peptide synthesis. *Trends Biotechnol.* **1987**, *5*, 164–170. [CrossRef]
35. Morihara, K.; Oka, T.; Tsuzuki, H. Semi-synthesis of human insulin by trypsin-catalysed replacement of Ala-B30 by Thr in porcine insulin. *Nature* **1979**, *280*, 412–413. [CrossRef] [PubMed]
36. Markussen, J. *Human Insulin*; MTP Press Limited: Lancaster, UK, 1987.
37. Kullmann, W. Proteases as catalysts for enzymic syntheses of opioid peptides. *J. Biol. Chem.* **1980**, *255*, 8234–8238. [PubMed]
38. Reddy, A.V. Thermolysin: A peptide forming enzyme. *Indian J. Biochem. Biophys.* **1991**, *21*, 10–15.
39. Bongers, J.; Heimer, E.P. Recent applications of enzymatic peptide synthesis. *Peptides* **1994**, *15*, 183–193. [CrossRef]
40. Dobry, A.; Fruton, J.S.; Sturtevant, J.M. Thermodynamics Of Hydrolysis Of Peptide Bonds. *J. Biol. Chem.* **1952**, *195*, 149–154.
41. Homandberg, G.A.; Mattis, J.A.; Laskowski, M. Synthesis of peptide bonds by proteinases. Addition of organic cosolvents shifts peptide bond equilibriums toward synthesis. *Biochemistry* **1978**, *17*, 5220–5227. [CrossRef]
42. Isowa, Y.; Ohmori, M.; Ichikawa, T.; Mori, K.; Nonaka, Y.; Kihara, K.-i.; Oyama, K.; Satoh, H.; Nishimura, S. The thermolysin-catalyzed condensation reactions of n-substituted aspartic and glutamic acids with phenylalanine alkyl esters. *Tetrahedron Lett.* **1979**, *20*, 2611–2612. [CrossRef]
43. Sealock, R.W.; Laskowski, M. Enzymic replacement of the arginyl by a lysyl residue in the reactive site of soybean trypsin inhibitor. *Biochemistry* **1969**, *8*, 3703–3710. [CrossRef]
44. Cassells, J.M.; Halling, P.J. Protease-catalysed peptide synthesis in aqueous-organic two-phase systems: Reactant precipitation and interfacial inactivation. *Enzym. Microb. Technol.* **1990**, *12*, 755–759. [CrossRef]
45. Zerner, B.; Bender, M.L. The Kinetic Consequences of the Acyl-Enzyme Mechanism for the Reactions of Specific Substrates with Chymotrypsin. *J. Am. Chem. Soc.* **1964**, *86*, 3669–3674. [CrossRef]
46. Schellenberger, V.; Jakubke, H.-D. Protease-catalyzed kinetically controlled peptide synthesis. *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 1437–1449. [CrossRef]
47. Luethi, P.; Luisi, P.L. Enzymic synthesis of hydrocarbon-soluble peptides with reverse micelles. *J. Am. Chem. Soc.* **1984**, *106*, 7285–7286. [CrossRef]
48. Jakubke, H.-D.; Kuhl, P.; Könnecke, A. Basic Principles of Protease-Catalyzed Peptide Bond Formation. *Angew. Chem. Int. Ed. Engl.* **1985**, *24*, 85–93. [CrossRef]
49. Schellenberger, V.; Schellenberger, U.; Mitin, Y.V.; Jakubke, H.-D. Characterization of the S′-subsite specificity of bovine pancreatic α-chymotrypsin via acyl transfer to added nucleophiles. *Eur. J. Biochem.* **1990**, *187*, 163–167. [CrossRef]
50. Imperiali, B.; Abeles, R.H. Extended binding inhibitors of chymotrypsin that interact with leaving group subsites S1′-S3′. *Biochemistry* **1987**, *26*, 4474–4477. [CrossRef]
51. Sakina, K.; Kawazura, K.; Morihara, K. Enzymatic synthesis of delta sleep-inducing peptide. *Int. J. Pept. Protein Res.* **1988**, *31*, 245–252. [CrossRef]
52. Schellenberger, V.; Schellenberger, U.; Jakubke, H.D.; Hänsicke, A.; Bienert, M.; Krause, E. Chymotrypsin-catalyzed fragment coupling synthesis of D-Phe(6)-GnRH. *Tetrahedron Lett.* **1990**, *31*, 7305–7306. [CrossRef]
53. Nishino, N.; Xu, M.; Mihara, H.; Fujimoto, T. Tryptic condensation combined with peptide segment synthesis–condensation strategy for the efficient synthesis of human growth hormone releasing factor (1–29) amide. *J. Chem. Soc. Chem. Commun.* **1992**, *10*, 648–650. [CrossRef]
54. Bongers, J.; Offord, R.E.; Felix, A.M.; Lambros, T.; Liu, W.; Ahmad, M.; Campbell, R.M.; Heimer, E.P. Comparison of Enzymatic Semisyntheses of Peptide Amides: Human Growth Hormone Releasing Factor and Analogs. *Biomed. Biochim. Acta* **1991**, *50*, S157–S162. [PubMed]

55. Morihara, K. Thermolysin Catalyzed Semisynthesis of Peptide Hormones by Introduction of Phe-NH2 or Tyr-NH2 at the Carboxyl Termini. *Biomed. Biochim. Acta* **1991**, *50*, S15–S18. [PubMed]
56. Henriksen, D.B.; Breddam, K.; Moeller, J.; Buchardt, O. Peptide amidation by chemical protein engineering. A combination of enzymic and photochemical synthesis. *J. Am. Chem. Soc.* **1992**, *114*, 1876–1877. [CrossRef]
57. Bongers, J.; Lambros, T.; Liu, W.; Ahmad, M.; Campbell, R.M.; Felix, A.M.; Heimer, E.P. Enzymic semisynthesis of a superpotent analog of human growth hormone-releasing factor. *J. Med. Chem.* **1992**, *35*, 3934–3941. [CrossRef]
58. Bongers, J.; Liu, W.; Lambros, T.; Breddam, K.; Campbell, R.M.; Felix, A.M.; Heimer, E.P. Peptide Synthesis Catalyzed by the Glu/Asp-specific Endopeptidase. Influence of the Ester Leaving Group of the Acyl Donor on Yield and Catalytic Efficiency. *Int. J. Pept. Protein Res.* **1994**, *44*, 123–129. [CrossRef]
59. Sahni, G.; Cho, Y.J.; Iyer, K.S.; Khan, S.A.; Seetharam, R.; Acharya, A.S. Semisynthetic hemoglobin A: Reconstitution of functional tetramer from semisynthetic alpha-globin. *Biochemistry* **1989**, *28*, 5456–5461. [CrossRef]
60. Groeger, C.; Wenzel, H.R.; Tschesche, H. Enzymatic semisynthesis of aprotinin homologues mutated in P′ positions. *J. Protein Chem.* **1991**, *10*, 245–251. [CrossRef]
61. Neet, K.E.; Koshland, D.E. The conversion of serine at the active site of subtilisin to cysteine: A chemical mutation. *Proc. Natl. Acad. Sci. USA* **1966**, *56*, 1606–1611. [CrossRef]
62. Wu, Z.P.; Hilvert, D. Conversion of a protease into an acyl transferase: Selenolsubtilisin. *J. Am. Chem. Soc.* **1989**, *111*, 4513–4514. [CrossRef]
63. Nakatsuka, T.; Sasaki, T.; Kaiser, E.T. Peptide segment synthesis catalyzed by the semisynthetic enzyme thiolsubtilisin. *J. Am. Chem. Soc.* **1987**, *109*, 3808–3810. [CrossRef]
64. Syed, R.; Wu, Z.P.; Hogle, J.M.; Hilvert, D. Crystal structure of selenosubtilisin at 2.0-.ANG. resolution. *Biochemistry* **1993**, *32*, 6157–6164. [CrossRef] [PubMed]
65. Abrahmsen, 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. [CrossRef] [PubMed]
66. Wong, C.H.; Chen, S.T.; Hennen, W.J.; Bibbs, J.A.; Wang, Y.F.; Liu, J.L.C.; Pantoliano, M.W.; Whitlow, M.; Bryan, P.N. Enzymes in organic synthesis: Use of subtilisin and a highly stable mutant derived from multiple site-specific mutations. *J. Am. Chem. Soc.* **1990**, *112*, 945–953. [CrossRef]
67. Gerisch, S.; Ullmann, G.; Stubenrauch, K.; Jakubke, H.D. Enzymatic peptide synthesis in frozen aqueous systems: Influence of modified reaction conditions on the peptide yield. *Biol. Chem. Hoppe Seyler* **1994**, *375*, 825–828.
68. Wehofsky, N.; Kirbach, S.W.; Haensler, M.; Wissmann, J.-D.; Bordusa, F. Substrate Mimetics and Freezing Strategy: A Useful Combination That Broadens the Scope of Proteases for Synthesis. *Org. Lett.* **2000**, *2*, 2027–2030. [CrossRef]
69. Haensler, M.; Wehofsky, N.; Gerisch, S.; Wissmann, J.D.; Jakubke, H.-D. Reverse catalysis of elastase from porcine pancreas in frozen aqueous systems. *Biol. Chem.* **1998**, *379*, 71–74.
70. Haensler, M.; Wissmann, H.-D.; Wehofsky, N. Enzymatic formation of Glu-Xaa and Asp-Xaa bonds using Glu/Asp-specific endopeptidase from Bacillus licheniformis in frozen aqueous systems. *J. Pept. Sci.* **2000**, *6*, 366–371. [CrossRef]
71. Jakubke, H.D.; Eichhorn, U.; Haensler, M.; Ullmann, D. Non-conventional enzyme catalysis: Application of proteases and zymogens in biotransformations. *Biol. Chem.* **1996**, *377*, 455–464.
72. Somalinga, B.R.; Roy, R.P. Volume Exclusion Effect as a Driving Force for Reverse Proteolysis: Implications For Polypeptide Assemblage in A Macromolecular Crowded Milieu. *J. Biol. Chem.* **2002**, *277*, 43253–43261. [CrossRef]
73. Tolbert, T.J.; Wong, C.-H. Subtilisin-Catalyzed Glycopeptide Condensation. *Methods Mol. Biol.* **2004**, *283*, 267–279.
74. Alden, R.A.; Birktoft, J.J.; Kraut, J.; Robertus, J.D.; Wright, C.S. Atomic coordinates for subtilisin BPN′ (or Novo). *Biochem. Biophys. Res. Commun.* **1971**, *45*, 337–344. [CrossRef]
75. Radisky, E.S.; Koshland, D.E. A clogged gutter mechanism for protease inhibitors. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 10316–10321. [CrossRef] [PubMed]
76. Kidd, R.D.; Sears, P.; Huang, D.-H.; Witte, K.; Wong, C.-H.; Farber, G.K. Breaking the low barrier hydrogen bond in a serine protease. *Protein Sci.* **1999**, *8*, 410–417. [CrossRef] [PubMed]
77. von der Osten, C.; Branner, S.; Hastrup, S.; Hedegaard, L.; Rasmussen, M.D.; Bisgård-Frantzen, H.; Carlsen, S.; Mikkelsen, J.M. Protein engineering of subtilisins to improve stability in detergent formulations. *J. Biotechnol.* **1993**, *28*, 55–68. [CrossRef]
78. Jackson, D.Y.; Burnier, J.; Quan, C.; Stanley, M.; Tom, J.; Wells, J.A. A designed peptide ligase for total synthesis of ribonuclease A with unnatural catalytic residues. *Science* **1994**, *266*, 243–247. [CrossRef] [PubMed]
79. Welker, E.; Scheraga, H.A. Use of Benzyl Mercaptan for Direct Preparation of Long Polypeptide Benzylthio Esters as Substrates of Subtiligase. *Biochem. Biophys. Res. Commun.* **1999**, *254*, 147–151. [CrossRef]
80. Atwell, S.; Wells, J.A. Selection for improved subtiligases by phage display. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9497–9502. [CrossRef]
81. Tan, X.-H.; Wirjo, A.; Liu, C.-F. An Enzymatic Approach to the Synthesis of Peptide Thioesters: Mechanism and Scope. *ChemBioChem* **2007**, *8*, 1512–1515. [CrossRef]
82. Yoshihara, H.A.I.; Mahrus, S.; Wells, J.A. Tags for labeling protein N-termini with subtiligase for proteomics. *Bioorganic. Med. Chem. Lett.* **2008**, *18*, 6000–6003. [CrossRef]
83. Wildes, D.; Wells, J.A. Sampling the N-terminal proteome of human blood. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 4561–4566. [CrossRef]

84. Crawford, E.D.; Seaman, J.E.; Agard, N.; Hsu, G.W.; Julien, O.; Mahrus, S.; Nguyen, H.; Shimbo, K.; Yoshihara, H.A.I.; Zhuang, M.; et al. The DegraBase: A Database of Proteolysis in Healthy and Apoptotic Human Cells. *Mol. Cell. Proteomics* **2013**, *12*, 813–824. [CrossRef] [PubMed]
85. Henager, S.H.; Chu, N.; Chen, Z.; Bolduc, D.; Dempsey, D.R.; Hwang, Y.; Wells, J.; Cole, P.A. Enzyme-catalyzed expressed protein ligation. *Nat. Methods* **2016**, *13*, 925–927. [CrossRef] [PubMed]
86. Weeks, A.M.; Wells, J.A. Engineering peptide ligase specificity by proteomic identification of ligation sites. *Nat. Chem. Biol.* **2018**, *14*, 50–57. [CrossRef] [PubMed]
87. Weeks, A.M.; Wells, J.A. Subtiligase-Catalyzed Peptide Ligation. *Chem. Rev.* **2020**, *120*, 3127–3160. [CrossRef]
88. Schmidt, M.; Huang, Y.-H.; Texeira de Oliveira, E.F.; Toplak, A.; Wijma, H.J.; Janssen, D.B.; van Maarseveen, J.H.; Craik, D.J.; Nuijens, T. Efficient Enzymatic Cyclization of Disulfide-Rich Peptides by Using Peptide Ligases. *ChemBioChem* **2019**, *20*, 1524–1529. [CrossRef]
89. Nguyen, T.X.; Cole, A.M.; Lehrer, R.I. Evolution of primate θ-defensins: A serpentine path to a sweet tooth. *Peptides* **2003**, *24*, 1647–1654. [CrossRef]
90. Schmidt, M.; Toplak, A.; Rozeboom, H.J.; Wijma, H.J.; Quaedflieg, P.J.L.M.; van Maarseveen, J.H.; Janssen, D.B.; Nuijens, T. Design of a substrate-tailored peptiligase variant for the efficient synthesis of thymosin-α1. *Org. Biomol. Chem.* **2018**, *16*, 609–618. [CrossRef]
91. Ton-That, H.; Liu, G.; Mazmanian, S.K.; Faull, K.F.; Schneewind, O. Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of *Staphylococcus aureus* at the LPXTG motif. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 12424–12429. [CrossRef]
92. Ilangovan, U.; Ton-That, H.; Iwahara, J.; Schneewind, O.; Clubb, R.T. Structure of sortase, the transpeptidase that anchors proteins to the cell wall of *Staphylococcus aureus*. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 6056–6061. [CrossRef]
93. Zong, Y.; Bice, T.W.; Ton-That, H.; Schneewind, O.; Narayana, S.V.L. Crystal Structures of Staphylococcus aureus Sortase A and Its Substrate Complex. *J. Biol. Chem.* **2004**, *279*, 31383–31389. [CrossRef]
94. Dramsi, S.; Trieu-Cuot, P.; Bierne, H. Sorting sortases: A nomenclature proposal for the various sortases of Gram-positive bacteria. *Res. Microbiol.* **2005**, *156*, 289–297. [CrossRef] [PubMed]
95. Mao, H.; Hart, S.A.; Schink, A.; Pollok, B.A. Sortase-Mediated Protein Ligation: A New Method for Protein Engineering. *J. Am. Chem. Soc.* **2004**, *126*, 2670–2671. [CrossRef] [PubMed]
96. Parthasarathy, R.; Subramanian, S.; Boder, E.T. Sortase A as a Novel Molecular "Stapler" for Sequence-Specific Protein Conjugation. *Bioconjug. Chem.* **2007**, *18*, 469–476. [CrossRef] [PubMed]
97. Pritz, S.; Wolf, Y.; Kraetke, O.; Klose, J.; Bienert, M.; Beyermann, M. Synthesis of Biologically Active Peptide Nucleic Acid−Peptide Conjugates by Sortase-Mediated Ligation. *J. Org. Chem.* **2007**, *72*, 3909–3912. [CrossRef]
98. Guo, X.; Wang, Q.; Swarts, B.M.; Guo, Z. Sortase-Catalyzed Peptide−Glycosylphosphatidylinositol Analogue Ligation. *J. Am. Chem. Soc.* **2009**, *131*, 9878–9879. [CrossRef]
99. Levary, D.A.; Parthasarathy, R.; Boder, E.T.; Ackerman, M.E. Protein-Protein Fusion Catalyzed by Sortase A. *PLoS ONE* **2011**, *6*, e18342. [CrossRef]
100. Strijbis, K.; Spooner, E.; Ploegh, H.L. Protein Ligation in Living Cells Using Sortase. *Traffic* **2012**, *13*, 780–789. [CrossRef]
101. Witte, M.D.; Theile, C.S.; Wu, T.; Guimaraes, C.P.; Blom, A.E.M.; Ploegh, H.L. Production of unnaturally linked chimeric proteins using a combination of sortase-catalyzed transpeptidation and click chemistry. *Nat. Protoc.* **2013**, *8*, 1808–1819. [CrossRef]
102. Kornberger, P.; Skerra, A. Sortase-catalyzed in vitro functionalization of a HER2-specific recombinant Fab for tumor targeting of the plant cytotoxin gelonin. *mAbs* **2014**, *6*, 354–366. [CrossRef]
103. Jia, X.; Kwon, S.; Wang, C.-I.A.; Huang, Y.-H.; Chan, L.Y.; Tan, C.C.; Rosengren, K.J.; Mulvenna, J.P.; Schroeder, C.I.; Craik, D.J. Semienzymatic Cyclization of Disulfide-rich Peptides Using Sortase A. *J. Biol. Chem.* **2014**, *289*, 6627–6638. [CrossRef]
104. Massa, S.; Vikani, N.; Betti, C.; Ballet, S.; Vanderhaegen, S.; Steyaert, J.; Descamps, B.; Vanhove, C.; Bunschoten, A.; van Leeuwen, F.W.B.; et al. Sortase A-mediated site-specific labeling of camelid single-domain antibody-fragments: A versatile strategy for multiple molecular imaging modalities. *Contrast Media Mol. Imaging* **2016**, *11*, 328–339. [CrossRef] [PubMed]
105. Zeng, M.; Ye, F.; Xu, J.; Zhang, M. PDZ Ligand Binding-Induced Conformational Coupling of the PDZ–SH3–GK Tandems in PSD-95 Family MAGUKs. *J. Mol. Biol.* **2018**, *430*, 69–86. [CrossRef] [PubMed]
106. Fauser, J.; Savitskiy, S.; Fottner, M.; Trauschke, V.; Gulen, B. Sortase-Mediated Quantifiable Enzyme Immobilization on Magnetic Nanoparticles. *Bioconjugate Chem.* **2020**, *31*, 1883–1892. [CrossRef] [PubMed]
107. Li, J.; Zhang, Y.; Soubias, O.; Khago, D.; Chao, F.-A.; Li, Y.; Shaw, K.; Byrd, R.A. Optimization of sortase A ligation for flexible engineering of complex protein systems. *J. Biol. Chem.* **2020**, *295*, 2664–2675. [CrossRef] [PubMed]
108. Hirakawa, H.; Ishikawa, S.; Nagamune, T. Ca2+-independent sortase-A exhibits high selective protein ligation activity in the cytoplasm of Escherichia coli. *Biotechnol. J.* **2015**, *10*, 1487–1492. [CrossRef]
109. Vorster, B.J.; Cullis, C.A.; Kunert, K.J. Plant Vacuolar Processing Enzymes. *Front. Plant Sci.* **2019**, *10*, 479. [CrossRef]
110. Bernath-Levin, K.; Nelson, C.; Elliott, A.G.; Jayasena, A.S.; Millar, A.H.; Craik, D.J.; Mylne, J.S. Peptide Macrocyclization by a Bifunctional Endoprotease. *Chem. Biol.* **2015**, *22*, 571–582. [CrossRef]
111. de Veer, S.J.; Kan, M.-W.; Craik, D.J. Cyclotides: From Structure to Function. *Chem. Rev.* **2019**, *119*, 12375–12421. [CrossRef]
112. Chaudhuri, D.; Aboye, T.; Camarero, J.A. Using backbone-cyclized Cys-rich polypeptides as molecular scaffolds to target protein–protein interactions. *Biochem. J.* **2019**, *476*, 67–83. [CrossRef]

113. Saska, I.; Gillon, A.D.; Hatsugai, N.; Dietzgen, R.G.; Hara-Nishimura, I.; Anderson, M.A.; Craik, D.J. An Asparaginyl Endopeptidase Mediates in Vivo Protein Backbone Cyclization. *J. Biol. Chem.* **2007**, *282*, 29721–29728. [CrossRef]
114. Nguyen, G.K.T.; Wang, S.; Qiu, Y.; Hemu, X.; Lian, Y.; Tam, J.P. Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. *Nat. Chem. Biol.* **2014**, *10*, 732–738. [CrossRef] [PubMed]
115. Haywood, J.; Schmidberger, J.W.; James, A.M.; Nonis, S.G.; Sukhoverkov, K.V.; Elias, M.; Bond, C.S.; Mylne, J.S. Structural basis of ribosomal peptide macrocyclization in plants. *Elife* **2018**, *7*, 1–22. [CrossRef] [PubMed]
116. Dall, E.; Fegg, J.C.; Briza, P.; Brandstetter, H. Structure and Mechanism of an Aspartimide-Dependent Peptide Ligase in Human Legumain. *Angew. Chem. Int. Ed.* **2015**, *54*, 2917–2921. [CrossRef] [PubMed]
117. Zauner, F.B.; Dall, E.; Regl, C.; Grassi, L.; Huber, C.G.; Cabrele, C.; Brandstetter, H. Crystal Structure of Plant Legumain Reveals a Unique Two-Chain State with pH-Dependent Activity Regulation. *Plant Cell* **2018**, *30*, 686–699. [CrossRef]
118. Zauner, F.B.; Elsässer, B.; Dall, E.; Cabrele, C.; Brandstetter, H. Structural analyses of Arabidopsis thaliana legumain γ reveal differential recognition and processing of proteolysis and ligation substrates. *J. Biol. Chem.* **2018**, *293*, 8934–8946. [CrossRef]
119. Jackson, M.A.; Gilding, E.K.; Shafee, T.; Harris, K.S.; Kaas, Q.; Poon, S.; Yap, K.; Jia, H.; Guarino, R.; Chan, L.Y.; et al. Molecular basis for the production of cyclic peptides by plant asparaginyl endopeptidases. *Nat. Commun.* **2018**, *9*, 2411. [CrossRef]
120. Hemu, X.; El Sahili, A.; Hu, S.; Wong, K.; Chen, Y.; Wong, Y.H.; Zhang, X.; Serra, A.; Goh, B.C.; Darwis, D.A.; et al. Structural determinants for peptide-bond formation by asparaginyl ligases. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 11737–11746. [CrossRef]
121. Harris, K.S.; Poon, S.; Quimbar, P.; Anderson, M.A. In Vitro and In Planta Cyclization of Target Peptides Using an Asparaginyl Endopeptidase from Oldenlandia affinis. In *Enzyme-Mediated Ligation Methods*; Nuijens, T., Schmidt, M., Eds.; Springer: New York, NY, 2019; pp. 211–235. [CrossRef]
122. Dall, E.; Zauner, F.B.; Soh, W.T.; Demir, F.; Dahms, S.O.; Cabrele, C.; Huesgen, P.F.; Brandstetter, H. Structural and functional studies of Arabidopsis thaliana legumain beta reveal isoform specific mechanisms of activation and substrate recognition. *J. Biol. Chem.* **2020**, *295*, 13047–13064. [CrossRef]
123. Elsässer, B.; Zauner, F.B.; Messner, J.; Soh, W.T.; Dall, E.; Brandstetter, H. Distinct Roles of Catalytic Cysteine and Histidine in the Protease and Ligase Mechanisms of Human Legumain As Revealed by DFT-Based QM/MM Simulations. *ACS Catal.* **2017**, *7*, 5585–5593. [CrossRef]
124. Korsinczky, M.L.J.; Schirra, H.J.; Rosengren, K.J.; West, J.; Condie, B.A.; Otvos, L.; Anderson, M.A.; Craik, D.J. Solution structures by 1H NMR of the novel cyclic trypsin inhibitor SFTI-1 from sunflower seeds and an acyclic permutant. *J. Mol. Biol.* **2001**, *311*, 579–591. [CrossRef]
125. Marx, U.C.; Korsinczky, M.L.J.; Schirra, H.J.; Jones, A.; Condie, B.; Otvos, L.; Craik, D.J. Enzymatic Cyclization of a Potent Bowman-Birk Protease Inhibitor, Sunflower Trypsin Inhibitor-1, and Solution Structure of an Acyclic Precursor Peptide. *J. Biol. Chem.* **2003**, *278*, 21782–21789. [CrossRef] [PubMed]
126. Thongyoo, P.; Roqué-Rosell, N.; Leatherbarrow, R.J.; Tate, E.W. Chemical and biomimetic total syntheses of natural and engineered MCoTI cyclotides. *Org. Biomol. Chem.* **2008**, *6*, 1462–1470. [CrossRef] [PubMed]
127. Xu, S.; Rall, K.; Bordusa, F. Enzymatic Coupling of Specific Peptides at Nonspecific Ligation Sites: Effect of Asp189Glu Mutation in Trypsin on Substrate Mimetic-Mediated Reactions. *J. Org. Chem.* **2001**, *66*, 1627–1632. [CrossRef] [PubMed]
128. Rall, K.; Bordusa, F. Substrate Mimetics-Specific Peptide Ligases: Studies on the Synthetic Utility of a Zymogen and Zymogen-Like Enzymes. *J. Org. Chem.* **2002**, *67*, 9103–9106. [CrossRef] [PubMed]
129. Franke, L.; Liebscher, S.; Bordusa, F. Engineering the oxyanion hole of trypsin for promoting the reverse of proteolysis. *J. Pept. Sci.* **2014**, *20*, 128–136. [CrossRef]
130. Liebscher, S.; Schöpfel, M.; Aumüller, T.; Sharkhuukhen, A.; Pech, A.; Höss, E.; Parthier, C.; Jahreis, G.; Stubbs, M.T.; Bordusa, F. N-Terminal Protein Modification by Substrate-Activated Reverse Proteolysis. *Angew. Chem. Int. Ed.* **2014**, *53*, 3024–3028. [CrossRef]
131. Liebscher, S.; Kornberger, P.; Fink, G.; Trost-Gross, E.-M.; Höss, E.; Skerra, A.; Bordusa, F. Derivatization of Antibody Fab Fragments: A Designer Enzyme for Native Protein Modification. *ChemBioChem* **2014**, *15*, 1096–1100. [CrossRef]
132. Meyer, C.; Liebscher, S.; Bordusa, F. Selective Coupling of Click Anchors to Proteins via Trypsiligase. *Bioconjug. Chem.* **2016**, *27*, 47–53. [CrossRef]
133. Liebscher, S.; Bordusa, F. Site-Specific Modification of Proteins via Trypsiligase. In *Bioconjugation Methods in Molecular Biology*; Massa, S., Devoogdt, N., Eds.; Humana: New York, NY, USA, 2019; Volume 2033.
134. Günther, R.; Thust, S.; Hofmann, H.-J.; Bordusa, F. Trypsin-specific acyl-4-guanidinophenyl esters for α-chymotrypsin-catalysed reactions. *Eur. J. Biochem.* **2000**, *267*, 3496–3501. [CrossRef]
135. Wehofsky, N.; Haensler, M.; Kirbach, S.W.; Wissmann, J.-D.; Bordusa, F. Effect of freezing on the enzymatic coupling of specific amino acid-containing peptide fragments. *Tetrahedron Asymmetry* **2000**, *11*, 2421–2428. [CrossRef]
136. Haensler, M.; Jakubke, H.-D. Nonconventional protease catalysis in frozen aqueous solutions. *J. Pept. Sci.* **1996**, *2*, 279–289. [CrossRef] [PubMed]
137. Morihara, K.; Oka, T. Peptide Bond Synthesis Catalyzed by Subtilisin, Papain, and Pepsin. *J. Biochem.* **1981**, *89*, 385–395. [CrossRef] [PubMed]
138. Wayne, S.I.; Fruton, J.S. Thermolysin-catalyzed peptide bond synthesis. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 3241–3244. [CrossRef] [PubMed]

139. Barbas, C.F.; Matos, J.R.; West, J.B.; Wong, C.H. A search for peptide ligase: Cosolvent-mediated conversion of proteases to esterases for irreversible synthesis of peptides. *J. Am. Chem. Soc.* **1988**, *110*, 5162–5166. [CrossRef]
140. Nguyen, G.K.T.; Kam, A.; Loo, S.; Jansson, A.E.; Pan, L.X.; Tam, J.P. Butelase 1: A Versatile Ligase for Peptide and Protein Macrocyclization. *J. Am. Chem. Soc.* **2015**, *137*, 15398–15401. [CrossRef]