Biochemical Society Transactions (2020)
https://doi.org/10.1042/BST20200001

PORTLAND PRESS

Review Article

Engineered peptide ligases for cell signaling and bioconjugation

Clara L. Frazier and Amy M. Weeks

Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706, U.S.A

Correspondence: Amy M. Weeks (amweeks@wisc.edu)

Enzymes that catalyze peptide ligation are powerful tools for site-specific protein bioconjugation and the study of cellular signaling. Peptide ligases can be divided into two classes: proteases that have been engineered to favor peptide ligation, and protease-related enzymes with naturally evolved peptide ligation activity. Here, we provide a review of key natural peptide ligases and proteases engineered to favor peptide ligation activity. We cover the protein engineering approaches used to generate and improve these tools, along with recent biological applications, advantages, and limitations associated with each enzyme. Finally, we address future challenges and opportunities for further development of peptide ligases as tools for biological research.

Introduction

The ability to site-specifically modify proteins with chemical probes and payloads that cannot be genetically encoded has advanced biological research by enabling drug discovery, synthesis of protein conjugates, application of advanced imaging methods, and chemoproteomic profiling of specific protein features. Protein N and C termini are attractive sites for site-specific protein bioconjugation because they occur exactly once in each polypeptide chain. While a number of chemical strategies to target protein termini have been developed [1–6], these often suffer from poor selectivity for protein N or C termini over lysine or aspartate/glutamate side chains, respectively. Based on their ability to perform molecular recognition of protein termini with absolute chemoselectivity, peptide ligase enzymes are an attractive alternative approach for site-specific modification of protein termini and represent powerful tools for probing biology (Figure 1).

Enzymes that catalyze peptide ligation can be categorized into two main classes: proteases that have been engineered or otherwise optimized to catalyze peptide ligation; and protease-related enzymes that have naturally evolved peptide ligation activity. Both classes of enzymes have been applied for site-specific bioconjugation to protein N and C termini and have been targeted with protein engineering to generate variants with higher catalytic activity and altered specificity. In this review, we focus on protein engineering efforts to optimize these enzymes as tools to advance biological research. We cover the advantages and limitations of each enzymatic technology and discuss recent applications of these tools for modification of individual proteins and for global profiling of new protein termini generated by cellular proteolytic signaling pathways. Finally, we discuss the remaining limitations of engineered peptide ligases and future opportunities for the application of engineered enzymes for site-specific modification of N and C termini to advance our understanding of biological signaling pathways.

Received: 11 April 2020
Revised: 19 May 2020
Accepted: 21 May 2020

Version of Record published:
15 June 2020

Engineering proteases with non-natural ligation activity

Proteases have been a particularly popular target for protein engineering based in part on their potential to catalyze peptide ligation, the microscopic reverse of the typical proteolytic reaction (Figure 2A). Attractive features of proteases as peptide ligation tools include their activity under physiological

Biochemical Society Transactions (2020)
https://doi.org/10.1042/BST20200001

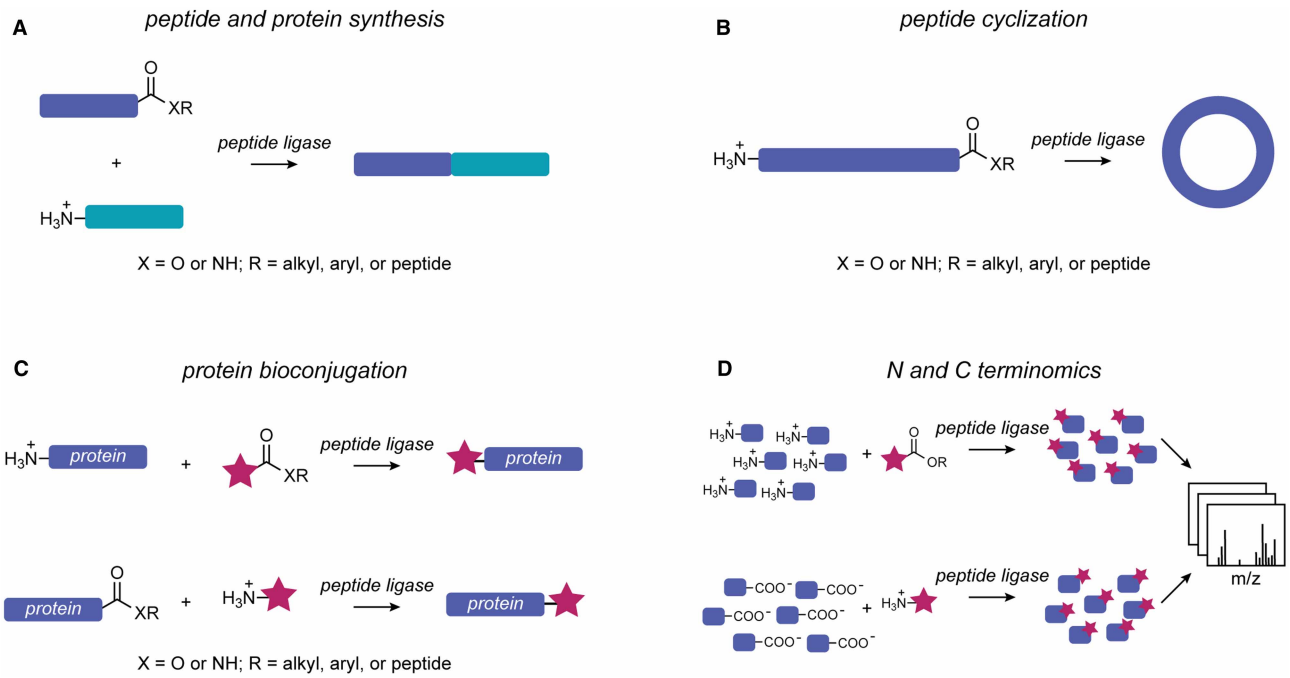

Figure 1. Applications of engineered peptide ligases in biological research.

conditions, the availability of both highly promiscuous [7] and highly sequence specific [8] scaffolds as starting points for protein engineering, and their ability to operate on peptide substrates in aqueous solution and in the absence of side chain protecting groups. Although proteases are subject to the reversible nature of the hydrolytic reaction, their direct application to peptide bond synthesis is limited by the position of the equilibrium between the proteolytic products and the ligation product [9,10]. Under physiological conditions, protease-catalyzed hydrolysis predominates over protease-catalyzed peptide ligation due to the higher thermodynamic stability of the hydrolytic products ($\Delta G^\circ \sim -3$ kcal/mol) [9]. However, under kinetically controlled conditions, the yield of peptide ligation is determined by the kinetic properties of the protease instead of the thermodynamic stabilities of the substrates and products ($\Delta G^{\ddagger}$) [11]. For example, in proteases that form an acyl-enzyme intermediate, acylation by esters kinetically outcompetes acylation by amides, minimizing protease-catalyzed hydrolysis of the ligation product and improving product yield on short time scales. Under these conditions, the ratio of hydrolytic and ligated products formed will depend on the rate of nucleophilic attack by water (hydrolysis) versus the ligation partner (aminolysis) rather than the thermodynamic stability of the products and reactants [10,11]. Protein engineering has been applied to shift the kinetic properties of proteases to favor the ligation reaction over the hydrolysis reaction under kinetic control. Protein engineering also has been used to modulate substrate specificity in protease scaffolds to allow for an increased array of potential substrates for ligation. These methods have been used to enhance ligation activity and to alter substrate specificity for a variety of natural proteases, including subtilisin (Figure 2B), trypsin (Figure 2C), and carboxypeptidase Y (Figure 2D).

Subtilisin variants

Subtilisin BPN' is a broad specificity serine protease produced by *Bacillus amyloliquefaciens* [12] that served as a starting point for the design of the engineered peptide ligase subtiligase. Subtiligase is a double mutant of subtilisin that was produced by mutating the catalytic Ser to Cys (S221C) [13] and by introducing a Pro-to-Ala mutation (P225A) to spatially accommodate the larger cysteine nucleophile [1]. The S221C mutation diminishes subtilisin's amidase activity, reducing secondary hydrolysis of the ligated peptide product, but maintains the enzyme's ability to form a thioacyl-enzyme intermediate from an ester substrate. The P225A mutation enhances the aminolysis to hydrolysis ratio by >100-fold, generating a variant that favors peptide ligase activity over the naturally evolved protease activity [1,14]. With these two mutations, subtiligase efficiently catalyzes

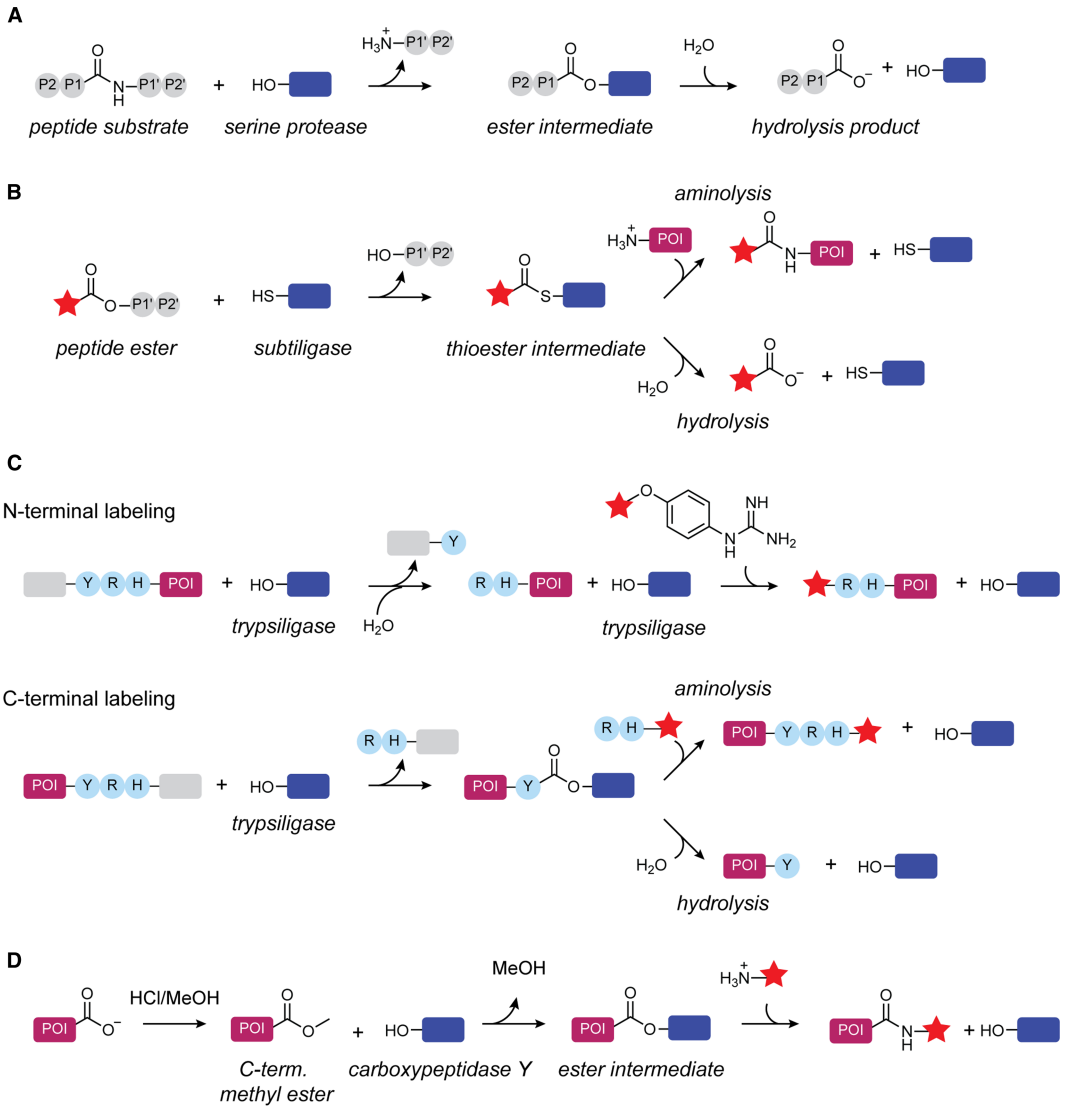

Figure 2. Repurposing proteases for peptide ligation.

(**A**) Serine protease mechanism. An amide bond is attacked by the catalytic Ser, generating an acyl-enzyme intermediate, which is hydrolyzed. (**B**) Subtiligase-catalyzed peptide ligation. A peptide ester forms a thioacyl-enzyme intermediate with the catalytic Cys, which can undergo hydrolysis or aminolysis. (**C**) N- and C-terminal trypsiligase-catalyzed peptide ligation reactions. In the N-terminal scheme, a protein of interest (POI) with an N-terminal YRH recognition motif is conjugated with a acyl-4-guanidinophenyl (OGp) ester. In the C-terminal scheme, a POI with a C-terminal YRH recognition motif is conjugated with an Arg-His acyl acceptor peptide. (**D**) Carboxypeptidase Y-catalyzed peptide ligation reaction. Carboxypeptidase Y catalyzes the conjugation of a C-terminal methyl ester with a labeled lysine derivative.

peptide bond formation between a peptide C-terminal ester and the N-terminal α-amine of a peptide or protein (Figure 2B). The peptide ligation activity of subtiligase was further optimized using phage display techniques that enabled screening of $>10^9$ enzyme variants [15]. The phage display approach led to the identification of subtiligase variants with a >2-fold improvement in peptide ligation activity.

Since the initial report of the subtiligase variant of subtilisin, numerous other properties of the enzyme, including stability and sequence specificity, have been altered through protein engineering. The stabiligase variant, which functions in the presence of detergents and chaotropic agents, was developed by introducing five previously identified mutations (M50F, N76D, N109S, K213R, and N218S) that stabilize subtilisin [14]. A

Biochemical Society Transactions (2020)
https://doi.org/10.1042/BST20200001

stabilized Ca^{2+}-independent subtiligase variant, peptiligase, was similarly generated by introducing the subtiligase mutations into a previously reported Ca^{2+}-independent subtilisin variant that was engineered for high stability [16]. Both stabiligase and peptiligase tolerate high temperatures, organic cosolvents, and denaturing agents, making them suitable for peptide ligation applications that require these conditions.

Protein engineering approaches have also been applied to alter subtiligase specificity. Although it is assumed that subtiligase variants have retained the broad specificity of subtilisin, they do exhibit some specificity for substrate binding and ligation. Subtiligase specificity has been well characterized using a variety of techniques, including substrate phage [14,15], synthetic peptide libraries [17,18], and proteome-derived peptide libraries [19]. While subtiligase generally maintains subtilisin's specificity on the non-prime side (according to Schechter and Berger nomenclature [20]), these techniques revealed sequence biases on the prime side that were not apparent from studies of subtilisin. Subtiligase prefers small amino acids at the P1′ position and aromatic or large hydrophobic amino acids at the P2′ position [14,17–19]. The proteomic identification of ligation sites (PILS) method, which utilizes highly diverse, proteome-derived peptide libraries as a pool of substrates for subtiligase, revealed that non-optimal sequences are accepted with varying degrees of efficiency. The PILS approach was further used to characterize the efficiency of peptide ligation for >25 000 substrate-subtiligase variant pairings, leading to the identification of 72 subtiligase mutants for labeling previously inefficient N-terminal sequences [19]. Additional subtiligase mutants that alter substrate specificity on both the prime and non-prime sides have also been developed by rational design and by screening variant activity against synthetic peptide libraries [1,14,21–23].

Based on its ability to catalyze peptide ligation, subtiligase has been deployed for a variety of N-terminal bioconjugation applications that have advanced the biological sciences. Subtiligase is an attractive tool for N-terminal modification because of its chemoselectivity for N-terminal α-amines over lysine ε-amines without restrictive N-terminal sequence requirements. In this context, subtiligase and subtiligase variants have been used as tools for synthesis of linear peptides [24] and for N-to-C-terminal peptide macrocyclization [25,26]. Because the main determinants of the ability of a peptide or protein N terminus to be modified by subtiligase or its variants are the accessibility of its N terminus and its ability to bind subtiligase in an extended conformation [14], subtiligase has also found broad applicability in modification of proteins with peptide esters. This strategy has been used to introduce reactive functional groups, fluorophores, cytotoxic drugs, polymers, and other biological molecules, such as DNA, into protein scaffolds [14,19]. Subtiligase has also been optimized as a catalyst for expressed protein ligation that eliminates the need for a Cys residue at the ligation junction. In this method, intein chemistry is used to generate a protein C-terminal thioester, which can be accepted as a substrate by subtiligase, enabling modification of the protein C terminus [27]. Subtiligase-catalyzed expressed protein ligation was used for synthesis of a monophosphorylated version of the tumor suppressor lipid phosphatase PTEN, revealing how site-specific phosphorylation inhibits PTEN's activity [28]. Beyond synthesis and modification of individual proteins and peptides, subtiligase has been used for global capture of protein N termini to study proteolytic signaling pathways. In this method, known as subtiligase N terminomics, subtiligase is used to globally modify N termini in a cell lysate with a peptide ester linked to a biotin affinity tag and a TEV protease cleavage sequence, allowing for the enrichment, selective elution, and analysis of protein N termini by LC–MS/MS [29]. This approach has been applied to study proteolytic cleavage events that occur in the context of apoptosis [29,30], viral [31] and bacterial infection [32], inflammation [33] and protein trafficking [34]. In a similar manner, subtiligase has also been used for protease substrate discovery, revealing specific protein substrates cleaved by inflammatory and apoptotic caspases [33,35,36] and a protease associated with Zika virus [31] in the context of cell lysate. Subtiligase applications have recently been reviewed in more detail [37].

Subtiligase-catalyzed N-terminal modification has the advantages that it relies on an enzyme that can be recombinantly expressed; that high yields can be achieved on a fast timescale; that it requires only a catalytic amount of enzyme; and that it is broadly sequence compatible. Despite the broad utility that these advantages provide, subtiligase also has several limitations that remain to be overcome. Subtiligase reaction yield is limited by peptide ester hydrolysis, so a large excess of peptide ester is often required to achieve high yields of ligation product. Although many subtiligase variants with altered specificity have been developed, none of these achieve truly comprehensive recognition of all N-terminal sequences, making selection of an appropriate mutant an important consideration [19]. However, the specificity mutants do retain broad specificity, so they are not applicable in situations in which the goal is to modify a single specific sequence in a complex background such as a cell lysate. While subtiligase N terminomics has led to the identification of many protease substrates, so far the technique has been applicable only in cell lysates and on the surface of live cells [38], but not to the interior

of live cells because available subtiligase substrates are cell-impermeable peptides. Future developments in protein engineering of subtiligase that enhance ligation efficiency and identify mutants with significantly broadened or narrowed specificity would therefore expand the applicability of subtiligase to N-terminal modification. Development of cell-permeable subtiligase substrates would allow subtiligase N terminomics to be performed inside living cells, enabling mapping of proteolytic signaling pathways with subcellular spatial resolution.

Trypsin variants

The serine protease trypsin has long been considered a promising enzyme for catalyzing peptide synthesis [39]. A peptide ligase variant of trypsin, trypsiligase, that retains the native Ser catalytic nucleophile was recently reported [40]. Trypsiligase was produced from trypsin through the introduction of four mutations, resulting in an enzyme with high specificity for the sequence YRH and dependence on Zn^{2+} ions (Figure 2C). Initial cleavage of the substrate occurs between Tyr and Arg. Trypsiligase exhibits substrate-activated catalysis in which it converts from a zymogen-like state to an active conformation that favors peptide ligation over hydrolysis only in the presence of the tripeptide motif YRH and Zn^{+2} ions [40]. Site specific protein tagging at both the N- and C-terminus using trypsiligase has been reported. When used for N-terminal labeling, the trypsiligase variant can recognize activated substrate mimetics such as peptidyl 4-guanidinophenyl esters (OGp), allowing for N-terminal protein tagging with diverse acyl groups (Figure 2C, top) [40]. Trypsiligase-catalyzed C-terminal modification procceds via a transpeptidase reaction between an introduced Y-RH motif and RH-X nucleophilic acyl acceptor peptide, where X represents either a tag or peptide sequence (Figure 2C, bottom). Trypsiligase-mediated bioconjugation has been applied for the site specific introduction of click anchors into antibodies [41,42].

Trypsiligase-catalyzed site-specific protein bioconjugation has the advantage that it is highly specific for the YRH motif, giving it the potential to label a single protein terminus in the context of a complex protein mixture such as a cell lysate or a live cell. However, this application has yet to be reported. Although this specificity provides an advantage for modification of a specific protein, it limits the applicability of trypsiligase for global profiling of protein termini, as the YRH motif is only found in in 0.5% of all known protein sequences. Trypsiligase has the limitations that a large of excess of acyl acceptor substrate is often required, and that the enzyme can catalyze secondary hydrolysis of the ligation product [43]. Future protein engineering efforts on trypsiligase to minimize secondary hydrolysis and to alter its substrate specificity would expand the utility of this enzyme for protein bioconjugation and the study of cell signaling.

Carboxypeptidase Y variants

Carboxypeptidases are a class of proteases that cleave amide bonds at the C-terminal end of their substrate proteins or peptides. In their natural contexts, carboxypeptidases contribute to a variety of processes, including providing nutrition for the cell by cleaving amino acids from extracellular peptides, participating in the general turnover of proteins, and protein maturation [44,45]. Carboxypeptidase Y (CPD-Y), isolated from *Saccharomyces cerevisiae*, is a well-studied carboxypeptidase with a natural P1′ preference for hydrolyzing large non-polar amino acid residues at the C terminus of proteins [46]. Despite its preference for particular P1′ residues, CPD-Y has been used in applications calling for broad-specificity C-terminal cleavage, such as C-terminal protein sequencing [47]. Under certain conditions, CPD-Y also exhibits transpeptidase activity, in which an amino acid or another nucleophile can be ligated to the C terminus of a protein (Figure 2D). This transpeptidation activity has been used to replace the C-terminal amino acids of proteins with a variety of labels, including fluorogenic lysine derivatives, cysteine derivatives, and click chemistry handles [48–51]. More recently, CPD-Y was used to develop a selective C-terminal biotinylation approach termed Profiling Protein C-Termini by Enzymatic Labeling (ProC-TEL) [52,53]. It was previously known that while CPD-Y favors hydrolysis at pH 5–7, the enzyme more readily catalyzes the transpeptidase reaction at pH > 9 [54]. By increasing the reaction pH, CPD-Y activity was shifted to favor incorporation of a biotinylated lysine derivative at the C-terminus of a variety of proteins, allowing for subsequent enrichment of biotinylated C-terminal peptides and their sequencing by LC–MS/MS [52]. Since its development, ProC-TEL has been utilized to study caspase-induced cleavage during apoptosis by monitoring the generation of neo-C termini [55]. Although CPD-Y mutants with increased ligation activity have been developed, CPD-Y has still not been adopted for widespread use due to its bias for particular sequences during ligation activity and poor ligation product yields

Biochemical Society Transactions (2020)
https://doi.org/10.1042/BST20200001

[50]. However, other natural carboxypeptidase variants may be promising protein engineering targets for tool development to produce non-specific C-terminal labeling enzymes [56].

Repurposing natural peptide ligases for bioconjugation

Since the first reports of repurposing proteases for peptide ligation, several classes of enzymes with naturally evolved peptide ligation activity have been discovered (Figure 3). These enzymes are mechanistically similar to proteases in that they typically use a catalytic Ser or Cys nucleophile to form an acyl-enzyme intermediate. This intermediate can then be hydrolyzed or can react with the N-terminal α-amine of a peptide to form a new amide bond. While the hydrolysis reaction predominates under physiological conditions in the proteases, naturally evolved peptide ligases favor peptide bond synthesis. However, in some cases, hydrolytic cleavage prevents the formation of ligation product in high yields. Protein engineering and other techniques have been applied to all of the major peptide ligase classes to optimize their activity for use as tools for protein bioconjugation and the study of cellular signaling.

Sortase variants

Sortases are a class of enzymes that catalyze the covalent attachment of proteins to Gram-positive bacterial cell wall peptidoglycans via a transpeptidation reaction [57]. *Staphylococcus aureus* sortase A is a transpeptidase that initially cleaves a substrate bearing an LPXTG motif between Thr and Gly to generate a Cys-linked thioacyl-enzyme intermediate. This intermediate is then intercepted by a pentaglycine peptide attached to the peptidoglycan precursor lipid II, forming a new peptide bond [58]. Because the key substrate recognition elements are the LPXTG and pentaglycine peptides, sortase A has been applied broadly for bioconjugation to proteins and peptides bearing one of these motifs (Figure 3A). To expand the applications of sortase A, protein engineering approaches have been applied to improve its catalytic activity and to modify substrate motif recognition. Yeast display and *in vitro* compartmentalization strategies were independently used develop a sortase A variants with a >100-fold increases in catalytic activity compared with wild-type sortase A [59,60]. The efficiency of sortase-mediated ligation has been improved by the development of proximity-based sortase-mediated ligation (PBSL) [61]. In PBSL, the target protein is co-immobilized with sortase A through a linker that is cleaved during the ligation reaction, increasing the efficiency of ligation to >95%, while leaving no trace of the immobilization tag. A yeast display approach was used for directed evolution of sortase A variants with the modified substrate recognition motifs LAXTG and LPXSG [62]. Characterization of sortase A homologs from

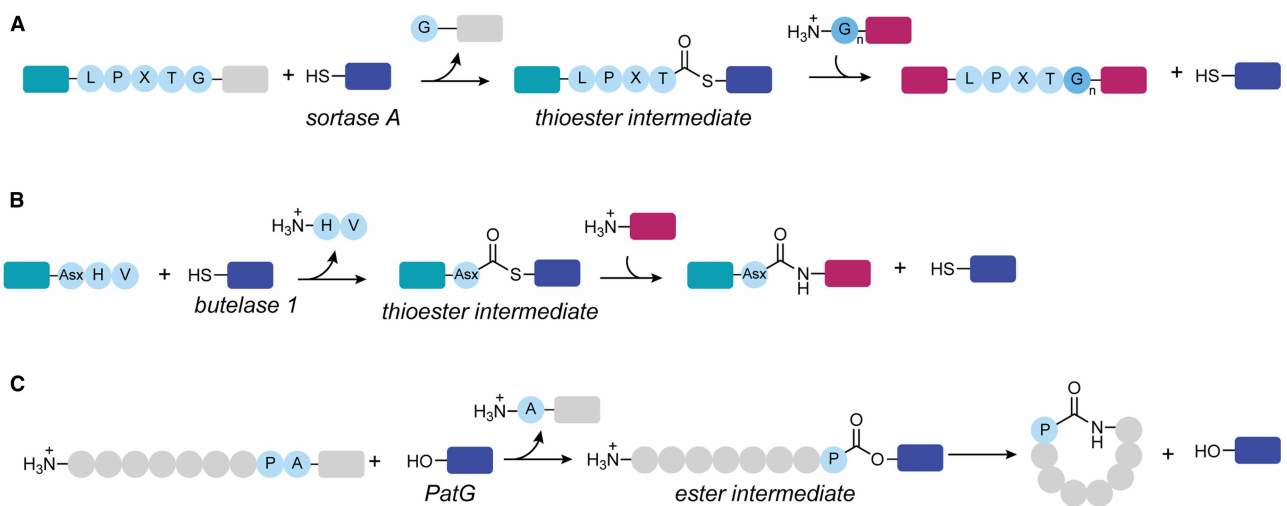

Figure 3. Reactions catalyzed by natural peptide ligases.

(**A**) Sortase A-catalyzed peptide ligation. Sortase A catalyzes ligation of an N-terminal peptide or protein bearing an LPXTG motif at its C terminus and a C-terminal peptide or protein bearing a (poly)Gly at its N terminus. (**B**) Butelase 1-catalyzed peptide ligation. Butelase 1 catalyzes ligation of a peptide or protein with an N/D-HV motif at its C terminus to the N terminus of a peptide or protein. (**C**) PatG catalyzed peptide cyclization. PatG cleaves after Pro, thiazole, thiazoline, or oxazoline in its recognition sequence and catalyzes the cyclization of the resulting linear peptide.

other bacterial species has also led to the identification of variants that modify sequences that are unreactive with the *S. aureus* enzyme. Other properties of sortase, including its stability and dependence on Ca^{2+}, have also been optimized through protein engineering [63,64].

Sortase-mediated ligation, also known as sortagging, has been used for a wide variety of bioconjugation applications, including modification of a protein of interest with fluorescent markers [65,66], synthesis of antibody conjugates [67–69], and introduction of chemical handles for click chemistry [70–72]. Sortase A has also been utilized for synthesis and modification of peptides, including spider venom peptides [73], cone snail toxins [74], cyclotides [75], relaxin peptides [76], and peptides labeled with biotin, PEG, and lipids [77]. Because both the N- and C-terminal substrates of sortase can be genetically encoded, and because its high sequence specificity ensures modification of only proteins with the cognate motif, sortase has been widely applied for *in vivo* protein bioconjugation. Sortase-mediated tagging has been used *in vivo* to tag proteins in *E. coli* [78]; to modify the surface of red blood cells with single-domain antibodies [79]; and for a variety of other tagging applications in yeast, *Toxoplasma gondii*, and human cell lines [80–82]. Importantly, sortase is able to modify naturally exposed surface glycine residues, eliminating the need to genetically engineer a pentaglycine tag for cell surface modification. Sortase A labeling has also been applied on an organismal level in *Caenorhabditis elegans,* demonstrating its potential for protein tagging in complex eukaryotic systems [83]. The broad applications of sortase-mediated labeling have previously been reviewed in more detail [84].

Sortase A-mediated ligation has the advantages of demonstrated applicability in living systems and high specificity for its cognate substrates. However, sortase specificity also presents a limitation for application of this enzyme in the context of global profiling of protein termini. Sortase has proven amenable to protein engineering to alter its properties, and future efforts in this direction may further enhance its activity and substrate scope, expanding the applicability of this enzyme for protein terminal modification.

Peptide asparaginyl ligases

Butelase 1, an Asp/Asn-specific peptide asparaginyl ligase (PAL) that is homologous to asparaginyl endopeptidases (AEPs), was recently isolated from the cyclic peptide producing plant *Clitoria ternatea* (Figure 3B) [85]. In its native context, butelase 1 catalyzes the cyclization of linear peptides to produce cyclotides that are necessary for host defense from herbivorous pests [86,87]. Butelase 1 recognizes peptides harboring a C-terminal D/N-H-V motif, cleaving between Asp/Asn and His to generate a thioacyl-enzyme intermediate. This intermediate can then react with the N-terminal α-amine of the same peptide to form a cyclized product. The N-terminal ligation partner is limited only by the exclusion of Pro and acidic residues at P1′, and the necessity of a hydrophobic residue at P2′. However, if Gly is present at P1′, any P2′ residue is accepted for macrocyclization [85]. This broad specificity has been exploited to allow butelase 1 to catalyze peptide ligation rather than cyclization by adding an excess of nucleophilic peptide to intercept the thioacyl-enzyme intermediate *in trans*. Applications of butelase 1 were initially limited due to the reversible nature of the ligation reaction [88]. In butelase 1-catalyzed ligation reactions, the HV dipeptide acts a competitive nucleophile with the desired substrate, requiring an excess of the desired substrate in solution. The use of thiodepsipeptides, in which the thiol leaving group is not accepted as a nucleophilic substrate by the enzyme, renders the reaction irreversible and has helped to overcome this challenge [89].

Butelase is a promising tool for peptide ligation due to its high peptide cyclization rates, with much higher efficiency than sortase [90]. Applications of butelase 1 to date have included cyclization of large bacteriocins up to 70 residues in length [91]; site-specific N-terminal labeling of proteins [89]; preparation of protein thioesters in conjunction with sortase [88,92,93]; and modification of the bacterial cell surface to study interactions with human cells [94]. However, applications of butelase 1 have been limited by its high sequence specificity and the inability to express the enzyme recombinantly, which limits the applicability of protein engineering approaches to alter its properties. Butelase 1 homologs from *Oldenlandia affinis* [95,96] and cyclotide-producing plants of the Violaceae family [97] that are amenable to recombinant expression have been identified, among others [98–100]. Recently, the *Oldenlandia affinis* asparaginyl endopeptidase has been successfully utilized for both N- and C-terminal protein labeling [101], demonstrating that these enzymes are promising targets for future protein engineering efforts.

Subtilisin-like peptide ligases

Naturally occurring subtilisin-like enzymes that catalyze peptide ligation were recently discovered, almost two decades after the engineered subtiligase variant of subtilisin was reported (Figure 3C). This enzyme family, the founding member of which is PatG, catalyzes N-to-C peptide macrocyclization in the context of biosynthesis of

Biochemical Society Transactions (2020)
https://doi.org/10.1042/BST20200001

cyanobactins and cyclotides, two classes of ribosomally synthesized and post-translationally modified peptide (RiPP) natural products [102,103]. PatG catalyzes proteolytic cleavage of the C-terminal motif PAYDGE between Pro and Ala, releasing the AYDGE peptide and forming an acyl-enzyme intermediate. This intermediate can then be intercepted by the N-terminal α-amine of the enzyme-bound peptide to form a cyclic product. Unlike the engineered peptide ligase subtiligase, PatG retains the Asp-His-Ser catalytic triad found in subtilisin proteases. A helix-loop-helix insertion in PatG compared with subtilisin proteases is believed to shield the acyl-enzyme intermediate from water, preventing hydrolysis and favoring the ligation reaction. The removal of this helix-loop-helix domain effectively silences macrocyclization activity in favor of proteolytic cleavage activity [104]. Unlike subtilisin, PatG binds its target peptide in a bent conformation, creating a requirement for a cyclized residue at the P1 position. This limits the substrate scope of PatG to peptides containing a P1 thiazoline or oxazoline derived from heterocyclization of Cys, Ser, or Thr. PatG also accepts Pro at P1 in synthetic substrates, but the catalytic efficiency of cyclization is much lower than with its native substrates. PatG has been applied for the cyclization of non-natural substrates between 3 and 22 residues in length [105,106]. A PatG homolog, PagG, that utilizes P1 Pro-containing peptides as its native substrates, was recently characterized and used to catalyze macrocyclization of more than 100 different peptides *in vitro* [107]. Therefore, aside from the requirement for a cyclic P1 residues, PatG-like enzymes are capable of synthesis of diverse macrocycles.

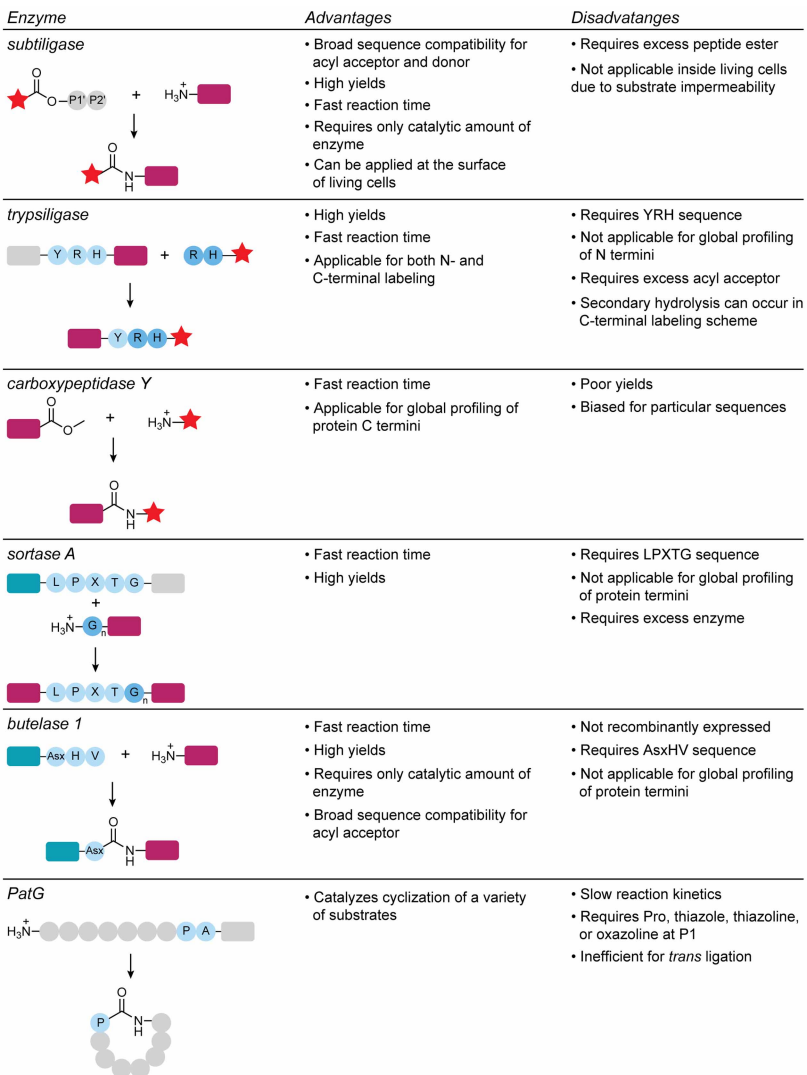

Figure 4. Comparison of engineered and naturally evolved peptide ligase enzymes.

PatG enzymes have been demonstrated as effective catalysts for cyclic peptides, an important class of drug candidates. However, PatG and its homologs exhibit unfavorably slow reaction kinetics compared with other peptide ligases. Additionally, PatG enzymes are inefficient catalysts of *trans* peptide ligation for synthesis of linear peptides. Future protein engineering efforts, in combination with genome mining for new homologs, may lead to development of enzymes with wider substrate tolerance and faster kinetics. Discovery of PatG variants with improved properties in these areas would likely lead to their wider adoption as tools for biology.

Future challenges and opportunities

The application of peptide ligases for site-specific modification of peptide and protein N and C termini has had a strong impact on a wide array of research areas in chemistry and biology. The current toolbox of natural and engineered peptide ligases enables site-specific modification of protein termini without the need to introduce an engineered epitope, tagging of specific sequences through introduction of a recognition motif, and global modification of N and C termini for chemoproteomic studies (Figure 4). However, several limitations to these tools remain to be overcome through protein engineering. While a number of terminal modification enzymes with broad N-terminal specificity have been characterized, few enzymes with broad C-terminal specificity have been identified, and none of these enzymes are able to modify free, unactivated C termini. These restrictions could be addressed through further genome mining for enzymes that possess these activities, or through protein engineering platforms such as phage [15] or yeast display [59], microfluidics [108] or chip-based [109] approaches, or phage-assisted continuous evolution [110]. With the exception of sortase, no peptide ligase enzymes have yet been developed that can function inside living cells, precluding their use to obtain subcellular spatial resolution of protein termini. This limitation arises primarily as a result of the need to use substrates that are not cell permeable or non-physiological reaction conditions. Development of new substrates and new peptide ligase enzymes, as well as further optimization of existing enzymes, could meet this challenge and would enable the study of cellular signaling pathways with subcellular resolution. New developments in protein engineering are therefore likely to drive new applications of engineered peptide ligases and to propel their impact on all areas of biological research.

Perspectives

- Importance of the field: Engineered peptide ligases catalyze site-specific modification of protein termini, enabling introduction of chemical probes and payloads for drug discovery, advanced imaging, protein synthesis, and chemoproteomic profiling, among many other applications.
- Summary of current thinking: The current toolbox of peptide ligases consists of engineered proteases and naturally occurring peptide ligases. These enzymes facilitate both site-specific modification of naturally occurring protein sequences as well as sequence-specific modification of proteins with engineered epitopes.
- Future directions: Improvements in enzyme specificity and activity will provide the field with more useful tools suitable for new applications. In particular, the development of broad specificity C-terminal modification enzymes, enzymes that can modify free C termini, and peptide ligases that can function as tools inside living cells would expand the utility of peptide ligases for site-specific bioconjugation and the study of cellular signaling.

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

A.M.W. acknowledges financial support from a Career Award at the Scientific Interface from the Burroughs Wellcome Fund [1017065] and a Steenbock Career Award from the University of Wisconsin.

Author Contributions

C.L.F. and A.M.W. wrote the manuscript and made the figures. Both authors discussed, edited, and approved the final manuscript.

Acknowledgements

We thank K. Radziwon, A. Amiridis, S. Coyle, and D. Sashital for helpful discussions and comments on the manuscript.

Abbreviations

AEP, asparaginyl endopeptidase; CPD-Y, carboxypeptidase Y; PAL, peptide asparaginyl ligase.

References

1 Abrahmsen, L., Tom, J., Burnier, J., Butcher, K.A., Kossiakoff, A. and Wells, J.A. (1991) Engineering subtilisin and its substrates for efficient ligation of peptide bonds in aqueous solution. *Biochemistry* **30**, 4151–4159 https://doi.org/10.1021/bi00231a007

2 Gilmore, J.M., Scheck, R.A., Esser-Kahn, A.P., Joshi, N.S. and Francis, M.B. (2006) N-terminal protein modification through a biomimetic transamination reaction. *Angew. Chem. Int. Ed.* **45**, 5307–5311 https://doi.org/10.1002/anie.200600368

3 Kleifeld, O., Doucet, A., auf dem Keller, U., Prudova, A., Schilling, O., Kainthan, R.K. et al. (2010) Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products. *Nat. Biotechnol.* **28**, 281–288 https://doi.org/10.1038/nbt.1611

4 Schilling, O., Barré, O., Huesgen, P.F. and Overall, C.M. (2010) Proteome-wide analysis of protein carboxy termini: C terminomics. *Nat. Methods* **7**, 508–511 https://doi.org/10.1038/nmeth.1467

5 Rosen, C.B. and Francis, M.B. (2017) Targeting the N terminus for site-selective protein modification. *Nat. Chem. Biol.* **13**, 697–705 https://doi.org/10.1038/nchembio.2416

6 Tanco, S., Gevaert, K. and Van Damme, P. (2015) C-terminomics: targeted analysis of natural and posttranslationally modified protein and peptide C-termini. *Proteomics* **15**, 903–914 https://doi.org/10.1002/pmic.201400301

7 Rheinnecker, M., Eder, J., Pandey, P.S. and Fersht, A.R. (1994) Variants of subtilisin BPN' with altered specificity profiles. *Biochemistry* **33**, 221–225 https://doi.org/10.1021/bi00167a029

8 Kapust, R.B., Tözsér, J., Copeland, T.D. and Waugh, D.S. (2002) The P1′ specificity of tobacco etch virus protease. *Biochem. Biophys. Res. Commun.* **294**, 949–955 https://doi.org/10.1016/S0006-291X(02)00574-0

9 Martin, R.B. (1998) Free energies and equilibria of peptide bond hydrolysis and formation. *Biopolymers* **45**, 351–353 https://doi.org/10.1002/(SICI)1097-0282(19980415)45:5<351::AID-BIP3>3.0.CO;2-K

10 Morihara, K. (1987) Using proteases in peptide synthesis. *Trends Biotechnol.* **5**, 164–170 https://doi.org/10.1016/0167-7799(87)90090-4

11 Barbas, C.F., Matos, J.R., West, J.B. and Wong, C.H. (1988) A search for peptide ligase: cosolvent-mediated conversion of proteases to esterases for irreversible synthesis of peptides. *J. Am. Chem. Soc.* **110**, 5162–5166 https://doi.org/10.1021/ja00223a042

12 Wells, J.A., Ferrari, E., Henner, D.J., Estell, D.A. and Chen, E.Y. (1983) Cloning, sequencing, and secretion of *Bacillus amyloliquefaciens* subtilisin in *Bacillus subtilis*. *Nucleic Acids Res.* **11**, 7911–7925 https://doi.org/10.1093/nar/11.22.7911

13 Nakatsuka, T., Sasaki, T. and Kaiser, E.T. (1987) Peptide segment synthesis catalyzed by the semisynthetic enzyme thiolsubtilisin. *J. Am. Chem. Soc.* **109**, 3808–3810 https://doi.org/10.1021/ja00246a064

14 Chang, T.K., Jackson, D.Y., Burnier, J.P. and Wells, J.A. (1994) Subtiligase: a tool for semisynthesis of proteins. *Proc. Natl Acad. Sci. U.S.A.* **91**, 12544–12548 https://doi.org/10.1073/pnas.91.26.12544

15 Atwell, S. and Wells, J.A. (1999) Selection for improved subtiligases by phage display. *Proc. Natl Acad. Sci. U.S.A.* **96**, 9497–9502 https://doi.org/10.1073/pnas.96.17.9497

16 Toplak, A., Nuijens, T., Quaedflieg, P.J.L.M., Wu, B. and Janssen, D.B. (2016) Peptiligase, an enzyme for efficient chemoenzymatic peptide synthesis and cyclization in water. *Adv. Synth. Catal.* **358**, 2140–2147 https://doi.org/10.1002/adsc.201600017

17 Groen, H., Meldal, M. and Breddam, K. (1992) Extensive comparison of the substrate preferences of two subtilisins as determined with peptide substrates which are based on the principle of intramolecular quenching. *Biochemistry* **31**, 6011–6018 https://doi.org/10.1021/bi00141a008

18 Nuijens, T., Toplak, A., Quaedflieg, P.J.L.M., Drenth, J., Wu, B. and Janssen, D.B. (2016) Engineering a diverse ligase toolbox for peptide segment condensation. *Adv. Synth. Catal.* **358**, 4041–4048 https://doi.org/10.1002/adsc.201600774

19 Weeks, A.M. and Wells, J.A. (2018) Engineering peptide ligase specificity by proteomic identification of ligation sites. *Nat. Chem. Biol.* **14**, 50–57 https://doi.org/10.1038/nchembio.2521

20 Schechter, I. and Berger, A. (1967) On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **27**, 157–162 https://doi.org/10.1016/S0006-291X(67)80055-X

21 Wells, J.A., Cunningham, B.C., Graycar, T.P. and Estell, D.A. (1987) Recruitment of substrate-specificity properties from one enzyme into a related one by protein engineering. *Proc. Natl Acad. Sci. U.S.A.* **84**, 5167–5171 https://doi.org/10.1073/pnas.84.15.5167

22 Estell, D.A., Graycar, T.P. and Wells, J.A. (1985) Engineering an enzyme by site-directed mutagenesis to be resistant to chemical oxidation. *J. Biol. Chem.* **260**, 6518–6521 PMID:3922976

23 Ballinger, M.D., Tom, J. and Wells, J.A. (1995) Designing subtilisin BPN′ to cleave substrates containing dibasic residues. *Biochemistry* **34**, 13312–13319 https://doi.org/10.1021/bi00041a006

24 Schmidt, M., Toplak, A., Rozeboom, H.J., Wijma, H.J., Quaedflieg, P.J.L.M., van Maarseveen, J.H. et al. (2018) Design of a substrate-tailored peptiligase variant for the efficient synthesis of thymosin-α 1. *Org. Biomol. Chem.* **16**, 609–618 https://doi.org/10.1039/c7ob02812a

25 Schmidt, M., Huang, Y.-H., Texeira de Oliveira, E.F., Toplak, A., Wijma, H.J., Janssen, D.B. et al. (2019) Efficient enzymatic cyclization of disulfide-rich peptides by using peptide ligases. *ChemBioChem* **20**, 1524–1529 https://doi.org/10.1002/cbic.201900033

26 Jackson, D.Y., Burnier, J.P. and Wells, J.A. (1995) Enzymic cyclization of linear peptide esters using subtiligase. *J. Am. Chem. Soc.* **117**, 819–820 https://doi.org/10.1021/ja00107a027
27 Tan, X., Yang, R. and Liu, C.F. (2018) Facilitating subtiligase-catalyzed peptide ligation reactions by using peptide thioester substrates. *Org. Lett.* **20**, 6691–6694 https://doi.org/10.1021/acs.orglett.8b02747
28 Henager, S.H., Henriquez, S., Dempsey, D.R. and Cole, P.A. (2019) Analysis of site-specific phosphorylation of PTEN by using enzyme-catalyzed expressed protein ligation. *ChemBioChem* **21**, 64–68 https://doi.org/10.1002/cbic.201900316
29 Mahrus, S., Trinidad, J.C., Barkan, D.T., Sali, A., Burlingame, A.L. and Wells, J.A. (2008) Global sequencing of proteolytic cleavage sites in apoptosis by specific labeling of protein N termini. *Cell* **134**, 866–876 https://doi.org/10.1016/j.cell.2008.08.012
30 Shimbo, K., Hsu, G.W., Nguyen, H., Mahrus, S., Trinidad, J.C., Burlingame, A.L. et al. (2012) Quantitative profiling of caspase-cleaved substrates reveals different drug-induced and cell-type patterns in apoptosis. *Proc. Natl Acad. Sci. U.S.A.* **109**, 12432–12437 https://doi.org/10.1073/pnas.1208616109
31 Hill, M.E., Kumar, A., Wells, J.A., Hobman, T.C., Julien, O. and Hardy, J.A. (2018) The unique cofactor region of zika virus NS2B-NS3 protease facilitates cleavage of key host proteins. *ACS Chem. Biol.* **13**, 2398–2405 https://doi.org/10.1021/acschembio.8b00508
32 Roncase, E.J., Moon, C., Chatterjee, S., González-Páez, G.E., Craik, C.S., O'Donoghue, A.J. et al. (2017) Substrate profiling and high resolution co-complex crystal structure of a secreted C11 protease conserved across commensal bacteria. *ACS Chem. Biol.* **12**, 1556–1565 https://doi.org/10.1021/acschembio.7b00143
33 Agard, N.J., Mahrus, S., Trinidad, J.C., Lynn, A., Burlingame, A.L. and Wells, J.A. (2012) Global kinetic analysis of proteolysis via quantitative targeted proteomics. *Proc. Natl Acad. Sci. U.S.A.* **109**, 1913–1918 https://doi.org/10.1073/pnas.1117158109
34 Calvo, S.E., Julien, O., Clauser, K.R., Shen, H., Kamer, K.J., Wells, J.A. et al. (2017) Comparative analysis of mitochondrial N-termini from mouse, human, and yeast. *Mol. Cell Proteomics* **16**, 512–523 https://doi.org/10.1074/mcp.M116.063818
35 Julien, O., Zhuang, M., Wiita, A.P., O'Donoghue, A.J., Knudsen, G.M., Craik, C.S. et al. (2016) Quantitative MS-based enzymology of caspases reveals distinct protein substrate specificities, hierarchies, and cellular roles. *Proc. Natl Acad. Sci. U.S.A.* **113**, E2001–E2010 https://doi.org/10.1073/pnas.1524900113
36 Agard, N.J., Maltby, D. and Wells, J.A. (2010) Inflammatory stimuli regulate caspase substrate profiles. *Mol. Cell Proteomics* **9**, 880–893 https://doi.org/10.1074/mcp.M900528-MCP200
37 Weeks, A.M. and Wells, J.A. (2020) Subtiligase-catalyzed peptide ligation. *Chem. Rev.* **120**, 3127–3160 https://doi.org/10.1021/acs.chemrev.9b00372
38 Weeks, A.M. and Wells, J.A. (2020) N-Terminal modification of proteins with subtiligase specificity variants. *Curr. Protoc. Chem. Biol.* **12**, e79 https://doi.org/10.1002/cpch.79
39 Riechmann, L. and Kasche, V. (1985) Peptide synthesis catalyzed by the serine proteinases chymotrypsin and trypsin. *Biochim. Biophys. Acta* **830**, 164–172 https://doi.org/10.1016/0167-4838(85)90024-x
40 Liebscher, S., Schöpfel, M., Aumüller, T., Sharkhuukhen, A., Pech, A., Höss, E. et al. (2014) N-terminal protein modification by substrate-activated reverse proteolysis. *Angew. Chem. Int. Ed. Engl.* **53**, 3024–3028 https://doi.org/10.1002/anie.201307736
41 Meyer, C., Liebscher, S. and Bordusa, F. (2016) Selective coupling of click anchors to proteins via trypsiligase. *Bioconjug. Chem.* **27**, 47–53 https://doi.org/10.1021/acs.bioconjchem.5b00618
42 Liebscher, S., Kornberger, P., Fink, G., Trost-Gross, E.-M., Höss, E., Skerra, A. et al. (2014) Derivatization of antibody Fab fragments: A designer enzyme for native protein modification. *ChemBioChem* **15**, 1096–1100 https://doi.org/10.1002/cbic.201400059
43 Liebscher, S. and Bordusa, F. (2019) Trypsiligase-catalyzed peptide and protein ligation. *Methods Mol. Biol.* **2012**, 111–133 https://doi.org/10.1007/978-1-4939-9546-2_7
44 Breddam, K. (1986) Serine carboxypeptidases. A review. *Carlsberg Res. Commun.* **51**, 83–128 https://doi.org/10.1007/BF02907561
45 Ambler, R.P. (1972) Enzymatic hydrolysis with carboxypeptidases. *Methods Enzymol.* **25**, 143–154 https://doi.org/10.1016/S0076-6879(72)25012-1
46 Remington, S.J. and Breddam, K. (1994) Carboxypeptidases C and D. *Methods Enzymol.* **244**, 231–248 https://doi.org/10.1016/0076-6879(94)44020-4
47 Patterson, D.H., Tarr, G.E., Regnier, F.E. and Martin, S.A. (1995) C-terminal ladder sequencing via matrix-assisted laser desorption mass spectrometry coupled with carboxypeptidase Y time-dependent and concentration-dependent digestions. *Anal. Chem.* **67**, 3971–3978 https://doi.org/10.1021/ac00117a024
48 Berne, P.F., Schmitter, J. and Blanquet, S. (1990) Peptide and protein carboxyl-terminal labeling through carboxypeptidase Y-catalyzed transpeptidation. *J. Biol. Chem.* **265**, 19551–19559 PMID:2246241
49 Lin, S. and Lowe, C.R. (2000) C-Terminal labeling of immunoglobulin G with a cysteine derivative by carboxypeptidase Y catalyzed transpeptidation. *Anal. Biochem.* **285**, 127–134 https://doi.org/10.1006/abio.2000.4723
50 Stennicke, H.R., Olesen, K., Sørensen, S.B. and Breddam, K. (1997) C-terminal incorporation of fluorogenic and affinity labels using wild-type and mutagenized carboxypeptidase Y. *Anal. Biochem.* **248**, 141–148 https://doi.org/10.1006/abio.1997.9998
51 Peschke, B., Zundel, M., Bak, S., Clausen, T.R., Blume, N., Pedersen, A. et al. (2007) C-terminally PEGylated hGH-derivatives. *Bioorg. Med. Chem.* **15**, 4382–4395 https://doi.org/10.1016/j.bmc.2007.04.037
52 Xu, G., Shin, S.B. and Jaffrey, S.R. (2011) Chemoenzymatic labeling of protein C-termini for positive selection of C-terminal peptides. *ACS Chem. Biol.* **6**, 1015–1020 https://doi.org/10.1021/cb200164h
53 Duan, W. and Xu, G. (2017) ProC-TEL: Profiling of protein C-termini by enzymatic labeling. *Methods Mol. Biol.* **1574**, 135–144 https://doi.org/10.1007/978-1-4939-6850-3_10
54 Hayashi, R., Bai, Y. and Hata, T. (1975) Kinetic studies of carboxypeptidase Y. I. Kinetic parameters for the hydrolysis of synthetic substrates. *J. Biochem.* **77**, 69–79 https://doi.org/10.1093/oxfordjournals.jbchem.a130720
55 Duan, W., Chen, S., Zhang, Y., Li, D., Wang, R., Chen, S. et al. (2016) Protein C-terminal enzymatic labeling identifies novel caspase cleavages during the apoptosis of multiple myeloma cells induced by kinase inhibition. *Proteomics* **16**, 60–69 https://doi.org/10.1002/pmic.201500356
56 Tanco, S., Tort, O., Demol, H., Aviles, F.X., Gevaert, K., Van Damme, P. et al. (2015) C-terminomics screen for natural substrates of cytosolic carboxypeptidase 1 reveals processing of acidic protein C termini. *Mol. Cell Proteomics* **14**, 177–190 https://doi.org/10.1074/mcp.M114.040360
57 Marraffini, L.A., Dedent, A.C. and Schneewind, O. (2006) Sortases and the art of anchoring proteins to the envelopes of gram-positive bacteria. *Microbiol. Mol. Biol. Rev.* **70**, 192–221 https://doi.org/10.1128/MMBR.70.1.192-221.2006

58 Mazmanian, S.K., Liu, G., Ton-That, H. and Schneewind, O. (1999) *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. *Science* **285**, 760–763 https://doi.org/10.1126/science.285.5428.760
59 Chen, I., Dorr, B.M. and Liu, D.R. (2011) A general strategy for the evolution of bond-forming enzymes using yeast display. *Proc. Natl Acad. Sci. U.S.A.* **108**, 11399–11404 https://doi.org/10.1073/pnas.1101046108
60 Gianella, P., Snapp, E.L. and Levy, M. (2016) An in vitro compartmentalization-based method for the selection of bond-forming enzymes from large libraries. *Biotechnol. Bioeng.* **113**, 1647–1657 https://doi.org/10.1002/bit.25939
61 Wang, H.H., Altun, B., Nwe, K. and Tsourkas, A. (2017) Proximity-based sortase-mediated ligation. *Angew. Chem. Int. Ed. Engl.* **56**, 5349–5352 https://doi.org/10.1002/anie.201701419
62 Dorr, B.M., Ham, H.O., An, C., Chaikof, E.L. and Liu, D.R. (2014) Reprogramming the specificity of sortase enzymes. *Proc. Natl Acad. Sci. U.S.A.* **111**, 13343–13348 https://doi.org/10.1073/pnas.1411179111
63 Jeong, H.J., Abhiraman, G.C., Story, C.M., Ingram, J.R. and Dougan, S.K. (2017) Generation of Ca^{2+}-independent sortase A mutants with enhanced activity for protein and cell surface labeling. *PLoS ONE* **12**, e0189068 https://doi.org/10.1371/journal.pone.0189068
64 Pelay-Gimeno, M., Bange, T., Hennig, S. and Grossmann, T.N. (2018) In situ cyclization of native proteins: structure-based design of a bicyclic enzyme. *Angew. Chem. Int. Ed. Engl.* **57**, 11164–11170 https://doi.org/10.1002/anie.201804506
65 Matsumoto, T., Tanaka, T. and Kondo, A. (2012) Sortase A-catalyzed site-specific coimmobilization on microparticles via streptavidin. *Langmuir* **28**, 3553–3557 https://doi.org/10.1021/la2047933
66 Antos, J.M., Chew, G.L., Guimaraes, C.P., Yoder, N.C., Grotenbreg, G.M., Popp, M.W. et al. (2009) Site-specific N- and C-terminal labeling of a single polypeptide using sortases of different specificity. *J. Am. Chem. Soc.* **131**, 10800–10801 https://doi.org/10.1021/ja902681k
67 Paterson, B.M., Alt, K., Jeffery, C.M., Price, R.I., Jagdale, S., Rigby, S. et al. (2014) Enzyme-mediated site-specific bioconjugation of metal complexes to proteins: sortase-mediated coupling of copper-64 to a single-chain antibody. *Angew. Chem. Int. Ed. Engl.* **53**, 6115–6119 https://doi.org/10.1002/anie.201402613
68 Beerli, R.R., Hell, T., Merkel, A.S. and Grawunder, U. (2015) Sortase enzyme-mediated generation of site-specifically conjugated antibody drug conjugates with high in vitro and in vivo potency. *PLoS ONE* **10**, e0131177 https://doi.org/10.1371/journal.pone.0131177
69 Massa, S., Vikani, N., Betti, C., Ballet, S., Vanderhaegen, S., Steyaert, J. et al. (2016) Sortase A-mediated site-specific labeling of camelid single-domain antibody-fragments: a versatile strategy for multiple molecular imaging modalities. *Contrast. Media. Mol. Imaging* **11**, 328–339 https://doi.org/10.1002/cmmi.1696
70 Gupta, K., Singh, S., Gupta, K., Khan, N., Sehgal, D., Haridas, V. et al. (2012) A bioorthogonal chemoenzymatic strategy for defined protein dendrimer assembly. *ChemBioChem* **13**, 2489–2494 https://doi.org/10.1002/cbic.201200559
71 Wang, L., Jiang, R., Wang, L., Liu, Y. and Sun, X.L. (2016) Preparation of chain-end clickable recombinant protein and its bio-orthogonal modification. *Bioorg. Chem.* **65**, 159–166 https://doi.org/10.1016/j.bioorg.2016.02.010
72 Witte, M.D., Theile, C.S., Wu, T., Guimaraes, C.P., Blom, A.E. and Ploegh, H.L. (2013) Production of unnaturally linked chimeric proteins using a combination of sortase-catalyzed transpeptidation and click chemistry. *Nat. Protoc.* **8**, 1808–1819 https://doi.org/10.1038/nprot.2013.103
73 Agwa, A.J., Blomster, L.V., Craik, D.J., King, G.F. and Schroeder, C.I. (2018) Efficient enzymatic ligation of inhibitor cystine knot spider venom peptides: Using sortase A to form double-knottins that probe voltage-gated sodium channel NaV1.7. *Bioconjug. Chem.* **29**, 3309–3319 https://doi.org/10.1021/acs.bioconjchem.8b00505
74 Kwon, S., Bosmans, F., Kaas, Q., Cheneval, O., Conibear, A.C., Rosengren, K.J. et al. (2016) Efficient enzymatic cyclization of an inhibitory cystine knot-containing peptide. *Biotechnol. Bioeng.* **113**, 2202–2212 https://doi.org/10.1002/bit.25993
75 Stanger, K., Maurer, T., Kaluarachchi, H., Coons, M., Franke, Y. and Hannoush, R.N. (2014) Backbone cyclization of a recombinant cystine-knot peptide by engineered Sortase A. *FEBS Lett.* **588**, 4487–4496 https://doi.org/10.1016/j.febslet.2014.10.020
76 Wang, J.H., Hu, M.J., Zhang, L., Shao, X.X., Lv, C.H., Liu, Y.L. et al. (2018) Exploring receptor selectivity of the chimeric relaxin family peptide R3/I5 by incorporating unnatural amino acids. *Biochimie* **154**, 77–85 https://doi.org/10.1016/j.biochi.2018.08.003
77 Cheng, X., Zhu, T., Hong, H., Zhou, Z. and Wu, Z. (2017) Sortase A-mediated on-resin peptide cleavage and in situ ligation: an efficient one-pot strategy for the synthesis of functional peptides and proteins. *Org. Chem. Front.* **4**, 2058–2062 https://doi.org/10.1039/c7qo00481h
78 Hirakawa, H., Ishikawa, S. and Nagamune, T. (2015) Ca^{2+}-independent sortase-A exhibits high selective protein ligation activity in the cytoplasm of *Escherichia coli*. *Biotechnol. J.* **10**, 1487–1492 https://doi.org/10.1002/biot.201500012
79 Shi, J., Kundrat, L., Pishesha, N., Bilate, A., Theile, C., Maruyama, T. et al. (2014) Engineered red blood cells as carriers for systemic delivery of a wide array of functional probes. *Proc. Natl Acad. Sci. U.S.A.* **111**, 10131–10136 https://doi.org/10.1073/pnas.1409861111
80 Swee, L.K., Lourido, S., Bell, G.W., Ingram, J.R. and Ploegh, H.L. (2015) One-step enzymatic modification of the cell surface redirects cellular cytotoxicity and parasite tropism. *ACS Chem. Biol.* **10**, 460–465 https://doi.org/10.1021/cb500462t
81 Strijbis, K., Spooner, E. and Ploegh, H.L. (2012) Protein ligation in living cells using sortase. *Traffic* **13**, 780–789 https://doi.org/10.1111/j.1600-0854.2012.01345.x
82 Shi, H., Shi, Q., Oswald, J.T., Gao, Y., Li, L. and Li, Y. (2018) Site-specific PEGylation of human growth hormone by mutated Sortase A. *Chem. Res. Chinese U* **34**, 428–433 https://doi.org/10.1007/s40242-018-8023-3
83 Wu, Q., Ploegh, H.L. and Truttmann, M.C. (2017) Hepta-mutant *Staphylococcus aureus* sortase A (SrtA7m) as a tool for in vivo protein labeling in *Caenorhabditis elegans*. *ACS Chem. Biol.* **12**, 664–673 https://doi.org/10.1021/acschembio.6b00998
84 Dai, X., Böker, A. and Glebe, U. (2019) Broadening the scope of sortagging. *RSC Adv.* **9**, 4700–4721 https://doi.org/10.1039/c8ra06705h
85 Nguyen, G.K., Wang, S., Qiu, Y., Hemu, X., Lian, Y. and Tam, J.P. (2014) Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. *Nat. Chem. Biol.* **10**, 732–738 https://doi.org/10.1038/nchembio.1586
86 Craik, D.J. (2012) Host-defense activities of cyclotides. *Toxins* **4**, 139–156 https://doi.org/10.3390/toxins4020139
87 Nguyen, G.K., Zhang, S., Nguyen, N.T., Nguyen, P.Q., Chiu, M.S., Hardjojo, A. et al. (2011) Discovery and characterization of novel cyclotides originated from chimeric precursors consisting of albumin-1 chain a and cyclotide domains in the fabaceae family. *J. Biol. Chem.* **286**, 24275–24287 https://doi.org/10.1074/jbc.M111.229922
88 Nguyen, G.K., Qiu, Y., Cao, Y., Hemu, X., Liu, C.F. and Tam, J.P. (2016) Butelase-mediated cyclization and ligation of peptides and proteins. *Nat. Protoc.* **11**, 1977–1988 https://doi.org/10.1038/nprot.2016.118

89 Nguyen, G.K., Cao, Y., Wang, W., Liu, C.F. and Tam, J.P. (2015) Site-specific N-terminal labeling of peptides and proteins using butelase 1 and thiodepsipeptide. *Angew. Chem. Int. Ed. Engl.* **54**, 15694–15698 https://doi.org/10.1002/anie.201506810
90 Nguyen, G.K., Kam, A., Loo, S., Jansson, A.E., Pan, L.X. and Tam, J.P. (2015) Butelase 1: a versatile ligase for peptide and protein macrocyclization. *J. Am. Chem. Soc.* **137**, 15398–15401 https://doi.org/10.1021/jacs.5b11014
91 Hemu, X., Qiu, Y., Nguyen, G.K. and Tam, J.P. (2016) Total synthesis of circular bacteriocins by butelase 1. *J. Am. Chem. Soc.* **138**, 6968–6971 https://doi.org/10.1021/jacs.6b04310
92 Cao, Y., Nguyen, G.K., Chuah, S., Tam, J.P. and Liu, C.F. (2016) Butelase-mediated ligation as an efficient bioconjugation method for the synthesis of peptide dendrimers. *Bioconjug. Chem.* **27**, 2592–2596 https://doi.org/10.1021/acs.bioconjchem.6b00538
93 Harmand, T.J., Bousbaine, D., Chan, A., Zhang, X., Liu, D.R., Tam, J.P. et al. (2018) One-pot dual labeling of IgG 1 and preparation of C-to-C fusion proteins through a combination of sortase A and butelase 1. *Bioconj. Chem.* **29**, 3245–3249 https://doi.org/10.1021/acs.bioconjchem.8b00563
94 Bi, X., Yin, J., Nguyen, G.K.T., Rao, C., Halim, N.B.A., Hemu, X. et al. (2017) Enzymatic engineering of live bacterial cell surfaces using butelase 1. *Angew. Chem. Int. Ed. Engl.* **56**, 7822–7825 https://doi.org/10.1002/anie.201703317
95 Harris, K.S., Durek, T., Kaas, Q., Poth, A.G., Gilding, E.K., Conlan, B.F. et al. (2015) Efficient backbone cyclization of linear peptides by a recombinant asparaginyl endopeptidase. *Nat. Commun.* **6**, 10199 https://doi.org/10.1038/ncomms10199
96 Yang, R., Wong, Y.H., Nguyen, G.K.T., Tam, J.P., Lescar, J. and Wu, B. (2017) Engineering a catalytically efficient recombinant protein ligase. *J. Am. Chem. Soc.* **139**, 5351–5358 https://doi.org/10.1021/jacs.6b12637
97 Hemu, X., El Sahili, A., Hu, S., Wong, K., Chen, Y., Wong, Y.H. et al. (2019) Structural determinants for peptide-bond formation by asparaginyl ligases. *Proc. Natl Acad. Sci. U.S.A.* **116**, 11737–11746 https://doi.org/10.1073/pnas.1818568116
98 Haywood, J., Schmidberger, J.W., James, A.M., Nonis, S.G., Sukhoverkov, K.V., Elias, M. et al. (2018) Structural basis of ribosomal peptide macrocyclization in plants. *eLife* **7**, e32955 https://doi.org/10.7554/eLife.32955
99 Bernath-Levin, K., Nelson, C., Elliott, A.G., Jayasena, A.S., Millar, A.H., Craik, D.J. et al. (2015) Peptide macrocyclization by a bifunctional endoprotease. *Chem. Biol.* **22**, 571–582 https://doi.org/10.1016/j.chembiol.2015.04.010
100 Jackson, M.A., Gilding, E.K., Shafee, T., Harris, K.S., Kaas, Q., Poon, S. et al. (2018) Molecular basis for the production of cyclic peptides by plant asparaginyl endopeptidases. *Nat. Commun.* **9**, 2411 https://doi.org/10.1038/s41467-018-04669-9
101 Rehm, F.B.H., Harmand, T.J., Yap, K., Durek, T., Craik, D.J. and Ploegh, H.L. (2019) Site-specific sequential protein labeling catalyzed by a single recombinant ligase. *J. Am. Chem. Soc.* **141**, 17388–17393 https://doi.org/10.1021/jacs.9b09166
102 Lee, J., McIntosh, J., Hathaway, B.J. and Schmidt, E.W. (2009) Using marine natural products to discover a protease that catalyzes peptide macrocyclization of diverse substrates. *J. Am. Chem. Soc.* **131**, 2122–2124 https://doi.org/10.1021/ja8092168
103 Donia, M.S. and Schmidt, E.W. (2011) Linking chemistry and genetics in the growing cyanobactin natural products family. *J. Chem. Biol.* **18**, 508–519 https://doi.org/10.1016/j.chembiol.2011.01.019
104 Koehnke, J., Bent, A., Houssen, W.E., Zollman, D., Morawitz, F., Shirran, S. et al. (2012) The mechanism of patellamide macrocyclization revealed by the characterization of the PatG macrocyclase domain. *Nat. Struct. Mol. Biol.* **19**, 767–772 https://doi.org/10.1038/nsmb.2340
105 Sardar, D., Lin, Z. and Schmidt, E.W. (2015) Modularity of RiPP enzymes enables designed synthesis of decorated peptides. *Chem. Biol.* **22**, 907–916 https://doi.org/10.1016/j.chembiol.2015.06.014
106 Oueis, E., Nardone, B., Jaspars, M., Westwood, N.J. and Naismith, J.H. (2017) Synthesis of hybrid cyclopeptides through enzymatic macrocyclization. *ChemistryOpen* **6**, 11–14 https://doi.org/10.1002/open.201600134
107 Sarkar, S., Gu, W. and Schmidt, E.W. (2020) Expanding the chemical space of synthetic cyclic peptides using a promiscuous macrocyclase from prenylagaramide biosynthesis. *bioRxiv* https://doi.org/10.1101/2020.03.17.996157
108 Agresti, J.J., Antipov, E., Abate, A.R., Ahn, K., Rowat, A.C., Baret, J.C. et al. (2010) Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. *Proc. Natl Acad. Sci. U.S.A.* **107**, 4004–4009 https://doi.org/10.1073/pnas.0910781107
109 Chen, B., Lim, S., Kannan, A., Alford, S.C., Sunden, F., Herschlag, D. et al. (2016) High-throughput analysis and protein engineering using microcapillary arrays. *Nat. Chem. Biol.* **12**, 76–81 https://doi.org/10.1038/nchembio.1978
110 Esvelt, K.M., Carlson, J.C. and Liu, D.R. (2011) A system for the continuous directed evolution of biomolecules. *Nature* **472**, 499–503 https://doi.org/10.1038/nature09929