

RESEARCH

Open Access



An efficient and easily obtainable butelase variant for chemoenzymatic ligation and modification of peptides and proteins

Avinash Kumar Singh^{1*†}, Anastasiia Antonenko^{1†}, Anna Kocyla¹ and Artur Krężel^{1*}

Abstract

The expanding field of site-specific ligation of proteins and peptides has catalyzed the development of novel methods that enhance molecular modification. Among these methods, enzymatic strategies have emerged as dominant due to their specificity and efficiency in modifying proteins under mild conditions. Asparaginyl endopeptidase is a group of cyclotide-producing cysteine proteases from plants. These plant cysteine proteases, known for their specificity, effectively recognize the tripeptide motif (Asx-Xaa-Yaa) and cleave at the C-terminal side of Asx residues, forming acyl-enzyme intermediates that facilitate transpeptidation. Butelase 1 stands out as the most efficient AEP for protein engineering, yet challenges in its expression and purification limit its accessibility for widespread research and industrial use. To address these challenges, we engineered a new, catalytically efficient variant of Butelase 1, Butelase AY, by mutating the gatekeeping residues Val237Ala and Thr238Tyr within the LAD-1 region. These modifications significantly enhanced the stability and yield of Butelase AY, allowing for successful application in various peptide and protein engineering tasks. Butelase AY was tested on the peptide GLGKY, the globular protein GFP, and the intrinsically disordered protein α -synuclein, effectively labeling them with a fluorescent probe. Notably, Butelase AY maintained its efficiency with substrates containing unnatural amino acids, making it a promising candidate for biorthogonal applications. Importantly, the mutations did not compromise the enzyme's specificity, as it continued to process model peptides and native protein substrates with N-term NHV recognition motifs effectively. In conclusion, Butelase AY presents a novel recombinant tool for diverse protein labeling and modifications, particularly in biorthogonal strategies. This innovation has the potential to expand applications in biotechnology and therapeutic development, ultimately revolutionizing protein engineering and its utility in synthetic biology.

Keywords Asparaginyl endopeptidase, Butelase, Gatekeeping residues, Protein engineering, Protein ligation, Unnatural amino acids

[†]Avinash Kumar Singh and Anastasiia Antonenko contributed equally to this work.

*Correspondence:
Avinash Kumar Singh
biotech.avi@gmail.com
Artur Krężel
artur.krezel@uwr.edu.pl

¹Present address: Department of Chemical Biology, Faculty of Biotechnology, University of Wrocław, Joliot-Curie 14a, Wrocław 50-383, Poland



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Introduction

The growing interest in site-specific ligation of proteins and peptides in both industry and academic research is driving demand for access to newer and advanced methods that enable complex modifications [1–5]. Several chemical, chemoenzymatic, and enzymatic strategies have been developed for site-specific ligation of proteins and peptides [1, 2, 6–9]. Among these, enzymatic labeling stands out as the preferred choice due to its applicability to a wide range of proteins under mild conditions and its exceptional specificity at the residue and/or site level [2, 4, 10]. Site-specific transpeptidases and ligases are crucial in these enzymatic and chemoenzymatic ligation strategies. For example, Sortase A, a cysteine transpeptidase, targets the pentapeptide motif LPXTG in its substrates [11–13]. It cleaves the peptide bond between threonine and glycine, forming an acyl-enzyme intermediate, which is subsequently resolved through nucleophilic replacement with a peptide containing an N-terminal glycine [14, 15]. Another notable example is asparaginyl endopeptidases (AEPs), a group of cyclotide-producing cysteine proteases derived from plants. AEPs have emerged as exceptional bioconjugation tools due to their ability to recognize a tripeptide motif (Asx-Xaa-Yaa) in their substrates. They cleave on the C-terminal side of the Asn (N) and Asp (D) residues, forming an acyl-enzyme intermediate. This intermediate can then ligate to an incoming nucleophile that has Gly at its N-terminus [16–20]. The substrate specificity of AEPs is quite relaxed at the Xaa position, allowing all amino acids except proline (Pro). In contrast, at the Yaa position, only large hydrophobic residues such as phenylalanine (Phe), isoleucine (Ile), leucine (Leu), methionine (Met), and tryptophan (Trp) are preferred [21]. Previous studies have found that dipeptides like GL and GV when installed at the N-terminus of peptide probes and proteins, serve as efficient acceptor nucleophiles [19–22]. Additionally, Cys247 has been identified as a “gatekeeper” residue in OaAEP1 from *Oldenlandia affinis*, influencing the type of nucleophile that can enter the active site [22]. Replacing Cys247 with an alanine residue significantly enhanced the catalytic efficiency, increasing the $k_{\text{cat}}/K_{\text{m}}$ constant from 215 to 34,209 $\text{M}^{-1}\times\text{s}^{-1}$. The variant OaAEP1_C247A has been employed in various applications, including protein and peptide labeling, cyclization, immobilization, chemoenzymatic and enzymatic protein synthesis, and native cell surface labeling [10, 17, 19, 23]. While the use of OaAEP1_C247A as a protein engineering tool is gaining traction, the availability of effective ligases remain limited. The discovery and successful application of OaAEP1_C247A have highlighted the potential for many undiscovered or understudied enzymes in nature that could further revolutionize this field. Therefore, the search for additional enzymes with

enhanced transpeptidation and ligation features remains a compelling pursuit.

Butelase 1 (referred to as Butelase hereafter) is another highly efficient asparaginyl endopeptidase (AEP) isolated from the butterfly pea *Clitoria ternatea*. It was the first Asn/Asp (Asx)-specific ligase discovered and recognizes a tripeptide motif Asx-His-Val at the C-terminus of a peptide. Butelase accommodates most N-terminal amino acid residues for transpeptidation and has demonstrated the ability to circularize several non-native substrates with diverse sequence compositions and sizes [24–25]. However, its practical utility as a protein engineering tool has been constrained by the challenges related to the expression and purification of the recombinant protein [25, 27, 28]. Like other AEPs, Butelase is expressed as an inactive proenzyme that requires low pH activation. Both Butelase (PDB ID: 6DHI) and OaAEP1 (PDB ID: 5H0I) exhibit identical structural folds (Fig. 1a), with most residues being conserved. Multiple sequence alignments of Butelase and OaAEP1 reveal significant conservation among the residues of asparaginyl endopeptidases (Fig. S1) [22]. In the catalytically efficient form of OaAEP1, OaAEP1_C247A, the gate-keeping cysteine has been mutated to alanine. A sequence alignment of Butelase and OaAEP1 shows high conservation among the neighboring residues of the gatekeeper (Fig. 1b). In Butelase, the gate-keeping residue is Val237, which, along with Trp236 and Thr238, constitutes the ligase-activity determinant-1 (LAD-1) region of Butelase [28]. Previous structural and mutagenesis studies have indicated that residues in the LAD-1 region are critical for the efficiency of an asparaginyl ligase. Typically, an aromatic and bulky amino acid, such as tyrosine (Tyr) or tryptophan (Trp), is present in the first position, followed by a hydrophobic amino acid in the second position [29]. The LAD-1 of OaAEP1_C247A is composed of $^{246}\text{TrpAlaTyr}^{248}$ ($^{246}\text{WAY}^{248}$), exhibiting two variations compared to Butelase (Fig. 1b and c). The structural superposition of Butelase and OaAEP1 crystal structures (Fig. S2) shows the presence of the Val237 at the extremity of the substrate channel, with its side chain protruding into the channel. This configuration may allow Val237 to function similarly to Cys247 in OaAEP1, acting as a final nucleophile filter by inducing specific conformational changes in the interacting regions of the substrate. Based on previous reports [22, 29], it appears that the side chain of Val237, in consort with Thr238, could influence not only ligase activity but also the catalytic efficiency of butelase by affecting substrate binding and positioning within the substrate channel. We hypothesized that mutating $^{237}\text{ValThr}^{238}$ to $^{237}\text{AlaTyr}^{238}$ would produce a Butelase variant with enhanced catalytic properties and a broader range of accepted substrates. In this study, we present a new, easy-to-purify, and catalytically efficient version

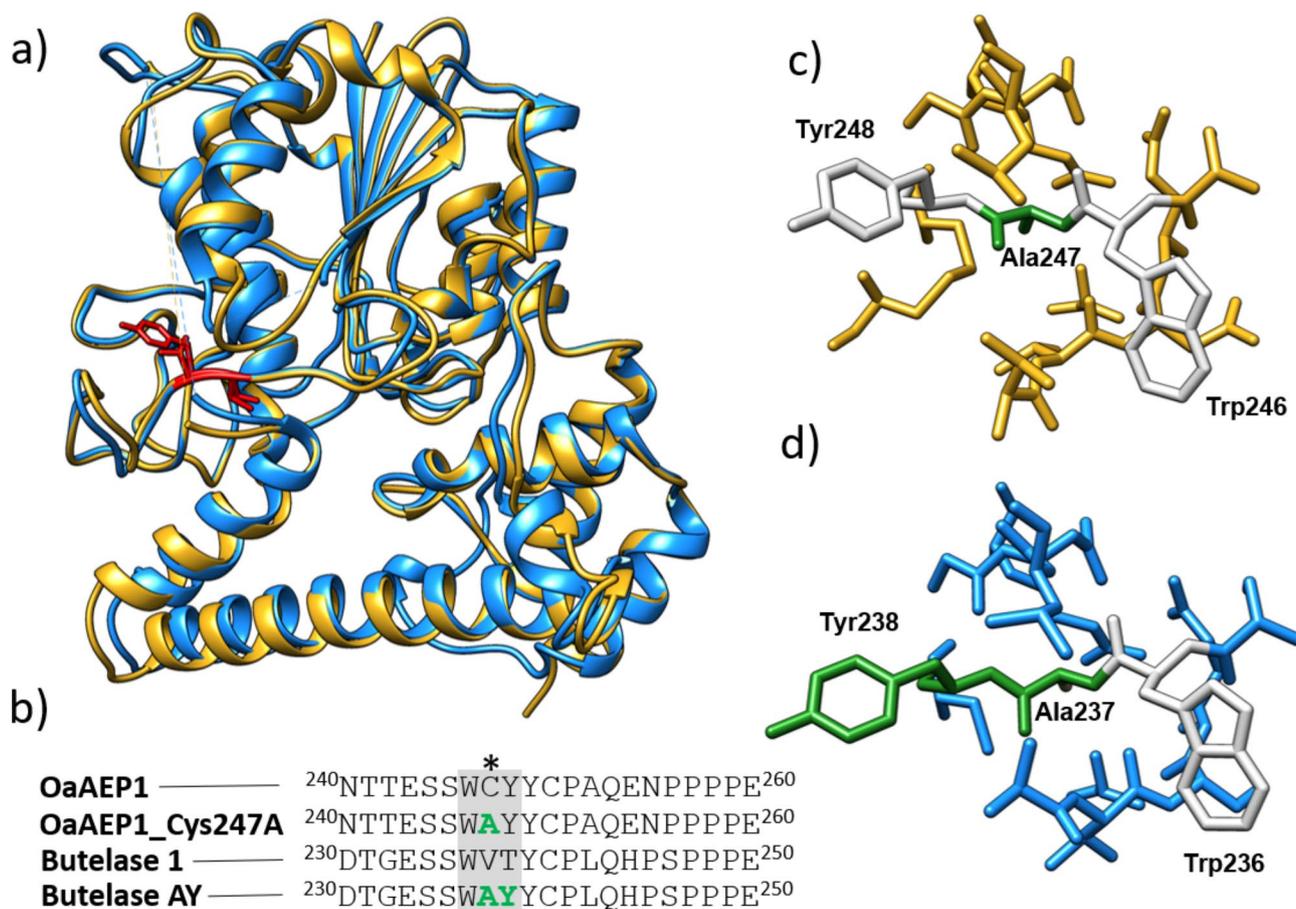


Fig. 1 Structural representation of Butelase and OaAEP1 and their mutants. **(a)** Superposition of Butelase (PDBID: 6DHI, blue color) and OaAEP1 (PDBID: 5H0I, sandy color). Red color-marked residues indicate a channel region of both enzymes (see Fig. S1 for a zoom); **(b)** Sequence alignment of the activity region of wild-type OaAEP1, OaAEP1_C247A mutant enzymes, Butelase 1, and double mutant Butelase AY. The grey bar indicates amino acid residues surrounding gatekeeping residues marked with the star. Green residues represent mutated positions. **(c)** and **(d)** Structural model of OaAEP1_Cys247A and Butelase AY with the group of atoms (marked with gold and blue color) present < 8Å from gatekeeping residues. Mutated residues are marked in green. Structural visualization and structural model development was done by UCSF Chimera 1.11.2

of Butelase, termed Butelase AY, obtained by mutating Val237Ala (the gatekeeping residue) and Thr238Tyr in the LAD-1 region of Butelase. The structural model of the mutant shows no conflicts or clashes among residues (Fig. 1d). We propose that Butelase AY can be effectively utilized for a wide range of peptide and protein engineering applications.

Materials and methods

Materials

Ethylenediaminetetraacetic acid sodium salt (EDTA), thioanisole, triisopropylsilane (TIPS), N-tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid (TES), Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) and sodium acetate were purchased from Sigma-Aldrich. The *N,N*-dimethylformamide (DMF), and HCl were from VWR Chemicals. Triton X-100 and 5 [6]-carboxyfluorescein (FAM) were bought from Merck. Sodium chloride (NaCl), dimethyl sulfoxide (DMSO),

diethyl ether, glycerol, acetic anhydride, 1-methyl-2-pyrrolidinone (NMP), trifluoroacetic acid (TFA), piperidine, dichloromethane (DCM), and acetonitrile (MeCN) were purchased from Avantor Poland. TentaGel S Ram, OxymaPure, DIC (*N,N*-diisopropyl carbodiimide), Fmoc-L-Leu-Wang resin, Fmoc-L-amino acids, unnatural amino acids, such as N-methyl-L-alanine (MetAla), α -aminoisobutyric acid (Aib), L-norvaline (NorVal), were purchased from Iris Biotech GmbH (Marktredwitz, Germany). PD-10 Desalting Columns were from GE Healthcare. HisPur™ Ni²⁺-NTA resin was from Thermo Scientific™. Sodium dodecyl sulfate (SDS), tryptone, yeast extract, LB Broth, agar, agarose, and isopropyl- β -D-1-thiogalactopyranoside (IPTG) were purchased from Lab Empire. Kanamycin monosulphate, 1,4-dithiothreitol (DTT), and Tris base were from Roth, pTYB21 vector and chitin resin from New England BioLabs, Chelex 100 resin from Bio-Rad, and 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid sodium salt (HEPES) were bought

from Bioshop. Ethanol was purchased from Sinopharm Chemical Reagent Co. Ltd. All aqueous solutions were prepared with Milli-Q water and were filtered using microporous membrane filters (0.22 and 0.45 μm) if required. Chelax 100 resin was used to remove the metal ion contamination from all pH buffers. For the culture of *E. coli*, Luria–Bertani (LB) medium and agar plates were used.

Plasmid construction and DNA manipulation

Plasmid pQE30-Butelase1 (Addgene #115430) was a gift from Joshua Mylne [27]. It was used as a template for the amplification of the Butelase gene with specific primers (Table S1). The Butelase was first subcloned in pET28 to form pET28-Butelase (deposited now in Addgene under #221848). The butelase gene was also subcloned in plasmid SUMO-MT3 between BamHI and HindIII sites to generate SUMO-Butelase [30]. Butelase AY was generated by site-directed mutagenesis (SDM) of SUMO-Butelase using appropriate primers (Table S1). Plasmid pBHRSF184 (Addgene #89482) encoding untruncated OaAEP1 enzyme, plasmid pJTRSF55 (Addgene #89701) encoding GL-GFP and plasmid pFGET19_Ulp1 (Addgene plasmid #64697) encoding Ulp1 protease were a gift from Hideo Iwai [31]. OaAEP1_C247A (Addgene #200307) was generated by SDM as described in our previous report [17]. Plasmid pETM11-SUMO-SNCA-GFP (Addgene #107292) was a gift from Dmytro Yushchenko and was used as a template to generate $\Delta 24$ α -synuclein using specific primers [32] (Table S1). The amplified $\Delta 24$ _ASN gene was cloned in plasmid eSrtA-MT3 (Addgene #200305) between NdeI and XhoI [30]. The resultant eSrtA- $\Delta 24$ -ASN plasmid has been submitted to Addgene (plasmid #221849).

Protein expression and purification

Butelase, butelase AY, OaAEP1 WT, and OaAEP1_C247A

The enzymes OaAEP1 WT, OaAEP1_C247A, Butelase, and Butelase AY (deposited now in Addgene under #221847) were expressed and purified similarly as described in our previous report [17]. Briefly, *E. coli* RIL cells were transformed with plasmids for protein expression, and a primary culture was inoculated in Super broth (SB) at 37 °C overnight with constant shaking at 180 rpm. The secondary culture was inoculated in SB by adding 1% (v/v) primary culture and grown at 37 °C until OD₆₀₀ reached 0.6–0.8. Protein expression was induced with 0.1 mM IPTG and incubated overnight at 20 °C with constant shaking at 120 rpm. The proteins were purified using Ni²⁺-NTA agarose beads and imidazole as described in our previous report [17]. To remove imidazole, the eluted proteins were desalted using phosphate buffer saline (PBS) on a PD-10 desalting column. The fusion protein was then digested at

24 °C for 2 h by adding 1 mM dithiothreitol (DTT) and 37 nM Ulp1 protease. N-terminally His-tagged SUMO and His₆-Ulp1 protease were removed using a second Ni²⁺-NTA purification. The proteins were then purified by size-exclusion chromatography and the fractions containing the proteins were collected. To the collected fraction, 1 mM EDTA and 0.5 mM tris(2-carboxyethyl) phosphine (TCEP) were added, and the pH was adjusted to 4.0 with glacial acetic acid. The resultant protein solution was then incubated at 37 °C for 5 to 16 h for self-cleavage auto-activation. However, the activation of the OaAEP1_C247A was carried out using the addition of 10% of WT OaAEP1 to the inactive enzyme (transactivation). Most of the contaminating proteins precipitated at this acidic pH during auto-activation (they were removed by centrifugation). The activity of the resulting enzymes was analyzed, and they were stored in aliquots at -80 °C. Protein concentrations were routinely determined by Nanodrop OneC (Thermo Scientific) using the respective molecular weight and molar absorption coefficients.

Ulp1 (Ubl-specific protease 1) and green fluorescent protein (GFP)

SUMO protease Ulp1 and GFP (with N-term GL) were expressed and purified using previously reported protocols [30, 31].

$\Delta 24$ α -synuclein ($\Delta 24$ _ASN) purification

$\Delta 24$ _ASN was expressed as a SUMO fusion protein 6 \times His-SUMO-eSrtA-LPNTG- $\Delta 24$ -ASN and was purified with Ni²⁺-NTA affinity purification followed by Ca²⁺-mediated self-cleavage with eSrtA as described previously for eSrtA-MT3 [30].

Peptide synthesis and purification

All peptides were synthesized via solid phase synthesis, using the Fmoc strategy on a Liberty Blue 2.0 microwave-assisted synthesizer (CEM). Synthesis of the fluorescent peptide, FAM-GKYINHV, was done by modifying the N-terminus of the linear peptide GKYNHV with 5 [6]-carboxyfluorescein (FAM) in an additional step – 2 h coupling with the reaction mixture of Oxyma, DIC, and 5 [6]-carboxyfluorescein (3 mol equiv. excess) at room temperature. The cleavage of the peptides from the resin was done by incubating the peptide resin with a mixture of TFA/TIS/water (95/2.5/2.5, v/v/v) for 2.5 h followed by precipitation in ice-cold diethyl ether. The precipitated crude peptide was freeze-dried and later purified by RP-HPLC (Varian) on a C18 preparative column, Gemini, 5 μm , 10 \times 250 mm, using a gradient of MeCN in 0.1% TFA/water. The purified peptide fractions were collected and lyophilized. The identity of the purified peptides was confirmed by ESI-MS. The list of synthesized peptides,

along with their calculated and experimental mass, is presented in Table S2.

RP-HPLC analysis of ligation reactions catalyzed by butelase AY

The peptide ligation assays were carried out in a 20 μ L of reaction mixture at 37 °C. The reaction mixture consisted of reaction buffer (50 mM sodium acetate pH 5.6, 50 mM NaCl, 1 mM EDTA, and 0.5 mM TCEP), 25 μ M Butelase AY, 0.5 mM GLPVSTKPVATRNHV or 0.25 mM N-terminal substrates (FAM-GKYINHV, YKVINHV) and 0.5 mM C-terminal substrates (GFP, Δ 24 ASN or G*VGKY-NH₂, where * indicates L-leucine (L), L-isoleucine (I), L-valine (V), N-methyl-L-alanine (MetAla), L-norvaline (NorVal), and α -aminoisobutyric acid (Aib) residues). The reactions were stopped with 0.1% TFA, and the reaction mixtures were analyzed by RP-HPLC on a C-18 column using a MeCN/0.1% TFA gradient from 5 to 35% in 40 min and from 35 to 85% in 20 min, and the absorbances were recorded at 220 and 280 nm. The % substrate conversion corresponds to the amount of substrate processed by the enzyme. The percentage substrate conversion for each enzyme was calculated by integrating the area under the HPLC peaks. ESI-MS confirmed the composition and identity of each HPLC peak. Experimental and calculated average molecular masses are provided in Table S3.

RP-HPLC analysis of ligation reactions catalyzed by OaAEP1_C247A

The peptide ligation assays were carried out in a 20 μ L of reaction mixture at 37 °C. The reaction mixture consisted of reaction buffer (50 mM sodium acetate pH 5.6, 50 mM NaCl, 1 mM EDTA, and 0.5 mM TCEP), OaAEP1_C247A (25 μ M), 0.25 mM N-terminal substrates (YKVINGL) and 0.5 mM C-terminal substrates (G*VGKY-NH₂, where * is leucine, valine, N-methyl-L-alanine (MetAla), L-norvaline (NorVal), and α -aminoisobutyric acid (Aib) residues). The reactions were stopped with 0.1% TFA, and the reaction mixtures were analyzed by RP-HPLC, as described in the previous section.

Structural modelling and visualization

The crystal structures of Butelase and OaAEP1 were visualized using UCSF Chimera 1.11.2. Additionally, the development and modulations of the structural models were conducted using the same software.

Results

Based on the structural superposition of the crystal structures of Butelase and OaAEP1 (Fig. 1a), multiple sequence alignment (Fig. 1b), and structural models of LAD-1 region generated by UCSF Chimera 1.11.2 (Fig. 1c and d), Val237 and Thr238 residues from Butelase were

selected for mutagenesis to Ala237 and Tyr238. pQE30-Butelase was subcloned in the SUMO-MT3 vector as a SUMO-fusion protein, as described in Materials and Methods [30]. Using specific Site-Directed Mutagenesis (SDM) primers (Table S1), the V237A and T238Y mutations were incorporated. The SUMO-fusion constructs of Butelase and Butelase AY and all other proteins were expressed and purified as described in Materials and Methods. The purification profile of Butelase AY was analyzed by SDS-PAGE (Fig. 2a). At the neutral pH of purification (pH=7.4), both Butelase WT and Butelase AY are mainly present as a dimer in solution, where both monomers are in their zymogenic form. In this zymogenic form, the C-terminal pro-domain forms a cap covering the core domain's active site. The pro-domain also contains the residues responsible for the dimerization of Butelase WT and Butelase AY [25, 27, 29, 33]. Dissociation of the cap domain occurs upon a shift to a low-pH environment where autocatalytic self-cleavage occurs at a flexible linker region and results in the enzymatically active monomeric Butelase WT and Butelase AY [27, 29, 33]. Autoactivation of Butelase WT and Butelase AY was carried out at pH 4.0, converting them to the catalytically active monomeric form. In the SDS-PAGE, the 75 kDa band corresponds to the SUMO-Butelase AY (and SUMO-Butelase WT) fusion protein. After activation, the catalytic Butelase AY can be seen as a 38 kDa protein on the SDS-PAGE (Fig. S3). During the autoactivation step, a lot of precipitation was observed for Butelase WT, but very little precipitation was seen for Butelase AY. The SEC profile of Butelase AY before and after activation (Fig. 2b) pointed out the presence of more than one autoactivation cleavage site as previously reported for other asparaginyl endopeptidases (i.e. OaAEP1, OaAEP1_C247A, Butelase WT, etc.) [17, 22]. One more interesting fact was observed that Butelase AY was activated by cis self-cleavage, which is different from the autoactivation of OaAEP1_C247A which requires 10% (v/v) of OaAEP1 WT for its activation (Fig. S4) [17]. The activity of Butelase AY was tested using peptide substrate YKVINHV along with two different peptidyl nucleophiles, GLGKY and GVGKY (Fig. 2c, Fig. S5). Butelase AY demonstrated activity and successfully recognized both peptide substrates. Unfortunately, no activity was observed for the butelase WT, despite three different purifications attempts. The loss of activity of Butelase WT may have occurred during the autoactivation step, during which a significant portion of the purified Butelase WT precipitated, leading to a reduced overall yield. In contrast, very little precipitation was observed during the autoactivation of Butelase AY, and the enzyme remained catalytically active. This observation suggests that the mutations in Butelase AY may enhance its stability. This is an

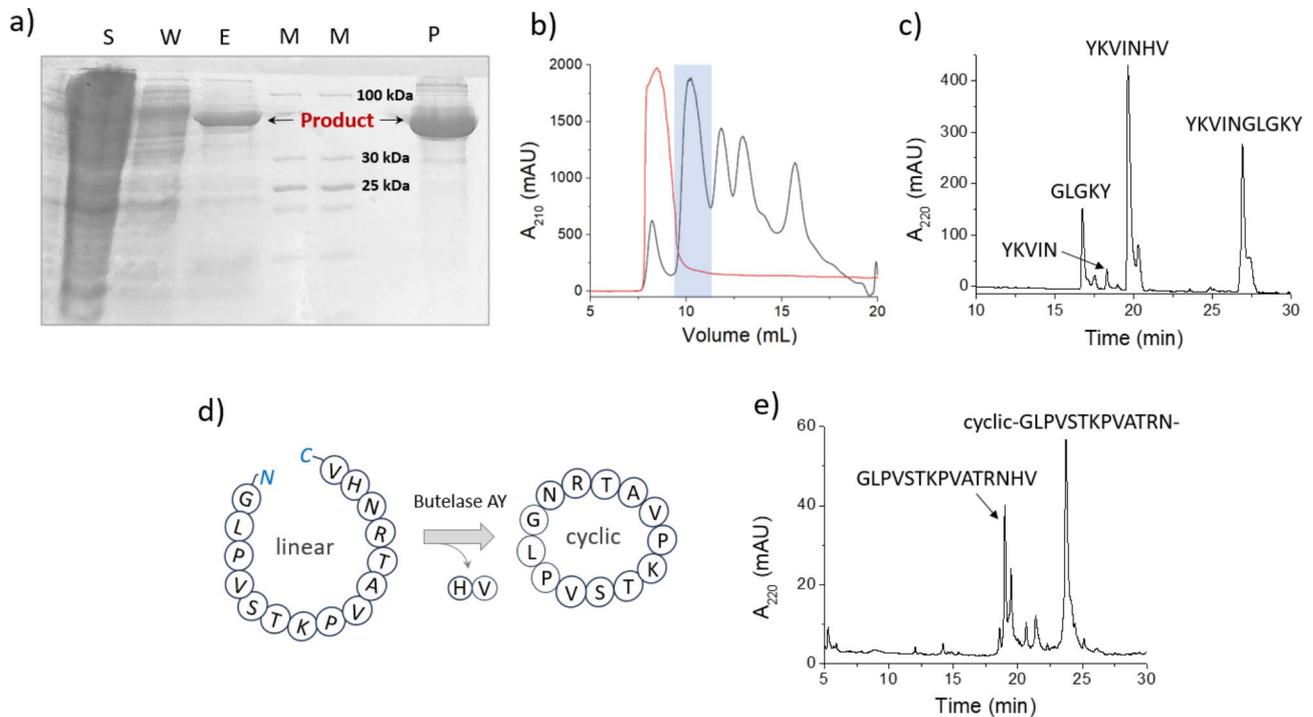


Fig. 2 Purification and activity profile of Butelase AY. **(a)** The SDS-PAGE purification profile of Butelase AY. The SDS-PAGE lanes represent the different fractions from the protein purification step. The fractions are S = supernatant, W = wash fraction, E = elution, M = protein ladder, and P = pellet fraction. **(b)** Size-exclusion chromatography profiles of Butelase AY before (in red) and after (black) activation showing Butelase AY dimer and the monomeric Butelase AY holo-enzyme (blue box) along with other cleaved protein/peptide fragments. **(c)** RP-HPLC analysis of the Butelase AY catalyzed reaction between YKVINHV and GLGKY peptides: the reaction was carried out using 0.25 mM YKVINHV, 0.5 mM GLGKY, and 25 μ M of Butelase AY at 37 $^{\circ}$ C for 1 h. The reaction was quenched by adding 0.1% TFA and was analyzed on a C18 analytical RP-HPLC column using a MeCN /0.1% TFA gradient from 5 to 35% in 40 min and from 35 to 85% in 20 min. The absorbance was recorded at 220 nm. **(d)** Schematic representation of the Butelase AY mediated cyclization of peptide GLPVSTKPVATRNVH and **(e)** A representative RP-HPLC profile of the cyclization reaction. Assays were performed at 37 $^{\circ}$ C for 60 min in the presence of 25 μ M Butelase AY

intriguing hypothesis that warrants further extensive research to draw any concrete conclusions.

Since Butelase is primarily involved in the synthesis and cyclization of several secondary metabolites in plants, we aimed to evaluate the cyclization efficiency of Butelase AY. The peptide cyclization activity of Butelase AY was assessed using the peptide substrate GLPVSTKPVATRNVH (Fig. 2d). The cyclization reactions were analyzed using RP-HPLC and mass spectrometry confirmed the successful cyclization of the peptide substrate. This peptide substrate, GLPVSTKPVATRNVH, along with a similar peptide substrate, GLPVSTKPVATRNVGL, has been previously used to evaluate and compare the cyclization efficiencies of OaAEP1_C247A, OaAEP1 WT and Butelase WT [22]. Therefore, using this peptide for cyclization will provide us with a good chance to compare the cyclization efficiency of Butelase AY with that of OaAEP1_C247A and other AEPs reported in the literature.

To better understand the impact of mutating these two gatekeeping residues (GK) in the LAD-1 region on the activity of Butelase AY and to experimentally compare its cyclization efficiency with OaAEP1_C247A, we

measured the kinetic parameters for the cyclization of GLPVSTKPVATRNVH at varying concentrations from 0.1 mM to 5 mM over a 60 min reaction (Fig. 2e, Fig. S5). The K_m value of Butelase AY was determined to be 0.75 mM, while the turnover rate (k_{cat}) was 10 s^{-1} (Fig. S6). The k_{cat} of 10 s^{-1} observed for Butelase AY remains in good agreement with the previously reported k_{cat} values for OaAEP1_C247A (13.9 s^{-1}) and Butelase WT (4.83 s^{-1}) [22], Table 1. In the previous reports, the kinetic parameters have been determined using nanomolar (nM) substrate concentrations while in the present study substrate concentration in the millimolar (mM) range was used. Since the reported values were determined from the nM range kinetic studies, those K_m values are significantly lower than the K_m value obtained in the present study. Notably, we also found that the expression and purification of active recombinant Butelase AY is significantly easier than that of Butelase WT. If Butelase AY demonstrates catalytic performance similar to or better than that of OaAEP1_C247A, it could enhance its applicability in protein engineering, labeling, imaging, and therapeutics.

Table 1 List of the available ligases and their variants from the asparaginyl endopeptidase (AEP) family

Enzyme	Recognition motif	Nucleophile	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1} \times s^{-1}$)	Reference
nButelase-1	NHV	GL	2.28	542,000	[25]
nButelase-1	NHV	GI	26.55	1,314,000	[26]
rButelase1	NHV	GL	2.15	9641	[37]
rButelase1	NHV	GL	2.53	11,048	[36]
rButelase 1	NHV	GL	4.83	99,793	[22]
Butelase AY	NHV	GL	10	13,330	This study
OaAEP1	NGL	GL	0.052	215	[22]
OaAEP1_C247A	NGL	GL	13.9	34,209	
OaAEP1	NGL	GL	0.99	6800	[38]
OaAEP3	NGL	GL	0.61	330,000	
OaAEP4	NGL	GL	0.76	984,000	
OaAEP5	NGL	GL	0.59	247,000	
OaAEP1_C247A	NGL	GL	1.52	10,400	[21]
MCoAEP2	NAL	GG	19.86	62.1	[39]
VyPAL2	NGL	GI	5.5	214,445	[40]
OaAEP1_C247A_aa55-351	NAL	RL	0.922	100,647	[41]
			11.6	2,376,384	

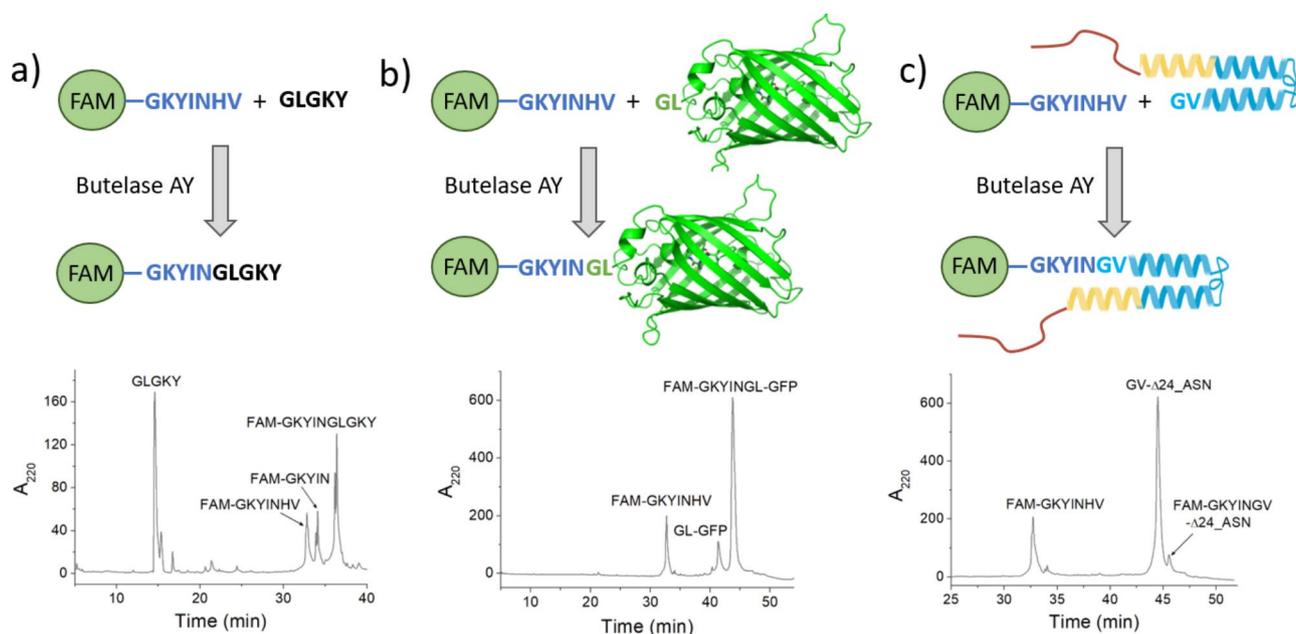


Fig. 3 Butelase AY mediated fluorescent labeling of peptides and proteins. Schematic representation and HPLC profile of the Butelase AY (25 μ M) catalyzed ligation reaction between 0.25 mM fluorescent probe (FAM-GKYVINHV) and 0.5 mM peptide GLGKY **(a)**, 0.5 mM green fluorescent protein **(b)**, and 0.5 mM Δ 24 α -synuclein (Δ 24_ASN) **(c)**. The reactions were incubated at 37 $^{\circ}$ C for 1 h. The reaction was quenched by adding 80 μ l of 0.1% TFA, and the reaction mixtures (sample final volume 100 μ l) were injected directly into a C18 analytical RP-HPLC column. The reaction was analyzed using a MeCN/0.1% TFA gradient from 5 to 35% in 40 min and from 35 to 85% in 20 min, and absorbance was recorded at 220 nm. The RP-HPLC peaks representing the substrates, hydrolysis by-product, and the final ligation product have been labeled

To explore the potential of Butelase AY as a peptide and protein engineering tool, we qualitatively demonstrated its capability to label model peptides as well as native proteins. We initiated the study by modifying the small model peptide, GLGKY, with a molecular probe FAM-GKYINHV, where FAM represents the 5 [6]-carboxyfluorescein scaffold moiety. The ligation reaction

was analyzed via RP-HPLC and the identity of the final product FAM-GKYINGLGKY was confirmed by ESI-MS (Fig. 3a).

Additionally, we applied Butelase AY to fluorescently modify two distinct native proteins, green fluorescent protein (GFP) and α -synuclein (ASN). These proteins were selected due to their contrasting structural and

functional properties. GFP is a well-characterized, easy-to-express, soluble globular protein that has been used extensively for protein engineering applications, whereas ASN is an intrinsically-disordered pathologically important aggregation-prone neuronal protein associated with neurodegenerative diseases, making it challenging to express and study. The ASN was selected to evaluate the scope of Butelase-mediated protein engineering towards intrinsically disordered proteins. The selection of the $\Delta 24$ variant of ASN was done because of the presence of GV at this position as well as its easy availability in our lab. GFP and a $\Delta 24$ _ASN containing Gly-Leu and Gly-Val, respectively, at their N-terminus were expressed and purified as described in the Material and Methods section. The Butelase AY-mediated modification of GL-GFP with the fluorescent probe FAM-GKYINHV was successfully achieved with an impressive efficiency of over 90% in a 1 h reaction (Fig. 3b). The identity of the fluorescently labeled GFP was confirmed using ESI-MS (Table S3). In contrast, the fluorescent labeling of $\Delta 24$ _ASN using Butelase AY was also successful, but with significantly lower efficiency—only 5% of the substrate converted to product (Fig. 3c). Notably, we observed precipitation of the substrate $\Delta 24$ _ASN during the 1 h incubation period. The identity of the fluorescently labeled $\Delta 24$ _ASN was also confirmed by ESI-MS (Table S3).

These results highlight the utility of Butelase AY in modifying both structured and intrinsically disordered proteins, showcasing its potential in peptide and protein engineering applications.

For protein engineering and site-specific labeling applications, the search for novel enzymes capable of recognizing substrates constituting not only natural amino acids, but also unnatural amino acids is gaining traction. The incorporation of unnatural amino acids and their successful recognition as a substrate can exponentially expand the biorthogonal applications of these protein engineering toolkits. To investigate the substrate preference and tolerance of Butelase AY and OaAEP1_C247A towards natural as well as unnatural amino acid-based substrates, we generated several peptide variants by incorporating leucine (Leu), α -aminoisobutyric acid (Aib), norvaline (NorVal), and N-methyl-L-alanine (MetAla) at the second position of the acceptor aminoglycine peptide G*GKY (Fig. 4a). These non-natural amino acids were introduced during solid-phase peptide synthesis. The transpeptidation assays were carried out with these peptide substrates (Table S2). The reactions were analyzed using RP-HPLC, and all final products were detected and confirmed by ESI-MS (Table S3). The same experiment with the same peptides was also carried out for OaAEP1_C247A. Butelase AY recognized all peptide substrates but exhibited different levels of proficiency. Among the unnatural amino acid-based

peptide substrates, Butelase AY recognized methyl alanine (MetAla), 2-aminoisobutyric acid (Aib), and norvaline (NorVal) at the second position and the % substrate conversion obtained was 21%, 88%, and 95%, respectively for a 24 h reaction (Fig. 4b and c and S7). In the case of OaAEP1_C247A, the best conversion rate among the unnatural amino acids was obtained for 2-aminoisobutyric acid (Aib), while norvaline (NorVal) was not recognized at all (Fig. 4b and d and S8, S9). The % substrate conversion obtained for MetAla and Aib was 56% and 74%, respectively. Leucine was recognized by both OaAEP1_C247A and Butelase AY equally well, and the % substrate conversion obtained was 89% and 96%, respectively. This differential recognition of norvaline—where it is recognized by Butelase AY but not by OaAEP1_C247A—opens avenues for designing biorthogonal experiments. By leveraging this unique substrate recognition pattern, researchers can explore specific labeling and modification strategies using norvaline-based substrates selectively with Butelase AY, enhancing the versatility of protein engineering applications.

Discussion

Butelase, the most effective plant asparaginyl endopeptidase identified to date, is typically derived from the seeds of *Clitoria ternatea* for its use in protein engineering endeavors [22, 24–26]. Numerous attempts have been made at expressing and purifying recombinant Butelase, employing various methods, but with limited minimal achievements [27–29]. The expression and purification of recombinant Butelase are challenging due to several factors. They are susceptible to denaturation and degradation during the purification process due to factors such as pH and temperature [27]. Butelase, like many other plant enzymes, may form inclusion bodies or aggregate when overexpressed in *E. coli*. This reduces the solubility and makes purification more challenging [27, 28]. Optimizing and engineering the recombinant construct of Butelase remains challenging due to the difficulty in producing an active recombinant enzyme. Consequently, despite being the most effective AEP identified to date, the challenges associated with producing recombinant Butelase prevent it from being included in the list of preferred protein engineering toolkits. Sortase, a family of cysteine transpeptidases, is another widely used enzymatic tool for peptide and protein engineering applications [12–15, 30]. Although the expression, purification, and reaction requirements for sortase are relatively easy, its poor catalytic efficiency, Ca^{2+} -dependency, and need for a larger, more flexible pentapeptide recognition motif (LPxTG) limits its use [12, 13]. Ca^{2+} -independent sortase variants have also been developed by mutagenesis [34]. Naturally occurring Ca^{2+} independent sortases have also been discovered and are being utilized in protein engineering

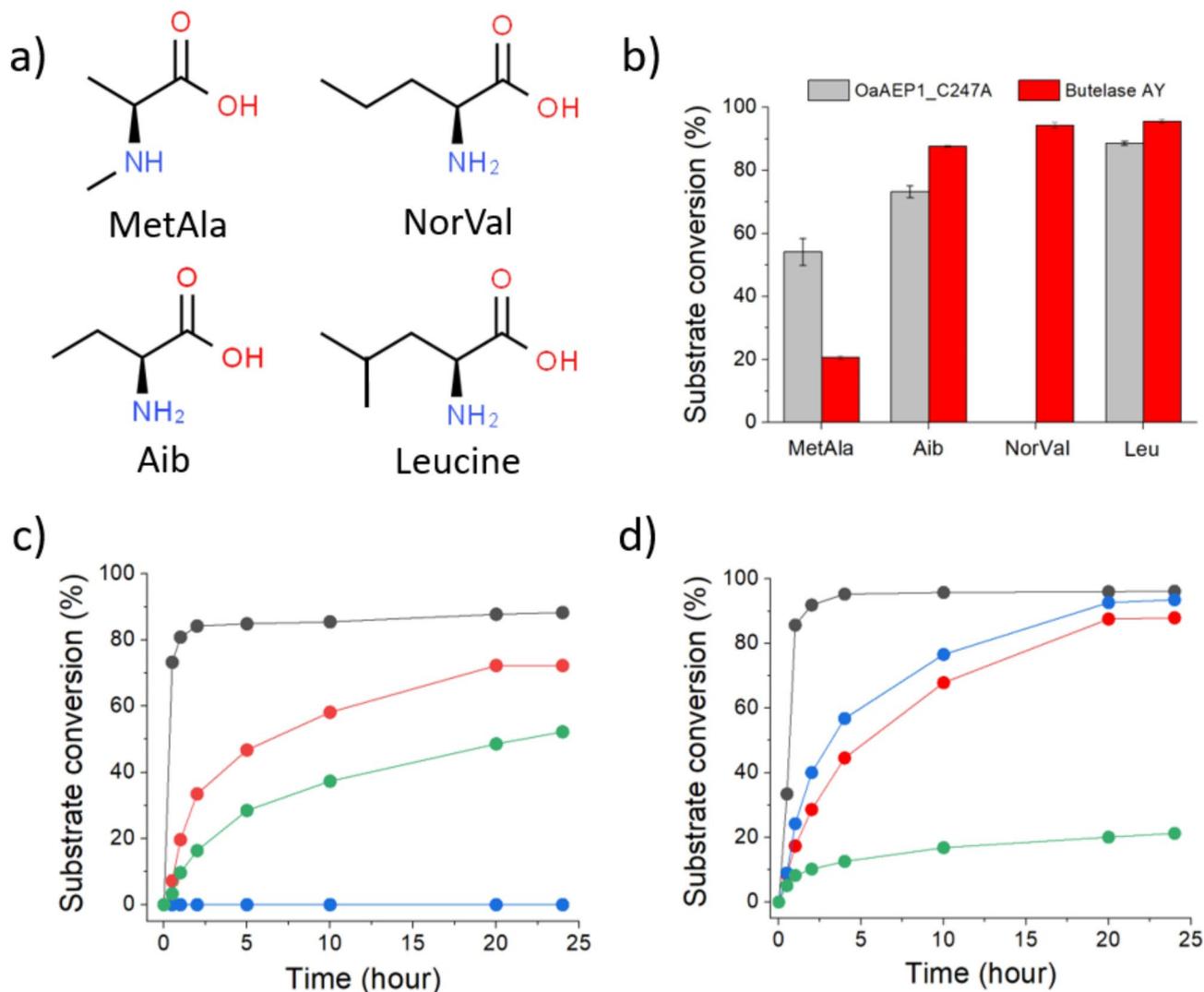


Fig. 4 Recognition and tolerance of non-natural amino acids at the second position of the amino-glycine acceptor peptide substrate (G*GKY) by Butelase AY and OaAEP1_C247A. Transpeptidation assays were carried out using 0.25 mM N-terminal substrate (YKVINHV for Butelase AY and YKLANGL for OaAEP1_C247A) as a model donor and 0.5 mM C-terminal G*GKY as an acceptor peptide in the presence of 25 μM ligase (Butelase AY or OaAEP1_C247A) at 37 °C. **(a)** Chemical structure of residue integrated at * position in G*GKY. **(b)** Comparison of the transpeptidation yield for Butelase AY and OaAEP1_C247A with various substrates after 24 h of reaction. The error bars represent Mean + SD for duplicate samples. **(c)** Time course of the transpeptidation reaction of YKLANGL with G*GKY in the presence of 25 μM OaAEP1_C247A. **(d)** Time course of the transpeptidation reaction of YKVINHV with G*GKY in the presence of 25 μM Butelase AY. Leucine (black), Aib (Red), NorVal (Blue), and MetAla (green) represent * in G*GKY peptide substrate in c and d. One replicate from duplicate readings was plotted for c and d

applications [12–13]. Recently, a promising plant AEP from *Oldenlandia affinis*, known as OaAEP1, along with its variant OaAEP1_C247A, has gained significant attention due to the relative ease of expressing and purifying active recombinant forms of both OaAEP1 and OaAEP1_C247A compared to Butelase [16–23]. OaAEP1_C247A features alanine at position 247 instead of cysteine as the gatekeeping residue in the LAD-1 region. This mutation enhances the enzyme's catalytic efficiency by 140-fold [22, 29]. OaAEP1 is a cysteine protease with ligase activity and functions as a post-translational processing enzyme in the biosynthesis of cyclic peptides. This ligase

activity of OaAEP1 and OaAEP1_C247A has enabled their wider application in peptide and protein modifications. Still, some issues with the recombinant expression, reaction reversibility and hydrolysis exist with the OaAEP1 and OaAEP1_C247A but significant advances are being made to address these issues [20–22, 35]. The AEPs, Butelase and OaAEP1, have a similar and highly conserved 3D structure consisting of a prodomain (cap), a core enzymatic domain, and a linker region. They also share strong sequence homology, especially in their core enzymatic domains, and their catalytic residues are identical [29]. The LAD-1 region and gatekeeping residues in

Butelase and OaAEP1 also show a strong conservancy. The Cys247 of OaAEP1, a gatekeeping residue in the LAD-1 region, is not conserved in Butelase [22, 29]. In OaAEP1, the Cys247 residue is located at the extremity of the substrate channel and functions as a final nucleophile filter (or “gatekeeper”), allowing an attack from the N-terminal amine substrate to complete the ligation reaction [22, 29]. The homologous residue Val237, along with Thr238 in Butelase, also supposedly plays a similar gatekeeping role. Mutation of these two residues, Val237Ala and Thr238Tyr, created a LAD-1 region identical to the LAD-1 of OaAEP1_C247A. The resultant mutant variant Butelase AY was expected to be an evolved version of Butelase with an improved and broadened conjugation efficiency for both peptides and proteins. The expression, purification, and activation of Butelase WT and Butelase AY were done parallelly. However, during activation, most of the Butelase WT protein precipitated. On the other hand, very little precipitation was observed for Butelase AY. When the ligation activity of these two enzymes was tested with peptide substrate YKVINHV and GLGKY, the product was only obtained for the Butelase AY. Several attempts to produce active Butelase WT failed, mostly due to the precipitation during the autocatalytic activation at acidic pH. Additional efforts were made to recover Butelase AY from the insoluble material using refolding protocols; however, these attempts were unsuccessful. The incorporation of these two mutations may have contributed to the stabilization of the enzyme during autocatalytic activation at acidic pH. Since these mutations are located in the active site channel, they are unlikely to directly affect the enzyme’s solubility. However, they could induce conformational changes that reduce the enzyme’s hydrophobicity, potentially enhancing solubility and decreasing precipitation. These aspects remain unresolved and require further investigation. For comparative studies, active OaAEP1_C247A was also expressed and purified [17, 19]. We noticed one very significant difference in the autoactivation of OaAEP1_C247A and Butelase AY. During the activation of the OaAEP1_C247A, we need to add 10% (v/v) of activated OaAEP1 WT for activating OaAEP1_C247A. Although many groups have not reported encountering such issues with the activation of OaAEP1_C247A. In the case of Butelase AY no such addition of activated Butelase WT is required. The Butelase AY was able to autoactivate itself. The catalytic efficiency of Butelase AY was determined using peptide GLPVSTKPVATRNVH. The calculated K_m value of Butelase AY for this peptide (0.1 mM – 5 mM) was 0.75 mM, while the turnover rate (k_{cat}) was 10 s^{-1} . The reported k_{cat} values of Butelase WT for GLPVSTKPVATRNVH is 4.83 s^{-1} and of OaAEP1 WT and OaAEP1_C247A for GLPVSTKPVATRNVH is 0.052 s^{-1} and 13.9 s^{-1} respectively [22]. The investigated Butelase

AY shows significant promise as a versatile and efficient tool for protein and peptide engineering, with exciting capabilities demonstrated in both natural and unnatural amino acid recognition. While challenges remain in its expression, purification, and activation, such as susceptibility to denaturation and aggregation, the study provides valuable foundational insights. These findings set the stage for further research to optimize production and fully realize the enzyme’s potential, highlighting its promising role in advancing protein engineering and synthetic biology. A recent report by Zhao et al. employed strain screening, co-expression of molecular chaperones, rare codon substitution, and RBS optimization to enhance the expression and activity of Butelase [36].

The labeling efficiency of Butelase AY was demonstrated by successfully ligating a fluorescein-labeled peptide, FAM-GKYINHV, to GLGKY. To further establish its potential as a protein labeling tool, we used full-length GFP with an N-terminal GL recognition motif as a substrate; achieving a labeling yield of more than 90% with the fluorescein-labeled peptide FAM-GKYINHV. GFP, a highly soluble globular protein with a stable and well-defined tertiary structure, served as an effective model. However, the ligation of FAM-GKYINHV and GL-GFP was only intended to qualitatively show the application of Butelase AY towards native globular proteins. We also aimed to investigate whether Butelase AY would recognize unstructured and intrinsically disordered proteins as substrates. To this end, we expressed and purified a $\Delta 24$ variant of the intrinsically disordered, aggregation-prone neuronal protein α -synuclein (ASN), which features a GV recognition motif at its N-terminus. While Butelase AY successfully recognized the $\Delta 24$ -GV-ASN and labeled it with the fluorescein-peptide, the product yield was only 5%, significantly lower than that for GFP. One potential reason for the low yield could be the presence of GV (instead of GL) recognition motif at the N-terminus, but since Butelase AY recognizes GL and GV substrates equally well, this is unlikely. Another possibility is the presence of a helical motif near the N-terminus of $\Delta 24$ -ASN, which may restrict the availability of GV for ligation. Additionally, the aggregation of $\Delta 24$ -ASN monomers could hinder ligation, as much of the protein may be present in the reaction mixture but unavailable for ligation due to aggregation.

Engineered Butelase AY variant exhibits significant potential as a versatile and efficient enzyme for protein and peptide engineering thus being valuable addition to the existing ligase pool (Table 1). Butelase AY successfully recognizes and processes both natural and unnatural amino acids, greatly expanding its utility in diverse applications and positioning it as a crucial tool for bioorthogonal strategies. This capability not only enhances its versatility in modifying complex biomolecules but also

opens new avenues for innovation in molecular engineering. However, further exploration of key parameters, such as substrate specificity and reaction rates under varying conditions, would provide a more complete understanding of its enzymatic properties. Expanding the kinetic profile in future research will enhance its utility and allow for broader and more precise applications in protein engineering and synthetic biology. The incorporation of specific mutations in Butelase AY has already shown potential for improved stability and solubility, and further optimization is likely to enhance these properties. These findings set the stage for further research to optimize production and fully realize the enzyme's potential, highlighting its promising role in advancing the field.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12934-024-02598-5>.

Supplementary Material 1

Acknowledgements

The research was supported by the National Science Centre of Poland under Opus grant no. 2019/33/B/ST4/02428 (to A.K.). Authors acknowledge Mr. Krzysztof Ciura's assistance during the initial work with Butelase WT.

Author contributions

Conceptualization: AKS, AK, and AA; Methodology: AKS and AA; Formal analysis and investigation: AKS, AA, AKocyla; Writing – original draft preparation: AKS, AA, and AK; Writing – review and editing: AKS and AK; Resources: AK; Supervision: AK. All authors read and approved the final manuscript.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors approved the manuscript.

Competing interests

The authors declare no competing interests.

Received: 7 September 2024 / Accepted: 23 November 2024

Published online: 30 November 2024

References

- Hanna CC, Kriegesmann J, Dowman LJ, Becker CFW, Payne RJ. Chemical synthesis and semisynthesis of lipidated proteins. *Angew Chemie Int Ed.* 2022;61:15.
- Thompson RE, Muir TW. Chemoenzymatic semisynthesis of proteins. *Chem Rev.* 2020;120:3051–126.
- Muir TW. Semisynthesis of proteins by expressed protein ligation. *Annu Rev Biochem.* 2003;72:249–89.
- Heck T, Faccio G, Richter M, Thöny-Meyer L. Enzyme-catalyzed protein cross-linking. *Appl Microbiol Biotechnol.* 2013;97:461–75.
- Hackenberger CPR, Schwarzer D. Chemoselective ligation and modification strategies for peptides and proteins. *Angew Chemie Int Ed.* 2008;47:10030–74.
- Munari F, Barracchia CG, Franchin C, Parolini F, Capaldi S, Romeo A, Bubacco L, Assfalg M, Arrigoni G, D'Onofrio M. Semisynthetic and enzyme-mediated conjugate preparations illuminate the ubiquitination-dependent aggregation of Tau Protein. *Angew Chemie Int Ed.* 2020;59:6607–11.
- Takaoka Y, Ojida A, Hamachi I. Protein organic chemistry and applications for labeling and engineering in live-cell systems. *Angew Chemie Int Ed.* 2013;52:4088–106.
- Krall N, Da Cruz FP, Boutoureira O, Bernardes GJL. Site-selective protein-modification chemistry for basic biology and drug development. *Nat Chem.* 2016;8:103–13.
- Adusumalli SR, Rawale DG, Thakur K, Purushottam L, Reddy NC, Kalra N, Shukla S, Rai V. Chemoselective and site-selective lysine-directed lysine modification enables single-site labeling of native proteins. *Angew Chemie Int Ed.* 2020;59:10332–6.
- Rehm FBH, Tyler TJ, de Veer SJ, Craik DJ, Durek T. Enzymatic C-to-C protein ligation. *Angew Chem Int Ed.* 2022;61:e202116672.
- Hendrickx APA, Budzik JM, Oh S-Y, Schneewind O. Architects at the bacterial surface - sortases and the assembly of pili with isopeptide bonds. *Nat Rev Microbiol.* 2011;9:166–76.
- Singh AK, Murmu S, Krężel A. One-step sortase-mediated chemoenzymatic semisynthesis of deubiquitinase-resistant Ub-peptide conjugates. *ACS Omega.* 2022;7:46693–701.
- Das S, Pawale VS, Dadireddy V, Singh AK, Ramakumar S, Roy RP. Structure and specificity of a new class of Ca²⁺-independent housekeeping sortase from *Streptomyces avermitilis* provide insights into its non-canonical substrate preference. *J Biol Chem.* 2017;292:7244–57.
- Suree N, Liew CK, Villareal VA, Thieu W, Fadeev EA, Clemens JJ, Jung ME, Clubb RT. The structure of the *Staphylococcus aureus* sortase-substrate complex reveals how the universally conserved LPXTG sorting signal is recognized. *J Biol Chem.* 2009;284:24465–77.
- Biswas T, Pawale VS, Choudhury D, Roy RP. Sorting of LPXTG peptides by archetypal sortase A: role of invariant substrate residues in modulating the enzyme dynamics and conformational signature of a productive substrate. *Biochemistry.* 2014;53:2515–24.
- Harris KS, Durek T, Kaas Q, Poth AG, Gilding EK, Conlan BF, Saska I, Daly NL, Van Der Weerden NL, Craik DJ, Anderson MA. Efficient backbone cyclization of linear peptides by a recombinant asparaginyl endopeptidase. *Nat Commun.* 2015;6:10199.
- Antonenko A, Singh AK, Mosna K, Krężel A. OaAEP1 ligase-assisted chemoenzymatic synthesis of full cysteine-rich metal-binding cyanobacterial metallothionein SmtA. *Bioconjug Chem.* 2023;34:719–27.
- Rehm FBH, Tyler TJ, Xie J, Yap K, Durek T, Craik DJ. Asparaginyl ligases: new enzymes for the protein engineer's toolbox. *ChemBioChem.* 2021;22:2079–86.
- Rehm FBH, Harmand TJ, Yap K, Durek T, Craik DJ, Ploegh HL. Site-specific sequential protein labeling catalyzed by a single recombinant ligase. *J Am Chem Soc.* 2019;141:17388–93.
- Rehm FBH, Tyler TJ, Yap K, Durek T, Craik DJ. Improved asparaginyl-ligase-catalyzed transpeptidation via selective nucleophile quenching. *Angew Chem Int Ed.* 2021;60:4004–8.
- Tang TMS, Cardella D, Lander AJ, Li X, Escudero JS, Tsai YH, Luk LYP. Use of an asparaginyl endopeptidase for chemo-enzymatic peptide and protein labeling. *Chem Sci.* 2020;11:5881–8.
- Yang R, Wong YH, Nguyen GKT, Tam JP, Lescar J, Wu B. Engineering a catalytically efficient recombinant protein ligase. *J Am Chem Soc.* 2017;139:5351–8.
- Harmand TJ, Pishesha N, Rehm FBH, Ma W, Pinney WB, Xie YJ, Ploegh HL. Asparaginyl ligase-catalyzed one-step cell surface modification of red blood cells. *ACS Chem Biol.* 2021;16:1201–7.
- Nguyen GKT, Cao Y, Wang W, Liu CF, Tam JP. Site-specific N-terminal labeling of peptides and proteins using butelase 1 and thiodepsipeptide. *Angew Chemie Int Ed.* 2015;54:15694–8.
- Nguyen GKT, Wang S, Qiu Y, Hemu X, Lian Y, Tam JP. Butelase 1 is an asx-specific ligase enabling peptide macrocyclization and synthesis. *Nat Chem Biol.* 2014;10:732–8.
- Nguyen GKT, Kam A, Loo S, Jansson AE, Pan LX, Tam JP. Butelase 1: a versatile ligase for peptide and protein macrocyclization. *J Am Chem Soc.* 2015;137:15398–401.
- James AM, Haywood J, Leroux J, Ignasiak K, Elliott AG, Schmidberger JW, Fisher MF, Nonis SG, Fenske R, Bond CS, Mylne JS. The macrocyclizing

- protease butelase 1 remains autocatalytic and reveals the structural basis for ligase activity. *Plant J.* 2019;98:988–99.
28. Hemu X, Zhang X, Nguyen GKT, To J, Serra A, Loo S, Sze SK, Liu CF, Tam JP. Characterization and application of natural and recombinant Butelase-1 to improve industrial enzymes by end-to-end circularization. *RSC Adv.* 2021;11:23105–12.
 29. Hemu X, El Sahili A, Hu S, Wong K, Chen Y, Wong YH, Zhang X, Serra A, Goh BC, Darwis DA, Chen MW, Sze SK, Liu CF, Lescar J, Tam JP. Structural determinants for peptide-bond formation by asparaginyl ligases. *Proc Natl Acad Sci USA.* 2019;116:11737–46.
 30. Singh AK, Krężel A. Calcium-assisted sortase a cleavage of SUMOylated metallothionein constructs leads to high-yield production of human MT3. *Microb Cell Fact.* 2023;22:125.
 31. Mikula KM, Tascón I, Tommila JJ, Iwai H. Segmental isotopic labeling of a single-domain globular protein without any refolding step by an asparaginyl endopeptidase. *FEBS Lett.* 2017;591:1285–94.
 32. Afitska K, Fucikova A, Shvadchak VV, Yushchenko DA. Modification of C terminus provides new insights into the mechanism of α -synuclein aggregation. *Biophys J.* 2017;113:2182–91.
 33. Zhao J, Ge G, Huang Y, Hou Y, Hu SQ. Study on activation mechanism and cleavage sites of recombinant Butelase-1 zymogen derived from *Clitoria ternatea*. *Biochimie.* 2022;199:12–22.
 34. Hirakawa H, Ishikawa S, Nagamune T. Design of Ca^{2+} -independent *Staphylococcus aureus* sortase A mutants. *Biotechnol Bioeng.* 2012;109:2955–61.
 35. Tang TMS, Luk LYP. Asparaginyl endopeptidases: enzymology, applications and limitations. *Org Biomol Chem.* 2021;19:5048–62.
 36. Zhao J, Song W, Huang Z, Yuan X, Huang Y, Hou Y, Liu K, Jin P, Hu SQ. Top-down overexpression optimization of butelase-1 in *Escherichia coli* and its application in anti-tumor peptides. *Int J Biol Macromol.* 2024;276:133933.
 37. Zhao J, Fan R, Jia F, Huang Y, Huang Z, Hou Y, Hu SQ. Enzymatic properties of recombinant ligase Butelase-1 and its application in Cyclizing Food-Derived Angiotensin I-Converting enzyme inhibitory peptides. *J Agric Food Chem.* 2021;69:5976–85.
 38. Harris KS, Guarino RF, Dissanayake RS, Quimbar P, McCorkelle OC, Poon S, Kaas Q, Durek T, Gilding EK, Jackson MA, Craik DJ, van der Weerden NL, Anders RF, Anderson MA. A suite of kinetically superior AEP ligases can cyclize an intrinsically disordered protein. *Sci Rep.* 2019;9:10820.
 39. Du J, Yap K, Chan LY, Rehm FBH, Looi FY, Poth AG, Gilding EK, Kaas Q, Durek T, Craik DJ. A bifunctional asparaginyl endopeptidase efficiently catalyzes both cleavage and cyclization of cyclic trypsin inhibitors. *Nat Commun.* 2020;11(1):1575.
 40. Zhang D, Wang Z, Hu S, Balamkundu S, To J, Zhang X, Lescar J, Tam JP, Liu CF. pH-Controlled protein Orthogonal Ligation using asparaginyl peptide ligases. *J Am Chem Soc.* 2021;143:8704–12.
 41. Tang J, Hao M, Liu J, Chen Y, Wufuer G, Zhu J, Zhang X, Zheng T, Fang M, Zhang S, Li T, Ge S, Zhang J, Xia N. Design of a recombinant asparaginyl ligase for site-specific modification using efficient recognition and nucleophile motifs. *Commun Chem.* 2024;7:87.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.