–Characterization of substrate specificity and novel autoprocessing mechanism of dipeptidase A from *Prevotella intermedia*

Mohammad Tanvir Sarwar[1], Yuko Ohara-Nemoto[1], Takeshi Kobayakawa[1], Mariko Naito[2], and Takayuki K. Nemoto[1, *]

[1] *Department of Oral Molecular Biology, Course of Medical and Dental Sciences, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki 852-8588, Japan*
[2] *Department of Microbiology and Oral Infection, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki 852-8588, Japan*

[*] Corresponding author. Takayuki K. Nemoto, Department of Oral Molecular Biology, Course of Medical and Dental Sciences, Nagasaki University Graduate School of Biomedical Sciences, 1-7-1 Sakamoto, Nagasaki 852-8588, Japan. Tel.: +81 95 819 7640; Fax: +81 95 819 7642

E-mail address: tanvirsarwariu@gmail.com (M.T. Sarwar), ynemoto@nagasaki-u.ac.jp (Y. Ohara-Nemoto), takecyan@nagasaki-u.ac.jp (T. Kobayakawa), mnaito@nagasaki-u.ac.jp (M. Naito), tnemoto@nagasaki-u.ac.jp (T.K. Nemoto).

Abstract

Prevotella intermedia, a gram-negative anaerobic rod, is frequently observed in subgingival polymicrobial biofilm from adults with chronic periodontitis. Peptidases in periodontopathic bacteria are considered to function as etiological reagents. *Pre. intermedia* OMA14 cells abundantly express an unidentified cysteine peptidase specific for Arg-4-methycoumaryl-7-amide (MCA). BAU17746 (locus tag, PIOMA14_I_1238) and BAU18827 (locus tag, PIOMA14_II_0322) emerged as candidates of this peptidase from the substrate specificity and sequence similarity with C69-family *Streptococcus gordonii* Arg-aminopeptidase. The recombinant form of the former solely exhibited hydrolyzing activity toward Arg-MCA, and BAU17746 possesses a 26.6% amino acid identity with the C69-family *Lactobacillus helveticus* dipeptidase A. It was found that BAU17746 as well as *L. helveticus* dipeptidase A was a P1-position Arg-specific dipeptidase A, although the *L. helveticus* entity, a representative of the C69 family, had been reported to be specific for Leu and Phe. The full-length form of BAU17746 was intramolecularly processed to a mature form carrying the N-terminus of Cys^{15}. In conclusion, the marked Arg-MCA-hydrolyzing activity in *Pre. intermedia* was mediated by BAU17746 belonging to the C69-family dipeptidase A, in which the mature form carries an essential cysteine at the N-terminus.

Keywords: dipeptidase A; cysteine peptidase, autoprocessing; *Prevotella intermedia*; substrate specificity; periodontal disease.

Introduction

Prevotella intermedia, a gram-negative anaerobic rod frequently noted in subgingival polymicrobial biofilms, was obtained from adults with chronic periodontitis (Rams and van Winkelhoff, 2017; Deng et al., 2017) and reported to be associated with periapical periodontitis (Gomes et al.,1994; 1996; Jacinto et al., 2003). This bacterium has also been detected in association with other oral conditions, including endodontic infections (Fukushima et al., 1990; Baumgartner et al., 1999), acute necrotizing

ulcerative (Loesche et al., 1982), and pregnancy gingivitis (Raber-Durlacher et al., 1994). In addition, others have implicated that *Pre. intermedia* is closely related to systemic diseases, such as atherosclerosis, cardiovascular disease, and stroke (Scannapieco et al., 2003), as well as preterm birth (Offenbacher et al., 2001). Because of its versatile nature, *Prevotella* species have also been observed in the upper respiratory tract, urogenital tract (Eiring et al., 1998), and large intestine (Hayashi et al., 2007).

Peptidases produced by periodontopathic bacteria provide nutrients for bacterial growth, and are important etiologic reagents involved in degradation of periodontal tissues and alterations of the host immune system and bioactive peptides (Potempa et al., 2000). Notably, various studies regarding Arg- and Lys-specific gingipains of *Porphyromonas gingivalis* have been presented (Chen et al., 1992; Scott et al., 1993; Pike et al., 1994; Pavloff et al., 1995; 1997), and more recent reports of dipeptide-producing enzymes, known as dipeptidyl-peptidases (DPPs), and a novel acylpeptidyl oligopeptidase are considered to be essential for growth and pathogenicity of *P. gingivalis* (Nemoto and Ohara-Nemoto, 2016; Nemoto et al., 2016; Ohara-Nemoto et al., 2017). In contrast to gingipains solely expressed in *P. gingivalis*, most DPP genes and their activities can be found in several periodontopathic bacteria including *Pre. intermedia* (Ohara-Nemoto et al., 2018). However, little is known about peptidases in *Pre. intermedia* yet, though it has been shown that the level of proteolytic activity in clinical strains was significantly higher than that in commensal strains isolated from healthy subjects (Yanagisawa et al., 2006). *Pre. intermedia* efficiently incorporates single amino acids and produces ammonia (Shah and Williams, 1987; Takahashi and Sato, 2001), thus peptidases related to release single amino acids are considered to be important.

In the present study, we identified the C69-family member dipeptidase A in *Pre. intermedia* BAU17746. Enzymatic and biochemical properties of this peptidase were compared with two additional C69-family members from *Lactobacillus helveticus* and *Tannerella forsythia*. The present results demonstrated that current understanding regarding the substrate specificity of the C69 family should be revisited. Moreover, we here report that the C69-family dipeptidase A is a unique cysteine peptidase with an essential cysteine at the N-terminus.

Results

Exopeptidase activities of *Pre. intermedia*

Peptidase activities of *Pre. intermedia* OMA14 were determined with phosphate-buffered saline (PBS)-washed and resuspended cells by use of a series of fluorogenic peptidyl-4-methycoumaryl-7-amide (MCA) substrates (Figure 1). In our previous study, DPP4 from the bacterial strain has been cloned and characterized (Ohara-Nemoto et al., 2017, 2018). In addition, Phe-Met-MCA was established as a specific substrate for DPP7 (Nemoto et al., 2018), which enables determination of DPP7 activity separately from that of DPP5 in bacterial cells. Figure 1A indicated that the activities of DPP7 for Phe-Met-MCA and DPP4 for Gly-Pro-MCA were predominant, while some hydrolysis of Lys-Ala-MCA by DPP5 was detected. No hydrolysis of Leu-Asp- or Leu-Glu-MCA was observed, in accord with lack of the DPP11 gene in the genome (Naito et al., 2016).

Among 6 aminoacyl-MCA substrates (Arg, Lys, Leu, Met, Ala, Phe), predominant hydrolysis of Arg-MCA was observed, while other activities were markedly limited. In order to determine whether the Arg-MCA-hydrolyzing entity is mediated by either an exopeptidase or endopeptidase, hydrolyses of various MCA substrates carrying Arg at the P1 position were examined. As shown in Figure 1C, *Pre. intermedia* predominantly hydrolyzed Arg-MCA, whereas degradation of Arg-Arg-, *t*-butyloxycarbonyl-[(2*S*)-2-amino-3-(benzyloxycarbonyl) propionyl (Boc)-FSR-, and Boc-RVRR-MCA was limited, and hydrolysis of the remaining 5 substrates [benzoylglycyl (Bz)-Arg-, benzyloxycarbonyl (Z)-Leu-Leu-, glutalyl (Glt)-Gly-Arg-, Z-VVR-, pyroglutamyl (Pyr)-RTKR-MCA] was negligible. These results suggested that Arg-MCA-hydrolyzing activity was mediated by an Arg-specific aminopeptidase.

To characterize this hydrolyzing activity, we examined its biochemical properties. The activity was steeply increased in the presence of 3 – 30 mM dithiothreitol, then elevation continued in a linear manner until 300 mM (Figure 2A), suggesting that this entity is a cysteine peptidase. Cellular Arg-MCA hydrolysis reached a plateau at pH 7.5 in phosphate buffer, though was further increased when

using Tris-HCl buffer (Figure 2B), suggesting that an alkaline pH level caused disruption of the cells. The Arg-MCA-hydrolyzing activity was scarcely detected in the culture supernatant. In contrast, that activity was observed in bacterial cells, with greater activity noted in the fraction comprising the cytosol and periplasm in the presence of dithiothreitol (Figure 2C).

Expression and processing of BAU17746

BAU17746 and BAU18827 in *Pre. intermedia* OMA14 have been annotated as C69-family cysteine peptidases (Naito et al., 2016) and showed amino acid sequence identity of 21.5% and 14.7%, respectively, to *Streptococcus gordonii* Arg-aminopeptidase (MER002648) (Goldstein et al., 2002), the representative molecule of the C69.002 cysteine peptidase family. The amino acid identity between BAU17746 and BAU18827 was low (13.4%) (Table 2). SignalP-5.0 search (http://www.cbs.dtu.dk/services/SignalP/) indicated the presence of a signal peptide in BAU18827 ($M^{1}KKLLLAVLAVCSFGAAKA^{19}$), but not in BAU17746, indicating that BAU17746 is a cytosol protein, but that BAU18827 is secreted.

There are two major subfamilies in the C69 family, i.e., C69.001 represented by *L. helveticus* dipeptidase A (Dudley et al., 1996; Vesanto et al., 1996) and C69.002 represented by *S. gordonii* Arg-aminopeptidase. Although other members, i.e., exocytosis regulator secernin and *Drosophila melanogaster* CG10098 protein are classified in other subfamilies C69.003-C69.006, their enzymatic properties are unknown (Way et al., 2002).

To examine peptidase activities of BAU17746 and BAU18827, their full-length recombinant forms with C-terminal His_6 tag were expressed in *Escherichia coli.* Purified proteins were separated with SDS-PAGE as 54- and 60-kDa species, respectively, in accordance with the calculated molecular masses (Figure 4A). We found that BAU17746 exhibited activity toward Arg-MCA, whereas that from BAU18827 did not show any hydrolysis (Figure 4B). Since Cys^{26} of *S. gordonii* Arg-aminopeptidase is an essential residue (Goldstein et al., 2002), we speculated that the equivalent Cys^{15} of BAU17746 is essential for its activity. In fact, both an N-terminal truncated form (BAU17746ΔN1-35) and the C15A mutant (BAU17746C15A) completely lost Arg-MCA-hydrolyzing activity.

Recombinant BAU17746 revealed a faint 56-kDa band in addition to a major 54-kDa band (Figure 4A). N-terminal sequencing analysis demonstrated the N-terminus of the 54-kDa species as Cys^{15}. On the other hand, the 56-kDa species was shown to be mainly composed of two species starting at Met^{1} of BAU17746 and Gly^{-3} derived from pQE60, in addition to inevitable contamination of predominant 54-kDa species with the N-terminus of Cys^{15} (Figure 4C). Thus, these results suggest that the 54-kDa species from BAU17746 was a mature form with Cys^{15} as the N-terminus. Interestingly, the 54-kDa processed species was never produced in the BAU17746C15A.

This phenomenon was clearly reproduced on glutathione *S*-transferase (GST)-BAU17746 fusion protein and its C15A mutant. The 80-kDa GST-BAU17746 was recovered as a minor species in the wild type, while 30- and 54-kDa species were dominant. An immunoblotting against anti-GST antibody demonstrated that 30- and 80-kDa were GST proteins. Accordingly, the 54-kDa species corresponded to mature BAU17746 (Figure 5A). The specific activity (pmol/min/μg protein) for Arg-MCA of this preparation accounted for 30% of the wild type, which seemed to be reasonable, because it contains a considerable amount of GST and a minor unprocessed 80-kDa fusion form (Figure 5A). In contrast, GST-BAU17746C15A carrying no peptidase activity was solely recovered as an 80-kDa species, confirming BAU17746-mediated processing at the Glu^{14}-Cys^{15} bond.

We next determined the mechanism of the cleavage at the Glu^{14}-Cys^{15} bond. Even though the 80-kDa GST-BAU17746C15A was co-incubated with BAU17746 mature form at 37 ˚C for 5h, the degradation of the 80-kDa species was neither observed, the 30-kDa GST was nor produced, and no increase in the 54-kDa species was observed (Figure 5B). This finding strongly suggested that the cleavage at the Glu^{14}-Cys^{15} bond was not intermolecularly processed by a mature form, but intramolecularly mediated by a proenzyme itself.

Arg-MCA-hydrolyzing property of BAU17746

The effects of reducing agents and protease inhibitors were examined with both *Pre. intermedia* cells and recombinant BAU17746. Arg-specific aminopeptidase activities were significantly enhanced with dithiothreitol and 2-mercaptoethanol in a dose-dependent manner, though the efficiency of

dithiothreitol was superior, reflecting its higher reducing potential (Table 3). Moreover, the activities were completely inhibited by monoiodoacetic acid, a cysteine-blocking agent. To the contrary, the effects of serine and metallopeptidase inhibitors on the cellular activity and BAU17746 are limited. Taken together, these results confirmed that the Arg-MCA-hydrolyzing activity observed in *Pre. intermedia* is mediated by a cysteine peptidase, BAU17746.

Recombinant BAU17746 hydrolyzed Arg-MCA most efficiently, while hydrolysis of other aminoacyl- and peptidyl-MCAs was scarcely observed (Figure 6A, B). The activity was highly dependent on dithiothreitol concentration (Figure 6C), and the pH profile showed a peak at pH 8.5 in Tris-HCl buffer, while the activity itself was higher at pH 7.7 in sodium phosphate buffer (Figure 6D), which indicated that the amino group of tris-hydroxymethane aminomethane functions as a very weak inhibitor of Arg-specific peptidase. Accordingly, the peptidase activity of the recombinant forms was measured in 50 mM sodium phosphate buffer at pH 7.7 containing 60 mM dithiothreitol in the following experiments. The k_{cat} and K_m values of the peptidase for Arg-MCA were 0.012 s^{-1} and 1.67 μM, respectively, thus that of k_{cat}/K_m was 0.0074 s^{-1} μM^{-1}.

We also examined the molecular status of the peptidase using size exclusion HPLC (Figure 6F). BAU17746 was eluted at the 450-kDa position with a preceding shoulder. In contrast, the cellular form was split into minor and major peaks of 350 and 120 kDa, respectively. Taken together, the peptidase seemed to mainly be presented as a dimer in an endogenous form, while recombinant peptidase showed a tendency to form at 8 – 10-mers. We previously reported a similar phenomenon with acylpeptidyl oligopeptidase of *P. gingivalis* (Nemoto et al., 2016).

Property of BAU17746 as dipeptidase A

Our findings of partial hydrolysis of Arg-Arg-, Boc-FSR-, and Boc-RVRR-MCA were not reproduced by BAU17746 (compare Figures 1 and 6), suggesting the possible presence of multiple Arg-specific peptidases in *Pre. intermedia.* In addition, the finding that BAU17746 never showed hydrolysis of Arg-Arg-MCA raised doubt about its aminopeptidase activity, because Arg-specific aminopeptidase should sequentially degrade Arg-Arg-MCA into an arginine and Arg-MCA, and finally, into two arginines and 7-amino-4-methylcoumarin.

An updated BLAST search indicated a 26.6% identity of amino acid sequence of BAU17746 to that of *L. helveticus* dipeptidase A (MER0002163), a representative C69.001-subfamily dipeptidase A/PepDA (Table 2). In addition, we found that *T. forsythia* putative dipeptidase (MER0284871) possesses 74.9% identity to BAU17746, which recorded the highest value except for those from other strains of *Pre. intermedia*. These findings suggested that even though BAU17746 was suspected to be an aminopeptidase, it might be a dipeptidase belonging to the C69.001, of which Arg-MCA is accepted as a dipeptide.

Since Leu-Leu was reported to be the preferred substrate for *L. helveticus* dipeptidase A (Dudley et al., 1996; 2012), we examined the dipeptidase activity of BAU17746 using a modified cadmium-ninhydrin assay with the substrates Arg-Leu and Arg-Phe. Those results conclusively demonstrated that BAU17746 hydrolyzed these dipeptides, while its inert mutants (ΔN1-35 and C15A) did not (Figure 7). Maximum dipeptidase activity was achieved at pH 8.3 in sodium phosphate buffer and at pH 8 in Tris-HCl buffer (data not shown), identical to the pH profile for Arg-MCA hydrolysis. Noticeably, dipeptidase activity was higher toward Arg-Leu and Arg-Phe than Leu-Leu, followed by Glu-Glu. Furthermore, hydrolysis of the tripeptide Leu-Leu-Leu was not observed (Figure 7B). The dipeptide Arg-Xaa was much more preferential than Arg-MCA for BAU17746, since the specific activities (20 nmol/min/μg) toward Arg-Leu and Arg-Phe (Figure 7A) were at least 500-fold higher than V_{max} (0.041 nmol/min/μg) for Arg-MCA (Figure 6E), presumably due to a non-suitable structure at the MCA at the P1' position. The low k_{cat} and k_{cat}/K_m values for Arg-MCA may be related to these findings. Conclusively, BAU17746 does not encode an aminopeptidase, but is a dipeptidase belonging to the C69 dipeptidase A family.

P1-position specificity of C69.001-family dipeptidase A

C69.001 dipeptidase A is represented by the *L. helveticus* entity. It was previously reported that *L. helveticus* dipeptidase A shows broad preference, except for Pro, at both the P1 and P1' position residues, and that bulky hydrophobic residues, such as Leu-Leu, Phe-Leu (Dudley et al., 1996; 2012),

and Met-Leu (Vesanto et al., 1996), were found to be most efficiently hydrolyzed. Hence, the present results raise a question with respect to the P1-position (N-terminal) specificity. To address this issue, substrate specificity *L. helveticus* dipeptidase A as well as a putative dipeptidase member from the periodontopathic bacterium *T. forsythia* possessing the highest identity (74.9%) to BAU17746 was examined. Recombinant dipeptidase A from *L. helveticus* and the putative dipeptidase of *T. forsythia* migrated to 60 and 55 kDa, respectively, in accordance with their calculated molecular masses (Figure 8A). Despite a previously reported preference for bulky hydrophobic amino acids, our findings showed *L. helveticus* dipeptidase A to be preferential for P1 Arg as *Pre. intermedia* peptidase (Figure 8B). In contrast, the *T. forsythia* dipeptidase was not preferential for the P1-position Arg, but rather for Leu (Figure 8C). As expected, these dipeptidases did not hydrolyze the tripeptide Leu-Leu-Leu or dipeptidyl-MCA substrates (data not shown). Accordingly, these results indicate that the C69.001 dipeptidase A is comprised of at least two types of dipeptidases, one preferential for dipeptides with the P1-position Arg and the other preferential for those with hydrophobic P1 residues.

P1'-position specificity of C69-family dipeptidase A

A previous study demonstrated that hydrolysis of Met-Ala (433%), Leu-Gly (272%), Phe-Gly (169%), Ala-Phe (133%), Arg-Asp (105%), and Lys-Ala (103%) was superior to that of Leu-Leu (100%) (Vesanto et al., 1996). Hence, we investigated the C-terminal preference of C69.001 dipeptidases toward dipeptides with P1 Arg and P1' Leu, Phe, Met, Ala, Gly, Arg, Lys, and Asp. Among them, *Pre. intermedia* dipeptidase A most efficiently hydrolyzed Arg-Gly, Arg-Phe, and Arg-Leu, and the other dipeptides were moderately hydrolyzed, except for Arg-Lys (Figure 9A). Arg-Phe was the best substrate for *L. helveticus* dipeptidase A, with other dipeptides, except for Arg-Lys and Arg-Asp, found to be moderately hydrolyzed. A similar preference regarding Arg-Xaa dipeptides was shown by *T. forsythia* dipeptidase A, except that a lower level of efficacy was observed with Arg-Arg and Arg-Lys (Figure 9B). Importantly, Leu-Leu was again the most superior among the examined substrates as compared to P1 Arg-containing dipeptides for *T. forsythia* dipeptidase. These results indicate that (i) P1 preference, *i.e.*, Arg or Leu, is altered between the *Pre. intermedia* and *L. helveticus* dipeptidase A and *T. forsythia* entity; (ii) P1' Lys is commonly unfavorable, and Arg and Asp are occasionally acceptable, and (iii) non-charged residues at the P1' position are relatively preferable for the three peptidases. These observations suggest that P1'-position residues are not as strictly selected as compared to P1 position residues. Taken together, we concluded that three dipeptides, Arg-Leu, Arg-Phe, and Arg-Gly, are appropriate substrates for *Pre. intermedia* and *L. helveticus* dipeptidase A, and that Leu-Leu is a good substrate for *T. forsythia* dipeptidase.

Discussion

The present results demonstrated that dominant Arg-MCA hydrolysis shown in *Pre. intermedia* is mediated by Arg-specific dipeptidase BAU17746 that belongs to the C69-family dipeptidase A. *Pre. Intermedia* dipeptidase A is composed of 479 amino acid residues with the essential Cys[15]. In our experiments, it was found to preferentially hydrolyze Arg-Gly, Arg-Phe, and Arg-Leu, and other dipeptides with P1 Arg, except for Arg-Lys, were moderately hydrolyzed.

The P1 Arg specificity of BAU17746 is apparently incompatible with the P1 Leu preference of *L. helveticus* dipeptidase A (Dudley et al., 1996; 2012; Vesanto et al., 1996). However, the present investigation revealed that *L. helveticus* dipeptidase A also exhibits a P1 Arg preference and further indicated that the lack of Arg preference noted in the previous study was due to usage of Arg-Asp, as we found that P1'-position Asp as well as Lys were least preferable for *L. helveticus* dipeptidase A (Figure 9B). Our results further demonstrated that all molecules annotated as the C69.001 dipeptidase A are not specific for P1 Arg, because *T. forsythia* dipeptidase A shows a preference for P1 Leu. Hence, this subfamily comprises at least two dipeptidase members with distinct specificities, *i.e.*, Arg and Leu at the P1 position.

Pre. intermedia and *T. forsythia* each possess one dipeptidase A gene, while *L. helveticus* possesses 5 genes of the C69.001 subfamily (Rawlings et al., 2012), of which only one (MER0233043/PepDA) has been enzymatically and biochemically characterized (Dudley et al., 1996; Vesanto et al., 1996; this

study). Based on the present findings showing multiple P1 preferences of the C69.001 subfamily, characterization of other members is a topic of interest.

Sequence-based classification of the C69.001 members into P1 Arg- and Leu-preferential subtypes seems inappropriate, because the sequence identity between altered P1 preferential *Pre. intermedia* and *T. forsythia* dipeptidase A is much greater than that between Arg-preferential *Pre. intermedia* and *L. helveticus* dipeptidase A (Table 2). We previously reported that the alteration in substrate specificity between DPP7 specific for hydrophobic residues and DPP11 for acidic ones is primarily defined by a single amino acid residue composing the active center (Ohara-Nemoto et al., 2011; Rouf et al., 2013). Similarly, we attempted to find an essential amino acid residue to produce P1 Arg or Leu specificity of the dipeptidases, though no such residue in the C69.001 subfamily was found in our examinations to date.

The present study also demonstrated that P1'-position residues of dipeptidase A are not stringent. Gly, Phe, and Leu at the P1' position of dipeptides showed relatively higher potentials, while moderate potentials were shown by the other residues. The ambiguity of a P1' residue in substrates has been previously demonstrated in three-dimensional structure analysis of *P. gingivalis* DPP11, which strictly binds to Asp and Glu residues at the P1 position, though an interaction with the P1'-position residue was found to be relaxed (Bezerra et al., 2017). Results of the present study revealed that Arg-Gly, Arg-Phe, and Arg-Leu are the best substrates for *Pre. intermedia* dipeptidase A (BAU17746), and Arg-Phe the best for *L. helveticus* dipeptidase A.

Our study revealed the unique N-terminal location of the essential catalytic cysteine residue (Cys^{15}) in a mature form of BAU17746. This indicates cleavage of the proenzyme at the Glu^{14}-Cys^{15} bond, which apparently occurs via the intramolecular proteolysis within the polypeptide chain of the zymogen. In this respect the BAU17746 zymogenic form exerted endopeptidase activity on itself to release the mature form with the exclusive dipeptidase activity. We should notify that essential cysteine residues of the C69.001 subfamily are located near the N-terminus of the sequences at 2^{nd} to 26^{th} positions from Met^{1} (Rawlings et al., 2012), which indicates that this subfamily members commonly possess essential cysteines at the N-termini of the mature forms. In fact, we noticed the existence of a higher molecular mass species in addition to the major one of *T. forsythia* dipeptidase, when a large amount of proteins was loaded (data not shown). Furthermore, it has been reported that the N-terminus of purified *S. gordonii* Arg-aminopeptidase appeared to be modified, because N-terminal sequencing was not successful (Goldstein et al., 2002). However, it is well known that cysteine residue is not easily identified by N-terminal sequencing (Figure 4C), and thus, we suspect that Cys^{26} is located at the N-terminus in a mature form of *S. gordonii* Arg-aminopeptidase. If so, an essential cysteine of the C69.002 subfamily is also located at the N-terminus.

By comparison with Penicillin V acylase, Class II glutamine amidotransferases, Penicillin G acylase, and proteasome catalytic subunits, Pei and Grishin (2003) previously proposed that *L. helveticus* dipeptidase A, which used to be classified in MEROPS to peptidases with unknown catalytic type U34, belong to the superfamily of N-terminal nucleophile hydrolases and insisted that the essential residue should be located at the N-terminus beyond the classes of cysteine, serine, and threonine peptidases. The findings on this study truly confirmed their hypothesis.

It should noted that the processing of the Glu^{14}-Cys^{15} bond of BAU17746 proenzyme is intramolecularly mediated by an endopeptidase activity of the proenzyme, although the mature BAU17746 is a dipeptidase. Our preliminary data suggest that Glu^{14} is most suitable for the auto-processing, and Asp^{14} is the second-best residue for this process (M.T.S and T.K.N, unpublished observation). However, a dipeptide Glu-Glu is not an appropriate substrate of BAU17746 (Figure 7B). Furthermore, the amino acid sequences of *Pre. intermedia* and *T. forsythia* dipeptidase A at the processing site were Pro^{12}-Ser-Glu-Cys^{15} (numbering in BAU17746) and that of *L. helveticus* is Gln^{3}-Ser-Glu-Cys^{6}. Therefore, the P1-position specificity of the dipeptidases is distinct from the specificity of the processing site specificity. However, this may be not surprising that some of the members in the N-terminal nucleophilic hydrolase superfamily are not peptidases (Pei and Grishin, 2003).

There are two C69-family members, BAU17746 and BAU18827, in *Pre. intermedia*. Although we speculated that both may represent dipeptidases with altered specificity, our attempts to detect aminoacyl-MCA hydrolysis as well as dipeptide hydrolysis of BAU18827 have not been successful to date. Similarly, there are two C69-family members in *T. forsythia*, of which one (MER0284871) is dipeptidase A characterized in this study, while the other (MER0285668) is currently listed as an unassigned member. The latter shows a 43.7% amino acid identity to BAU18827, indicating their close

kinship. Furthermore, there is only a single C69-family unassigned peptidase in *P. gingivalis* (PGN_1103), which has 47.5% identity to BAU18827. Thus, orthologues of *Pre. intermedia* BAU18827 are commonly present in *T. forsythia* and *P. gingivalis*, as well as other species of the phylum *Bacteroidetes*, although their activity remains unknown. Our purification data indicated that BAU18827 appeared to be not processed at the Ala^{19}-Cys^{20} peptide bond (data not shown), which might be related to no activity of recombinant BAU18827.

Although previous studies have shown that *Pre. intermedia* produces various kinds of proteases (Suido et al., 1996; Holdeman et al., 1977; Deschner et al., 2003; Potempa et al., 2009; Potempa and Pike, 2009; Yano et al., 2009), little is known about DPPs except for DPP4 (Shibata et al., 2003; Ohara-Nemoto et al., 2017). In the present study, we confirmed expressions of DPP4, DPP5, and DPP7, whereas DPP11 activity was not detected (Figure 1A). In fact, BLAST search results demonstrated that the DPP11 gene is absent in *Pre. intermedia*. Because the DPP11 and DPP7 genes are distributed in several different anaerobic bacteria, such as those of the phylum *Bacteroidetes* (Rouf et al., 2013), we speculate that the ancestor of *Pre. intermedia* also possessed the DPP11 gene that may have been lost during its evolution. Nevertheless, it is important to note that even though deletion of only a single DPP gene may significantly reduce efficiency for complete metabolism of polypeptides into dipeptides, that may be not have a significant effect on growth of this species, because *Pre. intermedia* can utilize free amino acids as well as dipeptides, in contrast to *P. gingivalis*, which shows dominant dipeptide utilization.

Materials and methods

Materials

pQE60 was purchased from Qiagen Inc. (Chatsworth, CA). pGEX4T-1, a Superdex 200 Increase 10/300 GL column, glutathione-Sepahrose 4B, low-molecular-weight markers, and full-range rainbow molecular weight markers were obtained from GE Healthcare (Buchinghamshire, England). Restriction enzymes and DNA-modifying enzymes came from Takara Bio (Tokyo, Japan) and New England Biolabs (Ipswich, MA), while KOD-Plus-Neo DNA polymerase came from Toyobo (Tokyo, Japan). Aminoacyl- and peptidyl-methylcoumaryl-7-amide (MCA) were obtained from the Peptide Institute (Osaka, Japan), Arg-Leu, Arg-Phe, Leu-Leu, and Glu-Glu from Bachem (Bubendorf, Switzerland), and Leu-Leu-Leu from Sigma-Aldrich Japan (Tokyo, Japan). Phe-Met-MCA was synthesized by Scrum (Tokyo, Japan), while Arg-Arg, Arg-Lys, Arg-Met, Arg-Ala, Arg-Asp, and Arg-Gly were synthesized by Biologica Co. (Nagoya, Japan). A ninhydrin reagent was obtained from Fuji Film-Wako Pure Chemicals (Osaka, Japan).

Bacterial cells

Pre. intermedia OMA14 was isolated from the periodontal pocket of a Japanese patient with periodontitis (Fukushima, 1992). *Pre. intermedia* OMA14 was selected as the strain can be manipulated for gene transfer, after which the genome project was completed (Naito et al., 2016). Bacteria were grown in anaerobic bacterial culture medium broth (Eiken Chemical Co., Ltd, Tokyo Japan) at 37 ˚C under an anaerobic condition. In an early stationary phase, bacterial culture was centrifuged at 9000 x*g* for 15 min at 4˚C, and the supernatant filtrated with a 0.45-μm filter was saved as culture supernatant at -80˚C until use. Bacterial cells were washed with ice-cold phosphate buffered saline (PBS) at pH 7.4, then resuspended in PBS to absorbance at 600 nm of 2 and stored at - 80 ˚C. In some experiments, bacterial cells were suspended in 20 mM Tris-HCl (pH 8.0) buffer containing 0.3 mg/ml of lysozyme, 0.1 M NaCl, 10 μg/ml of leupeptin, and 3 mM dithiothreitol. After two times of freezing and thawing, the cytosol-periplasm fraction was obtained by centrifugation at 20,400 x*g* for 30 min at 4 ˚C.

Construction of plasmids

Pre. intermedia OMA14, *L. helveticus* ATCC 15009, and *Tannerella forsythia* ATCC 25611 genomic DNA samples were prepared as previously reported (Ikeda et al., 2004). *L. helveticus* ATCC 15009

was obtained from RIKEN BRC JCM (Tsukuba, Japan). A DNA fragment of BAU17746 located at the 1238 position of the first chromosome of *Pre. intermedia* OMA14 (locus tag PIOMA14_I_1238) (Naito et al., 2016) was amplified by PCR using KOD-Plus-Neo with genomic DNA and an appropriate set of primers (5BAU17746M1Bam, 3BAU17746A479Bam) (Table 1). The PCR fragment was digested by *Bam*HI, then cloned into the *Bam*HI site of pQE60 and GEX4T-1, as previously described (Ohara-Nemoto et al., 2011), with the resulting plasmid designated as pQE60-BAU17746 and pGEX4T-1-BAU17746, respectively. Similarly, the DNA fragment of BAU18827 located at the 322 position of the second chromosome (PIOMA14_II_0322), genes for *L. helveticus* dipeptidase A (MER0002163, KEGG entry: K08659, Uniprot accession: Q48558), and *T. forsythia* putative dipeptidase (MER0284871) were cloned into pQE60, and designated as pQE60-BAU18827, pQE60-LhpepDA, and pQE60TfpepD, respectively. pQE60 plasmids encoding the N-terminal 35-residue deleted form of BAU17746 (pQE60-BAU17746ΔN1-35) and an amino acid substitution from Cys^{15} to Ala (pQE60-BAU17746C15A and pGEX4T-1-BAU17746C15A) were prepared by PCR with pQ60-BAU17746 and pGEXT4T-1-BAU17746, respectively, using appropriate primer sets (Table 1), as previously described (Ohara-Nemoto et al., 2011).

Expression and purification of recombinant proteins

Escherichia coli XL-1 Blue transformed with the expression plasmids were cultured in Luria-Bertani broth containing 75 μg/ml of ampicillin. Following an overnight culture, proteins were expressed using 0.2 mM isopropyl-thiogalactopyranoside at 30 ˚C for 4 h. After harvesting and lysing, the bacterial cell lysate was obtained by centrifugation at 15,000 x*g* for 40 min at 4 ˚C and recombinant proteins were purified by Talon affinity chromatography or glutathione-Sepharose 4B, as previously reported (Ohara-Nemoto et al., 2011; Nemoto et al., 1994). Proteins were stored at - 80 ˚C until use.

Measurement of peptidase activity with fluorogenic substrates

Peptidase activity was determined using aminoacyl- and peptidyl-MCA, as previously reported (Ohara-Nemoto et al., 2011; 2014). Briefly, an enzyme reaction was started by addition of recombinant proteins (100 – 200 ng) or a cell suspension (10 μl of A_{600} = 2.0) to a reaction mixture (200 μl) composed of 50 mM sodium phosphate buffer (pH 7.7), 5 mM EDTA, 60 mM dithiothreitol, and 20 μM aminoacyl- or peptidyl-MCA. After 30 min at 37 ˚C, fluorescence intensity was measured with excitation at 380 nm and emission at 460 nm. For some experiments, pH was varied by use of 50 mM sodium phosphate (pH 5 – 8) or Tris-HCl (pH 7 – 10), and the concentration of dithiothreitol was also varied from 0 – 300 mM.

To determine enzymatic parameters, recombinant BAU17746 was incubated with 0 – 10 μM Arg-MCA. Data were analyzed using a nonlinear regression curve fitted to the Michaelis-Menten equation with the GraphPad Prism software package (GraphPad Software Inc., La Jolla, CA). Values were calculated from 4 independent measurements.

Dipeptidase activity determination by ninhydrin assay

Dipeptidase activity was determined with 0.3 mM di- and tripeptides in a reaction mixture (200 μl) using a modified cadmium-ninhydrin method (Doi et al.,1981). BAU17746 and *T. forsythia* dipeptidase were incubated in a reaction mixture composed of 50 mM sodium phosphate (pH 8.3) and 10 mM dithiothreitol at 37 ˚C, while *L. helveticus* dipeptidase A was incubated in 50 mM MES (pH 6.0) at 55 ˚C with 10 mM dithiothreitol. After 30 min, 400 μl of cadmium-ninhydrin solution was added and the mixture was heated at 80 ˚C for 10 min. After cooling, absorbance at 508 nm was determined. Leucine was used as the standard for the ninhydrin reaction.

Size exclusion HPLC

Recombinant protein or bacterial lysate was subjected to size exclusion HPLC using an ÄKTA explorer 10S (GE Healthcare) with a Superdex 200 Increase 10/300 GL column (1.5 x 30 cm) equilibrated with 20 mM Tris-HCl (pH 8.0) and 50 mM NaCl. Samples were eluted at a rate of 0.5 ml/min at room temperature, with 0.5-ml fractions collected. Aliquots of the fractions were subjected to a peptidase

assay with Arg-MCA and SDS-PAGE.

SDS-PAGE, immunoblotting, and N-terminal sequencing

Purified proteins were separated by PAGE in the presence of 0.1% (w/v) of SDS with a polyacrylamide concentration of 10% (w/v), and then stained with Coomassie brilliant blue R250 (CBB or transferred to a polyvinylidene difluoride membrane (Merck-Millipore, Darmstadt, Germany). Immunoblotting was performed by using the anti-GST antiserum ($2x10^3$-fold dilution), and blots were visualized with alkaline phosphatase-conjugated anti-rabbit Ig(G+A+M) using 5-bromo-4-chloro-3-indolyl phosphate and nitro blue tetrazolium (Promega). An anti-GST antiserum was developed with recombinant GST as an antigen in rabbits with the method reported previously (Ohara-Nemoto et al., 2011; 2014). For the N-terminal sequencing, separated proteins were transferred to a Sequi-Blot membrane (Bio-Rad Laboratories, Hercules, CA), stained with CBB. The N-terminal sequences of CBB-stained bands were determined in Protein Research Institute (Osaka, Japan).

Protein concentration

Protein concentration was determined with a protein assay method (Bio-Rad), with bovine serum albumin used as the standard.

Acknowledgments

This study was supported by JSPS KAKENHI grants (JP19K10045 to T.K.N. and JP19K10071 to Y.O.-N.). M.T.S. is a graduate student supported by a Japanese Government (MEXT) Scholarship.

References

Baumgartner, J.C., Watkins, B.J., Bae, K.S., and Xia, T. (1999). Association of black-pigmented bacteria with endodontic infections, J. Endod. *25,* 413–415.

Bezerra, G.A., Ohara-Nemoto, Y., Cornaciu, I., Fedosyuk, S., Hoffmann, G., Round, A., Márquez, J.A., Nemoto, T.K., and Djinović-Carugo, K. (2017). Bacterial protease uses distinct thermodynamic signatures for substrate recognition. *Sci.* Rep. *7*, 2848 doi: 10.1038/s41598-017-03220-y.

Chen, Z., Potempa, J., Polanowski, A., Wikstrom, M., and Travis, J. (1992). Purification and characterization of a 50-kDa cysteine proteinase (gingipain) from *Porphyromonas gingivalis*. J. Biol. Chem. *267,*18896−188901.

Deng, Z.L., Szafrański, S.P., Jarek, M., Bhuju, S., and Wagner-Döbler, I. (2017). Dysbiosis in chronic periodontitis: key microbial players and interactions with the human host, Sci. Rep. *7*, 3703.

Deschner, J., Singhal, A., Long, P., Liu, C.C., Piesco, N., and Agarwal, S. (2003). Cleavage of CD14 and LBP by a protease from *Prevotella intermedia*, Arch. Microbiol. *179*, 430–436.

Doi, E., Shibata, D., and Matoba, T. (1981). Modified colorometric ninhydrin methods for peptidase assay, Anal. Biochem. *118*, 173-184.

Dudley, E.G., Husgen, A.C., He, W., and Steele, J.L. (1996). Sequencing, Distribution, and Inactivation of the Dipeptidase A Gene (*pepDA*) from *Lactobacillus helveticus* CNRZ32, J. Bacteriol. *178*, 701–704.

Dudley, E.G. and Steele, J.L. (2012). Dipeptidase DA. In Handbook of Proteolytic Enzymes, *3* edn (Rawlings, N.D. and Salvesen, G. S. eds), p3657-p3659, Elsevier, Amsterdam.

Eiring, P., Waller, K., Widmann, A., and Werner, H. (1998). Fibronectin and laminin binding of urogenital and oral prevotella species*,* Zentralbl, Bakteriol. *288*, 361–372.

Friedrich, V., Janesch, B., Windwarder, M., Maresch, D., Braun, M.L., Megson, Z. A., Vinogradov, E., Goneau, M.F., Sharma, A., Altmann, F., Messner, P., Schoenhofen, I. C., and Schäffer, C. (2017). *Tannerella forsythia* strains display different cell-surface nonulosonic acids: biosynthetic pathway characterization and first insight into biological implications. Glycobiology *27*, 342-357. doi: 10.1093/glycob/cww129.

Fukushima, H., Yamamoto, K., Hirohata, K., Sagawa, H., Leung, K.P., and Walker, C.B. (1990). Localization and identification of root canal bacteria in clinically asymptomatic periapical pathosis, J. Endod. *16*, 534–538.

Fukushima, H. (1992). Phenotypic characteristics and DNA relatedness in *Prevotella intermedia* and similar organisms, Oral Microbiol. Immunol. *7*, 60-64.

Goldstein, J.M., Nelson, D., Kordula, T., Mayo, J.A., and Travis, J. (2002). Extracellular Arginine Aminopeptidase from *Streptococcus gordonii* FSS2, Infect. Immun. *70*, 836–843.

Gomes, B.P., Drucker, D.B., and Lilley, J.D. (1994). Associations of specific bacteria with some endodontic signs and symptoms, Int. Endod. J. *27***,** 291–298.

Gomes, B.P., Lilley, J.D., and Drucker, D.B. (1996). Associations of endodontic symptoms and signs with particular combinations of specific bacteria, Int. Endod. J. *29,* 69–75.

Hayashi, H., Shibata, K., Sakamoto, M., Tomita, S., and Benno, Y. (2007). *Prevotella copri* sp. nov. and *Prevotella stercorea* sp. nov., isolated from human feces, Int. J. Syst. Evol. Microbiol. *57*, 941–946.

Holdeman, L.V., Cato, E.P., and Moore, W.E (1977). *Anaerobic laboratory manual. 4*, Blacksburg, V.A.: Virginia Polytechnic Institute and State University.

Ikeda, Y., Ohara-Nemoto, Y., Kimura, S., Ishibashi, K., and Kikuchi, K. (2004). PCR-based identification of *Staphylococcus epidermidis* targeting *gseA* encoding the glutamic-acid-specific protease, Can. J. Microbiol. *50*, 493–498.

Jacinto, R.C., Gomes, B.P., Ferraz, C.C., Zaia, A.A., and Filho, F.J. (2003). Microbiological analysis of infected root canals from symptomatic and asymptomatic teeth with periapical periodontitis and the antimicrobial susceptibility of some isolated anaerobic bacteria, Oral Microbiol. Immunol. *18*, 285–292.

Loesche, W.J., Syed, S.A., Laughon, B.E., and Stoll, J. (1982) The bacteriology of acute necrotizing ulcerative gingivitis, J. Periodontol. *53*, 223–230.

Naito, M., Ogura, Y., Itoh, T., Shoji, M., Okamoto, M., Hayashi, T., and Nakayama, K. (2016). The complete genome sequencing of *Prevotella intermedia* strain OMA14 and a subsequent fine-scale, intra-species genomic comparison reveals an unusual amplification of conjugative and mobile

transposons and identify a novel *Prevotella*-lineage-specific repeat, DNA Res. *23*, 11–19.
Nemoto, T., Ohara-Nemoto, Y., Shimazaki, S., and Ota, M., (1994). Dimerization characteristics of the DNA- and steroid-binding domains of the androgen receptor, J. Steroid Biochem. *50* 225-233.
Nemoto, T.K. and Ohara-Nemoto, Y. (2016). Exopeptidases and gingipains in *Porphyromonas gingivalis* as prerequisites for its amino acid metabolism, Jpn. Dental Sci. Rev. *52*, 22–29.
Nemoto, T.K., Ohara- Nemoto, Y., Bezerra, G.A., Shimoyama, Y., and Kimura, S. (2016). A *Porphyromonas gingivalis* periplasmic novel exopeptidase, acylpeptidyl oligopeptidase, releases N-acylated di- and tri-peptides from oligopeptides, J. Biol. Chem. *291*, 5913-5925.
Nemoto, T.K., Ono, T., and Ohara-Nemoto, Y. (2018). Establishment of potent and specific synthetic substrate for dipeptidyl-peptidase 7, Anal. Biochem. *548*, 78-81.
Offenbacher, S., Lieff, S., Boggess, K.A., Murtha, A.P., Madianos, P. N., Champagne, C.M., McKaig, R.G., Jared, H.L., Mauriello, S.M., Auten Jr, R.L., Herbert, W.N., and Beck, J.D. (2001). Maternal periodontitis and prematurity. Part I: obstetric outcome of prematurity and growth restriction, Ann. Periodontol. *6*, 164–174.
Ohara-Nemoto, Y., Shimoyama, Y., Kimura, S., Kon, A., Haraga, H., Ono, T., and Nemoto, T. K. (2011). Asp- and Glu-specific novel dipeptidyl peptidase 11 of *Porphyromonas gingivalis* ensures utilization of proteinaceous energy sources, J. Biol. Chem. *286*, 38115–38127.
Ohara-Nemoto, Y., Rouf, S.M., Naito, M., Yanase, A., Tetsuo, F., Ono, T., Kobayakawa, T., Shimoyama, Y., Kimura, S., Nakayama, K., Saiki, K., Konishi, K., and Nemoto, T.K. (2014). Identification and characterization of prokaryotic dipeptidyl-peptidase 5 from *Porphyromonas gingivalis*, J. Biol. Chem. *289*, 5436-5448.
Ohara-Nemoto, Y., Nakasato, M., Shimoyama, Y., Baba, T.T., Kobayakawa, T., Ono, T., Yaegashi, T., Kimura, S., and Nemoto, T.K. (2017). Degradation of incretins and modulation of blood glucose levels by periodontopathic bacterial dipeptidyl peptidase 4, Infect. Immun. *85* pii: e00277-17 doi: 10.1128/IAI.00277-17.
Ohara-Nemoto, Y., Shimoyama, Y., Nakasato, N., Nishimata, H., Ishikawa, T., Sasaki, M., Kimura, S., and Nemoto, T.K. (2018). Distribution of dipeptidyl peptidase (DPP) 4, DPP5, DPP7, and DPP11 in human oral microbiota-potent biomarkers indicating presence of periodontopathic bacteria, FEMS Microbiol. Lett. *365* doi: 10.1093/femsle/fny221.
Pavloff, N., Potempa, J., Pike, R.N., Prochazka, V., Kiefer, M.C., Travis, J., and Barr, P.J. (1995) Molecular cloning and structural characterization of the Arg-gingipain proteinase of *Porphyromonas gingivalis*. Biosynthesis as a proteinase-adhesin polyprotein. J. Biol. Chem. *270*, 1007-1010.
Pavloff, N., Pemberton, P.A., Potempa, J., Chen, W.C., Pike, R.N., Prochazka, V., Kiefer, M.C., Travis, J., and Barr P.J. (1997) Molecular cloning and characterization of Porphyromonas gingivalis lysine-specific gingipain. A new member of an emerging family of pathogenic bacterial cysteine proteinases. J Biol Chem. *272* 1595-1600.
Pei, J. and Grishin, N.V. (2003). Peptidase family U34 belongs to the superfamily of N-terminal nucleophile hydrolyases, Potein Sci. *12*, 1131-1135.
Pike, R., McGraw, W., Potempa, J., and Travis, J. (1994). Lysine- and arginine-specific proteinases from *Porphyromonas gingivalis.* Isolation, characterization, and evidence for the existence of complexes with hemagglutinins. J. Biol. Chem. *269* 406-411.
Potempa, J., Banbula, A., and Travis, J. (2000). Role of bacterial proteinases in matrix destruction and modulation of host responses, Periodontol. 2000 *24*, 153-192.
Potempa, M., Potempa, J., Kantyka, T., Nguyen, K. A., Wawrzonek, K., Manandhar, S.P., Popadiak, K., Riesbeck, K., Eick, S., and Blom, A.M. (2009). Interpain A, a cysteine proteinase from *Prevotella intermedia*, inhibits complement by degrading complement factor C3, PLoS Pathog. *5*. e1000316. doi: 10.1371/journal.ppat.1000316.
Potempa, J. and Pike, R.N. (2009). Corruption of innate immunity by bacterial proteases, J. Innate Immun.*1*, 70-87.
Raber-Durlacher, J.E.,Van Steenbergen, T.J., Van der Velden, U., De Graaff, J., and Abraham-Inpijn, L. (1994). Experimental gingivitis during pregnancy and post-partum: clinical, endocrinological, and microbiological aspects, J. Clin. Periodontol. *21*,549–558.
Rams, T.E. and van Winkelhoff, A.J. (2017). Introduction to clinical microbiology for the general dentist, Dent. Clin. North Am. *61***,** 179-197.

Rawlings, N.D., Barrett, A.J., and Bateman, A. (2012). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors, Nucleic Acids Res. *40* D343-D350 doi: 10.1093/nar/gkr987.

Rouf, S.M.A., Ohara-Nemoto, Y., Hoshino, T., Fujiwara, T., Ono, T., and Nemoto, T.K. (2013). Discrimination based on Gly and Arg/Ser at position 673 between dipeptidyl-peptidase (DPP) 7 and DPP11, widely distributed DPPs in pathogenic and environmental gram-negative bacteria, Biochimie *95*, 824-832.

Scannapieco, F.A., Bush, R.B., and Paju, S. (2003). Associations between periodontal disease and risk for atherosclerosis, cardiovascular disease, and stroke, A systematic review, Ann. Periodontol. *8*, 38–53.

Scott, C.F., Whitaker, E.J., Hammond, B.F., and Colman, R.W. (1993). Purification and characterization of a potent 70-kDa thiol lysyl-proteinase (Lys-gingivain) from *Porphyromonas gingivalis* that cleaves kininogens and fibrinogen, J. Biol. Chem. *268*, 7935-7942.

Shah, H.N. and Williams, R.A.D. (1987). Utilization of glucose and amino acids by *Bacteroides intermedius* ad *Bacteroides gingivalis*, Curr. Microbiol. *15*, 241–246.

Shibata, Y., Miwa, Y., Hirai, K., and Fujimura, S. (2003). Purification and partial characterization of a dipeptidyl peptidase from *Prevotella intermedia*, Oral Microbiol. Immunol. *18*, 196–198.

Suido, H., Nakamura, M., Mashimo, P.A., Zambon, J.J., and Genco, R.J. (1986). Arylaminopeptidase activities of oral bacteria. J. Dent. Res. *65,* 1336-1340.

Takahashi, N., and Sato, T. (2001). Preferential utilization of dipeptides by *Porphyromonas gingivalis,* J. Dent. Res. *80*, 1425-1429.

Vesanto, E., Peltoniemi, K., Purtsi, T., and Steele, J.L. (1996). Molecular characterization, over-expression and purification of a novel dipeptidase from *Lactobacillus helveticus*, Appl. Microbiol. Biotech. *45*, 638-645.

Way, G., Morrice, N., Smythe, C., and O'Sullivan, A.J. (2002). Purification and identification of secernin, a novel cytosolic protein that regulates exocytosis in mast cells, Mol. Cell. Biol. *13*, 3344-3354.

Yanagisawa, M., Kuriyama, T., Williams, D.W., Nakagawa, K., and Karasawa, T. (2006). Proteinase activity of prevotella species associated with oral purulent infection, Curr. Microbiol. *52*, 375–378.

Yano, T., Fukamachi, H., Yamamoto, M., and Igarashi, T. (2009). Characterization of L-cysteine desulfhydrase from *Prevotella intermedia*, Oral Microbiol. Immunol. 24, 485–492.

Table 1 Primers used for expression and *in vitro* mutagenesis.

Peptidase	Original plasmid	Primers	Sequence (5'-3')
BAU17746	pQE60	5BAU188277M1Bam	ACAAATGGATCCATGAAAGAAAATAAAGTT
		3BAU188277A479Bam	CGGCAGGGATCCTGCGCCGTGGAAAAGATA
BAU18827	pQE60	5BAU18827M1Bam	AATTATGGATCCATGAAGAAACTACTTTTA
		3BAU18827E546Bam	TTTAAGGGATCCTTCTGCAGGAACTTCAAA
BAU17746C15A	pQE60	5BAU17746C15A	*GCA*ACCACGATGATTGTGGGACAGGAAATG
		3BAU17746E14	TTCTGATGGTAAATTGCTTACCAAAAC
BAU17746Δ1-35	pQE60	5BAU17746E36Bam	GTGGCAGGATCCGAAGACTGGGACGCCATG
		3BAU17746A479Bam	CGGCAGGGATCCTGCGCCGTGGAAAAGATA
GST-BAU17746	pGEXT4T-1	5BAU188277M1Bam	ACAAATGGATCCATGAAAGAAAATAAAGTT
		3BAU188277A479Bam	CGGCAGGGATCCTGCGCCGTGGAAAAGATA
GST-BAU17746C15A	pQEX4T-1	5BAU17746C15A	*GCA*ACCACGATGATTGTGGGACAGGAAATG
		3BAU17746E14	TTCTGATGGTAAATTGCTTACCAAAAC
Lh dipeptidase A	pQE60	5LhDipepM1	ATGAAACAAACAGAATGTACTACTATCTT
		3LhDipepD474	GTCGAGCAAGTCGTACTTCA
Tf dipeptidase A	pQE60	5TFDipepAM1	TTTTGGATCCATGAATAGTAAAAAGATTAA
		3TfDipepAA481	GCTATGGGATCCCGCTCCGGCAAAGC

Restriction sites are underlined and mutated nucleotides are written in italics.

Table 2 Peptidases expressed and extensively discussed in this study.

Species and Strain	Genome	Locus tag	Protein	Family[a]	Reference	Amino acid identity (%)			
						PIOMA14_II_0322	NCTC9124_00164	Ihe_1838	BFO_RS06995
Pre. intermedia OMA14	AP014597	PIOMA14_I_1238	BAU17746/ Dipeptidase/ C69	C69.001	Naito et al., 2016	13.4	21.5	26.6	74.9
	AP014598	PIOMA14_II_0322	BAU18827/ Dipeptidase/ C69	C69 unassigned	Naito et al., 2016	100	14.7	12.0	12.9
S. gordonii NCTC9124	LR594041	NCTC9124_00164	Arg-aminopeptidase (RAP)/ abpB	C69.002	Golstein et al., 2002		100	18.4	20.2
L. helveticus CNRZ32	CP002081	lhe_1838	Peptidase DA/ PepDA	C69.001	Dudley et al., 1996; Vesanto et al., 1996			100	24.5
T. forsythia 92A2	CP003191	BFO_RS06995	C69	C69.001	Friedrich et al., 2017				100

[a] Classification by MEROPS

Table 3 Effects of peptidase inhibitors and activators on Arg-MCA hydrolysis of *Pre. intermedia* cells and recombinant BAU17746.

Reagent	Concentration	Cells (%)	BAU17746 (%)
-	-	100.0	100.0
Dithiothreitol	10 mM	1196.1	4080.6
	30 mM	1520.9	5690.9
2-Mercaptoethanol	10 mM	133.9	309.8
	30 mM	368.4	1832.1
Iodoacetic acid	5 mM	0	1.93
TLCK	50 μM	46.7	52.9
TPCK	85 μM	79.1	116.7
E-64	3 μM	83.8	94.6
Pefabloc	4 mM	46.9	64.3
Leupeptin	10 μM	38.4	84.03
PMSF	1 mM	112.5	27.04
EDTA	1 mM	196.6	152.4
EGTA	1 mM	225.4	157.5

One hundred activities of *Pre. intermedia* cells and recombinant protein without any additional reagents were set to 100%.

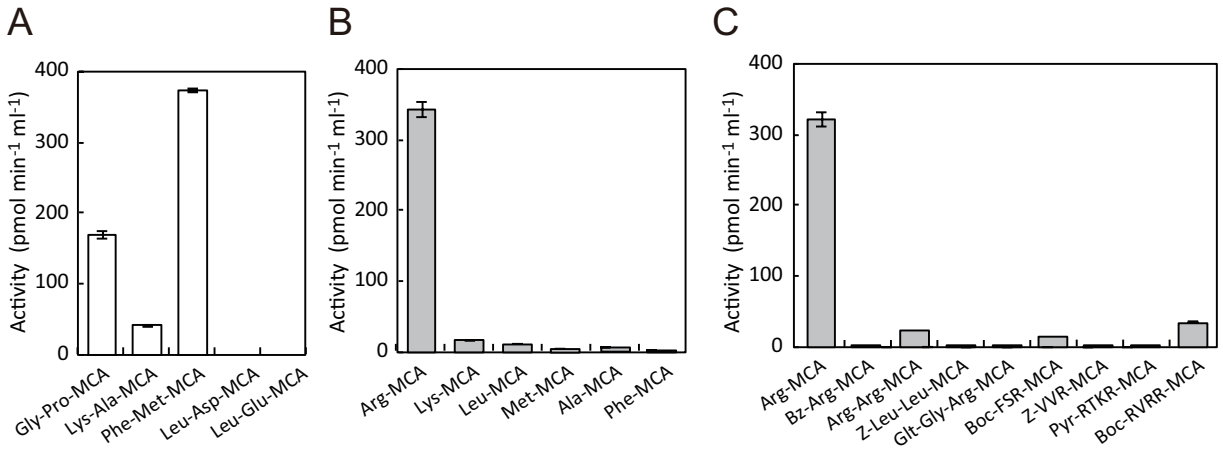


Figure 1 Exopeptidase activities of *Pre. intermedia.*
(A) Dipeptidyl-peptidase, (B) aminopeptidase, and (C) arginine-specific peptidase activities of cells were determined with MCA substrates. *Pre. intermedia* cell suspensions (10 µl) with $A_{600\ nm}$ = 2 were used and the activity is presented per ml of cells. Values are shown as the mean ± S.D. (n = 3).

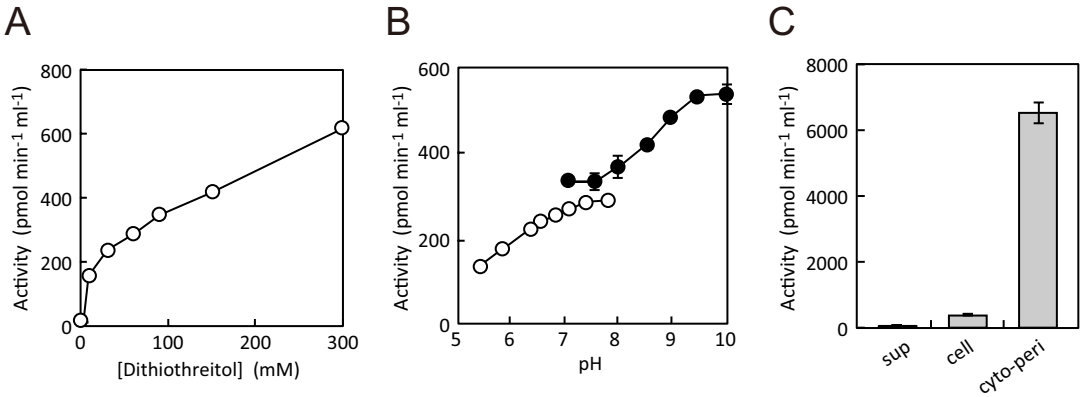
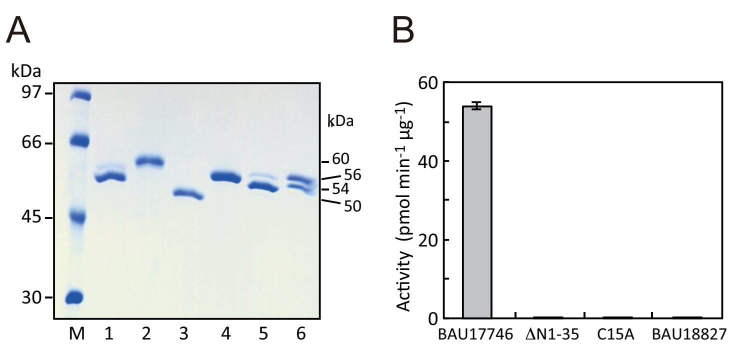


Figure 2 Effects of dithiothreitol, pH profile, and subcellular localization of Arg-MCA hydrolyzing activity.
(A) Arg-MCA-hydrolyzing activity of *Pre. intermedia* cells was determined in the presence of 0 – 300 mM dithiothreitol. (B) Arg-MCA-hydrolyzing activity of the cells was determined at pH 5 – 10 in sodium phosphate (open circle) and Tris-HCl (closed circle) buffers. (C) Arg-MCA-hydrolyzing activity was determined using culture supernatant (sup), cell suspension (cell), and cytosol-periplasm fraction (cyto-peri) samples. Values are shown as the mean ± S.D. (n = 3).

```
                                    56 kDa        54 kDa
Pi BAU17746                 1 -----------MKENKVLVSNLPSECTTMIVGQEMTSDGSMMVARSEDWDAMEAKN------YEIYEDTTEGATEYVAKDSPFRCELPKEALGYSALAPYCLPN- 87
Tf dipeptidase A            1 ---------MNSKKINEKGSIYPSECTTVIVGNKMTDDGSMIVARSEDWNTMFAKN------LEIYEDTAEGPETFVARDSAFRCELPREALGYTALAPYHLPG- 89
Lh dipeptidase A            1 -------------------MKQTECTTILVGKKASIDGSTMIARSEDGGRVIIPEGF---KVVNPEDQPKHYTSVISKQKIDDEDLAETPLRYTSAPDVSGKNG 82
Sg Arg-aminopeptidase       1 MKKILLRLMFLAMLVALLPVHVAQACSAFIVGKDLTADGSTLFGRTEDYPYAPDGGRHNQNYVVVPAKTYKDGDKIEDESNGFTYPHLANEMKYTAVYDSDRDNG 105
                                                       *::.:**:. : *** :..*:**                   .  :       ...           : *::       .

Pi BAU17746                88 ---HWGSAGFNTAGVGMSATESIFSSDEVLKYDPLVEN---GIAENAVFNVVLPYIRTAREGVERLGMLIEKHGIAEGFGIGFVDNKEIWYLETACGHRWLACRM 186
Tf dipeptidase A           90 ---HWGSAGFNTAGVGMSATESIFSNEKALQADPLVED---GVAENSVFNIVLPYVRTAREGVRRLGELIERHGIAEGFGIGFVDDSEIWYLETACGHRWLACRM 188
Lh dipeptidase A           83 ---IWGAAGINADNVAMTATETITTNSRIQGVDPILDPSEGGLGEEDFVTLTLPYLHSAFDGVKRVGYLVEKYGTYEMNGMAFSDKDNIWYLETIGGHHWIARRI 184
Sg Arg-aminopeptidase     106 SNGNFAAHGFNELGVAMTATVSATPNDKVLKEDPLVKD---GLPEASLVDLALPRAKTAREVIETVAKVLDEKGSAEGNIIVAADKNELWYMEILSGHQYVAIKF 207
                                             .*:** :   ....     **::.    *: *   .. :.**   ::* : :. :. ::::. *  *   :   *..::**:*   **:::* ::

Pi BAU17746               187 PKDQYFVTGNQSRFRLYDPLDTENYMASNDLLEFAEKHGLYDPKKGDFDFHEAYARDIKLDTTYNYPRVWGLQKMFSPSIKNDVTKNTFPVYAKAEKPISIADMR 291
Tf dipeptidase A          189 PEDKYFVTGNQSRFRKYDPNDTEHYMASADLIEFAREHGLYDPAKGEFDFHEAYARDIELDTTYNYPRVWGLQAMFSPQIKNDVARNTFPVFAEASKPLSLDDIR 293
Lh dipeptidase A          185 PDDAYVIAPNRLNIDTFDFDDSENFATASDLKDLIDEYHLN-PDREGYNMRHIFGSSTIKDAHYNNPRAWYIHNYFDPDFGGTPADQDQPFICRANRLISIEDIK 288
Sg Arg-aminopeptidase     208 PTNKYAVFANTYYLGHVDLNDKENVIASKDVEEVAKKADNYKTDKD-GNFMIAKSYGPDKYMERNRSRTYAGIKWMDPQAKVNYDDEAFDLLREPTDPNKKYTVH 311
                              * : * :  *   :    *  *.*:  :: *: :.  :     . :   ::     . .          * .*.:      :.*.         :   .  ..   .  ::

Pi BAU17746               292 HAFRFHYDNTEHDPYLHSNPKEP----------YRPVSIFRTTQTHILQVRPELPQAIGCVNYVAMGMADLGLFLPLYQGITSYPEAYTKGTDKASVDSAYWKFR 386
Tf dipeptidase A          294 RAFRFHYQGTEHDPYLHRNPKEK----------YRPVSIFRTTQTHILHVRPELPQAIGRVDYMSMGMAVLGVFLPLYQGITSYPEAYTKGTNRSSADSAYWKFR 388
Lh dipeptidase A          289 WAESSHYQDTPYDAYGDQGTPEQK-------KTFRPIGINRNFETHILQIRNDVPAEIAGVQWLAFGPNTFNSMLPFYTNVTTTPEAWQT-TPKFNLNKIFWLNK 385
Sg Arg-aminopeptidase     312 DVIAEQRNRFEHLPEYKADDLVEVGKKIDTNVYKYALGNENVIDAHVYQIKPNLPNAFGGVMWLGLAQSRNTPYVPFYGNVEDTYEAFKNRSTKYDGKSWYWTVW 416
                               .   : :   : .  . .                  .:.  .  ::*: ::: ::*  :. * ::.:.       :*:* .:    **: . : :  . .. :*

Pi BAU17746               387 KVMTLGMVNYNKYAPIIKEAYLKFEAETDQRQREMEEEYLRIYKTQPLRAKDMLQEFSDKILTKGMELADRLQEELFTQLTKDIQAEYLFHGA------------ 479
Tf dipeptidase A          389 KVQTLGMVDFNRYAPLIQSTYSRFEAETAQRQCEMEEQYLQIYKAQPIRARELLQSFSDKMLEAALDVTDKLTEKLFTCLAEDIQAEYRFAGA------------ 481
Lh dipeptidase A          336 LTAQLGDTNYRVYGELEDAFEQKSLAQCHKIQHETDKEVKNLSGKE---LQDKLIAANQKMSDTVYNNTVELLGQMVDEGHGLMTLKYDLLD------------- 474
Sg Arg-aminopeptidase     417 HIDQMVMKYPEIFGNSIQEKWKEKEAEWDKEQTERDAKYANYTDEQAKAAGPEVTKEALERSEKIFKEIKAVEAEMEEKIKKEKGQDADLVYTGYNKANLLADAK 521
                                  :     . :.    .     .  *:   : * * : :   .     :       :      :      .    :  ::             .  :

Pi BAU17746                   ---------------------------------------- 479
Tf dipeptidase A              ---------------------------------------- 481
Lh dipeptidase A              ---------------------------------------- 474
Sg Arg-aminopeptidase     522 ENSETKKSDSNQYGLWIVIAIVVVVGGFIAFKQFSNKKED 561
```

Figure 3 Alignment of amino acid sequences of *Pre. intermedia* BAU17746, *T. forsythia* putative dipeptidase, *L. helveticus* dipeptidase A, and *S. gordonii* Arg-aminopeptidase.
Amino acid sequences of *Pre. intermedia* BAU17746 (Pi BAU17746), *T. forsythia* putative dipeptidase (Tf dipeptidase A), *L. helveticus* dipeptidase A (Lh dipeptidase A), and *S. gordonii* Arg-aminopeptidase (Sg Arg-aminopeptidase) were compared. Identical amino acids are marked by an asterisk, and strongly and weakly similar amino acids are marked by colons and dots, respectively. Cysteines (red letters) essential for the activity are boxed (see text). N-terminal amino acid sequences were determined by sequencing of 54- and 56-kDa forms of BAU17746 are indicated by arrows. Thirty-five truncated residues in BAU17746ΔN1-35 are underlined with a broken line.



C

	Observed amino acids						Deduced sequence
Species	1	2	3	4	5	6	
54 kDa	X	Thr	Thr	Met	Ile	Val	Cys[15]-Thr-Thr-Met-Ile-Val[20]
56 kDa	Gly	Gly	Ser	Met	Lys	Glu	gly[-3]-gly-ser-Met-Lys-Glu[3]
	Met	Lys	Glu	Asn	Ile	Val	Met[1]-Lys-Glu-Asn-Ile-Val[6]
	(X)	Thr	Thr	(Met)	(Ile)	(Val)	Cys[15]-Thr-Thr-Met-Ile-Val[20]
	Ser	Ser			Glu		unidentified

Figure 4 Expression and activity of recombinant proteins.
(A) Recombinant proteins (1 µg) of BAU17746 (lanes 1 and 5), BAU18827 (lane 2), BAU17746ΔN1-35 (lane 3), BAU17746C15A (lane 4), and a mixture (0.5 µg each) of BAU17746 and BAU17746C15A (lane 6) were separated with SDS-PAGE. Lane M, low-molecular-weight marker. (B) Peptidase activities of BAU17746, BAU17746ΔN1-35, BAU17746C15A, and BAU18827 were determined with Arg-MCA. Values are shown as the mean ± S.D. (n = 3). (C) N-terminal sequencing of 54- and 56-kDa species of BAU17746. Vector-derived amino acids are shown with small letters.

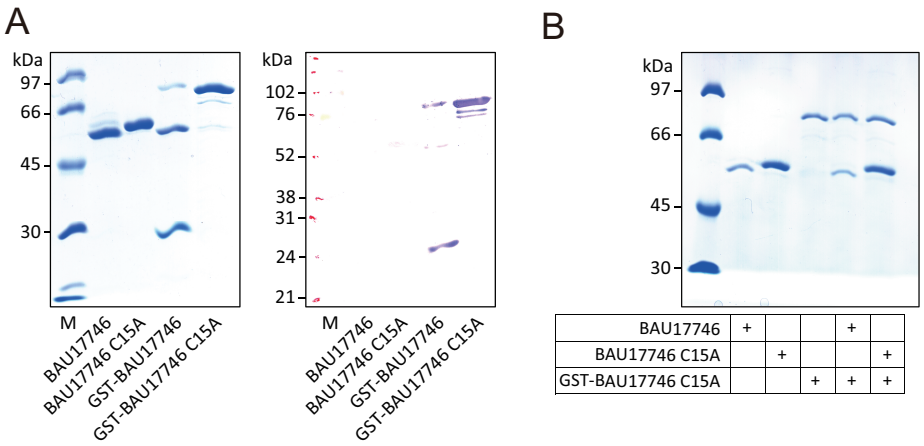


Figure 5 Effect of the Cys15Ala mutation on the processing of BAU17746.
(A) Recombinant proteins (left, 1 μg; right, 0.3 μg) were separated with SDS-PAGE and stained with CBB (left) or by immunoblotting with an anti-GST antibody (right). Low-molecular-weight markers (left) and full-length molecular markers (right) were used. (B) Recombinant proteins (0.2 mg/ml each, 40 μl) were incubated at 37 ˚C. After 5 h, proteins were denatured in the SDS buffer and proteins (1 μg of each) were separated on SDS-PAGE.

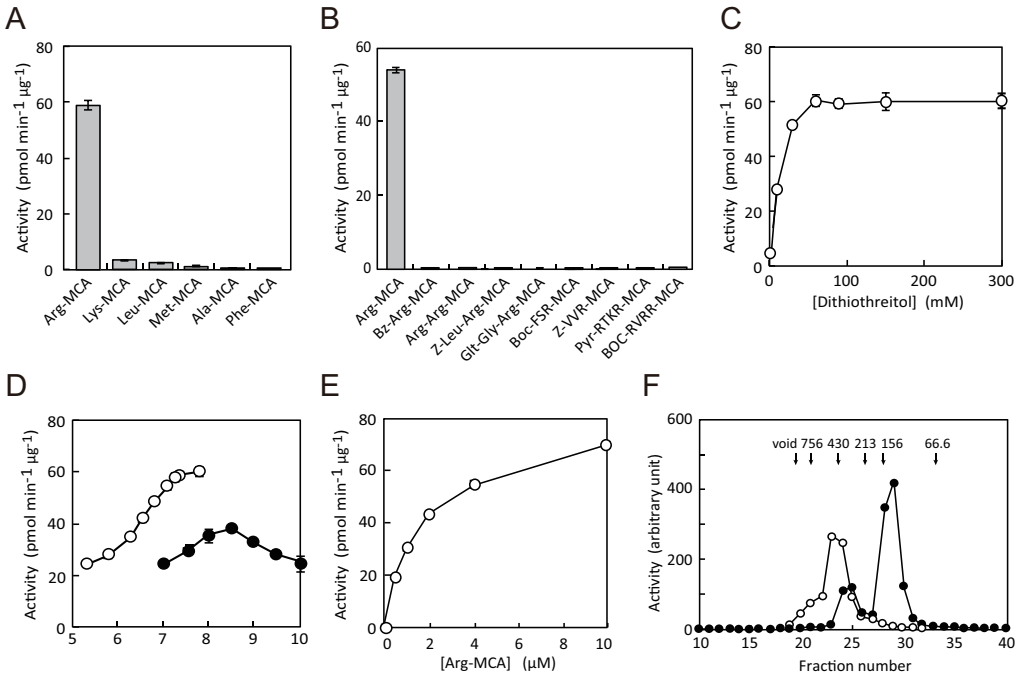


Figure 6 Properties of BAU17746.
(A) Aminopeptidase and (B) peptidase activities toward peptides containing Arg at the P1 position were determined with recombinant BAU17746. (C) Arg-MCA-hydrolyzing activity of BAU17746 was determined at 0 – 300 mM dithiothreitol. (D) pH profile of Arg-MCA-hydrolyzing activity with sodium phosphate (open circle) and Tris-HCl (closed circle) buffers. (E) Concentration dependence of Arg-MCA hydrolysis was determined. Values are shown as the mean ± S.D. (n = 3). (F) BAU17746 (0.5 mg/0.5 ml, open circle) and *Pre. intermedia* cell lysate containing cytosol and periplasm fractions (0.5 ml, closed circle) were separately subjected to size-exclusion HPLC with a Superdex 200 10/300 column. Activity was determined with Arg-MCA. Molecular markers shown running in parallel are blue dextran 2000 (void volume), thyroglobulin (756 kDa), ferritin (430 kDa), catalase (213 kDa), aldolase (156 kDa), and BSA (67 kDa).

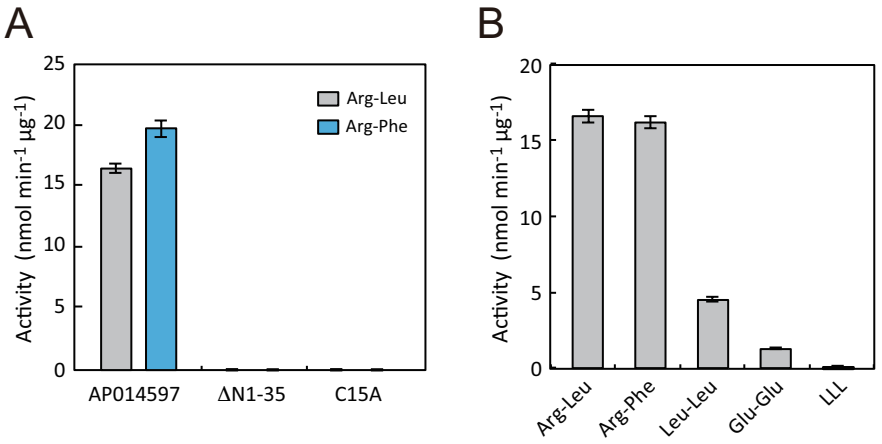


Figure 7 Dipeptidase activity of recombinant peptidases.
(A) BAU17746, BAU17746ΔN1-35, and BAU17746C15A dipeptidase activities were determined with 0.3 mM dipeptides using a ninhydrin method, as described in the Materials and methods. (B) Substrate specificity of BAU17746 was determined with 0.3 mM di- and tri-peptides. Values are shown as the mean ± S.D. (n = 3).

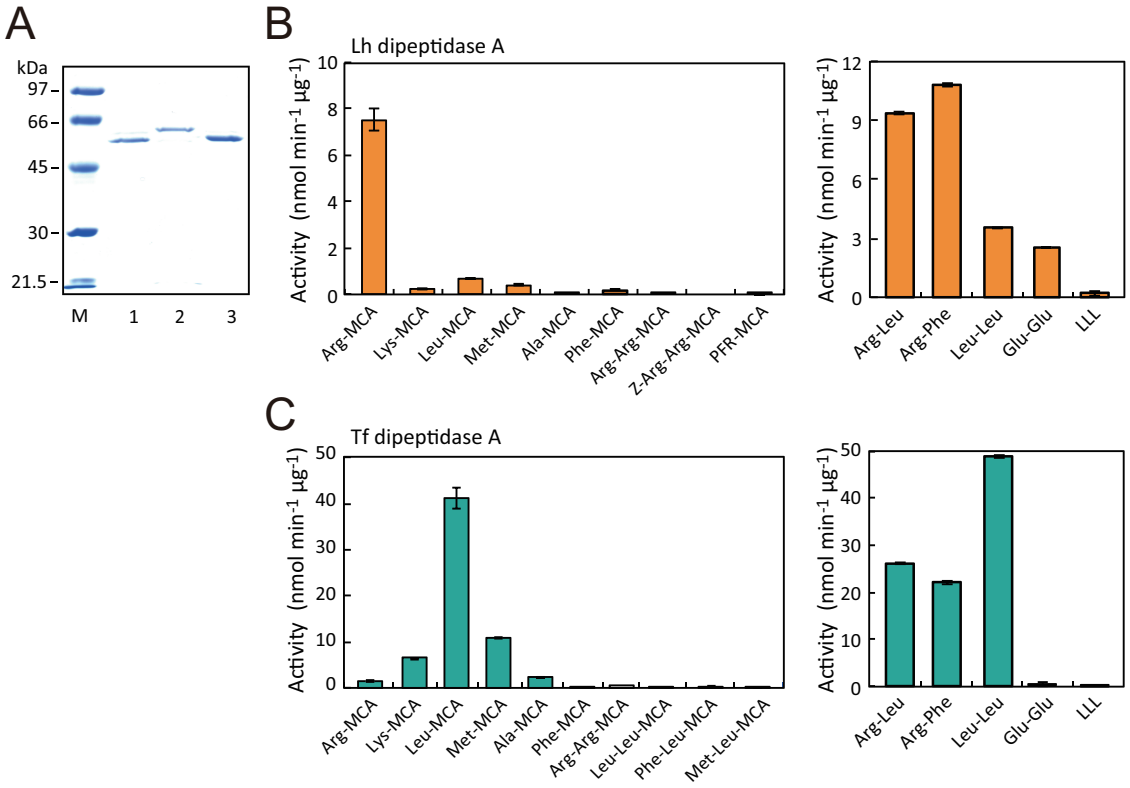


Figure 8 P1-position specificity of *L. helveticus* and *T. forsythia* dipeptidases.
(A) Recombinant proteins (1 μg) of BAU17746 (lane 1), *L. helveticus* dipeptidase A (lane 2), and *T. forsythia* dipeptidase (lane 3) were separated with SDS-PAGE. Lane M, low-molecular-weight marker. (B) Peptidase activity of *L. helveticus* dipeptidase A was determined with 20 μM aminoacyl-MCA (left) and 0.3 mM di- and tri-peptides (right). (C) Peptidase activity of *T. forsythia* dipeptidase was determined with 20 μM aminoacyl-MCA (left) and 0.3 mM di- and tri-peptides (right). Values are shown as the mean ± S.D. (n = 3).

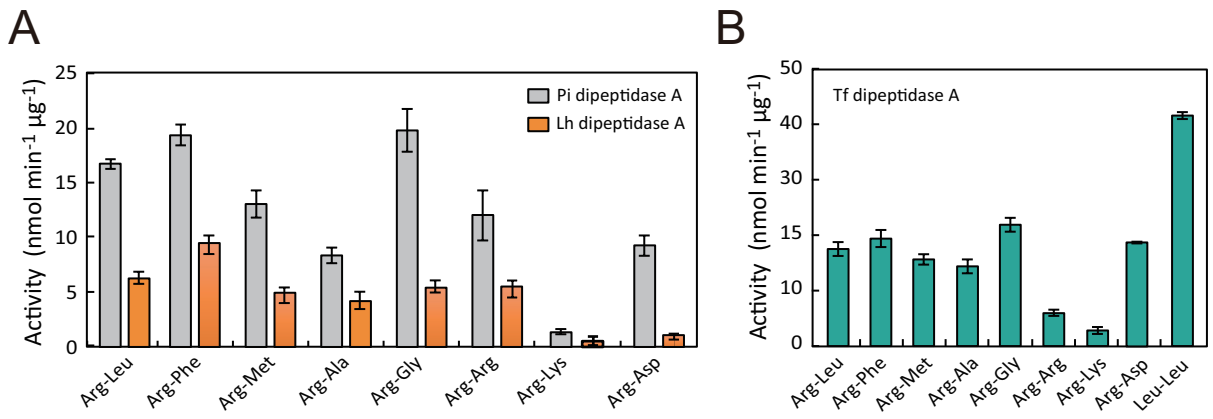


Figure 9 P1'-position specificity of recombinant dipeptidases.
Dipeptidase activities of (A) BAU17746 and *L. helveticus* dipeptidase A, and (B) *T. forsythia* dipeptidase were determined with 0.3 mM dipeptides harbouring Arg at the N-terminus. Additionally, the activity of *T. forsythia* dipeptidase A was determined with 0.3 mM Leu-Leu (presented in panel B). Values are shown as the mean ± S.D. (n = 3).