APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Sept. 2007, p. 5598–5606 Vol. 73, No. 17
0099-2240/07/$08.00+0 doi:10.1128/AEM.00642-07

Characterization of a *Bifidobacterium longum* BORI Dipeptidase Belonging to the U34 Family▽

Jeong Min Seo,[1,3] Geun Eog Ji,[2,3] Sang Hee Cho,[1] Myung Soo Park,[3,4] and Hyong Joo Lee[1]*

Department of Agricultural Biotechnology and Center for Agricultural Biomaterials, Seoul National University, 56-1 Sillim-dong, Gwanak-gu, Seoul 151-921, Korea[1]; Department of Food and Nutrition, Seoul National University, 56-1 Sillim-dong, Gwanak-gu, Seoul 151-742, Korea[2]; Research Institute, Bifido Inc., Ltd., Hongchon 250-804, Korea[3]; and Department of Hotel Culinary Art, Anyang Technical College, Anyang 430-749, Korea[4]

Received 20 March 2007/Accepted 23 June 2007

A dipeptidase was purified from a cell extract of *Bifidobacterium longum* BORI by ammonium sulfate precipitation and chromatography on DEAE-cellulose and Q-Sepharose columns. The purified dipeptidase had a molecular mass of about 49 kDa and was optimally active at pH 8.0 and 50°C. The enzyme was a strict dipeptidase, being capable of hydrolyzing a range of dipeptides but not tri- and tetrapeptides, *p*-nitroanilide derivatives of amino acids, or N- or C-terminus-blocked dipeptides. A search of the amino acid sequence of an internal tryptic fragment against protein sequences deduced from the total genome sequence of *B. longum* NCC2705 revealed that it was identical to an internal sequence of the dipeptidase gene (*pepD*), which comprised 1,602 nucleotides encoding 533 amino acids with a molecular mass of 60 kDa, and thereby differed considerably from the 49-kDa mass of the purified dipeptidase. To understand this discrepancy, *pepD* was cloned into an *Escherichia coli* expression vector (pBAD-TOPO derivative) to generate the recombinant plasmids pBAD-pepD and pBAD-pepD-His (note that His in the plasmid designation stands for a polyhistidine coding region). Both plasmids were successfully expressed in *E. coli*, and the recombinant protein PepD-His was purified using nickel-chelating affinity chromatography and reconfirmed by internal amino acid sequencing. The PepD sequence was highly homologous to those of the U34 family of peptidases, suggesting that the *B. longum* BORI dipeptidase is a type of cysteine-type N-terminal nucleophile hydrolase and has a β-hairpin motif similar to that of penicillin V acylase, which is activated by autoproteolytic processing.

Lactic acid bacteria (LAB) are characterized by their high demand for essential growth factors, such as peptides and amino acids. Milk does not contain sufficient free amino acids and peptides to allow the growth of LAB (31), which has led to LAB also possessing a complex system of proteinases and peptidases that enable them to use milk casein as a source of amino acids and nitrogen. The first step in casein degradation is mediated by proteases in the cell membrane that cleave casein into oligopeptides. LAB peptidases further degrade casein into smaller peptides and amino acids that can pass through the cell membrane (28, 30). The proteolytic activities of LAB, including yogurt bacteria and probiotic bacteria, have been studied extensively, and proteolytic enzymes have been isolated and characterized (5). Several endopeptidases have been purified from yogurt bacteria, including *Lactobacillus delbrueckii* subsp. *bulgaricus* B14 (4) and *Streptococcus thermophilus* CNRZ160 (8). Various dipeptidases have been purified from *S. thermophilus* (26) and *L. delbrueckii* subsp. *bulgaricus* B14 (30), and tripeptidases have been purified from *L. delbrueckii* subsp. *bulgaricus* B14 (2) and *S. thermophilus* CNRZ160 (26).

The components of the proteolytic systems of lactococci and some lactobacilli have been investigated, and many peptidases from LAB have been purified and characterized. In addition, several peptidase genes have been cloned and sequenced (18, 20).

However, very little is known about the proteolytic enzyme systems of *Bifidobacterium* spp., and no amino-, di-, and tripeptidase genes have been cloned from bifidobacteria. *Bifidobacterium* spp. are comparable to LAB in that their general aminopeptidase activity encompasses several dipeptidases and possibly also iminopeptidase and tripeptidases (3, 10, 22). Aminopeptidase and iminopeptidase activities were determined in a cell extract of *Bifidobacterium breve* by Cheng and Nagasawa (7), and *Bifidobacterium infantis*, *Bifidobacterium longum*, and *Bifidobacterium adolescentis* show amino-, di-, tri-, and carboxypeptidase activities (11). *B. longum* contains more than 20 predicted peptidases that could provide amino acids from proteinaceous substrates in the gastrointestinal tract and vagina, where carbohydrates are less abundant (27).

In recent years, bifidobacteria have been used in fermented dairy products and nutraceuticals for promoting intestinal health. Although bifidobacteria represent only 3 to 6% of the adult fecal flora, their presence has been associated with beneficial health effects, such as prevention of diarrhea, amelioration of lactose intolerance, and immunomodulation. These correlations have led to the widespread use of bifidobacteria as components of health-promoting foods (27). However, some probiotic bacteria, such as *Lactobacillus acidophilus* and *Bifidobacterium* spp., grow slowly in milk because of the absence of proteolytic activity (17).

Proteolytic enzymes, such as di-, tri-, and aminopeptidases, are expected to enhance fermentation because bifidobacteria have a high demand for essential growth factors, such as various amino acids. The importance of free amino acids and

* Corresponding author. Mailing address: School of Agricultural Biotechnology and Center for Agricultural Biomaterials, Seoul National University, 56-1 Sillim-Dong, Gwanak-gu, Seoul 151-921, Republic of Korea. Phone: 82-2-880-4853. Fax: 82-2-873-5095. E-mail: leehyjo@snu.ac.kr.

▽ Published ahead of print on 29 June 2007.

TABLE 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant characteristic(s)	Source or reference
Bacterial strains		
B. longum BORI	Source of a dipeptidase gene (*pepD*)	Infant feces (this study)
E. coli strains		
DH5α	Cloning host	Takara
LMG194	Expression host; F$^-$ Δ*lacX74 galE thi rpsL* Δ*phoA* (*pvuII*) Δ*ara-714 leu*::Tn*10*	Invitrogen
Plasmids		
pGEM-T Easy	TA cloning vector; Amp^r	Promega
pBAD-TOPO	Amp^r; pBR322 derivative under the control of *araC* promoter containing an inserted SpeI restriction site	Invitrogen
pBAD-TOPO derivative	Amp^r; pBAD-TOPO derivative containing an inserted SpeI restriction site	This study
pBAD-pepD	Amp^r; pBAD TOPO derivative containing *pepD*	This study
pBAD-pepD-His	Amp^r; pBAD TOPO derivative containing *pepD*-His followed by polyhistidine-tagged sequence	This study

peptides to the growth of bifidobacteria has been studied by Cheng and Nagasawa (7).

A dipeptidase is an enzyme that cleaves dipeptides into two amino acids. The dipeptidase activity of *B. longum* BORI is typically twofold (12) that of various other *B. longum* strains when using Ala-Ala dipeptide as a substrate. We decided to purify and clone the gene (*pepD*) coding for this dipeptidase in an *Escherichia coli* expression vector in order to overexpress this gene and purify the enzyme PepD. This paper describes the purification, characterization, and cloning of a dipeptidase from *B. longum* BORI and its expression in *E. coli*.

MATERIALS AND METHODS

Bacterial strains, media, and plasmids. The bacterial strains and plasmids used in this study are listed in Table 1. The *E. coli* strains DH5α and LMG194 (13) were used for cloning experiments and protein expression, respectively. Strains were purchased from Takara (Shiga, Japan) and Invitrogen (Carlsbad, CA) and cultured at 37°C in Luria-Bertani (LB) broth with vigorous shaking. Ampicillin was used at a concentration of 50 $\mu g \cdot ml^{-1}$ for selecting transformed bacteria. The plasmid pGEM-T Easy vector (Promega, Madison, WI) was used for cloning and amplifying the dipeptidase gene (*pepD*), and the plasmid pBAD-TOPO (Invitrogen) was used for protein expression (13). *B. longum* BORI was grown in brain heart infusion broth (Difco, Detroit, MI) supplemented with 0.5% glucose and 0.05% (final concentration) L-cysteine · HCl at 37°C. LB broth media and Bacto agar were purchased from Difco.

Reagents and enzymes. DEAE-cellulose and Q-Sepharose were purchased from Amersham Biosciences (Uppsala, Sweden). Di-, tri-, and oligopeptides, *p*-nitroanilide derivatives of amino acids and peptides, and *N*-carbobenzyloxy (CBZ)-blocked dipeptides were obtained from Bachem Feinchemikalien (Bubendorf, Switzerland) and Sigma Chemicals (St. Louis, MO). Restriction endonucleases, calf intestinal alkaline phosphatase, T4 DNA ligase, and LA-*Taq* polymerase were purchased from Promega and Takara and used according to the supplier's recommendations. Ampicillin and other chemicals were purchased from Sigma Chemicals.

Enzyme purification. All purification procedures were performed at 4°C. To purify the dipeptidase protein from *B. longum* BORI, cells were harvested by centrifugation at 10,000 × *g* for 30 min and washed three times with distilled water. The obtained cells were suspended in 50 mM Tris-HCl buffer (pH 7.0) and disrupted by sonication. The cell debris was removed by centrifugation at 10,000 × *g* for 20 min. Clear supernatant was concentrated by salting out with powdered ammonium sulfate at about 70% saturation. All purification procedures were performed at 4°C. The precipitate was dissolved in 50 mM Tris-HCl buffer (pH 7.0) and dialyzed against the same buffer. The dialysate was applied to a DEAE-cellulose column (2.8 by 60 cm) that had been equilibrated with 50 mM Tris-HCl buffer. The column was washed with the same buffer, and linear gradient elution at a flow rate of 2.0 $ml \cdot min^{-1}$ was applied with 250 ml of 50 mM Tris-HCl buffer (pH 7.0) and 250 ml of the same buffer containing 1.0 M KCl. In addition, the column was washed with 50 mM Tris-HCl buffer (pH 7.0) containing 1.0 M KCl. The fractions exhibiting dipeptidase activity were pooled and then dialyzed against 50 mM Tris-HCl buffer (pH 7.0). The dialysate was applied to a Q-Sepharose column (2.8 by 60 cm) that had previously been equilibrated with 50 mM Tris-HCl buffer. The column was washed with the same buffer, and linear gradient elution at a flow rate of 2.0 $ml \cdot min^{-1}$ was applied with 250 ml of 50 mM Tris-HCl buffer (pH 7.0) and 250 ml of the same buffer containing 1.0 M KCl. The fractions containing the dipeptidase activity were pooled, dialyzed against 50 mM Tris-HCl buffer (pH 7.0), and lyophilized.

SDS-PAGE. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed (Mini-Protein II; Bio-Rad Laboratories, Berkeley, CA) with a 10% polyacrylamide gel. Molecular mass SDS-PAGE protein standards supplied by Elpis Biotech (Taejeon, Korea) were used as reference proteins in the range from 100 to 20 kDa. SDS-PAGE was performed by the procedure of Laemmli (19).

Measurement of dipeptidase activity. Dipeptidase activity was quantified using the cadmium-ninhydrin assay (12) to follow the hydrolysis of Ala-Gln. The reaction mixture consisted of 50 μl of enzyme preparation, 50 μl of substrate (10 mM in methanol or deionized water), and 400 μl of reaction buffer (50 mM Tris-HCl buffer, pH 8.0). The mixture was incubated for 30 min at 37°C, and the reaction was terminated by the addition of 1 ml of cadmium-ninhydrin reagent. The contents were thoroughly mixed, heated at 84°C for 5 min, and cooled immediately on ice, and then the absorbance at 507 nm was measured. One unit of enzyme activity was defined as that required to increase the absorbance by one unit per minute under the assay conditions.

Measurement of protein concentration. The protein was measured using a Bradford protein assay kit with bovine serum albumin as the standard. The standard bovine serum albumin solution was purchased from Bio-Rad Laboratories, and the absorbance at 595 nm was measured.

Effects of pH and temperature on dipeptidase activity. The effects of pH values in the range 2 to 10 on enzyme activity were examined in buffers at 50 mM, comprising glycine-HCl (pH 2 to 3), maleic acid–Tris-NaOH (pH 4 to 6), Tris-HCl (pH 7 to 9), and glycine-NaOH (pH 10). To assess the effect of pH on enzyme stability, the enzyme was dissolved in buffers at 50 mM at pH values ranging from 2 to 10 and incubated for 30 min at 37°C. The residual activity was subsequently measured at 37°C with Ala-Gln as the substrate. The optimum temperature for dipeptidase activity was determined within the range from 10°C to 80°C in 50 mM Tris-HCl buffer (pH 8.0). The thermal stability of the enzyme was assessed by incubating the enzyme solutions for 10 min at temperatures ranging from 10°C to 80°C. The residual activity was subsequently measured at 37°C with Ala-Gln as the substrate.

N-terminal and internal amino acid sequence analysis. N-terminal amino acid sequencing of a protein blotted on a polyvinylidene difluoride membrane was analyzed at the Korea Basic Science Institute (Seoul, Korea) using a protein sequencer (PerkinElmer 491; Applied Biosystems, Foster City, CA). Edman degradation was performed using the standard program supplied by Applied Biosystems. The N-terminus-blocked proteins were further analyzed by internal amino acid sequencing. The blocked spots were cut out with a blade, subjected to in-gel trypsin digestion, and analyzed by reverse-phase high-performance liquid chromatography (SMART HPLC; Pharmacia Diagnostics, Uppsala, Sweden). Twenty-microliter aliquots of peptide samples were separated on a Sephasil C_{18} column (Amersham Biosciences)

TABLE 2. Purification of the dipeptidase from *B. longum* BORI

Purification step	Total amt of protein (mg)	Total activity (U)	Sp act (U · mg^{-1})	Recovery (%)	Purification ratio
Extraction from the cell	495.7	351	0.71	100	1.0
$(NH_4)_2SO_4$ precipitation	124.3	322	2.59	91.7	3.6
DEAE-cellulose chromatography	8.4	279	33.2	79.4	46.8
Q-Sepharose chromatography	1.8	139	77.2	26.8	109

(2.1 × 100 mm) equilibrated with acidified water containing 0.1% (wt/vol) trifluoroacetic acid. The column was washed with acidified water for 5 min, and the peptide sample was eluted with a linear gradient of acetonitrile (0 to 60% [vol/vol]) in acidified water over a 60-min period at a flow rate of 200 ml · min^{-1}. Fractions of 150 ml were collected.

Amplification of the *pepD* gene from *B. longum* BORI genomic DNA. PCR was used to amplify *pepD* from genomic DNA of *B. longum* BORI using GeneReleaser (Bioventure, Murfreesboro, TN). Two primers, pepD-F (forward, 5′-CACA**CCATGG**CCTGCACCACGATTCTG-3′) and pepD-R (reverse, 5′-TGGG**ACTAGT**CTAAAAGTCGGACATGTGGAAGC-3′), were designed based upon the *B. longum* NCC2705 genomic sequence (GenBank accession number NC_004307.2, gene BL0479), where the underlining indicates a translation start or stop codon. NcoI and SpeI restriction sites (shown in boldface type in the above sequences) were incorporated into pepD-F and pepD-R, respectively, in order to facilitate the cloning of *pepD* to the pBAD-TOPO derivative vector for subsequent screening. The pBAD-TOPO derivative was constructed by inserting the PCR product containing the SpeI site into the pBAD-TOPO vector (Invitrogen). Also, another primer, pepD-His-R (reverse, 5′-ATAC**ACTAGT**AAAGTCGGACATGTGGAAGC-3′; the SpeI restriction site shown in boldface type), was designed to amplify *pepD* for the expression of a C-terminal polyhistidine-tagged fusion protein. The pepD-His-R primer has no stop codon in order to continue the nucleotide sequence coding for the polyhistidine tag of the pBAD-TOPO derivative. The PCR product obtained from the pepD-F and pepD-R primer set was named pepD, and that obtained from the pepD-F and pepD-His-R primer set was named pepD-His, which is the C-terminal polyhistidine-tagged fusion protein. The additional C-terminal amino acid sequence deduced from the PepD-His-encoding gene was TSGPKGELEGKPIPNPLLGLDSTRTGHHHHHH, which is from amino acids 554 to 585 in the sequence of PepD-His (the underlining indicates the V5 epitope and polyhistidine tag region). The PCR was performed with 25 ng of DNA in a final volume of 50 μl containing deoxyribonucleoside triphosphate (each at 0.25 to 0.5 mM), oligonucleotides (50 pM), and 1.0 to 3.0 U of LA-*Taq* polymerase (Takara). Amplification was performed on a thermal cycler (PROGENE; Techne, Cambridge, United Kingdom) with 30 cycles of denaturation at 95°C for 30 s (5 min in the first cycle), annealing at 60°C for 60 s, and elongation at 72°C for 90 s (10 min in the last cycle).

General cloning techniques and sequence analysis. The PCR products encoding PepD and PepD-His were digested with NcoI and SpeI and then ligated with the NcoI-SpeI-digested pBAD expression vector (Invitrogen) to generate pBAD-pepD and pBAD-pepD-His, respectively. The ligation mixture was used in *E. coli* DH5α heat shock transformation. The transformed cells were spread on LB agar containing 50 μg · ml^{-1} ampicillin. After incubation at 37°C overnight, the transformants were selected, and their recombinant plasmids were purified (plasmid miniprep kit; QIAGEN, Hilden, Germany). The sequence of the cloned *pepD* gene was then determined (BigDye Terminator and ABI377 system; Applied Biosystems).

Expression of two recombinant dipeptidases of PepD and PepD-His in *E. coli*. The plasmids pBAD-pepD and pBAD-pepD-His (carrying *pepD* and *pepD*-His) proliferated in *E. coli* DH5α and were extracted, purified, and transformed again in *E. coli* LMG194, with the transformants screened on an agar plate containing 50 μg · ml^{-1} ampicillin. The transformant cells were cultured in LB broth containing 50 μg · ml^{-1} ampicillin, shaken overnight at 37°C, transferred into flasks containing 50 ml of fresh LB broth, and grown at 37°C with vigorous shaking. The *pepD* and *pepD*-His genes were expressed by induction with 0.02% L-arabinose, which was added to the culture broth of *E. coli* LMG194 harboring each plasmid at the beginning of the exponential phase (optical density at 600 nm of 0.5). Six hours after induction, in the case of *E. coli* LMG194 harboring pBAD-pepD-His, 50 ml of the culture solution was centrifuged at 10,000 × *g* for 10 min at 4°C, and the pelleted cells were resuspended in 5 ml of lysis buffer (50 mM Tris-HCl, 300 mM NaCl, and 10 mM imidazole, pH 8.0). In the case of *E. coli* LMG194 harboring pBAD-pepD, cell growth was remarkably retarded and hence was stopped 18 h after induction. The suspension was sonicated at least four times by a sonic dismembrator (Fisher Scientific, Pittsburgh, PA) using 30-s pulses that were applied at 1-min intervals. The lysate was clarified at 10,000 × *g* for 10 min at 4°C, and the soluble fraction was analyzed by SDS-PAGE.

Purification of PepD-His. The recombinant PepD-His (rPepD-His) was purified using an immobilized nickel-nitrilotriacetic acid (Ni-NTA) affinity column. The soluble fraction containing the recombinant protein was loaded onto the Ni-NTA column that had previously been equilibrated in the same lysis buffer, and the column was washed twice with five times the bed volume of 50 mM Tris-HCl, 300 mM NaCl, and 20 mM imidazole (pH 8.0). The recombinant protein was then eluted using elution buffer (50 mM Tris-HCl, 300 mM NaCl, and 250 mM imidazole; pH 8.0), and the eluted protein was analyzed by SDS-PAGE.

MALDI-TOF MS. Matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) mass spectrometry (MS) was performed (Voyager-DE STR biospectrometry workstation; Applied Biosystems) using ions formed with a N_2 pulsed laser beam (337 nm) accelerated to 25 kV.

RESULTS

Purification and biochemical properties of a native dipeptidase from *B. longum* BORI. The dipeptidase was purified about 109-fold from the cell extract by ammonium sulfate fractionation and two steps of column chromatography, for which the recovery in activity was about 26.8%. The specific dipeptidase activity observed with Ala-Gln as the substrate at each purification step is summarized in Table 2. SDS-PAGE of the enzyme obtained from the final Q-Sepharose chromatography step revealed a single protein band with an apparent molecular mass of 49 kDa (Fig. 1A).

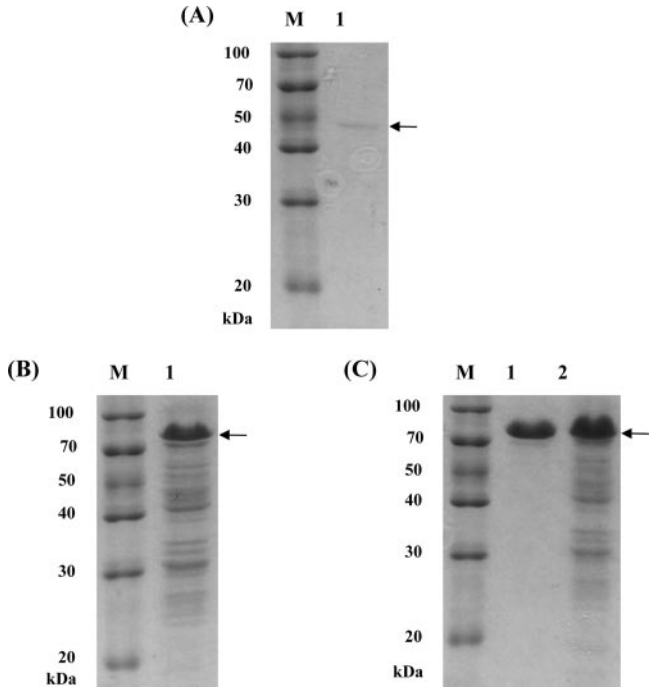

FIG. 1. Purification of *B. longum* BORI dipeptidase and expression of recombinant PepD and PepD-His in *E. coli*. (A) SDS-PAGE of the enzyme obtained from the final Q-Sepharose chromatography step. (B) SDS-PAGE of the induced rPepD in *E. coli* LMG194 harboring pBAD-pepD. Lanes: M, molecular mass standards; 1, disrupted cell extract containing rPepD. (C) SDS-PAGE of the induced rPepD-His in *E. coli* LMG194 harboring pBAD-pepD-His. Lanes: M, molecular mass standards; 1, purified rPepD-His from Ni-NTA affinity chromatography; 2, disrupted cell extract containing rPepD-His.

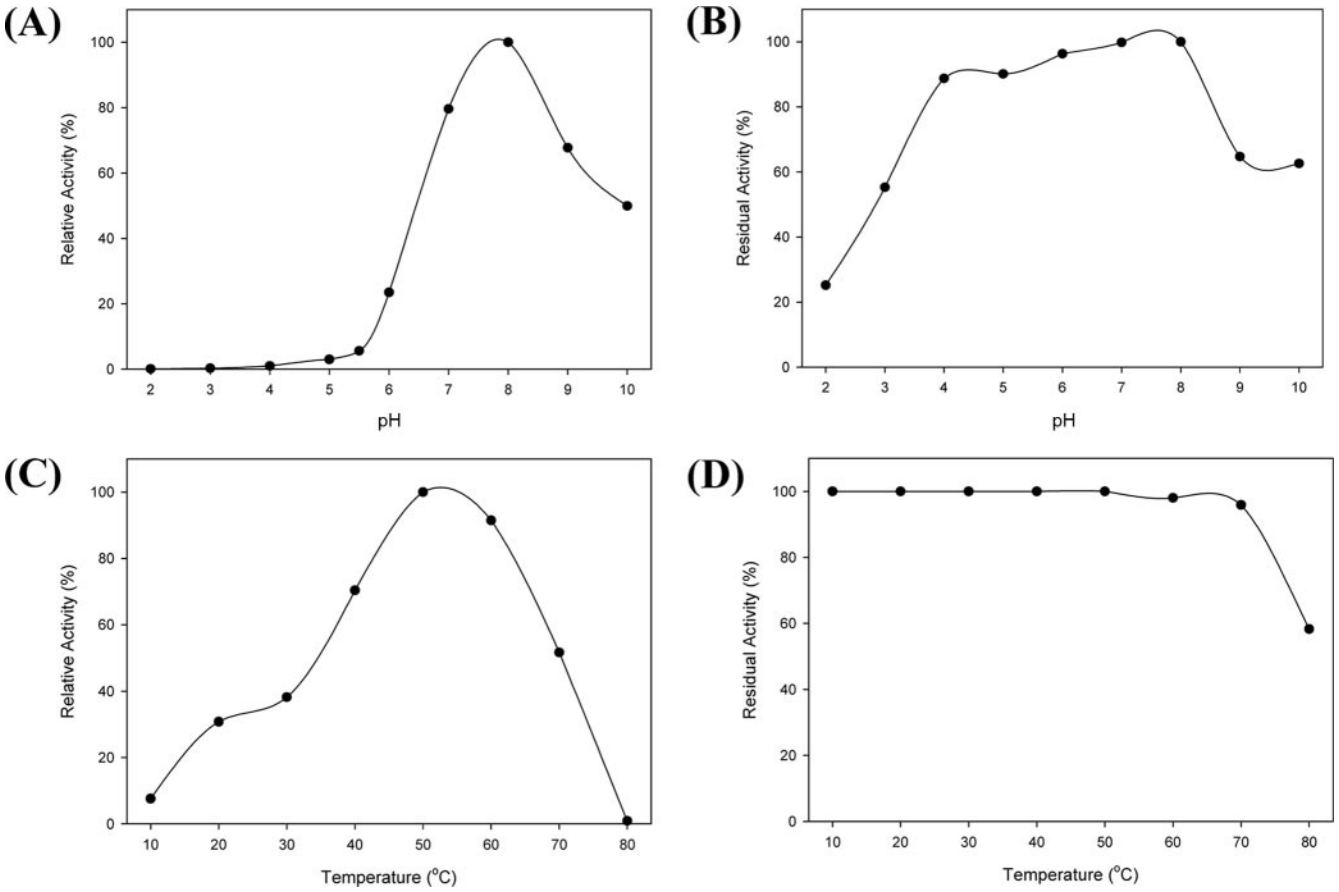

FIG. 2. Effects of environmental conditions on the activity of the dipeptidase purified from *B. longum* BORI determined on Ala-Gln. (A) pH dependence of enzyme activity as measured at 37°C. (B) Residual activity of the enzyme after 30 min of preincubation at various pH values as measured at 37°C for 30 min. (C) Effects of temperature on the activity of the enzyme at pH 8. (D) Residual activity of the enzyme after 10 min of preincubation at various temperatures as measured at 37°C for 30 min.

The dipeptidase enzyme showed high activity when the pH value was in the range 7.0 to 9.0, with optimal activity at pH 8.0 (Fig. 2A). Very low activity (less than about 10%) was observed below pH 5.5, and at pH 10.0, the activity was about 50% of the maximum. The enzyme was stable for 30 min over the pH range 4.0 to 8.0 (Fig. 2B). The effects of temperature on the enzyme activity and stability are shown in Fig. 2C and D, respectively. At pH 8.0, the enzyme showed high activity at temperatures from 40°C to 60°C, with maximal activity at 50°C, and it was about 30% and 50% of the maximum at 20°C and 70°C, respectively. The enzyme was stable at temperatures up to 70°C, but the activity decreased rapidly above 80°C.

The ability of the dipeptidase to hydrolyze different substrates is indicated in Table 3. The enzyme was most active toward Ala-Gln and displayed high activity toward the Ala-X dipeptides except for hydrophobic amino acids, such as Leu and Phe at their N termini. Its specificity also included dipeptides Leu-Leu and Phe-Gly. It showed low activity toward dipeptides containing aromatic amino acids except for Phe-Gly. It did not hydrolyze Leu-Asp, Phe-Phe, or N- or C-terminus-blocked dipeptides (CBZ-Phe-Gly, CBZ-Ala-Leu, or Leu-Leu-amide). Also, the enzyme did not hydrolyze tri-, tetra-, or pentapeptides or *p*-nitroanilide derivatives of dipeptides and peptides.

N-terminal amino acid sequencing of the purified dipeptidase was not successful. However, the amino acid sequence of a trypsin-generated fragment was determined as HLDLAV ENT. This sequence was found to be identical to a region of PepD deduced from the full genomic sequence of *B. longum* NCC2705. The internal amino acid sequence was identical to the sequence between amino acids 231 to 239 in *pepD* (gene identification number [gi] 23465071) (Fig. 3). Another dipeptidase gene, *pepDB* (gi 23465037), was found in the *B. longum* genomic sequence and had no region corresponding to the internal amino acid sequence determined from the purified dipeptidase. Since it was found that the purified dipeptidase was encoded by *pepD*, we planned to clone its gene to express and characterize it in *E. coli* using pBAD-TOPO derivative as an *E. coli* expression vector (Table 1).

Cloning and expression of the *pepD* gene from *B. longum* BORI in *E. coli*. The *pepD* gene was amplified by PCR using *B. longum* BORI chromosomal DNA as a target and the primer pair pepD-F and pepD-R. To determine the DNA sequence of *pepD*, an amplified version (1.6 kb) was subcloned into pGEM-T Easy vector. The cloned gene comprised a 1,602-bp sequence that was identical to the *pepD* sequence in GenBank under accession number NC_004307.2. These 1,602 nucleotides encoded a protein of 59.75 kDa comprising 553 amino acids (Fig. 3).

The *pepD* and *pepD*-His genes were subcloned into *E. coli*

TABLE 3. Substrate specificity of the dipeptidase purified from *B. longum* BORI

Substrate	Relative activity (%)
Ala-Ala	62.5
Ala-Gln[a]	100.0
Ala-Gly	68.9
Ala-Leu	14.0
Ala-Phe	14.0
Arg-Phe	6.3
Asp-Phe	4.9
Leu-Asp	0
Leu-Leu	71.0
Leu-Phe	12.5
Leu-Trp	10.2
Phe-Ala	6.7
Phe-Gly	74.1
Phe-Phe	0
Tyr-Leu	9.5
Tyr-Phe	3.8
CBZ-Phe-Gly	0
CBZ-Ala-Leu	0
Leu-Leu-amide	0
Ala-Ala-Ala	0
Ala-Leu-Gly	0
Leu-Gly-Gly	0
Leu-Leu-Leu	0
Pro-Phe-Gly-Lys	0
Ala-Ala-Ala-Ala	0
Ala-Ala-Ala-Ala-Ala	0
Leu-pNA[b]	0
Ala-pNA	0
Arg-Pro-pNA	0
Gly-Phe-pNA	0

[a] Ala-Gln hydrolysis was normalized to 100%.
[b] pNA, *p*-nitroanilide.

DH5α by using the pBAD-TOPO derivative vector as described in Materials and Methods. The two new plasmids containing the *pepD* and *pepD*-His genes, designated pBAD-pepD and pBAD-pepD-His, respectively, were sequenced to ensure that the PCR had not introduced any errors.

pepD was successfully expressed in *E. coli* LMG194 (Fig. 1B). The rPepD and rPepD-His proteins were expressed by induction with 0.02% L-arabinose. Further, rPepD-His was purified using a Ni-chelating resin column as described in Materials and Methods (Fig. 1C). Both rPepD and rPepD-His exhibited no dipeptidase activity, even though they were overexpressed in *E. coli*. SDS-PAGE revealed that rPepD and rPepD-His expressed in *E. coli* both had a molecular mass of about 80 kDa, which was not as expected from the theoretical molecular mass deduced from the *pepD* sequence. MALDI-TOF MS was used to determine whether this unexpected result was simply due to an unusual conformation or to the formation of a complex with another protein(s) when the protein was expressed in *E. coli*.

MALDI-TOF MS of the native dipeptidase purified from *B. longum* BORI produced a mass spectrum with a clear peak at 49,496 Da, as shown in Fig. 4A. This result was consistent with the molecular mass determined from SDS-PAGE (Fig. 1A). MALDI-TOF MS of purified dipeptidase (PepD-His) including the V5 epitope and polyhistidine residues also resulted in a characteristic mass spectrum, as shown in Fig. 4B, where the single fragment at a mass-to-charge ratio (*m*/*z*) of 62,932 corresponds to PepD-His. Actually, 63,191 is the theoretical *m*/*z* value deduced from the *pepD*-His gene sequence, which additionally encodes the V5 epitope and polyhistidine tag region. This discrepancy between the molecular masses is probably attributable to the MALDI-TOF MS procedures; hence, we consider that the amino acid sequence analysis and MALDI-TOF MS data were consistent. However, this result was not consistent with the molecular mass (about 80 kDa) determined from SDS-

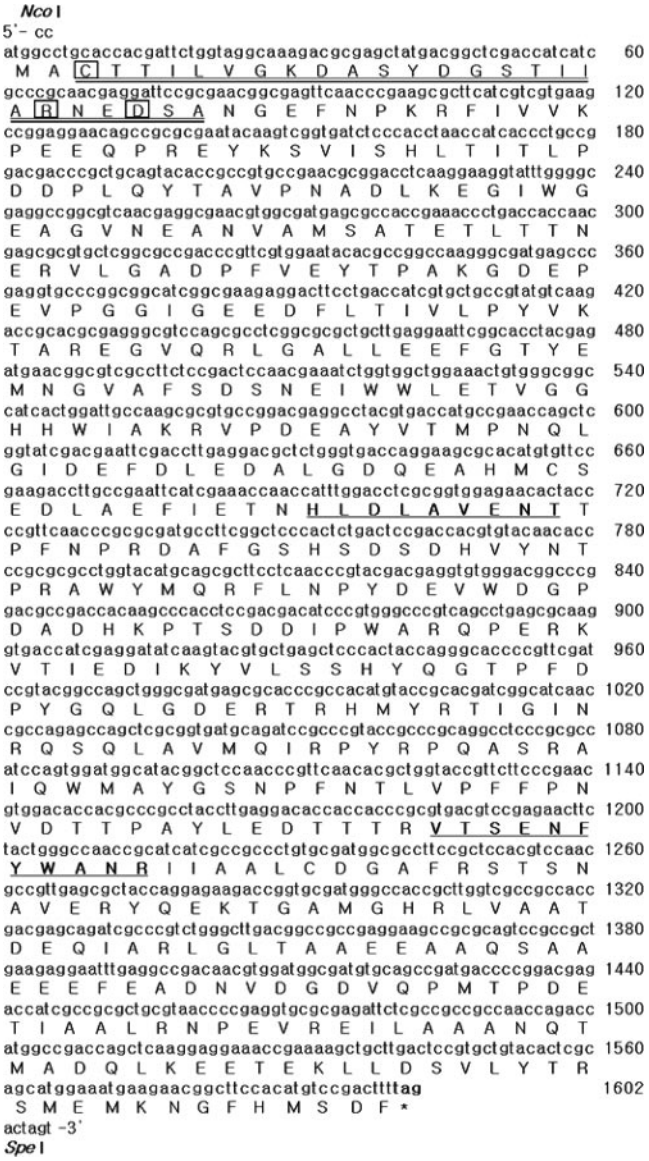

FIG. 3. Nucleotide sequences of 1.6-kb fragments and deduced amino acid sequence of *pepD* from *B. longum* BORI. The N-terminal conserved sequence (amino acids 3 to 27) of the U34 family are double underlined, and the catalytic cysteine and the Arg and Asp essential residues are boxed. The determined internal amino acid sequences of the purified enzyme and rPepD-His were identical at the boldface and underlined letters (amino acids 231 to 239 and 395 to 405 of the sequence deduced from *pepD*, respectively). The nucleotide sequence is available in the NCBI and GenBank databases under accession number NC_004307.2.

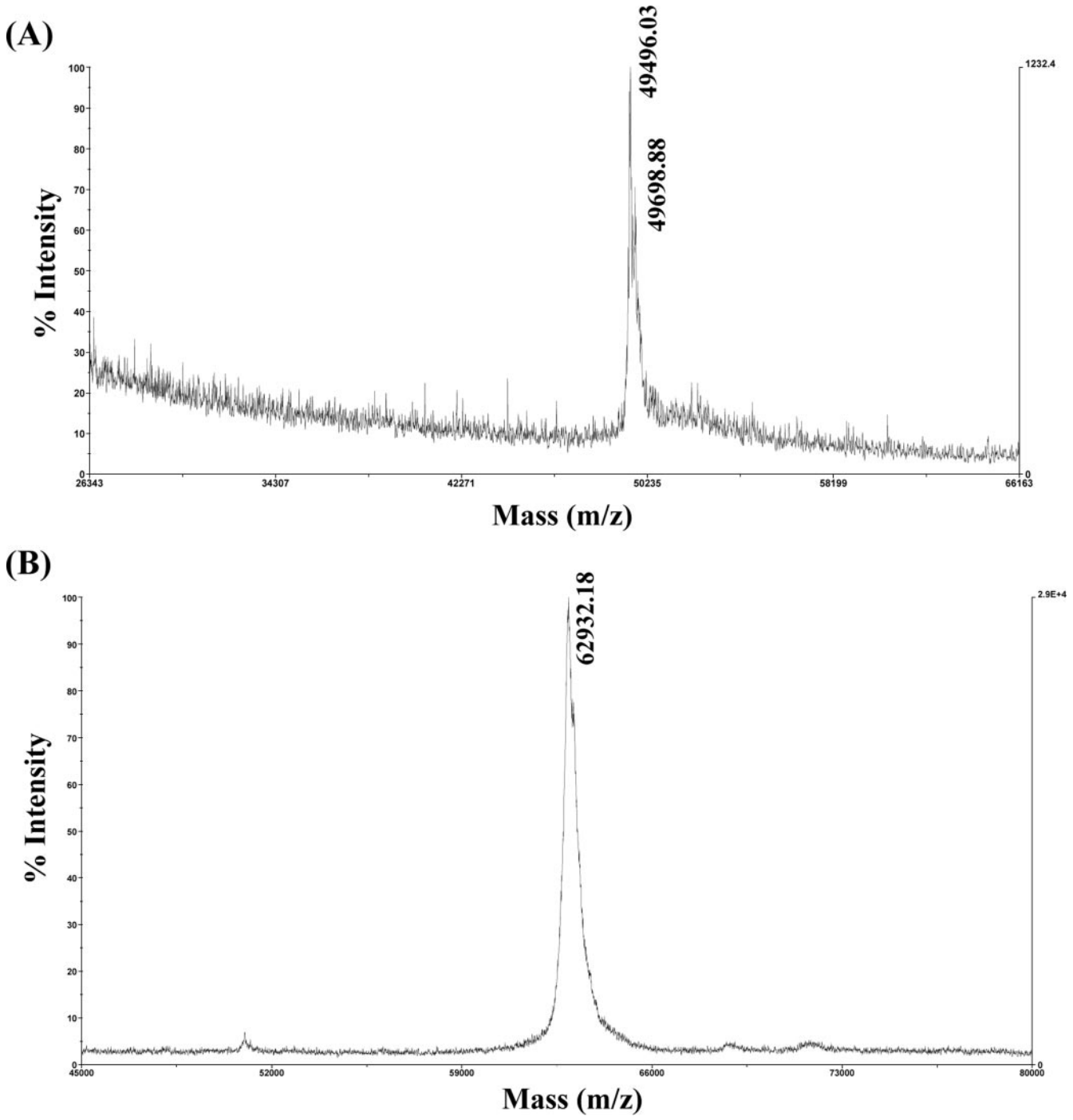

FIG. 4. MALDI-TOF MS spectra of native dipeptidase purified from *B. longum* BORI (A) and His-tagged dipeptidase expressed in *E. coli* (B).

PAGE (Fig. 1C), which indicates that the PepD protein has an unusual conformation. We were also unable to determine the N-terminal sequence of the purified rPepD-His as with the purified native enzyme. The internal amino acid sequence of rPepD-His was determined to be VTSENFYW ANR, which is identical to the sequence between amino acids 395 to 405 deduced from *pepD* (Fig. 3). Another dipeptidase gene, *pepDB*, had no region corresponding to the internal amino acid sequence determined from the purified rPepD-His (data not shown).

MALDI-TOF MS revealed that the molecular mass of native PepD from *B. longum* BORI was about 10 kDa lower than that of rPepD expressed in *E. coli*. These data suggest that the native dipeptidase is subject to a posttranslational modification involving protein splicing, unlike recombinant dipeptidases (rPepD and rPepD-His) expressed in *E. coli*.

TABLE 4. Multiple-sequence alignment of the first β-hairpin motif in the PVA/U34 family and other Ntn hydrolases

Row	gi no.	Species[a]	FRN[b]	Amino acid sequence[c]	Protein name
A	23465071	Bl	3	CTTILVGKDASYDGSTIIARNEDSA (533)	U34 dipeptidases and their close homologs
B	18202547	Lh	6	CTTILVGKKASIDGSTMIARSEDGG (474)	
C	15673540	Ll	11	CTTVLVGRKASIDGSTMIARNDDGH (474)	
D	15672234	Ll	3	CTTFLVGKKASLDGTTLIARNEDGG (459)	
E	1370209	Ls	10	CTSMMVGKNASIDGSTMISRNEDRF (473)	
F	15675831	Sp	26[d]	CTGFIIGKDLTKDGSLLYGRTEDLE (498)	
G	14029606	Sg	26[d]	CCGFIFGRQLTADGSTMFGRTEDYP (652)	
H	13938410	Hs	6	CDTFVALPPATVDNRIIFGKNSDRN (136)	
I	12852713	Mm	6	CDTFVALPPATVGNRVIFGKNSDRN (414)	
J	129549	Bs	4	CSSLSIRTTD---DKSLFARTMDFT (338)	PVAs and choloylglycine hydrolases
K	10732793	Lj	2	CTSIVYSSNN----HHYFGRNLDLE (326)	
L	16081005	Bsu	2	CTSLTLETAD---RKHVLARTMDFA (328)	
M	16804106	Lm	2	CTSITYTTK-----DHYFGRNFDYE (325)	
N	15923265	Sa	2	CTGFTIQTLN---NQVLIGRTMDYD (330)	
O	14751493	Hs	159	CTSIVAEDKK---GHLIHGRNMDFG (411)	Acid ceramidases
P	16758140	Rn	142	CTSIITEDGK---GHLLHGRNMDFG (394)	
Q	3025288	Ce	147	CTSIVAQTEDN--KDLYHARNLDFG (401)	
R	81428038	Ls	1	MLNWQQEAAKYKDQMLTDLTSLLK (467)	M20 dipeptidase
S	23465037	Bl	21	CTTILVGKNASYDGSTIIARDDDSG (562)	U34 family

[a] Species abbreviations: Bl, *B. longum*; Bs, *Bacillus sphaericus*; Bsu, *Bacillus subtilis*; Ce, *Caenorhabditis elegans*; Hs, *Homo sapiens*; Mm, *Mus musculus*; Lh, *Lactobacillus helveticus* CNRZ32; Lj, *Lactobacillus johnsonii*; Ll, *Lactococcus lactis*; Lm, *Listeria monocytogenes*; Ls, *Lactobacillus sakei*; Rn, *Rattus norvegicus*; Sa, *Staphylococcus aureus*; Sg, *Streptococcus gordonii*; Sp, *Streptococcus pyogenes*.

[b] FRN, first residue number.

[c] The numbers in parentheses to the right of amino acid sequences are the sequence lengths of each PVA/U34 family dipeptidase and Ntn hydrolase. The catalytic cysteine residue, Arg, and Asp are in boldface type. Positions occupied by mainly hydrophobic residues are underlined twice. Positions with mainly small residues (G, A, S, C, T, V, N, and D) are underlined once (23).

[d] Putative signaling sequence cleavage site.

Sequence analysis of PepD from *B. longum* BORI. The amino acid sequences deduced from the open reading frame were homologous to dipeptidases from various microorganisms, such as *B. adolescentis* (83% identity, gi 119025186), *Oenococcus oeni* PSU-1 (54% identity, gi 116491743), *B. longum* NCC2705 (52% identity, gi 23465037; Table 4, row S), *Lactobacillus casei* ATCC 334 (50% identity, gi 116493607), *Lactococcus lactis* subsp. *cremoris* SK11 (45% identity, gi 116512355), and *Lactobacillus helveticus* CNRZ32 (44% identity, gi 18202547). The sequence analysis also revealed that the G+C content of the *pepD* gene was 63.7%.

Application of the PSI-BLAST program (1) revealed that PepD is a member of the U34 family of peptidases. Although PepD exhibits only 44% identity with PepDA from *L. helveticus* (9), the N-terminus signature of PepD was highly homologous to that of PepDA (Table 4, row B), which is a member of the U34 family of peptidases.

On the other hand, the analysis using the PSI-BLAST program indicated that only amino acids 1 to 428 of the total 533 amino acids in the PepD sequence were conserved in the U34 peptidase family (data not shown). The theoretical molecular mass of these 428 amino acids was 48.3 kDa, which was similar to that of the native peptidase purified from *B. longum* BORI (Fig. 1A). Additionally, MALDI-TOF MS revealed that native pepD purified from *B. longum* BORI had a molecular mass of 49 kDa. These data suggest that the native dipeptidase is subject to a posttranslational modification involving protein splicing, unlike recombinant dipeptidases (rPepD and rPepD-His) expressed in *E. coli*.

DISCUSSION

An endopeptidase gene from *Bifidobacterium animalis* subsp. *lactis* has been identified by sequence similarity to the putative *B. longum* NCC2705 *pepO* gene (accession no. NC_004307). Overexpression of the *pepO* gene in *E. coli* BL21(DE3) revealed the characteristics of the intracellular endopeptidase from *B. animalis* subsp. *lactis* (15).

In this study, a dipeptidase from the cell extract of *B. longum* BORI was purified about 109-fold with an activity yield of 26.8%. The dipeptidase was found to be a monomer with a molecular mass of about 49 kDa. Almost no activity was detected below pH 5.0, but the enzyme showed high activity in alkaline conditions. The activity was about 50% of the maximum at 70°C but was very low at 80°C. Also, the dipeptidase retained 58.3% of its activity after 10 min of preincubation at 80°C. The enzyme was heat stable, unlike a dipeptidase from *Lactobacillus curvatus* DPC2024 (21). The sequence of the first 20 amino acids of the dipeptidase from *L. curvatus* DPC2024 was the same as that of a dipeptidase from *Lactobacillus sakei* subsp. *sakei* 23K, which is a member of the M20 family of peptidases (Table 4, row R). On the other hand, N-terminal amino acid sequencing of both the purified native dipeptidase from *B. longum* BORI and the purified rPepD-His in *E. coli*

was not successful, presumably because of a block, such as N-terminal posttranslational modification. We suggest that the initiator methionine of PepD was hydrolyzed, and N acetylation of the terminal alanine residue to yield acetyl-Ala-Cys-Thr-Thr-protein occurred similarly to that in cellular polypeptide Op18 (14).

Application of the PSI-BLAST program revealed that the PepD sequence was highly homologous to the sequences of the U34 family of peptidases. Also, the N-terminal conserved domain of U34 family members is highly homologous to that of penicillin V acylase (PVA) (Table 4) (23). Proteins of the PVA/U34 family are members of a large superfamily of amidohydrolases called the N-terminal nucleophile (Ntn) hydrolases, which exhibit variable sequences and great diversity in substrate specificity. They include enzymes such as choloylglycine hydrolases, acid ceramidases, and isopenicillin N acyltransferases and a subgroup of eukaryotic proteins with unclear function (16). The most representative sequence for all PVA/U34 family proteins resides in the motif containing the catalytic cysteine residue and corresponds to the N-terminal β-hairpin motif in the structure of *Bacillus sphaericus* PVA (Table 4, row J) (23, 29). Other common features of this motif include the hydrophobic pattern and positions occupied mainly by small residues near the catalytic cysteine. The β-hairpin motif is longer in close homologs of U34 family dipeptidases than in close homologs of PVAs (Table 4). Two other residues (Arg 22 and Asp 25) are also highly conserved in most PVA/U34 proteins. The cysteine side chain serves as the nucleophile, and the free αNH_2 serves as the proton donor and acceptor in the catalytic process. PVA/U34 proteins have been designated as cysteine-type Ntn hydrolases (23). These previous studies suggest that the PepD protein with the same catalytic cysteine also has an N-terminal β-hairpin structure in *B. longum* BORI and thus potentially adopts the Ntn hydrolase fold (Table 4, row A).

On the other hand, crystallization of PVA from *B. sphaericus* showed that the mature form of PVA was created by autoproteolytic processing originating from a nucleophile such as the catalytic cysteine residue (6). In addition, various Ntn hydrolases catalyze internal peptide bond rearrangements through an N→O or N→S acyl shift by a Ntn amino acid created by autoproteolytic processing (24, 25). These previous studies support that the native dipeptidase resulted from protein splicing such as autoproteolytic processing in *B. longum* BORI, since otherwise the molecular mass of rPepD-His expressed in *E. coli* would have been equal to the theoretical mass deduced from the *pepD* sequence. Hence, other factors might be involved in protein splicing in *B. longum* BORI, unlike the case in *E. coli*.

Taken together, these results suggest that the dipeptidase from *B. longum* BORI is a cysteine-type Ntn hydrolase and has a β-hairpin motif similar to that of PVA, which results in it being processed by protein splicing. Currently, the C-type U34 dipeptidases such as this PepD from *B. longum* BORI are assigned to the C69 family of peptidases.

PepD from *B. longum* BORI is the first of the U34 dipeptidases to be purified and characterized and to be potentially processed by protein splicing. Future studies should attempt to elucidate whether other factors are related to protein splicing and whether this enzyme can hydrolyze other substrates such as PVA.

ACKNOWLEDGMENTS

This work was supported by the 21C Frontier Microbial Genomics and Applications Center Program (grant MG01-2068754-10) of the Ministry of Science and Technology, Republic of Korea.

We also thank the Korea Basic Science Institute and the Seoul National University National Center of Interuniversity Research Facilities for performing internal amino acid sequencing and MALDI-TOF MS analyses, respectively.

REFERENCES

1. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
2. **Bockelmann, W., H. P. Beuck, S. Lick, and K. Heller.** 1995. Purification and characterization of a new tripeptidase from *Lactobacillus delbrueckii* ssp. *bulgaricus* B14. Int. Dairy J. **5:**493–502.
3. **Bockelmann, W., M. Fobker, and M. Teuber.** 1991. Purification and characterization of the X-prolyl-dipeptidyl-aminopeptidase from *Lactobacillus delbrueckii* subsp. *bulgaricus* and *Lactobacillus acidophilus*. Int. Dairy J. **1:**51–66.
4. **Bockelmann, W., T. Hoppe-Seyler, and K. J. Heller.** 1996. Purification and characterization of an endopeptidase from *Lactobacillus delbrueckii* subsp. *bulgaricus* B14. Int. Dairy J. **6:**1167–1180.
5. **Booth, M., V. Jennings, I. N. Fhaolain, and G. Ocuinn.** 1990. Prolidase activity of *Lactococcus lactis* subsp. *cremoris* AM2: partial purification and characterization. J. Dairy Res. **57:**245–254.
6. **Chandra, P. M., J. A. Brannigan, A. Prabhune, A. Pundle, J. P. Turkenburg, G. G. Dodson, and C. G. Suresh.** 2005. Cloning, preparation and preliminary crystallographic studies of penicillin V acylase autoproteolytic processing mutants. Acta Crystallogr. Sect. F **61:**124–127.
7. **Cheng, C. C., and T. Nagasawa.** 1985. Purification and some properties of an aminopeptidase from *Bifidobacterium breve*. Jpn. J. Zootech. Sci. **56:**257–266.
8. **Desmazeaud, M. J.** 1974. General properties and specificity of action of a neutral intracellular endopeptidase from *Streptococcus thermophilus*. Biochimie **56:**1173–1181. (In French.)
9. **Dudley, E. G., A. C. Husgen, W. He, and J. L. Steele.** 1996. Sequencing, distribution, and inactivation of the dipeptidase A gene (*pepDA*) from *Lactobacillus helveticus* CNRZ32. J. Bacteriol. **178:**701–704.
10. **Eggimann, B., and M. Bachmann.** 1980. Purification and partial characterization of an aminopeptidase from *Lactobacillus lactis*. Appl. Environ. Microbiol. **40:**876–882.
11. **Elsoda, M., A. Macedo, and N. F. Olson.** 1992. The peptide-hydrolase system of *Bifidobacterium* species. Milchwissenschaft **47:**87–90. (In German.)
12. **Folkertsma, B., and P. F. Fox.** 1992. Use of the Cd-ninhydrin reagent to assess proteolysis in cheese during ripening. J. Dairy Res. **59:**217–224.
13. **Guzman, L. M., D. Belin, M. J. Carson, and J. Beckwith.** 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. J. Bacteriol. **177:**4121–4130.
14. **Hanash, S. M., J. R. Strahler, J. V. Neel, N. Hailat, R. Melhem, D. Keim, X. X. Zhu, D. Wagner, D. A. Gage, and J. T. Watson.** 1991. Highly resolving two-dimensional gels for protein sequencing. Biochemistry **88:**5709–5713.
15. **Janer, C., F. Arigoni, B. H. Lee, C. Pelaez, and T. Requena.** 2005. Enzymatic ability of *Bifidobacterium animalis* subsp. *lactis* to hydrolyze milk proteins: identification and characterization of endopeptidase O. Appl. Environ. Microbiol.**71:**8460–8465.
16. **Kim, J. K., I. S. Yang, H. J. Shin, K. J. Cho, E. K. Ryu, S. H. Kim, S. S. Park, and K. H. Kim.** 2006. Insight into autoproteolytic activation from the structure of cephalosporin acylase: a protein with two proteolytic chemistries. Proc. Natl. Acad. Sci. USA **103:**1732–1737.
17. **Klaver, F. A. M., F. Kingma, and A. H. Weerkamp.** 1993. Growth and survival of bifidobacteria in milk. Neth. Milk Dairy J. **47:**151–164.
18. **Kunji, E. R., I. Mierau, A. Hagting, B. Poolman, and W. N. Konings.** 1996. The proteolytic systems of lactic acid bacteria. Antonie Leeuwenhoek **70:** 187–221.
19. **Laemmli, U. K.** 1970. Cleavage of structural proteins during assembly of the head of bacteriophage T4. Nature **227:**680–685.
20. **Law, J., and A. Haandrikman.** 1997. Proteolytic enzymes of lactic acid bacteria. Int. Dairy J. **7:**1–11.
21. **Magboul, A. A. A., and P. L. H. McSweeney.** 1999. Purification and characterization of a dipeptidase from *Lactobacillus curvatus* DPC2024. Food Chem. **67:**233–240.
22. **Meyer, J., and R. Jordi.** 1987. Purification and characterization of X-prolyl-dipeptidyl-aminopeptidase from *Lactobacillus lactis* and from *Streptococcus thermophilus*. J. Dairy Sci. **70:**738–745.
23. **Pei, J. M., and N. V. Grishin.** 2003. Peptidase family U34 belongs to the superfamily of N-terminal nucleophile hydrolases. Protein Sci. **12:**1131–1135.
24. **Perler, F. B.** 1998. Breaking up is easy with esters. Nat. Struct. Biol. **5:**249–252.

25. **Perler, F. B.** 1998. Protein splicing of inteins and hedgehog autoproteolysis: structure, function, and evolution. Cell **92:**1–4.
26. **Rabier, D., and M. J. Desmazeaud.** 1973. Inventory of different intracellular peptidase activities in *Streptococcus thermophilus*. Purification and properties of a dipeptide hydrolase and an aminopeptidase. Biochimie **55:**389–404. (In French.)
27. **Schell, M. A., M. Karmirantzou, B. Snel, D. Vilanova, B. Berger, G. Pessi, M. C. Zwahlen, F. Desiere, P. Bork, M. Delley, R. D. Pridmore, and F. Arigoni.** 2002. The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. Proc. Natl. Acad. Sci. USA **99:**14422–14427.
28. **Shihata, A., and N. P. Shah.** 2000. Proteolytic profiles of yogurt and probiotic bacteria. Int. Dairy J. **10:**401–408.
29. **Suresh, C. G., A. V. Pundle, H. SivaRaman, K. N. Rao, J. A. Brannigan, C. E. McVey, C. S. Verma, Z. Dauter, E. J. Dodson, and G. G. Dodson.** 1999. Penicillin V acylase crystal structure reveals new Ntn-hydrolase family members. Nat. Struct. Biol. **6:**414–416.
30. **Wohlrab, Y., and W. Bockelmann.** 1992. Purification and characterization of a dipeptidase from *Lactobacillus delbrueckii* subsp. *bulgaricus*. Int. Dairy J. **2:**345–361.
31. **Zourari, A., J. P. Accolas, and M. J. Desmazeaud.** 1992. Metabolism and biochemical characteristics of yogurt bacteria. A review. Lait **72:**1–34.