

# A new organic solvent tolerant protease from *Bacillus pumilus* 115b

Raja Noor Zaliha Raja Abd Rahman · Shalihah Mahamad ·  
Abu Bakar Salleh · Mahiran Basri

Received: 29 November 2006 / Accepted: 1 April 2007 / Published online: 10 May 2007  
© Society for Industrial Microbiology 2007

**Abstract** Five out of the nine benzene–toluene–ethylbenzene–xylene (BTEX) tolerant bacteria that demonstrated high protease activity on skim milk agar were isolated. Among them, isolate 115b identified as *Bacillus pumilus* exhibited the highest protease production. The protease produced was stable in 25% (v/v) benzene and toluene and it was activated 1.7 and 2.5-fold by *n*-dodecane and *n*-tetradecane, respectively. The gene encoding the organic solvent tolerant protease was cloned and its nucleotide sequence determined. Sequence analysis revealed an open reading frame (ORF) of 1,149 bp that encoded a polypeptide of 383 amino acid residues. The polypeptide composed of 29 residues of signal peptide, a propeptide of 79 residues and a mature protein of 275 amino acids with a calculated molecular mass of 27,846 Da. This is the only report available to date on organic solvent tolerant protease from *B. pumilus*.

**Keywords** Organic solvent tolerant protease · Cloning · *Bacillus pumilus* · Isolation

## Introduction

Proteases are among the most valuable catalysts used in food, pharmaceutical and detergent industries because they hydrolyze peptide bonds in aqueous environments and synthesize peptide bonds in microaqueous environments [13]. However, most enzymes including proteases are not stable in organic solvents. Microbial proteases that can function as catalysts in non-aqueous solvents offer new possibilities such as shifting of the thermodynamic equilibria in favor of synthesis, increasing the solubility of hydrophobic substrates, controlling specificity by solvent and improving thermal stability of enzymes. Organic solvents used as the media for enzyme reaction offer important advantages in industrial application of biocatalysis such as chemical processing, food-related conversions and analyses [11]. In addition, solvent-tolerant microorganisms were also useful in biotransformations with the whole cells in two-phase solvent–water systems [9]. For these reasons, the search for proteases, which are naturally stable in the presence of organic solvents is of great importance.

Organic solvent tolerant proteases have greatly expanded their potential for use in the syntheses of useful products. Studies are undertaken not only to optimize the enzyme production but also to enhance our immediate understanding of this enzyme. These studies could provide the foundation for future investigations of the tertiary structure and mechanism of action of the enzymes. Genetic engineering is instrumental in opening new opportunities for the construction of genetically modified microbial strains with selected enzymes properties.

Previously nine putative protease producers were isolated from contaminated soils at a wood factory in Selangor, Malaysia. In this paper, we describe the isolation of a

---

R. N. Z. R. A. Rahman (✉) · S. Mahamad · A. B. Salleh  
Enzyme and Microbial Technology Research,  
Faculty of Biotechnology and Biomolecular Sciences,  
Universiti Putra Malaysia, UPM Serdang,  
43400 Selangor, Malaysia  
e-mail: rnzaliha@biotech.upm.edu.my

M. Basri  
Faculty of Science, Universiti Putra Malaysia,  
UPM Serdang, 43400 Selangor, Malaysia

new organic solvent tolerant protease from *Bacillus pumilus* 115b.

## Materials and methods

### Bacterial strain

The microorganisms used were isolated from contaminated soils of a wood factory in Selangor, Malaysia. The nine isolates were benzene–toluene–ethylbenzene–xylene (BTEX) tolerant bacteria and were proven to be a polycyclic–aromatic–hydrocarbons (PAHs) degraders [18].

### Screening of organic solvent proteolytic microorganism

Five bacteria were screened qualitatively for their protease production on skim milk agar (SMA) containing (g/l): nutrient agar, 13.8 and skim milk powder, 12.0. The agar plates were incubated at 37°C for 24 to 48 h. Microorganisms that produced proteolytic enzyme formed clearing zones around the colonies on the SMA plate. Isolates that showed positive results on the SMA were then tested for their enzyme production in 100 ml liquid medium that consisted of (g/l): bacteriological peptone, 10.0; (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1.0; KH<sub>2</sub>PO<sub>4</sub>, 0.5; MgSO<sub>4</sub>·7H<sub>2</sub>O, 0.3; CaCl<sub>2</sub>·2H<sub>2</sub>O, 1.0; NaCl, 1.0; glycerol, 10.0 ml at pH 7.0. The bacteria inoculums (4 ml of 24 h culture) were inoculated into growth medium (100 ml) and incubated for 24 h under shaking condition (150 rpm) at 37°C. The culture was harvested from the medium by centrifuging at 10,000 rpm and 4°C for 10 min. The supernatant was then filtered with cellulose acetate membrane filter (pore size, 0.22 µm) to obtain the crude enzyme. Organic solvent stability of the enzyme was studied by incubating 3 ml of crude enzyme with 1 ml organic solvent (100% v/v) at 37°C for 30 min with shaking at 150 rpm. The organic solvents used were benzene and toluene with the final concentration of 25% (v/v) in the reaction mixture. Crude enzyme without organic solvent was used as a control. Each experiment was done in triplicates. Protease activity was measured by modification of the method as reported by Keay and Wildi [10].

### Assay of protease activity

The reaction was initiated by the addition of 1.0 ml enzyme previously preincubated at 37°C for 5 min to 1.0 ml casein 2.0% (w/v), pH 7.0. The reaction mixture was then incubated at 37°C for 10 min and the reaction was terminated by addition of 2.0 ml 0.4 M trichloroacetic acid (TCA). A vortex mixer was used to ensure homogenous mixing at various stages of the assay procedure. This

mixture was further incubated at 37°C for 20 min, followed by centrifugation at 13,000 ×g for 10 min and supernatant harvested. To 1.0 ml supernatant, 5.0 ml of 0.4 M Na<sub>2</sub>CO<sub>3</sub> and 1.0 ml of Folin Ciocalteu reagent : water (1:3 v/v) were added to yield a blue colour. The coloured mixture was incubated at 37°C for 20 min before the absorbance was read at 660 nm. Proteolytic (caseinolytic) activity was measured using the modification method as reported by Keay and Wildi [10]. One unit (U) of protease is equivalent to 0.5 µg tyrosine liberated by 1.0 ml enzyme solution under the assay conditions. The amount of tyrosine was determined from the tyrosine standard curve.

### Effect of organic solvents on the protease stability

The two best producers of organic solvent-tolerant protease were selected for further work. Crude enzymes from both isolates were obtained as described in the previous section. To study the stability of the protease in organic solvents, ten organic solvents with different log *P* were tested. Log *P* is defined as the logarithm of the partition coefficient *P*, of the solvent between octanol and water [12]. The organic solvents used were as ethyl acetate, benzene, 1-heptanol, toluene, 1-octanol, ethylbenzene, hexane, 1-decanol, isooctane, *n*-dodecane, *n*-tetradecane, 1-hexadecane. In the stability test, 1.0 ml of organic solvent (100% v/v) was added to 3.0 ml of the crude enzyme solution in a universal bottle and incubated at 37°C with shaking at 150 rpm for 30 min. The remaining proteolytic activity was measured as described above. Each experiment was performed in triplicates. Stability was expressed as the remaining proteolytic activity relative to the solvent-free controls (0%, v/v).

### Effect of temperature on the stability of protease

The crude enzyme was incubated at different temperature ranging from 37 to 60°C with shaking at 150 rpm in the absence of organic solvents. After 30 min, the sample was cooled rapidly in ice and assayed by using casein as substrate at 37°C. The remaining proteolytic activity was measured as described previously. Each experiment was performed in triplicates

### Identification and taxonomical studies

The isolate was identified as *Bacillus pumilus* strain 115b according to the method described in “Bergey’s manual of determinative bacteriology” [7] and also via 16S rDNA gene identification. The 16S rDNA gene was amplified by polymerase chain reaction (PCR) with two universal primers that were designed to amplify a 1538-bp segment of the 16S rDNA gene. The forward primer was: 5′-GAG TTT GAT CCT GGC TCA G-3′ and the reverse primer

was: 5'-CGG CTA CCT TGT TAC GAC TT-3'. PCR was carried out in 100 µl of mixture containing 5 µl of chromosomal DNA (≈100 ng), 3 µl of each 16S forward and reverse primer (20 pmol), 2 µl 10 mM dNTP mix, 2 µl (1.0 U) Taq DNA polymerase, 10 µl 10× Taq buffer, 8 µl 25 mM MgCl<sub>2</sub>, 67 µl sterile distilled water. After 4 min pre-denaturation at 94°C, 30 cycles of PCR including 1 min at 94°C (denaturation), 1 min at 58°C (annealing) and 1 min at 72°C (extension) were performed. This was followed by one cycle of 7 min at 72°C and held at 4°C. The reaction was amplified in thermocycler (GeneAmp PCR system 2,400, Perkin Elmer, Foster, CA). The PCR product was examined by electrophoresis and detected using ethidium bromide fluorescence. The purification of PCR product was carried out according to QIAquick Gel Extraction Kit (Qiagen, Germany) method. After purification, the PCR product was ligated into TOPO TA cloning vector (Introgen, USA) according to the manufacturer's instruction. The ligation mixture was transformed into *E. coli* and plated onto Lubria Bertani/Ampicillin agar containing X-Gal. Plasmids from positive colonies were extracted and sequenced. The sequence or DNA homology was matched with the Genbank database that was available in the net. The URL for this database is: <http://www.ncbi.nlm.nih.gov/BLAST/>.

#### Phylogenetic tree analysis

The phylogenetic tree was constructed based on comparison of 16S rDNA sequences of *Bacillus pumilus* 115b strain with other strains of *Bacillus* species that were extracted from GeneBank database (<http://www.ncbi.nlm.nih.gov>). All sequences were aligned with CLUSTALW (Multiple Sequence Alignment) that was obtained from: <http://seqtool.sdsc.edu/CGI/BW.cgi> and phylogenetic tree was constructed in GeneBee database (<http://genebee.msu.su>).

#### Amplification of protease gene by PCR

Four primers were designed based on nucleotide sequence from an alkaline serine proteinase of *Bacillus pumilus* TYO-67 [1] for detection and amplification of protease gene from *Bacillus pumilus* 115b. One pair of the primer, forward F1: 5'-ATG TGC GTG AAA AAG AAA AAT GTG -3' and reverse R1: 5'-TTA GTT AGA AGC TGC TTG AAC GTT-3' was designed based on open reading frame (ORF) and another pair, forward F2: 5'-GTC CTT GAT ACT GGT ATC CAC GCC-3' and reverse R2: 5'-AGG AGA TGC CAT CGA TGT TCC CGT-3' based on their conserved region nucleotide [412–977 (Val 138 -Pro 332)]. Another set of primer F4 (5'-CTA TCT CTA TTA AAC TGA AAA TAC-3') and R4 (5'-GGC GTG GAT

ACC AGT ATC AAG GAC 3') was then designed to amplify the upstream and downstream sequence of ORF gene.

Amplification of protease gene by PCR was performed in 100 µl containing 4 µl DNA template, 10 µl 10× PCR buffer, 6 µl 25 mM MgCl<sub>2</sub>, 2 µl 10 mM dNTP mix, 3 µl 20 pmol of each forward and reserve primers, 2 µl (1.0 U) Taq DNA polymerase and 70 µl sterile distilled water. The reaction mixture was amplified in a thermocycler and started with 4 min pre-denaturation at 94°C and followed by 30 PCR cycles, which comprised 1 min denaturation at 94°C, 2 min annealing at 55°C and 1 min extension at 72°C. The final extension was carried out at 72°C for 7 min and held at 4°C. The amplified PCR product was examined by electrophoresis and visualized under UV radiation after staining with ethidium bromide (1 mg/ml).

#### Cloning of organic solvent tolerant protease

The purified PCR product was cloned into pCR<sup>®</sup> 2.1-TOPO<sup>®</sup> vector by using the TOPO TA Cloning<sup>®</sup> (Introgen, USA). Fresh PCR product (4 µl) was ligated with 1 µl salt solution into 1 µl TOPO vector and incubated for 5 min at room temperature. Then, the ligation mixture was transferred into a vial containing chemically competent *E. coli* TOP 10 and mixed gently. The mixture was incubated on ice for 30 min and heat shocked for 30 s at 42°C without shaking, then immediately transferred back to ice. A 250 µl SOC medium was added to the tube and agitated horizontally at 37°C for 30 min. The 50 µl mixture was spread on skim milk/ampicillin agar and incubated overnight at 37°C. Plasmid obtained from positive colonies was then extracted and sequenced.

#### DNA sequencing and computer analysis of the protease gene

Recombinant plasmids containing inserts were sent for automated sequencing by using universal or synthetic oligonucleotides M13 forward and reverse sequencing primer. The sample was sequenced using ABI PRISM 377 Genetic Analyzer (Perkin-Elmer). The sequence and database similarity were analyzed by using BLAST from National Center of Biotechnology, Bethesda, MD (<http://www.ncbi.nlm.nih.gov>) and Biology Workbench (<http://biology.ncsa.uiuc.edu>). The translation of nucleotide sequence was performed with Expasy Molecular Biology server (<http://www.expasy.com>).

#### Nucleotide sequence accession number

The nucleotide sequence of *Bacillus pumilus* 115b protease gene and 16S rRNA were deposited to the GenBank

database under accession number AY743586 and AY740598, respectively.

## Results and discussion

### Isolation and screening of organic solvent tolerant proteolytic microorganisms

Since 1995, a few organic solvent tolerant proteases from *Pseudomonas* species were reported [6, 14]. However, only one report is available on *Bacillus* species producing an organic solvent-tolerant protease [5]. In the search of organic solvent tolerant protease producers, five bacteria were isolated. All five bacteria showed positive results by forming lysis zones around the colonies on SMA and secreted their enzyme production in liquid medium. Two best isolates were selected based on their high proteolytic activity in 25% (v/v) benzene and toluene. They were isolates 146 and 115b.

### Effect of organic solvents and temperature on the stability of crude protease

The effect of organic solvents on protease activity differs from protease to protease. Ten organic solvents with different log *P* were used in this study. Table 1 shows that protease from isolate 115b was inactivated in the presence of ethyl acetate (log *P* 0.68), benzene (log *P* 2.0), 1-heptanol (log *P* 2.4), toluene (log *P* 2.5), and ethylbenzene (log *P* 3.1). Similarly, protease from isolate 146 was not stable in organic solvents with log *P* values between 0.68 and 4.0.

**Table 1** Effect of various organic solvents on the stability of crude protease activity from isolate 115b and 146

Organic solvents	log <i>P</i>	Relative activity (%)	
		Isolate 115b	Isolate 146
None (cell-free supernatant)	–	100	100
Ethyl acetate	0.68	10	7
Benzene	2.00	80	46
1-Heptanol	2.40	54	74
Toluene	2.50	63	41
Ethylbenzene	3.10	90	75
Hexane	3.50	134	98
1-Decanol	4.00	128	144
Isooctane	4.50	139	129
<i>n</i> -Dodecane	6.60	166	155
<i>n</i> -Tetradecane	7.60	248	211

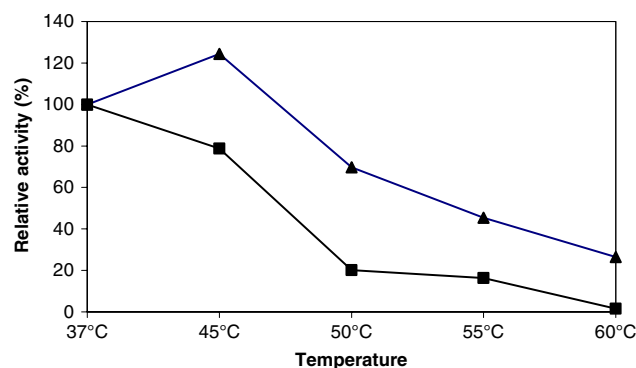
One ml organic solvent was added to 3 ml of crude enzyme (cell-free supernatant) and inoculated at 37°C, 150 rpm for 30 min. Each experiment was done in triplicate

However, when hexane, 1-decanol, isooctane, *n*-dodecane, and *n*-tetradecane, of which the log *P* values are equal to or more than 3.5 were added, the activities of both proteases were enhanced. The presence of *n*-dodecane (log *P* 6.6) and *n*-tetradecane (log *P* 7.6) increased the activity of isolates 115b and 146 to more than 1.5 and 2.0- fold, respectively. An organic solvent tolerant protease from *Pseudomonas aeruginosa* PST-01 exhibited a similar stability pattern whereby it was stable in the presence of organic solvents with the log *P* of 3.2 or higher. Protease from *Pseudomonas aeruginosa* strain K [4] was also activated in the presence of organic solvents, with log *P* values, equal to or higher than 4.0. The lower the log *P*, the greater the polarity value and the greater the toxicity of the solvent to the enzyme. Generally, solvents with log *P* values below 4 are considered extremely toxic as their degrees of partitioning into the aqueous layer are higher [3, 8]

Studies of the effect of temperature on the stability of both proteases were carried out at temperatures ranging from 37°C to 60°C (Fig. 1). Crude protease was incubated in water bath at different temperature for 30 min at 150 rpm in the absence of organic solvents. Protease from isolate 115b was very stable at 45°C but the stability decreased dramatically by 30% at 50°C. However, protease from isolate 146 was not stable at 45°C. The activity decreased by 21 and 80% at 45°C and 50°C, respectively. From the organic solvent and temperature stability tests, isolate 115b was shown to produce an organic solvent tolerant and thermostable protease. Isolate 115b was then identified and the gene encoding protease was cloned.

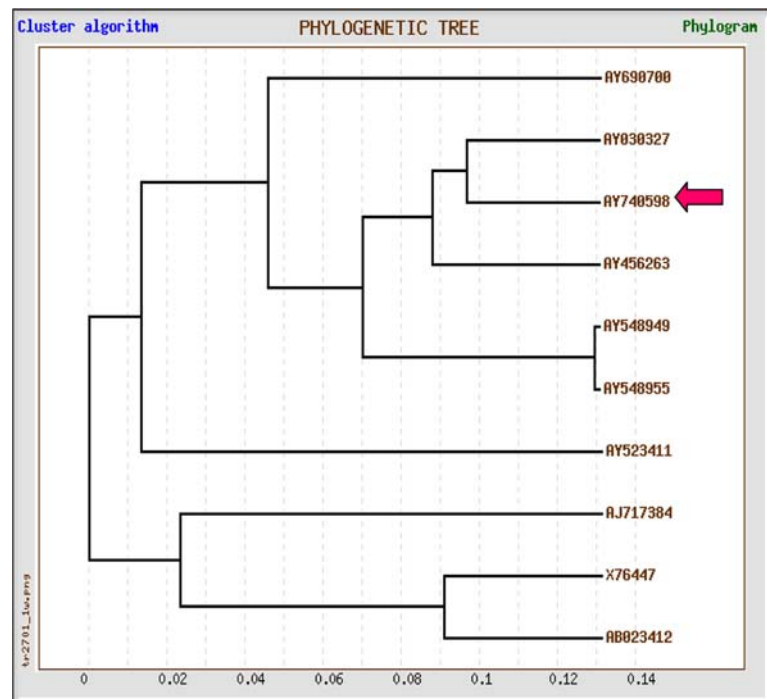
### Identification and phylogenetic analysis

Isolate 115b was characterized by physiological and biochemical characteristics and also via 16S rDNA sequence



**Fig. 1** Effect of temperature on stability of protease activity from isolate 115b and 146. Crude enzyme from isolate 115b (filled triangle) and isolate 146 (filled square) incubated at different temperature for 30 min. The treated enzyme solution was cooled rapidly in ice and the relative activity was measured under standard conditions

**Fig. 2** Rooted phylogenetic tree showing the relationship between isolate 115b to other *Bacillus* species. AY690700 (*Bacillus* sp. MH06), AY030327 (*B. pumilus* KL-052), AY740598 (*B. pumilus* 115b), AY456263 (*B. pumilus* DMSZ27), AY548949 (*B. pumilus* 8N-4), AY548955 (*B. pumilus* S9), AY523411 (*B. catenulatus*), AJ717384 (*B. firmus* CV93b), X76447 (*B. halmapalus* DSM 8723) and AB023412 (*B. cohnii* YN-2000). Arrow indicates 16S rDNA sequence from *B. pumilus* 115b



analysis. Colonies were small, irregular, convex, lobate and rough on nutrient agar. The bacterium was gram positive, motile, straight rods, and occurred in single or pairs with terminal spore. Biochemical tests indicated that isolate 115b showed a typical characteristic of the genus *Bacillus*.

Isolate 115b was further identified as *Bacillus pumilus* strain 115b via 16S rDNA gene sequencing. Comparison of its 16S rRNA sequence among available stains of *B. pumilus* showed a high homology (98%) to strain KL-052. A phylogenetic tree was constructed from the available 16S rDNA sequences of *B. pumilus* strains and a few selected mesophilic *Bacillus* species. Figure 2 shows all *B. pumilus* stains are clustered together and phylogenetically distant from other mesophilic *Bacillus* species. Among the available *B. pumilus* strains, *B. pumilus* 115b (AY740598) was found to be closely related to *B. pumilus* KL-052 (AY030327).

Cloning and sequence analysis of an organic solvent tolerant protease from *Bacillus pumilus* 115b.

The cloning of an organic solvent tolerant protease gene from *Bacillus pumilus* 115b, was performed by PCR using primers based on gene encoding of an alkaline serine proteinase of *B. pumilus* TYO-67 [1]. After PCR amplification of the DNA fragment, it was ligated into the pCR<sup>®</sup> 2.1-TOPO<sup>®</sup> cloning vector and transformed into *Escherichia coli* TOP 10 cells. The first set of primers forward F1 primer and reverse R1 primer was designed to amplify the full-length of open reading frame (ORF) of an organic

solvent tolerant protease gene from *B. pumilus* 115b. The other pair, forward F2 primer and reverse R2 primer was to amplify the consensus regent segment of protease gene. These sets of primer, F1 and R1, F2 and R2, F1 and R2, and F2 and R1 had successfully amplified 1152 bp, 588 bp, 999 bp and 741 bp of PCR products, respectively. Another set of primer F4 and R4 was then designed and was shown to successfully amplify the upstream and downstream sequence of ORF gene. The putative promoters (–10 and –35), ribosome-binding site of the mature protein and putative transcription termination that stopped the chain growth and released the mRNA and polymerase were encoded in this region.

The sequence analysis revealed an open reading frame (ORF) of 1,149 bp encoding 383 amino acids (Fig. 3). The calculated molecular mass of protein was 39,448 Da. A signal peptide (pre-sequence) of 29 amino acids and a long peptide (pro-sequence) of 79 amino acids preceded the sequence of mature proteins. The mature protein comprised 275 amino acids with a calculated molecular mass of 27,846 Da.

Signal and Shine–Dalgarno sequence were determined according to the rules described by Von Heijine [19] and Stormo et al. [17], respectively. The presumed putative promoter regions, –10 (ATAATC) and –35 (TTAAAC) sequenced resembled the consensus sequences for the promoter region recognized by lambda P<sub>R</sub> RNA polymerase of *Escherichia coli*. Dale and von Schantz [2] previously reported the promoter region to be ATAATG and TTGACT.

**Fig. 3** Nucleotide and deduced amino acid sequence of organic solvent tolerant protease *Bacillus pumilus* 115b. (Asterisk) indicate a possible Shine–Dalgarno sequence. The putative –35, –10 promoters, start and stop codon and the stem loop inverted repeat termination sequences were *bold*

```

-35          -10
ttaactgaaaatacagaataatc -91
      ****
aacggatcattctaatagactacggatgattattctgaaataagaaaaagggatgtgga -60
Start
atgtgcgtagaaaagaaaatgtgatgacaagtgttttatggctgtccccccttctgttt 60
M C V K K K N V M T S V L L A V P L L F 20
tcagctgggtttggaggaacctggcaaatgccgagacggtctccaaaacagatagtgaa 120
S A G F G G T M A N A E T V S K T D S E 40
aaaagctatatgttgggttttaagcctctgccaccacaaacagctctaagaacaagct 180
K S Y I V G F K A S A T T N S S K K Q A 60
gtcattcaaaatggtgaaaactagaaaaacaataccgcctcattaatgtgcacaaagtg 240
V I Q N G G K L E K Q Y R L I N A A Q V 80
aaaatgtctgaacaagccgcaaaaaacttgaacatgacccatgacattgcttacgtagaa 300
K M S E Q A A K K L E H D P S I A Y V E 100
gaagaccataaagcagaagcatatgcacaaaccgttccttatggaatccctcaaatcaaa 360
E D H K A E A Y A Q T V P Y G I P Q I K 120
gctccagctgtacacgctcaaggttataaagggtgctaattgtcaaagtagccgtccttgat 420
A P A V H A Q G Y K G A N V K V A V L D 140
actggaatcccagctgcacaccctgacttaaatgttcaggcggtgcgagcttcgcccct 480
T G I H A A H P D L N V A G G A S F V P 160
tcagagccaaatgccacccaagactttcaatcacatggaactcacgtagctggaaccatt 540
S E P N A T Q D F Q S H G T H V A G T I 180
gctgcccttgataacacaatgggtgttcttgggggtgctccaaacgcttcctctatatgct 600
A A L D N T I G V L G V A P N A S L Y A 200
gtgaaagtattagaccgtaatggtagcggacaatacacagctggattattagcggatttgaa 660
V K V L D R N G D G Q Y S W I I S G I E 220
tgggctgtagctaaataatggatgtcatcaatatgagcttaggcggaccaagcggttca 720
W A V A N N M D V I N M S L G G P S G S 240
acagcgcttaaaatgccgtggatacagcgaataatcgtggagtcggtgtgtggcggcc 780
T A L K N A V D T A N N R G V V V V A A 260
gcaggtaattctggctctagtggtcttagaagtacagttggctatcctgcaaaatcagat 840
A G N S G S S G S R S T V G Y P A K Y D 280
tctacaattgccgttgccaatgtaaacagtagcaatgtcagaaactcatcttcttagcgca 900
S T I A V A N V N S S N V R N S S S S A 300
ggtcctgaattagatgtttctgcacctgggtacttctattttaagtacagtccaagcagc 960
G P E L D V S A P G T S I L S T V P S S 320
ggatacacatcttactggaacatctatggcgtctcctcatgtagcaggagcagcagcg 1020
G Y T S Y T G T S M A S P H V A G A A A 340
cttattctttctaaaaaccggaatctaacaattcacagggttcgacgcttagaaaaat 1080
L I L S K N P N L T N S Q V R Q R L E N 360
acagcaacaccgcttggtgactcattctattacgaaaagggttaatacaagctcaagca 1140
T A T P L G D S F Y Y G K G L I N V Q A 380
gcttctaactaccaagtgatgataaaaaaatccggctgaataagccggatttttattta 1200
A S N Stop          ←—————→          ←—————→          383
ttactgctttttaggcgtttttctctttttccgctgcattttccatattccatgattgctc 1260
catcatggtcgggtgtccataacgaaaaatcttactagaaaacgcggttaccctgaga 1320
cagccctgattttgcaagcttttgaagaagcgttacgcctgcctcccctcttttg 1376

```

Homology search revealed that the *B. pumilus* 115b protease gene shared a high homology of 90% with two alkaline serine proteases from *B. pumilus* TYO-67 and *B. pumilus* UN-31-C-42. The essential amino acids in the catalytic center of subtilisin-like serine protease or subti-

lase are D<sup>140</sup>, H<sup>172</sup> and S<sup>329</sup> [16] and well conserved among them. Siezen and Leunissen [16] investigated the consensus segments of 20–40 amino acids around the above three residues in subtilases and divided them into families.

**Fig. 4** Comparison of the deduced amino acids sequence of 115b protease with other subtilisins. Subtilisin from *B. subtilis* (D25319); *Geobacillus stearothermophilus* (M64743); *B. subtilis* (AJ579472); *B. subtilis* ATCC 6633QK02 (AJ539133); *B. subtilis* K-54 (AF093112); *B. pumilus* TYO-67 (AB029082); *B. pumilus* (AJ539133); UN-31-C-42 (AY458140); *B. pumilus* 115b (AY743586); *B. licheniformis* 15413 (X91262); *Bacillus* sp. KSM-LD1 (AB096689). Subtilisins from *B. pumilus* were bold and arrow indicates putative three catalytic residues

Table with 2 columns: Accession numbers and Amino acid sequences. The sequences are aligned to show conserved regions. Three catalytic residues are highlighted in green and marked with red arrows. Conserved regions are marked with asterisks. The sequences include: D25319, M64743, AJ579472, AJ539133, AF093112, AB029082, AY458140, AY743586, X91262, AB096689. The sequences are: -RVRSKLWISLLFALTLI...; VISEKGGKVQKQFKYV...; KAPALHSQGYTGSNVK...; IAAALNNSIGVLGVP...; STALKTVVDKAVSSG...; SVGSELVDMAPGVSI...; ESTATYLGNSFYYGK...

The predicted amino acid protease sequence from 115b shared an extensive homology to the subtilisin family of protease. Homology search was performed by using NCBI

BlastP at <http://www.ncbi.nlm.nih.gov> and Biology Workbench 3.2 / Clustal W at <http://www.biology.ncsa.uiuc.edu>. A comparison of the deduced amino acids sequence

quence of 115b protease with other subtilisins from *Bacillus* species is as shown in Fig. 4. The consensus sequence are D<sup>140</sup>, H<sup>172</sup> and S<sup>329</sup> of 115b protease showed high identity with alkaline serine protease from *B. pumilus* TYO-67 and *B. pumilus* UN-31-C-42. Glycine residues G<sup>142</sup>, G<sup>173</sup>, G<sup>191</sup>, G<sup>262</sup>, and G<sup>327</sup> [17] that were highly conserved in most of subtilases were totally conserved in 115b protease. Therefore, this indicates that 115b protease belongs to the true subtilisin subgroup of the subtilisin family of the subtilase superfamily.

#### Analysis of amino acid sequence of an organic solvent tolerant protease gene

A comparison of the amino acid from *Bacillus pumilus* 115b protease with the *B. pumilus* TYO-67 protease and *B. pumilus* UN-31-C-42 protease indicated that there is a difference of 17 and 16 amino acids, respectively among them. Although these proteases shared a high homology, the characteristics of these proteases are somehow different to one another. Protease from *B. pumilus* 115b is tolerant to organic solvents and as far as we know, is the first of such protease reported from this *Bacillus* species. Meanwhile, alkaline serine protease from *B. pumilus* UN-31-C-42 was reported to possess a high dehairing activity and low collagen-degradation [15] while alkaline serine protease from *B. pumilus* TYO-67 was used in coagulating soybean milk [20]. Both proteases were not reported to be organic solvent tolerant enzymes.

ProtParam tool at <http://www.au.expasy.org/cgi-bin> was used to determine the amino acid composition of the organic solvent tolerant protease 115b. Out of the 383 amino acid residues deduced from *Bacillus pumilus* 115b, 26 residues were negatively charged (Asp + Glu), while 37 residues were positively charged (Arg + Lys + His). The total charged residues (Asp, Glu, Arg, Lys, His) were 63 amino acids or 16.44% of the total amino acids. The Ala, Ile, Leu, Met, Phe, Pro, Trp and Val, which were hydrophobic residues make up to 161 amino acids (42.04%) and the total uncharged residues (Asn, Cys, Gln, Gly, Ser, Thr, and Tyr) were 159 or 41.52%. Polarity plays an important role in the solubility of the enzyme, in which the polar amino acids may interact with water molecules. The instability index (II) of 115b protease gene was 29.59 and could be classified as a stable protein. Enzymes are considered as unstable when the instability index is higher than 40.

Analysis of the amino acid compositions indicated that Ala was the most abundant, accounting for up to 13% of the total amino acids in 115b protease sequence. Similarly, both *B. pumilus* TYO-67 protease and *B. pumilus* UN-31-C-42 protease were also reported to have an abundant of Ala in their amino acids sequences. Besides Ala, Ser also

made up a high proportion in the amino acid composition (11.2%) in 115b protease with a theoretical *pI* of 8.7 and therefore, 115b protease could be classified as alkaline serine protease.

The hydrophobic profile of 383 amino acid from 115b protease was performed by using ProtScale tools with Kyte and Doolittle method at <http://www.au.expasy.org/cgi-bin>. Hydrophobicity of a protein is a significant indicator in identifying signal sequences that encoded for a signal peptide and affect the stability of the protein folding. The amino acid residues from 10 to 22 (T-S-V-L-L-A-V-P-L-L-F-S-A), 256 to 260 (V-V-V-A-A) and 337 to 342 (G-T-S-M-A-S) were hydrophobic grooves. This first region was predicted to be a signal peptide that was predicted from deduced amino acids sequence with SignalP V2.0 (Expasy). The signal peptide varied between 10 and 22 amino acid residues in length and had no consensus sequence. The region M1 to A22 met these criteria of the signal peptide and was predicted to be the putative signal peptide. This region was very hydrophobic.

In conclusion, the study reported on the isolation of an organic solvent tolerant protease from *Bacillus pumilus* 115b and its stability in various organic solvents. The bacterium itself exhibited tolerance towards BTEX. This work may offer the opportunity to discover the molecular mechanism of the proteases in the presence of organic solvents. This study will eventually contribute to the application of biocatalysis in organic media.

**Acknowledgments** This research project was supported by research grants from the Ministry of Science, Technology and Innovation (MOSTI), Malaysia (IRPA Grant No: 09-02-04-0763-EA001).

#### References

1. Aoyama M, Toma C, Yasuda M, Iwanaga M (2000) Sequence of gene encoding an alkaline serine proteinase of *Bacillus pumilus* TYO-67. *Microbiol Immunol* 44(5):389–393
2. Dale JW, von Schantz M (2002) From genes to genomes: concepts and application of DNA technology. Wiley, England pp 5–304
3. De Bont JAM (1998) Solvent-tolerant bacteria in biocatalysis. *Trends Biotechnol* 16:493–499
4. Geok LP, Razak CNA, Rahman RNZA, Basri M, Salleh AB (2002) Isolation and screening of an organic solvent-tolerant protease producer. *Biochem Eng J* 36:63:1–5
5. Ghorbel B, Sellami-Kamoun A, Nasri M, (2003) Stability studies of protease from *Bacillus cereus* BG1. *Enzyme Microb Technol* 32:513–518
6. Gupta A, Roy I, Khare SK, Gupta MN (2005) Purification and characterization of a solvent stable protease from *Pseudomonas aeruginosa* PseA. *J Chromatogr A* 1069:155–161
7. Holt JG, Sneath PHA, Krieg NR, Staley JG, Williams ST (1994) Group 18: endospore-forming gram-positive rods and cocci. In: Hensyl WR (ed) *Bergey's manual of determinative bacteriology*. Library of congress cataloging-in-publication data, United States, pp 559–564



8. Isken S, De Bont JAN (1998) Active efflux of organic solvents by *Pseudomonas putida* is induced by solvents. J Bacteriol 180:6769–6772
9. Isken S, Derks A, Wolffs PFG, De Bont JAN (1999) Effect of organic solvent on the yield of solvent-tolerant *Pseudomonas putida* S12. Appl Environ Microbiol 65(6):2631–2635
10. Keay L, Wildi BS (1970) Proteases of genus Bacillus. I: Neutral proteases. Biotechnol Bioeng XII:179–212
11. Klivanov AM (1997) Why are enzymes less active in organic solvents than in water? Trends Biotechnol 15:97–101
12. Laane C, Boeren S, Vos K, Veeger C (1987) Rules for optimization of biocatalysis in organic solvents. Biotechnol Bioeng 30:81–87
13. Ogino H, Watanabe F, Yamada M, Nakagawa S, Hirose T, Noguchi A, Yasuda M, Ishikawa H (1999) Purification and characterization of organic solvent-stable protease from organic solvent-tolerant *Pseudomonas aeruginosa* PST-01. J Biosci Bioeng 87:61–68
14. Ogino H, Yasui K, Shiotani T, Ishihara T, Ishikawa H (1995) Organic solvent-tolerant bacterium, which secretes an organic solvent-stable proteolytic enzyme. Appl Environ Microbiol 61:4258–4262
15. Pan J, Huang Q, Zhang Y (2004) Gene cloning and expression of an alkaline serine protease with dehairing function from *Bacillus pumilus*. Curr Microbiol 49:165–169
16. Siezen RJ, Leunissen JAM (1997) Subtilases: the superfamily of subtilisin-like serine proteases. Protein Sci 6:501–523
17. Stormo GD, Schneider TD, Gold LM (1982) Characterization of translational initiation sites in *E. coli*. Nucleic Acids Res 10:2971–2996
18. Tan PS (1999) Screening and isolation of polycyclic aromatic hydrocarbons (PAHs) degrading bacteria from contaminated sites. BS Thesis, Universiti Putra Malaysia, Malaysia
19. Von Heijne G (1986) A new method for predicting signal sequence cleavage sites. Nucleic Acids Res 14:4683–4690
20. Yasuda M, Aoyama M, Sakaguchi M, Nakachi K, Kobatomo N (1999) Purification and characterization of a soybean-milk-coagulating enzyme from *Bacillus pumilus* TYO-67. Appl Environ Microbiol 65:474–479