

Identification of a Putative Keratinase Gene and Analysis of a Peptidase S8 Family from a Hyperthermophilic Isolate, *Fervidobacterium* sp. Strain FC2004 in Thailand

**Urarux Romruen¹, Juan M. Gonzalez²
and Wirojne Kanoksilapatham^{1*}**

¹*Department of Microbiology, Faculty of Science, Silpakorn University, Nakhon Pathom 73000, Thailand*

²*IRNAS-CSIC, Avda. Reina Mercedes 10, 41012 Sevilla, Spain*

**Corresponding author. E-mail: Kanoksilapatham_w@su.ac.th*

ABSTRACT

*A hyperthermophilic *Fervidobacterium* sp. strain FC2004 was previously isolated from a hot spring in Thailand. A putative gene encoding a proenzyme named in this study “ProA1” was identified from shotgun sequencing of genomic DNA of the strain FC2004. Predicted 3D-molecule contains a non-homologous signal peptide, a propeptide domain (PD), a catalytic domain (CD) with the typical catalytic triad residues of D169, H207 and S379, and a substrate binding domain (SD). Unlike fervidolysin and islandisin, the ProA1 completely lacks the SD2 domain. Phylogenetic analysis suggests that the ProA1 might be an intermediate isoform between high molecular weight and low molecular weight peptidase S8_subtilases. Although, the strain FC2004 is able to degrade pieces of native feather at high temperature, whether or not the mature ProA1 is active on keratin, the SD plays a role in hydrolyzing keratin substrate and the strain FC2004 might carry a second serine protease with two SDs remains to be investigated.*

Keywords: *Fervidobacterium*, Hyperthermophile, Keratinase, Serine protease, Thermostable enzyme

INTRODUCTION

Fervidobacterium sp. FC2004, a hyperthermophilic heterotroph belonging to order Thermotogales growing by peptide fermentation was isolated from a hot spring in Northern Thailand. The bacterium grew optimally at 75-80 °C, pH 7.5 and able to degrade native feather at growth temperature (Keawram and Kanoksilapatham, 2013). Proteases or peptide hydrolases can be categorized into four classes; cysteine, serine, metallo and aspartic proteases. Serine proteases consist of a serine residue in conserved catalytic triad which can be classified based on structures to be “chymotrypsin-like” and “subtilisin-like” (or known as “subtilase”). Superfamily of serine proteases has been divided into clans; subtilisin, thermitase, proteinase K, lantibiotic peptidase, kexin and pyrolysin (Siezen and Leunissen, 1997). The clan subtilisin comprises S8 (D, H and S) and S53 (E, D and S) subtilase families. The typical subtilisins are non-specific low molecular weight (20 - 45 kDa) enzymes and named in this study “small S8_subtilase”. Unlike other low molecular weight enzymes, molecular weight of keratinolytic enzymes from thermophilic *Fervidobacterium* spp. are approximately double in size or larger and named in this study “large S8_subtilase”. A cell bound keratinases (or fervidolysin) from a thermophilic *F. pennivorans* was revealed to be a serine protease with a molecular weight of 130 kDa (Friedrich and Antranikian, 1996). A 2.1-kb *fls* gene encoding fervidolysin, a 699 amino acids prosubtilase was cloned and expressed in *E. coli*. Autoproteolysis of subtilase precursor (73 kDa) was evidenced to a 58 kDa mature enzyme and a 14 kDa propeptide (Kluszens et al., 2002). A homomultimeric membrane bound keratinase (or islandisin)

native enzyme from *F. islandicum* (>200 kDa; 97 kDa subunits) was proved to be active optimally at 80 °C and very thermostable without detectable loss of activity after incubation at 80 °C for > 32 h (Nam et al., 2002; Godde et al., 2005).

In this study, a putative keratin-degrading gene was identified in the genomic DNA of a hyperthermophilic *Fervidobacterium* sp. strain FC2004. This study reports an intermediate size protease that is unique to this strain.

MATERIALS AND METHODS

Organism and cultivation

Fervidobacterium sp. strain FC2004 was employed. This strain is equivalent to strains JMC 18757, ATCC BAA-2483 and *Fervidobacterium thailandense* strain FC2004^T (Kanoksilapatham et al., 2016) Cultivation was performed in 480GM5 medium at 80 °C. Composition of 480GM5 medium is similar to the previously described 408G medium except that 5 g/L pancreatic digest of casein was used (Keawram et al., 2016).

Gene amplification

An ORF encoding a protease named in this study “ProtA1” was identified from shotgun sequencing of genomic DNA. Two specific primers, ProT1AF (5' CATA TGCG TAGA CCCG TTAA CGTC 3') and ProT1AR (5' TTAC TATA GTTC AACT TCAA TTTG CAC 3') were designed. PCR product sized of 1569 bp was amplified.

Sequence similarities and protein structure

A numbers of peptidase_S8 family serine protease system are identified in the complete genome sequences of *Fervidobacterium pennivorans* DSM 9078^T, *F. nodosum* Rt17-B1^T, *F. islandicum* AW1 and *F. pennivorans* DYC that can be grouped in this study according to sizes (>680 and <450 amino acids) and named “large S8_subtilase” and “small S8_subtilase” mentioned previously. All protein sequences were obtained from

NCBI. Two protein ID of large S8_subtilases including WP_104451857.1 (699 aa) and WP_014451869 (697 aa) from *F. pennivorans* DSM 9078^T, WP_011994230.1 (682 aa) and WP_011994224.1 (710 aa) from *F. nodosum* Rt17-B1^T, WP_052107197.1 (701 aa) and WP_033191969.1 (699 aa) from *F. islandicum* AW-1, and WP_064012406.1 (695 aa) and WP_064012395.1 (688 aa) from *F. pennivorans* DYC were chosen. WP_014451703.1 (439 aa), WP_011993735.1 (440 aa), and WP_033191846.1 (439 aa) representing small S8_subtilase sequences were employed. 3D protein structure was generated using SWISS-MODEL software (<http://swissmodel.expasy.org/>).

RESULTS

Amplification of keratinase gene

Shotgun sequencing from genomic DNA of *Fervidobacterium* sp. strain FC2004 revealed an ORF containing a putative preprosubtilase (522 amino acids). Nucleotide sequence is shown in Figure 1. BlastP analysis revealed remarkably high similarity to peptidase S8 from *F. pennivorans* DSM 9078^T (WP_064012406.1; 52 %), *F. islandicum* AW-1 (WP_052107197.1; 48 %) and *F. nodosum* Rt17-B1^T (WP_011994230.1; 48 %) and subtilase from *Thermosiphon africanus* (WP_004102466.1; 47%).

```

GTGCGTAGACCCGTTAACGCTCTTAATCTTGGTTTTTTGGTTCCATCGTAGCCACGTTTTCCCTGCGTGAAGTTTCCAGATACCAAACCGCTGGTGAAGAA 102
V R R P V N V L I L V F L V S I V A T F S C V K F P D T K T A G E E 34
GTGTTTCCTGAACAGTCTCGATCATCCGACAGCAGGTAAACTCTTGGTTGGATTTCCGAAGAAGCTGCGGTTTACGAACTTGCCACGCGAACTTGGTGGC 204
V F L N S L D H P H E Q G K L L V G F S E E A A V Y E L A R E L G A 68
GTGGTTTTGGGTGTGGATAAGGTTTTAAATTTGCGGCACCTCGGTGTGCGCGAAGAARCTGGATAAGGTTTACGCTAAACTGAAGAATTTAAGAATCGAAGGT 306
V V L G V D K V L K F A A L G V R E E L D K V Y A K L K N L R I E G 102
GTGACGTACGTTGAACCGAGTTATGTCAGAACTATTCACATCCGTTCTGAGAGTTCCCACTCCGCGAGTAACCGATGCTGGCACCGACCTGAAAAATCTTGAG 408
V T Y V E P S Y V R T I P S V L E L P T P R V T D A G T D L K I L E 136
AACGAGAGTAATAACCTTTGGGCTTGCAGTAACGGAAATCAAGAAGCGGTGGGAGCTTGGATTTACCGGTGACGCGCTAATCGTCGCGGTTTTGGACACG 510
N E S K Y L W G L Q V T E I K K A W E L G F T G D G V I V A V L D T 170
GGGGTAGACGGTACACATCCAGATCTCGCACCAACCGTGATAAAAGGTACAAACGCGGTCAACGGAAGCGAGATTGCACCGTCAACTGATAGTTCCGATAGGT 612
G V D G T H P D L A P N V I K G Y N A V N G S E I A P S T D S S I G 204
GGTGCACACGGAAACACATGTGGCCGATACGATGCGCCGCTTTTGGATGGAAAGGGAGTTGTTGGCGTTGACCGAAAAGCGAAGATCATGCCAGTGGTGGT 714
G A H G T H V A G T I A A V L D G K G V V G V A P K A K I M P V V I 238
TTCCGGAGTTGGTACGTCGCGGATGATAAAGTTGCAGAGGCTATCCGCTGGGCGGTACAAAACGGGGCCAAAGTGTGAGCAACTCTGGGGTGGGATGGGT 816
F G S W Y V G D D K V A E A I R W A V Q N G A K V L S N S W G G M G 272
TATTCGATGACGCTAAAACGGGCGATTGACTATGCGTTGGAGAACGTTGATGTTGTCGTCGCGCCGCTGGAACAGTTCGCTTACCAATCGTCATTGTAT 918
Y S M T L K R A I D Y A L E N G V V V V A A A G N S S A Y Q S S L Y 306
CCCGCAATATCCGGGCTTATCCAAGTTGGCGCGTTGAAAACGTTGAACCTCCGTTCACTACGAGTTTTTCAAACAGGAGCCCTTTGGTTTCGGTTGGA 1020
P A N Y P G V I Q V G A V E N G E P P F T T S F S N R S P L V S V G 340
GCACCTGGTAGATTGGTCTCTCAACCATGCCGATGCCCGGATCGGCTGGTTACGAGAGTGGTCTCTTGTATTCTCCGAGAACGSGGTTACTACGGGTTCC 1122
A P G R L V L S T M P M R G S A G Y E S G S L Y S S E N G G Y Y G F 374
ATGAGTGGTACGTCGATGGCTACACCATGTTTCCGGAATGGCTGCGCTTTTGTCTCAAAGTTTCCGACGCAAAACCTTGGCAAATCAGAAAACGATC 1224
M S G T S M A T P H V S G M A A L L L Q K F P S A K P W Q I R K L I 408
GAAAACGGTGCACGGATATAGATTACCTGGCTTGGACGAACAATCCGCTACGGTTTACTGAATGCCCGTCCCTGGGGCTTGGACTCTCAGAGATGA 1326
E N G A R D I D L P G L D E H S G Y G L L N A R S V G L E L P E N G 442
GCGCAACCGTGGTGAAGGTGAAAGTCAATGGAAGTAGTAACCGGTGCCATTGTTTCGTTGTTAGGAAAAACGGGATAAGTTACGCACGTAGGTAC 1428
A A N V V V K V K V N G K V V T G A I V S L L G K N G I S Y A R R Y 476
TTTGGTTGACCGATGTGGGAATCGCGAAGTTCCTGGGAATCGAIGTTGGAGAGTACAGAATGATTGTTTCAAGCCAGGAAAAATGCTTCGAGCAAGAAGTA 1530
F A L T D V G I A K F L G I D V G E Y R M I V S S Q E K C F E Q E V 510
AACATAAATTCGGACGTGCAAAATGAAGTTGAACATATAGta 1571
N I N S D V Q I E V E L * 522

```

Figure 1. Nucleotide and deduced amino acid sequences of a *proA1* from *Fervidobacterium* sp. strain FC2004. The ORF begins with a valine codon. Top lines represent nucleotide sequence. Bottom lines represent amino acid sequence. Underline represents primer binding site. Boxes represent catalytic triad amino acids. One letter abbreviation of amino acids is depicted.

3D structure analysis

Large S8_ and small S8_subtilases from *F. pennivorans*, *F. nodosum* and *F. islandicum* and ProA1 precursor were revealed for 3D-models. In general, all the large homologs (682 to 710 aa) including the fervedolysin (1r6v.1.A) contain a propeptide domain (PD), catalytic domain (CD) and two sandwich or substrate binding domains (SD1 and SD2) (Figure 2A, B, C, D, E, F). Unlike the fervedolysin/islandisin, all three small S8_subtilases were revealed lacking both sandwich domain (Figure 2 G, H and I). Interestingly, 3D-model of ProA1 contains an atypical SD lacks completely of the SD2 (Figure 2J).

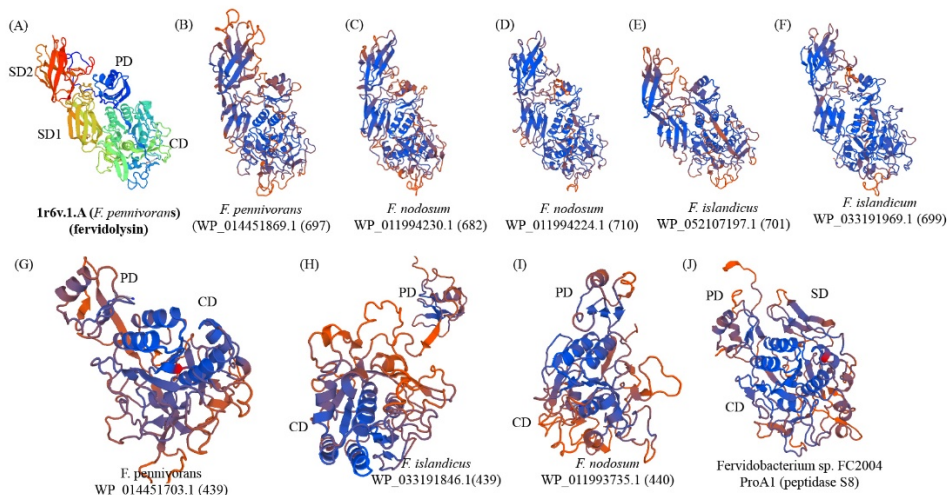


Figure 2. 3D-models of Peptidase_S8 family serine proteases (A) Model of the Ir6v.1.A template. The template is x-ray crystal structure of the mutant fervedolysin (A200/H200) (Kim et al., 2004). (B) - (F) Models of large S8_subtilases. (G) - (I) Models of small S8_subtilases. (J) Model of ProA1 (522 aa). Symbols: Species and protein IDs are as indicated in the images. Numbers in parenthesis indicate sizes (amino acids). PD represents propeptide domain, CD represents catalytic domain, SD represents sandwich or substrate binding domain.

Phylogenetic tree of S8_subtilase sequences categorized proProA1 within a branch formed by the large S8_homologs from *Fervidobacterium* rather than the small ones (Figure 3). Whether or not this is an intermediate peptidase_S8 family typical for this particular strain of *Fervidobacterium* in Thailand lineage or a truncated mutant of a large S8_homologs in genome of this strain is unclear.

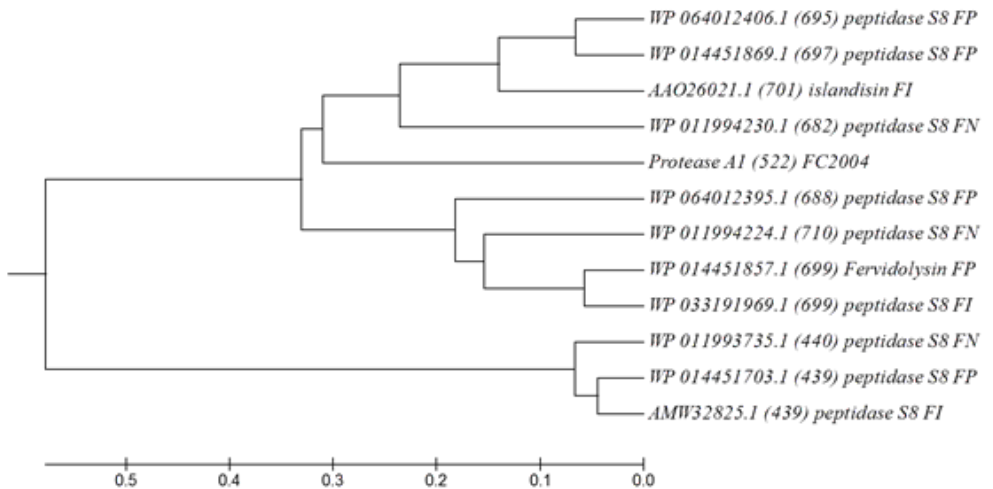


Figure 3. Phylogenetic tree of pro1A and other peptidase_S8 family serine proteases from *Fervidobacterium*. Abbreviations: FN, *Fervidobacterium nodosum*; FP, *F. pennivorans*; FI, and *F. islandicum*.

molecules of many propeptidases including the fervidolysin and the protein ID of WP_014451703.1, a small subtilase (Figure 5). In Figure 5C, catalytic triad residues of D169, H207 and S379 were revealed in CD of the ProA1.

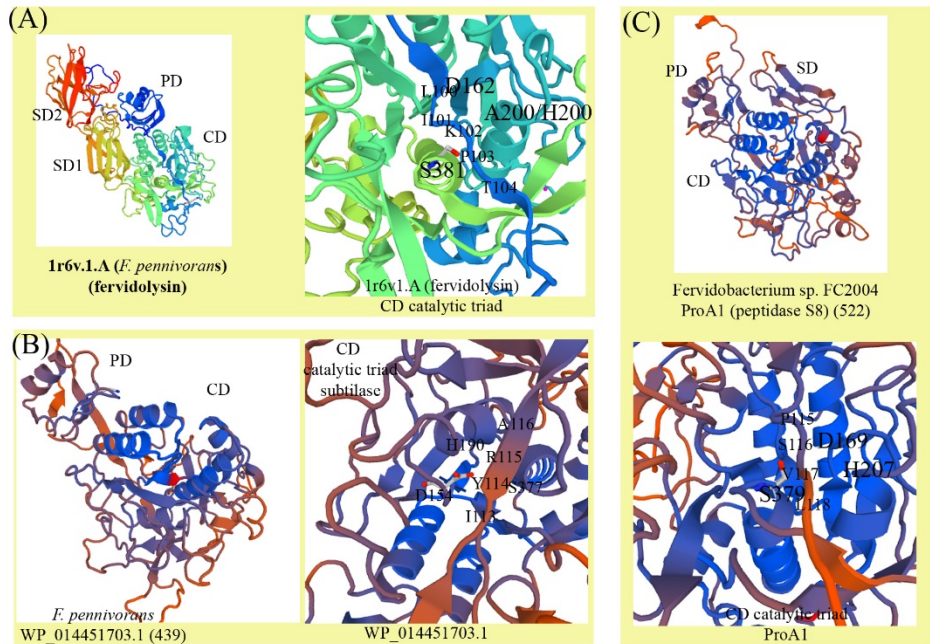


Figure 5. Catalytic triads and binding of C-terminal PD peptides. (A) Fervidolysin (1r6v.1.A) from *F. pennivorans*. (B) A small peptidase_S8 serine protease (WP_014451703.1) from *F. pennivorans*. (C) ProA1. D169, H207 and S379 are identified as the catalytic triad.

Alignment of C-terminal sequences of ProA1 (A431 to Y495 segment) to large S8_subtilases suggest a homologous sequence to SD1 and thus named “SD-like” in this study. However, the ProA1 lacks the second SD2 (Figure 6A). Modeling experiment using predicted SD-like region (G422 to Q502) reveals that models constructed from G442 to I498 and N441 to M497 are similar to the partial segment of a collagen adhesion protein (3kpt.1) and fibronectin binding protein (2x5p.1.A).

Superimposed images from the 1r6v.1.A, 3kpt.1 and 2x5p.1.A templates (Figures 6B and 6C) reveal good fit models suggesting a protein binding property of the SD-like. Although the strain FC2004 can degrade native feather, whether or not the mature ProA1 is active on keratin and the SD plays a role in hydrolyzing keratin substrate remained to be investigated.

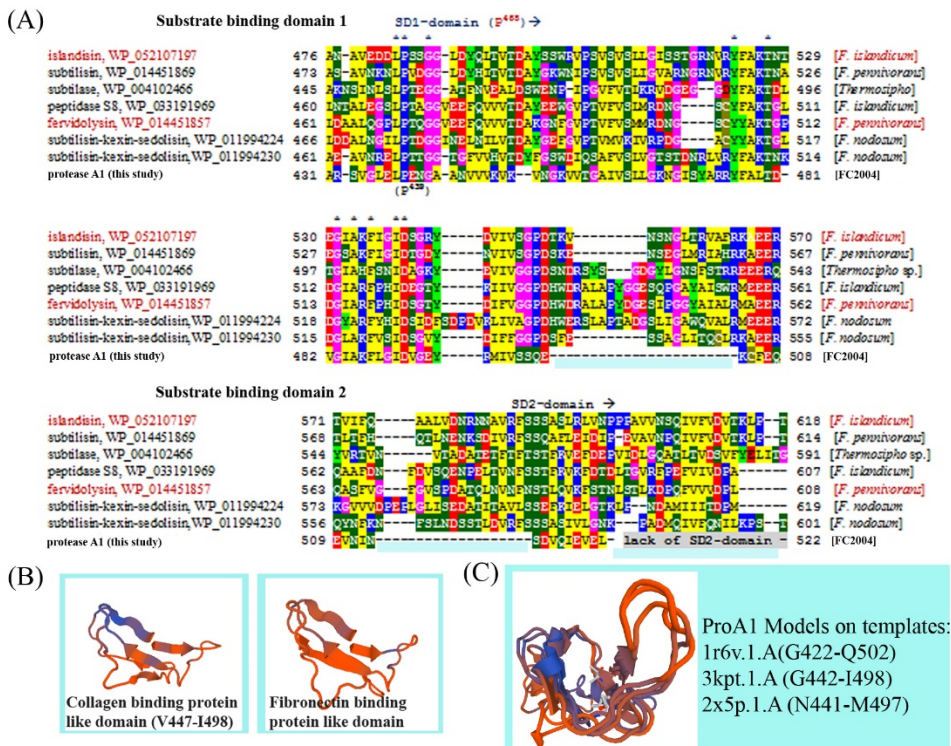


Figure 6. (A) Alignment of C-terminal regions of large S8_subtilases and ProA1. (B) 3D-model of ProA1 constructed using 3kpt.1.A (collagen binding protein) and 2x5p.1.A (fibronectin binding protein like domain). (C) Superimposed images constructed on three templates (1r6v.1.A, 3kpt.1.A and 2x5p.1.A). Symbols: yellow represents nonpolar amino acids (A, V, L, I, M) and F, red represents aspartic and glutamic acids, pink represent glycine, grey represents cysteine, green represents tyrosine, dark green represents polar amino acids (S, T, N,

Q), W and H and blue represents positively charged amino acids (K and R) and P. Dashes indicate spacing generated by alignment process. Numbers represent numbers of amino acids in a particular protein ID.

DISCUSSION & CONCLUSION

An ORF encoding a putative thermostable peptidase S8 family serine protease named in this study “ProA1” is identified in genome of a hyperthermophilic *Fervidobacterium* sp. FC2004 (Figure 1). ProA1 (522 amino acids) contains a signal peptide and a propeptide in molecule suggesting the ProA1 might be a membrane bound or extracellular subtilase precursor (Godde et al., 2005; Kim et al., 2004; Kluskens et al., 2002). Unlike keratinases and other S8_subtilases, mature ProA1 contains a SD-like (Figures 2, 6A). In Figure 3, phylogenetic analysis grouped ProA1 within the branch of fervidolysin-like molecules. The results imply that the ProA1 precursor is more closely related to the large S8_subtilases. Typical catalytic triad residues of D169, H207 and S379 in CD were identified suggesting ProA1 is an S8_family serine protease (Figure 5C). The possible function as substrate binding domain of the SD-like in mature ProA1 is evidenced (Figure 6). Whether or not the proProA1 is a truncated mutant peptidase S8_subtilase or an intermediate homolog (between the large and small S8) remained to be investigated. Comparative mapping of these alleles in the complete genome sequences of *F. pennivorans* DSM 9078^T (CP003260.1), *F. nodosum* Rt17-B1^T (CP000771.1) and *F. islandicum* AW1 (CP014334.1) might provide a clue. It is likely that strain FC2004 might contain other S8_subtilase homologs with two SDs and without SD in molecules. Although, strain FC2004 has ability to degrade native feather, the SD-like sequence might or might not be related to this phenotype.

In conclusion, the ProA1 of strain FC2004 is unique and have never been locating in the three complete genomes of *Fervidobacterium* (CP003260.1, CP000771.1 and CP014334.1). The enzyme ProA1 has potential applications in leather industries, textile processing and waste treatment from poultry industry. Also feather waste treatment at high temperature by strain FC2004 is considered a green process that possibly prevents contamination by dermatophytic fungi.

ACKNOWLEDGEMENTS

This work was supported by the following grants: The Scientific Promotion and Development Fund, Faculty of Science, Silpakorn University (SFR-SRG-2558-01) and the Silpakorn University Research and Development Institution from Thailand; and the Ministry of Economy and Productivity (Consolider CSD2009-00006) and the Andalusian Government (BIO288) from Spain with participation of FEDER funds.

REFERENCES

- Friedrich, A.B. and Antranikian, G. 1996. Keratin degradation by *Fervidobacterium pennavorans*, a novel thermophilic anaerobic species of the order Thermotogales. *Applied and Environmental Microbiology*. 62(8): 2875-2882.
- Godde, C., Sahn, K., Brouns, S.J.J., Kluskens, L.D., Oost, J.V.D., Vos, W.M. and Antranikian G. 2005. Cloning and expression of islandisin, a new thermostable subtilisin from *Fervidobacterium islandicum*, in *Escherichia coli*. *Applied and Environmental Microbiology*. 71(7): 3951-3958.
- Kanoksilapatham, W., Pasomsup, P., Keawram, P., Cuecas, A., Portillo, M.C. and Gonzalez, J. M. 2016. *Fervidobacterium thailandense* sp. nov., a novel extreme thermophilic bacterium isolated from a hot spring in Northern Thailand. *Internation*

- Journal of Systematic and Evolutionary Microbiology. DOI: 10.1099/ijsem.0.001463. Epub. Aug 26, 2016.
- Keawram, P., Pongsapukdee, V. and Kanoksilapatham, W. 2016. Isolation of *Thermotoga* spp. and *Fervidobacterium* spp., and characterization of *16S rRNA* genes of order Thermotogales: Unique lineage of hyperthermophiles thriving in 3 hot springs in Thailand. Silpakorn University Science and Technology Journal. 10(1): 9-20.
- Keawram, P. and Kanoksilapatham, W. 2013. Diversity of hyperthermophilic bacteria belonging to order Thermotogales thriving in three hot springs in Thailand: resources of genes encoding thermostable enzymes. Silpakorn University Science and Technology Journal. 7(2): 17-28.
- Kim, J.S., Kluskens, L.D., Vos, W.M., Huber, R. and Oost, J.V.D. 2004. Crystal structure of fervidolysin from *Fervidobacterium pennivorans*, a keratinolytic enzyme related to subtilisin. Journal of Molecular Biology. 335: 787-797.
- Kluskens, L.D., Voorhorst, W.G., Siezen, R.J., Schwerdtfeger, R.M., Antranikian, G., van der Oost, J. and de Vos, W.M. 2002. Molecular characterization of fervidolysin, a subtilisin-like serine protease from the thermophilic bacterium *Fervidobacterium pennivorans*. Extremophiles. 6(3): 185-194.
- Nam, G.W., Lee, D.W., Lee, H.S., Lee, N.J., Kim, B.C., Choe, E.A., Hwang, J.K., Suhartono, M.T. and Pyun, Y.R. 2002. Native-feather degradation by *Fervidobacterium islandicum* AW-1, a newly isolated keratinase-producing thermophilic anaerobe. Archives of Microbiology. 178: 538-547.
- Siezen R.J. and Leunissen, J.A.M. 1996. Subtilases: the superfamily of subtilisin-like serine proteases. Protein Science. 6: 501-523.