Biosci. Rep. (2012) / **32** / 609–618 (Printed in Great Britain) / doi 10.1042/BSR20120074

Original Paper

Bioscience Reports

An archaeal protein evolutionarily conserved in prokaryotes is a zinc-dependent metalloprotease

Yongmei HU*, Nan PENG*, Wenyuan HAN†, Yuxia MEI*, Zhengjun CHEN*, Xu FENG*, Yun Xiang LIANG*[1] and Qunxin SHE*†[1]

*State Key Laboratory of Agricultural Microbiology, College of Life Science and Technology, Huazhong Agricultural University, Wuhan 430070, People's Republic of China, and †Archaeal Centre, Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, Copenhagen Biocenter, DK-2200 Copenhagen N, Denmark

Synopsis

A putative protease gene (*tldD*) was previously identified from studying tolerance of *letD* encoding the CcdB toxin of a toxin–antidote system of the F plasmid in *Escherichia coli*. While this gene is evolutionarily conserved in archaea and bacteria, the proteolytic activity of encoded proteins remained to be demonstrated experimentally. Here we studied Sso0660, an archaeal TldD homologue encoded in *Sulfolobus solfataricus* by overexpression of the recombinant protein and characterization of the purified enzyme. We found that the enzyme is active in degrading azocasein and FITC–BSA substrates. Protease inhibitor studies showed that EDTA and *o*-phenanthroline, two well-known metalloprotease inhibitors, either abolished completely or strongly inhibited the enzyme activity, and flame spectrometric analysis showed that a zinc ion is a cofactor of the protease. Furthermore, the protein forms disulfide bond via the Cys^{416} residue, yielding protein dimer that is the active form of the enzyme. These results establish for the first time that tidD genes encode zinc-containing proteases, classifying them as a family in the metalloprotease class.

Key words: Archaea, metalloprotease, novel zinc-binding motif, Sso0660, *Sulfolobus*, TldD

Cite this article as: Hu, Y., Peng, N., Han, W., Mei, Y., Chen, Z., Feng, X., Liang, Y.X. and She, Q. (2012) An archaeal protein evolutionarily conserved in prokaryotes is a zinc-dependent metalloprotease. Biosci. Rep. **32**, 609–618

Bioscience Reports

INTRODUCTION

Most organisms typically employ 2–4 % of their genetic resources to code for proteases, enzymes that break down polypeptides or proteins by cleaving peptide bonds. For example, the hyperthermophilic acidophile *Sulfolobus solfataricus* encodes more than 70 putative proteases [1,2]. During the past decade *Sulfolobus* has been developed into a model system in studying archaeal molecular mechanisms such that major research progresses have been made in studying chromosome replication [3], DNA damage repair [4,5], cell cycle control [6] and regulation of gene transcription and protein translation [7–9] with this model. Although proteases are important players in cellular life, only a few proteases encoded in this archaeon have been characterized, including one extracellular protease and four intracellular enzymes [10–14]. Currently most proteases predicted for *S. solfataricus* remain to be characterized both for their biochemical properties and for their physiological functions.

We are interested in a pair of putative proteases, Sso0660 and Sso0661 encoded in the genome of *S. solfataricus* [2]. In the current protease database (MEROPS; http://merops.sanger.ac.uk/), both ORFs (open reading frames) are classified into the unknown protease family U62 [1]. These ORFs are homologous with the putative proteases encoded by *tldD*/*tldE* genes which were identified from screening for tolerance for *letD* on the *Escherichia coli* F plasmid coding for the CcdB toxin [15]. An insight into the functions of these putative proteases was gained from investigation of maturation of a peptide antibiotic and a toxin–antitoxin system both of which are plasmid-borne features in *E. coli*. Mutants defective in function of TldD or TldE (also called PmbA) accumulated the precursor of the antibiotic [16,17], and the encoded proteins were also implicated in the post-segregational killing by the F plasmid, which comprises two cognate factors, a stable toxin CcdB and a labile antitoxin CcdA. The cell death is 'addicted' to the short-lived CcdA such that cells rely on the *de novo* synthesis of CcdA to survive. While investigation of *tldD/E* deletion mutants suggested that both proteins could be involved in

Abbreviations used: AA, amino acid; DTT, dithiothreitol; E-64, *trans*-epoxysuccinyl-L-leucylamido-(4-guanidino)butane; ORF, open reading frame; LB, Luria–Bertani; NEM, *N*-ethylmaleimide; REU, relative fluorescence units.

[1] Correspondence may be addressed to either of these authors (email qunxin@bio.ku.dk or fa-lyx@163.com).

Table 1 Oligonucleotides used in the present study
Lower-case letters indicate restriction sites (underlined) and protection nucleotides of the PCR primers. The original codons in Sso0660 are listed under 'Sso0660 wild-type' in which the codons subjected to mutagenesis are highlighted in bold, with the start and end positions in the gene indicated. The mutated codons appear also in bold in which the substituted base(s) are underlined.

Name	Sequence 5′→3′	Source
PCR primer		
Sso0660fwd-Nde	taaagtcatatgTTTAAATACATTAAAAAAGCTG	The present study
Sso0660rev-Sal	tagtgtcgacTACATAGCCTCCCACCTT	The present study
Primers for site-directed mutagenesis		
Sso0660 wild-type	^{682}**CATGAG**GCTATAGGG**CAT**TTGAGTG^{706}	[2]
Mfwd1a (**H228F**)	**TTT**GAGGCTATAGGGCATTTGAGTG	The present study
Mfwd1b (**E229D**)	CAT**GAT**GCTATAGGGCATTTGAGTG	The present study
Mfwd1c (**H233Y**)	CATGAGGCTATAGGG**TAT**TTGAGTG	The present study
Mrev1	TGTAAAAATCCCTACTACCTCTGG	The present study
Sso0660 wild-type	^{1235}CTTCAGGCTTT**TGT**GAAAAGAATGG^{1259}	[2]
Mfwd2 (**C416G**)	CTTCAGGCTTT**GGT**GAAAAGAATGG	The present study
Mrev2	ACATTCCAAAATCCTTAGAAATCATC	The present study

degradation of the CcdA antitoxin *in vivo* [17], crystallographic analysis of *Thermotoga maritima* TldE (PmbA) failed to detect any co-ordinates for metal ions in the protein structure or any structural domain of a hydrolase [18]. Since there has not been any report on biochemical characterization of a TldD/E homologue in the current literature, whether or not any TldD or TldE encodes a protease remains to be tested.

Here we report for the first time that *S. solfataricus* Sso0660, a TldD homologue, encodes a metalloprotease and it contains an unusual zinc-binding motif and a C-terminal cysteine residue, both of which are of crucial importance to its protease activity.

MATERIALS AND METHODS

General DNA manipulation

Restriction and DNA modification enzymes were purchased from New England Biolabs, Fermentas or TransGen. Plasmid DNA was extracted from *E. coli* cells using an AxyPrep plasmid miniprep kit. Oligonucleotides used in the present study were synthesized from Invitrogen (listed in Table 1) where DNA sequencing of recombinant plasmids was also performed.

Strains, plasmids and medium

E. coli DH5α and Rosetta strains were used as host for DNA cloning and for producing recombinant protein respectively. Bacterial strains were cultured at 37 °C in LB (Luria–Bertani) broth containing 50 μg/ml kanamycin. Chloramphenicol was further supplemented to 17 μg/ml if applicable.

Sso0660 and Sso0661 genes were amplified from the *S. safataricus* P2 genome by PCR using Pyrobest DNA polymerase (Takara) and specific primers (Table 1). The resultant gene fragments were cloned to the *E. coli* expression vector pET30a, giving pET-660 and pET-661. Sequences of the cloned DNA fragments in the plasmids were confirmed by DNA sequencing.

Site-directed mutagenesis of Sso0660 gene

A PCR approach described previously [19] was employed to generate site-directed mutations in Sso0660. Two back-to-back primers were designed for generating each mutant gene (H228F, E229D, H233Y and C416G; Table 1), one of which carried the desired mutation (forward primer, fwd) whereas the other did not (reverse primer, rev). Full-length plasmids containing the designed mutations were amplified using the TransGen FastPfu DNA polymerase (TransGen Biotech) with pET-660 carrying the wild-type Sso0660 gene as the template. The resultant linear PCR products were phosphorylated with T4 DNA polynucleotide kinase and ligated with T4 DNA ligase to give circular plasmids, which were used to transform *E. coli* DH5α. Four resultant transformants were analysed for each cloning experiment. DNA sequencing of the mutant genes confirmed that all analysed plasmids carried the designed substitution mutations.

Expression and purification of Sso0660 recombinant protein

Expression plasmids derived from pET30a were transformed into the *E. coli* Rosetta, yielding strains for protein overexpression experiments. These strains were grown in LB broth at 37 °C. When the attenuance of the culture at 600 nm (D_{600}) reached 0.6–0.8, the synthesis of recombinant protein was induced by adding 0.8 mM IPTG (isopropyl β-D-thiogalactoside). The induction was at 16 °C for overnight. The cell mass was collected by centrifugation at 8000 ***g*** for 10 min and resuspended in the binding buffer of 50 mM sodium phosphate, 500 mM NaCl and 20 mM imidazole, pH 7.4. Cells were disrupted using a high-pressure homogenizer. After two passages, the resultant cell lysate was subjected to centrifugation at 12 000 ***g*** for 20 min, yielding soluble and insoluble

Table 2 Effect of protease inhibitors on the Sso0660 activity
The proteolytic reaction was conducted in the presence of each inhibitor at 55 °C for 1 h using the azocasein assay.

Inhibitor	Concentration (mM)	Relative activity (%)	Preferred protease target
No inhibitor	–	100	–
Pepstatin A	10^{-2}	87.4 ± 4.1	Aspartic protease
PMSF	5	116.2 ± 3.5	Serine protease
EDTA	5	0	Metalloprotease
o-Phenanthroline	10	10.4 ± 0.6	Metalloprotease
Bestatin	10^{-2}	92.4 ± 4.4	Aminopeptidase
E-64	10^{-2}	56.1 ± 2.6	Cysteine protease
Leupeptin	1	170.2 ± 1.4	Cysteine and serine protease
NEM	5	35.9 ± 3.9	Cysteine protease

fractions of the cell lysate. The recombinant protein in the soluble fraction was purified directly by affinity chromatography with a pre-packed HisTrap column as described previously [20]. Eluted fractions containing purified recombinant proteins were pooled together and dialysed against the buffer of 50 mM Tris/HCl, 10 % glycerol, pH 7.0, at 4 °C overnight. The protein preparations were analysed for homogeneity by SDS/PAGE. After adding glycerol to 25 % (v/v), the purified recombinant proteins were stored at − 20 °C until use. Only a small amount of the recombinant protein is present in the soluble fraction.

To purify the re-natured protein from the inclusion bodies, the insoluble fraction was washed with the buffer containing 50 mM sodium phosphate, 2 M urea, 1 mM EDTA, pH 8.0, and pelleted again (12 000 ***g*** at 4 °C for 10 min). Then, the pellet was dissolved with a denaturing buffer [50 mM sodium phosphate, 8 M urea, 5 mM DTT (dithiothreitol), 20 mM imidazole, pH 8.0], and the resultant suspension was subjected to protein purification by HisTrap affinity chromatography following the same procedure described above, except that all buffers used in this purification procedure contained 8 M urea. The purified protein was refolded in a stepwise dialysis against 50 mM sodium phosphate containing a decreasing concentration of urea (6, 4 and 2 M respectively), with each dialysis conducted for at least 4 h at 4 °C. The final dialysis was in 50 mM Tris/HCl, 10 % glycerol, pH 7.0. Remaining insoluble proteins were removed by centrifugation (12 000 ***g*** for 10 min at 4 °C), yielding re-natured proteins to be used for biochemical characterization.

Protease assays

Proteolytic activity of purified proteins was determined using either azocasein or FITC–BSA as substrate. In the azocasein assay, a reaction volume of 100 μl was prepared in a Microfuge tube, containing 2–10 μg of protease samples and 2 % azocasein dissolved in 50 mM Tris/HCl, pH 7.0. After incubation at 55 °C for 1 h, the reaction was stopped by adding 100 μl of 30 % trichloroacetic acid to precipitate unreacted substrates. After removing precipitates by centrifugation (12 000 ***g***, 5 min), the enzymatic products in the supernatant were converted into coloured products by adding 1 M NaOH and quantified at 440 nm using a NanoDrop ND-1000 spectrophotometer. The specific activity of the enzyme (units $\cdot$ mg^{-1} $\cdot$ h^{-1}) was expressed as the amount of enzyme required to produce an absorbance change of 1.0 under the assay conditions. Contents of purified recombinant proteins were estimated using a MicroBCA kit from Pierce following the manufacturer's instructions.

For the fluorescence protease assay, we essentially followed the method described previously [21,22]. Briefly, a reaction mixture of 60 μl was prepared including 10 μl of FITC–BSA, 20 μl of 50 mM Tris/HCl buffer, pH 7.0, and 30 μl of enzyme solution with 2–10 μg of enzyme. After incubation at 55 °C for 1 h, 120 μl of 5 % trichloroacetic acid was added to each reaction mixture to terminate the reaction. Unreacted substrates were removed as described above. The fluorescence intensity of each sample was measured using a Shimadzu RF-5301PC fluorospectrometer with λ_{ex} at 490 nm and λ_{em} at 525 nm. Proteolytic activity was expressed in RFU (relative fluorescence units) with 1 unit arbitrarily defined as the amount of enzyme required for producing 1 optical density change under the assay conditions (RFU $\cdot$ mg^{-1} $\cdot$ h^{-1}).

Effects of pH and temperature on enzyme activity

The pH range tested for was 5–9. Two different buffers were used to generate different pH values: 50 mM phosphate buffer (pH 5.0 and 6.0) and 50 mM Tris/HCl buffer (pH 7.0–9.0). The assay started with incubating 2–10 μg of purified enzyme of Sso0660 with 2 % azocasein in different buffers at 55 °C for 1 h. Proteolytic activity of each sample was measured as described above.

The optimal temperature for the recombinant Sso0660 enzyme (2–10 μg) was determined similarly. Reaction mixtures containing the purified enzyme, 2 % azocasein in 50 mM Tris/HCl buffer (pH 7.0) were prepared and incubated at various temperatures (40–100 °C) for 1 h. The amount of products was estimated by spectrophotometry.

Effects of metal ions and protease inhibitors

Nine metal ions including Zn^{2+}, Cu^{2+}, Ca^{2+}, Mg^{2+}, Co^{2+}, Fe^{3+}, Sr^{2+}, Mn^{2+} and Hg^{2+} were studied for their effects on Sso0660 enzymatic activity at the final concentration of 5 mM.

Table 3 Effect of metal ions on Sso0660 activity
Proteolytic reaction was conducted in the presence of each metal ion at 55 °C for 1 h using the azocasein assay.

Metal ion	Concentration (mM)	Relative activity (%)
None	–	100
Zn^{2+}	5	157 ± 4.3
Fe^{3+}	5	6.9 ± 3
Co^{2+}	5	8.8 ± 4.3
Mg^{2+}	5	2.6 ± 1.9
Mn^{2+}	5	62.7 ± 4.7
Cu^{2+}	5	63.3 ± 4.7
Ca^{2+}	5	99 ± 6.6
Sr^{2+}	5	17 ± 3.8
Hg^{2+}	5	84.7 ± 2.5

A master reaction was prepared from which aliquots were made. Metal ions were added to the aliquots individually, each with the final concentration of 5 mM. The resultant solutions were incubated at 55 °C for 1 h. The amount of products was estimated by spectrophotometry as described above.

Several known protease inhibitors (see Table 3 for details) were assayed to test for their inhibition on Sso0660. Aliquots of the enzyme were pre-incubated with each inhibitor at 55 °C for 1 h in the concentration indicated in Table 3. The residual activity of the enzyme was determined using the azocasein assay.

Flame spectrometric analysis of metal ion cofactor

Sso0660 wild-type protein and its C416G mutant were analysed for metal ion cofactor using a TAS-990 atomic absorption spectrometer (Beijing Puxi General Instrumental Company). All solutions used for this experiment were prepared in plasticware with ultrapure water. Approximately 800 μg protein was used in each determination and the absorption of zinc ion was recorded at a wavelength of 213.9 nm. The results were the averages of measurements obtained for three different batches of wild-type and mutant proteins.

RESULTS AND DISCUSSION

Sso0660 exhibited proteolytic activities

The *S. solfataricus* TldD homologue Sso0660 was overexpressed in *E. coli* as described in the Materials and methods section. As for many other thermophilic proteins expressed in a mesophilic host (e.g. [20]), Sso0660 recombinant protein forms predominantly inclusion bodies as a vast majority of the protein is present in the insoluble fraction of the cell lysate (Figure 1). We first purified the recombinant Sso0660 from the soluble fraction of the cell lysate and tested for its proteolytic activity. In an assay with azocasein as a substrate (azocasein assay), the purified Sso0660 protein showed a specific activity of 7.2 units $\cdot$ mg^{-1} $\cdot$ h^{-1}. Subsequently

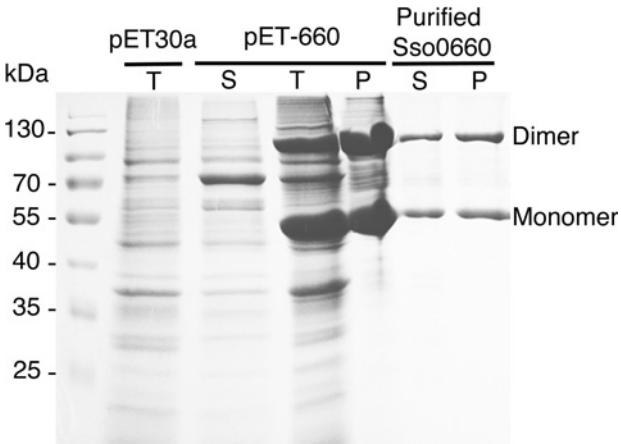

Figure 1 Purification of Sso0660 recombinant protein from *E. coli*
Samples were analysed by SDS/PAGE (12 % gel) and protein bands were visualized by staining with Coomassie Brilliant Blue R250. Cells containing the cloning vector (pET30a) and overexpression plasmid (pET-660) are indicated. T, total protein; S, supernatant – the soluble fraction of cell extracts; P, pellet – the insoluble fraction of cell extract. Samples were boiled for 5 min in an SDS-loading buffer to denature proteins before loading.

Sso0660 recombinant protein was also purified as a re-natured form from inclusion bodies. Protease assay with the re-natured protein showed a specific activity of 6.5 units $\cdot$ mg^{-1} $\cdot$ h^{-1}, which is slightly lower than the activity of the soluble protein. Testing both forms of recombinant Sso0660 protein with FITC-BSA as a substrate (fluorescence assay) yielded similar results, showing 1 230 000 and 1 125 000 RFU $\cdot$ mg^{-1} $\cdot$ h^{-1} for the soluble and re-natured forms of enzyme respectively. Thus we conclude that TldD has proteolytic activity.

We also produced recombinant protein for the *S. solfataricus* TldE homologue Sso0661 and tested the purified protein for protease activities as for Sso0660. The purified Sso0661 protein consistently gave a low level of protease activity that was approximately 7 % of the Sso0660 activity (results not shown). Moreover, whereas the purified Sso0660 recombinant protein appeared both as dimer and as monomer in a SDS/PAGE (Figure 1), the Sso0661 recombinant protein existed only as monomer (results not shown). These differences prompted us to examine their protein sequences more closely.

BLAST searches [23] of TldD/TldE homologues in the genomes of different *Sulfolobales* species, including *S. acidocaldarius* [24], *S. islandicus* [25,26], *S. tokodaii* [27], *Metallosphaera cuprina* [28] and *Acidianus hospitalis* [29] indicated that these organisms each encode 5 TldD/E homologues as for *S. solfataricus*. Four sets of distantly related TldD/Es were then chosen for sequence comparisons, including TldD and TldE (also named PmbA) of *E. coli* [30] and *T. maritima* [31] and the two pairs of TldD/Es (Sso0660, 0661, 1151 and 1152) of *S. solfataricus* [2]. Pairwise alignments of sequences of different members in TldD or TldE group revealed sequence similarity/identity of 19–26 %/35–46 % within each group. Multiple alignments of eight representative sequences using the Clustal

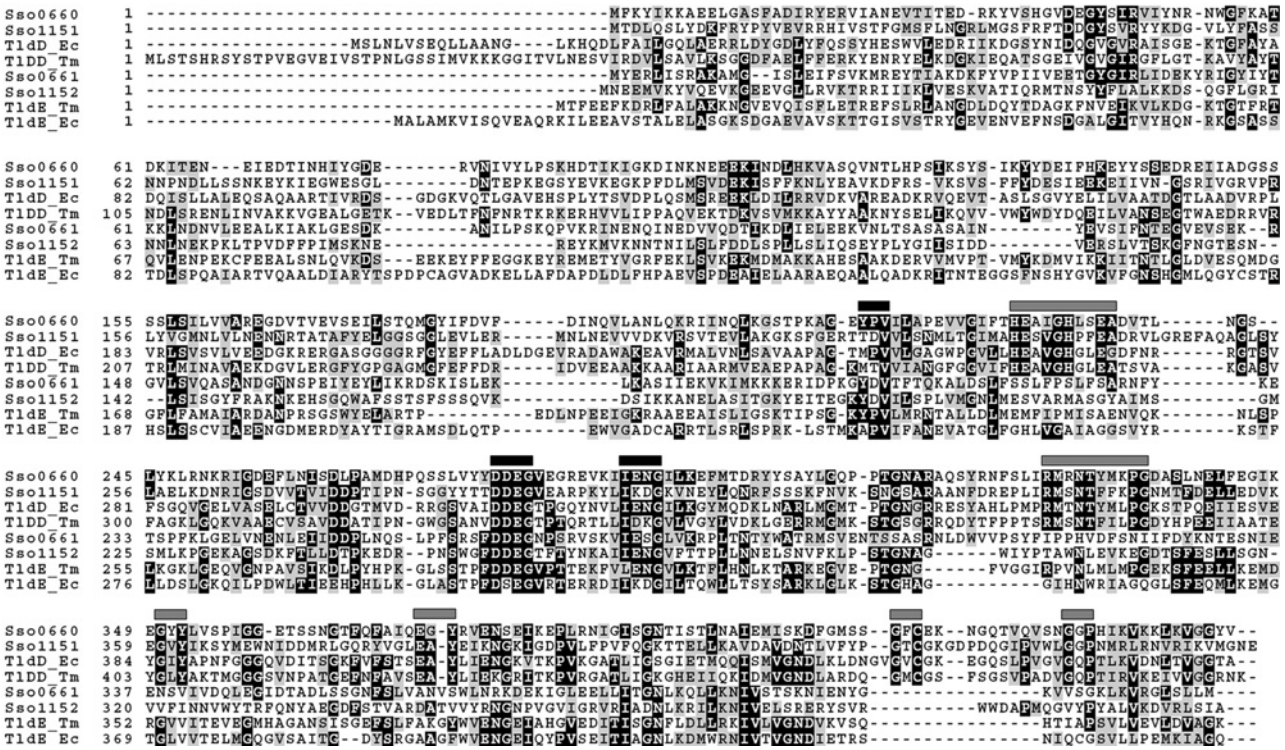

Figure 2 **A multiple sequence alignment of a few selected archaeal and bacterial TldD and TldE proteins**
Sso0660 and Sso1151 are *S. solfataricus* TldD homologues whereas Sso0661 and Sso1152, TldE homologues. TldD_Ec and TldE_Ec denote *E. coli* TldD and TldE; TldD_Tm and TldE_Tm: *Thermotoga maritima* TldD and TldE. These sequences were aligned to each other using Clustal X [32] and the obtained results were illustrated using a boxshade program (http://www.ch.embnet.org/software/BOX_form.html) with a threshold of 0.5 under which all the AA positions with $\geqslant$50% sequence identity are shaded in black whereas those with $\geqslant$50% sequence homology are highlighted in grey. Fully filled rectangles indicate the motifs conserved in both TldDs and TldEs and grey rectangles indicate the motifs conserved only in TldDs.

W program [32] revealed several stretches of conserved sequences most of which are located within the C-terminal halves of these proteins (Figure 2). Six TldD-specific motifs include HEXXXHXXE, RMXNTXXXPG and four short sequence motifs (Figure 2), with one of the short motifs being GxC which contains the only cysteine residue of Sso0660. These findings prompted us to focus on biochemical characterization of the Sso0660-encoded enzyme.

Sso0660 encoded a metalloprotease

As the re-natured Sso0660 recombinant protein was almost as active as the soluble recombinant protein both in the azocasein assay and in the fluorescent assay, re-natured recombinant protein was used in all subsequent experiments as a large yield of Sso0660 recombinant protein could be obtained from the inclusion bodies relatively easily.

First, we determined the optimal pH and temperature of the protease using the azocasein assay. When assayed for the pH range of 5–9, we found that Sso0660 is most active at pH 7 (Figure 3A). However, the enzyme also showed high activities at pH 6 and 9 (>75% of enzyme activity at pH 7), indicating that this enzyme is relatively insensitive to pH change.

The optimal temperature of the Sso0660 enzyme was also studied using the azocasein assay. As shown in Figure 3(B), the enzyme exhibits similar activities between 55 and 65 °C. Relatively high proteolytic activities were observed at all other tested temperatures (>60%), except 95 °C, for which the enzyme activity decreased to approximately 20% (Figure 3B). This indicates that Sso0660 recombinant enzyme is also relatively insensitive to the change of reaction temperature.

It should be noted that the temperature and pH optimums determined for Sso0660 with azocasein as the substrate may not be directly applicable to another substrate since optimal temperature and pH values of an enzyme can exhibit substrate-specificity, meaning that Sso0660 could exhibit different optimal values of temperature and pH when another substrate is to be used in the assay.

Next, we investigated the effects of various protease inhibitors on the activity of the Sso0660 protease. This was done by pre-incubation of the enzyme with each individual protease inhibitor at the indicated concentration at 55 °C for 1 h, and residual protease activities in the samples were determined using

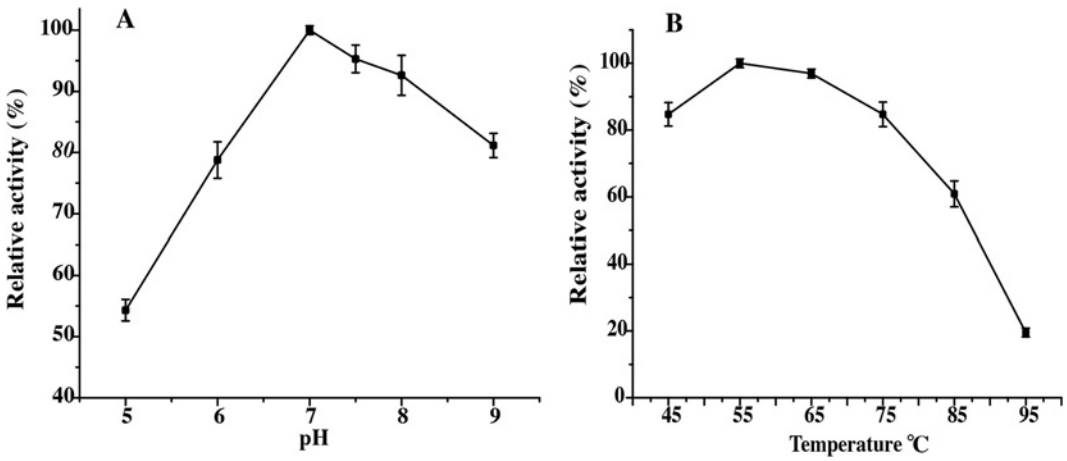

Figure 3 Temperature- and pH-dependence of Sso0660 protease activity

(**A**) The pH range was generated with two different buffer systems: 50 mM phosphate buffer (pH 5.0–6.0) and 50 mM Tris/HCl (pH 7.0–9.0). The activity of purified Sso0660 was assayed using the azocasein assay. Activities at different pH values were relative values to the activity at pH 7.0 (100%). (**B**) Azocasein (2%) was added to enzyme aliquots and incubated at the indicated temperatures for 1 h. Proteolytic products were estimated by the azocasein assay and relative activities of the enzyme were calculated referring to the activity at 55°C (100%).

the azocasein assay. PMSF and leupeptin did not show any inhibition. In contrast they stimulated the activity of the enzyme (Table 2), indicating that Sso0660 did not encode a typical serine protease. Neither pepstatin A nor bestatin inhibited the enzyme activity, demonstrating that the enzyme does not have an aspartic protease activity, nor does it have an aminopeptidase activity of leucine aminopeptidase, aminopeptidase B or tri-amino peptidase type.

When testing for the effects of the irreversible thiol inhibitor E-64 [*trans*-epoxysuccinyl-L-leucylamido(4-guanidino)butane] and NEM (*N*-ethylmaleimide), each inhibitor reduced the activity of Sso0660 to 36–56%. Since there is only one cysteine residue in Sso0660, this suggested that the cysteine residue could be important for Sso0660 activity. As both monomeric and dimeric forms of Sso0660 are present in the purified recombinant protein (Figure 1), it was of interest to study whether the monomer or the dimer could be the active form of the enzyme. The protein dimer was converted to monomer by treating the protein with DTT and the resultant monomeric protein was assayed for proteolytic activity. The protein showed a very low proteolytic activity (6%), suggesting that the protein dimer is the active form of the enzyme. When the only cysteine residue in Sso0660 was substituted with a glycine residue, the resultant mutant enzyme C416G is a monomeric protein, retaining 12% of the residual activity (Figure 4C). This indicates that the cysteine residue is very important to the enzyme activity of Sso0660.

We also treated the enzyme with EDTA or *o*-phenanthroline, two specific inhibitors of metalloprotease that inhibit the enzyme activity by depleting metal ions from the active centre of the enzyme. Measuring the residual proteolytic activity of the treated Sso0660 samples indicated that while *o*-phenanthroline strongly inhibited the activity, EDTA abolished the activity completely (Table 2). This demonstrates that Sso0660 encodes a metalloprotease.

To study the effects of metal ions on the enzyme activity, nine different metal ions, including Zn^{2+}, Cu^{2+}, Ca^{2+}, Mg^{2+}, Co^{2+}, Fe^{3+}, Sr^{2+}, Mn^{2+} and Hg^{2+}, were added to the enzyme reaction individually, all in an excess amount (5 mM). The activity of Sso0660 was assayed using azocasein as the substrate with the results summarized in Table 3. Only zinc ion further stimulated the enzyme activity (up to 160%). All other tested metal ions showed a negative effect on the enzyme activity and the strongest inhibitory effects were observed for Mg^{2+}, Fe^{3+} and Co^{2+} retaining only 3–9% residual activity whereas Mn^{2+}, Cu^{2+} and Hg^{2+} exhibited a moderate inhibitory effect (63–84%). To reveal if zinc ion could be a cofactor of this protease, flame spectrometric analysis of zinc ion was performed for the purified Sso0660 protein. We found that zinc ion is associated with the protein in approximately 1:1 molar ratio, indicating that Sso0660 is a metalloprotease with Zn^{2+} as the cofactor.

The HEAIGH motif comprises the active centre of the Sso0660 metalloprotease

Although Sso0660 lacks HEXXH, a common zinc-binding motif widely present in metalloproteases [33,34], it does contain the HEAIGH motif that could represent a variant version of the common zinc-binding motif. We conducted site-directed mutagenesis for the three conserved AA (amino acid) residues (His^{228}, Glu^{229} and His^{233}, Table 1) and constructed mutant genes for expressing H228F, E229D and H233Y mutant proteins. As for the wild-type enzyme, these mutant proteins also formed inclusion bodies. We purified re-natured mutant proteins from insoluble fractions and

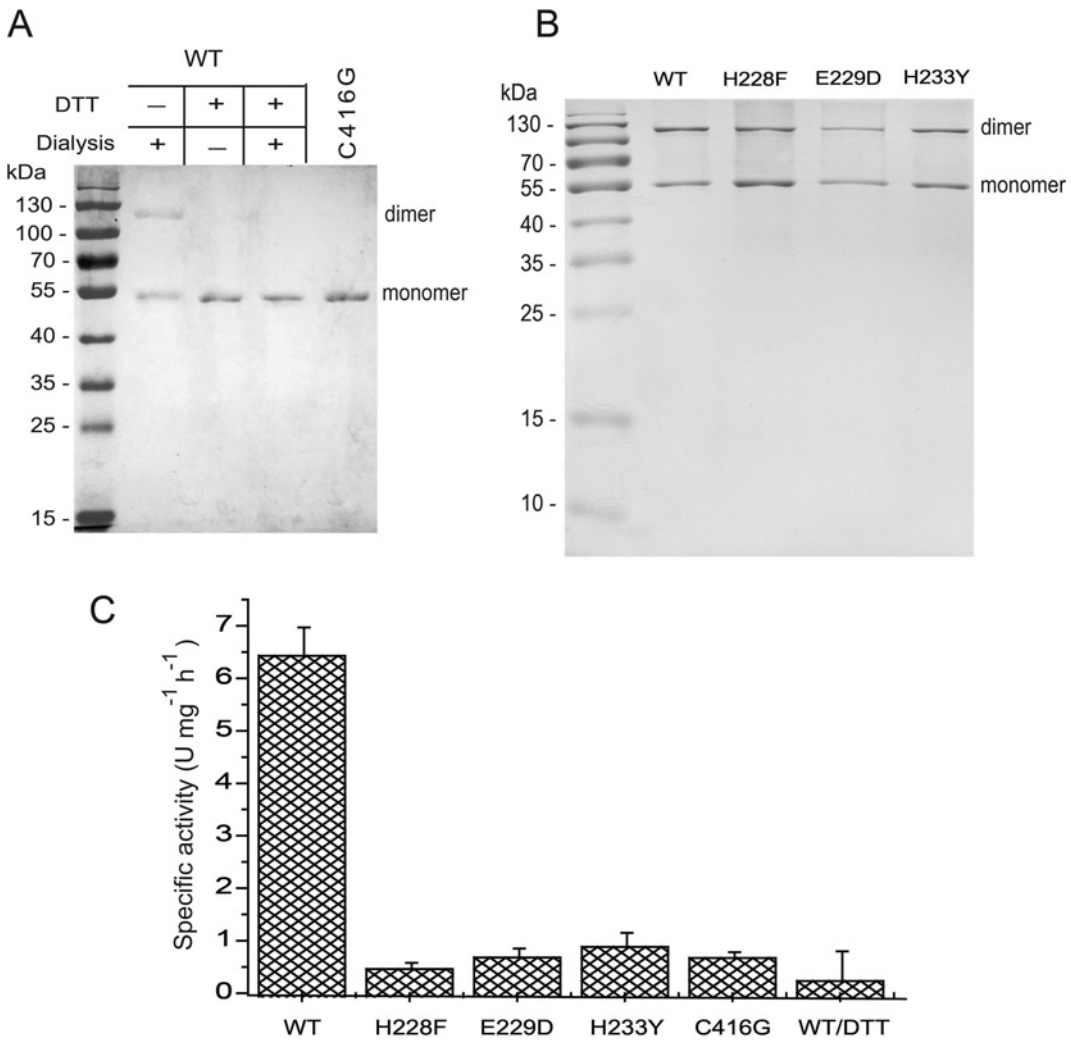

Figure 4 **Site-directed mutagenesis of Sso0660 HEXXXH motif and the C-terminal cysteine residue**

(**A**) Sso0660 WT (wild-type) enzyme was incubated at 55 °C for 1 h in the presence or absence of 5 mM DTT, which was then removed by dialysis against 50 mM Tris/HCl (pH 7.0). The treated samples were analysed by SDS/PAGE (12 % gel) and stained with Coomassie Brilliant Blue R250. C416G is a cysteine mutant of Sso0660. Positions of the monomer and dimer forms of the proteins are indicated. (**B**) Sso0660 WT and its mutant proteins (H228F, E229D and H233Y) were purified via histidine affinity-tag purification and analysed by SDS/PAGE (12 % gel). (**C**) Specific activity of the purified enzymes assayed with azocasein as the substrate. WT/DTT denotes a DTT-treated Sso0660 protein sample as shown in lane 3 of (**A**).

analysed them by SDS/PAGE. As shown in Figure 4(B), they formed both monomer and dimer as for the wild-type enzyme.

All three Sso0660 mutants were assayed for their protease activities using the azocasein assay. As shown in Figure 4(C), the activity was greatly decreased for three mutant enzymes, retaining 8.3–14.8 % of the activity of the wild-type enzyme. These results suggest that HEAIGH motif functions as the active centre of the metalloprotease.

The identified HEXXXH motif may comprise an unusual metal ion-binding motif in TldD proteins although it contains an AA insertion between the two histidine residues that co-ordinate the metal ion cofactor in the more common zinc-binding motif HExxH of metalloproteases. Interestingly, a similar motif with the same consensus functions in co-ordinating a zinc ion at the active centre of eukaryotic DPP (dipeptidyl peptidase) III enzymes [35,36]. In these eukaryotic enzymes, the interactions between the HEXXXH motif and the zinc ion is to be further stabilized by a glutamate residue located approximately 52 AAs downstream of the motif [37]. However, this glutamate residue does not appear to be conserved in TldDs, or at least not at the same location (Figure 2).

Since Cys^{416} is important to enzyme activity, we investigated if Cys^{416} could contribute to the zinc cofactor co-ordination at the active centre. The C416G mutant protein was subjected to flame spectrometric analysis for zinc ion component. We found that the mutant enzyme contains an equivalent or slightly higher

amount of zinc ion, indicating that Cys^{416} does not play a role in co-ordinating the zinc cofactor. Therefore, we reasoned that the importance of Cys^{416} in Sso0660 is solely for the formation of protein dimer.

It has been well documented that zinc-binding motifs require an additional site in metal ion cofactor co-ordination. In fact, differences in the usage of an additional site for metal ion co-ordination have been used as a criterion to classify metalloproteases [34,38]. Notably, there are several other AAs that are only conserved in TldD sequences (Figure 2). Probably one of the conserved AAs plays a function in zinc ion co-ordination in Sso0660. Taken together, our results indicate that TldDs constitute another group of HEXXXH motif-possessing metalloproteases, reassigning TldD enzymes to a family of protease belonging to the metalloprotease class.

The TldD superfamily of putative proteases is widespread in archaea and bacteria. While approximately 60% of known bacterial genomes encode TldD and TldE homologues, in the domain of Archaea only members in Haloarchaeota and Nanoarchaeota lack a TldD or PmbA/TldE. Many species have more than one pair of *tldD*/*tldE*-like genes as summarized in the peptidase database (MEROPS) (http://merops.sanger.ac.uk/cgi-bin/famsum?family = U62) [1]. Whereas the widespread occurrence of TldD and TldE argues for conserved functions for this superfamily of protease, *in vivo* functions of TldD and TldE have been investigated only for the maturation of antibiotic microcin B17 and for degradation of the antidote CcdA, both of which are plasmid-borne features. We reasoned that the observed effects of TldD/TldE deficiency could reflect the fact that these genetic elements hijacked a cellular process involving TldD/E. A prime candidate for the TldD/E-involved cellular processes is that these enzymes could be responsible for regulating the activities of chromosome-encoded toxins and antitoxins, which are highly abundant in archaea and bacteria [39]. To date, several toxin/antitoxin systems have been being characterized in bacteria (reviewed in [40]) and a pioneering experiment with an archaeal system suggests that archaeal toxin/antitoxin systems function in similar mechanisms [41]. Thus, further studies of these toxin–antitoxin systems will unravel possible functions of TldDs and TldEs in the regulation of toxin–antitoxin activities in archaeal and bacterial organisms.

Alternately, these proteases can directly regulate important enzymes by protein maturation or degradation. For example, a metal-dependent protease activity has recently been implicated in regulating the DNA reverse gyrase activity in *S. solfataricus* and thereby affecting genome integrity [42]. As many ORFs are predicted as metalloproteases in the MEROPS database [1], the involved protease has to be identified experimentally. Fortunately, very versatile genetic tools for *in vivo* study of gene functions are available for *S. acidocaldarius* [43] and for *S. islandicus* [44]. For the latter, the genetic tools developed include targeted gene inactivation, genetic complementation and protein overexpression [45–49]. Employing both biochemical and genetic methodologies to further investigate TldD/E encoded in *Sulfolobus* species should shed light on the functional roles of this important group of protease in Archaea.

AUTHOR CONTRIBUTION

Yongmei Hu, Wenyuan Han and Xu Feng conducted the experiments. All authors participated in experimental design and troubleshooting and data interpretation. Nan Peng, Yun Xiang Liang and Qunxin She provided research materials. Yongmei Hu, Wenyuan Han and Qunxin She wrote the paper.

ACKNOWLEDGEMENTS

We thank members of our laboratories for helpful discussions.

FUNDING

This work was supported the National Natural Science Foundation of China [grant numbers 31100050, 31100096, 31128011] and the Danish Council for Independent Research: Technology and Production Sciences [grant numbers 09-062332, 11-106683].

REFERENCES

1 Rawlings, N. D., Barrett, A. J. and Bateman, A. (2012) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res. **40**, D343–D350
2 She, Q., Singh, R. K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M. J., Chan-Weiher, C. C., Clausen, I. G., Curtis, B. A., De Moors, A. et al. (2001) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. Proc. Natl. Acad. Sci. U.S.A. **98**, 7835–7840
3 Beattie, T. R. and Bell, S. D. (2011) Molecular machines in archaeal DNA replication. Curr. Opin. Chem. Biol. **15**, 614–619
4 White, M. F. (2011) Homologous recombination in the archaea: the means justify the ends. Biochem. Soc. Trans. **39**, 15–19
5 Hong, Y., Chu, M., Li, Y., Ni, J., Sheng, D., Hou, G., She, Q. and Shen, Y. (2012) Dissection of the functional domains of an archaeal Holliday junction helicase. DNA Repair **11**, 102–111
6 Bernander, R. (2007) The cell cycle of *Sulfolobus*. Mol. Microbiol. **66**, 557–562
7 Benelli, D. and Londei, P. (2011) Translation initiation in Archaea: conserved and domain-specific features. Biochem. Soc. Trans. **39**, 89–93
8 Peng, N., Ao, X., Liang, Y. X. and She, Q. (2011) Archaeal promoter architecture and mechanism of gene activation. Biochem. Soc. Trans. **39**, 99–103
9 Jun, S. H., Reichlen, M. J., Tajiri, M. and Murakami, K. S. (2011) Archaeal RNA polymerase and transcription regulation. Crit. Rev. Biochem. Mol. Biol. **46**, 27–40
10 Hanner, M., Redl, B. and Stoffler, G. (1990) Isolation and characterization of an intracellular aminopeptidase from the extreme thermophilic archaebacterium *Sulfolobus solfataricus*. Biochim. Biophys. Acta **1033**, 148–153
11 Burlini, N., Magnani, P., Villa, A., Macchi, F., Tortora, P. and Guerritore, A. (1992) A heat-stable serine proteinase from the extreme thermophilic archaebacterium *Sulfolobus solfataricus*. Biochim. Biophys. Acta **1122**, 283–292
12 Colombo, S., D'Auria, S., Fusi, P., Zecca, L., Raia, C. A. and Tortora, P. (1992) Purification and characterization of a thermostable carboxypeptidase from the extreme thermophilic archaebacterium *Sulfolobus solfataricus*. Eur. J. Biochem. **206**, 349–357

13 Guagliardi, A., Cerchia, L. and Rossi, M. (2002) An intracellular protease of the crenarchaeon *Sulfolobus solfataricus*, which has sequence similarity to eukaryotic peptidases of the CD clan. Biochem. J. **368**, 357–363
14 Cannio, R., Catara, G., Fiume, I., Balestrieri, M., Rossi, M. and Palmieri, G. (2010) Identification of a cell-bound extracellular protease overproduced by *Sulfolobus solfataricus* in peptide-rich media. Protein Pept. Lett. **17**, 78–85
15 Murayama, N., Shimizu, H., Takiguchi, S., Baba, Y., Amino, H., Horiuchi, T., Sekimizu, K. and Miki, T. (1996) Evidence for involvement of *Escherichia coli* genes *pmbA*, *csrA* and a previously unrecognized gene *tldD*, in the control of DNA gyrase by *letD* (*ccdB*) of sex factor F. J. Mol. Biol. **256**, 483–502
16 Rodriguez-Sainz, M. C., Hernandez-Chico, C. and Moreno, F. (1990) Molecular characterization of *pmbA*, an *Escherichia coli* chromosomal gene required for the production of the antibiotic peptide MccB17. Mol. Microbiol. **4**, 1921–1932
17 Allali, N., Afif, H., Couturier, M. and Van Melderen, L. (2002) The highly conserved TldD and TldE proteins of *Escherichia coli* are involved in microcin B17 processing and in CcdA degradation. J. Bacteriol. **184**, 3224–3231
18 Rife, C., Schwarzenbacher, R., McMullan, D., Abdubek, P., Ambing, E., Axelrod, H., Biorac, T., Canaves, J. M., Chiu, H. J., Deacon, A. M. et al. (2005) Crystal structure of a putative modulator of DNA gyrase (pmbA) from *Thermotoga maritima* at 1.95 Å resolution reveals a new fold. Proteins **61**, 444–448
19 Peng, N., Xu, W., Wang, F., Hu, J., Ma, M., Hu, Y., Zhao, S., Liang, Y. X. and Ge, X. (2012) *Mitsuaria* chitosanase with unrevealed important amino acid residues: characterization and enhanced production in *Pichia pastoris*. Appl. Microbiol. Biotechnol., doi:10.1007/s00253-012-3901-y
20 Mei, Y., Peng, N., Zhao, S., Hu, Y., Wang, H., Liang, Y. X. and She, Q. (2012) Exceptional thermal stability and organic solvent tolerance of an esterase expressed from a thermophilic host. Appl. Microbiol. Biotechnol. **93**, 1965–1974
21 Twining, S. S. (1984) Fluorescein isothiocyanate-labeled casein assay for proteolytic enzymes. Anal. Biochem. **143**, 30–34
22 Zhu, H., Guo, D. C. and Dancik, B. P. (1990) Purification and characterization of an extracellular acid proteinase from the ectomycorrhizal fungus *Hebeloma crustuliniforme*. Appl. Environ. Microbiol. **56**, 837–843
23 Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S. and Madden, T. L. (2008) NCBI BLAST: a better web interface. Nucleic Acids Res. **36**, W5–9
24 Chen, L., Brugger, K., Skovgaard, M., Redder, P., She, Q., Torarinsson, E., Greve, B., Awayez, M., Zibat, A., Klenk, H. P. et al. (2005) The genome of *Sulfolobus acidocaldarius*, a model organism of the Crenarchaeota. J. Bacteriol. **187**, 4992–4999
25 Guo, L., Brugger, K., Liu, C., Shah, S. A., Zheng, H., Zhu, Y., Wang, S., Lillestol, R. K., Chen, L., Frank, J. et al. (2011) Genome analyses of Icelandic strains of *Sulfolobus islandicus*, model organisms for genetic and virus-host interaction studies. J. Bacteriol. **193**, 1672–1680
26 Reno, M. L., Held, N. L., Fields, C. J., Burke, P. V. and Whitaker, R. J. (2009) Biogeography of the *Sulfolobus islandicus* pan-genome. Proc. Natl. Acad. Sci. U.S.A. **106**, 8605–8610
27 Kawarabayasi, Y., Hino, Y., Horikawa, H., Jin-no, K., Takahashi, M., Sekine, M., Baba, S., Ankai, A., Kosugi, H., Hosoyama, A. et al. (2001) Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, *Sulfolobus tokodaii* strain7. DNA Res. **8**, 123–140
28 Liu, L. J., You, X. Y., Zheng, H., Wang, S., Jiang, C. Y. and Liu, S. J. (2011) Complete genome sequence of *Metallosphaera cuprina*, a metal sulfide-oxidizing archaeon from a hot spring. J. Bacteriol. **193**, 3387–3388
29 You, X. Y., Liu, C., Wang, S. Y., Jiang, C. Y., Shah, S. A., Prangishvili, D., She, Q., Liu, S. J. and Garrett, R. A. (2011) Genomic analysis of *Acidianus hospitalis* W1 a host for studying crenarchaeal virus and plasmid life cycles. Extremophiles **15**, 487–497
30 Blattner, F. R., Plunkett, G., 3rd, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F. et al. (1997) The complete genome sequence of *Escherichia coli* K-12. Science **277**, 1453–1462
31 Nelson, K. E., Clayton, R. A., Gill, S. R., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D., Nelson, W. C., Ketchum, K. A. et al. (1999) Evidence for lateral gene transfer between archaea and bacteria from genome sequence of *Thermotoga maritima*. Nature **399**, 323–329
32 Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. (1997) The Clustal_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25**, 4876–4882
33 Jongeneel, C. V., Bouvier, J. and Bairoch, A. (1989) A unique signature identifies a family of zinc-dependent metallopeptidases. FEBS Lett. **242**, 211–214
34 Hooper, N. M. (1994) Families of zinc metalloproteases. FEBS Lett. **354**, 1–6
35 Fukasawa, K., Fukasawa, K. M., Kanai, M., Fujii, S., Hirose, J. and Harada, M. (1998) Dipeptidyl peptidase III is a zinc metallo-exopeptidase. Molecular cloning and expression. Biochem. J. **329**, 275–282
36 Fukasawa, K., Fukasawa, K. M., Iwamoto, H., Hirose, J. and Harada, M. (1999) The HELLGH motif of rat liver dipeptidyl peptidase III is involved in zinc coordination and the catalytic activity of the enzyme. Biochemistry **38**, 8299–8303
37 Baral, P. K., Jajcanin-Jozic, N., Deller, S., Macheroux, P., Abramic, M. and Gruber, K. (2008) The first structure of dipeptidyl-peptidase III provides insight into the catalytic mechanism and mode of substrate binding. J. Biol. Chem. **283**, 22316–22324
38 Lew, R. A. (2004) The zinc metallopeptidase family: new faces, new functions. Protein Pept. Lett. **11**, 407–414
39 Pandey, D. P. and Gerdes, K. (2005) Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. Nucleic Acids Res. **33**, 966–976
40 Yamaguchi, Y., Park, J. H. and Inouye, M. (2011) Toxin-antitoxin systems in bacteria and archaea. Annu. Rev. Genet. **45**, 61–79
41 Maezato, Y., Daugherty, A., Dana, K., Soo, E., Cooper, C., Tachdjian, S., Kelly, R. M. and Blum, P. (2011) VapC6, a ribonucleolytic toxin regulates thermophilicity in the crenarchaeote *Sulfolobus solfataricus*. RNA **17**, 1381–1392
42 Valenti, A., Napoli, A., Ferrara, M. C., Nadal, M., Rossi, M. and Ciaramella, M. (2006) Selective degradation of reverse gyrase and DNA fragmentation induced by alkylating agent in the archaeon *Sulfolobus solfataricus*. Nucleic Acids Res. **34**, 2098–2108
43 Wagner, M., Berkner, S., Ajon, M., Driessen, A. J., Lipps, G. and Albers, S. V. (2009) Expanding and understanding the genetic toolbox of the hyperthermophilic genus *Sulfolobus*. Biochem. Soc. Trans. **37**, 97–101
44 She, Q., Zhang, C., Deng, L., Peng, N., Chen, Z. and Liang, Y. X. (2009) Genetic analyses in the hyperthermophilic archaeon *Sulfolobus islandicus*. Biochem. Soc. Trans. **37**, 92–96
45 She, Q., Deng, L., Zhu, H., Chen, Z., Dreibrøl, M., Awayez, M. and Liang, Y. X. (2008) Host-vector systems for hyperthermophilic archaeon *Sulfolobus*. In Microbes and the Environment: Perspective and Challenges (Liu, S. J. and Drake, H. L., eds), pp. 151–156, Science Press Beijing, Beijing
46 Deng, L., Zhu, H., Chen, Z., Liang, Y. X. and She, Q. (2009) Unmarked gene deletion and host-vector system for the hyperthermophilic crenarchaeon *Sulfolobus islandicus*. Extremophiles **13**, 735–746
47 Peng, N., Xia, Q., Chen, Z., Liang, Y. X. and She, Q. (2009) An upstream activation element exerting differential transcriptional activation on an archaeal promoter. Mol. Microbiol. **74**, 928–939

48 Zhang, C., Guo, L., Deng, L., Wu, Y., Liang, Y., Huang, L. and She, Q. (2010) Revealing the essentiality of multiple archaeal pcna genes using a mutant propagation assay based on an improved knockout method. Microbiology **156**, 3386–3397

49 Peng, N., Deng, L., Mei, Y., Jiang, D., Hu, Y., Awayez, M., Liang, Y. and She, Q. (2012) A synthetic arabinose-inducible promoter confers high levels of recombinant protein expression in hyperthermophilic archaeon *Sulfolobus islandicus*. Appl. Environ. Microbiol. **78**, 5630–5637

Received 20 July 2012/22 August 2012; accepted 24 August 2012
Published as Immediate Publication 5 September 2012, doi 10.1042/BSR20120074