

Research article

Open Access

Thermophile-specific proteins: the gene product of *aq_1292* from *Aquifex aeolicus* is an NTPase

Claudia Klinger, Michael Roßbach, Rebecca Howe and Michael Kaufmann*

Address: Institute for Neurobiochemistry, The Protein Chemistry Group, Witten/Herdecke University, Stockumer Str. 10, 58448 Witten, Germany

Email: Claudia Klinger - cklinger@uni-wh.de; Michael Roßbach - miro@uni-wh.de; Rebecca Howe - rebeccah@uni-wh.de;

Michael Kaufmann* - mika@uni-wh.de

* Corresponding author

Published: 23 September 2003

Received: 13 June 2003

BMC Biochemistry 2003, 4:12

Accepted: 23 September 2003

This article is available from: <http://www.biomedcentral.com/1471-2091/4/12>

© 2003 Klinger et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: To identify thermophile-specific proteins, we performed phylogenetic pattern searches of 66 completely sequenced microbial genomes. This analysis revealed a cluster of orthologous groups (COG1618) which contains a protein from every thermophile and no sequence from 52 out of 53 mesophilic genomes. Thus, COG1618 proteins belong to the group of thermophile-specific proteins (THEPs) and therefore we here designate COG1618 proteins as THEPIs. Since no THEPI had been analyzed biochemically thus far, we characterized the gene product of *aq_1292* which is THEPI from the hyperthermophilic bacterium *Aquifex aeolicus* (aaTHEPI).

Results: aaTHEPI was cloned in *E. coli*, expressed and purified to homogeneity. At a temperature optimum between 70 and 80°C, aaTHEPI shows enzymatic activity in hydrolyzing ATP to ADP + P_i with $k_{\text{cat}} = 5 \times 10^{-3} \text{ s}^{-1}$ and $K_{\text{m}} = 5.5 \times 10^{-6} \text{ M}$. In addition, the enzyme exhibits GTPase activity ($k_{\text{cat}} = 9 \times 10^{-3} \text{ s}^{-1}$ and $K_{\text{m}} = 45 \times 10^{-6} \text{ M}$). aaTHEPI is inhibited competitively by CTP, UTP, dATP, dGTP, dCTP, and dTTP. As shown by gel filtration, aaTHEPI in its purified state appears as a monomer. The enzyme is resistant to limited proteolysis suggesting that it consists of a single domain. Although THEPIs are annotated as "predicted nucleotide kinases" we could not confirm such an activity experimentally.

Conclusion: Since aaTHEPI is the first member of COG1618 that is characterized biochemically and functional information about one member of a COG may be transferred to the entire COG, we conclude that COG1618 proteins are a family of thermophilic NTPases.

Background

The comparison of protein coding sequences from complete genomes led to a classification, based on sequence similarities, that assigns proteins to clusters of orthologous groups (COGs) [1,2]. Phylogenetic patterns search [3] is a tool to retrieve COGs that contain protein sequences that match a certain predefined pattern of organisms. Recently, Forterre [4] used this tool to identify reverse gyrase as the only hyperthermophile-specific pro-

tein. Another survey identified 30 COGs enriched in hyperthermophilic procaryotes [5]. In contrast, we here extended a similar search with respect to a group consisting of both thermophiles and hyperthermophiles *i. e.* the currently 3 organisms with optimal growth temperatures between 55 and 80°C were included. This search retrieved COG1618 as one of the high ranking clusters containing a protein from every of the 13 thermophilic and, by including only one representative, the fewest sequences

Table 1: COGs containing the THEPs with the highest thermophile-specificity.

	Number of mesophilic genomes in COG						
	0	1	2	3	4	5	
Number of thermophilic genomes in COG	13	-/-	COG1618	COG3635	-/-	COG1355 COG1371	COG2078
	12	COG1980	-/-	-/-	COG1144 COG1820	COG1730	COG1503 COG1867
	11	COG1581 COG1888	COG1350 COG1909				
	10	COG1110					
	9	COG1318 COG1630 COG2250					

Given are the COGs that contain sequences from the most among the 13 thermophilic and the fewest among the 53 mesophilic genomes of the current COG database.

from mesophilic genomes of the current COG database. Thus, amongst others, the COG1618 proteins belong to a group of thermophile-specific proteins (THEPs) and in this report we designate them as THEP1s.

Orthologs typically have the same function, allowing the transfer of functional information from one member of a COG to the entire COG. To the best of our knowledge, no function could be established for any of the THEP1s thus far. On the other hand, Cort *et al.* predicted THEP1s being ATPases [6]. This paper confirms their hypothesis experimentally by characterizing the gene product of *aq_1292* from *Aquifex aeolicus* (aaTHEP1) as a first example representative for this protein family. *Aquifex aeolicus* is a hyperthermophilic bacterium growing optimally at 95 °C [7].

Results

Identification of COG1618

Among the 4,873 COGs of the current COG database, phylogenetic patterns search revealed 167 COGs that include a sequence from each of the 13 thermophilic genomes (thermophiles and hyperthermophiles). A further analysis showed that 63 COGs contain a sequence from every genome, indicating that those conserved proteins are probably essential for microbial life. Consequently, the remaining 104 COGs are more or less thermophile-specific proteins. To retrieve absolute thermophile-specific COGs the search parameters were chosen in the way that the output COGs must contain a sequence from every thermophilic and must not contain any sequence from a mesophilic genome. Whereas no COG exactly met this search condition, less thermophile-specific COGs with only one or a few exceptions from the exact search condition were detected (table 1). Among those, COG1618 is the only one that contains a sequence

from each thermophilic and only one further sequence from a mesophilic genome (MA3402).

Sequence analysis of THEP1s

CLUSTALW analysis of all THEP1s led to the design of suitable experiments to assay catalytic in vitro function of aaTHEP1. All proteins contain both the Walker A (GxxxxGK [ST]) and Walker B motif (4 × hydrophobic [DE]xxG) indicating that aaTHEP1 is a P-loop NTPase (figure 1). Since the distal [NT]KxD motif responsible for guanine specificity [8] is absent, aaTHEP1 probably do not belong to the GTPase superclass of the P-loop NTPases. Cort *et al.* stated that THEP1s "probably represent a novel family of ATPases in both sequence and structural terms" [6]. The authors determined the structure of MTH538 from *Methanobacterium thermoautotrophicum* by NMR and based on extensive sequence and structure comparisons proposed that the relationship between MTH583 and THEP1s might be similar to that between CheY and ATPase/phosphatase members of the HAD family as suggested by Ridder and Dijkstra [9].

Purification of recombinant aaTHEP1

The bulk of *E. coli* proteins could be removed by heating the crude cell extract for 10 min at 75 °C. As expected from the sequence-based theoretical pI of 9.88, aaTHEP1 binds to the matrix of a cation exchanger. Ion exchange followed by hydrophobic interaction chromatography finally resulted in a homogenous aaTHEP1 preparation that migrates slightly higher than the calculated molecular weight of 20,555 Da (figure 2).

Functional activity of aaTHEP1

As shown in figure 3, purified aaTHEP1 clearly catalyzes the hydrolysis of ATP to ADP and P_i. As can be seen, the longer the reaction mixture was incubated the more ³²P_i

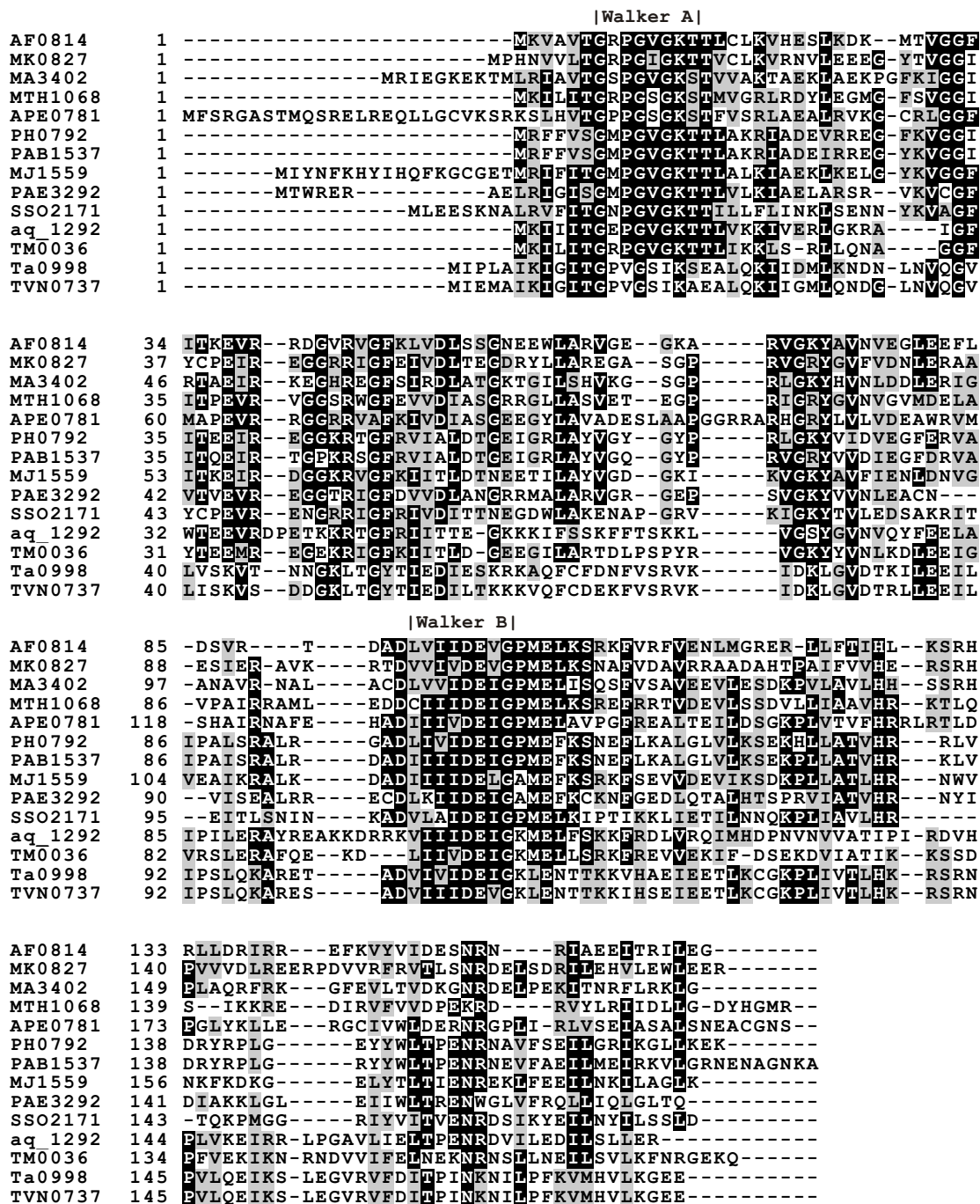


Figure 1
THEPIs contain Walker A and Walker B motifs. CLUSTALW analysis of all THEPIs. The positions of the Walker A and Walker B motifs are indicated. The shown sequences are from *Archaeoglobus fulgidus* (AF0814), *Methanopyrus kandleri* AV19 (MK0827), *Methanosarcina acetivorans* (MA3402), *Methanothermobacter thermautotrophicus* (MTH1068), *Aeropyrum pernix* (APE0781), *Pyrococcus horikoshii* (PH0792), *Pyrococcus abyssi* (PAB1537), *Methanococcus jannaschii* (MJ1559), *Pyrobaculum aerophilum* (PAE3292), *Sulfolobus solfataricus* (SSO2171), *Aquifex aeolicus* (aq_1292), *Thermotoga maritima* (TM0036), *Thermoplasma acidophilum* (Ta0998), and *Thermoplasma volcanium* (TVN0737).

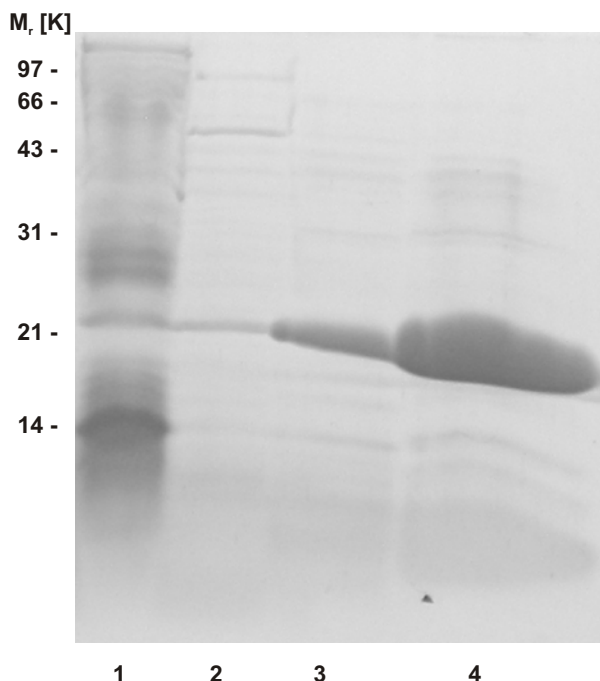


Figure 2
Recombinant aaTHEP1 can be expressed and purified. SDS-PAGE after different steps during purification of recombinant aaTHEP1. Lane 1: crude cell extract, lane 2: supernatant after heat treatment, lane 3: eluate after cation exchange chromatography, lane 4: eluate after hydrophobic interaction chromatography. To demonstrate the purity of the final preparation, lane 4 intentionally was overloaded.

was released from $[\gamma\text{-}^{32}\text{P}]\text{ATP}$. Since THEP1s are annotated as "predicted nucleotide kinases", we assayed aaTHEP1 for nucleoside diphosphate kinase and nucleoside monophosphate kinase activities. Using ATP as the phosphate donor and GDP (figure 4), GMP, AMP, and UMP (figure 5) as acceptors we could not detect the predicted phosphoryl transfer.

Temperature dependence

To show that aaTHEP1 is a thermophilic enzyme, thermal activity was determined by measuring ATP hydrolysis as catalyzed by the purified enzyme at different temperatures. Since spontaneous ATP hydrolysis also occurs at higher temperatures, those rates were measured and shown as well. As can be seen in figure 6, aaTHEP1 is still active at 90°C. As an optimum temperature with respect to the signal to noise ratio, 70°C was chosen for all further kinetic measurements.

Steady-state kinetics

aaTHEP1 catalyzes both ATP and GTP hydrolysis. Measuring the turnover rates at different substrate concentrations under steady-state conditions resulted in hyperbolic curves if presented in a double linear plot *i. e.* no cooperativity could be observed (figure 7). For that reason, steady-state kinetics could be further analyzed by fitting the data points to curves obeying the Michaelis-Menten equation retrieving k_{cat} - and K_{m} -values for each substrate. Compared to the hydrolysis of ATP, the maximum turnover rate k_{cat} for GTP is faster by a factor of 2. On the other hand, the enzyme's substrate affinity to ATP as represented by K_{m} exceeds that to GTP by one order of magnitude. The catalytic efficiency of an enzyme is defined as $k_{\text{cat}}/K_{\text{m}}$. Thus, the catalytic efficiency of ATP hydrolysis as catalyzed by aaTHEP1 significantly exceeds that of GTP hydrolysis. The specific activities for ATP and GTP hydrolysis corresponding to the k_{cat} -values given in figure 7 are 14.6 and 26.3 $\text{nmol min}^{-1} \text{mg}^{-1}$, respectively.

Inhibition of ATP-hydrolysis by other nucleotides

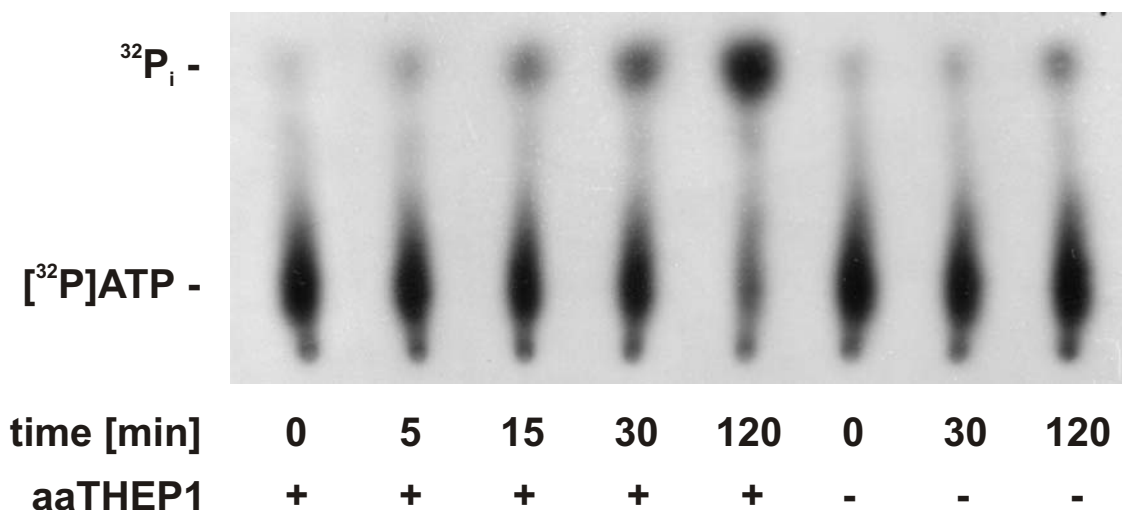
Competition experiments demonstrated that all nucleosidetriphosphates inhibit aaTHEP1 in hydrolyzing ATP. The data points fit to a model for competitive inhibition. As an example, inhibition by GTP is shown in figure 8. With the exception of GTP and dGTP, the K_{i} -values of all nucleotides are approximately in a similar range as the K_{m} for ATP (table 3). Whereas GDP also inhibits ATP-hydrolysis, GMP, AMP, and UMP show no inhibition (figures 4 and 5).

Possible homologous non-covalent interactions

Gel filtration on Superose 6 prep grade was performed under native conditions to detect possible aaTHEP1 oligomers. We observed a single peak at $K_{\text{av}} = 4.24$ corresponding to a relative molecular weight of 17,700 indicating that purified aaTHEP1 exists as a highly compact folded monomer (figure 9). In addition, several attempts to crosslink aaTHEP1 by glutaraldehyde resulted in a non-crosslinked protein whereas under the same experimental conditions β -tryptophan synthase resulted in a crosslinked dimer (data not shown).

Domain structure

A possible multidomain structure of aaTHEP1 was probed by limited proteolysis. As a control, the β -subunit of *E. coli* tryptophan synthase as a typical protein composed of distinct domains connected by a hinge region [10] was also proteolyzed. In contrast to β -tryptophan synthase, aaTHEP1 is resistant to proteolytic cleavage by both trypsin and endoproteinase Glu-C (figure 10). Compared to published experiments on the proteolytic cleavage of β -tryptophan synthase [10, 11], the experimental conditions were chosen in a way that fragmented β -tryptophan synthase (~ 30 and ~ 10 kDa fragments) already were further

**Figure 3**

aaTHEP1 is an ATPase. Autoradiography of thin-layer chromatograms showing samples containing $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ after different times of incubation at 70°C. Measurements were performed at 50 μM ATP. 0.5 μg of purified aaTHEP1 was used for each assay in 25 μl buffer.

degraded. Even under these conditions, aaTHEP1 remains stable although it contains 36 ($20 \times \text{K} + 16 \times \text{R}$) possible trypsin and 17 ($17 \times \text{E}$) possible endoproteinase Glu-C cleavage sites as predicted from the sequence.

Secondary structure analysis

To determine the overall content of secondary structure elements of purified aaTHEP1, far UV circular dichroism spectroscopy was performed. A typical spectrum is given in figure 11. Analysis of the CD spectrum using the Jasco Spectrum Analyzer software revealed an alpha helix content of 34.3 % and 33.6 % beta-sheets.

Discussion

Our approach is based on the assumption that proteins abundant in thermophilic and rare in mesophilic genomes are the most attractive targets for further biochemical investigations to understand the physiology specific for thermophiles in more detail. Consequently, our first goal was to identify such candidates via bioinformatic methods and as the most suitable protein THEP1 was selected for this study. The high ranking thermophile-specificity of THEP1 among procaryotes may be explained by an essential physiological role in thermophiles that is

of no functional relevance for almost all mesophilic microorganisms. As an alternative explanation, a function also present in mesophilic organisms could be carried out by a protein that was not able to adapt to higher temperatures and in the course of convergent evolution, THEP1 could have taken over this particular function. *Methanosarcina acetivorans* str. C2A is the only mesophilic organism containing THEP1 (MA3402). Since the genome of *M. acetivorans* reveals extensive metabolic and physiological diversity and there are thermophilic Methanosarcinae [12], one may take into consideration the possibility that *M. acetivorans* facultatively could be thermophilic. In addition to phylogenetic patterns search we also performed BLAST with the aaTHEP1 sequence. However, although no significant homologies to further sequences from mesophilic unicellular organisms could be detected, we discovered homologies to 6 multicellular eucaryotes. Consequently, THEP1s could belong to a class of proteins that are conserved in Archaea and Eukarya (with losses) and have been passed to thermophilic bacteria by lateral gene transfer.

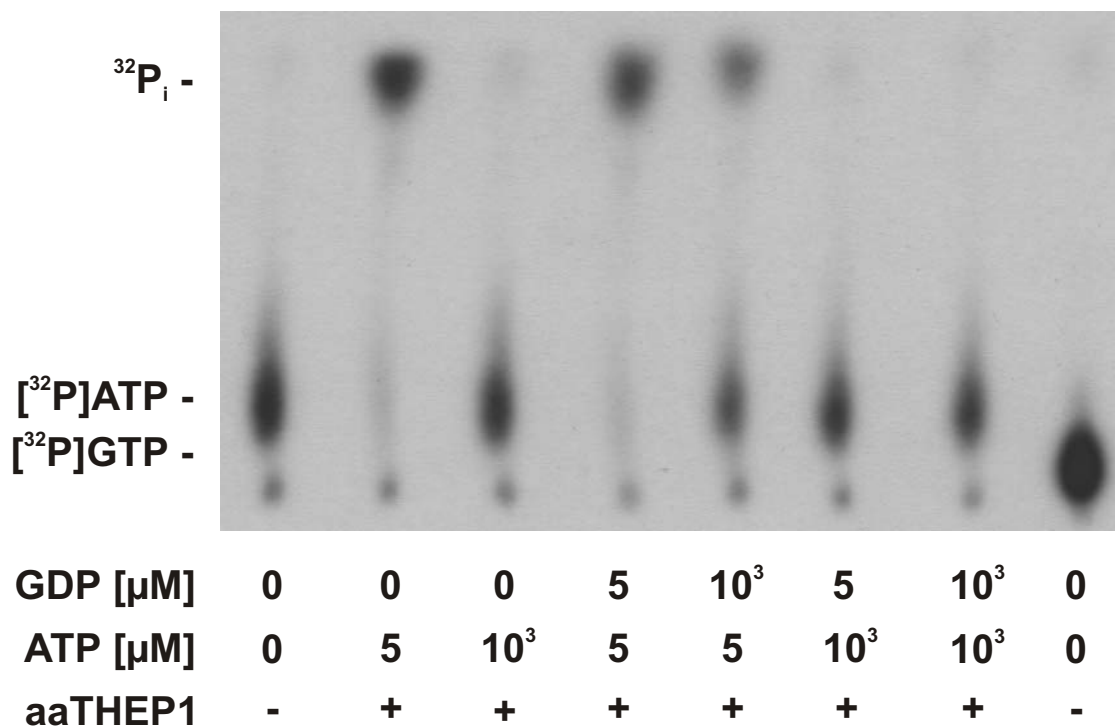


Figure 4
aaTHEP1 is no NDP kinase. Autoradiography of thin-layer chromatograms showing samples containing $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ after five minutes of incubation at 70°C and at different concentrations of ATP and GDP. 1 μg of purified aaTHEP1 was used for each assay in 25 μl buffer.

The present data show that aaTHEP1 catalyzes ATP and GTP hydrolysis *in vitro* as predicted by Cort *et al.* [6]. In contrast, the annotated nucleotide kinase activity could not be confirmed experimentally.

As expected, the observed turnover rates are too low to represent a physiological *in vivo* situation where the free energy of ATP hydrolysis is coupled to energy consuming tasks. *In vitro*, similar turnover rates of purified NTPase are published in the literature: $1.1 \times 10^{-2} \text{ sec}^{-1}$ for PilT from *A. aeolicus* [13], $1.2 \times 10^{-2} \text{ sec}^{-1}$ for TadA for *Actinobacillus actinomycetemcomitans* [14], $3 \times 10^{-3} \text{ sec}^{-1}$ for TrwD from *Escherichia coli* [15], and $6.9 \times 10^{-4} \text{ sec}^{-1}$ for PilQ from *Escherichia coli* [16]. *In vivo*, additional proteins might be needed to activate aaTHEP1 by protein-protein interactions and take up the free energy released by NTP-hydrolysis *e. g.* for motion, active transport or another energy

consuming cellular function. Alternatively, aaTHEP1 could catalyze an NTP driven thermodynamically unfavourable anabolic reaction of a yet undiscovered cosubstrate or play a role in cellular regulation.

ATP-hydrolysis could be inhibited by all other 7 nucleosidetriphosphates under investigation in a competitive manner which may be interpreted that in addition to GTP the other nucleotides are also substrates for isolated aaTHEP1. Possibly the enzyme lost a certain *in vivo* substrate specificity upon isolation. Since GDP inhibits ATP-hydrolysis whereas GMP, AMP and UMP do not, we propose that aaTHEP1 recognizes the β- and γ-phosphates rather than the nucleoside moieties.

A low activity of a recombinant protein may also be explained by trace amounts of *E. coli* enzymes still present

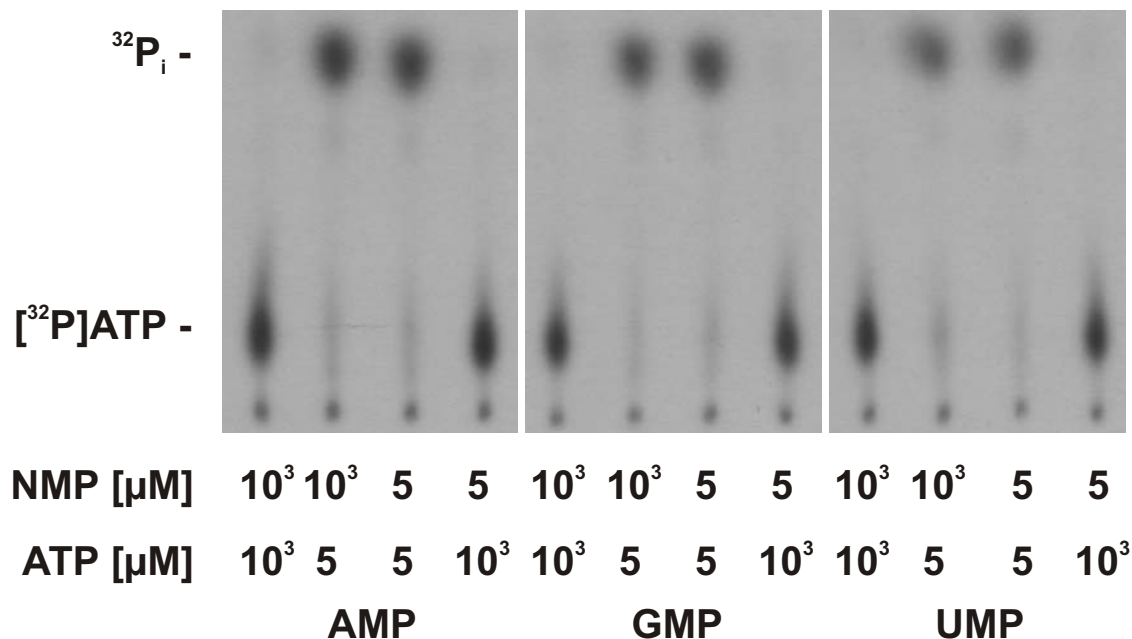


Figure 5
aaTHEP1 is no NMP kinase. Autoradiography of thin-layer chromatograms showing samples containing [γ -³²P]ATP after five minutes of incubation at 70°C and at different concentrations of ATP, AMP, GMP, and UMP, respectively. 1 μg of purified aaTHEP1 was used for each assay in 25 μl buffer.

after purification. For the measured NTPase activity of aaTHEP1 we exclude this possibility because we do not expect an *E. coli* enzyme exhibiting the same temperature dependence as shown in figure 6. There are even more indications that aaTHEP1 is a thermostable protein. The elution during gel filtration at a position corresponding to a lower relative molecular weight than the calculated 20,555 Da for a monomer and the resistance to limited proteolysis suggest that aaTHEP1 is a highly compact folded one-domain protein, a well known feature for thermophilic proteins. No cooperative behaviour in the kinetic experiments is also expected for a monomeric enzyme.

Although this study clearly defines a biochemical *in vitro* activity of aaTHEP1, there are many possible *in vivo* functions and bioinformatic analysis allows to make some predictions. Based on the genomic context, gene functions can be predicted by searching for the conservation of

operons and gene orders because genes found in gene strings, particularly in multiple genomes, can be legitimately assumed to be functionally linked [17]. For THEP1, we indeed detected 4 genomes where the THEP1-gene immediately is followed on the same strand by a COG1867 protein (N2, N2-dimethylguanosine tRNA methyltransferase). Furthermore, COG1867 also belongs to the group of THEPs indicating a functional link (table 1).

Conclusions

This study experimentally confirms the hypothesis of Cort *et al.* who suggested THEP1s being a novel family of ATPases in both sequence and structural terms [6]. On the other hand, we refute the theoretical prediction that THEP1s are nucleotide kinases. In addition to the experimental work, a list of further THEPs as potential targets to study microbial thermophily is provided (tables 1 and 2).

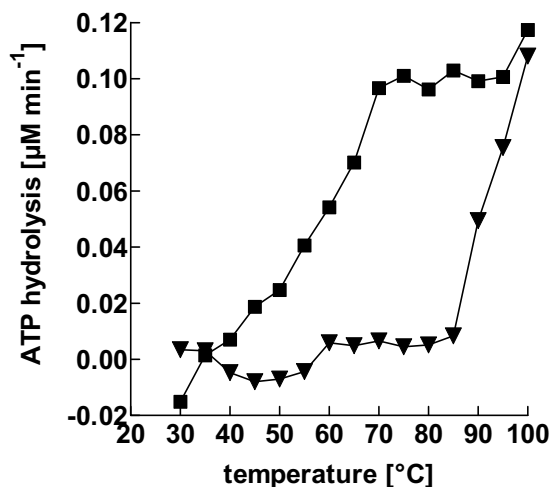


Figure 6
aaTHEP1 is a thermophilic enzyme. Temperature dependence of aaTHEP1 catalyzed ATP hydrolysis. Measurements were performed at 5 µM ATP in buffer A. ATP-hydrolysis was measured in the presence of aaTHEP1 (squares) and spontaneous ATP-degradation was determined in the absence of aaTHEP1 (triangles). 0.5 µg of purified aaTHEP1 was used for each assay in 25 µl buffer.

Methods

Cloning, expression and purification of aaTHEP1 from *A. aeolicus*

Genomic DNA from *A. aeolicus* was kindly provided by Dr. R. Huber, Regensburg, Germany. *aq_1292* was amplified by PCR using the primers 5'-CACCATGAAAATCAT-CATAACCGGTGA-3' and 5'-TTACCGCTCAAGAAGTGAGAGAAT-3'. The PCR fragment was inserted into the pre-linearized plasmid pET101/D-TOPO (Invitrogen). For propagation and maintenance, the *E. coli* strain TOP10 (Invitrogen) was used. The correct sequence of the insert as well as its orientation were verified by sequence analysis using the reverse T7-primer 5'-TAGTTATTGCTCAGCGGTGG-3'.

To express *aq_1292*, *E. coli* BL21 Star™ (DE3) was used. According to the instructions of the manufacturer, freshly transformed cells were grown by transferring the entire transformation mixture to 10 ml Luria-Bertani broth (LB) medium containing 50 µg/ml carbenicillin and 1% glucose. After growing for 4 h at 37°C, the preculture was added to 40 ml of fresh medium, grown for additional 16 h and then 30 ml were used to inoculate a 2 l main culture. At $A_{600} = 0.7$, protein expression was induced by 1 mM isopropyl β-D-thiogalactopyranoside (IPTG, Roth).

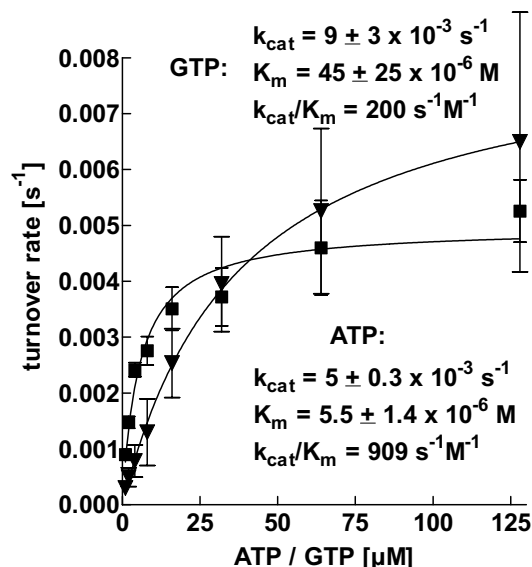


Figure 7
aaTHEP1 hydrolyzes ATP and GTP obeying the Michaelis-Menten-equation. Steady-state kinetics of ATP (squares) and GTP hydrolysis (triangles) as catalyzed by aaTHEP1 at 70°C. Each data point for ATP hydrolysis is the mean ± SD of 4, each data point for GTP-hydrolysis the mean ± SD of 3 independent measurements. Since the signal to noise ratio of the activity measurement dramatically increases at higher substrate concentrations, 128 µM was the highest concentration under investigation. 0.5 µg of purified aaTHEP1 was used for each assay in 25 µl buffer.

Cells were harvested at $A_{600} = 1.3$, yielding approximately 4 g after centrifugation at $6,000 \times g$ for 10 min. The pellet was stored at -20°C before use.

Buffer A for protein purification was 50 mM Tris/HCl, 25 mM $MgCl_2$, 5 mM KCl, 1 mM DL-dithiothreitol (DTT), 0.1 mM ethylenediaminetetraacetic acid (EDTA), pH 7.0. To disrupt the cell walls, 20 ml of buffer A containing 10 instead of 50 mM Tris, 5 mM sodium deoxycholate, 40,000 U/ml lysozyme, and 50 µg/ml DNaseI were added to 1 g of thawed cells. Lysis was performed by stirring the suspension at 20°C for 1 h and the cell debris were removed by centrifugation at $20,000 \times g$ for 30 min. To precipitate the bulk of *E. coli* proteins, the supernatant was heated to 75°C for 10 min and immediately chilled on ice. Denatured proteins were removed by centrifugation at $20,000 \times g$ and 4°C for 30 min. To remove nucleic acids and further purify recombinant aaTHEP1, the supernatant was loaded at 0.5 ml/min onto a 1 ml HiTrap™ SP HP

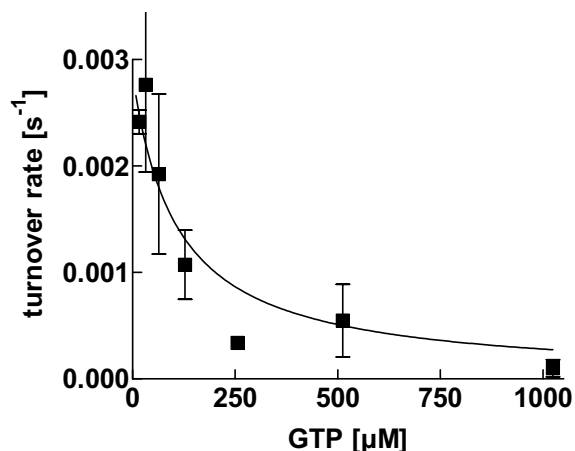


Figure 8
ATP hydrolysis is inhibited by GTP. Inhibition of aaTHEP1 catalyzed ATP hydrolysis by GTP. Each data point represents the mean \pm SD of two independent measurements. 0.5 μ g of purified aaTHEP1 were used for each assay in 25 μ l buffer.

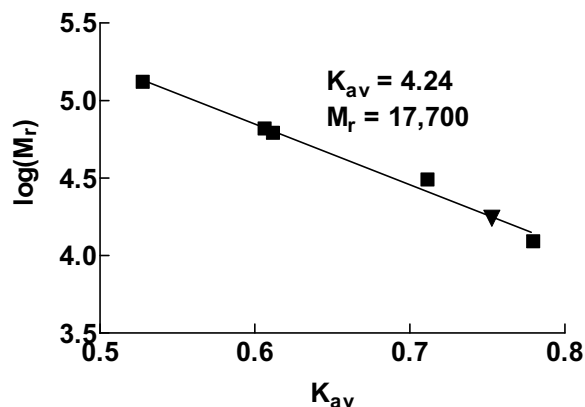


Figure 9
Isolated aaTHEP1 appears a monomer. Gel filtration on calibrated Superose™ 6 prep grade. Native marker protein peaks are represented as squares and aaTHEP1 is shown as a triangle in a linear regression of K_{av} versus $\log(M_r)$.

cation exchange column (Amersham Biosciences AB) equilibrated with buffer A. The column was washed with buffer A containing 150 mM KCl and aaTHEP1 was eluted at 500 mM KCl. To the collected peak a 4-fold volume of

Table 2: Functional predictions of high-scoring COGs.

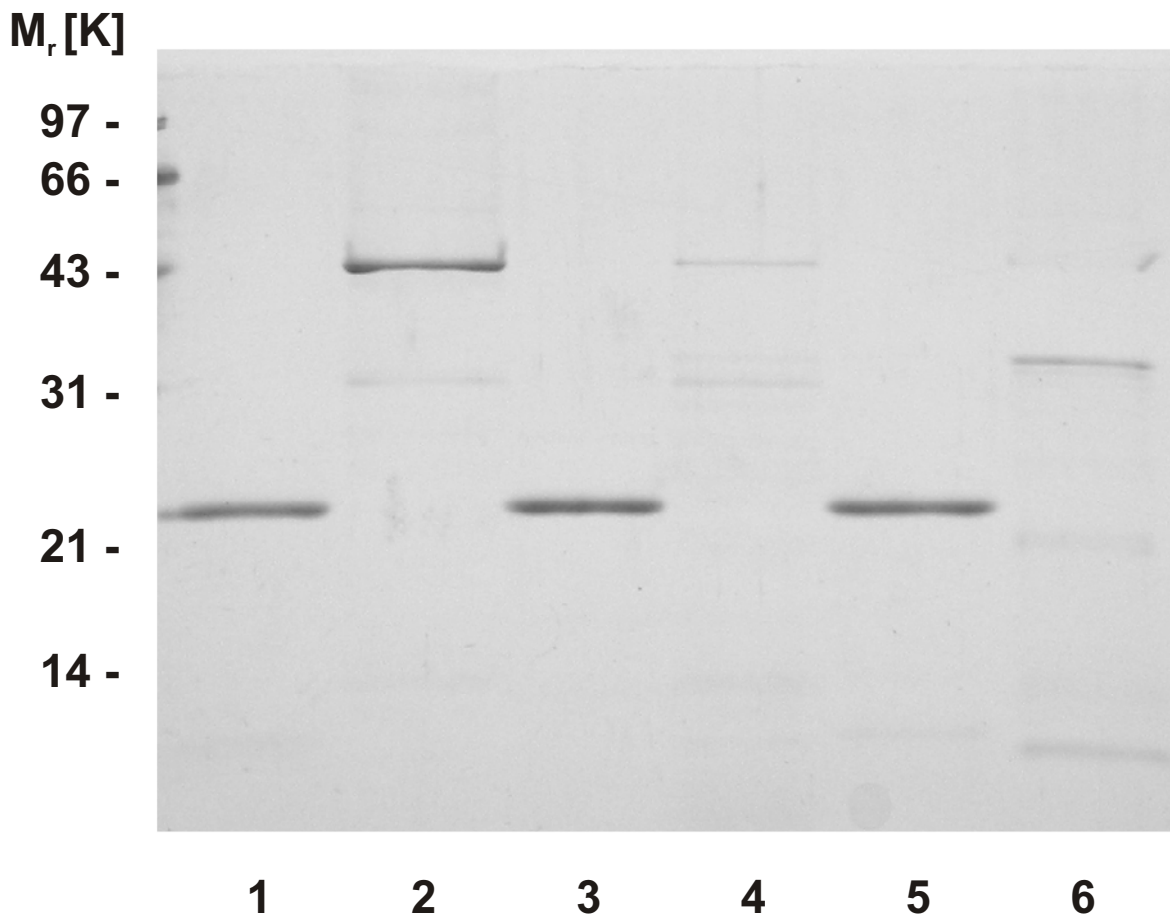
COG name	Predicted function
COG1618	Predicted nucleotide kinase
COG3635	Predicted phosphoglycerate mutase, AP superfamily
COG1355	Predicted dioxygenase
COG1371	Uncharacterized conserved protein
COG2078	Uncharacterized conserved protein
COG1980	Archaeal fructose 1,6-bisphosphatase
COG1144	Pyruvate:ferredoxin oxidoreductase and related 2-oxoacid:ferredoxin oxidoreductases, delta subunit
COG1820	N-acetylglucosamine-6-phosphate deacetylase
COG1730	Predicted prefoldin, molecular chaperone implicated in de novo protein folding
COG1503	Peptide chain release factor I (eRF1)
COG1867	N2, N2-dimethylguanosine tRNA methyltransferase
COG1581	Archaeal DNA-binding protein
COG1888	Uncharacterized protein conserved in archaea
COG1350	Predicted alternative tryptophan synthase beta-subunit (paralog of TrpB)
COG1909	Uncharacterized protein conserved in archaea
COG1110	Reverse gyrase
COG1318	Predicted transcriptional regulators
COG1630	Uncharacterized protein conserved in archaea
COG2250	Uncharacterized conserved protein related to C-terminal domain of eukaryotic chaperone, SACSIN

Descriptions are taken from the current COG-database.

buffer A containing 2.438 M $(\text{NH}_4)_2\text{SO}_4$ (62.5 % saturation) instead of KCl was added. After centrifugation at $20,000 \times g$ for 30 min, the resulting protein solution was applied to a 1 ml HiTrap™ (low sub) Phenyl Sepharose™ FF (Amersham Biosciences AB) equilibrated in buffer A containing 1.85 M $(\text{NH}_4)_2\text{SO}_4$ (50 % saturation) at 0.5 ml/min. The column was washed at 40 % and eluted at 20 % $(\text{NH}_4)_2\text{SO}_4$ saturation. Finally, $(\text{NH}_4)_2\text{SO}_4$ was removed from aaTHEP1 by using a NAP™ 5 column (Amersham Biosciences AB) equilibrated in buffer A. The purified protein could be stored in buffer A for several weeks at -20°C without significant loss in activity. Protein concentrations were determined by densitometry [18] of SDS-Gels [19] using the Sigma bovine serum albumin protein micro standard. Gels were analyzed via an imaging densitometer (Bio-Rad GS-700) and the Molecular Analyst software.

Steady state kinetics

To measure ATP or GTP hydrolysis, the release of $[\gamma\text{-}^{32}\text{P}]$ from $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ or $[\gamma\text{-}^{32}\text{P}]\text{GTP}$ was determined. Assays were performed using purified aaTHEP1 in 25 μ l buffer A. Aliquots of the reaction mixture were stopped after different times of incubation by adding 25 μ l of 40 % formic acid and the mixture was separated by thin layer chromatography on PEI-cellulose in 0.5 M potassium phosphate pH 3.9. Quantification was performed by scanning exposed and developed Hyperfilm™ – β max films

**Figure 10**

aaTHEP1 consists of a single domain. SDS-PAGE after limited proteolysis by trypsin and endoproteinase Glu-C. aaTHEP1 is shown on lanes 1, 3 and 5. As a control, the β -subunit of tryptophan synthase from *E. coli* is run on lanes 2, 4, and 6. Lanes 1 and 2 show the native proteins without the addition of proteases. Proteolysis by trypsin is seen on lanes 3 and 4 whereas proteolysis by endoproteinase Glu-C is shown on lanes 5 and 6.

(Amersham Biosciences AB) using an imaging densitometer (Bio-Rad GS-700) and the Molecular Analyst software. Each catalytic activity was determined by the average of end point measurements divided by time of at least three different times of incubation under steady-state conditions. Blank values determined in the absence of aaTHEP1 were subtracted from each data point. Unless stated otherwise, the experiments were performed at 70 °C in a Biometra T3 thermocycler. Competition experiments were measured at a constant concentration of 5 μ M ATP. Nonlinear regressions to determine k_{cat} - and K_m -values were performed with the GraphPad Prism™ software using

the Michaelis-Menten equation. K_i -values were determined by fitting the data points to

$$V = (V_{max} \times [ATP]) / ([ATP] + K_m \times (1 + [Inhibitor]/K_i)).$$

Nucleotide kinase activities were assayed at 5 and 1000 μ M ATP by adding 5 as well as 1000 μ M of GDP, UMP, AMP, and GMP respectively.

Gel filtration

To determine if purified aaTHEP1 is in mono- or oligomeric form, a 1 × 50 cm column filled with superose 6 prep grade (Amersham Biosciences AB) was calibrated

Table 3: Inhibition constants of ATP hydrolysis by nucleoside-3-phosphates.

	GTP	CTP	UTP	dATP	dGTP	dCTP	dTTP
K_i [μ M]	57 \pm 20	6 \pm 2	5 \pm 2	3 \pm 0.2	18 \pm 2	9 \pm 3	6 \pm 2

The K_i -values of inhibition of ATP hydrolysis as determined at 5 μ M ATP at 70°C are summarized. 0.5 μ g of purified aaTHEP1 was used for each assay in 25 μ l buffer.

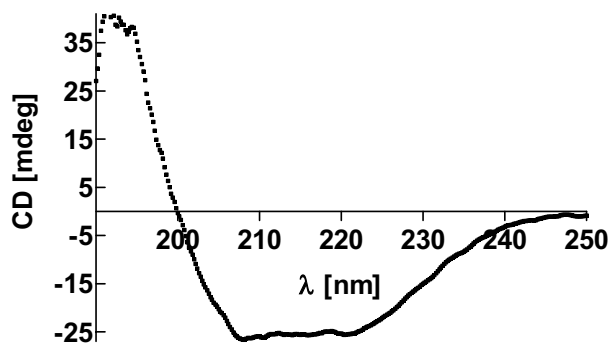


Figure 11
aaTHEP1 consists of 33.6% beta-sheets and 34.3% alpha-helix. Secondary structure analysis of purified aaTHEP1 by far UV circular dichroism spectroscopy.

with cytochrom C, carbonic anhydrase monomer and dimer [20] and bovine serum albumin monomer and dimer [21] as molecular weight standards. Proteins were run at 0.2 ml/min in buffer A containing 150 mM KCl. Approximately 80 μ g of aaTHEP1 or marker protein in 0.5 ml were applied to the column. To dissociate possible cold-induced unspecific aggregates, aaTHEP1 was heated for 10 min at 75°C prior to gel filtration.

Limited proteolysis

In order to probe the domain structure of aaTHEP1, limited proteolysis by trypsin and endoproteinase Glu-C was performed. The reactions were carried out in buffer A containing 4 μ g/ml protease and 40 μ g/ml protein substrate. After incubation at 20°C for 1 h, samples were analyzed by SDS-PAGE. The β -subunit of tryptophan synthase from *E. coli* was prepared as described earlier [22].

Circular dichroism

Far UV circular dichroism spectra were recorded using a J-810 CD spectralpolarimeter (Jasco) at 30°C. 2.5 μ g of purified aaTHEP1 was measured in 200 μ l of buffer A in an 1 mm quartz cuvette. Helix- and beta-sheet contents

were calculated using the Spectrum Analyzer software (Jasco).

Bioinformatics

COGs were analyzed by extended phylogenetic patterns search (EPPS) [23,24] using the march 5, 2003 release of the COG database. Multiple sequence alignments were performed at EMBL Outstation [25] using CLUSTALW 1.81 [26] and the data was visualized using BOXSHADE v3.21 [27]. BLAST searches were performed at NCBI [28].

Authors' contributions

CK carried out the molecular biology experiments, the purification of aaTHEP1, and the gel filtrations. MR performed the CD-measurements, RH performed the activity measurements and limited proteolysis experiments. MK carried out the bioinformatics, conceived of the study, and participated in its design and coordination. All authors read and approved the final manuscript.

Acknowledgements

We thank Andreas Savelsberg for his help in performing the kinetic measurements and Reiner Würdinger for assistance during the CD-measurements at Jasco GmbH Germany. We greatly appreciate discussions with Marina V. Rodnina on all aspects of the manuscript.

References

1. Tatusov RL, Koonin EV and Lipman DJ: **A genomic perspective on protein families.** *Science* 1997, **278**:631-637.
2. Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS, Kiryutin B, Galperin MY, Fedorova ND and Koonin EV: **The COG database: new developments in phylogenetic classification of proteins from complete genomes.** *Nucleic Acids Res* 2001, **29**:22-28.
3. **Phylogenetic patterns search** [<http://www.ncbi.nlm.nih.gov/COG/phylox.html>]
4. Forterre P: **A hot story from comparative genomics: reverse gyrase is the only hyperthermophile-specific protein.** *Trends Genet* 2002, **18**:236-237.
5. Makarova KS, Wolf YI and Koonin EV: **Potential genomic determinants of hyperthermophily.** *Trends Genet* 2003, **19**:172-176.
6. Cort JR, Yee A, Edwards AM, Arrowsmith CH and Kennedy MA: **Structure-based functional classification of hypothetical protein MTH538 from Methanobacterium thermoautotrophicum.** *J Mol Biol* 2000, **302**:189-203.
7. Deckert G, Warren PV, Gaasterland T, Young WG, Lenox AL, Graham DE, Overbeek R, Snead MA, Keller M, Aujay M, Huber R, Feldman RA, Short JM, Olsen GJ and Swanson RV: **The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.** *Nature* 1998, **392**:353-358.
8. Leipe DD, Wolf YI, Koonin EV and Aravind L: **Classification and evolution of P-loop GTPases and related ATPases.** *J Mol Biol* 2002, **317**:41-72.

9. Ridder IS and Dijkstra BW: **Identification of the Mg²⁺-binding site in the P-type ATPase and phosphatase members of the HAD (haloacid dehalogenase) superfamily by structural similarity to the response regulator protein CheY.** *Biochem J* 1999, **339** (Pt 2):223-226.
10. Kaufmann M, Schwarz T, Jaenicke R, Schnackerz KD, Meyer HE and Bartholmes P: **Limited proteolysis of the beta 2-dimer of tryptophan synthase yields an enzymatically active derivative that binds alpha-subunits.** *Biochemistry* 1991, **30**:4173-4179.
11. Hogberg-Raibaud A and Goldberg ME: **Preparation and characterization of a modified form of beta2 subunit of Escherichia coli tryptophan synthetase suitable for investigating protein folding.** *Proc Natl Acad Sci U S A* 1977, **74**:442-446.
12. Galagan JE, Nusbaum C, Roy A, Endrizzi MG, Macdonald P, FitzHugh W, Calvo S, Engels R, Smirnov S, Atnoor D, Brown A, Allen N, Naylor J, Stange-Thomann N, DeArellano K, Johnson R, Linton L, McEwan P, McKernan K, Talamas J, Tirrell A, Ye W, Zimmer A, Barber RD, Cann I, Graham DE, Grahame DA, Guss AM, Hedderich R, Ingram-Smith C, Kuettner HC, Krzycki JA, Leigh JA, Li W, Liu J, Mukhopadhyay B, Reeve JN, Smith K, Springer TA, Umayam LA, White O, White RH, Conway de Macario E, Ferry JG, Jarrell KF, Jing H, Macario AJ, Paulsen I, Pritchett M, Sowers KR, Swanson RV, Zinder SH, Lander E, Metcalf WW and Birren B: **The genome of M. acetivorans reveals extensive metabolic and physiological diversity.** *Genome Res* 2002, **12**:532-542.
13. Herdendorf TJ, McCaslin DR and Forest KT: **Aquifex aeolicus PilT, homologue of a surface motility protein, is a thermostable oligomeric NTPase.** *J Bacteriol* 2002, **184**:6465-6471.
14. Bhattacharjee MK, Kachlany SC, Fine DH and Figurski DH: **Non-specific adherence and fibril biogenesis by Actinobacillus actinomycetemcomitans: TadA protein is an ATPase.** *J Bacteriol* 2001, **183**:5927-5936.
15. Rivas S, Bolland S, Cabezon E, Goni FM and de la Cruz F: **TrwD, a protein encoded by the IncW plasmid R388, displays an ATP hydrolase activity essential for bacterial conjugation.** *J Biol Chem* 1997, **272**:25583-25590.
16. Sakai D, Horiuchi T and Komano T: **Atpase activity and multimer formation of PilQ protein are required for thin pilus biogenesis in plasmid R64.** *J Biol Chem* 2001, **276**:17968-17975.
17. Wolf YI, Rogozin IB, Kondrashov AS and Koonin EV: **Genome alignment, evolution of prokaryotic genome organization, and prediction of gene function using genomic context.** *Genome Res* 2001, **11**:356-372.
18. Wadsworth C: **A simple stained protein assay (SPA) for rapid total protein determination applied to spinal fluids.** *Scand J Immunol* 1976, **5**:1209-1213.
19. Laemmli UK: **Cleavage of structural proteins during the assembly of the head of bacteriophage T4.** *Nature* 1970, **227**:680-685.
20. Cleland JL and Wang DI: **Transient association of the first intermediate during the refolding of bovine carbonic anhydrase B.** *Biotechnol Prog* 1992, **8**:97-103.
21. Levi V and Gonzalez Flecha FL: **Reversible fast-dimerization of bovine serum albumin detected by fluorescence resonance energy transfer.** *Biochim Biophys Acta* 2002, **1599**:141-148.
22. Bartholmes P, Kirschner K and Gschwind HP: **Cooperative and noncooperative binding of pyridoxal 5'-phosphate to tryptophan synthase from Escherichia coli.** *Biochemistry* 1976, **15**:4712-4717.
23. Reichard K and Kaufmann M: **EPPS: mining the COG database by an extended phylogenetic patterns search.** *Bioinformatics* 2003, **19**:784-785.
24. Universität Witten/Herdecke: **EPPS** [<http://www.uni-wh.de/epps/>]
25. ClustalW [<http://www.ebi.ac.uk/clustalw/>]
26. Thompson JD, Higgins DG and Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** *Nucleic Acids Res* 1994, **22**:4673-4680.
27. BoxShade Server [http://www.ch.embnet.org/software/BOX_form.html]
28. NCBI BLAST Home Page [<http://www.ncbi.nlm.nih.gov/BLAST/>]

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

