# Cloning, sequencing and expression of VAT, a CDC48/p97 ATPase homologue from the archaeon *Thermoplasma acidophilum*

Vashu Pamnani<sup>a</sup>, Tomohiro Tamura<sup>a</sup>, Andrei Lupas<sup>a</sup>, Jürgen Peters<sup>a</sup>, Zdenka Cejka<sup>a</sup>, William Ashraf<sup>b</sup>, Wolfgang Baumeister<sup>a,\*</sup>

> <sup>a</sup>Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, Germany <sup>b</sup>Department of Biomedical Science, University of Bradford, Bradford, UK

> > Received 8 January 1997

Abstract A member of the AAA family of  $Mg^{2+}$ -ATPases from the archaeon *Thermoplasma acidophilum* has been cloned and expressed in *Escherichia coli*. The protein, VCP-like ATPase of *Thermoplasma acidophilum* (VAT), is a homologue of SAV from *Sulfolobus acidocaldarius* and CdcH of *Halobacterium salinarium*, and belongs to the CDC48/VCP/p97 subfamily. The deduced product of the *vat* gene is 745 residues long ( $M_r$  83 000), which has an optimal  $Mg^{2+}$ -ATPase activity at 70°C. Electron microscopy shows the purified protein to form single and double homo-hexameric rings. Although the symmetry is different, the appearance of the complexes formed of two rings resembles the 20S proteasome and Hsp60/GroEL.

© 1997 Federation of European Biochemical Societies.

Key words: AAA family; ATPase; CDC48; p97; Archaea; Thermoplasma acidophilum

## 1. Introduction

P-loop (or Walker A/B motif-containing [1]) nucleotide triphosphatases (NTPases) represent one of the largest protein superfamilies in nature, which includes ABC transporters [2], Ras and related GTPases [3], myosin [4], F<sub>1</sub>-ATPases [5], and nucleotide kinases [6]. Recently, a new family of P-loop NTPases has emerged, the AAA family (for ATPases Associated with a variety of cellular Activities) [7], some of whose best known members are the ATPase subunits of the 26S proteasome [8], FtsH [9], and NSF [10]. The hallmark of these proteins is a highly conserved domain of approximately 250 residues, which contains a second region of high sequence conservation C-terminal to the ATP-binding motifs. Members of this family (which may contain one or two ATPase domains) can be divided into groups on the basis of their sequence similarity, only one of which contains proteins with two fully conserved ATPase domains. The best characterized members of this group are CDC48 of Saccharomyces cerevisiae [11] and p97, also known as valosine-containing protein (VCP), of vertebrates [12]. Both proteins form homo-hexameric rings [13,14] and have an N-ethylmaleimide (NEM)-sensitive Mg<sup>2+</sup>-ATPase activity. They have recently been found to mediate homotypic membrane fusion during the cisternal regrowth of Golgi stacks [15] and of ER membranes [16]. Recently, CDC48 has also been found to interact with Ufd3p, a WD40-repeat protein required for ubiquitin-mediated proteolysis in *Saccharomyces cerevisiae* [17].

Archaeal homologues of CDC48/p97 have been sequenced in the vicinity of the reverse gyrase gene in *Sulfolobus acidocaldarius* [18] and of the sensory rhodopsin I gene in *Halobacterium salinarium* [19]. A further homologue has been identified by genomic sequencing in the methanogen, *Methanococcus jannaschii* [20]. However, none of the proteins have been isolated and their functions are unclear.

In this communication we describe the cloning, sequencing and expression of vat (VCP-like <u>A</u>TPase of <u>Thermoplasma</u>), a gene encoding a member of the CDC48/p97 subfamily from the archaeon *T. acidophilum*. Like CDC48 and p97, the purified VAT protein has Mg<sup>2+</sup>-dependent ATPase activity with a specific activity of 5.43 µmol/mg/h at an optimum temperature of approximately 70°C. Electron micrographs of the negatively stained protein show it to form hexameric rings, closely resembling the rings formed by p97 from *Xenopus laevis* [13] and CDC48 from *Saccharomyces cerevisiae* [14]. At higher concentrations, single rings tend to associate into double rings with an appearance similar to proteasomes and GroEL.

# 2. Materials and methods

## 2.1. DNA sequencing by PCR

*T. acidophilum* chromosomal DNA was purified by phenol-chloroform extraction. PCR was initially performed using the degenerate oligonucleotides that Schnall et al. [21] developed against the Walker A and B motifs of the conserved AAA module. After obtaining a 210 bp PCR product, primers were designed for sequencing the rest of the gene from a clone identified in a *T. acidophilum* genomic DNA library using the PCR cloning method [22,23]. Both strands of the gene were finally sequenced by the dideoxynucleotide method on an automated DNA sequencer (Applied Biosystems). The *vat* gene sequence was confirmed by direct PCR of chromosomal DNA using the same primers.

## 2.2. Cloning and expression of the vat gene

PCR was used to add an affinity tag coding for six histidine residues to the 3' end of the vat gene, as well as flanking EcoRI (3' end) and NdeI (5' end) restriction sites. The resulting construct was cloned into the NdeI and EcoRI sites of the expression vector pT7-7 to yield pT7-7-vat. Prior to the cloning step, the cleaved and purified vector was dephosphorylated with calf intestine alkaline phosphatase to ensure efficient sticky-end ligation. The plasmid was transformed into *E. coli* XL1 Blue cells and re-transformed into *E. coli* BL21 (DE3) cells carrying the plasmid pUBS520 [24]. The bacteria were grown in 2 liters of LB medium to an OD<sub>600</sub> of approximately 0.7 and optimally induced with 1 mM isopropyl- $\beta$ -D-thiogalactopyranoside for 8–10 h.

## 2.3. Purification of recombinant VAT

Lysis of bacterial cells was carried out with buffer A: 50 mM Na-

<sup>\*</sup>Corresponding author. Fax: (49) (89) 8578 2641. E-mail: pamnani@vms.biochem.mpg.de

*Abbreviations:* CDC48, cell division cycle protein 48; bp, base pair; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; RT-PCR, reverse transcription-PCR

phosphate, pH 8.0, 300 mM NaCl, 5 mM  $\beta$ -mercaptoethanol, 5 mM MgCl<sub>2</sub>, 1 mM ATP, containing 1 mg/ml lysozyme (Sigma, Deisenhofen, Germany). The lysate was fractionated by centrifugation (14000×g, 30 min) and the supernatant gently rotated with 2 ml of Ni-NTA resin (Diagen GmbH, Hilden, Germany) at 4°C for 1–2 h. The resin with the bound histidine-tagged VAT protein was packed into a column with an inner diameter of 6 mm and washed at 18 ml/h with buffer A, pH 8.0 and then with buffer A, pH 6.0 containing 10% glycerol, until the baseline was reached. VAT was finally eluted at room temperature with a 40 ml linear gradient of 0–300 mM imidazole in buffer A, pH 6.0. Fractions of 0.6 ml were collected and 10 µl were analyzed by SDS-PAGE. Imidazole fractions predominantly containing the recombinant VAT protein were pooled and dialysed for 24 h against 2 liters of 50 mM Tris-HCl, pH 8.0, 5 mM  $\beta$ -mer-

captoethanol, 100 mM NaCl, 5 mM MgCl<sub>2</sub>, 1 mM ATP at 4°C with three changes of buffer. Dialysed fractions were concentrated in Microcon-10 (Amicon, USA) and stored at 4°C in dialysis buffer. Protein concentration was determined by the Bradford method [25] using bovine serum albumin as a standard. Purity of the recombinant protein was monitored by SDS-PAGE (12.5% polyacrylamide) and oligomerization by PAGE (5% polyacrylamide) at room temperature according to Laemmli [26]. Gels were visualized by overnight Neuhoff staining [27].

### 2.4. Assay for ATPase activity

Duplicate reactions containing 1.5  $\mu$ g of VAT in a buffer containing 18 mM HEPES, pH 8.0, 10 mM  $\beta$ -mercaptoethanol, 5 mM MgCl<sub>2</sub>, 2 mM ATP were incubated for 30 min at various temperatures (45–

Ta VAT Hs CdcH Mj Cdc48 Sa SAV Sc Cdc48 Mm p97	1 1 1 1 1	A STORY A STORY AND	9 8 3 6
Ta VAT Hs CdcH Mj Cdc48 Sa SAV Sc Cdc48 Mm p97	33 30 29 64 57 48	KLNVEAGDYIEIIGQDGNS-ALAQVMPAYDISDDEIRIDGYIRKSIKVGIGDDVTVRKT-NVS   KLELFRGDTVLVKGKKRKD- <u>T</u> VLIVLIDDELEDGA <u>CRI</u> NRVVRNNLR <u>I</u> RLGDLVTIHPCPDIK	
Ta VAT Hs CdcH Mj Cdc48 Sa SAV Sc Cdc48 Mm p97	94 91 91 125 119 110	KADTLVLAPPEEÄSVQFGSDAAGMVKRQILKRPVVARDIVPVMSSTNHPFMRSPGQAIP 1 EAKKVVLAPTQPIRFGPGFEDFVKRKILGQVLSKGSKVTIGVLGTALT 1 PASKVVLAPTQPIRFDNSFVSVKDLMDKFLAKGETLPIPIYTGTLE 1 YATRISVLPIADTIEGIT-GNLFDVFLKPYFVEAYRPVRKGDHFVVRGGMRQVE 1	49 49 38 72 71 62
Ta VAT Hs CdcH Mj Cdc48 Sa SAV Sc Cdc48 Mm p97	150 150 139 173 172 163	LIAVETEPEGV-CLVTEDTDVELREEPISGFERTGGGITYEDIGGLENÊIQRVREMVELPM 2 FVVVSTTPAGP-VRVTDFTHVELKEEPVSEIKETK-VPDVTYEDIGGLKEEVKKVREMIELPM 1 LTVVNTQPSNY-VYVTGSTNIEIREEVKESSLAYPKVSWEDIGDLEEAKQKIREIVEWPM 2 FKVVD <u>V</u> EPEEY-AVVAQDIIF	10 99 32 33 24
Ta VAT Hs CdcH Mj Cdc48 Sa SAV Sc Cdc48 Mm p97	211 210 200 233 234 225	KHPQIFQKLGIEPPQGVLLHGPPGTGKTLLAKAVANETSASFFSIAGPEIISKYYGESEQQLR 2 RHPELFEKLGIEPPKGVLLVGPPGTGKTLLAKAVANEAGANFYVINGPEINSKYVGETEENLR 2 RHPELFQRLGIDPPKGVLLVGPPGTGKTLLARALRNEIGAYFITVNGPEINSKFYGESEQRIR 2 RHPQLFKAIGIKPPRGVLMYGPPGTGKTLMARAVANETGAFFFLINGPEVMSKMAGESESNLR 2 RHPALFKAIGVKPPRGTLLYGPPGTGKTLIARAVANETGAFFFLINGPEIMSKLAGESESNLR 2	273 272 262 295 296 287
Ta VAT Hs CdcH Mj Cdc48 Sa SAV Sc Cdc48 Mm p97	274 273 263 296 297 288	EIFEDAKDDSPSIIFIDELDSIAPKREDVTGEVERRVVAQLLTMMDGLEGRGQVIVIAATNRV 3 KIFEEAEENAPSIIFIDEIDAIAPKRDEATGEVERRLVAQLLTLMDGLKGRGQVVVIGATNRP 3 EIFKEAEENAPSIIFIDEIDAIAPKREDVTGEVERRVVAQLLTLMDGTKGRGRVIVIGATNRP 3 KAFEEAEKNAPAIIFIDEIDSIAPKRDKTNGEVERRVVSQLLTLMDGKARSNVVVIAATNRP 3 KAFEEAEKNAPAIIFIDELDAIAPKREKTHGEVERRVVSQLLTLMDGKQRAHVIVMAATNRP 3 Walker B	36 35 25 58 59 50
Ta VAT Hs CdcH Mj Cdc48 Sa SAV Sc Cdc48 Mm p97	337 336 326 359 360 351	α4 β5 α5   DAIDPALRRPGRFDREIEIGVPDRNGRKEILMIHTRNMPLGMSEEEKNKFLEEMADYTYGFVG 3   DAVDPALRRPGRFDREIEIGVPDEIGREEILKIHTRGMPLSDDVNLSTLADDTHGFVG 3   NALDPALRRPGRFDREIVIGVPDREGRKEILQIHTRNMPLSDDVDLDYLADVTHGFVG 3   DAIDPALRRPGRFDREIVIGVPDREGRKEILQIHTRNMPLAEDVDLDYLADVTHGFVG 3   DAIDPALRRFGRFDREIVIGVPDREGRKEILQIHTRNMPLAEDVDLDYLADVTHGFVG 3   DAIDPALRRFGRFDREIVIGVPQREGRKEILQVHTRNMPLAEDVDLDYLADVTHGFVG 4   NSIDPALRRFGRFDREVVIG 4	999 193 183 16 17 08
Ta VAT Hs CdcH Mj Cdc48 Sa SAV Se Cdc48 Mm p97	400 394 384 417 418 409	DomI ADLAALVRESAMNALRRYLPE IDLDKP - IPTEILEKMVVTEDDFKNALKSIEPSSLREVMV 4 ADIESLSKEAAMRALRRYLPE - IDLDEEDIPPSLIDRMIVKREDFKGALSEVEPSAMREVLV 4 ADLAALCKEAAMRALRRVLPS - IDLEAEEIPKEVLDNLKVTMDDFKEALKDVEPSAMREVLV 4 ADLAALAKEAAIYALRFVDEKKLNLDQPTIPAEIIKELKVSMNDFLNALKSIQPSLLREVVV 4 ADTASLCSEAAMQOIREKMDL - IDLDEDEIDAEVLDSLGVTMDNFFFALGNSNPSALRETVV 4	59 54 44 79 78 69

Fig. 1. Sequence alignment of VAT (GenBank accession number U78072) with other CDC48/p97 homologues. Residues conserved in a majority of sequences are shown in reverse type. *Dom I* and *II* denote the two ATPase domains. Walker A and B motifs and the 'second region of homology' as defined in [9] are marked. Above the alignment is given the consensus secondary structure for the two ATPase domains (open bars denote  $\alpha$ -helices and arrows  $\beta$ -strands). The five  $\beta$  strands and five  $\alpha$  helices of the mononucleotide-binding fold are numbered.

		$\beta 1$ $\alpha 1$	
Ta VAT Hs CdcH Mj Cdc48 Sa SAV Sc Cdc48 Mm p97	460 455 445 480 479 470	VPNVHWDDIGGLEDVKREIKETVELPLLKPDVFKRLGIRPSKGELLYGPPGVGKTLLAKAVA 522 LPKITWDDVGGLTEAKNNVKESVEWPLNQPEKFTRMGVEPPAGVLLYGPPGTGKTLMAKAVA 517 VPNVKWEDIGGLEEVKQELREAVEWPLKAKEVFEKIGVRPPKGVLLFGPPGTGKTLLAKAVA 507 VPKVNWNDIGGLDNVKQQLREAVEWPLKFFELFTKSGVTPPKGLLEGPPGTGKTLLAKAVA SVNVTWDDVGGLDEIKEELKETVEYPVLHPDQYTKFGLSPSKGVLFYGPPGTGKTLLAKAVA 541 VPQVTWEDIGGLEDVKRELQELVQYPVEHPDKFLKFGMTPSKGVLFYGPPGCGKTLLAKATA 532 Walker A	, , ,
Ta VAT Hs CdcH Mj Cdc48 Sa SAV Sc Cdc48 Mm p97	523 518 508 543 542 533	β2 ESNANFISIKGPEVLSKWVGESEKAIREIFKKAKQVAPAIVFLDEIDSIAPRRGTTSDSG ETNANFISVRGPOLLSKWVGESEKAIROIFRKARQVAPTVIFFDELDSIAPGRGQTGGNN ESGANFISVKGPEIFSKWVGESEKAIREIFRKARQSAPCIIFFDEIDAIAPKRGRDLSSA ESGANFISVKGPEIFSKWVGESEKAIREIFRKARQSAPCVIFFDEIDSIAPIRGLSDSG EVSANFISVKGPELLSMWVGESEKAIREIFDKARQAAPTVVFLDELDSIAKARGGSLGDAGG 604 ECQANFISTKGPELLSMWVGESEANVREIFDKARQAAPCVLFFDELDSIAKARGGNIGDGGG 595	3 3 3 4
		$\frac{\alpha 3}{\beta 4} \qquad \frac{\beta 4}{\alpha 4} \qquad \frac{\psi \text{alker B}}{\beta 5} \qquad \frac{\alpha 5}{\alpha 5}$	
Ta VAT Hs CdcH Mj Cdc48 Sa SAV Sc Cdc48 Mm p97	584 579 569 604 605 596	TERIVNQLLTSLDGIEVMNGVVVIGATNRPDIMDPALLRAGRFDKLIYIPPPDKEARLSILK SERVVNQLLTELDGLEEMEEVMVIAATNRPDIIDPALIRSGRFDRLVQVGQPGIEGREQILK TDKVVNQLLTELDGMEEPKDVVVIAATNRPDIIDPALLRPGRLDRVILVPVPDEKARLDIFK SERVVNQLLTELDGMEEPKDVVVIAATNRPDIIDPALLRPGRFDRLIVVPPDKTARFEILK SCRVVNQLLTEMDGMNAKKNVFVIGATNRPDQIDPAILRPGRLDQLIVVPLPDENARLSILN 666 ADRVINQILTEMDGMNSTKKNVFIIGATNRPDQIDPAILRPGRLDQLIVVPLPDENARLSILN (558	l 1 5 7
TaVAT	647	HTKNMPLAPDVDLNDIAQRTEGYVGADLENLCREAGMNAYRENPDATS 695	5
Hs CdcH Mj Cdc48 Sa SAV Sc Cdc48 Mm p97	642 632 667 668 659	HTQDTPLAADVSLRELAERADGYVGSDLANIAREAAIEALRDDEDADD 690 HTRSMNLAEDVNLEELAKKTEGYTGADIEALCREAAMLAVRESIGKPWDIEVKLRELINYLQ 694 HTKNVPLAEDVSLEDIAEKAEGYTGADLEALVREATINAMRSIYSMCDKQSRDECKGNMECY 723 QLRKTPLEPGLELTAIAKATQGFSGADLUVRAAKYAIKDSIEAHRQHEAEKEVKVEGED 730 NLRKS <b>EVAKDVDL</b> EFLAKMTNGFSGADLTEICQRACKLAIRESIESEIRRERERQTNPSAME 721	4 )
Ta VAT Hs CdcH Mj Cdc48 Sa SAV Sc Cdc48 Mm p97	696 691 695 730 731 722	693 5 I S G T F R A A A V E L N S V I K A T K E R E S A E A G E F S E L K NA I G K I I S V L S P A K E K I E A V E K E I D K F L 757 743 751 751 751 752 754 755 757 757 757 757 757 757	) 7 3
Ta VAT Hs CdcH Mj Cdc48	696 691 758	695 GVINKEELKPSEKDEAOKLAKYLKDILGKLKEMIDNIYELENKLNTLKEQVSAEEIDEIIKTT 820	)
Sa SAV Sc Cdc48	744 752	741 742 751	3 1
Mm p97	731		,
TaVAT	696	DomII - DOMII	
Hs CdcH Mj Cdc48 Sa SAV	691 821 744	NIIQRFTTSLDELKNILKDIESIRLKVSTKDVKIKKEHFMKALEKIKPSVSKEDMRVYEKLA 883	3 2
Sc Cdc48 Mm p97	752 731	ITKEHFAEAMKTAKRSVSDAELRRYEAYS 780 ITKEHFAEAMKTAKRSVSDAELRRYEAYS 780 ITKEHFAEAMRFARRSVSDAELRRYEAYS 780	
Τα VAT	725	TTMSKSVSERRKQLQDQGLYL 74	5
Hs CdcH Mj Cdc48	720 884	DQFKG <mark>S</mark> QGPNVNSRQGSEHIGFQ 742 DEYGRATSVEKKKEEGKEVI 902	2 3
Sa SAV Sc Cdc48 Mm p97	773 781 760	XELKRAIA   780     QQMKASRGQFSNFNFNDAPLGTTATDNANSNNSAPSGAGAAFGSNAEEDDDLYS   834     QTLQQSRG-FGSFRFPSGNQGGAGPSQGSGGGTGGSVYTEDNDDDLYG   800	4

Fig. 1 (continued).

80°C). Negative control reactions in the absence of the enzyme, also in duplicate, were performed in parallel. ATPase activity was assayed by measuring the release of phosphate from ATP, by the method of Lanzetta et al. [28].

#### 2.5. Electron microscopy and image processing

The expressed protein was negatively stained with 1% (w/v) uranyl acetate (pH 4.1) and viewed with a Philips EM 420 electron microscope at a magnification of  $\times 30\,000$ . The images were recorded under low-dose conditions. Suitable areas (2048×2048 pixels) of the micrographs were digitized with an Eikonix microdensitometer. The step size was 15 µm, which corresponds to a sampling of 0.5 nm per pixel at the specimen level. A total of 3350 images of the VAT particles were extracted interactively. The extracted images were aligned (cor-

# inter-image variations. 2.6. Sequence analysis

Homologues of sequenced open reading frames were identified in GenBank using the BLAST E-mail server at NCBI (blast@ncbi.nlm.nih.gov). Alignments were made in the PileUp program of the GCG package (Genetics Computer Group, Madison, WI, USA) and refined using MACAW [30]. The consensus secondary structure prediction was obtained for a complete alignment of AAA ATPase domains by the methods of Rost and Sander [31] and of Mehta et al. [32].

rected for the translation and rotation of the molecules) using cross-

correlation techniques. The aligned data set was subjected to an eigenvector-eigenvalue classification procedure [29] to detect significant

## 3. Results and discussion

# 3.1. Sequencing and cloning of the vat gene

We identified the gene encoding the CDC48/p97 homologue of T. acidophilum by PCR using degenerate oligonucleotides targeted to the conserved Walker A and B sequences that had been previously used to identify AAA ATPases in Saccharomyces cerevisiae [22]. A PCR fragment of approximately 210 bp was obtained, which gave a partial ORF with over 70% identity to Halobacterium salinarium CdcH and Sulfolobus acidocaldarius SAV. Following the sequencing of this fragment, the complete nucleotide sequence of the gene and of its 5' and 3' flanking regions was determined from the isolated clone (see Section 2). The sequence was also confirmed directly from genomic DNA using PCR. Two possible initiation codons but no putative ribosomal binding sites were detected in the sequence. The initiation codon giving the longer reading frame was identified as the correct one by N-terminal sequencing of the gene product. The vat gene codes for a polypeptide of 745 amino acid residues (Fig. 1) with a calculated molecular weight of 83050 and a pI of 5.4.

# 3.2. Sequence analysis of VAT

VAT can be aligned significantly over its entire length with proteins of the CDC48/p97 group. It is about 60% identical to SAV, CdcH, and the CDC48-homologue of *Methanococcus jannaschii*, and about 50% identical to CDC48 and p97. It contains two AAA ATPase domains from Ile<sup>187</sup> to Val<sup>459</sup> and from Val<sup>464</sup> to Lys<sup>718</sup> (Fig. 1). Sequence similarity to other proteins in the group is particularly high for the first of the two domains.

Both domains contain the Walker A and B motifs (Fig. 1) characteristic of the mononucleotide-binding fold. This fold consists of a doubly wound, five-stranded, parallel β-sheet and is one of the most common known folds [33], being found in mononucleotide-binding proteins (P-loop NTPases), dinucleotide-binding proteins (reductases and dehydrogenases), periplasmic binding proteins and bacterial two-component-system regulators. Invariably, the binding occurs at the C-terminal end of the  $\beta$ -sheet in a cleft formed by loops connecting the central strands  $\beta$ 1 and  $\beta$ 3 to helices on opposite sides of the sheet. In the mononucleotide-binding fold, the Walker A motif (G--G-GKT) corresponds to the loop between  $\beta$ 1 and the following  $\alpha$ -helix, with the glycines forming an anion hole, the lysine coordinating the  $\beta$  and  $\gamma$  phosphates of the nucleotide, and the threonine binding the  $Mg^{2+}$  ion of  $Mg^{2+}$ -NTP. The Walker B motif (four hydrophobic residues followed by aspartate) corresponds to  $\beta$ 3; the aspartate of this motif forms a salt bridge with the lysine of Walker A and provides a second coordination site for the Mg<sup>2+</sup> ion. A third coordination site is generally provided by a polar residue at the C-terminal end of  $\beta 2$ .

A secondary structure prediction for the AAA domains of VAT (Fig. 1) was generated by consensus from a complete alignment of AAA domains in current databases (A. Lupas, unpublished). Because more than 70 sequences spanning a wide range of identities were included in the alignment and because  $\alpha/\beta$  folds are more accurately predicted than other folds [34,35], the expected accuracy for the prediction is above 80% per residue. In addition, the prediction was validated by the known location of the Walker A and B motifs in the fold. The prediction indicates that in the AAA domain, the topol-

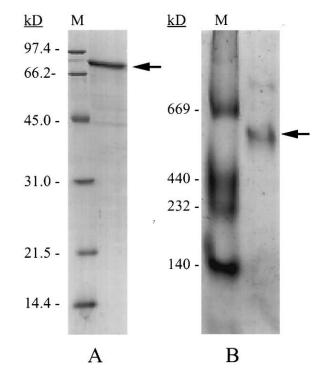


Fig. 2. Electrophoretic analysis of the purified recombinant VAT. A: 12.5% SDS-PAGE stained with colloidal Coomassie blue G [27]: M represents 1.5  $\mu$ g low molecular weight standards – phosphorylase b (97400), serum albumin (66200), ovalbumin (45000), carboanhydrase (31000), soy bean inhibitor (21500). VAT migrates at a molecular weight of approximately 83000. B: 5% Non-denaturing PAGE stained as above. M represents high molecular mass standards – thyroglobulin (669000), ferritin (440000), catalase (232000), lactate dehydrogenase (140000). VAT complex has an apparent molecular weight of approximately 570000. Arrows indicate position of VAT.

ogy of the central β-sheet is the canonical one of P-loop NTPases, rather than the one seen in RecA and  $F_1F_0$ -ATPase, as proposed by Yoshida and Amano [36]. In addition to the residues of the Walker A and B motifs, further functional residues of VAT located by this prediction are  $Glu^{259/536}$ , which is proposed to provide the third coordination site for the  $Mg^{2+}$  ion in the two domains, and  $Glu^{291/568}$  and  $Asp^{293/570}$ , which form a sequence motif related to the DEAD box [37] and which are proposed to initiate the inline attack on the  $\gamma$  phosphate by analogy to Gln61 in p21<sup>ras</sup> [38].

An unusual feature of VAT is a short insertional sequence, GMSEEEKNKF (residues 377–386), which replaces the sequence xDDVD conserved in the other CDC48/p97 proteins (Fig. 1). In order to show that the insertion is present on the messenger RNA level, first strand cDNAs were transcribed from total *T. acidophilum* RNA using random hexamer primers (data not shown). RT-PCR subsequently showed the insertion to be present. Its predicted location in a solventexposed loop at the opposite end from the nucleotide-binding pocket indicates that it may not be central to the function of the molecule.

A sequence matching a consensus nuclear localization signal (K-K/R/-x-K/R) [39] occurs between Lys<sup>554</sup> and Lys<sup>557</sup>. Although the function of such sequences remains puzzling in an organism lacking a nucleus, a similar sequence, functional in a eukaryotic system, has been identified in the 20S proteasome of *T. acidophilum* [40]. It is conceivable that elements of

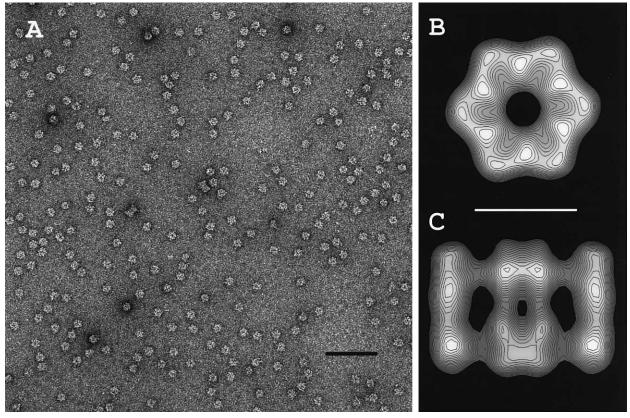


Fig. 3. Electron micrograph of negatively stained recombinant VAT protein (showing mainly top views of single particles) (A). Image averages of: top view of a single VAT particle (B) and side view of VAT dimers (C). VAT forms hexameric ring-shaped structures that may dimerize to form structures that resemble the proteasome. Scale bars represent 100 nm (in A) and 10 nm (in B and C).

a nuclear transport machinery (albeit fulfilling a different role) are present in *T. acidophilum*. A further noteworthy feature of the VAT sequence is a tyrosine at the next-to-last position, which is also found in all eukaryotic CDC48/p97 proteins, but not in other archaeal homologues (Fig. 1). In mammals, this tyrosine is one of the first cellular targets phosphorylated in response to T cell receptor activation [41] and is constitutively phosphorylated in Rous sarcoma virus-transformed fibroblasts [42].

# 3.3. Expression of the vat gene and purification of VAT

The vat gene was expressed using the T7 polymerase expression system in conjunction with another plasmid (pUBS520) encoding the *dnaY* gene for the rare arginine codons AGG and AGA. Since 48 out of the 57 arginine residues in VAT are encoded by rare codons, the adoption of the double expression system was carried out to enhance the otherwise very poor expression level of recombinant VAT. The gene from the isolated expression vector pT7-7vat was re-sequenced to confirm its integrity.

VAT was expressed with a fused  $(His)_6$  tag and purified via nickel-affinity chromatography. Expressed protein was eluted at approximately 150 mM imidazole and identified by N-terminal sequencing after application on a Laemmli SDS-polyacrylamide gel (Fig. 2A); its apparent molecular weight was estimated to 83 000. In native gel electrophoresis (Fig. 2B), the protein had an apparent molecular weight of 570 000 and on a Superose 6 gel filtration column (data not shown) an apparent molecular weight of approximately 530 000 suggesting, in both cases, an oligomeric state. These values indicate that under these conditions, VAT forms homo-hexameric complexes (predicted molecular weight: 498 000).

# 3.4. ATPase activity of VAT

Biochemical characterization studies showed that VAT possesses Mg<sup>2+</sup>-dependent ATPase activity with a specific activity of 5.43 µmol/mg/h in the presence of 18 mM HEPES, pH 8.0, 2 mM ATP, 5 mM Mg<sup>2+</sup> at an optimum temperature of ~70°C (Table 1), which is similar to the specific activity of p97 (~5 µmol/mg/h) [12] measured at 27°C. The optimum temperature of VAT was approximately 10°C above the growth temperature of *T. acidophilum* (~58–60°C).

# 3.5. Electron microscopy and image analysis of VAT

In electron micrographs (Fig. 3A), VAT was seen to form

Table 1 Special ATPase activity of recombinant VAT at variable temperatures

Temperature (°C)	Specific activity (µmol/mg/h)	
45	2.13	
45 55	3.60	
60	4.31	
65	5.22	
70	5.43	
80	4.30	

For experimental conditions, see Section 2. Values for specific activity were corrected for background hydrolysis of ATP. Highlighted values denote specific activity at optimum temperature. hexameric ring-shaped structures with a diameter of 15.5 nm, indistinguishable from those formed by *Xenopus* p97 [13] and CDC48 [14]. According to eigenvector-eigenvalue analysis, there is no indication that other than a six-fold symmetry exists in the VAT molecule. We also observed several particles that resembled the side views of the 20S proteasome when VAT was visualized at high protein concentrations. This observation has also been reported for p97 [13] and CDC48 [14]. It was proposed that these cylindrical particles represent side views of complexes formed by two stacked rings. We have no information whether or not the dodecameric complexes also occur in vivo.

Acknowledgements: The authors wish to thank R. Schnall (University of Munich) for initially providing the oligonucleotides against the Walker A and B motifs, R. Mattes from the University of Stuttgart for the kind gift of the pUBS520 plasmid encoding the *dnaY* gene for the expression of the *vat* gene, F. Lottspeich for N-terminal sequencing of the VAT protein and M. Boicu for DNA sequencing.

## References

- Walker, J.E., Saraste, M.J., Runswick, J.J. and Gay, N.J. (1982) EMBO J. 1, 945–951.
- [2] Hyde, S.C., Emsley, P., Hartshorn, M., Mimmack, M.M., Gileadi, U., Pearce, S.R., Gallagher, M.P., Gill, D.R., Hubbard, R.E. and Higgins, C.F. (1990) Nature 364, 362–365.
- [3] Bourne, H.R., Sanders, D.A. and McCormick, F. (1991) Nature 349, 117–127.
- [4] Rayment, I., Rypniewski, W.R., Schmidt-Bäse, K., Smith, R., Tomchick, D.R., Benning, M.M., Winkelmann, D.A., Wesenberg, G. and Holden, H.M. (1993) Science 261, 50–58.
- [5] Abrahams, J.P., Leslie, A.G.W., Lutter, R. and Walker, J.E. (1994) Nature 370, 621–628.
- [6] Schulz, G.E. (1992) Curr. Opin. Struct. Biol. 2, 61-67.
- [7] Erdmann, R., Wiebel, F.F., Flessau, A., Fytka, J., Beyer, A., Fröhlich, K.-U. and Kunau, W.-H. (1991) Cell 64, 499–510.
- [8] Fujiwara, T., Watanabe, T.K., Tanaka, K., Slaughter, C.A. and DeMartino, G.N. (1996) FEBS Lett. 387, 184–188.
- [9] Tomoyasu, T., Yuki, T., Morimura, S., Mori, H., Yamanaka, K., Niki, H., Hiraga, S. and Ogura, T (1993) J. Bacteriol. 175, 1344– 1351.
- [10] Wilson, D.W., Wilcox, C.A., Flynn, G.C., Chen, E., Kuang, W.-J., Henzel, W.J., Block, M.R., Ullrich, A. and Rothman, J.E. (1989) Nature 339, 355–359.
- [11] Fröhlich, K.U., Fries, H.-W., Rüdiger, M., Erdmann, R., Botstein, D. and Mecke, D. (1991) J. Cell Biol. 114, 443–453.
- [12] Peters, J.M., Walsh, M.J. and Franke, W.W. (1990) EMBO J. 9, 1757–1767.
- [13] Peters, J.M., Harris, J.R., Lustig, A., Müller, S., Engel, A., Volker, S. and Franke, W.W. (1992) J. Mol. Biol. 223, 557–571.
- [14] Fröhlich, K.U., Fries, H.-W., Peters, J.M. and Mecke, D. (1995) Biochim. Biophys. Acta 1253, 25–32.
- [15] Rabouille, C., Levine, T.P., Peters, J.-M. and Warren, G (1995) Cell 82, 905–914.

- [16] Latterich, M., Fröhlich, K.-U. and Schekman, R (1995) Cell 82, 885–893.
- [17] Ghislain, M., Dohmen, R.-J., Levy, F. and Varshavsky, A. (1996) EMBO J. 15, 4884–4899
- [18] Confalonieri, F., Marsault, J. and Duguet, M. (1994) J. Mol. Biol. 235, 396–401.
- [19] Bibikov, S. and Oesterhelt, D. (1994) Halobacterium salinarium. 6th International Conference on Retinal Proteins, Leiden, p. 67.
- [20] Bult, C.J., White, O., Olsen, G.J., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, J.A., Fitzgerald, L.M., Clayton, R.A., Gocayne, J.D., Kerlavage, A.R., Dougherty, B.A., Tomb, J.F., Adams, M.D., Reich, C.I., Overbeek, R., Kirkness, E.F., Weinstock, K.G., Merrick, J.M., Glodek, A., Scott, J.L., Geoghagen, N.S.M., Weidman, J.F., Fuhrmann, J.L., Nguyen, D., Utterback, T.R., Kelley, J.M., Peterson, J.D., Sadow, P.W., Hanna, M.C., Cotton, M.D., Roberts, K.M., Hurst, M.A., Kaine, B.P., Borodoovsky, M., Klenk, H-P., Fraser, C.M., Smith, H.O., Woese, C.R. and Venter, J.C. (1996) Science 273, 1058–1073.
- [21] Schnall, R., Mannhaupt, G., Stucka, R., Tauer, R., Ehnle, S., Schwarzlose, C., Vetter, I. and Feldmann, H. (1994) Yeast 10, 1141–1155.
- [22] Tamura, T., Tamura, N., Cejka, Z., Hegerl, R., Lottspeich, F. and Baumeister, W. (1996) Science 274, 1385–1389.
- [23] Alfandari, D. and Darribere, T. (1994) PCR Methods Appl. 4, 46–49.
- [24] Brinkmann, U., Mattes, R.E. and Buckel, P. (1989) Gene 85, 109-114.
- [25] Bradford, M.M. (1976) Anal. Biochem. 72, 248-254.
- [26] Laemmli, U.K (1970) Nature 227, 680-685.
- [27] Neuhoff, V., Arold, N., Taube, D. and Ehrhardt, W. (1988) Electrophoresis 9, 255–262.
- [28] Lanzetta, P.A., Alvarez, L.J., Reinach, P.S. and Candia, O.A. (1979) Anal. Biochem. 100, 95–97.
- [29] Van Heel, M. and Frank, J. (1981) Ultramicroscopy 6, 187-194.
- [30] Schuler, G.D., Altschul, S.F. and Lipman, D.J. (1991) Proteins 9, 180–190.
- [31] Rost, B. and Sander, C. (1993) J. Mol. Biol. 232, 584-599.
- [32] Mehta, P.K., Heringa, J. and Argos, P. (1995) Protein Sci. 4, 2517–2525.
- [33] Orengo, C.A., Flores, T.P., Taylor, W.R. and Thornton, J.M. (1993) Protein Eng. 6, 485–500.
- [34] Benner, S.A. (1992) Curr. Opin. Struct. Biol. 2, 402–412.
- [35] Niermann, T. and Krischner, K. (1991) Methods Enzymol. 202, 45–59.
- [36] Yoshida, M. and Amano, T. (1995) FEBS Lett. 359, 1-5.
- [37] Linder, P., Lasko, P.F., Leroy, P., Nielsen, P.J., Nishi, K., Schnier, J. and Slonimski, P.P. (1989) Nature 337, 121–122.
- [38] Pai, E.F., Krengel, U., Petsko, G.A., Goody, R.S., Kabsch, W. and Wittinghofer, A (1990) EMBO J. 9, 2351–2359.
- [39] Roberts, B. (1989) Biochim. Biophys. Acta 1008, 263-280.
- [40] Nederlof, P.M., Wang, H.R. and Baumeister, W. (1995) Proc. Natl. Acad. Sci. USA 92, 12060–12064.
- [41] Egerton, M. and Samuelson, L.E. (1994) J. Biol. Chem. 269, 11435–11441.
- [42] Schulte, R.J., Campbell, M.A., Fischer, W.H. and Sefton, B.M. (1994) J. Immunol. 153, 5465–5472.