

## Short sequence-paper

Characterisation of *vha26*, the *Drosophila* gene for a 26 kDa E-subunit of the vacuolar ATPase<sup>1</sup>Yiquan Guo, Zongsheng Wang, Andrew Carter, Kim Kaiser, Julian A.T. Dow<sup>\*</sup>

Division of Molecular Genetics, Institute of Biomedical and Life Sciences, University of Glasgow, Glasgow G11 6NU, UK

Received 13 March 1996; revised 10 May 1996; accepted 15 May 1996

---

Abstract

A *Drosophila melanogaster* gene and cDNA for the E-subunit of the V-ATPase were characterised. The deduced product has 226 amino acids and a molecular mass of 26.1 kDa. The gene is a single copy at 83B1-4 on chromosome 3R. The coding sequence is punctuated by three introns which do not align with those in *Neurospora*. The gene is ubiquitously expressed as an mRNA of 2.3 kb, but at lower levels in pupae.

**Keywords:** V-ATPase; E-subunit; Ion transport; Proton pump; Diptera; (*D. melanogaster*)

---

The vacuolar H<sup>+</sup>-ATPase (V-ATPase) is a ubiquitous proton pump which was originally characterised as an endomembrane component that pumps protons from the cytoplasm to the internal space of organelles. It also plays an important role in energising plasma membrane ion transport [1]. V-ATPases contain transmembrane and catalytic sectors, and are structurally related to the F-ATPases of bacteria, mitochondria and chloroplasts [2], and to a Na<sup>+</sup>-ATPase found in some prokaryotes [3]. The transmembrane sector, or protonophore (V<sub>o</sub>), contains 6 copies of a 16 kDa proteolipid subunit. This subunit is linked by other subunits to a headgroup (V<sub>i</sub>) made of three copies each of A and B subunits, which together form the catalytic core. Native gels of *M. sexta* goblet cell plasma membranes suggest a holoenzyme of 600–900 kDa, while denaturing SDS-PAGE gels reveal polypeptide subunits of 67, 56, 43, 40, 28, 20, 17, 16 and 14 kDa [4,5]. A similarly constructed V-ATPase has been isolated from a variety of sources, including plants, fungi, and mammals, and several corresponding cDNAs or genes have been cloned and characterised.

Recently, the 26–28 kDa E-subunit has been cloned from a number of phyla. It has been suggested that E-subunit may play an analogous role in the vacuolar ATPase to the  $\gamma$ -subunit in F-type ATPases [6], and as such should be considered to form part of the catalytic headgroup, although its precise function is not clear. The corresponding yeast gene, *vma4*, has been mutagenised [7], and showed a pH-sensitive lethal phenotype similar to other V-ATPase disruptant mutants. In mutants, the proteolipid inserted normally into the membrane, whereas the subunits of the catalytic sector did not assemble [8]. In vertebrate kidney, it has been suggested that E-subunits co-localise immunocytochemically with plasma membranes, rather than microsomes [9], implying that E-subunits may be important in assembly of the holoenzyme on the plasma membrane of certain epithelia. As a first step in clarifying this issue, we have characterised both the cDNA and gene of this subunit in *Drosophila melanogaster*, a species which is particularly suited to genetic analysis [5].

Two genomic and five cDNA clones (Figs. 1 and 2) were identified by homology with the *Manduca sexta* gene [10] using a cDNA clone kindly provided by H. Wiczorek. The 2.1 kb cDNA has an open reading frame corresponding to a 226 amino-acid polypeptide of *M*<sub>r</sub> 26 100 (Fig. 2). This cDNA clearly encodes a V-ATPase E-subunit, sharing 77% amino-acid identity with the E-subunit of *M. sexta* (insect), 63% identity with that of human but only 35% identity with that of yeast (Fig. 3). In

---

<sup>\*</sup> Corresponding author. Fax: +44 141 3305994; e-mail: j.a.t.dow@bio.gla.ac.uk.

<sup>1</sup> The sequence data reported in this paper have been submitted to the GenBank Data Libraries under the accession numbers U38198 (cDNA) and U38951 (genomic).

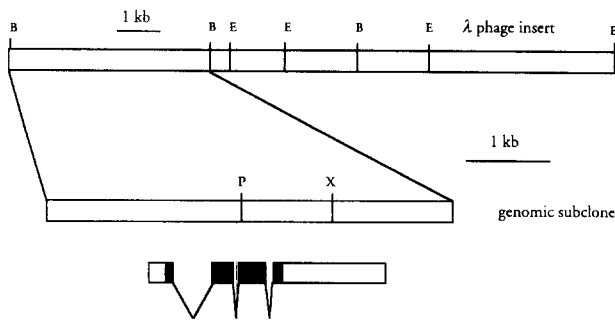


Fig. 1. Structure of the *D. melanogaster* V-ATPase E-subunit gene, *vha26*. A  $\lambda$ ZapII (Stratagene) oligo-dT primed cDNA library representing adult heads of *eya*– mutant *D. melanogaster* (S.R.H. Russell, unpublished) was screened by plaque hybridisation with a digoxigenin-random-primed probe of a cloned cDNA for the *M. sexta* E-subunit [21]. Positives were purified by a further round of plating. The resulting incomplete cDNA clone was then used to isolate a further 4 cDNA clones from a commercial (Promega)  $\lambda$ GEM-2 library of *D. melanogaster* head. (A) *D. melanogaster* genomic DNA library in the EMBL3 vector was similarly screened to identify two genomic clones, each containing 12 kb inserts, were isolated and restriction-mapped with E, *EcoRI*; (B) *BamHI*; X, *XbaI*; P, *PstI*. The 4.5 kb *BamHI* genomic fragment contained all the cDNA sequence and was subcloned in pBluescript SK<sup>–</sup>.

accordance with the nomenclature for other *D. melanogaster* V-ATPase loci, the gene has been named *vha26*. Although we cannot at present exclude the possibility that longer transcripts exist, the longest 5' UTR of the 5 cDNA clones is 77 bp. This length is in good agreement with those of 5' UTRs reported for other V-ATPase subunits in *Drosophila*, at 93 bp for the 67 kDa A subunit (Guo, Y. et al., unpublished results), 86 bp for the 55 kDa B subunit (Davies, S.A. et al., unpublished results), 116 bp for the 17 kDa c subunit [11], and 42 bp for the 14 kDa F subunit [12]. The sequence of the start site CAAAATG matches the consensus start site (C/A)AA(A/C)ATG perfectly [13]. The 3' UTR is 1307 bp long, with a canonical poly(A) AATAAA signal centred 26 bp upstream of the poly(A) tail.

The genomic DNA clone *dro26kg* contains the 2.1 kb cDNA sequence, punctuated by three small introns with in-frame boundaries (Fig. 1). This is the first description of a genomic sequence, and thus of intron placement in E subunit genes in animals. Intron placement frequently marks functional boundaries within proteins; however, the only other genomic sequence available, for *Neurospora crassa* [6], shows that, although the first introns of each sequence differ by only 4 residues in their position, precise intron placement is not conserved between animals and fungi; however, as further genomic sequences are obtained, they may be informative. As with the *N. crassa*

gene encoding *vma4*, no TATA or CAAT boxes could be seen upstream of the putative transcriptional start site in the available sequence for *vha26*; this is commonly the case for generally expressed genes.

The recent availability of deduced sequences from a broad range of phyla allows new insights into the protein structure. Although the primary sequence is poorly conserved across phyla, the substitutions are generally conservative, even in the distantly related archaeobacterial homologues (Fig. 3). Similarly, the predicted secondary structure is conserved; all members of the family appear to encode predominantly hydrophilic  $\alpha$ -helical proteins with conserved N- and C-termini, as noted previously [6]. However, despite the large evolutionary distance between insects and mammals, there is far closer conservation amongst these sequences than within those for either plants, fungi or prokaryotes (Fig. 3). Indeed, this dichotomy between animal and other phyla is greater than we have observed in the other V-ATPase subunits we have studied [12,14] (Davies, S.A. et al., unpublished results), suggesting that the E-subunit may have a distinctive role in animals (possibly for holoenzyme targeting to, or assembly in, plasma membranes of epithelia), which requires the absolute conservation of regions of primary sequence, and does not permit conservative substitution. For example, an extended 24-a.a. N-terminal motif ADVQKQIKHMAFIEQEANEKAEE is absolutely conserved in all animal sequences known across a 400 M year evolutionary span, but in the corresponding interval only 20 residues are conserved among plants, 16 among fungi and none among prokaryotes (Fig. 3). Further in the sequence, the motifs QRLKIMEYYEKKEKQ and QKKIQ(S/M)SN-(L/M)(L/M)NQARLKVL are absolutely conserved in animals, while being poorly conserved in plants; they also have a particularly high surface probability (as calculated by Mac Vector, IBI). Similarly, at the C-terminus, the motif NTLESRL(D/E)LI(A/S)QQ is conserved only in animals. As these genes are known to be single-copy both in *Manduca* [10] and *Drosophila* (Fig. 4; see below), it is likely that the same gene product serves both endomembrane and plasma membrane roles, so we speculate that in epithelia there may be as yet unidentified accessory proteins which bind such conserved domains.

Recently, it has been shown in *M. sexta* that V-ATPase activity can be controlled hormonally via reversible association and dissociation of the  $V_1$  headgroups from the  $V_0$  transmembrane sector [15], and that V-ATPases in *D. melanogaster* Malpighian tubules are controlled by cAMP and cGMP [16,17]. It may thus be significant that the insect genes share a C-terminal PKA/PKG phosphoryla-

Fig. 2. (p. 6) Genomic DNA, cDNA and deduced amino-acid sequences of *vha26*, the *D. melanogaster* V-ATPase E-subunit. Double-stranded sequencing of the excised phagemids and cloned genomic DNA fragment was performed according to the Sequenase<sup>™</sup> II protocol (USB) by generation of unidirectional deletions with the Erase-a-Base system (Promega), and also with the aid of synthetic oligo primers when required. The putative polyadenylation signal is underlined.

```

1 caacaaatacacatttttaccctcgcaatcgagggtcacactttcgtgaaatcatatgatcgatttgcagtgaaaattt 80
81 tcagacgttgggcagaaggcaaaagtaacttatcgttttccactttcctcgtgttggcgccggtttccaactcagttcg 160
161 gctgtgaatgtattagcttaataataatttcaattatttccagGCACGGTTGTTGTACGTGGGCTTCTTTAAACACTTGA 240
241 ATTTCCCTTTCGGTTTGTGTCAGTGAATAAATCAGTCAAA ATG GCA CTG AGC GAT GCT GAT GTA CAA AAG 309
1 M A L S D A D V Q K 10
310 CAG gtaattgaaaacttggattgggaacgggcaggcgatcaaggctcgtagggaacaagcaaaacgagagggttcgttt 388
11 Q 11
389 gcccttttgccttttgcatttgcctttgcaataaagatggcgaagtcattgggatctcccaggatcatgtaacttttcacc 468
469 gccagtagtccaattagactgacatccttccaaatcgcccggtcatttgggagttgcccaggatttggacatatttgtgtg 548
549 gctaataagacacatcaatttatttgtccagatagtttgcgtaaaaagtgagtaaaattcgtgtggtcatgtgacac 628
629 ggcccccgcatggagcaatgtgttggagcgagacgactagccctgcacccacactcgtactctctgtcacacgaccag 708
709 cgacccccctacgttatcaaaactttaacgaaaataaataagaggttaggtcttggacgtctcccttttccatttatcat 788
789 gtccagttatcatgtgacacacaggcaactactaacaacgagacgactgtttcag ATC AAG CAC ATG ATG GCG 859
12 I K H M M A 17
860 TTC ATT GAG CAG GAG GCC AAT GAG AAA GCC GAG GAG ATC GAT GCC AAG GCC GAG GAG GAG 919
18 F I E Q E A N E K A E E I D A K A E E E 37
920 TTC AAC ATT GAG AAG GGA CGC CTG GTC CAG CAG CAG CGT CTC AAG ATC ATG GAA TAC TAC 979
38 F N I E K G R L V Q Q Q R L K I M E Y Y 57
980 GAG AAG AAG GAG AAG CAA GTT GAG CTG CAG AAG AAG ATT CAG TCC TCC AAC ATG CTC AAC 1039
58 E K K E K Q V E L Q K K I Q S S N M L N 77
1040 CAG GCT CGT CTG AAG gtgcgtgtcgtccagttggtggccctaacaatataccggaacacacttatttttaaatcat 1114
78 Q A R L K 82
1115 tcgtaatgtaccctgtag GTG CTG AAA GTG CGC GAG GAC CAT GTG AGC AGC GTG CTG GAT GAT 1177
83 V L K V R E D H V S S V L D D 97
1178 GCC CGC AAG CGT CTC GGC GAG GTC ACC AAG AAT CAG TCC GAG TAC GAG ACT GTG CTG ACC 1237
98 A R K R L G E V T K N Q S E Y E T V L T 117
1238 AAG CTC ATC GTC CAG GGC CTG TTC CAG ATC ATG GAG CCC AAG GTG ATC CTG CGC TGC CGC 1297
118 K L I V Q G L F Q I M E P K V I L R C R 137
1298 GAG GTG GAC GTC CCC CTG GTA CGT AAC GTC CTG CCT GCC GCT GTG GAG CAA TAC AAG GCC 1357
138 E V D V P L V R N V L P A A V E Q Y K A 157
1358 CAG ATC AAT CAG AAC GTC GAG CTG TTC ATC GAC GAG AAA GAC TTC CTC TCT GCT GAT ACC 1417
158 Q I N Q N V E L F I D E K D F L S S N M L N 177
1418 TGC GGT GGT GTT GAG CTG CTG GCC CTC AAC GGA CGC ATC AAG gtgagtactgtcctttcgggtggag 1483
178 C G G V E L L A L N G R I K 191
1484 agagagcaatcccaactgatctaacaacaccattcag GTG CCC AAT ACG CTG GAG TCC AGA TTA GAC 1550
192 V P N T L E S R L D 201
1551 CTC ATT TCG CAG CAG CTG GTG CCC GAG AIT CGT AAC GCA CTT TTC GGC CGC AAC GTC AAT 1610
202 L I S Q Q L V P E I R N A L F G R N V N 221
1611 CGC AAA TTC ACC GAC TAA ATTCCTATAAGTGCAAAACAAAACATACTAACCAGAAAGAGAACCAGCATCAACAC 1684
222 R K F T D * 227
1685 CTATTTCAGCAGAACAGTTCAAGTTATTACACAGAGCTCCACCCACTAAATATTGAACCAAGTAAACTTATCTTTTGGC 1764
1765 AGTCAGGAGGCAACAGCTAGGATATATTGATTGTCAAAATCTTTTCCCGTTGTCTGTAAAGTGAATGAAACACTCA 1844
1845 AGAACATTTTCGGTCTTGTGTACGCAACAGTTTAAATAGTAACACACTAAACGCGCATATATATTCTCCGATATATATG 1924
1925 TCTGTATGCCAATACTTATTATATAGTTTAGAGGACACGATCCTAGGAGCATACGAAAGCATAATACGAAGTTTGTATAA 2004
2005 GTTTGTTCGTTTTCCTTTTACATATGCACATGTTTCTGAGCAGTAGGTCTAGATATGTGCTTATATTGTATACATACAT 2084
2085 TTAATAATTTTGCATACATTCTCTGTCCAGAATTTTATTTCAGTTTTCCTCTTTTATTGTATGATATTTCCTGTAGTC 2164
2165 TTTGTAACTTTTATATGTCTATGTCTGTTTATGTTTCGTAATTATCAAGTGACGTTTCAGGAGGAACAACGGCAGTGGAT 2244
2245 CGCCCCCTTTACAGACCCGCTGGCAGGTTGCGATGCGACACACAGCATTTGCTGTCAGCAAGCAGCAAAATGGACCTAA 2324
2325 ACCCCCGATTTCGCTTCTTCGAGGGCAACGACGCTTGTGCAACTGCCACTGGCTCAACGAAAGCCCCGAAAATCATCAA 2404
2405 TGCTGTGTTGTTGTGAGATACCGAGAGTAGAGAATACACACTGCTTAGCACGCGACACTTAATACCCATTTCATTACACAT 2484
2485 GCACCACGACGATGAAGTTTGCCAAGTAGCTAAGTTGTGACCTGACCATCAAGTGCAGCTTTTCACACCCCTCATATACT 2564
2565 ACTTAAAGAAAATATAGAAAATGGAATTAGTTTGTCAATTTAGGCCACTGCGAACTGCACCCGTTTCACCTGACGCT 2644
2645 GGGCCATCATATCAGGCTCTAAAAATCAACACACCATGTTCAAACACGACTAGCATACAGGAGCAGGAGCTACAGTAA 2724
2725 ATTTGAACCTTGTATTTCGCATGTTTCGCAATGTTTCATAGTGTATTCTTCAAGCTCATTTTCTAACCAGTTACCAAGTTC 2804
2805 AATATGATGAATAACTACAAGATTAGCAAAACAAATACAAGTAGCATATGCGTTATTATATACATCGAGTACTATATACA 2884
2885 TTACATGAATACAAAATGCAAGAAAATTTACTTTTAAACAAAATTTATGTTGATATAAAGACAGTATTTCACAAAACATA 2964
2965 AActtaactgtataacaacttctcttttgcattgttctaagtatcctctaaacaaagacatggggttaactattttaagaaa 3044
3045 ttcaatctaggactcaatagttctatagtagta 3076

```

tion site consensus (RKFT) at residues 222–225, although the target threonine is not preserved in other phyla.

Genomic Southern analysis at high stringency with *vha26* cDNA (Fig. 4) suggests that this gene is single-copy.

This is consistent with the in situ hybridisation to polytene chromosome squashes which identifies a single locus at 83B1–4 on the right arm of chromosome 3 (not shown). In principle, this finding would allow us to identify any

Drosophila	1	M	-	-	-	-	A	L	S	D	A	D	V	Q	K	Q	I	K	H	M	M	A	F	I	E	Q	E	A	N	E	K	A	E	E	-	D	A	K	A	E	E	E	F	N	I	E	K	G	R	L	45			
Manduca	1	M	-	-	-	-	A	L	S	D	A	D	V	Q	K	Q	I	K	H	M	M	A	F	I	E	Q	E	A	N	E	K	A	E	E	-	D	A	K	A	E	E	E	F	N	I	E	K	G	R	L	45			
Homo	1	M	-	-	-	-	A	L	S	D	A	D	V	Q	K	Q	I	K	H	M	M	A	F	I	E	Q	E	A	N	E	K	A	E	E	-	D	A	K	A	E	E	E	F	N	I	E	K	G	R	L	45			
Bos	1	M	-	-	-	-	A	L	S	D	A	D	V	Q	K	Q	I	K	H	M	M	A	F	I	E	Q	E	A	N	E	K	A	E	E	-	D	A	K	A	E	E	E	F	N	I	E	K	G	R	L	45			
Mus	1	M	-	-	-	-	G	L	R	H	A	D	V	Q	K	Q	I	K	H	M	M	A	F	I	E	Q	E	A	N	E	K	A	E	E	-	D	A	K	A	E	E	E	F	N	I	E	K	G	R	L	46			
Arabidopsi	1	M	-	-	-	-	N	D	G	D	V	S	R	Q	I	Q	Q	M	V	R	F	I	R	Q	E	A	E	E	K	A	N	E	-	S	V	P	A	E	E	E	E	F	N	I	E	K	L	43						
Spinacia	1	M	-	-	-	-	N	D	T	D	V	Q	K	Q	I	Q	Q	M	V	R	F	I	R	Q	E	A	E	E	K	A	N	E	-	S	V	A	A	E	E	E	E	F	N	I	E	K	L	43						
Mesembryan	1	M	-	-	-	-	N	D	T	D	V	Q	K	Q	I	Q	Q	M	V	R	F	I	R	Q	E	A	E	E	K	A	N	E	-	S	V	S	A	E	E	E	E	F	N	I	E	K	L	43						
Neurospora	1	M	S	-	Q	V	H	A	L	S	D	D	Q	M	G	Q	E	L	R	K	M	T	A	F	I	K	Q	E	A	E	E	K	A	N	E	-	Q	I	K	A	D	E	E	E	F	N	I	E	K	S	K	L	49	
Saccharomy	1	M	S	S	A	I	T	A	L	T	P	N	Q	V	N	D	E	L	N	K	M	Q	A	F	I	R	K	E	A	E	E	K	A	N	E	-	I	Q	L	K	A	D	E	E	E	F	N	I	E	K	T	N	L	50
Enterococc	1	M	R	-	-	-	L	I	N	V	N	P	T	R	M	E	L	T	R	L	K	Q	L	-	-	T	T	A	T	R	G	H	K	L	K	D	K	C	D	E	-	-	L	M	R	Q	F	I	L	42				
Methanosar	1	M	G	H	E	I	V	V	-	-	-	K	D	I	Q	E	G	A	R	A	E	V	S	R	I	K	A	E	G	D	A	K	A	S	E	I	L	N	E	A	K	E	I	-	-	-	-	-	-	39				
Haloferax	1	M	S	L	D	N	V	V	-	-	-	E	D	I	R	D	E	A	R	A	R	A	E	D	I	R	G	D	G	Q	E	Q	A	D	E	I	V	A	E	A	E	A	D	-	-	-	-	-	-	39				
Drosophila	46	V	Q	Q	R	L	K	I	M	E	Y	E	K	K	E	K	Q	V	-	E	L	Q	K	K	I	Q	S	S	N	M	L	N	Q	A	R	L	K	V	L	K	V	R	E	D	H	V	S	S	V	L	95			
Manduca	46	V	Q	Q	R	L	K	I	M	E	Y	E	K	K	E	K	Q	V	-	E	L	Q	K	K	I	Q	S	S	N	M	L	N	Q	A	R	L	K	V	L	K	V	R	E	D	H	V	R	N	V	L	95			
Homo	46	V	Q	T	Q	R	L	K	I	M	E	Y	E	K	K	E	K	Q	I	-	E	Q	Q	K	K	I	Q	S	S	N	M	L	N	Q	A	R	L	K	V	L	R	A	R	D	D	L	T	D	L	95				
Bos	46	V	Q	T	Q	R	L	K	I	M	E	Y	E	K	K	E	K	Q	I	-	E	Q	Q	K	K	I	Q	S	S	N	M	L	N	Q	A	R	L	K	V	L	R	A	R	D	D	L	T	D	L	95				
Mus	47	V	E	T	Q	R	L	K	I	M	E	Y	E	K	K	E	K	Q	I	R	Q	Q	K	K	I	Q	S	S	N	M	L	N	Q	A	R	L	K	V	L	R	A	R	D	D	L	T	D	L	97					
Arabidopsi	44	V	E	A	E	K	K	I	R	Q	D	Y	E	K	K	E	K	Q	A	-	D	V	R	K	I	D	Y	S	M	Q	L	N	A	S	R	I	K	V	L	Q	A	Q	D	I	V	N	A	M	K	93				
Spinacia	44	V	E	A	E	K	K	I	R	P	E	Y	E	K	K	E	K	Q	V	-	Q	V	R	K	I	E	Y	S	M	Q	L	N	A	S	R	I	K	V	L	Q	A	Q	D	I	V	N	A	M	K	93				
Mesembryan	44	V	E	A	E	K	K	I	R	Q	E	Y	E	K	K	E	K	Q	V	-	D	V	R	K	I	E	Y	S	M	Q	L	N	A	S	R	I	K	V	L	Q	A	Q	D	I	V	N	A	M	K	93				
Neurospora	50	V	R	Q	E	T	D	A	I	D	S	A	Y	A	K	K	F	K	Q	A	-	M	S	Q	I	T	R	S	T	M	A	N	K	T	R	L	B	V	L	G	A	R	G	E	L	D	E	F	99					
Saccharomy	51	V	R	N	E	T	N	I	D	G	N	E	K	S	K	L	K	K	A	-	L	S	Q	I	T	K	S	T	I	A	N	K	M	B	K	V	L	S	A	R	E	Q	S	L	E	R	I	F	100					
Enterococc	43	L	I	R	K	N	N	E	L	R	Q	A	I	E	K	-	E	T	Q	T	A	M	K	D	F	V	L	A	S	T	V	E	A	F	I	D	E	L	L	A	P	A	E	N	V	S	I	V	92					
Methanosar	40	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	86						
Haloferax	40	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	86						
Drosophila	96	D	D	A	R	K	R	L	G	E	V	T	K	N	Q	S	E	Y	E	T	V	L	T	K	L	I	V	Q	G	L	F	Q	I	M	E	P	K	V	L	R	C	R	E	V	D	V	P	L	V	R	N	146		
Manduca	96	D	E	A	R	K	R	L	A	E	V	P	K	D	I	K	L	Y	S	D	L	L	V	T	L	I	V	Q	A	L	F	Q	L	V	E	P	T	V	T	L	R	V	R	Q	A	D	K	A	L	V	E	S	146	
Homo	96	N	E	A	K	Q	R	L	S	K	V	V	K	D	T	T	R	Y	Q	V	L	L	D	G	L	V	L	Q	G	L	Y	Q	L	L	E	P	R	M	I	V	R	C	R	K	Q	D	F	P	L	V	K	A	146	
Bos	96	N	E	A	K	Q	R	L	S	K	V	V	K	D	T	T	R	Y	Q	V	L	L	D	G	L	V	L	Q	G	L	Y	Q	L	L	E	P	R	M	I	V	R	C	R	K	Q	D	F	P	L	V	K	A	146	
Mus	98	N	E	A	K	Q	R	L	S	K	V	V	K	D	T	T	R	Y	Q	V	L	L	D	G	L	V	L	Q	G	L	Y	Q	L	L	E	P	R	M	I	V	R	C	R	K	Q	D	F	P	L	V	K	A	148	
Arabidopsi	94	D	Q	A	A	K	D	L	L	N	V	S	R	D	E	Y	A	Y	K	Q	L	L	K	D	L	I	V	Q	C	L	L	R	L	K	E	P	S	V	L	L	R	C	R	E	D	I	G	L	V	E	A	144		
Spinacia	94	E	E	A	A	K	E	L	L	R	V	S	G	D	H	H	Y	K	R	L	L	K	E	L	V	Q	S	L	L	R	L	R	E	P	G	V	L	L	R	C	R	E	D	D	V	H	L	V	E	H	144			
Mesembryan	94	E	A	A	S	K	E	L	L	V	S	G	D	H	H	Y	K	R	N	L	L	K	E	L	V	Q	S	L	L	R	L	K	E	P	A	V	L	L	R	C	R	E	D	D	K	H	H	V	H	R	144			
Neurospora	100	E	A	A	S	A	Q	L	G	A	T	H	D	L	G	Y	K	D	I	L	R	D	L	E	G	E	Y	A	M	N	E	P	E	L	V	I	B	A	R	Q	A	D	Y	A	V	R	E	150						
Saccharomy	101	E	E	T	K	E	K	L	S	G	I	A	N	N	R	D	E	Y	K	P	I	L	Q	S	L	I	Y	E	A	T	L	K	L	E	P	K	A	T	V	K	A	L	E	R	D	V	I	E	S	151				
Enterococc	93	V	E	K	N	I	M	S	V	K	V	P	L	M	N	F	Q	Y	D	E	T	L	N	E	T	P	L	E	Y	G	-	L	H	S	N	A	E	L	D	S	I	D	G	F	T	Q	L	P	K	142				
Methanosar	87	N	Q	T	V	E	N	I	K	S	M	S	A	S	K	K	E	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	127					
Haloferax	87	E	Q	V	E	R	E	L	A	E	L	G	D	R	E	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-</																						

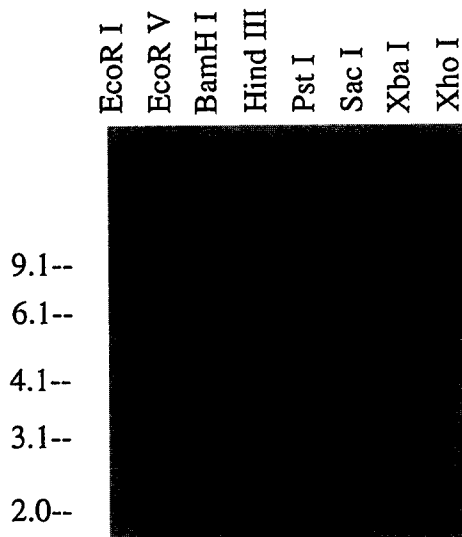


Fig. 4. Southern blot of *D. melanogaster* genomic DNA. Genomic DNA purified from wild-type *D. melanogaster* (Canton S) was cleaved with a range of restriction endonucleases, separated by electrophoresis in a 0.8% agarose gel, blotted to Hybond N (Amersham), and hybridised with a  $^{32}$ P-labelled random-primed probe of *vha26* cDNA. Prehybridisation was in Church buffer (7% SDS, 1% BSA, 1 mM EDTA, 0.25 M  $\text{Na}_2\text{HPO}_4$ , pH 7.2) at 65°C for 3 h, and hybridisation was in Church buffer overnight. The blot was then washed at 65°C in  $2\times$  SSPE, 0.1% SDS for 30 min;  $0.5\times$  SSPE, 0.1% SDS for 30 min; and finally in  $0.1\times$  SSPE, 0.1% SDS for 30 min, and exposed to X-ray film for 0.5 days.

plausible mutant phenotypes at this locus; however, at the time of writing, there are no genes nearby which we consider good candidates. The 188 kb 83B interval con-

tains three identified genes: *gorp*, a gene implicated in meiosis [18], *nmdaR*, a glutamate receptor [19], and a tRNA gene [20]. However there are also several lethal P-element insertions, suggesting that inactivation of the *vha26* locus by 'local jumping' of the P-element may be feasible, or even that an existing P-element insertion might already represent a lethal allele of this gene.

Northern blots of total RNA probed with *vha26* cDNA identify a single band equivalent to a transcript of approx. 2.3 kb (Fig. 5). We cannot exclude the possibility of transcriptional richness in this gene; however, only a single size of RNA was detected, the cDNAs differed only in the length of their 5' UTRs, and the genomic sequence identified so far does not contain alternative exons that could be spliced to yield a product of the same size. The simplest interpretation, therefore, is that a single mRNA species is transcribed from this gene. Equivalent levels of expression are found in adult head, thorax and abdomen (Fig. 5) as would be expected for a V-ATPase. However, the RNA is much reduced during pupation, as is the case with the *D. melanogaster* 67 kDa A subunit (Guo, Y. et al., unpublished results) but not the 14 kDa F subunit [12]. In *M. sexta*, it has been suggested that some of the V-ATPase subunits disappear as the midgut pump shuts down during larval moults [15]; it is possible that downregulation of certain critical mRNA species may be involved.

In summary, we have reported the first genomic sequence and chromosomal localisation for a V-ATPase E-subunit in an animal. Alignment with recently available sequences clearly shows this gene to be conserved across

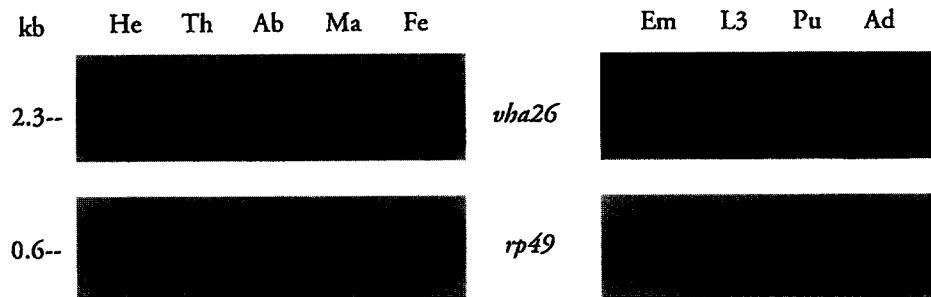


Fig. 5. Northern analysis of the *vha26* gene. Total RNA was isolated from Canton S embryos, larvae, pupae and adults, and from adult heads, thoraxes and abdomens, using RNazol™ [26]. The RNA was separated by electrophoresis in 1% formaldehyde-agarose/MOPS gels, blotted to nitrocellulose, and hybridised with a probe prepared as above. Prehybridisation was in Church buffer at 55°C for 3 h, and hybridisation was in Church buffer overnight. The blots were then washed at 55°C in  $2\times$  SSC, 0.1% SDS for 30 min;  $0.5\times$  SSC, 0.1% SDS for 30 min; and finally in  $0.1\times$  SSC, 0.1% SDS for 30 min. They were exposed to Fuji X-ray film for 1–3 days. Sizes were determined with respect to an RNA ladder (Gibco BRL). Left panel: adult tissues. He, head; Th, thorax; Ab, abdomen; Ma, males; Fe, females. Right panel: developmental Northern. Em, embryo; L3, third instar larva; Pu, pupa; Ad, adult. The lower panels show the same blots, stripped and reprobed with cDNA for the housekeeping gene *rp49*, as controls for differences in RNA loading.

Fig. 3. (p. 7) Alignment of V-ATPase E-subunit homologues, generated by Clustal V (Tompson et al., 1994) and hand-optimised. Residues identical to those in *Drosophila* are boxed, whereas conservative substitutions compared with the *D. melanogaster* sequence (PAM250 scoring matrix) are shaded. Bold 'I' bars denote introns in the *D. melanogaster* or *N. crassa* sequences. The species and accession numbers of the sequences aligned are: *Drosophila*, *Drosophila melanogaster* U38198 and U38951 (this paper); *Manduca*, *Manduca sexta* P31402 [10]; *Homo*, *Homo sapiens* P36543 [22]; *Bos*, *Bos taurus* P11019 [23]; *Mus*, *Mus musculus* U13841; *Arabidopsis*, *Arabidopsis thaliana* X92117 (Dietz, K.J. and Arbing, B., unpublished results); *Mesembryanthemum crystallinum* X92118 (Dietz, K.J. and Arbing, B., unpublished results); *Spinacia*, *Spinacia oleracea* X96785 (Dietz, K.J. and Arnold, J., unpublished); *Neurospora*, *Neurospora crassa* U17641 [6]; *Saccharomyces*, *Saccharomyces cerevisiae* M60663 [7]; *Enterococcus*, *Enterococcus hirae* sodium ATPase subunit D X76913 [24]; *Haloferax*, *Haloferax volcanii* *atpD* gene X79516 [25], *Methanosa*, *Methanosarcina mazei* *ahaE* gene MMU47274 (Wilms, R. et al., unpublished results).

eukaryote and prokaryote phyla, and it is possible to identify extended motifs diagnostic of either all members, or merely animal members, of the family. Expression studies suggest that the mRNA may fall into a subclass of V-ATPase subunits which is not expressed continually during the life of the insect. This characterisation of *vha26* is the first step in the further elucidation of the function of the subunit in an organismal context by *Drosophila* genetics.

We are grateful to Prof. H. Wiczorek for providing the cDNA clone of *Manduca sexta* V-ATPase E-subunit. This work was supported by MRC grant G9120579CB.

## References

- [1] Nelson, N. (1992) *J. Exp. Biol.* 172, 19–27.
- [2] Nelson, N. and Taiz, L. (1989) *Trends Biochem. Sci.* 14, 113–116.
- [3] Takase, K., Kakinuma, S., Yamato, I., Konishi, K., Igarashi, K. and Kakinuma, Y. (1994) *J. Biol. Chem.* 269, 11037–11044.
- [4] Wiczorek, H. (1992) *J. Exp. Biol.* 172, 335–343.
- [5] Dow, J.A.T. (1994) in *Organellar Proton-ATPases* (Nelson, N., ed.), pp. 75–102, R.G. Landes Company, Austin, TX.
- [6] Bowman, E.J., Steinhardt, A. and Bowman, B.J. (1995) *Biochim. Biophys. Acta* 1237, 95–98.
- [7] Foury, F. (1990) *J. Biol. Chem.* 265, 18554–18560.
- [8] Ho, M.N., Hill, K.J., Lindorfer, M.A. and Stevens, T.H. (1993) *J. Biol. Chem.* 268, 221–7.
- [9] Hemken, P., Guo, X.L., Wang, Z.Q., Zhang, K. and Gluck, S. (1992) *J. Biol. Chem.* 267, 9948–57.
- [10] Graf, R., Harvey, W.R. and Wiczorek, H. (1994) *Biochim. Biophys. Acta* 1190, 193–6.
- [11] Meagher, L., McLean, P. and Finbow, M.E. (1990) *Nucleic Acids Res.* 18, 6712.
- [12] Guo, Y., Kaiser, K., Wiczorek, H. and Dow, J.A.T. (1996) *Gene*, in press.
- [13] Cavener, D.R. (1987) *Nucleic Acids Res.* 15, 1353–1361.
- [14] Dow, J.A.T., Goodwin, S.F. and Kaiser, K. (1992) *Gene* 122, 355–360.
- [15] Sumner, J.P., Dow, J.A.T., Earley, F., Klein, U., Jäger, D. and Wiczorek, H. (1995) *J. Biol. Chem.* 270, 5649–5653.
- [16] Dow, J.A.T., Maddrell, S.H.P., Görtz, A., Skaer, N.V., Brogan, S. and Kaiser, K. (1994) *J. Exp. Biol.* 197, 421–428.
- [17] Dow, J.A.T., Maddrell, S.H.P., Davies, S.-A., Skaer, N.J.V. and Kaiser, K. (1994) *Am. J. Physiol.* 266, R1716–R1719.
- [18] Castrillon, D.H., Gonczy, P., Alexander, S., Rawson, R., Eberhart, C.G., Viswanathan, S., DiNardo, S. and Wasserman, S.A. (1993) *Genetics* 135, 489–505.
- [19] Ultsch, A., Schuster, C.M., Laube, B., Betz, H. and Schmitt, B. (1993) *FEBS Lett.* 324, 171–177.
- [20] Dunn, R., Delaney, A.D., Gillam, I.C., Hayashi, S., Tener, G.M., Grigliatti, T., Misra, V., Spurr, M.G., Taylor, D.M. and Miller, R.C., Jr. (1979) *Gene* 7, 197–215.
- [21] Gräf, R., Lepier, A., Harvey, W.R. and Wiczorek, H. (1994) *J. Biol. Chem.* 269, 3767–3774.
- [22] Van Hille, B., Vanek, M., Richener, H., Green, J.R. and Bilbe, G. (1993) *Biochem. Biophys. Res. Commun.* 197, 15–21.
- [23] Hirsch, S., Strauss, A., Masood, K., Lee, S., Sukhatme, V. and Gluck, S. (1988) *Proc. Natl. Acad. Sci. USA* 85, 3004–3008.
- [24] Solioz, M. and Davies, K. (1994) *J. Biol. Chem.* 269, 9453–9459.
- [25] Steinert, K., Kroth-Pancic, P.G. and Bickel-Sandkotter, S. (1995) *Biochim. Biophys. Acta* 1249, 137–144.
- [26] Chomczynski, P. and Sacchi, N. (1987) *Anal. Biochem.* 162, 156–159.