ELSEVIER

Biochimica et Biophysica Acta 1190 (1994) 193–196

BBA Biochimica et Biophysica Acta

Short Sequence-Paper

Cloning, sequencing and expression of cDNA encoding an insect V-ATPase subunit E

Ralph Gräf [a], William R. Harvey [b], Helmut Wieczorek [a,*]

[a] *Zoologisches Institut der Universität, Postfach 202136, D-80021 München (Germany);* [b] *Department of Biology, Temple University, Philadelphia, PA, USA*

(Received 28 October 1993)

Abstract

This work presents the first invertebrate cDNA sequence encoding subunit E of a V-ATPase. It was cloned by immuno-shotgun screening of a *Manduca sexta* (Insecta, Lepidoptera, Sphingidae) posterior larval midgut cDNA library. The amino acid sequence was 64% identical to that of the mammalian E-subunit and 34% to that of yeast. Southern and Northern blots suggested the existence of only one gene encoding the insect subunit E.

Key words: Vacuolar H^+-ATPase; ATPase, H^+-; Insect midgut; cDNA cloning; (*M. sexta*)

V-ATPases are multimeric proton pumps energizing membranes of acidic organelles and plasma membranes [1]. At least three subunits are homologous in all V-ATPases sequenced so far: two peripheral subunits, the 70 kDa-catalytic subunit A and the 60 kDa-regulatory subunit B, plus one 17 kDa-membrane integral proton channel [2]. Further polypeptides have been proposed to be subunits of V-ATPases, but so far it is not clear whether they are constitutive parts of all members of the V-ATPases family. For the 30 kDa-subunit E, only three amino acid sequences derived from cDNA's have been reported so far, two from mammals and one from yeast [3–5]. Despite this lack of quantitative evidence, subunit E is widely accepted as a general and constituent part of V-ATPases. Here we report the cloning, sequencing and expression of the first invertebrate subunit E and further support this prevailing view. Our source of subunit E was the V-ATPase of the model insect, *Manduca sexta* (Lepidoptera, Sphingidae). This V-ATPase occurs in the apical plasma membrane of larval midgut cells where it works in parallel with a K^+/nH^+ antiporter to constitute an electrogenic K^+ pump [6]. Several subunits of this V-ATPase have already been cloned and sequenced: subunit A [7], subunit B [8], the proteolipid subunit [9] and a novel 14 kDa subunit [10].

Shotgun screening of a lambda ZAP cDNA library from larval *Manduca sexta* posterior midgut [7] was performed with a rabbit polyclonal antiserum raised against the purified *M. sexta* plasma membrane V-ATPase holoenzyme [11]. Among 17 positive phage clones, 2 identical clones encoded a truncated polypeptide homologous to the bovine, human and yeast V-ATPase subunit E sequence [3–5] as revealed by DNA sequencing (Fig. 1; Sequenase Kit, USB) of pBluescript SK(−) clones obtained after in vivo excision [12]. We amplified missing cDNA 5′ ends by PCR as described previously [7] (see Fig. 1 for details). The 280 bp fragment obtained after *Eco*RI/*Sac*I digestion of the PCR product was subcloned into pBluescript II KS(−) (Stratagene) and gave rise to 64 bp of additional sequence information. Totally, the cloned cDNA sequence made up 1609 bp (Fig. 1c). The open reading frame encoded a hydrophilic protein (hydropathy plot not shown) with a calculated molecular mass of 26089 Da.

To confirm that the cloned subunit E corresponded to the electrophoretically characterized 28 kDa

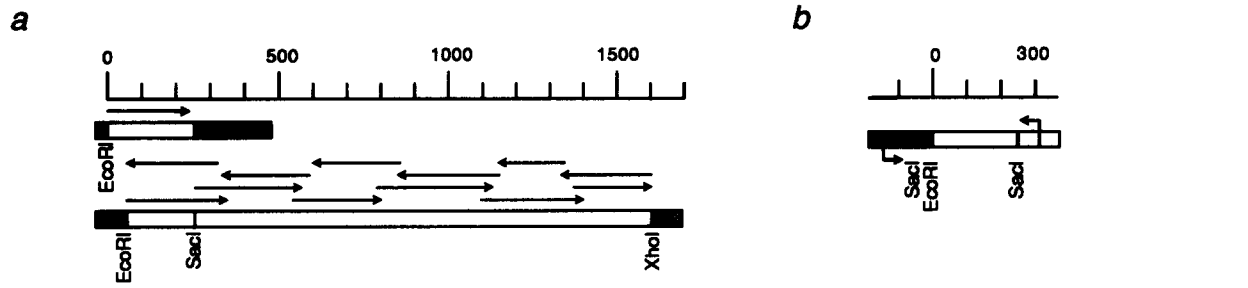
* Corresponding author. Fax: +49 895902450.

The nucleotide sequence reported in this paper has been submitted to the EMBL/GenBank Data Library under the accession number X67131.

SSDI 0005-2736(93)E0354-D

polypeptide of the purified V-ATPase [13,14], we expressed it in *Escherichia coli* TB-1. For this purpose, we cloned the complete subunit E coding sequence obtained by ligation of the 5′ cDNA sequence and the truncated cDNA sequence at the *Sac*I site (Fig. 1) into the pMAL c2 vector [10] (Protein Fusion and Purification System from NEB). Subunit E was expressed as a fusion protein together with *malE* encoded maltose binding protein (26-MBP) and cleaved with factor Xa at the fusion site. Since the recombinant subunit E protein and the native 28 kDa polypetide were strongly labeled by the monoclonal antibody 90–7 in immunoblots (Fig. 2) and since both appeared at exactly the same position in the gel, the native 28 kDa polypeptide is proven to be the V-ATPase subunit E.

Although the function of subunit E in V-ATPases has not been ascertained so far, results from yeast mutants [5,15] and immunological evidence from mam-



c

```
                                                                  ↓
GGCTCTTTATCGGTTGTTTTTCGTGTTAGTGGTGCTAAATAATAACCTACAGCCATGGCGCTCAGCGATGCCGATGTTCAAAAGCAGATC    90
                                        *  *          M  A  L  S  D  A  D  V  Q  K  Q  I      12
AAGCATATGATGGCCTTCATCGAACAAGAGGCCAATGAAAAGGCCGAAGAAATCGATGCTAAGGCGGAGGAGGAGTTCAACATCGAAAAG   180
K  H  M  M  A  F  I  E  Q  E  A  N  E  K  A  E  E  I  D  A  K  A  E  E  E  F  N  I  E  K     42
GGCCGTCTGGTTCAGCAGCAACGCCTCAAGATCATGGAATACTATGAAAAGAAAGAAAAACAAGTAGAGCTCCAGAAAAAGATTCAATCA   270
G  R  L  V  Q  Q  Q  R  L  K  I  M  E  Y  Y  E  K  K  E  K  Q  V  E  L  Q  K  K  I  Q  S     72
TCGAACATGCTGAACCAGGCTCGTCTGAAGGTACTAAAAGTGCGTGAGGACCACGTACGTAACGTACTGGACGAGGCCCGCAAGCGCCTG   360
S  N  M  L  N  Q  A  R  L  K  V  L  K  V  R  E  D  H  V  R  N  V  L  D  E  A  R  K  R  L    102
GCTGAGGTACCCAAGGACATCAAGCTGTACTCAGATCTCCTGGTCACTCTCATTGTGCAGGCTCTCTTCCAGCTTGTGGAGCCCACAGTG   450
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    132
ACCCTCCGCGTGCGACAAGCCGACAAAGCTCTGGTTGAATCTCTGCTTGGCAGAGCCCAACAGGATTACAAGGCCAAGATCAAGAAGGAT   540
T  L  R  V  R  Q  A  D  K  A  L  V  E  S  L  L  G  R  A  Q  Q  D  Y  K  A  K  I  K  K  D    162
GTCGTGCTCAAGATCGACAATGAGAACTTCCTTCCCCCTGACACCTGTGGAGGAATTGAGCTGATTGCTGCCAAGGGACGCATCAAGATC   630
V  V  L  K  I  D  N  E  N  F  L  P  P  D  T  C  G  G  I  E  L  I  A  A  K  G  R  I  K  I    192
AGCAACACGCTGGAGTCGCGCCTGGAGCTGATCGCGCAGCAGCTGCTGCCGGAGATCCGTAACGCGCTGTTCGGACGCAACCCTAACCGC   720
S  N  T  L  E  S  R  L  E  L  I  A  Q  Q  L  L  P  E  I  R  N  A  L  F  G  R  N  P  N  R    222
AAATTCACCGACTAAGCCCCAACGGTGACCCACATTCACTCGTTATACTAACAACATATTTAAATTTCATAATATGTAGAAGTAAAATAA   810
K  F  T  D  *                                                                                 226
TATTATAAGTGTTTTCTTTTGCCTAATTAATTATATCGTATAATGCGTACAATTATATCTGTATTAGTATGTATTGATGAATGGTCCCGT   900
TAAAGACACTCGTACAAATGCGATCATGTGGATTTCTGATTCTCGGCATGTCATTGTTCACTCGACTGTACATTAGTTTCGCCAATTAAC   990
CAACATGCAAATGTTAAATGTTTTATTATAACAATTCATTCGTATAGCATCATGTCTATGTATTTAGGGAACGTAGTTTATAGGATTTGA  1080
GAACAGTTAAGACGTTACCACATTTAGGCGAAATAATATATGTCGCTAAGTATTTCATAGAAATTGAGGCACCTATCTGATACAGTATTA  1170
TAGCACAATGACAGTTTGAGCAGTAGATGGTTAACGTATCTAAAACTATTGCGTTTCAAATCTTCCATATTTAATGTTACTGAATGTTTT  1260
GAATATTTATCGTGTTTACCTGTTTATTGGCTGATGTAGGAGGCATTGAGGTCTGTAGGTTTGAGACTTGATGGCAGCATCCCATGCTGT  1350
TTCTACTTAAATCGAAGCAAAGATTTGGCTGGTGGCAGTGTGGTGTCAAGTGCCAACTCTACAGCTGCATAACATTCCTACCTCGTGTTC  1440
CAACTGGCAATATGATATTTTGGCTTGCTTAGCAGATGCAAAGAGCACCTTAGCTCTTATAAAATAAGATACAAGTGTAATTTCCGTTAT  1530
GGCAATGTTAAAAATTCTTATGTAAATATATAGATGTGTTGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  1609
```

Fig. 1. cDNA sequencing and predicted amino acid sequence for the *Manduca sexta* V-ATPase subunit E. (*a*) cDNA sequencing strategy. The lower bar symbolizes the cDNA derived from immunoscreening, the upper one the PCR derived fragment. The *Sac*I site used for subcloning is marked. Each arrow indicates the direction of plasmid sequencing and the sequence length determined from one priming site using T3, T7 or custom synthesized primers. (*b*) Strategy for amplifying 5′-cDNA ends by PCR. Arrows indicate the priming directions of M13 (−48) primer (left) and sequence specific primer (right). The scales in (*a*) and (*b*) indicate the distance in bp, bars symbolize DNA, vector sequences are shaded and used restriction sites are marked. (*c*) cDNA sequence and predicted amino acid sequence. The 5′ end of the truncated cDNA clone derived from immunoscreening is indicated by an arrow. Stop codons are marked by asterisks, the PCR-priming site for amplification of the cDNA 5′ end is double underlined and the *Sac*I site representing the end of the PCR derived cDNA 5′ end clone is underlined.

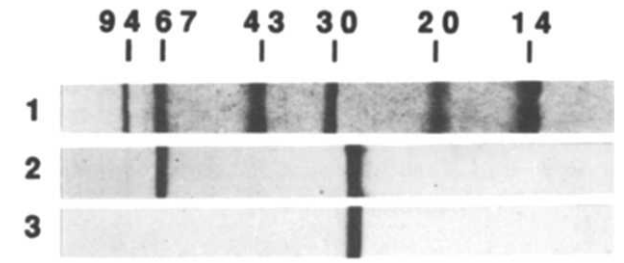


Fig. 2. Identity of subunit E expressed in *E. coli* with the 28 kDa polypeptide of the purified V-ATPase. Standard proteins, 26-MBP fusion protein cleaved with factor Xa proteinase at the fusion site (26-MBP-Xa) and purified V-ATPase were electrophoresed on SDS-polyacrylamide gels [11]. Protein amounts were determined as described [14]. 26-MBP (0.67 mg/ml) purified by affinity chromatography on amylose resin was cleaved by factor Xa proteinase (3 μg/ml) in 50 μl volumes containing 0.2 M NaCl, 1 mM Na_2EDTA, 10 mM Tris-HCl (pH 7.4). Samples containing 2% SDS and 2% mercaptoethanol were heated at 98°C for 3 min. *Lane 1*, 5 μg of standard proteins stained with Coomassie brilliant blue (molecular masses are indicated in kDa). *Lanes 2 and 3*, immuno-blots stained with the monoclonal antibody 90–7. *Lane 2*, 5 μg 26-MBP-Xa (the band at 67 kDa represents uncleaved 26-MBP); *lane 3*, 3 μg of purified *M. sexta* V-ATPase [13].

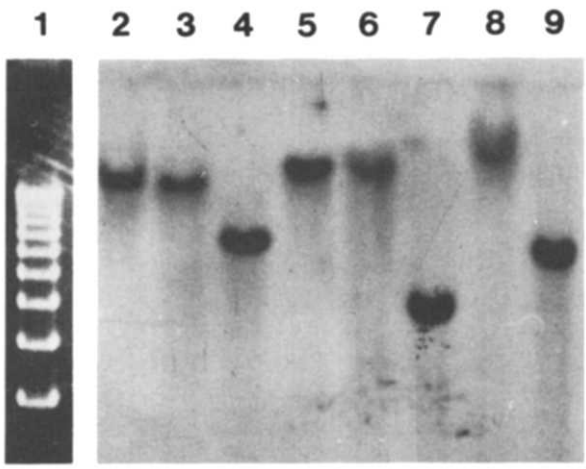


Fig. 3. Southern blot hybridization of genomic DNA with a subunit E specific probe. *Lane 1*, 1 μg of standard DNA fragments stained with ethidium bromide. The sizes of the standard fragments in bp are 12216, 11198, 10180, 9162, 8144, 7126, 6108, 5090, 4072 and 3054. *Lanes 2–9*, in each case 5 μg of genomic DNA isolated from fifth instar larval midguts was digested with restriction enzymes [10] and visualized after Southern hybridization [16] by Lumigen PPD-fluorography (Boehringer Mannheim protocol).The hybridization probe was a double strand DNA spanning base pairs 132 to 627, generated by PCR using digoxigenin (DIG) labeled dUTP (Boehringer Mannheim) [18]. Hybridization and stringency washing [10] was performed at 60°C. The following restriction enzymes were used: *lane 2*, *Apa*I; *lane 3*, *Bst*XI; *lane 4*, *Eco*RI; *lane 5*, *Eco*RV; *lane 6*, *Pst*I; *lane 7*, *Pvu*II; *lane 8*, *Sma*I; *lane 9*, *Xho*I.

malian kidney regarding differential expression of subunit E isoforms [3,4] argue in favour of a role in biogenesis and/or targeting of the holoenzyme. In *M. sexta* midgut tissue, the plasma membrane V-ATPase is certainly not the only V-ATPase since endosomal V-ATPases must also be present. If targeting is a function of the E-subunit then more than one gene should exist and/or more than one mRNA should occur. Southern blot hybridizations of genomic DNA [10,16] at 60°C as well as at 52°C (i.e., 21°C below T_m according to the equation of Meinkoth and Wahl [17]; data not shown) revealed only one restriction fragment for each of the eight digests (Fig. 3). Thus, only one gene encoding subunit E from all V-ATPase sources in *M. sexta* seems to exist. Northern blot hybridizations showed only one mRNA of approximately 1850 bases (data not shown). However, we cannot exclude the occurrence of a differential splicing mechanism leading to different mRNAs with the same length.

Although the sequence of the cloned *M. sexta* cDNA is without doubt that of an E-subunit, comparison of its amino acid sequence to the published subunit E sequences from bovine kidney [3], human kidney [4] and yeast [5] revealed significant differences (Fig. 4). Whereas the yeast/*M. sexta* sequence identity was only 34% and evenly distributed, the bovine/*M. sexta* sequence identity was 64% and clustered in two regions. The N-terminal region (amino acids 1 to 84) had 93% sequence identity and the C-terminal region (amino acids 190 to 226) had 76% identity but the middle region (amino acids 85 to 189) had only 38% identity.

The high overall sequence identity and the parallel clustering of identities in mammal and insect subunits

```
Manduca      MALSDADVQKQIKHMMAFIEQEANEKAEEIDAKA  34
Bovine       **********************************  34
Yeast   MSSAIT**TPNQ*NDELNK*Q***RK**E***K**QL**  39

Manduca EEEFNIEKGRLVQQQRLKIMEYYEKKEKQVELQKKIQSS  73
Bovine  **************T***************I*Q*****M*  73
Yeast   DQ*YE***TNI*RNEINN*DGNFKS*L*KAM*SQQ*TK*  78

Manduca NMLNQARLKVLKVREDHVRNVLDEARKRLAEVPKDIKLY 112
Bovine  *LM********RA*D*LITDL*N**KQ**SK*V**TTR* 112
Yeast   TIA*KM*****SA**QSLERIFE*TKEK*SGIANNRDE* 117

Manduca SDLLVTLIVQALFQLVEPTVTLRVRQADKALVESLLGRA 151
Bovine  QV**DG*VL*G*Y**L**RMIV*C*KQ*FP**KAAVQK* 151
Yeast   KPI*QS***E**LK*L**KAIVKALER*VD*I**MKDDI 156

Manduca QQDYKAKIKKD-VVLKIDNENFLPPD-TCGGIELIAAKG 188
Bovine  IPV**VAT*R*-*DVQ**Q*AY**EE-IA**V*IYNGDR 188
Yeast   MRE*GE*AQRAPLEEIVISNDY*NK*LVS**VVVSN*SD 195

Manduca R*KISNTLESRLELIAQQLLPEIRNALFGRNPNRKFTD 226
Bovine  K**V********D*****MM**V*G****A*A****** 226
Yeast   K*E*N****E**K*LSEEA**A**LE*Y*PSKT***F* 233
```

Fig. 4. Amino acid identity pattern of the *M. sexta*, bovine and yeast V-ATPase subunit E. Multiple sequence alignment performed with the Clustal program. Amino acid identities of the bovine and yeast E-subunit sequence, respectively, compared to the *M. sexta* sequence are indicated by asterisks. Hyphens indicate gaps introduced for better sequence alignment. The human sequence differing only in six amino acids from the bovine sequence was omitted in this alignment.

could reflect a response to a specific selective pressure during evolution. Since V-ATPases are prominent in plasma membranes of both mammals and insects but not in those of yeast whereas they are present in vacuolar membranes of all three taxa, it is possible that additional selective pressure for targeting the enzyme to the plasma membrane in mammals and insects in addition to its primitive targeting to vacuolar membranes, as in yeast, led to the high, clustered identity of mammal and insect E subunits.

This work was supported by the Deutsche Forschungsgemeinschaft (Wi 698), the European Economic Community (SC1*-CT90-0480), and the National Institutes of Health (AI 22444). We thank Ms. Alexandra Lepier and Dr. Michael G. Wolfersberger for critical comments.

1. References

[1] Harvey, W.R. (1992) J. Exp. Biol. 172, 1–17.
[2] Nelson, N. (1992) J. Exp. Biol. 172, 19–28.
[3] Hirsch, S., Strauss, A., Masood, K., Lee, S., Sukhatme, V. and Gluck, S. (1988) Proc. Natl Acad. Sci. USA 85, 3004–3008.
[4] Hemken, P., Guo, X.-L., Wang, Z.-Q., Zhang, K. and Gluck, S. (1992) J. Biol. Chem. 267, 9948–9957.
[5] Foury, F. (1990) J. Biol. Chem. 265, 18554–18560.
[6] Wieczorek, H. (1992) J. Exp. Biol. 172, 335–344.
[7] Gräf, R., Novak, F.J.S., Harvey, W.R. and Wieczorek, H. (1992) FEBS Lett. 300, 119–122.
[8] Novak, F.J.S., Gräf, R., Waring, R.B., Wolfersberger, M.G., Wieczorek, H. and Harvey, W.R. (1992) Biochim. Biophys. Acta 1132, 67–71.
[9] Dow, J.A.T., Goodwin, S.F. and Kaiser, K. (1992) Gene 122, 355–360.
[10] Gräf, R., Lepier, A., Harvey, W.R. and Wieczorek, H. (1994) J. Biol. Chem., in press.
[11] Wieczorek, H., Putzenlechner, M., Zeiske, W. and Klein, U. (1991) J. Biol. Chem. 266, 15340–15347.
[12] Short, J.M., Fernandez, J.M., Sorge, J.A. and Huse, W.D. (1988) Nucl. Acids Res. 16, 7583–7600.
[13] Schweikl, H., Klein, U., Schindlbeck, M. and Wieczorek, H. (1989) J. Biol. Chem. 264, 11136–11142.
[14] Wieczorek, H., Cioffi, M., Klein, U., Harvey, W.R., Schweikl, H. and Wolfersberger, M.G. (1990) Methods Enzymol. 192, 608–616.
[15] Ho, M.N., Hill, K.J., Lindorfer, M.A. and Stevens, T.H. (1993) J. Biol. Chem. 268, 221–227.
[16] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[17] Meinkoth, J. and Wahl, G. (1984) Anal. Biochem. 138, 267.
[18] Lion, T. and Haas, O.A. (1990) Anal. Biochem. 188, 335–337.