Volume 300, number 2, 119-122 FEBS 10840 March 1992

Cloning and sequencing of cDNA encoding the putative insect plasma membrane V-ATPase subunit A

Ralph Gräf, Frans J.S. Novak, William R. Harvey* and Helmut Wieczorek

Zoologisches Institut der Universität, W-8000 München, Germany

Received 2 February 1992

For the first time a cDNA encoding subunit A of an invertebrate V-ATPase has been sequenced. The cDNA library was prepared from larval midgut of the tobacco hornworm, *Manduca sexta*, and screened with monoclonal antibodies to the midgut plasma membrane subunit A. From the cDNA sequence the insect subunit A is predicted to consist of 617 amino acids with a relative molecular mass of 68,162. The predicted primary structure is similar to that of the published eukaryotic subunit A proteins (*Bos, Daucus, Saccharomyces* and *Neurospora*); it most closely resembles the bovine amino acid sequences with which it has an 83% sequence identity.

Vacuolar H^+-ATPase; Insect midgut; *Manduca*; cDNA cloning; Subunit structure; Potassium pump

1. INTRODUCTION

Vacuolar ATPases (V-ATPases) are a closely related family of proton pumps which, in eukaryotic cells, usually reside in the membranes of acidic organelles. However, they are also present in the plasma membranes of vertebrate renal epithelia and lepidopteran insect midgut epithelia [1,2]. The V-ATPase of the tobacco hornworm, *Manduca sexta*, is the first vacuolar-type proton pump shown to energize secondary active transport in an eukaryotic plasma membrane — a V-ATPase in parallel with a K^+/nH^+ antiporter [3,4] constitutes the well known 'potassium pump' of the midgut goblet cell apical membrane [5].

All of the V-ATPases purified to date share similar functional and structural features [1]. They are multimeric proteins with at least three common subunits: a catalytic subunit A, a regulatory subunit B, and a proton channel subunit with relative molecular masses of approximately 70,000, 60,000 and 17,000, respectively. Although cDNAs of several V-ATPase subunits had been sequenced, no sequence was available for an animal subunit A until those for bovine V-ATPases were reported recently [6,7]. We report here the first invertebrate cDNA sequence encoding a V-ATPase subunit A from the model insect, *Manduca sexta*. Although this insect cDNA appears to encode the subunit A of an ATPase which is the most prominent component of a plasma membrane, its primary structure is remarkably similar to that of cDNAs thought to encode subunit A in vacuolar membranes.

*Permanent address: Department of Biology, Temple University, Philadelphia, PA 19122, USA.

Correspondence address: H. Wieczorek, Zoologisches Institut der Universität, Luisenstraße 14, W-8000 München 2, Germany. Fax: (49) (89) 5902 450.

2. MATERIALS AND METHODS

2.1. *Construction and screening of the cDNA library*

Approximately 4 mg of posterior midgut total RNA was isolated by the AGPC-method [8] using 4 fifth instar larvae of *Manduca sexta*. PolyA$^+$ RNA was purified by oligo(dT) cellulose chromatography [9]. A cDNA library was constructed by directional cloning using the ZAP-cDNA synthesis kit from Stratagene, and by packaging 800 ng of λ-DNA with Gigapack II gold from Stratagene. The primary library consisted of 1.2×10^6 independent clones. The cDNA library was screened initially with rabbit polyclonal antiserum to the purified *Manduca* plasma membrane V-ATPase holoenzyme [4]. Seventy three of approximately 3.3×10^5 recombinant phage clones were positive. Three further screening steps with two different monoclonal antibodies (22F-9 and 20G-2) to the subunit A [10; Klein, Lepier, Förg-Brey and Wieczorek, in preparation] yielded 10 double positive clones.

2.2. *Subcloning and sequencing*

The pBluescript SK(−) phagemid portion containing subunit A-specific cDNA from the 10 positive clones was rescued by in vivo excision and infection of *E. coli* XL-1 Blue [11]. Restriction analysis of inserts revealed that at least 4 of the 10 clones differed from one another with respect to their insert length. The two longest clones obtained after *Eco*RI/*Xho*I double digestion showed 6 restriction fragments of 1000, 720, 390, 310, 210 and 100 bp. Seven appropriate restriction fragments from double and single digestions were subcloned separately. Convenient *Sac*I, *Kpn*I, *Bgl*II and *Bsp*106 restriction sites (Fig. 1a), determined when sequencing the 7 subclones, were used for the construction of 7 additional subclones, which were mainly derivatives of the long 1000 and 720 bp fragments. Restriction fragments for subcloning were separated by electrophoresis on 1% FMC Sea Plaque agarose and cloned into pBluescript II KS(+/−) by in-gel ligation-transformation according to the FMC protocol. XL-1 Blue and JM110 competent cells were prepared according to a standard $CaCl_2$ protocol. Plasmid DNA was usually isolated by a modified alkaline lysis miniprep procedure [12]. All DNA sequencing was per-

formed by the dideoxynucleotide chain termination method [13] using sequenase 2.0 [14] from USB. Most reactions were primed by T3 or T7 primers, some were primed by specific custom synthesized 17-mer or 20-mer oligonucleotides.

2.3. *Cloning of the cDNA 5'-end by PCR*

λ-DNA was prepared from the entire cDNA library by growing phages in LE392 cells and extracting phage DNA according to a standard protocol recommended by Stratagene. λ-DNA was digested with *Kpn*I and separated on a 1% FMC Sea Plaque agarose gel. A gel slice with DNA ranging between 3.5 and 5.8 kb containing insert and vector DNA was remelted and used directly as a PCR template. PCR was carried out according to Lion and Haas [15], using a 23-mer primer (5'-AGCGGATAACAATTTCACACAGG-3') binding 117 bp upstream of the *Eco*RI cloning site and an antisense strand 20-mer primer (position 435 to 454, Fig. 2) and Taq polymerase from Cetus. Annealing was performed at 63°C. Agarose gel analysis of the PCR product revealed 3 clearly visible bands of approximately 700, 600 and 530 bp in a faint smear of unspecifically amplified DNA. The whole reaction product was digested with *Sac*I which cleaves in the polylinker portion and, furthermore, cuts away a 67 bp 3'-fragment in the case of correct PCR products (Fig. 1b). Analysis of a 1% FMC Sea-Plaque agarose gel revealed that only the 600 and 530 bp fragments were cleaved to 450 and 380 bp fragments, respectively. Since the 380 bp fragment presumably corresponded to the cDNA already sequenced (see above), only the gel slice with the 450 bp fragment was used for ligation into pBluescript II KS(−) and transformed into *E. coli*. Plasmid DNA from 6 colonies was isolated and analysed with SacI for insert length, with *Hin*dIII for tandem inserts and with *Eco*RI/HI for insert orientation; it was then sequenced. The resulting *Manduca* subunit A cDNA sequence has been submitted to the EMBL Data Library and has been given accession number X64233.

3. RESULTS AND DISCUSSION

Fourteen subclones were obtained which allowed sequencing of both DNA strands with overlaps at the ends of individual subclones (Fig. 1a). The resulting sequence consisted of 2792 bp. Comparison of the derived amino acid sequence with published sequences for subunit A suggested that the initiator ATG was still missing. To obtain longer clones, and noting that only in-frame clones had been detected by immunoscreening, λ-DNA of the whole cDNA library was purified and used as a template for PCR amplification of all 5'-cDNA ends (Fig. 1b). Three clones found by this strategy were used for sequencing. The longest one (391 bp) was sequenced completely and turned out to be identical in the overlapping region with the previously sequenced, truncated, clone, but bearing an additional 68 bp 5'-sequence when compared to the sequence obtained initially (Figs. 1a and 2). This 391 bp clone contained an in-frame ATG codon which is preceded by a stop codon (Fig. 2) and is located within a sequence which matches closely the consensus sequence for the translation start site in eukaryotes [16]. Furthermore, the initiator Met codon was located similarly as in the bovine, carrot and *Neurospora* sequences. From each of the two shorter clones (386 and 378 bp) 100 bp were sequenced and found to match completely with the corresponding sequence of the longest clone mentioned above.

The polypeptide deduced from the cDNA (Fig. 2) consists of 617 amino acids and has an M_r of 68,162, agreeing well with that determined for subunit A of goblet cell apical membrane V-ATPase by SDS-PAGE [2,21]. The protein shows high amino acid identity with previously sequenced V-ATPase catalytic subunits [6,7,18,19,20], resembling most closely the bovine sequences (83% identity), followed by carrot (68%), yeast (65%) and *Neurospora* (63%). No striking differences with respect to hydropathy, predicted secondary structure or phosphorylation sites could be detected between the insect subunit A sequence and those of the other V-ATPases. An alanine residue next to the initiator methionine indicates methionine removal in the cell [22].

The eukaryotic V-ATPases discussed above are located in endomembranes. By contrast, the *Manduca* V-ATPase is located in the plasma membrane [2,17,21]. It is not certain that the cDNA sequence reported here encodes the plasma membrane V-ATPase since endomembrane V-ATPases are expected to occur in all eukaryotic cells and in the membranes of various organelles, and since we do not know if one or more genes encode the A subunits of endomembrane and plasma

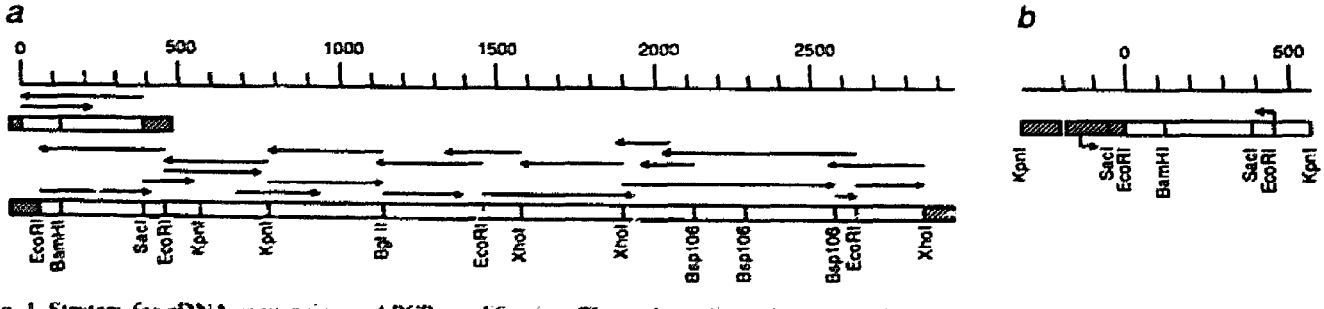

Fig. 1. Strategy for cDNA sequencing and PCR amplification. The scales indicate the distance in bp, bars symbolize DNA, vector sequences are shaded. *a*: cDNA sequencing strategy. The lower bar symbolizes the cDNA derived from immunoscreening, the upper one the cDNA PCR derived fragment. Each arrow indicates the direction of sequencing and the sequence length determined from one subclone. The restriction sites shown were used for subcloning, except the *Bam*HI site, which was needed only for restriction analysis. *b*: Strategy for amplifying 5'-cDNA ends by PCR. Arrows indicate the priming directions. The upstream *Kpn*I site is located 3.47 kb distant from the polylinker *Sac*I site as indicated by the gap in the bar.

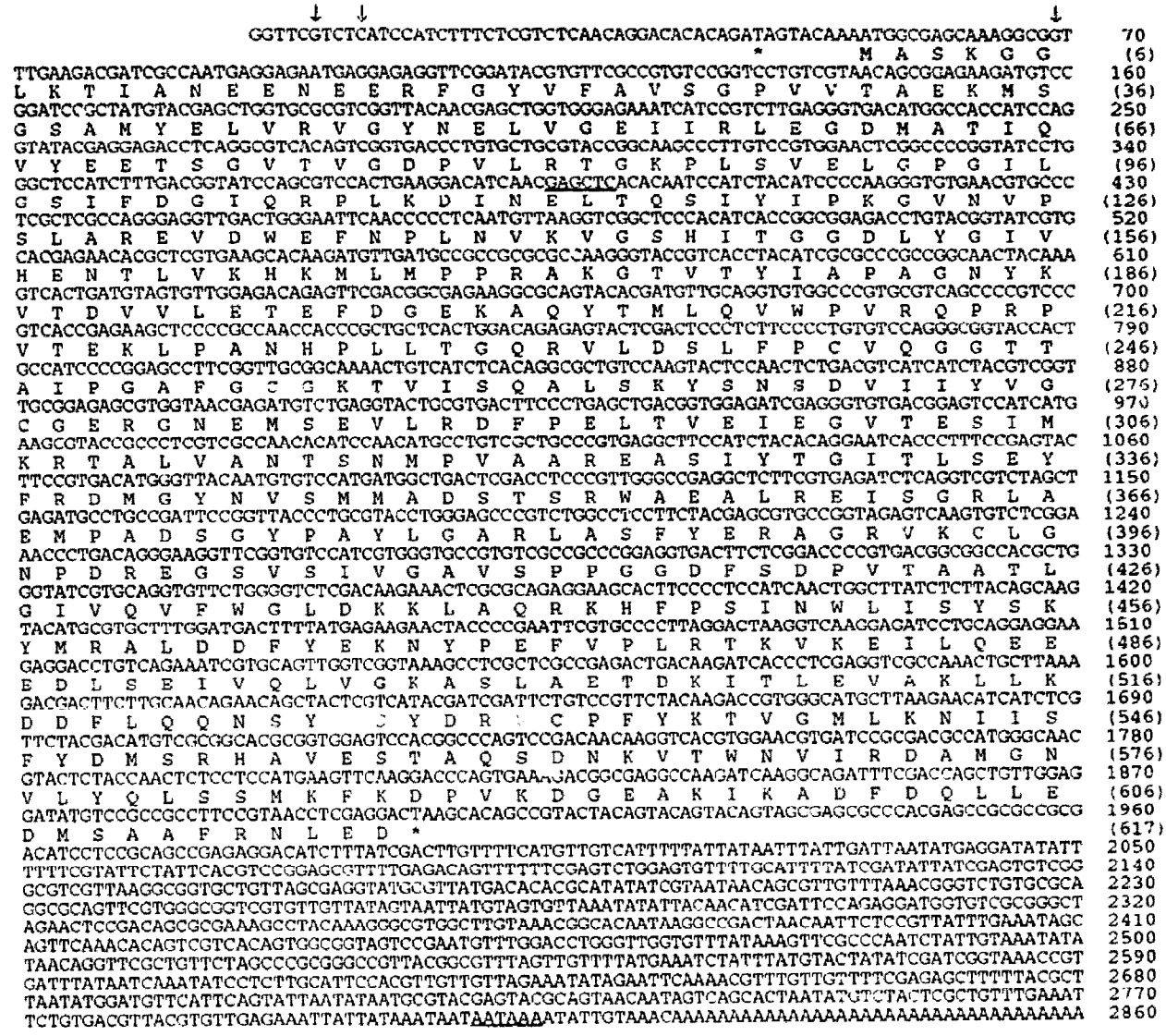

Fig. 2. cDNA and predicted amino acid sequence for the *Manduca* V-ATPase subunit A. The 5'-ends of the two shorter PCR clones and of the cDNA clone derived from immunoscreening are indicated by arrows. Asterisks indicate stop codons, the presumable polyadenylation site and the *Sac*I site representing the end of the PCR clones are underlined.

membrane V-ATPases in *Manduca*. Indeed, there is evidence from plants as well as from vertebrates suggesting the existence of more than one gene for subunit A [6,7,23]. Many published sequences for subunits of V-ATPases have been identified with the organelles from which the V-ATPase protein had been isolated (e.g. that of chromaffin granules or that of clathrin coated vesicles). Similarly, we have evidence suggesting that the cDNA which we have cloned encodes a subunit of the V-ATPase in the insect plasma membrane from which the protein was isolated. The V-ATPase is highly concentrated in the apical plasma membrane of goblet cells, one of the main cell types in the midgut. Indeed, the protein pattern of the purified goblet cell apical membrane as revealed by SDS-PAGE closely resembles that of the purified enzyme [2,4,21]. The monoclonal antibodies used to isolate the cDNA clones labeled the goblet cell apical membranes in immunocytochemical studies with only slight background [10; Klein, personal communication]. Both of these observations suggest that the plasma membrane enzyme is the principal V-ATPase in the midgut. Therefore, it seems likely that our midgut mRNA preparation was a rich source for the expression of plane membrane V-ATPase. Moreover, four independent clones from the midgut cDNA library exhibited an identical DNA sequence in the non-conserved, 5'-terminal region. This substantial evidence argues that, for the first time, the cDNA encoding subunit A of a plasma membrane V-ATPase has been sequenced.

Acknowledgements: This work was supported by the Deutsche Forschungsgemeinschaft (Wi 698), the European Economic Community (SCI*-CT90-0480), and the National Institute of Health (AI 22444). We thank Dr. Ulla Klein and Alexandra Lepier for critically reading the manuscript.

REFERENCES

[1] Forgac, M. (1989) Physiol. Rev. 69, 765–796.
[2] Schweikl, H., Klein, U., Schindlbeck, M. and Wieczorek, H. (1989) J. Biol. Chem. 264, 11136–11142.
[3] Wieczorek, H., Weerth, S., Schindlbeck, M. and Klein, U. (1989) J. Biol. Chem. 264, 11143–1114.
[4] Wieczorek, H., Putzenlechner, M., Zeiske, W. and Klein, U. (1991) J. Biol. Chem. 266, 15340–15347.
[5] Harvey, W.R., Cioffi, M. and Wolfersberger, M.G. (1983) Am. J. Physiol. 244, R163–R175.
[6] Pan, Y., Xu, J., Strasser, J.E., Howell, M. and Dean, G.E. (1991) FEBS Lett. 293, 89–92.
[7] Puopolo, K., Kumamoto, C., Adachi, I. and Forgac, M. (1991) J. Biol. Chem. 266, 24564–24572.
[8] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.
[9] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular cloning: A laboratory Manual, 2nd Edn. Cold Spring Harbor Laboratory Press, New York.
[10] Klein, U., Löffelmann, G. and Wieczorek, H. (1991) J. Exp. Biol. 161, 61–75.
[11] Short, J.M., Fernandez, J.M., Sorge, J.A. and Huse, W.D. (1988) Nucleic Acids Res. 16, 7583–7600.
[12] Birnboim, H.C. and Doly, J. (1979) Nucleic Acids Res. 7, 1513–1523.
[13] Sanger, F., Niklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[14] Tabor, S. and Richardson, C.C. (1987) Proc. Natl. Acad. Sci. USA 84, 4767–4771.
[15] Lion, T. and Haas, O.A. (1990) Anal. Biochem. 188, 335–337.
[16] Kozak, M. (1989) J. Cell. Biol. 108, 229–241.
[17] Wieczorek, H., Wolfersberger, M.G., Cioffi, M. and Harvey, W.R. (1986) Biochim. Biophys. Acta 857, 271–281.
[18] Zimniak, L., Dittrich, P., Gogarten, J.P., Kibak, H. and Taiz, L. (1988) J. Biol. Chem. 263, 9102–9112.
[19] Hirata, R., Ohsumi, Y., Nakano, A., Kawasaki, H., Suzuki, K. and Anraku, Y. (1990) J. Biol. Chem. 265, 6726–6733.
[20] Bowman, E.J., Tenney, K. and Bowman, B.J. (1988) J. Biol. Chem. 263, 13994–14001.
[21] Wieczorek, H., Cioffi, M., Klein, U., Harvey, W.R., Schweikl, H. and Wolfersberger, M.G. (1990) Methods Enzymol. 192, 608–616.
[22] von Heijne, G. (1987) Sequence Analysis in Molecular Biology, Academic Press.
[23] Starke, T., Linkkila, T.P. and Gogarten, J.P. (1991) Z. Naturforsch. C 46, 613–620.