Isolation and sequence of a cDNA clone encoding the 31-kDa subunit of bovine kidney vacuolar H⁺-ATPase

(proton pump/hydrogen ion transport/membrane protein structure/receptor-mediated endocytosis)

SHELDON HIRSCH*, ARNOLD STRAUSS^{†‡}, KHALID MASOOD*, STEPHEN LEE*, VIKAS SUKHATME[§],
AND STEPHEN GLUCK*^{¶||}

Departments of *Medicine, [†]Cell Biology and Physiology, [‡]Biological Chemistry, and [§]Pediatrics, Washington University School of Medicine, and Jewish Hospital, St. Louis, MO 63110; and [¶]Department of Medicine and The Howard Hughes Medical Institute, The University of Chicago Pritzker School of Medicine, Chicago, IL 60637

Communicated by Stuart Kornfeld, December 21, 1987 (received for review November 2, 1987)

ABSTRACT We have isolated a cDNA encoding the 31-kDa subunit of the bovine kidney vacuolar H⁺-ATPase. The composite sequence contains 1219 base pairs, which includes the entire 678-base-pair coding region. A lysine-rich sequence previously found in the Na⁺,K⁺-ATPase α subunit and the H⁺,K⁺-ATPase was identified in the 31-kDa subunit. An RNA blot and an immunoblot demonstrated variable 31-kDa subunit expression and immunoreactivity in different tissues; the highest levels were observed in kidney medulla and brain with both types of analysis. The isolation of a cDNA for the 31-kDa subunit is an important step in understanding this subunit's role in the function and regulation of the vacuolar H⁺-ATPase.

Proton-translocating ATPases acidify a variety of intracellular compartments in the vacuolar system (1). In the specialized acid-secreting kidney intercalated cell, the vacuolar H⁺-ATPase (proton pump) is also found in high concentrations on the plasma membrane, where its polarized distribution effects transepithelial proton secretion (2). Recently, this laboratory reported the isolation and functional reconstitution of the vacuolar H⁺-ATPase from bovine kidney microsomes (3, 4). The intact purified enzyme has a molecular mass of $\approx$ 580 kDa and has over 10 component subunits [i.e., one each of 70, 45, 42, 38, 33, 31, 15, 14, and 12 kDa and several of $\approx$ 56 kDa (3, 4)].

Vacuolar proton pumps partially purified from *Neurospora* vacuoles (5), plant tonoplast (6, 7), chromaffin granules (8), and coated vesicles (9) also are large molecular mass enzymes with subunits of $\approx$ 70, $\approx$ 56, $\approx$ 17 and, generally, $\approx$ 30 kDa (5, 6, 8, 9). Several features suggest that the complex structure of these enzymes and of the kidney H⁺-ATPase is similar to that of F₀F₁ H⁺-ATPases, such as the mitochondrial H⁺-ATPase. Rapid-freeze and thin-section electron micrographs of the vacuolar H⁺-ATPase show a 9.5-nm cytoplasmic domain ("stud") bound to the membrane (10), which is morphologically similar to F₀F₁ ATPase structures. Antibodies specific for the 70-, 56-, or 31-kDa kidney H⁺-ATPase subunits all localize to this cytoplasmic domain (37). The $\approx$ 70- (5) and $\approx$ 56-kDa (7) subunits have been implicated in the enzyme's catalytic activity, whereas the dicyclohexylcarbodiimide-binding $\approx$ 17-kDa subunit (5, 7) may form part of an integral membrane channel.

The cytosolic domains of several membrane proteins are necessary for sorting and control of removal by endocytosis (11, 12). The H⁺-ATPase's cytoplasmic domain might, in addition to having a role in catalytic activity, function in enzyme regulation or sorting to different membrane com-

partments. If sorting to different compartments requires diversity in proton pump structure, having isoforms or isoenzymes of H⁺-ATPase subunits is one possible mechanism for generating such diversity.

In this paper, we demonstrate a wide variation in the vacuolar H⁺-ATPase 31-kDa subunit's immunoreactivity in different tissues. As an initial step toward examining the possibility that variations in this subunit's expression or structure may be important in enzyme regulation or sorting, we report the isolation and sequence of a cDNA encoding this subunit.**

MATERIALS AND METHODS

Preparation and Screening of cDNA Library. RNA was isolated from bovine kidney medulla as described (13), and the poly(A) fraction was obtained by affinity chromatography on oligo(dT)-cellulose (14). The purified mRNA was used to prepare double-stranded cDNA as described (15). *Eco*RI linkers were added to the cDNA inserts, which were then ligated into *Eco*RI-digested λ gt11 and packaged *in vitro* (15).

The cDNA library was screened initially with rabbit polyclonal antiserum to bovine kidney H⁺-ATPase, as previously described (15). Approximately 300,000 recombinant phage plaques were analyzed, and three positive clones were isolated. The largest clone, C1 (1.1 kilobases), was labeled to high specific activity with [³²P]dCTP (16) and was used to reanalyze the kidney cDNA library by plaque hybridization (17).

Preparation of Fusion Protein. *Escherichia coli* strain Y1089 was infected with recombinant λ DNA from clones identified by antibody screening, and lysogenic colonies were selected (15). Cell lysates prepared from the recombinant lysogenic colonies or from a control lysogenic colony (15) were evaluated by immunoblotting as described previously (4, 37) by using preimmune rabbit serum, antiserum to the kidney proton pump, and antiserum monospecific for the 70-, 56-, or 31-kDa proton pump subunits.

Sequencing. Clones C1, R1, and R3 were cloned into the plasmid vector pGEM blue (Promega Biotec, Madison, WI), and the nucleotide sequence was determined by the dideoxy method of Sanger (18). The sequence of both C1 strands was determined in its entirety by using SP6 and T7 primers (Promega Biotec) to initiate sequencing and synthetic oligonucleotide primers to continue to the 3' ends (Fig. 1). Partial sequences of R1 and R3 were also obtained in this manner (Fig. 1).

^{||}To whom reprint requests should be addressed.

^{**}The sequence reported in this paper is being deposited in the EMBL/GenBank data base (Bolt, Beranek, and Newman Laboratories, Cambridge, MA, and Eur. Mol. Biol. Lab., Heidelberg) (accession no. J03244).

FIG. 1. Partial restriction map and sequencing strategy of cDNA clones C1, R1, and R3. The map is drawn to scale showing the 5' to 3' orientation of the mRNA strand. Closed circles represent SP6 or T7 primers; open circles denote synthetic oligonucleotide primers. Arrows designate direction and extent of sequence obtained from each primer. The asterisk indicates the location of the stop codon, and the hatched bar represents the coding region of the 31-kDa subunit mRNA. A partial restriction map of the composite cDNA is shown.

Expression of 31-kDa Protein. An *Nco*I–*Eco*RI restriction fragment of clone R1, which begins at the initiator methionine, was subcloned into the pAS (19) derivative pOTSN-co12 (obtained from A. Shatzman, Smith Kline & French). This plasmid has a λ *P1* promoter with an initiator methionine integrated into an *Nco*I site. The recombinant plasmid was constructed so that the R1 start codon replaced the start codon for the *P1* transcription unit and the entire R1 coding sequence could be transcribed. The recombinant plasmid was then used to transform competent *E. coli* strain N5151 (obtained from A. Shatzman), a λ lysogen with a temperature-sensitive mutation in the *CI* gene (19). Cell extracts for immunoblotting were obtained from the resulting strain after thermal induction at 42°C for 2 hr.

Immunoblots. Polypeptides were separated by NaDodSO₄/PAGE, transferred to nitrocellulose paper (20), and analyzed with polyclonal antiserum as described (4), except that 5% ovalbumin replaced bovine serum albumin in the blocking solution; bound antibody was detected by color development (21) after incubation with alkaline phosphatase-labeled goat anti-rabbit IgG (Southern Biotechnology, Birmingham, AL).

Tryptic Peptides. The 31-kDa subunit was excised from NaDodSO₄ gels, eluted, and dialyzed in an electrophoresis chamber (Schleicher & Schuell). The recovered protein (≈ 100 μ g) was digested with trypsin, and the tryptic peptides were separated by reverse-phase HPLC on a C₈ microbore column (Applied Biosystems, Foster City, CA) with a linear 0–70% acetonitrile gradient (0.5%/min) in 0.1% trifluoroacetic acid. Isolated peptides (≈ 75 pmol) were analyzed to completion by automated Edman degradation with a model 470A gas phase protein sequencer (Applied Biosystems) with on-line phenylthiohydantoin analysis (model 120A, Applied Biosystems).

RNA Blots. Total RNA was isolated from bovine tissues (13). Ten micrograms of RNA from each tissue was separated electrophoretically on a 1.5% agarose/formaldehyde gel and was transferred to a GeneScreen membrane (New England Nuclear). Hybridization was performed according to the manufacturer's instructions. The C1 insert was labeled (16) with [³²P]dCTP and used as a probe (6.0×10^5 cpm/ml). Final washing was with 0.3 M NaCl/0.03 M sodium citrate, pH 7.0/0.1% NaDodSO₄ at 50°C.

RESULTS

Approximately 300,000 recombinant clones from a bovine kidney medulla cDNA library were analyzed with rabbit polyclonal antiserum to the bovine kidney H⁺-ATPase. Three positive clones (C1, C4, and C6) were identified, and

pure plaques were isolated. Bacterial lysates prepared from clones C1 and C4 exhibited an ≈ 153 -kDa fusion protein on NaDodSO₄/PAGE not present in control lysate, which was recognized on immunoblot analysis by antibodies specific for the 31-kDa subunit (data not shown).

The nucleotide sequence of both C1 strands was determined, in its entirety, in the vector pGEM blue (Fig. 1). The amino acid sequences of four tryptic peptides generated from the purified 31-kDa subunit protein matched predicted sequences from C1, confirming that the clone encoded this subunit (Fig. 2). The sequence of C1 ended with a poly(A) tail, but it did not appear long enough [1073 base pairs (bp)] to comprise a full-length clone. C1 was then used to reanalyze the kidney library, and the hybridizing clones R1 and R3 were isolated. Sequence from R1 (1213 bp) added 75 bp of 5'-untranslated region and the first 65 nucleotides of the coding region to the C1 sequence. Additional partial sequence of R1 was identical to the corresponding C1 sequence. Partial sequence at the 3' end of R3, however, differed from the C1 sequence in that R3 had 6 additional nucleotides (AGGTGG) immediately prior to its poly(A) tail. This demonstrates that the poly(A) tail may be added at more than one site, as has been described previously in other gene transcripts (22).

The composite sequence (Fig. 2) contains 1219 bp, which includes the entire coding region. Although the putative 75-bp 5'-untranslated region does not contain a stop codon, several factors lend confidence that our choice of the initiator codon is correct. It is the first methionine, and it follows a G+C-rich (74%) region similar in G+C content to the 5'-untranslated regions of the Ca²⁺-ATPase (85%) (23) and the Na⁺, K⁺-ATPase α (72%) and β (74%) subunits (24, 25). It falls within a perfect nine nucleotide (CCGCCATGG) Kozak consensus sequence for initiation (26). The composite cDNA length (1219 bp) is close to the apparent mRNA length (1400 bp) on an RNA blot (see Fig. 6). Finally, when expressed in a bacterial system, the R1 coding region produced a polypeptide of 31 kDa, a size identical to that of the native subunit on NaDodSO₄/PAGE (Fig. 3).

Following the designated initiator codon, the open reading frame continues for a total of 678 nucleotides, encoding 226 amino acids. The mature protein has a predicted molecular mass of 26,139 Da. The larger apparent size on NaDodSO₄/PAGE (Fig. 3) is likely secondary to the polypeptide's extremely hydrophilic nature (see below).

The 3'-untranslated region contains 466 bp and ends with 16 adenosines. Starting 23 bp before the poly(A) tail in R3 is the sequence ATTAAA, a previously described (27) variant of the consensus polyadenylation signal (AATAAA) (28). The sequence ACACAACG follows immediately and is highly complementary to the downstream genomic consensus sequence YGTGTTY (where Y = pYrimidine) (29), which may be important for efficient cleavage and polyadenylation (29).

Hydropathy analysis using the program of Kyte and Doolittle (30) (Fig. 4) predicts that the protein is predominantly hydrophilic; none of the hydrophobic stretches are long enough to comprise a membrane-spanning α -helix. A Chou and Fasman analysis (31) (Fig. 4) predicts that the N-terminal half of the protein contains mostly α -helices. The C-terminal half is composed of alternating α -helical, β -sheet, and reverse turn structures, a structural motif common to ATP and dinucleotide binding domains (38).

A computer-aided search of the National Biomedical Research Foundation protein sequence data base^{††} using the FASTP program of Lipman and Pearson (32) did not identify long regions of sequence similarity between the 31-kDa

^{††}Protein Identification Resource (1987) Protein Sequence Database (Natl. Biomed. Res. Found., Washington, DC), Release 13.0.

```

      -66      -56      -46      -36      -26      -16      -6      5
CGCCGCGGAAGCTCCCGCCTCTCTGTAGGTCTGACCTCCGCGCGATCGTCCGCGCGCACGGCTTGCCGCGATGCG
      H   A
      15      25      35      45      55      65      75      85
CCTCAGCGATGCCGACGTGCAGAAAGCAGATTAAGCACATGATGGCTTTATTGAACAAGAAGCCAATGAGAAGCAGAAG
L S D A D V Q K Q I K H M H A F I E Q E A N E K A E
      95      105      115      125      135      145      155      165
AAATGTGATGCAAAGGCAGAAGAAGATTCAACATTGAGAAAGGTCCTCTGTCCAAACCCAAAGACTGAAGATTATGGAA
E I D A K A E E E F N I E K G R L V Q T Q R L K I M E
      175      185      195      205      215      225      235      245
TATTATGAGAAGAAAGAAAAGCAGATTGAGCAGCAGAAGAAAATTGATGTCCAATTGATGAATCAAGCGAGGCTCAA
Y Y E K K E K Q I E Q Q K K I Q H S N L M N Q A R L K
      255      265      275      285      295      305      315      325
AGTCCTCAGAGCGAGAGATGACCTTATCAGACCTACTAAATGAAGCAAAACAGAGACTCAGCAAAGTGGTAAAAGATA
V L R A R D D L I T D L L N E A K Q R L S K V V K D
      335      345      355      365      375      385      395      405
CAACCAGGTACCAAGTGTCTGCTGGATGGAAGTCTCCAGGTTTGTACCAAGTGTGAGAGCCCGGATGATCGTTCGC
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
      415      425      435      445      455      465      475      485
TGCAGAAACAGGATTTCCCTCTGGTGAAGGTCGAGTGCAAAAAGCAATCCCTGTGTACAAAGTTGCAACCAAAAGAGA
C R K Q D F P L V K A A V Q K A I P V Y K V A T K R D
      495      505      515      525      535      545      555      565
CGTTGATGTCCAGATTGACCAGGAGGCCTACCTGCCGAGGAGATAGCCGGCGGTGTCGAGATCTATAACGGGGACCGCA
V D V Q I D Q E A Y L P E E I A G G V E I Y N G D R
      575      585      595      605      615      625      635      645
AGATCAAGGTGTCCAACCACTCGAAAGTCGGCTGGACCTCATAGCCCAGCAGATGATGCCCGAAGTCCGGGGAGCCTTG
K I K V S N T L E S R L D L I A Q Q H M P E V R G A L
      655      665      675      685      695      705      715      725
TTTGGTGCAAAATGCCAACAGGAAGTTTGGACTAAGCCAGCAGGAGGTGCGTGACGAGGACGCTCCCGTCATTGGAAGA
F G A N A N R K F L D *
      735      745      755      765      775      785      795      805
AACGAGATGGCCTGTATAGCTTCTCTTTGCTGTCCCTGCCCTTATTGTTCAGTGGTGCTCTTCTGTAGCTAATGTTT
      815      825      835      845      855      865      875      885
TCCGACTGATGGGTAATGATGGGAGTCTCAGTTAGTACATTAGGAAACCTTCTCTAGCTTGTCTCAAAGTAGCACAGTT
      895      905      915      925      935      945      955      965
CCCGCGGTTCTGTAGCGTGCAAGTTGCGGGGCCAGGTGTGTGCGTGTCCGCGACCTGCTGCTGCTGGGCTCCACCACTC
      975      985      995      1005      1015      1025      1035      1045
CAGCCTACAGGCGGTGGGGTGGTAGACTCAGTCTCAAGTGCAGTGTGGGCTCATCCCTTCCGGCAGTAGCAGCTTCTC
      1055      1065      1075      1085      1095      1105      1115      1125
GCCTACGTTTGGACCTAGAAGTGTCTGTAATGTGATTAAAGTCTATTCCATGAATCTGCTCTATTATTAAAAACAC
      1135      1145
AACGTGGAGGTGGAAGAAAA

```

FIG. 2. Nucleotide sequence and deduced amino acid sequence of the 31-kDa subunit. Nucleotides are numbered positively in the 5' to 3' direction beginning with the initiator methionine and are numbered negatively in the 3' to 5' direction in the 5'-untranslated region. The predicted amino acid sequence is shown below the nucleotide sequence in the one-letter amino acid code. The start codon is boxed, and an asterisk denotes the stop codon. Sequences confirmed by automated Edman degradation of tryptic peptides (TP) are underlined. Only 6 of 16 adenosines of the poly(A) tail are shown.

sequence and any other reported sequence. However, the search did reveal that an 11 amino acid lysine-rich region of the 31-kDa subunit sequence is similar to sequences in the N-terminal hydrophilic domains of the Na^+ , K^+ -ATPase α subunit (24) and the H^+ , K^+ -ATPase (33) (Fig. 5). A similar sequence is also found in the rod portion of chicken embryonic muscle myosin heavy chain (34). The two pairs of lysines are present in the Na^+ , K^+ -ATPase β -subunit N-terminal domain (25), although the remaining sequence is not well conserved. Another sequence similar to a segment of the myosin heavy-chain rod domain (34) and a sequence similar to part of the Ca^{2+} -ATPase phosphorylation domain (23) were also found (Fig. 5).

The distribution and expression of the protein and mRNA were evaluated, respectively, with immunoblots of rat tissues and RNA blots of bovine tissues. Immunoreactivity to

the 70-, 56-, and 31-kDa subunits was demonstrated in every tissue tested (Fig. 6). The immunoreactivity of the 31-kDa subunit was highly variable, however, in marked contrast to the uniform reactivity of the 70- and 56-kDa subunits. High levels of 31-kDa subunit immunoreactivity were found only in kidney and brain (Fig. 6). Immunoblots of kidney cortex and medulla demonstrated substantially greater 31-kDa subunit immunoreactivity in the medulla (data not shown). The RNA blot (Fig. 7) revealed that the mRNA encoding the 31-kDa subunit was present in all tissues examined, with highest levels in kidney medulla and brain, in agreement with the immunoblot results.

DISCUSSION

We have isolated and determined the nucleotide sequence of a cDNA encoding the bovine kidney medulla vacuolar

FIG. 3. Expression of the 31-kDa protein in *E. coli*. An *Nco*I-*Eco*RI restriction fragment of clone R1 that begins at the start codon and ends with the poly(A) tail was ligated, in reading frame, to the plasmid vector pOTSNco12, which was then used to transform *E. coli* strain N5151. Cell lysates were extracted from the transformed bacteria and from nontransformed control *E. coli* strain N5151. (A) Cell lysate (25 μ g) from control (lane 1) or transformed *E. coli* (lane 2) were separated on a 12.5% NaDodSO₄ gel. Expression of a 31-kDa polypeptide was apparent in the transformed bacteria (arrowhead). (B) Immunoblot of lysates from control (lane 1) or transformed (lane 2) bacteria using monospecific anti-31-kDa subunit antiserum. An immunoreactive polypeptide at 31 kDa was observed in the lysate from transformed cells (arrowhead).

proton pump 31-kDa subunit. Although vacuolar proton pumps present in a variety of organelles are thought to be similar (1), their membrane distribution and density as well as their biochemical properties vary among cell types and among vacuolar compartments in a cell (1). In most cells, proton pumps are found largely in intracellular membrane compartments of the vacuolar system (1). However, in proton-transporting kidney epithelia, the enzyme is polarized in high concentrations to the luminal plasma membrane of some cells and to the basolateral plasma membrane or intracellular vacuolar membranes in other cells (37). The factors responsible for variations in targeting, membrane density, and physiologic properties of pumps are largely unknown, but one possibility is that these differences reflect structural heterogeneity of the enzyme. The demonstration of striking differences in the immunoreactivity of the 31-kDa subunit in different tissues, in contrast to the reasonably uniform immunoreactivity of the 70- and 56-kDa subunits, raises the possibility that there may be isoenzymes of the 31-kDa subunit. In support of this, a second transcript hybridizing to the C1 clone was detected on low-stringency

FIG. 4. Analysis of hydropathy and secondary structure. The hydropathy plot is based on the method of Kyte and Doolittle (30). The search length was six amino acids. Hydrophobicity increases with distance above the abscissa. A plot of predicted secondary structure obtained using the program of Chou and Fasman (31) is shown below the hydropathy plot.

FIG. 5. Regions of similarity of the 31-kDa subunit sequence to those of other ATPases. (A) Comparison of amino acid sequences from sheep kidney Na⁺,K⁺-ATPase α subunit (residues 20–30) and β subunit (residues 12–21) (24, 25), rat stomach H⁺,K⁺-ATPase residues (37–46) (33), and chicken embryonic skeletal muscle myosin heavy chain (residues 691–701) (34, 35) with sequence from the 31-kDa subunit (residues 59–69). (B) Sequence comparison of chicken muscle myosin heavy chain (residues 741–747) (34, 35) with the 31-kDa subunit (residues 25–31). (C) Sequence comparison of rabbit muscle sarcoplasmic reticulum Ca²⁺-ATPase (residues 428–437) (23) with the 31-kDa subunit (residues 96–105). Residues identical to the corresponding amino acid of the 31-kDa subunit are boxed; related amino acids (36) are enclosed by dashed lines.

RNA blots (S.H., A.S., and S.G., unpublished data). Whether isoenzymes of this subunit have a role in H⁺-ATPase sorting or regulation will need to be addressed in future studies.

The vacuolar H⁺-ATPase has a large cytosolic stalk-sphere domain that has recently been visualized in toad bladder and in the reconstituted, purified enzyme by a high-resolution rapid-freeze platinum shadowing technique and appears to consist of at least six subunits (10). Immunogold labeling of kidney epithelial cells with polyclonal antibodies specific to the 31-kDa subunit showed localization of the subunit to the enzyme's cytoplasmic domain (37). Consistent with this location, the 31-kDa sequence predicts a largely hydrophilic protein without any membrane-spanning domain.

FIG. 6. Immunoblot of rat tissues. Fifty micrograms of microsomal protein from each rat tissue and from bovine kidney was separated on a 12.5% NaDodSO₄ gel, transferred to nitrocellulose, and analyzed with rabbit polyclonal antiserum to the bovine kidney proton pump. Arrows denote the location of subunits at 70-, 56-, and 31-kDa. Enhanced immunoreactivity of the 31-kDa polypeptide was noted in kidney and brain (arrowheads).

FIG. 7. RNA blot from bovine tissues. Ten micrograms of total RNA from each bovine tissue was separated on a 1.5% agarose/formaldehyde gel, transferred to a GeneScreen membrane, and hybridized to ^{32}P -labeled clone C1. An ≈ 1.4 -kilobase transcript was present in all tissues and was accentuated in brain and kidney medulla. Locations of size markers and RNA are shown.

Subunits of the H^+ -ATPase's cytoplasmic head might be envisaged as having, minimally, two functional domains: one region that interacts with other H^+ -ATPase subunits and another region that is exposed to the cytosol. The predicted secondary structure of the 31-kDa subunit suggests two domains—a highly hydrophilic N-terminal half, which is $>90\%$ α -helix, and a C-terminal half with alternating secondary structure. The hydrophilic N-terminal half may represent a cytosol-facing domain, a possibility that needs to be examined by direct methods.

Regions of sequence similarity with several ATPases were present in the N-terminal half of the 31-kDa subunit. Of particular note is a lysine-rich sequence similar to highly conserved N-terminal lysine-rich sequences previously found in non-ATP-binding domains of Na^+ , K^+ -ATPase α subunit (24) and the H^+ , K^+ -ATPase (33) (Fig. 5). Whereas alignment of any two proteins may demonstrate "chance similarity," the presence of this lysine-rich sequence in several ATPases suggests functional significance. This highly charged region has been postulated to control access of Na^+ and H^+ to cytoplasmic cation-binding sites (24, 33). This domain in the 31-kDa subunit might control access of protons to the H^+ -ATPase's translocating site. Of note, however, is a similar lysine-rich sequence in chicken embryonic muscle myosin heavy chain (34) (Fig. 5). This 10 amino acid sequence, present in the rod portion [which is not involved in either ATP hydrolysis or actin binding (34)], has three lysines at its N terminus, a fourth at the C terminus, and an intervening sequence identical in four of five positions to that of the 31-kDa subunit. If this myosin domain functions in the same manner as the lysine-rich areas in the other ATPases, a function alternative to regulating cytoplasmic cation binding is likely.

The vacuolar H^+ -ATPase, with its multisubunit structure and access to many membrane compartments, is an attractive protein for investigating the role of the cytoplasmic

domain in membrane protein sorting. An approach using specific probes for examining the cellular distribution of individual subunits and putative isoenzymes of the H^+ -ATPase may lead to new insights in this area.

We thank Harold Sims, Daniel Kelly, and Thomas Chu for helpful discussions and advice; Lisa Krebs for technical assistance; Mark Frazier and Greg Grant for their expertise in protein sequencing; Drs. A. Shatzman, M. Rosenberg, and C. Debouk (Smith Kline & French) for providing pOTSNco12 vector and *E. coli* strain N5151; and Betty Ytzaina for expert secretarial assistance. This work was supported by a fellowship from the National Kidney Foundation (S.H.), National Institutes of Health Grant DK38848 (S.G.), and the Searle Scholars Program (S.G.).

1. Mellman, I., Fuchs, R. & Helenius A. (1986) *Annu. Rev. Biochem.* **55**, 663–700.
2. Steinmetz, P. R. (1986) *Am. J. Physiol.* **251**, F173–F187.
3. Gluck, S. & Caldwell, J. (1987) *J. Biol. Chem.* **262**, 15780–15789.
4. Gluck, S. & Caldwell, J. (1988) *Am. J. Phys.* **254**, F71–F79.
5. Bowman, E. J., Mandala, S., Taiz, L. & Bowman, B. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 48–52.
6. Mandala, S. & Taiz, L. (1985) *Plant Physiol.* **78**, 327–333.
7. Manolson, M. F., Rea, P. A. & Poole, R. J. (1985) *J. Biol. Chem.* **260**, 12273–12279.
8. Percy, J., Pryde, J. & Apps, D. (1985) *Biochem. J.* **231**, 557–564.
9. Xie, X. S. & Stone, D. (1986) *J. Biol. Chem.* **261**, 2492–2495.
10. Brown, D., Gluck, S. & Hartwig, J. H. (1987) *J. Cell Biol.* **105**, 1637–1648.
11. Lehrman, M. A., Goldstein, J. L., Brown, M. S., Russell, D. W. & Schneider, W. J. (1985) *Cell* **41**, 735–743.
12. Mostov, K. E., Bruyn Kops, A. & Deitcher, D. L. (1986) *Cell* **47**, 359–364.
13. Chirgwin, J. M., Przybyla, A. E., Macdonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
14. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1408–1412.
15. Huynh, T. V., Young, R. A. & Davis, R. W. (1985) in *DNA Cloning*, ed. Glover, D. M. (IRL, Oxford), Vol. 1, pp. 49–78.
16. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **123**, 6–13.
17. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
18. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
19. Shatzman, A. R. & Rosenberg, M. (1986) *Ann. N.Y. Acad. Sci.* **478**, 233–247.
20. Towbin, M., Staehelin, T. & Gordon, J. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4350–4354.
21. Blake, M. S., Johnston, K. H., Russell-Jones, G. J. & Gotschlich, E. C. (1984) *Anal. Biochem.* **136**, 175–179.
22. Biladello, J. J., Kelly, D. P., Roman, D. G. & Strauss, A. W. (1986) *Biochem. Biophys. Res. Commun.* **138**, 392–398.
23. MacLennan, D. H., Brandl, C. J., Korczak, B. & Green, N. M. (1985) *Nature (London)* **316**, 696–700.
24. Shull, G. E., Schwartz, A. & Lingrel, J. B. (1985) *Nature (London)* **316**, 691–695.
25. Shull, G. E., Lane, L. K. & Lingrel, J. B. (1986) *Nature (London)* **321**, 429–443.
26. Kozak, M. (1984) *Nucleic Acids Res.* **12**, 857–872.
27. Rosenfeld, M. G., Amara, S. G. & Evans, R. M. (1984) *Science* **225**, 1315–1320.
28. Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263**, 211–214.
29. McLauchlan, J., Gaffney, D., Whittan, J. L. & Clements, J. B. (1985) *Nucleic Acids Res.* **13**, 1347–1369.
30. Kyte, J. S. & Doolittle, R. F. (1983) *J. Mol. Biol.* **157**, 105–132.
31. Chou, P. Y. & Fasman, G. D. (1983) *J. Mol. Biol.* **47**, 251–276.
32. Lipman, D. J. & Pearson, W. R. (1985) *Science* **227**, 1435–1441.
33. Shull, G. E. & Lingrel, J. B. (1986) *J. Biol. Chem.* **261**, 16788–16791.
34. Kavinsky, C. J., Umeda, P. K., Sinha, A. M., Elzinga, M., Tong, S. W., Zak, R., Jakovcic, S. & Rabinowitz, M. (1983) *J. Biol. Chem.* **258**, 5196–5205.
35. McLachlan, A. D. & Karn, J. (1982) *Nature (London)* **299**, 226–231.
36. Robinson, A. C., Collins, J. F. & Donachie, W. D. (1987) *Nature (London)* **328**, 766.
37. Brown, D., Hirsch, S. & Gluck, S. (1988) *Nature (London)* **331**, 622–624.
38. Watson, J. D., Hopkins, N. M., Roberts, J. W., Steitz, J. A. & Weiner, A. M. (1987) *Molecular Biology of the Gene* (Benjamin/Cummings, Menlo Park, CA), 4th Ed., pp. 144–149.