We have identified cDNAs encoding three isoforms (a1, a2, and a3) of the 100-kDa a subunit of the mouse vacuolar proton-translocating ATPase (V-ATPase). The predicted protein sequences of the three isoforms are 838, 856, and 834 amino acids, respectively, and they display approximately 50% identity between isoforms. Northern blot analysis demonstrated that all three isoforms are expressed in most tissues examined. However, the a1 isoform is expressed most heavily in brain and heart, a2 in liver and kidney, and a3 in liver, lung, heart, brain, spleen, and kidney. We also identified multiple alternatively spliced variants for each isoform. Reverse transcriptase-mediated polymerase chain reaction revealed that one splicing variant of the a1 isoform (a1-I) was expressed only in brain, whereas two other variants (a1-II and a1-III) were expressed in tissues other than brain. These alternatively spliced forms differ in the presence or absence of 6–7 amino acid residues near the amino and carboxyl termini of the proteins encoded. The a3 isoform is also encoded by three alternatively spliced variants, two of which are predicted to encode a protein that is truncated near the border of the amino- and carboxyl-terminal domains of the a subunit and therefore lacks the integral transmembrane-spanning helices thought to participate in proton translocation. Expression of each isoform (with the exception of a1-I) was detectable at all developmental stages investigated, with a1-I absent only in day 7 embryos. The results obtained suggest that isoforms of the 100-kDa a subunit may contribute to tissue-specific functions of the V-ATPase.

The vacuolar (H+)-ATPases (or V-ATPases)1 function as ATP-dependent proton pumps to acidify intracellular compartments in eukaryotic cells. V-ATPases are present in a variety of intracellular compartments, including clathrin-coated vesicles, endosomes, lysosomes, Golgi-derived vesicles, chromatin granules, synaptic vesicles, and the central vacuoles of yeast. From the Department of Physiology, Tufts University School of Medicine, Boston, Massachusetts 02111

Molecular Cloning and Expression of Three Isoforms of the 100-kDa a Subunit of the Mouse Vacuolar Proton-translocating ATPase*

Tsuyoshi Nishi and Michael Forgac‡

From the Department of Physiology, Tufts University School of Medicine, Boston, Massachusetts 02111

1 The abbreviations used are: V-ATPase, vacuolar proton-translocating ATPase; EST, expression sequence tag; RT-PCR, reverse transcriptase-mediated polymerase chain reaction; RACE, rapid amplification of cDNA ends; bp, base pairs.
from Sigma.

Isolation of cDNA Clones—Clones for each of the three a subunit isoforms were identified in the expression sequence tag (EST) data base and obtained from the American Tissue Culture Collection (ATCC) (Manassas, VA). Clones for the a1, a2, and a3 isoforms are A444764, A444765, and A444415, respectively. The EST A444415 corresponds to the alternatively spliced variant of a3 that is predicted to encode a truncated form of the a subunit. Based on these sequences, we designed the primers to carry out 5′- and 3′-RACE in order to obtain full-length cDNAs for each isoform.

Identification of the 5′-End of a Subunit cDNAs—To determine the transcription initiation sites for each mRNA, we performed the 5′-RACE using the SMART cDNA amplification kit (CLONTECH, Palo Alto, CA). This system allows us to identify the complete 5′-end sequence of the cDNA. Reactions were performed using the manufacturer’s recommended protocol and mouse brain poly(A) mRNA (Ambion, TX). Amplification of fragments was performed using Pu Polymerase (Stratagene, CA), isoform-specific primers, and 10× Universal primer mix (CLONTECH). The resulting polymerase chain reaction products were subcloned into the pCR-Blunt-II-TOPO vector (Invitrogen) and sequenced.

Northern Blot Analysis—Mouse multiple tissue Northern blot membranes were purchased from CLONTECH. Isoform-specific DNA fragments were prepared from ATCC and sequenced. Sequence analysis revealed that the a1, a2, and a3 isoforms exist in mice, and obtained from the American Tissue Culture Collection (ATCC).

RESULTS

Isolation and Sequencing of cDNAs Encoding Isoforms of the V-ATPase a Subunit in Mice—A search of the data base of ESTs from mice revealed three different groups of sequences for the a subunit. Several EST clones from each group were obtained from ATCC and sequenced. Sequence analysis revealed that each EST clone corresponded to a fragment of the a subunit, 5′- and 3′-RACE was performed using mouse brain cDNA template in order to obtain full-length sequences for each of the a subunit isoforms. Fig. 1 shows the deduced amino acid sequence for each of the three a subunit isoforms identified (a1, a2, and a3) and the alignment with the two a subunit isoforms (Vph1p and Stv1p) from yeast. The cDNAs of a1, a2, and a3 are predicted to encode proteins of 838, 856, and 834 residues, respectively. The sequence identity (and similarity) between pairs of mouse a subunit isoforms was as follows: a1 and a2, 52% (69%); a1 and a3, 47% (66%); a2 and a3, 51% (67%). We have previously identified a number of residues of Vph1p that are important for activity or assembly of the V-ATPase complex in yeast (30, 31). All of the residues identified are conserved in all three mouse isoforms, with the exception of H729 (the Vph1p numbering), which is an asparagine residue in a1, a2, and a3. Construction of a phylogenetic tree using the mouse and yeast sequences (Fig. 2) reveals that the development of multiple isoforms of the a subunit appears to have occurred independently in yeast and mammals, so that none of the isoforms in mice is more closely related to either Vph1p or Stv1p in yeast. Homology between the mouse and yeast isoforms ranges from 33 to 39%.

Determination of the 5′-End of a Subunit Isoform cDNAs—To determine the transcription initiation site for each of the a subunit isoforms, the 5′-end of the mRNA for each isoform was determined by 5′-RACE analysis (Fig. 3). For both a2 and a3, two initiation sites were identified, with the major site occurring at −98 bp. By contrast, eight different initiation sites were identified for the a1 isoform, but none of them represented a dominant site.

Tissue Distribution of a Subunit Isoforms—To determine the tissue distribution of expression of the three a subunit isoforms, Northern blot analysis was performed on mRNA isolated from various mouse tissues using RNA probes specific for each isoform. As can be seen in Fig. 4a, the a1 isoform was expressed to some degree in all tissues, but mRNA levels for a1 were highest in brain and heart. The a2 isoform mRNA was detected in all tissues except skeletal muscle but was most abundant in liver. Similarly, mRNA for the a3 isoform was also detected in all tissues except skeletal muscle, with very prominent expression in liver. There were also significant levels of a3 mRNA detected in lung, heart, brain, spleen, and kidney. Thus, the three a subunit isoforms were expressed in a tissue-specific manner. In addition, Northern blot analysis revealed three bands that hybridized with the a2-specific probe and two bands that hybridized with the a3-specific probe. These results suggested that alternatively spliced forms of both the a2 and a3 isoforms existed in mice.

Southern Blot Analysis Using a Subunit Isoform-specific Probes—To test whether each of the a subunit isoforms was encoded by a single gene in mice, Southern blot analysis was performed using mouse genomic DNA and the same RNA probes employed in the Northern blot analysis. The results, shown in Fig. 5, reveal only a single band capable of hybridizing with each isoform-specific probe, indicating that each a subunit isoform is encoded by a single gene. Thus, the multiple bands observed by Northern analysis likely correspond to alternatively spliced forms.

Isolation of Alternatively Spliced Variants of the Mouse a1 Subunit Isoform—It has previously been reported that there exist two alternatively spliced forms of the a1 isoform in bovine brain (26). The full-length form is expressed exclusively in brain, whereas a truncated form containing a 6-amino acid deletion near the carboxyl terminus is expressed in all tissues examined, including brain, heart, liver, kidney, and spleen.

In the case of mice, the a1 isoform is encoded by three alternatively spliced variants (Fig. 6). Two of these variants (a1-II and a1-III), both cloned from liver, lack six amino acids near the carboxyl terminus relative to the other variant (a1-I, cloned from brain). The a1-I and a1-III variants, on the other hand, differ from a1-II in that they are missing seven amino acids near the amino terminus. To confirm the presence of these three variants in brain and liver, RT-PCR was performed using primers located on each side of the amino- and carboxyl-terminal deletions. As can be seen in Fig. 6b, brain mRNA gave...
a strong signal at the position corresponding to the amino-terminal deletion and the carboxyl-terminal insertion, indicating a predominance of the a1-I form in brain. By contrast, in liver, weak signals were observed for both the longer and shorter forms using the amino-terminal primers, but only the shorter form was observed using the carboxyl-terminal primers. This suggests that in liver, the a1-II and a1-III isoforms are the predominant species.

Identification of Alternatively Spliced Variants of the a3 Isoform—A search of the EST data base revealed two clones (AA244585 and AA444415) with homology to the a3 isoform cloned from brain. AA244585 is a partial sequence that predicts a protein identical to the a3 sequence beginning at residue 595 and continuing through the carboxyl terminus of the protein. AA444415 is also a partial sequence that predicts a protein identical to the a3 sequence from residue Glu-203 through residue Ser-341 (Fig. 7). After this point an insert is present in AA444415 that codes for a unique sequence of 30 amino acids followed by a stop codon. The predicted protein is therefore truncated after only 371 residues and thus completely lacks the

Expression of Mouse V-ATPase a Subunit Isoforms

**Fig. 1.** Amino acid sequence alignment of the 100-kDa a subunit isoforms in mice and yeast. Alignment of the three mouse a subunit isoform sequences (a1, a2, and a3) obtained in the current report and the two yeast a subunit isoform sequences (Vph1p and Stv1p) previously reported (23, 24) is shown. Identical residues are indicated by the shaded boxes. The nine putative transmembrane helices assigned from topographical analysis of Vph1p (34) are shown with a dark bar. The residues whose mutation has a significant effect on activity or assembly of the V-ATPase (30, 31) are indicated by asterisks. GenBank™ accession numbers assigned to a1, a2, and a3 are AF218249, AF218252, and AF218253, respectively.
hydrophobic carboxyl-terminal domain.

Tissue Variation in Expression of Alternatively Spliced Forms of a1 and a3—To further investigate the tissue distribution of alternatively spliced forms of the a subunit, RT-PCR was performed using isoform-specific primers and mRNA from various mouse tissues (Fig. 8). Using primers that amplify the carboxyl-terminal region of the a1 isoform, it can be seen that a1-I (containing the full carboxyl-terminal sequence) is expressed only in brain, whereas a1-II and a1-III (containing the carboxyl-terminal deletion) are expressed in all tissues except brain. Comparing the products obtained using mRNA from embryos at different developmental stages, it appears that the a1-II and a1-III isoforms are present in all developmental stages tested (7-, 11-, 15-, and 17-day embryos) whereas the a1-I isoform is absent in the 7-day embryos but present at later developmental stages.

Using primers directed against the amino-terminal region of the a2 isoform, a single RT-PCR product is observed in all tissues except skeletal muscle and in all developmental stages tested. Although the presence of multiple bands for a2 on Northern blots (Fig. 4) suggests the existence of multiple splice variants for this isoform as well, these variants must not differ in the 5'-region amplified using the primers employed in RT-PCR.

Using the a3-specific primers shown in Fig. 7, three products were observed for most tissues by RT-PCR (Fig. 8). The lowest band (a3-I) lacks both the insertions shown in Fig. 7 and thus corresponds to the a3 sequence shown in Fig. 7. The upper band (a3-II), which is almost 400 bp larger than a3-I, corresponds to the AA444415 sequence shown in Fig. 7. The upper band (a3-II) which was transferred poly(A) RNA (2 μg) isolated from mouse heart (lane 1), brain (lane 2), spleen (lane 3), lung (lane 4), liver (lane 5), skeletal muscle (lane 6), kidney (lane 7), and testis (lane 8) were hybridized with [32P]dUTP-labeled probes specific for isoform a1 (a), a2 (b), a3 (c), or glyceraldehyde-3-phosphate dehydrogenase (d) followed by washing and visualization as described under “Experimental Procedures.”

DISCUSSION

Previous studies in yeast have demonstrated the existence of two isoforms of the 100-kDa a subunit (Vph1p and Stv1p) that

![Fig. 2. Evolutionary relationship between the a subunit isoforms.](image)

![Fig. 3. Determination of the 5'-end of the mRNA encoding each isoform.](image)

![Fig. 4. Northern blot analysis of mRNA isolated from various mouse tissues using isoform-specific cDNA probes.](image)

![Fig. 5. Southern blot analysis of mouse genomic DNA using isoform-specific probes.](image)
appear to be targeted to different intracellular compartments (24). Sequencing studies in mammals have identified proteins homologous to Vph1p and Stv1p (25–29), but identification of isoforms has been complicated by the fact that the cDNAs cloned have generally been derived from different species. The most complete story thus far has been in bovine tissues, where two isoforms (a1 and a2) have been identified (27) and where a single alternative splicing variant of a1 has been described (26). Northern blotting indicated that the a1 isoform is expressed in brain, heart, kidney, liver, and spleen (26), whereas a2 is expressed mainly in kidney, lung, and spleen but only at low levels in brain and heart (27). An “osteoclast-specific” isoform of the a subunit has also been cloned from humans (28), although its precise relationship to a1 and a2 has been difficult to determine because of species differences.

The work presented in the current paper represents the first cloning and sequencing of three distinct isoforms of the a subunit in a single species. The mouse a1 and a2 isoforms reported are homologous to the a1 and a2 isoforms of bovines (26, 27), whereas the mouse a3 isoform is homologous to the osteoclast-specific isoform previously reported (28). Southern blot analysis indicates that a1, a2, and a3 are encoded by three distinct

![FIG. 6. Sequence and detection of alternatively spliced variants of the a1 isoform. a, amino acid sequence alignment of the alternatively spliced variants of the a1 isoform isolated from brain (a1-I) and liver (a1-II and a1-III). b, confirmation of expression of each isoform by RT-PCR. RT-PCR was performed on poly(A) mRNA isolated from either brain (lanes 1 and 3) or liver (lanes 2 and 4) using the primers indicated. Primers N-Fw and N-Rv were used for lanes 1 and 2 to detect the presence or absence of the amino-terminal insertion, whereas primers C-Fw and C-Rv were used for lanes 3 and 4 to detect the presence or absence of the carboxyl-terminal insertion. GenBankTM accession numbers assigned to a1-I, a1-II, and a1-III are AP219249, AP219250, and AP219251, respectively.](image)
genes in mice. The presence of eight transcription initiation sites in the a1 isoform 5'-untranslated region suggests that this form of the a subunit may serve a "housekeeping" function in cells (32). Northern blot analysis indicates that most tissues tested express all three isoforms of the a subunit although at levels that suggest tissue-specific functions. Thus the a1 isoform is most heavily expressed in brain and heart although lower levels of expression can be detected in all other tissues. The a2 isoform is predominant in liver and has significant expression in kidney, and lower levels of the isoform are detected in all other tissues except skeletal muscle. The a3 isoform is also heavily expressed in liver, and significant levels are also present in heart, brain, spleen, lung, and kidney.

The results obtained by Northern analysis differ somewhat from the previously reported expression patterns for the a subunit (26–29). Thus, greater differences are observed in expression levels of the a1 isoform between tissues, whereas for the a2 isoform, the highest expression is observed in liver, a

FIG. 7. Identification of alternatively spliced variants of the isoform. Nucleotide sequences of a3 and AA444415 (GenBankTM accession number AF218254) are shown together with the predicted amino acid sequences. AA444415 contains two inserted sequences that are indicated by dashes in the a3 sequence. The first insertion encodes a novel stretch of 30 amino acids followed by a stop codon (indicated by an asterisk). The splicing acceptor and donor sites are shown in the open boxes, and the nine putative transmembrane helices (34) are indicated in the shaded boxes.
Expression of Mouse V-ATPase a Subunit Isoforms

In contrast to the original report of an osteoclast-specific expression of a3 (28), we observe quite broad expression of a3 in a variety of tissues. It is possible that this difference may be due to the relative abundance of the a3 isoform in osteoclasts and other tissues, such that at equivalent levels of mRNA, the a3 isoform may appear to be osteoclast-specific. It is quite clear from the present work, however, that a3 is expressed not only in osteoclasts but in many tissues. We suggest that a3 may be expressed in many cell types that need to target V-ATPases to the plasma membrane, a role that a3 likely also plays in osteoclasts.

In addition to existing in multiple isoforms, the a subunit expressed in mammalian tissues also appears to exist in multiple splice variants. For the a1 isoform, there are three splice variants in mice that differ in the presence or absence of 6–7 amino acid residues near the amino and carboxyl termini of the protein. One of these (a1-I), which lacks the amino-terminal sequence but contains the carboxyl-terminal sequence, appears to be specifically expressed in brain. The carboxyl-terminal insert forms part of a PEST site, which has been found to facilitate more rapid degradation of proteins containing this sequence (33). The presence of this insert in the brain-specific variant of a1 suggests that this form may function in a capacity requiring rapid regulation, such as regulated secretion in neurons. Immunolocalization using antibodies specific for this sequence will be required to test this hypothesis.

In the case of the a3 isoform there are also three alternatively spliced variants, although the difference between these species is more dramatic than that between the variants of a1. Thus whereas the a3-I form encodes a full-length 100-kDa subunit, both the a3-II and a3-III forms are predicted to encode a truncated form of the protein containing only the amino-terminal soluble domain of the protein. The absence of the carboxyl-terminal hydrophobic domain, which has been shown to contain a number of buried charged residues important for proton translocation (30, 31), indicates that this truncated form of the a subunit cannot participate in the formation of a functional proton pump. On the other hand, the amino-terminal domain of the a subunit has recently been shown to reside on the cytoplasmic side of the membrane (34), where it can interact with subunits in the V₁ domain, it is possible that the truncated form of the a subunit could bind to free V₁ domains in the cytoplasm and function in regulating assembly of the V-ATPase. In fact, control of assembly of the V₁ and V₀ domains has been shown to play an important role in regulating vacuolar acidification in both yeast and insects (35, 13), and evidence for a cytoplasmic pool of V₁ domains has been obtained in mammalian cells (36). Western blot analysis using an antibody specific for the a3 isoform will allow determination of the stability and tissue distribution of the truncated form of the protein.

In summary, we have reported for the first time the complete sequence of three distinct isoforms of the V-ATPase a subunit in one species (mouse) and shown that these isoforms are the product of three unique genes. We have evaluated their tissue distribution by Northern analysis and shown that whereas most tissues express all three isoforms, the levels of expression appear to be tissue-specific. In addition, we have identified multiple alternatively spliced variants of both the a1 and a3 isoforms and analyzed their tissue distribution by RT-PCR. Two splice variants of the a3 isoform are predicted to encode truncated forms of the a subunit containing only the amino-terminal soluble domain of the protein.

Note Added in Proof—After submission of this paper, it was reported (Li, Y. P., Chen, W., Liang, Y., Li, E., and Stashenko, P. (1999) Nat. Genet. 23, 447–451) that disruption of the gene encoding the a3 isoform in mice leads to the loss of osteoclast-mediated extracellular acidification.

REFERENCES
1. Forgac, M. (1999) J. Biol. Chem. 274, 12951–12954
2. Stevens, T. H., and Forgac, M. (1997) Annu. Rev. Cell Dev. Biol. 13, 779–808
3. Margulies-Clark, E., Tenney, K., Bowman, E. J., and Bowman, B. J. (1999) J. Bioenerg. Biomembr. 31, 29–38
4. Kane, P. M. (1999) J. Bioenerg. Biomembr. 31, 49–56
5. Anraku, Y., Unemoto, N., Hirata, R., and Ohyu, Y. (1992) J. Bioenerg. Biomembr. 24, 395–406
6. Sze, H., Ward, J. M., and Lai, S. (1992) J. Bioenerg. Biomembr. 24, 371–382
7. Gluck, S. L. (1992) J. Bioenerg. Biomembr. 24, 351–360
8. Nelson, C. N. (1992) J. Bioenerg. Biomembr. 24, 407–414
9. Kibak, H., Taiz, L., Starke, T., Bernasconi, P., and Gogarten, J. P. (1992) J. Bioenerg. Biomembr. 24, 415–424
10. Swallow, C. J., Grinstein, S., and Rotstein, O. D. (1990) J. Biol. Chem. 265, 7645–7650
11. Chatterjee, D., Chakraborty, M., Leit, M., Neff, L., Janmala-Kekkoumpa, S., Fuchs, R., and Baron, R. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 6257–6261
12. Martinez-Zaguilan, R., Lynch, R., Martinez, G., and Gillies, R. (1993) Am. J. Physiol. 265, C1015–C1029
13. Sumner, J. P., Dow, J. A. T., Early, F. G. P., Klein, U., Jager, D., and Wieczorek, H. (1995) J. Biol. Chem. 270, 5649–5653
14. Weber, J., and Senior, A. E. (1997) Biochim. Biophys. Acta 1319, 19–58
15. Capaldi, R. A., Agerger, R., Willrings, S., and Gruber, G. (1996) J. Bioenerg. Biomembr. 28, 397–401
16. Futai, M., and Omote, H. (1996) J. Bioenerg. Biomembr. 28, 409–414
17. Bowman, E. J., Tenney, K., and Bowman, B. J. (1998) J. Biol. Chem. 263, 13994–14001
18. Zimmna, L., Dittrich, P., Gogarten, J. P., Kibak, H., and Taiz, L. (1988) J. Biol. Chem. 263, 9102–9112
19. Manos, D., Merviakury, Y., Hulmes, J. D., Pan, Y. C., Nelson, H., and Nelson, N. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 5521–5524
20. Manos, M. F., Proteau, D., Preston, R. A., Stenartic, A., Roberts, B. T., Hoyt, M., Presus, D., Mulholland, J., Botstein, D., and Jones, E. W. (1992) J. Biol. Chem. 267, 14294–14303
21. Manos, M. F., Wu, B., Proteau, D., Taillen, B. E., Roberts, B. T., Hoyt, M., A., and Jones, E. W. (1994) J. Biol. Chem. 269, 14064–14074
22. Permin, M. S., Fried, V. A., Stone, D. K., Xie, X. S., and Sudhof, T. C. (1991) J. Biol. Chem. 266, 3877–3881
23. Peng, S.-B., Crider, B. P., Xie, X.-S., and Stone, D. K. (1994) J. Biol. Chem. 269, 17262–17266
24. Peng, S.-B., Li, X., Crider, B. P., Zhou, Z., Andersen, P., Tsai, S. J., Xie, X.-S., and Sudhof, T. C. (1991) J. Biol. Chem. 266, 813–821
25. Lee, C.-K., Ghoshal, K., and Beaman, K. D. (1990) Mol. Immunol. 27, 1137–1144
26. Leng, X. H., Manolson, M., Liu, Q., and Forgac, M. (1996) J. Biol. Chem. 271, 22487–22493
27. Leng, X. H., Manolson, M., and Forgac, M. (1998) J. Biol. Chem. 273, 6717–6723
28. Ince, T. A., and Scott, K. W. (1995) J. Biol. Chem. 270, 30249–30252
29. Rogers, S., Wells, R., and Rechsteiner, M. (1986) Science 234, 364–368
30. Leng, X.-H., Nishi, T., and Forgac, M. (1999) J. Biol. Chem. 274, 14655–14661
31. Kane, P. M. (1995) J. Biol. Chem. 270, 17025–17032
32. Myers, M., and Forgac, M. (1993) J. Cell. Physiol. 156, 35–42

FIG. 8. Tissue- and developmental stage-specific expression of a subunit isoforms evaluated by RT-PCR. RT-PCR was performed using isoform-specific primers against a1 (a), a2 (b), and a3 (c) and mRNA isolated from heart (lane 1), brain (lane 2), spleen (lane 3), lung (lane 4), liver (lane 5), skeletal muscle (lane 6), kidney (lane 7), and testis (lane 8) as described under “Experimental Procedures.” Also shown is RT-PCR performed using mRNA isolated from mouse embryos at day 7 (lane 9), day 11 (lane 10), day 15 (lane 11), and day 17 (lane 12).