A 116-kDa polypeptide has recently been found to be a common component of vacuolar proton pumps isolated from a variety of sources. The 116-kDa subunit of the proton pump was purified from clathrin-coated vesicles of bovine brain, and internal sequences were obtained from proteolytic peptides. Oligonucleotide probes designed from these peptide sequences were utilized in polymerase chain reactions to isolate partial bovine cDNA clones for the protein. Sequences from these were then utilized to isolate rat brain cDNA clones containing the full-length coding region. RNA blots indicate the presence of an abundant 3.9-kilobase message for the 116-kDa subunit in brain, and primer extension analysis demonstrates that the cloned sequence is full-length. The rat cDNA sequences predict a hydrophilic amino-terminal half that is composed of >30% charged residues, and a hydrophobic carboxy-terminal half that contains at least six transmembrane regions. The structural properties of the 116-kDa proton pump polypeptide agree well with its proposed function in coupling ATP hydrolysis by the cytoplasmic subunits to proton translocation by the intramembranous components of the pump.

Acidification of intracellular organelles plays an important role in a number of basic cellular processes, including the dissociation of ligands from receptors after receptor-mediated endocytosis, intracellular targeting of newly synthesized enzymes or secretory products to lysosomes and secretory granules, maintenance of an acidic environment in organelles with enzymes that require a low pH for activity, and coupled transport of solutes across lysosomal and secretory vesicle membranes. This acidification is carried out by a class of ATP-driven proton pumps that are often referred to as vacuolar proton pumps (for reviews, see Refs. 1 and 2). Vacuolar proton pumps have been characterized in clathrin-coated vesicles (3, 4); chromaffin granules and synaptic vesicles (5-8); lysosomes (9, 10); endosomes (11, 12); Golgi-derived membranes (13, 14); and the vacuolar membranes of plants (15-18), Neurospora (19, 20), and yeast (21, 22).

The vacuolar proton pump purified from clathrin-coated vesicles contains eight to nine polypeptides with apparent molecular masses of 116, 70, 58, 40, 38, 34, 33, 19, and 17 kDa, respectively (23). The functions and structures of several of the proton pump's components have been identified, including the 70- and 58-kDa subunits, which are homologous to the $\alpha$ and $\beta$ subunits of the mitochondrial ATPase and function in ATP hydrolysis (24-28). In addition, the 40- and 33-kDa subunits were also shown to be part of the ATP hydrolytic sector (24, 29), and the 17-kDa subunit is thought to represent the proton channel of the pump (30, 31).

Initially, the 116-kDa polypeptide present in proton ATPases from clathrin-coated vesicles and chromaffin granules (1) was not found in vacuolar proton pumps that were partially purified from lower organisms; preparations from plants, fungi, and yeast were first thought to be composed of only three subunits of 70, 58, and 17 kDa, respectively. This led to the speculation that the additional polypeptides present in mammalian enzyme preparations were "accessory" to function. However, it is important to note that these preparations with apparently three subunits were not reconstituted in an active form and that the specific ATPase activity in most of these preparations was about one-tenth of that of the proton ATPase isolated from bovine brain clathrin-coated vesicles. Thus, two possibilities emerged: first, that striking differences indeed exist between vacuolar proton pumps harvested from mammals and simpler eukaryotes; or second, that the reported subunit compositions of yeast, plant, and fungal proton pumps reflect an incomplete purification or proteolytic degradation of key components. With regard to the latter, it is pertinent to note that the 116-kDa component is particularly susceptible to proteolysis (2).

Recently, the subunit compositions of the vacuolar proton pumps from yeast (22) and plants (18) have been re-examined. It was found that these proton ATPases have polypeptide compositions which include a 116-kDa component and are also otherwise similar to that of the clathrin-coated vesicle proton pump. In addition, these preparations of plant and yeast proton pumps have a specific activity of ATP hydrolysis (15-20 $\mu$mol of P$_i$/mg of protein/min) that closely matches...
that of the mammalian coated vesicle proton-translocating ATPase (23). Thus, at present, there is considerable evidence that the 116-kDa and other accessory subunits may be found in vacuolar proton pumps of most species.

Although it is the largest component of the proton pump, nothing is known about the primary structure of the 116-kDa polypeptide. Attempts at a biochemical definition of the role of the 116-kDa polypeptide in the function of the clathrin-coated vesicle proton pump have been partially successful. The ATPase activity of the purified and reconstituted vacuolar proton pump can be supported by either Mg2+ or Ca2+, but proton pumping is observed only in the presence of Mg2+ (24). Selective removal of the 116- and 38-kDa subunits from the pump results in a enzyme complex that retains the Ca2+-protons and has no Me-supported ATPase activity (24). Although it is the largest component of the proton pump, nothing is known about the primary structure of the 116-kDa polypeptide.

Synaptic vesicles are abundant secretory organelles of neurons that use the proton gradients and potential gradients established by the activity of a vacuolar proton pump to take up and store neurotransmitters (reviewed in Ref. 33). Recent evidence indicates that neurons contain a more than 20 fold higher vacuolar proton pump activity than most other tissues and that most of that proton pump activity in brain is localized to synaptic vesicles. Synaptic vesicles probably recycle via clathrin-coated vesicles (34), and it seems likely that a sizable proportion of brain clathrin-coated vesicles represent the endocytotic leg of the synaptic vesicle pathway.

We have now molecularly cloned the 116-kDa component of the rat brain vacuolar proton pump. It has a structure that is composed of two domains: a hydrophilic amino-terminal domain in which almost one-third of the amino acids are positively or negatively charged, and a hydrophobic carboxyl-terminal region which contains at least six transmembrane regions and is composed of >50% hydrophobic amino acids. These structural characteristics of the 116-kDa polypeptide agree well with its proposed function in coupling ATP hydrolysis to proton movement through the intramembranous proton channel.

**EXPERIMENTAL PROCEDURES**

**Purification and Peptide Sequencing of Bovine Brain Proton Pump**—The vacuolar proton pump was purified from bovine brain clathrin-coated vesicle fractions as described (23). The 116-kDa component was isolated by preparative gel electrophoresis, electroeluted from the gel, and digested with trypsin (28). Tryptic fragments were isolated by reverse-phase high-performance liquid chromatography, and eight well-separated tryptic fragments were selected for amino acid sequencing by automated Edman degradation. A total of 10 peptide sequences were determined because some proteolytic fragments yielded more than one sequence. Of these peptide sequences, all of the eight sequences derived as major signals and one of the sequences present as a minor signal were found in the amino acid sequence translated from the rat mRNA (Fig. 3). Numbers to the right indicate positions of RNA molecular size standards in kilobases (kb).

**cDNA Cloning and Sequencing**—After initial attempts failed to isolate cDNA clones encoding the 116-kDa polypeptide by screening a total of 4 × 10⁶ plaques of different bovine cDNA libraries with degenerate oligonucleotides, a strategy based on the polymerase chain reaction (PCR) was adopted. DNA from a bovine retina cDNA library in Agt10 (kind gift of Dr. J. Nathans, Johns Hopkins University) was purified and used as a template for PCR using an oligonucleotide designed on the basis of a peptide sequence (GAGGATCCCA-CIGAGGACATGATGACGIGGIGA/C/T/T/A, where I denotes inosine and the 2 bases in parentheses identify a redundant position) and a second oligonucleotide complementary to a sequence flanking the cloning site in the Agt10 vector. PCR was performed in the presence of a 1 μM concentration of both oligonucleotides for 35 cycles of 90 s each at 95 °C followed by 4 min at 68 °C. Reaction products were cleaved with BamHI and EcoRI, purified by polyacrylamide gel electrophoresis, and cloned into M13 vectors as described (36). One fragment of the >10 PCR product fragments sequenced was identified as being correct because it contained sequences predicted by the peptide sequence beyond the oligonucleotide sequence used for PCR.

An oligonucleotide (complementary to residues 721–749; Fig. 3) was then designed on the basis of the sequence of the PCR product. After RNA blots had indicated that this oligonucleotide specifically reacted with the 116-kDa subunit message in rat brain (see below), the oligonucleotide was used to screen a rat brain cDNA library as described (37, 38). Rat brain was chosen as a source because of our long-standing work on rat brain synaptic vesicles, which contain the majority of the rat brain proton pump. A total of 19 hybridization-positive clones were isolated, nine of which were mapped and partially sequenced. The clone with the largest insert (pP116-1) was completely sequenced but was found to contain a 5'-fusion to the poly(A) tail of an unrelated unidentified cDNA and not to be full-length. Accordingly, a second clone (pP116-17), which was found to be full-length, was sequenced; and several other clones were partially sequenced to test specific questions (see below).

DNA sequencing was performed by the dideoxy chain termination method. The abbreviations used are: PCR, polymerase chain reaction; bp, base pair(s).
FIG. 4. Hydrophilicity plot of amino acid sequence of 116-kDa proton pump polypeptide.

RESULTS

The clathrin-coated vesicle proton pump was purified from bovine brain as described and, importantly, was separated from the unrelated 116-kDa ATPase (41), which copurifies at early stages of purification. The 116-kDa component of the vacuolar proton pump was purified by preparative gel electrophoresis, and tryptic fragments were isolated from it by high-performance liquid chromatography (data not shown). These were then used to obtain internal amino acid sequences to allow cDNA cloning.

Several attempts at isolation of cDNA clones encoding the 116-kDa proton pump component from multiple bovine cDNA libraries using oligonucleotide probes designed on the basis of peptide sequences were unsuccessful. We then used a cloning method based on the polymerase chain reaction (see "Experimental Procedures"). The single correct PCR fragment that we obtained was identified on the basis of its sequence, and an oligonucleotide complementary to this sequence was used to probe RNA blots of rat brain (Fig. 1). A single message of ~3.9 kilobases was specifically recognized by this oligonucleotide. The 116-kDa message was fairly abundant in brain as judged by the strength of the hybridization signal and was at least 20 times more abundant in brain than in any other tissue tested (Fig. 1 and data not shown).

The oligonucleotide was then used to screen rat brain cDNA libraries. Of the 19 hybridization-positive clones that were isolated, several were studied in detail by restriction enzyme mapping and sequencing. The localization and sequencing strategy of the three most important clones are shown in Fig. 2. The nucleotide sequence of the inserts of these clones encoding the 116-kDa proton pump subunit are aligned below the translated amino acid sequence and printed in boldface type. The 18-nucleotide and 6-amino acid sequence (at positions 706-721, respectively) that was absent from one cDNA clone (pP116-5) but present in the other two is also shown in boldface type. The in-frame termination codon in the 5'-untranslated region and the translation termination codon as well as the polyadenylation consensus sequence are underlined.

The nucleotide sequence of the inserts of the cDNAs encoding the 116-kDa proton pump polypeptide contains 3865 bp excluding the 3'-poly(A) tail. The poly(A) tail of the cDNAs is preceded by a canonical polyadenylation signal (underlined in Fig. 3), of which there is no other copy in the genome. The 116-kDa message was fairly abundant in brain as judged by the strength of the hybridization signal and was at least 20 times more abundant in brain than in any other tissue tested (Fig. 1 and data not shown).

The oligonucleotide was then used to screen rat brain cDNA libraries. Of the 19 hybridization-positive clones that were isolated, several were studied in detail by restriction enzyme mapping and sequencing. The localization and sequencing strategy of the three most important clones are shown in Fig. 2. Of these cDNA clones, only pP116-17 was found to be full-length (see below). The nucleotide and translated amino acid sequences derived from the cDNA clones are shown in Fig. 3.
The cDNA is full-length as judged by the fact that its size corresponds well to the message size observed on RNA blots and that it contains an in-frame termination codon 5' to the initiator methionine (underlined in Fig. 3). In addition, we performed primer extension analysis of the 5'-end of the message for the 116-kDa subunit using rat brain RNA. These experiments demonstrated the presence of a single transcription start site beginning ~70 bp 5' to the 5'-end of our cDNA (data not shown).

The cDNA sequence exhibits a single large open reading frame, the translation of which predicts the synthesis of an 838-amino acid protein with an $M_r$ of 96,267 (Fig. 3). The sequence context of the putative initiator methionine agrees well with the consensus sequences for initiator methionines (42). All of the peptide sequences obtained from the purified bovine protein (except for one peptide derived from contaminating keratin) were found in the amino acid sequence translated from the rat cDNA sequence (shown below the translated amino acid sequence in Fig. 3). This suggests that the cDNA indeed encodes the 116-kDa proton pump component and that this polypeptide is well conserved evolutionarily. The fact that most of the peptide sequences of the 116-kDa component were derived from its amino-terminal half together with its relatively long 3'-untranslated region, may explain the difficulty in cloning its cDNA.

Two cDNAs were fully sequenced, and one was partially sequenced (Fig. 2). All sequences determined were identical with the exception of an 18-bp sequence that was missing in pPI16-5, but was present in the other two clones (nucleotides 2146-2165; boldface type in Fig. 3). The absence of this 18-bp sequence could be due to either an alternative splicing event or a cloning artifact. However, the fact that this is an in-frame deletion that is identically present in a partial human cDNA suggests alternative splicing of the 116-kDa protein.

The predicted amino acid sequence of the 116-kDa proton pump subunit is very hydrophilic in the first 390 amino acids (30.1% charged residues and 38.2% hydrophobic residues) and rather hydrophobic in the following 448 amino acids (51.0% hydrophobic and 17.1% charged residues). Hydrophobicity plots (43) indicate that the amino-terminal half of the 116-kDa polypeptide contains no transmembrane regions, whereas extended hydrophobic sequence stretches are present in the carboxyl-terminal half (Fig. 4). According to the criteria of Kyte and Doolittle (43) and of Klein et al. (44), at least six transmembrane regions are present in this region (Table I). The putative transmembrane regions are separated by spacer regions of different lengths and hydrophilicity, and additional transmembrane regions may be present that are formed by the less hydrophobic connecting sequences (the two most probable of these are also listed in Table I).

Only a single $N$-linked glycosylation consensus sequence is present in the carboxyl-terminal half of the protein (residues 489-491), indicating that if the 116-kDa component is glycosylated as reported (32) and if this glycosylation is $N$-linked, then these residues are probably intravesicular. No ATP-binding consensus sequence was found in any part of the 116-kDa structure.

**DISCUSSION**

This study reports the purification and amino acid sequencing of the 116-kDa polypeptide of the vacuolar proton pump from bovine brain. cDNA clones of the 116-kDa component were isolated from rat brain, and its complete primary structure was determined. Mounting evidence now indicates that the 116-kDa component may constitute a major and invariant subunit of the vacuolar proton pump that is not only a functionally essential component of the vertebrate pumps (24), but is also present in lower unicellular eukaryotes and plants (18, 22). The 17- and 116-kDa components are the components of the proton pump that are most heavily labeled by hydrophobic reactive probes (45). These results suggest that the 116-kDa polypeptide may be part of the proton-conducting, intramembranous complex of the vacuolar proton pump. This view is also supported by functional studies which demonstrate that selective removal of the 38- and 116-kDa components of the proton pump abolishes proton pumping. Under these conditions, the Mg$^{2+}$-ATPase activity of the pump is depressed, whereas its Ca$^{2+}$-activated (24). Although partial biochemical definition of the ATP hydrolytic subcomplex and the proton pore has been achieved, reconstitution of the coupling factors responsible for linking these two functions has not yet been possible, and we speculate that the 116-kDa polypeptide may serve such a role.

The primary structure of the 116-kDa polypeptide predicts a 838-residue protein composed of two large domains, each of which constitutes approximately half of the protein: a highly charged hydrophilic amino-terminal domain and a hydrophobic carboxyl-terminal domain that contains multiple membrane-spanning regions. At least six transmembrane regions are present in the carboxyl-terminal half of the 116-kDa polypeptide as judged by the criteria of Klein et al. (44) and Kyte and Doolittle (43); and two more membrane-spanning regions seem probable, although they do not contain enough hydrophobicity to allow an unequivocal assignment (Table I).
Three of the putative transmembrane regions contain negatively charged residues, and one contains a positively charged residue. If these charged residues are truly in the middle of the membrane, either they could function in interactions between intramembranous helices or they may have role in the movement of protons by the pump.

The putative transmembrane regions are interconnected by sequences of different lengths and character. The longest of the connecting sequences is found at residues 655–741. This sequence is probably differentially spliced since two classes of cDNAs containing or missing an 18-bp segment in this region could be isolated. If the hydrophilic amino terminus is cytoplasmic (see below) and if the 116-kDa subunit indeed contains eight transmembrane regions, then this sequence would be cytoplasmic and may have relevance for the functional properties of the pump.

No significant homology was observed between the sequence of the 116-kDa subunit and the sequences in the current data banks with the exception of a small sequence of this similarity encompassing 8 consecutive identical amino residues (Sequences are shown in single-letter amino acid code and are flanked by their respective residue numbers; identical residues are marked by asterisks.) The functional significance of this similarity encompassing 8 consecutive identical amino acids is currently unclear.

A distinct two-domain structure was postulated for the 116-kDa component of the vacuolar pump based on its primary structure. The presence of multiple membrane-spanning regions in the carboxyl-terminal half of the 116-kDa polypeptide agrees well with its possible participation in the intramembranous proton-translocating complex (24, 32). The 116-kDa subunit also contains a large and highly charged amino-terminal domain of unknown function that may interact with the ATPase domain.

No signal sequence can be found preceding the amino-terminal domain, and the 116-kDa polypeptide is fully susceptible to proteolysis in intact vesicles (45). This suggests that the amino-terminal domain of the 116-kDa polypeptide is cytoplasmic and may directly interact with the cytoplasmic subunits of the pump in addition to the intramembranous subunits. Although such an interaction was not demonstrated in cross-linking studies of the proton pump, it could easily have been missed because of the inherent limitations of the cross-linking approach and because these studies did not utilize antibodies to visualize specific cross-linked products (32). If the soluble amino-terminal domain of the 116-kDa polypeptide indeed interacts with the cytoplasmic subunits of the pump, it would be in an ideal position to mediate the coupling between ATP hydrolysis by the cytoplasmic 70- and 58-kDa subunits and proton translocation by the intramembranous subunits, including perhaps its own transmembrane regions. Such a role would also agree very well with the functional studies demonstrating that selective removal of the 38- and 116-kDa subunits from the pump transforms its Mg2+-activated ATPase into an uncoupled Ca2+-activated ATPase (24).