Generation of truncated brain AE3 isoforms by alternate mRNA processing

Catherine W. Morgans and Ron R. Kopito*
Department of Biological Sciences, Stanford University, Stanford CA 94305-5020, USA
*Author for correspondence

SUMMARY

AE3 gene is a member of the AE anion exchanger gene family that is expressed primarily in brain and heart. The principal product of the AE3 gene in rodent brain, FL-AE3p, when expressed in heterologous cell lines, gives rise to chloride-dependent changes in intracellular pH consistent with its operation as an anion exchanger. We have identified two novel isoforms of mouse AE3 that are generated by tissue-specific alternate RNA processing. One of these isoforms encodes a polypeptide, 14-AE3p, that corresponds to a portion of the NH2-terminal cytoplasmic domain of AE3. 14-AE3p lacks the entire transmembrane domain that - in FL-AE3p - forms the anion exchange channel. Immunoblots with antibodies to the NH2- and COOH- termini confirm that FL-AE3 and 14-AE3 are expressed in rat brain as 160 kDa and 74 kDa polypeptides, respectively. Unlike FL-AE3p, however, 14-AE3p is insoluble in non-ionic detergent, suggesting a possible association with the cytoskeleton.

Key words: anion exchanger, alternate splicing, AE3, cytoskeleton

INTRODUCTION

AE3 is a member of the AE gene family of anion exchangers (reviewed by Kopito (1990)) that is expressed primarily in brain neurons and heart (Kopito et al., 1989; Kudrycki et al., 1990). Like the other members of this family, which include the erythrocyte ‘band 3’ protein (AE1p) and a broadly distributed isoform, AE2p (reviewed by Kopito (1990)), AE3p catalyzes the transmembrane exchange of chloride and bicarbonate (Kopito et al., 1989; Lee et al., 1991). By virtue of this activity the AE proteins contribute to the maintenance of cellular ionic homeostasis and the regulation of intracellular volume and pH (Kopito, 1990). All known AE family members share a common structural organization consisting of an ~600 amino acid COOH-terminal membrane-associated domain (which spans the membrane 12-14 times) and a NH2-terminal domain of variable size (Kopito, 1990). The former domain is highly conserved among the AE family members in both primary sequence and predicted secondary structure. The 81% sequence identity between the COOH-terminal domains of AE1p and AE3p is reflected in the functional similarity between the two proteins. Expression of the COOH-terminal 645 amino acids of AE3p in COS or human embryonic kidney (HEK) cells (Lee et al., 1991) leads to a measurable increase in Cl-/HCO3- exchange activity, demonstrating that this domain is sufficient to mediate anion exchange (Kopito et al., 1989).

By contrast, the function of the NH2-terminal domain has been established only for AE1p (Bennett and Stenbuck, 1980). In the erythrocyte, the ~43 kDa NH2-terminal domain of AE1p faces the cytoplasm, where it interacts with several cytoplasmic proteins including hemoglobin, glycolytic enzymes and ankyrin (reviewed by Low (1986)). This latter interaction is significant as it forms the primary linkage between the phospholipid bilayer and the subcortical spectrin/actin cytoskeleton, which confers on erythrocytes their characteristic morphological and rheological properties (reviewed by Bennett (1990)). The high-affinity (Bennett and Stenbuck, 1980) ankyrin-binding site has not been mapped on AE1p, although indirect evidence suggests that it is somewhat removed from the extreme NH2 terminus (reviewed by Low (1986)). The cytoplasmic domain of AE3p shares significant (59%) sequence identity with the corresponding domain of AE1p, raising the possibility that AE3p may interact with components of the neuronal cytoskeleton or membrane skeleton. Such an association has yet to be established, although several ankyrin isoforms have been identified in neurons (Otto et al., 1991; reviewed by Bennett (1992)). We have recently demonstrated that, when co-expressed in transfected HEK cells, AE3p forms a stable complex with the repeat domain of ankyrin (Morgans and Kopito, 1993). The AE3p cytoplasmic domain is also distinguished by the presence of an ~260 residue, proline-rich extension at the extreme NH2 terminus. Thus, while the COOH-terminal domains of AE1p and AE3p are highly conserved in both sequence and function in anion exchange, the role of the more divergent NH2-terminal, cytoplasmic domain of AE3p remains unknown.

The AE3 gene is expressed in brain and heart and, at much lower levels, in several other tissues including kidney, stomach, skeletal muscle and testes (Kopito et al., 1989; Linn et al., 1992). The predominant AE3 transcript in heart is ~1 kb shorter than that in brain (Kopito et al., 1989;
Kudrycki et al., 1990) due to tissue-specific alternative processing of four exons at the 5′ end of the AE3 gene (Linn et al., 1992). Cardiac AE3p is therefore expected to differ from brain AE3p only at the extreme NH2-terminal portion of the cytoplasmic domain.

Alternative mRNA processing is widely used by eukaryotic cells to increase the coding power of a gene, and can be used to generate proteins with quite different biological activities. We report here the identification of two novel AE3 transcripts, which arise from tissue-specific alternative mRNA processing. Both these sequences encode AE3 polypeptides that are truncated within the cytoplasmic domain. One of these truncated AE3 polypeptides is expressed in brain as a 74 kDa polypeptide that is present within a high molecular mass, detergent-insoluble complex, suggesting a possible association with the cytoskeleton.

### MATERIALS AND METHODS

#### Genomic clone isolation and sequencing

A mouse B10A genotyping screen was screened with a HindIII/Smal fragment from plasmid pRK128 (nucleotides 2128 to 2510) of the cDNA sequence (Kopito et al., 1989) for the 5′ end of the AE3 gene, and the entire insert of prK133 (nucleotides 413 to 815) for the 5′ end of the gene. From 8.5×10^3 plaques, 21 duplicate positive clones were found and of these, six were purified. A 2.1 kb XhoI restriction fragment of one of the clones was found by Southern blotting (Maniatis et al., 1982) to contain both alternatively spliced sequences. This fragment was subcloned into a Bluescript pSK+ vector (Stratagene) and called pCM105. Nested deletions of the insert were generated using exonuclease III and S1 nuclease and then sequenced with the dyeoxy chain termination method as previously described (Kopito et al., 1989).

#### S1 nuclease mapping

Total RNA and polyadenylated RNA were prepared as previously described (Kopito et al., 1989). S1 nuclease mapping was performed as described by Kopito et al. (1987). Briefly, RNA samples were hybridized with 10,000 to 50,000 cpm of 32P-labeled cDNA probe and then digested with 75 units of S1 nuclease per sample for 1 hour. Protected fragments were resolved by gel electrophoresis on a 6% acrylamide/7.5 M urea gel followed by electroelution from the gel. The probe used to analyze the 311 base fragment (Fig. 2b), the first 9 nucleotides contain a HindIII restriction site, and the remaining 20 nucleotides match nucleotides 505 to 525 of the genomic sequence. The sequence of oligo 2 is 5′-CCCAAGCTTTAGCCCTGTCCTCAGGCCGG-3′. The first 9 nucleotides of this sequence also contain a HindIII restriction site, and the last 11 amino acids of 311-AE3 or the last 12 amino acids of 14-AE3 were coupled to keyhole limpet hemocyanin (KLH), and to bovine serum albumin (BSA). New Zealand white rabbits were immunized with KLH-coupled peptides, then boosted once a month with alternating injections of peptide-BSA and peptide-KLH. The sera were tested by ELISA and by immunofluorescent staining of COS-7 cells transfected with 311-AE3 or 14-AE3. Antibody 7210 was affinity-purified by incubation overnight at 4°C with nitrocellulose strips onto which FL-AE3 had been electrophoretically transferred followed by elution in 0.1 M glycine, pH 2.5, for 10 minutes at room temperature. The eluate was neutralized by the addition of 50 µl/ml 1 M Tris-HCl, pH 9.5, and then concentrated and dialyzed against PBS in a Centricon-30 (Amicon) that had been preabsorbed with 2% BSA. The antibody was stored in PBS/1% BSA/0.02% NaN3 at 4°C. Antibody 7015 was affinity-purified on an ACA-22 Ultragel column (LKB) to which the 14-AE3 C-terminal peptide had been coupled. It was eluted with 0.1 M glycine, pH 2.5. The eluate was treated as described above for 7210.

#### Cultured cell growth and transfection conditions

COS-7 cells were grown and transfected in Dulbecco’s modified medium supplemented with 10% fetal calf serum in a humidified incubator at 37°C and in 5% CO2. They were plated at a density of 5×10^3 per 100 mm culture dish 18 hours prior to transfection, and were then transfected with 8 µg DNA per dish as previously described (Oprian et al., 1987). At 48 to 56 hours after transfection, the cells were either fixed and stained for immunofluorescence, or processed for immunoblotting by dissolving in Laemmli sample buffer (Laemmli, 1970).

#### Triton X-100 extraction of rat brain homogenate

Brains from Sprague-Dawley male rats from SIM were homogenized in a Potter-Elvijer homogenizer in 10 ml H buffer (0.32 M sucrose, 10 mM Tris, pH 7.4, 1 mM EDTA, 0.1 mM EGTA) per brain with 10 strokes of a loose-fitting Teflon pestle at 500 rpm. Homogenate corresponding to 100 µg of protein was diluted to 100 µl in H buffer containing 1% Triton X-100. The mixture was incubated on ice for 30 minutes, then centrifuged for 15 minutes at 10,000 g in a microfuge at 4°C. The supernatant and pellet were diluted to 200 µl in Laemmli sample buffer for western blot analysis.

#### Immunoblotting

Samples were electrophoresed on a 7.5% acrylamide/SDS gel
Alternate AE3 mRNAs encode truncated polypeptides

Screening of a mouse brain cDNA library for full-length AE3 clones (Kopito et al., 1989) resulted in the isolation of two additional classes of cDNA that are identical to full-length AE3, with the exception of insertions of 311 bp and 14 bp at positions 1325 and 1646, respectively. The novel clones were designated 311-AE3 and 14-AE3. To determine whether these cDNAs were derived from the same AE3 gene, a 2.1 kb genomic subclone containing this region was isolated (Fig. 1A). Alignment of this sequence with that of the published FL-AE3 cDNA (Kopito et al., 1989) indicates that it contains 5 exons, designated A-E (Fig. 1A,B). These exons correspond to exons 7-11 of the rat AE3 genomic sequence (Linn et al., 1992). Each of these exons is flanked by consensus splice donor and acceptor sequences (Smith et al., 1989). Analysis of the genomic sequence revealed that the 311 bp insertion identified in several cDNA clones corresponds to an unspliced intron between exons ‘B’ and ‘C’. The 14 bp insertion arises from the use of an alternate splice acceptor site between exons ‘D’ and ‘E’ at the end of intron ‘e’ (Fig. 1B). We conclude that all three AE3 transcripts arise from a common AE3 gene by alternative mRNA processing.

Conceptual translation of the amino acid sequences encoded by 311-AE3 and 14-AE3 indicated that both insertions disrupt the AE3 reading frame (Fig. 1B). The 14 bp insertion causes a frame-shift resulting in termination at a TGA 16 codons downstream of the insertion site. The polypeptide encoded by 311-AE3 (311-AE3p) is predicted to terminate at a TAG 12 codons downstream of the unused splice donor site. Thus, by contrast to the 135 kDa polypeptide predicted from FL-AE3 (FL-AE3p), 311-AE3 and 14-AE3 are predicted to encode polypeptides of 43 kDa and 56 kDa, respectively. These polypeptides, which correspond to the NH2-terminal domain of AE3, are predicted to lack the highly conserved membrane-associated domain shared by all three members of the AE gene family. They are thus not predicted to be integral membrane proteins and are not likely to participate directly in ion transport.

Tissue-specific differential processing of AE3 transcripts

The presence of both alternate sequences in AE3 mRNA was investigated by S1 nuclease protection analysis (Fig. 2). A 358 bp probe containing the 14 bp alternate sequence was 5’ end-labeled and hybridized with RNA from mouse brain, heart and kidney prior to digestion with S1 nuclease and resolution on a polyacrylamide gel (Fig. 2A). This probe contained 100 nt of vector sequence at the 3’ end in addition to the AE3 sequence (Fig. 2C), which permitted discrimination between the undigested probe (358 nt) and the fully protected probe (258 nt). Two protected fragments were detected in all three tissues: a 258 nt fragment corresponding to 14-AE3, and a 220 nt fragment corresponding to transcripts lacking the 14 bp insert. Protected fragments of identical size were detected in all AE3-expressing tissues including brain, heart and kidney (Fig. 2A), and testes and skeletal muscle (data not shown). In all cases the 220 bp fragment was the predominant species detected. Densitometric analysis of the bands, however, showed that while only 3% of AE3 mRNA in brain and heart contains the 14 alternate bases, this isoform accounts for 33% of transcripts detected in kidney mRNA. The different ratios of the two bands between tissues suggest the involvement of tissue-specific regulation of the splicing of intron ‘e’.

The relative abundance of 311-AE3 transcript was assessed by S1 nuclease protection using a uniformly labeled 460 bp cDNA probe spanning from the middle of exon ‘A’ to the middle of intron ‘c’ and including 156 bp of vector sequence (Fig. 2C). This probe was predicted to generate 304 bp and 218 bp bands corresponding to transcripts either containing or lacking the 311 bp intron, respectively. This analysis generated rather broad bands centered around the expected sizes (Fig. 2B). The significance of these diffuse bands is unclear, but was reproducible. Such heterogeneity of S1 nuclease digestion may reflect a degree of ‘nibbling’ of heteroduplex DNA and has been observed by others (Green and Roeder, 1980). Densitometric analysis of the autoradiograms indicates that the 311 base sequence is present in approximately 19% of heart, 33% of brain and 68% of kidney AE3 transcripts. Taken together, these data suggest that both 311-AE3 and 14-AE3 transcripts are present at steady-state in polyadenylated RNA and that their relative abundance is regulated by tissue-specific factors.

To confirm the S1 nuclease protection data on the presence of the 311 bases in AE3 transcripts, cDNA from brain, heart, and kidney was analyzed by the polymerase chain reaction (PCR) using two primers, corresponding to AE3 sequences flanking the 311 bp intron (see Materials and Methods). The major products of the PCR were 316 and 627 bp, which correspond to the FL-AE3 and 311-AE3 sequences, respectively (Fig. 3A). These sizes were confirmed by performing the PCR reaction on plasmids containing cloned copies of FL-AE3 and 311-AE3 cDNA. The results show that RNA samples from all tissues sources contained both types of transcript. The relative intensities of the two bands, however, could not be taken as a quantitative index of the relative abundance of the two mRNA species since the ratios were not preserved upon re-amplification of the PCR products (data not shown). The PCR results do argue though that the ‘smear’ observed in the S1 analysis of 311-AE3 is not likely to be due to heterogeneous...

(Laemmli, 1970), then electrophoretically transferred to a nitrocellulose filter. The filter was incubated for 30 minutes in 150 mM NaCl, 10 mM Tris-HCl, pH 7.5, 0.05% Tween-20, 5% w/v nonfat dry milk (TBSTM) to block non-specific binding, then in TBSTM containing the primary antibody for either 2 hours at room temperature or 4°C overnight. Following one 30 second wash and three 15 minutes washes in TBST (TBSTM with no milk), the filter was incubated in donkey anti-rabbit antibody coupled to horseradish peroxidase (Amersham) diluted 1:2000 in TBSTM for 30 minutes at room temperature, and then washed as before. The filter was then incubated for 1 minute in equal volumes of the two ECL reagents (Amersham), pressed between two pieces of blotting paper to remove excess liquid, and exposed to Kodak XAR-5 film.
ity in the location of the actual splice sites since the PCR-generated bands are homogeneous.

**Expression of AE3 isoforms in brain**

In order to assess the expression of individual AE3 polypeptides isoform-specific polyclonal antibodies were generated for use in immunoblotting and immunohistochemistry. Antisera were raised in rabbits to synthetic peptides corresponding to the unique COOH termini predicted from the DNA sequences of the individual isoforms (Fig 4A). Additionally, antibody 7210 (Kopito et al., 1989) was raised against a fusion protein encoding the NH2-terminal ~23 kDa region of AE3 common to all three AE3 isoforms. All antisera were affinity-purified.

Immunoblots of rat brain homogenate probed with antibody 7210 revealed the presence of two immunoreactive
bands (Fig. 4B, lane 2). The strongest signal was at ~160 kDa, which corresponds well with the predicted size of glycosylated FL-AE3p, and is in agreement with the size of the protein detected in cells transfected with the FL-AE3 cDNA (Kopito et al., 1989). The other band, at 74 kDa, is larger than predicted from either of the two alternate AE3 isoforms, but is identical in mobility to the band detected in COS (Fig. 4B, lane 1) or human embryonic kidney cells (HEK-293; data not shown). Significantly, a band of indistinguishable mobility was also observed on immunoblots of rat brain homogenate probed with antibody 7015 (Fig 4B, lane 3), which was raised against the unique COOH-terminal peptide predicted from the 14-AE3p sequence. These data confirm that both FL-AE3p and 14-AE3p are expressed as stable polypeptides in rat brain.

By contrast to the results with 14-AE3p, we obtained no evidence supporting the existence of 311-AE3p in rat brain. Only two bands were detected in immunoblots probed with antibody 7210 (Fig. 4B, lane 2), against the common fragment. No bands were observed in immunoblots probed with the 7017 antibody (data not shown). Although it is possible that the antibody, which was raised against the mouse sequence, does not cross-react with rat, these results suggest that in contrast to FL-AE3 and 14-AE3, 311-AE3 is not translated into a stable polypeptide in rat brain.

**Detergent solubility of brain AE3 isoforms**

Since the 14-AE3p sequence shares significant homology to the cytoskeleton-binding domain of AE1, we used non-ionic detergent extraction to detect possible association of the brain AE3 isoforms with the cytoskeleton. Crude rat brain homogenate was extracted in 1% Triton X-100 in isotonic buffer, then separated into ‘soluble’ and ‘insoluble’ fractions by centrifugation at 10,000 g. These fractions were then immunoblotted with antibody 7210 (Fig. 5). Virtually all of the material migrating at 160 kDa, which corresponds to FL-AE3p, was soluble under these conditions. In contrast, all of the detectable 14-AE3p was present in the insoluble pellet. No 14-AE3p was detected on immunoblots of purified brain nuclei (data not shown), indicating that the 14-AE3p in the pellet was not introduced by contaminating nuclei.

**DISCUSSION**

We have identified two novel isoforms of AE3 generated...
by alternate RNA processing of the mouse AE3 gene. Both of these transcripts are predicted to encode prematurely truncated AE3 polypeptides that lack the membrane-associated domain which in the full-length AE3 polypeptide (FL-AE3p), forms the anion exchange channel. We demonstrate that one of these alternate AE3 isoforms, 14-AE3p, is translated into a 74 kDa polypeptide in rat brain that may interact with the cytoskeleton.

The mouse AE3 genomic sequence corresponds well with the genomic sequence of rat AE3 (Linn et al., 1992). The location of the intron/exon boundaries is identical, although the introns diverge more in both sequence and length. The AE3 genomic sequence confirms that the FL-AE3, 311-AE3 and 14-AE3 transcripts all derive from a common gene by alternate mRNA processing. The 311 bp sequence is an unspliced intron since it contains consensus splice donor and acceptor sequences at its ends and lies between two known exons. In the rat, the equivalent of the 311 intron is 358 bp in length, and is also present in a subset of AE3 transcripts (Kudrycki et al., 1990). The sequence of this intron is, however, rather poorly conserved between rat and mouse; it is not predicted to encode a polypeptide similar to 311-AE3p. Because of this poor conservation of the 311-AE3 sequence and our inability to detect a corresponding protein, it is likely that the 311-AE3 transcript represents a splicing intermediate. However, similar use of an unspliced intron to create a truncated protein has been recently observed in the processing of the mRNA encoding the α-1 subunit of the Na⁺,K⁺-ATPase, another ion transporting, integral membrane protein (Medford et al., 1991).

In contrast to 311-AE3, the sequence of 14-AE3 is highly conserved between rat and mouse. The sequences of both alternate splice acceptors at the junction between intron ‘e’ and exon ‘E’ are perfectly conserved in the rat (where exon ‘E’ corresponds to rat exon 10). Moreover, the nonsense codon in the alternate reading frame is conserved between rat and mouse, as is the sequence intervening between the splice acceptor and the premature termination of the 14-AE3p peptide. Use of this alternative splice acceptor site in the rat should give rise to a 14-AE3p polypeptide of nearly identical sequence to mouse. This prediction is supported by immunoblots of rat brain showing that the mouse 14-AE3p specific antibody reacts with the appropriate polypeptide from rat brain. The use of alternate acceptor and donor sequences is common in the processing of viral mRNAs and has also been observed in the processing of several Drosophila genes (Smith et al., 1989). In vertebrates, the use of alternate donor sequences occurs in the processing of the mRNA for the two prohormones for gastrin-releasing peptide (Spindel et al., 1986). The splicing of the 14 bp sequence belongs in this class of alternate splicing and involves a choice between two potential acceptor sequences.

The 14-AE3p isoform corresponds to the NH₂-terminal 485 codons of FL-AE3p with a unique 20 codon COOH-terminus generated from the insertion of the extra 14 bp and the consequent shift in reading frame. The resulting 14-AE3p protein is predicted to have a molecular mass of 56 kDa and to lack the entire transmembrane domain, which - in FL-AE3p - forms the anion exchange channel (Kopito et al., 1989). Immunoblots probed with antibodies to both the NH₂- and COOH-termini of the predicted 14-AE3p sequence confirm that this alternatively spliced AE3 polypeptide is expressed in rat brain. The mobility (74 kDa) of the rat brain 14-AE3p band is slower than predicted from the 14-AE3 cDNA sequence (56 kDa), but is identical to the 14-AE3p expressed in transiently transfected COS cells, arguing that the 74 kDa band is the mobility of authentic 14-AE3p. Since 14-AE3p lacks both a consensus sequence for Asn-linked glycosylation and hydrophobic membrane-spanning segments that serve as ER insertion signals for FL-AE3p, it is unlikely that the slower mobility is due to the covalent addition of oligosaccharide residues. The presence of a high proportion of charged amino acids in this fragment (36%) may account for its anomalous migration on SDS-PAGE.

These data confirm the presence in rat brain of at least two isoforms of the AE3 gene. Full length AE3 (FL-AE3p) is present as an ~160 kDa polypeptide associated with a detergent-soluble membrane fraction. Based on its strong homology with AE1 and on its activity when expressed in heterologous cells (Kopito et al., 1989; Lee et al., 1991), FL-AE3p probably functions as a plasma membrane anion exchanger. By contrast, 14-AE3p, derived from the AE3 gene by alternate RNA splicing, is not predicted to be an integral membrane protein, as it lacks the entire membrane-spanning domain of the other AE family members. It is therefore unlikely to participate directly in ion transport, although we cannot exclude a possible role in regulating the activity of other membrane proteins. The portion of FL-AE3p to which 14-AE3p corresponds is highly polar and predicted to be cytoplasmic. We have recently shown that the NH₂-terminal cytoplasmic domain of FL-AE3p can interact with the ‘repeat’ domain of ankyrin, suggesting a possible link between FL-AE3p and the spectrin-based membrane cytoskeleton (Morgans and Kopito, 1993). The difference in detergent solubility between 14-AE3p and FL-AE3p may reflect a difference in the affinity of such an interaction. A possible role for 14-AE3p could be to modulate the interaction between integral plasma membrane proteins like FL-AE3p and the underlying membrane cytoskeleton.

REFERENCES


(Received 11 June 1993 - Accepted 26 August 1993)