Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 735–739, January 1989
Neurobiology

Alternative mRNA splicing generates multiple forms of peptidyl-glycine α-amidating monooxygenase in rat atrium

(heart/post-translational processing/peptide biosynthesis/secretory granule)

DORIS A. STOFFERS, CRISTINA BARTHEL-ROSA GREEN, AND BETTY A. EIPPER

Department of Neuroscience, The Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, MD 21205

Communicated by Paul Talalay, October 18, 1988 (received for review September 15, 1988)

ABSTRACT **Peptidyl-glycine α-amidating monooxygenase (PAM; EC 1.14.17.3) catalyzes the conversion of a variety of glycine-extended peptides into biologically active α-amidated product peptides in a reaction dependent on copper, ascorbate, and molecular oxygen. We have isolated and sequenced cDNAs representing the two major classes of PAM mRNA in the adult rat heart atrium. The two types of cDNA, rPAM-1 and rPAM-2, are identical except for the deletion of a 315-base-pair segment within the protein coding region in rPAM-2, suggesting that rPAM-1 and rPAM-2 arise by alternative splicing. Northern analysis using a cDNA probe derived from within the 315-base-pair region deleted in rPAM-2 visualized the larger of the PAM mRNAs in adult rat atrium and not the smaller, indicating that the presence or absence of this 315-nucleotide segment is a major feature distinguishing the two size forms of PAM mRNA. The 105 amino acid segment that distinguishes the two forms of atrial PAM contains a consensus N-glycosylation site and a paired basic amino acid site of potential importance in endoproteolytic processing. Comparison of the nucleotide sequences of rat, frog, and bovine PAM cDNAs reveals an extremely well conserved segment in the 3′ untranslated region. The high degree of conservation in amino acid sequence throughout the catalytic, intragranular, and cytoplasmic domains of rat atrium, bovine pituitary, and frog skin PAM suggests that both the catalytic and noncatalytic domains of the protein subserve important functions.**

Production of bioactive peptides from inactive protein precursors involves a series of post-translational modifications which occur during transit through the secretory pathway (1–3). Many peptides require an α-amidated carboxyl terminus for full biological potency. Peptidyl-glycine α-amidating monooxygenase (PAM; EC 1.14.17.3) catalyzes the conversion of carboxyl-terminal glycine-extended peptides to α-amidated product peptides in a copper-, ascorbate- and molecular oxygen-dependent reaction (4, 5).

Bovine neurointermediate pituitary (6, 7) and frog skin (8–10) PAMs have been purified and cDNAs encoding them have been cloned. The bovine PAM cDNA encodes a 108-kDa protein with a single putative transmembrane domain (7). The catalytic domain resides in the amino-terminal third of the precursor, and endoproteolytic cleavage at a subset of the 10 pairs of basic amino acid residues in the precursor may generate the soluble forms of PAM purified from the bovine neurointermediate pituitary (7). Consistent with the presence of a hydrophobic transmembrane domain in the PAM precursor, PAM activity could be solubilized from pituitary membrane fractions (11). The distribution of PAM activity between soluble and particulate fractions was found to be tissue specific (11).

Surprisingly high levels of particulate PAM activity were associated with secretory granule-enriched fractions from bovine and rat heart atrium (12). Northern analysis indicated the presence of two major forms of PAM mRNA in rat atrium (12). While the role of the atrium as an endocrine organ is now well recognized (13, 14), the function of PAM in the atrium remains to be determined; atrial natriuretic factor is not α-amidated and atrial substrates for PAM have not yet been identified.

PAM is one of the few enzymes involved in the biosynthesis of bioactive peptides that has been purified and whose cDNA has been cloned. PAM activity is regulated in a tissue-specific manner (15). In AtT-20 corticotropic tumor cells, levels of PAM mRNA vary in parallel with hormone in response to treatment with glucocorticoid or corticotropin-releasing hormone (7, 16). The broad tissue distribution of PAM activity (17, 18) and its key position in the biosynthetic pathway of many peptides makes it an attractive model for investigating the mechanisms involved in tissue-specific regulation of peptide processing.

We isolated cDNA clones encoding rat PAM to facilitate further studies of the physiological role and regulation of this enzyme. Atrium was selected as the source of RNA because high levels of at least two forms of PAM mRNA are expressed, permitting elucidation of the basis and functional consequences of the presence of these different forms. Here we report the sequences of two distinct types of cDNA encoding rat atrial PAM.* The forms of PAM mRNA presumably derive from alternative splicing of a single primary transcript. We compare the nucleotide and predicted amino acid sequences of rat PAM with those of bovine pituitary (7) and frog skin (9, 10) PAM to identify highly conserved regions of potential functional importance.

MATERIALS AND METHODS

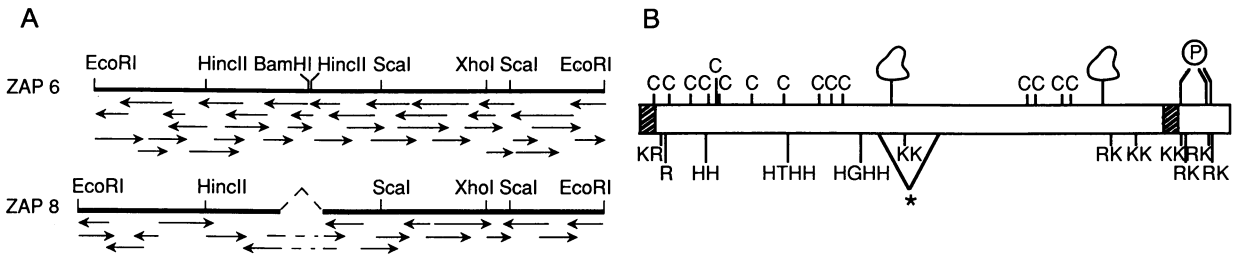
Cloning of cDNAs Encoding Rat Atrial PAM. Total RNA prepared from adult male rat atria (Sprague–Dawley) (19) was subjected to oligo(dT)-cellulose chromatography. Double-stranded cDNA was synthesized from 2.8 μg of poly(A)$^+$ RNA by using the BRL cDNA synthesis system with oligo(dT) as the first strand primer (20). cDNAs larger than 1 kilobase (kb) were selected by gel filtration on Bio-Gel A-50m (Bio-Rad), ligated to 1 μg of *Eco*RI-digested λZAP arms (Stratagene), and packaged by using the Gigapack Gold *in vitro* λ packaging kit (Stratagene). A titer of 2.9×10^6 phage was obtained when the library was plated on *Escherichia coli* BB4 cells.

For screening, 180,000 phage were plated at a density of 30,000 phage per 150-mm plate. Triplicate nitrocellulose replicas were screened with nick-translated *Eco*RI cDNA

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: PAM, peptidyl-glycine α-amidating monooxygenase; AE, amidating enzyme.

*The sequence reported in this paper is being deposited in the EMBL/GenBank data base (accession no. J04161).



C

```
   1 cgaatgatgactgacgcgggtctgcgcgatagccctcacagccccctggcattgcgagcgataaggaccctgctcctccagccctgtggctgccaccggctctggccgctctggaggGCGG
 121 GCGGCTCCAGCGGGCGGGCGGACGGACTCCAGAGCCGCTTCCCACGGCGTCGGGAGCGGACCGAGCGGACCCGGCTGGCCGCGCATCGCCCTCGCCCGGGGCCATGAAGTAGCTGCGGCCG
 241  CCAGCCCTGCCATCCGCGGCCCGGCCACTGCCCGCTGCCCTGGTCCTGCGCGGACATGGCCGGACGCGCCCGCAGCGGTCTGCTACTGCTGCTGCTGGGGCTGCTCGCCCTGCAGAGC
                                                                                                    MetAlaGlyArgAlaArgSerGlyLeuLeuLeuLeuLeuGlyLeuLeuAlaLeuGlnSer
                                                                                                                                  20
 360 AGCTGCCTGGCCTTCAGAAGCCCACTTTCTGTCTTTAAGAGGTTTAAAGAAACTACCAGATCATTTTCCAATGAATGCCTTGGTACCATTGGACCAGTCACCCCTCTTGATGCATCAGAT
     SerCysLeuAlaPheArgSerProLeuSerValPheLysArgPheLysGluThrThrArgSerPheSerAsnGluCysLeuGlyThrIleGlyProValThrProLeuAspAlaSerAsp
                                                                  40                                                          60
 480 TTTGCGCTGGATATTCGCATGCCTGGGGTTACACCTAAAGAGTCTGACACATACTTCTGCATGTCCATGCGTCTGCCTGTGGATGAGGAAGCCTTCGTGATTGACTTCAAGCCTCGTGCC
     PheAlaLeuAspIleArgMetProGlyValThrProLysGluSerAspThrTyrPheCysMetSerMetArgLeuProValAspGluGluAlaPheValIleAspPheLysProArgAla
                                                                  80                                                         100
 600 AGCATGGATACTGTCCACCATATGCTGCTGTTTGGATGCAATATGCCCTCGTCCACTGGAAGTTACTGGTTTTGTGATGAAGGAACCTGTACAGATAAAGCCAATATTCTATATGCCTGG
     SerMetAspThrValHisHisMetLeuLeuPheGlyCysAsnMetProSerSerThrGlySerTyrTrpPheCysAspGluGlyThrCysThrAspLysAlaAsnIleLeuTyrAlaTrp
                                                                 120                                                         140
 720 GCAAGGAATGCTCCCCCCACCCGGCTCCCGAAAGGTGTTGGATTCAGAGTTGGAGGAGAAACTGGAAGCAAATACTTCGTCCTTCAAGTTCACTATGGCGATATCAGTGCTTTTCGAGAT
     AlaArgAsnAlaProProThrArgLeuProLysGlyValGlyPheArgValGlyGluGluThrGlySerLysTyrPheValLeuGlnValHisTyrGlyAspIleSerAlaPheArgAsp
                                                                 160                                                         180
 840 AATCACAAAGACTGCTCTGGCGTGTCCGTACATCTCACACGTGTGCCCCAGCCTTTAATTGCGGGCATGTACCTTATGATGTCTGTTGACACTGTCATACCACCAGGAGAGAAAGTAGTG
     AsnHisLysAspCysSerGlyValSerValHisLeuThrArgValProGlnProLeuIleAlaGlyMetTyrLeuMetMetSerValAspThrValIleProProGlyGluLysValVal
                                                                 200                                                         220
 960 AATGCTGACATTTCGTGCCAATACAAAATGTATCCAATGCATGTGTTTGCCTACAGAGTCCACACTCACCATTTAGGTAAGGTGGTGAGCGGATACAGAGTAAGAAACGGACAGTGGACA
     AsnAlaAspIleSerCysGlnTyrLysMetTyrProMetHisValPheAlaTyrArgValHisThrHisHisLeuGlyLysValValSerGlyTyrArgValArgAsnGlyGlnTrpThr
                                                                 240                                                         260
1080 CTGATTGGACGCCAGAACCCCCAGCTGCCACAGGCTTTCTACCCTGTGGAACACCCCGTTGATGTTACTTTTGGTGATATACTGGCAGCCAGATGTGTGTTCACTGGTGAAGGGAGGACA
     LeuIleGlyArgGlnAsnProGlnLeuProGlnAlaPheTyrProValGluHisProValAspValThrPheGlyAspIleLeuAlaAlaArgCysValPheThrGlyGluGlyArgThr
                                                                 280                                                         300
1200 GAGGCCACCCATATCGGCGGCACTTCTAGTGACGAAATGTGTAACCTGTACATCATGTATTACATGGAAGCCAAATATGCACTTTCCTTCATGACCTGTACAAAGAACGTGGCTCCAGAT
     GluAlaThrHisIleGlyGlyThrSerSerAspGluMetCysAsnLeuTyrIleMetTyrTyrMetGluAlaLysTyrAlaLeuSerPheMetThrCysThrLysAsnValAlaProAsp
                                                                 320                                                         340
1320 ATGTTCAGAACTATCCCAGCAGAGGCCAATATCCCAATTCCTGTCAAACCGGACATGGTTATGATGCACGGGCATCACAAAGAAGCAGAAAACAAAGAAAAGAGTGCTTTAATGCAGCAG
     MetPheArgThrIleProAlaGluAlaAsnIleProIleProValLysProAspMetValMetMetHisGlyHisHisLysGluAlaGluAsnLysGluLysSerAlaLeuMetGlnGln
                                                                 360                                                         380
1440 CCAAAACAGGGAGAGGAAGAAGTATTAGAGCAGGGTGATTTCTATTCACTGCTTTCCAAGCTGCTAGGAGAAAGGGAAGATGTTCATGTGCACAAGTATAATCCTACAGAAAAGACAGAA
     ProLysGlnGlyGluGluGluValLeuGluGlnGlyAspPheTyrSerLeuLeuSerLysLeuLeuGlyGluArgGluAspValHisValHisLysTyrAsnProThrGluLysThrGlu
                                                                 400                                                         420
1560 TCTGGGTCAGACCTGGTAGCTGAGATTGCAAACGTGGTCCAGAAAAAGGACCTTGGTCGGTCTGACGCCAGAGAAGGTGCAGAGCATGAGGAATGGGGTAATGCTATCCTAGTCAGAGAC
     SerGlySerAspLeuValAlaGluIleAlaAsnValValGlnLysLysAspLeuGlyArgSerAspAlaArgGluGlyAlaGluHisGluGluTrpGlyAsnAlaIleLeuValArgAsp
                                                                 440                                                         460
1680 AGGATCCACAGATTCCACCAGCTAGAGTCAACTCTGAGGCCAGCTGAGAGCAGAGCTTTCTCGTTCCAGCAGCCTGGCGAAGGCCCTTGGGAACCAGAACCCTCAGGACATTTCCATGTG
     ArgIleHisArgPheHisGlnLeuGluSerThrLeuArgProAlaGluSerArgAlaPheSerProGlnGlnProGlyGluGlyProTrpGluProGluProSerGlyAspPheHisVal
                                                                 480                                                         500
1800 GAAGAAGAACTGGACTGGCCTGGAGTGTACTTGTTACCAGGCCAGGTTTCTGGGGTGGCCCTGGATTCTAAGAATAACCTAGTGATTTTCCACAGAGGTGACCATGTTTGGGATGGAAAC
     GluGluGluLeuAspTrpProGlyValTyrLeuLeuProGlyGlnValSerGlyValAlaLeuAspSerLysAsnAsnLeuValIlePheHisArgGlyAspHisValTrpAspGlyAsn
                                                                 520                                                         540
1920 TCTTTTGACAGCAAGTTTGTTTACCAGCAAAGAGGTCTTGGGCCAATTGAAGAAGACACCATCCTGGTCATTGACCCAAATAATGCTGAAATCCTCCAGTCCAGTGGCAAGAACCTGTTT
     SerPheAspSerLysPheValTyrGlnGlnArgGlyLeuGlyProIleGluGluAspThrIleLeuValIleAspProAsnAsnAlaGluIleLeuGlnSerSerGlyLysAsnLeuPhe
                                                                 560                                                         580
2040 TATTTACCACACGGCTTGAGCATAGATACAGATGGAAATTATTGGGTCACAGATGTGGCTCTCCACCAGGTGTTCAAATTGGACCCGCATAGCAAAGAAGGCCCTCTCTTAATTCTGGGA
     TyrLeuProHisGlyLeuSerIleAspThrAspGlyAsnTyrTrpValThrAspValAlaLeuHisGlnValPheLysLeuAspProHisSerLysGluGlyProLeuLeuIleLeuGly
                                                                 600                                                         620
2160 AGGAGCATGCAACCTGGGAGTGACCAAAATCATTTCTGCCAGCCCACCGATGTGGCTGTGGAGCCCAGTACTGGAGCTGTCTTCGTGTCAGACGGTTACTGTAACAGTCGGATTGTGCAG
     ArgSerMetGlnProGlySerAspGlnAsnHisPheCysGlnProThrAspValAlaValGluProSerThrGlyAlaValPheValSerAspGlyTyrCysAsnSerArgIleValGln
                                                                 640                                                         660
2280 TTTTCACCAAGCGGAAAGTTCGTCACCCAGTGGGGAGAAGAGTCCTCTGGAAGCAGTCCTAGGCCAGGCCAGTTCAGTGTTCCTCACAGTTTGGCCCTTGTGCCTCATTTGGACCAGTTG
     PheSerProSerGlyLysPheValThrGlnTrpGlyGluGluSerSerGlySerSerProArgProGlyGlnPheSerValProHisSerLeuAlaLeuValProHisLeuAspGlnLeu
                                                                 680                                                         700
2400 TGTGTGGCAGACAGGGAAAATGGCCGAATCCAATGCTTCAAAACTGACACCAAAGAATTTGTGAGAGAGATTAAGCACGCATCATTTGGAAGGAATGTCTTTGCCATTTCATATATACCA
     CysValAlaAspArgGluAsnGlyArgIleGlnCysPheLysThrAspThrLysGluPheValArgGluIleLysHisAlaSerPheGlyArgAsnValPheAlaIleSerTyrIlePro
                                                                 720                                                         740
2520 GGTTTCCTCTTTGCCGTAAACGGGAAGCCTTACTTTGGAGACCAAGAGCCCGTGCAAGGATTTGTGATGAACTTTTCCAGTGGGGAAATTATAGACGTCTTCAAGCCAGTACGCAAGCAC
     GlyPheLeuPheAlaValAsnGlyLysProTyrPheGlyAspGlnGluProValGlnGlyPheValMetAsnPheSerSerGlyGluIleIleAspValPheLysProValArgLysHis
                                                                 760                                                         780
2640 TTCGACATGCCTCATGATATTGTGGCTTCTGAAGATGGGACTGTGTACATTGGAGACGCACACACAAACACCGTGTGGAAGTTCACCCTGACTGAAAAAATGGAGCATCGGTCAGTTAAA
     PheAspMetProHisAspIleValAlaSerGluAspGlyThrValTyrIleGlyAspAlaHisThrAsnThrValTrpLysPheThrLeuThrGluLysMetGluHisArgSerValLys
                                                                 800                                                         820
2760 AAGGCTGGCATTGAAGTCCAGGAAATCAAAGAAGCCGAGGCAGTTGTTGAACCCAAAGTGGAGAACAAACCCACCTCCTCAGAATTGCAGAAGATGCCAAGAGAAACAGAAACTGAGCACA
     LysAlaGlyIleGluValGlnGluIleLysGluAlaGluAlaValValGluProLysValGluAsnLysProThrSerSerGluLeuGlnLysMetGlnGluLysGlnLysLeuSerThr
                                                                 840                                                         860
2880 GAGCCCGGCTCGGGAGTGTCCGTGGTTCTCATTACAACCCTTCTGGTTATTCCTGTGCTGGTCCTGCTGGCCATTGTCATGTTTATTCGGTGGAAAAAATCAAGGGCCTTTGGAGATCAT
     GluProGlySerGlyValSerValLeuIleThrThrLeuLeuValIleProValLeuValLeuLeuAlaIleValMetPheIleArgTrpLysLysSerArgAlaPheGlyAspHis
                                                                 880                                                         900
3000 GACCGCAAGCTCGAGTCAAGTTCTGGAAGAGTCCTGGGAAGATTCCGAGGAAAGGGAAGCGGCGGCTTAAATCTGGGAAATTTCTTTGCAAGTCGAAAAGGCTACAGCAGAAAAGGGTTT
     AspArgLysLeuGluSerSerGlyArgValLeuGlyArgPheArgGlyLysGlySerGlyGlyLeuAsnLeuGlyAsnPhePheAlaSerArgLysGlyTyrSerArgLysGlyPhe
                                                                 920                                                         940
3120 GACCGAGTGAGCACAGAGGGGAGTGACCAAGAGAAAGATGAGGACGACGGAACTGAGTCTGAAGAGGAGTACTCGGCCCCGCTGCCCAAGCCTGCACCTTCCTCCTGAGCCTCCAGCCTT
     AspArgValSerThrGluGlySerAspGlnGluLysAspGluAspAspGlyThrGluSerGluGluGluTyrSerAlaProLeuProLysProAlaProSerSer
                                                                 960
3240 CGCCCGGGTAGCTGGACTGAGGTTTACCAGGATGCCCAGACTCCTTCCCCTTTAGCGCGTGTAAAGTTCTGTGCATTTGATTGTAAACTGTACTCGTCAGTGTGGGACTGTACACACCTT
3360 ATTTACTTCATTTGGCTCCGTTGGCTTCTGTTCTCTAGGTGAGGAGTTCCCCACCAGTTCACTCCAGTGCCATTGTCTTTATATGAACTTAGCGTAGAGAAGCCGCCCTCCTCTTCCAAG
3480 GTAGCGCTCCAACCCCCGAGGGAAGTTTAGCTCATTCACATTTGGAGACGTTTTAGTTGGTGGATGTAAATAGCCCTATTCTCTGCTTGAACACAGTATTCTCCCAGTCCACACCCATCG
3600 CCAGTGTCTTTCTTTGGTGCCTTTCCTGTTCAGCATTCTCAGCCTGTGGCAGTGAAGAGAACCAACCTGCCACACGACGAAAAGCTGCTAAATCTCCTTCTATTTTTTTAAAATCACTAA
3720 CATTATATTGCAATGAGAGAAATTTTAAAAAGTCTCTATTTAAATTCTTTTTTTAAATTTCTCCTCAGTTGGTGTGTTTCCGGGATGTCTTATTTTTAGATGGTTACACTGTTAGAACAC
3840 TATTTTTCAGAATCTGAATGTAATTTGTGTAATAAAGTGTTTTCAGA
```

FIG. 1. Nucleotide and predicted protein sequence for rat atrium PAM. (*A*) Diagram outlining the relationship between ZAP 6 and ZAP 8 clones, selected restriction endonuclease sites, and the sequencing strategy. (*B*) Diagram outlining key features of the proteins encoded by rPAM-1 and -2 (missing region denoted by *): hatched areas, putative signal sequence and transmembrane domain; C, cysteine residues; irregular closed curves, potential glycosylation sites; Ⓟ, potential phosphorylation site; K (lysine) and R (arginine), potential endoproteolytic cleavage sites; histidine-rich regions are indicated by H with other single-letter amino acid symbols. (*C*) cDNA sequence of ZAP 6 (uppercase letters) and deduced protein sequence. The 315-bp region missing in ZAP 8 is enclosed in brackets. The additional 116 nucleotides present at the 5′ end of ZAP 8 (lowercase letters) have been added to the 5′ end of the ZAP 6 sequence. The highly conserved 86-bp region of the 3′ end is underlined. The putative polyadenylylation signal is boxed.

fragments spanning the entire bovine PAM cDNA [base pairs (bp) 1–781, 782–1503, and 1504–3724] (7). Of 252 positive plaques identified by the 2.2-kb fragment, 55 hybridized to all three probes. Twenty of these 55 were plaque-purified, and

18 were rescued by R408 helper phage (Stratagene) as positive cDNA insert-containing Bluescript (SKM13−) plasmids. Plasmid DNA was characterized by restriction mapping. By restriction analysis with *Bam*HI and *Hin*cII, the plasmids were separated into two major classes, rPAM-1 (7 of 18) and rPAM-2 (5 of 18). One insert from each class was chosen for further characterization: ZAP 6 (rPAM-1) and ZAP 8 (rPAM-2).

Nested sets of deletions extending inward from the 5′- and 3′-termini of ZAP 6 were generated with an exonuclease III/mung bean nuclease kit (Stratagene). Fragments subcloned in Bluescript and inserts bearing deletions generated by using unique restriction sites were used to direct sequencing to remaining regions of ZAP 6 and to sequence ZAP 8. ZAP 6 was sequenced on both strands in its entirety by the dideoxy chain termination method (21) using the Sequenase kit (United States Biochemical). ZAP 8 was sequenced on at least one strand in its entirety and on both strands over the presumptive splice junction and over the additional 117 bases at its 5′ end.

Northern Blot Analysis. Total atrial RNA was prepared, fractionated on 1.0% agarose gels containing formaldehyde, transferred to Nytran (Schleicher & Schuell), and analyzed as described (7). cDNA probes were derived from bovine PAM cDNA (*Eco*RI/*Eco*RI bp 782–1503 fragment) and rat PAM ZAP 6 cDNA (*Xmn* I/*Hin*cII bp 1517–1708 fragment). Nick-translated probes were added to the prehybridization solution (10^6 cpm/ml) and hybridization continued at 42°C for 24–36 hr. Molecular weights were estimated by comparison with an RNA "ladder" (BRL) and with the positions of ribosomal RNA bands.

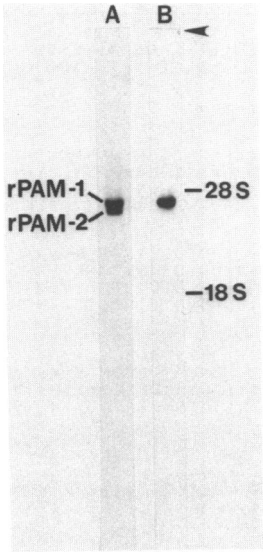


FIG. 2. Northern analysis of adult rat atrium RNA. Ten micrograms of total atrium RNA was fractionated on a denaturing agarose gel as described (7). After hybridization with the bovine PAM cDNA fragment and exposure to film (64 hr, −70°C, with intensifying screen), the filter was stripped and reprobed with the rPAM-1-specific probe (24 hr, −70°C, with intensifying screen). The point of sample application is marked by an arrowhead. With the use of longer denaturing gels, the apparent sizes of the two major forms of atrial PAM mRNA are 4.2 ± 0.1 kb and 3.8 ± 0.1 kb instead of 3.8 ± 0.1 kb and 3.6 ± 0.1 kb (12).

RESULTS

We have cloned and sequenced two distinct cDNAs for PAM from an adult rat atrium cDNA library. The first cDNA (ZAP 6), which is of the rPAM-1 type, is 3770 bp long, contains a 180-bp G+C-rich (79%) 5′ untranslated region followed by a single 2931-bp open reading frame, and ends with a 659-bp A+T-rich (57%) 3′ untranslated region (Fig. 1*C*). The nucleotide sequence around the initiator methionine satisfies the criteria set forth by Kozak (22). Although this cDNA clone does not contain a poly(A) tail, it does have a consensus polyadenylylation signal (Fig. 1*C*) and ends only five nucleotides before the poly(A) tail in the bovine PAM cDNA (7). The second cDNA (ZAP 8), which is of the rPAM-2 type, is 3551 bp long; it contains 20 fewer nucleotides at its 3′ end and an additional 116 nucleotides at its 5′ end (lowercase letters in Fig. 1*C*). In addition, ZAP 8 is missing 315 nucleotides that are found in ZAP 6 (bp 1475–1789); this deletion leads to the disappearance of several restriction sites (Fig. 1*A*). Sequence analysis shows no other differences between the two major types of cDNA.

Northern analysis of adult rat atrium RNA was carried out to determine whether the two types of cDNA identified correspond to the major species of PAM mRNA observed (Fig. 2). Hybridization with a bovine PAM cDNA probe derived from a region common to both types of rat PAM cDNA revealed the expected two major forms of PAM mRNA (4.2 ± 0.1 kb and 3.8 ± 0.1 kb). When the same blot was stripped and rehybridized with a probe specific for rPAM-1, only the larger of the rPAM mRNAs was recognized, suggesting that this 315-nucleotide deletion accounts in part for the difference between the two major size forms of atrial PAM mRNA.

The rPAM-1 type cDNA encodes a 976 amino acid protein of 108,650 Da (Fig. 1*B*). A hydrophobic segment at the amino terminus satisfies the consensus rules for signal sequences, with cleavage most likely to occur after amino acid 25 (23). The signal sequence is followed by a 10 amino acid putative propeptide [by analogy with the bovine enzyme (7)], an 831 amino acid catalytic/intragranular domain, a 24 amino acid hydrophobic putative transmembrane domain, and an 86 amino acid hydrophilic putative cytoplasmic tail. Notable features include the presence of 8 pairs of basic amino acid residues, 3 clusters of histidine residues, 2 potential sites for N-glycosylation, 14 cysteine residues in the catalytic/intragranular domain, and 3 potential phosphorylation sites in the putative cytoplasmic tail (Fig. 1*B* and Fig. 3). The rPAM-2 type cDNA encodes an 871 amino acid protein of 96,800 Da, which is missing 105 amino acids present in the catalytic/intragranular domain of the rPAM-1 type precursor (amino acids 393–497 of rPAM-1). The deleted region contains a pair of basic amino acids (Lys-Lys at positions 436–437), 9 single arginine residues, and 1 of the potential N-glycosylation sites. The position of the deletion in rPAM-2 corresponds to the region missing from frog AE-II when it is compared to rPAM-1 and bovine λPAM-1. FASTP searches (24) of the National Biomedical Research Foundation Protein Identification Resource data base on May 19, 1988, identified no proteins with significant homology to the entire rat PAM sequence or to this 105 amino acid segment.

With the sequence of PAM cDNAs from three species, it is possible to identify features conserved at the nucleotide and amino acid level. Comparisons here are made between rPAM-1 and bovine λPAM-1 and between rPAM-2 and frog AE-II. The nucleotide identity is 80% between rat and bovine PAM and 60% between rat and frog PAM. Conservation is most marked at the extreme 3′ end of the 3′ untranslated region, with 98% identity between rat and bovine sequences in an 86-bp segment (Fig. 1*C*) (82% identity for rat/frog). In the protein coding region, the nucleotide sequences of rat and bovine PAM exhibit 84% identity, while rat and frog exhibit 66% identity. The nucleotide sequence in the 5′ untranslated region of the PAM cDNAs shows the least conservation (57% identity for rat/bovine; 39% for rat/frog).

At the amino acid level, sequence conservation is fairly uniform throughout the protein, with overall identity of 88% between rat and bovine PAM and 66% between rat and frog

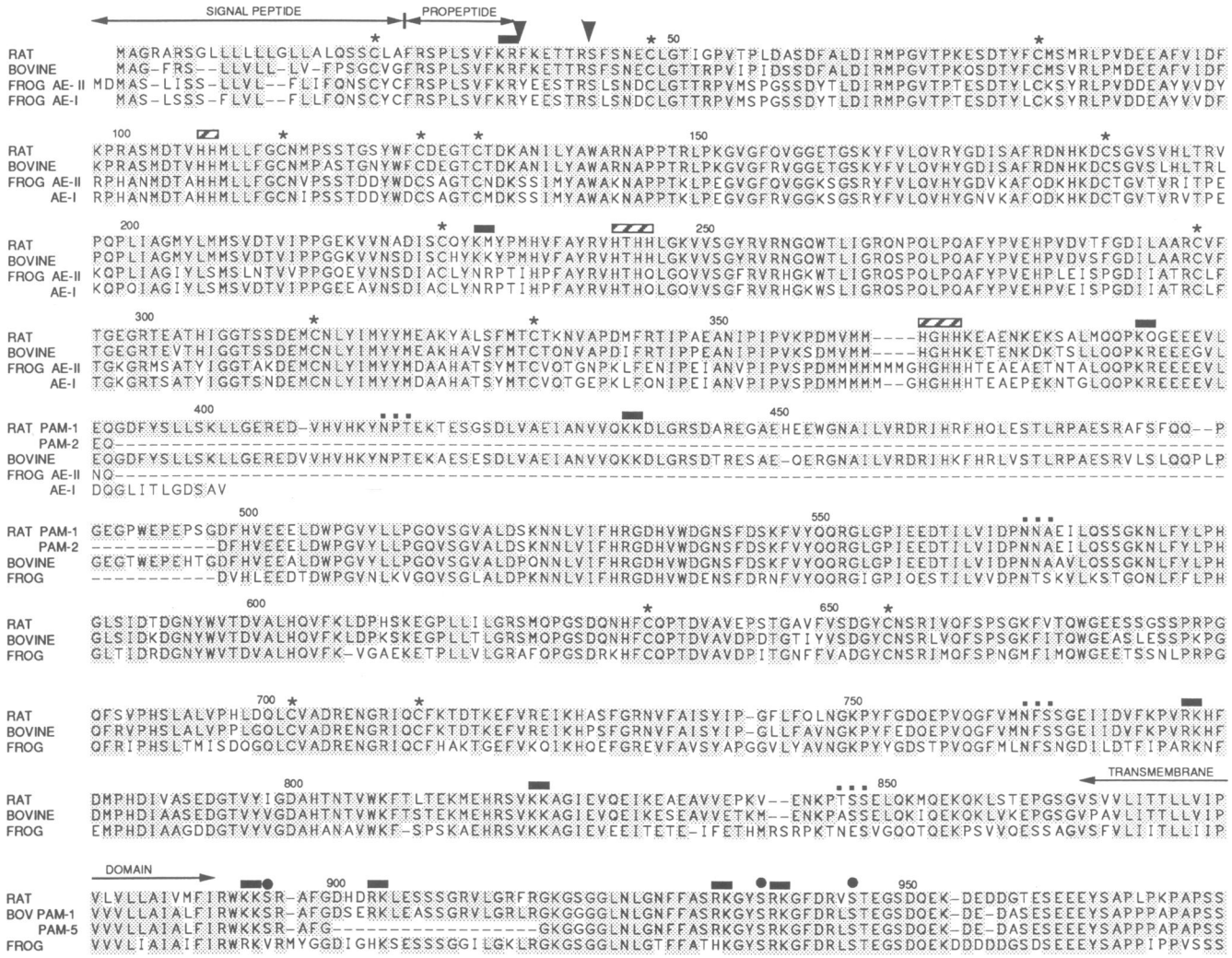


FIG. 3. Comparison of rat, bovine, and frog PAMs. The sequences of the proteins encoded by rat atrium rPAM-1 and rPAM-2, bovine intermediate pituitary λPAM-1 and λPAM-5 (7), and frog skin amidating enzyme (AE)-I and AE-II (9, 10) are compared. The numbering given is that for the rPAM-1 sequence. Amino acid symbols are shaded only if they are identical in at least two of the three species. The positions of several conserved features are noted: *, cysteine residues; ▨, clusters of histidine residues; ▬, pairs of basic amino acids; ■■■, potential sites for N-glycosylation; and ●, potential sites for phosphorylation. Initial alignment was pairwise by computer (Microgenie); final alignment was manual.

PAM. The signal sequences are relatively poorly conserved (60% identity for rat/bovine). In contrast, the sequence of the putative propeptide is completely conserved in all three species. Soluble PAM purified from bovine neurointermediate pituitary included proteins with amino-terminal sequences beginning immediately after the Lys-34–Arg-35 sequence at the carboxyl terminus of the propeptide domain and proteins beginning after Arg-41 (arrowheads in Fig. 3) (7). In the amino-terminal half of the catalytic/intragranular domain, which constitutes the catalytic domain, amino acid sequence identity is 90% between rat and bovine and 65% between rat and frog. In the carboxyl-terminal half of the catalytic/intragranular domain, sequence identity is similar (85% and 67%, respectively). The amino acid sequence of the putative transmembrane domain is substantially better conserved than that of the signal sequence (80% rat/bovine; 71% rat/frog), and the hydrophilic putative cytoplasmic tail is also well conserved (82% rat/bovine; 71% rat/frog).

Conservation of several specific features is notable (Fig. 3). The positions of all 14 cysteine residues within the catalytic/intragranular domain are conserved in all three species; these residues may be important in maintaining the conformation of the enzyme within the secretory granule or at the cell surface after secretion. The two clusters of histidine residues postulated to play a role in the interaction of bovine PAM with copper (7) are conserved in all three species, as is the additional His-His sequence at residues 207–208.

Endoproteolytic cleavage at paired basic amino acids was postulated to play a role in the tissue-specific generation of soluble PAM (7). Eight of the 10 pairs of basic amino acids in the bovine PAM precursor are conserved in rPAM-1. The potential endoproteolytic cleavage sites not conserved include bovine PAM Lys-225–Lys-226, which is absent from rat and frog PAM, and bovine PAM Lys-379–Arg-380, which is conserved in frog but not in rat PAM. Cleavage of the bovine PAM precursor at the latter site was postulated to generate a 38,000-Da form of PAM in bovine pituitary (PAM-B) (7). Consistent with this species specificity, a 39,000-Da form of PAM has also been purified from frog skin (8), while PAM activity purified from a rat medullary thyroid carcinoma has a mass of 75,000 Da (25). The Lys-432–Lys-433 sequence postulated to serve as the endoproteolytic cleavage site for production of bovine pituitary PAM-A is conserved in rPAM-1 but is in the region deleted from rPAM-2 and frog PAM. Proteolytic cleavage of bovine pituitary PAM also appears to occur after single arginine residues (7). Endoproteolytic cleavage of this type may be particularly relevant in the atrium, since production of atrial natriuretic factor involves

cleavage after a single arginine residue (14); the region deleted in rPAM-2 contains nine single arginine residues. The remaining paired basic sites in the intragranular domain are conserved in rat, bovine, and frog PAM, as are three of the four paired basic amino acid sequences in the putative cytoplasmic domain.

Two of the three potential phosphorylation sites in the putative cytoplasmic tail (Ser-937 and Ser-945) are conserved in all three species, while the third site (Ser-895) is conserved only between rat and bovine (7). Rat PAM-1 contains both potential N-glycosylation sites found in bovine PAM (Asn-415 and Asn-765); the first of these sites is in the region deleted in rPAM-2 and frog PAM, while the second site is conserved in frog PAM.

DISCUSSION

Two types of cDNA, rPAM-1 and rPAM-2, were identified in a rat atrium cDNA library. The ratio between the number of clones of each type (rPAM-1/rPAM-2 = 1.4) correlates well with the relative levels of the two major forms of PAM mRNA seen on Northern blots of adult rat atrium (Fig. 2; ref. 12). cDNAs of each type were sequenced and differ by an in-frame deletion of a 315-bp segment from the protein coding region of rPAM-1. Southern analysis of rat and bovine genomic DNA indicates that a single complex gene encodes PAM (60–80 kb and at least eight introns in the bovine PAM gene; D.A.S., unpublished observations). Taken together, these results strongly suggest that one gene gives rise to both mRNAs by alternative splicing. Use of multiple sites of initiation and polyadenylylation as well as differing extents of polyadenylylation requires further investigation. The two types of PAM cDNA found in bovine neurointermediate pituitary are also probably products of alternative RNA splicing; a 54-bp segment in the region coding for the cytoplasmic domain of λPAM-1 is deleted in λPAM-5 (7). In contrast, the two types of PAM cDNA found in the frog must be products of separate genes (9, 10). The exact alignment of rPAM-2 and frog AE-II suggests conservation of the exon/intron structure of the gene encoding PAM in this region. Alternative splicing is a frequent means of generating protein diversity in postmitotic tissues such as muscle and nerve (26), and it is anticipated that additional forms of PAM mRNA will be identified in both rat and bovine tissues.

The PAM precursors encoded by the two rat cDNAs should differ from each other and from bovine and frog PAM in their ability to generate soluble PAM activity. It is possible that the rat PAM precursors yield enzymes with different substrate specificities and catalytic properties. Alternative splicing may provide a mechanism whereby a single gene can encode a processing enzyme utilized by many different neuroendocrine cells to produce a variety of amidated peptides and respond to multiple regulatory inputs. Consistent with this prediction, Western blot analysis of membrane-associated atrial PAM reveals a complex pattern of bands of molecular mass 94 to 125 kDa, and the forms of PAM mRNA expressed during atrial and ventricular development undergo a complex series of tissue-specific and developmentally regulated changes (31).

The marked conservation of the 86-bp segment in the 3′ untranslated region of the PAM cDNAs examined suggests that this region has a specific function. Sequences within the 3′ untranslated region may affect mRNA stability (27, 28) and translational efficiency (29, 30). The relatively high degree of conservation throughout the rat, bovine, and frog PAM proteins implies that the catalytic and noncatalytic domains of the molecule are likely to serve specific functions. The cytoplasmic tail might serve a regulatory role; e.g., phosphorylation of a site in the cytoplasmic domain might regulate the activity of the catalytic domain or subcellular routing through the conserved membrane-spanning and intragranular domains. When the structure of the gene encoding PAM has been elucidated it will be possible to identify additional mRNA species and proteins derived from this gene and examine their tissue-specific expression and regulation.

We thank Richard Mains, L'Houcine Ouafik, and Ian Dickerson for their advice and encouragement. This work was supported by grants from the National Institutes of Health (DK32949 and GM07309) and the National Institute on Drug Abuse (DA00098).

1. Eipper, B. A., May, V., Cullen, E. I., Sato, S. M., Murthy, A. S. N. & Mains, R. E. (1987) in *Psychopharmacology: The Third Generation of Progress*, ed. Meltzer, H. Y. (Raven, New York), pp. 385–400.
2. Docherty, K. & Steiner, D. F. (1982) *Annu. Rev. Physiol.* **44,** 625–638.
3. Gainer, H., Russell, J. T. & Loh, Y. P. (1985) *Neuroendocrinology* **40,** 171–184.
4. Eipper, B. H. & Mains, R. E. (1988) *Annu. Rev. Physiol.* **50,** 333–344.
5. Bradbury, A. F. & Smyth, D. G. (1987) *Eur. J. Biochem.* **169,** 579–584.
6. Murthy, A. S. N., Mains, R. E. & Eipper, B. A. (1986) *J. Biol. Chem.* **261,** 1815–1822.
7. Eipper, B. A., Park, L. P., Dickerson, I. M., Keutmann, H. T., Thiele, E. A., Rodriguez, H., Schofield, P. R. & Mains, R. E. (1987) *Mol. Endocrinol.* **1,** 777–790.
8. Mizuno, K., Sakata, J., Kojima, M., Kangawa, K. & Matsuo, H. (1986) *Biochem. Biophys. Res. Commun.* **137,** 984–991.
9. Mizuno, K., Ohsuye, K., Wada, Y., Fuchimura, K., Tanaka, S. & Matsuo, H. (1987) *Biochem. Biophys. Res. Comun.* **148,** 546–552.
10. Ohsuye, K., Kitano, K., Wada, Y., Fuchimura, K., Tanaka, S., Mizuno, K. & Matsuo, H. (1988) *Biochem. Biophys. Res. Commun.* **150,** 1275–1281.
11. May, V., Cullen, E. I., Braas, K. M. & Eipper, B. A. (1988) *J. Biol. Chem.* **263,** 7550–7554.
12. Eipper, B. A., May, V. & Braas, K. M. (1988) *J. Biol. Chem.* **263,** 8371–8379.
13. deBold, A. J. (1985) *Science* **230,** 767–770.
14. Needleman, P. & Greenwald, J. E. (1986) *N. Engl. J. Med.* **314,** 828–834.
15. Mains, R. E., Myers, A. C. & Eipper, B. A. (1985) *Endocrinology* **116,** 2505–2515.
16. Mains, R. E. & Eipper, B. A. (1984) *Endocrinology* **115,** 1683–1690.
17. Eipper, B. A., Myers, A. C. & Mains, R. E. (1985) *Endocrinology* **116,** 2497–2504.
18. Sakata, J., Mizuno, K. & Matsuo, H. (1986) *Biochem. Biophys. Res. Commun.* **140,** 230–236.
19. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299.
20. Gubler, U. & Hoffman, B. J. (1983) *Gene* **25,** 263–269.
21. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
22. Kozak, M. (1986) *Cell* **44,** 283–292.
23. von Heijne, G. (1985) *J. Mol. Biol.* **184,** 99–105.
24. Lipman, D. J. & Pearson, W. R. (1985) *Science* **227,** 1435–1441.
25. Mehta, N. M., Gilligan, J. P., Jones, B. N., Bertelsen, A. H., Roos, R. A. & Birnbaum, R. S. (1988) *Arch. Biochem. Biophys.* **261,** 44–54.
26. Andreadis, A., Gallego, M. E. & Nadal-Ginard, B. (1987) *Annu. Rev. Cell Biol.* **3,** 207–242.
27. Caspasso, O., Bleecker, G. C. & Heintz, N. (1987) *EMBO J.* **6,** 1815–1831.
28. Shaw, G. & Kamen, R. (1986) *Cell* **46,** 659–667.
29. Kruys, V., Wathelet, M., Poupart, P., Contreras, R., Fiers, W., Content, J. & Huez, G. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 6030–6034.
30. Dickey, L. F., Wang, Y.-H., Shull, G. E., Wortman, I. A., III, & Theil, E. C. (1988) *J. Biol. Chem.* **263,** 3071–3074.
31. Ouafik, L'H., May, V., Keutmann, H. T., & Eipper, B. A. (1989) *J. Biol. Chem.* **264,** in press.