

# Alternative usage of different poly(A) addition signals for two major species of mRNA encoding human aromatase P-450

Katsumi Toda, Masako Terashima\*, Yasuhiro Mitsuuchi, Yasuhisa Yamasaki, Yuichi Yokoyama<sup>+</sup>, Shunji Nojima<sup>°</sup>, Hiroshi Ushiro, Tomoho Maeda\*, Yasutake Yamamoto<sup>+</sup>, Yusuke Sagara<sup>°</sup> and Yutaka Shizuta

*Departments of Medical Chemistry, \*Radiology, <sup>+</sup>Internal Medicine and <sup>°</sup>Obstetrics and Gynecology, Kochi Medical School, Nankoku, Kochi 781-51, Japan*

Received 4 February 1989; revised version received 8 March 1989

Two cDNA clones for human placental aromatase P-450 (P-450<sub>AROM</sub>) have been isolated and sequenced. The insert of one clone (2894 bp) contains an open reading frame encoding a protein consisting of 503 amino acid residues together with a 49 bp 5'-untranslated stretch and a 1336 bp 3'-noncoding region to which a poly(A) tract is attached. Three potential poly(A) addition signals are detected in this 3'-noncoding region. The other clone contains a shorter cDNA insert, the nucleotide sequence of which overlaps with most of the sequence of the longer cDNA insert (nucleotides 36-2355) except for one nucleotide substitution. The 3'-noncoding region of this shorter cDNA is only 846 bp in length, but a poly(A) tract is also attached to its 3'-terminus. Northern blot analysis of human placental RNA reveals the presence of two major mRNA species of 3.4 and 2.9 kb when probes excised from the overlapping region of these two cDNAs are employed. The 2.9 kb mRNA is not detected, however, when a fragment of the non-overlapping region of the longer cDNA is used as a probe. It is therefore concluded that the two major species of P-450<sub>AROM</sub> mRNA are formed as a consequence of alternative processing of precursor mRNA(s).

Monooxygenase; Aromatase P-450; Estrogen synthesis; cDNA; Poly(A) addition signal; (Human placenta)

## 1. INTRODUCTION

The conversion of androgens to estrogens is catalyzed by a specific monooxygenase [1], called aromatase P-450 (P-450<sub>AROM</sub>) or estrogen synthetase [2-4]. The mechanism of P-450<sub>AROM</sub> expression has received much attention, because various androgens are converted to different species of estrogen in a tissue-specific fashion [5-7] and the P-450<sub>AROM</sub> expression is regulated by a variety of factors such as cyclic AMP, glucocorticoids and growth factors [8].

*Correspondence address:* Y. Shizuta, Department of Medical Chemistry, Kochi Medical School, Nankoku, Kochi 781-51, Japan

The nucleotide sequence presented here has been submitted to the EMBL/GenBank database under the accession no. Y07508

Based on the tissue-specific synthesis of different estrogens, Osawa et al. [9] have suggested the existence of two or more forms of P-450<sub>AROM</sub> differing in substrate specificities. Recent cloning and sequencing of partial-length cDNAs for human P-450<sub>AROM</sub> have led to the suggestion that there exist three different P-450<sub>AROM</sub> mRNAs [10] or at least two P-450<sub>AROM</sub> genes [11]. More recently, full-length cDNAs for P-450<sub>AROM</sub> have been isolated from chicken ovarian [12] and human placental [13,14] cDNA libraries and sequenced.

As a first step to investigate estrogen-dependent diseases in humans, we here report the isolation and sequencing of two P-450<sub>AROM</sub> cDNA clones derived from human placenta. The results provide evidence that the occurrence of two major species of mRNA for P-450<sub>AROM</sub> is mainly ascribed to the alternative RNA processing using different poly(A) addition signals.

## 2. MATERIALS AND METHODS

### 2.1. Materials

The oligo(dT)-primed human placental cDNA library constructed in  $\lambda$ gt11 from poly(A)<sup>+</sup> RNA [15] was kindly supplied by Dr Y. Ebina, University of Tokushima. Two oligonucleotide probes (44-mer, 5'-AG CAC CAG GAT CAG GTA GTG GAT GGG GTT CAG CAT CTC CAG CAC-3'; 35-mer, 5'-TT CTG GAT GCT CTC CAC ACA CTG GCC CTG CAG GGT-3') were synthesized by the modified triester method [16]. The base sequences of both probes were designed on the basis of partial amino acid sequences of human placental P-450<sub>AROM</sub> [17], using published data for codon usage frequencies [18]. All other materials including restriction endonucleases, [ $\alpha$ -<sup>32</sup>P]dCTP and [ $\gamma$ -<sup>32</sup>P]ATP were obtained from commercial sources.

### 2.2. Cloning procedures

The cDNA library on *E. coli* Y1088 was screened by hybridization with the <sup>32</sup>P-labeled oligonucleotide probes. Nitrocellulose filters containing plaque replicas were hybridized at 38°C with the 35-mer probe for at least 14 h in 3 × SSC, 10 × Denhardt's solution containing denatured herring DNA (100 µg/ml). With the 44-mer as probe, hybridization was performed at 40°C. The filters were washed at the same

temperature as that used for hybridization for 20 min with 6 × SSC containing 0.1% SDS, and then with 3 × SSC containing 0.1% SDS. They were further washed at 55°C for 10 min with 3 × SSC containing 0.1% SDS.

### 2.3. Nucleotide sequencing

The cDNAs were digested with appropriate restriction endonucleases and the fragments were subcloned into M13mp18 or 19 [19]. Nucleotide sequencing was conducted by the dideoxy chain-termination method [20] according to the strategy shown in fig.1.

### 2.4. Northern blot analysis

Total RNA was extracted from human placenta as described by Chirgwin et al. [21]. Poly(A)<sup>+</sup> RNA was enriched therefrom by oligo(dT)-cellulose chromatography [22]. Total RNA (20 µg) or poly(A)<sup>+</sup> RNA (4 µg) was subjected to agarose gel electrophoresis and transferred to nitrocellulose filters [23,24]. Hybridization was carried out using <sup>32</sup>P-labeled probes (probes A-C in fig.1). The DNA size markers used were fragments of pUC119 DNA cut with *Dra*I (2454 bp), *Pvu*II (2840 bp) or *Hin*dIII (3162 bp) and linearized plasmid DNA (4009 bp) consisting of pUC118 DNA (3162 bp) and the *Eco*RI-*Kpn*I fragment of pES-1 (847 bp).

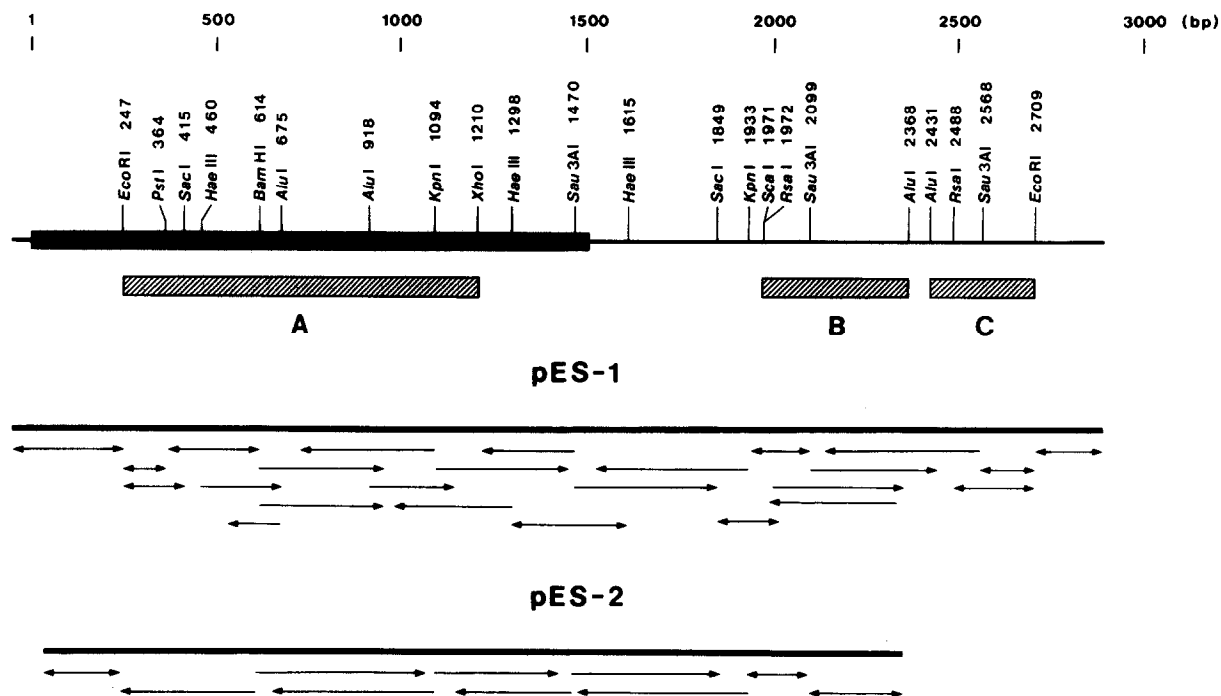
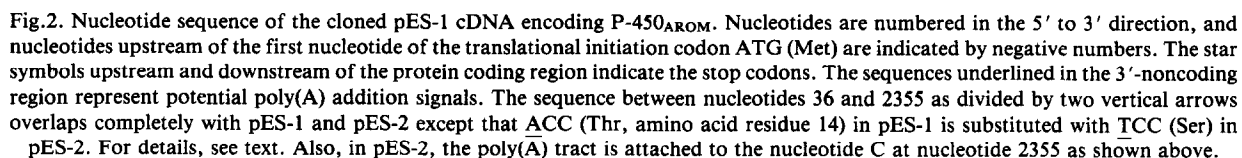


Fig.1. Restriction map of and sequencing strategy for the cloned cDNAs encoding human placental P-450<sub>AROM</sub>. The restriction map displays only relevant restriction endonuclease sites with numbers which indicate the 5'-terminal nucleotide position generated by cleavage (for nucleotide numbering, see fig.2). The protein coding region is indicated by a closed box. The hatched boxes represent the restriction fragments of the cDNA inserts employed as hybridization probes for the experiments in fig.3: (A) *Eco*RI(247)-*Xho*I(1210) fragment; (B) *Sac*I(1971)-*Alu*I(2368) fragment; (C) *Alu*I(2431)-*Eco*RI(2709) fragment. The direction and the extent of sequence determinations are indicated by horizontal arrows under each cDNA insert.



### 3. RESULTS

Screening of the human placental cDNA library in  $\lambda$ gt11 ( $\sim 5 \times 10^4$  plaques) with the 35-mer probe resulted in the detection of about 200 positive clones. Characterization of 5 randomly chosen clones indicated that one clone contained a 2.4 kbp insert that resembled those reported by Simpson et al. [10] and Chen et al. [11]. Upon rescreening of the same library with an *EcoRI-BamHI* fragment of the 2.4 kbp insert as a probe, 198 positive clones were obtained. Of these, 23 clones were randomly isolated and selected with the 44-mer synthetic probe. It was thus possible to isolate two cDNA clones, pES-1 and pES-2, containing 2.9 and 2.3 kbp inserts, respectively.

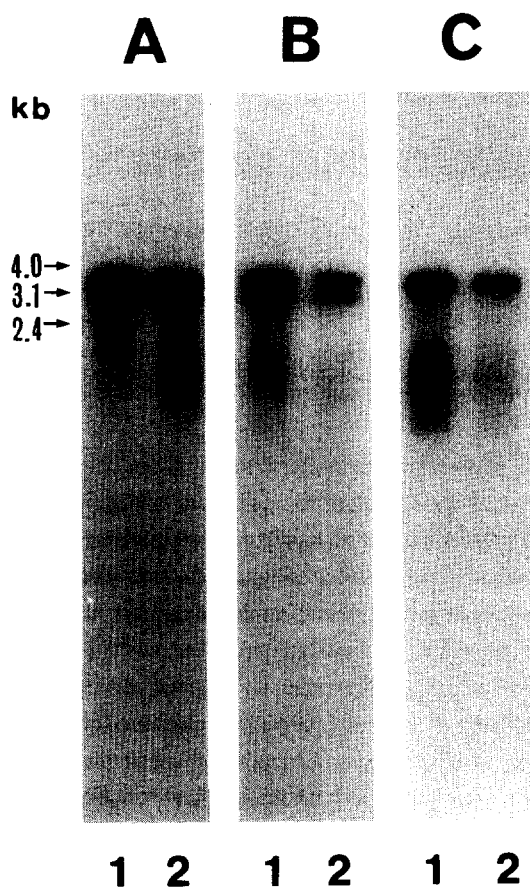


Fig.3. Autoradiograms of blot hybridization analysis of human placental RNA. Total RNA (1) or poly(A)<sup>+</sup> RNA (2) was hybridized with probes A-C as described in the legend to fig.1. The horizontal arrows with numerals indicate the position of nucleotide size markers (see section 2).

Fig.1 shows the restriction map and sequencing strategy for the two cDNAs cloned. The nucleotide sequence determined for the pES-1 insert, which consists of 2894 bp (excluding the poly(A) tract), is presented in fig.2. The translation initiation site of pES-1 cDNA can be assigned as the methionine codon at nucleotides 1-3, because this is the first ATG codon appearing downstream of the stop codon TAA (nucleotides -33 to -31) that is in-frame with the largest open reading frame. This is also supported by the fact that the sequence around the ATG triplet agrees well with the favored sequence that flanks the functional initiation codon:  $\hat{A}XXAUGG$ , where X is any nucleotide [25]. The open reading frame starting from this initiation codon consists of 1509 nucleotides (encoding a polypeptide of 503 amino acid residues) followed by the stop codon TAG (nucleotides 1510-1512). A poly(A) tract is attached to the 3'-terminus of the 1336 bp 3'-noncoding region of this cDNA. On the other hand, the cDNA insert of pES-2 is shorter (2320 bp) and corresponds to the segment of pES-1 cDNA from nucleotides 36 to 2355 (see legend to fig.2). The two nucleotide sequences are identical with each other except for a single base substitution (A  $\rightarrow$  T) at nucleotide 40 in pES-2 cDNA. An important finding is that a poly(A) tract is also attached to the 3'-terminus of pES-2 cDNA, even though its 3'-noncoding region is only 846 bp in length. Moreover, three potential poly(A) addition signals (underlined in fig.2) can be detected in the 3'-noncoding region of pES-1 cDNA, but the third signal is lacking in pES-2 cDNA. The primary structure of the polypeptide deduced from the cDNA sequence is also shown in fig.2.

To determine the size of mRNA(s) encoding P-450<sub>AROM</sub>, total RNA and poly(A)<sup>+</sup> RNA from human placenta were analyzed by Northern blot hybridization using probes A-C, which had been cut out of pES-1 cDNA (see fig.1). When probe A or B was used, two major mRNA bands of 3.4 and 2.9 kb were detected (fig.3). The 2.9 kb mRNA was not detected, however, when a fragment of the non-overlapping region of pES-1 cDNA (probe C) was used. These results indicate that pES-1 and pES-2 cDNAs are derived from the 3.4 and 2.9 kb mRNAs, respectively, and that these two major species of P-450<sub>AROM</sub> mRNA are formed as a consequence of alternative RNA processing using two different poly(A) addition signals.

#### 4. DISCUSSION

A comparison of the nucleotide sequences of pES-1 and pES-2 cDNAs with those of human P-450<sub>AROM</sub> cDNAs as determined by other workers [10,11,13,14] indicates that all the sequences are very similar to one another in the protein coding region; changes detectable are substitutions of at most 2 nucleotides. However, nucleotide substitutions, deletions and additions are more marked (up to 14 nucleotides) in the 3'-noncoding region of the cDNAs (Toda, K., unpublished). The minor differences in the coding region are probably due to allelic variants, but the reason for significant changes in the 3'-noncoding region is unclear.

Evans et al. [26] reported the occurrence of three mRNAs (3.0, 2.7 and 2.4 kb) for human placental P-450<sub>AROM</sub>. They also isolated two P-450<sub>AROM</sub> cDNAs of different length and suggested that the two major mRNAs are formed by differential RNA processing [10]. Chen et al. [11] also reported a partial sequence of human P-450<sub>AROM</sub> cDNA and mapped the corresponding gene on chromosome 15. They further suggested the possibility that there exist at least two different P-450<sub>AROM</sub> genes and that the two major mRNAs in human placenta are probably derived from the different genes. More recently, Harada [14] has suggested that different transcription initiation or alternative splicing of the 5'-untranslated region of precursor mRNA is responsible for the formation of two major species of P-450<sub>AROM</sub> mRNA in human placenta.

Here, we have obtained direct evidence that the occurrence of two major species of P-450<sub>AROM</sub> mRNA is mainly ascribed to alternative RNA processing using different poly(A) addition signals. It should be noted, however, that we have also obtained a unique cDNA clone; a portion of its 5'-untranslated region (490 bp) is different from that reported by Harada [14] (Toda, K. et al., unpublished). Although nothing is known of the transcription start site of the P-450<sub>AROM</sub> gene, the existence of this unique cDNA clone may support Harada's suggestion [14] to a certain extent. Furthermore, we have isolated another type of P-450<sub>AROM</sub> cDNA clone containing a 4 kbp insert and found that the nucleotide sequence of its protein coding region is identical with that of pES-1 cDNA, but its 3'-noncoding region is markedly different from that of pES-1 cDNA (Toda, K. et al.,

in preparation). This clone may correspond to the P-450<sub>AROM</sub> cDNA clones having 4 kbp inserts recently isolated by McPhaul et al. [12] and by Corbin et al. [13]. These findings suggest that the expression of human placental P-450<sub>AROM</sub> is more complicated than previously anticipated. It is, therefore, absolutely necessary to isolate and characterize genomic DNA(s) for P-450<sub>AROM</sub> in order to elucidate the regulation of its expression.

*Acknowledgements:* We are grateful to Dr Y. Ebina of the University of Tokushima for providing the cDNA library, Dr Y. Osawa of Medical Foundation of Buffalo, Inc., for helpful advice on this work, and M. Horiuchi and N. Isobe for secretarial assistance. This work was supported in part by Toray Science Foundation and by Grants-in-Aid for Science and Cancer Research from the Ministry of Education, Science and Culture of Japan.

#### REFERENCES

- [1] Hayaishi, O. (1974) *Molecular Mechanisms of Oxygen Activation*, Academic Press, New York.
- [2] Thompson, E.A. jr and Siiteri, P.K. (1974) *J. Biol. Chem.* 249, 5364-5372.
- [3] Thompson, E.A. jr and Siiteri, P.K. (1974) *J. Biol. Chem.* 249, 5373-5378.
- [4] Fishman, J. and Goto, J. (1981) *J. Biol. Chem.* 256, 4466-4471.
- [5] Nimrod, A. and Ryan, K.J. (1975) *J. Clin. Endocrinol. Metab.* 40, 367-372.
- [6] Naftolin, F., Ryan, K.J., Davies, I.J., Reddy, V.V., Flores, F., Petro, Z. and Kuhn, M. (1975) *Recent Prog. Horm. Res.* 31, 295-319.
- [7] Longcope, C., Pratt, J.H., Schneider, S.H. and Fineberg, S.E. (1978) *J. Clin. Endocrinol. Metab.* 46, 146-152.
- [8] Evans, C.T., Corbin, C.J., Saunders, C.T., Merrill, J.C., Simpson, E.R. and Mendelson, C.R. (1987) *J. Biol. Chem.* 262, 6914-6920.
- [9] Osawa, Y., Tochigi, B., Higashiyama, T., Yarbrough, C., Nakamura, T. and Yamamoto, T. (1982) *Cancer Res. (suppl.)* 42, 3299s-3306s.
- [10] Simpson, E.R., Evans, C.T., Corbin, C.J., Powell, F.E., Ledesma, D.B. and Mendelson, C.R. (1987) *Mol. Cell. Endocrinol.* 52, 267-272.
- [11] Chen, S., Besman, M.J., Sparkes, R.S., Zollman, S., Klisak, I., Mohandas, I., Hall, P.F. and Shively, J.E. (1988) *DNA* 7, 27-38.
- [12] McPhaul, M.J., Noble, J.F., Simpson, E.R., Mendelson, C.R. and Wilson, J.D. (1988) *J. Biol. Chem.* 263, 16358-16363.
- [13] Corbin, C.J., Graham-Lorence, S., McPhaul, M., Mason, J.I., Mendelson, C.R. and Simpson, E.R. (1988) *Proc. Natl. Acad. Sci. USA* 85, 8948-8952.
- [14] Harada, N. (1988) *Biochem. Biophys. Res. Commun.* 156, 725-732.

- [15] Ebina, Y., Ellis, L., Jarnagin, K., Edery, M., Graf, L., Clauser, E., Ou, J., Masiarz, F., Kan, Y.W., Goldfine, I.D., Roth, R.A. and Rutter, W.J. (1985) *Cell* 40, 747-758.
- [16] Ito, H., Ike, Y., Ikuta, S. and Itakura, K. (1982) *Nucleic Acids Res.* 10, 1755-1769.
- [17] Grantham, R., Gautier, C., Gouy, M., Jacobzon, M. and Mercier, R. (1981) *Nucleic Acids Res.* 9, 43-74.
- [18] Chen, S., Shively, J.E., Nakajin, S., Shinoda, M. and Hall, P.F. (1986) *Biochem. Biophys. Res. Commun.* 135, 713-719.
- [19] Messing, J. (1983) *Methods Enzymol.* 101, 20-78.
- [20] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.
- [21] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) *Biochemistry* 18, 5294-5299.
- [22] Aviv, H. and Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* 69, 1408-1412.
- [23] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- [24] Davis, L.G., Dibner, M.D. and Battey, J.F. (1986) *Basic Methods in Molecular Biology*, Elsevier, Amsterdam.
- [25] Kozak, M. (1981) *Nucleic Acids Res.* 9, 5233-5252.
- [26] Evans, C.T., Ledesma, B.D., Schulz, T.Z., Simpson, E.R. and Mendelson, C.R. (1986) *Proc. Natl. Acad. Sci. USA* 83, 6387-6391.