Vol. 202, No. 1, 1994 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
July 15, 1994 Pages 49-57

STRUCTURAL COMPARISON OF HUMAN AND RAT PROSTATE-SPECIFIC ACID PHOSPHATASE GENES AND THEIR PROMOTERS: IDENTIFICATION OF PUTATIVE ANDROGEN RESPONSE ELEMENTS[1]

Pirjo Virkkunen§, Pirjo Hedberg§, Jorma J. Palvimo#, Elli Birr§, Katja Porvari§, Minna Ruokonen§, Päivi Taavitsainen§, Olli A. Jänne#, and Pirkko Vihko§[2]

§Biocenter Oulu and Department of Clinical Chemistry, University of Oulu, Kajaanintie 50, FIN-90220 Oulu, Finland

#Department of Physiology, University of Helsinki, FIN-00014 Helsinki, Finland

Received May 5, 1994

SUMMARY: Structural comparison of human and rat prostate-specific acid phosphatase (hPAP and rPAP) genes indicate that the exon number is different between these species. The hPAP gene contains 10 exons, whereas the rPAP gene has 11 exons. However, exons 2-9 of the genes are identical in size. The 5' regions of the two genes show 71 % identity in the most homologous region +1 to +340. The 5' untranslated regions of the human and rat genes are 50 and 49 nucleotides long, respectively. An *Alu* sequence is present upstream from the proximal promoter of the hPAP gene. Five putative androgen response elements altogether were localized in both the human and rat gene, one of which is conserved in location and sequence between the two genes. Two of these elements in both genes, the conserved one in the proximal promoter region and another one in intron 1, were shown to bind androgen receptor efficiently *in vitro*.

Acid phosphatases (orthophosphoric monoester phosphohydrolases, EC 3.1.3.2) are a group of isoenzymes that hydrolyze phosphomonoesters under acidic conditions. Human prostatic acid phosphatase (hPAP) and rat prostatic acid phosphatase (rPAP) are glycoproteins synthesized in the epithelial cells of the prostate gland, from which they are released into the prostatic fluid (1,2). To study the regulation and the prostate-specific expression of hPAP (3) and rPAP genes, we have undertaken the isolation and cloning of the genes[2]. By comparing the 5' flanking and promoter sequences of the hPAP and rPAP genes, it is possible to locate common regions necessary for prostate-specific gene expression. Because PAP is an androgen-regulated protein

[1]The nucleotide sequence data reported in this paper will appear in the EMBL GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers X74961-X74968 (hPAP) and X74969-X74978 (rPAP).

[2]Corresponding author: Fax: 358-81-3155631.

Abbreviations: AR, androgen receptor; ARE/GRE/PRE, androgen/glucocorticoid/progesterone response element; LAP, lysosomal acid phosphatase; PAP, prostatic acid phosphatase; h, human; r, rat.

0006-291X/94 $5.00

(4,5), its gene is also suitable for studying the mechanism of androgen action in the prostate. This approach could also provide information on the role of androgens in prostatic cancer. A few genes have been postulated to be regulated by androgens in the prostate, both in a positive and in a negative fashion. Examples of these genes are the rat C3(1) gene (6) and the human prostate-specific antigen (PSA) gene (7). Human PSA, like hPAP, is a prostate-specific secretory protein, but its mRNA is up-regulated by androgens in the LNCaP cell line, whereas hPAP mRNA is down-regulated (8), which makes it interesting to compare the promoter structures of the corresponding genes. In the present study, we have clarified the exon-intron structures, characterized the promoter regions and identified potential elements for transcription factor binding in the hPAP and rPAP genes. Also, binding of androgen receptor (AR) to putative androgen response elements (AREs) of the genes was studied.

MATERIALS AND METHODS

Isolation of genomic clones. Genomic libraries made of human lymphocyte DNA (Stratagene) and placental DNA (Clontech) were screened (9) using cDNA fragments for hPAP (10) as probes. Together three clones were isolated. A rat genomic library (Clontech) made of liver DNA was screened using rPAP cDNA fragments (11) as probes. Three positive genomic clones (Gr5', Gr7 and Gr11) were obtained.

DNA sequencing. The gene fragments were subcloned into plasmid vectors for double-stranded DNA sequencing (12), using Sequenase. The organization of the subcloned regions was verified by restriction enzyme mapping or by partial sequencing.

Identification of transcription start sites. Primer extension analysis was performed as previously described (13), using avian myeloblastosis virus reverse transcriptase (Promega). Ribonuclease protection assay was done as described (13).

Computer analyses. Location of putative transcription factor binding sites in the 5' sequences of the genes were carried out using a sequence analysis program from the University of Wisconsin Genetics Computer Group (UWGCG) (14).

Electrophoretic mobility shift assay. Electrophoretic mobility shift assays were carried out essentially as previously described (15,16).

RESULTS

Exon-intron structures

Sequence data from our genomic clones together with recently published data (17) established that the hPAP gene contains 10 exons, whereas the rPAP gene contains 11 exons similar to human lysosomal acid phosphatase (hLAP) (18). The seventh exon was not covered by our genomic clones (Fig.1A-B), however, the size of this exon is 133 bp on the grounds of rPAP cDNA. Comparison of the exon sizes between the genes of three acid phosphatases are shown in Table 1. Introns 1-6 in hPAP gene are >6000 bp, ~3500 bp, 458 bp, ~5400 bp, ~4900 bp and ~2370 bp in size, respectively. The sizes of the introns in rPAP gene were not determined, but all exon/intron splice junctions comply with the gt/ag rule, except for gc/ag in intron 3, as in the hPAP gene.

Transcription start sites

One major product corresponding to position 50 nucleotides 5' from the ATG codon was obtained in the transcription start site analysis of the hPAP gene (Fig. 2A). Similarly, the major

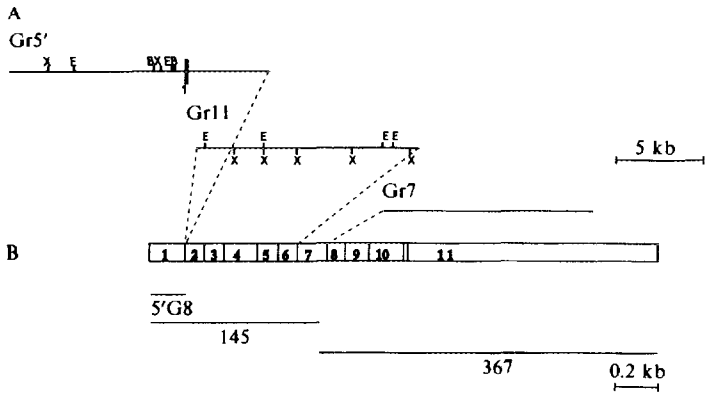


FIG. 1. Structure of rPAP gene. (A) Genomic clones Gr5', Gr11, and Gr7 of the rPAP gene with restriction enzyme cleavage sites used in subcloning, *Bam*HI (B), *Eco*RI (E) and *Xba*I (X). The exons covered by each of the clones are indicated by dotted lines. The location of the first exon is marked by a vertical line and the number 1. The sizes of the introns were not determined. (B) Rat PAP cDNA corresponding to the middle-sized rPAP mRNA (11). The cDNA fragments used as probes in this study are indicated by horizontal lines.

protected fragment corresponded to position 49 nucleotides upstream from the translation initiation codon in the case of the rPAP gene (Fig. 2B). In addition, some other signals were detected from both genes. Repetition of the experiments confirmed that 5' untranslated regions of hPAP and rPAP mRNAs are 50 and 49 nucleotides long, respectively.

5' sequences

We have sequenced about 3.9 and 1.8 kb of the 5' flanking regions of the hPAP and rPAP genes, respectively. Computer analysis of the sequence revealed many sequence motifs reminiscent of reported consensus sequences for transcription factor binding sites (Fig. 3). When comparing the hPAP and the rPAP 5' flanking sequences by aligning the ATG codons, the

TABLE 1
Comparison of exons in the hPAP, rPAP and hLAP genes [a]

	hPAP		rPAP		hLAP	
exon	bp	covers	bp	covers	bp	covers
1	170	5'-utr, sp, 8 aa	166	5'-utr, sp, 8 aa	>118	5'-utr, sp, 8 aa
2	96		96		96	
3	87		87		87	
4	153		153		153	
5	99		99		99	
6	93		93		90	
7	133		133		133	
8	83		83		83	
9	104		104		107	
10	2098	63 aa, 3'-utr	170	56 aa	176	58 aa
11	-	-	>1173	3 aa, 3'-utr	947	44 aa (cd + tmd), 3'-utr

[a] aa, amino acid; utr, untranslated region; sp, signal peptide; cd, cytoplasmic domain; tmd, transmembrane domain.

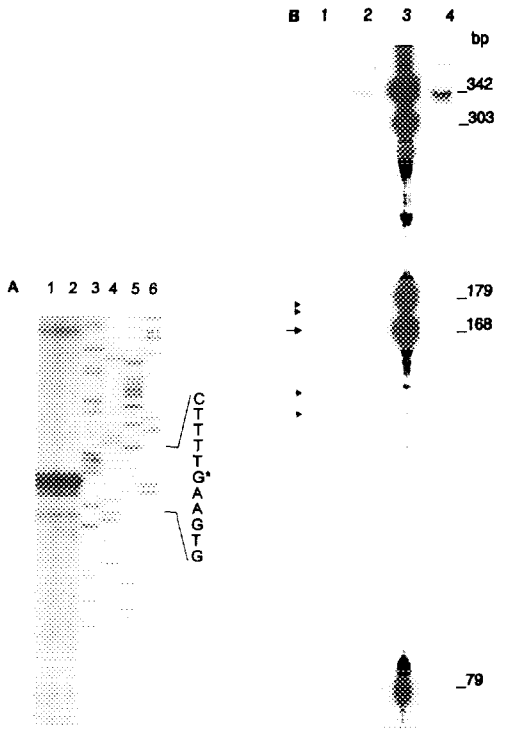


FIG. 2. Identification of transcription start sites. (**A**) Primer extension analysis of the 5' end of the hPAP gene. A synthetic 20-nucleotide ^{32}P-labelled primer, complementary to nucleotides 8-27 downstream from the initiation codon ATG, was annealed to 20 μg of poly(A)$^+$RNA from human prostatic carcinoma (lane 1) or hyperplasia (lane 2) tissue at 30 °C. The products were analyzed by electrophoresis on a sequencing gel. Lanes 3-6 show sequencing reactions with the same primer. The asterisk depicts the main transcription start site in the sequence, 50 nucleotides upstream from the ATG codon. (**B**) Ribonuclease protection assay of the rPAP gene. The ^{32}P-labelled RNA probe covering nucleotides -199 to +123 of the rPAP gene, was hybridized to 100 μg (lane 1) and 150 μg (lane 2) of total rat prostate RNA at 55 °C and 60 °C, respectively, the unhybridized RNA was digested, and the protected products were analyzed on a sequencing gel. Lanes 3 and 4 contain the RNA size standards and the tRNA control, respectively. The transcription start sites are marked with arrows; the main one is 49 nucleotides upstream from the ATG codon (bold arrow).

locations of the TATA box and the CAAT box, one of the putative binding sites of the transcription factor AP-2, and that of a putative binding site of the androgen/glucocorticoid/progesterone receptors are conserved. Immediately upstream from the proximal promoter of the hPAP gene, from -561 to -280, there is a complete dimeric *Alu* element, which has 85 % homology with the human *Alu* consensus sequence (19). The overall homology of the aligned 900 bp sequences around the ATGs is 54 % (Fig. 3), but in the most homologous region, about 340 bp downstream from the transcription start sites, it is 71 %. The computer analyses revealed altogether five potential AREs in both the hPAP gene and the rPAP gene, in the region from -1700 to +1800 (Table 2). Similarly in the both genes, three putative AREs are reminiscent of the half consensus TGTTCT and two have the reversed half consensus AGAACA.

As for the tissue-specific regulation of the hPAP and rPAP genes, it is of interest that the first intron of the former contains a sequence motif 5'-ATTTTAGGATGATT-3' (+665 to +678) and

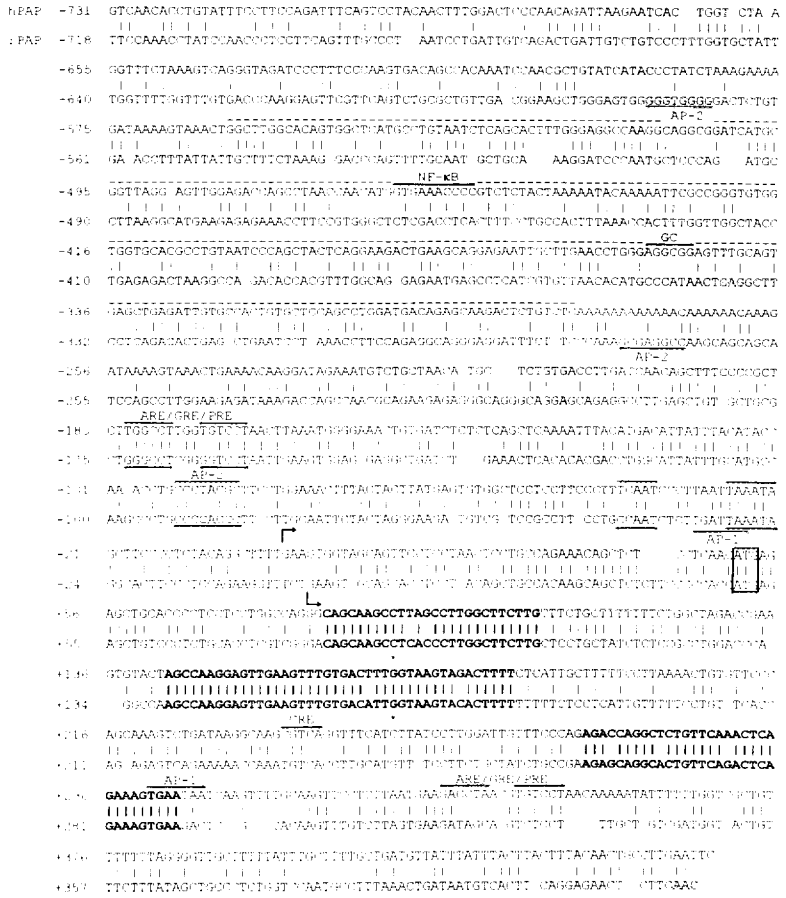


FIG. 3. Alignment of the 5' sequences of hPAP and rPAP genes. The TATA boxes and the CAAT boxes are overlined (hPAP) and underlined (rPAP). The putative transcription factor binding sites are marked by abbreviations of the elements or by the names of the factors above (hPAP) and under (rPAP) the sequence; AP-2 transcription factor, NF-κB transcription factor, GC box, ARE/GRE/PRE, AP-1 transcription factor, cAMP response element (CRE). The ATG translation initiation codons are boxed, and the transcription start sites are shown by arrows above (hPAP) and under (rPAP) the sequence. An *Alu* element in the hPAP gene is marked by a dashed line above the sequence. Three long stretches of conserved nucleotides are indicated by bold text and the last nucleotides of the first exons by asterisks.

TABLE 2
Putative ARE/GRE/PREs in the hPAP and rPAP genes

	sequence	location
hPAP gene	5'-AGAACAcaaTCCTCC-3'	-1631/-1617
	5'-GTAAACcatTGTTCT-3'	-1576/-1562
	5'-TGGCCTtggTGTCCT-3'	-178/-164
	5'-GACCTAacgTGTCCT-3'	+336/+350
	5'-AGAACAactATTTGT-3'	+611/+625
rPAP gene	5'-AGAACAggaAGCCGA-3'	-1612/-1598
	5'-GCAAACcatTGTTCT-3'	-1150/-1136
	5'-GGGCCTcggGGTCCT-3'	-174/-160
	5'-AGAACAcagAATTTC-3'	+586/+600
	5'-TTGATGctcTGTCCT-3'	+1366/+1380
consensus	5'-GGTACAnnnTGTTCT-3'	

the first intron of the latter a motif 5'-TTTAGTTGATCATT-3' (+623 to +636), which both resemble the element 5'-TTTATAGGATGTTT-3' in the first intron of the rat prostatic binding protein (PBP) C3(1) gene that was recently reported to bind a prostate-specific nuclear protein (20).

Androgen receptor binding to putative androgen response elements of the hPAP and rPAP genes

Interaction of androgen receptor with three putative AREs of the hPAP (-1576/-1562, -178/-164, and +336/+350) and rPAP (-1612/-1598, -174/-160 and +586/+600) genes (Table 2), were studied by electrophoretic mobility shift assay. Oligonucleotide -185/-157 of the hPAP gene was found to bind AR most efficiently (Fig. 4A, lane 7). Similarly, the oligonucleotide -181/-153 covering the putative ARE at the proximal promoter of the rPAP gene showed a clear AR-ARE complex (Fig. 4B, lane 7). In addition to the receptor, an unknown protein of the AR preparation bound the oligonucleotide -185/-157 of the hPAP gene. This complex was not, however, affected by the receptor-specific antibody. Oligonucleotide +329/+357 covering the putative ARE in the first intron of the hPAP gene (Fig. 4A, lane 4) and the corresponding oligonucleotide +579/+607 of the rPAP gene (Fig. 4B, lane 4), showed specific binding in the presence of the AR specific antibody, which stabilizes the interaction between AR and ARE (16). Very weak AR binding were seen with the third oligonucleotides (hPAP: -1581/-1554; rPAP: -1619/-1591) even in the presence of the antibody (Fig. 4A and 4B, lane 14).

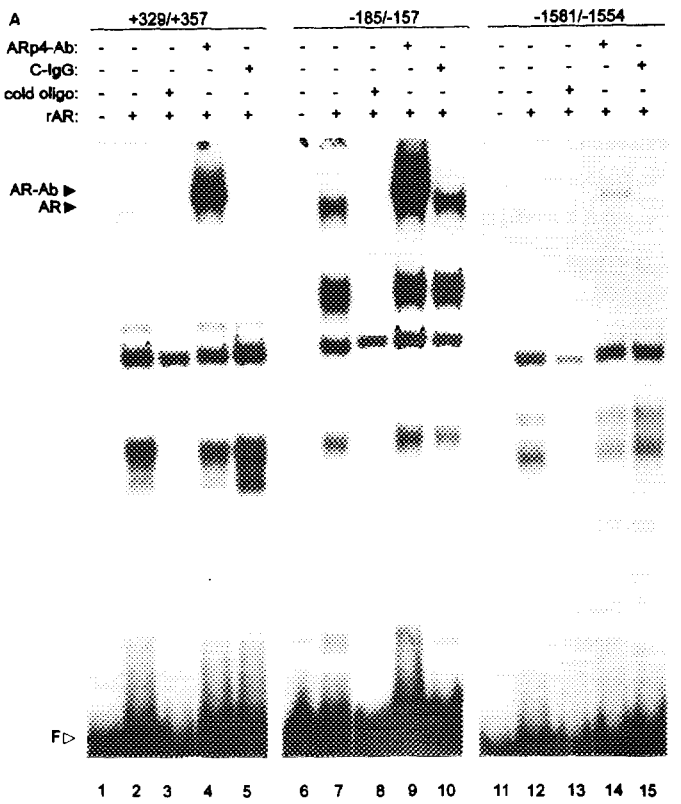
DISCUSSION

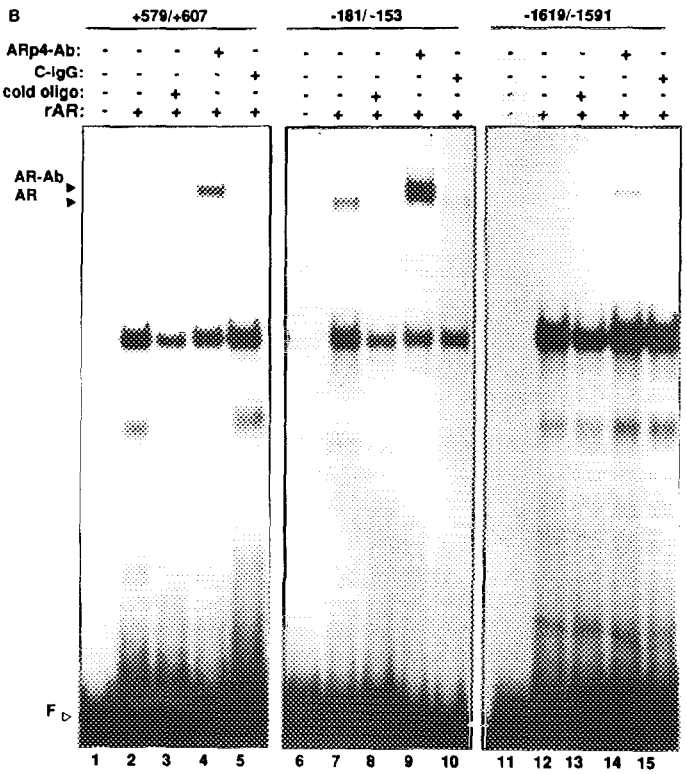
In the present study we have shown that the exon-intron and proximal promoter structures of human PAP and rat PAP genes are very similar. However, evolution of this gene seems to be different between these species when the number of exons is concerned. Furthermore, the proximal promoters of the genes contain conserved regions (Fig. 3), but a noticeable difference is the 5' upstream *Alu* sequence. There are examples of upstream *Alu* sequences containing negative and positive regulatory elements, which interact with specific nuclear factors (21,22). The existence of such regulatory elements in the *Alu* repeat of the hPAP gene are not yet clarified.

The functionality of the potential ARE/GRE/PREs and the binding sites of transcription factors, which were recognized in the hPAP and rPAP gene sequences in this study, are under

---

FIG. 4. Electrophoretic mobility shift assay with full-length androgen receptor and putative androgen response elements of the hPAP (**A**) and rPAP (**B**) genes. Partially purified rAR (16), was incubated with ^{32}P-labelled double-stranded oligo-nucleotides corresponding to indicated regions of the hPAP gene (+329 to +357, -185 to -157, and -1581 to -1554) and the rPAP gene (+579 to +607, -181 to -153, and -1619 to -1591) and the resulting complexes were analyzed. Lanes 1, 6, and 11, free probe; lanes 2, 7, and 12, rAR in the presence of a labelled probe; lanes 3, 8, and 13, rAR in the presence of a labelled probe and a 100-fold molar excess of unlabelled probe; lanes 4, 9, and 14, rAR incubated with a labelled probe and an antibody (ARp4-Ab) against N-terminal region of androgen receptor; lanes 5, 10, and 15, rAR incubated with a labelled probe and normal rabbit IgG (C-IgG). Closed triangles indicate specific androgen receptor- and androgen receptor-antibody-DNA-complexes and open triangles the position of free probe (F).


A
+329/+357
-185/-157
-1581/-1554
ARp4-Ab:
C-IgG:
cold oligo:
rAR:
AR-Ab
AR
F
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15


B
+579/+607
-181/-153
-1619/-1591
ARp4-Ab:
C-IgG:
cold oligo:
rAR:
AR-Ab
AR
F
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

investigation. The possible function of the ARE/GRE/PREs as mediators of the known androgen regulation of hPAP and rPAP genes is especially interesting. In particular, the conserved location and sequence of the putative ARE at position -178/-174 in the hPAP/rPAP genes suggest an important role for this element. The three AREs for which AR shows affinity, might have a synergistic effect in androgen regulation of the genes. There are examples of other genes that have several AREs, such as rat prostatic binding protein (PBP) genes (23), mouse sex-limited protein (Slp) gene (24), and rat prostatic probasin (PB) gene (25). In the latter two of these genes, each of the AREs were found to be functionally important for maximum hormone response, although, e.g. in the case of Slp gene, the *in vitro* binding affinities of AR for the individual elements were remarkably different. Thus, also the AREs of the hPAP or rPAP gene may function in cooperation with each other and with other transcriptional elements. In addition, the gene of a third human prostate-specific secretory protein, prostate-specific antigen (PSA), contains an ARE at almost the same position from the transcription start site (-170) as the conserved putative ARE of the hPAP and rPAP genes (-178/-174). This ARE was shown to be essential for androgen-regulated transcription of the PSA gene (7,26).

The first intron of both the hPAP and the rPAP gene contains a sequence motif that resembles the 14-nucleotide element of the rat PBP C3(1) gene that was recently proposed to mediate, at least in part, the prostate-specific expression of the gene (20). Whether these C3(1)-like motifs of the two PAP genes are involved in the regulation of the tissue-specific expression of the genes, remains to be found out.

ACKNOWLEDGMENTS: We thank Mrs. Mirja Mäkeläinen for her expert technical assistance. This work was supported by the Sigrid Jusélius Foundation, the Finnish Cancer Foundation, and the Research Council for Medicine of the Academy of Finland. The Department of Clinical Chemistry, University of Oulu, is a WHO Collaborating Center for research in reproduction supported by the Ministries of Education, Health and Social Affairs, and Foreign Affairs, Finland.

REFERENCES

1. Yam, L.T., Li, C.Y., and Lam, K.W. (1980). *In* "Male Accessory Sex Glands" (E. Spring Mills, E.S.E. Hafez, Eds.), pp.183-196, Elsevier, North Holland.
2. Vanha-Perttula, T., Niemi, R., and Helminen, H.J. (1972) Invest. Urol. 9: 345-352.
3. Solin, T., Kontturi, M., Pohlmann, R., and Vihko, P. (1990) Biochim. Biophys. Acta. 1048: 72-77.
4. Tenniswood, M., Bird, C.E., and Clark, A.F. (1976) Can. J. Biochem. 54: 350-357.
5. Bolton, N., Lahtonen, R., Vihko, P., Kontturi, M., and Vihko, R. (1981) Prostate 2: 409-416.
6. Page, M.J., and Parker, M.G. (1982) Mol. Cell. Endocrinol. 27: 343-355.
7. Riegman, P.H.J., Vlietstra, R.J., van der Korput, J.A.G.M., Romijn, J.C., and Trapman, J. (1989) Biochem. Biophys. Res. Comm. 159: 95-102.
8. Henttu, P., Liao, S., and Vihko, P. (1992) Endocrinology 130: 766-772.
9. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). "Molecular Cloning: a Laboratory Manual", 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
10. Vihko, P., Virkkunen, P., Henttu, P., Roiko, K., Solin, T., and Huhtala, M-L. (1988) FEBS Lett. 236: 275-281.
11. Roiko, K., Jänne, O.A., and Vihko, P. (1990) Gene 89: 223-229.
12. Sanger, F., Nicklen, S., and Coulson, A.P. (1977) Proc. Natl. Acad. Sci. USA 74: 5463-5467.

13. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (1989). "Current Protocols in Molecular Biology", Vol. 1, Greene Publishing Associates and Wiley-Interscience, NY.
14. Devereaux, J., Haeberli, P., and Smithies, O. (1984) Nucleic Acids Res. 12: 387-395.
15. Palvimo, J.J., Eisenberg, L.M., and Jänne, O.A. (1991) Nucleic Acids Res. 19: 3921-3927.
16. Xie, Y-B., Sui, Y-P., Shan, L-X., Palvimo, J.J., Phillips, D.M., and Jänne, O.A. (1992) J.Biol.Chem. 267: 4939-4948.
17. Sharief, F.S., and Li, S.S-L. (1992) Biochem. Biophys. Res. Comm. 184: 1468-1476.
18. Geier, C., von Figura, K., and Pohlmann, R. (1989) Eur. J. Biochem. 183: 611-616.
19. Schmid, C.W., and Jelinek, W.R. (1982) Science 216: 1065-1070.
20. Celis, L., Claessens, F., Peeters, B., Heyns, W., Verhoeven, G., and Rombauts, W. (1993) Mol. Cell. Endocrinol. 94: 165-172.
21. Saffer, D.J., and Thurston, S.J. (1989) Mol. Cell. Biol. 9: 355-364.
22. Brini, A.T., Lee, G.M., and Kinet, J-P. (1993) J. Biol. Chem. 268: 1355-1361.
23. Claessens, F., Rushmere, N.K., Davies, P., Celis, L., Peeters, B., and Rombauts, W.A. (1990) Mol. Cell. Endocrinol. 74: 203-212.
24. Adler, A., Scheller, A., Hoffman, Y., and Robins, D.M. (1991) Mol. Endocrinol. 5: 1587-1596.
25. Rennie, P.S., Bruchovsky, N., Leco, K.J., Sheppard, P.C., McQueen, S.A., Cheng, H., Snoek, R., Hamel, A., Bock, M.E., MacDonald, B.S., Nickel, B.E., Chang, C., Liao, S., Cattini, P.A., and Matusik, R.J (1993) Mol. Endocrinol. 7: 23-36.
26.Riegman, P.H.J., Vlietstra, R.J., van der Korput, J.A.G.M., Brinkmann, A.O., and Trapman, J. (1991) Mol. Endocrinol. 5: 1921-1930.