The EMBO Journal Vol.2 No.5 pp.769-774, 1983

Rat prostatic steroid binding protein: DNA sequence and transcript maps of the two C3 genes

H.C.Hurst and M.G.Parker*

Imperial Cancer Research Fund, P.O.Box 123, Lincoln's Inn Fields, London WC2A 3PX, UK

Communicated by W.F.Bodmer
Received on 22 February 1983

In the rat there are two non-allelic genes C3(1) and C3(2) for the C3 polypeptide of prostatic steroid binding protein. We have cloned and sequenced both genes and show that only C3(1) is responsible for the production of authentic C3. Although there is a marked difference in their transcriptional activity, the two genes share extensive DNA sequence homology there being only one base difference from nucleotide −235 to within the first intron. Transcript mapping has shown that there are two distinct C3 transcripts which share a unique 3′ terminus but have 5′ termini 38 bases apart each preceded by a 'TATA' box homology. Interestingly, an identical repetitive element is present just upstream of both genes. Both families of transcripts, which are produced in a ratio of 18:1, are coordinately regulated by testosterone.

Key words: DNA sequence/repetitive sequences/promoters/duplicated genes/steroid hormones

Introduction

Prostatic steroid binding protein is the predominant protein secreted into rat prostatic fluid (Heyns and DeMoor, 1977). The protein is an oligomer containing C1, C2 and C3 polypeptides (Heyns *et al.*, 1978) whose rates of synthesis are stimulated markedly by testosterone (Parker *et al.*, 1978). In common with other classes of steroid hormone (Gorski and Gannon, 1976; Higgins and Gehring, 1978), the mechanism whereby androgens regulate the expression of steroid binding protein is *via* effects on rates of transcription but, in addition, they also modulate RNA turnover (Page and Parker, 1982). There are two non-allelic genes, C3(1) and C3(2) for the C3 polypeptide (Parker *et al.*, 1983) and DNA clones containing each of these genes have been introduced into an heterologous, androgen-responsive cell line, namely mouse mammary tumour cells (S115), to investigate the regulatory role played by testosterone. The genes for C3 were accurately transcribed and the steady-state levels of mRNA were stimulated by testosterone in cell-lines that contained only 1-2 integrated copies of the gene (Page and Parker, 1983).

The elucidation of the interaction of the C3 genes with testosterone and its receptor depends on a knowledge of their DNA sequence organisation. In this paper we report the DNA sequence of C3(1) and C3(2) and mapping studies of prostatic C3 mRNA which together indicate that only C3(1) is responsible for C3 production in the ventral prostate. This difference in transcriptional activity of the two genes is remarkable because the DNA sequence from nucleotide −235 to within the first intron shows only one base change and both genes were expressed similarly in heterologous cells. Finally, we show the presence of an identical repetitive element just upstream of both genes.

*To whom reprint requests should be sent.

Results

DNA sequencing

The overall organisation of the C3 genes represented in clones λ6 and λ11B has previously been described (Parker *et al.*, 1983) and is summarised in Figure 1 which also shows the strategy used to sequence the two genes. Most of the sequence was obtained *via* the cloning of specific genomic restriction fragments into the vectors M13 mp8 and 9. As the λ6 clone lacks the 5′ *Pst*I site present in λ11B it was not possible to clone directly suitably sized fragments around this site. This was overcome by constructing a series of deletion mutants, made by resecting with *Bal*31 nuclease from unique restriction sites. In addition, we did some shotgun cloning of both genes into M13 mp7 using each of the enzymes *Alu*I, *Hae*III or *Sau*3A to digest the genomic DNA. We have sequenced both strands of the coding region and immediate 5′-flanking region and all sequences were determined at least twice. Sequence information was sorted and stored using the computer program DBUTIL (Staden, 1980) and compared with other sequences using the SEQ program (Brutlag *et al.*, 1982).

A comparison of the two genomic sequences is shown in Figure 2. C3(1) proved to be totally compatible with the cDNA clone, pA34 (Parker *et al.*, 1983) whereas C3(2) showed some crucial differences, but is nevertheless 97.7% homologous with mRNA as represented by pA34. Both genes are interrupted by two introns of 1.7 kb and 0.69 kb, confirming the previous interpretation of R-looping and Southern blotting experiments (Parker *et al.*, 1983). The introns occur at the same point in both genes and can be seen to possess flanking sequences that agree well with published consensus sequences for splice sites (Breathnach and Chambon, 1981; Mount, 1982).

The DNA sequences of the first exon which encodes the complete signal peptide are identical in both genes. Sequences corresponding to the C3 protein (Peeters *et al.*, 1981) are en-

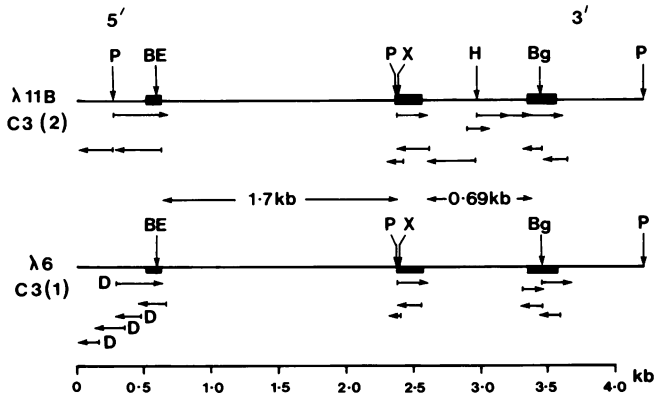

Fig. 1. DNA sequencing strategy. The organisation of the two C3 genes as represented by the clones λ11B and λ6 is presented with exons shown as solid blocks. The restriction enzyme sites shown are: X = *Xba*I, P = *Pst*I, H = *Hind*III, BE = *Bst*EII, Bg = *Bgl*II. Arrows indicate the positions of each DNA strand which has been sequenced. A 'D' indicates that the sequence was derived from a deletion mutant made by resecting a subclone with *Bal*31 nuclease.

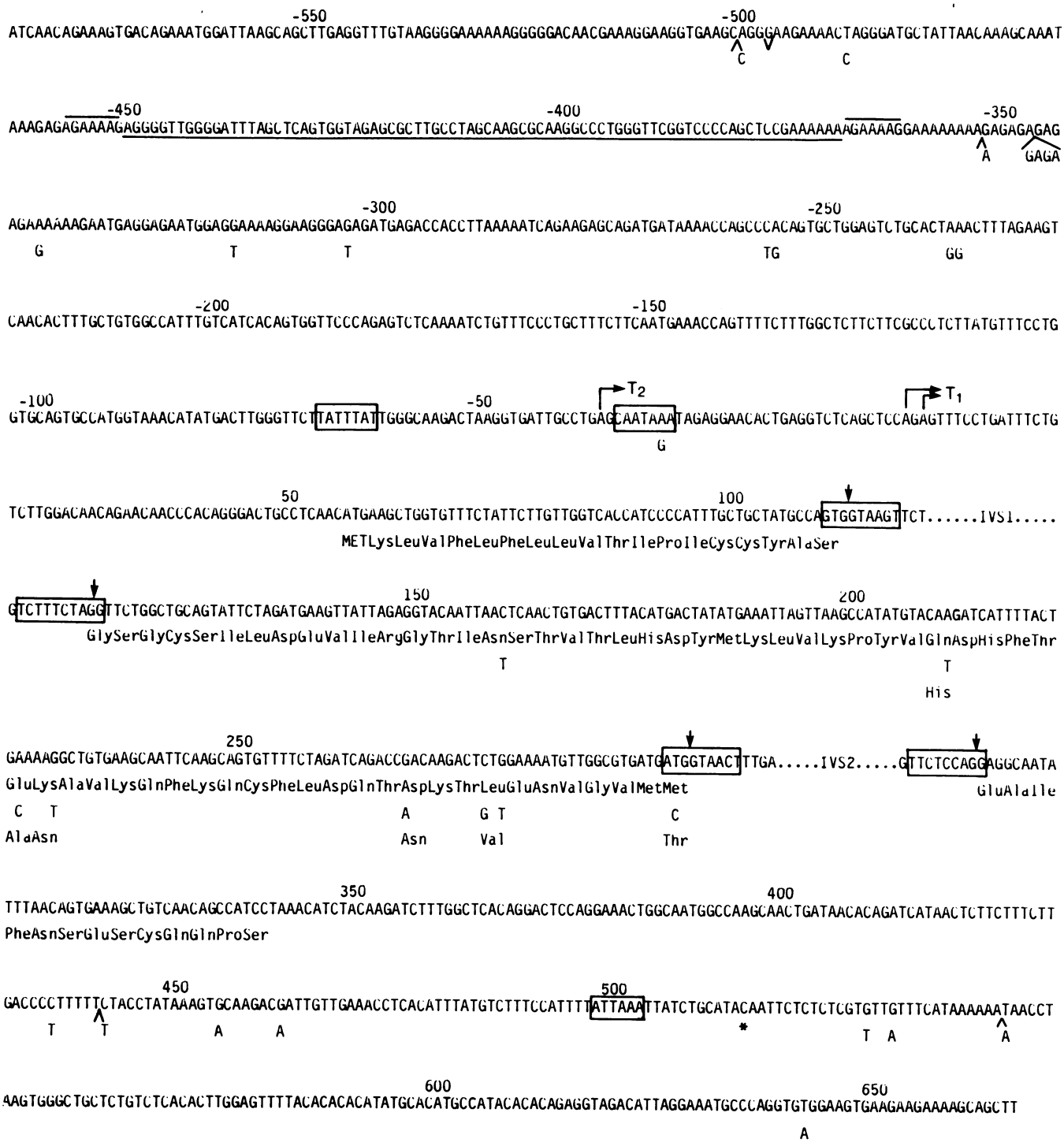

Fig. 2. DNA sequence of two C3 genes. The nucleotide sequence for the non-coding strand of the C3(1) gene and corresponding amino acid sequence is presented with differences in the C3(2) gene shown below. Deletions are marked (∨) and insertions are marked (∧). TATA boxes, splice acceptor and donor sites and the polyadenylation signal are boxed. Transcription regions T_1 and T_2 are labelled and the 3′ transcription terminus marked with an asterisk (*). The rat repetitive element is underlined and its flanking 7 bp direct repeat overlined. Vertical arrows (↓) delineate the exact exon-intron junctions, though only enough intervening sequence to compare with the consensus splice donor [$^{C}_{A}$AGGT$^{G}_{A}$AGT] and acceptor [(T)$_{(C)n}$ NCAGG] sequences (Mount, 1982) is shown. The numbering corresponds with that given for pA34 (Parker *et al.*, 1983) and does not include any intervening sequence.

coded by the second and a small portion of the third exons. Differences within the coding portions of the two genes are confined to the second exon, with C3(2) showing eight base changes compared with C3(1) and the cDNA clone. The protein sequence of C3 (Peeters *et al.*, 1981) indicated that it must be coded for by C3(1) and not C3(2) which shows six amino acid differences.

A total of 580 bp of 5′-flanking DNA of the two genes have been compared and shown to be 96.6% homologous, while 150 bp of 3′-flanking DNA show 97.3% homology. We have complete intervening sequence data for the small intron of C3(2) only. However, over the regions for which data

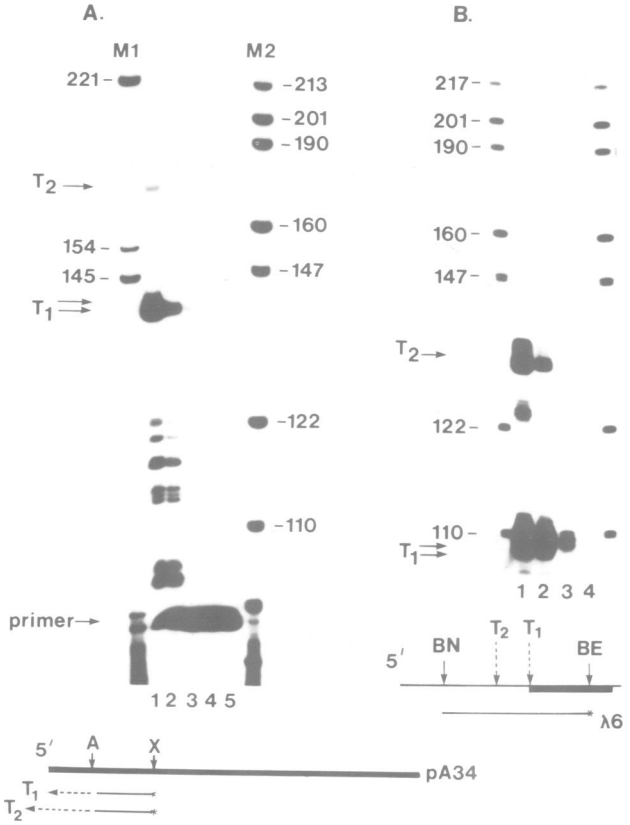

Fig. 3. Mapping of prostatic C3 transcripts. **(A)** Primer extension (Parker *et al.*, 1982) was carried out using a ^{32}P-labelled 68-base primer generated from the cDNA clone pA34. The clone was cut with *Xba*I(X), labelled using polynucleotide kinase (Maxam and Gilbert, 1980), and recut with *Alu*I(A) before strand separation. All samples contained 30 000 c.p.m. primer and 20 μg RNA comprising yeast RNA plus ventral prostrate poly(A)$^+$ RNA as follows: **track 1**, 1.0 μg; **track 2**, 0.1 μg; **track 3**, 0.01 μg; **track 4**, 0.001 μg; **track 5**, control. Marker tracks M1 and M2 are ^{32}P-labelled *Hinf*I and *Hpa*I digests of pAT153, respectively. **(B)** S1 mapping was carried out using a 190-bp, ^{32}P-labelled probe generated from a subclone of the first exon of C3(1). The clone was cut with *Bst*EII (BE), labelled using polynucleotide kinase (Maxam and Gilbert, 1980) and recut with *Bst*N1 (BN) before gel purification. All samples contained 50 000 c.p.m. probe and 25 μg RNA comprising yeast RNA plus ventral prostate poly(A)$^+$ RNA as follows: **track 1**, 1.0 μg; **track 2**, 0.1 μg; **track 3**, 0.01 μg; **track 4**, control. The markers are ^{32}P-labelled *Hpa*II digested pAT153.

is available from both genes (not shown), the intervening sequences share ~95% homology. Such extensive homology suggests that the two C3 genes are the result of a recent duplication of a single gene.

Finally, examination of the 5′-flanking sequences revealed an element of 88 bp flanked by a 7-bp direct repeat (Figure 2) which resembles the Alu-like moderately repetitive elements found dispersed in mammalian DNA (reviewed by Jelinek and Schmid, 1982). It is identical in both C3(1) and C3(2) and is flanked by extensive AG-rich sequences.

Mapping of 5′ and 3′ termini of C3 mRNA

Since the two C3 genes are so similar it was interesting to try to determine whether one or both are expressed *in vivo*. To investigate this, prostatic C3 mRNA, which comprises ~10% total poly(A)$^+$ RNA from normal adult animals, was mapped to the genomic sequence. The 5′ ends of C3 mRNA were determined using primer extension and S1 nuclease mapping with ^{32}P-labelled probes as described in Figure 3. Both techniques show clearly that there are two major C3 transcription regions, designated T_1 and T_2. However, there were so-called ladders produced in the primer extension reactions probably due to premature termination and the doublet and triplet bands produced in S1 mapping experiments may be due to over digestion of the DNA/RNA hybrids by S1 nuclease.

T_1 appears to be heterogeneous and was examined more closely by primer extension under chain termination conditions. The resulting sequence is compared with the corresponding genomic sequence run alongside in Figure 4A. It is clear that transcription from T_1 maps within the sequence 5′ CCAGAG 3′ although the mRNA sequence is a little unclear at the 5′ end, presumably caused by the methylation of ribonucleotides at the cap site. The T_1 doublet probably represents equal transcription from both A residues within this sequence. The second A also coincides with the 5′ end of the cDNA clone pA34. Examination of the DNA sequence flanking T_1 (Figure 2) shows that a sequence of 5′ CAATAAATATA 3′ occurs 36 bases further upstream in C3(1). This closely resembles, in both sequence and position relative to the cap site, the 'TATA' box believed to form part of the promoter in other eucaryotic protein coding genes (Breathnach and Chambon, 1981). In C3(2), however, the sequence is mutated to 5′ CAATAGATA 3′. Interestingly, this A to G transversion is the only nucleotide from −235 to within the first intervening sequence where the two genes differ.

The transcriptional start site T_2 is 36/38 bases upstream from the T_1 doublet. From Figure 4A (track 5) allowing for the slope of the gel, T_2 appears to have a 5′ A. This maps 33 bases downstream of the sequence 5′ TATTTAT 3′ present in both C3(1) and C3(2) (Figure 2) and also found preceding the transcription start sites of SV40 early (Benoist and Chambon, 1981) and adenovirus 5 E1a (Maat *et al.*, 1980) genes. A primer prepared from genomic DNA whose 3′ end maps between T_1 and T_2 was used for primer extension sequencing of the T_2 transcript (Figure 4b). The resulting sequence is not very clear in terms of assigning the precise 5′ end of the transcript, but is interesting because it is clear that a transcript initiating at T_2 is similar to the sequence 5′ CAATAAATA 3′ found in C3(1) upstream of T_1 and not to the mutated form 5′ CAATAGATA 3′ found in C3(2). This indicates that prostatic mRNA, initiated from T_2, may derive from C3(1) alone.

Since the sequenced cDNA is incomplete at its 3′ end, lacking a poly(A) tail (Parker *et al.*, 1983), we have mapped the 3′ ends of C3 transcripts to the genomic DNA. A 300-bp ^{32}P-labelled probe was generated by labelling at the *Bgl*II site within the third exon using Klenow DNA polymerase (Maniatis *et al.*, 1982). This was used in an S1 mapping experiment and resulted in the production of a single protected band of 165 bp (not shown) which maps 12 bp downstream of a sequence 5′ ATTAAA 3′ found in the non-coding DNA of both genes.

Effect of testosterone on C3 mRNA

The ratio of C3 mRNA which had been transcribed from T_1 and T_2 was 18:1 as judged by scanning of autoradiographs from primer extension and S1 mapping gels. To see if this ratio altered with the hormonal status of the tissue, primer extension reactions were carried out on rat prostate nuclear RNA from normal animals and animals 3 days after castra-

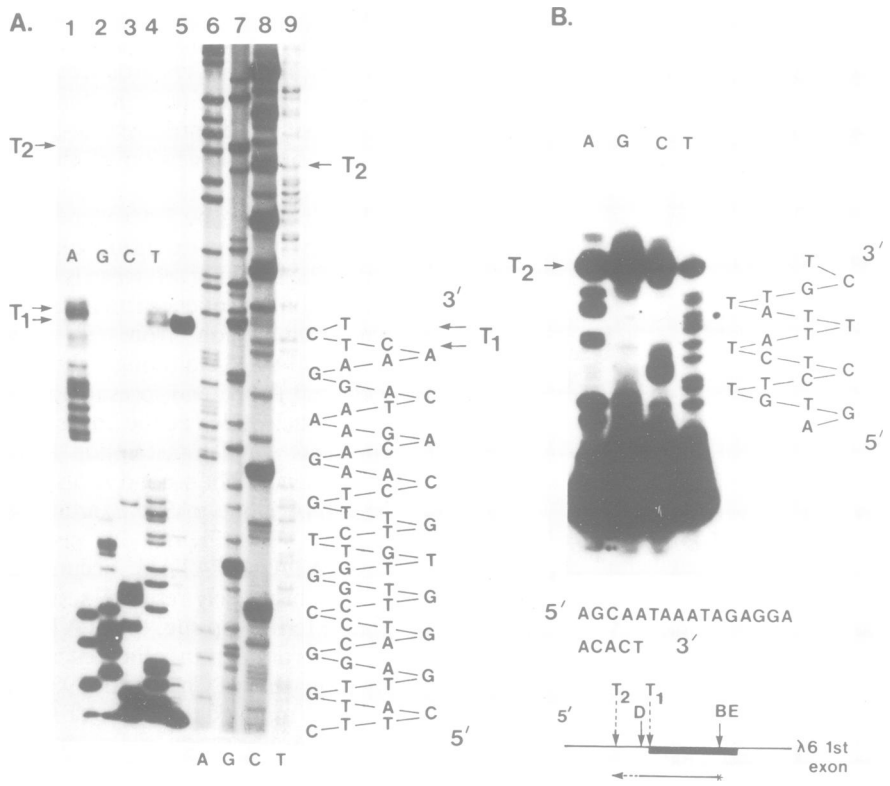

Fig. 4. Sequence analysis of C3 transcripts. **(A)** The sequence of transcripts initiating at T_1 was determined by primer extension analysis under chain termination conditions using the ^{32}P-labelled *Xba*I-*Alu*I primer described in Figure 3A. **Track 1–4** each represent 10 000 c.p.m. primer hybridised to 2.0 μg ventral prostate poly(A)$^+$ RNA. **Track 5** shows primer extension under normal conditions of 0.25 μg poly(A)$^+$ RNA hybridised to 30 000 c.p.m. primer. **Tracks 6–9** show the sequence of the genomic coding strand. The 2.1-kb *Pst* fragment of λ11B was cloned into M13 mp 7 and the non-coding strand clone dideoxy sequenced using the *Xba*I-*Alu*I primer. The sequence obtained and reproduced here therefore represents the coding strand. **(B)** The sequence of transcripts initiating at T_2 was determined using a ^{32}P-labelled 95-base primer generated from a subclone for the first exon of C3(1) by labelling with polynucleotide kinase (Maxam and Gilbert, 1980) at the *Bst*EII site (BE) and recutting with DdeI (D) before strand separation. Each track represents 20 000 c.p.m. of primer hybridised to 12.5 μg poly(A)$^+$ RNA and primer extended under chain termination conditions. The sequence produced, representing the coding strand, is shown adjacent to the gel with the complementary non-coding sequence written underneath.

tion with or without 24 h of testostrone (Figure 5). Scanning of the autoradiograph shows that both transcripts fall to 1/5th of normal levels on castration and that treatment with testosterone results in a doubling of these levels. Hence, the relative ratios of the transcripts do not alter during hormonal manipulation and androgens have an equal effect on steady-state levels of both transcripts.

Discussion

In the rat there are two non-allelic genes for the C3 polypeptide of prostatic steroid binding protein (Parker *et al.*, 1983). Here we report the DNA sequences of these two genes, which support our preliminary suggestion that they result from the duplication of an ancestral gene (Parker *et al.*, 1983), and mapping studies of prostatic C3 mRNA which reveal several interesting features in the DNA sequence.

From mapping studies, we show that there are two major transcription initiation regions, T_1 and T_2, each of which is preceded by a 'TATA' box homology. However, there is no sequence which shares extensive homology with the so-called CAAT box that is found 70–90 nucleotides in front of several eucaryotic genes (Breathnach and Chambon, 1981). Mapping of the 3′ end of RNA transcripts indicated that there was a unique termination site which is 12 bp downstream from the sequence ATTAAA. Although this sequence differs slightly from the putative polyadenylation signal observed in most other eucaryotic protein-coding genes (Proudfoot and Brownlee, 1976) it is found in the chicken lysozyme (Jung *et al.*, 1980) and mouse pancreas α-amylase (Hagenbuchle *et al.*, 1980) genes.

In view of the structure of prostatic steroid binding protein which consists of two subunits, one containing the polypeptides C1 and C3 and one containing the polypeptides C2 and C3, the question arises as to whether both C3 genes are expressed. The evidence indicates that C3(1) is responsible for production of C3 which has been sequenced (Peeters *et al.*, 1981) and that C3(2) is transcribed poorly, if at all. Firstly, minor forms of C3 have not been detected (Peeters *et al.*, 1981). Secondly, although we identified R-loops between mRNA and λ6 DNA, we have never identified R-loops with the entire C3 gene in λ11B DNA (Parker *et al.*, 1983). Instead, only incomplete R-loops were obtained in which the second exon in C3(2) failed to hybridise and presumably such R-loops represented mismatching between C3(1) RNA transcripts and C3(2) DNA. Thirdly, mapping studies showed that RNA which was initiated at T_2 was transcribed from C3(1), although we cannot completely rule out the possibility that there are a small number of C3(2) transcripts produced at levels below the resolution of the method used. Finally, the state of DNA methylation within and flanking the genes has been shown to correlate inversely with transcriptional activity (Razin and Friedman, 1981; Felsenfeld and McGhee, 1982). In agreement with this, we have shown that in the ventral prostate C3(2) was methylated and C3(1) was competely un-

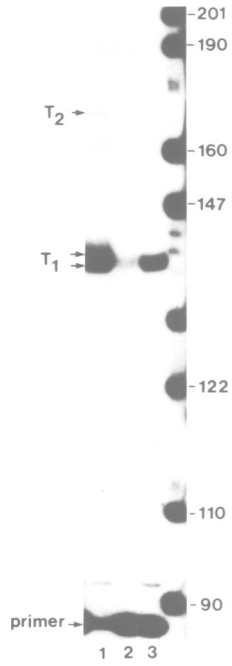

Fig. 5. Effect of androgens on C3 transcript levels. Primer extension was carried out using the ^{32}P-labelled *Xba*I-*Alu*I primer as described in Figure 3A. 35 000 c.p.m. per track of primer was hybridised to 2.5 μg rat ventral prostate nuclear RNA derived as follows: **track 1**, normal animals; **track 2**, animals 3 days after castration; **track 3**, 3-day castrates after 24 h with testosterone. Nuclear RNA samples were provided by M.Page. Markers are ^{32}P-labelled *Hpa*II digested pAT153.

methylated at CCGG sites within and flanking the genes (White and Parker, 1983) suggesting that only C3(1) is transcribed *in vivo*.

As noted previously (Parker *et al.*, 1982) the genes for C1 and C2 share 76% DNA sequence homology and probably result from gene duplication. Since their organisation is identical to that for the two C3 genes it is conceivable that all four genes are derived from a single ancestral gene. However, there are no extensive DNA sequence homologies between all of them to support this possibility. Nevertheless, the genes for C1, C2 and C3 are regulated similarly by testosterone and therefore we have examined their 5′-flanking regions for DNA sequences which may have a functional role in the hormonal control of their expression. Surprisingly there are significant regions of homology with the 18-bp consensus sequence postulated as a progesterone receptor-DNA binding site (Mulvihill *et al.*, 1982) around the cap site of all four genes. These homologies which ranged from 65 to 76%, a value which increased to 72–80% when only the central 16 bp of the consensus sequence were considered, were located relative to the cap site at position 1 as follows: −112 to −94 and 72 to 91 in C1; −182 to −163 and 65 to 84 in C2; −177 to −162, −120 to −107, 6 to 25 and 90 to 111 in both C3 genes. It is noteworthy that there is no evidence to suggest that progesterone modulates the expression of prostatic steroid binding protein.

Inspection of the DNA sequences upstream of the C3 genes revealed two other interesting features. First, there is an Alu-like repetitive element (Jelinek and Schmid, 1982) between nucleotides −450 to −370 upstream of both genes. It shares 98% DNA sequence homology with similar repetitive elements which have been found within rat growth hormone gene (Barta *et al.*, 1981) and α-tubulin pseudogene (Leminschka and Sharp, 1982) and, in addition, it is homologous to a 160-base RNA species expressed in rat brain (Sutcliffe *et al.*, 1982). Since this RNA was not detected in other tissues examined the authors speculated that it has a role in tissue-specific differentiation. Indeed, a model has been proposed (Davidson and Britten, 1979) for the regulation of gene activity which required the existence of repetitive sequence elements adjacent to and involved in the activation of structural genes. In view of this, and since the repetitive elements which are upstream of the C3 gene contain an excellent consensus sequence with the RNA polymerase III promoter (Sakonju *et al.*, 1980; Bogenhagen *et al.*, 1980), it is conceivable that the element is transcribed in the ventral prostate and is of functional significance.

Secondly, there are two dyad symmetries, which may form stem and loop structures, just upstream of the C3(1) gene. Such symmetries have been shown to be part of the promoter for the herpes thymidine kinase gene (McKnight, 1982) and the *Drosophila* heat shock genes where their functional significance has been demonstrated (Pelham, 1982). They are located in C3(1) at −115 to −109 (TATGTTT) and −88 to −82 (AAACATA) which also contains the consensus for the progesterone receptor binding site and at −68 to −61 (TATTTATT) and −32 to −25 (AATAAATA).

In spite of the marked difference in transcriptional activity of the genes for C3 in ventral prostate they are expressed similarly in heterologous cells and cell-free systems. Both C3(1) and C3(2) produced C3 transcripts with 5′ termini corresponding to T_1 and T_2 under hormonal control in mouse mammary gland tumour cells (S115 cells) at ~100–1000 copies/integrated gene/cell (Page and Parker, 1983). In addition, C3(2) produced transcripts *in vitro* of which only those initiated at T_1 and T_2 were sensitive to low doses of α-amanitin (H.Hurst, unpublished observations). Since the DNA used for both of these studies was unmethylated we conclude that the C3(2) gene is inactive in rat ventral prostate because it is hypermethylated. Furthermore, in view of the similarity of the DNA sequence between nucleotide −235 and the first intron, we suggest that signals which may regulate DNA methylation in C3(1) and C3(2) do not reside within the proximal 5′-flanking DNA.

Finally, since the genes for C3 are accurately transcribed and their expression regulated by testosterone in S115 cells (Page and Parker, 1983) it will now be possible to test the functional significance of the DNA sequences described above.

Materials and methods

M13 subcloning and dideoxy sequencing

Restriction fragments from subclones of λ6 and λ11B in pAT153 were eluted from acrylamide gels, ligated into suitably restricted replicative forms of M13 mp7, 8 or 9 (Messing *et al.*, 1981; Messing and Vieira, 1982) and transfected into JM103 (Messing *et al.*, 1981). Recombinant phage were selected and grown up (Sanger *et al.*, 1980) to yield single-stranded DNA.

Chain termination sequencing reactions were performed as described by Sanger and co-workers (Sanger *et al.*, 1977; Schreier and Cortese, 1979) using either a commercially available 15-base universal primer or, occasionally, single-stranded fragments prepared from genomic subclones. The DNA samples were electrophoresed on 0.3 mm 6% acrylamide, 7 M urea gels (Sanger and Coulson, 1978) with the acrylamide cross-linked to the back plate (Garoff and Ansorge, 1981) to allow easier manipulation of the gels. The gels were fixed in 10% acetic acid, dried and autoradiographed at room temperature without screens.

Transcript mapping

The 5′ and 3′ ends of the C3 transcripts were mapped using either primer extension or S1 mapping techniques. Both involved the hybridisation of terminally labelled DNA fragments (Maxam and Gilbert, 1980; Maniatis *et al.*, 1982) to total poly(A)$^+$ RNA prepared from rat ventral prostate (Parker and Mainwaring, 1977). Protocols for the generation of individual ^{32}P-labelled primers and probes are given in the text and figure legends. All primer extension reactions were performed as previously described (Parker *et al.*, 1982). S1 nuclease mapping was performed essentially as described by Kamen and coworkers (Favaloro *et al.*, 1980). 5′ or 3′ ^{32}P-labelled double-stranded probes were hybridised with poly(A)$^+$ RNA in formamide buffer at 52°C overnight. S1 nuclease digestion of the hybrids was for 1 h at 37°C in buffer containing 100 U/ml S1 nuclease. Products from S1 and primer extension mapping were electrophoresed on 6% sequencing gels (Sanger and Coulson, 1978; Garoff and Ansorge, 1981). Autoradiographs of the gels were scanned using a Joyce Loeble Chromoscan 3 densitometer to determine the relative intensities of the observed bands.

Acknowledgements

We would like to thank Shelagh Redmond and Dr.M.Jones for help and advice with the M13 cloning and dideoxy sequencing, and Drs.G.Peters, M.Fried and J.Williams for critical reading of the manuscript.

References

Barta,A., Richards,R.I., Baxter,J.D. and Shine,J. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 4867-4871.
Benoist,C. and Chambon,P. (1981) *Nature*, **290**, 304-310.
Bogenhagen,D.F., Sakonju,S. and Brown,D.D. (1980) *Cell*, **19**, 27-35.
Breathnach,R. and Chambon,P. (1981) *Annu. Rev. Biochem.*, **50**, 349-383.
Brutlag,D.L., Clayton,J., Friedland,P. and Kedes,L.H. (1982) *Nucleic Acids Res.*, **10**, 279-294.
Davidson,E.H. and Britten,R.J. (1979) *Science (Wash.)*, **204**, 1052-1059.
Favaloro,J., Treisman,R. and Kamen,R.I. (1980) *Methods Enzymol.*, **65**, 718-749.
Felsenfeld,G. and McGhee,J. (1982) *Nature*, **296**, 602-603.
Garoff,H. and Ansorge,W. (1981) *Anal. Biochem.*, **115**, 450-457.
Gorski,J. and Gannon,F. (1976) *Annu. Rev. Physiol.*, **38**, 425-450.
Hagenbuchle,O., Bovey,R. and Young,R.A. (1980) *Cell*, **21**, 179-187.
Heyns,W. and DeMoor,P. (1977) *Eur. J. Biochem.*, **78**, 221-230.
Heyns,W., Peeters,B., Mous,J., Rombouts,W. and DeMoor,P. (1978) *Eur. J. Biochem.*, **89**, 181-186.
Higgins,S.J. and Gehring,U. (1978) *Adv. Cancer Res.*, **28**, 313-393.
Jelinek,W.R. and Schmid,C.W. (1982) *Annu. Rev. Biochem.*, **51**, 813-844.
Jung,A., Sipple,A.E., Grezim,M. and Schutz,G. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5759-5763.
Lemischka,I. and Sharp,P.A. (1982) *Nature*, **300**, 330-335.
Maat,J., van Boveren,C.P. and van Ormandt,H. (1980) *Gene*, **10**, 27-38.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) in *Cloning: A Laboratory Manual*, published by Cold Spring Harbor Laboratory Press, NY, pp. 113-116.
Maxam,A. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499-560.
McKnight,S.L. (1982) *Cell*, **31**, 355-365.
Messing,J., Crea,R. and Seeburg,P.H. (1981) *Nucleic Acids Res.*, **9**, 309-321.
Messing,J. and Vieira,J. (1982) *Gene*, **19**, 269-276.
Mount,S.M. (1982) *Nucleic Acids Res.*, **10**, 459-472.
Mulvihill,E.R., LePennec,J. and Chambon,P. (1982) *Cell*, **24**, 621-632.
Page,M.J. and Parker,M.G. (1982) *Mol. Cell. Endocrinol.*, **27**, 343-355.
Page,M. and Parker,M. (1983) *Cell*, **32**, 495-502.
Parker,M.G. and Mainwaring,W.I.P. (1977) *Cell*, **12**, 401-407.
Parker,M.G., Needham,M., White,R., Hurst,H. and Page,M. (1982) *Nucleic Acids Res.*, **10**, 5121-5132.
Parker,M.G., Scrace,G.T. and Mainwaring,W.I.P. (1978) *Biochem. J.*, **170**, 115-121.
Parker,M.G., White,R., Hurst,H., Needham,M. and Tilly,R. (1983) *J. Biol. Chem.*, **258**, 12-15.
Peeters,B., Rombauts,W., Mous,J. and Heyns,W. (1981) *Eur. J. Biochem.*, **115**, 115-121.
Pelham,H.R.B. (1982) *Cell*, **30**, 517-528.
Proudfoot,N. and Brownlee,G.G. (1976) *Nature*, **263**, 211-214.
Razin,A. and Friedman,J. (1981) *Prog. Nucleic Acid Res. Mol. Biol.*, **25**, 33-52.
Sakonju,S., Bogenhagen,D.F. and Brown,D.D. (1980) *Cell*, **19**, 13-25.
Sanger,F. and Coulson,A.R. (1978) *FEBS Lett.*, **87**, 107-110.
Sanger,F., Coulson,A.R., Barrell,B.G., Smith,A.J.H. and Roe,B.A. (1980) *J. Mol. Biol.*, **143**, 161-178.
Sanger,F., Nicklen,S. and Coulson,A.R. (1974) *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.
Schreier,P.H. and Cortese,R. (1979) *J. Mol. Biol.*, **129**, 169-172.
Staden,R. (1980) *Nucleic Acids Res.*, **8**, 3673-3694.
Sutcliffe,J.G., Milner,R.J., Bloom,F.E. and Lerner,R.A. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 4942-4946.
White,R. and Parker,M. (1983) *J. Biol. Chem.*, in press.