Vol. 185, No. 1, 1992 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
May 29, 1992 Pages 404-413

STRUCTURE OF THE MOUSE ACTIVIN RECEPTOR TYPE II GENE

Martin M. Matzuk[1,2] and Allan Bradley[1]

[1]Institute for Molecular Genetics and [2]Department of Pathology
Baylor College of Medicine
Houston, Texas 77030

Received April 13, 1992

SUMMARY: Knowledge of the structures of the activin receptor genes is crucial to our understanding of the role of the activins, inhibins, and their receptors in developmental and physiological processes. The type II activin receptor (ActRc) has been shown to be a transmembrane protein with putative serine/threonine kinase activity. Using a human ActRc type II cDNA as a probe, 18 overlapping λ clones containing portions of the mouse ActRc type II gene were isolated from a mouse 129SvE genomic library. Sequence analysis of the exons, exon-intron boundaries, and 5' and 3' non-translated regions as well as Southern blot analysis of mouse genomic DNA were used to establish the structure of the mouse ActRc type II gene. The mouse ActRc type II gene is encoded by 11 exons and spans >66 kilobases. Two large introns (introns 1 and 4) contribute the majority of the gene size and are found to delineate exons which encode important domains of the activin receptor. Analysis of the 5' region of the gene reveals several putative transcription factor binding sites which may be important for the complex transcriptional regulation of this gene.

The activins and inhibins are developmentally and physiologically important dimeric growth factors (1-3) with structural homology to several other proteins including the transforming growth factor (TGF) β's (4). Both activin and inhibin have been shown to cause responses in many of the same cell types and tissues and in most cases, the effects of these growth factors are antagonistic (1). Previous affinity-labeling studies have shown that activin (5,6), similar to TGF-β (7,8), binds to both a 50-60 kD protein as well as a 70-80 kD protein, which by virtue of their size, have been termed type I and type II receptors. During the past year, ActRc type II cDNA's have been isolated from mouse (5,8), Xenopus (9), and human (10). Massague's group (8) has shown that at least in the mouse, there is a second

0006-291X/92 $4.00

type II ActRc (IIB) which must be encoded by a separate gene than that isolated by Mathews and Vale (5). These different type II ActRc's (5,8-10) are all transmembrane glycoproteins with putative serine/threonine specificity and show homology to the C. elegans gene, *daf-1* (11) and the recently isolated TGF-β type II receptor (12). The type IIB ActRc precursor mRNA also appears to undergo alternative splicing giving rise to ActRc isoforms which continue to bind activin (8).

To understand more about the activin receptor and to predict whether the type II receptor may undergo alternative splicing similar to the type IIB receptor, we have cloned the mouse activin receptor type II gene. This is the first cloning of a gene for this family of receptors. Knowledge of the structure of the type II gene will aid in our future developmental studies and will help us to understand the complex transcriptional regulation of this gene.

MATERIALS AND METHODS

Library Screening

A mouse genomic DNA library was constructed by Shan Lu and Dr. Richard Behringer (M.D. Anderson Cancer Center) from mouse 129SvE DNA using the vector EMBL-3 (Stratagene). 1.1 x 10^6 recombinant λ plaques were plated onto a ER1647 bacterial lawn. Plaques were lifted onto duplicate Hybond-N nylon membranes filters (Amersham). Filters were prehybridized for several hours in 50% Formamide, 5X SSPE (0.75 M NaCl, 50 mM NaH_2PO_4, 4mM Na_2EDTA) and 1% SDS at 42°C and hybridized overnight in 100 ml of the same solution containing 1.5 x 10^8 cpm total of the human type II ActRc cDNA (10) probe. The probe was radiolabeled using (α-^{32}P)dCTP (Amersham) and the T7QuickPrime labeling kit (Pharmacia). The filters were washed in 1X SSC (0.15 M NaCl, 15 mm sodium citrate) and 0.1% SDS twice for 15 minutes and twice for 30 minutes at 65°C prior to autoradiography.

λ genomic clones

18 positive λ genomic clones were identified on initial screening of the genomic library. These λ clones were purified to homogeneity by secondary and tertiary screening using the human ActRc type II cDNA probe. DNA from individual positive λ recombinant isolates were digested with the restriction endonucleases Asp 718 and Sal I and analyzed by gel electrophoresis to orient the genomic DNA with respect to the λ arms. The DNA was transferred onto GeneScreen Plus nylon membrane (Dupont-NEN) and probed with portions of the human ActRc type II cDNA to determine the position of the λ clones with respect to the 5' and 3' end of the gene and to determine overlap. Fragments of several of these genomic clones were subcloned into the appropriate sites of the vectors pBluescript SK(+) or

pBluescript SK(-) (Stratagene) for restriction endonuclease analysis and sequence analysis.

DNA sequence analysis

All sequence analysis was carried out by the Nucleic Acids Core in the Institute for Molecular Genetics at Baylor College of Medicine using automated fluorescent DNA sequencing technology (13). Computational analysis of the 5' region of the mouse ActRc type II gene was performed using the Signal Scan computer analysis program (14).

RESULTS AND DISCUSSION

Isolation and characterization of the mouse ActRc type II gene

A mouse 129SVE genomic library was screened with the coding sequence of the human ActRc type II cDNA (10). The human and mouse type II ActRc's share 95% nucleotide sequence homology in the coding region and differ by only two conservative amino acids (5,10). 18 λ recombinant genomic clones were isolated and 6 of these were used to construct a structural map of the mouse ActRc type II gene as shown in Figure 1. The gene consists of 11 exons and spans >66 kb. Intron 1 (>40 kb; see below) and intron 4 (12.9 kb) contribute the majority of the size of the gene. The other 8 introns are ≤2.5 kb in length. As shown, all 11 exons are contained on the λ clones. However, two λ clones (λ-17 and λ-25) which contain 15.0 and 11.6 kb of intron 1 sequence (total sequence = 26.6 kb), respectively, do not overlap. Southern blot analysis of mouse 129SvE genomic DNA using a probe isolated from the end of λ-26 suggested that

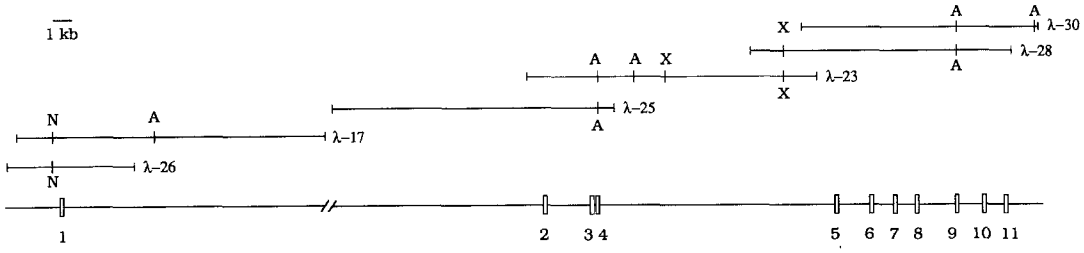


Figure 1. Structure of the mouse ActRc type II gene. The DNA inserts present in overlapping λ recombinant clones are shown at the top along with unique restriction endonucleases which were important in construction of the linear map shown at the botton. The boxes in the map denote exons with the exon number shown below each exon and the line connecting each box denotes introns. The break in intron 1 is where λ clones 17 and 25 terminate without overlapping. Asp 718 sites are present in exons 4 and 9 as shown. The location of the exons and the intron sizes are drawn to scale; the exon sizes are approximate. A, Asp 718; X, XhoI, N, NotI.

there is a gap between these clones of at least 13.5 kb (data not shown). This suggests that the minimum size of intron 1 is approximately 40 kb. Thus, intron 1 encompasses >60% of the size of the mouse ActRc type II gene and the entire gene is >66 kb in length.

Analysis of exons, introns, and exon/intron boundaries

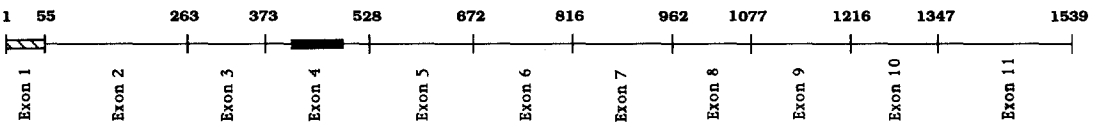
All 11 exons of the mouse ActRc type II genes were sequenced in their entirety and the nucleotide positions of the introns compared to the cDNA coding sequence are shown in Figure 2A. Compared to the mouse ActRc type II cDNA sequence (5), there are no nucleotide differences and therefore the sequence is not repeated here. Exons 2-10 vary in size from 110 bp to 208 bp and satisfy the observation that internal exons are usually <300 bp (15). The coding region of exons 1 and 11 are 55 and 195 bp, respectively.

55 of the 57 nucleotides, which code for the signal peptide, are encoded in exon 1. A portion of exon 4 encodes the entire transmembrane sequence of the ActRc while the remainder of exon 4 encodes small regions of the extracellular and intracellular sequences. The majority of the extracellular sequence (the region which binds the activin dimers), is encoded by exons 2 and 3. The majority of the intracellular domain and all of the sequence which encodes the putative serine/threonine kinase domain, is encoded in exons 5 to 11.

As shown in Figure 2B, the introns vary in size from 152 bp to >40kb. All of the exon-intron boundaries satisfy the GT-AG intron donor-acceptor splice rule although there are a number of nucleotides in the remainder of the junction sequences which are found infrequently at rodent intron/exon boundaries (see Figure 2B). It is particularly peculiar that there are a number of adenosine nucleotides seen in the 3' intron sequences which is normally pyrimidine-rich. Comparison of the intron positions of the ActRc type II gene to the location of the alternative splicing events seen with the type IIB ActRc (8) reveals that the exon/intron boundaries are identical (introns 3 and 4). These data suggest that the structures of the mouse ActRc type II and type IIB genes are likely similar with identical exon/intron boundaries, and it is likely that other type II receptor genes for ligands of this family will have similar structures.

As mentioned above, there are 2 large introns in the mouse ActRc type II gene. The locations of these large introns appear to separate the gene into important structural regions. The >40 kb intron 1 and the 12.9 kb intron 4 separate exons which code for the signal peptide (exon 1), the extracellular ligand binding domain and the transmembrane region (exons 2-4), and the putative serine/threonine kinase domain (exons 5-11). Comparison of the gene structure of *daf-1* (11) to our gene reveals that the

A



B

EXON	(SIZE)	EXON 3′	INTRON (SIZE)	EXON 5′
Exon	1 (>55bp)	TCA G Ser G	gtaggtggaggagcc...intron 1 (≥40kb)....aagttttctcttaccttacag	GT GCT ly Ala
Exon	2 (208bp)	GAC AG Asp Ar	gtaagaatgtattta...intron 2 (~2.5kb)...ttctttctttttttaaaatag	G ACT g Thr
Exon	3 (110bp)	CAG C Gln P	gtaagtgcatagtga...intron 3 (152bp)....ttcctcttccctctttgcag	CC ACT ro Thr
Exon	4 (155bp)	ACT CAA Thr Gln	gtaagttactatcct...intron 4 (~12.9kb)..attattattatttatctctag	GAC CCA Asp Pro
Exon	5 (144bp)	ATA CAG Ile Gln	gtatgtttattgaaa...intron 5 (~1.7kb)...gctttcatttctttactctag	GAC AAA Asp Lys
Exon	6 (144bp)	GAA AAG Glu Lys	gtataacgacataat...intron 6 (~1.2kb)...tcccacctctcctttctttag	GGC TCA Gly Ser
Exon	7 (146bp)	CAC AG His Ar	gtagaacaaaatctg...intron 7 (~1.0kb)...attttatttttatgtttgaag	G GAC g Asp
Exon	8 (115bp)	GGG CAG Gly Gln	gtaagaacaacagct...intron 8 (~2.2kb)...tgatttcaaaatctttttttag	GTT GGT Val Gly
Exon	9 (139bp)	GAT G Asp G	gtaagggaaaaagct...intron 9 (~1.4kb)...tctccttcttttacatttcag	GA CCC ly Pro
Exon	10 (131bp)	CAT GCA His Ala	gtaagtttatacact...intron 10(~1.0kb)...tgactattcttctcctcttag	GGA ATG Gly Met
Exon	11 (>192bp)			

Figure 2. Location of introns and exon-intron boundaries. A) The nucleotide preceeding the intron is shown. Nucleotide 1 is the beginning of the translated protein and nucleotide 1539 is the last nucleotide of the translated sequence. The hatched box is the region that codes for the signal peptide and the solid box is where the transmembrane region is located. B) Exon-intron boundaries and the size of each intron are shown. The consensus donor sequence in rodents is (C/A)AG/GTUAGT and the consensus acceptor sequence is YYYYYYYYYYNCAG/G (Y, pyrimidine; U, purine; N, any nucleotide; Ref. 16). Nucleotides which appear ≤11% of the time in splice donor and acceptor sequences (16) are underlined.

largest intron in the *daf-1* gene is also intron 1 which similarly separates the exon encoding the signal peptide from the extracellular domain. The position of these large introns may have been important evolutionarily in the creation of diversity of the many receptors which will likely be discovered to bind the diverse TGF-β family of ligands.

Analysis of 5' and 3' non-translated sequences

2800 nucleotides of 5' non-translated mouse ActRc type II gene sequence were generated in our laboratory. Mathews and Vale (5) had cloned two different mouse ActRc type II cDNA fragments, one of which was a partial cDNA which started 621 bp upstream of the translation start sites (Figure 3, underlined). Since there are no satisfactory 3' splice acceptor sites upstream of where their two cDNA's overlap, and because internal exons are usually <300 bp (15), the mouse ActRc type II gene would not be expected to have a 5' non-translated exon.

Primer extension and nuclease protection studies to analyze this region were complicated by the fact that the region between residues -349 and -1110 is very highly GC rich (79% GC residues). Therefore, design of primers, which were not highly GC rich, or synthesis of labeled RNA were both difficult. However, one primer generated a long primer extension product of 116 nucleotides using total RNA isolated from mouse testes (data not shown). It is difficult to interpret this result since it ends in a very GC rich sequence (see Figure 3), just downstream of a putative CAP site (indicated) and may therefore be the result of a strong "pause" by the reverse transcriptase. This result though would predict the "start site" of transcription at residue -752 and suggest that exon 1 is 807 bp. This "start site" would lie an additional 131 nucleotides upstream of one of the cDNA's isolated by Mathews and Vale (5). Taking into account that the polyadenylation addition site lies 648 bp 3' of the end of translation (see below) and that the protein coding sequence is 1539 nucleotides, these results would predict that the ActRc type II mRNA is 2939 nucleotides not including the poly(A) tail. These results are therefore consistent with the predicted 3.0 kb size of the mouse ActRc type II mRNA produced in the testes (5).

Further analysis of the 5' region of the mouse ActRc type II gene using a program to predict transcription factor binding sites reveals several interesting findings (Figure 3). Within this region, there are 6 Sp1 binding sites (17), 7 GATA-1 binding sites (18), 5 PU boxes (19), and 1 site that resembles the glucocorticoid/progesterone receptor binding site (20). Among the remaining sites, there are also 2 sequences whch are identical to the core insulin enhancer sequence (21), 1 site identical to the Zif268 consensus sequence (22), and 26 sites which resemble the consensus sequence for interferon-γ induction (not shown; ref. 23). There are no TATA or CCAAT boxes in the vicinity of the putative transcription start site.

Although the appearance of a number of these putative sites could be a chance event, it is interesting to speculate on the presence of several of these. Activin has been suggested to play an important role in regulation of adrenal corticosteroid synthesis and adrenal growth at the hypothalamic,

```
-2800 AGGGGTGGCT ATGACTGAAT TACATAGCAT TTGTAGAAAT TTGTCAAACA TAAGTAATAT
        PU>                                          <GATA-1
-2740 GGAGGAAAAA GGGCAGCCTT AATAATTTCA GCTATTTCTA TCTTCAGTGT GAGCTTTGGA
                                <GATA-1
-2680 ACTGTAGTGG GGATTAATAA CACTATCTTT TAAAGCCACT ATTTCTGGAT TTTTTTTTT
-2620 TTAAAGTACA TTCTGGGCTC TGGGTGGAGC TCATTTGCTT GCTAAGTATG CATGGGGTTC
-2560 TAGGTTGGAT CCCCAGGACC ACATAAATAC AACTCAGTAG TGCATGTCAG TAATTGCAGC
-2500 AATTACAAAG TAGAGGCAGG AGGATCAGAA GTTCAAGGTC ATCTTAGGCT CTCTAGGGAG
-2440 TTCGAAGCTA CCTTGAGCTC CACAAGAGCC CATCCCGGTG CCTCCCAGAA GCACATTTAA
      <GR/PR
-2380 GAACAGAATG CATAGCTGTT ATGACTGGTT GTCACTAATT TCTCCCAGGG GTTTACTTTA
-2320 ACTCTTTGTT AGCAAAGAAT GCTCTTATTC AGAATTATTC TTGAAGTTCT TCTTCATTAC
                                                          PU>
-2260 TTTAGAAAAT AATTTTACGT TCACTCATAT GTCAGACTAG GAGGTGAGAG GAAACTAATA
-2200 TGCAACTCTT GTTGCTTCAC TTGCTCTGTT TTTAGCGATG AAGAAACATC GAGACGATGT
-2140 AGTGACAAGA CAACAGTAGG GCCTGATCCA TCCCTGGGAG AAAGGAGGGT AGATGGTTCT
-2080 AAAGTAAACA CACTCTTCCA TTTTAAGACT GTTAGTGGCA ACTCTTCGTT TCTGTCCAGC
-2020 CTAAACTGTT TGCAGAACTA CATTCGATTT ACTGAGGCAT CCTGAAGACC AGTTTGGGAT
-1960 TTTGTTGCAT AATACCATCA ATTCTGAGCA GGGCTTGAGA ACTGGAAATC TTGGAGAGAG
                                                          GATA-1>
-1900 TTTAAAGAGC CACAATGATT ATTAACCCTG ACAAGTAGTA GAAATAGGCA GATAATTATA
-1840 GGATGTACTT AATAGACAGA AATATAGAGC AGTTTATAGG AAGAGAAAAA GAAAACAAAT
                    INS-C>
-1780 ACAAGTGAGT TTGTGGAAAA ATGTGGGGAT CAGTGTACAC TAAAACATTA ATTACCCATC
-1720 CCACAGTTGC TTCCTTTTTT TCTTGAAAAT GTCTCCAAAC ACAGAGGTTA CTACTGTATT
                                          <GATA-1    PU>
-1660 AAGAAAAAAG GGCAAACTTC TAGGTATTTA CCTATCTACT GAGGAAGCAT TCCTGTCAAA
                                INS-C>
-1600 ACCTTCAGAA TGCTGGAATA AAGTGGAAAA CAAGATCTCC CCCTTTATTA TACCAAATTG
-1540 TGCTAAATGG CAGACGAACT GACTCTAGCA ATCATCCTTT GCACGCTATA GCCAACTTCG
-1480 AATATCCCAA ATAGCCGAGC TTCTGGAGGT AGGGCTCACA TTTTTACATT CTGTATGTTT
-1420 CGGTGCTTAA CTGAACACCC GATCAACAGT TCAACATGGC TTACGGGGTA AAAACAGGTC
                          GATA-1>
-1360 TTACAAGAGG CTAAGAATTG ATAACACCAC CGTCTCTGAC GGTCCAATAC ACTGAAATCT
-1300 GCGATTTTCA ACTTCAAGCC TAATACAGAC TTTCACGGCA ACTAGGAGCA AACCTGTTTG
-1240 CAACACAATG CTTTGTGTAT ACAAGGTGTG GTGGATAGGA TCCGAACCTC ACATCGAGCT
    <GATA-1
-1180 ATCTGGCGAG GGTGTCCTGA CCTGGCGGAC AGCCCTTCCT GCCTTTATTT TTTAGCCCAA
                                Zif268>
-1120 AGACAAACGT GGGGAGGCCT CGGCGGCGGG GGCGCGGCCG GCTCGCGAGC ACCGAGCTCG
                                <SP1
-1060 CTCGGGCAGC AGCGGGTGCG GCTCCCGCCC ACTCGGAGCA AGTCGGCGCG GCCTCGGCGT
            PU>
-1000 CCCCCAGAGG AAAGGCTGGG CGCCGCAGTC GGCCTTAGCA TCGCCCATCC CGCTCGGGCG
                          SP1>
 -940 CCGCGGCTCT CGCGCCCGGG CGGNCACCGC ACTGGGCACA CCGGGCGTCT CGGGGCCGCG
 -880 GCGGGCACCC GCAGGGCCCG CCGTCCCGTC TGCTCGGTCT CGGTTCCCGC AGGCCTCGCG
                                <SP1                          CAP>
 -820 GAGCGCGCTC CGGATCCGGC CCCGCCCCTC CGGCTCGCCC CGCCTCCCGG ACAATCTCCT
                <SP1
 -760 CGCGCTCGCC GCCCGGAACG CGCGCTGCGC CCGGCCGCCG CTCCCGCCGC CGCCGCCACC
 -700 GCCGCCGCGG AGCGTTTCTC CGCCGCGCCT GGCTTCCAGC TTCAGGACGG AACCGGAAGT
          SP1>
 -640 GTAGGGGGCG GGGCCGGCGG AGGCCAGGAG ACCGAAAACG CGGCCGAGCC GGGAGCCGGG
 -580 AGCCGGAGCC ACAGCCTGGA CCAGAACTTG GCCGCCGCCT GCGCCACCGT CTCTCCCGTC
 -520 GCCGTCTCCC GCGGCCGCCT CGCACCGCCA CCGCCGCGAG CTCGGCTACT GTGGTCCGCG
 -460 GACTTTGGGT GTCCGGGTTG AAGGTCGGTC CGGACGTCGG AGCAGCTCGG CTCCCGGACC
        SP1>
 -400 GTCCGGGCGG CTGCGGGCGC CGCCGCCGCC GCCTCGTCGC CTCAGCCAGG CGTTTTGTTC
 -340 GCTGAGACGC GAGAGGCGAG CGGAGTGACA GCTGATTTTG ACAGTGATTC CAACCCGCTT
 -280 TTGTTGTTGT TGGCTTTTCG TTGTTTGGTT TGGTGTGTTG TGCGTGTATG TGGCGGTTTT
 -220 TCCCCCGTGG TACATTTTTT TAAATTGTTG TGCCTTTTTT TTAAATCCTT TTTTTTTCAA
 -160 CCCTGTGAAC GTTTTTTGCT TCCCCCTCCC CCCCGCCTCG GGGCTTCCGA ATATGTTTTA
                                                  PU>
 -100 TGACGGTTGA TTTTACACCA GGAGGTTTGT CTCCGAGGAA GACCCAGGGA ACTGGATATC
  -40 TAGCGAGAAC TTCCTACGGC TTCTCCGGCG CCTCGGGAAA ATGGGAGCTG CTGCAAAGTT
                          <GATA-1
   21 GGCGTTCGCC GTCTTTCTTA TCTCTTGCTC TTCAG/GTGC
```

Figure 3. The 5' sequence of the mouse ActRc type II gene. Numbers refer to distances upstream (negative) or downstream (positive) of the ATG initiation codon. Putative transcription factor binding sites are underlined with the factor name above the sequence. Arrowheads to the right denote that the sequence shown matches the consensus sequence; arrowheads to the left denote the complementary sequence matches the consensus sequence. Also underlined are the following: possible CAP site (-769); possible "start site" of transcription (-752); nucleotide where the published cDNA (5) sequence likely ends (-621); and the ATG initiation codon (+1). The "/" following nucleotide 55 denotes the exon 1/intron 1 boundary.

pituitary, and adrenal levels (24) and the ActRc is synthesized in pituitary corticoticotrophs (5). Therefore, the presence of a glucocorticoid receptor binding site in the promoter region of the ActRc type II gene might allow for glucocorticoid regulation of the receptor and be critical for modulating the effects of activin on the hypothlamic-pituitary-adrenal axis.

Since activin has been shown to have an effect at several steps to stimulate hematopoiesis (1), the presence of several GATA-1 binding sites in the ActRc promoter sequence could be important in regulation of receptor expression in erythropoietic cell types (18) and be important for differentiation of meosderm into hematopoietic tissue (25).

The other factors which may regulate ActRc gene expression in other tissues are: 1) The Zif268 protein, which has been shown to be induced by estradiol and retinoic acid (26,27) and may, therefore, be important in ActRc expression in the uterus and in embryonic stem cells which respond to activin and retinoic acid (28); 2) the insulin core enhancer (21) binding protein which may modulate receptor expression in the pancreas where activin has been shown to stimulate insulin secretion (29); 3) the lymphocyte proteins which are responsible for γ-interferon induction (23)

```
1528 GAATCTAGTC TATGATGGTG GCACCGTCTG TACACACTGA GGACTGGGAC TCTGAACTGG
1588 AGCTGCTAAG CTAAGGAAAG TGCTTAGTTG ATTTTCTGTG TGAAATGAGT AGGATGCCTC
1648 CAGGACATGT ACGCAAGCAG CCCCTTGTGG AAAGCATGGA TCTGGGAGAT GGATCTGGGA
1708 AACTTACTGC ATCGTCTGCA GCACAGATAT GAAGAGGAGT CTAAGGGAAA AGCTGCAAAC
1768 TGTAAAGAAC TTCTGAAAAT GTACTCGAAG AATGTGGCCC TCTCCAAATC AAGGATCTTT
1828 TGGACCTGGC TAATCAAGTA TTTG*AAAAC TGACATCAGA TTTCTTAATG TCTGTCAGAA
1887 GACACTAATT CCTTAAATGA ACTACTGCTA TTTTTTTTAA ATCAAAAACT TTTCATTTCA
1947 GATTTTAAAA AGGGTAACTT TTTATTGCAT TTGCTGTTGT TTCTATAAAT GACTATTGTA
2007 ATGCCAACAT GACACAGCTT GTGAATGTGT AGTGTGCTGC TGTTCTGTGT ACATAGTCAT
2067 CAAAGTGGGG TACAGTAAAG AGGCTTCCAA GCATTACTTT AACCTCCCTC AACAAGGTAT
2127 ACCTCAGTTC CACGGTTGCT AAATTATAAA ATTGAAAACA CTAACAGAAT TTGAATAAAT
2187 CAGTCCATGT TTTATAACAA GGT**ATTAC AAATTCACTG TGTTATTTAA GAAAAAATGG
2245 TAAGCTATGC TTAGTGCCAA TAGTAAGTGG CTATTTGTAA AGCAGTGTTT TAGCTTTTCT
2305 TCTACTGGCT TGTAATTTAG GGAAAACAAA GTGCTGTCTT TGAAATGGAA AAGAATATGG
2365 TGTCACCCTA CCCCCCATAC TTATATCAAG GTCCCAAAAT ATT/GCATAC TTAAGTAGTA
2424 CTTTTTTCTT TAAGGTGTGC TGTGTTTGGG AAATATTTGA AAACATAAAG CATGATTTAA
2484 AATTTTTTAA GTGAGCTGTG ACACTGGAAA GCTCTTCATT TTTCTTTTAA AATAGATTTT
2544 TTTTTCCTAT TTATATATAT AAAATGACGG TGTATTTCTT CTTTTCACCA AACAGTGTGT
2604 GGGATATTCT TATCACTGTT TTATGATCAC CTCAGGAAGT GTCATTACTC AGAATTCTTC
2664 ACTCTGCTTC TAGACTTGTA ACTTTATCAC TGTACTGCGT GGTGCCATCT TGTCAGAGTA
2724 ATATTTGATG TCTGTGATAT GGAGAGAATT AGCTTAGGAT AAAGGTAGTG ACCTTTAAGA
2784 ACAGATTTCA GATATTACTG CTTTAAGACA AATCAGGGAT AACAAGTTGA ACTTATAATA
2844 TGCAATGACA TTAAGAGGTA ACCAGTGCTG GTCTAAGGAA CACAACCCAG CTTCCTCCCC
2904 AAATTACTTT TACAGCTAAT TTTAGATAGN TCATGCCTTA TCCTTACGAA GAAAGTGGAA
2964 TTGATGGCAT GTAGGTACCA NAGTGCTTTA AAAAACATTA TTACATTAGA ATATAATTGG
3024 NTCTTCCCTA AAGCTT
```

Figure 4. The 3' sequence of mouse ActRc type II gene. Numbers refer to distance from the ATG initiation codon of the mouse ActRc cDNA. The stop codon (1540), the polyadenylation signal (2180), an Eco R1 site (2655), and a Kpn I site (2977) are underlined. Nucleotide differences from the published cDNA sequence (5) are as follows: deletion of a C (* after 1851); C for G (1929); C for T (2145); deletion of a TA (two *'s after 2209); and an A insertion (2333). The "/" after 2407 is where the cDNA and gene sequences diverge completely.

and binding to the PU boxes (19) may play similar roles in regulating ActRc expression in lymphocytes which have been shown to have altered γ-interferon production in response to inhibin and activin (30). Future studies using the ActRc type II promoter will lead to a better understanding of the complex regulation of the mouse ActRc type II gene in these various cells and will aid in our understanding of the role of the type II activin receptor in mammaliam development and physiology.

As mentioned above, the polyadenylation signal sequence ends 648 bp 3' of the end of translation (Figure 4). This is at the same position predicted in the mouse cDNA sequence and thus exon 11 is 840 nucleotides for the mRNA produced in the testis and may be 3 kb larger for some other tissues (5). Comparison of our 3' non-translated gene sequence to that of the previously published cDNA sequence (5) reveals some minor nucleotide differences and a major difference after the polyadenylation signal (Figure 4 and legend). 868 bp after the end of translation, the cDNA and gene sequences diverge. To ensure that our sequence was not due to a rearrangement event, we used a portion of this sequence (the Eco R1 to Kpn I fragment) in Southern blot analysis. Comparison of the restriction enzyme pattern using this probe and a more 5' coding sequence probe reveals that our sequence is structurally intact (data not shown).

ACKNOWLEDGMENTS

This work was supported by a grant from the Lalor Foundation (M.M.) and NIH grants 1K11HD00960-01 (M.M.) and HD24613 (A.B.). We thank Drs. Richard Behringer, Thomas Cooper, Glenn Friedrich, Barbara Incerti, Phillipe Soriano, and David Wheeler for aid and advice in these endeavors. We also thank Lisa M. Matzuk and Sylvia Perez for aid in manuscript preparation. The nucleotide sequence data reported in this paper will appear in the EMBL, GenBank, and DDBJ Nucleotide Databases.

REFERENCES

1. Vale, W., Hsueh, A., Rivier, C., and Yu, J. (1990) in Peptide Growth Factors and Their Receptors II (M. B. Sporn and A.B. Roberts, eds.), pp. 211-248. Springer-Verlag, Berlin, Germany.
2. Thomsen, G., Woolf, T., Whitman, M., Sokol, S., Vaughan, J., Vale, W., and Melton, D.A. (1990) Cell 63, 485-493.
3. Mitrani, E., Ziv, T., Thomsen, G., Shimoni, Y., Melton, D.A., and Bril, A. (1990) Cell 63, 495-501.
4. Roberts, A.B. and Sporn, M.B. (1990) in Peptide Growth Factors and Their Receptors I (M. B. Sporn and A.B. Roberts, eds.), pp. 419-472. Springer-Verlag, Berlin, Germany.
5. Mathews, L.S. and Vale, W.W. (1991) Cell 65, 973-982.

6. Hino, M., Tojo, A., Miyazono, K., Miura, Y., Chiba, S., Eto, Y., Shibai, H., and Takaku, F. (1989) J. Biol. Chem. 264, 10309-10314.
7. Massague, J., Heino, J. and Laiko, M. (1991) in Clinical Applications of TGF-β (G.R. Bock and J. Marsh, eds.), pp59-74. John Wiley and Sons, Inc., New York.
8. Attisano, L., Wrana, J.L., Cheifetz, S., and Massague, J. (1992) Cell 68, 97-108.
9. Kondo, M., Tashiro, K., Fujii, G., Asano, M., Miyoshi, R., Yamada, R., Muramatsu, M., and Shiokawa, K. (1991) Biochem. Biophys. Res. Commun. 181, 684-690.
10. Matzuk, M.M. and Bradley, A. (1992) Biochim. Biophys. Acta., 1130, 105-108.
11. Georgi, L.L., Albert, P.S., and Riddle, D.L. (1990) Cell 61, 635-645.
12. Lin, H.Y., Wang, X.-F., Ng-Eaton, E., Weinberg, R.A., and Lodish, H.F. (1992) Cell 68, 775-785.
13. Smith, L.M., Sanders, J.Z., Kaiser, R.J., Hughes, P., Dodd, C., Connell, C.R., Heiner, C., Kent, S., and Hood, L.E. (1986) Nature 321, 674-679.
14. Prestridge, D.S. (1991) CABIOS 7, 203-206.
15. Robberson, B.L., Cote, G.J., and Berget, S.M. (1990) Molec. Cell. Biol. 10, 84-94.
16. Senapathy, P., Shapiro, M.B., and Harris, N.L. (1990) Meth. Enzymol. 183, 252-278.
17. Mitchell, P.J. and Tjian, R. (1989) Science 245, 371-378.
18. Evans, T., Reitman, M., and Felsenfeld, G. (1988) Proc. Natl. Acad. Sci. USA 85, 5976-5980.
19. Klemsz, M.J., McKercher, S.R., Celada, A., Van Beveren, C., and Maki, R.A. (1990) Cell 61, 113-124.
20. Beato, M., Chalepakis, G., Schauer, M., and Slater, E.P. (1989) J. Steroid Biochem. 32, 737-748.
21. Ohlsson, H., and Edlund, T. (1986) Cell 45, 35-44.
22. Christy, B. and Nathans, D. (1989) Proc. Natl. Acad. Sci. USA 86, 8737-8741.
23. Yang, Z. , Sugawara, M., Ponath, P.D., Wessendorf, L., Banerji, J., Li, Y., and Strominger, J. (1990) Proc. Natl. Acad. Sci. USA 87, 9226-9230.
24. Plotsky, P.M., Kjaer, A., Sutton, S.W., Sawchenko, P.E., and Vale, W. (1991) Endocrinol. 1228, 2520-2525.
25. Zon, L.I., Mather, C., Burgess, S., Bolce, M.E., Harland, R.M., and Orkin, S.H. (1991) Proc. Natl. Acad. Sci. USA 88, 10642-10646.
26. Suva, L.J., Harm, S.C., Gardner, R.M., and Thiede, M.A. (1991) Molec. Endocrinol. 5, 829-835.
27. Suva, L.J., Ernst, M., and Rodan, G.A. (1991) Molec. and Cell. Biol. 5, 2503-2510.
28. van den Eijnden-van Raaij, A.J.M., van Achterberg, T.A.E., van der Kruijssen, C.M.M., Piersma, A.H., Huylebroeck, D., de Laat, S.W., and Mummery, C.L. (1991) Mech. Develop. 33, 157-166.
29. Totsuka, Y., Tabuchi, M., Kojima, I., Shibai, H., and Ogata, E. (1988) Biochem. Biophys. Res. Commun. 156, 335-339.
30. Petraglia, F., Sacerdote, P., Cossarizza, A., Angioni, S., Genazzani, A.D., Franceschi, C., Muscettola, M., and Grasso, G. (1991) J. Clin. Endocrinol. Metab. 72, 496-502.