

Structure, characterization, and expression of the rat oxytocin receptor gene

(G-protein-coupled receptor/uterus/acute-phase response elements/initiation of labor)

FLORENCE ROZEN*, CATERINA RUSSO*, DENIS BANVILLE†, AND HANS H. ZINGG*‡

*Laboratory of Molecular Endocrinology, Royal Victoria Hospital, McGill University, 687 Pine Avenue West, Montreal, PQ, Canada H3A 1A1; and

†Biotechnology Research Institute, 600 Royalmount, Montreal, PQ, Canada H4P 2R2

Communicated by Donald W. Pfaff, The Rockefeller University, New York, NY, October 10, 1994

ABSTRACT The multiple hormonal and neurotransmitter functions of the nonapeptide oxytocin are mediated by specific oxytocin receptors (OTRs). In most target tissues, the number of OTRs is strongly regulated. Specifically, in the uterus, a dramatic OTR upregulation precedes the onset of parturition. To study the molecular mechanisms underlying OTR regulation, we have isolated and characterized recombinant bacteriophage λEMBL3 genomic clones containing the rat OTR gene, using sequence information derived from a human myometrial OTR cDNA. The rat OTR gene spans >20 kb and contains three exons. A 97-bp intron is in the 5' untranslated region and a >12-kb intron interrupts the coding region between transmembrane domains 6 and 7. The promoter region lacks an apparent TATA or CCAAT box but contains multiple putative interleukin-response elements [six NF-IL6 (C/EBP β) and four APRF (STAT3) binding motifs], supporting the notion that interleukins may mediate labor induction via transcriptional activation of the OTR gene. The predicted amino acid sequence is 93% identical to the human OTR sequence but only 48% and 38% identical to the rat V1 and V2 vasopressin receptor sequences, respectively. At parturition, the OTR gene is highly expressed in the rat uterus and gives rise to at least three transcripts (2.9, 4.8, and 6.7 kb) which differ in the length of their 3' untranslated regions.

The hypothalamic nonapeptide oxytocin (OT) exerts multiple biological actions both as a hormone and as a neurotransmitter. These include uterine contractions during parturition, milk ejection from the lactating mammary gland, and induction of specific mating and maternal behaviors (1–3). In most target sites, the number of OT binding sites is highly regulated (1, 4). In the uterus of each mammalian species studied, OT binding sites undergo a 10- to >100-fold upregulation during pregnancy, reaching the maximum prior to the onset of parturition (1, 5). Indeed, it has been proposed that this dramatic increase represents the trigger for parturition (5). The recent cloning of cDNAs encoding the human OT receptor (OTR) (6) and the human (7), rat (8), and pig (9) V₂ and rat V₁ (10) vasopressin receptors indicates that these receptors are a subgroup of the family of G-protein-linked receptors. As a first step toward an understanding of the molecular mechanisms underlying the dynamic control of OTR expression, we have characterized the rat OTR gene and its 5' flanking region.[§]

MATERIALS AND METHODS

Screening and Sequencing Procedures. A λEMBL3 genomic library (Clontech) was subdivided into 40 pools, and aliquots of each pool were analyzed by PCR using pairs of oligonucleotides allowing the specific amplification of different regions of the rat OTR gene (Fig. 1). Phages from positive

pools were replated and screened with probe 1 or probe 2 (Fig. 1) by established procedures (11). Phage DNA was purified by polyethylene glycol precipitation and analyzed by Southern blotting. Exon-containing restriction fragments were subcloned in pBluescript KS (Stratagene) and sequenced on both strands by the dideoxy chain-termination method using primers corresponding to flanking vector sequences or internal rat DNA sequences.

Rapid Amplification of cDNA Ends (RACE). RACE was performed essentially as described (12). cDNA was prepared by using 0.25 unit of avian myeloblastosis virus reverse transcriptase (Life Sciences, St. Petersburg, FL), 5 μg of random primers and 5 μg of poly(A)⁺ RNA from a rat uterus at parturition. The cDNA was G-tailed by 0.6 unit of terminal deoxynucleotidyltransferase (GIBCO/BRL) using 0.05 mM dGTP. Two consecutive PCRs were performed with two nested sets of primers. In pair 1, the forward primer was 5'-GCCGCCGCTCGAGTCGACATCGAT(C)₁₄-3' and the reverse primer (R1) was complementary to nt 1987–2010 (see Fig. 2). In pair 2, the forward primer was as above but without the C tail and the reverse primer (R2; see Fig. 3B) was complementary to nt 1957–1977. Amplification products were subcloned with the TA cloning kit (Invitrogen). Additional RACE experiments were performed with the same protocol, but using reverse primers that were complementary to sequences closer to the cap site. In the second experiment, the R1 and R2 priming sites corresponded to nt 1674–1693 and 1636–1655; in the third experiment, the R1 and R2 sites corresponded to nt 1636–1655 and 1593–1612, respectively.

RNase Mapping of Cap Site. A 406-bp template for *in vitro* cRNA transcription was synthesized by PCR using subclone 5-1 (Fig. 1) as a template, a forward primer corresponding to nt 1494–1513, and a reverse primer which contained a T7 polymerase binding site attached to an additional 20 nt complementary to 1854–1873. *In vitro* transcription (11) used 50 units of T7 RNA polymerase and 2 μCi of [α -³²P]UTP (8000 Ci/mmol; 1 Ci = 37 GBq). The resulting 380-nt transcript was purified by electrophoresis and hybridized to 10 μg of uterine poly(A)⁺ RNA in 40 mM Pipes, pH 6.4/0.4 M NaCl/1 mM EDTA/80% formamide for 18 hr at 45°C. After digestion with RNase A (40 ng/μl) and RNase T₁ (2 ng/μl) at 30°C for 45 min, the protected fragments were electrophoresed in a denaturing 5% polyacrylamide gel along with *in vitro* synthesized RNA size markers and a DNA sequencing ladder.

RT-PCR. RNA was extracted from uteri of parturient rats with 4 M guanidinium thiocyanate and purified by ultracentrifugation through 5.7 M CsCl and adsorption to oligo(dT)-cellulose (Pharmacia) (13). RT was performed with 250 ng of

Abbreviations: APRE, acute-phase response element; APRF, acute-phase response factor; IL, interleukin; OT, oxytocin; OTR, OT receptor; RACE, rapid amplification of cDNA ends; RT, reverse transcription; UTR, untranslated region.

[‡]To whom reprint requests should be addressed.

[§]The sequence reported in this paper has been deposited in the GenBank data base (accession nos. U15169 and U15280).

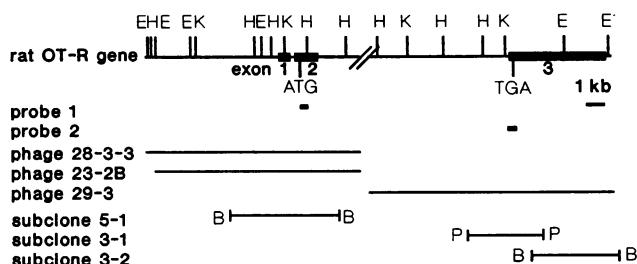


FIG. 1. Structure and restriction map of the rat OTR gene. Exons 1–3 are indicated by heavy bars. E, *Eco*RI; K, *Kpn* I; H, *Hind*III; B, *Bam*HI; P, *Pst* I; ATG and TGA, translational start and stop sites, respectively. The extent of phage clones, plasmid subclones, and probes used are indicated. Probe 1 corresponded to a 257-bp genomic fragment encoding the N-terminal part of the rat OTR. The amplification primer corresponded to nt 1–27 (forward primer) and the reverse complement of nt 231–257 (reverse primer) of the published human OTR cDNA (6). Probe 2 was a 188-bp fragment encoding part of the C terminus of the rat OTR. The amplification primers corresponded to nt 3621–3640 (forward primer) and 3789–3808 (reverse primer) of the rat OTR gene (Fig. 2). The primer sequences used for the generation of probe 2 were derived from a partial cDNA clone generated by reverse transcriptional (RT)–PCR using rat uterine RNA and primers corresponding to nt 2820–2839 (Fig. 2) of the rat OTR gene (forward primer) and to the reverse complement of nt 1145–1165 of the published human OTR cDNA (5) (reverse primer). The primer pairs used to generate probes 1 and 2 were also used in the initial screening of 40 individual phage pools derived from a rat genomic library (see *Materials and Methods*).

RNA, random or sequence-specific primers, and 2 units of Gene Amp rTth reverse transcriptase (Perkin–Elmer/Cetus). PCR amplification was performed with the same enzyme after chelation of Mn²⁺ by 0.75 mM EGTA and addition of 25 mM MgCl₂ and specific amplification primers (see *Results and Discussion*). Amplification products were analyzed by electrophoresis in 0.8% agarose.

RESULTS AND DISCUSSION

Isolation of Phage Clones Encoding the Rat OTR Gene. Screening of a rat λEMBL3 genomic library with a PCR-derived 257-bp genomic fragment encoding the N-terminal part of the rat OTR (probe 1; Fig. 1) resulted in the isolation of two nearly identical but distinct phage clones (28-3-3 and 23-2B; Fig. 1). A *Bam*HI fragment of phage 28-3-3 (subclone 5-1) contained the first 306 codons. Comparison with a partial rat uterine OTR cDNA sequence indicated that the coding region was interrupted by an intron within codon 307. Since the 3' end of the coding region was not present in phage 28-3-3, the library was rescreened with probe 2 and a single positive phage clone (phage 29-3) was isolated which contained the 3' end of the intron, the remaining 245 nt of coding region, and the entire 3' untranslated region (UTR). Phages 28-3-3 and 29-3 contained 3.1 and 9 kb of intronic sequence, respectively (Fig. 1). Since cross-hybridization experiments indicated that phages 28-3-3 and 29-3 did not overlap, the size of the intron present within codon 307 is >12 kb. Southern blot analysis of total rat genomic DNA showed that a PCR-derived probe encompassing codons 10–252 hybridized to single *Hind*III, *Eco*RI, and *Bam*HI fragments, indicating that the gene is unique in the rat genome.

Analysis of the Coding Region. The predicted amino acid sequences of the rat and the human OTR are 92.8% identical. Two consensus sequences for N-linked glycosylation as well as numerous potential phosphorylation sites are conserved (Fig. 2). The amino acid sequence identity with the closely related rat V1a and V2 vasopressin receptors is only 47.7% and 38.1%, respectively. As in the human V2 vasopressin receptor gene (14), the coding region is interrupted by an intron in the region

that encodes the third putative extracellular loop. Despite limited nucleotide sequence identity surrounding the splicing site, amino acid alignment indicates that the two introns are located in exactly homologous positions and are in the same phase.

Analysis of the 5' UTR. RNase mapping with a 380-nt uniformly labeled cRNA probe complementary to nt 1873–1494 gave a major protected band of 281 nt, indicating that the major start site is at nt 1593 (Fig. 3C). Transcriptional start sites were also mapped by RACE. Three different two-step amplification schemes were used, each utilizing two nested primer pairs (Fig. 3B and *Materials and Methods*). In six of the clones analyzed, the cDNA end was located at nt 1593, in agreement with the RNase protection experiments (Fig. 3A). However, four additional cDNA clones were identified ending at nt 1582, 1585, 1603, and 1620. These data suggest that initiation may also occur at sites close to nt 1593. In two cDNA clones obtained by RACE, nt 1709–1805 were missing. This finding, together with the fact that this segment is flanked by sequences matching the exon/intron boundary consensus sequence (15), indicates that this segment corresponds to an intron that is, however, not consistently spliced out in the uterus. The 5' UTR also contains a short open reading frame of 7 codons.

Analysis of the Promoter Region. There are no typical TATA or CCAAT homologies in the vicinity of the transcriptional start sites (Fig. 2). However, the proximal promoter region contains APREs known to be involved in mediating the effects of interleukins (ILs): (i) the T(G/T)NNGNAA(G/T) motif, recognized by NF-IL6; (ii) the CTGGGA motif, recognized by the acute-phase response factor (APRF); and (iii) the TGGCA motif, an element recognized by an NF-1-related factor that is found associated with IL-1-responsive promoters (16). The NF-IL6 binding motifs are present six times (twice in reverse orientation) and are centered 72, 150, 319, 332, 947, and 1035 bp upstream of the main cap site. The transcription factor NF-IL6, which is identical to C/EBPβ, was initially identified as the factor mediating IL-1 induction of the IL-6 gene and was later found to induce the expression of a number of acute-phase proteins in response to IL-1, IL-6, tumor necrosis factor, and nerve growth factor (17). The second type of APRE (CTGGGA) occurs four times (twice in reverse orientation). Three motifs are centered 14, 536, and 1141 bp upstream of the cap site and one motif is within intron 1. This motif represents the binding site for APRF (or STAT3), a member of the STAT family of transcription factors (18). The motif is present and functional in the rat fibrinogen gene and a series of other IL-6-regulated acute-phase response genes (18). Whereas NF-IL6 induction is *ras*-dependent, APRF activation involves a *ras*-independent pathway (18).

The presence of APREs in the OTR gene promoter suggests that the acute induction of OTR expression at the onset of parturition may be a phenomenon mechanistically similar to the fast induction of acute-phase response genes. This notion is strengthened by the observations that the uterus is populated by macrophages and other specific lymphocytes (19, 20). Specifically, at term, nearly half of the decidual cells are of bone marrow origin (20). IL-1β released from macrophages stimulates the production and release of IL-6 by uterine stromal cells (21, 22). Moreover, IL-1 is a central pathophysiological mediator of infection-induced premature delivery (23), and preterm delivery can be prevented by an IL-1 antagonist in mice (24). We speculate that under physiological as well as pathophysiological conditions, inflammatory cytokines are important inducers of labor and that this mechanism involves the cytokine-induced transcriptional activation of the OTR gene.

Despite the fact that estrogens induce a strong upregulation of uterine and hypothalamic OT binding sites (1, 2) and uterine and hypothalamic OTR mRNA (unpublished observations),

1 tggaccaagttgcgttccactgtctttagaattcacttcactactacactgaaaatagacaagcgccttatccctggteagggtcaaggatgtcagtgtgccttcacc
 121 aggaccttaggggggtcaactaggcaagcagccccctctttcggttccacacagactgagaattaaacaactgtggatgtcacttcgttcatgc
 241 agggcaactagagaacttagtttgcgttccatgcgttcaatgcacttgcataagaatcattttgtcacaagaccacacaactggcgttagggaaaggagcc
 361 tgggtgaagtgcgtactaacgccttgcggaaaggaaatggctgttctcactcggcacttccatgcggatgttgccttccatgcgttcaactcagggtcacc
 481 agatccttcattcaggcatggtgaaacaggaatattagacttgcgttgcacttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
 601 agcggtaaagcggaccacaaatcccggaaatattgg
 721 gcatgtcagcggaaaggacacttcaacaacaaatgg
 841 cagcccttcatatgg
 961 gatacaatccaaagcaagaagaaatgggttttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
 1081 taatcgggggactgaggaaacactgtcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
 1201 catttctcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
 1321 atctggcttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
 1441 gggatcacagggt
 1561 gcccgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
 1681 ACTGGCTCGTGCCTGAGTGAGACTCAGttaatcccgg
 1801 ttccggAGGAGGACAGGCAGGGTAGACACCTGGACCCGAGCACCGCCCCACGGCTGGCCCTTTAAAGAATCGTAAGACTTAGTGTCCCCTGGAGG
 1921 ATGGAGGCCAGCCAGCAGCAACTGGACTGTGAGTTGGACCTCGGGAGTGGAGTGGCCACCCGGACCGCCACGGCCACCCACAGCAACGAGG
 1 MetGluGlyThrProAlaAsnTrpSerValGluLeuAspLeuGlySerGlyValProProGlyGluGluGlyAsnArgThrAlaGlyProGlnArg
 2041 GTGGAGGTGGCGCTCTGTGCTCATCTGTTCTGGCTCTGAGTGGCAACCGCTGTGCTGGCCTGCGCACCACCGACTCGGCCCTCTCTCTCATGA
 41 ValGluValAlaValLeuCysLeuIleLeuPheLeuAlaLeuSerGlyAsnAlaCysValLeuLeuAlaLeuArgThrThrArgHisLysHis
 2161 CTGAGCATCCGTGACCTGGTGTGCTGCTCAGTGTGCTCGACCTTCGGCAGCTGCTGGACATCACCTCCCGCTTATGGCCCGACCTGCTG
 81 LeuSerIleArgAspLeuValAlaValPheGlnLeuLeuLeuTrpAspIleThrPheArgPheTyrGlyProAspLeuLeuCysArgLeuVal
 2281 GTGGGCATGTCGCCCTCACCTACCTGCTGCTGATCGTGCACCGCTGTGCTGGCCATCTGCCAGCCGCTGCCCTCGCAGCCGCCAACCG
 121 ValGlyMetPheAlaSerThrTyrLeuLeuLeuMetSerLeuAspArgCysLeuAlaIleCysGlnProLeuArgSerLeuArgArgThrAsp
 2401 TGGCTTGCTGCCAGTGTCCGGCACTTCTCGTGTGCGAACGAGTGGCGAGCGAGTCTCGTGTGCTGGCGGTCTTCATCCAAACCTGG
 161 TrpLeuGlyCysLeuValAlaSerAlaProGlnHisIlePheSerLeuArgGluValAlaAspGlyValPheAspCysTrpAlaValPh
 2521 GTCACATGGATCACGCTGCCGTACATTGTACCGGTATCGTGTGCTGGCCGCTGCTACGGCCTCATCAGCTTCAGAATCTGAGACTCA
 201 ValThrTrpIleThrLeuAlaValTyrIleValProValIleValLeuAlaAlaCysTyrGlyLeuIleSerPheLysIleTrpGlnAsn
 2641 GCTGCCGAGGGAAATGACGCCGGGGTGGAGCTGGCGCTGGCACGGGTCAAGCTATCTCAAGGCAAATCTGGCACAGGTTAGTGG
 241 AlaAlaGluGlyAsnAspAlaAlaGlyGlyAlaGlyArgAlaAlaLeuAlaArgValSerSerValLysLeuIleSerAlaLysIle
 2761 CTGGCCTTCATCGTGTGCTGGACACCTTCTCTCGTGCAGATGTGGAGCGTTGGAGCTCATGGCCTCAAGGAAGGAGtagtgtgg
 281 LeuAlaPheIleValCysTrpThrProPhePhePheValGlnMetTrpSerValTrpAspValAsnAlaProLysGlu
 VI
 2881 acttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
 3001 actacaggatccatccacaaatccacaaatccacatccatccatccatccatccatccatccatccatccatccatccatccatccatcc
 3121 gcaggcgactgtctttacagagagaaacgaaatgggttt
 3241 ggggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttaggtttag
 3361 tcatggggcaggtatagagagagatgttt
 ...->12 kb /---->
 3481 gtcaatgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
 307 3601 TTTCATCATGCCATGCTCTGGCCAGCTCACAGCTGCTGCAACCCGTGGATCTACATGCTCTTCACAGGTACCTCTGTCAGGGCTTCT
 310 PheIleIleAlaMetLeuLeuAlaSerLeuAsnSerCysCysAsnProTrpIleTyrMetLeuPheThrGlyHisLeuValGlnArgPhe
 VII
 3721 TACCTGAAGGGCAGCCGCCGGCAGACGAGTGTGAGCAAGAGAGCAACTCATCCACCTTGTCTGAGCGTCCAGCTAAGCCAGGGAGCT
 350 TyrLeuLysGlySerArgProGlyGluThrSerValSerLysSerAsnSerSerThrPheValLeuSerArgArgSerSerGlnProSer
 3841 GCCATAGGCCGACGGGGCAGCCGAGGG
 3961 AGTGGCTGAGGGGTGTGCTATTTCTGACGG
 4081 GACTCAAGTCTCTGG
 4201 AGTGGCTGACAAGGAGTGGGGAAATGAGGCTGGGGACGGAGTGGTGTGTAATTTCGGGCTGACCTGAGTGTGTTGGAGCAGAT
 4321 CCGGAAACTGAGCTGG
 4441 TGAGGG
 4561 CT
 4681 AGTGAAGGGGGCTCCATCATAAAAGTCCCAAGATGAGGAAAGGAGGGTTAGGATAACAAAGGGAATGTAACCTCGCCCTCAAG
 4801 AATAGATAATGGGGGACCTTCGCTGGAGACATAATTGTTGATGATCTCAGGAGATTTGATGATCTGAGGAAATTGATAAAG
 4921 AAAATCCCAAGGAGATGGGAATTGG
 5041 CACTCAGAGAAGGCTGTGCTGAGCAAGTGTGAGAACCCAACTTGTGAGCACCTCAGGGCTCATGTAGAGGAGCAGGCTCAG
 5161 TCCCCTACATCACAGGGGCCGCGCCAGGACTGATCACTTGTGAGTCAGGCAAAAGACTCATCTCAAGGAAATTAAGGTTGG
 5281 CCTTGGGCTCTCATGATCTGGGATGACACATCACAGGATGACATCACACACACACACACACACACACACACACACACAC
 5401 AAAAAAATTTAGAGCTGAGTCTGGGAAAACCCAAATGATGGATGACTTCTGGGCTATTTCTGAGTGTGAGTGTGAGTGTG
 5521 ATGCTGGGCTGTAGATGCGTCTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGG
 5641 CTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGG
 5761 TCTGTGGCTGGTACTGTGAGTCTGGGAGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
 5881 AACCCACTCTGCCCTCCCTCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
 6001 GTCCAATGTGTCAGGGAGGG
 6121 GCTAGCAGTGTCCACCTCTGCTTGTGAGTGTGAGGAGTAAACAGGTTCTGGGCTTGGGCTTGGGCTTGGGCTTGGG
 6241 GTGGGAGGCTGTAGCATGTCAGTGTGAGTGTGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
 6361 ATCTGACCCGG
 6481 TTACATAAGTGTGAGTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
 6601 CTCTAGGGCTTCCAAACACACTCCCTCCCACTTGTGCGGCTGCAACTCTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG
 6721 TAGGGTTGG
 6851 CCTACATTTAGGAAATCTGAGGCCCCCTTGGGACAAATCTGAGGCTTCACTTCAAGGAGGAGGAGGAGGAGGAGGAGGAG
 6961 TTGTTATTCTCATGAG
 7081 TGCTTCTTACAGCTGCTTGTGATAATTTGGGCTTCTTCTGTTATTTCTGTTCTTCTGTTCTTCTGTTCTTCTGTTCTT
 7201 CTGGCAGGCTCTGGGCTCATCTCTCGTGTGACACGTCAGTGGCTGGGCTTGGGAGGGGGGGGGGGGGGGGGGGGGGG
 7321 AGGATCTTGAGGCAAATGTTGCTGTGTTCTAGTGGAAATGGGAAATGGGAAATGGGAAATGGGAAATGGGAAAGCAGT
 7441 TGAGGTGTTGTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG

FIG. 2. (Figure continues on the opposite page.)

7561 ACGGCTGGTGGTGCGCACAAGCTGCATGAAAGGTTAGTGTGTTCCGCTCCGTGCTGAGCCAGAGCGGGATTCCCTGGGGCTGAAGTCGGGCCATTTCAGCAGTCCCCGG
 7681 GATGCTCACCTGCCTCACTCAGTCACCTCAGCTCCTCATGAAAGCAGCAGAAGCTGGAAACATTCTCATATGGAAACACCTATTGGCAAGCAACAAGTAGAGATCTGGAAGGA
 7801 AACAAAATTAAGTTCTGCAATTGAGGACCAACACTGAAGTCTGGGTTTCAAGAAACTAGAAAAGGACACACACATGGGGCTTACAGACAGACATAGTGTCCCCTGGTCAGC
 7921 CGAACACCTTATTGAGCAGCACCTCTCATGCTGGCAGTCACCGCTCTCCCTAGCATGAGGAGACGGGCTACCCCTAGAAGACTGGTGTGGTTGTGCTTACAGTACATGT
 8041 GATGCGAACAGCTTGTGATAGCTGGAACTGAGGAAATCTTAAAGCATGAGCTGCCAGCAGCCCCATGGAGCAGACCTGGGTGACACCTGGCCCTATCCCCCTGCCAGCC
 8161 GCACCATGACTCCCCATTGAGGTTGAAGCTGAGGAAATCTTAAAGCATGAGCTGCCAGCAGCCCCATGGAGCAGACCTGGGTGACACCTGGCCCTATCCCCCTGCCAGCC
 8281 TTCCCCACAGCCAGCCGGCTTAGCTCCCTGAGCACCCACCTTAATGTTGCTGCGACGCCCCAGATGAGGAGACAGGAAACAGCAGGAGACAGGAGACGGGCAAGTGAAGTAGA
 8401 TGTCGCTGATGACTGACTTCATGAGGAGCTCTTAATGAGGAGACAGGAGACAGGAGACAGGAGACAGGAGACAGGAGACAGGAGACAGGAGACAGGAGACAGGAGACAGGAG
 8521 GCATGTCAGACAGCAGGGTAGCAGCAGCTGGAGACGTTGAAGTGTGACTCACCCCTCACAGCCATGAGCTGAGTGCCTGGGCTGGCAATCACATCTTCAAAA
 8641 TCCACCTGAGGAGCAGAACAGGAAAGGTTAACACAGCAGCTGGGCTTTCTAGTGTGCTGCTGATCGGAATCTGGGGTGAAGTCCAGGGACATGGTTAG
 8761 CAGCTTAAGAGGACTGTGCTGAGAACAGCTGAGTTGCTCTTACAGATGAGCTGGCTACAGGAAATCTGAGCTGGCTACAGGAAACAGGAAATCTGAGCTGGCTACAGGAA
 8881 GGGTCCCAGATTACAGAATGAATATAATTGGAAACAGCACTGGAGGAGCTGGTGTGAGTTCAGGAGCAGGCTGGTCTACAAATGAAAGTCCAGGACAGCAG
 9001 CTGGCTACAAATGAAAGTTCAGGACAGCCAGGGCTGTTACACTGATAATAAAACAAACAAATAAGCAACACAAAAACAGGAAAGGAGAAACTCTGTTAGTGTGTTGAG
 9121 CCTCTTGATGGTGTCTGACAGGGCTGAGGAGCTGTTACATGAGTGTGCAATATGGACAATAATATTCTGCGAGCCCTGCTGACTTGCCATGACTGAGTACAAGGACTCTG
 9241 AGGCTGTTAGCAGCTTCATGATGTTAGAGGAGCTCCAGAGTGTGAAATGCTGTTCAACTGCACTGAGTCAAGTGCAGTGGCTGATAATTGGATCC

FIG. 2. Nucleotide and deduced amino acid sequences of the rat OTR gene. Potential binding sites for known transcription factors are indicated. APRE, acute-phase response element; TGACC or GGTCA, half of the classical estrogen response element palindrome; stars, transcriptional initiation sites. The main initiation site is at position 1593 and was therefore defined as the beginning of exon 1. A short upstream open reading frame is underlined. Amino acids that differ between the rat and the human sequence are indicated in bold type. The predicted transmembrane domains are underlined. ♦, Potential glycosylation site; ~, potential phosphorylation site. In the 3' UTR, sequence clusters of ≥ 6 nt conserved between the rat and the human gene are indicated in bold type. Potential polyadenylation signals are printed in bold and underlined. ATTAA elements (A+U-rich elements, AREs) are italicized and underlined.

there is no classical palindromic estrogen response element in the 5' flanking region analyzed here. This does not exclude, however, a potential direct effect of estrogens on OTR gene expression. In the case of the chicken ovalbumin gene, activation by estrogens is mediated by several widely spaced half-palindromes (TGACC or GGTCA), with one such element located in the promoter proximal region (25). As shown here, the OTR gene 5' flanking region contains a total of six widely spaced half-palindromes, with one centered 25 bp 5' to the cap site.

Other features include a potential Sp1 binding site (in the 5' UTR), a (GT)₂₆ dinucleotide repeat, and an AP1 site. Preliminary transfection experiments indicate that a 2.2-kb promoter fragment of the rat OTR gene can direct the expression

of a luciferase reporter gene in myometrium-derived hamster cells, but not in rat fibroblasts (unpublished observations).

Analysis of the 3' UTR. To rule out the presence of additional introns and to delineate the polyadenylation site, RT-PCR was performed with a series of primer pairs distributed along the 3' part of the gene (Fig. 4). With uterine cDNA as a template, all primer pairs resulted in the predicted PCR

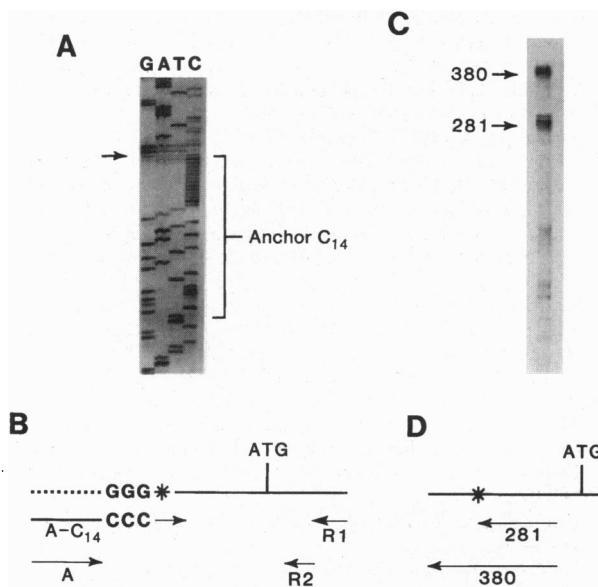


FIG. 3. Determination of transcriptional start sites by 5' RACE and RNase protection analysis. (A) Dideoxy sequencing reactions of RACE product (sense strand). The sequence corresponding to the upstream primer (anchor plus C₁₄ tail) is bracketed. Arrow, transcriptional initiation site, corresponding to position 1593 in Fig. 2. (B) RACE strategy. The first amplification step used G-tailed cDNA, a downstream primer (R1), and an upstream primer (anchor plus C tail). The second step used primer R2 and the anchor sequence alone as primers. (C) RNase protection analysis. Protected fragments were electrophoresed in a denaturing 5% polyacrylamide gel. The sizes of the undigested probe (380 nt) and the main protected fragment (281 nt) are indicated. (D) Outline of RNase protection strategy. The cRNA probe (380 nt) and the main protected fragment (281 nt) are indicated.

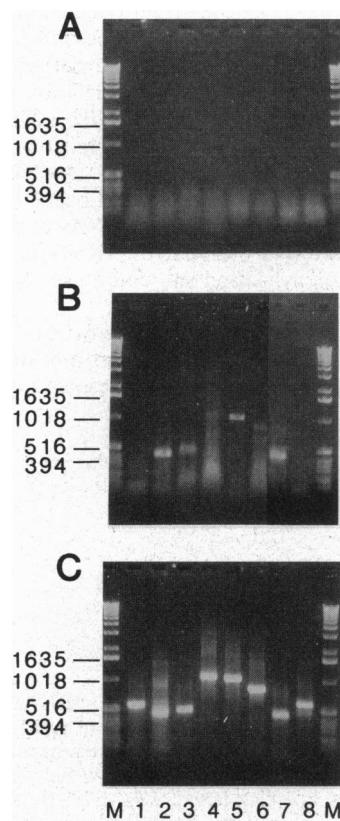


FIG. 4. RT-PCR analysis of the 3' part of the rat OTR gene. Uterine poly(A)⁺ RNA (250 ng) was reverse transcribed and amplified by using primer pairs designed to amplify the following sequence fragments (numbering according to Fig. 2, predicted sizes in parentheses): lanes 1, 3484–4041 (558 bp); lanes 2, 3621–4041 (421 bp); lanes 3, 4642–5136 (495 bp); lanes 4, 5302–6388 (1087 bp); lanes 5, 6592–7656 (1065 bp); lanes 6, 7724–8570 (847 bp); lanes 7, 8585–9031 (447 bp); lanes 8, 8585–9187 (6034 bp). (A) Mock RT-PCR. No reverse transcriptase was present during the RT reaction (negative control). (B) As in A, but 2 units of Gene Amp rTh reverse transcriptase was added to the RT mixture. (C) As in B, but the RT step was omitted and 10 ng of genomic phage 29-3 DNA was used as template (positive control). Lanes M show size markers.

amplification products with the exception of pair 1, which contained a forward primer located in the 3' end of intron 2, and pair 8, which contained a reverse primer complementary to nt 9158–9187. A reverse primer located 156 bp further upstream, in conjunction with the same forward primer as the one used in pair 8, resulted in a band of the predicted size (Fig. 4, lane 7), suggesting that the OTR gene transcript ends between the two reverse priming sites used in pairs 7 and 8, respectively. Indeed, there is a consensus site for polyadenylation centered at 9052 that is preceded by a CACTG motif corresponding to the consensus CA(C/U)UG (26) and followed by a GT-rich cluster starting at 9104. Use of this polyadenylation site would result in a transcript of 6.6 kb [not including the poly(A) tail].

Expression of the OTR Gene. As shown by Northern blot analysis (Fig. 5, lane 1), the OTR gene is strongly expressed in the rat uterus at the moment of parturition. The largest transcript detected was estimated at 6.7 kb. A probe derived from a sequence downstream of the proposed polyadenylation site did not detect any OTR gene transcripts (lane 3). These data agree with the RT-PCR results described above. In addition, smaller bands were observed, most notably at 2.9 kb (lane 1). Since multiple potential polyadenylation signals are present in the very long 3' UTR (Fig. 2), a likely possibility was that the transcripts differed with respect to the lengths of their 3' UTR. When blots were hybridized with a probe complementary to nt 5302–5795, the larger bands, but not the 2.9-kb band, were detected (Fig. 5, lane 2). These results suggest that the length heterogeneity of the OTR transcripts arises through alternative use of different polyadenylation sites, although other, less likely explanations, such as degradation or specific cleavage of the mRNA, cannot be ruled out. The 3' UTR contains six AUUUA motifs. These A+U-rich elements (AREs) impart mRNA instability and occur typically in unstable transcripts (27). Since OTR mRNAs of different length would contain different numbers of AREs, differential selection of the polyadenylation site could represent a means of modulating transcript stability.

Our analysis of the rat OTR gene provides a framework for the elucidation of the molecular mechanisms underlying OTR gene expression and, thus, of the unresolved question of the

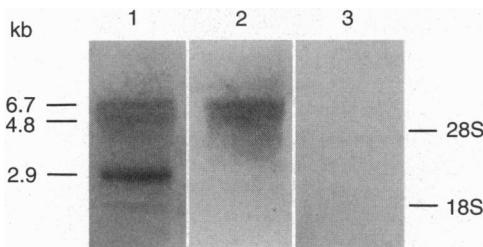


FIG. 5. Northern blot analysis of OTR mRNA. Each lane contained 50 µg of poly(A)⁺ RNA extracted from a parturient rat uterus. The probes were generated by PCR and corresponded to the following sequence segments (numbering according to Fig. 2): Lane 1, 4642–5136 (beginning of 3' UTR); lane 2, 5302–5795 (middle of 3' UTR); lane 3, 9101–9351 (past the last polyadenylation site). The migration positions of 28S and 18S rRNA are indicated at right. Size estimates based on RNA size markers are indicated at left.

mechanisms triggering the onset of parturition under physiological and pathophysiological conditions.

We thank Drs. R. Rehfuss and B. Turcotte for advice, helpful discussions, and critical reading of the manuscript. We thank Ms. R. Goltzman for valuable help and Ms. L. Chin for secretarial assistance. The work was supported by a grant from the Medical Research Council, Canada. F.R. holds a Royal Victoria Hospital Research Institute Fellowship and H.H.Z. holds a Scientist Award from the Medical Research Council, Canada.

1. Fuchs, A.-R. (1985) in *Oxytocin: Clinical and Laboratory Studies*, eds. Amico, J. & Robinson, A. G. (Elsevier, New York), pp. 207–235.
2. McCarthy, M. M., Kow, L. M. & Pfaff, D. W. (1992) *Ann. N.Y. Acad. Sci.* **652**, 70–82.
3. McCarthy, M. M., Kleopoulos, S. P., Mobbs, C. V. & Pfaff, D. W. (1994) *Neuroendocrinology* **59**, 432–440.
4. Schuhmacher, M., Coirini, H., Johnson, A. E., Flanagan, L. M., Frankfurt, M., Pfaff, D. W. & McEwen, B. (1993) *Regul. Pept.* **45**, 115–119.
5. Soloff, M. S., Alexandrova, M. & Fernstrom, M. J. (1979) *Science* **204**, 1313–1315.
6. Kimura, T., Tanizawa, O., Mori, K., Brownstein, M. J. & Okyama, H. (1992) *Nature (London)* **356**, 526–529.
7. Birnbaumer, M., Seibold, A., Gilbert, S., Ishido, M., Barberis, C., Antaramian, A., Brabet, P. & Rosenthal, W. (1992) *Nature (London)* **357**, 333–335.
8. Lolait, S. J., O'Carroll, A.-M., McBride, O. W., Konig, M., Morel, A. & Brownstein, M. J. (1992) *Nature (London)* **357**, 336–339.
9. Gorbulev, V., Buchner, H., Alkhundova, A. & Fahrenholz, F. (1993) *Eur. J. Biochem.* **215**, 1–7.
10. Morel, A., O'Carroll, A.-M., Brownstein, M. J. & Lolait, S. J. (1992) *Nature (London)* **356**, 523–526.
11. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K., eds. (1994) *Current Protocols in Molecular Biology* (Greene and Wiley, New York).
12. Frohman, M. A., Dush, M. K. & Martin, G. R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8998–9002.
13. Zingg, H. H., Goodman, R. H. & Habener, J. F. (1984) *Endocrinology* **115**, 90–94.
14. Seibold, A., Brabet, W., Rosenthal, W. & Birnbaumer, M. (1992) *Am. J. Hum. Genet.* **51**, 1078–1083.
15. Mount, S. M. (1982) *Nucleic Acids Res.* **10**, 459–472.
16. Abraham, L. J., Bradshaw, A. D., Shiels, B. R., Northemann, W., Hudson, G. & Fey, G. H. (1990) *Mol. Cell. Biol.* **10**, 3483–3491.
17. Akira, S. & Kishimoto, T. (1992) *Immunol. Rev.* **127**, 25–50.
18. Akira, S., Nishio, Y., Inoue, M., Wang, X.-J., Wei, S., Matsusaka, T., Yoshida, K., Sudo, T., Naruto, M. & Kishimoto, T. (1994) *Cell* **77**, 63–71.
19. Hunt, J. S. (1994) *Biol. Reprod.* **50**, 461–466.
20. Vince, G. S., Starkey, P. M., Jackson, M. C., Sargent, I. L. & Redman, C. W. G. (1990) *J. Immunol. Methods* **132**, 181–189.
21. Tabibzadeh, S. (1991) *Endocr. Rev.* **12**, 272–290.
22. Dudley, D. J., Trautman, M. S., Araneo, B. A., Edwin, S. S. & Mitchell, M. D. (1992) *J. Clin. Endocrinol. Metab.* **74**, 884–889.
23. Romero, R., Sepulveda, W., Mazor, M., Brandt, F., Cotton, D. B., Dinarello, C. A. & Mitchell, M. D. (1992) *Am. J. Obstet. Gynecol.* **167**, 863–872.
24. Romero, R. & Tartakovsky, B. (1992) *Am. J. Obstet. Gynecol.* **167**, 1041–1045.
25. Kato, S., Tora, L., Yamauchi, J., Masushige, S., Bellard, M. & Chambon, P. (1992) *Cell* **68**, 731–742.
26. Birnstiel, M. L., Busslinger, M. & Struhl, K. (1985) *Cell* **41**, 349–359.
27. Sachs, A. B. (1993) *Cell* **74**, 413–421.