Biochemistry

 Volume 31, Number 36 September 15, 1992

Articles

Analysis of the Promoter Region of the Rat D_2 Dopamine Receptor Gene

Takashi Minowa, Mari T. Minowa, and M. Maral Mouradian*

Experimental Therapeutics Branch, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20892

Received May 1, 1992; Revised Manuscript Received June 25, 1992

ABSTRACT: To investigate the regulatory processes involved in the expression of the D_2 dopamine receptor gene, a rat genomic clone was isolated using a 21-mer oligonucleotide probe made of exon 1 sequences. A 1.3-kb region including all of exon 1, its 5′-flanking region, and part of intron 1 was sequenced. S1 nuclease analysis indicated three consecutive nucleotides as the main transcription start sites; several weaker sites were also noted between 321 and 363 nucleotides upstream from the 3′ end of exon 1. The promoter region lacks TATA and CAAT boxes and is rich in G+C content with several putative Sp1 binding sites. Transient expression assays using chimeric constructs of D_2 promoter deletion mutants–chloramphenicol acetyltransferase gene in the neuroblastoma cell line NB41A3 which expresses D_2 binding sites indicated strong transcription enhancing activity between nucleotides −75 and −30 and silencing activity between nucleotides −217 and −76. DNase I footprinting studies using nuclear extract from NB41A3 suggested Sp1 binding to its consensus sequence at nucleotide −48 but inhibition of Sp1 binding at nucleotide −86 by the extract. The D_2 promoter could not induce transcription of the heterologous CAT gene in C6 glioma, embryonal NIH 3T3, or hepatic Hep G2 cells. It is concluded that the rat D_2 gene shares with the human D_{1A} dopamine receptor gene several features typical of "housekeeping" genes but they are both tissue-specific, regulated genes. Unlike the D_{1A} gene, however, the D_2 gene has a strong preference for transcription initiation to three consecutive nucleotides. The homology of the D_2 sequence between nucleotides −6 and +11 to the "initiator" sequence might be of functional significance in determining this preference.

Central dopamine neurotransmission has a key role in several brain processes including the generation of coordinated motor function, hypothalamic–pituitary axis regulation, and probably affect and cognition (Creese & Fraser, 1987). Dopamine exerts its biochemical and cellular effects by interacting with its cell surface receptors that belong to a large superfamily of receptors coupled to guanine nucleotide binding proteins (Freissmuth et al., 1989). Prior to the advent of molecular cloning techniques, dopamine receptors were classified into two categories, D_1 and D_2. To date, many more different dopamine receptor cDNAs and some of their cognate genes have been cloned (Bunzow et al., 1988; Chio et al., 1990; Dal Toso et al., 1989; Dearry et al., 1990; Gandelman et al., 1991; Giros et al., 1989; Grandy et al., 1989; Mack et al., 1991; Miller et al., 1990; Monsma et al., 1989b, 1990; O'Malley et al., 1990; Rao et al., 1990; Selbie et al., 1989; Sokoloff et al., 1990; Stormann et al., 1989; Sunahara et al., 1990; 1991; Tiberi et al., 1991; Van Tol et al., 1991; Weinshank et al., 1991; Zhou et al., 1990). Despite this progress, little is known about the molecular events that regulate the transcription of these genes with the exception of that coding for the D_{1A} receptor. We recently reported on the promoter structure of the human D_{1A} gene and found features typical of "housekeeping" genes, such as lack of TATA and CAAT boxes, rich G+C content, and multiple transcription initiation sites, even though this promoter induces transcription in a cell-type-specific manner (Minowa et al., 1992). Elucidation of factors involved in the transcriptional regulation of genes coding for dopamine receptors will assist our understanding of how the expression of these receptors could be altered by various pathologic and pharmacologic states.

The D_2 dopamine receptor has been traditionally recognized to be the primary mediator of the motor, endocrine, and

* Address correspondence to this author at Bldg. 10, Room 5C116, National Institutes of Health, Bethesda, MD 20892 [phone (301) 496-7872; fax (301) 496-6609].

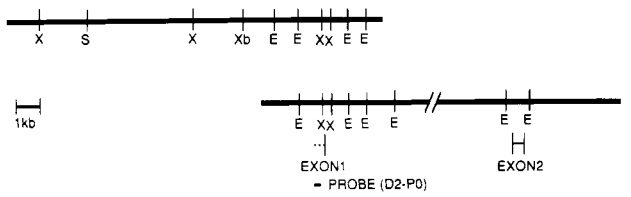


FIGURE 1: Restriction map of the rat D_2 gene. Upper line is clone λD2G7 isolated in the present study; lower line represents the rat D_2 genomic clone reported previously (O'Malley et al., 1990). Exons 1 and 2 are mapped but the 5′ boundary of the first exon had not been determined previously. The 21-mer oligonucleotide D2-P0 made of exon 1 sequence was used as probe to screen the genomic library. Restriction sites: E, *Eco*RI; S, *Sac*I; X, *Xho*I; Xb, *Xba*I.

behavioral effects of dopaminergic transmission (Schachter et al., 1980; Seeman, 1981). The rat gene coding for the D_2 receptor spans at least 50 kilobases (kb),[1] is interrupted by several introns, and can generate two alternatively spliced mature messages that differ by 87 nucleotides (Chio et al., 1990; Dal Toso et al., 1989; Giros et al., 1989; Miller et al., 1990; Monsma et al., 1989b; O'Malley et al., 1990; Rao et al., 1990). However, no experimental evidence has been presented thus far addressing the transcription start site or promoter activity of this gene. Furthermore, available data do not exclude the possibility of another exon located upstream to the previously described first exon. To analyze the structure and function of the promoter of the rat D_2 receptor gene, we began by cloning exon 1 and defining its 5′ end; we then determined the transcription start sites and studied the ability of deletion mutants of the putative promoter to transcribe a heterologous gene in various cell types.

EXPERIMENTAL PROCEDURES

Screening of Genomic Library and DNA Sequencing. A rat genomic library constructed in the λ phage vector EMBL3/SP6/T7 (Clontech) was screened with a 21-mer oligonucleotide probe, D2-P0 (5′-GGAGGGGCGGCCGTGCGTGGA-3′) (Figures 1 and 2),the sequence of which is part of the previously described first exon of the rat D_2 dopamine receptor gene (Bunzow et al., 1988; O'Malley et al., 1990), extending between nucleotides 234 and 254 in Figure 2. Screening was done under medium-stringency hybridization conditions; filters were washed with 0.1 × SSC (1 × SSC is 0.15 M NaCl and 0.015 M sodium citrate)/0.1% SDS at 42 °C. Among the seven positive clones identified, one (λD2G7) had a 2-kb *Eco*RI fragment which was found to contain sequences from the first exon of the rat D_2 gene on the basis of restriction analysis and Southern blots both with probe D2-P0 used for library screening and with D2-P1 (5′-CCGGGGCCAGAGAAG-GCG-3′, complementary to bases 287–304) (Figure 2). This *Eco*RI fragment was subcloned in pUC19 (Bethesda Research Laboratories) and in pGEM-3Zf(−) (Promega), giving plasmids pG7E and pG7E2, respectively. Both strands of a 1.3-kb region of the inserts in pG7E and pG7E2 were sequenced with M13 primers (United States Biochemical and New England Biolabs) and with synthesized primers on the basis of previous sequence information. Sequencing was done by the dideoxynucleotide chain-termination method (Sanger et al., 1977) with Sequenase (United States Biochemical) using 7-deaza-dGTP.

S1 Nuclease Analysis. S1 nuclease protection assays were carried out according to standard procedures (Sambrook et al., 1989) with slight modification. Single-stranded DNA of pG7E2 was used as template in primer extension reaction with a ^{32}P-5′-end labeled 30-mer oligonucleotide primer D2-P2 (5′-AGGACGAGGCGGGGCGGGAGGACGGGAGCG-3′, complementary to nucleotides 69–98) (Figure 2) followed by *Eco*RI digestion. The resulting probe was purified on 8.3 M urea/4% polyacrylamide gel, boiled for 10 min with 50 μg of total RNA or 2 μg of poly $(A)^+$ RNA from rat striatum, slowly cooled, and annealed overnight at 30 °C. Following hybridization, the nucleotides were digested with 300 units of S1 nuclease (United States Biochemical) at 30 °C for 60 min. The protected fragments were electrophoresed in 6% polyacrylamide gel containing 8.3 M urea parallel to DNA sequencing reaction using pG7E2 as template and the same oligonucleotide primer, D2-P2, used for preparing the S1 mapping probe.

Analysis of 5′-End Sequences of D_2 Receptor mRNA Using Reverse Transcription–Polymerase Chain Reaction (RT–PCR). Rapid amplification of cDNA 5′ ends (5′-RACE) was performed using a modification of a previously described method (Frohman et al., 1988). Briefly, first strand cDNA was synthesized using 3 μg of rat striatum poly $(A)^+$ RNA and 10 pmol of the D_2 receptor gene-specific primer, D2-P2, with 10 units of AMV reverse transcriptase (Invitrogen). After removal of primer by Centricon 100 (Amicon), poly[d(A)] tail was added to the 3′ ends of cDNAs by terminal deoxynucleotidyltransferase (United States Biochemical). The cDNAs was subjected to PCR amplification using the D_2-specific primer D2-P3 (5′-CGGGGACCAGGCGGGAA-3′, complementary to bases 47–63) (60 pmol) as the 3′ primer and an 18-mer adaptor (5′-GCCGCATGCGAATTCACC-3′) (60 pmol) and adaptor with poly[$d(T)_{17}$] tail (6 pmol) both as 5′ primers. PCR products were electrophoresed in 2% agarose gel containing formaldehyde followed by Southern hybridization with the oligonucleotide probe D2-PR (5′-AGTGCCGGAGCTGGTCGC-3′, identical to nucleotides 1–18) (Figure 2) under high-stringency conditions. Subsequently, to clone the D_2-specific product, a partial cDNA library was constructed from the PCR products in pUC19 and screened with D2-PR, again under vigorous washing conditions. Twelve positive clones of 200 screened were isolated. Recombinant plasmids were analyzed by restriction enzymes *Eco*RI and *Hin*dIII and by Southern hybridization with D2-PR. Seven of these 12 clones were further analyzed by DNA sequencing primed with D2-P3 or M13 universal primer.

Plasmid Constructions. A 1.1-kb *Eco*RI–*Xho*I fragment in pG7E was excised using *Xho*I and the *Pst*I site of the vector and subcloned between the *Pst*I and *Sal*I sites of pCAT-Basic plasmid (Promega) upstream of the translation initiation site of the chloramphenicol acetyltransferase (CAT) gene, giving rise to pCATD2-852. Four additional 5′-deletion mutants were constructed from pCATD2-852 using two *Acc*III sites, a *Sma*I site, and an *Eco*47III site, yielding constructs pCATD2-394, -217, -75, and -29, respectively (Figure 4). Length of the inserts and the 5′–3′ orientation relative to the CAT gene were verified by restriction analysis and sequencing. pCAT-Basic plasmid, which lacks a promoter and an enhancer, was used as a negative control, and pCAT-Control (Promega), which has both the SV40 promoter and enhancer, was used as a positive control for CAT assays.

Cell Culture and Transient Expression Assays. The murine neuroblastoma cell line NB41A3, the rat glioma C6, the mouse embryonal NIH 3T3, and the human hepatoblastoma Hep G2 were all cultured in Dulbecco's modified Eagle's medium

[1] Abbreviations: bp, base pair(s); kb, kilobase; PCR, polymerase chain reaction; RACE, rapid amplification of cDNA ends; CAT, chloramphenicol acetyltransferase.

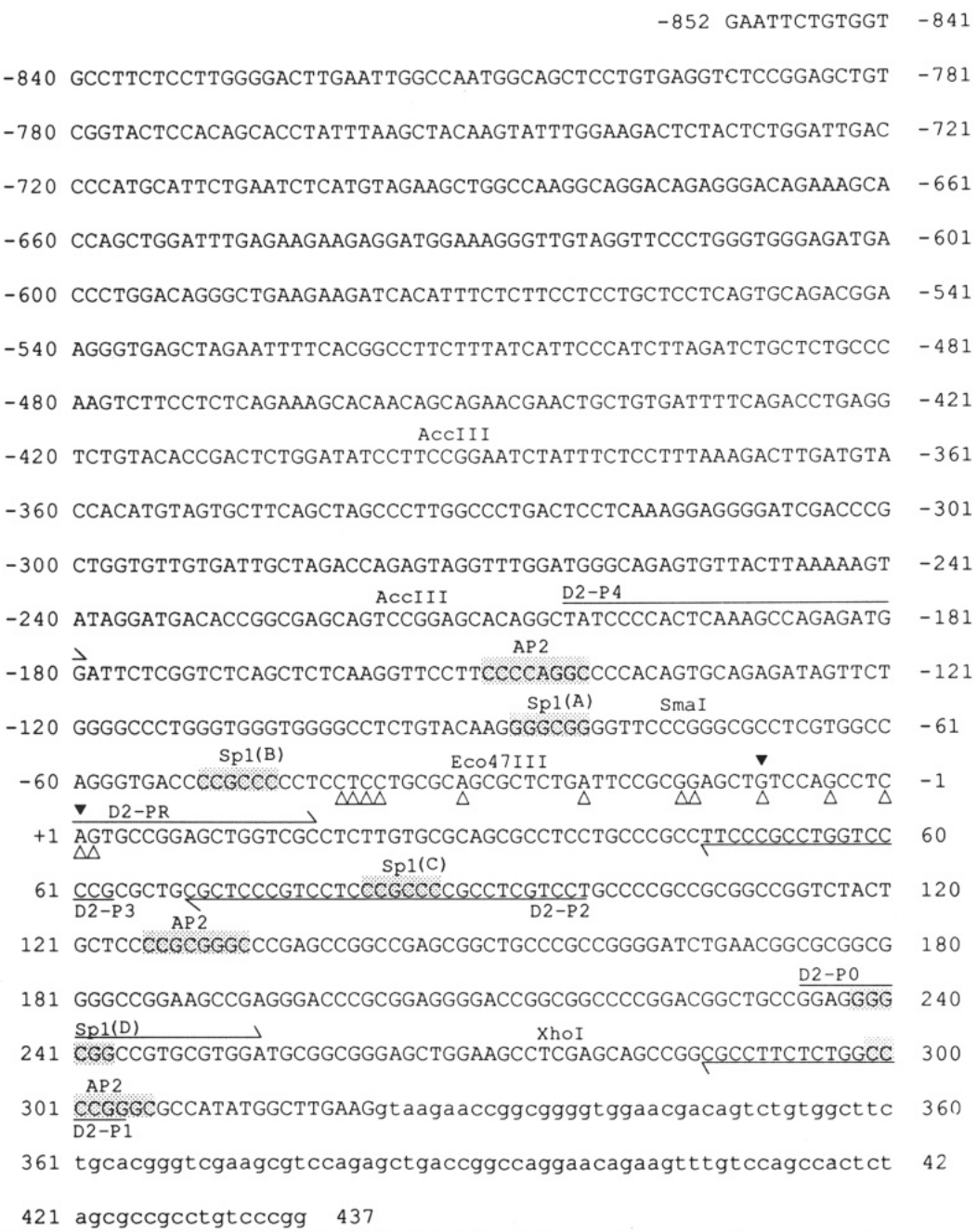


FIGURE 2: Nucleotide sequence of the 1.3-kb *Eco*RI–*Xho*I region of the rat D_2 gene. The segment in lower case characters downstream from base 322 indicates intron 1. Open arrowheads underneath the sequence indicate transcription initiation sites determined by S1 nuclease analysis. Solid arrowheads above the sequence refer to transcription initiation sites determined by 5′-RACE. Oligonucleotides and their orientation used in library screening and in RT-PCR are shown as horizontal arrows above and below the sequence. Restriction sites used for making deletion mutants for CAT constructs are indicated above the sequence. Consensus sequences for Sp1 (designated A–D) and for AP2 binding sites are shaded.

(Mediatech) supplemented with 10% fetal bovine serum at 37 °C in a humidified atmosphere containing 10% CO_2. All cell lines were obtained from ATCC. Transfections were carried out using the $CaPO_4$ coprecipitation method (Graham & Van Der Eb, 1973) with 10 μg of the test pCAT plasmid and 10 μg of the β-galactosidase expression vector pRAS β-GAL (Umesono & Evans, 1989) per 100-mm culture plate. All plasmids used in transfections were purified by two CsCl centrifugations. Following transfection, cell lysate [normalized according to β-galactosidase activity (Edlund et al., 1985)] was incubated with 0.15 μCi of [^{14}C]chloramphenicol (Amersham), 0.5 mM acetyl-CoA, and 0.25 M Tris-HCl, pH 7.8, in a final volume of 160 μL at 37 °C for 30 min. Incubation mix was extracted with ethyl acetate and analyzed by thin-layer chromatography. Radioactivity was quantitated by liquid scintillation counter.

S1 Mapping with RNA from Transfected Cells. Poly (A)$^+$ RNA from NB41A3 cells was isolated both from nontransfected cells and following transfection with pCATD2-75. S1 mapping probe was prepared by PCR with ^{32}P-5′-end labeled D2-P2 as the 3′ primer and unlabeled D2-P4 (5′-TATCCCCACTCAAAGCCAGAGATGG-3′, identical to nucleotides −204 to −180) as the 5′ primer using pG7E as template and purified by denaturing gel electrophoresis. S1 mapping was performed using 12 μg of poly (A)$^+$ RNA as described above.

DNase I Footprinting. Nuclear extract from NB41A3 cells was prepared using a modification of Dignam's procedure (Dignam et al., 1983; Henninghausen & Lubon, 1987). Purified human Sp1 protein was purchased from Promega

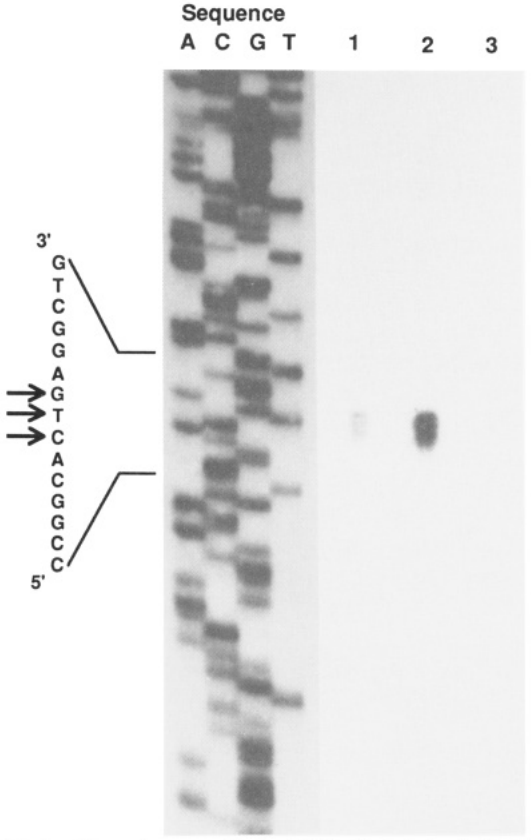


FIGURE 3: S1 nuclease analysis. 5′-End-labeled, single-stranded genomic DNA probe (complementary to nucleotide −852 to 98) was hybridized with 2 μg of rat striatum poly (A)$^+$ RNA (lane 1), 50 μg of rat striatum total RNA (lane 2), and 50 μg of yeast tRNA (lane 3) and digested with S1 nuclease. The protected fragments were electrophoresed in 8.3 M urea/6% polyacrylamide gel. The left lanes show the DNA sequence of the minus strand determined with the same oligonucleotide primer (D2-P2, complementary to nucleotides 69–98) used for preparing the S1 mapping probe. Arrows refer to the nucleotides corresponding to the strong signals. Several weaker signals of slower gel mobility were identified on the autoradiogram; their locations are indicated on the nucleotide sequence in Figure 2.

(La Jolla, CA). DNA probe for footprinting was prepared by PCR using unlabeled 5′ primer D2-N1 (5′-GAGATG-GTTCTGGGGCCCTGGGT-3′, identical to nucleotides −131 to −109) and ^{32}P-5′-end labeled 3′ primer D2-N2 (5′-GGCGACCAGCTCCGGCACTGA-3′, complimentary to nucleotides −2 to +19) with pG7E as template. The generated probe was purified by polyacrylamide gel electrophoresis. DNase I footprinting was performed according to an established method (Henninghausen & Lubon, 1987).

RESULTS

Cloning and Nucleotide Sequence of Exon 1 and the 5′-Upstream Region of the Rat D_2 Dopamine Receptor Gene. A rat D_2 genomic clone (λD2G7) was isolated by screening a library with a 21-mer oligonucleotide probe (D2-P0) based on the sequence of the previously described first exon. A partial restriction map of the insert in λD2G7 is shown in Figure 1 and compared with the map of a rat D_2 genomic clone reported previously (O'Malley et al., 1990). The restriction map of the 2-kb *Eco*RI fragment is identical between the two clones. A 1.3-kb *Eco*RI–*Xho*I region of this fragment, which was subsequently determined to include the entire first exon, was completely sequenced (Figure 2). The exon 1 sequence is virtually identical to those of rat D_2 receptor cDNAs (Bunzow et al., 1988; Miller et al., 1990; Monsma et al., 1989b; Rao et al., 1990). We conclude that λD2G7 has part of the rat D_2 receptor gene.

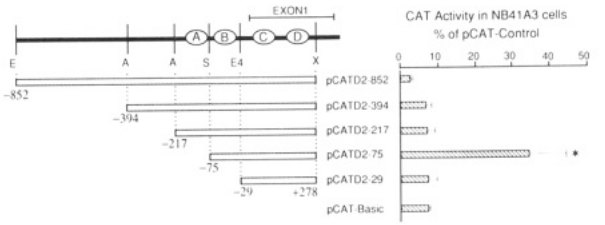


FIGURE 4: Deletion mutants of the rat D_2 gene 5′-flanking region used for transient expression assays and their transcriptional activity in NB41A3 cells. The black bar on top shows restriction map of the D_2 upstream region. Ellipses A–D represent the putative Sp1 binding sites (see Figure 2). Restriction sites: A (*Acc*III), E (*Eco*RI), E4 (*Eco*47III), S (*Sma*I), X (*Xho*I). Shaded bars represent the fragments subcloned upstream of the CAT gene in pCAT-Basic in the 5′–3′ orientation. Nucleotide numbers are shown under the bars. The constructs generated are designated pCAT-D2-852, -394, -217, -75, and -29. Hatched bars on the right show promoter activity of each pCAT construct in NB41A3 cells relative to the positive control plasmid, pCAT-Control. Promoter activity of pCAT-Basic is also shown. Transfection efficiencies were normalized according to β-galactosidase activity. The data shown are means ± SEM of three independent experiments. (Asterisk) ANOVA $p < 0.002$ for difference from all other clones.

Table I: Stimulation of CAT Activity in Different Cell Lines[a]

clone	C6	NIH 3T3	Hep G2
pCAT-D2-852	2.7 ± 0.2	1.6 ± 0.8	2.7 ± 0.6
pCAT-D2-394	4.1 ± 0.6	2.3 ± 0.8	5.2 ± 1.1
pCAT-D2-217	2.3 ± 0.3	2.5 ± 1.4	6.0 ± 1.4
pCAT-D2-75	1.3 ± 0.4	4.3 ± 1.3	4.9 ± 0.3
pCAT-D2-29	0.5 ± 0.0	0.9 ± 0.1	1.4 ± 0.1
pCAT-Basic	1.0 ± 0.1	1.5 ± 0.5	4.1 ± 0.7

[a] Deletion mutants of the D_2 promoter and the negative control plasmid, pCAT-Basic, are indicated on the left. Acetylation results are expressed as percent of the positive control plasmid, pCAT-Control, in each cell line. Transfection efficiencies were normalized according to β-galactosidase activity. Origins of cell lines are as follows: C6, rat glioma; NIH 3T3, mouse embryonal; Hep G2, human hepatoblastoma. The data shown are means ± SEM of two independent experiments.

Comparison between our genomic clone and a rat D_2 cDNA which has the longest 5′ stretch reported (Miller et al., 1990) indicated an abrupt diversion in the sequence identity in the 36 nucleotides at the 5′ end of that cDNA (upstream of nucleotide −2 in Figure 2). To test if this additional sequence represented another upstream exon in the D_2 gene, we performed reverse transcription coupled with PCR amplification of D_2 message from rat striatum poly (A)$^+$ RNA. While a specific product was obtained with the 5′ primer, D2-PR, which has a sequence common between our genomic clone and the reported cDNA clone, no such product could be amplified with a 5′ primer which had the unidentified sequence in the cDNA (5′-CGGAATCGAGAAAGAGCTATC-3′) (data not shown). Therefore, the 36 bases in the 5′ end of that cDNA clone may be due to recombination during cDNA library construction and are not sequences from another exon of the D_2 gene located further upstream.

The 5′ Ends of Exon 1 Are the Transcription Start Sites of the Rat D_2 Gene. The 5′ end of exon 1 was determined by S1 nuclease analysis. S1 mapping with a 5′-end-labeled single-stranded DNA probe and rat striatal RNA showed three strong signals and several weak ones with slower gel mobility (Figure 3). The positions of all of these signals are indicated in the nucleotide sequence (Figure 2). The finding of several signals rather than a single S1 signal suggests that these nucleotides represent transcription start sites and not splice acceptor site. Furthermore, the nucleotide sequence around these signals is

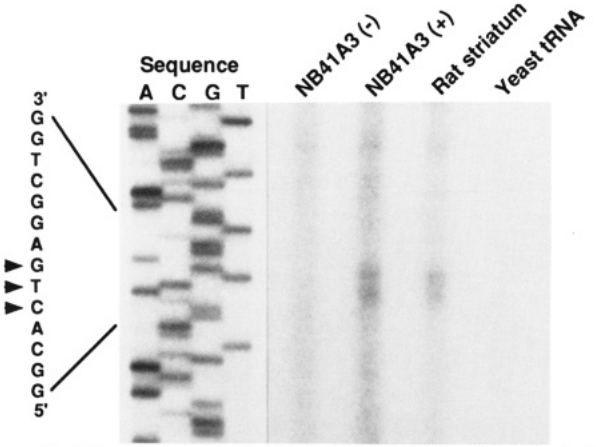


FIGURE 5: S1 nuclease analysis using RNA from transfected NB41A3 cells. A 5′-end-labeled S1 mapping probe (complementary to nucleotide −204 to 98) was generated by PCR with D2-P2 as the labeled 3′ primer and D2-P4 as the unlabeled 5′ primer using pG7E as template. Poly $(A)^+$ RNA was isolated from NB41A3 (12 μg) cells both without transfection (−) and following transfection with pCATD2-75 (+). Parallel experiments were done with rat striatum total RNA (50 μg) and yeast tRNA (50 μg). S1 analysis was done as described under Experimental Procedures and in the legend to Figure 3. The left lanes show the DNA sequence of the minus strand determined with primer D2-P2. Arrowheads show the positions of the signals seen both with transfected cells and with rat striatum.

generally G+C-rich and contains several putative Sp1 binding sites which are features of TATA-less promoters.

The possibility that the 5′ ends of exon 1 are indeed the transcription start sites was further confirmed by sequencing the 5′ ends of D_2 cDNAs using rapid amplification of cDNA ends (5′-RACE). The results of the RACE procedure were in general agreement with those of S1 mapping. Gel electrophoresis of RACE products revealed a smear, but Southern hybridization with the D_2-specific probe, D2-PR, showed a single band of approximately 100 bases (data not shown). The RACE products were then ligated with linearized pUC19, and 200 clones were screened with D2-PR. Twelve positive clones were obtained, and all contained about 100 base pair (bp) inserts which hybridized with D2-PR. Among the seven positive clones sequenced, the 5′ ends of two were the same adenine that corresponded to one of the strong S1 mapping signals, and the 5′ ends of another five clones matched with one of the weak S1 signals (Figure 2). These results indicate that there is no exon further upstream to exon 1 in the rat D_2 gene and that this gene has multiple transcription initiation sites located between 321 and 363 nucleotides upstream from the 3′ end of the first exon. We have designated the adenine that corresponds to one of the strong S1 signals and is also one of the 5′ cDNA ends generated by RACE as +1 (Figure 2).

Transcriptional Activity of the Putative Promoter of the Rat D_2 Gene in Transiently Transfected NB41A3 Cells. To determine the promoter region of the rat D_2 gene, several 5′-deletion mutants made of overlapping restriction fragments of the 5′-flanking region were subcloned in the promoter-less vector pCAT-Basic (Figure 4). In the mouse neuroblastoma cell line, NB41A3 which expresses binding sites for D_2 ligands (Monsma et al., 1989a), strong transcriptional activity was observed with pCATD2-75 but not with the other constructs (Figure 4; Table I). To identify the transcription start site of this recombinant CAT gene in transient expression assays, S1 nuclease analysis was performed using RNA from transfected NB41A3. The results indicated that transcription of the recombinant D_2 promoter-CAT gene in this cell line is

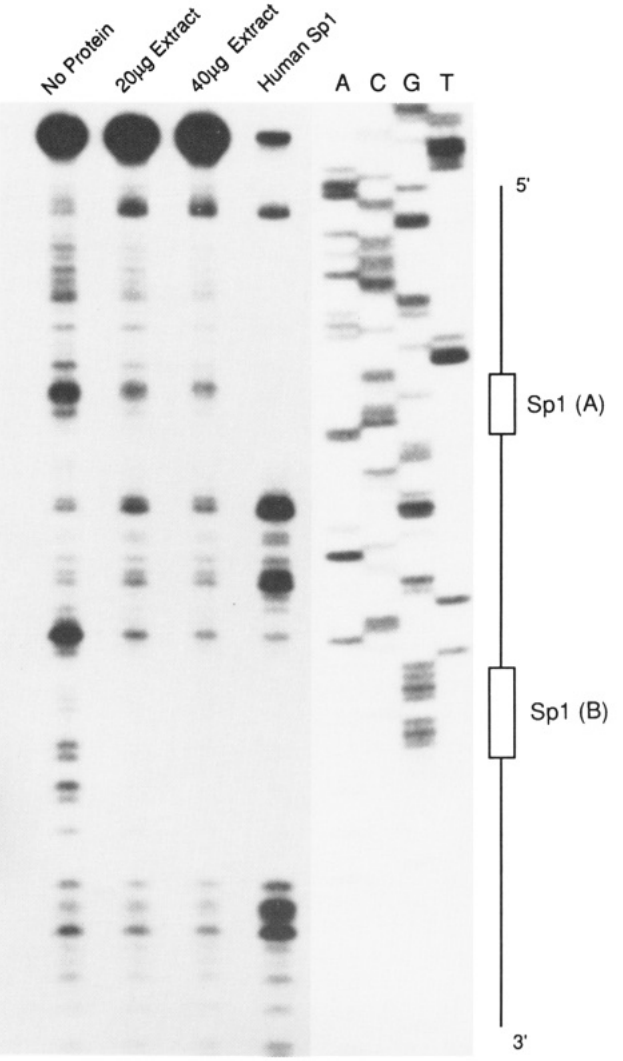
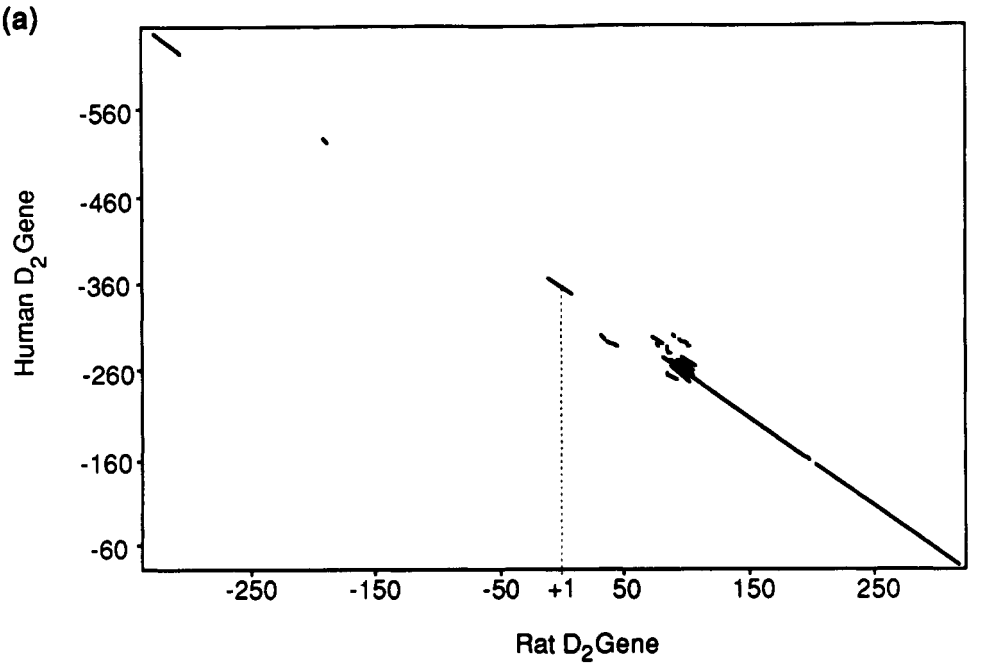


FIGURE 6: DNase I footprinting of the rat D_2 promoter region. The probe is 150 bp extending between nucleotides −131 and +19, labeled at the 5′ end of the noncoding strand at +19. The DNase I cleavage pattern is shown in the absence of protein, in the presence of 20 or 40 μg of nuclear extract from NB41A3 cells, and in the presence of 2 footprinting units of purified human Sp1. The sequence ladder shown on the right is generated with the same labeled primer used to prepare the footprinting probe. The boxes represent the positions of the putative Sp1 binding sites A and B (see Figure 4). Parallel experiments were performed with the DNA probe labeled on the coding strand, yielding consistent results.

initiated at the same points as is the D_2 gene in vivo in the rat striatum (Figure 5). These results strongly suggest that the 353-bp *Sma*I-*Xho*I fragment of the rat D_2 gene contains the promoter of this gene.

DNase I Footprinting Analysis of the 5′-Upstream Region of the Rat D_2 Gene. The results of the CAT assays suggest that a strong positive modulator of the D_2 promoter is localized between nucleotides −75 and −30. To identify the cis-acting elements and trans-acting factors interacting with this region, we performed DNase I footprinting using nuclear extract from NB41A3 cells and a DNA probe extending from nucleotides −131 to +19 which contains two putative Sp1 binding sites (sites A and B; see Figure 4). While purified human Sp1 protein strongly bound to both Sp1 consensus sequences, the nuclear extract showed strong binding only to site B, which is located between nucleotides −75 and −30 (Figure 6).

Cell-Type-Specific Transcriptional Activity of the Rat D_2 Promoter. To determine cell-type specificity of the rat D_2 promoter, the rat glioma C6, mouse embryonal NIH 3T3,



(b)

```
RAT    GCCCCACAGTGCAGAGATAGTTCTGGG-GCCCTGGGTGGGTGGGGCCTCTGTACAAGGGG  -86
       |||| |    | |    ||  | |||  | | | |  |||   ||  || |    ||||
HUMAN  ACCCCTCCTACCCGTTCCAGGCCGGGGATCGCCGAGGAGGTACAGCTCCTTTGGTGGGGG -443

RAT    CGGGGTTCCCGGGC-GCCTCGTGGCCAGGGTGACCCCGCCCCCT-CCTCCTGCG--CAGC  -30
       |||  |  || | | |||  || |  || ||||  | | ||| |||||  |||  |  |
HUMAN  GCGGGGGCGGGGCCTGTCTCAGGGGC--GGGGACCGGGGCACCTCCCTCCCGCGCTCCCC -385

                                      +1
RAT    GCTCT-GATTCCGCGGAGCTGTCCAGCCTCAGTGCCGGAGCTGGTCGCCTCTTGTGCGCA  +30
       || || |    ||| ||||||||||||| |||||||| || | ||  |||||  | |||||
HUMAN  GCGCTCGGGCGCGCAGAGCTGTCCAGCTTCAGTGCC-GAACCGGCAGCCTCACGCGCGCA -326

RAT    --GCGCCTCCTGC-CCGCCTTCCCGC-CTGGTCCCCGCGCTGCGCTCCCGTCCTCCCGCC  +86
         ||||| ||| | || | | ||||| ||    || || | || | |||  |  || |||
HUMAN  CCGCGCCGCCTCCGCCCCGTCCCCGCGCTCCCTCCTGC-CCGCCCGCCCCGCGCCCGGCC -267

RAT    CCGCCTCGTCCTGCCCCGCCGCGGCCGGTCTACTGCTCCCCGCGGGCCCGAGCCGGCCGA +146
       ||||| || || |||||||||||||| ||| ||||||||||||||||| ||||||||||
HUMAN  CCGCCCCGCCCCGCCCCGCCGCGGCCCGTCCACTGCTCCCCGCGGGCCAGAGCCGGCCGA -207
```

FIGURE 7: Sequence comparison of the 5′-upstream region between the rat and human D_2 dopamine receptor genes. (a) Homology analysis was done with the program MacVector (International Biotechnologies). Lines indicate more than 65% homology with a window size of 30 bases. (b) Sequence comparison between the rat and human genes. The numbering system for the rat gene is the same adopted in the present study as in Figure 2. The numbering system for the human gene is the same as previously published (Gandelman et al., 1991), whereby the adenine of the first ATG codon is +1 since the transcription start site for the human gene has not been defined. Putative Sp1 binding sites in both genes are underlined.

and human hepatoblastoma Hep G2 were transfected with all five CAT constructs (Figure 4; Table I). Essentially no transcriptional activity was found in these cell lines tested compared with the negative control plasmid, pCAT-Basic.

DISCUSSION

To analyze the promoter region of the gene encoding the D_2 dopamine receptor, we cloned a rat genomic fragment containing part of this gene. We sequenced a 1.3-kb *Eco*RI–*Xho*I region which includes all of exon 1, part of intron 1, and 852 nucleotides upstream of exon 1. The sequence of overlapping segments between our clone and previously reported rat D_2 cDNA or genomic clones is virtually identical (Bunzow et al., 1988; Miller et al., 1990; Monsma et al., 1989b; Rao et al., 1990).

Experiments performed to determine transcription initiation site using S1 nuclease analysis and sequencing the 5′ ends of D_2 cDNAs indicated that the previously described exon 1 is indeed the first exon of this gene. The present experiments also revealed that the rat D_2 gene has multiple transcription start sites located between 321 and 363 nucleotides upstream of the 3′ end of the first exon. This assignment of transcription initiation sites and the known 3′ extent of the rat D_2 cDNA are consistent with the mRNA size of 2.7–2.9 kb (Bunzow et al., 1988; Giros et al., 1989; Monsma et al., 1989b).

Functional analysis of the promoter region of the D_2 gene revealed that pCATD2-75 demonstrates significant transcriptional activity in the neuroblastoma cell line NB41A3, which expresses binding sites for D_2 ligands (Monsma et al., 1989a). Furthermore, transcription of this chimeric gene starts at the same points as does the D_2 gene in the rat brain, indicating that the authentic D_2 promoter drives CAT gene expression in the transfection experiments (Figure 5). Thus, we conclude that the 1.3-kb region we sequenced contains the promoter of the rat D_2 gene. The poor promoter activities of pCATD2-852, -394, and -217 might explain the low abundance of the D_2 message in NB41A3 cells. In fact, we could detect a D_2-specific mRNA signal in these cells by solution hybridization/ribonuclease protection assay but not by northern hybridization (data not shown).

Analysis of the sequence upstream of the transcription start sites indicated that the promoter region of the rat D_2 gene lacks a TATA box and a CAAT box. A CCAAT sequence is found at −814, which is not the expected position of a CAAT box relative to transcription initiation site. In addition, this promoter is highly rich in G+C content, reaching 80% in some portions, and has multiple putative Sp1 binding sites (Figure 2). These sequence features are typical for promoters of housekeeping genes (Dynan, 1986). Yet, the rat D_2 gene is regulated and has a highly tissue-specific pattern of expression. Transient transfections of NB41A3 cells with chimeras of D_2 upstream fragments-CAT gene indicated that the D_2 promoter is positively modulated by a cis-acting element located between nucleotides −75 and −30 and negatively modulated by elements located between nucleotides −217 and −76 and between nucleotides −852 and −394. Furthermore, DNase I footprinting experiments suggested that Sp1 in nuclear extract of NB41A3 cells binds to its cognate sequence at nucleotide −48 (site B) but not to the same core sequence at nucleotide −86 (site A). These observations strongly support the possibility that the Sp1 binding site B is crucial for transcription of the D_2 gene. Whether lack of binding at site A is due to an as yet unidentified factor in the extract inhibiting Sp1 binding and whether this interference relates to the silencer activity of the −217 to −76 portion of the gene require further investigation. In contrast to NB41A3, none of the CAT constructs showed evidence of transcription in either the C6 glioma cells, the NIH 3T3 embryonal cells, or the Hep G2 hepatoblastoma cells, indicating that the D_2 promoter is virtually silent in these cells.

The housekeeping gene-like features of the D_2 promoter are shared with several other tissue-specific genes including that encoding the human D_{1A} dopamine receptor (Sehgal et al., 1988; Tsukamoto et al., 1991; Sauerwald et al., 1990; Minowa et al., 1992). However, distinct differences are noted between the D_2 and D_{1A} promoters. While transcription of the D_{1A} gene starts apparently equally at multiple points, the D_2 gene has a clear preference to begin transcription at the three consecutive nucleotides between −1 and +2, as indicated by the strong bands protected by S1 nuclease (Figure 3). Furthermore, 8 of the 17 nucleotides between −6 and +11 of the D_2 gene are identical to the "initiator" sequence, identified in the murine terminal deoxynucleotidyltransferase gene (5′-GCCCTCATTCTGGAGAC-3′, initiation site is underlined). The initiator, for which no strong consensus sequence has thus far been identified, is functionally similar to TATA elements in influencing the direction of transcription as well as defining its site of initiation (O'Shea-Greenfield & Smale, 1992). This initiator-like sequence in the rat D_2 gene may be critical in starting transcription preferentially at the three nucleotides.

Sequence comparison of the D_2 gene first exon and its 5′-flanking region between human (Gandelman et al., 1991) and rat showed decreasing homology with increasing distance from the 3′ end of exon 1 (Figure 7). However, two highly homologous segments were identified in this generally nonhomologous region of the two genes. Interestingly, one of these segments is around the transcription start sites of the rat gene. In addition, several consensus DNA sequences for Sp1 are found at corresponding sites between the D_2 genes of both species. These observations may suggest that the human D_2 gene has transcription initiation site(s) and promoter structure similar to those of its rat counterpart.

We conclude that the rat D_2 gene, similar to the human D_{1A} gene (Minowa et al., 1992), is transcribed from a TATA-less promoter and is expressed in a tissue-specific manner. Unlike the D_{1A} gene, however, transcription of the D_2 gene is not initiated from all possible sites at equal efficiency. An initiator-like element in the D_2 gene might be the determinant of this feature. Whether the promoters of other dopamine or catecholamine receptor genes share these features remains to be determine]d.

REFERENCES

Bunzow, J. R., Van Tol, H. H. M., Grandy, D. K., Albert, P., Salon, J., Christie, M., Machida, C. A., Neve, K. A., & Civelli, O. (1988) *Nature 336*, 783–787.

Chio, C. L., Hess, G. F., Graham, R. S., & Huff, R. M. (1990) *Nature 343*, 266–269.

Creese, I., & Fraser, C. M. (1987) *Receptor Biochemistry and Methodology: Dopamine Receptors*, Vol. 8, Liss, New York.

Dal Toso, R., Sommer, B., Ewert, M., Herb, A., Pritchett, D. B., Bach, A., Shivers, B. D., & Seeburg, P. H. (1989) *EMBO J. 8*, 4025–4034.

Dearry, A., Gingrich, J. A., Falardeau, P., Fremeau, R. T., Jr., Bates, M. D., & Caron, M. G. (1990) *Nature 347*, 72–76.

Dignam, J. D., Leboviz, R. M., & Roeder, R. G. (1983) *Nucleic Acids Res. 11*, 1475–1489.

Dynan, W. S. (1986) *Trends Genet. 2*, 196–197.

Edlund, T., Walker, M. D., Barr, P. J., & Rutter, W. J. (1985) *Science 230*, 912–916.

Freissmuth, M., Casey, P. J., & Gilman, A. G. (1989) *FASEB J. 3*, 2125–2131.

Frohman, M. A., Dush, M. K., & Martin, G. R. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8998–9002.

Gandelman, K.-Y., Harmon, S., Todd, R. D., & O'Malley, K. L. (1991) *J. Neurochem. 56*, 1024–1029.

Giros, B., Sokoloff, P., Martres, M-P., Riou, J.-F., Emorine, L. J., & Schwartz, J.-C. (1989) *Nature 342*, 923–926.

Graham, F. L., & Van Der Eb, A. J. (1973) *Virology 52*, 456–467.

Grandy, D. K., Marchionni, M. A., Makam, H., Stofko, R. E., Alfano, M., Frothingham, L., Fischer, J. B., Burke-Howie, K. J., Bunzow, J. R., Server, A. C., & Civelli, O. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9762–9766.

Hennighausen, L., & Lubon, H. (1987) *Methods Enzymol. 152*, 721–735.

Mack, K. J., Todd, R. D., & O'Malley, K. L. (1991) *J. Neurochem. 57*, 795–801.

Miller, J. C., Wang, Y., & Filer, D. (1990) *Biochem. Biophys. Res. Commun. 166*, 109–112.

Minowa, M. T., Minowa, T., Monsma, F. J., Jr., Sibley, D. R., & Mouradian, M. M. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 3045–3049.

Monsma, F. J., Jr., Brassard, D. L., & Sibley, D. R. (1989a) *Brain Res. 492*, 314–324.

Monsma, F. J., Jr., McVittie, L. D., Gerfen, C. R., Mahan, L. C., & Sibley, D. R. (1989b) *Nature 342*, 926–929.

Monsma, F. J., Jr., Mahan, L. C., McVittie, L. D., Gerfen, C. R., & Sibley, D. R. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6723–6727.

O'Malley, K. L., Mack, K. J., Gandelman, K.-Y., & Todd, R. D. (1990) *Biochemistry 29*, 1367–1371.

O'Shea-Greenfield, A., & Smale, S. T. (1992) *J. Biol. Chem. 267*, 1391–1402.

Rao, D. D., McKelvy, J., Kebabian, J., & MacKenzie, R. G. (1990) *FEBS Lett. 263*, 18–22.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Sauerwald, A., Hoesche C., Oschwald, R., & Kilimann, M. W. (1990) *J. Biol. Chem. 265*, 14932–14937.

Schachter, M., Bedard, P., Debono, A. G., Jenner, P., Marsden, C. D., Price, P., Parkes, J. D., Keenan, J., Smith, B., Rosenthaler, J., Horowski, R., & Dorrow, R. (1980) *Nature 286*, 157–159.

Seeman, P. (1981) *Pharmacol. Rev. 32*, 229–313.

Sehgal, A., Patil, N., & Chao, M. (1988) *Mol. Cell. Biol. 8*, 3160–3167.

Selbie, L. A., Hayes, G., & Shine, J. (1989) *DNA 8*, 683–689.

Sokoloff, P., Giros, B., Martres, M.-P., Bouthenet, M.-L., & Schwartz, J.-C. (1990) *Nature 347*, 146–151.

Stormann, T. M., Gdula, D. C., Weiner, D. M., & Brann, M. R. (1989) *Mol. Pharmacol. 37*, 1–6.

Sunahara, R. K., Niznik, H. B., Weiner, D. M., Stormann, T. M., Brann, M. R., Kennedy, J. L., Gelernter, J. E., Rozmahel, R., Yang, Y., Israel, Y., Seeman, P., & O'Dowd, B. F. (1990) *Nature 347*, 80–83.

Sunahara, R. K., Guan, H.-C., O'Dowd, B. F., Seeman, P., Laurier, L. G., Ng, G., George, S. R., Torchia, J., Van Tol, H. H. M., & Niznik, H. B.(1991) *Nature 350*, 614–619.

Tiberi, M., Jarvie, K. R., Silvia, C., Falardeau, P., Gingrich, J. A., Godinot, N., Bertrand, L., Yang-Feng, T. L., Fremeau, R. T., Jr., & Caron, M. G. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 7491–7495.

Tsukamoto, T., Toyama, R., Itoh, H., Kozasa, T., Matsuoka, M., & Kaziro, Y. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 2974–2978.

Umesono, K., & Evans, R. M. (1989) *Cell 57*, 1139–1146.

Van Tol, H. H. M., Bunzow, J. R., Guan, H.-C., Sunahara, R. K., Seeman, P., Niznik, H. B., & Civelli, O. (1991) *Nature 350*, 610–614.

Weinshank, R. L., Adham, N., Macchi, M., Olsen, M. A., Branchek, T. A., & Hartig, P. R. (1991) *J. Biol. Chem. 266*, 22427–22435.

Zhou, Q.-Y., Grandy, D. K., Thambi, L., Kushner, J. A., Van Tol, H. H. M., Cone, R., Pribnow, D., Salon, J., Bunzow, J. R., & Civelli, O. (1990) *Nature 347*, 76–80.