Genomic structure and analysis of promoter sequence of a mouse μ opioid receptor gene

BON H. MIN*[†], LANCE B. AUGUSTIN*, RODERICK F. FELSHEIM*, JAMES A. FUCHS[‡], AND HORACE H. LOH*[§]

*Department of Pharmacology, University of Minnesota, Minneapolis, MN 55455; [‡]Department of Biochemistry, University of Minnesota, St. Paul, MN 55108; and [†]Department of Biochemistry, Kon-Kuk University, Choongju-city, Korea 380-701

Communicated by James F. Bonner, May 26, 1994

We have isolated mouse μ opioid receptor ABSTRACT genomic clones (termed MOR) containing the entire amino acid coding sequence corresponding to rat MOR-1 cDNA, including additional 5' flanking sequence. The mouse MOR gene is >53 kb long, and the coding sequence is divided by three introns, with exon junctions in codons 95 and 213 and between codons 386 and 387. The first intron is >26 kb, the second is 0.8 kb, and the third is >12 kb. Multiple transcription initiation sites were observed, with four major sites confirmed by 5' rapid amplification of cDNA ends and RNase protection located between 291 and 268 bp upstream of the translation start codon. Comparison of the 5' flanking sequence with a transcription factor database revealed putative cis-acting regulatory elements for transcription factors affected by cAMP, as well as those involved in the action of gluco- and mineralocorticoids, cytokines, and immune-cell-specific factors.

Opioid drugs, which are widely known clinically as painkillers and for their abuse potential, initiate their actions through cell surface receptors. In regions of the central nervous system, and in cell lines such as NG108-15 neuroblastoma \times glioma cells and SH-SY5Y cells, these opioid receptors are coupled to G proteins that mediate opioid actions such as inhibition of adenylyl cyclase and regulation of ion channels (1–7). The recent cloning of the δ opioid receptor (8, 9) from NG108-15 cells, followed by cloning of the μ opioid receptor (10–13) and the κ opioid receptor (14–17) from rodent brain, has confirmed that these receptors have a predicted amino acid sequence containing seven putative membrane-spanning regions that is typical of other G-protein-coupled receptors. The sequences of these cDNAs confirmed the existence of at least three types of opioid receptors (μ , δ , and κ) observed in pharmacological and biochemical studies over the last two decades (41). However, nothing is known about the genomic structure of the opioid receptor gene(s).

There is ample pharmacological evidence to suggest that μ opioid receptors in brain might be involved in tolerance and dependence as well as analgesia since μ receptors have high affinities for addictive and analgesic opioid drugs such as morphine (18). There is also evidence suggesting the existence of μ opioid receptors in other tissues (or cell types) that mediate other pharmacologic actions of opiates (18-21). At present, little is known about the molecular and cellular mechanisms of opioid actions in these tissues; however, regulation of the expression of opioid receptors could be related to some of these phenomena. To study the regulation of opioid receptor gene transcription, we have cloned a mouse μ opioid receptor gene.[¶] We report here its exonintron structure, transcription initiation sites, and identification and analysis of the promoter region DNA sequences that may contribute to the regulation of μ opioid receptor gene expression.

MATERIALS AND METHODS

Isolation of Mouse μ Opioid Receptor Genomic Clones. A mouse genomic library in Lambda DASHII (Stratagene, cat. no. 945301) was screened for μ opioid receptor genomic DNA sequences (termed MOR). Phage containing MOR genomic sequences were isolated by *in situ* plaque hybridization using three probes derived from rat MOR-1 cDNA (10): a 1.4-kb *Eco*RI-*Hin*dIII fragment containing the entire MOR coding region, a 200-bp *Eco*RI-*Stu* I fragment of 5' untranslated region (5'-UTR), and a 208-bp polymerase chain reaction (PCR) product generated with primers homologous to MOR-1 cDNA 3' coding sequence (bases 1376-1400) and complementary to MOR-1 cDNA 3'-UTR (bases 1549-1583). Positive phage clones were isolated as single plaques after tertiary screening and further characterized by restriction mapping, Southern blot analysis, and DNA sequencing.

DNA Sequence Analysis. Appropriate DNA fragments from restriction enzyme-digested genomic clones were subcloned into pMOB plasmid vector (22) for transposon-facilitated DNA sequencing using a TN1000 kit (Gold Biotechnology, St. Louis). DNA sequencing reactions were performed using the Sequenase kit (United States Biochemical). Both strands of DNA were sequenced using synthetic oligonucleotide primers corresponding to rat MOR-1 cDNA (10), as well as G186 and G187 transposon-specific primers provided with the TN1000 kit. Nucleotide and amino acid sequence analysis was performed using the IntelliGenetics Suite of molecular biology software (IntelliGenetics).

RNase Protection Assay. Poly(A)⁺ RNA was prepared from mouse brain using the FastTrack mRNA isolation kit (Invitrogen). To generate a complementary RNA (cRNA) probe, a Pst I-Pvu II mouse genomic DNA fragment from phage λ MOR2 (see *Results*), containing sequence 93-445 bp upstream of the MOR translation start codon (see Fig. 3), was subcloned into pBluescript KS(+) (Stratagene). Plasmid DNA (1 μ g) was linearized by EcoRI digestion and transcribed in vitro using T7 RNA polymerase and $[\alpha^{-32}P]CTP$ to make the antisense RNA probe. An aliquot of the antisense RNA probe $(5 \times 10^5 \text{ cpm})$ was hybridized to various amounts of mouse brain $poly(A)^+$ RNA (see Fig. 3), as well as to yeast RNA used as a negative control, in 5 M guanidine isothio-cyanate as described (23). The mRNA³²P-labeled cRNA hybrids were diluted 1:40 with RNase buffer and digested with various concentrations of an RNase A/RNase T1 mixture (RPAII kit, Ambion, Austin, TX). After ethanol precipitation, fragments protected from RNase digestion were electrophoresed on a 6% polyacrylamide/7 M urea denaturing gel alongside dideoxynucleotide-sequenced DNA for size

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: RACE, rapid amplification of cDNA ends; UTR, untranslated region; cRNA, complementary RNA. [§]To whom reprint requests should be addressed.

⁴The sequences reported in this paper have been deposited in the GenBank data base (accession nos. U10558, U10559, U10560, and U10561).

estimation. The gel was exposed to X-Omat AR x-ray film (Kodak) at -70° C with an intensifying screen.

5' Rapid Amplification of cDNA Ends (5' RACE). 5' RACE was performed using the components of a 5' RACE kit (GIBCO/BRL). Mouse brain poly(A)⁺ RNA (1 µg) (Clontech) was reverse-transcribed with Superscript RNase H⁻ reverse transcriptase at 42°C using a primer complementary to MOR-1 cDNA (bases 610-632). After removal of excess nucleotides and primer on a GlassMax spin column, the purified cDNA was homopolymeric-tailed with dCTP using terminal deoxynucleotidyl transferase. Oligo(dC)-tailed cDNA was then amplified by PCR using the oligo(dC)-tailspecific primer, provided in the 5' RACE kit, and a nested gene-specific primer complementary to MOR-1 cDNA (bases 536-557). Amplified PCR products were cloned in the pAMP1 plasmid vector using the uracil DNA glycosylase/ CloneAMP system (GIBCO/BRL). Transformants were screened by colony hybridization with a ³²P-end-labeled oligonucleotide probe specific to the 5' end of rat MOR-1 cDNA (bases 15-39). Plasmids from positive colonies were further characterized by DNA sequencing to determine transcription start sites.

RESULTS

Exon-Intron Structure of the MOR Gene. To obtain genomic clones containing the MOR gene, 10⁶ phage from an amplified mouse genomic library were screened with a 1.4-kb EcoRI-HindIII fragment of rat MOR-1 cDNA. All of 70 positive primary clones were collected and rescreened with a 200-bp EcoRI-Stu I fragment from the 5'-UTR and the 1.4-kb coding region probe. Two positive clones that hybridized with both probes were plaque-purified. Based on restriction endonuclease maps, the two clones contained identical 18.3-kb inserts. Southern blot analysis of BamHIdigested DNA from one of these clones (λ MOR2) with the 200-bp 5'-UTR probe revealed a 0.85-kb band identical in size to a *Bam*HI-digested mouse genomic DNA band identified by Southern blot analysis (data not shown). Nucleotide sequence of the 0.85-kb BamHI fragment revealed an exonintron junction coinciding with nt 493/494 of the rat MOR-1 cDNA (Fig. 1). Phage that hybridized with the 1.4-kb coding region probe, but not with the 200-bp 5'-UTR probe, were also plaque-purified. Southern blot analysis of HindIIIdigested DNA of one of these clones (λ MOR10) identified 1.6-kb and 7.8-kb HindIII fragments that hybridized with the 1.4-kb cDNA probe. Comparison of λ MOR10 nucleotide sequence with that of rat MOR-1 cDNA confirmed that the 1.6-kb and 7.8-kb AMOR10 HindIII fragments contained the second and third exons, respectively (Fig. 1). Since the λ MOR10 genomic clone did not contain the 3' end of the coding region, we screened the genomic library to obtain a phage containing the fourth exon. A 208-bp PCR product overlapping the MOR translation stop codon was generated from rat MOR-1 cDNA and used as a probe. A positive clone $(\lambda MOR11)$ was isolated that contained a 3.7-kb Sal I-BamHI fragment that hybridized with the 3'-end probe. The nucleotide sequence of the 3.7-kb Sal I-BamHI fragment identified the fourth exon (Fig. 1). To determine whether the three phage clones overlapped, we used Southern blot hybridization. A restriction fragment generated from the 3' end of λ MOR2 did not hybridize to λ MOR10, and a probe generated from the 3' end of λ MOR10 did not hybridize to λ MOR11 (data not shown). Although these clones do not overlap, the three phage clones (λ MOR2, λ MOR10, and λ MOR11) contained the entire amino acid coding sequence for the mouse μ opioid receptor (Figs. 1 and 2).

The first intron extends 8.2 kb from the 3' end of exon 1 in the λ MOR2 insert and includes 3.0 kb on the 5' side of exon 2 in the λ MOR10 insert. To obtain an overlapping clone, we probed the genomic library with a 200-bp Sal I-HindIII fragment from the 3' end of the λ MOR2 insert. The resulting clone λ MOR3 contained a 16.3-kb genomic insert that hybridized with the 200-bp 3' end of the λ MOR2 insert. However, a 5'-end probe from λ MOR10 did not hybridize with λ MOR3. Therefore, intron 1 is >26 kb (Fig. 1). Intron 2 is ≈ 0.8 kb. The size of intron 3 remains to be determined, but it is at least 12 kb because the inserts in λ MOR10 and λ MOR11 do not overlap (Fig. 1). All of the amino acid coding sequence was determined (Fig. 2), as were all coding region splice junction sequences. All of the intron boundaries conformed to consensus donor and acceptor sequences (5'-GT-AG-3') for RNA splicing. Junctions between exons were found in the first intracellular loop (amino acid 95), the second extracellular loop (amino acid 213), and the cytoplasmic C-terminal region (amino acid 386/387). The mouse MOR

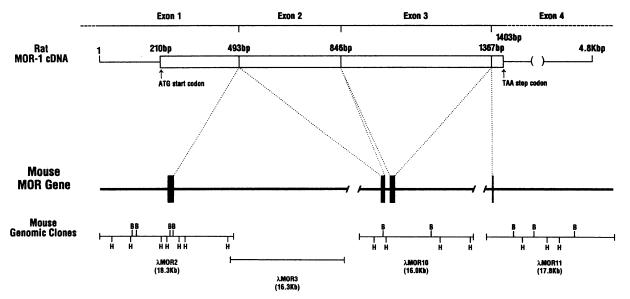


FIG. 1. Structure of the mouse MOR gene. The center line represents mouse genomic DNA, with vertical bars indicating exons determined by sequence comparison with rat MOR-1 cDNA (top line). Diagonal breaks in the genomic DNA indicate unknown intron distances. Genomic inserts from the four mouse MOR phage clones isolated in this study are drawn to scale relative to the genomic DNA, with their *Bam*HI (B) and *Hind*III (H) restriction sites indicated.

Mouse	ATGGACAGCAGCGCC	GCCCAGGGAACATC	AGOGACTGCTCTGAC CCCTTAGC	CCTGCA AGTTGCTCCCCAGCA 75/25
	MDSSA	GPGNI	S D C S D P L A	
Rat	T	T		Q
Mouse		N L S H V	GATGGCAACCAGTCC GACCCATGC D G N O S D P C	
Rat	PGSWL		DGNQS DPC	GPNRTGL - L
Mouse	GCCGCGGGGGCCACAGC	CTGTGCCCTCAGACC	GCAGCCCTTCCATG GTCACAGCO	
Theater	GGSHS	LCPQT	G S P S M V T A	ITIMALY
Rat	N D -			<u></u>
			< Exon	1 Exon 2>
Mouse	TCTATCGTGTGTGTGTA	GTGGGCCTCTTTGGA	AACTTCCTGGTCATG TATGTGATT	
	sıvcv	VGLFG	NFLVMYVI	VRYTKMK
Rat	<u></u>		<u> </u>	<u>-</u>
		TM1.		
Mouse	ACTGCCACCAACATC		GCTCTGGCAGATGCC TTAGCCACT	
	ΤΑΤΝΙ	YIFNL	ALADA LAT	STLPFQS
Rat	<u>-</u>			
			11/2	
Mouse			GGAAACATCCTCTGC AAGATCGTC	
	VNYLM	GTWPF		Ι S Ι D Y Y N
Rat			- T <u></u>	<u></u>
Mouse	ATGTTCACCAGTATC		ATGAGTGTAGACCGC TACATTGCC	
	MFTSI	FTLCT	M S V D R Y I A	V С Н Р V К А
Rat	TMB		<u> </u>	
Mouse		00000333300003333	ATTGTCAATGTCTGC AACTGGATC	
mouse	L D F R T	P R N A K		
Rat	LDFKI	PRNAK	IVNVCNWI	LSSAIGL
Rac				TM4
Mouse	COCOTTA A TOTAL TO A TO	Exo	1 2 Excel 3> Aggcaggggggccata gattgcacc	
Thuse	PVMFM	ATTKY		L T F S H P T
D - 4				
Rat Mouse	TOGTACTOGGAGAAC	CTOCTCAAAATCTGT		
Rat Mouse	TOGTACTOOGAGAAC W X W E N	CTOCTCAAAAATCTOT	GTCTTCATCTTCGCC TTCATCATC	CCGGTC CTCATCATCACTGTG 750/250
			GTCTTCATCTTCGCC TTCATCATG	CCGGTC CTCATCATCACTGTG 750/250
Mouse			GTCTTCATCTTCGCC TTCATCATG	CCGGTC CTCATCATCACTGTG 750/250
Mouse	W Y W E N 	L L K I C	GTCTTCATCTTCGCC TTCATCATG	COGGTC CTCATCATCACTOTG 750/250 PVLIITV TM5
Mouse Rat	W Y W E N 	L L K I C	GTCTTCATCTTCGCC TTCATCATG V F I F A F I M 	COGGTC CTCATCATCACTOTG 750/250 PVLIITV TM5
Mouse Rat	W Y W E N TGITTATGGACTGATG	L L K I C	GTCTTCATCTTCGCC TTCATCATG V F I F A F I M 	COGGTC CTCATCATCACTGTG 750/250 P V L I I T V I
Mouse Rat Mouse	W Y W E N TGTTATGGACTGATG C Y G L M 	L L K I C 	GTCTTCATCTTCGCC TTCATCATG V F I F A F I M 	CCGGTC CTCATCATCACTGTG 750/250 P V L I I T V I TMS AAAGAA AAGGACAGGAACCTG 825/275 K E K D R N L
Mouse Rat Mouse Rat	W Y W E N TGTTATGGACTGATG C Y G L M 	L L K I C 	GTETICATETTOSCE TICATEATE V F I F A F I M AGTETECECATECTES TOSSECTOR S V R M L S G S GTOSCEGTATETTATE GTETECTOS	CCOGTC CTCATCATCACTGTG 750/250 P V L I I T V TM5 AAAGAA AAGGACAGGAACCTG 825/275 K E K D R N L
Mouse Rat Mouse Rat	W Y W E N TGITATGGACTGATG C Y G L M CGCAGGATCACCOGG	L L K I C 	GTCTTCATCTTCGCC TTCATCATC V F I F A F I M AGTCTCCCCATCCTG TCGGGCTCC S V R M L S G S GTGGCTGTATTTATT GTCTGCTGG V A V F I V C W	CCGGTC CTCATCATCACTOTG 750/250 P V L I I T V I TM5 AAAGAA AAGGACAGGAACCTG 825/275 K E K D R N L
Mouse Rat Mouse Rat Mouse	W Y W E N TGITATGGACTGATG C Y G L M CGCAGGATCACCOGG	L L K I C 	GTCTTCATCTTCGCC TTCATCATC V F I F A F I M AGTGTCCCCATGCTG TCGGGCTCC S V R M L S G S GTGGCTGTATTTATT GTCTCCTGG	CCGGTC CTCATCATCACTOTG 750/250 P V L I I T V I TM5 AAAGAA AAGGACAGGAACCTG 825/275 K E K D R N L
Mouse Rat Mouse Rat Mouse	W Y W E N TGTTATOGACTGATG C Y G L M COCAGGATCACCOGG P R I T R ATCATCAAAGCACTG	L L K I C 	GTGTTCATCTTCGCC TTCATCATC V F I F A F I M AGTGTCCGCATGCTG TCGGGCTCC S V R M L S G S GTGGCTGTATTTATT GTGCGTGG V A V F I V C W 	COCGGTC CTCATCATCACTGTG 750/250 P V L I I T V - I
Mouse Rat Mouse Rat Mouse Rat Mouse	W Y W E N TGTTATOGACTGATG C Y G L M COCAGGATCACCOGG P R I T R 	L L K I C 	GTGTTCCATCTTCGCC TTCATCATC GTGTTCCGCATGCTG TCGGGCTCC S V R M L S G S GTGGCCGGTATTTTATT GTGTCGTGG V A V F I V C W 	CCGGTC CTCATCATCACTGTG 750/250 P V L I I T V I TM5 AAAGAA AAGGACAGGACCTG 825/275 K E K D R N L ACCCCC ATCCACATCTATGTC 900/300 T P I H I Y V
Mouse Rat Mouse Rat Mouse Rat Mouse Rat	W Y W E N TGTTATGGACTGATG C Y G L M COCAGGATCACCOGG P R I T R ATCATCAAAGCACTG I K A L 	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	AGRETCOCCATGCTC TCATCATC V F I F A F I M AGRETCCCCATGCTC TCOGGCTCC S V R M L S G S GTGGCTGTATTTATT GTCTGCTGC V A V F I V C W TM6 ACCACTTTCCAGACT GTTTCCTGG T T F Q T V S W	CCCCCC TCCATCATCACTCGTC 750/250 P V L I I T V - I
Mouse Rat Mouse Rat Mouse Rat Mouse	W Y W E N 	L L K I C 	GTGTTCCATCTTCGCC TTCATCATC GTGTTCCCCATGCTG TCGGGCTCC S V R M L S G S GTGGCCTGTATTTATT GTCTGCTGG V A V F I V C W 	CCCGGTC CTCATCATCACTGTG 750/250 P V L I I T V TM5 AAAGAA AAGGACAGGAACCTG 825/275 K E K D R N L
Mouse Rat Mouse Rat Mouse Rat Mouse Rat	W Y W E N TGTTATGGACTGATG C Y G L M COCAGGATCACCOGG P R I T R ATCATCAAAGCACTG I K A L 	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	GTGTTCCATCTTCGCC TTCATCATC GTGTTCCCCATGCTG TCGGGCTCC S V R M L S G S GTGGCCTGTATTTATT GTCTGCTGG V A V F I V C W 	CCCCCC TCCATCATCACTCGTC 750/250 P V L I I T V - I
Mouse Rat Mouse Rat Mouse Rat Mouse Rat	W Y W E N TGTTATOGACTGATG C Y G L M COCAGGATCACCCGG P R I T R ATCATCAAAGACTGC I I K A L TACACAAACAGCTGC Y N S C	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	GTGTTCCATCTTCGCC TTCATCATC GTGTTCCCCATGCTG TCGGGCTCC S V R M L S G S GTGGCCTGTATTTATT GTCTGCTGG V A V F I V C W 	CCCGGTC CTCATCATCACTGTG 750/250 P V L I I T V TM5 AAAGAA AAGGACAGGAACCTG 825/275 K E K D R N L
Mouse Rat Mouse Rat Mouse Rat Mouse Rat Mouse Rat	W Y W E N TGTTATGGACTGATG C Y G L M COCAGGATCACCOGG P R I T R ATCATCAAAGCACTG I I K A L TACACAAACAGCTGC Y T N S C TMC	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	AGIGITICCATCITUGCC TICATCATCA S V F I F A F I M AGIGITICCGCATGCTG TCGGGCTCC S V R M L S G S GIGGCTGTATTTATT GTGTGCTG V A V F I V C W AU F I V C W AU F I V C W AU F I V S W ACCACTTTCCAGACT GTTTCCTGG T T F Q T V S W AU F I D E N P ACCACTTCCTGGAT GAAACTTCC	COCOGTC CTCATCATCACTCATC 750/250 P V L I I T V TMS AAAGAA AAGGACAGGAACCTG 825/275 K E K D R N L AAGGA ATCCACATCTATGTC 900/300 T P I H I Y V CACTTC TGCATTGCCTTGGGT 975/325 H F C I A L G CACTTC TGCATTGCCTTGGGT 975/325 H F C I A L G CACTTC TGCATTGCCTTGGGT 1050/350 K R C F R E F
Mouse Rat Mouse Rat Mouse Rat Mouse Rat	W Y W E N TGTTTATGACTGATG C Y G L M GGCAGGATCACCOGG P R I T R ATCATCAAAGCACTGG I I K A L TACAAAAGCACTGC Y T N S C TMA TGCATCCCAACTICC	L L K I C 	$\begin{array}{c ccccc} \hline & & & & & & & \\ \hline & & & & & \\ \hline & & & &$	$\begin{array}{cccccccccccccccccccccccccccccccccccc$
Mouse Rat Mouse Rat Mouse Rat Mouse Rat Mouse	W Y W E N TGTTATGGACTGATG C Y G L M COCAGGATCACCOGG P R I T R ATCATCAAAGCACTG I I K A L TACACAAACAGCTGC Y T N S C TMC	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	GTCTTCATCTTCGCC TTCATCATC GTCTTCCATCTTCGCC TTCATCATC V F I F A F I M AGRETCCCCATGCTG TCGGGCTCC S V R M L S G S GTGGCTGTTTTTTTT GTCTGCTGC V A V F I V C W TTM ACCACTTTCCAGACT GTTTCCTGC T T F Q T V S W TATGCGTTCCCGCAT GAAACTTC Y A F L D E N F CAAAACTTCCTGCTCGAT GAAACTTC Y A F L D E N F CAAAACTTCCTGCTCGAT GAAACTTC Q N S A R I R Q	$\begin{array}{cccccccccccccccccccccccccccccccccccc$
Mouse Rat Mouse Rat Mouse Rat Mouse Rat Mouse Rat	W Y W E N TGTTTATGACTGATG C Y G L M GGCAGGATCACCOGG P R I T R ATCATCAAAGCACTGG I I K A L TACAAAAGCACTGC Y T N S C TMA TGCATCCCAACTICC	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	$\begin{array}{cccccccccccccccccccccccccccccccccccc$
Mouse Rat Mouse Rat Mouse Rat Mouse Rat Mouse Rat	W Y W E N TGTTTATGACTGATG C Y G L M COCAGGATCACCOGG P R I T R ATCATCAAAAGCACTG I I K A L 	L L K I C 	$\begin{array}{c cccccc} - & - & - & - & - & - & - & - & - & - $	COCGTC CTCATCATCACTCTC TG P V L I I T V TM5 ANAGA AAGGACAGGAACCTG $925/275$ K E K D R N L ACCCCC ATCCACATCTATGTC $900/300$ T P I H I Y V CACTTC TGCATTGCCTTGGGT $975/325$ H F C I A L G ANACCA TGCTTTTTTGGAGAGTTC $1050/350$ K R C F R E F AACCCT AGGGAACACCCTCC $1125/375$ N T R E H P S
Mouse Rat Mouse Rat Mouse Rat Mouse Rat Mouse	W Y W E N TGTTATOGACTGATG C Y G L M COCAGGATCACCCGG P R I T R ATCATCAAAGCACTGC I K A L TACACAAACACTGC Y N S C TM TGCATCCCAACTTCC C I P T S ACCGCTAATACAGTG	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	GIGICICICATCINGCC TICATCATC GIGITICCATCINGCC TICATCATC V F I F A F I M 	$ \begin{array}{ccccccc} - & - & - & - & - & - & - & - & - & - $
Mouse Rat Mouse Rat Mouse Rat Mouse Rat Mouse Rat	W Y W E N TGTTTATGACTGATG C Y G L M COCAGGATCACCOGG P R I T R ATCATCAAAAGCACTG I I K A L 	L L K I C 	GIGICICICATCINGCC TICATCATC GIGITICCATCINGCC TICATCATC V F I F A F I M 	COCGTC CTCATCATCACTCTC TG P V L I I T V TM5 ANAGA AAGGACAGGAACCTG $925/275$ K E K D R N L ACCCCC ATCCACATCTATGTC $900/300$ T P I H I Y V CACTTC TGCATTGCCTTGGGT $975/325$ H F C I A L G ANACCA TGCTTTTTTGGAGAGTTC $1050/350$ K R C F R E F AACCCT AGGGAACACCCTCC $1125/375$ N T R E H P S

coding sequence has 94% homology to rat MOR-1 cDNA, and analysis of amino acid sequences deduced from genomic sequence revealed that 10 amino acids are different from those encoded by rat MOR-1 cDNA (Fig. 2).

Determination of the Transcription Initiation Site. Transcription initiation sites were determined by comparing RNase protection analysis of poly(A)⁺ mouse brain RNA prepared in our laboratory with the results of 5'-RACE performed on a commercial preparation of mouse brain $poly(A)^+$ RNA. RNase protection analysis suggested the existence of four major transcription start sites: ³²P-labeled cRNA generated from a 353-bp Pst I-Pvu II genomic DNA fragment located between 445 bp and 93 bp upstream of the MOR ATG start codon (Figs. 3B and 4) protected four distinct bands of mouse brain poly(A)⁺ RNA from RNase digestion (Fig. 3). With longer exposure, several additional faint bands were observed. No protected bands were observed when yeast RNA was used as a negative control. Among the 5'-RACE clones obtained, 48 were selected that hybridized to a probe homologous to the 5' end of rat MOR-1 cDNA. The most 5' end observed in these 48 clones was 377 bases upstream of the ATG translation start codon. However, repeatedly observed 5' ends at 291 bases (five clones). 284 bases (three clones), 276 bases (four clones), and 268 bases (three clones) upstream of the translation start codon suggested four major start sites. The heterogeneous sizes of fragments protected from RNase digestion, in agreement with the heterogeneity observed in the 5'-RACE analysis. suggested multiple transcription initiation sites within the MOR promoter. For purposes of navigating relative to the

FIG. 2. Amino acid coding sequence of the mouse MOR gene, deduced from comparison of mouse genomic DNA sequence with rat MOR-1 cDNA sequence. Exon boundaries are indicated above the nucleotide sequence. Seven putative transmembrane (TM) domains are indicated by solid lines. Amino acid differences between mouse genomic and rat cDNA sequence are indicated by alternative amino acids in the rat polypeptide line.

promoter, we have assigned position +1 to the base located 268 bases upstream of the ATG translation start codon (Fig. 4). This represents the smallest protected fragment in the RNase protection experiment, and the shortest multiple 5' end observed in the 5'-RACE experiment.

Nucleotide Sequence Analysis of the 5' Flanking Region. To identify sequence elements in the MOR promoter region that might be involved in regulation of transcription, we sequenced 1.8 kb of DNA upstream from the ATG translation start codon and compared this sequence to the Transcription Factors Database (TFD 7.3, QUEST program, IntelliGenetics) at 100% homology. Several putative binding sites for transacting transcriptional regulatory proteins were found in this sequence (Fig. 4). A CCAAT box is present at positions -140 to -136; however, no consensus TATA box was found between this CCAAT box and the start of transcription. A consensus binding sequence for the Sp1 transcription factor, an important element of many cellular and viral promoters that do not contain TATA boxes (24), was found at positions -99 to -91. Two consensus binding sequences for the AP-2 transcription factor (25, 26) were found, one immediately upstream of the CCAAT box at positions -155 to -145 and the other inverted at positions -665 to -672. A consensus AP-1 (27) site is located at positions -767 to -761. Additional consensus sequences found in the promoter region are a glucocorticoid/mineralocorticoid response element (28) located at positions -1441 to -1427, immune-cell-specific element Pu.1 (29) located at positions -442 to -434, and cytokine response elements NF-IL6 (30) and NF-GMb (31) located at positions -1226 to -1218 and -384 to -378,

9084 Pharmacology: Min et al.

Proc. Natl. Acad. Sci. USA 91 (1994)

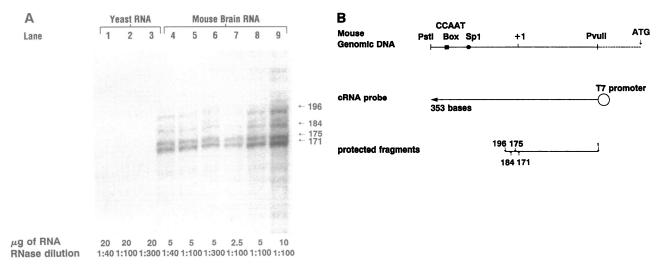


FIG. 3. RNase protection analysis of transcription initiation sites. (A) Autoradiograph of electrophoresed mouse brain $poly(A)^+$ RNA protected from RNase digestion by a cRNA probe generated from the 353-bp *Pst* I–*Pvu* II mouse genomic DNA fragment indicated in B. Reaction products in lanes 1–3 also contained 20 μ g of yeast RNA, and those in lanes 4–9 contain mouse brain RNA in the amounts indicated. All nine reaction mixtures contained the same volume of RNase at the dilutions indicated. Sizes of protected fragments, estimated by a dideoxynucleotide-sequenced DNA ladder (data not shown), are indicated to the right of the four major bands in A and on the bottom line of B.

respectively. Sequences within 1 bp of consensus cAMP response elements (32, 33) were located at positions -248 to -241 and -1483 to -1476.

DISCUSSION

We have isolated three mouse μ opioid receptor genomic clones corresponding to the entire amino acid coding sequence in rat MOR-1 cDNA, determined all exon sequence and splice junctions upstream of the translation stop codon, and identified multiple transcription start sites and potential binding sites for transcriptional regulatory factors 5' of the start codon. Since the three clones do not overlap, we are unable to determine the size of all the introns or of the complete MOR-1 gene, but the latter is >53 kb. Moreover, since Northern blots indicate that the size of rat MOR mRNA includes at least 9 kb 3' of the reading frame (12), there may be additional introns and exons in the MOR-1 genome not identified in this work.

The positions of splice sites are at amino acids 95, 213, and 386/387 in the open reading frame. The first of these sites, occurring at the junction between the first putative transmembrane region and the first intracellular loop, is of particular interest, for we have evidence of splicing at the equivalent position in other opioid receptor, or opioid receptor related, clones. Thus an intron at this junction is also found in genomic DNA for δ opioid receptor (L.B.A., unpublished data) and in another clone isolated in our laboratory (RBC) that has $\approx 60\%$ amino acid homology to δ , μ , and κ opioid receptors (the same level that exists between any two of the latter) but does not

-1532	
-1332	GRE/MRE
-1442	GREMRE G <u>ATCACAAAATGATCT</u> GATAGCAAAGGCAGCAGTAAGCACCACAAGTCAACCTGTCATTCACAAGAAGGOGGTAAAGATGGTTCTGTGTG
-1352	TCCTGGGGAGACCGATGAATGTGAGCACACACACATTTCAGCTCAGCCATATGAACTGACACACTTATTCCCGAAGAGAATATATTTTAAAGGT NF-IL6
-1262	NT*110 ATTITIAGCAGCCTCTAGAAAACCTTATCTACATTAC <u>TGCTGAAATG</u> ACACCCCTTTAAOGCAAAAGOOGGACAAATAATACTGAACGTTGTC
-1172	CAAAAGTCTACAATAGTTGAGCTGTGTGTGTTATCTTTTTTCAGAATGATTGCTTAAAAGTAAAAATATATGTTCCTTCTAAGCTTGCTT
-1082	CCCACAGTCTGTGAGAATTCAGTTAAACTTCTACAACCAAATGOGAGGGGGGACATTTATTTCCAGTATTTATAGTTAAAACTATGOGTA
-992	TGCGCATGGCATGCATATGCATACACAGTGTTGAGACGAAACCACACTACTGCTGAATTTACGATTGTATTCCTTCAGAAGCACTTGTGA
-902	ACTTAGACATGAGATCTAACATTTTTTTTTTTTTTTTTT
-812	GTTTCTTTTGGATTCTCTATAGIGGGAAGAAAAAACATTTCTGGT <u>TCATTCA</u> GTAACCTGCTCCCCAGAAGCTCGGGAGGAGGGGAGG
-722	ALZ TGOGAAAGCTTROGATAAATCTTCCTAAGCAGGTCCAGGAGAGTAGACCT <u>CGCATGGG</u> TRGTGGGTTCCTTATTTACAGGTTATCACCCC
-632	TCCTCCCTCACATCTCTAATGTGTTCTTATGCTTCTTCTCACTAGGAAAAAAATCAGATGCAAATAGACCTTCTAACCTAAGAGAAATT
-542	GAAGGAGTGGGGGGGCACATGAAAAAGGCTTCTTTCACAAAAATTTTTCTCTTTTAGAAAAAGACAATTGTTCTTTGAACAGGTTTGTGGGG Pu.1 NF-GMb
-452	GTIGATTTTT <u>TCACCAACT</u> AAAOOOOCCTTTIGTTTGTTTGTTTGTTTGTTTGTTTGTTTGCCC <u>TCACATA</u> TOOOGATGTQGGTAA
-362	AGGACAACATTAATAACCCTCCAGTCCTAATTAATTGCATATGGTTAGTTGAGTCGGAAGAGTGTTGAGGTGATAAGGGTCATTGAGATTT CRE
-272	TAACAGACCAACGACTTTCCTGGCTGATGCAAACAACCTCAAACAACCTAATGCATGTTGTTACAAAGAAACTTAGAGTCTCTCTTCCCA AP2 CCAAT box Sp1
-182	CCCTCT <u>CTGCAGCATCCCCCCCTCCCCCCCCTACCCCAAT</u> TTACACTCCTTTACACGGAAGTGCATAGAGACTGAGGAGGCTG
-92	ATTCRASTICCTTIGCGCACAATCCACTCCTTCTCTCCTCCCCCCCCTCTAGCCTCTGGATCCCTCACAGCCCATGCTCCCCTC
-2	CCACTCAGAGAGTGGGGCTTTGGGGATGCTAAGGATGCGCCTCCGTGTACTTCTAAGGTGGGAGGGGGCTACAAGCAGAGGAGAATATCG
+89	GACGCTCAGACGTTCCATTCTGCCTGCCGCTCTTCTCTGGTTCCACTAGGGCCTTGTCCTTGTAAGAAACTGACGGAGCCTAGGGCAGCTG
+179	TGAGAGGAAGAGGCTGGGGCGCCTGGAACCCGAACACTCTTGAGTGCTCTCAGTTACAGCCTACGGAGTCCGCAGCAAGCA
+269	ATEGACAGCAGC M D S S

CDP

FIG. 4. Nucleotide sequence of MOR gene 5' flanking region. Sequence is numbered relative to position +1 representing the shortest major band in the RNase protection experiment. The ATG translational start codon is underlined at position +269. The Pst I and Pvu II restriction sites used to subclone the RNase protection probe are indicated centered at positions -173 and +176, respectively. Sites of consensus sequence for possible cis-acting transcriptional regulatory elements are underlined and labeled above the sequence, CRE, cAMP response element; GRE/MRE, glucocorticoid/mineralocorticoid response element.

MDSS

express opioid binding when transfected into eukaryotic cells (M. J. Wick and H.H.L., unpublished data).

The existence of an intron at this position raises the possibility of an alternative form of MOR (as well as of the δ opioid receptor and other related clones) in which additional amino acids have been inserted into the first intracellular loop. There is precedent for such a splice variant, as the third cytoplasmic loop of the D₂ dopamine receptor exists in at least two forms, one of which has an additional 29 amino acids (34), and such differences could be related to either differences in ligand binding and/or interactions with second messengers.

The multiple transcription initiation sites observed in this study may be associated with the absence of a well-defined TATA box. Nevertheless, some definition of the start of transcription is provided by the sequence in the promoter region, since four major initiation sites spanning 24 bases were observed \approx 280 bp upstream of the ATG translation start codon in both RNase protection assays and 5'-RACE studies. In addition to a CCAAT box and Sp1 site, the putative promoter region also contains consensus sequences for AP-1 and AP-2 transcription factors and potential cAMP response elements. Since μ opioid receptors are known to be coupled to adenylyl cyclase in brain (35, 36) and in some clonal cell lines (19, 21), the finding of these motifs suggests a possible feedback mechanism between receptor activation and receptor expression. We speculate that such feedback could play an important role in down-regulation of μ opioid receptors by chronic agonist treatment (21).

We also found response elements for glucocorticoid/ mineralocorticoid and cytokine response elements (NF-IL6 and NF-GMb). The presence of corticoid response elements is consistent with interactions between opioids and this system and in particular might explain a known fact that corticoid levels can exert long-lasting effects on morphine antinociception (37). The presence of the cytokine response elements is also intriguing, because numerous associations between opioids and the immune system have been documented in recent years (38). In particular, opioids are known to alter the growth, proliferation, or activity of several types of immune cells, consistent with the possibility that they may in fact act as cytokines in these cells. Work from our own laboratory supports this conclusion. Thus morphine treatment of bone marrow stem cells reduced macrophage colonystimulating factor-induced proliferation of these cells (39, 40). While it has been difficult to demonstrate the presence in immune cells of receptors that fulfill all the classical pharmacological criteria of opioid receptors, PCR screening in this laboratory has identified the MOR-1 mRNA in human neutrophils and rat thymocytes (S. Roy and H.H.L., unpublished data).

Rat MOR-1 cDNA was kindly provided by Dr. Lei Yu (Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis). We also thank Dr. A. Smith for helping to prepare this manuscript and Ms. Stephanie Fuchs for her technical assistance. This research was supported by National Institute on Drug Abuse Grants DA00564, DA01583, DA05695, and K05-DA-70554 and the F. & A. Stark Fund of the Minnesota Medical Foundation.

- Attali, B., Saya, D., Nah, S. Y. & Vogel, Z. (1989) J. Biol. Chem. 264, 347–353.
- 2. Blume, A. J. (1978) Proc. Natl. Acad. Sci. USA 75, 1713–1717.
- 3. Fedynyshyn, J. P. & Lee, N. M. (1989) Brain Res. 476, 102-109.
- Koski, G. & Klee, W. A. (1981) Proc. Natl. Acad. Sci. USA 78, 4185–4189.

- Law, P. Y., Wu, J., Koehler, J. E. & Loh, H. H. (1981) J. Neurochem. 36, 1834–1846.
- Nestler, R. J., Erdos, J. J., Terwilliger, R., Duman, R. S. & Tallman, J. F. (1989) Brain Res. 476, 230-239.
- 7. North, R. A., Williams, J. T., Surprenant, A. & Christie, M. J. (1987) Proc. Natl. Acad. Sci. USA 84, 5487-5491.
- Evans, C. J., Keith, D. E., Jr., Morrison, H., Magendzo, K. & Edwards, R. H. (1992) Science 258, 1952–1955.
- Kieffer, B. L., Befort, K., Gaveriaux-Ruff, C. & Hirth, C. G. (1992) Proc. Natl. Acad. Sci. USA 89, 12048–12052.
- Chen, Y., Mestek, A., Liu, J., Hurley, J. A. & Yu, L. (1993) Mol. Pharmacol. 44, 8-12.
- 11. Wang, J. B., Imai, Y., Eppler, M., Gregor, P., Spivak, C. E. & Uhl, G. R. (1993) Proc. Natl. Acad. Sci. USA 90, 10230-10234.
- 12. Thompson, R. C., Mansour, A., Akil, H. & Watson, S. J. (1993) Neuron 11, 903–913.
- 13. Fukuda, K., Kato, S., Mori, K., Nishi, M. & Takeshima, H. (1993) FEBS Lett. 327, 311–314.
- 14. Chen, Y., Mestek, A., Liu, J. & Yu, L. (1993) Biochem. J. 295, 625-628.
- Yasuda, K., Raynor, K., Kong, H., Breder, C. D., Takeda, J., Reisine, T. & Bell, G. I. (1993) Proc. Natl. Acad. Sci. USA 90, 6736-6740.
- Li, S., Zhu, J., Chen, C., Chen, Y., Deriel, J. K., Ashby, B. & Liu-Chen, L. (1993) Biochem. J. 295, 629-633.
- Meng, F., Xie, G., Thompson, R. C., Mansour, A., Goldstein, A., Watson, S. J. & Akil, H. (1993) Proc. Natl. Acad. Sci. USA 90, 9954–9958.
- Lord, J. A. H., Waterfield, A. A., Hughes, J. & Kosterlitz, H. W. (1977) Nature (London) 267, 495-499.
- 19. Frey, F. A. & Kebabian, J. W. (1984) Endocrinology 115, 1797-1804.
- Ward, S. J., LoPresti, D. & James, D. W. (1986) J. Pharmacol. Exp. Ther. 238, 625-631.
- Yu, V. C., Eiger, S., Duan, D. S., Lameh, J. & Sadee, W. (1990) J. Neurochem. 55, 1390-1396.
- Strathmann, M., Hamilton, B. A., Mayeda, C. A., Simon, M. I., Meyerowitz, E. M. & Palazzolo, M. J. (1991) Proc. Natl. Acad. Sci. USA 88, 1247-1250.
- 23. Thompson, J. & Gillespie, D. (1987) Anal. Biochem. 163, 281-291.
- 24. Pugh, B. F. & Tjian, R. (1990) Cell 61, 1187-1197.
- 25. Williams, T. & Tjian, R. (1991) Genes Dev. 5, 670-682.
- 26. Imagawa, M., Chiu, R. & Karin, M. (1987) Cell 51, 251-260.
- 27. Lee, W., Mitchell, P. & Tjian, R. (1987) Cell 49, 741-752.
- 28. Beato, M. (1989) Cell 56, 335-344.
- Klemsz, M. J., McKercher, S. R., Celada, A., Van Beveren, C. & Maki, R. A. (1990) Cell 61, 113-124.
- Akira, S., Isshiki, H., Sugita, T., Tanabe, O., Kinoshita, S., Nishio, Y., Nakajima, T., Hirano, T. & Kishimoto, T. (1990) *EMBO J.* 9, 1897–1906.
- Shannon, M. F., Gamble, J. R. & Vadas, M. A. (1988) Proc. Natl. Acad. Sci. USA 85, 674–678.
- Montminy, M. R., Sevarino, K. A., Wagner, J. A., Mandel, G. & Goodman, R. H. (1986) Proc. Natl. Acad. Sci. USA 83, 6682-6686.
- 33. Leza, M. A. & Hearing, P. (1988) J. Virol. 62, 3003-3013.
- 34. Eidne, K. A., Taylor, P. L., Zabavnik, J., Saunders, P. T. K.
- & Inglis, J. D. (1989) Nature (London) **342**, 865. 35. Beitner, D. B., Duman, R. S. & Nestler, E. J. (1989) Mol. Pharmacol. **35**, 559-564.
- Pharmacol. 35, 559–564.
 36. Chneiweiss, H., Glowinski, J. & Premont, J. (1988) J. Neuro-
- sci. 8, 3376-3382. 37. Ratka, A., Sutanto, W. & De Kloet, E. R. (1988) Neuroendo-
- crinology **48**, 439–444. 38. Sibinga, N. E. S. & Goldstein, A. (1988) Annu. Rev. Immunol.
- 6, 219–249. 39. Roy, S., Ramakrishnan, S., Loh, H. H. & Lee, N. M. (1991)
- Eur. J. Pharmacol. 195, 359-363.
 40. Roy, S., Ge, B.-L., Ramakrishnan, S., Lee, N. M. & Loh, H. H. (1991) FEBS Lett. 287, 93-96.
- 41. Loh, H. H. & Smith, A. P. (1990) Annu. Rev. Pharmacol. Toxicol. 30, 123-147.