Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 8565–8569, November 1989
Neurobiology

Olfactory marker protein gene: Its structure and olfactory neuron-specific expression in transgenic mice

(neuron promoter/intronless/lack of TATA and CAAT/Thy-1)

E. DANCIGER*, C. METTLING†, M. VIDAL‡, R. MORRIS†, AND F. MARGOLIS*§

*Department of Neurosciences, Roche Institute of Molecular Biology, Roche Research Center, Nutley, NJ 07110; and †Laboratory of Neurobiology, Norman and Sadie Lee Center, and ‡Laboratory of Gene Structure and Expression, National Institute for Medical Research, The Ridgeway, Mill Hill, London, NW7 1AA, England

Communicated by Sidney Udenfriend, July 24, 1989

ABSTRACT **Olfactory marker protein (OMP) genomic clones were isolated from a Charon 4A phage λ rat genomic library. A 16.5-kilobase (kb) fragment of the rat genome containing the gene was isolated and characterized. Sequence analysis of the gene showed the absence of introns and the lack of CAAT and TATA boxes in the 5′ flanking region. The transcription initiation site was mapped, and two sites 55 and 58 base pairs upstream of the ATG were observed. The 5′ flanking region is rich in G+C residues and contains a G+C-rich motif as well as direct and inverted repeats. Functional OMP regulatory sequences were demonstrated in transgenic mice. An 11-kb chimeric gene was constructed in which the coding region for OMP was replaced with that for Thy-1.1. In *Thy-1.2* mice carrying this transgene, Thy-1.1 was expressed solely by olfactory receptor neurons and their axons and terminals in the olfactory bulb.**

The vertebrate olfactory pathway consists of the bipolar olfactory receptor neurons in the nasal neuroepithelium, their axons, and synaptic terminations in the synaptic glomeruli of the olfactory bulb that is located intracranially. These sensory neurons have the unusual ability to be replaced from progenitor cells that persist in the adult neuroepithelium.

Olfactory marker protein (OMP) is a highly abundant 19-kDa cytoplasmic protein that is expressed almost exclusively in mature, functioning, olfactory neurons but not in the neural precursor basal cells. OMP is initially expressed in olfactory neurons early in the last trimester of gestation at about the time when innervation of the bulb begins. Although its biological function is unknown, OMP has a wide species distribution, being present in many vertebrate classes including man (see ref. 1 for a review). Its tissue-specific expression in the receptor cells might imply involvement either in processes of regeneration or transduction.

The cDNA for OMP was cloned, characterized and sequenced in our laboratory (2). In the present study, we used this clone to isolate and characterize the gene coding for OMP. This analysis shows that the gene is devoid of introns and lacks traditional upstream CAAT and TATA boxes. To evaluate whether the upstream sequences are functional in tissue-specific expression, transgenic mice were created by using *Thy-1.1* as a reporter gene.

MATERIALS AND METHODS

Cloning of the Gene for OMP. A rat genomic library in Charon 4A λ phage was screened with a ^{32}P-labeled OMP cDNA probe. The library (a gift from J. Bonner and L. Jagodzinski), consisted of a partial *Hae* III digest ligated into *Eco*RI sites of the phage and contained 1.8×10^6 independent clones. Of 5×10^5 phages screened, those containing OMP sequences were isolated, digested with restriction endonucleases, fractionated by electrophoresis in 1% agarose, transferred to nylon membranes (Biotrans; ICN), and hybridized to probes from both the 5′ and the 3′ regions of OMP cDNA. An 11-kilobase (kb) *Eco*RI fragment obtained from a 16.5-kb insert of one of the phages isolated was subcloned into pBluescript plasmid (Stratagene) and mapped by restriction enzyme analysis.

DNA Sequence Analysis. Two *Hin*dIII–*Sac* I fragments (2.6 kb and 1.2 kb) were subcloned into M13mp18 and sequenced by the Sanger method (3). Deletions were generated for the 2.6-kb fragment to provide a series of overlapping clones by using the Cyclone Biosystem Kit (IBI) according to the protocol recommended by the manufacturer. The 1.2-kb fragment was sequenced by the Sanger method with synthetic primers. The overlap between these clones was sequenced by using the 11-kb *Eco*RI fragment subcloned into pBluescript with synthetic primers as above.

S1 Nuclease Analysis. Total cellular RNA (10 μg), isolated from rat olfactory epithelium or cerebral hemispheres by the guanidinium isothiocyanate procedure (4), was hybridized with a 5′-end-labeled *Hin*dIII–*Acc* I probe. The genomic fragment was end-labeled with [γ-^{32}P]ATP (5000 Ci/mmol; 1 Ci = 37 GBq) and T4 polynucleotide kinase (Boehringer–Mannheim). The RNA and the probe were precipitated together and resuspended in 15 μl of hybridization buffer (80% formamide/0.4 M NaCl/40 mM Pipes, pH 6.4/1 mM EDTA). After 15 hr of hybridization at 60°C, the samples were diluted with 300 μl of ice-cold buffer containing 40 mM potassium acetate (pH 4.5), 2.5 mM $ZnCl_2$, 300 mM NaCl, 20% (vol/vol) glycerol and 250 units of S1 nuclease (Boehringer–Mannheim). The S1 nuclease digest proceeded for 30 min at 30°C and was stopped by addition of 1 μl of 0.5 M EDTA. Protected fragments were resolved by electrophoresis in 6% polyacrylamide/urea gel and visualized by autoradiography.

Primer Extension Analysis. Poly(A)$^+$ RNA from rat olfactory epithelium (5 μg) was hybridized to ^{32}P-labeled synthetic oligonucleotides (0.2 pmol) complementary to nucleotides +104 to +121 in the coding region or to nucleotides −113 to −130 in the 5′ upstream region. After precipitation of the RNA and the primer, the samples were resuspended in 18.5 μl of reverse transcriptase buffer (50 mM Tris·HCl, pH 8.0/50 mM KCl/5 mM $MgCl_2$/5 mM dithiothreitol), heated for 3 min at 95°C, and allowed to cool gradually to 35°C (≈30 min). After the annealing process, the following components were added: 3.5 μl of dNTPs (4 mM each), 1 μl of RNasin (Bethesda Research Laboratories), and 2 μl of reverse transcriptase (Boehringer–Mannheim), and the extension reaction proceeded for 60 min at 42°C. Extended fragments were

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.

Abbreviation: OMP, olfactory marker protein.

§To whom reprint requests should be addressed.

analyzed by electrophoresis in 6% polyacrylamide/urea gel and visualized by autoradiography.

Production and Characterization of Transgenic Mice with a Chimeric Gene. A hybrid construct where the entire OMP gene coding sequence was removed from the 11-kb *Eco*RI fragment and replaced by the *Thy-1.1* coding sequence was made as follows: the 1.5-kb *Bst*EII–*Bam*HI *Thy-1.1* genomic fragment was ligated to the 294-base pair (bp) *Hin*dIII (position −240, see Fig. 3)–*Nco* I (position +54) OMP gene fragment by using a 64-bp linker reconstituting the missing Thy 1.1 amino acids and was cloned in the *Hin*dIII and *Bam*HI sites of the pEMBL19 vector (5). The 5′ and 3′ flanking OMP sequences were introduced by cloning the 5.2-kb *Eco*RI–*Hin*dIII OMP fragment in the *Hin*dIII site of this plasmid with an *Eco*RI–*Hin*dIII linker and by cloning the 3.2-kb *Kpn* I–*Eco*RI OMP fragment into the *Kpn* I and *Eco*RI sites. This 10.3-kb fusion *Eco*RI fragment was introduced into CBA × C57B1 mice, which endogenously express only the *Thy-1.2* allele.

For immunohistochemistry, deeply anesthetized mice were fixed by intracardiac perfusion for 30 min with 0.5% paraformaldehyde/0.5% glutaraldehyde in 100 mM phosphate buffer (pH 6) followed by 153 mM ethanolamine, and their olfactory mucosa, bulbs, and brain were removed and processed through polyester wax as before (6). Serial 5-μm sections in both coronal and sagittal planes as far caudal as medulla hind brain were alternatively incubated with monoclonal antibody OX7 [to identify Thy-1.1 (7)], 30H12 [to identify Thy-1.2 (7)], and a goat antibody against rat OMP (8). After the sections were washed and incubated with appropriate horseradish peroxidase-conjugated secondary antibodies, the binding was visualized by reaction with 3,3′-diaminobenzidine tetrahydrochloride and H_2O_2. Nuclei were counterstained with thionin. Sections from liver and kidney were similarly treated.

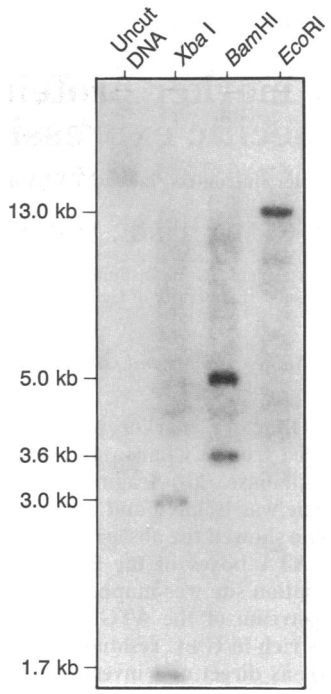

FIG. 2. Southern analysis of rat genomic DNA. Genomic DNA isolated from rat thymus (10 μg) was digested with *Eco*RI, *Xba* I, and *Bam*HI. The southern blot was probed with ^{32}P-labeled OMP cDNA. Hybridization (overnight, 37°C) was carried out in 50% formamide/4× SSPE/sodium dodecyl sulfate/0.1% low fat milk/10% dextran sulfate/50 μg of denatured salmon sperm DNA per ml. After hybridization, the blot was washed with: 2× SSC/0.1% sodium dodecyl sulfate at 23°C and then at 50°C with 0.1× SSC/0.1% sodium dodecyl sulfate (1× SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7).

RESULTS

Isolation of the OMP Gene. Upon screening 5 × 10^5 plaques from a Charon 4A rat genomic library with OMP cDNA as a probe, 12 phages were isolated. A phage hybridizing to probes both from the 5′ and the 3′ regions of OMP cDNA was further characterized (λ-OMP). An 11-kb *Eco*RI fragment obtained from the 16.5-kb phage insert (λ-OMP), was subcloned into pBluescript plasmid, and a partial restriction map was constructed (Fig. 1); upon hybridization to OMP-cDNA, a 2.6-kb *Hin*dIII–*Sac* I fragment was found to contain the entire cDNA information. A Southern blot of rat genomic DNA digested with *Eco*RI, *Xba* I, and *Bam*HI was hybridized to OMP-cDNA (Fig. 2). The bands obtained agreed with the restriction map of λ-OMP and imply that the OMP gene is a single-copy gene. In addition to the strong signals, a few of low intensity were also seen; these latter bands originate from hybridization with a probe spanning the 3′ noncoding region of the OMP cDNA. To confirm that OMP is a single-copy gene the library was rescreened with OMP cDNA as the probe in low-stringency conditions (35% formamide/4 × SSPE/1% low fat milk/50 μg of denatured salmon sperm per ml; 1× SSPE = 0.15 M NaCl/10 mM phosphate, pH 7.4/1 mM EDTA). The phages isolated from the genomic library were all generated from the same genomic locus. When they were probed with fragments flanking the OMP gene [1.2-kb *Sac* I–*Hin*dIII 5′ to the gene and 3.0-kb *Xba* I–*Eco*RI 3′ to the gene (Fig. 1)], all hybridized to either one or both probes.

Sequence of the Rat Gene for OMP. Two *Hin*dIII–*Sac* I fragments, the 2.6-kb fragment hybidizing to OMP cDNA and the adjacent 5′ upstream fragment, were subcloned into

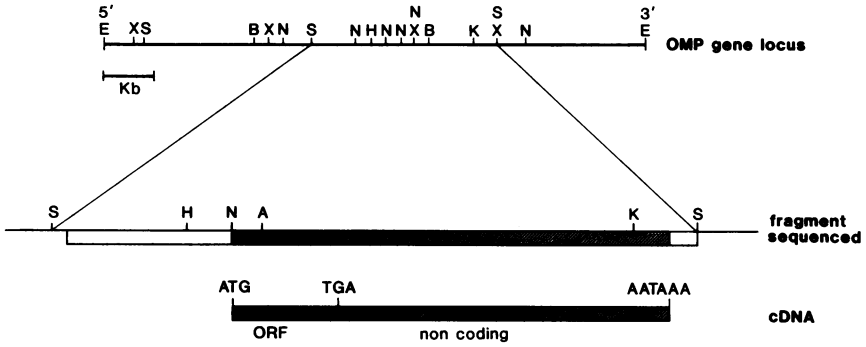

FIG. 1. Partial restriction map of the rat OMP gene locus and its comparison to OMP cDNA. Partial map of the 11-kb *Eco*RI insert obtained from the λ-OMP phage. E, *Eco*RI; X, *Xba* I; S, *Sac* I; B, *Bam*HI; N, *Nco* I; H, *Hin*dIII. Schematic comparison is shown between the 2.6-kb genomic *Hin*dIII–*Sac* I fragment and OMP cDNA.

M13mp18 and sequenced by dideoxy chain termination (3). Deletions were generated to provide a series of overlapping clones by using the exonuclease activity of phage T4 polymerase (Cyclone Biosystem Kit; IBI). The 1.2-kb *Hin*dIII–*Sac* I fragment was sequenced with synthetic primers. The sequence of the 3122-nucleotide-long genomic fragment is present in Fig. 3. The sequence obtained was colinear with the cDNA sequence (2), except for a few mismatches in the 3′ noncoding region. This discrepancy could be a true polymorphism or could be due to reverse transcriptase artifacts in the cDNA cloning. At the 3′ end, 16 bp downstream of the polyadenylylation signal, there are two G/T clusters that are found also in several other genes (9) and are believed to be important for the 3′ processing of precursors of polyadenylated mRNAs. In the region 5′ to the transcription initiation sites (as determined by S1 nuclease analysis and primer extension to be discussed below), no typical promoter motifs such as TATA and CAAT boxes were found. This upstream region is G+T rich and the G+C content (61.4%) is similar to that observed in the coding region of the OMP cDNA (58.9%) as compared with the 3′ noncoding region (53.1%). The presumed OMP promoter region contains a GC rich motif (GGCGGG), direct repeats and inverted repeats.

```
-801  CTCAGTTCTG CCATTTGTCC ACAATCACTG TCTGGGTTCC ACCTCCCACT GACACTCGCC AGTATGTAGA
-731  TCTCCAGCGT CCTCCCCGGC CTCACATCTC CCAGGGGAGG TGGAGGATCA GTTTAGGTGA AATACCTGGT
-661  AGGATTTTGG TGGCAGTGAG AGCCAGGCCC ATAGCTATGT GGGGCAGTAT GCGGTTGGAT CAATCAGACC
-591  ACAGCCTCTG GGAGCCGAGC CAGCCATCTG TCTGGCAGAT GATTGGGATT TGAGAGCTGC AGGTTCAGAT
-521  GGGAGGTGAC AGCTGGGCTG GGTCCTGGTG ATGATAAAGG AGAGGGAGAA ACCCAGGGCA TCTGACGGGA
-451  CCTGGCAGGG ACTTCGAGAG AATGAGGTGG GGAGGAAGCA ACCATGGTAA GTCGGCTTGG CTGACTACAG
-381  AGAACGAATG TATGCTACTG GTGCTGCCCC TCCCTGCCCC CTCTTCCTGG GCAGTCTCCA GTTACCTCCA
-311  TGTGTCTGTC ACCCCACCTG TCTCCCAACA GCCCTGTGGA GTATTCTGTT CTTCACAAAC AAGCAAACCT
-241  CAAGCTTGCC ACTAGCACTG TAGTCAAGGT GGTTGCCACA GCAGTTGATA CCCATGCTCT GGTCCCCAAG
-171  GAGCCTGTCA CCCTCCAGCC TGCCTACGGC ACGGCTTTGC CACTTAGAAG GCAGTTGGAC ACACACTCAT
-101  GTGTCCCCTG TTCTGAGAAC TGGGTGGGGC CAGGAAGGCT GGAAAGGGAG GCGGGCCTTC AGGTGGCCTC
 -31  TTCTCGTGGC ACCTGAGGCT CCAGCCCACT TGATTCCCTG ACGTCTGTGG CAGTGGTGGC AGTGGCAACA
  39  GCTGTAGCAC TTGGGCCATG GCAGAGGACG GGCCACAGAA GCAGCAGCTG GATATGCCGC TGGTTCTGGA
 109  CCAGGACCTG ACTAAGCAGA TGCGGCTCCG AGTAGAGAGC CTGAAGCAGC GCGGGGAGAA GAAGCAGGAT
 179  GGTGAGAAGC TGCTCCGGCC GGCTGAGTCT GTCTACCGCC TTGATTTCAT CCAGCAGCAG AAGCTGCAGT
 249  TCGATCACTG GAACGTGGTT CTGGACAAGC CGGGCAAGGT CACCATCACG GGCACCTCGC AGAACTGGAC
 319  GCCAGACCTC ACCAACCTCA TGACACGCCA GCTGCTGGAC CCTGCTGCCA TCTTCTGGCG CAAGGAAGAC
 389  TCCGATGCCA TGGATTGGAA TGAGGCAGAC GCCCTGGAGT TTGGGGAGCG CCTTTCTGAC CTGGCCAAGA
 459  TCCGCAAGGT CATGTATTTC CTCATCACCT TTGGCGAGGG TGTGGAGCCC GCCAACCTAA AGGCCTCTGT
 529  GGTGTTTAAC CAGCTCTGAT GGCAGCCGTG GCTCGCTGCC TTCGGCCCCA ACTCTCCCTT GGCTGGACCT
 599  CCTAGCTCAT GTGTATTTTG GAAACATTCT TCTAGCTGTT CCTTCTGTGC TCATCTTGGC TAGAGGTCCC
 669  CTGAGTGCTA CACCCGCTCT TTTTCCCTGG TGTCAGTGCC ACGGCTCACA GGGATGTCCC ATGGCTTCAT
 739  AGTCTAGAAG CTGGACGCTG CTACTCTAGA CAGTAGAGGC CTTTTGGGTC CATGTGGCCA GAGGGATGAG
 809  CCTCTTGGCC ACCTGCCATC TCTGCTTTAT TGTGGTGAAG AACAGGATTG AGAGAGAAAA GAGACTGACC
 879  AAGAAATGCC AACGGCCATC ATGATTCCTC CCTTTGGGGA CAAGAGGCTG AGACTGGACA GGAACACCTT
 949  CCAGGGATCC GGGGGAGAAG GCTTTTCCCT GCTGGCCAAA GCTGGAACCA GGAGGTGAAT ACCCAGCTGC
1019  ACATCGGCAG CAGGAAGGTG CTTCTTCCAG TGTTGGCATC AGCCCGCGGT GACCTTAGGG CCTTCCAGAC
1089  ACTTGGCGGG ATGACAGCAG GGCTTGATCT GACTGGTTTT CCAGGTCTGG CCCCGGTTTT TATGGAGTCG
1159  TGAGAGAACG CGTAGAAACG GAAACAGCCC TAAGCTACCT ATACTCATAA GTATATTGAG AAATAGCCTG
1229  ACTGTATCTG TATGGATGTG TGCCTGAGAG CTATTCCTAG TCAGGCGTAA GGCTAACTCT AGTTTAATTG
1299  TTGAGCTGGT ACTGGTTTGT GGGCTTGGTG GAAGTGACCC TGGCTAAGCC TTTCTTGGTA CAGTGCTCTT
1369  TGAACTGGGG GACTGAGGCT CAAATGGTGA AGCAGAGAAC TGCATTAGAG GGGTCCAGGA CTTTGAGCTA
1439  GAAACACTTC CATTAGGAAG GCTGGCATTT GCTGATCACT GGTCATGGTG TTGTATCCTG CTGTCACCTC
1509  CCTGGGCATT TCCTAGCTGT TTTCCTGGAG TGAGGGCACA TGGGTAAGGC TTGGGGGCAG TTATGATGCC
1579  TGACATCTGA TGTGTGCTGG AGCTGTCCGG CTATGATGCC TAGTACTGGC CTCAGAGGCT GTCCAGGACA
1649  GCCACTCAGC AAACTGACAG AAAAACTATG GCACAGTTAT CAGCAGATTC AACCCTGCCC CAAGTCTCAT
1719  TGTGCCTCAC CCTGCACATC CTGAGAGCCT CTCATTGAGG AGAAGCCTCA CCTGTCACCT AGCATCAGCC
1789  AGGGGCACCC CAGCAAAGCC CTCGACTCTG TCTCAAGGCT GCCTCCATTG GGGATACCAA GATCTGAAGG
1859  TAGGAGTTGT CCCGGTGGGT GGGTTGTTAG AAGGGCAAGC CCTAGCTTAG ATTCAGGATT CTGGAGGCAG
1929  GAGATAGGCT GTGGTACCTA CCGGTCTCTT CTACTGGTGC CTCTGCACTG GCCAGTGCTT CTGCACTTTG
1999  CTGACTCTAG GGAGCCATAC CTAGACAGAC CTACCTTTCT GCTTCTCTTT CTGCCTCTCC CTACAGCTTT
2069  AGAGACTCCT TTCACACTGC CAGACCCCCA ATTCTGTCTC ACTCCATTTG CCCTATGGGG ACAGTTGTGT
2139  CTCTGCTGTG CCTGTCACAC AATAAAGACT GTATGCCCTC CCCTCTGTGG TGTGGTGTGT GCGTCTCCAT
2209  GTGGGGGTTT GTTTATCTTG CTGCAAGCCC ATAGCATCAG AAACCTTCTC TCGTGGGCCC CGTGGTTCTG
2279  AGGAACTCCA CTCTGCCAGC CTTAAGTTTG AGCAGTGAGC TC
```

FIG. 3. Sequence of rat OMP gene fragment. The sequence of the 3122-nucleotide-long genomic fragment is presented. Underlined are (*i*) the postulated upstream binding site for transcription factor Sp1, (*ii*) the ATG initiation codon, (*iii*) the TGA termination codon, and (*iv*) the AATAAA polyadenylylation signal and the G+T clusters beyond the signal. The site of the poly(A) tail at the 3′ end of OMP mRNA occurs at nucleotide 2180. *, Transcription initiation sites.

Mapping of the Transcription Initiation Site. S1 nuclease and primer extension analyses were used to map the transcription initiation site. The *Hin*dIII–*Acc* I DNA fragment used for S1 nuclease analysis spans the region between −235 and +212. The fragment was end-labeled and hybridized with 10 μg of total RNA from olfactory epithelium or cerebral hemispheres. Two major fragments of 212 and 219 bp were protected (Fig. 4), identifying two transcription start sites 55 and 62 bp upstream of the ATG where translation begins. To further establish this 5′ end, the transcription initiation site of OMP mRNA was mapped by primer extension. ^{32}P-labeled 18-mer synthetic oligonucleotides complementary to nucleotides +104 to +121 and −113 to −130 were hybridized with 5 μg of poly(A)$^+$ RNA from olfactory epithelium. The primers were extended with reverse transcriptase, and the extended products were sized by electrophoresis in a denaturing gel (Fig. 5). Extended fragments were observed only with the primer complementary to the OMP coding region (Fig. 5, lane 2). Two extended fragments were observed, 121 bp and 124 bp in length, demonstrating the presence of two initiation sites 55 and 58 bp upstream of the ATG. The same primer was used to sequence the gene. The sequence was analyzed on a denaturing gel in lanes adjacent to those containing the primer-extended fragments, enabling us to determine the exact nucleotides corresponding to the transcription initiation sites (data not shown). Comparing the data from both methods, we were able to determine two transcription initiation sites corresponding to guanosine and cytidine residues: the first is at position +1 and the other is at −3 (Fig. 3).

Functional Identification of Sequences Determining Cellular Specificity in the OMP Gene. Functional OMP gene regulatory sequences were demonstrated in transgenic mice. A hybrid gene was constructed in which the *Thy-1.1* coding region, third intron, and part of the 3′ untranslated region was flanked by OMP sequences. Upstream, 5.5 kb of OMP gene sequence was fused to *Thy-1.1* at the ATG; downstream, 3.2 kb of OMP gene sequence containing the polyadenylylation signal but only 200 of its 1600-nucleotide 3′ untranslated region was used. The OMP gene-*Thy-1.1* chimeric gene was injected into pronuclei of oocytes of a host mouse strain that is *Thy-1.2* in genotype. This allelic difference arises from a single amino acid substitution and can be distinguished by monoclonal antibodies (10). Twelve transgenic animals were obtained, and four lines were analyzed. Immunohistochemical localization of *Thy-1.1* in *Thy-1.2* host transgenic mice was restricted to the olfactory receptor neurons and their axons in the fiber and glomerular layers of the olfactory bulb (Fig. 6). These neurons do not normally express either *Thy-1.1* allele. No *Thy-1.1* expression was seen elsewhere in brain or in liver or kidney. The level of Thy-1.1 determined by RIA(6) in liver and kidney of transgenic mice was <0.01% and for the central nervous system minus bulb was <0.05% of that seen in control mice. However, in the transgenic olfactory bulb the level was 15–35% of that observed in control mouse central nervous system. Therefore, this pattern of *Thy-1.1* expression is consistent with that expected for olfactory neuron specific expression determined by the OMP gene.

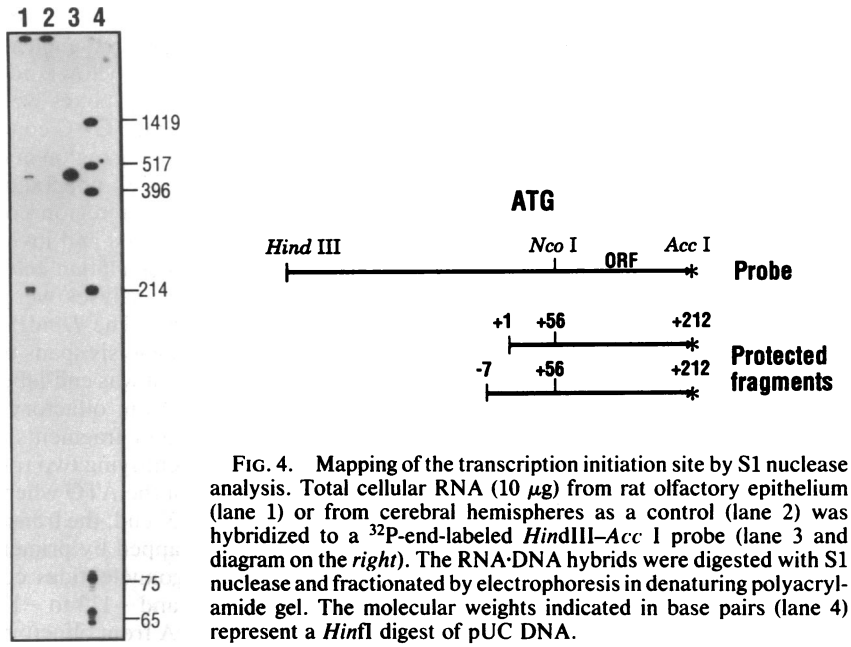

FIG. 4. Mapping of the transcription initiation site by S1 nuclease analysis. Total cellular RNA (10 μg) from rat olfactory epithelium (lane 1) or from cerebral hemispheres as a control (lane 2) was hybridized to a ^{32}P-end-labeled *Hin*dIII–*Acc* I probe (lane 3 and diagram on the *right*). The RNA·DNA hybrids were digested with S1 nuclease and fractionated by electrophoresis in denaturing polyacrylamide gel. The molecular weights indicated in base pairs (lane 4) represent a *Hin*fI digest of pUC DNA.

DISCUSSION

The olfactory marker protein (OMP) is expressed only in the nervous system and almost exclusively in one population of neurons, the mature olfactory receptor neurons (11). In these neurons OMP is phylogenetically conserved and developmentally regulated (1). We have now demonstrated that the functional gene for OMP lacks CAAT and TATA boxes and, as it is colinear with its mRNA, is intronless as well.

Very few examples of functional vertebrate genes are known with these characteristics. Genes for the adrenergic receptor (12), muscarinic acetylcholine receptor (13), interferon (14), phosphoglycerate kinase (15), *JUN* protooncogene (16), and histone (17) are reported to be devoid of introns. Those for epidermal growth factor receptor (18), Ha-*ras* protooncogene (19), 3-hydroxy-3-methylglutaryl-CoA reductase (20), hypoxanthine phosphoribosyltransferase (21), adenosine deaminase (22), Thy-1 (23), T-cell antigen receptor-associated T3δ (24), and DNA polymerase β chain (25) lack CAAT and TATA boxes. OMP is devoid of both properties. Since these characteristics are also generally associated with nonfunctional pseudogenes, it was essential to evaluate whether the genomic fragment that we had isolated was biologically functional. To determine whether this isolated gene contained the requisite information enabling it to be expressed according to the appropriate biological pattern, we utilized transgenic mice.

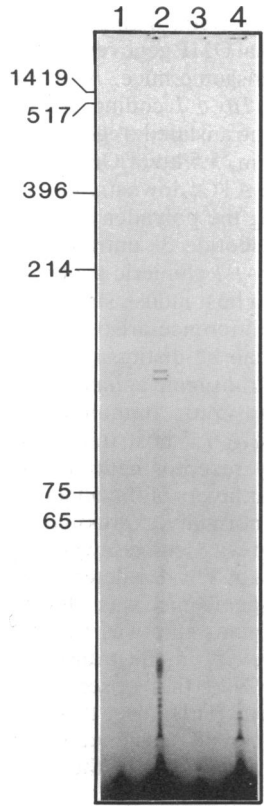

FIG. 5. Mapping of the transcription initiation site by primer extension analysis. Poly(A)$^+$ RNA (5 μg) from olfactory epithelium was hybridized to ^{32}P-labeled synthetic oligonucleotides (10^7 cpm/pmol) complementary to: primer A, nucleotides +104 to +121 in the coding region or primer B, nucleotides −113 to −130 from the 5′ upstream region beyond the transcription start site. The extended fragments synthesized by reverse transcriptase were analyzed on a denaturing polyacrylamide gel. Lanes 1, primer A; 2, extended fragment using primer A; 3, primer B; 4, extended fragment using primer B.

Southern blot analysis of rat genomic DNA was in agreement with the restriction map of λ-OMP, implying the gene for OMP is a single-copy gene. Support for this finding also derives from the mapping of OMP gene only to mouse chromosome VII (D. Pravtcheva and F. Ruddle, personal communication).

Sequence analysis of the gene revealed that it is intronless. The 5′ nontranslated region has two transcription initiation sites as determined by S1 nuclease and primer extension analysis. The promoter region lacks characteristic TATA and CAAT boxes, is rich in G+C content (61.4%), and contains a G+C-rich motif, direct repeats, and inverted repeats. When the region (−120) to (+37) was analyzed for secondary structure by the method of Zuker and Stiegler (26), a predicted hairpin displays a high free energy of −54.3 kcal/mol. The hairpin includes 37 bp beyond the transcription start site. Since some characteristics of this promoter (G+C-rich and lack of TATA and CAAT boxes) are typical of many "housekeeping" genes, this predicted stable secondary structure may be involved in preventing OMP gene expresion in all other cells and thus restrict its tissue-specific expression to olfactory neurons.

A chimeric gene, generated from presumptive regulatory sequences derived from the rat OMP gene and coding sequences from the *Thy-1.1* gene, was used to create transgenic mice. The *Thy-1* allelic system allows it to be used as a reporter gene, since transgenic Thy-1.1 product can be unambiguously distinguished from endogenous Thy-1.2 by antibodies. In addition, Thy-1 is not normally expressed by primary olfactory neurons, and so its appearance on their axons can only be due to the influence of the OMP regulatory elements. On the other hand, Thy-1 is normally expressed on

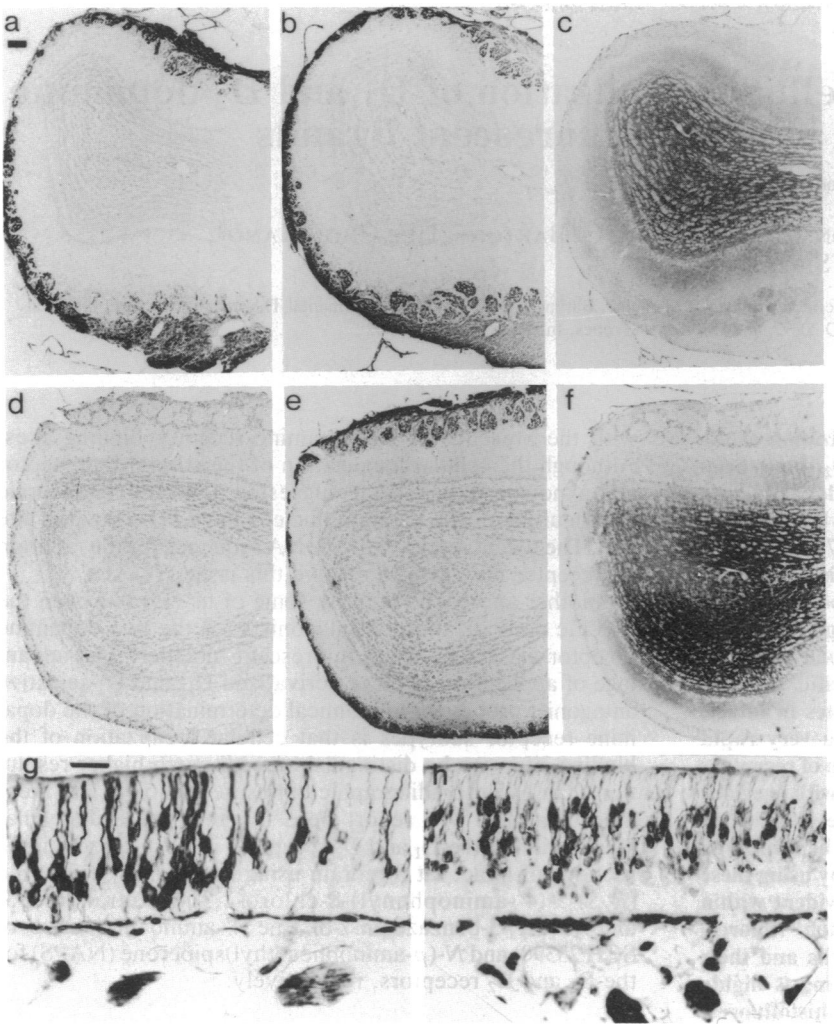

FIG. 6. Immunohistochemical localization of Thy.1.1 (*a*, *d*, and *g*), OMP (*b*, *e*, and *h*), and the endogenous Thy-1.2 (*c* and *f*) in transgenic (*a–c*, *g* and *h*) and control (*d–f*) mice. [Bar = 200 μm for olfactory bulbs (*a–f*) and 20 μm for olfactory mucosa (*g* and *h*).] The axons of the primary olfactory neurons and their terminals, which stain heavily for OMP, are also heavily labeled for Thy-1.1 in the transgenic but not control mouse. Cell nuclei are counterstained with thionin (e.g., *d*, which is immunohistochemically negative for Thy-1.1).

all other neurons (7), whereas the chimeric gene is not expressed elsewhere in the brain, indicating the normal *Thy-1* neural regulatory elements are missing from this construct.

The ability of the chimeric gene to direct Thy-1.1 protein expression in transgenic animals according to the pattern expected for OMP demonstrates that the genomic clone we have isolated includes at least the regulatory sequences controlling the olfactory cell-specific pattern of expression of the OMP gene.

We are now in a position to identify the minimal essential sequence information necessary to achieve the biological specificity of OMP expression and to determine the mechanisms by which this is attained. Understanding the processes involved in this gene regulation offers the possibility of identifying other genes selectively expressed in these neurons and associated with their neurogenesis and detection, transduction, and transmission of olfactory information.

We thank M. Grillo for providing the RNA and K. Ianuzzi for synthesizing the oligonucleotides. Drs. J. Bonner and L. Jagodzinski generously provided the rat genomic library in Charon 4A.

1. Margolis, F. L. & Getchell, T. V. (1988) *Molecular Neurobiology of the Olfactory System* (Plenum, New York).
2. Rogers, K., Dasgupta, P., Gubler, U., Grillo, M., Khew-Goodall, Y.-S. & Margolis, F. L. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 1704–1708.
3. Sanger, F. & Coulson, A. R. (1978) *FEBS Lett.* **87**, 107–110.
4. Chirgwin, J. M., Przybyla, A. E., Macdonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
5. Dente, L., Cesareni, G. & Cortese, R. (1983) *Nucleic Acids Res.* **11**, 1645–1655.
6. Morris, R. J. & Barber, P. C. (1983) *J. Histochem. Cytochem.* **31**, 263–274.
7. Morris, R. J. (1985) *Dev. Neurosci.* **7**, 133–166.
8. Keller, A. & Margolis, F. L. (1975) *J. Neurochem.* **24**, 1101–1106.
9. Birnstiel, M. L., Busslinger, M. & Strub, K. (1985) *Cell* **41**, 349–359.
10. Williams, F. F. & Gagnon, J. (1982) *Science* **216**, 696–703.
11. Baker, H., Grillo, M. & Margolis, F. L. (1989) *J. Comp. Neurol.* **285**, 246–261.
12. Emorine, L. J., Marullo, S., Delavier-Klutchko, C., Kaveri, S. V., Durieu-Trautman, O. & Strosberg, A. D. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 6995–6999.
13. Bonner, T. I., Young, A. C., Bronn, M. R. & Buckley, J. J. (1988) *Neuron* **1**, 403–410.
14. Nagata, S., Mantei, N. & Weissmann, C. (1980) *Nature (London)* **287**, 401–410.
15. McCarrey, J. R. & Thomas, K. (1987) *Nature (London)* **326**, 501–505.
16. Hattori, K., Angel, P., LeBeau, M. M. & Karin, M. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 9148–9152.
17. Schaffner, W., Kunz, G., Daetwyler, H., Telford, J., Smith, H. O. & Birnstiel, M. L. (1978) *Cell* **14**, 655–671.
18. Ishii, S., Xu, Y.-H. Stratton, R. H., Roe, B. A., Merlino, G. T. & Paston, I. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4920–4924.
19. Ishii, S., Merlino, G. T. & Paston, I. (1985) *Science* **230**, 1378–1381.
20. Reynolds, G. A., Basu, S. K., Osborne, T. F., Chin, D. J., Gil, G., Brown, M. S., Goldstein, J. L. & Luskey, K. L. (1984) *Cell* **38**, 275–285.
21. Melton, D. W., Konecki, D. S., Brennaud, J. & Casky, C. J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 2147–2151.
22. Valerio, D., Duyvesteyn, M. G. C., Dekker, B. M. M., Weeda, G., Berkvens, Th. M., Van der Voorn, L., Van Ormondt, H. & Van der Eb, A. J. (1985) *EMBO J.* **4**, 437–443.
23. Giguere, V., Isobe, K.-I. & Grosveld, F. (1985) *EMBO J.* **4**, 2017–2024.
24. Tunnacliffe, A., Sims, J. E. & Rabbits, T. H. (1986) *EMBO J.* **5**, 1253–1260.
25. Yamaguchi, M., Hayoshi, Y. & Matsukage, A. (1988) *Nucleic Acids Res.* **16**, 8773–8787.
26. Zuker, M. & Stiegler, P. (1981) *Nucleic Acids Res.* **9**, 133–148.