Biochimica et Biophysica Acta, 1152 (1993) 197–200

BBAMEM 70732

Short Sequence-Paper

Cloning and analysis of the 5′ flanking sequence of the rat *N*-methyl-D-aspartate receptor 1 (NMDAR1) gene

Guang Bai * and John W. Kusiak

National Institute on Aging, National Institutes of Health, Molecular Neurobiology Unit, Gerontology Research Center, 4940 Eastern Avenue, Baltimore, MD 21224 (USA)

(Received 21 June 1993)

Key words: *N*-Methyl-D-aspartate recepter 1 gene; Transcription start site; Promoter; Housekeeping gene

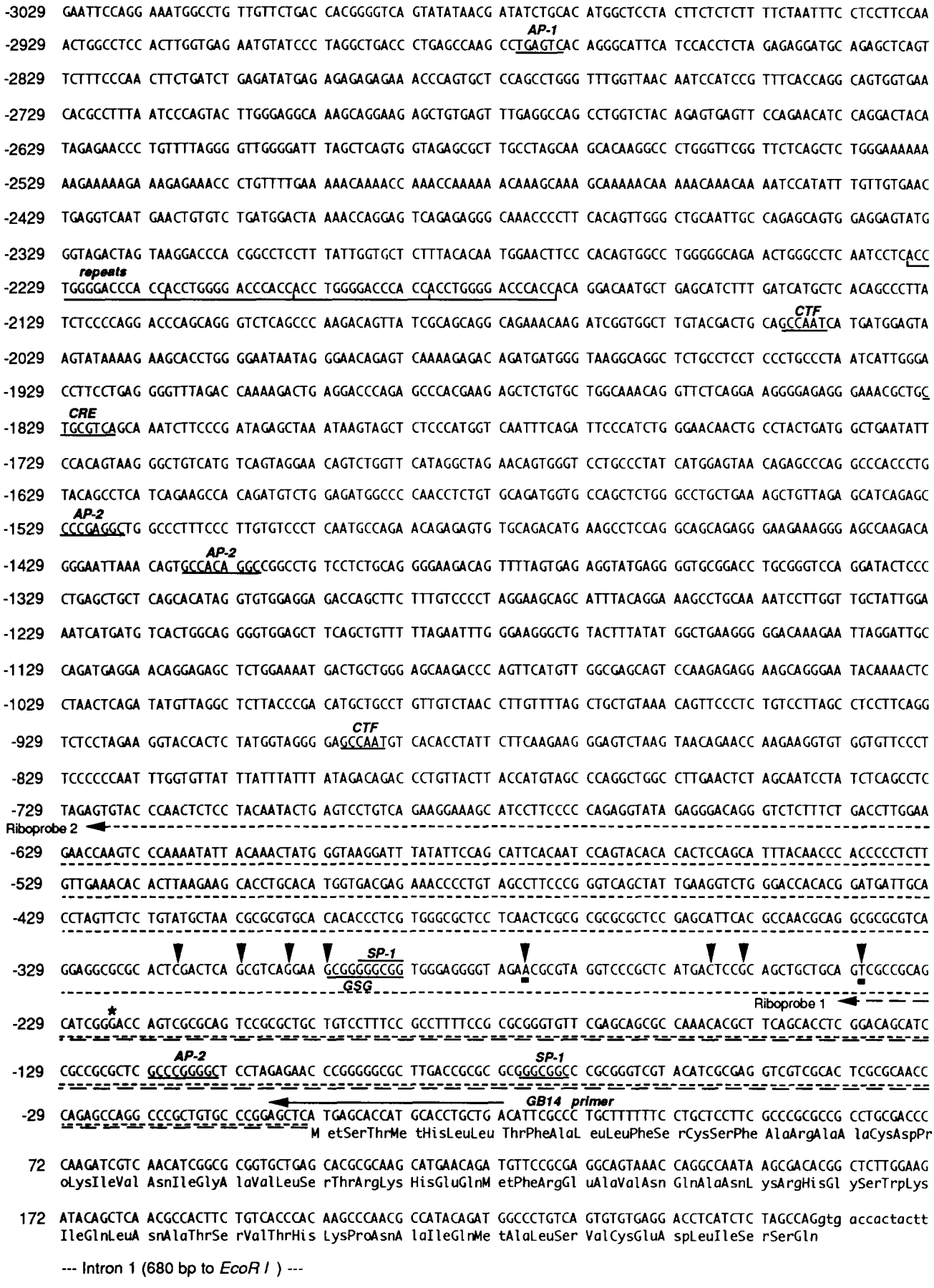
We cloned and analyzed a 3.8 kb *Eco*RI fragment of the rat NMDAR1 gene. It contains 3 kb of promoter/enhancer region, exon 1 and a portion of intron 1. Two major transcription start sites were identified at −276 and −238 from the first nucleotide in codon 1. One GSG and two SP1 motifs, but no TATA/CAAT boxes, exist in the region proximal to the transcription start sites. Our results suggest that NMDAR1 has the characteristics of a housekeeping gene and may be regulated by immediate-early genes.

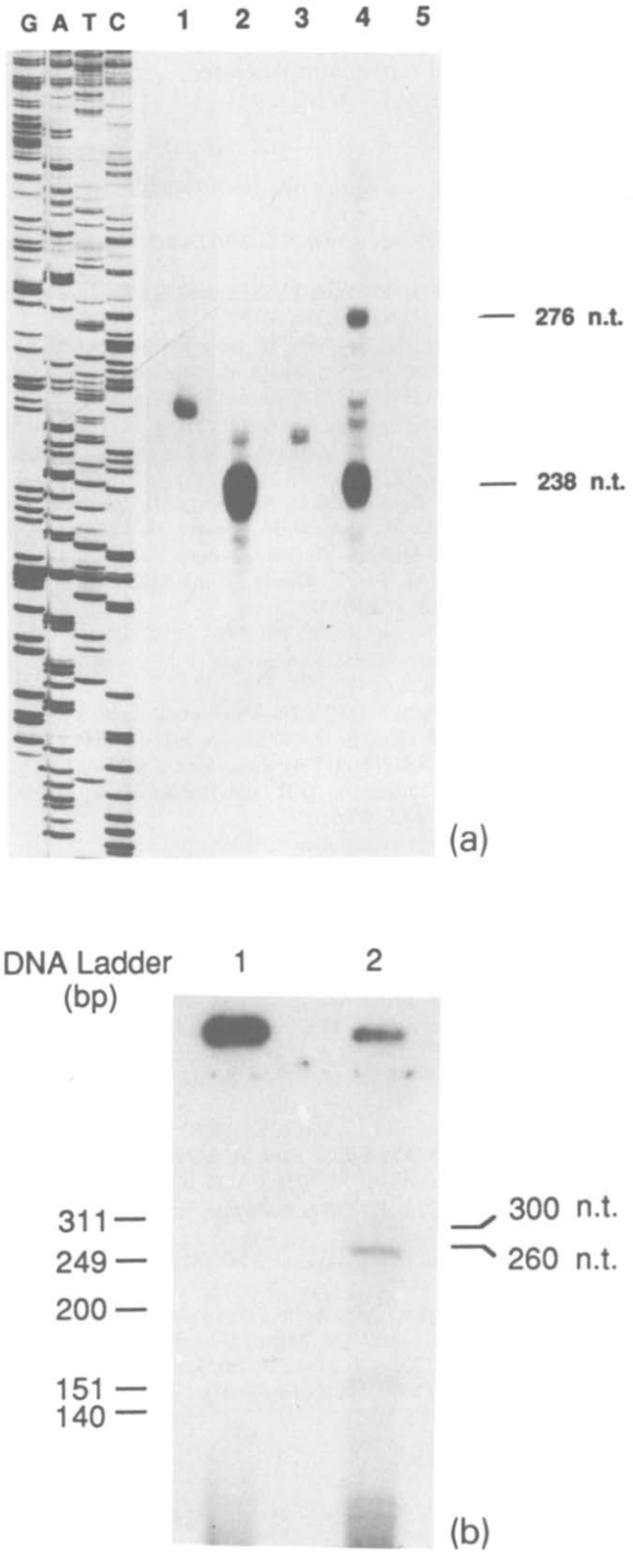
N-Methyl-D-aspartate (NMDA) gated ion channels are one subtype of glutamate receptor in the mammalian central nervous system. They are distinguished from other glutamate receptors by their voltage dependence and high permeability to Ca^{2+} ion. These features of NMDA receptors may underlie their role in synaptic plasticity, learning and memory, and excitotoxicity [1,2]. Recent molecular biology studies revealed that NMDA receptors are composed of two types of subunits, i.e., NMDAR1 [3] and R2 (A, B, C, D) [4–6]. NMDAR1 mRNA is widely expressed in the brain while the distribution of NMDAR2 mRNAs is more limited [4,5]. Data from expression studies of micro-injected mRNA in *Xenopus* oocytes demonstrated that the NMDAR1 subunit is the only one which is capable for forming functional channels. NMDAR2 subunits have channel activity only when co-expressed with the NMDAR1 subunit [4–6]. Further, the amino acid sequence of NMDAR1 is highly conserved from rat to human with 99% homology [7]. Since little is known about its regulation at the transcription level, as an initial approach, we isolated and characterized the 5′ flanking region of the NMDAR1 gene.

Using polymerase chain reaction (PCR), we generated from rat genomic DNA a 1.7 kb product containing sequence from codon 2 to 86, the next intron, and part of the following exon to codon 112 of the NMDAR1 gene [3,6]. In the intron there is an *Eco*RI site which allowed us to use the 5′ portion of the PCR-generated DNA (about 800 bp) as a probe to screen a lambda phage library harboring partial *Eco*RI-digested rat genomic DNA. We identified a clone with a 12 kb insert which was further digested by *Eco*RI into a 9.2 kb fragment and a 3.8 kb fragment. The latter hybridized to the probe after blotting onto nylon membranes. The DNA fragments were cloned and sequenced using a fluorescent *Taq* dye deoxy terminator cycle sequencing kit and a 373A ABI automatic DNA sequencer. Results showed that the 3.8 kb fragment (Fig. 1) contains 3029 bp upstream of the translation start site (+1), an exon covering the region from nucleotide −223 to codon 86 of NMDAR1 cDNA and the next intron up to the *Eco*RI site. The 9.2 kb fragment contains the following five exons and six introns (data not shown). The last 42 bp at the 5′ end of the published cDNA sequence [3], i.e., from −224 to −265, was not found in the 3.8 kb fragment. The isolated genomic sequence lacks a consensus splice acceptor site (Py_6NCAGN) in this region [8] suggesting that DNA recombination may have occurred during cDNA library construction. We defined the 5′ end of NMDAR1 mRNA by using RNase protection with riboprobes generated from the genomic sequence from nucleotide −1 and by primer extension with primer GB14 (Fig. 1). After hybridizing riboprobe 2 to rat brain RNA and digesting with cloned RNase ONE, which non-selectively breaks the phosphodiester bonds of single-stranded RNA [9], two major protected bands

* Corresponding author. Fax: +1 (410) 5588317.

The nucleotide sequence data reported in this paper have been submitted to the EMBL/GenBank Data Libraries under the accession number L19708.


-3029 GAATTCCAGG AAATGGCCTG TTGTTCTGAC CACGGGGTCA GTATATAACG ATATCTGCAC ATGGCTCCTA CTTCTCTCTT TTCTAATTTC CTCCTTCCAA
AP-1
-2929 ACTGGCCTCC ACTTGGTGAG AATGTATCCC TAGGCTGACC CTGAGCCAAG CCTGAGTCAC AGGGCATTCA TCCACCTCTA GAGAGGATGC AGAGCTCAGT
-2829 TCTTTCCCAA CTTCTGATCT GAGATATGAG AGAGAGAGAA ACCCAGTGCT CCAGCCTGGG TTTGGTTAAC AATCCATCCG TTTCACCAGG CAGTGGTGAA
-2729 CACGCCTTTA ATCCCAGTAC TTGGGAGGCA AAGCAGGAAG AGCTGTGAGT TTGAGGCCAG CCTGGTCTAC AGAGTGAGTT CCAGAACATC CAGGACTACA
-2629 TAGAGAACCC TGTTTTAGGG GTTGGGGATT TAGCTCAGTG GTAGAGCGCT TGCCTAGCAA GCACAAGGCC CTGGGTTCGG TTCTCAGCTC TGGGAAAAAA
-2529 AAGAAAAAGA AAGAGAAACC CTGTTTTGAA AAACAAAACC AAACCAAAAA ACAAAGCAAA GCAAAAACAA AAACAAACAA AATCCATATT TGTTGTGAAC
-2429 TGAGGTCAAT GAACTGTGTC TGATGGACTA AAACCAGGAG TCAGAGAGGG CAAACCCCTT CACAGTTGGG CTGCAATTGC CAGAGCAGTG GAGGAGTATG
-2329 GGTAGACTAG TAAGGACCCA CGGCCTCCTT TATTGGTGCT CTTTACACAA TGGAACTTCC CACAGTGGCC TGGGGGCAGA ACTGGGCCTC AATCCTCACC
repeats
-2229 TGGGGACCCA CCACCTGGGG ACCCACCACC TGGGGACCCA CCACCTGGGG ACCCACCACA GGACAATGCT GAGCATCTTT GATCATGCTC ACAGCCCTTA
CTF
-2129 TCTCCCCAGG ACCCAGCAGG GTCTCAGCCC AAGACAGTTA TCGCAGCAGG CAGAAACAAG ATCGGTGGCT TGTACGACTG CAGCCAATCA TGATGGAGTA
-2029 AGTATAAAAG AAGCACCTGG GGAATAATAG GGAACAGAGT CAAAAGAGAC AGATGATGGG TAAGGCAGGC TCTGCCTCCT CCCTGCCCTA ATCATTGGGA
-1929 CCTTCCTGAG GGGTTTAGAC CAAAAGACTG AGGACCCAGA GCCCACGAAG AGCTCTGTGC TGGCAAACAG GTTCTCAGGA AGGGGAGAGG GAAACGCTGC
CRE
-1829 TGCGTCAGCA AATCTTCCCG ATAGAGCTAA ATAAGTAGCT CTCCCATGGT CAATTTCAGA TTCCCATCTG GGAACAACTG CCTACTGATG GCTGAATATT
-1729 CCACAGTAAG GGCTGTCATG TCAGTAGGAA CAGTCTGGTT CATAGGCTAG AACAGTGGGT CCTGCCCTAT CATGGAGTAA CAGAGCCCAG GCCCACCCTG
-1629 TACAGCCTCA TCAGAAGCCA CAGATGTCTG GAGATGGCCC CAACCTCTGT GCAGATGGTG CCAGCTCTGG GCCTGCTGAA AGCTGTTAGA GCATCAGAGC
AP-2
-1529 CCCGAGGCTG GCCCTTTCCC TTGTGTCCCT CAATGCCAGA ACAGAGAGTG TGCAGACATG AAGCCTCCAG GCAGCAGAGG GAAGAAAGGG AGCCAAGACA
AP-2
-1429 GGGAATTAAA CAGTGCCACA GGCCGGCCTG TCCTCTGCAG GGGAAGACAG TTTTAGTGAG AGGTATGAGG GGTGCGGACC TGCGGGTCCA GGATACTCCC
-1329 CTGAGCTGCT CAGCACATAG GTGTGGAGGA GACCAGCTTC TTTGTCCCCT AGGAAGCAGC ATTTACAGGA AAGCCTGCAA AATCCTTGGT TGCTATTGGA
-1229 AATCATGATG TCACTGGCAG GGGTGGAGCT TCAGCTGTTT TTAGAATTTG GGAAGGGCTG TACTTTATAT GGCTGAAGGG GGACAAAGAA TTAGGATTGC
-1129 CAGATGAGGA ACAGGAGAGC TCTGGAAAAT GACTGCTGGG AGCAAGACCC AGTTCATGTT GGCGAGCAGT CCAAGAGAGG AAGCAGGGAA TACAAAACTC
-1029 CTAACTCAGA TATGTTAGGC TCTTACCCGA CATGCTGCCT GTTGTCTAAC CTTGTTTTAG CTGCTGTAAA CAGTTCCCTC TGTCCTTAGC CTCCTTCAGG
CTF
-929 TCTCCTAGAA GGTACCACTC TATGGTAGGG GAGCCAATGT CACACCTATT CTTCAAGAAG GGAGTCTAAG TAACAGAACC AAGAAGGTGT GGTGTTCCCT
-829 TCCCCCCAAT TTGGTGTTAT TTATTTATTT ATAGACAGAC CCTGTTACTT ACCATGTAGC CCAGGCTGGC CTTGAACTCT AGCAATCCTA TCTCAGCCTC
-729 TAGAGTGTAC CCAACTCTCC TACAATACTG AGTCCTGTCA GAAGGAAAGC ATCCTTCCCC CAGAGGTATA GAGGGACAGG GTCTCTTTCT GACCTTGGAA
Riboprobe 2
-629 GAACCAAGTC CCAAAATATT ACAAACTATG GGTAAGGATT TATATTCCAG CATTCACAAT CCAGTACACA CACTCCAGCA TTTACAACCC ACCCCCTCTT
-529 GTTGAAACAC ACTTAAGAAG CACCTGCACA TGGTGACGAG AAACCCCTGT AGCCTTCCCG GGTCAGCTAT TGAAGGTCTG GGACCACACG GATGATTGCA
-429 CCTAGTTCTC TGTATGCTAA CGCGCGTGCA CACACCCTCG TGGGCGCTCC TCAACTCGCG CGCGCGCTCC GAGCATTCAC GCCAACGCAG GCGCGCGTCA
SP-1
-329 GGAGGCGCGC ACTCGACTCA GCGTCAGGAA GCGGGGGCGG TGGGAGGGGT AGAACGCGTA GGTCCCGCTC ATGACTCCGC AGCTGCTGCA GTCGCCGCAG
GSG
Riboprobe 1
-229 CATCGGGACC AGTCGCGCAG TCCGCGCTGC TGTCCTTTCC GCCTTTTCCG CGCGGGTGTT CGAGCAGCGC CAAACACGCT TCAGCACCTC GGACAGCATC
AP-2
SP-1
-129 CGCCGCGCTC GCCCCGGGGCT CCTAGAGAAC CCGGGGGCGC TTGACCGCGC GCGGGCGGCC CGCGGGTCGT ACATCGCGAG GTCGTCGCAC TCGCGCAACC
GB14 primer
-29 CAGAGCCAGG CCCGCTGTGC CCGGAGCTCA TGAGCACCAT GCACCTGCTG ACATTCGCCC TGCTTTTTTC CTGCTCCTTC GCCCGCGCCG CCTGCGACCC
M etSerThrMe tHisLeuLeu ThrPheAlaL euLeuPheSe rCysSerPhe AlaArgAlaA laCysAspPr
72 CAAGATCGTC AACATCGGCG CGGTGCTGAG CACGCGCAAG CATGAACAGA TGTTCCGCGA GGCAGTAAAC CAGGCCAATA AGCGACACGG CTCTTGGAAG
oLysIleVal AsnIleGlyA laValLeuSe rThrArgLys HisGluGlnM etPheArgGl uAlaValAsn GlnAlaAsnL ysArgHisGl ySerTrpLys
172 ATACAGCTCA ACGCCACTTC TGTCACCCAC AAGCCCAACG CCATACAGAT GGCCCTGTCA GTGTGTGAGG ACCTCATCTC TAGCCAGgtg accactactt
IleGlnLeuA snAlaThrSe rValThrHis LysProAsnA laIleGlnMe tAlaLeuSer ValCysGluA spLeuIleSe rSerGln
--- Intron 1 (680 bp to EcoR I) ---



were seen, 238 (strong) and 276 (less intense) nucleotides (Fig. 2a). Riboprobe 1 protected only a 238 nucleotide band. As indicated in Fig. 1, all protected bands are large enough to extend beyond the point of divergence between cDNA and genomic sequences. In primer extension experiments (Fig. 2b), an intense band of 260 nucleotides was extended from primer GB14, indicating a termination of reverse transcription at nucleotide −238, which corresponds exactly to the strongest protected band in RNase protection. Above this band, a weaker band of 300 nucleotides was seen, representing another termination at nucleotide −278 which is close to the less intense protected band of 276 nucleotides. In genomic blots with single digested rat DNA, the PCR-generated probe hybridized to a single band for each digestion (data not shown). This excludes a possible polymorphism of the NMDAR1 gene in this region. Therefore, we believe that the correct 5′ flanking sequence of NMDAR1 gene was isolated and this gene has a cluster of transcription start points located mainly around −276 and −238.

The analysis of putative DNA motifs in the 5′ flanking region did not reveal any canonical TATA/CAAT boxes in proximity to the major transcription start sites. However, this region is highly GC rich (71% from +1 to −356) and contains two SP1 motifs (GGCGGG) at −77 and −295 and one GSG motif (GCG_5CG) at −299 [10,11]. Putative AP-2 elements occur twice following the consensus palindromic sequence, $GCCN_3GGC$, at −1407 and −111 [12]. Another AP-2 at −1522 matched exactly the sequence of the remote AP-2 in human metallothionein II_A [13]. There are two

Fig. 2. Identification of 5′ end of NMDAR1 mRNA. Total RNA of rat brain was extracted using the guanidinium isothiocyanate/CsCl method. (a) RNA mapping. The riboprobes were synthesized with a Promega kit from genomic DNA subcloned in pGEM3Zf(+). The RNase protection was performed using RPA II kit from Ambion with substitution of cloned RNase ONE (Promega) for RNase A + T_1. The dideoxynucleotide sequencing DNA ladder was prepared using a Sequenase 2 kit (USB). The size of protected bands was corrected by the migration of riboprobe 1 which has the sequence shown in Fig. 1 and an additional 10 nucleotides (n.t.) from the vector. Lane 1, riboprobe 1 without digestion; lane 2 and 3, riboprobe 1; lane 4 and 5, riboprobe 2; lane 2 and 4, rat brain RNA (10 μg each); lane 3 and 5, yeast RNA (10 μg each). (b) Primer extension. An antisense oligonucleotide, GB14, nucleotides +22 to −5 (Fig. 1), was kinased at the 5′ end with [γ-^{32}P]ATP (Amersham) and extended by AMV reverse transcriptase using Promega's primer extension system. End-labeled DNA was used as a molecular size marker. Lane 1, yeast RNA (10 μg); lane 2, rat brain RNA (10 μg).

Fig. 1. Nucleotide sequence of the 5′ flanking region and first exon of rat NMDAR1 gene. The transcription start sites are indicated by arrows and the major sites are marked with a solid square underneath. * indicates the divergence point of published cDNA. The antisense riboprobes are illustrated with dashed lines and the primer GB14 is represented by a solid arrow above the sequence. All putative motifs were searched by homology with the compilation in Refs. 10–13 and are underlined. Lower cases represent intronic sequence.

CTF and one CRE motifs far from the transcription start site [11,13]. At −2943 there is a single AP-1 consensus core sequence [11]. In addition to these motifs, a 15-bp sequence of $AC_2TG_4AC_3AC_2$ is directly repeated four times from −2232 to −2173.

In view of the absence of CAAT/TATA boxes, the presence of GC islands, and transcription starting from multiple sites, NMDAR1 may belong to the category of housekeeping genes [14]. It is very interesting that, although housekeeping genes were originally proposed to be constitutively expressed with poor tissue selectivity, these characteristics have been found in a number of neurotransmitter receptor genes, such as human dopamine $D1_A$[15] and $D2_A$ [16], murine $GABA_A$ receptor δ subunit [17], chicken $\alpha 7$ nicotinic acetylcholine receptor [18]. These genes are expressed in clearly tissue-specific patterns and can be up- or down-regulated by certain physiological or pharmacological factors. Therefore, our observations suggest that NMDAR1 may be constitutively expressed with neuronal specificity and with the potential for being regulated.

The GSG motif is recognized by a group of immediate-early genes (IEG), called NGFI-A (or EGR1), B, C and EGR2, 3. Particularly, NGFI-A, B and C can be induced by nerve growth factor (NGF) in PC12 cells [19–21]. The AP-1 consensus sequence is a target of other IEG like c-fos and c-Jun. The occurrence of AP-1 and GSG motifs in the promoter of the NMDAR1 gene suggests that this gene may be regulated by some IEG products. On the other hand, the activation of NMDA receptor has been shown to turn on several IEG, such as c-fos and NGFI-A [22]. Therefore, NMDAR1 gene may have a dual relation to IEG. Since PC12 cells express endogenous NMDAR1 mRNA [23] and bear the potential to respond to NGF stimulation [20], we are examining the possibility that the NMDAR1 gene can be regulated by NGF-induced IEG in PC12 cells.

Cyclic AMP modulates transcription through AP-2 and CRE consensus sequences which are present in the promoter region of the NMDAR1 gene [13]. It has been observed that the activation of NMDA receptor in rat hippocampal CA1 cells results in elevation of cAMP levels [24]. However, whether cAMP is able to further influence the transcription level of NMDAR1 message through this putative motif is still an interesting question and remains to be tested.

References

1 Collingridge, G.L. and Singer, W. (1990) Trends Pharmacol. Sci. 11, 290–196.
2 Meldrum, B. and Garthwaite, J. (1990) Trends Pharmacol. Sci. 11, 379–387.
3 Moriyoshi, K., Masu, M., Ishii, T., Shigemoto, R., Mizuno, N. and Nakanishi, S. (1991) Nature 354, 31–37.
4 Ishii, T., Moriyoshi, K., Sugihara, H., Sakurada, K., Kadotani, H., Yokoi, M., Akazawa, C., Shigemoto, R., Mizuno, N., Masu, M. and Nakanishi, S. (1993) J. Biol. Chem. 268, 2836–2843.
5 Monyer, H., Sprengel, R., Schoepfer, R., Herb, A., Higuchi, M., Lomeli, H., Burnashev, N., Sakmann, B. and Seeburg, P.H. (1992) Science 256, 1217–1221.
6 Kutsuwada, T., Kashiwabuchi, N., Mori, H., Sakimura, K., Kushiya, E., Araki, K., Meguro, H., Masaki, H., Kumanishi, T., Arakawa, M. and Mishina, M. (1992) Nature 358, 36–41.
7 Karp, S.J., Masu, M., Eki, T., Ozawa, K. and Nakanishi, S. (1993) J. Biol. Chem. 268, 3728–3733.
8 Sharp, P.A. (1987) Science 235, 766–771.
9 Meador, J.3., Cannon, B., Cannistraro, V.J. and Kennell, D. (1990) Eur. J. Biochem. 187, 549–553.
10 Faisst, S. and Meyer, S. (1992) Nucleic Acids Res. 20, 3–26.
11 Mitchell, P.J. and Tjian, R. (1989) Science 245, 371–378.
12 Williams, T. and Tjian, R. (1991) Genes Dev. 5, 670–682.
13 Roesler, W.J., Vandenbark, G.R. and Hanson, R.W. (1988) J. Biol. Chem. 263, 9063–9066.
14 Dynan, W.S. (1986) Trends Genet. 2, 196–197.
15 Minowa, T., Minowa, M.T. and Mouradian, M.M. (1992) Biochemistry 31, 8389–8396.
16 Gandelman, K.Y., Harmon, S., Todd, R.D. and O'Malley, K.L. (1991) J. Neurochem. 56, 1024–1029.
17 Sommer, B., Poustka, A., Spurr, N.K. and Seeburg, P.H. (1990) DNA Cell Biol. 9, 561–568.
18 Matter, S.L., Hernandez, M.C., Roztocil, T., Ballivet, M. and Matter, J.M. (1992) EMBO J. 11, 4529–4538.
19 Morgan, J.I. and Curran, T. (1991) Annu. Rev. Neurosci. 14, 421–451.
20 Crosby, S.D., Puetz, J.J., Simburger, K.S., Fahrner, T.J. and Milbrandt, J. (1991) Mol. Cell. Biol. 11, 3835–3841.
21 Patwardhan, S., Gashler, A., Siegel, M.G., Chang, L.C., Joseph, L.J., Shows, T.B., Le, B.M. and Sukhatme, V.P. (1991) Oncogene 6, 917–928.
22 Cole, A.J., Saffen, D.W., Baraban, J.M. and Worley, P.F. (1989) Nature 340, 474–476.
23 Sucher, J.J., Deitcher, D.L., Bading, H., Greenberg, M.E., Cepko, C.L., Lipton, S.A. (1992) Soc. Neurosci. 18, 378.
24 Chetkovich, D.M., Gray, R., Johnston, D. and Sweatt, J.D. (1991) Proc. Natl. Acad. Sci. USA 88, 6467–6471.