

# Gene Structure of the Human Metabotropic Glutamate Receptor 5 and Functional Analysis of Its Multiple Promoters in Neuroblastoma and Astrogloma Cells\*

Received for publication, December 5, 2002, and in revised form, May 9, 2003  
Published, JBC Papers in Press, June 3, 2003, DOI 10.1074/jbc.M212380200

Corrado Corti<sup>‡§</sup>, Richard W. E. Clarkson<sup>¶</sup>, Luca Crepaldi<sup>§</sup>, Cinzia F. Sala<sup>§</sup>, John H. Xuereb<sup>‡</sup>, and Francesco Ferraguti<sup>‡||\*\*</sup>

From the <sup>‡</sup>Cambridge Brain Bank Laboratory, Department of Pathology, University of Cambridge, Level 3 Laboratory Block Addenbrooke's Hospital, Hills Road, CB2 2QQ Cambridge, the <sup>¶</sup>Department of Pathology, University of Cambridge, Tennis Court Road, CB2 1QP Cambridge, United Kingdom, the <sup>§</sup>Department of Pharmacology, University of Innsbruck, Peter Mayr Strasse 1a, A-6020 Innsbruck, Austria, and the <sup>||</sup>Department of Biology, Psychiatry Centre of Excellence in Drug Discovery, GlaxoSmithKline Medicines Research Centre, Via Fleming 4, 37135 Verona, Italy

The metabotropic glutamate receptor 5 (mGluR5) has a discrete tissue expression mainly limited to neural cells. Expression of mGluR5 is developmentally regulated and undergoes dramatic changes in association with neuropathological disorders. We report the complete genomic structure of the mGluR5 gene, which is composed of 11 exons and encompasses ~563 kbp. Three clusters of multiple transcription initiation sites located on three distinct exons (IA, IB, and II), which undergo alternative splicing, have been identified. The 5'-flanking regions of these exons were isolated and, using a luciferase reporter gene assay, shown to possess active promoter elements in SKN-MC neuroblastoma and U178-MG astrogloma cells. Promoter IA was characterized by a CpG island; promoter IB contained a TATA box, and promoter II possessed three active Oct-1-binding sites. Preferential luciferase activity was observed in SKN-MC concomitant with differential DNA binding activity to several responsive elements, including CREB, Oct-1, C/EBP, and Brn-2. Exposure to growth factors produced enhanced expression of promoters IB and II in astrogloma cells and activation of NF- $\kappa$ B. These results suggest that alternative 5'-splicing and usage of multiple promoters may contribute regulatory mechanisms for tissue- and context-specific expression of the mGluR5 gene.

Glutamate, the main excitatory neurotransmitter in the brain, exerts a variety of physiological roles through the activation of multiple receptor proteins (1). These have been categorized into two main classes: ionotropic receptors, which are ligand-gated cation-permeable ion channels, and metabotropic receptors (mGluRs),<sup>1</sup> which can couple to several intracellular

second messengers through heterotrimeric G-proteins. Each of these classes is comprised of several highly homologous receptors each showing a selective distribution in the brain (2). The characteristic expression pattern of glutamate receptors raises some interesting questions regarding the regulatory sequences and molecular mechanisms that determine their cell-specific expression. This information resides in the genomic structure of each receptor gene and in the way it responds to environmental cues. Numerous studies have investigated the genomic structure and genetic regulation of ionotropic glutamate receptor subunits (3), whereas very little is known about mGluRs (4–7). Transcript and protein expression of the mGluR5 subtype have been shown recently to undergo dramatic changes as a consequence of both physiological and pathological conditions. In rodents, this receptor is particularly enriched in telencephalic areas including the isocortex, hippocampus, caudate/putamen, and olfactory bulb (8); and unlike most other mGluRs, it is expressed in both neuronal and glial cells (9). During postnatal development the expression of mGluR5 has been shown to be either up- or down-regulated depending on the brain region (9–12). Exposure of cultured cortical astrocytes to specific growth factors was shown to produce a large up-regulation of mGluR5 expression (5, 13, 14). Recently, we and others have described up-regulation of mGluR5 protein in rodent-reactive astrocytes *in vivo* following neurodegenerative conditions (15–17) in which the release of growth factors has been largely documented (18). Modulation of mGluR5 mRNA expression levels in numerous rat subcortical areas has also been described after challenge with addictive or hallucinating drugs (19–21), thus suggesting that transcriptional modulation of mGluR5 might be linked to drug abuse.

The human mGluR5 gene (*GRM5*) has been mapped to chromosome 11 cytogenetic position 11q14 (22), and although a preliminary exon/intron arrangement of the *GRM5* was reported (22), the 5'- and 3'-ends of the gene were not identified. As a consequence, no evidence of the mechanism and regulation of transcription for *GRM5* have yet been provided.

Here we report the entire genomic structure of *GRM5*, and we demonstrate that its transcriptional activity is driven by at least three distinct promoters. The transcriptionally competent regions for each promoter were determined by means of a luciferase reporter gene in both neuroblastoma and astrogloma cells; the *cis* regulatory elements within these promoter regions were identified based on gel-shift analysis. In addition, we provide evidence that growth factor-mediated transcriptional up-regulation of mGluR5 only affects two *GRM5* promot-

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

\*\* To whom correspondence should be addressed: Dept. of Pharmacology, University of Innsbruck, Peter Mayr Strasse 1a, A-6020 Innsbruck, Austria. Tel.: 43-512-507-3704; Fax: 43-512-507-2783; E-mail: francesco.ferraguti@uibk.ac.at.

<sup>1</sup> The abbreviations used are: mGluR, metabotropic glutamate receptor; RACE, rapid amplification of cDNA terminal ends; RT, reverse transcription; BA, Brodmann area; ADM, astrocyte-defined medium; EGF, epidermal growth factor; bFGF, basic fibroblast growth factor; EMSA, electrophoretic mobility shift assay; *GRM5*, human mGluR5 gene; TIS, transcription initiation sites; UTR, untranslated region; CHO, Chinese hamster ovary cells; TM, transmembrane.

ers. Hence, this study provides the basis for a better understanding of the regulatory mechanisms of *GRM5* transcription in both neuronal and glial cells under physiological and pathological conditions.

#### EXPERIMENTAL PROCEDURES

**5'- and 3'-Rapid Amplification of cDNA Ends (RACE)**—All oligonucleotide primers used in this work were provided by the Human Molecular Genetic Group, Department of Pathology, University of Cambridge, or purchased at MWG-Biotech and are listed in Table I. The transcription initiation and polyadenylation sites were determined by 5'- and 3'-RACE using human hippocampal Marathon-Ready cDNA (Clontech). For the 5'-cDNA RACE, two gene-specific primers (A-Rev and B-Rev), designed according to the public mGluR5 cDNA sequence I21436, were used in combination with adapter primers (AP-1 and AP-2; Clontech). PCRs were carried out for 35 cycles with a denaturing step at 94 °C (1 min), followed by annealing at 56 °C (1 min) and extension at 72 °C (1.5 min). For the 3'-cDNA RACE a similar procedure was performed using two specific gene primers (C-For and D-For) in combination with AP-1 and AP-2 adapter primers. PCR products were subcloned into the pCRII-TOPO vector (Invitrogen), and nucleotide sequence analysis of these and all other cloned DNA fragments were confirmed determining both strands by means of a dye terminator cycle sequencing ready-reaction kit (ABI Prism, PerkinElmer Life Sciences).

**RNAse Protection Assay**—RNAse protection assays were conducted using the RPA III kit (Ambion). In order to identify the 5'-ends of the human mGluR5 mRNAs, two antisense riboprobes were generated by *in vitro* transcription using T3 polymerase: probe IA, spanning both exons IA and II (from bp -335 to -1 upstream from the ATG), and probe IB, encompassing exons IB and II (from -382 to -1). The antisense probes were radiolabeled using [ $\alpha$ -<sup>32</sup>P]UTP (Amersham Biosciences, specific activity 800 Ci/mmol) and purified onto a 5% acrylamide, 8 M urea gel and eluted at 65 °C for 2 h into the kit elution buffer (Ambion). Total RNA was extracted from human hippocampus and temporal cortex (kindly provided by Prof. G. Sperk, Department of Pharmacology, University of Innsbruck, Austria) with Trizol reagent (Invitrogen). The quality of the RNA samples was assessed by means of the Agilent Bioanalyzer (Agilent Technologies). Annealing reactions ( $2 \times 10^5$  cpm of radiolabeled probe per reaction) were carried out in 10  $\mu$ l of hybridization solution at 42 °C overnight, using 50  $\mu$ g of human hippocampal or cortical total RNA, or the same amount of yeast total RNA as a negative control. The annealed products were treated with a 1:100 dilution of RNase A (2.5 units/ml)/RNase T1 (100 units/ml) mix for 30 min at 37 °C, precipitated, and resolved onto a 5% acrylamide, 8 M urea gel. RNA Century Marker Plus Template Set (Ambion) was used to generate a radiolabeled molecular weight ladder. The gel was subsequently dried and exposed to X-Omat AR film (Eastman Kodak Co.).

**Sequence Analyses and Data Base Searches**—Genomic sequences for the human mGluR5 gene were retrieved by BLAST 2.0 (23) analysis using public cDNA (GenBank<sup>TM</sup> accession numbers D28538, I21436, and I21437) and RACE-derived sequences. *GRM5* genomic sequences were identified with the accession numbers NT\_024168, AP000626, and AC026078. Consensus sequences for poly(A)<sup>+</sup> signals and 3'-end cleavage were analyzed using the POLYAH software (BCM Search Launcher, Baylor College of Medicine). Sequence analyses of all PCR products were performed with the alignment program Gel Assemble-GCG Package (GCG, Wisconsin Package version 8, Genetics Corporate Group, Madison, WI) or with FASTA. Transcription factor recognition sites were searched by MatInspector TRANSFAC and TFSEARCH programs (24). Promoter prediction analysis was carried out by means of the Promoter Prediction software by Neural Network (25).

**Screening of the Human Genomic P1-derived Artificial Chromosome (PAC) Library**—The human genomic DNA PAC library RPC11 (kindly provided by Peter de Jong and the Human Genomic Mapping Project, Hinxton, UK) was screened by PCR using a set of oligonucleotide primers laid on exon II of the *GRM5* gene (HR5-For and HR5-Rev). PCRs were performed as described previously (6).

**RT-PCR Analysis of Alternative 5'-Untranslated Exons of *GRM5***—Poly(A)<sup>+</sup> RNA was extracted from human caudate, hippocampus, Brodmann area 7 (BA7), BA11, BA17, cerebellar cortex, and thalamus (provided by the Cambridge Brain Bank Laboratory, University of Cambridge) by means of a Oligotex suspension (Qiagen). Reverse transcription (RT) reactions were performed with 1  $\mu$ g of poly(A)<sup>+</sup> RNA using the First Strand cDNA Synthesis kit (Amersham Biosciences). PCRs were carried out as reported previously (6) using 1  $\mu$ l of template cDNA for each brain area. When nested PCRs were

carried out, 2  $\mu$ l of PCR products were used as template for the successive reactions.

The amplification of exon IA was obtained using the following set of primers: HR5-ExIA.AFor and HR5-Rev for first round, and HR5-ExIA.BFor and II-C.Rev for the second round. Consecutively, a nested PCR was carried out with primers HR5-ExIA.BFor and HR5-ExIA.DRev, which span exon IA only. The amplification of exon IB was obtained from a first round of PCR using the oligonucleotide primers IB-G.For and HR5-Rev. A second set of primers was used for the nested PCR (IB-G.For and II-C.Rev) to amplify a fragment that spanned exon IB and exon II. Amplification of a fragment contained in exon II was carried out with the set of oligonucleotide primers HR5-For and HR5-Rev, which spanned the translation initiation codon. Amplification of  $\beta$ -actin was carried out as positive control.

**Northern Blot Analysis**—Human brain MTN Blot II and V (Clontech), containing ~2  $\mu$ g of poly(A)<sup>+</sup> RNA/line, were hybridized with a mGluR5 probe (bp 143–752) labeled with [<sup>32</sup>P]dCTP by random priming. Hybridizations were performed at 68 °C overnight in ExpressHyb Solution (Clontech). Washes were carried out at 50 °C following the manufacturer's instructions. Blots were then exposed to BioMax MR1 films (Kodak).

**Reporter Gene Constructs**—Reporter gene constructs were prepared by PCR amplification, using either the PAC clone 33C21 or human caudate cDNA as templates. First strand cDNA synthesis was carried out on poly(A)<sup>+</sup> RNA extracted from caudate as described above. Promoter regions were amplified and subcloned into the *NheI/XhoI* sites of the luciferase reporter vector pGL3-Basic (Promega) upstream from the firefly luciferase gene. *GRM5* promoter-reporter gene plasmids used in this study were prepared as shown in Table II.

**Detection of mGluR5 in Cell Lines by RT-PCR**—The human astrogloma cell lines H4, T98G, Tp265, Tp336, Tp356, Tp483, U87-MG, and U178-MG (kindly provided by Prof. V. P. Collins, Department of Pathology, University of Cambridge), the human neuroblastoma cell line SKN-MC, and Chinese hamster ovary (CHO) cells were used to extract poly(A)<sup>+</sup> RNA as described above. Reverse transcription was carried out using the First Strand cDNA synthesis kit as described above. PCRs were then performed using the set of oligonucleotide primers (HR5-For and HR5-Rev) covering a region common to all different mGluR5 mRNA isoforms. Poly(A)<sup>+</sup> RNA from human BA17 was used as a positive control.

**Transient Transfections and Luciferase Assay**—U178-MG astrogloma cells were cultured in Ham's F-12 medium (Invitrogen), and SKN-MC cells were cultured in MEM with the addition of non-essential amino acids (Invitrogen). All media were supplemented with 10% dialyzed fetal bovine serum (Invitrogen), 2 mM glutamine, 100 units/ml penicillin, and 100 units/ml streptomycin (Invitrogen). Cells were maintained in a humidified 5% CO<sub>2</sub> atmosphere at 37 °C.

Reporter gene constructs were transfected into U178-MG, SKN-MC, and CHO cell lines by means of LipofectAMINE (Invitrogen) according to the manufacturer's instructions. Briefly,  $1.5 \times 10^5$  U178-MG,  $3 \times 10^5$  SKN-MC, and  $1.5 \times 10^5$  CHO cells were plated in 6-well tissue culture dishes and transfected the following day, using 0.8  $\mu$ g of DNA and 4  $\mu$ l of LipofectAMINE/well. To normalize for transfection efficiency pRL-RSV (Promega), coding for *Renilla* luciferase, was co-transfected in a 1:50 ratio with each reporter gene construct. Cells were incubated with the DNA-LipofectAMINE complex for 4 h, washed, and maintained in their respective medium for 48 h. The medium was then removed, and cells were lysed by application of 250  $\mu$ l of lysis buffer (Promega) followed by mechanical scraping. firefly and *Renilla* luciferase activity were measured using the Dual Luciferase Reporter Assay System (Promega). In each lysate, the activity of the firefly luciferase was normalized for that of the *Renilla* luciferase. Basal luciferase activity was measured in the extract of cells transfected with the pGL3-Basic (Promega). Background activity was determined in cells transfected with 0.8  $\mu$ g of pGEM-4Z (Promega). Luminescence was measured using a Sirius luminometer (Berthold Detection System).

The effect of growth factors on transcriptional activity of *GRM5* promoters in U178-MG cells was evaluated by exposing the cells, transiently transfected with *GRM5*-reporter genes, to astrocyte-defined medium (ADM) containing growth factors. Twenty four hours after transfection with reporter gene constructs, the medium was removed and substituted with serum free-ADM containing the following components: Ham's F-12 medium, transferrin 50  $\mu$ g/ml, D-biotin 10 ng/ml, sodium selenite 5.2 ng/ml, fibronectin 1.5 ng/ml, heparan sulfate 0.5  $\mu$ g/ml, epidermal growth factor (EGF) 10 ng/ml, basic fibroblast growth factor (bFGF) 5 ng/ml, insulin 5  $\mu$ g/ml. EGF and bFGF were from Invitrogen; all other chemicals were purchased from Sigma. Cells were exposed to

TABLE I

## Oligonucleotide sequences and their use in experimental methods

Only one strand of each complementary pair of oligonucleotides used for EMSA is shown.

Primers	Sequence	Experimental methods
A-Rev	5'-CTACGAACAAGCGATGTCCT-3'	First round 5'-RACE
B-Rev	5'-ATGTCCTACGTTGAGTCGCA-3'	Second round 5'-RACE
C-For	5'-TCAATGGCAGCTGCTGATGG-3'	First round 3'-RACE
D-For	5'-GTCTAAATCATGTTACTTGG-3'	Second round 3'-RACE
HR5-For	5'-ATGGTCTCTGTTGATCC-3'	PAC library Screening
HR5-Rev	5'-GTACCTTCTCATATGTC-3'	PAC library Screening
IA-A.For	5'-AAGCTCACTCACAGCTTCC-3'	PRM region exon IA
IA-B.Rev	5'-GTGACAGCCGACGACG-3'	PRM region exon IA
IA-D.Rev	5'-TTCGTGGTAACTACAGAC-3'	Ex IA (Rev) for 5'-UTR
IA-C.For	5'-GTCTGCCTGCTGTCAC-3'	Ex IA (For) for 5'-UTR
II-C.Rev	5'-TTAGGAAAGGAGTCAAGCC-3'	Ex II Rev for 5'-UTR
IB-A.For	5'-GAGGATCTTCTGCAATTTGG-3'	PRM region exon IB
IB-B.Rev	5'-AGAACAGCCGAGCTGAG-3'	PRM region exon IB
IB-D.Rev	5'-AGTGAGACTCGGTCTAGC-3'	Ex IB (Rev) for 5'-UTR
IB-G.For	5'-GGAATGAGCTTGAGATCATCTTG-3'	Ex IB (For) for 5'-UTR
II-D.For	5'-CTACCTGGCAGAAATCTCC-3'	PRM region exon II
II-E.Rev	5'-GTGGTCATGATCTTAAACC-3'	PRM region exon II
HR5-ExIA.AFor	5'-GCAGCCGTTCTGCCTG-3'	Exon IA amplification
HR5-ExIA.BFor	5'-CGTTCGCTGCTGTCAC-3'	Exon IA amplification
HR5-ExIA.DRev	5'-CGCGCCTGAGCCCTG-3'	Exon IA amplification
HR5-ExII.BFor	5'-GACCTCTTTCAGAATCTCAGC-3'	Exon II amplification
β-Actin.For	5'-TGAAACCTAAGGCCAACCCGTG-3'	β-Actin amplification
β-Actin.Rev	5'-CTCATAGCTCTTCTCCAGG-3'	β-Actin amplification
Brn2 WT	5'-TCTGCTCCTGCATAAATCATAGGGCC-3'	EMSA
CREB WT	5'-AGAGATTGCTGACGTCAGAGAGTAG-3'	EMSA
Oct-1 WT	5'-TGTGAGATGCAAAATCAAGAA-3'	EMSA
CEBP WT	5'-TGCAGATTGCGCAATCTGCA-3'	EMSA
Spl WT	5'-ATTGATGCGGGGGGGGGGAGC-3'	EMSA
Apl WT	5'-CCGTTGATGACTCAGCCGAA-3'	EMSA
NFK-B WT	5'-AGTTGAGGGGACTTCCACGGC-3'	EMSA
Spl	5'-AACCCCGCCGGCGCTGCA-3'	EMSA Exon IA (-127/-108)
Ap-4(1)	5'-CTCCGCAACAGCTTCTCTCT-3'	EMSA Exon IA (-275/-256)
Ap-4(2)	5'-CTGCTCCCGAGCTTCTCT-3'	EMSA Exon IA (-318/-279)
CREB	5'-TCCAGCAGTACGCCCCCTTC-3'	EMSA Exon IA (-394/-375)
Ap-1(1)	5'-CACCTCTGAATCATCTCCGG-3'	EMSA Exon IA (-443/-424)
Oct-1(1)	5'-GTCCGAGAAAATGCAAAATGG-3'	EMSA Exon IA (-335/-316)
Oct-1(2)	5'-AACGGGAATAAATATCAT-3'	EMSA Exon IA (-511/-492)
Oct-1(3)	5'-AGAACCTATTATCTACCTG-3'	EMSA Exon IA (-771/-752)
Brn2	5'-GGTATTGTTAATAAGTAACA-3'	EMSA Exon IA (-493/-474)
CEBP(1)	5'-AAAGTTGCGCAAAAGACTGT-3'	EMSA Exon IB (-110/-91)
CREB(1)	5'-CTGAGATGACCGTAATAA-3'	EMSA Exon IB (-267/-248)
NFK-B	5'-AGACAACCTGGGACTACCAT-3'	EMSA Exon IB (-386/-367)
Oct-1(1)	5'-TCATGGATATCTCCTCTGAG-3'	EMSA Exon IB (-329/-310)
Ap-1(1)	5'-ATTAGGTAGTCAGGGGCA-3'	EMSA Exon IB (-516/-497)
Brn2(1)	5'-GGCAAGGAAAATAAGTGCCA-3'	EMSA Exon IB (-549/-530)
CEBP(2)	5'-GTTAGATGACGAAAGG-3'	EMSA Exon IB (-576/-556)
Brn2(2)	5'-AGTGTATTATTTATTTA-3'	EMSA Exon IB (-664/-645)
Oct-1(2)	5'-AAAGTCATATGCTATTTTTT-3'	EMSA Exon IB (-685/-666)
CEBPB(2)	5'-TGAGTAGGAAAATCCAAACC-3'	EMSA Exon IB (-773/-754)
Ap-1(2)	5'-TGGAAATGAGTGAAGAAAAT-3'	EMSA Exon IB (-779/-760)
Ap-1(3)	5'-CAAGATCCTGACTGTGCTG-3'	EMSA Exon IB (-824/-805)
Spl	5'-TAAGCAAGGCAAGTCAAGA-3'	EMSA Exon II (-236/-217)
Oct-1(1)	5'-TGTTAATATCTGGGTGCG-3'	EMSA Exon II (-400/-381)
Oct-1(2)	5'-CTAATGAAAATTTAACAACA-3'	EMSA Exon II (-454/-444)
Oct-1(3)	5'-TTAATACATGCTCAAGGCTAA-3'	EMSA Exon II (-494/-474)
Oct-1(4)	5'-TAAATTTTATTTTGAATTG-3'	EMSA Exon II (-519/-500)

ADM plus growth factors for 24 h, and then the luciferase activity was assessed as described above.

**Electrophoretic Mobility Shift Assay (EMSA)**—EMSA was performed as described previously (26). Briefly, nuclear extracts were harvested from  $1 \times 10^7$  cells in 20 mM Hepes, 0.4 M NaCl, 1 mM dithiothreitol, 1 mM EDTA, 1 mM EGTA, and 3  $\mu$ g of total protein ( $\sim 2 \mu$ l) incubated with 0.2 mg/ml bovine serum albumin, 50  $\mu$ g/ml poly(dI-dC), 20 mM Hepes, 100 mM NaCl, 1 mM dithiothreitol, 0.05% Nonidet P-40, 0.15 pmol of  $^{32}$ P-labeled oligonucleotide, and 7.5 pmol of cold competitor oligonucleotide in a final volume of 25  $\mu$ l for 30 min at room temperature. Reactions were run on 6% bisacrylamide gels in  $1 \times$  TAE, pH 7.5, at 200 V for 3 h and exposed for autoradiography. Oligonucleotide sequences are listed in Table I.

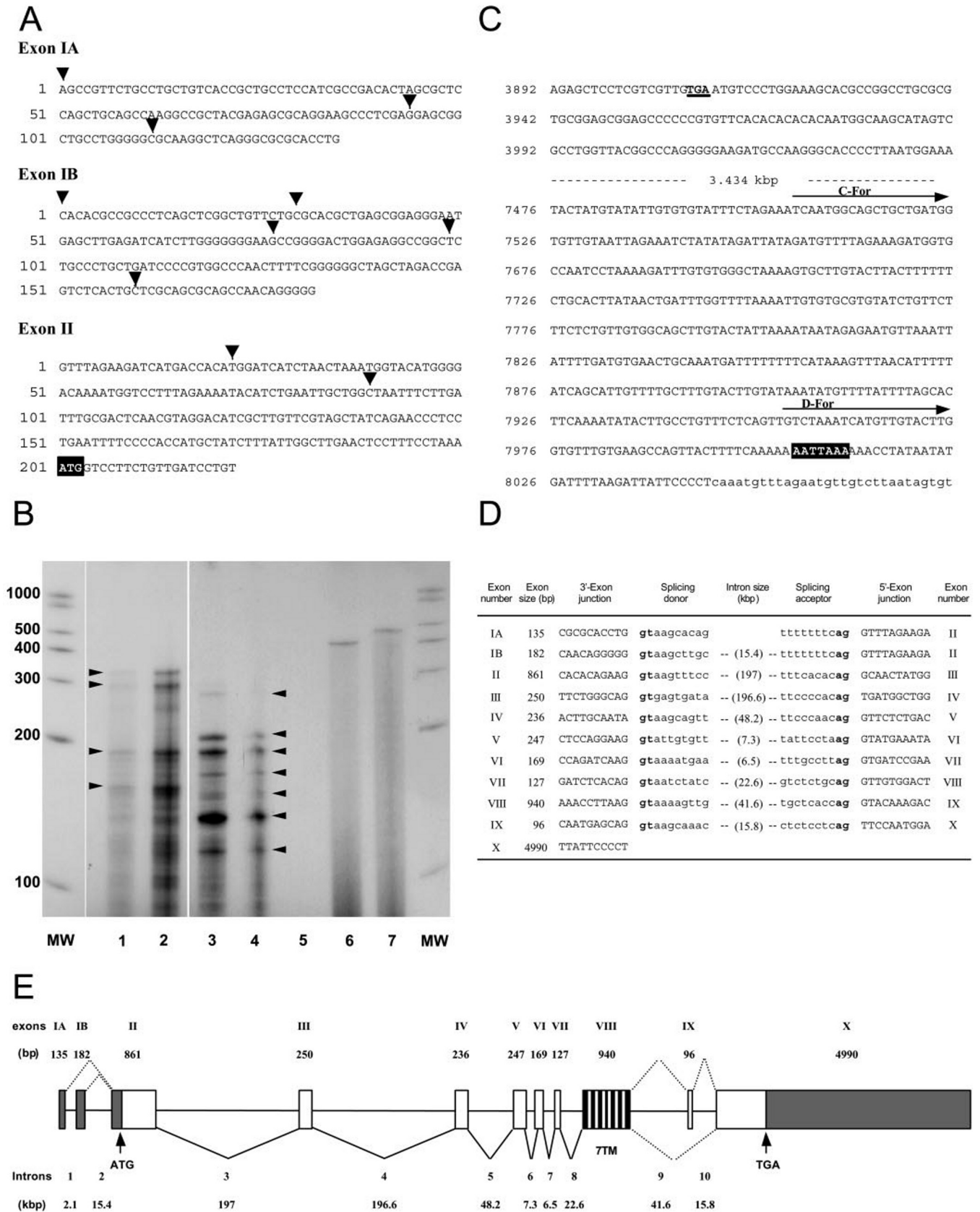
## RESULTS

**Characterization of the Human mGluR5 Gene (GRM5)**—To define the 5' extent of the mGluR5 mRNA, 5'-RACE was performed on human hippocampal poly(A)<sup>+</sup> RNA. Ten different populations of clones were obtained out of 120 5'-RACE clones isolated and subjected to DNA sequencing, which indicated the presence of multiple transcription initiation sites. These were found to belong to three clusters of alternative mRNA forms (Fig. 1A). One form contained a newly identified 5'-end se-

quence, named exon IA, consisting of 135 bp, in which three distinct transcription initiation sites (TISs) were found (Fig. 1A). BLAST analysis identified a human EST sequence (B1826234), obtained from human brain stem mRNA, containing exon IA. This sequence extends the 5'-end of exon IA by an additional 34 nucleotides, suggesting that alternative TISs might be present in other brain areas. The second mRNA form appeared to be generated from 5 distinct TISs within an exon located downstream from exon IA and named exon IB (Fig. 1A). Exon IB is separated from exon IA by an intronic sequence of 2.1 kbp in length, and it extends by 13 additional bps at the 5'-end of the mGluR5 mRNA sequence I21436. The third 5'-alternative mRNA form was found to initiate within the exon containing the codon for the initiation of translation, named exon II, and to be generated by two TISs located 22 and 90 bp downstream from the 5'-intron/exon junction (Fig. 1A). The presence of multiple transcription initiation sites in the three alternatively spliced first exons was confirmed by RNase protection assays carried out on total human RNA extracted from both hippocampus and temporal cortex (Fig. 1B).

The 3'-region of the *GRM5* gene was also mapped by RACE. 3'-RACE confirmed the putative polyadenylation signal AAT-TAAA, located 4076 bp downstream from the translation termination codon, as a transcription termination signal (Fig. 1C). BLAST analysis of human mGluR5 cDNA sequences (from both data bases and sequencing of 5'- and 3'-RACE products) with genomic sequences allowed us to assign exon/intron boundaries within the gene (Fig. 1, D–E). Northern blot analysis of different human adult brain areas, carried out under high stringency conditions, revealed a hybridization band of  $\sim 8.5$  kbp consistent with the size of mGluR5 transcripts deduced from the complete *GRM5* structure reported here. High mGluR5 mRNA expression was observed in the hippocampus, amygdala, isocortex (frontal lobe, temporal lobe, and occipital pole), caudate-putamen, and thalamus, whereas low expression levels were seen in the cerebellum and medulla oblongata (Fig. 2). Only a very weak signal was detected in spinal cord and corpus callosum (Fig. 2). An additional band of  $\sim 2.5$  kbp was observed in most of the brain areas analyzed, which corresponds to a novel metabotropic glutamate receptor 5-related gene (27).

The genomic structures of rat and mouse *Grm5* were also mapped (data not shown) by using publicly available mRNA and genomic sequences, as well as the 5'-end organization of the rat gene reported previously (10). Comparison of the human *GRM5* with that of the rat and mouse showed similar organization and a high degree of homology at the level of their exonic sequences (Fig. 3, A and B). However, in the human gene an exon corresponding to the rat and mouse exon II could not be identified. In order to recognize conserved segments, a comparison between the human (AP000626) and mouse (AC122517) genomic sequences, covering the whole 5'-untranslated region (UTR) and the first translated exon, was performed using the PipMaker Web Server (bio.cse.psu.edu). As shown in Fig. 3, regions of conserved sequences with identity above 75% are present only in correspondence of exon IA and IB and of the first translated exon. A sequence sharing 65% identity with rat exon II was found in the contig NT\_033240, lying within the region spanned by the *GRM5* gene. However, this sequence was found to be located within a LINE2 repetitive element. No human ESTs corresponding to this region could be retrieved by BLAST analysis. In addition, none of the clones obtained from the 5'-RACE appeared to contain this putative exon. RT-PCR analysis from several human brain areas (caudate, hippocampus, BA7, BA11, BA17, cerebellar cortex, and thalamus) also failed to identify transcripts containing a homologue sequence to rat exon II (data not shown).



**FIG. 1. Genomic organization of the human GRM5 gene.** *A*, nucleotide sequence of the three alternative first exons IA, IB, and II as defined by 5'-RACE. The location of the transcription initiation sites is indicated by arrowheads. The codon for translation initiation in exon II is shown on a black background. Nucleotide residues are numbered beginning from the first transcription initiation site. *B*, identification of 5'-ends by RNase protection assay in 50  $\mu$ g of RNA of human adult hippocampus and temporal cortex. Probe IA (lanes 1 and 2) and IB (lanes 3 and 4) both identified several transcription initiation sites in either the hippocampus (lanes 1 and 2) or temporal cortex (lanes 3 and 4). Arrowheads indicate the position of the major protected probe fragments. Lane 5 shows yeast total RNA hybridized with probe IA followed by RNase digestion. Lanes 6 and 7 show the undigested probes IA and IB, respectively. *C*, determination of mGluR5 polyadenylation signal. The nucleotide sequence of the

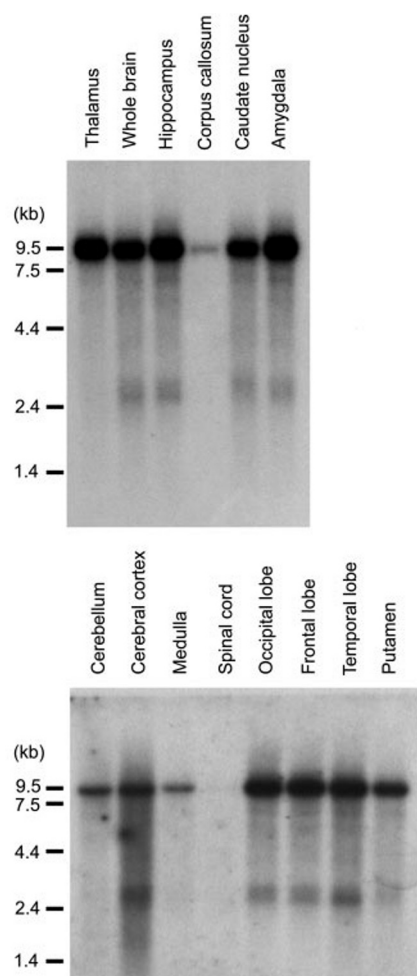


FIG. 2. Northern blot analysis of human mGluR5 mRNA expression in several brain areas. The probe used for Northern hybridization was derived from a cDNA fragment corresponding to nucleotides 143–752 of mGluR5, GenBank™ accession number D28538, common to all splice isoforms. A principal band of 8.5 kbp can be observed in all tissues analyzed, consistent with the size of mGluR5 transcripts deduced from the complete *GRM5* structure. A second band of ~2.5 kbp can be observed in most of the brain areas analyzed, which corresponds to a novel metabotropic glutamate receptor 5-related gene (27).

In conclusion, the complete *GRM5* gene appears to span ~563 kbp in length and to consist of 11 exons and 10 introns (Fig. 1E). Intron/exon splice junctions, as shown in Fig. 1D, conformed to the GT-AG rule of splice donor/acceptor sites. Exon II contains the translation initiation codon, whereas exon X contains the entire 3'-untranslated region. Exon size varies from 96 (exon IX) to 4990 bp (exon X), whereas intron size varies from 2.1 to >196 kbp. Exon IX corresponds to the 96-bp cassette coding for the 32 amino acids inserted in the mGluR5b isoform.

**Expression of the Alternative First Exons of the *GRM5* Gene**—Regional expression of the 5'-alternatively spliced mGluR5 mRNA isoforms containing either exon IA or exon IB was investigated by RT-PCR analysis followed by nested PCRs

on poly(A)<sup>+</sup> RNA extracted from several human brain areas. Amplification of mGluR5 exon IA and of the non-discriminatory exon II was obtained in all brain areas analyzed (Fig. 4). PCR products containing exon IB were detected in BA7, BA17, caudate, hippocampus, thalamus, and cerebellar cortex but not in BA11 (Fig. 4).

**In Silico Analysis of Promoter Regions**—The identification by 5'-RACE analysis of three distinct mGluR5 mRNA populations, which possess different 5'-UTRs due to their initiation from different exons, hinted at the existence of three independent promoters. Therefore, the genomic sequence (~1.0 kbp) upstream from the first TIS of exons IA, IB, and II was analyzed by using the Neural Network Promoter Scan (25). For exon IA a putative promoter region with a score of 0.98, spanning the nucleotide sequence between -128/-80 upstream from the first TIS, was identified (Fig. 5). This region was contained into a CpG island that extended into the exon IA sequence. Several putative transcription factor consensus sequences were also identified in this region (Fig. 5).

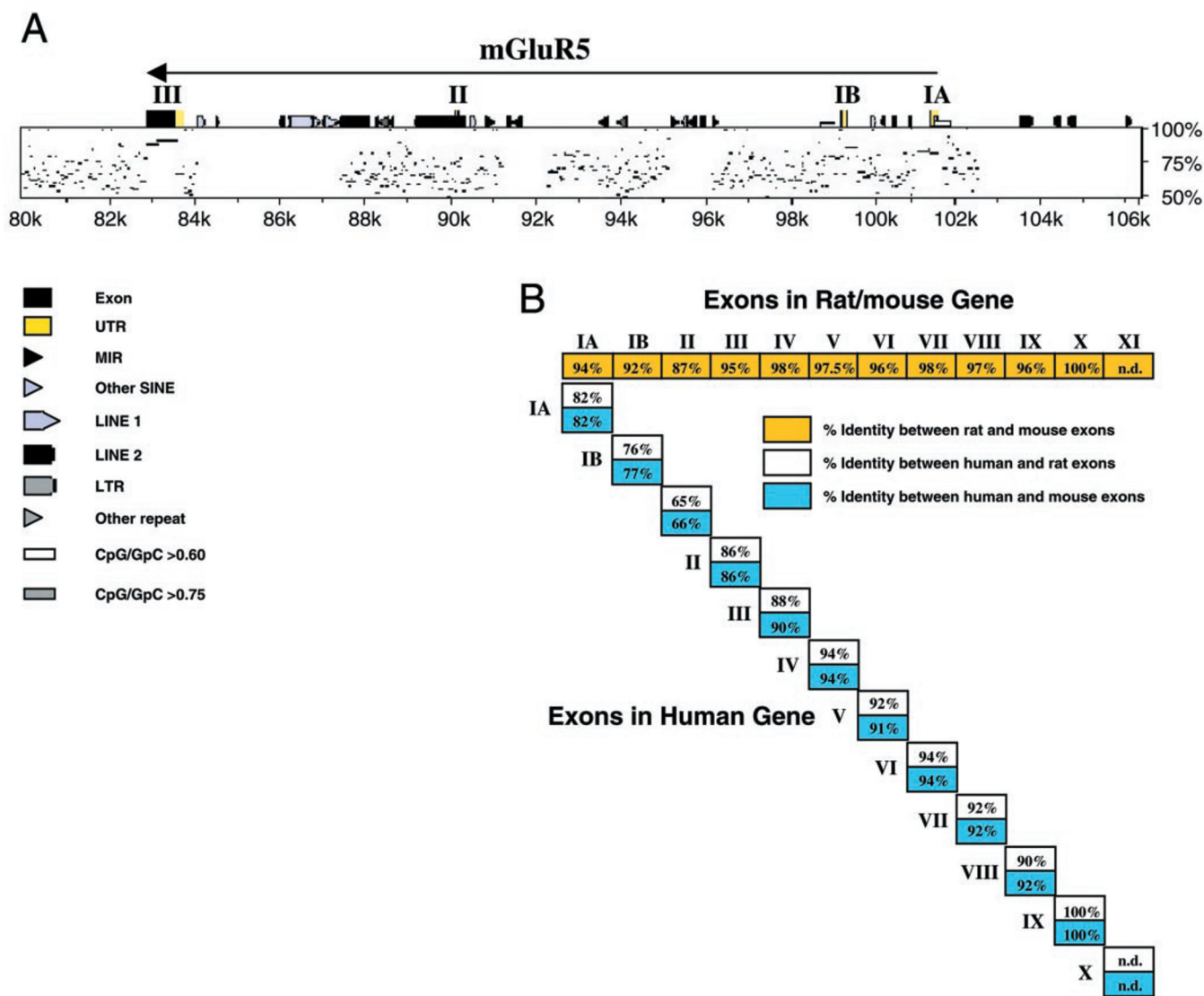
Promoter *in silico* analysis of the exon IB 5'-flanking region identified a promoter element that spanned the nucleotide sequence -96/-47 upstream from the first TIS (score 0.96, Fig. 6). This region contained a consensus sequence for a TATA box at position -84. Several putative transcription factor consensus sequences were also identified, and an additional TATA box was found at -475 (Fig. 6).

A similar analysis of the 5'-flanking region upstream from exon II failed to identify a putative promoter element. However, consensus sequences for several transcriptional regulatory sites, such as C/EBP, Sp1, GATA-1, GATA-2 and Oct-1, were identified in the 5'-intronic region flanking exon II (Fig. 7).

**DNA Binding Analysis of Promoter Regions**—Twenty five putative transcription factor-binding sites, identified by sequence analysis and closely matching the cognitive consensus sequences, were analyzed by electrophoretic mobility shift assay on nuclear extract of SKN-MC and U178-MG cells (Fig. 8). Thirteen functional consensus binding sites were confirmed (boxed sites in Figs. 5–7). The promoter region IA contained functional sites with specificity for AP-1, CREB, and Oct-1. The promoter region IB contained specific CREB, Oct-1, C/EBP, and Brn-2 sites; NF- $\kappa$ B was also identified as a functional site although the signal obtained was weak (data not shown). The promoter region upstream from exon II contained three active Oct-1 consensus sequences. All sites exhibited lower DNA binding affinity than their cognitive consensus sequence (Fig. 8).

**Analysis of Acting Elements within *GRM5* Promoter Regions**—In order to identify the genomic *cis*-elements responsible for transcriptional regulation and cell-specific expression of *GRM5*, the 5'-intronic regions of the gene flanking exon IA, IB, and II were subcloned into the pGL3-basic plasmid. Their ability to drive expression of a reporter gene was assessed in the human neuroblastoma SKN-MC and astrogloma U178-MG cells. These cell lines were selected on the basis of their transcriptional permissibility for the *GRM5* gene. To this end, the expression of mGluR5 mRNA was assessed by RT-PCR in 8 different human astrogloma cell lines and in the neuroblastoma SKN-MC (Fig. 9), using a set of oligonucleotide primers

3'-untranslated region is shown with the termination codon in **boldface** and underlined. The polyadenylation signal was determined by 3'-RACE on human hippocampal cDNA. The position of the gene-specific primers (*C-For* and *D-For*) is indicated by *arrows*. The polyadenylation signal is shown on a *black background*. Capital letters represent coding sequences. The reported sequence is numbered according to the mRNA sequence deposited in the GenBank™ with the accession number D28538. *D*, exon/intron arrangement of the *GRM5* gene and relative sequences at the splice junctions. Splice sites (shown in **boldface**) conform to the gt/ag rule of canonical donor and acceptor splice sites. *E*, schematic diagram of *GRM5* exon/intron organization. Exons are *boxed*, and the untranslated regions in exons IA, IB, II, and X are *shaded*. The seven transmembrane spanning regions (7TM) contained in exon VIII are shown as *vertical stripes*. The position of the translational start (ATG) and stop (TGA) codons are marked by *arrows*. The size of the exons and introns given in base pairs (bp) and kilobase pairs (kbp), respectively, are not drawn to scale.



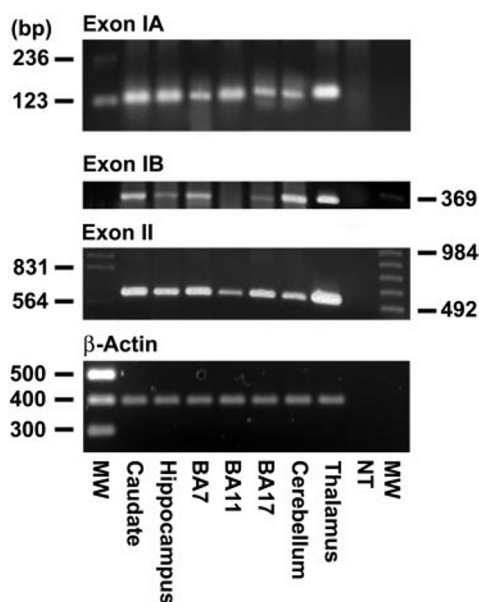
**FIG. 3. Comparison of the genomic structure of *GRM5* between human and rodents (rat and mouse).** A, graphic representation of the comparison between the mouse (GenBank™ accession number AC122517) and human (GenBank™ accession number AP000626) genomic sequences covering the entire 5'-end of the mGluR5 gene as obtained from the PipMaker Web Server. Conserved sequences with identity above 75% are only present in correspondence of exons (yellow) IA, IB, and III. Here, exon numbering corresponds to that of the mouse *Grm5*. B, schematic diagram showing the percentage of identity as follows: between homologous *Grm5* exons in rat and mouse (orange boxes), between homologous *GRM5* (human) and *Grm5* rat exons (white boxes), and between homologous *GRM5* and *Grm5* mouse exons (blue boxes). Because the polyadenylation signal in both the rat and mouse *Grm5* has not been identified, the relative degree of identity between last exons has not been determined (n.d.).

common to all of the different mGluR5 isoforms. CHO cell mRNA was used as a negative control, whereas mRNA extracted from the human brain cortical area BA17 was used as a positive control (Fig. 9). As shown in Table II, *GRM5* promoter-reporter gene constructs containing either the putative promoter regions only or extending into their relative 5'-untranslated sequences up to the translation initiation codon were transiently transfected in U178-MG and SKN-MC cells and tested for their transcriptional activity.

The relative transcriptional activity of the promoter upstream from exon IA (−1128/+21) in SKN-MC cells (mean ± S.E.;  $6.5 \pm 0.2$ -fold over basal luciferase activity) was twice that detected in U178-MG cells ( $3.11 \pm 0.14$ ; Fig. 10A). This showed a possible correlation with the lack of binding to the TIS-proximal Oct-1 site (Oct-1 (1)) in U178-MG cells compared with SKN-MC cells (Fig. 8). Insertion of the 5'-UTR of exon IA and exon II (−1128/+335) had no effect on promoter activity (Fig. 10A). In order to confirm the identity of the *in silico* identified core promoter region, the CpG island was either partially or

completely deleted, and the resultant promoter activity was compared with the full-length form (Fig. 10A). In SKN-MC cells, plasmids −1128/−108 and −1128/−254 showed no transcriptional activity, whereas in U178-MG cells luciferase activity was reduced but not entirely abolished, even when the *in silico* identified promoter region (−128/−80) was entirely removed (plasmid −1128/−254; Fig. 10A).

Transient transfection of the plasmid containing the genomic region upstream from exon IB (−977/+29) resulted in a 19-fold increase of luciferase activity in SKN-MC cells (Fig. 10B). Conversely, this plasmid drove much lower reporter gene activity in U178-MG cells (Fig. 10B). This pronounced cell-specific effect correlated with a loss of binding to C/EBP, Brn-2, and CREB sites in U178-MG compared with SKN-MC cells (Fig. 8). Insertion of the 5'-UTR up to the ATG (plasmid −977/+383) resulted in a markedly reduced luciferase activity in both cell types (Fig. 10B), thus suggesting the presence of negative regulatory elements in the exon IB sequence that repress transcription in both neuroblastoma and astrogloma cells. In order



**FIG. 4. RT-PCR analysis of mGluR5 5'-UTR structures in various adult human brain areas.** RNA (1  $\mu$ g) extracted from post-mortem brain tissue (caudate, hippocampus, Brodmann area 7 (BA7), Brodmann area 11 (BA11), Brodmann area 17 (BA17), thalamus (the specific nucleus was not identified), and cerebellar cortex), obtained from the Cambridge Brain Bank Laboratory (with full ethical approval for collection, storage, and usage from the Cambridge Local Ethical Review Committee), was reverse-transcribed and subjected to PCR with sets of primers to detect exon IA (HR5-ExIA.AFor and HR5-Rev), exon IB (IB-G.For and H5-Rev), or transcripts containing exon II (HR5-For and HR5-Rev). To increase the specificity of the reactions, nested PCRs were carried out for the amplification of exon IA (HR5-ExIA.BFor and HR5-ExIA.DRev) and exon IB (IB-G.For and II-C.Rev). To control for possible contaminations, the template was omitted in one reaction (NT).  $\beta$ -Actin amplification was carried out as positive control for RT-PCR. DNA molecular weight markers (MW) are indicated in bps.

to characterize the functional significance of the core regulatory elements present in the promoter IB region, the TATA box and the region just upstream of the TIS ( $-977/-52$ ) were deleted (Fig. 10B). In SKN-MC cells, loss of this latter sequence resulted in a 4-fold reduction in luciferase activity, whereas deletion of the TATA box (plasmid  $-977/-103$ ) virtually abolished the transcriptional activity ( $1.71 \pm 0.06$ , Fig. 10B). In U178-MG cells, loss of 50 bp upstream from the TIS ( $-977/-52$ ) showed no effect, whereas deletion of the TATA box (plasmid  $-977/-103$ ) drastically reduced the transcription of the reporter gene ( $1.38 \pm 0.28$ ; Fig. 10B).

The putative promoter region upstream from exon II ( $-795/-11$ ) elicited a 14-fold increase over basal luciferase activity in neuroblastoma cells (Fig. 10C). This plasmid exhibited lower transcriptional activity in U178-MG cells ( $5.2 \pm 1.45$ ; Fig. 10C) compared with that obtained in SKN-MC cells, which is similar to the effect observed with the other *GRM5* promoters (Fig. 10C). This also correlated with the loss of transcription factor binding in U178-MG cells to three Oct-1 sites in promoter II (Fig. 8). Inclusion of the 5'-UTR sequence up to the ATG ( $-795/+179$ ) had no effect on transcriptional activity (Fig. 10C). Deletion of the  $-394/-11$  region, which includes the first Oct-1 element and the Sp1 site, completely abolished luciferase reporter gene activity in both SKN-MC and U178-MG cells (Fig. 10C).

To examine further cell-specific transcriptional activity of *GRM5* promoters upstream from exon IA, exon IB, and exon II, CHO cells were transfected with these reporter gene plasmids. The promoter region upstream from exon IA ( $-1128/+21$ ) produced a  $2.6 \pm 0.1$ -fold increase of reporter gene activity. Promoter IB ( $-977/+29$ ) did not elicit a significant luciferase ac-

tivity ( $1.1 \pm 0.03$ ), whereas the promoter region upstream from exon II ( $-795/-11$ ) showed a transcriptional activity ( $4.9 \pm 0.1$ ) comparable with that observed in U178-MG cells.

**Effect of Growth Factors on *GRM5* Transcriptional Activity**—Previous studies (13, 14) have shown that transcripts for mGluR5 are up-regulated in cortical astrocytes challenged with growth factors and cytokines, which included EGF and bFGF. Therefore, we evaluated in transiently transfected astrogloma cells cultured in serum-free astrocyte-defined medium and exposed for 24 h to non-physiological concentrations of EGF and bFGF the transcriptional activity of *GRM5* promoters. Exposure of U178-MG to EGF and bFGF had no effect on promoter IA ( $-1128/+21$  and  $-1128/+335$ ) activity but resulted in a 2-fold up-regulation of promoter IB ( $-977/+29$ ;  $p < 0.01$ , unpaired *t* test) activity (Fig. 11A). The presence of the 5'-UTR region of exon IB ( $-977/+383$ ), which may contain a silencer element, prevented the growth factor-mediated marked increase in transcriptional activity (Fig. 11A). In cells transfected with promoter II constructs, a statistically significant up-regulation of reporter gene activity ( $p < 0.01$ , unpaired *t* test; Fig. 11A) was detected only when the 5'-UTR was present, although a trend toward higher luciferase activity could also be observed with the promoter region  $-795/-11$ . Hence, only promoters IB and II appear to be involved in up-regulating mGluR5 transcripts in response to growth factor exposure. Because EGF and bFGF are both known to induce gene expression through up-regulation of transcription factors such as CREB, NF- $\kappa$ B, and Oct-1 (28), these sites were tested by EMSA on nuclear extracts of U178-MG cells in the presence or absence of both growth factors. Oct-1 and CREB activity was not changed by exposure to growth factor, whereas NF- $\kappa$ B was clearly induced by the application of EGF and bFGF (Fig. 11B).

## DISCUSSION

**Genomic Structure of *GRM5***—We report here that the mature forms of mGluR5 mRNA are generated from three alternatively spliced first exons each containing several distinct TISs. The identification of the TISs and polyadenylation signal allows mapping the entire *GRM5* gene. This gene consists of 11 exons and encompasses  $\sim 563$  kbp. The genomic structure of *GRM5* diverges substantially from that of other GRMs, namely *GRM2*, *GRM3*, and *GRM6*. These latter GRM genes have their TISs located on only one exon and always have an intronic interruption of the transmembrane (TM) domain in the third extracellular loop between the 6th and 7th TM spanning regions (6, 29, 30). Conversely, a preliminary analysis of the genomic structure of the other group I mGluR gene, *GRM1*, showed a very similar organization to that of *GRM5*. Transcripts of *GRM1*, like those of *GRM5*, appear to be generated by different alternatively spliced first exons, and the location of introns within the translated sequence occurs at highly conserved regions between the two genes but never within the TM domain.<sup>2</sup>

The genomic organization of the human, rat, and mouse genes encoding mGluR5 appears to be largely conserved, although one important difference was observed between these species. The rat and mouse *Grm5* contain a constitutive untranslated exon, previously named exon II (10), which could not be detected in the human gene by bioinformatic and transcript analyses. In addition, no human ESTs with homology to rat or mouse exon II could be identified. On the other hand, support for the expression of alternatively spliced transcripts containing exon II in adult mouse brain is provided by the identification of two mouse EST sequences, namely BB625841 (from diencephalon) and BB580390 (from cortex),

<sup>2</sup> F. Ferraguti and C. Corti, unpublished results.

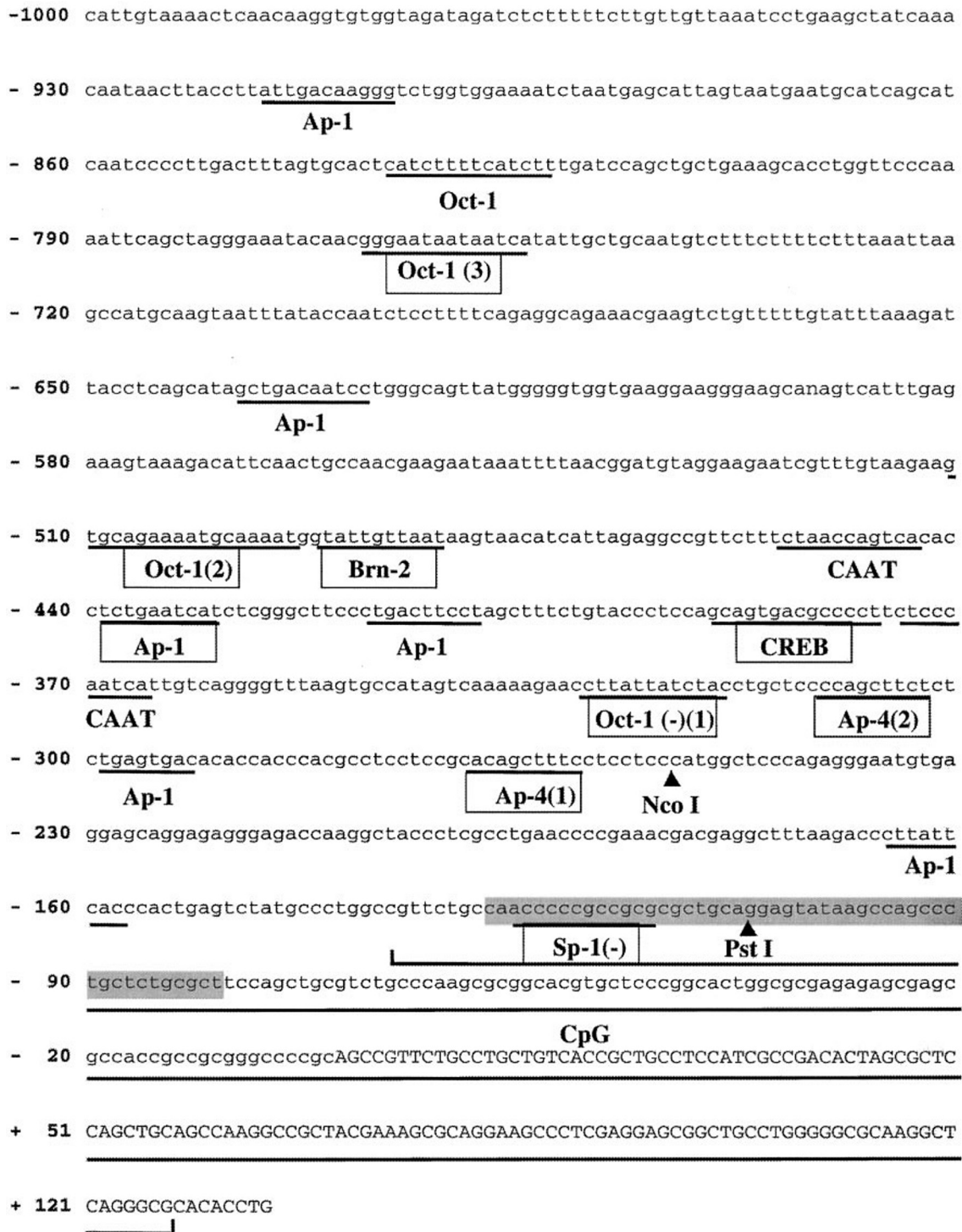


FIG. 5. Nucleotide sequence of the 5' genomic region flanking exon IA of the *GRM5* gene. Nucleotide numbering is relative to the first transcription initiation site. The transcribed sequence is indicated in *capital letters*. Potential regulatory elements, as identified by the transcription factor binding site data bases TRANSFAC and TFSEARCH, are *underlined* and identified by the appropriate names. Different putative binding sites for the same transcription factor tested by EMSA are identified by *numbers in parentheses*. The full list of putative transcription factor binding sites tested is given in Table I. Transcription factor-binding sites confirmed by EMSA in SKN-MC or U178-MG cells are *boxed*. The putative promoter region identified by Neural Network Promoter Scan is shown on a *gray background*. A putative CpG island is indicated by a *square bracket*. Enzyme restriction sites used for sequential deletion of the promoter are indicated by *arrowheads*.

showing >80% identity to rat exon II and flanking, with their 5'-sequences, exon IA and exon IB, respectively. In conclusion, in human brain, unlike in rodents, *GRM5* is alternatively spliced to produce mRNA isoforms containing at their 5'-UTR: (i) exon IA-exon II, (ii) exon IB-exon II, or (iii) just exon II, in which the human exon II is the homologue of exon III in rodents.

In this study, we provide evidence that these alternative first exons are commonly used in the initiation of mGluR5 transcripts in various, but specialized, brain areas. Among the various regions analyzed, exon IB could not be amplified in the prefrontal cortical area BA11, despite the use of a nested PCR protocol. The lack of expression of this isoform in BA11, although it was detected in BA7 and BA17, may indicate in the



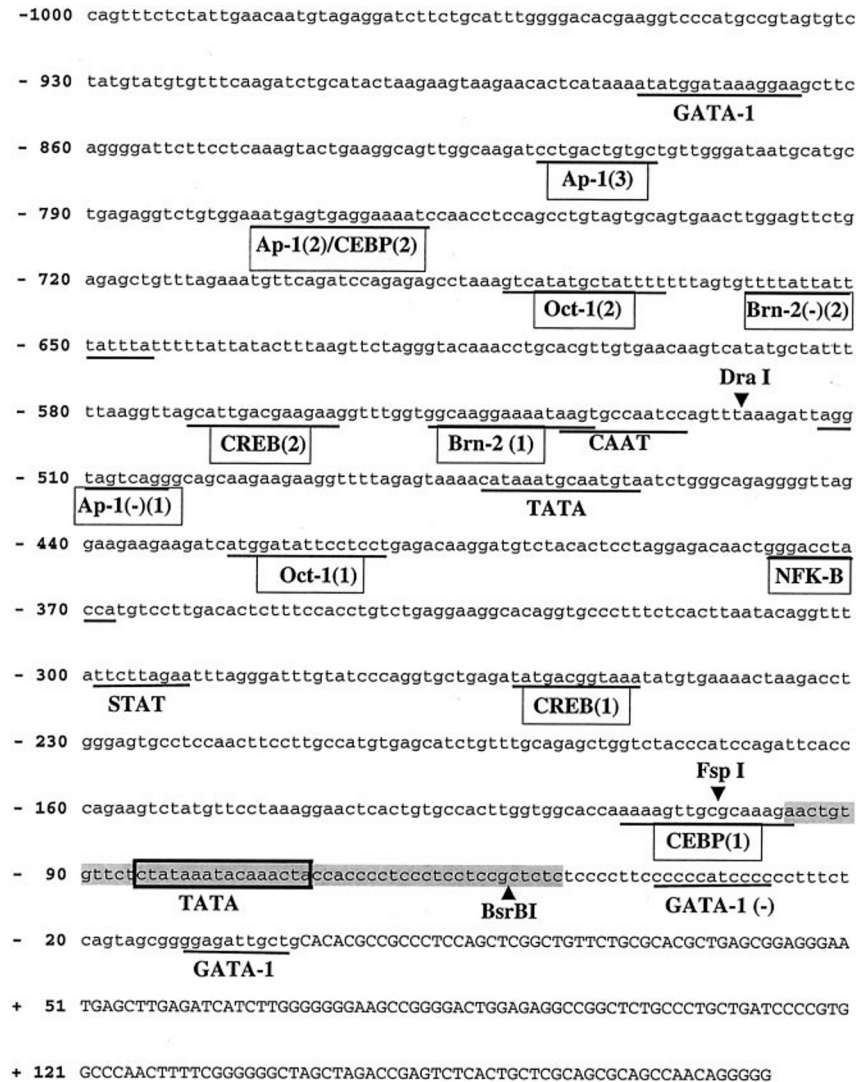


FIG. 6. Nucleotide sequence of the 5' genomic region flanking exon IB of the GRM5 gene. Nucleotide numbering is relative to the first transcription initiation site. The transcribed sequence is indicated in *capital letters*. Potential regulatory elements, as identified by the transcription factor binding site data bases TRANSFAC and TFSEARCH, are *underlined* and identified by the appropriate names. Different putative binding sites for the same transcription factor tested by EMSA are identified by *numbers in parentheses*. The full list of putative transcription factor binding sites tested is given in Table I. Transcription factor-binding sites confirmed by EMSA in SKN-MC or U178-MG cells are *boxed*. The putative promoter region identified by Neural Network Promoter Scan is shown on a *gray background*. A consensus sequence for a TATA box present in the putative promoter region is indicated with a *shaded box*. Enzyme restriction sites used for sequential deletion of the promoter are indicated by *arrowheads*.

specific pattern of cortical connectivity of this area a source for selective promoter regulation of this gene.

**Transcriptional Regulation of GRM5**—Little is known on the *cis* and *trans* regulatory elements that control the transcriptional activity of mGluRs, although many studies have reported modifications in the level of expression of these genes as a consequence of a number of pathophysiological conditions.

To date, the identification of the promoter region of mGluRs is limited to the GRM6 (29) and GRM3 (6) genes, and the characterization of the transcriptional activity has been defined only for the GRM3 promoter (7). One of the primary findings of this study is that transcription of GRM5 is driven by at least three different promoter regions located upstream from three alternatively spliced first exons. The promoter region IA is characterized by a CpG island similarly to the previously described GRM6 promoter region (29). We have discovered a classic TATA element, which is associated with transcriptional initiation in a variety of eukaryotic promoters (31), in the promoter region upstream (−84) from exon IB. Conventionally, a TATA box is present within 25–35 bp upstream from the 5'

site of transcription initiation. However, in many neuron-specific genes the TATA box is located much further away (32–34). Interestingly, Yamaguchi and Nakanishi (10) also reported the presence of a TATA box upstream from the 5'-end of the rat exon IB. The promoter region upstream from exon II does not contain a TATA box, CAAT box, or CpG islands, but it possesses active sequences for Oct-1.

Multiple promoters have been identified for a number of genes, in particular for those that have complex tissue-specific patterns and multiple contexts of transcriptional modulation by different signals (35). For instance, the gene encoding brain-derived neurotrophic factor contains at least four distinct promoters that are responsible for regulation in different context, such as  $[Ca^{2+}]_i$  influx and neuronal activity (36, 37). Other examples of multiple promoters and their complex usage include those for the calcitonin receptor (38), somatostatin 2 receptor (39), estrogen receptor- $\alpha$  (40), and acetylcholinesterase (41). Although expression of mGluR5 occurs primarily in neurons and glial cells of the central nervous system and in peripheral sensory (42) and enteric neurons (43), it has been

-1000 taaaataatttgagatccttggctggcaagttgataacaaagtaagtctcttacacaattcccctaagt

- 930 gatcaaatcactaatttgtcaattctcttttacaattatgctagaatttgccatggaaaatttgggaca

- 860 aaatttcttcttcccaggaatttaaattgaaatagatgatgacattcatagctataactaaatttctacc

- 790 tggcagaatagttccaacatacatccttttagcctttgagccttagtgatgtggaattttttctggat

- 720 cagggctagaggactgattactgtgaattgtattgcaagactctgtacagtttatttaacatgaaacacc  
**CEBP**

- 650 caccttcttcatcttgtttttgttatgtcatccacattaattagacaataacttatttattctctgagaaga  
Oct-1 **Nsp I** **GATA-1** **Oct-1(4)**

- 510 tttgtaattgcataattaattacatgctcaaggctaaagatttcaaaaggaaagcctaatagaaaatttaa  
**Oct-1(3)** **Ssp I** **Oct-1(2)**

- 440 caacagttatccacacttccattaccaagagatcaccaatgtttaatatctgggggtgcatacttcttga  
**Oct-1(1)**

- 370 agttttctgttcatggcgtagttttgcttctaggatgctctgtcatctctgctgagagtacaatacagt

- 300 tctgcccgccacacacaagatttggttttatgtagtgaggtgaacaaaaaatactattaatctaagca  
**Sp1**

- 230 aggcagggtcaagaagctacatcagctgcttcatcaaatcaaaaggaagcagacaaaggaaagtcagaa  
**Ap-4**

- 160 acgtttttgaattatttttcttggttatctgcatcatcaaaatgttttctctgctataatggattatag  
**GATA-3**

- 90 aacttccccaaactagttgcatttatgtttgtaagaccatggtatttttttttccaggtttagaaga  
**STAT** **CEBP** **GATA-3**

- 20 tcatgaccaccatgaccacaTGGATCATCTAACTAAATGGTACATGGGGACAAAATGGTCCTTTAGAAAA

+ 51 TACATCTGAATTGCTGGCTAATTTCTTTGATTGCGACTCAACGTAGGACATCGCTTGTTCGTAGCTATCA

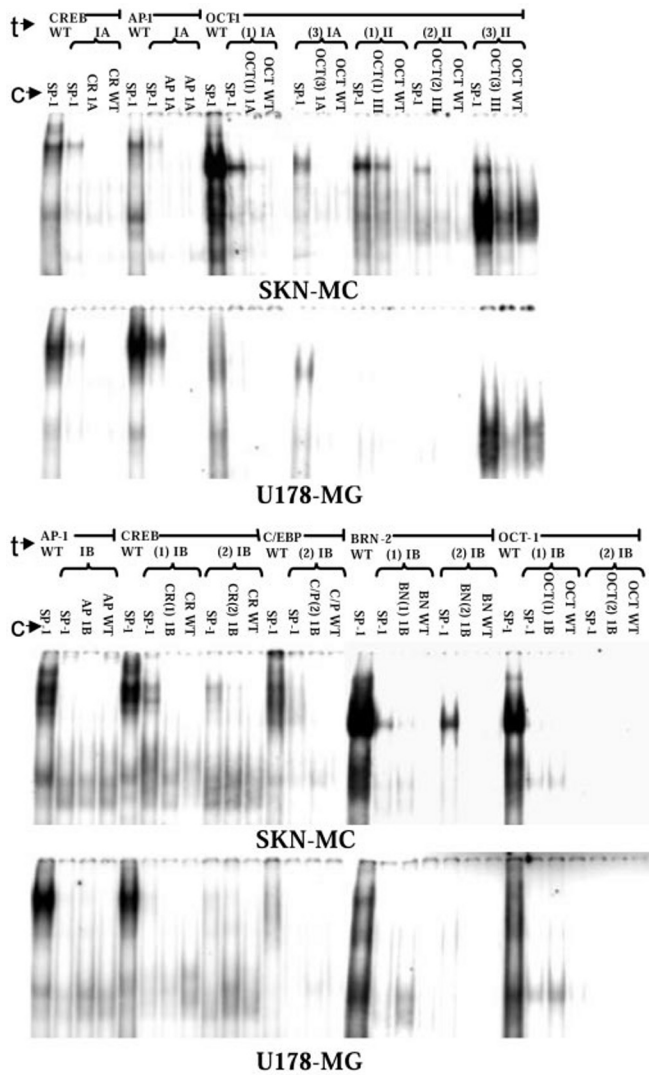
+ 121 GAACCTCCTGAATTTCCCCACCATGCTATCTTTATTTGGCTTGAACCTCTTCCTAAAATGGTCCTTCT

FIG. 7. Nucleotide sequence of the 5' genomic region flanking exon II of the *GRM5* gene. Nucleotide numbering is relative to the first transcription initiation site. The transcribed sequence is indicated in *capital letters*. Potential regulatory elements, as identified by the transcription factor binding site data bases TRANSFAC and TFSEARCH, are *underlined* and identified by the appropriate names. Different putative binding sites for the same transcription factor tested by EMSA are identified by *numbers in parentheses*. The full list of putative transcription factor binding sites tested is given in Table I. Transcription factor-binding sites confirmed by EMSA in SKN-MC or U178-MG cells are *boxed*. Enzyme restriction sites used for sequential deletion of the promoter are indicated by *arrowheads*. The translational start codon is shown in *boldface and underlined*.

reported recently also in pancreatic beta cells (44), pinealocytes (45, 46), and spermatozoa (47). In addition, mGluR5 expression in CA3 hippocampal principal cells is dramatically increased by behavioral acquisition of a classical conditioning (Pavlovian conditioning) procedure (48) and by repeated electroconvulsive shocks (49). The presence of three clusters of TISs and three promoters within the 5'-end of the *GRM5* gene delineates the

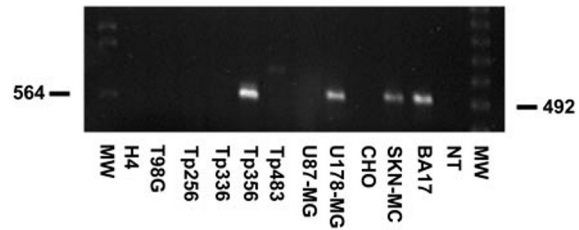
intrinsic complexity of this gene, which may be needed to respond to many different biological contexts.

Functional analysis of *GRM5* promoters demonstrated that all these genomic sequences contain active promoter elements when expressed in SKN-MC and U178-MG cells. Among all three promoters, promoter IA was less active in neuroblastoma cells but displayed comparable transcriptional activity to IB



**FIG. 8. EMSA for putative transcription factor-binding sites in the *GRM5* promoter regions.** Nuclear extracts from SKN-MC and U178-MG cells were incubated with labeled oligonucleotides (*t*) for known transcription factor consensus binding sites (*WT*) and putative binding sites identified from the promoter sequences (see Table I). Each binding reaction was incubated in the presence of 50-fold excess non-specific Sp-1 (*Sp-1*), self, or consensus (*WT*) cold competitor oligonucleotide (*c*). The amount of labeled oligonucleotide used per reaction was not limiting, as defined by the relative quantity of unbound probe. For space considerations, unbound probe is not presented in this figure.

and II promoters in astroglia cells. Promoter IB was found to be the most active in SKN-MC cells and showed, in the 5'-UTR encoded by exon IB, a sequence that appears to exercise repression of transcription. None of the known position-dependent silencer sequences was identified in the *GRM5* exon IB 5'-UTR. Evidence for a negative regulation of transcriptional/translational activity by a silencer located in the 5'-UTR region was also described for the *GRM3* promoter (7). In addition, several ionotropic glutamate receptor subunit gene promoters share this characteristic (3). Some promoters of these genes (*Grin1*, *Grin2b*, *Grin2c*, and *Gria2*) contain the RE1/NRSE silencer element, which contributes to neuronal specificity, or other silencer elements as yet unidentified, which reduce promoter activity, as in the case of the *Gria1* and *Grik5* genes (3). Because mGluR5 is expressed in a restricted subset of neurons, interplay between the proximal silencer and distal enhancers may be required to regulate the expression of this transcript correctly. However, because our read-out assay can be affected



**FIG. 9. Expression of mGluR5 in astroglia and neuroblastoma cell lines.** Reverse-transcribed RNA extracted from the astroglia cell lines H4, T98G, Tp265, Tp336, Tp356, Tp483, U87-MG, and U178-MG and from the neuroblastoma SKN-MC was subjected to PCR using oligonucleotide primers (HR5-For and HR5-Rev) designed within exon II and spanning the translation initiation codon. By using the same set of primers, RT-PCR of RNA extracted from Chinese hamster ovary cells (*CHO*) was used as negative control, whereas RNA from human Brodmann area 17 (*BA17*) was used as positive control. *NT*, template omitted in the reaction. DNA molecular weight markers (*MW*) are indicated in bps.

**TABLE II**  
*GRM5-luciferase reporter gene constructs*

Intronic region upstream of Exon IA	Exon IA	Intronic region upstream of Exon IB	Exon IB	Intronic region upstream of Exon II	Exon II	Amplified from	Restriction site used to join fragments
-1128	+21/Luc					PAC 33C21	
-1128					+336/Luc	PAC 33C21 & Caudate	BerBI
-1128	-108/Luc					PAC 33C21	
-1128	-254/Luc					PAC 33C21	
		-977	+29/Luc			PAC 33C21	
		-977			+363/Luc	PAC 33C21 & Caudate	NaeI
		-977	-52/Luc			PAC 33C21	
		-977	-103/Luc			PAC 33C21	
				-795	-11/Luc	PAC 33C21	
				-795	+179/Luc	PAC 33C21	
				-795	-395/Luc	PAC 33C21	
				-795	-485/Luc	PAC 33C21	

by translational impediments, such as RNA stem-loop structures, we cannot at present rule out that the observed reduction in luciferase activity by the inclusion of the 5'-UTR of exon IB is the result of impaired translation. Elements negatively affecting mRNA translation have been described previously (50) in 5'-UTRs of numerous genes including those of ionotropic glutamate receptor subunits. Promoter II, similarly to the other two *GRM5* promoters, also showed higher reporter gene activity in neuroblastoma cells, which was shown to be dependent on the binding of Oct-1 to sequences in the promoter.

The luciferase activity of two of the *GRM5* promoter regions, namely IA and II, detected in CHO cells, used as negative control, matched the degree of activity observed in astroglia cells. This may represent constitutive expression determined by the core promoter regions. The rate of transcription can be partially or completely inhibited by the action of a silencer as well as increased by the presence of enhancer(s) and by the activation of *trans* regulatory elements, which frequently contribute to determine tissue-specific expression (51, 52).

Overall, reporter gene activity was always higher in neuroblastoma cells, a finding that is supported by the different pattern of transcription factor binding observed in U178-MG and SKN-MC cells. Astroglia cells showed a loss of binding to C/EBP, CREB, and Brn-2 compared with neuroblastoma cells. The higher luciferase activity detected in SKN-MC cells is also consistent with the preferential expression of mGluR5 in neurons rather than astroglial cells in brain slices (9, 53).

Deletion of the region most proximal to the TIS of exon IB resulted in a dramatic reduction of transcriptional activity in SKN-MC cells but not in U178-MG cells. A similar result was obtained for the *GRM3* promoter, in which deletion of a corresponding region also caused a reduction of its transcriptional activity in neuroblastoma but not astroglia cells (7). The regions proximal to the TISs in the *GRM3* and IB *GRM5* promoter are both characterized by the presence of consensus sequences for GATA zinc finger transcription factor-binding

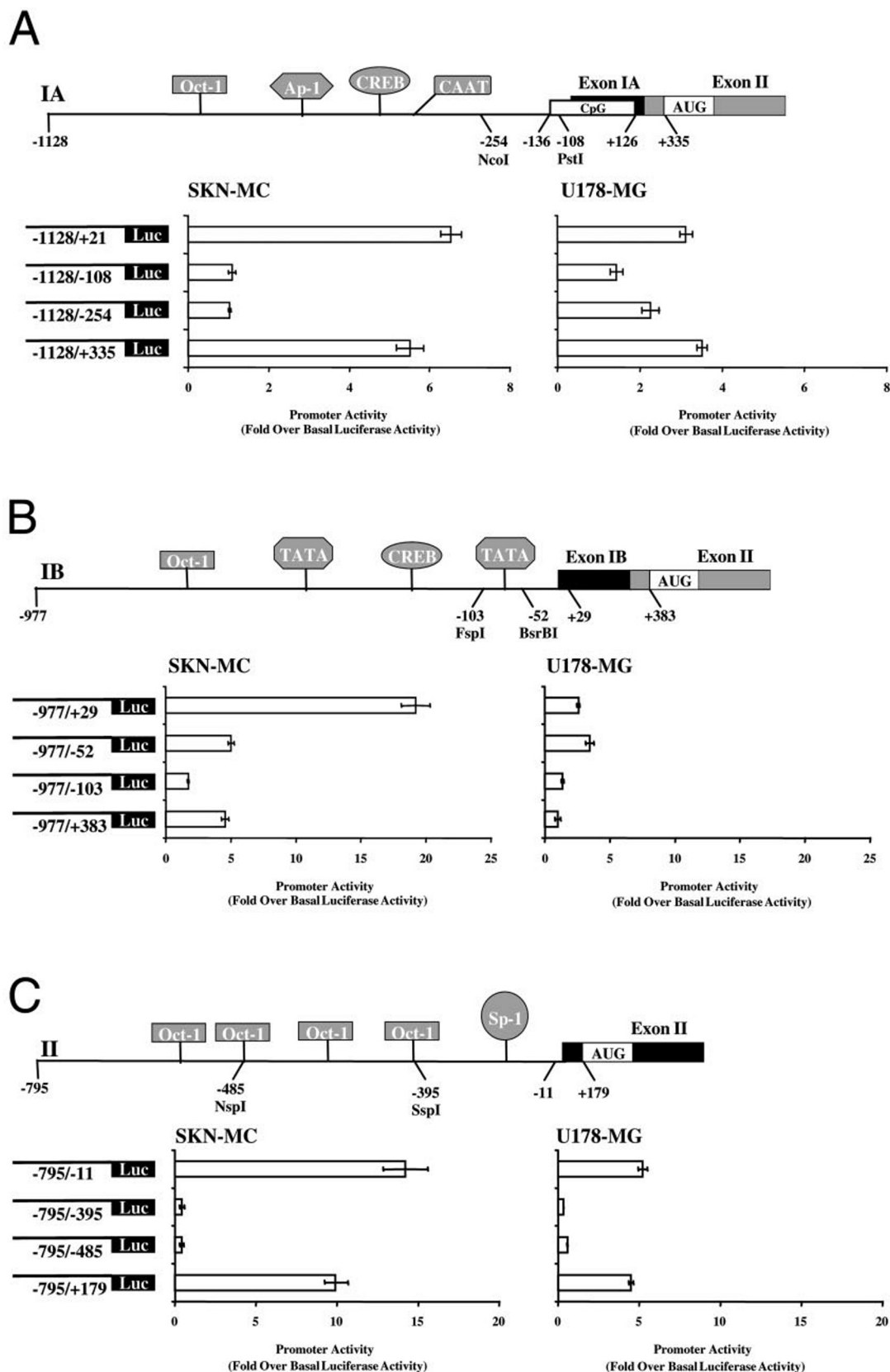


FIG. 10. **Transcriptional activity of *GRM5* alternative promoters.** A, schematic representation of the structure of promoter IA and of *GRM5*-exonIA/luciferase reporter constructs, and corresponding activities after transfection into SKN-MC and U178-MG cells. Luciferase activity of each construct is displayed as fold over basal activity obtained with the pGL3-basic plasmid. Data are normalized for transfection efficiency to *Renilla* luciferase expression driven by the co-transfected pRL-RSV plasmid. B, schematic representation of the structure of promoter IB and of *GRM5*-exonIB/luciferase reporter constructs, and corresponding activities after transfection into SKN-MC and U178-MG cells. C, schematic representation of the structure of promoter II and of *GRM5*-exonII/luciferase reporter constructs, and corresponding activities after transfection into SKN-MC and U178-MG cells. Data represent the mean  $\pm$  S.E. of at least three independent transfections/experiments performed in triplicate.

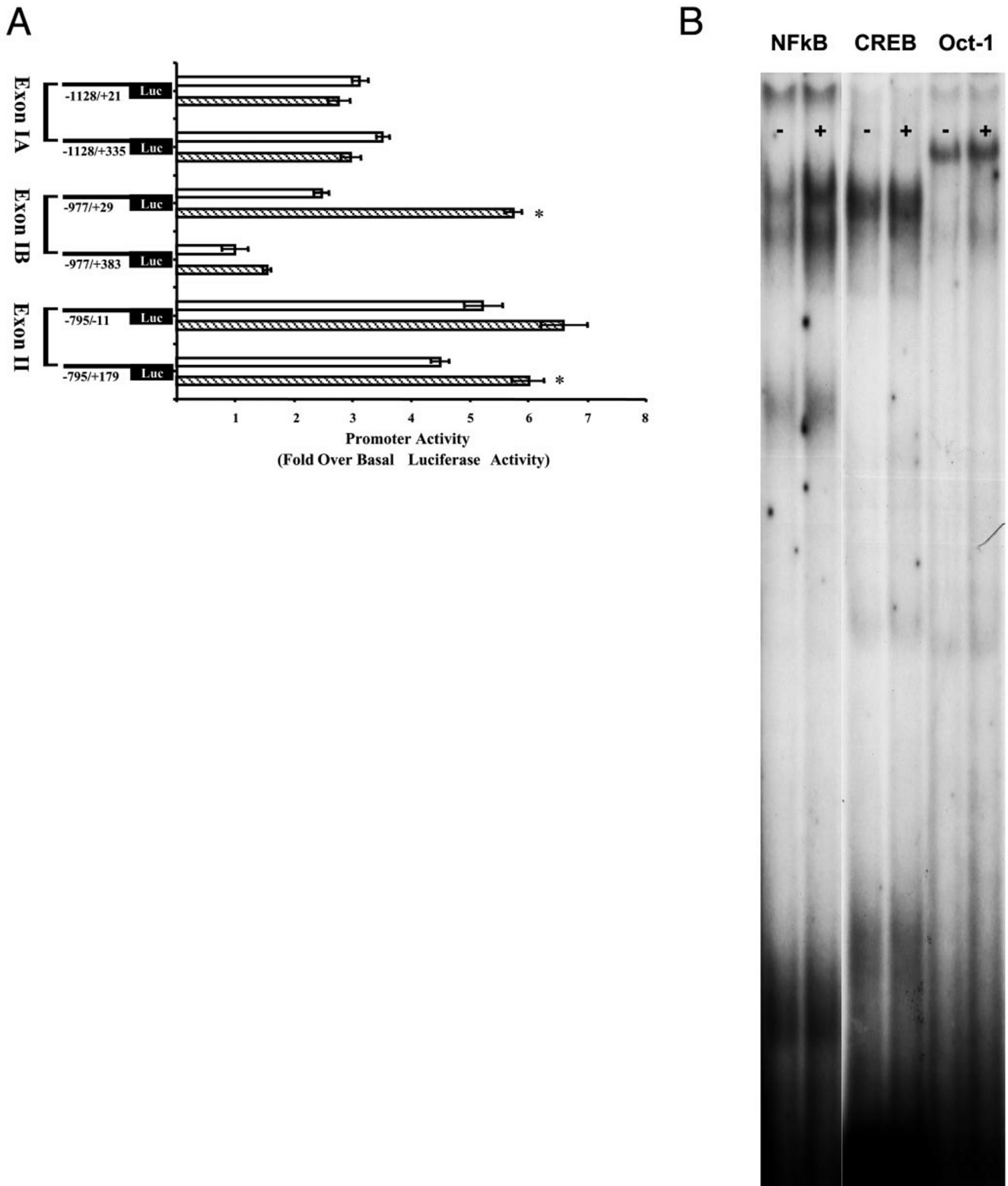


FIG. 11. **Effects of EGF and bFGF on U178-MG cells.** *A*, *GRM5* promoter activity after exposure to the growth factors EGF and bFGF. U178-MG cells transiently transfected with the *GRM5* promoter-reporter gene constructs (indicated on the left side), after being transferred into ADM, were exposed to EGF (10 ng/ml) and bFGF (5 ng/ml) for 24 h and successively harvested. *Open bars* represent *GRM5* promoter activities in complete medium; *hatched bars* represent *GRM5* promoter activities in ADM plus growth factors. Data are normalized for transfection efficiency to *Renilla* luciferase expression driven by the co-transfected pRL-RSV plasmid. Data represent the mean  $\pm$  S.E. of at least three independent transfections/experiments performed in triplicate. \*,  $p < 0.01$  between transfected cells exposed to complete medium or ADM plus growth factors, unpaired  $t$  test. *B*, EMSA for Oct-1, CREB, and NF- $\kappa$ B-binding sites. Nuclear extracts from U178-MG cells cultured in presence (+) or absence (-) of growth factors were incubated with labeled oligonucleotides for wild type transcription factor consensus binding sites.

sites. GATA transcription factors, which constitute a family comprised of several members (54), are known to impart cell-specific transcription (55). Hence, in neuroblastoma cells,

which naturally express some of these transcriptional regulatory factors (56), the activity of mGluR promoters may be selectively enhanced.

The utilization of three alternative promoters that transcribe isoforms with 5'-UTR of different length raises questions concerning the functional significance of these heterogeneous mGluR5 transcripts. One possible explanation for the generation of multiple mGluR5 5'-UTRs may be that mRNAs with different untranslated exons can differ either in their stability or efficacy of translation. For instance, when the translation of mRNAs from the same gene with both long and short 5'-UTRs has been compared, the short 5'-UTRs were usually found more efficiently translated (57).

**Functional Implications**—Transcription of *GRM5* has been shown to undergo developmental regulation (10, 53, 58). In the early post-natal periods both transcripts and proteins of mGluR5 are markedly down-regulated in numerous brain areas, and this was shown to affect both exon IA- and exon IB-containing mRNAs in the rat brain (10). The presence of a CpG island in the IA promoter and of consensus sequences actively binding to Brn-2 in the IB promoter suggests a prominent transcriptional activity for these two promoters during development. Promoters containing CpG islands are thought to be associated with replication origins and with transcriptional activity during embryogenesis (59). Similarly, Brn-2 has been detected in a restricted subset of neurons and shown to be critical to early embryonic development of the central nervous system (60).

Numerous studies have reported changes in the expression of mGluR5 in various pathophysiological conditions. In particular, growth factor-mediated up-regulation of mGluR5 has been described in astrocytes both *in vitro* and *in vivo* (10, 13, 15–17). Here we demonstrate that the growth factors EGF and bFGF modulate *GRM5* transcriptional activity in astroglia cells through promoters IB and II, whereas IA promoter activity remains unaffected. In a previous study (10) on cultured rat cortical astrocytes, application of these growth factors induced up-regulation of mGluR5 transcripts containing both exon IA and exon IB. This difference might be accounted for by the lack of a promoter region upstream from rat exon III and the organizational differences between the human and the rat genes. Alternatively, the effect of growth factor on the transcriptional regulation of *GRM5* may be different in astroglia cells and astrocytes, either because of the nature of the two cultures, a lack of relevant transcription factors in astroglia cells, or because of the experimental paradigm used. In our experimental conditions growth factors were applied for only 24 h, whereas rat cultured astrocytes were usually exposed for at least 3 days (5, 14). The growth factors EGF and bFGF have been shown to activate transcription factors such as CREB, NF- $\kappa$ B, and Oct-1 (28, 61). U178-MG cells showed high levels of CREB and Oct-1 DNA binding activity to the wild type sequences both in the presence or absence of EGF and bFGF. Conversely, EGF and bFGF induced a specific NF- $\kappa$ B activation to a degree that is likely to be of physiological relevance for the cells. Therefore, although relative binding of NF- $\kappa$ B was low, it is possible that this may contribute to the up-regulation of *GRM5* promoter IB activity in response to growth factors, particularly in context of other transcription factors binding to the full-length promoter.

Differential *in vivo* modulation of mGluR5 mRNA expression was shown to take place in several rat brain areas after both acute and chronic challenge with drugs of addiction and/or hallucinating drugs (19–21), which indicate adaptation of mGluR5 transcriptional activity in response to these stimuli. Indeed, the promoter region upstream from exon IA and exon IB show active binding for *trans*-elements such as CREB and Ap-1, transcription factors known to be up-regulated by drugs of abuse (62). Interestingly, mice carrying targeted disruption

of *Grm5* did not self-administer cocaine (63). Taken together, these findings suggest that modifications of mGluR5 transcriptional activity contribute to the occurrence of addiction, withdrawal, and sensitization to drugs of abuse.

In conclusion, this study constitutes the first reported analysis of the 5' genomic region and DNA regulatory sequences directing the transcription of a metabotropic glutamate receptor that is widely expressed in the mammalian nervous system. Multiple promoters and alternative splicing are frequently used mechanisms to create diversity and flexibility in the regulation of gene expression. This complexity would account at least in part for the tissue and developmental differences observed in the level mGluR5 receptor expression, and for the changes to which it undergoes in response to other physiological and pathological cues.

**Acknowledgments**—We thank Prof. Peter Collins (University of Cambridge) for kindly providing the human astroglia cell lines; Prof. Günther Sperk (Dept. of Pharmacology, University of Innsbruck) who generously made available samples of post-mortem brain tissue for Trizol RNA extraction; Dr. Nicola Hall (University of Cambridge) for assistance with the luciferase assay; Dr. Koichi Ichimura (University of Cambridge) and Federico Faggioni (GSK Verona) for excellent technical help. We also thank Dr. Jim Hagan (GlaxoSmithKline) for reading a previous version of the manuscript.

#### REFERENCES

- Nakanishi, S., Nakajima, Y., Masu, M., Ueda, Y., Nakahara, K., Watanabe, D., Yamaguchi, S., Kawabata, S., and Okada, M. (1998) *Brain Res. Brain Res. Rev.* **26**, 230–235
- Shigemoto, R., and Mizuno, N. (2000) *Handb. Chem. Neuroanat.* **18**, 63–98
- Myers, S. J., Dingleline, R., and Borges, K. (1999) *Annu. Rev. Pharmacol. Toxicol.* **39**, 221–241
- Ueda, Y., Iwakabe, H., Masu, M., Suzuki, M., and Nakanishi, S. (1997) *J. Neurosci.* **17**, 3014–3023
- Minoshima, T., and Nakanishi, S. (1999) *J. Biochem. (Tokyo)* **126**, 889–896
- Corti, C., Sala, C. F., Yang, F., Corsi, M., Xuereb, J. H., and Ferraguti, F. (2000) *J. Neurogenet.* **14**, 207–225
- Corti, C., Xuereb, J. H., Corsi, M., and Ferraguti, F. (2001) *Biochem. Biophys. Res. Commun.* **286**, 381–387
- Romano, C., Sesma, M. A., McDonald, C. T., O'Malley, K., Van den Pol, A. N., and Olney, J. W. (1995) *J. Comp. Neurol.* **355**, 455–469
- van den Pol, A. N., Romano, C., and Ghosh, P. (1995) *J. Comp. Neurol.* **362**, 134–150
- Yamaguchi, S., and Nakanishi, S. (1998) *J. Neurochem.* **71**, 60–68
- Ghosh, P. K., Baskaran, N., and van den Pol, A. N. (1997) *Brain Res. Dev. Brain Res.* **102**, 1–12
- Cai, Z., Schools, G. P., and Kimelberg, H. K. (2000) *Glia* **29**, 70–80
- Miller, S., Romano, C., and Cotman, C. W. (1995) *J. Neurosci.* **15**, 6103–6109
- Nakahara, K., Okada, M., and Nakanishi, S. (1997) *J. Neurochem.* **69**, 1467–1475
- Ulas, J., Satou, T., Ivins, K. J., Kesslak, J. P., Cotman, C. W., and Balazs, R. (2000) *Glia* **30**, 352–361
- Ferraguti, F., Corti, C., Valerio, E., Mion, S., and Xuereb, J. (2001) *Exp. Brain Res.* **137**, 1–11
- Aronica, E., van Vliet, E. A., Mayboroda, O. A., Troost, D., da Silva, F. H., and Gorter, J. A. (2000) *Eur. J. Neurosci.* **12**, 2333–2344
- Ridet, J. L., Malhotra, S. K., Privat, A., and Gage, F. H. (1997) *Trends Neurosci.* **20**, 570–577
- Abe, S., Suzuki, T., Ito, T., Yamaguchi, M., Baba, A., Hori, T., Kurita, H., Shiraishi, H., and Okado, N. (2001) *Neuropsychopharmacology* **25**, 173–184
- Freeman, W. M., Brebner, K. A., Lynch, W. J., Robertson, D. J., Roberts, D. C. S., and Vrana, K. E. (2001) *Neuroscience* **108**, 371–380
- Mao, L., and Wang, J. Q. (2001) *Synapse* **41**, 230–240
- Devon, R. S., Anderson, S., Teague, P. W., Muir, W. J., Murray, V., Pelosi, A. J., Blackwood, D. H., and Porteous, D. J. (2001) *Mol. Psychiatry* **6**, 311–314
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25**, 3389–3402
- Heinemeyer, T., Chen, X., Karas, H., Kel, A. E., Kel, O. V., Liebich, I., Meinhardt, T., Reuter, I., Schacherer, F., and Wingender, E. (1999) *Nucleic Acids Res.* **27**, 318–322
- Reese, M. G., and Eeckman, F. H. (1995) in *Proceedings of The Seventh International Genome Sequencing and Analysis Conference, Hilton Head Island, South Carolina, September 16–20, 1995* (Venter, J. C., and Doyle, D., eds) Vol. 1, No. 1, p. 45
- Clarkson, R. W., Heeley, J. L., Chapman, R., Aillet, F., Hay, R. T., Wyllie, A., and Watson, C. J. (2000) *J. Biol. Chem.* **275**, 12737–12742
- Bates, B., Xie, Y., Taylor, N., Johnson, J., Wu, L., Kwak, S., Blatcher, M., Gulukota, K., and Paulsen, J. E. (2002) *Brain Res. Mol. Brain Res.* **109**, 18–33
- Sung, J. Y., Shin, S. W., Ahn, Y. S., and Chung, K. C. (2001) *J. Biol. Chem.* **276**, 13858–13866
- Hashimoto, T., Inazawa, J., Okamoto, N., Tagawa, Y., Bessho, Y., Honda, Y., and Nakanishi, S. (1997) *Eur. J. Neurosci.* **9**, 1226–1235
- Joo, A., Shibata, H., Ninomiya, H., Kawasaki, H., Tashiro, N., and Fukumaki,

- Y. (2001) *Mol. Psychiatry* **6**, 186–192
31. Stargell, L. A., and Struhl, K. (1996) *Trends Genet.* **12**, 311–315
32. Storrington, J. M., Charest, A., Cheng, P., and Albert, P. R. (1999) *J. Neurochem.* **72**, 2238–2247
33. Nedivi, E., Basi, G. S., Akey, I. V., and Skene, J. H. (1992) *J. Neurosci.* **12**, 691–704
34. Mori, N., Stein, R., Sigmund, O., and Anderson, D. J. (1990) *Neuron* **4**, 583–594
35. Ayoubi, T. A., and Van De Ven, W. J. (1996) *FASEB J.* **10**, 453–460
36. Shieh, P. B., Hu, S. C., Bobb, K., Timmusk, T., and Ghosh, A. (1998) *Neuron* **20**, 727–740
37. Timmusk, T., Palm, K., Metsis, M., Reintam, T., Paalme, V., Saarma, M., and Persson, H. (1993) *Neuron* **10**, 475–489
38. Anusaksathien, O., Laplace, C., Li, X., Ren, Y., Peng, L., Goldring, S. R., and Galson, D. L. (2001) *J. Biol. Chem.* **276**, 22663–22674
39. Kraus, J., Woltje, M., Schonwetter, N., and Holtt, V. (1998) *FEBS Lett.* **428**, 165–170
40. Griffin, C., Flouriot, G., Sonntag-Buck, V., Nestor, P., and Gannon, F. (1998) *Endocrinology* **139**, 4614–4625
41. Atanasova, E., Chiappa, S., Wieben, E., and Brimijoin, S. (1999) *J. Biol. Chem.* **274**, 21078–21084
42. Walker, K., Reeve, A., Bowes, M., Winter, J., Wotherspoon, G., Davis, A., Schmid, P., Gasparini, F., Kuhn, R., and Urban, L. (2001) *Neuropharmacology* **40**, 10–19
43. Liu, M., and Kirchgessner, A. L. (2000) *J. Neurosci.* **20**, 3200–3205
44. Brice, N. L., Varadi, A., Ashcroft, S. J., and Molnar, E. (2002) *Diabetologia* **45**, 242–252
45. Frati, C., Marchese, C., Fisichella, G., Copani, A., Nasca, M. R., Storto, M., and Nicoletti, F. (2000) *J. Cell. Physiol.* **183**, 364–372
46. Yatsushiro, S., Yamada, H., Hayashi, M., Tsuboi, S., and Moriyama, Y. (1999) *Neuroreport* **10**, 1599–1603
47. Storto, M., Sallèse, M., Salvatore, L., Poulet, R., Condorelli, D. F., Dell'Albani, P., Marcello, M. F., Romeo, R., Piomboni, P., Barone, N., Nicoletti, F., and De Blasi, A. (2001) *J. Endocrinol.* **170**, 71–78
48. Riedel, G., Casabona, G., Platt, B., Macphail, E. M., and Nicoletti, F. (2000) *Neuropharmacology* **39**, 1943–1951
49. Smialowska, M., Szcwyczyk, B., Branski, P., Wieronska, J. M., Palucha, A., Bajkowska, M., and Pilc, A. (2002) *Neuropharmacology* **42**, 1016–1023
50. Wood, M. W., Van Dongen, H. M., and Van Dongen, A. M. (1996) *J. Biol. Chem.* **271**, 8115–8120
51. Ogbourne, S., and Antalis, T. M. (1998) *Biochem. J.* **331**, 1–14
52. Maniatis, T., Goodbourn, S., and Fischer, J. A. (1987) *Science* **236**, 1237–1245
53. Romano, C., van den Pol, A. N., and O'Malley, K. L. (1996) *J. Comp. Neurol.* **367**, 403–412
54. Molkentin, J. D. (2000) *J. Biol. Chem.* **275**, 38949–38952
55. Orkin, S. H. (1998) *Int. J. Dev. Biol.* **42**, 927–934
56. George, K. M., Leonard, M. W., Roth, M. E., Lieuw, K. H., Kioussis, D., Grosfeld, F., and Engel, J. D. (1994) *Development* **120**, 2673–2686
57. Nielsen, F. C., Gammeltoft, S., and Christiansen, J. (1990) *J. Biol. Chem.* **265**, 13431–13434
58. Minakami, R., Iida, K., Hirakawa, N., and Sugiyama, H. (1995) *J. Neurochem.* **65**, 1536–1542
59. Antequera, F., and Bird, A. (1999) *Curr. Biol.* **9**, R661–R667
60. Sugitani, Y., Nakai, S., Minowa, O., Nishi, M., Jishage, K., Kawano, H., Mori, K., Ogawa, M., and Noda, T. (2002) *Genes Dev.* **16**, 1760–1765
61. Obata, H., Biro, S., Arima, N., Kaieda, H., Kihara, T., Eto, H., Miyata, M., and Tanaka, H. (1996) *Biochem. Biophys. Res. Commun.* **224**, 27–32
62. Nestler, E. J. (2000) *Nat. Genet.* **26**, 277–281
63. Chiamulera, C., Epping-Jordan, M. P., Zocchi, A., Marcon, C., Cottiny, C., Tacconi, S., Corsi, M., Orzi, F., and Conquet, F. (2001) *Nat. Neurosci.* **4**, 873–874