Publications


Banday AR, Azim S, Tabish M. Two Novel N-terminal coding exons of Prkar1b gene of mouse- Identified using novel *in silico* pipeline confirmed by molecular biology techniques. *Communicated*

Banday AR, Azim S, Tabish M. Identification and expression analysis of three novel splice variants of PKA catalytic beta subunit gene in mouse: combinatorial *in silico* and molecular biology approaches. *Communicated*

Banday AR, Azim S, Tabish M. Prediction and identification of new splice variant of Prkaca gene in mouse. *Communicated*

Banday AR, Azim S, Ola M, Tabish M. PKA at interfaces of different disease status: an overview. *Communicated*

Abstracts in international conferences


Alternative promoter usage and differential expression of multiple transcripts of mouse Prkar1a gene

Abdul Rouf Banday · Shafquat Azim · Mohammad Tabish

Received: 14 February 2011 / Accepted: 19 May 2011
© Springer Science+Business Media, LLC. 2011

Abstract Prkar1a gene encodes regulatory type 1 alpha subunit (RIα) of cAMP-dependent protein kinase (PKA) in mouse. The role of this gene has been implicated in Carney complex and many cancer types that suggest its involvement in physiological processes like cell cycle regulation, growth and/or proliferation. We have identified and sequenced partial cDNA clones encoding four alternatively spliced transcripts of mouse Prkar1a gene. These transcripts have alternate 5’ UTR structure which results from splicing of three exons (designated as E1a, E1b, and E1c) to canonical exon 2. The designated transcripts T1, T2, T3, and T4 contain 5’ UTR exons as E1c, E1a + E1b, E1a, and E1b, respectively. The transcript T1 corresponded to earlier reported transcript in GenBank. In silico study of genomic DNA sequence revealed three distinct promoter regions namely, P1, P2, and P3 upstream of the exons E1a, E1b, and E1c, respectively. P1 is non-CpG-related promoter but P2 and P3 are CpG-related promoters; however, all three are TATA less. RT-PCR analysis demonstrated the expression of all four transcripts in late postnatal stages; however, these were differentially regulated in early postnatal stages of 0.5 day, 3 day, and 15 day mice in different tissue types. Variations in expression of Prkar1a gene transcripts suggest their regulation from multiple promoters that respond to a variety of signals arising in or out of the cell in tissue and developmental stage-specific manner.

Keywords Regulatory type 1 alpha subunit · Alternate promoter · Differential expression · Transcription factors · Alternative splicing

Abbreviations
PKA Cyclic AMP-dependent protein kinase
RIα Regulatory type 1 alpha subunit
PN Postnatal
T Transcript
P Promoter
RT-PCR Reverse transcriptase polymerase chain reaction
RACE Rapid amplification of cDNA ends

Introduction
Cyclic AMP-dependent protein kinase (PKA) is the enzyme that phosphorylates the downstream molecules on activation by secondary messenger called cAMP. This phosphorylation plays a key regulatory role in different cellular events like transcription, metabolism, cell cycle progression, and apoptosis. The PKA holoenzymes contain two catalytic (C) subunits bound to homo or heterodimers of either regulatory type I (RI) or regulatory type II (RII) subunit. There are four different R subunits (Iα, Iβ, IIα, and IIβ) and three different C subunits (Cα, Cβ, and Cγ), each encoded by a separate gene. These holoenzymes and individual R subunits show different biochemical properties and also play specific physiological functions.

Among all the subunits of PKA, RIα is studied with immense importance in cancer cases. Increase in the expression of RIα has been found to be associated with...
neoplastic cell growth in case of cancers of breast [1], ovary [2], lung [3], and colon [4]. Downregulation of R1z expression with sequence-specific oligonucleotide results in inhibition of growth and modulation of cAMP signaling in cancer cells like LS-174T, HCT-15, and Colo-205 colon carcinoma cells; A-549 lung carcinoma cells; LNCaP prostate adenocarcinoma cells; Molt-4 leukemia cells; and Jurkat T lymphoma cells. Also inhibition of R1z expression has shown compensatory changes in expression of the isoforms of R and C subunits and cAMP signaling in a cell type-specific manner [5]. Recently, in case of cholangiocarcinoma (CCA), abrogation of PRKAR1A gene (gene symbol for human) expression showed significant CCA cell growth inhibition, oncogenic signaling, and coupled apoptosis induction, suggesting potential of this gene as a drug target for CCA therapy [6]. In Carney complex, mutations leading to the loss of function of one R1z allele give a phenotype displaying variable defects in cellular growth control [7]. In common variable immunodeficiency (CVI) patients, hyperactivation of PKA type I contribute to T cell dysfunction [8]. One of the studies has demonstrated that isoform-specific antagonists of PKA regulatory subunits may be targeted for the novel drug discovery to be used in many cancer types [9].

Therefore, the study of isoforms of PKA regulatory subunits would be significant for designing of isoform-specific drugs. Alternative splicing mechanism is one of the main sources for generation of isoforms. For R1z subunit gene, 11 splice variants in Drosophila [10, 11] and three in human [12] have been studied as per records available in GenBank. To analyze the isoforms resulting from alternative splicing, we have studied the mouse PKA R1z subunit gene as mouse is the best model for correlating gene structure and function in human. In GenBank single transcript was reported for mouse Prkar1a gene with accession number NM_021880 [13]. However, partial sequences of 5′ UTR exons have been previously described [14] which were shown to be different by our methodology using 5′ RACE kit and further supported by EST data. In this article, we have demonstrated multiple transcripts that arise due to alternative splicing and alternative promoter usage in 5′ UTR of Prkar1a gene. We detected three novel transcript variants of Prkar1a gene of mouse, which differ from the previously reported transcript in first exon and studied their expression profile in brain, heart, skeletal muscle, and liver across different postnatal development stages. Three promoter regions were also characterized. The results provide evidence for a complex regulation of Prkar1a expression from three promoters which showed considerable differences for transcription factor-binding sequence motifs. We have also found that the transcripts of R1z subunit gene of human and mouse share less homology for 5′ UTR exons, could be considered as different, but have the same coding exons.

Materials and methods

Materials

Mice (C57BL/6J) were purchased from NII (National Institute of Immunology), New Delhi, India and bred in-house. All the animals were housed according to the Institutional Animal Care and Use Committee and Guidelines of the Committee on Care and Use of Laboratory Animal Resources, National Research Council (Department of Health, Education and Welfare; National Institutes of Health). The total RNA extraction kit was purchased from iNtRON Biotechnology, Inc. (Gyeonggi-do, Korea). M-MuLV-Reverse Transcriptase, PCR nucleotide mix, 1 kb DNA ladders were purchased from Fermentas Life Sciences, USA. The TOPO-TA cloning kit II was obtained from Invitrogen Corp. (Carlsbad, CA). The Plasmid DNA miniprep kit and Qiaquick PCR gel purification kits were purchased from Qiagen, Inc. (Santa Clarita, CA). The 5′ RACE kit was purchased from Clontech (Palo Alto, CA). Primers were custom synthesized from MWG Biotech, Pvt. Ltd., India. All the other chemicals used in the experiments were of molecular biology grade.

Preparation of RNA from different tissues of mouse

Total cellular RNA was isolated from mouse tissues like brain, liver, skeletal muscle, and heart using RNA extraction kit according to manufacturer’s instructions. Finally, the RNA was re-dissolved in diethyl pyrocarbonate-treated water and quantitated spectrophotometrically. Integrity of RNA was confirmed by ethidiumbromide staining after denaturing agarose gel electrophoresis. RNA prepared was either used immediately or stored at −80°C.

Primers

The following oligonucleotide primers were used:

PF1: 5′-GCGACTCGGTCTTTAATCTTCCACTCTA (from exon E1a)
PF2: 5′-AGTCGCCACCTGTACCCGGAATCGT (from exon E1b)
PF3: 5′-CTATCGCAGAGTGGATGAGGCT (from exon E1c)
PR1: 5′-CTCCTTGATCAATCACATAGAAGTTATCC (from exon E6)
PR2: 5′-AATGTCACCTTCTCCTGTTATCAAGTTATCC (from exon E4/E5 junction)
**Reverse transcriptase (RT)-PCR**

To confirm the RACE data, we performed RT-PCR analysis in touchdown mode. Tissue-specific total RNA (2 μg) was primed with gene-specific primer PR1 for the synthesis of cDNA using M-MuLV-Reverse Transcriptase at 42°C for 1 h in a total volume of 20 μl. Subsequently, 2 μl of the single-stranded cDNA thus synthesized was amplified by touchdown PCR with an appropriate set of primers using PCR master mix in a total volume of 50 μl. PCR was performed for 35 cycles as denaturation at 94°C, 35 cycles at 94°C for 30 s, 62°C for 30 s, and 72°C for 30 s were performed, followed by a final extension at 72°C for 7 min in an Amplicator Block (Taurus Scientific, USA). The RACE product was visualized on 1.5% agarose gels. Several bands were excised from the gel, purified, and subcloned into the TOPO-TA vector (a plasmid vector). *E. coli* JM109-competent cells were transformed. Transformed colonies were grown overnight at 37°C. Plasmid DNA was purified using a plasmid purification kit. Plasmids containing the insert were sequenced using an automatic sequencer using either M13 forward or reverse primers.

**Semi-nested RT-PCR**

To further validate the genuine PCR product, semi-nested PCR was performed. After the RT reaction, as described above, PCR amplification was carried out for 30 cycles (as above) using gene-specific internal reverse primer PR2 and first-exon-specific primers as down and upstream primers, respectively. The resulting RT-PCR product (1 μl) was used as a template for further amplification by PCR using the same upstream primer but downstream primer from the junction of exon-4 and exon-5. The resulting semi-nested PCR product (8 μl) was then subjected to electrophoresis on a 1.5% (w/v) agarose gel, stained with ethidium bromide and photographed on a UV transilluminator.

**Subcloning and sequencing**

The PCR amplified products were electrophoresed on 1.5% (w/v) agarose gels. Anticipated ethidium-bromide-stained bands were cut from the gels and DNA was recovered using a Qiaquick PCR gel purification kit. The purified DNA was subcloned using the plasmid cloning vector. *E. coli* JM109-competent cells were transformed. Transformed colonies were grown overnight at 37°C and plasmid DNA was purified for sequencing. Plasmid containing the insert was sequenced using an automatic sequencer using either M13 forward or reverse primers.

**Bioinformatics tools and databases used**

Homology and similarity searches of the nucleotide sequences were performed using the BLASTN non-redundant database (http://www.ncbi.nlm.nih.gov/BLAST). Gene/Exon finding tools were used at HMMgene (http://www.cbs.dtu.dk/services/HMMgene), Genescan (http://genes.mit.edu/GENSCAN.html), GeneSplicer (http://cbcb.umd.edu/software/GeneSplicer/), FGENESH (http://www.softberry.com/softberry.html). Alignment analysis was carried out using the Gene stream Align tool (http://www2.igh.cnrs.fr/bin/align-guess.cgi) and ClustalW tool available at www.ebi.ac.uk/clustalw. Promoter regions were characterized using tools such as TFSEARCH (http://www.cbrc.jp/research/db/TFSEARCH.html), TFBIND (http://tfbind.hgc.jp/), SIGSCAN (http://www-bimas.cit.nih.gov/molbio/signal/), CpG Island Searcher (http://www.uscnorris.com/cpgislands2/cpg.aspx) and CpGPlot (http://www.ebi.ac.uk/Tools/emboss/cpgplot/index.html). Expressed sequence tag search was done at DNA bank of Japan (DDBJ) using BLASTN program (http://blast.ddbj.nig.ac.jp/top-e.html).

**Results**

Four transcripts of mouse Prkar1a gene that differ in the 5' UTR exons

The 5' UTR of mouse regulatory type 1 alpha gene is complex having multiple upstream non-coding exons. Using 5'-RACE followed by cloning and subsequent sequencing, we found four unique clones having the heterogeneous structure of non-coding exons present upstream of first coding exon (designated as E2). Sequence analysis indicated that all of these clones contain a common

**PR3: 5'-AACAGCACATTTCTTTTGATGCGCCTTGCC**

(from exon E4)

5'-Rapid amplification of cDNA ends (RACE)

5'-RACE reactions were performed with 3 μg of total RNA using a 5' RACE kit. The total RNA was annealed with gene-specific reverse primer 5'-AACAGCACATTTCTTTTGATGCGCCTTGCC-3' (PR3) according to the instructions provided with the kit. 5'-RACE followed by PCR was performed with cycling conditions as follows: after an initial denaturation for 3 min at 94°C, 35 cycles at 94°C for 30 s, 62°C for 30 s, and 72°C for 30 s were performed, followed by a final extension at 72°C for 7 min in an Amplicator Block (Taurus Scientific, USA). The RACE product was visualized on 1.5% agarose gels. Several bands were excised from the gel, purified, and subcloned into the TOPO-TA vector (a plasmid vector). *E. coli* JM109-competent cells were transformed. Transformed colonies were grown overnight at 37°C. Plasmid DNA was purified using a plasmid purification kit. Plasmids containing the insert were sequenced using an automatic sequencer using either M13 forward or reverse primers.

**Semi-nested RT-PCR**

To further validate the genuine PCR product, semi-nested PCR was performed. After the RT reaction, as described above, PCR amplification was carried out for 30 cycles (as above) using gene-specific internal reverse primer PR2 and first-exon-specific primers as down and upstream primers, respectively. The resulting RT-PCR product (1 μl) was used as a template for further amplification by PCR using the same upstream primer but downstream primer from the junction of exon-4 and exon-5. The resulting semi-nested PCR product (8 μl) was then subjected to electrophoresis on a 1.5% (w/v) agarose gel, stained with ethidium bromide and photographed on a UV transilluminator.

**Subcloning and sequencing**

The PCR amplified products were electrophoresed on 1.5% (w/v) agarose gels. Anticipated ethidium-bromide-stained bands were cut from the gels and DNA was recovered using a Qiaquick PCR gel purification kit. The purified DNA was subcloned using the plasmid cloning vector. *E. coli* JM109-competent cells were transformed. Transformed colonies were grown overnight at 37°C and plasmid DNA was purified for sequencing. Plasmid containing the insert was sequenced using an automatic sequencer using either M13 forward or reverse primers.

**Bioinformatics tools and databases used**

Homology and similarity searches of the nucleotide sequences were performed using the BLASTN non-redundant database (http://www.ncbi.nlm.nih.gov/BLAST). Gene/Exon finding tools were used at HMMgene (http://www.cbs.dtu.dk/services/HMMgene), Genescan (http://genes.mit.edu/GENSCAN.html), GeneSplicer (http://cbcb.umd.edu/software/GeneSplicer/), FGENESH (http://www.softberry.com/softberry.html). Alignment analysis was carried out using the Gene stream Align tool (http://www2.igh.cnrs.fr/bin/align-guess.cgi) and ClustalW tool available at www.ebi.ac.uk/clustalw. Promoter regions were characterized using tools such as TFSEARCH (http://www.cbrc.jp/research/db/TFSEARCH.html), TFBIND (http://tfbind.hgc.jp/), SIGSCAN (http://www-bimas.cit.nih.gov/molbio/signal/), CpG Island Searcher (http://www.uscnorris.com/cpgislands2/cpg.aspx) and CpGPlot (http://www.ebi.ac.uk/Tools/emboss/cpgplot/index.html). Expressed sequence tag search was done at DNA bank of Japan (DDBJ) using BLASTN program (http://blast.ddbj.nig.ac.jp/top-e.html).

**Results**

Four transcripts of mouse Prkar1a gene that differ in the 5’ UTR exons

The 5’ UTR of mouse regulatory type 1 alpha gene is complex having multiple upstream non-coding exons. Using 5’-RACE followed by cloning and subsequent sequencing, we found four unique clones having the heterogeneous structure of non-coding exons present upstream of first coding exon (designated as E2). Sequence analysis indicated that all of these clones contain a common

**PR3: 5'-AACAGCACATTTCTTTTGATGCGCCTTGCC**

(from exon E4)
sequence from exon E2. Aligning these sequences with the available genomic DNA sequence of the RIα gene downloaded from MGI database (MGI: 104878, Fig. 1a), confirmed that these sequences are part of the genomic DNA of the RIα gene representing its 5′ UTR. Three exons designated as E1a, E1b, and E1c were located upstream of first coding exon E2 according to the organization in DNA sequence of RIα gene (Fig. 1a). The non-coding exon, E1c corresponded to the non-coding exon of previously described transcript available in pubmed with GenBank accession no. NM_021880. Further analysis revealed that the four clones are the result of alternate exons in the 5′ UTR and presence of multiple transcription start sites. However, one of the newly identified transcript contains

**Fig. 1** Schematic diagram of genomic structure of the mouse Prkar1a gene and mRNA transcripts expressed in adult mouse brain, heart, liver, and skeletal muscle as determined by 5′-RACE, RT-PCR and sequencing. a The exon–intron organization on gene and position of gene on chromosome 11. Boxes refer to exons (size of boxes indicate relative sizes of exons) and interconnecting lines as introns. The 5′ UTR of the Prkar1a gene consists of at least three different exon 1 variants: E1a, E1b, and E1c (size and genomic positions relative to ATG at position +1 are shown). The small stanches of nucleotides below indicate the nucleotides present at exon–intron boundaries. b These exon 1 variants are spliced (alone or in combination) to a common exon 2 splice acceptor region (position –6, relative to ATG at position +1) are shown. The small stanches of nucleotides indicate the nucleotides present at exon–intron boundaries. c These exon 1 variants are spliced (alone or in combination) to a common exon 2 splice acceptor region (position –6, relative to ATG at position +1). ORF is located in exon 2. d Four transcripts were detected in brain, heart, liver, and skeletal muscle of mouse: T1, T2, T3, and T4. In transcript isoform T2, exon E1a is spliced with exon E1b and then to the common splice acceptor region of exon 2. d 5′ RACE of mouse mRNA. 5′ Rapid amplification of cDNA ends was independently performed three times using RIα-subunit encoding gene-specific internal primer and RNA isolated from mouse tissues. The products were resolved on a 1.5% agarose gel and stained with ethidium bromide. M indicates DNA size marker, +RT indicates reverse transcriptase was added, and –RT indicates no reverse transcriptase was added to the reaction mixture. Excised products were subcloned into TOPO-TA plasmid and sequenced. Same band pattern was reproduced from all four tissues. e, f RT-PCR and semi-nested PCR amplified products electrophoresed on 1.5% agarose gel using exon-specific forward primers. Lane 1 indicates marker (1 kb ladder). Lanes 2, 3, and 4 represent amplified products obtained using forward primers for exons E1a, E1b, and E1c, respectively. Second lane shows two bands, above one represents transcript T2 and below one transcript T3. Anticipated sizes of the products obtained in RT-PCR are 796 bp and 612 bp (lane 2), 719 bp (lane 3), and 643 bp (lane 4) using exon-specific forward primers PF1, PF2, and PF3, respectively with reverse PR1. Internal reverse primer PR2 was used for semi-nested PCR and bands of anticipated size were obtained. Arrows indicate size of marker and corresponding product size. RT-PCR products were subcloned in TOPO-TA plasmid and bands of anticipated size were obtained. g Alignment of nucleotide sequence of four transcripts. The 5′ UTR exons are indicated with the name of exon. The accession numbers are given at the end of each transcript variant. The sequence shown is up to the sixth exon. Sequence from exon-2 onward is the same in all the transcripts. Translational initiation codon is bold and underlined in exon-2.
Fig. 1 continued
two non-coding exons in which E1a is spliced with E1b and then spliced to the common acceptor site in exon 2 (Fig. 1b, c). The new transcripts were named as T2, T3, and T4 considering published transcript as T1 as seen in 5′ RACE experiment (Fig. 1d).

In order to further validate the presence of these newly identified exons with transcripts, RT-PCR was done using forward primers specific to these 5′ UTR exons and reverse primers from downstream known exons to maximize the length of transcript. Three bands (lanes 2 and 3 marked with arrows) were obtained at indicated sizes using the forward primers FP1 (corresponding to the exon E1a) and FP2 (corresponding to the E1b), with reverse primer PR1 (Fig. 1e—lanes 2 and 3). As a control, we used the forward primer from exon E1c (FP3) with the same reverse primer PR1, and found a single band of the anticipated size (Fig. 1e, lane 4). Two bands were obtained in lane 2 with forward primer FP1 as expected, one corresponding to transcript T2, and other corresponding to transcript T3 validating the presence of splicing event between exons E1a and E1b. The new transcript T2 was the largest one containing both E1a and E1b upstream exons. The seminestem PCR also proved the presence of these transcripts. The expected products were obtained using another reverse primer (PR2) internal to the previous reverse primer (Fig. 1f). The presence of these new transcripts having different exon structures at the 5′ end was confirmed by three independent RT-PCR experiments performed using different RNA preparations, and by cloning and sequencing the products. The cDNA sequences of these new transcripts were submitted in the GenBank and the assigned accession numbers are T2 (HQ412666), T3 (HQ412667), and T4 (HQ412668).

The two new exons E1a and E1b have TG/GT locus for splicing donor sites indicating minor variation from consensus splice donor site, like AG/GT (Fig. 1a). However, the first two and last two bases in the intron between E1a and E1b are GT and AG, respectively. Also, a splice acceptor site at the 5′ end of first coding exon is always CAGA which indicated that the same protein product is generated from all the three transcripts. This is also clear from multiple sequence alignment of four transcripts which showed 100% sequence homology from exon 2 (Fig. 1g).

Mouse Prkar1a gene is expressed from three distinct promoter regions

Using an alternative approach of bioinformatics tools, three different promoter regions were characterized in the 5′ UTR of Prkar1a gene designated as P1, P2, and P3 (Fig. 2a). Earlier it has been established that promoter regions of the RIβ and RIIfβ subunit genes of PKA are TATA less and G/C rich, and initiation of transcription occurs at multiple sites [15, 16]. Using CpG island detection tools such as CpG Island Searcher and CpGPlot, we found that no CpG island is present in or near upstream of exon E1a; however, there is the presence of CpG islands in and near upstream of exon E1b and E1c (Fig. 2b). These CpG islands have an average GC content of about 65–70% and observed/expected CpG ratio >60%. It is noteworthy that all the three promoter regions are TATA less, but have CAAT window upstream of transcription start site that matches with conventional CAAT box. The position of CAAT window varies upstream of transcription start site of exons E1a, E1b, and E1c at positions −346, −414, and −273, respectively. These findings supported the occurrence of transcript T4 having first exon as E1b. An interesting fact is that E1b exon has E-box consensus sequence of CACGTG spanning from position +40 to +45 downstream of transcription start site. This E-box may act either as an enhancer or a suppressor depending on the binding protein. It is well established that Max/Myc dimer binding with E-box activates transcription whereas the Max/Mad dimer suppresses transcription of these genes [17]. The transcription from P2 promoter might also be regulated by Max/Myc dimer.

In order to visualize the transcription regulation from promoter regions, in silico analysis was performed by using tools like TFSEARCH, TFBIND, and SIGSCAN for finding transcription factor-binding consensus sequence motifs (Table 1). Analysis has identified a complex array of recognition sequences that are involved in different cellular processes while responding to variety of signals in and out of a cell. The most important observation was that all three promoters have some binding sites in common, e.g., for Sp1 (steroidogenic factor 1) which is expressed in all mammalian cells and is involved in regulating the transcriptional activity of genes implicated in most cellular processes [18, 19]. CREB (cAMP response element-binding) which decreases or increases transcription [20], GR (glucocorticoid receptor) responds to cortisol, and other glucocorticoids [21], LF-A1 (Liver factor A1) binds promoter regions of several liver-specific genes [22], NF-I (Nuclear factor 1) recognizes the sequence CAAT box and is implicated in eukaryotic transcription, as well as DNA replication [23], etc. However, there are different binding sites for different transcription factors specific to each promoter, e.g., AP-2 (Activator protein-2), which is known to have its roles in vertebrate development, apoptosis, and cell-cycle control [24], can bind to promoter P3; TFαD (Transcription factor II D), is one of several general transcription factors that make up the RNA polymerase II pre-initiation complex [25], can bind to promoter P1; and F2F (Footprint II-binding factor) is believed to be a
transcriptional repressor [26] can bind to promoter P2. These observations clearly describe the diverse regulations of Prkar1a gene transcription.

Differential expression of Prkar1a gene in mouse heart, brain, skeletal muscle, and liver

To give some insights into the regulatory behavior of alternate promoters, the expression patterns of the alternative transcript isoforms were analyzed in four major organs (heart, brain, skeletal muscle, and liver) by RT-PCR at different postnatal development stages of mouse like 0.5 day, 3 day, 15 day, and 60 day named as PN0.5, PN3, PN15, and PN60, respectively. The identity of the resulting amplicons was confirmed by DNA sequencing (not shown). Only few tissues from early and late postnatal stages were studied because it was beyond the scope of the study presented here to study all tissues at different developmental stages of mouse. The RT-PCR analysis showed ubiquitous pattern across these organs in adult stages of development. However, earlier stages of postnatal development showed variation in their expression patterns (Fig. 3). Under our experimental conditions, at PN60 all four transcripts were amplified from the RNAs isolated from heart, brain, skeletal muscle, and liver. The transcript T3 showed the most apparent expression at all postnatal stages which may be because its transcription was under the promoter P1 that possesses binding site for transcription factor TF IID. During the developmental stages, the transcript T1 was found to be very weakly expressed at PN0.5 and PN3 stages; however, in brain and liver at stage PN15, it showed relatively higher expression. In adult, at stage PN60, all four tissues examined, i.e., heart, brain, skeletal muscle, and liver, showed higher expression of transcript T1. Interestingly, at early postnatal stages, transcript T4 was amplified from muscle tissue only. The transcript T2 was not detected from early postnatal stages. However, both T2 and T4 transcripts were detected at late postnatal stage PN60 in all four tissue types examined (Fig. 3). This indicated that the splicing event between exons E1a and E1b occurs mostly in adult mice, the consequence of which is unknown. The absence of CpG Islands in promoter P1 may be the reason behind the ubiquitous type of expression patterns of transcripts expressed from this promoter. In
order to further identify tissue and developmental stage-specific expression, ESTs were searched for all four transcript isoforms using DDBJ blast program. We retrieved several ESTs that showed considerable match with these transcript isoforms. The accession numbers of ESTs, along with their presence in tissue types and respective developmental stages are given in Table 2. The EST data supported our observations of differential and ubiquitous type of Prkar1a gene expression in temporal manner. However, due to lack of EST coverage from all tissues, it was not conclusive to correlate data comparatively. No EST was found against transcript T3 belonging to embryonic developmental stages which may be due to the absence of or very rare expression of T3 transcript during embryonic development. In brain, the role of R1a subunit might be more diverse and complex because all four transcripts were expressed in adult mice brain tissue types (visual cortex, cerebellum, lateral ventricle wall, corpora quadrigemina, etc.). One of the most important observations from EST data was that the splicing event between exon E1a and E1b occurs not only in adult mice, but also in embryo. Among all four transcripts, a large number of ESTs with many tissue types were retrieved for transcript T1 which clearly indicated its abundant expression. These facts evidently suggest complex regulation of transcription of Prkar1a from multiple promoter regions that we envisaged during in silico analysis of these promoter regions.

Discussion

Our results confirmed the presence of four R1α transcript variants in different tissues of adult mouse like brain, liver, heart and skeletal muscle. These transcripts of R1α gene arise due to alternate promoter usage and alternative RNA splicing of three distinct non coding exons in 5’ UTR. The consensus sequences at the splice-donor/acceptor site are believed to be the structural features that determine whether a specific splice site is used although these consensus sequences may not be always present [27]. Analysis of the exon–intron junction region of all new exons indicated structural conservation however, splice donor site varies slightly from consensus where A is replaced by T (TG/GT instead of AG/GT). The presence of more than 75% consensus sequences in the splice-donor/acceptor site provides evidence for the described splicing events. Thus, our data indicated that there are four possible transcript variants for mouse Prkar1a gene including one reported earlier, which are expressed in heart, liver, brain, and skeletal muscle. Interestingly, all four transcript variants differ only in the
non-coding exons. In all variants, coding exons are present unchanged encoding identical proteins. The 5' RACE and bioinformatics analysis has revealed multiple transcription sites controlled by distinct promoter regions. The initiation of transcription at several sites, generating transcripts differing in their 5' UTR could render the expression of this important gene less vulnerable to mutations which could modify its expression. Also the sequence composition in the upstream region for promoters were found TATA less and G/C rich, which is in accordance with the promoter regions of the genes for mouse RIβ, and mouse/rat RIIβ as described previously [15, 16]. However, all three promoters have varying GC content ratio making P1 as non-CpG-related promoter while P2 and P3 as CpG-related promoters. The presence of CpG islands may play an important role in gene regulation, for example, in causing gene silencing or transcriptional repression by hypermethylation of the CpG Island [28, 29]. It is noteworthy that in silico study revealed characteristic signature sequences in all the three promoter regions which are potential binding sites for different transcription factors. This supports the existence of multiple promoters for Prkar1a gene. The existence of multiple promoters may be due to the need for the expression of this gene in many but not all cell types and/or at different times during embryogenesis. The individual promoters can be regulated independently so that the gene can be transcribed in different and overlapping cell types or in response to different signaling systems. In silico analysis of the promoter elements together with expression profiling have shown diagnostic differences; however, in vitro analysis is needed to sort out and establish consensus over the significance of these differences. Another consequence of using different promoters is the addition of a different 5' UTR sequence to the mRNA. The sequences present in E1a, E1b, and E1c could also potentially play a very selective role in translational regulation of Prkar1a gene and/or may contain recognition sites for RNA binding proteins involved in mRNA translocation, degradation, or stability [30]. Particularly, the transcript T2 that contains two 5' exons spliced to each other could play role in post-transcriptional regulation of translation, functions that rely on a combination of primary and secondary structures of the RNA, such as containing a susceptible site for miRNA or hairpin formation [31].

**Table 1** Comparative in silico prediction of different transcription factor binding sites of promoter regions and their position relative to transcription start sites using TFSEARCH, TFBIND, SIGSCAN tools

<table>
<thead>
<tr>
<th>Transcription factor</th>
<th>Target sequence</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP-2</td>
<td>TGGGGA</td>
<td></td>
<td></td>
<td>-92</td>
</tr>
<tr>
<td>ATF</td>
<td>TGACGT</td>
<td></td>
<td></td>
<td>-9</td>
</tr>
<tr>
<td>CAC-binding-pro</td>
<td>GGTGGG</td>
<td>-71,-378</td>
<td></td>
<td>-95,+33</td>
</tr>
<tr>
<td>EBP45</td>
<td>GTTTGCTTT</td>
<td>-198</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GATA-1</td>
<td>TTATATCT</td>
<td>-256,-82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GATA-2</td>
<td>AGATAAA</td>
<td></td>
<td></td>
<td>-193</td>
</tr>
<tr>
<td>CREB</td>
<td>TGACG</td>
<td>-189</td>
<td>-378</td>
<td>-148,-10</td>
</tr>
<tr>
<td>F2F</td>
<td>TAAAAAT</td>
<td></td>
<td></td>
<td>-527</td>
</tr>
<tr>
<td>GR</td>
<td>CAGAG</td>
<td>-229,+83</td>
<td></td>
<td>-25,+19</td>
</tr>
<tr>
<td>GR</td>
<td>TGTCCC</td>
<td></td>
<td></td>
<td>-519</td>
</tr>
<tr>
<td>GR</td>
<td>GACACA</td>
<td>-296</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H4TF-1</td>
<td>GATTTTC</td>
<td></td>
<td>-428,-207</td>
<td></td>
</tr>
<tr>
<td>LF-A1</td>
<td>GGGCA</td>
<td>-237,-29</td>
<td>-350</td>
<td>-177</td>
</tr>
<tr>
<td>MEP-1</td>
<td>TGCGCTC</td>
<td></td>
<td></td>
<td>-103</td>
</tr>
<tr>
<td>NF-1</td>
<td>GCCA</td>
<td></td>
<td></td>
<td>-46</td>
</tr>
<tr>
<td>NF-1</td>
<td>TCCA</td>
<td>-284</td>
<td>+278</td>
<td></td>
</tr>
<tr>
<td>NF-1/L</td>
<td>TGCCA</td>
<td>-108,</td>
<td>-386</td>
<td></td>
</tr>
<tr>
<td>NF-E</td>
<td>CTGTC</td>
<td>-303</td>
<td>-394,+23</td>
<td></td>
</tr>
<tr>
<td>PEA3</td>
<td>AGGAAA</td>
<td>-148</td>
<td></td>
<td>-236</td>
</tr>
<tr>
<td>Sp1</td>
<td>CCGCCC</td>
<td></td>
<td></td>
<td>-41</td>
</tr>
<tr>
<td>Sp1</td>
<td>GGGCGGG</td>
<td>-236</td>
<td></td>
<td>-34</td>
</tr>
<tr>
<td>TFHD</td>
<td>TTCAAA</td>
<td></td>
<td></td>
<td>-364</td>
</tr>
</tbody>
</table>

Signs (-,+) indicates position upstream or downstream of transcription start site.
transcripts like T1, T2, and T4 showed differential expression where as T2 was heavily expressed in adult mice (Fig. 3). In later stages, expression of all four transcripts was ubiquitous that appeared unchanged. These results suggest developmental stage-specific expression of Prkar1a gene. Expressed sequence tags matching these transcripts also supported our RT-PCR analysis. However, EST data revealed that transcript T2 is also expressed in embryo and that transcript T1 is abundantly present at different embryonic developmental stages. These findings

<table>
<thead>
<tr>
<th>Transcripts</th>
<th>EST Accession No.</th>
<th>Tissues (cell types)</th>
<th>Developmental stages</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>CX238268</td>
<td>Lateral ventricle wall</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>CN676072</td>
<td>Embryonic stem (ES) cell</td>
<td>Embryo</td>
</tr>
<tr>
<td></td>
<td>BY350563</td>
<td>Whole joints</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>BY325193</td>
<td>Synovial fibroblasts</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>BY295910</td>
<td>Visual cortex</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>BY171043</td>
<td>Bone marrow macrophage</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>BY099397</td>
<td>Kidney</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>BU922352</td>
<td>Retina</td>
<td>14.5 days embryo</td>
</tr>
<tr>
<td></td>
<td>BE850859</td>
<td>Mammary gland</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>BB665656</td>
<td>Oviduct</td>
<td>2 days pregnant adult female</td>
</tr>
<tr>
<td></td>
<td>BB654792</td>
<td>Whole embryo</td>
<td>9 days</td>
</tr>
<tr>
<td></td>
<td>CJ182565</td>
<td>Rathke’s pouches</td>
<td>14.5 days embryo</td>
</tr>
<tr>
<td></td>
<td>CA894890</td>
<td>Neural stem cell</td>
<td>Differentiated</td>
</tr>
<tr>
<td></td>
<td>BY782095</td>
<td>Whole body</td>
<td>17.5 days embryo</td>
</tr>
<tr>
<td></td>
<td>BY747531</td>
<td>Thymus thymic cells</td>
<td>2 days neonate</td>
</tr>
<tr>
<td></td>
<td>BY736666</td>
<td>Blastocyst</td>
<td>Embryo</td>
</tr>
<tr>
<td></td>
<td>BY726571</td>
<td>Corpora quadrigemina</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>BY705244</td>
<td>Liver</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>BB610002</td>
<td>Lung</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>BY350231</td>
<td>Whole joints</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>BY333602</td>
<td>Synovial fibroblasts</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>CN663942</td>
<td>Whole embryo</td>
<td>13.5 days embryo</td>
</tr>
<tr>
<td></td>
<td>CA540754</td>
<td>Whole Embryo</td>
<td>7.5 days embryo</td>
</tr>
<tr>
<td></td>
<td>BY229339</td>
<td>Olfactory sensory neurons</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>BY064378</td>
<td>Kidney</td>
<td>17 days embryo</td>
</tr>
<tr>
<td></td>
<td>BB850806</td>
<td>Inner ear</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>BB864160</td>
<td>B lymphocyte (B cells)</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>BY036616</td>
<td>Liver</td>
<td>14 days embryo</td>
</tr>
<tr>
<td></td>
<td>CF167794</td>
<td>Germ Cell</td>
<td>Embryonic</td>
</tr>
<tr>
<td>T2</td>
<td>BB610165</td>
<td>Cerebellum</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>BY132531</td>
<td>Whole body</td>
<td>17.5 days embryo</td>
</tr>
<tr>
<td></td>
<td>BY153772</td>
<td>Pancreas islet cells</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>BB567873</td>
<td>Head</td>
<td>12 days embryo</td>
</tr>
<tr>
<td>T3</td>
<td>DN174114</td>
<td>Lateral Ventricle Wall</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>BY129491</td>
<td>Brain</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>BB610165</td>
<td>Cerebellum</td>
<td>Adult</td>
</tr>
<tr>
<td>T4</td>
<td>BU709493</td>
<td>Whole brain</td>
<td>Develop. stage: embryo 15.5</td>
</tr>
<tr>
<td></td>
<td>CB194067</td>
<td>Embryonic limb, maxilla and mandible (pooled)</td>
<td>Embryo day 12.5, 13.5, 14.5, and 15.5</td>
</tr>
<tr>
<td></td>
<td>BI657582</td>
<td>Tumor, gross tissue</td>
<td>5 months</td>
</tr>
<tr>
<td></td>
<td>BY267453</td>
<td>Visual cortex</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>BY026492</td>
<td>Mammary gland</td>
<td>Adult</td>
</tr>
<tr>
<td></td>
<td>BY153772</td>
<td>Pancreas islet cells</td>
<td>Adult</td>
</tr>
</tbody>
</table>
further support the importance of R1α gene for progression of embryogenesis in mouse.

In human R1α, gene initiation at different transcription start sites lead to three distinct, alternatively spliced transcripts coding for identical proteins [12]. Multiple sequence alignment of human and mouse R1α gene transcripts showed that 5' UTR exons of mouse and human are different and share less homology (Fig. 4). However, all of these transcripts have almost 100% homology from exon 2. This observation indicated two possibilities that either 5' UTR of human and mouse R1α gene are less conserved in sequence or that the 5' UTR exons that we reported here have not been identified in human and vice versa. The pairwise alignment of human R1α 5' UTR exons with genomic DNA of mouse and that of mouse R1α 5' UTR exons with genomic DNA of human using NCBI BLAST program showed no significant match at exon level. This fact excluded the possibility of unreported transcripts in human of the same type of sequence that we found in mouse. Also, the 5' UTR of human and mouse R1α gene showed 68% homology at nucleotide level which indicates less conservation of nucleotide sequence. It was also found that no splicing event is reported between human R1α 5' UTR exons similar to that of mouse where such an event produces transcript T2.

These data suggest complex regulation of R1α subunit gene transcription in mouse and human to produce the same protein product. The study of transcripts belonging to mouse and human R1α gene will be helpful while investigation of diseases and other functional aspects in mouse as model, in particular to those cases which pertain to regulation of this gene. This study seeks to understand the mechanism and significance of splicing event that gives rise to transcript T2 in mouse and also whether such splicing event is present in human R1α subunit gene.

Acknowledgments The authors are thankful to the Department of Biotechnology (DBT), the Ministry of Science and Technology, New Delhi, India for providing generous funding to project No. BT/PR-10917/BID/07/254/2008. The necessary facility provided by A.M.U., Aligarh, India is also thankfully acknowledged.

References

gene encoding the protein kinase A type I-alpha regulatory subunit in patients with the Carney complex. Nat Genet 26:89–92
Differentially expressed three non-coding alternate exons at 5# UTR of regulatory type I beta subunit gene of mouse

Abdul Rouf Banday, Shafquat Azim & Mohammad Tabish
Your article is protected by copyright and all rights are held exclusively by Springer Science+Business Media B.V.. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author’s version for posting to your own website or your institution’s repository. You may further deposit the accepted author’s version on a funder’s repository at a funder’s request, provided it is not made publicly available until 12 months after publication.
Differentially expressed three non-coding alternate exons at 5′ UTR of regulatory type I beta subunit gene of mouse

Abdul Rouf Banday · Shafquat Azim · Mohammad Tabish

Received: 23 May 2011 / Accepted: 17 June 2011
© Springer Science+Business Media B.V. 2011

Abstract Prkar1b gene encodes regulatory type I, beta subunit (RIβ) of cAMP dependent protein kinase A in mouse. Among the various isoforms of regulatory and catalytic subunits that comprise mammalian PKA, RIβ subunit is considered to be one of the important subunits for neuronal functions. This is involved in multiple forms of synaptic plasticity, and influences memory and learning by maintaining hippocampal long-term potentiation (LTP). Deficient expression of this gene has been implicated in autoimmune disease systemic lupus erythematosus (SLE). We have identified two novel non-coding exons of the Prkar1b gene (designated as exon 1A and exon 1B), which are spliced to the canonical exon 2 and constitute the 5′ untranslated region giving rise to three alternative transcript isoforms. We have also confirmed the expression of the previously known first exon (designated as exon 1C) with known transcript published earlier. The transcripts containing exons 1A, 1B and 1C are differentially regulated during the development and tissue types. In silico study of more than 20 kb nucleotide sequence upstream of known translational initiation codon revealed three distinct promoter regions named as PA, PB, and PC upstream of the exon 1A, exon 1B and exon 1C respectively. PB is non-CpG related promoter but PA and PC are CpG related promoters, however all three promoters are TATA less. Further analysis showed that these promoters possess potential signature sequences for common as well as different transcription factors suggesting complex regulation of Prkar1b gene.

Keywords Alternative splicing · Alternate promoter · Differential expression · Exon · Regulatory type I beta subunit · Transcription factors · RT-PCR

Introduction

In eukaryotic cells, cAMP regulates many different cellular functions. Its effects are in most cases mediated by cAMP-dependent protein kinase A (PKA). PKA enzyme is tetramer that contains two catalytic (C) subunits bound to homo- or heterodimers of either regulatory type I (RI) or regulatory type II (RII) subunit. At least four regulatory (RIα, RIβ, RIIα, and RIIβ) and three catalytic (Cα, Cβ and Cγ) subunits have been characterized in mammals. The activation of PKA is controlled by the regulatory (R) subunits of the holoenzyme. Different anatomical distribution of the regulatory isoforms may contribute to determine the specificity of diverse effects of cAMP [1].

Signaling events mediated by PKA are critical for many neuronal functions [2]. PKA is involved in multiple forms of synaptic plasticity [3–5], synaptic transmission, memory and learning [6]. Among the various isoforms of regulatory and catalytic subunits that comprise mammalian PKA, the regulatory type I beta (RIβ) subunit is considered to be one of the important subunits for neuronal functions. RIβ subunit has been found to be important for mossy fiber long-term potentiation (LTP) [7], hippocampal long-term depression and depotentiation [8, 9], inflammation and sensitization of primary afferent nociceptors [10]. In autoimmune disease systemic lupus erythematosus (SLE) 25% of subjects show no detectable RIβ protein in their T cells [11]. One of the studies has demonstrated that isoform specific antagonists of PKA R subunits may be targeted for the novel drug discovery to be used in many cancer types...
[12]. Alternative splicing generates variants that are tissue or developmental stage specific and can be differentially affected by pharmacological agents to either up regulate or down regulate the expression as per the requirements of disease condition [13]. But this is possible only when spliced variants are known for any given gene. Owing to this we proposed that there could be alternatively spliced variants in case of mouse R1β gene which can become therapeutic targets.

The PKA R1β subunit encoding gene named Prkar1b is present on chromosome 5 in mouse. To identify the possible spliced variants arising from R1β subunit encoding gene, we have studied the mouse Prkar1b gene as mouse is the best model for correlating gene structure and function with human. In GenBank, single transcript was reported for mouse. Previously, the promoter region of mouse Prkar1b gene has been described as TATA-less flanked by 1.5 kilobases of 5′-upstream sequence and 1.8-kilobase intron [14]. Here, we have demonstrated three transcripts that arise due to alternative splicing in pre-mRNA and alternative promoter usage in 5′ UTR of Prkar1b gene. We have also located two more promoter regions which lead to the expression of three novel transcript variants of Prkar1b gene. These variants differ from the previously reported transcript with respect to first exon. One of the new exons was found to be part of previously described promoter (PDP) region [14]. We have also studied their expression pattern in brain, heart, skeletal muscle and liver across different postnatal development stages.

Materials and methods

Materials

Mice (C57BL/6 J) were obtained from NII (National Institute of Immunology), New Delhi, India and bred in house. All animals were housed according to the Institutional Animal Care and Use Committee and Guidelines of the Committee on Care and Use of Laboratory Animal Resources, National Research Council (Department of Health, Education and Welfare; National Institutes of Health). The total RNA extraction kit was purchased from iNtRON Biotechnology, Inc. (Gyeonggi-do, Korea). M-MuLV-Reverse Transcriptase, PCR nucleotide mix, 1 kb DNA ladders were purchased from Fermentas Life Sciences, USA. The TOPO-TA cloning kit was obtained from Invitrogen Corp. (Carlsbad, CA). The Plasmid DNA miniprep kit and Qiaquick PCR gel purification kits were purchased from Qiagen, Inc. (Santa Clarita, CA). The 5′ RACE kit was purchased from Clontech (Palo Alto, CA). Primers were custom synthesized from Sigma-Aldrich chemicals Pvt. Ltd. (Bangalore, India). All other chemicals used in the experiments were of molecular biology grade.

Preparation of RNA from different tissues of mouse

Total cellular RNA was isolated from various mouse tissues using RNA extraction kit according to manufacturer’s instructions. Finally, the RNA was re-dissolved in diethyl pyrocarbonate-treated water and quantitated spectrophotometrically. Integrity of RNA was confirmed by ethidiumbromide staining after denaturing agarose gel electrophoresis. RNA prepared was either used immediately or stored at −80°C.

Primers

Primer positions are indicated in parenthesis. The following oligonucleotide primers were used:

PF1: 5′-CGAA CCT TGA GGG CGA GCC AGA AGA (from exon 1A)
PF2: 5′-ATGGAC TCT GTT TCC ACA GGA GCA GAT G (from exon 1B)
PF3: 5′-CTA AGA GAA GCA AGT GTA ATT GGC TAG CG (from exon 1C)
PR1: 5′-CAC ATA TAC ATC TAC TTC TCC TTG GGA (from junction of exons 6 and 7)
PR2: 5′-ATG TCA CTT CTC TCG TTG TCG TCC AGG (from junction of exons 4 and 5)
PR3: 5′-GTC ATG GTC TTA TAG TCC TTG GGA ATG (from exon 4)

5′ Rapid amplification of cDNA ends (RACE)

5′ RACE reactions were performed with 3 μg of total RNA isolated from mouse brain using a 5′ RACE kit. The total RNA was annealed with gene-specific reverse primer PR3 according to the instructions provided with the kit. 5′ RACE products were PCR amplified with cycling conditions as follows: after an initial denaturation for 3 min at 94°C, 30 cycles were performed with denaturation at 94°C for 30 s, annealing at 66°C for 30 s, extension at 72°C for 30 s followed by a final extension at 72°C for 8 min in an Amplicator Block (Taurus Scientific, USA). The RACE product was visualized on 1.3% agarose gels. Several bands were excised from the gel, purified and subcloned into the TOPO vector. E. coli JM109-competent cells were transformed. Transformed colonies were grown overnight at 37°C. Plasmid DNA was purified using a plasmid purification kit. Plasmids containing the insert were sequenced using an automatic sequencer using either M13 forward or reverse primers.
Reverse transcriptase (RT)-PCR

To confirm the RACE data, we performed RT–PCR analysis in touch down mode. Tissue specific total RNA (2 μg) was primed with oligo-dT for the synthesis of cDNA using M-MuLV-Reverse Transcriptase at 42°C for 1 h in a total volume of 20 μl. Subsequently, 2 μl of the single-stranded cDNA thus synthesized was amplified by touchdown PCR with an appropriate set of primers using the PCR master mix in a total volume of 50 μl. PCR was performed for 30 cycles as; initial denaturation at 94°C for 4 min followed by denaturation 93°C, 30 s; annealing 70°C, 30 s with a decrease of 0.3°C per cycle, extension 72°C for 45 s, with final extension at 72°C for 8 min. The PCR products obtained were subjected to electrophoresis on a 1.3% (w/v) agarose gel, stained with ethidium bromide and photographed on a UV transilluminator.

Semi-nested RT-PCR

To further validate the genuine PCR product, semi-nested PCR was performed. After the RT reaction, as described above, PCR amplification was carried out for 30 cycles (as above) using gene-specific reverse primer (PR1) and first-exon-specific primers as down- and upstream primers, respectively. The resulting RT-PCR product (1 μl) was used as a template for further amplification by PCR for 30 cycles using the same upstream primer but downstream primer from internal exon (PR2). The resulting semi-nested PCR product (8 μl) was then subjected to electrophoresis on a 1.3% (w/v) agarose gel, stained with ethidium bromide and photographed on a UV transilluminator.

Subcloning and sequencing

The PCR amplified products were electrophoresed on 1.3% (w/v) agarose gels. Anticipated ethidium-bromide-stained bands were cut from the gels and DNA was recovered using a Qiaquick PCR gel purification kit. The purified DNA was either directly sequenced or subcloned into TOPO vector. E. coli JM109-competent cells were transformed. Transformed colonies were grown overnight at 37°C and plasmid DNA was purified for sequencing. Plasmid containing the insert was sequenced using an automatic sequencer using either M13 forward or reverse primers.

Computational analysis

Homology and similarity searches of the nucleotide sequences were performed using the BLASTN nonredundant database (http://www.ncbi.nlm.nih.gov/BLAST). Gene/Exon finding tools were used at HMMgene (http://www.cbs.dtu.dk/services/HMMgene), Genescan (http://genes.mit.edu/GENSCAN.html), GeneSplicer (http://cbcb.umd.edu/software/GeneSplicer/), FGENESH (http://www.softberry.com/berry.phtml). Alignment analysis was carried out using the Gene stream Align tool (http://www2.igh.cnrs.fr/bin/align-guess.cgi) and ClustalW tool available at http://www.ebi.ac.uk/clustalw. Promoter regions were characterized using tools such as TFSEARCH (http://www.ebrc.jp/research/db/TFSEARCH.html), TFBIND (http://tfbind.hgc.jp/), SIGSCAN (http://www-bimas.cit.nih.gov/molbio/signal/), CpG Island Searcher (http://www.uscnorris.com/cpgislands2/cpg.aspx) and CpGPlot (http://www.ebi.ac.uk/Tools/emboss/cpgplot/index.html).

Results and discussion

Diversity at 5’ end of mouse Prkar1b transcripts

It has been reported that mouse Prkar1b consists of 11 exons and 10 introns [14]. The structure of this gene appears obscure at 5’ end as there is large segment of DNA flanking upstream of known transcription start site not belonging to any gene. This fact prompted us to analyze the 5’ transcriptional regulatory region of this gene. To this end we sought to define the 5’ end of Prkar1b encoding transcript by applying the 5’ RACE as described in material and method section. Using 5’ RACE followed by cloning and subsequent sequencing, we found three unique clones having the heterogeneous structure at the 5’ end. Sequence analysis indicated that 5’ region of Prkar1b transcripts is diverse but all of the transcripts contain a common sequence from exon 2. Aligning these sequences with the available genomic DNA sequence of the Prkar1b gene along with 20 kb flanking sequence of nucleotides upstream of known transcription site downloaded from MGI database (MGI: 104878, Fig. 1), confirmed that these sequences are part of the genomic DNA of the Prkar1b gene representing its 5’ UTR. Three non-coding exons appearing as variants of exon 1 designated as exon 1A, exon 1B and exon 1C were located upstream of first coding exon (exon 2) according to the organization in DNA sequence of Prkar1b gene (Fig. 1). The non-coding exon 1C corresponded to the non-coding exon of previously described transcript available in pubmed with GenBank accession number NM_008923 [14]. The new transcripts were named as Mr1b-A and Mr1b-B considering published transcript as Mr1b-C according to the presence of 5’ exon 1 variant. The translational initiation codon ‘methionine’ lies in exon 2 in case of all three transcript variants which are depicted in Fig. 1. The number of nucleotides in 5’ UTR of each transcript varies from 106 bp to 217 bp as shown in
The initiation of transcription at several sites, generating transcripts differing in their 5′ UTR could render the expression of this important gene less vulnerable to mutations which could modify its expression.

In order to further validate the presence of newly identified exons with transcripts by 5′ RACE, RT-PCR was done using forward primers specific to newly identified exons from 5′ UTR and reverse primers from downstream known coding exons to maximize the length of transcripts. As shown in Fig. 2b, using common reverse primer PR1, two bands were amplified from forward primer PF2 (lane 2) and one band was amplified from forward primer PF1 (lane 3). As a positive control, we used the forward primer PF3 from the 1st exon published earlier [14] with same reverse primer, and found a single band of the anticipated 718 bp size (Fig. 2b, lane 1). One band in lane 2 was of expected size (770 bp) as anticipated from the result of 5′ RACE however, a second band appeared below showing a smaller variant. Sequencing the second band revealed that a product of 688 bp size was amplified because one more transcript variant existed in which a part of exon 1B was skipped off. This transcript variant was named as Mr1b-B’ as being a variant of transcript Mr1b-B.

In order to remove PCR artifacts, we performed semi-nested PCR as described in materials and methods section. The semi-nested PCR reaction also resulted in the amplification of expected size products that were obtained by using another reverse primer (PR2) internal to the previous reverse primer PR1 (Fig. 2c). The presence of these new transcripts having different exon structure at the 5′ end was confirmed by three independent RT-PCR experiments performed using different RNA preparations, and by cloning and sequencing the products. The sequences of these new transcripts were submitted to the GenBank and were assigned accessions numbers: Mr1b-A (GenBank accession no. JF720023), Mr1b-B (GenBank accession no. JF720024) and Mr1b-B’ (GenBank accession no. JF720025). Multiple sequence alignments of all four transcripts of Prkar1b gene including one published earlier [14] showed 100% sequence homology from exon 2 (Fig. 2d).

The transcripts of Prkar1b gene arise due to alternate promoter usage and alternative RNA splicing of three distinct non-coding exons in 5′ UTR. The nucleotide sequences present in exon 1A, exon 1B and exon 1C could also potentially play a very selective role in translational regulation of Prkar1b gene and/or may contain recognition
sites for RNA binding proteins involved in mRNA translocation, degradation or stability [15].

Analysis of the exon–intron junction region of exon 1A and exon 1B for splicing donor–acceptor sites indicated the presence of conserved consensus sequences (Fig. 3). The two new exons 1A and 1B have consensus AG/GT locus for splicing donor sites (Fig. 3). Also, a splice acceptor site at the 5' end of first coding exon is always GGAG which indicated that the same protein product could be generated from all the three transcripts. The presence of consensus...
sequences in the splice-donor/acceptor site provides evidence for the described splicing events. This is also clear from multiple sequence alignment of four transcripts which showed 100% sequence homology from exon 2 (Fig. 2d).

Three distinct promoter regions and expression of mouse Prkar1b gene

Earlier it was found that regulation of mouse Prkar1b gene is controlled by a sequence of genomic DNA comprised of a TATA-less promoter flanked by 1.5 kb of 5'-upstream sequence and 1.8 kb downstream sequence [14]. Alignment between the nucleotide sequences of newly found exons and PDP revealed that exon 1B belongs to PDP as shown in Fig. 1. Using an alternative approach of bioinformatics tools the immediate upstream regions of exon 1A, exon 1B and exon 1C were characterized as promoter regions designated as PA, PB and PC respectively. Thus PDP contains PB, exon 1B and PC sequence regions. Using CpG island detection tools like CpG Island Searcher and CpG Plot, no CpG island was established in or near upstream of exon 1B, however there was presence of CpG islands in and near upstream of exon 1A and exon 1C (Fig. 4). These CpG islands have an average GC content of about 65–70% and observed/expected CpG ratio greater than 60%. The presence of CpG islands may play an important role in gene regulation, for example, in causing gene silencing or transcriptional repression by hypermethylation of the CpG Island [16, 17]. There was an E-box consensus sequence of CACGTG at position -538
upstream of transcription start site of exon 1B. This E-box may act either as an enhancer or a silencer depending on the binding protein. It is well established that Max/Myc dimer binding to E-box activates transcription whereas the Max/Mad dimer binding to E-box suppresses transcription of the genes [18]. Thus, in silico studies indicated that Max/Myc or Max/Mad dimer can regulate transcription of Prkar1b gene from promoter PB.

As three promoter regions lead to the expression of the transcript encoding same protein, there may be possibility of differences in their transcription factor (TF) binding sites so that they can respond to variety of signals in and out of the cell. For such analysis we performed in silico analysis using tools and databases like TFSEARCH, TFBIND and SIGSCAN. Analysis revealed a complex array of recognition sequence motifs for different types of TFs (Fig. 3). It was interesting to observe that the three promoters PA, PB and PC show discrepancy for TFs responsible for constitutive and conditional expression of Prkar1b gene. All three promoters have some binding sites in common for some of the TFs though number vary for example the general TF Sp1 (steroidogenic factor 1) and NF-I (Nuclear factor I) remain constitutively active. Sp1 is expressed in all mammalian cells and is involved in regulating the transcriptional activity of genes implicated in most cellular processes [19, 20]. NF-I is involved in eukaryotic transcription as well as DNA replication [21]. However, there are different binding sites for different transcription factors specific to each promoter e.g. signal dependent TF CREB (cAMP response element-binding) which decreases or increases transcription [22] has binding site in promoter region PA and this site may be involved in auto-regulation of transcription of Prkar1b gene by its catalytic subunit counterpart; GR (glucocorticoid receptor) TF which is a conditionally active and responds to cortisol and other glucocorticoids [23] has number of binding sites in promoter region PB; AP-2 (Activator protein-2) which is known to have role in vertebrate development, apoptosis and cell-cycle control [24], can bind to promoter PB and PC. The PB promoter seems to be more regulatory in nature as it has number of different TF binding sites than PA and PC. The promoter

Fig. 4 CpG Plot generated by tools available at EBI used for CpG Island prediction. The predicted promoter region sequence was used for all three promoters. CpG Islands were present in the regions of promoter PA and PC with observed/expected ratio >0.60 and G + C percentage >55.00. However, the putative CpG Islands were found in PC only.
PB has binding sites for GATA, PIT-1, Hox etc. which take part in developmental and cell specific expression and once expressed require no additional activation. This fact prompted us to find out whether expression from PB was differentially regulated or not. To this end we studied the expression patterns of 5’ UTR exons by RT–PCR using exon specific forward primer and reverse primer (PR2) in four major organs namely heart, brain, skeletal muscle and liver at three different postnatal development stages of mouse like 3 day, 15 day, 60 day namely PN3, PN15, PN60 respectively (Fig. 5). Only few tissues from early and late postnatal stages were studied because it was beyond the scope of the work presented here to study all tissues at different developmental stages of mouse. As envisaged from in silico analysis the expression from exon 1B was found to be differentially expressed. The exon 1B containing transcript was only expressed in brain at early PN stages however, at PN60 (adult mice) exon 1B was expressed in all tissues. Exon 1A and exon 1B were expressed in all tissues at all PN stages examined. Few lanes have unexpected extra bands that were due to non specific binding of primers as confirmed by sequencing. Marker (1 kb DNA ladder) was always run to correlate the amplified product size.

Conclusion

Our results confirmed the presence of four Prkar1b transcript variants in different tissues of adult mouse like brain, liver, heart and skeletal muscle where transcript variants Mr1b-A and Mr1b-C were ubiquitously expressed however, the transcript variants Mr1b-B and Mr1b-B’ were expressed in brain at PN3 and PN15 only and expressed in all tested tissues at PN60. The study of transcripts belonging to mouse will be helpful while investigation of diseases and other functional aspects in mouse as model, in particular to those cases which pertain to regulation of this gene or cloning experiments in expression vectors with one of the three putative promoter regions required as per the factors under which expression has to be investigated.
Acknowledgments  The authors are thankful to the Department of Biotechnology (DBT), Ministry of Science and Technology, New Delhi, India for providing generous funding to project No. BT/PR-10917/BID/07/254/2008. The necessary facilities provided by A.M.U., Aligarh, India are thankfully acknowledged.

References

Alternatively Spliced Three Novel Transcripts of \textit{gria1} in the Cerebellum and Cortex of Mouse Brain

Shafquat Azim · Abdul Rouf Banday · Mohammad Tabish

Received: 22 July 2011 / Revised: 30 August 2011 / Accepted: 8 September 2011
© Springer Science+Business Media, LLC 2011

Abstract  Glutamate receptor type 1 (GluR1) subunit of \(\alpha\)-amino-3-hydroxy-5-methyl-4-isoxazole propionate receptors plays an important role in the expression of long-term potentiation and memory formation. GluR1 is encoded by \textit{gria1} gene containing 16 exons and 15 introns in mouse. Previous studies have reported two alternatively spliced variants of this subunit. These flip and flop variants differ enormously in their properties as well as expression. In our studies, we report the presence of three new transcripts of this gene present in the cerebellum and cortex of mouse brain produced by alternative splicing at 5' end. Four new exons are reported; N1 is located in 5' untranslated region, N2 is located in the 1st intronic region while N3 and N4 are located in the 2nd intronic region. The properties of these new exons encoding N-terminal variants are highly diverse. N1, N3 and N4 are coding while N2 is a non-coding exon and results in a truncated transcript. The existence of N2 exon containing transcript is further supported by the presence of an Expressed Sequence Tag from the database. The translated amino acid sequences of these transcripts differ in the presence of signal peptide as well as in their phosphorylation and acetylation pattern. The differences in their properties might be involved in receptor modulation.

Keywords  Glutamate · \(\alpha\)-amino-3-hydroxy-5-methyl-4-isoxazole propionate receptors · Cerebellum · Cortex · Alternative splicing · Flip and flop variants

Introduction

The \(\alpha\)-amino-3-hydroxy-5-methyl-4-isoxazole propionate receptors (AMPARs) are of fundamental importance in the brain as they are responsible for the majority of fast excitatory synaptic transmission, synapse formation and stabilization and synaptic plasticity. AMPARs are comprised of four subunits (GluR1–GluR4). Amino acid sequences for these subunits share approximately 70% sequence homology. The genes encoding these subunits may undergo alternative splicing in two distinct regions, resulting in long or short C-termini, and flip or flop variants in an extracellular domain [1]. Two variants designated as flip and flop are produced and they differ in 38 amino acids that lie between the third and fourth putative transmembrane domain [2]. These variants differ in their channel kinetics, pharmacological properties, sensitivity to certain allosteric modulators and exhibit regional and cell-specific expression, developmental regulation, and modification by physiological insults, lesions and disease [3–11]. AMPARs are homomers of identical subunits or heteromers of different subunits. The N-terminal is important for receptor assembly and receptor targeting [12, 13]. Subunits compatible only at the NTD interacts with each other and are functional only if there is compatibility in the C-terminal half of the subunits. Such hetero-oligomerization produces functional receptors. This suggests that there might be diversity in the N-terminus of the subunits which facilitates the various combinations of the subunits. Moreover, AMPARs have widespread and varied distribution in the brain [14].

The presence of flip-flop variants suggests that alternative splicing plays an important role in generating diversity which facilitates their differential expression and varied subunit combinations. The other family of glutamate
receptors includes NMDARs. The \textit{NR2B} subunit encoding gene undergoes alternative splicing to produce 5 variants with different N-termini [15–17]. These subunits belonging to the different families of the glutamate receptor share considerable sequence homology. This provided us the impetus to study the N-terminal variants of \textit{grial} subunit of AMPAR. In this study we have analyzed the potential new 5' end exons which can splice with the internal exons to produce new transcript variants of \textit{grial} gene in mouse brain. The novel exons of these new transcripts might play a role in differential targeting of the receptor in different areas of the brain. We have identified three new transcripts of \textit{grial} gene that arise due to alternative splicing of four new exons encoding GluR1 subunit with three different N-termini. These new exons N1, N2, N3 and N4 were identified from the 5' untranslated region (5'UTR), 1st intron and 2nd intron. They can splice with each other or internal exons to produce three transcript variants T1, T2 and T3.

Materials and Methods

Materials

Mice (A/J) were bred in house according to the Institutional Animal Care and Use Committee and Guidelines of the Committee on Care and Use of Laboratory Animal Resources, National Research Council (Department of Health, Education and Welfare; National Institutes of Health). The total RNA extraction kit was purchased from iNtRON Biotechnology, Inc. (Gyeonggi-do, Korea). M-MuLV-Reverse Transcriptase, PCR nucleotide mix, 100 bp PCR DNA ladders were purchased from GenScript (USA). The TOPO-TA cloning kit II was obtained from Invitrogen Corp. (Carlsbad, CA). The Plasmid DNA miniprep kit and Qiaquick PCR gel purification kits were purchased from Qiagen, Inc. (Santa Clarita, CA). Primers were custom synthesized from Sigma Aldrich Chemicals Pvt. Ltd., India. All other chemicals used in the experiments were of molecular biology grade.

RNA Preparation

Total cellular RNA from cerebellum and cortex of 2 months old A/J mouse was prepared using the easy spin™ (DNA free) Total RNA Extraction Kit (iNtRON Biotechnology, Inc) according to the manufacturer’s instructions. The eluted RNA was quantitated spectrophotometrically and RNA integrity was checked by ethidium bromide staining on denaturing agarose gel electrophoresis.

Primers

Genomic sequence of \textit{grial} gene was downloaded from NCBI with accession number GenBank ID NM001113325. Primers were designed using the downloaded sequence. The following oligonucleotides primers were used:

<table>
<thead>
<tr>
<th>Exon</th>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1 specific forward primer</td>
<td>MCTRGRIA1</td>
<td>GCC GTA CAT CTG TCT TTT CTG CAC CG</td>
</tr>
<tr>
<td>N1 specific forward primer</td>
<td>MN1GRIA1</td>
<td>CAT TCT ACA GAT ACA TGC TGA GCT CAG GG</td>
</tr>
<tr>
<td>N2 specific forward primer</td>
<td>MN2GRIA1</td>
<td>CAT TCC TGG CTG CCA TCC TCT CAA CCT G</td>
</tr>
<tr>
<td>N3 specific forward primer</td>
<td>MN3GRIA1</td>
<td>GCT CAG GGC AGA ATG CAG TTG ATT GGT C</td>
</tr>
<tr>
<td>Reverse primer from the junction of 5th and 6th exon</td>
<td>MREV1GRIA1</td>
<td>GTC AAT GTC CAT GAA GCC CAG GTT GCC</td>
</tr>
<tr>
<td>Reverse primer from 5th exon</td>
<td>MREV2GRIA1</td>
<td>GTG GTA CCC GAT GCC GTC CTT TTC TAG C</td>
</tr>
</tbody>
</table>

Rapid Amplification of cDNA Ends (5'RACE)

The presence of alternatively spliced transcripts was examined through 5' rapid amplification of cDNA ends (RACE). 2 µg of total RNA from cerebellum and cortex were amplified using a 5'RACE kit and \textit{grial} specific reverse primer (MREV1GRIA1) from the junction of exon 5 and exon 6. The amplification was performed according to the instructions provided with the kit. The RACE product was fractionated by electrophoresis using 1.2% agarose gel. Several bands were excised from the gel, purified and sub-cloned into the TOPO vector, and sequenced as described later.

Reverse Transcriptase (RT)-PCR

Multiple transcripts encoded by \textit{grial} gene were identified using RT-PCR. RNA (2 µg) from the cortex and cerebellum were primed with oligo (dT)$_{18}$ and single-stranded cDNA was synthesized using the RevertAid™ H-Minus Reverse Transcriptase at 43°C for 1 h in a total volume 20 µl. 1 µl of the single-stranded cDNA was then added to PCR system for amplification using appropriate set of primers in a total vol. 50 µl. PCR product was amplified using touchdown PCR for 30 cycles. The cycle was performed as; denaturation at 94°C for 4 min, 1 cycle; 93°C,
30 s; annealing 66°C, 30 s with a decrease of 0.3°C per cycle, extension 72°C for 45 s, with final extension at 72°C for 8 min. Final product was subjected to electrophoresis on a 1.2% agarose gel, stained with ethidium bromide and photographed on a UV transilluminator.

Semi-Nested RT-PCR

To confirm the results obtained after first round of PCR and reconfirm the genuine PCR product, semi-nested PCR was performed. After the RT reaction, as described above, PCR amplification was carried out for 30 cycles (as above) using gene-specific reverse primer (MREV1GRIA1 from the junction of exon 5 and exon 6) and first-exon-specific primer (MREV2GRIA1 internal to the reverse primer used in 1st PCR). 10 μl of the semi-nested PCR product was then subjected to electrophoresis on a 1.2% (w/v) agarose gel, stained with ethidium bromide and photographed on a UV transilluminator.

Subcloning and Sequencing of RT-PCR Products

The RACE and semi-nested PCR amplified products were electrophoresed on 1.2% (w/v) agarose gels. Anticipated ethidium-bromide-stained bands were cut from the gels and DNA was purified using a PCR gel purification kit. The purified DNA was subcloned using the plasmid cloning vector. E. coli JM109-competent cells were transformed. Transformed colonies were grown overnight at 37°C and plasmid DNA was purified for sequencing. Plasmids containing the insert were sequenced by automatic sequencer using either M13 forward or reverse primers [18].

Bioinformatic Analysis

Homology and similarity searches of the obtained nucleotide sequences were performed using the BLASTN non-redundant database (http://www.ncbi.nlm.nih.gov/BLAST). Alignment analysis was carried out using the Gene stream Align tool (http://www2.igh.cnrs.fr/bin/align-guess.cgi) and ClustalW tool available at www.ebi.ac.uk/clustalw [19]. TFBind (http://tfbind.hgc.jp/), TFSEARCH (http://www.cbrc.jp/research/db/TFSEARCH.html) and WWW signal scan (http://www-bimas.cit.nih.gov/molbio/signal/) were used to study the transcription factors binding to the upstream region of the exons. Some of the properties of the amino acid sequences coded by the new exons were analyzed using ExPASy tools (http://ca.expasy.org/).

Results and Discussion

Three New and Differentially Expressed Transcripts of gria1

AMPARs are expressed in various subunit combinations. The four subunits (GluR1–GluR4) of the receptor assemble in various stoichiometries to form functional ion channels [20]. The different subunit composition determines channel gating kinetics, conductance, vesicular trafficking and targeting to specific synapse [21–23]. In mouse, gria1 gene is located on chromosome 11 and a large 5’UTR of almost 20 kb is present, the 1st intron is 3 kb and the 2nd intron is 150 kb in size as depicted in Fig 1a. The second intron is unusually large. Mouse database searches revealed that this subunit is coded by a single gene consisting of 16 exons and 15 introns. cDNA corresponding to gria1 is present in the database with Accession number NM001113325 that encodes for a protein of 907 amino acids. In order to investigate the presence of possible new splice variants 5’RACE was performed. Total RNA isolated from cerebellum and cortex was amplified using a 5’RACE kit using the primers spanning exon 5 and exon 6. We were able to amplify multiple transcripts from the RNA preparations of mouse brain as shown in Fig 2. Following 5’RACE, sequencing of the products was done as described in “Materials and Methods” section. The cDNA sequences containing the reverse primer sequence corresponding to the part of 5th and 6th exon was further analyzed for their 5’ ends. We obtained several clones containing new 5’ end sequences. The analysis of the sequencing data revealed the identification of three new cDNA of gria1 differing at the 5’ ends which were not reported earlier. The alignment of these sequences with the genomic sequence confirmed that these cDNA were transcribed from gria1 gene and the new exons were derived from 5’UTR, 1st intron and 2nd intron of the gene. Three new sequences differing in their N-terminus were produced from the usage of four novel exons. These exons were designated as N1, N2, N3 and N4 and the resulting transcripts as T1, T2 and T3. N1 exon derived from 5’UTR splices with the 1st reported exon (E1) producing a variant (T1) which encodes for an additional 19 amino acid residues to the published N-terminus of GluR1. N2 exon was found to be non-coding and splices with the 2nd reported exon (E2) to produce a variant (T2) which is truncated at the N-terminal by 69 amino acids. N3 exon splices with a novel N4 exon which then splices with the 3rd reported exon (E3) producing a different variant (T3) which is 40 amino acid residues shorter as compared to the reported sequence. N4 exon is also located in the 2nd intron and is 42,173 bp downstream of exon N3. These transcripts variants produced through alternative splicing are depicted in Fig 1b. The presence of
N2 variant was further supported through Expressed Sequence Tag (EST). EST with accession number BY719132 from adult male olfactory brain encodes for a sequence identical to N2 variant. These analyses were further confirmed by RT-PCR using upstream and downstream primer (given in “Materials and Methods”) from the 1st newly identified exons and from the junction of 5th and 6th exon (MREV1GRIA1) respectively. The RT-PCR results revealed bands of expected sizes as predicted from 5'0 RACE sequence data. These results were further confirmed through semi-nested PCR using the downstream primer MREV2GRIA1 (from exon 5th) which was designed from a sequence internal to the first reverse primer (MREV1GRIA1). The agarose gel electrophoresis of the semi-nested PCR revealed bands of expected sizes for T1 (752 bp), T2 (684 bp) and T3 (573 bp) in both cerebellum and cortex of mouse brain (Fig 3). These semi-nested RT-PCR products were also cloned and sequenced. Sequencing confirmed that these cDNAs were produced through alternative splicing of gria1 gene and the transcripts were formed through the use of N1, N2 and N3 exons in combination with the internal exons as discussed above. The sequences were submitted to the GenBank and following accession numbers JN180918, JN180919 and JN180920 were obtained for T1, T2 and T3 respectively.

The study of the structure of these exons showed that the donor–acceptor residues associated with these new exons are found to be conserved. The donor “GT” and the
acceptor “AG” residues are conserved in all these new exons (Fig 4). In order to understand the differential expression of these variants we analyzed the 5' upstream of these new exons. In silico analysis employing programs like TFBIND, TFSEARCH and SIGSCAN of the upstream region of these newly predicted exons showed that a number of transcription factor (TF) binding sites are available as depicted in Fig 5. The 5' upstream region of N1, N2 and N3 exons contain TATA Binding Protein (TBP) sites. TBP is a key component of eukaryotic transcription initiation machinery [24]. The locations of other general TFs are also shown in the Fig 5. SP-1 TF binding site is present in the upstream of all these exons. This TF is involved in the gene expression in early development of an organism [25]. Multiple Nuclear Factor-1 (NF-1) binding sites are present preferably this is associated with the upstream region of the genes that are highly expressed [26]. Upstream of E1, N1 and N2 exons report Myc-associated zinc-finger protein (MAZ) TF binding site. MAZ and SP-1 are often found regulating the same gene and may even bind to the same sites [27]. Activating protein-2 (AP-2) TF binding site is located upstream of N2 and AP-2 play important regulatory roles in vertebrate development, apoptosis and cell-cycle control [28]. Pit-1 TF is a pituitary-specific TF while serum response factor (SRF) is ubiquitous nuclear protein that stimulates both cell proliferation and differentiation. A number of TFs bind to CCAAT box such as NF-Y, CP-1, CDP, GATA-1, and NFE3. Some TFs are tissue specific while others are constitutively expressed. These analyses suggest that these transcripts may involve various TF activities which might be responsible for their differential activity in different regions of the mouse brain.

Comparative Analysis of the New Variants

In order to understand the structure and functions of these variants we analyzed their conceptually translated amino
Fig. 5 Analysis of the upstream region of these exons using web-based programs TFBIND, TFSEARCH and SIGSCAN. The upstream region of the exons is represented by rectangular boxes and the binding positions of various transcription factors (TFs) are shown. 600 bp upstream sequences are analyzed to locate the position of TFs. Analysis revealed that many brain specific TFs can potentially bind to upstream regions of the newly identified exons.

![Diagram showing the analysis of upstream regions and TF binding positions.](image)

Fig. 6 Multiple sequence alignment (ClustalW) of the deduced amino acid sequences of the published transcripts (E1) and the newly identified transcripts (N1, N2, and N3). (+), colon (:), and dash (−) represent identical, similar, and no residue respectively. The accession numbers are provided at the end of the sequences. All sequences are identical from third exon onwards. Accession numbers are given at the end of each sequence.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Accession Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td></td>
</tr>
<tr>
<td>N1</td>
<td></td>
</tr>
<tr>
<td>N2</td>
<td></td>
</tr>
<tr>
<td>N3</td>
<td></td>
</tr>
<tr>
<td>E1</td>
<td></td>
</tr>
<tr>
<td>N1</td>
<td></td>
</tr>
<tr>
<td>N2</td>
<td></td>
</tr>
<tr>
<td>N3</td>
<td></td>
</tr>
<tr>
<td>E1</td>
<td></td>
</tr>
<tr>
<td>N1</td>
<td></td>
</tr>
<tr>
<td>N2</td>
<td></td>
</tr>
<tr>
<td>N3</td>
<td></td>
</tr>
<tr>
<td>E1</td>
<td></td>
</tr>
<tr>
<td>N1</td>
<td></td>
</tr>
<tr>
<td>N2</td>
<td></td>
</tr>
<tr>
<td>N3</td>
<td></td>
</tr>
</tbody>
</table>

Accession Numbers:
- E1
- N1
- N2
- N3

(Additional sequence data and accession numbers are provided in the full text of the article.)
Table 1 A comparative analysis of the amino acid sequence of the translated transcripts

<table>
<thead>
<tr>
<th>Properties</th>
<th>E1 + E2</th>
<th>N1 + E1</th>
<th>E2 + E3</th>
<th>N3 + N4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amino acid sequence</td>
<td>MPYIFAFFCTGFLGAVVGANFPNNIQI&lt;sup&gt;G&lt;/sup&gt;RFQGAQCLLQLQMPYIFAFFC</td>
<td>MLSSGLGKGRACILLQLQGFGLAVGVGPFNNIQI</td>
<td>MTYRFCSQFSKGVYAIFGFEYRRTVNMLTSFCGALHVCFITPSFVPDTSNQFVLQLRPEDQALISIDHY</td>
<td>MQLIGLSMEMGYDALQETKANRENQNMALGLAV</td>
</tr>
<tr>
<td>No of aa</td>
<td>73</td>
<td>46</td>
<td>84</td>
<td>34</td>
</tr>
<tr>
<td>pI</td>
<td>4.90</td>
<td>9.50</td>
<td>6.69</td>
<td>4.41</td>
</tr>
<tr>
<td>Mw</td>
<td>8,170.37</td>
<td>4,907.89</td>
<td>9,936.43</td>
<td>3,740.34</td>
</tr>
<tr>
<td>Signal peptide</td>
<td>Yes</td>
<td>Signal anchor</td>
<td>Non-secretory protein</td>
<td>Non-secretory protein</td>
</tr>
<tr>
<td>Mitochondrial targeting</td>
<td>No</td>
<td>No</td>
<td>Yes cleavage site at 24</td>
<td>No</td>
</tr>
<tr>
<td>sequence</td>
<td>VVG-AN</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NetOglyc</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>NetNglyc</td>
<td>MPYIFAFFCTGFLGAVVGANFNQIQI&lt;sup&gt;G&lt;/sup&gt;RFQGAQCLLQLQGFGLAVGVGPFNNIQI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NetPhos</td>
<td>S&lt;sup&gt;66&lt;/sup&gt;, S&lt;sup&gt;67&lt;/sup&gt;, T&lt;sup&gt;71&lt;/sup&gt; Ser:2 Thr:1 Tyr:0</td>
<td>No</td>
<td>Yes S&lt;sup&gt;66&lt;/sup&gt;T&lt;sup&gt;34&lt;/sup&gt;, T&lt;sup&gt;48&lt;/sup&gt;, Y&lt;sup&gt;14&lt;/sup&gt; Ser:1 Thr:2 Tyr:1</td>
<td>Yes, S&lt;sup&gt;7&lt;/sup&gt;</td>
</tr>
<tr>
<td>NetAcet</td>
<td>No</td>
<td>Yes&lt;sup&gt;3&lt;/sup&gt;S</td>
<td>Yes&lt;sup&gt;3&lt;/sup&gt;T</td>
<td>No</td>
</tr>
<tr>
<td>NetPhosK</td>
<td>PKC T&lt;sup&gt;71&lt;/sup&gt;</td>
<td>PKC S&lt;sup&gt;4&lt;/sup&gt;</td>
<td>PKC T&lt;sup&gt;5&lt;/sup&gt;</td>
<td>No</td>
</tr>
<tr>
<td>Myristoylation</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

The number as superscript indicates the position of the amino acid residue undergoing modification. The 1st two exons of the new transcripts are analyzed in order to study the difference associated with the N-termini variation of these transcripts. The sequence ‘NIS’ in bold letter defines the glycosylation site.
acids. Multiple sequence alignment employing web-based ClustalW of these translated sequences revealed that they differed only in their N-terminal (Fig 6). A number of properties of these variants were studied and compared using ExPASy tool. Table 1 compares some of the properties of these translated exon sequences and provides a number of differences in the properties of these new exons. The phosphorylation and acetylation pattern of these sequences are very different. N1 and N2 contain 3S and 2T respectively which can be potentially acetylated. However, no such residues are present in E1 and N3 variants. Although a majority of eukaryotic proteins, almost 50% of yeast proteins and 80% of the human proteins are N-terminally acetylated (Nt-acetylated), the function of this modification is largely unknown [29, 30]. For some proteins such as actin, Nt-acetylation has been reported to affect protein functionality, e.g. non-acetylated actin is less efficient at assembling microfilaments [31]. Recently, the other role suggested for acetylation is that Nt-acetylation of a protein can function as a degradation signal (degron) playing a role in protein turnover and homeostasis [29]. N1 and N2 variants have potential PKC dependent phosphorylation sites. Protein phosphorylation of neuronal glutamate receptors by cyclic AMP-dependent protein kinase (PKA) and protein kinase C (PKC) may regulate their function and play a role in some forms of synaptic plasticity [32]. E1 contains a signal peptide while N1 contains a signal anchor. Many integral proteins contain signal non-cleavable transmembrane signal sequence called signal anchor sequence which anchors the protein to the membrane [33–35]. N2 contains a mitochondrial targeting sequence. Glutamate excitotoxicity often leads to mitochondrial dysfunctioning associated with apoptotic and necrotic neuronal cell death [36]. N3 does not contain such signal sequence. The detailed study of this variant might provide new pathways to study the regulation in neuronal function and play a role in some forms of synaptic plasticity [32].

The diverse molecular properties of the variants might be responsible for their differential expression in cerebellum and cortex of the mouse brain as shown by the RT-PCR result in our study. Our studies provide the 1st experimental evidence for the presence of multiple alternatively spliced transcripts variants of the gria1 gene encoding different N-termini. The new exons are produced from the large intronic regions of the gene. The large introns are frequently associated with the evolution of a gene. These transcripts being differentially expressed in cerebellum and cortex of mouse brain suggest that the N-termini may be associated with differential targeting of the receptor to different regions of the brain. The properties of these variants are highly heterogeneous and might play a role in providing functional, structural and molecular diversity to the AMPA receptors. A detailed study of these variants might throw light on the mechanism of synaptic plasticity, neuronal development and various neuropathological conditions involving AMPARs.

Acknowledgments The authors are thankful to CSIR, New Delhi, India for providing senior research fellowship to SA. Partial funding from the Department of Biotechnology (DBT), Ministry of Science and Technology, New Delhi and necessary facilities provided by AMU, Aligarh, India is thankfully acknowledged.

References