CHAPTER 6

This chapter deals with the identification and confirmation of alternatively spliced transcript variants of *htr4* gene encoding HTR4 receptor.

*Azim S, Banday AR, Tabish M. Identification of alternatively spliced multiple transcripts of 5-hydroxytryptamine receptor in mouse. Brain Res Bull. 2011*
6.1. Introduction

5-Hydroxytryptamine receptor (HTR) more frequently called serotonin receptors bind to the neurotransmitter serotonin. HTR includes seven different receptor classes classified as HTR1-7. HTR3 is a ligand gated ion channel (LGIC) while the rest of the receptors belong to the G-protein coupled receptors family [Norum JH et al. 2003; Pauwels PJ 2000] (Figure 6.1). Seventeen different genes encode serotonin receptor family members. Out of these five of the genes belonging to HTR1 class contain a single exon thereby lacking alternatively spliced variants. Two genes encoding HTR5 contain two exons and alternative splicing events have not been reported. However, alternatively spliced variants have been reported for HTR2 [De Lucchini S et al. 2001; Guest PC et al. 2000; Xie E et al. 1996], HTR3 [Bruss M et al. 2000], HTR4 [Ray AM et al. 2009; De Maeyer JH et al. 2008; Vilaro MT et al. 2002; Bender E et al. 2000; Blondel O et al. 1998; Gerald C et al. 1995], HTR6 [Olsen MA et al. 1999] and HTR7 [Gellynck E et al. 2008; Krobert KA et al. 2001] in various organisms. htr4 gene in mouse is found to undergo alternative splicing at its C-terminus to produce four variants HTR4 (a,b,c,f). HTR4 (a) and HTR4 (b) are heavily expressed in brain. HTR4 (e) and HTR4 (f) are myenteric plexus specific [De Maeyer JH et al. 2008; Liu MT et al. 2009]. Even in other vertebrates, HTR4 is extensively spliced to generate several isoforms. In rats, four variants are reported HTR(a), HTR(b), HTR (e) and HTR (c1) exhibiting differential expression where HTR4 (a) is expressed more in brain and dorsal root ganglion, HTR4 (b) and HTR (c1) are abundant in gastrointestinal tract however, HTR (c1) does not exhibit constitutive receptor activity [Ray AM et al. 2009]. In humans, there are eight alternatively spliced variants (a-g, and n) of HTR4 receptor which differ in their C-terminal tail length and structure, however the ninth variant contains an insertion of 14 amino acids in the second extracellular loop [De Maeyer JH et al. 2008; Vilaro MT et al. 2002; Bender E et al. 2000; Blondel O et al. 1998]. In porcine, eleven variants have been reported [De Maeyer JH et al. 2008] which differ at the level of receptor
palmitoylation [Ponimaskin EG et al. 2002; Ponimaskin EG et al. 2001],
downstream coupling [Robert SJ and Lezoualc’h F 2008; Pindon A et al.
2002], receptor desensitization [Pindon A et al. 2004; Mialet J et al. 2003] and
constitutive receptor activity [Claeysen S et al. 1999].

**Figure 6.1: Structure of HTR.** There are seven classes of serotonin receptors. Except HTR3
which is a LGIC all others are GPCR. These receptors contain 7 transmembrane segments.
This figure has been adopted from CNSforum.com.

HTR4 is involved in cognitive function and memory consolidation [Vilaro MT
et al. 2005], enteric nervous system growth and maintenance [Liu MT et al.
2009], increasing heart rate, atrial contractile force, hastening atrial relaxation
[Bach T et al. 2001]. It is responsible for feeding [Jean A et al. 2007] and stress
responses [Bockaert J et al. 2004]. HTR4 is expressed in a variety of tissues
namely brain [Dumuis A et al. 1988a; Dumuis A et al. 1988b], heart [Kaumann
AJ and Levy FO, 2006; Blondel O et al. 1997], and intestine [Mader R et.al
2006] where it performs diverse functions. This receptor serves as a potential
target for the development of therapeutic agents involved in the treatment of
numerous mental illness including Alzheimer’s disease [Maillet M et al. 2004],
feeding disorders such as anorexia nervosa [Jean A et al. 2007], and stress
related disorders such as anxiety [Bockaert J et al. 2004] and depression [Lucas
Chapter 6: Introduction

G et al. 2007. *htr4* gene is located on chromosome number 18 and spans 145kb. The gene encoding beta adrenergic (β2-AR, *adrb2* gene) receptor is embedded within this gene. *htr4* and *adrb2* can interact at transcriptional as well as post transcriptional levels [Hernandez MC and Janusonis S, 2010]. *adrb2* is intronless and spans 2.2 kb. These two receptors can form homodimers and heterodimers which can integrate serotonergic and noradrenergic signals. The expressions of both these receptors are interdependent at transcriptional level. Both HTR4 and β2-AR are implicated in autism spectrum disorders, depression and Alzheimer’s disease. The requirement for a homodimeric or heterodimeric receptor assembly by a cell might involve transcriptional regulation of the genes. This suggests that the N-terminal variants of the receptors produced through alternative splicing might be involved in facilitating their assembly. As *adbr2* is intronless the major site of regulation appears to be HTR4. Alternative splicing provides transcriptional regulation through the use of new exons, exon skipping or alternate promoter usage. β2-AR has been implicated in cardio inotropy. It is hypothesized that β2-AR-coupled PI3K constrain increases in cardiac inotropy through cyclic adenosine monophosphate-dependent phosphodiesterase activation although the underlying mechanism remains elusive [Gregg CJ et al. 2010]. HTR4 receptors are present in human and porcine atrial myocytes while they are reported to be absent from the heart of small laboratory animals [De Maeyer JH et al. 2008].

Although a number of C-terminal variants have been reported which display differential expression, no N-terminal variants has been demonstrated so far. Keeping in mind that the β2-AR is located in the 5' untranslated region (5'UTR) of the *htr4* gene, we decided to study the nature of 5'UTR of *htr4* gene and search for potential transcripts encoding N-terminal variants which might be expressed in brain and heart. Moreover, the restriction upon the use of small laboratory animals to study the functions of HTR4 in heart propelled us to take up the study of *htr4* transcripts in heart. Our study provides the first
experimental evidence for the presence and identification of four new transcripts encoding HTR4 variants differing at their 5'ends from mouse brain. Five novel exons (N1, N2, N3, N4 and N5) splice with internal exons in different combinations to produce four new transcript variants designated as T1, T2l, T2s and T3. All the four transcripts are expressed in mouse brain. T3 transcript is unique being expressed in heart making it a good candidate for further studies which can help to understand the role of the receptor in heart. The existence of T3 transcript in heart can aid in designing a specific drug targeting various cardiovascular disorders.
6.2. Materials and Method

6.2.1. RNA preparation

Total cellular RNA of brain and heart of 2 months old mouse was prepared using the easy spinTM (DNA free) Total RNA Extraction Kit according to the manufacturer’s instructions. The eluted RNA was quantitated spectrophotometrically and RNA integrity was checked by ethidium bromide staining after denaturing agarose gel electrophoresis.

6.2.2. Primers

Genomic sequence of htr4 gene was downloaded from NCBI with accession number GenBank ID NM008313. Primers were designed using the downloaded sequence. The oligonucleotides primers were used are shown in Table 6.1.

Table 6.1: Anticipated RT-PCR products. The transcripts produced, exons in amplicon, exon specific primers, semi-nested PCR product size obtained with external reverse primer (MHTR4E4REV1: 5’ CAC ATC AAC TAT GCC GAT GTT CCA G) internal reverse primer (M5HTR4REV2: 5’ CCA GCC TTG CAT TAT GGG GAG AAA AG) were

<table>
<thead>
<tr>
<th>1st Exon</th>
<th>Transcripts</th>
<th>Name and exon specific primer sequence</th>
<th>Exons in amplicon</th>
<th>Semi-nested PCR product size (bp) with Rev2</th>
<th>PCR product size (bp) with Rev1</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>T1</td>
<td>MHTR4N1F : 5’ CTC AGT GGG ACA AAA GCC CTG CAG</td>
<td>N1-N2-E2-E3-E4</td>
<td>630</td>
<td>654</td>
</tr>
<tr>
<td>N3</td>
<td>T2l</td>
<td>MHTR4N3F : 5’ GTA AAC ACA TGT CTT GTT GCA TAC AAC ATG AC</td>
<td>N3-N4-E2-E3-E4</td>
<td>615</td>
<td>639</td>
</tr>
<tr>
<td>N3</td>
<td>T2s</td>
<td>MHTR4N3F : 5’ GTA AAC ACA TGT CTT GTT GCA TAC AAC ATG AC</td>
<td>N3-N4-E3-E4</td>
<td>489</td>
<td>513</td>
</tr>
<tr>
<td>N5</td>
<td>T3</td>
<td>MHTR4N5F : 5’ CCC GAA GCC TGT CCT GT CAA TAG</td>
<td>N5-E2-E3-E4</td>
<td>502</td>
<td>526</td>
</tr>
<tr>
<td>E1</td>
<td>Published</td>
<td>MHTR4CTRF : 5’ AAT CCA ACG TCC TCA TGC CCA TTT CCT GT</td>
<td>E1-E2-E3-E4</td>
<td>513</td>
<td>537</td>
</tr>
</tbody>
</table>
6.2.3. 5′ RACE

In order to detect the presence of possible alternatively spliced transcript variants of \textit{htr4} gene, 5′ RACE was performed using 2µg of total RNA from brain and heart of mouse, 5′ RACE kit and \textit{htr4} specific reverse primer M5HTR4REV1 from the junction of exon 4 and exon 5. The amplification was done following the instructions with the kit. The amplified RACE product was fractionated by electrophoresis using 1.2% agarose gel. Several bands were excised from the gel, purified and subcloned into the TOPO vector and sequenced as described later.

6.2.4. Reverse transcriptase (RT)-PCR

The total isolated RNA was amplified using RT-PCR which led to the identification of several alternatively spliced transcript variants from both brain and heart of mouse. 2µg of total RNA from both the tissues were primed with oligo (dT)$_{18}$ and single-stranded cDNA was synthesized using the RevertAid$^\text{TM}$ H-Minus Reverse Transcriptase at 43ºC for 1h in a total volume 20µl. 1µl of the synthesized single-stranded cDNA was used for PCR amplification using appropriate sets of primers in a total volume of 50µl employing touchdown PCR. The cycle were performed as; denaturation at 94ºC for 4min, 1cycle; 30 cycles were repeated as denaturation 93ºC, 30sec; annealing 66ºC, 30sec with a decrease of 0.3ºC per cycle, extension 72ºC for 45sec, final extension at 72ºC for 8min. Final product was subjected to electrophoresis on a 1.2% agarose gel, stained with ethidium bromide and photographed on a UV transilluminator.

6.2.5. Semi-nested RT-PCR

To further confirm the results obtained after the first round of PCR, semi-nested PCR was performed. 1µl of the RT-PCR product was further amplified for 30 cycles PCR reaction using a gene-specific reverse primer (from the junction of exon 4-exon 5) and first-exon-specific primer as down and upstream primers respectively. This was followed by a second round of PCR utilizing
1µl of the PCR product obtained above and same upstream primer while using a different downstream primer designed from the sequence located internally to the first reverse primer (from exon 4). The amplification was carried out for 30 cycles as above. 10µl of the PCR product was then subjected to 1.2% (w/v) agarose gel electrophoresis, stained with ethidium bromide and photographed on a UV transilluminator.

### 6.2.6. Subcloning and sequencing of RT-PCR products

The resulting bands of RACE, RT-PCR and semi-nested PCR products were fractionated and excised from 1.2% agarose gel. These products were purified using 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 using plasmid purification kit. Plasmids containing the insert were sequenced in an automatic sequencer using either M13 forward or reverse primers [Sanger F et al. 1977].

### 6.2.7. Bioinformatic analysis

Homology and similarity searches of the obtained nucleotide sequences were performed using the BLASTN nonredundant 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 [Altschul SF et al. 1997]. Various bioinformatics tools were further used to reach conclusive results. The gene coding for *htr4* along with 5’ UTR region in mouse was downloaded from Mouse Genome Informatics (MGI www.informatics.jax.org/). Some of the properties of the amino acid sequences coded by the new exons were analysed using ExPASy tools (http://ca.expasy.org/).
6.3. Results

6.3.1. Five new coding exons making four novel alternatively spliced transcripts of htr4 gene

The genomic DNA of htr4 gene is 145 kb and is reported to contain 6 exons and 5 introns. adrb2 gene is located upstream of 5htr4 (Figure 6.2A). A single transcript containing all six exons was identified earlier (accession number NM008313) encoding HTR4 protein [Claeysen S et al. 1996] as shown in Figure 6.2B. Later, four transcript variants were identified that arise due to alternative splicing of the 3′ end exons making four HTR4 protein isoforms having different C-terminals [De Maeyer JH et al. 2008]. In order to detect and identify the transcripts variants of htr4 utilizing different 5′ exons and thereby producing variants that differ at N-terminals 5′ RACE was performed. RNA isolated from the heart of mouse was amplified by 5′ RACE (Figure 6.3A) and sequenced as discussed in materials and method. 5′ RACE product from heart when electrophoresed on 1.2% agarose gel displayed a single band (Figure 6.3A). Sequence analysis of the product revealed that it codes for cDNA with a different sequence at the 5′ end from the reported sequence. This confirmed a novel single variant of htr4 from the heart of the mouse encoding new N-terminal. The reported transcript having first exon E1 was not amplified from the heart even after several independent experiments suggesting that heart does not contain the transcript coding for the prototype HTR4 receptor. This result is in accordance with the previous studies where htr4 transcript is not found in the heart of small laboratory animals [De Maeyer JH et al. 2008].

However several bands were observed when RNA was used from mouse brain (Figure 6.3B). These bands were excised, purified and sub-cloned in E. coli JM109 strain using TOPO vector. Bacterial clones were screened through colony PCR using the primers spanning the multiple cloning sites. Both the strands were sequenced for the clones having inserts in the plasmid. The
sequences thus obtained were further analyzed for homology searches using the nonredundant database BlastN and multiple sequence alignment using ClustalW to find if these sequences were part of \textit{htr4} gene. Four new transcripts were identified having new sequences present at the 5' ends in addition to the one reported earlier. Homology searches revealed that these new sequences were derived from 5' UTR and 1\textsuperscript{st} intron of the gene. This result was further confirmed by aligning the new 5' ends of the nucleotide sequences with that of the 5' UTR and 1\textsuperscript{st} intron sequence of \textit{htr4} gene. Five new exons were identified from 5' UTR and 1\textsuperscript{st} intron of the gene. The new exons designated as N1 and N2 are located in the 5' UTR while N3, N4 and N5 arise from the 1\textsuperscript{st} intron of the gene (Figure 6.2B). N1 splices with N2 which then splices with exon E2 of the prototype transcript maintaining the reading frame and producing a transcript designated as T1. Similarly exon N3 first splices with exon N4 which then splices either with E2 or E3 to produce two new transcript variants designated as T2\textsubscript{l} (T2 large) and T2\textsubscript{s} (T2 small) respectively. Exon N5 splices with E2 to produce a transcript variant T3. The reading frame was maintained in all these transcript variants and splicing pattern of these exons are depicted in Figure 6.4. EST searches did not give positive hit with these new sequences.

\textbf{6.3.2. Expression of four new transcripts of \textit{htr4} in brain and heart of mouse}

The results obtained from 5' RACE were further confirmed through RT-PCR and semi-nested PCR. New exon sequences identified by 5' RACE were used to design first exon-specific primers and were used in combination with either reverse primer (M5HTR4REV1) to perform RT-PCR or internal reverse primer (M5HTR4REV2) for semi-nested PCR. The anticipated size of semi-nested PCR products of transcripts when amplified with the forward primer from first exon in combination with internal reverse primer M5HTR4REV2 is summarized in Table 6.1. RT-PCR was done as described in materials and methods section using 2µg of total RNA isolated from the mouse brain, first
exon specific primer and M5HTR4REV1 as upstream and downstream primers respectively. 1µl of this RT-PCR product was further amplified in a semi-nested PCR to remove non-specific amplification using a new internal downstream primer M5HTR4REV2. Amplified products were fractionated by agarose gel electrophoresis as shown in Figure 6.5A. It was interesting to observe the bands of anticipated sizes. This experiment was performed three times independently and amplification of the new transcripts along with the previously published transcript was observed. To further confirm the products, DNA bands were excised from the gel and sequenced as described earlier. Sequencing data and their analysis supported the results obtained by 5′ RACE. All these new sequences were finally submitted to the GenBank with their accession numbers JN418927, JN418928, JN418929 and JN418930 for T1, T2l, T2s and T3 respectively.

To understand the expression profile of transcript variants in the heart of mouse, we performed RT-PCR and semi-nested PCR amplification using RNA isolated from heart as was done for mouse brain mentioned above. Amplified products were electrophoresed on agarose gel electrophoresis as shown in Figure 6.5B. It was fascinating to observe the amplification of T3 transcripts of anticipated size. We repeated this experiment five times independently and every time only T3 transcript was observed. Amplification of other transcripts was not seen in heart. Further to this, DNA band corresponding to T3 transcript was confirmed by sequencing. RT-PCR and semi-nested PCR clearly demonstrated the expression of all four newly identified transcripts and the transcript identified earlier in brain whereas the expression of only T3 transcript was seen in heart as well. However, the search for ESTs corresponding to these transcripts failed to hit any positives. Therefore, our experiments provide the first experimental evidence for the existence of transcripts with novel 5′ end sequences.
6.3.3. An *in silico* analysis of the splice junction and deduced amino acid sequence of the new transcript variants

The genomic sequence of *htr4* was studied to find out if the donor-acceptor residues of the new exons are conserved or not. The donor “GT” dinucleotide is conserved in all the newly identified exons and the acceptor “AG” is conserved in all the newly identified and previously reported exons except an acceptor “GG” dinucleotide present at the 5′ end of exon E6 as depicted in Figure 6.6.

In order to understand the functional significance of these new alternatively spliced transcripts, the deduced amino acid sequence corresponding to the variants were analysed using some bioinformatics tools, the most common web-based ExPASY tool, which predicts some of the features related to the peptide sequence. The analysis provided a number of differences in the conceptual proteins encoded by four new transcripts and is summarized in Table 6.2. Multiple sequence alignment using web-based ClustalW of the HTR4 protein sequences from mouse showed that this receptor is highly homologous across the species Figure 6.7 and the N-terminals of these sequences are highly comparable. The differences are majorly confined to the C-terminals. ClustalW alignment of conceptually translated amino acid sequences of these new transcripts revealed that they differ in their first two exons while the sequences from third exon is identical (Figure 6.8).
Figure 6.2: htr4 gene and transcript variants.

(A) Genomic organization of htr4 gene in mouse: The gene is located on chromosome 18 with a large 5' UTR of almost 145 kb in size. adrb2 is located upstream of htr4 gene.

(B) Known transcript of htr4. Earlier known transcript for this gene with accession number NM008313 consisted of six exons. Exons are shown in filled barrels; connecting straight lines are introns. The positions of primers are shown above the exons. Size of introns and exons are not to scale.
Figure 6.3: Agarose gel electrophoresis of 5' RACE products.

(A) Agarose gel electrophoresis of 5' RACE product from heart. The amplified products from the heart of the mouse were electrophoresed on 1.2% gel and stained with ethidium bromide. Lanes are marked as M, -RT and +RT where M stands for 100bp DNA ladder, -RT for RACE products in absence of reverse transcriptase, +RT denotes RACE products in presence of reverse transcriptase.

(B) Agarose gel electrophoresis of 5' RACE product from brain. The amplified products were electrophoresed on 1.2% gel and stained with ethidium bromide. Lanes indicate M, -RT and +RT where M stands for 100bp DNA ladder, -RT for RACE products in absence of reverse transcriptase, +RT denotes RACE products in presence of reverse transcriptase.
Figure 6.4: Splicing pattern of all the newly identified transcripts.

(A) Exon-intron organization of htr4. All exons reported earlier (filled barrel) and newly identified in this study (open barrel) are shown with interconnecting line as intron. The 1st intron is exceptionally large 71119 bp in size while the 2nd, 3rd, 4th and 5th intron are 1359 bp, 14493 bp, 9360 bp and 27051 bp respectively. (B) T1 transcript variant. The newly identified exon N1 splices with other newly identified exon N2 which then splices with the published E2. (C) T2l transcript variant. N3 splices with N4 which then splices with E2. (D) T2s transcript variant. N3 splices with N4 which then splices with E3. (E) T3 transcript variant. N5 exon splices with E2 to produce a new transcript variant. The splicing pattern of the exons is shown by bent lines. The positions of the primers are shown by arrows above the exons. Sizes of introns and exons are not to scale.
Figure 6.5: Agarose gel electrophoresis of the semi-nested PCR products. Amplified products were electrophoresed using 1.2% agarose gel. 100 bp DNA ladder was run as standard DNA marker and labeled as ‘M’ for all the gels.

(A) Total RNA isolated from brain was used as template for RT-PCR and semi-nested PCR amplified with reverse primer M5HTR4REV2. Lanes 1-4 show the amplified products (shown with arrowheads) using first exon specific forward primer as summarized in table 1. Lane 1 indicates the transcript T1 (630 bp), lane 2 for transcript T2l (615 bp) and T2s (489 bp), lane 3 for the transcript published earlier (513 bp) as positive control and lane 4 for transcript T3 (502 bp).

(B) Total RNA isolated from brain was used as template for RT-PCR and semi-nested PCR amplified with reverse primer M5HTR4REV2 with forward primer from 1st exons. Lane 1: amplified product corresponding to T3 transcript (502 bp). Lane 2: published transcript. The amplification of transcripts T1, T2l, T2s and published earlier failed from heart. The products were further confirmed by sequencing as described earlier.
Figure 6.6: Splice donor-acceptor sites of first few exons of the transcripts.

(A) The published transcript of *htr4*; (B) identified T1 transcript; (C) T2l transcript; (D) T2s transcript; (E) T3 transcript. The junction donor “GT” residues are conserved in all the exons. The acceptor “AG” residues are conserved in all the newly identified exons as well as previously published exon with the exception that the acceptor residue at the 5’ end of exon 6 is “GG”. Name of the exons is denoted above the sequence and the junctional residues are represented as bold and capitalized font.
Figure 6.7: Multiple sequence alignment (ClustalW) of the amino acid sequences from mouse, rat, humans and porcine.

The accession numbers for the amino acid sequences of htr4 in database are: mouse, NP032339; rat, NP036985; human, NP001035259; porcine, NP001001267. Mouse sequence is 87%, 88% 96% identical to 5HTR4 sequences from humans, porcine and rat respectively.
### Figure 6.8: Multiple sequence alignment (ClustalW) of the deduced amino acid sequences of the published transcript (C) and the newly identified transcripts (T1, T2l, T2s and T3).

All sequences are identical from third exon onwards. Accession numbers are given at the end of each sequence. (*), colon (:) and dash (-) represents identical, similar and no residue respectively.
Table 6.2: A comparative analysis of the amino acid sequence of the translated transcripts. The newly identified exons in combination with the internal published exon with which the new exon splices are analyzed in order to study the difference associated with the N-terminals variation of these isoforms. The analysis was done using the web-based ExPASy Tools.

<table>
<thead>
<tr>
<th>Properties</th>
<th>E1-E2 (C)</th>
<th>N1-N2-E2 (T1)</th>
<th>N3-N4-E2 (T2)</th>
<th>N3-N4-E3 (T2s)</th>
<th>N9-E2 (T3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amino acid (AA)</td>
<td>MDKLDANVS SNEGRFSVEK VVLLTLFLAV VILMAILGNL LVMVAVCRD RQLR</td>
<td>MGTLALW HSACPGV AASRTAGT PKLAGAVEP AEPLGVR SRRMLCVPS SNEGRFSV EKVLLTF LAVVILMA ILGNLILVM VAVCRD RQLR</td>
<td>MTLIKICNK AYNVLMQ L RASENFYT RASLMAQYC YCSTYEYR YSSNEGFRS VEKVVLLTF LAVVILMAI LGNLLVMV AVCRD RQLR</td>
<td>MTLIKICNKA YNVLMQLR ASENFYTK CSSLMAQYC EYSTYEYR YSSNEGFRS VEKVVLLTF LAVVILMA ILGNLLVM AVCRD RQLR</td>
<td>MHIIPKA VLSNRKLS YSFVESVE MVLSSNEG LTLTFLAVI</td>
</tr>
<tr>
<td>No. AA</td>
<td>51</td>
<td>47</td>
<td>86</td>
<td>111</td>
<td>62</td>
</tr>
<tr>
<td>pI</td>
<td>7.95</td>
<td>9.60</td>
<td>9.14</td>
<td>8.17</td>
<td>10.95</td>
</tr>
<tr>
<td>Mw</td>
<td>5661.83</td>
<td>9288.14</td>
<td>9854.76</td>
<td>12765.15</td>
<td>6935.56</td>
</tr>
<tr>
<td>Signal peptide</td>
<td>Signal anchor</td>
<td>Signal anchor</td>
<td>Non-secretory</td>
<td>Signal anchor</td>
<td></td>
</tr>
<tr>
<td>Acetylation</td>
<td>No</td>
<td>G(^2),(^{20,23})</td>
<td>T(^2)</td>
<td>T(^3)</td>
<td>No</td>
</tr>
<tr>
<td>Net-O-glyc</td>
<td>No</td>
<td>No</td>
<td>T(^4),(^{20,23})</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Net-N-glyc</td>
<td>NVS</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Ser:2, Thr:0, Tyr:0</td>
<td>MGLTALW HSACPGV AASRTAGT PKLAGAVEP AEPLGVR SRRMLCVPS SNEGRFSV EKVLLTF LAVVILMA ILGNLILVM VAVCRD RQLR</td>
<td>MTLIKICNK AYNVLMQ L RASENFYT RASLMAQYC YCSTYEYR YSSNEGFRS VEKVVLLTF LAVVILMAI LGNLLVMV AVCRD RQLR</td>
<td>MHIIPKA VLSNRKLS YSFVESVE MVLSSNEG LTLTFLAVI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net PhosK PKC at T(^{24})</td>
<td>MGLTALW HSACPGV AASRTAGT PKLAGAVEP AEPLGVR SRRMLCVPS SNEGRFSV EKVLLTF LAVVILMA ILGNLILVM VAVCRD RQLR</td>
<td>MTLIKICNK AYNVLMQ L RASENFYT RASLMAQYC YCSTYEYR YSSNEGFRS VEKVVLLTF LAVVILMAI LGNLLVMV AVCRD RQLR</td>
<td>MHIIPKA VLSNRKLS YSFVESVE MVLSSNEG LTLTFLAVI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PKC at T(^{62})</td>
<td>PKC at T(^{59})</td>
<td>PKC at T(^{59})</td>
<td>PKA at S(^{91})</td>
<td>PKC at T(^{35})</td>
<td></td>
</tr>
</tbody>
</table>

The position of amino acid undergoing post-translational modification such as acetylation, O-glycosylation, N-glycosylation and phosphorylation by specific kinases are indicated as superscript.

The seronin (S), threonine (T) and tyrosine (Y) residues which are predicted as the phosphorylation sites are underlined and are in bold font.
6.4. Discussion

htr4 gene encodes for 5 hydroxytryptamine (serotonin) receptor 4 in mouse located on chromosome 18 and spans 145 kb. The genomic structure analysis of htr4 showed that adrb2 is located 145 kb upstream of the first coding exon E1 of htr4 gene hence the 5' UTR of this gene is very large (Figure 6.2A). Mouse database searches revealed that this subunit is coded by a single gene consisting of 6 exons and 5 introns. cDNA corresponding to htr4 is present in the database that encodes for a protein consisting of 388 amino acids (Figure 6.2B). The 1st intron is relatively large (71119 bp) in size while the 2nd, 3rd, 4th and 5th intron are 1359 bp, 14493 bp, 9360 bp and 27051 bp respectively (Figure 6.4A). Sequencing data from 5' RACE as well as RT-PCR clearly demonstrated the identification of four new transcripts T1, T2l, T2s, and T3 not reported earlier. Expression studies revealed that all the four transcripts along with the variant published earlier were present in mouse brain. Though the transcripts encoding HTR4 receptor was reported from the heart of various organisms including human, porcine and rat where they play important role but no such transcript was detected earlier from mouse heart [De Maeyer JH et al. 2008]. Expression analysis for all four newly identified transcripts in mouse heart confirmed the presence of only T3 transcript. No other htr4 transcript variants were detected even after several repetitions. Thus for the first time we have identified a htr4 transcript from the heart of the mouse. All new variants differ in their 1st two exons making unique N-terminals for HTR4 variants. The complete transcript sequences including the 3' ends when expressed at the level of protein would produce receptor of 426 amino acids for T1, 423 for T2l, 381 for T2s, 399 for T3. T2s is 126 nucleotides shorter than T2l which is due to the skipping of exon 2 in T2s transcript. The donor-acceptor residues present at the splice site are conserved in all the newly identified and previously reported exons except the acceptor residue at 5' end of exon 6 (Figure 6.6). First published exon E1 can be replaced by the new exons or their combinations in
alternative fashion increasing the number of variants of HTR4 and amplifying the diversity of the receptor.

Table 6.2 is a comparative analysis of the encoded amino acid sequences of 1st two exons in published and newly identified transcripts. T2l and T3 contain signal anchor sequence similar to that of the published sequence. However T1 contains a signal peptide making it secretory protein whereas T3 is a non-secretory protein. Signal anchor sequences are non-cleavable sequences which facilitate the anchoring of the protein to the membrane [Berntd U et al. 2009; Martoglio B and Dobberstein B 1998; Ng DT et al.1996]. While most of the receptor protein and secreted proteins contain a cleavable signal peptide which guide them to endoplasmic reticulum and are further transported through golgi apparatus and incorporated into the membrane. T1, T2l and T2s variants contain a putative acetylation residue. The exact function of N-terminal acetylation is not yet known, though recently it is proposed to participate in protein turnover and possibly homeostasis thereby serving as a degradation signal [Hwang CS et al. 2010]. The published sequence has a potential N-terminal N-glycosylation site (\textsuperscript{7}NVS) and T1 variant is predicted to undergo O-glycosylation at threonine residues at position 4, 20 and 23. It is predicted that the other variants do not have potential O- or N-glycosylation sites. All the variants except T2s reported to have a threonine residue which can be potentially phosphorylated by PKC. T2s has a serine residue that can be phosphorylated by PKA. PKA- mediated phosphorylation of \( \beta_2\)-AR switches its predominant coupling from Gs to Gi [Zamah AM et al. 2002]. Therefore, phosphorylation of GPCRs modulates their signaling pathways. Properties summarized in Table 6.2 are suggestive of large diversity among these variants. They share some of the properties while differ in many thus increasing the molecular diversity of the receptor. Amino acid sequence of mouse HTR4 is 87%, 88% and 96% identical to HTR4 sequence from humans, porcine and rat respectively (Figure 6.7). The N-terminals of these receptors are highly similar while they diverge in sequences at their C-terminals. The high homology of the
receptors suggests that our studies in mouse can be extrapolated to humans as such studies in humans have not been reported so far. The presence of a unique transcript in heart with a novel N-terminals can serve as potential therapeutic target. The published protein sequence of HTR4 and the deduced amino acid sequences of all the transcripts were also aligned. These sequences are identical from 3rd exon onwards differing at the N-terminal only (Figure 6.8). As the protein sequences of HTR4 receptor are highly conserved across the species, there are chances for the presence of similar variants in humans, porcine and rats. The existence of multiple variants will further increase our therapeutic targets while designing drugs for cardiovascular disorders, depression and various other neuropathologies involving HTR4 receptors.

Our studies provided the first experimental evidence for the existence of a new transcript T3 expressed in mouse heart. This sequence is not identical to the previously reported sequence. This transcript variant was also detected in brain. Three additional variants T1, T2l and T2s were reported from brain which differs from the published sequence in first two exons. No N-terminal variants were reported earlier for this gene despite the presence of long 5' UTR region and 1st intron. These variants might be important for differential expression of the HTR4 receptor. Their assembly and co-assembly with β2-AR might be governed by these new N-terminals. This may also facilitate the post-transcriptional regulation of the two receptors. A detailed study of these variants is essentially required to understand the varied function of HTR4 receptor and understand their mechanism of action.