Characterization of a novel transcript variant of human \textit{STAU1} gene*

Nie Fa-hui$^{1,2}$, Yao Hai-feng$^1$, Qi Rui-feng$^3$, Li Xin$^4$*, and Wu Cai-bin$^2$*

$^1$State Key Laboratory of Pollution Control and Resource Reuse, Tongji University, Shanghai, PR China; $^2$Institute of Environment Engineering Research, East China Jiao Tong University, Nanchang, PR China; $^3$Institute of Protein Research, Tongji University, Shanghai, PR China; $^4$Institute of Biomedical Science, Fudan University, Shanghai, PR China

Received: 09 March, 2008; revised: 11 August, 2008; accepted: 15 August, 2008
available on-line: 20 September, 2008

Human \textit{STAU1} is one member of the family of double-stranded RNA (dsRNA)-binding proteins. It is thought to function in transporting mRNA, controlling translation and eliciting mRNA decay in neurons, and to function in infection of influenza virus and human immunodeficiency virus type 1 (HIV-1). Four transcripts coding two isoforms have been identified before. In this study, we have isolated a novel transcript of \textit{STAU1}, coding a novel isoform that has six amino acids more (SFPLKQ) than isoform a. In order to examine the tissue distribution of this novel isoform, we have performed RT-PCR experiments and the analysis showed that it was highly expressed in heart, liver, kidney and pancreas.

Keywords: \textit{STAU1}, expression pattern

INTRODUCTION

\textit{Drosophila} Staufen, homologous human \textit{STAU1}, is the first RNA-binding protein proven to play a role in RNA localization. Staufen and its homologous genes play a key role in many biological phenomena. One of the best known examples is determination of the anteroposterior axis during the early development of \textit{Drosophila melanogaster}. Staufen plays a role in the transportation and location of \textit{bicoid} and \textit{oskar} mRNAs to the anterior and posterior poles, respectively (Schuldt \textit{et al.}, 1998). The specific localization of mRNAs is also essential in the definition of cell asymmetries in cell division and differentiation. Staufen transports \textit{prospero} mRNA to basal area of fly mitotic neuroblasts (Li \textit{et al.}, 1997). Cell communication events are also affected by the localization of specific mRNAs. Human \textit{Stau1} can interact with many proteins and locates in the dendrites of differentiated neuroblasts (Kiebler & DesGroseillers, 2000). Human \textit{STAU1} functions in infection by influenza virus and human immunodeficiency virus type 1 (HIV-1). It can interact with NS1 protein of influenza virus in infected cells (Falcón \textit{et al.}, 1999), and it can also interact with the NC domain of HIV-1 p55Gag and may be make the virus generate infectious viral particles (Chatel-Chaix \textit{et al.}, 2004). In addition to these biological processes, \textit{Drosophila} Staufen function in the translation derepression of \textit{oskar} mRNA once the mRNA has been located (Micklem \textit{et al.}, 2000). Human \textit{STAU1} causes the decay of many mRNAs through binding Upf1 and interacting with the 3' untranslated regions of mRNAs (Kim \textit{et al.}, 2005).

\textit{Drosophila} Staufen has five double-stranded RNA-binding domains (dsRBD) (Gibson \textit{et al.}, 1994). Each domain contains a 65- to 68-amino-acid consensus sequence and folds into a compact \textalpha\beta\beta\alpha\beta structure. The analysis of crystal structure shows that dsRBD recognizes the shape of A-form dsRNA through two regions between \textbeta2-\textbeta1 and \textbeta3-\textalpha2 (Ramos \textit{et al.}, 2000). The composition of Staun-containing RNA has been identified in neuroblasts, including ribosomes and ribosome components (ribosomal protein S6, L12, FMRP, PABP,.

Corresponding authors: Li Xin, lxstaroad@yahoo.com or Wu Cai-bin, wucabinn@ecjtu.jx.cn

*The nucleotide sequence reported in this paper has been submitted to GenBank under accession number AY536098.

\textbf{Abbreviation:} dsRBD, double-stranded RNA-binding domain; G3PD, glyceraldehyde-3-phosphate dehydrogenase gene; NS1, nonstructural protein; RT-PCR, reverse transcriptional PCR; Stau, Staufen protein.
S40, S60 etc); proteins from the cytoskeleton, motor proteins and regulatory proteins (tubulins, actin, internexin, myosin, IQGAP-1, cdc42, rac1, RasGAP, dynein, kinesin etc.); and other RNA-binding proteins normally localized to the nucleus (nucleolin, RNA helicase A, hnRNP U etc.) (Villace et al., 2004). In addition, STAU1 can bind protein phosphatase-1 in primary hippocampal neurons (Monshausen et al., 2002), Upf1 and Arf1 proteins in HeLa cells, NS1 protein of influenza virus and pr55<sup>gap</sup> protein of HIV-1 (Falcon et al., 1999; Chatel-Chaix et al., 2004; Kim et al., 2005). Therefore, the component is very complex and human Staun may be plays vital roles.

Four transcript variants resulting from human STAU1 gene and encoding two isoforms (a, b) with different N-terminal ends have been described (Marión et al., 1999). Three of these variants encode the same isoform (a), however, differ in their 5' UTR. In contrast to Drosophila Staufen, human STAU1 has only four dsRBDS. However, the significance of the various transcripts is not known today. In a recent study, we have isolated a novel transcript variant of STAU1 through a large-scale sequencing analysis of the 18-week human fetal brain cDNA library. We termed it STAU1 transcript variant T5 (isoform c). Here, we report the cloning and initial characterization of this new transcript variant, together with data in its genomic structure and mRNA tissue distribution.

**MATERIALS AND METHODS**

The cloning and bioinformatics analysis. cDNA library construction and DNA sequencing were performed as previously described (Li et al., 2005). DNA sequence homology searches and comparisons were performed using BLAST-N and BLAST-X at the National Center for Biotechnology Information (NCBI) network service (http://www.ncbi.nlm.nih.gov/blast). BLAST-N searching in the human genome was performed to identify the chromosomal localization and the gene structure of this novel gene. The predicted amino-acid sequence of the STAU1 T5 was compared against the profile entries to find the occurrence of known profiles (http://www.expasy.ch/pfscan). To identify the chromosomal localization and the gene structure, multiple alignments were performed by the GeneDoc program (http://www.psc.edu/biomed/genedoc/). Other databases from Genbank, Swiss Pro, PDB, and EXPASY were also used. The software includes Gene Runner, Primer Premier 5.0 et al.

Assessment of human STAU1 T5 mRNA tissue distribution. Human Multiple Tissue cDNA (MTC) panels (CLONTECH) were used as PCR templates according to the manufacturer’s protocol. Thirty six PCR cycles for STAU1 T5 and 30 PCR cycles for G3PDH (as control) were performed using Taq Plus polymerase (Sangon) in the following program: 0.5 min at 94°C, 0.5 min at 65°C, 1.0 min at 72°C. The PCR products of STAU1 T5 and G3PDH were then electrophoresed on a 2% agarose gel. PCR primers are indicated from 5' to 3' as the following: STAU1 T5 sense CTCTCTCGGCTCCCGTTCCTTT; STAU1 T5 antisense ACCGTGGTCAAGGGAAAAGCTCG; G3PDH sense: TAGAAGGTCCGGAATCAGGATTGGT; G3PDH antisense CATGTGGGCCCATGAGGTCCACCAC. The sense and antisense primers of human STAU1 T5 span 688 bp in the cDNA from 61 to 748 bp.

**RESULTS AND DISCUSSION**

Through the large-scale cDNA sequencing, we cloned a novel transcript variant from the constructed human fetal brain cDNA library. This cDNA is composed of an open reading frame from nucleotide 363 to 1868, encoding a 502 amino-acid protein with a molecular mass of 55 kDa (Genbank: AY546099) (Fig. 1). The deduced protein shows high identity with the mouse and Drosophila Staufen2 proteins. Alignment was performed between STAU1 T5, other two isoforms in human, and its orthologs from mouse and Drosophila. Like other Staufens, the deduced protein of T5 also contains four dsRNA-binding domains (dsRBDS) (Fig. 2). Universally, dsRBD2 contains an insertion sequence at an identical position in all species, and the domain does not bind RNA when this insertion is deleted. Surprisingly, we found another insertion of 18 bp in the sequence, which is deduced into six amino acids between dsRBD1 and dsRBD2. This new insertion may suggest a new kind of mechanism or method of RNA localization during the cell cycle or embryo development. Moreover, through ortholog analysis (not shown), the insertion of the 18 bp sequence is also present in certain Staufen protein in Drosophila, which suggests that this protein may be conserved in evolution and that this insertion may have something to do with the specificity of its binding in different tissues of human and Drosophila. However, on its own Staufen binds mRNA without apparent specificity, indicating that the specificity might be provided by auxiliary factors.

Four transcript variants resulting from alternative splicing of the STAU1 gene and encoding two isoforms with different N-terminal ends have been reported. Three of these variants encode the same isoform, however, differ in their 5' UTR.
Novel transcript variant of human STAU1 gene
Figure 1. The transcript variant cDNA and deduced amino-acid sequences of STAU1 T5.

The nucleotides are numbered at the left. The extra amino-acid residues of Stau1 T5 compared to other isoforms are boxed.
Figure 2. Alignment of STAU1 c and a (NP_004593), b (NP_059347) and its orthologs in mouse Stau1 (NP_035620) and Drosophila (NP_476751).

The alignment was performed by GeneDoc program (http://www.psc.edu/biomed/genedoc/): Black (100% similarity), grey (80–90% similarity); light grey (60–70% similarity). Domains of the 5 dsRBDs are underlined.

Table 1. The exon-intron analysis of the human STAU1 T5 gene.

<table>
<thead>
<tr>
<th>Exon Number</th>
<th>Length (bp)</th>
<th>Splicing acceptor</th>
<th>Splicing donor</th>
<th>Intron Number</th>
<th>Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>249</td>
<td>tc...tcag</td>
<td>CAGCAGCCCAggtggggtgc</td>
<td>1</td>
<td>13846</td>
</tr>
<tr>
<td>2</td>
<td>75</td>
<td>tattttcag</td>
<td>TTTAAGAGGgttataa</td>
<td>2</td>
<td>20123</td>
</tr>
<tr>
<td>3</td>
<td>139</td>
<td>tttccctag</td>
<td>ATCCCCCAgagtttgttt</td>
<td>3</td>
<td>2185</td>
</tr>
<tr>
<td>4</td>
<td>166</td>
<td>ttatactag</td>
<td>GAGGCTGGGgagggatttc</td>
<td>4</td>
<td>15650</td>
</tr>
<tr>
<td>5</td>
<td>99</td>
<td>atttctag</td>
<td>GAATTGCGGgtaggtaat</td>
<td>5</td>
<td>11227</td>
</tr>
<tr>
<td>6</td>
<td>231</td>
<td>ttatcactag</td>
<td>CATATCGCAAgtggiaat</td>
<td>6</td>
<td>1139</td>
</tr>
<tr>
<td>7</td>
<td>144</td>
<td>ttgtcctag</td>
<td>TGTGATCGAggtggcgc</td>
<td>7</td>
<td>2963</td>
</tr>
<tr>
<td>8</td>
<td>147</td>
<td>cttgtgcag</td>
<td>AGAGAGAAGGgttgtgct</td>
<td>8</td>
<td>1573</td>
</tr>
<tr>
<td>9</td>
<td>76</td>
<td>cttttctag</td>
<td>ATGGGCGAAGgttgaag</td>
<td>9</td>
<td>232</td>
</tr>
<tr>
<td>10</td>
<td>320</td>
<td>tccctctccag</td>
<td>GGATCAGAGAgttactgtt</td>
<td>10</td>
<td>528</td>
</tr>
<tr>
<td>11</td>
<td>123</td>
<td>tcccccctag</td>
<td>CCAATGACTGAGgttgaag</td>
<td>11</td>
<td>1258</td>
</tr>
<tr>
<td>12</td>
<td>86</td>
<td>tccctcaag</td>
<td>ACCAAGGCGAGgttgaag</td>
<td>12</td>
<td>888</td>
</tr>
</tbody>
</table>

Intron and exon nucleotide sequences are shown in lower-case and upper-case letters, respectively. Bold italics lettering indicates donor and acceptor splice site.
Figure 3. Tissue distribution of human STAU1 T5 mRNA.
Reverse transcription-PCR analysis of human cDNA for STAU1 T5 and G3PDH (as a control). Prenormalized cDNAs from sixteen human adult tissues were purchased from CLONTECH and employed as templates in PCR reactions containing STAU1 T5 and G3PDH-specific primers described in Materials and Methods.

Isoform (a) associates with 40S and 60S ribosomal subunits and colocalizes with rough endoplasmic reticulum in neuroblasts (Wickham et al., 1999). Transcript variant T5 has an insertion of 18 bases at position 729 compared to other four transcript variants of STAU1. This introduces an insertion of six amino acids in the deduced protein, causing T5 to encode a protein isoform (c) of 502 amino-acids protein. So we termed it human Stau1 variant 5 and the deduced protein isoform (c). All the sequences of the exon-intron junctions are consistent with the AG-GT rule (Table 1).

Expression pattern of the STAU1 T5 was analyzed by RT-PCR. The data demonstrate that it is especially expressed in certain tissues we used. The expression level in heart, liver, kidney and pancreas are relatively high, while there is slightly lower expression in placenta, colon (Fig. 3). It is different from the other four transcripts that are expressed in heart, brain, pancreas, skeletal muscles, liver, placenta, lung, kidney (Wickham et al., 1999). Therefore, further studies will be necessary to define the precise roles of three isoforms.

REFERENCE


