

# Identification of alternative splicing form of Stat2

Tetsuya Sugiyama<sup>a,b</sup>, Yukihiro Nishio<sup>a</sup>, Tadamitsu Kishimoto<sup>a</sup>, Shizuo Akira<sup>a,\*</sup>

<sup>a</sup>Institute for Molecular and Cellular Biology, Osaka University, 1-3 Yamada-oka, Suita, Osaka 565, Japan

<sup>b</sup>Faculty of Pharmaceutical Sciences, Osaka University, 1-6 Yamada-oka, Suita, Osaka 565, Japan

Received 25 December 1995; revised version received 26 January 1996

**Abstract** We identified alternatively spliced forms of Stat2 in human and mouse mRNAs. The spliced forms are generated by reading through the intron between exon 20 and 21, which correspond to the region encoding SH2 domain. The spliced forms contain a stop codon in SH2 domain, and therefore give rise to a short form of Stat2 when the mRNAs are translated. The putative translated proteins lack half of the SH2 domain, the tyrosine phosphorylation site required for dimerization (or oligomerization) and DNA binding, and C-terminal activation domain. The significance of these spliced forms is discussed.

**Key words:** Stat2; IFN- $\alpha/\beta$  signaling; Alternative splicing; SH2 domain

## 1. Introduction

Cytokine receptor signals are mediated by Jak family of tyrosine kinases and Stats (signal transducers and activators of transcription) [1]. IFN- $\alpha$  signaling involves phosphorylation of Stat1 [2] and Stat2 [3,4] through Jak1 [5] and Tyk2 [6], leading to formation of the transcription factor IFN-stimulated gene factor 3 (ISGF3) [7,8] that binds to the IFN-stimulated response elements (ISREs). The ISGF3 transcription factor is composed of three subunits: Stat1 $\alpha/\beta$  [2], Stat2 and p48 DNA binding protein [9]. On the other hand, IFN- $\gamma$  induces formation of the  $\gamma$ -activated transcription factor (GAF) consisting of Stat1 $\alpha$  homodimer, which binds to the IFN- $\gamma$  activation site (GAS) elements.

A src homology 2 (SH2) domain of each Stat family member has a crucial role in dimerization (or oligomerization) of the activated Stat [10] as well as in recognition of the phosphorylation of the activated receptor [11,12]. Ligand-specific activation of Stats is shown to be determined by specific docking interaction between a Stat SH2 domain and a phosphorylation-containing sequence in the receptor.

Several alternatively spliced forms of Stat family are reported. Alternative splicing of Stat1 at 3' end results in Stat1 $\alpha$  (p91) and Stat1 $\beta$  (p84) [13,14]. Stat1 $\beta$ , which lacks C-terminal 39 amino acid residues of Stat1 $\alpha$  has no activity in IFN- $\gamma$  dependent transcription, although Stat1 $\alpha$  is able to mediate the transcription signal [15]. The Stat3 $\beta$  [16] which lacks C-terminal 55 amino acid residues of Stat3/APRF [17,18] but contains novel 7 amino acid residues at its C-terminus was recently cloned by use of two hybrid systems. Unlike Stat3, Stat3 $\beta$  can synergize with *c-jun* in transcription activity. Some other reports by use of Western blotting techniques suggested

the existence of other isoforms of Stat3 [19] and Stat5 [20]. However, there is no previous report about Stat2 (p113) isoform.

In order to identify novel members of Stat family expressed in mouse liver, RT-PCR (reverse transcription-polymerase chain reaction) amplification was performed using degenerative primers derived from the amino acid sequences highly conserved among Stat family members (data not shown). We report here the identification of alternatively spliced forms of Stat2 transcript.

## 2. Materials and methods

### 2.1. Cell culture and isolation of RNA

The human hepatoblastoma cell line Hep3B [21] and the human fibroblast cell line HeLa were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum, high glucose (4.5 g/l), penicillin, and streptomycin. HeLa cells were stimulated with IFN- $\alpha$  (1000 U/ml) (Collaborative Research Inc., Bedford, MA) for 0, 0.5, 2 and 6 h. The cells were washed with PBS (phosphate-buffered saline) and lysed with the denaturing solution (4 M guanidine thiocyanate, 25 mM sodium citrate (pH 7), 0.5% sarcosyl, 0.1 M 2-mercaptoethanol), and total RNAs at various time points were prepared by acid guanidine-phenol chloroform (AGPC) method [22]. The contamination of genomic DNA was eliminated by RNase-free DNase I (Stratagene, La Jolla, CA) treatment.

Mouse total RNAs were prepared from various tissues as described above.

### 2.2. RT-PCR analysis

First-strand cDNA was prepared from 1–2  $\mu$ g total RNA of various mouse tissues and the human cell lines using an oligo d(T)<sub>15–18</sub> primer and the Superscript II reverse transcriptase (Life Technologies Inc., Gaithersburg, MD). To amplify the mouse cDNA of Stat2 SH2 region, a sense primer m1L (5'-GGCATTACTTGTTCTGGGTG-GAGCACC-3') at exon 20 and an antisense primer m2L (5'-CACTTCCTGGTGTAGGGCTGCACTGAG-3') at exon 21 were used in RT-PCR. To amplify the human cDNA of Stat2, a sense primer h1 (5'-AGTTGGTACATGACCACC-3') at exon 19 and an antisense primer h2 (5'-AGCGCAGTGGGTTTCAG-3') at exon 21 were used in RT-PCR. To amplify only the alternatively spliced form of human Stat2, a sense primer h1 (described above) and an antisense primer hint (5'-AGTCCCAAACCAGCTGAG-3') at the normal intron were used in RT-PCR. All polymerase chain reactions (PCR) were performed in a final volume of 50  $\mu$ l containing cDNA aliquot corresponding to 25 or 150 ng of total RNAs, 20 mM Tris (pH 8.4), 50 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 1 mM of each dNTP, 40 pmol of each primer, and 1.25 U Taq polymerase (Takara Shuzo Inc., Japan). The PCR conditions to amplify human cDNAs were 35 cycles of 20 s at 94°C, 30 s at 62°C, and 1 min at 72°C by Gene Amp PCR System 9600 thermal cycler (Perkin Elmer, CT). The PCR conditions to amplify mouse cDNAs were 35 cycles of 20 s at 94°C, 30 s at 67°C, and 30 s at 72°C. All PCR reactions described in this study had an initial denaturing step of 1 min at 94°C and a terminal extension step of 7 min at 72°C. All PCR products were electrophoresed in a 2% agarose gel stained with ethidium bromide and photographed.

The amplified clones were isolated from the gels and subsequently cloned into T-vector (Novagen, Madison, WI). Sequencing of the subclones was performed by DNA autosequencer 373A (Applied Biosystems Inc., Foster City, CA) using dye termination method.

\*Corresponding author. Fax: (81) (6) 877-1955.

**Abbreviations:** IFN, interferon; Stat, signal transducers and activators of transcription; Jak, Janus kinase; ISGF3, IFN-stimulated gene factor 3.

### 2.3. Genomic organization of SH2 domain of Stat2

Mouse and human genomic DNAs were isolated by a general procedure [23] and subjected to PCR. To clone mouse genomic DNA, a sense primer m1L at exon 20 and an antisense primer m2L at exon 21 were used in PCR. To clone human genomic DNA, a sense primer h1L (5'-ATTACCTGCTCCTGGGTGGAGCACCAG-3') at exon 20 and an antisense primer h2L (5'-CACCTCCTTCGTGTACGGTTG-CACAGAG-3') at exon 21 were used in PCR. The PCR conditions to amplify human and mouse genomic DNAs were 35 cycles of 20 s at 94°C, 30 s at 68°C, and 1 min at 74°C after hot start by Cetus 480 thermal cycler (Perkin Elmer). Both PCR products are electrophoresed in a 1% agarose gel and isolated, subsequently the purified PCR products were subcloned into T-vector and sequenced as described above.

## 3. Results

### 3.1. Alternative splicing of Stat2 transcript

We identified alternatively spliced forms of Stat2 transcript by use of RT-PCR technique. Several extra bands were observed in PCR products derived from mouse tissue cDNAs using primers m1L at exon 20 and m2L at exon 21 (Fig. 1A). Each band was cloned and subjected to sequencing analysis. Comparison of the sequences of these PCR products with the genomic DNA sequence around the exon/intron boundaries demonstrated that mouse Stat2 existed in three alternatively spliced forms which we named type 'a', 'b', and 'c' (Fig. 2A). The major PCR product was derived from a normally spliced form of Stat2 (p113, type 'a'), two extra products from novel alternatively spliced forms of Stat2 (type 'b' and 'c'). However, both of the novel identified forms type 'b' and 'c' contain a translation stop codon at the same position which should yield the short form of Stat2. These novel identified mRNAs were distributed in various tissues as well as Stat2 type 'a' mRNA (Fig. 1A), although Stat2 type 'a' mRNA was expressed much abundantly in most of the tissues.

The extra band of Stat2 was also observed in PCR products derived from human hepatoblastoma cell line Hep3B cDNA using primers h1 at exon 19 and h2 at exon 21 (Fig. 1B). The

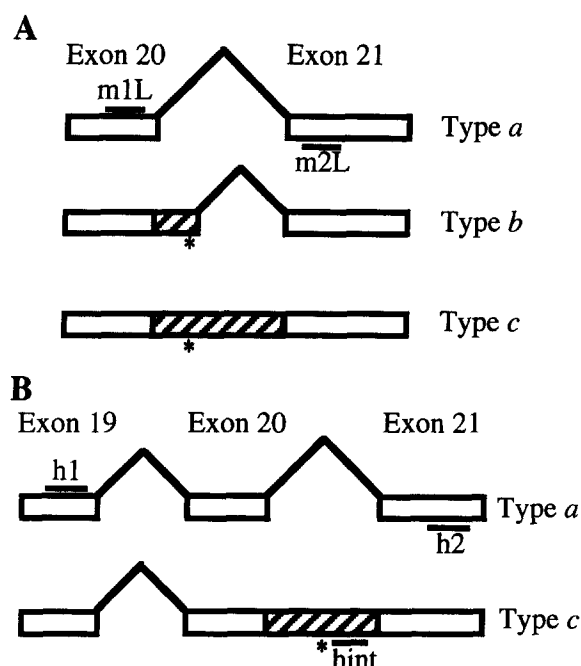


Fig. 2. (A) Alternative splicing of mouse Stat2 transcript. The open box represents an exon. The hatched box is considered an exon if it is not spliced out. Short bars indicate the position of primers m1L and m2L used for RT-PCR analysis. An asterisk (\*) indicates a translation stop codon. (B) Alternative splicing of human Stat2 transcript. Short bars indicate the position of primers h1, h2, and hint used for RT-PCR analysis. Primers h1 and h2 were used to amplify type 'a' and 'c' cDNAs. Primers h1 and hint were used to amplify only type 'c' cDNA.

sequencing of the extra band showed that it corresponded to mouse type 'c' mRNA (Fig. 2B). Human type 'c' mRNA was also detected in other cell lines such as fibroblast cell line HeLa and plasmacytoma cell line U266 (data not shown).

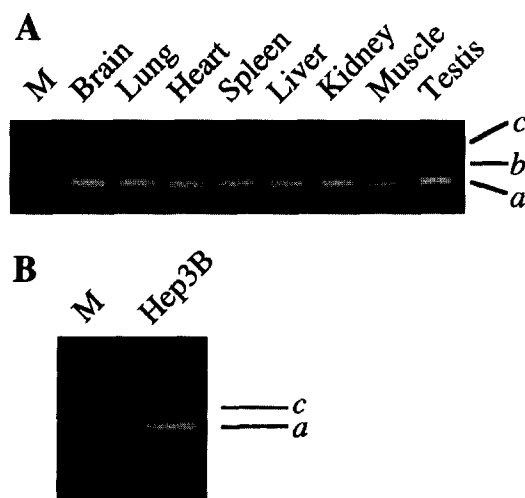


Fig. 1. (A) Tissue distribution of mouse Stat2 and the isoform mRNAs. Three alternatively spliced forms of Stat2 (band 'a', 'b' and 'c') were detected by use of RT-PCR. The size of the PCR products 'a', 'b' and 'c' were 80, 160, and 269 bp, respectively. M =  $\Phi$ X174 *Hae*III digest marker. (B) Two forms of human Stat2 were detected in Hep3B cell line by use of RT-PCR. An extra 'c' band was observed. The size of the PCR products 'a' and 'b' were 303 and 497 bp, respectively.

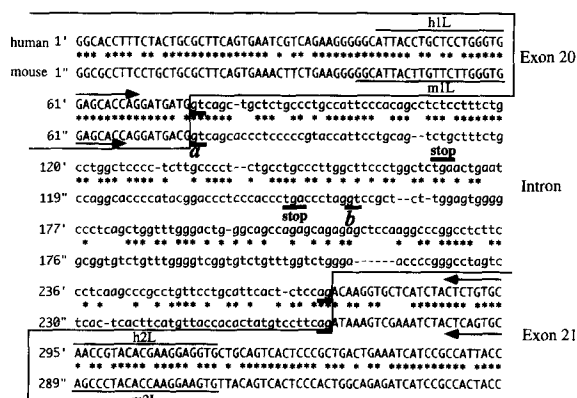


Fig. 3. Comparison of human and mouse Stat2 genomic DNA sequences around exons 20 and 21. Partial genomic DNA sequences between exons 20 and 21 are aligned, introducing gaps (—) to maximize nucleic acid identity. The squared sequences written by large characters represent exons 20 and 21 and the other sequences written by small characters represent an intron. Arrows indicate the sequence of primers used for cloning of genomic DNA. Under lines indicate splicing signals, GU-AG. Splicing signals 'a' and 'b' were used for type 'a' and 'b' in mouse, respectively.

### 3.2. The genomic structure and putative amino acid sequence of human and mouse Stat2 isoforms around exon 20 and 21

The human and mouse genomic DNA sequences were compared around exon 20 and 21 boundaries (Fig. 3). The sequences of exon 20 and 21 were highly conserved, while intron DNA sequences between exon 20 and 21 were 70% homologous. All the intron extremities follow the GT-AG rule [24]. The mouse intron DNA sequence contains two splice donor sites (a and b) and one splice acceptor site, whereas the human intron DNA sequence contains one pair of splicing signal. This difference of splicing signals accounts for the existence of three alternatively spliced forms in mouse and of two forms in human.

The putative amino acid sequences of Stat2 (Fig. 4A) and the short form (Fig. 4B) were compared between human and mouse. In the short form of Stat2, 231 C-terminal amino acid residues were replaced by novel 32 amino acids in human and 22 amino acids in mouse, respectively. The additional C-terminal sequences of the short form were not conserved, although normal Stat2 amino acid sequences were highly conserved between human and mouse.

### 3.3. Induction of Stat2 and the alternatively spliced messages

HeLa cells were stimulated by IFN- $\alpha$ . PCR amplifications were performed using two sets of primers because a large amount of Stat2 made it difficult to detect type 'c' mRNA. RT-PCR analysis using one set of primers h1 at exon 19 and h2 at exon 21 detected type 'a' mRNA and a faint band corresponding to type 'c' mRNA (Fig. 5A) whereas RT-PCR using the other set of primers h1 at exon 19 and hint at the intron between exon 20 and 21 detected only type 'c' mRNA (Fig. 5B). The type 'c' mRNA was induced in response to IFN- $\alpha$  as well as that of normal Stat2 mRNA.

## 4. Discussion

The alternatively splicing patterns of Stat2 described in this paper is different from those of the other Stat isoforms [13,14,16]. The previously identified Stat isoforms are variants which lack their C-terminal amino acid sequences, but still conserving the SH3 and SH2 domain and tyrosine phosphorylation site. On the other hand, novel identified Stat2 isoforms are internally spliced ones, which would generate short forms lacking crucial regions for the Stat signaling. Such a splicing is reported in Ash/Grb2 [25], which is the adapter molecule with two SH3 and one SH2 domain between protein tyrosine kinase receptors and Ras. One of the splicing isoforms of Ash/

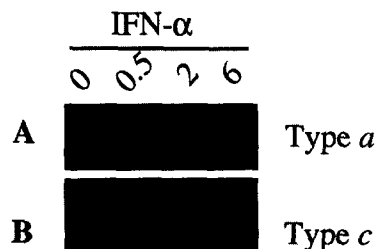


Fig. 5. Induction of human Stat2 in HeLa cell line in response to IFN- $\alpha$ . Two alternatively spliced forms of Stat2 were detected by RT-PCR. (A) Primers h1 and h2 were used to amplify the type 'a' and 'c' cDNA aliquots corresponding to 25 ng of total RNA. Only the type a products were clearly visible. (B) Primers h1 and hint were used to amplify only the type 'c' cDNA aliquots corresponding to 150 ng of total RNA.

Grb2 (Ash-s) generates a short form lacking SH2 and second SH3 domain. Ash-s containing first SH3 domain is considered to exert an inhibitory effect by competition [25,26].

The biological role of the predicted short forms derived from alternatively spliced Stat2 message is unclear. The Stat1 and Stat2 lacking the SH2 domain and/or the tyrosine phosphorylation site are unable to form of ISGF3 complex [10], suggesting that the predicted short form could not participate in the formation of ISGF3 complex. The short form conserving N-terminal half region might exert an inhibitory effect as well as Ash-s does.

We have tried to detect the predicted short forms employing western blotting techniques using a polyclonal antibody (Santa Cruz Biotechnology Inc., Santa Cruz, CA) that recognize the N-terminal site of human Stat2, but we could not ensure. Stat3 $\beta$  mRNA is reported to be expressed in a low amount and detected only by use of RT-PCR [16] like our case. Although the predicted short form of Stat2 would be expressed in low amount, it is possible that the protein has any function like Stat3 $\beta$  does.

We further investigated alternate splicing of Stat2 in other regions and obtained implication of existence of a few more isoforms around exon 15 and 18, which appear to be derived by skipping of the intervening exons (data not shown). Moreover, exon-intron organizations of Stat2 and Stat3 are quite similar and the boundary of exon 20 and exon 21 is located at the same position of amino acid sequences. However, Stat3 was not alternatively spliced.

In conclusion, we demonstrated that a small part of Stat2 transcripts is alternatively spliced, which would generate a short form lacking half of the SH2 domain, tyrosine phosphorylation site, and C-terminal activation domain. The precise role of the short form needs to be further elucidated.

**Acknowledgements:** We would like to thank Ms. K. Kubota for her secretarial assistance. This work was supported in part by a Grant-in-Aid from the Ministry of Education, Science and Culture, Japan.

## References

- [1] Ihle, J.N. (1995) Science 377, 591–594.
- [2] Schindler, C., Fu, X.-Y., Imbrota, T., Aebersold, R. and Darnell Jr., J.E. (1992) Proc. Natl. Acad. Sci. USA 89, 7836–7839.
- [3] Fu, X.-Y., Schindler, C., Imbrota, T., Aebersold, R. and Darnell Jr., J.E. (1992) Proc. Natl. Acad. Sci. USA 89, 7840–7843.
- [4] Leung, S., Qureshi, S.A., Kerr, I.M., Darnell Jr., J.E. and Stark, G.R. (1995) Mol. Cell. Biol. 15, 1312–1317.
- [5] Muller, M., Briscoe, J., Laxton, C., Gushin, D., Ziemiecki, A.,

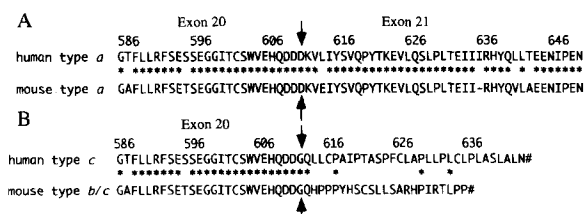


Fig. 4. Comparisons of human and mouse Stat2 putative amino acid sequences. (A) Putative amino acid sequences of human and mouse Stat2 around exons 20 and 21 were compared. Conserved sequences are indicated by an asterisk (\*) and a translation stop codon is indicated by a sharp (#). An arrow indicates the boundary between exon 20 and the following intron if it is spliced out. (B) Putative amino acid sequences of human and mouse Stat2 short forms were compared.

- Silvennoinen, O., Harpur, G., Barbieri, G., Witthuhn, B.A., Schindler, C., Pellegrini, S., Wilks, A.F., Ihle, J.N., Stark, G.R. and Kerr, I.M. (1993) *Nature* 366, 129–135.
- [6] Velazquez, L., Fellous, M., Stark, G.R. and Pellegrini, S. (1992) *Cell* 70, 313–322.
- [7] Darnell Jr., J.E., Kerr, I.M. and Stark, G.R. (1994) *Science* 264, 1415–1420.
- [8] Schuai, K., Horvath, C.M., Huang, L.H., Qureshi, S.A., Cowburn, D. and Darnell Jr., J.E. (1994) *Cell* 76, 821–828.
- [9] Veals, S.A., Schindler, C., Leonard, D., Fu, X.-Y., Aebersold, R., Darnell Jr., J.E. and Levy, D.E. (1992) *Mol. Cell. Biol.* 12, 3315–3324.
- [10] Heim, M.H., Kerr, I.M., Stark, G.R. and Darnell Jr., J.E. (1995) *Science* 267, 1347–1349.
- [11] Waksman, G., Shoelson, S.E., Pant, N., Cowburn, D. and Kuriyan, J. (1993) *Cell* 72, 779–790.
- [12] Pawson, T. and Gish, G.D. (1992) *Cell* 71, 359–362.
- [13] Schindler, C., Fu, X.-Y., Improta, T., Aebersold, R. and Darnell Jr., J.E. (1992) *Proc. Natl. Acad. Sci. USA* 89, 7836–7839.
- [14] Yan, R., Qureshi, S., Zhong, Z., Wen, Z. and Darnell Jr., J.E. (1995) *Nucleic Acids Res.* 23, 459–463.
- [15] Shuai, K., Stark, G.R., Kerr, I.M. and Darnell Jr., J.E. (1993) *Science* 261, 1744–1746.
- [16] Schaefer, T.S., Sanders, L.K. and Nathans, D. (1995) *Proc. Natl. Acad. Sci. USA* 92, 9097–9101.
- [17] Akira, S., Nishio, Y., Inoue, M., Wang, X.J., Wei, S., Matsusaka, T., Yoshida, K., Sudo, T., Naruto, M. and Kishimoto, T. (1994) *Cell* 77, 63–71.
- [18] Zhong, X., Wen, Z. and Darnell Jr., J.E. (1994) *Science* 264, 95–98.
- [19] Zhong, Z., Wen, Z. and Darnell Jr., J.E. (1994) *Proc. Natl. Acad. Sci. USA* 91, 4806–4810.
- [20] Azam, M., Erdjument-Bromage, H., Kreider, B.L., Xia, M., Quelle, F., Basu, R., Saeis, C., Tempst, P., Ihle, J. and Schindler, C. (1995) *EMBO J.* 14, 1402–1411.
- [21] Knowles, B.B., Howe, C.C. and Aden, D.P. (1980) *Science* 209, 479–499.
- [22] Chomcnsi, P. and Sacchi, N. (1987) *Anal. Biochem.* 162, 156–159.
- [23] Laird, P.W., Zijderveld, A., Linders, K., Rudnicki, M.A., Jaenish, R. and Berns, A. (1991) *Nucleic Acids Res.* 19, 4293.
- [24] Shapiro, M.B. and Senapathy, P. (1987) *Nucleic Acids Res.* 15, 7155–7174.
- [25] Watanabe, K., Fukuchi, T., Hosoya, H., Shirasawa, T., Matuoka, K., Miki, H. and Takenawa, T. (1995) *J. Biol. Chem.* 270, 13733–13739.
- [26] Tanaka, M., Gupta, R. and Mayer, B.J. (1995) *Mol. Cell. Biol.* 15, 6829–6837.