

## A Dual Reporter Splicing Assay Using HaloTag-containing Proteins

Koichi Oshima<sup>a,b,g</sup>, Takahiro Nagase<sup>a</sup>, Kohsuke Imai<sup>c</sup>, Shigeaki Nonoyama<sup>c</sup>, Megumi Obara<sup>d</sup>, Tomoyuki Mizukami<sup>d</sup>, Hiroyuki Nunoi<sup>d</sup>, Hirokazu Kanegane<sup>e</sup>, Futoshi Kuribayashi<sup>f</sup>, Shin Amemiya<sup>g</sup>, Osamu Ohara<sup>\*a,b</sup>

<sup>a</sup>Department of Human Genome Research, Kazusa DNA Research Institute, Kisarazu, Japan

<sup>b</sup>Laboratory for Immunogenomics, Research Center for Allergy and Immunology, RIKEN, Yokohama Institute, Yokohama, Japan

<sup>c</sup>Department of Pediatrics, National Defense Medical College, Tokorozawa, Japan

<sup>d</sup>Department of Pediatrics, Miyazaki Medical College, Miyazaki, Japan

<sup>e</sup>Department of Pediatrics, Graduate School of Medicine, University of Toyama, Toyama, Japan

<sup>f</sup>Department of Biochemistry, Kawasaki Medical School, Kurashiki, Japan

<sup>g</sup>Department of Pediatrics, Saitama Medical University, Japan

**Abstract:** To evaluate the effects of genetic variations on mRNA splicing, we developed a minigene-based splicing assay using reporter genes encoding luciferase and the multifunctional HaloTag protein. In addition to conventional RT-PCR analysis, splicing events can be monitored in this system using two parameters: luciferase activity and signals derived from HaloTag-containing proteins bound to a fluorescent ligand following SDS-PAGE. The luciferase activity reflects the accumulated amounts of successfully spliced HaloTag-luciferase fusion products, whereas the amounts and sizes of HaloTag-containing proteins provide quantitative insights into precursor, correctly spliced, and aberrantly spliced mRNA species. Preliminary experiments confirmed that the dual reporter minigene assay can provide estimates of overall splicing efficiency based on the levels of protein products. We then used the minigene assay to analyze a case of chronic granulomatous disease that was caused by a G>C mutation at position +5 in the 5'-splice donor site of intron 5 of the *CYBB* gene. We found that the G>C mutation affected *CYBB* mRNA splicing by changing a delicate balance of splicing efficiencies of introns 4, 5, and 6.

**Keywords:** Chronic granulomatous disease, *CYBB* gene, HaloTag fusion protein, luciferase reporter assay, splicing, genetic mutation.

### INTRODUCTION

Determining whether mutations at splice sites cause a disease generally involves comparative reverse transcription-PCR (RT-PCR) analysis of the patient and the patient's family members using appropriate RNA samples. The necessary RNA samples, however, are not always available in practice. Thus, alternative approaches with surrogate minigene constructs have been widely used to test the effects of mutations on splicing [1-3]. Whereas conventional (or end-point) RT-PCR is the method of choice to monitor splicing events in most minigene-based approaches, this technique does not allow robust quantitative analysis and routine estimates of overall splicing efficiency. Although quantitative RT-PCR may overcome some of these limitations, designing a

quantitative RT-PCR system for precursor, correctly spliced, and aberrantly spliced products without a priori knowledge of the splicing events is extremely challenging.

Reporter-based minigene assays can also be used to quantitatively evaluate the effects of genetic variation on splicing because the reporter gene allows monitoring of splicing events based on the levels of protein products. Previously described reporter-based minigene systems, however, can detect only normally spliced products and not aberrant products generated because of cryptic splice-site activation, exon skipping, or intron retention, for example. To overcome this problem, we developed a dual reporter system that includes luciferase and the new tag protein HaloTag (molecular mass, approximately 33 kDa; [4]). Luciferase activity can be used as a quantitative measure of correctly spliced minigene products, whereas the HaloTag moiety, which can be easily detected just by adding a fluorescent HaloTag ligand to cell lysates prior to SDS-PAGE, allows us

\*Address correspondence to this author at the Department of Human Genome Research, Kazusa DNA Research Institute, Kisarazu, Chiba Japan; Tel: +81-438-52-3913; Fax: +81-483-52-3914; E-mail: [ohara@kazusa.or.jp](mailto:ohara@kazusa.or.jp)

to quantify unspliced, correctly spliced, and aberrantly spliced minigene products without protein blotting [4].

## MATERIALS AND METHODOLOGY

### Characterization of Abnormal Phenotypes in a Patient with Chronic Granulomatous Disease (CGD)

All experiments with human samples were reviewed and approved by the appropriate Institutional Review Boards. Production of reactive oxygen species in neutrophils obtained from a patient with CGD was analyzed using flow cytometry, the fluorescent probe dihydrorhodamine 123 (DHR-123), and phorbol myristate acetate as the stimulus. Neutrophils were also examined for surface expression of cytochrome  $b_{558}$  using 7D5 anti-cytochrome  $b_{558}$  monoclonal antibodies and phycoerythrin-conjugated secondary antibodies. The stained neutrophils were analyzed using FACScan and CellQuest software (Becton Dickinson, San Jose, CA). We also analyzed the *CYBB* gene of the patient after obtaining informed consent. Coding exons and flanking intronic sequences were amplified in PCRs using a Multiplex PCR Assay Kit (Takara Bio Inc., Shiga, Japan) (Supplementary Table 1). PCR products were purified using an ExoSAP-IT kit (GE Healthcare, Amersham, UK), and directly sequenced using an ABI 3130 DNA sequencer (Life Technologies, Carlsbad, CA). cDNA was created using neutrophils from the patient, his mother, and a healthy donor, and subjected to RT-PCR analysis of a region spanning exons 4 to 8 (upstream primer: CGAAGACAACCTGGACAGGAA-TC; downstream primer: GTGATGACCACCTTCTGTGATGAT). We directly sequenced amplified products excised from agarose gels.

### Construction of Exon/Intron Cassette-trapping Plasmids Carrying the HaloTag and *Renilla* Luciferase Genes

We first amplified the coding region of the *Renilla* luciferase gene from pRL-CMV (Promega, Madison, WI) using a high-fidelity polymerase (KOD plus ver. 2 PCR kit; Toyobo, Osaka, Japan). The PCR product was cloned into the *SgfI* and *PmeI* sites of pFN21A to create pFN21A-RL (Promega). The exon/intron cassette of interest was prepared using KOD plus ver. 2 PCR kit (Toyobo) with 20-bp flanking extensions that were homologous to the vector ends, based on instructions from an In-Fusion kit (Clontech Laboratories, Mountain View, CA). The cassette was designed to terminate translation in an intronic region when the cassette was inserted between the HaloTag and luciferase genes, whereas in-frame splicing at authentic splicing sites in the cassette generated HaloTag-luciferase fusion proteins. The exon/intron cassette was purified on an agarose gel and cloned into the *SgfI* site of pFN21A-RL using an In-Fusion protocol (Clontech Laboratories). Detailed experimental conditions are described in Supplementary Table 2 and the overall scheme is shown in Fig. (1).

Wild-type exon/intron cassettes were prepared in PCRs using commercially available human genomic DNA from a healthy donor (Clontech Laboratories) as a template. Unless stated otherwise, mutant exon/intron cassettes were prepared from patient genomic DNA (Supplementary Table 2). In some cases, specific nucleotide substitutions were introduced

using PCR-assisted overlap-extension mutagenesis [5]. The resulting reporter plasmids were subjected to DNA sequencing using a BigDye Terminator Ver. 3 Kit (Life Technologies) to confirm the sequences of the inserted fragments.

### Transfection

Unless stated otherwise, the splicing assay was conducted using human embryonic kidney 293T cells that were seeded in 100  $\mu$ l of growth medium (Dulbecco's modified Eagle medium-high glucose supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin-amphotericin B solution; Gibco Invitrogen, Carlsbad, CA, USA) at  $2 \times 10^4$  cells/well in 96-well plates. Twenty-four hours after plating, cells were transfected with 100 ng of reporter plasmid and 2 ng of pGL4.13 encoding the firefly luciferase reporter (Promega) using FuGENE HD (Roche Diagnostics, Indianapolis, IN, USA) according to the manufacturer's recommendations. After 16–24 h, HaloTag fusion proteins were labeled with 5  $\mu$ M HaloTag tetramethylrhodamine (TMR) ligand (Promega) and incubated for an additional 20–24 h in a 5% CO<sub>2</sub> incubator at 37°C. The percentage of TMR-positive cells was determined using a Biozero fluorescence microscope (Keyence, Osaka, Japan) with a TRITC filter setting (OP-66837, Keyence).

### Luciferase Reporter Assay

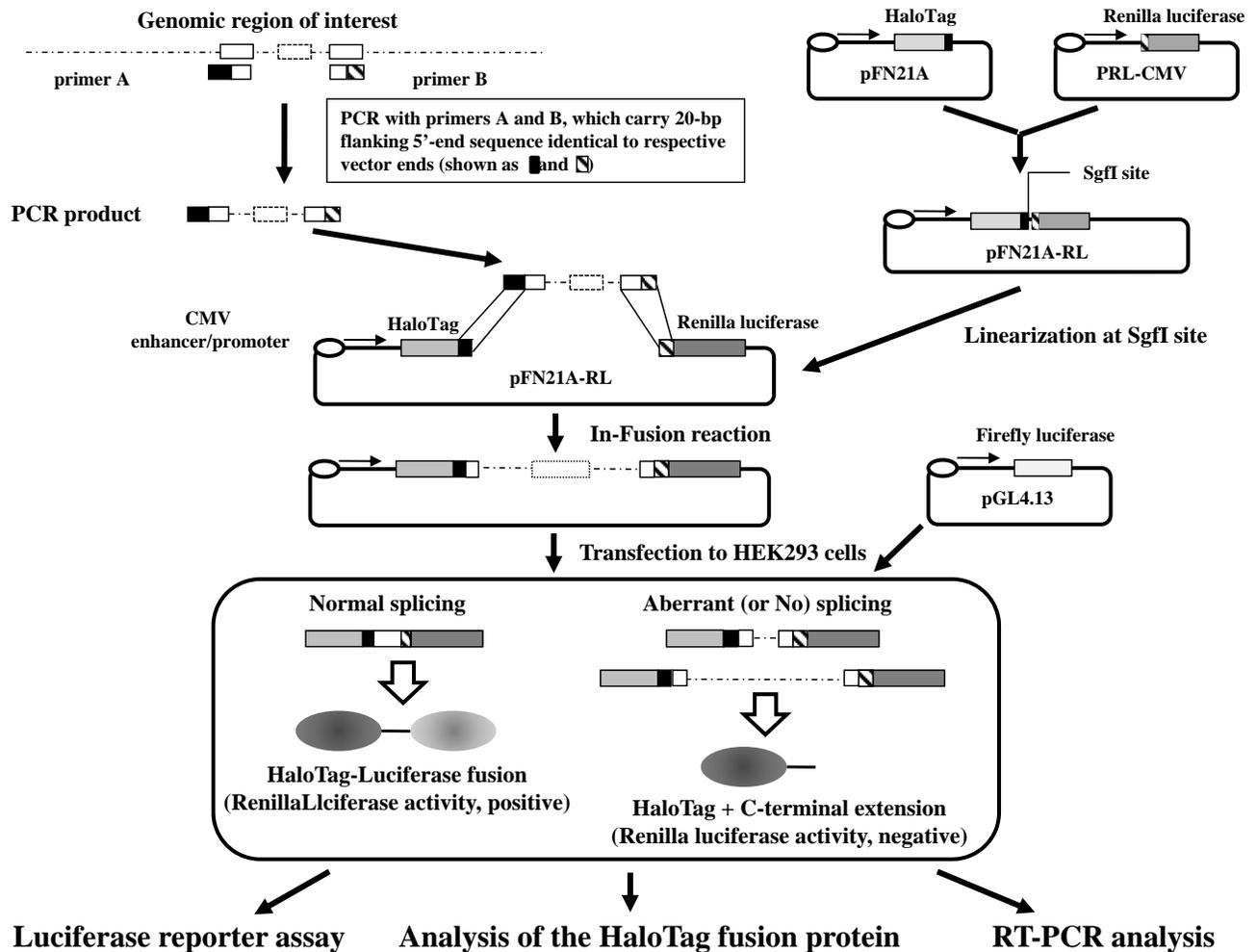
Dual luciferase assays were performed using the Dual-Luciferase Reporter Assay System (Promega) 44 h following transfection. After mixing 10  $\mu$ l of cell extracts with 100  $\mu$ l of Luciferase Assay Reagent II in a 96-well plate, firefly luciferase activity was measured using a GloMax 96 Microplate Luminometer (Promega). Following the first measurement, *Renilla* luciferase activities were measured with the same luminometer after 100  $\mu$ l of Stop and Glo reagent was added (Promega). All samples were assayed in triplicate. Average *Renilla* luciferase activities for the wild-type and mutant exon/intron cassettes in the same experimental trial were compared after normalizing the samples based on firefly luciferase activities to compensate for variations in transfection efficiency.

### Analysis of HaloTag Fusion Proteins

Extracts from TMR ligand-treated cells prepared as described for the luciferase reporter assay were separated on a 12% polyacrylamide gel. TMR fluorescent signals were detected and quantified using a FLA-3000 system (Fujifilm, Tokyo, Japan) and MultiGauge image analysis software (Fujifilm).

### RT-PCR Analysis

cDNA was synthesized using the ImProm II Reverse Transcription System (Promega) and total RNA (approximately 200 ng) that was prepared with an RNeasy Plus Micro Kit (Qiagen, Valencia, CA) 48 h after cells were transfected. The resulting cDNA samples served as templates in 20- $\mu$ l PCR reactions using a Multiplex PCR Assay Kit (Takara Bio, Kyoto, Japan) according to the manufacturer's instructions. The upstream and downstream PCR primers and RT-PCR conditions are described in Supplementary Table 3.



**Fig. (1). Experimental overview**

The inserted DNA fragments were designed to contain introns with or without mutations and the adjacent exons. The primers used to amplify the fragments in PCRs included 20-bp sequences that overlapped with the terminal sequences of the digested plasmid DNA. An In-Fusion reaction was used to ligate each DNA fragment to an exon/intron cassette-trapping plasmid containing the *Renilla* luciferase gene and HaloTag gene (pFN21A-RL). After the construct was transfected into mammalian cells, successful splicing produced HaloTag–luciferase fusion proteins and high levels of luciferase activity, whereas aberrant splicing or no splicing did not.

PCR products were separated and quantified using a MCE-202 MultiNA Microchip Electrophoresis System (Shimadzu Biotech, Kyoto, Japan).

### Statistical Analysis

Data are expressed as means  $\pm$  standard errors. Unless stated otherwise, statistical significance was determined using Student's *t*-tests and defined as *P* values  $<$  0.05.

## RESULTS

### 1. Development of a Dual Reporter–based Splicing Assay to Evaluate the Effects of Splice-site Mutations

#### a. A Dual Reporter–based Splicing Assay

Fig. (1) illustrates the overall scheme that was used to construct a dual reporter minigene plasmid. Each exon/intron cassette was amplified from genomic DNA, and inserted into

a dual reporter minigene vector. The resultant plasmids were introduced into cells together with a plasmid encoding firefly luciferase to normalize transfection efficiencies in each experiment. HaloTag proteins were monitored using a membrane-permeable fluorescent HaloTag ligand (TMR-HaloTag ligand). Because binding between the TMR-HaloTag ligand and the HaloTag moiety is highly stoichiometric (1:1 ratio), the fluorescence intensities of protein bands after SDS-PAGE likely reflected relative levels of HaloTag proteins after cell lysis. Thus, luciferase activity was proportional to the amount of properly spliced HaloTag–luciferase fusion proteins, whereas unspliced and aberrantly spliced products were quantified based on fluorescent signals from protein bands after SDS-PAGE and labeling of HaloTag-containing proteins. Conventional RT-PCR analysis can also be performed with this system; however, although RT-PCR analysis is useful for characterizing spliced transcripts, quantification of transcripts using conventional RT-PCR assays is limited because of differences in PCR amplification efficiencies

for various mRNA species. In particular, because unspliced mRNA is frequently large enough to considerably lower amplification efficiencies, end-point RT-PCR assays are often used to estimate the relative amounts of alternatively spliced mRNA. Although quantitative RT-PCR assays may provide a solution to this problem, this technique is somewhat laborious owing to the need to prepare control DNA to quantify all spliced and unspliced products.

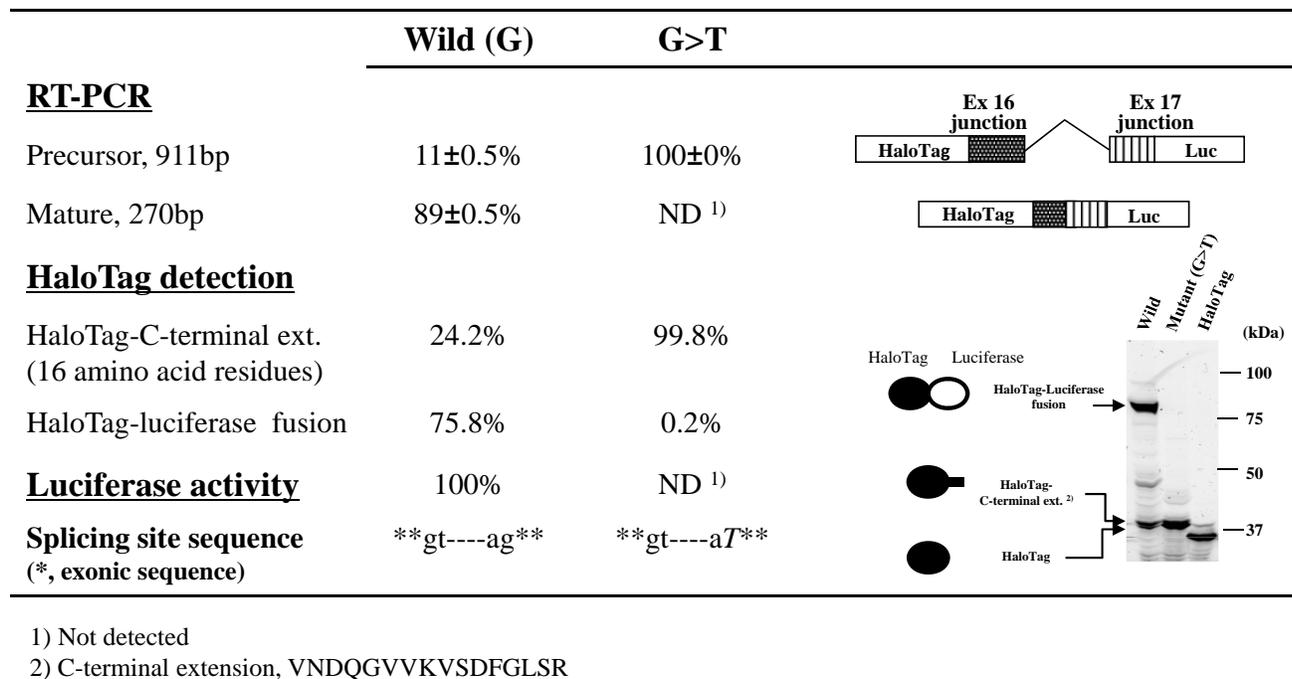
**b. Evaluation of a G>T Mutation at Position -1 of the 3'-splicing Acceptor Site of Intron 16 in the Gene Encoding Bruton's Tyrosine Kinase (BTK)**

As a proof-of-concept experiment, we first examined the effects of a G>T mutation at position -1 of the 3'-splicing acceptor site of intron 16 of the *BTK* gene, which was identified as a causative mutation for X-linked agammaglobulinemia based on a genetic test performed at Kazusa DNA Research Institute. This mutation disrupts the canonical GT-AG rule of splicing boundaries and, as expected, resulted in complete loss of mature RT-PCR products owing to a lack of splicing of intron 16 (Fig. 2). Consistent with this result, we observed almost complete suppression of luciferase activity and very low levels of the HaloTag-luciferase fusion protein with the mutant minigene construct. The lower fluorescent band shown in Fig. (2) was likely a product of minigene precursor mRNA, because the molecular mass estimated based on its electrophoretic mobility was consistent with the deduced molecular mass of the translation product of the precursor mRNA (*i.e.*, a HaloTag protein with a carboxy-terminal extension of 16 amino-acid residues). The splicing efficiency of this *BTK* exon/intron cassette in the minigene

construct was estimated from the relative fluorescent HaloTag band intensities with and without luciferase. Although end-point RT-PCR can also be used to estimate splicing efficiency, this method can produce biased results due to differences in amplification efficiencies. Because measuring fluorescence intensities of HaloTag bands is simple and direct, the dual reporter assay appears to provide more consistent analysis of overall splicing efficiency compared with end-point RT-PCR assays.

**c. Evaluation of Mutations at Position +5 of the 5'-splicing Donor Site of Intron 6 in the Gene Encoding Wiskott-Aldrich Syndrome Protein (WAS)**

We next examined the effects of nucleotide changes (G>A, G>T, or G>C) at position +5 of the 5'-splicing donor site of intron 6 in the *WAS* gene. Although only the G>A change has been described in case reports [6-14], we examined the effects of all possible nucleotide changes at this position. Fluorescence images of polyacrylamide gels indicated that all mutant minigene constructs generated a band in addition to those derived from mature and precursor mRNA (Fig. 3). Judging from the size of the additional band, it did not include a luciferase moiety and was likely an aberrantly spliced product. Sequencing the RT-PCR products revealed that all of the mutations activated a cryptic splicing donor site in the dual reporter splicing assay, as has been observed *in vivo* [6-14]. The nucleotide sequence of the RT-PCR products suggested that the protein in the additional band included a carboxy-terminal extension of 26 amino-acid residues. This result indicated that SDS-PAGE analysis can be used to identify aberrant splicing without DNA sequencing.



**Fig. (2). Effects of mutations at splice sites in the *BTK* gene on splicing events**

Splicing events in *BTK* transcripts were examined using RT-PCR products, luciferase activity, and HaloTag fusion proteins. We examined the effects of a G>T change at position -1 of the 3'-splicing acceptor site in intron 16 of the *BTK* gene. The experiment was carried out in triplicate and RT-PCR products were quantified as weighted ratios obtained from a MCE-202 MultiNA Microchip Electrophoresis System. Schematic representations of the RT-PCR products are shown on the right in the figure.

	Wild (G)	G>A	G>T	G>C	
<b>RT-PCR</b>					
Precursor, 859bp	40±1.0%	70±0.6%	70±1.1%	63±0.6%	
Aberrant <sup>2)</sup> , 251bp	ND <sup>1)</sup>	26±0.3%	24±1.0%	35±0.6%	
Mature, 213bp	60±1.0%	4.0±0.2%	5.0±0.1%	3.0±0.1%	
<b>HaloTag detection</b>					
HaloTag-C-terminal ext. (15 amino acid residues)	39.6%	62.2%	73.2%	54.0%	
HaloTag-C-terminal ext. (26 amino acid residues)	1.6%	36.5%	23.8%	44.6%	
HaloTag-luciferase fusion	58.8%	1.3%	3.0%	1.3%	
<b>Luciferase activity</b>	100%	16±2.2%	26±4.0%	12±0.6%	
<b>Splicing site sequence</b> (*, exonic sequence)	**gtgcg---	**gtgcA---	**gtgcT---	**gtgcC---	

1) Not detected

2) Aberrant activation of a cryptic splicing site in intron 6 as reported

3) C-terminal extension caused by the cryptic splicing site activation, KEVR(H/L/P)ADSSLCLWMDGPSSGSALPRSP

4) C-terminal extension, KEVR(H/L/P)ADSSLCLWMDG

**Fig. (3). Effects of mutations at splice sites in the WAS gene on splicing events**

Splicing events in *WAS* transcripts were assessed using RT-PCR products, luciferase activity, and HaloTag fusion proteins. We used exon/intron cassette-trapping plasmids with or without mutations (G>A, G>T, G>C, or wild-type) at position +5 of the 5'-splicing donor site in intron 6 of the *WAS* gene. The experiment was carried out in triplicate and RT-PCR products were quantified as weighted ratios obtained from a MCE-202 MultiNA Microchip Electrophoresis System. Schematic representations of the RT-PCR products are shown on the right in the figure.

Furthermore, our dual reporter assay revealed that, the cryptic site was used even with the wild-type minigene construct, whereas this splicing event was not detected in a conventional RT-PCR assay. Moreover, the protein products revealed that mutation of this site increased not only the amounts of aberrantly spliced mRNA at the cryptic splicing site but also those of precursor mRNA. Finally, although no marked differences in the effects of the G>A, G>T, and G>C mutations were detected using a conventional RT-PCR assay, our dual reporter assay showed that the G>C mutation produced the greatest effects on normal splicing, whereas the G>T mutation produced the smallest effects.

## 2. Using Dual Reporter Splicing Assays to Characterize an Intronic Mutation in the *CYBB* Gene of a Patient with CGD

### a. Identification of a Mutation in a Patient with CGD

The patient was a 2-year-old boy with no family history of CGD who had been hospitalized several times because of *Bacillus Calmette-Guérin* vaccine-induced regional axillary lymphadenitis, liver abscesses, and bacterial lymphadenitis. Flow cytometry with DHR-123 showed reduced reactive oxygen species production in activated neutrophils from the patient (Fig. 4A). On the other hand, the patient's neutrophils showed markedly decreased expression of cytochrome *b<sub>558</sub>* and the mother's neutrophils exhibited a mosaic pattern (Fig.

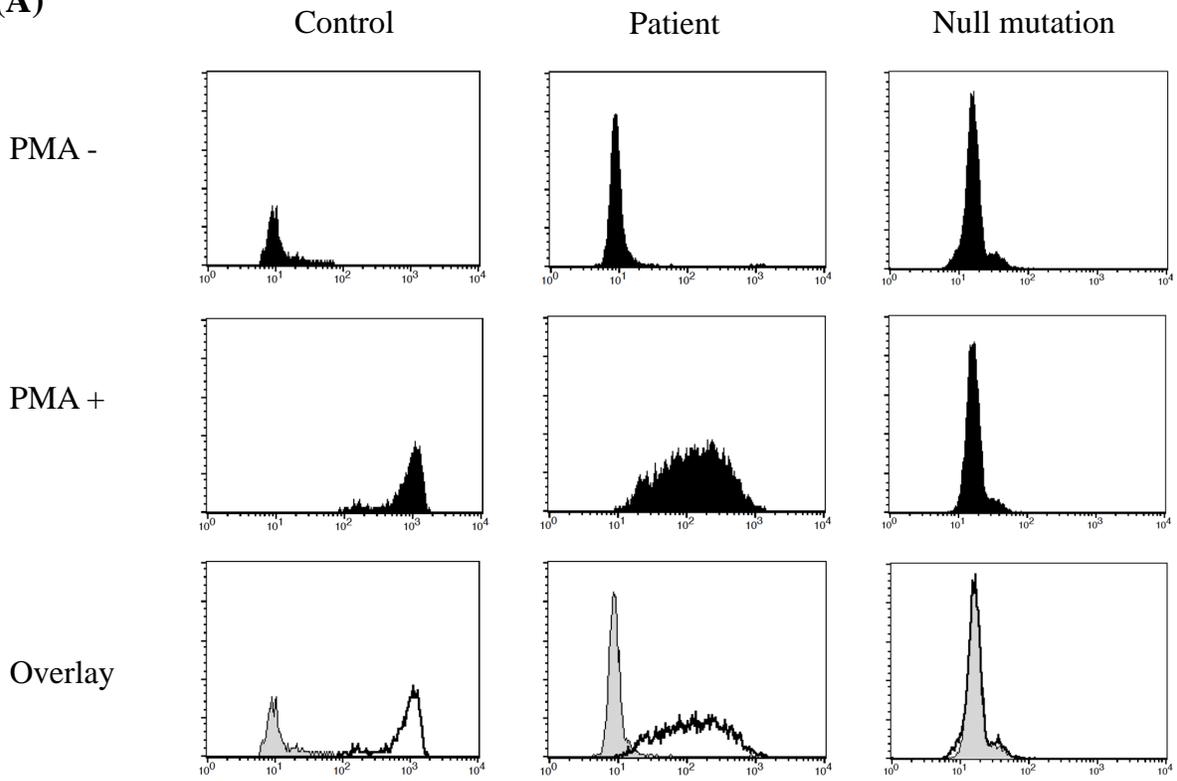
4B); because of random X chromosome inactivation, female carriers of genes that cause X-linked CGD often exhibit this type of cytometric pattern. After obtaining informed consent, protein coding exons and short flanking intronic regions of the *CYBB* gene were directly sequenced. We identified a single G>C nucleotide substitution at position +5 in the 5'-splice donor site of intron 5 (Fig. 4C), a site at which a single nucleotide G>A substitution was previously identified [15].

We then characterized the effects of this newly identified mutation in the *CYBB* gene using clinical samples. We performed RT-PCRs to amplify exon 4 to exon 8 in cDNA created from the neutrophils of the patient, his mother, and a healthy donor (Fig. 5). The RT-PCR products were separated on an agarose gel and directly sequenced, revealing exon skipping of *CYBB* exon 5 or exons 5 and 6 in the patient, whereas only heterozygous skipping of exon 5 was observed in his mother. These results indicated that the mutation caused CGD due to exon skipping. We then evaluated the effects of the mutation on *CYBB* splicing events quantitatively using the new dual reporter minigene assay.

### b. Evaluating Mutations at Position +5 of the 5'-splicing Donor Site of Intron 5 in the *CYBB* Gene Using Dual Reporter Minigene Assays

We used our splicing assay to examine splicing of *CYBB* intron 5 flanked by short regions of exons 5 and 6 with and without a nucleotide change at position +5 of the 5'-splicing

**(A)**



**(B)**

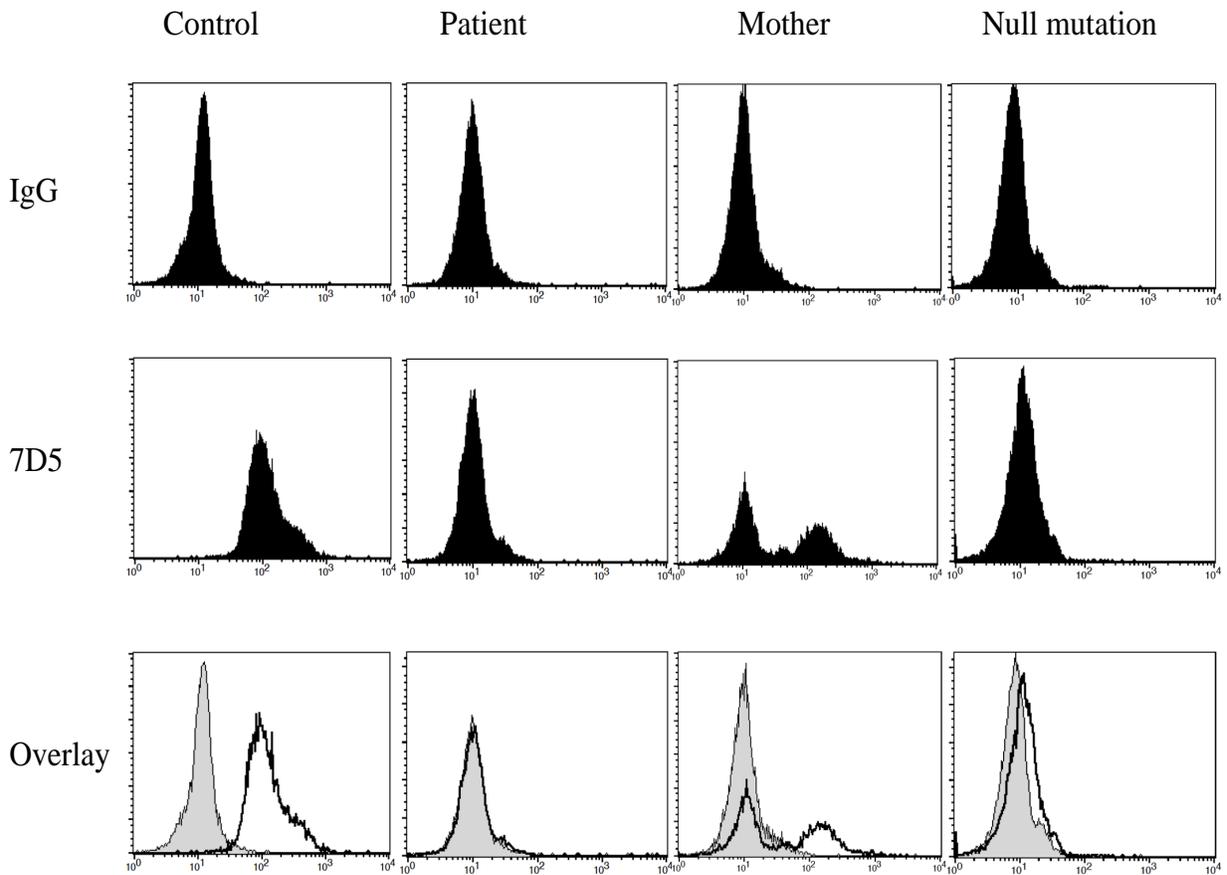
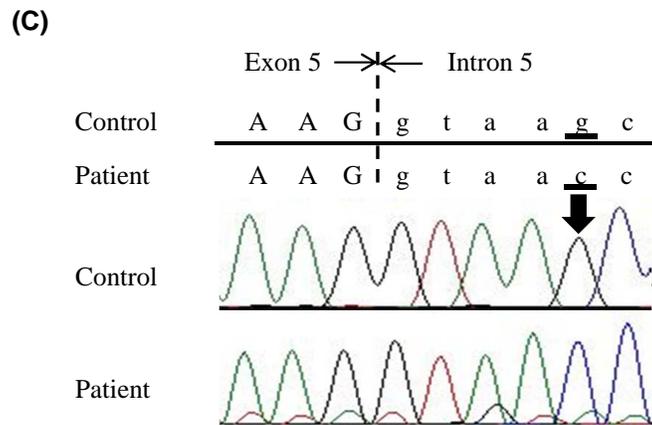


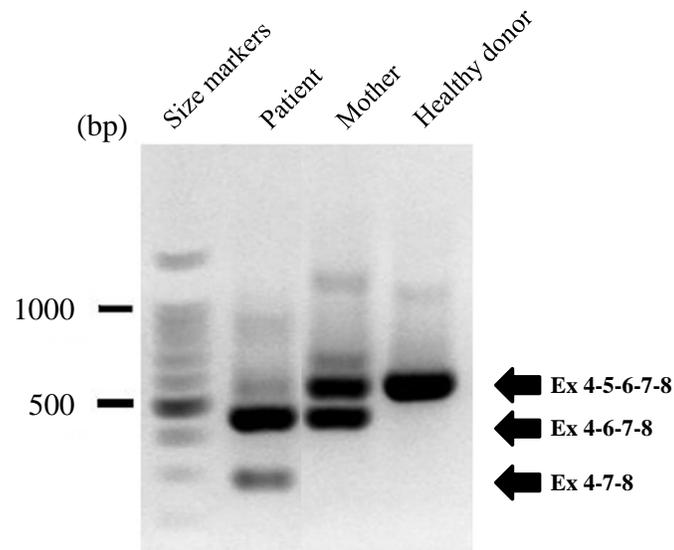
Fig. (4). Contd....

**Fig. (4). Patient phenotypes and direct sequencing of the *CYBB* gene**

**A.** Flow cytometry with DHR-123 demonstrated that reactive oxygen species were produced at low levels in neutrophils obtained from the patient. The effects of a null mutation (del425\_426CT, 142S>fsX) are shown as a negative control.

**B.** Flow cytometry with anti-cytochrome  $b_{558}$  monoclonal antibodies showed markedly decreased surface expression of cytochrome  $b_{558}$  on the patient's neutrophils.

**C.** Direct sequencing of the *CYBB* gene indicated that the patient had a homozygous single G>C nucleotide substitution at position +5 in the consensus 5'-donor splice site of intron 5.

**Fig. (5). RNA analysis of neutrophils from a patient with CGD, his mother, and a control subject**

Single-stranded cDNA was synthesized from neutrophil RNA obtained from the patient, his mother, and a control subject. cDNA was PCR-amplified using a sense primer specific for exon 4 and an antisense primer specific for exon 8 of the *CYBB* gene. The products were separated on a 3% agarose gel and stained with ethidium bromide. DNA size markers are shown on the left and the numbers on the right indicate differently spliced products. Although some normally splice transcripts were identified, most PCR products amplified from the patient lacked exon 5 or exons 5 and 6. Normally spliced transcripts and transcripts lacking exon 5 were identified in the mother.

donor site (exon/intron cassettes containing a single intron). Although these experimental conditions were similar to those used in the preliminary experiment with the *WAS* gene (Fig. 3), RT-PCR and luciferase assays indicated that changing the nucleotide from G to another residue did not lower the splicing efficiency. Interestingly, the absolute luciferase activity observed with the minigene construct containing the G>C change was reproducibly highest among the minigene constructs (Fig. 6). In addition, the splicing efficiency observed with this cassette appeared high because the amount of Hal-

oTag fusion protein without luciferase (< 5% of total protein; Fig. 6), which was likely transcribed from precursor mRNA, were lower than results observed with cassettes containing the *BTK* or *WAS* gene (24.2% or 39.6% of total protein, respectively). HaloTag fusion proteins without luciferase appeared as a doublet on polyacrylamide gels even though we expected that a single product of approximately that size was encoded by the minigene construct. The mutant cassettes resulted in slight but detectable increases in the amount of luciferase-deficient HaloTag products (Fig. 6), suggesting a

	Wild (G)	G>C	G>T	G>A	
<b>RT-PCR</b>					
Precursor, 2383bp	ND <sup>1)</sup>	ND <sup>1)</sup>	ND <sup>1)</sup>	ND <sup>1)</sup>	
Mature, 243bp	100±0%	100±0%	100±0%	100±0%	
<b>HaloTag detection</b>					
HaloTag-C-terminal ext. (23 amino acid residues)	4.9%	16.7%	14.1%	17.1%	
HaloTag-luciferase fusion	95.1%	83.3%	85.9%	82.9%	
<b>Luciferase activity</b>					
	100%	155±29.1%	119±10.5%	101±7.9%	
<b>Splicing site sequence (* , exonic sequence)</b>					
	**gtaag---	**gtaaC---	**gtaaT---	**gtaaA---	

1) Not detected

2) C-terminal extension, ESVLNFARKAIKVS(T/I/N)LSLSDLIL

**Fig. (6). Effects of splice-site mutations in *CYBB* exon/intron cassettes containing a single intron**

Splicing events identified based on RT-PCR products, luciferase activity, and HaloTag fusion proteins are shown. We used exon/intron cassette-trapping plasmids with or without mutations (G>C, G>T, G>A, or wild-type) at position +5 of the 5'-splicing donor site in intron 5 of the *CYBB* gene. The experiment was carried out in triplicate and levels of RT-PCR products were quantified as the weighted ratios obtained using a MCE-202 MultiNA Microchip Electrophoresis System. Schematic representations of the RT-PCR products are shown on the right in the figure.

small decrease in splicing efficiency with the exon/intron cassette carrying a mutation at the +5 position of *CYBB* intron 5. Thus, the mutation found in the patient with CGD did not markedly affect the splicing of intron 5.

### c. Examining the Effects of Mutations at Position +5 of the 5'-splicing Donor Site of *CYBB* Intron 5 Using Dual Reporter Minigene Constructs with Multiple Introns

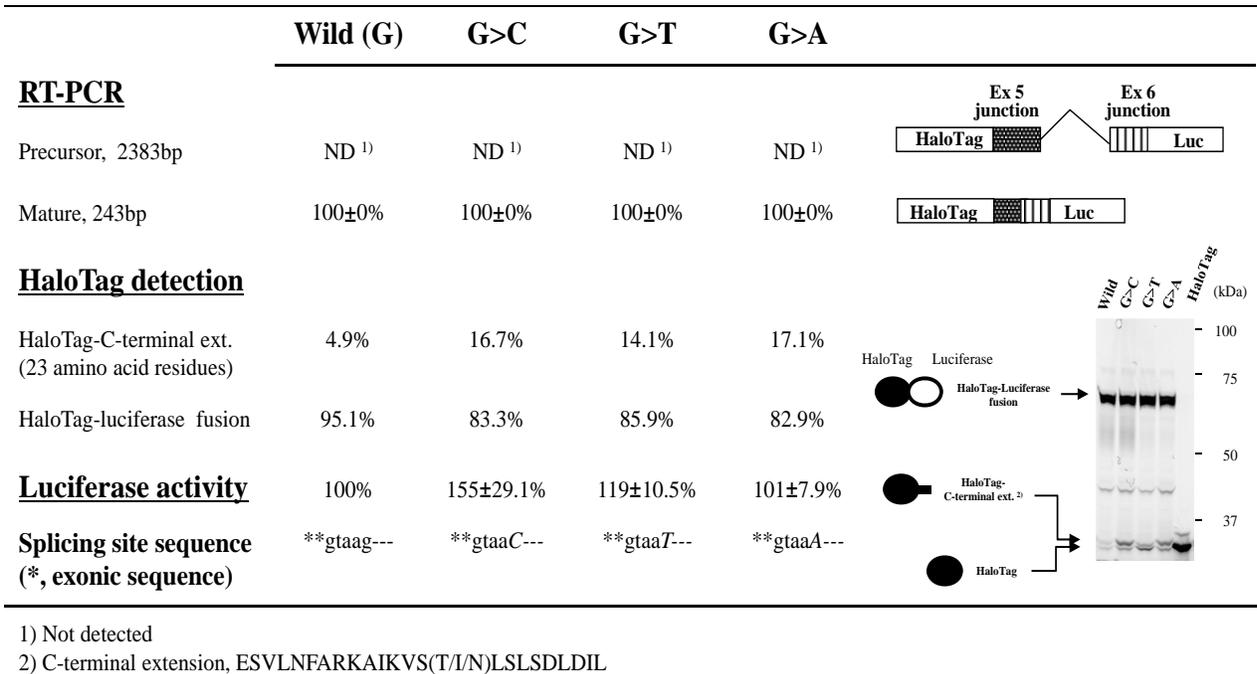
Evaluating transcripts from the patient's neutrophils suggested that the G>C substitution resulted in skipping of exon 5 (major) or exons 5 and 6 (minor; Fig. 5). To reproduce exon skipping using the minigene construct, we next examined two additional exon/intron cassettes: the first cassette contained intron 4–exon 5–intron 5 flanked by short regions of exons 4 and 6 (3950 bp), whereas the second cassette carried intron 4–exon 5–intron 5–exon 6–intron 6 flanked by short regions of exons 4 and 7 (6940 bp). We examined four cassettes containing introns 4 and 5, each with a different nucleotide (G, C, T, or A) at position +5 of the 5'-splicing donor site of *CYBB* intron 5. On the other hand, only the G>C mutation was examined in the exon/intron cassettes containing three introns.

For mutant exon/intron cassettes containing two or three introns, most RT-PCR products showed the expected exon skipping, whereas both wild-type cassettes also generated a considerable amount of RT-PCR products that were shortened by exon skipping (Figs. 7 and 8). The results were not consistent with the RT-PCR analyses of blood samples from the patient and family members, which did not show exon skipping for the wild-type *CYBB* gene, whereas the mutant *CYBB* allele generated aberrant mRNA without exon 5 (ma-

nor) or exons 5 and 6 (minor). Analyzing luciferase activities obtained with the exon/intron cassettes containing multiple introns demonstrated that amounts of mature mRNA encoding HaloTag–luciferase fusions decreased by an order of magnitude when one additional intron was included. Thus, SDS-PAGE analysis of HaloTag fusion proteins was of little use in this case because the main protein products produced from the exon/intron cassettes were encoded by partially spliced and/or precursor mRNA. On the other hand, RT-PCR assays allowed us to analyze only spliced mRNA because the introns were too large to be effectively amplified, resulting in a rough estimate of relative amounts of spliced variants derived from each exon/intron cassette. Therefore, only the luciferase assay measured the overall splicing efficiency of each exon/intron cassette (Figs. 7 and 8). The results indicated that the G>C mutation at position +5 of the 5'-splicing donor site of intron 5 markedly reduced the amounts of successfully spliced products relative to those translated after exon skipping, whereas notable amounts of luciferase activities were detected with the mutant exon/intron cassettes (Figs. 7 and 8). It should be also noted that skipping exon 5 was rare in the minigene construct shown in Fig. (8), whereas it was more common than skipping exons 5 and 6 in the patient sample (Fig. 5).

## DISCUSSION

Minigene-based splicing assays should quantitatively examine unspliced, aberrantly spliced, and correctly spliced products to determine relative splicing efficiencies with and without mutations. We have developed a dual reporter assay that uses luciferase and HaloTag proteins. Because splicing



**Fig. (7). Effects of splice-site mutations in *CYBB* exon/intron cassettes containing two introns**

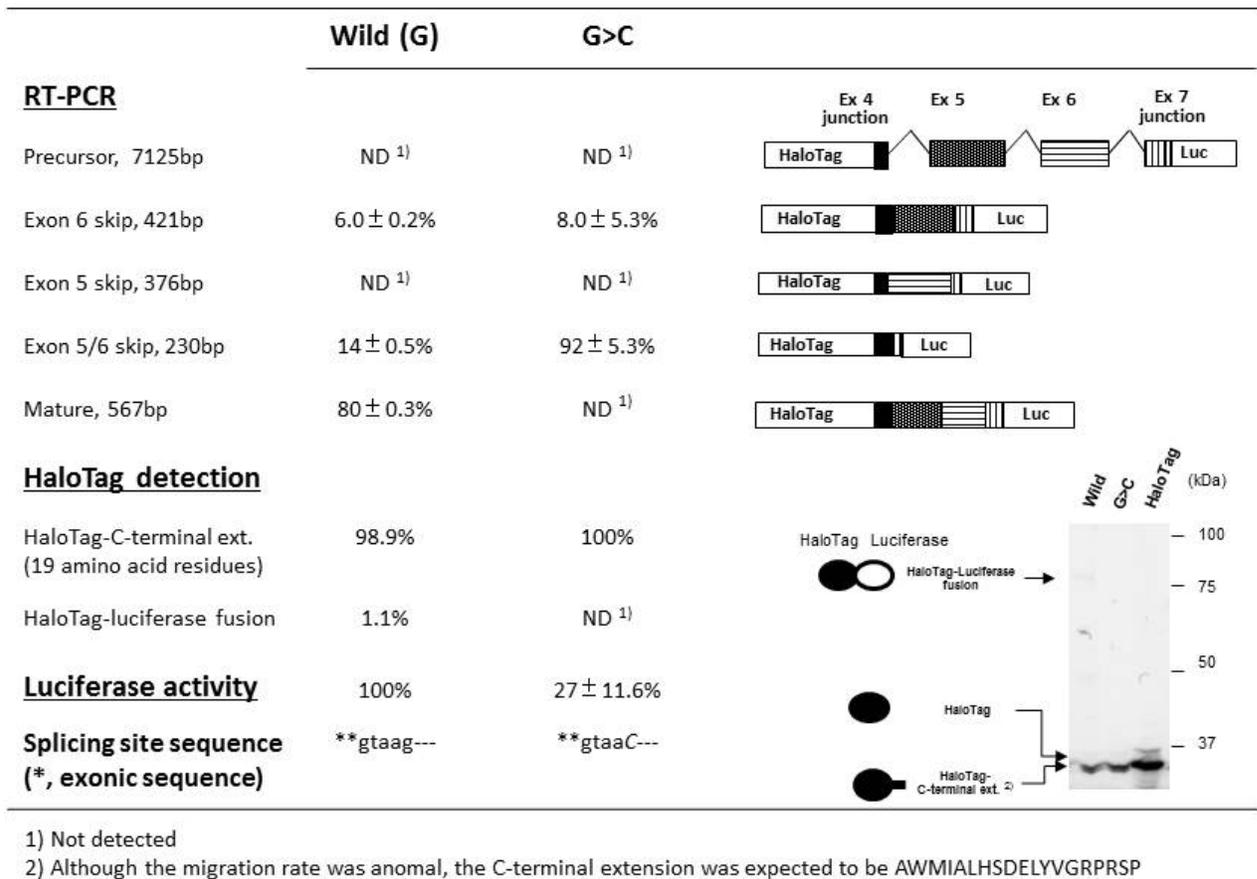
Splicing events identified based on RT-PCR products, luciferase activity, and HaloTag fusion proteins are shown. We used exon/intron cassette-trapping plasmids with or without mutations (G>C, G>T, G>A, or wild-type) at position +5 of the 5'-splicing donor site in intron 5 of the *CYBB* gene. The experiment was carried out in triplicate and levels of RT-PCR products were quantified as the weighted ratios obtained using a MCE-202 MultiNA Microchip Electrophoresis System. Schematic representations of the RT-PCR products are shown on the right in the figure.

determines the sequence and length of any carboxy-terminal extension on the HaloTag protein, separating HaloTag fusion proteins encoded by unspliced, aberrantly spliced, and correctly spliced mRNA *via* SDS-PAGE allowed us to estimate the amounts of various protein products by measuring the fluorescent intensities of the protein bands. Although the relationships between electrophoretic mobilities of proteins during SDS-PAGE and the proteins' molecular masses are complex, the degree of protein migration usually allowed us to determine which band contained each of the possible minigene products. In particular, protein products translated from correctly spliced mRNA were easily identified based on the 36-kDa *Renilla* luciferase moiety at the carboxy terminus, which could be independently evaluated using luciferase activity.

The advantage of the HaloTag reporter was demonstrated using two examples (Figs. 2 and 3). Splicing events were best detailed by integrating data from RT-PCRs, luciferase assays, and fluorescence gel imaging of HaloTag proteins. In each of the preliminary experiments, the data indicated that a single base substitution resulted in a substantial reduction in successful splicing. Of note, splicing efficiencies (*i.e.*, the percentages of spliced products) estimated using RT-PCRs and HaloTag-based SDS-PAGE assays were similar, whereas the splicing efficiency observed with the *BTK* exon/intron cassette was considerably different from that detected with the *WAS* exon/intron cassette; although mutation of the canonical GT-AG splicing junction in the *BTK* gene almost completely disrupted splicing (Fig. 2), mutation at position

+5 of the 5'-splicing donor site of the *WAS* gene activated a cryptic splicing site in the intron and lowered the splicing efficiency from 60% to 30–40% based on RT-PCR and SDS-PAGE analyses (Fig. 3). For larger introns, however, precursor mRNA may not be amplified in RT-PCRs, as in the case of the *CYBB* gene (Figs. 6–8). Instead, the HaloTag moiety and a fluorescently labeled ligand can be used to monitor overall splicing efficiencies without biases caused by differences in the lengths of intronic sequences. This is the most prominent advantage of the HaloTag moiety as a reporter in the minigene-based splicing assay.

In the last experiment, we characterized the effects of a newly identified mutation in the *CYBB* gene of a patient with CGD using our minigene-based splicing assay. When we examined the effects of this mutation using an exon/intron cassette containing only intron 5, no inhibition of splicing was observed. The data shown in Fig. (6) indicated that splicing efficiency observed with this exon/intron cassette was highest among the examined cassettes (approximately 95%) and the mutation at position +5 of the splicing donor site slightly lowered the splicing efficiency. On the contrary, luciferase activity levels showed slight increases in the amounts of successfully spliced HaloTag–luciferase fusion proteins (Fig. 6), suggesting that overall expression of HaloTag–luciferase fusion proteins was enhanced by the mutation. The results indicated that the mutation did not markedly affect splicing efficiency of intron 5. We, however, observed exon skipping in *CYBB* transcripts from the patient. Thus, using the minigene-based assay and cassettes with a single



**Fig. (8). Effects of splice-site mutations in *CYBB* exon/intron cassettes containing three introns**

Splicing events identified based on RT-PCR products, luciferase activity, and HaloTag fusion proteins are shown. We used exon/intron cassette-trapping plasmids with or without a mutation (G>C or wild-type) at position +5 of the 5'-splicing donor site in intron 5 of the *CYBB* gene. The experiment was carried out in triplicate and levels of RT-PCR products were quantified as the weighted ratios obtained from a MCE-202 MultiNA Microchip Electrophoresis System. Schematic representations of the RT-PCR products are shown on the right in the figure.

intron to monitor splicing events based on the “intron definition model” [16] was inappropriate for this case. Instead, we examined the effects of the G>C mutation on splicing using exon/intron cassettes carrying two and three introns (Figs. 7 and 8, respectively). Significant exon skipping was observed even with the wild-type cassette, suggesting that inclusion of exon 5 in transcripts is intrinsically opportunistic; the results shown in Figs (7 and 8) indicated that inclusion of exon 5 seemed to be governed by a delicate balance of splicing efficiencies of flanking introns. This splicing profile cannot be explained based solely on the local genomic context of the splicing site. For example, although mutations at position +5 of the 5'-splicing donor site have frequently been reported to produce a range of effects in various genes [17-19], the identified mutation in this case was not the clear cause of the splicing abnormalities based on an evaluation with genomic analysis software (MutationTaster; <http://www.mutation-taster.org/>), which makes predictions based on local sequences around the splicing sites.

Other data obtained using the minigene assay were also somewhat discrepant from the patient data; in the minigene system, skipping exons 5 and 6 was more prominent than

skipping exon 5 or exon 6 (Fig. 8). These results likely reflected, at least in part, lower rates of splicing for each intron when the transcripts were transiently expressed in the minigene system compared with *in vivo* processes. Thus, the levels of spliced products in the minigene assay depend not only on splicing efficiencies of various introns but also on the number of splicing events required to generate the products; mRNA that is derived from the cassettes and lacks exon 5 or 6 is generated through two splicing steps, whereas a single splicing event can produce mRNA without exons 5 and 6 (Fig. 8). In addition, nonsense-mediated mRNA decay can affect steady-state levels of aberrantly spliced products *in vivo*. Because total HaloTag-specific signals were almost constant regardless of changes in splicing patterns, nonsense-mediated mRNA decay likely does not play a prominent role in our minigene assay. The results obtained using the minigene assay, however, were qualitatively consistent with those observed *in vivo*. In fact, the Human-transcriptome Database for Alternative Splicing ([www.h-invitational.jp/h-dbas/](http://www.h-invitational.jp/h-dbas/)) shows that exons 2, 5, and 6 of the *CYBB* gene are sometimes skipped, even in wild-type transcripts.

**CONFLICT OF INTEREST DISCLOSURE**

The authors have no financial conflicts of interest regarding the material discussed in the manuscript.

**ACKNOWLEDGEMENTS**

This work was supported by a Grant-in-Aid for Scientific Research; the program for Fostering Regional Innovation from the Ministry of Education, Culture, Sports, Science, and Technology of Japan; and a grant for research on intractable diseases from the Ministry of Health, Labour, and Welfare of Japan. We also thank the Kazusa DNA Research Institute (Chiba, Japan) and the RIKEN Research Center for Allergy and Immunology (Yokohama, Japan) for continued support.

**SUPPLEMENTARY MATERIAL**

Supplementary material is available on the publisher's website along with the published article.

**REFERENCES**

- [1] Cooper TA. Use of minigene systems to dissect alternative splicing elements. *Methods* 2005; 37: 331-40.
- [2] Singh G, Cooper TA. Minigene reporter for identification and analysis of cis elements and trans factors affecting pre-mRNA splicing. *Biotechniques* 2006; 41: 177-81.
- [3] Andreassi C, Jarecki J, Zhou J, *et al.* Aclarubicin treatment restores SMN levels to cells derived from type I spinal muscular atrophy patients. *Hum Mol Genet* 2001; 10: 2841-9.
- [4] Los GV, Wood K. The HaloTag: a novel technology for cell imaging and protein analysis. *Methods Mol Biol* 2007; 356: 195-208.
- [5] Ho SN, Hunt HD, Horton RM, *et al.* Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* 1989; 77: 51-9.
- [6] Falet H, Hoffmeister KM, Neujahr R, Hartwig JH. Normal Arp2/3 complex activation in platelets lacking WASp. *Blood* 2002; 100: 2113-22.
- [7] Imai K, Morio T, Zhu Y, *et al.* Clinical course of patients with WASP gene mutations. *Blood* 2004; 103: 456-64.
- [8] Inoue H, Kurosawa H, Nonoyama S, *et al.* X-linked thrombocytopenia in a girl. *Br J Haematol* 2002; 118: 1163-5.
- [9] Jin Y, Mazza C, Christie JR, *et al.* Mutations of the Wiskott-Aldrich Syndrome Protein (WASP): hotspots, effect on transcription, and translation and phenotype/genotype correlation. *Blood* 2004; 104: 4010-9.
- [10] Kwan SP, Hagemann TL, Radtke BE, *et al.* Identification of mutations in the Wiskott-Aldrich syndrome gene and characterization of a polymorphic dinucleotide repeat at DXS6940, adjacent to the disease gene. *Proc Natl Acad Sci USA* 1995; 92: 4706-10.
- [11] Lemahieu V, Gastier JM, Francke U. Novel mutations in the Wiskott-Aldrich syndrome protein gene and their effects on transcriptional, translational, and clinical phenotypes. *Hum Mutat* 1999; 14: 54-66.
- [12] Proust A, Guillet B, Picard C, *et al.* Detection of 28 novel mutations in the Wiskott-Aldrich syndrome and X-linked thrombocytopenia based on multiplex PCR. *Blood Cells Mol Dis* 2007; 39: 102-6.
- [13] Remold-O'Donnell E, Cooley J, Shcherbina A, *et al.* Variable expression of WASP in B cell lines of Wiskott-Aldrich syndrome patients. *J Immunol* 1997; 158: 4021-5.
- [14] Zhu Q, Watanabe C, Liu T, *et al.* Wiskott-Aldrich syndrome/X-linked thrombocytopenia: WASP gene mutations, protein expression, and phenotype. *Blood* 1997; 90: 2680-9.
- [15] Rae J, Newburger PE, Dinauer MC, *et al.* X-Linked chronic granulomatous disease: mutations in the CYBB gene encoding the gp91-phox component of respiratory-burst oxidase. *Am J Hum Genet* 1998; 62: 1320-31.
- [16] Hertel KJ. Combinatorial control of exon recognition. *J Biol Chem* 2008; 283: 12111-5.
- [17] Buratti E, Chivers M, Kralovicova J, *et al.* Aberrant 5' splice sites in human disease genes: mutation pattern, nucleotide structure and comparison of computational tools that predict their utilization. *Nucleic Acids Res* 2007; 35: 4250-63.
- [18] Sahashi K, Masuda A, Matsuura T, *et al.* *In vitro* and *in silico* analysis reveals an efficient algorithm to predict the splicing consequences of mutations at the 5' splice sites. *Nucleic Acids Res* 2007; 35: 5995-6003.
- [19] Kralovicova J, Vorechovsky I. Global control of aberrant splice-site activation by auxiliary splicing sequences: evidence for a gradient in exon and intron definition. *Nucleic Acids Res* 2007; 35: 6399-413.

Received: March 02, 2012

Revised: May 09, 2012

Accepted: May 16, 2012

© Oshima *et al.*; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.