Exon skipping by mutation of an authentic splice site of c-*kit* gene in *W/W* mouse

Shin-Ichi Hayashi*, Takahiro Kunisada, Minetaro Ogawa, Kikuo Yamaguchi and Shin-Ichi Nishikawa
Department of Pathology, Institute for Medical Immunology, Kumamoto University Medical School, 2-2-1 Honjo, Kumamoto 860, Japan

Received December 11, 1990; Revised and Accepted January 31, 1991 EMBL accession nos X57211, X57212

ABSTRACT

The murine mutation dominant *white spotting (W)* is in the proto-oncogene, c-*kit*. The receptor tyrosine kinase encoded by this gene has pleiotropic effects on murine development including hemopoietic cells, pigment cells, and germ cells. In this study, mutation in *W* homozygous mouse was identified as a single base substitution (GT→AT) at the 5'-splice donor site of the exon which encodes the transmembrane domain. Two types of aberrant exon skipping resulted from this mutation, occurred in a tissue specific manner. Either transcript lost the exon coding for transmembrane region and therefore the product might not be functional for signal transduction. Any unusual cryptic splice sites were not activated by this mutation as beta-globin gene in beta-thalassaemia. In addition, twelve base pair sequence of the 3'-end of the exon prior to the exon coding for transmembrane domain was found to be alternatively spliced. These findings should provide the genetic base for not only the receptor function but the splicing mechanism.

INTRODUCTION

More than 30 independent mutations have been discovered at the *W* locus of mice, which is now known to include the c-*kit* proto-oncogene (1–5). Mutations of this type III tyrosine kinase receptor gene differ in severity, but affect development of hemopoietic cells, melanocytes and germ cells (1, 6). Analysis of genomic and/or cDNA from some of these mutations has revealed deletions and insertions, as well as nucleotide changes affecting the tyrosine kinase domain (3, 7–10). This provides a unique opportunity to study structure-function relationships in a large series of naturally occurring mutations of a mammalian gene. We have now found that at least one of these mutations can be informative about mRNA splicing mechanisms.

Recent findings indicate an essential role for conserved sequences at 5' and 3' splice junctions (11, 12). Genetic variants in these have been found to reduce or abolish splicing and, in some cases, to result in the activation of cryptic splice sites (13–18). A previous study of the cDNA from *W* mutant mice revealed a shortened transcript, which lacked 234 base pairs (bp) corresponding to the transmembrane domain of c-*kit* (9). By investigation of the genomic sequence of c-*kit* of this mutant, we now report that a point mutation accounts for this disorder. The obligatory GT within the donor splice junctional site of the exon for the transmembrane region is changed to an AT. Of particular interest was the finding that this mutation revealed in two aberrantly spliced transcripts, which were expressed in a tissue specific manner. These transcripts were produced by two patterns of exon skipping. Such findings are relevant to mechanisms which normally regulate pre-mRNA splicing events. We also found the alternatively spliced sequence corresponding to 3 amino acids at the 3' end of the exon prior to the exon coding for the transmembrane region.

MATERIALS AND METHODS

Mice

WB-*W*/+ mice were purchased from Shizuoka Animal Laboratory Co. Ltd. (Shizuoka, Japan). *W/W* homozygous and +/+ littermates were obtained by mating of *W*/+ heterozygous parents. Seven days after birth, *W/W* and +/+ mice were identified by their skin color.

Preparation of IL-3 dependent mast cells

Mast cells were prepared from 7 day-old bone marrow in RPMI-1640 (Gibco, Grand Island, NY) supplemented with 10% calf serum (Hyclone, Logan, UT), 5×10^{-5} M 2-mercapto-ethanol and 50 U/ml murine recombinant IL-3 as described (19).

Preparation and analysis of RNA

Total RNA was prepared from newborn brains and IL-3 dependent cultured mast cells. mRNA was prepared by using oligo(dT) column, if necessary (20). For Northern blot analysis of c-*kit* transcript, 20 μg of each total RNA was separated by 1.2% agarose–formaldehyde gel electrophoresis, and blotted onto nitrocellulose membrane (Hybond-C, NEN). *Ava*I–*Sal*I fragment of c-*kit* cDNA clone (pUC19Nb2-10) was labeled by random primer methods (Pharmacia, Uppsala, Sweden) and used

* To whom correspondence should be addressed

as a hybridization probe. Relative amount of c-*kit* transcript was measured by Bio-Image Analyzer BA2000 (Fuji Film, Tokyo, Japan).

DNA isolation and analysis

Cellular DNA from 7 day-old brain and whole body was isolated by the standard technique (20). Genomic DNA was digested by restriction enzymes, separated by 1% agarose gel electrophoresis and blotted onto Hybond-C membrane by alkaline transfer. The filter baked at 80°C for 2 hrs and hybridized overnight at 42°C. *Mva*I−*Dde*I fragment of pUC19Nb2-10 was used as a hybridization probe.

Isolation of genomic DNA for c-*kit*

Murine genomic library ligated to lambdaEMBL3 arm was purchased from Clonetech (Palo Alto, CA). The recombinants were grown on *E. coli* strain LE392 as a host and were plated at 5×10^4/dish. After transfer to nitrocellulose filters and denaturation of the bound DNA (20), plaques were screened for the presence of c-*kit* gene by using the ^{32}P-labeled *Sph*I−*Dra*I and *Dra*I−*Sph*I fragments from cDNA for c-*kit* as probes.

cDNA synthesis

cDNA was prepared from 10 μg total RNA or 2 μg poly(A$^+$) RNA by using a cDNA Synthesis Kit (BRL) with oligo(dT) primer then used as a template for polymerase chain reaction (PCR) amplification of the c-*kit* coding region (20).

PCR amplification and determination of nucleotide sequences

Oligonucleotide primers for PCR amplification were synthesized by an automated DNA synthesized (ABI) (21). The primers were as follows: Sense primers; #05 GAGCTCAGAGTCTAGCG CAGCCAC (1→26); #S11 GAATGGATCCAGGAAAAA-GCCGAGG (272→296); #S1 CCCCAAACCCGAGCACC AGC (1057→1076); #02 CTCCAAGAATTGTATTCACA (2374→2393); #S13 CAACTGCAGGGTGGAGTGTAAGG CCTCC (1492→1519); #S10 GCCTTCTTTAACTTTGCA TTTAAAG (1541→1565); #W1 CATTTAAAGGTAACAA CAAAG (1557→1565+12); #S200 ATGGCATGCTCCAG-TGTGTG (1305→1325); #S130 ACACTCTGTTCACGCC-GCTG (1581→1600); #S30 CCTTATGATCACAAATGGGAG (1754→1774); #S1000 GTTCCCTGAATGTGCCATGA (intron a, see Figure 4), Antisense primers; #AS5 CCTCGGC-TTTTTCCTGGATCCATTC (296→272); #AS2 GCTGGT GCTCGGGTTTGGGG (1076→1057); #AS1 CTTGGAGTCGAC CGGCATCC (2374→2393); #AS01 TGTTGGAC-TTGGGTTTCTGC (2976→2957); #AS21 CAATGATCCCC-ATCGCGCCAGCTG (1643→1620); #AS6 CCGTGCATGC GCCAAGCAGG (2009→1990); #AS120 ACTCAGCC-TGTTTCTGGGAA (1795→1776); #AS30 CCTTCCCGAA-GGCACCAGCT (1832→1813); #AS111 GAGTGTGGGCCT GGATTTGCT (1583→1565); #AS150 GCTGTACTGTCGG-TATTACA (intron a). The condition used was; 2.5 U *Taq* DNA polymerase (BRL, Bethesda, MD), 100 pmol of primers in 50 μl 2 mM $MgCl_2$, 50 mM KCl, 25 mM TAPS buffer (pH 9.3), 1 MM DTT and 25−30 cycles in a Program Temp Control System PC-500 (ASTEC, Kitakyushu, Japan). Amplified sequences were purified by agarose gel electrophoresis, digested with appropriate restriction enzymes or converted to blunt end products by using DNA blunting kit (Takara, Tokyo, Japan), subcloned into Bluescript (pBS) M13$^+$ (Stratagene, San Diego, CA) or pUC19, and sequenced by dideoxy method (Sequenase, UBS, Cleveland, OH).

RESULTS

Detection of two aberrant c-*kit* transcripts

IL-3 dependent mast cells propagated from bone marrow and newborn brain were used as two sources of RNA from mutant *W/W* and wildtype control mice. A series of oligonucleotides (Figure 1) and the polymerase chain reaction (PCR) were then used to prime and amplify cDNA. As expected, shortened products were obtained from mutant brain RNA with two sets of primers [labeled 3 (S200-AS6) and 4 (S1-AS1) in Figure 1]. This region (*Ava*I−*Sal*I fragment) was subcloned into Bluescript M13$^+$ and sequenced, confirming the 234 bp deletion of the transmembrane region previously reported by Nocka *et al.* (9).

A more detailed analysis was performed with two additional pairs of primers (S10-AS1 and W1-AS1). Again, evidence for shortened transcripts was found in the mutant RNA samples. One identical amplified band was produced in RNAs isolated from either brain or bone marrow cells and with either set of primers (Figure 2). However, of particular interest was the finding of one unique, and slightly larger, amplification product only in

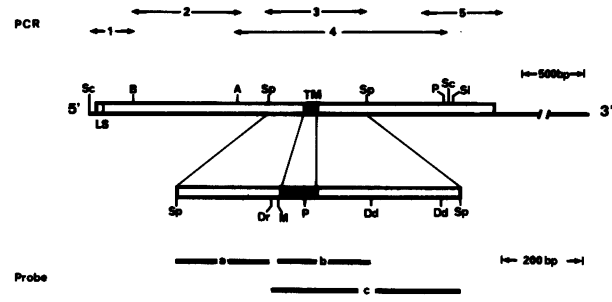

Figure 1. Schematic representation of the murine c-*kit* cDNA (4). The coding sequence is indicated by a box. The locations of recognition sites for restriction enzymes are given as follows; A, *Ava*I; B, *Bam*HI; Dd, *Dde*I; Dr, *Dra*I; M, *Mva*I; P, *Pvu*II; Sc, *Sac*I; Sp, *Sph*I. The coding sequence of c-*kit* was divided into five overlapping segments as indicated by solid lines with arrows over the coding sequence for PCR amplification. Segment 1 was amplified by primer 05 and AS5, segment 2 by S11 and AS2, segment 3 by S200 and AS6, segment 4 by S1 and AS1, and segment 5 by 02 and AS01. The probes designated by solid lines under the coding sequence were used later for Southern blot analysis (b), and for screening genomic library (a and c). LS and TM indicate leader sequence and transmembrane region, respectively.

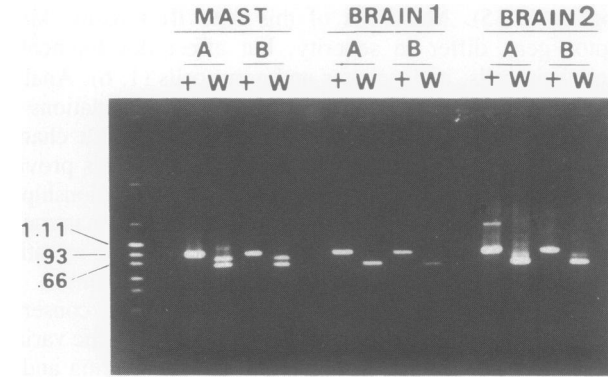

Figure 2. Size analysis of PCR products from cDNA of *W/W* and +/+ littermate. cDNAs for PCR templates were prepared from total RNA of IL-3 dependent mast cells (MAST) and whole brains (BRAIN1), and from poly(A)$^+$ RNA of brain (BRAIN2). Fragments were defined by the combination of primers S10 and AS1 (A) or W1 and AS1 (B) (See Figure 4).

RNA from mast cells of *W/W* anemic mice. The two products were subcloned and sequenced, revealing a deletion of only 107 bp at the transmembrane region in the larger one, as compared to 234 in the small fragment. Thus, one of two mutated transcripts may be expressed in a tissue specific manner.

Southern blot analyses

Genomic DNA of normal and mutant mice was digested with a series of restriction enzymes and subjected to Southern blot analysis (Figure 3). A *Mva*I−*Dde*I fragment corresponding to the transmembrane region was employed as a probe and one of the restriction sites, *Pvu*II, is located in the same gene segment (see Figure 1). However, no evidence for rearrangement or deletion was found that could account for the loss of 234 bp from the transcribed mRNA. This finding indicated that a mutation must have occurred in or near the exon encoding the transmembrane region and that this resulted in abnormal mRNA splicing. One possibility is that the deleted 234 bp constituted an exon and that the 107 bp deletion resulted from activation of a cryptic splice site within that exon. Candidate sequences for a splice acceptor site were present immediately upstream of the 107 bp deletion. As another possibility, the coding sequences could have been derived from two or more exons in the transmembrane domain with two or more types of exon skipping in the mutant species.

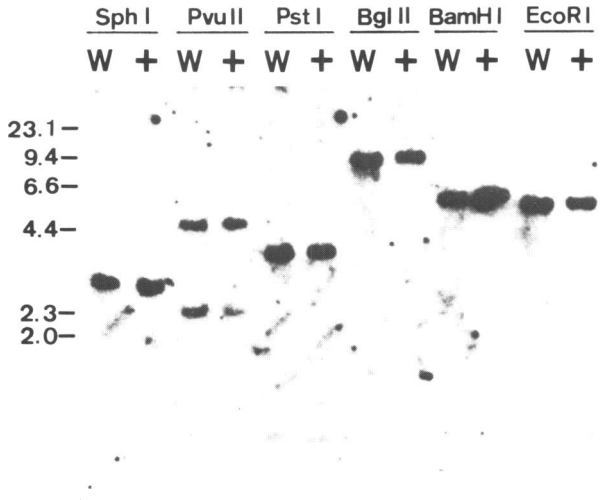

Figure 3. Southern hybridization analysis of the organization of the c-*kit* coding region. The probe was the *Mva*I−*Dde*I fragment of the cloned cDNA of c-*kit* (pUC19Nb2−10) as shown in Figure 1.

Single-base changes in the splice donor site downstream of the transmembrane exon affect splicing

To assess these possibilities, we utilized the *Sph*I−*Dra*I fragment of c-*kit* to isolate four clones from a mouse genomic DNA library in EMBL3. The *Eco*RI-fragment (5 kb) containing the transmembrane exon was subcloned into pUC19 and the intron sequences determined with a series of oligonucleotides used as primers. Each intron contained typical consensus sequences, with GT and AG dinucleotides at the 5′ and 3′ intron boundaries as well as polypyrimidine tracts and putative branch points within 24−39 nucleotides of the 3′ splice sites (11, 22). An intron was found in the region corresponding to the probable deletion in the *W/W* transcripts (Figure 4). This result indicated that the 107 bp deletion could have resulted from a transmembrane region exon skipping and the 234 bp deletion from skipping of two exons in this area, plus the following exon. The 107 bp deletion would generate a stop codon 12 bp downstream because of a frame shift, whereas the larger deletion would still be in-frame. The skipping would have occurred at authentic splice sites.

Appropriate primers were then synthesized and used to amplify genomic DNA fragments from control and *W/W* mutant mice. These were subcloned into Bluescript M13$^+$ and sequenced at the exon-intron junctional regions. All sequences from *W/W* DNA exactly corresponded to those from normal wild type animals, with the single exception of an adenosine residue substituted for a guanosine as the first nucleotide at the 5′ boundary of the intron which follows the transmembrane exon (Figure 5).

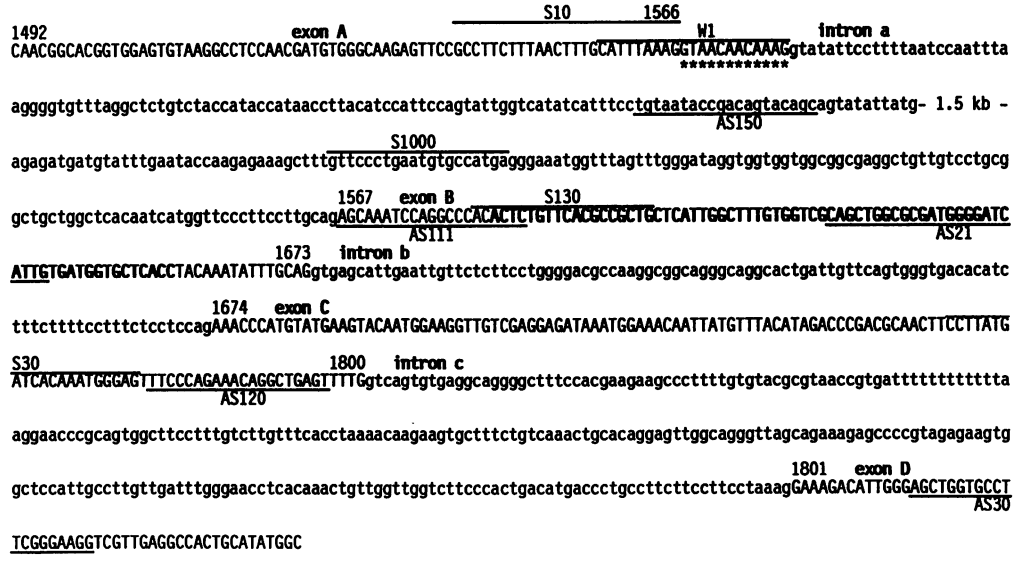

Figure 4. A part of the nucleotide sequence of the c-*kit* gene including transmembrane region. The alternatively spliced twelve nucleotide sequence is marked by *. The sequence corresponding to transmembrane region is indicated by bold letters. Sequences for the sense and antisense strand PCR primers are indicated by overlines and underlines, respectively. The sequence corresponding to its cDNA is numbered as described by Qiu *et al.* (4).

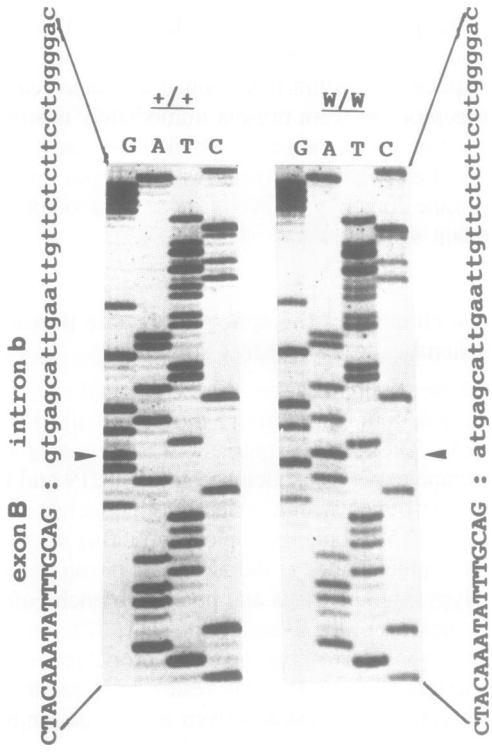

Figure 5. A single base pair substitution found in c-*kit* of *W/W* mouse. PCR amplification was carried out by using genomic DNA of +/+ normal littermate or *W/W* mice as templates and primers used were S130 and AS120. Amplified PCR fragments were subcloned into pUC19 and sequenced.

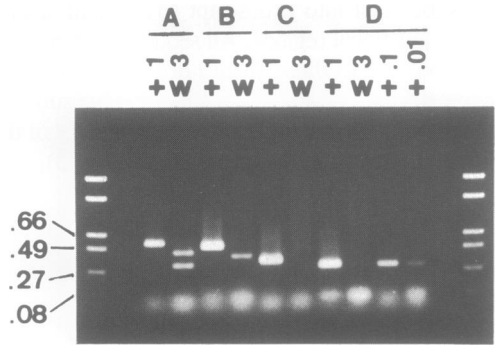

Figure 6. Possible cryptic sites near the mutation point were not activated in *W/W* mast cells. The amount of cDNA from IL-3 mast cell was indicated based on the total RNA used for the template of reverse transcriptase (μg). Each amplified sequence were defined by primers S200 and AS30 (A), AS120 (B), AS21 (C) or AS111 (D).

Conversion of the GT to an AT at this position probably caused this segment to no longer be recognizable as a splice donor site. Previous examples of such mutations include the beta-N-acetylhexosaminidase A gene in Tay-Sachs disease (13), the phenylalanine hydroxylase gene in phenylketonuria (14), and the factor IX gene in Hemophilia B (15). In these cases, either normal or aberrant transcripts were not detected. The beta-globin gene is mutated in one form of beta-thalassaemia, such that three cryptic 5′ splice sites surrounding an authentic site are activated to yield aberrant transcripts (16, 17). Northern blot analysis revealed that mast cells from *W/W* mice have only 31−37% the

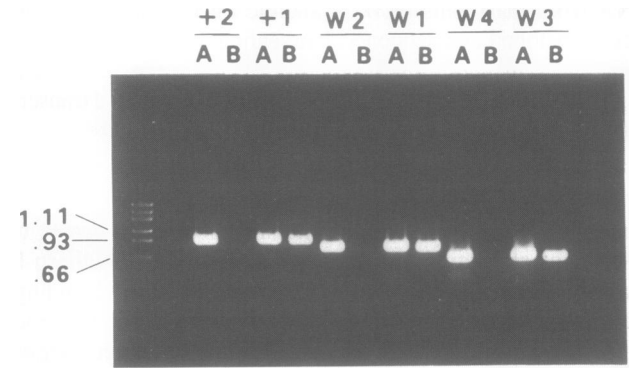

Figure 7. Presence of alternatively spliced products in normal or *W/W* c-*kit* transcripts. PCR reaction was carried out by using cDNAs of IL-3 mast cell as template with S200 and AS6 primers. Each amplified product was subcloned into pUC19. Second round PCR was carried out by primers S10 and AS6 (A) or W1 and AS6 (B). As indicated in Figure 8, S10 can recognize both of the alternatively spliced products and W1 can only recognize the larger product containing 12 bp sequence.

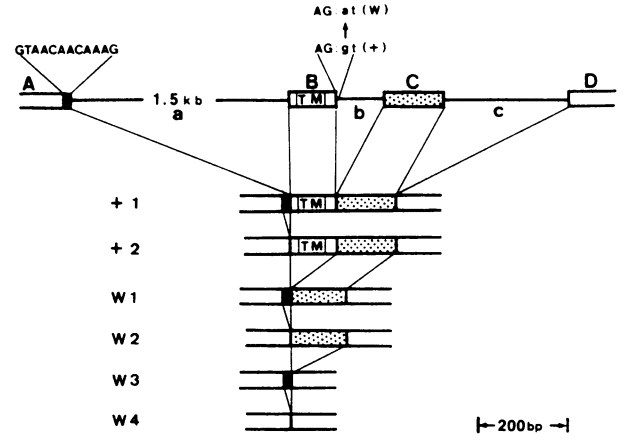

Figure 8. Schematic representation of the intron-exon structure of c-*kit* gene and its cDNA surrounding the transmembrane region.

abundance of c-*kit* transcripts as mast cells derived from normal control bone marrow (data not shown). Thus, the mutation may reduce the efficiency, as well as the authenticity of transcription and splicing.

Cryptic splice sites are not activated by the *W/W* mutation

A total of 8 candidates were identified which might function as cryptic splice sites, 5 were GT sequences in the transmembrane exon and 3 were GT sequences located in the following intron. Appropriate primers (sense primer S200 and antisense primer AS111, see Figure 4) were used to amplify these regions from graded amounts of cDNA prepared from mast cells of normal and mutant mice (Figure 6). Even when 300 times as much *W/W* was used, no product was amplified corresponding to use of these cryptic sites. Thus, only exon skipping results from the single nucleotide substitution of *W/W* mice.

Utilization of a 5′ donor site for alternative mRNA splicing

By analyzing sequences surrounding the transmembrane region, we found a previously unreported segment (4). This 12 bp sequence is in the exon which is immediately upstream of the

transmembrane region exon and it is expressed in normal as well as aberrant transcripts. A PCR analysis of brain and mast cell cDNA's revealed two isoform splice products (Figures 2 and 7). From the chromosomal sequence analysis described above, the two cDNA's must be generated by alternative use of two 5′ splice donor sequences in the intron downstream of the transmembrane exon. When we randomly picked up the clones by using PCR (primers S200 and AS6), in 24 clones derived from cDNA, nine were positive and 15 were negative for the 12 bp segment.

DISCUSSION

A single nucleotide substitution in the c-*kit* gene was found to be the mutation of *W/W* anemic mice. The change is precisely at the splice donor consensus sequence of the intron juxtaposed to the 3′ of the exon encoding the transmembrane domain of this receptor tyrosine kinase. Analysis of genomic DNA organization and sequencing of PCR products from cDNA revealed that two patterns of exon skipping occur in processing of c-*kit* transcripts. Those derived from brain and half of those recovered from mast cells skipped not only the transmembrane exon (B), but the next exon (C) downstream (Figure 8). Our findings define the molecular mechanism which is the basis of cellular defects in one of a large series of mutants known to occur at this important locus. The results are also informative about sequence constraints for normal spliceosome recognition (11, 12, 22).

A schematic diagram of the transmembrane region of the c-*kit* locus and surrounding area is shown in Figure 8. Evidence has been obtained for 4 exons (A to D) and 3 introns (a to c). These are utilized for production of two transcripts in normal mice and 4 different transcripts in mutant *W/W* mice. The intron between exons B and C is relatively short (99 nucleotides), but apparently sufficient to influence the efficiency and authenticity of mRNA splicing (23). Although single exon skipping might be thought to be easier for the splicing machinery, double exon skipping was the dominant pattern observed. Only in mast cells, which expressed c-*kit* in very high levels, was the single exon loss apparent.

In general, when 5′ splice donor sites are inactivated through mutation, levels of transcripts dramatically decrease (13–15, 18). It is thought that the life-time of primary transcripts vary due to differences in intronic sequences (11). Mutations in introns which affect the efficiency and prolong the process of splicing should decrease the abundance of primary transcript, relative to mature mRNA (11). In the case of one mutation of the beta-globin gene, cryptic splice sites are utilized, resulting in aberrant transcripts (16, 17). However, the *W/W* mutation of c-*kit* results in exon skipping, rather than use of cryptic splice sites. The abundance of transcripts was reduced, but still approximately one third normal in mutant cells. The efficiency of splicing over one or two exons may only be reduced by this magnitude because intron (b) is relatively short.

It has been suggested that when comparable splice sites are present in *cis*, the splicing machinery favors use of the internal ones (24). We identified an alternative splice site in exon (A). However, the relative utilization of internal versus external sites in this case was 3:5. The human c-*kit* cDNA sequence contains an identical 12 bp sequence in the same position (25). This insert (Gly-Asn-Asn-Lys) immediately upstream of the transmembrane region, may change the conformation of the c-*kit* molecule, its affinity for ligand, and/or its tyrosine kinase function. It is known that alternatively spliced products of the insulin receptor and G-CSF genes give rise to products with dramatically different affinities for ligand or receptor (26, 27). More information is needed about the functional consequences of the various mutations of the c-*kit* gene.

The severe defect in c-*kit* expression on mast cells of *W/W* mice has now been exploited to prepare monoclonal antibodies to this transmembrane receptor tyrosine kinase (M.Ogawa *et al.* submitted). Some of these reagents block the normal function of this receptor on cells in vitro (S.Nishikawa *et al.* submitted). The role of c-*kit* in development can thus be studied and it is now possible to determine if mutant gene products appear on cell surfaces, or are thwarted at translational stages.

ACKNOWLEDGMENTS

We thank Dr Paul W.Kincade (Oklahoma Medical Research Foundation) for critical review of this manuscript; Ms Chizu Furukawa for her technical assistance. This work was supported by grants from RIKEN, from the Ministry of Education, Science, and Culture in Japan, and from the Mochida Memorial Foundation for Medical and Pharmaceutical Research.

REFERENCES

1. Russell,E.S. (1979) *Adv. Genet.* **20**, 357–453.
2. Chabot,B., Stephenson,D.A., Chapman,V.M., Besmer,P. and Bernstein,A. (1988) *Nature* **355**, 88–89.
3. Geissler,E.N., Ryan,M.A. and Housman,D.E. (1988) *Cell* **55**, 185–192.
4. Nocka,K., Majumder,S., Chabot,B., Ray,P., Cervone,M., Bernstein,A. and Besmer,P. (1989) *Gene Dev.* **3**, 816–826.
5. Qiu,F., Ray,P., Brown,K., Barker,P.E., Jhanwar,S., Ruddle,R.H. and Besmer,P. (1988) *EMBO J.* **7**, 1003–1011.
6. Kitamura,Y. (1989) *Ann. Rev. Immunol.* **7**, 59–76.
7. Tan,J.C., Nocka,K., Ray,P., Trakman,P. and Besmer,P. (1990) *Science* **247**, 209–212.
8. Reith,A.D., Rottapel,R., Giddens,E., Brady,C., Forrester,L. and Bernstein,A. (1990) *Gene Dev.* **4**, 390–400.
9. Nocka,K., Tan,J.C., Chiu,E., Chu,T.Y., Ray,P., Traktman,P. and Besmer,P. (1990) *EMBO J.* **9**, 1805–1813.
10. Lyon,M.F., Glenister,P.H., Loutit,J.F., Evans,E.P. and Peters,J. (1984) *Genet. Res.* **44**, 161–168.
11. Padgett,R.A., Grabowski,P.J., Konarska,M.M., Seiler,S. and Sharp,P.A. (1986) *Ann. Rev. Biochem.* **55**, 1119–1150.
12. Jacob,M. and Gallinaro,H. (1989) *Nucl. Acids Res.* **17**, 2159–2180
13. Myerowitz,R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 3955–3959.
14. DiLella,A.G., Marvit,J., Lidsky,A.S., Guttler,F. and Woo,S.L.C. (1986) *Nature* **322**, 799–803.
15. Ress,D.J.G., Rizza,C.R. and Brownlee,G.G. (1985)*Nature* **316**, 643–645.
16. Wieringa,B., Meyer,F., Reiser,J. and Weissmann,C. (1983) *Nature* **301**, 38–43.
17. Treiman,R., Orkin,S.H. and Maniatis,T. (1983) *Nature* **302**, 591–596.
18. Esumi,H., Takahashi,Y., Sato,S., Nagase,S. and Sugimura,T. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 95–99.
19. Nakano,T., Sonoda,T., Hayashi,C., Yamatodani,A., Kanayama,Y., Yamamura,T., Asai,H., Yonezawa,T., Kitamura,Y. and Galli,S.J. (1985) *J. Exp. Med.* **162**, 1025–1043.
20. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
21. Saiki,R.K., Gelfand,D.H., Stoffel,S., Scharf,S.J., Higuchi,R., Horn,G.T., Mullis,K.B. and Erlich,H.A. (1985) *Science* **239**, 487–491.
22. Smith,C.W.J., Porro,E.B., Patton,J.G. and Nadal-Ginard,B. (1989) *Nature* **342**, 243–247.
23. Wieringa,B., Hofer,E. and Weissmann,C. (1984) *Cell* **37**, 915–925.
24. Reed,R. and Maniatis,T. (1986) *Cell* **46**, 681–690.
25. Yarden,Y., Kuang,W.-J., Yang-Feng,T., Coussens,L., Munemitsu,S., Dull,T.J., Chen,E., Schlessinger,J., Francke,U. and Ullrich,A. (1987) *EMBO J.* **6**, 3341–3351.
26. Mosthaf,L., Grako,K., Dull,T.J., Coussens,L., Ullrich,A. and McClain,D.A. (1990) *EMBO J.* **9**, 2409–2413.
27. Nagata,S., Tsuchiya,M., Asano,S., Yamamoto,O., Hirata,Y., Kubota,N., Oheda,M., Nomura,H. and Yamazaki,T. (1986) *EMBO J.* **5**, 575–581.