Dissertation

Studies on the mechanisms for posttranscriptional regulation of gene expression of Friend murine leukemia virus

2014

Soka University

Graduate School of Engineering

Akihito Machinaga

Studies on the mechanisms for posttranscriptional regulation of gene expression of Friend murine leukemia virus

September, 2014

Akihito Machinaga

CONTENTS

Chapter 2

A 38 bp region within the *gag* gene of Friend murine leukemia virus is crucial for splicing at the correct 5' and 3' splice sites

INTRODUCTION

Friend murine leukemia virus (MLV) is a member of simple retroviruses in the *Retroviridae* family, with a genome that contains (from 5' to 3') a 5' LTR, 5' leader sequence, *gag*, *pol*, *env*, and 3' LTR. The *gag* gene encodes the structural proteins of the virion and the *pol* gene encodes a protease, reverse transcriptase and integrase. The *env* gene encodes the Env protein, which has a surface domain (SU) and a transmembrane domain (TM). There is a 5' splice site (5'ss) in the 5' leader sequence of the *gag* gene and a 3' splice site (3'ss) in the 3' end of the *pol* gene. Both full-length unspliced and spliced mRNAs are produced in MLV-infected cells. Gag and Pol proteins are translated from unspliced mRNAs and the Env protein is translated from spliced mRNA (Coffin et al., 1997). The MLV Env protein plays important roles both in viral adsorption to host cells and in induction of neuropathogenic disease in MLV-infected mouse and rat hosts (DesGroseillers et al., 1984; Masuda et al., 1993; Masuda et al., 1992; Paquette et al., 1989; Szurek et al., 1988; Takase-Yoden and Watanabe, 1997; Yuen et al., 1986). In previous studies, we showed that the expression level of neuropathogenic A8-MLV Env is correlated with neuropathogenicity (Takase-Yoden et al., 2006; Takase-Yoden and Watanabe, 2005). Therefore, definition of the regulatory mechanism of Env expression and production of *env*-mRNA is important for understanding the functions of the Env protein. We previously found that splicing is important for increasing *env*-mRNA stability and translation (Yamamoto and Takase-Yoden, 2009). However, the detailed mechanism for gene expression of MLV due to splicing is still not clear.

Processing of the 3' ends of mRNA is a factor in regulation of protein expression or translation efficiency. The mRNAs of eukaryotic cells have 3' poly(A) tails that are produced by a two-step reaction involving endonucleolytic cleavage and subsequent poly(A) tail synthesis (Danckwardt et al., 2008; Millevoi and Vagner, 2010). The specificity and efficiency of 3' end processing is determined by the binding of multiprotein complexes to specific elements at the 3' end of the pre-mRNA. The length of the poly(A) tail is similar (approximately 250 A nucleotides) in different mRNAs, and is thought to be determined by interaction between the nuclear poly(A)-binding proteins (PABPN1, PABP2), cleavage/polyadenylation specificity factor (CPSF) and poly(A)-polymerase (PAP). After nuclear export of mRNA, PABPN1 is replaced by the cytosol

poly(A)-binding protein (PABPC), which interacts with the translation initiation factor eIF4G to stimulate translation and regulate mRNA stability (Kuhn and Wahle, 2004; Mangus et al., 2003). Furthermore, PABPC can interact with translation termination factor eRF3, indicating that the poly(A) tail has a role in translation termination and possibly ribosome recycling (Cosson et al., 2002). mRNA polysome structure is another factor that regulates protein expression or translation efficiency. Multiple ribosomes can bind to an mRNA to form a polysome structure, enabling many ribosomes to progress along an mRNA simultaneously to synthesize the same protein. In eukaryotic cells, mRNA bound with a cluster of ribosomes (polysomes) circularize primarily by interaction between the poly(A) binding protein and the translation initiation factor to bind the mRNA 5' end (Gierer, 1963; Warner et al., 1963; Warner et al., 1962). mRNA circularization produces a cycle of rapid ribosome recycling.

Splicing in eukaryotic cells involves a series of reactions catalyzed by the spliceosome. The spliceosome is a ribonucleoprotein complex that removes introns from precursor mRNA (Kramer, 1996; Smith and Valcarcel, 2000). For splicing to occur, a branch point is required within the intron in addition to a 5'ss and a 3'ss. The sequences of the 5'ss, 3'ss and branch point are short and have been poorly conserved during evolution; we have frequently identified these sequences in genes. The presence of the 5'ss, 3'ss, and branch point in genes is not sufficient for the proper recognition of exonic and intronic sequences and regulation of splicing. Other regulatory *cis*-elements are required for the recognition of splice sites; these elements are classified according to their location and function and are known as exonic and intronic splicing enhancers (ESEs and ISEs, respectively) or exonic and intronic splicing silencers (ESSs and ISSs, respectively) (Blencowe, 2000; Cartegni et al., 2002; Schaal and Maniatis, 1999; Singh and Valcarcel, 2005). The best characterized ESEs and ISEs have purine-rich sequences that recruit splicing activators from the SR protein family (Graveley, 2000). The splicing activators bind to ESEs and ISEs and interact with spliceosomal factors recruited to the 5'ss and 3'ss. As a result, recognition of appropriate 5'ss and 3'ss is promoted (Busch and Hertel, 2012; Graveley, 2000; Lin and Fu, 2007; Manley and Krainer, 2010; Zhou and Fu, 2013). The best characterized ESSs and ISSs are recognized by members of the hnRNP family including polypyrimidine track-binding protein (PTB) (Black, 2003; Busch and Hertel, 2012). Several mechanisms have been proposed for ESS- or ISS-mediated splicing

suppression: (a) binding of hnRNPs to the ESS blocks the interaction of splicing factors such as SR proteins; (b) the propagation of hnRNP or PTB binding from a site of high-affinity located in an exon or an intron, respectively, block the binding of a splicing factor such as an SR protein; (c) an hnRNP inhibits exon definition when bound to the exon, or inhibits intron definition when bound to the intron; (d) and, for hnRNP A1, H and PTB, interactions between bound hnRNPs loop out portions of a pre-mRNA to promote exon skipping or stimulate intron definition (Blanchette and Chabot, 1999; Martinez-Contreras et al., 2007; Martinez-Contreras et al., 2006; McNally et al., 2006; Singh and Valcarcel, 2005; Spellman and Smith, 2006; Wagner and Garcia-Blanco, 2001). In human immunodeficiency virus type-1 (HIV-1), which produces many types of mRNA by alternative splicing, the *cis*-regulatory sequences responsible for recognition of splice sites, such as ESEs/ISEs or ESSs/ISSs, have been identified (Amendt et al., 1994; Amendt et al., 1995; Exline et al., 2008; Jacquenet et al., 2001; Kammler et al., 2001; Kammler et al., 2006; Madsen and Stoltzfus, 2005; Si et al., 1997; Si et al., 1998; Staffa and Cochrane, 1995; Stoltzfus, 2009). Alternative splicing of HIV-1 is regulated by these elements through the recruitment of splicing regulatory proteins belonging to the SR protein family (SF2/ASF, SC35, 9G8, SRp40, SRp55, and SRp75) or to the hnRNP protein family (hnRNP A/ B, hnRNP H) (Bilodeau et al., 2001; Caputi et al., 2004; Caputi and Zahler, 2002; Domsic et al., 2003; Jacquenet et al., 2005; Stoltzfus and Madsen, 2006; Tange et al., 2001; Tange and Kjems, 2001; Tranell et al., 2010; Zahler et al., 2004; Zhu et al., 2001).

MLV and other simple retroviruses have no regulatory genes, such as those that control gene expression, including splicing events, in lentiviruses. There are, however, several reports of splicing regulation in simple retroviruses. In Moloney MLV, it was reported that the CA-encoding region within the *gag* gene is a negative *cis*-regulatory element and that the region immediately upstream of the 3'ss is a positive *cis*-regulatory element that regulates expression of the unspliced RNA (Hoshi et al., 2002; Oshima et al., 1998; Oshima et al., 1996). Rous sarcoma virus and other members of the avian leukosis virus family contain a unique *cis*-acting element, termed the negative regulator of splicing (NRS), that acts globally to repress splicing to both the *env* and *src* splice sites (Arrigo and Beemon, 1988; Cook and McNally, 1999; Fogel and McNally, 2000; Giles and Beemon, 2005; Katz et al., 1988; McNally and Beemon, 1992; McNally et al., 1991; Stoltzfus and

Fogarty, 1989). Cellular factors, such as the serine/arginine-rich (SR) protein and heteronuclear ribonucleoprotein (hnRNP) H, are recruited to the NRS of Rous sarcoma virus, resulting in the regulation of recognition of splice sites (Fogel and McNally, 2000; McNally and McNally, 1996, 1998). We have also shown that the positive and negative regulatory regions within the intron that controls Env expression in Fr-MLV at the level of *env* mRNA expression do so by controlling splicing efficiency (Yamamoto and Takase-Yoden, 2009). In addition, the proximal sequence of the 5'ss of MLV contains *cis* elements that regulate splicing efficiency (Abbink and Berkhout, 2008; Choo et al., 2013; Kraunus et al., 2004; Zychlinski et al., 2009). Overall, however, the molecular mechanisms that regulate splicing in MLV, including the selection of splice sites, are not well understood.

In this thesis, we focused on splicing to understand mechanisms for posttranscriptional regulation of gene expression of MLV. In Chapter 1, effects of splicing of Friend MLV *env*-mRNA on its 3' end processing and polysome structure formation were analyzed. The results showed that splicing of MLV promoted the efficiency of complete polyadenylation of *env*-mRNA and the formation of *env*-mRNA polysome structures. These splicing-dependent phenomena were not observed with expression vectors in which the *env* gene was replaced by the *luciferase* (*luc)* gene. In Chapter 2, *cis*-elements that regulate splicing were analyzed. It was indicated that the 38 nt fragment within *gag* plays an important role in splicing at the correct 5'ss and 3'ss of MLV. Further analyses showed that a region (1183–1611 nt) upstream of the 38 nt fragment also contains sequences that positively or negatively influence splicing at the correct splice sites. In addition, the 38 nt fragment appears to exert this function in cooperation with a region located just upstream of the 3'ss. In each Chapter, possible molecular mechanisms for posttranscriptional regulation of gene expression of MLV were discussed.

Chapter 1

Effects of splicing of Friend murine leukemia virus *env*-mRNA on its 3' end processing and polysome structure formation

MATERIALS AND METHODS

Cell culture and transfection

Hela cells were grown in Dulbecco's Modified Eagle's Medium - low glucose (Sigma-Aldrich) supplemented with 10% fetal calf serum (MP Biomedicals), 50 units penicillin (Gibco)/ml and 50 μg streptomycin (Gibco)/ml, at 37°C in 5% CO_2. NIH3T3 cells were grown using the same conditions as Hela cells except NIH3T3 cells were incubated in 7% CO2. The cells (1 x 10^6 cells) were transfected with 8 μg *env* or *luciferase (luc)* expression vectors using Lipofectamine 2000 Reagent (Invitrogen) diluted with Opti MEM (Invitrogen) according to the manufacturer's instructions. Construction of the vectors used in this study was described previously (Yamamoto and Takase-Yoden, 2009). However, vectors m1 and m1gpL were previously designated proA8m1 and proA8gpL, respectively. G to T (nt 2608), G to T (nt 2614) and G to T (nt 2629) mutations were introduced into the *pol* region in m1 and m1gpL to suppress progeny virus production. m1 and m1gpL also had A to T (nt 2126) and T to A (nt 2777) point mutations.

Determination of poly(A) tail length

Hela cells were transfected with vectors and total RNA was extracted 24 h post-transfection using an RNeasy Mini Kit (QIAGEN) according to the manufacturer's instructions. After treatment with RNase-free DNase (QIAGEN), 4 μg RNA was added to each ligation reaction. As shown in Figure 1, the anchor primer oligo1 (5'-GGGACAGCCTATTTTGCTAG-3') was ligated to the 3' end of the RNA and reverse transcription (RT) then was carried out using the poly(dT)+oligo2 primer (5'-CTAGCAAAATAGGCTGTCCCTTTTTTTTTT-3'), which has the oligo1 antisense sequence and 10 T nucleotides, or oligo2 primer (5'-CTAGCAAAATAGGCTGTCCC-3'), which is complementary to the oligo1 sequence. Nested PCR was carried out to amplify the viral mRNA

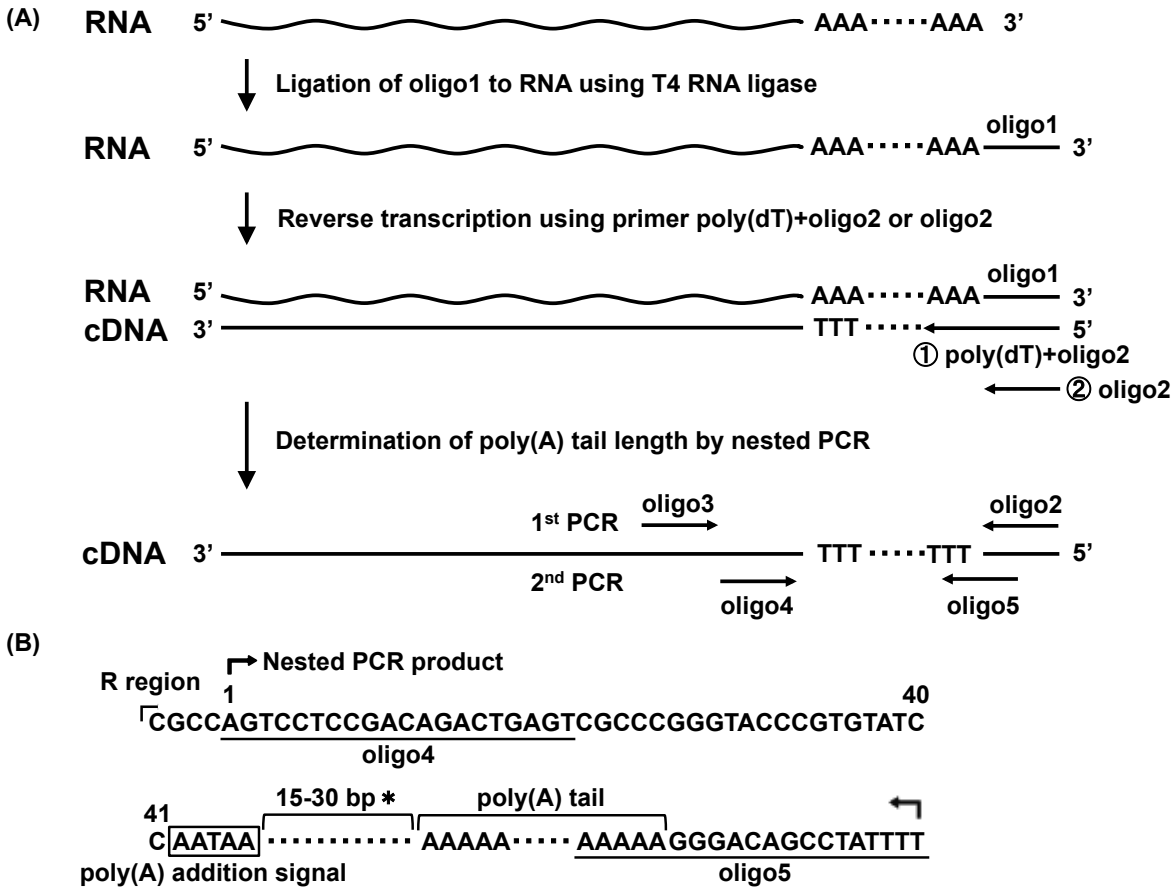


Figure 1 Experimental design in this study. (A) Experimental design to determine the poly(A) tail length of mRNA by nested PCR. (B) Putative structure of the nested PCR product. *: It is generally known that cleavage of the 3' end of mRNA for poly(A) tail addition usually occurs 15-30 nt downstream of the poly(A) addition signal (Elkon et al., 2013).

poly(A) tail. In the first PCR, a forward oligo3 primer targeting the 3' end of U3 in the 3' LTR (5'-GCCCTATAAAAGAGCTCACAACC-3') and a reverse oligo2 primer were used. After purification of the first PCR products using a MicroElute Clean-Up Column (FAVORGEN), a second PCR was performed. In the second PCR, a forward oligo4 primer targeting the 5' end of the R region in the 3' LTR (5'-AGTCCTCCGACAGACTGAGT-3') and a reverse oligo5 primer targeting the 3' end of the oligo2 and poly(dT) sequence (5'-AAAATAGGCTGTCCCTTTTT-3') were used. To determine the poly(A) tail length of *gapdh*-mRNA as a control, in the first PCR, a forward primer hGAPDH3end (5'-ACCACACTGAATCTCCCCT-3') targeting the upstream region of the poly(A) signal and a reverse oligo2 primer were used. After purification of the first PCR products using a MicroElute Clean-Up Column (FAVORGEN), a second PCR was performed. In the second PCR, a forward hGAPDH3end2 (5'-CATGTAGACCCCTTGAAGAG-3') primer targeting the downstream region of hGAPDH3end and a reverse oligo5 primer were used. The resulting PCR products were separated by electrophoresis on a 5% polyacrylamide gel in TBE buffer [50 mM Tris-HCl (pH 8.0), 48.5 mM boric acid, 2 mM EDTA] and stained with ethidium bromide. The size of DNA contained in the band was determined by reference to a standard curve generated using the known sizes of the marker DNAs and their mobilities. To analyze the sequences of the bands separated by electrophoresis, the bands were extracted from the gels using a QIAquick Gel Extraction Kit (QIAGEN) and cloned into a pGEM-T-easy vector (pGEM-T-easy Vector System; Promega). The sequences of the cloned fragments in T-easy vectors were amplified using T7 (5'-GTAATACGACTCACTATAGGGC-3') or sp6 (5'-ATTTAGGTGACACTATAGAA-3') primers and a BigDye Terminator v 3.1 Cycle Sequencing Kit (Applied Biosystems). The sequences were analyzed using an ABI PRISMOR 3100 Genetic Analyzer (Applied Biosystems).

Detection of cDNA synthesized from *env*- and *luc*-mRNA by PCR

To detect the cDNAs from *env*- and *luc*-mRNA in the RT reaction products, after RT synthesis using the poly(dT)+oligo2 primer, PCR with forward primer s1 (5'-GAGACCCTTGCCCAGGGA-3') and reverse primer s2 (5'-TGCCGCCAACGGTCTCC-3') was performed using Go Taq (Promega). The PCR products were separated by electrophoresis on

2% agarose gels in TBE buffer and stained with ethidium bromide. Control samples without the cDNA synthesis step did not yield these specific bands.

Quantitation of cDNA synthesized from vector mRNAs by real-time PCR

To evaluate the copy numbers of cDNAs from *env*- and *luc*-mRNA in the RT reaction products after RT synthesis using the poly(dT)+oligo2 primer, quantitative real-time PCR was performed as described previously (Yamamoto and Takase-Yoden, 2009). Briefly, the primers and probe used to quantitate cDNA of *env*- and *luc*-mRNA were forward s1-primer, reverse s2-primer, and TaqMan ss-probe (5'-CACCACCGGGAGCTCATTTACAGGCAC-3'). These primers were designed to amplify the exon-exon junction region of *env*- and *luc*-mRNA. The copy numbers of cDNAs corresponding to *env*- and *luc*-mRNA were determined by standard curves that were created using serially diluted splA8 plasmid and splA8L plasmid, respectively. The primers and probe used to quantitate cDNA corresponding to total RNA from the m1 and splA8 vectors were forward e1-primer (5'-AGGACCTCGGGTCCCAATAG-3'), reverse e2-primer (5'-TTAGGTAGCGGGAACGAAAGTT-3'), and TaqMan e-probe (5'-CCGAACCCCGTCCTGGCAGAC-3'). The copy numbers of cDNAs corresponding to total RNAs from the m1 and splA8 vectors were determined using standard curves that were created using serially diluted splA8 plasmid. The primers and probe used to quantitate total RNA from the m1gpL and splA8L vectors were forward L1-primer (5'-CGGCTTCGGCATGTTCA-3'), reverse L2-primer (5'-TACATGAGCACGACCCGAAA-3), and TaqMan L-probe (5'-CACGCTGGGCTACTTGATCTGCGG-3'). The copy numbers of cDNAs corresponding to total RNA from the m1gpL and splA8L vectors were determined using standard curves that were created using serially diluted splA8L plasmid. To quantitate cDNA from *gapdh*-mRNA, TaqMan Human GAPDH Control Reagents containing primer sets and a probe (Applied Biosystems) were used. The copy number of cDNA from *gapdh*-mRNA was determined by standard curves that were created by using serially diluted *gapdh* T-easy vector containing a fragment of the human *gapdh* gene. The copy number of cDNA from mRNA of vectors was normalized relative to the copy number of cDNA from cellular *gapdh*-mRNA. Negative control samples without the cDNA synthesis step did not show specific amplification.

Quantitation of vector DNA in transfected cells

To quantitate the vector DNA in transfected cells, cellular genomic DNA was extracted using a DNeasy Blood and Tissue Kit (QIAGEN) according to the manufacturer's instructions. A sample of the resulting DNA was analyzed by real-time PCR using specific primers and the TaqMan probe with a 7500 Real-Time PCR System (Applied Biosystems). The primers and probe used to quantitate the DNA introduced by m1 and splA8 were forward e1-primer, reverse e2-primer, and TaqMan e-probe. The copy numbers of m1 and splA8 DNAs were determined using standard curves that were created using serially diluted m1 plasmid and splA8 plasmid, respectively. The primers and probe used to quantitate the DNA introduced by m1gpL and splA8L were forward L1-primer, reverse L2-primer, and TaqMan L-probe. The copy numbers of DNAs of m1gpL and splA8L were determined using standard curves that were created using serially diluted m1gpL plasmid and splA8L plasmid, respectively. The amount of *gapdh* DNA was measured as an internal control using TaqMan Human GAPDH Control Reagents containing primer sets and a probe (Applied Biosystems). The copy number of *gapdh* DNA was determined using a standard curve that was created using serially diluted *gapdh* T-easy vector plasmids. The copy number of DNA of introduced plasmid was normalized relative to the copy number of *gapdh* DNA.

Fractionation of cell lysates by sucrose density gradient centrifugation

Polysome fractions were obtained by fractionation of cell extracts by sucrose density gradient centrifugation as described previously (Esposito et al., 2010). Briefly, transfected or infected NIH3T3 cells (6×10^6) were incubated in medium containing 100 μg cycloheximide/ml for 15 min. The cells then were lysed in 1 ml hypotonic lysis buffer [1.5 mM KCl, 2.5 mM $MgCl_2$, 5 mM Tris-HCl (pH 7.4), 1% Triton X-100, 1% sodium deoxycholate, 100 μg cycloheximide/ml, 1 mM dithiothreitol, 100 units RNase inhibitor (TaKaRa)/ml]. After 10 min on ice, lysates were centrifuged at 10,000 x g for 10 min and the resulting cytosol-containing supernatant was removed and layered onto a 10-50% sucrose density gradient in buffer [80 mM NaCl, 5 mM $MgCl_2$, 20 mM Tris-HCl (pH 7.4), 1 mM dithiothreitol]. After ultracentrifugation in an SW 41 rotor in an Optima XL-90 ultracentrifuge (Beckman) at 30,000 rpm for 3 h at 4°C, 16 fractions were obtained. The RNA in each fraction was extracted using TRIzol LS Reagent (Invitrogen) and measured by

absorbance at 260 nm using Gene Quant pro (Amersham Biosciences).

Detection of ribosomal RNA

Extracted RNA from each fraction was denatured in sample buffer [50% (v/v) formamide, 1x MOPS （20 mM MOPS, pH 7.0, 8.3 mM NaAc, 1 mM EDTA), 15% (v/v) formaldehyde, 3.35% (v/v) ethidium bromide] and analyzed by electrophoresis on a 1% denaturing agarose gel. The 28S and 18S rRNAs were detected as ethidium bromide-stained bands on the gels.

Quantitation of *env*-, *luc*-, and *gapdh*-mRNA in each fraction by real-time RT-PCR

Extracted RNA from each fraction was treated with RNase-free DNase (QIAGEN). Equal volume samples of each fraction were used as template for reverse transcription (RT) using an oligo(dT) primer (Invitrogen). In its volume of the fraction that had a largest absorbance peak at 260 nm, 2 μg RNA was contained. A portion of the resulting cDNA was amplified by real-time PCR using a 7500 Real-Time PCR System (Applied Biosystems). The primers and probe used to quantitate *env*- and *luc*-mRNA were forward s1-primer, reverse s2-primer, and TaqMan ss-probe. To quantitate *gapdh* mRNA, TaqMan Rodent GAPDH Control Reagents containing primer sets and probe (Applied Biosystems) were used. Standard curves used to calculate the amount of mRNA were produced using serial dilutions of splA8 plasmid and *gapdh* T-easy vector containing a fragment of the Rodent *gapdh* gene. Negative control samples without the cDNA synthesis step did not show specific amplification.

RESULTS

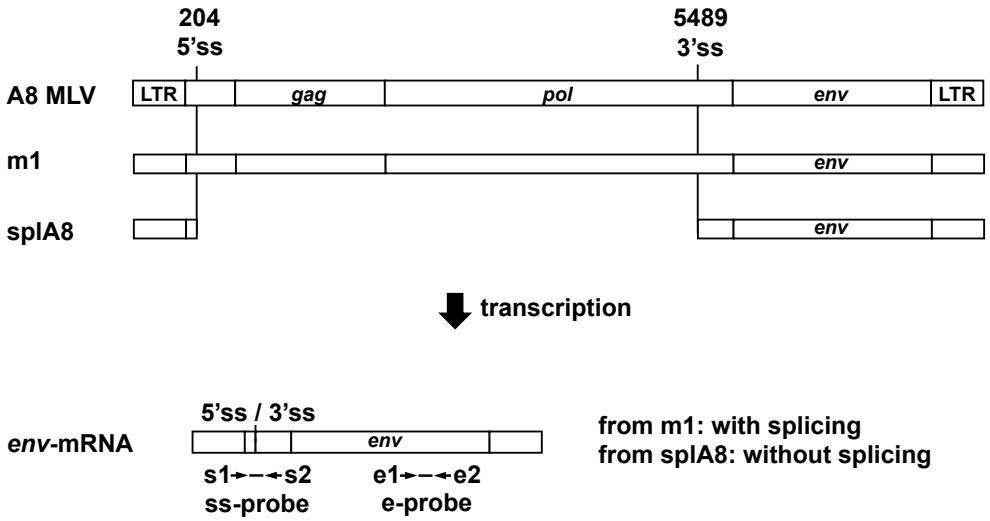
Deteremination of poly(A) tail length of MLV mRNA

The poly(A) tail length of mRNA from MLV was analyzed by the methods as described in Figure 1A. The putative structure of the nested PCR product is also presented in Figure 1B. The m1 vector carried the full-length A8-MLV provirus genome and generated spliced *env*-mRNA (Fig. 2A). Electrophoretic analysis of the products of nested PCR using oligo4 and oligo5 primers showed a smear in the 100-200 bp range in the mRNA from samples obtained from m1-transfected cells (lane 1 of Fig. 2B). Sequence analysis showed that these smears contained multiple A nucleotides 11-21 nt downstream of the poly(A) addition signal (AAUAAA). As a control, the poly(A) tail length of *gapdh*-mRNA was determined in m1-transfected cells. There were no differences in the poly(A) tail smear patterns of the *gapdh*-mRNAs from m1- and mock-transfected cells (bottom of Fig. 2B).

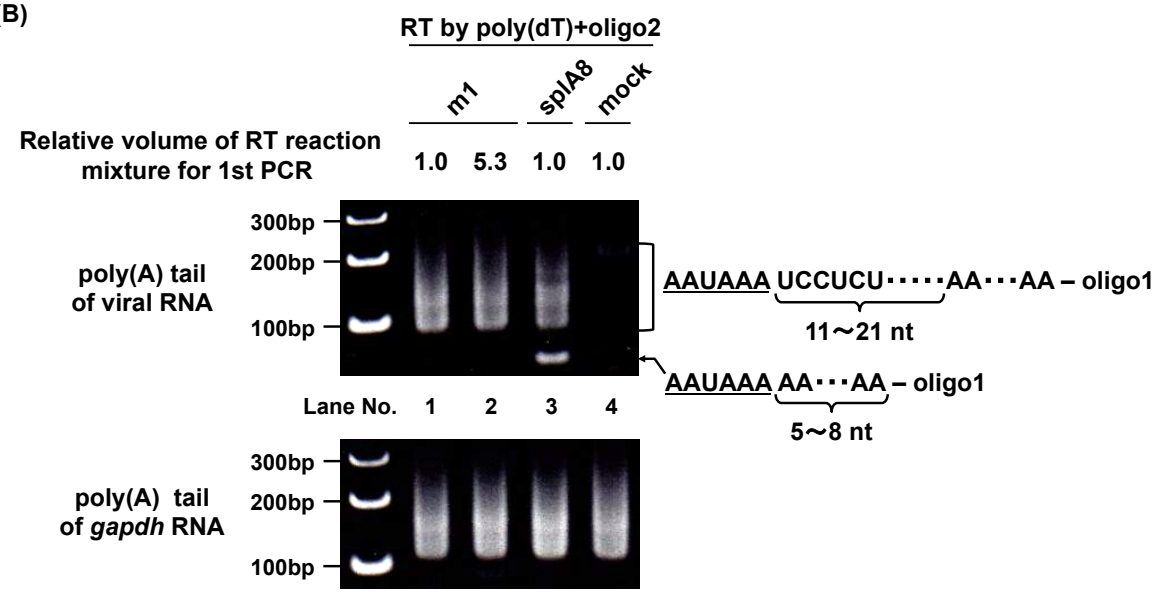
Effects of splicing on the poly(A) tail length of *env*-mRNA

To examine whether *env*-mRNA splicing affected polyadenylation of the mRNA, the poly(A) tail length of spliced *env*-mRNA and unspliced *env*-mRNA were compared using the m1 and splA8 vectors (Fig. 2A). The splA8 vector was designed to generate unspliced *env*-mRNA by deletion of the intron region in m1. Electrophoretic analysis of the nested PCR products showed a smear in the 100-200 bp range for mRNA obtained from splA8-transfected cells (lane 3 of Fig. 2B). Sequence analysis showed that these smears contained multiple A nucleotides 11-21 nt downstream of the poly(A) addition signal (AAUAAA). Interestingly, an approximately 70 bp band was also detected in RNA from splA8-transfected cells. Sequence analysis showed that this band came from mRNA in which 5-8 A nucleotides were attached just downstream of the poly(A) addition signal and the oligo1 sequence (Fig. 2B). This band was not observed in m1-transfected cells, as shown in lane 1 of Figure 2B. In this experimental system, the length of the mRNA poly(A) tail could be determined in both unspliced full-length MLV mRNA and in spliced *env*-mRNA from m1-transfected cells. Therefore, to confirm that the first strand of cDNA synthesized by RT using the poly(dT)+oligo2 primer contained *env*-mRNA, PCR was performed using s1 and s2 primers.

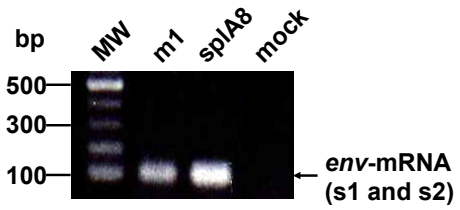
(A)


204
5'ss
5489
3'ss
A8 MLV
LTR
gag
pol
env
LTR
m1
env
splA8
env
transcription
5'ss / 3'ss
env-mRNA
env
s1
s2
ss-probe
e1
e2
e-probe
from m1: with splicing
from splA8: without splicing

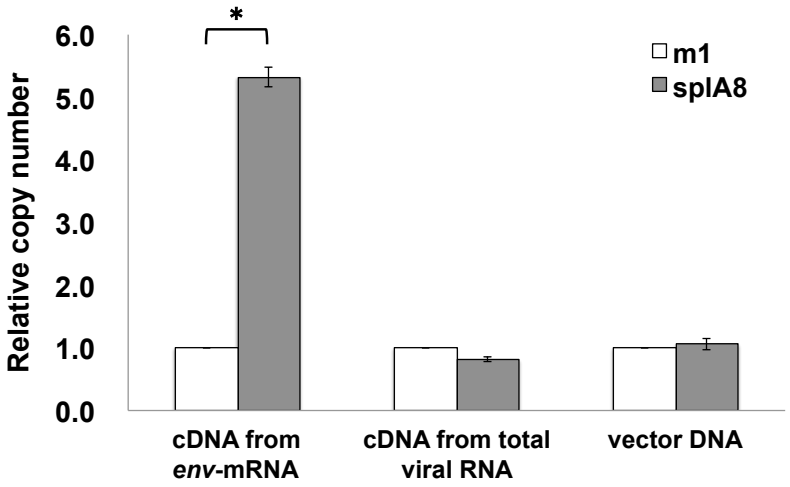
(B)


RT by poly(dT)+oligo2
m1
splA8
mock
Relative volume of RT reaction mixture for 1st PCR
1.0 5.3 1.0 1.0
300bp
200bp
100bp
poly(A) tail of viral RNA
AAUAAA UCCUCU·····AA···AA – oligo1
11∼21 nt
AAUAAA AA···AA – oligo1
5∼8 nt
Lane No. 1 2 3 4
300bp
200bp
100bp
poly(A) tail of gapdh RNA

(C)


bp
MW
m1
splA8
mock
500
300
100
env-mRNA
(s1 and s2)

(D)


*
6.0
5.0
4.0
3.0
2.0
1.0
0.0
Relative copy number
m1
splA8
cDNA from
env-mRNA
cDNA from total
viral RNA
vector DNA

(E)

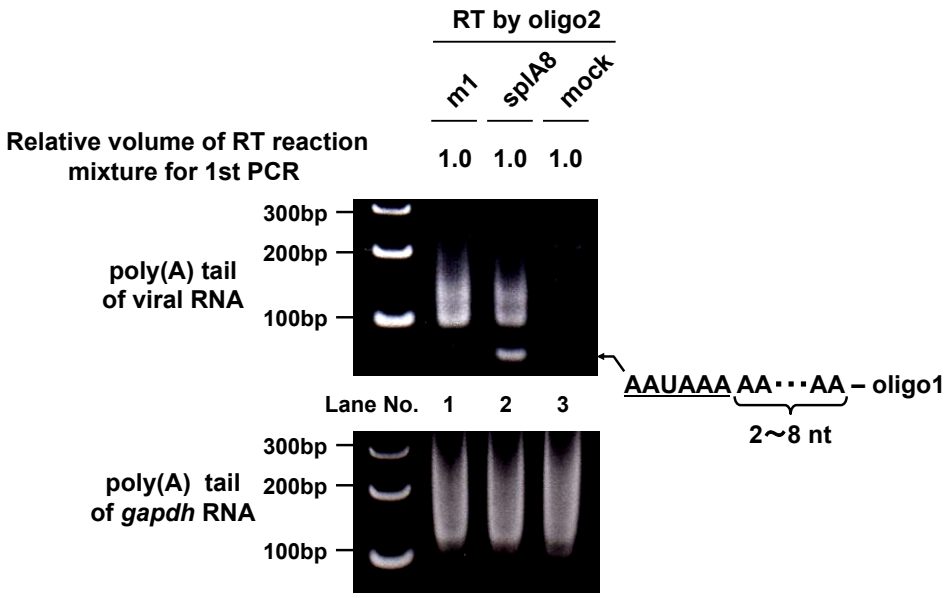

RT by oligo2
m1
splA8
mock
Relative volume of RT reaction
mixture for 1st PCR
1.0 1.0 1.0
300bp
200bp
100bp
poly(A) tail
of viral RNA
AAUAAA AA･･･AA – oligo1
2～8 nt
Lane No. 1 2 3
300bp
200bp
100bp
poly(A) tail
of gapdh RNA

Figure 2 Analysis of poly(A) tail length of mRNA from m1- and splA8-transfected cells. (A) Structures of the MLV provirus genome, the vectors used in this study, and *env*-mRNA. 5'ss, 5' splice site; 3'ss, 3' splice site. (B) Determination of poly(A) tail length of mRNA from m1- and splA8-transfected cells by using the poly(dT)+oligo2 primer for the RT reaction. The poly(A) tail length of the mRNA from each sample was analyzed by nested RT-PCR as described in Materials and Methods and Fig. 1. The PCR products were separated by electrophoresis. As a control, the *gapdh*-mRNA poly(A) tail length was analyzed in m1- and splA8-transfected cells. This figure is representative of 5 independent experiments that yielded similar results. Sequence analysis of the indicated bands in (B) was performed by TA-cloning as described in Materials and Methods. (C) Detection of cDNA synthesized from *env*-mRNA obtained from m1- and splA8-transfected cells with a poly(dT)+oligo2 primer. (D) Measurement of the copy number of cDNA synthesized from *env*-mRNA and total viral RNA with the poly(dT)+oligo2 primer, and the copy number of vector DNA in m1- and splA8-transfected cells by real-time PCR. The amount of cDNA from *env*-mRNA was measured by quantitative real-time PCR using s1 and s2 primers and an ss-probe. The amount of cDNA from total viral RNA and vector DNA was measured by quantitative real-time PCR using e1 and e2 primers and an e-probe. The mean values from 3 independent experiments and SEM are shown. Statistical comparison was performed using the *t* test. *: $p<0.01$ (E) Determination of the poly(A) tail length of mRNA from m1- and splA8-transfected cells by using the oligo2 primer for the RT reaction. The poly(A) tail length of the mRNA from each sample was analyzed by nested RT-PCR as described in Materials and Methods and Fig. 1. As a control, the *gapdh*-mRNA poly(A) tail length was analyzed in m1- and splA8-transfected cells. This figure is representative of 3 independent experiments that yielded similar results.

These primers were designed to amplify the 94 bp fragment containing the splicing junction region in the cDNA from *env*-mRNA. As shown in Figure 2C, 94 bp bands were detected in cDNA of *env*-mRNA from m1- and splA8-transfected cells. All 94 bp bands were confirmed to come from *env*-mRNA by sequence analysis (data not shown). In addition, the copy number of cDNA synthesized from *env*-mRNA was evaluated by quantitative real-time PCR using s1 and s2 primers and the ss-probe. The copy number of cDNA from splA8 *env*-mRNA was 5.3-fold more than that synthesized from m1 *env*-mRNA (Fig. 2D). The copy number of cDNA synthesized from total viral RNA was also evaluated by quantitative real-time PCR using e1 and e2 primers and an e-probe, and it was not significantly different from results obtained for RNA from m1- and splA8-transfected cells. The transfection efficiency, which was measured by the copy number of plasmid DNA in transfected cells, was not significantly different in m1- and splA8-transfected cells. To confirm the finding that the approximately 70 bp band was detected only in splA8-transfected cells that contained unspliced *env*-mRNA but not in m1-transfected cells that contained spliced *env*-mRNA, the amount of template cDNA for the 1st PCR was normalized by the amount of cDNA synthesized from *env*-mRNA; the volume of the RT reaction mixture containing m1 cDNA used as a template for the first PCR was 5.3 times that of the RT reaction mixture containing splA8 cDNA. The nested PCR reactions produced a smear in the 100-200 bp range for mRNA from m1-transfected cells, but the approximately 70 bp band was not detected (lane 2 of Fig. 2B).

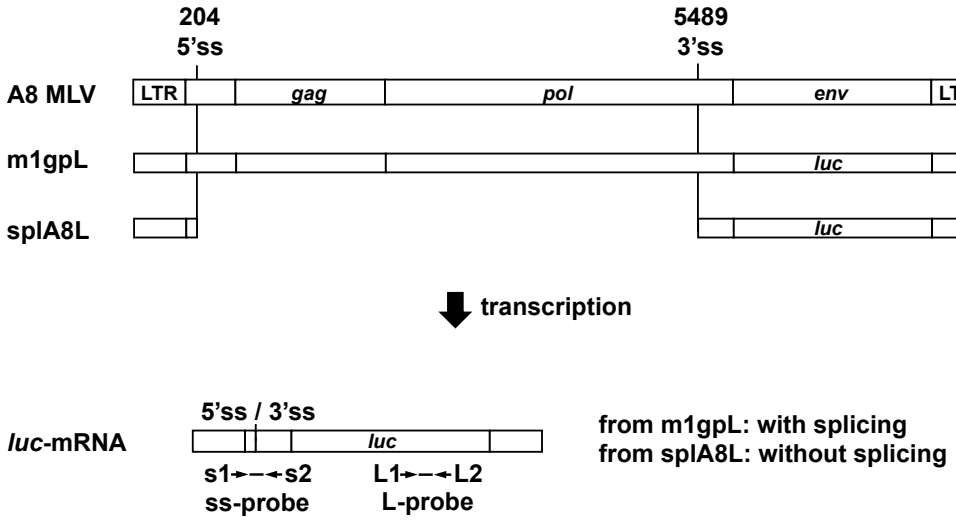
Since the poly(dT)+oligo2 primer used in the RT reaction contains 10 T nucleotides, mRNA with less than 10 A nucleotides cannot be reverse transcribed. To examine whether other kinds of abnormal mRNAs with less than 10 A nucleotides existed in the transcripts of m1 and splA8, the oligo2 primer, which has no T nucleotide, was used for the RT reaction and then nested PCR was performed. Electrophoretic analysis of the nested PCR products showed a smear in the 100-200 bp range in the mRNA of the samples from m1-transfected cells (lane 1 of Fig. 2E). In RNA from splA8-transfected cells, bands in the 100-180 bp range were detected (lane 2 of Fig. 2E). Sequence analysis showed that these smears contained multiple A nucleotides, similar to the smear that was detected when the poly(dT)+oligo2 primer was used for the RT reaction (data not shown). The approximately 70 bp band was also detected in splA8-transfected cells (lane 2 of Fig. 2E). Sequence analysis showed that this band came from mRNA in which 2-8 A nucleotides were attached just

downstream of the poly(A) addition signal and the oligo1 sequence (Fig. 2E). There was no difference in the poly(A) tail smear patterns of control *gapdh*-mRNA from m1-, splA8-, and mock-transfected cells in the RT reactions using the oligo2 primer (the bottom panel of Fig. 2E).

Effects of splicing on the poly(A) tail length of *luc*-mRNA

Previous studies that utilized m1gpL and splA8L vectors in which the *env* gene in m1 and splA8 was replaced by the *luc* gene (Fig. 3A) demonstrated that splicing of *luc*-mRNA did not influence its stability or translation efficiency (Yamamoto and Takase-Yoden, 2009). The effect of *luc*-mRNA splicing on its polyadenylation in cells transfected with m1gpL and splA8L was investigated. As shown in Fig. 3B, 94 bp bands were produced from cDNA synthesized from *luc*-mRNA obtained from m1gpL- and splA8L-transfected cells. The copy number of cDNA from *luc*-mRNA was also measured by quantitative real-time PCR using s1 and s2 primers and an ss-probe. The copy number of cDNA synthesized from *luc*-mRNA from splA8L-transfected cells was 1.9-fold more than that from m1gpL-transfected cells (Fig. 3C). The copy number of cDNA synthesized from total vector RNA was measured by real-time PCR using L1 and L2 primers and an L-probe, and it did not differ significantly between m1gpL- and splA8L-transfected cells. The transfection efficiency also did not differ significantly between m1gpL- and splA8L-transfected cells. As shown in Figure 3D (lanes 1-3 of the upper left panel), when the poly(dT)+oligo2 primer was used for the RT reaction, the nested PCR products showed a smear in the 100-200 bp range for mRNA from all samples obtained from m1gpL- and splA8L-transfected cells, but the approximately 70 bp band was not detected. There was also no difference in the poly(A) tail smear patterns of control *gapdh*-mRNA from m1gpL-, splA8L-, and mock-transfected cells (lower left panel of Fig. 3D). When the oligo2 primer was used for the RT reaction, nested PCR produced a smear in the 100-200 bp range in the mRNA of the samples from m1gpL- and splA8L-transfected cells, but the approximately 70 bp band was not detected (upper right panel of Fig. 3D). There was also no difference in the poly(A) tail smear patterns of control *gapdh*-mRNA from m1gpL-, splA8L-, and mock-transfected cells (lower right panel of Fig. 3D).

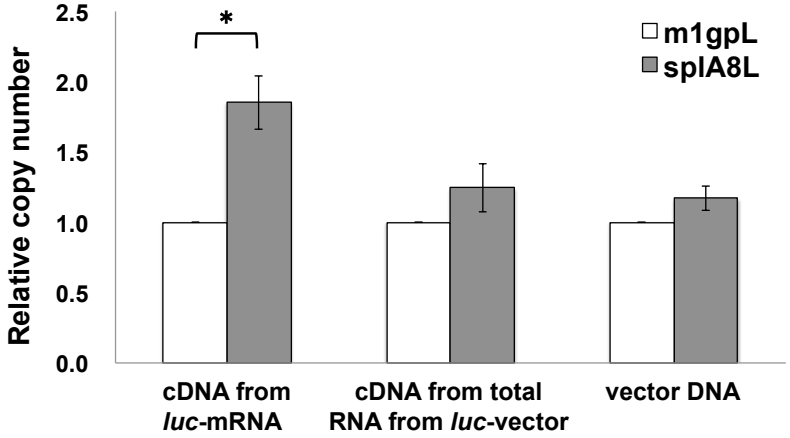
(A)


204
5'ss
5489
3'ss
A8 MLV
LTR
gag
pol
env
LTR
m1gpL
luc
splA8L
luc
transcription
luc-mRNA
5'ss / 3'ss
luc
s1
s2
ss-probe
L1
L2
L-probe
from m1gpL: with splicing
from splA8L: without splicing

(B)


bp
MW
m1gpL
splA8L
mock
500
300
100
luc-mRNA
(s1 and s2)

(C)


Relative copy number
2.5
2.0
1.5
1.0
0.5
0.0
*
m1gpL
splA8L
cDNA from
luc-mRNA
cDNA from total
RNA from luc-vector
vector DNA

(D)

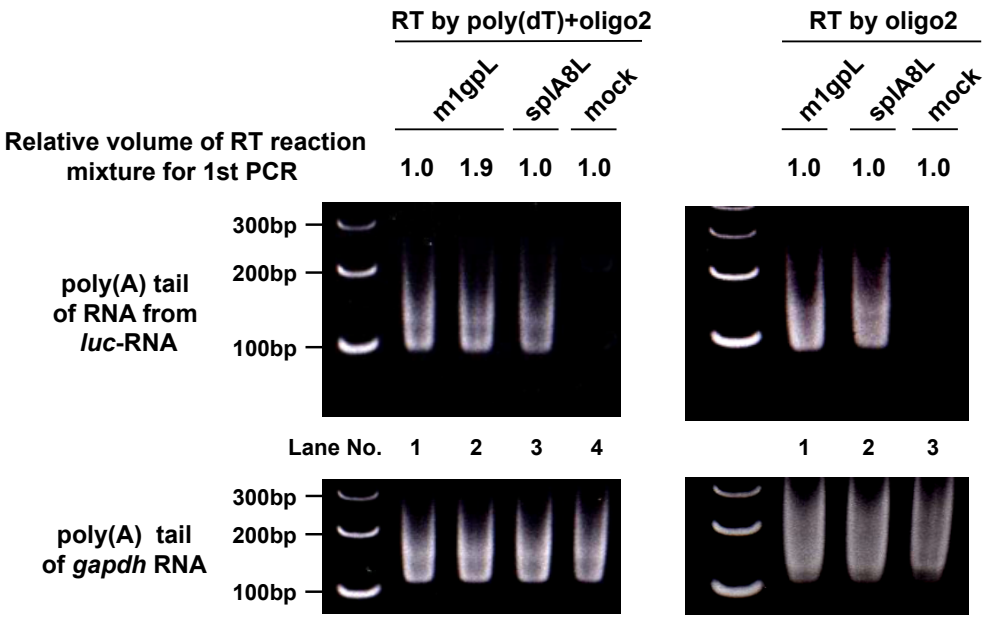

RT by poly(dT)+oligo2
RT by oligo2
m1gpL
splA8L
mock
m1gpL
splA8L
mock
Relative volume of RT reaction mixture for 1st PCR
1.0 1.9 1.0 1.0
1.0 1.0 1.0
300bp
200bp
100bp
poly(A) tail of RNA from luc-RNA
Lane No. 1 2 3 4
1 2 3
300bp
200bp
100bp
poly(A) tail of gapdh RNA

Figure 3 Analysis of poly(A) tail length of mRNA from m1gpL- and splA8L-transfected cells. (A) Structures of the MLV provirus genome, the vectors used in this study, and *luc*-mRNA. 5'ss, 5' splice site; 3'ss, 3' splice site. (B) Detection of cDNA synthesized from *luc*-mRNA obtained from m1gpL- and splA8L-transfected cells with a poly(dT)+oligo2 primer. (C) Measurement of the copy number of cDNA synthesized from *luc*-mRNA and total RNA from the *luc*-vector with a poly(dT)+oligo2 primer, and the copy number of vector DNA in m1gpL- and sp1A8L-transfected cells by real-time PCR. The amount of cDNA from *luc*-mRNA was measured by quantitative real-time PCR using s1 and s2 primers and an ss-probe. The amount of cDNA from vectors and vector DNA was measured by quantitative real-time PCR using L1 and L2 primers and an L-probe. The mean values from 3 independent experiments and SEM are shown. Statistical comparison was performed using the *t* test. *: $p<0.01$ (D) Determination of the poly(A) tail length of mRNA from m1gpL- and sp1A8L-transfected cells by using the poly(dT)+oligo2 primer and the oligo2 primer for the RT reaction. The poly(A) tail length of the mRNA from each sample was analyzed by nested RT-PCR as described in Materials and Methods and Fig. 1. As a control, the *gapdh*-mRNA poly(A) tail length was analyzed in m1gpL- and splA8L-transfected cells. This figure is representative of 5 (RT by poly(dT)+oligo2 primer) or 3 (RT by oligo2 primer) independent experiments that yielded similar results.

Effects of splicing on formation of *env*-mRNA polysome structures

Since the MLV genome has a 5'ss in the 5' leader sequence and a 3'ss in the 3' end of the *pol* gene, both full-length unspliced mRNA and spliced *env*-mRNA are produced from the provirus. We have previously shown that splicing of *env*-mRNA promotes translation of the Env protein (Yamamoto and Takase-Yoden, 2009). Formation of polysome structures of mRNA is generally correlated with mRNA translation efficiency. To investigate whether splicing of MLV affected formation of *env*-mRNA polysome structures, we used the m1 and splA8 vectors to compare the amount of *env*-mRNA associated with polysome structures in spliced and unspliced *env*-mRNA. Lysates of NIH3T3 cells transfected with m1 or splA8 were separated by centrifugation on linear 10–50% sucrose density gradients. After extraction of RNA from each fraction, the distribution of total RNA and ribosomal RNA (rRNA) was analyzed by measurement of absorbance at 260 nm and agarose gel electrophoresis, respectively (Fig. 4). mRNA in polysome structures was found in higher density fractions, while mRNA that was not in polysome structures was in lower density fractions, as reported previously (Akimitsu et al., 2007; del Prete et al., 2007; Holetz et al., 2007; Nashchekin et al., 2006; Otulakowski et al., 2004; Ryu et al., 2008). In lysates of m1-transfected cells, there was a large absorbance peak at 260 nm in fractions 6-9 (Fig. 4A). Agarose gel electrophoresis analysis showed that these fractions contained most of the 28S and 18S rRNAs in the lysate. There were small peaks at 260 nm in higher density fractions 11-16. These fractions also contained 28S and 18S rRNA, in agreement with these fractions containing polysomes. As a control, the distribution of *gapdh*-mRNA was examined by real-time RT-PCR. There were two peaks of *gapdh*-mRNA, one in lower density fraction 8 and the other in higher density fractions 13-15 (Fig. 4A). These results were agreement with a previous report (Akimitsu et al., 2007). The distribution of *env*-mRNA was examined by real-time RT-PCR using s1 and s2 primers and the ss-probe (Fig. 2A). As shown in Fig. 4A, most *env*-mRNA was in fractions 13-16, which were polysome fractions. From quantitative analysis of these real-time PCR results, 69% of *env*-mRNA in m1-transfected cells was in the polysome structures in fractions 11-16. A similar experimental analysis of lysates of splA8-transfected cells showed that a large amount of *env*-mRNA was in fractions 1-8 (Fig. 4B). Although 24% of *env*-mRNA in splA8-transfected cells was in polysome structures (fractions 10-16), this was significantly less ($p<0.01$) than the 69% in polysome structures in m1-transfected

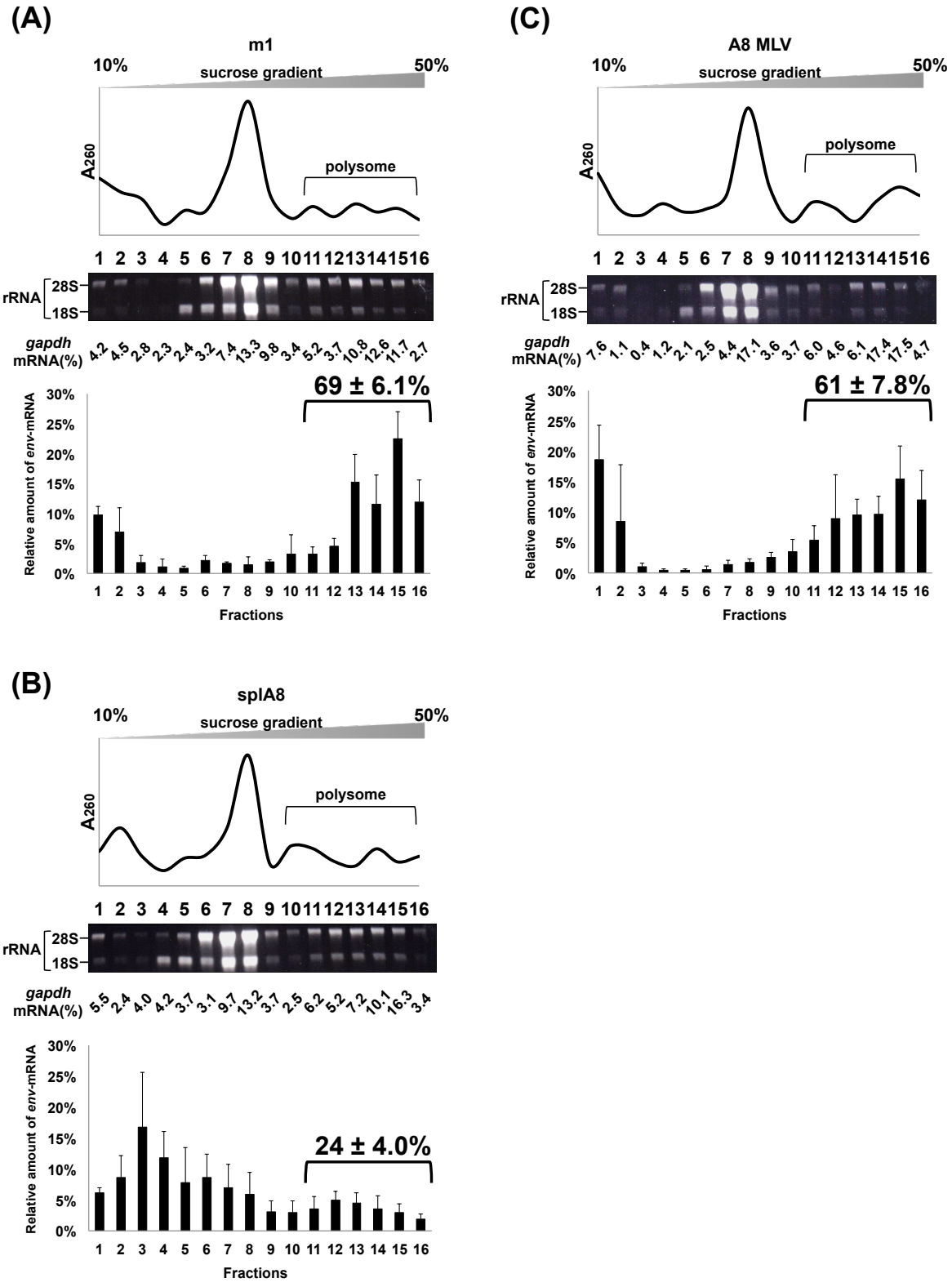

(A)
m1
10%
sucrose gradient
50%
A260
polysome
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
rRNA
28S
18S
gapdh mRNA(%)
4.2 4.5 2.8 2.3 2.4 3.2 7.4 13.3 9.8 3.4 5.2 3.7 10.8 12.6 11.7 2.7
69 ± 6.1%
Relative amount of env-mRNA
30%
25%
20%
15%
10%
5%
0%
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Fractions
(B)
splA8
10%
sucrose gradient
50%
A260
polysome
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
rRNA
28S
18S
gapdh mRNA(%)
5.5 2.4 4.0 4.2 3.7 3.1 9.7 13.2 3.7 2.5 6.2 5.2 7.2 10.1 16.3 3.4
24 ± 4.0%
Relative amount of env-mRNA
30%
25%
20%
15%
10%
5%
0%
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Fractions
(C)
A8 MLV
10%
sucrose gradient
50%
A260
polysome
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
rRNA
28S
18S
gapdh mRNA(%)
7.6 1.1 0.4 1.2 2.1 2.5 4.4 17.1 3.6 3.7 6.0 4.6 6.1 17.4 17.5 4.7
61 ± 7.8%
Relative amount of env-mRNA
30%
25%
20%
15%
10%
5%
0%
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Fractions

Figure 4 Polysome profiles of cells transfected with m1 and splA8, or infected with MLV. Fractionation of ribosomes from cells transfected with (A) m1 and (B) splA8, or infected with (C) MLV. Cell lysates were centrifuged through a 10-50% sucrose density gradient and fractionated, and RNA was extracted from each fraction. Ribosomal RNA was analyzed by electrophoresis on a 1% denaturing agarose gel. To quantitate *gapdh* mRNA, TaqMan Rodents GAPDH Control Reagents containing primer sets and probe were used. The results of the distribution of total RNA, ribosomal RNA, and *gapdh*-mRNA were representative of 3 or 4 experiments. Similar results were obtained in these experiments. The amount of *env*-mRNA was measured by real-time RT-PCR using s1 and s2 primers and ss-probe (Fig. 2A), and the amount in each fraction relative to the total amount of *env*-mRNA in all fractions was calculated. The bottom graphs show the *env*-mRNA distribution from 3 or 4 independent experiments, with the mean ± standard error of each fraction.

cells. In A8-MLV-infected cells (Fig. 4C), 61% of *env*-mRNA was in polysome structures (fractions 11-16), which was not significantly different from the 69% in polysome structures in m1-transfected cells.

Analyses of polysome structure formation of *luc*-mRNA

We analyzed the effect of splicing of *luc*-mRNA on polysome structure formation in cells transfected with m1gpL and splA8L (Fig. 3A). The experiments in Fig. 4 were done using m1gpL- and splA8L- transfected cells to express *luc*-mRNA. The distribution of *luc*-mRNA was examined by real-time RT-PCR using s1 and s2 primers and the ss-probe (Fig. 3A). In m1gpL-transfected cells, 68% of *luc*-mRNA was in polysome fractions 11-16 and, in splA8L-transfected cells, 60% of *luc*-mRNA was in polysome fractions 11-16 (Fig. 5A and B). There was no significant difference in these amounts of *luc*-mRNA in polysome structures. In addition, the amount of unspliced *luc*-mRNA in polysome fractions was higher than that of unspliced *env*-mRNA ($p<0.01$).

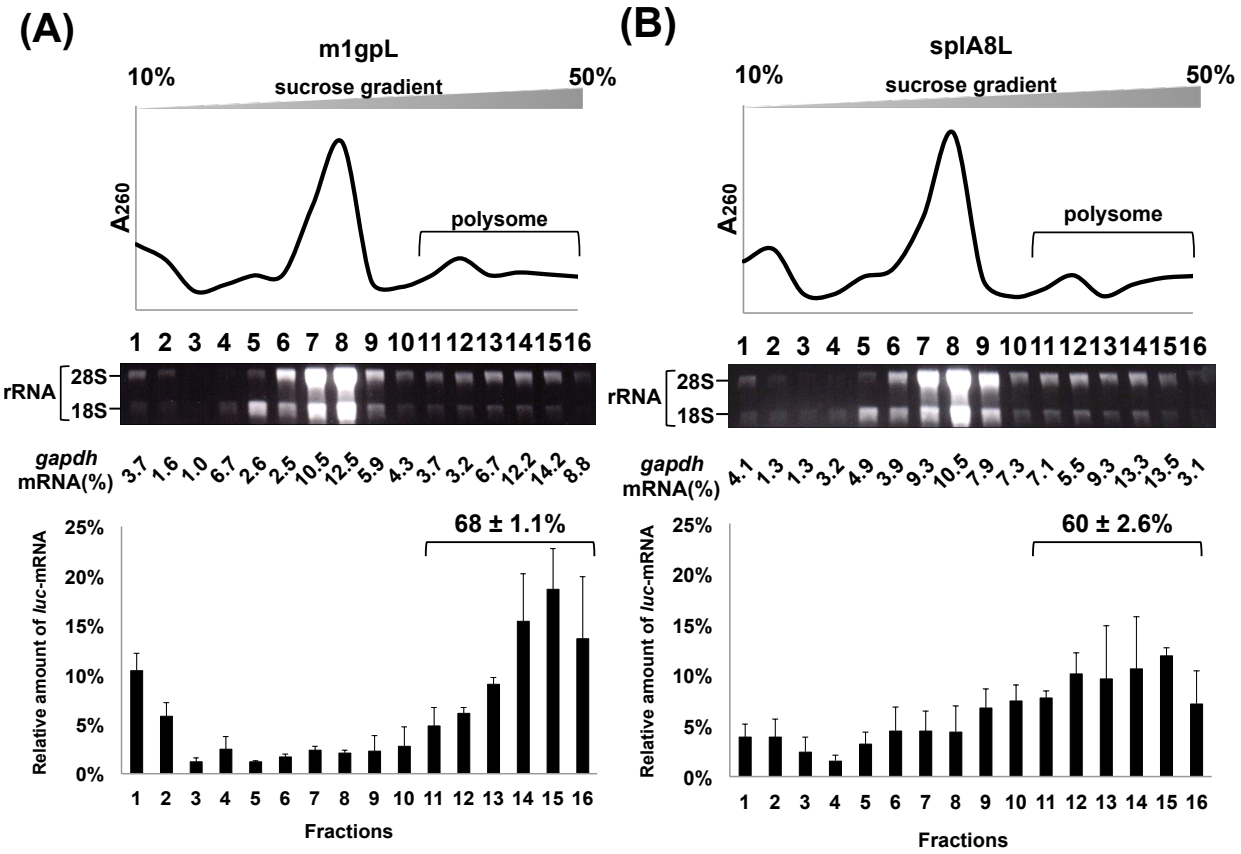


Figure 5 Polysome profiles of cells transfected with m1gpL and splA8L. Fractionation of ribosomes from cells transfected with (A) m1gpL and (B) sp1A8L. The experiments in Fig. 4 were done using m1gpL- and sp1A8L -transfected cells to express *luc*-mRNA. The amount of cDNA from *luc*-mRNA was measured by real-time RT-PCR using s1 and s2 primers and ss-probe. The bottom graphs show the *luc*-mRNA distribution from 3 independent experiments, with the mean ± standard error of each fraction.

DISCUSSION

In complete polyadenylation was detected in a fraction of the unspliced *env*-mRNA products (Fig. 2B and 2E). Sequence analysis of the approximately 70 bp band showed that a fraction of the unspliced *env*-mRNA was cleaved just downstream of the poly(A) addition signal and 2-8 A nucleotides were added (Fig. 2B and 2E). Since the oligo5 primer, which was used for the 2nd PCR, contains 5 T nucleotides, the 1st PCR product with less than 5 A nucleotides cannot be amplified. Thus it is possible that other kinds of abnormal *env*-mRNA, such as *env*-mRNA having 1 or no A nucleotide just downstream of the poly(A) addition signal or *env*-mRNA that was cleaved upstream of the poly(A) addition signal, existed in a fraction of the unspliced *env*-mRNA. Since cleavage of the 3' end of mRNA for poly(A) tail addition usually occurs 15-30 nt downstream of the poly(A) addition signal (Elkon et al., 2013), the 3' ends of the approximately 70 bp bands fraction of unspliced *env*-mRNA were probably abnormally processed unspliced *env*-mRNA. The mechanism(s) by which splicing affects processing of the 3' end of *env*-mRNA is not well understood, but there are known to be functional interactions between spliceosomes and the 3' end-processing machinery of eukaryotic mRNA (Danckwardt et al., 2007). Splicing factors that associate with the 3' terminal intron also interact with downstream polyadenylation factors to promote both 3' end cleavage/polyadenylation and terminal intron splicing (Awasthi and Alwine, 2003; Danckwardt et al., 2007; Kyburz et al., 2006; Li et al., 2001; McCracken et al., 2002; Millevoi et al., 2006). In Rous sarcoma virus and avian sarcoma virus, splicing factors promote 3' end processing through the *cis*-acting negative regulator of splicing (NRS) elements (Miller and Stoltzfus, 1992; Wilusz and Beemon, 2006). Wilusz et al. proposed a model in which heteronuclear ribonucleoprotein (hnRNP) H and serine/arginine-rich (SR) proteins compete to bind to the NRS. In this model, bound SR proteins bridge between the NRS and 3'LTR, promoting recruitment of the 3' end processing machinery. Therefore, splicing factors might assist in processing of the 3' end of *env*-mRNA through interactions with factors involved in 3' end processing. On the other hand, electron microscopic analyses of nascent transcripts from *Drosophila* embryos has shown that splicing occurred co-transcriptionally (Beyer and Osheim, 1988). In the current era, it is generally thought that RNA processing, including splicing, is coupled to transcription (Luco et al., 2011).

Thus, one cannot exclude the possibility that the abnormal *env*-mRNA found in a fraction of the unspliced *env*-mRNA was produced by pre-termination of transcription. When the oligo2 primer was used in the RT reaction to analyze the poly(A) tail length, the size of the smear band (100-180 bp) that was detected in splA8-transfected cells was smaller than that (100-200 bp) in m1-transfected cells (Fig. 2E). It is possible that the amount of mRNAs with a longer poly(A) tail may be larger in spliced *env*-mRNA than in unspliced *env*-mRNA. However, since in this experimental system the length of the mRNA poly(A) tail could be determined for either the unspliced full-length MLV mRNA or the spliced *env*-mRNA from m1-transfected cells, this requires further analyses. The data presented herein, together with previous reports, suggest that splicing of MLV plays an important role in complete polyadenylation of *env*-mRNA, although the mechanisms are still unknown. Interestingly, when the *env* gene in m1 and splA8 was replaced by the *luc* gene, splicing did not affect the 3' end structure of *luc*-mRNA (Fig. 3). These data suggested that there were positive *cis*-elements within the *env* region to complete the polyadenylation of *env*-mRNA by splicing.

Our previous study showed that following introduction of splA8 to cells, the Env expression level was 0.2-fold lower than that for cells with m1. Furthermore, the half-life of *env*-mRNA from splA8 was significantly lower than that for m1 (Yamamoto and Takase-Yoden, 2009). Recently, it was reported that the nuclear export receptor, nuclear RNA export factor 1 (NXF1), was involved in nuclear export of RNA transcripts, especially unspliced mRNA, of gamma retroviruses including xenotropic murine leukemia virus-related virus (XMRV) and MLV (Sakuma et al., 2014). A conserved *cis*-acting element was identified in the *pol* gene of gamma retroviruses, named the cytoplasmic accumulation element (CAE). The 3'ss of XMRV (5638 nt) existed within the CAE (5607-5752 nt), but in Friend-MLV A8, the 3'ss at 5489 nt was actually used and it existed upstream of the CAE (5620-5766 nt) (Machinaga and Takase-Yoden, 2014). Since unspliced *env*-mRNA, which was produced in splA8 transfected-cells, had a complete CAE region, nuclear export of the unspliced *env*-mRNA was unaffected. In fact, our previous study revealed that there was no difference in the distribution of nuclear and cytoplasmic of *env*-mRNA from m1- and splA8-transfected NIH3T3 cells, which produced spliced *env*-mRNA and unspliced *env*-mRNA, respectively (Yamamoto and Takase-Yoden, 2009). In the present study, incomplete

polyadenylation was detected in a fraction of the unspliced *env*-mRNA obtained from splA8-transfected cells (Fig. 2B and 2E). In retroviruses, polyadenylation of the 3' end of mRNA plays an important role in translation initiation and mRNA stability (Karn and Stoltzfus, 2012; Schrom et al., 2013). It is possible that the appearance of incomplete polyadenylation in unspliced *env*-mRNA is partially correlated with a reduction of its stability and translation efficiency.

We also investigated the correlation between splicing and formation of *env*-mRNA polysome structures. The results showed that the fraction of spliced *env*-mRNA in polysome structures was significantly greater than of unspliced *env*-mRNA (Fig. 4). This indicated that splicing promoted the formation of *env*-mRNA polysome structures. To our knowledge, this is the first report showing that splicing of a viral mRNA increased its polysome structure formation. However, it has been reported that mRNAs of the T-lymph cell receptor beta chain and beta globin genes expressed from vectors containing their introns formed more polysome structures than mRNAs expressed from vectors containing truncated introns (Nott et al., 2004). It is known that, during pre-mRNA splicing in the nucleus, several proteins bind to a region 20-24 nucleotides upstream of mRNA exon–exon junctions to form exon junction complexes (EJC) (Le Hir and Andersen, 2008; Le Hir et al., 2000; Reichert et al., 2002; Schmidt et al., 2009; Tange et al., 2004). The EJC is transported with the mature mRNA to the cytoplasm and remains associated with the mRNA-binding proteins until the mRNA is translated, and is involved in various cellular processes, including nucleocytoplasmic mRNA export, subcellular localization, quality control, and translation (Dostie and Dreyfuss, 2002; Giorgi and Moore, 2007; Kashima et al., 2006; Kim et al., 2001; Le Hir et al., 2001; Le Hir and Seraphin, 2008; Lejeune et al., 2002; Tange et al., 2004). It has been proposed that EJC may promote formation of polysome structures, thereby enhancing translation (Diem et al., 2007; Lee et al., 2009; Nott et al., 2004; Wiegand et al., 2003). Diem et al. reported that a 29 kDa protein, PYM, that binds EJC proteins in the cytoplasm also binds, via a separate domain, to the 40S ribosomal subunit and the 48S preinitiation complex, and suggested that PYM functions as a bridge between EJC-associated spliced mRNAs and the translation machinery to enhance mRNA translation (Diem et al., 2007). Therefore, for MLV mRNA, *env*-mRNA that is associated with EJC or that is spliced may form larger polysome structures. When the *env* gene in m1 and splA8 was replaced by the *luc* gene, splicing did not affect polysome structure formation of *luc*-mRNA (Fig. 5) and the degree of

polysome structure formation of unspliced *luc*-mRNA was higher than that of unspliced *env*-mRNA (Fig. 4B and 5B). Thus, it was suggested that there were *cis*-elements within the *env* region that contribute to polysome structure formation of mRNA.

As described above, this study showed that splicing of MLV *env*-mRNA contributed to promote the efficiency of complete polyadenylation of *env*-mRNA and increased the fraction of *env*-mRNA in polysome-structures. Since it is generally known that processing of the 3' ends of mRNA and mRNA polysome structure are factors in the regulation of protein expression or translation efficiency, the promotion of complete polyadenylation and polysome structure formation of *env*-mRNA by splicing might partially explain up-regulation of Env protein expression as a result of splicing. However, there still remains the possibility that other factors such as recruitment of translation factors or ribosome recycling also contribute to promote the translation efficiency of *env*-mRNA by splicing.

Chapter 2

A 38 nt region and its flanking sequences within *gag* of Friend murine leukemia virus are crucial for splicing at the correct 5' and 3' splice sites

MATERIALS AND METHODS

Construction of vectors

To construct the d3 vectors, AatII and SphI were used to carry out a restriction digest of m1. The enzyme sites were blunted using a DNA Blunting Kit (TaKaRa), and blunt-end ligation was then performed. To construct the d3+1026 vector, the *Hind*III-*Bgl*II fragment of m1 was cloned by PCR using a forward primer containing an AatII restriction site in the 5' terminus, (5'-GAGACGTCAGCTTTAGCAGTAGAC-3'), and a reverse primer containing the *Sph*I restriction site in the 5' terminus (5'-TGGCATGCTTTTCAGCTTCCCTCA-3'). The *Art*II-*Sph*I fragment was recombined using the restriction site of m1.

The vectors, B1, B2, B3, B2-a, B2-b, B2-c, B3-d, B3-e, B3-f, B3-d1, B3-d2, B3-d3, and B3-d4, which have serially truncated *Hind*III-*Bgl*II regions of m1, were constructed using a KOD-Plus-Mutagenesis Kit (TOYOBO). Point mutations, G to T (2608 nt), G to T (2614 nt) and G to T (2629 nt), were introduced into the *pol* gene of m1 in order to suppress the production of progeny (Yamamoto and Takase-Yoden, 2009). The mutations A to T (2126 nt) and T to A (2777 nt) were also introduced into m1. To construct the B1 vector, the 879-1182 bp fragment was deleted from m1 by inverse PCR using the primer sets listed in Table 1. The B2, B3, B2-a, B2-b, B2-c, B3-d, B3-e, B3-f, B3-d1, B3-d2, B3-d3, and B3-d4 vectors were constructed by deleting the 1183-1541 bp, 1542-1904 bp, 1183-1296 bp, 1297-1443 bp, 1444-1541 bp, 1542-1649 bp, 1650-1770 bp, 1771-1904 bp, 1542-1569 bp, 1570-1591 bp, 1592-1611 bp, and 1612-1649 bp fragments from m1, respectively, using the same method as for B1. The primers used for generation of these vectors are listed in Table 1. To construct the B3-d4inv38 vector, two-stage inverse PCR was performed using the 2 sets of primers shown in Table 1.

TABLE 1. List of primers used for construction of vectors.

Name of vectore	Forward(5'→3') primer	Reverse(5'→3') primer
B1	cctgactcttccccaatggtatc	gctaaagcttcccaggtcacgat
B2	taatgatgcttttcccttggaacgtcc	gataccattggggaagagtcagg
B3	gtgagggaagctgaaaagatct	atatcattgggcagctgagttggg
B2-a	tactggccattttcctcctctg	gataccattggggaagagtcagg
B2-b	accctgctgacgggagaagaaa	ttgaaactgtccattccctccc
B2-c	taatgatgcttttcccttggaacgtcc	ccctaatagctgttggcagtca
B3-d	gggcagaagccccaccaattt	atatcattgggcagctgagttggg
B3-e	gaggacccagggcaagaaaccaa	gcgctttggagacccgctagga
B3-f	gtgagggaagctgaaaagatct	agggtcataaggagtgtatctgcg
B3-d1	gactgggactacaacacccaac	atatcattgggcagctgagttggg
B3-d2	gaggtaggaaccacctagtcca	gggacgttccaagggaaaagca
B3-d3	cactatcgccagttgctcctag	gttgggtgttgtagtcccagtc
B3-d4	gggcagaagccccaccaattt	gactaggtggttcctacctcgt
B3-d4inv38	1.ccgctatcacgggcagaagccccaccaatt	1.ggtttcgcggactaggtggttcctacc
	2.cctcgttgaccgctatcacgggcagaa	2.atcgcccagaggtttcgcggactaggtggt
57	tagcgggtctccaaaacgcg	ggagcaactggcgatagt
MSV	tagcgggtctccaaaacgcg	tgagcaactggcgatagtgg
FeLV	tattagcgggtctccgcggggcgggcagaag	gcaactggcgataaaggactagg
XMRV	tagcgggtctccaaaacgcg	agagcaactggcggtagagg
CasBrE	tagcgggtctccaaaacgcg	agagcaactggcgatagagg
d3+771	ccacaggtattgggaaccga	gcgctttggagacccgctagga
d3+293	cactatcgccagttgctcctag	acgtctcccagggttgcggc
d3+467	cctgactcttccccaatggtatc	acgtctcccagggttgcggc
d3+353	tactggccattttcctcctctg	acgtctcccagggttgcggc
d3+206	accctgctgacgggagaagaaa	acgtctcccagggttgcggc
d3+108	taatgatgcttttcccttggaacgtcc	acgtctcccagggttgcggc
d3+80	gactgggactacaacacccaac	acgtctcccagggttgcggc
d3+58	gaggtaggaaccacctagtcca	acgtctcccagggttgcggc

To construct 57 MLV, Moloney murine sarcoma virus (MSV), Feline leukemia virus (FeLV), xenotropic murine leukemia virus-related virus (XMRV), and CasBrE vectors, point mutations were inserted into d3+1026 by inverse PCR using the primer pairs shown in Table 1.

To construct the d3+771 vector, the 1650-1904 bp fragment of d3+1026 was deleted by inverse PCR using the primer set shown in Table 1. The d3+293, d3+467, d3+353, d3+206, d3+108, d3+80, and d3+58 vectors were constructed by deleting the 879-1611 bp, 879-1182 bp, 879-1296 bp, 879-1443 bp, 879-1541 bp, 879-1569 bp, and 879-1591 bp fragments from d3+1026, respectively, using the same method as for d3+771. The primer pairs listed in Table 1 were used to construct the vectors. To construct d3+38, a 38 bp fragment with an *Aat*II restriction site in the 5' terminus and a *Sph*I restriction site in the 3' terminus was prepared by annealing the following synthetic single-stranded DNAs: (5'-CCACTATCGCCAGTTGCTCCTAGCGGGTCTCCAAAGCGCGCATG-3') and (5'-CGCGCTTTGGAGACCCGCTAGGAGCAACTGGCGATAGTGGACGT-3'). To construct d3+inv38 vectors, a fragment with the inverse sequence of the 38 bp fragment and with an *Aat*II restriction site in the 5' terminus and a *Sph*I restriction site in the 3' terminus was prepared by annealing the following synthetic single-stranded DNAs: (5'-CCGCGAAACCTCTGGGCGATCCTCGTTGACCGCTATCACGCATG-3'), and (5'-CGTGATAGCGGTCAACGAGGATCGCCCAGAGGTTTCGCGGACGT-3'). These fragments were inserted into the *Aat*II and *Sph*I restriction sites of m1.

To construct the d4+58 and d4+1026 vectors, *Sph*I and *Nde*I were used to carry out a restriction digest of d3+58 and d3+1026, respectively. The enzyme sites were blunted using a DNA Blunting Kit (TaKaRa), and blunt-end ligation was then performed.

Cell cultures and transfections

NIH3T3 cells were grown in Dulbecco's Modified Eagle Medium (DMEM; Cellgro) supplemented with 10% (v/v) fetal calf serum (FCS; MP Biomedicals), 50 units penicillin (GIBCO)/ml, and 50 μg streptomycin (GIBCO), at 37°C in a 7% CO_2 atmosphere. Cells were seeded with 1×10^6 cells in a 6 cm dish with growth medium minus penicillin-streptomycin. They were transfected the next day with 8 μg of vectors using Lipofectamine 2000 Reagent (Invitrogen)

diluted by OPTI-MEM (Invitrogen) according to the manufacturer's instructions.

RT-PCR analysis of viral spliced mRNA in transcripts

Total cellular RNA was isolated from transfected cells using RNeasy Mini Kit (QIAGEN) according to the manufacturer's instructions. After treatment with RNase-free DNase (QIAGEN), 2 μg of RNA were added to the reverse transcription (RT) reaction, which used an oligo (dT) primer (Invitrogen). The spliced mRNA was detected by PCR using Go Taq (Promega) and specific primers. The primers for detecting spliced mRNA containing the splicing junction region were: s1 forward primer (5'-GAGACCCTTGCCCAGGGA-3') and s2 reverse primer (5'-TGCCGCCAACGGTCTCC-3'). The *env* coding region was detected by PCR, using KOD-plus (TOYOBO) and the s1 forward primer and the u3 reverse primer (5'-TGCGGCTATCAGGCTAAGCAACTTGGT-3'). PCR products were separated on 2% or 1% agarose gels in TBE or TAE buffer, respectively, and stained with ethidium bromide. Negative control samples without the cDNA synthesis step did not yield specific bands.

Sequence analysis of splice variants

To analyze the sequence of splice variants, the relevant electrophoretic band was extracted from the agarose gel using a QIAquick Gel Extraction Kit (QIAGEN), and cloned into a pGEM-T-easy vector (pGEM-T-easy Vector System; Promega). The sequences of the cloned fragment in T-easy vector were amplified using T7 (5'-GTAATACGACTCACTATAGGGC-3'), or sp6 (5'-ATTTAGGTGACACTATAGAA-3') primers, and a BigDye Terminator v 3.1 Cycle Sequencing Kit (Applied Biosystems). The sequences were analyzed using an ABI PRISMOR 3100 Genetic Analyzer (Applied Biosystems).

RESULTS

Narrowing down the region within the *Hind*III-*Bgl*II fragment (879-1904 bp) that is crucial for splicing at the correct 5'ss and 3'ss

During splicing of the Friend-MLV gene, the 205-5488 nt region is normally spliced out and, subsequently, *env* mRNA is produced (Fig. 6). We previously investigated the role of the intron within the Friend-MLV gene in Env expression using vectors with serially truncated introns. We found that when the *Hind*III-*Bgl*II (879-1904 bp) fragment was deleted from the m1 vector, which contains the full-length A8-Friend-MLV proviral gene sequence, abundant splice-variants were observed among the transcripts (Yamamoto and Takase-Yoden, 2009). The m1 vector was transfected into NIH3T3 cells and generation of spliced mRNA was examined after 48 h by RT-PCR using the s1 and s2 primers (Fig. 6 and 7A). These primers were designed to amplify a 94 bp fragment containing the splicing junction region from the cDNA of normal spliced transcripts. As shown in Figure 7B, the 94 bp band was detected in the transcripts of m1. In contrast, in the transcripts of the B vector (called the pA8d1b vector in our previous report) with a truncated *Hind*III-*Bgl*II fragment from m1, a band of approximately 300 bp was detected (Fig. 7B). Sequence analysis showed that the 300 bp band came from an mRNA splice variant in which the 205-807 nt (D1-A3) region and the 2042-5488 nt (D6-A13) regions were spliced out (mRNA-D1-A3/D6-A13). These results are in agreement with those of our previous study (Yamamoto and Takase-Yoden, 2009). We analyzed the importance of the *Hind*III-*Bgl*II fragment in splicing using the d3 vector, in which the majority of the intronic *Aat*II-*Sph*I (366-5139 bp) fragment was deleted from m1 (Fig. 7A). The d3 vector yielded the splice variant mRNA-D1-A1/D12-A13 and unspliced transcripts (Fig. 7B and 7C). When the *Hind*III-*Bgl*II fragment was inserted between 366 bp and 5139 bp of the d3 vector (to produce the d3+1026 vector) only produced correctly spliced mRNA (Fig. 7B).

We first sought to narrow down the region within the *Hind*III-*Bgl*II fragment that is crucial for splicing at the correct 5'ss and 3'ss of Friend-MLV. To this end, we constructed B1, B2, and B3 vectors with serially truncated 1 kbp (*Hind*III-*Bgl*II) fragments from the entire sequence of m1 (Fig. 7A). Among these vectors, B2 and B3 yielded a small amount of mRNA splice variants in addition to correctly spliced mRNA (Fig. 7A and 7B). Details of the splice variants B2 and B3 are described

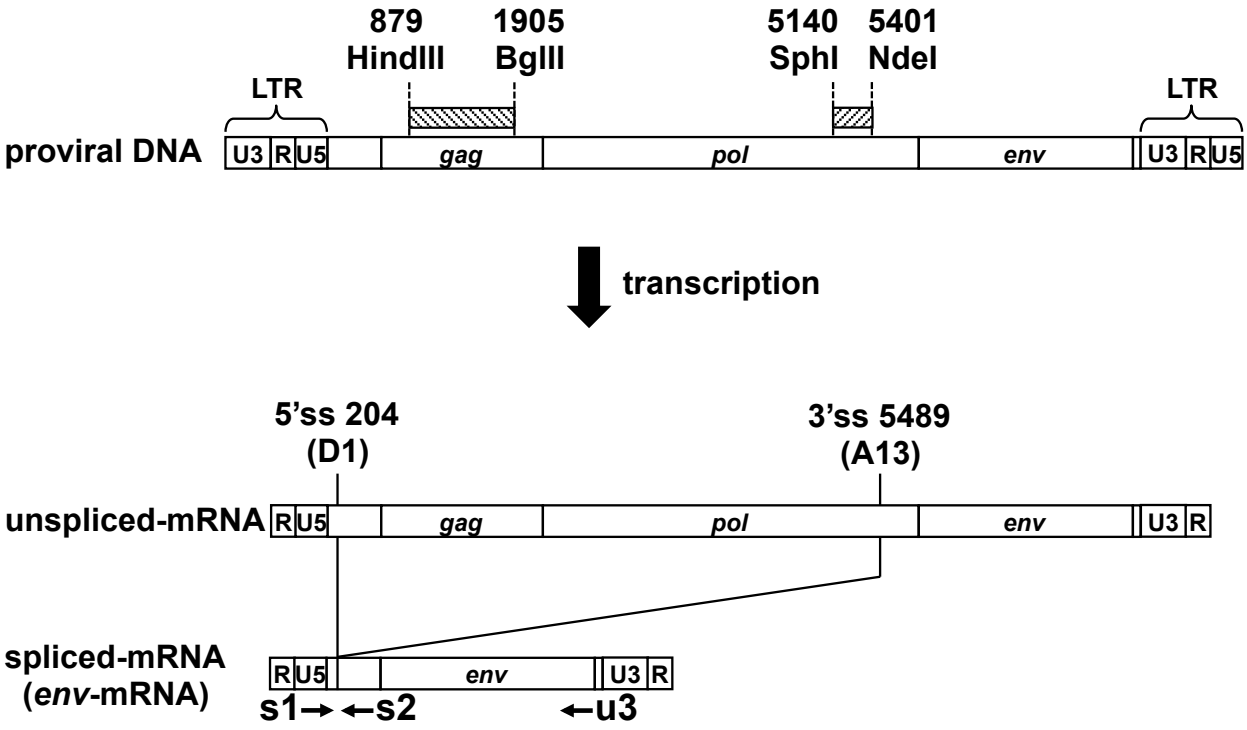
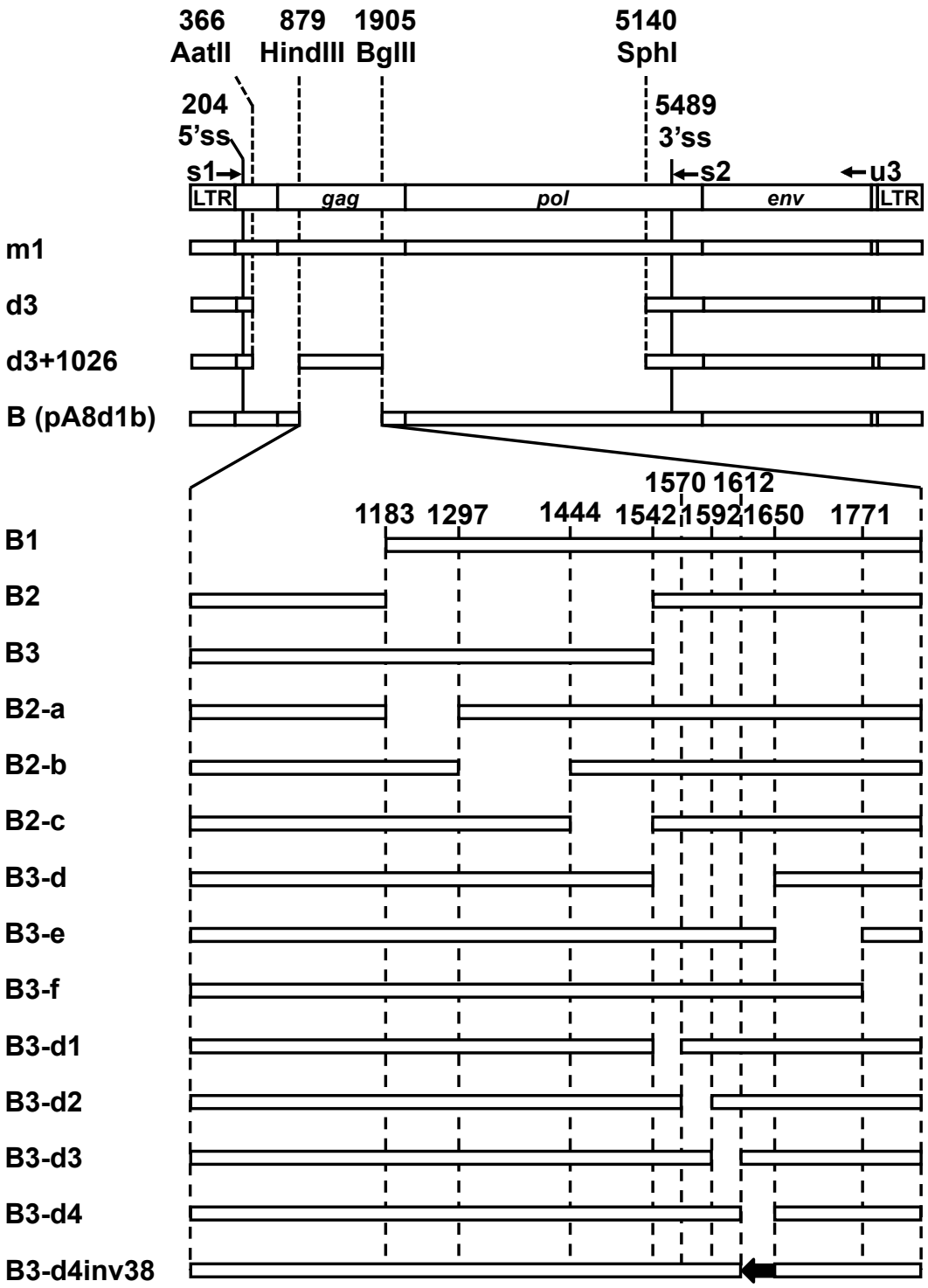


Figure 6 Structure of proviral DNA of MLV and the primers used to detect *env*−mRNA by RT−PCR. Throughout these figures, native 50ss (204nt) and 30ss (5489 nt) are designated D1 and A13, respectively (see splice site numbers in Fig.11a). The numbering of nucleotides is based on the transcript. The s1 and s2 primers were designed to amplify a 94 bp fragment from the cDNA of spliced transcripts containing the splice junction region and the s1 and u3 primers to amplify a 2.4 kbp fragment from cDNA of spliced *env*−mRNA.

(A)


366
AatII
879
HindIII
1905
BglII
5140
SphI
204
5'ss
5489
3'ss
s1
s2
u3
LTR
gag
pol
env
LTR
m1
d3
d3+1026
B (pA8d1b)
1570
1612
1183
1297
1444
1542
1592
1650
1771
B1
B2
B3
B2-a
B2-b
B2-c
B3-d
B3-e
B3-f
B3-d1
B3-d2
B3-d3
B3-d4
B3-d4inv38

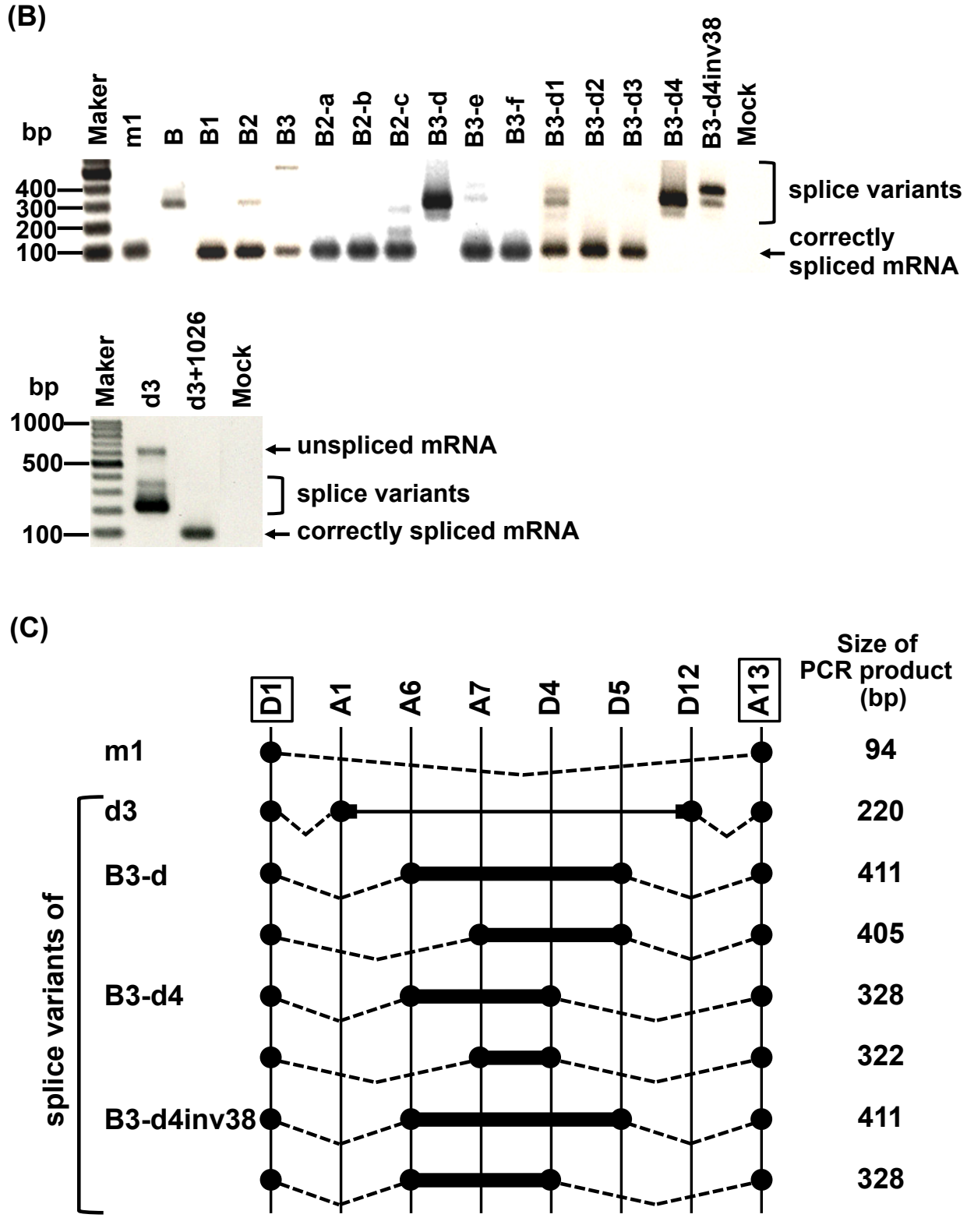

(B)
bp
Maker
m1
B
B1
B2
B3
B2-a
B2-b
B2-c
B3-d
B3-e
B3-f
B3-d1
B3-d2
B3-d3
B3-d4
B3-d4inv38
Mock
400
300
200
100
splice variants
correctly
spliced mRNA
bp
Maker
d3
d3+1026
Mock
1000
500
100
unspliced mRNA
splice variants
correctly spliced mRNA
(C)
Size of
PCR product
(bp)
D1
A1
A6
A7
D4
D5
D12
A13
m1
d3
B3-d
B3-d4
B3-d4inv38
splice variants of
94
220
411
405
328
322
411
328

Figure 7 Analysis of the crucial for splicing at the correct 5'ss and 3'ss of Fr-MLV. (A) Structures of d3, d3+1026, and vectors with a serially truncated *Hind*III-*Bgl*II fragment of m1. (B) Detection of the splice junction region of *env* mRNA in the transcripts of vectors. (C) Structures of splice variants in transcripts of d3, B3-d, B3-d4, and B3-d4inv38 and sizes of the PCR products. In (B), the m1 vector and the vectors with the serially truncated *Hind*III-*Bgl*II fragment were designed to generate both unspliced mRNA and spliced mRNA; however, bands corresponding to unspliced mRNA were not detected in the transcripts of all tested vectors under the conditions used in the present experiments. Sequence analysis confirmed that the 94 bp band came from normally spliced transcripts. In (C), closed circles represent the splice sites that were utilized in splice variants (see splice site numbers in Figure 11A), thick solid-lines represent the exon region in splice variants, and broken lines represent regions that were spliced out. The splice variants not included in (C) are as follows: B2, mRNA-D1-A4/D2-A13 that was produced by the splicing out of 205-1006 nt (D1-A4) and 1201-5488 nt (D2-A13) region; B3, mRNA-D1-A6/D6-A13; B2-c, mRNA-D1-A6/1444nt (artificial splice site that arose due to vector construction)-A13; B3-e, mRNA-D1-A6 or A7/D4-A13; B3-d1, mRNA-D1-A6, A7, or A8/D4-A13.

in the legend to Figure 7. We sought to further narrow down the 0.7 kb (1183-1904 bp) region by constructing the B2-a to B3-f vectors that are serially truncated in this region. The transcripts of the B2-a, B2-b, and B2-f included only correctly spliced mRNA (Fig. 7B). The transcripts of the B2-c and B3-e vectors included abundant correctly spliced mRNA and a few mRNA splice variants (see the legend to Fig. 7 for details). It is noteworthy that B3-d included only mRNA splice variants as described in Figure 7C. To further narrow down the 0.1 kb (1542-1649 bp) fragment, we next constructed the B3-d1 to B3-d4 vectors that were serially truncated in the fragment. In B3-d1, the transcripts contained normally spliced mRNA and a few mRNA splice variants (see the legend to Fig. 7 for details). In B3-d2 and B3-d3, the transcripts contained only correctly spliced mRNA and no splice variants were detected. Interestingly, B3-d4, in which the 1612-1649 bp fragment was deleted from m1, only yielded mRNA splice variants. The transcripts from the B3-d4inv38 vector carrying a reverse sequence of the 1612-1649 bp fragment included only splice variants. These findings showed that the 38 nt region (1612-1649 nt) contained the important elements that regulate splicing at the correct 5'ss and 3'ss.

We examined whether the spliced mRNA of the tested vectors contained the entire sequence of the *env* coding region by RT-PCR using the s1 and u3 primers (Fig. 6). These primers were designed to amplify a 2.4 kbp band in the spliced mRNA of m1. We found that the spliced mRNAs of all of the tested vectors contained the entire sequence of the *env* coding region (data not shown).

Comparison of the 38 nt fragment among gamma retroviruses

The 38 nt fragment sequence is well conserved among simple retroviruses such as gamma retroviruses, although a few nucleotides differ from A8-MLV (Fig. 8A). To determine whether the differences in sequences of the 38 nt fragment in gamma retroviruses influenced splicing, we replaced the 38 bp fragment of the d3+1026 vector with the 38 bp fragment from Friend-MLV 57, Moloney murine sarcoma virus (MSV), feline leukemia virus (FeLV), xenotropic MLV-related virus (XMRV), or CasBrE MLV (Fig. 8A). The derived vectors were transfected into NIH3T3 cells and their mRNA transcripts were analyzed by RT-PCR using the s1 and s2 primers. In the transcripts of A8, 57, MSV, XMRV, and CasBrE, only the 94 bp band of the correctly spliced transcript was observed (Fig. 8B). In FeLV, a few mRNA splice variants were detected in addition

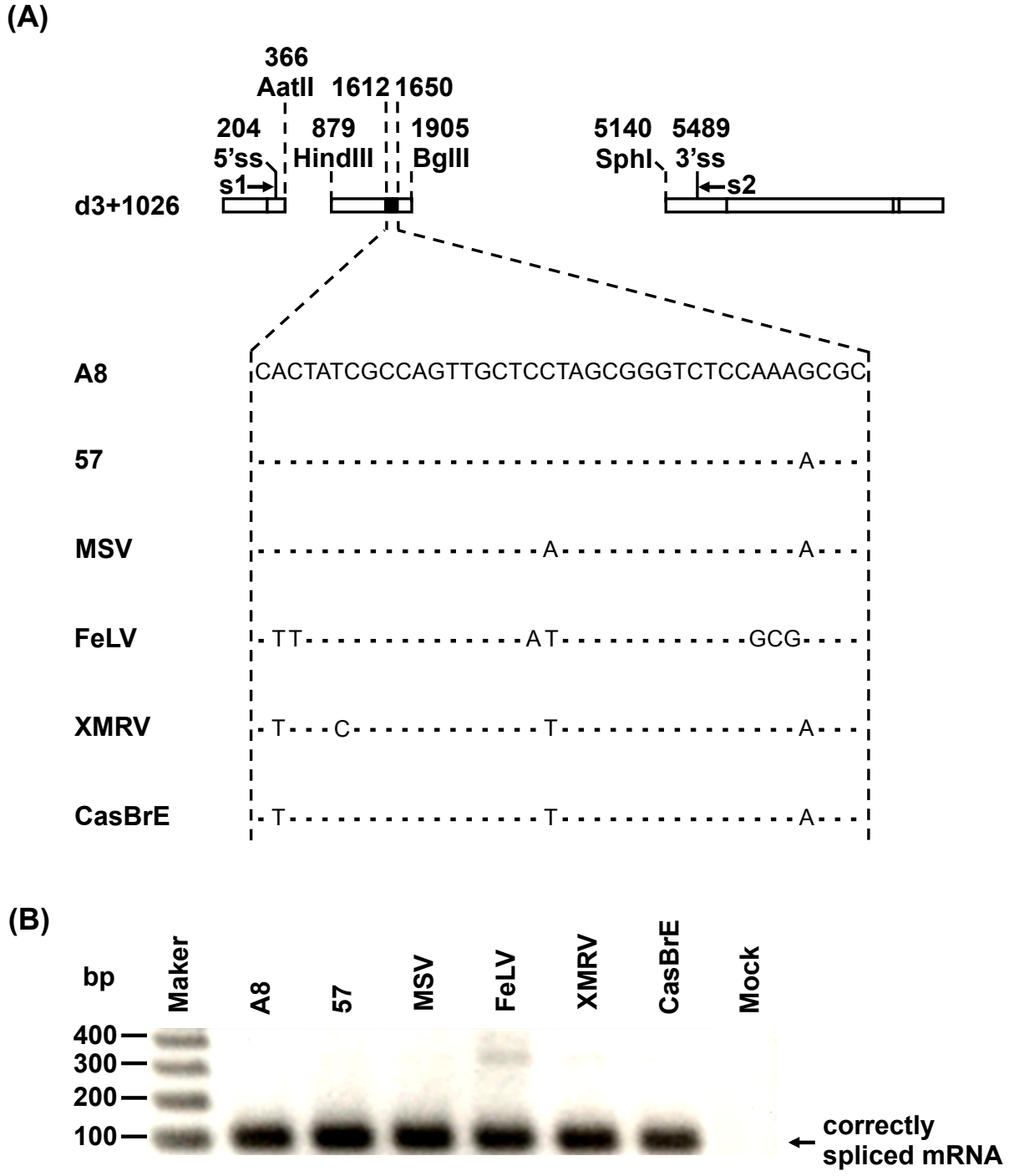


Figure 8 Comparison of the 38 nt fragment among gamma retroviruses. (A) Structures of vectors in which the 38 bp fragment of d3+1026 was replaced by the 38 bp fragment derived from Fr-MLV 57, MSV, FeLV, XMRV, or CasBrE. (B) Detection of the splice junction region of *env* mRNA in the transcripts of vectors.

to correctly spliced mRNA.

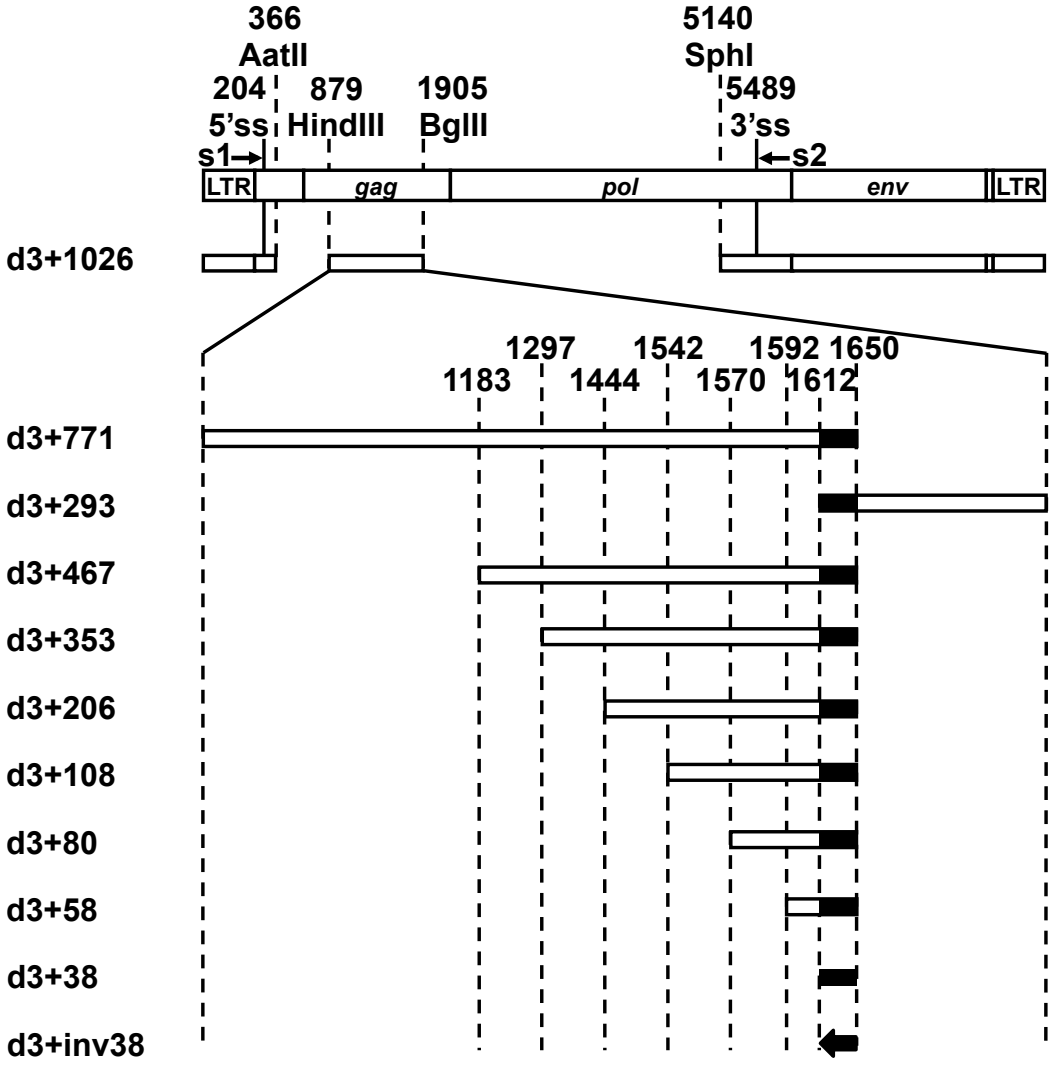
Effects of the flanking sequence of the 38 nt fragment on splicing at the correct 5' and 3' splice sites

As shown in Figure 7, the 38 nt region was demonstrated to play an important role in splicing at the correct splice sites. However, the flanking sequences of the 38 nt region, such as the 1183-1541 nt region, the 1542-1569 nt region, and the 1650-1770 nt region, also positively influenced splicing (B2, B3-d1, and B3-e vectors in Fig. 7), as described in the first section of the Results and the first paragraph of the Discussion. In order to further define the region for control of correct splicing within the flanking sequences of the 38 nt fragment, we constructed vectors with serially truncated *Hind*III-*Bgl*II fragments derived from the d3+1026 vector (Fig. 9A). The d3+771 and d3+293 vectors have the upstream and downstream flanking regions, respectively, of the 38 bp fragment of d3+1026. In d3+771, correctly spliced mRNA was detected (Fig. 9B). By contrast, in d3+293, variant transcripts (data not shown) were detected in addition to correctly spliced mRNA. Next, we constructed a series of vectors, d3+467 to d3+38, in which the 879-1611 bp region of d3+771 was serially truncated, to narrow down the critical flanking region within the 879-1611 nt fragment for correct splicing (Fig. 9A). Serial truncation of the upstream region of the 38 bp fragment resulted in the production of transcripts with splicing variants in d3+353, d3+206, d3+80, and d3+38 in addition to correctly spliced mRNA (Fig. 9B). No correctly spliced transcripts were present in the spliced mRNA products of d3+inv38 carrying a reverse sequence of the 38 bp fragment; only correctly spliced mRNA was detected in the transcripts of d3+467, d3+108, and d3+58.

Effect of the upstream region of 3'ss of MLV on function of the 38 nt fragment and its flanking sequence

In a previous study, we showed that the *Sph*I-*Nde*I (5140-5400 nt) fragment located approximately 100 nt upstream of the 3'ss could influence splicing efficiency and the appearance of splice variants (Yamamoto and Takase-Yoden, 2009) (Fig. 6). When this *Sph*I-*Nde*I fragment was deleted from m1, a splice variant appeared in which the 205-1360 nt (D1-A6) and 1595-5488 nt

(A)


366
AatII
5140
SphI
204
5'ss
879
HindIII
1905
BglII
5489
3'ss
s1
s2
LTR
gag
pol
env
LTR
d3+1026
1297
1542
1592
1650
1183
1444
1570
1612
d3+771
d3+293
d3+467
d3+353
d3+206
d3+108
d3+80
d3+58
d3+38
d3+inv38

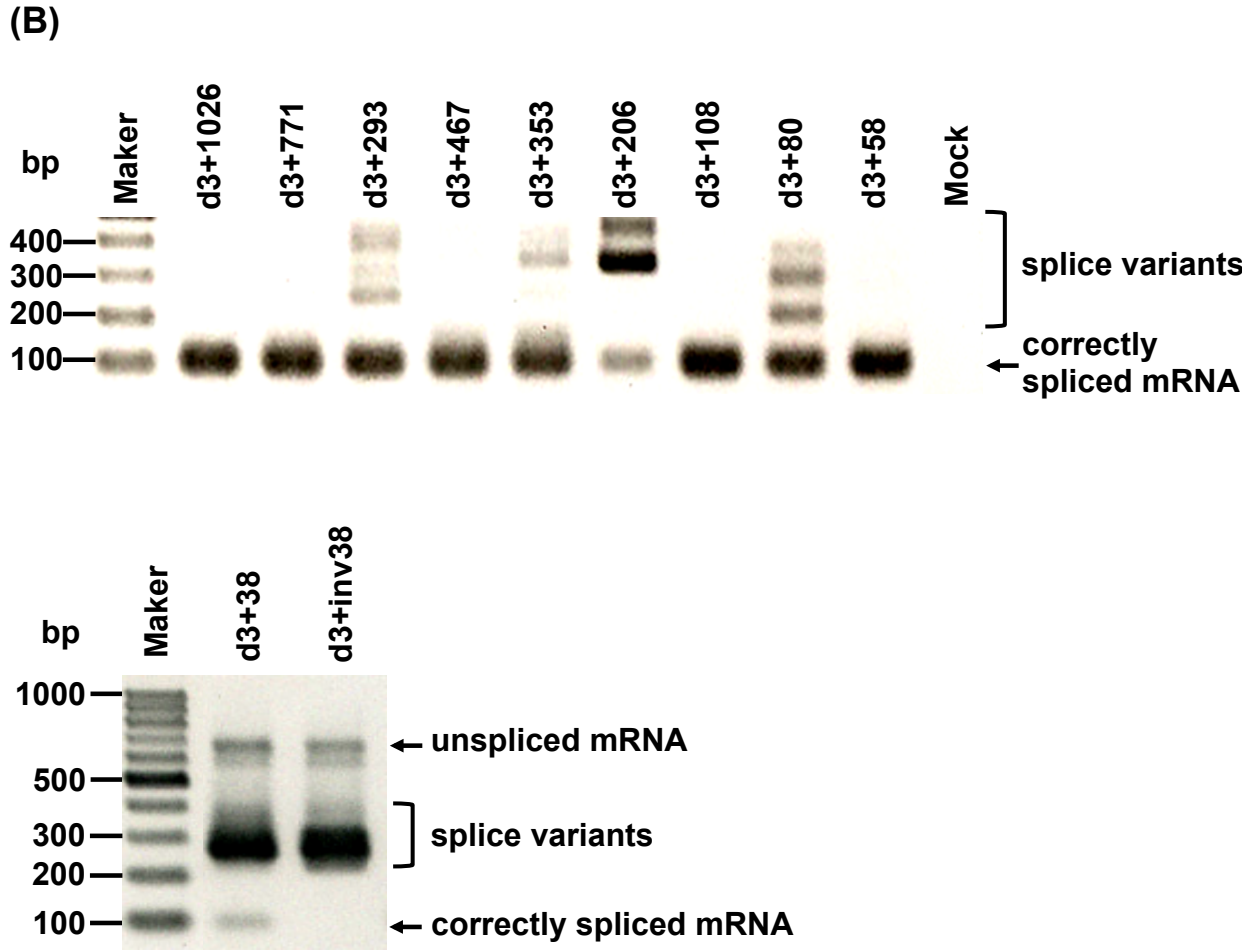


Figure 9 Effects of the flanking sequence of the 38 nt fragment on splicing at the correct 5' and 3' splice sites. (A) Structures of vectors with serial truncation of the *Hind*III-*Bgl*II fragment of d3+1026. (B) Detection of the splice junction region of *env* mRNA in the transcripts of the vectors. In (A), solid lines represent the 38 bp fragment.

(D4-A13) regions were spliced out (data not shown). Interestingly, the structure of the splice variant was identical to that of a splice variant observed in the transcripts of B3-d4, in which the 38 bp fragment was deleted from m1 (Fig. 7). The vectors developed for the experiments shown in Figure 9 contain the *Sph*I-*Nde*I region. To examine whether this region influenced the control of splicing by the 38 nt fragment and its flanking sequence, as shown in Figure 10A, we deleted it from d3+58 and d3+1026 (Fig. 9); the resulting transcripts were all correctly spliced. By contrast, abundant splice variants were obtained from the transcripts of d4+58 and d4+1026 (Fig. 10B and 10C). The findings revealed that a synergistic interaction between the 5140-5400 nt region located upstream region of the 3'ss and the 38 nt fragment and its flanking sequence is required for splicing at the correct 5'ss and 3'ss.

Mapping of cryptic 5'ss and cryptic 3'ss on the MLV gene

As described in earlier sections, our experiments with vectors carrying truncated fragments from m1 yielded a range of splice variants. On the basis of the sequence data of these splice variants, we mapped the splice sites that were actually used as 'D' (5'ss) and 'A' (3'ss). The native 5'ss (204 nt) and 3'ss (5489 nt), designated D1 and A13, respectively, are indicated by the boxed regions in Figure 11A. Cryptic 5'ss were identified at D2 (1200 nt), D3 (1366 nt), D4 (1594 nt), D5 (1677 nt), D6 (2041 nt), and D12 (1200 nt); cryptic 3'ss were found at A1 (285 nt), A3 (808 nt), A4 (1007 nt), A6 (1361 nt), A7 (1367 nt), A8 (1411 nt), and A9 (1435 nt). A few sites, D4, D5, A6, and A7, were frequently used to produce splice site variants in the tested vectors. The sequences of these splice sites are shown in Figure 11B. All of the fragments that were spliced out from the tested vectors contained GU at the 5'ss and AG at the 3'ss, and had a polypyrimidine-tract-like sequence located before the 3' end of the fragment to be spliced. In addition, we found consensus sequences, GAGGUAAGC and $Y_{(11)}$AAG, where Y=T or C, that could act as 5'ss and 3'ss, respectively ($D7^{\#}$-$11^{\#}$, $D13^{\#}$-$17^{\#}$, $A2^{\#}$, $A5^{\#}$, $A10^{\#}$-$12^{\#}$, and $A14^{\#}$ in Fig. 11A). However, these predicted sites were not used by all of our tested vectors. The consensus sequences of the branch points for splicing, YNYURAY, where Y=T or C, N=A, T, G or C, and R=A or G, were also mapped (black ellipses in Fig. 11A). Predicted branch-points were located in the upstream region of each cryptic 3'ss.

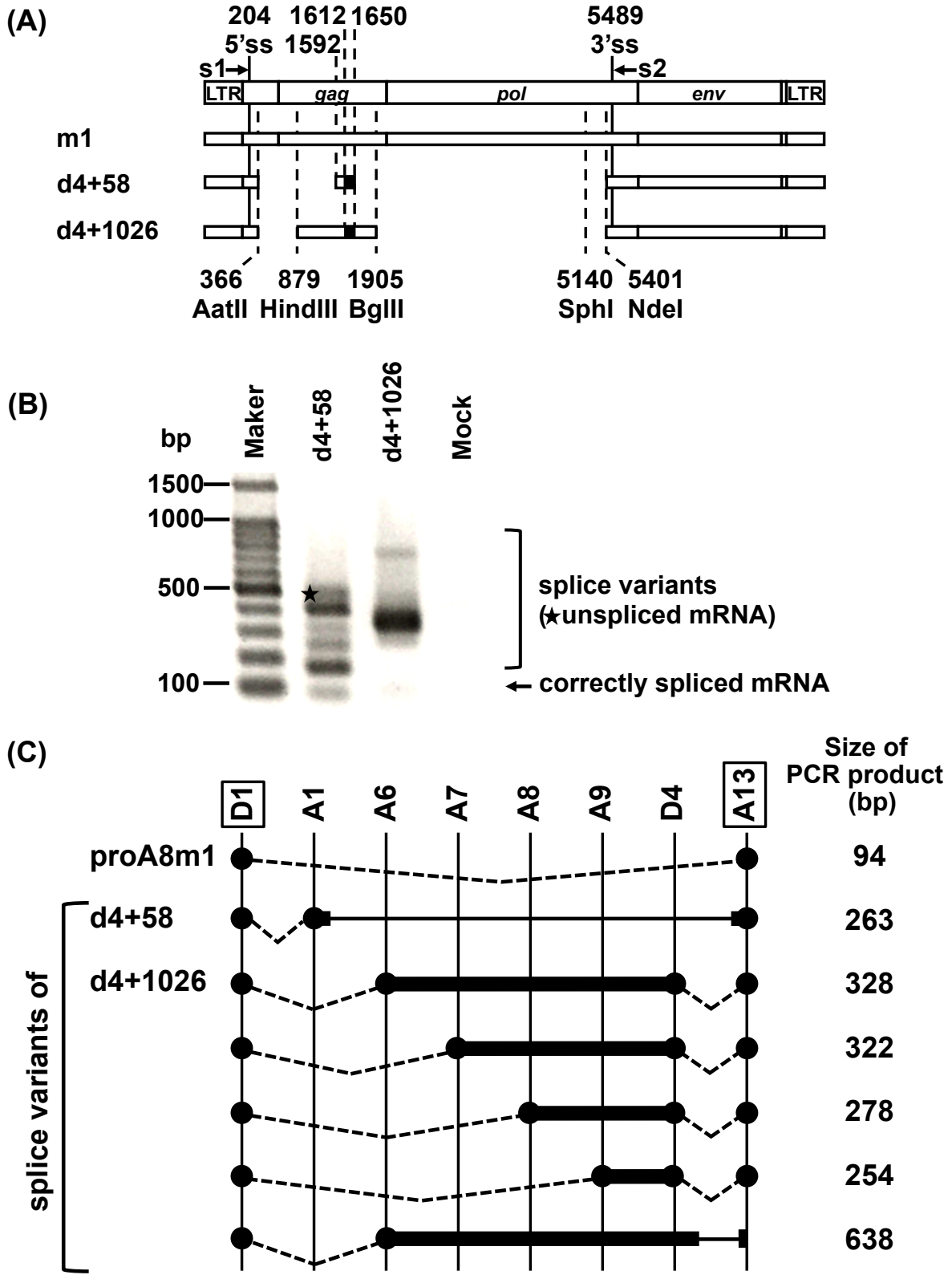
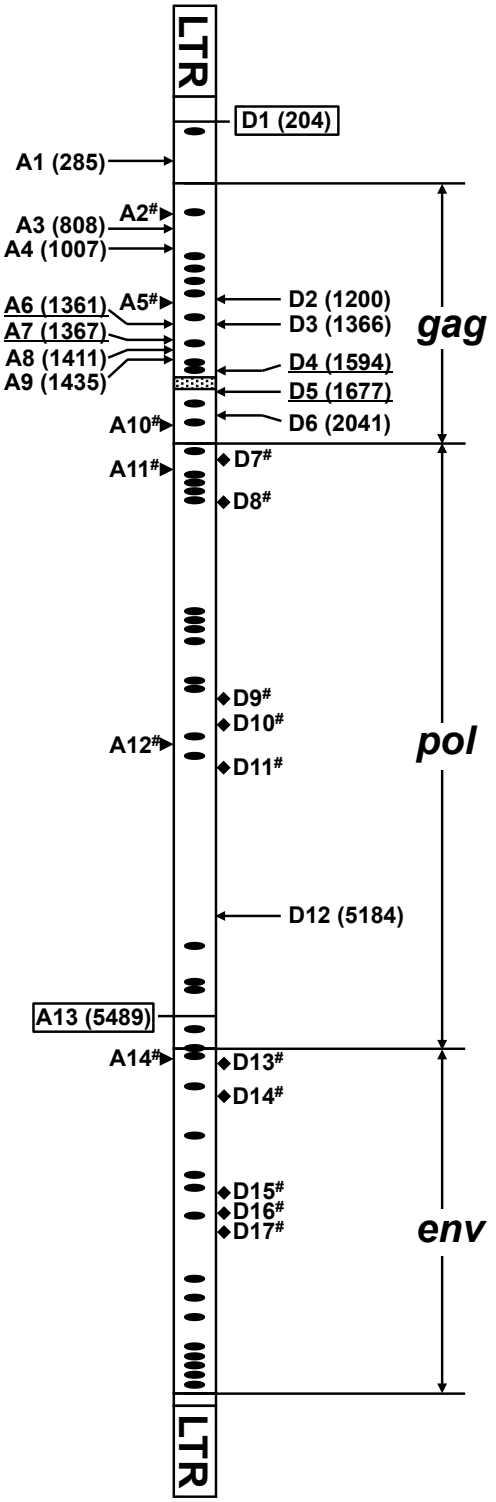

(A)
204
1612
1650
5489
5'ss
1592
3'ss
s1
s2
LTR
gag
pol
env
LTR
m1
d4+58
d4+1026
366
879
1905
5140
5401
AatII
HindIII
BglII
SphI
NdeI
(B)
bp
Maker
d4+58
d4+1026
Mock
1500
1000
500
100
splice variants
(★unspliced mRNA)
correctly spliced mRNA
(C)
Size of
PCR product
(bp)
D1
A1
A6
A7
A8
A9
D4
A13
proA8m1
d4+58
d4+1026
splice variants of
94
263
328
322
278
254
638

Figure 10 Analysis of the *Sph*I-*Nde*I region influenced the control of splicing by the 38 nt fragment and its flanking sequence. (A) Structures of vectors produced by deletion of the *SphI-Nde*I fragment from d3+58, and d3+1026 (see Figures 4). (B) Detection of the splice junction region of *env* mRNA in the transcripts of the vectors. (C) Structures of splice variants in transcripts of d4+58, and d4+1026, and sizes of the PCR products. In (C), closed circles represent the splice sites that were utilized in splice variants (see splice site numbers in Figure 14A), thick solid lines represent the exonic region in splice variants, broken lines represent regions that were spliced out, and thin solid lines represent the regions that were not contained in the vectors.

(A)



(B)

	Sequences
D1	(GAG / GUAAGC)
D2	(AUG / GUAUCC)
D3	(CAG / GUAAAU)
D4	(GAG / GUAGGA)
D5	(AAG / GUAAAA)
D6	(UGA / GUAAGU)
D12	(AAG / GUAAGU)

	Sequences
A13	(CCCCUCUCUCCAAG / CUCA) polypyrimidine tract
A1	(UGUGUCUGUACUAG / UUGG)
A3	(ACAUUAUUACACAG / GUUA)
A4	(UCUUCCCCCUUCAG / CCCC)
A6	(CCUCUUUCUCCGAG / GACC)
A7	(CUCCGAGGACCCAG / GUAA)
A8	(UCCUUACUCAUCAG / CCCA)
A9	(AUGACUGCCAACAG / CUAU)

Figure 11 Mapping of cryptic 5'ss and cryptic 3'ss on the MLV gene. (A) Mapping of the splice sites that were actually used in splice variants, the predicted splice sites, and the predicted branch points for splicing. Numbers in parentheses represent the positions of splice sites. The consensus sequences for 50ss and 30ss (GAGGUAAGC and Y(11)AAG: Y.T or C, respectively) were also mapped and indicated by "#." Black ellipses represent the consensus sequences of branch points for splicing. Splice sites that were frequently used in splice variants are underlined. The shadowed region on the MLV gene represents the 38 nt fragment. (B) Sequences of the splice sites of splice variants. The splice sites that were used in splice variants were mapped and designated "D" (5'ss) and "A" (3'ss). Artificial splice sites that arose as a result of vector construction are not included. Native 5'ss (D1) and 3'ss (A13) are indicated by the

DISCUSSION

The genome of Friend-MLV contains a 5'ss located at 205 nt and a 3'ss located at 5489 nt. In virus-infected cells, *env* mRNA is produced by a single splicing event. In previous studies, we have shown that when the *Hind*III-*Bgl*II (879-1905 bp) fragment within the *gag* gene was deleted from the m1 vector that has the entire Friend-MLV sequence, then cryptic splicing of mRNA occurred (Yamamoto and Takase-Yoden, 2009). In the present study, we first attempted to narrow down the region within the *Hind*III-*Bgl*II fragment that is crucial for splicing at the correct 5'ss and 3'ss of Friend-MLV. The results indicated that a 38 nt fragment (1612-1649 nt) in *gag* contained the important elements that regulate splicing at the correct 5'ss and 3'ss, and that the orientation of the 38 nt fragment was crucial to its function (Fig. 7A and 7B). However, elements that positively contribute to splicing at the correct splice sites seem to exist in the *Hind*III-*Bgl*II region except for the 38 nt fragment. In the transcripts of B2, which contained the 38 nt fragment, some mRNA splice variants were observed. There were no splice variants among transcripts of B2-a or B2-b; however, splice variants were present among the transcripts of B2-c. The latter had an artificial splice site at 1444 nt that arose due to vector construction and which was used to produce the variants. Therefore, although we could not further delimit the 1183-1541 nt region with regard to sequences essential to splicing, our results indicated that the entire 1183-1541 nt region might play a positive role in splicing at the correct splice sites. We believe that the presence of normally spliced mRNA and a few splice variants in the transcripts from B3, which did not carry the 38 nt fragment, might be due to the presence of the 1183-1541 nt region in this vector. Similarly, as the transcripts of B3-e and B3-d1 contained a few mRNA splice variants, we suggest that the 1650-1770 nt region and the 1542-1569 nt region positively contributed to splicing at the correct splice sites. The vectors tested in this study were deletion mutants created within Gag protein coding region, which is known to be a relatively conserved region in gamma retroviruses. Therefore, we cannot exclude the possibility that other regions in addition to the 38 nt fragment contribute to splicing through changes in pre-mRNA features such as exon length and pre-mRNA secondary structures.

The effects of the 38 nt fragment on splicing are not restricted to NIH3T3 cells but also occur in F10, RS-A, 293T, U927, HeLa, and Jurkat cells (data not shown). When the *env* gene of m1 was

replaced by a *luciferase* gene, the effects of the 38 nt fragment on splicing were also the same as the vector having an *env* gene (data not shown). In addition, as shown in Figure 8, the analysis using vectors with the 38 nt fragment, i.e., 57-MLV, MSV, FeLV, XMRV, and CasBrE, showed that this region has a splicing function in all gamma retroviruses. Thus, the 38 nt fragment seems to play a universal role in splicing. FeLV had some splicing variants in addition to correctly spliced mRNA, suggesting that differences in the nucleotide sequence of the 38 nt region between FeLV and other viruses might have influenced its function in splicing.

We analyzed flanking sequences upstream of the 38 nt fragment to determine whether they influenced splicing (see Fig. 9). Our results showed that sequences with a positive or negative influence on splicing were scattered through the 1183-1611 nt upstream flanking region. The vector d3+353, produced by deletion of the 1183-1296 bp sequence from d3+467, yielded a small number of splice variants. In d3+206, produced by deletion of the 1297-1443 bp region from d3+353, abundant splice variants were recovered. These results indicate that the 1183-1443 nt region has a positive influence on splicing. When the 1444-1541 bp fragment was deleted from d3+206 to produce d3+108, only correct splicing was found; this suggested that the 1444-1541 nt region negatively influenced splicing. Deletion of the 1542-1569 bp fragment from d3+108 to produce d3+80 resulted in cryptic splicing; further deletion of the 1570-1591 bp fragment from d3+80 to give d3+58, resulted in the appearance of correctly spliced products. These results suggested that the 1542-1569 nt region and the 1570-1591 region positively and negatively influenced splicing, respectively. When the 1592-1611 bp fragment was deleted from d3+58 to produce d3+38, abundant splice variants appeared, suggesting that the 1592-1611 nt region positively influenced splicing. In the vectors used in this experiment, most of the intronic region was deleted. As stated above, it is also possible that changes in pre-mRNA features are also involved in regulating correct splicing. In addition, our results also indicated that the 38 nt fragment and its upstream sequences appear to exert the splicing function in cooperation with the 5140-5400 nt region located upstream region of the 3'ss, because when this fragment was deleted from the vectors that carried the 38 bp fragment and its flanking sequences and that yielded correctly spliced mRNA, abundant splice variants appeared (see Fig. 10).

We assume that cellular factors are recruited to the 38 nt region for its function in splicing,

because it is known that regulatory *cis*-elements recruit splicing activators or repressors for the recognition of splice sites in human immunodeficiency virus type 1 (Bilodeau et al., 2001; Caputi et al., 2004; Caputi and Zahler, 2002; Domsic et al., 2003; Jacquenet et al., 2005; Stoltzfus and Madsen, 2006; Tange et al., 2001; Tange and Kjems, 2001; Tranell et al., 2010; Zahler et al., 2004; Zhu et al., 2001). We were not able to identify candidate consensus sequences for the binding of such factors in the 38 nt region. However, we did identify several consensus sequences for the binding of splicing-related factors, such as hnRNP family proteins, SR family proteins, and PTB, within the upstream region (1183-1611 nt) of the 38 nt fragment (data not shown). Therefore, it is possible that 38 nt region interacts with upstream *cis*-elements that regulate splicing and that this interaction has an important role in the selection of splice sites. Although the mechanisms for this putative interaction are not yet clear, the 38 nt region might interact with cellular factors that bind to the *cis*-elements through its tertiary structure or via new host factors recruited into the 38 nt region.

As shown in Figure 11, the cryptic 5'ss and cryptic 3'ss, which were used in splice variants, clustered near the 38 bp fragment of MLV gene. Among these sites, we identified a few cryptic splicing sites, D4, D5, A6, and A7 that were frequently used. It has been reported that in cells infected with replication-competent Moloney MLV, a novel subgenomic 4.4 kb RNA is found. The subgenomic RNA is produced by splicing at an alternative 5'ss (1597 nt) within the *gag* region and the canonical 3'ss before the *env* gene (Dejardin et al., 2000; Houzet et al., 2003). Interestingly, the new 5'ss of Moloney MLV is identical to the D4 (1594 nt) of Friend-MLV. Since the cryptic splicing sites D4, D5, A6, and A7 were easily utilized in the MLV genome, the 38 nt fragment might play an important role in the attenuation of the activities of these potential cryptic splice sites.

CONCLUSION

In this thesis, we focused on splicing to understand mechanisms for posttranscriptional regulation of gene expression of MLV. First, we studied effects of splicing of Friend MLV *env*-mRNA on its 3' end processing and polysome structure formation using *env* expression vectors that produce spliced *env*-mRNA and unspliced *env*-mRNA. Since it is known that 3' processing of mRNA such as polyadenylation is an important factor in translational efficiency, we investigated whether splicing of *env*-mRNA affected its polyadenylation. The results showed that incomplete polyadenylation was detected in a fraction of the unspliced *env*-mRNA products, indicating that splicing of Friend MLV plays an important role in the efficiency of complete polyadenylation of *env*-mRNA. Formation of polysome structures of mRNA is also generally correlated with mRNA translation efficiency, thus we investigated whether splicing of MLV affected formation of *env*-mRNA polysome structures. It was revealed that more spliced than unspliced *env*-mRNA formed polysome structures. These splicing-dependent phenomena were not observed with expression vectors in which the *env* gene was replaced by the *luc* gene. Second, we analyzed *cis*-elements that regulate splicing of MLV. In previous studies, we showed that if the *Hind*III-*Bgl*II (879-1904 bp) fragment within *gag* is deleted from the m1 vector, which carries the entire Friend-MLV sequence, then cryptic splicing of *env*-mRNA occurs. To identify the genomic segment(s) in this region that are essential to correct splicing, we constructed vectors with a serially truncated *Hind*III-*Bgl*II fragment. The vector, in which a 38 bp fragment (1612-1649 bp) was deleted or reversed in m1, only produced splice variants. The results indicated that a 38 nt region within *gag* contained important elements that positively regulated splicing at the correct splice sites. Further analyses of a series of vectors carrying the 38 bp fragment and its flanking sequences showed that a region (1183-1611 nt) upstream of the 38 nt fragment also contained sequences that positively or negatively influenced splicing at the correct splice sites. We deleted the *Sph*I-*Nde*I (5140-5400 bp) fragment just upstream of the 3'ss from vectors that carried the 38 bp fragment and its flanking sequences and which yielded correctly spliced mRNA; interestingly, these deleted vectors showed cryptic splicing. The results suggested that the 5140-5400 nt region located just upstream of the 3'ss was required for the splicing function of the 38 nt fragment and its flanking

sequences. The results of this thesis showed new mechanisms for posttranscriptional regulation of gene expression of MLV, in which splicing of MLV promoted the efficiency of complete polyadenylation of *env*-mRNA and the formation of *env*-mRNA polysome structures in an *env* gene-dependent manner and the 5' and 3' splice sites of MLV are selected correctly through the function of the 38 nt region within *gag* gene.

ACKNOWLEDGMENTS

The author gratefully acknowledges the continuing guidance and encouragement of Prof. Sayaka Takase-Yoden, Department of Bioinformatics, Faculty of Engineering, Soka University.

The author is deeply indebted to Prof. Shoko Nishihara and Prof. Yuri Aoyama of the Department of Bioinformatics, Faculty of Engineering, Soka University, for reviewing the manuscript.

Thanks are also due to the members of Takase laboratory, and the members of the Department of Bioinformatics, Faculty of Engineering, Soka University.

REFERENCES

Abbink, T.E., Berkhout, B., 2008. RNA structure modulates splicing efficiency at the human immunodeficiency virus type 1 major splice donor. J Virol 82, 3090-3098.

Akimitsu, N., Tanaka, J., Pelletier, J., 2007. Translation of nonSTOP mRNA is repressed post-initiation in mammalian cells. EMBO J 26, 2327-2338.

Amendt, B.A., Hesslein, D., Chang, L.J., Stoltzfus, C.M., 1994. Presence of negative and positive *cis*-acting RNA splicing elements within and flanking the first tat coding exon of human immunodeficiency virus type 1. Mol Cell Biol 14, 3960-3970.

Amendt, B.A., Si, Z.H., Stoltzfus, C.M., 1995. Presence of exon splicing silencers within human immunodeficiency virus type 1 *tat* exon 2 and *tat-rev* exon 3: evidence for inhibition mediated by cellular factors. Mol Cell Biol 15, 4606-4615.

Arrigo, S., Beemon, K., 1988. Regulation of Rous sarcoma virus RNA splicing and stability. Mol Cell Biol 8, 4858-4867.

Awasthi, S., Alwine, J.C., 2003. Association of polyadenylation cleavage factor I with U1 snRNP. RNA 9, 1400-1409.

Beyer, A.L., Osheim, Y.N., 1988. Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. Genes Dev 2, 754-765.

Bilodeau, P.S., Domsic, J.K., Mayeda, A., Krainer, A.R., Stoltzfus, C.M., 2001. RNA splicing at human immunodeficiency virus type 1 3' splice site A2 is regulated by binding of hnRNP A/B proteins to an exonic splicing silencer element. J Virol 75, 8487-8497.

Black, D.L., 2003. Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem 72, 291-336.

Blanchette, M., Chabot, B., 1999. Modulation of exon skipping by high-affinity hnRNP A1-binding sites and by intron elements that repress splice site utilization. EMBO J 18, 1939-1952.

Blencowe, B.J., 2000. Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases. Trends Biochem Sci 25, 106-110.

Busch, A., Hertel, K.J., 2012. Evolution of SR protein and hnRNP splicing regulatory factors. Wiley Interdiscip Rev RNA 3, 1-12.

Caputi, M., Freund, M., Kammler, S., Asang, C., Schaal, H., 2004. A bidirectional SF2/ASF- and SRp40-dependent splicing enhancer regulates human immunodeficiency virus type 1 *rev*, *env*, *vpu*, and *nef* gene expression. J Virol 78, 6517-6526.

Caputi, M., Zahler, A.M., 2002. SR proteins and hnRNP H regulate the splicing of the HIV-1 tev-specific exon 6D. EMBO J 21, 845-855.

Cartegni, L., Chew, S.L., Krainer, A.R., 2002. Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat Rev Genet 3, 285-298.

Choo, Y.C., Seki, Y., Machinaga, A., Ogita, N., Takase-Yoden, S., 2013. The 0.3-kb fragment containing the R-U5-5'leader sequence of Friend murine leukemia virus influences the level of protein expression from spliced mRNA. Virol J 10, 124.

Coffin, J.M., Hughes, S.H., Varmus, H.E., 1997. The Interactions of Retroviruses and their Hosts. Retroviruses. Cold Spring Harbor, NY.

Cook, C.R., McNally, M.T., 1999. Interaction between the negative regulator of splicing element and a 3' splice site: requirement for U1 small nuclear ribonucleoprotein and the 3' splice site branch point/pyrimidine tract. J Virol 73, 2394-2400.

Cosson, B., Couturier, A., Chabelskaya, S., Kiktev, D., Inge-Vechtomov, S., Philippe, M., Zhouravleva, G., 2002. Poly(A)-binding protein acts in translation termination via eukaryotic release factor 3 interaction and does not influence [PSI(+)] propagation. Mol Cell Biol 22, 3301-3315.

Danckwardt, S., Hentze, M.W., Kulozik, A.E., 2008. 3' end mRNA processing: molecular mechanisms and implications for health and disease. EMBO J 27, 482-498.

Danckwardt, S., Kaufmann, I., Gentzel, M., Foerstner, K.U., Gantzert, A.S., Gehring, N.H., Neu-Yilik, G., Bork, P., Keller, W., Wilm, M., Hentze, M.W., Kulozik, A.E., 2007. Splicing factors stimulate polyadenylation via USEs at non-canonical 3' end formation signals. EMBO J 26, 2658-2669.

Dejardin, J., Bompard-Marechal, G., Audit, M., Hope, T.J., Sitbon, M., Mougel, M., 2000. A novel subgenomic murine leukemia virus RNA transcript results from alternative splicing. J Virol 74, 3709-3714.

del Prete, M.J., Vernal, R., Dolznig, H., Mullner, E.W., Garcia-Sanz, J.A., 2007. Isolation of

polysome-bound mRNA from solid tissues amenable for RT-PCR and profiling experiments. RNA 13, 414-421.

DesGroseillers, L., Barrette, M., Jolicoeur, P., 1984. Physical mapping of the paralysis-inducing determinant of a wild mouse ecotropic neurotropic retrovirus. J Virol 52, 356-363.

Diem, M.D., Chan, C.C., Younis, I., Dreyfuss, G., 2007. PYM binds the cytoplasmic exon-junction complex and ribosomes to enhance translation of spliced mRNAs. Nat Struct Mol Biol 14, 1173-1179.

Domsic, J.K., Wang, Y., Mayeda, A., Krainer, A.R., Stoltzfus, C.M., 2003. Human immunodeficiency virus type 1 hnRNP A/B-dependent exonic splicing silencer ESSV antagonizes binding of U2AF65 to viral polypyrimidine tracts. Mol Cell Biol 23, 8762-8772.

Dostie, J., Dreyfuss, G., 2002. Translation is required to remove Y14 from mRNAs in the cytoplasm. Curr Biol 12, 1060-1067.

Elkon, R., Ugalde, A.P., Agami, R., 2013. Alternative cleavage and polyadenylation: extent, regulation and function. Nat Rev Genet 14, 496-506.

Esposito, A.M., Mateyak, M., He, D., Lewis, M., Sasikumar, A.N., Hutton, J., Copeland, P.R., Kinzy, T.G., 2010. Eukaryotic polyribosome profile analysis. J Vis Exp.

Exline, C.M., Feng, Z., Stoltzfus, C.M., 2008. Negative and positive mRNA splicing elements act competitively to regulate human immunodeficiency virus type 1 vif gene expression. J Virol 82, 3921-3931.

Fogel, B.L., McNally, M.T., 2000. A cellular protein, hnRNP H, binds to the negative regulator of splicing element from Rous sarcoma virus. J Biol Chem 275, 32371-32378.

Gierer, A., 1963. Function of aggregated reticulocyte ribosomes in protein synthesis. J Mol Biol 6, 148-157.

Giles, K.E., Beemon, K.L., 2005. Retroviral splicing suppressor sequesters a 3' splice site in a 50S aberrant splicing complex. Mol Cell Biol 25, 4397-4405.

Giorgi, C., Moore, M.J., 2007. The nuclear nurture and cytoplasmic nature of localized mRNPs. Semin Cell Dev Biol 18, 186-193.

Graveley, B.R., 2000. Sorting out the complexity of SR protein functions. RNA 6, 1197-1211.

Holetz, F.B., Correa, A., Avila, A.R., Nakamura, C.V., Krieger, M.A., Goldenberg, S., 2007. Evidence of P-body-like structures in Trypanosoma cruzi. Biochem Biophys Res Commun 356, 1062-1067.

Hoshi, S., Odawara, T., Oshima, M., Kitamura, Y., Takizawa, H., Yoshikura, H., 2002. *cis*-Elements involved in expression of unspliced RNA in Moloney murine leukemia virus. Biochem Biophys Res Commun 290, 1139-1144.

Houzet, L., Battini, J.L., Bernard, E., Thibert, V., Mougel, M., 2003. A new retroelement constituted by a natural alternatively spliced RNA of murine replication-competent retroviruses. EMBO J 22, 4866-4875.

Jacquenet, S., Decimo, D., Muriaux, D., Darlix, J.L., 2005. Dual effect of the SR proteins ASF/SF2, SC35 and 9G8 on HIV-1 RNA splicing and virion production. Retrovirology 2, 33.

Jacquenet, S., Ropers, D., Bilodeau, P.S., Damier, L., Mougin, A., Stoltzfus, C.M., Branlant, C., 2001. Conserved stem-loop structures in the HIV-1 RNA region containing the A3 3' splice site and its *cis*-regulatory element: possible involvement in RNA splicing. Nucleic Acids Res 29, 464-478.

Kammler, S., Leurs, C., Freund, M., Krummheuer, J., Seidel, K., Tange, T.O., Lund, M.K., Kjems, J., Scheid, A., Schaal, H., 2001. The sequence complementarity between HIV-1 5' splice site SD4 and U1 snRNA determines the steady-state level of an unstable *env* pre-mRNA. RNA 7, 421-434.

Kammler, S., Otte, M., Hauber, I., Kjems, J., Hauber, J., Schaal, H., 2006. The strength of the HIV-1 3' splice sites affects Rev function. Retrovirology 3, 89.

Karn, J., Stoltzfus, C.M., 2012. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. Cold Spring Harb Perspect Med 2, a006916.

Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., Ohno, M., Dreyfuss, G., Ohno, S., 2006. Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. Genes Dev 20, 355-367.

Katz, R.A., Kotler, M., Skalka, A.M., 1988. *cis*-acting intron mutations that affect the efficiency of avian retroviral RNA splicing: implication for mechanisms of control. J Virol 62,

2686-2695.

Kim, V.N., Yong, J., Kataoka, N., Abel, L., Diem, M.D., Dreyfuss, G., 2001. The Y14 protein communicates to the cytoplasm the position of exon-exon junctions. EMBO J 20, 2062-2068.

Kramer, A., 1996. The structure and function of proteins involved in mammalian pre-mRNA splicing. Annu Rev Biochem 65, 367-409.

Kraunus, J., Schaumann, D.H., Meyer, J., Modlich, U., Fehse, B., Brandenburg, G., von Laer, D., Klump, H., Schambach, A., Bohne, J., Baum, C., 2004. Self-inactivating retroviral vectors with improved RNA processing. Gene Ther 11, 1568-1578.

Kuhn, U., Wahle, E., 2004. Structure and function of poly(A) binding proteins. Biochim Biophys Acta 1678, 67-84.

Kyburz, A., Friedlein, A., Langen, H., Keller, W., 2006. Direct interactions between subunits of CPSF and the U2 snRNP contribute to the coupling of pre-mRNA 3' end processing and splicing. Mol Cell 23, 195-205.

Le Hir, H., Andersen, G.R., 2008. Structural insights into the exon junction complex. Curr Opin Struct Biol 18, 112-119.

Le Hir, H., Gatfield, D., Izaurralde, E., Moore, M.J., 2001. The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. EMBO J 20, 4987-4997.

Le Hir, H., Izaurralde, E., Maquat, L.E., Moore, M.J., 2000. The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. EMBO J 19, 6860-6869.

Le Hir, H., Seraphin, B., 2008. EJCs at the heart of translational control. Cell 133, 213-216.

Lee, H.C., Choe, J., Chi, S.G., Kim, Y.K., 2009. Exon junction complex enhances translation of spliced mRNAs at multiple steps. Biochem Biophys Res Commun 384, 334-340.

Lejeune, F., Ishigaki, Y., Li, X., Maquat, L.E., 2002. The exon junction complex is detected on CBP80-bound but not eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodeling. EMBO J 21, 3536-3545.

Li, Y., Chen, Z.Y., Wang, W., Baker, C.C., Krug, R.M., 2001. The 3'-end-processing factor CPSF

is required for the splicing of single-intron pre-mRNAs in vivo. RNA 7, 920-931.

Lin, S., Fu, X.D., 2007. SR proteins and related factors in alternative splicing. Adv Exp Med Biol 623, 107-122.

Luco, R.F., Allo, M., Schor, I.E., Kornblihtt, A.R., Misteli, T., 2011. Epigenetics in alternative pre-mRNA splicing. Cell 144, 16-26.

Machinaga, A., Takase-Yoden, S., 2014. A 38 nt region and its flanking sequences within *gag* of Friend murine leukemia virus are crucial for splicing at the correct 5' and 3' splice sites. Microbiol Immunol 58, 38-50.

Madsen, J.M., Stoltzfus, C.M., 2005. An exonic splicing silencer downstream of the 3' splice site A2 is required for efficient human immunodeficiency virus type 1 replication. J Virol 79, 10478-10486.

Mangus, D.A., Evans, M.C., Jacobson, A., 2003. Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. Genome Biol 4, 223.

Manley, J.L., Krainer, A.R., 2010. A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). Genes Dev 24, 1073-1074.

Martinez-Contreras, R., Cloutier, P., Shkreta, L., Fisette, J.F., Revil, T., Chabot, B., 2007. hnRNP proteins and splicing control. Adv Exp Med Biol 623, 123-147.

Martinez-Contreras, R., Fisette, J.F., Nasim, F.U., Madden, R., Cordeau, M., Chabot, B., 2006. Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing. PLoS Biol 4, e21.

Masuda, M., Hoffman, P.M., Ruscetti, S.K., 1993. Viral determinants that control the neuropathogenicity of PVC-211 murine leukemia virus in vivo determine brain capillary endothelial cell tropism of the virus in vitro. J Virol 67, 4580-4587.

Masuda, M., Remington, M.P., Hoffman, P.M., Ruscetti, S.K., 1992. Molecular characterization of a neuropathogenic and nonerythroleukemogenic variant of Friend murine leukemia virus PVC-211. J Virol 66, 2798-2806.

McCracken, S., Lambermon, M., Blencowe, B.J., 2002. SRm160 splicing coactivator promotes transcript 3'-end cleavage. Mol Cell Biol 22, 148-160.

McNally, L.M., McNally, M.T., 1996. SR protein splicing factors interact with the Rous sarcoma

virus negative regulator of splicing element. J Virol 70, 1163-1172.

McNally, L.M., McNally, M.T., 1998. An RNA splicing enhancer-like sequence is a component of a splicing inhibitor element from Rous sarcoma virus. Mol Cell Biol 18, 3103-3111.

McNally, L.M., Yee, L., McNally, M.T., 2006. Heterogeneous nuclear ribonucleoprotein H is required for optimal U11 small nuclear ribonucleoprotein binding to a retroviral RNA-processing control element: implications for U12-dependent RNA splicing. J Biol Chem 281, 2478-2488.

McNally, M.T., Beemon, K., 1992. Intronic sequences and 3' splice sites control Rous sarcoma virus RNA splicing. J Virol 66, 6-11.

McNally, M.T., Gontarek, R.R., Beemon, K., 1991. Characterization of Rous sarcoma virus intronic sequences that negatively regulate splicing. Virology 185, 99-108.

Miller, J.T., Stoltzfus, C.M., 1992. Two distant upstream regions containing *cis*-acting signals regulating splicing facilitate 3'-end processing of avian sarcoma virus RNA. J Virol 66, 4242-4251.

Millevoi, S., Loulergue, C., Dettwiler, S., Karaa, S.Z., Keller, W., Antoniou, M., Vagner, S., 2006. An interaction between U2AF 65 and CF I(m) links the splicing and 3' end processing machineries. EMBO J 25, 4854-4864.

Millevoi, S., Vagner, S., 2010. Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. Nucleic Acids Res 38, 2757-2774.

Nashchekin, D., Zhao, J., Visa, N., Daneholt, B., 2006. A novel Ded1-like RNA helicase interacts with the Y-box protein ctYB-1 in nuclear mRNP particles and in polysomes. J Biol Chem 281, 14263-14272.

Nott, A., Le Hir, H., Moore, M.J., 2004. Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. Genes Dev 18, 210-222.

Oshima, M., Odawara, T., Hanaki, K., Igarashi, H., Yoshikura, H., 1998. *cis* Elements required for high-level expression of unspliced Gag-containing message in Moloney murine leukemia virus. J Virol 72, 6414-6420.

Oshima, M., Odawara, T., Matano, T., Sakahira, H., Kuchino, Y., Iwamoto, A., Yoshikura, H., 1996. Possible role of splice acceptor site in expression of unspliced *gag*-containing

message of Moloney murine leukemia virus. J Virol 70, 2286-2295.

Otulakowski, G., Rafii, B., O'Brodovich, H., 2004. Differential translational efficiency of ENaC subunits during lung development. Am J Respir Cell Mol Biol 30, 862-870.

Paquette, Y., Hanna, Z., Savard, P., Brousseau, R., Robitaille, Y., Jolicoeur, P., 1989. Retrovirus-induced murine motor neuron disease: mapping the determinant of spongiform degeneration within the envelope gene. Proc Natl Acad Sci U S A 86, 3896-3900.

Reichert, V.L., Le Hir, H., Jurica, M.S., Moore, M.J., 2002. 5' exon interactions within the human spliceosome establish a framework for exon junction complex structure and assembly. Genes Dev 16, 2778-2791.

Ryu, D.K., Kim, S., Ryu, W.S., 2008. Hepatitis B virus polymerase suppresses translation of pregenomic RNA via a mechanism involving its interaction with 5' stem-loop structure. Virology 373, 112-123.

Sakuma, T., Davila, J.I., Malcolm, J.A., Kocher, J.P., Tonne, J.M., Ikeda, Y., 2014. Murine leukemia virus uses NXF1 for nuclear export of spliced and unspliced viral transcripts. J Virol 88, 4069-4082.

Schaal, T.D., Maniatis, T., 1999. Multiple distinct splicing enhancers in the protein-coding sequences of a constitutively spliced pre-mRNA. Mol Cell Biol 19, 261-273.

Schmidt, U., Im, K.B., Benzing, C., Janjetovic, S., Rippe, K., Lichter, P., Wachsmuth, M., 2009. Assembly and mobility of exon-exon junction complexes in living cells. RNA 15, 862-876.

Schrom, E.M., Moschall, R., Schuch, A., Bodem, J., 2013. Regulation of retroviral polyadenylation. Adv Virus Res 85, 1-24.

Si, Z., Amendt, B.A., Stoltzfus, C.M., 1997. Splicing efficiency of human immunodeficiency virus type 1 tat RNA is determined by both a suboptimal 3' splice site and a 10 nucleotide exon splicing silencer element located within tat exon 2. Nucleic Acids Res 25, 861-867.

Si, Z.H., Rauch, D., Stoltzfus, C.M., 1998. The exon splicing silencer in human immunodeficiency virus type 1 Tat exon 3 is bipartite and acts early in spliceosome assembly. Mol Cell Biol 18, 5404-5413.

Singh, R., Valcarcel, J., 2005. Building specificity with nonspecific RNA-binding proteins. Nat Struct Mol Biol 12, 645-653.

Smith, C.W., Valcarcel, J., 2000. Alternative pre-mRNA splicing: the logic of combinatorial control. Trends Biochem Sci 25, 381-388.

Spellman, R., Smith, C.W., 2006. Novel modes of splicing repression by PTB. Trends Biochem Sci 31, 73-76.

Staffa, A., Cochrane, A., 1995. Identification of positive and negative splicing regulatory elements within the terminal tat-rev exon of human immunodeficiency virus type 1. Mol Cell Biol 15, 4597-4605.

Stoltzfus, C.M., 2009. Chapter 1. Regulation of HIV-1 alternative RNA splicing and its role in virus replication. Adv Virus Res 74, 1-40.

Stoltzfus, C.M., Fogarty, S.J., 1989. Multiple regions in the Rous sarcoma virus *src* gene intron act in *cis* to affect the accumulation of unspliced RNA. J Virol 63, 1669-1676.

Stoltzfus, C.M., Madsen, J.M., 2006. Role of viral splicing elements and cellular RNA binding proteins in regulation of HIV-1 alternative RNA splicing. Curr HIV Res 4, 43-55.

Szurek, P.F., Yuen, P.H., Jerzy, R., Wong, P.K., 1988. Identification of point mutations in the envelope gene of Moloney murine leukemia virus TB temperature-sensitive paralytogenic mutant ts1: molecular determinants for neurovirulence. J Virol 62, 357-360.

Takase-Yoden, S., Wada, M., Watanabe, R., 2006. A viral non-coding region determining neuropathogenicity of murine leukemia virus A8 is responsible for envelope protein expression in the rat brain. Microbiol Immunol 50, 197-201.

Takase-Yoden, S., Watanabe, R., 1997. Unique sequence and lesional tropism of a new variant of neuropathogenic friend murine leukemia virus. Virology 233, 411-422.

Takase-Yoden, S., Watanabe, R., 2005. A 0.3-kb fragment containing the R-U5-5' leader sequence is essential for the induction of spongiform neurodegeneration by A8 murine leukemia virus. Virology 336, 1-10.

Tange, T.O., Damgaard, C.K., Guth, S., Valcarcel, J., Kjems, J., 2001. The hnRNP A1 protein regulates HIV-1 *tat* splicing via a novel intron silencer element. EMBO J 20, 5748-5758.

Tange, T.O., Kjems, J., 2001. SF2/ASF binds to a splicing enhancer in the third HIV-1 *tat* exon and stimulates U2AF binding independently of the RS domain. J Mol Biol 312, 649-662.

Tange, T.O., Nott, A., Moore, M.J., 2004. The ever-increasing complexities of the exon junction

complex. Curr Opin Cell Biol 16, 279-284.

Tranell, A., Fenyo, E.M., Schwartz, S., 2010. Serine- and arginine-rich proteins 55 and 75 (SRp55 and SRp75) induce production of HIV-1 *vpr* mRNA by inhibiting the 5'-splice site of exon 3. J Biol Chem 285, 31537-31547.

Wagner, E.J., Garcia-Blanco, M.A., 2001. Polypyrimidine tract binding protein antagonizes exon definition. Mol Cell Biol 21, 3281-3288.

Warner, J.R., Knopf, P.M., Rich, A., 1963. A multiple ribosomal structure in protein synthesis. Proc Natl Acad Sci U S A 49, 122-129.

Warner, J.R., Rich, A., Hall, C.E., 1962. Electron Microscope Studies of Ribosomal Clusters Synthesizing Hemoglobin. Science 138, 1399-1403.

Wiegand, H.L., Lu, S., Cullen, B.R., 2003. Exon junction complexes mediate the enhancing effect of splicing on mRNA expression. Proc Natl Acad Sci U S A 100, 11327-11332.

Wilusz, J.E., Beemon, K.L., 2006. The negative regulator of splicing element of Rous sarcoma virus promotes polyadenylation. J Virol 80, 9634-9640.

Yamamoto, N., Takase-Yoden, S., 2009. Friend murine leukemia virus A8 regulates Env protein expression through an intron sequence. Virology 385, 115-125.

Yuen, P.H., Tzeng, E., Knupp, C., Wong, P.K., 1986. The neurovirulent determinants of ts1, a paralytogenic mutant of Moloney murine leukemia virus TB, are localized in at least two functionally distinct regions of the genome. J Virol 59, 59-65.

Zahler, A.M., Damgaard, C.K., Kjems, J., Caputi, M., 2004. SC35 and heterogeneous nuclear ribonucleoprotein A/B proteins bind to a juxtaposed exonic splicing enhancer/exonic splicing silencer element to regulate HIV-1 *tat* exon 2 splicing. J Biol Chem 279, 10077-10084.

Zhou, Z., Fu, X.D., 2013. Regulation of splicing by SR proteins and SR protein-specific kinases. Chromosoma 122, 191-207.

Zhu, J., Mayeda, A., Krainer, A.R., 2001. Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. Mol Cell 8, 1351-1361.

Zychlinski, D., Erkelenz, S., Melhorn, V., Baum, C., Schaal, H., Bohne, J., 2009. Limited

complementarity between U1 snRNA and a retroviral 5' splice site permits its attenuation via RNA secondary structure. Nucleic Acids Res 37, 7429-7440.