

# Studies on the regulation of gene expression by a 0.3-kb fragment containing the R-U5-5' leader sequence of murine leukemia virus

## マウス白血病ウイルスの R-U5-5'リーダ配列を含む 0.3-kb 領域による遺伝子発現制御に関する研究

09D5602 朱瑩青 指導教授 高瀬明

### SYNOPSIS

マウス白血病ウイルス (MLV) は、ウイルスゲノムに *gag*, *pol*, *env* 遺伝子のみをコードしているシンプルレトロウイルスである。MLV のプロウイルスからは、スプライシングを受けていない全長の mRNA と、5'リーダ配列内に存在するスプライシング供与部位、および、*pol* 遺伝子内に存在するスプライシング受容部位でスプライシングを受けた *env*-mRNA の 2 種類の mRNA が生成される。神経病原性 A8 と非神経病原性 57 の様々な組換えウイルスを用いた研究から、病原性を決定する遺伝子は *env* であることが明らかにされている。さらに、R-U5-5'リーダ配列を含む 0.3-kb 非コード領域も病原性発現に重要な領域であることが明らかにされており、この領域は Env タンパク質の発現量に影響を及ぼすことが示されているが、その機序は分かっていない。A8 の 0.3-kb 領域と A8-*env* を有する神経病原性 R7f ウイルス、および、R7f の 0.3-kb 領域を 57 の配列に置き換えた非病原性 Rec5 ウイルスを NIH3T3 細胞に感染させ、ウイルスタンパク質の発現量、mRNA 量、および、ウイルス産生量を経時的に調べた。その結果、R7f 感染細胞における Env タンパク質の発現量および *env*-mRNA 量は、Rec5 に比べて高かったが、Gag タンパク質の発現量、全 mRNA 量、および、ウイルス産生量は同様であった。0.3-kb 領域がどのような機構で Env タンパク質の発現量を制御しているのかを明らかにするために、R7f および Rec5 の *env* 遺伝子をルシフェラーゼ遺伝子に置き換えた R7f-L ベクター、および、Rec5-L ベクターを作製し、NIH3T3 細胞に導入した。R7f-L のルシフェラーゼ活性は、Rec5-L よりも 2 倍高かった。R7f-L および Rec5-L から転写された全 mRNA 量は同じであったが、R7f-L の spliced-mRNA 量は、Rec5-L に比べて、2 倍多かった。また、R7f-L 導入細胞における spliced-mRNA の核-細胞質存在比は、Rec5-L 導入細胞の場合と同じであった。R7f-L の mRNA の polyA 鎖の長さも、Rec5-L の場合と同様であった。これらの結果は、0.3-kb 領域は、mRNA の転写量、spliced *luc*-mRNA の核外輸送効率、および、polyA 鎖長には影響しないが、スプライシングの効率に影響を及ぼすことにより、Env タンパク質の発現量を制御することを示唆している。0.3-kb 領域における A8 と 57 の遺伝子配列の違いは 17 塩基である。これらのうち、どの塩基がルシフェラーゼ活性に影響を及ぼすかを、点変異導入ベクターを用いて解析した。その結果、0.3-kb 領域内の 17 個の置換塩基のうち、5'末端側より 7 個の塩基が遺伝子発現に重要であることが示された。mRNA の 2 次構造予測の結果、これらの塩基は、5'スプライス部位の認識に関与することが示唆されているステムループ構造の中に存在することが示された。

Keywords: マウス白血病ウイルス、遺伝子発現、非コード領域、スプライシング

### 1. Introduction

Murine Leukemia Virus (MLV) belongs to the simple retrovirus family. The simple retroviruses are characterized by a coding structure in which the *gag*, *pol* and *env* genes are flanked by two long terminal repeats (LTRs), a 5'LTR and 3'LTR. Proteins responsible for the constitution of the inner structures of the virion are encoded by the *gag* gene, which includes the matrix, capsid and nucleocapsid proteins. The *pol* gene encodes the enzymatic proteins, i.e. the reverse transcriptase, protease, integrase and RNase H and the *env* gene encodes the proteins protruding out from the viral particle surface, namely the surface (SU) and transmembrane (TM) proteins (1). Transcription begins from the R region of the 5'LTR and ends at the polyadenylation signal located at the R region at the other end of the 3'LTR. A 5'splice site (5'ss) is located in the 5'leader sequence and a 3'splice site (3'ss) is located at the 3' end of the *pol* gene. Only a singly spliced mRNA is usually found in simple retroviruses. Gag and Pol proteins are translated from the unspliced full-length viral mRNA, and the Env protein is translated from the spliced *env*-mRNA (1). In contrast, it is reported

that HIV-1, which belongs to the complex retrovirus family, could generate up to 40 different spliced RNAs using four 5'ss and nine 3'ss (2).

In a previous study of a neuropathogenic variant of Friend MLV clone A8, we showed that a high level of expression of A8-Env protein in rat brain was correlated with neuropathogenicity (3-5). Studies with chimeras constructed from the A8 virus and the non-neuropathogenic Fr-MLV clone 57 identified a 0.3-kb fragment containing the R-U5-5'leader sequence as an important determinant of neuropathogenicity, in addition to the *env* gene of A8 as the primary determinant (3). Chimeric virus Rec5, which contains the A8-*env* gene on the background of 57, did not exhibit neuropathogenicity. In contrast, the chimeric virus R7f, which contains a 0.3-kb fragment of A8 and the A8-*env* gene on the background of 57, induced spongiform neurodegeneration. It has been shown that the 0.3-kb fragment influences the expression levels of Env protein in both cultured cells and rat brain (3, 4). This fragment contains functional domains, such as a signal for poly (A) addition to mRNA that works in the 3' long terminal repeat (LTR), a primer binding site (PBS)

for reverse transcription, and a 5' ss. However the steps of gene expression at which the 0.3-kb fragment may influence Env expression have not yet to be elucidated. The goal of this study is to elucidate the molecular mechanisms for the regulation of gene expression by the 0.3-kb fragment containing the R-U5-5' leader sequence of MLV.

## **2. Materials and Methods**

### **2.1 Time course analysis of viral proliferation**

DNA clone of R7f and Rec5 was constructed as previously described (3). NIH3T3 cells were infected by R7f and Rec5 virus at moi 1. Infected 1 dpi to 5 dpi cells were harvested and downstream assay were performed accordingly.

### **2.2 Focal Immunoassay**

Virus titer was determined by infecting *M. dunnii* cells at  $2 \times 10^4$  cells in 6 well plates. Env protein on cell surface was stained with monoclonal antibody 85-1. Cells were fixed in 0.5% glutaraldehyde and horse peroxidase-conjugated sheep anti-Mouse IgG whole antibody (GE Healthcare) was used as secondary antibody.

### **2.3 Immunoblot analysis**

Env protein and Gag protein, was detected using goat anti-Rauscher MLV gp70 and anti-AKR p30<sup>Gag</sup> (Quality Biotech Incorporated Resource Laboratory). Actin was detected using rabbit anti-beta-actin (Santa Cruz Biotechnology) as a loading control. Horseradish peroxidase-conjugated anti-goat IgG antibody and anti-rabbit IgG (Santa Cruz Biotechnology) were used as a secondary antibody. Bands were visualised after developing with ECL plus reagents (Amersham Bioscience Corp.).

### **2.4 Vector construction and Cell Transfection**

Luciferase expression vectors were constructed by replacing the viral *env* gene with *luciferase* gene to generate the R7f-L, Rec5-L, R7fa-L and R7fb-L vectors. A series of point mutation vectors were constructed by PCR mutagenesis. For cell transfection, plasmid was diluted in OPTI-MEM (Invitrogen) and transfected using Lipofectamine 2000 Reagent (Invitrogen). All downstream analysis is carried on 48-hours post transfected NIH3T3 cells unless otherwise stated.

### **2.5 Luciferase Assay**

Luciferase activity was determined using Dual-Luciferase Assay System (Promega) following manufacturer's protocol, and normalized to the activity of Renilla luciferase activity of co-transfected pRL-SV40.

### **2.6 Genomic DNA and RNA extraction and analysis**

DNA and RNA extraction was carried out using DNase Blood and Tissue Kit (Qiagen) and RNase Mini Kit (Qiagen), respectively. 2ug of RNA were reverse transcribed using OligodT<sub>20</sub> primer and SSIII reverse transcribing kit (Invitrogen). Quantification of viral DNA, plasmid DNA, total mRNA, and spliced-variant mRNA

were performed with real-time RT-PCR using specific primers and the Taqman probe. As internal control, *gapdh* DNA and *gapdh*-mRNA were detected.

### **2.7 Cell Fractionation**

Nuclear and cytoplasmic fractions were obtained using PARIS kit (Ambion) according to the manufacturer's manual. As a control for the fractionation, cellular ribosomal RNA from each section was electrophoresed on a 1% agarose gel in morpholinepropane-sulfonic acid (MOPS) buffer.

### **2.8 Determination of PolyA-tail length**

Total RNA extracted from 24 hours post-transfected HeLa-cells were ligated with RV3PC-anchor primer. Reverse transcription is then carried with an antisense sequence of the RV3PC-anchor primer. Resulted cDNA were subjected to PCR and electrophoresed and visualized in ethidium-bromide (Et-BR) staining.

## **3. Results and Discussion**

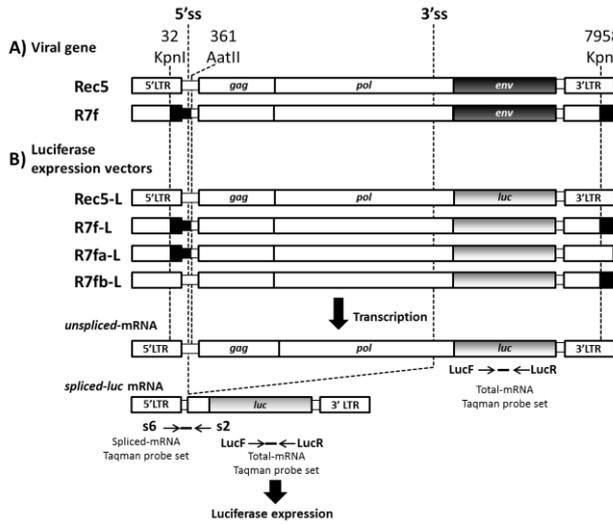
### **3.1 Kinetic analysis of the effects of the 0.3-kb fragment on viral gene expression**

The effect of the 0.3-kb fragment were studied based on two chimeric viruses, Rec5 and R7f where each carries the 0.3-kb fragment of 57 and A8 respectively and an A8-*env* gene (Fig 1A). On viral growth, we observed no difference in the level of viral production between R7f and Rec5 throughout the 5 days of incubation with viral titer at  $10^4$  at 1 dpi and reaching a lag phase above  $10^6$  from 3 dpi onwards for both R7f and Rec5 (data not shown). Results for kinetic analysis of viral protein and viral mRNA are summarized in Table 1. Immunoassay showed that the expression level of Env protein in R7f-infected cells at 3 dpi and 5 dpi was three- to four-fold higher than that of Rec5. In contrast, the amount of Gag protein in R7f-infected cells was similar to that in Rec5-infected cells. Furthermore, we found that amount of total viral mRNA was similar in R7f- and Rec5-infected cells over the course of the observation period. The amount of spliced *env*-mRNA however was seen twice higher in R7f compared to Rec5 at 3 dpi ( $p < 0.01$ ) and at 5 dpi ( $p < 0.005$ ). These analysis suggested that the 0.3-kb fragment influenced the expression level of the Env protein by regulating the amount of spliced *env*-mRNA rather than the amount of total viral mRNA or viral production.

### **3.2 Influence of the 0.3-kb fragment in protein expression and amount of mRNA**

To further analyze the molecular mechanism governed by the 0.3-kb fragment, we constructed the Luciferase expression vectors (Fig. 1B). The luciferase activity of R7f-L increased by 2-fold than that of Rec5-L ( $p < 0.001$ ), (Fig. 2A). The luciferase activity of R7fa-L, which carries the 0.3-kb fragment of A8 at the 5'LTR, was the same as that of R7f-L, and the luciferase activity of R7fb-L, which carries the 0.3-kb fragment of A8 at the 3'LTR, was the

same as that of Rec5-L. These results showed that 0.3-kb fragment exhibits its function dominantly from the 5'LTR instead of the 3'LTR. Furthermore, the amount of total mRNA and spliced mRNA were measured. The amount of spliced mRNAs paralleled the luciferase activity where R7f-L and R7fa-L exhibited about 2-fold increase compared to that of Rec5-L ( $p < 0.001$ ) and R7fb-L (Fig.2B). The amount of total transcripts was also measured and all expression vectors showed comparable amount (Fig.2B).



**Fig. 1.** Structures of viral genes (A) and luciferase expression vectors (B). In the viral genomes, solid regions are sequence derived from the A8 virus and open region are sequence derived from 57 virus. The numbering of nucleotides is based on the transcript.

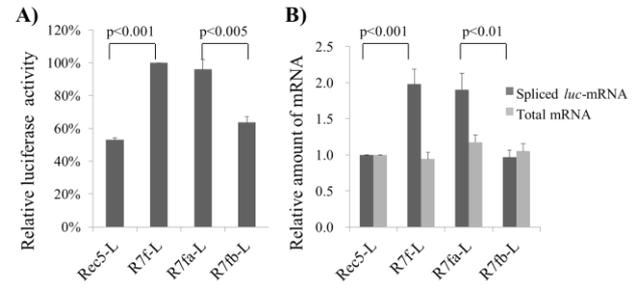
Virus	dpi	Relative amount <sup>a</sup> of		Copy number <sup>b</sup> of		The ratio <sup>c</sup> of env-mRNA per total viral mRNA
		Env protein	Gag protein	env-mRNA	total viral mRNA	
R7f	1	1.0 ± 0.0	1.0 ± 0.0	1.6 ± 0.3 × 10 <sup>4</sup>	2.9 ± 0.6 × 10 <sup>5</sup>	0.06 ± 0.00
	3	8.3 ± 2.3*	2.9 ± 0.3	1.7 ± 0.4 × 10 <sup>5</sup>	1.4 ± 0.5 × 10 <sup>6</sup>	0.14 ± 0.02**
	5	10.3 ± 1.2**	4.1 ± 0.9	3.0 ± 0.4 × 10 <sup>5</sup>	2.4 ± 0.5 × 10 <sup>6</sup>	0.13 ± 0.01****
Rec5	1	0.5 ± 0.1	1.1 ± 0.3	1.6 ± 0.5 × 10 <sup>4</sup>	2.6 ± 0.7 × 10 <sup>5</sup>	0.06 ± 0.01
	3	1.9 ± 0.2	2.4 ± 0.6	7.4 ± 1.8 × 10 <sup>4</sup>	1.1 ± 0.3 × 10 <sup>6</sup>	0.08 ± 0.01
	5	3.1 ± 0.7	3.0 ± 1.0	1.4 ± 0.2 × 10 <sup>5</sup>	2.0 ± 0.4 × 10 <sup>6</sup>	0.07 ± 0.01

**Table 1** Comparison of viral protein expression levels, env-mRNA and total viral mRNA levels, and the ratios of the amount of env-mRNA to that of total viral mRNA. The means of three ( $\pm$ SEM)<sup>a</sup> and four ( $\pm$ SEM)<sup>b,c</sup> independent experiments are shown. Statistical comparisons were performed using the *t* test. \* $p < 0.05$  vs Env of Rec5 at 3 dpi; \*\* $p < 0.005$  vs Env of Rec5 at 5 dpi; \*\*\* $p < 0.01$  vs the ratio of Rec5 at 3dpi; \*\*\*\* $p < 0.01$  vs the ratio of Rec5 at 5 dpi.

### 3.2 Influence of the 0.3-kb fragment towards the poly (A) tail length and nuclear-cytoplasmic distribution of luc-mRNA

In general, polyA-tail length is correlated with stability of mRNA and efficiency of translation. Therefore, we compared polyA-tail length of mRNA of R7f-L and Rec5-L. Differences polyA-tail length between the total transcripts from R7f-L and Rec5-L were not detected (data not shown). The distribution of luc-mRNA in the nucleus and cytoplasm of NIH3T3 cells with introduced Rec5-L and R7f-L were seen to be equal where only about 15% of luc-mRNA were detected in the cytoplasmic fraction and largely 85% were maintained in the nuclear fraction (data not shown). *Gapdh*-mRNA as the distribution control was

equal in cells transfected with Rec5-L and R7f-L and showed distribution predominantly in the cytoplasmic fraction as expected. These results suggest that the role of 0.3-kb was seen to contribute largely to the ability of transcripts to be spliced and that an enhanced nuclear export of mRNA or polyA-tail length did not explain in the increase of protein expression.



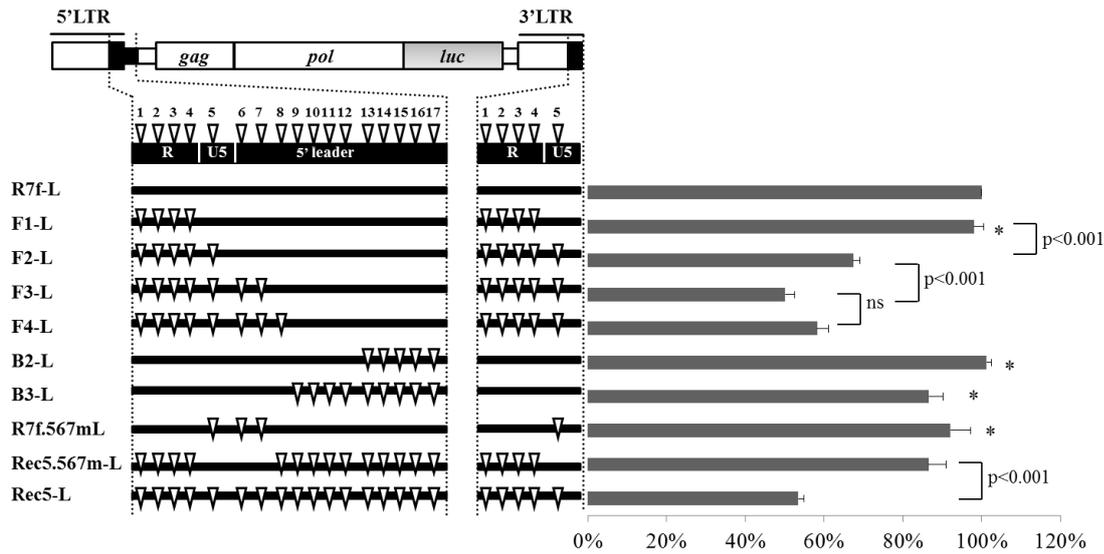
**Fig. 2** Relative luciferase activity (A) and mRNA analysis (B)

### 3.3 Point mutation analysis

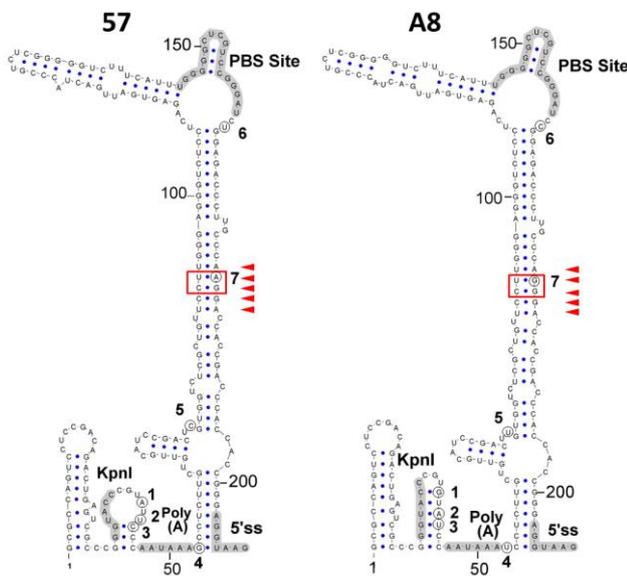
Following the results of characterizing gene expression of Rec5-L and R7f-L, we set out to determine if any of the nucleotides within the 0.3-kb fragment are key(s) to the observed splicing and luciferase expression effects. A series of point mutation vectors where A8 sequences are gradually mutated into 57 sequences were constructed and their luciferase activities were determined (Fig.3). There was no significant difference in the luciferase activity of F1-L, B2-L and B3-L compared to R7f-L. The luciferase activity of F2-L was lower than that of F1-L ( $p < 0.001$ ), and the luciferase activity of F3-L, which showed same luciferase activity as FL-4, was lower than that of F2-L ( $p < 0.001$ ). These decreases correspond to the 5<sup>th</sup>, 6<sup>th</sup> and 7<sup>th</sup> nucleotide being mutated into 57 sequences. We asked if the 5th, 6th and 7th nucleotides alone could contribute to the regulation of luciferase activity. Towards this end, we constructed R7f.567m-L and Rec5.567m-L. The luciferase activity of R7f.567 m-L remained at about 95% and could not be brought down to parallel that of Rec5-L, while its exact reverse vector, Rec5.567 m-L, had a significantly increased luciferase activity (86%) that was higher than that of Rec5-L ( $p < 0.001$ ). These results suggested that the 1st to 7th nucleotide of the 0.3-kb fragment were important regulators of the luciferase protein expression level.

### 3.4 Secondary structure analysis

To explain how the 1st to 7th nucleotides might be important for luc-mRNA expression, we mapped out the secondary structure formed by the sequence containing the 1st to 7th nucleotides of the 0.3-kb fragment of the A8 and 57 sequences. Figure 4 illustrates the major functional secondary structures of MLV. At first glance, there is not a striking difference between the two secondary structures generated, despite the 7 nucleotides that differ between the A8 and 57 sequences. The most visible changes actually occur upstream from the polyadenylation signal, where the 1st, 2nd, and 3rd nucleotides are incorporated into a stem structure in the A8 sequence, thereby lengthening the stem



**Fig.3** Point mutation vectors on left side of chart and its corresponding relative activities of luciferase on the right. Number 1 to 17 denotes each of the nucleotide that differs between A8 and 57 sequences. Mutations from A8 to 57 are indicated by triangles. The mean values from 4-7 independent experiments and the SEM are shown. Statistical comparison was done using the *t* test. ns: differences were not significant. \*: differences were not significant versus R7f-L.



**Fig. 4.** Predicted secondary structure formed by the sequence containing the 1<sup>st</sup> to 7<sup>th</sup> nucleotides of the 0.3-kb fragment of the A8 sequence [accession no. D88386] and the 57 sequence [accession no. X02794]. This representation shows the results of an MFOLD simulation on the basis of previous studies (6) and the figure was drawn using VARNA software. Nucleotides that differ between A8 and 57 are circled and numbered from 1 to 7. Important regulatory signals are highlighted. PolyA: polyadenylation signal; PBS: primer binding site; 5'ss: 5' splice site. Restriction enzyme KpnI is shown.

structure compared to the 57 sequence. The site with the smallest conformational change contains the 5th, 6th and 7th nucleotides. These three nucleotides reside within a stem-loop structure that protrudes out into the PBS. Kraunus et al. (7) reported that the higher complementarity of bases facing each other in the boxed motif (Fig.4) decreased the splicing efficiency. This suggests that the 7th nucleotide plays an important role in luciferase expression by participating in the splicing step.

#### 4. Conclusion

This study has tried to elucidate the molecular mechanisms for the regulation of gene expression by a 0.3-kb fragment containing the R-U5-5' leader sequence of MLV. In both the kinetic studies using infectious viruses and the analysis using luciferase expression vectors, the 0.3-kb fragment of A8 was shown to produce higher amount of spliced-mRNAs as well as protein expression compared to its counterpart the 57 0.3-kb fragment. The 0.3-kb fragment influenced the protein expression level from spliced-mRNA by regulating the efficiency of splicing, rather than transcription, poly (A) addition to mRNA, or nuclear export of spliced-mRNA. Furthermore, seven nucleotides that apparently contribute to regulation of gene expression have been identified. Interestingly, these nucleotides reside within the stem-loop structure that has been speculated to limit recognition of the 5'ss.

#### 5. References

- Coffin JM, Hughes SH, Varmus HE. 1997. Retroviruses. Cold Spring Harbor Laboratory Press, New York.
- Delgado E, Carrera C, Nebreda P, Fernandez-Garcia A, Pinilla M, Garcia V, Perez-Alvarez L, Thomson MM. 2012. Identification of new splice sites used for generation of rev transcripts in human immunodeficiency virus type 1 subtype C primary isolates. PLoS One 7:e30574.
- 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.
- Takase-Yoden S, Wada M, Watanabe R. 2006. A viral non-coding region determining neuropathogenicity of murine leukemia virus. Microbiol Immunol 50:197-201.
- Watanabe R, Takase-Yoden S. 2006. Neuropathology induced by infection with Friend murine leukemia viral clone A8-V depends upon the level of viral antigen expression. Neuropathology 26:188-195.
- Mougel M, Tounekti N, Darlix JL, Paoletti J, Ehresmann B, Ehresmann C. 1993. Conformational analysis of the 5' leader and the gag initiation site of Mo-MuLV RNA and allosteric transitions induced by dimerization. Nucleic Acids Res 21:4677-4684.
- Kraunus J, Zychlinski D, Heise T, Galla M, Bohne J, Baum C. 2006. Murine leukemia virus regulates alternative splicing through sequences upstream of the 5' splice site. Journal of Biological Chemistry 281:37381-37390.