RNA Mapping of Mutant Myotonic Dystrophy Protein Kinase 3'-Untranslated Region Transcripts

Min-Sun Song and Seong-Wook Lee*

Department of Molecular Biology, BK21 Graduate Program for RNA Biology, Institute of Nanosensor and Biotechnology, Dankook University, Yongin 448-701, Korea

Abstract

Myotonic dystrophy type 1 (DM1), which is a dominantly inherited neurodegenerative disorder, results from a CTG trinucleotide repeat expansion in the 3'-untranslated region (3'-UTR) of the myotonic dystrophy protein kinase (DMPK) gene. Retention of mutant DMPK (mDMPK) transcripts in the nuclei of affected cells has been known to be the main cause of pathogenesis of the disease. Thus, reducing the RNA toxicity through elimination of the mutant RNA has been suggested as one therapeutic strategy against DM1. In this study, we suggested RNA replacement with a *trans*-splicing ribozyme as an alternate genetic therapeutic approach for amelioration of DM1. To this end, we identified the regions of mDMPK 3'-UTR RNA that were accessible to ribozymes by using an RNA mapping strategy based on a *trans*-splicing ribozyme library. We found that particularly accessible sites were present not only upstream but also downstream of the expanded repeat sequence. Repair or replacement of the mDMPK transcript with the specific ribozyme will be useful for DM1 treatment through reduction of toxic mutant transcripts and simultaneously restore wild-type DMPK or release nucleus-entrapped mDMPK transcripts to the cytoplasm.

Keywords: group I intron, myotonic dystrophy protein kinase, RNA mapping, RNA replacement, *trans*-splicing ribozyme, trinucleotide repeat expansion

Introduction

DM1 is an autosomal dominant genetic and the most common neuromuscular disease in adults, affecting approximately 5 to 20 out of 100,000 among the worldwide population including Asians (Harper, 2001). The DM mutation is a CTG trinucleotide expansion in the 3'-UTR of the DMPK gene on chromosome 19q (Brook et al., 1992). The severity of disease is proportional to the number of the expanded CTG repeats. The number of repeats is in the range of 5-35 in the normal population, and DM alleles harbor more than 50 and up to several thousand repeats in DM1 patients. The pathogenesis of DM1 is directly mediated by a toxic mDMPK RNA that is entrapped in the affected cell nucleus. A number of pieces of evidence suggest that pleiotropic effects of aberrant interactions between the mDMPK transcripts that are retained in the nucleus and RNA-binding proteins, such as muscleblind-like 1, change the metabolism of target mRNAs, including splicing (Wheeler, 2008). Additively, decreased DMPK protein levels in DM1 patients due to loss of cytoplasm-localized mRNA for translation could cause physiological and functional alterations in phosphorylation of target proteins, including CUG-binding protein 1, and hence has been suggested as a contributor to the pathology of the disease (Kaliman and Llagostera, 2008).

There is no cure for DM yet. However, several approaches to reduce the mutant RNA-mediated toxicity through elimination of nucleus-trapped toxic mDMPK transcripts have been proposed as specific therapies for reversal of the disease phenotype. They include approaches with antisense, small catalytic RNA, or siRNA, which can reduce the level of mDMPK RNA by directly targeting the expanded CUG repeats of the mDMPK transcript (Furling et al., 2003; Krol et al., 2007; Langlois et al., 2003; Mahadevan et al., 2006). However, high specificity for degradation of the mutant alleles, non-toxicity, and lack of off-target effects will be required to minimize nonspecific reduction of the normal allele and cellular toxicity. RNA repair or RNA replacement has been described as a new approach to human gene therapy against genetic disorders, which relies on the *Tetrahymena* group I-based *trans*-splicing ribozyme (Sullenger and Gilboa, 2002). This ribozyme could cleave a target 5'-exon RNA and *trans*-ligate an exon that is tagged at its 3'-end onto the cleaved target RNA both *in vitro* and in bacteria and mammalian cells (Been and Cech, 1986; Sullenger and Cech, 1994; Jones et al., 2006). Therefore, the *trans*-splicing ribozyme can reduce the level of a specific target RNA and simultaneously induce transgene activity specifically in the target RNA-expressing cells. Indeed, *trans*-splicing ribozymes have been successfully developed to replace mutant transcripts that are associated with several human genetically inherited diseases, including DM1 (Lan et al.,

1998; Phylactou *et al.* 1998; Rogers *et al.,* 2002). Moreover, we have shown that *trans*-splicing ribozyme targeting and replacing the human telomerase reverse transcriptase (hTERT) RNA could specifically and efficiently regress human tumors not only in cell cultures but also in animals without any toxicity in normal tissues or any off-target splicing effects (Hong *et al.,* 2008; Jeong *et al.,* 2008; Kwon *et al.,* 2005; Song and Lee, 2008; Song *et al.,* 2009). Thus, a *trans*-splicing ribozyme that can repair toxic mDMPK RNA with the wild-type one or replace the nucleus-entrapped mutant transcripts with RNA that can be released into the cytoplasm would be a potential tool for DM1 treatment. To this end, we have developed mDMPK 3'-UTR-specific *trans*-splicing ribozymes by utilizing an RNA mapping procedure that identifies the most accessible ribozyme sites.

## Mapping of the trans-splicing ribozyme

Any uridine (U) residue in a target RNA can be targeted for splicing by the *trans*-splicing ribozyme through alteration of the internal guide sequence (IGS), which is the 5'-exon binding site of the ribozyme, to make it complementary to the target sequence (Long *et al.,* 2003; Sullenger, 1995). However, the ribozyme can actually access only a limited number of uridines on the target RNA due to complicated tertiary structure formation of the substrate RNA (Lan *et al.,* 2000). To develop most potential mDMPK 3'-UTR-specific ribozymes, therefore, identification of the regions of the target transcript that are accessible to ribozymes should

be performed first. To this end, we employed an RNA mapping strategy based on RNA tagging (Jones *et al.,*1996) and a *trans*- splicing ribozyme library, which has been described in detail elsewhere (Fig. 1) (Lan *et al.,* 1998; Lan *et al.,* 2000; Kwon *et al.,* 2005; Ryu *et al.,* 2003; Song and Lee, 2007).

We constructed the GN5 ribozyme library, which is based on the *Tetrahymena* group I intron and contains a randomized IGS. Ribozymes in the GN5 library can bind to the substrate RNA, cleave at any accessible U residue, and transfer a 3'-exon that is tagged onto the ribozyme to the end of the cleaved 5'-target product. We used a portion of the *lacZ* gene as a 3'-exon molecular tag in the GN5 library. mDMPK 3'-UTR RNA was synthesized using *in vitro* transcription of PCR-amplified vector encoding the mDMPK 3'-UTR gene with 200 CTG repeats (kindly supplied by Dr. Leonidas Phylactou at the Cyprus Institute of Neurology and Genetics). In order to map the mDMPK 3'-UTR RNA, GN5 library RNA (10 nM) was incubated at 37°C for 3 hrs under
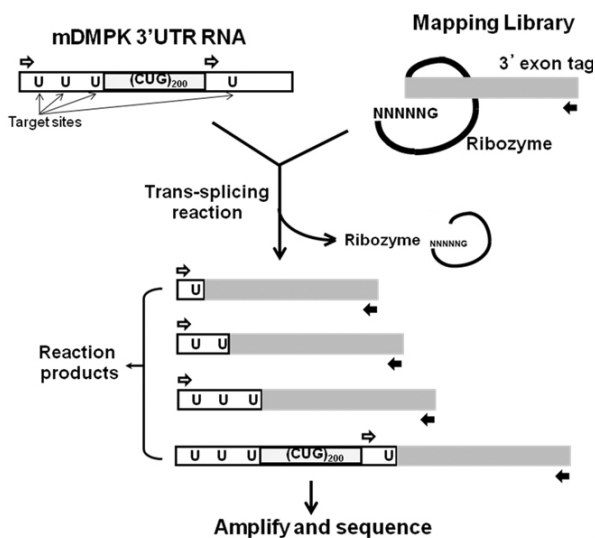
Fig. 1. RNA mapping procedure for the identification of accessible sites in mDMPK 3'-UTR RNA using a *trans*-splicing ribozyme library.
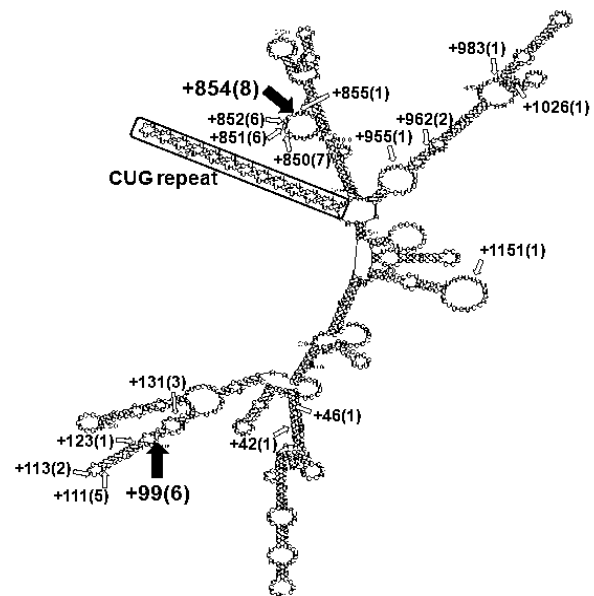

Fig. 2. The structure of the ribozyme-accessible sites in mDMPK 3'-UTR RNA. Nucleotide positions of the accessible uridines identified from the mapping study are indicated by arrows and nucleotide numbers in the secondary structure of mDMPK 3'-UTR RNA, as predicted by the mfold computer simulating program (Zuker, 2003). The number of clones that harbored a given uridine at the splice site is shown in parentheses. Repeated sequences of CUG, which are represented in a box, form a stable hairpin structure. For brevity, 25 repeats of the CUG sequence were drawn in this figure. The extended 200 CUG repeat itself also formed a long hairpin stem configuration.

splicing conditions (50 mM HEPES, pH 7.0, 150 mM NaCl, 5 mM MgCl₂) in the presence of guanosine (200 μM) and *in vitro*-generated target RNA (100 nM). The *trans*-splicing reaction products were then reverse-transcribed at 37°C for 30 min in the presence of argininamide (10 mM) with a 3'primer that was specific for the ribozyme's 3'-exon *lacZ* sequence (5'-ATGTGCTGCAAG GCGATT-3') (Jones *et al.* 1996). Produced cDNAs were amplified by PCR for 30 cycles using the same 3'primer and a 5'primer that was specific for the 5'-end of the target RNA (5'-ACCCTAGAACTGTCTTCGACTCC-3') or downstream region of the expanded CTG sequence (5'-GGAATTCGGGGGATCACAGACCAT-3'). The amplified *trans*- splicing products were then cloned into the pUC19 vector, and sequences of several clones were determined.

Ribozyme-accessible uridines in mDMPK 3'-UTR RNA

When we analyzed sequences of the splicing junction sites of the *trans*-spliced products amplified with the 5'-end of the target RNA as the 5' primer, several uridines, only in the upstream of the expanded trinucleotide repeat sequence, appeared to be particularly accessible (Fig. 2). Long hairpin structures of the expanded trinucleotide repeat might hamper PCR amplification of

trans-spliced products that result from the splicing reaction with U residues downstream of the trinucleotide repeat sequence. To identify accessible U residues in the target RNA that were present downstream of the trinucleotide repeat, we also amplified *trans*-splicing products with the 5'primer that was specific for right downstream part of the trinucleotide repeat sequence. Then, several uridines downstream of the expanded CUG repeat sequence were also shown to be particularly accessible. In conclusion, the most accessible site upstream of the CUG repeat sequence was located in the uridine at position 99 of the mDMPK 3'-UTR RNA. In addition, the uridine at position 854 of the target RNA was the most accessible site that was present downstream of the CUG repeat. Interestingly, most of the accessible sites, including positions 99 and 854, which were identified using this mapping method, were generally present in the loop or bulge region of the RNA's secondary structure, as predicted by a computer program (Fig. 2). In contrast, no accessible sites were found in the long hairpin-formed CUG repeat region. In previous studies, we showed that the relative *trans*-splicing reaction efficacy at different sites in a target RNA in cells and *in vitro* was closely associated with the predicted accessibility determined by mapping results (Ryu *et al.*, 2003; Ryu and Lee, 2004; Kwon *et al.*, 2005). Therefore, the ribozymes that target the ac-
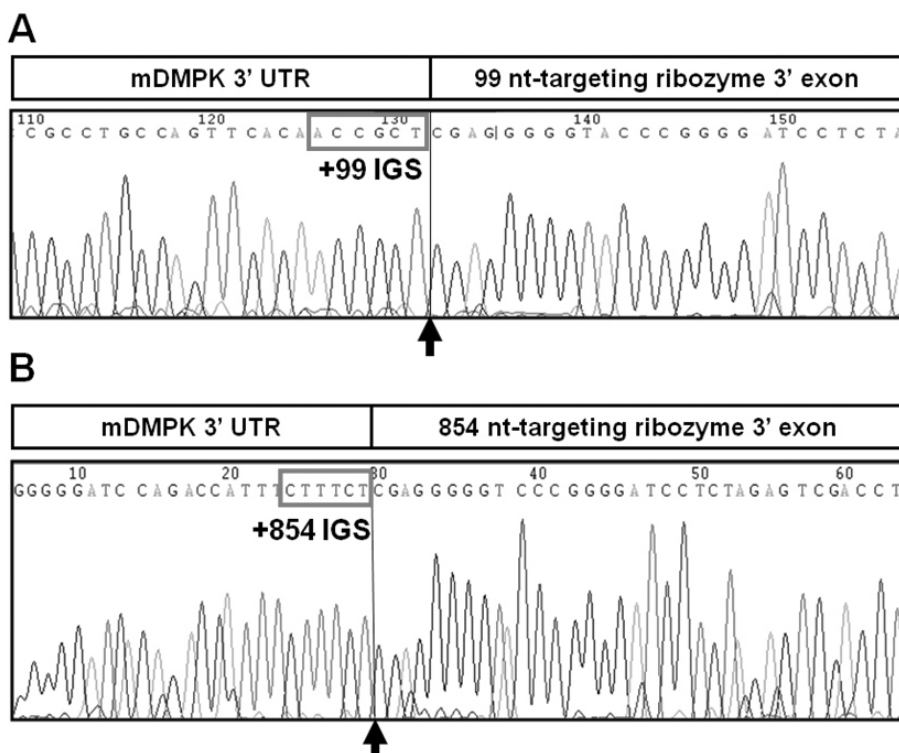
Fig. 3. A representative sequence of the *trans*-spliced transcripts of Rib99 (A) or Rib854 (B). The expected sequence around the splicing junction is shown with an arrow, and the ribozyme recognition sequence in mDMPK 3'-UTR RNA is indicated within a box.

cessible sites identified here are anticipated to be more active than those that target other sites.

A specific ribozyme that targeted position 99 or 854 of the mDMPK 3'-UTR RNA (Rib99 or Rib854, respectively) was constructed by *in vitro* transcription of cDNA template. We made cDNA template by PCR of the pT7L-21 plasmid with a 5' primer that contained the T7 promoter and each ribozyme's IGS, and a 3' primer that was specific for the 3'-exon sequence, as described (Sullenger and Cech, 1994; Jones *et al.*, 1996). The pT7L-21 plasmid encodes a slightly shortened version of the native group I intron from *Tetrahymena*. The IGS of the L-21 *trans*-splicing ribozyme (5'-GGAGGG-3') was exchanged with 5'-GGCGGT-3' in Rib99 or with 5'-GGAAAG-3' in Rib854. The specific ribozyme (100 nM) was then incubated with the substrate RNA (10 nM) under splicing conditions. RT-PCR analyses were performed using a 3' primer that was specific for the 3'-exon tag sequence and a 5' primer that was specific for the mDMPK 3'-UTR RNA. The amplified fragment could be seen at the expected size of 197 bp or 130 bp when incubating the target RNA with Rib99 or Rib854, respectively (data not shown). Sequence analyses of the amplified *trans*-spliced products revealed that both Rib99 (targeting the upstream region of the expanded CUG repeat) and Rib854 (targeting the downstream site of the expanded CUG repeat) accurately employed the *trans*-splicing reaction of their 3'-exon tag onto the target RNA at the predicted reaction site (Fig. 3). These studies indicate that specific ribozyme targeting of predicted accessible sites can replace mDMPK 3'-UTR RNA with a 3'-exon that is tagged at the 3'-end of the ribozyme with high fidelity through targeted *trans*-splicing.

Recently, several gene therapy protocols, based on inhibitory RNA or short oligonucleotides that target the toxic mDMPK transcript, have been described for DM1 therapy (Wheeler, 2008). For the successful application of the RNA- or oligonucleotide-mediated strategies to clinical DM1 treatment, specificity, safety, and efficacy, in addition to delivery issues, are important criteria to be considered. In this study, we developed *trans*-splicing ribozymes that can potentially repair the mutant DMPK transcript with the wild-type one or replace the mutant RNA with a therapeutically relevant transcript. In previous studies, we have shown that hTERT-specific *trans*-splicing ribozymes could specifically and effectively target and replace the hTERT RNA *in vivo* as well as in tumor cells without toxic effects on normal tissues and without nonspecific off-target splicing to other RNAs besides the target transcript (Hong *et al.*, 2008; Jeong *et al.*, 2008; Song *et al.*, 2009). The *trans*-splicing ribozyme that targets upstream of the expanded CUG

repeats (Rib99) that was developed in this study could be used as a repair tool to restore the wild-type DMPK protein. A major advantage of the RNA repair approach with a *trans*-splicing ribozyme for DM1 therapy is that the ribozyme could restore and maintain the regulated expression of normal DMPK protein and that it could simultaneously reduce the production of toxic mutant DMPK RNA. However, targeting sites in the mDMPK 3'-UTR are also present in the wild-type DMPK RNA. Therefore, one limitation of the RNA repair approach is the possibility of targeting normal alleles. Although targeting and replacing normal DMPK RNA with a *trans*-splicing ribozyme will not lead to reduction of normal protein, the efficacy of repair of mutant transcript can decrease. Therefore, high-efficiency targeting to the target RNA will be a critical factor for a successful therapeutic outcome. Another developed ribozyme that targets downstream of the expanded CUG repeats (Rib 854) could be useful to replace the toxic RNA that is entrapped in the nucleus with RNA that can be released into the cytoplasm. Recent findings showed that mDMPK 3'-UTR that is aberrantly retained in the nucleus could be transported extensively to the cytoplasm when it has attached to its 3'-end the woodchuck posttranscriptional regulatory element (WPRE), which can naturally accumulate hepatitis transcripts in the cytoplasm, resulting in an increase in myoblast differentiation whose initiation is failed in mutant 3'-UTR-expressing cells (Mastroyiannopoulos *et al.*, 2005 and 2008). Therefore, RNA replacement of the mDMPK 3'-UTR RNA with RNA with WPRE could stimulate cytoplasmic release of the mDMPK transcript, relieving the pathogenic effects that are induced by the toxic RNA.

Another group also suggested that the *trans*-splicing ribozyme approach may form a therapeutic strategy for diseases that are associated with repeat expansions, including DM1 (Phylactou *et al.*, 1998). However, the authors constructed and characterized a ribozyme that targeted a randomized selected site. An essential factor that influences the efficacy of RNA-mediated gene therapeutic approaches is the accessibility of inhibitory elements for binding to the target RNA. RNA mapping studies, which were performed in this study, therefore, should be preceded to select the most accessible sites in the target RNA, develop the most potent RNA inhibitors, and optimize intracellular anti-mDMPK RNA activity. Noticeably, this study showed that very stable RNAs, such as mDMPK 3'-UTR, were successfully targeted and mapped with the *trans*-splicing ribozyme. The mapping procedure that was performed in this study could be broadly useful for the identification of the most accessible sites in specific disease-associated transcripts to develop therapeutic strategies with RNA in-

hibitors.

Acknowledgments

References

Been, M., and Cech, T. (1986). One binding site determines sequence specificity of Tetrahymena pre-rRNA self-splicing, trans-splicing, and RNA enzyme activity. *Cell* 47, 207-216.

Brook, J.D., McCurrach, M.E., Harley, H.G., Buckler, A.J., Church, D., Aburatani, H., Hunter, K., Stanton, V.P., Thirion, J.P., Hudson, T., Sohn, R., Zemelman, B., Snell, R.G., Rundle, S.A., Crow, S., Davies, J., Shelbourne, P., Buxton, J., Jones, C., Juvonen, V., Johnson, K., Harper, P.S., Shaw, D.J., and Housman, D.J. (1992). Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* 68, 799-808.

Furling, D., Doucet, G., Langlois, M.A., Timchenko, L., Belanger, E., Cossette, L., and Puymirat, J. (2003). Viral vector producing antisense RNA restores myotonic dystrophy myoblast functions. *Gene Ther.* 10, 795-802.

Harper, P.S. (2001). Myotinic dystrophy. 3rd ed. London: W.B. Saunders.

Hong, S.H., Jeong, J.S., Lee, Y.J., Jung, H.I., Cho, K.S., Kim, C.M., Kwon, B.S., Sullenger, B.A., Lee, S.W., and Kim, I.H. (2008). *In vivo* reprogramming of hTERT by trans-splicing ribozyme to target tumor cell. *Mol. Ther.* 16, 74-80.

Jeong, J.S., Lee, S.W., Hong, S.H., Lee, Y.J., Jung, H.I., Cho, K.S., Seo, H.H., Lee, S.J., Park, S., Song, M.S., Kim, C.M., and Kim, I.H. (2008). Antitumor effects of systemically delivered adenovirus harboring *trans*-splicing ribozyme in intrahepatic colon cancer mouse model. Clin. *Cancer Res.* 14, 281-290.

Jones, J.T., Lee, S.W., and Sullenger, B.A. (1996). Tagging ribozyme reaction sites to follow *trans*-splicing in mammalian cells. *Nature Med.* 2, 643-648.

Kaliman, P., and Llagostera, P. (2008). Myotonic dystrophy protein kinase (DMPK) and its role in the pathogenesis of myotonic dystrophy 1. *Cell. Signal.* 20, 1935-1941

Krol, J., Fiszer, A., Mykowska, A., Sobczak, K., de Mezer, M., and Krzyzosiak, W.J. (2007). Ribonuclease dicer cleaves triplet repeat hairpins into shorter repeats that silence specific targets. *Mol. Cell.* 25, 575-586.

Kwon, B.S., Jung, H.S., Song, M.S., Cho, K.S., Kim, S.C., Kimm, K., Jeong, J.S., Kim, I.H., and Lee, S.W. (2005). Specific regression of human cancer cells by ribozyme-mediated targeted replacement of tumor-specific transcript. *Mol. Ther.* 12, 824-834.

Lan, N., Howrey, R.P., Lee, S.W., Smith, C.A., and Sullenger, B.A. (1998). Ribozyme-mediated repair of sickle β-globin mRNAs in erythrocyte precursors. *Science* 280, 1593-1596.

Lan, N., Rooney, B.L., Lee, S.W., Howrey, R.P., Smith, C.A., and Sullenger, B.A. (2000). Enhancing RNA repair efficiency by combining *trans*-splicing ribozymes that recognize different accessible sites on a target RNA. *Mol. Ther.* 2, 245-255.

Langlois, M.A., Lee, N.S., Rossi, J.J., and Puymirat, J. (2003). Hammerhead ribozyme-mediated destruction of nuclear foci in myotonic dystrophy myoblasts. *Mol. Ther.* 7, 670-680.

Long, M.B., Jones, J.P. 3rd, Sullenger, B.A., and Byun, J. (2003). Ribozyme-mediated revision of RNA and DNA. *J. Clin. Invest.* 112, 312-318.

Mahadevan, M.S., Yadava, R.S., Yu, Q., Balijepalli, S., Frenzel-McCardell, C.D., Bourne, T.D., and Phillips, L.H. (2006). Reversible model of RNA toxicity and cardiac conduction defects in myotonic dystrophy. *Nat. Genet.* 38, 1066-1070.

Mastroyiannopoulos, N.P., Chrysanthou, E., Kyriakides, T.C., Uney, J.B., Mahadevan, M.S., and Phylactou, L.A. (2008). The effect of myotonic dystrophy transcript levels and location on muscle differentiation. *Biochem Biophys. Res. Commun.* 377, 526-531

Mastroyiannopoulos, N.P., Feldman, M.L., Uney, J.B., Mahadevan, M.S., and Leonidas A.P. (2005). Woodchuck post-transcriptional element induces nuclear export of myotonic dystrophy 3' untranslated region transcripts. *EMBO Rep.* 6, 458-463.

Phylactou, L.A., Darrah, C., and Wood, M.A.J. (1998). Ribozyme-mediated *trans*-splicing of a trinucleotide repeat. *Nat. Genet.* 18, 378-381.

Rogers, C.S., Vanoye, C.G., Sullenger, B.A., and George, A.L. (2002). Functional repair of a mutant chloride channel using a *trans*-splicing ribozyme. *J. Clin. Invest.* 110, 1783-1798.

Ryu, K.J., and Lee, S.W. (2004). Comparative analysis of intracellular *trans*-splicing ribozyme activity against hepatitis C virus internal ribosome entry site. *J. Microbiol.* 42, 361-364.

Ryu, K.J., Kim, J.H., and Lee, S.W. (2003). Ribozyme-mediated selective induction of new gene activity in hepatitis C virus internal ribosome entry site-expressing cells by targeted *trans*-splicing. *Mol. Ther.* 7, 386-395.

Song, M.S., and Lee, S.W. (2007). An RNA mapping strategy to identify ribozyme-accessible sites on the catalytic subunit of mouse telomerase. *Genom. Inform.* 5, 32-35.

Song, M.S., and Lee, S.W. (2008). *In vivo* target RNA specificity of *trans*-splicing phenomena by the group I intron. *Genom. Inform.* 6, 84-86.

Song, M.S., Jeong, J.S., Ban, G. Lee, J.H., Won, Y.S., Cho, K.S., Kim, I.H., and Lee, S.W. (2009). Validation of tissue-specific promoter-driven tumor-targeting *trans*-splicing ribozyme system as a multifunctional cancer gene therapy device *in vivo*. *Cancer Gene Ther.* 16, 113-125.

Sullenger, B.A. (1995). Revising messages traveling along the cellular information highway. *Chem. Biol.* 2, 249-253.

Sullenger, B.A., and Cech, T.R. (1994). Ribozyme-mediated repair of defective mRNA by targeted *trans*-splicing. *Nature* 317, 619-622.

Sullenger, B.A., and Gilboa, E. (2002). Emerging clinical ap-

plications of RNA. *Nature* 418, 252-258.

Wheeler, T.M. (2008). Myotonic dystrophy: therapeutic strategies for the future. *Neurotherapeutics* 5, 592-600.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucl. Acids Res.* 31, 3406-3415.