

Selection of the Most Potent Specific On/Off Adaptor-Hepatitis Delta Virus Ribozymes for Use in Gene Targeting

Michel V. Lévesque, Samuel G. Rouleau, and Jean-Pierre Perreault

The Hepatitis Delta Virus (HDV) ribozyme, which is well adapted to the environment of the human cell, is an excellent candidate for the future development of gene-inactivation systems. On top of this, a new generation of HDV ribozymes now exists that benefits from the addition of a specific on/off adaptor (specifically the SOFA-HDV ribozymes) which greatly increases both the ribozyme's specificity and its cleavage activity. Unlike RNAi and hammerhead ribozymes, the designing of SOFA-HDV ribozymes to cleave, *in trans*, given RNA species has never been the object of a systematic optimization study, even with their recent use for the gene knockdown of various targets. This report aims at both improving and clarifying the design process of SOFA-HDV ribozymes. Both the ribozyme and the targeted RNA substrate were analyzed in order to provide new criteria that are useful in the selection of the most potent SOFA-HDV ribozymes. The crucial features present in both the ribozyme's biosensor and blocker, as well as at the target site, were identified and characterized. Simple rules were derived and tested using hepatitis C virus NS5B RNA as a model target. Overall, this method should promote the use of the SOFA-HDV ribozymes in a plethora of applications in both functional genomics and gene therapy.

Introduction

NUCLEIC ACID-BASED GENE inactivation systems present a promising avenue for new drug development. To date, most of the efforts in this domain involve technologies that are based on RNA interference mechanism, even though these mechanisms have been demonstrated to both exhibit off-target effects and to trigger immunological responses (Haasnoot et al., 2007; Robbins et al., 2008; Tamura and Nagasaki, 2010). An alternative to this approach is to use nucleic acid molecules, such as either ribozymes (RNA enzyme, Rz) or DNazymes (DNA enzyme, Dz) (D'Anjou et al., 2004; Peracchi, 2004; Schubert and Kurreck, 2004), that possess their own catalytic activities. Several small self-cleaving RNA motifs (ex. the *hammerhead* ribozymes) have been transformed into *trans*-acting systems in which 1 molecule, the ribozyme, possesses the ability to recognize and, subsequently, cleave a second molecule, the substrate (Bergeron and Perreault, 2002; D'Anjou et al., 2004; Muller-Kuller et al., 2009; Tedeschi et al., 2009). Among these catalytic RNAs, the Hepatitis Delta Virus ribozyme (HDV Rz) has, for a long time, been the unique example of a ribozyme derived from an RNA species that is naturally found in human cells (ie, infected hepatocytes) (Asif-Ullah et al., 2007). Consequently, the HDV ribozyme possesses several properties desirable for a gene

inactivation tool, including both its natural ability to function in the presence of human proteins and at the physiological magnesium concentration (1 mM Mg²⁺). Thus, the HDV ribozyme appears to be well adapted to the human cell environment, as is demonstrated by its outstanding stability *in cellulo* (Levesque et al., 2002). However, its use in gene-knockdown systems has been relatively neglected due to a lack of substrate specificity. The original HDV ribozyme folds into a double-pseudoknot secondary structure in which the substrate recognition is limited to the formation of a 7-base pair (bp) helix (ie, the stem I) called the recognition domain (Fig. 1). To overcome this hurdle, molecular engineering was used to develop a target-dependent activation module named the specific on/off adapter (SOFA) (Bergeron et al., 2005; Bergeron and Perreault, 2005). The SOFA module has the ability to switch the ribozyme from an inactive (*off* state) to an active conformation (*on* state) solely in the presence of the appropriate substrate (Fig. 1). More specifically, the SOFA module includes a blocker (BI) that maintains the ribozyme in an inactive conformation through the formation of a short helix with the recognition domain. This interaction competes with the binding of any substrate that could be potentially bound by the recognition domain. The SOFA module also includes a biosensor (Bs) that extends the base pairing with the target. The formation of a duplex between the biosensor

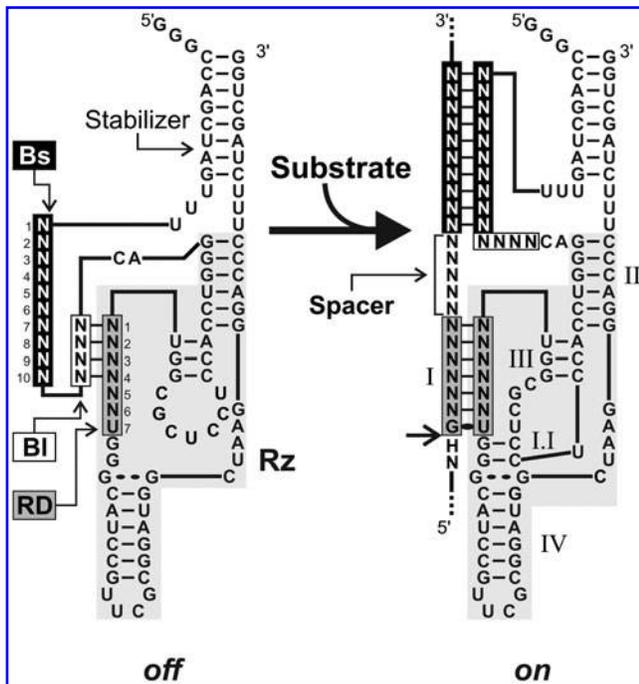


FIG. 1. Secondary structure representations of the specific on/off adaptor-hepatitis delta virus ribozyme (SOFA-HDV Rz) in both the *off* (left) and *on* (right) conformations. The HDV ribozyme is highlighted in gray in both structures, whereas the cognate substrate is drawn only for the *on* conformation, as it is the modulator of the conformational switch. The 3 key variable features of the ribozyme, the recognition domain, the biosensor, and the blocker, are identified by RD (black on gray), Bs (white on black), and BI (black on white), respectively. The stabilizer and the spacer (the short sequence in the target that is located between the recognition domain and the biosensor binding site) are also indicated. The roman numerals I, II, III, and IV identify the corresponding stems or stem loops in the HDV ribozyme, whereas the arabic numbers indicate the positions of each nucleotide in the Bs and the RD. The variable nucleotides (ie, those that can be A, C, G, or U) of the ribozyme are indicated by the boxed letter N. All invariable nucleotides are unboxed and are constant from 1 ribozyme to the other, regardless of the target sequence. The short arrow indicates the cleavage site.

and the target RNA concomitantly causes the disruption of the blocker-recognition domain duplex, thereby causing the switch to the active conformation. Essentially, the blocker acts as a safety lock for the ribozyme in which the key to opening it is the recognition of the target RNA by the biosensor. Finally, the SOFA module also includes a stabilizer that consists of an additional stem which brings both ribozyme's 5'- and 3'-strand ends together in order to mimic the stem II of the HDV ribozyme, a feature which has been demonstrated to be a key reason for its outstanding stability (Levesque et al., 2002). Importantly, the presence of the stabilizer does not affect the cleavage activity of the SOFA-HDV Rz.

The arrangement of the SOFA module greatly increases both the ribozyme's specificity and its cleavage activity in terms of its potential applications in the fields of both gene therapy and functional genomics (Bergeron et al., 2005; Bergeron and Perreault, 2005). Several examples of the development of gene inactivation systems in both prokaryotic and

eukaryotic cells have been reported (Fiola et al., 2006; Robichaud et al., 2008; Levesque et al., 2010; D'Anjou et al., 2011; Laine et al., 2011; Levesque and Perreault, 2011, in press). In the course of these studies, more than 100 distinct SOFA-HDV ribozymes were designed and tested *in vitro* to identify the best ones with which to target various RNA substrates. The lesson learned from these efforts was that the ribozymes exhibited cleavage activities that ranged from 0% to almost 100%, thus suggesting that the design process was flawed in some way. That said, these SOFA-HDV ribozymes formed a relatively large library of catalytic RNAs for which the cleavage activity data were known. Analysis of the cleavage activity as a function of the nucleotide composition led to the formulation of hypotheses identifying the determinants that are important for the development of efficient ribozymes. Here, several of these hypotheses are verified to establish new criteria for improving the SOFA-HDV ribozyme design process. These experiments support the importance of a critical selection of the ribozymes before any *in vitro* testing. Moreover, this work led to the elucidation of important features present in both the biosensor and blocker domains, as well as in the target site, that could be detrimental to cleavage activity.

Materials and Methods

SOFA-HDV ribozyme and substrate construction

The DNA templates of both the SOFA-HDV ribozymes and their substrates were generated by a polymerase chain reaction (PCR)-based strategy using different pairs of complementary oligonucleotides as previously described (Levesque et al., 2010; Levesque and Perreault, 2011, in press). Briefly, 2 DNA oligonucleotides were used for the production of the SOFA-HDV ribozymes: the universal reverse primer (5'-CCAGC TAGAAAGGGTCCCTTAGCCATCCGCGAACGGATGCC C-3') and the SOFA-HDV RzX (where X identifies the specific ribozyme) sense primer [5'-TAATACGACTCACTATAGGG CCAGCTAGTTT(N)_{10Bs}(N)_{4BI}CAGGGTCCACCTCCTCGCG GT(N)_{6RD}TGGGCATCCGTTCGCGG-3', where N represents A, C, G, or T and Bs, BI, and RD indicate the biosensor, the blocker sequence, and the recognition domain, respectively]. The SOFA-HDV RzX sense primer is specific for each ribozyme, and it enables the incorporation of the T7 RNA polymerase promoter, whereas the universal reverse primer was used with all ribozymes. The 5' to 3' elongation of the DNA sequence that produced a double-stranded DNA template was performed using *Pwo* DNA polymerase (Roche Diagnostics). Similarly, the substrate DNA template was produced by a combination of 2 complementary oligonucleotides, 1 of which contained the T7 RNA polymerase promoter at its 5'-end. All PCR reactions were ethanol precipitated before *in vitro* transcription.

RNA synthesis

Both the SOFA-HDV ribozymes and their corresponding substrates were synthesized by run-off transcriptions as previously described (Levesque et al., 2010). Specifically, transcriptions were performed in the presence of purified T7 RNA polymerase (10 µg), pyrophosphatase (0.01 U, Roche Diagnostics), and PCR product (2–5 µM) in a buffer containing 80 mM HEPES-KOH (pH 7.5), 24 mM MgCl₂, 2 mM spermidine, 40 mM dithiothreitol, and 5 mM of each ribonucleotide

triphosphates (rNTP) in a final volume of 100 μ L at 37°C for 2 hours. After completion, the reaction mixtures were treated with RQ1 DNase (Promega) at 37°C for 20 minutes. The RNAs were then purified by phenol/chloroform extraction and ethanol precipitation. The resulting pellets were dissolved in equal volumes of ultrapure water and loading buffer [95% formamide, 10 mM ethylenediaminetetraacetic acid (EDTA) (pH 8.0), 0.025% xylene cyanol, and 0.025% bromophenol blue]. The samples were then fractionated through either 8% or 20% denaturing polyacrylamide gels (PAGE, 19:1 ratio of acrylamide to bisacrylamide) in a buffer containing 45 mM Tris-borate (pH 7.5), 8 M urea, and 2 mM EDTA. The RNA products were visualized by ultraviolet shadowing. The bands corresponding to the correct sizes for both the SOFA-HDV ribozymes and their corresponding substrates were cut out of the gel, and the transcripts were then eluted overnight at 4°C in elution buffer (500 mM ammonium acetate, 10 mM EDTA, and 0.1% sodium dodecyl sulfate). The eluted transcripts were then ethanol precipitated, washed, dried, and dissolved in ultrapure water. The RNA was quantified by absorbance at 260 nm. For the self-cleavage experiments, 1 μ L of [α -³²P]UTP (3,000 Ci/mmol; New England Nuclear) was added to each *in vitro* transcription reaction. The transcripts were purified (phenol/chloroform extraction and ethanol precipitation) as just described, and samples corresponding to 1% of each reaction were fractionated through 10% denaturing PAGE and then exposed to a Phosphor Screen to visualize the results (Molecular Dynamics).

RNA substrate labeling

The RNA substrates used in the different cleavage reactions were 5'-end labeled as previously described (Levesque et al., 2010). Briefly, the purified transcripts were dephosphorylated by mixing 50 pmol of RNA with 1 U of Antarctic phosphatase (New England Biolabs) in a final volume of 10 μ L containing the buffer provided with the enzyme, and they were then incubated for 30 minutes at 37°C. The enzyme was then inactivated by incubation at 65°C for 8 minutes. The dephosphorylated RNAs (5 pmol) were then 5'-end labeled by incubation for 1 hour at 37°C with 3 U of T4 polynucleotide kinase (USB Corp.) and 3.2 pmol of [γ -³²P]ATP (6,000 Ci/mmol; New England Nuclear) in the reaction buffer provided with the enzyme. The reactions were stopped by the addition of 2 volumes of loading buffer. The 5'-end-labeled RNA substrates were fractionated by 20% denaturing PAGE. The RNAs were detected by autoradiography, cut out of the gel, and eluted as just described.

In vitro cleavage assays

The *in vitro* cleavage assays were performed under single turnover conditions in which a trace amount of 5'-end-labeled substrate (<1 nM) was incubated at 37°C with a final concentration of 100 nM of SOFA-HDV ribozyme. The reactions were performed in a total volume of 30 μ L. For each experimental condition, 3 μ L of the proper 5'-end-labeled substrate at a concentration less than 10 nM, 3 μ L of 1 μ M SOFA-HDV ribozyme, 3 μ L of cleavage reaction buffer (500 mM Tris-HCl, pH 7.5), and 18 μ L of RNase-free water were mixed and preincubated at 37°C for 25 minutes. The cleavage reactions were initiated by the addition of 3 μ L of 100 mM MgCl₂. At different time intervals, 2 μ L aliquots were transferred in a new tube containing 10 μ L of loading buffer to stop the reac-

tion. All aliquots were then fractionated on a 20% denaturing PAGE under the conditions just described. The results were visualized using a Phosphor Screen and were quantified using the ImageQuant software (Molecular Dynamics). A control reaction in which the ribozyme's RNA was replaced by water was performed for each RNA substrate used, and a sample at the last time interval was kept for background subtraction. The cleavage percentage was calculated (cleaved product counts over cleaved plus uncleaved products counts) for each time point, and the k_{obs} was calculated using GraphPad Prism 5. Briefly, the rate of cleavage (k_{obs}) was obtained by fitting the data to the equation $A_t = A_{\infty} (1 - e^{-kt})$, where A_t is the percentage of cleavage a time t , A_{∞} is the maximum percent cleavage, and k is the rate constant (k_{obs}).

Bioinformatics analysis

To estimate the potential interaction between the biosensor and the stem-loop III, the software RNAhybrid (<http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/>) was used. The RNAhybrid is a tool for finding the minimum free energy of hybridization (MFE) of 2 RNAs, and it is primarily used for miRNA target prediction. Briefly, the sequence corresponding to the stem-loop III (5'-ACCUCUCGCGGU-3') was considered the target input, whereas the sequences corresponding to the linker (ie, the sequence between the biosensor and the stabilizer stem), the biosensor, the blocker, and the CA nucleotides of junction I/II (5'-UUUN_{Bs10}N_{B14}CA-3') of each ribozyme were taken as the microRNA input in the software. The MFE was used to approximate the stability of the potential interaction in each case, and the proposed duplex served as a template for the drawing of the potential ribozyme's secondary structure presented in Fig. 2. All other sequence analyses were performed using Microsoft Excel and simple logic functions. The sorting of all of the potential cleavage sites within that hepatitis C virus (HCV) NS5B gene (1,773 nucleotides, AJ242654) was performed manually, looking for H₁G₊₁ sites. Each potential site was named for the position number of the corresponding G₊₁. The extracted sequences started 4 nucleotides upstream of each cleavage site (G₊₁), and they finished 22 nucleotides downstream to include all the important components for a target site of the SOFA-HDV ribozyme. The corresponding biosensor, blocker, and recognition domain were designed for each potential site. For the comparison of our 2 groups of ribozymes targeting NS5B, the unpaired t test was performed using GraphPad Prism 5.0.

Results

Previous experiments with the HDV ribozyme demonstrated the existence of a strong relationship between its structure and its cleavage activity (Reymond et al., 2009). Consequently, the same logic will be used here to refine the design process for the identification of the SOFA-HDV ribozymes possessing the greatest potential for targeting a given RNA substrate. Initially, a set of 3 SOFA-HDV Rzs was selected with which to test the various hypotheses and to ensure that the conclusions drawn can be applied to other Rzs. The Rzs in question (ie, SOFA-HDV-Rz-Rev1, SOFA-HDV-Rz-NS527, and SOFA-HDV-Rz-HBV513) are directed against the human immunodeficiency virus, the influenza A, virus and the hepatitis B virus (HBV), respectively. The target sequence of each ribozyme is presented in Table 1.

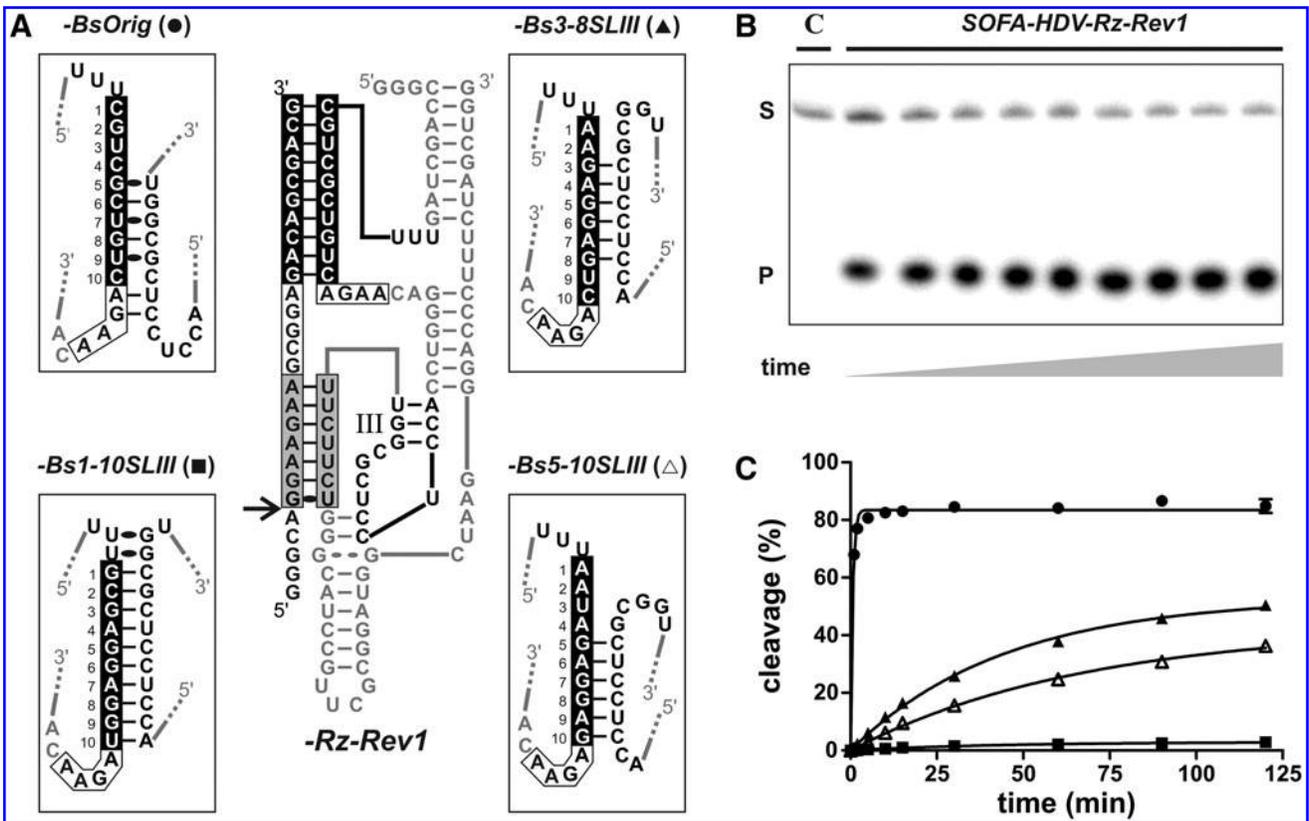


FIG. 2. Analysis of the base-pairing potential between the biosensor and the stem-loop III. (A) Representation of the SOFA-HDV-Rz-Rev1 and 3 biosensor mutants. The original SOFA-HDV Rz designed to target the HIV RNA is presented with its 27nt model substrate. The arrow indicates the cleavage site. The predicted interactions between the biosensor (Bs) and the stem-loop III (SLIII) are shown for the original SOFA-HDV-Rz-Rev1, and for 3 different biosensor mutants. (B) Example of an *in vitro* cleavage assay. The reaction was performed in the presence of trace amounts of 5'-end-labeled substrate, 100 nM of SOFA-HDV-Rz-Rev1, and 10 mM MgCl₂, except for the control condition in which no ribozyme was added (*first lane*). Aliquots were taken at various time points (1, 2, 5, 10, 15, 30, 60, 90, and 120 minutes) and fractionated on a 20% denaturing polyacrylamide gel. The full-length substrate (27 nt) and the 5' cleavage product (5 nt) are identified by the letters S and P, respectively. (C) Graphic representation of the cleavage percentage as a function of time. The quantification of 2 independent experiments is reported for the SOFA-HDV-Rz-Rev1-BsOrig (●), -Bs1-10SLIII (■), -Bs3-8SLIII (▲), and -Bs5-10SLIII (△) ribozymes.

Interaction of the biosensor with the catalytic core

The HDV ribozyme structure contains single-stranded regions that should fold precisely for efficient catalysis of the cleavage reaction to occur, and that can be used to negatively modulate the ribozyme (Reymond et al., 2009). Specifically, both the loop III and the junction IV/II regions are relatively long stretches of unpaired bases (ie, 7 and 5 nucleotides, respectively). It was hypothesized that these regions may interact with the biosensor sequence that is also single stranded in the absence of the substrate. This would impair substrate binding, and, consequently, reduce any potential cleavage by the ribozyme. This hypothesis received indirect support from analysis of the SOFA-HDV ribozyme library which revealed that a few SOFA-HDV ribozymes exhibited a low cleavage activity when significant complementary existed between stem-loop III and the biosensor. To verify this hypothesis, different biosensors that could potentially base pair with the stem-loop III were designed for each of the 3 models of SOFA-HDV Rzs, and their cleavage activities were then assessed for each under single turnover conditions ($[Rz] \gg [S]$). Briefly, a trace amount of 5'-end ³²P-labeled substrate (<1 nM) was

incubated in the presence of 100 nM SOFA-HDV ribozyme in a buffer containing 10 mM MgCl₂ for 25 minutes at 37°C. The reactions were stopped by the addition of loading buffer, and were then fractionated on PAGE. A typical autoradiogram is shown in Fig. 2 for SOFA-HDV-Rz-Rev1. As expected, the ribozyme with its original biosensor is highly active, with more than 80% cleavage being achieved within 5 minutes. Replacement of the biosensor by a sequence that is predicted to form 10 bp with the stem-loop III (-Bs1-10SLIII) causes the activity of the ribozyme to be dramatically reduced. In fact, even after 2 hours of incubation, less than 5% of cleavage was observed. Even when the biosensor sequence was modified to allow the formation of only 6 bp, the cleavage activity was still significantly lower than that of the original SOFA-HDV-Rz-Rev1, regardless of whether or not the base pairing occurred between the nucleotides located in either positions 3 to 8 (-Bs3-8SLIII) or positions 5 to 10 (-Bs5-10SLIII). However, it should be noted that in these latter 2 cases, the reductions were less drastic than that observed with the 10 bp.

To achieve an accurate analysis of the data, a prediction of the base pairing potential between the biosensor and the stem-loop III for each of the ribozymes tested was performed. Since

TABLE 1. CLEAVAGE ACTIVITIES AND THERMODYNAMIC PARAMETERS OF THE SPECIFIC ON/OFF ADAPTOR-HEPATITIS DELTA VIRUS RIBOZYMES MUTATED IN THE BIOSENSOR DOMAIN

Ribozymes	Substrate sequences ^a				Cleavage activity ^b		Bs interaction ^c
	Upstream sequence	RD	Spacer	Biosensor	k_{obs} (min^{-1})	Cleavage (%)	MFE (kcal/mol)
-Rev1-Bsorig	5'-GGCA	GGAAGAA	GCGGA	GACAGCGACG-3'	1.606 ± 0.082	85	-10.6
-Bs1-10SLIII	5'-GGCA	GGAAGAA	GCGGA	acCucCucgc-3'	0.024 ± 0.004	3	-27.0
-Bs3-8SLIII	5'-GGCA	GGAAGAA	GCGGA	GACucCucuu-3'	0.023 ± 0.001	50	-15.1
-Bs5-10SLIII	5'-gGGCA	GGAAGAA	GCGGA	cuCcuCuAuu-3'	0.015 ± 0.001	36	-15.6
-NS527-Bsorig	5'-ggUCCA	GGACAUU	CUAA	UGAGGAUGUC-3'	0.420 ± 0.020	90	-5.6
-Bs1-10SLIII	5'-ggUCCA	GGACAUU	CUAA	accuccUcgC-3'	0.061 ± 0.025	10	-26.4
-Bs3-8SLIII	5'-ggUCCA	GGACAUU	CUAA	UGcuccUcUu-3'	0.040 ± 0.003	72	-14.9
-Bs5-10SLIII	5'-gggUCCA	GGACAUU	CUAA	cuccucUGUu-3'	0.119 ± 0.007	58	-15.4
-HBV513-Bsorig	5'-gggcuaa	GCGUCGC	aaaa	AUCUCAAUUCU-3'	0.726 ± 0.063	88	-9.2
-Bs1-10SLIII	5'-gggcuaa	GCGUCGC	aaaa	acCUCcucgc-3'	0.027 ± 0.003	21	-26.4
-Bs3-8SLIII	5'-gggcuaa	GCGUCGC	aaaa	AUCUCcucCU-3'	0.437 ± 0.038	77	-15.4
-Bs5-10SLIII	5'-gggcuaa	GCGUCGC	aaaa	cUCcucAUCU-3'	0.309 ± 0.031	80	-15.4

^aRD and Biosensor indicate the binding sites, in the substrate, for their respective ribozyme's domain. The capital letters indicate the nucleotides coming from the original target sites.

^bCleavage activity was determined in the presence of 100 nM of SOFA-HDV ribozyme and trace amounts of substrate. The percentage of cleavage was determined after 2 hours at 37°C with a variation of less than 5% for at least 2 independent assays.

^cPotential interactions between the biosensor and the stem-loop III as predicted using the RNAhybrid software (MFE).

SOFA-HDV, specific on/off adaptor-hepatitis delta virus; MFE, minimum free energy of hybridization.

the prediction of the complete secondary structure of the SOFA-HDV ribozyme is not possible using common RNA folding software, the RNAhybrid software (<http://bibiserv.techfak.uni-bielefeld.de/rnahybrid>) was used, and the predictions of hybridization potential between the stem-loop III and the biosensor were performed (see Materials and Methods). This strategy bypassed the main limiting factor of these folding software, which is the miss-prediction of pseudoknot, a major structural element of the HDV ribozyme (Masquida et al., 2010). It is, in fact, an intrinsic weakness of such programs. The sequence of the biosensor was flanked in 5' by the UUU linker, and in 3' by the blocker followed by the CA (the junction between the blocker and the stem II; ie, 5'-UUUBs₁₀Bl₄CA-3') in each case. It was important to consider the flanking sequences in the prediction to include their potential contributions to the biosensor-stem-loop III interaction. These flanking sequences were added, because they are single-stranded nucleotides that could potentially contribute to the harmful interaction, especially in the case of the blocker, as there is already evidence supporting this possibility (data not shown). The second RNA sequence required for the interaction prediction was, in all cases, that of the stem-loop III (ie, 5'-ACCUCUCGCGGU-3'). The predicted interactions for the 4 biosensors of the SOFA-HDV-Rz-Rev1 are presented in Fig. 2, whereas the resulting MFE values are reported in Table 1. Concurrently, k_{obs} and the cleavage percentage after 2 hours, which are the selected indicators of the rate and the activity level, of the ribozyme respectively, were determined. The data are compiled in Table 1 for the 3 SOFA-HDV ribozymes and for their respective biosensor mutants. Briefly, if 6 bp formed between the biosensor and the sequence of the stem-loop III, the resulting ribozyme exhibited a significantly lower cleavage activity, regardless of the position of this base pairing and of the sequences of the biosensor and recognition domains. Although, this limit seems to apply to the majority of the tested ribozymes, there are exceptions

such as the SOFA-HDV-HBV513-Bs3-8SLIII or -Bs5-10SLIII. In these cases, the percentages of cleavage after 2 hours are less affected as compared with other ribozymes, whereas their k_{obs} are still reduced by half. This higher residual activity may come from the optimal cleavage site of the substrate with an upstream sequence CUAA or from other interactions that decrease the strength of the inhibitory interaction. The 4 other nucleotides of the biosensor and the sequence of the recognition domain may also contribute to this higher activity. The occurrence of 10 bp was detrimental to the cleavage activity (Table 1). However, the closer to the blocker the interaction occurs, the stronger the effect tends to be (Table 1 and data not shown).

Similar experiments were also performed to verify whether or not the biosensor can interact with the junction IV/II. When the sequence of the biosensor included a stretch of nucleotides that can base pair with the 5 nucleotides of the junction, no reduction of the cleavage activity was observed (data not shown). In fact, it was shown that in order to observe a significant inhibition of the ribozyme, the biosensor should be complementary to 10 consecutive nucleotides in the region spanning the 3' strand of stem II as well as junction IV/II, a situation which is highly improbable (1 out of 4¹⁰ nucleotides). Even if this was the case, the reduction was less important than that observed for 10 bp formed with the stem-loop III, probably because it also requires the unfolding of stem II.

In summary, only the intramolecular base pairing between the biosensor and the stem-loop III appears to be an important factor reducing the potential of the SOFA-HDV ribozyme. Detailed analysis of the SOFA-HDV ribozymes in the library developed from previous projects did not result in the identification of other domains whose interactions could explain why some ribozymes did not exhibit significant cleavage levels.

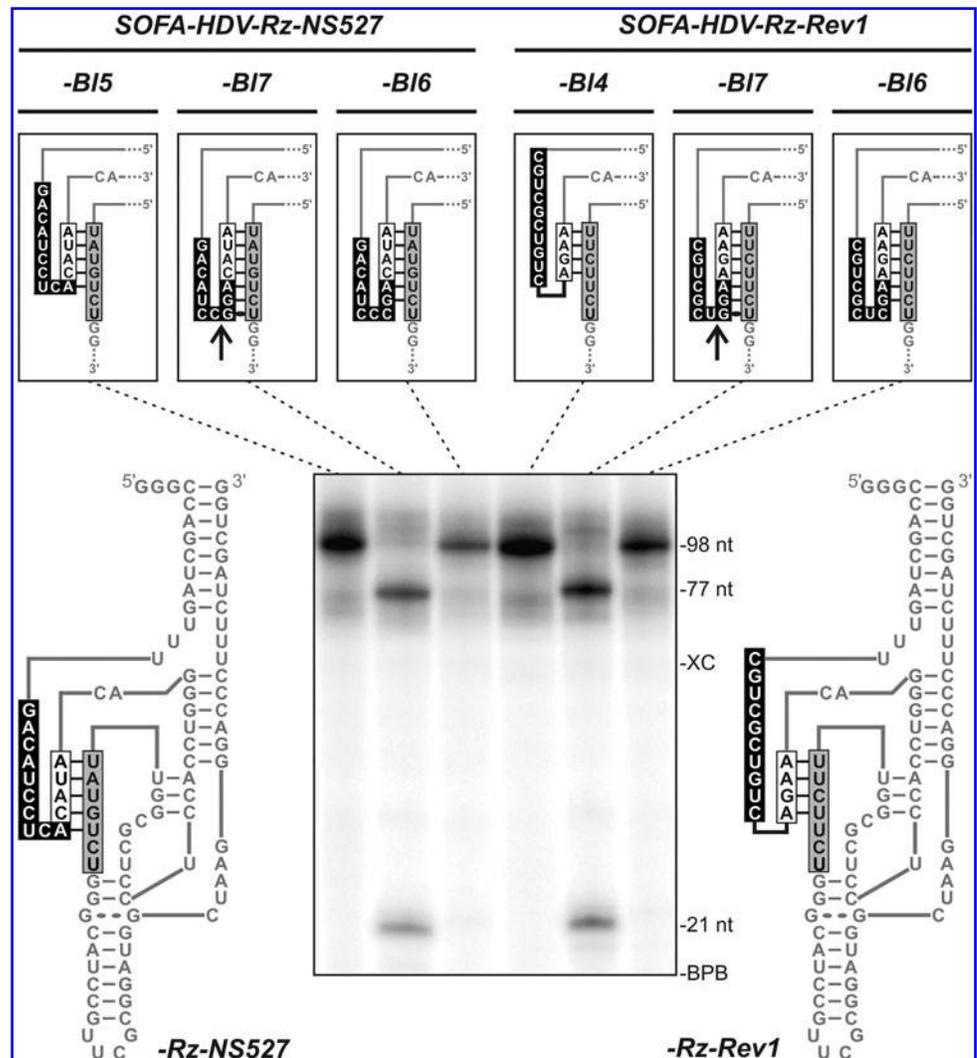
The size of the blocker domain

The underlying concept of the SOFA module relies on both the formation of the blocker stem and on the annealing of the

biosensor to the target RNA. Basically, the blocker acts as an element competing with the substrate for access to the recognition domain, thus increasing the ribozyme's fidelity. Consequently, the longer the blocker stem is, the lower the activity of a given ribozyme tends to be, although that activity should be more specific (Bergeron et al., 2005). However, the biosensor and the blocker are contiguous domains within the SOFA module. Consequently, the last nucleotide of the biosensor could also extend the blocker simply by base pairing with the 3' region of the recognition domain. More specifically, in some cases, the 9th and 10th nucleotides of a 10-nucleotide-long biosensor could form base pairs with the 6th and 5th nucleotides of the recognition domain, respectively, thus creating a longer blocker stem. This may happen relatively frequently (ie, a 1 out of 4 chance of creating a 1bp extension). For example, the biosensor of SOFA-HDV-Rz-NS527 forms a 5 bp blocker due to the presence of an extra AU bp (Fig. 3, left side). Thus, this ribozyme remains highly active either because 5 bp are insufficient to inactivate it significantly, or because, in this particular case, the resulting blocker stem is not stable enough to have a significant impact on the activity (Table 1). More importantly, the addition of 2 bp to the blocker stem may have a critical impact on the cleavage ac-

tivity of the ribozyme. Some examples of this phenomenon were found in a library created in the course of a study aiming at designing SOFA-HDV ribozymes targeting HCV RNA strands (Levesque et al., 2010). All the ribozymes harboring a 6 bp blocker sequences resulting from an extension of the blocker by 2 bp that involved residues considered as being within the biosensor exhibited a low level of cleavage activity. If a Wobble bp is present in the 7th bp, which can occur 25% of the time, this would result in a self-cleaving sequence (ie, a *cis*-acting HDV ribozyme). To verify this hypothesis, both the SOFA-HDV-Rz-NS527 and -Rz-Rev1 biosensors were mutated in order to create blocker stems of 6 bp in length. One nucleotide was mutated in SOFA-HDV-Rz-NS527 (which possessed an original blocker of 5 bp) to create SOFA-HDV-Rz-NS527-BI6, whereas 2 nucleotides were changed in the -Rz-Rev1 (which possessed an original blocker of 4 bp) to form SOFA-HDV-Rz-Rev1-BI6. These SOFA-HDV ribozymes forming 6 bp blocker stems were compared with their corresponding original SOFA-HDV ribozymes. One more mutation was added to both to include a GU Wobble bp at the bottom of the recognition domain, creating SOFA-HDV-Rz-NS527-BI7 and -Rz-Rev1-BI7, respectively (see Fig. 3). The SOFA-HDV ribozymes were synthesized by *in vitro* transcription in the presence

FIG. 3. Self-cleavage demonstration with 2 different SOFA-HDV Rz. The SOFA-HDV-Rz-NS527 (*left*) and the SOFA-HDV-Rz-Rev1 (*right*) are presented on each side of the gel. For both ribozymes, the original sequence (-BI5 or -BI4) and 2 variants with blockers of either 7 (-BI7) or 6 (-BI6) bp (mutations in the last 2 or 3 nucleotides of the Bs) were *in vitro* transcribed in the presence of [α - 32 P]UTP. The full-length transcripts (98 nt) and both the 5'(21 nt) and 3'(77 nt) self-cleavage products can be seen on the gel. The sequences of the biosensor, the blocker, and the recognition domain for each SOFA-HDV ribozyme are shown in boxes in a representation of an *off* conformation. The arrows indicated the putative self-cleavage sites. The positions of the xylene cyanol (XC) and bromophenol blue (BPB) dyes are indicated adjacent to the gel.



of [α - 32 P]UTP to be able to follow all of the RNA species that could be produced during the reactions, including the self-cleavage products. The full-length SOFA-HDV ribozyme is 98 nt in size (see Fig. 3). If a 7bp blocker that includes a GU Wobble which is formed between the eighth nucleotide of the biosensor and the seventh nucleotide of the recognition domain is present, self-cleavage produced 2 shorter RNA products and there is almost no full length ribozyme present (see Fig. 3, SOFA-HDV-Rz-NS527-B17 and -Rev1-B17). Ribozymes of this type are not interesting, as they will never be useful for the development of a gene-inactivation system. No self-cleavage products are detected when the last guanosine residue is replaced by a cytosine (Fig. 3-Rz-NS527_B16 and -Rz-Rev1_B16). The 2 other possibilities that were tested for the last base pair revealed that only the presence of an AU bp resulted in self-cleavage, and even then only at a low level (data not shown). Overall, our results show that the identity of the last nucleotides of the biosensor need to be analyzed to detect and remove any Rz that forms an extended blocker stem of 6 bp or more. This conclusion is also supported by previously reported data (see Bergeron et al., 2005).

Structure of the target

The structure of the target is obviously an important feature to consider when targeting an RNA molecule *in trans*. The formation of an intermolecular RNA duplex between the ribozyme and the target RNA can be compromised through competition from unfavorable intramolecular base pairing present in the targeted region. The importance of a site's accessibility to a ribozyme has been demonstrated in several studies. In the case of the HDV ribozymes, various protocols for the identification of the more susceptible regions have also been reported (Bergeron and Perreault, 2002). However, more recently, it was found that the addition of the SOFA motif reduces the importance of the accessibility hurdle (Bergeron and Perreault, 2005). Most likely, the presence of the biosensor domain that forms 10 bp with the substrate helps disrupt any initial secondary structure present in the target RNA, at least in terms of binding energy. The subsequent binding of both the biosensor and the ribozyme's recognition domain to the substrate might then occur in a cooperative manner. However, analysis of previous studies aimed at developing SOFA-HDV ribozymes targeting various RNA substrates does not lead to a clear conclusion on this issue. For example, it was a lot easier to obtain active, even very active (ie, >70% cleavage activity *in vitro*), SOFA-HDV ribozymes targeting the 3'-end of the HCV RNA negative strand than ones targeting its counterpart of the positive strand that folds into a more stable secondary structure (ie, the 5'-UTR and its internal ribosome entry site motif) (Levesque et al., 2010).

To clarify the situation, an experiment was designed that independently addressed the importance of the accessibilities of the biosensor and recognition domains. Several substrates that include identical sequences for the binding of both the biosensor and recognition domains of SOFA-HDV-Rz-HBV513 were synthesized (Fig. 4A). This ribozyme was selected because of its outstanding activity *in vitro* (Fig. 4B). The reference substrate (513ss) was designed to fold mainly into a single-stranded RNA that contains both binding sites in its 5'-end portion. The sequence of the 3' portion of the various substrates was mutated in such a way that it was either

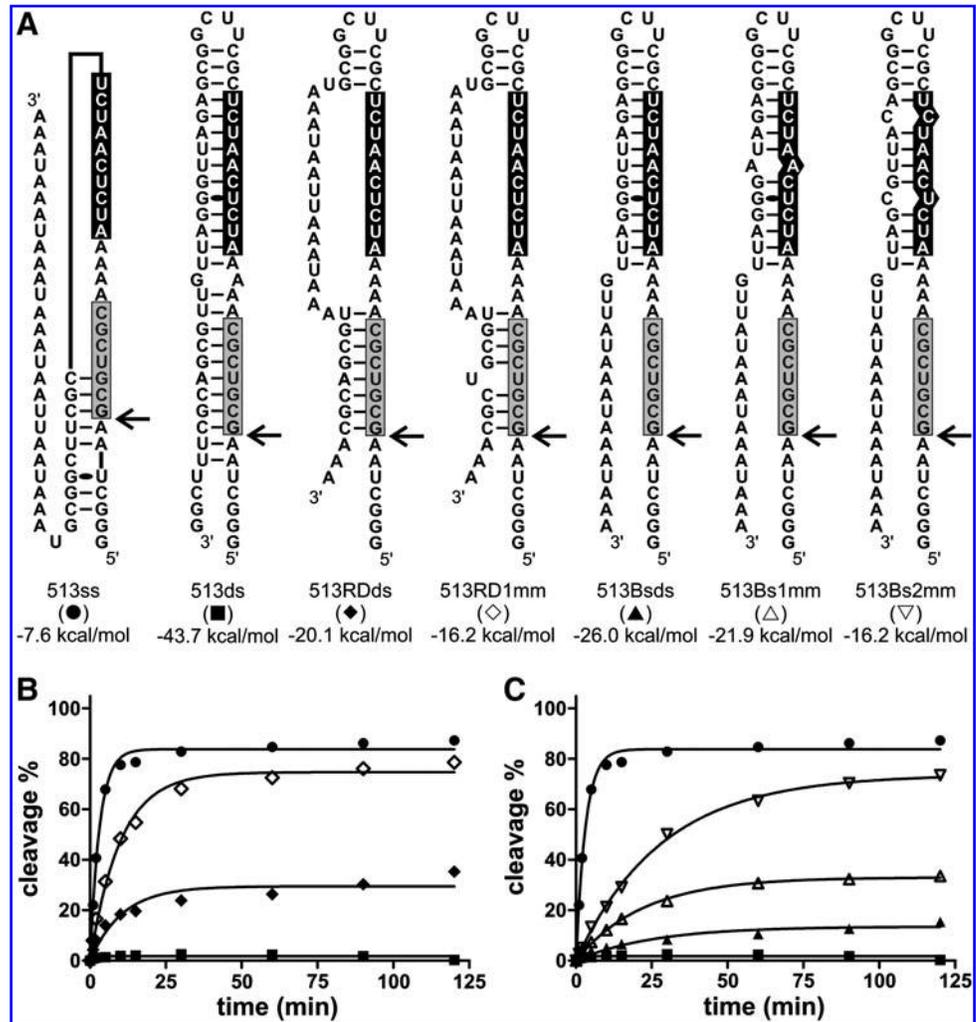
complementary, or not, to either the biosensor or the recognition domain binding sites. In other words, when perfectly complementary, it should limit (or prevent) the binding of the corresponding ribozyme domain, and *vice versa*. The secondary structures of each substrate, as predicted by Mfold (Zuker, 2003), are illustrated in Fig. 4A. The SOFA-HDV-Rz-HBV513 exhibited high cleavage activity on the almost completely single-stranded substrate (513ss), but it could not efficiently cleave the substrate that is primarily double stranded (513ds) (Fig. 4B; 87% and <1% of cleavage, respectively). The presence of base pairs in the recognition domain binding site was enough to strongly reduce the activity of the ribozyme (Fig. 4B, 513RDds, 35% cleavage). A single mismatch located in the middle of the base-pairing region of the recognition domain was sufficient to cause a significant improvement in the ribozyme's cleavage activity (Fig. 4B, 513RD1mm, 78% cleavage). The influence of base pairing at the level of the biosensor's binding site on the ribozyme's cleavage was also evaluated (Fig. 4C). When this region was included in a double-stranded structure, and one of the recognition domains was single stranded, the ribozyme exhibited only residual activity (Fig. 4C, 513Bsds 15% cleavage). Mutations that enabled the formation of either 1 (513Bs1mm) or 2 (513Bs2mm) mismatches restored the activity in a Δ G dependent manner (Fig. 4C, 33% and 73% cleavage, respectively). In fact, the activity of the ribozyme on this set of substrates seems to be inversely correlated with the stability of the structure present in the target RNA, in agreement with the predicted impact of the substrate's structure.

This experiment showed that even if both the biosensor and the recognition domains are located in close proximity to one another and may interact with the substrate in a cooperative manner, the accessibility of the substrate is important for the ribozyme to exhibit high level of cleavage activity. However, predicting the accessibility of a targeted site for a ribozyme is not a simple task, and fixing a threshold for the design of SOFA-HDV ribozymes is even more complicated. For example, the experiments in which the HCV RNAs were targeted showed that it is still hard to reliably predict the accessibility to the SOFA-HDV ribozyme, even when using a coupled bioinformatic-biochemical approach (Levesque et al., 2010). However, in the course of this work, it also appeared to be clear that the addition of the SOFA module significantly reduces the substrate accessibility hurdle, because several HDV ribozymes did not cleave at all, whereas their SOFA-HDV ribozyme counterparts did, albeit at different levels (F.P. Brière and J.P. Perreault, data not shown). Thus, it was concluded that it is important to select target sites that show a minimum of accessibility to be able to design SOFA-HDV ribozymes with a relatively good potential for success. Subsequently, an *in vitro* cleavage assay should permit the screening of the most potent ribozymes. However, it is important to keep in mind that an RNA may fold differently *in vivo* than *in vitro* and may, therefore exhibit a different accessibility for a given cleavage site.

A simple procedure for selecting the more potent SOFA-HDV ribozymes

The work just described, as well as the accumulated data from several studies recently performed aiming at developing a gene-inactivation system based on SOFA-HDV ribozymes

FIG. 4. Evaluation of the target site's accessibility. (A) The predicted secondary structures of the different substrates for SOFA-HDV-Rz-HBV513. Seven substrates that adopt different structures, due to sequence changes in the 3' region, are illustrated. The substrate 513ss (●) is mostly single stranded, whereas 513ds (■) is mainly double stranded. The substrates 513RDds (◆) and 513RD1mm (◇) are double stranded in the recognition domain's binding site with 0 or 1 mismatches, respectively. The substrates 513Bsds (▲), 513Bs1mm (△), and 513Bs2mm (▽) are double stranded in the biosensor's binding site with 0, 1, or 2 mismatches, respectively. The ΔG of each of the structure, as predicted by Mfold, is shown below the name of each substrate. The arrows indicate the cleavage sites. (B) and (C) Graphic representations of the *in vitro* cleavage assays performed with each substrate in the presence of 100 nM of SOFA-HDV-Rz-HBV513.



(Fiola et al., 2006; Robichaud et al., 2008; Levesque et al., 2010; D'Anjou et al., 2011; Laine et al., 2011), identified some of the determinants that appear to be important while considering the design of more potent tools for gene inactivation. Importantly, we believe a point has been reached that where considering only simple rules based on the sequences of both the ribozyme and the target, RNA should be sufficient to remove a significant proportion of the latter, specifically those with limited potential. Previous work performed with minimal RNA cleaved by the HDV ribozyme has shown that the identities of the nucleotides located in positions -1 and -2 of the substrate, that is to say those in the positions adjacent to the cleavage site, significantly influence the cleavage level (Deschenes et al., 2000). Specifically, the presence of a guanosine in position -1 should be avoided, and the presence of 2 consecutive pyrimidines in positions -1 and -2 is not favorable. Moreover, the experiments just described have shown that the presence of long base-pairing regions between the biosensor and the stem-loop III, as well as extended blocker sequences, should be avoided. Together, this provides 3 simple criteria for removing SOFA-HDV ribozymes that have a lesser chance of success before any target structure analysis.

To validate these simple rules, an experiment designing SOFA-HDV ribozymes targeting the RNA coding for the

RNA polymerase HCV (ie, the NS5B gene, a coding region 1,773 nucleotides in size) was performed. First, all potential target sites along the NS5B RNA strand were selected on the basis of the presence of H_{-1}/G_{+1} sites (H stands for A, C or U). This excludes the presence of a guanosine in position -1 and satisfies the necessity for the presence of a guanosine (required for the formation of the Wobble bp) located in position $+1$. This selection led to the identification of 319 potential cleavage sites along the NS5B RNA. For each site, the sequences of the substrate complementary to both the biosensor and recognition domains of the SOFA-HDV ribozyme were always defined by considering a spacer of 5 nucleotides in length (ie, the spacer is the substrate sequence located between the regions bound by the biosensor and the recognition domains, see Fig. 1). This spacer size, which was shown to be optimal when between 3 and 7 nucleotides in length, was kept constant to restrict the number of variables in the experiment. Subsequently, a simple scoring system that includes 2 values was developed to classify the potential SOFA-HDV ribozymes. First, each of the possible nucleotides combination in position -2 and -1 was assigned an arbitrary score (ie, specifically $CC = -15$; UC, UU and $CU = -10$; $CA = -5$; UA, AU and $GU = 0$; GC, AC , and $GA = +5$; $AA = +10$) from the least to the most suitable ones based on previously reported

data for the HDV ribozyme's preferred nucleotides upstream of the cleavage site (Deschenes et al., 2000). However, it should be noted that this has never been verified for SOFA-HDV ribozymes. Second, the predicted MFE value of the potential interaction between the biosensor and the stem-loop III as calculated using RNAhybrid was associated to the SOFA-HDV ribozymes. The sum of both values was considered as the score for a given SOFA-HDV ribozyme, and the classification of the 319 potential SOFA-HDV ribozymes from the most to the least potent was then obtained.

From this list, the 10 ribozymes with the highest scores, along with the 10 with the lowest scores, were selected for further analyses. The 2 clusters of SOFA-HDV- ribozymes were synthesized by run-off transcription. Two ribozymes of the lowest scores showed unexpected self-cleavage during the transcription without the formation of a perfect stem I. To avoid any bias that could be generated by the use of self-cleaving ribozymes in a trans-cleavage reaction, we removed these 2 ribozymes from the study. To exclude the impact of the target structure, short substrates 32 nucleotides long, of which 26 nucleotides were derived from the NS5B gene, were used (Table 2). At the 5'-end, 6 nucleotides were added for technical purposes without any impact on the cleavage activity, as they corresponded to positions -5 to -10 from the cleavage site and are too far from the cleavage site to affect the cleavage efficiency. Subsequently, cleavage assays were performed *in vitro* over an incubation time of 3 hours. The cleavage rates (k_{obs}) and percentages of cleavage after 3 hours are compiled in (Fig. 5). A significant difference is observed between the 2 clusters. All ribozymes from the high-score cluster targeted

substrates with 2 adenosines located in positions -1 and -2 [ie, A_2A_{-1}/G_{+1} , the best reported sequence for an HDV ribozyme cleavage site; (Deschenes et al., 2000)]; had a relatively weak base pairing probability between the biosensor and the stem-loop III according to the MFE values; and possessed a blocker that was no more than 5 bp long (including blocker extension achieved through complementarity with the biosensor) (see Table 2). With the exception of 1 ribozyme (SOFA-HDV-Rz-NS5B-222), all of these SOFA-HDV ribozymes exhibited a relatively high level of cleavage, with the k_{obs} varying from 0.49 to 0.558 minute^{-1} and the cleavage percents being greater than 66% (8 out of 9 were greater than 78%). The exception, SOFA-HDV-Rz-NS5B-222, cleaved its substrate only moderately, with a k_{obs} of 0.024 minute^{-1} and cleavage that was limited to only 10% after 3 hours, for unknown reasons. In the case of the ribozymes with the lowest scores, the best value observed for k_{obs} was 0.060 minute^{-1} , and the highest cleavage percentage was 70%. More importantly, most of these SOFA-HDV ribozymes did not achieve a k_{obs} higher than 0.025 minute^{-1} , which is 20-fold lower than what was observed for the best ribozymes. All the SOFA-HDV ribozymes from this cluster targeted substrates with 2 cytosines located in positions -1 and -2 (C_2C_{-1}/G_{+1}), except for SOFA-HDV-Rz-NS5B-266 that had a C_2U_{-1}/G_{+1} and is, in fact, the most active of these ribozymes. Some of these ribozymes may possess the potential of base pairing between the biosensor and the stem-loop III, as indicated by the relatively high predicted MFE values, and one seems to possess an extended blocker of 6 bp in length. These facts might partially explain the low levels of cleavage activity observed. However,

TABLE 2. DATA FROM THE SPECIFIC ON/OFF ADAPTOR-HEPATITIS DELTA VIRUS RIBOZYMES TARGETING HEPATITIS C VIRUS NS5B

Ribozymes		Substrate sequences ^a			Prediction ^b			Cleavage activity ^c	
Name	Upstream sequence	RD	Spacer	Biosensor	Scores	MFE (kcal/mol)	Bl (bp)	k_{obs} (min^{-1})	Cleavage (%)
-NS5B_207	5'-gggaaaUCAAA	GGAGAUG	AAGGC	GAAGGCGUCC-3'	0.6	-9.4	4	0.093 ± 0.008	78
-NS5B_222	5'-gggaaaCGAA	GGCGUCC	ACAGU	UAAGGCUAAA-3'	2.4	-7.6	4	0.024 ± 0.008	10
-NS5B_237	5'-gggaaaUUAA	GGCUAAA	CUUCU	AUCCGUGGAG-3'	1.6	-8.4	5	0.558 ± 0.080	80
-NS5B_270	5'-gggaaaGUAA	GCUGACG	CCCCC	ACAUCGGCC-3'	-0.9	-10.9	5	0.412 ± 0.033	87
-NS5B_342	5'-gggaaaGCAA	GGCCGUU	AACCA	CAUCCGCUCC-3'	-0.3	-10.3	4	0.132 ± 0.006	88
-NS5B_372	5'-gggaaaGGAA	GGACUUG	CUGGA	AGACACUGAG-3'	2.0	-8.0	4	0.049 ± 0.006	66
-NS5B_465	5'-gggaaaGCAA	GCCAGCU	CGCCU	UAUCGUUUC-3'	3.2	-6.8	4	0.240 ± 0.021	88
-NS5B_894	5'-gggaaaUGAA	GGCCGCU	GCGGC	CUGUCGAGCU-3'	1.0	-9.0	4	0.109 ± 0.014	79
-NS5B_1265	5'-gggaaaGCAA	GGAUGAU	CCUGA	UGACUCAUUU-3'	2.4	-7.6	5	0.055 ± 0.006	82
-NS5B_1592	5'-gggaaaGUAA	GGACCAA	GCUCA	AACUCACUCC-3'	-1.6	-11.6	4	0.071 ± 0.004	86
-NS5B_266	5'-gggaaaGCCU	GUAAGCU	GACGC	CCCCACAUUC-3'	-28.5	-18.5	5	0.060 ± 0.003	70
-NS5B_346	5'-gggaaaGGCC	GUAUACC	ACAUC	CGCUCGUGU-3'	-26.1	-11.1	4	0.025 ± 0.001	60
-NS5B_359	5'-gggaaaAUCC	GCUCCGU	GUGGA	AGGACUUGCU-3'	-24.3	-9.3	6	0.022 ± 0.004	44
-NS5B_665	5'-gggaaaACCC	GCUGUUU	UGACU	CAACGGUCAC-3'	-28.2	-13.2	4	0.013 ± 0.002	67
-NS5B_833	5'-gggaaaCGCC	GGUGCCG	CGCGA	GCGGUGUACU-3'	-29.5	-14.5	4	0.0030 ± 0.0001	30
-NS5B_1157	5'-gggaaaACCC	GUGACCC	CACCA	CCCCCUUGC-3'	-34.1	-19.1	5	0.017 ± 0.001	42
-NS5B_1620	5'-gggaaaUCCC	GGCUGCG	UCCCA	GUUGGAUUUA-3'	-27.3	-12.3	5	0.018 ± 0.001	33
-NS5B_1703	5'-gggaaaGCCC	GACCCCG	CUGGU	UCAUGUGGUG-3'	-27.0	-12.0	4	0.016 ± 0.001	48

^aRD and Biosensor indicate the binding sites, in the substrate, for their respective ribozyme's domain. The capital letters indicate the original nucleotides coming from the NS5B gene.

^bBioinformatic prediction including the arbitrary scoring values (Scores), the potential interaction between the biosensor, and the SLIII as predicted using the RNAhybrid software (MFE) and the number of base pairs forming the blocker stem (Bl) including any extension caused by complementarity between the biosensor and the recognition domain.

^cCleavage activity was determined in the presence of 100 nM of SOFA-HDV ribozyme and trace amounts of substrate. The percentage of cleavage was determined after 3 hours at 37°C with a variation of less than 5% for at least 2 independent assays.

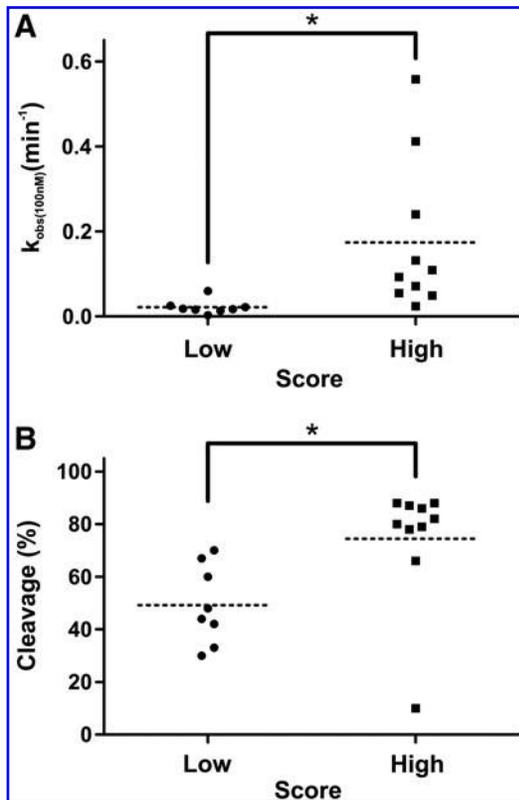


FIG. 5. Analysis of the kinetic parameters of the SOFA-HDV ribozymes targeting the NS5B sequences. **(A)** and **(B)** are the cleavage rates (k_{obs}) and the percentages of cleavage after 3 hours, respectively, determined for SOFA-HDV ribozymes of both the low- and high-score clusters. The *dashed lines* represent the mean of each group, whereas the * indicates P values less than 0.05.

it is not obvious that the negative effects are additive, at least according to the cleavage activities of the tested ribozymes. More importantly, when an unpaired t test was performed to establish whether both clusters of SOFA-HDV ribozymes were significantly different, the p -values obtained were found to be lower than 0.05 for both the k_{obs} and the cleavage percentage (Fig. 5). Consequently, it is possible to select the most potent SOFA-HDV ribozymes and to discard the others.

Discussion

Significant effort has been directed toward optimizing the design processes of the different RNA silencing tools (eg, siRNA and shRNA) and of the most commonly used catalytic RNAs (eg, *hammerhead* ribozyme) (Hendry et al., 2004; Vorobjeva et al., 2006; Levesque et al., 2007; Ashihara et al., 2010; Laitala-Leinonen, 2010; Matveeva et al., 2010; Muhonen and Holthofer, 2010; Zhou and Rossi, 2010). This study is the first that is aimed gaining knowledge about the features which should be taken into consideration when developing a gene-inactivation system, more specifically one based on the SOFA-HDV ribozyme, a catalytic RNA that has shown great promise for cleaving various RNA targets of clinical interest (Fiola et al., 2006; Robichaud et al., 2008; Levesque et al., 2010; D'Anjou et al., 2011; Laine et al., 2011). To improve the design process, and thereby ensure a more efficient production of

promising SOFA-HDV ribozymes, the key determinants for the selection of the ribozymes with the greatest potential were identified by considering the substrate and the ribozyme as distinct RNA molecules that influence the resulting cleavage activity.

In terms of the ribozyme itself, the goal is to avoid working with molecules that lack significant cleavage activity, a situation that may occur when either the ribozyme adopts an inactive structure, or itself impairs its catalytic ability. The HDV ribozyme folds into a highly compact structure that seems to follow a relatively limited number of alternative folding pathways. However, the addition of the SOFA module implies the presence of the biosensor, which initially is a relatively long single-stranded domain, but can potentially intramolecularly base pair with other parts of the ribozyme. The experiments performed here unambiguously demonstrated that a biosensor sequence possessing good complementarity to the stem-loop III inhibited the resulting SOFA-HDV ribozyme's activity (see Fig. 2). Initially, during the folding pathway of the HDV ribozyme, the stem-loop III is located outside the catalytic core (Reymond et al., 2010), and, therefore, was shown to be relatively available for interaction with an oligonucleotide (Ananvoranich and Perreault, 2000). This is in agreement with the observation that the inactive structure of an allosteric HDV ribozyme was caused by the presence of a sequence located at the top of the stem II that interacts with the stem-loop III region (Beaudoin and Perreault, 2008). Conversely, when the biosensor sequence was complementary to junction IV/II, the other mainly single-stranded region of the catalytic core, the cleavage activity was not impaired. The biosensor and the junction IV/II are most likely located in opposite orientations, at least according to the tridimensional structure of the HDV ribozyme (Reymond et al., 2010). The second situation investigated that can result in the production of an inactive ribozyme was the extension of the blocker stem. Suitable blocker domains are expected to be formed by 4 nt of base pairing with the recognition domain (Bergeron et al., 2005). A longer blocker of 5 bp may, or may not, limit cleavage activity depending on the relative stabilities of the bp. In the case of the 6 bp blocker, the competition of the substrate with the blocker for binding to the recognition domain appears to be impossible most, if not all, of the time. This observation receives physical support from the analysis of several SOFA-HDV ribozymes present in the library that exhibited low levels of cleavage activity while permitting the presence of longer blocker stems resulting from the presence of 1 or 2 additional base pairs derived from the positions located adjacent to the biosensor. More importantly, blocker sequences of 7 bp that included a Wobble base pair at the end were detrimental, because they resulted in ribozyme self-cleavage (see Fig. 3). Retrospective analysis of the SOFA-HDV ribozymes produced to target HCV yielded 2 self-cleaving ribozymes, which might help in explaining their low observed cleavage activities (Levesque et al., 2010). This type of situation can be easily predicted during the design of a SOFA-HDV ribozyme, and, therefore, can be simply avoided.

With regard to the substrate, the situation is more complex with the important parameters being both the sequence and the structure. The experiments illustrated in Fig. 4 revealed that the structure of the substrate is important. Clearly, when the sequences bound by both the recognition and the biosensor domains are accessible (ie, are single-stranded), the

cleavage level is higher. However, the binding of both domains appears to act in a cooperative manner, thus reducing the importance of this requirement as compared with what is observed in other ribozymes (including the original HDV ribozyme) for which this was shown to be of critical importance. A single mismatch in the substrate's structure located in the middle of the domains bound by the ribozyme was sufficient to cause good binding and, subsequently, significant cleavage activity. Various bioinformatic and biochemical approaches have been developed to identify the most accessible sites along a target RNA (Amarzguioui et al., 2000; Bergeron and Perreault, 2002; Mercatanti et al., 2002; Marin and Vanicek, 2011). However, even today, the ability to determine the target site's accessibility for a ribozyme is limited; hence, the development of the SOFA module and the SOFA-HDV ribozyme, as it could potentially reduce the importance of this hurdle.

Conversely, the identity of the nucleotides located adjacent to the cleavage site appears to be of primary importance. The analysis of the SOFA-HDV ribozymes designed to target small RNA substrates derived from the NS5B RNA revealed that the presence of the required nucleotides in the positions located in -1 and -2 of the cleavage site for the HDV ribozyme remained important on the addition of the SOFA module. More precisely, the presence of a guanosine residue cannot be tolerated in position -1 , and it is preferable to avoid the presence of 2 consecutive pyrimidines in positions -1 and -2 . In addition, the presence of 2 consecutive adenosines in these positions seems to be the best combination, although no systematic analysis in the context of the SOFA-HDV ribozymes has been performed.

In conclusion, substrate regions that appear to be relatively accessible, which harbor the sequence $A_{-2}A_{-1}/G_{+1}$ and are targeted by a SOFA-HDV ribozyme possessing both a relatively weak base-pairing probability between the biosensor and the stem-loop III and a blocker of no more than 5 bp in length (potential blocker extension), appear to be the best probable situation for the development of a gene-inactivation system. Among the 319 possible SOFA-HDV ribozymes targeting the NS5B RNA (1,773 nucleotides), 105 sites (33%) would have been removed by simply considering that the presence of any $Y_{-1}Y_{-2}G_{+1}$ site should result in that site being discarded (ie, Y is a pyrimidine). Moreover, if a threshold MFE value of -15 kcal/mol, as predicted by RNAhybrid for the potential base pairing between the biosensor and stem-loop III, was considered, another 21 SOFA-HDV ribozymes (7%) would have been discarded. Finally, concerning the extension of the blocker stem, 121 SOFA-HDV ribozymes (38%) have an extended blocker, including 52 (16%) that would have formed a blocker of 6 or more bp and that, therefore, should have been removed. Thus, if it is considered that a potential ribozyme can be discarded only once, 155 ribozymes (48%) would have been retained by this selection process. Interestingly, if only the SOFA-HDV ribozymes targeting the preferential cleavage site ($A_{-2}A_{-1}/G_{+1}$) are considered, only 28 sites (9%) remain for the subsequent steps without even considering the accessibility data. Analysis of the library of SOFA-HDV ribozymes tested up to date revealed that the best ones are in agreement with these requirements. Clearly, the proposed selection process should now be considered as a gold standard for the further development of a gene-inactivation tool based on SOFA-HDV ribozyme, although some testing will always be needed for the final selection of the ribozymes to be used in the cell.

Acknowledgments

The authors would like to acknowledge Dominique Lévesque and Noura Mazloum for their technical assistance. This work was supported by a grant from the Canadian Institute of Health Research (CIHR, grant MOP-44002) to J.P.P. The RNA group is supported by grants from Université de Sherbrooke. M.V.L. is the recipient of a predoctoral fellowship from the Fonds de Recherche en Santé du Québec. J.P.P. holds the Canada Research Chairs in both genomics and catalytic RNAs, and is a member of the Centre de Recherche Clinique Étienne Lebel.

Author Disclosure Statement

No competing financial interests exist.

References

- AMARZGUIOUI, M., BREDE, G., BABAIE, E., GROTLI, M., SPROAT, B., and PRYDZ, H. (2000). Secondary structure prediction and *in vitro* accessibility of mRNA as tools in the selection of target sites for ribozymes. *Nucleic Acids Res.* **28**, 4113–4124.
- ANANVORANICH, S., and PERREAULT, J.P. (2000). The kinetics and magnesium requirements for the folding of anti-genomic delta ribozymes. *Biochem. Biophys. Res. Commun.* **270**, 600–607.
- ASHIHARA, E., KAWATA, E., and MAEKAWA, T. (2010). Future prospect of RNA interference for cancer therapies. *Curr. Drug Targets* **11**, 345–360.
- ASIF-ULLAH, M., LEVESQUE, M., ROBICHAUD, G., and PERREAULT, J.P. (2007). Development of ribozyme-based gene-inactivations; the example of the hepatitis delta virus ribozyme. *Curr. Gene Ther.* **7**, 205–216.
- BEAUDOIN, J.D., and PERREAULT, J.P. (2008). Potassium ions modulate a G-quadruplex-ribozyme's activity. *RNA* **14**, 1018–1025.
- BERGERON, L.J., and PERREAULT, J.P. (2002). Development and comparison of procedures for the selection of delta ribozyme cleavage sites within the hepatitis B virus. *Nucleic Acids Res.* **30**, 4682–4691.
- BERGERON, L.J., and PERREAULT, J.P. (2005). Target-dependent on/off switch increases ribozyme fidelity. *Nucleic Acids Res.* **33**, 1240–1248.
- BERGERON, L.J., REYMOND, C., and PERREAULT, J.P. (2005). Functional characterization of the SOFA delta ribozyme. *RNA* **11**, 1858–1868.
- D'ANJOU, F., BERGERON, L.J., LARBI, N.B., FOURNIER, I., SALZET, M., PERREAULT, J.P., and DAY, R. (2004). Silencing of SPC2 expression using an engineered delta ribozyme in the mouse betaTC-3 endocrine cell line. *J. Biol. Chem.* **279**, 14232–14239.
- D'ANJOU, F., ROUTIER, S., PERREAULT, J.P., LATIL, A., BONNEL, D., FOURNIER, I., SALZET, M., and DAY, R. (2011). Molecular validation of PACE4 as a target in prostate cancer. *Transl. Oncol.* **4**, 157–172.
- DESCHENES, P., LAFONTAINE, D.A., CHARLAND, S., and PERREAULT, J.P. (2000). Nucleotides -1 to -4 of hepatitis delta ribozyme substrate increase the specificity of ribozyme cleavage. *Antisense Nucleic Acid Drug Dev.* **10**, 53–61.
- FIOLA, K., PERREAULT, J.P., and COUSINEAU, B. (2006). Gene targeting in the Gram-Positive bacterium *Lactococcus lactis*, using various delta ribozymes. *Appl. Environ. Microbiol.* **72**, 869–879.

- HAASNOOT, J., WESTERHOUT, E.M., and BERKHOUT, B. (2007). RNA interference against viruses: strike and counter-strike. *Nat. Biotechnol.* **25**, 1435–1443.
- HENDRY, P., MCCALL, M.J., STEWART, T.S., and LOCKETT, T.J. (2004). Redesign and chemically-modified hammerhead ribozymes with improved activity and serum stability. *BMC Chem. Biol.* **4**, 1.
- LAINE, S., SCARBOROUGH, R.J., LEVESQUE, D., DIERLAURENT, L., SOYE, K.J., MOUGEL, M., PERREAULT, J.P., and GATIGNOL, A. (2011). *In vitro* and *in vivo* cleavage of HIV-1 RNA by new SOFA-HDV ribozymes and their potential to inhibit viral replication. *RNA Biol.* **8**, 343–353.
- LAITALA-LEINONEN, T. (2010). Update on the development of microRNA and siRNA molecules as regulators of cell physiology. *Recent. Pat. DNA Gene Seq.* **4**, 113–121.
- LEVESQUE, D., BRIERE, F.P., and PERREAULT, J.P. (2007). A modern mode of activation for nucleic acid enzymes. *PLoS One* **2**, e673.
- LEVESQUE, D., CHOUFANI, S., and PERREAULT, J.P. (2002). Delta ribozyme benefits from a good stability *in vitro* that becomes outstanding *in vivo*. *RNA* **8**, 464–477.
- LEVESQUE, M.V., LEVESQUE, D., BRIERE, F.P., and PERREAULT, J.P. (2010). Investigating a new generation of ribozymes in order to target HCV. *PLoS One* **5**, e9627.
- LEVESQUE, M.V., and PERREAULT, J.P. (2011). Target Induced SOFA-HDV Ribozyme. *Methods Mol. Biol.* (in press).
- MARIN, R.M., and VANICEK, J. (2011). Efficient use of accessibility in microRNA target prediction. *Nucleic Acids Res.* **39**, 19–29.
- MASQUIDA, B., BECKERT, B., and JOSSINET, F. (2010). Exploring RNA structure by integrative molecular modelling. *N. Biotechnol.* **27**, 170–183.
- MATVEEVA, O.V., KANG, Y., SPIRIDONOV, A.N., SAETROM, P., NEMTSOV, V.A., OGURTSOV, A.Y., NECHIPURENKO, Y.D., and SHABALINA, S.A. (2010). Optimization of duplex stability and terminal asymmetry for shRNA design. *PLoS One* **5**, e10180.
- MERCATANTI, A., RAINALDI, G., MARIANI, L., MARIANGONI, R., and CITTI, L. (2002). A method for prediction of accessible sites on an mRNA sequence for target selection of hammerhead ribozymes. *J. Comput. Biol.* **9**, 641–653.
- MUHONEN, P., and HOLTHOFER, H. (2010). Bioinformatic approaches to siRNA selection and optimization. *Methods Mol. Biol.* **623**, 93–107.
- MULLER-KULLER, T., CAPALBO, G., KLEBBA, C., ENGELS, J.W., and KLEIN, S.A. (2009). Identification and characterization of a highly efficient anti-HIV pol hammerhead ribozyme. *Oligonucleotides* **19**, 265–272.
- PERACCHI, A. (2004). Prospects for antiviral ribozymes and deoxyribozymes. *Rev. Med. Virol.* **14**, 47–64.
- REYMOND, C., BEAUDOIN, J.D., and PERREAULT, J.P. (2009). Modulating RNA structure and catalysis: lessons from small cleaving ribozymes. *Cell Mol. Life Sci.* **66**, 3937–3950.
- REYMOND, C., LEVESQUE, D., BISAILLON, M., and PERREAULT, J.P. (2010). Developing three-dimensional models of putative-folding intermediates of the HDV ribozyme. *Structure* **18**, 1608–1616.
- ROBBINS, M., JUDGE, A., AMBEGIA, E., CHOI, C., YAWORSKI, E., PALMER, L., MCCLINTOCK, K., and MACLACHLAN, I. (2008). Misinterpreting the therapeutic effects of small interfering RNA caused by immune stimulation. *Hum. Gene Ther.* **19**, 991–999.
- ROBICHAUD, G.A., PERREAULT, J.P., and OUELLETTE, R.J. (2008). Development of an isoform-specific gene suppression system: the study of the human Pax-5B transcriptional element. *Nucleic Acids Res.* **36**, 4609–4620.
- SCHUBERT, S., and KURRECK, J. (2004). Ribozyme- and deoxyribozyme-strategies for medical applications. *Curr. Drug Targets* **5**, 667–681.
- TAMURA, A., and NAGASAKI, Y. (2010). Smart siRNA delivery systems based on polymeric nanoassemblies and nanoparticles. *Nanomedicine (Lond)* **5**, 1089–1102.
- TEDESCHI, L., LANDE, C., CECCHETTINI, A., and CITTI, L. (2009). Hammerhead ribozymes in therapeutic target discovery and validation. *Drug Discov. Today* **14**, 776–783.
- VOROBEVA, M., ZENKOVA, M., VENYAMINOVA, A., and VLASSOV, V. (2006). Binary hammerhead ribozymes with improved catalytic activity. *Oligonucleotides* **16**, 239–252.
- ZHOU, J., and ROSSI, J.J. (2010). Aptamer-targeted cell-specific RNA interference. *Silence* **1**, 4.
- ZUKER, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415.

Address correspondence to:

Prof. Jean-Pierre Perreault

RNA Group/Groupe ARN

Département de Biochimie

Faculté de médecine et des sciences de la santé

Université de Sherbrooke

Sherbrooke

Québec

Canada J1H 5N4

E-mail: jean-pierre.perreault@usherbrooke.ca

Received for publication April 07, 2011; accepted after revision June 06, 2011.