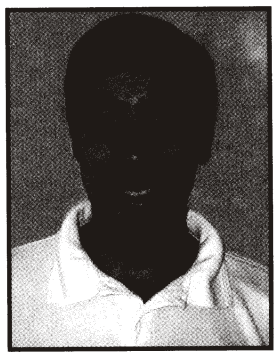
Ribozyme-Based Gene-Inactivation Systems Require A Fine Comprehension of their Substrate Specificities; the Case of Delta Ribozyme

Lucien Junior Bergeron[1], Jonathan Ouellet[1] and Jean-Pierre Perreault*

RNA Group/Groupe ARN, Département de Biochimie, Faculté de médecine, Université de Sherbrooke, Sherbrooke, Québec, J1H 5N4, Canada



Abstract: The ability of ribozymes (i.e. RNA enzymes) to specifically recognize and subsequently catalyze the cleavage of an RNA substrate makes them attractive for the development of therapeutic tools for the inactivation of both viral RNAs and mRNAs associated with various diseases. Several applicable ribozyme models have been tested both *in vitro* and in a cellular environment, and have shown significant promise. However, several hurdles remain to be surpassed before we generate a useful gene-inactivation system based on a ribozyme. Among the most important requirements for further progress are a better understanding of the features that contribute to defining the substrate specificity for cleavage by a ribozyme, and the identification of the potential cleavage sites in a given target RNA. The goal of this review is to illustrate the importance of both of these factors at the RNA level in the development of any type of ribozyme based gene-therapy. This is achieved by reviewing the recent progress in both the structure-function relationships and the development of a gene-inactivation system of a model ribozyme, specifically *delta* ribozyme.

Keywords: Ribozyme, Substrate specificity, Hepatitis *delta* virus, Gene-inactivation, RNA structure-function.

1. INTRODUCTION

Conventional medicinal chemistry consists of the development of a specific inhibitor for a given protein. In contrast, by taking action on the nucleic acid, gene-therapy aims to prevent the synthesis of this protein. The ability of ribozymes (i.e. RNA enzymes; Rz) to specifically recognize and subsequently catalyze the cleavage of an RNA substrate makes them attractive for the development of therapeutic tools for the inactivation of both viral RNAs and RNAs associated with various diseases. The sequencing of the human genome has clearly shown that more than 90% of the genome is not transcribed into mRNA molecules (i.e. the products of the genes). Therefore, it should be more precise, in terms of specificity, to target the RNA rather than the DNA. The specificity of recognition of a ribozyme for its substrate is determined by Watson-Crick base pairs. In addition, the specificity of ribozyme cleavage depends on numerous other features including: the nucleotides close to the cleavage site on the target RNA; intramolecular interactions; and, *trans*-acting factors. Each type of ribozyme (i.e. *delta*, hairpin, hammerhead, etc.) possesses a distinct substrate specificity that determines what it will cleave. When considering developing a gene-inactivation system that is functional *in vivo*, it is imperative to remember that an ideal ribozyme would cleave only the desired mRNA species without affecting any other RNA species or biological mechanisms of the cell.

*Address correspondence to this author at the RNA Group/Groupe ARN, Département de Biochimie, Faculté de médecine, Université de Sherbrooke, Sherbrooke, Québec, J1H 5N4, Canada; Tel: (819) 564-5310; Fax: (819) 564-5340; E-mail: Jean-Pierre.Perreault@USherbrooke.ca

[1]These authors have an equal contribution.

2. *DELTA* RIBOZYME: A MODEL CATALYTIC RNA

Hepatitis *delta* virus (HDV) is unique among animal viruses because it contains a single-stranded circular RNA genome of 1,7 kilobases [for a review see ref. 1]. This genome replicates through a DNA-independent rolling circle mechanism. The genomic circular monomer is replicated by a host polymerase into linear, multimeric, antigenomic strands, which are then cleaved and circularized, yielding antigenomic circular monomers. Using the latter RNA as template, the same three steps are then repeated to produce the progeny. Both the genomic and antigenomic strands possess self-catalytic RNA motifs (i.e. ~85 nucleotides, nt) that are responsible for cleaving the multimeric strands and releasing the monomers (i.e. *cis*-acting ribozymes; see Fig. (**1A**)) [for reviews see ref. 2,3]. Like other small, active, self-cleaving motifs, *delta* motifs catalyze the cleavage of one of its own phosphodiester bonds, yielding reaction products containing a 5'-hydroxyl and a 2',3'-cyclic phosphate termini. It has been possible to separate the self-cleaving sequence into two molecules in order to develop *trans*-acting systems where one molecule, identified as a ribozyme, possesses the catalytic properties required to successively cleave several molecules of substrate (i.e. turnover) [for a review see ref. 4]. The most frequent way to separate the RNA self-cleaving motif is to remove the J1/2 junction, as illustrated in figure **1** [5]. In this example, the P1 stem is the recognition domain and is formed by Watson-Crick base-pairing of the substrate to the ribozyme. It is noteworthy that cleavage by a *cis*-acting *delta* ribozyme is ~50-fold faster than that by the equivalent *trans*-acting version [2]. This difference in rate constant has been attributed to a higher entropic penalty for the bimolecular reaction as compared to the unimolecular one, and to a kinetically less stable catalytic center resulting from the absence of the J1/2

0929-8673/03 $41.00+.00

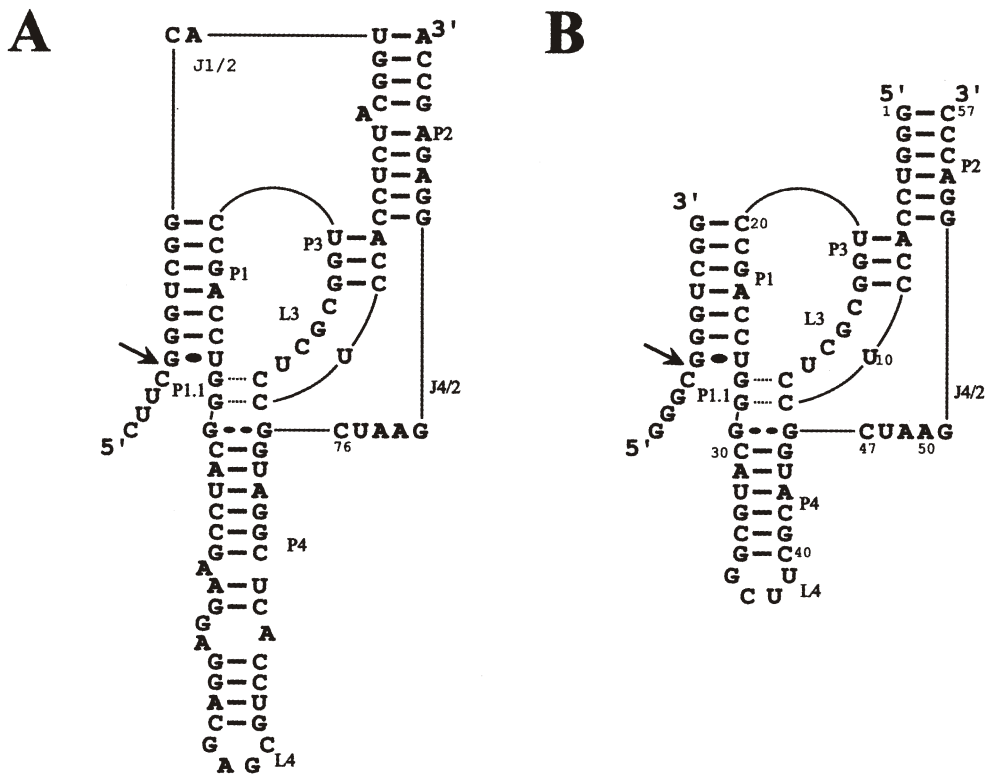


Fig. (1). Secondary structures and sequences of the antigenomic *delta* ribozymes. **A**) A *cis*-acting natural variant. **B**) A *trans*-acting version derived from the *cis*-acting ribozyme [see ref. 18 for more details]. Paired regions are numbered P1, P1.1 (dotted lines), P2, P3, and P4; and joining regions are named according to the paired elements they link. The arrow indicate the cleavage sites. The homopurine base pairs at the top of the P4 stem are represented by two dots, and the wobble base pairs by one.

crossover, which could add rigidity to the tertiary structure [2].

Delta ribozyme possesses several unique features, all of which are related to the fact that it is the only catalytic RNA motif that has been discovered in humans. Thus, it was postulated that *delta* ribozymes have evolved so as to possess optimal activity in this cellular environment [6]. For example, *delta* ribozyme is naturally efficient at low magnesium concentrations (~1 mM, i.e. human-like conditions), and is the only ribozyme known to be fully active in the presence of calcium [7]. This contrasts with all other ribozymes that are currently attracting interest as potential tools for gene inactivation as these are all derived from RNAs isolated from non-mammalian species. Clearly, *delta* ribozyme appears to be well adapted to the human cell environment, and is therefore ideal for the development of a gene-inactivation system. Consequently, a proper understanding of its substrate specificity for ribozyme cleavage is of primary importance. In order to achieve this task, several aspects of the molecular mechanism of *delta* ribozyme activity were reviewed.

3. MOLECULAR MECHANISM OF *DELTA* RIBOZYME ACTIVITY

In order to understand in detail the specificity of an enzyme, a fine comprehension of its molecular mechanism, including the folding pathway that leads to the formation of an active ternary complex, is critical. Since a review of the relevant *delta* ribozyme literature is difficult because several different versions of the ribozyme have been characterized (e.g. *trans*- versus *cis*-acting forms), here we exclusively refer, in terms of nucleotide numbering, to a 57 nt version derived from the antigenomic strand of the HDV illustrated in figure **1B**. In this numbering system C_{76} of the antigenomic *cis*-acting motif is referred to C_{47} (see Fig. (**1A**) and Fig. (**1B**)).

3.1. Secondary Structure

Trans-acting *delta* ribozymes of either polarity have frequently been designed and used to study the impact of mutations on cleavage activity [e.g. 5,8-10]. The pseudoknot secondary structure best fits the experimental data for both the antigenomic and genomic RNAs [1,2]. According to this structure, the *trans*-acting *delta* ribozyme consist of two stems (P1 and P2), two stems-loops (P3-L3 and P4-L4), one pseudoknot (P1.1) and one single-stranded junction (J4/2) (Fig. (**1B**)). The *cis*-acting version possesses an additional junction (J1/2) and its P2 stem is a pseudoknot, therefore it adopts a nested double-pseudoknot secondary structure (Fig. (**1A**)). The P2 stem and the P4-L4 hairpin, which is invariably preceded by a homopurine base pair, are structural features with non-specific sequences. These stems are located on each side of the catalytic center. The P3 stem length (i.e.

a 3 base pairs (bp) region with a preference for GC pairing) is critical, suggesting that it positions the L3 loop in a specific and critical orientation. The single-stranded L3 and J4/2 regions are essential for catalysis, while the J1/4 junction may be either single- or double-stranded (a pseudoknot referred as the P1.1 stem (Fig. (**1**)). This additional pseudoknot is formed by 2 base pair interactions, and was initially observed in the crystal structure of a self-cleaved product of the genomic version [11,12]. The existence of this pseudoknot was further supported by a mutational analysis of both *cis*- and *trans*-acting ribozymes [13-15].

The global structure comprises five helical segments joined together to generate two parallel co-axial stacks [3]. One includes the P1, the P1.1 and the P4 stems, while the other is formed by the P2 and P3 stems. The proximity of the N3 of the C_{47} to the 5'-OH group of G_{+1} (2.7 angstroms) suggested a role for this cytosine in the chemical reaction [11]. The crystal structure also provided several other important observations [11,12]. Firstly, the presence of a trefoil turn with G_{46}, C_{47} and U_{48} being the three leaflets. The consequence of this trefoil is a deep projection of C_{47} into the active site of the ribozyme that extrudes U_{48} to the solvent and causes G_{46} to stack with A_{49}. This observation received physical evidence from fluorescence [16], cross-linking [17], and binding shift experiments [18]. Secondly, the identification of several 2'-hydroxyl groups that contribute to a network of interactions within the catalytic center, including the formation of a ribose zipper between residues from the J4/2 junction and the P3 stem. The involvement of several of these 2'-hydroxyl groups was confirmed by a systematic analysis using 2'-H modifications at specific locations of the antigenomic *trans*-acting *delta* ribozyme [19]. In the latter study, the ribose zipper appears as a motif specific to the crystal structure. More generally, it is important to notice that all of these interactions occur near the cleavage site (that is to say at the bottom of the P1 stem). Therefore, the P1 stem is essential for the tertiary interactions within the catalytic core that result in a proper folding pathway, rather than being only a simple binding site for the substrate.

3.2. Folding Pathway

Kinetic experiments under single- and multiple-turnover conditions, cleavage inhibition, gel-shift assays and other experiments have allowed to the description of the minimal reaction pathway for *delta* ribozyme (see Fig. (**2**)) for a schematic representation) [18]. Initially, the substrate (S) and the ribozyme (Rz) engage in formation of the P1 helix (RzS). Next, upon the addition of a magnesium ion, a conformational transition has been shown to be essential for the formation of an active RzS' complex. One of the consequences of this conformational transition is to bring the P1 stem within the catalytic core. Substrate binding is a two-step process, and both steps are involved in determining the substrate specificity for cleavage by the ribozyme. After binding, the chemical step occurs (i.e. a transesterification). The 5'-end product appears to be released immediately upon completion of the chemical step, while the 3'-end product allowed the observation of product release inhibition under certain conditions. It has been suggested that both ground state destabilisation of the RzS complex, as well as the formation of a very stable ribozyme-3' product complex (RzP), were two important features driving the kinetic pathway [2].

Elucidation of the intimate details of the folding pathway remains a complex problem. However, some information has been obtained using various approaches. For example, an assay based on ribonuclease H accessibility to DNA-RNA duplexes formed by specific oligonucleotides bound to the single-stranded domain of the *delta* ribozyme revealed several features [20]. It was shown that the P1 stem can form in the presence of NaCl, while the P1.1 stem forms only in the presence of divalent metal ions. More importantly, these experiments suggested that the folding steps, and most likely the formation of the P1.1 stem, were the rate-limiting ones of the kinetic mechanism.

According to the crystal structure of the self-cleaving *delta* ribozyme, no tightly bound metal ion was located within the catalytic RNA [11]. However, a previous study of the antibiotic inhibition of *delta* ribozyme activity suggested that both neomycin and magnesium are likely to displace lead ion(s), leading to the hypothesis that this catalytic RNA does indeed possess a (or several) metal ion coordination site(s) [21]. It has been proposed that a *delta* ribozyme composed of three RNA oligomer strands is bound by three magnesium ions [22]. In addition, an absolute requirement for the presence of divalent metal ions was shown for *delta* ribozyme catalysis at physiological pH [23]. Furthermore, a specific magnesium-induced cleavage at position G_{52} at the bottom of the P2 stem was observed to occur solely within catalytically active RzS complexes; whereas, neither the binding site(s), nor the function of this metal ion have been elucidated [24]. Recently, the NMR spectrum of a version of *delta* ribozyme bound to an inactive substrate was found to be unmodified by the addition of magnesium ions [25]. Conversely, a post cleavage complex, the ribozyme bound to the 3'-end product (RzP), shows important differences upon the addition of magnesium ion. The largest structural

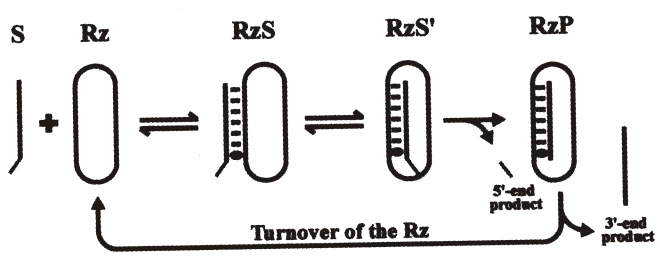


Fig. (2). Schematic representation of the kinetic pathway of *delta* ribozyme cleavage.

difference driven by the magnesium appears to be its binding at the bottom of the P1 stem (at the G-U wobble base pair) [25]. It appears reasonable to suggest that a magnesium ion bound at the bottom of the P1 stem either contributes to the chemical step, or stabilizes a structural rearrangement.

Several differences were noted between the RzS complexes pre- and post-cleavage. For example, NMR studies confirmed the presence of an additional base pair, G_{15}-U_{10}, within the L3 loop of the post-cleavage complex only [25]. This base pair, originally identified in the crystal [11], is a reverse wobble base pair. Moreover, the P1 stem in the post-cleavage complex appears to be more stable than that in the pre-cleavage one [2]. The P1.1 stem is most likely critical for and central to the conformational transition leading to the formation of the active ternary complex (RzS'), but is not the only important interaction. Furthermore, fluorescence resonance energy transfer experiments using donors and acceptors located at both extremities (i.e. the end of P2 and P4 stems) indicate that the space between them varies during the kinetic pathway [26]. The ribozyme-substrate ternary complex (RzS') appears to be contracted when compared to the ribozyme-3'product complex (RzP). Clearly, an important network of tertiary interactions takes place after formation of the P1 stem in order to produce a highly ordered catalytic center. Since the P1 stem is an essential feature of the catalytic center, it necessarily influences the substrate specificity of *delta* ribozyme cleavage.

3.3. Chemical Step

Although it is not the main topic of this review, a few words must be said concerning the recent progress made in the understanding of the chemistry of *delta* ribozyme catalysis. Upon proper folding leading to the formation of a functional RzS complex, RNA components participating in the catalysis should be in the appropriate conformation to perform the in-line attack that leads to the formation of a trigonal-bipyramidal intermediate in the transition state [2]. *Delta* ribozyme was the first RNA for which a general acid-base catalysis was proposed [27-30], however the detailed mechanism of this cleavage remains controversial. It has been demonstrated that imidazol buffer rescued the activity of a mutant *delta* ribozyme possessing C_{47} altered to U_{47} [27,28]. This demonstration suggests that C_{47} acts as a general base. In contrast, it has been shown that C_{47} act as a general acid, with the Rz-bound hydrated metal hydroxide acting as a general base [29,30]. In the latter report, it was also demonstrated that in the absence of bivalent cation, the presence of NaCl at very high concentrations supports the cleavage activity, although only at a pH near 5.0. The relative rate constant for the chemical step in the *delta* ribozyme reaction is comparable to that for cleavage by RNase A [29]. This observation suggests that general acid-base catalysis has the potential of providing enormous cleavage rate enhancements for this ribozyme.

4. SUBSTRATE SPECIFICITY FOR RIBOZYME CLEAVAGE

Substrate recognition by *delta* ribozyme is based primarily on the formation of the P1 stem by Watson-Crick base pairs (Fig. (3)), that is to say the initial step of the folding pathway. Early studies demonstrated that the P1 stem required one G-U wobble base pair followed by six non-specific Watson-Crick base pairs [2,5]. As a result, *delta* ribozyme can be considered as an example of an RNA restriction enzyme for which it is easily possible to change the recognition specificity because the nature of the base pair interactions is known (i.e. A and U as well as G and C base-pair together, respectively) [6]. This affirmation is true for all ribozymes and antisense RNAs. In contrast, restriction enzymes (i.e. protein) cannot have their specificities so easily modified because the nature of the interactions between the enzyme (i.e. protein) and its DNA substrate remains unknown and is highly unpredictable based on our current knowledge.

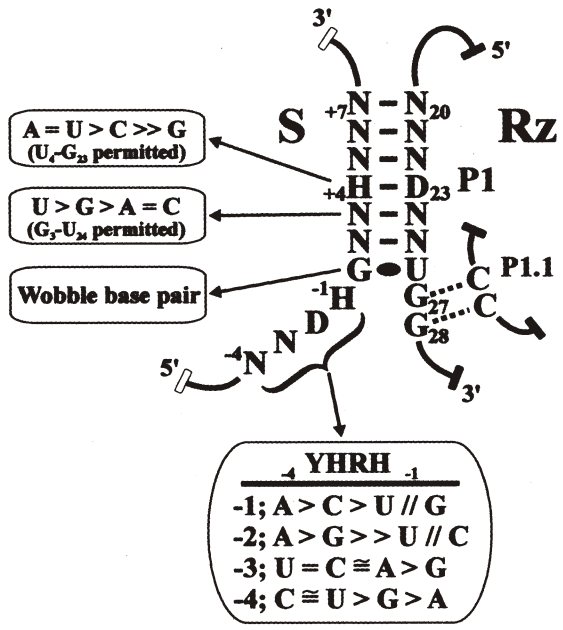


Fig. (3). Nucleotide sequences that contribute to the substrate specificity of *delta* ribozyme cleavage. The order of cleavage efficiency, as determined with a collection of small model substrates, is indicated for some of the most important positions. Order of reactivity is identified for several positions. **S** and **Rz** represent substrate and ribozyme, respectively. The base pairs of the P1.1 stem are represented by dotted lines.

Although, several studies addressing the questions related to the substrate specificity of *delta* ribozymes in both *cis*- and *trans*-acting forms have been performed [e.g. see ref. 6,31-36], most experiments were carried out by either randomly changing the base pairing combinations or by introducing nucleotides that interfere with the Watson-Crick base pairing between the substrate and the ribozyme in the P1 stem. No information existed on how each nucleotide of the substrate affects both the cleavage activity and the reaction kinetics as most investigations were performed at only one or two positions at a time.

In order to achieve this task our group undertook a systematic study aiming to identify most, if not all, of the features that contribute to defining the substrate specificity of one version of the *trans*-acting antigenomic *delta* ribozyme (see Fig. (3)). Initially, a collection of substrates

and ribozymes including modifications affecting only the P1 stem were synthesized and the resulting cleavage activities determined [38,39]. The introduction of any mismatches at positions +3 and +4 resulted in a complete lack of cleavage by the original ribozyme, while mismatches at other positions of the P1 stem severely decreased the level of activity [38] (see Fig. (**3**)). The presence of base pairs between the substrate and the ribozyme at positions S_{+3}-Rz_{24} and S_{+4}-Rz_{23} has been shown to be essential for cleavage activity. In addition, the combination of the nucleotides in positions S_{+4}-Rz_{23} significantly influences the level of cleavage; for example, the base pair G_{+4}-C_{23} results in an unproductive ribozyme-substrate complex [39]. These nucleotides in the middle of the P1 stem were proposed to be essential for both substrate binding and subsequent steps in the cleavage pathway [38,39]. A systematic evaluation of all possible combinations of natural bases in positions S_{+1}-Rz_{26} revealed that only the presence of wobble base pairs, including the G-U found in all natural variants, results in an efficient cleavage (unpublished data; M.C. April and J.P. Perreault). This is in agreement with the results of mutagenesis experiments using a *cis*-acting ribozyme that showed the importance of the nature of the base pair in position S_{+1}-Rz_{26} [40]. Together, these findings suggest that factors more complex than simple base pairing interactions, such as tertiary structure interactions, could play an important role in the substrate specificity of *delta* ribozyme cleavage. Thus, the P1 stem is more than only a recognition domain, it also participates, through tertiary interactions, in the folding pathway that leads to the formation of an active RzS complex.

The presence of a single nucleotide located immediately 5' of the cleavage site is sufficient for cleavage [6]. At position –1, the order of reactivity was found to be A > C >U. The presence of a guanosine produced a substrate uncleavable by the ribozyme used in this systematic analysis (Fig. (**3**)) [41]. Most likely the presence of a G_{-1} either significantly modifies the chemical environment near the scissile phosphate, or base pairs with C_{47} thereby preventing this base from playing its role in the chemical mechanism (see above). A base preference slightly different in position –1 was reported for another version of the *delta* ribozyme, but in all studies the presence of a guanosine was detrimental [36,37]. Shift of the cleavage site to position -2 was observed only in an exceptional construct [36], and therefore is not a common phenomenon with *delta* ribozymes.

During the development of *delta* ribozymes cleaving RNA encoding the HDV antigen, evidence suggesting that the substrate sequence upstream of the cleavage site is essential in the selection of an appropriate cleavage site was gathered [42]. In order to understand the role of this single-stranded region in efficient cleavage, a collection of small substrates that possessed single and multiple mutations in positions –4 to –1 was studied [41]. Some of these substrates were uncleavable, while others showed a >60-fold difference in relative specificity between the least and most efficiently cleaved substrates. The base of each position from –4 to –1 contributes differently to the ability of a substrate to be cleaved. The optimal sequence for positions –4 to –1 was determined to be $_{-4}$YHRH$_{-1}$ (where Y indicates C or U; H indicates A, C or U; and R indicates A or G; see Fig. (**3**)). Although this portion of the substrate is not a part of the recognition domain, it plays a crucial role as an external determinant of the ability of a substrate to be cleaved, and could impose steric effects that limit the cleavage activity. The mechanisms responsible for the effects of this domain are still unknown. It appears that this region of the substrate forms base pairs with the guanosine residues of the J1/4 junction, thereby preventing the formation of the P1.1 pseudoknot and consequently leading to the formation of unproductive RzS complexes [15]. This phenomenon may be thought of as a proof reading mechanism. Following the formation of the P1 stem, a substrate that possesses the appropriate sequence in positions -4 to -1 supports the formation of the P1.1 stem and consequently the cleavage takes place. Conversely, if the substrate does not possess the appropriate sequence (i.e. a sequence that prevents the formation of the P1.1 stem), cleavage does not occur. Thus, changing the composition of the residues forming the P1.1 stem allows one to modify the substrate specificity [15]. It should be noted, however, that the ribozyme with the original P1.1 stem (i.e. one formed by two C-G bp) was found to be the most active. Finally, the presence of secondary structures (e.g. a hairpin) in close proximity to the cleavage site was shown to have an inhibitory effect on cleavage [41]. In contrast, no sequence from this domain is known to modulate the substrate specificity of *delta* ribozyme cleavage, with the exception of the relatively large RNA motif in the extension in 3'-end of the binding domain (i.e. after the P1 domain; unpublished data; J. Lehoux, S. Ananvoranich and J.P. Perreault).

These studies demonstrated that at least 11 contiguous nucleotides of the substrate (i.e. positions -4 to +7) contribute to determining the ability of an RNA molecule to be cleaved by *delta* ribozyme [15,41]. In general, hammerhead and hairpin ribozymes are designed so as to bind their substrates through the formation of two small helices of 5-6 bp each, therefore involving 10–12 nt in defining their substrate specificities. Clearly, the *delta* ribozyme specificity is as strict as that of any other ribozyme.

5. *DELTA* RIBOZYMES AS RESTRICTION ENDONUCLEASES *IN VIVO*

The possibility of designing a ribozyme to cleave any specific target RNA has made them valuable tools in both basic research (e.g. to identify the function of a protein due to their knock-down) and therapeutic applications [for recent reviews see ref. 43,44]. Ribozymes that specifically bind to and then cleave targeted RNAs may provide a way to selectively inactivate genes (see Fig. (**4**)), particularly in the case of organisms in which gene-inactivation by homologous recombination is not feasible. In addition, when the knock-out of a gene's expression is fatal to a cell, it is possible, using a specific ribozyme, to attenuate the expression of the protein to a precise level. The resulting cleavage products are then hydrolyzed by cellular endo- and exonucleases, thereby eliminating these unstable RNA molecules (Fig. (**4**)). Ribozyme efficiency can be monitored at various levels (RNA, proteins, etc.) depending the purpose. Thus, ribozymes function as sequence-selective

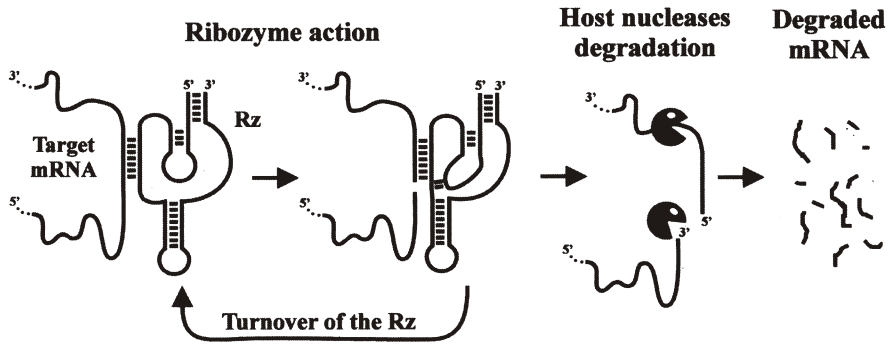


Fig. (4). Illustration of the mRNA inactivation by a ribozyme from its binding to its degradation by the host nucleases.

endoribonucleases that cleave their targets upon proper binding.

Delta ribozyme offers several unique features that are appealing for the development of a gene-inactivation system. These features are all related to the fact that it is derived from an RNA species found in human hepatocytes. For example, *delta* ribozyme possesses a long half-life regardless of both the cell line tested and the means of transfection used [6]. More precisely the *delta* ribozyme half-life is an order of magnitude longer than that of a hammerhead ribozyme, derived from an RNA species discovered in plants, under the same conditions [6]. Analysis of the hydrolysis pattern of *delta* ribozyme mutants and of the hammerhead ribozyme have led to the suggestion that the presence of the single-stranded regions at both the 5'- and 3'-extremities of an uncapped RNA species constitutes a serious drawback in terms of stability. In the case of *delta* ribozyme both the 5'- and 3'-ends are located within the P2 stem, and this appears to be sufficient to greatly stabilize the molecule. Therefore, molecular engineering does not appear to be required in order to optimize the potential of *delta* ribozyme for further use as a tool in human cells. This is an important advantage in favour of *delta* ribozyme, since the engineering of a natural ribozyme for a given criteria (e.g. the stability) has been shown to be detrimental for its other characteristics [45,46].

To date, the amount of work aimed at developing gene-inactivation therapies using *delta* ribozyme is relatively limited. *Delta* self-catalytic RNA (i.e. the *cis*-acting motif) has been used in cells to release a hammerhead ribozyme that subsequently acted in *trans* on a herpes simplex virus mRNA [47], as well as for the synthesis of an RNA with the precise 3'-end termini required in the production of defective interfering particles of vesicular stomatitis virus and HCV RNA fragments [48,49]. *Delta* ribozymes have also been used to cleave a natural mRNA in *trans*. More precisely, the specific cleavage *in vitro* of both the mRNA encoding the *delta* antigen (HDAg) and the pregenome RNA strands of the hepatitis B virus in *trans* by *delta* ribozymes have been reported [42,50], showing their potential for the further development of a system for inhibiting gene expression. This potential for cleavage in *trans* was subsequently demonstrated *in vivo* using three RNA targets [51-53]. Firstly, *delta* ribozymes were designed to target the junction of a BCR-ABL chimeric mRNA that causes chronic myelogenous leukemia [51]. This study compared the cleavage activity *in vitro*, as well as *in vivo*, of various ribozymes expressed in a transitory fashion in cell culture. *Delta* ribozyme possessing an activity *in vitro* almost three orders of magnitude lower than that of a hammerhead ribozyme but it still exhibited an activity in cells similar to that seen *in vitro*, while that of the hammerhead ribozyme was significantly reduced [51]. This result supports the idea that *delta* ribozyme is better adapted to the cellular environment. Secondly, *delta* ribozymes were shown to efficiently cleave RNA transcripts encoding uracil phosphoribosyl-transferase by transitory transfection of the pathogen *toxoplasma gondii* [52]. *Delta* ribozyme appeared to be slightly more efficient than RNA interference, which in turn was shown to be more effective than the antisense RNA approach. Finally, a stable human cell line expressing specific *delta* ribozymes allowed the observation of a complete knock-down of the subtilisin pro-protein convertase 2 (SPC2) mRNA [53]. The bioactivity of the SPC2 protein was also confirmed to be absent in the cell line by both HPLC and radioimmunoassay. In this case, the appropriate controls were also performed in order to definitively prove that it was the ribozyme activity, and not only an antisense effect, that was responsible. Thus, the initial experiments using *delta* ribozymes as a tool for the development of gene-inactivation systems were successful.

6. SUBSTRATE SPECIFICITY WITH LONG RNA MOLECULES

The works described above show the potential of *delta* ribozyme to target natural RNA for various purposes, as well as identifying several differences in this type of cleavage as compared that of a small, model substrate. There are many factors that could result in the inefficient cleavage of targeted RNAs in the cellular environment, including inadequate expression or delivery of ribozyme minigenes to the cell and inappropriate cellular localization of both the ribozyme and target RNA, among others [54]. More specifically at least three issues have to be addressed when discussing the differences in substrate specificity between natural RNAs and small model substrates: i. the selection of sites to target; ii. the factors influencing the formation of an active RzS complex; and, iii. the verification of whether or not the rules governing the specificity for the ribozyme cleavage of small substrates can be generalized to long substrates.

6.1. Selection of Target Sites

In order to produce the anticipated effect, a ribozyme must cleave its substrate and yield a product deprived of

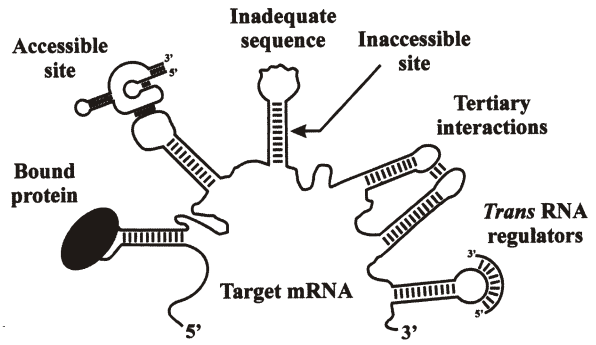


Fig. (5). Schematic representation of various factors influencing the identification of a site accessible for ribozyme cleavage.

function. In the case of an mRNA, potential cleavage sites should be located near the 5'-end region in order to ensure that cleavage results in the production of an inactive, smaller peptide. In contrast, targeting an RNA virus is less constrained because cleavage throughout the genome should control viral propagation. Moreover, target sequences located in single-stranded regions of an mRNA have a higher potential to be cleaved because they should be more accessible to ribozyme attack than those found in double-stranded regions (Fig. (**5**)) [42,50,54,55]. Within the double-stranded regions, the ribozyme might compete unfavourably with intramolecular base-pairing in order to bind its substrate. In addition, in order to be accessible a site must not be embedded in a compact tertiary structure, nor bound by a protein. For example, the tertiary structure of a large target RNA could inhibit ribozyme association by sterical hindrance [56]. This is important because the binding of a ribozyme to its target sequence is the rate-limiting step of the kinetic pathway in a cellular environment [57].

The prediction of the most suitable sites to target in a natural RNA molecule is not obvious. In order to achieve this task several procedures, based on four different approaches, have been developed for *delta* ribozyme. Firstly, structural probing of the RNA target, both via nuclease probing and chemical modification mapping, has been shown to be of limited accuracy in addition to being fastidious to perform and complex to analyze [for an example see ref. 42]. Secondly, bioinformatic packages allow the prediction of secondary structures from which it is then possible to deduce the region most susceptible to attack by a ribozyme [for an example see ref. 58]. Unfortunately, these tools have failed for long targets, at least in the case of *delta* ribozyme [42,50]. Moreover, since accessibility is a function of many complex factors, including the tertiary structures of both the ribozyme and RNA target, it is currently beyond our ability to make accurate predictions *in silico* [50]. Thirdly, empirical approaches that include the formation of a ribozyme-target complex have been used. Several procedures, such as gel-shift assays and RNase H mapping using DNA oligonucleotides that mimic the binding domain of a ribozyme, or simply a ribozyme cleavage assay, have been developed [for examples see ref. 42,50]. These procedures are useful in order to validate previous predictions; however, since the sequence of the targeted site is defined, these procedures require a tremendous number of assays/errors in order to identify the best sites. Finally, approaches based on the use of combinatorial libraries either mimicking the binding domain (i.e. including binding shift assay or RNase H hydrolysis), or using randomized ribozymes, were tested [for an example see ref. 50]. As both the target site accessibility and the ability to form an active RzS complex (or oligonucleotide-RNA target complex) are two interdependent factors whose relationship is complex, referring to a combinatorial library of the P1 stem appears to be the best way to address such a relationship [50].

Obviously, both the secondary and tertiary structures are essential components driving the accessibility of the potential target sites. These two components have a lot of potential to increase the substrate specificity of *delta* ribozyme. In addition, RNA targets that bind ribonucleoproteins would be occluded and therefore difficult to cleave (see Fig. (**5**)) [59]. This may contribute to increasing the substrate specificity. These ribonucleoproteins may also influence the folding of an RNA, resulting in an inaccurate *in vitro* selection [60]. Therefore, the development of an *in vivo* combinatorial library based approach appears to be a promising avenue for the accurate prediction of the sites with the most potential to target [for an example with the hammerhead ribozyme see ref. 59]. Alternatively, it appears possible to engineer ribozymes that are not dependent on the structure of the target. This has been made possible by coupling a domain, which recruits an RNA helicase that opens any secondary structure of a given mRNA, to the ribozyme [61]. However, this approach, which has not yet been tested for *delta* ribozyme, might be useless when we allow for the fact that the secondary and tertiary structures of an RNA species contribute to increasing the substrate specificity of ribozyme cleavage. Clearly, the emergence of new methodologies for the identification of accessible sites to be targeted by ribozymes is required.

6.2. RzS Complex Formation Including a Long RNA Species

The binding of a ribozyme to its substrate depends on the formation of Watson-Crick base pairs. The formation of the first couple of base pairs, referred to as the nucleation event, results in a base-specific complex between two

complementary strands and is assumed to be the rate-limiting step of duplex formation [62]. This initial event is then be followed by a faster, cooperative, zippering mechanism leading to the double-stranded complex. As compared to a small substrate, a longer RNA target implies a high number of tertiary interactions and tightly secondary interactions and strong RNA motifs. Therefore, the binding rates of long, structured, RNA targets would be expected to be lower than those of short, unstructured, RNA molecules. Since the RzS complex with a *delta* ribozyme depends on the short (7 bp) P1 stem, the specificity, as well as its formation efficiency, may be questionable. It has been proposed that the discrimination against mismatches in the stem formed between a ribozyme and a substrate decreases with increasing binding-arm length [63]. Furthermore, it has been suggested that the greater the length of the recognition sequence, the higher the probability that the target site has been sequestered by secondary and tertiary structural interactions that limit ribozyme accessibility [64]. The presence of long stems between a substrate and hammerhead ribozyme, whose binding domains may be extended as desired, produces a stable ribozyme-product complex. In this case product dissociation becomes the rate-limiting step of the kinetic pathway, thereby significantly reducing the turnover of the ribozyme. Consequently, it seems reasonable to postulate that the 7-bp P1 stem may be ideal for both avoiding formation of an RzS complex that includes a mismatch with a wrong substrate, and in providing a good turnover of the ribozyme.

Theoretically, a *trans*-cleaving *delta* ribozyme can be designed to cleave any mRNA species in a sequence-specific fashion but the full activity is not guaranteed even if the targeting site is accessible. For example, sequences with a high G+C content have the potential to hybridize more efficiently with the complementary arms of their ribozymes, thereby likely increasing the efficiency of the cleavage [46]. However, it should be noted that certain P1 stem sequence compositions result in misfolding of the ribozyme leading to a loss of activity.

6.3. Specificity with a Long mRNA Substrate

The specificity of *delta* ribozyme cleavage was described using small model substrates. It is of interest to ask whether or not the established rules are applicable with a long, natural RNA target? Analysis of the collections of *delta* ribozymes that target either the mRNAs encoding the antigen of the hepatitis *delta* virus, or the viral RNA of the hepatitis B virus, are useful in answering this question [42,50]. Regardless of the accessibility of the target site, most of the cleavage sites identified in these studies respect the rules. For example, they do not possess consecutive pyrimidines at positions –2 and –1, or, if they do, they are inefficiently cleaved. Exceptions to the rules were also observed. For example, a ribozyme exhibited a cleavage of HBV even though two cytosines were located at positions –2 and –1, revealing differences in targeting between small and long substrates. There seems to be a complex situation in which several factors, including the secondary structure motifs in close proximity to a potential site and long interactions, modulate cleavage efficiency. An example of this is the possibility that the two cytosines in the above ribozyme may be involved in either secondary or tertiary structures with the HBV RNA, and consequently do not compete for P1.1 stem formation. The effect of a mismatch in the P1 stem at positions +2 and +7 was also tested with a long substrate. While such mismatches can support the cleavage of a small substrate, albeit at a reduced level [38], no cleavage of HBV genome can be detected (unpublished data, L.J. Bergeron and J.P. Perreault). It seems that the integrity of the RzS complex, and more specifically that of the P1 stem, is even more important in the cleavage of a natural RNA species than in that of a small substrate. More importantly, all indications suggest that the substrate specificity for *delta* ribozyme cleavage, which is based on at least 11 contiguous nucleotides located in a accessible region, are sufficient to ensure the development of a gene-inactivation system.

7. *DELTA* RIBOZYME SPECIFICITY VS HUMAN RNA SPECIES

Does *delta* ribozyme possess sufficient substrate specificity to specifically target any RNA species into the human cell? It is possible to estimate the probability of finding a specific potential sequence to be targeted by *delta* ribozyme in a human cell. The length of the genome is estimated to be ~3.0 billion bp. Since ~5% of the genome is transcribed, this implies that ~150 million bases form the population of mRNAs. However, along a mRNA species it has been estimated that only ~20% (i.e. ~30 million bases) of the sequence are accessible for binding by short complementary molecules like ribozymes or antisenses [65]. Taking into account that *delta* ribozyme cleavage is based on 11 specific contiguous nucleotides, this predicts that it cleaves once every 4.2 million bases (i.e. 4^{11}). Thus, a specific *delta* ribozyme has the potential to cleave 7 different sites within the total population of mRNAs. This number of 7 is valuable only when considering the expression of the entire human genome at a precise time, obviously not a probable event. In fact, only ~5000 genes are estimated to be expressed in a cell at any given moment. This corresponds to approximately one seventh of the genome, which reduces the number of potential cleavage sites, to only 1. It should be noted that untranslated RNA species were not included in these calculations as, even though these RNA species are not negligible, they are usually highly structured and not suitable for ribozyme targeting. In addition, several other factors (e.g. ribozyme compartmentalization and protein-mRNA interactions) have not been taken into account in these calculations (Fig. (**5**)). Most of these factors might contribute to increasing the substrate specificity of *delta* ribozyme. For all of these reasons, we are tempted to speculate that *delta* ribozyme has the ability to recognize and cleave unique mRNA species in the human cell.

Several recent *in vivo* results strengthen these theorical calculations. In the case of *delta* ribozymes targeting the SPC2 mRNA, the expression of other members of this family of proteins was shown to be unaltered by the presence of the catalytic RNA, even though they have a high degree of similarity [53]. Moreover, in a study of molecular stability, various versions of *delta* ribozyme were transfected by different means into several cell lines. No phenotype,

such as higher cell mortality, was observed [6]. This observation receives additional support from several studies, currently being performed, aimed at targeting various natural RNA in different cell lines (unpublished data; L.J. Bergeron, K. Fiola, D. Lévesque and J.P. Perreault). All of these studies support the idea that *delta* ribozyme does not non-specifically cleave other cellular RNAs. However, it must be kept in mind to these are indirect indications. New emerging technologies such as microarrays will be helpful in definitively answering this primordial question.

8. ACKNOWLEDGMENTS

This work was supported by grants from both the Canadian Institute of Health Research (CIHR) and Health Canada to J.P.P. The RNA group is supported by a grant from the CIHR. L.J.B. was the recipient of a pre-doctoral fellowship from CIHR and J.O. from the Fonds FCAR (Québec). J.P.P. is an investigator from the Canadian Institute of Health Research (CIHR).

9. REFERENCES

[1] Mercure, S.; Lafontaine, D.; Roy, G.; Perreault, J.P. *Médecine/Science*, **1997**, *13*, 662-667.
[2] Shih, I.h.; Been, M.D. *Annu. Rev. Biochem.*, **2002**, *71*, 887-917.
[3] Doherty, E.A.; Doudna, J.A. *Annu. Rev. Biochem.*, **2000**, *69*, 597-615.
[4] Been, M.D.; Wickham, G. S. *Eur. J. Biochem.*, **1997**, *247*, 741-753.
[5] Perrotta, A.T.; Been, M.D. *Nature*, **1991**, *350*, 434-436.
[6] Lévesque, D.; Choufani, S.; Perreault, J.P. *RNA*, **2002**, *8*, 464-477.
[7] Wu, H.N.; Lin, Y.J.; Lin, F.P.; Makino, S.; Chang, M.F.; Lai, M. *Proc. Natl. Acad. Sci. USA*, **1989**, *86*, 1831-1835.
[8] Been, M.D.; Perrotta, A.T.; Rosenstein, S.P. *Biochemistry*, **1992**, *31*, 11843-11852.
[9] Perrotta, A.T.; Been, M.D. *Nucleic Acids Res.*, **1993**, *21*, 3959-3965.
[10] Thill, G.; Vasseur, M.; Tanner, N.K. *Biochemistry*, **1993**, *32*, 4254-4262.
[11] Ferré-D'Amaré, A.R.; Zhou, K; Doudna, J.A. *Nature*, **1998**, *395*, 399-404.
[12] Ferré-D'Amaré, A.R.; Doudna, J.A. *J. Mol. Biol.*, **2000**, *295*, 541-556.
[13] Wadkins, T.S.; Perrotta, A.T.; Ferré-D'Amaré, A.R.; Doudna, J.A.; Been, M.D. *RNA*, **1999**, *6*, 720-727.
[14] Nishikawa, F.; Nishikawa, S. *Nucleic Acids Res.*, **2000**, *28*, 925-931.
[15] Deschênes, P.; Ouellet, J.; Perreault, J.; Perreault, J.P. *Nucleic Acids Res.*, **2003**, *31*, 2087-2096.
[16] Harris, A.D.; Rueda, D.; Walter, G.N. *Biochemistry*, **2002**, *41*, 12051-12061.
[17] Bravo, C.; Lescure, F.; Laugaa, P.; Fourrey, J.L.; Favre, A. *Nucleic Acids Res.*, **1996**, *24*, 1351-1359.
[18] Mercure, S.; Lafontaine, D.; Ananvoranich, S.; Perreault, J.P. *Biochemistry*, **1998**, *37*, 16975-16982.
[19] Fiola, K.; Perreault, J.P. *J. Biol. Chem.* **2002**, *277*, 26508-26516.
[20] Ananvoranich, S.; Perreault, J.P. *Biochem. Biophys. Res. Commun.*, **2000**, *270*, 600-607.
[21] Rogers, J.; Chang, A. H.; Ahsen, U. V.; Schroeder, R.; Davies, J. *J. Mol. Biol.*, **1996**, *259*, 916-925.
[22] Sakamoto, T.; Tanaka, Y.; Kuwabara, T.; Kim, M.H.; Kurihara, Y.; Katahira, M.; Uesugi, S. *J. Biochem. (Tokyo)*, **1997**, *121*, 1123-1128.
[23] Murray, J. B.; Seyhan, A. A.; Walter, N. G.; Burke, J. M.; Scott, W. G. *Chem. Biol. (Lond.)*, **1998**, *5*, 587-595.
[24] Lafontaine, D.A.; Ananvoranich, S.; Perreault, J.P. *Nucleic Acids Res.*, **1999**, *27*, 3236-3243.
[25] Takana, Y.; Hori. T.; Tagaya, M.; Sakamoto. T.; Kurihara. Y.; Katahira. M.; Uesugi. S. *Nucleic Acids. Res.*, **2002**, *30*, 766-774.
[26] Pereira, J.B.M.; Harris, A.D.; Rueda, D.; Walter, G.N. *Biochemistry*, **2002**, *41*, 730-740.
[27] Perrota, T.A.; Shih, H.I.; Been, D.M. *Science*, **1999**, *286*, 123-126.
[28] Shih, H.I.; Been, D.M. *Proc. Natl. Acad. Sci. USA*, **2001**, *98*, 1489-1494.
[29] Nakano, I.S.; Chadalavada M.D.; Bevilacqua C.P. *Science*, **2000**, *287*, 1493-1497.
[30] Nakano, I.S.; Bevilacqua C.P. *J. Am. Chem. Soc.*, **2001**, *123*, 11333-11334.
[31] Kumar, P.K.R.; Suh, Y.A.; Taira, K.; Nishikawa, S. *FASEB J.*, **1993**, *7*, 124-129.
[32] Kawakami, J.; Yada, K.; Suh, Y.A.; Kumar, P.K.R.; Nishikawa, F.; Maeda, H.; Taira, K.; Nishikawa, S. *FEBS Lett.*, **1996**, *394*, 132-136.
[33] Wu, H.N.; Huang, Z.S. *Nucleic Acids Res.*, **1992**, *20*, 5937-5941.
[34] Wu, H.N.; Wang, Y.J.; Hung, C.F.; Lee, H.J.; Lai, M.M.C. *J. Mol. Biol.*, **1992**, *223*, 233-245.
[35] Wu, H.N.; Lee, J.Y.; Huang, H.W.; Huang, Y.; Hsueh, T.G. *Nucleic Acids Res.*, **1993**, *21*, 4193-4199.
[36] Nishikawa, F.; Fauzi, H.; Nishikawa, S. *Nucleic Acids Res.*, **1997**, *25*, 1605-1610.
[37] Perrotta, A.T.; Been, M.D. *Biochemistry*, **1992**, *31*, 16-21.
[38] Ananvoranich, S; Perreault, J.P. *J. Biol. Chem.*, **1998**, *273*, 13182-13188.
[39] Ananvoranich, S; Lafontaine, D.A.; Perreault, J.P. *Nucleic Acids Res.*, **1999**, *27*, 1473-1479.
[40] Perrotta, A.T.; Been, M.D. *Nucleic Acids Res.*, **1996**, *24*, 1314-1321.
[41] Deschênes, P.; Lafontaine, D.A.; Charland, S.; Perreault, J.P. *Antisense Nucleic Acid Drug Dev.*, **2000**, *10*, 53-61.
[42] Roy, G.; Ananvoranich, S.; Perreault, J.P. *Nucleic Acids Res.*, **1999**, *27*, 942-948.
[43] Lewin, A.S.; Hauswirth, W.W. *Trends Mol. Med.*, **2001**, *7*, 221-228.
[44] Michienzi, A.; Rossi, J.J. *Methods Enzymol.*, **2001**, *341*, 581-596.
[45] Wang, L.; Ruffner, D.E. *J. Am. Chem. Soc.*, **1998**, *120*, 7684-7690.
[46] Amarzguioui, M.; Brede, G.; Babaie, E.; Grotli, M.; Sproat, B.; Prydz, H. *Nucleic Acids Res.*, **2000**, *28*, 4113-4124.
[47] Pattnaik, A.K.; Ball, L.A.; LeGrone, A.W.; Wertz, G.W. *Cell*, **1992**, *69*, 1011-1020.
[48] Chowrira, B.H.; Pavco, P.A.; McSwiggen, J.A. *J. Biol. Chem.*, **1994**, *269*, 25856-25864.
[49] McCormick, C.J.; Rowlands, D.J.; Harris, M. *J. Gen. Virol.*, **2002**, *83*, 383-394.
[50] Bergeron, L.J.; Perreault, J.P. *Nucleic Acids Res.*, **2002**, *30*, 4682-4691.
[51] Kato, Y.; Kuwabara, T.; Warashina, M.; Toda, H.; Taira, K. *J. Biol. Chem.*, **2001**, *276*, 15378-15385.
[52] Al-anouti, F.; Ananvoranisch, S. *Antisense Nucleic Acid Drug Dev.*, **2002**, *12*, 275-281.
[53] D'Anjou, F.; Bergeron, L.J.; Perreault, J.P.; Day, R. *J. Biol. Chem.*, **2003**, submitted.
[54] Yu, Q.; Pecchia, D.B.; Kingsley, S.L.; Heckman, J.E.; Burke, J.M. *J. Biol. Chem.*, **1998**, *273*, 23524-23533.
[55] Birikh, K.R.; Berlin, Y.A.; Soreq, H.; Eckstein, F. *RNA*, **1997**, *3*, 429-437.
[56] Campbell, T.B.; McDonald, C.K.; Hagen, M. *Nucleic Acids Res.*, **1997**, *25*, 4985-4993.
[57] Kawasaki, H.; Taira, K. *Prosphorus, Sulfur and Silicon*, **2002**, *177*, 1645-1649.
[58] Matthews, D.H.; Burkard, M.E.; Freier, S.M.; Wyatt, J.R.; Turner, D.H. *RNA*, **1999**, *5*, 1458-1469.
[59] Pierce, M.L.; Ruffner, D.E. *Nucleic Acids Res.*, **1998**, *26*, 5093-5101.
[60] Michienzi, A.; Rossi, J.J. *Methods Enzymol.*, **2001**, *341*, 581-596.
[61] Kawasaki, H.; Taira, K. *EMBO Reports*, **2002**, *3*, 443-450.
[62] Scherr, M.; Lebon, J.; Castanotto, D.; Cunliffe, H.E.; Meltzer, P.S.; Ganser, A.; Riggs, A.D.; Rossi, J.J. *Molecular Therapy*, **2001**, *4*, 454-460.
[63] Bramlage, B.; Luzi, E.; Eckstein, F. *TibTech*, **1998**, *16*, 434-438.
[64] Lloyd, B.H.; Giles, R.V.; Spiller, D.G., Grzybowski, J.; Tidd, D.M.; Sibson, D.R. *Nucleic Acids Res.*, **2001**, *29*, 3664-3673.
[65] Nielsen, P.E. *Pharmacology Toxicology*, **2000**, *86*, 3-7.