

Université de Sherbrooke

ÉTUDE DU RIBOZYME SOFA-HDV COMME OUTIL MOLÉCULAIRE :
APPLICATION ET OPTIMISATION

Par

M. Michel Lévesque

Département de Biochimie

Thèse présentée à la Faculté de médecine et des sciences de la santé

en vue de l'obtention du grade de

philosophiae doctor (Ph.D.) en Biochimie

4 novembre 2013



Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 978-0-499-00429-1

Our file Notre référence

ISBN: 978-0-499-00429-1

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Membres du jury d'évaluation :

Dr Jean-Pierre Perreault (Biochimie, Université de Sherbrooke)

Dr Gerardo Ferbeyre (Biochimie, Université de Montréal)

Dr Brendan Bell (Microbiologie, Université de Sherbrooke)

Dr Martin Bisailon (Biochimie, Université de Sherbrooke)

TABLE DES MATIÈRES

TABLE DES MATIÈRES	III
LISTE DES TABLEAUX	VII
LISTE DES FIGURES	VIII
LISTE DES ABRÉVIATIONS	XII
RÉSUMÉ	1
INTRODUCTION	3
1. Modulation de gènes et thérapie génique : cibler l'ARN	3
1.1. Approches pour modifier l'activité d'un gène	4
1.2. Importance & intérêts de cibler l'ARN.....	6
2. Les outils moléculaires d'acides nucléiques	8
2.1. Outils non catalytiques	10
2.1.1. Oligonucléotides antisens (ASON).....	10
2.1.2. Interférence par l'ARN.....	15
2.2. Les outils catalytiques	28
2.2.1. Ribozyme de la ribonucléase P.....	31
2.2.2. Ribozyme « <i>hammerhead</i> »	38
2.2.3. Ribozyme « <i>hairpin</i> »	42
2.2.4. Ribozyme du virus de l'hépatite <i>delta</i>	45
3. Le ribozyme SOFA-HDV comme outil moléculaire	50
3.1. Concept du module de spécificité SOFA	51
3.2. Spécificité du ribozyme SOFA-HDV	53
3.3. Utilisations du ribozyme SOFA-HDV	56
4. Objectifs du projet	59
4.1. Application du ribozyme SOFA-HDV pour cibler des virus	59
4.2. Optimisation de la méthode pour la production de ribozyme SOFA-HDV .	60
RÉSULTATS	62
Chapitre 1 : Ciblage du VHC avec des ribozymes SOFA-HDV	62

ARTICLE : <i>Investigating a New Generation of Ribozymes in Order to Target HCV</i>	62
Avant-propos :	62
Résumé	63
Abstract	64
Introduction.....	65
Results	69
Discussion	80
Materials and Methods	88
Acknowledgments	93
References	94
Chapitre 2 : Ciblage du VIH avec des ribozymes SOFA-HDV	99
ARTICLE : <i>A Conserved Target Site in HIV-1 Gag RNA is Accessible to Both HDV Ribozyme and Short Hairpin RNA Activities</i>	99
Avant-propos :	99
Résumé	100
Abstract	101
Introduction.....	102
Results	104
Discussion	116
Materials and methods	120
Acknowledgements	122
References	122
Supplementary tables.....	127
Supplementary figure	128
Supplementary methods.....	129
Chapitre 3 : Sélection des ribozymes SOFA-HDV avec le meilleur potentiel	138
ARTICLE : <i>Selection of The Most Potent SOFA-HDV Ribozymes For Use in Gene Targeting</i>	138
Avant-propos :	138

Résumé	139
Abstract	140
Introduction.....	141
Materials and Methods	144
Results	148
Discussion	164
Acknowledgments	168
References	168
Chapitre 4 : Méthode de design pour la production de ribozymes SOFA-	
HDV	171
ARTICLE : <i>Target Induced SOFA-HDV Ribozyme</i>.....	171
Avant-propos :	171
Résumé	172
Abstract	173
1. Introduction.....	174
2. Materials.....	176
3. Methods.....	179
4. Notes	192
Acknowledgments	194
References	194
DISCUSSION	197
1. Applications du ribozyme SOFA-HDV.....	197
1.1. Cibler les ARN du VHC	197
1.2. Cibler les ARN du VIH.....	201
2. L'équation multifactorielle de l'activité des ribozymes SOFA-HDV.....	204
2.1. L'ARN cible a un impact à plus d'un niveau.....	205
2.2. Les composantes du ribozyme et leur influence sur l'activité du ribozyme	
.....	213
3. Comment pré-sélectionner les ribozymes SOFA-HDV	226
4. L'attrait majeur du module SOFA : la spécificité.....	230
4.1. Que savons-nous sur la spécificité du ribozyme SOFA-HDV?	230

4.2. Une spécificité comparable ou meilleure que les autres outils moléculaires?	232
5. Où se situe le ribozyme SOFA-HDV?	234
CONCLUSION	239
REMERCIEMENTS	240
LISTE DE RÉFÉRENCES	243
ANNEXES	266
1. Carte du vecteur d'expression pCMV-EGFP-phH1	266
2. Lettres d'approbation des co-auteurs	267

LISTE DES TABLEAUX

Chapitre 1

Chapitre 1, Tableau 1. Compilation of the data from the RNase H hydrolysis of HCV (+) RNA..... 71

Chapitre 1, Tableau 2. Compilation of the *in vitro* cleavage of HCV (+) strand by the SOFA-HDV ribozyme data..... 72

Chapitre 1, Tableau 3. Compilation of the *in vitro* cleavage of the HCV (-) strand by SOFA-HDV ribozyme data..... 79

Chapitre 2

Chapitre 2, Tableau S1. SOFA-HDV-Rz target sites and DNA coding sequences 127

Chapitre 3

Chapitre 3, Tableau 1. Cleavage activities and thermodynamic parameters of the SOFA-HDV ribozymes mutated in the biosensor domain..... 152

Chapitre 3, Tableau 2. Data from the SOFA-HDV ribozymes targeting HCV NS5B. 163

Dicussion

Discussion, Tableau 1. Liste partielle^a des gènes prédits par Ribosubstrate comme potentiellement ciblés par le SOFA-HDV-Rz1498. 232

LISTE DES FIGURES

Introduction

Introduction, Figure 1. Schéma de l'expression génique et des cibles potentielles de modulation.....	5
Introduction, Figure 2. Mécanismes d'action simplifiés des principaux outils d'acides nucléiques pour le ciblage d'ARN.....	9
Introduction, Figure 3. Schéma de la maturation des ARNm et des cibles potentielles des ASON.	13
Introduction, Figure 4. Illustration des modifications chimiques utilisées dans la synthèse des ASON.	15
Introduction, Figure 5. Biogénèse et fonction des miARN.	18
Introduction, Figure 6. Vue d'ensemble de l'interférence par l'ARN.	21
Introduction, Figure 7. Illustration des différents points d'entrée dans la voie de l'ARNi pour le ciblage génique.....	22
Introduction, Figure 8. Schéma de la modulation de la voie des miARN.	28
Introduction, Figure 9. Schéma de la maturation des ARN de transfert et représentation de la structure secondaire du ribozyme RNase P de <i>E. coli</i>	33
Introduction, Figure 10. Schéma des stratégies pour le ciblage de gènes via l'activité RNase P.	35
Introduction, Figure 11. Schéma de la réplication en cercle roulant symétrique et représentation de la structure secondaire du ribozyme « <i>hammerhead</i> ».	39
Introduction, Figure 12. Représentation de la structure secondaire du ribozyme « <i>hairpin</i> ».....	43
Introduction, Figure 13. Représentation de la structure secondaire du ribozyme HDV et schéma de la spécificité de son site de coupure.....	48
Introduction, Figure 14. Concept du ribozyme SOFA-HDV et représentation de sa structure secondaire.	52
Introduction, Figure 15. Démonstration de la spécificité de coupure du ribozyme SOFA-HDV.....	56

Chapitre 1

Chapitre 1, Figure 1. Schematic representations of both the HCV genome and the SOFA-HDV ribozyme.	66
Chapitre 1, Figure 2. Schematic representation of the 3-step procedure used for the identification of the sites possessing the greatest targeting potential in the HCV 5'-UTR.	70
Chapitre 1, Figure 3. Typical autoradiogram of an 8 % polyacrylamide gel performed in order to analyze SOFA-HDV ribozyme cleavage <i>in vitro</i>	75
Chapitre 1, Figure 4. Analysis of the inhibition of the HCV replicon by SOFA-HDV ribozymes.	77
Chapitre 1, Figure 5. Analysis of the cleavage activity of the SOFA-HDV ribozyme collections.	82

Chapitre 2

Chapitre 2, Figure 1. Schematic representation of HIV-1 RNA and the SOFA-HDV-Rz.	104
Chapitre 2, Figure 2. Target site identification.	Erreur ! Signet non défini.
Chapitre 2, Figure 3. Screen for SOFA-HDV-Rz activity against HIV-1 production.	Erreur ! Signet non défini.
Chapitre 2, Figure 4. Catalytic activity and specificity of SOFA-HDV-Rz1498.	Erreur ! Signet non défini.
Chapitre 2, Figure 5. Activity of shRNA1498 against HIV-1 production.	Erreur ! Signet non défini.
Chapitre 2, Figure 6. Activity of SOFA-HDV-Rz1498 and shRNA1498 against HIV-1 production from diverse viral strains.	Erreur ! Signet non défini.
Chapitre 2, Figure 7. Specificity of SOFA-HDV-Rz1498 and shRNA1498....	Erreur ! Signet non défini.
Chapitre 2, Figure 8. Activity of SOFA-HDV-Rz1498 and shRNA1498 against HIV-1 replication.	Erreur ! Signet non défini.

Chapitre 2, Figure S 1. Stable Jurkat cell lines analysis.	128
---	-----

Chapitre 3

Chapitre 3, Figure 1. Secondary structure representations of the SOFA-HDV ribozyme in both the off (left) and on (right) conformations.	143
--	-----

Chapitre 3, Figure 2. Analysis of the base-pairing potential between the biosensor and the stem-loop III.	150
--	-----

Chapitre 3, Figure 3. Self-cleavage demonstration with 2 different SOFA-HDV Rz.	155
--	-----

Chapitre 3, Figure 4. Evaluation of the target site's accessibility.	158
---	-----

Chapitre 3, Figure 5. Analysis of the kinetic parameters of the SOFA-HDV ribozymes targeting the NS5B sequences.	163
---	-----

Chapitre 4

Chapitre 4, Figure 1. Detailed representation of the SOFA-HDV ribozyme.	176
--	-----

Chapitre 4, Figure 2. Scheme of the design process for the SOFA-HDV ribozyme.	180
--	-----

Chapitre 4, Figure 3. Details on spacer flexibility and biosensor design.	184
--	-----

Discussion

Discussion, Figure 1. Représentation détaillée du ribozyme SOFA-HDV lié à son substrat.	206
--	-----

Discussion, Figure 2. Représentations de trois ribozymes de l'expérience préliminaire liés au substrat de 11 nucléotides.	216
--	-----

Discussion, Figure 3. Comparaison de l'activité des ribozymes tRNAval-SOFA-HDV-Rz et de leur version SOFA-HDV-Rz.	218
--	-----

Discussion, Figure 4. Comparaison de l'effet de différentes séquences leader. .	219
---	-----

Discussion, Figure 5. Illustrations de la structure secondaire des ribozymes SOFA-HDV avec différents stabilisateurs et de leur activité <i>in vitro</i>	221
Discussion, Figure 6. Expression et distribution cellulaire de trois ribozymes SOFA-HDV avec des stabilisateurs différents.	225

LISTE DES ABRÉVIATIONS

(NH ₄) ₂ SO ₄	Sulfate d'ammonium
[-]sTRSV	<i>Negative strand of the satellite RNA of tobacco ringspot virus</i>
μM	Micromolaire
μm	Micromètre
3'	Extrémité 3'
3' UTR	Région 3' non traduite, <i>3' untranslated region</i>
3'-OH	3'-hydroxyle
5'	Extrémité 5'
5' UTR	Région 5' non traduite, <i>5' untranslated region</i>
5'-OH	5'-hydroxyle
A	Adénosine
ADN	Acidedésoxiribonucléique
AGO	Protéine de la famille Argonaute
APP	<i>Amyloid Protein Precursor</i>
ARN	Acide ribonucléique
ARNdb	ARN double brin
ARNi	Interférence par l'ARN
ARNm	ARN messenger
ARNt	ARN de transfert
ASON	<i>Antisens oligonucleotide</i>
B	C, G ou U
BI	Blocqueur
Bs	Biosenseur
BSA	<i>Bovine serum albumin</i>
C	Cytosine
CFU	<i>Colony-Forming Unit</i>
CPM	<i>Counts Per Minute</i>
DNAzyme	ADN catalytique
dNMP	Déoxyribonucléoside-5'-monophosphate
dNTP	Déoxyribonucléoside-5'-triphosphate
DTT	Dithiothréitol
E.coli	<i>Escherichia coli</i>
EDTA	Acide éthylène diamine tétraacétique
EGS	<i>External Guide Sequence</i>
EXP5	Exportin 5
FBS	<i>Fetal Bovine Serum</i>
FDA	<i>Food and Drug Administration</i>
G	Guanine
GFP	<i>Green Fluorescent Protein</i>

GTP	Guanosine-5'-triphosphate
H	A, C ou U
h	<i>Hour</i>
HCl	Acide chlorhydrique
HCMV	<i>Human cytomegalovirus</i>
Hdag	<i>Hepatitis Delta antigen</i>
HDV	<i>Hepatitis Delta Virus</i>
HEPES	Acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique
HIV	<i>Human Immunodeficiency Virus</i>
HSV-1	<i>Herpes Simplex Virus 1</i>
IFN- α	<i>Interferon-α</i>
IRES	<i>Internal Ribosomal Entry Site</i>
KCl	Chlorure de potassium
kobs	Constante de vitesse observée
KOH	Hydroxyde de potassium
KSHV	<i>Kaposi's Sarcoma-Associated HerpesVirus</i>
LANL	<i>Los Alamos National Laboratory</i>
LTR	<i>Long Terminal Repeat</i>
LTSV	<i>Lucerne Transient Streak Virus</i>
M	Molaire
M1GS	Ribozyme M1 fusionné à un ARN guide
MFE	<i>Minimum Free Energy</i>
Mg ²⁺	Ion magnésium
MgCl ₂	Chlorure de magnésium
MgSO ₄	Sulfate de magnésium
miARN	MicroARN
min	Minute
MOI	<i>Multiplicity Of Infection</i>
N	A, C, G ou U
nm	Nanomètre
nM	Nanomolaire
nt	Nucléotide
NTP	Ribonucléoside-5'-triphosphate
ORF	Open Reading Frame
ex.	Exemple
p ³²	Isotope radioactif, phosphore 32
PAGE	<i>Polycrylamide gel electrophoresis</i>
pb, bp	Paire de base, <i>base pair</i>
pCMV	Promoteur du cytomégalovirus
PCR	<i>Polymerase Chain Reaction</i>
phH1	Promoteur H1 humain
pmol	Picomole
pol	Polymérase

poly(A)	Poly adénosines
pré-ARNm	ARN pré-messenger
pré-ARNt	Précurseur d'ARNt
pré-miARN	Précurseur de microARN
pri-miARN	MicroARN primaire
R	Purine
RD	Domaine de reconnaissance, <i>Recognition Domain</i>
RISC	<i>RNA-Induced Silencing Complex</i>
RLU	<i>Relative Light Unit</i>
RNAi	<i>RNA interference</i>
RNase	Ribonucléase
RQI	<i>RNA Quality Index</i>
RSV	<i>Rous Sarcoma Virus</i>
RT	<i>Reverse Transcriptase</i>
Rz	Ribozyme
S	Substrat
SDS	Dodécylsulfate de sodium
SEM	<i>Standard Error of the Mean</i>
shARN	Petit ARN en épingle à cheveux, <i>small hairpin RNA</i>
siARN	Petit ARN interférent, <i>small interfering RNA</i>
SLIII	<i>Stem-Loop III</i>
SNP	<i>Single-Nucleotide Polymorphism</i>
snRNA	<i>Small nuclear RNA</i>
SOFA	<i>Specific On/oFf Adapter</i>
Sp	Espaceur, <i>spacer</i>
SSC	<i>Saline-Sodium Citrate</i>
TBE	Tampon Tris-Borate-EDTA
Tris	Trishydroxyméthylaminométhane
tRNAval	ARNt de la valine, <i>valine tRNA</i>
U	Uracile
VHC	Virus de l'Hépatite C
VHD	Virus de l'Hépatite <i>Delta</i>
VIH	Virus de l'Immunodéficience Humaine
Y	Pyrimidine

RÉSUMÉ

ÉTUDE DU RIBOZYME SOFA-HDV COMME OUTIL MOLÉCULAIRE : APPLICATION ET OPTIMISATION

Par

Michel Lévesque
Département de Biochimie
Université de Sherbrooke

Thèse présentée à la faculté de médecine et des sciences de la santé en vue de
l'obtention du grade *de philosophiae doctor* (Ph.D.) en Biochimie

Les avancées en biologie moléculaire et cellulaire des dernières décennies ont permis de redéfinir le rôle de l'ARN au sein des cellules de tous les domaines du vivant. Initialement cantonné dans un rôle de support transitoire de l'information génétique du génome (ADN) en direction des effecteurs ou molécules actives (protéines et métabolites), l'ARN est maintenant associé à toutes les sphères de la biologie. Les molécules d'ARN peuvent agir autant comme génome (virus), comme molécules adaptatrices (ARNt), comme messager de l'information génétique (ARNm), comme enzyme (ribozyme) ou encore comme molécules régulatrices en *cis* (riboswitch) ou en *trans* (miARN). Nous savons aussi que la grande majorité du génome des cellules eucaryotes est transcrite à un moment ou un autre. Ces implications de l'ARN en font une cible de choix pour la recherche en génomique fonctionnelle, ainsi que pour des applications thérapeutiques. C'est pourquoi, depuis la découverte des ARN catalytiques et de l'interférence à l'ARN, beaucoup d'efforts ont été consacrés pour développer un éventail d'outils moléculaires permettant d'inhiber l'expression de gènes d'intérêt. Le défi qui se dessine aujourd'hui est le développement d'outils plus spécifiques et plus efficaces, entre autres parce que la variété d'ARN qu'un inhibiteur peut rencontrer est beaucoup plus grande qu'initialement estimée. De plus, les données recueillies lors d'essais cliniques montrent la nécessité de combiner un très grand potentiel avec une spécificité accrue. Les travaux de cette thèse se concentrent sur un outil moléculaire qui a le potentiel de répondre positivement à ce défi : le ribozyme SOFA-HDV. Mon projet de recherche visait à démontrer le potentiel de cet ARN catalytique pour le ciblage de gènes *in cellulo* et de développer son application.

Tout d'abord, j'ai démontré que le ribozyme SOFA-HDV pouvait être utilisé pour inhiber la fonction d'un ARN *in cellulo*. Cette étude a également mis en évidence l'usage de ce ribozyme comme agent antiviral, avec le virus de l'hépatite C comme modèle. Le développement de nouvelles thérapies plus performantes avec peu d'effets secondaires demeure un enjeu important. Au moment de publier ces travaux, en plus d'être le premier exemple exhaustif de l'utilisation du ribozyme SOFA-HDV *in cellulo*, notre étude contenait le plus grand nombre de ribozymes

jamais testés contre le VHC en une seule publication. Bien que modeste, l'effet observé démontre que ce ribozyme peut inhiber la réplication d'un virus dans un modèle *in cellulo*. Nos résultats exposent aussi la différence d'accessibilité entre les ARN de polarités positive et négative du VHC *in cellulo*.

Par la suite, j'ai participé au développement de ribozymes SOFA-HDV ciblant le virus de l'immunodéficience humaine (VIH) dans le cadre d'une collaboration. Parmi les ribozymes testés, nous en avons identifié un dont l'activité catalytique réduit la réplication du VIH de plus de 50 % dans un modèle cellulaire. Dans le cadre de cette étude, nous avons identifié un site hautement favorable pour le ciblage par un ribozyme SOFA-HDV ou par un shARN. Des données suggèrent aussi une spécificité élevée du ribozyme SOFA-HDV. Des tests d'inhibition avec différentes souches du VIH montrent que l'activité du ribozyme est affectée avec un seul mésappariement entre le biosenseur (élément du module SOFA responsable de reconnaissance du substrat) et son site de liaison.

Finalement, dans un esprit d'intégration des connaissances recueillies au fil des différents projets impliquant le ribozyme SOFA-HDV, je me suis intéressé à leur processus de sélection pour le ciblage génique. J'ai démontré l'impact de la séquence du biosenseur sur l'activité du ribozyme. J'ai également illustré l'auto-coupeure possible lorsque la séquence du biosenseur crée un prolongement du bloqueur (élément du module SOFA agissant comme verrou) ainsi que l'impact de la structure du substrat autant au niveau des sites de liaison du domaine de reconnaissance que du biosenseur. Ces nouveaux éléments combinés aux données antérieures sur le ribozyme HDV original m'ont permis d'élaborer une marche à suivre pour la présélection des ribozymes SOFA-HDV selon leur potentiel comme ciseaux moléculaires.

En conclusion, ces travaux ont contribué à mettre de l'avant le potentiel du ribozyme SOFA-HDV pour des applications de ciblage de gènes, plus particulièrement pour des cibles virales. De ce fait, il existe maintenant des exemples concrets de l'utilisation de ce ribozyme en cellules humaines. Tout indique que la spécificité du module SOFA est préservée *in cellulo* et serait avantageusement comparable à d'autres technologies. Globalement, cette thèse devrait rendre l'utilisation du ribozyme SOFA-HDV plus accessible et favoriser son développement comme outil moléculaire.

Mots-clés : Ribozyme, outil moléculaire, méthode de design, ciblage génique, spécificité, virus de l'hépatite C, virus de l'immunodéficience humaine.

INTRODUCTION

1. Modulation de gènes et thérapie génique : cibler l'ARN

L'information permettant aux cellules des organismes vivants de fonctionner est conservée dans leur génome. Ces instructions sont organisées dans l'ADN génomique sous la forme d'une succession de gènes. Par exemple, pour le génome humain, on estime qu'il renferme près de 25 000 gènes codant pour des protéines (International Human Genome Sequencing Consortium, 2004). La nature, l'arrangement et le nombre de ces derniers sont au centre de la diversité des organismes vivants. Par contre, la complexité et la diversité des processus biologiques nécessaires à la vie ne se limitent pas uniquement à l'expression de la partie codante du génome qui représente moins de 2 % de la taille totale de celui-ci (ENCODE Project Consortium et al., 2012). Contrairement à ce que fût la pensée dominante depuis les débuts de la biologie moléculaire et de la génétique, il appert que la très grande majorité du génome est transcrit d'une manière ou d'une autre (ENCODE Project Consortium et al., 2007). Ainsi une panoplie de nouveaux ARN non-codants sont présentement étudiés pour leurs rôles dans différents processus biologiques (pour une revue voir (Gambari et al., 2011; Harries, 2012)). L'intégration de tous ces éléments est la source de l'inimaginable diversité de la vie, mais c'est surtout une condition *sine qua none* pour le maintien de l'homéostasie des organismes vivants, des plus simples bactéries jusqu'aux animaux les plus imposants.

L'équilibre des cellules peut être brisé par le dérèglement d'un ou plusieurs de ces éléments résultant en la mort ou en un changement phénotypique pour les cellules touchées. La régulation aberrante de gènes est au cœur d'un éventail de pathologies diverses comme des cancers, des infections virales, des maladies auto-immunes, des inflammations chroniques ou encore certaines problématiques associées au vieillissement (Abdellatif, 2012; Azmi et al., 2011; Fowler et al., 2011; McGivern et Lemon, 2011; Sun et al., 2012). Ainsi, la surexpression d'un

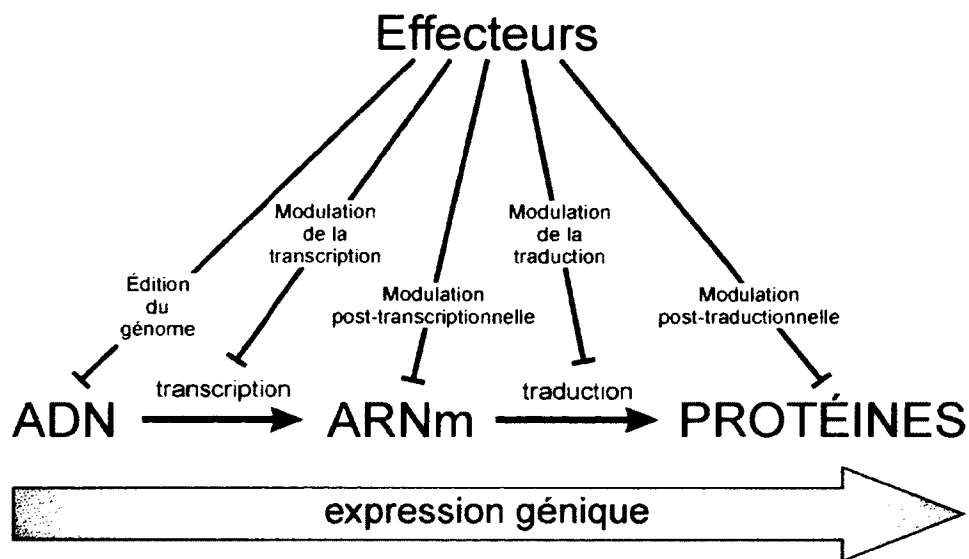
oncogène, par exemple, peut activer la prolifération d'une cellule et lui donner un phénotype cancéreux (Katoh et Katoh, 2009; Myant et Sansom, 2011).

Dans de nombreux cas, il est commun de penser que des méthodes permettant de rétablir le niveau d'expression normal d'un gène peuvent renverser un phénotype anormal et potentiellement offrir une solution à ces pathologies. De plus, le contrôle de l'expression d'un gène d'intérêt est un outil indispensable en recherche fondamentale. Entre autres, cela permet de décortiquer des voies métaboliques et de mettre à jour la complexité de certains processus cellulaires.

1.1. Approches pour modifier l'activité d'un gène

L'activité des gènes peut être modulée à moult niveaux. Entre autres, la chromatine, la transcription, l'ARN (ARN messenger (ARNm) ou ARN non-codant), la traduction ou la protéine sont autant d'éléments dans l'expression d'un gène au niveau desquels une panoplie d'évènements de régulation agissent. Ce sont également autant de cibles potentielles pour affecter l'activité d'un gène (Introduction, Figure 1). Selon l'étape visée, les approches possibles sont bien différentes, chacune ayant ses propres particularités et auront un effet soit sur la quantité, l'activité ou sur l'état du produit final d'un gène (protéine ou ARN dans le cas des ARN non-codants).

Tout d'abord, à l'échelle de l'ADN, il est possible de moduler l'activité transcriptionnelle du gène. Par exemple, des facteurs de transcription artificiels peuvent être développés et permettent de stimuler ou d'inhiber l'expression de gènes cellulaires (Gommans et al., 2005; Katoh et Katoh, 2009). L'ajout d'un nouveau gène et le remplacement ou la suppression d'un gène défectueux directement dans le génome sont également des solutions potentielles pour certaines maladies (Su et al., 2012).



Introduction, Figure 1. Schéma de l'expression génique et des cibles potentielles de modulation.

Le génome (ADN) encode pour des gènes qui sont transcrits (ARNm) puis traduits en protéines. Il existe différents effecteurs, dont certains seront discutés en détails dans cette thèse, pouvant être utilisés pour affecter l'expression d'un gène au long du chemin de l'expression génique. Ainsi, le génome, les ARN et les protéines sont des cibles pour ces différentes technologies, mais aussi les étapes de transfert d'un type de biomolécules à un autre comme la transcription ou la traduction. Inspirée de (Kurzynska-Kokorniak et al., 2009).

Au niveau de l'ARNm, il est possible d'influencer sa stabilité ou sa traduction et ainsi changer la quantité de protéines produites à partir d'un locus (Patil et al., 2008). Pour les ARN non-codants tel que les micro ARN (miARN), en plus d'affecter leur stabilité, il est possible de moduler leur fonction en réduisant leur disponibilité ou encore en empêchant l'arrimage avec leurs cibles ou leurs partenaires cellulaires (Gambari et al., 2011; Tyagi et al., 2011). Dans les deux cas, une grande partie des technologies agissant au niveau de l'ARN sont faites d'acides nucléiques ou de leurs dérivés et seront discutées en détails dans la prochaine section.

Finalement, le produit fini des gènes, les protéines, peuvent être directement visées par des ligands (Xu et al., 2012). Il est entre autres possible de stimuler ou de bloquer l'activation d'une protéine (agoniste et antagoniste), d'inhiber l'activité d'un enzyme (inhibiteur compétitif ou allostérique) ou encore

d'empêcher son interaction avec d'autres protéines cellulaires (bloc stérique) (Hantschel et al., 2012; Sloan et al., 2009; Xu et al., 2012). Ces effecteurs sont généralement de petits composés chimiques ou peptidiques. Dans certains contextes précis, s'ajoute la possibilité d'utiliser des anticorps comme agents de reconnaissance de la protéine cible (Wolchok et al., 2010). Globalement, ces dernières stratégies ne visent pas particulièrement à retirer une protéine d'une cellule, mais bien d'en modifier l'activité ou la fonctionnalité contrairement aux stratégies qui s'attaquent directement à la genèse des protéines au niveau de l'ADN ou de l'ARN.

1.2. Importance & intérêts de cibler l'ARN

De toutes ces avenues, le ciblage des molécules d'ARN est le centre d'intérêt des travaux présentés dans cette thèse. Les principales caractéristiques expliquant le choix de l'ARN comme cible, seront développées dans cette section. Il est important de noter ici que mon intérêt se tourne vers des situations où le but est de moduler, principalement réduire, le niveau d'activité d'un gène contrairement à certaines conditions où il serait nécessaire de modifier son activité (ex. : le remplacement d'un gène défectueux par un allèle fonctionnel).

Tout d'abord, cibler l'ARNm et induire sa dégradation permet de bloquer l'activité du gène en évitant la production de la protéine qu'il encode. De cette façon, il est possible d'annihiler la source du problème avant qu'il n'existe. Contrairement aux stratégies visant à modifier directement le génome telles que la recombinaison ou encore la réduction de l'activité transcriptionnelle du locus via des facteurs de transcription artificiels, la dégradation de l'ARNm peut être spécifique à un ou plusieurs isoformes selon le besoin. En effet, la très grande majorité des gènes transcrits dans une cellule humaine sont épissés alternativement (Kelemen et al., 2013). Ainsi, lorsque l'on inhibe la transcription au niveau du promoteur, tous les messagers produits à ce site seront affectés et ultimement l'expression d'un groupe de protéines sera perturbée. De la même

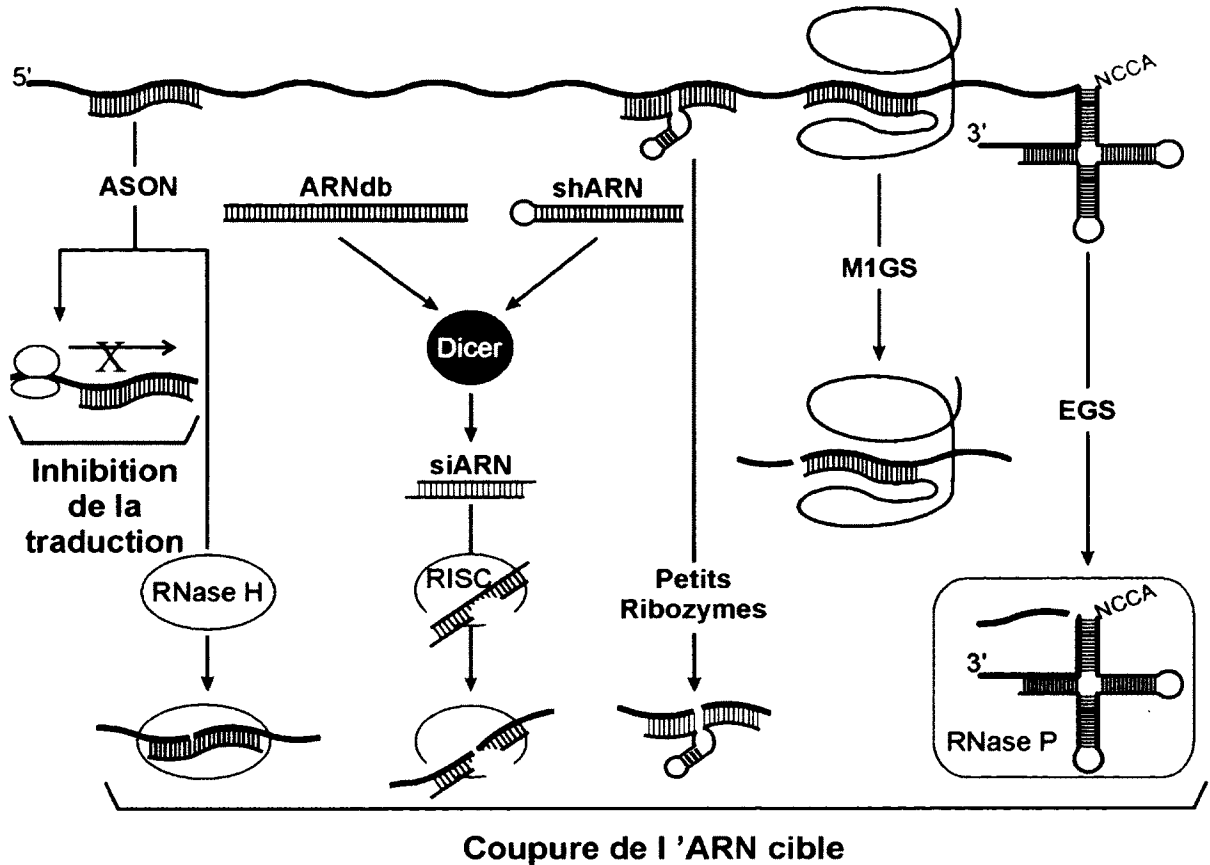
façon, le remplacement du locus par recombinaison pourrait déséquilibrer la proportion de chaque isoforme ou même empêcher complètement l'expression de certaines. Avec un choix judicieux de l'endroit visé dans l'ARNm, il est possible d'éviter ces problématiques. Par exemple, la jonction entre deux exons d'un ARNm mature, spécifique à un isoforme, permet de le discriminer par rapport aux autres messagers produits à partir du même pré-ARNm (Robichaud et al., 2008). Si le site choisi est spécifique à un sous-groupe d'isoformes, il sera alors possible d'influencer l'expression d'un groupe précis d'isoformes provenant de la même région du génome. Cette logique s'applique également aux cas de gènes fusionnés que l'on retrouve par exemple dans certaines cellules cancéreuses (Cobaleda et Sanchez-Garcia, 2000).

Il est possible de cibler spécifiquement l'ARNm d'une protéine partageant une très grande similitude avec une autre, alors qu'au niveau protéique cela demeure un défi presque inatteignable dans certains cas. L'une des explications possibles vient du fait que deux séquences identiques d'acides aminés n'impliquent pas obligatoirement des séquences nucléotidiques conservées. Ainsi, les éléments essentiels des sites actifs d'enzymes très similaires n'engendrent pas nécessairement des séquences géniques identiques, mais seront difficiles à discriminer au niveau peptidique. De la même façon, l'identification d'un anticorps spécifique pour un épitope d'intérêt n'est pas une simple tâche. Il peut aussi arriver que les anticorps réagissent à des motifs peptidiques partagés par plusieurs protéines. La réactivité croisée représente donc une limitation importante étant donné le manque d'outils pour la contrôler. À l'opposé, les technologies pour cibler l'ARN sont moins sujettes à ces limitations étant donné que la majorité d'entre elles reconnaissent leur cible ARN via la formation de paires de bases (pb) Watson-Crick (Bennett et Swayze, 2010). De plus, il est possible de penser à des méthodes simples et efficaces pour prédire et contrôler leur spécificité (Lucier et al., 2006). Au final, cette analyse pourrait être couplée à l'étape du design et de la présélection des molécules effectrices, ce qui n'est pas le cas pour les inhibiteurs agissant au niveau des protéines.

Finalement, les travaux des dernières années sur le transcriptome cellulaire et surtout sur les mécanismes de régulation qui y sont associés ont exposés une nouvelle série de cibles thérapeutiques potentielles («*druggable target*») (Bennett et Swayze, 2010). Au-delà de tous les ARNm codants pour des protéines, de plus en plus d'ARN non-codants sont mis en lien avec des pathologies (Gambari et al., 2011; Harries, 2012). La plupart des méthodologies développées pour inhiber ou induire la dégradation des ARNm sont généralement adaptables aux ARN non-codants. Ainsi, il est possible de viser spécifiquement tous les ARN produits dans une cellule ce qui représente un nombre impressionnant de cibles autant dans une optique de thérapie que pour la recherche fondamentale. Bien que les petits composés chimiques représentent à ce jour la majorité des molécules pharmaceutiques, les stratégies comme celles ciblant les ARN sont une alternative dont le potentiel reste encore à être exploré.

2. Les outils moléculaires d'acides nucléiques

Les outils moléculaires disponibles pour inhiber ou induire la dégradation d'ARN sont majoritairement faits d'acides nucléiques ou d'un dérivé de ceux-ci (Introduction, Figure 2). Pour certaines approches, la simplicité de prédiction des interactions par paires de bases est l'élément clé alors que pour d'autres c'est la sélection *in vitro* qui est la pierre angulaire de l'approche. La production de petits composés chimiques affectant l'activité d'ARN d'intérêt demeure à ce jour une problématique complexe et difficile à résoudre par sélection traditionnelle ou par design rationnel ((Bennett et Swayze, 2010). De leur côté, les approches décrites dans cette section peuvent être générées à la fois par design rationnel ou par sélection. Elles sont considérées comme plutôt rapides et efficaces à produire permettant l'essai de molécules tests à l'intérieur de courts délais.



Introduction, Figure 2. Mécanismes d'action simplifiés des principaux outils d'acides nucléiques pour le ciblage d'ARN.

Les oligonucléotides antisens («*antisense oligonucleotides*», ASON) bloquent la traduction de l'ARNm ou recrutent l'activité RNase H cellulaire pour couper l'ARNm. Les petits ARN interférents (siARN) induisent la coupure de l'ARN cible via l'activité catalytique du complexe RISC auquel ils sont incorporés. Ces siARN peuvent être produit à partir d'ARN double brin (ARNdb) ou d'ARN «hairpin» (shARN). Les petits ribozymes et le ribozyme de la RNase P de *E. coli* (M1GS) reconnaissent et coupent leur ARN cible sans la participation de facteurs cellulaires. Finalement, l'activité endogène de la RNase P cellulaire est dirigé spécifiquement vers un ARN d'intérêt par l'entremise d'un ARN guide (EGS). Modifiés de (Asif-Ullah et al., 2007).

Pour les besoins de cette section, les outils moléculaires les plus significatifs seront répartis en deux groupes, selon que l'effecteur ait ou pas une activité catalytique (Introduction, Figure 2). Ainsi, dans la catégorie des outils moléculaires non catalytiques se retrouveront les approches d'oligonucléotides antisens («*antisense oligonucleotides*», ASON) et de l'interférence à l'ARN (ARNi). Dans un

second temps, l'éventail des *ciseaux moléculaires* faits d'acide nucléique (plus particulièrement faits d'ARN) et possédant leur propre activité catalytique sera introduit. Des notions importantes pour le ciblage génique par des outils d'acides nucléiques seront présentées au fil de ces sections. L'idée est d'aborder ces éléments généralisables par l'entremise d'exemples associés à un outil ou à un autre.

2.1. Outils non catalytiques

Il existe plusieurs technologies différentes basées sur des molécules effectrices d'acides nucléiques non catalytiques. Les plus communes sont l'utilisation d'oligonucléotide antisens et de petits ARN interférents pour inhiber l'expression d'un gène (Drude et al., 2007; Scherer et al., 2007). Des acides nucléiques antisens peuvent également être employés pour moduler l'épissage d'un gène (Du et Gatti, 2009). Finalement, la liaison d'aptamères sélectionnés *in vitro* peut être exploitée pour inhiber l'activité d'un gène (Zhou et al., 2012). De par la nature de mes travaux, je me concentrerai sur l'inhibition de l'expression de gène par les oligonucléotides antisens et les petits ARN interférents.

2.1.1. Oligonucléotides antisens (ASON)

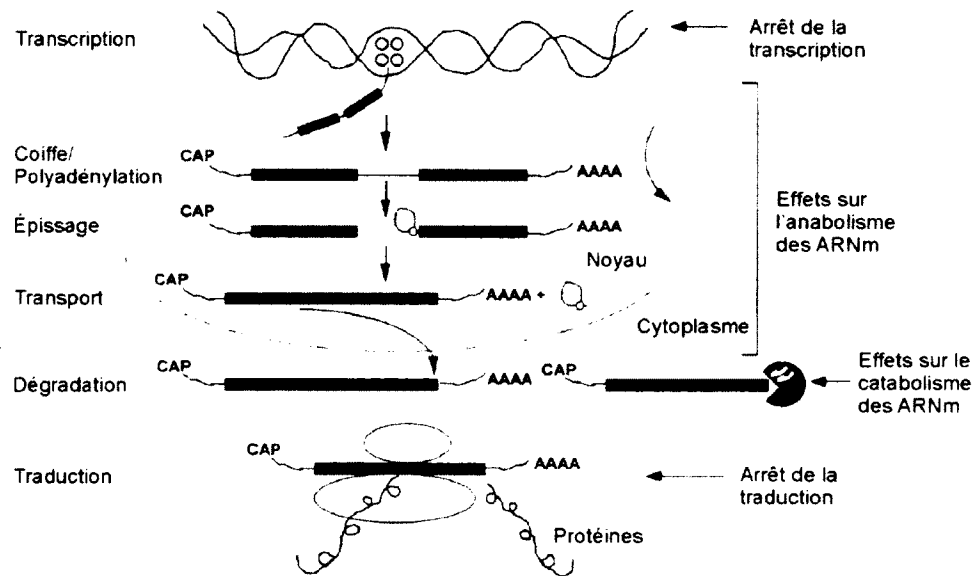
La première utilisation d'oligodésoxyribonucléotides comme outil moléculaire pour l'inhibition d'un gène remonte à la fin des années 70. En utilisant des oligodésoxyribonucléotides de 13 nucléotides complémentaires (antisens) aux séquences terminales répétées de la longue répétition terminale («*long terminal repeat*», LTR) du virus du sarcome de Rous («*Rous sarcoma virus*», RSV), Zamecnik et Stephenson furent les premiers à mettre de l'avant le concept d'«*antisens oligonucleotide*» (ASON) (Stephenson et Zamecnik, 1978; Zamecnik et Stephenson, 1978). Avec leurs travaux, ils ont pu démontrer que la traduction *in vitro* en extrait cellulaire ainsi que la réplication du RSV *in cellulo* pouvait être

inhibées par les ASON. Ce n'est qu'avec l'automatisation de la synthèse des oligonucléotides au cours des années 80 que le plein potentiel de cette approche a commencé à se matérialiser. La prémisse au cœur de cette méthode repose sur la complémentarité (pb Watson-Crick) entre un oligomère nucléotidique (oligonucléotide), initialement fait d'ADN, et un acide nucléique d'intérêt, tel qu'un ARNm (Introduction, Figure 2). La formation du duplex permet donc l'inhibition du transfert de l'information génétique de l'ADN vers les protéines via différents mécanismes qui seront décrits plus loin. Cette façon de faire peut être utilisée dans une approche de design rationnel et demeure relativement peu coûteuse selon la nature de la chimie incorporée dans ces molécules effectrices (synthèse chimique d'oligonucléotide relativement court, généralement entre 13 et 25 nt) (Dias et Stein, 2002). Avec les connaissances de plus en plus étoffées sur la séquence du génome humain ainsi que sur celles de plusieurs pathogènes d'importance, cette technologie revêt un potentiel fort intéressant comme agent thérapeutique et comme outil pour l'étude fonctionnelle de gènes. D'autant plus que, virtuellement, cette stratégie devrait permettre de cibler spécifiquement n'importe quelle séquence d'intérêt.

Il existe principalement deux classes d'oligonucléotides antisens définies selon le mécanisme utilisé pour moduler l'expression génique. Le principal mode d'action des ASON repose sur le recrutement de l'activité ribonucléasique RNase H. En fait, ce mécanisme d'action est le mieux étudié de cette approche. D'ailleurs, c'est sur cette activité que repose la majorité des drogues en développement qui utilisent des ASON (Bennett et Swayze, 2010; Dias et Stein, 2002). L'enzyme RNase H reconnaît les hétéroduplexes ADN/ARN et induit l'hydrolyse de la molécule d'ARN tout en relâchant la molécule d'ADN intacte (Introduction, Figure 2) (Eder et Walder, 1991). La coupure de l'ARN génère des extrémités 5'-phosphate et 3'-OH. Il est important de noter que l'activité d'hydrolyse sera possible seulement si l'oligonucléotide est composé d'ADN ou de certains de ses dérivés chimiques. Ces modifications chimiques au niveau des oligonucléotides antisens seront traitées un peu plus loin dans cette section. La famille d'enzyme RNase H est présente dans toutes les cellules de mammifères (expression

ubiquitaire) signifiant que l'inhibition via la stimulation dirigée de cette activité pourrait être accomplie dans tous les types cellulaires humains (Cerritelli et Crouch, 2009; Dias et Stein, 2002). Chez l'humain, il existe deux types de RNase H, soit les RNase H1 et RNase H2. À priori, leurs rôles physiologiques se retrouveraient au niveau de la réplication et de la réparation de l'ADN génomique. La RNase H est retrouvée à la fois au niveau du cytoplasme et du noyau de la cellule supportant l'hypothèse d'une implication autre qu'à proximité de l'ADN génomique (Crooke, 1999). Dans le cas des ASON, il a été démontré que c'est la RNase H1 qui est responsable de la coupure des ARN cibles (Wu et al., 2004). Virtuellement, l'hydrolyse par la RNase H peut être dirigée n'importe où sur l'ARNm cible puisque cela dépend uniquement de la formation de l'hétéroduplexe entre l'oligonucléotide et sa cible. C'est d'ailleurs une différence notable avec la seconde classe d'ASON. En effet, les oligonucléotides de cette deuxième catégorie agissent comme un bloqueur stérique perturbant un processus biologique nécessaire à l'expression du gène cible. La liste de processus potentiellement modifiable par la liaison d'un antisens à un ARN est très longue. (Introduction, Figure 3) Sommairement, il est possible d'inhiber l'épissage d'un pré-ARNm, de bloquer la traduction d'un ARNm ou encore de détruire une structure essentielle pour la fonction de l'ARN cible simplement par l'occupation d'une région de l'ARN par l'oligonucléotide antisens. À cela s'ajoute la possibilité d'induire la déstabilisation de l'ARN cible via l'appariement de l'oligonucléotide. Par exemple, l'ajout de la structure coiffe à l'extrémité 5' ou de la queue de poly adénosines (poly(A)) en 3' des messagers peuvent être inhibés par la liaison d'un oligonucléotide entraînant une dégradation rapide des ARNm ciblés. Pour générer ces effets, l'appariement de l'antisens doit se faire à des endroits bien précis de l'ARN visé (Bennett et Swayze, 2010; Crooke, 1999). L'exemple le plus courant et probablement le plus efficace est le blocage de l'initiation de la traduction qui évidemment sera plus efficace lorsqu'une séquence près du codon d'initiation sera liée par l'oligonucléotide qu'ailleurs dans les exons. Malgré un mécanisme d'action proposé rationnellement acceptable, il y a peu d'évidences qui le démontrent *in cellulo* et surtout *in vivo*. L'utilisation d'oligonucléotides antisens comme bloc

stérique demeure encore à ce jour moins répandu, surtout du côté thérapeutique. Cette limitation provient en partie de la plus grande difficulté encourue pour la détection des protéines (primordiale dans les cas où le niveau d'ARNm ne varie pas) et du nombre de site de liaison potentiel plus restreint pour l'action par bloc stérique (Bennett et Swayze, 2010).



Introduction, Figure 3. Schéma de la maturation des ARNm et des cibles potentielles des ASON.

Les oligonucléotides antisens ont le potentiel d'agir à toutes les étapes de la vie des ARNm. Ainsi, la transcription et l'ensemble des étapes de maturation de pré-ARNm en ARNm sont des cibles potentielles pour l'action des ASON. Ces effecteurs peuvent également stimuler la dégradation des ARNm matures ou simplement réprimer la traduction. Adapté de (Crooke, 1999).

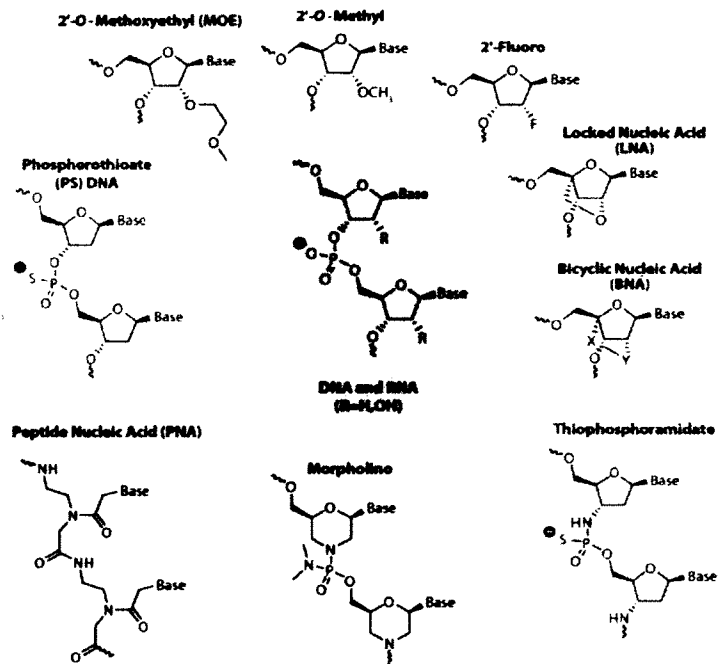
Les oligonucléotides d'ADN discutés jusqu'à maintenant ont un potentiel restreint dans un environnement biologique en raison de leur sensibilité aux endonucléases et aux exonucléases, généralement 3' vers 5' (Akhtar et al., 1991; Eder et al., 1991; Wickstrom, 1986). La dégradation de ces oligomères d'ADN réduit la quantité de molécules disponibles pour générer l'effet antisens spécifique au gène d'intérêt, mais en plus il a été démontré que les produits de dégradation pouvaient être responsables de changements phénotypiques au niveau des cellules traitées. En fait, certains produits de dégradation comme les dNMP

(désoxyribonucléoside-5'-monophosphate) semblent cytotoxiques et exercent un effet antiprolifératif sur les cellules (Vaerman et al., 1997)

Dans le but de trouver des solutions à ces problèmes d'effets indésirables, mais aussi pour augmenter le potentiel, l'efficacité et la spécificité des oligonucléotides antisens, l'incorporation d'une multitude de modifications chimiques fut entreprise. Ainsi, il existe maintenant une kyrielle de modifications chimiques pouvant être incluses dans les oligonucléotides. Certains types sont dérivés de molécules d'ADN ou d'ARN alors que d'autres sont constitués d'un squelette plus artificiels. Il est important de noter que la nature de la chimie incorporée dans un oligonucléotide a un impact majeur sur son mode d'action et sur la méthode de livraison des molécules actives à l'intérieur des cellules (Bennett et Swayze, 2010; Dias et Stein, 2002). (Introduction, Figure 4)

Les oligonucléotides antisens les plus étudiés, en raison de leur grande stabilité face aux nucléases, présentent un squelette de phosphorothioates. De plus, leur synthèse est relativement simple et consiste à remplacer un atome d'oxygène, non impliqué dans le lien phosphodiester, par un atome de soufre au niveau de chaque phosphate du squelette. Les oligomères phosphorothioates ont tendance à déstabiliser l'hélice du duplexe oligonucléotide-ARN. Par contre leur solubilité est très élevée et ils possèdent une excellente activité antisens. Il est important de noter que cette chimie, dérivée de l'ADN, permet de recruter l'activité RNase H contre l'ARN cible. Plusieurs études *in cellulo* et *in vivo* ont démontré le potentiel de cette technologie et ont permis d'amener certains candidats à l'étape des essais cliniques. D'ailleurs la première, et unique, drogue antisens approuvée par la «*Food and Drug Administration*» (FDA, Rockville, MD, USA), le Vitravene (Isis Pharmaceuticals, Carlsbad, CA, USA) aussi connue sous le nom de Fomivirsén, est issue de ce type de squelette (Bennett et Swayze, 2010; Dias et Stein, 2002; Roehr, 1998). Malgré un certain succès, les phosphorothioates ont été démontrés pour déclencher des effets non spécifiques, liés entre autres à leur interaction de façon longueur dépendante avec certaines protéines cellulaires (Benimetskaya et al., 1997; Fennewald et Rando, 1995; Guvakova et al., 1995;

Kitajima et al., 1999; Stein, 1997). Il a même été proposé que ces effets non spécifiques puissent contribuer à un phénotype anti-cancer. Bref, en dépit de ces effets non-spécifiques, les phosphorothioates demeurent les oligonucléotides les plus étudiés en essais cliniques. De plus la plupart des molécules antisens présentement en développement contiennent en tout ou en partie cette modification chimique (Bennett et Swayze, 2010).



Introduction, Figure 4. Illustration des modifications chimiques utilisées dans la synthèse des ASON.

L'ADN et l'ARN diffèrent seulement au niveau de la position 2' du ribose indiqué ici par le R (cercle). Dans l'encadré du haut (vert pâle) sont présentés les chimiques qui contiennent un lien phosphodiester alors que dans la partie du bas (bleu pâle) sont regroupés celles où le lien entre les nucléotides est complètement différent de l'ADN ou l'ARN. Tirée de (Bennett et Swayze, 2010).

2.1.2. Interférence par l'ARN

L'interférence par l'ARN est un processus biologique utilisant de l'ARN double brin (ARNdb) pour guider le ciblage spécifique d'un ARNm et ainsi inhiber l'expression de la protéine correspondante. Cette machinerie de régulation est présente dans la

majorité des cellules eucaryotes. La découverte et la caractérisation récente (Fire et al., 1998) de ce mécanisme de régulation génique dépendant de l'ARN a permis aux professeurs Craig Mello et Andrew Fire d'être honorés du prix Nobel de médecine 2006 (Nobel Media AB. , 2006). L'expérience clé de leurs travaux fut d'introduire des ARNdb dans le nématode *Caenorhabditis elegans* et d'observer des phénotypes associés à une perte d'expression du gène homologue à ces ARNdb. En résumé, ils ont pu déduire que l'ARNdb pouvait atténuer l'expression d'un gène, que le gène affecté et l'ARNdb injecté dans l'animal devait être complémentaire. Ils ont aussi proposé que l'interférence à l'ARN devait être un processus catalytique puisque l'injection d'une très faible quantité d'ARNdb était suffisante pour induire le phénotype d'inhibition. Leurs travaux fournissaient également une possible explication à un phénomène observé chez les plantes depuis plusieurs années nommé «*postranscriptional gene silencing*». Finalement, ils ont même pu spéculer sur la potentielle utilisation de l'ARNdb par les organismes pour la régulation physiologique de gènes, proposition que l'on pourrait maintenant associer à la régulation par les microARN (miARN).

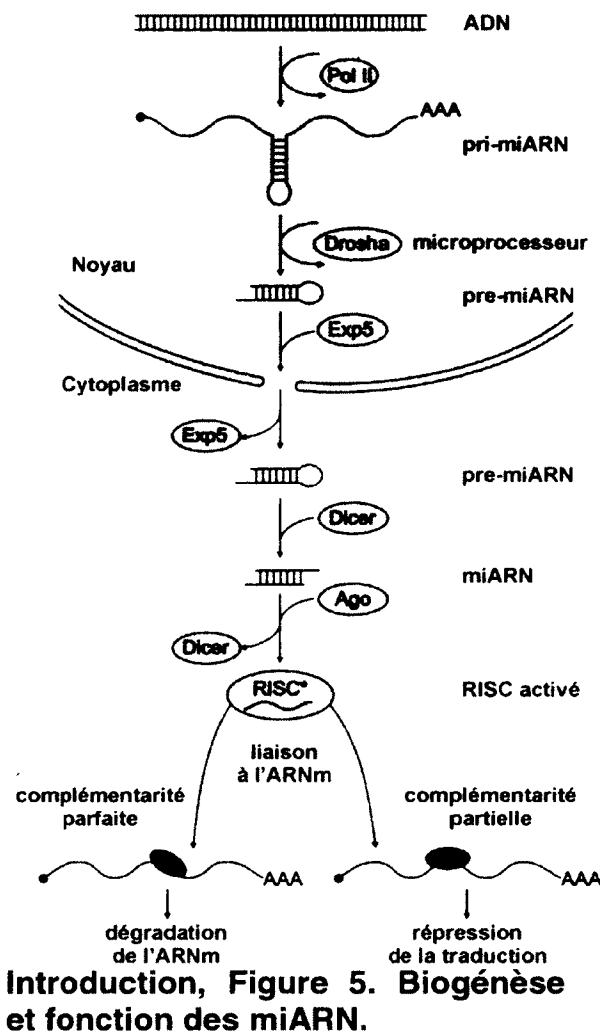
Depuis ce temps, leurs travaux et ceux de bien d'autres groupes de recherche ont permis de mieux comprendre ce processus cellulaire et d'en développer une série d'applications. Dans les segments qui suivent, le mécanisme général de l'ARN interférence tel que retrouvé chez les cellules de mammifères sera décrit. Par la suite, les deux principales méthodes utilisées afin de tirer parti de cette machinerie cellulaire pour l'inhibition de gènes seront sommairement exposées tout en soulignant certaines particularités de chaque approche.

Mécanisme général de l'interférence par l'ARN

L'interférence par l'ARN (ARNi, de l'anglais «*RNA interference*» ou *RNAi*) désigne le processus cellulaire ayant comme joueur clé l'ARN et qui crée de «l'interférence» dans l'expression d'un ou plusieurs gènes. Tel que mentionné précédemment, la source voir le signal induisant ce mécanisme biologique est

l'ARNdb. Cet ARNdb peut être d'origines diverses (endogènes ou exogènes) et son incorporation dans cette voie métabolique est influencée par différents éléments. La formation d'une structure d'ARN en hélice, majoritairement formée de paires de bases Watson-Crick, représente la base de cette signature. Il est possible de séparer en deux grandes catégories les ARN interférents pouvant stimuler ce mécanisme cellulaire chez les mammifères. D'un côté, il y a les miARN qui sont des substrats endogènes de la machinerie d'interférence par l'ARN. La cellule les exploite pour assurer la régulation de ses propres gènes. De l'autre côté, il y a les petits ARN interférents (siARN, de l'anglais «*small interfering RNA*» ou *siRNA*), qui peuvent être d'origine endogène ou exogène. Ces siARN sont ordinairement produits à partir de molécules d'ARN formant de plus longs duplexes d'ARN.

Malgré qu'il y ait des différences notoires entre les deux groupes de petits ARN, les destinées des miARN et des siARN s'entrecroisent à l'intérieur du mécanisme global d'interférence par l'ARN. Les miARN sont transcrits dans le noyau de la cellule à partir de son génome. Généralement transcrits par l'ARN polymérase II, les miARN primaires (pri-miARN) peuvent être longs de quelques kilobases (Lee et al., 2004). Ces transcrits sont ensuite maturés dans le noyau en précurseurs de miARN (pre-miARN) par l'endonucléase de type RNase III, Drosha, du complexe Drosha-DGCR8 connu sous le nom de microprocesseur (de l'anglais «*microprocessor*», Introduction, Figure 5). Les pre-miARN résultants forment des tiges boucles de 60 à 70 nucléotides de long incluant de nombreux mésappariements et boucles internes dans leur région double brin (Denli et al., 2004; Han et al., 2004; Lee et al., 2003). Ces pre-miARN sont activement transportés au cytoplasme via le transporteur dépendant du GTP, Exportin 5 (EXP5) (Lund et al., 2004).

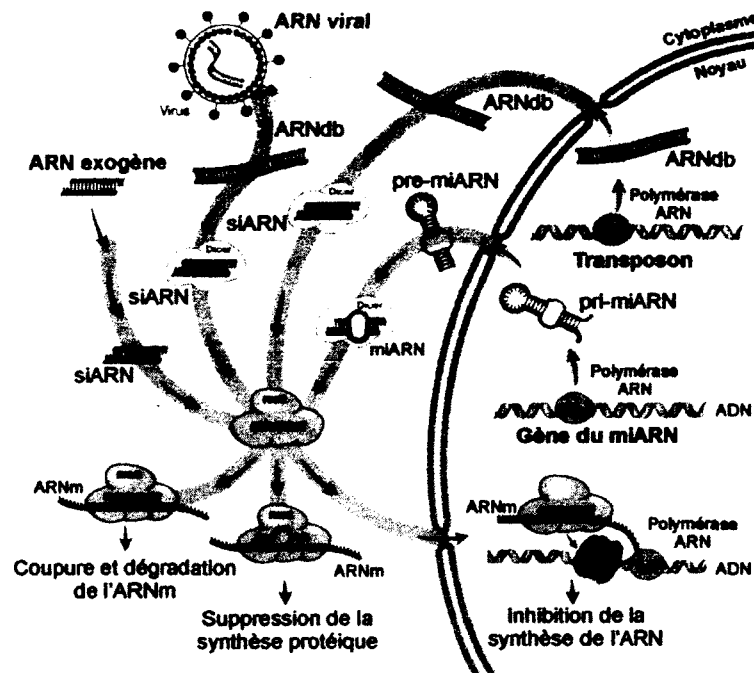


Les miARN sont produits par un processus comprenant de multiples étapes débutant par la transcription d'un ARN contenant une portion double brin (dont une tige-boucle locale). Ce miARN primaire (pri-miARN) est généralement le produit de l'ARN polymérase II (coiffe et queue poly(A)). Par la suite, le précurseur de miARN (pre-miARN) de 50 à 70 nt (tige-boucle) sera excisé par la RNase Drosha (complexe microprocesseur). Son transport au cytoplasme est assuré par l'Exportin-5 (EXP5). À ce moment, le pre-miARN sera coupé par la RNase Dicer et il y aura relâche de court miARN mature sous la forme de duplexe. Ces duplexes sont ensuite incorporés dans le «*RNA induced silencing complex*» (RISC) et un brin sera sélectionné pour diriger la spécificité de l'activité du complexe. Dépendamment de la complémentarité du miARN avec l'ARNm cible ainsi que de la nature du RISC (plusieurs protéines AGO différentes), il y aura dégradation de l'ARNm ou répression de la traduction. Modifiée de (Kurzynska-Kokorniak et al., 2009).

Une fois au cytoplasme, les pre-miARN en forme d'épingle à cheveux («*hairpin*») sont reconnues par la protéine Dicer. Cette protéine lie des ARNdb présents dans la cellule et les aiguille dans la voie de l'ARNi. Tout comme Drosha, Dicer est une endonucléase de type RNase III agissant sous la forme d'un dimère (Zhang et al., 2004). De cette activité résulte la production de petits ARN interférents (siARN). C'est à cette étape que les miARN et les siARN empruntent un chemin commun (Introduction, Figure 6) (Hutvagner et al., 2001). En effet, c'est l'enzyme Dicer qui est responsable de retirer la boucle des pre-miARN pour produire des duplexes prêt à être utilisé dans l'étape suivante (Zhang et al., 2004).

Le mot siARN réfère également à des ARN qui entreraient dans la voie de l'ARNi par l'entremise de la protéine Dicer ou directement dans le complexe RISC («*RNA-Induced Silencing Complex*»). Les miARN/siARN produits par Dicer sont des duplexes d'ARN comprenant entre 19 et 25 paires de bases (Han et al., 2004; Hutvagner et al., 2001; Lee et al., 2003; Zhang et al., 2004). Ces duplexes portent une signature caractéristique issue de la coupure par Dicer, c'est-à-dire des extrémités 3' avec deux nucléotides protubérants (Zhang et al., 2004). Ces petits ARN sont ensuite transférés à l'un des complexes multiprotéiques RISC. À ce moment, un seul brin du duplexe miARN ou siARN est transmis au RISC et devient le brin guide. Sa contrepartie est désignée comme étant le brin passager et il est généralement dégradé après le chargement du RISC (Rand et al., 2005). La sélection du brin guide repose principalement sur la stabilité thermodynamique des extrémités du duplexe de petits ARN (Khvorova et al., 2003; Schwarz et al., 2003). L'extrémité 5' la moins stable sera sélectionnée et le brin guide sera lié au domaine PIWI de la protéine Ago2. C'est la complémentarité entre le brin guide et l'ARN cible qui dicte la spécificité. La composante majeure du complexe RISC est une protéine de la famille Argonaute (AGO). Chez l'humain, la protéine Ago2 est la seule protéine connue de la famille qui possède l'activité catalytique permettant de couper l'ARN cible («*slicer activity*») (Liu et al., 2004; Meister et al., 2004). Ce sont donc les complexes RISC renfermant la protéine Ago2 qui sont les plus importants pour le ciblage spécifique d'ARNm tel qu'il est discuté dans cette thèse. En règle générale, pour induire la coupure de l'ARNm via l'activité «*slicer*», il doit y avoir une complémentarité quasi parfaite entre l'ARN guide (siARN) et l'ARNm. Dans les cas où l'ARNm et le brin guide ne sont que partiellement complémentaire, il peut y avoir inhibition de la traduction du messenger sans qu'il ne soit coupé par le complexe RISC (pas d'activité «*slicer*») (Meister et Tuschl, 2004; Tang, 2005). D'ailleurs, les miARN utilisent principalement l'inhibition de la traduction pour réguler l'expression de leurs gènes cibles. Dans cette éventualité, il peut tout de même y avoir déstabilisation de l'ARNm après l'arrêt de la traduction entre autre via le recrutement de facteurs qui s'attaquent à la structure coiffe de l'extrémité 5'

et la queue de poly(A) à l'extrémité 3' de l'ARNm (Bagga et al., 2005; Fabian et al., 2009; Fabian et al., 2010). La reconnaissance de ces ARN cibles repose en bonne partie sur l'appariement de la région 5' du petit ARN. Plus précisément, se sont les nucléotides 2 à 8, région appelée la «*seed region*» qui fixent la spécificité d'un miARN pour ses cibles (Lewis et al., 2005). Les autres positions peuvent être partiellement complémentaires à leurs cibles (Bagga et al., 2005). Ainsi, un miARN peut reconnaître et moduler un large éventail de gènes cibles. De la même façon, un ARNm peut contenir des sites de liaisons pour plusieurs miARN différents ce qui permet une très grande complexité de régulation génique via la voie des miRNA. L'addition de sites de liaisons pour un miARN donné augmente aussi la sensibilité de l'ARNm face à ce miARN. Plus il y a de possibilités pour un miARN de se lier à un messenger donné, plus l'expression de ce messenger sera affectée par une variation de la quantité cellulaire du miARN (Lewis et al., 2005). Il est important de noter que les miARN connus jusqu'à maintenant ciblent majoritairement les régions 3' non traduites (3' UTR) des ARN messagers (Bartel, 2004). En revanche, lorsqu'il y a complémentarité parfaite entre le brin guide et leur cible, les ARN interférents (particulièrement les siARN) sont capables de réguler l'expression d'un messenger via un site de liaison unique pouvant être situé n'importe où sur l'ARNm mature (Elbashir et al., 2001).

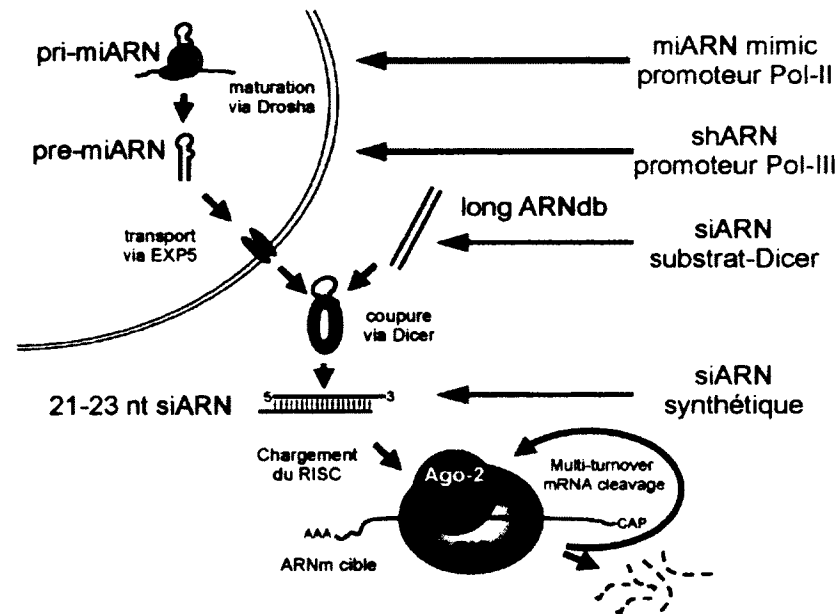


Introduction, Figure 6. Vue d'ensemble de l'interférence par l'ARN.

La protéine Dicer et le complexe RISC jouent un rôle central dans le fonctionnement de l'interférence par l'ARN. Les ARN double brins (ARNdb) d'origines diverses (p.ex. : miARN et siARN synthétiques) convergent vers ces deux éléments clés et inhibent l'expression de leurs cibles via des mécanismes similaires. Cette figure est un schéma ainsi, les complexes de Dicer et RISC peuvent varier selon le processus cellulaire. Adaptée de (Nobel Media AB., 2006).

L'approche des siARN

La première méthode pour stimuler la voie d'interférence par l'ARN consiste à introduire directement dans les cellules cibles des duplexes d'ARN (siARN) pouvant être reconnus et incorporés dans ce mécanisme. Ainsi, il sera possible d'inhiber l'expression d'un gène d'intérêt selon la complémentarité entre le siARN et l'ARNm visé (Introduction, Figure 7).



Introduction, Figure 7. Illustration des différents points d'entrée dans la voie de l'ARNi pour le ciblage génique.

Le schéma résume la voie de l'interférence par l'ARN (ARNi) chez les mammifères (en noir). Les points d'entrée pour les techniques utilisant des petits ARN interférents (siARN) sont précisés à droite (bleu) ainsi que pour les techniques basées sur la production de shARN («*short hairpin RNA*») à partir d'ADN artificiel (rouge). Adaptée de (Aagaard et Rossi, 2007).

Généralement, les duplexes sont longs de 21-22 pb additionnés de deux nucléotides protubérants en 3' de chaque brin (Hutvagner et al., 2001). Ils imitent donc les produits de coupures de Dicer pour prendre contact avec la machinerie cellulaire d'interférence par l'ARN. De études suggèrent que des duplexes plus long (25-27 pb) seraient plus efficaces pour induire l'inhibition des gènes cibles (Kim et al., 2005; Rose et al., 2005; Siolas et al., 2005). Leur liaison suivie d'une coupure par Dicer faciliteraient leur incorporation au complexe RISC. Le brin guide sera sélectionné et chargé par le complexe RISC selon les mêmes règles que pour les petits ARN endogènes (Introduction, Figure 7). Une complémentarité parfaite du brin guide avec sa cible permettra à la protéine Ago2 d'induire la coupure du messenger approximativement au centre de la région appariée. Après coupure, le

complexe RISC chargé pourra répéter l'opération puisque l'ARN guide demeure intact alors que seul l'ARN cible subit l'activité «*slicer*» (Aagaard et Rossi, 2007; Lares et al., 2010; Singh et al., 2011).

Dans un contexte d'expériences *in cellulo*, cette façon de faire représente une option intéressante, simple et efficace pour inhiber un gène d'intérêt. Les critères pour le choix des séquences à utiliser ont été améliorés par le passé (Amarzguioui et Prydz, 2004; Heale et al., 2005; Khvorova et al., 2003; Lee et al., 2003; Patzel et al., 2005; Reynolds et al., 2004; Scaggiante et al., 2011). La plupart du temps, un nombre restreint de molécules doit être essayé avant d'obtenir l'effet recherché sans toutefois garantir l'identification de molécules hyperfonctionnelles (Aagaard et Rossi, 2007)(Khvorova et al., 2003; Lee et al., 2003). Il demeure tout de même certaines limites et points à améliorer pour le travail *in vivo* ou pour une utilisation thérapeutique de cette technologie. En laboratoire, la livraison des siARN se fait le plus souvent par transfection via la formation de complexes liposomes-siARN (Aagaard et Rossi, 2007). Très peu de spécificité pour un type cellulaire précis peut être obtenue de cette façon et donc le potentiel de cette approche pour les thérapies reste minime sans avancement technologique. Beaucoup de modes de livraison ont été ou sont présentement développés pour acheminer des siARN (et autres acides nucléiques) aux cellules cibles à l'intérieur d'un animal. En bref, l'idée est de complexer les siARN avec des molécules permettant la liaison aux cellules cibles et d'induire leur entrée (pour une revue voir (Lares et al., 2010; Singh et al., 2011)). Outre la livraison, qui représente la principale limite pour le développement de thérapies, la stabilité des siARN influence grandement leur activité. Ce concept est encore plus important lorsqu'il est question de les injecter dans le sang d'un animal par exemple. À l'instar des oligonucléotides antisens présentés dans la section précédente, de nombreuses modifications chimiques ont été développées pour protéger les siARN de la dégradation et pour améliorer leur potentiel (Introduction, Figure 4). La panoplie de dérivés chimiques maintenant associés aux siARN a également permis d'atténuer

certains effets négatifs de ces macromolécules. Ainsi, certaines modifications contribuent à réduire la reconnaissance des siARN par le système immunitaire et la toxicité pour les cellules alors que d'autres augmentent la spécificité des siARN (pour une revue voir (Aagaard et Rossi, 2007; Lares et al., 2010; Singh et al., 2011)). Ce dernier élément est très important puisque de nombreuses études ont démontré la présence d'effets non-spécifiques, «*off-target effects*» (Fedorov et al., 2006; Jackson et al., 2003). Ces ciblage non-spécifiques sont principalement associés à trois éléments. D'abord, à l'image des miARN, l'appariement partiel du siARN au 3' UTR des ARNm peut modifier l'expression de plusieurs gènes. Puis, il peut y avoir déclenchement d'une réponse immunitaire (ex. : les «*Toll like receptors*», TLR) par les siARN et/ou leur véhicule de livraison, ce qui aura un effet marqué sur les tissus touchés. Finalement, la présence d'une grande quantité de siARN dans la cellule pourrait saturer la machinerie d'ARNi endogène, entraînant ainsi une modification du patron de petits ARN endogènes et de leur effets (Birmingham et al., 2006; Fedorov et al., 2005; Jackson et al., 2006; Jackson et Linsley, 2010; Lim et al., 2005; Lin et al., 2005).

Comme l'a si bien discuté Patricia F. Dimond, dans un article paru le 11 juin 2012 dans la section *Insight & Intelligence*TM du site internet de la revue GEN (*Genetic Engineering & Biotechnology News*), l'ARN interférence accuse encore un retard sur les technologies antisens sur le périlleux chemin du développement de drogues (Dimond, 2012). Les premiers essais cliniques avec des siARN, en 2004, ont pavés la voie à plusieurs autres (pour une revue voir (Lares et al., 2010). Malheureusement, il n'y a pas à ce jour de drogue basée sur l'ARNi approuvée pour une quelconque thérapie.

L'approche des shARN

L'autre méthode de ciblage génique basée sur le mécanisme d'interférence par l'ARN rappelle la synthèse des miARN par les cellules. Pour cette approche, les

petits ARN interférents sont générés par la machinerie cellulaire à partir de matrices d'ADN exogènes (Introduction, Figure 7).

Initialement, la séquence des siARN est clonée dans un vecteur d'expression qui doit être acheminé au noyau de la cellule cible. En amont de la séquence d'intérêt, le promoteur d'une ARN polymérase cellulaire ADN dépendante permet la transcription d'un précurseur de siARN replié en épingle à cheveux (tige-boucle) que l'on nomme shARN («*short hairpin RNA*» ou «*shRNA*»). Les shARN ainsi produits dans le noyau de la cellule peuvent, à l'image des pre-miARN, être exportés directement au cytoplasme de façon EXP5 dépendante. On pourrait ainsi les considérer comme des pre-shARN. Alternativement, il est possible d'exprimer les shARN sous forme d'un précurseur (pri-shARN) qui serait mûré par le microprocesseur (Drosha-DGCR8) avant d'être envoyé au cytoplasme. Au cytoplasme, les shARN seront reconnus par Dicer et coupés en siARN qui à leur tour seront transférés au complexe RISC. Le brin guide sera alors sélectionné et le complexe RISC chargé pourra inhiber l'expression du gène cible (Introduction, Figure 7) (Singh et al., 2011).

Avec cette approche, le premier élément à choisir est le vecteur d'expression. Voici donc un bref survol des possibilités (pour une revue plus complète des modes de livraisons et des stratégies pour l'expression des shARN voir (Aagaard et Rossi, 2007; Lares et al., 2010; Sliva et Schnierle, 2010)). Pour les travaux *in cellulo* les plus courants, la transfection d'un plasmide encodant le shARN et son promoteur suffira pour inhiber le gène d'intérêt. Pour des cas plus complexes ou simplement pour obtenir une plus grande proportion de cellules exprimant le shARN, il est opportun d'utiliser des vecteurs viraux. Il existe plusieurs types de vecteurs viraux qui ont été développés autant pour la recherche fondamentale que pour le développement de thérapies. Dans tous les cas, les vecteurs viraux sont modifiés pour ne pas être infectieux ou du moins pour ne pas déclencher d'infections nocives pour l'hôte. Le potentiel de chacun pour le développement de thérapie varie selon la pathologie à traiter et les cellules visées

(Lares et al., 2010). Un autre facteur important pour la production de shARN efficaces est le promoteur permettant l'expression du petit ARN. Les premiers vecteurs d'expression pour les shARN utilisaient des promoteurs de l'ARN polymérase III (ex. : promoteur U6 ou H1). Ces promoteurs favorisent une expression ubiquitaire et l'accumulation de grandes quantités de shARN (Makinen et al., 2006a). À première vue, cela représente une qualité recherchée pour l'inhibition d'un gène d'intérêt causant une maladie. Par contre, il a maintenant été démontré qu'il est possible de saturer la machinerie cellulaire d'ARNi, principalement à l'étape du transport via EXP5, par une surexpression de shARN (Grimm et al., 2006; Yi et al., 2005). Cette situation devient néfaste pour la cellule puisqu'il y a un dérèglement dans l'expression de miARN et par conséquent dans l'expression d'une multitude de gènes. Une alternative est l'utilisation de promoteurs de l'ARN polymérase II. La sélection d'un promoteur actif uniquement dans les cellules cibles permet d'ajouter un niveau de spécificité à la technique. Il est aussi possible d'utiliser des promoteurs inductibles pour mieux contrôler le niveau d'expression des shARN. D'ailleurs, il existe maintenant des systèmes inductibles autant pour les promoteurs des ARN polymérases II et III. Bien que mieux adaptés à la recherche fondamentale, il serait envisageable de les utiliser en thérapie dans certaines conditions. Peu importe quelle stratégie est utilisée pour livrer les shARN aux cellules cibles, il est impossible de tirer parti de modifications chimiques pour corriger ou atténuer les effets secondaires dus à un manque de spécificité. Par contre, cette stratégie permet une inhibition soutenue du gène cible, surtout lorsqu'il y a intégration dans le génome des cellules ciblées. Au final, le choix entre siARN ou shARN et leurs multiples déclinaisons doit se faire selon le contexte précis de chaque expérience ou de chaque pathologie. Il existe donc une multitude d'options pour une panoplie de problèmes.

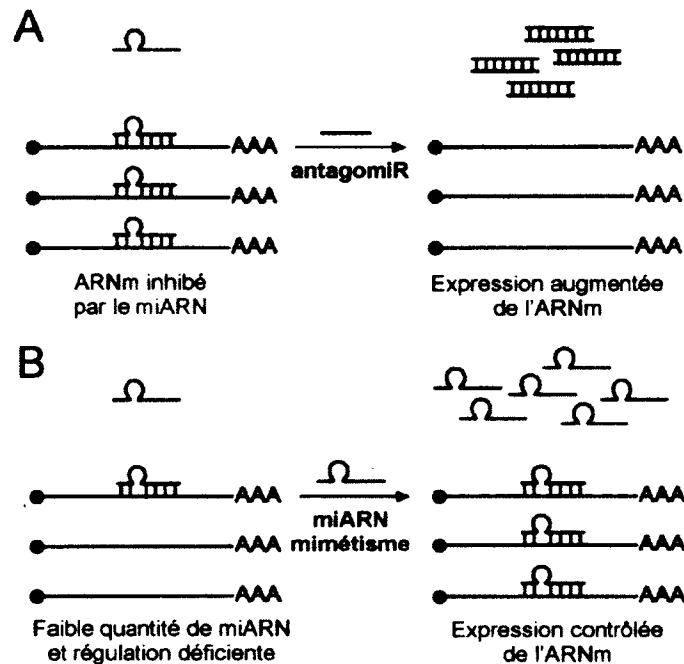
Les anti-miARN et les miARN-mimic

Pour terminer cette section sur l'interférence par l'ARN, voici une autre façon originale d'exploiter la puissance de cette machinerie cellulaire pour modifier l'expression génique de cellules cibles. Ainsi, plutôt que de cibler un ou des messagers pour inhiber leur expression, c'est l'activité des miARN endogènes qui sera influencée à la baisse ou à la hausse selon le besoin et la méthode.

Les miARN représentent des cibles de choix pour modifier le phénotype des cellules en raison de leur implication dans la régulation d'une multitude de gènes et donc d'une panoplie de processus biologiques. Par exemple, certains miARN auraient un caractère oncogénique alors que d'autres agiraient comme suppresseurs de tumeurs (Chan et al., 2005; He et al., 2005; Johnson et al., 2005). À l'aide d'oligonucléotides antisens, tels que présentés précédemment, il est possible de bloquer l'action d'un miARN (Introduction, Figure 8) (Kurzynska-Kokorniak et al., 2009). Ces antisens peuvent être désignés comme anti-miARN ou encore comme des antagonirs (*antagonist-miRNAs*) (Kurzfeldt et al., 2005). Le potentiel de cette stratégie a entre autres été démontré dans un modèle de souris où l'inhibition de miR-122 (miARN-122) a permis de mettre à jours son implication dans le métabolisme des lipides ((Esau et al., 2006). Cette méthode a permis d'élucider la fonction *in vivo* de plusieurs miARN (Collison et al., 2011; Fontana et al., 2008; Kurzfeldt et al., 2005). L'implication de miARN dans le cycle de vie de certains virus pourrait également être une cible thérapeutique à développer (Kurzynska-Kokorniak et al., 2009; Pan et al., 2009; Sall et al., 2008).

À l'inverse d'inhiber la fonction d'un miARN, il est aussi possible de rétablir ou d'augmenter l'activité d'un miARN (Introduction, Figure 8) qui serait en trop faible quantité dans la cellule pour effectuer son travail correctement ou encore pour inhiber la réplication d'un virus (Arbuthnot, 2011; Murakami et al., 2009; Tsai et Yu, 2010). Des stratégies similaires à celles développées pour les siARN (Tsuda

et al., 2006) ou pour l'expression de shARN (Zeng et al., 2002) sont appropriées pour cette tâche.



Introduction, Figure 8. Schéma de la modulation de la voie des miARN.

Les ARN messagers (ARNm) sont illustrés en noir avec un point (coiffe 5') et trois A (queue 3' poly(A)). Les différents microARN (miARN) sont représentés par une courte ligne présentant un renflement. A) L'inhibition de l'expression d'un ARNm par un miARN (rouge) peut être bloquée par une molécule d'acide nucléique complémentaire tel qu'un antagomiR (bleu). La liaison de l'antagomiR au miARN permet une augmentation de l'expression de l'ARNm. B) L'inhibition insuffisante de l'expression d'un ARNm par un miARN (noir) peut être complétée par l'ajout de miARN mimétiques (vert). Le miARN mimétique est dessiné selon la séquence d'un miARN matrice et il utilise la machinerie cellulaire pour contrôler l'expression de son ARNm cible. Adaptée de (Kurzynska-Kokorniak et al., 2009).

2.2. Les outils catalytiques

Les travaux présentés jusqu'ici dans cette thèse illustrent l'importance de l'ARN bien au-delà de sa simple fonction de messenger. En effet, depuis que les professeurs Sidney Altman et Thomas R. Cech ont démontré que l'ARN pouvait jouer le rôle de catalyseur pour des réactions biochimiques, l'ARN ne peut plus

être cantonné dans un rôle de support de l'information génétique entre l'ADN, source d'hérédité, et les protéines, biocatalyseurs de la vie (Nobel Media AB. , 1989). En 1982, le groupe de recherche du professeur Thomas R. Cech a pu démontrer qu'une molécule d'ARN pouvait se couper et raccorder les bouts sans la présence de protéine (Kruger et al., 1982). À ce moment, ils décrivaient l'épissage d'un ARN chez *Tetrahymena thermophila*, maintenant identifié comme l'intron du groupe I. L'année suivante, les travaux de l'équipe du professeur Sidney Altman, sur la maturation des ARN de transfert chez la bactérie *Escherichia coli* (*E. coli*), démontrèrent qu'une molécule d'ARN pouvait posséder toutes les caractéristiques d'une enzyme (Guerrier-Takada et al., 1983). En fait, cinq ans après avoir démontré la nécessité de la présence d'une sous-unité ARN pour la fonction de l'enzyme RNase P (Stark et al., 1978), ils ont pu démontrer que la portion ARN de la RNase P catalysait la coupure de la séquence *leader* en 5' des ARN de transfert (ARNt) (Guerrier-Takada et al., 1983). Ces découvertes majeures leur permirent de recevoir conjointement le prix Nobel de chimie en 1989, pour la découverte des ARN catalytiques ou ribozymes (Rz) (Nobel Media AB. , 1989).

Ces enzymes faites de ribonucléotides ont maintenant été identifiés un peu partout dans la nature sous diverses formes (Salehi-Ashtiani et al., 2006; Webb et al., 2009). À cela s'ajoute une pléthore de ribozymes issus de sélections *in vitro* et catalysant une myriade de réactions biochimiques différentes (Joyce, 2007). Des expériences de sélection *in vitro* ont aussi été faites avec de l'ADN et ont permis d'obtenir des molécules d'ADN catalytique («*DNAzymes*») pouvant être utilisées pour couper un ARN d'intérêt et inhiber l'expression d'un gène (pour une revue voir (Chan et Khachigian, 2009; Silverman, 2005)). Dans tous les cas, ils possèdent des structures secondaires et tertiaires bien définies qui leur permettent de catalyser une foule de réactions biochimiques. Ils représentent d'ailleurs un excellent modèle pour l'étude des relations structures fonctions de l'ARN (Reymond et al., 2009a). Parmi toutes les activités possibles des ribozymes, l'hydrolyse d'un

lien phosphodiester est la plus pertinente pour les travaux présentés dans cette thèse et sera donc le focus de la section des ARN catalytiques.

L'idée du ciblage spécifique d'un ARNm est apparue dès la découverte de l'activité catalytique des ARN et plus particulièrement avec celle du ribozyme de la RNase P. Aujourd'hui, il existe une multitude de technologies impliquant divers motifs catalytiques pour induire la dégradation d'ARN messagers. Avant d'entrer dans les détails des ribozymes les plus pertinents pour le ciblage spécifique de gènes, voici quelques précisions sur ces ribozymes. D'abord, la majorité, à l'exception du ribozyme de la RNase P, sont naturellement des ribozymes agissant en *cis*, c'est-à-dire que l'ARN à être coupé (le substrat) et la structure catalytique se retrouve dans une seule et même molécule. Évidemment, ces ribozymes ont dû être modifiés pour agir en *trans*, (substrat et ARN catalytique séparés) pour être utilisés comme ciseaux moléculaires. Finalement, la littérature réfère généralement à deux catégories de ribozymes. Les grands ribozymes («*large ribozymes*») font plus de 300 nucléotides de long et incluent les ribozymes des introns des groupes I et II ainsi que celui de la RNase P. Les ribozymes inférieurs à 300 nucléotides sont regroupés parmi les petits ribozymes («*small ribozymes*»). Les motifs catalytiques de ce dernier groupe partagent d'autres éléments en dehors de leur taille restreinte. Ils sont tous impliqués dans le processus de réplication de génome d'ARN dans lequel ils sont retrouvés. Ils catalysent également la même réaction de trans-estérification qui génère des extrémités 5'-hydroxyle (5'-OH) et 2',3'-monophosphate cyclique. Par contre, chaque motif possède une structure secondaire caractéristique et par conséquent ses particularités au niveau de la mécanistique de la réaction catalysée (Puerta-Fernandez et al., 2003). Les ribozymes particulièrement intéressants pour le développement d'outils moléculaires comprennent le ribozyme de la RNase P, le ribozyme en épingle à cheveux (Rz «*hairpin*»), le ribozyme en tête de marteau (Rz «*hammerhead*») et le ribozyme du Virus de l'Hépatite *Delta* (Rz HDV). Il est important de noter qu'une foule de travaux ont permis d'élucider la structure et le mécanisme d'action de ces

ribozymes, mais que ces notions ne seront que très peu abordées cette thèse. Dans les pages qui suivent, chacun de ces ribozymes sera brièvement présenté individuellement, l'emphase portera sur le ciblage de gènes.

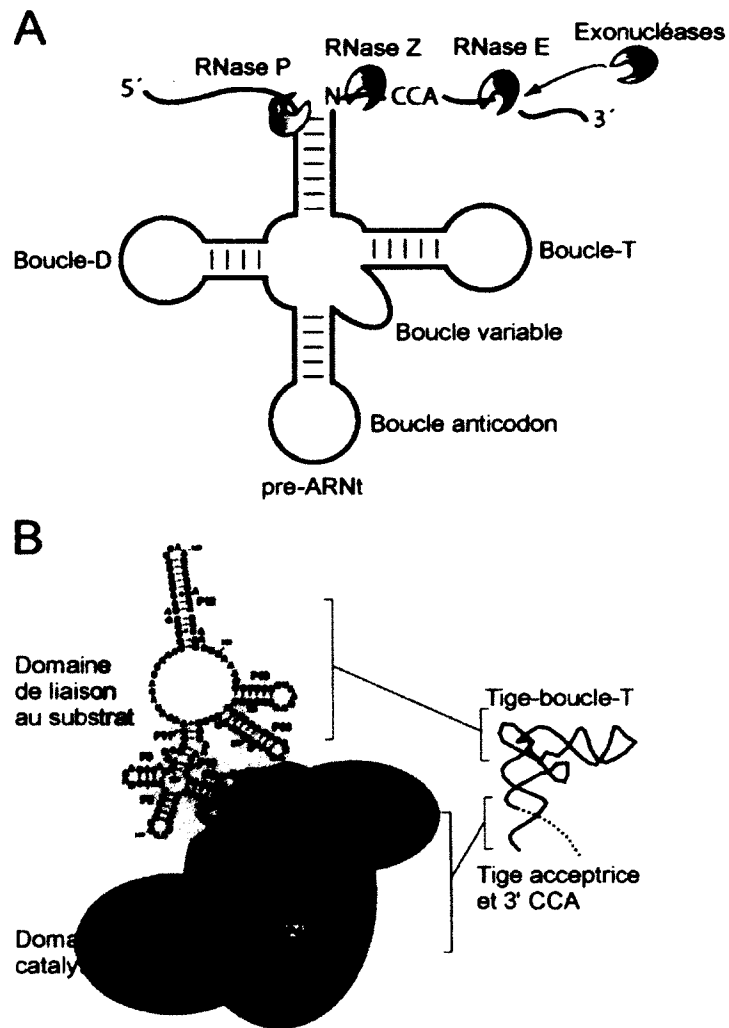
2.2.1. Ribozyme de la ribonucléase P

Parmi les activités enzymatiques de la biologie des ARN, l'activité de la ribonucléase P (RNase P) est l'une des plus conservées des cellules vivantes. Depuis les premières observations de cette activité jusqu'à tout récemment, la vie sans activité RNase P était considérée comme impossible. Le premier exemple d'une cellule n'utilisant pas cette activité pour la biogénèse de ses ARNt a d'ailleurs été démontré pour la première fois plus de 35 ans après sa découverte originale (Randau et al., 2008). La RNase P représente une modulatrice clé de la maturation en plusieurs étapes des transcrits primaires d'ARNt. Le rôle de cette activité est de retirer la séquence *leader* en 5' des précurseurs d'ARNt (pré-ARNt) alors que d'autres endonucléases et exonucléases sont responsables de rogner l'extrémité 3' pour former des ARNt matures (Introduction, Figure 9A) (Lundblad et Altman, 2010)). D'autres substrats plus marginaux ont également été rapportés dans la littérature. Parmi ceux-ci, on retrouve des ARN régulateurs de phages, des ARN d'opérons bactérien, des ARN du signal de reconnaissance de particules («*signal recognition particle RNAs*») ou encore des riborégulateurs (pour une revue voir (Marvin et Engelke, 2009)).

La particularité de l'holoenzyme RNase P est la contribution majeure de la sous-unité ARN pour la catalyse. Comme je l'ai mentionné précédemment, les travaux du Dr. Altman ont permis de démontrer que cet enzyme était en fait un ribozyme (Guerrier-Takada et al., 1983). Des bactéries aux eucaryotes supérieurs, l'activité catalytique est réalisée par la partie ARN de l'enzyme. Chez les bactéries l'enzyme RNase P est formée simplement de deux sous-unités, l'une ARN et l'autre protéique. Plus précisément, la RNase P de *E. coli* consiste en un ARN de

377 nucléotides de long nommé M1 (Introduction, Figure 9B) et en une sous-unité protéique de 14 kDa, la protéine C5 (Christian et al., 2002). Bien que les deux sous-unités soient nécessaires pour l'activité *in vivo*, une concentration ionique plus élevée permet l'activité catalytique *in vitro* de la sous-unité ARN seule (Kim et Liu, 2007). L'ARN M1 est donc un véritable enzyme fait d'ARN. Ce ribozyme reconnaît son substrat, le coupe et se régénère pour une ronde subséquente de coupure. Dans ce cas précis, le rôle de la protéine C5 est de stabiliser des structures tertiaires du ribozyme ainsi que d'augmenter la spécificité de l'holoenzyme pour les pre-ARNt (Kim et Liu, 2007). À l'autre extrémité de l'éventail des RNase P, l'holoenzyme des eucaryotes est beaucoup plus complexe. Ainsi, l'ARN H1 représente la sous-unité ARN de l'enzyme humaine. Ce ribozyme s'associe avec une dizaine de sous-unités protéiques pour réaliser sa tâche. La stabilisation de l'ARN, la localisation cellulaire ou la contribution au site actif de l'enzyme représentent des rôles potentiels de ces protéines (Kim et Liu, 2007).

Depuis sa découverte, l'enzyme RNase P a été vu comme un potentiel outil moléculaire pour l'inhibition dirigée de l'expression de gènes. Pour moi, ce ribozyme se démarque quelque peu des autres ribozymes qui seront discutés dans cette thèse. D'une part, la reconnaissance du substrat se fait essentiellement au niveau de la structure plutôt que par la séquence primaire de l'ARN cible. En fait, c'est la structure du pré-ARNt, particulièrement la tige acceptrice, qui est reconnue par le ribozyme (Introduction, Figure 9B). D'autre part, dans son contexte naturel, ce ribozyme doit effectuer plusieurs rondes de catalyse. Il est donc parfaitement adapté pour la réalisation d'un haut taux de *turnover*. À contrario, les autres ribozymes qui seront discutés plus loin ont été naturellement sélectionnés pour la réalisation d'un évènement de coupure unique lors de la réplication d'un génome ARN infectieux.



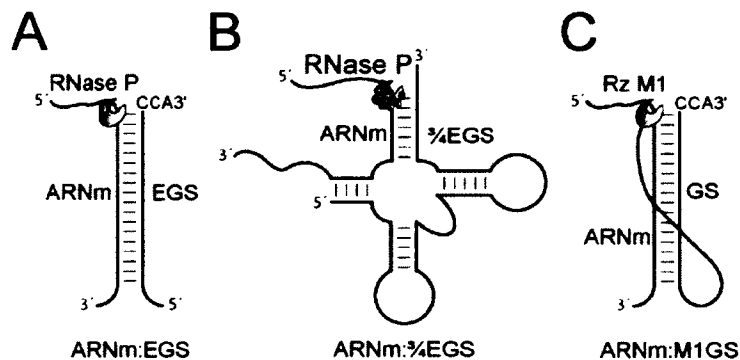
Introduction, Figure 9. Schéma de la maturation des ARN de transfert et représentation de la structure secondaire du ribozyme RNase P de *E. coli*.

A) Les différentes activités catalytiques impliquées dans la maturation des ARN de transfert (ARNt) sont représentées sous une forme «pacman». L'ARN de la RNase P (pacman bleu) et la ou les protéines de la RNase P (symbole brun) sont responsables de la maturation de l'extrémité 5' des pre-ARNt. Adaptée de (Lundblad et Altman, 2010). **B)** La structure secondaire du ribozyme de la RNase P de *E. coli* ainsi que les deux principaux domaines sont représentés. Le domaine de liaison au substrat (gris pâle) et le domaine catalytique (gris foncé) sont formés de plusieurs hélices désignées par la lettre P suivie d'un numéro. Le croquis de droite situe la liaison du pre-ARNt au ribozyme. Adaptée de (Christian et al., 2002).

À partir de cette activité catalytique conservée, deux stratégies ont pu être mises de l'avant pour le ciblage de gènes. L'une est basée sur le recrutement de l'activité RNase P endogène des cellules cibles alors que l'autre est basée sur l'utilisation du ribozyme M1 de *E. coli*. Dans les deux cas, l'idée est de recréer une structure semblable à la tige acceptrice d'un ARNt et de permettre la liaison avec le ribozyme de la RNase P. En fait, on pourrait même dire que cette stratégie de mimétisme de la structure était déjà utilisée dans la nature par les autres substrats non-conventionnels qui se replient de façon à être reconnus et coupés par le ribozyme de la RNase P (Marvin et Engelke, 2009). Cette activité catalytique peut donc être dirigée contre différentes cibles à condition de respecter certains critères structuraux. La spécificité du ciblage génique de ces deux stratégies est dirigée vers une séquence précise par l'entremise d'un ARN guide. L'appariement de l'ARN guide avec l'ARN cible crée une structure secondaire qui correspond à un substrat du ribozyme (Introduction, Figure 10). Le complexe ainsi formé sera reconnu par le ribozyme et la catalyse aura lieu. Les deux stratégies, utilisent un ARN guide, communément appelé EGS pour «*external guide sequence*», mais leur design est différent selon la méthode. Toutefois, l'activité de chaque approche dépend de l'accessibilité du site de liaison visé, comme tous les mécanismes impliquant la reconnaissance d'un ARN via la formation de paires de bases. Ainsi, la structure secondaire et tertiaire de l'ARN cible ainsi que la présence de protéines liées à l'ARN affectent l'efficacité du ciblage. Les prédictions *in silico* de structures secondaires, la cartographie *in vitro* et *in cellulo* ainsi que différentes méthodes de sélection *in vitro* sont utilisées pour le choix des meilleurs candidats (pour une revue voir (Kim et Liu, 2007; Lundblad et Altman, 2010)).

Tout d'abord, l'activité RNase P cellulaire peut être utilisée afin d'inhiber un gène d'intérêt. Dans ce cas, l'effecteur se limite à un ARN guide (EGS). Ce dernier doit être complémentaire au gène visé et permettre, une fois en complexe avec celui-ci, le recrutement de l'holoenzyme RNase P endogène. La coupure de l'ARN cible est catalysée par la machinerie cellulaire sans qu'un ribozyme exogène

n'intervienne. Ce système de ciblage génique fonctionne en bactéries ou en cellules de mammifères. L'enzyme bactérienne peut reconnaître un complexe ARNm-EGS double brin avec un EGS minimaliste portant une extrémité 3'-CCA simple brin (Introduction, Figure 10A). Par contre, les modèles humains nécessitent que la structure de l'EGS soit plus complexe et que la localisation intracellulaire de la cible soit compatible avec celle de l'enzyme. Dans ce cas, le complexe ARNm-EGS doit en plus reproduire les tiges T et D de l'ARNt alors que la tige de l'anticodon n'est pas nécessaire (Introduction, Figure 10B). En général, l'utilisation d'EGS bénéficie de l'efficacité de l'enzyme cellulaire et la taille de l'effecteur à introduire dans les cellules demeure relativement petite.



Introduction, Figure 10. Schéma des stratégies pour le ciblage de gènes via l'activité RNase P.

A) La coupure d'un ARNm par la RNase P bactérienne avec un ARN guide (EGS). **B)** La coupure d'un ARNm par la RNase eucaryote nécessite une EGS plus complexe. Dans ce cas, l'ARN guide (3/4EGS) forme une structure semblable à celle d'un ARNt. **C)** Le ribozyme bactérien (M1) fusionné à son ARN guide (GS) forme un ribozyme (M1GS) qui peut reconnaître son ARN cible et le couper sans la participation de l'activité RNase P endogène. Adapté de (Lundblad et Altman, 2010).

La seconde stratégie développée utilise les propriétés catalytiques du ribozyme bactérien M1 couplé à un ARN guide. Ainsi, l'EGS est fusionné au ribozyme M1 de *E. coli* pour former un ribozyme M1GS (Liu et Altman, 1995) (Introduction, Figure 10C). C'est donc le ribozyme de fusion M1GS qui doit être

exprimé dans les cellules cibles pour induire la coupure spécifique des ARN complémentaires à la séquence guide. Cette méthode plus facile d'approche permet de cibler un gène sans avoir besoin de recruter l'activité endogène de la cellule. La portion guide permet la liaison spécifique du messenger cible via la formation de paires de bases. Étant donné la liaison covalente de l'ARN guide et du ribozyme, cet appariement permet le rapprochement du ribozyme M1 à son ARN cible (Kim et Liu, 2007).

Le ciblage d'un ARN par l'une ou l'autre des technologies basées sur la RNase P nécessite ou du moins profite de certains déterminants (Lundblad et Altman, 2010). Par exemple, les séquences adjacentes au site de coupure dans l'ARN cible doivent répondre à certaines contraintes. Pour le ribozyme M1GS, le site idéal de coupure correspond à une pyrimidine suivi d'une guanosine ($Y_{-1} * G_{+1}$, avec * représentant le site de coupure). Une hélice intermoléculaire entre l'ARN guide et le substrat favorise une activité de coupure maximale. Dans le cas de la RNase P humaine, s'ajoute le besoin d'avoir un uracile en position +8 du site de coupure. Un excès d'ARN EGS par rapport à l'ARN cible est également un facteur de réussite pour l'inhibition de gènes. D'autres éléments sont cruciaux pour l'activité inhibitrice. Pour la fusion M1GS, la longueur de la séquence séparatrice entre les deux motifs doit être comprise entre 25 et 65 nucléotides de long. Dans ce sens, une optimisation par sélection *in vitro* peut permettre d'augmenter autant l'efficacité des ribozymes M1GS que celle des ARN EGS (Kilani et al., 2000; Zhou et al., 2002).

Altman et ses collègues ont démontré le potentiel des EGS comme agents antibactériens (McKinney et al., 2001). Pour leurs travaux, ils ont utilisé des EGS pour cibler des gènes bactériens essentiels. Cette étude a démontré la très grande spécificité des EGS. En fait, ils ont été en mesure de cibler spécifiquement des gènes de *E. coli* ou de *Salmonella typhimurium* encodant pour des protéines identiques mais ayant différents ARNm (McKinney et al., 2001). Ils ont ainsi démontré la capacité des EGS à spécifiquement inhiber la croissance d'une

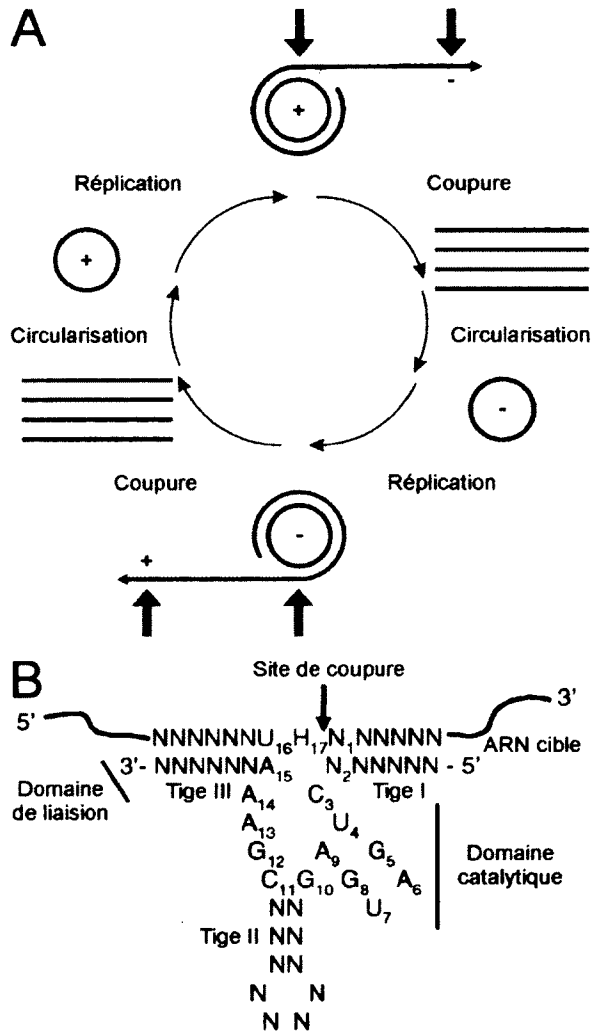
espèce bactérienne, ce qui permettrait le traitement de certaines infections sans affecter la flore bactérienne normale des patients. De nombreux essais ont aussi été réalisés pour tirer parti du potentiel énorme de ces ribozymes afin d'inhiber la réplication de différents virus d'importance pour la santé humaine. La réplication de virus tels que le virus de l'immunodéficience humaine (VIH), le virus de l'influenza A humain ou encore une série d'herpèsvirus («*human cytomegalovirus*», HCMV; «*herpes simplex virus 1*», HSV-1 et «*Kaposi's sarcoma-associated herpesvirus*», KSHV) a été inhibée avec succès par l'expression de ribozymes M1GS ou d'ARN EGS. D'ailleurs, des expériences de sélection *in vitro* ont permis de produire des variants du ribozyme M1GS dirigés contre le HSV-1 avec une activité de coupure 20 fois supérieure (Kilani et al., 2000). La même approche appliquée au développement d'EGS plus performants a permis d'accroître leur activité antivirale contre le HSV-1 et le HCMV (Zhou et al., 2002). Les deux stratégies utilisant la RNase P ont également été mises à profit pour combattre différents types de cancers (Cobaleda et Sanchez-Garcia, 2000). Finalement, le ribozyme de la RNase P et ses technologies dérivées représentent des outils intéressants pour étudier la fonction des gènes et pour décortiquer des voies cellulaires (Lundblad et Altman, 2010).

En terminant, il est important de noter que la différence de taille entre les deux effecteurs (EGS 25 à 60 nt et M1GS ~400 nt) pourrait influencer leurs utilisations futures. En effet, il est envisageable de produire des ARN EGS contenant des modifications chimiques (Ma et al., 2000) pour augmenter leur stabilité et favoriser leur livraison alors que les ribozymes M1GS devront toujours être pourvus aux cellules cibles via un vecteur d'expression. Il semble clair que les technologies issues de la RNase P montrent un potentiel pour le ciblage de gènes autant pour des applications thérapeutiques que pour la recherche fondamentale. Il reste encore beaucoup de questions à résoudre et de perfectionnement à faire. Même en l'absence d'essai clinique pour ces approches, il y a maintenant un

premier exemple de leur potentiel dans un modèle animal qui laisse entrevoir un bel avenir pour ces technologies (Bai et al., 2008).

2.2.2. Ribozyme «*hammerhead*»

Le ribozyme en tête de marteau (Rz «*hammerhead*») a d'abord été découvert comme un motif catalytique présent dans différents pathogènes de plantes à génome ARN. Le tout premier exemple de ce motif a été identifié dans le génome du virusoïde du «*Lucerne Transient Streak Virus*» (LTSV) en 1987 (Forster et Symons, 1987). Dans le génome monocaténaire d'ARN circulaire du virusoïde, le ribozyme «*hammerhead*» agit en *cis* (c.à.d. ribozyme et substrat inclus dans le même ARN) et permet ainsi la scission des copies de génomes multimériques en monomères lors de la réplication en cercle roulant (Puerta-Fernandez et al., 2003) (Introduction, Figure 11A). Le ribozyme «*hammerhead*» a été le premier petit ribozyme à être converti en ribozyme agissant en *trans* (c.à.d. ribozyme et substrat dans deux ARN séparés) (Haseloff et Gerlach, 1988; Uhlenbeck, 1987). Tous les ribozymes «*hammerhead*» possèdent une structure secondaire caractéristique (Introduction, Figure 11B) (Tedeschi et al., 2009)). La structure du ribozyme «*hammerhead*» en complexe avec le substrat se résume à une hélice intramoléculaire au niveau du ribozyme (hélice II) et deux hélices intermoléculaires (hélices I et III) formées par l'interaction avec le substrat (Puerta-Fernandez et al., 2003). Les régions simples brins sont très bien conservées et contiennent la majorité des nucléotides importants pour une activité de coupure optimale. La spécificité du ribozyme peut être changée en modifiant les séquences des bras de liaisons («*binding arms*») formant les hélices I et III (Introduction, Figure 11B). La séquence consensus du site de coupure la plus utilisée afin de générer des ribozymes «*hammerhead*» pour le ciblage de gènes est NUH (N= A, C, G ou U; H= A, C ou U), avec la coupure en aval du triplet. La séquence consensus moins stricte, NHH, peut également être utilisée (Drude et al., 2007; Puerta-Fernandez et al., 2003).



A) La réplication en cercle roulant symétrique est une stratégie utilisée par des agents infectieux à génome ARN circulaire. Les ARN circulaires de polarité positive (bleu) servent de matrices pour la synthèse d'ARN de polarité négative (rouge) sous la forme de concatémères. Ces intermédiaires de réplication sont ensuite coupés pour libérer des monomères. Dans plusieurs cas cette coupure dépend d'un motif auto-catalytique (ribozyme en *cis*). Les monomères sont circularisés pour ensuite compléter le cycle de réplication. **B)** La structure secondaire du ribozyme «hammerhead» en *trans* contient 3 tiges principales. Le site de coupure consensus (rouge) se situe entre 2 tiges (dans ce cas-ci, tige I et III) qui se forment lors de l'appariement du ribozyme à son ARN cible. Les nucléotides variables dans le ribozyme sont marqués d'un N. Les nucléotides devant être conservés pour maximiser l'activité de coupure sont précisés et se retrouvent dans le domaine catalytique ou à proximité du site de coupure. (N signifie A, C, G ou U et H signifie A, C ou U). Adaptée de (Tedeschi et al., 2009).

Introduction, Figure 11. Schéma de la réplication en cercle roulant symétrique et représentation de la structure secondaire du ribozyme «hammerhead».

Le ribozyme «*hammerhead*» est parmi les ribozymes les plus utilisés comme ciseaux moléculaires. Son activité de coupure élevée, sa capacité d'effectuer plusieurs *turnovers* et sa petite taille (aussi court que 35 nt) ont suscité l'intérêt de plusieurs groupes de recherche, incluant certaines compagnies pharmaceutiques (Drude et al., 2007; Puerta-Fernandez et al., 2003)). En effet, il est aussi bien possible de produire des ribozymes «*hammerhead*» par synthèse chimique que via un vecteur d'expression. Généralement, le design des ribozymes «*hammerhead*» pour le ciblage d'ARNm débute par l'identification des sites potentiels de coupure (ex. : NUH) à l'intérieur de la séquence visée. Par la suite, les ribozymes sont choisis pour cibler les meilleurs sites selon différents critères comme l'accessibilité (mesurée biochimiquement ou simplement prédite), la thermodynamique ou la sélection *in vitro*. Un peu à l'image de l'interférence à l'ARN, il y a plusieurs méthodes éprouvées permettant de présélectionner les meilleurs candidats avec un certain succès selon les critères mentionnés ci-haut (Denman, 1993; Lieber et Strauss, 1995; Lieber et Strauss, 1995; Mercatanti et al., 2002; Mercatanti et al., 2012; Pan et al., 2001).

Les applications des ribozymes «*hammerhead*» sont multiples, mais le ciblage de gène est la plus pertinente pour cette thèse. En effet, des ribozymes ont été développés pour couper en *trans* une pléthore d'ARN cibles (pour une revue voir (Tedeschi et al., 2009)). Des cancers aux infections virales, en passant par des maladies génétiques ou autres désordres métaboliques, de nombreux gènes candidats ont été ciblés avec cette technologie (Citti et Rainaldi, 2005). Plusieurs des meilleurs candidats ont été testés lors d'essais cliniques. Parmi les plus récentes études, on retrouve l'investigation du ribozyme OZ1 («*Johnson & Johnson Research Proprietary Limited*», Sydney, NSW, Australie) comme inhibiteur du VIH ainsi que l'ANGIOZYME («*Sirna Therapeutics*», San Francisco, CA, États-Unis) comme agent anti-cancer (Mitsuyasu et al., 2009; Morrow et al., 2012). Le premier cible directement l'ARN viral et est exprimé à partir d'un vecteur

viral («*Moloney murine leukemia virus-based, replication-incompetent gammaretroviral vector*», LNL6) servant à transduire *ex vivo* des cellules autologues hématopoïétiques progénitrices CD34+ afin de les réintroduire dans le sang du sujet. Le second est dirigé contre l'ARNm du VEGFR-1 («*vascular endothelial growth factor receptor-1*») et est livré directement sous sa forme active (ribozymes synthétiques contenant plusieurs modifications chimiques). Malheureusement, comme pour plusieurs autres ribozymes testés auparavant, les essais cliniques n'ont pu démontrer une activité thérapeutique suffisante et significative qui justifierait la poursuite des travaux. La tolérance aux traitements n'était pas limitante, ce qui est encourageant pour le développement de technologies semblables. En fait, la capacité d'acheminer aux cellules cibles une dose suffisante de molécules actives a semblé limitante, du moins dans le cas de l'ANGIOZYME, illustrant les grands défis toujours présents au niveau de la livraison des drogues à base d'acides nucléiques.

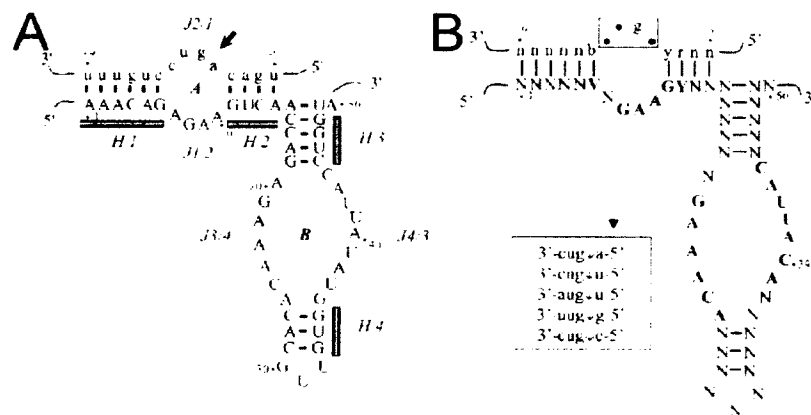
Des stratégies ingénieuses ont été mises en œuvres afin d'augmenter le potentiel des ribozymes «*hammerhead*». Parmi celles-ci, on retrouve les maxizymes («*minimized, activated, x-shaped, intelligent ribozyme*») (Shiota et al., 2004). Ces ribozymes sont formés de deux unités *minizymes* (Rz «*hammerhead*» sans la tige-boucle II) qui, seuls, sont généralement moins actifs que les ribozymes «*hammerhead*» mais qui en dimère démontre des activités très élevées. Un autre exemple d'une stratégie prometteuse est l'utilisation de ribozymes multi-unités (Chen et al., 1992). Ainsi, différents sites de coupures peuvent être ciblés par une seule molécule contenant plusieurs unités de ribozymes plutôt que par une série de ribozymes individuels. Finalement, certains groupes ont développé des molécules chimériques mariant le motif catalytique d'un côté à un autre motif d'ARN. Ces fusions visent à améliorer la distribution intracellulaire des ribozymes, d'augmenter leur potentiel inhibiteur en créant une synergie entre deux motifs inhibiteurs ou encore à créer des ribozymes allostériques dont l'activité catalytiques sera spécifique aux cellules visées (pour une revue voir : Puerta-

Fernandez et al., 2003; Scherer et al., 2007). Un autre élément adressé par la communauté scientifique est la faible stabilité de ce motif dans un environnement biologique. En plus de développer des vecteurs d'expression permettant la production constante d'une importante quantité de ribozymes dans les cellules cibles, la synthèse chimique automatisée a été développée (Citti et Rainaldi, 2005) afin de produire des ribozymes contenant des modifications chimiques permettant d'échapper à la dégradation par des nucléases (pour une revue voir (Citti et Rainaldi, 2005; Drude et al., 2007; Drude et al., 2007)(Citti et Rainaldi, 2005; Drude et al., 2007; Drude et al., 2007)r (Drude et al., 2007)).

2.2.3. Ribozyme «*hairpin*»

Il existe un second motif d'ARN capable d'auto-coupeure retrouvé parmi les pathogènes ARN de plantes. Cet ARN catalytique est le ribozyme en épingle à cheveux (Rz «*hairpin*»). La première observation de l'activité du motif «*hairpin*» date du milieu des années 80 (Gerlach et al., 1986). Toutefois, ce n'est qu'en 1989 qu'il a été proposé qu'un motif d'ARN différent du motif «*hammerhead*» était responsable de la coupeure du brin négatif de l'ARN satellite du virus des taches en anneaux du tabac («*negative strand of the satellite RNA of tobacco ringspot virus*», [-]sTRSV,) (Haseloff et Gerlach, 1989). Rapidement, le ribozyme «*hairpin*» du [-]sTRSV a été isolé et caractérisé (Feldstein et al., 1989; Hampel et Tritz, 1989) puis le nom «*hairpin catalytic RNA*» a été proposé (Hampel et al., 1990). Encore aujourd'hui, ce motif a seulement été retrouvé dans trois organismes, tous des ARN satellites de virus infectants des plantes (Shippy et al., 1999). Le rôle de ces ARN catalytiques est comparable à celui des ribozymes «*hammerhead*». Lors de la réplication en cercle roulant de ces ARN satellites, ils permettent la formation de monomères de génomes à partir des concatémères d'ARN transcrits (Introduction, Figure 11A). Encore une fois, il s'agit d'une activité de coupeure en *cis*.

Tout comme pour le ribozyme «hammerhead», le ribozyme «hairpin» a été converti en ribozyme agissant en *trans*. (Feldstein et al., 1989; Hampel et Tritz, 1989). La version la plus communément utilisée pour le ciblage de gènes est celle du [-]sTRSV (Figure 12A) qui fait environ 50 nucléotides de long, mais certaines versions atteignent jusqu'à 70 nucléotides (Drude et al., 2007). Le site de liaison du ribozyme sur l'ARN cible est de 14 nucléotides. La structure secondaire du ribozyme «hairpin» en complexe avec son substrat forme deux domaines bien définis se repliant indépendamment (Puerta-Fernandez et al., 2003). Chaque domaine est formé de deux régions hélicoïdales séparées par une boucle interne (Introduction, Figure 12A). D'une part, le domaine A est constitué de deux hélices intermoléculaires (hélices 1 et 2) formées par le ribozyme et le substrat séparées par la boucle interne A. D'autre part, le domaine B est l'agencement exclusif de nucléotides du ribozyme formant deux hélices intramoléculaire entrecoupées d'une longue boucle interne B. L'étude approfondie de ce ribozyme a révélé la singularité de son mécanisme réactionnel (pour une revue voir (Fedor, 2009)). L'ultime exemple de sa différence, bien que non pas le seul, réside dans sa capacité de catalyser aussi efficacement la réaction de coupure que la réaction de ligation (Fedor, 2000).



Introduction, Figure 12. Représentation de la structure secondaire du ribozyme «hairpin».

A) La séquence et la structure du ribozyme «hairpin» du [-]sTRSV sont illustrées. Les 2 domaines (A et B) du ribozyme ainsi que les 4 hélices (H1 à H4) et jonctions

(J1/2, J3/4, J4/3 et J2/1) sont identifiés. Les nucléotides du ribozyme sont numérotés (1 à 50) indépendamment de ceux du substrat (-5 à +9). **B)** La séquence consensus du site de coupure dans le substrat et les nucléotides essentiels (gras) à l'activité de coupure dans le ribozyme sont représentés (N signifie A, C, G ou U; H signifie A, C ou U; Y signifie C ou U, R signifie A ou G; B signifie C, G ou U; V signifie A, C ou G). Les séquences de la région J2/1 du substrat qui sont les plus favorables à la coupure sont représentées dans l'encadré. La mutation d'optimisation U39C est aussi montrée. Les flèches indiquent les sites de coupure. Adaptée de (Puerta-Fernandez et al., 2003).

Encore une fois, une foule d'applications utilisent le ribozyme «*hairpin*» bien que comparativement au ribozyme «*hammerhead*», l'utilisation du ribozyme «*hairpin*» pour le ciblage d'ARN d'intérêt est plus modeste. La capacité du ribozyme «*hairpin*» à reconnaître et induire la coupure d'un ARN cible à l'intérieur d'une cellule a été rapportée (Vitiello et al., 2000). Parmi ses applications, on retrouve l'inhibition de différents virus tel que le VIH, le virus de l'hépatite C (VHC) ou le virus de l'hépatite B (VHB) (Barroso-DelJesus et al., 2005; Jarczak et al., 2005; Klebba et al., 2000; Song et al., 2002). La capacité de ce ribozyme à inhiber l'expression de gènes d'intérêt a été évaluée pour plusieurs cibles de natures différentes (Cavusoglu et al., 2002; Leavitt et al., 2000). Plus spécifiquement, son potentiel comme molécule anticancéreuse a été mis à l'épreuve à quelques reprises (Chang et al., 2009; Halatsch et al., 2000; Li et al., 2007). En plus de ces applications à caractère thérapeutique, le ribozyme «*hairpin*» a également été utilisé pour réaliser des projets de génomique fonctionnelle. Ceux-ci ont entre autres permis l'identification de cofacteurs importants pour la traduction via le motif IRES du VHC (Kruger et al., 2000) ou encore de facteurs impliqués dans la transformation ou la prolifération cellulaire (Li et al., 2000; Welch et al., 2000).

Comme c'est le cas pour la majorité des outils d'acides nucléiques visant l'inhibition de l'expression d'un gène d'intérêt, des méthodes de design ont été mises au point. La sélection des sites les plus prometteurs et le design des ribozymes les plus efficaces pour les cibler sont les enjeux majeurs de ces méthodes. L'une des pistes de solutions est l'utilisation d'une librairie combinatoire

de ribozymes «*hairpin*» coupant en *cis*, permettant d'identifier les sites les plus accessibles dans un ARN cible et de sélectionner le ribozyme le mieux adapté pour ces sites (Barroso-DelJesus et Berzal-Herranz, 2001; Barroso-delJesus et al., 2004; Romero-Lopez et al., 2004). Sinon, de façon plus classique, la séquence minimale pour le site de coupure du ribozyme est définie comme BN*GUC (où B= C, G ou U et * indique le site de coupure) (Shippy et al., 1999). Une version plus stricte du site de coupure ainsi que les nucléotides importants pour la réaction de coupure ont été identifiés (Introduction, Figure 12B) (Puerta-Fernandez et al., 2003). Cette séquence est particulièrement critique pour l'activité de coupure en *trans* du ribozyme «*hairpin*» alors que pour l'activité de coupure en *cis* ou la ligation les contraintes sont légèrement différentes. Il est également possible d'optimiser le ribozyme «*hairpin*» pour favoriser une meilleure efficacité catalytique (kcat/Km) en modifiant la longueur de l'hélice 1 entre le ribozyme et son ARN cible (entre 5 et 10 bp) (Hampel et al., 1997). Dans certains cas, la transformation de la boucle 3 par une tétra-boucle (tetraloop) permet également d'améliorer les propriétés catalytiques des ribozymes (Lian et al., 1999). Globalement, les paramètres influençant l'activité des ribozymes «*hairpin*» comme outils moléculaires *in cellulo* sont similaires aux autres technologies impliquant la formation de paires de bases entre les molécules effectrices et leurs ARN cibles. Ainsi la proximité (sous localisation cellulaire) et la disponibilité du site de liaison (structure et liaison d'autres facteurs comme des protéines) s'ajoutent à la stabilité et à l'activité intrinsèque du ribozyme pour déterminer son efficacité *in cellulo*.

2.2.4. Ribozyme du virus de l'hépatite *delta*

Le ribozyme HDV est le nom donné aux motifs catalytiques retrouvés dans les ARN du virus de l'hépatite *delta* (VHD). Ce virus dépend du virus de l'hépatite B (VHB) pour sa propagation puisqu'il exploite les protéines de celui-ci pour former son enveloppe. Le VHD est donc un virus satellite infectant spécifiquement les

hépatocytes humains. Il possède un génome ARN monocaténaire circulaire d'approximativement 1700 nucléotides (Wang et al., 1986). La séquence du virus, riche en guanosine et cytosine, se replie en une structure non-branchée en forme de tige («*rod-like*») avec près de 75 % d'appariement intramoléculaire. Le VHD est le seul virus infectant les animaux avec un génome d'ARN circulaire et il s'apparente aux viroïdes affectant les plantes (Kos et al., 1986). À la différence des viroïdes qui ne sont pas traduits, le VHD encode pour une protéine, l'antigène HDag («*Hepatitis Delta antigen*»), responsable d'encapsider le génome viral (pour une revue voir (Hughes et al., 2011)). À l'image de ces pathogènes de plantes au génome circulaire, il se réplique par un mécanisme en cercle roulant symétrique. Cela signifie que la réplication de l'ARN génomique produit un ARN complémentaire, l'ARN anti-génomique. Le virus utilise un motif d'ARN catalytique dans chaque polarité pour assurer la formation d'ARN monomérique à partir de l'ARN multimérique transcrit. Ces deux ribozymes possèdent une structure et un mécanisme équivalents qui se distinguent des ribozymes «*hammerhead*» et «*hairpin*».

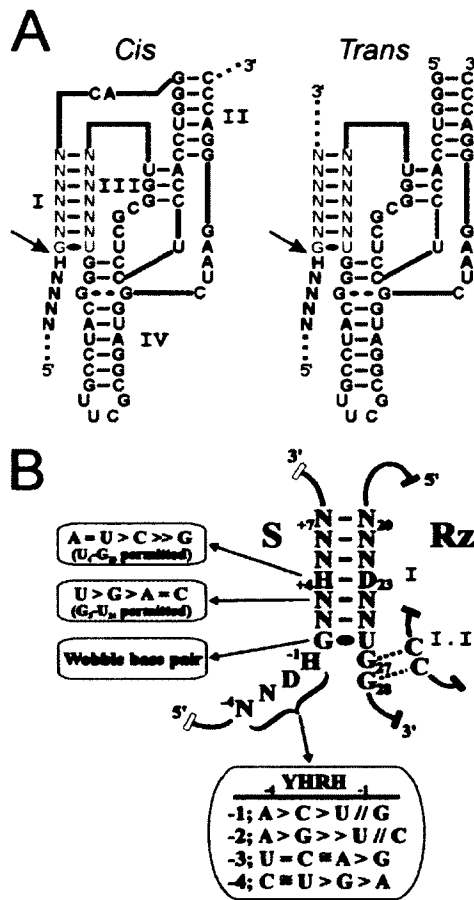
Les ribozymes HDV ont été identifiés pour la première fois à la fin des années 80 (Kuo et al., 1988; Sharmeen et al., 1988). Tout comme les ribozymes présents chez les pathogènes de plantes, les ribozymes HDV fonctionnent naturellement en *cis*. Bien que non identiques, les structures du ribozyme génomique et de l'anti-génomique sont très similaires (Fedor, 2009). Puisque mes travaux portent sur l'utilisation du ribozyme anti-génomique, je me concentrerai sur celui-ci. Ainsi, pour le reste de cette thèse l'expression «ribozyme HDV» réfèrera au motif anti-génomique sauf en cas d'une mention contraire.

La conversion des motifs catalytiques du VHD en ribozymes agissant en *trans* n'a pas été aussi «simple» que pour les ribozymes «*hammerhead*» et «*hairpin*». En fait, malgré une structure secondaire et tertiaire encore mal comprise, Branch et Robertson ont été les premiers à générer des ribozymes HDV pouvant reconnaître et couper en *trans* une seconde molécule d'ARN (Branch et

Robertson, 1991). Encore aujourd'hui, le ribozyme HDV demeure moins étudié que les ribozymes «*hammerhead*» et «*hairpin*», tel que l'indique le nombre de publications en lien avec chacun. Par contre, sa structure est maintenant comprise alors que le détail du mécanisme de la réaction chimique est le principal sujet de débat (Golden, 2011). Le sentier de repliement en plusieurs étapes du ribozyme HDV est d'ailleurs l'un des mieux décrits dans la littérature (Reymond et al., 2009b; Reymond et al., 2010). Brièvement, le ribozyme forme quatre tiges reliées entre elles par des nucléotides non-appariés (boucles ou jonctions) (Introduction, Figure 13A) (Reymond et al., 2010). On retrouve deux tiges : la tige I est le domaine de reconnaissance du substrat pour le ribozyme en *trans* alors que la tige II forme un pseudonoeud dans la version *cis* du ribozyme. Les deux autres séquences double brin forment des tiges-boucles (tige III et IV). Outre les boucles III et IV, on retrouve trois régions séquences simple brin correspondant aux jonctions entre les tiges (I/II, I/IV et IV/II). De plus, lors du repliement vers la forme active du ribozyme, deux nucléotides de la boucle III s'apparient à la jonction I/IV pour former un pseudonoeud essentiel à l'activité, la tige I.I. Finalement, il est important de noter que la cytosine 76 (C₇₆ ou C₇₅ pour la version génomique), désignée comme la cytosine catalytique, est impérative pour l'activité de coupure (Ferre-D'Amare et Doudna, 2000; Golden, 2011). La façon la plus simple de générer un ribozyme catalytiquement inactif est de muter cette cytosine en guanosine ou adénosine (Mercure et al., 1998). Ce type de mutant conserve sa capacité de liaison, mais reste incapable de couper son substrat.

Les différents travaux sur la relation structure-fonction du ribozyme HDV ont également permis de faire de l'ingénierie sur celui-ci pour moduler son activité de coupure. Pour contrôler l'activité du ribozyme, ces stratégies utilisent des combinaisons de modules d'ARN et de différents cofacteurs tels que des ions (Beaudoin et Perreault, 2008), des petites molécules chimiques (Kertsburg et Soukup, 2002), des oligonucléotides (Rouleau et al., 2012) ou encore le substrat lui-même (Bergeron et Perreault, 2005). Mes travaux de thèse portent d'ailleurs sur

ce dernier exemple où l'activité du ribozyme est dépendante de la liaison à son substrat spécifique (plus de détails à la section suivante)



du ribozyme HDV et schéma de la spécificité de son site de coupure.

A) La structure et la séquence du ribozyme HDV pour la version en *cis* (gauche) et en *trans* (droite). La séquence du site de liaison du ribozyme avec son substrat (bleu) peut varier (N signifie A, C, G ou U). Les hélices qui forment l'architecture du motif catalytique sont identifiées par les chiffres romains (I, II, III et IV). La tige I représente le domaine de reconnaissance du ribozyme. **B)** Représentation des séquences de nucléotides qui contribuent à la spécificité de substrat du ribozyme HDV. Pour chaque position où il y a une préférence pour le ribozyme, un encadré indique l'activité de coupure en ordre décroissant. (N signifie A, C, G ou U; H signifie A, C ou U; D signifie A, G, ou U; Y signifie C ou U, R signifie A ou G). La tige I.I, essentielle à l'activité de coupure, est aussi identifiée sur cet agrandissement du site de coupure. Adaptée de (Bergeron et al., 2003).

Introduction, Figure 13. Représentation de la structure secondaire

La grande stabilité *in cellulo* du ribozyme HDV, près de 100 heures sous certaines conditions (Levesque et al., 2002) et son adaptation à l'environnement des cellules de mammifères lui confèrent un avantage théorique qui ne s'est pas encore transposé en une utilisation accrue pour le ciblage génique. Il reste, encore aujourd'hui, moins employé à cette fin que les ribozymes «*hammerhead*» et «*hairpin*», mais il a tout de même le potentiel d'être un outil moléculaire

performant. Tout comme pour les autres ribozymes, sa capacité à couper un ARN cible *in vitro* et *in cellulo* a été démontrée au cours des 20 dernières années (Al-Anouti et Ananvoranich, 2002; Bergeron et Perreault, 2002; D'Anjou et al., 2004; Kato et al., 2001; Lu et al., 2011; Roy et al., 1999; Sheng et al., 2004; Wang et al., 2010). Le ribozyme HDV reconnaît son substrat via la formation de la tige I. Cette séquence de reconnaissance contient seulement sept nucléotides. Le résultat est donc une spécificité limitée, qui ne semble pas suffisante pour discriminer un ARNm parmi la myriade de séquences contenues dans une cellule. Il a été proposé qu'une séquence doit avoir 13-14 nucléotides de long pour être statistiquement unique (Herschlag, 1991; Peracchi, 2004). L'analyse de l'impact des nucléotides adjacents au site de coupure a permis de raffiner les éléments qui définissent un site de coupure optimal. D'une part, une étude de mutagenèse des positions -1 à -4 en amont du site de coupure a permis d'identifier les séquences favorisant l'activité de coupure maximale (Deschenes et al., 2000). D'autre part, les nucléotides au centre du domaine de reconnaissance du ribozyme semblent également influencer l'activité de coupure des ribozymes HDV en *trans* (Ananvoranich et al., 1999). L'imposition de contraintes pour le site de coupure par ces deux éléments, signifie que la spécificité du ribozyme va plus loin que la simple formation de sept paires de bases (Introduction, Figure 13B) (Bergeron et al., 2003). De ce fait, il est aussi possible de sélectionner des sites de coupures plus propices pour la coupure par le ribozyme HDV. Au-delà de mieux définir les déterminants d'un bon site de coupure, des travaux ont aussi testé différentes méthodes pour la sélection de ribozyme HDV pouvant couper efficacement un ARN d'intérêt (Bergeron et Perreault, 2002). Dans cette étude, trois approches ont été développées pour l'identification de ribozymes HDV pouvant cibler des ARN du VHB. Premièrement, ils ont utilisé une approche combinée de prédiction bioinformatique (RNA structure/OligoWalk) couplée à un essai biochimique (oligonucléotide plus RNase H). En deuxième lieu, ils ont utilisé une librairie combinatoire d'oligonucléotides pour effectuer des réactions de coupure de l'ARN cible par la RNase H. Finalement, ils ont généré une librairie combinatoire de

ribozymes HDV qu'ils ont ensuite utilisée pour couper l'ARN cible du VHB. Ultiment, pour chacune de ces approches, il faut procéder à des essais de coupure *in vitro* avec les ribozymes HDV dirigés contre les sites identifiés. Globalement, cela permet d'évaluer l'accessibilité des sites cibles ainsi que la capacité à former des complexes ribozyme-substrat actifs dans les conditions expérimentales utilisées. Bien que relativement efficace pour trouver des ribozymes actifs *in vitro*, cette technique pour le ciblage de gène *in cellulo* avec le ribozyme HDV a été peu mise à l'épreuve.

En terminant, il n'y a pas encore d'essais cliniques évaluant une technologie basée sur le ribozyme HDV. La spécificité du ribozyme sous sa forme la plus simple fut un frein pendant longtemps. Toutefois, un module qui répond brillamment à cette limite a été développé avant mon arrivée au laboratoire. Mes travaux de thèses portent sur le développement de ce nouvel outil pour le ciblage génique. Ainsi, la dernière partie de l'introduction de cette thèse sera donc consacrée aux détails de ce module de spécificité.

3. Le ribozyme SOFA-HDV comme outil moléculaire

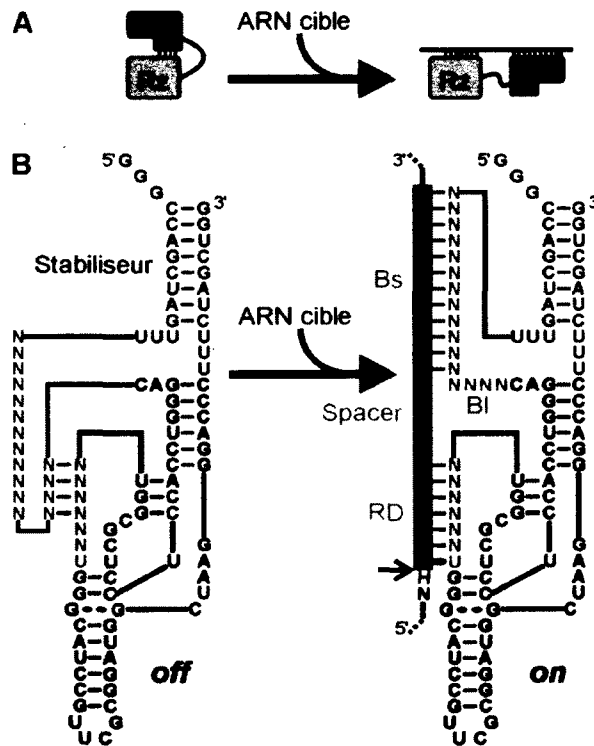
Dans la section précédente, j'ai présenté le ribozyme HDV comme étant un petit ribozyme adapté aux rigueurs des cellules de mammifères ou à tout le moins aux conditions des hépatocytes humains. La principale ombre au tableau de son utilisation pour le ciblage génique (recherche fondamentale et thérapeutique) est son manque présumé de spécificité. Dans le passé, beaucoup d'efforts ont été mis au laboratoire pour amenuiser cette limite et rendre le ribozyme HDV plus spécifique. Les travaux de thèse du Dr Lucien Jr Bergeron constituent une grande avancé sur la question. En fait, il a développé un module qu'il a fusionné au ribozyme HDV pour rendre son activité dépendante de la liaison à son substrat spécifique (Bergeron et Perreault, 2005). Ce module se nomme SOFA, l'acronyme

de «*specific on/off adapter*». Les ribozymes HDV équipés de ce dispositif de spécificité sont appelés des ribozymes SOFA-HDV.

3.1. Concept du module de spécificité SOFA

Tout d'abord, la stratégie développée pour augmenter la spécificité du ribozyme consiste à emprisonner le ribozyme dans une conformation inactive en absence de son substrat spécifique (Introduction, Figure 14A) (Bergeron et Perreault, 2005). La clé pour le passage à une conformation active est la reconnaissance du substrat spécifique du ribozyme en question. Le design rationnel de ce module de spécificité SOFA repose sur trois éléments distincts ayant chacun un rôle bien précis (Introduction, Figure 14B). Initialement, le bloqueur (Bl) s'apparie sur le domaine de reconnaissance (RD), anciennement nommé la tige I, du ribozyme HDV. Cette hélice de quatre paires de bases empêche la liaison et la coupure subséquente d'un ARN qui serait simplement complémentaire (7 pb) au RD. Sous cette forme, le ribozyme est dans une conformation inactive («*off*»). Par la suite, le biosenseur (Bs) se charge de reconnaître le substrat du ribozyme par complémentarité avec celui-ci. La liaison du Bs au substrat entraîne la formation d'une tige (entre 10-15 pb) et favorise le rapprochement du motif catalytique avec son substrat précis. Cet appariement a pour résultat la relâche du bloqueur et donc, le passage vers la conformation active («*on*») du ribozyme. La coupure de l'ARN cible est strictement dépendante de son appariement quasi-parfait au Bs et au RD du ribozyme et ce conjointement (Bergeron et Perreault, 2005). La complémentarité d'un seul des deux domaines de liaison du ribozyme avec un ARN quelconque n'est pas suffisante pour permettre l'activité de coupure (Bergeron et Perreault, 2005). Les nucléotides de l'ARN cible situés entre les sites de liaison du RD et du Bs sont désignés comme le «*spacer*» (Introduction, Figure 14B). La longueur de celui-ci (entre 1 et 10 nucléotides) est flexible et permet d'accommoder des besoins particuliers. Finalement, le dernier élément du module

SOFA est le stabilisateur. Cette tige se forme avec les extrémités du ribozyme ce qui permet de les joindre et de favoriser la stabilité du ribozyme (Introduction, Figure 14B). L'idée est de recréer une structure rappelant celle de la tige II du ribozyme HDV de type sauvage puisque cette structure a été identifiée comme l'un des éléments principaux expliquant la très grande stabilité du ribozyme HDV (Levesque et al., 2002). Selon les données recueillies par Bergeron *et al.*, le stabilisateur n'influencerait pas significativement l'activité de coupure du ribozyme (Bergeron et al., 2005).



Introduction, Figure 14. Concept du ribozyme SOFA-HDV et représentation de sa structure secondaire.

A) Schéma illustrant le concept du module de spécificité SOFA. Le ribozyme HDV (gris) est maintenu dans une conformation inactive (gauche) par le bloqueur (rouge). L'ARN cible permet l'activation du ribozyme via la liaison du biosenseur (Bs, vert) et la relâche du bloqueur (droite). **B)** La structure secondaire du ribozyme SOFA-HDV est formée de celle du ribozyme HDV (gris) additionné du module SOFA (stabilisateur, biosenseur et bloqueur). Les extrémités 5' et 3' sont jointes par une tige (stabilisateur). Le bloqueur (BI, rouge) inactive le ribozyme par

sa liaison avec le domaine de reconnaissance (RD, bleu) en l'absence de l'ARN cible. Le biosenseur (Bs, vert) permet la reconnaissance du substrat spécifique via la formation de paires de bases. L'activation du ribozyme nécessite par la relâche du RD par le BI. Les sites de liaisons dans l'ARN cible sont marqués d'encadrés de couleur correspondant aux éléments de liaison du ribozyme (Bs, vert et RD, bleu) alors que le «*spacer*» est en gris foncé. Les nucléotides variables sont identifiés par N (où N signifie A, C, G ou U) ou H (où H signifie A, C ou U). La petite flèche pointe le site de coupure. Adapté de (Bergeron et Perreault, 2005).

De façon générale le module SOFA, particulièrement l'ajout du biosenseur, a permis d'augmenter l'activité de coupure du ribozyme HDV (Bergeron et Perreault, 2005). Il semble que favoriser la proximité entre le motif catalytique et son substrat spécifique soit un élément clé pour favoriser l'activité de coupure du ribozyme HDV. D'ailleurs, un autre groupe de recherche a utilisé un stratagème très semblable au biosenseur pour stimuler l'activité du ribozyme HDV (Hori et al., 2001; Hori et al., 2006). Ils n'ont toutefois pas inclus un élément permettant d'empêcher la liaison à des substrats imprévus ce qui permet au ribozyme SOFA-HDV de se démarquer significativement en terme de spécificité. Une autre caractéristique majeure du concept SOFA est qu'il peut virtuellement être adapté à n'importe quel outil fait d'acides nucléiques. En effet, il a été adapté avec succès aux ribozymes «*hammerhead*» et «*hairpin*» ainsi que pour le déoxyribozyme 10-23 (pour la coupure d'un ARN cible) (Levesque et al., 2007). Le ribozyme «*hairpin*» (en conditions de ligation) et un ribozyme artificiel sélectionné pour coiffer un ARN («*capping ribozyme*») ont également été additionnés d'un module SOFA (Levesque et al., 2007). Ces travaux ont démontré qu'avec de simples ajustements, le module SOFA est un concept exportable à une multitude de situations pour des réactions *in vitro*.

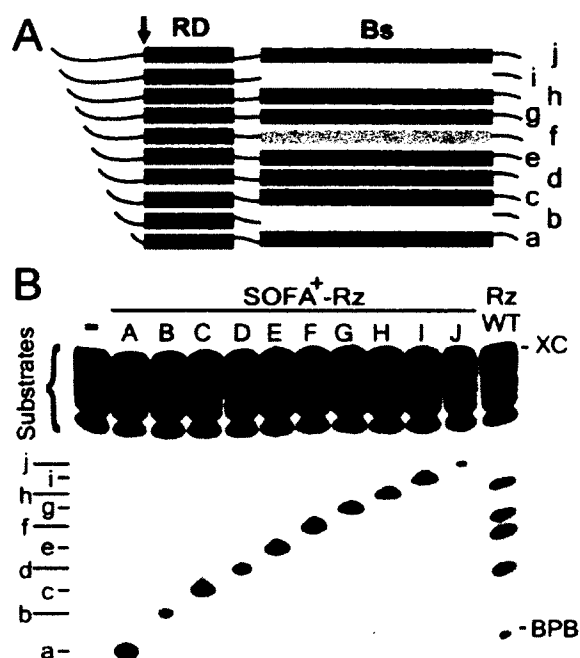
3.2. Spécificité du ribozyme SOFA-HDV

Dans la présentation du ribozyme SOFA-HDV, l'accent est volontairement mis sur la spécificité de ce ribozyme d'ingénierie moléculaire. Cette spécificité provient de

différents éléments, dont certains ont été déduits du ribozyme HDV alors que d'autres sont propres à la version SOFA. Comme je l'ai mentionné précédemment, les travaux sur le ribozyme HDV, et plus particulièrement sur ses préférences au niveau de son site de coupure, ont permis d'élargir les déterminants de sa spécificité au-delà de la formation de sept paires de bases. Ainsi, la séquence en amont du site de coupure et au milieu du RD (tige I dans le ribozyme HDV original) ont été montrés pour influencer l'activité de coupure du ribozyme HDV (Ananvoranich et al., 1999; Deschenes et al., 2000) (Introduction, Figure 13B). Étant donné que le ribozyme SOFA-HDV est un ribozyme HDV avec un module supplémentaire, il est justifié de penser que les règles établies pour le ribozyme HDV sont également valides pour la version SOFA. Bien que logique, cette hypothèse n'a jamais été testée avant mes travaux, même qu'avec le gain significatif d'activité apporté par le module SOFA, la définition d'un site potentiel de coupure a été minimisée au dinucléotide $H_{-1}G_{+1}$ (où H= A, C ou U). De plus, la nature des nucléotides +3 et +4 n'était plus prise en compte (communication personnelle, Michel V Lévesque). En contre partie, dans ces conditions, un plus grand nombre de sites potentiels de coupure pouvaient être testés pour une séquence donnée.

Évidemment la spécificité du ribozyme SOFA-HDV va bien au-delà des éléments du ribozyme HDV. Plusieurs expériences (essais de coupure *in vitro*) ont été effectuées pour démontrer la contribution majeure du module SOFA à la spécificité du ribozyme (Bergeron et Perreault, 2005; Bergeron et al., 2005). Dans les lignes qui suivent, un court résumé soulignant les points majeurs à retenir de ces travaux sera présenté. Comme je l'ai mentionné auparavant, il a été démontré que l'activité du ribozyme SOFA-HDV dépend de la liaison conjointe du biosenseur et du domaine de reconnaissance. Lorsque l'une ou l'autre de ces séquences n'est pas complémentaire avec le substrat, le ribozyme demeure inactif. Une expérience clé démontre clairement cette affirmation et ce, de façon très élégante. Dans cette expérience, une série de couples substrat-ribozyme a été utilisée (Introduction,

Figure 15A). La séquence du RD était constante pour tous les couples alors que leur Bs était différent. Chaque ribozyme a été exposé à un mélange des différents substrats. L'expérience a démontré que seulement les paires parfaitement complémentaires généraient un produit de coupure alors que le ribozyme HDV de type sauvage était en mesure de couper la grande majorité de ces substrats (Introduction, Figure 15B). Lorsque la constante de vitesse observée de coupure (k_{obs}) a été mesurée pour ces couples de substrat-ribozyme, la même conclusion s'est dressée. Uniquement les substrats parfaitement complémentaires aux ribozymes SOFA-HDV (RD et Bs) étaient coupés efficacement. La différence entre la coupure d'un substrat idéal et celle d'un substrat de complémentarité par le même ribozyme peut atteindre jusqu'à trois ordres de grandeurs (généralement entre 25 et 300 fois plus faible) (Bergeron et al., 2005). Une analyse encore plus poussée de la spécificité du ribozyme SOFA-HDV, a permis de montrer que la présence de seulement un mésappariement entre le biosenseur et son site de liaison dans le substrat est suffisant pour entraîner une baisse marquée de l'activité de coupure (Bergeron et al., 2005). En fait, l'activité était de 4 à 15 fois plus élevée pour le substrat parfait que pour ceux avec un mésappariement. De plus, cette différence augmente avec le nombre de mésappariements. Finalement, la distance entre le site de liaison du RD et du Bs influence l'activité de coupure. La liaison du biosenseur au-delà d'une certaine distance («*spacer*» de plus 10 nucléotides) ne permet pas ou très peu l'activation du ribozyme. Le «*spacer*» optimal est de un à cinq nucléotides. Par exemple, si un ARN cellulaire possédait des séquences complémentaires au RD et au Bs, ces séquences devraient être à moins de 10 nucléotides de distances pour être un substrat potentiel du ribozyme SOFA-HDV. Cela réduit le nombre de cibles non-spécifiquement visées («*off-targets*») qui pourraient être reconnues. À partir de ces données *in vitro*, un outil de prédiction de cibles non-spécifiques a été mis au point (Lucier et al., 2006). Brièvement, cette application web permet de retracer dans une banque de donnée (ex. : une liste d'ARNm humains) les séquences nucléotidiques ayant la plus forte probabilité d'être ciblées par un ribozyme SOFA-HDV donné.



Introduction, Figure 15. Démonstration de la spécificité de coupure du ribozyme SOFA-HDV.

A) Représentation de la série substrats utilisés pour les réactions de coupure *in vitro*. Les substrats sont identifiés par des lettres minuscules (a à j) alors que le site de coupure est marqué par la flèche. Le site de liaison du RD (bleu) est constant pour chaque substrat alors que le site de liaison du Bs (différentes couleurs) est spécifique à chacun. Le produit de coupure 5' (ligne à gauche du site de coupure) est de longueur croissante du substrat « a » au substrat « j ». **B)** Essai de coupure d'un mélange de 10 substrats radio-marqués en 5' (a à j) soit par un ribozyme SOFA-HDV (SOFA⁺-Rz) spécifique (A à J) ou par le ribozyme HDV original (Rz WT). Les réactions ont été réalisées à 37°C pour une durée de 3 heures. Le contrôle (-) a été fait en absence de ribozyme. Adapté de (Bergeron et Perreault, 2005).

3.3. Utilisations du ribozyme SOFA-HDV

Au cours de mes études graduées, plusieurs projets impliquant des ribozymes SOFA-HDV sont arrivés à terme. Ainsi, au moment d'écrire cette thèse, les connaissances sur le ribozyme SOFA-HDV et son utilisation pour le ciblage de

gènes sont bien plus vastes qu'elles ne l'étaient à mon arrivée dans le laboratoire. Dans cette courte section, je vais ressortir les éléments majeurs qui étaient déjà connus à mes débuts ou qui sont importants pour les résultats présentés dans cette thèse. Les travaux ayant été publiés durant mes études gradués seront plutôt abordés dans la discussion.

Les travaux initiaux sur le ribozyme SOFA-HDV ont permis de démontrer que ce nouvel outil moléculaire avait un potentiel amélioré pour le ciblage de gènes (Bergeron et Perreault, 2005). En fait, le ribozyme SOFA-HDV a été utilisé pour couper *in vitro* différents substrats ARN. Durant son développement, le ribozyme a été confronté à une flopée de petits substrats (moins de 100 nucléotides). Une fois au point, ce produit d'ingénierie moléculaire a été utilisé pour couper des ARN de plus grandes tailles (plus de 1000 nucléotides). En fait, il a été testé pour couper des ARN provenant des virus VHB et VHC. Dans l'ensemble, cela démontrait le potentiel de ce ribozyme pour cibler ces virus. Par la suite, un ribozyme SOFA-HDV a été utilisé dans un essai d'activité *in cellulo* pour démontrer que le potentiel observé *in vitro* pouvait être transposé à un contexte cellulaire (Bergeron et Perreault, 2005). Dans cette première expérience *in cellulo* pour le ribozyme SOFA-HDV, l'ARN cible était le gène C du VHB. Dans ces conditions de co-transfections, l'inhibition observée fut d'environ 50 %.

Dans cette première expérience *in cellulo*, les ribozymes SOFA-HDV étaient livrés aux cellules par transfection de plasmides. Leur expression était gouvernée par le promoteur CMV et le ribozyme était transcrit sous la forme d'un ARNm. Comme je l'ai discuté précédemment, pour les autres technologies de ciblage génique, la livraison des molécules effectrices est un enjeu majeur. Les ribozymes SOFA-HDV peuvent être livrés aux cellules par les mêmes stratégies développées pour les autres techniques (Sliva et Schnierle, 2010). La majorité des travaux sur le ribozyme SOFA-HDV ont été réalisés soit par transfection d'ADN plasmidique ou par transduction avec des virus. Des essais préliminaires ont été réalisés pour la transfection directe de ribozymes en ARN, mais sans succès (données non-

publiées de Francis P Brière et Jean-Pierre Perreault). L'un des éléments clé pour l'expression des ribozymes est le choix du promoteur. Outre le promoteur CMV (pCMV) ou d'autres promoteurs forts de l'ARN polymérase II, l'utilisation d'un promoteur de l'ARN polymérase III est très populaire pour l'expression de ribozymes ou de shARN. En fait, pour les travaux présentés dans cette thèse, deux stratégies d'expression ont été utilisées. D'abord, pour les ribozymes SOFA-HDV ciblant le VHC, la livraison a été réalisée avec des lentivirus permettant l'expression des ribozymes via le promoteur de l'ARNt^{val} (Levesque et al., 2010). L'utilisation de ce promoteur permet la production de ribozymes fusionnés à l'ARNt^{val} qui pourront être exportés efficacement au cytoplasme ((Kato et al., 2001) une condition primordiale pour cibler l'ARN du VHC qui est confiné à ce compartiment cellulaire. D'autre part, les ribozymes ciblant le VIH ont été livrés par la transfection du vecteur plasmidique psiRNA-hH1GFPzeo (Invitrogen) (Scarborough et al., Soumis). Avec ce vecteur, les ribozymes SOFA-HDV ont été exprimés via le promoteur humain de l'ARN de la RNase P (pH1). Dans ces conditions, le ribozyme SOFA-HDV ne contient pas de séquences excédentaires et devrait être restreint au noyau étant donné l'absence de signal d'export.

4. Objectifs du projet

L'ingénierie du ribozyme SOFA-HDV a établi les bases d'un système d'inactivation de gènes original et spécifique. Le développement de cet outil pour cibler des ARN d'intérêts nécessite l'utilisation à plus grande échelle de ribozymes *in cellulo* ainsi que l'élaboration d'une stratégie pour la sélection des meilleurs candidats. Les travaux réalisés ces dernières années permettent une meilleure compréhension du fonctionnement de ce ribozyme et l'échafaudage d'une méthode claire et efficace pour son utilisation. Ainsi, mon objectif principal de recherche est d'appliquer le ribozyme SOFA-HDV pour le ciblage génique et de synthétiser une marche à suivre pour la sélection de ceux-ci.

4.1. Application du ribozyme SOFA-HDV pour cibler des virus

En 2006, lorsque j'ai débuté mes études graduées, un nombre très restreints d'études avait testé l'activité de coupure du ribozyme HDV dans un milieu cellulaire. L'utilisation de sa version SOFA se limitait à une simple preuve de concept (Bergeron et Perreault, 2005). Il y avait bien quelques travaux en cours, mais ceux-ci concernaient un nombre limité de ribozymes SOFA-HDV (Fiola et al., 2006; Robichaud et al., 2008). L'évaluation de ribozymes SOFA-HDV à plus grande échelle n'avait pas été abordée et représentait une limite importante dans l'émergence de cette technologie. Les virus représentent des modèles intéressants pour le développement d'outils de ciblage génique. Dans le cas du VHC et du VIH, l'intérêt est d'autant plus grand, car il n'existe toujours pas de cure pour traiter ces infections. De plus, de nombreux patients sont porteurs de virus résistants aux thérapies actuellement proposées. Ainsi, l'application du ribozyme SOFA-HDV pour le ciblage *in cellulo* d'ARN viral est un élément important pour illustrer le potentiel de cet outil moléculaire.

L'objectif général pour ce projet était d'effectuer le criblage d'un grand nombre de ribozymes SOFA-HDV pour leur activité de coupure *in cellulo*. De façon plus spécifique, j'avais comme objectif de dessiner des ribozymes SOFA-HDV et ainsi tester la méthode de design mise au point pour le ribozyme HDV. Par la suite, je voulais évaluer la capacité du ribozyme SOFA-HDV à cibler spécifiquement des ARN viraux *in vitro* et *in cellulo*. Mon but ultime étant de tester le potentiel du ribozyme SOFA-HDV pour l'inhibition de la réplication d'un virus.

4.2. Optimisation de la méthode pour la production de ribozyme SOFA-HDV

Les observations et données collectées avec les projets d'application du ribozyme SOFA-HDV fournissent une base d'information primordiale pour la compréhension du ribozyme SOFA-HDV (Bergeron et Perreault, 2005; Fiola et al., 2006; Laine et al., 2011; Levesque et al., 2010; Motard et al., 2011). Avec son module de spécificité, ce ribozyme est forcément plus complexe que son prédécesseur le ribozyme HDV. Il est donc réaliste de penser qu'un plus grand nombre de facteurs doivent être pris en compte pour la production des ribozymes SOFA-HDV. Évidemment, les méthodes de design pour le ribozyme HDV original ne tiennent pas compte des éléments du module SOFA. Plus particulièrement, la séquence du biosenseur et du bloqueur ne sont pas évaluées par ces méthodes alors qu'elles sont déterminées par la séquence de l'ARN cible. De plus, le criblage de tous les ribozymes possibles pour un ARN donné n'est pas une tâche réaliste et efficace pour la sélection des ribozymes SOFA-HDV. Cette façon de faire est fastidieuse et nécessite un investissement maximal. Il est donc crucial de revoir notre méthode de design et de l'adapter à la réalité plus complexe du ribozyme SOFA-HDV.

Mon objectif principal pour ce second volet de mes travaux était d'optimiser la façon de produire des ribozymes SOFA-HDV pour le ciblage génique. Pour atteindre cet objectif, je devais d'abord analyser la relation entre l'activité *in vitro* de l'éventail de ribozyme SOFA-HDV testés en fonction de leur séquence et de celle

de leur site cible. Je devais aussi évaluer l'influence des différents éléments du module SOFA sur l'activité du ribozyme. La réalisation de ces deux objectifs spécifiques devait me permettre de déterminer les paramètres favorisant l'activité du ribozyme SOFA-HDV et pouvant servir de critères de présélection. Finalement, je devais élaborer une méthode de design des ribozymes SOFA-HDV permettant de limiter le nombre de ribozymes à tester pour un ARN cible donné.

RÉSULTATS

Chapitre 1 : Ciblage du VHC avec des ribozymes SOFA-HDV

ARTICLE : *Investigating a New Generation of Ribozymes in Order to Target HCV*

Michel V. Lévesque, Dominique Lévesque, Francis P. Brière and Jean-Pierre Perreault

Article publié dans : *PLoS ONE* (2010) 10;5(3):e9627

Avant-propos :

J'ai réalisé et analysé la majorité des travaux présentés dans l'article. Entre autres, j'ai effectué tous les essais de coupure *in vitro* et les essais *in cellulo*. Pour la préparation de l'article, j'ai monté toutes les figures et participé activement à la rédaction du manuscrit sous la supervision de Jean-Pierre Perrault.

Résumé

Pendant longtemps, les approches basées sur les acides nucléiques visant le contrôle de la propagation du virus de l'hépatite C (VHC) ont été considérées comme ayant un fort potentiel. Dans ce but, les ribozymes (c.à.d. enzymes d'ARN), qui reconnaissent spécifiquement et subséquentment catalysent la coupure de leur substrat ARN, représentent un outil moléculaire attirant. Ici, les propriétés uniques d'une nouvelle génération de ribozymes sont mises à profit pour le développement d'une technologie efficace et durable, basée sur un ribozyme ciblant l'ARN de polarité positive du VHC. Ces ribozymes résultent de la fusion d'un module de spécificité SOFA («*specific on/off adaptor*») au domaine ribozyme dérivé du virus de l'hépatite *Delta* (HDV). Le premier élément permet l'activité de coupure seulement en présence du substrat ARN désiré, alors que le second était le premier ARN catalytique rapporté pour fonctionner naturellement dans les cellules humaines, spécifiquement dans les hépatocytes. Dans le but de maximiser les chances de succès, une approche étape par étape a été utilisée pour le design et la sélection des ribozymes. Cette approche inclut l'utilisation de méthodes bioinformatiques et biochimiques pour l'identification de sites possédant le meilleur potentiel pour le ciblage, et subséquentment des tests *in vitro* pour l'activité de coupure des ribozymes SOFA-HDV correspondant. Ces efforts ont mené à une amélioration significative du design des ribozymes. L'habileté des ribozymes SOFA-HDV résultants à inhiber la réplication du VHC a ensuite été examinée à l'aide d'un réplicon exprimant la luciférase. Bien que certains ribozymes ont montré un haut niveau de coupure *in vitro*, aucun apparaît comme un potentiel inhibiteur *in cellulo* à long terme. L'analyse des récentes découvertes de la biologie cellulaire du VHC pourrait expliquer cet échec, ainsi qu'amener des idées sur les limites potentielles de l'utilisation de drogues faites d'acides nucléiques pour le contrôle de la propagation du VHC. Finalement, les conclusions ci-hauts reçoivent le support d'expériences réalisées sur une collection de ribozymes SOFA-HDV dirigés contre l'ARN de polarité négative du VHC.

Abstract

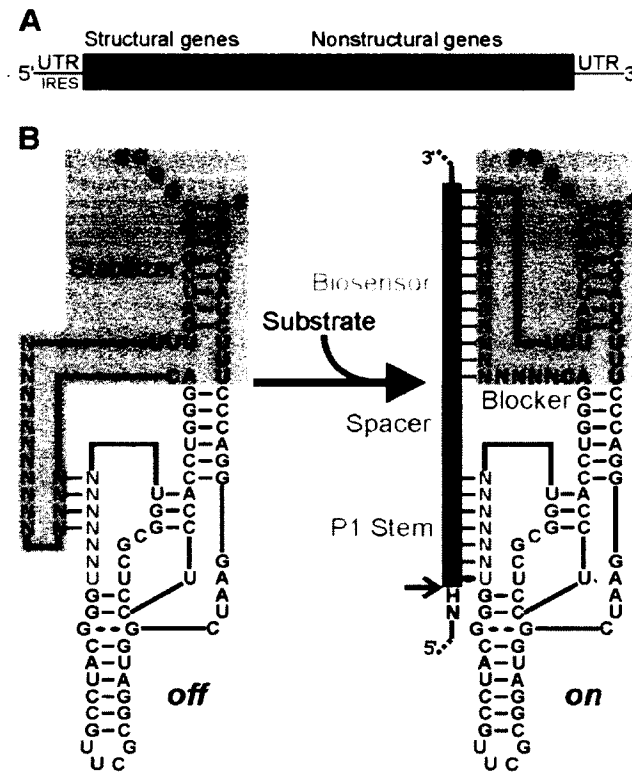
For a long time, nucleic acid-based approaches directed towards controlling the propagation of Hepatitis C Virus (HCV) have been considered to possess high potential. Towards this end, ribozymes (i.e. RNA enzymes) that specifically recognize and subsequently catalyze the cleavage of their RNA substrate present an attractive molecular tool. Here, the unique properties of a new generation of ribozymes are taken advantage of in order to develop an efficient and durable ribozyme-based technology with which to target HCV (+) RNA strands. These ribozymes resulted from the coupling of a specific *on/off* adaptor (SOFA) to the ribozyme domain derived from the Hepatitis Delta Virus (HDV). The former switches cleavage activity “*on*” solely in the presence of the desired RNA substrate, while the latter was the first catalytic RNA reported to function naturally in human cells, specifically in hepatocytes. In order to maximize the chances for success, a step-by-step approach was used for both the design and the selection of the ribozymes. This approach included the use of both bioinformatics and biochemical methods for the identification of the sites possessing the greatest potential for targeting, and the subsequent *in vitro* testing of the cleavage activities of the corresponding SOFA-HDV ribozymes. These efforts led to a significant improvement in the ribozymes’ designs. The ability of the resulting SOFA-HDV ribozymes to inhibit HCV replication was further examined using a luciferase-based replicon. Although some of the ribozymes exhibited high levels of cleavage activity *in vitro*, none appears to be a potential long term inhibitor in cellulo. Analysis of recent discoveries in the cellular biology of HCV might explain this failure, as well as provide some ideas on the potential limits of using nucleic acid-based drugs to control the propagation of HCV. Finally, the above conclusions received support from experiments performed using a collection of SOFA-HDV ribozymes directed against HCV (–) strands.

Introduction

Hepatitis C Virus (HCV) is the major worldwide cause of both blood transfusion-associated and sporadic non-A, non-B hepatitis. Estimates place the number of HCV-infected individuals worldwide at 170 million, representing nearly 3 % of the world's population [1,2]. In 20-30 % of these patients the HCV infection is acute and the virus is naturally cleared. The majority (70-80 %) remain chronically infected, leading to various clinical outcomes including an asymptomatic carrier state with normal or almost normal liver functions, acute hepatitis and, about 50 % of the time, chronic hepatitis. Today, chronic HCV infections can only be treated with pegylated interferon- α (IFN- α) [1,2]. Unfortunately, about 70 % of compliant patients experience a relapse, and only 25 % maintain low serum alanine aminotransferase levels. Although a combination therapy including the purine nucleoside analogue ribavirin improves the rate of success, only 40 % of patients infected with HCV achieve a sustained response. Moreover, in clinical practice many patients do not qualify for IFN- α therapy for several reasons. Hence, only a minority of patients with chronic hepatitis C can be successfully treated and, consequently, projections indicate that the mortality rate from hepatocellular carcinoma associated with chronic hepatitis C will further increase in the next 15-20 years [1,3]. Clearly, it is an urgent mission to search for innovative strategies with which to treat/cure HCV infected patients.

HCV, a member of the *Flaviviridae* family, is an enveloped virus with a plus-stranded RNA genome ~9,6 kb in length (Figure 1A) [4]. Six major genotypes and more than 50 minor subtypes have been distinguished. The high replicative activity of the virus, together with the lack of a proofreading function of the viral RNA polymerase, provides the basis for the significant genetic variability of HCV. This genome carries a single long open reading frame (ORF) which encodes a polyprotein of ~3,010 amino acids that is subsequently cleaved both co- and post-translationally into the mature viral proteins. The ORF is flanked by structured 5' and 3' untranslated regions (UTRs) that are important for both the translation of the

polyprotein and the replication of the genome. For example, the 5' UTR is highly conserved among different HCV isolates and contains an internal ribosomal entry site (IRES) essential for the cap-independent translation of the viral RNA [5].



Chapitre 1, Figure 1. Schematic representations of both the HCV genome and the SOFA-HDV ribozyme.

(A) HCV genomic RNA organization including the structured 5' and 3' UTRs. (B) Secondary structure of both the *off* (left) and *on* (right) conformations of the SOFA-HDV ribozyme. The SOFA module is highlighted in grey. The P1 domain (ribozyme binding domain), the biosensor, the blocker and the stabilizer domain are indicated blue, green, red and brown, respectively. The spacer region separating the substrate' sequences binding by the P1 and biosensor domains of the ribozyme is indicated in grey. The arrow indicates the cleavage site. All structures have been previously described [29,30].

Several antiviral strategies are being explored for the treatment of HCV infections [6,7]. For example, antiviral inhibitors that block essential viral enzymes, such as the NS3-4 protease or the NS5B polymerase, have led to the discovery of

several compounds that are currently at different stages of development [8-10]. Moreover, gene-inactivation approaches, including the use of antisense oligonucleotides and, more recently, RNA interference (RNAi), also show some promise in terms of being able to inhibit both HCV replication and gene expression in experimental systems [11-15]. However, the RNAi approach has been shown to trigger an immunological response, which poses an important limitation in terms of further therapeutical applications [16-18]. Ribozymes (Rz), RNA molecules that catalyze the cleavage of RNA substrates, are an interesting alternative to the RNAi approach to gene inactivation. Both ribozymes and deoxyribozymes have been demonstrated to block both viral replication and gene expression in experimental systems both *in vitro* and *in cellulo* [19-22].

Among the different ribozymes, Hepatitis Delta Virus ribozyme (HDV Rz) was for a long time the sole example derived from an RNA species naturally found in human cells (i.e. the infected hepatocytes) [23,24]. As a result, HDV Rz offers several unique properties as a potential tool, including the natural ability to function in the presence of human proteins and at the physiological magnesium concentration (1 mM Mg²⁺). Furthermore, HDV Rz has a long half-life, regardless of both the cell line tested and the means of transfection used [23]. Thus, HDV Rz appears to be well adapted to the human cell environment and is therefore an interesting potential candidate for the development of a gene-inactivation system. This potential has subsequently been demonstrated *in vitro* as well as *in cellulo* by the use of HDV ribozyme to cleave various natural mRNA [24-28]. The HDV Rz folds into a model pseudoknot secondary structure in which substrate recognition is based solely on the formation of the P1 stem, a stem that involves only 7 base pairs (bp) (Figure 1B). Recently, molecular engineering has led to the development of a novel target-dependent riboswitch that increases HDV Rz fidelity [29]. This latter ribozyme possesses a specific *on/off* adapter (SOFA) that switches the cleavage activity from *off* (i.e. like a “safety lock”) to *on* solely in the presence of the desired RNA substrate. The SOFA module is composed of three domains: a

blocker, a biosensor and a stabilizer [30]. The blocker sequence inhibits the cleavage activity of the ribozyme by binding, in *cis*, the P1 domain of the ribozyme portion. This provides an inactive SOFA-HDV ribozyme, i.e. *off* conformation. The biosensor must bind its complementary sequence on the substrate in order to unlock the SOFA module, thereby permitting the folding of the catalytic core into the *on* conformation. Both the blocker and the biosensor have been shown to increase the substrate specificity of the ribozyme's cleavage by several orders of magnitude compared with the wildtype HDV ribozyme [29]. This is due mainly to the addition of the biosensor domain that increases the binding strength of the HDV ribozyme to its target but it is also due to the fact that the blocker domain interacts with the P1 domain and decreases its binding capacity. It is important to note that the sequences of the substrate binding both the P1 and the biosensor domains are not contiguous, but rather separated by a small region named spacer that usually varies from 4 to 7 nt for optimal design [30]. Finally, the presence of the stabilizer (i.e. the stem that brings together both the 5' and 3' extremities), which has no effect on the cleavage activity, stabilizes the SOFA-HDV ribozymes in vivo against ribonucleases [29]. This new development provides a highly specific and improved HDV-based molecular tool that displays significant potential for application in the fields of both functional genomics and gene therapy.

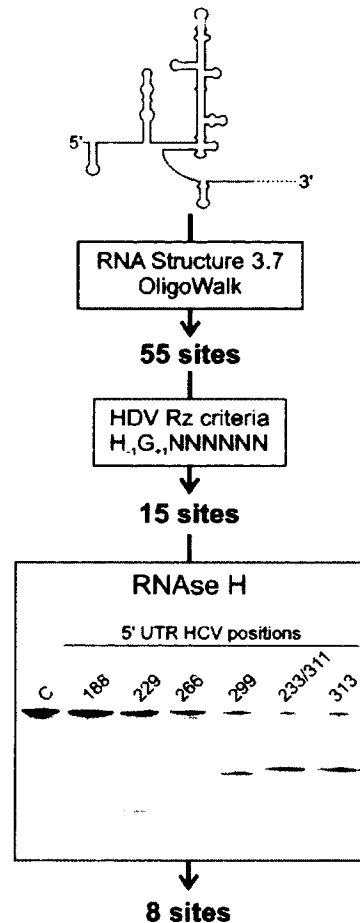
In the present work, the unique properties of the HDV Rz that are improved by the presence of the SOFA adapter are used in order to attempt to develop a ribozyme-based technology with which to target HCV. Sites with greatest potential for targeting in the HCV strands of both polarities were identified, followed by a screen to find the most active SOFA-HDV ribozymes.

Results

Identification of potential target sites in the HCV (+) strand RNA

Since HCV is known to exhibit several genotypes and many subgenotypes, the identification of the most highly conserved sequences within its genome is critical to the development of nucleic acid based drugs that possess the ability to specifically bind to most, if not all, of the sequence variants. Consequently, only the first 341 nucleotides (nt) of the HCV (+) strand that correspond to the 5' UTR, a region that is known to be highly conserved, were considered for targeting. Moreover, only the HCV (+) strand was analyzed because it has been reported to be significantly more abundant than its (-) counterpart [31,32].

The initial step in developing a ribozyme capable of specifically cleaving the HCV (+) RNA was the identification of the sites most accessible for targeting. Sites located in single-stranded regions should exhibit a higher ribozyme binding potential than those located in double-stranded regions [33]. In order to address this issue, a bioinformatics approach was initially used to identify the most accessible sites (see Figure 2). Using the RNA folding software RNA Structure 3.7 [34], a series of potential secondary structures that can be formed by the 341-nt RNA sequence were obtained. The twenty most stable structures were then analyzed using the OligoWalk software, which provides, in terms of energy, the degree of accessibility of short sequences along the different structures [34]. A window size of 7 nt, corresponding to the binding domain of the HDV Rz, was used (i.e. the P1 stem). Using a cut-off value of smaller than -12 kcal/mol for the predicted binding constant of the substrate to the ribozyme, this analysis identified 55 potentially accessible sites.



Chapitre 1, Figure 2. Schematic representation of the 3-step procedure used for the identification of the sites possessing the greatest targeting potential in the HCV 5'-UTR.

Step 1 involved a bioinformatic analysis that included both the prediction of the secondary structure and the identification of the 7 nt stretches most likely to be bound by the ribozyme's P1 region using both the RNA structure 3.7 and Oligowalk softwares. Step 2 involved the selection of the sequences that fulfill the HDV ribozyme requirements. Step 3 involved the RNase H hydrolysis assay. The autoradiogram shown corresponds to a typical 5% polyacrylamide gel performed for the analysis of 6 potential sites. The positions of proposed cleavage sites are identified at the top of the gel. The negative control performed in the absence of any oligonucleotide is indicated by the letter C.

The resulting sites were then further analyzed to see if they fulfilled the criteria essential for efficient HDV R_z cleavage. Specifically, the first nucleotide downstream of the cleavage site (position +1) has to be a guanosine in order to allow formation of a critical GU Wobble bp with the ribozyme (position +1, Figure 1B), and the nucleotide upstream of the cleavage site should not be a guanosine (position -1, Figure 1B). Application of these constraints shortened the list to 15 potential target sites (Figure 2). Next, biochemical assays were performed in order to validate the bioinformatic predictions. Specifically, RNase H hydrolysis assays were performed *in vitro* using 7-nt oligonucleotides corresponding to the binding domain of the HDV R_z and a 575-nt HCV-derived RNA species corresponding to the 5' region of the genome. RNase H specifically cleaves the RNA half of an RNA-

DNA duplex. In this analysis, the most accessible sites, that is, those that bind the oligonucleotides, should be cleaved by the RNase H. Five'-end-labeled HCV-derived transcripts were pre-incubated with the oligonucleotides prior to being incubated with RNase H and analyzed samples. Figure 2 shows an example of autoradiogram performed with a 5 % polyacrylamide gel. The time of migration used provides a resolution suitable for the analysis of the sites located between positions 188 to 313. In this example, 5 out of the 6 oligonucleotides generated specific hydrolysis products, although at different levels. It is important to note that because the sites 233 and 311 share the same 7 nt sequence, one oligonucleotide permitted the evaluation at both sites. Efficient hydrolysis was detected at position 311 while only a faint band at position 233. Several electrophoresis with various migration times were performed in order to evaluate the 14 different oligonucleotides. These experiments permitted the validation of 8 potential sites and data was compiled at Table 1. These 8 sites are located throughout the 5' UTR, albeit with a predominance within the region from 299 to 318 in which four of the eight are located.

Chapitre 1, Tableau 1. Compilation of the data from the RNase H hydrolysis of HCV (+) RNA.

Cleavage positions	Oligonucleotide sequences	RNase H hydrolysis
15	CGCCCC	-
110	TGGAGGC	-
117	GGGGTCC	+
176	TCCTGGC	-
184	CCCGGTC	--
188	AGGACCC	-
217	TCCAGGC	-
229	CACGCC	-
233	GGGGCAC	+
235	CGGGGGC	-
266	GCGACCC	+
299	AGCACCC	-
311	GGGGCAC	+
313	CCGGGGC	-
318	ACCTCCC	+

doi:10.1371/journal.pone.0009627.t001

The site located at position 188 was retained as a control for the subsequent step. The analysis of the 5' to 3' distribution of the selected sites revealed a bias that lowered representation in the region spanning positions 1 to 200. Only the site at position 117 was retrieved within this region. In order to obtain more sites in the region spanning positions 1 to 200, 8 other sites were selected. The selection for these new sites was based on the presence of a high level of sequence conservation as well as their potential accessibility according to the reported crystal structure of the HCV IRES [5] (see Table 2). Together, these analyses provided a collection of 17 potential sites for targeting in the HCV (+) strands.

Chapitre 1, Tableau 2. Compilation of the *in vitro* cleavage of HCV (+) strand by the SOFA-HDV ribozyme data.

Cleavage positions	Rz P1 sequences	Rz Biosensor sequences	Cleavage %
117	GGGGUCU	UGGCUCUCCCGG	02 ± 1
188	AGGACCU	AUUGAGCGGGUU	02 ± 1
229	CACGCCU	UAGCAGUCUCGC	07 ± 1
233	GGGGCAU	CGGCUAGCAGUC	01 ± 1
266	GCGACCU	GUACCACAAGGC	10 ± 1
299	AGCACCU	CUCCCGGGGCAC	13 ± 3
311	GGGGCAU	GGUCUACGAGAC	07 ± 1
313	CCGGGGU	ACGGUCUACGAG	33 ± 2
318	ACCUCCU	GGUGCACGGUCU	29 ± 1
21	GAGUGUU	GGGAGUGAUCUA	21 ± 1
60	UGAAGAU	GCUAGACGCUUU	39 ± 3
68	UUUCUGU	ACGCCAUGGCUA	35 ± 1
82	CCAUGGU	CGACACUCAUAC	04 ± 1
135	AUGGCUL	CGGUUCCGAGA	07 ± 1
143	AGACCAU	GUACUCACCGGU	02 ± 1
170	CAAUUCU	AAGAAAGGACCC	01 ± 1
200	UUGAUCU	UCCAGGCAUUGA	04 ± 1

The sequences for each SOFA-HDV Rz corresponding to their P1 and biosensor recognition domains are listed. The upper section includes the SOFA HDV Rz targetable sites as identified by bioinformatic and biochemical procedures while the lower section includes those based on the reported secondary and crystal structures.

doi:10.1371/journal.pone.0009627.t002

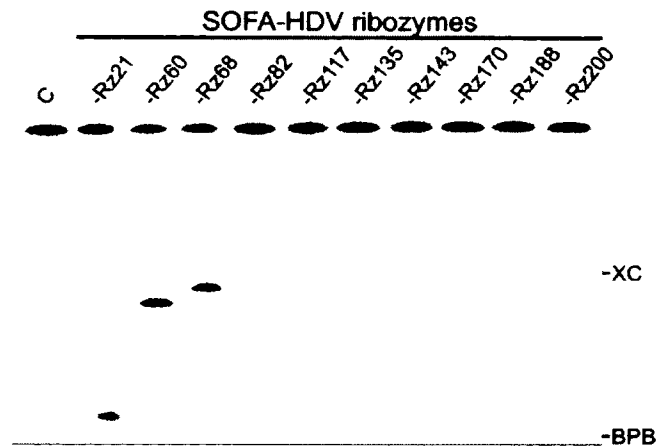
Designing of the SOFA-HDV ribozymes

The various selected target sequences of the HDV ribozymes were all highly conserved among the major HCV genotypes (data not shown), a prerequisite for the development of a viable HDV Rz-based therapeutical approach. Preliminary sequence homology searches between the HDV Rz and a bank of human transcripts revealed that the ribozymes could potentially target several human transcripts (version 35.1 of the human genome). These potential off-target effects resulted from the limited size of the recognition domain present in the original HDV Rz (only 7 nt). It has been estimated that a minimum sequence of 13-14 base pairing nucleotides is required for the specific targeting of an unique RNA species in the human transcriptome [35]. In order to improve the substrate specificity of each ribozyme, a SOFA module was designed for each (see the description of the SOFA in the Introduction section and in reference [29]). In each case the biosensor sequences added are presented in Table 2. The sequence located between the ribozyme and the biosensor binding domains on the HCV target, the so-called spacer domain, was always kept between 4 and 7 nt in length as this length has been shown to be optimal [30], except in the cases SOFA-HDV-Rz170 and -Rz188. The former includes a spacer of 11 nt, while the latter possesses one of 10 nt, and both were required in order to improve the biosensor binding site's conservation level. As a result, the designed SOFA-HDV ribozymes should target only the desired sequence's region, and should therefore be specific enough to target only HCV. Moreover, the addition of the biosensor sequence of the SOFA domain permits the differentiation between the recognition sites of the SOFA-HDV-Rz233 and -Rz311 ribozyme even though they share the same sequence for the P1 region of the ribozyme domain.

***In vitro* cleavage assay of the potential SOFA-HDV ribozymes**

The next step was the synthesis of a collection of SOFA-HDV ribozymes and the identification of those that demonstrate the highest levels of cleavage activity *in*

vitro. The ability of each SOFA-HDV ribozyme to cleave 5'-radiolabeled 575-nt HCV RNA was tested under single-turnover conditions ($[Rz] \gg [S]$). A typical example is illustrated in Figure 3. The time of migration used for this specific electrophoresis was suitable for the resolution of the sites located in the 5' end of the HCV IRES (positions 1 to 200). In this example, the SOFA-HDV-Rz21, -Rz60 and -Rz68 ribozymes exhibited high levels of cleavage. All others cleaved less efficiently. Five ribozymes exhibited cleavage levels $>20\%$, a level that we considered relatively high for the cleavage of a long and highly structured RNA species. Several electrophoresis using different migration time were performed to monitor the cleavage level at the various positions (raw data not shown) and the data were compiled at Table 2. Overall, 7 out of 17 SOFA-HDV ribozymes exhibited significant cleavage of the substrate (i.e. cleavage percentage $>10\%$; Table 2). Clearly, these results suggest that the time invested in identifying the most accessible sites was productive, as without it this number would certainly have been smaller. The 5 SOFA-HDV ribozymes that exhibited a cleavage level below 2% were discarded at this step. This includes as expected SOFA-HDV-Rz188 that was the negative control. The other 12 SOFA-HDV ribozymes were conserved for the subsequent step.



Chapitre 1, Figure 3. Typical autoradiogram of an 8 % polyacrylamide gel performed in order to analyze SOFA-HDV ribozyme cleavage *in vitro*.

The experiments were performed using 5'-end-labeled HCV transcripts in the presence of an excess of SOFA-HDV ribozyme. The SOFA-HDV ribozymes are identified at the top of the gel. The negative control performed in the absence of ribozyme is indicated by the letter C. The positions of the xylene cyanol (XC) and bromophenol blue (BPB) marker dyes are indicated.

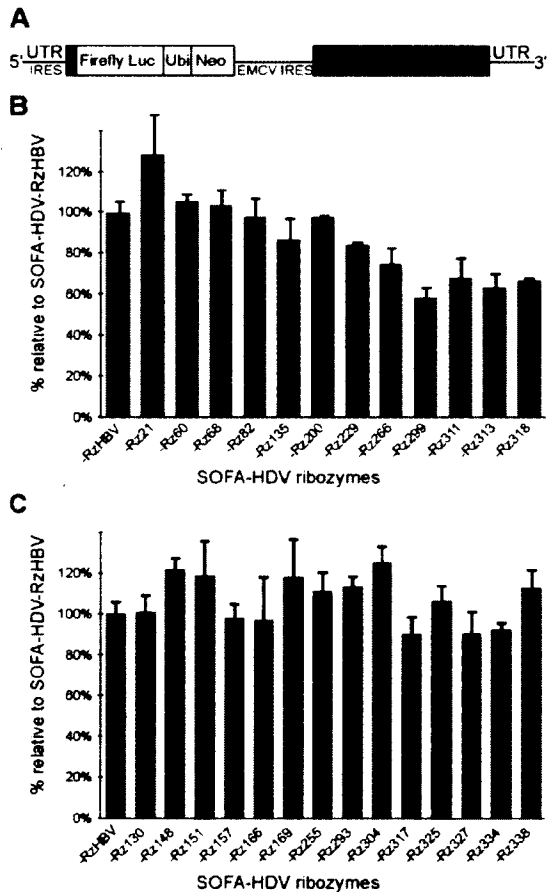
Evaluation of the SOFA-HDV ribozymes cleaving the HCV (+) strand *in cellulo*

The potential abilities of all SOFA-HDV ribozymes to inhibit an HCV replicon were tested. Prior to this experiment, the ribozymes were cloned adjacent to a tRNA^{Val} in modified pLenti plasmids. The resulting plasmids were used to produce lentivirus which in turn were used to transduce Huh-7 cells (see Materials and Methods). The use of the tRNA^{Val} as both promoter and leader sequence has previously been proven to be successful for both the production and the localization of HDV ribozymes [26]. It permits the production of a large amount of the chimeric RNA by the host RNA pol III. In addition, it is well known that the tRNA motif directs the expressed ribozyme to the cytoplasm, where it hybridizes to the targeted mRNA [36, 37]. The expression of all chimeric tRNA^{Val}:SOFA-HDV ribozymes was confirmed at the time of the replicon inhibition assay (i.e. 48 h post transduction, data not shown). Results from preliminary expression experiments showed that 48 to 72 h post transfection was the optimal window for the expression of the

ribozymes. The excess of chimeric tRNA^{Val}:SOFA-HDV ribozymes over targeted HCV strands remained undetermined. However, detection of tRNA^{Val}:SOFA-HDV ribozymes was performed by Northern blot hybridizations and confirmed their relatively large abundance while HCV strand were barely detectable using the same technique (data not shown). These results support the notion that the ribozymes were in excess compared to the HCV substrate, although this does not constitute direct evidence.

It is important to note that during an experiment neither G-418 nor antibiotic were used in order to reduce interference with cell transduction as well as to avoid the positive selection of HCV replicon containing cells while measuring the effect of the SOFA-HDV ribozymes. The Huh-7 cells containing the subgenomic bicistronic replicon Luc-ubi-neo-ET, which includes the firefly luciferase ORF that provides a quick enzymatic assay for drug screening, was used (Figure 4A) [38]. Similar replicon systems were used in many other studies in order to screen various nucleic acid-based approaches including siRNA [e.g. 39-43]. A colorimetric assay was performed for normalization purposes by considering the total protein concentration. Each SOFA-HDV ribozyme tested was compared to an irrelevant one possessing a recognition sequence that should not be able to interact with either the HCV RNA, nor with any human mRNA (as determined by a homology search). The latter SOFA-HDV ribozyme was originally engineered to target the hepatitis B virus (SOFA-HDV-RzHBV) [29]. Its use permitted determination of the relative percentage of luciferase detected in the presence of each SOFA-HDV ribozyme, as compared to that observed in the presence of the irrelevant one which was arbitrarily set at a luciferase/protein ratio of 100 %. At least two or three biological replicates were performed for all ribozymes, and typical results are illustrated in Figure 4B. Simple cell exposition to polybrene, the transduction reagent, or even the expression of the tRNA^{Val} alone, led to a variation of +/- 20 % (data not shown). Consequently, it was decided that a ribozyme must decrease the luciferase/protein ratio by at least 20 % in order for it to be considered a significant

inhibition of replication. Using this criteria, 7 SOFA-HDV ribozymes were rejected (i.e. SOFA-HDV-Rz21, -Rz60, -Rz68, -Rz82, -Rz135, -Rz200, and -Rz229). One ribozyme exhibited intermediate inhibition level at 25 % (i.e. SOFA-HDV-Rz266). Most importantly, 4 ribozymes exhibited cleavage activities that led to >30 % reductions in the luciferase/protein ratio: the SOFA-HDV-Rz311, -Rz313 and -Rz318 reduced the level between 30 and 40 %, while SOFA-HDV-Rz299 caused even more significant reduction (i.e. a 42 % luciferase/protein).



(A) Schematic representation of the HCV replicon used (described previously [38]). **(B)** and **(C)**. Histograms of the relative luciferase activities detected for the HCV (+) and (-) strands, respectively. Luciferase activity was detected in the presence of all SOFA-HDV ribozymes tested, and was reported relative to that determined for an irrelevant SOFA-HDV ribozyme whose level was arbitrarily set at 100 %. The latter SOFA-HDV ribozyme was designed to target the hepatitis B virus and does not possess the sequences required in order to recognize the HCV strands of either the (+) or the (-) polarity. The characterization of this SOFA-HDV ribozyme was previously reported [29].

Chapitre 1, Figure 4. Analysis of the inhibition of the HCV replicon by SOFA-HDV ribozymes.

The experiment was repeated several times, and always produced similar results and experimental variation when the same lentivirus preparation was used, while use of different viral preparations increased the experimental variation (data not shown). The use of different lentivirus concentrations did not result in improved inhibition of the replicon. Similarly, prolonged exposition of the cells to the lentiviruses did not lead to more important reduction of the luciferase activity. In summary, some SOFA-HDV ribozymes exhibited sufficiently high cleavage levels of the IRES that they significantly reduced the luciferase activity; however, the observed level of replicon inhibition was never high enough to warrant further development of the SOFA-HDV ribozyme as an antiviral drug against HCV. In other words, although the new generation of ribozyme, specifically the SOFA-HDV ribozyme, provides a highly specific and improved tool with significant potential for practical application, its potential for controlling HCV replication appears limited.

Designing SOFA-HDV ribozymes that cleave the HCV (–) strand

One possible explanation for the limited inhibitory effect of the SOFA-HDV ribozyme against HCV (+) strands might be the fact that the replication complex is compartmentalized by a lipid bilayer membrane [32,44-45]. This localization considerably reduces the accessibility of the HCV RNA (+) strand to the SOFA-HDV ribozyme during replication. In fact, the action of the ribozymes may take place mainly, if not exclusively, during the translation of the (+) strand that encodes the polyprotein. If this is indeed the case, it should be almost impossible to target the HCV (–) strand *in cellulo* because they are non-coding, and, therefore, are not accessible during the translation step. In order to verify this hypothesis, SOFA-HDV ribozymes specifically targeting the 3'-end of the HCV (–) strand that is the counterpart of the 5' UTR found in the (+) strand were designed and tested for cleavage activity both *in vitro* and *in cellulo*.

Instead of applying a complex design strategy, HCV sequence alignment was used to identify both highly conserved sequences in the HCV genome and potential sites for HDV ribozyme cleavage (i.e. those harboring the sequence H₋₁G₊₁; with H being C, U or A). The 3'-end region of the HCV (-) strand was shown to fold into a secondary structure, but not one as stable as the IRES adopted by the HCV (+) polarity strand [46]. Application of this criteria led to the identification of 14 potential sites, all of which were located in the last 341 nt of the HCV (-) strand (see Table 3). The cleavage activities of all of the corresponding SOFA-HDV ribozymes were assayed *in vitro* using a 575-nt transcript derived from the HCV (-) strand that was the counterpart of the one used for the (+) polarity above. All SOFA-HDV ribozymes exhibited cleavage (>1 %, Table 3). Three ribozymes exhibited a limited level of < 10 %, seven exhibited a significant level of between 10 and 20 % and four exhibited a cleavage level over 20 %. Specifically, SOFA-HDV-Rz166⁻, -Rz327⁻, -Rz304⁻ and -Rz293⁻ exhibited cleavage activities of 24 %, 43 %, 70 % and 85 %, respectively (see Table 3). Observation of such significant levels of cleavage for the HCV (-) strand provides a good indication that it is less stable than its (+) counterpart.

Chapitre 1, Tableau 3. Compilation of the *in vitro* cleavage of the HCV (-) strand by SOFA-HDV ribozyme data.

Cleavage positions	Rz P1 sequences	Rz Biosensor sequences	Cleavage %
130	CCCUCU	CAGCCUCCAGG	04 ± 3
146	AGUGGU	CCUCCGGGAGA	15 ± 2
151	GGUCUG	CCCGGAGAGCC	14 ± 3
57	CGGAGU	AGAGCCAUAGUG	04 ± 3
166	UGAGUAU	GUGGUCUGCGGA	24 ± 5
169	GUACACU	GUCUGCGGAAC	17 ± 1
255	GCUAGCU	GCGUGCCCCGC	16 ± 7
293	UACUGGU	GCGAAAGGCCU	85 ± 6
304	AGGGUGU	UUGUGGUACUGC	70 ± 7
317	GUGCCCU	UGAUAGGGUGCU	12 ± 3
325	GGAGGUU	UGCUUGCGAGUG	14 ± 3
327	AGGUU	CUUACGAGUGCC	43 ± 5
334	GUAGACU	GUGCCCCGGGAG	15 ± 5
336	AUGUGU	GUUGGAGGUGU	09 ± 2

The sequences for each SOFA-HDV Rz corresponding to the P1 and biosensor recognition domains are listed. Nucleotide positions are numbered from 3' to 5' in order to facilitate concordance with the data obtained for the HCV (-) strand.

doi:10.1371/journal.pone.0066377.t003

The cleavage of the HCV (–) strand by all of these SOFA-HDV ribozymes was then accessed using both the replicon and the lentivirus systems as described above. Only 3 SOFA-HDV ribozymes out of the 14 reduced the level of luciferase activity of approximately 10 %. Specifically, SOFA-HDV-Rz317[–], -Rz327[–] and -Rz334[–] showed reductions of 10 %, 10 % and 8 %, respectively, all just below the cut-off previously established for considering a ribozyme as being active. In other words, even the ribozymes that possess an outstanding cleavage activity *in vitro* did not efficiently cleave the HCV (–) *in cellulo*. Together, these data support the hypothesis that the HCV (–) strand appears to be inaccessible for cleavage by the SOFA-HDV ribozymes.

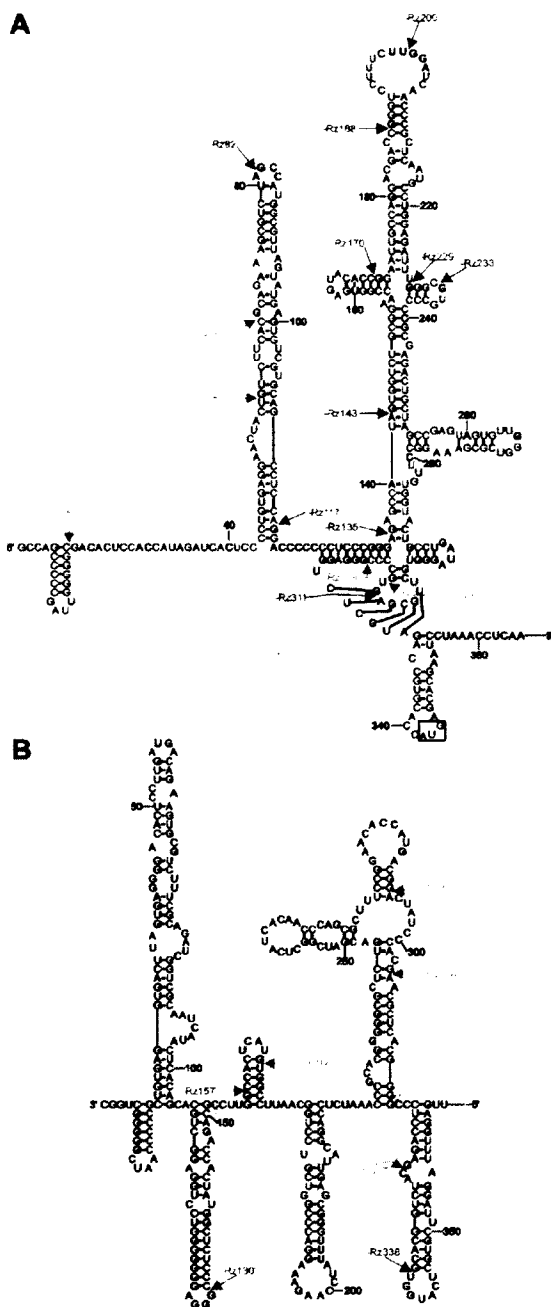
Discussion

Designing SOFA-HDV ribozymes

We report here the first extensive design of SOFA-HDV ribozymes directed against two distinct RNA targets (considering both the (+) and (–) HCV strands as distinct RNA molecules). Previous studies directed towards the development of a gene-inactivation system based on this nucleic acid approach always included a minimal number of ribozymes (i.e. only one to five ribozymes per target [25,28,29]). In the present study, a total of 31 SOFA-HDV ribozymes targeting the HCV strands were designed and their cleavage activities accessed *in vitro*. Seventeen of these SOFA-HDV ribozymes targeted the HCV (+) strand, while fourteen targeted the (–) counterpart. Although not a very large collection of ribozymes, this number should be sufficient to provide some information on the features that must be considered in the design step. In this report the analysis was limited to the data from *in vitro* cleavages because inclusion of the *in cellulo* results implicates too many other important features, including several that remain to be identified.

In order to facilitate the analysis, the cleavage positions of all SOFA-HDV ribozymes tested *in vitro* are illustrated on the proposed secondary structures of both the (+) and (-) HCV strands (Figure 5). Ribozymes that showed either limited or no cleavage activity are in red, while those exhibiting moderate or relatively high levels of cleavage activity are in yellow and green, respectively.

The structure of a target is obviously an important feature to consider when targeting an RNA molecule *in trans*. The formation of double-stranded RNA within the target region can compete with any ribozyme binding that involves unfavourable intramolecular base pairings. This study presents an elegant example of this situation. In the case of the HCV (+) strand, only 7 out of the 17 SOFA-HDV ribozymes exhibited cleavage activities of greater than 10 % (Figure 5A). Six of these ribozymes targeted sites located either side of the highly structured IRES motif. Specifically, 3 SOFA-HDV ribozymes cleaved site before position 68, and 3 after position 299. Conversely, the corresponding region of the HCV (-) strand appears to fold into a less stable secondary structure [46], resulting in 11 out of 14 SOFA-HDV ribozymes that exhibited cleavage activities of over 10 % cleaving in this region (Figure 5B). The higher accessibility of the HCV (-) strand, as compared to the (+) strand, also received support from the discovery of two ribozymes that exhibited outstanding cleavage levels when the length of target is taken into consideration (i.e. SOFA-HDV-Rz293⁻ and -Rz304⁻).



Chapitre 1, Figure 5. Analysis of the cleavage activity of the SOFA-HDV ribozyme collections.

(A) and (B). The cleavage activity levels observed *in vitro* are illustrated on the proposed secondary structures of both the HCV (+) and (-) strands, respectively. The secondary structures shown are as previously reported [13,40]. The SOFA-HDV ribozymes exhibiting either no, or relatively low cleavage activities (< 10%), are indicated in red. Those exhibiting moderate cleavage activities (between 10% and 20%) are indicated in yellow, while those exhibiting relatively high cleavage activities (> 20%) are indicated in green. The arrows indicate the proposed cleavage position of each SOFA-HDV ribozyme.

The previous elucidation of the secondary structure was of great help in designing the ribozymes directed against both strands. In the case of the HCV (+) strand, the use of bioinformatics coupled to validation by the RNase H hydrolysis

was not really more productive, in terms of the number of SOFA-HDV ribozymes that exhibited significant cleavage activity, that was simple analysis of the secondary structure (i.e. 4 and 3 SOFA-HDV ribozymes out of 8 in each case, respectively). However, it is important to remember that the computer-assisted approach coupled to the biochemical validation was performed on the original HDV Rz and only then was the SOFA motif added. It should be noticed that it would be irrelevant to perform RNase H assays using long oligonucleotides including the complementary sequence for the P1, the spacer and the biosensor sequences. In that case the spacers would also be hybridized favoring significantly the formation of the duplexes, therefore introducing an important bias. Alternatively one might imagine using two distinct oligonucleotides at the same time in the RNase H assay (one for the P1 domain and the other for the biosensor domain). In the later case, the resolution of the electrophoresis would not be sufficient to allow distinction between the binding of each oligonucleotide to the target. However, it has previously been reported that the addition of the SOFA motif reduces the importance of the accessibility hurdle [30]. Most likely the presence of the biosensor domain that forms 11-12 base pairs with the substrate contributes significantly, certainly in terms of binding energy. It is easy to imagine that the subsequent binding of the biosensor and the ribozyme's P1 region to the substrate might occur in a cooperative manner. Finally, we noted that the more accessible the complementary sequence of the biosensor (i.e. the more it is located in a single-stranded region), the more efficient the resulting SOFA-HDV ribozyme tended to be. However, this is not a simple question of adding the base pairs involved in the formation of both the P1 and biosensor domains. For that reason it is complex to consider it during the *in silico* step of the SOFA-HDV ribozyme' design. There are many factors that influence the binding of a ribozyme to a target that are not trivial to identify. For example there is all the tertiary structure of the ribozyme by itself that will influence the result, which cannot be taken into account and, therefore, may influence the accuracy of the predictions and contributes to some discrepancies.

Observation of the correlation between the sequence composition of the biosensor and the P1 strand of the ribozyme domain, as well as of their binding sequences within the substrate, with the cleavage efficiency prompts several conclusions. With the exception of sequences rich in adenosine and uridine in either of these two domains, these analyses did not reveal any bias towards higher cleavage levels. Moreover, the presence of the biosensor domain seems to reduce the importance of the composition of the P1 domain. Previously reported analyses of HDV ribozyme cleavage using a collection of small substrates revealed that the identity of the base pairs involving the pairing of substrate position +4 with the ribozyme significantly influenced the activity level [47]. For example, the presence of a guanosine residue in the substrate that base paired with a cytosine in the ribozyme resulted in the formation of an unproductive ribozyme-substrate complex. These nucleotides, located in the middle of the P1 stem, have been proposed to be essential for both substrate binding and the subsequent steps in the cleavage pathway [47,48]. The fact that both the SOFA-HDV-Rz68 targeting the HCV (+) strand, and -Rz169⁻ targeting the HCV (-) strand, each of which possesses a cytosine in this position and exhibits a relatively high level of cleavage activity, suggests that the requirements for the efficient HDV ribozyme targeting of small substrates may differ from those of the SOFA-HDV ribozyme targeting long RNA molecules. Even so, the observed substrate specificity preferences found for positions -1 to -4 adjacent to the cleavage site using a small substrate agree in part with the results observed for the collection of SOFA-HDV ribozymes targeting the HCV strands. Previously, a collection of small substrates possessing both single and multiple mutations in positions -1 to -4 was studied [49]. Some of these substrates were found to be uncleavable, while others demonstrated almost two orders of magnitude of difference, in terms of relative specificity, between the least and the most efficiently cleaved substrates. The nucleotides located at each of these four positions contribute differently to the ability of a substrate to be cleaved [50]. The optimal sequence for positions -4 to -1 was determined to be -4YHRH-1 (where Y indicates C or U; H indicates A, C or U; and R indicates A or G [49]).

Even if this portion of the substrate is not part of the recognition domain (i.e. the base-pairing domain), it has been proposed to play a crucial role as an external determinant of the ability of a substrate to be cleaved in addition to potentially imposing steric hindrances that limit the cleavage activity. Several analyses were performed in order to verify whether or not the cleavage activity of SOFA-HDV ribozymes targeting HCV obeyed the same specificity requirements (raw data not shown). The identities of the bases located in positions -1 and -2 appeared to significantly influence the cleavage level observed. The presence of two consecutive pyrimidines in these positions appears to be detrimental, although to varying degrees. In terms of positions -3 and -4, the nucleotides identities were less important, similar to what is observed for small substrates.

In brief, the cleavage site of a substrate for SOFA-HDV ribozyme should harbor a guanosine in position +1, no guanosine in position -1, should be in agreement with the previously determined preferable nucleotides in positions -1 to -2, and, as much as possible, should be accessible for binding. However, even if *in vitro* experiments suggest the potential for success, there is no guarantee that this will be the case *in vivo*. Moreover, there is some, but not a perfect, correlation between the level of cleavage observed *in vitro* with one detected *in vivo*. For example, SOFA-HDV-Rz60 exhibited the best cleavage activity of the HCV (+) strand *in vitro* but it was not the case *in cellulo* (i.e. 39 % *in vitro* compared to only 5 % *in cellulo*; see Table 2 and Figure 4). Similarly, the best cutter of the HCV (-) strand *in vitro* was not the one that exhibited the higher level of cleavage *in cellulo* (i.e. SOFA-HDV-Rz293⁻ exhibited 85 % of cleavage *in vitro* while it was not inhibiting *in cellulo*). The opposite solution was also observed. For example, the SOFA-HDV-Rz299 exhibited only 13 % of cleavage *in vitro* but was the best one *in cellulo* with a 42 % of reduction. There are several parameters *in cellulo* that may influence significantly the cleavage level and are not occurring *in vitro* or taken into consideration during the design of the SOFA-HDV ribozyme. Some parameters are related to the target (e.g. the presence of interacting proteins either masking

cleavage site or modifying the structure of the substrate) while others are related to the ribozyme by itself (i.e. the turnover and half-life of each ribozymes may differed), to name only these examples. Clearly, this is a complex problem and there is not a simple solution.

The SOFA-HDV ribozymes may be of limited interest for HCV treatment

This report shows that targeting the HCV replicon system using SOFA-HDV ribozymes was not really productive, even when great attention was devoted towards optimizing their design. The *in cellulo* experiments were repeated several times using different conditions, including the use of various lentivirus preparations, transfection with the ribozyme and use of other HCV IRES-mediated translation systems, to name only a few examples. Even though inhibition of HCV replication was observed, it was only at a relatively limited level. The suitable inhibition level required for further development of a therapeutical approach is at least 80 % inhibition. Anything less would likely not lead to viral clearance. Therefore, none of the SOFA-HDV ribozymes examined appeared to be a potential candidate for further development in HCV treatment.

One potential explanation for the limited cleavage activity might be the bicistronic replicon system used to screen the ribozyme *in cellulo*, although it has been used successfully in many studies [38-43]. It was showed that a picornavirus IRES can hand over components of the translation apparatus to another translation start site on the same reporter RNA in cis. [51]. By this translation enhancement such an internal picornavirus IRES reduces possible effects of mutations or inhibitors of translation. However, this study did not verify if the reduction of inhibition can be observed when using cleaved mRNA (i.e. resulting from the action of a ribozyme). Therefore, there is no certitude that the camouflage phenomenon may have masked the action of some of the tested SOFA-HDV ribozymes. Moreover, this possibility cannot be evoked in order to explain the low level of

cleavage activity observed *in cellulo* for the collection of SOFA-HDV ribozymes targeting the (–) strand. In that case it has to occur on the opposite strand that is used for the translation.

Another potential explanation might be intrinsic to the SOFA-HDV ribozymes or the fact that the ribozymes were linked to a tRNA for their expression, which may impair their cleavage ability. However, there is increasing evidence that the HCV viral genome replicates in a peculiar membranous structure located around the endoplasmic reticulum [52-54]. Electron microscopy observations revealed that cells supporting the replication possess an unusual membranous structure, and that part of the endoplasmic reticulum membrane is notably deformed [44]. This structure has a complicated shape, as well as a botryoidal structure in some cases, and is also seen in liver specimen taken from HCV infected patients. This structure was named a membranous web, and is common to Flaviviridae. Both (+) and (–) HCV strands were shown to be synthesized in a membrane-protected complex [44]. Moreover, HCV strands were shown to be resistant to micrococcal nuclease treatment, a treatment that completely hydrolyzed ribosomal RNA [44]. In contrast, HCV strands were fully hydrolyzed by the nuclease treatment only when the experiment was performed in the presence of a detergent which disrupted the membranous structures. Moreover, it has been shown that the core recruits NS proteins, HCV RNAs and the replication complex to lipid droplet-associated membranes. In fact, it has been shown that lipid droplets play a central role in the production of infectious HCV particles [44,45]. This situation considerably reduces the accessibility of both the HCV (+) and (–) strands to the SOFA-HDV ribozyme during replication, in agreement with the limited inhibition observed for all of tested SOFA-HDV ribozymes. Furthermore, this fact also explains why other nucleic acid based approaches directed towards controlling the propagation of HCV, including antisense, ribozymes and deoxyribozymes [19-22], and for which much effort has been devoted, has yielded no commercial research program to date even though they remains active. It has even been shown that the entire HCV (–) strand was

resistant to RNA interference [55]. These studies also support the notion that the cellular localization of the HCV strands significantly limits the possibility of controlling their propagation by targeting the RNA directly, certainly not the (-) strand. In the case of siRNAs designed to target the HCV (+) strands, the best of them achieved ~60 % silencing [11-14]. Contrastly, one study did claim viral clearance based on siRNA targeting [15]. One possible explanation for the latest result may be because siRNA take advantage on the cellular protein machinery, which may have for effect to reduce the hurdle of the HCV RNA strand accessibility. Altogether, this collection of data suggests that much more work is required in order to achieve development of a therapeutic approach directed against HCV, including finding strategies that can release the viral strands from their lipid droplets. At least, nucleic acid-based drugs should be tested with an infection system in which the HCV RNA accessibility problem could be overcome. Specifically, replicon containing cells may already possess membranous alterations due to replication complexes. Alternatively, the use of HCV-JFH1 virus cell culture system [41] may provide an open window of time for targeting the HCV within the early stage of replication. When initiating its infection, the viral RNA should be more exposed to cellular translation system in order to produce viral proteins required for RNA replication. At this point, HCV RNA strands should be more available to our targeting component and the resulting inhibition higher.

Materials and Methods

SOFA-HDV ribozymes and DNA constructs

SOFA-HDV ribozymes were constructed using a PCR-based strategy that included two complementary and overlapping oligonucleotides as described previously [29]. Briefly, two DNA oligonucleotides were used: the universal reverse primer (5'-CCAGCTAGAAAGGGTCCCTTAGCCATCCGCGAACGGATGCCC-3') and the SOFA-HDV ribozyme sense primer (5'-TAATACGACTCACTATAGGGCCAGC

TAGTTT(N)₁₁₋₁₂BS(N)₄BL CAGGGTCCACCTCCTCGCGGT(N)₆P1 TGGGCATCCGTT CGCGG-3'), where N represents A, C, G or T and BS, BL and P1 indicate the biosensor, the blocker sequence and the P1 recognition domain, respectively. The SOFA-HDV ribozyme sense primer is specific for each ribozyme and permits the incorporation of the T7 RNA promoter, while the universal antisense 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). The PCR products were ethanol precipitated, dissolved in water and were then used either for *in vitro* transcription (see below) or in a second PCR reaction (for cloning purposes). The tRNA^{Val}:SOFA-HDV ribozyme chimeric constructs were generated as described previously [26]. Briefly, the first PCR products were used as templates in order to produce a double-stranded DNA sequence containing a KpnI restriction site in 5'. For this second PCR, the universal reverse primer was used in combination with a universal SOFA KpnI forward primer (5'-CGGGGTACCGGGCCAGCTAGTTT-3') in a filling reaction using Pwo DNA polymerase. This PCR product was ethanol-precipitated prior to digestion with KpnI (New England Biolabs). The digested amplicons were ligated to KpnI/EcoRV co-digested ptRNA^{Val} vector derived from pcDNA3.1 (Invitrogen) as described previously [26]. The resulting constructions produce tRNA^{Val}-driven SOFA-HDV ribozymes targeting various regions of either the HCV 5' UTR or its complementary sequence in the (-) strand. In order to construct lentiviral vectors for each ribozyme, the entire tRNA^{Val}:SOFA-HDV ribozyme was amplified by PCR using purified Pwo DNA polymerase and a tRNA-EcoRI forward primer (5'-TATTGAATTCACCGTTGGTTTCCGTAG-3') coupled with a SOFA-Rz-XhoI reverse primer (5'-ATAACTCGAGAAAAAAGATCCAGCTAGAAAGGG-3'). EcoRI-XhoI co-digested PCR products and pLentiV5-U6 (Invitrogen) were ligated together, producing a lentiviral vector that expresses the tRNA^{Val}:SOFA-HDV ribozyme (the co-digestion of pLentiV5-U6 removes the U6 promoter). All constructions were verified by DNA sequencing.

Plasmid pHCVA was constructed by cloning the 1348-nt HCV 5' sequence from pHCV-1b [56] into the HindIII/BamHI sites of pcDNA3 (Invitrogen). Plasmid pHCV(-) was constructed by subcloning the HindIII/KpnI restriction fragment of pHCVA into pcDNA3.1(-) (Invitrogen).

RNA synthesis by *in vitro* transcription

Both the ribozyme and the HCV-derived transcripts were synthesized by run-off transcription as described previously [29]. Briefly, transcriptions were performed in the presence of purified T7 RNA polymerase (10 µg), RNAGuard (24 U, GE Healthcare), pyrophosphatase (0.01 U, Roche Diagnostics) and either KpnI-linearized pHCVA (5 µg), HindIII-linearized pHCV(-) (5 µg) or PCR product (2 to 5 µM) in a buffer containing 80 mM HEPES-KOH, pH 7.5, 24 mM MgCl₂, 2 mM spermidine, 40 mM DTT and 5 mM of each NTP in a final volume of 100 µL at 37°C for 2 h. Upon completion, the reaction mixtures were treated with DNase RQ1 (Promega) at 37°C for 20 min, and the RNA 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 EDTA, pH 8.0, 0.025 % xylene cyanol and 0.025 % bromophenol blue). The samples were then fractionated through either 5 % or 8 % denaturing polyacrylamide gels (PAGE, 19:1 ratio of acrylamide to bisacrylamide) in buffer containing 45 mM Tris-borate, pH 7.5, 8 M urea, and 2 mM EDTA. The reaction products were visualized by ultraviolet (UV) shadowing. The bands corresponding to the correct sizes for both the ribozymes and the HCV-derived RNAs were cut out and the transcripts eluted overnight at room temperature in elution buffer (500 mM ammonium acetate, 10 mM EDTA, 0.1 % SDS). The transcripts were ethanol precipitated, washed, dried and dissolved in ultrapure water. The RNA was quantified by absorbance at 260 nm.

RNA substrate labeling

In order to generate 5'-end-labeled RNA substrate, 50 pmol of purified transcripts were dephosphorylated by adding 1 U of Antarctic phosphatase (New England Biolabs) and incubating for 1 h at 37°C in a final volume of 10 µL containing 50 mM Bis-Propane, pH 6.0, 1 mM MgCl₂, 0.1 mM ZnCl₂ and 40 U of RNAGuard. The enzyme was then inactivated by incubating at 65°C for 8 min. The dephosphorylated RNAs (5 pmol) were then 5'-end labeled by incubation for 1 h at 37°C with 3 U of T4 polynucleotide kinase (USB) and 3.2 pmol of [γ -³²P]ATP (6000 Ci/mmol; New England Nuclear) in the provided reaction buffer. The reactions were stopped by the addition of two volumes of loading buffer. The 5'-end-labeled RNA substrates were fractionated on 5 % PAGE gels and the RNAs detected by autoradiography, cut out and eluted as described above.

Determination of the most accessible sites within the HCV RNA

The selection of potential target sites was performed as previously described [33]. Briefly, the bioinformatic approach includes three steps. First, the RNA Structure 3.7 software was used to both fold the HCV RNA and to predict the most stable secondary structure of the first 341 nucleotides, in terms of energy, of the genotype 1b variant (NCBI database accession number: AJ238799). Secondly, using the OligoWalk software [34], the accessibility for the binding of complementary 7-mer oligonucleotides mimicking the HDV Rz binding domain was assessed *in silico* on the predicted structures. Thirdly, the accessibility of the resulting sites *in vitro* was verified by ribonuclease H probing. The reactions were performed using 5'-end-labeled HCV-derived 575-nt transcripts (~0.1 µM, ~50,000 CPM) in combination with 7-mer oligonucleotides (5 µM) complementary to a specific sequence on the substrate. Each oligonucleotide was used in the presence of the substrate, and was pre-incubated for 10 min at 25°C in a final volume of 8 µL containing 20 mM Tris-HCL, pH 7.5, 20 mM KCl, 10 mM MgCl₂, 0.1 mM EDTA and 0.1 mM DTT.

RNase H (0.5 units, Ambion) was then added (2 μ L) and the samples were incubated at 37°C for 30 min. The reactions were quenched by the addition of one volume of loading buffer, loaded on denaturing 5-10 % PAGE gels and, after electrophoresis exposed to PhosphorImager screens (Storm apparatus, Molecular Dynamics).

***In vitro* ribozyme cleavage assays**

The conditions for SOFA-HDV ribozyme cleavage assays have been described previously [29]. Briefly, 5'-end-labeled 575-nt HCV-derived transcripts of 575 nt (~50,000 CPM) were mixed with SOFA-HDV Rz (10 pmol) in a 20 μ L reaction containing 50 mM Tris-HCl, pH 7.5 and 10 mM MgCl₂, and were then incubated at 37°C for 3 h. The reactions were stopped by the addition of one volume of loading buffer, fractionated on denaturing 5-8 % PAGE gels which were then exposed to PhosphorImager screens.

Lentivirus production, cell culture and in cellulo assays

Lentivirus were produced as described previously [57]. Briefly, 5x10⁶ HEK 293-FT cells were seeded in 100 mm tissue culture dishes. The next day, the media was replaced by Opti-MEM (Gibco-BRL) media prior to transfection of 6 μ g each of pLenti tRNA^{Val}:SOFA Rz, pLP1, pLP2 and pLP/VSVG (Invitrogen) with Lipofectamine 2000 (Invitrogen). After a 4 h incubation, the Opti-MEM was changed for DMEM containing 10 % FBS. At 72 h post-transfection, the viral supernatant was filtered through a 0.45 μ m Millipore filter. Each viral stock was titrated by counting the number of CFU produced by transduction of HT1080 cells with different dilutions of the viral stock.

Huh-7 cells containing the subgenomic bicistronic replicon Luc-ubi-neo-ET (kindly provided by R. Bartenschlager [37]) were cultured in 100 mm cell culture dishes in Dulbecco's Modified Eagle Medium (DMEM, Wisent) supplemented with 10 % FBS (Wisent), 1 mM sodium pyruvate, 2 mM L-glutamine, 250 µg/mL G-418 (Wisent). The cells were passaged twice weekly at a dilution of 1:4-1:8 depending on the rate of cell growth. Prior to transduction, Huh-7 cells containing the bicistronic replicon were seeded in 24 well plates at a confluence of 5×10^4 cells per well. From that point and until the end of the experiment, Huh-7 cells containing the replicon were never in contact with media containing G-418 or antibiotic. The next day, the media was replaced by 250 µL of filtered lentivirus stock (MOI $\gg 1$) containing 6 µg/mL hexadimethrine bromide (polybrene, Sigma Aldrich) and the plates then incubated at 37°C for 48 h.

The dual-luciferase reporter assay system (Promega) was used in order to monitor the firefly luciferase as reported elsewhere [58]. Briefly, the cells were washed with 1X PBS and then were lysed in 150 µL of Passive Lysis Buffer for 20 min at room temperature. The luciferase activity was measured by the addition of 20 µL of lysate to 100 µL of luciferase assay reagent II (Promega) in a 5 mL test tube followed by reading on a Berthold Lumat LB9501 luminometer (Berthold Technologies). For each lysate, the total protein concentration was determined using the Bradford assay (Bio-Rad). Data was expressed as a ratio of RLU per µg of total protein, and was then normalized for each ribozyme by dividing by the result obtained with a control ribozyme (i.e. SOFA-HDV-RzHBV). Both the mean value and the standard deviation were calculated for each ribozyme.

Acknowledgments

The authors would like to thank R. Bartenschlager for kindly providing the bicistronic HCV replicon.

References

1. Moradpour D, Brass V, Gosert R, Wölk B, Blum HE (2002) Hepatitis C: molecular virology and antiviral targets. *Trends Mol Med* 8: 476-482.
2. Chisari FV (2005) Unscrambling hepatitis C virus-host interactions. *Nature* 436: 930-932.
3. Griffiths J, Nix B (2002) Modeling the hepatitis C virus epidemic in France using the temporal pattern of hepatocellular carcinoma deaths. *Hepatology* 35: 709-715.
4. Lindenbach BD, Rice CM (2005) Unravelling hepatitis C virus replication from genome to function. *Nature* 436: 933-938.
5. Spahn CM, Kieft JS, Grassucci RA, Penczek PA, Zhou K, et al. (2001) Hepatitis C virus IRES RNA-induced changes in the conformation of the 40S ribosomal subunit. *Science* 291: 1959-1962.
6. De Francesco R, Migliaccio G (2005) Challenges and successes in developing new therapies for hepatitis C. *Nature* 436: 953-960.
7. Pereira AA, Jacobson IM (2009) New and experimental therapies for HCV. *Nat Rev Gastroenterol Hepatol* 6:403-411.
8. Lamarre D, Anderson PC, Bailey M, Beaulieu P, Bolger G, et al. (2003) An NS3 protease inhibitor with antiviral effects in humans infected with hepatitis C virus. *Nature* 426: 186-189.
9. Beaulieu PL, Tsantrizos YS (2004) Inhibitors of the HCV NS5B polymerase: new hope for the treatment of hepatitis C infections. *Curr Opin Investig Drugs* 5: 838-850.
10. Olsen DB, Eldrup AB, Bartholomew L, Bhat B, Bosserman MR, et al. (2004) A 7-deaza-adenosine analog is a potent and selective inhibitor of hepatitis C virus replication with excellent pharmacokinetic properties. *Antimicrob Agents Chemother* 48: 3944-3953.
11. Krönke J, Kittler R, Buchholz F, Windisch MP, Pietschmann T, et al. (2004) Alternative approaches for efficient inhibition of hepatitis C virus RNA replication by small interfering RNAs. *J Virol* 78: 3436-3446.
12. Kapadia SB, Brideau-Andersen A, Chisari FV (2003) Interference of hepatitis C virus RNA replication by short interfering RNAs. *Proc Natl Acad Sci USA* 100: 2014-2018.
13. Yokota T, Sakamoto N, Enomoto N, Tanabe Y, Miyagishi M, et al. (2003) Inhibition of intracellular hepatitis C virus replication by synthetic and vector-derived small interfering RNAs. *EMBO Rep* 4: 602-608.

14. Wilson JA, Jayasena S, Khvorova A, Sabatino S, Rodrigue-Gervais IG, et al. (2003) RNA interference blocks gene expression and RNA synthesis from hepatitis C replicons propagated in human liver cells. *Proc Natl Acad Sci USA* 100: 2783-2788.
15. Randall G, Grakoui A, Rice CM (2003) Clearance of replicating hepatitis C virus replicon RNAs in cell culture by small interfering RNAs. *Proc Natl Acad Sci USA* 100: 235-240.
16. Marques JT, Williams BR (2005) Activation of the mammalian immune system by siRNAs. *Nat Biotechnol* 23: 1399-1405.
17. Sledz CA, Holko M, de Veer MJ, Silverman RH, Williams BR (2003) Activation of the interferon system by short-interfering RNAs. *Nat Cell Biol* 5: 834-839.
18. Kim DH, Longo M, Han Y, Lundberg P, Cantin E, et al. (2004) Interferon induction by siRNAs and ssRNAs synthesized by phage polymerase. *Nat Biotechnol* 22: 321-325.
19. Welch PJ, Tritz R, Yei S, Leavitt M, Yu M, et al. (1996) A potential therapeutic application of hairpin ribozymes: in vitro and in vivo studies of gene therapy for hepatitis C virus infection. *Gene Ther* 3: 994-1001.
20. Trepanier J, Tanner JE, Momparler RL, Le ON, Alvarez F, et al. (2006) Cleavage of intracellular hepatitis C RNA in the virus core protein coding region by deoxyribozymes. *J Viral Hepat* 13: 131-138.
21. Ohkawa K, Yuki N, Kanazawa Y, Ueda K, Mita E, et al. (1997) Cleavage of viral RNA and inhibition of viral translation by hepatitis C virus RNA-specific hammerhead ribozyme in vitro. *J Hepatol* 27: 78-84.
22. Macejak DG, Jensen KL, Jamison SF, Domenico K, Roberts EC, et al. (2003) Inhibition of hepatitis C virus (HCV)-RNA-dependent translation and replication of a chimeric HCV poliovirus using synthetic stabilized ribozymes. *Hepatology* 31: 769-776.
23. Lévesque D, Choufani S, Perreault JP (2002) Delta ribozyme benefits from a good stability in vitro that becomes outstanding in vivo. *RNA* 8: 464-477.
24. Asif-Ullah M, Lévesque M, Robichaud G, Perreault JP (2007) Development of ribozyme-based gene-inactivations; the example of the hepatitis delta virus ribozyme. *Curr Gene Ther* 7: 205-216.
25. Fiola K, Perreault JP, Cousineau B (2006) Gene targeting in the gram-positive bacterium *Lactococcus lactis*, using various delta ribozymes. *Appl Environ Microbiol* 72: 869-879.
26. D'Anjou F, Bergeron LJ, Larbi NB, Fournier I, Salzet M, et al. (2004) Silencing of SPC2 expression using an engineered delta ribozyme in the mouse betaTC-3 endocrine cell line. *J Biol Chem* 279: 14232-14239.

27. Al-Anouti F, Ananvoranich S (2002) Comparative analysis of antisense RNA, double-stranded RNA, and delta ribozyme-mediated gene regulation in *Toxoplasma gondii*. *Antisense Nucleic Acid Drug Dev* 12: 275-281.
28. Robichaud GA, Perreault JP, Ouellette RJ (2008) Development of an isoform-specific gene suppression system: the study of the human Pax-5B transcriptional element. *Nucleic Acids Res* 36: 4609-4620.
29. Bergeron LJ, Perreault JP (2005) Target-dependent on/off switch increases ribozyme fidelity. *Nucleic Acids Res* 33: 1240-1248.
30. Bergeron LJ, Reymond C, Perreault JP (2005) Functional characterization of the SOFA delta ribozyme. *RNA* 11: 1858-1868.
31. Fong TL, Shindo M, Feinstone SM, Hoofnagle JH, Di Bisceglie AM (1991) Detection of replicative intermediates of hepatitis C viral RNA in liver and serum of patients with chronic hepatitis C. *J Clin Invest* 88: 1058-1060.
32. Quinkert D, Bartenschlager R, Lohmann V (2005) Quantitative analysis of the hepatitis C virus replication complex. *J Virol* 79: 13594-13605.
33. Bergeron LJ, Perreault JP (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.
34. Mathews DH, Burkard ME, Freier SM, Wyatt JR, Turner DH (1999) Predicting oligonucleotide affinity to nucleic acid targets. *RNA* 5: 1458-1469.
35. Peracchi A (2004) Prospects for antiviral ribozymes and deoxyribozymes. *Rev Med Virol* 14: 47-64.
36. Kato Y, Kuwabara T, Warashina M, Toda H, Taira K (2001) Relationships between the activities in vitro and in vivo of various kinds of ribozyme and their intracellular localization in mammalian cells. *J Biol Chem* 276: 15378-15385.
37. Bertrand E, Castanotto D, Zhou C, Carbonnelle C, Lee NS, et al. (1997) The expression cassette determines the functional activity of ribozymes in mammalian cells by controlling their intracellular localization. *RNA* 3: 75-88.
38. Vrolijk JM, Kaul A, Hansen BE, Lohmann V, Haagmans BL, et al. (2003) A replicon-based bioassay for the measurement of interferons in patients with chronic hepatitis C. *J Virol Methods* 110: 201-209.
39. Seo MY, Abrignami S, Houghton M, Han JH (2003) Small interfering RNA-mediated inhibition of hepatitis C virus replication in the human hepatoma cell line Huh-7. *J Virol* 77: 810-812.
40. Kroke J, Kittler R, Buchholz F, Windisch MP, Pietschmann T, et al. (2004) Alternative approaches for efficient inhibition of hepatitis C virus RNA replication by small interfering RNAs. *J Virol* 78: 3436-3446.

41. Sakamoto N, Tanabe Y, Yokota T, Satoh K, Sekine-Osajima Y, et al. (2008) Inhibition of hepatitis C virus infection and expression in vitro and in vivo by recombinant adenovirus expressing short hairpin RNA. *Hepatology* 23: 1437-1447.
42. Watanabe T, Umehara T, Yasui F, Nakagawa SI, Yano J, et al. (2007) Liver target delivery of small interfering RNA to the HCV gene by lactosylated cationic liposome. *J Hepatol* 47: 744-750.
43. Yokota T, Sakamoto N, Enomoto N, Tanabe Y, Miyagishi M, et al. (2003) Inhibition of intracellular hepatitis C virus replication by synthetic and vector-derived small interfering RNAs. *EMBO Reports* 4: 602-608.
44. Egger D, Wölk B, Gosert R, Bianchi L, Blum HE, et al. (2002) Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. *J Virol* 76: 5974-5984.
45. Miyanari Y, Hijikata M, Yamaji M, Hosaka M, Takahashi H, et al. (2003) Hepatitis C virus non-structural proteins in the probable membranous compartment function in viral genome replication. *J Biol Chem* 278: 50301-50308.
46. Smith RM, Walton CM, Wu CH, Wu GY (2002) Secondary structure and hybridization accessibility of hepatitis C virus 3'-terminal sequences. *J Virol* 76: 9563-9574.
47. Ananvoranich S, Lafontaine DA, Perreault JP (1999) Mutational analysis of the antigenomic trans-acting delta ribozyme: the alterations of the middle nucleotides located on the P1 stem. *Nucleic Acids Res* 27: 1473-1479.
48. Ananvoranich S, Perreault JP (1998) Substrate specificity of delta ribozyme cleavage. *J Biol Chem* 273: 13182-13188.
49. Deschênes P, Lafontaine DA, Charland S, Perreault JP (2000) Nucleotides -1 to -4 of hepatitis delta ribozyme substrate increase the specificity of ribozyme cleavage. *Antisense Nucleic Acid Drug Dev* 10: 53-61.
50. Bergeron LJ, Ouellet J, Perreault JP (2003) Ribozyme-based gene-inactivation systems require a fine comprehension of their substrate specificities: the case of delta ribozyme. *Curr Med Chem* 10: 2589-2597.
51. Jünemann C, Song Y, Bassili G, Goergen D, Henke J, et al (2007) Picornavirus internal ribosome entry site elements can stimulate translation of upstream genes *J Biol Chem* 282: 132-141.
52. Ogawa K, Hishiki T, Shimizu Y, Funami K, Sugiyama K, et al. (2009) Hepatitis C virus utilizes lipid droplet for production of infectious virus. *Proc Jpn Acad Ser B Phys Biol Sci* 85: 217-228.
53. McLauchlan J (2009) Hepatitis C virus: viral proteins on the move. *Biochem Soc Trans* 37: 986-990.

54. Miller S, Krijnse-Locker J (2008) Modification of intracellular membrane structures for virus replication. *Nat Rev Microbiol* 6: 363-374.
55. Smith RM, Smolic R, Volarevic M, Wu GY (2007) Positional effects and strand preference of RNA interference against hepatitis C virus target sequences. *J Viral Hepat* 14: 194-212.
56. Hicham Alaoui-Ismaili M, Gervais C, Brunette S, Gouin G, Hamel M, et al. (2000) A novel high throughput screening assay for HCV NS3 helicase activity. *Antiviral Res* 46: 181-193.
57. Boudreau F, Lussier CR, Mongrain S, Darsigny M, Drouin JL, et al. (2007) Loss of cathepsin L activity promotes claudin-1 overexpression and intestinal neoplasia. *FASEB J* 21: 3853-3865.
58. Liu Q, Fu H, Sun F, Zhang H, Tie Y, et al. (2008) miR-16 family induces cell cycle arrest by regulating multiple cell cycle genes. *Nucleic Acids Res* 36: 5391-5404.

Chapitre 2 : Ciblage du VIH avec des ribozymes SOFA-HDV

ARTICLE : A Conserved Target Site in HIV-1 Gag RNA is Accessible to Both HDV Ribozyme and Short Hairpin RNA Activities

Robert J. Scarborough, **Michel V. Lévesque**, Etienne Boudrias-Dalle, Ian C. Chute, Sylvanne M. Daniels, Rodney J. Ouellette, Jean-Pierre Perreault, and Anne Gagnol

Article soumis à : *Molecular Therapy* (novembre 2013)

Avant-propos :

J'ai réalisé une partie des travaux présentés dans l'article. Concrètement, j'ai effectué les essais de coupure *in vitro* et l'analyse de l'expression des ribozymes SOFA-HDV par buvardage Northern. J'ai conçu la majeure partie de l'expérience de puce à ADN «*microarray*» et j'ai effectué l'analyse de ces résultats. J'ai rédigé les parties du manuscrit concernant les expériences que j'ai réalisées et conçues. De plus, j'ai produit les figures (ou panneaux de figures) correspondant à ces résultats et celles des ribozymes. Finalement, j'ai participé activement à la révision de l'ensemble du manuscrit.

Résumé

Les ARN ayant une activité antisens ont le potentiel d'être utilisés comme élément d'une thérapie génique contre le VIH-1 ou comme une alternative aux thérapies reposant sur des drogues s'appuyant sur de petites molécules. Dans cette étude, l'ARN du VIH-1 a été criblé afin d'identifier des sites cibles hautement conservés et accessibles pour des ribozymes basés sur le motif du virus de l'hépatite delta. En utilisant un criblage quantitatif de l'effet sur la production de VIH-1, nous avons identifié un site cible hautement conservé dans la séquence codant pour la protéine Gag, qui était particulièrement accessible pour l'activité catalytique du ribozyme. Nous avons aussi démontré que ce site est hautement accessible à l'interférence par l'ARN dirigé via un shARN, suggérant qu'il est possiblement accessible pour la liaison d'autres ARN antisens utilisant différents modes d'action. Afin d'explorer le potentiel clinique de ce site cible, nous fournissons des évidences qu'il est structurellement conservé dans diverse souches virales et qu'il est suffisamment différent du transcriptome humain pour limiter les effets hors-cibles des thérapies antisens. Il est aussi montré qu'un ribozyme du virus de l'hépatite delta muni d'un module de spécificité SOFA (specific on/off adaptor) est plus spécifique en comparaison à un shARN pour une mutation dans le site cible. Finalement, nous confirmons que le ribozyme et le shARN peuvent inhiber la réplication du VIH-1 lorsqu'exprimés dans une lignée cellulaire de lymphocytes T.

Abstract

Antisense-based RNAs have the potential to be used as part of HIV-1 gene therapy or as an alternative to small-molecule drug therapies. In this study, HIV-1 RNA was screened to identify more conserved and accessible target sites for ribozymes based on the hepatitis delta virus motif. Using a quantitative screen for effects on HIV-1 production we identified a highly conserved target site in the Gag coding sequence that was uniquely accessible to ribozyme catalytic activity. We also demonstrate that this target site is highly accessible to short hairpin directed RNA interference, suggesting that it may be available for the binding of antisense RNAs with different modes of action. To explore the clinical potential of this target site, we provide evidence that it is structurally conserved in diverse viral strains and that it is sufficiently different from the human transcriptome to limit off target effects from antisense therapies. It is also shown that a specific on/off adaptor hepatitis delta virus ribozyme is more specific compared to a short hairpin RNA for a mutation in its target site. Finally we confirm that both our Rz and shRNA can inhibit HIV replication when expressed in a T lymphocyte cell line.

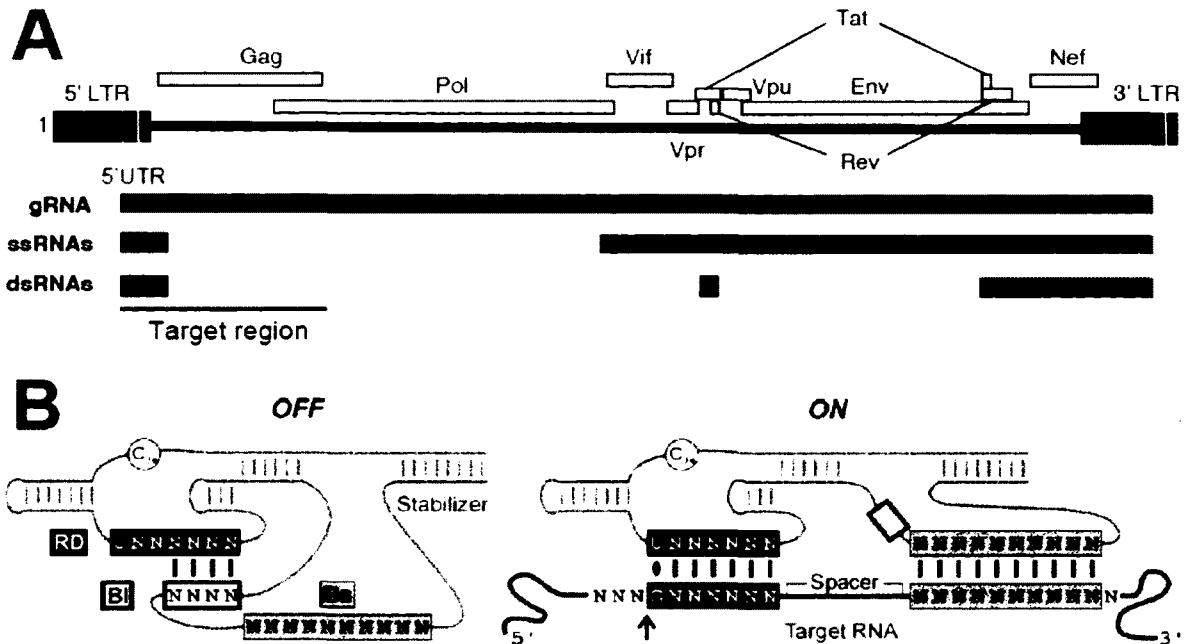
Introduction

Over 30 small molecules are available for the treatment of Human Immunodeficiency Virus 1 (HIV-1) infection, targeting the viral enzymes reverse transcriptase (RT), protease and integrase, as well as the cellular entry co-receptor, CCR5¹. Although treatment of HIV-1 with combination small molecule therapy is effective in preventing acquired immune deficiency syndrome, it is not able to eradicate the virus and is associated with a number of short- and long-term side effects². Alternative therapeutic strategies for long-term viral suppression with low adverse effects are needed and small RNAs represent a growing class of molecules with the potential to complement or replace current therapies. They are being evaluated for use in *ex vivo* gene therapy³ and with advances that have been made in their systemic delivery⁴, may soon be evaluated for use in combination drug therapy. Many small RNAs, including antisense oligonucleotides (ASOs), ribozymes (Rzs), decoys, aptamers, small nuclear (sn) RNAs, and small interfering (si) or short hairpin (sh) RNAs have been designed with diverse target sites in the HIV-1 replication cycle⁵. Antisense-based RNAs (ASOs, Rzs, snRNAs, sh/si RNAs) have a unique ability to target HIV-1 RNA and can be designed to inhibit viral production with high specificity. As the HIV-1 RNA genome is highly structured⁶, a limited number of target sites are easily accessible to antisense molecules, and the structure has been shown to severely restrict the activity of some molecules against the same target site in different HIV-1 variants⁷. The development of new antisense-based RNAs and the identification of conserved antisense-accessible regions in HIV-1 RNA will help add to the currently limited number of viral targets exploited for therapy.

The pandemic group of HIV-1 (Group M) has diverged significantly into multiple genetic subtypes or clades, each comprising several different variants⁸. An important consideration for the design of therapies targeting HIV-1 is that they will be active against the majority of these variants. Detailed sequence conservation data has been provided in two genome wide screens for active siRNAs and

shRNAs^{9, 10}, and in a screening of the terminal exon of HIV-1 RNA for the identification of modified U1 snRNAs¹¹. Although several active Rzs have been identified, a similar approach to screen the HIV-1 genome for conserved Rz target sites has not yet been described in detail. Instead, most Rzs evaluated have targeted sequences in the Tat/Rev exon 1 open reading frame (ORF) and the highly conserved 5' untranslated region (UTR)¹²⁻²¹, present in all splice variants of HIV-1 RNA (**Fig. 1a**). Additional target sites in the Pol^{17, 18, 22} and Nef¹⁸ ORFs have also been identified, however a large proportion of potential Rz target sites remain untested.

Rzs with activity against HIV-1 production have been made by modifying hammerhead, hairpin or bacterial RNase P motifs^{12, 15-17, 19, 20}. The HDV Rz represents an alternative small Rz motif, that has evolved to function in human cells and has the potential to be used for the development of more active and stable therapeutic Rzs²³. To overcome the limitation in specificity of HDV Rzs, the SOFA (Specific On/oFf Adaptor) module was engineered^{24, 25} (**Fig. 1b**). Several SOFA-HDV-Rzs have been identified with activity against human^{26, 27}, viral^{24, 28, 29} and bacterial³⁰ RNAs, including three Rzs that we have evaluated targeting the Tat/Rev ORF of HIV-1 RNA²¹. As with several other studies^{14-16, 19} we were unable to demonstrate a strong catalytic activity of our Rzs in different HIV-1 production assays, and we set out to screen additional target sites in other parts of HIV-1 RNA (**Fig. 1a**) to identify candidates with improved Rz activity. One SOFA-HDV-Rz targeting a novel site in the Gag ORF with catalytic activity against HIV-1 production was identified, and an shRNA targeting the same site was shown to be an extremely potent and active inhibitor of viral production. We provide evidence that the identified target site is highly conserved and specific for HIV-1 RNA, and our results suggest that it is an excellent target site for the development of antisense molecules with therapeutic potential.



Chapitre 2, Figure 1. Schematic representation of HIV-1 RNA and the SOFA-HDV-Rz.

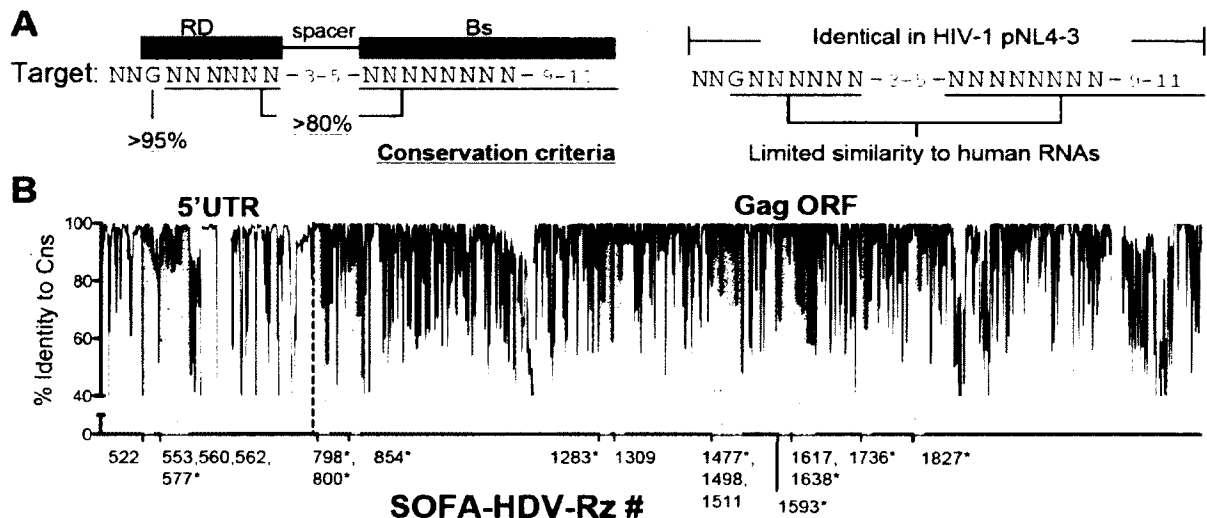
a) The full length genomic (g), singly-spliced (ss) and doubly-spliced (ds) RNA species of HIV-1 are illustrated. Reading frames for all HIV-1 proteins are shown above the different RNAs and the 5' region, used to identify SOFA-HDV-Rz target sites is underlined. b) The SOFA-HDV-Rz is illustrated in both its *OFF* and *ON* conformations. In the *OFF* conformation, the SOFA blocker (BI) base pairs with the last 4 nts of the recognition domain (RD). When the SOFA biosensor (Bs) base pairs with a specific target sequence, the RD is released from the BI sequence and binds at 3 to 5 nts upstream from the Bs binding site in the *ON* conformation. The first nt in the target site (n+1) must be a G, forming a wobble base pair with the RD U. The cleavage site is indicated with an arrow and the nt C₇₆, which can be mutated to disable the catalytic activity of the SOFA-HDV-Rz, is shown as a circle in the Rz backbone.

Results

Identification of SOFA-HDV-Rz target sites in HIV-1 RNA

HIV-1 sequence conservation was analyzed to identify target sites that are relevant for the majority of HIV-1 strains. Conservation estimates have previously been reported in 19 to 21 nucleotide (nt) frames to identify siRNAs or shRNAs^{9, 10}, and at

the single nt level to identify or characterize Rz³¹, snRNA¹¹ and shRNA³² target sites. Estimates at the nt level (Supplementary File S1) were used to identify highly conserved SOFA-HDV-Rz target sites that were identical in HIV-1 strain pNL4-3 and for which the corresponding Rz had a low potential to target human RNAs³³ (**Fig. 2a**). Consistent with previous studies, several highly conserved target sites were identified in the 5'LTR U5 region within the 5'UTR (**Fig. 2b**). The Gag ORF had much lower overall conservation, however we were able to identify four highly conserved and, with some exceptions to our conservation criteria, nine moderately conserved target sites in this region (**Fig. 2b**). Of the Rzs that we had previously evaluated targeting the Tat/Rev exon1 ORF of HIV-1 RNA²¹, both Tat1 and Tev1 were moderately conserved. Details on the target sequences used in this study are illustrated in Table S1.



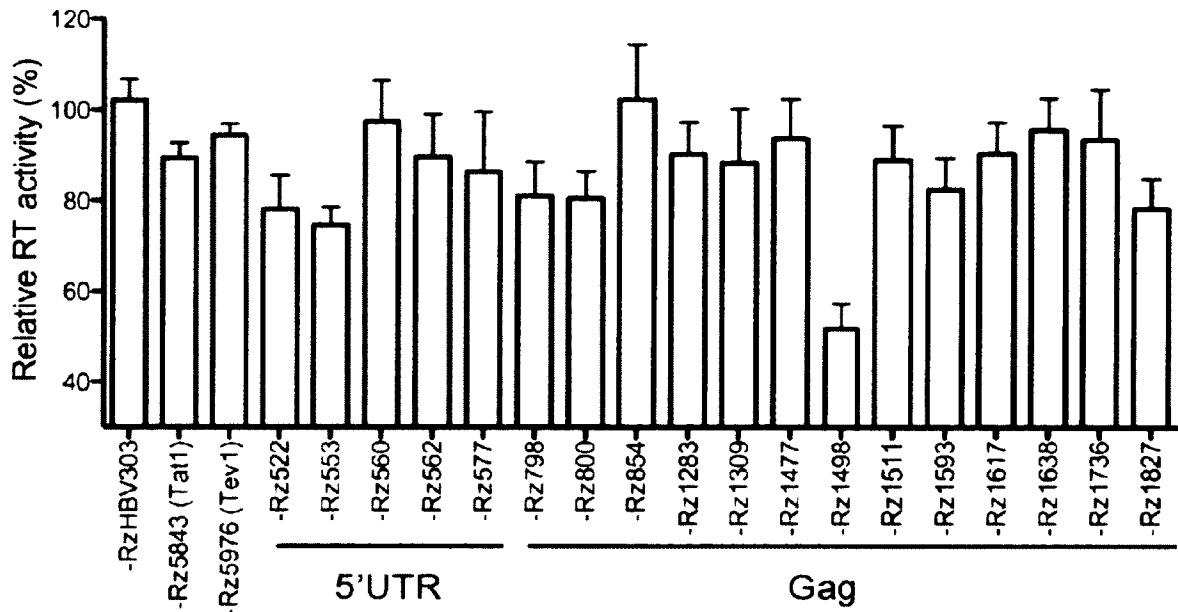
Chapitre 2, Figure 2. Target site identification.

a) Left, conservation criteria used to identify SOFA-HDV-Rz target sites in HIV-1 RNA are illustrated. The number of nts between the RD and the Bs (spacer, 3-5), and or the length of the Bs (9-11) were adjusted to avoid poorly conserved positions or to improve the specificity of the corresponding Rz. A spacer of 4 nt and Bs length of 10 nt were used as the default positioning. Right, target sites were excluded if they were not identical in HIV-1 pNL4-3, or if they had a score of less than 20 for any human RNA sequence using Ribosubstrates software³³. b) Sequence conservation for the consensus (Cns) sequence in the 5' region of HIV-1

RNA is shown with the identified Rz binding sites shaded in grey. SOFA-HDV-Rzs were named according to the first nt in their HIV-1 pNL4-3 binding site. The dashed line represents the separation between the 5'UTR and Gag ORF. SOFA-HDV-Rz target sites that did not meet our conservation criteria, but were moderately conserved are indicated with a *.

SOFA-HDV Rz screen for activity against HIV-1 production

The activity of each SOFA-HDV-Rz was then evaluated by co-transfection with HIV-1 strain pNL4-3 in HEK293T cells, using conditions similar to those reported for other Rzs^{17, 19, 20}. An irrelevant Rz targeting Hepatitis B Virus RNA (SOFA-HDV-RzHBV, adapted from SOFA- δ Rz-303²⁴) was used as a negative control and previously identified SOFA-HDV-Rzs Tat1 and Tev1²¹ were used as positive controls. Of all of the Rzs tested only one Rz (SOFA-HDV-Rz-1498) demonstrated a significant improvement compared to SOFA-HDV-Rz Tat1 and Tev1 (**Fig. 3**), suggesting that this target site may be especially accessible to SOFA-HDV Rz catalytic activity.



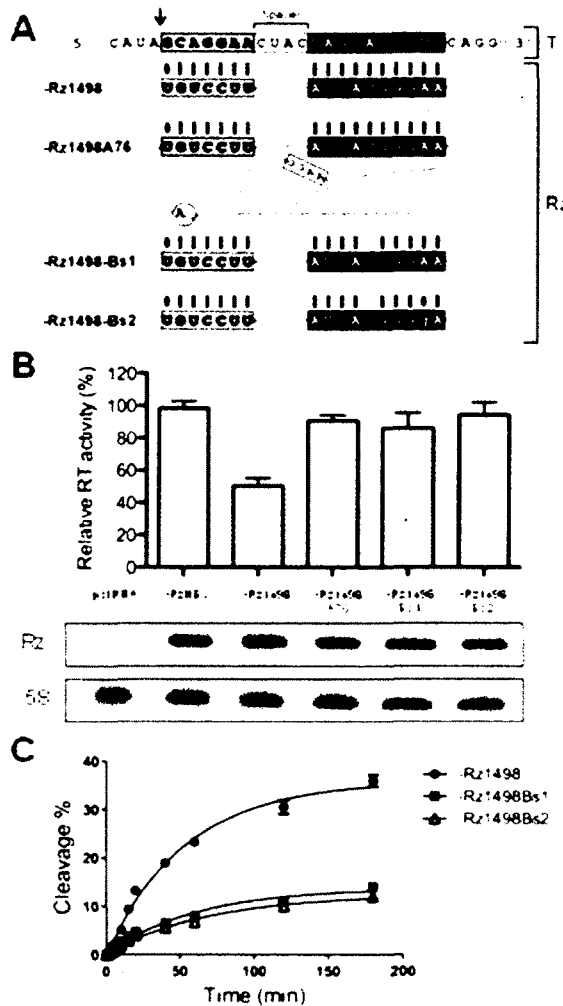
Chapitre 2, Figure 3. Screen for SOFA-HDV-Rz activity against HIV-1 production.

HEK293T cells were co-transfected with HIV-1 pNL4-3 plasmid DNA (75 ng) and one of the indicated psiRNA SOFA-HDV-Rz expression plasmids (750 ng). Viral production was estimated 48 h following transfection by measuring the activity of HIV-1 RT in culture supernatants. Data are expressed as a fraction of RT activity in the empty vector (psiRNA) co-transfected cells and represent the mean \pm SEM (n=5-10).

SOFA-HDV-Rz1498 is catalytically active and specific for its target site

Several Rzs have been shown to exhibit significant antisense activity with only limited catalytic activity in co-transfection assays^{14-16, 19}. To determine whether or not the activity of SOFA-HDV-Rz1498 was related to its catalytic activity, we generated an inactive mutant (SOFA-HDV-Rz1498A76), in which a C to A mutation at position 76 in its backbone disables its cleaving capability without affecting its binding to its target³⁴ (**Fig. 4a**). SOFA-HDV-Rz1498A76 provided only a very weak inhibition of HIV-1 production at similar expression levels to SOFA-HDV-Rz1498 (**Fig. 4b**), suggesting that the catalytic activity, rather than the antisense activity, is primarily responsible for the observed effects of SOFA-HDV-Rz1498. Variants with

either a single or double mutation in their biosensor sequence were also generated to evaluate the specificity of SOFA-HDV-Rz1498 (SOFA-HDV-Rz1498Bs1 and SOFA-HDV-Rz1498Bs2, **Fig. 4a**). Neither variant was able to provide a strong inhibition of HIV-1 production (**Fig. 4b**), demonstrating that the activity of SOFA-HDV-Rz1498 is also highly specific for its target site. These mutants were also less active in a single turnover *in vitro* cleavage assay (**Fig. 4c**), suggesting that the reduced activity in cells is mainly due to a reduced cleavage capacity.



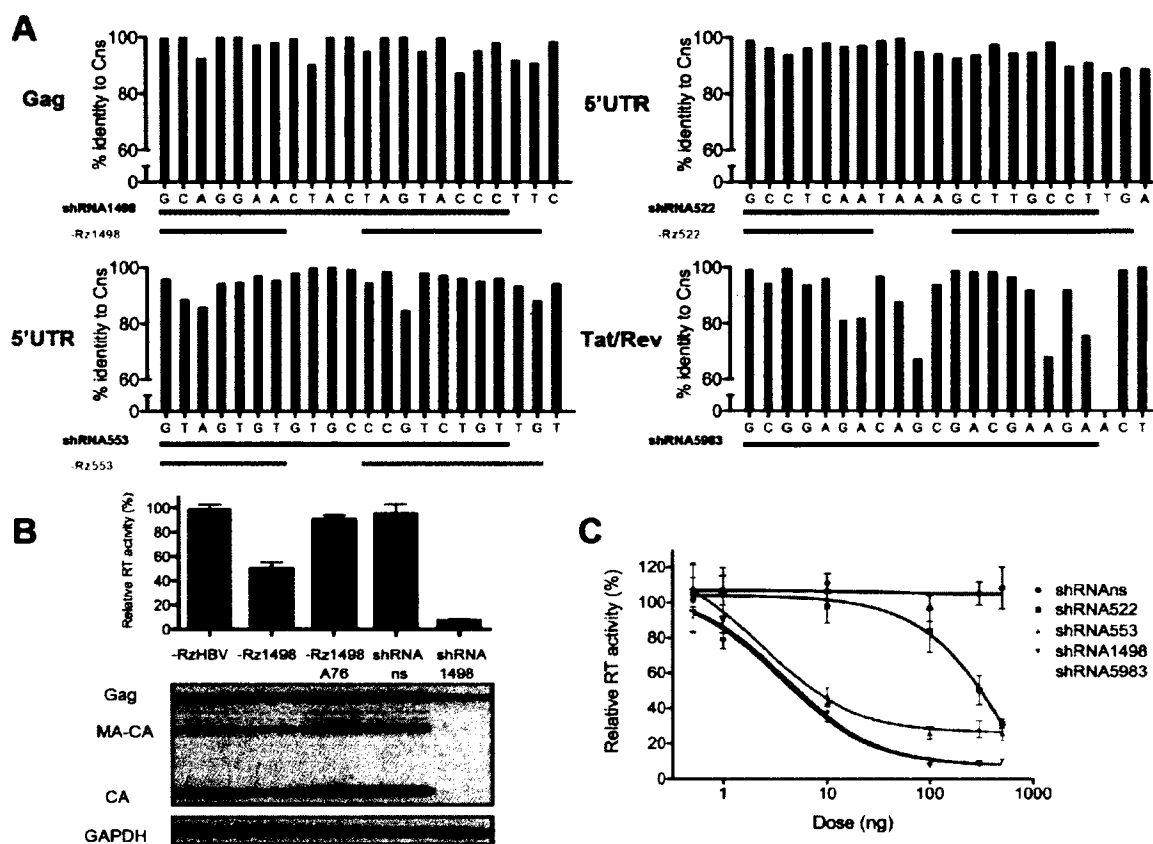
Chapitre 2, Figure 4. Catalytic activity and specificity of SOFA-HDV-Rz1498.

a) Schematic representation of the SOFA-HDV-Rz1498 target site (T) and variants (Rz). SOFA-HDV-Rz1498A76 has a C to A mutation in the Rz backbone, -Rz1498Bs1 and -Rz1498Bs2 have 1 or 2 nt variants in the biosensor (Bs), indicated in lower case. b) Effects on viral production for each Rz were determined as in Figure 3 (n=6-10), the relative expression of Rz and 5S RNA loading control for the different conditions are shown below for one representative set of transfections. c) Single turnover *in vitro* cleavage activities for SOFA-HDV-Rz1498, -Rz1498Bs1 and -Rz1498Bs2 were determined with a small substrate RNA as described previously³⁴.

shRNA1498 is a potent and highly active inhibitor of HIV-1 production

The activity of both a Rz and an shRNA targeting the same site in HIV-1 RNA has been previously evaluated for a target site in the pol ORF. In that case, only the Rz exhibited significant activity in reducing viral replication²². To determine whether the SOFA-HDV-Rz1498 target site was also accessible to shRNA activity, we designed shRNA1498, with > 80% sequence conservation for each nt in its target site (Fig. 5a, Gag). Compared to a nonsense shRNA (shRNAs) as well as SOFA-

HDV-Rz1498 and its controls, shRNA1498 was able to completely inhibit viral production and intracellular expression of the p24 capsid protein (**Fig. 5b**). To compare the potency and activity of shRNA1498 to previously characterized shRNAs, we used shRNA522 and shRNA553¹⁰, targeting two of the most highly conserved target sites in the HIV-1 genome (**Fig. 5a**, 5'UTR), and shRNA5980, representing a construct in clinical development³⁵ targeting a moderately conserved region of the Tat/Rev exon1 ORF (**Fig. 5a**, Tat/Rev). At the highest level of expression, all four shRNAs were able to inhibit HIV-1 production with shRNA1498 and shRNA5980 providing a greater inhibition compared to shRNA522 and shRNA553 (**Fig. 5c**). By using a range of expression levels it was found that shRNA1498, shRNA5980 and shRNA553 were all extremely potent with IC50 values below 5 ng, while shRNA522 was much less potent with an IC50 value of 702 ng. These results demonstrate that, in addition to its accessibility to a SOFA-HDV-Rz, the 1498 target site in the Gag ORF is also accessible to an shRNA, with the same potency and activity as the previously described highly active shRNA5980, and improved target site conservation.



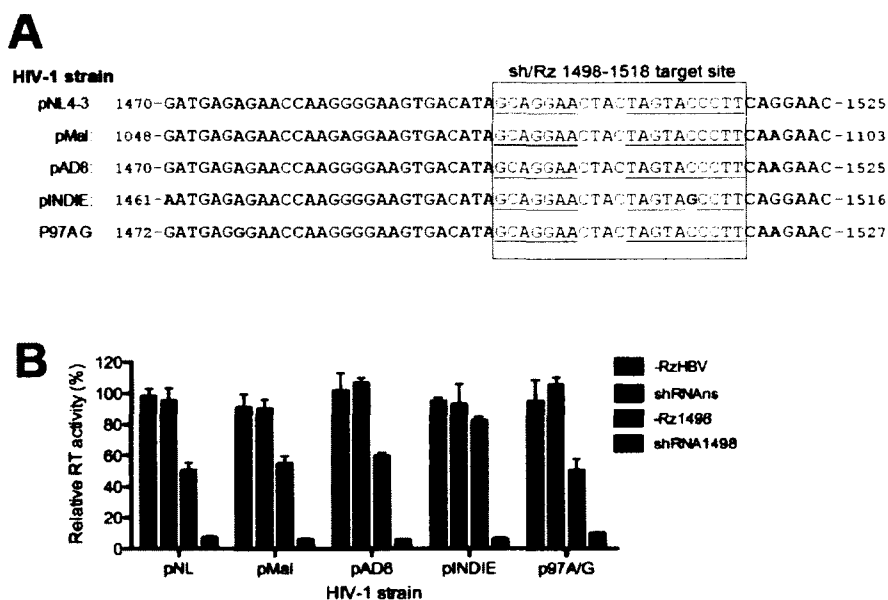
Chapitre 2, Figure 5. Activity of shRNA1498 against HIV-1 production.

a) The consensus (Cns) sequence of SOFA-HDV-Rz and shRNA target sites in the Gag, 5'UTR and Tat/Rev ORFs are shown in relation to their conservation at the nt level. b) The activity of shRNA1498 and shRNAs was evaluated as in Fig. 3 (n=6), and is shown in comparison to the values obtained for SOFA-HDV-Rz1498, -Rz1498A76 and -RzHBV. The relative expression of HIV-1 Gag polyprotein (GAG, p55), matrix-capsid intermediate (MA-CA, p39), and capsid (CA, p24) proteins as well as GAPDH loading control are shown for one representative transfection. c) HEK293T cells were co-transfected with HIV-1 pNL4-3 plasmid DNA (100 ng) and one of the indicated shRNA expression plasmids (1-500 ng). Viral production was determined as in Fig. 3 (n=4-8).

SOFA-HDV-Rz1498 and shRNA1498 are active against diverse viral strains

As the 1498 target site was shown to be accessible to both Rz and shRNA activity in HIV-1 strain pNL4-3, we next evaluated whether or not their activity was

conserved in diverse viral strains representing clade B (pMal³⁶ and pAD8³⁷), clade C (pINDIE³⁸) and clade A/G (p97AG³⁹). Both SOFA-HDV-Rz1498 and shRNA1498 were able to inhibit HIV-1 production from diverse viral strains (pMal, pAD8, p97AG) with nt variants in proximity to their target sites compared to pNL4-3 (Fig. 6), suggesting that the structure of the target site is equally accessible in these strains. Consistent with the effects of mutations in the SOFA-HDV-Rz Bs binding site (Fig. 4), SOFA-HDV-Rz1498 was weakly active against the Clade C strain (pINDIE) with a single nt variant within its Bs binding site (Fig. 6), confirming its high specificity. In contrast, shRNA1498 was active against this strain, suggesting that the structure of the target site is also conserved in the pINDIE strain and that shRNA1498 can tolerate a single nt mismatch in its binding site.



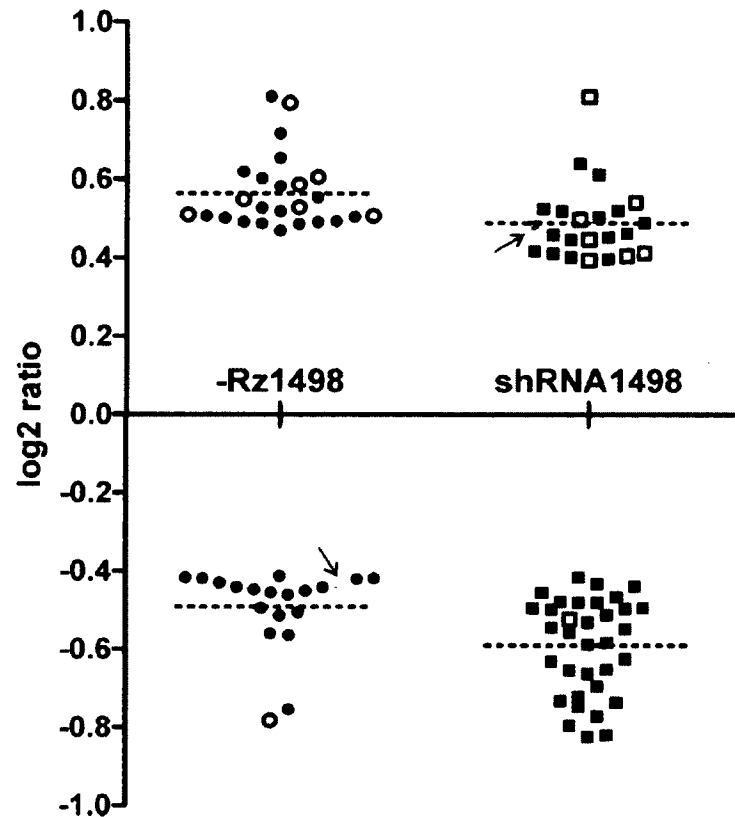
Chapitre 2, Figure 6. Activity of SOFA-HDV-Rz1498 and shRNA1498 against HIV-1 production from diverse viral strains.

a) The sequence in and around the shRNA1498 and SOFA-HDV-Rz1498 target site is shown for HIV-1 pNL4-3 (M19921), pMal (K03456), pAD8 (AF004394), pINDIE (AB023804.1) and p97GH-AG1 (AB049811.1). The overlapping target site for shRNA1498 and SOFA-HDV-Rz1498 is highlighted in blue, with both the RD (7nt) and Bs (10nt) binding sites underlined. Nt variations compared to the consensus sequence are highlighted in red and the start and end positions for

each sequence are shown according to their annotation in Genbank. **b)** The activity of the indicated Rz and shRNA expression plasmids against the different HIV-1 strains was determined as in Fig. 3 (n=4-10).

SOFA-HDV-Rz1498 and shRNA1498 have minimal off target effects on human RNAs

An important consideration for the design of antisense therapies is that they will be specific for their target site. To evaluate the specificity of both SOFA-HDV-Rz1498 and shRNA1498 in human cells, the relative expression levels of cellular mRNAs in HEK293T cells cotransfected with HIV-1 pNL4-3 were measured using human gene expression microarrays. Prior to the gene expression profiling, the inhibition of viral production was confirmed for each condition (data not shown) and agreed with results presented in **Fig. 5b**. All microarray experiments were performed as triplicate dye swaps and the results were expressed as the log₂ ratio of SOFA-HDV-Rz1498 or shRNA1498 compared to the empty vector cotransfected cells (Supplementary File S2). Only a small set of genes showed consistent variation in the presence of SOFA-HDV-Rz1498 or shRNA1498 and the extent of the changes for all of them were limited (**Fig. 7**). As none of the genes that were down regulated included sequences with similarities to the 1498 target site (data not shown), the observed changes are likely unrelated to the specific activities of either SOFA-HDV-Rz1498 or shRNA1498. These results suggest that both molecules are able to specifically target HIV-1 RNA with minimal off-target effects on human genes.



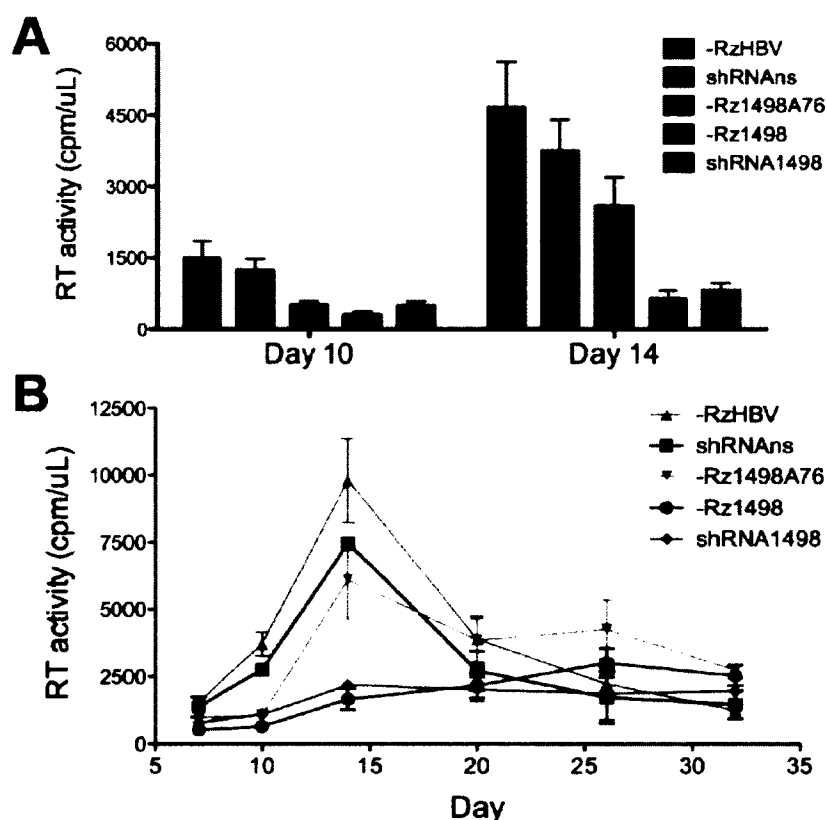
Chapitre 2, Figure 7. Specificity of SOFA-HDV-Rz1498 and shRNA1498.

The log₂ ratio of the most significantly up and down regulated genes are shown for SOFA-HDV-Rz1498 and shRNA1498 in relation to the empty vector transfected cells. Identical genes that were up- or down-regulated in both conditions are outlined in a circle or square respectively. One gene that was down-regulated by SOFA-HDV-Rz1498 and up-regulated by shRNA1498 is indicated with an arrow. The gene identities and log₂ ratio values are provided in Supplementary File S2.

SOFA-HDV-Rz1498 and shRNA1498 inhibit HIV-1 replication in a T cell line

To evaluate the potential for SOFA-HDV-Rz1498 and shRNA1498 to inhibit HIV-1 replication, we constructed Jurkat T cell lines stably expressing the active molecules or their controls. All cell lines had a similar distribution of GFP expression from the integrated plasmids and proliferated at similar levels (Figure S1). Following infection with HIV-1 pNL4-3, both SOFA-HDV-Rz1498 and

shRNA1498 expressing cells were able to efficiently suppress viral replication compared to cells expressing SOFA-HDV-RzHBV and shRNAs (Fig. 8). A significant suppression of viral replication was also observed in cells expressing the catalytically inactive SOFA-HDV-Rz1498A76. These results confirm that the 1498 target site is susceptible to Rz and shRNA activity both in HEK293T cells and in a T lymphocyte cell line.



Chapitre 2, Figure 8. Activity of SOFA-HDV-Rz1498 and shRNA1498 against HIV-1 replication.

a) Jurkat T cells expressing the indicated SOFA-HDV-Rzs and shRNAs were infected with HIV-1 pNL4-3. The average RT activity (cpm/ μ L) across four independent infections performed in triplicate (n=12) is shown for days 10 and 14 following infection. b) Time course of a representative infection (n=3) followed out to 32 days post infection.

Discussion

Rzs specifically targeting HIV-1 RNA were among the first gene therapy agents tested in a clinical setting⁴⁰, and the only agent tested in a phase II gene therapy trial¹³. While the safety of Rz therapy has been demonstrated, the moderate activity observed in clinical trials emphasize the need to find more active molecules for use in combination gene therapy. Approaches that have been described to identify new Rz candidates include the use of: i) RNA Polymerase III (Pol III) promoters to achieve higher levels of expression^{14, 15}, ii) *in vivo* screens to identify accessible target sites^{17, 22}, iii) chimeric Rzs to increase activity^{12, 16, 17, 19}, and iv) alternative motifs such as modified RNase P and HDV Rzs^{20, 21}. In this study we have used the RNA Pol III H1 promoter to screen SOFA-HDV-Rzs targeting the 5'UTR and Gag ORF of HIV-1 RNA for activity against viral production. Several candidate Rzs that had the potential to target the majority of HIV-1 strains were identified (**Fig. 2**) and one of these (SOFA-HDV-Rz1498) had significant catalytic activity against viral production (**Fig. 4**). The effects of both SOFA-HDV-Rz1498 and its inactive control on viral replication (**Fig. 8**) are consistent with those of other Rzs evaluated under similar conditions^{12, 17, 18, 22, 41}. As there is great variability in the conditions used to evaluate HIV-1 replication, it is difficult to compare the activity of SOFA-HDV-Rz1498 to Rzs evaluated in other studies. In the more quantitative co-transfection assay, the approximately 50% reduction in viral production observed for SOFA-HDV-Rz1498 (**Figs. 3 and 4**) is superior to that of HH and Hp Rzs expressed alone from Pol III promoters^{14, 15}, however it is less than the effects observed for recently described RNaseP²⁰ and chimeric HH or Hp^{17, 19} Rzs. While improvements in the design or expression of SOFA-HDV-Rz1498 may be required to compete with other Rzs targeting HIV-1 RNA, the results from our screen (**Fig. 3**) suggest that the target site we identified in the Gag ORF may be particularly accessible to Rz catalytic activity. This is consistent with results from other Rz screens that identified only a limited number of active molecules from a large group

of potential candidates^{17, 22, 42} and show how important target site accessibility is to Rz activity.

According to the secondary structure published for the full length genomic RNA of HIV-1 pNL4-3⁶, the target site we identified in the Gag ORF lies predominantly in a 27 nt single-stranded loop enumerated 1033 to 1059 based on its position in pNL4-3 HIV-1 RNA (1487 to 1513 in pNL4-3 DNA). This target site was highly conserved at the nt level for both a SOFA-HDV-Rz and an shRNA (**Fig. 5a**, Gag), and we show that it is highly accessible to both Rz (**Fig. 3**) and shRNA (**Fig. 5c**) activity. Based on previous results showing that Pol III promoted Rzs accumulate in the nucleus¹⁴, our results suggest that the target site is accessible both in the nucleus and in the cytoplasm, where shRNA activity takes place. Furthermore, our results in diverse strains (**Fig. 6**) suggest that its accessibility may also be conserved in both of these compartments.

In addition to being active against the majority of HIV-1 circulating strains, an important consideration for the design of HIV-1 therapies is that they will be well tolerated. This is particularly important for gene therapy, as the proposed procedures involve the permanent expression of candidate molecules in all cells derived from hematopoietic stem cells³. Due to their potential to saturate the RNAi machinery⁴³, an important safety consideration for shRNA therapy is that they will be active at the lowest possible expression level. We show that in agreement with previous results⁴⁴ co-transfection in HEK293T cells can be used to provide a highly reproducible quantitative evaluation of both activity and potency for different shRNAs (**Fig. 5c**). Although both si and sh molecules targeting the 522 and 553 sites have previously been shown to have similar activities^{9, 10}, our results demonstrate that shRNA553 has much greater potency than shRNA522. This is consistent with a Rz screen of the 5'LTR which identified a target site similar to the 553 target site as being particularly accessible to Rz activity⁴². Our results demonstrate that like the 5980 target site, the 1498 target site can be used to design a potent shRNA with superior maximum activity compared to both

shRNA522 and shRNA553 (**Fig. 5c**). As the 1498 target site is also highly conserved (**Fig. 5a**), our results suggest that shRNA1498 represents a top candidate for RNA therapy, with the potential to be highly active at sub-saturating levels.

A unique liability for all antisense molecules is their potential to target cellular RNAs with complementary target sites. While target sites can be compared to the human genome, it is not always possible to predict non-specific targeting, as many antisense technologies do not require perfect homology to exert their effects. This is the first study to evaluate the specificity of SOFA-HDV-Rzs in human cells, and in agreement with *in vitro* studies^{24, 25} we found that a single nt mismatch either in SOFA-HDV-Rz1498 (**Fig. 4b**) or its binding site (pINDIE, **Fig. 6**) resulted in a loss of activity against HIV-1 production. In contrast, an imperfectly matched snoRNA-hammerhead Rz construct targeting HIV-1 RNA was active against HIV-1 production¹⁷ and several shRNAs have been shown to tolerate single and even double mismatches in HIV-1 RNA^{32, 45}. Consistent with previous results³², a single nt mismatch at position 17 did not affect the activity of shRNA1498 (**Fig. 6**), suggesting that the SOFA-HDV-Rz is more specific for its target site compared to an shRNA for a mutation at this position. This mutation represents the least conserved nt from our conservation estimates (**Fig. 5a**, Gag) and our results suggest that shRNA1498 would remain active against HIV-1 variants carrying this mutation. As it has been shown that mutations in the 5' and 3' extremities of siRNA target sites have less effects on their activity compared to those that occur in the central regions^{32, 45}, future screens of the HIV-1 genome for highly conserved siRNA targets may benefit from using differential cut-off values at the nt level, as we have done in this study for SOFA-HDV-Rz target sites (**Fig. 2a**). These considerations may also assist in predicting potential off target effects with programs such as the Ribosubstrates software³³ used in this study. Despite their different specificities and sequence requirements for off-target effects, neither SOFA-HDV-Rz1498 nor shRNA1498 had major effects on relative cellular mRNA

levels when compared to controls (**Fig. 7**), suggesting that the 1498 target site is sufficiently different from the human transcriptome to avoid off-targeting by small antisense RNAs.

Although shRNA1498 was several times more active compared to SOFA-HDV-Rz1498 against viral production in HEK293T cells (**Fig. 5b**), it had identical effects on viral kinetics in T lymphocytes (**Fig. 8**). This suggests that Rzs may be more active against viral replication relative to their effects against viral production. Indeed, several Rzs have been shown to have catalytic activity against viral replication^{12, 17, 18, 22, 41}, while only two have been shown to have catalytic activity against viral production^{17, 20}. As this is the first study to directly compare the activity of a Rz and an shRNA targeting the same site in both assays, additional comparisons will be required to determine whether this effect is generally observed for other Rzs. Direct comparisons of our molecules to others in primary cell models of HIV-1 replication will also be required to identify the top candidates from each class for use in therapy.

Overall, our data strongly suggest that the 1498 target site in HIV-1 RNA is an ideal target for small antisense RNAs in terms of conservation, accessibility, and specificity, and support a potential clinical application for both SOFA-HDV-Rz1498 and shRNA1498. Moreover, physical evidence for a unique specificity of the SOFA-HDV-Rz among small RNA antisense molecules is provided, suggesting that these Rzs can be used for the development of highly specific tools for research purposes and therapeutic applications.

Materials and methods

***In silico* identification of SOFA-HDV-Rz target sites**

Percent identity was calculated in reference to the consensus sequence at each nt in the 5' end of HIV-1 RNA from position 450 to 2300 in molecular clone pNL4-3 (GenBank: M19921). Sequence alignments were obtained using the Quick Align tool on the Los Alamos National Laboratory (LANL) website (www.hiv.lanl.gov) for all complete HIV-1 sequences (up to 1850 sequences at the time of analysis). Alignments were analyzed using Jalview Sequence editor⁴⁶ and conservation data were exported to Excel for target site identification⁴⁷ (Supplementary File S1). Nucleotide BLAST⁴⁸ was used to identify the position of the target site in HIV-1 pNL4-3 and exclude target sites for which this strain was not identical to the consensus sequence. Ribosubstrate software was then used to evaluate the potential for the corresponding SOFA-HDV-Rzs to target human RNAs as previously described³³. SOFA-HDV-Rzs with a score of 20 or less for any human RNA were excluded, where 0 represents a perfect match.

Plasmid constructs

All SOFA-HDV Rzs and shRNAs were expressed from the H1 promoter in the vector psiRNA-H1GFP::Zeo (InvivoGen, San Diego, CA). SOFA-HDV-Rz inserts were generated using an overlapping PCR strategy^{26, 29, 34} and shRNA inserts were generated by annealing complementary oligonucleotides. Sequences for shRNA522 and shRNA553 inserts were obtained from a previous study¹⁰ and using an identical design shRNAs (adapted from siControl⁹), shRNA5980 (adapted from sh1⁴⁹) and shRNA1498 inserts were designed. The cloning strategies and sequences of all oligonucleotides used for the generation of plasmid inserts are provided in the Supplementary Methods and Table S1.

HIV-1 production assay

psiRNA constructs and an HIV-1 molecular clone were co-transfected into HEK293T cells at different weight ratios and their effects on HIV-1 production were assessed 48 h later. The production of virus from co-transfected cells was estimated by measuring the activity of HIV-1 RT enzyme in culture supernatants and by the detection of HIV-1 Capsid (p24) and Gag (p55) proteins in cell lysates using previously described methods^{21, 47, 50}. Cell culture and transfection conditions for HIV-1 production assays, as well as northern blot methods to evaluate the relative expression of SOFA-HDV-Rzs in HEK293T cells are described in the Supplementary Methods.

Gene expression profiling

Gene expression levels relative to control transfections were analyzed by human gene expression microarrays (ACRI proprietary slides). The triplicate dye-swap experiments are described in the Supplementary methods.

HIV-1 infection assay

Stable Jurkat T lymphocytes were generated by electroporation of psiRNA constructs followed by selection with Zeocin (InvivoGen, San Diego, CA). The relative expression of psiRNA constructs in the stable cell populations was estimated by measuring GFP expression from the integrated vector with a FACS Calibur flow cytometer (BD Biosciences, San Jose, CA), and proliferation was determined by counting live cells by Trypan blue (Wisent, St Bruno, Canada) exclusion using a hemocytometer. Prior to infection, SOFA-HDV-Rz or shRNA expressing cells were plated in 6 well plates at 2×10^5 cells per well. Viral replication was monitored by measuring RT activity in culture supernatants at

various days post infection. Cell culture and selection conditions are described in the Supplementary Methods.

Acknowledgements

The authors would like to thank Daniel Leger (Atlantic Cancer Research Institute) for gridding and analysis of the microarray data, Mark Wainberg (McGill University) for the provision of HIV-1 molecular clones and Samantha Burugu (McGill University) for review of the manuscript. This study was supported by the Canadian Institutes of Health Research (CIHR) [CIHR: HOP-93434 and DCB-120266 to A.G., MOP-44022 to J-P.P]: R.J.S, M.V.L and S.M.D. were recipients of a Banting and Best Canada Graduate Scholarship, a doctoral fellowship from the Fonds de Recherche en Santé du Québec and a Vanier Canada Graduate Scholarship respectively. J-P.P. held the Canada Research Chair in Genomics and Catalytic RNA and now holds the Chaire de l'Université de Sherbrooke en structure et génomique de l'ARN. J-P.P. is member of the Centre de Recherche Clinique Etienne-Le Bel.

References

1. De Clercq, E. (2010). Antiretroviral drugs. *Curr Opin Pharmacol* **10**: 507-15.
2. Le Douce, V., Janossy, A., Hallay, H., Ali, S., Riclet, R., Rohr, O. *et al.* (2012). Achieving a cure for HIV infection: do we have reasons to be optimistic? *J Antimicrob Chemother* **67**: 1063-74.
3. Scherer, L.J. and Rossi, J.J. (2011). Ex vivo gene therapy for HIV-1 treatment. *Hum Mol Genet* **20**: R100-7.
4. Burnett, J.C. and Rossi, J.J. (2012). RNA-based therapeutics: current progress and future prospects. *Chem Biol* **19**: 60-71.
5. Haasnoot, J. and Berkhout, B. (2009). Nucleic acids-based therapeutics in the battle against pathogenic viruses. *Handb Exp Pharmacol*: 243-63.

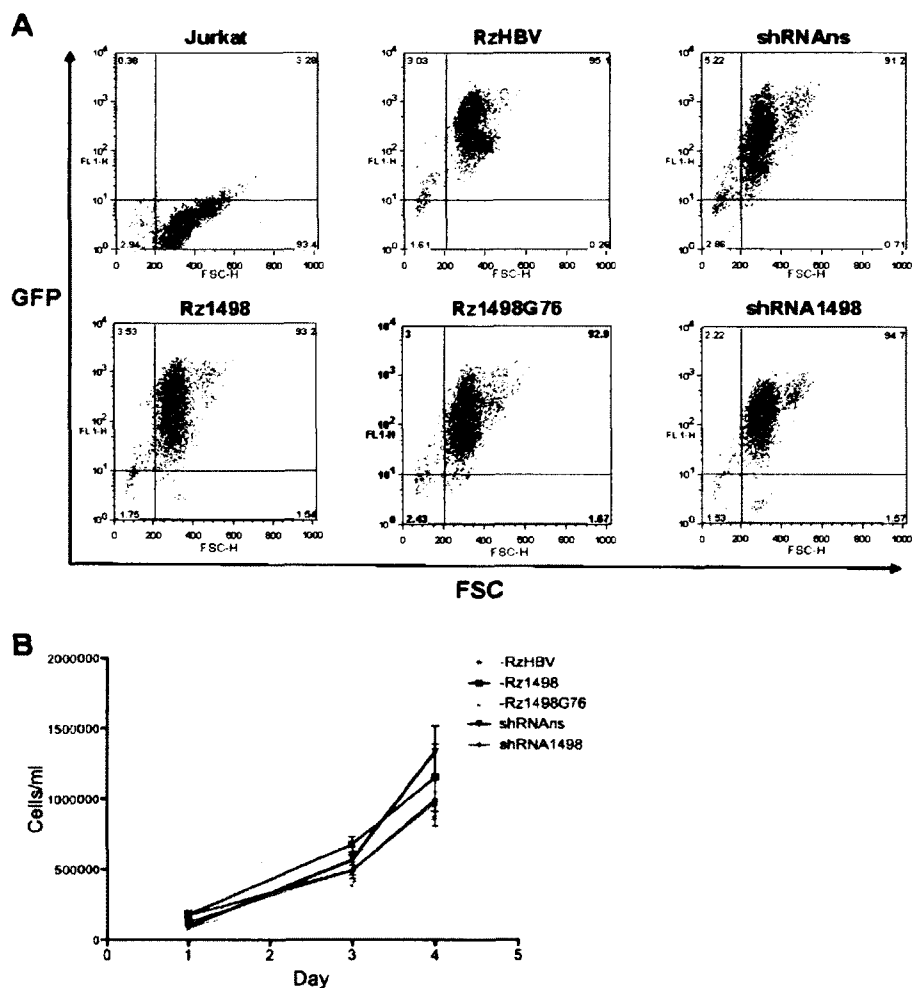
6. Watts, J.M., Dang, K.K., Gorelick, R.J., Leonard, C.W., Bess, J.W., Jr., Swanstrom, R. *et al.* (2009). Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature* **460**: 711-6.
7. Westerhout, E.M., Ooms, M., Vink, M., Das, A.T. and Berkhout, B. (2005). HIV-1 can escape from RNA interference by evolving an alternative structure in its RNA genome. *Nucleic Acids Res* **33**: 796-804.
8. Hemelaar, J., Gouws, E., Ghys, P.D. and Osmanov, S. (2011). Global trends in molecular epidemiology of HIV-1 during 2000-2007. *AIDS* **25**: 679-89.
9. Naito, Y., Nohtomi, K., Onogi, T., Uenishi, R., Ui-Tei, K., Saigo, K. *et al.* (2007). Optimal design and validation of antiviral siRNA for targeting HIV-1. *Retrovirology* **4**: 80.
10. McIntyre, G.J., Groneman, J.L., Yu, Y.H., Jaramillo, A., Shen, S. and Applegate, T.L. (2009). 96 shRNAs designed for maximal coverage of HIV-1 variants. *Retrovirology* **6**: 55.
11. Sajic, R., Lee, K., Asai, K., Sakac, D., Branch, D.R., Upton, C. *et al.* (2007). Use of modified U1 snRNAs to inhibit HIV-1 replication. *Nucleic Acids Res* **35**: 247-55.
12. Michienzi, A., Cagnon, L., Bahner, I. and Rossi, J.J. (2000). Ribozyme-mediated inhibition of HIV 1 suggests nucleolar trafficking of HIV-1 RNA. *Proc Natl Acad Sci U S A* **97**: 8955-60.
13. Mitsuyasu, R.T., Merigan, T.C., Carr, A., Zack, J.A., Winters, M.A., Workman, C. *et al.* (2009). Phase 2 gene therapy trial of an anti-HIV ribozyme in autologous CD34+ cells. *Nat Med* **15**: 285-92.
14. Good, P.D., Krikos, A.J., Li, S.X., Bertrand, E., Lee, N.S., Giver, L. *et al.* (1997). Expression of small, therapeutic RNAs in human cell nuclei. *Gene Ther* **4**: 45-54.
15. Puerta-Fernandez, E., Barroso-del Jesus, A., Romero-Lopez, C., Tapia, N., Martinez, M.A. and Berzal-Herranz, A. (2005). Inhibition of HIV-1 replication by RNA targeted against the LTR region. *AIDS* **19**: 863-70.
16. Chang, Z., Westaway, S., Li, S., Zaia, J.A., Rossi, J.J. and Scherer, L.J. (2002). Enhanced expression and HIV-1 inhibition of chimeric tRNA(Lys3)-ribozymes under dual U6 snRNA and tRNA promoters. *Mol Ther* **6**: 481-9.
17. Unwalla, H.J., Li, H., Li, S.Y., Abad, D. and Rossi, J.J. (2008). Use of a U16 snoRNA-containing ribozyme library to identify ribozyme targets in HIV-1. *Mol Ther* **16**: 1113-9.
18. Hotchkiss, G., Maijgren-Steffensson, C. and Ahrlund-Richter, L. (2004). Efficacy and mode of action of hammerhead and hairpin ribozymes against various HIV-1 target sites. *Mol Ther* **10**: 172-80.

19. Sánchez-Luque, F.J., Reyes-Darias, J.A., Puerta-Fernández, E. and Berzal-Herranz, A. (2010). Inhibition of HIV-1 replication and dimerization interference by dual inhibitory RNAs. *Molecules* **15**: 4757-72.
20. Zeng, W., Chen, Y.C., Bai, Y., Trang, P., Vu, G.P., Lu, S. *et al.* (2012). Effective Inhibition of Human Immunodeficiency Virus 1 Replication by Engineered RNase P Ribozyme. *PLoS One* **7**: e51855.
21. Lainé, S., Scarborough, R.J., Lévesque, D., Didierlaurent, L., Soye, K.J., Mougel, M. *et al.* (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**: 343-53.
22. Müller-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**: 265-72.
23. Asif-Ullah, M., Lévesque, 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**: 205-16.
24. Bergeron, L.J. and Perreault, J.P. (2005). Target-dependent on/off switch increases ribozyme fidelity. *Nucleic Acids Res* **33**: 1240-8.
25. Bergeron, L.J., Reymond, C. and Perreault, J.P. (2005). Functional characterization of the SOFA delta ribozyme. *RNA* **11**: 1858-68.
26. 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**: 4609-20.
27. D'Anjou, F., Routhier, S., Perreault, J.P., Latil, A., Bonnel, D., Fournier, I. *et al.* (2011). Molecular Validation of PACE4 as a Target in Prostate Cancer. *Transl Oncol* **4**: 157-72.
28. Lévesque, M.V., Lévesque, D., Brière, F.P. and Perreault, J.P. (2010). Investigating a new generation of ribozymes in order to target HCV. *PLoS One* **5**: e9627.
29. Motard, J., Rouxel, R., Paun, A., von Messling, V., Bisailon, M. and Perreault, J.P. (2011). A novel ribozyme-based prophylaxis inhibits influenza A virus replication and protects from severe disease. *PLoS One* **6**: e27327.
30. 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**: 869-79.
31. DeYoung, M.B. and Hampel, A. (1997). Computer analysis of the conservation and uniqueness of ribozyme-targeted HIV sequences. *Methods Mol Biol* **74**: 27-36.

32. Lee, S.K., Dykxhoorn, D.M., Kumar, P., Ranjbar, S., Song, E., Maliszewski, L.E. *et al.* (2005). Lentiviral delivery of short hairpin RNAs protects CD4 T cells from multiple clades and primary isolates of HIV. *Blood* **106**: 818-26.
33. Lucier, J.F., Bergeron, L.J., Brière, F.P., Ouellette, R., Elela, S.A. and Perreault, J.P. (2006). RiboSubstrates: a web application addressing the cleavage specificities of ribozymes in designated genomes. *BMC Bioinformatics* **7**: 480.
34. Lévesque, M.V. and Perreault, J.P. (2012). Target-induced SOFA-HDV ribozyme. *Methods Mol Biol* **848**: 369-84.
35. DiGiusto, D.L., Krishnan, A., Li, L., Li, H., Li, S., Rao, A. *et al.* (2010). RNA-based gene therapy for HIV with lentiviral vector-modified CD34(+) cells in patients undergoing transplantation for AIDS-related lymphoma. *Sci Transl Med* **2**: 36ra43.
36. Peden, K., Emerman, M. and Montagnier, L. (1991). Changes in growth properties on passage in tissue culture of viruses derived from infectious molecular clones of HIV-1LAI, HIV-1MAL, and HIV-1ELI. *Virology* **185**: 661-72.
37. Theodore, T.S., Englund, G., Buckler-White, A., Buckler, C.E., Martin, M.A. and Peden, K.W. (1996). Construction and characterization of a stable full-length macrophage-tropic HIV type 1 molecular clone that directs the production of high titers of progeny virions. *AIDS Res Hum Retroviruses* **12**: 191-4.
38. Mochizuki, N., Otsuka, N., Matsuo, K., Shiino, T., Kojima, A., Kurata, T. *et al.* (1999). An infectious DNA clone of HIV type 1 subtype C. *AIDS Res Hum Retroviruses* **15**: 1321-4.
39. Takahoko, M., Tobiume, M., Ishikawa, K., Ampofo, W., Yamamoto, N., Matsuda, M. *et al.* (2001). Infectious DNA clone of HIV type 1 A/G recombinant (CRF02_AG) replicable in peripheral blood mononuclear cells. *AIDS Res Hum Retroviruses* **17**: 1083-7.
40. Rossi, J.J., June, C.H. and Kohn, D.B. (2007). Genetic therapies against HIV. *Nat Biotechnol* **25**: 1444-54.
41. Sun, L.Q., Wang, L., Gerlach, W.L. and Symonds, G. (1995). Target sequence-specific inhibition of HIV-1 replication by ribozymes directed to tat RNA. *Nucleic Acids Res* **23**: 2909-13.
42. Bramlage, B., Luzi, E. and Eckstein, F. (2000). HIV-1 LTR as a target for synthetic ribozyme-mediated inhibition of gene expression: site selection and inhibition in cell culture. *Nucleic Acids Res* **28**: 4059-67.

43. Castanotto, D., Sakurai, K., Lingeman, R., Li, H., Shively, L., Aagaard, L. *et al.* (2007). Combinatorial delivery of small interfering RNAs reduces RNAi efficacy by selective incorporation into RISC. *Nucleic Acids Res* **35**: 5154-64.
44. Low, J.T., Knoepfel, S.A., Watts, J.M., ter Brake, O., Berkhout, B. and Weeks, K.M. (2012). SHAPE-directed discovery of potent shRNA inhibitors of HIV-1. *Mol Ther* **20**: 820-8.
45. Sabariego, R., Giménez-Barcons, M., Tàpia, N., Clotet, B. and Martínez, M.A. (2006). Sequence homology required by human immunodeficiency virus type 1 to escape from short interfering RNAs. *J Virol* **80**: 571-7.
46. Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M. and Barton, G.J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**: 1189-91.
47. Scarborough, R., Lévesque, M.V., Perreault, J.P. and Gagnon, A. (In Press). Therapeutic applications of Ribozymes and Riboswitches: Design and evaluation of clinically relevant SOFA HDV Ribozymes targeting HIV RNA. *Methods Mol Biol*.
48. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. *et al.* (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389-402.
49. Li, M.J., Kim, J., Li, S., Zaia, J., Yee, J.K., Anderson, J. *et al.* (2005). Long-term inhibition of HIV-1 infection in primary hematopoietic cells by lentiviral vector delivery of a triple combination of anti-HIV shRNA, anti-CCR5 ribozyme, and a nucleolar-localizing TAR decoy. *Mol Ther* **12**: 900-9.
50. Clerzius, G., Gélinas, J.F., Daher, A., Bonnet, M., Meurs, E.F. and Gagnon, A. (2009). ADAR1 interacts with PKR during human immunodeficiency virus infection of lymphocytes and contributes to viral replication. *J Virol* **83**: 10119-28.

Supplementary figure



Chapitre 2, Figure S 1. Stable Jurkat cell lines analysis.

A) 10,000 events are shown for the different cell lines, Green fluorescence levels versus Forward scatter (FSC) is plotted, showing the level of green fluorescent protein (GFP) expression and approximate cell size (FSC) for the different Jurkat cell lines. **B)** Live cell counts \pm SEM (n=2) out to four days after plating different cell lines at 1×10^5 cells/mL.

Supplementary methods

Plasmid constructs:

SOFA-HDV-Rz inserts were generated by PCR using DNA primers A and B (2 μ M), C and D (25 nM):

A: 5'-TATAAGTTCTGTATGAGTTCACGGAAGACCGACCT↓CGGGCCAGCTAGT
TT-3'

B: 5'-CAACAACAGTGTTCCGGATGAACTGATGCTATGAAGACTCCAAA↓AACCA
GCTAGAAAGGGTC-3'

C: 5'-CCAGCTAGAAAGGGTCCCTTA_gCCATCCGCGAACGGATGCC-3'

D: 5'-TAATACGACTCACTATAGGGCCAGCTAGTTT(Bs)(BI)CAGGGTCCACCTC
CTCGCGGT(RD)GGGCATCCGTTTCGCG-3'

Primers A and B include *Bbs*I recognition sites (bold, cut site indicated with an arrow), C and D contain a common reverse and specific forward ribozyme sequence respectively (overlapping sequences underlined). The nucleotide in primer C that was mutated to produce a catalytically inactivate SOFA-HDV Rz is shown as lower case (G to T mutation in this study). Variable Bs, BI and RD sequences in primer D are shown in Supp Table S1 for each SOFA-HDV-Rz evaluated.

shRNA inserts were generated by annealing complementary sense (S) and antisense (AS) oligonucleotides (1.25 μ M each in 75 mM NaCl, 40 μ L, 2 min at 80 °C, cooled to 37 °C):

shRNA522 S: 5'-ACCTCGCCTCAATAAAGCTTGCCTTCCTCGAGCAAGGCAAGC
TTTATTGAGGCTT-3'

shRNA522 AS: 5'-CAAAAAGCCTCAATAAAGCTTGCCTTGCTCGAGGAAGGCAA
GCTTTATTGAGGCG-3'

shRNA553 S: 5'-ACCTCGTAGTGTGTGCCCGTCTGTTTCCTCGAGCAACAGACG
GGCACACACTACTT-3'

shRNA553 AS: 5'-CAAAAAGTAGTGTGTGCCCGTCTGTTGCTCGAGGAACAGAC
GGGCACACACTACG-3'

shRNAs S: 5'-ACCTCGTACCGCACGTCATTCGTATCCTCGAGCATACGAATGA
CGTGCGGTACTT-3'

shRNAs AS: 5'-CAAAAAGTACCGCACGTCATTCGTATGCTCGAGGATACGAAT
GACGTGCGGTACG-3'

shRNA5980 S: 5'-ACCTCGCGGAGACAGCGACGAAGAGGCTCGAGGCTCTTCG
TCGCTGTCTCCGCTT-3'

shRNA5980 AS: 5'-CAAAAAGCGGAGACAGCGACGAAGAGCCTCGAGCCTCTT
CGTCGCTGTCTCCGCG-3'

shRNA1498 S: 5'-ACCTCGCAGGAACTACTAGTACCCTACTCGAGAAGGGTACT
AGTAGTTCCTGCTT-3'

shRNA1498 AS: 5'-CAAAAAAGCAGGAACTACTAGTACCCTTCTCGAGTAGGGTA
CTAGTAGTTCCTGCG-3'

The nucleotides flanking the core loop sequence (bold) for the HIV-1 specific shRNA522, shRNA553, shRNA5980 and shRNA1498 are complementary to the 21st nt of their target site in case of differential processing as previously described (manuscript reference #10). *Bbs* I cut site overhangs in the complementary oligonucleotides are underlined.

SOFA-HDV-Rz and shRNA inserts were ligated into *Bbs* I (Fermentas) digested psiRNA-H1GFP::Zeo (InvivoGen) expression plasmid. All constructs were confirmed by sequencing using a primer located in the H1 promoter: 5'-TCTACGGGGTCTGACGC-3'

Cell culture:

HEK293T cells were maintained in Dulbecco's modified Eagle's medium with high glucose (Hyclone) supplemented with 10 % fetal bovine serum (Hyclone), 50 U/mL Penicillin and 50 µg/mL Streptomycin (Gibco). Jurkat T cells were maintained in Roswell Park Memorial Institute 1640 (Hyclone) supplemented with 10 % heat

inactivated (55°C, 30min) fetal bovine serum (Hyclone), 50 U/mL Penicillin and 50 µg/mL Streptomycin (Gibco).

Transient transfection:

24 h prior to transfection, HEK293T cells were plated at 2×10^5 cells/mL in 24 well plates for evaluation of viral production, or 6 well plates for intracellular protein production, Rz expression levels and gene expression profiling. For viral production, 75 ng of HIV-1 plasmid DNA and 750 ng of psiRNA plasmid DNA (1:10) were cotransfected. For lower ratios, 100 ng of pNL4-3 plasmid DNA and 0.5 to 500 ng of psiRNA plasmid DNA (500:1 to 1:5) were used, topped up to 800 ng total DNA with pBluescript SK(+) (Stratagene). For evaluation of protein levels, Rz expression levels and gene profiling, 200 ng of HIV-1 plasmid DNA and 2 µg of psiRNA plasmid DNA (1:10) were cotransfected. All transfections were carried out using TransIT reagent (Mirus) according the manufacturers instructions. Supernatants or cell extracts were harvested 48 h after transfection.

Northern blot:

Total RNA extracts were harvested from transfected cells using Trizol reagent (Invitrogen) according to the manufacturers instructions. Total RNA samples were dissolved in RNase-free water and quantified (Nanvovue, Roche). For each condition, 10 µg of total RNA mixed with 2 volumes of loading buffer was resolved on 8 % denaturing PAGE. The RNA was then transferred (90 min, 4 °C, 200 mA) to a nylon membrane (Amersham Hybond-N+, GE Healthcare) in 0.5× TBE using a Trans-blot cell (Bio-Rad). Membranes were UV crosslinked prior to pre-hybridization at 42 °C in CHURC buffer (1 % (W/V) BSA, 1 mM EDTA, 500 mM phosphate buffer and 7 % (W/V) SDS). After 2 to 4 h, the radio-labeled ProbeSOFA was added, and the membranes were hybridized at 42 °C overnight.

Before analysis, the membranes were washed twice, 5 min each, in wash buffer #1 (2× SSC, 0.1 % SDS) and twice, 15 min each, in wash buffer #2 (0.1× SSC, 0.1 % SDS), all at 42 °C. The results were visualized using a Phosphor Screen. Then, to proceed to the Probe5S hybridization as described above, the membranes were stripped by incubation in wash #2 at 80 °C for 20 min to remove the ProbeSOFA. Only 5 to 10 % of the labeled Probe5S was used for the hybridization of one membrane. The sequence of each probe is presented in the RNA labeling section.

DNA templates for *in vitro* transcription:

SOFA-HDV-Rz DNA templates were generated through a PCR-based strategy with DNA oligonucleotides C and D shown in Plasmid constructs. The SOFA-HDV-Rz sense primer (D) provided the T7 RNA polymerase promoter needed for subsequent *in vitro* transcription. The fully double-stranded DNA sequences were produced using *Pwo* DNA polymerase (Roche Diagnostics). Similarly, the substrate DNA template was produced by a combination of two complementary oligonucleotides, subA and subB:

subA: 5'-taatacgactcactataGGGCATAGCAGGAACTACTAGTACCCTTGGGTCCG
CAGGGTCCACCTCC-3'

subB: 5'-GGGTCCCTTAGCCATGCGAAGCCGCATGCCAGGTCCGACCGCGA
GGAGGTGGACCCTGCCGACCC-3'

SubA contained the T7 RNA polymerase promoter (lower case) at its 5'-end and part of a *cis*-acting HDV ribozyme (underlined sequence) at its 3'-end. The subB primer is complementary (underlined sequence) to the forward primer and completes the *cis*-acting HDV ribozyme. The *cis*-acting HDV ribozyme permits the production of a precise 3'-end as described in Supplementary reference (1). The

final RNA substrate, SubMin1498, corresponds to the sequence in bold. All PCR reactions were ethanol precipitated prior to *in vitro* transcription.

RNA synthesis for *in vitro* transcription:

The SOFA-HDV-Rz and the substrate RNA were synthesized by run-off transcriptions as described previously (manuscript reference #34). Briefly, transcriptions were performed in the presence of purified T7 RNA polymerase (10 µg), pyrophosphatase (0.01 U, Roche Diagnostics) and PCR product (2 to 5 µM) in a buffer containing 80 mM HEPES-KOH (pH 7.5), 24 mM MgCl₂, 2 mM spermidine, 40 mM DTT and 5 mM of each NTP in a final volume of 100 µL at 37 °C for 2 h. The reaction mixtures were then treated with RQ1 DNase (Promega) at 37 °C for 20 min. After phenol/chloroform extraction the RNAs were ethanol precipitated. The pellets were dissolved in equal volumes of ultrapure water and loading buffer (95 % formamide, 10 mM 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 buffer containing 45 mM Tris-borate (pH 7.5), 8 M urea, and 2 mM EDTA. The RNA products were visualized by ultraviolet (UV) shadowing. The bands corresponding to the correct sizes for both the SOFA-HDV-Rzs and the substrate were cut out of the gel and the RNAs eluted overnight at 4 °C in elution buffer (500 mM ammonium acetate, 10 mM EDTA, 0.1 % SDS). The samples were again ethanol precipitated, washed, dried and dissolved in ultrapure water. The RNA was quantified by absorbance at 260 nm and diluted to the desired concentration (Rz 1 µM and Substrate 5 µM).

RNA and Probes Labelling:

The RNA substrate used in cleavage reactions was 5'-end labeled as described previously (manuscript reference #34). Briefly, the purified RNA substrate was 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 then incubated for 30 min at 37 °C. Incubation at 65 °C for 8 min was used to inactivate the enzyme. The 5'-end labeling reaction was performed with 5 pmol of dephosphorylated RNAs that were incubated for 1 h at 37 °C with 3 U of T4 polynucleotide kinase (USB) and 3.2 pmol of [γ -³²P]ATP (6000 Ci/mmol; New England Nuclear) in the reaction buffer provided with the enzyme. The reactions were stopped by the addition of two volumes of loading buffer prior to fractionation by 20 % denaturing PAGE. The RNAs were detected by autoradiography, cut out of the gel and eluted as described above. For the northern DNA probes, either 5 pmol of the oligonucleotide ProbeSOFA or 10 pmol of the oligonucleotide Probe5S were labeled in the same conditions as described for the dephosphorylated RNA substrate. The quantity of [γ -³²P]ATP was doubled for the ProbeSOFA labeling reaction. The labeled oligonucleotides were purified with ProbeQuant[™] G50 Micro Columns (GE Healthcare) and used directly for the northern blot, their sequences were:

ProbeSOFA: 5'-GAAAGGGTCCCTTAGCCATCCGCGAACGGATGCCC-3'

Probe5S: 5'-AAAGCCTACAGCACCCGGTATTCCC-3'.

***In vitro* cleavage assay:**

An excess of ribozyme compared to the substrate (single turnover conditions) was used for the *in vitro* cleavage assays as published before (manuscript reference #34). Specifically, a trace amount of 5'-end-labeled substrate (< 1 nM) was incubated at 37 °C with a final concentration of 100 nM of the selected SOFA-HDV-

Rz. For each experimental condition, 3 μL of the 5'-end-labeled SubMin1498 substrate (concentration under 10 nM), 3 μL SOFA-HDV-Rz (1 μM), 3 μL of cleavage reaction buffer (500 mM Tris-HCl, pH 7.5) and 18 μL of RNase free water were mixed and pre-incubated at 37 °C for 25 min. The cleavage reactions were initiated by the addition of 3 μL MgCl_2 (100 mM) to reach a final volume of 30 μL . Aliquots (2 μL each) were sampled at different time intervals and transferred to a new tube containing 10 μL of loading buffer to stop the reaction. To determine the percentage of cleavage, all recovered samples were fractionated on a 20 % denaturing PAGE under the conditions described above. The results were visualized using a Phosphor Screen, and were quantified using the ImageQuant software (Molecular Dynamics). The control condition reaction was performed in the absence of ribozymes (replaced by water) and its last time interval sample was used to subtract the background. For each time point, the percentage of cleavage was calculated (cleaved product counts over cleaved + uncleaved products counts). The k_{obs} were then calculated using GraphPad Prism 5 for each ribozyme. 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 at time t , A_{∞} is the maximum percent cleavage and k is the rate constant (k_{obs}).

RNA extraction and mRNA analysis:

Total RNA extracts were harvested from transfected cells using Trizol reagent (Invitrogen) according to the manufacturers instructions and purified using an RNeasy column (Qiagen). Quality of total RNA samples was assessed using the Experion bioanalyzer system with RNA Stdsens chips and associated reagents (BIO-RAD). All RNA used in these experiments had an RNA Quality Index (RQI) value greater than nine. 1 μg of each total RNA sample was amplified using the Amino Allyl MessageAmp II aRNA amplification kit and subsequently labeled with AlexaFluor 555 or 647 (Life Technologies). Quantity and quality of amplified aRNA

was assessed using a Nanodrop spectrophotometer and the Experion bioanalyzer. Samples were compared in triplicate dye swap experiments, with 1.5 μg of each labeled, fragmented aRNA (3 μg total per slide) hybridized to proprietary human cDNA microarray slides. These arrays consist of roughly 35000 spots, representing roughly 17000 different 50-mer oligonucleotides spotted in duplicate on Nexterion-E epoxy microarray slides (Schott).

Hybridizations were performed in Ambion SlideHyb #2 buffer (Life Technologies) at 42 °C for 16 h using the automated TECAN 4800 Hybridization station (TECAN). Following hybridization, slides were scanned at 10 μm resolution using an Axon GenePix 4200AL scanner (Molecular Devices, Sunnyvale, CA) and gridded using SpotReader (Niles Scientific). Fine tuning of spot rejection was subsequently done by visual inspection of the gridded image and of a scatter plot of $M = \log(532/635)$ versus $A = \log(532) + \log(635)$, special attention being paid to outliers. A GPR file was generated that was subsequently processed to flag spots with a signal to noise ratio of less than 5.

Analysis was done with Acuity 4.0 (Axon Instruments). The data was normalized using Lowess and analyzed by various statistical and visual methods such as SOM (Self Organizing Maps), t-Test, PCA (Principal Component Analysis) and volcano plot.

Generation and characterization of stable Jurkat T cells

1×10^6 Jurkat T cells were electroporated at 250 mV for 10 msec (GenePulserII, Biorad) with 15 μg of psiRNA plasmids and cultured in 5 mL RPMI overnight. Cells were transferred to 20 mL culture media containing 600 $\mu\text{g}/\text{mL}$ Zeocin (InvivoGen) and cultured for 4 weeks with selection media changed every 3-4 days. Cells were split at various times during the selection to maintain a high cell density in the first 2 weeks of selection, and a low cell density in the last 2 weeks of selection. At the

end of the selection period cells were frozen, or cultured for an additional day in the absence of Zeocin prior to infection with HIV-1 virions. The expression of GFP in the different cell populations was determined using a FACS Calibur flow cytometer (BD) and data analysis was performed using FlowJo Version 8.7 (Tree Star). Cell proliferation was determined by plating cells at 1×10^5 cells/mL in a 6 well plate followed by counting live cells diluted in Trypan blue (Wisent) with a hemocytometer. Duplicate wells were counted for each cell line out to four days.

Virus infection:

A 20 mL culture of HEK293T cells was transfected with 20 μ g of HIV-1 pNL4-3 plasmid DNA. The supernatant was harvested 48 h later, cleared of cell debris by centrifugation, and stored at -80°C in 1 mL aliquots. SOFA-HDV-Rz and shRNA expressing Jurkat T cells were plated in 6 well plates at 2×10^5 cells/ well and infected with the HEK293T cell supernatant corresponding to 6×10^5 cpm equivalent of HIV-1 RT activity per well. Viral replication was monitored by measuring RT activity in the culture supernatants at various days post infection. All infections were performed in triplicate wells, cells were split twice a week.

Supplementary references

1. Avis, J.M., Conn, G.L. and Walker, S.C. (2012) Cis-acting ribozymes for the production of RNA in vitro transcripts with defined 5' and 3' ends. *Methods Mol Biol*, 941, 83-98.

Chapitre 3 : Sélection des ribozymes SOFA-HDV avec le meilleur potentiel

ARTICLE : *Selection of The Most Potent SOFA-HDV Ribozymes For Use in Gene Targeting*

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

Article publié dans : *Nucleic Acid Therapeutics* (2011) 21(4): 241-252

Avant-propos :

J'ai fait la conception de toutes les expériences présentées dans l'article ainsi que l'analyse des résultats. J'ai généré tous les résultats présentés dans les figures, alors que certaines expériences avaient initialement été faites par Samuel Rouleau (expériences avec le biosenseur, Figure 2 et Tableau 1). J'ai rédigé le papier en entier avec la supervision de Jean-Pierre Perreault (analyses, rédaction, figures et tableaux).

Résumé

Le ribozyme du Virus de l'Hépatite Delta («*Hepatitis Delta Virus*», HDV), lequel est bien adapté à l'environnement des cellules humaines, est un excellent candidat pour le développement de futurs systèmes d'inactivation de gènes. De plus, il existe maintenant une nouvelle génération de ribozymes HDV qui bénéficie de l'addition d'un module de spécificité («*Specific On/off Adaptor*», SOFA), à savoir les ribozymes SOFA-HDV. Cet ajout augmente grandement la spécificité du ribozyme et son activité. Contrairement à l'ARNi et aux ribozymes «*hammerhead*», la conception des ribozymes SOFA-HDV pour couper *en trans* un ARN donné n'a jamais fait l'objet d'une étude d'optimisation systématique, et ce malgré leur récente utilisation pour l'inactivation génique de différentes cibles. Cet article vise à améliorer et clarifier le processus de design (la sélection) des ribozymes SOFA-HDV. Le ribozyme et le substrat ARN ont tous deux été analysés afin de fournir de nouveaux critères utiles dans la sélection des ribozymes SOFA-HDV avec le meilleur potentiel. Les éléments cruciaux présents dans le biosenseur et le bloqueur du ribozyme, ainsi que dans le site cible, ont été identifiés et caractérisés. Des règles simples furent dérivées et testées en utilisant l'ARN de NS5B du VHC comme cible modèle. Dans l'ensemble, cette méthode devrait promouvoir l'utilisation des ribozymes SOFA-HDV dans une pléthore d'applications, autant pour la génomique fonctionnelle que pour la thérapie génique.

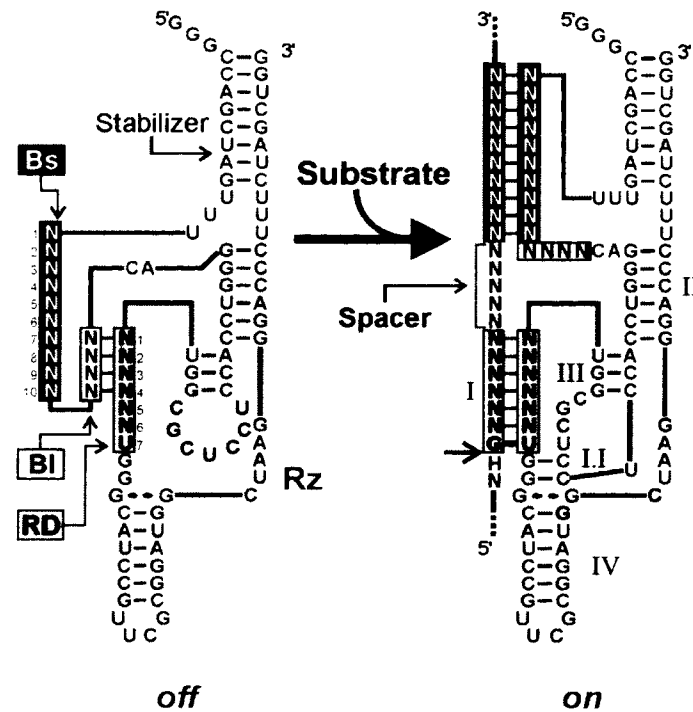
Abstract

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 both to improve and to clarify 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 in the target site, were identified and characterized. Simple rules were derived and tested using HCV 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 one 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 (i.e. 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 Mg^{2+}). Thus, the HDV ribozyme appears to be well adapted to the human cell environment, as is demonstrated by its outstanding stability in cellulose (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 (i.e. the stem I) called the recognition domain (Fig. 1). In order to overcome this hurdle, molecular engineering was used to develop a target-dependent activation module named the Specific On/oFF Adapter (SOFA) (Bergeron and Perreault, 2005; Bergeron et al.,

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) which 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 potentially could be bound by the recognition domain. The SOFA module also includes a biosensor (Bs) which extends the base-pairing with the target. The formation of a duplex between the biosensor 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 that 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 ribozyme.



Chapitre 3, Figure 1. Secondary structure representations of the SOFA-HDV ribozyme in both the off (left) and on (right) conformations.

The HDV ribozyme is highlighted in grey in both structures, while the cognate substrate is drawn only for the *on* conformation since it is the modulator of the conformational switch. The three key variable features of the ribozyme; the recognition domain, the biosensor and the blocker, are identified by RD (black on grey), Bs (white on black) and Bl (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, I.I, II, III and IV identify the corresponding stems or stem-loops in the HDV ribozyme, while the arabic numbers indicate the positions of each nucleotide in the Bs and the RD. The variable nucleotides (i.e. 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 one ribozyme to the other, regardless of the target sequence. The short arrow indicates the cleavage site.

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 and Perreault, 2005; Bergeron et al., 2005). Several examples of the development of gene inactivation systems in both prokaryotic and eukaryotic cells have been reported (D'Anjou et

al., in press; Fiola et al., 2006; Laine et al., 2011; Levesque et al., 2010; Levesque and Perreault, in press; Robichaud et al., 2008). In the course of these studies, over 100 distinct SOFA-HDV ribozymes were designed and tested *in vitro* in order 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 %, 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 was 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 in order to establish new criteria for improving the SOFA-HDV ribozyme design process. These experiments support the importance of a critical selection of the ribozymes prior to 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 PCR-based strategy using different pairs of complementary oligonucleotides as described previously (Levesque et al., 2010; Levesque and Perreault, in press). Briefly, two DNA oligonucleotides were used for the production of the SOFA-HDV ribozymes: the universal reverse primer (5'-CCAGCTAGAAAGGGTCCCTTAGCCATCCGCGAACGGATGCCC-3') and the SOFA-HDV RzX (where X identifies the specific ribozyme) sense primer (5'-TAATACGACTCACTATAGGGCCAGCTAGTTT(N)_{10Bs}(N)_{4Bl}CAGGGTCCACCTCC TCGCGGT(N)_{6RD}TGGGCATCCGTTTCGCGG-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 enables the incorporation of the T7 RNA polymerase promoter, while 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 two complementary oligonucleotides, one of which contained the T7 RNA polymerase promoter at its 5'-end. All PCR reactions were ethanol precipitated prior to *in vitro* transcription.

RNA synthesis

Both the SOFA-HDV ribozymes, and their corresponding substrates, were synthesized by run-off transcriptions as described previously (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 to 5 µM) in a buffer containing 80 mM HEPES-KOH (pH 7.5), 24 mM MgCl₂, 2 mM spermidine, 40 mM DTT and 5 mM of each ribonucleotide triphosphates (rNTP) in a final volume of 100 µL at 37°C for 2 h. After completion, the reaction mixtures were treated with RQ1 DNase (Promega) at 37°C for 20 min. 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 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 buffer containing 45 mM Tris-borate (pH 7.5), 8 M urea, and 2 mM EDTA. The RNA products were visualized by ultraviolet (UV) 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 then eluted overnight at 4°C in elution buffer (500 mM ammonium acetate, 10 mM EDTA, 0.1 % SDS). 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 µL of [α -³²P]UTP (3000 Ci/mmol; New England Nuclear) was added to each *in vitro* transcription reaction. The transcripts were purified (phenol/chloroform extraction and ethanol precipitation) as described above, and samples corresponding to 1 % of each reaction were fractionated through 10 % denaturing polyacrylamide gels and then exposed to a Phosphor Screen in order to visualize the results (Molecular Dynamics).

RNA substrate labeling

The RNA substrates used in the different cleavage reactions were 5'-end labeled as described previously (Levesque et al., 2010). In brief, 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 µL containing the buffer provided with the enzyme, and then incubating for 30 min at 37°C. The enzyme was then inactivated by incubation at 65°C for 8 min. The dephosphorylated RNAs (5 pmol) were then 5'-end labeled by incubation for 1 h at 37°C with 3 U of T4 polynucleotide kinase (USB Corp.) and 3.2 pmol of [γ -³²P]ATP (6000 Ci/mmol; New England Nuclear) in the reaction buffer provided with the enzyme. The reactions were stopped by the addition of two 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 described above.

***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 (< 1nM) 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 µL. For each experimental condition, 3 µL of the proper 5'-end-labeled substrate at a concentration lower than 10 nM, 3 µL of 1 µM SOFA-HDV ribozyme, 3 µL of cleavage reaction buffer (500 mM Tris-HCl, pH 7.5) and 18 µL of RNase free water were mixed and pre-incubated at 37°C for 25 min. The cleavage reactions were initiated by the addition of 3 µL of 100 mM MgCl₂. At different time intervals, 2 µL aliquots were transferred in a new tube containing 10 µL of loading buffer in order to stop the reaction. All aliquots were then fractionated on a 20 % denaturing PAGE gel under the conditions described above. The results were visualised 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

In order 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. RNAhybrid is a tool for finding the minimum free energy of hybridization (mfe) of two RNAs, and is primarily used for miRNA target prediction. Briefly, the sequence corresponding to the stem-loop III (5'-ACCUCCUCGCGGU-3') was

considered as the target input, while the sequences corresponding to the linker (i.e. 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_{Bl4}CA-3') of each ribozyme were taken as the microRNA input in the software. The minimum free energy of hybridization 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 Figure 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 HCV NS5B gene (1773 nucleotides, AJ242654) was performed manually, looking for H₋₁G₊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 finished 22 nucleotides downstream in order to include all of 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 two groups of ribozymes targeting NS5B, the unpaired t test was performed using GraphPad Prism 5.0.

Results

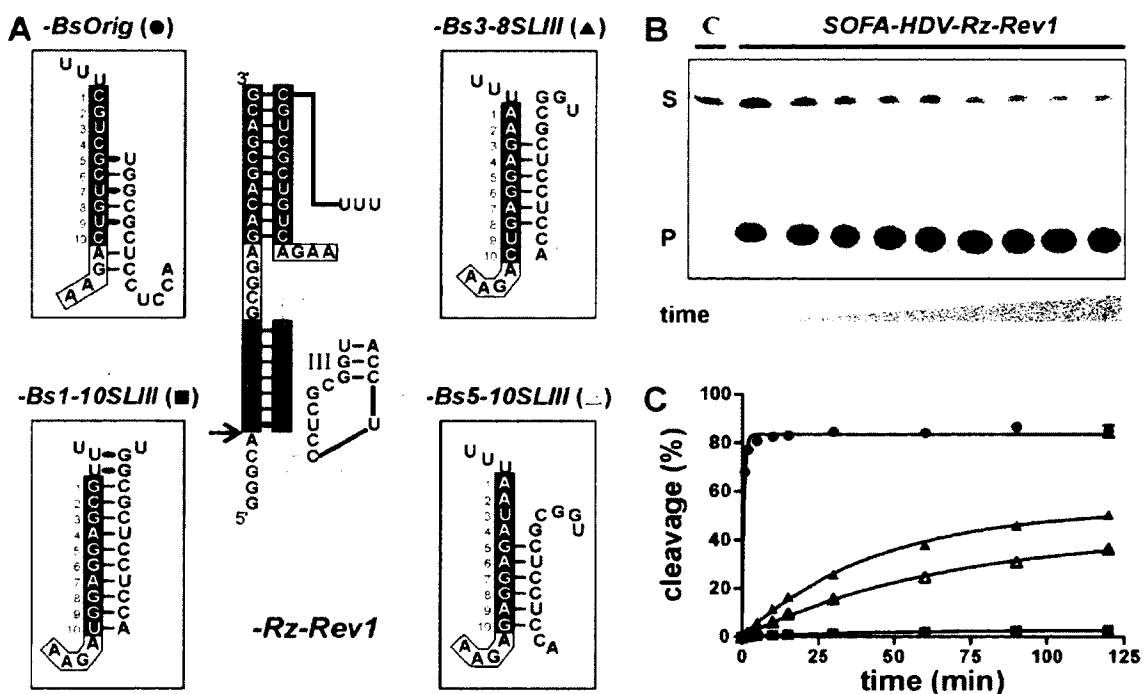
Previous experiments with the HDV ribozyme demonstrated the existence of a strong relationship between its structure and its cleavage activity (Reymond et al., 2009). Consequently, the same logic will be used here in order to refine the design process for the identification of the SOFA-HDV ribozymes possessing the greatest potential for targeting a given RNA substrate. Initially, a set of three SOFA-HDV ribozymes was selected with which to test the various hypotheses and to insure that the conclusions drawn can be applied to other ribozymes. The ribozymes in question (i.e. SOFA-HDV-Rz-Rev1, SOFA-HDV-Rz-NS527 and SOFA-HDV-Rz-HBV513) are directed against the human immunodeficiency virus (HIV), the

influenza A virus and the hepatitis B virus (HBV), respectively. The target sequence of each ribozyme is presented in Table 1.

Interaction of the biosensor with the catalytic core

The HDV ribozyme structure contains single-stranded regions that must 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 (i.e. 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, reduces 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. In order to verify this hypothesis, different biosensors that could potentially base pair with the stem-loop III were designed for each of the three model SOFA-HDV ribozymes, and their cleavage activities were then assessed for each under single turnover conditions ($[Rz] \gg [S]$). Briefly, a trace amount of 5'-end ^{32}P -labeled substrate ($< 1\text{nM}$) was incubated in the presence of 100 nM SOFA-HDV ribozyme in a buffer containing 10 mM MgCl_2 for 25 min at 37°C . The reactions were stopped by the addition of loading buffer, and were then fractionated on polyacrylamide gels. A typical autoradiogram is shown in Figure 2 for SOFA-HDV-Rz-Rev1. As expected, the ribozyme with its original biosensor (-BsOrig) is highly active, with more than 80 % cleavage being achieved within 5 min. 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 h of incubation, less than 5 % of cleavage was observed. Even when

the biosensor sequence was modified in order 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 two cases the reductions were less drastic than that observed with the 10 bp.



Chapitre 3, Figure 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 min) and fractionated on a 20 % denaturing PAGE 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.

In order 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 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 softwares, 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; i.e. 5'-UUUBs₁₀Bl₄CA-3') in each case. It was important to consider the flanking sequences in the prediction in order 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 (i.e. 5'-ACCUCCUCGCGGU-3'). The predicted interactions for the four biosensors of the SOFA-HDV-Rz-Rev1 are presented in Figure 2, while the resulting minimum free energies of hybridization (mfe) values are reported in Table 1. Concurrently, k_{obs} and the cleavage percentage after 2 hr, 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 three 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 like the SOFA-HDV-HBV513-Bs3-8SLIII or -Bs5-10SLIII. In these cases, the percentages of cleavage after 2 hours are less affected as compared to other ribozymes while 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 CUA A or from other interactions that decrease the strength of the inhibitory interaction. The four 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).

Chapitre 3, Tableau 1. Cleavage activities and thermodynamic parameters of the SOFA-HDV 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 (%)	MFI (kcal/mol)
-Rev1-Bsorig	5'-GGCA	GGAAGAA	GCCGA	GACAGCGACC-3'	1.606 ± 0.082	85	-10.6
-Bs1-10SLIII	5'-GGCA	GGAAGAA	GCCGA	acCucCuegc-3'	0.024 ± 0.004	3	-27.0
-Bs3-8SLIII	5'-GGCA	GGAAGAA	GCCGA	GACucCucuu-3'	0.023 ± 0.001	50	-15.1
-Bs5-10SLIII	5'-ggGCA	GGAAGAA	GCCGA	cuCcuCuAuu-3'	0.015 ± 0.001	36	-15.6
-NS527-Bsorig	5'-ggUCCA	GGACAU A	CUAA	UGAGGAUGLC-3'	0.420 ± 0.020	90	-5.6
-Bs1-10SLIII	5'-ggUCCA	GGACAU A	CUAA	accucUcgC-3'	0.061 ± 0.025	10	-26.4
-Bs3-8SLIII	5'-ggUCCA	GGACAU A	CUAA	UGcucUcUu-3'	0.040 ± 0.003	72	-14.9
-Bs5-10SLIII	5'-ggUCCA	GGACAU A	CUAA	cucucUcUu-3'	0.119 ± 0.007	58	-15.4
-HBV513-Bsorig	5'-gggcuaa	GCGUCCG	aaaa	AUCUC AUCL-3'	0.726 ± 0.063	88	-9.2
-Bs1-10SLIII	5'-gggcuaa	GCGUCCG	aaaa	acCUCcucgc-3'	0.027 ± 0.003	21	-26.4
-Bs3-8SLIII	5'-gggcuaa	GCGUCCG	aaaa	AUCUCcucCU-3'	0.137 ± 0.038	77	-15.4
-Bs5-10SLIII	5'-gggcuaa	GCGUCCG	aaaa	cCUCcucAUCL-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.

Similar experiments were also performed in order to verify whether or not the biosensor can interact with the junction IV/II. When the sequence of the biosensor includes a stretch of nucleotides that can base-pair with the five 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 must 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 4^{10} 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.

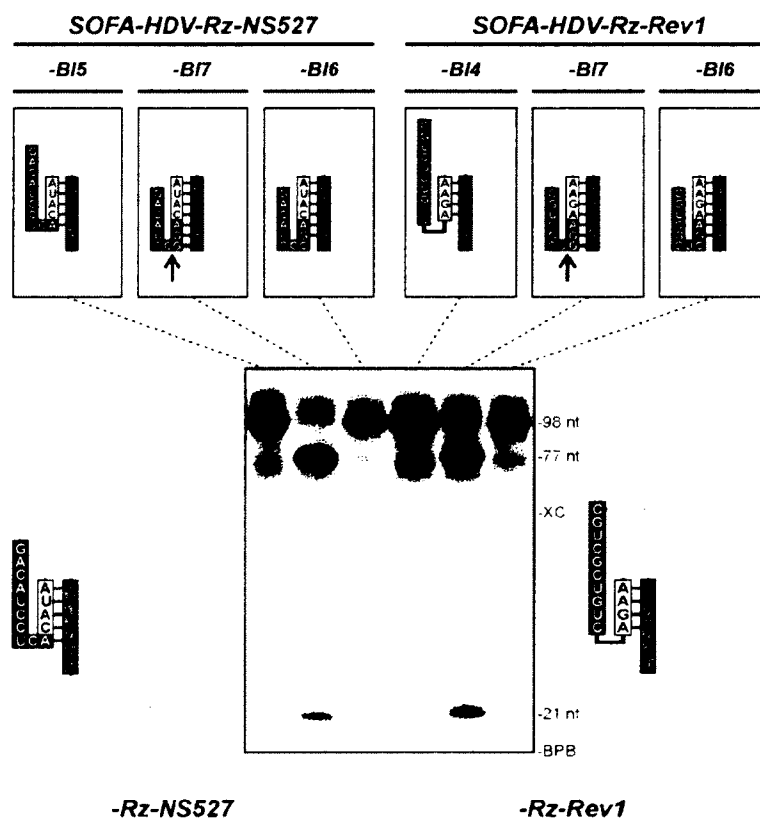
In summary, only the intramolecular base-pairing between the biosensor and the stem-loop III appears to be an important factor reducing the potential of the SOFA-HDV ribozyme. Detailed analysis of the SOFA-HDV ribozymes in the library developed from previous projects did not result in the identification of other domains whose interactions could explain why some ribozymes that 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 the 5th nucleotides of the recognition domain, respectively, creating a longer blocker stem. This may happen relatively frequently (i.e. a 1 out of 4 chance of creating a 1 bp 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 activity of the ribozyme. Some examples of this phenomenon were found in a library created in the course of a study aiming to design SOFA-HDV ribozymes targeting *Hepatitis C Virus* (HCV) RNA strands (Levesque et al., 2010). All of the ribozymes harbouring 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 (i.e. a *cis*-acting HDV ribozyme). In order 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) in order to create SOFA-HDV-Rz-NS527-BI6, while two nucleotides were changed in the -Rz-Rev1 (which possessed an original blocker of 4 bp) in order 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 in order 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 of [α -³²P]UTP in order 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 7 bp blocker that include a GU Wobble that is formed between the 8th nucleotide of the biosensor and the 7th nucleotide of the recognition domain is present, self-cleavage produced two shorter RNA products and there is almost no full length ribozyme present (see Fig. 3, SOFA-

HDV-Rz-NS527-BI7 and -Rev1-BI7). 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_BI6 and -Rz-Rev1_BI6). The two 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 analysed in order to detect and remove any ribozyme that forms an extended blocker stem of 6 base pairs or more. This conclusion is also supported by previously reported data (see Bergeron et al., 2005).



Chapitre 3, Figure 3. Self-cleavage demonstration with 2 different SOFA-HDV Rz.

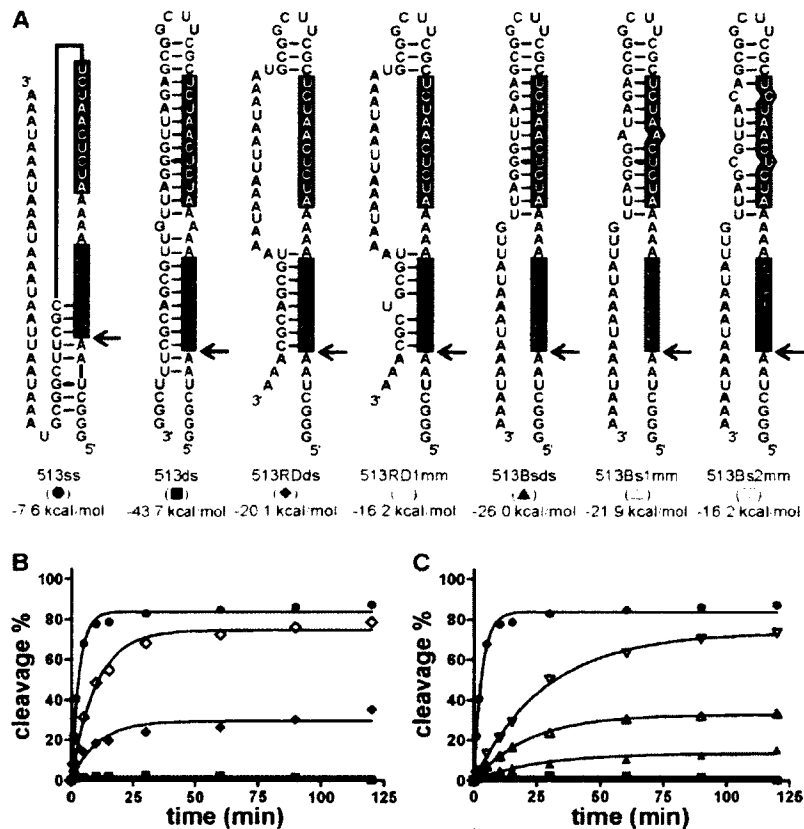
The SOFA-HDV-Rz-NS527 (left) and the SOFA-HDV-Rz-Rev1 (right) are presented on each side of the gel. For both ribozymes, the original sequence (-BI5 or -BI4) and two variants with blockers of either 7 (-BI7) or 6 (-BI6) bp (mutations in

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 (BBP) dyes are indicated adjacent to the gel.

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 unfavourable 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 to 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 (i.e. >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 which folds into a more stable secondary structure (i.e. the 5'-UTR and its internal ribosome entry site motif) (Levesque et al., 2010).

In order to clarify the situation, an experiment was designed which 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 Figure 4A. 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 ΔG 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.



Chapitre 3, Figure 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, while, 513ds (■) is mainly double-stranded. The substrates 513RD (◇) 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.

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, while 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 in order 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 described above, as well as the accumulated data from several studies performed recently aiming to develop a gene-inactivation system based on SOFA-HDV ribozymes (D'Anjou et al., in press; Fiola et al., 2006; Laine et al., 2011; Levesque et al., 2010; Robichaud et al., 2008), identified some of the determinants that appear to be important to consider in 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 must be avoided, and the presence of two consecutive pyrimidines in positions -1 and -2 is not favorable. Moreover, the experiments described above 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 three simple criteria for removing SOFA-HDV ribozymes that have a lesser chance of success prior to any target structure analysis.

In order to validate these simple rules, an experiment designing SOFA-HDV ribozymes targeting the RNA coding for the RNA polymerase HCV (i.e. the NS5B gene, a coding region 1773 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 (i.e. the spacer is the substrate sequence located between the regions bound by the biosensor and the recognition domains, see Figure 1). This spacer size, which was shown to be optimal when between 3 and 7 nucleotides in length, was kept constant in order to restrict the number of variables in the experiment. Subsequently, a simple scoring system that includes two values was developed in order to classify the potential SOFA-HDV ribozymes. First, each possible nucleotides combination in position -2 and -1 was assigned an arbitrary score (i.e. 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 two clusters of SOFA-HDV- ribozymes were synthesized by run-off transcription. Two ribozymes of the lowest scores were showing 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 two ribozymes from the study. In order to exclude the impact of the target structure, short substrates 32 nucleotides long, of which 28 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 since 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 h. The cleavage rates (k_{obs}) and percentages of cleavage after 3 hr are compiled in (Fig.5). A significant difference is observed between the two clusters. All ribozymes from the high score cluster targeted substrates with two adenosines located in positions -1 and -2 (i.e. $A_{-2}A_{-1}/G_{+1}$, 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 one 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 min^{-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 min^{-1} and cleavage that was limited to only 10 % after 3 h, for unknown reasons. In the case of the ribozymes with the lowest scores, the best value observed for k_{obs} was 0.060 min^{-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 min^{-1} , which is 20 fold lower than what was observed for the best ribozymes. All of the SOFA-HDV ribozymes from this cluster targeted substrates with two cytosines located in positions -1 and -2 ($\text{C}_{-2}\text{C}_{-1}/\text{G}_{+1}$), except for SOFA-HDV-Rz-NS5B-266 that had a $\text{C}_{-2}\text{U}_{-1}/\text{G}_{+1}$ 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, 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 in order to establish if 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.

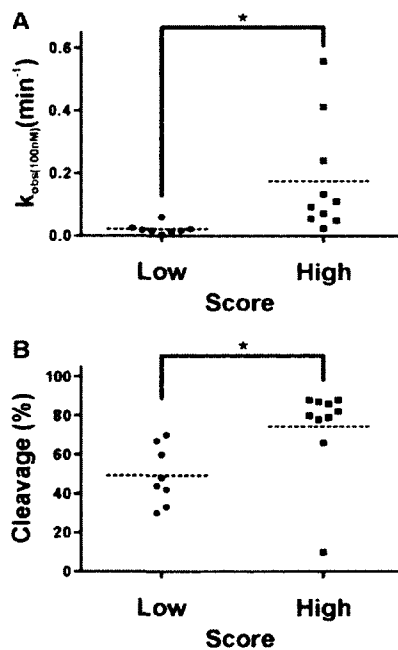
Chapitre 3, Tableau 2. Data from the SOFA-HDV ribozymes targeting HCV NS5B.

Ribozymes	Substrate sequences ^a				Prediction ^b			Cleavage activity ^c	
	Name	Up-stream sequence	RD	Spacer	Biosensor	Scores	MFI (kcal/mol)	BI (bp)	k_{obs} (min ⁻¹)
-NS5B_207	5'-gggaaaUCAA	GGAG AUG	AAGGC	GAAGGCCG UCC-3'	0.6	-9.4	4	0.093 ± 0.008	78
-NS5B_222	5'-gggaaaCGAA	GGCGUCC	ACAGU	UAAGGCCUAAA-3'	2.4	-7.6	4	0.024 ± 0.008	10
-NS5B_237	5'-gggaaaULAA	GGCUAAA	CUUCU	AUCCGUGGAG-3'	1.6	-8.4	5	0.558 ± 0.080	80
-NS5B_270	5'-gggaaaGLAA	GCUGACG	CCCCC	ACAUUCGGCC-3'	0.9	-10.9	5	0.412 ± 0.033	87
-NS5B_342	5'-gggaaaGCAA	GGCCGUU	AACCA	CAUCCGGUCC-3'	0.3	-10.3	4	0.132 ± 0.006	88
-NS5B_372	5'-gggaaaGAA	GGACUUG	CU GGA	AGACACU GAG-3'	2.0	-8.0	4	0.049 ± 0.006	66
-NS5B_465	5'-gggaaaGCAA	GCCAGCU	CGCCU	UAUCGU AUUC-3'	3.2	-6.8	4	0.240 ± 0.021	88
-NS5B_894	5'-gggaaaUGAA	GGCCGCU	GCCGC	CU GUCGAGCU-3'	1.0	-9.0	4	0.109 ± 0.014	79
-NS5B_1265	5'-gggaaaGCAA	GGAU GAL	CCLGA	LGACU CAUU-3'	2.4	-7.6	5	0.055 ± 0.006	82
-NS5B_1592	5'-gggaaaGLAA	GGACCAA	GCUCA	AACUCACUCC-3'	-1.6	-11.6	4	0.071 ± 0.004	86
-NS5B_266	5'-gggaaaGCCL	GU AAGCU	GACCC	CCCCACAUUC-3'	-28.5	-18.5	5	0.060 ± 0.003	70
-NS5B_346	5'-gggaaaGGCC	GUUAACC	ACAUC	CGUCUCCGUGU-3'	-26.1	-11.1	4	0.025 ± 0.001	60
-NS5B_359	5'-gggaaaAUCC	GCUUCGU	GU GGA	AGGACUUGCU-3'	-24.3	-9.3	6	0.022 ± 0.004	44
-NS5B_665	5'-gggaaaACCC	GCLGULL	LGACU	CAACGGLCAC-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	CCCCCULUGC-3'	-34.1	-19.1	5	0.017 ± 0.001	42
-NS5B_1620	5'-gggaaaUCCC	GGCU GCG	LCCCA	GUUGGAUUUA-3'	-27.3	-12.3	5	0.018 ± 0.001	33
-NS5B_1703	5'-gggaaaGCCC	GACCCCG	CLCGU	UC AUGUGGUG-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 S1 III as predicted using the RNAhybrid software (MFI) and the number of base pairs forming the blocker stem (BI) 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.



Chapitre 3, Figure 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 hr, respectively, determined for SOFA-HDV ribozymes of both the low and high score clusters. The dashed lines represent the mean of each group, while the * indicates p values smaller than 0.05.

Discussion

Significant effort has been directed towards optimizing the design processes of the different RNA silencing tools (e.g. siRNA and shRNA) and of the most commonly used catalytic RNAs (e.g. *hammerhead* ribozyme) (Ashihara et al., 2010; Hendry et al., 2004; Laitala-Leinonen, 2010; Levesque et al., 2007; Matveeva et al., 2010; Muhonen and Holthofer, 2010; Vorobjeva et al., 2006; Zhou and Rossi, 2010). This study is the first aiming to gain knowledge about the features that 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 (D'Anjou et al., in press; Fiola et al., 2006; Laine et al., 2011; Levesque et al., 2010; Robichaud et al., 2008). In order 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. 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 of 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 depend 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 one or two 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 two self-cleaving ribozymes, which might help in explaining their low observed cleavage activities (Levesque et al., 2010). This type of situation can

easily be predicted during the design of a SOFA-HDV ribozyme, and, therefore, can be simply avoided.

With respect to the substrate, the situation is more complex with the important parameters being both the sequence and the structure. The experiments illustrated in Figure 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 (i.e. are single-stranded), the cleavage level is higher. However, the binding of both domains appears to act in a cooperative manner, reducing the importance of this requirement as compared what is observed with 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 in order to identify the most accessible sites along a target RNA (Amarzguioui et al., 2000; Bergeron and Perreault, 2002; Marin and Vanicek, 2011; Mercatanti et al., 2002). 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 since 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 upon 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 two consecutive pyrimidines in positions -1 and -2. In addition, the presence of two 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 harbour 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 (1773 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 (i.e. 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 Mazloun 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.

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 antigenomic 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., Routier, S., Perreault, J.P., Latil, A., Bonnel, D., Fournier, I., Salzet, M. and Day, R. in press. Molecular validation of PACE4 as a target in prostate cancer. *Transl Oncol.*

- 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.
- 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 counterstrike. *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., Didierlaurent, 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.
- 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. in press. Target Induced SOFA-HDV Ribozyme. *MiMB*.
- 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., Marangoni, 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., Bisailon, 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.
- Vorobjeva, 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.

Chapitre 4 : Méthode de design pour la production de ribozymes SOFA-HDV

ARTICLE : *Target Induced SOFA-HDV Ribozyme*

Michel V. Lévesque and Jean-Pierre Perreault

Article publié dans : *Methods in Molecular Biology* (2012) 848: 369-384

Avant-propos :

J'ai effectué l'analyse et l'intégration de nos données sur le design des ribozymes pour en faire une méthode. J'ai rédigé en entier ce manuscrit et préparé les figures de l'article sous la supervision de Jean-Pierre Perreault.

Résumé

Les petits ribozymes agissant en *cis* ont été convertis en ribozymes actifs en *trans* ayant la capacité de couper des substrats d'ARN. Parmi eux, le ribozyme du Virus de l'Hépatite Delta («*Hepatitis Delta Virus*», HDV) est un des rares exemples qui est dérivé d'une espèce d'ARN que l'on retrouve dans les cellules humaines. Conséquemment, il possède l'habileté naturelle de fonctionner en présence de protéines humaines et une stabilité exceptionnelle en cellules humaines, deux avantages significatifs pour son utilisation. Le développement d'un module additionnel de spécificité («*specific on/off adaptor*», SOFA) a mené à la production d'une nouvelle génération de ribozyme HDV avec une spécificité améliorée qui fournit un outil avec un potentiel significatif pour un futur développement dans les domaines de la génomique fonctionnelle et de la thérapie génique. Des systèmes d'inactivation génique basés sur le ribozyme SOFA-HDV ont été rapportés autant dans des cellules procaryotes qu'eucaryotes. Ici, une approche «*pas à pas*» pour le design efficace de ribozymes SOFA-HDV hautement spécifiques, avec un investissement minimum de temps et d'effort, est décrite.

Abstract

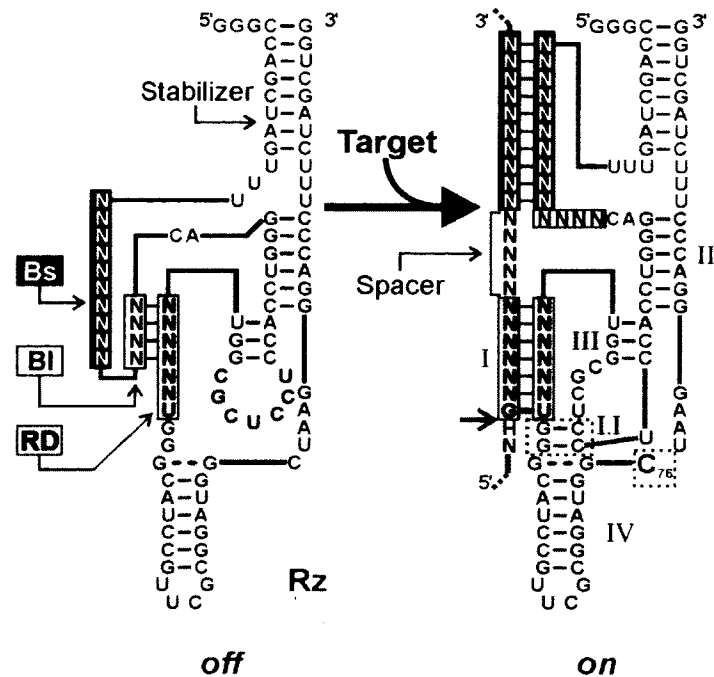
Small cis-acting ribozymes have been converted into trans-acting ribozymes possessing the ability to cleave RNA substrates. The Hepatitis Delta Virus (HDV) ribozyme is one of the rare examples of these that is derived from an RNA species that is found in human cells. Consequently, it possesses the natural ability to function in the presence of human proteins in addition to an outstanding stability in human cells, two significant advantages in its use. The development of an additional specific on/off adaptor (SOFA) has led to the production of a new generation of HDV ribozymes with improved specificities that provide a tool with significant potential for future development in the fields of both functional genomics and gene therapy. SOFA-HDV ribozyme-based gene-inactivation systems have been reported in both prokaryotic and eukaryotic cells. Here, a step by step approach for the efficient design of highly specific SOFA-HDV ribozymes with a minimum investment of time and effort is described.

1. Introduction

Ribozymes (RNA enzyme, Rz), are commonly found as self-cleaving RNA motifs that are essential in the life cycles of infectious RNA that replicate via a rolling circle mechanism. These self-cleaving RNA motifs include both the hammerhead and hairpin structures that are found in the viroid and viroid-like satellite RNAs infecting plants. A self-cleaving ribozyme is also retrieved in the hepatitis delta virus (HDV) that infects several eukaryotes including humans. All of these cis-acting RNA motifs have been converted into trans-acting ribozymes possessing the ability to specifically recognize and subsequently catalyze the cleavage of an RNA target. As a consequence, they have become attractive tools in the development of gene-inactivation systems (for reviews (1-4)).

For a long time, the only example of a ribozyme derived from an RNA species naturally found in human cells (i.e. infected hepatocytes) was the HDV ribozyme (5, 6). Its evolution in human cells confers to it several unique properties for its use as a potential tool, including the natural ability to function both in the presence of human proteins and at the physiological magnesium concentration (i.e. ~1 mM magnesium). Due to its outstanding stability in human cells, the HDV ribozyme is an interesting potential candidate for the development of a gene-knockdown system (7). However, the potential development of such a gene-inactivation system based on the HDV ribozyme has been relatively neglected because it suffers from a lack of substrate specificity when used as a molecular tool (8). Specifically, its substrate specificity depends on the formation of stem I that includes only 7 base pairs (Fig. 1, recognition domain, RD), while a total of 13 to 14 base pairs has been estimated to be required in order to ensure the targeting of a unique RNA species from the human transcriptome (9). In order to overcome this hurdle, a module named SOFA (Specific On/oFf Adaptor) was engineered for the HDV ribozyme (Fig. 1) (10). The SOFA module switches the cleavage activity from the off to the on state solely in the presence of its cognate target. Initially, the recognition domain site forms a short duplex with an inserted sequence element

(the blocker, BI), thereby locking the ribozyme in an inactive conformation (off) by increasing the energetic barrier for non-specific base-pairing interactions and thus reducing the potential for off-target cleavages. A second inserted sequence element (the biosensor, Bs) extends the base-pairing with the target in order to favour the binding of the genuine target, and the formation of this duplex concomitantly results in the disruption of the short duplex involving the blocker and the recognition domain, switching the ribozyme into an active conformation (on). In other words, the blocker acts as a safety lock for the ribozyme in which the key is the recognition of the target RNA by the biosensor. The combined work of the biosensor and the blocker has been shown to increase the substrate specificity of the ribozyme's cleavage by several orders of magnitude when compared with the wild type version (10, 11). Finally, the last component of the SOFA module is a stabilizer stem that brings both the 5' and 3' strand ends together. This additional stem has no effect on the cleavage activity of the SOFA-HDV ribozyme, but data suggests that it is important for its molecular stability in cellulo (7, 10). Importantly, experiments performed in *Lactococcus lactis* confirmed that SOFA-HDV ribozyme retained the property of enzyme turnover, meaning that one molecule of ribozyme can successively cleave several target molecules (12). To our knowledge, the SOFA-HDV ribozyme constitutes the first example of a ribozyme bearing a target-dependent module that is activated by its RNA target, an arrangement which greatly diminishes non-specific effects. This new approach provides a specific and improved tool with significant potential for application in the fields of both functional genomics and gene therapy. The development of several gene-inactivation systems based on SOFA-HDV ribozyme in both prokaryotic and eukaryotic cells have been reported (10, 12-15). These studies have permitted the identification of the important features that must be considered in an optimized design process of SOFA-HDV ribozymes. Here, a simple method for designing SOFA-HDV ribozymes, highlighting the key points that need to be considered in order to improve the rate of success, is described.



Chapitre 4, Figure 1. Detailed representation of the SOFA-HDV ribozyme.

The secondary structures of both the *off* (left) and *on* (right) conformations of the SOFA-HDV ribozyme are shown. The SOFA-HDV ribozyme (Rz) is illustrated in both cases, while the target is only shown for the *on* conformation. The HDV ribozyme is highlighted in grey in each case. The recognition domain, the biosensor and the blocker are identified by RD (black on grey), Bs (white on black) and BI (black on white), respectively. The stabilizer and the spacer (the region in the target that is localized between the recognition domain and the biosensor binding site) are also indicated. The roman numerals I, I.I, II, III and IV identify the corresponding stems or stem-loops in the HDV ribozyme. The variable nucleotides (i.e. those that can be A, C, G or U) of the ribozyme are indicated by the boxed letter N (black for Bs, white for BI and dark grey for RD). All unboxed nucleotides are invariable and are constant from one ribozyme to the other, regardless of the target sequence. Positions where mutations can be inserted in order to produce a catalytically inactive SOFA-HDV ribozyme (I.I pseudoknot and the catalytic cytosine 76) are boxed with dashed lines on the *on* conformation. Finally, the short arrow indicates the cleavage site.

2. Materials

2.1. Specificity analysis

Ribosubstrate: <http://www.riboclub.org/ribosubstrates>

2.2. Synthesis of DNA template

1. SOFA-HDV RzX primer at a concentration of 100 μ M:
5'-TAATACGACTCACTATAGGGCCAGCTAGTTT(N)_{10Bs}(N)_{4Bl}CAGGGTC
CACCTCCTCGCGGT(N)_{6RD}TGGGCATCCGTTTCGCGG-3'.
2. Universal SOFA reverse primer at a concentration of 100 μ M:
5'-CCAGCTAGAAAGGGTCCCTTAGCCATCCGCGAACGGATGCCC-3'.
3. 10X PCR Buffer: 200mM Tris-HCl (pH 8.8), 100 mM KCl, 100 mM
(NH₄)₂SO₄ and 1 % Triton X-100.
4. 100 mM MgSO₄ stock solution.
5. dNTP stock solution: 10 mM of each dATP, dCTP, dGTP and dTTP.
6. Pwo DNA polymerase (Roche Diagnostics).
7. 3 M sodium acetate (pH 5.2).
8. 100 % and 70 % ice-cold ethanol.

2.3. *In vitro* transcription

1. NTP stock solution: 25 mM of each ATP, CTP, GTP and UTP.
2. 5X transcription buffer: 400 mM HEPES-KOH (pH 7.5), 120 mM MgCl₂, 10
mM spermidine and 200 mM DTT.
3. Pyrophosphatase (1 U/ μ L, Roche Diagnostics) diluted 1:100 (V:V) in a
dilution buffer containing 50 % glycerol, 20 mM Tris-HCl (pH 8.0), 10 mM
NaCl, 1 mM DTT, 1 mM EDTA, 100 μ g/mL BSA and 0.03 % NP40.
4. RNaseOUT (40 U/ μ L, Invitrogen) diluted 1:2 (V:V) in a buffer containing
50 % glycerol, 20 mM Tris-HCl (pH 8.0), 50 mM KCl, 0.5 mM EDTA and 8
mM DTT.
5. Purified T7 RNA polymerase (5 μ g/ μ L).
6. RQ1 RNase-free DNase (1 U/ μ L, Promega).
7. Phenol and chloroform.

8. Loading buffer: 98 % formamide, 10 mM EDTA, 0.025 % xylene cyanol and 0.025 % bromophenol blue.
9. Elution buffer: 500 mM ammonium acetate, 1 mM EDTA and 0.1 % SDS.

2.4. Target labeling

1. 10X Antarctic Phosphatase Reaction Buffer (New England Biolabs).
2. Antarctic Phosphatase (5 U/ μ L, New England Biolabs).
3. 10X T4 polynucleotide kinase (PNK) (USB).
4. T4 PNK (30 U/ μ L, USB) diluted 1/10 in T4 PNK dilution buffer (USB).
5. [γ -³²P]ATP (6000 Ci/mmol, PerkinElmer).
6. Glycogen (20 mg/mL, Roche Diagnostics).

2.5. Cleavage assays

1. 500 mM Tris-HCl (pH 7.5).
2. 100 mM MgCl₂.

2.6. Denaturing polyacrylamide gel electrophoresis

1. Denaturing gels: 8 M urea and 19:1 ratio of acrylamide/bisacrylamide in 1X TBE buffer (89 mM Tris base, 89 mM boric acid and 2 mM EDTA).

2.7. Autoradiography and RNA quantification

1. Hyperfilm MP films (GE Healthcare).
2. Phosphoscreen (GE Healthcare).
3. ImageQuant software (GE Healthcare).

3. Methods

The step-by-step protocol illustrated in Fig. 2 provides for a fast and simple selection of SOFA-HDV ribozymes directed against any target of choice. This protocol is composed of three main modules that are then subdivided into various steps. Briefly, the modules include, successively, the identification of the potential targeting sites, the design of the SOFA-HDV ribozymes and, lastly, the selection of the most prominent ones for further development.

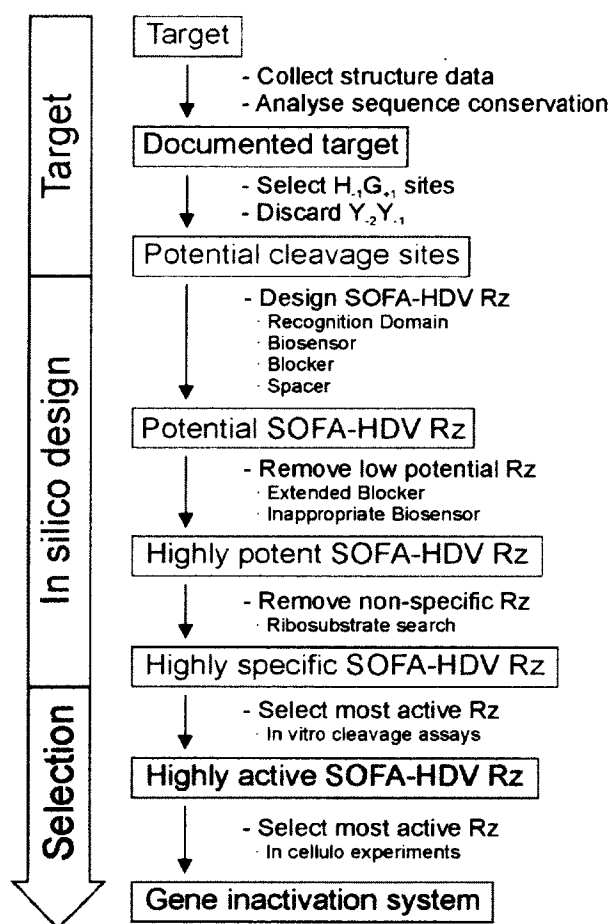
3.1. Target module

The first step is solely related to the selected RNA target. It includes all of the prerequisite analyses and considerations that must be taken into account before moving on with the designing of SOFA-HDV ribozymes. This step can be subdivided into two parts: the documentation of the target and the identification of potential sites to be targeted.

3.1.1. Documentation of the target

Initially it is paramount to both gather all available information on the target and to establish all considerations that might influence the targeting of a given RNA species. For example, the targeting of a viral RNA with the aim of controlling the propagation of this pathogen may benefit from a sequence analysis in order to identify any stretches of highly conserved nucleotides that could then be used for ribozyme binding. This consideration increases the chances of targeting a wide range of viral genotypes. In addition, targeting more conserved nucleotides should reduce the potential escape of any resistant mutants (16, 17). Conversely, the identification of sequences specific to a unique mRNA belonging to a family of

relatively well conserved proteins may lead to a reduction in the level of a single protein species even though several related ones are similarly expressed in a cell.



Chapitre 4, Figure 2. Scheme of the design process for the SOFA-HDV ribozyme.

The large arrows represent the three major, sequential, design steps. The target (white box), the in silico design (light grey box) and the selection (dark grey box) modules are divided into different analyses that are described in section 3. The horizontal boxes indicate what is obtained after each procedure, while the text on the right side of each small arrow describes the key element in the process.

Effective cleavage at a specific site depends on its accessibility to the ribozyme. In principle, target sites located in single-stranded regions of an RNA species possess a higher potential because they should be more accessible for ribozyme binding than those located in double-stranded regions (18-20). Within the double-stranded regions, the ribozyme might compete unfavourably with intramolecular base pairing when trying to bind the target. Therefore, any available data on the secondary and tertiary structures of the target might be of great value. In the past, bioinformatic and biochemical procedures, as well as a combination of both, have been used (18, 20). However, most of these approaches remain exhausting and may not necessarily be relevant if they are performed in test tube, as many factors influence the structure of an RNA species in vivo. Therefore, when no structural data is available, many years of work has led us to suggest limiting these investigations to only bioinformatic analysis and instead compensating by designing and testing a larger number of ribozymes in the subsequent steps. Moreover, the biosensor has been demonstrated to act as a facilitator (i.e. unwinds the secondary structure in the neighbourhood of the target site after its binding), thereby reducing the importance of the accessibility of the target sites (10).

3.1.2. Identification of potential sites

The resulting sequences (i.e. conserved, specific and mostly accessible) are then further analyzed in order to identify those that fulfill the essential criteria for efficient HDV ribozyme cleavage. The minimum requirement for an RNA to be cleaved by an HDV-derived ribozyme is the presence of a H-1G+1 at the cleavage site, where H can be A, C or U (8). More specifically, the first nucleotide downstream of the cleavage site (position +1) must be a guanosine in order to allow for formation of the essential GU Wobble base pair with the ribozyme. At position -1, the presence of a guanosine residue is detrimental to cleavage. A systematic analysis of the sequences in positions -1 to -4 of a collection of small targets revealed that each of

these nucleotides contributes differently to the ability of a given target RNA to be cleaved (21). Further analysis using longer RNA targets indicates that the nucleotides located in positions -1 and -2 may significantly influence the cleavage level. Clearly, it is of interest to select the more favourable ones when identifying potential sites in order to improve the chances of success (14). Specifically, the presence of two consecutive pyrimidines in these positions (i.e -1 and -2) must be avoided as it is detrimental for the cleavage. Based upon these analyses, a list of all potential target sites is generated, and the sequences surrounding the cleavage sites from positions -2 up to at least +27, which includes all of the features important for designing a ribozyme, are extracted and used in the next step.

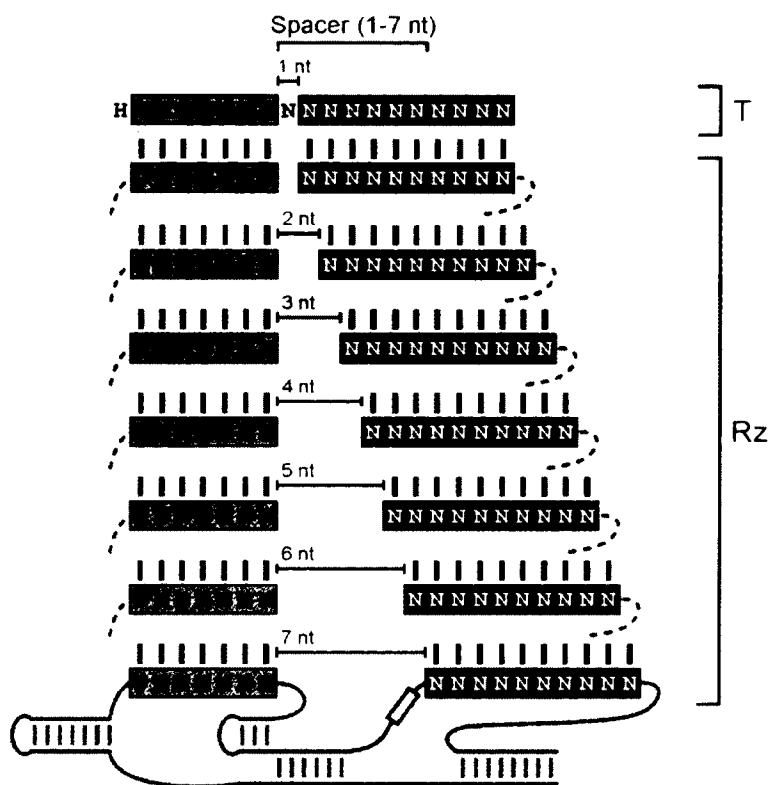
3.2. Ribozyme design module

The second module is directed towards both the design and the pre-selection of ribozymes *in silico*.

3.2.1. Designing SOFA HDV ribozymes

Over the years, the sequence of the SOFA-HDV ribozyme has been optimized in order to favour a high cleavage level (10, 13). Although the backbone of this RNA molecule is largely conserved, there are three domains that are considered to be variable in terms of the designing of ribozymes: the recognition domain, the biosensor domain and the blocker domain. The recognition domain is complementary to nucleotides +1 to +7 of the target and includes the uridine residue that forms a G-U Wobble base pair (Fig. 1). In other words, the recognition domain is determined by the cleavage site ($H_{-1}G_{+1}$). Subsequently, the sequence composing the biosensor, which interacts with the target downstream of the recognition domain, must be determined (Fig. 3). The regions of the target interacting with both the recognition domain and biosensor are separated by a

spacer sequence that was included in order to avoid the stacking of both domains, as they are in close proximity to one another. This stacking should it occur, could be detrimental to the product release (Fig. 1 and 3). Previous experiments demonstrated that it is preferable to have a spacer of at least 1 nucleotide, but no longer than 7. The optimal length is 3 nucleotides, but anything between 1 and 5 nucleotides works well (Fig. 3)(13). Therefore, a 3-nucleotide spacer is generally used in the initial design; and consequently, the next 10 nucleotides of the target can be considered as being the region bound by its complementary biosensor. A length of 10 nucleotides for the biosensor has been shown to be optimal for efficient binding without preventing product release and thus conserving the turnover property of the ribozyme. Finally, the sequence of the blocker is determined. It consists of 4 nucleotides that are complementary to the 5' part of the ribozyme's recognition domain (Fig. 1) (see Notes 1 and 2 for details).



Chapitre 4, Figure 3. Details on spacer flexibility and biosensor design.

Representation of a potential cleavage site in a target (T), and of the corresponding SOFA-HDV ribozymes (Rz). The length in nucleotides (nt) of each spacer is indicated. The corresponding recognition and biosensor domains are shown below the target RNA. The optimal ribozyme with a spacer of 3 nt is highlighted in a light grey box. The sequences of the recognition domain, and of the biosensor binding sites on both the target and the ribozyme, are identified using the same color code as in Fig. 1 (i.e. the recognition domain is in black letters within grey boxes and the biosensor is in white letters in black boxes).

3.2.2. Removal of low potential SOFA-HDV ribozymes

The various series of SOFA-HDV ribozymes developed in the past have led to the elucidation of guidelines for the pre-selection of the ribozymes possessing the greatest potential prior to the investment of any time in laboratory experiments.

The length of the blocker sequence that base pairs with the recognition domain has been shown to affect the cleavage activity ((13) and unpublished data). The formation of more than 4 base pairs with the recognition domain may

significantly reduce the cleavage activity by either preventing the switch from the off to the on conformation, or by generating a cis-acting ribozyme possessing the ability to self-cleave and therefore reduce the quantity of active SOFA-HDV ribozyme. The additional nucleotides that might be involved in these extra base pairs are in fact those located at the 3' end of the ribozyme's biosensor. Therefore, it is important to analyse the biosensor's sequence in terms of the possibility of finding such base-pairing. When it occurs, the simplest solution is to consider slightly displacing the biosensor's binding sequence by repeating the previous step. If no solution appears to be possible, the given SOFA-HDV ribozyme must be removed from the collection.

Previous experiments have shown that significant complementarity between either the recognition domain, or the biosensor, with the sequence composing stem-loop III may also impair the cleavage activity of a given ribozyme. Specifically, whenever the possibility exists of forming at least 6 consecutive base pairs, or of any stability equivalent (e.g. 7 base pairs including a bulge), it is preferable to either remove or redesign the SOFA-HDV ribozyme in question because it will most likely be inactive (unpublished data).

3.2.3. Specificity analysis

When working towards the development of a gene-inactivation system, a key question to ask is how to design target-specific ribozymes that do not produce any side effects. There are several potential sources of these side effects including, for example, the triggering of immunological responses. Moreover, in many cases, the side effects may result from the cleavage of non-desired mRNA species (i.e. off-target effects). In order to evaluate this possibility, the Ribosubstrates, (<http://www.riboclub.org/ribosubstrates>) software was developed. This integrated software searches in selected cDNA databases for all potential targets of a given SOFA-HDV ribozyme (22). These potential targets include not only mRNAs having

perfect matches with the catalytic RNA in question, but also for those interactions that include Wobble base pairs and/or mismatches. The results generated permit a rapid selection of the sequences suitable as targets for SOFA-HDV ribozymes. We suggest removing from the list, at a minimum, all ribozymes that would potentially recognize another target. Further selection criteria may also take into consideration the occurrence of any mismatches between the target and the ribozyme's binding domain. The presence of only one mismatch in the recognition domain is significantly more detrimental than the one occurring in the biosensor.

3.3. Ribozyme screening module

In this final module, an experimental design using *in vitro* cleavage assays for selecting the best candidate, is presented. In addition, some important elements to be taken into consideration while using SOFA-HDV ribozyme in cellulo are discussed.

3.3.1. In vitro cleavage assay

3.3.1.1. Preparations of DNA templates

Most of the time, SOFA-HDV ribozymes are produced by *in vitro* transcription from PCR generated DNA templates as described previously (14). Briefly, the PCR-based strategy includes two complementary and overlapping oligonucleotides. The sense primer, namely the SOFA-HDV RzX primer (where the X corresponds to the identification of the specific ribozyme, that is to say the cleavage site's position), is specific for each ribozyme and is composed of a T7 promoter followed by the SOFA-HDV ribozyme's sequence, stopping after loop IV (5'-TAATACGACTCACTA TAGGGCCAGCTAGTTT(N)_{10BS}(N)_{4BI}CAGGGTCCACCTCCTCGCGGT(N)_{6RD}TGG GCATCCGTTTCGCGG-3', where N represents A, C, G or T and Bs, BI and RD indicate the biosensor, the blocker and the recognition domain, respectively). The

reverse primer, specifically the Universal SOFA reverse primer, is universal to all ribozymes (5'-CCAGCTAGAAAGGGTCCCTTAGCCATCCGCGAACGGATGCCC-3'). The underlined nucleotides represent the overlapping sequences between the primers. Those two DNA primers are then used in a filling PCR reaction in order to produce the template for the run off transcription. The target's DNA template is obtained either from a digested plasmid, or through a PCR strategy as described below (see Note 3).

1. Mix 2 μL of both the SOFA-HDV RzX and the SOFA reverse primers (100 μM each) with 10 μL of 10X PCR buffer, 2 μL of 100 mM MgSO_4 , 2 μL of 10 mM dNTP, 81.5 μL of sterilized deionised water and 0.5 μL of Pwo DNA polymerase (2.5 U) in a final reaction volume of 100 μL .
2. Run the PCR for 12 cycles (1 cycle = 45 sec at 95°C, 45 sec at 55°C and 45 sec at 72°C).
3. Transfer the reaction to a tube containing 10 μL of sodium acetate. Add 2,5 volumes of ice-cold 100 % ethanol and centrifuge at 16 200x g for 25 min. Remove the supernatant and wash the pellet by adding 150 μL of 70 % ethanol and centrifugating at 16 200x g for 5 min.
4. After removing the supernatant, quickly dry the pellet and then dissolve it in 52 μL of RNase free water.

3.3.1.2. Preparation of RNA ribozymes and targets

Both the targets and the SOFA-HDV ribozymes are produced by run-off transcription as described below.

1. To the 52 μL of DNA template from above, add 24 μL of 25 mM NTP, 20 μL of 5X transcription buffer, 1 μL of diluted pyrophosphatase (0.01 U), 1

μL of diluted RNaseOUT (20 U) and 2 μL of purified T7 RNA polymerase (10 μg) in a final reaction volume of 100 μL .

2. Incubate at 37°C for 2 h.
3. To each reaction, add 3 μL of DNase RQ1 (3U) and incubate for 20 min at 37°C.
4. Add 0.5 volumes of both phenol and chloroform, vortex and centrifuge at 16 200x g for 10 min. Transfer the upper aqueous phase to a new tube containing 10 μL of sodium acetate (3M, pH 5.2). Add 2,5 volumes of ice-cold 100 % ethanol, vortex and centrifuge at 16 200x g for 25 min. Discard the supernatant and wash the pellet by adding 150 μL of 70 % ethanol followed by centrifugation at 16 200x g for 5 min.
5. After discarding the supernatant, quickly dry the pellet.
6. Dissolve the pellet in 40 μL of RNase free water. Add 80 μL of loading buffer and fractionate the RNA on a standard 8 M urea denaturing polyacrylamide gel (see Note 4). Visualize the RNA by UV shadowing (see Note 5) and excise the band. Transfer the band to a 1,5 mL tube and add 500 μL of elution buffer. Elute the RNA over night at 4°C on a rotating shaker.
7. Transfer the eluate to a fresh 2 mL tube containing 50 μL of sodium acetate (3M, pH 5.2). Add 2,5 volumes of ice-cold 100 % ethanol and centrifuge at 16 200x g for 25 min. Discard the supernatant and wash the pellet by adding 250 μL of 70 % ethanol followed by centrifugation at 16 200x g for 5 min.
8. Discard the supernatant and quickly dry the pellet.
9. Dissolve the pellet in RNase free water and quantify the RNA by absorbance at 260 nm.

3.3.1.3. RNA target labelling

Target RNA could either be transcribed in the presence of radiolabeled nucleotides ($[\alpha\text{-}^{32}\text{P}]\text{UTP}$), or be labelled at either the 5' or the 3'-end in the presence of either $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ or $[\alpha\text{-}^{32}\text{P}]\text{Cp}$, respectively, and then be further purified by denaturing PAGE (10, 14, 23). Generally, the 5'-end labeling protocol described below is used (see also Note 6).

1. Mix 5 μL of 5 μM *in vitro* transcribed target (25 pmol) with 3 μL of RNase free water, 1 μL of 10X Antarctic Phosphatase Reaction Buffer and 1 μL of Antarctic Phosphatase (5 U).
2. Incubate the 10 μL dephosphorylation reaction at 37°C for 30 min. To stop the reaction, incubate the reaction at 65°C for 7 min, then store it on ice for 5 min.
3. Transfer 2 μL (5 pmol of dephosphorylated RNA) of the previous reaction to a new tube containing 14 μL of RNase free water and 2 μL of 10X T4 Polynucleotide Kinase (PNK) Reaction Buffer. Add 1 μL of $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ and 1 μL of diluted T4 PNK (3 U).
4. Incubate the 20 μL reaction at 37°C for 1 h. Add 30 μL of loading buffer and gel purify the 5'-end radiolabeled RNA as described in section 3.3.1.2. The RNA is visualised by autoradiography, cut out of the gel and eluted overnight with 500 μL of elution buffer at 4°C on a rotating shaker.
5. Add 1 μL of glycogen to the eluate followed by 50 μL of sodium acetate (3M, pH 5.2) and 2,5 volumes of 100 % ethanol. Centrifuge at 16 200x g for 25 min and discard the supernatant. Wash the pellet with 150 μL of 70 % ethanol followed by centrifugation at 16 200x g for 5 min and finally dissolve it in 500 μL of RNase free water for a final concentration of less than 10 nM.

3.3.1.4. Cleavage assays

Usually the cleavage reactions are carried out under single turnover conditions in which the ribozyme concentration greatly exceeds that of the target ($[\text{SOFA-HDV Rz}] \gg [\text{S}]$) (10, 14). The following detailed procedure describes a 10 μL end point cleavage reaction.

1. Mix 6 μL of RNase free water with 1 μL of both 500 mM Tris-HCl (pH 7.5) and 100 mM MgCl_2 prior to the addition of 1 μL of radiolabeled target from section 3.3.1.3 (see Note 7).
2. Start the reaction by adding 1 μL of 1 μM of SOFA-HDV ribozyme and then incubating at 37°C for 2 h.
3. Stop the reaction by adding 20 μL of loading buffer.
4. The completed reactions are fractionated on a 8 M urea denaturing PAGE gel of the appropriate concentration, and are then exposed to a Phosphoscreen.
5. The activity of each SOFA-HDV ribozyme is expressed as a percentage of the cleaved products over the total quantity of target ($100 \times \frac{\text{cleaved}}{\text{cleaved} + \text{uncleaved}}$) as determined using the ImageQuant (Molecular Dynamics) software.

Based on our experience, SOFA-HDV ribozymes that exhibit cleavage percentages over 75 % under the conditions described above are considered as being active, and possess the greatest potential for further experiments. Conversely, SOFA-HDV ribozymes that exhibit cleavage activities below 20 % are not suitable for targeting experiments. All of the SOFA-HDV ribozymes that exhibit median levels of cleavage activity could give good results in cellulo, and may therefore be considered depending on the desired number of ribozymes to test.

3.3.2. *In celluloin vivo* assays

The next step in the establishment of a fully functional gene inactivation system is to test the selected SOFA-HDV ribozymes either in *cellulo* or *in vivo*. One important point to mention is that each case needs to be analyzed individually in order to design a suitable experiment. Depending on both the target and the cellular model, two major aspects should be considered: the appropriate controls and the expression system.

Firstly, two types of controls are recommended in all experiments: a catalytically inactive ribozyme and an irrelevant, active ribozyme. The catalytically inactive ribozyme differentiates between any possible antisense effects and the catalytic cleavage of the target. Such a ribozyme can be produced by mutating either the catalytic cytosine to a guanosine (C76G), and/or the residues of the two GC base pairs that form the I.I pseudoknot to four uridines (Fig. 1). The resulting SOFA-HDV ribozyme has the ability to bind the target, but does not exhibit cleavage. The other control is an active SOFA-HDV ribozyme that is irrelevant to the target, that is to say the sequence of the recognition domain and/or the biosensor are not complementary to the target. This control allows the measurement of the overall impact of the presence of a SOFA-HDV ribozyme in the cell without cleavage of the target. This control ribozyme should be analysed with Ribosubstrate to confirm its irrelevancy to not only the target, but also to any important cellular genes. It is crucial that this ribozyme has a minimal impact on the overall life of the cell.

Finally, the selection of the expression system is crucial for the establishment of a gene inactivation system. According to the model of choice, either DNA or RNA transfection, as well as transduction with lentivirus encoding the ribozymes, could be used in *in celluloin vivo* experiments. When expressing SOFA-HDV ribozymes from a DNA template, different options are available. First, the type

of promoter has a significant impact on the results. A tissue specific RNA polymerase II promoter will favor expression in a specific cell type, while an RNA polymerase III promoter should result in a more efficient transcription that yields a greater amount of SOFA-HDV ribozyme. The cellular localization of the target entails a specific transport of the ribozyme inside the cells. The direct use of any polymerase III promoter without any additional sequence is a good way to produce SOFA-HDV ribozymes that are restricted to the cell's nucleus. For cytoplasmic export of the ribozyme, a 5' capped and 3' polyadenylated mRNA-like expression driven by the RNA polymerase II is a good option. It is also possible to use hybrid ribozymes consisting of a fusion between SOFA-HDV ribozyme and an RNA motif that permits an active transport into the cytoplasm. A good example of this is the use of tRNA^{val} as a leader sequence for SOFA-HDV ribozyme (14). Data from our laboratory suggest that hY RNA are also good candidates for driving the cellular localization of ribozymes (unpublished data, J. Perreault and J.P. Perreault) (see Note 8). Finally a cocktail of ribozymes targeting two or three different sites may also be used in conjunction to knockdown a gene during in cellulo assays.

In summary, no universal protocol governs the details of an in cellulo experiment with a ribozyme. Each unique context possesses its own constraints that define the optimal strategy.

4. Notes

1. The biosensor can be elongated up to 12 nucleotides in order to increase the binding affinity without too significant effect on the ribozyme's turnover.
2. The stabilizer can be changed if undesired interactions can occur between the target and the stabilizer. As this part of the SOFA module only plays a structural role, the only limitation is to avoid base-pairing with other parts of both the ribozyme and the target.

3. The target's DNA template must include a T7 RNA polymerase promoter upstream of the targeted sequence. If the template is a linearized plasmid, about 5 μg of DNA should be used in a 100 μL run-off transcription reaction.
4. For SOFA-HDV ribozyme purification, in general 8 % PAGE is used and the electrophoresis is halted when the xylene cyanol reaches the bottom of the gel. For targets, the concentration of the gel needs to be adapted to their length.
5. In order to visualize RNA using UV shadowing, simply put the gel (covered with laboratory plastic wrap) on a TLC plate (silica gel, Whatman) and expose it to 254 nm UV light.
6. End labelling gives results that are easier to quantify as only two different products need to be quantified on the gel, and the calculation is therefore more accurate. 5'-end labelling is usually favoured; however, when the target cannot be efficiently labelled in 5', 3'-end labelling involving the ligation of $[\alpha\text{-}^{32}\text{P}]\text{Cp}$ is performed. Alternatively, internal labelling by direct incorporation of $[\alpha\text{-}^{32}\text{P}]\text{UTP}$ during the transcription can be performed.
7. When testing a series of SOFA-HDV ribozymes targeting one RNA, a master mix of water, radiolabeled target, Tris-HCl and MgCl_2 is prepared in order to favour greater uniformity between the reactions.
8. The use of an RNA polymerase III promoter for the expression of SOFA-HDV ribozymes limits in the variety of sequences that can be used. All ribozymes that contain four or more consecutive uridines should be discarded because this number of uridines forms a transcription termination signal. The three uridines of the linker that are located between the stabilizer and the biosensor increase the possibility of creating a

transcription terminator. In this particular case, the linker could be mutated in order to avoid having four consecutive uridines present.

Acknowledgments

We thank Jonathan Perreault and Gilles Boire for information about hY RNA. This work was supported by grants from Canadian Institute of Health Research (CIHR; grant numbers MOP-44002 and EOP-38322) to J.P.P. The RNA group is supported by grants both from CIHR and Université de Sherbrooke. M.V.L. was the recipient of a predoctoral fellowship from the Fonds de Recherche en Santé du Québec. J.P.P. holds the Canada Research Chairs in genomics and catalytic RNAs, and is a member of the Centre de Recherche Clinique Étienne Lebel.

References

1. Bagheri, S., and Kashani-Sabet, M. (2004) Ribozymes in the Age of Molecular Therapeutics. *Curr. Mol. Med.* 4, 489-506.
2. Schubert, S., and Kurreck, J. (2004) Ribozyme- and Deoxyribozyme-Strategies for Medical Applications. *Curr. Drug Targets.* 5, 667-681.
3. 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.
4. Tedeschi, L., Lande, C., Cecchettini, A., and Citti, L. (2009) Hammerhead Ribozymes in Therapeutic Target Discovery and Validation. *Drug Discov. Today.* 14, 776-783.
5. Teixeira, A., Tahiri-Alaoui, A., West, S., Thomas, B., Ramadass, A., Martianov, I., Dye, M., James, W., Proudfoot, N. J., and Akoulitchev, A. (2004) Autocatalytic RNA Cleavage in the Human Beta-Globin Pre-mRNA Promotes Transcription Termination. *Nature.* 432, 526-530.
6. Salehi-Ashtiani, K., Luptak, A., Litovchick, A., and Szostak, J. W. (2006) A Genomewide Search for Ribozymes Reveals an HDV-Like Sequence in the Human CPEB3 Gene. *Science.* 313, 1788-1792.

7. 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.
8. Bergeron, L. J., Ouellet, J., and Perreault, J. P. (2003) Ribozyme-Based Gene-Inactivation Systems Require a Fine Comprehension of their Substrate Specificities; the Case of Delta Ribozyme. *Curr. Med. Chem.* 10, 2589-2597.
9. Peracchi, A. (2004) Prospects for Antiviral Ribozymes and Deoxyribozymes. *Rev. Med. Virol.* 14, 47-64.
10. Bergeron, L. J., and Perreault, J. P. (2005) Target-Dependent on/off Switch Increases Ribozyme Fidelity. *Nucleic Acids Res.* 33, 1240-1248.
11. Bergeron, L. J., Reymond, C., and Perreault, J. P. (2005) Functional Characterization of the SOFA Delta Ribozyme. *RNA*. 11, 1858-1868.
12. 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.
13. Bergeron, L. J., Reymond, C., and Perreault, J. P. (2005) Functional Characterization of the SOFA Delta Ribozyme. *RNA*. 11, 1858-1868.
14. 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.
15. 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.
16. von Eije, K. J., ter Brake, O., and Berkhout, B. (2008) Human Immunodeficiency Virus Type 1 Escape is Restricted when Conserved Genome Sequences are Targeted by RNA Interference. *J. Virol.* 82, 2895-2903.
17. Haasnoot, J., Westerhout, E. M., and Berkhout, B. (2007) RNA Interference Against Viruses: Strike and Counterstrike. *Nat. Biotechnol.* 25, 1435-1443.
18. 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.
19. Ryu, K. J., and Lee, S. W. (2004) Comparative Analysis of Intracellular Trans-Splicing Ribozyme Activity Against Hepatitis C Virus Internal Ribosome Entry Site. *J. Microbiol.* 42, 361-364.

20. Doran, G., and Sohail, M. (2006) Systematic Analysis of the Role of Target Site Accessibility in the Activity of DNA Enzymes. *J. RNAi Gene Silencing*. 2, 205-214.
21. 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.
22. Lucier, J. F., Bergeron, L. J., Briere, F. P., Ouellette, R., Elela, S. A., and Perreault, J. P. (2006) RiboSubstrates: A Web Application Addressing the Cleavage Specificities of Ribozymes in Designated Genomes. *BMC Bioinformatics*. 7, 480.
23. Ouellet, J., and Perreault, J. P. (2004) Cross-Linking Experiments Reveal the Presence of Novel Structural Features between a Hepatitis Delta Virus Ribozyme and its Substrate. *RNA*. 10, 1059-1072.

DISCUSSION

1. Applications du ribozyme SOFA-HDV

La conception n'est que l'étape initiale du long processus de développement d'une nouvelle technologie. Après la réalisation d'une preuve de concept, c'est l'application de cette technologie qui permet d'en établir la crédibilité et le potentiel. Mes travaux de recherche sur le ribozyme SOFA-HDV portaient sur cette étape. Le ciblage du VHC à l'aide de ribozymes SOFA-HDV représente le premier exemple d'application de cette technologie dans un modèle de cellules de mammifères. D'autre part, la collaboration avec le laboratoire de la Dre Gatignol, a permis d'identifier un ribozyme ayant un fort potentiel comme agent antiviral contre le VIH. À cela s'ajoute également plusieurs autres projets de ciblage d'ARN de natures diverses qui ont été réalisés pendant mes études dans le laboratoire ou par l'entremise de collaborations externes. Ce grand nombre de ribozymes testés fournit une source impressionnante d'information sur l'activité du ribozyme SOFA-HDV et des facteurs qui l'influencent.

1.1. Cibler les ARN du VHC

Tout d'abord, mes travaux initiaux d'application du ribozyme SOFA-HDV visaient l'inhibition de la réplication du VHC. Le développement de ces ribozymes pour cibler l'ARN de polarité positive (ARN brin (+)) et l'ARN de polarité négative (ARN brin (-)) du VHC a présenté plusieurs défis. Ces efforts nous ont permis de produire la première étude testant un grand nombre de ribozymes SOFA-HDV dirigés contre une cible d'intérêt. Ces 31 ribozymes testés *in vitro* (17 ARN brin (+) et 14 ARN brin (-)), dont 26 ribozymes testés *in cellulo* (12 ARN brin (+) et 14 ARN brin (-)), sont significativement plus nombreux que les études précédentes, limitées à un maximum de 5 ribozymes *in vitro* et 4 ribozymes *in cellulo* (Bergeron et

Perreault, 2005; Fiola et al., 2006; Robichaud et al., 2008). Au moment d'écrire cette thèse, une seule étude a surpassé ce sommet avec 32 sites de coupure, pour plus de 40 ribozymes SOFA-HDV testés *in vitro* et un nombre comparable de ribozymes SOFA-HDV testés *in cellulo* (Motard et al., 2011). Au-delà du grand nombre de ribozymes testés contre les ARN du VHC, ces travaux ont identifié deux ribozymes SOFA-HDV parmi les plus actifs contre un long ARN *in vitro*. Les ribozymes SOFA-HDV-Rz293⁻ et SOFA-HDV-Rz304⁻ avaient la plus haute activité de coupure observée (85 % et 70 % de coupure après trois heures d'incubation, respectivement). À ma connaissance, seulement le ribozyme SOFA-HDV-RzTev1 a atteint un tel niveau de coupure *in vitro* sur un long substrat (>100 nt de long) (Laine et al., 2011).

L'un des sous-objectifs du projet était de valider la méthode de design qui avait été mise au point pour la production de ribozymes HDV conventionnels (Bergeron et Perreault, 2002). Elle proposait une combinaison d'analyses bioinformatiques et d'essais RNase H pour identifier les sites de coupures les plus accessibles pour le ribozyme HDV. Cette technique a permis de sélectionner huit sites potentiels de coupure dans le 5' UTR de l'ARN(+) du VHC (Chapitre 1, Tableau 1). Après l'analyse de l'activité de coupure des ribozymes SOFA-HDV correspondants, seulement quatre ont démontré une activité de coupure supérieure à 10 % (Chapitre 1, Tableau 2). En comparaison, trois des huit sites sélectionnés sur la base de la conservation de la séquence et d'une évaluation visuelle de l'accessibilité basée sur la structure connue du 5' UTR de l'ARN brin (+) du VHC répondaient à ce même critère d'activité. Il faut aussi noter que dans le groupe des huit premiers sites identifiés, seulement deux ribozymes dépassaient les 20 % de coupure alors que dans l'autre groupe, ce nombre s'élevait à trois. La prédiction de l'accessibilité du site de liaison du domaine de reconnaissance (RD, identifié P1 dans l'article) du ribozyme HDV ne semble donc pas améliorer significativement le rendement *in vitro* des ribozymes SOFA-HDV subséquents. Des résultats semblables ont été obtenus avec des ribozymes SOFA-HDV dirigés

contre l'ARNm de la protéine précurseur de l'amyloïde («*amyloid protein precursor*», APP) (Ben Aissa et al., 2012). Cette observation ne signifie pas que l'accessibilité ne joue aucun rôle dans l'activité de coupure des ribozymes SOFA-HDV *in vitro* et *in cellulo*. Au contraire, nos résultats tendent à confirmer un lien entre l'activité de coupure observée *in vitro* et la structure secondaire de l'ARN ciblé. L'aspect de l'accessibilité en lien avec l'activité de coupure des ribozymes SOFA-HDV sera discuté plus en détails dans la section 2.1.3.

Après l'identification de ribozymes SOFA-HDV aptes à couper efficacement les ARN du VHC *in vitro*, nous nous sommes attaqués à l'évaluation de ces ribozymes dans un modèle cellulaire de réplication du VHC. Pour ce faire, nous avons utilisé un vecteur lentiviral permettant l'expression de ribozymes hybrides tRNA^{val}-SOFA-HDV. Cette stratégie a été jugée essentielle pour permettre la production d'une grande quantité de ribozymes dans le cytoplasme des cellules, étant donné la localisation exclusivement cytoplasmique du VHC. Malheureusement, les effets d'inhibition observés *in cellulo* sont demeurés modestes et ceci pour les rares cas où un effet a pu être mesuré. Au mieux, cet effet atteignait les 30 à 40 % de réduction pour les quatre ribozymes. Ceci est une faible baisse de l'activité luciférase mesurée, surtout dans un contexte où nous avons jugé non significatif toutes variations de ± 20 %. Ces résultats sous optimaux pour le développement d'une thérapie visant le VHC sont tout de même en accord sur certains points avec la littérature.

D'abord, il est intéressant de noter qu'aucun de nos ribozymes visant l'ARN brin (-) n'a démontré d'activité significative sur la réplication du VHC. Pourtant, cette stratégie semblerait avantageuse en raison du ratio inégal entre la quantité d'ARN brin (+) et d'ARN brin (-) lors de la réplication du virus, à la fois dans les modèles de cellules en culture (Komurian-Pradel et al., 2004) et dans les hépatocytes de patients infectés (Yuki et al., 2006). Nous pouvons supposer que le ciblage de l'ARN brin (-), présent en plus faible quantité, représente la meilleure approche pour perturber la réplication du virus, mais nos résultats semblent

contredire cette idée. Une explication de l'absence d'activité des ribozymes SOFA-HDV contre l'ARN brin (-) pourrait provenir du site même de la réplication du VHC, qui est plutôt enclavé par un arrangement plus ou moins ordonné de membranes lipidiques formé dans les cellules infectées (Romero-Brey et al., 2012). Ainsi, les complexes de réplication se retrouvent dans un réseau de membranes («*membranous web*») issu du réticulum endoplasmique. L'ARN brin (-) est principalement un intermédiaire de réplication alors que l'ARN brin (+) sert autant pour la réplication, la traduction que comme génome viral (Pawlotsky et al., 2007). Les rôles distincts de ces deux ARN justifient des localisations cellulaires différentes. D'un côté, l'ARN brin (-) sera confiné aux sites de réplication, alors que l'ARN brin (+) devra minimalement voyager hors de ces complexes pour être traduit. Cette bio-accessibilité dissemblable pourrait expliquer la sensibilité différente des deux ARN viraux face aux ribozymes SOFA-HDV. Une étude récente a également démontré qu'il est quasi impossible de cibler l'ARN brin (-) à l'aide de shARN asymétriques dans un modèle de réplication virale, alors que les mêmes shARN étaient actifs dans un système de surexpression d'un gène rapporteur (Lisowski et al., 2013).

Il semble se dessiner une autre tendance, cette fois dans l'efficacité des ribozymes SOFA-HDV ciblant l'ARN brin (+). Le quatuor de ribozymes ayant les effets les plus marqués (SOFA-HDV-Rz-299, -Rz311, -Rz313 et -Rz318), ciblent une région de 42 nucléotides (positions 299 à 341) en amont du codon d'initiation de la traduction dans l'IRES du virus (Chapitre 1, Figure 5). Il semble que cette séquence ne soit pas uniquement ciblée avantageusement par les ribozymes SOFA-HDV alors que deux siARN se sont montrés efficaces pour cibler ces mêmes nucléotides. Les sites d'appariements du siARN IRES-331 (positions 322 à 340) et du siARN 5U5 (positions 286 à 304) chevauchent largement les sites reconnus par nos ribozymes (Seo et al., 2003; Yokota et al., 2003). Un autre groupe de recherche a également publié un ribozyme «*hammerhead*» dirigé contre les nucléotides 313 à 337 (Sakamoto et al., 1996). Il est difficile de savoir pourquoi

cette région pourrait être avantageusement ciblée par les ribozymes SOFA-HDV. Il est possible que ce biais dans l'inhibition de la traduction virale soit le fruit d'un effet amplifié par la proximité de l'AUG ou encore que cette région soit tout simplement plus propice à l'action de technologies antisens. Il ne faut pas négliger l'impact de l'accessibilité de cette région de l'ARN viral, bien qu'à première vue, cette portion de l'IRES du VHC ne semble pas particulièrement accessible. En terminant, je crois qu'il serait pertinent d'aller évaluer le potentiel des ribozymes SOFA-HDV contre d'autres régions de l'ARN viral. Il semble y avoir des sites beaucoup plus favorables à l'interférence par l'ARN viral que le 5' UTR (Sagan et al., 2010). Ces sites pourraient probablement être ciblés dans un modèle de réplication *in cellulo* sans effectuer un large criblage de ribozymes SOFA-HDV.

1.2. Cibler les ARN du VIH

Le second projet d'application du ribozyme SOFA-HDV auquel j'ai contribué avait pour objectif de neutraliser la réplication du VIH. L'article présenté au Chapitre 2 est en fait la seconde tentative de ciblage des ARN du VIH par des ribozymes SOFA-HDV (Laine et al., 2011). Cette première étude, qui se concentrait à cibler les gènes Tat et Rev du VIH, avait identifié quatre ribozymes pouvant affecter la biologie du virus sans toutefois présenter tous les contrôles démontrant un effet catalytique des ribozymes SOFA-HDV.

À partir de ces résultats, nous avons voulu explorer d'autres gènes du VIH comme cibles potentielles pour des ribozymes SOFA-HDV. L'approche utilisée pour le design des ribozymes était comparable à celle du projet visant le virus de l'Influenza A (Motard et al., 2011). Plus précisément, une analyse détaillée de la conservation nucléotidique dans l'ARN de la région 5' UTR jusqu'à la fin du cadre de lecture du gène *Gag* a été conduite sur les nombreux variants du VIH (banque de données du Los Alamos National Laboratory, www.hiv.lanl.gov). Cette étape initiale était primordiale pour cibler le VIH en raison de la fréquence élevée de

l'incorporation de mutations dans son génome qui favorise l'émergence de virus résistants (pour une revue voir (Bennasser et al., 2007; Haasnoot et al., 2007). Au final, l'activité inhibitrice de la réplication du VIH a été testée pour 18 ribozymes SOFA-HDV spécifiques pour le 5' UTR et le gène *Gag* du virus. Parmi ce groupe, seul le ribozyme SOFA-HDV-Rz1498 a montré une activité significative (~50 % d'inhibition). Les travaux subséquents ont permis de mieux caractériser ce ribozyme prometteur et de démontrer la qualité de son site de liaison pour cibler le VIH.

Premièrement, nous avons démontré que l'effet du SOFA-HDV-Rz1498 sur la réplication du VIH est majoritairement dû à son activité catalytique puisque le ribozyme mutant inactif (SOFA-HDV-Rz1498A76, RD et Bs inchangés mais cytosine catalytique mutée en adénosine) montre un niveau comparable d'activité de transcription inverse (RT) à celui du ribozyme contrôle (SOFA-HDV-RzHBV). C'est donc dire que l'activité inhibitrice du SOFA-HDV-Rz1498 provient de la coupure des ARN cibles et non d'un simple effet antisens. C'est la première fois qu'un tel effet est observé pour un ribozyme SOFA-HDV ciblant la réplication d'un virus. Dans les études précédentes, soit ce mutant inactif n'a pas été testé (Levesque et al., 2010) ou soit l'activité totale observée était l'addition de l'activité catalytique et antisens (Laine et al., 2011; Motard et al., 2011). En fait, dans la majorité des projets utilisant des ribozymes SOFA-HDV, cet aspect n'a pas été évalué (Ben Aissa et al., 2012; D'Anjou et al., 2011; Levesque et al., 2010; Robichaud et al., 2008) ou bien les ribozymes démontraient une activité mixte (catalyse et antisens) (Laine et al., 2011; Motard et al., 2011). Les deux seuls exemples semblables au ribozyme SOFA-HDV-Rz1498 sont le SOFA-HDVRz-303 (co-expression du ribozyme avec le gène C du VHB) (Bergeron et Perreault, 2005) et le SOFA-HDVRz-E232 (ciblant le gène EF-Tu dans la bactérie *E. coli*) (Fiola et al., 2006). Il est important de mentionner que cette nouvelle tentative de cibler les ARN du VIH m'apparaît plus robuste que la précédente (Laine et al., 2011), malgré que l'activité du SOFA-HDV-Rz1498 puisse sembler inférieure à celle du SOFA-

HDV-RzTev1 (~50 % d'inhibition et ~80 %). Mon argument repose sur la présence de meilleurs contrôles dans nos expériences que dans la précédente étude. L'effet mesuré dans l'étude de 2011 surestimait probablement l'impact des ribozymes, car la présence d'un ribozyme non pertinent («*irrelevant*») comme le SOFA-HDV-RzHBV ou d'un ribozyme inactif était suffisante pour générer un effet significatif sur l'activité de RT mesurée en comparaison avec le vecteur vide (communication personnelle de Robert J. Scarborough). D'ailleurs, cette hypothèse est supportée par les résultats présentés au chapitre 2 (**Erreur! Source du renvoi introuvable.**), où l'effet du SOFA-HDV-RzTev1 est de l'ordre de 20 %, une valeur bien en-deçà de celle présentée précédemment (Laine et al., 2011). La principale différence entre les deux études est le système d'expression des ribozymes (ARN polymérase II (Laine et al., 2011) vs ARN polymérase III (Scarborough et al., Soumis)). Il semble que la production de ribozymes sous la forme d'un messager (coiffe en 5' et poly(A) en 3') affecterait la biologie du virus sans lien avec le ribozyme lui-même. Probablement que l'utilisation du promoteur CMV ne représente pas une bonne option pour mesurer l'efficacité d'un outil moléculaire dirigé contre le VIH (Konopka et al., 2000). Au final, le vecteur psiRNA-hH1GFPZEO (Invivogen) semble être un outil plus approprié pour l'expression de ribozymes SOFA-HDV.

Deuxièmement, l'activité du ribozyme SOFA-HDV-Rz1498 laissait supposer que nous avons identifié un site favorable à la liaison du ribozyme et à son activité catalytique. La production d'un shARN (shRNA1498) reconnaissant la même séquence que le SOFA-HDV-Rz1498 a non seulement permis de confirmer que cette région des ARN du VIH est particulièrement accessible pour la liaison d'acides nucléiques, mais également d'identifier un shARN très efficace contre la réplication du VIH. Nos travaux réalisés avec différentes souches du VIH ont également confirmé que ce site est conservé en terme de séquence, mais également au niveau de son accessibilité. Le ciblage du site 1498 est donc

prometteur pour le développement d'une stratégie de contrôle de la réplication du VIH basée sur des outils moléculaires faits d'acides nucléiques.

Finalement, nous nous sommes intéressés à la fiabilité du SOFA-HDV-Rz1498 sous différents angles. Brièvement, à l'étape du design nous avons d'abord favorisé la sélection de ribozymes avec un faible risque d'affecter le transcriptome des cellules en rejetant un certain nombre de ribozymes sur la base d'une prédiction bioinformatique produite par le programme Ribosubstrate (Lucier et al., 2006). De cette façon, les candidats ayant les plus fortes probabilités d'avoir un effet toxique sur les cellules ont pu être rejetés. L'absence de phénotype marqué lors de l'expression des ribozymes testés en cellules et l'analyse plus approfondie de l'effet du SOFA-HDV-Rz1498 tendent à justifier la pertinence de cette étape. Par la suite, nous avons analysé la spécificité du ribozyme SOFA-HDV-Rz1498 *in cellulo*, puisque la fiabilité ou la sûreté d'un ciseau moléculaire repose principalement sur cet élément de son activité. Ce point est discuté en détails à la section 4 de la discussion. En résumé, l'ensemble des données recueillies dans ce projet démontre le fort potentiel du site 1498 du gène *Gag* pour le ciblage du VIH par des effecteurs d'acides nucléiques et la spécificité élevée du ribozyme SOFA-HDV-Rz1498.

2. L'équation multifactorielle de l'activité des ribozymes SOFA-HDV

Au cours du projet de ciblage du VHC, mais également dans la plupart des projets similaires, nous avons observé une disparité dans le niveau d'activité *in vitro* des ribozymes SOFA-HDV. Cette variation dans l'efficacité de coupure des ribozymes même sur des substrats très courts (± 30 nucléotides) où l'accessibilité du site de liaison devrait être moins limitante, laissait croire que notre processus de sélection pouvait être amélioré. Mon hypothèse était donc que l'activité intrinsèque des ribozymes était affectée par les nucléotides variables, qui sont spécifiques à chaque couple « ribozyme-site » de coupure. Ainsi, j'ai entrepris d'améliorer la

façon de faire en visant la présélection plus efficace des ribozymes SOFA-HDV avant de les tester. Il faut comprendre que l'emphase de mes travaux a été mise sur l'identification des ribozymes capables d'effectuer efficacement la catalyse en raison de leur séquence et de celle qu'ils ciblent plutôt que sur l'évaluation de l'accessibilité de leur site de coupure dans l'ARN cible. Ce dernier aspect est un domaine en soit que j'ai abordé sans toutefois qu'il occupe le centre de mes travaux. En considérant tous les travaux d'analyse de la relation structure-fonction du ribozyme HDV ainsi que ceux portant sur sa spécificité, il était possible de générer des hypothèses solides sur des facteurs pouvant influencer l'activité des ribozymes SOFA-HDV. De plus, l'accumulation de données comprenant l'activité de coupure *in vitro*, la combinaison de séquences du RD, Bs et BI, ainsi que la séquence de l'ARN cible pour un grand nombre de ribozymes SOFA-HDV, nous a fourni l'information nécessaire pour tester certaines hypothèses. Ce sont ces analyses qui sont à l'origine des manuscrits présentés aux chapitres 3 et 4.

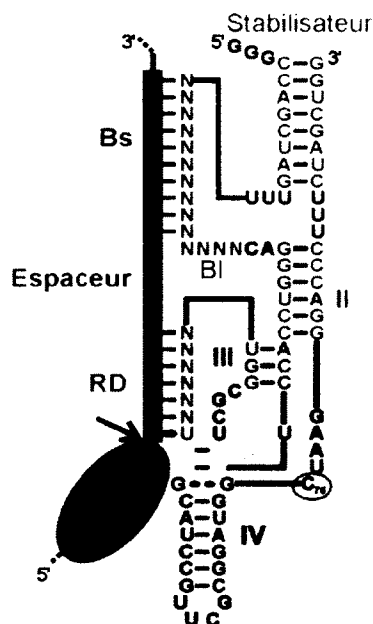
2.1. L'ARN cible a un impact à plus d'un niveau

L'activité de coupure associée à un ribozyme SOFA-HDV et à un site de coupure dépend d'éléments présents dans deux ARN distincts, le ribozyme et l'ARN cible. Dans cette section sur l'ARN cible, je vais d'abord discuter de l'influence des nucléotides à proximité du site de coupure, puis commenter sur l'impact possible de l'espaceur («*spacer*») et finalement aborder le sujet de l'accessibilité du site de liaison.

2.1.1. L'entourage du site de coupure influence la coupure

Tout d'abord, j'ai voulu vérifier la pertinence des éléments de spécificité du ribozyme HDV qui ont été décrits précédemment dans la littérature (Introduction, Figure 13B). Les analyses précédentes démontraient que la nature de certains nucléotides, positions -1 à -4, ainsi que les positions +3 et +4 par rapport au site de

coupure, contribuaient à définir la spécificité du ribozyme de type sauvage (pour une revue voir : Bergeron et al., 2003). En fait, l'activité de coupure du ribozyme varie selon la nature des nucléotides à ces positions dans le substrat et donc certains sites de coupure pourraient être meilleurs que d'autres. Avant d'inclure ces paramètres dans les critères de présélection des ribozymes, j'ai voulu évaluer si les effets observés sur le ribozyme HDV demeuraient valides pour le ribozyme SOFA-HDV. Plutôt que de refaire une série d'expériences pour répondre à cette question, je me suis tourné vers l'analyse des ribozymes SOFA-HDV que j'avais déjà testés. Ainsi, j'ai épluché la séquence de chaque ribozyme du projet VHC pour voir s'il y avait un lien entre les meilleurs ribozymes et les nucléotides présents aux positions -1 à -4 (en amont) et +3 et +4 (en aval) de leur site de coupure (Discussion, Figure 1).



Discussion, Figure 1. Représentation détaillée du ribozyme SOFA-HDV lié à son substrat.

La structure secondaire du ribozyme SOFA-HDV (ARN de droite) lié à son substrat (ARN de gauche) est illustrée. Au niveau du substrat, les nucléotides en amont du site de coupure (cercle gris) sont identifiés par la séquence H.₁N.₂N.₃N.₄ (où le chiffre en indice indique la position par rapport au site de coupure). Les sites de

liaisons du biosenseur (Bs, N noir sur fond vert) et du domaine de reconnaissance (RD, N noir sur fond bleu) sont séparés par l'espaceur (N noir sur fond gris). Au niveau du ribozyme, les nucléotides du stabilisateur (lettres grises), du Bs (N verts), du bloqueur (Bl, N rouges) et du RD (N bleus) sont dessinés. L'encadré gris marque le ribozyme HDV avec ses tiges I.I (jaune), II (orange), III (rose) et IV (brun) identifiées par des nucléotides de couleurs. La cytosine catalytique (C₇₆, cercle rouge) est également identifiée. Tous les nucléotides sont représentés par des lettres (A, C, G, U, N ou H). Spécifiquement, les nucléotides variables sont identifiés par N (N signifie A, C, G, ou U) ou par H (H signifie A, C ou U). Le site de coupure est marqué d'une flèche.

Les meilleurs ribozymes avaient effectivement plus souvent les nucléotides préférés du ribozyme HDV aux positions -1 et -2. Avec le système de pointage présenté au chapitre 3, les ribozymes avec une activité de coupure plus grande que 20 % avait un pointage moyen plus de deux fois plus grand que ceux ayant une activité de coupure inférieure à 20 %. En fait, la leçon générale à tirer est que des doublets de pyrimidines (C ou U) sont des sites de coupures moins favorables pour l'obtention de ribozymes avec des activités de coupure au dessus de la moyenne. Le lien avec les positions -3 et -4 n'était pas visible, mais ce n'était pas vraiment surprenant puisque les précédents travaux montraient un effet beaucoup plus modeste de ces positions sur l'activité du ribozyme HDV (Deschenes et al., 2000; Deschenes et al., 2003). La principale explication de l'effet des nucléotides en amont du site de coupure est que s'ils sont complémentaires à la jonction I/IV (GG), il y aura compétition pour la formation de la tige I.I essentielle à l'activité (Discussion, Figure 1) (Bergeron et al., 2003; Deschenes et al., 2003).

Mes observations au niveau des positions +3 et +4 ne supportent aucune corrélation entre l'identité des nucléotides et l'activité des ribozymes SOFA-HDV. Ainsi, il appert que le domaine de reconnaissance du ribozyme est relativement flexible et peut accommoder assez bien un large éventail de séquences. L'étude de ces positions pour le ribozyme HDV a été réalisée dans le contexte d'un seul substrat où les cinq autres positions du RD demeuraient inchangées (Ananvoranich et al., 1999). Il est donc possible que les résultats passés étaient

liés au reste de la séquence de sept nucléotides du domaine de reconnaissance, plutôt qu'à un fait général pour ce motif catalytique.

Globalement, la tendance observée parmi les ribozymes ciblant le VHC semble également se transposer à la pléthore de ribozymes testés parmi les nombreux projets de ciblage (Ben Aissa et al., 2012; Fiola et al., 2006; Laine et al., 2011; Levesque et al., 2010; Motard et al., 2011; Robichaud et al., 2008; Scarborough et al., soumis). L'analyse de ces ribozymes en un tout est toutefois difficile et devrait être marquée d'un bémol étant donné la grande variation dans les conditions expérimentales (ex. : la concentration des ribozymes, la taille des substrats ou les différents expérimentateurs) utilisées dans ces projets. L'information à retenir est donc que les nucléotides de l'ARN cible en amont du site de coupure (-1 et -2) influencent l'activité de coupure des ribozymes SOFA-HDV.

2.1.2. Le «*spacer*» peut lui aussi changer la donne

Les nucléotides de l'ARN cible entre le site de liaison du RD et du Bs (le «*spacer*») sont également un facteur dans la réussite d'un ribozyme (Discussion, Figure 1). Les travaux de Lucien Jr Bergeron ont très bien démontré que la longueur du «*spacer*» influence l'activité du ribozyme (Bergeron et Perreault, 2005; Bergeron et al., 2005). Au-delà d'une quinzaine de nucléotides entre le site de liaison du RD et du Bs, l'activité de coupure du ribozyme SOFA-HDV devient très faible. Il a aussi été démontré que l'activation du ribozyme par la liaison du biosenseur à un site distant (ex. : 80 ou 500 nucléotides en aval) ne pouvait se faire. D'ailleurs, cet élément ajoute à la spécificité du ribozyme SOFA-HDV, puisqu'un ARN ne représente pas une cible si des séquences éloignées l'une de l'autre sont complémentaires au RD et au Bs d'un ribozyme donné. Les travaux publiés en 2005 démontraient que, dans un interval allant de un à sept nucléotides, le ribozyme était relativement efficace mais que la longueur optimale entre les deux sites de liaison serait de trois nucléotides (Bergeron et al., 2005). Il ne faut pas

oublier que ces données ont été recueillies dans une expérience spécifique sans que les résultats ne soient validés pour plusieurs ribozymes SOFA-HDV différents. Pour la majorité des projets de ciblage de gènes réalisés jusqu'à maintenant, les ribozymes ont été dessinés avec un «*spacer*» de base de cinq nucléotides correspondant au design original du ribozyme (Introduction, Figure 14B) (Bergeron et Perreault, 2005). Dans certains cas, la stratégie présentée au chapitre 4 (Chapitre 4, Figure 3) a été utilisée pour respecter les critères de conservation de séquence dans l'ARN cible. Je crois que l'utilisation d'une longueur de base de trois à cinq nucléotides est la bonne stratégie pour le criblage d'une banque de ribozymes dirigés contre un ARN d'intérêt. Une fois l'identification d'un ou plusieurs candidats intéressants, il serait justifié de tester différentes possibilités. Des données préliminaires non publiées (Robert Scarborough, communication personnelle) obtenues pour le site de coupure 1498 du projet VIH démontrent une activité *in cellulo* variable pour des ribozymes dessinés avec différentes longueurs de «*spacer*». Dans ce cas particulier, le «*spacer*» de quatre nucléotides du SOFA-HDV-Rz1498 semble optimal. En terminant cette section, je veux mentionner que des essais de coupure avec un ribozyme SOFA-HDV contre des substrats ayant différentes séquences de «*spacer*» suggèrent une influence sur l'activité de coupure. Entre autres, ces essais préliminaires évoquent un risque d'inhibition du ribozyme dans le cas d'une complémentarité étendue (> 4 pb) qui prolongerait l'hélice du RD par la formation de paires de bases entre les nucléotides de la tige-boucle III (Discussion, Figure 1). En considérant les probabilités de rencontrer un «*spacer*» avec cette séquence précise et l'effet d'inhibition modéré, il ne semble pas approprié d'analyser ce point lors du design des ribozymes. Bien que ce ne soit pas un facteur clé dans l'activité des ribozymes SOFA-HDV, il demeure intéressant de savoir qu'en plus de la longueur du «*spacer*», la séquence de celui-ci peut modifier l'efficacité de la coupure par le ribozyme.

2.1.3. L'accessibilité du site de coupure est un élément clé

Sans l'ombre d'un doute, nous pouvons affirmer que l'accessibilité du site de liaison a un impact sur l'activité d'un ribozyme ou de tout autre outil d'acides nucléiques antisens (Low et al., 2012; Lundblad et Altman, 2010; Sagan et al., 2010). Certains travaux présentés dans cette thèse illustrent cette réalité. Ils apportent également des informations qui seront éventuellement importantes pour aborder la prédiction d'accessibilité pour les ribozymes SOFA-HDV.

Tout d'abord, l'analyse de l'activité de coupure des ribozymes dirigés contre l'ARN brin (-) du VHC montre que ces derniers sont, en moyenne, plus actifs que ceux visant l'ARN brin (+) (Chapitre 1, Tableau 3). La comparaison de la structure des deux ARN ciblés supporte l'idée que celle de l'ARN brin (-) est moins compacte que celle de l'ARN brin (+) (Chapitre 1, Figure 5) (Smith et al., 2002). En regardant de plus près, il appert que les meilleurs ribozymes des deux groupes ciblent davantage des sites plus accessibles (moins de nucléotides formant des paires de bases selon les structures publiées). La prédiction d'accessibilité (bioinformatique et biochimique, Chapitre 1, Figure 2) aurait dû être plus fructueuse qu'elle ne l'a finalement été. Le potentiel de cette méthode est principalement affaibli par le fait qu'elle tient compte uniquement du site de liaison du ribozyme HDV (RD) sans évaluer l'accessibilité du site de liaison du biosenseur. L'adaptation de cette méthode pour tenir compte de cette dualité dans la reconnaissance du substrat est difficile à envisager comme nous l'avons mentionné dans la discussion de l'article du chapitre 1. La prédiction de l'accessibilité d'un site pour le ribozyme SOFA-HDV était d'autant plus complexe, que la dynamique de liaison du ribozyme à son substrat en fonction de la structure de ce dernier n'avait jamais été étudiée.

Les résultats obtenus dans le projet de ciblage du VIH supportent également l'importance de l'accessibilité pour l'activité du ribozyme SOFA-HDV. Le site de coupure du ribozyme le plus actif identifié dans le criblage *in cellulo* initial

semble être dans une région accessible de l'ARN du VIH. Le shRNA1498 ciblant la même séquence et une série de petits ARN interférents se liant près de cette région démontrent des activités élevées (Low et al., 2012; McIntyre et al., 2009; Naito et al., 2007). De plus, la structure proposée pour cette région du génome viral est majoritairement simple brin (Watts et al., 2009).

L'étude des éléments importants pour le design de ribozymes SOFA-HDV plus efficaces a mené à la question de l'accessibilité de leur site de liaison. Avant de réaliser ces expériences (Chapitre 3, Figure 4), il n'était pas clair si l'accessibilité du site de liaison du domaine de reconnaissance ou du biosenseur pouvait influencer l'activité de coupure individuellement. Dans un premier temps, l'activité du ribozyme SOFA-HDV dépend clairement de l'accessibilité générale de son site de liaison puisqu'un substrat principalement double brin n'est pas coupé efficacement. Une structure au niveau du site de liaison du RD est suffisante pour réduire significativement l'activité de coupure. Dans un second temps, la même conclusion peut être tirée pour le site de liaison du Bs. Puisque l'accessibilité des deux sites de liaison contribue à l'activité de coupure, on comprend mieux pourquoi les prédictions du projet de ciblage du VHC n'ont pas donné les résultats escomptés. De plus, dans ces deux cas, l'inhibition dépend de la stabilité (ΔG) de la structure prédite de l'ARN cible. En effet, l'introduction de mésappariements dans les structures encombrant l'un ou l'autre des sites de liaisons permet de rehausser l'activité de coupure du ribozyme. Une autre donnée importante est que l'activité du ribozyme ne dépend pas seulement des 17 (c'est-à-dire 7 plus 10) nucléotides auxquels il s'apparie. Nous observons une différence dans l'activité du ribozyme lorsque la stabilité globale de l'ARN cible était modifiée sans réduire le nombre de paires de bases au niveau des sites de liaison (données non-présentées). Ainsi, l'accessibilité d'un site va au-delà des nucléotides qui sont impliqués dans l'appariement du ribozyme. Ce phénomène avait déjà été rapporté pour le ribozyme HDV (Swiatkowska et al., 2007). Dans cette même étude, les auteurs n'ont pu obtenir de ribozymes très actifs malgré la sélection de sites

accessibles par deux méthodes (bibliothèque combinatoire de ribozymes ou oligonucléotide avec RNase H). Ils ont conclu que le ciblage d'un ARN structuré était plus difficile qu'un ARN linéaire malgré la pleine accessibilité du site de liaison. Cela rejoint les conclusions obtenues avec mon expérience modèle qui, selon moi, représenterait un excellent gabarit pour étudier plus en détails l'influence de la structure de l'ARN cible.

À ce point-ci, il est clair que l'accessibilité de la cible est une composante importante pour l'activité des ribozymes SOFA-HDV. Par contre, comment prédire si un site de coupure sera favorable pour son ribozyme correspondant demeure une question ouverte. Les méthodes de prédiction développées dans le passé pour le ribozyme HDV se sont montrées peu efficaces (Ben Aissa et al., 2012; Levesque et al., 2010). De son côté, la prédiction bioinformatique n'est pas encore tout à fait au point (Sagan et al., 2010). La principale raison vient du manque de précision des outils de prédiction de structures secondaires des longs ARN (Shareghi et al., 2012), mais les progrès sont fréquents. Les méthodes biochimiques semblent donner de bons résultats pour les ARN interférents (Sagan et al., 2010). On peut penser qu'une méthode adaptée au contexte du ribozyme SOFA-HDV donnerait des résultats aussi satisfaisants. Les outils existant pour les siARN ou les shARN pourraient être utilisés sans modifications pour les ribozymes SOFA-HDV, puisque le site 1498 (VIH) est bien accessible pour les deux technologies.

En terminant cette section, je pense qu'il faut garder en tête que cet élément n'est pas si simple à évaluer. Certaines techniques pourront sûrement donner d'excellents résultats, mais représenteront également un investissement important de temps et de ressources. De plus, la majorité des techniques visant à identifier l'accessibilité de futurs sites de ciblage ne tient pas compte des réalités intracellulaires, comme la présence de protéines liées aux ARN et leur influence sur le repliement de ceux-ci. Je crois qu'il faudra toujours maintenir un équilibre entre les efforts mis pour le design et pour les expériences. Il est probablement

utopique de penser identifier le meilleur candidat sans en avoir directement testé un certain nombre. C'est d'ailleurs une leçon qui peut être tirée des nombreux travaux avec les oligonucleotides antisens, où les méthodes de «*tiling*» ou de sélection à partir d'une banque aléatoire sont encore utiles pour dénicher les meilleurs candidats (Bennett et Swayze, 2010).

2.2. Les composantes du ribozyme et leur influence sur l'activité du ribozyme

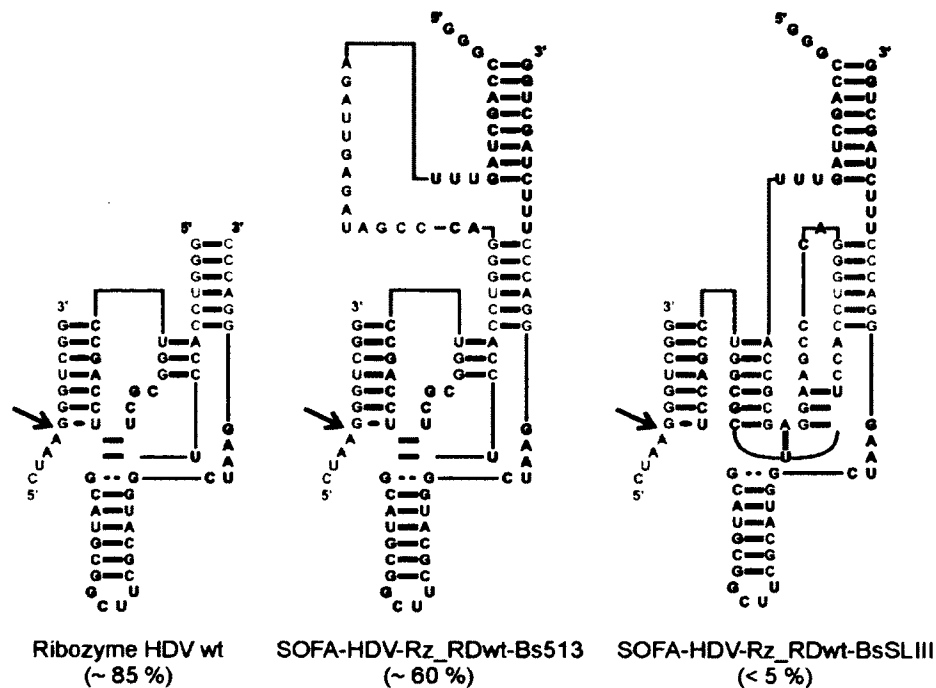
Certes la structure, ou plutôt l'accessibilité de la région ciblée, influence l'activité du ribozyme. Par contre, ce critère n'est pas suffisant pour expliquer les succès ou non d'un ribozyme SOFA-HDV. Le plus bel exemple de cette réalité se trouve dans les résultats obtenus avec le projet de ciblage du VIH. Nous avons observé une inhibition avec deux shARN (shRNA522 et shRNA553), ce qui semble indiquer deux sites relativement accessibles dans l'ARN viral. Les ribozymes SOFA-HDV ciblant ces deux mêmes sites n'ont pas montré une activité significative, contrairement au SOFA-HDV-Rz1498 et au shRNA1498. La principale différence entre ces deux cas est la piètre activité de coupure *in vitro* des ribozymes SOFA-HDV-Rz553 et -Rz522. Selon moi, il existe un facteur autre que l'accessibilité pour qu'un ribozyme puisse avoir un effet sur sa cible en cellule. Je définirais ce facteur comme l'activité intrinsèque du ribozyme. Cette activité intrinsèque représente plus ou moins l'activité maximale qu'un ribozyme peut avoir dans le cas où son site de liaison est parfaitement accessible. Plusieurs éléments interviennent pour fixer cette activité. Certains proviennent de l'ARN cible et ont été discutés dans la section précédente, alors que d'autres sont liés à la séquence du ribozyme. Dans la section qui suit, je vais donc discuter de ces éléments qui influencent l'activité de coupure des ribozymes SOFA-HDV et qui proviennent de leur séquence.

2.2.1. Les biosenseurs ne sont pas tous égaux

Mon analyse s'est d'abord tournée vers un élément du module SOFA, le biosenseur (Bs). L'idée d'évaluer la possibilité que certaines séquences de biosenseurs puissent être néfastes à l'activité du ribozyme provient de diverses observations. Entre autres, l'étude du G-quartzyme, un ribozyme HDV fusionné à un aptamère formant une structure G-quadruplexe, a révélé un nouveau mécanisme d'inhibition du ribozyme par une séquence externe au motif catalytique (Beaudoin et Perreault, 2008). Dans cette étude, un ribozyme allostérique activable par la présence d'ions potassium a été créé. En l'absence d'ions potassium, un segment de l'aptamère se liait au centre catalytique (tige-boucle III, Discussion, Figure 1) du ribozyme et empêchait le bon repliement du ribozyme HDV. L'ajout d'ions potassium permettait à l'aptamère de former une structure G-quadruplexe qui libérait le centre catalytique et permettait l'activité. Cette expérience démontrait qu'il est possible d'affecter l'activité du ribozyme par l'entremise d'une séquence en *cis* qui perturbe le bon repliement du centre catalytique.

Initialement, j'ai testé l'hypothèse du biosenseur inhibiteur avec un modèle simple de ribozyme. Pour ce faire, j'ai créé un système où le ribozyme serait constitutivement actif dû à un bloqueur non complémentaire au RD (Discussion, Figure 2). Pour ces expériences initiales, le substrat ARN était très court (11 nucléotides) et ne contenait pas de site de liaison pour le Bs. Ainsi, il n'y avait pas de différences dans la liaison au substrat entre les ribozymes mutants et les essais pouvaient être réalisés avec un seul substrat radio-marqué. De cette façon l'effet du biosenseur sur la catalyse du ribozyme était isolé et mesuré dans un essai de coupure standardisé. Dans ces conditions expérimentales, j'ai testé une panoplie de différentes combinaisons de nucléotides plus ou moins complémentaires à d'autres parties du ribozyme HDV (Discussion, Figure 2 et données non-présentées). De cette façon, nous avons pu identifier la tige-boucle III comme la plus sensible pour ce type d'interaction. Ce modèle était très simplifié et ne tenait

pas compte du fait que la liaison du substrat au biosenseur pourrait compétitionner cette interaction néfaste. Pour cela, nous avons plutôt opté pour le modèle expérimental présenté au chapitre 3 (Chapitre 3, Figure 2 et Chapitre 3, Tableau 1). Dans l'ensemble, toutes ces expériences ont démontré que la complémentarité entre le biosenseur et le centre catalytique du ribozyme HDV avait un effet inhibiteur. Cet effet était plus marqué pour les nucléotides en 3' du biosenseur et s'étendait même à la séquence du bloqueur (données non-présentées). C'est d'ailleurs pour cette raison que la séquence utilisée pour l'analyse de complémentarité («*RNAhybrid*») comprenait plus que les 10 à 12 nucléotides du biosenseur. Ainsi, j'ai ajouté au biosenseur les trois U en amont du biosenseur (entre le stabiliser et le Bs), le bloqueur et le CA (entre le bloqueur et la tige II) (Discussion, Figure 1). Cette séquence 5'-UUUBsBICA-3' décrite au chapitre 3 est davantage pertinente, étant donné qu'elle incluait tous les nucléotides qui sont disponibles pour former des paires de bases selon la structure secondaire attendue des ribozymes SOFA-HDV. Lorsqu'on analyse les valeurs de MFE présentées au chapitre 3, mais également celles obtenues pour plusieurs autres ribozymes SOFA-HDV (différents projets), il semble qu'une valeur inférieure à -15 kcal/mol soit peu favorable à l'activité du ribozyme. Cette valeur pourrait représenter une limite à ne pas dépasser lors de la sélection d'un excellent ribozyme SOFA-HDV.



Discussion, Figure 2. Représentations de trois ribozymes de l'expérience préliminaire liés au substrat de 11 nucléotides.

La séquence et la structure secondaire attendue du ribozyme HDV de type sauvage (Ribozyme HDV wt, gauche) et de deux ribozymes SOFA-HDV (centre et droite) utilisés dans l'expérience préliminaire d'inhibition de l'activité par le biosenseur, sont illustrées. Le substrat de 11 nucléotides (gris) peut être lié par le RD (bleu) de chaque ribozyme. Les deux ribozymes SOFA ont le même bloqueur (rouge) qui ne peut s'apparier au RD ce qui leur permet de reconnaître et de couper le court substrat sans l'appariement du Bs (vert). Le biosenseur du ribozyme SOFA-HDV-Rz_RDwt-Bs513 provient du ribozyme SOFA-HDV-Rz513 dessiné pour cibler le VHB (Bergeron et Perreault, 2005) alors que celui du ribozyme SOFA-HDV-Rz_RDwt-BsSLIII est complémentaire à la tige-boucle III du ribozyme HDV. Le pourcentage de coupure calculé après une incubation de deux heures pour chaque ribozyme est indiqué sous leur nom respectif.

2.2.2. Le bloqueur, une petite séquence au potentiel dévastateur

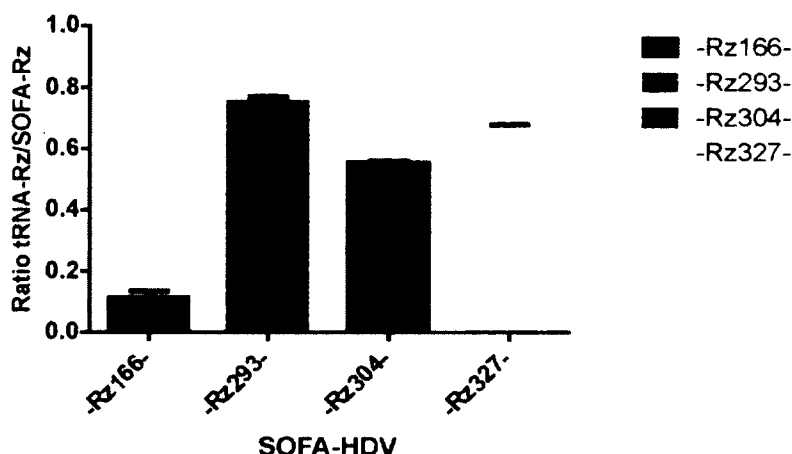
Le second élément étudié est le bloqueur. Il faut savoir que les travaux initiaux sur le ribozyme SOFA-HDV démontraient que l'activité du ribozyme diminuait avec

l'extension du bloqueur (Bergeron et al., 2005). Selon ces mêmes travaux, la longueur optimale du bloqueur est de quatre paires de bases. Toutefois, il est possible que les derniers nucléotides du biosenseur prolongent l'hélice du bloqueur. À un certain point, si cette tige s'allonge jusqu'à faire sept paires de bases avec le domaine de reconnaissance, il y a possibilité d'auto-coupeure (Chapitre 3, Figure 3). C'est d'ailleurs un problème que j'ai rencontré par hasard en testant différents mutants, sans avoir pris la peine de vérifier cet élément. J'ai également observé à plusieurs reprises ce qui semblait être de l'auto-coupeure lors de la préparation de ribozymes SOFA-HDV par transcription *in vitro*. Sans pouvoir différencier hors de tout doute l'auto-coupeure d'un arrêt de la transcription, la taille des ARN obtenus, trois transcrits (court, moyen et pleine longueur), concordait avec une coupure au niveau du biosenseur. Malheureusement, dans chacun des cas, il m'a été impossible de retracer une structure canonique de ribozyme en *cis* pouvant expliquer facilement mes observations. Par contre, tout cela démontre qu'avant de se lancer dans le clonage d'une série de ribozymes SOFA-HDV pour un projet de ciblage génique en cellules, il est important d'une part, de vérifier la longueur possible du bloqueur lors du design des ribozymes et d'autre part, de confirmer l'absence d'auto-coupeure par transcription *in vitro*.

2.2.3. L'impact des séquences leaders et du stabilisateur sur l'activité

Les résultats quelque peu limités pour le ciblage du VHC justifiaient la vérification de l'activité de coupure *in vitro* des ribozymes hybrides (tRNA^{val}-SOFA-HDV). Avec la stabilité du motif HDV, je pensais que l'ajout d'une séquence leader en 5' n'aurait peu ou pas d'effet sur l'activité de coupure. J'ai donc testé l'activité des quatre meilleurs ribozymes (% coupure *in vitro* à la fin de la réaction) ciblant l'ARN brin (-) dans les mêmes conditions présentées dans l'article (Chapitre 1, Tableau 3). Le ratio de l'activité de coupure de la version hybride sur l'activité de sa version

SOFA-HDV-Rz montre que tous les ribozymes hybrides sont moins actifs (réduction entre 25 et 90 %, Discussion, Figure 3).

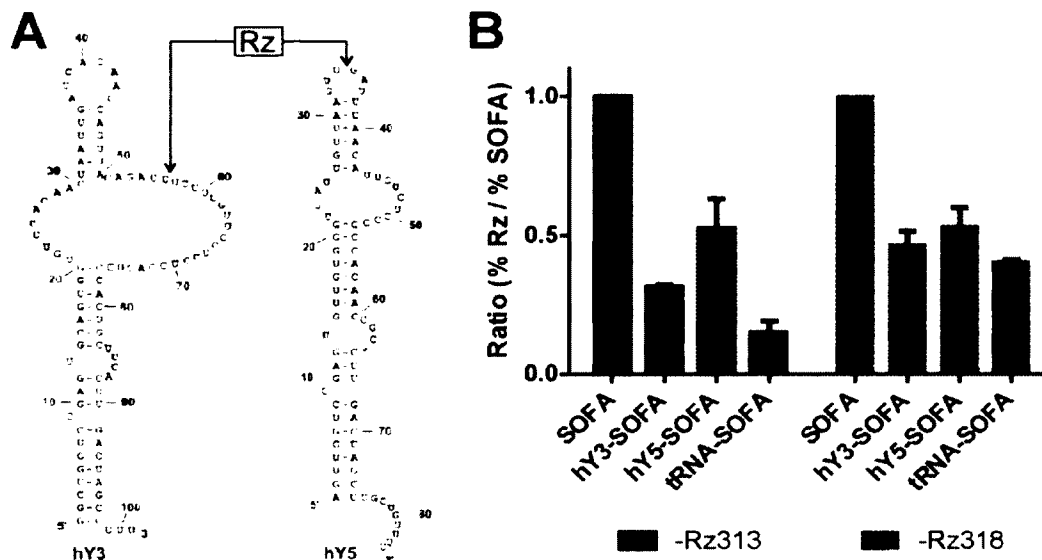


Discussion, Figure 3. Comparaison de l'activité des ribozymes tRNAval-SOFA-HDV-Rz et de leur version SOFA-HDV-Rz.

L'histogramme présente le ratio des pourcentages de coupure *in vitro* à la fin de la réaction («*end-point*», 3 heures) pour comparer la version hybride (tRNAval-SOFA-HDV-Rz) et la version originale (SOFA-HDV-Rz). L'expérience a été faite pour quatre ribozymes (-Rz166⁻, -Rz293⁻, -Rz304⁻ et -Rz327⁻) dessinés pour cibler l'ARN brin(-) du VHC. Un ratio inférieur à 1 signifie que le ribozyme est moins actif que sa version SOFA-HDV-Rz. Les réactions de coupures ont été réalisées dans les mêmes conditions que celles présentées au chapitre 1 (Chapitre 1, Tableau 3).

La question qui suivait était de savoir si cette baisse de l'activité était spécifique aux fusions avec l'ARN de transfert de la valine. Ainsi, j'ai vérifié si d'autres ARN structurés avaient le même genre d'effet sur l'activité du ribozyme. Les ARN hY (hY1, hY2, hY3 et hY5) sont de petits ARN cellulaires non codants, transcrits par l'ARN polymérase III et dont le rôle reste partiellement compris (Langley et al., 2010; Sim et Wolin, 2011). L'ARN hY3 serait exporté au cytoplasme alors que l'ARN hY5 resterait majoritairement au noyau (Sim et Wolin, 2011). Donc, leur fusion au ribozyme SOFA-HDV pourrait représenter une alternative pour l'expression et la localisation de ces ribozymes (Discussion, Figure 4A). Pour tester l'effet de leur fusion au ribozyme, j'ai utilisé deux autres ribozymes SOFA-HDV ciblant cette fois l'ARN brin (+) du VHC. Encore une fois, l'activité des

ribozymes hybrides est significativement réduite comparativement au ribozyme SOFA-HDV seul (Discussion, Figure 4B). On voit également que l'impact sur l'activité de coupure n'est pas le même pour chaque ribozyme.



Discussion, Figure 4. Comparaison de l'effet de différentes séquences leader.

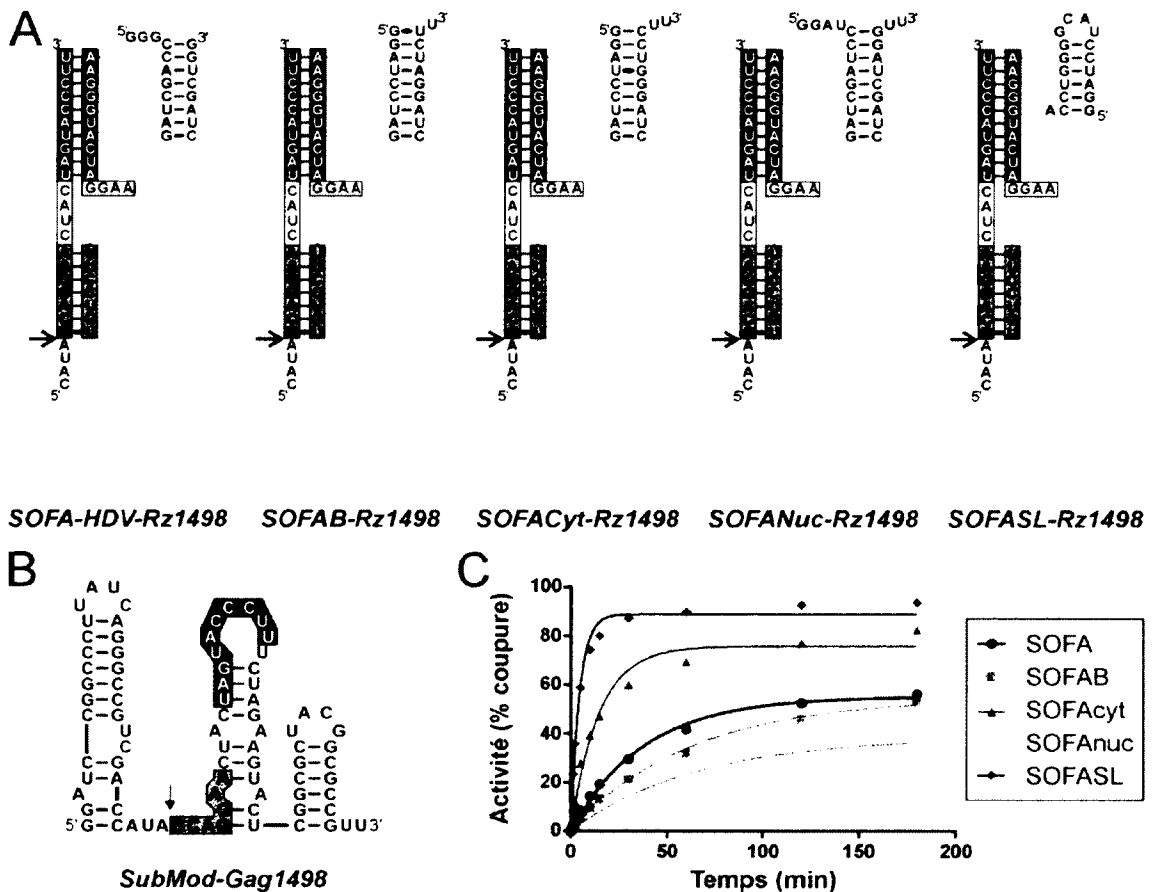
A) Structures secondaires des ARN hY3 et hY5. Le site d'insertion des ribozymes SOFA-HDV sur chaque ARN hY est indiqué par les flèches reliées au Rz encadré. Les structures secondaires ont été inspirées de (Boulanger et al., 1995). **B)** Histogramme comparant l'activité de coupure *in vitro* de quatre versions de ribozymes SOFA. Les ratios sont donnés par le pourcentage de coupure à la fin de la réaction d'une version de ribozymes divisés par l'activité de coupure de la version SOFA-HDV-Rz. L'expérience a été réalisée avec deux ribozymes SOFA-HDV (-Rz313 et -Rz318) dessinés pour cibler l'ARN brin(+) du VHC. Un ratio inférieur à 1 signifie que le ribozyme est moins actif que sa version SOFA-HDV-Rz. Les réactions de coupures ont été faites dans les mêmes conditions que celles présentées au chapitre 1 (Chapitre 1, Figure 3 et Chapitre 1, Tableau 2).

Puisque la fusion du ribozyme SOFA-HDV avec des « motifs de localisation » semblait générer une baisse de l'activité de coupure, je me suis demandé si on ne pouvait pas tirer parti du stabilisateur pour moduler la localisation du ribozyme (voir section suivante). J'ai dessiné deux variants de stabilisateur créant le ribozyme SOFACyt et le ribozyme SOFANuc (Discussion, Figure 5A). Pour des besoins de

clonage, j'ai aussi créé un ribozyme SOFA-HDV avec un stabilisateur débutant par le site de l'enzyme de restriction *BamH1* (GGAUCC) que j'ai nommé SOFAB (Discussion, Figure 5A). Ce stabilisateur est de même taille que l'original, mais ne possède pas de nucléotides simples brins en 5' et contient une paire de base «*Wobble*» qui empêche d'avoir deux sites de coupure *BamH1*. J'ai effectué des essais de coupure sur un petit substrat peu structuré avec lequel le ribozyme SOFA montre une bonne activité de coupure (Discussion, Figure 5B). On peut remarquer que les ribozymes SOFA-HDV-, SOFAB-, SOFACyt- et SOFANuc-Rz1498 ont des activités de coupure bien différentes (Discussion, Figure 5C). Étonnamment, le ribozyme SOFA-HDV-Rz1498 n'était pas le plus actif, ce qui ajoute un nouveau facteur dans l'activité du ribozyme. Les travaux précédents avançaient que le stabilisateur n'avait aucun effet sur l'activité de coupure (Bergeron et al., 2005), mais mes résultats semblent montrer le contraire. Mes conditions expérimentales sont probablement plus sensibles (court substrat de 83 nt vs long substrat de 1190 nt, marquage au P³² en 5' vs marquage interne, concentration du ribozyme 100 nM vs 1 µM et essai de coupure dans le temps vs essai de type «*end-point*»).

L'activité du SOFACyt plus élevée que celle du SOFA régulier m'a donné l'idée de tester un ribozyme qui serait encore plus efficace pour la coupure. J'ai émis l'hypothèse que l'union des extrémités 5' et 3' du ribozyme pouvait nuire à la formation de l'hélice entre le biosenseur et le substrat. En fait, l'appariement du biosenseur (10 pb) représente presque un tour d'hélice (11 pb) et il me semblait plus difficile d'enrouler le biosenseur autour du substrat avec les extrémités jointes. Ainsi, j'ai dessiné le ribozyme SOFASL dans lequel une tige boucle en 5' remplace le stabilisateur, alors que le 3' se termine à la fin de l'hélice II comme pour le ribozyme HDV original. L'activité de ce ribozyme est effectivement plus élevée que pour tous les autres designs de stabilisateurs (Discussion, Figure 5C). L'activité des ribozymes SOFA-HDV, SOFAB et SOFASL a également été mesurée pour la version Tev1, ainsi que sur des substrats plus courts ou plus longs (données non-

présentées). Dans tous les cas, l'activité de ces ribozymes était différente, mais le meilleur ribozyme demeurerait toujours le SOFASL. Il serait intéressant d'investiguer davantage les possibilités au niveau du stabilisateur. Il y a peut-être un certain potentiel pour augmenter l'efficacité d'un ribozyme donné en ajustant ou sélectionnant le meilleur stabilisateur pour ce ribozyme.



Discussion, Figure 5. Illustrations de la structure secondaire des ribozymes SOFA-HDV avec différents stabilisateurs et de leur activité *in vitro*.

A) Structures secondaires des ribozymes SOFA-HDV avec différents stabilisateurs. Le biosenseur (Bs, boîte noire), le bloqueur (Bl, boîte blanche) et le domaine de reconnaissance (RD, boîte grise) sont montrés pour chaque ribozyme. **B)** Structure secondaire prédite du substrat ARN utilisé pour les essais de coupure *in vitro*. Les sites de liaisons du RD (gris) et du Bs (noir) sont encadrés alors que le site de coupure est marqué d'une flèche. **C)** Représentation graphique du pourcentage de coupure en fonction du temps. Les essais de coupure ont été faits

en présence d'un excès de ribozymes dans les mêmes conditions utilisées au chapitre 3 (Chapitre 3, Figure 2).

2.2.4. Expression et distribution des ribozymes SOFA-HDV *in cellulo*

L'élaboration d'une stratégie basée sur le ribozyme SOFA-HDV pour le ciblage d'un gène d'intérêt excède l'étape de la sélection des meilleurs candidats. La façon d'amener le ribozyme dans la cellule est un facteur déterminant qui ne doit pas être négligé. Pour le moment, la taille (± 100 nucléotides) du ribozyme SOFA-HDV restreint l'acheminement sous forme d'ARN synthétique avec ou sans modifications chimiques en raison du coût de synthèse élevé. L'utilisation de vecteurs d'expression où les ribozymes sont transcrits par l'une des polymérases cellulaires est donc la principale option. Le choix du vecteur et du promoteur doit être fait de façon à permettre la production efficace de ribozymes dans les cellules à l'étude. D'un type cellulaire à un autre, le niveau de transcription à partir d'un promoteur peut varier significativement. En fait, il a été montré que l'efficacité des shARN exprimés avec un promoteur de l'ARN polymérase II varie énormément d'un type cellulaire à un autre, principalement parce que le niveau de transcription de ces promoteurs n'est pas constant dans chaque type cellulaire (Lebbink et al., 2011)(Boden et al., 2003)(Boden et al., 2003). Il a aussi été démontré dans le passé que l'étape limitante dans le ciblage d'un ARN dans une cellule était principalement la mise en contact des ribozymes et de leur substrat (Kato et al., 2001). Il faut donc permettre aux ribozymes de se retrouver dans le même compartiment cellulaire que l'ARN visé.

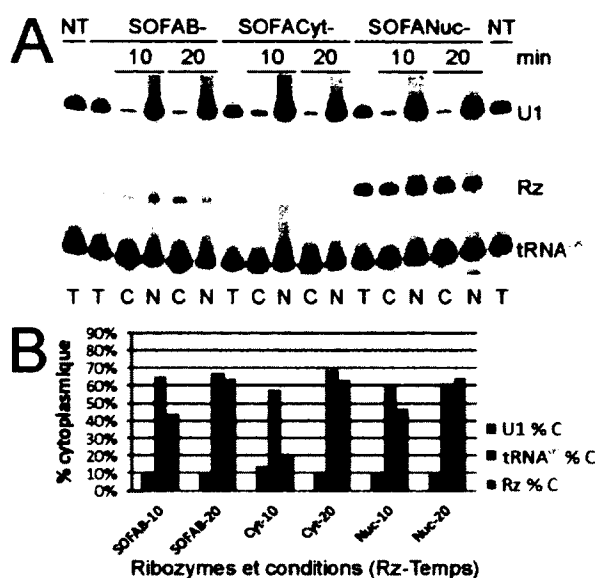
Au fil des années, différents promoteurs (ARN pol II ou ARN pol III) ont été utilisés pour diriger l'expression des ribozymes SOFA-HDV en cellules de mammifères. Les promoteurs pCMV (Bergeron et Perreault, 2005; Laine et al., 2011), ptRNA^{val} (D'Anjou et al., 2011; Levesque et al., 2010), pU6 (Ben Aissa et al., 2012) et pH1 (Motard et al., 2011; Robichaud et al., 2008; Scarborough et al., Soumis) ont tous permis le ciblage spécifique d'un ARN d'intérêt via l'activité d'un

ribozyme SOFA-HDV. Les promoteurs de l'ARN polymérase III (pU6 et pH1) m'apparaissent comme les meilleures options pour l'expression des ribozymes SOFA-HDV. Comme je l'ai démontré, l'ajout de motifs externes au ribozyme semble importuner son activité, du moins *in vitro*. Il semble donc préférable que les ribozymes soient transcrits en minimisant les ajouts (ex : coiffe, queue poly(A) ou ARNt). De plus, des études sur l'expression des shARN ont montré que l'utilisation de tels promoteurs est la meilleure option (Boden et al., 2003).

Il est parfois nécessaire de générer l'accumulation des ribozymes au niveau du cytoplasme, comme c'était le cas pour le ciblage du VHC. Il est plutôt bien accepté dans la littérature que l'expression de ribozymes ou d'aptamères par l'entremise d'un promoteur de l'ARN polymérase III résulte en une localisation majoritairement nucléaire des transcrits (Bertrand et al., 1997; Kato et al., 2001). Par contre, il existe certains motifs (ex. : les ARNt ou des tiges-boucles comme les pre-miARN) capables d'être recrutés par des voies d'export vers le cytoplasme. L'un des facteurs les mieux connus est le transporteur Exportin 5, qui reconnaît des tiges-boucles aux bouts francs ou avec une extrémité 3' protubérante et qui permet leur export vers le cytoplasme (Gwizdek et al., 2003; Lund et al., 2004). Il a été démontré que la fusion d'un ribozyme et d'un ARNt (ptRNA^{val}) permet l'export des ribozymes hybrides vers le cytoplasme (Kato et al., 2001; D'Anjou et al., 2004). Nous avons utilisé cette stratégie pour le projet VHC, puisqu'elle avait déjà été utilisée avec succès pour le ribozyme HDV (D'Anjou et al., 2004). Lors du ciblage du VHC, en plus d'avoir confirmé l'expression de chaque ribozyme par buvardage Northern, j'ai réalisé des expériences préliminaires de fractionnement cellulaire qui ont confirmé que la majeure partie des ribozymes tRNA^{val}-SOFA-HDV se retrouvaient au cytoplasme.

Comme mentionné dans la section précédente, j'ai tenté d'utiliser le stabilisateur comme motif de localisation. L'idée était de favoriser (SOFACyt) ou de défavoriser (SOFANuc) la reconnaissance des ribozymes par l'Exportin 5, en me basant sur les motifs reconnus par cette protéine lors de l'export de certains

transcrits de l'ARN polymérase III (Gwizdek et al., 2003; Zeng et Cullen, 2004). Ce faisant, le ribozyme SOFACyt aurait pu s'accumuler au cytoplasme et le ribozyme SOFANuc au noyau. La localisation de trois versions du stabilisateur a été testée dans une expérience préliminaire de fractionnement cellulaire semblable à celles réalisées pour les ribozymes tRNA^{val}-SOFA-HDV (Discussion, Figure 6). Ces remaniements du stabilisateur n'ont malheureusement pas pu moduler la localisation tel qu'espéré au départ, étant donné que les ribozymes SOFANuc et SOFAB semblent se retrouver au niveau du cytoplasme. L'explication de ce résultat peut être de multiples natures. Au point de départ, l'export par l'entremise de l'Exportin 5 n'était qu'une hypothèse de travail. Les modifications apportées au stabilisateur ne sont peut être pas adéquates ou encore ce transporteur n'est peut être pas un élément clé dans les déplacements intracellulaires de ces ribozymes. Pour sa part, le ribozyme SOFACyt était si peu exprimé, que sa distribution cellulaire ne pouvait être déduite avec précision. Le fait le plus intéressant et le plus important de l'ensemble de cette expérience est que les trois ribozymes utilisés semblent avoir des niveaux d'expression ou d'accumulation très différents (voir extrait total, Discussion, Figure 6). Nous avons également observé des variations similaires, mais de moindre amplitude, dans l'expression des ribozymes tRNA^{val}-SOFA-HDV dirigés contre le VHC (données non-présentées). Il faut donc retenir que des changements dans la séquence des ribozymes, autant dans les régions variables (RD, Bs ou Bl) que dans les régions invariables (ex : stabilisateur ou motif d'export), peuvent affecter le niveau d'expression des ribozymes en cellules, en plus d'affecter l'activité de coupure (voir section précédente). Cette donnée pourrait représenter un élément d'optimisation afin d'augmenter l'expression d'un ribozyme SOFA-HDV qui serait, par exemple, développé pour une thérapie.



Discussion, Figure 6. Expression et distribution cellulaire de trois ribozymes SOFA-HDV avec des stabilisateurs différents.

Détection (A) et quantification (B) par buvardage Northern des ribozymes SOFA-HDV avec des stabilisateurs variés. Des extraits d'ARN total (T), de la fraction cytoplasmique (C) ou de la fraction nucléaire (N) ont été faits à partir de cellules HEK293 exprimant l'un des trois différents ribozymes (SOFAB, SOFACyt et SOFANuc, pour la séquence de chaque ribozyme, voir Discussion, Figure 5A). L'expression transitoire des ribozymes a été faite à partir du promoteur de la polymérase III, pH1 (pCMV-EGFP-pH1, Annexe 1). Les ARN U1 (noyau) et ARN^{lys} (cytoplasme) ont été utilisé pour contrôler la qualité du fractionnement cellulaire. Le temps d'incubation (min) des cellules dans le tampon de lyse est indiqué pour chaque fraction. Un temps d'incubation trop long augmente la contamination de la fraction cytoplasmique par des ARN du noyau alors qu'un temps trop court aura l'effet inverse sur la fraction nucléaire. Des extraits d'ARN provenant de cellules non-transfectées (NT) ont servis de contrôle pour la spécificité des sondes. Le pourcentage cytoplasmique de chaque ARN (U1, ARN^{lys} et ribozyme) a été calculé pour chacune des conditions en divisant le signal de la fraction cytoplasmique par la somme des fractions cytoplasmique et nucléaire ($\% \text{ cytoplasmique} = 100 \times C \div (C+N)$). Les conditions du buvardage Northern sont les mêmes que celles présentées au chapitre 2 (Erreur ! Source du renvoi introuvable.).

En plus des différences d'expression que l'on pourrait associer à des stabilités intracellulaires différentes, il y a un autre facteur critique pour l'expression des ribozymes SOFA-HDV par les promoteurs de l'ARN polymérase III. En fait,

chez les mammifères, le terminateur de transcription de cette ARN polymérase est une série de quatre T ou plus (Dieci et al., 2007). Avec les trois U qui servent de lien entre le stabilisateur et le biosenseur (Discussion, Figure 1), il suffit que le biosenseur commence par un U pour qu'il y ait un terminateur de la transcription dans l'ADN. D'ailleurs, nous avons retiré un ribozyme du projet VHC, car il n'était pas du tout exprimé en raison de son biosenseur qui commençait par un U. La liaison entre le stabilisateur et le biosenseur pourrait sûrement être mutée de façon à permettre l'expression d'un tel ribozyme. Par contre, lorsque ce sont le biosenseur, le bloqueur ou le domaine de reconnaissance qui contiennent quatre U consécutifs ou plus, il n'y a pas vraiment de solution. Il est donc très important de s'assurer qu'il n'y a pas plus de trois U consécutifs dans les ribozymes avant de les tester en cellules avec ce genre de système d'expression.

En terminant cette section, il est important de rappeler que le choix du système d'expression est un élément clé dans l'efficacité d'un système de ciblage génique, peu importe l'outil moléculaire utilisé. Pour le moment, le promoteur pH11 semble la meilleure option de départ pour les ribozymes SOFA-HDV. Par contre, l'activité du promoteur U6 serait plus élevée (Makinen et al., 2006b). On doit donc garder en tête qu'une fois un ribozyme intéressant identifié, tester différents modes d'expression pourrait aider à augmenter l'inhibition observée.

3. Comment pré-sélectionner les ribozymes SOFA-HDV

Bien que le chapitre 4 ne présente pas de résultat biochimique, il demeure pertinent pour cette thèse. En fait, il présente la méthode pour la pré-sélection des ribozymes SOFA-HDV dans le but de cibler un ARN d'intérêt. Cet article résume bien où nous en sommes avec la technologie du ribozyme SOFA-HDV. Ainsi, dans les lignes qui suivent, la méthode proposée sera discutée de manière à faire ressortir comment aborder un nouveau projet de ciblage de gène avec ce ribozyme.

Le point de départ sera toujours de bien s'informer sur l'ARN que l'on veut cibler. Quelles sont ses particularités dont il faut tenir compte? Par exemple, est-ce que nous ciblons un ARNm parmi plusieurs isoformes différents? Lorsqu'on décide de cibler spécifiquement un ARNm ou une partie de celui-ci, il faut s'assurer que la séquence n'est pas partagée par plusieurs autres ARN. Il faut également tenir compte de l'origine de cet ARN. Comme je l'ai mentionné dans les chapitres 1 et 2, certains virus accumulent beaucoup de mutations au fil des cycles de réplication. Il faut donc privilégier des régions hautement conservées pour éviter l'apparition rapide de mutants résistants à la coupure du ribozyme. De plus, d'autres informations peuvent être utiles au moment de la sélection des ribozymes et du système d'expression, comme la structure de l'ARN cible ou sa localisation cellulaire.

Lorsqu'on connaît bien la cible, il faut identifier tous les sites potentiels de coupure du ribozyme SOFA-HDV dans cet ARN. Une liste préliminaire devrait contenir tous les sites $H_{-1}G_{+1}$. Par exemple, pour un ARN tel que la portion codante du gène NS5B du VHC (1773 nt), cette liste s'élevait à 319 sites possibles (chapitre 3). En tenant compte que les sites $Y_{-2}Y_{-1}G_{+1}$ sont moins favorables à la coupure par le ribozyme, il est possible de retirer un bon nombre de cette liste préliminaire (105 sites dans le cas de NS5B). C'est seulement à ce stade qu'il est nécessaire de dessiner les ribozymes SOFA-HDV correspondant aux sites potentiels de coupure. L'analyse des biosenseurs et des bloqueurs devrait permettre de réduire la liste de 20 à 25 %. Globalement, dans l'exemple de l'ARN NS5B, notre méthode permettait de réduire de près de la moitié le nombre de ribozymes intéressants. Avec un critère encore plus sévère du côté du site de coupure ($A_{-2}A_{-1}G_{+1}$), il aurait été possible d'identifier 28 sites à fort potentiel, ce qui correspondrait à plus ou moins 10 % des sites possibles. Pour différentes raisons, je ne crois pas qu'un critère aussi sévère soit la meilleure option. D'abord, pour un virus comme le VHC une grande majorité des sites possibles au départ aurait été retirée en raison d'un manque de conservation de la séquence. De plus, il y a

plusieurs exemples de ribozymes SOFA-HDV qui ont été relativement efficaces *in cellulo* sans un site de coupure optimal ($A_{-2}A_{-1}G_{+1}$). En résumé, je crois qu'à l'étape de présélection *in silico*, il faut trouver le bon niveau de stringence en fonction de notre cible. Dans le cas d'une cible très courte ou peu conservée, il sera préférable d'être moins strict alors que ce sera l'inverse pour un ARNm de plusieurs milliers de nucléotides.

Aussi efficace qu'elle puisse être, la présélection des ribozymes doit être suivie d'une étape qui teste l'activité des ribozymes pour identifier lequel ou lesquels génèrent l'effet recherché. À ce stade, il est possible de tester l'activité des ribozymes *in vitro* pour ensuite tester les meilleurs candidats dans un essai *in cellulo*. Alternativement, les ribozymes présélectionnés peuvent être directement testés dans l'essai *in cellulo*. Bien qu'il n'ait pas été possible d'observer une corrélation directe entre le niveau d'activité de coupure *in vitro* et l'activité *in cellulo* des ribozymes SOFA-HDV au cours des récents projets de ciblage, je vois tout de même un intérêt pour les essais *in vitro*. Dans un monde idéal, le ribozyme le plus performant *in vitro* serait automatiquement le plus actif une fois en cellules, mais ce ne fût pas le cas pour les projets auxquels j'ai contribué. Il y avait tout de même un élément intéressant à retenir : aucun ribozyme incapable de couper son ARN cible *in vitro* n'a été en mesure d'inhiber l'expression de sa cible dans un essai cellulaire. De cette façon, je crois qu'il ne faut pas rechercher spécifiquement le ribozyme avec le plus grand pourcentage de coupure ou la vitesse la plus élevée, mais simplement des ribozymes qui sont capables de couper leur ARN cible au-delà d'un certain niveau minimum. Il apparaît que l'activité de coupure *in vitro* des ribozymes actifs en cellules est supérieure à un minimum de 10 à 20 % de coupure. Je propose donc la notion d'un seuil d'activité qu'il semble nécessaire afin d'avoir une chance de réussite une fois en cellules. L'idée est de savoir si le ribozyme est en mesure d'adopter sa conformation active en fonction de sa séquence et de celle du site qu'il cible.

De plus, je ne suis pas convaincu de la pertinence de toujours utiliser les substrats d'ARN les plus longs pour ces essais *in vitro*. Comme la question est de savoir si les ribozymes sont en mesure d'effectuer la réaction catalytique, le substrat devrait être plutôt minimaliste de façon à ce que la structure des substrats ne soit pas un enjeu pour l'activité des ribozymes. Un autre groupe travaillant sur le ribozyme «*hairpin*» est arrivé à une conclusion semblable (Shippy et al., 1999). Il est difficile de savoir si la structure d'un ARN *in vitro* est la même que celle retrouvée *in cellulo*. Dans certains cas, l'ARN cible est trop long et on doit se limiter à une seule portion sans savoir l'impact sur l'architecture globale du site visé. Selon moi, une analyse approfondie de la structure *in vitro* représente beaucoup trop de travail pour l'impact que cela apporte à la sélection des ribozymes. De plus, l'interaction de l'ARN avec des protéines et différents mécanismes cellulaires (ex. : la maturation, la traduction ou le transport) influencent grandement l'accessibilité relative le long d'un ARN en cellules. Je ne crois pas qu'il y ait un modèle *in vitro* parfait pour reproduire les conditions d'accessibilité retrouvées à l'intérieur d'une cellule, alors il est préférable de garder un plus grand nombre de ribozymes à évaluer *in cellulo*. Qui plus est, nous avons démontré à plusieurs reprises qu'il est possible de tester un grand nombre de ribozymes en cellules (Levesque et al., 2010; Motard et al., 2011; Scarborough et al., soumis). La comparaison des projets de ciblage du VHC et du VIH supportent également ma proposition. Dans le cas du VIH, contrairement au VHC, nous avons directement testé *in cellulo* la série de ribozymes présélectionnés. Par la suite, des essais de coupures plus détaillés ont été réalisés avec notre candidat vedette (les autres ribozymes ont été testés *in vitro*, données non-présentées). L'énergie a été dirigée en faveur des essais d'inhibition du virus plutôt que sur un nombre élevé d'expériences préliminaires *in vitro*. Les résultats de ce projet ont été tout aussi bons, sinon meilleurs, que pour le VHC. En terminant, les essais de coupures *in vitro* permettent d'enlever de la liste les ribozymes qui sont totalement inactifs ou qui ont tendance à s'auto-couper.

4. L'attrait majeur du module SOFA : la spécificité

La spécificité du ribozyme SOFA-HDV a été étudiée *in vitro* de façon élégante (Bergeron et Perreault, 2005; Bergeron et al., 2005). L'un de mes objectifs parallèle à l'application de cette technologie et au développement d'une méthode de design des ribozymes était d'en évaluer la spécificité *in cellulo*. Bien que cet élément n'ait pas été étudié dans le cadre d'une étude dédiée à cette question, une série d'évidences présentées dans cette thèse supportent les données *in vitro*.

4.1. Que savons-nous sur la spécificité du ribozyme SOFA-HDV?

La spécificité suspecte du ribozyme HDV était la question centrale qui a mené au développement du module SOFA (Bergeron et Perreault, 2005). Le résultat fut impressionnant, tant la démonstration de la spécificité de coupure *in vitro* du ribozyme SOFA-HDV était sans équivoque et unique en son genre (Bergeron et al., 2005). Trois expériences clés rapportées dans le chapitre 2 appuient ces études précédentes dans un contexte cellulaire.

Premièrement, l'introduction d'une mutation dans le biosenseur du ribozyme SOFA-HDV-Rz1498 s'est montrée suffisante pour enrayer son activité inhibitrice de la réplication du VIH (**Erreur ! Source du renvoi introuvable.**). Cette première expérience est une bonne indication que la spécificité observée *in vitro* se poursuit en cellules. La limite principale de cette expérience vient du fait que les mutations sont introduites dans la séquence du ribozyme. Comme certains travaux de cette thèse tendent à le montrer, un changement de séquence dans le biosenseur peut affecter significativement l'activité de coupure du ribozyme. La façon la plus représentative d'évaluer la spécificité est de confronter un ribozyme à différents ARN cibles (mutations dans le substrat) comme c'est le cas dans une cellule (un ribozyme parmi une pléthore d'ARN cellulaires). C'est à ce point que les deux

autres expériences interviennent. D'abord, l'expérience d'inhibition d'une série de variants du VIH, dont un possédant une mutation au niveau du site de liaison du biosenseur (sous-type C, pINDIE), montre aussi qu'un seul mésappariement est suffisant pour réduire significativement l'impact du SOFA-HDV-Rz1498 sur la réplication du virus (**Erreur ! Source du renvoi introuvable.**). Il y a donc une grande homologie entre la spécificité du ribozyme SOFA-HDV *in cellulo* et *in vitro*, où un mésappariement réduisait l'activité de 4 à 15 fois (Bergeron et al., 2005). Ensuite, l'expérience de puce à ADN ne montre pas d'impact majeur sur l'expression des gènes cellulaires, supportant ainsi le faible potentiel de non-spécificité du ribozyme (**Erreur ! Source du renvoi introuvable.**). En fait, l'expression d'un nombre très restreint de gènes a été affectée par l'expression du ribozyme SOFA-HDV et les variations n'atteignaient pas un facteur deux, à la hausse ou à la baisse. Il faut noter ici que l'analyse du ribozyme SOFA-HDV-Rz1498 avec l'application web Ribosubstrate a identifié des gènes pouvant être des cibles secondaires pour le ribozyme, dont trois avec des appariements quasi parfaits (trois paires de bases «*wooble*» au niveau du biosenseur, Discussion, Tableau 1). L'expression d'aucun de ces trois gènes ne semble avoir été affectée lors de cette expérience. En fait, seulement deux gènes prédits par Ribosubstrate ont vu leur profil d'expression changer et c'était à la hausse. Leur pointage était supérieur à 300, signifiant la présence d'au moins trois mésappariements entre le ribozyme et l'ARNm. À l'intérieur de ses limites, cette expérience est très encourageante pour la poursuite des travaux avec ce ribozyme, mais également pour l'utilisation du ribozyme SOFA-HDV en général.

Discussion, Tableau 1. Liste partielle^a des gènes prédits par Ribosubstrate comme potentiellement ciblés par le SOFA-HDV-Rz1498.

Result: SOFA-HDV-Rz1498

Search parameters	
BS sequence (5'→3')	AAGGGTACTA
BS mismatch	3
P1 sequence (5'→3')	TTCTGT
Minimum spacer length	1
Maximum spacer length	10
cDNA bank	Human NCBI build 36.2 mRNAs

Perfect match in HIV							
Id	Gene name	P1 sequence perfect match	BS sequence perfect match	Spacer	BS5 mismatch	BS5 wobble	Score
HIV-1	Gag	GCAGGAA	TAGTACCCTT	4	0	0	0
Bs with wobble							
Id	Gene name	P1 sequence match	BS sequence match	Spacer	BS5 mismatch	BS5 wobble	Score
2571	NM_000817	GCAGGAA	TAGTGTCTT	1	0	3	30
221477	NM_152734	GCAGGAA	TGGTGCCTT	7	0	3	34
1238	NM_001296	GCAGGAA	TGGTGTCTT	8	0	3	34
Bs with mismatches and wobbles							
Id	Gene name	P1 sequence match	BS sequence match	Spacer	BS5 mismatch	BS5 wobble	Score
2044	NM_004439	GCAGGAA	AAGTACC TTT	2	1	1	110
2044	NM_182472	GCAGGAA	AAGTACC TTT	2	1	1	110
440823	XM_001132128	GCAGGAA	TAGTATCCTA	4	1	1	110
8816	NM_003861	GCAGGAA	GAGTACC TTT	9	1	1	114
391697	XR_016964	GCAGGAA	GAGTAC TCTT	7	1	1	114
10390	NM_001007794	GCAGGAA	AAGTGTCTT	1	1	2	120
10390	NM_006090	GCAGGAA	AAGTGTCTT	1	1	2	120
9697	NM_012288	GCAGGAA	TAGTGTCTT	3	1	2	120
160419	NM_152589	GCAGGAA	TTGTGCCTT	5	1	2	120
134829	NM_001010852	GCAGGAA	TAGTCTCTT	5	1	2	120
9927	NM_014874	GCAGGAA	TGGGACCTT	5	1	2	120
56692	NM_020242	GCAGGAA	GAGTGTCTT	5	1	2	120
730078	XM_001132314	GCAGGAA	AAGTGTCTT	1	1	2	120
10188	NM_001010938	GCAGGAA	AGGTGCCTT	7	1	2	124
10188	NM_005781	GCAGGAA	AGGTGCCTT	7	1	2	124
115795	NM_173472	GCAGGAA	TGGTACTCTT	9	1	2	124
150946	XM_097977	GCAGGAA	TGGAGCCTT	7	1	2	124

^a Seulement les 20 gènes avec le pointage le plus faible (plus près de la cible parfaite) sont montrés dans ce tableau.

4.2. Une spécificité comparable ou meilleure que les autres outils moléculaires?

Les trois expériences clés que nous avons réalisées et qui nous informent sur la spécificité du ribozyme SOFA-HDV peuvent paraître limitées. Il faut toutefois savoir que cela représente tout de même plus de données que ce qui est connu pour

l'ensemble des autres ribozymes utilisés comme ciseaux moléculaires. En fait, il y a bien peu d'études de la spécificité des ribozymes à l'intérieur des cellules, contrairement aux siARN ou shRNA (Fedorov et al., 2006; Grimm et al., 2006; Jackson et al., 2003; Tschuch et al., 2008; Yi et al., 2005). J'ai retrouvé principalement deux démonstrations de la spécificité d'autres ribozymes dans la littérature. Premièrement, il y a le ciblage d'un ARNm issu d'une recombinaison anormale ayant lieu dans certains cancers entre les gènes BCR et ABL par des ribozymes M1GS (Cobaleda et Sanchez-Garcia, 2000). Ce remaniement génomique donne lieu à deux fusions qui diffèrent par la portion du gène BCR qui se fusionne en 5' du gène ABL. Les ribozymes M1GS dessinés pour cibler l'une ou l'autre des jonctions étaient spécifiques à leur cible, même si les deux sites de coupures incluaient une même séquence de neuf nucléotides (portion du gène ABL). Dans un second temps, des études démontrent que les ribozymes «*hammerhead*» et «*hairpin*» peuvent discriminer un polymorphisme d'un seul nucléotide («*single-nucleotide polymorphism*», SNP) lorsque situé au niveau du site de coupure minimal *in vitro* (Drenser et al., 1998; Millington-Ward et al., 1999) et *in vivo* (Lewin et al., 1998). Par contre, bien peu d'études ont été menées au niveau des bras de liaison. Il est toutefois probable que l'effet des variations de séquences dans ces régions de l'ARN cible sera moins drastique qu'à proximité du site de coupure. Pour sa part, le ribozyme «*hammerhead*» tolérerait mal des mésappariements (Hertel et al., 1996), ce qui suggérerait une spécificité accrue *in cellulo*. Cela demeure toutefois hypothétique, car il ne semble pas y avoir de données sur le sujet. Outre ces quelques études, certaines rapportent, par exemple, peu de signes de cytotoxicité au niveau des cellules ayant reçues des ARN guides (EGS) pour cibler un ARNm avec la RNase P endogène (Kim et Liu, 2007). Dans l'ensemble, cela laisse un vide dans la littérature concernant la spécificité des ribozymes *in cellulo*.

Bien qu'à plusieurs reprises il a été rapporté que l'interférence par l'ARN (siARN ou shARN) donnait lieu à du ciblage non-spécifique, le shRNA1498 semble

relativement spécifique. L'effet de l'expression de ce shARN affecte le patron d'expression des gènes de la cellule d'une manière similaire à celle du ribozyme SOFA-HDV-Rz1498. Il est donc possible de penser que la spécificité de ces deux outils dépend en bonne partie de la séquence précise du site 1498, qui semble relativement spécifique aux ARN du VIH. Par contre, le shRNA1498 demeure aussi efficace face au sous type C (pINDIE) que face aux autres variants du VIH. Ainsi, le ribozyme SOFA-HDV est plus sensible aux mutations. Ces résultats sont informatifs pour le potentiel de ces outils comme agent thérapeutique. D'un côté, le shRNA1498 semble capable de contrer aussi efficacement différents sous-type du virus, même lorsqu'une mutation est présente à son site de liaison. De l'autre côté, cela démontre l'importance d'analyser la conservation de la séquence lors du design de ribozyme pour cibler des ARN à haut taux de mutations. L'expérience de puce à ADN devrait être répétée dans le futur avec d'autres couples ribozyme-shARN pour permettre d'approfondir la comparaison des deux technologies. Il serait intéressant de savoir si les ribozymes et les shARN ciblant un site identique génèrent toujours un effet similaire sur l'expression générale des gènes cellulaires.

5. Où se situe le ribozyme SOFA-HDV?

Le principal élément de comparaison des technologies de ciblage génique est le niveau d'inhibition des ARN ciblés. Dans cette catégorie, le champion est assurément l'interférence par l'ARN. La plupart du temps, les ribozymes sont moins efficaces, bien que certains ont atteint des niveaux d'inhibition qui n'ont rien à envier à ceux de siARN ou de shARN. Il y a aussi l'exemple d'un ribozyme plus performant que son ARN interférent correspondant (Muller-Kuller et al., 2009). Le ribozyme SOFA-HDV n'est peut être pas plus efficace que les ARN interférents, mais il est comparable aux autres ribozymes.

Tel que présenté dans l'introduction, beaucoup de travaux ont été faits pour améliorer les performances des technologies utilisant les M1GS ou EGS et des

ribozymes «*hammerhead*» ou «*hairpin*» (pour une revue voir (Puerta-Fernandez et al., 2003; Schubert et Kurreck, 2004)). Je vois le ribozyme SOFA-HDV comme un outil fraîchement développé sur lequel très peu d'optimisation a été faite. Après les preuves de concepts et l'établissement d'une méthode de design, l'optimisation de ses performances est le prochain défi. Comme il a été fait pour les ribozymes M1GS et «*hairpin*», la sélection de domaines catalytiques plus performants pourrait apporter une piste de solution (Barroso-DelJesus et Berzal-Herranz, 2001; Kilani et al., 2000; Schubert et Kurreck, 2004). Des expériences de sélection *in vitro* sur le corps catalytique du ribozyme HDV ont déjà démontré qu'il est possible de changer certains de ces nucléotides sans affecter l'activité de coupure et même dans certains cas de l'augmenter (Nehdi et Perreault, 2006; Nehdi et al., 2007). En fait, une mutation issue de l'expérience de sélection *in vitro* a été testé avec un ribozyme SOFA-HDV ciblant le VHC, mais il n'y avait pas de gain d'activité de coupure *in vitro*. Je crois donc que ce genre d'optimisation devrait se faire à partir d'un candidat relativement actif tel que le ribozyme SOFA-HDV-Rz1498. La séquence du ribozyme (Bs, le BI et le RD) devrait être tenue en compte puisque l'idée est d'obtenir un ribozyme dont tous les éléments sont optimaux pour fonctionner ensemble.

Un autre élément qui a été évalué, aussi bien pour les ribozymes que pour les ARN interférents, est la longueur des domaines de liaisons ((Dias et Stein, 2002; Hampel et al., 1997; Peracchi, 2004; Schubert et Kurreck, 2004). Dans le cas du ribozyme SOFA-HDV, nous savons que d'allonger le domaine de reconnaissance (RD) n'est pas une option (Nishikawa et al., 1999). D'un autre côté, différentes longueurs de biosenseur (Bs) ont été testées *in vitro* et il semble qu'un biosenseur plus long favorise la coupure (données non-présentées et (Bergeron et Perreault, 2005)). L'effet de la longueur du biosenseur n'a jamais été évalué exhaustivement *in cellulo*. Une augmentation du nombre de paires de bases entre le ribozyme et son substrat devrait normalement favoriser la coupure du premier turnover, mais nuire à la relâche du produit et donc à la réalisation de

multiples turnovers. A première vue, la capacité d'effectuer plusieurs turnovers est importante pour le ciblage d'ARN en cellules. Le ribozyme SOFA-HDV peut effectuer plus d'un turnover *in vitro* et *in cellulo* (Bergeron et Perreault, 2005; Fiola et al., 2006). Dans le cas spécifique du ribozyme SOFA-HDV-Rz1498, son habileté à couper *in vitro* plusieurs molécules de substrats s'est montrée plutôt limitée en condition d'excès de substrat par rapport au ribozyme (données non-montrées). Ce ribozyme démontre tout de même un effet de plus de 50 % sur la réplication du VIH. Il est donc possible d'envisager qu'une augmentation de la longueur du biosenseur, même au dépend du potentiel de multiples turnovers, pourrait augmenter l'activité d'un ribozyme SOFA-HDV *in cellulo*. Afin de réduire l'impact sur le turnover d'une trop forte association entre le ribozyme et sa cible, le biosenseur allongé pourrait, par exemple, inclure une courte séquence non-appariée. Un concept similaire s'est montré efficace pour un DNAzyme (Yi et Liu, 2011). Dans cette étude, le bras de liaison 5' du «10-23 DNA enzyme» a été allongé en prenant soin d'inclure une hernie («*bulge*») entre la séquence de liaison initiale et l'extension. Cette stratégie a permis d'augmenter l'activité du DNAzyme en question. Au final, l'intérêt d'une activité plus élevée de l'effecteur, quel qu'il soit, est de permettre la déduction de sa concentration active et, du même coup, de favoriser son utilisation autant en rapport aux effets secondaires qu'à sa livraison (Aagaard et Rossi, 2007).

Sans nécessairement changer le ribozyme, il y a plusieurs stratégies qui pourraient être intéressantes pour l'utilisation des ribozymes SOFA-HDV. L'effet inhibiteur contre un virus, mais aussi contre un ARNm, pourrait être augmenté par l'expression de deux ribozymes ou plus en multiplexe dirigés contre la cible (Liu et al., 2009). Un premier ribozyme pourrait aussi viser un ARN viral alors qu'un second cible un gène cellulaire essentiel pour la réplication (Bian et al., 2012). La combinaison de plusieurs technologies (ribozyme, shARN et aptamère) est aussi une façon d'augmenter le potentiel d'une stratégie antivirale (Li et al., 2005). Il serait intéressant de mesurer l'efficacité des ribozymes SOFA-HDV dans un tel

contexte. Une autre avenue qui mériterait d'être investiguée est l'efficacité des ribozymes SOFA-HDV livrés aux cellules sous forme d'ARN plutôt qu'avec un vecteur d'expression. Il y a déjà de nombreux exemples de ciblage génique avec des ribozymes ou des EGS synthétiques, dont certaines études cliniques (Drude et al., 2007; Ma et al., 2000). L'activité et la stabilité de ces effecteurs ont été améliorées par l'incorporation judicieuse de modifications chimiques. Évidemment, avec la taille actuelle du ribozyme SOFA-HDV et le coût très élevé de production pour de tels ribozymes synthétiques, cette option n'est pas encore envisageable dans le cadre d'une thérapie. Par contre, les technologies évoluent sans cesse et il est difficile de prédire ce que l'avenir réserve à ce domaine toujours d'actualité qu'est la synthèse d'acides nucléiques. Des recherches fondamentales, pré-cliniques et cliniques présentement en cours, pourraient mener à des avancées majeures au niveau de la réduction du coût de production, des qualités biochimiques ou de la livraison des drogues à base d'acides nucléiques (Bennett et Swayze, 2010). En y pensant bien, il y a quelques années, le séquençage de plusieurs millions d'ARN n'était pas réaliste, alors que maintenant c'est une option facilement accessible.

Finalement, les expériences utilisant les ribozymes SOFA-HDV ont surtout été menées *in vitro* ou *in cellulo*. Il y a bien eu des exemples *in vivo*, mais cela demeure bien limité (D'Anjou et al., 2011; Motard et al., 2011). Le ribozyme SOFA-HDV-Rz1498 dirigé contre le VIH possède un bon potentiel antiviral. Des ribozymes «*hammerhead*» ont déjà été testés dans des modèles *ex vivo* pour l'inhibition du VIH (Rossi, 2000), même que des essais cliniques ont déjà été réalisés avec le ribozyme «*hammerhead*» OZ1 (Mitsuyasu et al., 2009). Ces modèles seraient parfaits pour valider le potentiel du ribozyme SOFA-HDV. Alternativement, le concept du module SOFA pourrait être utilisé sur les meilleurs ribozymes «*hammerhead*» identifiés dans ces travaux. La preuve de concept du module SOFA sur différents ARN catalytiques a été faite *in vitro*, mais n'a pas été investiguée *in cellulo* ou *in vivo*. Les bénéfices du module SOFA pourraient être

avantageux, alors que les ribozymes testés en essais cliniques se sont avérés être peu efficaces (Mitsuyasu et al., 2009; Morrow et al., 2012).

CONCLUSION

Depuis leur découverte, les ribozymes suscitent de l'intérêt comme outils moléculaires et thérapeutiques. Leur potentiel pour différentes fonctions, dont le ciblage génique, ne fait aucun doute. Des essais cliniques montrent qu'ils sont généralement bien tolérés, mais que leur effet demeure limité. Le développement d'une nouvelle génération de ribozymes efficaces et spécifique est l'un des éléments nécessaire pour leur utilisation thérapeutique.

Les travaux de cette thèse ont permis de faire progresser le ribozyme SOFA-HDV comme outil moléculaire. D'un côté, nous avons présentés deux études où un grand nombre de ribozymes ont été criblés pour cibler des ARN viraux du VHC et du VIH. Ces travaux montrent qu'il est possible d'identifier un ribozyme SOFA-HDV efficace pour inhiber la réplication d'un virus dans un modèle cellulaire pertinent. De l'autre côté, une analyse de l'éventail de ribozymes SOFA-HDV testés jusqu'à maintenant nous a permis de mieux comprendre ce ciseau moléculaire. Une étude plus approfondie nous a permis d'identifier des éléments clés dans l'activité du ribozyme, dont l'intégration fournit une méthode pour la pré-sélection des ribozymes adaptée aux réalités du module de spécificité. Dans l'ensemble, les travaux présentés dans cette thèse supportent le potentiel pour le ciblage génique et aspirent à confirmer la grande spécificité du ribozyme SOFA-HDV.

En terminant, la recherche d'agents thérapeutiques à base d'acides nucléiques est plus que jamais d'actualité. L'interférence par l'ARN semble voler la vedette aux outils plus anciens comme les oligonucléotides antisens ou les ribozymes, mais les développements, entre autres au niveau de la livraison, profiteront à toutes ces technologies. L'arrivée de nouvelles innovations telles que le module SOFA aura certainement un impact sur les thérapies futures.

REMERCIEMENTS

Je voudrais tout d'abord remercier le Dr Jean-Pierre Perreault pour son excellente supervision. La première entrevue que j'ai passée avec lui a été un point tournant important dans ma formation au baccalauréat en biochimie, qui jusqu'alors n'était qu'une étape vers un autre champ d'étude. En m'accueillant pour mon premier stage coopératif à l'hiver 2004, il m'a permis de découvrir le monde stimulant de la recherche. En tant que directeur de recherche, il m'a donné l'opportunité de travailler dans un environnement scientifique de qualité, mais surtout de développer ma façon de réfléchir et de voir la science. Il avait toujours les mots dans les moments difficiles pour me convaincre de me retrousser les manches et de travailler encore plus fort. Jean-Pierre représente mon modèle par excellence de ténacité, de persévérance et d'ardeur au travail. Je le remercie sincèrement pour son support et pour tous les bons moments qu'il m'a permis de vivre. Je serai toujours fier d'avoir été un membre du labo JPP.

J'aimerais remercier spécialement Dominique Lévesque, ancien assistant de recherche au laboratoire. Dans les hauts comme dans les bas de la vie d'un étudiant gradué, une oreille attentive vaut son pesant d'or. Pour toutes ces fois où je me suis confié à toi à propos de tout et de rien, pour tous ces moments où tu trouvais les mots pour garder mon moral à flot et pour toutes ces discussions sans queue ni tête qui nous faisaient oublier nos soucis, Dominique, je te remercie. Je ne pourrai jamais assez te dire comment j'ai apprécié travailler avec toi. Tu étais une encyclopédie sur pattes, un vrai «wiki Lévesque» du laboratoire. Tes trucs et conseils nous ont tous aidés à un moment ou à un autre. J'espère que nos chemins se croiseront de nouveau dans le futur, car je ne peux pas me passer de toutes tes anecdotes plus rocambolesques les unes que les autres.

Je prends aussi le temps de remercier François Bolduc, assistant de recherche au laboratoire, pour les excellentes discussions scientifiques. Pour moi,

François incarne le calme et la sagesse. Ses réflexions toujours posées et son assiduité au travail sont contagieuses. J'ai eu beaucoup de plaisir à travailler avec toi, mais également de te côtoyer en dehors du laboratoire. Que ce soit au hockey «cosom», en voyage à Seattle ou encore pour déguster du bon vin, ce fût un privilège d'être à tes cotés. J'espère qu'on gardera le contact pour échanger autant sur la science que le hockey sans oublier de partager un bon vin.

Au cours des huit dernières années, j'ai eu l'occasion de croiser le chemin d'un nombre impressionnant de personnes dans le laboratoire. Certains ont eu un impact direct sur mes travaux de recherches ou sur mon évolution en tant que scientifique. Je remercie les Drs Lucien Jr Bergeron et Gilles Robichaud de m'avoir supervisé à mes deux premiers stages au laboratoire. J'ai goûté à la recherche pour la première fois en votre compagnie, vous avez su rendre l'expérience tellement attrayante que je ne me suis pas arrêté depuis. Je veux aussi remercier trois anciens étudiants au PhD, les Drs Jonathan Ouellet, Jonathan Perreault et Atef Nehdi, qui, avec Lucien, m'ont vraiment montré ce qu'était avoir la piqure pour la science. Les «lab meetings» en leur compagnie étaient incroyablement stimulant pour un novice comme moi. J'aimerais remercier le Dr Jean-Luc Parent de m'avoir accueilli dans son laboratoire pour effectuer un stage en co-direction avec le Dr Jean-Pierre Perreault. Je veux aussi remercier Jean-Denis Beaudoin, mon partenaire de PhD. Ses qualités de scientifique ont toujours été appréciées autant que son coté enjoué et souriant. J'ai eu beaucoup de plaisir à jouer au hockey avec toi un soir et à discuter de science le lendemain matin. J'espère qu'on gardera le contact encore longtemps. Lors de mes études graduées, j'ai eu le plaisir de superviser deux excellents étudiants, Noura Mazloum et Samuel Rouleau. J'ai bien aimé vous transmettre mes connaissances tout en goûtant à nouveau aux plaisirs du «coaching». Je dois aussi vous remercier pour votre apport à l'avancement de mes travaux, tout particulièrement pour Samuel et sa contribution à l'un de mes articles. J'aimerais également remercier les membres de mon comité d'évaluation d'examen pré-doctoral, les Drs Xavier Roucou, Guylain

Boissonneault, Andrew Mouland et Jean-Pierre Perreault. Cette étape peut être un supplice pour certains étudiants, mais pour moi ce fût une expérience extrêmement stimulante et formatrice.

J'aimerais remercier tous les membres présents et passés du laboratoire. Je pourrais probablement écrire plusieurs lignes pour la majorité d'entre vous, mais je devrai me contenter d'un gros MERCI. Donc, sans ordre particulier je dis merci à Patricia Landry, Audrey Dubé, Marie-Claude April, Cédric Reymond, Francis P Brière, Marie F Soulière, Julie Motard, Patrick Saint-Pierre, François d'Anjou, Wenxing Xu, Rachel Jodoin, Tamara Giguère, Lubos Bauer, Francis Phaneuf, Anne-Marie Lamontagne, Charith A Purushothama, Jean-Pierre S Glouzon, Mahdi Laaref, Jean-Michel Garant et à tous ceux que j'aurais pu oublier.

Je prends aussi le temps de remercier les organismes subventionnaires pour le financement de mes travaux et particulièrement le Fonds de recherche en santé du Québec (FRSQ) pour ma bourse d'étude au PhD. Je remercie également mes collaborateurs, Robert Scarborough, Ian Chute, la Dre Anne Gagnon et le Dr Rodney Ouellette.

Finalement, j'aimerais remercier du fond du cœur ma mère, Suzanne Boucher, mon père, Vital Lévesque, ma blonde Andréane Cartier et mon petit garçon d'amour, Ludovick Lévesque pour tout leur support inconditionnel pendant toutes ces années d'études. Sans leur aide morale et financière, je ne serais pas l'auteur de cette thèse. Merci de m'avoir toujours épaulé dans tous mes projets et de toujours être là pour moi. Je vous aime plus que tout.

LISTE DE RÉFÉRENCES

- Aagaard, L. et Rossi, J.J. (2007) RNAi therapeutics: principles, prospects and challenges. *Advanced Drug Delivery Reviews* 59(2-3): 75-86.
- Abdellatif, M. (2012) Differential expression of microRNAs in different disease states. *Circulation Research* 110(4): 638-650.
- Akhtar, S., Kole, R. et Juliano, R.L. (1991) Stability of antisense DNA oligodeoxynucleotide analogs in cellular extracts and sera. *Life Sciences* 49(24): 1793-1801.
- Al-Anouti, F. et Ananvoranich, S. (2002) Comparative analysis of antisense RNA, double-stranded RNA, and delta ribozyme-mediated gene regulation in *Toxoplasma gondii*. *Antisense & Nucleic Acid Drug Development* 12(4): 275-281.
- Amarzguioui, M. et Prydz, H. (2004) An algorithm for selection of functional siRNA sequences. *Biochemical and Biophysical Research Communications* 316(4): 1050-1058.
- Ananvoranich, S., Lafontaine, D.A. et Perreault, J.P. (1999) Mutational analysis of the antigenomic trans-acting delta ribozyme: the alterations of the middle nucleotides located on the P1 stem. *Nucleic Acids Research* 27(6): 1473-1479.
- Arbuthnot, P. (2011) MicroRNA-like antivirals. *Biochimica Et Biophysica Acta* 1809(11-12): 746-755.
- Asif-Ullah, M., Levesque, M., Robichaud, G. et Perreault, J.P. (2007) Development of ribozyme-based gene-inactivations; the example of the hepatitis delta virus ribozyme. *Current Gene Therapy* 7(3): 205-216.
- Azmi, A.S., Beck, F.W., Bao, B., Mohammad, R.M. et Sarkar, F.H. (2011) Aberrant epigenetic grooming of miRNAs in pancreatic cancer: a systems biology perspective. *Epigenomics* 3(6): 747-759.
- Bagga, S., Bracht, J., Hunter, S., Massirer, K., Holtz, J., Eachus, R. et Pasquinelli, A.E. (2005) Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell* 122(4): 553-563.
- Bai, Y., Trang, P., Li, H., Kim, K., Zhou, T. et Liu, F. (2008) Effective inhibition in animals of viral pathogenesis by a ribozyme derived from RNase P catalytic

- RNA. *Proceedings of the National Academy of Sciences of the United States of America* 105(31): 10919-10924.
- Barroso-DelJesus, A. et Berzal-Herranz, A. (2001) Selection of targets and the most efficient hairpin ribozymes for inactivation of mRNAs using a self-cleaving RNA library. *EMBO Reports* 2(12): 1112-1118.
- Barroso-delJesus, A., Puerta-Fernandez, E., Romero-Lopez, C. et Berzal-Herranz, A. (2004) An experimental method for selecting effective target sites and designing hairpin ribozymes. *Methods in Molecular Biology (Clifton, N.J.)* 252: 313-325.
- Barroso-DelJesus, A., Puerta-Fernandez, E., Tapia, N., Romero-Lopez, C., Sanchez-Luque, F.J., Martinez, M.A. et Berzal-Herranz, A. (2005) Inhibition of HIV-1 replication by an improved hairpin ribozyme that includes an RNA decoy. *RNA Biology* 2(2): 75-79.
- Bartel, D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116(2): 281-297.
- Beaudoin, J.D. et Perreault, J.P. (2008) Potassium ions modulate a G-quadruplex-ribozyme's activity. *RNA (New York, N.Y.)* 14(6): 1018-1025.
- Ben Aissa, M., April, M.C., Bergeron, L.J., Perreault, J.P. et Levesque, G. (2012) Silencing of amyloid precursor protein expression using a new engineered delta ribozyme. *International Journal of Alzheimer's Disease* 2012: 947147.
- Benimetskaya, L., Loike, J.D., Khaled, Z., Loike, G., Silverstein, S.C., Cao, L., el Khoury, J., Cai, T.Q. et Stein, C.A. (1997) Mac-1 (CD11b/CD18) is an oligodeoxynucleotide-binding protein. *Nature Medicine* 3(4): 414-420.
- Bennasser, Y., Yeung, M.L. et Jeang, K.T. (2007) RNAi therapy for HIV infection: principles and practicalities. *BioDrugs : Clinical Immunotherapeutics, Biopharmaceuticals and Gene Therapy* 21(1): 17-22.
- Bennett, C.F. et Swayze, E.E. (2010) RNA targeting therapeutics: molecular mechanisms of antisense oligonucleotides as a therapeutic platform. *Annual Review of Pharmacology and Toxicology* 50: 259-293.
- Bergeron, L.J., Ouellet, J. et Perreault, J.P. (2003) Ribozyme-based gene-inactivation systems require a fine comprehension of their substrate specificities; the case of delta ribozyme. *Current Medicinal Chemistry* 10(23): 2589-2597.
- Bergeron, L.J. et Perreault, J.P. (2002) Development and comparison of procedures for the selection of delta ribozyme cleavage sites within the hepatitis B virus. *Nucleic Acids Research* 30(21): 4682-4691.

- Bergeron, L.J. et Perreault, J.P. (2005) Target-dependent on/off switch increases ribozyme fidelity. *Nucleic Acids Research* 33(4): 1240-1248.
- Bergeron, L.J., Reymond, C. et Perreault, J.P. (2005) Functional characterization of the SOFA delta ribozyme. *RNA (New York, N.Y.)* 11(12): 1858-1868.
- Bertrand, E., Castanotto, D., Zhou, C., Carbonnelle, C., Lee, N.S., Good, P., Chatterjee, S., Grange, T., Pictet, R., Kohn, D. et al. (1997) The expression cassette determines the functional activity of ribozymes in mammalian cells by controlling their intracellular localization. *RNA (New York, N.Y.)* 3(1): 75-88.
- Bian, Z., Xiao, A., Cao, M., Liu, M., Liu, S., Jiao, Y., Yan, W., Qi, Z. et Zheng, Z. (2012) Anti-HBV efficacy of combined siRNAs targeting viral gene and heat shock cognate 70. *Virology Journal* 9: 275-422X-9-275.
- Birmingham, A., Anderson, E.M., Reynolds, A., Ilsley-Tyree, D., Leake, D., Fedorov, Y., Baskerville, S., Maksimova, E., Robinson, K., Karpilow, J. et al. (2006) 3' UTR seed matches, but not overall identity, are associated with RNAi off-targets. *Nature Methods* 3(3): 199-204.
- Boden, D., Pusch, O., Lee, F., Tucker, L., Shank, P.R. et Ramratnam, B. (2003) Promoter choice affects the potency of HIV-1 specific RNA interference. *Nucleic Acids Research* 31(17): 5033-5038.
- Boulanger, C., Chabot, B., Menard, H.A. et Boire, G. (1995) Autoantibodies in human anti-Ro sera specifically recognize deproteinized hY5 Ro RNA. *Clinical and Experimental Immunology* 99(1): 29-36.
- Branch, A.D. et Robertson, H.D. (1991) Efficient trans cleavage and a common structural motif for the ribozymes of the human hepatitis delta agent. *Proceedings of the National Academy of Sciences of the United States of America* 88(22): 10163-10167.
- Cavusoglu, E., Chen, I., Rappaport, J. et Marmur, J.D. (2002) Inhibition of tissue factor gene induction and activity using a hairpin ribozyme. *Circulation* 105(19): 2282-2287.
- Cerritelli, S.M. et Crouch, R.J. (2009) Ribonuclease H: the enzymes in eukaryotes. *The FEBS Journal* 276(6): 1494-1505.
- Chan, C.W. et Khachigian, L.M. (2009) DNazymes and their therapeutic possibilities. *Internal Medicine Journal* 39(4): 249-251.
- Chan, J.A., Krichevsky, A.M. et Kosik, K.S. (2005) MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. *Cancer Research* 65(14): 6029-6033.

- Chang, S., Chen, W., Yang, J., Xie, Y. et Sheng, W. (2009) Antitumor activity of an adenovirus harboring two therapeutic genes, anti-VEGF ribozyme and human IL-24, in colon cancer. *Molecular Medicine Reports* 2(5): 693-700.
- Chen, C.J., Banerjea, A.C., Harmison, G.G., Haglund, K. et Schubert, M. (1992) Multitarget-ribozyme directed to cleave at up to nine highly conserved HIV-1 env RNA regions inhibits HIV-1 replication--potential effectiveness against most presently sequenced HIV-1 isolates. *Nucleic Acids Research* 20(17): 4581-4589.
- Christian, E.L., Zahler, N.H., Kaye, N.M. et Harris, M.E. (2002) Analysis of substrate recognition by the ribonucleoprotein endonuclease RNase P. *Methods (San Diego, Calif.)* 28(3): 307-322.
- Citti, L. et Rainaldi, G. (2005) Synthetic hammerhead ribozymes as therapeutic tools to control disease genes. *Current Gene Therapy* 5(1): 11-24.
- Cobaleda, C. et Sanchez-Garcia, I. (2000) In vivo inhibition by a site-specific catalytic RNA subunit of RNase P designed against the BCR-ABL oncogenic products: a novel approach for cancer treatment. *Blood* 95(3): 731-737.
- Collison, A., Herbert, C., Siegle, J.S., Mattes, J., Foster, P.S. et Kumar, R.K. (2011) Altered expression of microRNA in the airway wall in chronic asthma: miR-126 as a potential therapeutic target. *BMC Pulmonary Medicine* 11: 29.
- Crooke, S.T. (1999) Molecular mechanisms of action of antisense drugs. *Biochimica Et Biophysica Acta* 1489(1): 31-44.
- D'Anjou, F., Bergeron, L.J., Larbi, N.B., Fournier, I., Salzet, M., Perreault, J.P. et Day, R. (2004) Silencing of SPC2 expression using an engineered delta ribozyme in the mouse betaTC-3 endocrine cell line. *The Journal of Biological Chemistry* 279(14): 14232-14239.
- D'Anjou, F., Routhier, S., Perreault, J.P., Latil, A., Bonnel, D., Fournier, I., Salzet, M. et Day, R. (2011) Molecular Validation of PACE4 as a Target in Prostate Cancer. *Translational Oncology* 4(3): 157-172.
- Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F. et Hannon, G.J. (2004) Processing of primary microRNAs by the Microprocessor complex. *Nature* 432(7014): 231-235.
- Denman, R.B. (1993) Using RNAFOLD to predict the activity of small catalytic RNAs. *BioTechniques* 15(6): 1090-1095.
- Deschenes, P., Lafontaine, D.A., Charland, S. et Perreault, J.P. (2000) Nucleotides -1 to -4 of hepatitis delta ribozyme substrate increase the specificity of

- ribozyme cleavage. *Antisense & Nucleic Acid Drug Development* 10(1): 53-61.
- Deschenes, P., Ouellet, J., Perreault, J. et Perreault, J.P. (2003) Formation of the P1.1 pseudoknot is critical for both the cleavage activity and substrate specificity of an antigenomic trans-acting hepatitis delta ribozyme. *Nucleic Acids Research* 31(8): 2087-2096.
- Dias, N. et Stein, C.A. (2002) Antisense oligonucleotides: basic concepts and mechanisms. *Molecular Cancer Therapeutics* 1(5): 347-355.
- Dieci, G., Fiorino, G., Castelnovo, M., Teichmann, M. et Pagano, A. (2007) The expanding RNA polymerase III transcriptome. *Trends in Genetics : TIG* 23(12): 614-622.
- Dimond. (2012) With Oligonucleotide Drugs, Antisense Arrives but RNAi Has a Long Way to Go. <http://www.genengnews.com/keywordsandtools/print/3/27435/>.
- Drenser, K.A., Timmers, A.M., Hauswirth, W.W. et Lewin, A.S. (1998) Ribozyme-targeted destruction of RNA associated with autosomal-dominant retinitis pigmentosa. *Investigative Ophthalmology & Visual Science* 39(5): 681-689.
- Drude, I., Dombos, V., Vauleon, S. et Muller, S. (2007) Drugs made of RNA: development and application of engineered RNAs for gene therapy. *Mini Reviews in Medicinal Chemistry* 7(9): 912-931.
- Du, L. et Gatti, R.A. (2009) Progress toward therapy with antisense-mediated splicing modulation. *Current Opinion in Molecular Therapeutics* 11(2): 116-123.
- Eder, P.S., DeVine, R.J., Dagle, J.M. et Walder, J.A. (1991) Substrate specificity and kinetics of degradation of antisense oligonucleotides by a 3' exonuclease in plasma. *Antisense Research and Development* 1(2): 141-151.
- Eder, P.S. et Walder, J.A. (1991) Ribonuclease H from K562 human erythroleukemia cells. Purification, characterization, and substrate specificity. *The Journal of Biological Chemistry* 266(10): 6472-6479.
- Elbashir, S.M., Lendeckel, W. et Tuschl, T. (2001) RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes & Development* 15(2): 188-200.
- ENCODE Project Consortium, Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C. et Snyder, M. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489(7414): 57-74.
- ENCODE Project Consortium, Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M.,

- Dermitzakis, E.T. et al. (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 447(7146): 799-816.
- Esau, C., Davis, S., Murray, S.F., Yu, X.X., Pandey, S.K., Pear, M., Watts, L., Booten, S.L., Graham, M., McKay, R. et al. (2006) miR-122 regulation of lipid metabolism revealed by in vivo antisense targeting. *Cell Metabolism* 3(2): 87-98.
- Fabian, M.R., Mathonnet, G., Sundermeier, T., Mathys, H., Zipprich, J.T., Svitkin, Y.V., Rivas, F., Jinek, M., Wohlschlegel, J., Doudna, J.A. et al. (2009) Mammalian miRNA RISC recruits CAF1 and PABP to affect PABP-dependent deadenylation. *Molecular Cell* 35(6): 868-880.
- Fabian, M.R., Sundermeier, T.R. et Sonenberg, N. (2010) Understanding how miRNAs post-transcriptionally regulate gene expression. *Progress in Molecular and Subcellular Biology* 50: 1-20.
- Fedor, M.J. (2000) Structure and function of the hairpin ribozyme. *Journal of Molecular Biology* 297(2): 269-291.
- Fedor, M.J. (2009) Comparative enzymology and structural biology of RNA self-cleavage. *Annual Review of Biophysics* 38: 271-299.
- Fedorov, Y., Anderson, E.M., Birmingham, A., Reynolds, A., Karpilow, J., Robinson, K., Leake, D., Marshall, W.S. et Khvorova, A. (2006) Off-target effects by siRNA can induce toxic phenotype. *RNA (New York, N.Y.)* 12(7): 1188-1196.
- Fedorov, Y., King, A., Anderson, E., Karpilow, J., Ilsley, D., Marshall, W. et Khvorova, A. (2005) Different delivery methods-different expression profiles. *Nature Methods* 2(4): 241.
- Feldstein, P.A., Buzayan, J.M. et Bruening, G. (1989) Two sequences participating in the autolytic processing of satellite tobacco ringspot virus complementary RNA. *Gene* 82(1): 53-61.
- Fennewald, S.M. et Rando, R.F. (1995) Inhibition of high affinity basic fibroblast growth factor binding by oligonucleotides. *The Journal of Biological Chemistry* 270(37): 21718-21721.
- Ferre-D'Amare, A.R. et Doudna, J.A. (2000) Crystallization and structure determination of a hepatitis delta virus ribozyme: use of the RNA-binding protein U1A as a crystallization module. *Journal of Molecular Biology* 295(3): 541-556.

- Fiola, K., Perreault, J.P. et Cousineau, B. (2006) Gene targeting in the Gram-Positive bacterium *Lactococcus lactis*, using various delta ribozymes. *Applied and Environmental Microbiology* 72(1): 869-879.
- Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E. et Mello, C.C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391(6669): 806-811.
- Fontana, L., Fiori, M.E., Albin, S., Cifaldi, L., Giovinazzi, S., Forloni, M., Boldrini, R., Donfrancesco, A., Federici, V., Giacomini, P. et al. (2008) Antagomir-17-5p abolishes the growth of therapy-resistant neuroblastoma through p21 and BIM. *PLoS One* 3(5): e2236.
- Forster, A.C. et Symons, R.H. (1987) Self-cleavage of plus and minus RNAs of a virusoid and a structural model for the active sites. *Cell* 49(2): 211-220.
- Fowler, T., Sen, R. et Roy, A.L. (2011) Regulation of primary response genes. *Molecular Cell* 44(3): 348-360.
- Gambari, R., Fabbri, E., Borgatti, M., Lampronti, I., Finotti, A., Brognara, E., Bianchi, N., Manicardi, A., Marchelli, R. et Corradini, R. (2011) Targeting microRNAs involved in human diseases: a novel approach for modification of gene expression and drug development. *Biochemical Pharmacology* 82(10): 1416-1429.
- Gerlach, W.L., Buzayan, J.M., Schneider, I.R. et Bruening, G. (1986) Satellite tobacco ringspot virus RNA: biological activity of DNA clones and their in vitro transcripts. *Virology* 151(2): 172-185.
- Golden, B.L. (2011) Two distinct catalytic strategies in the hepatitis delta virus ribozyme cleavage reaction. *Biochemistry* 50(44): 9424-9433.
- Gommans, W.M., Haisma, H.J. et Rots, M.G. (2005) Engineering zinc finger protein transcription factors: the therapeutic relevance of switching endogenous gene expression on or off at command. *Journal of Molecular Biology* 354(3): 507-519.
- Grimm, D., Streetz, K.L., Jopling, C.L., Storm, T.A., Pandey, K., Davis, C.R., Marion, P., Salazar, F. et Kay, M.A. (2006) Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature* 441(7092): 537-541.
- Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. et Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* 35(3 Pt 2): 849-857.

- Guvakova, M.A., Yakubov, L.A., Vlodayvsky, I., Tonkinson, J.L. et Stein, C.A. (1995) Phosphorothioate oligodeoxynucleotides bind to basic fibroblast growth factor, inhibit its binding to cell surface receptors, and remove it from low affinity binding sites on extracellular matrix. *The Journal of Biological Chemistry* 270(6): 2620-2627.
- Gwizdek, C., Ossareh-Nazari, B., Brownawell, A.M., Doglio, A., Bertrand, E., Macara, I.G. et Dargemont, C. (2003) Exportin-5 mediates nuclear export of minihelix-containing RNAs. *The Journal of Biological Chemistry* 278(8): 5505-5508.
- Haasnoot, J., Westerhout, E.M. et Berkhout, B. (2007) RNA interference against viruses: strike and counterstrike. *Nature Biotechnology* 25(12): 1435-1443.
- Halatsch, M.E., Schmidt, U., Botefur, I.C., Holland, J.F. et Ohnuma, T. (2000) Marked inhibition of glioblastoma target cell tumorigenicity in vitro by retrovirus-mediated transfer of a hairpin ribozyme against deletion-mutant epidermal growth factor receptor messenger RNA. *Journal of Neurosurgery* 92(2): 297-305.
- Hampel, A., DeYoung, M.B., Galasinski, S. et Siwkowski, A. (1997) Design of the hairpin ribozyme for targeting specific RNA sequences. *Methods in Molecular Biology (Clifton, N.J.)* 74: 171-177.
- Hampel, A. et Tritz, R. (1989) RNA catalytic properties of the minimum (-)sTRSV sequence. *Biochemistry* 28(12): 4929-4933.
- Hampel, A., Tritz, R., Hicks, M. et Cruz, P. (1990) 'Hairpin' catalytic RNA model: evidence for helices and sequence requirement for substrate RNA. *Nucleic Acids Research* 18(2): 299-304.
- Han, J., Lee, Y., Yeom, K.H., Kim, Y.K., Jin, H. et Kim, V.N. (2004) The Drosha-DGCR8 complex in primary microRNA processing. *Genes & Development* 18(24): 3016-3027.
- Hantschel, O., Grebien, F. et Superti-Furga, G. (2012) The growing arsenal of ATP-competitive and allosteric inhibitors of BCR-ABL. *Cancer Research* 72(19): 4890-4895.
- Harries, L.W. (2012) Long non-coding RNAs and human disease. *Biochemical Society Transactions* 40(4): 902-906.
- Haseloff, J. et Gerlach, W.L. (1988) Simple RNA enzymes with new and highly specific endoribonuclease activities. *Nature* 334(6183): 585-591.
- Haseloff, J. et Gerlach, W.L. (1989) Sequences required for self-catalysed cleavage of the satellite RNA of tobacco ringspot virus. *Gene* 82(1): 43-52.

- He, L., Thomson, J.M., Hemann, M.T., Hernando-Monge, E., Mu, D., Goodson, S., Powers, S., Cordon-Cardo, C., Lowe, S.W., Hannon, G.J. et al. (2005) A microRNA polycistron as a potential human oncogene. *Nature* 435(7043): 828-833.
- Heale, B.S., Soifer, H.S., Bowers, C. et Rossi, J.J. (2005) siRNA target site secondary structure predictions using local stable substructures. *Nucleic Acids Research* 33(3): e30.
- Herschlag, D. (1991) Implications of ribozyme kinetics for targeting the cleavage of specific RNA molecules in vivo: more isn't always better. *Proceedings of the National Academy of Sciences of the United States of America* 88(16): 6921-6925.
- Hertel, K.J., Herschlag, D. et Uhlenbeck, O.C. (1996) Specificity of hammerhead ribozyme cleavage. *The EMBO Journal* 15(14): 3751-3757.
- Hori, T., Guo, F., Tanaka, Y. et Uesugi, S. (2001) Design and properties of trans-acting HDV ribozymes with extended substrate recognition regions. *Nucleic Acids Research. Supplement* (2001) (1)(1): 201-202.
- Hori, T., Guo, F. et Uesugi, S. (2006) Addition of an extra substrate binding site and partial destabilization of stem structures in HDV ribozyme give rise to high sequence-specificity for its target RNA. *Nucleosides, Nucleotides & Nucleic Acids* 25(4-6): 489-501.
- Hughes, S.A., Wedemeyer, H. et Harrison, P.M. (2011) Hepatitis delta virus. *Lancet* 378(9785): 73-85.
- Hutvagner, G., McLachlan, J., Pasquinelli, A.E., Balint, E., Tuschl, T. et Zamore, P.D. (2001) A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science (New York, N.Y.)* 293(5531): 834-838.
- International Human Genome Sequencing Consortium. (2004) Finishing the euchromatic sequence of the human genome. *Nature* 431(7011): 931-945.
- Jackson, A.L., Bartz, S.R., Schelter, J., Kobayashi, S.V., Burchard, J., Mao, M., Li, B., Cavet, G. et Linsley, P.S. (2003) Expression profiling reveals off-target gene regulation by RNAi. *Nature Biotechnology* 21(6): 635-637.
- Jackson, A.L., Burchard, J., Schelter, J., Chau, B.N., Cleary, M., Lim, L. et Linsley, P.S. (2006) Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA (New York, N.Y.)* 12(7): 1179-1187.

- Jackson, A.L. et Linsley, P.S. (2010) Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. *Nature Reviews Drug Discovery* 9(1): 57-67.
- Jarczak, D., Korf, M., Beger, C., Manns, M.P. et Kruger, M. (2005) Hairpin ribozymes in combination with siRNAs against highly conserved hepatitis C virus sequence inhibit RNA replication and protein translation from hepatitis C virus subgenomic replicons. *The FEBS Journal* 272(22): 5910-5922.
- Johnson, S.M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K.L., Brown, D. et Slack, F.J. (2005) RAS is regulated by the let-7 microRNA family. *Cell* 120(5): 635-647.
- Joyce, G.F. (2007) Forty years of in vitro evolution. *Angewandte Chemie (International Ed. in English)* 46(34): 6420-6436.
- Kato, Y., Kuwabara, T., Warashina, M., Toda, H. et Taira, K. (2001) Relationships between the activities in vitro and in vivo of various kinds of ribozyme and their intracellular localization in mammalian cells. *The Journal of Biological Chemistry* 276(18): 15378-15385.
- Kato, Y. et Kato, M. (2009) Hedgehog target genes: mechanisms of carcinogenesis induced by aberrant hedgehog signaling activation. *Current Molecular Medicine* 9(7): 873-886.
- Kelemen, O., Convertini, P., Zhang, Z., Wen, Y., Shen, M., Falaleeva, M. et Stamm, S. (2013) Function of alternative splicing. *Gene* 514(1): 1-30.
- Kertsburg, A. et Soukup, G.A. (2002) A versatile communication module for controlling RNA folding and catalysis. *Nucleic Acids Research* 30(21): 4599-4606.
- Khvorova, A., Reynolds, A. et Jayasena, S.D. (2003) Functional siRNAs and miRNAs exhibit strand bias. *Cell* 115(2): 209-216.
- Kilani, A.F., Trang, P., Jo, S., Hsu, A., Kim, J., Nepomuceno, E., Liou, K. et Liu, F. (2000) RNase P ribozymes selected in vitro to cleave a viral mRNA effectively inhibit its expression in cell culture. *The Journal of Biological Chemistry* 275(14): 10611-10622.
- Kim, D.H., Behlke, M.A., Rose, S.D., Chang, M.S., Choi, S. et Rossi, J.J. (2005) Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nature Biotechnology* 23(2): 222-226.
- Kim, K. et Liu, F. (2007) Inhibition of gene expression in human cells using RNase P-derived ribozymes and external guide sequences. *Biochimica Et Biophysica Acta* 1769(11-12): 603-612.

- Kitajima, I., Unoki, K. et Maruyama, I. (1999) Phosphorothioate oligodeoxynucleotides inhibit basic fibroblast growth factor-induced angiogenesis in vitro and in vivo. *Antisense & Nucleic Acid Drug Development* 9(2): 233-239.
- Klebba, C., Ottmann, O.G., Scherr, M., Pape, M., Engels, J.W., Grez, M., Hoelzer, D. et Klein, S.A. (2000) Retrovirally expressed anti-HIV ribozymes confer a selective survival advantage on CD4+ T cells in vitro. *Gene Therapy* 7(5): 408-416.
- Komurian-Pradel, F., Perret, M., Deiman, B., Sodoyer, M., Lotteau, V., Paranhos-Baccala, G. et Andre, P. (2004) Strand specific quantitative real-time PCR to study replication of hepatitis C virus genome. *Journal of Virological Methods* 116(1): 103-106.
- Konopka, K., Lee, N.S., Rossi, J. et Duzgunes, N. (2000) Rev-binding aptamer and CMV promoter act as decoys to inhibit HIV replication. *Gene* 255(2): 235-244.
- Kos, A., Dijkema, R., Arnberg, A.C., van der Meide, P.H. et Schellekens, H. (1986) The hepatitis delta (delta) virus possesses a circular RNA. *Nature* 323(6088): 558-560.
- Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E. et Cech, T.R. (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* 31(1): 147-157.
- Kruger, M., Beger, C., Li, Q.X., Welch, P.J., Tritz, R., Leavitt, M., Barber, J.R. et Wong-Staal, F. (2000) Identification of eIF2B γ and eIF2 γ as cofactors of hepatitis C virus internal ribosome entry site-mediated translation using a functional genomics approach. *Proceedings of the National Academy of Sciences of the United States of America* 97(15): 8566-8571.
- Krutzfeldt, J., Rajewsky, N., Braich, R., Rajeev, K.G., Tuschl, T., Manoharan, M. et Stoffel, M. (2005) Silencing of microRNAs in vivo with 'antagomirs'. *Nature* 438(7068): 685-689.
- Kuo, M.Y., Sharmeen, L., Dinter-Gottlieb, G. et Taylor, J. (1988) Characterization of self-cleaving RNA sequences on the genome and antigenome of human hepatitis delta virus. *Journal of Virology* 62(12): 4439-4444.
- Kurzynska-Kokorniak, A., Jackowiak, P., Figlerowicz, M. et Figlerowicz, M. (2009) Human- and virus-encoded microRNAs as potential targets of antiviral therapy. *Mini Reviews in Medicinal Chemistry* 9(8): 927-937.
- Laine, S., Scarborough, R.J., Levesque, D., Didierlaurent, L., Soye, K.J., Mougel, M., Perreault, J.P. et 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 Biology* 8(2).
- Langley, A.R., Chambers, H., Christov, C.P. et Krude, T. (2010) Ribonucleoprotein particles containing non-coding Y RNAs, Ro60, La and nucleolin are not required for Y RNA function in DNA replication. *PloS One* 5(10): e13673.
- Lares, M.R., Rossi, J.J. et Ouellet, D.L. (2010) RNAi and small interfering RNAs in human disease therapeutic applications. *Trends in Biotechnology* 28(11): 570-579.
- Leavitt, M.C., Yu, G., Zhou, C. et Barber, J.R. (2000) Inhibition of interleukin-1beta (IL-1beta) production in human cells by ribozymes against IL-1beta and IL-1beta converting enzyme (ICE). *Antisense & Nucleic Acid Drug Development* 10(6): 409-414.
- Lebbink, R.J., Lowe, M., Chan, T., Khine, H., Wang, X. et McManus, M.T. (2011) Polymerase II promoter strength determines efficacy of microRNA adapted shRNAs. *PloS One* 6(10): e26213.
- Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S. et al. (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425(6956): 415-419.
- Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H. et Kim, V.N. (2004) MicroRNA genes are transcribed by RNA polymerase II. *The EMBO Journal* 23(20): 4051-4060.
- Levesque, D., Briere, F.P. et Perreault, J.P. (2007) A modern mode of activation for nucleic acid enzymes. *PloS One* 2(7): e673.
- Levesque, D., Choufani, S. et Perreault, J.P. (2002) Delta ribozyme benefits from a good stability in vitro that becomes outstanding in vivo. *RNA (New York, N.Y.)* 8(4): 464-477.
- Levesque, M.V., Levesque, D., Briere, F.P. et Perreault, J.P. (2010) Investigating a new generation of ribozymes in order to target HCV. *PloS One* 5(3): e9627.
- Lewin, A.S., Drenser, K.A., Hauswirth, W.W., Nishikawa, S., Yasumura, D., Flannery, J.G. et LaVail, M.M. (1998) Ribozyme rescue of photoreceptor cells in a transgenic rat model of autosomal dominant retinitis pigmentosa. *Nature Medicine* 4(8): 967-971.
- Lewis, B.P., Burge, C.B. et Bartel, D.P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 120(1): 15-20.

- Li, L.H., Guo, Z.J., Yan, L.L., Yang, J.C., Xie, Y.F., Sheng, W.H., Huang, Z.H. et Wang, X.H. (2007) Antitumor and antiangiogenic activities of anti-vascular endothelial growth factor hairpin ribozyme in human hepatocellular carcinoma cell cultures and xenografts. *World Journal of Gastroenterology : WJG* 13(47): 6425-6432.
- Li, M.J., Kim, J., Li, S., Zaia, J., Yee, J.K., Anderson, J., Akkina, R. et Rossi, J.J. (2005) Long-term inhibition of HIV-1 infection in primary hematopoietic cells by lentiviral vector delivery of a triple combination of anti-HIV shRNA, anti-CCR5 ribozyme, and a nucleolar-localizing TAR decoy. *Molecular Therapy : The Journal of the American Society of Gene Therapy* 12(5): 900-909.
- Li, Q.X., Robbins, J.M., Welch, P.J., Wong-Staal, F. et Barber, J.R. (2000) A novel functional genomics approach identifies mTERT as a suppressor of fibroblast transformation. *Nucleic Acids Research* 28(13): 2605-2612.
- Lian, Y., De Young, M.B., Siwkowski, A., Hampel, A. et Rappaport, J. (1999) The sCYMV1 hairpin ribozyme: targeting rules and cleavage of heterologous RNA. *Gene Therapy* 6(6): 1114-1119.
- Lieber, A. et Strauss, M. (1995) Selection of efficient cleavage sites in target RNAs by using a ribozyme expression library. *Molecular and Cellular Biology* 15(1): 540-551.
- Lim, L.P., Lau, N.C., Garrett-Engele, P., Grimson, A., Schelter, J.M., Castle, J., Bartel, D.P., Linsley, P.S. et Johnson, J.M. (2005) Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* 433(7027): 769-773.
- Lin, X., Ruan, X., Anderson, M.G., McDowell, J.A., Kroeger, P.E., Fesik, S.W. et Shen, Y. (2005) siRNA-mediated off-target gene silencing triggered by a 7 nt complementation. *Nucleic Acids Research* 33(14): 4527-4535.
- Lisowski, L., Elazar, M., Chu, K., Glenn, J.S. et Kay, M.A. (2013) The anti-genomic (negative) strand of Hepatitis C Virus is not targetable by shRNA. *Nucleic Acids Research* 41(6): 3688-3698.
- Liu, F. et Altman, S. (1995) Inhibition of viral gene expression by the catalytic RNA subunit of RNase P from *Escherichia coli*. *Genes & Development* 9(4): 471-480.
- Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L. et Hannon, G.J. (2004) Argonaute2 is the catalytic engine of mammalian RNAi. *Science (New York, N.Y.)* 305(5689): 1437-1441.

- Liu, Y.P., von Eije, K.J., Schopman, N.C., Westerink, J.T., ter Brake, O., Haasnoot, J. et Berkhout, B. (2009) Combinatorial RNAi against HIV-1 using extended short hairpin RNAs. *Molecular Therapy : The Journal of the American Society of Gene Therapy* 17(10): 1712-1723.
- Low, J.T., Knoepfel, S.A., Watts, J.M., ter Brake, O., Berkhout, B. et Weeks, K.M. (2012) SHAPE-directed discovery of potent shRNA inhibitors of HIV-1. *Molecular Therapy : The Journal of the American Society of Gene Therapy* 20(4): 820-828.
- Lu, Y., Gu, J., Jin, D., Gao, Y. et Yuan, M. (2011) Inhibition of telomerase activity by HDV ribozyme in cancers. *Journal of Experimental & Clinical Cancer Research : CR* 30: 1-9966-30-1.
- Lucier, J.F., Bergeron, L.J., Briere, F.P., Ouellette, R., Elela, S.A. et Perreault, J.P. (2006) RiboSubstrates: a web application addressing the cleavage specificities of ribozymes in designated genomes. *BMC Bioinformatics* 7: 480.
- Lund, E., Guttinger, S., Calado, A., Dahlberg, J.E. et Kutay, U. (2004) Nuclear export of microRNA precursors. *Science (New York, N.Y.)* 303(5654): 95-98.
- Lundblad, E.W. et Altman, S. (2010) Inhibition of gene expression by RNase P. *New Biotechnology* 27(3): 212-221.
- Ma, M., Benimetskaya, L., Lebedeva, I., Dignam, J., Takle, G. et Stein, C.A. (2000) Intracellular mRNA cleavage induced through activation of RNase P by nuclease-resistant external guide sequences. *Nature Biotechnology* 18(1): 58-61.
- Makinen, P.I., Koponen, J.K., Karkkainen, A.M., Malm, T.M., Pulkkinen, K.H., Koistinaho, J., Turunen, M.P. et Yla-Herttuala, S. (2006a) Stable RNA interference: comparison of U6 and H1 promoters in endothelial cells and in mouse brain. *The Journal of Gene Medicine* 8(4): 433-441.
- Makinen, P.I., Koponen, J.K., Karkkainen, A.M., Malm, T.M., Pulkkinen, K.H., Koistinaho, J., Turunen, M.P. et Yla-Herttuala, S. (2006b) Stable RNA interference: comparison of U6 and H1 promoters in endothelial cells and in mouse brain. *The Journal of Gene Medicine* 8(4): 433-441.
- Marvin, M.C. et Engelke, D.R. (2009) RNase P: increased versatility through protein complexity? *RNA Biology* 6(1): 40-42.
- McGivern, D.R. et Lemon, S.M. (2011) Virus-specific mechanisms of carcinogenesis in hepatitis C virus associated liver cancer. *Oncogene* 30(17): 1969-1983.

- McIntyre, G.J., Groneman, J.L., Yu, Y.H., Jaramillo, A., Shen, S. et Applegate, T.L. (2009) 96 shRNAs designed for maximal coverage of HIV-1 variants. *Retrovirology* 6: 55-4690-6-55.
- McKinney, J., Guerrier-Takada, C., Wesolowski, D. et Altman, S. (2001) Inhibition of *Escherichia coli* viability by external guide sequences complementary to two essential genes. *Proceedings of the National Academy of Sciences of the United States of America* 98(12): 6605-6610.
- Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G. et Tuschl, T. (2004) Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Molecular Cell* 15(2): 185-197.
- Meister, G. et Tuschl, T. (2004) Mechanisms of gene silencing by double-stranded RNA. *Nature* 431(7006): 343-349.
- Mercatanti, A., Lande, C. et Citti, L. (2012) A computational approach to predict suitable target sites for trans-acting minimal hammerhead ribozymes. *Methods in Molecular Biology (Clifton, N.J.)* 848: 337-356.
- Mercatanti, A., Rainaldi, G., Mariani, L., Marangoni, R. et Citti, L. (2002) A method for prediction of accessible sites on an mRNA sequence for target selection of hammerhead ribozymes. *Journal of Computational Biology : A Journal of Computational Molecular Cell Biology* 9(4): 641-653.
- Mercure, S., Lafontaine, D., Ananvoranich, S. et Perreault, J.P. (1998) Kinetic analysis of delta ribozyme cleavage. *Biochemistry* 37(48): 16975-16982.
- Millington-Ward, S., O'Neill, B., Kiang, A.S., Humphries, P., Kenna, P.F. et Farrar, G.J. (1999) A mutation-independent therapeutic strategem for osteogenesis imperfecta. *Antisense & Nucleic Acid Drug Development* 9(6): 537-542.
- Mitsuyasu, R.T., Merigan, T.C., Carr, A., Zack, J.A., Winters, M.A., Workman, C., Bloch, M., Lalezari, J., Becker, S., Thornton, L. et al. (2009) Phase 2 gene therapy trial of an anti-HIV ribozyme in autologous CD34+ cells. *Nature Medicine* 15(3): 285-292.
- Morrow, P.K., Murthy, R.K., Ensor, J.D., Gordon, G.S., Margolin, K.A., Elias, A.D., Urba, W.J., Weng, D.E., Rugo, H.S. et Hortobagyi, G.N. (2012) An open-label, phase 2 trial of RPI.4610 (Angiozyme) in the treatment of metastatic breast cancer. *Cancer* 118(17): 4098-4104.
- Motard, J., Rouxel, R., Paun, A., von Messling, V., Bisailon, M. et Perreault, J.P. (2011) A novel ribozyme-based prophylaxis inhibits influenza A virus replication and protects from severe disease. *PloS One* 6(11): e27327.

- Muller-Kuller, T., Capalbo, G., Klebba, C., Engels, J.W. et Klein, S.A. (2009) Identification and characterization of a highly efficient anti-HIV pol hammerhead ribozyme. *Oligonucleotides* 19(3): 265-272.
- Murakami, Y., Aly, H.H., Tajima, A., Inoue, I. et Shimotohno, K. (2009) Regulation of the hepatitis C virus genome replication by miR-199a. *Journal of Hepatology* 50(3): 453-460.
- Myant, K. et Sansom, O.J. (2011) Wnt/Myc interactions in intestinal cancer: partners in crime. *Experimental Cell Research* 317(19): 2725-2731.
- Naito, Y., Nohtomi, K., Onogi, T., Uenishi, R., Ui-Tei, K., Saigo, K. et Takebe, Y. (2007) Optimal design and validation of antiviral siRNA for targeting HIV-1. *Retrovirology* 4: 80.
- Nehdi, A., Perreault, J., Beaudoin, J.D. et Perreault, J.P. (2007) A novel structural rearrangement of hepatitis delta virus antigenomic ribozyme. *Nucleic Acids Research* 35(20): 6820-6831.
- Nehdi, A. et Perreault, J.P. (2006) Unbiased in vitro selection reveals the unique character of the self-cleaving antigenomic HDV RNA sequence. *Nucleic Acids Research* 34(2): 584-592.
- Nishikawa, F., Roy, M., Fauzi, H. et Nishikawa, S. (1999) Detailed analysis of stem I and its 5' and 3' neighbor regions in the trans-acting HDV ribozyme. *Nucleic Acids Research* 27(2): 403-410.
- Nobel Media AB. (1989) Press Release: The 1989 Nobel Prize in Chemistry. http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1989/press.html.
- Nobel Media AB. (2006) The 2006 Nobel Prize in Physiology or Medicine - Advanced Information. http://www.nobelprize.org/nobel_prizes/medicine/laureates/2006/advanced.html.
- Pan, Q., Tilanus, H.W., Janssen, H.L. et van der Laan, L.J. (2009) Prospects of RNAi and microRNA-based therapies for hepatitis C. *Expert Opinion on Biological Therapy* 9(6): 713-724.
- Pan, W.H., Devlin, H.F., Kelley, C., Isom, H.C. et Clawson, G.A. (2001) A selection system for identifying accessible sites in target RNAs. *RNA (New York, N.Y.)* 7(4): 610-621.
- Patil, C.S., Liu, M., Zhao, W., Coatney, D.D., Li, F., VanTubergen, E.A., D'Silva, N.J. et Kirkwood, K.L. (2008) Targeting mRNA stability arrests inflammatory bone loss. *Molecular Therapy : The Journal of the American Society of Gene Therapy* 16(10): 1657-1664.

- Patzel, V., Rutz, S., Dietrich, I., Koberle, C., Scheffold, A. et Kaufmann, S.H. (2005) Design of siRNAs producing unstructured guide-RNAs results in improved RNA interference efficiency. *Nature Biotechnology* 23(11): 1440-1444.
- Pawlotsky, J.M., Chevaliez, S. et McHutchison, J.G. (2007) The hepatitis C virus life cycle as a target for new antiviral therapies. *Gastroenterology* 132(5): 1979-1998.
- Peracchi, A. (2004) Prospects for antiviral ribozymes and deoxyribozymes. *Reviews in Medical Virology* 14(1): 47-64.
- Puerta-Fernandez, E., Romero-Lopez, C., Barroso-delJesus, A. et Berzal-Herranz, A. (2003) Ribozymes: recent advances in the development of RNA tools. *FEMS Microbiology Reviews* 27(1): 75-97.
- Rand, T.A., Petersen, S., Du, F. et Wang, X. (2005) Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. *Cell* 123(4): 621-629.
- Randau, L., Schroder, I. et Soll, D. (2008) Life without RNase P. *Nature* 453(7191): 120-123.
- Reymond, C., Beaudoin, J.D. et Perreault, J.P. (2009a) Modulating RNA structure and catalysis: lessons from small cleaving ribozymes. *Cellular and Molecular Life Sciences : CMLS* 66(24): 3937-3950.
- Reymond, C., Bisailon, M. et Perreault, J.P. (2009b) Monitoring of an RNA multistep folding pathway by isothermal titration calorimetry. *Biophysical Journal* 96(1): 132-140.
- Reymond, C., Levesque, D., Bisailon, M. et Perreault, J.P. (2010) Developing three-dimensional models of putative-folding intermediates of the HDV ribozyme. *Structure (London, England : 1993)* 18(12): 1608-1616.
- Reynolds, A., Leake, D., Boese, Q., Scaringe, S., Marshall, W.S. et Khvorova, A. (2004) Rational siRNA design for RNA interference. *Nature Biotechnology* 22(3): 326-330.
- Robichaud, G.A., Perreault, J.P. et Ouellette, R.J. (2008) Development of an isoform-specific gene suppression system: the study of the human Pax-5B transcriptional element. *Nucleic Acids Research* 36(14): 4609-4620.
- Roehr, B. (1998) Fomivirsen approved for CMV retinitis. *Journal of the International Association of Physicians in AIDS Care* 4(10): 14-16.
- Romero-Brey, I., Merz, A., Chiramel, A., Lee, J.Y., Chlanda, P., Haselman, U., Santarella-Mellwig, R., Habermann, A., Hoppe, S., Kallis, S. et al. (2012) Three-dimensional architecture and biogenesis of membrane structures

- associated with hepatitis C virus replication. *PLoS Pathogens* 8(12): e1003056.
- Romero-Lopez, C., Barroso-delJesus, A., Puerta-Fernandez, E. et Berzal-Herranz, A. (2004) Design and optimization of sequence-specific hairpin ribozymes. *Methods in Molecular Biology* (Clifton, N.J.) 252: 327-338.
- Rose, S.D., Kim, D.H., Amarzguioui, M., Heidel, J.D., Collingwood, M.A., Davis, M.E., Rossi, J.J. et Behlke, M.A. (2005) Functional polarity is introduced by Dicer processing of short substrate RNAs. *Nucleic Acids Research* 33(13): 4140-4156.
- Rossi, J.J. (2000) Ribozyme therapy for HIV infection. *Advanced Drug Delivery Reviews* 44(1): 71-78.
- Rouleau, S.G., Jodoin, R., Bisailon, M. et Perreault, J.P. (2012) Programming a highly structured ribozyme into complex allostery using RNA oligonucleotides. *ACS Chemical Biology* 7(11): 1802-1806.
- Roy, G., Ananvoranich, S. et Perreault, J.P. (1999) Delta ribozyme has the ability to cleave in trans an mRNA. *Nucleic Acids Research* 27(4): 942-948.
- Sagan, S.M., Naseri, N., Luebbert, C. et Pezacki, J.P. (2010) The efficacy of siRNAs against hepatitis C virus is strongly influenced by structure and target site accessibility. *Chemistry & Biology* 17(5): 515-527.
- Sakamoto, N., Wu, C.H. et Wu, G.Y. (1996) Intracellular cleavage of hepatitis C virus RNA and inhibition of viral protein translation by hammerhead ribozymes. *The Journal of Clinical Investigation* 98(12): 2720-2728.
- Salehi-Ashtiani, K., Luptak, A., Litovchick, A. et Szostak, J.W. (2006) A genomewide search for ribozymes reveals an HDV-like sequence in the human CPEB3 gene. *Science (New York, N.Y.)* 313(5794): 1788-1792.
- Sall, A., Liu, Z., Zhang, H.M., Yuan, J., Lim, T., Su, Y. et Yang, D. (2008) MicroRNAs-based therapeutic strategy for virally induced diseases. *Current Drug Discovery Technologies* 5(1): 49-58.
- Scaggiante, B., Dapas, B., Farra, R., Grassi, M., Pozzato, G., Giansante, C., Fiotti, N. et Grassi, G. (2011) Improving siRNA bio-distribution and minimizing side effects. *Current Drug Metabolism* 12(1): 11-23.
- Scarborough, R.J., Levesque, M.V., Boudrias-Dalle, E., Chute, I.C., Daniels, S., Ouellette, R.J., Perreault, J. et Gagnon, A. (2008) A conserved target site in HIV-1 Gag RNA is accessible to both SOFA-HDV Ribozyme and short hairpin RNA activities. .

- Scherer, L., Rossi, J.J. et Weinberg, M.S. (2007) Progress and prospects: RNA-based therapies for treatment of HIV infection. *Gene Therapy* 14(14): 1057-1064.
- Schubert, S. et Kurreck, J. (2004) Ribozyme- and deoxyribozyme-strategies for medical applications. *Current Drug Targets* 5(8): 667-681.
- Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N. et Zamore, P.D. (2003) Asymmetry in the assembly of the RNAi enzyme complex. *Cell* 115(2): 199-208.
- Seo, M.Y., Abrignani, S., Houghton, M. et Han, J.H. (2003) Small interfering RNA-mediated inhibition of hepatitis C virus replication in the human hepatoma cell line Huh-7. *Journal of Virology* 77(1): 810-812.
- Shareghi, P., Wang, Y., Malmberg, R. et Cai, L. (2012) Simultaneous prediction of RNA secondary structure and helix coaxial stacking. *BMC Genomics* 13 Suppl 3: S7-2164-13-S3-S7.
- Sharmeen, L., Kuo, M.Y., Dinter-Gottlieb, G. et Taylor, J. (1988) Antigenomic RNA of human hepatitis delta virus can undergo self-cleavage. *Journal of Virology* 62(8): 2674-2679.
- Sheng, J., Al-Anouti, F. et Ananvoranich, S. (2004) Engineered delta ribozymes can simultaneously knock down the expression of the genes encoding uracil phosphoribosyltransferase and hypoxanthine-xanthine-guanine phosphoribosyltransferase in *Toxoplasma gondii*. *International Journal for Parasitology* 34(3): 253-263.
- Shiota, M., Sano, M., Miyagishi, M. et Taira, K. (2004) Ribozymes: applications to functional analysis and gene discovery. *Journal of Biochemistry* 136(2): 133-147.
- Shippy, R., Lockner, R., Farnsworth, M. et Hampel, A. (1999) The hairpin ribozyme. Discovery, mechanism, and development for gene therapy. *Molecular Biotechnology* 12(1): 117-129.
- Silverman, S.K. (2005) In vitro selection, characterization, and application of deoxyribozymes that cleave RNA. *Nucleic Acids Research* 33(19): 6151-6163.
- Sim, S. et Wolin, S.L. (2011) Emerging roles for the Ro 60-kDa autoantigen in noncoding RNA metabolism. *Wiley Interdisciplinary Reviews.RNA* 2(5): 686-699.

- Singh, S., Narang, A.S. et Mahato, R.I. (2011) Subcellular fate and off-target effects of siRNA, shRNA, and miRNA. *Pharmaceutical Research* 28(12): 2996-3015.
- Siolas, D., Lerner, C., Burchard, J., Ge, W., Linsley, P.S., Paddison, P.J., Hannon, G.J. et Cleary, M.A. (2005) Synthetic shRNAs as potent RNAi triggers. *Nature Biotechnology* 23(2): 227-231.
- Sliva, K. et Schnierle, B.S. (2010) Selective gene silencing by viral delivery of short hairpin RNA. *Virology Journal* 7: 248.
- Sloan, L.A., Fillmore, M.C. et Churcher, I. (2009) Small-molecule modulation of cellular chaperones to treat protein misfolding disorders. *Current Opinion in Drug Discovery & Development* 12(5): 666-681.
- Smith, R.M., Walton, C.M., Wu, C.H. et Wu, G.Y. (2002) Secondary structure and hybridization accessibility of hepatitis C virus 3'-terminal sequences. *Journal of Virology* 76(19): 9563-9574.
- Song, Y.H., Lin, J.S., Liu, N.Z., Kong, X.J., Xie, N., Wang, N.X., Jin, Y.X. et Liang, K.H. (2002) Anti-HBV hairpin ribozyme-mediated cleavage of target RNA in vitro. *World Journal of Gastroenterology : WJG* 8(1): 91-94.
- Stark, B.C., Kole, R., Bowman, E.J. et Altman, S. (1978) Ribonuclease P: an enzyme with an essential RNA component. *Proceedings of the National Academy of Sciences of the United States of America* 75(8): 3717-3721.
- Stein, C.A. (1997) Controversies in the cellular pharmacology of oligodeoxynucleotides. *Ciba Foundation Symposium* 209: 79-89; discussion 89-93.
- Stephenson, M.L. et Zamecnik, P.C. (1978) Inhibition of Rous sarcoma viral RNA translation by a specific oligodeoxyribonucleotide. *Proceedings of the National Academy of Sciences of the United States of America* 75(1): 285-288.
- Su, C.H., Wu, Y.J., Wang, H.H. et Yeh, H.I. (2012) Nonviral gene therapy targeting cardiovascular system. *American Journal of Physiology. Heart and Circulatory Physiology* 303(6): H629-38.
- Sun, X., Bromley-Brits, K. et Song, W. (2012) Regulation of beta-site APP-cleaving enzyme 1 gene expression and its role in Alzheimer's disease. *Journal of Neurochemistry* 120 Suppl 1: 62-70.
- Swiatkowska, A., Dutkiewicz, M. et Ciesiolka, J. (2007) Structural features of target RNA molecules greatly modulate the cleavage efficiency of trans-acting delta ribozymes. *Biochemistry* 46(18): 5523-5533.

- Tang, G. (2005) siRNA and miRNA: an insight into RISCs. *Trends in Biochemical Sciences* 30(2): 106-114.
- Tedeschi, L., Lande, C., Cecchettini, A. et Citti, L. (2009) Hammerhead ribozymes in therapeutic target discovery and validation. *Drug Discovery Today* 14(15-16): 776-783.
- Tsai, L.M. et Yu, D. (2010) MicroRNAs in common diseases and potential therapeutic applications. *Clinical and Experimental Pharmacology & Physiology* 37(1): 102-107.
- Tschuch, C., Schulz, A., Pscherer, A., Werft, W., Benner, A., Hotz-Wagenblatt, A., Barrionuevo, L.S., Lichter, P. et Mertens, D. (2008) Off-target effects of siRNA specific for GFP. *BMC Molecular Biology* 9: 60-2199-9-60.
- Tsuda, N., Ishiyama, S., Li, Y., Ioannides, C.G., Abbruzzese, J.L. et Chang, D.Z. (2006) Synthetic microRNA designed to target glioma-associated antigen 1 transcription factor inhibits division and induces late apoptosis in pancreatic tumor cells. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research* 12(21): 6557-6564.
- Tyagi, A.C., Sen, U. et Mishra, P.K. (2011) Synergy of microRNA and stem cell: a novel therapeutic approach for diabetes mellitus and cardiovascular diseases. *Current Diabetes Reviews* 7(6): 367-376.
- Uhlenbeck, O.C. (1987) A small catalytic oligoribonucleotide. *Nature* 328(6131): 596-600.
- Vaerman, J.L., Moureau, P., Deldime, F., Lewalle, P., Lamineur, C., Morschhauser, F. et Martiat, P. (1997) Antisense oligodeoxyribonucleotides suppress hematologic cell growth through stepwise release of deoxyribonucleotides. *Blood* 90(1): 331-339.
- Vitiello, D., Pecchia, D.B. et Burke, J.M. (2000) Intracellular ribozyme-catalyzed trans-cleavage of RNA monitored by fluorescence resonance energy transfer. *RNA (New York, N.Y.)* 6(4): 628-637.
- Wang, C.X., Lu, Y.Q., Qi, P., Chen, L.H. et Han, J.X. (2010) Efficient inhibition of hepatitis B virus replication by hepatitis delta virus ribozymes delivered by targeting retrovirus. *Virology Journal* 7: 61-422X-7-61.
- Wang, K.S., Choo, Q.L., Weiner, A.J., Ou, J.H., Najarian, R.C., Thayer, R.M., Mullenbach, G.T., Denniston, K.J., Gerin, J.L. et Houghton, M. (1986) Structure, sequence and expression of the hepatitis delta (delta) viral genome. *Nature* 323(6088): 508-514.

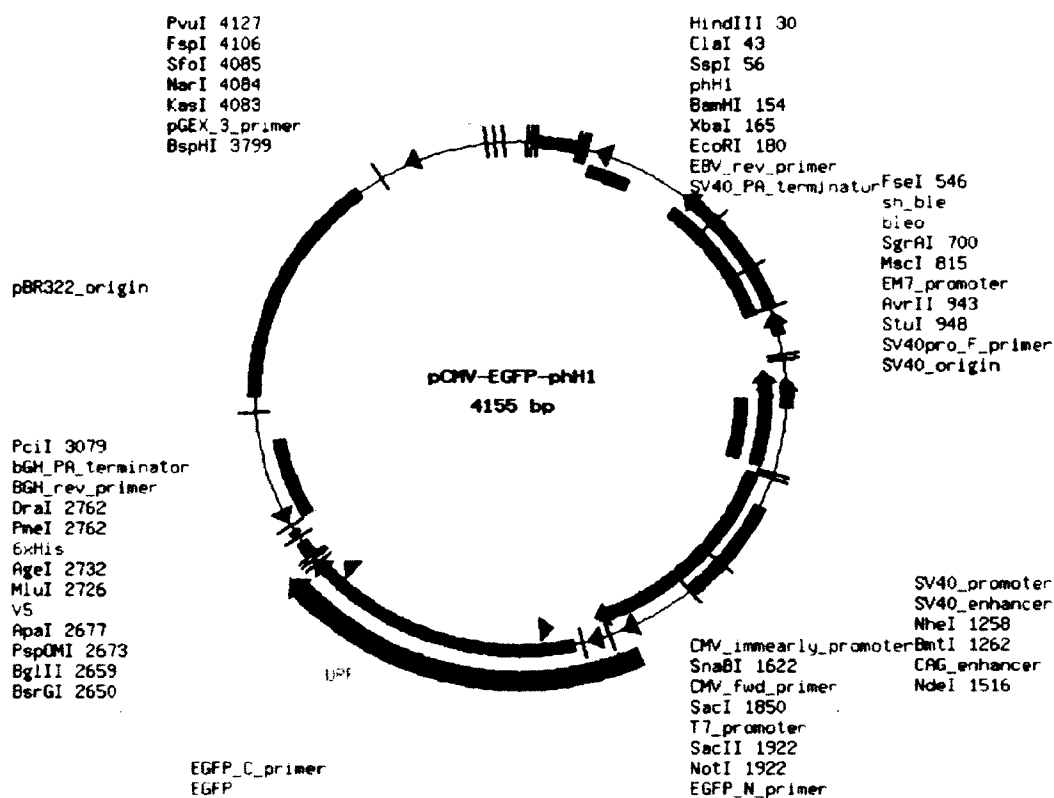
- Watts, J.M., Dang, K.K., Gorelick, R.J., Leonard, C.W., Bess, J.W., Jr, Swanstrom, R., Burch, C.L. et Weeks, K.M. (2009) Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature* 460(7256): 711-716.
- Webb, C.H., Riccitelli, N.J., Ruminiski, D.J. et Luptak, A. (2009) Widespread occurrence of self-cleaving ribozymes. *Science (New York, N.Y.)* 326(5955): 953.
- Welch, P.J., Marcusson, E.G., Li, Q.X., Beger, C., Kruger, M., Zhou, C., Leavitt, M., Wong-Staal, F. et Barber, J.R. (2000) Identification and validation of a gene involved in anchorage-independent cell growth control using a library of randomized hairpin ribozymes. *Genomics* 66(3): 274-283.
- Wickstrom, E. (1986) Oligodeoxynucleotide stability in subcellular extracts and culture media. *Journal of Biochemical and Biophysical Methods* 13(2): 97-102.
- Wolchok, J.D., Yang, A.S. et Weber, J.S. (2010) Immune regulatory antibodies: are they the next advance? *Cancer Journal (Sudbury, Mass.)* 16(4): 311-317.
- Wu, H., Lima, W.F., Zhang, H., Fan, A., Sun, H. et Crooke, S.T. (2004) Determination of the role of the human RNase H1 in the pharmacology of DNA-like antisense drugs. *The Journal of Biological Chemistry* 279(17): 17181-17189.
- Xu, Y., Jiang, Y.F. et Wu, B. (2012) New agonist- and antagonist-based treatment approaches for advanced prostate cancer. *The Journal of International Medical Research* 40(4): 1217-1226.
- Yi, J. et Liu, C. (2011) Efficient silencing of gene expression by an ASON-bulge-DNAzyme complex. *PloS One* 6(4): e18629.
- Yi, R., Doehle, B.P., Qin, Y., Macara, I.G. et Cullen, B.R. (2005) Overexpression of exportin 5 enhances RNA interference mediated by short hairpin RNAs and microRNAs. *RNA (New York, N.Y.)* 11(2): 220-226.
- Yokota, T., Sakamoto, N., Enomoto, N., Tanabe, Y., Miyagishi, M., Maekawa, S., Yi, L., Kurosaki, M., Taira, K., Watanabe, M. et al. (2003) Inhibition of intracellular hepatitis C virus replication by synthetic and vector-derived small interfering RNAs. *EMBO Reports* 4(6): 602-608.
- Yuki, N., Matsumoto, S., Tadokoro, K., Mochizuki, K., Kato, M. et Yamaguchi, T. (2006) Significance of liver negative-strand HCV RNA quantitation in chronic hepatitis C. *Journal of Hepatology* 44(2): 302-309.
- Zamecnik, P.C. et Stephenson, M.L. (1978) Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide.

Proceedings of the National Academy of Sciences of the United States of America 75(1): 280-284.

- Zeng, Y. et Cullen, B.R. (2004) Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. *Nucleic Acids Research* 32(16): 4776-4785.
- Zeng, Y., Wagner, E.J. et Cullen, B.R. (2002) Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Molecular Cell* 9(6): 1327-1333.
- Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E. et Filipowicz, W. (2004) Single processing center models for human Dicer and bacterial RNase III. *Cell* 118(1): 57-68.
- Zhou, J., Bobbin, M.L., Burnett, J.C. et Rossi, J.J. (2012) Current Progress of RNA Aptamer-Based Therapeutics. *Front. Genet.* 3.
- Zhou, T., Kim, J., Kilani, A.F., Kim, K., Dunn, W., Jo, S., Nepomuceno, E. et Liu, F. (2002) In vitro selection of external guide sequences for directing RNase P-mediated inhibition of viral gene expression. *The Journal of Biological Chemistry* 277(33): 30112-30120.

ANNEXES

1. Carte du vecteur d'expression pCMV-EGFP-phH1



Annexes, Figure 1. Carte du vecteur d'expression pCMV-EGFP-phH1.

Le vecteur pCMV-EGFP-phH1 a été construit à partir du vecteur pBudCE4.1 (Invitrogen). Le promoteur hh1 (phH1) a été inséré à la place du promoteur CMV (pCMV). Le pCMV a ensuite remplacé le promoteur EF-1 α pour l'expression de la protéine EGFP. Les ribozymes exprimés par le phH1 sont clonés avec les sites de restriction *BamHI* et *EcoRI*.

2. Lettres d'approbation des co-auteurs



PERMISSION TO INCORPORATE A CO-AUTHORED ARTICLE INTO A THESIS OR DISSERTATION

I (or We), the undersigned, co-author(s) of the article titled:
Investigating a New Generation of Ribozymes in Order to
Target HCV

acknowledge that said article will be included as a constituent part of the
master's thesis doctoral dissertation

of the following student (name): Michel Lévesque
registered in the program: Biochimie
offered by the following Faculty: Faculté de Médecine et des Sciences
of the Université de Sherbrooke. de la santé

In witness whereof, I (we) have signed this agreement in a sufficient number of copies*

Signature: Date: 20-07-2013
Name: Dominique Lévesque
Contact information: Département de Biochimie, Université de
Sherbrooke, Sherbrooke, Québec, Canada

Signature: Date: 21-07-2013
Name: Francis P. Brière
Contact information: Département de Biochimie, Université de
Sherbrooke, Sherbrooke, Québec, Canada

Signature: Date: 19-07-2013
Name: Jean-Pierre Perreault
Contact information: Département de Biochimie, Université de
Sherbrooke, Sherbrooke, Québec, Canada

Signature: _____ Date: _____
Name: _____
Contact information: _____

*One copy for the student, one copy for each of the signatories, and one copy for the University Library (such copy to be included with the thesis or dissertation at the time of the final submission).



**PERMISSION TO INCORPORATE A CO-AUTHORED ARTICLE INTO A
THESIS OR DISSERTATION**

I (or We), the undersigned, co-author(s) of the article titled:
A Conserved Target Site in HIV-1 Gag RNA is Accessible to
Both HDV Ribozyme and Short Hairpin RNA Activities

acknowledge that said article will be included as a constituent part of the
master's thesis doctoral dissertation

of the following student (name): Michel Lévesque
registered in the program: Biochimie
offered by the following Faculty: Faculté de Médecine et des Sciences
of the Université de Sherbrooke. de la santé

In witness whereof, I (we) have signed this agreement in a sufficient number of copies*

Signature: [Signature] Date: July 19, 2013
Name: Robert J. Scarborough
Contact information: Lady Davis Institute for Medical Research,
Department of Microbiology & Immunology, McGill University,
Montréal, Qc, Canada

Signature: [Signature] Date: 2013-02-10
Name: Etienne Boudrias-Dalle
Contact information: Lady Davis Institute for Medical Research,
Department of Microbiology & Immunology, McGill University,
Montréal, Qc, Canada

Signature: [Signature] Date: 2013-07-19
Name: Sylvaine M. Daniels
Contact information: Lady Davis Institute for Medical Research,
Department of Microbiology & Immunology, McGill University,
Montréal, Qc, Canada

Signature: [Signature] Date: July 19, 2013
Name: Anne Gagnon
Contact information: Lady Davis Institute for Medical Research,
Department of Microbiology & Immunology, McGill University,
Montréal, Qc, Canada

*One copy for the student, one copy for each of the signatories, and one copy for the University Library (such copy to be included with the thesis or dissertation at the time of the final submission).



**PERMISSION TO INCORPORATE A CO-AUTHORED ARTICLE INTO A
THESIS OR DISSERTATION**

I (or We), the undersigned, co-author(s) of the article titled:
**A Conserved Target Site in HIV-1 Gag RNA is Accessible to
Both HDV Ribozyme and Short Hairpin RNA Activities**

acknowledges that said article will be included as a constituent part of the
master's thesis doctoral dissertation

of the following student (name): Michel Lévesque
registered in the program: Biochimie
offered by the following Faculty: Faculté de Médecine et des Sciences
of the Université de Sherbrooke. de la santé

In witness whereof, I (we) have signed this agreement in a sufficient number of copies*

Signature: [Signature] Date: 23/07/13
Name: Ian C. Chute
Contact information: Atlantic Cancer Research Institute, Moncton,
New Brunswick, Canada

Signature: [Signature] Date: 22/07/13
Name: Rodney J. Ouellette
Contact information: Atlantic Cancer Research Institute, Moncton,
New Brunswick, Canada

Signature: [Signature] Date: 19-07-2013
Name: Jean-Pierre Perveault
Contact information: Département de Biochimie, Université de
Sherbrooke, Sherbrooke, Québec, Canada

Signature: _____ Date: _____
Name: _____
Contact information: _____

*One copy for the student, one copy for each of the signatories, and one copy for the University Library (such copy to be included with the thesis or dissertation at the time of the final submission).



PERMISSION TO INCORPORATE A CO-AUTHORED ARTICLE INTO A THESIS OR DISSERTATION

I (or We), the undersigned, co-author(s) of the article titled:
Selection of The Most Potent SOFA-HDV Ribozymes For Use
in Gene Targeting

acknowledge that said article will be included as a constituent part of the

master's thesis doctoral dissertation

of the following student (name): Michel Lévesque
registered in the program: Biochimie
offered by the following Faculty: Faculté de Médecine et des Sciences
of the Université de Sherbrooke. de la santé

In witness whereof, I (we) have signed this agreement in a sufficient number of copies*

Signature: [Signature] Date: 19-07-2013

Name: Samuel G Rouleau

Contact information: Département de Biochimie, Université de
Sherbrooke, Sherbrooke, Québec, Canada

Signature: [Signature] Date: 19-07-2013

Name: Jean-Pierre Perrault

Contact information: Département de Biochimie, Université de
Sherbrooke, Sherbrooke, Québec, Canada

Signature: _____ Date: _____

Name: _____

Contact information: _____

Signature: _____ Date: _____

Name: _____

Contact information: _____

*One copy for the student, one copy for each of the signatories, and one copy for the University Library (such copy to be included with the thesis or dissertation at the time of the final submission).



**PERMISSION TO INCORPORATE A CO-AUTHORED ARTICLE INTO A
THESIS OR DISSERTATION**

I (or We), the undersigned, co-author(s) of the article titled:
Target Induced SOFA-HDV Ribozyme

acknowledge that said article will be included as a constituent part of the
master's thesis doctoral dissertation

of the following student (name): Michel Lévesque
registered in the program: Biochimie
offered by the following Faculty: Faculté de Médecine et des Sciences
of the Université de Sherbrooke. de la santé

In witness whereof, I (we) have signed this agreement in a sufficient number of copies*

Signature: _____ Date: 19-07-2013
Name: Jean-Pierre Perreault
Contact information: Département de Biochimie, Université de
Sherbrooke, Sherbrooke, Québec, Canada

Signature: _____ Date: _____
Name: _____
Contact information: _____

Signature: _____ Date: _____
Name: _____
Contact information: _____

Signature: _____ Date: _____
Name: _____
Contact information: _____

*One copy for the student, one copy for each of the signatories, and one copy for the University Library (such copy to be included with the thesis or dissertation at the time of the final submission).