

Centre Armand-Frappier Santé Biotechnologie

**INGÉNIERIE DES RIBOZYMES POUR DÉVELOPPER UN CIRCUIT
LOGIQUE APPLICABLE COMME NOUVEL OUTIL DE
DIAGNOSTIC ET THÉRAPIE**

Par

Sabrine Najeh

Thèse présentée pour l'obtention du grade de Philosophiae Doctor (Ph.D.)
en Biologie

Jury d'évaluation

**Président du jury et
Examineur interne**

Dr. Kessen Patten
INRS-Centre Armand-Frappier
Santé-Biotechnologie

Examineur externe

Dr. Pascal Chartrand
Département de Biochimie
Et Médecine Moléculaire
Université de Montréal

Examinatrice externe

Dr. Pascale Legault
Département de Biochimie
Et Médecine Moléculaire
Université de Montréal

Directeur de recherche

Dr. Jonathan Perreault
INRS-Centre Armand-Frappier
Santé-Biotechnologie

Co-Directeur de recherche

Dr. Yves St-Pierre
INRS-Centre Armand-Frappier
Santé-Biotechnologie

Dédicace

Je dédie cette œuvre à ma famille, les personnes qui ont toujours cru en moi et que sans leur soutien et leur encouragement, je n'aurais pas pu réussir mes études et d'atteindre ce stade.

Une dédicace spéciale à mon grand-père que j'aurais beaucoup aimé qu'il soit encore parmi nous, mais il a perdu son combat contre le cancer.

Liste des publications

Total : 4 publications : **1** publié, **2** soumis et **1** en préparation

Computer-Aided Design of Active Pseudoknotted Hammerhead Ribozymes. Najeh S, Zandi K, Djerroud S, Kharma N, Perreault J. Methods Mol Biol. 2021;2167:91-111. doi: 10.1007/978-1-0716-0716-9_7.(Publié)

Computational design and experimental verification of pseudoknotted ribozymes. Sabrina Najeh, Kasra Zandi, Nawwaf Kharma and Jonathan Perreault. (Accepté dans RNA)

A Web Service for Automated Design of Multiple Types of Ribozymes Targeting RNA: from Minimal Hammerhead to Aptazyme. Sabrina Najeh, Nawwaf Kharma², Thomas Vaudry-Read, Anita Haurie, Christopher Paslawski, Ankit Ranja, Daniel Adams, Steve Ferreira and Jonathan Perreault.(soumis)

Dual selection for aptazymes for biosensor development purpose, Sabrina Najeh, Jonathan Perreault, Yves St-Pierre. (En préparation)

Résumé

La découverte du pouvoir catalyseur des ARN a entraîné un changement du dogme de la biologie. Les protéines ne constituaient plus les seules entités responsables des activités enzymatiques dans les cellules. De multiples études ont contribué à l'identification des fonctions ainsi qu'à la compréhension des structures de ces enzymes à ARN. Ces ribozymes se sont avérés impliqués dans de multiples processus biologiques, allant de la réplication virale jusqu'à la synthèse protéique chez les eucaryotes et les procaryotes. Ceux impliqués dans la réplication virale sont responsables de la libération de la nouvelle copie de génome par l'auto-clivage de cette dernière. Les ribozymes ayant cette activité appartiennent à la famille des petits ARN auto-clivants. Les ARN de cette famille ont attiré beaucoup d'attention en raison des différentes façons dont ils peuvent être employés dans la régulation de l'expression génique. Ceci a été fait par l'ingénierie des séquences de ces ribozymes avec des méthodes manuelles et des sélections *in vitro*. Le tout a été possible grâce à de nombreuses recherches ayant permis la détermination des structures des ribozymes auto-clivants, aboutissant à une meilleure compréhension des motifs structuraux de l'ARN en général. Ainsi, l'ingénierie de ces ribozymes et d'autres ARN a été préconisée pour leur application dans des thérapies géniques contre des gènes viraux et des gènes impliqués dans le cancer. Les essais de ces ribozymes ont atteint des phases cliniques ; ils ont par contre échoué. De ce fait, lorsque l'on prend en considération les motifs structuraux ainsi que l'environnement de l'application, l'optimisation de l'ingénierie des ribozymes pourrait aider à surmonter de tels échecs. Dans cette optique, plusieurs outils bioinformatiques ont été développés pour la génération automatisée des ARN synthétiques ayant des structures et des fonctions prédéfinies.

Dans le cadre de mon projet, nous nous sommes intéressés aux ribozymes en tant que potentiels outils thérapeutiques et diagnostiques. Notre hypothèse est qu'en combinant des ribozymes thérapeutiques avec des biosenseurs, nous pourrions développer un circuit logique applicable comme thérapie intelligente. Cette thérapie permet en parallèle de détecter des molécules indiquant ou non la présence d'un cancer et dépendamment de ces signaux de prendre la décision de tuer la cellule ou non. Au cours de la première étape du projet, nous nous sommes penchés sur deux nouveaux programmes de conception de ribozymes. Le premier, qui est nommé « Enzymer », est destiné à la conception de ribozymes auto-clivants ayant des pseudonœuds, un motif de structure secondaire indispensable à l'activité de plusieurs ARN. Nous avons utilisé ce programme

pour la génération de deux types de ribozymes dans le but de prouver l'efficacité de cet outil ainsi que le développement d'un « *switch OFF* » pour contrôler l'expression d'une protéine fluorescente chez les bactéries. Le deuxième outil était le service web « Ribosoft 2.0 », destiné à la conception de différents types de ribozymes actifs en *trans*. Nous avons eu recours à ce logiciel dans le contexte de choix de ribozymes efficaces pour cibler des ARNm de protéines impliquées dans le cancer. En plus de l'utilisation de ces outils pour nos objectifs de recherche, nous avons participé à leur optimisation à l'aide de nos résultats expérimentaux.

En deuxième lieu, nous avons profité de la flexibilité des ribozymes pour développer des biosenseurs détectant des protéines et des miARN qui peuvent être appliqués comme outils de diagnostics. Les ribozymes liant les miARN ont également été conçus en utilisant des méthodes computationnelles. Quant aux ribozymes détectant les protéines, une sélection *in vitro* a été appliquée ; ce qui a mené à la génération d'aptazymes, qui sont la combinaison d'un aptamère et d'un ribozyme. Ces aptazymes ont été combinés avec un aptamère et des billes fluorescentes, produisant ainsi des biosenseurs détectant des protéines qui peuvent être utilisées comme biomarqueurs de certains types de cancers.

Abstract

The discovery of the catalytic power of RNAs led to a change in the dogma of biology. Proteins were no longer the only entities responsible for enzymatic activities in cells. Multiple studies have contributed to the identification of the functions and the understanding of the structures of the RNA enzymes. The ribozymes have been shown to be involved in multiple biological processes, ranging from viral replication to protein synthesis in eukaryotes and prokaryotes. Those involved in viral replication are responsible for releasing of the new genome copy by the self-cleavage of the latter. Ribozymes with this activity belong to the family of self-cleaving small RNAs. RNAs from this family have attracted a lot of attention because of the different ways they can be employed to regulate gene expression. This was done by engineering the sequences of these ribozymes with manual methods and *in vitro* selections. All this has been possible thanks to numerous researches that have allowed the determination of the structures of self-cleaving ribozymes, leading to a better understanding of the structural motifs of RNA in general. Thus, the engineering of these ribozymes and other RNAs has been advocated for their application in gene therapies against viral genes and genes involved in cancer. Trials of these ribozymes have reached clinical phases; however, they failed. Therefore, when considering the structural motifs as well as the application environment, optimizing ribozyme engineering could help overcome such failures. In this perspective, several bioinformatics tools have been developed for the automated generation of synthetic RNAs with predefined structures and functions.

As part of my project, we were interested in ribozymes as potential therapeutic and diagnostic tools. Our hypothesis is that by combining therapeutic ribozymes with biosensors, we could develop a logic circuit applicable as an intelligent therapy. This therapy makes it possible, in parallel, to detect molecules indicating or not the presence of cancer and, depending on these signals, to make the decision to kill the cell or not. During the first stage of the project, we looked at two new ribozyme design programs. The first, which is called "Enzymer", is intended for the design of self-cleaving ribozymes with pseudoknots, a secondary structure motif essential for the activity of several RNAs. We used this program for the generation of two types of ribozymes in order to prove the effectiveness of this tool as well as the development of an "OFF switch" to control the expression of a fluorescent protein in bacteria. The second tool was the "Ribosoft 2.0" web service, intended for the design of different types of trans-active ribozymes. We used this

software in the context of choosing effective ribozymes to target mRNAs of proteins involved in cancer. In addition to the use of these tools for our research objectives, we participated in their optimization using our experimental results.

During the second aim of my project, we have taken advantage of the modularity of ribozymes to develop biosensors detecting proteins and miRNAs that can be applied as diagnostic tools. The miRNA-binding ribozymes were also designed using computational methods. As for the ribozymes detecting the proteins, an *in vitro* selection strategy was used to generate aptazymes, which are the combination of an aptamer and a ribozyme. These aptazymes were combined with an aptamer and fluorescent beads, allowing to produce biosensors detecting proteins that can be used as biomarkers of certain types of cancer.

Table des matières

Liste des publications	II
Résumé.....	III
Table des matières	VII
Liste des figures	XI
Liste des tableaux	XIII
Liste des abréviations.....	XIV
1. Revue de littérature	1
1.1. Aperçu de découvertes des ARN non-codants	1
1.2. Ribozyme.....	3
1.2.1. Ribozyme en tête de marteau.....	4
1.2.2. Ribozyme Pistol.....	10
1.3. Ribozymes allostériques/ <i>Riboswitchs</i>	11
1.3.1. Riboswitchs.....	12
1.3.2. Naturels : Ribozyme GImS.....	13
1.3.3. Ribozymes allostériques synthétiques	15
1.3.4. Aptazyme.....	15
1.3.5. Ribozymes allostériques liant des oligonucléotides	21
1.4. Outils de conception de ribozymes	22
1.4.1. Sélection de nouveaux ribozymes <i>in vitro</i>	23
1.4.2. SELEX	26
1.4.3. Limites de sélection <i>in vitro</i>	29
1.4.4. Conception informatique des ARN	30
1.4.5. Outils de conception de ribozymes actifs en <i>trans</i>	32
1.4.6. Outils de conception de ribozymes actifs en <i>cis</i>	33
1.5. Applications et utilisations de ribozymes.....	38
1.5.1. But thérapeutique	38
1.5.2. Thérapie antivirale	40
1.5.3. Biosenseurs et circuit logique	41
1.5.4. Besoin de circuits logiques dans le cancer.....	42
2. Contexte de recherche et problématique.....	45
3. Hypothèse et objectifs	48
3.1. Hypothèse	48
3.1.1. Sous-hypothèses.....	48
3.2. Objectifs	48

4. Article 1: Computer-Aided Design of Active Pseudoknotted Hammerhead Ribozymes	51
4.1. Mise en contexte.....	51
4.2. Contribution des auteurs.....	55
4.3. Abstract.....	56
4.4. Introduction.....	57
4.5. Materials.....	61
4.6. Methods.....	62
4.6.1. Installation and set-up of Enzymer.....	62
4.6.2. Designing ribozymes using Enzymer.....	64
4.6.3. <i>In vitro</i> testing of cis acting ribozymes.....	69
4.7. Notes.....	76
5. Article 2: Computational design and experimental verification of pseudoknotted ribozymes	82
5.1. Mise en contexte.....	82
5.2. Contributions des auteurs.....	84
5.3. Abstract.....	85
5.4. Introduction.....	87
5.5. Results.....	90
5.5.1. Synthetic hammerhead ribozymes with a pseudoknot for the stem I-II interaction.....	90
5.5.2. Pseudoknotted hammerhead ribozymes overlapping coding sequence.....	94
5.5.3. Synthetic glmS ribozymes activated by glucosamine-6-phosphate.....	97
5.6. Discussion.....	101
5.7. Materials and methods.....	105
5.8. Acknowledgments.....	110
5.9. Supplementary Information.....	112
6. Article 3: A Web Service for Automated Design of Multiple Types of Ribozymes Targeting RNA: from Minimal Hammerhead to Aptazyme	125
6.1. Mise en contexte.....	125
6.1. Contributions des auteurs.....	128
6.2. Abstract.....	129
6.3. Introduction.....	131
6.4. Materials and methods.....	134
6.5. Results.....	143
6.5.1. Comparison of end-point cleavage of pistol, extended HHRZ and pseudoknotted HHRZ.....	144
6.5.2. Kinetics measurements.....	145

6.5.3.	FMN Aptazymes	148
6.6.	Discussion.....	150
6.7.	Supplementary material	155
7.	Portes logiques basés sur des ribozymes allostériques contrôlés par des miARN ..	164
7.1.	Mise en contexte.....	164
7.2.	Circuit logique.....	164
7.2.1.	Les fonctions logiques.....	164
7.2.2.	Les portes logiques (<i>logic gates</i>)	166
7.2.3.	Circuit logique en biologie.....	166
7.2.4.	Les micros ARN	167
7.2.5.	Les profils d'expression de miARN	168
7.3.	Des ribozymes liant jusqu'à trois miARN	169
7.3.1.	Portes logiques ET et NON-ET	169
7.4.	Des ribozymes ayant une séquence complémentaire au miARN en 3'	174
7.4.1.	Matériels et méthodes	175
7.4.2.	Résultats.....	176
8.	Article 4: Dual selection for aptazymes for biosensor development purpose.....	179
8.1.	Mise en contexte.....	179
8.1.	Contributions des auteurs	181
8.2.	Abstract	182
8.3.	Introduction.....	184
8.4.	Materials and methods.....	186
8.5.	Results.....	198
8.5.1.	Validation of library (C)	198
8.5.2.	SELEX	198
8.5.3.	Sequencing results.....	199
8.5.4.	Reversing the C6 back to A6 consensus core	201
8.5.5.	Aptazyme activity is modulated by galectins	201
8.5.6.	Aptazyme-based biosensor development	204
8.6.	Discussion.....	206
9.	Discussion générale.....	210
9.1.	La conception des ARN synthétiques incluant des pseudonoeuds	210
9.2.	Conception d'un système d'inhibition de gènes.....	213
9.3.	Importance de la validation expérimentale	216
9.4.	Les aptamères : un exemple concret du manque de données expérimentales .	217
9.5.	Le choix des ribozymes	218
9.6.	Ribozymes allostériques	222

9.6.1. Portes logiques.....	222
9.6.2. Aptazymes en biosenseurs	224
10. Conclusion et perspectives	229
11. Références	232
Annexes.....	266
Annexe I.....	266
Annexe II.....	272
Annexe III.....	282

Liste des figures

Figure 1.1	Les états de structure du ribozyme en tête de marteau.....	5
Figure 1.2	Structure et séquence conservées du ribozyme en tête de marteau.	7
Figure 1.3	Les deux structures du ribozyme en tête de marteau.....	8
Figure 1.4	Les trois types de ribozyme en tête de marteau.	9
Figure 1.5	Ribozyme en tête de marteau actif en trans	10
Figure 1.6	Structure secondaire du ribozyme pistol.....	11
Figure 1.7	Mécanisme des riboswitches	12
Figure 1.8	Structure tertiaire et secondaire du ribozyme GImS	14
Figure 1.9	Activité d'un aptazyme	16
Figure 1.10	Changement structural de l'aptamère.....	18
Figure 1.11	Ribozyme allostérique inhibé par la liaison d'un oligonucléotide	22
Figure 1.12	Librairie de sélection d'un ribozyme actif en <i>trans</i>	25
Figure 1.13	Protocole de SELEX.....	27
Figure 1.14	Chronologie de différentes méthodes de sélection et conception de ribozymes.....	38
Figure 3.1	Circuit logique possible = thérapie intelligente.....	50
Figure 4.1	Utilisation de « Enzymer ».....	52
Figure 4.2	Pseudoknot types as illustrated by various hammerhead ribozymes stem I and II ..	57
Figure 4.3	Ribozyme in cis and trans acting versions	73
Figure 4.4	Visualization of radioactive bands for purification	75
Figure 4.5	Assembly PCR	77
Figure 4.6	Size range of markers and RNA in PAGE.....	79
Figure 5.1	Ribozymes designed according to Mouse gut metagenome ribozyme structure	92
Figure 5.2	<i>Yarrowia lipolytica</i> ribozymes overlapping RFP coding sequence.	96
Figure 5.3	GImS ribozymes	99
Figure 5.4	Inline probing of GImS ribozymes	101
Figure 5.5	Secondary structures of ribozymes.....	114
Figure 5.6	Ribozymes activity.....	119
Figure 5.7	Discrepancies between the consensus from (McCown et al. 2011) and our model	123
Figure 5.8	Cleavage extent of <i>cis</i> and <i>trans</i> acting ribozymes.....	124
Figure 6.1	Conception de ribozymes actifs en trans par Ribosoft2.0.....	126
Figure 6.2	Hammerhead ribozyme structure.	133
Figure 6.3	Complete overview of Ribosoft (V2) program activity.....	135
Figure 6.4	Input specifications provided by user.	136
Figure 6.5	Output designs presented to user	137
Figure 6.6	<i>Trans</i> -acting ribozymes designed by Ribosoft 2.0	140

Figure 6.7	End point cleavage assays of ribozymes against Gal-7.	145
Figure 6.8	Scatter plots showing the kobs of each ribozyme compared to the score of each one of the parameters given by Ribosoft 2.0.....	147
Figure 6.9	End point cleavage tests of activated and inhibited version of <i>trans</i> acting FMN aptazymes against Gal-7 mRNA in presence and absence of FMN.....	149
Figure 6.10.	Examples of Bcl-2 mRNA cleavage results	155
Figure 6.11	État du circuit logique	163
Figure 7.1	Activité du micro-ARN.	168
Figure 7.2	Structure d'un ribozyme 'porte logique ET' liant trois miARN	171
Figure 7.3	Les portes logiques 'ET' à deux et trois entrées	173
Figure 7.4	Mode d'action des ribozymes et des miARN d'un ribozyme contrôlé par un miARN	175
Figure 7.5	Test de clivage des ribozymes contre l'ARNm de la luciférase	177
Figure 7.6	Ribozymes détecteurs de miARN	178
Figure 8.1	Library sequence design	187
Figure 8.2	Affinity Selection steps	191
Figure 8.3	Activity based selection steps	193
Figure 8.4	Aptazyme-mango aptamer complex functionalized on streptavidin bead.....	196
Figure 8.5	Libraries activity in presence of magnesium or manganese	198
Figure 8.6	Predicted structures of two aptazymes binding Gal-7.....	200
Figure 8.7	C6 aptazymes and with reverted mutation (C6A)	201
Figure 8.8	Aptazyme cleavage activity	203
Figure 8.9	Confocal microscope images of functionalized beads after cleavage reactions.....	205
Figure 8.10	Composantes du circuit final	209

Liste des tableaux

Table 1.1	Différents logiciels basés sur l'approche de 'inverse folding'	37
Table 4.1	Pseudoknotted ribozymes	59
Table 4.2	Example of hammerhead sequences.....	70
Table 4.3	Example of a Time table.....	80
Table 5.1	RNA sequence data generated by Enzymer for all ribozymes.	107
Table 5.2	Sequences of trans-acting ribozymes, their substrates and corresponding oligonucleotides	112
Table 5.3	Oligonucleotide sequence for cis-acting ribozymes	120
Table 6.1	Oligonucleotides sequences	158
Table 6.2	Parameters scores of chosen ribozymes	161
Table 7.1	Table de vérité de la fonction ET.....	165
Table 7.2	Table de vérité de la fonction OU	165
Table 7.3	Table de vérité de la fonction NON	165
Table 8.1	Summary of libraries and selection	188
Table 8.2	Sequences for library amplification	188
Table 8.3	SELEX rounds and conditions.....	191
Table 8.4	Negative selection conditions.....	192
Table 8.5	Aptazymes sequences	194
Table 8.6	Primers and biotinylated oligo sequence.....	197
Table 8.7	Frequency of chosen sequences (from sequencing results).....	199
Table 8.8	Inactive aptazymes with reverted mutation	201
Table 9.1	Résumé des systèmes utilisés en thérapies et leurs avancés.....	220
Table 9.2	Comparaison des aptamères/aptazymes et les anticorps	224

Liste des abréviations

32P	Radioisotope du phosphore
A	Adénine
ADN	Acide DésoxyriboNucléique
ADNc	ADN complémentaire
AdoCbl	AdénosylCobalamine
APS	Ammonium PerSulfate
ARN	Acide RiboNucléique
ARNm	ARN messenger
ARNnc	ARN non-codant
ARNr	ARN ribosomal
ARNt	ARN de transfert
ASO	Oligonucleotides AntiSens
ATP	Adénosine TriPhosphate
Bcl-2	<i>B-cell lymphoma 2</i>
BSA	Albumine de sérum bovin
BLAST	Basic Local Alignment Search Tool
C	Cytidine
c-di-GMP	di-Guanosine MonoPhosphate cyclique
CLEC2	<i>C-type Lectin Like receptor II</i>
CRISPR	<i>Clustered Regulatory Interspaced Short Palindromic Repeats</i>
CTE	Élement de Transport Constitutif
CTP	Cytidine triphosphate
DMEM	<i>Dulbecco's Modified Eagle Medium</i>
dNTP	Désoxyribonucléotide triphosphate
dsDNA	AND double brin
DTT	Dithiothreitol
ECL	<i>Enhanced Chemiluminescence</i>
EDTA	Ethylène Diamine Tétra-Acétique
ELISA	<i>Enzyme-Linked Immunosorbent Assay</i>
FLC	<i>Full Length Circularization</i>

FMN	Flavin Mononucléotide
FRET	<i>Forster Resonance Energy Transfer</i>
G	Guanine
Gal-7	Galectine-7
GlcN6P	Glucosamine-6-Phosphate
GTP	Guanosine TriPhosphate
HCV	Virus de l'Hépatite C
HDV	Virus de l'Hépatite Delta
HEPES	Acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique
Hhd	Hammerhead
HHRz	Ribozyme en tête de marteau
HIV	Virus d'Immunodéficience Humaine
IGEPAL	Octylphénoxy polyéthoxyéthanol
IPTG	Isopropyl β -D-1-thiogalactopyranoside
Kd	Constante de dissociation
Km	Constante de <i>Michaelis</i>
LB	<i>Luria Broth</i>
lncARN	Long ARN non-codant
MFE	Énergie minimale libre
miARN	micro-ARN
MMP9	Métalloprotéinase matricielle 9
Nct	Nucléotides
NGS	Séquençage à haut débit
PEG	Polyéthylène glycol
PVDF	Difluorure de polyvinylidène
RBS	Site de liaison du Ribosome
RFP	Protéine fluorescente rouge
RISC	<i>RNA-induced silencing complex</i>
Rnase	RiboNucléase
PABPN1	Polyadenylate-Binding Nuclear Protein 1
rNTP	Ribonucléoside tri-phosphate

RPM	Rotation Par Minute
SAM	S-AdénosylMéthionine
SD	Shine Dalgarno
SDS-PAGE	<i>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</i>
SELEX	<i>Systemic Evolution of Ligand by Exponential enrichment</i>
siARN	Petit ARN interférents
SINE	<i>Short Interspersed Nuclear Elements</i>
snoARN	Petit ARN nucléolaire
TAE	Tris Acétate EDTA
TALEN	<i>Transcription Activator-Like Effector Nuclease</i>
TBE	Tris Borate EDTA
TBS	Tris-buffered saline
TEMED	TEtraMéthyl Ethylène Diamine
TPP	Thiamine pyrophosphate
U	Uridine
UTP	Uridine triphosphate
UTR	Région non-traduite
VEGF	Facteurs de croissance de l'endothélium
VEGFR-1	Récepteur de facteurs de croissance de l'endothélium-1
ZFN	<i>Zinc Finger Nucleases</i>

1. Revue de littérature

1.1. Aperçu de découvertes des ARN non-codants

L'historique de l'ARN remonte au début des années 1900, lorsqu'il a été découvert dans les levures. Pendant longtemps, les ARN cellulaires, par opposition aux ARN d'autres origines comme les génomes viraux, ont été considérés comme ayant seulement le rôle de messagers entre l'ADN, qui est le porteur de l'information génétique, et les protéines, qui sont les entités fonctionnelles dans les cellules. Une séquence d'ARN messager (ARNm) est formée par des régions nommées « exons », qu'on appelle codantes, car ce sont les régions traduites en protéine ; et ce qui reste se nomme de l'ARN non codant (ARNnc). Les ARNnc incluent les régions 3' et 5' UTR (*untranslated regions*) et les introns, de même que le reste des ARN non traduits en protéine. Ces régions ont longtemps été considérées comme un surplus non utilisé pour une fonction biologique dans la cellule (Palazzo & Lee, 2015). Cette idée excluait les ARN ribosomiaux (ARNr) (Lamfrom *et al.*, 1966; Palade, 1955) et les ARN de transfert (ARNt) (Hoagland *et al.*, 1958; Rich & RajBhandary, 1976) découverts plusieurs années avant que le code génétique ne soit déchiffrés dans les années 60 (Crick *et al.*, 1961; Khorana *et al.*, 1966) . Cependant, la majorité du génome humain, incluant les régions non-traduites des pre-ARNm, n'avait pas de rôles connus jusqu'à des travaux beaucoup plus récents. Ce n'est qu'à partir de 1980 que de nouvelles recherches ont permis l'identification et la définition de plusieurs classes d'ARN ayant différentes fonctions.

L'une des découvertes les plus importantes a été celle du groupe de Cech (1982) (Cech, 1986; Kruger *et al.*, 1982): il a prouvé la présence d'une activité catalytique de l'ARN en essayant d'identifier les molécules impliquées dans le processus de l'épissage. Une autre découverte à la même période a permis d'identifier l'activité catalytique de la composante ARN de la RNase P (Guerrier-Takada *et al.*, 1983). Ceci a mené à l'identification de plusieurs familles d'ARN ayant des activités catalytiques semblables à celle des enzymes (Pyle, 1993). Ces ARN ont été nommés « ribozymes », parmi ceux-ci, les ARNr ont été confirmés comme étant entièrement responsables de la formation des liaisons peptidiques entre les acides aminés au cours de la traduction (Noller *et al.*, 1992). D'ailleurs, l'existence des ARN catalytiques a été la raison de l'apparition de la théorie du « monde d'ARN » (« *the RNA world* ») (Cech, 2012; Neveu *et al.*, 2013), ce qui suggère que la

molécule à l'origine l'apparition de la vie n'était pas de l'ADN, mais plutôt de l'ARN catalytique. Ce dernier aurait à la fois eu l'information génétique nécessaire et la capacité de s'auto-répliquer.

Hormis l'activité catalytique de l'ARN, son implication dans la régulation de l'expression génique, sous différentes formes et dans plusieurs processus, a été démontrée. Cette régulation peut avoir lieu par une activité en *cis* comme celle des *riboswitchs* qui sont capables d'inhiber ou d'activer la transcription ou la traduction du gène en amont ou en aval de leurs séquences ; ceci en réponse à la liaison d'un ligand spécifique (Bedard *et al.*, 2020). Elle peut aussi se faire en *trans*, comme par l'effet des miARN (micro-ARN) (Dexheimer & Cochella, 2020) et des petits ARN interférents (siRNA) (Dana *et al.*, 2017) qui reconnaissent des séquences d'ARN spécifiques et inhibent leurs traductions à l'aide du complexe RISC (*RNA-induced silencing complex*). D'autres ARN, comme les ARN guides du système CRISPR (*Clustered Regulatory Interspaced Short Palindromic Repeats*), agissent en se liant à l'ADN par complémentarité de séquence, ce qui permet le recrutement d'une protéine Cas9 qui effectue la coupure de cet ADN (Adli, 2018). Une autre stratégie est adoptée par les lncARN (*long non-coding RNA*), qui contrôlent la transcription par le recrutement des facteurs de transcription afin d'activer l'expression d'un gène adjacent ou en formant des complexes avec d'autres facteurs, en empêchant ainsi la liaison de leurs séquences cibles (Bhat *et al.*, 2016). Toutes ces découvertes de différentes fonctionnalités et de classes d'ARN ont également permis une meilleure compréhension de plusieurs maladies, puis ont ouvert la porte à de nouvelles cibles et de nouveaux schémas thérapeutiques (Bhatti *et al.*, 2021).

En plus des fonctions particulières qui définissent chaque famille ou classe d'ARNnc, il a été démontré que les structures secondaires et tertiaires des ARN avaient un rôle majeur dans cette classification. En effet, un ARN non codant appartenant à une certaine classe est défini par sa structure (Pang *et al.*, 2006). Alors, en cherchant de nouveaux ARN ayant la même fonctionnalité, habituellement la première étape consiste en la recherche des ARN ayant la même structure conservée (Huang *et al.*, 2016) (Boccaletto *et al.*, 2018). Cette dépendance de l'ARN par rapport à sa structure a mené au développement d'ARN synthétiques qui conservent l'activité d'un ARN naturel en leur donnant la structure fonctionnelle et la séquence nécessaire à l'activité, tout en ajoutant des modifications qui permettent de répondre aux conditions et d'effectuer les actions désirées et prédéfinies (Isaacs *et al.*, 2006). Ce domaine comporte plusieurs applications dans la recherche

fondamentale, car les ARNnc synthétiques ont permis de construire des outils de contrôle de l'expression génique personnalisée, et ce, dans le but d'avoir un *knock-out* ou *knock-down* d'un gène désiré afin de déterminer son activité ou son effet sur une autre molécule (Lagana *et al.*, 2014). Pour la recherche appliquée, les approches sont diverses et nombreuses, allant du simple contrôle d'un gène à la construction de circuits formés par de nombreux ARN ayant différentes fonctionnalités (Kim & Franco, 2020). Outre les ARN naturels et synthétiques qui ont des fonctions connues, des ARN avec de nouvelles fonctions ont été développés ou sélectionnés *in vitro*. Parmi les exemples ayant beaucoup d'intérêt et d'applications différentes, nous retrouvons les aptazymes, formés par la combinaison de ribozymes et d'aptamères. En réalité, depuis leur découverte, les structures de ribozymes ont été sujettes à plusieurs approches d'ingénierie afin de les inclure dans différentes applications.

1.2. Ribozyme

Les ribozymes sont des ARNnc ayant des activités catalytiques semblables à celles effectuées par des enzymes, d'où vient leur appellation de « ribozyme » (Walter & Engelke, 2002). Les ribozymes sont classés selon la réaction effectuée. Un ribozyme peut être un ARN solitaire agissant par lui-même en tant que molécule, n'ayant pas besoin de protéine chaperonne, ni de cofacteur. L'autre type d'ARN catalytique représente des ARN jouant le rôle du cœur catalytique d'un complexe ARN-protéine. Pour nommer l'un des premiers ribozymes découverts, nous citons la RNase P, qui est un ARN formant un complexe avec des protéines (Guerrier-Takada *et al.*, 1983). Pour cette raison, ces ARN ont toujours été considérés comme une entité structurale non catalytique qui aurait simplement joué le rôle de support aux protéines effectuant la réaction et de guide pour reconnaître et lier les cibles (Akaboshi *et al.*, 1980). Cette idée a changé après l'étude de Thomas Cech, qui a montré que le responsable de la réaction d'épissage n'est rien d'autre que l'intron de l'ARNm. Il a été le premier à publier une étude mentionnant une activité d'auto-clivage de l'ARN (Kruger *et al.*, 1982). La confirmation est venue quelques années plus tard avec l'étude démontrant l'activité des « Intron de type I » responsable de la réaction de clivage durant l'épissage de l'ARNm (Cech, 1986). La même histoire se répète avec la RNase P; étant une vraie enzyme responsable de la maturation des ARNt, son cœur catalytique a été prouvé comme étant la partie ARN du complexe (Guerrier-Takada *et al.*, 1983). Ces deux découvertes ont reçu le prix Nobel de chimie en 1989. La confirmation a également été affirmée pour l'ARN 23S de la grande sous-unité

ribosomale, responsable de l'activité peptidyl transférase, qui est effectuée par l'ARN et non par la composante protéique (Nissen *et al.*, 2000; Noller *et al.*, 1992). Plusieurs familles d'ARN catalytique ont été identifiées depuis ; cela continue encore aujourd'hui (Weinberg *et al.*, 2019a). Ces recherches ont montré que les ribozymes existent chez toutes les classes d'êtres vivants et avec différentes fonctions, principalement la ligation et le clivage (Hammann *et al.*, 2012; Roth *et al.*, 2014a; Webb *et al.*, 2009).

Parmi les ribozymes les plus étudiés et utilisés se trouvent, les petits ribozymes auto-clivants (en tête de marteau, *twister* et *twister sister*, pistol, HDV (*hepatitis delta virus*), Varkud Satellite, hairpin, etc.). (Ferre-D'Amare & Scott, 2010). Ils présentent l'avantage d'avoir une petite taille, donc plus faciles à manipuler en laboratoire. Dans ce projet, notre travail a porté principalement sur le ribozyme en tête de marteau, puis sur quelques autres types.

1.2.1. Ribozyme en tête de marteau

Le ribozyme en tête de marteau est un petit ARN catalytique, découvert en 1986 (Prody *et al.*, 1986) dans l'ARN satellite du « *tobacco ringspot virus* ». Il constituait le troisième ARN catalytique à être découvert et demeure depuis le ribozyme le plus étudié dans le domaine de la catalyse d'ARN. Il a été identifié chez presque tous les règnes d'êtres vivants, incluant les procaryotes (bactéries et archaea, de même que les virus, viroïdes et bactériophages) et les eucaryotes mono et pluricellulaires (de la Pena & Garcia-Robles, 2010; Hammann *et al.*, 2012) (Perreault *et al.*, 2011b). Les ribozymes des eucaryotes sont généralement positionnés dans des introns et des séquences répétées en tandem, comme dans le cas de ribozymes identifiés dans le génome humain. Bien que la fonction *in vivo* de ces ribozymes demeure encore inconnue, parmi ceux-ci, deux sont conservés chez d'autres vertébrés ; leur activité a été prouvée *in vitro* (Hammann *et al.*, 2012).

1.2.1.1. Fonction biologique et mécanisme d'action

L'auto-clivage du ribozyme en tête de marteau suit le même principe de la dégradation spontanée de l'ARN par la réaction de « *in-line attack* ». Cette réaction, aussi nommée transestérification, se déroule entre les nucléotides du cœur catalytique, qui sont le G12, le C17, le G8 et le C1.1. Le clivage aura lieu dans le site de clivage RUH, où le H est souvent un C. Tout d'abord, « N1 » du nucléotide G12 est déprotoné par une molécule d'eau. Le G12 ionisé devient la base générale dans la réaction d'acide-base impliquée dans le clivage et cause à son tour la déprotonation du groupement 2'-OH du C17. Dans

la structure, les deux nucléotides sont alignés de façon à permettre une substitution nucléophile (S_N2), d'où vient le nom « *in-line attack* ». Le C17 transformé en « nucléophile » du site catalytique attaque le groupement phosphate de C 1.1 (N 1.1 dans **Figure 1.1**). La réaction se poursuit quand le groupement 2'-hydroxyl du G8, qui a cédé un proton au 5'-O du C1.1, par le biais de la catalyse acide, regagne un proton en arrachant celui que le G12 avait gagné. Le clivage aura finalement lieu entre le C et le N1.1 du site actif en générant une extrémité 5'-OH et un phosphate cyclique 2',3' (Schultz *et al.*, 2014) (**Figure 1.1**).

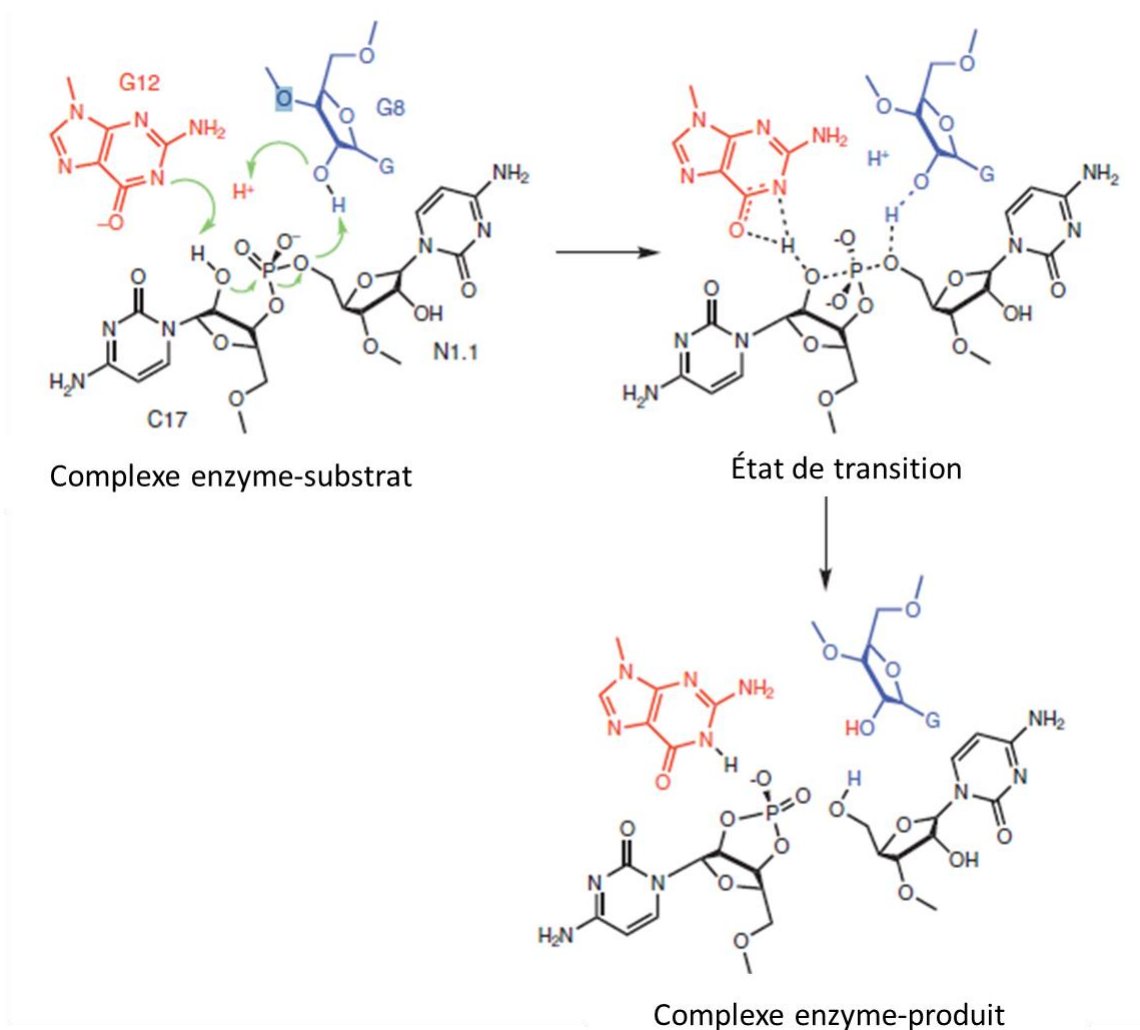


Figure 1.1 Les états de structure du ribozyme en tête de marteau.

Tiré de (Scott, 2010)

L'état de transition est montré par le mouvement du proton H^+ des nucléotides en position de *inline*, par arrachement, ce qui fini par la coupure de la liaison phosphodiester entre le C17 et le N1.1.

Malgré la présence de séquences de ribozymes chez différents types de génomes, la fonction biologique de ces derniers demeure inconnue pour la grande majorité d'entre eux. Les seuls génomes chez lesquels les ribozymes ont un rôle bien défini sont les virus satellites et les viroïdes qui dépendent des ribozymes pour libérer la nouvelle copie du génome. Lors de la réplication du génome circulaire du virus, les copies générées font partie de la même molécule d'ARN multimérique. Ceci se produit jusqu'à la formation de la structure active du ribozyme qui procède à l'auto-clivage de l'ARN, générant des copies individuelles du génome (Scott, 2010). Chez les autres espèces, des recherches proposent que la présence de ribozymes à proximité des régions codantes suggère leur implication dans le contrôle de l'expression génique (de la Pena & Garcia-Robles, 2010; Wurmthaler *et al.*, 2018). Un exemple s'applique sur des séquences de ribozymes conservées chez des mammifères ; leurs structures actives sont formées dans la région 3' UTR d'un gène codant la protéine CLEC2 (C-type lectin type II). Il a été démontré que ces ribozymes sont actifs *in vitro* ; le clivage de l'ARN *in vivo* avait résulté par une diminution considérable du taux d'expression de la protéine (Martick *et al.*, 2008; Scott *et al.*, 2013).

1.2.1.2. Structure du ribozyme

La version minimale du ribozyme a été identifiée en 1987 (Uhlenbeck, 1987) (Haseloff & Gerlach, 1989). Elle s'avère être une région centrale correspondant au cœur catalytique et incluant le site de clivage GUC mais en général GUH (H= tous les nucléotides sauf G). Cette séquence de 15 nucléotides rarement variables est entourée par trois tiges ayant des séquences non conservées, à l'exception de trois paires de bases; deux sur la tige III et une sur la tige II (Scott *et al.*, 1996) (**Figure 1.2**) et (**Figure 1.3.A**). Cette version de ribozyme a été nommée minimale, car elle contenait seulement la séquence minimale nécessaire pour effectuer la réaction de clivage, mais avec une vitesse de coupure assez faible. En effet, le ribozyme minimal ne représente qu'un « *turnover* » d'une molécule/minute ainsi qu'une K_m (constante de Michaelis-Menten) de 10 μM . Cette faible activité a été expliquée par l'absence d'arrangement des bases du cœur catalytique dans une position de « *in-line-attack* » (McKay, 1996) (Blount & Uhlenbeck, 2005). Pour compenser cette absence, le ribozyme minimal a besoin d'une concentration élevée d'ions métalliques bivalents comme le Mg^{2+} (10 à 100 mM) (Dahm & Uhlenbeck, 1991). En raison de cette dépendance en ions divalents pour effectuer la réaction du clivage, pendant plusieurs années les ribozymes ont été considérés comme des métallo-enzymes. Ceci a

généralisé l'idée selon laquelle les ions métalliques sont impliqués dans la réaction acide-base résultant de la coupure de la liaison phosphodiester (Dahm *et al.*, 1993).

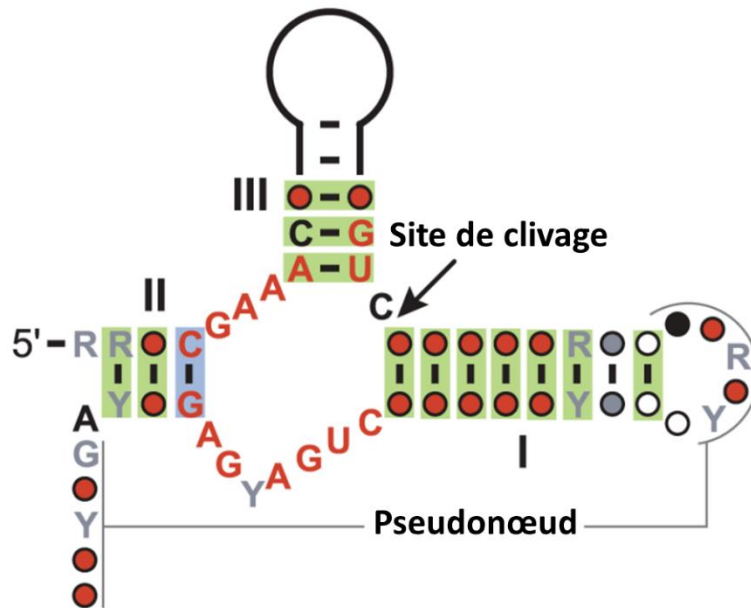


Figure 1.2 Structure et séquence conservées du ribozyme en tête de marteau.

Le cœur catalytique est formé par quinze nucléotides conservés (en rouge), entouré par trois tiges/tiges boucles. La flèche indique le site de clivage GUC ici (mais GUH en général). Le pseudonœud est une interaction tertiaire entre la boucle d'une tige avec une région simple brin d'une autre tige du ribozyme. Adapté de (Perreault *et al.*, 2011)

Le concept de métalloenzymes a été démenti par la preuve que certains types de ribozymes conservent une activité de clivage en l'absence d'ions métalliques. Dans ce cas, ces derniers sont remplacés par d'autres ions chargés positivement comme le Li^+ (Khvorova *et al.*, 2003). Ce qui est venu confirmer l'absence de cette relation ; fut la détermination de la structure d'une version mille fois plus active que la minimale. Cette structure était celle du ribozyme complet trouvé chez *Schistosoma mansoni*, qui a été publiée en 2006 et qui avait une séquence plus longue que la version minimale (Martick & Scott, 2006). En effet, la différence que cette structure présentait était la présence d'un surplus de séquences à la tige I et II, permettant des interactions tertiaires absentes chez la version minimale. Cette interaction peut prendre l'allure d'un pseudonœud qui se forme par l'appariement de bases entre deux régions, simples brins sur les tiges I et II (**Figure 1.3.B**). L'interaction tertiaire stabilise la structure fonctionnelle du ribozyme en permettant le réarrangement des bases du cœur catalytique en position du « *in-line attack* » (Martick & Scott, 2006). Ainsi, il a été prouvé que le besoin en Mg^{2+} , dans le cas de la version minimale, est plutôt pour ses charges positives, qui neutralisent les charges négatives et

diminuent la répulsion électrostatique entre les nucléotides du cœur catalytique déprotoné. Dans le cas de la version complète, une concentration moins élevée (de l'ordre de micromolaire) est suffisante pour avoir un taux de clivage et un *turnover* plus élevés. Le besoin en Mg^{2+} est en fait compensé par le changement conformationnel causé par la structure tertiaire (Khvorova *et al.*, 2003), qui s'est avérée essentielle. Ainsi, la déstabilisation de cette interaction via des mutations dans la séquence formant le pseudonœud mène à une diminution remarquable, ou même à une perte totale de clivage (Perreault *et al.*, 2011b).

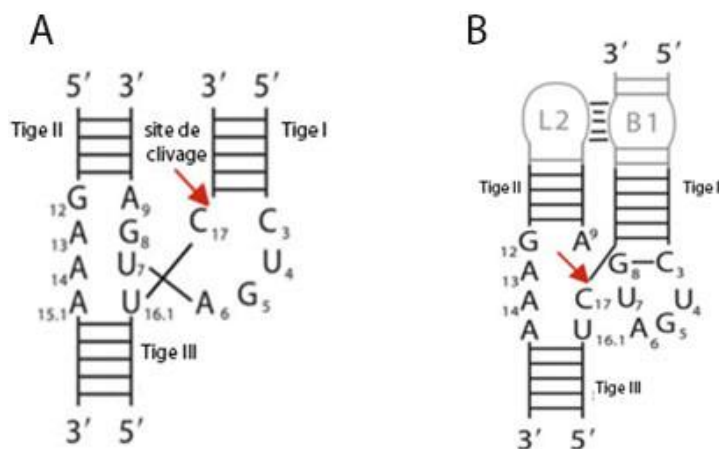


Figure 1.3 Les deux structures du ribozyme en tête de marteau

A : La structure du ribozyme minimal contenant le cœur catalytique, le site de clivage (flèche rouge) et les tiges I, II et III. **B** : la structure complète du ribozyme ayant les mêmes composantes que le ribozyme minimal avec des séquences additionnelles sur les tiges I et II avec B1 : bulge (renflement) 1, L2 : boucle 2 qui forment l'interaction tertiaire stabilisant la structure du ribozyme. Adapté de (Scott *et al.*, 2013).

1.2.1.3. Modification de la structure du ribozyme

Le ribozyme en tête de marteau a été étudié sous tous ses angles ; différentes versions ont été trouvées actives. Le positionnement des extrémités 3' et 5' définit le type de ribozyme. Puisqu'il y a trois tiges, les ribozymes sont de type I, II ou III, si les extrémités 3' et 5' se trouvent sur la tige I, II ou III, respectivement (Hamman *et al.*, 2012) (**Figure 1.4**).

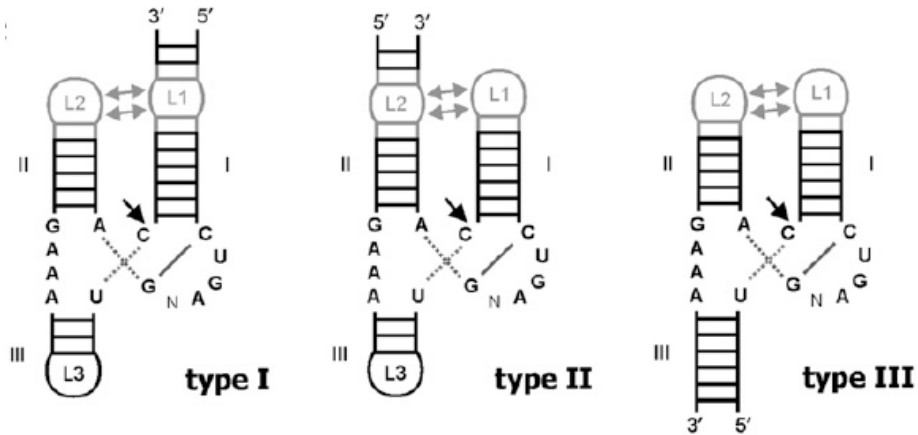


Figure 1.4 Les trois types de ribozyme en tête de marteau.

Le site de clivage est indiqué par la flèche, les numéros de tiges sont indiqués par I, II et III, L1, L2 et L3 sont des boucles. Tiré de (Hammann *et al.*, 2012)

Bien que le ribozyme en tête de marteau naturel soit actif *en cis*, sa séquence peut être modifiée pour qu'il soit actif *en trans* afin de le rendre actif contre d'autres ARN d'une façon spécifique. Ceci est effectué en conservant la séquence du cœur catalytique et la tige II, mais les tiges I et III auront des séquences simples brins conçues pour être complémentaires à la séquence de l'ARN visé. La séquence cible contient le site de clivage RUH (où R est un A ou G, et H est n'importe quelle base sauf G), mais préférentiellement GUC (Scott *et al.*, 2013) (**Figure 1.5**).

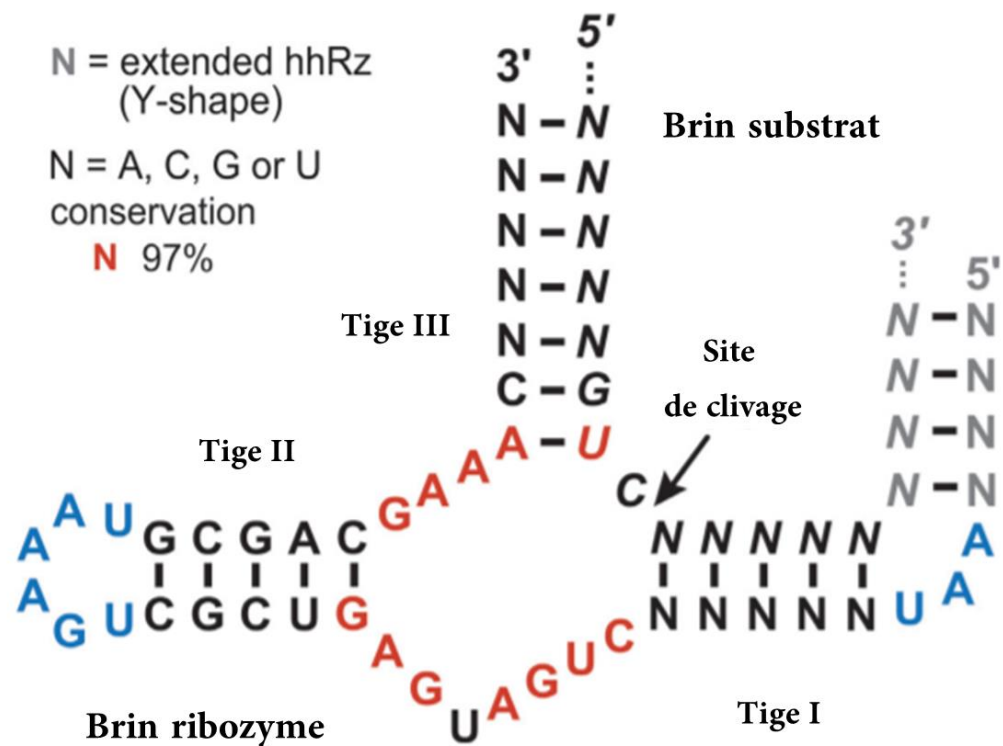


Figure 1.5 Ribozyme en tête de marteau actif en trans

Les N représentent les nucléotides modifiables pour avoir une complémentarité avec la cible. En rouge : les nucléotides du cœur catalytique conservés. En bleu : les nucléotides impliqués dans l'interaction tertiaire. Adapté de (Kharma *et al.*, 2016)

En effet, ce principe a été largement utilisé pour le contrôle des gènes dans plusieurs études, dans une optique de thérapie génique (Citti & Rainaldi, 2005; Mulhbacher *et al.*, 2010). Des ribozymes ont été conçus contre des gènes de virus tels que l'hépatite C (HCV) (Sakamoto *et al.*, 1996) et le virus de l'immunodéficience humaine (VIH) (Bai *et al.*, 2002). Ils ont également été appliqués contre des gènes impliqués dans le cancer, tel que le C-myc, surexprimé dans certains cancers (Cheng *et al.*, 2000). Des ribozymes actifs en *trans* ont été sélectionnés dans des cellules afin de trouver une meilleure alternative aux ribozymes naturels, moins efficaces dans les essais intracellulaires (Huang *et al.*, 2019).

1.2.2. Ribozyme Pistol

Plusieurs ribozymes, dont le Pistol, ont été découverts en 2015 via une analyse génomique comparative ; ils se sont ajoutés à la liste des ribozymes ayant une activité auto-catalytique (Weinberg *et al.*, 2015a). Le ribozyme Pistol possède une structure de trois tiges (P1, P2 et P3) connectées entre elles par deux boucles et un pseudonœud. Ce

dernier se forme entre les boucles 1 et 2 ; il joue un rôle dans la stabilisation de la structure dans une géométrie permettant la formation du site actif. La région entre la tige P1 et P2 contient trois adénosines conservées qui forment un « *A-minor motif* » interagissant avec la tige P1. Ceci permet la fixation de la tige P2 dans une conformation spatiale, qui donne aussi lieu à la formation du site actif constitué des boucles 1 et 2 (**Figure 1.6**). Le triple A du « *A minor motif* » fait partie des dix résidus conservés du ribozyme Pistol.

La réaction d'auto-clivage effectuée par le ribozyme Pistol est une réaction de transestérification, comme celle du ribozyme en tête de marteau. Pour ce ribozyme, le site de clivage proposé est un GU (G53-U54) (Joseph *et al.*, 2020; Nguyen *et al.*, 2017a).

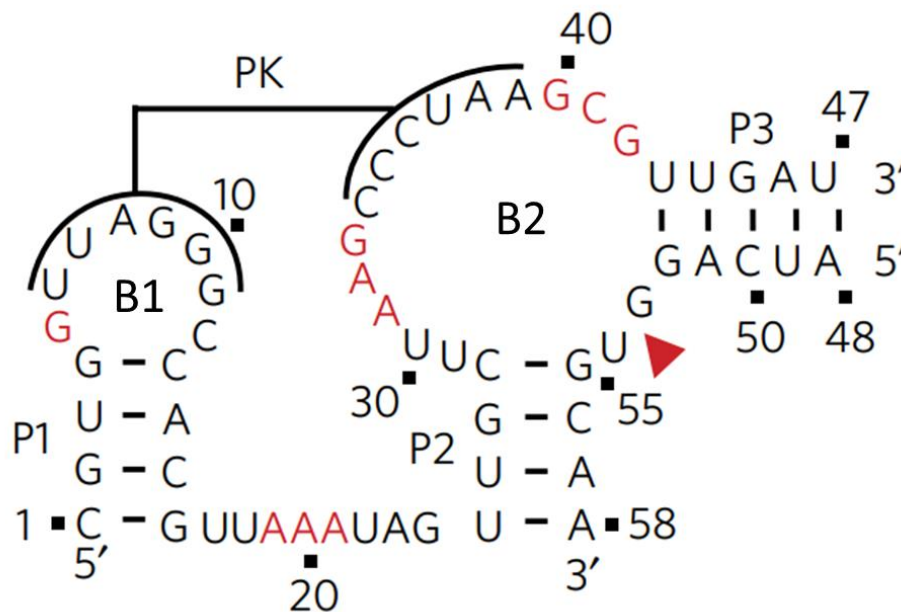


Figure 1.6 Structure secondaire du ribozyme pistol

La structure du ribozyme Pistol est formée par trois hélices (P1, P2 et P3), deux boucles (B1 et B2) et un pseudonœud indiqué par PK. Les nucléotides conservés sont colorés en rouge et le site de clivage est indiqué par le petit triangle rouge. Adapté de (Ren *et al.*, 2016)

1.3. Ribozymes allostériques/*Riboswitchs*

Un ribozyme allostérique est un ARN ayant une dépendance à une molécule spécifique qui cause un changement conformationnel de sa structure afin d'effectuer sa réaction de clivage. Après avoir montré que les ribozymes ne sont pas des métalloenzymes, l'idée selon laquelle il s'agit des molécules allostériques a été éliminée (Khvorova *et al.*, 2003). Les premiers ribozymes allostériques étaient synthétiques, sélectionnés *in vitro* ou par méthodes computationnelles, pendant les années quatre-vingt-dix. Quelques années plus

tard, dans le contexte de découverte d'une nouvelle famille d'ARN, nommée *riboswitches*, un premier *riboswitch* ayant une activité d'autoclivage a été découvert.

1.3.1. Riboswitchs

Les *riboswitchs* se trouvent typiquement dans les régions 3' ou 5' UTR d'un gène donné ; ils sont capables de contrôler leur expression. Dans le cas des *riboswitchs* en 5', ils contrôlent le gène en aval en formant un *switch* ON ou OFF dépendamment de la présence ou de l'absence d'un ligand spécifique. En effet, un *riboswitch* a une structure formée par deux régions : une partie nommée « aptamère », liant le ligand et une plateforme d'expression qui subissent des changements conformationnels selon la liaison ou non du ligand par l'aptamère. Ce changement de structure induit soit l'achèvement ou l'inhibition de la transcription, ou la traduction du gène en aval du *riboswitch* (**Figure 1.7**). De plus, chez les eucaryotes, le mode d'action principal des *riboswitchs* constitue le contrôle d'épissage de l'ARNm. Ce mécanisme conduit à la maturation d'ARNm qui est alors traduit en une protéine complète, ou alors il donne une version défectueuse (Breaker, 2011; Nahvi *et al.*, 2002). Un type unique de *riboswitch* est celui ayant un ribozyme comme plateforme d'expression, contrôlé par la glucosamine-6-phosphate.

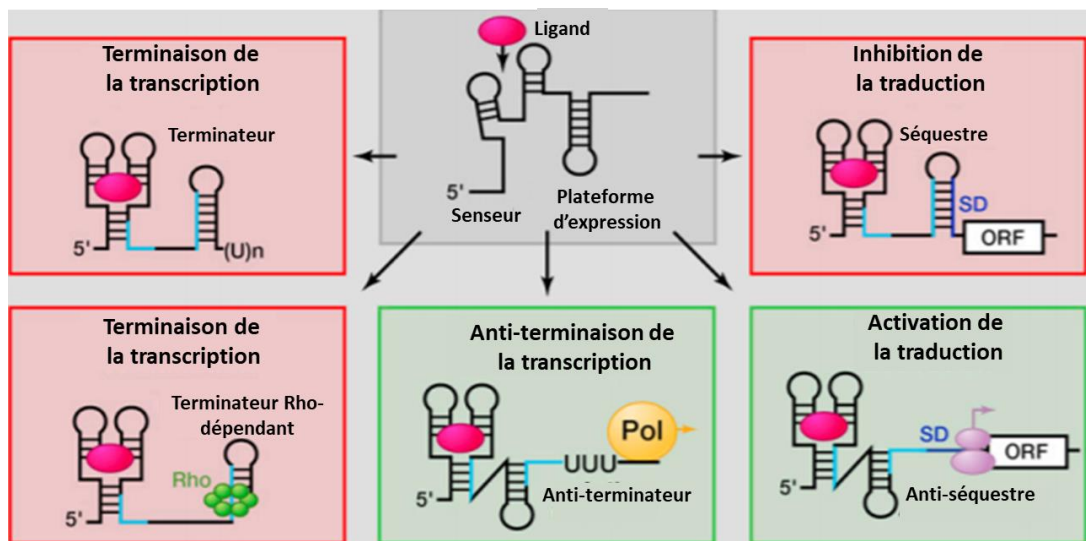


Figure 1.7 Mécanisme des riboswitchs

Les différents modes de contrôle d'expression génique par les riboswitchs. Gauche : la liaison du ligand au riboswitch cause la formation d'un terminateur de transcription mettant fin à cette dernière. Milieu : la structure d'anti-terminateur se forme lors de la capture du ligand par l'aptamère du riboswitch. Droite : le site SD (Shine Dalgarno) peut être séquestré ou libéré par l'effet du changement structural résultant de l'activation du riboswitch lorsque son ligand est présent. Adapté de (Serganov & Nudler, 2013)

1.3.2. Naturels : Ribozyme GImS

La famille des petits ARN auto-clivants comprend un ribozyme allostérique naturel, le ribozyme GImS. Le ribozyme, identifié grâce à une recherche bioinformatique, a été trouvé dans le génome des bactéries gram-positif en 5'-UTR du gène codant l'enzyme *glucosamine 6 phosphate synthase* (Ferre-D'Amare, 2010). Ce ribozyme est aussi considéré comme un *riboswitch*, car il contrôle l'expression du gène en aval. Il se lie spécifiquement avec la glucosamine 6 phosphate (GlcN6P), ce qui induit la réaction d'auto-clivage. La coupure de l'ARN aboutit à l'inhibition de l'expression de la glucosamine 6 phosphate synthase, enzyme essentiel à la synthèse du GlcN6P. Bien que la liaison du GlcN6P au ribozyme ne cause aucun changement structural (il n'est donc pas allostérique au sens strict), le ribozyme GImS est considéré comme un *riboswitch* étant donné son implication dans la régulation génique. Dans les faits, la fixation du GlcN6P par le ribozyme induit l'auto-clivage de ce dernier, ce qui se résulte par la perte du groupement triphosphate protecteur de l'ARNm. Ainsi, l'exposition d'un 5'OH issu de la coupure rend l'ARNm reconnaissable par la RNase J1 qui procède alors à sa dégradation (Ferre-D'Amare, 2010).

Par rapport aux autres petit-ribozymes auto-clivants, le ribozyme GImS présente une structure assez complexe comprenant trois pseudonœuds. La structure secondaire est formée par quatre hélices (P1-P4) formant trois empilements coaxiaux et des régions interconnectées formant trois pseudonœuds (P2.1, P2.2 et P3.1) dont l'un est formé par la connexion, ou par l'interaction entre la région simple brin en 3' du ribozyme et la boucle de l'hélice P3 (Figure 1.7). Les deux autres pseudonœuds forment le cœur catalytique par l'empilement entre des régions de P1-P2.2-P2 et P2.1. Le site de clivage est positionné au premier nucléotide en 5' (Klein & Ferre-D'Amare, 2006; Savinov & Block, 2018) (**Figure 1.8**). Le GlcN6P agit comme cofacteur du ribozyme grâce au groupement 2'-amine, essentiel à la catalyse de la réaction de clivage. Ceci a été démontré par le remplacement du GlcN6P par du glucose 6 phosphate (Glc6P) qui agit comme compétiteur de la GlcN6P. Précisément, le Glc6P peut former la liaison nécessaire avec le cœur catalytique. De ce fait, il agit comme inhibiteur de la liaison du GlcN6P, mais sa présence n'induit pas le clivage, dû à l'absence du groupement amine (McCarthy *et al.*, 2005). L'implication du groupement amine dans la catalyse de la réaction de clivage, effectuée par le ribozyme GImS, a aussi été prouvée par l'activité du ribozyme en présence de glucosamine, tris, L-sérine et éthanolamine, bien que l'activité dans ces conditions était faible par rapport à

celle en présence du GlcN6P (McCarthy *et al.*, 2005). Finalement, il a été démontré que le GlcN6P agit comme acide général dans la réaction acide-base impliquée dans le clivage (Viladoms & Fedor, 2012).

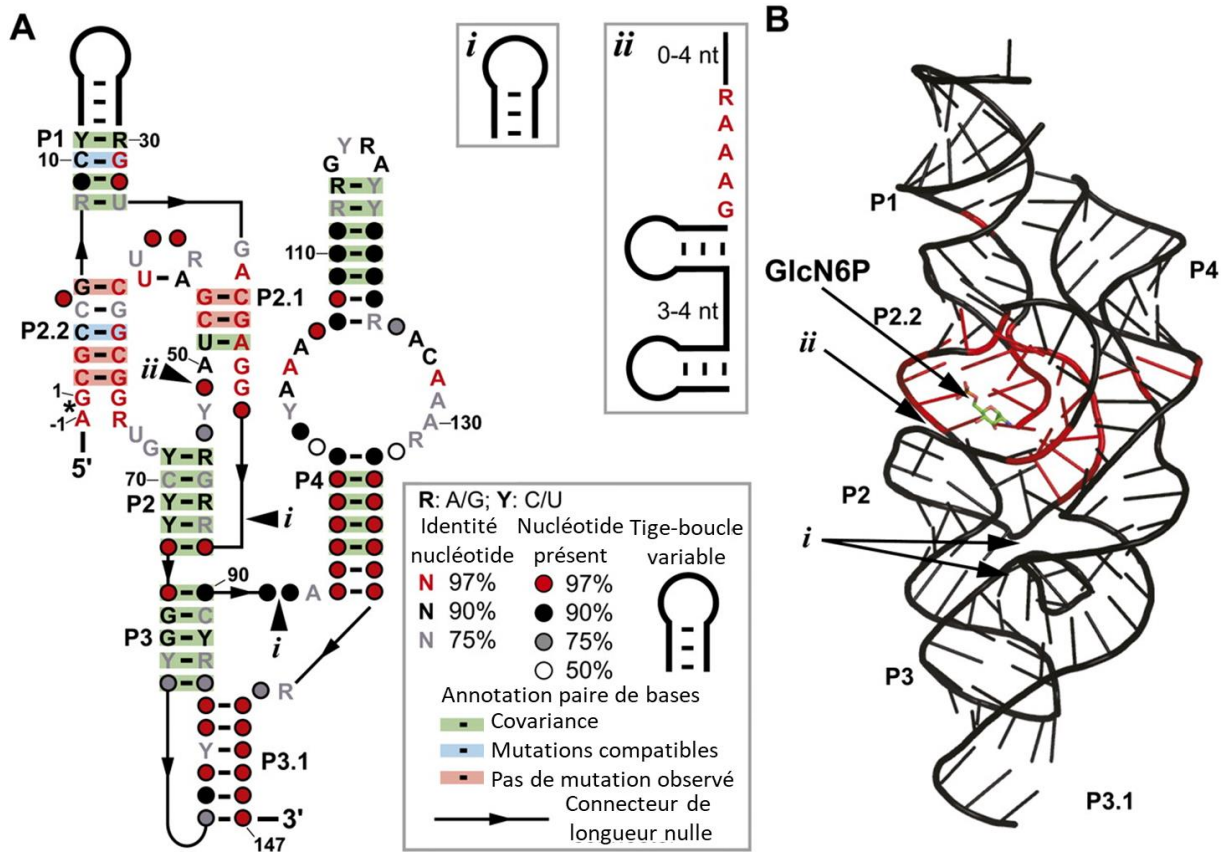


Figure 1.8 Structure tertiaire et secondaire du ribozyme GImS

A : Présentation de la structure secondaire consensus du ribozyme GImS ainsi que des nucléotides conservés (en rouge et noir). **B** : Structure tertiaire formée par le ribozyme en présence de son ligand spécifique GlcN6P. * indique le site de clivage- P1 à P4 : tiges, P2.1, P2.2, P3.1 : pseudonœuds. Adapté de (McCown *et al.*, 2011)

Finalement, ces découvertes ont abouti à la construction de *riboswitchs* synthétiques qui sont la combinaison d'un aptamère et d'une plateforme d'expression ou d'un autre ARN (Etzel & Morl, 2017; Groher & Suess, 2014; Mayo *et al.*, 2020; Yokobayashi, 2019a). La plateforme d'expression peut être un ARN semblable à ceux existant naturellement dans les *riboswitchs* (Desai & Gallivan, 2004), mais d'autres combinaisons sont possibles - notamment avec un ribozyme (Mustafina *et al.*, 2020; Strobel *et al.*, 2020; Wieland *et al.*, 2010; Wieland & Hartig, 2008) nommé aptazyme.

1.3.3. Ribozymes allostériques synthétiques

Les ribozymes auto-clivants présentent les avantages d'être des molécules flexibles, facilement modifiables pour agir sur les cibles choisies ; à l'exception du GImS, il s'agit d'ARN qui ne dépendent d'aucun cofacteur, ligand ou protéine chaperonne pour leur activité. Bien que l'indépendance des ribozymes soit un point puissant pour leur utilisation *in vitro*, ceci peut être problématique lors de leur application *in vivo*, dans le cas où l'on voudrait contrôler leur activité. Afin d'avoir un contrôle sur le clivage effectué par un ribozyme, plusieurs stratégies sont développées. La façon la plus simple étant d'utiliser des promoteurs inductibles pour pouvoir décider quand permettre l'expression du ribozyme (Agop-Nersesian *et al.*, 2008).

Comme il a été démontré que les ribozymes dépendent de la formation de structures secondaires et tertiaires bien définies afin d'avoir une activité catalytique, la modulation de la structure est la clé pour contrôler leur activité. L'autre stratégie utilise aussi l'ingénierie de la séquence du ribozyme ; plutôt qu'utiliser un aptamère, des oligonucléotides ont été utilisés pour moduler sa structure (Penchovsky & Breaker, 2005) (Penchovsky, 2014). Après le développement de la technique SELEX (*Systematic Evolution of Ligands by EXponential enrichment*) permettant d'obtenir des séquences d'aptamères liant des cibles spécifiques, Tang et collaborateurs (Tang & Breaker, 1997) ont développé les premiers ribozymes allostériques dont l'activité est modulée par un aptamère et son ligand. Ce type de ribozymes allostériques a été nommé aptazymes.

1.3.4. Aptazyme

Les aptazymes constituent une classe d'ARN inductibles, ou *riboswitchs* synthétiques, établis pour la première fois par le groupe de Breaker en 1997. Il s'agit de la combinaison d'un aptamère et d'un ribozyme. L'aptamère est le domaine détecteur capable de lier un ligand spécifique. Le ribozyme, pour sa part, est le domaine catalytique qui effectue une réaction de clivage. L'activité du ribozyme devient alors dépendante du ligand de l'aptamère. En effet, la structure fonctionnelle du ribozyme est altérée par la séquence de l'aptamère en l'absence du ligand. En présence de ce dernier, le changement structural de l'aptamère rétablit et stabilise la structure active du ribozyme (**Figure 1.9**) (Tang & Breaker, 1997) (Mustafina *et al.*, 2020; Strobel *et al.*, 2020; Wieland *et al.*, 2010; Wieland & Hartig, 2008).

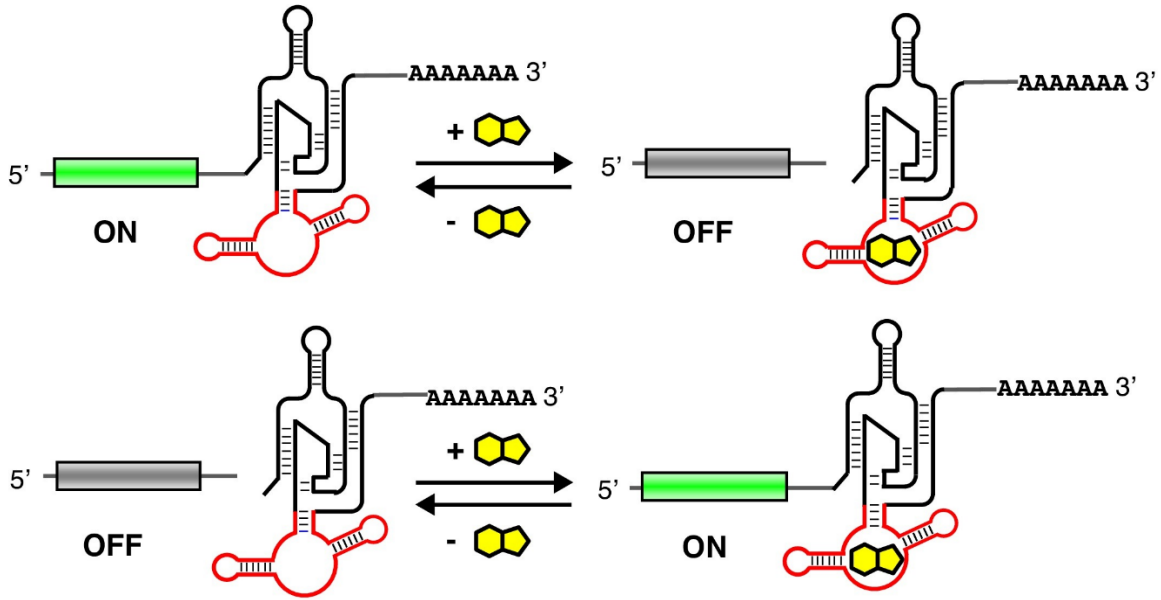


Figure 1.9 Activité d'un aptazyme

L'aptazyme composé par un aptamère (en rouge) lié à la tige d'un ribozyme (en noir) peut fonctionner de deux façons. En haut, l'aptazyme est actif en l'absence du ligand de l'aptamère. Par contre, lorsqu'il est présent, le changement conformationnel de l'ARN déstabilise le ribozyme en inhibant alors la réaction de clivage. En bas, l'aptazyme est activé par la liaison du ligand par l'aptamère, car la structure formée permet la stabilisation de la forme active du ribozyme. Tiré de (Yokobayashi, 2019a)

1.3.4.1. Domaine senseur-Aptamère

Avant que les aptamères fassent partie des *riboswitches*, le nom d'aptamère a été donné en premier aux oligonucléotides d'ADN ou d'ARN, « simple brin » synthétiques. Ils sont sélectionnés *in vitro* pour lier sélectivement et avec une grande affinité un ligand spécifique (Figure 1.10). La technique de sélection est nommée SELEX (*Systematic Evolution of Ligands by EXponential enrichment*) (Ellington & Szostak, 1990; Tuerk & Gold, 1990b), et est également utilisée pour les aptazymes. Le nom définit la fonction de la molécule, car il s'agit de la combinaison de deux mots latins : « aptus » signifiant « s'adapter » et « meros » qui veut dire « partie ». Les aptamères sont l'objet de beaucoup d'intérêt, étant donné la panoplie d'applications dans lesquelles ils peuvent être impliqués, et grâce à leurs habiletés en lien avec les différentes modifications chimiques. L'importance des aptamères vient du fait qu'ils représentent plusieurs avantages, comme le faible coût de synthèse, dû à leur petite taille, ainsi qu'à l'avancement technologique permettant la production des acides nucléiques dans un temps très court et avec grande facilité. Leur implication dans les *riboswitches* synthétiques témoigne d'une preuve de leur modularité. Plusieurs autres applications montrent la versatilité de l'utilisation des

aptamères. Ces derniers sont d'une importance majeure pour le développement des aptazymes, d'où la majorité des aptazymes publiés ont été construits en utilisant des aptamères existants. Ainsi, les applications et les modifications des aptamères sont aussi applicables sur les aptazymes.

- **Aptamères/aptazymes comme biosenseurs**

L'aisance de marquage des aptamères, à l'aide des fluorophores, par exemple, a permis leur utilisation dans des essais de développement d'outils de diagnostics. Ceci a été appliqué pour la détection des molécules impliquées dans des maladies comme le cancer par la liaison d'un biomarqueur (protéine, récepteur cellulaire...) ou d'une cellule complète, permettant leur distinction avec d'autres cellules (lignée cellulaire MCF7, cellules souches...) (Zhang *et al.*, 2019b; Zhou *et al.*, 2014; Zhuo *et al.*, 2017) (Odeh *et al.*, 2019). Plusieurs aptamères ont été sélectionnés et marqués pour permettre la détection d'une panoplie de bactéries pathogènes (*Escherichia coli* entérotoxigène (ETEC) K88, *Lactobacillus acidophilus*, *Staphylococcus aureus*, *Shigella sonnei*, *Salmonella tythimirium*, etc) (Chang *et al.*, 2021; Chen *et al.*, 2021b; Davydova *et al.*, 2016; Hamula *et al.*, 2011; Li *et al.*, 2021b; Wan *et al.*, 2021), de parasites (*Leishmania* spp., *Plasmodium* spp,...) (Birch *et al.*, 2015; Cui *et al.*, 2021; Hassan *et al.*, 2021; Lo *et al.*, 2021; Ospina-Villa *et al.*, 2018; Ospina-Villa *et al.*, 2016; Oteng *et al.*, 2020) et de virus (Hépatites B et C, VIH, SARS coronavirus) (Bellecave *et al.*, 2008; Han *et al.*, 2021a; Khanali *et al.*, 2021; Kruger *et al.*, 2021; Mandal *et al.*, 2021; Zou *et al.*, 2019). Grâce à leur capacité de conjugaison à d'autres molécules ou surfaces, les aptamères ont été impliqués dans le développement de biosenseurs détectant des polluants environnementaux (pesticides, herbicides) ou d'autres molécules néfastes à la santé humaine dans certains contextes tels que : antibiotiques, toxines et métaux lourds (Farzin *et al.*, 2017; McConnell *et al.*, 2020b; Nguyen *et al.*, 2017b; Palchetti & Mascini, 2008; Zhang *et al.*, 2018). Pour toutes les applications citées ci-dessus, les aptamères peuvent être considérés comme des biosenseurs s'ils sont conjugués à une molécule permettant d'avoir un signal visuel, comme un fluorophore, ou des nanoparticules d'or, par exemple (Modh *et al.*, 2018; Villalonga *et al.*, 2020). La détection se fait par plusieurs techniques parmi lesquelles on peut citer le FRET (*Forster Resonance Energy Transfer*) (Pehlivan *et al.*, 2019; Zheng *et al.*, 2020).

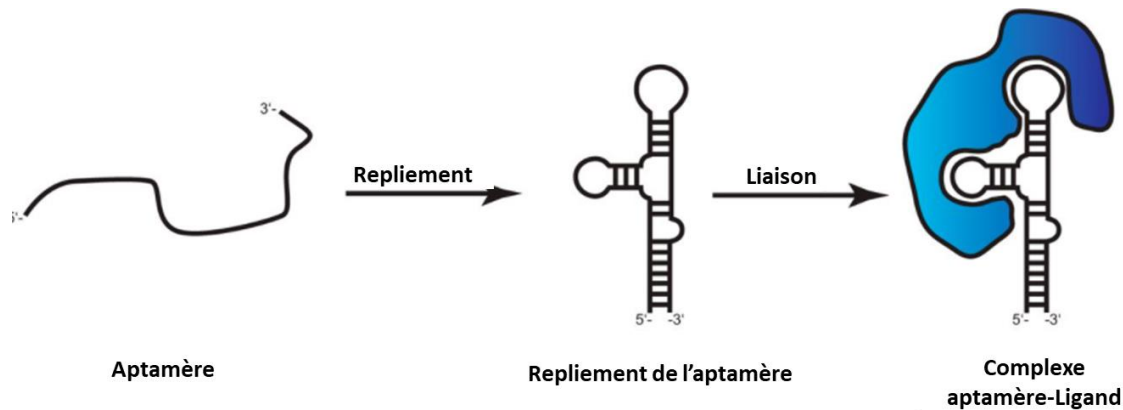


Figure 1.10 Changement structural de l'aptamère

L'aptamère se replie en une structure tertiaire bien définie. C, ce qui lui confère la capacité de lier son ligand spécifique. Adapté de (Wolter & Mayer, 2017)

- **Les aptamères/aptazymes dans les thérapies**

Des études ont montré l'efficacité de l'utilisation des aptamères dans plusieurs applications où des anticorps sont normalement utilisés, incluant : immunobuvardage, sélection de cellules exprimant un certain antigène, ELISA, etc. (Famulok *et al.*, 2007a). La liaison de l'aptamère à son ligand avec une affinité et spécificité très élevées lui confère la capacité de jouer un rôle d'inhibiteur. Un aptamère englobant son ligand ou liant une région d'une macromolécule résulte, par une inhibition totale ou partielle, de l'activité de la cible en bloquant son site actif ou en rentrant en compétition avec son récepteur. Une autre stratégie appliquée est l'utilisation des aptamères comme des véhicules de livraison d'un agent thérapeutique (Filippi *et al.*, 2020; Hori *et al.*, 2018; Zhu & Chen, 2018). Lors de la migration d'un aptamère, en cherchant sa cible, il peut être conjugué à une autre séquence d'acide nucléique (siRNA par exemple), d'agent chimiothérapeutique (cisplatine, doxorubicine), de toxines et de nanoparticules (Li *et al.*, 2013; Ray & White, 2010; Sivakumar *et al.*, 2019; Thiel & Giangrande, 2010; Zhang *et al.*, 2011). Un exemple d'application d'aptamère comme inhibiteur qui est utilisé actuellement comme thérapie est l'aptamère d'ARN liant le VEGF (*Vascular Endothelial Growth Factor*). Il inhibe la fixation de la VEGF à ces récepteurs. Il est utilisé comme thérapie pour une pathologie oculaire vasculaire et il s'agit du premier aptamère à être approuvé par la FDA (*Food and Drug Administration*) (Ng *et al.*, 2006).

- **Modifications des aptamères**

Afin d'améliorer l'efficacité des aptamères et d'augmenter la durée de leur demi-vie, plusieurs conjugaisons et modifications chimiques ont été appliquées (Adachi & Nakamura, 2019; Elskens *et al.*, 2020; Odeh *et al.*, 2019). De façon à protéger les acides nucléiques de la dégradation une fois dans les liquides physiologiques pleins de nucléases, les extrémités 3' et 5' sont protégées par leurs conjugaisons à des molécules comme la biotine, ou par le *capping* 3', avec la désoxy-thymidine inversée (Shum & Tanner, 2008). Les aptamères modifiés avec du polyéthylène glycol (PEG) ou du cholestérol en 5' montrent une demi-vie plus longue et une excrétion rénale moins rapide (Prodeus *et al.*, 2015) (Kratschmer & Levy, 2017; Odeh *et al.*, 2019). Puisque le 2'OH du ribose est le groupement attaqué par les ribonucléases, nous observons l'utilisation des nucléotides modifiés en position 2' (2'Fluoro, 2'O méthyl, 2' amino) (Green *et al.*, 1995; Levi-Acobas *et al.*, 2019; Padilla & Sousa, 1999; Wilds & Damha, 2000). Les analogues de nucléotides ayant un pont méthylène entre le 2'O et 4'C du sucre (Locked nucleic acids) (Schmidt *et al.*, 2004) constitue également une stratégie très répandue pour améliorer la résistance aux nucléases (Karlsen & Wengel, 2012; Veedu & Wengel, 2010), Kratschmer & Levy, 2017). Au-delà des modifications chimiques, les aptamères ont subi des changements structuraux, comme la circularisation d'aptamères bivalents (Jiang *et al.*, 2018) ou la formation de polymères par la liaison de plusieurs copies d'un aptamère (Vorobyeva *et al.*, 2016) (Bouhedda *et al.*, 2020).

1.3.4.2. **Domaine catalytique**

Différents types de ribozymes ont été utilisés pour développer des aptazymes. Les ribozymes en tête de marteau étaient les premiers à être expérimentés dans ce but (Rehm & Hartig, 2014; Soukup & Breaker, 1999b) (Zhong *et al.*, 2016) (Pu *et al.*, 2020). Les ribozymes en épingle à cheveux (Hall *et al.*, 2007; Najafi-Shoushtari & Famulok, 2005), les ribozymes HDV (Kertsburg & Soukup, 2002; Nomura *et al.*, 2013; Strobel *et al.*, 2020), les introns du groupe I (Thompson *et al.*, 2002) pistol (Kobori *et al.*, 2017), *twister* (Felletti *et al.*, 2016a; Mustafina *et al.*, 2020) et les ribozymes synthétiques (Helm *et al.*, 2005) ont été moins utilisés. Il y a deux façons de produire des aptazymes : par sélection ou par conception rationnelle.

1.3.4.3. Sélection *in vitro*

On peut sélectionner des aptazymes qui répondent à de nouvelles molécules en ayant recours à la technique SELEX. Le principe est d'insérer une séquence de nucléotides aléatoires dans une tige ou une partie qui a un rôle critique dans la stabilisation de la structure et de l'activité du ribozyme. Des cycles de sélection seront effectués afin d'obtenir des ribozymes actifs ou inactifs aux dépens de la présence du ligand choisi (Ferguson *et al.*, 2004; Koizumi *et al.*, 1999b; Robertson & Ellington, 1999; Robertson & Ellington, 2001; Robertson *et al.*, 2004). Cette technique a été moins utilisée que celle de l'insertion d'un aptamère connu dans la tige d'un ribozyme par conception rationnelle.

1.3.4.4. Conception rationnelle

L'autre façon d'obtenir des aptazymes est d'utiliser un aptamère connu pour l'insérer dans la séquence du ribozyme. Dans ce cas aussi, il y a plusieurs critères à considérer pour trouver la position la plus convenable pour insérer l'aptamère et les meilleures modifications de séquences à inclure (Ogawa, 2011; Ogawa, 2014; Pu *et al.*, 2020; Stifel *et al.*, 2019a; Tang & Breaker, 1997; Zhong *et al.*, 2016b).

Un aptazyme peut être structuré de deux façons différentes :

- L'aptamère est inséré directement dans la séquence du ribozyme (Tang & Breaker, 1997);
- L'aptamère est lié au ribozyme par une séquence qui joue le rôle d'un « *linker* », ou d'un pont. Cette séquence permettra la modulation de la structure du ribozyme selon la présence ou l'absence du ligand. Dans ce cas, le *linker* joue le rôle de transmetteur d'information entre l'aptamère et le ribozyme. La modulation de la structure de l'aptamère affectera celle du *linker*, ce qui entraînera alors un changement de conformation du ribozyme, qui sera soit activé ou inactivé, selon la nouvelle réorganisation (Mustafina *et al.*, 2020; Soukup & Breaker, 1999b; Stifel *et al.*, 2019a; Wurmthaler *et al.*, 2019).

1.3.4.5. Application des aptazymes

Les aptazymes ont montré leur efficacité dans plusieurs études *in vitro* et *in vivo* pour le contrôle de l'expression génique dans des contextes différents, tels que le contrôle de la réplication des virus utilisés comme vecteurs pour livrer des thérapies géniques en formant un « *safety switch* » (Ketzer *et al.*, 2014). Chez les bactéries, des aptazymes ont

été insérés dans les régions 5' UTR et 3' UTR du gène à contrôler. Ce principe a été appliqué à des cellules humaines, et même chez des souris (Felletti & Hartig, 2017). Les ribozymes allostériques étaient aussi utilisés comme biosenseurs (Du *et al.*, 2021; Penchovsky, 2013; Rueda & Walter, 2006; Yang *et al.*, 2017). Bien que les aptazymes montraient une efficacité assez élevée dans les essais *in vitro*, ils rencontraient toutefois des obstacles, comme la courte demi-vie dans les liquides physiologiques et les milieux intracellulaires. Pour cela, les DNAzyme, qui sont des enzymes d'ADN synthétique capables de couper l'ADN et l'ARN, ont été utilisés dans une plus grande majorité d'applications et de développements de biosenseurs basés sur des aptazymes. Les DNAzyme présentent une meilleure alternative étant donné qu'ils ont une meilleure stabilité et un coût de synthèse plus faible, en comparaison avec l'ARN (Walter & Stahl, 2020). Malgré cela, les DNAzymes sont limités à des applications *in vitro* et des milieux extracellulaires, puisqu'ils ne peuvent pas être exprimés directement dans les cellules.

1.3.5. Ribozymes allostériques liant des oligonucléotides

Après avoir établi la preuve montrant que la séquence d'un ribozyme peut être modifiée pour l'ajout d'aptamères, l'idée d'utiliser des oligonucléotides pour le même but a été publiée pour la première fois par Penchovsky et Breaker (Penchovsky & Breaker, 2005). L'étude avait démontré qu'à l'aide d'une approche automatisée, en prenant en considération la structure conservée du ribozyme, une région avec une séquence aléatoire ou définie pour être complémentaire à un oligonucléotide ayant une séquence donnée est rajoutée à la tige II, par exemple (**Figure 1.11**). Cette région doit lier un oligonucléotide donné, mais d'une façon à perturber ou stabiliser la structure active du ribozyme.

Pour ce genre de conception, la génération de séquences de ribozymes passe par deux étapes représentant les deux états que le ribozyme doit pouvoir adopter.

- L'état inactif en absence de l'oligonucléotide : les séquences générées doivent former une structure désorganisée et inactive, mais stable dans cet état;
- L'état actif en présence de l'oligonucléotide : les séquences qui passent la première étape de sélection sont filtrées pour conserver seulement celles qui forment une structure répondant aux critères d'un ribozyme actif, stable et capable de lier « l'entrée » (« *input* »).

Dans le cas contraire, où la présence de l'oligonucléotide doit inhiber l'activité du ribozyme, le même principe est conservé, mais en inversant les conditions de sélection.

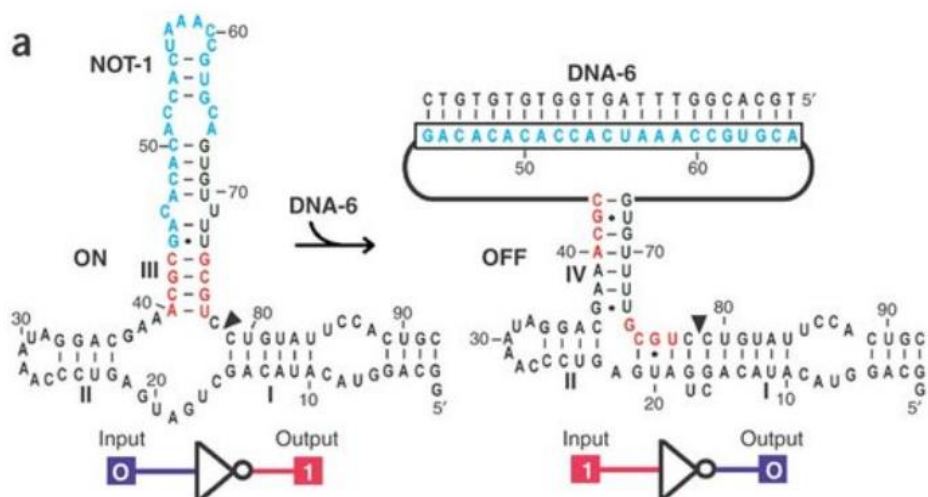


Figure 1.11 Ribozyme allostérique inhibé par la liaison d'un oligonucléotide

La séquence (en bleu) complémentaire à l'oligonucléotide est insérée dans la tige III. A : en absence de l'oligonucléotide, la séquence bleue forme une tige en boucle par interaction avec la séquence de la tige III. Ceci permet la formation de la structure active du ribozyme capable d'effectuer la réaction de clivage. B : en présence de l'oligonucléotide, il interagit avec la séquence bleue, ce qui déstabilise la structure du ribozyme qui devient alors inactif. Tiré de (Penchovsky & Breaker, 2005).

Cette stratégie a permis d'utiliser les ribozymes pour construire des fonctions et des circuits logiques en réponse à une seule ou plusieurs « entrées » (*inputs*). En effet, Penchovsky (Penchovsky, 2014) a construit différents types de fonctions logiques (YES, AND, OR, NOT) par la détection d'un ou de plusieurs oligonucléotides. De plus, des ribozymes capables de distinguer des oligonucléotides avec des tailles variables ont été sélectionnés (Penchovsky, 2012).

1.4. Outils de conception de ribozymes

Grâce à leur activité catalytique et à l'universalité de leurs applications chez les procaryotes et eucaryotes, durant presque 40 ans d'études, les ribozymes ont été utilisés dans différentes applications. Par ailleurs, les ribozymes naturels sont actifs en *cis*, mais la séquence de la majorité des petits ribozymes auto-clivants est modifiable en une version *trans*. En effet, un ribozyme doit répondre à deux conditions pour être actif ; posséder la séquence conservée, qui inclut généralement le cœur et le site catalytique, et former la structure conservée. Ainsi, la génération ou la conception de nouvelles

séquences synthétiques de ribozymes peut être faite manuellement en respectant ces deux règles.

En effet, la famille des petits ribozymes auto-clivants ne contient que le ribozyme Varkud Satellite ayant une activité en *trans*, mais ils sont presque tous facilement modifiables en version *trans*. Pour ce faire, il suffit de diviser l'ARN en deux régions : région ribozyme et région substrat. Le ribozyme inclut le cœur catalytique, une ou plusieurs tiges complètes, dépendamment du type du ribozyme, et une ou deux régions simple brin. Dans la majorité des cas, le substrat est une séquence simple brin contenant le site de clivage flanqué de part et d'autre par des séquences complémentaires aux deux régions simple brin du ribozyme. L'hybridation entre les deux parties permettra la formation de la structure fonctionnelle du ribozyme, activant ainsi le clivage.

1.4.1. Sélection de nouveaux ribozymes *in vitro*

Les découvertes des ribozymes auto-clivants ont commencé en 1983 et se sont poursuivies encore aujourd'hui. De plus, dans le but de rendre les ribozymes malléables aux applications voulues, la sélection de nouveaux ribozymes est indispensable. Cette sélection a été faite en premier *in vitro*, soit de façon à générer des ribozymes plus efficaces que ceux existant naturellement, ou pour obtenir des ribozymes avec de nouvelles fonctions. Ceci a commencé avec la découverte de certains ribozymes qui peuvent être divisés en deux brins : enzyme et substrat. Ce fut à la fin des années quatre-vingt (Forster & Symons, 1987b; Hampel & Tritz, 1989; Uhlenbeck, 1987), ce qui avait abouti à la conception des premiers ribozymes actifs en *trans*. Les premiers designs étaient faits manuellement pour la génération de ribozymes ayant des mutations ponctuelles (Haseloff & Gerlach, 1988; Jeffries & Symons, 1989; Lai *et al.*, 1996; Puttaraju *et al.*, 1993) (Koizumi *et al.*, 1989; Scherr *et al.*, 1997), des ribozymes HDV (Prasad *et al.*, 1992), des ribozymes antiviraux (Inokuchi *et al.*, 1994; Rossi *et al.*, 1990; Rossi & Sarver, 1990; Russell, 1990; Sarver *et al.*, 1990; Weinberg *et al.*, 2007) et d'autres contre des ARNm (Cameron & Jennings, 1989). Les ribozymes issus de ces designs s'étaient avérés actifs, mais une méthode pour la génération de plus de candidats donnant une probabilité plus élevée d'avoir de meilleurs candidats était la sélection *in vitro*.

1.4.1.1. Sélection de ribozymes actifs en trans

La sélection de nouveaux ribozymes *in vitro* a commencé par celle de ribozymes en *trans*. Cette sélection consiste en une itération de cycles de manière à obtenir une séquence ou une activité désirée. Au début, par exemple la séquence d'un ribozyme hairpin, ayant prouvé son activité contre l'ARN de HIV, est altérée par des mutations dans la région complémentaire à la cible. Cette altération a permis de sélectionner des séquences de ribozymes ayant une meilleure activité de clivage à la suite de quelques cycles de criblage (Joseph & Burke, 1993). Ensuite, un criblage pour des régions accessibles sur des cibles, comme des ARNm, était fait par l'utilisation d'une librairie de séquences aléatoires ayant seulement deux nucléotides complémentaires au site de clivage comme séquence fixe. Dans l'exemple de recherche des sites potentiels à cibler par des ribozymes en tête de marteau, deux nucléotides AC ont été fixés (**Figure 1.12**). Le processus de sélection se fait par l'incubation des ARN radio-marqués de la librairie avec l'ARNm cible, suivie d'une migration sur un gel natif. Les ARN capables de lier la cible migrent plus lentement que ceux libres, qui sont, à la fin, à un niveau plus bas sur le gel. Ces processus sont répétés au moins cinq ou six fois, après lesquels les séquences sont clonées et séquencées. Enfin, la décision du choix des sites se fait après vérification de la structure secondaire de l'ARNm afin de désigner les régions ayant plus d'accessibilité. Les séquences sélectionnées représentent alors les brins de ribozymes actifs en *trans* capables de s'hybrider sur l'ARNm cible (Lieber & Strauss, 1995; Norris *et al.*, 2000). L'autre moyen de sélection est de construire une librairie de ribozymes avec le même principe, mais plutôt que deux nucléotides, seulement la tige II ainsi que le cœur catalytique sont conservés. Dans ce cas, la sélection se fait en se basant sur l'activité en premier. Les régions coupées sont récupérées, les ADN copies sont générés, amplifiés et séquencés ; quant aux, séquences obtenues, elles permettront de déterminer celles de ribozymes actifs (Campbell & Cech, 1995; Lieber *et al.*, 1997; Lieber & Strauss, 1995; Nakamaye & Eckstein, 1994). Une autre stratégie de sélection de ribozymes actifs en *trans* est de fixer l'ARN cible sur une colonne et faire passer les ribozymes de la librairie à l'intérieur. Ceci permet tout d'abord de garder seulement les ribozymes qui lient la cible par complémentarité. À la suite d'une incubation dans un tampon de clivage, seuls les ribozymes actifs tomberont une fois qu'ils auront coupé l'ARN cible. Le processus est

répété plusieurs fois jusqu'à l'obtention d'un ribozyme ayant une activité supérieure à celle du WT (Nishikawa *et al.*, 1995; Nishikawa *et al.*, 1996).

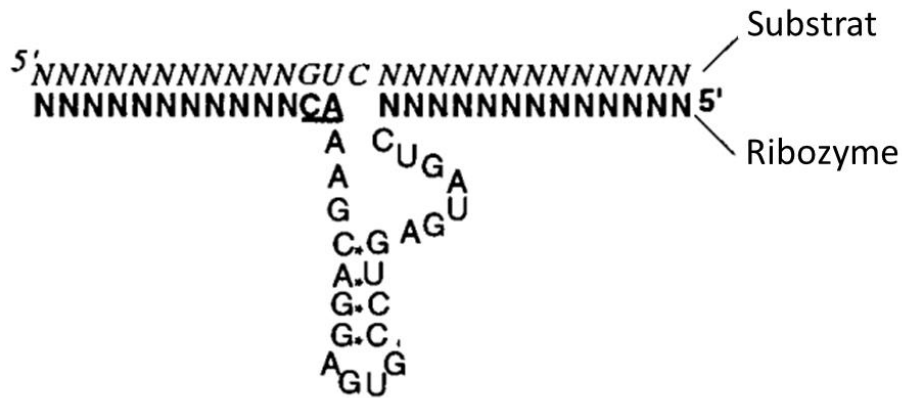


Figure 1.12 Librairie de sélection d'un ribozyme actif en *trans*

Les tiges I et III du ribozyme ont des séquences aléatoires (présentées avec des N) sauf deux nucléotides CA permettant l'hybridation au site GUC qui est le site de clivage. Adapté de (Lieber & Strauss, 1995)

1.4.1.2. Sélection de nouveaux ribozymes

En parallèle à la sélection de ribozymes avec des activités connues, pendant la même période, plusieurs laboratoires ont travaillé sur la sélection de ribozymes avec de nouvelles fonctions. Ceci était encouragé par les découvertes de différentes activités catalytiques de ribozymes semblables aux protéines, ce qui a conduit à la volonté de sélection de nouveaux ARN non existants en nature avec d'autres activités enzymatiques.

Le principe de sélection était toujours le même, soit celui de la répétition de cycles de sélection par l'incubation d'une librairie d'ARN avec des séquences aléatoires, avec une cible donnée afin d'obtenir une activité désirée. Cette technique avait permis la sélection de ribozymes ayant une activité peptidyl transférase, semblable à celle des ribosomes, sans l'aide de protéine (Zhang & Cech, 1997), d'autres à l'activité de polynucléotides kinase (Lorsch & Szostak, 1994), de ligase (Bartel & Szostak, 1993) et de plusieurs autres réactions enzymatiques (Beaudry *et al.*, 2000; Gold *et al.*, 1995; Pan, 1997; Tuschl *et al.*, 1998). Toutes ces sélections ont été véritablement produites en utilisant le principe d'une technique, développée en 1990, pour la sélection de séquences d'acides nucléiques fonctionnelles liant un ligand spécifique, soit « les aptamères ». Cette technique a été

nommée SELEX. Ensuite, elle a été adaptée pour la sélection de ribozymes et d'aptazymes.

1.4.2. SELEX

1.4.2.1. Principe

La technique SELEX a été développée en 1990 (Ellington & Szostak, 1990; Tuerk & Gold, 1990b). Cette technique consiste en l'utilisation d'une librairie de séquences aléatoires (typiquement $\sim 10^{15}$ séquences différentes) d'une taille de 20 à 100 nucléotides flanqués par deux séquences constantes, et ce afin d'effectuer des cycles de sélection. La librairie d'acides nucléiques est incubée en présence du ligand d'intérêt. Les ADN/ARN liant la molécule cible sont sélectionnés et amplifiés par PCR afin de passer au prochain cycle, et celles non liées sont éliminées (**Figure 1.13**). Avec l'avancement des tours de sélection, les conditions deviennent plus strictes (plus de lavages, diminution du temps d'incubation et de la concentration du ligand, etc.) afin d'augmenter la stringence de la liaison avec le ligand (Darmostuk *et al.*, 2015). La sélection est terminée lorsqu'un enrichissement de séquences est atteint. Les séquences enrichies sont caractérisées afin de choisir « l'aptamère final ». Ceci se fait en déterminant les constantes de dissociation (K_D). Le K_D le plus faible appartient généralement à la séquence ayant la meilleure affinité au ligand.

Dans le cas de ribozymes, la sélection est achevée une fois que l'activité catalytique voulue est atteinte. Dans le cas d'aptazymes, les deux critères d'affinité au ligand et de la capacité d'effectuer la réaction catalytique sont considérés pour mettre fin au SELEX (Piganeau, 2009; Piganeau *et al.*, 2001; Robertson & Ellington, 1999; Roth & Breaker, 2004).

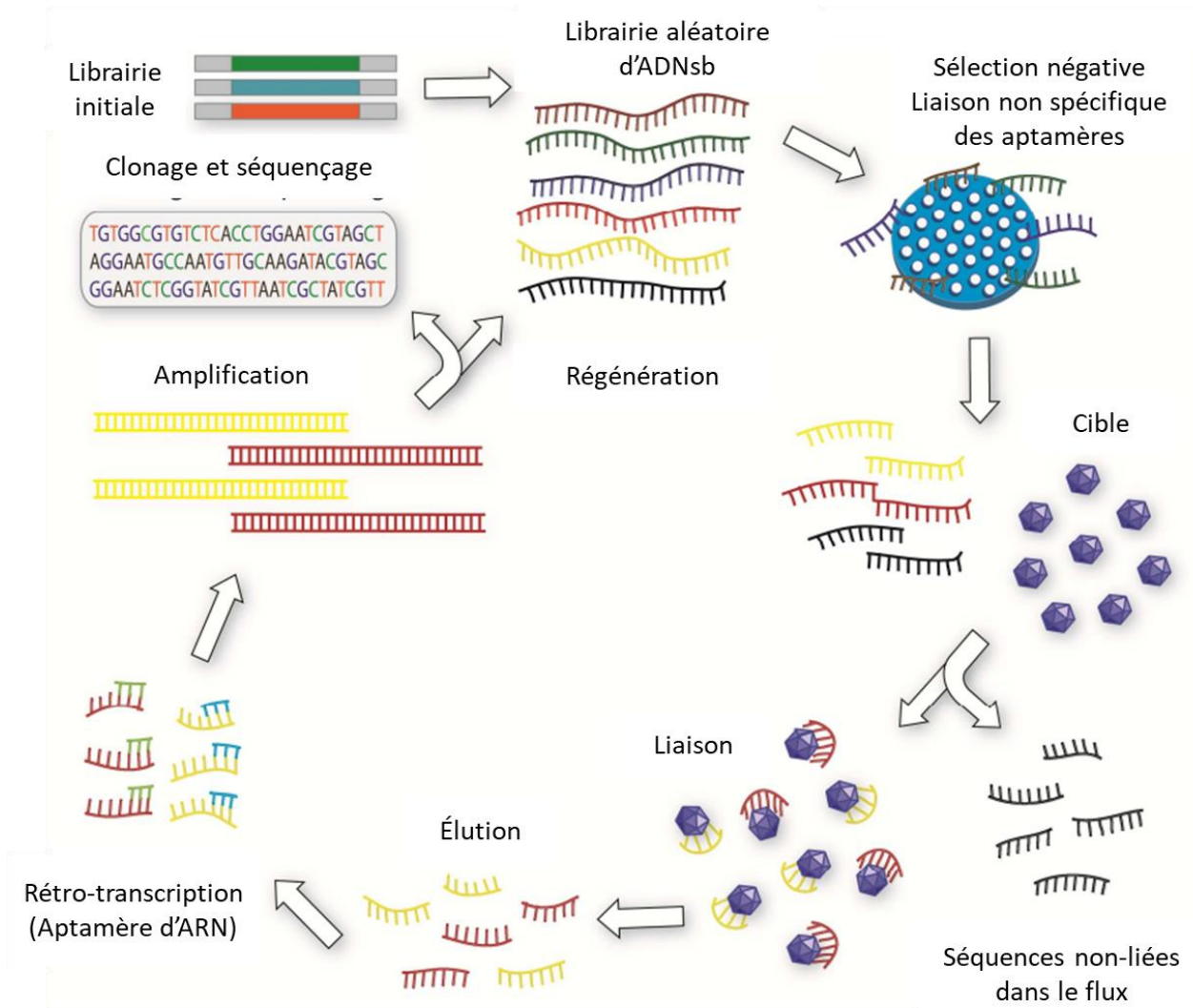


Figure 1.13 Protocole de SELEX

Le ligand fonctionnalisé sur des billes magnétiques est incubé avec la librairie d'ADNs (simple brin) ou ARN ayant des séquences aléatoires. La séparation se fait par un champ magnétique permettant de lier les billes magnétiques ainsi que les ARN liés, les oligonucléotides non liés passent dans le flux. Les ARN liés sont ensuite élués et rétro-transcrits ce qui permet la génération de l'ADNc qui est amplifié par PCR permettant le passage au prochain cycle. Adapté de (Darmostuk *et al.*, 2015)

1.4.2.1. Ligands et switch de structure

Le criblage de nouveaux ribozymes a commencé au début des années 90 et continue jusqu'à aujourd'hui (Arriola & Muller, 2020; Berzal-Herranz *et al.*, 1992; Chapman & Szostak, 1994; Chen *et al.*, 2009; Fukuda *et al.*, 2014; Ghaem Maghami *et al.*, 2019; Huang *et al.*, 2019; Olson & Muller, 2012; Piganeau, 2009; Piganeau *et al.*, 2001; Robertson & Ellington, 1999; Roth & Breaker, 2004) (Dagenais & Legault, 2021; Magde

et al., 2021; Walton *et al.*, 2020). Le principe d'itération est le même, mais les méthodes de sélection ont évolué au cours du temps. Ceci dépend de la fonction ainsi que du ligand à lier. En effet, le ligand d'un aptamère ou d'un aptazyme peut être une petite molécule (acides aminés, métaux, vitamines et métabolites), une macromolécule (protéines, polysaccharides, etc.) ou des cellules complètes (bactéries, cellules cancéreuses, etc.) en ciblant une molécule spécifique à la surface cellulaire (Farjami *et al.*, 2013; Krasitskaya *et al.*, 2019; Labib *et al.*, 2012; Lee *et al.*, 2013; Porchetta *et al.*, 2012). Dans le cas des cellules, la cible réelle est vraisemblablement une macromolécule (Zhou *et al.*, 2012). Ainsi, le SELEX ne représente pas une seule méthode bien définie, mais plutôt un principe général qui est suivi pour obtenir un oligonucléotide d'une façon différente dépendamment du ligand, du type et de la fonction de l'ARN à sélectionner. De plus, avec l'avancement technologique, le processus de sélection a subi plusieurs modifications et de perfectionnement. De nouvelles approches ont été développées, comme « *the one step SELEX* » (Arnold *et al.*, 2012; Lauridsen *et al.*, 2012), *Capillary Electrophoresis SELEX* (CE-SELEX) (Hamedani & Muller, 2016), *Capture SELEX* (Boussebayle *et al.*, 2019), *in vivo SELEX* (Cheng *et al.*, 2013) } (Chen *et al.*, 2009; Huang *et al.*, 2019; Olson & Muller, 2012), *RAPID* (*aptamer isolation via Dual cycles*) et SELEX, utilisant une plateforme nommée MEDUSA (*Microplate based Enrichment Device Used for the Selection of Aptamers*) (Chau *et al.*, 2020).

La capacité de liaison spécifique des molécules est conférée aux aptamères/aptazymes par le repliement en une structure tertiaire stable et bien définie pour chaque aptamère (**Figure 1.10**) (Tuerk & Gold, 1990b) (Famulok *et al.*, 2007a). Cette structure spécifique ne se forme souvent qu'en présence du ligand, de sorte qu'un changement conformationnel se produit entre l'état libre et l'état lié (Figure 1.13). De plus, la liaison entre l'aptamère et son ligand peut être accomplie par des liaisons hydrogène, des liaisons Van der Waals et des interactions électrostatiques (Gulbakan *et al.*, 2018; Pendergrast *et al.*, 2005).

1.4.2.2. Analyse des résultats de SELEX

L'enrichissement des séquences est détecté par le séquençage de l'ADN issu de différentes générations de SELEX. La méthode conventionnelle est le clonage de l'ADN de la dernière génération de SELEX et de quelques dizaines de plasmides extraits de colonies résultant de ce clonage ; et ce dernier est envoyé au séquençage Sanger. Ceci

donne une idée générale des exemplaires de séquences enrichies choisies aléatoirement, ce qui n'est jamais assez pour identifier tous les candidats et ainsi trouver le meilleur aptamère. Une nouvelle façon qui permet d'avoir une vue exhaustive des séquences et de l'évolution de leur enrichissement est le séquençage à haut débit avec les différentes plateformes disponibles (Illumina, 454 de Roche et IonTorrent). L'ADN amplifié de toutes les générations issues de chaque cycle de SELEX, en commençant par la librairie initiale, peut être séquencé. Ceci permet d'analyser de nombreuses séquences et de choisir les candidats ayant le plus grand nombre de représentants, c'est-à-dire « lectures » ou « reads » (Komarova *et al.*, 2020; Yokobayashi, 2020). Bien que le séquençage haut débit ait l'avantage de générer beaucoup de données (des milliers, voire des millions de séquences), l'analyse de résultats de séquençage n'est pas une tâche facile à faire, manuellement. Des outils bioinformatiques facilitant ce processus ont surtout été développés pour les séquences d'aptamères, ce qui peut être aussi utilisé pour les aptazymes. Citons comme exemple Galaxy (Thiel & Giangrande, 2016), AptaSuite (Hoinka *et al.*, 2018), AptaPLEX (Hoinka & Przytycka, 2016) et FASTAptamer (Alam *et al.*, 2015). Ces programmes autorisent généralement la formation des « Cluster » de séquences uniques, mais similaires selon leurs abondances ou selon la structure primaire qu'ils forment. Il y a des outils comme APTANI2 (Caroli *et al.*, 2020) qui offrent même des prédictions des structures secondaires et des motifs de liaison du ligand puis ils comparent les séquences en fonction de cette caractéristique.

1.4.3. Limites de sélection *in vitro*

Bien qu'il y ait beaucoup d'avancement dans les méthodes de sélection *in vitro* et une apparition de plusieurs essais à haut débit qui ont pour but d'accélérer le processus, le progrès reste limité. Ceci est dû au grand nombre de paramètres à prendre en considération lors de la conception de ribozymes et de librairies de sélection. À ces paramètres s'ajoutent de nouvelles variables, comme la révélation des motifs structuraux indispensables à l'activité de plusieurs ribozymes, qui ont été ignorés plusieurs années après leur découverte. Le dévoilement des motifs, comme les pseudonœuds n'a pas changé uniquement les données de structures, mais également les facteurs environnementaux permettant l'activité de ribozymes. De ce fait, en variant la séquence, les conceptions manuelles peuvent résulter par des possibilités très vastes de structures secondaires non actives, ce qui complique la tâche. Par conséquent, des méthodes de conception permettant d'accélérer l'obtention de résultats ainsi que l'optimisation des

designs étaient une nécessité. En vue de ce besoin, des outils bioinformatiques permettant la conception automatique de nouvelles séquences de différents types de ribozymes ont été développés. Ce développement a touché à la fois les ribozymes actifs en *cis* et en *trans*, avec des programmes spécialisés pour ces derniers ou pour des ARN en général.

1.4.4. Conception informatique des ARN

Les essais de développement de programmes informatiques pour l'ARN ont débuté par des outils de prédiction de structure secondaire, quelques années avant la découverte de ribozymes (Pipas & McMahon, 1975; Studnicka *et al.*, 1978; Tinoco *et al.*, 1973; Tinoco *et al.*, 1971). Néanmoins, pendant ce temps, la structure d'ARN n'était pas encore très bien comprise, mais les petits ARN catalytiques ont dévoilé plusieurs secrets par rapport à ces structures. Avec les données expérimentales fournies par résolution de plusieurs structures de ribozymes et des motifs de structures secondaires, la relation entre la structure globale et la fonction d'un ARN était déterminée (Dethoff *et al.*, 2012; Jaeger *et al.*, 1993; Lu & Chang, 2016; Mortimer *et al.*, 2014; Shah & Brunger, 1999; Wan *et al.*, 2011). En réalité, l'ARN est formé par des sous-unités structurales au moyen de divers types de paires de bases. En effet, le repliement des ARN sur eux-mêmes résulte principalement de la formation des motifs des tiges avec boucles terminales et/ou internes, des pseudonœuds, des *kissing loop*, des jonctions, etc. Ces motifs, ainsi que l'interaction entre eux formant la structure tridimensionnelle, sont sous l'emprise des lois de la thermodynamique. Ceci est réalisé en déterminant des paramètres comme l'entropie « ΔS », l'énergie libre incluant « *the melting temperature* (T_m) » et « *the Gibbs free energy* » ou l'enthalpie libre « ΔG ». L'importance de ces paramètres vient du fait que leurs valeurs définissent la stabilité d'une structure formée (Turner, 1996). Les valeurs dépendent des paires de bases (canoniques ou non canoniques) formées, mais aussi des nucléotides proches, de leurs positions et interactions ainsi que des conditions de l'environnement de l'ARN (Broda *et al.*, 2005; Jacek Nowakowski and Ignacio Tinoco, 1997; Turner, 1996; Wu *et al.*, 2009).

C'est en se basant sur ces paramètres et facteurs qu'un grand nombre d'outils de prédiction et de conception de nouvelles séquences d'ARN naturelles et synthétique synthétiques se sont développés. En fait, la prédiction de structures de nouvelles séquences est basée sur les données expérimentales existantes constituant les valeurs des paramètres de stabilités thermodynamiques. Le but de ce genre de prédiction est la

minimisation de la valeur de l'énergie libre ('*minimum free energy*' (MFE)) correspondant à la structure obtenue. Par le fait que des données existantes fournissent les MFE de différentes paires de bases et motifs, dépendamment de la nouvelle séquence entrée, plusieurs versions de structures sont générées et évaluées par le calcul de paramètres thermodynamiques. La structure choisie est celle présentant la valeur de MFE la plus faible (Doshi *et al.*, 2004; Mathews, 2004; Mathews & Turner, 2006; Sato *et al.*, 2021; Wu *et al.*, 2009). Ce type de prédiction peut être aussi combiné avec une approche d'analyse comparative basée sur la comparaison de séquences similaires, en supposant qu'elles forment des structures similaires (Gardner & Giegerich, 2004; Reuter & Mathews, 2010; Seetin & Mathews, 2012).

Dans le cas d'une conception de nouvelles séquences, le principe ne diffère pas beaucoup. La différence est qu'à défaut d'entrer une séquence, c'est la structure qui est prédéfinie. Dans ce cas également, ce sont les mêmes paramètres thermodynamiques qui déterminent la séquence à choisir. Ceci est dû au fait que l'évaluation de chaque nouvelle séquence s'effectue par une prédiction de structure qui passe par une étape de comparaison à la structure désignée au début (Aguirre-Hernandez *et al.*, 2007; Dirks *et al.*, 2004; Dotu *et al.*, 2014; Rodrigo & Jaramillo, 2014).

Les outils de conception développés sont nombreux, mais une majorité a utilisé des algorithmes basés sur la MFE ou la fonction de partition. En effet, cette dernière, pour le repliement d'ARN, constitue la probabilité de différentes possibilités de repliement d'un ARN en une certaine structure à une température définie. Ceci est exécuté par la considération non seulement d'une seule structure optimale ayant la MFE la plus faible, mais aussi de toutes les structures sous-optimales formées par chaque possibilité de formation de paires de bases. La fonction de partition fait la somme des prédictions de sous-structures afin de donner comme résultat la prédiction d'une seule structure optimisée (Bonhoeffer *et al.*, 1993; Huang *et al.*, 2009; Lorenz & Clote, 2011; Mathews, 2004; McCaskill, 1990; Zhang *et al.*, 2020b).

Les méthodes computationnelles diffèrent et évoluent jusqu'à nos jours, mais le but principal reste le même, soit d'obtenir des prédictions efficaces qui remplaceront les techniques expérimentales ardues. C'est dans ce but que ces outils de prédiction étaient et sont toujours utilisés dans le cadre de générations de nouvelles séquences de ribozymes.

1.4.5. Outils de conception de ribozymes actifs en *trans*

Pendant les années de sélection de ribozymes actifs en *trans in vitro*, seulement des outils de prédiction comme mfold (Zuker, 2003b) étaient employés. Ceci se produisait dans le cadre de vérifications de structures secondaires des ARN cibles de ribozymes, afin de vérifier l'accessibilité des régions complémentaires aux brins du ribozyme (Campbell & Cech, 1995; Lieber *et al.*, 1997; Lieber & Strauss, 1995; Nakamaye & Eckstein, 1994; Norris *et al.*, 2000).

C'est seulement après plus d'une dizaine d'années qu'un premier outil est venu remplacer l'étape de recherche *in vitro* des cibles potentielles de ribozymes en tête de marteau et SOFA-HDV (qui est une version modifiée du ribozyme HDV). « Ribosubstrate » permet d'effectuer la recherche des cibles possibles de ribozymes ayant déjà leurs séquences définies (Lucier *et al.*, 2006a). La recherche s'appuie sur une analyse par l'outil informatique Blast avec des bases de données d'ADNc (ADN complémentaire) de plusieurs génomes intégrés dans le système. Ce Blast permet de prendre en considération des séquences très courtes (quelques nucléotides), des paires de bases non canoniques « *wobble* » (G-U) et des mésappariements. Pendant ce temps, ces critères assemblés n'étaient présents dans aucun autre outil de Blast existant. Le but principal de « Ribosubstrate » était de fournir les séquences cibles, tout en assurant leur spécificité à un seul substrat pour éviter l'effet de « *OFF targets* », un problème majeur des ribozymes et d'autres acides nucléiques fonctionnels anti-sens (Lucier *et al.*, 2006a).

Quelques années plus tard, dans le but de limiter la charge du travail expérimental pour la sélection de ribozymes actifs, le premier outil de conception automatisé de ribozymes a été développé, en 2012. Le programme nommé ALADDIN (*SeArch computing tool for hAmmerhead ribozyme DesigN*) (Mercatanti *et al.*, 2012) était désigné pour le ribozyme en tête de marteau minimal. Il prenait en considération à la fois la recherche du site de clivage accessible sur la cible et la conception des brins de ribozymes complémentaires à cette dernière. Ce qui manquait dans ce programme était la vérification de la spécificité du ribozyme généré, donc de l'activité hors cible. De plus, à l'exception de la séquence de l'ARN à dégrader et sa structure secondaire, le logiciel ALADDIN ne prenait en considération aucun paramètre en relation avec l'environnement de la réaction.

Parallèlement au développement de ces programmes, les avancements technologiques ont permis de mieux comprendre les motifs structuraux de l'ARN en

étudiant les ribozymes, surtout celui en tête de marteau. Avec ses avancements les logiciels existants avant 2015 n'étaient plus suffisants pour répondre au besoin de conception de ribozyme d'une façon efficace car ces outils n'étaient pas complètement automatisés et en général, aucun paramètre ne pouvait être ajusté. De plus, les pseudonœuds se sont avérés essentiels dans l'activité de plusieurs ARN et donc il est important de les inclure dans les outils de conception.

Parmi les motifs structuraux, les pseudonœuds représentent un exemple d'interaction tertiaire qui se passe entre deux ou plusieurs tiges de ribozymes comme celui en tête de marteau, HDV, pistol, et GImS. Ceci a engendré une amélioration des algorithmes de conception et de prédiction de nouveaux ARN. Le pseudonœud n'était pris en considération pour des ribozymes actifs en *trans* que dans le cadre du service web « Ribosoft1.0 », développé en 2015. Ce programme a tenté de couvrir, en plus, des paramètres déjà utilisés par les outils publiés avant, ceux qui étaient ignorés par ceux-là, comme l'accessibilité du site de clivage du ribozyme présent sur la cible et ce dépendamment de la structure secondaire de l'ARN à dégrader (Kharna *et al.*, 2016). « Ribosoft1.0 » permet de concevoir les ribozymes en tête de marteau ayant la structure minimale ou complète. Cet outil inclut déjà la structure secondaire de chaque version du ribozyme avec les nucléotides conservés. Il génère des séquences qui doivent se replier dans la structure donnée en répondant à des critères choisis par l'utilisateur. Pendant le processus de conception, les séquences générées passent par des étapes de filtration selon des paramètres calculés par Ribosoft comme la qualité de la structure des ribozymes conçus, la température d'hybridation entre les brins du ribozyme et la cible, ainsi que la spécificité des ribozymes destinés à l'utilisation *in vivo*.

Finalement, malgré les avantages que Ribosoft présentaient par rapport aux logiciels existants, son inconvénient était le choix de ribozymes à concevoir, limité à deux types de ribozymes en tête de marteau. Ainsi, pour mieux profiter de la diversité des fonctions et structures des ribozymes et aptazymes connus, un serveur permettant la conception de n'importe quel ribozyme connus serait souhaitable.

1.4.6. Outils de conception de ribozymes actifs en *cis*

La conception automatisée des ribozymes auto-clivants a commencé avant celle des ribozymes en *trans* par le laboratoire de Breaker, dans le but de générer des ribozymes allostériques répondant à des oligonucléotides (Penchovsky, 2012; Penchovsky, 2014;

Penchovsky & Breaker, 2005; Penchovsky & Kostova, 2013). Ce type de conception a utilisé des algorithmes basés sur la fonction de partition, afin de permettre l'obtention de toutes les possibilités de structures obtenues par l'hybridation du ribozyme avec l'oligonucléotide menant à la meilleure structure finale (Penchovsky, 2014). D'autres types de ribozymes étaient conçus en ayant recours à des programmes existants sans développement de nouvel algorithme comme des ribozymes « *Hairpin* » générés par l'utilisation de ViennaRNA package (Petkovic *et al.*, 2015) ainsi que des ribozymes RNase P allostériques (Ender *et al.*, 2021; Petkovic *et al.*, 2015). Toutes ces approches ne nécessitaient pas la modification d'une grande partie des séquences du ribozyme, mais plutôt d'une seule région. Celle modifiée devait répondre au seul critère fixé par le développeur, soit stabiliser ou déstabiliser la structure active du ribozyme selon le but de l'expérience. Pour des conceptions plus avancées, où la plus grande majorité de la séquence devrait changer, et des motifs structuraux plus complexes, qui sont différents de structure simple de tige-boucle, une approche nommée repliement inverse (« *inverse folding* ») est utilisée (Zhong *et al.*, 2016b). Cette approche préconise l'utilisation de la structure secondaire comme matrice pour générer de nouvelles séquences se repliant exactement dans la même structure. Elle permet de prendre en considération des motifs comme les pseudonœuds, présents dans un grand nombre de ribozymes, pour lesquels la prédiction et la conception sont complexes (Perrotta & Been, 1991; Yamada & Tanaka, 2020). Cette complexité de conception de pseudonœuds vient du manque de données sur les paramètres thermodynamiques de leur formation. Ces paramètres essentiels à leur conception ne sont pas très bien compris et sont actuellement encore étudiés (Cao & Chen, 2006; Mandic *et al.*, 2019; Shi *et al.*, 2018). De plus, les pseudonœuds se forment par l'interaction entre une boucle et un autre motif de région simple brin d'ARN ; par contre, les données de modèles d'énergie de boucles sont limitées. Pour ces raisons, les outils de prédictions et de conceptions se basant seulement sur la MFE sont incapables de prédire ce motif et lorsqu'ils le font, le taux de fausses prédictions est élevé (Hajdin *et al.*, 2013).

Pour régler ces problèmes, le repliement inverse tient compte de différentes contraintes structurales qui sont parfois difficiles à prédire (Churkin *et al.*, 2018). Cette capacité du repliement inverse est due à l'utilisation d'algorithme assemblant la fonction de partition et la MFE. Ceci fait en sorte que les probabilités de structures prédites ne sont pas limitées à une seule structure définie seulement par la MFE. En fait, la conception par cette

approche commence par la génération d'une séquence « *seed* » qui sera mutée dans chaque position où des nucléotides non conservés existent. Les cycles de mutations se poursuivent jusqu'à ce que la séquence réponde aux critères de la structure entrée au début (Churkin *et al.*, 2018). Dans le contexte de ces conceptions, un paramètre nommé « *ensemble defect* » est pris en considération. Il définit le nombre moyen de nucléotides mésappariés à un équilibre thermodynamique. En réduisant la valeur de ce paramètre, la probabilité d'obtention de la bonne structure augmente (Zandi *et al.*, 2016). Dans le cas de la (MFE), qui permet de comparer la structure de la nouvelle séquence à l'originale, la différence de MFE des paires de bases formées par la nouvelle séquence et celle de la structure visée doit être minimisée.

La représentation de la structure dans un algorithme de « *inverse folding* » peut être produite par des symboles qui sont des « () » correspondants aux paires de bases, des points « . » indiquant les nucléotides des régions simple brin. Finalement, pour la conception de structures avec des pseudonœuds, « [] » définissent les nucléotides formant des paires de bases générant ce motif (Zandi *et al.*, 2016). Alors, le principe de la conception est de conserver la séquence de nucléotides indispensables à l'activité d'un ARN et à la modification de tout le reste. Dans le cas de ribozymes, afin de maintenir l'activité, les nucléotides cruciaux à la réaction de clivage (cœur catalytique) sont conservés.

La première utilisation du repliement inverse a été effectuée dans le cadre du *Vienna RNA package* (Hofacker, 2004), qui est un ensemble de programmes permettant la conception, la computation et l'analyse des séquences et structures d'ARN ; ceci est suivi par le développement de plusieurs autres outils (**Table 1.1**). En se basant sur la suite de logiciel Vienna RNA, plusieurs programmes ont été développés, permettant la conception de différents types d'ARN, y compris les ribozymes. Un exemple de service web disponible est le RNAiFold, ayant prouvé son efficacité par la conception des ribozymes en tête de marteau qui étaient actifs *in vitro* (Dotu *et al.*, 2014).

Ce qui manque dans RNAiFold est la capacité de prendre en considération les pseudonœuds qui ont un rôle majeur dans la formation et la stabilisation de la structure active de plusieurs ribozymes (Chadalavada *et al.*, 2002; Penedo *et al.*, 2004; Peselis & Serganov, 2014; Rastogi *et al.*, 1996; Staple & Butcher, 2005; Vusurovic *et al.*, 2017). De nouveaux algorithmes ont été développés et les ribozymes synthétiques générés ont été

actifs. Le premier est l'algorithme Enzymer, qui a la capacité de concevoir n'importe quel type d'ARN ayant un ou plusieurs pseudonœuds (Zandi *et al.*, 2016). Un autre algorithme se montrait efficace pour la conception des ribozymes HDV synthétique ayant un double pseudonœud. Il a été prouvé que les ribozymes sont actifs *in vitro*, et la comparaison de leurs structures avec celle du WT a été faite par *in-line-probing* afin de montrer que les ribozymes conçus se replient dans la même structure (Yamagami *et al.*, 2019).

L'intérêt de prouver l'activité de nouveaux ribozymes ou d'ARN conçus, ainsi que la preuve qu'ils forment la bonne structure, permet, en premier lieu, d'évaluer l'efficacité des outils de conception. De plus, les données expérimentales permettent de définir les problèmes et défauts de conceptions d'un outil bioinformatique, en identifiant les défauts de structures ou d'activités. Ce retour joue le rôle d'un feedback qui permet d'optimiser les paramètres de conception, à plusieurs reprises jusqu'à l'obtention des meilleures conceptions possibles.

Table 1.1 Différents logiciels basés sur l'approche de 'inverse folding'

Program	Service web	Code source	Extension de pseudo-noeuds	Capacité cible multiple	Remarques
AntaRNA	+	+	+	-	
RNAiFold	+	+	-	+	Expérience de laboratoire en biologie
RNAinverse	+	+	-	-	Premier program développé- Expérience de laboratoire en biologie
NUPACK	+	+	-	-	Optionnel : Structure multibrins de cible- Expérience de laboratoire en biologie
INFO-RNA	+	+	-	-	
RNA-SSD	+	-	-	-	
Frnakenstein	-	+	-	+	
ERD	+	+	+	-	
MODENA	-	+	+	+	
IncaRNAfbinv	+	+	-	-	Sélection de fragment activée. Expérience dans la détection d'ARN
RNAfbinv	-	+	-	-	Sélection de fragment activée
RNAexinv	-	+	-	-	Pas de sélection de fragment par l'utilisateur
IncaRNAtion	-	+	-	-	Approche globale-locale
RNA-ensign	-	+	-	+	Approche globale
NanoFolder	+	-	+	-	Nanostructures - ARN multibrin - expérience de laboratoire en biologie
CDSfold	+	+	-	-	Conception de séquences codant des protéines
RNAdesign	-	+	-	-	
EternaBot	+	-	-	-	Règles de conception définies par les joueurs d'Eterna
RNA-redesign	+	-	-	-	Tridimensionnel : Backbone fixe

Adapté de de (Churkin et al., 2018)

Finalement, les ribozymes synthétiques sont passés par différents stades de sélection et conception ; ils évoluent actuellement avec l'avancement technologique (résumé dans la

figure 1.12). Tous ces essais ont pour but d'avoir de meilleurs ribozymes utilisables dans des applications réelles.

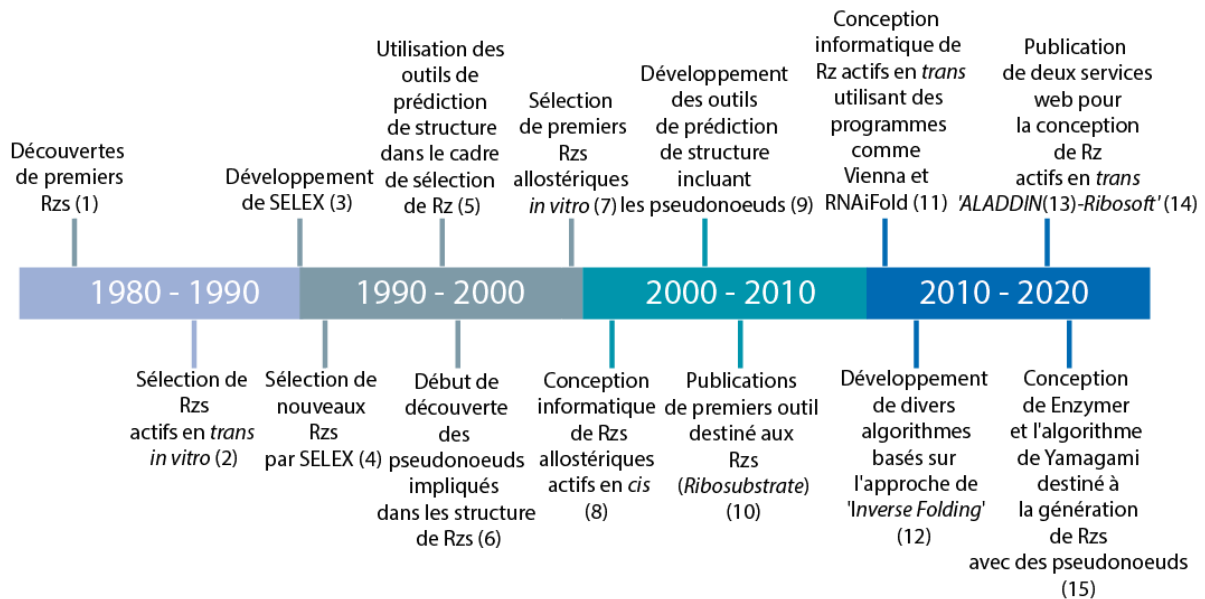


Figure 1.14 Chronologie de différentes méthodes de sélection et conception de ribozymes.

*Rzs : Ribozymes. Références (1) (Kruger *et al.*, 1982), (2) (Lieber & Strauss, 1995), (3) (Ellington & Szostak, 1990), (4) (Bartel & Szostak, 1993), (5) (Nakamaye & Eckstein, 1994), (6) (Jeng *et al.*, 1996), (7) (Tang & Breaker, 1997), (8) (Penchovsky & Breaker, 2005), (9) (Hofacker, 2004), (10) (Lucier *et al.*, 2006a), (11) (Petkovic *et al.*, 2015), (12) (Dotu *et al.*, 2014), (13) (Mercatanti *et al.*, 2012), (14) (Kharmat *et al.*, 2016), (15) (Zandi *et al.*, 2016) (Yamagami *et al.*, 2019)

1.5. Applications et utilisations de ribozymes

Grâce à leur activité de clivage et à leur petite taille de ribozymes, ils ont été utilisés dans une large gamme d'applications à différentes fins. En fait, le ribozyme en tête de marteau est l'ARN catalytique le plus étudié, ce qui a beaucoup aidé à la compréhension de la catalyse de l'ARN. C'est grâce à cette compréhension qu'il y a eu le développement de méthodes expérimentales et bioinformatiques de conception de nouveaux ribozymes pour permettre de concevoir plus facilement des ribozymes qui répondent exactement au besoin de l'application, en particulier lorsqu'il s'agit de cibler un ARN pour qu'il soit coupé.

1.5.1. But thérapeutique

Les ribozymes actifs en *trans* ont été utilisés dans un but thérapeutique en ciblant des gènes impliqués dans plusieurs maladies comme différents types de cancers, de même

que des maladies musculaires ou neuronales. Par exemple, dans le cas de la dystrophie musculaire oculopharyngée, les ribozymes attaquent spécifiquement l'ARNm responsable de la production d'une protéine aberrante comme la PABPN1 (Kharma *et al.*, 2016).

Concernant les cancers, plusieurs essais ont été faits *in vitro*, dans des cellules et *in vivo*, où les ribozymes ont montré une efficacité dans la dégradation des ARNm cibles causant la diminution de l'expression, par exemple pour HER-2 (*human epidermal growth factor receptor 2*), impliqué dans le cancer du sein (He *et al.*, 2010). Aussi une combinaison de trois ribozymes en tête de marteau avait montré une forte efficacité dans la réduction de l'expression de survivine, résultant une diminution du taux de survie des cellules cancéreuses chez des souris (Fei *et al.*, 2008).

D'autres ribozymes ont été capables de dépasser les tests de laboratoire pour passer aux essais cliniques. Afin de pouvoir utiliser les ribozymes *in vivo*, il fallait améliorer leur stabilité (résistance à la dégradation) pour augmenter leur demi-vie. Dans ce but, plusieurs modifications chimiques ont été incorporées dans les ARN, comme la substitution du 2'-OH par 2'-O-méthyl et 2'-deoxy-2'-C-allyl uridine. Ceci est le cas du ribozyme nommé « Angiozyme » qui diminue l'angiogenèse. Le ribozyme cible le VEGFR-1 (*Vascular Endothelial Growth Factor Receptor-1*) et il a été testé chez des patients ayant le cancer des reins (Weng *et al.*, 2005) et le cancer du sein (Schubert & Kurreck, 2004). En plus de l'utilisation des ribozymes contre les gènes impliqués directement dans le développement et la survie des cellules cancéreuses, des ribozymes allostériques ont été utilisés pour contrôler l'expression des vecteurs et des gènes de suicides à exprimer dans les cellules cibles. En guise d'exemple, un aptazyme contrôlé par la théophylline a été inséré dans le 3'UTR du gène de suicide HSV-TK (*Herpes simplex virus-thymidine kinase*). Ceci a été élaboré afin de développer un système de suicide cellulaire dose-dépendant, en réponse à la quantité de la théophylline introduite (Morrow *et al.*, 2012; Zhang *et al.*, 2017).

Ces études cliniques sont terminées sans mener à un traitement utilisable. Bien que les ribozymes aient présenté une cytotoxicité faible et qu'ils étaient assez tolérés par les patients, leur efficacité demeurait faible. Cette faiblesse a été remarquée par une réduction basse de l'expression des gènes ciblés et la durée de leurs activités a été assez courte. Ce manque d'efficacité est lié au fait que, pour plusieurs études, la version minimale a été utilisée. Ceci a abouti à l'abandon de leur utilisation et à la concentration sur d'autres outils, comme les petits ARN interférents (siRNA), qui ont montré une grande efficacité (Schubert & Kurreck, 2004).

En fait, les siRNA, liés à un complexe protéique nommé RISC, sont capables de reconnaître des ARNm cibles et de s'hybrider par complémentarité de séquences sur ceux-ci. Cette liaison induit alors la dégradation de l'ARNm par l'endonucléase Argonaute. Il s'agit d'un mécanisme naturel, mais il était adapté à l'utilisation comme thérapie par la conception de séquences synthétiques de siRNA pour des cibles de choix. Ces conceptions sont passées par plusieurs étapes d'optimisation avant d'atteindre le stade auquel ils sont aujourd'hui (Kim *et al.*, 2019; Maguregui & Abe, 2020; Zhang *et al.*, 2021b). Ces ARN étaient d'une grande efficacité et plusieurs sont utilisés dans des études cliniques qui sont encore en cours, alors que d'autres sont déjà approuvés comme thérapies (Garrelfs *et al.*, 2021) (Honor *et al.*, 2021) (Milani *et al.*, 2019). Néanmoins, comme pour toute autre thérapie génique, ils font face à quelques obstacles liés au mode de livraison, à la courte demi-vie et parfois au manque de spécificité (Hu *et al.*, 2020).

Alors, bien que les ribozymes aient fini par échouer par manque d'efficacité, l'optimisation des paramètres de conceptions et une meilleure compréhension de leurs interactions avec leurs cibles aideront à les proposer, encore une fois, comme alternative de thérapie.

1.5.2. Thérapie antivirale

De la même façon que les ribozymes ont été utilisés contre les gènes du cancer et d'autres maladies, ils ont été étudiés comme des thérapies antivirales (Wang *et al.*, 2010). Un exemple de ribozyme qui a atteint les phases cliniques est l'heptazyme, ciblé contre la région 5'UTR du génome du virus de l'hépatite C (Anonyme, 2000). Par ailleurs, plusieurs études ont démontré l'efficacité de ribozymes (en tête de marteau, hairpin et HDV) ciblant le virus de l'immunodéficience humaine (Fedoruk-Wyszomirska *et al.*, 2015; Scarborough & Gagnon, 2015; Scarborough *et al.*, 2014). Quelques ribozymes ont atteint les études cliniques (Sullenger & Gilboa, 2002), incluant deux ribozymes *hairpin* ciblant la région 5'UTR du génome de VIH (Wong-Staal *et al.*, 1998). Malgré l'échec des essais cliniques déjà peu nombreux, les études sur les ribozymes se poursuivent. Un exemple récent est la proposition d'un système qui combine des ribozymes avec des *nanocarriers* afin de cibler le virus SARS-COV2 en supposant que ceci aura de meilleurs résultats par rapport aux anciennes études par l'utilisation d'un ribozyme chimère d'ARN/ADN (Donmus *et al.*, 2021).

Les ribozymes ont fait face à plusieurs obstacles, commençant par la méthode de livraison, un problème commun entre toutes les thérapies géniques testées. L'autre paramètre qui a beaucoup joué sur l'activité des ribozymes est l'accessibilité de leurs

cibles. Les structures des cibles dans les milieux physiologiques sont complexes et ceci crée une compétition entre le ribozyme, qui doit s'hybrider sur l'ARN afin d'effectuer le clivage, et les différentes régions de la cible elle-même.

1.5.3. Biosenseurs et circuit logique

Plusieurs ribozymes allostériques répondant à différents types de molécules et oligonucléotides ont servi de biosenseurs et d'éléments de circuit logique. Ceci s'explique par le fait que la liaison du ligand par l'aptazyme induit ou inhibe une réaction de clivage qui peut être rendue détectable par l'ajout de la fluorescence, par exemple (Kuwabara *et al.*, 2000). Cette détection permet un suivi en temps réel de la liaison et du clivage. Les ribozymes de Penschovsky ont été développés de façon à opérer différentes fonctions logiques. En réalité, compte tenu de la structure formée ainsi que du nombre de séquences « *input* », la fonction booléenne doit être déterminée avant que les nouvelles séquences soient conçues. Parmi les fonctions logiques développées, nous retrouvons les « ET, OU, NON et OUI » (Penschovsky, 2014). Ce principe a entraîné la détection des séquences spécifiques de miARN permettant d'émettre une fluorescence comme réponse générée par le clivage (Hartig *et al.*, 2004). Penschovsky a aussi développé des circuits qui détectent deux à trois *inputs* pour construire des multiplexeurs et décodeurs en décelant la taille de la cible formée par une séquence répétée. Dans cette étude, il propose que le système puisse être appliqué à des maladies comme la dystrophie musculaire oculopharyngée causée par une altération génique dans le nombre de séquences de triplet de nucléotides répétés (Penschovsky, 2012).

Outre les oligonucléotides, les ribozymes allostériques liant des protéines ou des métabolites peuvent généralement fonctionner comme biosenseurs pour la détection de ces molécules. La détection de l'état de phosphorylation de l'enzyme ERK2 (nommée aussi MAPK1 : *Mitogen-activated protein kinase 1*) représente un exemple d'application, celle-ci étant active lorsque phosphorylée. Également, elle est impliquée dans la malignité des cellules cancéreuses. Le ribozyme est seulement actif lorsqu'il lie la protéine phosphorylée et qu'il est en mesure de la distinguer des protéines similaires dans un lysat de cellules (Vaish *et al.*, 2002). Un exemple récent est celui d'un aptazyme détectant la thiamine pyrophosphate, considérée comme biomarqueur sanguin de la maladie d'Alzheimer (Du *et al.*, 2021).

Une utilisation différente de ribozymes compose le développement d'un système de détection de conditions environnementales. Un ribozyme en tête de marteau a été utilisé pour bloquer ou permettre l'expression d'un ARN thermomètre, en relation avec la température. Le ribozyme bloque la séquence de Shine Dalgarno ; à des températures basses, le clivage est permis, ce qui libère le SD. La traduction peut alors avoir lieu. Lorsque la température est augmentée, le ribozyme perd son activité catalytique et le SD demeure bloqué, ce qui inhibe la traduction (Saragliadis *et al.*, 2013).

En fait, les ribozymes allostériques, selon leurs réponses aux ligands correspondants, peuvent être considérés comme des portes logiques « OUI » et « NON ». En plus de la réponse simple résultant de la réaction de clivage d'un seul ribozyme, la combinaison de plusieurs ribozymes, ensemble ou avec d'autres types d'ARN, permet la construction de circuits logiques ayant des entrées multiples et effectuant différentes fonctions logiques (Auslander *et al.*, 2018).

Beaucoup de ribozymes allostériques sont testés ou développés dans le but d'agir comme biosenseurs, mais jusqu'à maintenant, aucune application réelle n'est trouvée sur le marché. Les raisons sont multiples, en commençant par la sensibilité de l'ARN à la température et aux nucléases qui causent facilement sa dégradation. De plus, le processus de sélection, caractérisation et optimisation d'un biosenseur, est plutôt long. Dans le but de résoudre ce problème, une équipe de l'Université Stanford a développé une méthode automatisée de sélection de biosenseurs « DRIVER » (*De novo Rapid In Vitro Evolution of RNA biosensors*). Ceci a été combiné à une étape de caractérisation avec un essai haut débit, basé sur l'activité de clivage et une analyse par séquençage haut débit qu'ils ont nommée « CleaveSeq » (Townshend *et al.*, 2021a).

1.5.4. Besoin de circuits logiques dans le cancer

Plusieurs exemples de circuits logiques et d'utilisation de biologie synthétique ont été appliqués sur de nombreux types de cancers (Liu *et al.*, 2020; Matsuura *et al.*, 2018; Quan *et al.*, 2019; Song *et al.*, 2019). Ceci est dû au fait que le cancer regroupe une panoplie de maladies ayant en commun la caractéristique de prolifération non contrôlée de cellules malignes touchant n'importe quel organe. De ce fait, un seul traitement ou moyen de diagnostic ne peut s'appliquer à tous les types de cancers. C'est dans cette optique que des thérapies et des outils de diagnostics basés sur des acides nucléiques naturels, modifiés ou synthétiques, combinés ou non à des protéines, ont été développés (Boscaino

et al., 2019; Chen *et al.*, 2022; Nissim *et al.*, 2017; Yang & Ding, 2021). La combinaison de plusieurs de ces ARN formant des circuits logiques peut potentiellement être applicable au diagnostic ou à la thérapie de cancer. L'espoir est que ces circuits offrent une personnalisation et une spécificité meilleures que celles des chimio- et radiothérapies, qui sont parmi les traitements les plus utilisés contre le cancer. La spécificité vient du fait qu'un circuit logique pourrait détecter simultanément plusieurs biomarqueurs et cibler plusieurs oncogènes, par exemple. Ce genre de diagnostic ou thérapie sont d'une importance élevée, surtout dans le cas de certains types de cancers ayant de mauvais pronostics et une forte mortalité, et pour lesquels il n'y a toujours pas de thérapies ciblées. Parmi ces types de cancers, on compte celui du sein triple négatif, et le cancer séreux de haut grade de l'ovaire. Ce sont parmi les sous-types de cancers ayant un taux de survie inférieur à 50 % (Dube-Delarosbil & St-Pierre, 2018).

Afin de développer des thérapies ciblées ou personnalisées, la détection de multiples biomarqueurs est utile. Parmi les molécules souvent utilisées comme biomarqueurs sont les protéines, et récemment des profils d'expression de miARN. Des protéines comme MMP9 (métalloprotéinase matricielle 9) (Farina & Mackay, 2014), VEGF (facteur de croissance de l'endothélium vasculaire) (Goel & Mercurio, 2013) et galectines (Kaur *et al.*, 2016) sont surexprimées dans plusieurs types de cancers. Par conséquent, leur présence simultanée dans un échantillon peut indiquer l'existence d'une néoplasie. Ces molécules combinées peuvent être utilisées comme un signal qui permet d'identifier spécifiquement un certain type de cancer. Ce signal pourrait permettre la personnalisation d'un diagnostic d'identification des types et sous-types de cancers. Les profils de sous- et surexpression de miARN sont différents d'un sous-type de cancer à un autre, ce qui permet alors de les distinguer s'ils sont utilisés en diagnostic (Anelli *et al.*, 2021; Sokilde *et al.*, 2019; Wang *et al.*, 2017; Wang *et al.*, 2019; Yerukala Sathipati & Ho, 2018; Zhu *et al.*, 2014).

Parmi les biomarqueurs qui attirent l'attention récemment, ce sont les galectines (Bartolazzi, 2018; Girotti *et al.*, 2020). Il s'agit en fait d'une famille de protéines de 15 membres ayant en commun une structure comprenant au moins un domaine de reconnaissance de carbohydrate (CRD). La plupart des membres de cette famille ont une affinité à un sucre β -galactoside (Vladoiu *et al.*, 2014) et peuvent lier les glycolipides et les glycoprotéines présents à la surface cellulaire et sur la matrice extracellulaire.

Cette liaison permet une activation ou une inhibition de certaines voies de signalisations. Il a été démontré que dans plusieurs types de cancer, l'expression des galectines est dérégulée. Cette dérégulation peut résulter par un effet pro- ou anti-cancer dépendamment du type du cancer ainsi que de la localisation des protéines dans les cellules. Dans le cas où elles participent dans l'agressivité du cancer, elles agissent notamment sur l'adhésion et la migration cellulaire ainsi que sur l'apoptose et la résistance aux thérapies. De plus, elles sont en mesure de moduler la réponse immunitaire (Brinchmann *et al.*, 2018; Chou *et al.*, 2018; Thiemann & Baum, 2016) (Li *et al.*, 2021a). L'intérêt des galectines est récent par rapport aux autres protéines oncogènes et chaque membre de la famille a ses propres caractéristiques lui accordant un rôle dans certains types de cancer (Girotti *et al.*, 2020).

Comme exemple la galectine 7 est caractérisée par une surexpression marquée dans les cellules épithéliales touchées par un cancer en comparaison aux tissus sains (Advedissian *et al.*, 2017). De ce fait, cette galectine peut être utilisée comme un biomarqueur des cancers comme ceux du sein et de l'ovaire, ainsi qu'une cible de thérapie (Advedissian *et al.*, 2017; Sewgobind *et al.*, 2021). Des études ont montré que l'inhibition de l'activité de la galectine 7 peut mener à l'atténuation de son rôle pro-cancer (Vladoiu *et al.*, 2015).

2. Contexte de recherche et problématique

La complexité du cancer vient du fait qu'il existe de très nombreux types et sous-types, de sorte qu'une seule thérapie ou méthode de diagnostic ne peut être appliquée pour les guérir tous. Deux raisons majeures rendent le taux de mortalité de certains types de cancers très élevé, soit le retard de diagnostic, puis des traitements qui ne sont pas assez efficaces ni spécifiques. Le problème vient du fait qu'un cancer, par exemple le cancer séreux de l'ovaire de haut grade ou du sein triple négatif, est généralement diagnostiqué lorsque les tumeurs ont déjà atteint un stade avancé. Cela entraîne un taux de survie beaucoup plus faible par rapport à d'autres sous-types ayant un meilleur pronostic. De plus, actuellement, la détection de présence de tumeur ou de cellules malignes est réalisée de différentes façons et avec diverses méthodes. L'une de celles-ci est la détection de biomarqueurs de cancer, révélant comme problème le fait que dans la majorité des cas, ce sont des molécules présentes et sécrétées aussi par des cellules normales. Ceci engendre un manque de précision. L'autre problème des méthodes de détection est le temps requis pour l'obtention de résultat.

Des solutions comme l'application d'outils de diagnostic préliminaires donnent des résultats plus rapides en utilisant des biomarqueurs spécifiques comme des protéines, et des miARN sont nécessaires. Une partie de la solution est d'utiliser des outils ayant des temps de production et de réaction plus courts que les méthodes disponibles qui sont généralement basés sur des anticorps ou la détection de l'expression de certains gènes. Bien que les anticorps soient l'outil standard utilisé actuellement, le processus de leur sélection et synthèse reste toujours assez long, demandant et coûteux. Parmi les outils proposés récemment, il y a les traitements ainsi que les méthodes de détection basées sur les acides nucléiques, qui constituent dans notre projet des ribozymes conventionnels et allostériques. Notre choix de ribozymes est basé sur le fait qu'ils peuvent être utilisés à la fois pour cibler des gènes impliqués dans le cancer et pour détecter les biomarqueurs. Bien que les ribozymes semblent être des molécules assez inefficaces comme une thérapie qui traite le cancer, du moins si on se fie aux échecs précédents, l'idée du projet est de prouver un concept qui implique les ribozymes comme outil facile à utiliser pour détecter et cibler différentes molécules. La combinaison de plusieurs ribozymes senseurs et effecteurs, qui sont facilement générés à l'aide des outils bio-informatiques et des sélections *in vitro*, peut aboutir à la formation d'un *toolbox* utilisable d'une façon personnalisée dépendamment du type du cancer ou de la maladie ciblée. De plus, même

si les ribozymes ne seront pas nécessairement assez efficaces pour traiter une maladie à eux-seuls, leur utilisation peut être dans le but d'atténuer l'avancement d'un cancer ou de générer une meilleure réponse aux autres thérapies.

Les ribozymes ont montré leur efficacité dans une multitude d'applications *in vitro* et *in vivo*, mais en arrivant au stade des études cliniques, ils ont échoué. Cet échec dû à la qualité et au type des ribozymes utilisés qui sont des ribozymes minimaux, était probablement la raison majeure derrière l'abandon des ribozymes. De plus, les découvertes et le développement de nouvelles molécules et méthodes applicables dans la régulation de l'expression génique ont attiré plus d'attention ces dernières années. Parmi ces méthodes, nous citons CRISPR, les petits ARN interférant, les miARN, *zinc-finger* et TALENS. Ce sont tous des systèmes qui, contrairement aux ribozymes, impliquent des protéines comme agents effecteurs.

Les ribozymes demeurent des candidats potentiels pour des applications futures, car ils présentent les avantages d'être facilement modulés, d'indépendance de cofacteurs et de possibilités de contrôle de leur activité par l'ajout d'un senseur qui lie un ligand ou un oligonucléotide. Leur petite taille permet d'en combiner plusieurs afin d'en faire un circuit qui répondrait à la fonction voulue par l'utilisateur ou en fonction de l'application future. Toutes ces caractéristiques permettent leur utilisation dans des applications complexes - comme la thérapie génique-, et le diagnostic de maladie complexe -comme le cancer-.

Pour ce faire, les ribozymes conventionnels ou allostériques doivent être modifiés selon l'application désirée. Qu'il s'agisse de les diviser en deux pour créer des ribozymes en *trans* ou de sélectionner un aptazyme, ce sont toutes des modifications importantes avec un risque de réduire drastiquement l'activité. Cette ingénierie affecte la séquence, donc la structure représentant la clé pour le bon fonctionnement d'un ribozyme. Actuellement, aucun logiciel ne permet la conception automatisée de ribozymes synthétiques en prenant en considération des paramètres structuraux et environnementaux.

Dans mon travail, nous avons collaboré avec une équipe de l'Université Concordia, pour le développement et l'optimisation de programmes de conception de ribozymes actifs en *cis* et en *trans*, conventionnels et allostériques. Afin de montrer l'efficacité de ces programmes, nous avons choisi d'appliquer les ribozymes comme agent de régulation de l'expression génique chez des bactéries et certaines cellules cancéreuses. La cible des ribozymes actifs en *trans* étant l'ARNm des protéines ayant un rôle dans la survie des

cellules cancéreuses, deux protéines ont été choisies, la Bcl-2 connue par son activité anti-apoptotique dans une majorité de types de cancers, et la galectine 7 afin de détecter l'effet de la diminution de son expression sur la croissance cellulaire. Aussi, nous avons utilisé les ribozymes comme biosenseurs pour la détection des miARN. Autre que les outils bioinformatiques, la sélection *in vitro*, qui est la méthode traditionnelle de sélection de nouveaux ribozymes et aptamères, constituait notre choix pour la sélection d'aptazymes liant des protéines impliquées dans certains types de cancers.

3. Hypothèse et objectifs

3.1. Hypothèse

Les différentes approches d'ingénierie de ribozymes permettent la construction de biosenseurs et de circuits logiques applicables dans de potentielles nouvelles thérapies et outils de diagnostics.

3.1.1. Sous-hypothèses

- 1- De meilleurs programmes permettraient plus de diversité de types et structures de ribozymes conventionnels ou allostériques incluant ou non des pseudonœuds.
- 2- Des ribozymes allostériques pourraient détecter des biomarqueurs de cancer.¹

3.2. Objectifs

- **Objectif 1**

Génération de ribozymes ayant des pseudonœuds actifs en cis pour les appliquer comme régulateurs de l'expression génique en cis.

Pour réaliser cet objectif, nous avons eu recours au programme Enzymer qui permet la conception des ARN ayant des pseudonœuds. Ceci nous a permis de tester des ribozymes conventionnels et un allostérique et d'en utiliser comme un '*switch-OFF*' de l'expression d'un gène rapporteur exprimant une protéine fluorescente.

- **Objectif 2**

Évaluer l'activité de plusieurs types de ribozymes actifs en *trans* générés d'une façon automatisée à l'aide du service web Ribosoft2.0. Ceci est dans le but de les appliquer comme une stratégie anti-sens pour des gènes impliqués dans le cancer du sein. Aussi, essayer de trouver une concordance en comparant les résultats expérimentaux avec les paramètres calculés par Ribosoft 2.0.

Pour atteindre cet objectif nous avons testé une totalité de 60 ribozymes en *trans*, conventionnels et allostériques. Enfin, les ribozymes ayant montré la meilleure activité *in vitro* ont été choisis pour être testés dans des cellules eucaryotes contre l'ARNm d'une protéine impliquée dans le cancer.

- **Objectif 3**

Évaluer le fonctionnement de ribozymes allostériques contrôlés par un ou plusieurs miARN générés d'une façon automatisée dans le but de les utiliser comme porte logique.

Afin d'atteindre cet objectif une librairie de ribozymes théoriquement capables de lier plusieurs miARN simultanément a été testée. De plus des ribozymes conçus par Ribosoft2.0 liant un seul miARN ont aussi été testés *in vitro* et dans des cellules eucaryotes.

- **Objectif 4**

Sélectionner des ribozymes allostériques liant des protéines impliquées dans plusieurs types de cancers afin de les appliquer comme biosenseurs ou outil de diagnostic.

La méthode SELEX a été utilisée pour sélectionner des aptazymes qui étaient activés ou inactivés par les protéines. Nous avons pu montrer que ces aptazymes peuvent être transformés en biosenseurs pouvant être applicables comme outils de diagnostic.

En conclusions, les ribozymes auto-clivants sont l'élément principal de mon travail de doctorat. La réussite de toute application basée sur des ribozymes synthétiques ou modifiés commence par le choix du meilleur ribozyme adapté à nos besoins. Dans ce but nous avons collaboré avec un laboratoire spécialisé dans la biologie computationnelle afin de mettre en place de nouveaux programmes et outils permettant de concevoir différents types de ribozymes (actifs en *cis* ou en *trans*, conventionnels ou allostériques...). Notre but ultime était de construire un circuit logique en assemblant les différents types de ribozymes sélectionnés dans le but d'avoir une thérapie génique intelligente (**Figure 3.1**).

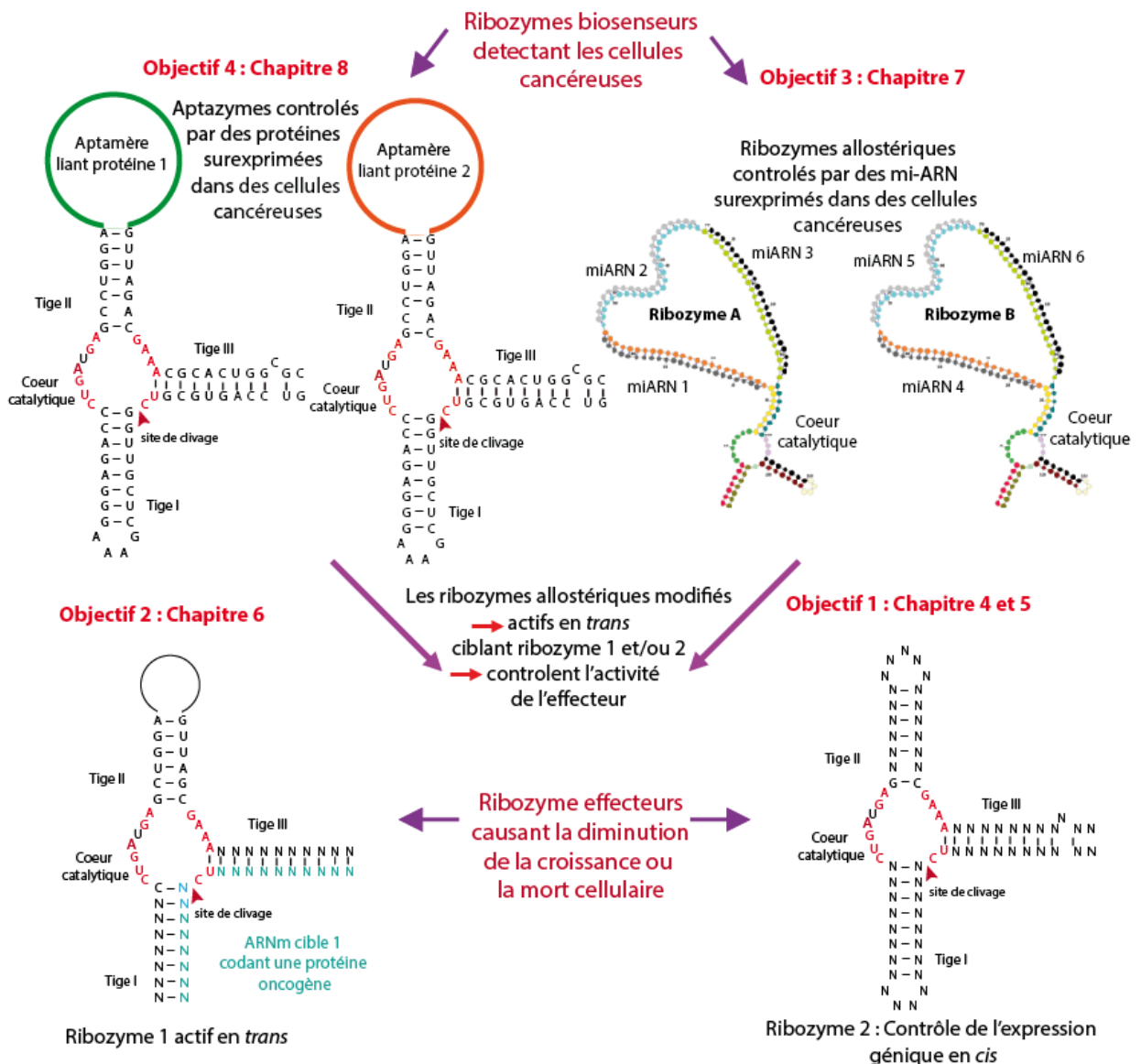


Figure 3.1 Circuit logique possible = thérapie intelligente

Les différents types de ribozymes sélectionnés dans le cadre de ce projet peuvent être assemblés pour former un circuit logique utilisable à la fois pour la détection et l'attaque des cellules cancéreuses. Les ribozymes allostériques liant les miARN et ceux contrôlés par des protéines surexprimées dans certains types de cancer, assemblés agiront comme biosenseurs pour permettre au circuit de décider s'il est dans une cellule cancéreuse ou saine. Les ribozymes actifs en *trans* et en *cis* agiront comme effecteurs qui dégradent les ARNm des protéines essentielles à la survie des cellules cancéreuses. Les ribozymes allostériques sont modifiables pour être actifs en *trans* et attaquer les ribozymes effecteurs en jouant le rôle de 'Safety Switch'. Ceci est applicable dans le cas où les ribozymes allostériques sont inactifs en présence de leurs ligands mais actifs en leurs absences.

4. Article 1: Computer-Aided Design of Active Pseudoknotted Hammerhead Ribozymes

4.1. Mise en contexte

Dans le but de leur application comme outil thérapeutique ou leur inclusion dans un circuit logique, les ribozymes subissent de multiples modifications. Bien que ces derniers montrent une certaine flexibilité de structure et séquence pour tolérer ces modifications. Par contre, certains éléments doivent être pris en considération. Comme exemple, le cœur catalytique ne tolère aucune mutation et la structure secondaire doit être conservée. Afin d'éviter de réduire l'activité d'un ribozyme avec une modification non convenable, un outil de conception efficace et rapide est utile.

Ce chapitre de la thèse porte sur la conception de ribozymes ayant une activité d'auto-clivage. L'élément principal de l'article présenté dans cette section est l'habilité d'un nouvel outil informatique «Enzymer» à générer des ribozymes ayant un ou plusieurs pseudonœuds. Ce dernier est un motif de structure secondaire existant dans un grand nombre de ribozymes auto-clivants et ayant un rôle majeur dans la stabilisation de leurs structures. Nos résultats montrent que les nouveaux ribozymes adoptent une conformation active comparable à celle des ribozymes utilisés comme matrices. Nos tests ont prouvé l'activité des ARN conçus en se basant sur les matrices de différents types de ribozymes. Enzymer nous permettra par la suite de concevoir et modifier les ribozymes allostériques à inclure dans le circuit logique selon les besoins de l'application (**Figure 4.1**)

But d'utilisation de Enzymer

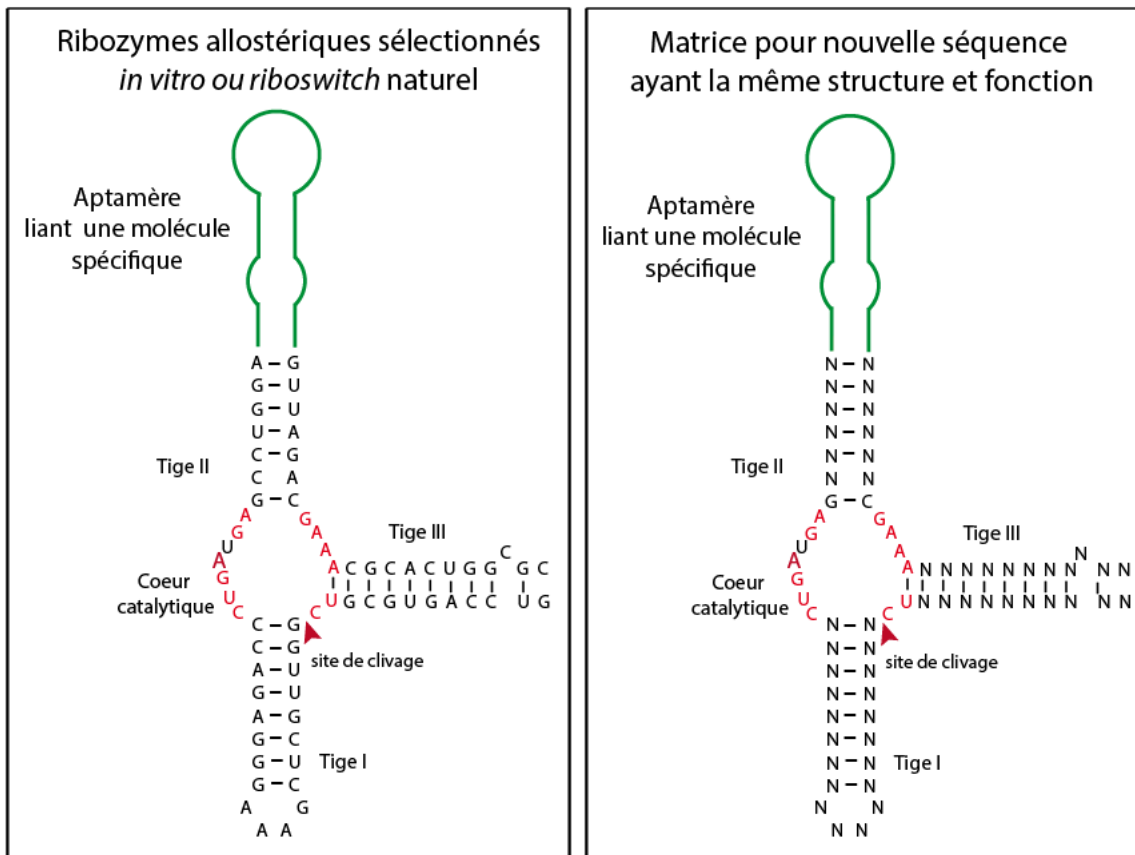


Figure 4.1 Utilisation de « Enzymer »

La structure secondaire et les nucléotides conservés d'un ARN naturel ou sélectionné *in vitro* peuvent être utilisés comme entrée pour Enzymer. Cette structure secondaire peut inclure des pseudonoeudss. Ceci est la particularité d'Enzymer, car il est capable de générer des séquences ayant des structures assez complexes comprenant un ou plusieurs pseudonoeuds. Dans l'exemple présenté, en utilisant la structure secondaire d'un ribozyme allostérique comme matrice, de nouvelle séquence liant le même ligand et effectuant la même réaction peuvent être générées. Le but d'utiliser un logiciel comme Enzymer est de limiter le travail expérimental habituellement nécessaire pour valider l'activité des ARN modifiés. De plus les séquences obtenues par Enzymer sont optimisées selon le besoin de l'application.

Computer-Aided Design of Active Pseudoknotted Hammerhead Ribozymes

Conception assistée par ordinateur de ribozymes en tête de marteau à pseudonœud

Sabrine Najeh¹, Kasra Zandi^{2,3}, Samia Djerroud¹, Nawwaf Kharma^{2,3}, and Jonathan Perreault¹

¹Centre Armand-Frappier Santé Biotechnologie, Institut national de la recherche scientifique (INRS), 531 boulevard des Prairies, Laval, QC, H7V 1B7, Canada

²Software Engineering & Computer Science Dept., Concordia University, 1455 boul. de Maisonneuve O., Montreal, QC, H3G 1M8, Canada

³Electrical & Computer Eng. Dept., Concordia University, 1455 boul. de Maisonneuve O., Montreal, QC, H3G 1M8, Canada

#To whom correspondence should be addressed. Tel: +450 687 5010 (Ext 4411); Fax: +450 686 5301; Email: jonathan.perreault@iaf.inrs.ca

#To whom correspondence should be addressed. Tel: +514 848 3100; Fax: +514 848 3802; E_mail: nawafkharma@gmail.com or kharma@ece.concordia.ca

Running head: Pseudoknotted ribozyme design

Sabrine Najeh and Kasra Zandi have contributed equally to this work.

Article publié:

Computer-Aided Design of Active Pseudoknotted Hammerhead Ribozymes. Najeh S, Zandi K, Djerroud S, Kharma N, Perreault J. *Methods Mol Biol.* 2021;2167:91-111. doi: 10.1007/978-1-0716-0716-9_7.

4.2. Contribution des auteurs

Sabrine Najeh : Établissement du plan expérimental avec Jonathan Perreault. Rédaction de l'introduction et toute la partie en relation avec la validation expérimentale de l'activité de ribozymes. Révision et optimisation de la méthodologie d'utilisation de 'Enzymer' décrite dans l'article en collaboration avec Samia Djerroud.

Kasra Zandi : Développement de 'Enzymer' dans le cadre de son projet de doctorat et rédaction de la partie en relation avec l'algorithme.

Nawwaf Kharma : Supervision du travail de Kasra et révision du manuscrit

Jonathan Perreault : Supervision et assistance au travail de Sabrine Najeh et Kasra Zandi et révision du manuscrit.

4.3. Abstract

Pseudoknots are important motifs for stabilising the structure of functional RNAs. As an example, pseudoknotted hammerhead ribozymes are highly active compared to minimal ribozymes. The design of new RNA sequences that retain the function of a model RNA structure includes taking in account pseudoknots presence in the structure, which is usually a challenge for bioinformatics tools. Our method includes using 'Enzymer', a software for designing RNA sequences with desired secondary structures that may include pseudoknots. Enzymer implements an efficient stochastic search and optimization algorithm to sample RNA sequences from low ensemble defect mutational landscape of an initial design template to generate an RNA sequence that is predicted to fold into the desired target structure.

Keywords: pseudoknot, hammerhead ribozyme, glmS ribozyme, pistol ribozyme, twister ribozyme, HDV ribozyme, VS ribozyme, inverse folding, RNA structure, synthetic biology

Résumé

Les pseudonœuds sont des motifs importants pour stabiliser la structure des ARN fonctionnels. À titre d'exemple, les ribozymes à tête de marteau ayant un pseudonœud sont beaucoup plus actifs que les ribozymes minimaux. La conception de nouvelles séquences d'ARN qui conservent la fonction d'une structure d'un ARN modèle, inclut la prise en compte de la présence de pseudonœuds dans la structure, ce qui est généralement un défi pour les outils bioinformatiques. Notre méthode comprend l'utilisation de « Enzymer », un logiciel permettant de concevoir des séquences d'ARN avec les structures secondaires désirées, pouvant inclure des pseudonœuds. Enzymer met en œuvre un algorithme de recherche et d'optimisation stochastique efficace pour échantillonner des séquences d'ARN à partir d'un paysage mutationnel à faible *ensemble defect* d'un modèle de conception initial afin de générer une séquence d'ARN qui devrait se replier dans la structure cible souhaitée.

4.4. Introduction

Non-coding RNAs have multiple roles in regulation, protein synthesis and even as enzymes, in many cases to perform RNA cleavage reactions. For the latter, beyond the primary sequence there is a tight relation between the RNA structure and its function. Folding in secondary and tertiary structures brings closer different regions of the RNA molecule to achieve a conformation essential for a certain reaction or function (Gupta et al., 2012).

In many cases, this tertiary interaction forms through base pairs in a pseudoknot. Pseudoknots are widely present in many types of RNA and are essential for the function of numerous RNAs by stabilising RNA structure. One or more pseudoknots can exist in a given RNA structure. Several pseudoknot configurations also exist, the most prevalent being the H-type where H is Hairpin loop with nucleotides base pairing with bases outside of the stem, but not with a terminal loop. The pairing of bases from a hairpin loop with another hairpin loop forms a HH-type pseudoknot (also called kissing-loop), and pairing with a single strand region of an internal loop, a multistem junction or a bulge is called an HL-type pseudoknot (Han & Byun, 2003) (Staple & Butcher, 2005) (**Figure 4.2**).

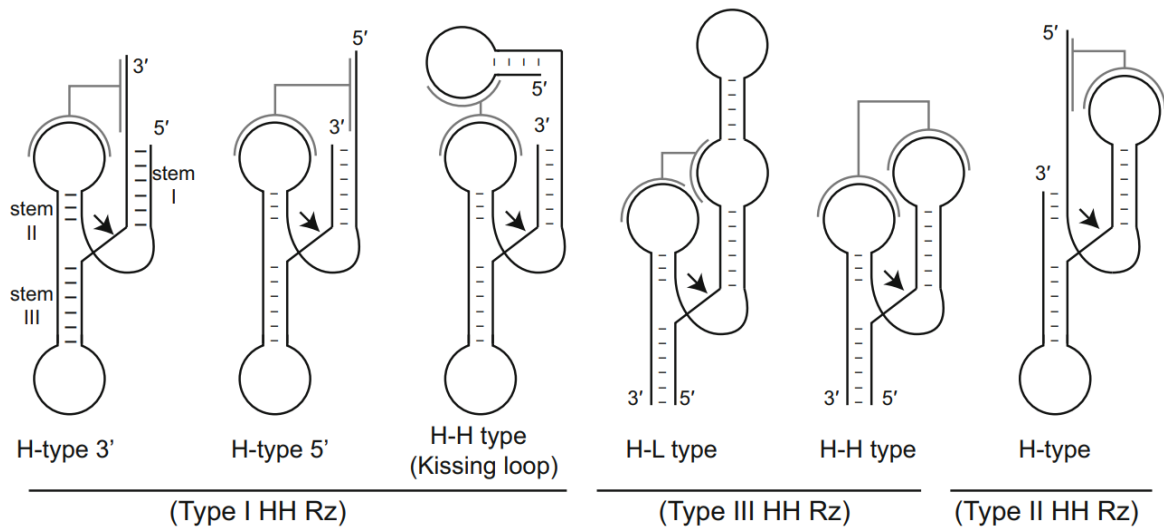


Figure 4.2 Pseudoknot types as illustrated by various hammerhead ribozymes stem I and II

Pseudoknot types as illustrated by various hammerhead ribozymes stem I and II interactions. Note that H-H type pseudoknots are also commonly called kissing-loops.

An RNA that is widely used to study the structure-function relationship in RNAs is the hammerhead ribozyme. This ribozyme is a catalytic self-cleaving RNA that was initially

discovered in plant satellite viruses and viroids (Forster & Symons, 1987a). In the context of rolling circle replication, each copy of the genome has a ribozyme that cleaves the multimeric genomic RNA in genome monomers. Following two decades of research with a minimal version of the hammerhead ribozyme, a critical interaction between stem I and II was discovered in so-called “extended” hammerhead ribozymes. Both have the same general secondary structure which is a catalytic core of conserved nucleotides surrounded by three stems. What makes the difference between the minimal and the extended ribozyme is the presence of nucleotides on stem I and II that allow a long-range interaction between these two stems. This interaction confers to the full length version of the ribozyme a 1000 fold higher cleavage rate compared to the minimal one by permitting the correct positioning of the nucleotides in the catalytic core, which is attained by the minimal version through inefficient sampling of numerous sub-optimal tertiary structures (De la Pena *et al.*, 2003; Khvorova *et al.*, 2003). It was later found that this interaction in the case hammerhead ribozymes can be a pseudoknot that brings stem I and II closer for high activity (**Figure 4.2**) (Perreault *et al.*, 2011a). A major advantage of this type of interaction over the previously characterized tertiary interaction is that it lends itself better to design ribozymes because of the predictability of base pairing.

Besides hammerhead ribozymes, many other types of ribozyme structures include pseudoknots of different types. In addition to their natural function, one or different types of self-cleaving ribozymes can be used separately or in combination in different applications. To have the desired function, it is important to make the right choice of the ribozyme to use. **Table 4.1** includes some examples of ribozyme applications previously published.

Table 4.1 Pseudoknotted ribozymes

Ribozyme	Pseudoknot type	Applications
Hammerhead	Different types	Upstream of an RNA of interest to have a discrete 5' (H and H-H pseudoknot-Type I HHRz) and downstream of it to have a 3' end (H-H and H-L pseudoknots-Type III HHRz) (Tang <i>et al.</i> , 2018) or to control gene expression – used to develop and select allosteric ribozymes (aptazymes) (H-L type) (Sack <i>et al.</i> , 2019)
GlmS	Three H-type pseudoknot : nested pseudoknot	Pull down of RNA-protein complexes (Di Tomasso <i>et al.</i> , 2016)- Control of gene expression using Glc6P. Insertion of GlmS ribozyme in a gene (5'UTR or 3'UTR) to silence it. (Niu <i>et al.</i> , 2018)
HDV	Two H-type pseudoknot : nested pseudoknot	Very stable (Levesque <i>et al.</i> , 2002) - use downstream of a desired RNA to have a discrete 3' overhang <i>in vitro</i> and <i>in vivo</i> for RNA or RNA-protein complex where the RNA needs a precise 3' sequence. (Gao & Zhao, 2014)
VS	Kissing-loop (H-H type pseudoknot)	Target specifically a stem-loop structure (Lacroix-Labonte <i>et al.</i> , 2016)
pistol	Kissing-loop (H-H type pseudoknot)	Gene expression control - Aptazyme selection based on pistol ribozyme (Kobori <i>et al.</i> , 2017)
twister	Two H-type pseudoknots	Expression platform for artificial riboswitches detecting two inputs simultaneously (Felletti <i>et al.</i> , 2016b)
Rnase P	Two pseudoknots : different types	Gene silencing- Using a guide sequence that hybridise with the target mRNA to form a structure similar to the natural target of RNase P, M1 RNA or tRNA. So it targets RNA depending on the structure not a specific sequence (Cobaleda & Sanchez-Garcia, 2001)
Group I introns	H-type	Mutations repair and RNA reprogramming by <i>trans-splicing</i> a target RNA (Fiskaa & Birgisdottir, 2010)
Group II introns	Two pseudoknots	Gene knockout using retrotransposition-activated selectable marker to target a desired DNA (Zhong <i>et al.</i> , 2003).

We use Enzymer (Zandi *et al.*, 2016), our previously developed software, for designing RNA sequences with desired secondary structures and nucleotide content. Notably, Enzymer can include tertiary contacts including pseudoknots in the design process. Enzymer uses the NUPACK package (Zadeh *et al.*, 2011a) to compute the partition function of a given RNA sequence using a particular energy model, from a possible set of models (Dirks & Pierce, 2003). The partition function captures the statistical characteristics

of individual RNA strands at thermodynamic equilibrium. Using the partition function, NUPACK derives the matrix of base pair probabilities, followed by prediction of the secondary structure of the RNA sequence. The available energy models are the Serra and Turner model (Serra & Turner, 1995), which we refer to as rna1995, and the Mathews model (Mathews et al., 1999), which we refer to as rna1999. In addition to these two energy models, NUPACK implements the Dirks energy model (Dirks & Pierce, 2003), which captures the energy parameters related to formation of pseudoknots. Enzymer implements a stochastic adaptive optimization algorithm to iteratively sample from the low ensemble defect (Zadeh et al., 2011b) mutational landscape of an initial RNA seed until a sequence with the desired predicted secondary structure attributes is sampled. Ensemble defect is a numerical value, which represents the average number of incorrectly paired nucleotides at thermodynamic equilibrium. Using the matrix of base pair probabilities generated by NUPACK, Enzymer computes the ensemble defect over the ensemble of all possible secondary structures including pseudoknots. The sequence design objective of Enzymer is to minimize the normalized ensemble defect of an RNA sequence, given a target secondary structure. Users can enforce sequence constraints by providing a design template as an input. The design template includes a set of nucleotides that may represent evolutionary conserved RNA motifs with known function. For instance, one can use a hammerhead motif and a corresponding consensus secondary structure inferred from multiple sequence alignment analyses (or from a structure confirmed by biochemical or crystal structure) and design sequences that are predicted to fold into the consensus secondary structure with low ensemble defect, while simultaneously preserving the hammerhead motif. It is expected that sequences with lower ensemble defect values and having the conserved functional motif, will show the desired functionality.

Regarding the pseudoknotted-ribozymes listed in **Table 4.1**, it should be noted that kissing-loops (such as those found in two of the hammerhead ribozymes from **Figure 4.2**, as well as in RNase P and group II introns) cannot be designed by Enzymer. To our knowledge, there are no algorithms that allow the computation of a partition function for kissing-loops, which Enzymer uses to design structures with optimized folding. There are however other tools available to do inverse folding of such pseudoknots by calculating their Minimum Free Energy (MFE), such as RNAiFold (Dotu et al., 2014), which uses a mix of partition function and MFE calculation to design structures.

4.5. Materials

Reagents

1. PCR mix: 200 μ M dNTPs, 1X Taq reaction buffer (20 mM Tris-HCl pH 8.8, 10 mM $(\text{NH}_4)_2\text{SO}_4$, 10 mM KCl, 2 mM MgSO_4 , 0.1% Triton X-100), Taq polymerase, 1 μ M of each primer and template DNA.
2. Radioactive transcription mix: 10 μ l DNA produced by PCR, rNTPs (2 mM each of rATP, rGTP, rCTP and 0.8 mM UTP), 1X transcription buffer (80 mM HEPES-KOH pH 7.5, 24 mM MgCl_2 , 40 mM DTT, 2 mM spermidine), 1 U/ μ l of inorganic pyrophosphatase, 1 U/ μ l RNases (A, B and C) inhibitor, 1 U/ μ l T7 RNA polymerase (produced in Lab) and top up to 50 μ l with milliQ water. For each reaction add 0.5 μ l of α -UT³²P (3000 Ci/mmol-10 mCi/ml) and 1 μ l DNase I (RNase free, 2000 U/ml).
3. Dephosphorylation mix: 1 U antarctic phosphatase, 1X buffer (50 mM bis tris propane-HCl pH 6, 1 mM MgCl_2 , 0.1 mM ZnCl_2) and 1U of RNases (A, B and C) inhibitor.
4. Labelling mix: 1 U of T4 polynucleotide kinase, 1 μ l γ -AT³²P (6000 Ci/mmol- 10 mCi/ml), 1X of the enzyme buffer (70 mM tris-HCl pH 7.6, 10 mM MgCl_2 , 5mM DTT) and 1U of RNases (A, B and C) inhibitor.
5. Cleavage reaction mix: 50 mM Tris-HCl pH 7.5, 100 mM NaCl and 25 mM KCl
6. Tris acetate EDTA (TAE) 50X: 2M Tris-HCl pH 7.6, 1 M acetic acid and 50 mM EDTA [28]
7. 6X native buffer: 40% sucrose, 0.05 % bromophenol blue and 0.05% xylene cyanol
8. Tris borate EDTA (TBE) 10X: 890 mM Tris, 890 mM boric acid and 20 mM EDTA [28]
9. Denaturing TBE: Urea (1922,32 g), TBE 10X (400 ml), top up to 4 L with water
10. Acrylamide: Urea (1922,32 g), TBE 10X (400 ml), 760 g Acrylamide, 40 g bis-acrylamide, top up to 4 L with water
11. Acrylamide gels: Acrylamide 20%, urea 8M, TBE 1X, 0.4 μ l/ml N,N,N',N'-Tetramethylethylenediamine (TEMED) and 8 μ l/ml ammonium persulfate (APS)
12. Agarose gel: 2 % Agarose, TAE 1X

13. 2X denaturing formamide buffer: 95% formamide, 10 mM EDTA, 0.025% bromophenol blue and 0.025% xylene cyanol

14. Elution buffer: EDTA 0.1 mM, NaCl 0.3 M, SDS 0.001%.

Equipment

1. Thermocycler
2. Spectrophotometer
3. Heating block
4. Electrophoresis racks
5. Typhoon or other radioactive scanning apparatus.

2.3 Software

1. Enzymer (and dependencies)
2. Image Quant
3. GraphPad

4.6. Methods

4.6.1. Installation and set-up of Enzymer

Enzymer is a python 2 programming language application, developed on OS X, tested on OS X and Ubuntu operating systems, and is available for download from a code repository at <https://bitbucket.org/casraz/enzymer/src/master/>. We have included the NUPACK package with Enzymer to facilitate the installation and configuration process. The instructions of how to install and use Enzymer assume the user has working knowledge of the UNIX terminal. This manual also assumes git (Anonyme), pip (Anonyme) and python 2.7 are installed on the user's computer and that the user has appropriate read-write privileges.

1. The first step in using Enzymer is to download it by cloning its code repository and then navigating to the root directory. In a bash terminal, run the following statement in bold :

```
bash-3.2$ git clone https://casraz@bitbucket.org/casraz/enzymer.git
Cloning into 'enzymer'...
remote: Counting objects: 584, done.
remote: Compressing objects: 100% (395/395), done.
remote: Total 584 (delta 81), reused 532 (delta 52)
Receiving objects: 100% (584/584), 10.53 MiB | 3.93 MiB/s, done.
Resolving deltas: 100% (81/81), done.
Checking connectivity... done.
```

2. Then verify that the download was successful:

```
bash-3.2$ ls -l
total 0
drwxr-xr-x 16 kasra staff 544 7 May 20:51 enzymer
```

1. Prior to navigating to the source code folder, it is necessary to prepare the python environment that the application will use to install its required dependencies. Start a python interpreter and make sure python 2 is run by default. Then call the exit() function to exit the interpreter:

```
bash-3.2$ python
Python 2.7.10 (default, Feb 7 2017, 00:08:15)
[GCC 4.2.1 Compatible Apple LLVM 8.0.0 (clang-800.0.34)] on darwin
Type "help", "copyright", "credits" or "license" for more information.
>>> exit()
```

2. You will have to create a python virtual environment (see **Note 1**), which will allow us to manage Enzymer dependencies. Run the following command:

```
bash-3.2$ virtualenv env_enzymer
New python executable in /Users/kasra/Desktop/env_enzymer/bin/python
Installing setuptools, pip, wheel...done.
```

3. The above command creates a folder for a python 2 virtual environment called *env_enzymer*. Validate the successful creation of the virtual environment.

```
bash-3.2$ ls -l
total 0
drwxr-xr-x  6 kasra  staff  204  7 May 21:07 env_zymer
drwxr-xr-x 16 kasra  staff  544  7 May 20:51 enzymer
```

4. To activate *env_zymer*, run the following command:

```
bash-3.2$ source env_zymer/bin/activate
(env_zymer) bash-3.2$
```

5. Notice that the bash command prompt has changed indicating that you have activated *env_zymer*. You are now ready to navigate to the right locations within the Enzymer folder to install Enzymer's required dependencies (see **Note**

```
(env_zymer) bash-3.2$ cd enzymer/
(env_zymer) bash-3.2$ pip install -r requirements.txt
Collecting numpy (from -r requirements.txt (line 1))
Downloading ...
Installing collected packages: numpy
Successfully installed numpy-1.16.3
```

4.6.2. Designing ribozymes using Enzymer

The entry point to Enzymer is the python file called *Enzymer.py*, which resides in the root folder of the application.

1. Run the following instruction to see the user manual:

```
(env_zymer) bash-3.2$ python Enzymer.py -h
usage: Enzymer.py [-h] [--input INPUT]

How to design an RNA secondary structure using Enzymer
```

2. To run Enzymer, one must provide an input file in which the required run-time parameters including the target secondary structure and the design template are provided. We have included two sample input files entitled *sample_input_with_pseudoknot.json* and *sample_input_without_pseudoknot.json*. Each of these input files contain a single python dictionary containing the required parameters. Users can create as many input files as desired and use each as many times as needed. However, only one input file can be used at a time. Here is an example of the input files:

```
(env_zymer) bash-3.2$ cat sample_input_with_pseudoknot.json
{
  "target_structure": ".....[[[...((([])]..)))]....",
  "design_template": "oooooooooooooooooooooooooooooooooooooooo",
  "seed": "",
  "target_normalized_ensemble_defect": "0.01",
  "design_temperature": "37",
  "max_iterations": "400",
  "dangles": "all",
  "energy_model": "rna1999",
  "NUPACK": "./nupack3.0"
}
```

The following is the definition of each parameter (i.e., called key in the python dictionary) of the input file:

1. target_structure: deals with the dot bracket notation of the target secondary structure with the following alphabet:
 1. single stranded position: “.”
 2. pseudoknot base pairs (opening and closing, respectively): “[“ and “]”
 3. regular base pairs (opening and closing, respectively): “(“ and “)”
2. design_template: the design template uses the alphabet A, U, G, C, o. Letter ‘o’ means that the corresponding position in the sequence can be any nucleotide, whereas the other letters imply the final sequence must contain these exact letters. The length of the design template must be equal to the length of the target structure.
3. seed: provides the means to start the design process from a given seed (a WT sequence for instance). If the provided seed is the empty string (i.e., “ ”) then a

random seed will be used. If the seed and the design template do not agree on a certain position, the nucleotide provided by the template will be chosen to conserve (i.e., keep constant) during the design process. If the value provided is not the empty string, then the length of this parameter must be equal to the length of the target structure.

4. `target_normalized_ensemble_defect`: normalized ensemble defect can be set to any floating point value between 0.00 and 1.00, inclusively. The optimization process will terminate when the desired sequence is found, that is one with the predicted normalized ensemble defect value.
5. `design_temperature`: the temperature in Celsius at which the energy parameters are used. The choice of this parameter depends on the choice of energy model. The energy model `rna1995` accepts temperatures between 0°C and 100°C, inclusively, while the `rna1999` energy model only allows for 37°C.
6. `max_iterations`: the number of sampling steps the optimization algorithm takes during the design process. We recommend that this parameter be set to 10 x size of the target structure. If the target ensemble defect is reached during the optimization process, the algorithm will terminate and return the results.
7. `dangles`: the dangling end treatment as described by NUPACK manual. The default is “all”.
8. `energy_model`: the energy models used for computation of the partition function. As mentioned, the options are `rna1995` and `rna1999`. The choice of energy model will have an impact on the `design_temperature`. Both energy models include the Dirks energy contributions related to the formation of pseudoknots.
9. `NUPACK`: the path to the NUPACK package relative to the root of the application. A copy of NUPACK is included with Enzymer and the path to it is: `“./nupack3.0”`. If during the first time you run Enzymer, errors related to NUPACK are thrown, we

recommend downloading NUPACK version 3.0.6 from www.nupack.org and replace the existing unpacked folder with the one downloaded. Then NUPACK package will require compilation. Call the make command inside the ./nupack folder to compile. The compilation process assumes a C compiler such as gcc is installed on user's workstation.

Note that all these parameters are mandatory: values must be provided by the user.

To run Enzymer using `sample_input_with_pseudoknot.json` as input file, enter the command line in bold:


```
(env_enzymmer) bash-3.2$ python Enzymer.py --input sample_input_with_pseudoknot.json
```

```
-----  
Input design parameters are:  
-----
```

```
{  
  'DANGLES': 'all',  
  'MATERIAL': 'rna1999',  
  'MAX_ITERATIONS': 400,  
  'NUPACK': './nupack3.0',  
  'SEED': '',  
  'STRUCTURE': '.....[[[...((([])]..)))]....',  
  'TARGET_NORMALIZED_ENSEMBLE_DEFECT': 0.01,  
  'TEMPERATURE': '37',  
  'TEMPLATE': 'oooooooooooooooooooooooooooooooooooooooooooo'  
}
```

```
Initializing Enzymer . . .
```

```
seed -> GGUAUAAGAUGAUGGCGGUACGGCUCGAAGCCGUUAG  seed GC ->  
0.540540540541
```

```
oooooooooooooooooooooooooooooooooooooooooooo  
oooooooooooooooooooooooooooooooooooooooooooo
```

```
iteration: 1  defect: 0.110726756757  mutations: 3  time: 07:13:05
```

```
sequence: GGUAUAAGAUGAGGGGCGUACAGCUGCAAGCUGAUAG
```

```
iteration: 2  defect: 0.110726756757  mutations: 1  time: 07:13:05
```

```
.
```

```
iteration: 25  defect: 0.0864916216216  mutations: 2  time:  
07:13:09
```

```
sequence: GGUAUAAGAUGAUGGGCGAACCGCUGCAAGCGGAAAG
```

```
-----  
target structure -> .....[[[...((([])]..)))]....
```

```
sequence -> GGUAUAAGAUGAUGGGCCAACCGCGGCAAGCGGAAAG
```

```
sequence normalized ensemble defect -> 0.00559189189189
```

```
free energy -> -12.800000000000000
```

```
sequence G/C content -> 0.567567567568
```

```
temperature -> -T 37
```

```
material -> rna1999
```

```
max iterations -> 400
```

```
used iterations 25
```

```
C_design in seconds 4.30981206894
```

```
C_eval in seconds 0.192348003387
```

```
C_design / C_eval 22.4063259979
```

Once the program starts, it displays relevant information about each iteration of the design process, such as the current normalized ensemble defect of the current RNA sequence,

the number of mutations that will be applied to the current sequence and the local time stamp. Notice that as the number of iterations grows, the normalized ensemble defect of the sampled RNA decreases. Iteration continues until either the target `normalized_ensemble_defect` value is reached or the maximum number of iterations (`max_iterations`) is exhausted. Once either one of the two termination criteria is met, Enzymer will print out the final RNA sequence together with relevant information such as folding energy and other performance measures. A summary of the run will be found in the `./results` folder in a file with the same as the input file but with a time stamp suffix attached to it.

4.6.3. *In vitro* testing of cis acting ribozymes

To prove the efficiency of Enzymer to design pseudoknotted RNAs by inverse folding, hammerhead ribozyme structure can be used as a model to generate sequences to assay and to compare their cleavage activity to the WT ribozyme. The cleavage activity of the cis acting ribozymes can be detected during transcription, but to determine the rate constant that gives a better idea on the ribozyme activity and the cleavage efficiency the use of trans acting version of ribozymes is preferred (Section 3.4).

✓ **DNA template production**

The first step to start with is the production of the DNA templates for the transcription to generate the RNAs.

1. Thaw the frozen reactants at room temperature and once thawed put them on ice.
2. Prepare the PCR master mix by putting the buffer, dNTPs and the water together; add the enzyme last (see Note 3) .
3. Depending on the DNA template, if the primer is the same for many DNA oligonucleotides, add the primers to the master mix, this can be the case for the forward primer consisting of the T7 promoter (like for the hammerhead ribozymes given as examples in **Table 4.2**). Otherwise add the primers individually before the next step.
4. Add 1 μ l of the DNA oligonucleotide used as a template (see Notes 4, 5 and 6).

- Run your PCR. For the hammerhead ribozyme sequences in **Table 4.2**, the T_m (annealing temperature) is 49°C since the T7 promoter sequence is used as a primer.

Table 4.2 Example of hammerhead sequences

Name	Sequence
T7	TAATACGACTCACTATAGG
HHRz 6	AATAGGTAGAGGTCCGACCGAGAAAAAGATCTCGGTTTCGCTCGTTTTTCG GCCTAGCTCATCAGGGACCGGGGGGCGTCCCTATAGTGAGTCGTATTA
HHRz 14	GGTGTGTTTTGGGTTCGACCGCCTTATCTGACCGGTTTCGCGTTTTTTTA GCCGTGCTCATCAGGACCCTTTTTGGCCATTCCCTATAGTGAGTCGTATTA

3.3.2 Agarose gel electrophoresis and DNA precipitation:

When the PCR is about to finish (10 to 15 min before it ends), prepare an agarose gel to verify the size of the PCR products

- Prepare a 2% agarose gel (assuming PCR products <150 bp) and run 5 µl of each of your samples.
- Visualise the bands on the gel (in the GelDoc under UV light) to verify the PCR worked.

Precipitate the remainder of the PCR products with ethanol or purify them using a column to get better results for transcription.

✓ **Radioactive transcription**

The cis-acting ribozymes will normally cleave during transcription, to be able to visualise and quantify the cleavage result; radioactive uridine triphosphate (UT³²P) is used (see Note 7).

- Thaw the frozen reactants at room temperature then put them on ice.
- Prepare the radioactive transcription mix for the transcription (except UT³²P and DNase I) and add the enzyme last.
- Add UT³²P to the mix with the amount of 0.5 µl for each reaction and mix gently to avoid contaminating the pipet.

4. For each reaction add 10 μl of the PCR product from section 3.3.2 step 3 of the corresponding ribozyme.
5. Incubate at 37°C for 2 to 3 hours.

Add 1 μl of DNase I (RNase free) and incubate for 30 min at 37°C.

✓ **RNA precipitation**

1. Each reaction from section 3.3.3 is precipitated with 0.1 volume of NaOAc 3 M and 2.5 volume of 100% ethanol for 30 min at -80°C.
2. Centrifuge for 30 min at 14000 rpm in a microcentrifuge at 4°C.
3. Discard the supernatant in the radioactive liquid trash and wash the pellet with 500 μl 70% ethanol.
4. Centrifuge for 10 min at 14000 rpm in a microcentrifuge at 4°C.
5. Discard the supernatant and let the pellets dry behind a plexiglass shield (see Note 8).
6. Add 20 μl of sterilised milliQ water (RNase free water).

Keep the RNAs at -20°C for a long time use (see Note 9).

✓ **Cleavage reaction detection on polyacrylamide gel**

1. Prepare the glass plates with the spacers and a comb with a number of well corresponding to the samples number. Clean them with water and 70% ethanol. The smaller glass plate should be treated with a silicon water repellent, that prevents the gel from sticking to the plate (see Note 10).
2. Prepare a 10% polyacrylamide gel by mixing an equal volume of 20% denaturing acrylamide with 8 M TBE 1X urea. The polymerisation reaction is catalysed by the addition of 0.4 μl of TEMED and 8 μl 10% APS/ml acrylamide (see Notes 11 and 12).
3. Let the gel polymerise and keep a small amount of the liquid in a beaker to be able to know when the gel is solidified.

4. Put the gel in the electrophoresis rack and fill with 1 X TBE. The TBE level should be high enough to cover the wells.
5. For the RNA samples from section 3.3.4, take 2 μ l of each RNA and dilute it in 8 μ l of water and 10 μ l of 2X denaturing formamide buffer. This is done because we are using a transcription product that is usually highly concentrated in salt and free ribonucleotides which makes the RNA bands distort and make them less sharp on the gel. So to have clean bands on the gel it is better to dilute the RNA (unless the gel is meant for RNA purification). Load the samples on the gel and run it.
6. Take off the small glass plate and cover the gel with two layers of plastic film (e.g. saran wrap).
7. Expose the gel to a clean phosphor screen for 1 hour, if this is not enough expose it for a longer time (typically overnight).

For visualisation, we use the PhosphorImager Typhoon FLA 9500.

✓ ***In vitro* testing of trans acting ribozymes**

To determine the rate constant (k_{obs}) for each ribozyme, the cleavage activity is followed during a time course by taking aliquots from the reaction mix. Since the ribozymes cleave during transcription it is difficult to get enough of non-cleaved RNA after transcription of the cis-acting version. To solve this problem each ribozyme sequence can be divided in two parts: 1-the ribozyme part with the catalytic core, stem II, the 5' side of stem I and 3' side of stem III; 2-the substrate with the cleavage site as well as the 3' side of stem I and 5' side of stem III (**Figure 4.3**).

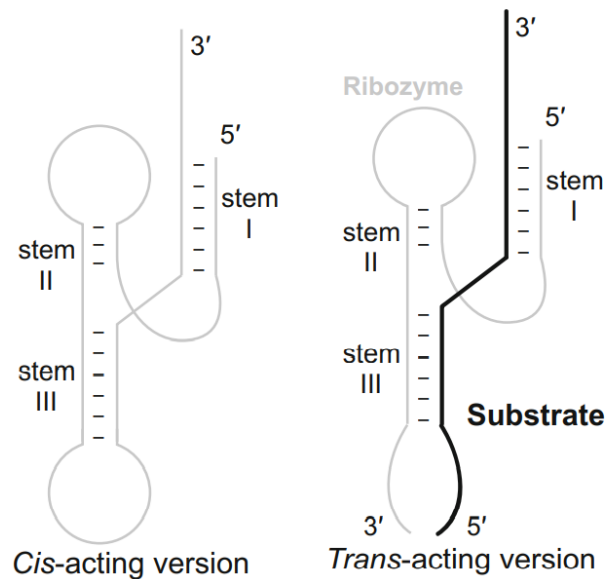


Figure 4.3 Ribozyme in cis and trans acting versions

To have a *trans* acting version, the ribozyme sequence is divided into a substrate part and a ribozyme part without affecting the ribozyme structure by opening the loop of stem III. Grey, ribozyme part; black, substrate part.

✓ **Non-radioactive transcription**

For the trans acting ribozymes and their substrates, cold RNAs are produced using the same protocol as in sections 3.3.1 to 3.3.3 without the use of radioactive $UT^{32}P$ in section 3.3.3. Instead, for the radioactive transcription mix, remove $UT^{32}P$ and use 2 mM of UTP in the rNTP mix instead of 0.8 mM.

✓ **RNA extraction**

1. For each sample from section 3.4.1 a phenol/chloroform extraction is performed before loading on gel. Add 50 μ l of phenol-chloroform mix to the resulting transcription and vortex for 1 minute. Centrifuge at 14000 rpm in a microcentrifuge for 2 minutes.
2. Replace the aqueous (top) phase in a new microcentrifuge tube.
3. Precipitate RNA as in section 3.3.4.
4. Let the pellet dry.

5. During the time the pellet is drying, prepare a 10% polyacrylamide gel as in Section 3.3.5 steps 1-4.
 6. When the gel is polymerized, dilute each RNA in 10 μ l milliQ water and 10 μ l 2X denaturing formamide buffer. Load the RNAs on the gel and run it
 7. Take off the glass plates and put the gel in plastic wrap that allows the passage of UV light (e.g. saran wrap).
 8. Use UV shadowing to visualize the RNA bands. The xylene cyanol and bromophenol blue can be used as weight markers depending on the gel size (**Figure 4.6**).
 9. Mark the bands and cut them with a sterilized razor blade.
 10. To elute the RNA put each band in a microcentrifuge tube in presence of 400 μ l of elution buffer (see Note 13).
 11. Put the tubes on a rotator for 1 to 2 hours at room temperature or overnight at 4°C.
- Recover the elution buffer and precipitate the RNAs as in section 3.3.4.

✓ **Substrate labelling**

To be able to visualise the cleavage reaction, the substrate strand should be radioactively labelled.

1. Incubate each substrate from section 3.4.2 in the dephosphorylation mix at 37°C for 20 minutes. To inactivate the enzyme put the mix at 65°C for 5 minutes.
2. From the dephosphorylation reaction take ~5 to 100 pmoles and incubate it with the labelling mix and 1 μ l of AT³²P (6000Ci/mmol- 10mCi/ml) for 1 hour at 37°C.
3. Purify the RNA substrates, to which you added 1 volume of denaturing formamide buffer, on a 10% acrylamide gel prepared as in Section 3.3.5 steps 1-4 (after running the gel, the TBE used to run the gel should be thrown in the radioactive liquid waste).
4. In order to localise the bands correctly on the gel, radioactive markers can be used. The markers we use are radioactive spots we make using dye mixed in an almost

empty bottle of radioactivity source. You can put pieces of Whatman paper in the four corners of the gel (uncovered from the upper glass plate) and put few microliters of the radioactive dye on each spot (**Figure 4.4**). To prevent contamination of the phosphor screen two layers of plastic wrap are used to cover the gel and the screen is exposed for 5 to 10 minutes.

5. After scanning the phosphor screen the bands are cleaved using the real-scale print out (or film) of the gel.

The radioactive RNA are eluted and precipitated as in section 3.4.2 steps 9-11 (see Note 14).

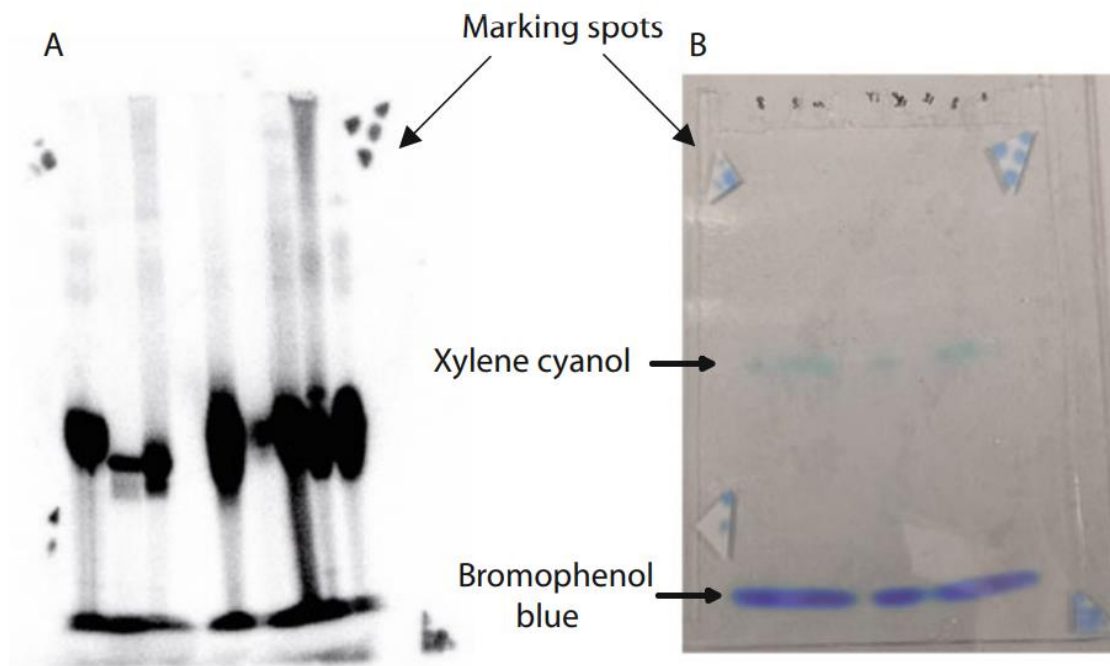


Figure 4.4 Visualization of radioactive bands for purification

- (a) Image of the gel scanned with the phosphorimager. (b) Real gel photo with the markers.

✓ **Kinetics measurements**

1. Pre-incubate each ribozyme reaction mix at 85°C for 1 min to denature completely the RNA then put it immediately on ice so that it folds randomly or in a stable structure

that is not optimal. When the RNA is incubated at 37°C with Mg²⁺, it will fold correctly.

2. Add the substrate and take the first aliquot (T0) of the reaction just before adding the Mg²⁺.
3. The moment the Mg²⁺ is added, start the timer.
4. Prepare all the tubes you need to put your aliquots in advance with 10 µl of denaturing formamide buffer. At the desired time (see Notes 15, 16 and 17) take your aliquot and mix it with the formamide buffer to stop the reaction and put the sample on ice.
5. Run the samples on a 10% acrylamide gel as in Section 3.3.5, except that samples already contain denaturing formamide buffer.
6. Expose the phosphor screen to the gel overnight.

Scan the screen and calculate the kobs by using the ImageQuant software to measure the intensity of the bands, then the calculations can be done on excel. Plot your data on a curve using the 'one phase decay' equation ($Y = (Y_0 - \text{Plateau}) * e^{-K * X} + \text{Plateau}$) from the enzyme kinetic plots in GraphPad.

4.7. Notes

1. Virtual environments are used to manage software dependencies. It allows the external python libraries that Enzymer is relying on to be managed in isolation of the Python packages that already exist.
2. Trouble shooting installation: Make sure the python version use is correct. If you receive an error message related to installation of a package called numpy, try to install the package by running:

2.

```
(env_zymer) bash-3.2$ pip install numpy
```

```
Collecting numpy
```

```
Using cached
```

```
https://files.pythonhosted.org/packages/6e/36/e8369aa628b29f50211ba82daec31cc110f6627feca160bc11b0e4ee1191/numpy-1.16.3-cp27-cp27m-macosx_10_6_intel.macosx_10_9_intel.macosx_10_9_x86_64.macosx_10_10_intel.macosx_10_10_x86_64.whl
```

```
Installing collected packages: numpy
```

3. We have found that addition of 1X Q solution (Qiagen) to the reaction mixture can help for amplifications that fail on the first attempt.
4. While 0.2 μM is often enough to get a PCR amplicon for analytical purpose, to prepare DNA templates for transcription it is preferable to use 1 μM .
5. For each PCR a negative control with the primers only and without a template is useful to ensure that the amplification is occurring correctly. If only oligonucleotides are used as template (see Note 6), the negative control can be a PCR mix with all the oligonucleotides, but with no enzyme.
6. Assembly PCR: If we need to amplify a long ribozyme sequence for which we do not have the full length template, two or four oligonucleotides with overlapping sequence between each, can be used to make a PCR assembly (**Figure. 4.5**).

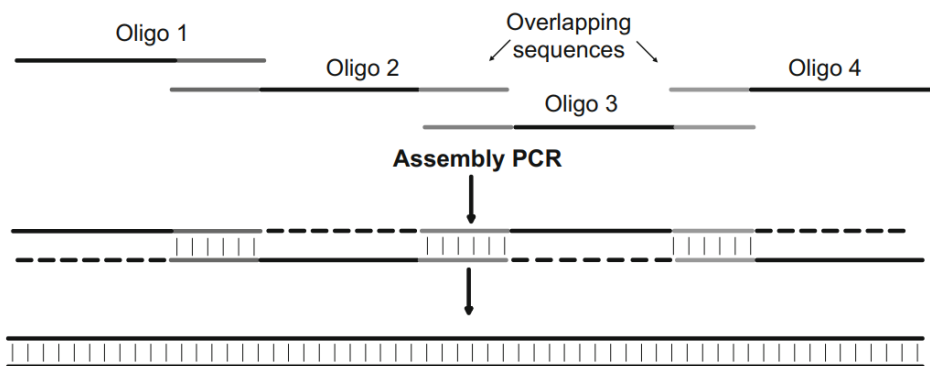


Figure 4.5 Assembly PCR

Oligonucleotides with overlapping sequences can be used by the polymerase as primers.

7. The ribozymes are usually highly active in presence of Mg^{2+} , which is present in the transcription buffer at 24 mM. If you need to get the full length ribozyme after transcription (to do kinetic assays for instance) and lessen the self-cleavage reaction occurring during transcription, a buffer with a smaller amount of Mg^{2+} can be used. This is also something to take in consideration when doing all the experiments with ribozymes. The use of trans acting versions circumvents this problem by allowing to prepare RNA and resuspend it in water without Mg^{2+} , this is discussed in more detail in Section 3.4.

8. After centrifugation of radioactive RNA, use a Geiger counter to ensure that radioactivity is in the pellet and not in the supernatant. If you have more radioactivity in the supernatant centrifuge the RNA for a longer time.

9. Unlike the common ideas about RNA sensitivity to degradation, when the RNA is not contaminated with RNase it is stable at $-20^{\circ}C$ for a long period of time, especially if it is not frequently exposed to cycles of freezing and thawing. It is however preferable to keep it on ice when it is taken outside the freezer.

10. The product we use in the laboratory to prepare the glass plates for the polyacrylamide gels is RAINX, which is a product sold for car windshield and is as efficient as the products bought from specialized vendors for our purpose.

11. For polyacrylamide gels, a sequencing gel size of 33 X 39 cm would be good enough to separate the RNA fragments to have single nucleotide resolution up to 100 bases after running it for 2 to 3 hours

12. The polyacrylamide gel percentage depends on the RNA size so the bigger the RNA size is, the lower the acrylamide percentage should be, to make it easier for the RNA to move in the gel. As an example, the cis acting Hammerhead ribozymes we used had a length of 82 nt so we used a 10% gel (**Figure 4.6**) (Sambrook, 1989).

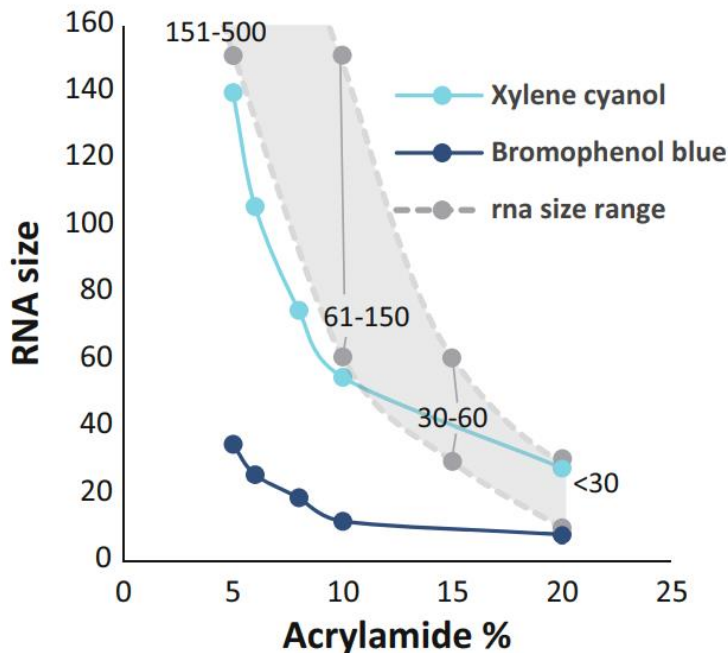


Figure 4.6 Size range of markers and RNA in PAGE

The bromophenol blue and xylene cyanol can be used as approximate weight markers of RNA considering the acrylamide percentage. For example in a 10% polyacrylamide gel, an RNA of 61 or more nucleotides should be above the xylene cyanol. The recommended size range of RNA for each gel % is indicated in gray. For 20% gel, it is recommended to purify RNA with a size range between 10-30 nt so that it will not be lower than the bromophenol blue (~8 nt).

13. For the RNA elution from the gel, depending on gel thickness, it might be good to cut the gel in pieces or crush it for better elution. For gels of 0.8 mm or less and RNAs of 100 bases or less, 1 h should be enough to recover most of the RNA from the gel.

14. Each time the centrifuge is used for the precipitation of radioactive samples, it should be checked for contamination. Usually this is true for all the material used when manipulating radioactivity, after finishing the experiment everything should be checked for contamination.

15. Ribozymes Kinetics: For relatively fast cleaving ribozymes the times used to stop the reaction could be for example 30s, 1 min, 2 min, 4 min, 8 min, 16 min, 32 min and 64 min. For slow cleaving ribozymes the reaction can be followed for 24 hours by leaving the reaction overnight (see Note 16).

16. For the overnight reactions, a bigger volume of the reaction should be prepared to prevent the evaporation and the condensation of the liquid in the tube lid. A heating lid or mineral oil can also be used.

17. To follow several ribozyme kinetics, either use an excel sheet as a time table or use one timer for each ribozyme. Here is an example of such a time sheet (**Table 4.3**).

Table 4.3 Example of a Time table

Time	Rz1	Rz2	Rz3
00:00	Stop t0		
00:30	Stop t1		
1:00	Stop t2		
2:00	Stop t3		
4:00	Stop t4		
4:30		Stop t0	
5:00		Stop t1	
5:30		Stop t2	
6:00			
6:30		Stop t3	
7:00			
8:00	Stop t5		
8:30		Stop t4	
9:00			Stop t0
9:30			Stop t1
10:00			Stop t2
11:00			Stop t3
12:00			
14:00			Stop t4
15:00			
16:00	Stop t6		
16:30		Stop t5	

18. For complex RNA structures Enzymer may be able to make the design of a new sequence, but there is a high chance that its activity will be much lower than the wild-type. In previous work GlmS ribozyme designs were made by Enzymer but the best ribozyme obtained had 50% less activity (Sabrine Najeh, 2019). It is usually possible to choose a small ribozyme for a given application (see Table 1 and Note 19 for additional comments on applications).

19. In addition to the examples referenced in **Table 4.1**, as an example of application, a cis acting hammerhead ribozyme can be used to control the expression of a downstream gene. This can be done by inserting the ribozyme sequence so that it overlaps the coding sequence of the gene. Two conditions should be respected in this case: the amino acid sequence must be the same, even if we change the nucleotides sequence, and the catalytic core and the structure of the ribozyme needs to be conserved. In this context Enzymer, as an inverse folding tool, can help design a ribozyme with the desired criteria. Similarly, we can foresee that different ncRNA functions could be combined, such as a riboswitch and a ribozyme to make allosteric ribozymes. The ability to design sequences from scratch based on known structures provides multiple opportunities for synthetic biology.

5. Article 2: Computational design and experimental verification of pseudoknotted ribozymes

5.1. Mise en contexte

Dans l'article précédent nous avons détaillé la méthodologie utilisée pour la conception ainsi que la validation expérimentale de l'activité de ribozymes synthétiques ayant des pseudonœuds. Dans le présent manuscrit, nous présentons la confirmation expérimentale, via l'évaluation de l'activité, de l'adoption de la bonne structure par les ribozymes conçus par Enzymer. Les résultats de cet article démontrent l'efficacité de l'algorithme de conception de ribozyme du programme Enzymer décrit dans notre premier article de méthodologie. C'est un article qui propose un programme pour la génération, pas seulement des ARN ayant un seul pseudonœud, mais de plusieurs, ce qui a été validé avec le ribozyme GImS qui a trois pseudonœuds.

Le programme Enzymer permet la conception de séquences de ribozymes (ou autres structures d'ARN) très divergentes de la séquence originale, ce qui permet d'incorporer d'autres contraintes de séquence. Par exemple, nous avons fusionné un ribozyme pour qu'il chevauche en grande partie une séquence codante et on peut envisager, dans l'optique de la mise au point de portes logiques par exemple, que cela pourra être utile pour la construction de vecteurs qui seront destinés au développement de systèmes « intelligents » pour le diagnostic ou la thérapie de cancer. La particularité d'Enzymer est de tenir compte des pseudonœuds permettant de s'approcher un peu plus des modèles de structures de nombreux ARN.

Computational design and experimental verification of pseudoknotted ribozymes

Conception computationnelle et vérification expérimentale de ribozymes à pseudonœud

Sabrina Najeh^{1*}, Kasra Zandi^{2*}, Nawwaf Kharma³ and Jonathan Perreault^{1#}

¹INRS - Institut Armand-Frappier, 531 boulevard des Prairies, Laval, QC, H7V 1B7, Canada

²Software Engineering & Computer Science Dept., Concordia University, 1455 boul. de Maisonneuve O., Montreal, QC, H3G 1M8, Canada

³Electrical & Computer Eng. Dept., Concordia University, 1455 boul. de Maisonneuve O., Montreal, QC, H3G 1M8, Canada

*authors contributed equally as indicated in RNA

#To whom correspondence should be addressed. Tel: +450 687 5010 (Ext 4411); Fax: +450 686 5301; Email: jonathan.perreault@iaf.inrs.ca

État du manuscrit : Ce manuscrit a été accepté le 27 mai dans le journal RNA, mais initialement rejeté dans sa version originale. La version ici fait suite aux commentaires et suggestions des réviseurs et la resoumission a été faite en date de 2 Mars 2022 au même journal (MS ID#:RNA/2022/079148).

5.2. Contributions des auteurs

Sabrine Najeh : mise en place du plan expérimentale avec Jonathan Perreault et réalisation de toutes les expériences présentées dans l'article. Analyse et interprétation des résultats de même que la rédaction de toutes les parties en relation avec la validation expérimentale et la discussion.

Kasra Zandi : Développement de 'Enzymer' qui était publié en 2016, rédaction de la partie en relation avec l'algorithme, design des ribozymes testés dans ces travaux.

Nawwaf Kharma : Supervision du travail de Kasra Zandi et révision du manuscrit.

Jonathan Perreault : Supervision et assistance du travail de Sabrine Najeh et Kasra Zandi par la participation dans l'interprétation des résultats et révision du manuscrit.

5.3. Abstract

The design of new RNA sequences that retain the function of a model RNA structure is a challenge in bioinformatics because of the structural complexity of these molecules. RNA can fold into secondary and tertiary structures by forming stem loops and pseudoknots. A pseudoknot is a set of base pairs between a region within a stem-loop and nucleotides outside of this stem-loop; this motif is very important for numerous functional structures. It is important for any computational design algorithm to take into account these interactions to give a reliable result for any structures that include pseudoknots. In our study, we experimentally validated synthetic ribozymes designed by Enzymer which implements algorithms allowing for the design of pseudoknots. The ribozymes tested were the hammerhead and the glmS which have a self-cleaving activity that allows them to liberate the new RNA genome copy during rolling-circle replication or to control the expression of the downstream genes, respectively. We demonstrated the efficiency of Enzymer by showing that the pseudoknotted hammerhead and glmS ribozymes sequences it designed were extensively modified compared to wild-type sequences and were still active.

Keywords: RNA inverse folding, pseudoknot, secondary structure, hammerhead ribozyme, glmS ribozyme

Résumé

La conception de nouvelles séquences d'ARN qui conservent la fonction d'une structure d'un ARN modèle est un défi en bioinformatique, en raison de la complexité structurelle de ces molécules. L'ARN peut se replier en structures secondaires et tertiaires en formant des tiges-boucles et des pseudonœuds. Un pseudonœud est un appariement de bases entre une région à l'intérieur d'une tige-boucle et des nucléotides à l'extérieur de cette tige-boucle; ce motif est très important pour de nombreuses structures fonctionnelles. Il est important que tout algorithme de conception informatique tienne compte de ces interactions pour donner un résultat fiable pour toutes les structures qui incluent des pseudonœuds. Dans notre étude, nous avons validé expérimentalement des ribozymes synthétiques conçus par Enzymer qui implémente des algorithmes permettant la conception de pseudonœuds. Les ribozymes sont des ARN catalytiques qui ont des activités similaires à celles des enzymes. Les ribozymes testés étaient le ribozyme en tête de marteau et le ribozyme glmS, qui ont, une activité d'auto-clivage qui leur permet de libérer la nouvelle copie du génome d'ARN lors de la réplication circulaire ou de contrôler l'expression des gènes en aval, respectivement. Nous avons démontré l'efficacité d'Enzymer en montrant que les séquences conçues avec des pseudonœuds, de ribozymes en tête de marteau et glmS, étaient considérablement modifiées par rapport aux séquences WT et étaient actives.

5.4. Introduction

Natural functions of noncoding RNAs (ncRNAs) are numerous, but they also have various applications in engineering biological systems (Liang *et al.*, 2011);(Cameron *et al.*, 2014)(Auslander *et al.*, 2010; Prommana, 2013; Zalatan *et al.*, 2015), therapeutics (Ruder, 2011) (Lienert *et al.*, 2014) (Esposito *et al.*, 2014) as well as in nano-technology (Afonin *et al.*, 2015)(Grabow & Jaeger, 2014; Shu *et al.*, 2015). Some ncRNAs with complex structures, such as riboswitches that act as receptors binding specific metabolites, are involved in controlling gene expression (Serganov & Nudler, 2013). Other intricate RNA structures can confer catalytic activities to RNA, like for the small self-cleaving ribozymes. The best known example is the hammerhead ribozyme (HHRz), which is involved in producing single-copy genomes out of the multimeric RNA resulting from rolling circle replication of viroids (Prody *et al.*, 1986). There are also hundreds of examples of different hammerhead-type ribozymes for which RNA self-cleavage still has no obvious function (Hammann *et al.*, 2012). The core of the so-called minimal version of the hammerhead ribozyme has been studied extensively over more than three decades, atomic resolution structures are available (Martick & Scott, 2006), as well as extensive biochemical data strongly supporting the requirement of the core consensus sequence (Nelson & Uhlenbeck, 2008). In the case of the hammerhead ribozyme, a tertiary interaction between stem I and II was found to have a critical impact on its activity (De la Pena *et al.*, 2003; Khvorova *et al.*, 2003). Different configurations of pseudoknots were later found to hold stem I and II to achieve the same impact and also make the ribozymes over 1000 fold more active than a version without this interaction (Perreault *et al.* 2011). Another example of ribozyme, the glmS ribozyme, has a clear regulatory function. This ribozyme cleaves its 5'-UTR (UnTranslated Region) by using glucosamine-6-phosphate (GlcN6P) as a cofactor, leading to degradation of the mRNA and repression of genes involved in GlcN6P synthesis when the latter is in sufficient concentrations (Collins *et al.*, 2007; Winkler *et al.*, 2004). While not studied as extensively as the hammerhead ribozyme, the archetypal model for RNA structure/function studies, several atomic resolution structures of glmS ribozyme are also available (Cochrane *et al.*, 2009; Klein, 2006), as well as a wealth of experimental data (Ramesh & Winkler, 2014).

It is widely accepted that the function of ncRNAs is attributed to their structure (Leontis *et al.*, 2006; Mortimer, 2014), which is determined by the nucleotide composition of the RNA polymer. An RNA strand folds to form an array of secondary structures which then form

stable tertiary structures (Batey *et al.*, 1999). The diverse range of functions of ncRNAs, as well as the relationship between their sequence-structure and function, highlight the importance of methods for analysis and design of ncRNAs with desired structural attributes.

Formation of the secondary structure is the first step in RNA folding. It starts by forming the stem-loops before getting to the formation of pseudoknots to have the final tertiary structure (Batey *et al.*, 1999). For an RNA sequence, a secondary structure is defined by a set of base-pairs (ϕ_i, ϕ_j) where $1 < i < j < n$, such that positions i and j are paired. For two base-pairs (ϕ_i, ϕ_j) and (ϕ_k, ϕ_l) , a non-nested loop or a pseudoknot forms if either of the nesting rules $i \leq k \leq j \leq l$ or $k \leq i \leq j \leq l$ is violated. Pseudoknots are abundant in nature, (Zandi *et al.*, 2016) exist in specific types (Condon, 2004) and are known to play key roles in the functionality of active RNAs (Staple & Butcher, 2005), including rRNAs (Powers & Noller, 1991), riboswitches (Gilbert *et al.*, 2008) and ribozymes (Harris *et al.*, 2015). Experimental methods, such as nuclear magnetic resonance (NMR) spectroscopy (Varani & Tinoco, 1991) or X-Ray crystallography (Muchmore *et al.*, 1996), which are used for determining RNA structure, are complex and time-consuming. Hence, computational approaches provide an attractive alternative to study and analyze RNA structures. In addition, computational methods can generate thousands of RNA sequences in short amounts of time, at relatively low cost, while providing the means to predict key sequence attributes useful for further analysis and experimentation.

The classical definition of computational RNA structure prediction or RNA folding (Nussinov & Jacobson, 1980) is to find the set of base-pairs which are predicted to exist in RNA's most stable structure called the minimum free energy (MFE) structure at thermodynamic equilibrium. The MFE structure can be predicted by computational methods using some energy model. Over the past few decades, several energy models (Freier *et al.*, 1986) (Serra & Turner, 1995) ;(Dirks & Pierce, 2003; Mathews *et al.*, 1999) as well as computational structure prediction methods such as RNAfold (Hofacker, 2003b), mfold (Zuker, 2003a), HotKnots (Ren *et al.*, 2005), RNAstructure (Reuter & Mathews, 2010), IPknot (Sato, 2011), pKiss (Theis, 2010) and NUPACK-analyze (Zadeh *et al.*, 2011a) have been developed. On the other hand, the RNA inverse folding problem is to find a sequence with an MFE structure that precisely matches a desired target structure. Besides the classical MFE as the design criteria, other criteria such as Boltzmann probability or ensemble defect optimization criteria have also been used. In the

context of RNA inverse folding, Boltzmann probability quantifies the probability of an RNA sequence folding into a given structure, where the ensemble defect quantifies the expected number of incorrectly paired nucleotides at thermodynamic equilibrium. The value of ensemble defect should be minimized as much as possible to avoid misfoldings in double and single stranded regions of structure (Dirks *et al.*, 2004). It has been shown that the sequences designed to maximize the Boltzmann probability or minimize ensemble defect tend to be thermodynamically more stable than those designed to satisfy the MFE criteria (Zadeh *et al.*, 2011a; Zandi *et al.*, 2016).

The RNA design problem is computationally complex (Schnall-Levin, 2008) and to find solutions, most existing algorithms resort to heuristics and a combination of local, global and stochastic search methods. Generally, a random seed RNA sequence is first generated, then the seed is iteratively mutated until the predicted folding attributes of the design candidate converges to the desired values. Our recent work, Enzymer (Zandi *et al.*, 2016) utilizes an adaptive weighted sampling strategy to design RNA secondary structure with low ensemble defect. The *in silico* simulations showed that Enzymer generates RNAs that are thermodynamically more stable, have higher Boltzmann probability of folding into the desired target and have lower ensemble defect than those generated by other state of the art pseudoknot designer methods such as MODENA (Taneda, 2011) and antaRNA (Kleinkauf, 2015).

The wealth of existing computational RNA secondary structure design methods, where each method utilizes a different design criterium and sequence optimization strategy, makes it a difficult task for the experimental biologists to choose the right method for their particular RNA design purpose. For instance, some methods such as NanoTiler (Bindewald, 2008) or NUPACK-design are more suited for the design of non-functional RNA nano structures, while other methods such as Frnakenstein are specialized in designing RNA switches (Lyngso *et al.*, 2012). Another decisive factor in the applicability of an RNA designer method is its ability to realize all the structural elements necessary for a particular design objective. For instance, only NanoTiler (Bindewald, 2008), MODENA, antaRNA and Enzymer can handle the pseudoknots. Another important consideration is related to the applicability of the underlying energy model used by each design method. Still another important consideration is the availability of experimental evidence in support of the applicability of a design method for a specific design purpose. Ultimately, it is wet-lab experimental data that provide the most reliable measure of the usefulness of a design

method. To the best of our knowledge, as summarized by (Churkin, 2017), there seems to be no comprehensive report in the literature that provides experimental evidence on the applicability of all inverse RNA folding methods to the design of pseudoknotted ncRNAs.

In this study, we demonstrate that Enzymer can be used as a reliable method for the design of pseudoknotted ribozymes. We used Enzymer to reengineer three naturally occurring ribozymes: a self-cleaving HHRz from a mouse gut metagenome, a self-cleaving HHRz of *Yarrowia lipolytica* (Perreault, 2011) as well as a self-cleaving glmS ribozyme (Klein, 2006). For each ribozyme, we obtained the minimal required catalytic core extracted from multiple sequence alignment data and used the catalytic core as a design template for Enzymer.

5.5. Results

5.5.1. Synthetic hammerhead ribozymes with a pseudoknot for the stem I-II interaction

The activity of an RNA depends on its folding in defined secondary and tertiary structures (Bhartiya & Scaria, 2016). To have a functional synthetic RNA, the structure of the original RNA used as template should be conserved. In our study, we used Enzymer to design sequences that fit the functional structures of given templates. To validate the efficiency of Enzymer's designed sequences, we tested experimentally two types of self-cleaving ribozymes; HHRz and glmS ribozymes.

Numerous HHRz were previously designed "by hand" by many researchers, where for typical minimal HHRzs, all three stems had four or more base pairs, with few exceptions (Kuwabara *et al.*, 1999). It was noted that cleavage rate measurement was performed with 25 mM Mg²⁺. On the other hand, we have chosen, on purpose, more challenging target models, i.e. with a stem II of only two canonical base pairs. Two HHRz structures from different genomes and context were used as templates: one from the mouse gut metagenome and the other from *Yarrowia lipolytica* (Perreault *et al.*, 2011b). Both of them are type I hammerhead with a pseudoknot occurring between stem I and the loop of stem II. The generated ribozymes should fold in the same secondary structure as the chosen ribozymes by respecting the position of canonical base pairs and single stranded regions. Enzymer does not change the size of the sequences generated as they should be exactly the same as the template sequence. The two base-pair stem II of the mouse gut metagenome HHRz (**Fig. 5.1A, B**), makes it a good model to test Enzymer since

misfolding could easily prevent stem II from folding correctly following a faulty design. Using the secondary structure of this HHRz (**Fig. 5.1A, B**), we selected the 14 sequences with the lowest ensemble defects. At least two such sequences were selected among ten designed per set of differing parameters, namely different optimal temperatures for folding, which we thought of as a way to have structures with a gradient of stability. In this way, HHRz 1 to 6 were set for optimal folding at 37°C, HHRz 13 and 14 at 17°C, 11 and 12 at 27°C, 7 and 8 at 47°C and 9 and 10 at 57°C. For self-cleavage tests to compare their activity to the wild type, all ribozymes were assayed during transcription at 37°C.

96%. The rest of them have variable end-point cleavage efficiency similar or slightly lower to the wild type, except ribozyme 1 that clearly cleaved less (56%) (**Fig. 5.1C,D**). For more precise measurements of the activity of the ribozymes, we followed the kinetics of each one of them during 64 minutes. For these assays, *trans*-acting ribozymes were used to facilitate measures of rate constants independently of transcription. The ribozyme sequences were divided into two parts: the template and the ribozyme part (**Table 5.2** and **Fig. 5.5**). The results show that all the ribozymes were able to cleave the substrate, but with different rates and efficiencies. The k_{obs} (rate constant) of the wild type ribozyme was the highest (3 min^{-1}) followed by HHRz 13 (1.6 min^{-1}), HHRz 11 (0.54 min^{-1}) and HHRz 12 (0.54 min^{-1}). Most of the ribozymes had cleavage rates 10 to 100 fold slower (ranging roughly from 0.03 to 0.3 min^{-1}) with no apparent correlation with the end-point cleavage (full curves in supplementary **Fig. 5.6**).

Interestingly, WT HHRz was the fastest cleaving ribozyme, although HHRz 13 has a very similar activity, other ribozymes cleave much more slowly, even hundreds of fold more slowly. This contrasts with the computed ensemble defects (**Table 5.1**) which are optimized for the designed ribozymes, and thus very low, as opposed to the WT for which the calculated ensemble defect was higher than most designed ribozymes.

In addition to these results and to prove that ribozymes are able to fold in the correct tertiary structure, HHRz 13 activity was compared to a mutated version (disrupted pseudoknot) and a version where the mutations were compensated. As we expected, the mutated version lost the cleavage activity by about 20 fold ($k_{\text{obs}}=0.07 \text{ min}^{-1}$) and when we compensated the mutation, the activity is recovered ($k_{\text{obs}}=1.5 \text{ min}^{-1}$). However, cleavage efficiency or maximum fraction of substrate cleaved is not entirely restored, likely due to the fact that contrary to the original sequence designed by Enzymer, this one is not optimized. For another test of a ribozyme generated by Enzymer, but without the iterations that lead to optimized ensemble defect and thus named “Non-Optimized” (N/O), there was less cleavage detected than most other designs and it had the smallest k_{obs} , $2 \times 10^{-3} \text{ min}^{-1}$, which is 100 fold less than the average of the other 11 designs for the same model (**Fig. 5.1E, Table 5.3 and Fig. 5.6**). This indicates that, with this model-ribozyme, ensemble defect optimization was important to obtain efficient ribozymes that have high k_{obs} and high fraction cleaved, as without this optimization, even if all base pairing rules were followed, the cleavage rate was a few orders of magnitude lower. Moreover, the much slower HHRz 13 with the mutant pseudoknot, compared to its compensatory mutant, which cleaves

almost as fast as the “WT HHRz 13”, support that Enzymer was able to design active hammerhead ribozymes using the inverse folding approach by conserving the pseudoknotted structure.

5.5.2. Pseudoknotted hammerhead ribozymes overlapping coding sequence

Next, we wanted to highlight how Enzymer could be used to include additional sequence constraints to obtain efficient ribozymes. As illustrated by the N/O control in Figure 1, the “potential capability” to fold in the active structure is not enough to ensure it will do it efficiently. We thus designed ribozymes flanking the beginning of the coding sequence for the Red Fluorescent Protein (RFP). As a template, we used a type I hammerhead ribozyme found in the *Yarrowia lipolytica* genome (Perreault *et al.*, 2011b) (**Fig. 5.2A,B**). Sequence constraints include the conserved catalytic core, the Shine-Dalgarno (SD) sequence and the first 10 codons of RFP (**Fig. 5.2A**). Identity of the rest of the nucleotides could vary in order to accommodate these constraints and permit proper folding of the HHRz, while ensuring that we did no modifications to the RFP protein. The aim is to have a non-defective protein and to be sure that if there is a decrease in the protein’s expression it is caused by the ribozyme’s activity and not a mutation in the protein’s sequence, not even a synonymous substitution. It should be noted that while stem II was designed to form two base pairs (as in the WT), the conserved G-C base pair of stem II was changed to U-A because “A” was a constraint from the SD sequence at that position. Such transversions are expected to reduce rate by >> 100 fold (Ruffner *et al.*, 1990), this is thus in principle a very detrimental change. Nevertheless, this study was in the minimal HHRz model and in this pseudoknotted version, the extent of cleavage was similar to that of WT and the rate was reduced by “only” 30 fold for HHRz YL1 compared to WT according to our results.

To have uncleaved size markers, we mutated both HHRz YL ribozymes in the catalytic core without affecting the coding sequence. The ribozymes were tested for their cleavage efficiency during *in vitro* transcription and their activity was compared to the wild type. Both ribozymes were active with fractions cleaved of 76% for the HHRz YL1 and 83% for the YL2 which are close to the cleavage efficiency of the wild type (86%) (**Fig. 5.2C**). The ribozymes were divided into two parts in the same way as for the mouse gut metagenome ribozymes (**Table 5.2** for DNA sequences). The ribozyme WT was able to cleave 54% of the substrate after 8 minutes compared to the two ribozymes designed by enzymer (YL1

and YL2) that attained only 28 and 25% of cleavage rate after 64 minutes. While end-point cleavage was lower, it was still within a few folds of the WT, on the other hand, k_{obs} was orders of magnitude lower.

In spite of an activity much lower than for the ribozymes from figure 5.1, expression of RFP was repressed in presence of the active version of the ribozyme (RzYL1) as compared to the inactive mutant version (RzYL1m), as illustrated by the lower RFP fluorescence (**Fig. 5.2D**). The background, represented by *E. coli* without any RFP construct, was relatively high, raw fluorescence results are therefore presented in **Fig. 5.2D** to avoid misinterpretation. Otherwise, subtracting the background from both the active RzYL1 and its inactive version suggest an eight-fold decrease in expression due to ribozyme cleavage. An example of RFP expression without any potentially hindering secondary structures in its 5'-UTR (such as HHRz stems and pseudoknot) is also provided for indicative purpose.

5.5.3. Synthetic glmS ribozymes activated by glucosamine-6-phosphate

To be able to see if Enzymer could successfully design structures more complex than the HHRz, we used the glmS ribozyme as another template. This ribozyme has a multi-stem structure, three pseudoknots and many highly conserved nucleotides (**Fig. 5.3A, B**). To design such a ribozyme we used the glmS ribozyme from *Thermoanaerobacter tengcongensis* (Klein & Ferre-D'Amare, 2006).

We used two glmS sequences generated by Enzymer to compare their activity to the wild type glmS ribozyme. These ribozymes were transcribed in presence or absence of the GlcN6P, where activity could be observed only in the former case. The wild type has a cleavage end-point of 76% and glmS2 designed by Enzymer has a 45% cleavage efficiency while no activity was detected for glmS1 in these conditions (**Fig. 5.3C**). We also assayed glmS ribozymes after purification to evaluate rate constants (**Fig. 5.3D**). While co-transcriptional cleavage was not observed for glmS1, we could detect a small amount of cleavage for glmS1 in these conditions, but both glmS1 and glmS2 cleaved much more slowly than the WT (1.7 min^{-1} vs 5.2×10^{-7} and 0.052 min^{-1}) (**Fig. 5.6**). Still, as for the wild type, self-cleavage glmS2 is completely GlcN6P dependent since we could not detect cleavage in its absence, showing that in addition to self-cleaving capabilities, the designed glmS2 also have a competent binding pocket for GlcN6P.

To better understand the low cleavage rate of GlmS2 compared to the WT ribozyme, we wanted to test if the several GGGs, present in multiple positions in the GlmS2 sequence, form a G-quadruplex. If this is the case, a G-quadruplex might affect the formation of the

correct active structure of the ribozyme. Our results have shown no almost no difference in activity and cleavage rate of GlnS2 in presence of either KCl or NaCl (0.052 and 0.041 min^{-1}) which was an indication that either no G-quadruplex is formed or if does, it does not affect greatly the structure.

5.4C). GImS2 ribozyme has more similarity in the structure with the WT in the part closer to the 3' (after G73) compared to GImS1. These results may explain the cleavage activity of this sequence compared to glmS1. This difference may be the reason for the instability of glmS1's structure causing the cleavage reaction's absence or the inability of the ribozyme to bind the cofactor correctly.

Enzymer could design pseudoknotted functional ribozymes. Even the very significantly reduced rate of cleavage in *trans* for RzYL1 appears to be sufficient for biologically significant activity according to RFP reduced expression with the high *cis* cleavage activity. We also demonstrate for the first time that the Dirks energy model (Dirks & Pierce, 2003) which was developed for characterization of pseudoknotted RNA shapes, can be effectively used to model and design functional ribozymes. The fact that most HHRz designs have much lower k_{obs} , compared to WT, suggests that the ensemble defect is better to optimize the end point cleavage %, which means the *cis* acting versions of ribozymes, for which most HHRz designs were comparable to WT, rather than cleavage rate, for which only HHRz 13 was comparable to WT.

Previous attempts reengineered the glmS ribozyme (Lau & Ferre-D'Amare, 2016; Lau *et al.*, 2017; Lau, 2013), where sequences differed from the WT by as few as 3 mutations, with many other examples having ~10 mutations and as much as 26 mutations, although the majority of these were not assayed individually; or ~3 to ~5 mutations on average in work from Link and co-workers (Link *et al.*, 2006). The glmS ribozyme sequences we designed differed from their model natural template by 92 and 79 bases (out of 152) for glmS1 and glmS2, respectively and they have the least sequence similarity with the WT ribozyme sequence for any designed glmS ribozyme, as well as the least overall sequence similarity with natural sequences as well (**Table 5.4**). Interestingly, both glmS1 and glmS2 diverge slightly from the more recent alignment consensus found in (McCown *et al.* 2011), with 5 positions for glmS1 and 7 for glmS2, that differ from positions with at least 90% of conservation. Interestingly, one of the P2 base pair from our model is not conserved in the McCown 2011 consensus (see **Fig. 5.7**). The fact that our designs have a G-C bp rather than a G-U (differences indicated in **Fig. 5.4** with a star next to the base pair), like the WT template used from Klein *et al.* 2016 at that position would stabilize a bp that could potentially be only transient in the folding pathway leading to catalysis. Together, this suggests that rather than being a faulty Enzymer design, glmS1 could be co-transcriptionally inactive because the provided model lacked important information. Indeed, glmS1 was generated in our early attempts to design a sequence for this complex ribozyme and as such it lacked some constraints (colored in blue in **Table 5.1**). For instance, the different potential to form purine stacks in an internal loop of P4, which is described as the “floor” of the GlcN6P binding site by Klein *et al.* 2006 would likely impact ribozyme activity if it affects binding GlcN6P or its positioning relative to the cleavage site. As opposed to HHRz, for which the catalytic core has been studied extensively and is very

well defined, we may have overlooked important constraints in the glmS ribozyme, explaining that the two tested designs had a considerably decreased activity compared to WT. It is the case especially for GImS1 ribozyme that had very low to absent cleavage activity compared to the GImS2 that had a good cleavage efficiency and structure similar to the WT. Such constraints could be related to the target structure itself, but could also include additional structures corresponding to intermediates during catalysis: pre-bound, bound pre-cleaved, cleaved states. Previous studies (Klein et al. 2006; Cochrane et al. 2009) indicate global rigidity of this RNA and no significant conformational change between these states, the fact that the cleaved state remains unchanged, in spite of losing a portion of sequence may be important as the designs and corresponding ensemble defects were computed for the full sequence only. A more recent study shows that the ribozyme folding in the structure allowing the binding of the ligand is unstable and even the cofactor binding to its binding pocket is unable to stabilize this conformation. These findings suggest that the structure is too dynamic and that there should be an equilibrium between the structure's stability and the ligand binding for the cleavage to occur efficiently. This might help explain the low activity of the two glmS ribozymes designed by Enzymer (Savinov & Block, 2018). Alternatively, it could also highlight challenges that still need to be overcome for inverse folding. For instance, the WT sequence likely has a favorable folding landscape to achieve the "target structure" while the best way for Enzymer to simulate this is to have a minimal ensemble defect. Multiple instances of kinetic traps or misfolded dead-ends exist in the literature, which have been shown to be reduced by specific folding intermediates (Brown *et al.*, 2004; Cao & Chen, 2007). If the strategy of using ensemble defect to reduce such misfolded RNAs seems to work, at least to some extent, it may also have limitations. Taking in account multiple substructures may represent one of the future challenges of inverse folding, but represents a challenge for experimentalists as well.

Interestingly, when comparing the rate constants with the design temperatures we observe an inverse correlation between the ribozyme design temperature and their activity, as opposed to end-point cleavage assays. The most active designed ribozymes were: HHRz 13 followed by HHRz 12 and HHRz 11, respectively designed for 17°C, 27°C and 27°C (details in **Table 5.5** and **Fig. 5.5**). This may illustrate an advantage of having less stable (less rigid) structures, which could concurrently, also potentially reduce likelihood of misfolded dead-ends. The use of *trans*-cleaving ribozymes facilitates the measurement of

time series to obtain rate constants, but may not always be representative of the *cis*-cleaving version, which is normally expected to cleave more efficiently.

To our knowledge, only two groups actually tested ribozymes designed through inverse folding. The first group (Dotu *et al.*, 2014), where they used a type III HHRz from a portion of the plus polarity strand of Peach Latent Mosaic Viroid (PLMVd). To make their design they used RNAiFold and they tested the generated ribozymes experimentally. These ribozymes were active, but RNAiFold does not take into account pseudoknots. On the other hand, others have designed pseudoknotted ribozymes, like the HDV ribozyme, but without experimental validation (Taneda, 2012). More recently, (Yamagami *et al.*, 2019) have designed active ribozymes (~50 fold slower than WT), which do have pseudoknots, but this work, to our knowledge, the only other experimental validation of pseudoknotted ribozymes designed by inverse folding and they only focused on HDV ribozymes, which have a different topology than glmS and the two HHRz types that we have used as templates for designs. In that article, Yamagami and co-workers had fixed base identity for several positions, reducing the sequence space to search from 53 to 43 Ns (64, 38 and 101 Ns for HHRz from mouse gut, HHRz from *Yarrowia* and glmS, respectively in this study). Among the fixed base identity, five “As” were assigned to the loop 4 of HDV, interestingly, beyond our sequence constraints, multiple contiguous “As” were often noticed in regions designed to be single stranded. Perhaps this occurred as minimization of ensemble defect that might preferentially assign A instead of U or G to form single stranded regions. This is due to the fact that Us or Gs have more base pairing capability due to wobble G-U base pairs in addition to G-C. Also Cs are probably avoided because they can form stronger base pairs with G than As with U.

The ability of Enzymer to design sequences for complex active structures with many sequence constraints, such as for the *Yarrowia lipolytica* HHRz example, paves the way for combinations of sequence elements. Indeed, we were able to design functional ribozyme structures with unprecedented non-ribozyme sequence constraints (where >40% of the ribozyme sequence was fixed to fit the RFP sequence). Interestingly, we noted that glmS2 had several GGGs, including four which were close that would theoretically have the potential to form a G-quadruplex according to G4Hunter (Brazda *et al.*, 2019). Even if our comparison of activity in NaCl vs KCl suggests that G-quadruplexes do not actually form, it nevertheless highlights the potential uses of being able to push the boundaries in terms of divergence from the natural sequences. With proper design, many

useful arrangements can be made to engineer new regulatory elements by overlapping important sequences, such as coding sequence or splicing sites, with structures, such as ribozymes or riboswitches, to gain new functions. As an example, a riboswitch controlled by two or more aptamers with an expression platform than can be a ribozyme or one similar to the platforms of natural riboswitches.

5.7. Materials and methods

We generated a population of candidate sequences for each ribozyme. For each ribozyme, we sorted the generated sequences by their predicted normalized ensemble defect value, chose a small set for *in-vitro* studies and measured their catalytic activities. We also designed a non-HHRz ribozyme: the GlcN6P-dependent glmS ribozyme.

- **Designing ribozymes**

We obtained the secondary structures as well as the minimum catalytic core (reviewed in Nelson 2008) for the HHRzs from (Perreault, 2011). For the glmS ribozyme, we used the secondary structure as well as the minimum catalytic core from the model presented by (Klein, 2006), but in its natural *cis*-acting form. Given a secondary structure and a set of nucleotides representing the minimal catalytic core, the step to initialize a design template is to generate a seed sequence using the catalytic core and then using the letter “o” for any position which is not part of the catalytic core, or highly conserved nucleotides, of the ribozymes. In other words, positions considered to be involved only in secondary structure (standard base pairs, including pseudoknots) as well as single stranded “non-essential positions” are open for change of nucleotide identity. To account for the three dimensional structure essential for ribozyme activity, all conserved positions considered important for tertiary contacts are fixed in the provided model.

We ran Enzymer using the design templates as input and set the maximum number of iterations to 600. We generated 14, 2 and 2 sequences for the mouse gut metagenome hammerhead, *Yarrowia lipolytica* hammerhead and the glmS ribozymes respectively in independent trials. For most trials, we chose the Mathews parameters (Mathews *et al.*, 1999), but for optimal folding at temperatures other than 37°C we used the Serra and Turner energy parameters (Serra & Turner, 1995). For all cases, we used the additional parameters for pseudoknots from the Dirks and Pierce model (Dirks & Pierce, 2003). More details on how to use Enzymer are available in (Najeh *et al.*, 2021) and on the algorithm in (Zandi *et al.*, 2016). In short, users need to provide a secondary structure model with

sequence constraints, as well as parameters as described above. The sequence constraints will typically correspond to the conserved catalytic core of the ribozyme (or of the binding pocket of riboswitches for instance), as well as any other nucleotides which participate in non-canonical base pairs that are important to properly fold in the active tertiary conformation. For the HHRz, these bases are very well defined, but for glmS these constraints were determined from a combination of sequence conservation (McCown *et al.*, 2011) and information from the structure (Klein & Ferre-D'Amare, 2006). The sequences we generated are presented in **Table 5.1** together with their design templates (which are also included in each relevant figure), additional details on sequence generation are available (supplementary text).

- **Transcription**

The synthesis of the *cis* acting ribozymes by *in vitro* transcription was made in presence of radioactive Uridine Triphosphate (alpha-³²P-labeled UTP) to be able to visualize and quantify the self-cleavage activity during the transcription. The reaction of 50 µl contained the DNA produced by PCR, 2 mM rNTPS (2 mM each of rATP, rGTP, rCTP) and 0.8 mM UTP, 1X transcription buffer (80 mM HEPES-KOH pH 7.5, 24 mM MgCl₂, 40 mM DTT, 2 mM spermidine), 1 U/µl of inorganic pyrophosphatase (Sigma Aldrich, USA), 1 U/µl RNase inhibitor (Thermofisher), 1 U/µl T7 RNA polymerase (produced by a lab at University of Sherbrooke) and milliQ water for 2 hours at 37°C. Then 1 U/µl of DNase I (RNase free) (NEB) was added to each reaction and they were incubated for 30 min at 37°C. For the glmS ribozymes, for each ribozyme we had two reactions with or without 1 mM of GlcN6P. The RNAs are then precipitated and a sample is taken and run on a denaturing polyacrylamide gel to visualize the cleavage reaction and measure the cleavage rate of the ribozymes during transcription.

For the *trans* acting ribozymes (divided into two parts: the ribozyme and the substrate) and the glmS ribozymes used for the in-line probing, transcription was made in absence of radioactivity. RNAs were extracted by phenol/chloroform extraction followed by ethanol precipitation. The RNAs were later separated on 10% denaturing polyacrylamide gel, the corresponding bands were cut and RNAs were eluted overnight at 4°C.

- **RNA 5' labelling**

After elution, the RNAs were precipitated by ethanol and the substrates were dephosphorylated using 1U antarctic phosphatase (NEB) in presence of 1X buffer (50 mM bis tris propane-HCl, pH 6, 1 mM MgCl₂, 0.1 mM ZnCl₂) during 20 min at 37°C. The enzyme was inactivated by heating at 65°C during 5 min. The RNA labelling was done using 1U of T4 polynucleotide kinase at 37°C during 1h in presence of gamma-³²P-labelled-ATP and 1X of the enzyme buffer (70 mM tris-HCl pH 7.6, 10 mM MgCl₂, 5mM DTT). The labelled substrates were purified on a 10% denaturing polyacrylamide gel and precipitated with ethanol.

For the glmS ribozymes the same conditions were used except for the buffer where the tris-HCl was replaced by the HEPES solution in order to inhibit the self-cleavage of the ribozyme in presence of the tris.

- **Self-cleavage analysis**

An aliquot of 1 μ l of each ribozyme transcription was taken and diluted in 9 μ l of milliQ water. To each aliquot an equal volume of the 2X dye formamide buffer (95% formamide, 10 mM EDTA, 0.025% bromophenol blue and 0.025% xylene cyanol blue) was added. The full length and the cleaved ribozymes were separated by electrophoresis on a 10% polyacrylamide gel in 1X TBE buffer (89 mM Tris, 89 mM boric acid, 0.02M EDTA). The gel was exposed to a storage phosphor screen for (30 min to 1 hour) and scanned with a Typhoon FLA9500 (GE Health Care) and quantified with ImageQuant.

- **Kinetics measurements**

Hammerhead ribozymes: Each *trans* acting ribozyme (5 pmol/ μ l) was pre-incubated in presence of the cleavage buffer (50 mM Tris pH 7.5, 100 mM NaCl and 25 mM KCl) at 85° for 1 min then snap cooled on ice. Trace amounts of radiolabelled substrate and 4 pmol/ μ l of non-radiolabelled substrate (to ensure similar concentrations for all ribozymes) were added to the reaction that was initiated by adding 1 mM MgCl₂ and incubated at 37°. The timer starts at the moment the MgCl₂, then the reaction is followed during 64 minutes. Aliquots of the reaction were taken at different times (20s, 40s, 1', 2', 4', 8', 16', 32' and 64') and the reaction is stopped by putting the aliquot in the stop dye (2X formamide buffer). The RNA were fractionated on a 10% denaturing PAGE and the results visualized and quantified as described. The cleavage rates (k_{obs}) were calculated for each ribozyme using GraphPad by using the equation of 'one phase decay'.

GlmS ribozymes: The kinetic measurements for the glmS ribozymes were done using the *cis* acting versions of the ribozymes. The denaturation-renaturation of these ribozymes was tried in two ways. The ribozymes were denaturated by incubating them at 85°C for 1 min then either directly put on ice. In the second assay after denaturation the RNA were left to cool at room temperature and this gave a better result for the ribozymes cleavage rate.

To start the reactions, 1 mM of GlcN6P was added and reactions were incubated at 37°C for one day. For these ribozymes, in addition to time points mentioned above for the hammerhead ribozymes, aliquots were also taken after 2 hours, 4 hours and 24 hours.

- **In-line probing of glmS ribozymes**

The radiolabeled ribozymes were incubated for 16 hours at 37°C in the in-line probing buffer (2x)(100 mM Tris-HCl, pH 8.3, 200 mM KCl and 40 mM MgCl₂) and in absence of GlcN6P. The stop dye was added to stop the reaction. Two controls were used to identify the positions of the nucleotides on the gel :

The T1 reaction :RNA is incubated in presence of 0.25 M sodium citrate, pH 5.0, and formamide for 2 min at 56°C. Then 1.5 µl of diluted RNase T1 (1/100) is added and the reaction is incubated at 56°C during 5 min.

OH reaction :RNA was put in presence of 2 µl of Na₂CO₃ buffer (10×) (0.5 M Na₂CO₃, pH 9.0, and 10 mM EDTA).

After the two controls were prepared, RNA fractions were separated on an 8% denaturing polyacrylamide gel.

- **RFP repression**

The ribozyme overlapping the RFP sequence (Rz YL1) and its inactive version (Rz YL1m) were cloned in pUC57 with the inducible Lac promoter, upstream of the rest of the RFP sequence (ordered from Biomatic). The plasmids were transformed in *E. coli* BL21DE3. The plasmid pSEVA3b67Rb (ordered from Addgene in *E. coli* DH10B), including a version of RFP without ribozymes was used as positive control to ensure proper fluorescence detection. As a negative control, we used *E. coli* BL21DE3 without any plasmid, and therefore lacking RFP, to evaluate background fluorescence.

The bacteria were pre-cultured in Lysogeny Broth (LB) overnight with ampicillin (100 µg/ml) for pUC57 and chloramphenicol (6 µg/ml) for pSEVA3b67Rb. Then each culture was diluted in M9 minimal mineral media to an OD₆₀₀ of 0.2. When the OD₆₀₀ was around 0.5-0.6, expression was induced using 1 mM IPTG. After 3 hours induction the fluorescence intensity was measured using a Typhoon FLA9500 and was normalized with the OD₆₀₀.

5.8. Acknowledgments

Natural Sciences and Engineering Council of Canada (NSERC) [418240 to J.P.]. J.P. is a junior 1 FRQS research scholar. S.N. received a fellowship from the Armand-Frappier foundation and FRQNT. The authors also wish to thank Elvira Rwasamanzi for technical help.

5.9. Supplementary Information

Table 5.2 Sequences of trans-acting ribozymes, their substrates and corresponding oligonucleotides

Ribozymes	Oligo Sequences	RNA sequences
Rz WT Mouse	TTTGGGTAGTTTTCGCCGGCTCTCACCGTTGCTCATC AGGGGATTTATTTCGGTACCCCTATAGTGAGTCGTAT TA	ggGGUACCGAAUAAAUCGCCUGAUGAGC AACGGUGAGAGCCGGCGAAACUACCCAA A
Sb WT Mouse	TTATGGTACTATCCCGACTACCCTTGCTTATAGTGA GTCGTATTA	ggCAAGGGUAGUCGGGAUAGUACCAUAA
HHRz1	GAATATAACTTTTCGCTACGGTGTAACTGGCTCATC AGGACCCTTTTCGCTACGACCTATAGTGAGTCGTAT TA	ggUCGUAGCGAAAAGGGUCCUGAUGAGC CAGUUACACCGUAGGCGAAAGUUUAUU C
HHSb1	AGTATATATTGGGTCGAGTTATAATGCCTATAGTGA GTCGTATTA	ggCAUUUAACUCGACCCAAUUAUUAUCU
HHRz2	GTTCTACAGTTTTCGCTTTTATTTCAGAGTTGCTCATC AGGGCGGAAGGGCTCCGTTTCTATAGTGAGTCGTAT TA	ggAACGGAGCCCUCCGCCUGAUGAGC AACUCUGAAUAAAAGCGAAACUGUAGAA C
HHSb2	GTATAGAAAACCGCCGACCGTAGGTACCTATAGTGA GTCGTATTA	ggUACCUACGGUCGGCGUUUUCUAUAC
HHRz3	TATGCTTGCTTTTCGGTTTTTTTTCTAGCTTCCCTCATC AGGCCGCTTTTTTGCTAGTTTCTATAGTGAGTCGTAT TA	ggAAUAGCAAAAAGCGGCCUGAUGAGG AAGCUAGAAAAAACCGAAAGCAAGCAU A
HHSb3	GTGTTTTTTTTGCGGCGAGCAAGCGTACCTATAGTGA GTCGTATTA	ggUACGCUUGCUCGCCGCAAAAAAACAC
HHRz5	TTCACGAAGTTTTCGCTTTTTTTGTCCCTCGCTCATC AGGGCCCTTTTTGGGACTTCCCTATAGTGAGTCGTAT TA	ggAAGUCCCAAAAAGGGCCUGAUGAGC GAGGGACAAAAAAGCGAAACUUCGUGA A
HHSb5	TTTTTTTTCTCGGGCCGACTTCGCTTCTCTATAGTGA GTCGTATTA	ggAGAGCGAAGUCGGCCCCGAGAAAAAAA
HHRz7	TGAACCAGTTTTTCGTTTTTTTTTCGCGGGTCACTCATC AGGCCCTCTCTCCCGCTTCCCTATAGTGAGTCGTAT TA	ggAAGCGGGAGAGAGGGGCCUGAUGAGU GACCCGCGAAAAAACGAAACCUGGUUC A
HHSb7	CACTTCTCCTGGGGCGACCTGGCAGCCCTATAGTGA GTCGTATTA	ggGCUGCCAGGUCGCCCCAGGAGAAGUG
HHRz8	CATGGGCCCTTTTCGCTTTTTTTCCACCTCGCTCATC AGGGCACTTTTTGGTGGTTTCCCTATAGTGAGTCGTAT TA	ggAACCACCAAAAAGUGCCCUGAUGAGC GAGGUGGAAAAAAGCGAAAGGGCCCAU G
HHSb8	CTTTTTTTTTGTGCCGAGGACCCGTTTCTATAGTGA GTCGTATTA	ggAACGGGUCCUCGGCACAAAAAAAAG
HHRz9	TCTAGTGCCTTTTCGTTTTTTTTCCGACTCACTCATC AGGCCGCTTCTTGTCCGCTCCTATAGTGAGTCGTAT TA	ggAGCGGACAAGAAGCGGCCUGAUGAGU GAGUCCGAAAAAACGAAAGGCACUAG A
HHSb9	CGTTTTATTTGCGGCGAGGCACTCTACCTATAGTGA GTCGTATTA	ggUAGAGUGCCUCGCCGCAAAUAAAACG
HHRz10	GTTCCGACCTTTTCGCTTTTTTTCCAGGTGCTCATC AGGGCCCTTTTTTCTGGTTTCCCTATAGTGAGTCGTAT TA	ggAACCAGGAAAAAGGGCCUGAUGAGC GACCUGGAAAAAAGCGAAAGGUCCGAA C
HHSb10	CGTTTTTTTTGGGCGAGGTCCGCTCCCTATAGTGA GTCGTATTA	ggGAGCGGACCUCGGCCCAAAAAAACG

HHRz11	TCAATCCCGTTTTGCTTTTTTTGAGACTCGCTCATC AGGGCCTTTTTGTCTCTCCCTATAGTGAGTCGTAT TA	ggGAGAGACAAAAAGGCCUGAUGAGC GAGUCUAAAAAAGCGAAACGGGAUUG A
HHSb11	TCTTTCCTCCAGGCCGACGGGACATACCTATAGTGA GTCGTATTA	ggUAUGUCCCGUCGGCCUGGAGGAAAGA
HHRz12	GTCATGGAGTTTTGCTTTTTTTGGTCCGTGCTCATC AGGGCCCTTTTTGGACCTTCTATAGTGAGTCGTAT TA	ggAAGGUCCAAAAAGGCCUGAUGAGC ACGGACCAAAAAACGCGAAACUCCAUGA C
HHSb12	GGTTTTTTGTGGGCCGACTCCACTCTCCTATAGTGA GTCGTATTA	ggAGAGUGGAGUCGGCCACAAAAAACC
HHRz13	CTCGTGCTGTTTTGCTTTTTTTGAGGGTTCGCTCATC AGGAACCTTTTTCCCTCTTCTATAGTGAGTCGTAT TA	ggAAGAGGGAAAAAGGUUCCUGAUGAGC GACCCUAAAAAAGCGAAACAGCACGA G
HHSb13	TTCTTTTTTTGGTTCGACAGCACTTCTCCTATAGTGA GTCGTATTA	ggAAAGUGCUGUCGAACCAAAAAAAGAA
HHRz13 mut	CTCGTGCTGTTTTGCTTTTTTTGGGGTTCGCTCATC AGGAACCTTTTTCTCTTTTCTATAGTGAGTCGTAT TA	ggAAAAGAGAAAAAGGUUCCUGAUGAGC GAACCCAAAAAAGCGAAACAGCACGA G
HHRz13 mut com	CTCGTGCTGTTTTGCTTTTTTTAGGATTCGCTCATC AGGAACCTTTTTATCCTTCTCCTATAGTGAGTCGTAT TA	ggAAAGGAUAAAAAGGUUCCUGAUGAGC GAAUCCUAAAAAAGCGAAACAGCACGA G
RzWT YL	GCCCCACTGTTGATGTTTCGACGCTAGTCATGGGT TTGTTCTCATCAGGGCAGCCAGTCCCCCCTATAGT GAGTCGTATTA	ggGGGGGACUGGCUGCCCUGAUGAGAAC AAACCCAUGACUAGCGUCGAAACAUCAA CAGUGGGGGC
Sb WT YL	TTATGACTAGTGGCTGCCGACACCAACACCTATAGT GAGTCGTATTA	ggUGUUGGUGUCGGCAGCCACUAGUCAU AA
Rz YL1	ATGCTCCGTTTTCTCCTTCATCAAATCTCTTTGGATC ATCAGGGAAGACGACACTATCCTATAGTGAGTCGTA TTA	ggAUAGUGUCGUCUCCCCUGAUGAUCCA AAGAGAUUUGAUGAAGGAGAAACGGGAGC AU
Sb YL1	TCTTTGATAACGTCTTCCGACGAAGCCCCTATAGTG AGTCGTATTA	ggGGCUUCGUCGGAAGACGUUAUCAAAG A
RzYL2	ATGCCTCGTTTTCTCCTTCATCAAATTTTTTTGGATC ATCAGGGAAGACGGGTATATCCTATAGTGAGTCGTA TTA	ggAUUAUACCCGUCUCCCCUGAUGAUCCA AAAAAUUUGAUGAAGGAGAAACGAGGC AU
SbYL2	TCTTTGATAACGTCTTCCGACGAAGCCCCTATAGTG AGTCGTATTA	ggGGCUUCGUCGGAAGACGUUAUCAAAG A

For the hammerhead ribozyme designs HHRz_1 to HHRz_14, *trans*-acting ribozyme assays were performed by opening the middle of the loop of stem III. Related information is also available in **Table 5.2**. Note that two “G” bases are added in 5′ (to allow efficient transcription), this is also true for the substrate of the *trans*-acting version.

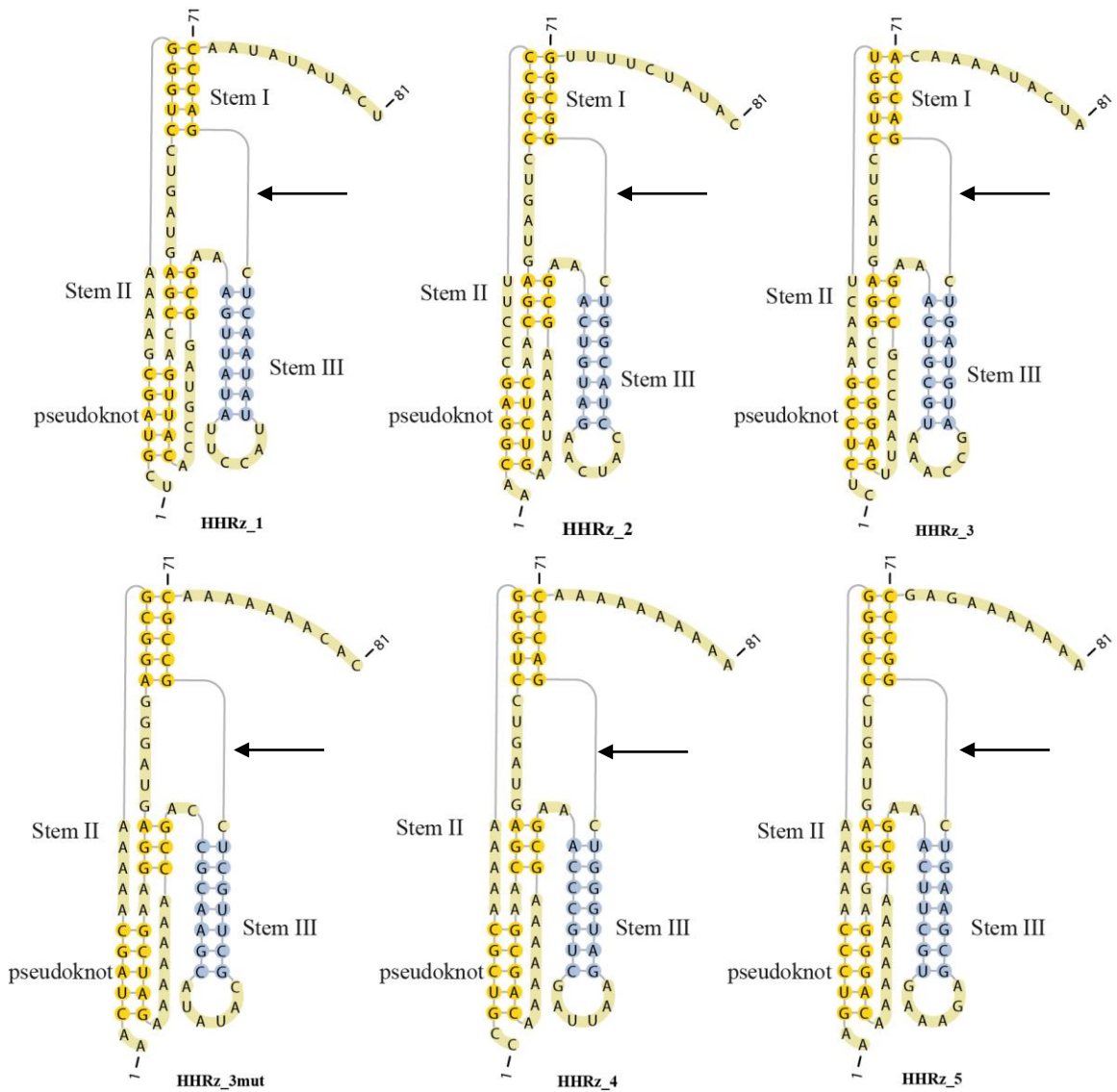
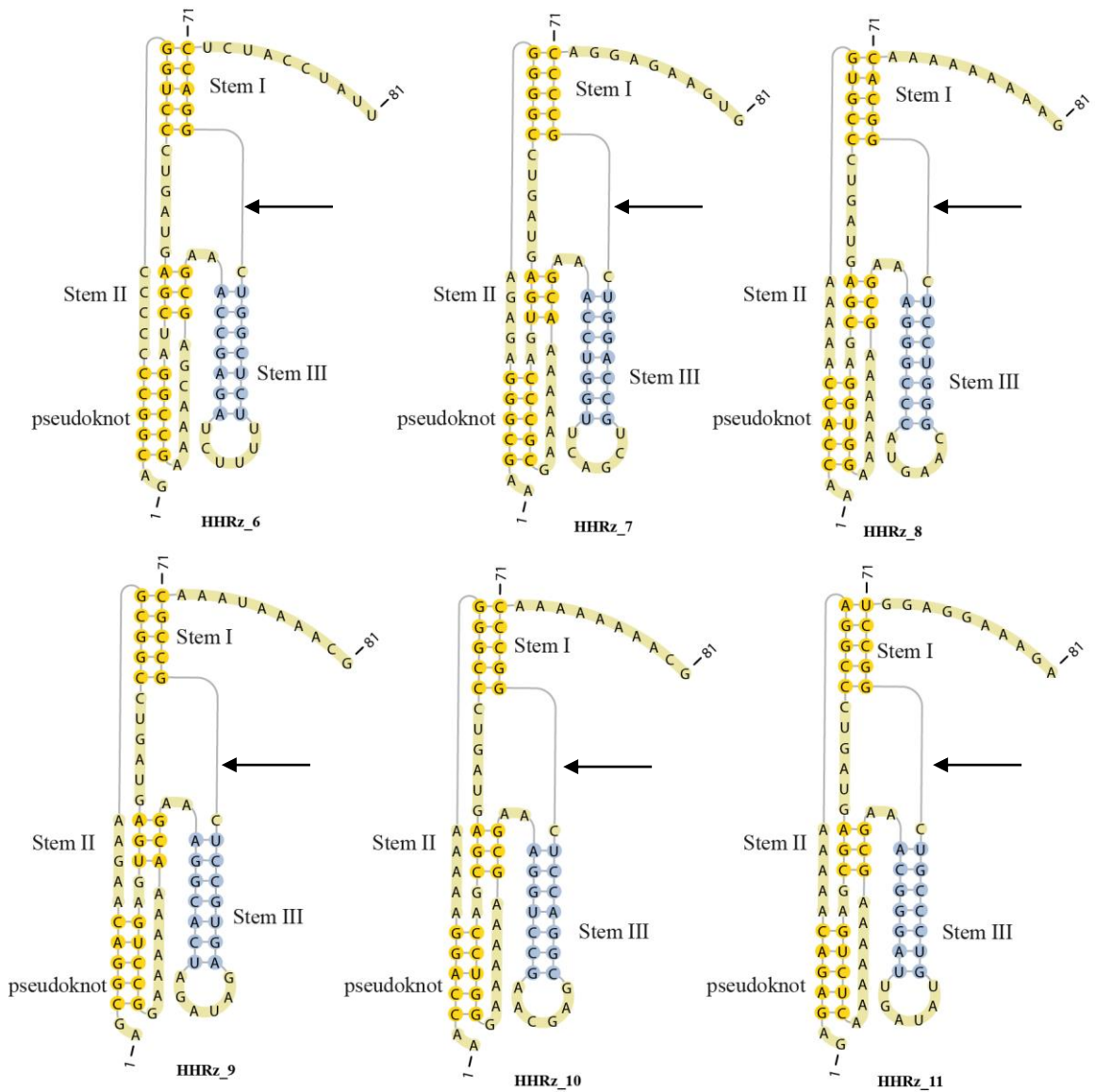


Figure 5.5 Secondary structures of ribozymes

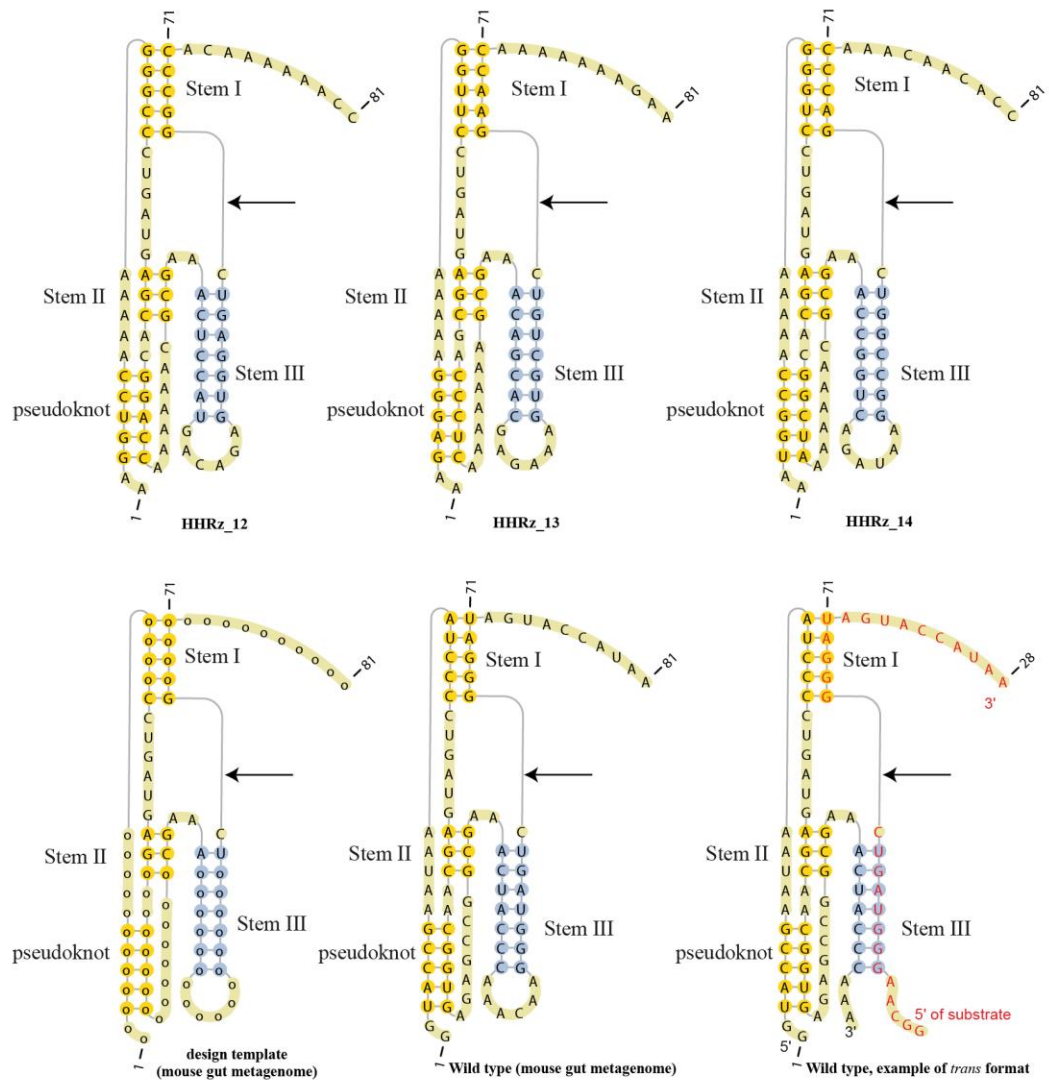
Stems are indicated as per the convention, in addition to the pseudoknot that contributes to bringing stems I and II closer. Beige nucleotides are single stranded, while yellow and blue are double stranded by design. Arrows indicate the cleavage site.

(figure continues on the following pages, a *trans*-acting HHRz example is presented in 2 pages).



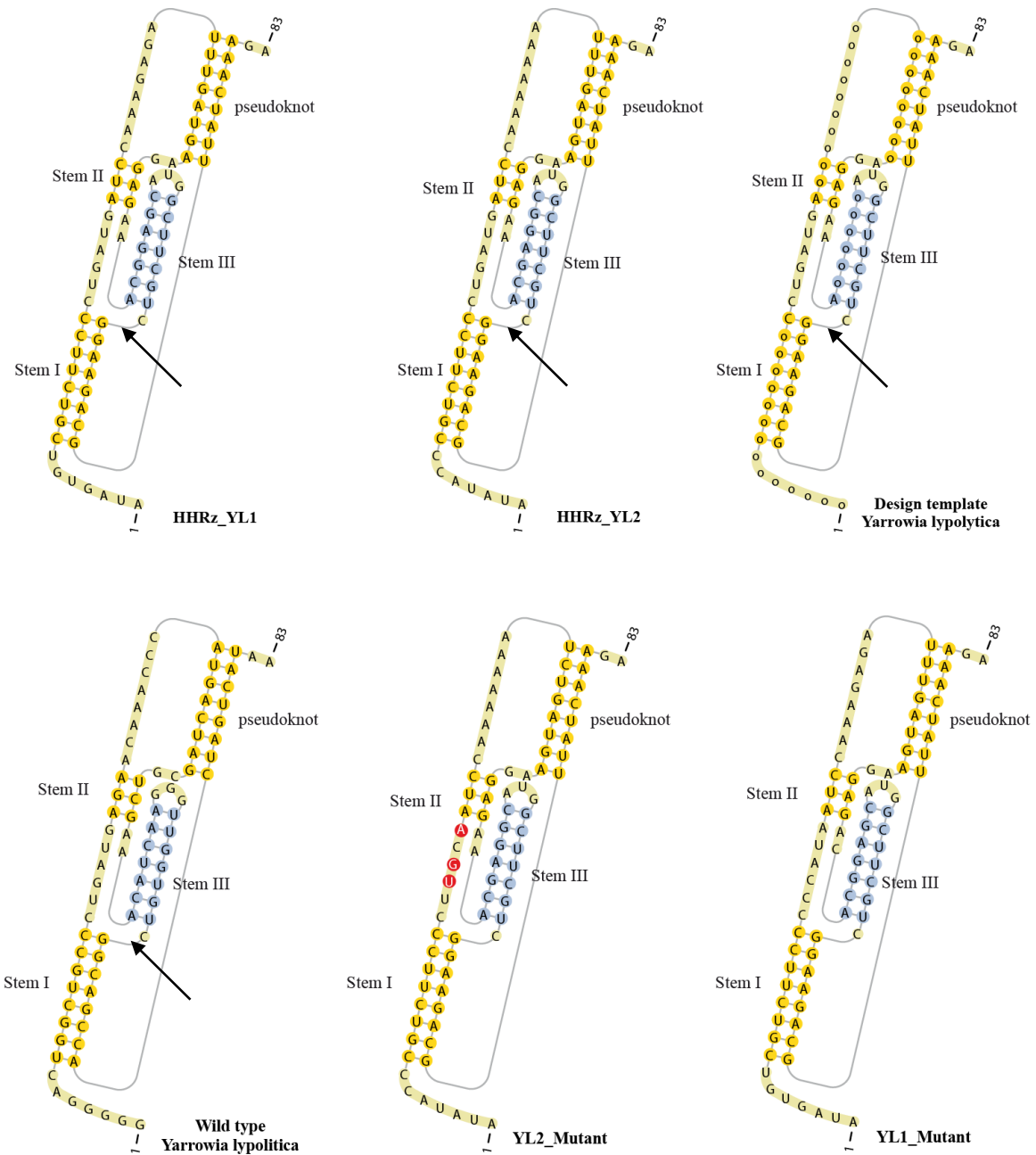
(continued) Figure 5.5. Secondary structures of ribozyme designs. Stems are indicated as per the convention, in addition to the pseudoknot that contributes to bringing stems I and II closer. Beige nucleotides are single stranded, while yellow and blue are double stranded by design. Arrows indicate the cleavage site.

(figure continues on the following pages, a *trans*-acting HHRz example is presented next page).



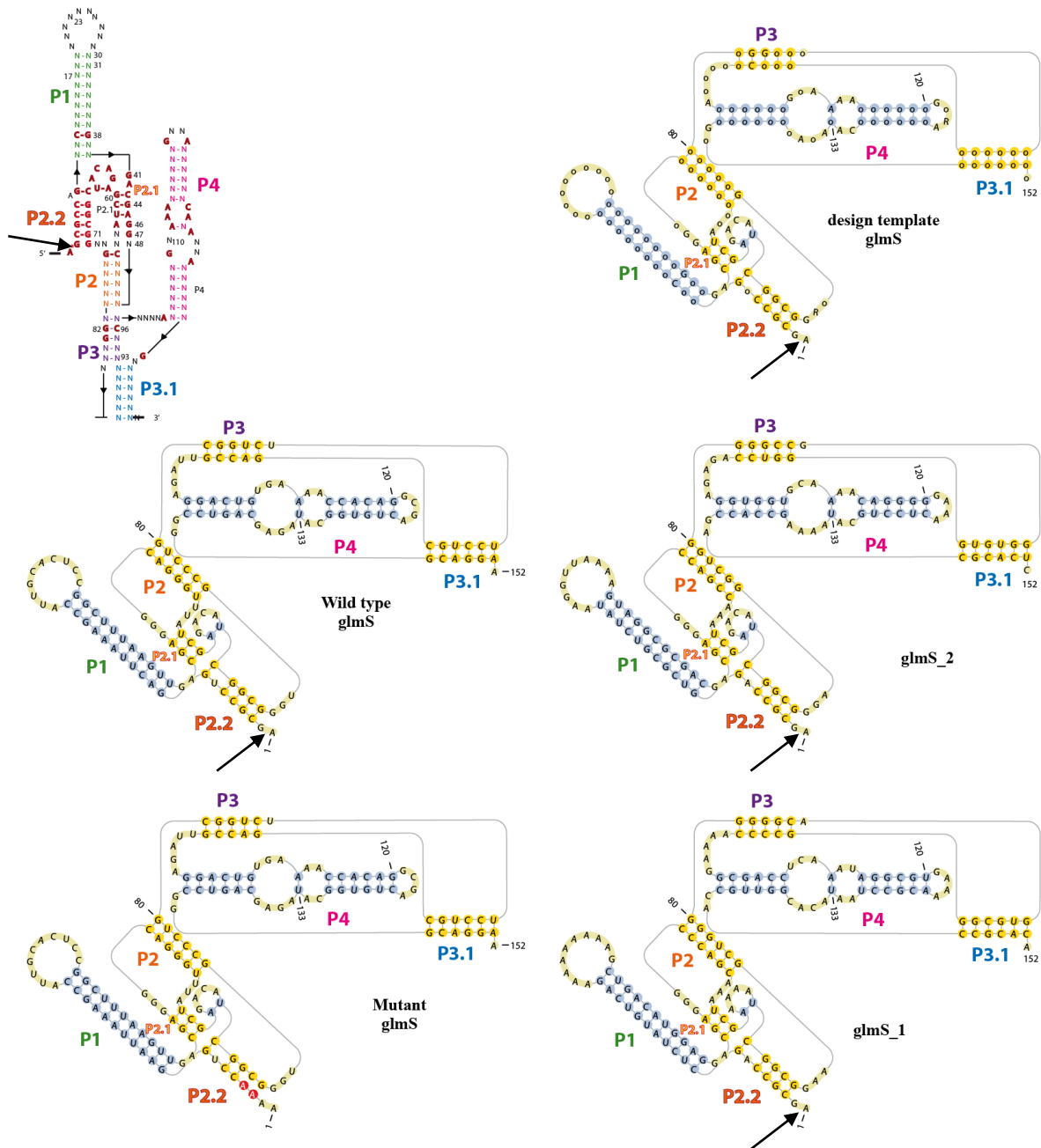
(continued) Figure 5.5. Secondary structures of ribozyme designs. Stems are indicated as per the convention, in addition to the pseudoknot that contributes to bringing stems I and II closer. Beige nucleotides are single stranded, while yellow and blue are double stranded by design. Arrows indicate the cleavage site. For the design template, open circles correspond to positions free to vary while letters correspond to nucleotide constraints. The last structure (bottom right) provides an example of the *trans*-acting configuration, where the substrate portion of the RNA is colored in red.

(figure continues on the following pages).



(continued) Figure 5.5. Secondary structures of ribozyme designs. Stems are indicated as per the convention, in addition to the pseudoknot that contributes to bringing stems I and II closer. Beige nucleotides are single stranded, while yellow and blue are double stranded by design. Arrows indicate the cleavage site. For the design template, open circles correspond to positions free to vary while letters correspond to nucleotide constraints. For the mutant, mutated positions are circled in red.

(figure continues on the following page).



(continued) Figure 5.5. Secondary structures of ribozyme designs. Stems are indicated as per the convention shown on the top left. Beige nucleotides are single stranded, while yellow and blue are double stranded by design. Arrows indicate the cleavage site. For the design template, open circles correspond to positions free to vary while letters correspond to nucleotide constraints. For the mutant, mutated positions are circled in red.

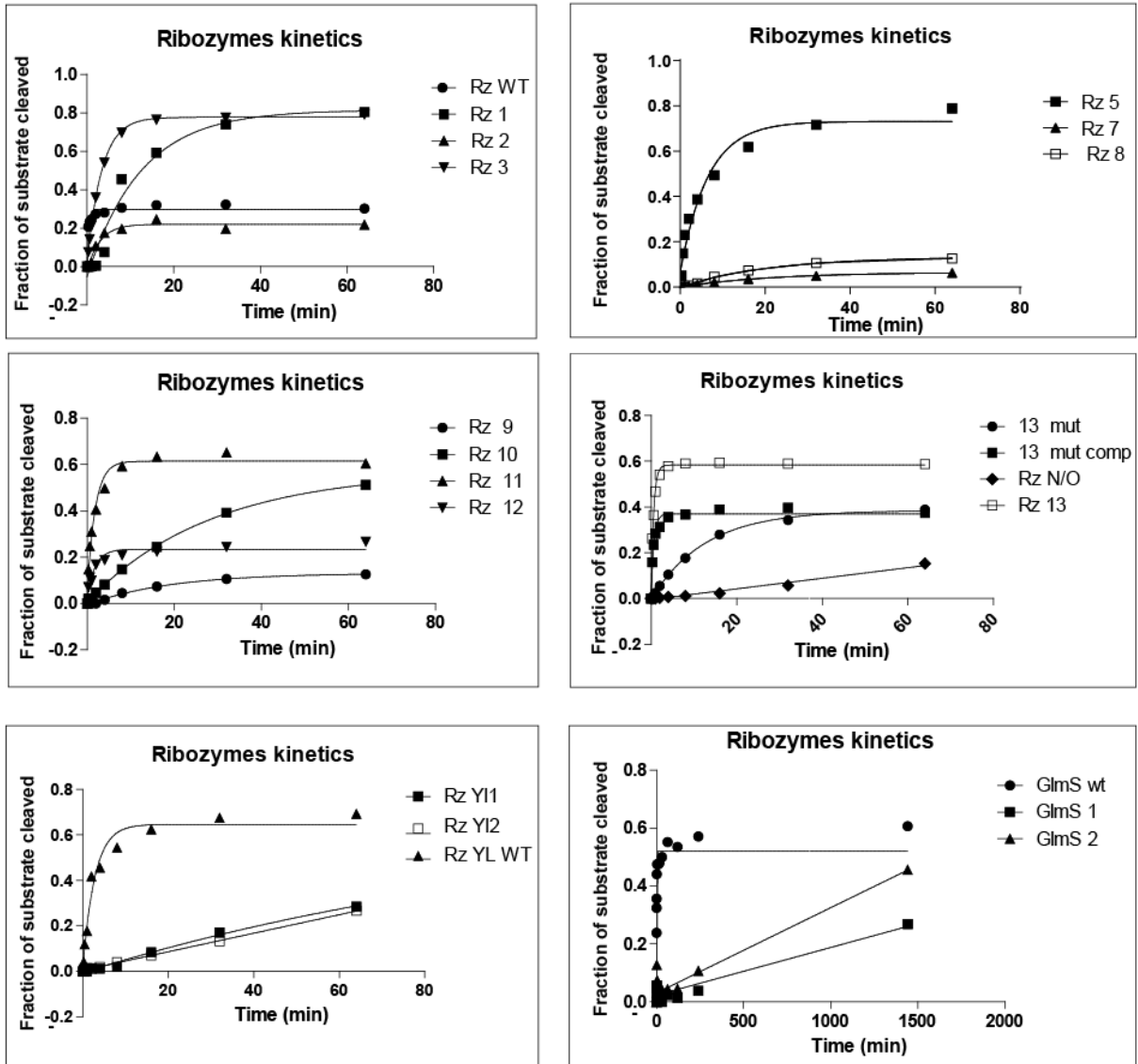


Figure 5.6 Ribozymes activity

Cleavage over time is plotted for all ribozymes assayed. Note that this is for the trans-acting versions (except for glmS). It can be noted that the WT (from the HHRz found in the mouse gut metagenome) has a much lower % cleavage, with an apparent plateau around 30%. It is nevertheless the HHRz with the best cleavage rate (3 min). The Rz13 presenting the best Kobs compared to all the other ribozymes shows a higher cleavage % than the WT with a plateau at almost 60%.

Table 5.3 Oligonucleotide sequence for cis-acting ribozymes

Name	Sequence
T7	TAATACGACTCACTATAGG
WT	TTATGGTACTATCCCAGACTACCCCTTGTGGGTAGTTTCGCCGGCTCTCA CCGTTGCTCATCAGGGGATTTATTCGGTACCCCTATAGTGAGTCGTATTA
HHRz 1	AGTATATATTGGGTTCGAGTTATAATGGAATATAACTTTCGCCTACGGTGT AACTGGCTCATCAGGACCCTTTTTCGCTACGACCTATAGTGAGTCGTATTA
HHRz 2	GTATAGAAAACCGCCGACCGTAGGTAGTTCTACAGTTTCGCTTTTTATTCA GAGTTGCTCATCAGGGCGGAAGGGCTCCGTTCCATAGTGAGTCGTATTA
HHRz 3	GTGTTTTTTTTGCGGCGAGCAAGCGTATATGCTTGCCTTTCGGTTTTTTTTCT AGCTTCCATCATCAGGCCGCTTTTTGCTAGTTCCATAGTGAGTCGTATTA
HHRz 4	TTTTTTTTTTGGGTTCGACCCATCTTAATCGACGGGTTTCGCTTTTTTTGT CGCTTGCATCATCAGGACCCTTTTTTCGACGGCCATAGTGAGTCGTATTA
HHRz 5	TTTTTTTTCTCGGGCCGACTTCGCTCTTTCACGAAGTTTCGCTTTTTTTGT CCCTCGCTCATCAGGGCCCTTTTTGGGACTTCCATAGTGAGTCGTATTA
HHRz 6	AATAGGTAGAGGTCCGACCGAGAAAAAGATCTCGGTTTCGCTCGTTTTTCG GCCTAGCTCATCAGGGACCGGGGGGGCCGTCCTATAGTGAGTCGTATTA
HHRz 7	CACTTCTCCTGGGGCGACCTGGCAGCTGAACCAGGTTTCGTTTTTTTTCGC GGTCACTCATCAGGCCCTCTCTCCCGCTTCCATAGTGAGTCGTATTA
HHRz 8	CTTTTTTTTTGTGCCGAGGACCCGTTTCATGGGCCCTTTCGCTTTTTTTCC ACCTCGCTCATCAGGGCACTTTTTGGTGGTTCCATAGTGAGTCGTATTA
HHRz 9	CGTTTTATTTGCGGCGAGGCACCTATCTAGTGCCTTTCGTTTTTTTTCCG GACTCACTCATCAGGCCGCTTCTTGTCCGCTCCATAGTGAGTCGTATTA
HHRz 10	CGTTTTTTTTGGGCCGAGGTCCGCTCGTTCGGACCTTTCGCTTTTTTTCC AGGTGCTCATCAGGGCCCTTTTTCCGTTCCATAGTGAGTCGTATTA
HHRz 11	TCCTTCCCTCCAGGCCGACGGGACATATCAATCCCGTTTCGCTTTTTTTGA GACTCGCTCATCAGGGCCCTTTTTGTCTCTCCCTATAGTGAGTCGTATTA
HHRz 12	GGTTTTTTGTGGGCCGACTCCACTCTGTTCATGGAGTTTCGCGTTTTTTGG TCCGTGCTCATCAGGGCCCTTTTTGGACCTTCCATAGTGAGTCGTATTA
HHRz 13	TTCTTTTTTTGGTTCGACAGCACTTTCCTCGTGTCTGTTTCGCTTTTTTTGA GGGTCGCTCATCAGGAACCTTTTTCCCTCTTCCATAGTGAGTCGTATTA
HHRz 14	GGTGTGTTTGGGTTCGACCGGCCATTATCTGACCGGTTTCGCGTTTTTTTA GCCGTGCTCATCAGGACCCTTTTGGCCATTCCATAGTGAGTCGTATTA
HHRz mut	GTGTTTTTTTTGCGGCGAGCAAGCGTATATGCTTGCCTGCGGTTTTTTTTCT AGCTTCCATCCTCCGCTTTTTGCTAGTTCCATAGTGAGTCGTATTA
WT YL	TAATACGACTCACTATAGGGGGGGACTGGCTGCCCTGATGAGAACAAACCCAT GACTAGCGTCGAAACATCAACAGTGGGGGCTGTTGGTGTTCGGCAGCCACTAGT CATAA
R WT YL	TTATGACTAGTGGCTGCCGACAC
Mut YL	TAATACGACTCACTATAGGGGGGGACTGGCTGCCCCAGTGAGAACAAACCCAT GACTAGCGTCGAAACATCAACAGTGGGGGCTGTTGGTGTTCGGCAGCCACTAGT CATAA
HHRz05	ACCCGAGCTCTTTACACTTTATGCTTCCGGCTCGTATGTTATATACCCGCTTT CCCTGATGATCCAAAAAATTTGATGAAGGAGAAACGAGGCATGGCTTCGTCG GAAGACGTTATCAAAGACGTTATCAAAGAGTTCATGCGTTTCAAAGTTCGTAT GGAAGGTTCCGTTAACATGC
HHRz 05 mut	TTTACACTTTATGCTTCCGGCTCGTAATATACCCGCTTCCCTTGCAATCCAA AAAAATCTGATGAAGGAGAAACGAGGCATGGCTTCGTCGGAAGACGTTATCAA AGA
F HHRz05	TAATACGACTCACTATAGGATATAACCCGCTTCCCTGATG
R HHRz05	GCATGTTAACGGAACCTTCATACGAACTTTGAAACGC
F HHRz05 mut	TAATACGACTCACTATAGGATATAACCCGCTTCCCTTGC

F HHRz00	TAATACGACTCACTATAGGATAGTGTCGTCTTCCCTGATGA
R HHRz00	TCTTTGATAACGTCTTCCGAC
F HHRz Mut	TAATACGACTCACTATAGGATAGTGTCGTCTTCCCGATAAT

GlmS WT	AGCGCCTGGACTTAAAGCCATTGCACTCCGGCTTTAAGTTGACGAG GGCAGGGTTTATCGAGACATCGGCGGGTGCCCTGCGGTCTTCCCTGC GACCGTTAGAGGACTGTGAAAACCACAGGCGACTGTGGCATAGAGC AGTCCGGGCAGGAA
F GlmS WT	TAATACGACTCACTATAGGGAAAAAAAAAAAAAAAAAAAAAAG CGCCTGGACTTAAAGCCA
R GlmS WT	TTCTTGCCCGGCTGCTCTATGC
F GlmS cleaved	TAATACGACTCACTATAGGGCCTGGACTTAAAGCCATTGCACTCCGG
F GlmS mut	TAATACGACTCACTATAGGGAAAAAAAAAAAAAAAAAAAAAAC CTGGACTTAAAGCCA
GlmS 1	TGTGCGGTGGCAACCGTGTATTAGGCGTTTTTCACGCCATTTGAGG TCGCCTTTTGGGGCCCGCACTGCCCGCCAGCTTCCGCCGATTTTT CGATTTGCTGGGCCCTCGTCCATGTCAGCTTTTTTTTTCTGAC ATAGAGCTGGCGCTCGCTATAGTGAGTCGTATTA
F GlmS 1	TAATACGACTCACTATAGGGAAAAAAAAAAAAAAAAAAAAAAG CGCCAGCUCUAUGUCA
R GlmS 1	TGTGCGGTGGCAACCGTGTATTAGG
F GlmS 6	TAATACGACTCACTATAGGGAAAAAAAAAAAAAAAAAAAAAGCGCCAGGTGCGG TCTATAAGGTTAAAAGTAGGCGCGACGACGAGGGCCAGCCAAATCGAGACATCG GCGGGAGGCTGG
R GlmS 6	GAGTGCCTCGGTGGCTTTTATGCAGGAGTTTCCCCCTGTTTTGCACCACCTCTC TGGACCCACACCCGGCCCCCAGCCTCCCGCCGATGTCTCGATTTGGCTGGCCCT CGTCGTGCGGCC

- **Supplementary details on sequence design by Enzymer**

The design goal of the algorithm is to interactively evolve the initial template by lowering the predicted normalized ensemble defect of the candidate sequence at each iteration. The algorithm was set to stop after 600 iterations or when the predicted normalized ensemble defect reaches 0.01. In the final sequence, the initial catalytic core remained unchanged. When we run Enzymer, when the run is finished, it returns a single final sequence. The sequences generated internally are of no value and Enzymer does not keep track of them.

Sequences were generated at different temperatures of 17, 27, 37, 47 and 57 (according to Dirks model). For each temperature about 10 sequences were generated, which were sorted from each temperature group based on their ensemble defect and then from each

group two or more were picked for experiments (as described in main text), but the ensemble defects were all very close. For the YL and glmS they were all designed at 37. The final ensemble defects were very close.

Table 5.4. Details for k_{obs} determination

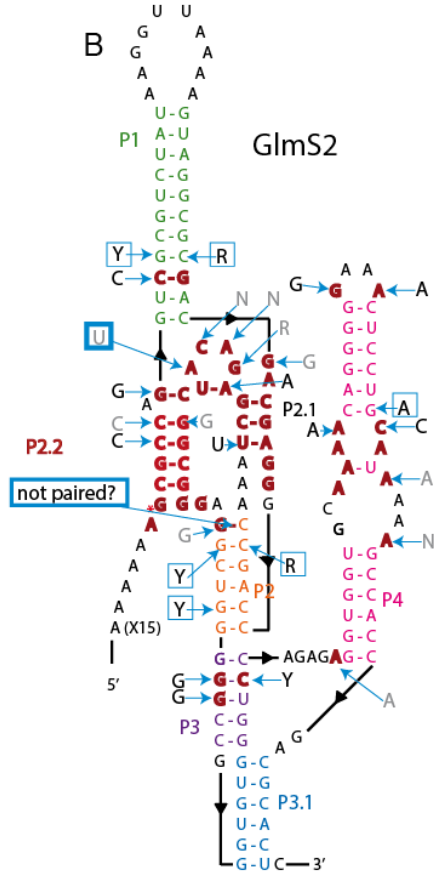
Detailed results of rate determination are provided in an excel sheet

Table 5.5. GlmS1 and glmS2 alignment with natural glmS ribozyme sequences

Detailed results of glmS1 and glmS2 alignments with natural glmS ribozymes are provided in an excel sheet.

A AGCGCCNGNNCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGNGACGAGGNNNNNNNNNNNAUCGAGACAUCGGCGGGANG
 AGCGCCoGooCooooooooooooooooooooooooooooooooooGooGACGAGGoooooooooAUCGAGACAUCGGCGGRooG
 ..[[[[[.((((((((((((.....)))))))))..[[[...((((((...]]])((...))]]]].....)

NNNNNGGNNNNNNNNNNNNNNNNNNAAAANNNNNGAAAANNNNNNNCANANANNNNNNNGNNNNNNNN
 ooooooooooCoooooooooAoooooooooAAAooooooooGNRAooooooooCAoAoooooGooooooooo
))))((((. [[[[[[]))). ((((((.....)))))..))))).]]]]]]..



- X conserved 90% or more
- X conserved 75% or more
- N identity not conserved

Figure 5.7 Discrepancies between the consensus from (McCown et al. 2011) and our model

Based on the structure and consensus from (Klein et al. 2006). The structure of glmS2 is shown as in Figure 3 with the red nucleotides representing all the identities resulting from constraints from the model shown in A (used to generate glmS2). All the constrained positions for which the level of conservation is lower than 97% as the consensus (see color coding) of McCown et al. are shown. Positions with blue boxes highlight all the positions for which a discrepancy is observed for the base identity. Blue boxes with thicker lines highlight more notable discrepancies, notably: the U within the tetraloop of the single bp between P2.1 and P2.2 (which is nevertheless an A in the sequence used to determine the structure in Klein et al. 2006, as well as the top base pair of P2, which is absent in the consensus from McCown et al. 2011. In the latter case, the somewhat conserved G seems to be part of the junction between P2.2 and P2 instead of forming a base pair. It can also be noted in the structure determined from the crystal that this bp is not aligned as well as some other bp. Forming a more stable G-C bp might thus be detrimental, compared to a G-U bp, potentially only transient and captured within the snapshot of the structure determined from the crystal.

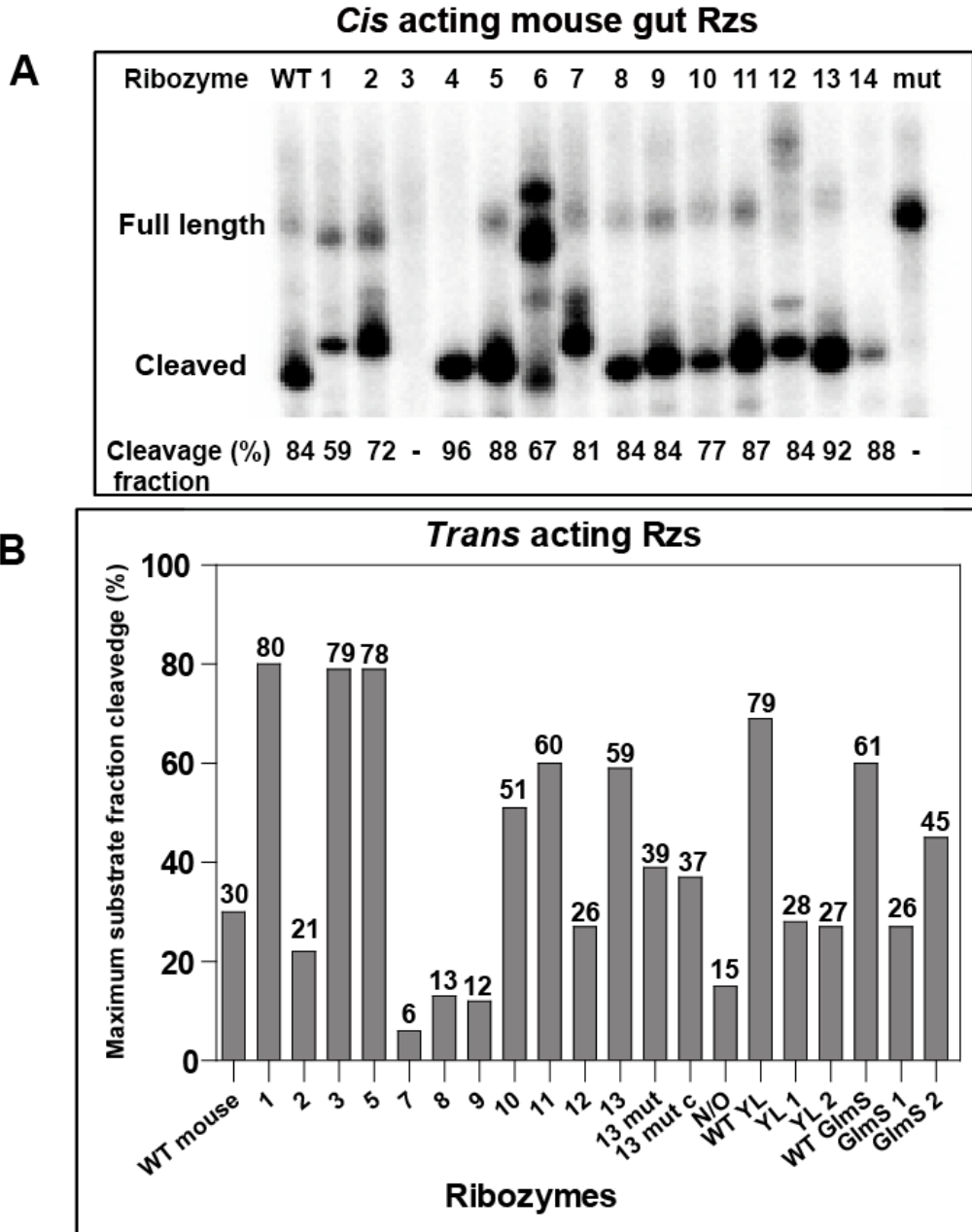


Figure 5.8 Cleavage extent of *cis* and *trans* acting ribozymes

(A) Cleavage activity during transcription, of all selected mouse gut-hammerhead ribozymes including Rzs 4, 6 and 14 that were eliminated from the kinetic tests because of the transcription failure of the *trans* versions. (B) Cleavage extent of all the tested ribozymes resulting from the last aliquot taken at 64 minutes during kinetics following. The used ribozymes were all *trans*-acting.

6. Article 3: A Web Service for Automated Design of Multiple Types of Ribozymes Targeting RNA: from Minimal Hammerhead to Aptazyme

6.1. Mise en contexte

Après avoir montré que nous avons un outil solide pour la génération des ribozymes actifs en *cis* et qu'on peut les impliquer dans le contrôle de l'expression génique, ou les modifier pour les inclure dans notre circuit logique, dans le sixième chapitre nous nous sommes concentrés sur les ribozymes actifs en *trans*. Dans le même contexte de contrôle de l'expression génique, les ribozymes agissant en *trans* présentent une meilleure alternative dans le cas de contrôle de l'expression génique car ils peuvent être utilisés contre pratiquement n'importe quelle séquence d'ARN.

L'un des objectifs du projet était de déterminer que les ribozymes peuvent être appliqués dans un but thérapeutique par la réduction du taux d'expression d'une cible impliquée dans un type de cancer. Pour se faire, plusieurs conditions et paramètres doivent être pris en considération car les ribozymes sont destinés à être utilisés dans des cellules ou un organisme vivant. Pour cela nous présentons dans ce chapitre un nouveau service web 'Ribosoft 2.0' développé pour la conception des ribozymes allostériques et conventionnels, tous actifs en *trans*. Notre travail visait la génération des matrices des modèles de ribozymes utilisées par Ribosoft 2.0 ainsi qu'à toute validation expérimentale de l'activité de ribozymes générés permettant l'optimisation des paramètres de conceptions et de la qualité des résultats donnés.

Les résultats générés nous ont permis de choisir les meilleurs candidats de ribozymes à utiliser contre les ARNm cibles de notre projet qui sont l'ARNm de la Gal-7 et celui de la Bcl-2 (résultats additionnels présentés en **annexe I**). Ces ribozymes joueront le rôle des effecteurs de notre circuit logique (**Figure 6.1**).

Matrice des Ribozymes conçus par Ribosoft 2.0

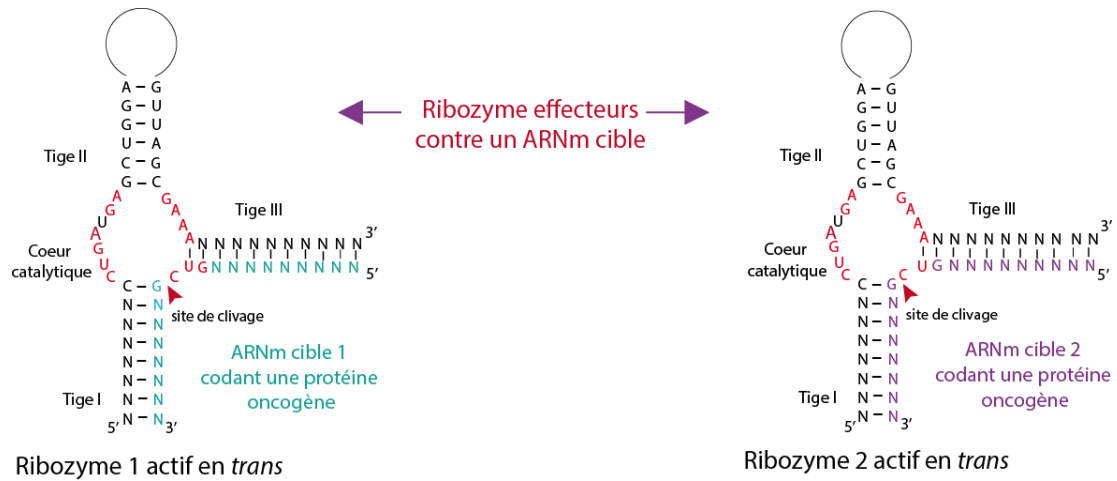


Figure 6.1 Conception de ribozymes actifs en trans par Ribosoft2.0

Ribosoft2.0 permet la conception de différents types de ribozymes actifs en *trans*. La matrice utilisée est la structure secondaire du brin enzyme du ribozyme. La séquence à concevoir est celle de deux tiges simple brin qui s'hybrident par complémentarité de séquence sur l'ARNm cible. Les ribozymes issus de Ribosoft 2.0 sont testés pour leurs activités et les meilleurs sont sélectionnés pour jouer le rôle d'effecteur dans le circuit logique. Ces ribozymes seront capables d'inhiber l'expression des protéines oncogènes afin de causer la mort ou la diminution de la croissance des cellules cancéreuses.

A Web Service for Automated Design of Multiple Types of Ribozymes Targeting RNA: from Minimal Hammerhead to Aptazyme

Un service web pour la conception automatisée de multiples types de ribozymes ciblant les ARN : de ribozymes en tête de marteau minimal jusqu'aptazyme

Sabrine Najeh^{1,2}, Nawwaf Kharma^{2,3*}, Thomas Vaudry-Read³, Anita Haurie³, Christopher Paslawski⁴, Ankit Ranja³, Daniel Adams³, Steve Ferreira⁴ and Jonathan Perreault^{1,2*}

1 INRS-Centre Armand-Frappier Santé Biotechnologie, 531 Boul. Des Prairies, Laval, Quebec, Canada H7V 1B7

2 Centre for Applied Synthetic Biology (CASB), Concordia University, 7141 Sherbrooke Street W., Montreal, Quebec, Canada H4B 1R6

3 Electrical and Computer Engineering Dept., Concordia University, 1455 De Maisonneuve Blvd. W., Montreal, Quebec, Canada H3G 1M8

4 Computer Science & Software Engineering Dept., Concordia University, 1455 De Maisonneuve Blvd. W., Montreal, Quebec, Canada H3G 1M8

*To whom correspondence should be addressed:

Phone: 450-687-5010 x4411, fax: 450-686-5301, e-mail: jonathan.perreault@inrs.ca or phone: 514-8482424 x3117, fax: 514-8482802, e-mail: kharma@ece.concordia.ca

État du manuscrit : Ce manuscrit a été soumis au journal *Nucleic Acids Research*, mais initialement rejeté. Une nouvelle version révisée a été soumise au *International Journal of Molecular Sciences* le 14 Mars 2022 (ijms-1659069)

6.1. Contributions des auteurs

Sabrine Najeh : En collaboration avec Jonathan Perreault, a fourni les matrices (structures secondaires de ribozymes et substrats) utilisées par Ribosoft2.0. La planification du travail expérimental présenté dans l'article (avec plus de données non présentées aussi) ainsi que sa réalisation ont été fait par moi. L'analyse et l'interprétation des résultats, la préparation des figures et la rédaction de tout l'article sauf la partie de méthodologie décrivant l'algorithme de Ribosoft2.0, ont été faites par moi. J'ai participé au développement et à l'optimisation de Ribosoft2.0 avec mes commentaires ainsi que avec mes résultats et analyses comme utilisatrice du service web pour le rendre plus convivial.

Jonathan Perreault : Supervision et assistance de Sabrine Najeh avec beaucoup d'implication dans le développement du service web en agissant comme collaborateur de l'équipe du Pr.Nawwaf Kharma où il a participé dans l'optimisation du service web.

Nawwaf Kharma : Superviseur du groupe d'étudiants qui ont développé le service web et rédaction de la partie de méthodologie concernant Ribosoft 2.0.

Thomas Vaudry-Read, Anita Haurie, Christopher Paslawski, Daniel Adams, Steve Ferreira : L'équipe qui a développé Ribosoft2.0 dans le cadre de leur projet de fin d'étude. Thomas Vaudry-Read est la personne qui restait en contact après la fin de ses études et qui faisait les optimisations et les ajouts demandés avec Ankit Rajan après chaque révision ou les essais d'améliorations du service web.

6.2. Abstract

Ribosoft 2.0 is the second version of a web service to design different types of *trans*-acting conventional and allosteric ribozymes. The web service is publicly available at <https://ribosoft2.fungalgenomics.ca/>. Ribosoft 2.0 uses template secondary structures that can be submitted by users to design ribozymes in accordance with parameters provided by the user. The generated designs specifically target a transcript (or, generally, an RNA sequence) given by the user. Herein, sixty ribozymes of different types were tested on two different mRNAs, with a majority shown to be active. We have also generated and tested the activity of the first *trans*-acting aptazyme designed in silico.

Keywords: RNA, ribozyme, allosteric ribozyme, pistol ribozyme, pseudoknot, *trans*-acting ribozyme, automated design, synthetic biology.

Résumé

Ribosoft 2.0 est la deuxième version d'un service Web permettant de concevoir différents types de ribozymes conventionnels et allostériques actifs *trans*. Le service Web est accessible à <https://ribosoft2.fungalgenomics.ca/>. Ribosoft 2.0 utilise des modèles de structures secondaires qui peuvent être soumis par les utilisateurs pour concevoir des ribozymes conformément aux paramètres fournis par l'utilisateur. Les conceptions générées ciblent spécifiquement un transcrit (ou, généralement, une séquence d'ARN) donné par l'utilisateur. Ici, soixante ribozymes de différents types ont été testés sur deux ARNm différents, dont une majorité s'est avérée active. Nous avons également généré et prouvé l'activité des premiers aptazymes actifs en *trans* conçu in silico.

6.3. Introduction

Ribozymes are non-coding RNA with catalytic activities. Ribozymes were discovered in 1981 in a study demonstrating that the effector of the splicing mechanism was an RNA molecule rather than a protein, and those RNAs were designated group I introns (Cech et al., 1981; Kruger et al., 1982). This event led to the discovery of many other ribozymes, including a group of short RNAs sharing a common 'transesterification' reaction. This type of reaction occurs at a specific site, causing the self-cleavage of the RNA through a nucleophilic attack, generating an RNA molecule with a 2'-3' cyclic phosphate and a 5' hydroxyl group. RNAs belonging to this group are called 'small self-cleaving RNAs', and they are classified into different families based on the secondary and tertiary structure that they adopt, in order to affect the cleavage reaction (Ferre-D'Amare & Scott, 2010). The hammerhead ribozyme (HHRz) was the first small self-cleaving ribozyme to be discovered and is the one most studied. HHRz was first discovered in plant-infecting viroids, before subsequently being discovered in all domains of life, including prokaryotes and eukaryotes, ranging from single-cell organisms to mammals including humans (Hammann et al., 2012). Since the discovery of HHRz, many different families of short self-cleaving RNAs have been identified, including hepatitis delta virus (Nakano et al., 2003), GImS (Winkler et al., 2004), hairpin (Hampel, 1998), Varkud satellite (Saville & Collins, 1990), twister (Roth *et al.*, 2014b), twister sister, hatchet and pistol (Weinberg et al., 2015b) ribozymes. Some of these ribozymes are involved in viral and viroid replication cycle (Bai et al., 2002; Mishra et al., 2016), while the natural function of others remains less-well defined. Ribozymes were also used for gene expression control purposes by their inclusion into mRNA sequences, downstream or upstream of coding sequences (13). They have also been used as independent trans-cleaving RNA molecules that target mRNAs anywhere, including in their coding sequences (Prommana et al., 2013; Scanlon et al., 1991).

Self-cleaving ribozymes possess many potential advantages for RNA targeting tool development and gene expression control. In general, this advantage is due to their small size and activity which is often independent of protein helpers, making them easier to test *in vitro*. Besides their ability to self-cleave (*cis*-cleavage), many ribozymes can be conceptually divided into two parts: a ribozyme strand and a substrate strand, with no loss in activity. The binding site of the ribozyme strand can be easily modified to recognize and

hence, cleave other RNA targets, at specific sites (trans-cleavage) (Haseloff & Gerlach, 1988).

The majority of small self-cleaving RNAs do not need cofactors for their activity and they are active by themselves. This latter characteristic can be a limiting factor, when such ribozymes are used in environments where their activity must be controlled. To respond to this need, allosteric ribozymes were developed (Tang & Breaker, 1997). One method used to develop ribozymes with allosteric (modulator-dependent) activity is to link the ribozyme sequence to an aptamer, which is known to bind efficiently and specifically to a desired target (Gu *et al.*, 2012; Nomura *et al.*, 2012a; Stifel *et al.*, 2019b). In this context, an aptamer will interfere and disturb ribozyme folding to disrupt its activity. Binding of the cofactor to the aptamer sequence will alleviate this interference and subsequently rescue ribozyme activity. The combination of a ribozyme and an aptamer is often called 'aptazyme'. Generally speaking, aptazyme activity depends on the presence or absence of the cognate molecule of the aptamer, where the structure of the aptamer bound to its ligand can either disrupt or stabilize the active structure of the ribozyme module. Aptazymes were used in different applications, such as the control of gene expression and viral vector replication (Ketzer *et al.*, 2012; Ogawa & Maeda, 2008; Rehm *et al.*, 2015; Stifel *et al.*, 2019b; Win & Smolke, 2007b)

HHRz has a conserved structure of three stem-loops with a catalytic core of 17 nucleotides specific to the cleavage site GUH (**Figure 6.2**). To generate a trans-acting version of a HHRZ, stem II and the catalytic core must be conserved, stem I and III are modified to be perfectly complementary to the target substrate (Carbonell *et al.*, 2011). Nevertheless, making a new design cannot be achieved by simply changing the ribozyme sequence; other parameters should be taken into consideration. The ability of a modified sequence to fold into the hammerhead ribozyme's active structure is an essential condition of catalytic activity. Also, the lengths of stems I and III must offer sufficient target specificity, while simultaneously allowing for some ribozyme turnover. Making stems I and III too long will inhibit turnover and excessively short stems raises the risk of reduced activity and increased off-targets effects. Hence, even if manual designs can often be functional, they are much less likely to be optimal. The manual process is also time-consuming and offers limited design choices. A bioinformatics tool taking into account all the important parameters to automatically generate designs would greatly improve this process, reduce

human effort (and error), as well as open the door to many ribozyme and aptazyme applications.

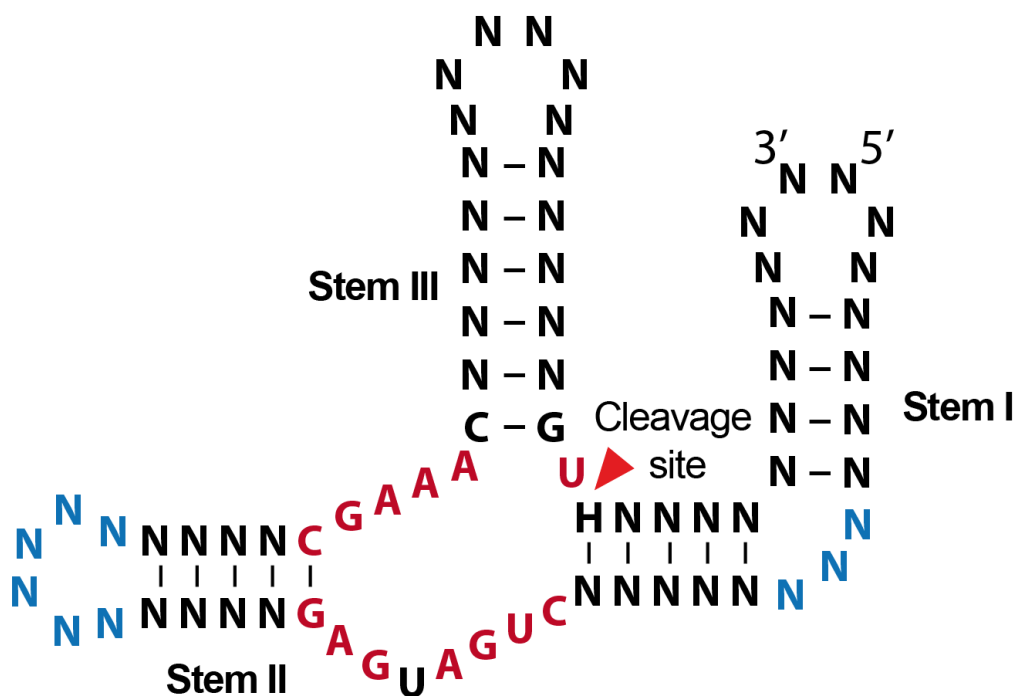


Figure 6.2 Hammerhead ribozyme structure.

Three stem loops surround the conserved catalytic core (red nucleotides). The cleavage site is GUH and the cleavage occur between the U and H. The blue sequences are nucleotides that form a tertiary interaction (a pseudoknot for example) that occurs between stem I and stem II.

In this paper, we present Ribosoft 2.0, an improved version of Ribosoft (Kharmas et al., 2016) for the design of trans-acting ribozymes. The first version allowed the generation of solely the minimal and the full-length versions of hammerhead ribozyme. This new version of the web service offers the ability to design different conventional and allosteric ribozymes, with known secondary structures. As opposed to similar web servers, designs can be made for almost any known ribozyme. Different types of ribozymes can be designed from the web service: different types of HHRZ (extended or pseudoknotted version), pistol, twister, twister sister (Park et al., 2019a; Weinberg et al., 2019b) and aptazymes controlled by FMN (Soukup & Breaker, 1999a). This program is able to generate sequences of active ribozymes against any desired target RNA. During the process of generating new sequences, Ribosoft 2.0 evaluates different parameters. The ribozymes structure is validated by RNA fold (Lorenz et al., 2011) to make sure the new sequence folds into the same structure as the template. The accessibility of the cleavage site is calculated using the target RNA's secondary structure. Finally, the specificity of the

ribozyme sequence (off-targets effects) is evaluated by BLASTing the target sequence on NCBI (Altschul et al., 1990). We used Ribosoft 2.0 to design ribozymes of different types, to target mRNAs encoding Bcl2 and Galectin7 proteins as example targets. The choice of these mRNA is due to the implication of Bcl-2 and Gal-7 in many diseases and in the aggressivity of many cancer types as well as the resistance to cell death and therapies. Our aim is to find efficient ribozymes that can be used to target their mRNAs, which can be used to lower their expression in cancer cells in the goal of slowing down the cells growth or to obtain a better response to chemo- and radiotherapy.

6.4. Materials and methods

- **Software Methodology**

The backend of the Ribosoft 2.0 web service relates to the correct processing, via programmed algorithms, of the biological information provided by the end-user or accessed by the program.

Known as an activity diagram, **Figure 6.3** provides an overview of the whole program, with each vertical section representing a well-defined functionality. The front end, which is best presented as a Graphical User Interface (or GUI) deals with end-user interaction: input from the user (**Figure 6.4**) and ranked design output (**Figure 6.5**). The core information processing modules of Ribosoft 2.0 are all part of the back-end and the algorithm library, each component of these two sections (e.g., generate design or evaluate accessibility) are described, using natural language and pseudo-code below.

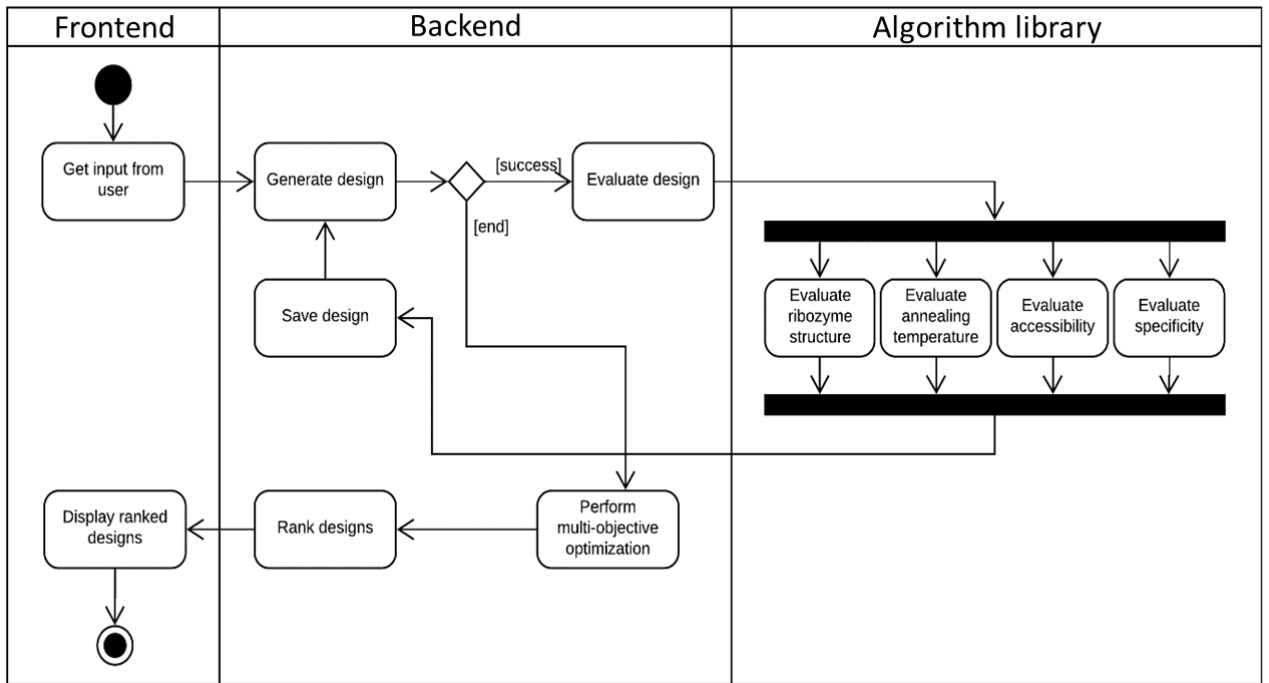


Figure 6.3 Complete overview of Ribosoft (V2) program activity

Create a new request

Ribozyme

Ribozyme structure

RNA sequence

Input sequence [?](#)

CUUGAAGUGGUU...

or [?](#)

Open reading frame start index [?](#)

Open reading frame end index [?](#)

Target regions [?](#)

5'UTR Open reading frame (ORF) 3'UTR

Target environment

In-vitro In-vivo

Environment properties

Temperature (°C)

Na⁺ (mM)

Probe (nM) [?](#)

Advanced settings

Figure 6.4 Input specifications provided by user.

Job details

Request information

Status message None

Ribozyme Theophylline Aptazyme

RNA input

```

    UGGCGCACGUCUGGGAGAACGGGUACGAUAACCGGGAGAUAGUGAUGAAGUACAUCAUUAUAAGCUGUCGCAGAGGGG
    CUACGAGUGGGGAUCGGGAGAUUGGGCGCCGCGCCCGGGGGCCGCCCGCACCGGGCAUCUUCUCUCCCAGCCCCG
    GGCACACGCCCCAUCCAGCCGCAUCCCGGGACCCCGGUCGCCAGGACCUCCGCCGUCGAGACCCCGGCUCCCCCGCGCCG
    CGCGGGGCCUCGCGUCAGCCCGGUGCCACCUUGGUCACCUGACCCUCCGCCAGGCCGGCGGACGACUUCUCCCAGCCGCUA
    
```

Open reading frame start index 0 **Open reading frame end index** 0

5' **Open reading frame** 3'

Target environment In-vivo

Assembly Homo sapiens (taxonomy [9606](#), species [9606](#)) using assembly [GRCh38.p12](#)

Specificity method Cleavage and hybridization

Temperature (°C) 37

Na⁺ (mM) 100

Probe (nM) 0.05

Created 2019-10-07 2:34:22 PM

Advanced settings

Progress

Completed

Designs [Refine Results](#) ▼

▲ Rank	Ribozyme Sequence	Desired Temperature Score	Highest Temperature Score	Specificity Score	Accessibility Score	Structure Score
1	CGGCAGGUCCUGGCCUGAUGAGCCUGGAUACCAGCCGAAAGGCCCU UGGCAGUAGACGAAACGGGUCCCC	187.408	224.408	215.96	23.6	51.472
1	CGGCAGGUCCUGGCCUGAUGAGCCUGGAUACCAGCCGAAAGGCCCU UGGCAGUAGACGAAACGGGUCCCCG	191.367	228.367	230.585	20.8	51.3
1	CGGCAGGUCCUGGCCUGAUGAGCCUGGAUACCAGCCGAAAGGCCCU UGGCAGUAGACGAAACGGGUCCCCG	196.034	233.034	227.936	20.6	51.28
1	CGGCAGGUCCUGGCCUGAUGAGCCUGGAUACCAGCCGAAAGGCCCU UGGCAGUAGACGAAACGGGUCCCCG	199.885	236.885	251.972	20.6	51.202
1	CGGCAGGUCCUGGCCUGAUGAGCCUGGAUACCAGCCGAAAGGCCCU UGGCAGUAGACGAAACGGGUCCCCGGA	200.865	237.865	275.369	20.6	50.82

Figure 6.5 Output designs presented to user

- **Back-end and Algorithm Library**

A. Design Generation

Every instance (specific design) of a general template representing the used ribozyme is generated. Hence, the program discards all ribozyme instances that do not hybridize (strongly enough) to the binding regions around legitimate cleavage-sites on the provided transcript (of the targeted gene/sequence). The instances that remain are then evaluated for four different fitness terms: ribozyme structure, annealing temperature, cut-site accessibility and specificity (to the target transcript and not others within the chosen organism). Hence, the program stores all evaluated designs with their fitness terms in a database.

B. Design Objectives evaluation

B1. Evaluate Ribozyme Structure.

```
Fold ribozyme sequence // this generates multiple foldings
Score = 0
FOR each folding
    Calculate distance between folding and ideal folding
    Score = Score + distance * probability of folding
END FOR
RETURN Score
```

B2. Evaluate Annealing Temperature.

```
Determine limits of binding regions
Score = 0
FOR each binding region
    Calculate annealing temperature of binding region
    Score = Score + annealing temperature
END FOR
RETURN Score
```

B3. Evaluate Accessibility.

```
Fold whole transcript, without constraints
Calculate the minimum free energy of the resulting folding
(MFE)
Determine the limits of the binding regions on the
transcript
```

```
Fold whole transcript with constraints (forcing open all
binding regions)
Calculate the minimum free energy of the constrained
folding (MFE_R)
Delta = MFE_R - MFE
RETURN Delta
```

B4. Evaluate Specificity.

Form a search string using the binding regions of the ribozyme
BLAST the search string for chosen organism to:

(1) find the *number* of affected transcripts

(2) **FOR** each transcript:

find the *percentage* of overlap with the search string

Calculate a *Score* reflecting (1) and (2)

RETURN *Score*

C. Multi-Objective Optimization & Ranking

Multi-objective optimization is the last step in the design process, receiving input from the job and outputs a ranked list of candidates for the user. It is important that this step be quick and accurate, as it is the final step to complete a job. The advised method for ranking is Pareto Dominance, for its simplicity and generality.

- **Generation and selection of ribozymes for experimental assays**

Before designing ribozyme sequences, different types of ribozymes were added to Ribosoft 2.0 and are now available in the dropdown list. This was done by providing models such as that of the pistol ribozyme (**Figure 6.6A**) as well as four different hammerhead ribozymes, including aptazymes (**Figure 6.6** and supplementary material). Afterwards, we used Ribosoft 2.0 as any user should by simply adding our sequence as shown in **Figure 6.6**. Parameters for environment properties that we used as input were: 37°C, 100 mM NaCl and 0.05 nM probe, with 0.05 for advanced settings of all parameters determined by Ribosoft 2.0, this number relates to a chosen range to consider values equivalent with regards to the Pareto Dominance.

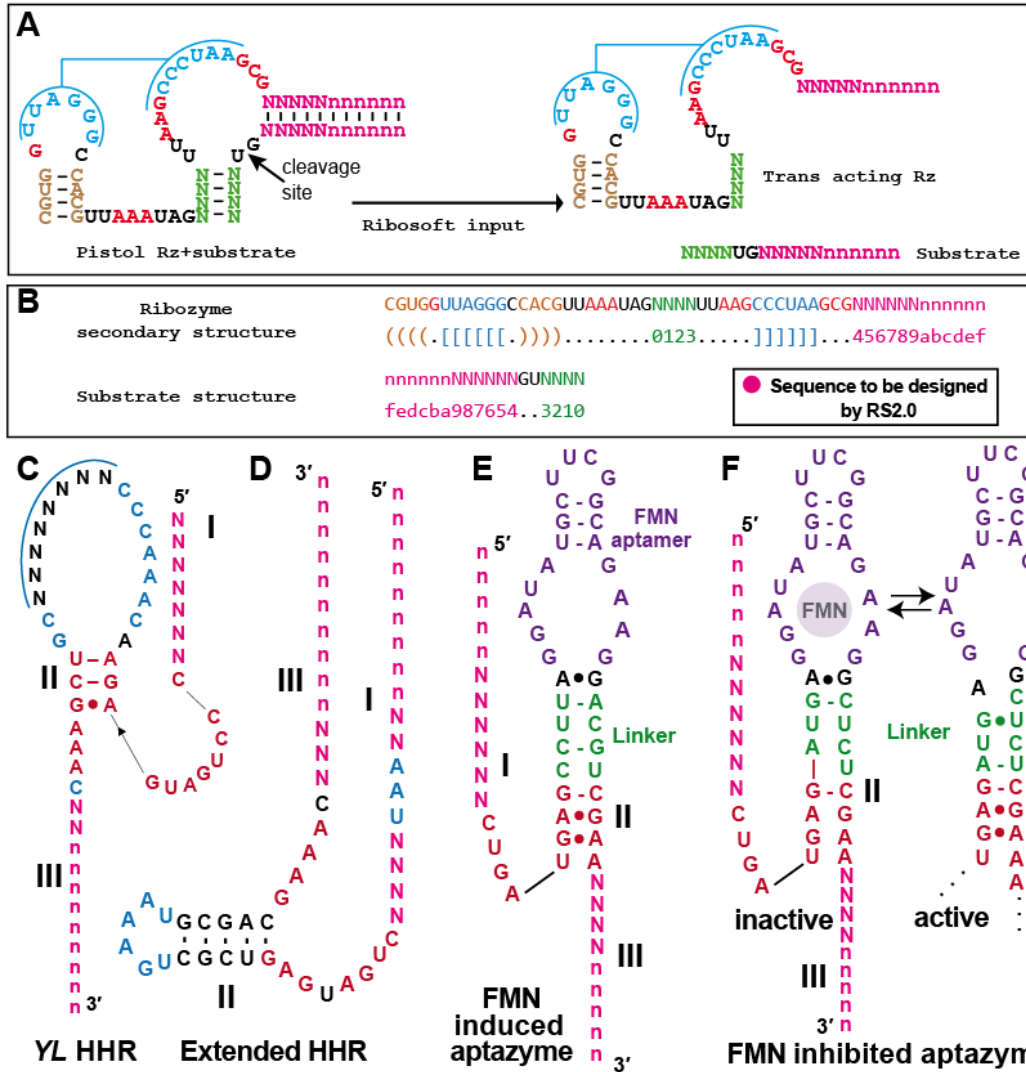


Figure 6.6 *Trans*-acting ribozymes designed by Ribosoft 2.0

(A) Input of new ribozyme models for Ribosoft2.0. Left: Pistol ribozyme structure is formed of three stems and two loops. In red is the conserved sequence of the catalytic core. Green and pink nucleotides are the sequences to be modified by Ribosoft depending on the target sequence. Blue color shows the sequences that form a pseudoknot. Right: Input of Ribosoft 2.0 : Pistol ribozyme secondary structure divided in a ribozyme strand (with the catalytic core) and a substrate strand containing the cleavage site. (B) Each strand structure is given to Ribosoft 2.0 as input where '.' is a single stranded part, '(' is a base pair and '[' is a pseudoknotted part. To figure out the complementarity between ribozyme and substrate strands, the sequences letters and numbers are used in the same order from both sides. (C), (D), (E) and (F) Secondary structure of the ribozyme strands. Templates used by Ribosoft 2.0 to design new sequences (C) HHRZ YL, (D) extended HHRZ, (E) FMN-induced aptazyme and (F) FMN-inhibited aptazyme, with both inactive (FMN-bound, according to (Famulok *et al.*, 2007b), used by Ribosoft to design sequences) and active structures pictured. For all the ribozymes the catalytic core (red nucleotides) and one of the stems was conserved (stem II for all the HHRZ and P1 for the pistol ribozyme). The nucleotides in pink 'Ns' are the single stranded stems that will be designed by RS 2.0 to be complementary to the target sequence surrounding the cleavage site of each ribozyme.

Sixty ribozyme sequences designed by Ribosoft 2.0 were chosen to be assayed. All the ribozyme sequences were chosen to pair the highest and lowest scores calculated by Ribosoft 2.0 when designing the new sequences, for the following Ribosoft 2.0 parameters: 1- “temperature”, 2- “accessibility” and 3- “structure”; “specificity” was not considered as *in vitro* characterization of ribozymes would not provide any relevant data in that regard. Theoretically, for the accessibility and the structures, the higher the score is, the better the ribozyme’s activity is supposed to be. For these ribozymes, 30 were designed to target Gal-7 mRNA and 30 to target Bcl2 mRNA. (Supplementary **Table 6.1**).

- **Preparation of transcription template**

DNA templates used for the ribozyme production were purchased as ssDNA oligos (Alpha DNA or IDT), double stranded DNA templates were produced by primer extension. Oligonucleotides were incubated with a PCR mix containing (100 μ M dNTP, 1X thermopol Taq reaction buffer, 1X of Q solution (Qiagen) and 1 μ M of each primer). A complete list of oligonucleotide sequences is available in supplementary **Table 6.1**).

The expression vector (Pet22b G7h) was used as a template for the amplification of the Gal-7 gene by PCR. The forward primer used harbors the T7 promoter. dsDNA product, was purified and used as template for *in vitro* transcription, using T7 RNA polymerase, to generate Gal-7 mRNA of 419 nucleotides size.

- **Reverse Transcription-PCR**

Total RNA was extracted from Jurkat cells (10^6 cells) using Trizol then treated with DNase I to eliminate the genomic DNA. Bcl2 cDNA was generated using the one-step RT-PCR Qiagen Kit containing 1U of Enzymes mix (Omniscript and Sensiscript Reverse Transcriptases and HOTstart Taq DNA polymerase, provided one-step RT-PCR buffer (5x) Q solution (5x), 2 mM dNTP mix and complete to 50 μ l with RNase free water. Primers used were designed to amplify a specific segment of the Bcl2 cDNA.

- **Transcription**

The ribozymes synthesis was made by *in vitro* transcription in presence of the PCR product for each ribozyme. Each PCR product was incubated with 3 mM rNTPS (2 mM each of rATP, rGTP, rCTP and rUTP), 1X transcription buffer (80 mM HEPES-KOH pH 7.5, 24 mM MgCl₂, 40 mM DTT, 2 mM spermidine), 1 U/ μ l of inorganic pyrophosphatase Roche), USA, 1 U/ μ l RNase inhibitor (Thermofisher), 1 U/ μ l T7 RNA polymerase and Milli-Q water for 2 hours at 37°C.

After reaction completion, 1 U/ μ l of DNase I (RNase free) (NEB) was added to each reaction followed by 30 min incubation at 37°C for the elimination of the DNA templates.

RNA products were extracted with phenol/chloroform, then precipitated with ethanol overnight at -20°C. They were, afterwards, fractionated on a 10% denaturing polyacrylamide gel. The band corresponding to the expected size of each ribozyme was cut and the RNAs were extracted overnight in the elution buffer (EDTA 0.1 mM, NaCl 0.3 M, SDS 0.001%) at 4°C. Finally, the RNAs were precipitated by ethanol, dried, resuspended in water and their concentration measured using a nanodrop spectrophotometer.

To be able to visualize the cleavage activity of the ribozymes, the target RNAs were either externally (5') radiolabeled, when this was not possible, due to target RNA length and structure, we used internally radiolabeled RNA targets made by radioactive transcription.

- **Radioactive transcription**

The radioactive transcription for target RNAs (Gal-7, Bcl2) was done using the same procedure for the cold transcription except for the use of 0.5 μ l alpha-³²P-labeled UTP (3000 Ci/mmol-10 mCi/ml) (Perkin Elmer) and 0.8 mM UTP (instead of 2 mM). After the DNase I treatment, the RNAs were precipitated with ethanol and purified using a 6% denaturing polyacrylamide gel.

- **RNA 5' labeling**

Unlabeled purified RNA targets were dephosphorylated using 1 U antarctic phosphatase (NEB) incubated with 1X of provided buffer during 20 min at 37°C and then 5 min at 65°C to inactivate the enzyme. For labeling, RNA was incubated with 1 U T4 polynucleotide kinase (NEB) in presence of its buffer (70 mM tris-HCl pH 7.6, 10 mM MgCl₂, 5 mM DTT) and γ -³²P-ATP. The reaction was incubated for 1 h at 37°C. The 5'-radiolabeled RNA targets were purified on a 6% denaturing polyacrylamide gel, eluted overnight at 4°C and then precipitated.

- **End point cleavage assays**

Each ribozyme was pre-incubated in the cleavage buffer (50 mM Tris pH 7.5, 100 mM NaCl and 25 mM KCl) at 85°C for 1 min and then snap-cooled on ice. Afterwards, each corresponding target was added to the reaction as well as MgCl₂ (10 mM) to start the reaction. For the FMN activated and FMN inhibited aptazymes 200 μM of FMN is also added to the reaction. The incubation lasted one hour at 37°, and was stopped by adding an equal volume of the 2X stop dye (formamide buffer: 95% formamide, 10 mM EDTA, 0.025% bromophenol blue and 0.025% xylene cyanol blue).

The resulting cleaved RNAs were fractionated on a 6% denaturing polyacrylamide gel. Then the gel was exposed to a storage phosphor screen overnight. The phosphor screen was scanned with a Typhoon FLA9500 (GE Health Care) and the cleavage extent was quantified using ImageQuant.

- **Cleavage time course**

To be able to determine the cleavage efficiency and rate of each ribozyme, we followed the cleavage reaction for 64 minutes and aliquots were taken during this time.

The reaction was prepared in the same way for the end-point cleavage assay, but before the MgCl₂ is added to the reaction an aliquot for the T₀ is taken. Once MgCl₂ is added in the reaction, the timer is started and aliquots are taken at 30s, 1', 2', 4', 8', 16', 32' and 64'. The reaction of each aliquot is stopped with the addition of the stop dye. The results were visualized on 6% denaturing polyacrylamide gels exposed overnight to a storage phosphor screen. The quantification of band intensity was done using ImageQuant and the cleavage rates (k_{obs}) were calculated with GraphPad where the analysis was done with the 'one phase decay' equation.

6.5. Results

To evaluate efficiency and reliability of our web service for RNA design, we validated experimentally the sequences it generates. Ribosoft 2.0 accepts the secondary structure of different types of ribozymes (conventional and allosteric) as input. From these, we have chosen to design 60 ribozymes of five different types. These are the HHRZ with a pseudoknot (from *Yarrowia lipolytica* genome), extended HHRZ (from the negative strand of tobacco ringspot virus (sTRSV) modified by (Saksmerprome et al., 2004)), pistol, as well as FMN activated and FMN-inhibited hammerhead ribozymes. To determine the

efficiency of Ribosoft 2.0 in designing active ribozymes, we targeted these ribozymes against Gal-7 and Bcl2 mRNAs. We also aimed to figure out if there is a correlation between the ribozyme activity and the parameters calculated by Ribosoft 2.0 (for ribozyme structure, accessibility and highest temperature scores). For each parameter, we picked a ribozyme with the highest score and another with the lowest score, in both cases targeting Bcl2 and Gal-7 mRNAs (supplementary **Table 6.2 (S2)**), which encode for an anti-apoptotic protein (Belka & Budach, 2002; Shamas-Din et al., 2013) and a lectin (Vladoiu et al., 2014), respectively, both with pro-cancer impact (Dewson & Kluck, 2010; Yip & Reed, 2008).

6.5.1. Comparison of end-point cleavage of pistol, extended HHRZ and pseudoknotted HHRZ

For hammerhead and pistol ribozymes, 36 ribozymes were tested. Results show that 15 of 18 ribozymes tested against Gal-7 mRNA were active (**Figure 6.7**), with cleavage efficiency ranging from 22 to 93%. In total, close to half of these ribozymes, 7 out of 18, cleaved their designated targets at over 50% efficiency, and similar results were obtained with 18 ribozymes against a 719 bases fragment of Bcl2 mRNA or smaller fragments (examples presented in **Figure 6.10**). These are relatively good results for ribozymes acting on long mRNA-like targets (Bertrand et al., 1997; Langlois et al., 2003; Lévesque et al., 2010; Roy et al., 1999), since less active ribozymes often require more than 1 hour to result in cleavage efficiencies exceeding 50% (Campbell et al., 1997).

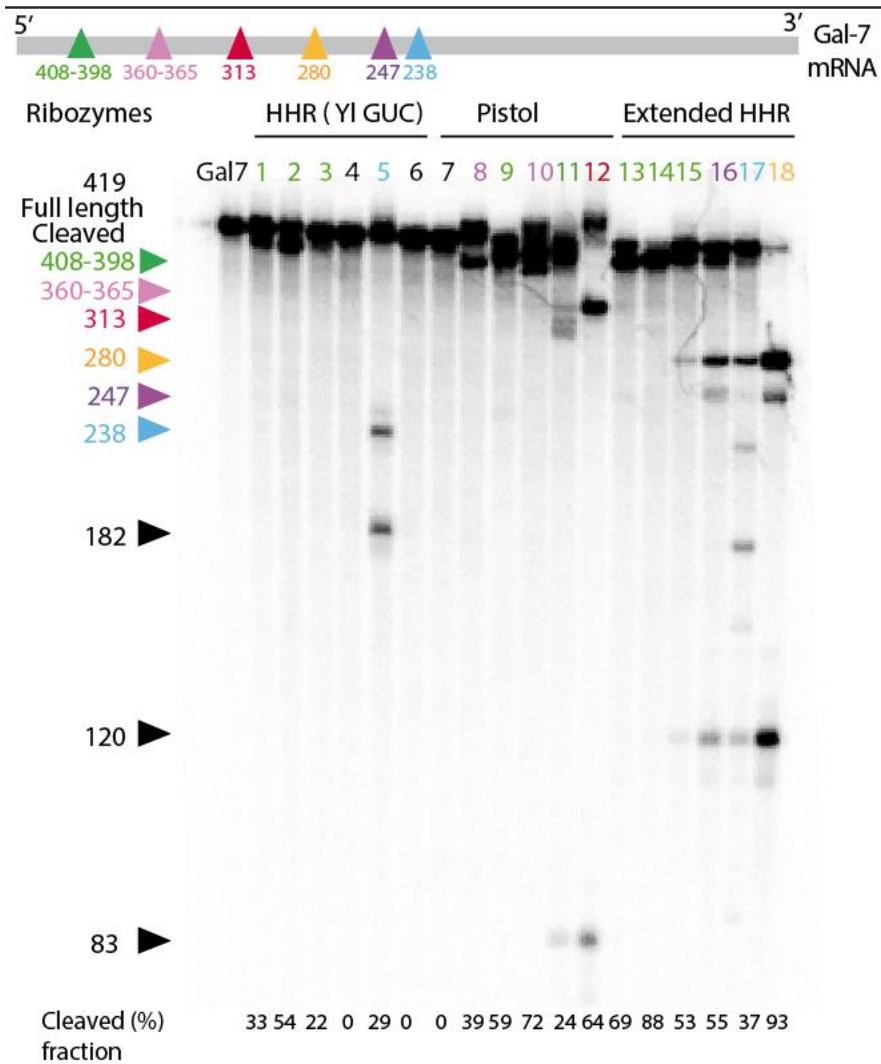


Figure 6.7 End point cleavage assays of ribozymes against Gal-7.

Gal-7 mRNA labelled with alpha-³²P-UTP was incubated with each of the eighteen ribozymes designed by Ribosoft (Pistol and HHRZ ribozymes). The black triangles correspond to the cleaved fragments of each target. The percentages in the bottom of the gel show the cleavage extent of each ribozyme. This cleavage extent is done by calculating the sum of the intensity of the cleaved bands divided by the total of the intensity of the full length and the cleaved bands. For the lanes where more than three bands are observed, the ones below expected band size were eliminated because they are the result of cleavage of a partially degraded RNA. Color code: each color indicates the cleavage of the 5' portion generated from cleavage. The active ribozymes are presented in the same color as their cleavage site

6.5.2. Kinetics measurements

The end-point cleavage assays are good indicators of ribozyme activity, but they are not enough to determine the cleavage rate of a ribozyme. This rate depends on the speed of the reaction, and significant differences can be observed between ribozymes, when comparing them for rate vs cleavage extent. In order to assess the rate constant of each active ribozyme, the kinetics were followed for 64 minutes in the same conditions used for

the end-point cleavage assays. The kinetics measurements were essentially performed in order to compare each pair of ribozymes for a given parameter evaluated by Ribosoft 2.0 (e.g. “accessibility”) with the highest and the lowest score. This would in principle allow us to figure out if there is a correlation between those parameters and the cleavage efficiency of a ribozyme.

The kinetics were followed as well (all k_{obs} are in **Table 6.2**). The cleavage rates (k_{obs}) were as high as 1.8 min⁻¹ and most active ribozymes cleaved at > 0.2 min⁻¹. Given these results, we could see little to no correlation between ribozyme type and ribozyme activity (**Figure 6.8**), or with any of the other scores determined by Ribosoft 2.0 because ribozymes with the highest k_{obs} are distributed across different types and different accessibility/ temperature/structure scores.

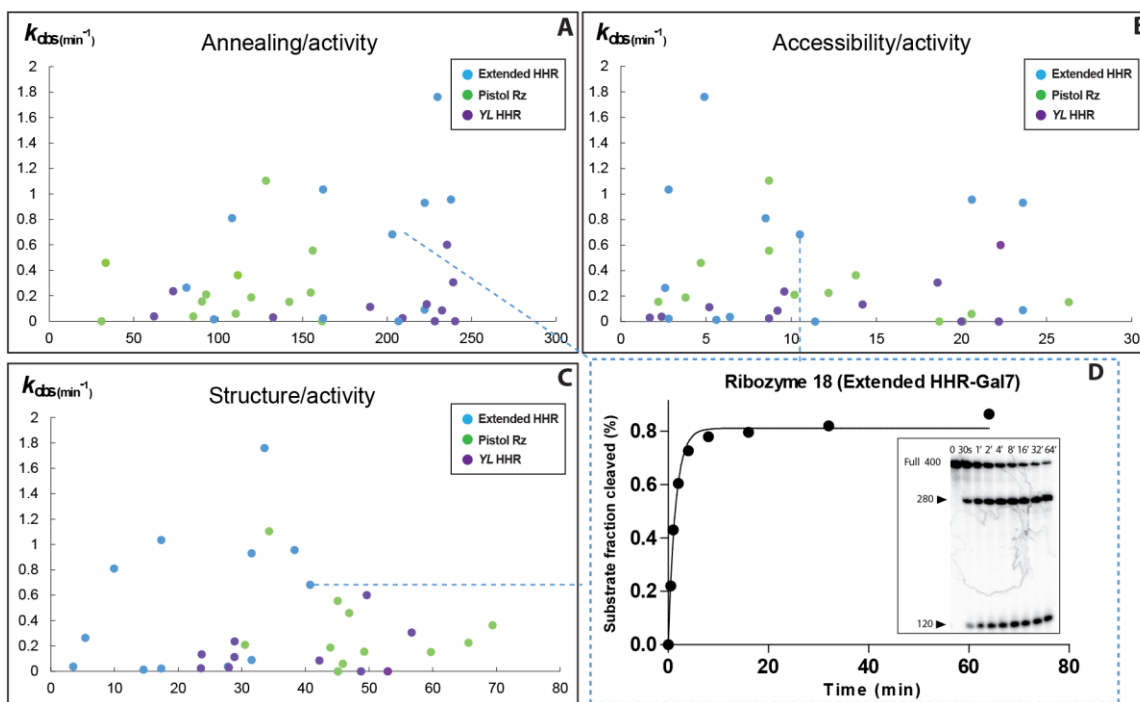


Figure 6.8 Scatter plots showing the k_{obs} of each ribozyme compared to the score of each one of the parameters given by Ribosoft 2.0

(A) The temperature score relevant for annealing of the ribozyme binding arms to its complementary target. (B) The accessibility score calculated according to accessibility of the cleavage site depending on the secondary structure of the target and (C) the structure score calculated according to the structure formed by the new ribozyme sequence. In all three cases, the units of the X axis are arbitrary and correspond to their respective scores. (D) Example of a one phase decay curve (from ribozyme 18) used to calculate the k_{obs} of each ribozyme. The insert shows the corresponding gel of ribozyme kinetics used to measure the cleavage extent of the target during a time course over 64 minutes.

6.5.3. FMN Aptazymes

For FMN activated and FMN inhibited aptazymes, twelve ribozymes targeting Gal-7 mRNA were tested in the presence and absence of 200 μ M of FMN. For the FMN-activated version of the aptazyme, one of the six ribozymes was active under both conditions (presence and absence of FMN), with cleavage efficiency 57%, while five were inactive. As for the six FMN-inhibited ribozymes, four of the six ribozymes tested were active in the absence of FMN (giving cleavage efficiencies of 10, 13, 38 and 58%), while their activity was completely suppressed when FMN was added to the reaction (**Figure 6.9**). Although only four ribozymes have shown the desired activity in presence/absence of FMN, the result still indicates that Ribosoft 2.0 was not only able to design different trans-acting ribozymes, but for the first time, design inducible trans-acting ribozymes as well.

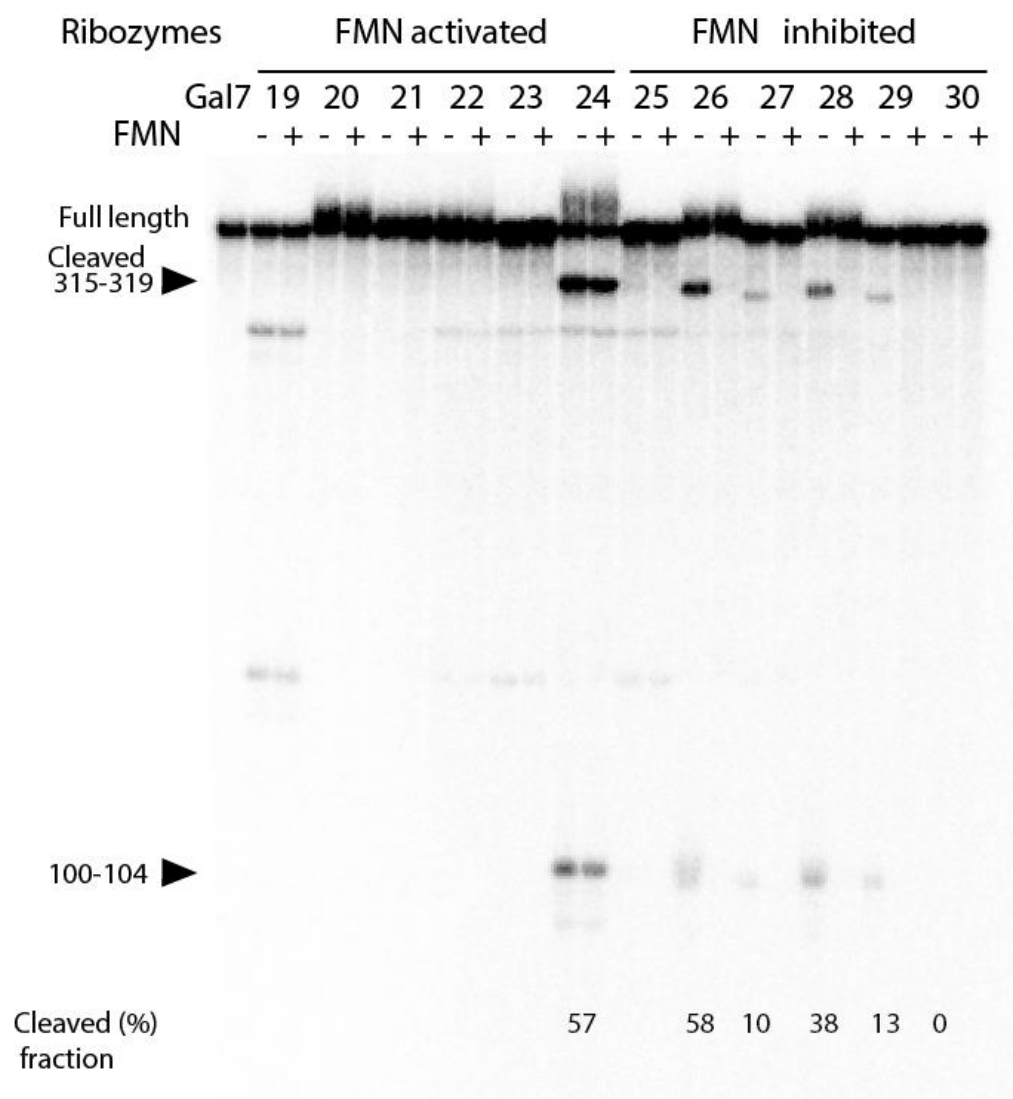


Figure 6.9 End point cleavage tests of activated and inhibited version of *trans* acting FMN aptazymes against Gal-7 mRNA in presence and absence of FMN

The activated version of the aptazyme did not show any activity except for aptazyme 19 that is active in presence and absence of FMN. The inhibited version: four aptazymes (26, 27, 28 and 29) have shown the desired activity where they were active in absence of FMN and they lost completely their activity when FMN was present in the reaction solution.

6.6. Discussion

Ribosoft 2.0 is a web service with a user-friendly interface offering an easy and simple way to design specific types of ribozymes and, if desired, allows the entry of advanced parameter values. Our results show that the service is an efficient tool for the generation of different types of active ribozymes and aptazymes targeting any desired RNA in an automated fashion, quickly and correctly. This renders Ribosoft 2.0 as the only publicly available web service with such facility. There is no other bioinformatics tool allowing for the generation of this number of different structures simultaneously. To our knowledge, the only tool handling more than one type of ribozyme is 'Ribosubstrate' (Lucier *et al.*, 2006a) but it does not automatically design ribozymes. Rather, it requires prior determination of the target site and ribozyme sequence and then uses BLAST to look for off-targets of SOFA-HDV, hammerhead ribozymes or siRNAs. Ribosoft 2.0 has advantages over 'Aladdin' (Mercatanti *et al.*, 2012), a bioinformatics tool that generates solely hammerhead ribozyme designs, relying on the target sequence as the only input. Moreover, the only parameter calculated with this tool is the annealing temperature (T_m) of right and left arms of the ribozyme; it does not check for off-targets. In this case, each of the mentioned programs misses an important part of the design, either the specificity parameter or the automation of the design. In our case, we developed a tool combining the options given by both programs, and with advantageous extensions. Many of these extensions were already provided by the first version of Ribosoft (Kharma *et al.*, 2016) that we used to generate HHRZ against the mRNA of PABPN1, which were active *in vitro*, in mammalian cells and in *C. elegans* (Abu-Baker *et al.*, 2019).

In the new version of the web service as well as the old one, the design is made, taking in account the specificity of a ribozyme by blasting sequences in NCBI. In addition, Ribosoft 2.0 computes three other parameter. The accessibility of the cleavage sites is calculated for all sites compatible with the sequence constraints provided by the ribozyme model. The program uses the secondary structure of the target to check for single stranded regions. To reduce "computational costs" (prohibitive for very large structures), the structure is predicted by multiple sequence windows of (check size), this makes prediction faster with a reasonable trade-off, since the vast majority of secondary structures are local and do not occur between regions separated by several hundred bases. It also checks for the quality of the newly designed ribozyme structure and, lastly, it provides a score reflecting the annealing temperature between the target and the ribozyme's arms.

Regarding a relationship between scores calculated by Ribosoft 2.0 and ribozyme activity, since the software already screens out all ribozymes likely to be inactive, it excludes ribozymes from a wide range of parameter values and predicted efficiency, hindering our attempt to evaluate such a correlation. Instead, we proved that most ribozymes designed by Ribosoft 2.0 (except aptazymes) were active (**Figure 6.7 and 6.10**), i.e. the program has generated 8-9 active ribozymes out of every 10 designs, independently of parameter values. Actually, all of the chosen ribozymes were ranked #1 by Ribosoft, so this ranking was not particularly useful. This can be explained by the filters of Ribosoft 2.0, which already eliminate all the ribozymes having minor defaults and lower scores for ribozymes. In the end, this sorts through sequences to select only the highly active ribozymes, even those with the lowest scores.

The main addition to the Ribosoft 2.0 software is the possibility of adding almost any type of ribozyme. It is even possible to include the presence of a pseudoknot within the “ribozyme strand”, although folding of pseudoknots is not formally computed. As for pseudoknots between the “substrate” and “ribozyme” strands, such as for the validated HHRZs from *Y. lipolytica*, they are taken into account via the hybridization of each binding arm. To our knowledge, Ribosoft 2.0 allows for the first time to automatically design allosteric ribozymes against desired targets. Indeed, we generated the first trans-acting aptazymes and demonstrated activity of the FMN-inhibited HHRZ aptazymes in presence of FMN. Known examples of trans-acting aptazymes act in the following way: the output of the cleavage reaction works as the effector or an input of another reaction. As an example, a guanine aptazyme that generates a pri-miRNA as an output acts in combination with the RISC complex to degrade a target mRNA (Nomura *et al.*, 2012b). In our case, the aptazyme works directly as a degradation machine of the target RNA in a sequence specific manner. This is a novelty provided by our software giving the users the opportunity to have on the same RNA molecule a detector of a specific molecule and an effector against a specific target sequence, which we can easily modulate depending on the target. This option will not be limited only to aptazymes, but it can also work for other types of allosteric ribozymes controlled by the presence of a certain input like an RNA or DNA sequence. Despite the fact that our results show only four active aptazymes corresponding to the given condition, they still indicate that for a few easily designed aptazymes users are very likely to get design sequences fitting their needs. Indeed, four out of six of the FMN-inhibited aptazymes against Gal-7 mRNA function as expected. As for the FMN-activated aptazymes, the fact that none of them worked may indicate that the

structure model described by Soukup and Breaker (Soukup & Breaker, 1999c) does not correspond exactly to the active structure of the aptazyme, which would lead Ribosoft 2.0 to inappropriately score the structure of the ribozyme. These RNAs can be used in many applications like the replacement of the cis-acting aptazymes included upstream or downstream of the target gene in many previous examples of aptazyme use in synthetic biology (Nomura *et al.*, 2012a; Townshend *et al.*, 2015; Yokobayashi, 2019b; Zhong *et al.*, 2016a). By having trans-acting aptazymes, the function of aptazymes can be applied to prokaryotic and eukaryotic cells, including mammalian and human cells, without the need of combining different macromolecules (cis-acting aptazyme and RISC complex with miRNA or siRNA).

The results we got from aptazymes were limited because all of the aptazymes against Blc2 mRNA were inactive (data not shown). Also some aptazymes against Gal-7 mRNA were active but did not show any difference in activity in the presence or absence of FMN. A possible explanation is that these aptazymes were developed using minimal HHRZs which are 100 fold less active than full-length HHRZs. Besides, the input secondary structure given to Ribosoft 2.0 for the design of these aptazymes is only representative of the active state of the aptazyme in presence or absence of a specific ligand. The structure for the other state (i.e. inactive state) might be an important information that Ribosoft 2.0 needs, but it is missing. Both structure states will be needed to make the comparison between the two states, and design ribozymes with properly optimized structures. In that regard, the native conventional HHRZ is the most studied catalytic RNA and it has a well characterized structure, contrary to aptazymes which are built based on aptamers selected *in vitro* and usually having poorly characterized structures. In other words, the biggest limitation of Ribosoft 2.0 for the design of aptazymes might not be the algorithm underlying the software, but rather the lack of comprehensive characterization of the structure of most selected aptamers and aptazymes as well as the secondary structure of the target. This can be circumvented by using well-studied examples, by characterizing aptamers/aptazymes that bind ligands of interest, or simply by including information on secondary structure only for the “ribozyme portion” and leaving the aptamer portion as if it was single-stranded. Nevertheless, the published FMN aptazyme structures allowed us to demonstrate the applicability of Ribosoft 2.0 for aptazymes. In the future, any aptazyme or other allosteric ribozymes with a known structure can be used as input to Ribosoft 2.0, as the software allows for future addition to new ribozyme templates.

Acknowledgements

The authors wish to thank Atef Nehdi for helpful discussion.

Funding

S. Najeh received fellowships from Natural Sciences and Engineering Council of Canada (NSERC) SynBioApps program [CREATE-511601-2018], Fondation Armand-Frappier and Fonds de Recherche du Québec-Nature-Technologie (FRQNT). J.P. is a junior 2 FRQS research scholar. Other funding, including for open access charge: NSERC [RGPIN-2019-06403].

6.7. Supplementary material

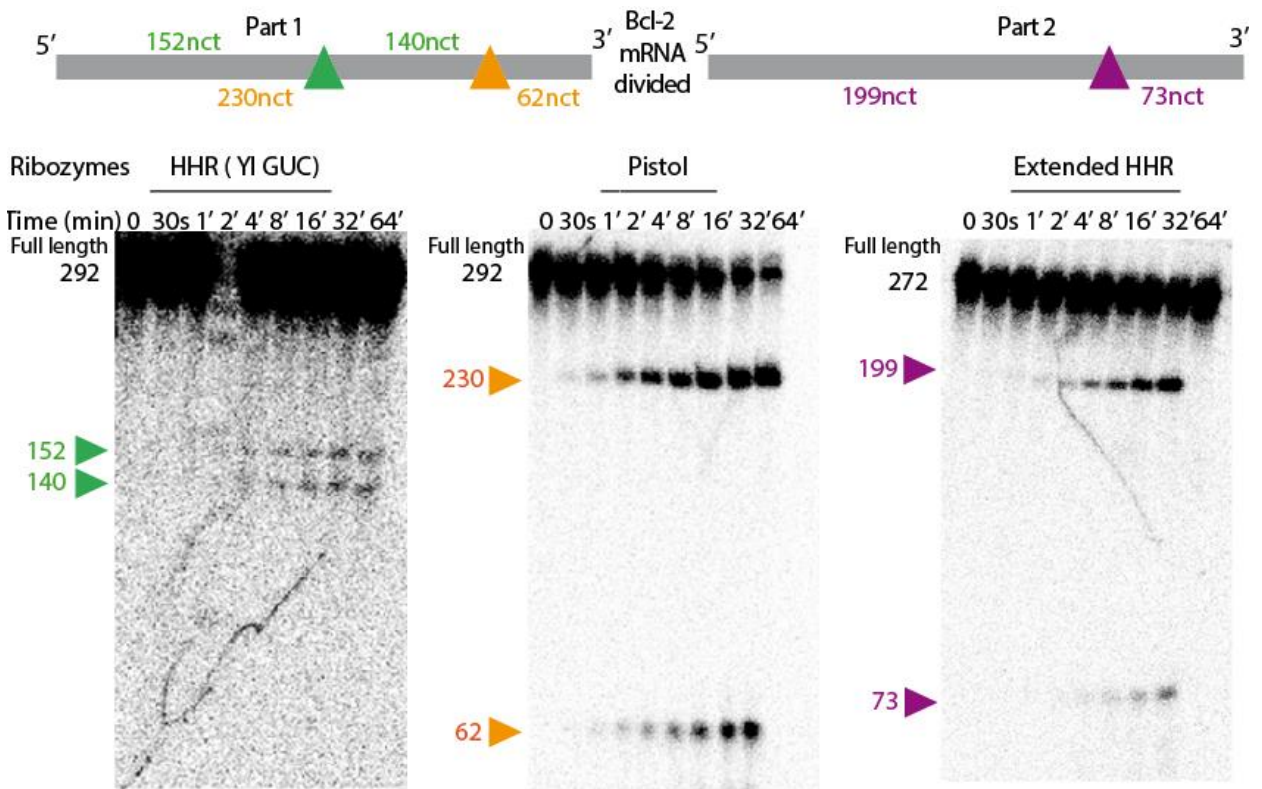


Figure 6.10. Examples of Bcl-2 mRNA cleavage results

Parts of Bcl-2 mRNA were used to test the ribozymes activity. One example of kinetics gel for each ribozyme type is presented (YL GUC-Pistol-Extended HHR from left to right). YL GUC ribozyme and pistol ribozyme cleave in different sites on the same mRNA part and the extended HHR cleaves in a different region.

Ribozyme models

Yarrowia lipolytica hammerhead ribozyme (cis version has a pseudoknot)

Rz template:

```

NNNNNNNCUGAUGAGAACAAACCCNNNNNNNCGUCGAAACNNNNNNNNNN
01234567.....((.....89abcdef..))...ghijklmnopqrstu
  
```

Substrate template:

```

nnnnnnnnnnNNGUCGNNNNNNNNNNNNNNNN
utsrqponmlkjihg.76543210fedcba98
  
```

Cut site: 16

(after 16th character of the template)

Cut site: 5

(after 5th character of the template)

Notes:

- 1-Sequence that corresponds to the pseudoknot in the *cis* version is highlighted in pink.
- 2-The base pair in red is not watson-crick (A-G), so it could simply be considered as two conserved nucleotides with no base-pairing, the other two base pairs (in turquoise) are part of stem II.
- 3-The nucleotides for which identity is absolutely essential are in italic. If for some reason it is convenient to change identity of other nucleotides (while preserving structure), it is possible.
- 4-Sequence in 5' of stem I (in gray), as seen in the initial model sequence, is not important, but for in vitro transcription, it is normally required to add "GGG"
- 5-For a GUC cleavage site, we should add a fixed "C" after the A at the beginning of stem III (in green) and another one just after the "U" in the substrate.

Extended Hammerhead ribozyme

Ribozyme template:

nnnnnnnnNNAAU**NNNNNN**CUGAUGA**GUCGC**UGAAAU**GCGAC**GAA**AC**NNNNnnnnnnnnnn
 0123456789...abcde.....((((.....)))...fghijklmnopqrst

Substrate template:

nnnnnnnnnnNNNGUCNNNNNNNnnnnnnnnn
 tsrqponmlkjihgf.edcba9876543210

Cut-site: 16
(after the base 16 on the substrate template)

Pistol ribozyme

Rz template:

CGUGG**UUAGGG****CCACG**UUAUAUAG**NNNN**UUAAG**CCCUAA**GCG**NNNNNN**nnnnnnn
 ((((. [[[[[(.))))) 0123]]]]] 456789abcdef

Substrate template:

nnnnnnnnNNNNNN**GU****NNNN**
 fedcba987654..3210

Cut-site number: 13
(meaning between base 13 and 14 of template)

Hammerhead induced by FMN :

Rz template :

nnnnnnNNNNNNCUGAUGA**GCCUUAGGAUAUGCUUCGGCAGAAGGACGUC**GAA**AC**NNNNnnnnnn
 0123456789a.....(.(.....(((.....))).....).)....bcdefghijkl

Target template :

nnnnnNNNNGUANNNNNNnnnnn
lkjihgfedcb.a9876543210

cut site= 12

Hammerhead inhibited by FMN:

Rz template:

nnnnnNNNNNNCUGAUGA GAUGAGGAUAUGCUUCGGCAGAAGGCUCUCGAAACNNNNnnnnn
0123456789a.....(..(.....(((.....))).....)...)...bcdefghijkl

Target template:

nnnnnNNNNGUANNNNNNnnnnn
lkjihgfedcb.a9876543210

cut site= 12

Note: As for extended HHR and *Yarrowia lipolytica* HHR, stem I, II (or linker between aptamer and HHR) and III are highlighted in gray, turquoise and green, respectively. The FMN aptamer is highlighted in yellow.

Table 6.1 Oligonucleotides sequences

Rz #	Ribozymes	Oligo sequence
1	YL GUC1 (Gal-7)	AAGTTTCGACGCCGAGGGCGGGTTTGTTCATCAGCTCACTGCCCTATAGTGAGTCGTATTA
2	YL GUC2 (Gal-7)	ACGTCCCCCACAAGTTTCGACGCCGAGGGCGGGTTTGTTCATCAGCTCACTGCCCTATAGTGAGTCGTATTA
3	YL GUC3 (Gal-7)	CAAGTTTCGACGCCGAGGGCGGGTTTGTTCATCAGCTCACTGCCCTATAGTGAGTCGTATTA
4	YL GUC4 (Gal-7)	CCCGGCTGGACACGTTTCGACGCTTCAACGGGTTTGTTCATCAGGGAGGTGGCCCTATAGTGAGTCGTATTA
5	YL GUC5 (Gal-7)	CGGAGGTGGTTCGACGCAAGGAGCGGGTTTGTTCATCAGTTCAACAGCCCTATAGTGAGTCGTATTA
6	YL GUC6 (Gal-7)	CCCGGCTGGACACGTTTCGACGCTTCAACGGGTTTGTTCATCAGGGAGGTGGCCCTATAGTGAGTCGTATTA
7	Pistol1 (Gal-7)	CAGGTTCCATCGCTTAGGGCTTAAAAACCTATTTAACGTGGCCCTAACCCAGCCCTATAGTGAGTCGTATTA
8	Pistol2 (Gal-7)	CGTGCCTGGTGGAGCGCTTAGGGCTTAAGGGCTATTTAACGTGGCCCTAACCCAGCCCTATAGTGAGTCGTATTA
9	Pistol3 (Gal-7)	TCCCCCACAACGCTTAGGGCTTAACCTCCTATTTAACGTGGCCCTAACCCAGCCCTATAGTGAGTCGTATTA
10	Pistol4 (Gal-7)	GGAGGTGGGCGGGACCGCTTAGGGCTTAAGCAGCTATTTAACGTGGCCCTAACCCAGCCCTATAGTGAGTCGTATTA
11	Pistol5 (Gal-7)	CATGTCCAACCGCTTAGGGCTTAACCCCTATTTAACGTGGCCCTAACCCAGCCCTATAGTGAGTCGTATTA
12	Pistol6 (Gal-7)	TGGTTGGGACGCCACGCTTAGGGCTTAAACCACTATTTAACGTGGCCCTAACCCAGCCCTATAGTGAGTCGTATTA
13	Extend1 (Gal-7)	CAAGTTTCGTCGCATTTTCAGCGACTCATCAGCTCACATTTGCCCTATAGTGAGTCGTATTA
14	Extend2 (Gal-7)	ACGTCCCCCACAAGTTTCGTCGCATTTTCAGCGACTCATCAGCTCACATTTGCCCGAGGGCCCTATAGTGAGTCGTATTA
15	Extend3 (Gal-7)	ATTATAAGCTGTTTCGTCGCATTTTCAGCGACTCATCAGGCAGAATTGGCCCTATAGTGAGTCGTATTA
16	Extend4 (Gal-7)	GGGACCCGGTTCGTCGCATTTTCAGCGACTCATCAGGCCAGATTGACCTCGCCCCCTATAGTGAGTCGTATTA
17	Extend5 (Gal-7)	GTGGTTTCGTCGCATTTTCAGCGACTCATCAGTTCAAATTCACCCTATAGTGAGTCGTATTA
18	Extend6 (Gal-7)	TGCTCATCATCGCGTTTCGTCGCATTTTCAGCGACTCATCAGAGACGATTACGGCCCTATAGTGAGTCGTATTA
19	FMN act1 (Gal-7)	CCATGTTTCGACGTCCTTCTGCCGAAGCATATCCTAAGGCTCATCAGAACCTGCCCTATAGTGAGTCGTATTA
20	FMN act2 (Gal-7)	GGACGCCAGTTTCGACGTCCTTCTGCCGAAGCATATCCTAAGGCTCATCAGCCACCACTTCCCCCTATAGTGAGTCGTATTA
21	FMN act3 (Gal-7)	CCCAGTTTCGACGTCCTTCTGCCGAAGCATATCCTAAGGCTCATCAGCCACCACTTCCCCCTATAGTGAGTCGTATTA
22	FMN act4 (Gal-7)	GACGCCAGTTTCGACGTCCTTCTGCCGAAGCATATCCTAAGGCTCATCAGCCACCACTTCCCCCTATAGTGAGTCGTATTA
23	FMN act5 (Gal-7)	CCCAGTTTCGACGTCCTTCTGCCGAAGCATATCCTAAGGCTCATCAGCCACCACTTCCCCCTATAGTGAGTCGTATTA
24	FMN act6 (Gal-7)	AGGTTCCATGTTTCGACGTCCTTCTGCCGAAGCATATCCTAAGGCTCATCAGAACCTGCCCTATAGTGAGTCGTATTA
25	FMN inh1 (Gal-7)	CCATGTTTCGAGAGCCTTCTGCCGAAGCATATCCTCATCTCATCAGAACCTGCCCTATAGTGAGTCGTATTA
26	FMN inh2 (Gal-7)	GGACGCCAGTTTCGAGAGCCTTCTGCCGAAGCATATCCTCATCTCATCAGCCACCACTTCCCCCTATAGTGAGTCGTATTA
27	FMN inh3 (Gal-7)	CCCAGTTTCGAGAGCCTTCTGCCGAAGCATATCCTCATCTCATCAGCCACCACTTCCCCCTATAGTGAGTCGTATTA

28	FMN inh4 (Gal-7)	GACGCCAGTTTTCGAGAGCCTTCTGCCGAAGCATATCCTCATCTCATCAGCCACCACTTCCCTATAGTGAGTCGTATTA
29	FMN inh5 (Gal-7)	CCCAGTTTTCGAGAGCCTTCTGCCGAAGCATATCCTCATCTCATCAGCCACCACTTCCCTATAGTGAGTCGTATTA
30	FMN inh6 (Gal-7)	AGGTTCCATGTTTTCGAGAGCCTTCTGCCGAAGCATATCCTCATCTCATCAGAACCTGCCCTATAGTGAGTCGTATTA
31	YL GUC1b (Bcl2)	ATGTTTTCGACGTGCACCTGGGGTTTTGTTCTCATCAGCAGCCAGCCCTATAGTGAGTCGTATTA
32	YL GUC2b (Bcl2)	TCCCGGGACCCGGTTTTCGACGCTCGCCGGGGTTTTGTTCTCATCAGGCCAGGACCCCTATAGTGAGTCGTATTA
33	YL GUC3b (Bcl2)	ATAAGCTGTTTTCGACGGCTACGAGGGGTTTTGTTCTCATCAGGCAGAGGGCCCTATAGTGAGTCGTATT A
34	YL GUC4b (Bcl2)	GGTGCCACCTGTGGTTTTCGACGCCTCCGCCGGGTTTTGTTCTCATCAGCACCTGACCCCTATAGTGAGTCGTATTA
35	YL GUC5b (Bcl2)	TGAGTTCCGGTGGGGTTTTCGACGGGAGAGCGGGGTTTTGTTCTCATCAGATGTGTGTCCCTATAGTGAGTCGTATTA
36	YL GUC6b (Bcl2)	ATCCCGGGACCCGGTTTTCGACGCTCGCCGGGGTTTTGTTCTCATCAGGCCAGGACCCCTATAGTGAGTCGTATTA
37	Pistol1b (Bcl2)	TAGTGATGAACGCTTAGGGCTTAAACATCTATTTAACGTGGCCCTAACCCAGCCCTATAGTGAGTCGTATTA
38	Pistol2b (Bcl2)	GTGGGATGCGGGAGATCGCTTAGGGCTTAAAGGGCTATTTAACGTGGCCCTAACCCAGCCCTATAGTGAGTCGTATTA
39	Pistol3b (Bcl2)	TCTCCTGGCTCGCTTAGGGCTTAACTCTCTATTTAACGTGGCCCTAACCCAGCCCTATAGTGAGTCGTATTA
40	Pistol4b (Bcl2)	GCTCTTCAGGGACGGGCGCTTAGGGCTTAAAGAACCTATTTAACGTGGCCCTAACCCAGCCCTATAGTGAGTCGTATTA
41	Pistol5b (Bcl2)	ACATCGCCCTCGCTTAGGGCTTAAAGGATCTATTTAACGTGGCCCTAACCCAGCCCTATAGTGAGTCGTATTA
42	Pistol6b (Bcl2)	CGCATCCCGGGACCCGGCTTAGGGCTTAAAGCCCTATTTAACGTGGCCCTAACCCAGCCCTATAGTGAGTCGTATTA
43	Extend1b (Bcl2)	GCTGTTTTCGTCGCATTTTCAGCGACTCATCAGTCTGAATTAGCCCTATAGTGAGTCGTATTA
44	Extend2b (Bcl2)	TCCCGGGACCCGGTTTTCGTCGCATTTTCAGCGACTCATCAGGCCAGATTGACCTCGCCGCCCTATAGTGAGTCGTATTA
45	Extend3b (Bcl2)	ATTATAAGCTGTTTTCGTCGCATTTTCAGCGACTCATCAGGCAGAATTGGCCCTATAGTGAGTCGTATTA
46	Extend4b (Bcl2)	GGGACCCGGTTTTCGTCGCATTTTCAGCGACTCATCAGGCCAGATTGACCTCGCCCCCTATAGTGAGTCGTATTA
47	Extend5b (Bcl2)	GATGTTTTCGTCGCATTTTCAGCGACTCATCAGGCCCATTTGCCCTATAGTGAGTCGTATTA
48	Extend6b (Bcl2)	TTCCCGGAGATGTTTTCGTCGCATTTTCAGCGACTCATCAGCAGCCATTAGCTGCCCTATAGTGAGTCGTATTA
49	FMN act1 (Bcl2)	TGAAGTTTTCGACGTCTTCTGCCGAAGCATATCCTAAGGCTCATCAGCATCCACCCTATAGTGAGTCGTATTA
50	FMN act2 (Bcl2)	TGTGGAAGTGTTCGACGTCTTCTGCCGAAGCATATCCTAAGGCTCATCAGCGGCCCCAGCACCCCTATAGTGAGTCGTATTA
51	FMN act3 (Bcl2)	TGATGAAGTTTTCGACGTCTTCTGCCGAAGCATATCCTAAGGCTCATCAGCATCCACCCTATAGTGAGTCGTATTA
52	FMN act4 (Bcl2)	GTGGAAGTGTTCGACGTCTTCTGCCGAAGCATATCCTAAGGCTCATCAGCGGCCCCCTATAGTGAGTCGTATTA
53	FMN act5 (Bcl2)	ACTGAGTTTTCGACGTCTTCTGCCGAAGCATATCCTAAGGCTCATCAGCCTGAACCCTATAGTGAGTCGTATTA
54	FMN act6 (Bcl2)	GGAAGTGTTCGACGTCTTCTGCCGAAGCATATCCTAAGGCTCATCAGCGGCCCCAGCCCTATAGTGAGTCGTATTA
55	FMN inh1 (Bcl2)	TAATACGACTCACTATAGGGTTTCAGGCTGATGAGCCTTAGGATATGCTTCGGCAGAAGGACGTCGAAACTCAGT
56	FMN inh2 (Bcl2)	TGTGGAAGTGTTCGAGAGCCTTCTGCCGAAGCATATCCTCATCTCATCAGCGGCCCCAGCACCCCTATAGTGAGTCGTATTA

57	FMN inh3 (Bcl2)	TGATGAAGTTTTCGAGAGCCTTCTGCCGAAGCATATCCTCATCTCATCAGCATCCATTACCCTATAGTG AGTCGTATTA
58	FMN inh4 (Bcl2)	GTGGAAGTGTTCGAGAGCCTTCTGCCGAAGCATATCCTCATCTCATCAGCGGGCCCCCTATAGTGAG TCGTATTA
59	FMN inh5 (Bcl2)	ACGGGGTTTTCGAGAGCCTTCTGCCGAAGCATATCCTCATCTCATCAGCGATAACCCTATAGTGAGTCG TATTA
60	FMN inh6 (Bcl2)	ATGACTGAGTTTTCGAGAGCCTTCTGCCGAAGCATATCCTCATCTCATCAGCCTGAACCGGGCCCCCTATA GTGAGTCGTATTA
61	For Gal-7	TAATACGACTCACTATAGGATGTCCAACGTCCCCACAAG
62	Rev Gal-7	TCAGAAGATCCTCACGGAGTC
63	For Bcl2	TAATACGACTCACTATAGGTGGCGCACGCTGGGAGAACG
64	Rev Bcl2	TCACTTGTGGCCCAGATAGG

Table 6.2 Parameters scores of chosen ribozymes

Rz #	Ribozymes	Parameters	Score
1	YL GUC1 (Gal-7)	Tm min	62.055
2	YL GUC2 (Gal-7)	Tm max	232.478
3	YL GUC3 (Gal-7)	Accessibility min	1.7
4	YL GUC4 (Gal-7)	Accessibility max	22.2
5	YL GUC5 (Gal-7)	Structure min	23.593
6	YL GUC6 (Gal-7)	Structure max	48.722
7	Pistol1 (Gal-7)	Tm min	30.933
8	Pistol2 (Gal-7)	Tm max	161.474
9	Pistol3 (Gal-7)	Accessibility min	3.8
10	Pistol4 (Gal-7)	Accessibility max	26.3
11	Pistol5 (Gal-7)	Structure min	34.292
12	Pistol6 (Gal-7)	Structure max	69.411
13	Extend1 (Gal-7)	Tm min	81.307
14	Extend2 (Gal-7)	Tm max	229.866
15	Extend3 (Gal-7)	Accessibility min	2.8
16	Extend4 (Gal-7)	Accessibility max	23.6
17	Extend5 (Gal-7)	Structure min	3.525
18	Extend6 (Gal-7)	Structure max	40.736
19	FMN act1 (Gal-7)	Tm min	98.405
20	FMN act2 (Gal-7)	Tm max	207.06
21	FMN act3 (Gal-7)	Accessibility min	8.4
22	FMN act4 (Gal-7)	Accessibility max	22.4
23	FMN act5 (Gal-7)	Structure min	29.651
24	FMN act6 (Gal-7)	Structure max	52.461
25	FMN inh1 (Gal-7)	Tm min	98.405
26	FMN inh2 (Gal-7)	Tm max	207.06
27	FMN inh3 (Gal-7)	Accessibility min	8.4
28	FMN inh4 (Gal-7)	Accessibility max	22.4
29	FMN inh5 (Gal-7)	Structure min	29.478
30	FMN inh6 (Gal-7)	Structure max	51.639
31	YL GUC1b (Bcl2)	Tm min	73.379
32	YL GUC2b (Bcl2)	Tm max	240.264
33	YL GUC3b (Bcl2)	Accessibility min	5.2
34	YL GUC4b (Bcl2)	Accessibility max	22.3
35	YL GUC5b (Bcl2)	Structure min	23.711
36	YL GUC6b (Bcl2)	Structure max	56.681

37	Pistol1b (Bcl2)	Tm min	33.491
38	Pistol2b (Bcl2)	Tm max	155.947
39	Pistol3b (Bcl2)	Accessibility min	2.2
40	Pistol4b (Bcl2)	Accessibility max	20.6
41	Pistol5b (Bcl2)	Structure min	30.511
42	Pistol6b (Bcl2)	Structure max	65.625
43	Extend1b (Bcl2)	Tm min	97.492
44	Extend2b (Bcl2)	Tm max	237.865
45	Extend3b (Bcl2)	Accessibility min	2.8
46	Extend4b (Bcl2)	Accessibility max	23.6
47	Extend5b (Bcl2)	Structure min	9.931
48	Extend6b (Bcl2)	Structure max	45.15
49	FMN act1 (Bcl2)	Tm min	85.063
50	FMN act2 (Bcl2)	Tm max	204.685
51	FMN act3 (Bcl2)	Accessibility min	3.3
52	FMN act4 (Bcl2)	Accessibility max	14.6
53	FMN act5 (Bcl2)	Structure min	31.579
54	FMN act6 (Bcl2)	Structure max	55.275
55	FMN inh1 (Bcl2)	Tm min	85.063
56	FMN inh2 (Bcl2)	Tm max	204.685
57	FMN inh3 (Bcl2)	Accessibility min	3.3
58	FMN inh4 (Bcl2)	Accessibility max	14.6
59	FMN inh5 (Bcl2)	Structure min	26.714
60	FMN inh6 (Bcl2)	Structure max	51.762

Ribozymes are listed by pairs of high/low values for the indicated parameter computed by Ribosoft 2.0.

Ribozymes testing in cells: some highly active ribozymes were chosen and cloned into plasmids to be tested in cells. Protocols and results are presented in 'Appendix I'.

Récapitulatif

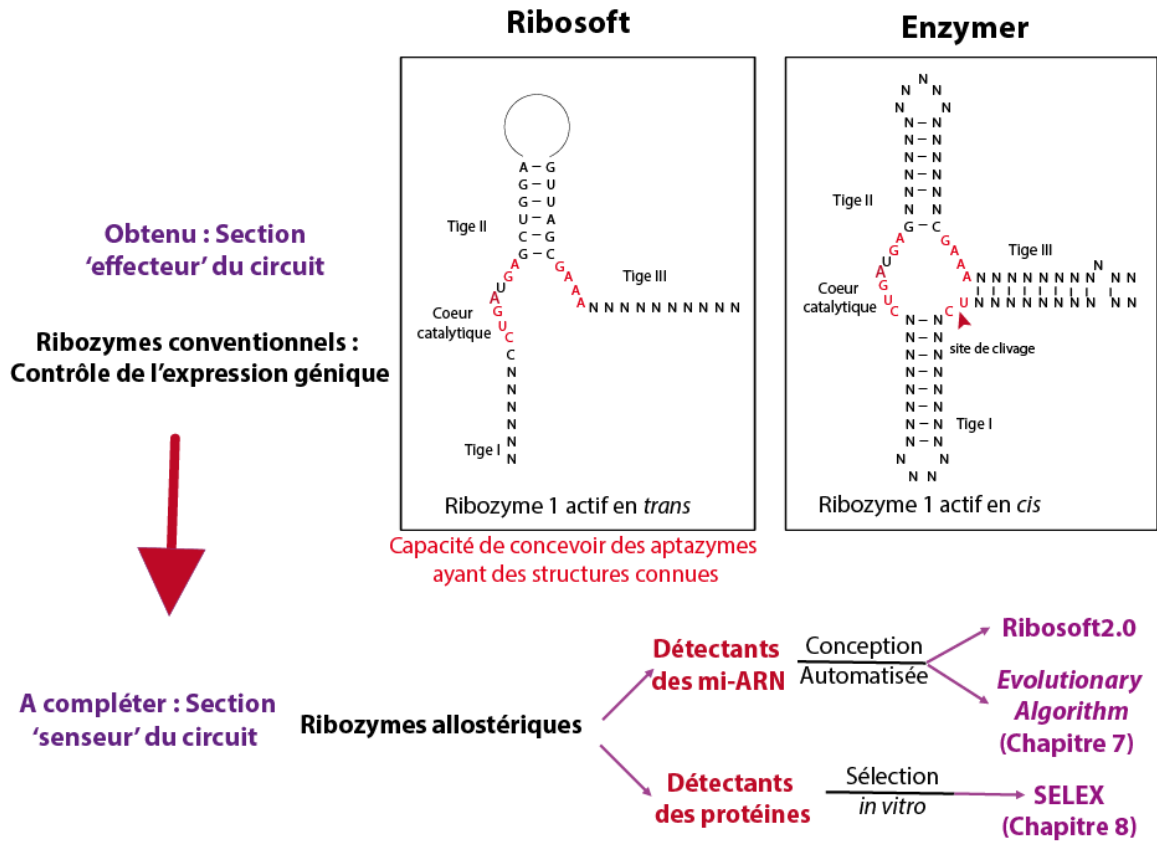


Figure 6.11 État du circuit logique

Les logiciels de conception ont permis jusqu'ici de produire des ribozymes ayant la capacité d'agir comme outil de contrôle de l'expression génique. Pour compléter le circuit, les ribozymes senseurs, seront soit conçus par outil bioinformatique ou sélectionnés *in vitro*.

7. Portes logiques basés sur des ribozymes allostériques contrôlés par des miARN

7.1. Mise en contexte

Dans les chapitres précédents, nous avons décrit des outils de conception de ribozymes, surtout conventionnels. Nous avons également, précisé que les exemples de ribozymes allostériques mentionnés étaient contrôlés par des métabolites. Dans ce chapitre, nous décrirons des ribozymes allostériques induits par la liaison d'oligonucléotides comme entrées. Dans notre projet, ce sont des miARN qui jouent le rôle d'entrée pour induire ou inhiber la réaction du clivage du ribozyme. Ces ribozymes ont été conçus pour agir comme porte logique répondant à un ou plusieurs miARN.

Dans ce chapitre, nous présenterons deux hypothèses : la première porte sur l'utilisation d'un seul ribozyme pour lier deux ou trois miARN afin de créer une porte logique répondant à une fonction logique. Cette liaison inhibe ou induit l'activité du ribozyme, ce qui peut être utilisé pour générer un signal indiquant alors la présence de ces trois miARN. La deuxième concerne la liaison d'un seul miARN par un ribozyme, entraînant la dégradation de ce dernier.

7.2. Circuit logique

Le développement des circuits logiques utilisés dans le cas des ribozymes et d'autres ARNnc est le résultat des avancements en génie génétique. Ces progrès ont permis d'utiliser différentes composantes pour des systèmes de plus en plus complexes, permettant de combiner des opérations logiques pour constituer des circuits (Singh, 2014; Xiang *et al.*, 2018).

7.2.1. Les fonctions logiques

Les fonctions booléennes mathématiques sont utilisées pour développer des opérations logiques représentées sous la forme B^n où « n » est le nombre de données d'entrées qui peuvent être soit « Vraie » (égale à 1) ou « Fausse » (égale à 0) ; leur combinaison donnera une valeur « Vraie » ou « Fausse » (Benenson, 2009). Plusieurs opérateurs ont été mis en place à partir de ces deux valeurs, essentiellement, ET, OU, NON, Non-ET, NON-OU et OU exclusif (Tinder, 2000).

- **Conjonction-ET (AND)**

Pour deux variables « a » et « b », la valeur « a et b » n'est VRAIE que lorsque « a » est vrai et « b » est vrai.

Table 7.1 Table de vérité de la fonction ET

A	B	a et b
0	0	0
1	0	0
0	1	0
1	1	1

- **Disjonction-OU (OR)**

C'est lorsque a OU b est vrai, de même que lorsque a ET b sont tous deux vrais.

Table 7.2 Table de vérité de la fonction OU

A	B	a ou b
0	0	0
1	0	1
0	1	1
1	1	1

- **Négation-NON (NOT)**

Dans ce cas, Y n'est vrai que lorsque a est faux.

Table 7.3 Table de vérité de la fonction NON

A	Non a
0	1
1	0

7.2.2. Les portes logiques (*logic gates*)

Les fonctions logiques sont utilisées sous forme de portes logiques dans les circuits numériques en utilisant deux ou plusieurs données d'entrées pour aboutir à un seul résultat comme sortie (Voigt, 2006).

7.2.3. Circuit logique en biologie

Le principe des circuits logiques a été exploré par les biologistes dans un domaine nommé « la biologie synthétique » pour résoudre des problèmes liés au vivant (santé, environnement, etc.). Ceci est exécuté en contrôlant l'expression génique et en essayant d'attribuer de nouvelles fonctions aux cellules vivantes (Endy, 2005). La découverte de l'opéron Lac a montré que les cellules utilisent déjà la logique mathématique pour réguler l'expression génique (Jacob & Monod, 1961). Le principe est alors passé de la régulation simple de l'expression génique au développement de systèmes complexes. Ces systèmes sont composés de plusieurs parties biologiques fonctionnelles pour obtenir de nouvelles fonctions chez un organisme donné (Andrianantoandro *et al.*, 2006; Benenson, 2012; Goni-Moreno & Nikel, 2019).

En général, les parties biologiques (*building blocks*) qui sont utilisées pour développer les circuits logiques sont : des promoteurs, des protéines capables de lier l'ADN (répresseurs, facteurs de transcription, etc.), des recombinaisons, le système CRISPR et des ARN régulateurs (Brophy & Voigt, 2014; Nielsen *et al.*, 2016; Sudarsan *et al.*, 2006). Ces parties peuvent jouer l'un des quatre rôles trouvés généralement dans un dispositif génétique :

- 1- Un senseur : partie qui détecte un signal dans l'environnement.
- 2- Un régulateur : partie qui effectue l'action de régulation génique.
- 3- Un actuateur ou actionneur : partie qui émet un signal détectable considéré comme « *output* » du circuit.
- 4- Un adaptateur : partie qui lie les composantes du circuit ensemble (Wang *et al.*, 2013).

Un exemple de nouvelles fonctions générées par la biologie synthétique est celui des aptazymes, qui ont été conçus par Win et Smolke en 2007. Ils peuvent répondre à des signaux intra- ou extracellulaires pour effectuer les opérations logiques : ET, OU, NON-OU et NON-ET (Win & Smolke, 2007a) (Shapiro & Gil, 2008). Les *antiswitchs* sont formés par l'assemblage d'un aptamère et d'un domaine anti-sens. Ce dernier fonctionne aux

dépens de la présence ou de l'absence du ligand de l'aptamère, pour créer un *switch* ON ou OFF (Bayer & Smolke, 2005; Isaacs *et al.*, 2006). Des exemples plus récents sont les *Toehold switches* (Green *et al.*, 2014) et les petits ARN activant la transcription (*STAR : small transcription activating RNAs*) (Chappell *et al.*, 2015) qui agissent comme des riborégulateurs, liant des ARN plutôt que des métabolites ou des protéines (Kim & Franco, 2020). Les circuits synthétiques ont aussi été utilisés dans les cellules eucaryotes à plusieurs fins. Citons comme exemples : le contrôle de l'expression génique dans les cellules cancéreuses par un ordinateur biologique (Benenson *et al.*, 2004) et une thérapie génique basée sur une porte logique ET combinant deux CARs (*Chimeric Antigen Receptor*) (Kloss *et al.*, 2013). Diverses autres applications existent, comme le développement des vaccins, l'étude de signalisation cellulaire et la régulation génique (Lienert *et al.*, 2014) (Kim & Franco, 2020).

Récemment, plusieurs études ont utilisé des combinaisons de miARN comme entrée pour des portes logiques (Wu *et al.*, 2019; Xie *et al.*, 2011; Yang *et al.*, 2021; Yue *et al.*, 2020; Yue *et al.*, 2019). Ceci s'explique par leur rôle comme marqueur dans la reconnaissance de maladies selon les profils d'expressions.

7.2.4. Les micros ARN

Les miARN sont des ARNnc simple brin d'une taille d'environ 22 nucléotides. Ils existent chez la plupart des eucaryotes et ils sont inclus dans la régulation de plus de 30 % des gènes codants des protéines (Macfarlane & Murphy, 2010). Ces ARN sont transcrits sous forme de pri-miARN qui passent par deux étapes de clivage pour générer les miARN matures qui joignent le complexe RISC (*RNA-induced silencing complex*). Un miARN guide, via la séquence *seed*, le complexe RISC en liant, par complémentarité de base, une séquence d'ARNm cible. Ceci engendre la dégradation de la cible par le complexe lorsque le nombre de paires de bases qui se forment entre le miARN et l'ARNm est grand. Cette liaison peut aussi inhiber la traduction par la déstabilisation de la liaison du ribosome à l'ARNm, lorsque la complémentarité entre les deux ARN n'est pas parfaite (**Figure 7.1**). Il s'agit d'une vue générale du mode d'action des miARN, car leurs mécanismes d'action sont multiples et sont encore étudiés afin que l'on en ait une meilleure compréhension (Bartel, 2004; Oliveto *et al.*, 2017; Pu *et al.*, 2019).

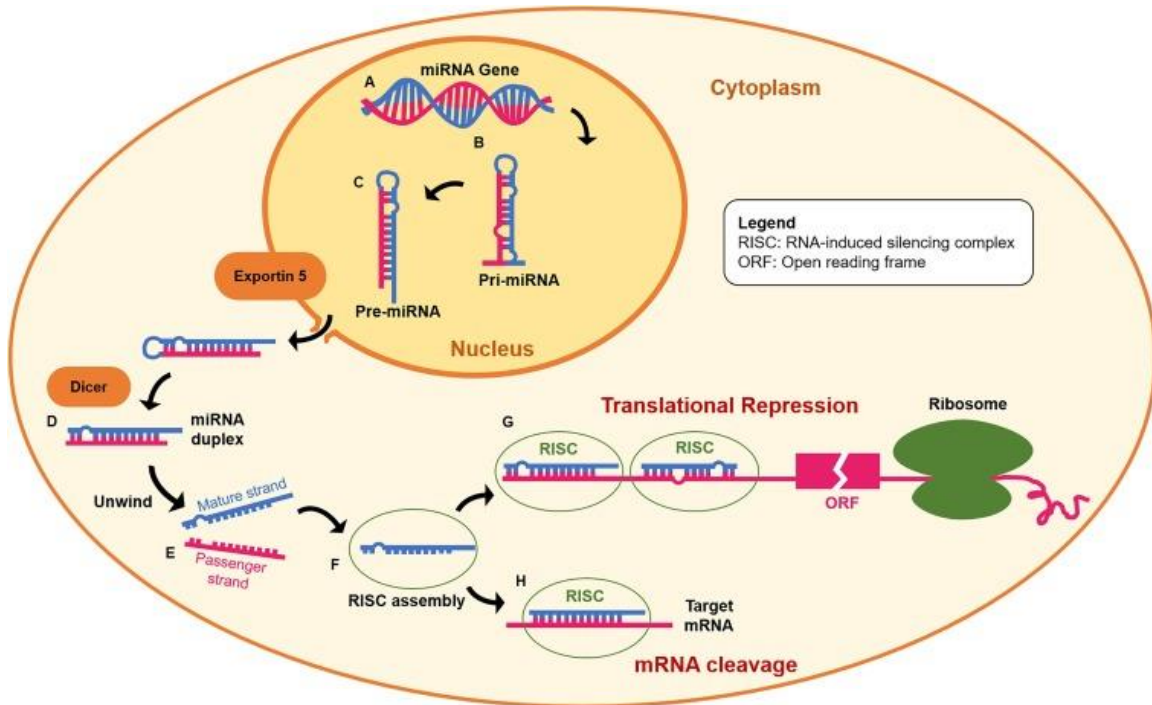


Figure 7.1 Activité du micro-ARN.

Les pre-miARN transportés au cytoplasme par l'exportin5 sont clivés par Dicer et le brin mature rejoint le complexe RISC. Le miARN dirige le complexe RISC vers l'ARNm cible afin d'empêcher sa traduction : lorsque la complémentarité entre le miARN (en bleu) et la région cible de l'ARNm (en rose) est parfaite, le complexe RISC cause la dégradation de l'ARNm. Quand la complémentarité entre les deux séquences est imparfaite, la liaison du complexe RISC à l'ARNm empêche le passage du ribosome en altérant alors l'étape de la traduction. Tiré de (Lee *et al.*, 2019)

Étant donné que les miARN se trouvent impliqués dans le contrôle de l'expression des gènes responsables de plusieurs mécanismes biologiques (par exemple, la croissance cellulaire), la dérégulation de leur expression est impliquée dans plusieurs cancers (Chen *et al.*, 2021a; Ferragut Cardoso *et al.*, 2021; Ng *et al.*, 2021). En fait, la surexpression ou la sous-expression de certains miARN affecte la croissance des cellules cancéreuses positivement ou négativement compte tenu des fonctions de leurs cibles (Macfarlane & Murphy, 2010) (Golhani *et al.*, 2021; Peng & Croce, 2016).

7.2.5. Les profils d'expression de miARN

Le taux d'expression des miARN est variable d'un type cellulaire à un autre ; c'est aussi le cas de leur expression d'un type ou sous-type de cancer à un autre. Cette différence veille à ce que les profils d'expression des miARN, incluant leurs sous- et surexpressions, puissent être utilisés pour distinguer différents sous-types de cancers de cellules saines (Anelli *et al.*, 2021; Sokilde *et al.*, 2019; Wang *et al.*, 2017; Wang *et al.*, 2019; Yerukala

Sathipati & Ho, 2018; Zhu *et al.*, 2014). Cette expression altérée permet aux miARN d'agir comme des oncogènes ou suppresseurs de tumeurs, ce qui les définit comme biomarqueurs potentiels de diagnostic et cibles intéressantes de thérapie (Berindan-Neagoe *et al.*, 2014). En se basant sur cette hypothèse, plusieurs recherches ont été produites pour démontrer que les miARN extracellulaires, circulant dans le sang, exosomaux, etc. sont utilisables pour la détection de maladies. En ce qui a trait au cancer, ils sont utilisables pour la détermination non seulement du type du cancer, mais aussi de son état de développement, de sa progression et de ses métastases, puis de la réponse des tumeurs aux thérapies (Alves Dos Santos *et al.*, 2020; Chen *et al.*, 2021c; Duan *et al.*, 2021; Gherman *et al.*, 2020; Kachris *et al.*, 2021; Liu & Chu, 2020; Monsellato *et al.*, 2020; Yang *et al.*, 2019).

En nous basant sur ce principe, nous avons décidé d'utiliser des portes logiques pouvant lier entre un à trois miARN en parallèle. Cette liaison n'est pas envisageable par un simple ajout de séquence au ribozyme, mais plutôt par un choix dès l'étape de conception de la fonction logique, qui sera effectuée par le ribozyme en liant les miARN.

7.3. Des ribozymes liant jusqu'à trois miARN

7.3.1. Portes logiques ET et NON-ET

Initialement une librairie de 1000 ribozymes allostériques répondant à un, deux ou trois différent miARN a été conçue avec un outil informatique en développement par un collaborateur à l'université Concordia, Nicolas Kamel, du laboratoire du Pr. Nawwaf Kharma. L'algorithme qui a été utilisé pour la conception de ces ribozymes utilise une approche évolutionnaire. Cette approche permet de générer une séquence *seed* qui passe par une évolution sous forme d'itérations de mutations. Chaque itération résulte en une 'génération' de séquences qui passent par une étape d'évaluation du repliement de la structure formée après les mutations. Ce repliement est comparé en présence des inputs et en leur absence. Ces itérations et évaluations permettent de chercher le meilleur repliement possible, et les meilleurs candidats sont choisis, et leurs séquences sont optimisées jusqu'à l'obtention du repliement désiré en fonction de la porte logique visée. Les miARN agissant comme entrées pour les ribozymes allostériques, lors de leurs liaisons à des séquences complémentaires sur les ribozymes, stabiliseront les structures actives de ces derniers. En fait, les ribozymes en absence des miARN forment une structure altérée qui bloque le cœur catalytique, ce qui empêche le clivage. La liaison des

miARN permet alors le relâchement de la séquence bloquée permettant ainsi le bon repliement du ribozyme ce qui engendre l'activation de sa réaction catalytique (**Figure 7.2**). Des essais ont été effectués (présenté en **Annexe II**) avec les différentes combinaisons possibles (**Table S4**) mais le séquençage a échoué, les résultats obtenus étaient principalement des artéfacts issus de polymères d'amorces d'amplification. Plusieurs problèmes techniques (tel que des ligations différentes pour les ARN intacts et les ARN coupés) devront être surmontés pour réussir ces essais. Ceci ne nous a pas fournis des résultats clairs pour conclure si notre hypothèse est validée ou non. Pour cette raison, il fallait tester au moins quelques séquences pour avoir une idée préliminaire sur l'activité de ce genre de portes logiques. Nous avons alors pris l'opportunité pour tester des *design* optimisés, après quelques améliorations du programme. Ces nouvelles portes logiques sont supposées lier un, deux ou trois miARN et répondre aux opérations logiques ET et NON-ET. Sept ribozymes ont été conçus pour être seulement actifs en présence de trois miARN (miARN18, 135 et 155 (**Table S3**)), c'est-à-dire une 'opération ET'. Les 14 autres séquences répondent chacune à deux miARN et ils sont actifs en présence de deux miARN seulement dans le cas de l'opération 'ET' ou d'un seul seulement dans le cas de l'opération 'NON-ET'. (**Table S5 en annexe II**). Les miARN choisis sont surexprimés dans le cancer du sein et ils ont été choisis à titre d'exemple dans le but de les utiliser seulement dans la preuve de concept.

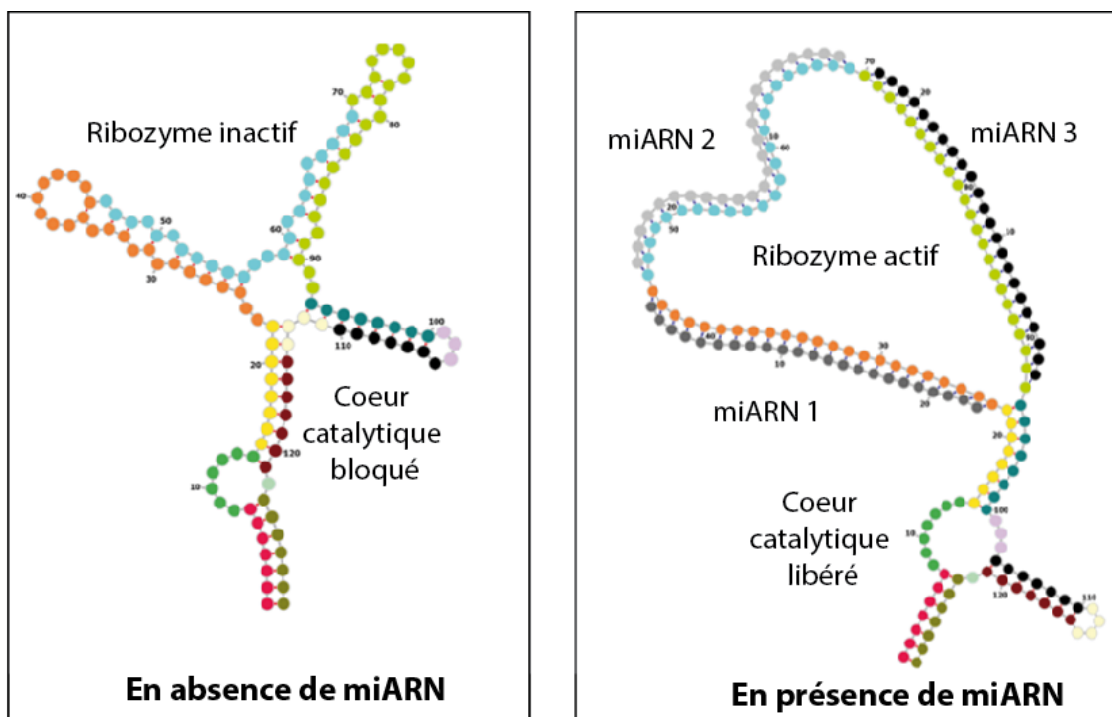


Figure 7.2 Structure d'un ribozyme 'porte logique ET' liant trois miARN

Gauche : En absence des miARN le cœur catalytique du ribozyme est bloqué par le mauvais repliement de la structure du ribozyme. Droite : la liaison des miARN au ribozyme cause un déplacement des brins qui bloquent le cœur catalytique en libérant ce dernier ce qui active le ribozyme effectuant la réaction d'auto-clivage.

7.3.1.1. Matériels et méthodes

- Préparation de l'ADN et ARN (marquage radioactif)

Les matrices d'ADN utilisées pour amplifier les ribozymes ont été commandées de chez Twist. Plusieurs matrices contenaient deux séquences de ribozymes, et des séquences adaptatrices ont été utilisées comme amorces sens et pour les amorces inverses, une séquence spécifique à chaque ribozyme a été utilisée. L'amplification de la matrice de chaque amorce de ribozyme a été faite par PCR avec les amorces correspondantes (Table S6 en annexe).

Les tailles de produits de PCR ont été vérifiées par migration sur gel d'agarose 2%, ensuite l'ADN a été précipité à l'éthanol afin de le préparer à l'étape de transcription. Les ribozymes ont été produits par transcription *in vitro* en présence de l'UTP radioactif. Les ARN produits ont été précipités à l'éthanol et purifiés sur gel de polyacrylamide dénaturant de 6%.

- **Test de clivage avec les différentes combinaisons**

A la suite de la précipitation des ARN à l'éthanol, les tests de clivage ont été effectués. Pour les ribozymes ayant trois entrées, huit combinaisons différentes ont été essayées pour chaque ribozyme. Pour ceux ayant deux entrées, quatre combinaisons ont été testées pour chaque. Les réactions ont été effectuées dans le tampon de clivage et en présence de miARN correspondants, pendant une heure, à 37°C. Les résultats ont été visualisés sur un gel de polyacrylamide dénaturant de 6%.

7.3.1.2. Résultats

Au total 21 ribozymes contrôlés par deux ou trois miARN ont été testés. Ceux activés par seulement un miARN, mais pas leur combinaison (NON-ET), n'avaient pas montré les résultats attendus. Les ribozymes étaient soit activés ou inactivés indépendamment des miARN incubés avec eux alors les résultats ne sont pas montrés ici.

La nomenclature choisie pour les portes logiques «ET» est : X AND #X' ou le X est le nombre de miARN à lier par le ribozyme et le X' est le numéro du ribozyme allant de 1 à 7. Parmi les sept ribozymes conçus pour être activés par deux miARN, les deux ribozymes 2 AND #3 et 2 AND# 6 avaient montré une activité de clivage en absence de miARN ou en présence d'un seul miARN, mais la combinaison des miARN 18 et 155 semble inhiber cette activité. Ceci est l'opposé de ce qui était attendu de ces ribozymes (**Figure 7.3 A**).

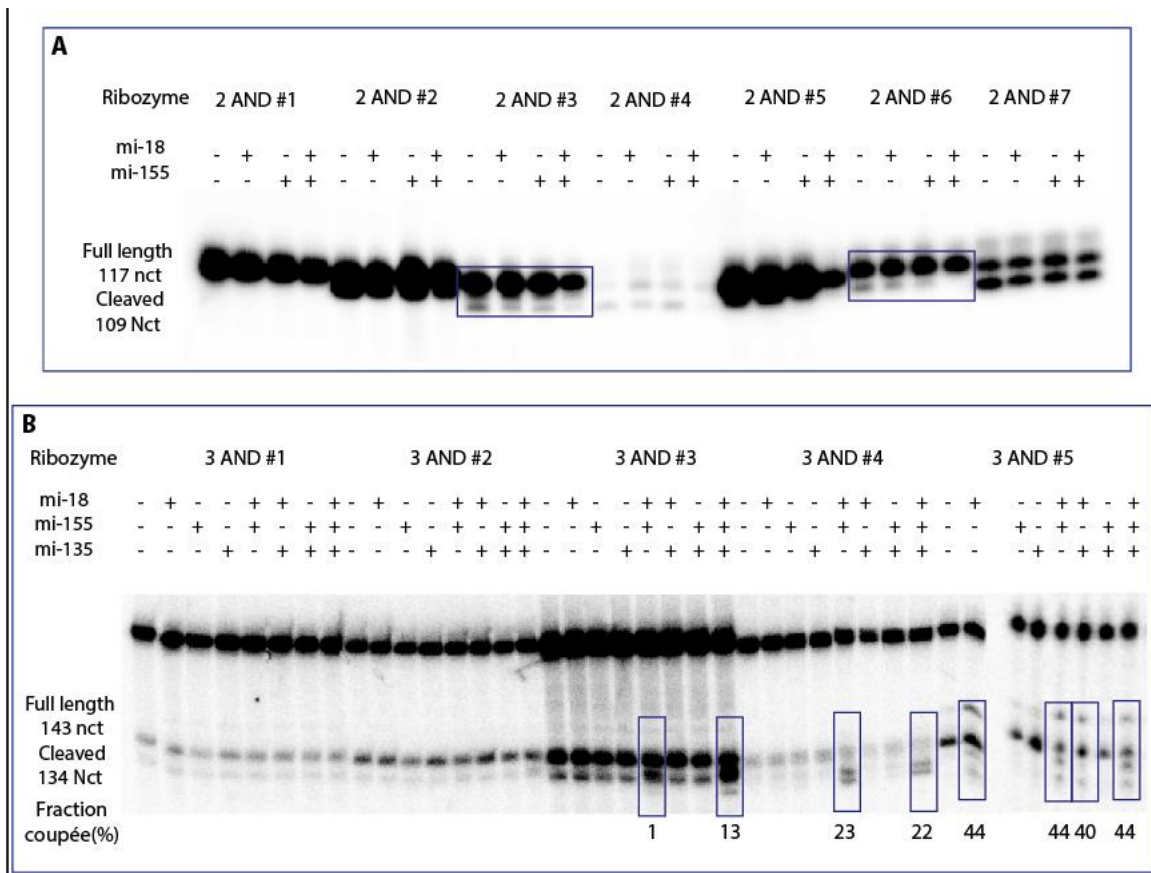


Figure 7.3 Les portes logiques 'ET' à deux et trois entrées

(A) Les ribozymes 2 AND (ET) #3 et 2 AND (ET) #6 montrent une faible activité de clivage en présence d'un seul miARN ou en absence d'entrées. (B) Les taux de clivage (en bas du gel en (%)) ont été calculés seulement pour les ribozymes présentant une activité en réponse à certains miARN. Les différentes bandes qui apparaissent sont probablement dues à des régions riches en U de ces ribozymes ce qui a causé problème lors des transcriptions radioactives avec de l'UTP radioactif (où l'on met moins d'UTP dans la réaction de transcription).

Parmi les sept ribozymes liant trois miARN pour être activés, deux ribozymes avaient montré un résultat intéressant. Le ribozyme 3 AND #4 a été activé quand les miARN 18 et 155 étaient combinés, une bande était aussi visible pour la combinaison miARN 18+miARN 155+miARN 135, mais pas miARN18+miARN 135.

Le ribozyme 3 AND #5 a été activé à chaque fois qu'il était en présence du miARN 18 et l'activité de clivage était plus élevée lorsque combiné avec le miARN 155.

Dans le cas du ribozyme 3 AND#3, une faible activité de clivage est observée en présence des miARN 18 et 155 et lorsque le miARN 135 est rajouté à la combinaison, l'activité est devenue plus importante. **(Figure 7.3 B)**

Bien que les réponses de ces ribozymes ne correspondent pas parfaitement aux fonctions logiques pour lesquelles ils ont été conçus, l'activité obtenue suggère qu'une optimisation du modèle pourrait mener à des séquences ayant la fonction attendue.

7.4. Des ribozymes ayant une séquence complémentaire au miARN en 3'

Lier des miARN pour activer les ribozymes semble avoir un potentiel pour contrôler leur activité. Par contre une complémentarité parfaite entre une région du ribozyme avec un miARN pourrait engendrer la dégradation du ribozyme si le miARN est lié au complexe RISC. Alors pour tester cette hypothèse et dans le but d'avoir des ribozymes inactivés par la présence de miARN, nous avons conçus des ribozymes actifs en *trans* capables de lier un miARN (matrice présenté en **annexe III**). Pour visualiser l'activité de ribozymes dans les cellules nous avons choisi d'utiliser la luminescence en choisissant comme cible l'ARNm de la luciférase. Les ribozymes ont été conçus par Ribosoft 2.0 pour dégrader l'ARNm de la luciférase et lier un miARN par complémentarité de séquence à une région ajoutée en 3'. Un exemple de matrice pour le ribozyme liant le miARN et la séquence cible est présenté en annexe et une présentation schématique est dans la **figure 7.4**.

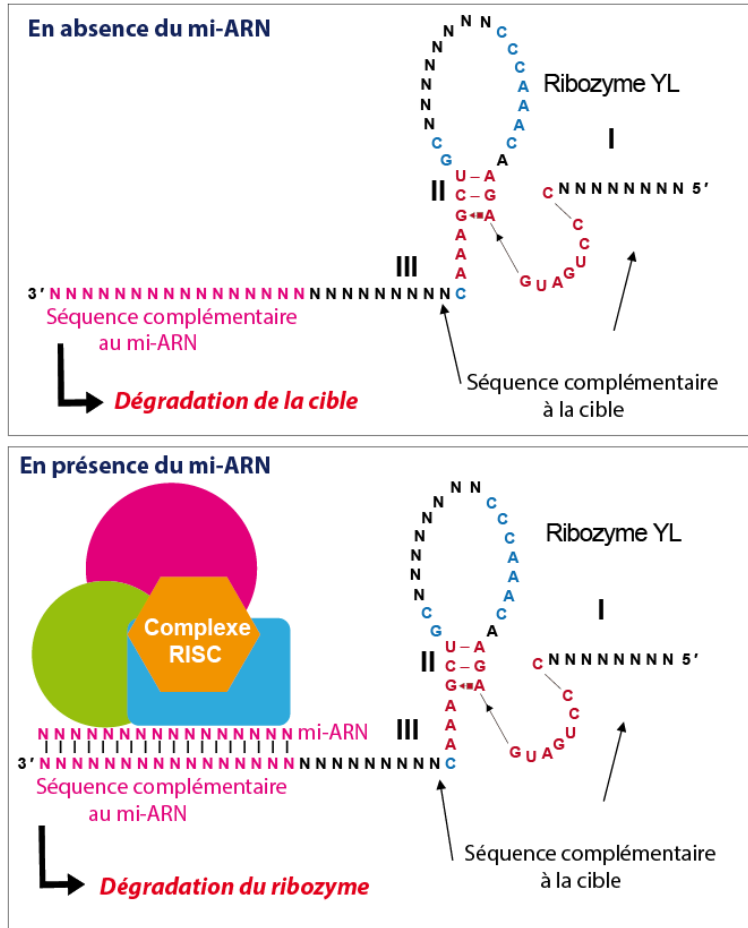


Figure 7.4 Mode d'action des ribozymes et des miARN d'un ribozyme contrôlé par un miARN

Le ribozyme est actif et capable de dégrader sa cible en absence du miARN (haut). Le miARN accompagné par le complexe RISC lorsque présent, se lie par complémentarité parfaite au ribozyme causant alors sa dégradation (bas).

L'hypothèse est basée sur le fait que lorsqu'un miARN lie une séquence par complémentarité de séquence parfaite, le complexe RISC dégrade l'ARN ciblé. Dans ce cas le ribozyme liant un miARN avec une complémentarité parfaite sera dégradé et en absence de miARN, il sera intact et actif. Les séquences choisies sont présentées dans la table S7.

7.4.1. Matériels et méthodes

- Test de clivage de l'ARNm de la luciférase *in vitro* : Design des séquences

Chaque séquence de ribozyme a été conçue pour avoir une séquence complémentaire à un miARN en 3'. De plus la séquence du promoteur de l'ARNt, utilisé pour l'expression et le transport du ribozyme dans les cellules, a été ajouté en 5' de la séquence du ribozyme.

Ceci a été fait dans le but de vérifier l'activité des ribozymes même en présence de ces séquences afin de confirmer que les séquences en surplus n'interfèrent pas avec la structure active du ribozyme. Ceci permettra de prouver que l'absence ou l'inhibition d'activité des ribozymes *in vivo* est due à l'effet du complexe RISC et non pas aux autres facteurs.

- **Préparation des ribozymes**

Les séquences de ribozymes ont été amplifiées par PCR d'assemblage en utilisant la Taq polymérase (NEB). En fait, les séquences des ribozymes étaient assez longues, donc pour chaque ribozyme quatre oligos ayant des régions chevauchantes ont été utilisés pour générer la matrice finale utilisée pour synthétiser l'ARN (**Table S8**). En total, neuf séquences, liant chacun des trois miARN mir-18a, mir-135 et mir-155, ont été choisies. Les produits de PCR ont été précipités à l'éthanol et les ARN ont été synthétisés par transcription *in vitro* à l'aide de la T7 polymérase. Les ARN ont été purifiés par extraction phénol/chloroforme suivie d'une migration sur gel de polyacrylamide dénaturant.

- **Préparation de la cible et marquage**

La cible, qui est l'ARNm de la luciférase a été amplifiée à partir du plasmide '*pCMV-Green Renilla Luc Vector for Luciferase Assays*' de chez Thermofisher en utilisant les oligos : sens **TAATACGACTCACTATAGGGATCGGCATGGGCAAGAG** et inverse : **GATCAGGGCCAGTTCTTCC**. Le produit de PCR a été précipité à l'éthanol, et l'ARN radioactif a été produit par transcription *in vitro* en présence de l'UTP³² et purifié sur gel de polyacrylamide dénaturant 6%.

- **Test de clivage**

Pour le test de clivage, chaque ribozyme a été incubé avec la cible dans le tampon de clivage (10 mM MgCl₂, 100 mM NaCl, 50 mM tris HCl pH 7.4 et 25 mM KCl) en présence de 4 µl de la cible radiomarquée pendant une heure à 37°C. La réaction a été arrêtée par l'ajout du tampon de formamide dénaturant (10 mM EDTA, 95 % formamide, 0.025% bleu de bromophénol et 0.025% bleu de xylène cyanol).

7.4.2. Résultats

Les ribozymes choisis comprenaient trois ribozymes liant chacun un des miARN 18, 135 ou 155. Le test de clivage de l'ARNm de la luciférase était fait en absence de miARN pour valider leur activité dans un milieu ne contenant pas les miARN. Ceci a été fait aussi pour vérifier que le surplus de séquences en 5' et 3' ne causent pas la formation d'une structure altérée empêchant le cœur catalytique de se former correctement.

Après une heure d'incubation huit parmi neuf ribozymes étaient actifs avec un taux de clivage élevé (entre 78% et 96%) (**Figure 7.5**). Ceci nous a permis de conclure que les structures formées par les ribozymes se replient correctement. Puisque les ribozymes étaient assez efficaces, nous avons aussi choisi de tester une séquence liant le miARN135 dans les cellules en présence du miARN correspondant et un autre non spécifique.

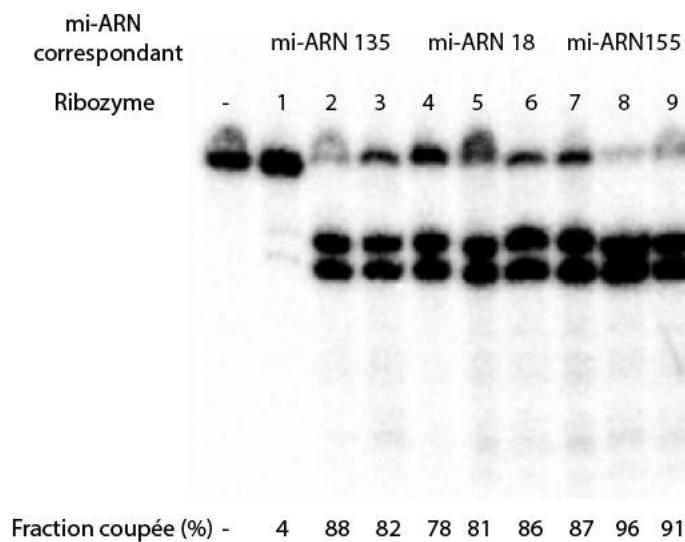


Figure 7.5 Test de clivage des ribozymes contre l'ARNm de la luciférase

Les ribozymes ont été testés contre l'ARNm de la luciférase radiomarqué. Le taux de clivage est présenté en (%) en bas du gel.

Finalement trois ribozymes liant chaque un miARN (135, 155 ou 18) ont été cloné dans des plasmides afin d'évaluer notre hypothèse dans des cellules exprimant ces miARN. Les résultats préliminaires obtenus étaient négatifs (présenté dans l'**annexe III**).

8. Article 4: Dual selection for aptazymes for biosensor development purpose

8.1. Mise en contexte

Dans les chapitres antérieurs nous avons décrit des méthodes rationnelles et automatisées de générations de nouveaux ribozymes conventionnels ou allostériques. Ceci a été jumelé à des essais expérimentaux pour montrer qu'une majorité de ces ribozymes s'est avérée active au moins *in vitro*. Bien que les résultats des tests de clivage de ces ribozymes ne fussent pas une réussite dans les cellules, nous avons plusieurs autres candidats qui peuvent être testés afin de trouver le ribozyme susceptible de jouer le rôle de 'l'effecteur' de notre circuit. De même, des ribozymes 'senseurs' détectant des miARN ont été obtenus de la même façon. Le manuscrit suivant décrira une méthode uniquement expérimentale dans le but de l'obtention de nouveaux ribozymes allostériques. En fait, pour le but de sélection de nouveaux ARN avec une capacité de lier une cible comme un métabolite ou une protéine aucun outil informatique existant actuellement n'est capable d'atteindre ce but. Ceci fait que le travail expérimental à ce stade est irremplaçable.

A ce stade du projet nous pouvons passer à la dernière partie de notre projet, déterminé comme : un senseur qui détectera les protéines surexprimées dans les cellules cancéreuses. Dans le cadre de notre projet, le senseur est un ribozyme allostérique et plus précisément un aptazyme. Cet aptazyme est une combinaison d'un aptamère et d'un ribozyme qui est capable, simultanément, de détecter une molécule d'intérêt et de réagir en réponse à cette détection par une activation ou une inhibition d'une réaction catalytique. Dans ce chapitre, nous proposons une nouvelle méthode de sélection de ribozymes allostériques inactivés ou activés en liant la protéine cible. Nous montrons aussi que dans un contexte de diagnostic les aptazymes obtenus peuvent être utilisés comme des biosenseurs des biomarqueurs de certains cancers.

Dual selection for aptazymes for biosensor development purpose

Double sélection des aptazymes à des fins de développement de biosenseurs

Sabrine Najeh¹, Jonathan Perreault¹, Yves St-Pierre¹

¹Centre Armand-Frappier Santé Biotechnologie, Institut national de la recherche scientifique (INRS), 531 boulevard des Prairies, Laval, QC, H7V 1B7, Canada

État du manuscrit : Les résultats du manuscrit seront utilisés pour déposer un brevet, alors ce manuscrit sera soumis suite au dépôt du brevet.

8.1. Contributions des auteurs

Sabrine Najeh: A l'aide de Jonathan Perreault, développement de la nouvelle méthode de sélection. La planification du travail expérimental présenté et sa réalisation, le développement du biosenseur, l'analyse et interprétation des résultats ainsi que la rédaction de l'article ont été mon travail.

Jonathan Perreault : Supervision et assistance de Sabrine Najeh et révision de l'article.

Yves St-Pierre : Supervision et assistance de Sabrine Najeh et révision de l'article.

8.2. Abstract

Allosteric ribozymes are adaptable to any desired application by the fact that they can be selected for any type of molecule. They can be applied as gene regulation tools, logic gates as well as biosensors. Multiple methods were developed for their selection ranging from *in vitro* and *vivo* isolation to rational and computational design. The selection can aim to obtain a ribozyme induced by the ligand binding or the opposite. In this paper, we describe a new selection method that take in account both the affinity of the ligand and the cleavage activity of the ribozyme. We also present the results of using a mutated catalytic core of a hammerhead ribozyme, for the selection of inactive aptazymes that can be applied as biosensors.

Résumé

Les ribozymes allostériques sont adaptables à toute application souhaitée par le fait qu'ils peuvent être sélectionnés pour tout type de molécule. Ils peuvent être appliqués comme outils de régulation de l'expression génique, des portes logiques ainsi que des biosenseurs. Plusieurs méthodes ont été développées pour leur sélection, allant de la sélection *in vitro* et *in vivo* à la conception rationnelle et computationnelle. La sélection peut viser à obtenir un ribozyme induit par la liaison du ligand ou l'inverse. Dans cet article, nous décrivons une nouvelle méthode de sélection qui prend en compte à la fois l'affinité du ligand ainsi que l'activité de clivage du ribozyme. Nous présentons également les résultats de l'utilisation d'un cœur catalytique muté d'un ribozyme en tête de marteau, pour la sélection d'aptazymes inactifs pouvant être appliqués comme biosenseurs.

8.3. Introduction

Small self-cleaving ribozymes represent a powerful tool in synthetic biology due to the versatility of their use and the ease of their modifications. They have multiple applications, from controlling gene expression to their use in complicated logic circuits (Park *et al.*, 2019b). The control of their activity becomes an advantage, to ensure a safety lock when needed and permit sensing. For this purpose, allosteric ribozymes were developed (Araki *et al.*, 1998). The allosteric regulation of a molecule requires the binding of an 'effector' molecule in a site other than the active site. This binding generates a conformational change in the structure resulting in the activation or the inhibition of the molecule's function (Wallden *et al.*, 2007). This mechanism is used by many enzymes involved in different metabolic and cell signaling pathways, representing one of the firm and solid control mechanisms that cells have to ensure the well-being and the homeostasis needed for survival (Bu & Callaway, 2011).

Until the end of the 90s, all catalytic RNAs discovered were not subject to any allosteric regulation. In 1997, the Breaker lab came up with the idea of applying the allosteric regulation principle to catalytic RNAs to have a control on their activity (Tang & Breaker, 1997). The development of SELEX (systematic evolution of ligand by exponential enrichment) (Tuerk & Gold, 1990a) had a major role in this progress since it allowed the selection of aptamers to bind any desired molecule or cell. The Breaker lab used rational design to implement already selected aptamers into a minimal hammerhead ribozyme (HHRZz) stem. The idea came from the fact that aptamers can potentially have two states of structure depending on the binding of the ligand. By inserting an aptamer into a ribozyme's sequence, the structure could be manipulated, thereby affecting its activity. This combination was named aptazyme by Robertson and Ellington in 1999 (Robertson & Ellington, 1999).

While the construction of aptazymes was optimized, many findings helped understand the importance of allosteric regulation. One of these discoveries was 'riboswitches', formed by two elements, a sensor part binding specifically a metabolite and the expression platform acting in accordance to the change in the aptamer structure (Garst *et al.*, 2011). In prokaryotic cells, the mechanism of riboswitches is simpler than the riboswitch-mediated modulation of splicing in eukaryotic cells (Garst *et al.*, 2011; Li & Breaker, 2013).

In either case, trying to convert synthetic aptamers into riboswitches to control the expression of a new target was not a simple task. The discovery of two natural allosteric ribozymes showed that inducible ribozymes could be a simpler way to reach the goal of universally active synthetic riboswitches. The first ribozyme was 'GlmS' and is induced by glucosamine-6-phosphate for cleavage (Savinov & Block, 2018). Even though the binding of Glc-6-P does not induce a structural change in the ribozyme structure, the GlmS ribozyme acts as if it was allosteric because its dependence on a cofactor. The other ribozyme is a member of the group I intron and its activity is controlled by c-di-GMP (Chen et al., 2011).

All these reasons combined, paved the way to the development of different methods to select and design new aptazymes. Ligand-dependent HHRZs were designed using rational methods by engineering the HHRZ sequence to fuse the aptamer in the right position for inducible activity (Amontov & Jaschke, 2006; Soukup & Breaker, 1999a; Tang & Breaker, 1997; Vaish *et al.*, 2002). Another strategy used a communication module by inserting a random sequence between the HHRz and the aptamer and selecting those with the desired response from the aptazyme (Auslander et al., 2010; Pu et al., 2020). The communication module could also be created by inserting mutations into the HHRz stem holding the aptamer (Wieland & Hartig, 2008). Finally, to select aptazymes responding to molecules for which no aptamers exist, a random sequence of up to 50 nucleotides was inserted in one of the HHRz stems (Ferguson et al., 2004; Koizumi et al., 1999a; Robertson et al., 2004).

Aptazymes were developed for different kinds of applications because they presented many advantages. As an example, allosteric ribozymes formed solely by RNA when used in cells have a low immunogenicity compared to proteins and antibodies. They are smaller in size and easier to express due to the absence of post-transcriptional and post-translational modifications. Moreover, they have the ability to function either in *cis*, cleaving the RNA directly in which it is inserted, or in *trans*, acting on a different target. Ligand-responsive ribozymes have been used in different applications mainly gene expression control in prokaryotic cells (Felletti *et al.*, 2016a; Wieland *et al.*, 2009; Wieland & Hartig, 2008) and mammalian cells (Beilstein *et al.*, 2015; Nomura *et al.*, 2013; Zhong *et al.*, 2016b) and even *in vivo* (Rehm et al., 2015; Win & Smolke, 2008). Aptazymes were also used as biosensors for different types of molecules, owing that to their ability to translate the detection of the ligand into a readable signal generated by the ribozyme's

activity (Ogawa, 2009; Rueda & Walter, 2006). Actually, a lot of DNA and RNA-based aptazymes were developed as biosensors for the detection of metabolites, environmental conditions, pollutants, proteins, etc (Gong et al., 2015; Jaffrey, 2018) (Debais et al., 2020; Frommer et al., 2015; Porter et al., 2017; Rueda & Walter, 2006).

In this paper, we are presenting a dual selection method of active and inactive aptazymes detecting proteins included in cancer and the spike protein from MERS-COV. These aptazymes can potentially be used as a biosensor for the detection of diseases and also to control gene expression in cells. The novelty we are presenting is the 'inactive' version of selection library, which has a mutated catalytic core of the HHRz shifting its activity from cleavage in presence of magnesium to cleavage in presence of manganese. The advantage of our method is the ability to get aptazymes that are highly active in absence of the target, but are inhibited in its presence, which is more challenging to perform with the existing methods. This is due to the fact that most selections made for the allosteric activity of HHRzs are done in presence of magnesium (present in the transcription buffer) which increase chances that highly active HHRzs cleave before even being able to detect the target. This leads to having aptazymes with low activity when the ligand is not in the media.

We performed fifteen rounds of SELEX, to select both protein induced and inhibited versions of aptazymes. Chosen enriched aptazymes sequences were tested for their activity. Those selected aptazymes were then used with fluorogenic aptamers to develop biosensors.

8.4. Materials and methods

- **Library construction**

We built two libraries: one for active aptazymes in presence of the target and one for the inactive aptazymes. The design was based on the structure of the theophylline aptazyme that was originally formed by the combination of a minimal HHRz and the theophylline aptamer inserted in stem II of the HHRz (Soukup & Breaker, 1999b). In our design, we replaced the theophylline aptamer by 50 random nucleotides which construct our library for the selection (**Figure 8.1**).

For the ribozyme, it is originally active in presence of 10 mM of Mg^{2+} . A recent study from our group has shown the discovery of a HHRz active in presence of Mn^{2+} but not Mg^{2+} (Naghdi et al., 2020). The only difference in this HHRz is a variation in the catalytic core

in position A6 (CUGANGA) that is in this case a C. In the same study, the mutation 'A6C' was performed in another HHRz, also leading to a HHRz, inactive with magnesium and active with manganese. Using this mutation, we were able to build two libraries with this one nucleotide difference in the catalytic core.

The libraries were named Library A for the WT ribozyme (CUGANGA) and library C for the mutated one (CUGCNGA) (**Figure 8.1**). The library A was used for the selection of aptazymes activated by the ligand, and the library C to select aptazymes inactivated by the ligand. Library C is used for inactive aptazyme selection to prevent cleavage during transcription (which contains 24 mM MgCl₂), giving time for the ligand to bind the aptazyme before the cleavage reaction. A summary of libraries and the activity of selected aptazymes is in **Table 8.1**.

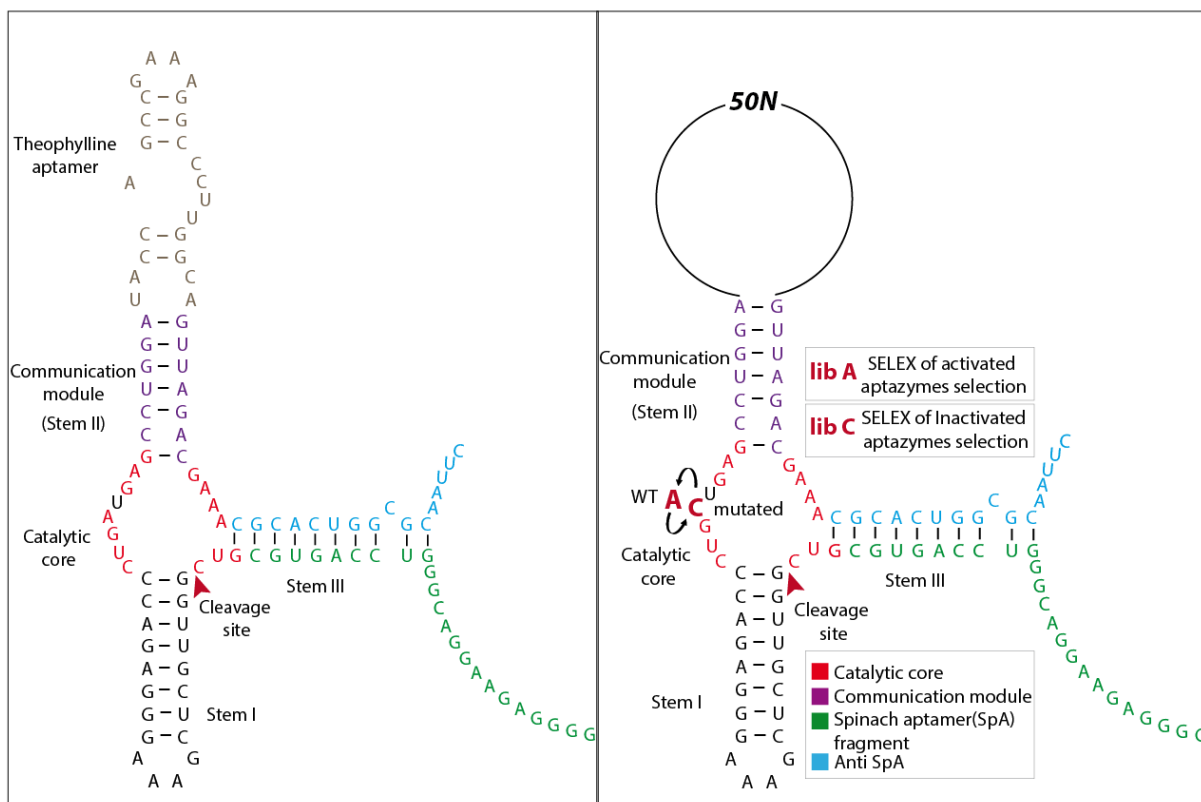


Figure 8.1 Library sequence design

Left: the original sequence of Theophyllin aptazyme with modified Stem I. Right: Our library design where theophylline aptamer was replaced with 50 random nucleotides. The design was the same for Lib (A) and Lib (C) except for the A6 position (arrows).

After transcription, when selecting for an aptazyme inactive in presence of the target, the ones that cleave will be discarded and those kept inactive are recovered and used for the

next selection cycle. The risk of using conventional aptazymes that cleave in presence of Mg^{2+} is that only slow ribozymes will be selected. Indeed, transcription buffer is concentrated in magnesium, which enhances the cleavage of “standard” active aptazymes. Also, the highly active aptazymes will cleave immediately after being transcribed, too fast to have the opportunity to bind the ligand supposed to inactivate the aptazyme. This implies that most of the good candidates of highly active aptazymes could be lost because they cannot be selected after the cleavage.

Hence, by using the library C with the A6C mutation, no cleavage will happen during transcription, even if the aptazymes have the active structure. Therefore, the aptazymes will have sufficient time to fold into their structure and bind the target. After two hours of transcription, the manganese can be added to the reaction and in this case, those that were inactivated when binding the target will stay intact and we can recover them. With this design, we have higher chances of selecting highly active aptazymes that are truly inhibited by the target.

Table 8.1 Summary of libraries and selection

Library	Activity of selected aptazymes	Nomination
Library A (Lib A)	Active aptazymes in presence of ligand	Ligand\uparrowRz
Library C (Lib C)	Inactive aptazymes in presence of ligand	Ligand\downarrowRz

- **Amplification of library**

The oligonucleotides for the libraries were ordered from IDT with 50 random nucleotides (in loop II) flanked by the core, stem III and a part of stem I. Two primers were ordered to amplify the library, the forward had a T7 promoter and the sequence completing stem I as well as a part binding to the library and a reverse primer covering the stem I (**Table 8.2**).

Table 8.2 Sequences for library amplification

Name	Sequence
Forward Library A	TAATACGACTCACTATAGGGGGGAGAAGGACGGGTCCAGTGCCTCGG TTGTTTCGAAAGGACGACCCCTGATGAGCCTGG
Forward Library C	TAATACGACTCACTATAGGGGGGAGAAGGACGGGTCCAGTGCCTCGG TTGTTTCGAAAGGACGACCCCTGCTGAGCCTGG
Library	GACGACCCCTGATGAGCCTGG (50N) TTAGACGAAACGCACCTGGCG
Reverse primer	CGCCAGTGCCTTCGTCTAA

RNA complete library sequence	GGGGAGAAGGACGGGUCCAGUGCUCGGUUGUUCGAAAGGGACGACC CUGAUGAGCCUGG (50N) UUAGACGAAACGCACUGGCGCAAUUC
--------------------------------------	--

Brown: T7 promoter, Green: stem I (spinach aptamer part), Red: catalytic core, black: stem III, violet: stem II (communication module)

The forward primer allows the regeneration of the cleaved part after each selection round. Both libraries (at 25 nmoles each) were amplified by PCR using the corresponding set of primers at a T_m of 51.3° C for lib (C) and 61.4 °C for lib (A). For amplification, for each reaction 1 U/μl of Hot start Taq DNA polymerase (QIAGEN) was used in a reaction mix in presence of its buffer (1X), Q solution (1X), 200 μM of dNTPs and 1 μM of each oligonucleotide. The size of PCR products was validated with 2% agarose gel. Using ethanol precipitation, 8 ml of PCR products for each library were pooled together and resuspended in 1.2 ml of Milli-Q water. The same protocol was used for the enrichment of the selected aptazymes after every SELEX round, but with the use of 100 μl of PCR product for the following rounds.

- **RNA library preparation**

For the RNA library named G0, 400 μl of PCR product was used for a total volume of 4 ml of *in vitro* transcription reactions. For the enzyme, a home produced T7 enzyme was used with transcription buffer (80 mM HEPES-KOH pH 7.5, 24 mM MgCl₂, 40 mM DTT, 2 mM spermidine) in presence of 3 mM of rNTPs mix (rATP, rGTP, rCTP and rUTP) and 1 U/μl of rnae inhibitors (NEB) and 1 U/μl of inorganic pyrophosphatase (Roche). After 2h30 incubation at 37°C, DNase I (NEB) was added to the reaction to degrade the DNA template for 30 minutes. The produced RNA was extracted using double phenol/chloroform extraction then precipitated with 100% ethanol. The precipitated RNAs were purified on 10% denaturing polyacrylamide gel (PAGE). For the library A, the band corresponding to the uncleaved aptazymes was cut for elution in 0.3 M NaCl elution buffer overnight at 4°C.

- **Validation of Library C cleavage in manganese**

To ensure the activity of aptazymes in presence of manganese instead of magnesium, purified RNA was incubated in cleavage buffer (100 mM NaCl, 50 mM Tris-HCl, pH 7 et 25 mM KCl), either in absence of metals or in presence of 10 mM Mg²⁺ or Mn²⁺. For comparison purpose, library A was incubated in presence or absence of Mg²⁺. All reactions were incubated at 37°C for 60 minutes. The reactions were stopped using 2X denaturing formamide dye (95% formamide, 10 mM EDTA, 0.025% bromophenol blue and 0.025%

xylene cyanol) and RNA fractions were separated on 10 % denaturing PAGE. To be able to visualize the results the gel was stained in TBE 1X buffer with 10 µl of gel stain (Transit) for 20 minutes and it was visualized under UV with a gel doc (Biorad).

- **CNBr agarose magnetic beads functionalization**

Four proteins (Human Galectin 7 (Gal-7) and Galectin 1 (Gal-1) from Yves St-Pierre's lab (loaded with lactose), MERS-COV Spike protein and a fragment of this same protein) were coupled separately to the CNBr agarose magnetic beads using the manufacturer's protocol with a concentration of 10 mg/ml of proteins for 300 mg/ml of beads. The functionalization was validated by a Bradford test where 20 µl of beads were incubated in 1 ml of Bradford buffer. The beads were visualized under optical microscope to see the color change from white for the negative control to blue in presence of proteins.

- **Affinity selection**

For the first selection cycle (G0), 1.5 nmoles of RNA (from library A or library C)(containing in total $\sim 10^{14}$ molecules, but considering that the template PCR had many copies for each template and similarly that the transcription produces many RNA molecules from each template, the actual sequence diversity is probably closer to $\sim 10^{12}$) was incubated with 100 µl of functionalized beads with the corresponding protein in 700 µl of SELEX buffer (20 mM Tris-HCl (pH:7.4), 2 mM CaCl₂, 5 mM KCl, 100 mM NaCl, 1 mM MgCl₂) and 1 µl of RNase inhibitor for two hours on a rotator. To remove the unbound RNAs, the tubes were put on a magnetic stand for two minutes allowing the beads to stick to tube walls leaving the supernatant containing the unbound RNAs. The supernatant was removed and replaced with 1 ml of SELEX buffer to wash the beads for 20 minutes while on rotation. This allows the weakly bound RNAs to also be removed to keep only those tightly bound to the target. After the washing step, the supernatant is again removed

To elute the bound RNAs, the beads were incubated for 15 minutes at 75 C in presence of RNase free water. These eluted RNAs were reverse-transcribed and amplified in one reaction using a one-step RT-PCR kit from QIAGEN. The reaction contained 1 µl of one-step RT-PCR enzyme mix with its buffer at 1X concentration, 200 µM of dNTP mix, 1X of Q solution and the volume was completed with the eluted RNA and RNase free water. Cleaved RNAs were regenerated by the forward primer because it included the missing portion during the PCR step (**Figure 8.2**). After checking the PCR products sizes using 2% agarose gels, the DNA was precipitated with ethanol then transcribed *in vitro* using T7

polymerase to obtain the G1 RNAs for the second cycle. Then, half of the DNA produced by each PCR after every cycle was used for transcription, the other half was archived. The rest of the rounds conditions are in **Table 8.3**.

Table 8.3 SELEX rounds and conditions

Round	Beads (ligand)	Incubation	Washes
1	100 μ l (300 mg/ml)	Two hours	1 : 20 minutes
2	20 μ l (300mg/ml)	Two hours	1 : 20 minutes
3	20 μ l	1h45 minutes	1: 25 minutes
4	18 μ l	1h30 minutes	1 : 30 minutes
5	16	1h15	1 : 35 minutes
6	14	1h	2 : 30 minutes each
7	12	45 minutes	2 : 30 minutes each
8	10	30 minutes	2 : 30 minutes each
9	10	15 minutes	2 : 30 minutes each

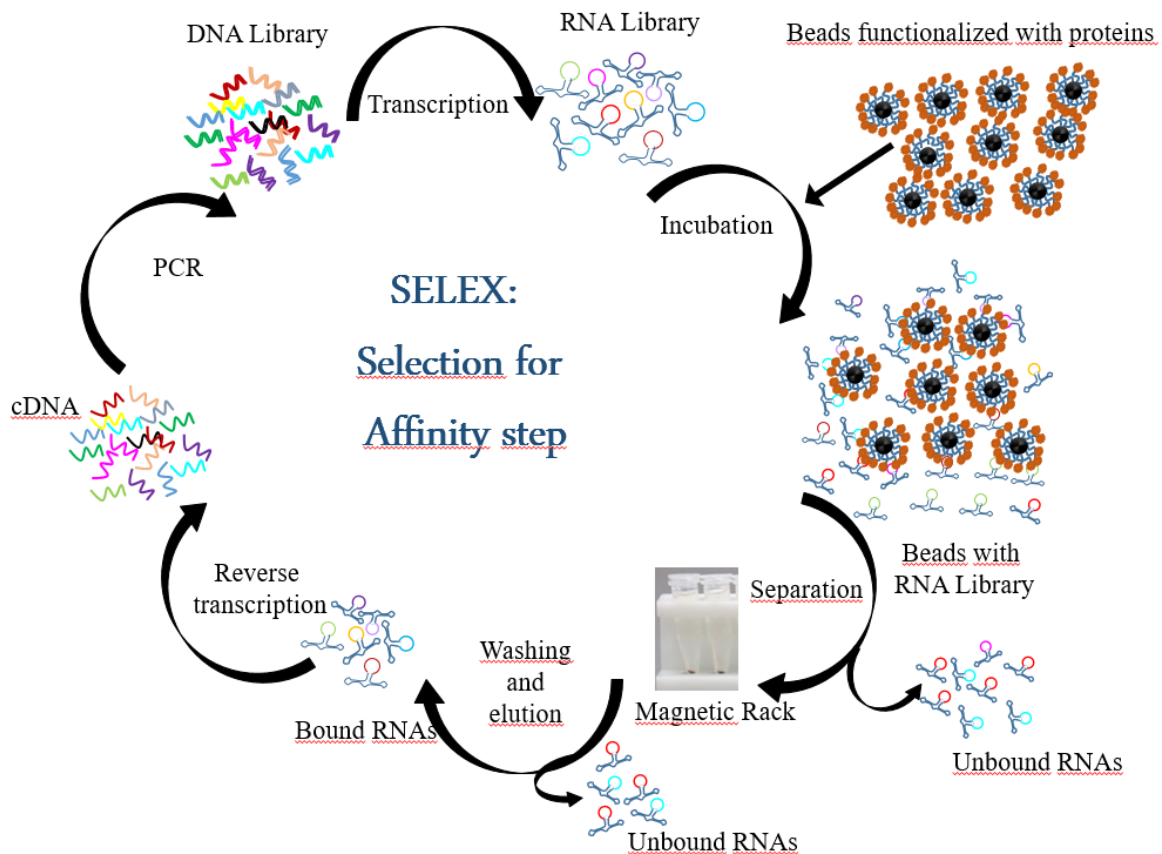


Figure 8.2 Affinity Selection steps

- **Negative selection**

After round 8, a counter selection using naked beads was applied. After round 9 negative selection was applied using different proteins than the final ligand for the aptazyme. the negative selection was applied after every cycle (conditions in **Table 8.4**)

Table 8.4 Negative selection conditions

Round	Protein selected for	Counter selection protein
After round 8	Gal-7-Gal-1	Naked beads
After round 9	Gal-7	Gal-1
After round 9	Gal-1	Gal-7
After round 10	Gal-7	Spike protein
After round 10	Gal-1	Spike protein
After round 11	Gal-7	Gal-1
After round 11	Gal-1	Gal-7
After round 12	Gal-7- Gal-1	Lysozyme
After round 13	Gal-7- Gal-1	Spike
After round 14	Gal-7- Gal-1	Lysozyme

- **Allosteric Selection**

After nine cycles of selection for the affinity only, cycles ten to fifteen were done based on the allosteric activity of the aptazymes. For selection of aptazymes induced by ligand, one step selection was used. During transcription, 10 µl of functionalized beads with target ligand, was incubated with the reaction mix containing DNA of library A. After degrading the DNA template using DNase I (NEB), RNA was extracted and fractioned on 10 % denaturing PAGE. The band corresponding to the cleaved RNA length was cut from the gel and eluted overnight at 4°C or one hour at room temperature. After precipitation, the RNA was resuspended in 25 µl of SELEX buffer and used for the negative binding selection step in presence of a non-specific protein. The unbound RNA, present in the supernatant is used for the RT-PCR to produce the enriched DNA for the next cycle.

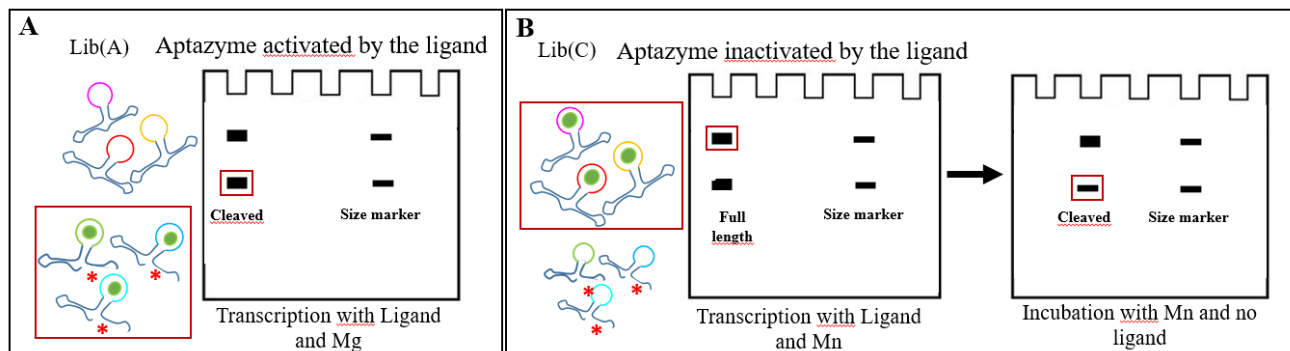


Figure 8.3 Activity based selection steps

The red boxes indicate the bands cut from the gel for the purpose of retrieving the chosen RNA. **A:** for lib(A) the selection is done in one step where the cleaved bands are cut and eluted. **B:** For lib(C) RNAs eluted from the first gel corresponding to uncleaved ribozymes (left) are incubated with only manganese to pick those active in absence of the ligand (left).

For selection of aptazymes inhibited by ligand, two steps of selection were used. At first, after 2 hours of incubation with the target protein during transcription, 10 mM of Mn^{2+} was added to the reaction mix, allowing the cleavage reaction to occur if unwanted active aptazymes were present. After running the RNAs on PAGE, the uncleaved RNAs were selected and purified. To select only aptazymes that were originally active, but were inactivated by binding the target protein, the purified RNAs were incubated in the cleavage buffer (100 mM NaCl, 50 mM Tris-HCl pH 7 and 25 mM KCl) and in presence of 10 mM Mn^{2+} without the protein 30 min at 37°C. After that step, following separation on gel, the cleaved RNAs were precipitated after elution and used for the negative binding selection step with non-specific proteins. To detect sequence enrichment, the PCR product of different rounds (G0, G9, G12 and G15) were sent for deep sequencing by MiSeq (Genome Quebec) and the enriched sequences were tested for their affinity and their cleavage activity.

- **Sequencing results analysis**

MiSeq (Illumina) results have shown an enrichment of a few sequences in the last round. Two enriched sequences for inactivated aptazymes were chosen to be tested for their activity in presence of the targets. The chosen sequences were enriched for both Gal-7 and Gal-1. For the activated aptazymes, one sequence was picked for each protein (**Table 8.5**).

Table 8.5 Aptazymes sequences

Name	Sequence
Gal-7/1↓Rz 1	GGGGAGAAGGACGGGTCCAGTGCCTCGGTTGTTCGAAAGGACGACCC TGCTGAGCCTGGGTGTATCTACTTGAAGAGAGCCACATTGGTTTGC CTCTGTGTGTGCACCTTAGACGAAACGCACTGGCG
Gal-7/1↓Rz 2	GGGGAGAAGGACGGGTCCAGTGCCTCGGTTGTTCGAAAGGACGACCC TGCTGAGCCTGGTACCTACAGACAGGAAGCCGAAATCATGCGTGTT TCCACACGGTGCCTTAGACGAAACGCACTGGCG
Gal-7↑Rz	GGGGAGAAGGACGGGTCCAGTGCCTCGGTTGTTCGAAAGGACGACCC TGATGAGCCTGGAAGCTGATACACTAGTGCCTTCGAGCTCCAATTAT GGCGAGAACGTGGCGTTAGACGAAACGCACTGGCG
Gal1↑Rz	GGGGAGAAGGACGGGTCCAGTGCCTCGGTTGTTCGAAAGGACGACCC TGATGAGCCTGGATAGGCAGAGCACCATTTTTGCGCTCAACGTGTGG AATGACTTTGGCTGGTTAGACGAAACGCACTGGCG
Forward	GGGGAGAAGGACGGGTCCAGTGC
Reverse	CGCCAGTGCCTTTCGTCTAA

Color code: red : catalytic core, green:SpA, blue : anti SpA, dark red : aptamer sequence. Arrows : ↓ inhibited, ↑ induced.

- **A6C mutation reversing**

The C6 mutation in the aptazymes sequences was reverted to A to check if our hypothesis of selecting new ribozymes using the C in position 6, would lead to aptazymes active in standard conditions, i.e. with magnesium rather than manganese. The sequences were amplified using a primer with A6. The templates for the mutated and the WT versions of the aptazyme sequences were transcribed in presence of magnesium (present in transcription buffer). The cleavage activity was detected by running the RNA, after ethanol precipitation, on a 6% denaturing polyacrylamide gel.

- **Aptazyme activity validation**

The aptazyme DNA sequences were ordered from Twist biosciences and amplified by PCR using the primers in **Table 8.5**. For PCR the T_m used was 56°C and Taq DNA polymerase and its buffers from NEB were used. After size verification by agarose gel electrophoresis, PCR products were precipitated by ethanol and used as templates for radioactive transcription. Transcription was done using [α -³²P]UTP and T7 polymerase in presence of transcription buffer and rNTPs (10 mM except for UTP that was used at 4 mM concentration) for 2h30 at 37°C. Transcription was performed in absence of manganese to make sure that no cleavage can occur. RNA was precipitated by ethanol then purified on a 6% denaturing polyacrylamide gel. The RNA was then eluted from the gel overnight at 4°C. The RNA was again precipitated with ethanol.

Cleavage activity was tested in presence of increasing protein concentrations. Because, inactive aptazyme sequences were enriched for both galectins, each sequence was tested with both proteins. The galectins concentrations ranged from 100 nM up to 4 μ M. Each reaction contained the cleavage buffer (10 mM $MgCl_2$, 100 mM NaCl, 50 mM Tris HCl, pH 7.4, 25 mM KCl), 4 μ l of radiolabelled RNA, 2 μ l of protein (10X of the corresponding concentration) and 4 mM of $MnCl_2$ for inactivated aptazymes or 10 mM $MgCl_2$ for the activated ones. Before adding the divalent ion to the reaction, all the components were mixed and the manganese or magnesium was added immediately before the incubation at 37°C for one hour. To detect the cleavage activity, the reaction products were fractioned on a 6% denaturing polyacrylamide gel that was later exposed overnight to a phosphor screen.

- **Biosensor design**

To use the aptazymes as a diagnostic tool, we wanted to prove that we can make it work as a biosensor. To do so, the mango aptamer sequence was added to each aptazyme via a linker sequence on the 5' strand of stem I (**Figure 8.4**).

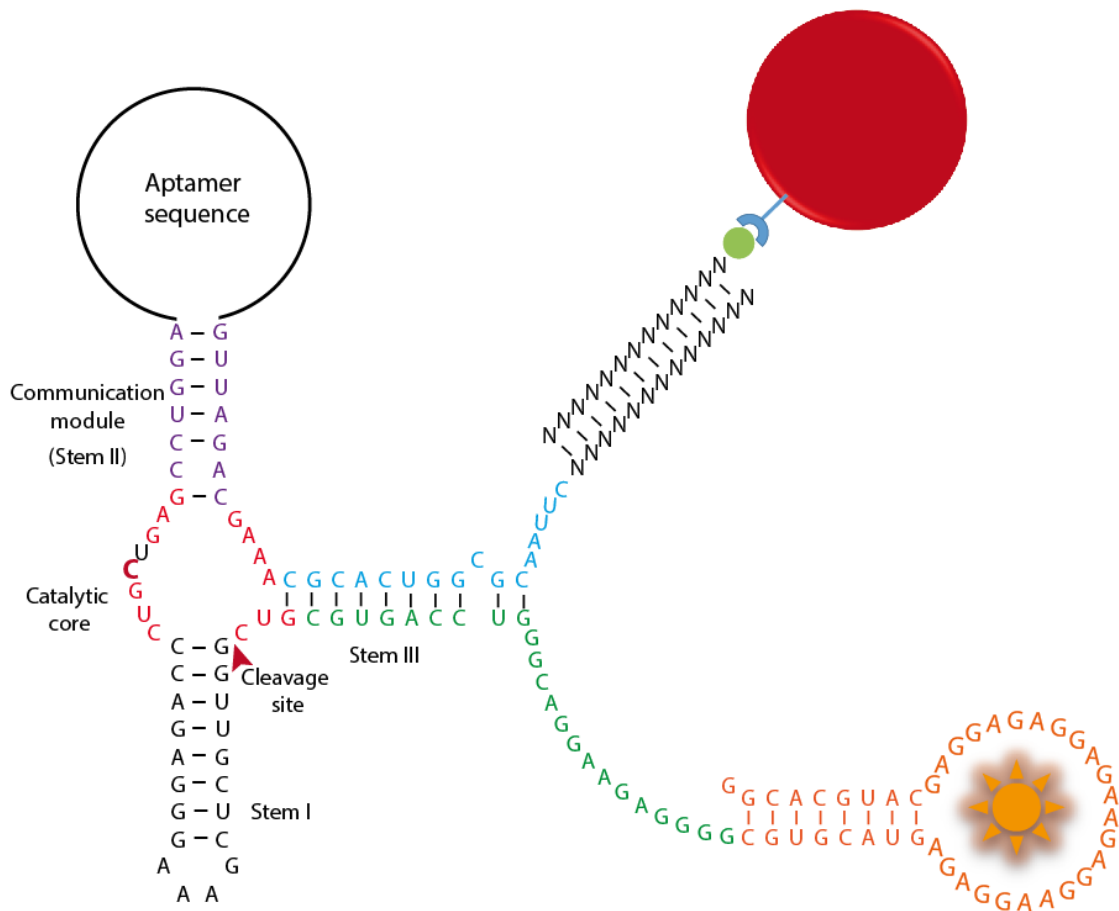


Figure 8.4 Aptazyme-mango aptamer complex functionalized on streptavidin bead

A fluorescent bead (red) functionalised with a biotinylated oligo with a complementary sequence to the aptazymes 3' part of stem I. Mango aptamer (orange sequence) was bound to stem I from the 5' side.

The concept consists in using the fluorescence of mango aptamer as an indicator of the presence or absence of the target protein. The idea is that whenever the ribozyme is activated, the mango aptamer sequence is separated from the aptazyme due to the cleavage reaction. To be able to remove the cleaved sequences, each aptazyme and mango aptamer complex was immobilized on an internally dyed bead. When mango aptamer is still attached to the aptazyme, the fluorescence of TO1-biotin will be detected simultaneously with the beads' fluorescence. In the opposite case, when the mango aptamer is released after cleavage, only the beads' fluorescence will be detected.

- **PCR and transcription**

The DNA matrices used were the same as above. To add the mango aptamer sequence a forward primer was used with an overhang containing the aptamer sequence and a linker

sequence (Table 8.6). The reverse aptamer had an overhang with a sequence complementary to a biotinylated oligonucleotide. The biotinylated oligo is used as a bridge to bind the aptazymes to the streptavidin beads.

Table 8.6 Primers and biotinylated oligo sequence

Name	Sequence
Forward Apta+Mango	TAATACGACTCACTATAGGGGGCACGTACGAAGGAGAGGAGAGGAA GAGGAGAGTACGTGCAGGACGGGGAGAAGGACGGGTCCAGTGC
Reverse Apta	CGTATTCTGCCGCCATCTTACTCGCCGCCAGTGCGTTTCGTCTAA
Beads binding oligo	CGCTCATTCTACCGCCGTCTTATGC-Biotin

- **Streptavidin beads functionalization**

Quantumplex streptavidin beads from BangLabs, are five populations of 4.4 µm beads, internally dyed with different degrees of starfire red. The beads were functionalized with a biotinylated oligo that have a complementary sequence to an added sequence on the 3' strand of stem I of each aptazyme. The functionalization was done according to the manufacturer's protocol. Each beads population was later incubated with one of the aptazyme-mango complexes in a buffer containing 250 mM Tris HCl, pH 7.5, 25 mM KCl, 100 mM NaCl. The hybridization reaction was done at 37°C for 20 minutes. Two washing steps were applied using TTE buffer (20 mM Tris HCl, pH 7.5 and 0.002 % Tween) to remove any unbound sequences.

- **Cleavage tests and fluorescence detection**

For each aptazyme, the functionalized beads were incubated in presence of cleavage buffer (100 mM NaCl, 50 mM Tris HCl pH 7.5), 2 µl of TO1-Biotin and an excess concentration of the target protein (6 µM). Also for specificity evaluation purpose the complexes were incubated with spike protein (non-specific).

Aptazymes binding both galectins were incubated with either Gal-7, Gal-1, spike or without any protein. The Gal-7-activated aptazyme and the one activated by Gal-1 were each incubated with or without the corresponding protein. To activate the cleavage reaction, 10 mM of MgCl₂ was added to the aptazymes with the WT sequence and 4 mM of MnCl₂ to the inactivated aptazymes just before the incubation at 37°C.

8.5. Results

8.5.1. Validation of library (C)

Both WT and mutated libraries were transcribed and incubated with either magnesium or manganese. The WT library was completely inactive in absence of magnesium but active in 10 mM magnesium. On the other hand, the mutated library was inactive in presence and absence of magnesium, but active with manganese (**Figure 8.5**). The C6 core variant was characterized by our group and we saw it as an opportunity for the selection of inactive aptazymes using this strategy.

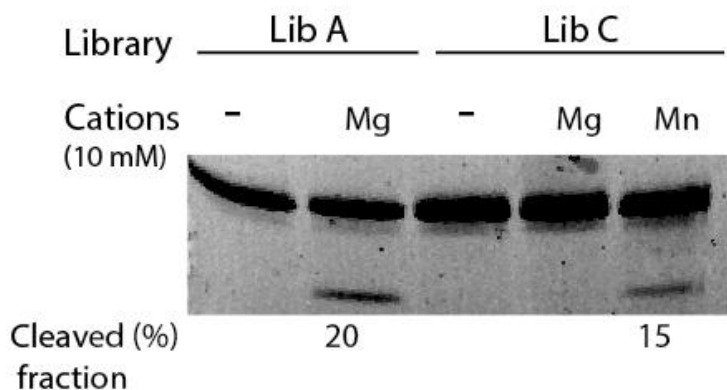


Figure 8.5 Libraries activity in presence of magnesium or manganese

8.5.2. SELEX

The selection for inactive aptazymes in presence of Gal-7 and Gal-1 was carried out following fifteen cycles, nine cycles for affinity selection followed by six cycles for cleavage activity selection which includes activity counter selection with ligand and activity selection in the absence of ligand.

The affinity selection was performed by magnetic affinity purification with the target protein ligands. During the nine affinity-based rounds, transcription in absence of ligand showed no cleavage for the library A. After the ninth round, during transcription with the ligands, cleavage was detected for the library A, and for the mutated library when manganese was added to the reaction in the last 30 minutes.

For the library C, when the ribozymes inhibited in presence of proteins were recovered from the gel and incubated with manganese in absence of proteins, cleavage was detected, which implies that there was an enrichment of sequences that are able to bind

the proteins and be inactivated in their presence. This also indicated our strategy to select aptazymes was successful.

This is the first study that employs a library with a mutated ribozyme core that gets activated with a component other than the target ligand itself. This strategy circumvents the problem of selecting inactive aptazymes. Otherwise, if the consensus core is used, the final selected aptazyme would likely have a poor cleavage efficiency. This is because they would be selected among the ribozymes that cleaved the slowest during transcription.

8.5.3. Sequencing results

Several artefacts (wrong size of the library, primer dimers and polymers) revealed by deep sequencing complicated a thorough and quantitative analysis of the selected sequences. Nevertheless, the relevant sequences from the last round of SELEX are shown in **Table 8.7**. Some sequences were the same for both Gal-7 and Gal-1, which indicates that they probably bind a common region in the proteins as there was 36% of similarity in the amino acid sequences of both proteins with a region of six successive identical amino acids. Structure prediction of these aptazyme sequences are in **Figure 8.6**.

Table 8.7 Frequency of chosen sequences (from sequencing results)

Name	Sequence	Enrichment
Gal-7/1 ↓Rz 1 (Lib C-inhibited)	GGGGAGAAGGACGGGTCCAGTGCCTCGGTTGTTTCGAAAGGACGACC CTGCTGAGCCTGGGTGTATCTACTTGGAAAGAGGCCACATTGGTTT GCCTCTGTGTGTGCACCTTAGACGAAACGCCTGGCG	279/707 Gal-7 4/13 Gal1
Gal-7/1 ↓Rz 2 (Lib C-inhibited)	GGGGAGAAGGACGGGTCCAGTGCCTCGGTTGTTTCGAAAGGACGACC CTGCTGAGCCTGGTCACCTACAGACAGGAAGCCGAAATCATGCGTG TTCCACACGGTGCCTTAGACGAAACGCCTGGCG	5889/7104 Gal-7 4/17 Gal1
Gal-7 ↑Rz (Lib A-activated)	GGGGAGAAGGACGGGTCCAGTGCCTCGGTTGTTTCGAAAGGACGACC CTGATGAGCCTGGAAGCTGATACACTAGTGCCTTCGAGCTCCAATT ATGGCGAGAACGTGGCGTTAGACGAAACGCCTGGCG	2/30
Gal-1 ↑Rz (Lib A-activated)	GGGGAGAAGGACGGGTCCAGTGCCTCGGTTGTTTCGAAAGGACGACC CTGATGAGCCTGGATAGGCAGAGCACCATTTTTGCGCTCAACGTGT GGAATGACTTTGGCTGGTTAGACGAAACGCCTGGCG	2/33

Color code ; red : catalytic core, green:SpA, blue : anti SpA, dark red : aptamer sequence

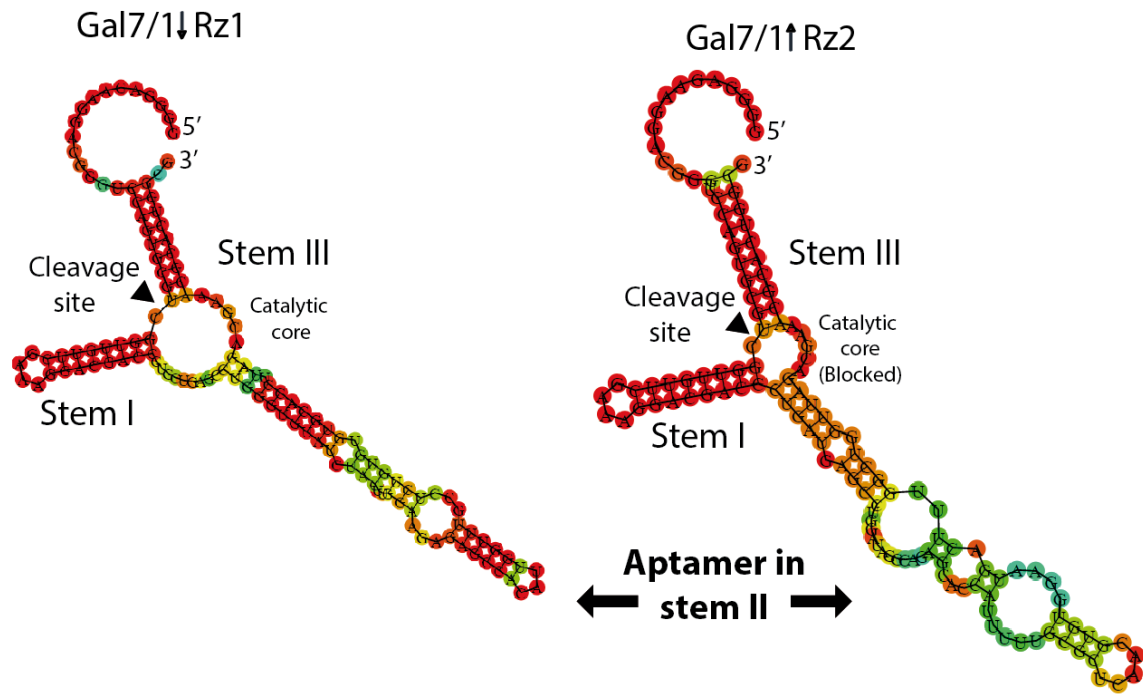


Figure 8.6 Predicted structures of two aptazymes binding Gal-7

Structure predictions obtained by Vienna RNA tool. The colors indicate the base pairing probabilities: Orange to red: higher probability of base-pairing, green-yellow: lower probability of base pairing. The Gal-7/1↓Rz 1 (Left) forms an active structure where the catalytic core is correctly formed in absence of the ligand. Whilst Gal-7↑Rz's (right) catalytic core is not fully free to have an active cleavage which will probably be rearranged with the binding of the ligand allowing the activation of the aptazyme.

8.5.4. Reversing the C6 back to A6 consensus core

To validate the activity of ribozymes when the C6 is reverted to A, the inhibitable aptazyme **Gal-7/1**↓**Rz1**, **Gal-7/1**↓**Rz2** as well as **Gal-7/1**↓**Rz1C6A** and **Gal-7/1**↓**Rz2 C6A** were tested during transcription (presence of magnesium)(**Table 8.8**). The gel results have shown high cleavage rates during transcription for sequences reverted to A6 (71 and 66%), while C6 had a low activity in absence of manganese (21%) (**Figure 8.7**). This result confirmed our hypothesis that selecting with the A6C mutation can be used for selection and reverting it later would allow the use of these aptazymes in living organisms since they are active in presence of magnesium at physiological concentration once the mutation is reverted.

Table 8.8 Inactive aptazymes with reverted mutation

Selected inactive aptazyme	Aptazyme with reverted mutation
Gal-7/1 ↓ Rz1	Gal-7/1 ↓ Rz1 C6A
Gal-7/1 ↓ Rz2	Gal-7/1 ↓ Rz2 C6A

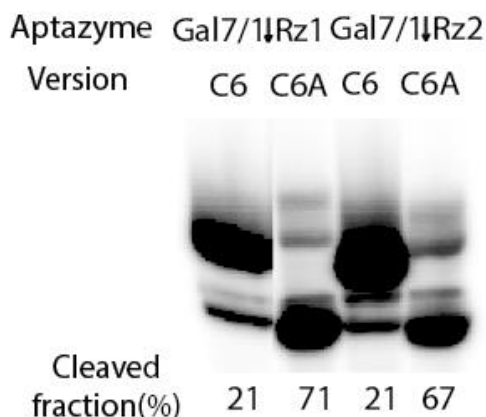


Figure 8.7 C6 aptazymes and with reverted mutation (C6A)

Aptazymes binding both galectins were tested with the mutated core (C6) and the reverted mutation (C6A) for activity during transcription.

8.5.5. Aptazyme activity is modulated by galectins

To confirm the activity of the selected aptazymes, after RNA synthesis, each sequence was incubated with increased concentrations of each protein. The aptazyme **Gal-7/1**↓**Rz1** showed an activity of 50 to 60% without protein and in presence of concentrations of gal-7 below 3 μM. When Gal-7 concentration was 3 and 4 μM a decrease in cleavage activity was detected (41 and 32%). This indicates that the target was able to inhibit the

aptazyme's cleavage activity. For the same aptazyme, 0.4 μM Gal-1 partially inhibited activity (**Figure 8.8 A**). These results are preliminary because the gel showed a lot of RNA degradation that might be caused by an RNase contamination.

Activity of the second aptazyme **Gal-7/1 \downarrow Rz2** decreases with both Gal-1 and Gal-7 at 2 μM (cleavage from 40-48% to 10-15%) (**Figure 8.8 B**). For the inducible aptazymes, two sequences were tested, one for Gal-7 and the other for Gal-1. The aptazymes **Gal1 \uparrow Rz** labeling was not very successful so the results were unclear (not shown). In the case of aptazyme **Gal-7 \uparrow Rz**, the cleavage rate increased at almost 15% (from 31% to 45%) in presence of high concentrations of Gal-7 (**Figure 8.8 C**).

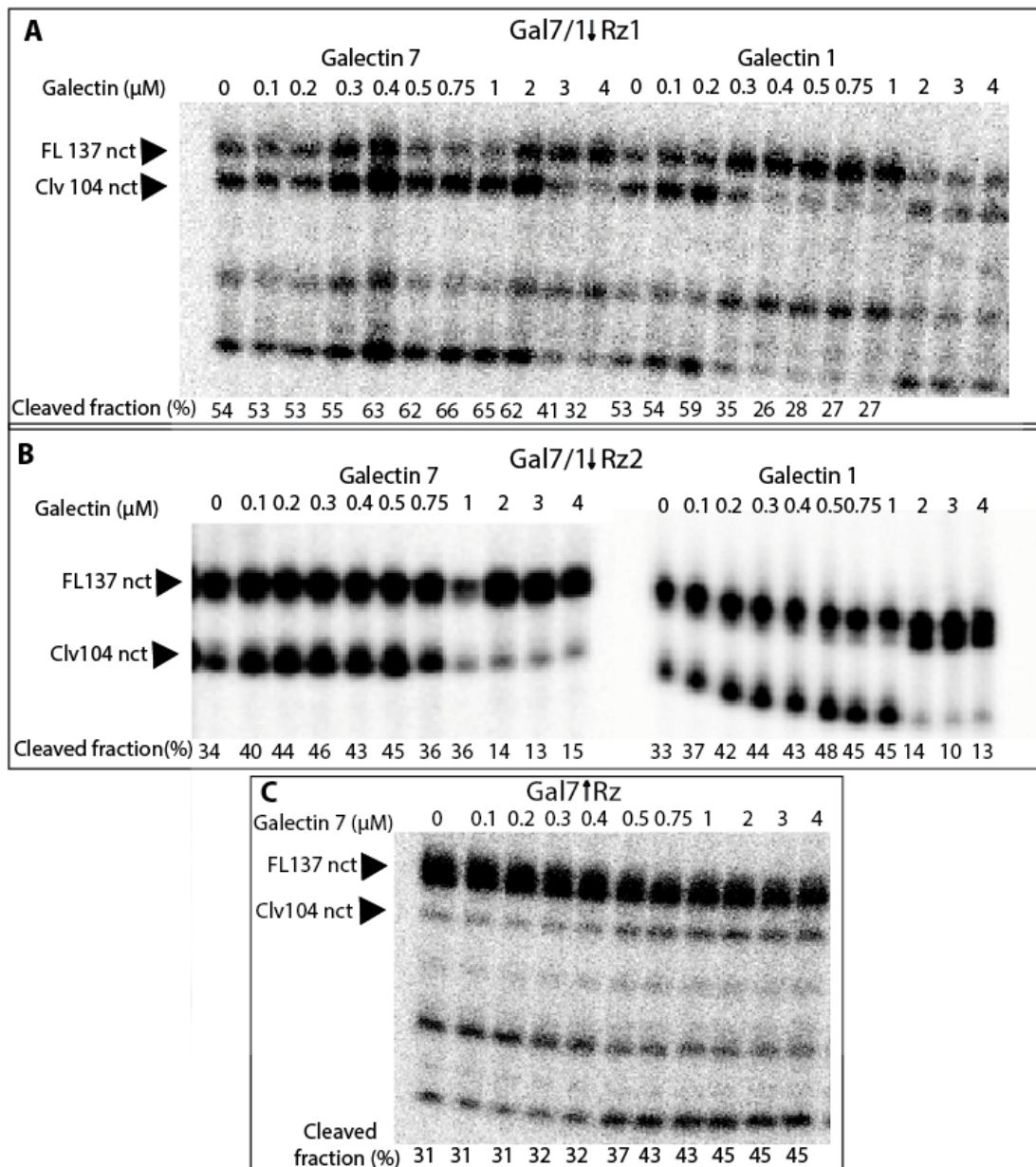


Figure 8.8 Aptazyme cleavage activity

A and B : Aptazymes Gal-7/1↓Rz1 and Gal-7/1↓Rz2 were tested for self-cleavage in presence of either Gal-7 or Gal-1. **(C)** Aptazyme Gal-7↑Rz was tested for cleavage in presence of Gal-7. Aptazymes Gal-7/1↓Rz1 and Aptazymes Gal-7↑Rz were labeled with alpha-³²P-UTP and aptazyme Gal-7/1↓Rz2 was labeled with gamma-³²P-ATP .The cleavage rate is presented in the bottom of the gels in (%). FL : Full length- Clv : Cleaved. The presence of multiple bands is the result of a probable non specific degradation generating a smaller product size for both the FL and the cleaved regions. The different distance between the bands from one gel to another is due to the gel size and the running time.

8.5.6. Aptazyme-based biosensor development

The idea behind building the biosensor is to have a system that allows multiplexing. The use of a fluorescent bead and a fluorescent aptamer allow the detection of two signals, one to identify the molecule present with the color of the bead and one to quantify the concentration of the detected molecule. This will be possible once beads of different colors are used for each detected molecule or with a different degree of the same color. This idea is based on the Luminex system, which uses beads of different degrees of red color on which primary antibodies are functionalized. The antibodies bind a specific ligand and the revelation of its presence is done by the use of a secondary antibody linked to a fluorophore. Even if we have tried to identify only one analyte at a time, in principle, the beads used in this project allow multiplexing. Future work will explore multiplexing possibilities with aptazymes.

The functionalization of DNA-aptazyme-mango aptamer complex on the beads was confirmed by mango fluorescence detection when the RNA is present. The confocal microscope images (**Figure 8.9**) show that the aptazyme are responding as expected.

In the case of the aptazymes induced by **Gal-7 \uparrow Rz** and **Gal1 \uparrow Rz**, in absence of the proteins, no cleavage occurred resulting in the detection of the enhanced TO1-Biotin fluorescence indicating the presence of mango aptamer on the beads. On the other hand, when the proteins are added to the cleavage reaction, the TO1-biotin fluorescence is lost confirming the separation of mango aptamer from the DNA-aptazyme-mango complex functionalized on the beads (**Figure 8.9 A-B**).

Both inhibitable aptazymes **Gal-7/1 \downarrow Rz1** and **2** were inactive in presence of manganese and 6 μ M of Gal-7 or Gal-1, while they were active in absence of the proteins (**Figure 8.9 C-D**). This change in activity was detected when mango aptamer fluorescence remains bound to the beads when the proteins are present in the solution, proving that the cleavage did not occur. The opposite was seen when the functionalized beads were incubated without galectins, then washed, no signal was detected, meaning that the cleavage had taken place, resulting in the loss of the mango aptamer at the washing step.

The results showed here are preliminary but encouraging. The beads used are made of polystyrene so their manipulation was quite hard and many were lost during transcription as well as many showed aggregation. This explains the different bead numbers showed on the **Figure 8.9** and the orange or yellow halo around the beads.

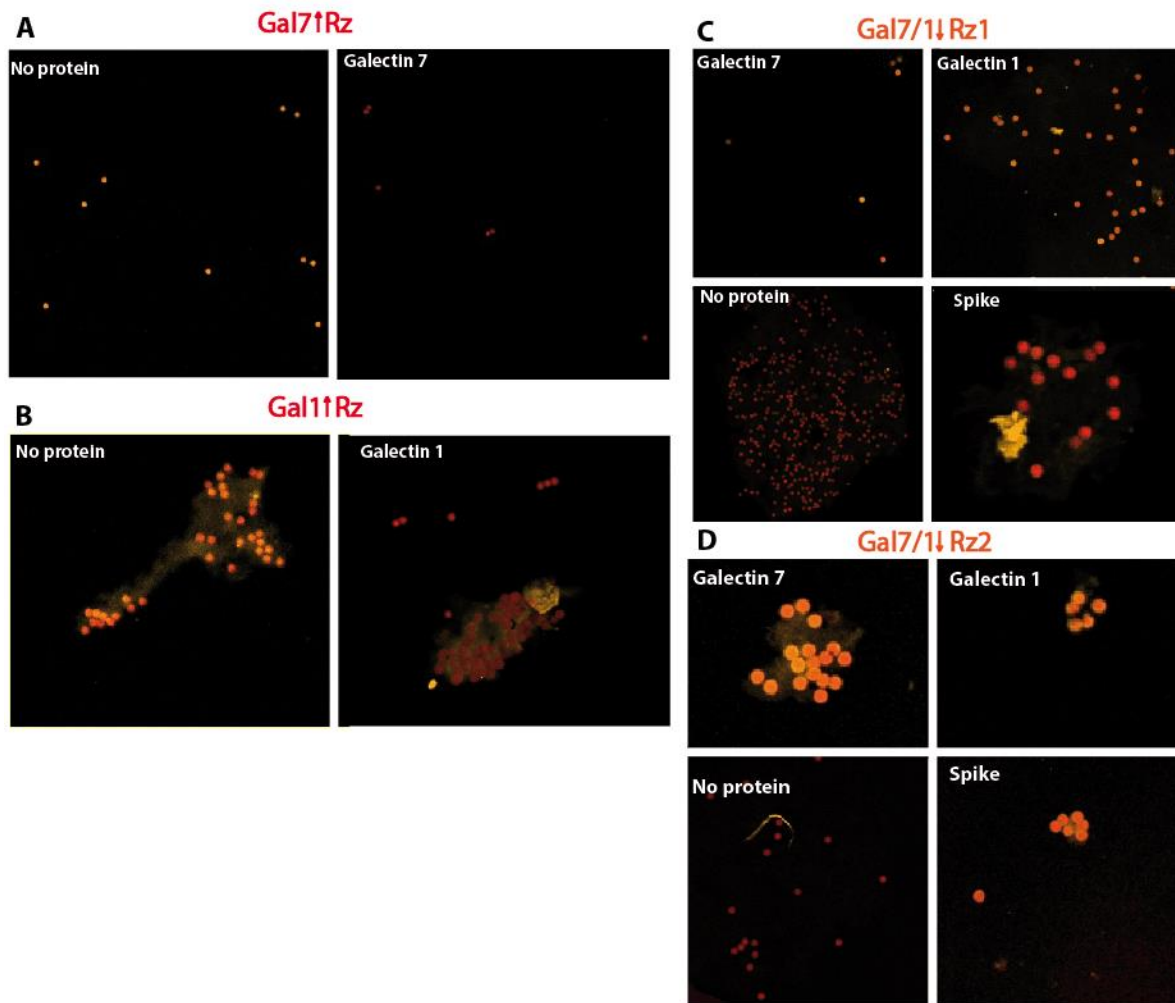


Figure 8.9 Confocal microscope images of functionalized beads after cleavage reactions

Images were taken using the same parameters using Cy5 wavelength for detection (635nm) of the starfire red (beads fluorescence) and Cy3 wavelength (569 nm) for TO1-Biotin (mango aptamer fluorescence). Both fluorescences were superposed for each picture taken to show the presence of beads functionalised with the RNA having or not the mango aptamer.

To confirm the specificity of the aptazymes' activity, a non-specific protein, Spike protein from MERS-COV-2 was used as a negative control. Spike was incubated at the same concentration with the different bead populations functionalized with the different aptazymes. The results supported their specificity when the aptazymes acted the same way in presence of spike as in absence of galectin proteins in case of aptazyme **Gal-7/1↓Rz1** (**Figure 8.9-C**), whereas aptazymes 2 seems to be less specific (**Figure 8.9-D**).

This test is not enough to define the specificity of the aptazymes, but it indicates that, at least, the aptazyme **Gal-7/1↓Rz1** does not bind to any protein since it does not binds to

Spike (from MERS-CoV). To determine the real level of specificity of these aptazymes, their affinity to other galectins and proteins present in the same cell type should be evaluated.

8.6. Discussion

Aptazymes make a strong gene regulation tool due to the flexibility of their design and the versatility of their applications in prokaryotic and eukaryotic cells. They also have the advantage to be active in *cis* or in *trans*, and in both cases the cleaved target can be easily modified (Ribosoft2.0 Article 3).

One method to build an RNA aptazyme is to select a new aptamer while using a known ribozyme in which a random sequence is inserted (Ferguson *et al.*, 2004; Koizumi *et al.*, 1999b; Robertson & Ellington, 1999; Robertson & Ellington, 2001; Robertson *et al.*, 2004; Townshend *et al.*, 2021a). This method was less used compared to the use of known aptamer to insert in the ribozymes sequence (Ogawa, 2011; Ogawa, 2014; Pu *et al.*, 2020; Stifel *et al.*, 2019b; Tang & Breaker, 1997; Zhong *et al.*, 2016b). This method has the advantage of selecting an aptazyme for any desired target depending on the final application and not being limited to the library of the existing aptamers. On the other hand, many published aptamers are very dependent to their selection environment (Byun, 2021). This affect their activity in further applications, making it hard to choose an aptamer sequence to use for making a new aptazyme. By using a selection of a complete structure already combining the sensor and the effector parts, these risks are limited. Added to that, the activity of the final aptazyme can be established from the start, and the conditions will be decided and optimized to result in the best combination possible, responding to the application's needs.

Bringing this method one step further in this paper we present a new method of selecting aptazymes inactivated by the target, by ensuring they have high cleavage activity when the ligand is absent. We show the results of using a variant-core ribozyme library activated by manganese to ensure the absence of cleavage during transcription. To our knowledge this is the first study to prove it is possible to control the time at which the cleavage is activated during transcription when the aptazyme ligand is present. This possibility gives the RNAs in solution time to fold into their correct shape when bound to the target. Then, when the manganese is added, the cleavage will occur only for the ribozymes activated by the proteins or unbound to any ligand.

When all ribozymes in solution are active only when manganese is added, it is as if we are giving all the sequences an equal opportunity of binding the target which will define their final shape. This would not be possible if ribozymes were able to cleave with magnesium, where the ribozymes with the highest activity cleave as soon as they are synthesized. These RNAs would not have the chance of binding the desired targets due to their cleavage. Meanwhile, using the A6C mutant selection strategy we increase the probability of picking aptazymes that are highly active in absence of targets, but are inactivated in its presence.

With our results we showed that the obtained aptazymes are clearly active, but some optimization still needs to be done to have a better specificity and a better ON/OFF switch. We also showed that we could reverse the mutation to have active ribozymes in presence of magnesium. These results show that our selected aptazymes can be used for *in vitro* applications when mutated, and *in vivo* application by simply switching back to the consensus catalytic core sequence.

The aptamer translates the presence of the target into a structural change in the ribozyme's overall conformation, triggering or inhibiting a cleavage reaction as a result. This change can be translated into a detectable signal as exemplified by the biosensors based on aptazymes. Most of the existing aptazymes were selected to be activated by the aptamer's target. In this paper, we present our biosensors based on the detection of Gal-7 and Gal-1 that are highly expressed in many aggressive cancer types.

These proteins are expressed in both intra- and extra-cellular media. They are currently detected by techniques based on the use of antibodies, like ELISA and Luminex assays, that are efficient enough in this purpose.

In this project, the use of these proteins was just an example allowing the proof of concept in detection of disease biomarkers with an RNA or nucleic acid based tool. The aim is to have a less time consuming system that allows multiplexing with the use of less material, and thus cheaper

The biosensor is based on the fluorescence emitted by mango aptamer when the target is bound to the aptazyme. The use of a combination of fluorescent beads and a fluorescent aptamer allows the detection with different methods, like microscopy, fluorescence reading and flow cytometry. The latter technique would allow the detection of multiple proteins in the same solution if beads of different colors are used for each

target. This is similar to the multiplex immunoassay (Luminex) capable of sorting up to 100 different targets. The advantage of using aptazymes instead of the traditional multiplex immunoassay is the reaction time where the cleavage reaction time can be only 30 minutes and with optimized aptazymes it can even be lowered. Also, the existing Luminex multiplex immunoassay kit comes already with antibodies functionalized on the beads, which first limits the targets to detect as well as the lifetime of the kit. In our suggested method, the user needs to order only the beads and functionalize them with the desired target aptazyme that can be easily synthesized in a lab when needed. The only limit of our technique is the use of RNA that is unstable in physiological liquids, but this can be fixed by adding modified ribonucleotides that protect RNAs from ribonucleases attacks making its half-life longer. The stability problem can partially be circumvented by using DNAzymes. However, DNAzymes also come with their limitations. For instance, DNA cannot be “expressed” in cells as RNA-based aptazymes.

Récapitulatif final

Moitié senseur du circuit

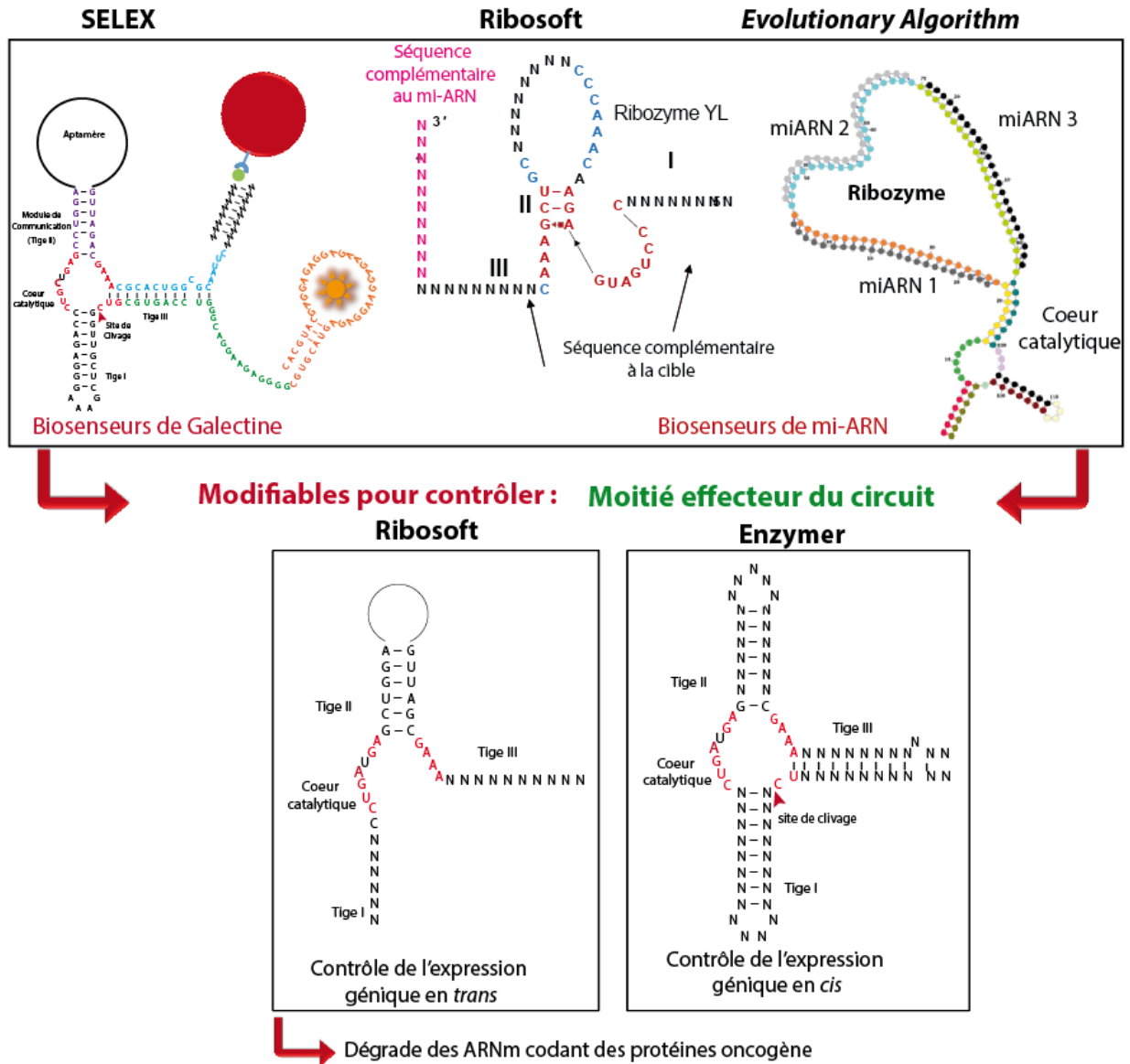


Figure 8.10 Composantes du circuit final

Au final, l'activité de toutes les composantes du circuit logique a été démontrée, mais seulement *in vitro*. L'activité des ribozymes dans les cellules n'a pas pu être prouvée à ce stade, ni le lien entre les ribozymes senseurs et ribozymes effecteurs.

9. Discussion générale

Les ribozymes ont été l'objet de plusieurs études depuis leur découverte. Ceci est dû aux multiples fonctions catalytiques qu'ils possèdent ainsi qu'à leur rôle majeur comme modèles dans l'étude des structures secondaires et tertiaires de l'ARN. Bien que la découverte du premier ARN catalytique ait eu lieu il y a presque 40 ans, le dévoilement de nouveaux ribozymes par de nouvelles méthodes se poursuit jusqu'à nos jours (Chen & Luptak, 2021; Olzog *et al.*, 2021). Les ribozymes présentent différentes fonctions comme la ligation et l'épissage, mais la fonction la plus étudiée et appliquée est celle du clivage d'ARN (Weinberg *et al.*, 2019b). Leur activité dépend de certaines conditions. Le paramètre majeur étant celui de conservation de la séquence du cœur catalytique ; le reste de la séquence peut être modifié à condition que la structure formée soit conforme à celle du ribozyme sauvage. Ceci a permis d'intégrer les ribozymes dans plusieurs applications. Cette intégration a été possible grâce à la conception et à la sélection de nouveaux ribozymes actifs en *cis* et en *trans*. Même si, en théorie, il semble facile de modifier la séquence d'un ARN afin qu'il réponde aux critères de son utilisation, lors de l'application, ce n'est pas toujours le cas.

Ceci est causé par le fait que, pour une molécule d'ARN avec une séquence donnée, il n'y a pas qu'une seule possibilité de repliement, mais il y en a plusieurs. Ce repliement dépend de la séquence et est sujet au changement ou à la stabilisation par d'autres facteurs. Comme exemple de facteurs, la présence des cations mono et divalents réduisant la répulsion électrostatique entre les nucléotides, des protéines chaperonnes et de l'encombrement moléculaire, permettent une stabilisation de la structure. Un changement de ces facteurs peut produire des mésappariements menant à un mauvais repliement. Par conséquent, l'activité de l'ARN mal replié peut être réduite ou altérée, puisque la bonne conformation d'un ARN est essentielle à sa fonction. De ce fait, lors de la conception de nouvelles séquences, il faut s'assurer de prendre en considération cette condition de respecter la structure secondaire. En conséquence, plusieurs logiciels de prédiction de structures ainsi que de conception de nouvelles séquences d'ARN fonctionnels permettent de réduire les possibilités de mauvais repliements.

9.1. La conception des ARN synthétiques incluant des pseudonœuds

L'un des motifs structuraux complexes et difficiles à inclure dans les programmes bioinformatiques de prédiction et de conception est le pseudonœud. Nous rappelons que

ce dernier se définit comme l'interconnexion d'une région simple brin d'une tige boucle avec une autre région simple brin, à l'extérieur de cette tige sur la même molécule d'ARN, par l'appariement de base (Brierley *et al.*, 2007; Pleij *et al.*, 1985; Staple & Butcher, 2005). L'importance des pseudonœuds vient de leur rôle dans la stabilisation des structures actives d'une panoplie d'ARN fonctionnels comme les ribozymes, les *riboswitches* et des régions des ARN viraux (Staple & Butcher, 2005).

Dans notre projet, nous avons pu valider l'efficacité d'un premier logiciel de conception de nouvelles séquences d'ARN à partir de structures connues impliquant des pseudonœuds. « Enzymer », un algorithme publié en 2016 (Zandi *et al.*, 2016), utilise l'approche du repliement inverse pour engendrer des séquences d'ARN actifs. Nous avons été en mesure de confirmer l'activité des ribozymes en tête de marteau ainsi que la formation du pseudonœud par l'un des ribozymes prédit. Ceci a été prouvé par l'absence de l'activité de clivage lorsque les nucléotides responsables de la formation du pseudonœud ont été mutés. De plus, la conception non optimisée de nouvelles séquences par Enzymer a également généré des ARN avec une faible activité. Un exemple ayant une structure plus complexe, le ribozyme GlmS comportant trois pseudonœuds, a été utilisé comme matrice pour la génération de nouvelles séquences. Ceci a été la première conception par **repliement inverse** d'un ribozyme similaire au GlmS ayant montré une activité. Les études qui ont été produites précédemment sur le GlmS impliquaient des bibliothèques mutées et des sélections *in vitro*. (Lau & Ferre-D'Amare, 2013; Lau & Ferre-D'Amare, 2016; Lau *et al.*, 2017; Link *et al.*, 2006) Elles n'ont pas permis de sélectionner des versions aussi divergentes de la séquence originale (du moins, pas validées individuellement) que celles que nous avons testées. Cet exemple montre très bien l'importance des motifs de structures secondaires, car les pseudonœuds sont impliqués dans la formation du cœur catalytique du ribozyme WT (Peselis & Serganov, 2014; Soukup, 2006). Dans ce cas, un échec de conception des pseudonœuds pourrait engendrer un échec total de l'activité des nouveaux ribozymes.

L'intérêt montré aux pseudonœuds comme motifs de structure tertiaire dans le cas de ribozymes auto-clivants est dû au fait qu'ils peuvent être utilisés comme troisième site de liaison. Les pseudonœuds sont formés par des paires de bases canoniques, ce qui facilite leurs prédictions ainsi que leurs conceptions par rapport à d'autres motifs de structures tertiaires formés par des paires de bases non canoniques. Ceci leur donne l'avantage d'être plus faciles à inclure dans plusieurs logiciels de prédiction et de conception. Dans

le cas de ribozymes en tête de marteau ayant un pseudonœud entre la tige I et II, comme celui de *Yarrowia lipolytica*, pour la version *trans* du ribozyme, ils ont l'avantage de posséder un site de liaison de plus à la cible. Cela résulte par trois sites de liaison à la cible offrant plus de spécificité, limitant l'activité en dehors de la cible. Ainsi, ce site jouera également le rôle d'un verrou de sécurité, car l'absence de la cible complémentaire résultera d'une activité réduite, voire absente du ribozyme. De ce fait, « Enzymer » a l'avantage d'inclure ce motif pour ce type d'application, et plusieurs autres.

Le développeur de « Enzymer » a déjà montré l'efficacité, *in-silico*, de son programme, en comparant des scores obtenus pour les paramètres du design avec ceux du WT. Ces résultats étaient similaires ou meilleurs à ceux d'autres algorithmes ayant le même but. Néanmoins, pour les biologistes, et même pour les bioinformaticiens, des résultats expérimentaux sont requis pour assurer l'applicabilité des algorithmes. Dans le cadre de notre projet, l'obtention de ribozymes actifs a été un succès démontré par des données expérimentales, ce qui nous confirme la possibilité de compter sur Enzymer pour des applications futures.

Malgré cela, Enzymer présente aussi des limites partagées avec la majorité de logiciels disponibles actuellement, soit l'incapacité de prendre en considération des motifs de structures tertiaires formés par des paires de bases non-canoniques. De plus, il sera utile qu'un tel logiciel évalue l'inclusion des nucléotides avec des groupements modifiés. Ceci serait particulièrement utile dans deux cas, le premier étant la nécessité de modification des ARN afin d'augmenter leur résistance à la dégradation par le remplacement du groupement 2'OH du sucre par un 2'F, 2' O méthyl ou un 2' amine. Ce genre de modification peut affecter la structure de l'ARN (Harcourt *et al.*, 2017; Tanzer *et al.*, 2019), alors un logiciel qui permet de prendre en considération ses modulations aura beaucoup d'intérêt dans le domaine. En deuxième lieu, les nucléotides modifiés existent naturellement dans certains ARN, comme les ARNt, les ARNr, certains lncARN, certains miARN, etc (McCown *et al.*, 2020). Lors de la conception de nouvelles séquences il est important de conserver ces modifications car leur absence risque d'affecter la structure ainsi que la fonction de ces ARN, bien que des développements récents permettent de tenir compte des méthylation en N6 des adénines (Szabat *et al.*, 2022), les logiciels de repliements de structures ne peuvent pour l'instant pas prendre en considération la majorité des bases modifiées.

En fait, plusieurs logiciels et algorithmes de conception de nouvelles séquences d'ARN avec l'approche de repliement inverse ont été publiés, avec une grande majorité ne prenant pas en considération les pseudonœuds (Hofacker, 2003a; Honer zu Siederdisen *et al.*, 2013; Levin *et al.*, 2012; Lyngso *et al.*, 2012) (Dotu *et al.*, 2014; Drory Retwitzer *et al.*, 2016; Retwitzer *et al.*, 2020). Parmi les logiciels en mesure de concevoir les pseudonœuds, peu ont validé l'activité des ARN conçus en laboratoire (antaRNA (Kleinkauf *et al.*, 2015), MODENA (Taneda, 2015), nanofolder (Bindewald *et al.*, 2011), ERD (Esmaili-Taheri & Ganjtabesh, 2015) (Dotu *et al.*, 2014) et (Yamagami *et al.*, 2019). Seule l'efficacité des ARN conçus par les deux derniers logiciels a été validée avec des expériences de laboratoire publiées dans le même article (Yamagami *et al.*, 2019). L'importance de cette partie expérimentale vient du fait qu'elle nous permet non seulement de valider les designs des ARN, mais aussi d'identifier les problèmes et les défauts des algorithmes à corriger pour obtenir des programmes réellement applicables et plus performants. Ceci ouvre la porte à des applications futures utilisant des ARN plus complexes comportant des pseudonœuds plus complexes et variés. En fait, il existe plus de 200 pseudonœuds naturels ayant des topologies très variées, classifiées en six types, résumés dans la base de données *pseudobase* (<https://rnavlab.utep.edu/database>). Un exemple ayant une structure et une fonction complexe est l'intron du groupe I contrôlé par le *riboswitch* cyclic-di-GMP (Lee *et al.*, 2010). Cette combinaison unique et rare possède une structure comportant deux pseudonœuds et une fonction très intéressante qui permet un contrôle étroit de l'expression génique. L'obtention des logiciels de conception assez rigoureux et efficaces permettra de produire des séquences se repliant en de telles structures, et même de créer des structures avec de nouvelles fonctionnalités. Un ARN rassemblant deux ou plusieurs aptamères pour des ligands différents et une plateforme d'expression ou un ribozyme représentera une combinaison parfaite d'un circuit logique appliqué pour des maladies complexes comme le cancer.

9.2. Conception d'un système d'inhibition de gènes

Les systèmes d'inhibition de gènes actuels sont multiples avec plusieurs logiciels gratuits permettant leurs conceptions et prédictions. Ces systèmes ont de multiples paramètres à considérer en fonction de leur type et de l'application. CRISPR est le système qui a le plus de succès depuis plusieurs années et il continue à surprendre les chercheurs avec la vaste panoplie d'applications dans laquelle il est capable de s'intégrer. Il s'agit d'un système efficace avec un taux de succès assez élevé dans des tests *in vivo*, mais son

défaut est la complexité du suivi de son destin intracellulaire incluant son activité et la durée de vie de l'ARN guide (Shivram *et al.*, 2021). Ceci vient de deux faits : la façon de contrôler l'expression des ARN guides (ARNg) ainsi que l'amélioration de leur résistance à la dégradation une fois qu'ils sont exprimés dans la cellule. Ceci se produit par l'insertion d'autres ARN en aval ou en amont, dans le but d'avoir un *switch ON/OFF* afin de limiter le fonctionnement du système CRISPR à un moment, ou dans un environnement désiré (Collantes *et al.*, 2021; de Solis *et al.*, 2016; Wang *et al.*, 2021c). Actuellement, il existe plusieurs logiciels disponibles (résumé dans la **Table 9.1**) pour la conception des ARNg, mais aucun ne prend en considération ces ajouts de séquences. Dans d'autres systèmes, comme pour les miARN, les petits ARN interférents et les ribozymes, un autre paramètre s'ajoute, soit l'accessibilité de la région d'ARN cible (Kiryu *et al.*, 2011; Sagan *et al.*, 2010; Tafer *et al.*, 2008). C'est évidemment dans cette optique que le service web Ribosoft 2.0 de conception des ARNnc, a été développé. Ribosoft 2.0 est apte à concevoir des ribozymes conventionnels et allostériques actifs en *trans* ; l'ajout de l'option de conception des ARN guide du système CRISPR est planifié. Ce service web a la capacité de générer les ribozymes en prenant en considération des paramètres liés à celui-ci ainsi qu'à la cible. En fait, depuis la première version du service web (Kharma *et al.*, 2016), nous avons montré l'effet important de la structure de la cible sur l'activité du ribozyme avec un paramètre nommé « accessibilité ». Celui-ci indique la position du site du clivage et d'hybridation du ribozyme sur l'ARN à dégrader, qui est replié dans sa structure finale. La nouvelle version du logiciel a conservé les mêmes paramètres, en rajoutant la possibilité de concevoir beaucoup plus qu'un seul type de ribozyme.

Nous avons utilisé cette nouvelle version pour générer des ribozymes conventionnels de deux types (en tête de marteau et pistol) et des ribozymes allostériques actifs en *trans* contrôlés par la FMN. Une majorité des ribozymes conventionnels ont été actifs montrant l'efficacité de la conception de ce type de ribozymes par Ribosoft2.0. En ce qui concerne les aptazymes, seulement ceux inhibés par la FMN ont montré une activité spécifique en présence du ligand. Ceci fait de Ribosoft2.0 un logiciel unique qui, permet non seulement d'utiliser pratiquement n'importe quel ribozyme connu comme outil, mais qui peut également fournir de multiples informations lors de la génération des résultats. Ceci rend le choix des séquences par l'utilisateur beaucoup plus facile. Avec nos résultats présentés dans le premier chapitre, nous avons pu montrer l'efficacité des ARN conçus par ce logiciel. De plus, l'obtention de ribozymes allostériques actifs en *trans* avec une génération automatisée est une première dans le domaine. Il est important de mentionner ici aussi

que les résultats expérimentaux ont permis d'optimiser le logiciel à plusieurs reprises montrant ainsi, comme mentionné ci-plus haut, l'importance de ces données pour de meilleurs résultats futurs. Ceci a été fait par l'évaluation de l'activité de ribozymes de différents types à plusieurs reprises au fur et à mesure que le logiciel est entrain d'être développé et optimisé (résultats non pertinents dans le cadre du projet donc non montrés). De plus, grâce à nos résultats, présentés dans le sixième chapitre, les développeurs étaient capables de découvrir des défauts dans le système de classification donc de filtration de ribozymes utilisé par Ribosoft2.0 qui ne donnait presque que des ribozymes classifiés '1' ce qui indique qu'ils sont supposés être les meilleurs.

En conclusion, Ribosoft2.0 est un logiciel puissant ayant la capacité d'inclure la structure de n'importe quel ribozyme, tant que sa structure est connue. La flexibilité, tout comme la fiabilité de Ribosoft2.0, nous ont permis de concevoir des ribozymes liant des miARN d'une façon simple et rapide. Bien que les résultats obtenus fussent seulement *in vitro*, ceci représente déjà un avancement. En effet, quoique l'ajout de séquences additionnelles de 20 bases soit très simple en principe, le faire de façon rapide en s'assurant que cet ajout n'interfère pas avec l'activité du ribozyme constitue un exemple parmi d'autres de l'utilité de Ribosoft2.0.

En parallèle au ribosoft2.0, un algorithme d'apprentissage automatique a été développé par le même laboratoire dans le but de concevoir des ribozymes actifs en *trans*, activés par la liaison d'une séquence spécifique. L'exemple publié est l'activation d'un ribozyme par la liaison d'une séquence ayant un nombre défini de répétitions d'un codon. Cette liaison induit la restructuration du ribozyme et permet la liaison et la dégradation du transcrit cible, soit l'ARNm muté de PABPN1, causant la maladie « Dystrophie musculaire oculopharyngée » (Nicolas Kamel, 2020). Ceci est une autre preuve que la conception automatisée ouvre la porte à une large gamme d'applications. De plus, elle offre une option plus sûre et sécuritaire quant à l'usage des ribozymes pour le contrôle de l'expression génique.

Malgré tout ces avantages de ribozymes présentés ci-dessus et l'utilité de Ribosoft2.0 comme outil à utiliser pour obtenir des ribozymes applicables dans des thérapies, le progrès reste limité actuellement. Bien que l'optimisation des séquences et structures des ribozymes a été faite par Ribosoft2.0, on n'était pas capable de montrer leurs activités dans les cellules. Ceci implique qu'on n'a pas encore atteint le niveau d'amélioration désiré qui permettra de compter les ribozymes comme outil de thérapie. Il y a actuellement

un manque d'information sur certains paramètres qui affectent l'activité des ribozymes *in vivo*. Comme exemple, la vraie structure des ARN cibles ainsi que leur interaction dans les cellules avec d'autres ARN et protéines sont des paramètres mal compris et partiellement ignorés par les logiciels de conceptions. La cassette d'expression de ribozyme utilisée implique l'ajout de grands bouts de séquences de part et d'autres des tiges du ribozyme, ce qui peut facilement affecter sa structure et évidemment son activité. De plus, le destin des ribozymes dans les cellules après leur expression n'est pas caractérisé.

En conclusion, les ribozymes sont une alternative très intéressante aux outils utilisés actuellement qui dépendent des protéines et qui touchent à l'ADN, mais leur arrivée au même stade que ces outils demandera encore des efforts pour mieux comprendre et définir les paramètres essentiels à optimiser.

9.3. Importance de la validation expérimentale

Dans un cas de prédiction ou de conception, l'élément principal sur lequel le résultat s'appuie est la similarité avec une structure donnée et une concordance avec des paramètres précisés dès le début. Bien que ces données soient déjà prises d'un exemple naturel, les résultats générés demeurent une simulation de la réalité qu'il faut valider. Le manque de validation empêche d'avoir une optimisation et de découvrir les anomalies d'un algorithme.

Notre choix d'un certain système ou algorithme vient du fait qu'il est fiable. Notre expérience avec des ribozymes initialement prévus pour la thérapie génique nous amène à penser que l'une des raisons derrière l'échec d'études cliniques impliquant des ribozymes ou d'autres ARN non codants est le fait que l'on n'a pas encore saisi parfaitement les paramètres à respecter pour avoir de meilleurs résultats. Parmi ces paramètres on trouve, l'interaction de l'ARN effecteur avec sa cible dans le milieu intracellulaire, la structure formée par cet ARN ainsi que sa cible, les interactions non spécifiques qui peuvent avoir lieu, la durée de la demi-vie de ces ARN, etc. Ceci nous ramène encore au point de départ, soit au manque de données expérimentales, alors à la nécessité d'améliorations notables dans l'utilisation de ces données pour les transformer en modèles prédictifs.

En fait, davantage de données expérimentales aideraient à avoir de meilleures prédictions dans le futur. N'importe quel système de prédiction, ou une approche de « *machine*

learning », par exemple, a besoin d'être implémenté avec des données réelles, ce qui signifie, dans notre cas, des résultats expérimentaux. C'est tout à fait similaire à la façon dont le cerveau humain fonctionne, c'est-à-dire que; notre compréhension est basée sur l'expérience, peu importe s'il s'agit d'une expérience vécue par soi-même ou par d'autres. De plus, la créativité en science dépend des connaissances scientifiques pour être en mesure de concevoir de nouvelles idées et d'avoir la capacité et le sens critique permettant d'évaluer ces idées. Tout cela est applicable sur les logiciels de conception et de prédiction, car la biologie synthétique est basée sur ce lien entre la biologie avec ces données expérimentales et les nouvelles technologies. Ces dernières permettent de résoudre des problèmes ou d'ouvrir de nouvelles portes en lien avec notre domaine.

9.4. Les aptamères : un exemple concret du manque de données expérimentales

Les aptamères constituent un autre exemple qui montre la grande importance des données. Une recherche sur Pubmed révèle plus de 16 000 publications concernant ceux-ci. Par contre, si l'on se penche sur les exemples concrets des aptamères utilisés dans des applications réelles comme thérapies ou biosenseurs, on n'en trouve que très peu. Le site ClinicalTrials.gov montre que seulement 50 aptamères sont utilisés dans des études cliniques. Il y a aussi quelques exemples d'aptamères utilisés comme biosenseurs, ainsi que des compagnies offrant des plateformes de sélection et d'utilisation d'aptamères comme SOMAlogic et Neoventures (Kaur *et al.*, 2018). Ceci vient du fait que les informations publiées concernent plutôt la méthode de sélection, l'affinité et l'activité. Les données sur les structures sont manquantes. La grande majorité des aptamères a été seulement publiée avec des prédictions de structures, et non avec des structures cristallines ni avec d'autres essais pour confirmer leur structure secondaire. Ceci laisse un grand point d'interrogation sur la vraie conformation de la molécule ainsi que son interaction avec son ligand et plusieurs autres paramètres comme la température de dénaturation « Tm » et l'énergie libre. Des essais sont exécutés pour développer des programmes informatiques permettant de prédire les types d'interactions entre une certaine molécule et des séquences aléatoires d'ARN ou ADN. Ce sont des programmes qui visent à faire des sélections *in silico* afin d'alléger l'étape de SELEX qui est la partie la plus délicate, mais aussi la plus importante dans le processus de trouver un nouvel aptamère, mais le progrès reste limité (Ahirwar *et al.*, 2016; Ishida *et al.*, 2020; Kilgour *et al.*, 2021; Sanchez-Luque *et al.*, 2014; Sepehri Zarandi *et al.*, 2020; Yokoyama *et al.*,

2017). La sensibilité du processus de SELEX vient des artéfacts de PCR qui peuvent affecter la qualité de la librairie de sélection à cause des longs cycles d'amplifications répétés. De plus, une petite contamination du ligand spécifique avec un autre non spécifique diminuera la spécificité de l'aptamère finale. Plusieurs autres facteurs entrent en jeu, comme la dépendance de l'aptamère finale des conditions de sélection, les longs cycles de SELEX dans le cas des librairies d'ARN qui rend la tâche assez ardue, etc.

En conclusion, les modèles de prédictions de structures existantes ont été développés avec l'information disponible. Plus nous aurons un grand nombre de données, meilleure sera la prédiction et, à un certain moment, nous serons en mesure d'avoir un système de prédiction fiable au point de ne plus avoir besoin de faire les tests expérimentaux (ou du moins, pas autant). Le meilleur exemple est le logiciel « AlphaFold2 », développé par *DeepMind* de *Google*, qui a permis, grâce à des millions de données rassemblées au fil des années, de prédire les structures de protéines à partir de leurs séquences d'acides aminés (Callaway, 2021). Le programme le plus proche de celui-là, dans le cas de l'ARN, est « ARES » (*Atomic Rotationally Equivariant Scorer*) (Townshend *et al.*, 2021b), qui calcule des notations sur des prédictions de structures tertiaires des ARN ainsi que des modèles structurels, mais qui a besoin de beaucoup plus de données expérimentales afin d'atteindre le même niveau de précision.

9.5. Le choix des ribozymes

La question qui se pose en présence des échecs des ribozymes dans des applications réelles, est de savoir pourquoi choisir les ribozymes, et non d'autres types d'ARN dont l'usage est plus répandu, comme les siRNA et CRISPR?

Les ribozymes présentent une grande flexibilité, ce que nous avons été capables de prouver avec les divers ribozymes utilisés dans cette étude : les ribozymes conventionnels, ribozymes contrôlés par un ou plusieurs oligonucléotides qui sont, dans notre cas, des miARN, et finalement, les aptazymes liant des protéines.

De plus, ces ribozymes sont des ARN qu'on a pu tester *in vitro*, chez les bactéries et les cellules humaines, ce qui montre l'universalité de leur application. Un autre avantage est leur indépendance de cofacteurs ou de protéines chaperonnes, ce qui rend le test d'un grand nombre de séquences dans les essais en dehors de cellules beaucoup moins cher que les autres types d'ARN, générant le même résultat de « clivage », les rendant compatibles avec une approche d'évolution *in vitro* comme le SELEX. De ce fait,

l'automatisation des tests de clivage ou de sélection en développant des essais haut débit est beaucoup plus facile (Townshend *et al.*, 2021a). Ceci implique qu'avec les logiciels de conceptions capables de générer des centaines de séquences à tester, en combinant cela aux résultats générés par des essais de clivage haut débit, nous serons capables d'optimiser les conceptions et ainsi d'avoir des ARN plus performants.

En contrepartie, les petits ARN interférants, les miARN ainsi que CRISPR ont besoin de tout un système formé par des complexes protéiques qui coûtent assez cher pour des essais *in vitro*. De plus, des systèmes d'expression acellulaire sont utilisés afin de produire les protéines fonctionnelles du CRISPR par exemple, ce qui nécessite un processus assez demandant en matériel et en temps (Jiang *et al.*, 2021a; Lehr *et al.*, 2021). D'autres systèmes comme TALENS (*Transcription Activator-Like Effector Nucleases*) (Bhardwaj & Nain, 2021) et ZFN (*Zinc Finger Nucleases*) (Chandrasegaran, 2017) ainsi que les oligonucléotides antisens (ASO) (Bennett, 2019) ne sont testables que dans les cellules comme première étape. D'un autre côté, exprimer les ribozymes dans les cellules peut être fait à l'aide de vecteurs, mais aussi avec un ARN modifié, augmentant la demi-vie de l'ARN. Ceci n'est pas une option pour les approches basées sur CRISPR, sauf si les cellules expriment déjà la protéine, soit la partie active du système.

Malgré le fait qu'en théorie, les ribozymes s'avèrent être l'outil parfait d'inhibition génique, ils semblent avoir perdu la course en comparaison avec les systèmes mentionnés ci-dessus. Néanmoins, cette comparaison est injuste ; un ribozyme est une entité fonctionnelle par elle-même en tant qu'ARN, mais pour tous les autres complexes, l'ARN ou l'ADN n'est que le guide de livraison d'une enzyme. Ceci peut représenter une raison majeure de l'échec de l'activité d'un ARN, car sa demi-vie très courte, soit de quelques minutes (Chen *et al.*, 2008), ne peut être comparable à des nucléases ayant des demi-vies de plusieurs heures (Chen *et al.*, 2016). De plus, les ribozymes passés aux études cliniques étaient des ribozymes minimaux avec des structures beaucoup moins caractérisées et moins bien comprises qu'aujourd'hui. Tout cela avait coïncidé avec les grandes découvertes des siRNA et CRISPR qui semblaient être des outils plus puissants. Actuellement, aucun essai clinique impliquant des ribozymes en tant qu'effecteur n'a été fait en presque 20 ans, alors que les autres outils ont eu et continuent d'avoir beaucoup de succès ; des centaines d'études se réalisent encore là-dessus (résumé dans la **Table 9.1**). Cette réalité ne nous a pas empêchés de travailler sur les ribozymes. Au contraire,

avec nos travaux, nous essayons de trouver les meilleures façons pour exploiter la puissance des ribozymes en apprenant des études réalisées dans le passé.

Table 9.1 Résumé des systèmes utilisés en thérapies et leurs avancés

Système	Avantages	Inconvénients	Outils de prédiction et conception	Avancées
Ribozymes clivant des cibles Découverte en 1983	Touche l'ARN sans aucun effet sur le génome et les cellules germinales Faible immunogénicité et toxicité cellulaire Activité indépendante de toute assistance. Expression facilement contrôlable	Faible efficacité dans les cellules	Quelques outils de conception de ribozymes et de prédiction de cibles (Ribosubstrate (Lucier <i>et al.</i> , 2006b), Ribosoft (Kharna <i>et al.</i> , 2016), Enzymer (Zandi <i>et al.</i> , 2016), (Yamagami <i>et al.</i> , 2019)	Plusieurs ribozymes ont atteint des essais cliniques mais aucun parmi eux n'a été approuvé pour un traitement. Les recherches académiques utilisant des ribozymes pour des thérapies géniques (Han <i>et al.</i> , 2021b) ou anti-viral continuent
CRISPR Découverte en 1993 (Groenen <i>et al.</i> , 1993)	Réparation ou élimination complète du gène responsable de la maladie avec une efficacité et spécificité élevée. Seulement la séquence d'ARN guide doit être modifiée pour chaque cible différente (Li <i>et al.</i> , 2020) Facilité de conception d'ARNg.	Risque de fautes irréparables dans le cas de off-target car le système agit sur l'ADN dans la majorité des cas. Risque élevé de dégradation hors cible. Obligation d'avoir le site PAM (<i>Protospacer Adjacent Motif</i>) dans la cible. Tolérance assez élevée des mésappariements (Bhardwaj & Nain, 2021)	Plusieurs laboratoires et compagnies (IDT DNA, horizon discovery, synthego, benchling, Thermofisher) ont développé des programmes de conception d'ARN guide, la prédiction de l'efficacité du système ainsi que les cibles non spécifiques (Cas OFFinder (Bae <i>et al.</i> , 2014), CRISPRscan (Moreno-Mateos <i>et al.</i> , 2015), CRISPR multitargeter (Prykhodzij <i>et al.</i> , 2015), etc)	Aucun traitement n'est encore approuvé mais plusieurs études sont à l'étape des essais cliniques (Karamperis <i>et al.</i> , 2021; Maeder <i>et al.</i> , 2019) (ClinicalTrials.gov: NCT03545815, NCT04037566, NCT04244656, NCT 04178382, NCT03855631, NCT04122742, NCT03655678, NCT03745287)
Petits ARN interférents Découvert en 1998 (Fire <i>et al.</i> , 1998)	Machinerie de reconnaissance est déjà présente dans les cellules animales. Efficacité élevée (Hu <i>et al.</i> , 2020)	Risque d'activité hors cible. Risque d'immunogénicité élevée pour les siARN non modifié (Hu <i>et al.</i> , 2020)	De multiples compagnies ont leurs propres outils de conception de siRNA et de prédiction de cible (eurofinsgenomics, Thermofisher, genscript, invivogen, benchling, IDTdna) et il y a d'autres logiciels disponibles (si-Fi (Luck <i>et al.</i> , 2019), DSIR (Vert <i>et al.</i> , 2006)	Trois traitements approuvés par la FDA et commercialisés : 1-Lumasiran pour Hyperoxalurie primaire de type 1 (Garrelfs <i>et al.</i> , 2021) 2-Givosiran contre la porphyrie hépatique aiguë (Honor <i>et al.</i> , 2021)

				3- Patisiran amylose héréditaire à transthyréine (Milani <i>et al.</i> , 2019)
Oligonucléotide anti-sens (ASO)	Différents mécanismes d'action. Efficacité élevée. Dépend des systèmes naturels existant chez les cellules cibles. (Bennett <i>et al.</i> , 2019)	Risque élevé hors cible. Besoin de nucléotides modifiés pour limiter ses effets négatifs comme la toxicité, l'immunogénicité et l'activité hors cible (Di Fusco <i>et al.</i> , 2019)	Des compagnies comme genelink, benchling et IDT fournissent des outils de conception, mais il y a peu d'outils disponibles en ligne qui sont publiés par des laboratoires de recherche (PFRED (pour les siRNA et les ASO (Sciabola <i>et al.</i> , 2021)	Trois traitements approuvés par la FDA : 1-Traitement de l'atrophie musculaire spinale (Neil & Bisaccia, 2019) 2-Traitement de l'atrophie musculaire de duchenne (Anonyme, 2018) 3-Traitement contre l'amylose héréditaire à transthyréine (Mathew & Wang, 2019)
<i>Zinc-fingers nucleases</i> Découvert en 1985 (Miller <i>et al.</i> , 1985)	Spécificité élevée. Petite taille des éléments de système par rapport aux systèmes TALENS et CRISPR (Li <i>et al.</i> , 2020)	Besoin de plus d'ingénierie de protéines (Li <i>et al.</i> , 2020)	<i>Zinc finger tools</i> est un service web pour la recherche de cible et la conception de protéine zinc-finger (Mandell & Barbas, 2006). <i>ZFNgenome</i> est un outil de recherche de cible dans des organismes modèles (Reyon <i>et al.</i> , 2011) et il fait partie des outils offerts par le <i>zinc finger consortium</i> (http://www.zincfingers.org/software-tools.htm)	Plusieurs études sont en phases cliniques actuellement, mais aucun traitement n'est encore approuvé (NCT04212858, NCT03666871, NCT03617198)
TALENS La découverte des composantes du système et sa compréhension a pris 10 ans (Boch & Bonas, 2010)	Haute spécificité et efficacité (Hu <i>et al.</i> , 2015)	Construction de vecteurs portant le système est assez complexe et long (Zhang <i>et al.</i> , 2019a)	Thermofisher fournit l'outil ' <i>Invitrogen TrueDesign Genome Editor</i> ' qui est utilisable pour TALEN ou CRISPR. <i>TAL Effector Nucleotide Targeter 2.0</i> est un outil de cornell university qui offre différentes fonctions par rapport aux cibles du système TALEN (Doyle <i>et al.</i> , 2012). E-TALEN est un service de conception (Heigwer <i>et al.</i> , 2013)	Des études cliniques sont en cours par exemple ciblant le papilloma virus (Hu <i>et al.</i> , 2015) et d'autres pour des CAR-T (ClinicalTrials.gov: NCT04142619, NCT03190278, and NCT04150497). Utilisé par la compagnie Cellectis pour la production des UCAR-T (universal chimeric antigen

				receptor T) et approuvé par la FDA
miARN Découvert en 1993 (Lee <i>et al.</i> , 1993)	Autre que la dégradation d'ARN cible par RISC, les miARN synthétiques agissent comme antagonistes de miARN naturels surexprimés. Un seul miARN peut avoir de multiples cibles (Ishida & Selaru, 2013)	Touche plusieurs cibles : risque d'activité hors cible très élevée (Ishida & Selaru, 2013) Stimulation non spécifique du système immunitaire (Ahmadzada <i>et al.</i> , 2018)	Moins de compagnies offrent des outils de conception de miARN synthétiques (Horizon, Thermofisher), et peu d'autres sont publiés (WMD3 (Schwab <i>et al.</i> , 2006), (Ossowski <i>et al.</i> , 2008)	Des inhibiteurs de miARN et des miARN mimics sont à l'étape des essais cliniques (Anastasiadou <i>et al.</i> , 2021; Gallant-Behm <i>et al.</i> , 2019). Actuellement aucun traitement n'est encore approuvé ou commercialisé

9.6. Ribozymes allostériques

9.6.1. Portes logiques

Dans notre projet, nous avons présenté différentes façons d'avoir des ribozymes allostériques contrôlés par une ou plusieurs entrées comme les miARN. D'une part, des ribozymes ont été conçus *in silico* pour lier un à trois miARN. Ces ribozymes n'ont pas répondu parfaitement comme prévu, au lieu d'être actifs en présence de trois miARN, le clivage a eu lieu lorsqu'un seul ou deux miARN ont été ajouté à la réaction. Le concept d'oligonucléotides comme « entrée » (*input*) contrôlant l'activité des ribozymes, par la modulation de leurs structures, a déjà été abordée par Robert Penchovsky. Il était apte à générer des designs effectuant plusieurs fonctions logiques (YES, NOT, AND, OR) (Penchovsky, 2012; Penchovsky, 2013; Penchovsky, 2014; Penchovsky & Breaker, 2005; Penchovsky & Kostova, 2013). Ces designs sont basés sur le ribozyme en tête de marteau (versions minimale et complète) et un grand nombre d'entre eux sont actifs *in vitro*. Ceci a constitué le point de départ des conceptions faites par notre collaborateur à l'Université Concordia.

La nouveauté que nous montrons alors n'est pas celle des ribozymes allostériques, mais plutôt l'utilisation d'une porte logique liant des miARN impliqués dans le cancer du sein triple négatif. Bien que les portes logiques testées n'aient pas répondu parfaitement comme souhaité, une partie de la preuve de concept est exécutée. Une optimisation des ribozymes est nécessaire, par exemple en utilisant des ribozymes complets qui auront une meilleure activité dans les cellules. Malgré les manques de ces designs, nos résultats

montrent que la conception de ribozymes constituant des portes logiques complexes est envisageable dans un futur rapproché. Une étape importante serait ainsi franchie pour l'utilisation des miARN comme biomarqueurs d'identification des cellules cancéreuses. Ces ribozymes ayant la capacité de reconnaître de multiples miARN en parallèle peuvent être employés comme biosenseurs par le simple ajout d'un fluorophore et d'un *quencher*. En effet, une étude publiée récemment est venue confirmer notre hypothèse sur l'utilisation des miARN comme signaux d'activation d'une porte logique appliquée à une thérapie génique (Yue *et al.*, 2020). Bien que ce concept diffère du nôtre, deux miARN ont été utilisés ensemble afin de causer un déplacement de brin, permettant ainsi la libération de la doxorubicin et d'un siRNA. Une autre étude a confirmé notre hypothèse par l'emploi de la liaison de trois miARN sur une nanostructure d'ADN afin de distinguer des lignées cellulaires et différents types de cancers. La porte logique « AND », sous forme d'un hexahedron formé par l'assemblage des brins d'ADN, émet trois signaux fluorescents différents lors de l'attachement de trois miARN (Yang *et al.*, 2021). Des circuits logiques plus complexes prenant en considération un plus grand nombre de miARN ont également été construits en jouant sur les taux d'expression génique en ayant recours à la fonction naturelle des miARN qui lient des régions sur l'ARNm (Xie *et al.*, 2011) ou en utilisant des nanoparticules pour détecter des miARN intracellulaires (Yue *et al.*, 2019). Une porte logique basée sur un DNAzyme, qui est similaire dans son fonctionnement aux ribozymes, a été développée pour être activée en liant simultanément deux miARN, ce qui génère une amplification du signal (Wu *et al.*, 2019).

En fait, ces études concordent avec la vague d'usage du concept de circuits logiques qui deviennent de plus en plus impliqués dans le cadre de diagnostic et de thérapie du cancer (Boscaino *et al.*, 2019; Cai *et al.*, 2021; Jiang *et al.*, 2021b; Li *et al.*, 2021c; Liu *et al.*, 2021; Liu *et al.*, 2020; Miao & Tang, 2021; Wang *et al.*, 2021a; Wang *et al.*, 2021b; Zhang *et al.*, 2020a; Zhang *et al.*, 2021a; Zhao *et al.*, 2021). En dehors du contexte des circuits construits par des acides nucléiques, des portes logiques AND sont appliquées en immunothérapie par la combinaison obligatoire de deux antigènes afin d'activer les cellules T. Ceci engendre une spécificité renforcée et limite l'activité hors cible ; il s'agit d'un principe qui a été appliqué dans des études précliniques (Bonfa *et al.*, 2020; Feldmann *et al.*, 2020; Savanur *et al.*, 2021).

9.6.2. Aptazymes en biosenseurs

Les protéines, en étant les biomarqueurs considérés idéaux pour une détection précoce ou une réapparition d'un cancer, n'ont jamais perdu leur importance comme cible de nouvelles méthodes de détection et diagnostic (Chatterjee & Zetter, 2005; Koo *et al.*, 2019; Pavlou *et al.*, 2013). Bien qu'il y ait un progrès remarquable dans les outils de détection de ces protéines, il reste toujours une place à l'amélioration. En effet, la majorité des techniques de détection de ces biomarqueurs sont basées sur des anticorps (Antal-Szalmas *et al.*, 2013; Dong *et al.*, 2017; Furuya *et al.*, 2019; Kim *et al.*, 2012; Ling *et al.*, 2008; Soler *et al.*, 2014; Vanarsa *et al.*, 2021; Wei *et al.*, 2009; Yu *et al.*, 2019; Zangar *et al.*, 2006). Ces molécules ont montré leur efficacité au cours des années, mais elles représentent des inconvénients, ce qui ouvre la porte à la recherche d'une meilleure solution, en l'occurrence les aptamères. Avec leur petite taille, la rapidité de leur sélection et synthèse ainsi que plusieurs autres avantages (résumé dans la **Table 9.2**), les aptamères ont pu remplacer les anticorps dans diverses applications (Bauer *et al.*, 2019; Chen & Yang, 2015; Kukova *et al.*, 2019; Lai *et al.*, 2017; Nezlin, 2016; Toh *et al.*, 2015).

Table 9.2 Comparaison des aptamères/aptazymes et les anticorps

	Aptamères/Aptazymes (ADN/ARN)	Anticorps
Avancés	<p>Premier aptamère sélectionné en 1990 par SELEX, avant la découverte d'aptamères naturels faisant partie des <i>riboswitchs</i> en 2003.</p> <p>Depuis, des centaines d'aptamères ont été sélectionnés et de multiples méthodes de SELEX ont été développées.</p> <p>Un seul aptamère approuvé cliniquement (pegatanib) mais plusieurs sont en phase clinique d'évaluation.</p> <p>Des plateformes chez des compagnies comme SOMAlogic et Neoventures sont complètement basées sur les aptamères dans des</p>	<p>Première production d'anticorps monoclonaux en 1975.</p> <p>1992 : premier anticorps approuvé en application thérapeutique et actuellement 133 anticorps thérapeutiques ont été approuvés selon le site de <i>antibodies society</i> d'où il y avait une montée exponentielle remarquée au cours des 10 dernières années.</p> <p>Utilisés dans une panoplie de types de kit et essais immunologiques pour la détection et diagnostique (immunobuvardage, ELISA, immunoprécipitation, cytométrie en flux, etc)</p>

	<p>but de diagnostics, biosenseurs et découvertes de biomarqueurs.</p> <p>Novaptech offre des essais basés sur les aptamères d'une façon personnalisée.</p> <p>Utilisé pour contrôler l'activité des enzymes comme la HotStart taq ADN polymérase de Qiagen.</p>	
<p>Avantages</p>	<p>Petite taille résultant en une immunogénicité faible et une meilleure pénétration intracellulaire.</p> <p>Affinité et spécificité élevée.</p> <p>Distinction entre des molécules très similaires</p> <p>Coût de synthèse et production pas trop cher.</p> <p>Synthèse complètement chimique et rapide générant des séquences identiques à chaque fois éliminant la possibilité d'avoir des variations d'un lot à un autre.</p> <p>Sélection en quelques semaines.</p> <p>Possibilité de sélectionner des aptamères contre des petites molécules ainsi que macromolécules et des cellules</p> <p>Aptes à des modifications chimiques résultant par une résistance élevée aux conditions difficiles (résistance aux nucléases, températures élevée.</p> <p>(Byun, 2021; Toh <i>et al.</i>, 2015)</p> <p>Détection directe par l'ajout d'un fluorophore à l'aptamère</p> <p>Une séquence est tout ce qu'il faut pour conserver l'aptamère aussi longtemps que désiré car les</p>	<p>Haute affinité et spécificité</p> <p>Des dizaines de méthodes et kits utilisant les anticorps sont déjà établis et leur efficacité est assez élevée.</p> <p>Utilisé depuis longtemps, alors des plateformes et des machines sont déjà installées dans des milliers de laboratoires autour du monde rendant leurs utilisations plus faciles que d'inventer de nouvelles techniques et méthodes de caractérisation, optimisation ou détection.</p> <p>Montre une efficacité importante dans le traitement de plusieurs sortes de maladies, incluant différent types de cancer et des maladies auto-immunes, ce qui est prouvé par le grand nombre d'anticorps utilisés actuellement comme thérapie approuvée (Lu <i>et al.</i>, 2020)</p>

	<p>aptamères sont synthétisés chimiquement et non par des cellules comme dans le cas des anticorps.</p> <p>Facile à modifier ou insérer dans d'autres molécules comme les ribozymes et les ARN guide afin de créer des aptazymes et des <i>riboswitchs</i>.</p>	
Inconvénients	<p>Dépendants des conditions de sélection ce qui peut diminuer son efficacité dans l'environnement de son application</p> <p>Nécessite des étapes de caractérisation et optimisation après la sélection pour s'assurer de leur activité en dehors des conditions de sélection</p> <p>Plusieurs modifications sont nécessaires pour augmenter la demie vie d'un aptamère surtout un aptamère ARN.</p> <p>Ils sont rapidement éliminés de la circulation sanguine ce qui nécessite l'ajout d'un PEG par exemple.</p> <p>Nombre élevé d'artéfacts qui peuvent subvenir pendant le SELEX (Byun, 2021; Lakhin <i>et al.</i>, 2013; Rozenblum <i>et al.</i>, 2016)</p>	<p>Taille assez grande</p> <p>Du matériel biologique et des organismes vivants nécessaires pour les anticorps monoclonaux ce qui crée un risque de variation entre les lots de chaque production et soulève des questions éthiques.</p> <p>Aptitude aux modifications limitée.</p> <p>Durée de sélection et production de quelques mois.</p> <p>Un anticorps secondaire, une enzyme et un substrat sont nécessaires pour la détection dans la plupart des techniques utilisant des anticorps.</p> <p>Une fois dénaturé, un anticorps est irrécupérable.</p> <p>Susceptible aux conditions environnementales comme changement de pH et température</p> <p>Immunogénicité élevée</p> <p>Faible taux de pénétration tissulaire</p> <p>Conservation limitée en cas où les cellules sont perdues ou endommagées (Byun, 2021; Chames <i>et al.</i>, 2009)</p>

En considérant tous ces arguments, en plus des miARN, nous avons choisi les protéines (Gal-7 et Gal-1) comme signal de détection d'un cancer et des aptamères comme outils

de détection. Nos aptamères ont été sélectionnés en combinaison avec des ribozymes, et ce afin de les inclure dans des circuits logiques ou pour les appliquer comme biosenseurs. Nous avons été capables de sélectionner *in vitro* des ARN catalytiques dont l'activité est modulée par la liaison de protéines choisies. Bien que la spécificité des séquences obtenues était partielle, elle peut être améliorée par l'ajout de quelques cycles de sélection ou la mutation des aptamères dans certaines positions. Pour cette raison, nous avons développé une méthode de sélection appuyée sur deux étapes : une pour l'affinité et une pour l'activité de clivage. En fait, il n'y avait pas beaucoup de méthodes de SELEX d'aptazymes publiées qui impliquaient deux étapes. L'étude de Breaker (Ferguson *et al.*, 2004), utilisant une sélection basée seulement sur l'activité de clivage ou une méthode hybride semblable à la nôtre (Ferguson *et al.*, 2004), a montré que les aptazymes issus de la méthode hybride représentent de meilleurs candidats. Notre méthode a permis d'obtenir un ARN ayant les deux fonctionnalités de détection et de clivage. Ceci implique qu'une seule molécule d'ARN puisse être utilisée pour ses deux fonctions combinées (biosenseur ou régulation de l'expression génique), ou seulement comme un aptamère en ignorant la partie ribozyme. Ces biosenseurs pourront être applicables dans une panoplie de maladies détectable à l'aide des biomarqueurs protéiques. De plus, le système est facilement modulable, alors il peut être utilisé pour détecter plusieurs molécules autres que les protéines, comme des polluants environnementaux, des toxines, des pathogènes, etc.

L'originalité de cette partie de notre travail vient de l'utilisation d'une mutation naturelle du cœur catalytique du ribozyme en tête de marteau « A6C », nous autorisant à avoir un contrôle sur son activité lors de la sélection. Par la suite, la réversion de la mutation permet de conserver la fonction du ribozyme dans des conditions plus standards, par exemple dans le cas de son utilisation dans un organisme vivant. Enfin, nous avons montré la possibilité de transformer les aptazymes obtenus en biosenseurs utilisables, *in vitro* ou potentiellement *in vivo*, par l'ajout d'un aptamère fluorescent à la tige III. En fait, les aptazymes basés sur l'ARN ont surtout été appliqués comme outils de contrôle de l'expression génique en *cis* par leur intégration devant des gènes d'intérêt (Chen *et al.*, 2010; Mustafina *et al.*, 2020; Win & Smolke, 2007b; Winkler *et al.*, 2004; Wittmann & Suess, 2011), mais l'ont rarement été comme biosenseur (Porter *et al.*, 2017; Rueda & Walter, 2006) (Gong *et al.*, 2015; Jaffrey, 2018) (Townshend *et al.*, 2021a). Ceci est dû à la fragilité de l'ARN, très sensible à la température, aux RNases ainsi qu'à l'activité de ribozymes non mutés dans les liquides physiologiques contenant des concentrations

importantes de magnésium. Ces problèmes peuvent actuellement être réglés par l'inclusion de ribonucléotides modifiés dans l'ARN sans affecter l'activité du ribozyme afin d'améliorer sa demi-vie. Une alternative utilisée constitue les enzymes à ADN « DNAzyme », qui montrent une meilleure résistance dans les conditions physiologiques et une efficacité de clivage aussi intéressante. En effet, plusieurs DNAzyme ont été transformés en biosenseur (Gong *et al.*, 2015; Khan *et al.*, 2021; McConnell *et al.*, 2020a; Sun *et al.*, 2010).

Malgré la multitude d'applications dans lesquelles les aptamères et aptazymes ont du potentiel, il reste encore un long chemin à parcourir avant d'atteindre le niveau d'utilisation d'anticorps ou de les remplacer. Les aptazymes en *cis*, en tant que régulateurs d'expression génique, ne sont pas facilement applicables par le fait qu'ils doivent s'insérer dans la séquence du gène d'intérêt, ce qui n'est pas possible pour les cellules humaines, par exemple. Quant à leur usage comme biosenseur, il y a davantage d'espoir, avec l'avancement technologique et la facilité d'automatisation de plusieurs essais ; de meilleures méthodes de sélection et de caractérisation des aptamères et aptazymes aideront de plus en plus à l'amélioration de ces outils.

10. Conclusion et perspectives

L'objectif principal de ce projet était initialement d'impliquer les ribozymes ayant une activité de clivage dans des applications en lien avec la thérapie génique ciblant le cancer, puis pour des applications plus variées, afin de prouver leur efficacité et utilité.

Dans un premier temps, nous avons utilisé des outils bioinformatiques, développés pour la conception de ribozymes, dans le but d'évaluer leur fiabilité. Nous avons pu montrer qu'un grand nombre de ribozymes générés par ces logiciels présentent une activité de clivage assez élevée *in vitro*. Ceci a impliqué des ribozymes de différents types, auto-clivants, actifs en *trans*, conventionnels et allostériques, ayant ou non des pseudonœuds dans leurs structures. Les résultats obtenus nous ont permis de donner de l'information en retour aux développeurs de ces programmes, ce qui a servi à en améliorer la conception et leurs présentations aux utilisateurs.

Parmi les ribozymes testés, plusieurs peuvent être employés dans des applications futures. Par exemple, nous avons pu prouver sa faisabilité c'est-à-dire le contrôle en *cis* de l'expression génique en insérant le ribozyme dans la séquence codante du gène ciblé, et ce sans affecter la séquence des codons. Nous avons prouvé le concept avec un ribozyme conçu par Enzymer qui a permis de diminuer l'expression de la protéine RFP chez des cellules procaryotes. Ce même principe est également ouvert à l'application chez les eucaryotes, et un contrôle allostérique peut être appliqué par l'utilisation d'un aptazyme liant la molécule désirée et convenable pour l'application.

En ce qui concerne les aptazymes, grâce au service web Ribosoft2.0, nous avons été en mesure de concevoir, pour la première fois, des aptazymes actifs en *trans* à partir des séquences connues d'aptazymes sélectionnés *in vitro*. Ceci ouvre la porte à plusieurs applications qui n'étaient pas possibles *in vivo*, malgré le fait que le concept ait été prouvé *in vitro*. En fait, bien que les aptazymes aient prouvé leur fiabilité dans le contrôle de l'expression génique allostérique, leur activité en *cis* limite beaucoup leurs usages dans les cellules vivantes. Ceci est dû à la complexité de l'insertion des ribozymes devant le gène d'intérêt ; ainsi, l'obtention d'aptazyme, ayant une activité en *trans*, résout ce problème et rajoute la possibilité de choisir le meilleur site à dégrader sur la cible.

À côté de ces aptazymes, dans un autre contexte d'utilisation d'outils bioinformatiques, nous avons obtenu du même laboratoire des séquences de portes logiques basées sur des ribozymes liants d'un jusqu'à trois miARN. Nos tests *in vitro* ont abouti à des résultats

encourageants de quelques ribozymes qui montraient une activité en présence de certaines combinaisons spécifiques de miARN. Bien que ces résultats ne soient pas parfaitement en concordance avec le but initial, mais avec quelques optimisations des designs, des portes logiques ayant l'activité désirée seront obtenues.

Un tel design de ribozyme pourrait permettre d'obtenir des outils utilisables en thérapie comme en diagnostic. En liant trois miARN en parallèle, un ribozyme pourrait servir comme outil de diagnostic par l'ajout d'un fluorophore qui n'émettra le signal qu'en présence de ces trois entrées. Ceci peut s'appliquer *in vitro* comme dans les cellules et des tissus. Dans un but thérapeutique, l'expression intracellulaire d'une thérapie génique comme un ribozyme, un siRNA, etc. peut être contrôlée par une porte logique répondant à plusieurs miARN, ce qui augmente la spécificité d'une telle thérapie.

En plus de ces portes logiques complexes, nous avons utilisé Ribosoft2.0 pour générer des ribozymes en *trans* ayant une séquence complémentaire à un seul miARN à la fois. Ces ribozymes ont été actifs *in vitro*, mais nous n'avons pas pu prouver leur activité dans les cellules. Malgré cela, ces ribozymes présenteront un outil thérapeutique dont l'activité ne serait pas contrôlée allostériquement, mais plutôt par la présence du miARN lié au complexe RISC causant la dégradation du ribozyme.

En dernier lieu, nous avons pu développer un biosenseur pour la détection de protéines Gal-7 et Gal-1 surexprimées dans le cas de plusieurs types de cancers. Ceci a été fait à l'aide d'une sélection d'aptazymes *in vitro*, de leur couplage à la fin à un aptamère fluorescent, ainsi qu'à des billes fluorescentes. Ce travail peut mener au développement d'un outil de diagnostic applicable à la fois pour des analyses d'échantillons de patients, par exemple, ou en milieu intracellulaire par la transfection de l'ADN codant pour l'aptazyme. Étant donné que la méthode de sélection utilisée dans notre projet a été basée sur des billes magnétiques, le processus de SELEX pourrait être automatisé afin d'être apte à sélectionner plus que 90 ligands à la fois. Ces derniers n'ont pas forcément besoin d'être des protéines, car des aptazymes existent pour un large éventail de molécules, et même pour des cellules. Ceci ouvre la porte à l'idée d'emploi de ce principe dans d'autres pathologies, comme des maladies infectieuses, ou dans un contexte environnemental en adaptant le système en fonction de la situation. Comme exemple, des aptazymes à ADN, qui sont basés sur des DNAzyme, seront une meilleure alternative pour les applications dans des milieux extracellulaires ou des échantillons environnementaux.

En conclusion, notre travail ne représente qu'un début de plusieurs études à suivre. Nous avons pu prouver de nouveaux concepts, puis démontrer l'importance et l'efficacité des outils bioinformatiques dans la génération des ARN synthétiques. La diversité d'outils présentés dans cette thèse ouvre la porte à des études chez les procaryotes et les eucaryotes ainsi que dans des milieux intracellulaires comme extracellulaires.

11. Références

- Abu-Baker A, Kharma N, Perreault J, Grant A, Shekarabi M, Maios C, Dona M, Neri C, Dion PA, Parker A, Varin L & Rouleau GA (2019) RNA-Based Therapy Utilizing Oculopharyngeal Muscular Dystrophy Transcript Knockdown and Replacement. *Molecular therapy. Nucleic acids* 15:12-25.
- Adachi T & Nakamura Y (2019) Aptamers: A Review of Their Chemical Properties and Modifications for Therapeutic Application. *Molecules* 24(23).
- Adli M (2018) The CRISPR tool kit for genome editing and beyond. *Nature communications* 9(1):1911.
- Advedissian T, Deshayes F & Viguier M (2017) Galectin-7 in Epithelial Homeostasis and Carcinomas. *International journal of molecular sciences* 18(12).
- Afonin KA, Lindsay B & Shapiro BA (2015) Engineered RNA Nanodesigns for Applications in RNA Nanotechnology. *DNA RNA Nanotechnol* 1(1):1-15.
- Agop-Nersesian C, Pfahler J, Lanzer M & Meissner M (2008) Functional expression of ribozymes in Apicomplexa: towards exogenous control of gene expression by inducible RNA-cleavage. *International journal for parasitology* 38(6):673-681.
- Aguirre-Hernandez R, Hoos HH & Condon A (2007) Computational RNA secondary structure design: empirical complexity and improved methods. *BMC bioinformatics* 8:34.
- Ahirwar R, Nahar S, Aggarwal S, Ramachandran S, Maiti S & Nahar P (2016) In silico selection of an aptamer to estrogen receptor alpha using computational docking employing estrogen response elements as aptamer-alike molecules. *Scientific reports* 6:21285.
- Ahmadzada T, Reid G & McKenzie DR (2018) Fundamentals of siRNA and miRNA therapeutics and a review of targeted nanoparticle delivery systems in breast cancer. *Biophysical reviews* 10(1):69-86.
- Akaboshi E, Guerrier-Takada C & Altman S (1980) Veal heart ribonuclease P has an essential RNA component. *Biochemical and biophysical research communications* 96(2):831-837.
- Alam KK, Chang JL & Burke DH (2015) FASTAptamer: A Bioinformatic Toolkit for High-throughput Sequence Analysis of Combinatorial Selections. *Molecular therapy. Nucleic acids* 4:e230.
- Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. *Journal of molecular biology* 215(3):403-410.
- Alves Dos Santos K, Clemente Dos Santos IC, Santos Silva C, Gomes Ribeiro H, de Farias Domingos I & Nogueira Silbiger V (2020) Circulating Exosomal miRNAs as Biomarkers for the Diagnosis and Prognosis of Colorectal Cancer. *International journal of molecular sciences* 22(1).
- Amontov S & Jaschke A (2006) Controlling the rate of organic reactions: rational design of allosteric Diels-Alderase ribozymes. *Nucleic acids research* 34(18):5032-5038.
- Anastasiadou E, Seto AG, Beatty X, Hermreck M, Gilles ME, Stroopinsky D, Pinter-Brown LC, Pestano L, Marchese C, Avigan D, Trivedi P, Escolar DM, Jackson AL & Slack FJ (2021) Cobomarsen, an Oligonucleotide Inhibitor of miR-155, Slows DLBCL Tumor Cell Growth In Vitro and In Vivo. *Clinical cancer research : an official journal of the American Association for Cancer Research* 27(4):1139-1149.
- Andrianantoandro E, Basu S, Karig DK & Weiss R (2006) Synthetic biology: new engineering rules for an emerging discipline. *Molecular systems biology* 2:2006 0028.
- Anelli L, Zagaria A, Specchia G, Musto P & Albano F (2021) Dysregulation of miRNA in Leukemia: Exploiting miRNA Expression Profiles as Biomarkers. *International journal of molecular sciences* 22(13).
- Anonyme (<https://git-scm.com/>. <https://git-scm.com/>)

- Anonyme (<https://pypi.org/project/pip/>. <https://pypi.org/project/pip/>)
- Anonyme (2000) Heptazyme. *Drug news & perspectives* 13(2):112.
- Anonyme (2018) Eteplirsén treatment for Duchenne muscular dystrophy: Exon skipping and dystrophin production. *Neurology* 91(13):637.
- Antal-Szalmás P, Nagy B, Jr., Debreceni IB & Kappelmayer J (2013) Measurement of Soluble Biomarkers by Flow Cytometry. *Ejifcc* 23(4):135-142.
- Araki M, Okuno Y, Hara Y & Sugiura Y (1998) Allosteric regulation of a ribozyme activity through ligand-induced conformational change. *Nucleic acids research* 26(14):3379-3384.
- Arnold S, Pampalakis G, Kantiotou K, Silva D, Cortez C, Missailidis S & Sotiropoulou G (2012) One round of SELEX for the generation of DNA aptamers directed against KLK6. *Biological chemistry* 393(5):343-353.
- Arriola JT & Muller UF (2020) A combinatorial method to isolate short ribozymes from complex ribozyme libraries. *Nucleic acids research* 48(20):e116.
- Auslander D, Auslander S, Pierrat X, Hellmann L, Rachid L & Fussenegger M (2018) Programmable full-adder computations in communicating three-dimensional cell cultures. *Nature methods* 15(1):57-60.
- Auslander S, Ketzner P & Hartig JS (2010) A ligand-dependent hammerhead ribozyme switch for controlling mammalian gene expression. *Molecular bioSystems* 6(5):807-814.
- Bae S, Park J & Kim JS (2014) Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* 30(10):1473-1475.
- Bai J, Banda N, Lee NS, Rossi J & Akkina R (2002) RNA-based anti-HIV-1 gene therapeutic constructs in SCID-hu mouse model. *Molecular therapy : the journal of the American Society of Gene Therapy* 6(6):770-782.
- Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116(2):281-297.
- Bartel DP & Szostak JW (1993) Isolation of new ribozymes from a large pool of random sequences [see comment]. *Science* 261(5127):1411-1418.
- Bartolazzi A (2018) Galectins in Cancer and Translational Medicine: From Bench to Bedside. *International journal of molecular sciences* 19(10).
- Batey RT, Rambo RP & Doudna JA (1999) Tertiary Motifs in RNA Structure and Folding. *Angew Chem Int Ed Engl* 38(16):2326-2343.
- Bauer M, Strom M, Hammond DS & Shigdar S (2019) Anything You Can Do, I Can Do Better: Can Aptamers Replace Antibodies in Clinical Diagnostic Applications? *Molecules* 24(23).
- Bayer TS & Smolke CD (2005) Programmable ligand-controlled riboregulators of eukaryotic gene expression. *Nature biotechnology* 23(3):337-343.
- Beaudry A, DeFoe J, Zinnen S, Burgin A & Beigelman L (2000) In vitro selection of a novel nuclease-resistant RNA phosphodiesterase. *Chemistry & biology* 7(5):323-334.
- Bedard AV, Hien EDM & Lafontaine DA (2020) Riboswitch regulation mechanisms: RNA, metabolites and regulatory proteins. *Biochimica et biophysica acta. Gene regulatory mechanisms* 1863(3):194501.
- Beilstein K, Wittmann A, Grez M & Suess B (2015) Conditional control of mammalian gene expression by tetracycline-dependent hammerhead ribozymes. *ACS synthetic biology* 4(5):526-534.
- Belka C & Budach W (2002) Anti-apoptotic Bcl-2 proteins: structure, function and relevance for radiation biology. *International journal of radiation biology* 78(8):643-658.
- Bellecave P, Cazenave C, Rumi J, Staedel C, Cosnefroy O, Andreola ML, Ventura M, Tarrago-Litvak L & Astier-Gin T (2008) Inhibition of hepatitis C virus (HCV) RNA polymerase by

- DNA aptamers: mechanism of inhibition of in vitro RNA synthesis and effect on HCV-infected cells. *Antimicrobial agents and chemotherapy* 52(6):2097-2110.
- Benenson Y (2009) Biocomputers: from test tubes to live cells. *Molecular bioSystems* 5(7):675-685.
- Benenson Y (2012) Synthetic biology with RNA: progress report. *Current opinion in chemical biology* 16(3-4):278-284.
- Benenson Y, Gil B, Ben-Dor U, Adar R & Shapiro E (2004) An autonomous molecular computer for logical control of gene expression. *Nature* 429(6990):423-429.
- Bennett CF (2019) Therapeutic Antisense Oligonucleotides Are Coming of Age. *Annual review of medicine* 70:307-321.
- Bennett CF, Krainer AR & Cleveland DW (2019) Antisense Oligonucleotide Therapies for Neurodegenerative Diseases. *Annual review of neuroscience* 42:385-406.
- Berindan-Neagoe I, Monroig Pdel C, Pasculli B & Calin GA (2014) MicroRNAome genome: a treasure for cancer diagnosis and therapy. *CA: a cancer journal for clinicians* 64(5):311-336.
- Bertrand E, Castanotto D, Zhou C, Carbonnelle C, Lee NS, Good P, Chatterjee S, Grange T, Pictet R & Kohn D (1997) The expression cassette determines the functional activity of ribozymes in mammalian cells by controlling their intracellular localization. *Rna* 3(1):75-88.
- Berzal-Herranz A, Joseph S & Burke JM (1992) In vitro selection of active hairpin ribozymes by sequential RNA-catalyzed cleavage and ligation reactions. *Genes & development* 6(1):129-134.
- Bhardwaj A & Nain V (2021) TALENs-an indispensable tool in the era of CRISPR: a mini review. *Journal, genetic engineering & biotechnology* 19(1):125.
- Bhartiya D & Scaria V (2016) Genomic variations in non-coding RNAs: Structure, function and regulation. *Genomics* 107(2-3):59-68.
- Bhat SA, Ahmad SM, Mumtaz PT, Malik AA, Dar MA, Urwat U, Shah RA & Ganai NA (2016) Long non-coding RNAs: Mechanism of action and functional utility. *Non-coding RNA research* 1(1):43-50.
- Bhatti GK, Khullar N, Sidhu IS, Navik US, Reddy AP, Reddy PH & Bhatti JS (2021) Emerging role of non-coding RNA in health and disease. *Metabolic brain disease* 36(6):1119-1134.
- Bindewald E, Afonin K, Jaeger L & Shapiro BA (2011) Multistrand RNA secondary structure prediction and nanostructure design including pseudoknots. *ACS nano* 5(12):9542-9551.
- Bindewald EaG, Calvin and Boyle, Brett and O'Connor, Mary and Shapiro, Bruce A (2008) Computational strategies for the automated design of RNA nanoscale structures from building blocks using NanoTiler. *Journal of Molecular Graphics and Modelling* 27:299--308.
- Birch CM, Hou HW, Han J & Niles JC (2015) Identification of malaria parasite-infected red blood cell surface aptamers by inertial microfluidic SELEX (I-SELEX). *Scientific reports* 5:11347.
- Blount KF & Uhlenbeck OC (2005) The structure-function dilemma of the hammerhead ribozyme. *Annual review of biophysics and biomolecular structure* 34:415-440.
- Boccaletto P, Magnus M, Almeida C, Zyla A, Astha A, Pluta R, Baginski B, Jankowska E, Dunin-Horkawicz S, Wirecki TK, Boniecki MJ, Stefaniak F & Bujnicki JM (2018) RNArchitecture: a database and a classification system of RNA families, with a focus on structural information. *Nucleic acids research* 46(D1):D202-D205.
- Boch J & Bonas U (2010) Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annual review of phytopathology* 48:419-436.

- Bonfa G, Blazquez-Roman J, Tarnai R & Siciliano V (2020) Precision Tools in Immuno-Oncology: Synthetic Gene Circuits for Cancer Immunotherapy. *Vaccines* 8(4).
- Bonhoeffer S, McCaskill JS, Stadler PF & Schuster P (1993) RNA multi-structure landscapes. A study based on temperature dependent partition functions. *European biophysics journal : EBJ* 22(1):13-24.
- Boscaino V, Fiannaca A, La Paglia L, La Rosa M, Rizzo R & Urso A (2019) MiRNA therapeutics based on logic circuits of biological pathways. *BMC bioinformatics* 20(Suppl 9):344.
- Bouhedda F, Fam KT, Collot M, Autour A, Marzi S, Klymchenko A & Ryckelynck M (2020) A dimerization-based fluorogenic dye-aptamer module for RNA imaging in live cells. *Nature chemical biology* 16(1):69-76.
- Boussebayle A, Groher F & Suess B (2019) RNA-based Capture-SELEX for the selection of small molecule-binding aptamers. *Methods* 161:10-15.
- Brazda V, Kolomaznik J, Lysek J, Bartas M, Fojta M, Stastny J & Mergny JL (2019) G4Hunter web application: a web server for G-quadruplex prediction. *Bioinformatics* 35(18):3493-3495.
- Breaker RR (2011) Prospects for riboswitch discovery and analysis. *Molecular cell* 43(6):867-879.
- Brierley I, Pennell S & Gilbert RJ (2007) Viral RNA pseudoknots: versatile motifs in gene expression and replication. *Nature reviews. Microbiology* 5(8):598-610.
- Brinchmann MF, Patel DM & Iversen MH (2018) The Role of Galectins as Modulators of Metabolism and Inflammation. *Mediators of inflammation* 2018:9186940.
- Broda M, Kierzek E, Gdaniec Z, Kulinski T & Kierzek R (2005) Thermodynamic stability of RNA structures formed by CNG trinucleotide repeats. Implication for prediction of RNA structure. *Biochemistry* 44(32):10873-10882.
- Brophy JA & Voigt CA (2014) Principles of genetic circuit design. *Nature methods* 11(5):508-520.
- Brown TS, Chadalavada DM & Bevilacqua PC (2004) Design of a highly reactive HDV ribozyme sequence uncovers facilitation of RNA folding by alternative pairings and physiological ionic strength. *Journal of molecular biology* 341(3):695-712.
- Bu Z & Callaway DJ (2011) Proteins move! Protein dynamics and long-range allostery in cell signaling. *Advances in protein chemistry and structural biology* 83:163-221.
- Byun J (2021) Recent Progress and Opportunities for Nucleic Acid Aptamers. *Life (Basel)* 11(3).
- Cai Z, Fu Y, Qiu Z, Wang Y, Wang W, Gu W, Li Z, Wu S & Gao F (2021) Multitarget Reaction Programmable Automatic Diagnosis and Treatment Logic Device. *ACS nano* 15(12):19150-19164.
- Callaway E (2021) DeepMind's AI predicts structures for a vast trove of proteins. *Nature* 595(7869):635.
- Cameron DE, Bashor CJ & Collins JJ (2014) A brief history of synthetic biology. *Nature reviews. Microbiology* 12(5):381-390.
- Cameron FH & Jennings PA (1989) Specific gene suppression by engineered ribozymes in monkey cells. *Proceedings of the National Academy of Sciences of the United States of America* 86(23):9139-9143.
- Campbell TB & Cech TR (1995) Identification of ribozymes within a ribozyme library that efficiently cleave a long substrate RNA. *RNA* 1(6):598-609.
- Campbell TB, McDonald CK & Hagen M (1997) The effect of structure in a long target RNA on ribozyme cleavage efficiency. *Nucleic acids research* 25(24):4985-4993.
- Cao S & Chen SJ (2006) Predicting RNA pseudoknot folding thermodynamics. *Nucleic acids research* 34(9):2634-2652.
- Cao S & Chen SJ (2007) Biphasic folding kinetics of RNA pseudoknots and telomerase RNA activity. *Journal of molecular biology* 367(3):909-924.

- Carbonell A, Flores R & Gago S (2011) Trans-cleaving hammerhead ribozymes with tertiary stabilizing motifs: in vitro and in vivo activity against a structured viroid RNA. *Nucleic acids research* 39(6):2432-2444.
- Caroli J, Forcato M & Bicciato S (2020) APTANI2: update of aptamer selection through sequence-structure analysis. *Bioinformatics* 36(7):2266-2268.
- Cech TR (1986) A model for the RNA-catalyzed replication of RNA. *Proceedings of the National Academy of Sciences of the United States of America* 83(12):4360-4363.
- Cech TR (2012) The RNA worlds in context. *Cold Spring Harbor perspectives in biology* 4(7):a006742.
- Cech TR, Zaugg AJ & Grabowski PJ (1981) In vitro splicing of the ribosomal RNA precursor of Tetrahymena: involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell* 27(3 Pt 2):487-496.
- Chadalavada DM, Senchak SE & Bevilacqua PC (2002) The folding pathway of the genomic hepatitis delta virus ribozyme is dominated by slow folding of the pseudoknots. *Journal of molecular biology* 317(4):559-575.
- Chames P, Van Regenmortel M, Weiss E & Baty D (2009) Therapeutic antibodies: successes, limitations and hopes for the future. *British journal of pharmacology* 157(2):220-233.
- Chandrasegaran S (2017) Recent advances in the use of ZFN-mediated gene editing for human gene therapy. *Cell & gene therapy insights* 3(1):33-41.
- Chang D, Zakaria S, Esmaeili Samani S, Chang Y, Filipe CDM, Soleymani L, Brennan JD, Liu M & Li Y (2021) Functional Nucleic Acids for Pathogenic Bacteria Detection. *Accounts of chemical research* 54(18):3540-3549.
- Chapman KB & Szostak JW (1994) In vitro selection of catalytic RNAs. *Current opinion in structural biology* 4:618-622.
- Chappell J, Takahashi MK & Lucks JB (2015) Creating small transcription activating RNAs. *Nature chemical biology* 11(3):214-220.
- Chatterjee SK & Zetter BR (2005) Cancer biomarkers: knowing the present and predicting the future. *Future Oncol* 1(1):37-50.
- Chau THT, Mai DHA, Pham DN, Le HTQ & Lee EY (2020) Developments of Riboswitches and Toehold Switches for Molecular Detection-Biosensing and Molecular Diagnostics. *International journal of molecular sciences* 21(9).
- Chen A & Yang S (2015) Replacing antibodies with aptamers in lateral flow immunoassay. *Biosensors & bioelectronics* 71:230-242.
- Chen AG, Sudarsan N & Breaker RR (2011) Mechanism for gene control by a natural allosteric group I ribozyme. *RNA* 17(11):1967-1972.
- Chen CC & Luptak A (2021) Hunting for human ribozymes. *Nature chemical biology* 17(5):507-508.
- Chen CY, Ezzeddine N & Shyu AB (2008) Messenger RNA half-life measurements in mammalian cells. *Methods in enzymology* 448:335-357.
- Chen D, Yang X, Liu M, Zhang Z & Xing E (2021a) Roles of miRNA dysregulation in the pathogenesis of multiple myeloma. *Cancer gene therapy* 28(12):1256-1268.
- Chen J, Fu S, Zhang C, Liu H & Su X (2022) DNA Logic Circuits for Cancer Theranostics. *Small* 10.1002/sml.202108008:e2108008.
- Chen Q, Gao R & Jia L (2021b) Enhancement of the peroxidase-like activity of aptamers modified gold nanoclusters by bacteria for colorimetric detection of Salmonella typhimurium. *Talanta* 221:121476.
- Chen W, Li Z, Deng P, Li Z, Xu Y, Li H, Su W & Qin J (2021c) Advances of Exosomal miRNAs in Breast Cancer Progression and Diagnosis. *Diagnostics (Basel)* 11(11).

- Chen W, Smeekens JM & Wu R (2016) Systematic study of the dynamics and half-lives of newly synthesized proteins in human cells. *Chem Sci* 7(2):1393-1400.
- Chen X, Denison L, Levy M & Ellington AD (2009) Direct selection for ribozyme cleavage activity in cells. *RNA* 15(11):2035-2045.
- Chen YY, Jensen MC & Smolke CD (2010) Genetic control of mammalian T-cell proliferation with synthetic RNA regulatory systems. *Proceedings of the National Academy of Sciences of the United States of America* 107(19):8531-8536.
- Cheng C, Chen YH, Lennox KA, Behlke MA & Davidson BL (2013) In vivo SELEX for Identification of Brain-penetrating Aptamers. *Molecular therapy. Nucleic acids* 2:e67.
- Cheng J, Luo J, Zhang X, Hu J, Hui H, Wang C & Stern A (2000) Inhibition of cell proliferation in HCC-9204 hepatoma cells by a c-myc specific ribozyme. *Cancer gene therapy* 7(3):407-412.
- Chou FC, Chen HY, Kuo CC & Sytwu HK (2018) Role of Galectins in Tumors and in Clinical Immunotherapy. *International journal of molecular sciences* 19(2).
- Churkin A, Retwitzer MD, Reinharz V, Ponty Y, Waldispuhl J & Barash D (2018) Design of RNAs: comparing programs for inverse RNA folding. *Briefings in bioinformatics* 19(2):350-358.
- Churkin AaR, Matan Drory and Reinharz, Vladimir and Ponty, Yann and Waldispuhl, Jerome and Barash, Danny (2017) Design of RNAs: comparing programs for inverse RNA folding. *Briefings in bioinformatics* :bbw120.
- Citti L & Rainaldi G (2005) Synthetic hammerhead ribozymes as therapeutic tools to control disease genes. *Current gene therapy* 5(1):11-24.
- Cobaleda C & Sanchez-Garcia I (2001) RNase P: from biological function to biotechnological applications. *Trends in biotechnology* 19(10):406-411.
- Cochrane JC, Lipchock SV, Smith KD & Strobel SA (2009) Structural and chemical basis for glucosamine 6-phosphate binding and activation of the glmS ribozyme. *Biochemistry* 48(15):3239-3246.
- Collantes JC, Tan VM, Xu H, Ruiz-Uriguen M, Alasadi A, Guo J, Tao H, Su C, Tyc KM, Selmi T, Lambourne JJ, Harbottle JA, Stombaugh J, Xing J, Wiggins CM & Jin S (2021) Development and Characterization of a Modular CRISPR and RNA Aptamer Mediated Base Editing System. *The CRISPR journal* 4(1):58-68.
- Collins JA, Irnov I, Baker S & Winkler WC (2007) Mechanism of mRNA destabilization by the glmS ribozyme. *Genes & development* 21(24):3356-3368.
- Condon AaD, Beth and Rastegari, Baharak and Zhao, Shelly and Tarrant, Finbarr (2004) Classifying RNA pseudoknotted structures. *Theoretical Computer Science* 320:35--50.
- Crick FH, Barnett L, Brenner S & Watts-Tobin RJ (1961) General nature of the genetic code for proteins. *Nature* 192:1227-1232.
- Cui W, Shen X, Wang C, Bibi A, Cudjoe O, Zhao L, Yu L, Du J, Xu Y, Chen X, Shen J & Wang W (2021) Direct enzyme-linked aptamer assay (DELAA) for diagnosis of toxoplasmosis by detection of SAG1 protein in mice and humans. *Acta tropica* 226:106255.
- Dagenais P & Legault P (2021) In Vitro Selection of Varkud Satellite Ribozyme Variants that Cleave a Modified Stem-Loop Substrate. *Methods Mol Biol* 2167:61-77.
- Dahm SC, Derrick WB & Uhlenbeck OC (1993) Evidence for the role of solvated metal hydroxide in the hammerhead cleavage mechanism. *Biochemistry* 32(48):13040-13045.
- Dahm SC & Uhlenbeck OC (1991) Role of divalent metal ions in the hammerhead RNA cleavage reaction. *Biochemistry* 30(39):9464-9469.
- Dana H, Chalbatani GM, Mahmoodzadeh H, Karimloo R, Rezaiean O, Moradzadeh A, Mehmandoost N, Moazzen F, Mazraeh A, Marmari V, Ebrahimi M, Rashno MM, Abadi SJ

- & Gharagouzlo E (2017) Molecular Mechanisms and Biological Functions of siRNA. *International journal of biomedical science : IJBS* 13(2):48-57.
- Darmostuk M, Rimpelova S, Gbelcova H & Ruml T (2015) Current approaches in SELEX: An update to aptamer selection technology. *Biotechnology advances* 33(6 Pt 2):1141-1161.
- Davydova A, Vorobjeva M, Pyshnyi D, Altman S, Vlassov V & Venyaminova A (2016) Aptamers against pathogenic microorganisms. *Critical reviews in microbiology* 42(6):847-865.
- De la Pena M, Gago S & Flores R (2003) Peripheral regions of natural hammerhead ribozymes greatly increase their self-cleavage activity. *The EMBO journal* 22(20):5561-5570.
- de la Pena M & Garcia-Robles I (2010) Intronic hammerhead ribozymes are ultraconserved in the human genome. *EMBO reports* 11(9):711-716.
- de Solis CA, Ho A, Holehonnur R & Ploski JE (2016) The Development of a Viral Mediated CRISPR/Cas9 System with Doxycycline Dependent gRNA Expression for Inducible In vitro and In vivo Genome Editing. *Frontiers in molecular neuroscience* 9:70.
- Debiais M, Lelievre A, Smietana M & Muller S (2020) Splitting aptamers and nucleic acid enzymes for the development of advanced biosensors. *Nucleic acids research* 48(7):3400-3422.
- Desai SK & Gallivan JP (2004) Genetic screens and selections for small molecules based on a synthetic riboswitch that activates protein translation. *Journal of the American Chemical Society* 126(41):13247-13254.
- Dethoff EA, Chugh J, Mustoe AM & Al-Hashimi HM (2012) Functional complexity and regulation through RNA dynamics. *Nature* 482(7385):322-330.
- Dewson G & Kluck R (2010) Bcl-2 family-regulated apoptosis in health and disease. *Cell health and Cytoskeleton* 2:9-22.
- Dexheimer PJ & Cochella L (2020) MicroRNAs: From Mechanism to Organism. *Frontiers in cell and developmental biology* 8:409.
- Di Fusco D, Dinallo V, Marafini I, Figliuzzi MM, Romano B & Monteleone G (2019) Antisense Oligonucleotide: Basic Concepts and Therapeutic Application in Inflammatory Bowel Disease. *Frontiers in pharmacology* 10:305.
- Di Tomasso G, Jenkins LMM & Legault P (2016) ARiBo pull-down for riboproteomic studies based on label-free quantitative mass spectrometry. *RNA* 22(11):1760-1770.
- Dirks RM, Lin M, Winfree E & Pierce NA (2004) Paradigms for computational nucleic acid design. *Nucleic acids research* 32(4):1392-1403.
- Dirks RM & Pierce NA (2003) A partition function algorithm for nucleic acid secondary structure including pseudoknots. *Journal of computational chemistry* 24(13):1664-1677.
- Dong YX, Cao JT, Liu YM & Ma SH (2017) A novel immunosensing platform for highly sensitive prostate specific antigen detection based on dual-quenching of photocurrent from CdSe sensitized TiO₂ electrode by gold nanoparticles decorated polydopamine nanospheres. *Biosensors & bioelectronics* 91:246-252.
- Donmus B, Unal S, Kirmizitas FC & Turkoglu Lacin N (2021) Virus-associated ribozymes and nano carriers against COVID-19. *Artificial cells, nanomedicine, and biotechnology* 49(1):204-218.
- Doshi KJ, Cannone JJ, Cobaugh CW & Gutell RR (2004) Evaluation of the suitability of free-energy minimization using nearest-neighbor energy parameters for RNA secondary structure prediction. *BMC bioinformatics* 5:105.
- Dotu I, Garcia-Martin JA, Slinger BL, Mechery V, Meyer MM & Clote P (2014) Complete RNA inverse folding: computational design of functional hammerhead ribozymes. *Nucleic acids research* 42(18):11752-11762.

- Doyle EL, Booher NJ, Standage DS, Voytas DF, Brendel VP, Vandyk JK & Bogdanove AJ (2012) TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic acids research* 40(Web Server issue):W117-122.
- Drory Retwitzer M, Reinharz V, Ponty Y, Waldispuhl J & Barash D (2016) incaRNAfbinv: a web server for the fragment-based design of RNA sequences. *Nucleic acids research* 44(W1):W308-314.
- Du X, Cheng X, Li W, Ge Z, Zhong C, Fan C & Gu H (2021) Engineering Allosteric Ribozymes to Detect Thiamine Pyrophosphate in Whole Blood. *Analytical chemistry* 93(9):4277-4284.
- Duan X, Qiao S, Li D, Li S, Zheng Z, Wang Q & Zhu X (2021) Circulating miRNAs in Serum as Biomarkers for Early Diagnosis of Non-small Cell Lung Cancer. *Frontiers in genetics* 12:673926.
- Dube-Delarosbil C & St-Pierre Y (2018) The emerging role of galectins in high-fatality cancers. *Cellular and molecular life sciences : CMLS* 75(7):1215-1226.
- Ellington AD & Szostak JW (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature* 346(6287):818.
- Elskens JP, Elskens JM & Madder A (2020) Chemical Modification of Aptamers for Increased Binding Affinity in Diagnostic Applications: Current Status and Future Prospects. *International journal of molecular sciences* 21(12).
- Ely A, Naidoo T & Arbuthnot P (2009) Efficient silencing of gene expression with modular trimeric Pol II expression cassettes comprising microRNA shuttles. *Nucleic acids research* 37(13):e91.
- Ender A, Etzel M, Hammer S, Findeiss S, Stadler P & Morl M (2021) Ligand-dependent tRNA processing by a rationally designed RNase P riboswitch. *Nucleic acids research* 49(3):1784-1800.
- Endy D (2005) Foundations for engineering biology. *Nature* 438(7067):449-453.
- Esmaili-Taheri A & Ganjtabesh M (2015) ERD: a fast and reliable tool for RNA design including constraints. *BMC bioinformatics* 16:20.
- Esposito CL, Catuogno S & de Franciscis V (2014) Aptamer-mediated selective delivery of short RNA therapeutics in cancer cells. *Journal of RNAi and gene silencing : an international journal of RNA and gene targeting research* 10:500-506.
- Etzel M & Morl M (2017) Synthetic Riboswitches: From Plug and Pray toward Plug and Play. *Biochemistry* 56(9):1181-1198.
- Famulok M, Hartig JS & Mayer G (2007a) Functional aptamers and aptazymes in biotechnology, diagnostics, and therapy. *Chemical reviews* 107(9):3715-3743.
- Famulok M, Hartig JS & Mayer G (2007b) Functional aptamers and aptazymes in biotechnology, diagnostics, and therapy. *Chemical reviews* 107(9):3715-3743.
- Farina AR & Mackay AR (2014) Gelatinase B/MMP-9 in Tumour Pathogenesis and Progression. *Cancers* 6(1):240-296.
- Farjami E, Campos R, Nielsen JS, Gothelf KV, Kjems J & Ferapontova EE (2013) RNA aptamer-based electrochemical biosensor for selective and label-free analysis of dopamine. *Analytical chemistry* 85(1):121-128.
- Farzin L, Shamsipur M & Sheibani S (2017) A review: Aptamer-based analytical strategies using the nanomaterials for environmental and human monitoring of toxic heavy metals. *Talanta* 174:619-627.
- Fedoruk-Wyszomirska A, Szymanski M, Glodowicz P, Gabryelska M, Wyszko E, Estrin WJ & Barciszewski J (2015) Inhibition of HIV-1 gp41 expression with hammerhead ribozymes. *The Biochemical journal* 471(1):53-66.

- Fei Q, Zhang H, Fu L, Dai X, Gao B, Ni M, Ge C, Li J, Ding X, Ke Y, Yao X & Zhu J (2008) Experimental cancer gene therapy by multiple anti-survivin hammerhead ribozymes. *Acta biochimica et biophysica Sinica* 40(6):466-477.
- Feldmann A, Hoffmann A, Bergmann R, Koristka S, Berndt N, Arndt C, Rodrigues Loureiro L, Kittel-Boselli E, Mitwasi N, Kegler A, Lamprecht C, Gonzalez Soto KE & Bachmann M (2020) Versatile chimeric antigen receptor platform for controllable and combinatorial T cell therapy. *Oncoimmunology* 9(1):1785608.
- Felletti M & Hartig JS (2017) Ligand-dependent ribozymes. *Wiley interdisciplinary reviews. RNA* 8(2).
- Felletti M, Stifel J, Wurmthaler LA, Geiger S & Hartig JS (2016a) Twister ribozymes as highly versatile expression platforms for artificial riboswitches. *Nature communications* 7:12834.
- Felletti M, Stifel J, Wurmthaler LA, Geiger S & Hartig JS (2016b) Twister ribozymes as highly versatile expression platforms for artificial riboswitches. *Nature communications* 7.
- Ferguson A, Boomer RM, Kurz M, Keene SC, Diener JL, Keefe AD, Wilson C & Cload ST (2004) A novel strategy for selection of allosteric ribozymes yields RiboReporter sensors for caffeine and aspartame. *Nucleic acids research* 32(5):1756-1766.
- Ferragut Cardoso AP, Banerjee M, Nail AN, Lykoudi A & States JC (2021) miRNA dysregulation is an emerging modulator of genomic instability. *Seminars in cancer biology* 76:120-131.
- Ferre-D'Amare AR (2010) The glmS ribozyme: use of a small molecule coenzyme by a gene-regulatory RNA. *Quarterly reviews of biophysics* 43(4):423-447.
- Ferre-D'Amare AR & Scott WG (2010) Small self-cleaving ribozymes. *Cold Spring Harbor perspectives in biology* 2(10):a003574.
- Filippi L, Bagni O & Nervi C (2020) Aptamer-based technology for radionuclide targeted imaging and therapy: a promising weapon against cancer. *Expert review of medical devices* 17(8):751-758.
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE & Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391(6669):806-811.
- Fiskaa T & Birgisdottir AB (2010) RNA reprogramming and repair based on trans-splicing group I ribozymes. *New biotechnology* 27(3):194-203.
- Forster AC & Symons RH (1987a) Self-cleavage of plus and minus RNAs of a virusoid and a structural model for the active sites. *Cell* 49(2):211-220.
- Forster AC & Symons RH (1987b) Self-cleavage of virusoid RNA is performed by the proposed 55-nucleotide active site. *Cell* 50(1):9-16.
- Freier SM, Kierzek R, Jaeger JA, Sugimoto N, Caruthers MH, Neilson T & Turner DH (1986) Improved free-energy parameters for predictions of RNA duplex stability. *Proceedings of the National Academy of Sciences of the United States of America* 83(24):9373-9377.
- Frommer J, Appel B & Muller S (2015) Ribozymes that can be regulated by external stimuli. *Current opinion in biotechnology* 31:35-41.
- Fukuda M, Kurihara K, Yamaguchi S, Oyama Y & Deshimaru M (2014) Improved design of hammerhead ribozyme for selective digestion of target RNA through recognition of site-specific adenosine-to-inosine RNA editing. *RNA* 20(3):392-405.
- Furuya H, Pagano I, Chee K, Kobayashi T, Wong RS, Lee R & Rosser CJ (2019) Comparison of Commercial ELISA Kits, a Prototype Multiplex Electrochemoluminescent Assay, and a Multiplex Bead-Based Immunoassay for Detecting a Urine-Based Bladder-Cancer-Associated Diagnostic Signature. *Diagnostics (Basel)* 9(4).

- Gallant-Behm CL, Piper J, Lynch JM, Seto AG, Hong SJ, Mustoe TA, Maari C, Pestano LA, Dalby CM, Jackson AL, Rubin P & Marshall WS (2019) A MicroRNA-29 Mimic (Remlarsen) Represses Extracellular Matrix Expression and Fibroplasia in the Skin. *The Journal of investigative dermatology* 139(5):1073-1081.
- Gao Y & Zhao Y (2014) Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. *Journal of integrative plant biology* 56(4):343-349.
- Gardner PP & Giegerich R (2004) A comprehensive comparison of comparative RNA structure prediction approaches. *BMC bioinformatics* 5:140.
- Garrelfs SF, Frishberg Y, Hulton SA, Koren MJ, O'Riordan WD, Cochat P, Deschenes G, Shashalavsky H, Saland JM, Van't Hoff WG, Fuster DG, Magen D, Mochhala SH, Schalk G, Simkova E, Groothoff JW, Sas DJ, Meliambro KA, Lu J, Sweetser MT, Garg PP, Vaishnav AK, Gansner JM, McGregor TL & Lieske JC (2021) Lumasiran, an RNAi Therapeutic for Primary Hyperoxaluria Type 1. *The New England journal of medicine* 384(13):1216-1226.
- Garst AD, Edwards AL & Batey RT (2011) Riboswitches: structures and mechanisms. *Cold Spring Harbor perspectives in biology* 3(6).
- Ghaem Maghami M, Scheitl CPM & Hobartner C (2019) Direct in Vitro Selection of Trans-Acting Ribozymes for Posttranscriptional, Site-Specific, and Covalent Fluorescent Labeling of RNA. *Journal of the American Chemical Society* 141(50):19546-19549.
- Gherman A, Balacescu L, Gheorghe-Cetean S, Vlad C, Balacescu O, Irimie A & Lisencu C (2020) Current and New Predictors for Treatment Response in Metastatic Colorectal Cancer. The Role of Circulating miRNAs as Biomarkers. *International journal of molecular sciences* 21(6).
- Gilbert SD, Rambo RP, Van Tyne D & Batey RT (2008) Structure of the SAM-II riboswitch bound to S-adenosylmethionine. *Nature structural & molecular biology* 15(2):177-182.
- Girotti MR, Salatino M, Dalotto-Moreno T & Rabinovich GA (2020) Sweetening the hallmarks of cancer: Galectins as multifunctional mediators of tumor progression. *The Journal of experimental medicine* 217(2).
- Goel HL & Mercurio AM (2013) VEGF targets the tumour cell. *Nature reviews. Cancer* 13(12):871-882.
- Gold L, Polisky B, Uhlenbeck O & Yarus M (1995) Diversity of oligonucleotide functions. *Annual review of biochemistry* 64:763-797.
- Golhani V, Ray SK & Mukherjee S (2021) Role of MicroRNAs and Long Non-Coding RNAs in Regulating Angiogenesis in Human Breast Cancer- A Molecular Medicine Perspective. *Current molecular medicine* 10.2174/1566524022666211217114527.
- Gong L, Zhao Z, Lv YF, Huan SY, Fu T, Zhang XB, Shen GL & Yu RQ (2015) DNAzyme-based biosensors and nanodevices. *Chem Commun (Camb)* 51(6):979-995.
- Goni-Moreno A & Nikel PI (2019) High-Performance Biocomputing in Synthetic Biology- Integrated Transcriptional and Metabolic Circuits. *Frontiers in bioengineering and biotechnology* 7:40.
- Grabow WW & Jaeger L (2014) RNA self-assembly and RNA nanotechnology. *Accounts of chemical research* 47(6):1871-1880.
- Green AA, Silver PA, Collins JJ & Yin P (2014) Toehold switches: de-novo-designed regulators of gene expression. *Cell* 159(4):925-939.
- Green LS, Jellinek D, Bell C, Beebe LA, Feistner BD, Gill SC, Jucker FM & Janjic N (1995) Nuclease-resistant nucleic acid ligands to vascular permeability factor/vascular endothelial growth factor. *Chemistry & biology* 2(10):683-695.

- Groenen PM, Bunschoten AE, van Soolingen D & van Embden JD (1993) Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method. *Molecular microbiology* 10(5):1057-1065.
- Groher F & Suess B (2014) Synthetic riboswitches - A tool comes of age. *Biochimica et biophysica acta* 1839(10):964-973.
- Gu HZ, Furukawa K & Breaker RR (2012) Engineered Allosteric Ribozymes That Sense the Bacterial Second Messenger Cyclic Diguanosyl 5'-Monophosphate. *Analytical chemistry* 84(11):4935-4941.
- Guerrier-Takada C, Gardiner K, Marsh T, Pace N & Altman S (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* 35(3 Pt 2):849-857.
- Gulbakan B, Barylyuk K, Schneider P, Pillong M, Schneider G & Zenobi R (2018) Native Electrospray Ionization Mass Spectrometry Reveals Multiple Facets of Aptamer-Ligand Interactions: From Mechanism to Binding Constants. *Journal of the American Chemical Society* 140(24):7486-7497.
- Gupta A, Rahman R, Li K & Gribskov M (2012) Identifying complete RNA structural ensembles including pseudoknots. *RNA biology* 9(2):187-199.
- Hajdin CE, Bellaousov S, Huggins W, Leonard CW, Mathews DH & Weeks KM (2013) Accurate SHAPE-directed RNA secondary structure modeling, including pseudoknots. *Proceedings of the National Academy of Sciences of the United States of America* 110(14):5498-5503.
- Hall B, Hesselberth JR & Ellington AD (2007) Computational selection of nucleic acid biosensors via a slip structure model. *Biosensors & bioelectronics* 22(9-10):1939-1947.
- Hamedani NS & Muller J (2016) Capillary Electrophoresis for the Selection of DNA Aptamers Recognizing Activated Protein C. *Methods Mol Biol* 1380:61-75.
- Hamann C, Luptak A, Perreault J & de la Pena M (2012) The ubiquitous hammerhead ribozyme. *RNA* 18(5):871-885.
- Hampel A (1998) The hairpin ribozyme: discovery, two-dimensional model, and development for gene therapy. *Progress in nucleic acid research and molecular biology* 58:1-39.
- Hampel A & Tritz R (1989) RNA catalytic properties of the minimum (-)STRSV sequence. *Biochemistry* 28(12):4929-4933.
- Hamula CLA, Zhang H, Li F, Wang Z, Chris Le X & Li XF (2011) Selection and analytical applications of aptamers binding microbial pathogens. *Trends in analytical chemistry : TRAC* 30(10):1587-1597.
- Han C, Li Q, Ji H, Xing W, Zhang L & Zhang L (2021a) Aptamers: The Powerful Molecular Tools for Virus Detection. *Chemistry, an Asian journal* 16(11):1298-1306.
- Han J, Lee Y, Yeom KH, Nam JW, Heo I, Rhee JK, Sohn SY, Cho Y, Zhang BT & Kim VN (2006) Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* 125(5):887-901.
- Han KS & Byun Y (2003) PseudoViewer2: visualization of RNA pseudoknots of any type. *Nucleic acids research* 31(13):3432-3440.
- Han SR, Lee CH, Im JY, Kim JH, Kim JH, Kim SJ, Cho YW, Kim E, Kim Y, Ryu JH, Ju MH, Jeong JS & Lee SW (2021b) Targeted suicide gene therapy for liver cancer based on ribozyme-mediated RNA replacement through post-transcriptional regulation. *Molecular therapy. Nucleic acids* 23:154-168.
- Harcourt EM, Kietrys AM & Kool ET (2017) Chemical and structural effects of base modifications in messenger RNA. *Nature* 541(7637):339-346.
- Harris KA, Lunse CE, Li S, Brewer KI & Breaker RR (2015) Biochemical analysis of pistol self-cleaving ribozymes. *RNA* 21(11):1852-1858.

- Hartig JS, Grune I, Najafi-Shoushtari SH & Famulok M (2004) Sequence-specific detection of MicroRNAs by signal-amplifying ribozymes. *Journal of the American Chemical Society* 126(3):722-723.
- Haseloff J & Gerlach WL (1988) Simple RNA enzymes with new and highly specific endoribonuclease activities. *Nature* 334(6183):585-591.
- Haseloff J & Gerlach WL (1989) Sequences required for self-catalysed cleavage of the satellite RNA of tobacco ringspot virus. *Gene* 82(1):43-52.
- Hassan EM, Dixon BR, Sattar SA, Stalker A, Ormeci B & DeRosa MC (2021) Highly sensitive magnetic-microparticle-based aptasensor for *Cryptosporidium parvum* oocyst detection in river water and wastewater: Effect of truncation on aptamer affinity. *Talanta* 222:121618.
- He P, Zhu D, Hu JJ, Peng J, Chen LS & Lu GX (2010) pcDNA3.1(-)-mediated ribozyme targeting of HER-2 suppresses breast cancer tumor growth. *Molecular biology reports* 37(3):1597-1604.
- Heigwer F, Kerr G, Walther N, Glaeser K, Pelz O, Breinig M & Boutros M (2013) E-TALEN: a web tool to design TALENs for genome engineering. *Nucleic acids research* 41(20):e190.
- Helm M, Petermeier M, Ge B, Fiammengo R & Jäschke A (2005) Allosterically activated Diels-Alder catalysis by a ribozyme. *Journal of the American Chemical Society* 127(30):10492-10493.
- Hoagland MB, Stephenson ML, Scott JF, Hecht LI & Zamecnik PC (1958) A soluble ribonucleic acid intermediate in protein synthesis. *The Journal of biological chemistry* 231(1):241-257.
- Hofacker IL (2003a) Vienna RNA secondary structure server. *Nucleic acids research* 31(13):3429-3431.
- Hofacker IL (2003b) Vienna RNA secondary structure server. *Nucleic acids research* 31:3429--3431.
- Hofacker IL (2004) RNA secondary structure analysis using the Vienna RNA package. *Current protocols in bioinformatics / editorial board, Andreas D. Baxevanis ... [et al.]* Chapter 12:Unit 12 12.
- Hoinka J, Backofen R & Przytycka TM (2018) AptaSUITE: A Full-Featured Bioinformatics Framework for the Comprehensive Analysis of Aptamers from HT-SELEX Experiments. *Molecular therapy. Nucleic acids* 11:515-517.
- Hoinka J & Przytycka T (2016) AptaPLEX - A dedicated, multithreaded demultiplexer for HT-SELEX data. *Methods* 106:82-85.
- Honer zu Siederdisen C, Hammer S, Abfalter I, Hofacker IL, Flamm C & Stadler PF (2013) Computational design of RNAs with complex energy landscapes. *Biopolymers* 99(12):1124-1136.
- Honor A, Rudnick SR & Bonkovsky HL (2021) Givosiran to treat acute porphyria. *Drugs Today (Barc)* 57(1):47-59.
- Hori SI, Herrera A, Rossi JJ & Zhou J (2018) Current Advances in Aptamers for Cancer Diagnosis and Therapy. *Cancers* 10(1).
- Hu B, Zhong L, Weng Y, Peng L, Huang Y, Zhao Y & Liang XJ (2020) Therapeutic siRNA: state of the art. *Signal transduction and targeted therapy* 5(1):101.
- Hu Z, Ding W, Zhu D, Yu L, Jiang X, Wang X, Zhang C, Wang L, Ji T, Liu D, He D, Xia X, Zhu T, Wei J, Wu P, Wang C, Xi L, Gao Q, Chen G, Liu R, Li K, Li S, Wang S, Zhou J, Ma D & Wang H (2015) TALEN-mediated targeting of HPV oncogenes ameliorates HPV-related cervical malignancy. *The Journal of clinical investigation* 125(1):425-436.

- Huang FW, Qin J, Reidys CM & Stadler PF (2009) Partition function and base pairing probabilities for RNA-RNA interaction prediction. *Bioinformatics* 25(20):2646-2654.
- Huang J, Li K & Gribskov M (2016) Accurate Classification of RNA Structures Using Topological Fingerprints. *PLoS one* 11(10):e0164726.
- Huang X, Zhao Y, Pu Q, Liu G, Peng Y, Wang F, Chen G, Sun M, Du F, Dong J, Cui X, Tang Z & Mo X (2019) Intracellular selection of trans-cleaving hammerhead ribozymes. *Nucleic acids research* 47(5):2514-2522.
- Inokuchi Y, Yuyama N, Hirashima A, Nishikawa S, Ohkawa J & Taira K (1994) A hammerhead ribozyme inhibits the proliferation of an RNA coliphage SP in Escherichia coli. *The Journal of biological chemistry* 269(15):11361-11366.
- Isaacs FJ, Dwyer DJ & Collins JJ (2006) RNA synthetic biology. *Nature biotechnology* 24(5):545-554.
- Ishida M & Selaru FM (2013) miRNA-Based Therapeutic Strategies. *Current anesthesiology reports* 1(1):63-70.
- Ishida R, Adachi T, Yokota A, Yoshihara H, Aoki K, Nakamura Y & Hamada M (2020) RaptRanker: in silico RNA aptamer selection from HT-SELEX experiment based on local sequence and structure information. *Nucleic acids research* 48(14):e82.
- Jacek Nowakowski and Ignacio Tinoco J (1997) RNA Structure and Stability. *Seminars in VIROLOGY* 8:153–165.
- Jacob F & Monod J (1961) Genetic regulatory mechanisms in the synthesis of proteins. *Journal of molecular biology* 3:318-356.
- Jaeger JA, SantaLucia J, Jr. & Tinoco I, Jr. (1993) Determination of RNA structure and thermodynamics. *Annual review of biochemistry* 62:255-287.
- Jaffrey SR (2018) RNA-Based Fluorescent Biosensors for Detecting Metabolites in vitro and in Living Cells. *Adv Pharmacol* 82:187-203.
- Jeffries AC & Symons RH (1989) A catalytic 13-mer ribozyme. *Nucleic acids research* 17(4):1371-1377.
- Jeng KS, Daniel A & Lai MM (1996) A pseudoknot ribozyme structure is active in vivo and required for hepatitis delta virus RNA replication. *Journal of virology* 70(4):2403-2410.
- Jiang N, Ding X & Lu Y (2021a) Development of a robust Escherichia coli-based cell-free protein synthesis application platform. *Biochemical engineering journal* 165:107830.
- Jiang Q, Yue S, Yu K, Tian T, Zhang J, Chu H, Cui Z & Bi S (2021b) Endogenous microRNA triggered enzyme-free DNA logic self-assembly for amplified bioimaging and enhanced gene therapy via in situ generation of siRNAs. *Journal of nanobiotechnology* 19(1):288.
- Jiang Y, Pan X, Chang J, Niu W, Hou W, Kuai H, Zhao Z, Liu J, Wang M & Tan W (2018) Supramolecularly Engineered Circular Bivalent Aptamer for Enhanced Functional Protein Delivery. *Journal of the American Chemical Society* 140(22):6780-6784.
- Joseph NN, Roy RN & Steitz TA (2020) Molecular dynamics analysis of Mg(2+) -dependent cleavage of a pistol ribozyme reveals a fail-safe secondary ion for catalysis. *Journal of computational chemistry* 41(14):1345-1352.
- Joseph S & Burke JM (1993) Optimization of an anti-HIV hairpin ribozyme by in vitro selection. *The Journal of biological chemistry* 268(33):24515-24518.
- Kachris S, Papadaki C, Rounis K, Tsitoura E, Kokkinaki C, Nikolaou C, Sourvinos G & Mavroudis D (2021) Circulating miRNAs as Potential Biomarkers in Prostate Cancer Patients Undergoing Radiotherapy. *Cancer management and research* 13:8257-8271.
- Karamperis K, Tsoumpeli MT, Kounelis F, Koromina M, Mitropoulou C, Moutinho C & Patrinos GP (2021) Genome-based therapeutic interventions for beta-type hemoglobinopathies. *Human genomics* 15(1):32.

- Karlsen KK & Wengel J (2012) Locked nucleic acid and aptamers. *Nucleic acid therapeutics* 22(6):366-370.
- Kaur H, Bruno JG, Kumar A & Sharma TK (2018) Aptamers in the Therapeutics and Diagnostics Pipelines. *Theranostics* 8(15):4016-4032.
- Kaur M, Kaur T, Kamboj SS & Singh J (2016) Roles of Galectin-7 in Cancer. *Asian Pacific journal of cancer prevention : APJCP* 17(2):455-461.
- Kertsburg A & Soukup GA (2002) A versatile communication module for controlling RNA folding and catalysis. *Nucleic acids research* 30(21):4599-4606.
- Ketzer P, Haas SF, Engelhardt S, Hartig JS & Nettelbeck DM (2012) Synthetic riboswitches for external regulation of genes transferred by replication-deficient and oncolytic adenoviruses. *Nucleic acids research* 40(21):e167.
- Ketzer P, Kaufmann JK, Engelhardt S, Bossow S, von Kalle C, Hartig JS, Ungerechts G & Nettelbeck DM (2014) Artificial riboswitches for gene expression and replication control of DNA and RNA viruses. *Proceedings of the National Academy of Sciences of the United States of America* 111(5):E554-562.
- Khan S, Burciu B, Filipe CDM, Li Y, Dellinger K & Didar TF (2021) DNAzyme-Based Biosensors: Immobilization Strategies, Applications, and Future Prospective. *ACS nano* 15(9):13943-13969.
- Khanali J, Azangou-Khyavy M, Asaadi Y, Jamalkhah M & Kiani J (2021) Nucleic Acid-Based Treatments Against COVID-19: Potential Efficacy of Aptamers and siRNAs. *Frontiers in microbiology* 12:758948.
- Kharmar N, Varin L, Abu-Baker A, Ouellet J, Najeh S, Ehdaevand MR, Belmonte G, Ambri A, Rouleau G & Perreault J (2016) Automated design of hammerhead ribozymes and validation by targeting the PABPN1 gene transcript. *Nucleic acids research* 44(4):e39.
- Khorana HG, Buchi H, Ghosh H, Gupta N, Jacob TM, Kossel H, Morgan R, Narang SA, Ohtsuka E & Wells RD (1966) Polynucleotide synthesis and the genetic code. *Cold Spring Harbor symposia on quantitative biology* 31:39-49.
- Khvorova A, Lescoute A, Westhof E & Jayasena SD (2003) Sequence elements outside the hammerhead ribozyme catalytic core enable intracellular activity. *Nature structural biology* 10(9):708-712.
- Kilgour M, Liu T, Walker BD, Ren P & Simine L (2021) E2EDNA: Simulation Protocol for DNA Aptamers with Ligands. *Journal of chemical information and modeling* 61(9):4139-4144.
- Kim B, Park JH & Sailor MJ (2019) Rekindling RNAi Therapy: Materials Design Requirements for In Vivo siRNA Delivery. *Adv Mater* 31(49):e1903637.
- Kim J & Franco E (2020) RNA nanotechnology in synthetic biology. *Current opinion in biotechnology* 63:135-141.
- Kim YW, Bae SM, Kim IW, Liu HB, Bang HJ, Chaturvedi PK, Battogtokh G, Lim H & Ahn WS (2012) Multiplexed bead-based immunoassay of four serum biomarkers for diagnosis of ovarian cancer. *Oncology reports* 28(2):585-591.
- Kiryu H, Terai G, Imamura O, Yoneyama H, Suzuki K & Asai K (2011) A detailed investigation of accessibilities around target sites of siRNAs and miRNAs. *Bioinformatics* 27(13):1788-1797.
- Klein DJ & Ferre-D'Amare AR (2006) Structural basis of glmS ribozyme activation by glucosamine-6-phosphate. *Science* 313(5794):1752-1756.
- Klein DJ & Ferre-D'Amare AR (2006) Structural basis of glmS ribozyme activation by glucosamine-6-phosphate. *Science* 313:1752--1756.
- Kleinkauf R, Houwaart T, Backofen R & Mann M (2015) antaRNA--Multi-objective inverse folding of pseudoknot RNA using ant-colony optimization. *BMC bioinformatics* 16:389.

- Kleinkauf RaH, Torsten and Backofen, Rolf and Mann, Martin (2015) antaRNA--Multi-objective inverse folding of pseudoknot RNA using ant-colony optimization. *BMC bioinformatics* 16:389.
- Kloss CC, Condomines M, Cartellieri M, Bachmann M & Sadelain M (2013) Combinatorial antigen recognition with balanced signaling promotes selective tumor eradication by engineered T cells. *Nature biotechnology* 31(1):71-75.
- Kobori S, Takahashi K & Yokobayashi Y (2017) Deep Sequencing Analysis of Aptazyme Variants Based on a Pistol Ribozyme. *ACS synthetic biology* 6(7):1283-1288.
- Koizumi M, Hayase Y, Iwai S, Kamiya H, Inoue H & Ohtsuka E (1989) Design of RNA enzymes distinguishing a single base mutation in RNA. *Nucleic acids research* 17(17):7059-7071.
- Koizumi M, Kerr JN, Soukup GA & Breaker RR (1999a) Allosteric ribozymes sensitive to the second messengers cAMP and cGMP. *Nucleic acids symposium series* 10.1093/nass/42.1.275(42):275-276.
- Koizumi M, Soukup GA, Kerr JNQ & Breaker RR (1999b) Allosteric selection of ribozymes that respond to the second messengers cGMP and cAMP. *Nature structural biology* 6(11):1062-1071.
- Komarova N, Barkova D & Kuznetsov A (2020) Implementation of High-Throughput Sequencing (HTS) in Aptamer Selection Technology. *International journal of molecular sciences* 21(22).
- Koo KM, Mainwaring PN, Tomlins SA & Trau M (2019) Merging new-age biomarkers and nanodiagnosics for precision prostate cancer management. *Nature reviews. Urology* 16(5):302-317.
- Krasitskaya VV, Chaukina VV, Abroskina MV, Vorobyeva MA, Ilminskaya AA, Kabilov MR, Prokopenko SV, Nevinsky GA, Venyaminova AG & Frank LA (2019) Bioluminescent aptamer-based sandwich-type assay of anti-myelin basic protein autoantibodies associated with multiple sclerosis. *Analytica chimica acta* 1064:112-118.
- Kratschmer C & Levy M (2017) Effect of Chemical Modifications on Aptamer Stability in Serum. *Nucleic acid therapeutics* 27(6):335-344.
- Kruger A, de Jesus Santos AP, de Sa V, Ulrich H & Wrenger C (2021) Aptamer Applications in Emerging Viral Diseases. *Pharmaceuticals (Basel)* 14(7).
- Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschling DE & Cech TR (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* 31(1):147-157.
- Kukova LZ, Mansour SG, Coca SG, de Fontnouvelle CA, Thiessen-Philbrook HR, Shlipak MG, El-Khoury JM & Parikh CR (2019) Comparison of Urine and Plasma Biomarker Concentrations Measured by Aptamer-Based versus Immunoassay Methods in Cardiac Surgery Patients. *The journal of applied laboratory medicine* 4(3):331-342.
- Kuwabara T, Warashina M, Nakayama A, Ohkawa J & Taira K (1999) tRNAVal-heterodimeric maxizymes with high potential as geneinactivating agents: simultaneous cleavage at two sites in HIV-1 Tat mRNA in cultured cells. *Proceedings of the National Academy of Sciences of the United States of America* 96(5):1886-1891.
- Kuwabara T, Warashina M & Taira K (2000) Allosterically controllable ribozymes with biosensor functions. *Current opinion in chemical biology* 4(6):669-677.
- Labib M, Zamay AS, Kolovskaya OS, Reshetneva IT, Zamay GS, Kibbee RJ, Sattar SA, Zamay TN & Berezovski MV (2012) Aptamer-based viability impedimetric sensor for bacteria. *Analytical chemistry* 84(21):8966-8969.

- Lacroix-Labonte J, Girard N, Dagenais P & Legault P (2016) Rational engineering of the Neurospora VS ribozyme to allow substrate recognition via different kissing-loop interactions. *Nucleic acids research* 44(14):6924-6934.
- Lagana A, Shasha D & Croce CM (2014) Synthetic RNAs for Gene Regulation: Design Principles and Computational Tools. *Frontiers in bioengineering and biotechnology* 2:65.
- Lai JC, Horng HE & Hong CY (2017) Multiplex Immunoassays Utilizing Differential Affinity Using Aptamers Generated by MARAS. *Scientific reports* 7(1):6397.
- Lai YC, Lee JY, Liu HJ, Lin JY & Wu HN (1996) Effects of circular permutation on the cis-cleavage reaction of a hepatitis delta virus ribozyme: application to trans-acting ribozyme design. *Biochemistry* 35(1):124-131.
- Lakhin AV, Tarantul VZ & Gening LV (2013) Aptamers: problems, solutions and prospects. *Acta naturae* 5(4):34-43.
- Lamfrom H, McLaughlin CS & Sarabhai A (1966) Direction of reading the genetic message in reticulocytes. *Journal of molecular biology* 22(2):355-358.
- Langlois M-A, Lee NS, Rossi JJ & Puymirat J (2003) Hammerhead ribozyme-mediated destruction of nuclear foci in myotonic dystrophy myoblasts. *Molecular Therapy* 7(5):670-680.
- Lau MW & Ferre-D'Amare AR (2013) An in vitro evolved glmS ribozyme has the wild-type fold but loses coenzyme dependence. *Nature chemical biology* 9(12):805-810.
- Lau MW & Ferre-D'Amare AR (2016) In vitro evolution of coenzyme-independent variants from the glmS ribozyme structural scaffold. *Methods* 106:76-81.
- Lau MW, Trachman RJ, 3rd & Ferre-D'Amare AR (2017) A divalent cation-dependent variant of the glmS ribozyme with stringent Ca(2+) selectivity co-opts a preexisting nonspecific metal ion-binding site. *RNA* 23(3):355-364.
- Lau MWaF-DA, Adrian R (2013) An in vitro evolved glmS ribozyme has the wild-type fold but loses coenzyme dependence. *Nature chemical biology* 9:805--810.
- Lauridsen LH, Shamaileh HA, Edwards SL, Taran E & Veedu RN (2012) Rapid one-step selection method for generating nucleic acid aptamers: development of a DNA aptamer against alpha-bungarotoxin. *PLoS one* 7(7):e41702.
- Lee ER, Baker JL, Weinberg Z, Sudarsan N & Breaker RR (2010) An allosteric self-splicing ribozyme triggered by a bacterial second messenger. *Science* 329(5993):845-848.
- Lee RC, Feinbaum RL & Ambros V (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* 75(5):843-854.
- Lee SJ, Kwon YS, Lee JE, Choi EJ, Lee CH, Song JY & Gu MB (2013) Detection of VR-2332 strain of porcine reproductive and respiratory syndrome virus type II using an aptamer-based sandwich-type assay. *Analytical chemistry* 85(1):66-74.
- Lee SWL, Paoletti C, Campisi M, Osaki T, Adriani G, Kamm RD, Mattu C & Chiono V (2019) MicroRNA delivery through nanoparticles. *Journal of controlled release : official journal of the Controlled Release Society* 313:80-95.
- Lehr FX, Kuzembayeva A, Bailey ME, Kleindienst W, Kabisch J & Koepl H (2021) Functionalizing Cell-Free Systems with CRISPR-Associated Proteins: Application to RNA-Based Circuit Engineering. *ACS synthetic biology* 10(9):2138-2150.
- Leontis NB, Lescoute A & Westhof E (2006) The building blocks and motifs of RNA architecture. *Current opinion in structural biology* 16(3):279-287.
- Levesque D, Choufani S & Perreault JP (2002) Delta ribozyme benefits from a good stability in vitro that becomes outstanding in vivo. *RNA* 8(4):464-477.
- Lévesque MV, Lévesque D, Brière FP & Perreault J-P (2010) Investigating a new generation of ribozymes in order to target HCV. *PLoS One* 5(3):e9627.

- Levi-Acobas F, Katolik A, Rothlisberger P, Cokelaer T, Sarac I, Damha MJ, Leumann CJ & Hollenstein M (2019) Compatibility of 5-ethynyl-2'-F-ANA UTP with in vitro selection for the generation of base-modified, nuclease resistant aptamers. *Organic & biomolecular chemistry* 17(35):8083-8087.
- Levin A, Lis M, Ponty Y, O'Donnell CW, Devadas S, Berger B & Waldispuhl J (2012) A global sampling approach to designing and reengineering RNA secondary structures. *Nucleic acids research* 40(20):10041-10052.
- Li CH, Chang YC, Chan MH, Yang YF, Liang SM & Hsiao M (2021a) Galectins in Cancer and the Microenvironment: Functional Roles, Therapeutic Developments, and Perspectives. *Biomedicines* 9(9).
- Li D, Liu L, Huang Q, Tong T, Zhou Y, Li Z, Bai Q, Liang H & Chen L (2021b) Recent advances on aptamer-based biosensors for detection of pathogenic bacteria. *World J Microbiol Biotechnol* 37(3):45.
- Li H, Yang Y, Hong W, Huang M, Wu M & Zhao X (2020) Applications of genome editing technology in the targeted therapy of human diseases: mechanisms, advances and prospects. *Signal transduction and targeted therapy* 5(1):1.
- Li L, Ren Y, Wen X, Guo Q, Wang J, Li S, Yang M & Wang K (2021c) Endogenous miRNA-Activated DNA Nanomachine for Intracellular miRNA Imaging and Gene Silencing. *Analytical chemistry* 93(41):13919-13927.
- Li S & Breaker RR (2013) Eukaryotic TPP riboswitch regulation of alternative splicing involving long-distance base pairing. *Nucleic acids research* 41(5):3022-3031.
- Li X, Zhao Q & Qiu L (2013) Smart ligand: aptamer-mediated targeted delivery of chemotherapeutic drugs and siRNA for cancer therapy. *Journal of controlled release : official journal of the Controlled Release Society* 171(2):152-162.
- Liang JC, Bloom RJ & Smolke CD (2011) Engineering biological systems with synthetic RNA molecules. *Molecular cell* 43(6):915-926.
- Lieber A, Rohde M & Strauss M (1997) Selection of efficient ribozyme cleavage sites in target RNAs. *Methods Mol Biol* 74:45-50.
- Lieber A & 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.
- Lienert F, Lohmueller JJ, Garg A & Silver PA (2014) Synthetic biology in mammalian cells: next generation research tools and therapeutics. *Nature reviews. Molecular cell biology* 15(2):95-107.
- Ling J, Wiederkehr U, Cabiness S, Shroyer KR & Robinson JP (2008) Application of flow cytometry for biomarker-based cervical cancer cells detection. *Diagnostic cytopathology* 36(2):76-84.
- Link KH, Guo L & Breaker RR (2006) Examination of the structural and functional versatility of glmS ribozymes by using in vitro selection. *Nucleic acids research* 34(17):4968-4975.
- Liu L, Liu P, Ga L & Ai J (2021) Advances in Applications of Molecular Logic Gates. *ACS omega* 6(45):30189-30204.
- Liu X & Chu KM (2020) Exosomal miRNAs as circulating biomarkers for prediction of development of haematogenous metastasis after surgery for stage II/III gastric cancer. *Journal of cellular and molecular medicine* 24(11):6220-6232.
- Liu Y, Huang W & Cai Z (2020) Synthesizing AND gate minigene circuits based on CRISPRReader for identification of bladder cancer cells. *Nature communications* 11(1):5486.
- Lo Y, Cheung YW, Wang L, Lee M, Figueroa-Miranda G, Liang S, Mayer D & Tanner JA (2021) An electrochemical aptamer-based biosensor targeting Plasmodium falciparum histidine-rich protein II for malaria diagnosis. *Biosensors & bioelectronics* 192:113472.

- Lorenz R, Bernhart SH, Zu Siederdisen CH, Tafer H, Flamm C, Stadler PF & Hofacker IL (2011) ViennaRNA Package 2.0. *Algorithms for molecular biology* 6(1):26.
- Lorenz WA & Clote P (2011) Computing the partition function for kinetically trapped RNA secondary structures. *PloS one* 6(1):e16178.
- Lorsch JR & Szostak JW (1994) In vitro evolution of new ribozymes with polynucleotide kinase activity. *Nature* 371(6492):31-36.
- Lu RM, Hwang YC, Liu IJ, Lee CC, Tsai HZ, Li HJ & Wu HC (2020) Development of therapeutic antibodies for the treatment of diseases. *Journal of biomedical science* 27(1):1.
- Lu Z & Chang HY (2016) Decoding the RNA structurome. *Current opinion in structural biology* 36:142-148.
- Lucier JF, Bergeron LJ, Briere FP, Ouellette R, Abou Elela S & Perreault JP (2006a) RiboSubstrates: a web application addressing the cleavage specificities of ribozymes in designated genomes. *BMC bioinformatics* 7.
- Lucier JF, Bergeron LJ, Briere FP, Ouellette R, Elela SA & Perreault JP (2006b) RiboSubstrates: a web application addressing the cleavage specificities of ribozymes in designated genomes. *BMC bioinformatics* 7:480.
- Luck S, Kreszies T, Strickert M, Schweizer P, Kuhlmann M & Douchkov D (2019) siRNA-Finder (si-Fi) Software for RNAi-Target Design and Off-Target Prediction. *Frontiers in plant science* 10:1023.
- Lyngso RB, Anderson JW, Sizikova E, Badugu A, Hyland T & Hein J (2012) Frnakenstein: multiple target inverse RNA folding. *BMC bioinformatics* 13:260.
- Macfarlane LA & Murphy PR (2010) MicroRNA: Biogenesis, Function and Role in Cancer. *Current genomics* 11(7):537-561.
- Maeder ML, Stefanidakis M, Wilson CJ, Baral R, Barrera LA, Bounoutas GS, Bumcrot D, Chao H, Ciulla DM, DaSilva JA, Dass A, Dhanapal V, Fennell TJ, Friedland AE, Giannoukos G, Gloskowski SW, Glucksmann A, Gotta GM, Jayaram H, Haskett SJ, Hopkins B, Horng JE, Joshi S, Marco E, Mepani R, Reyon D, Ta T, Tabbaa DG, Samuelsson SJ, Shen S, Skor MN, Stetkiewicz P, Wang T, Yudkoff C, Myer VE, Albright CF & Jiang H (2019) Development of a gene-editing approach to restore vision loss in Leber congenital amaurosis type 10. *Nature medicine* 25(2):229-233.
- Magde D, Akoopie A, Magde MD, Jr. & Muller UF (2021) Water/Oil Emulsions with Controlled Droplet Sizes for In Vitro Selection Experiments. *ACS omega* 6(33):21773-21783.
- Maguregui A & Abe H (2020) Developments in siRNA Modification and Ligand Conjugated Delivery To Enhance RNA Interference Ability. *Chembiochem : a European journal of chemical biology* 21(13):1808-1815.
- Mandal M, Dutta N & Dutta G (2021) Aptamer-based biosensors and their implications in COVID-19 diagnosis. *Analytical methods : advancing methods and applications* 13(45):5400-5417.
- Mandell JG & Barbas CF, 3rd (2006) Zinc Finger Tools: custom DNA-binding domains for transcription factors and nucleases. *Nucleic acids research* 34(Web Server issue):W516-523.
- Mandic A, Hayes RL, Lammert H, Cheng RR & Onuchic JN (2019) Structure-Based Model of RNA Pseudoknot Captures Magnesium-Dependent Folding Thermodynamics. *The journal of physical chemistry. B* 123(7):1505-1511.
- Martick M, Horan LH, Noller HF & Scott WG (2008) A discontinuous hammerhead ribozyme embedded in a mammalian messenger RNA. *Nature* 454(7206):899-902.
- Martick M & Scott WG (2006) Tertiary contacts distant from the active site prime a ribozyme for catalysis. *Cell* 126(2):309-320.

- Mathew V & Wang AK (2019) Inotersen: new promise for the treatment of hereditary transthyretin amyloidosis. *Drug design, development and therapy* 13:1515-1525.
- Mathews DH (2004) Using an RNA secondary structure partition function to determine confidence in base pairs predicted by free energy minimization. *RNA* 10(8):1178-1190.
- Mathews DH, Sabina J, Zuker M & Turner DH (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *Journal of molecular biology* 288(5):911-940.
- Mathews DH & Turner DH (2006) Prediction of RNA secondary structure by free energy minimization. *Current opinion in structural biology* 16(3):270-278.
- Matsuura S, Ono H, Kawasaki S, Kuang Y, Fujita Y & Saito H (2018) Synthetic RNA-based logic computation in mammalian cells. *Nature communications* 9(1):4847.
- Mayo ML, Eberly JO, Crocker FH & Indest KJ (2020) Modeling a synthetic aptamer-based riboswitch biosensor sensitive to low hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) concentrations. *PloS one* 15(11):e0241664.
- McCarthy TJ, Plog MA, Floy SA, Jansen JA, Soukup JK & Soukup GA (2005) Ligand requirements for glmS ribozyme self-cleavage. *Chemistry & biology* 12(11):1221-1226.
- McCaskill JS (1990) The equilibrium partition function and base pair binding probabilities for RNA secondary structure. *Biopolymers* 29(6-7):1105-1119.
- McConnell EM, Cozma I, Morrison D & Li Y (2020a) Biosensors Made of Synthetic Functional Nucleic Acids Toward Better Human Health. *Analytical chemistry* 92(1):327-344.
- McConnell EM, Nguyen J & Li Y (2020b) Aptamer-Based Biosensors for Environmental Monitoring. *Frontiers in chemistry* 8:434.
- McCown PJ, Roth A & Breaker RR (2011) An expanded collection and refined consensus model of glmS ribozymes. *RNA* 17(4):728-736.
- McCown PJ, Ruzskowska A, Kunkler CN, Breger K, Hulewicz JP, Wang MC, Springer NA & Brown JA (2020) Naturally occurring modified ribonucleosides. *Wiley interdisciplinary reviews. RNA* 11(5):e1595.
- McKay DB (1996) Structure and function of the hammerhead ribozyme: an unfinished story. *RNA* 2(5):395-403.
- Mercatanti A, Lande C & Citti L (2012) A computational approach to predict suitable target sites for trans-acting minimal hammerhead ribozymes. *Methods Mol Biol* 848:337-356.
- Miao P & Tang Y (2021) Cascade Strand Displacement and Bipedal Walking Based DNA Logic System for miRNA Diagnostics. *ACS central science* 7(6):1036-1044.
- Milani P, Mussinelli R, Perlini S, Palladini G & Obici L (2019) An evaluation of patisiran: a viable treatment option for transthyretin-related hereditary amyloidosis. *Expert opinion on pharmacotherapy* 20(18):2223-2228.
- Miller J, McLachlan AD & Klug A (1985) Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *The EMBO journal* 4(6):1609-1614.
- Mishra P, Furey C, Balaraman V & Fraser MJ (2016) Antiviral Hammerhead Ribozymes Are Effective for Developing Transgenic Suppression of Chikungunya Virus in *Aedes aegypti* Mosquitoes. *Viruses* 8(6).
- Modh H, Scheper T & Walter JG (2018) Aptamer-Modified Magnetic Beads in Biosensing. *Sensors (Basel)* 18(4).
- Monsellato I, Garibaldi E, Cassinotti E, Baldari L, Boni L, Elmore U, Delpini R, Rosati R, Perinotti R, Alongi F, Bertocchi E, Gori S, Ruffo G, Pernazza G, Pulighe F, De Nisco C, Morpurgo E, Contardo T, Mammano E, Perna F, Menegatti B, Coratti A, Buccianti P, Balestri R, Ceccarelli C, Cavaliere D, Solaini L, Ercolani G, Traverso E, Fusco V, Torri V & Orecchia S (2020) Expression levels of circulating miRNAs as biomarkers during multimodal

- treatment of rectal cancer - TiMiSNAR-mirna: a substudy of the TiMiSNAR Trial (NCT03962088). *Trials* 21(1):678.
- Moreno-Mateos MA, Vejnar CE, Beaudoin JD, Fernandez JP, Mis EK, Khokha MK & Giraldez AJ (2015) CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nature methods* 12(10):982-988.
- Morrow PK, Murthy RK, Ensor JD, Gordon GS, Margolin KA, Elias AD, Urba WJ, Weng DE, Rugo HS & Hortobagyi GN (2012) An open-label, phase 2 trial of RPI.4610 (Angiozyme) in the treatment of metastatic breast cancer. *Cancer* 118(17):4098-4104.
- Mortimer SA, Kidwell MA & Doudna JA (2014) Insights into RNA structure and function from genome-wide studies. *Nature reviews. Genetics* 15(7):469-479.
- Mortimer SAaK, Mary Anne and Doudna, Jennifer A (2014) Insights into RNA structure and function from genome-wide studies. *Nature reviews. Genetics* 15:469.
- Muchmore SW, Sattler M, Liang H, Meadows RP, Harlan JE, Yoon HS, Nettekheim D, Chang BS, Thompson CB, Wong SL, Ng SL & Fesik SW (1996) X-ray and NMR structure of human Bcl-xL, an inhibitor of programmed cell death. *Nature* 381(6580):335-341.
- Mulhbachter J, St-Pierre P & Lafontaine DA (2010) Therapeutic applications of ribozymes and riboswitches. *Current opinion in pharmacology* 10(5):551-556.
- Mustafina K, Fukunaga K & Yokobayashi Y (2020) Design of Mammalian ON-Riboswitches Based on Tandemly Fused Aptamer and Ribozyme. *ACS synthetic biology* 9(1):19-25.
- Naghdi MR, Boutet E, Mucha C, Ouellet J & Perreault J (2020) Single Mutation in Hammerhead Ribozyme Favors Cleavage Activity with Manganese over Magnesium. *Non-coding RNA* 6(1).
- Nahvi A, Sudarsan N, Ebert MS, Zou X, Brown KL & Breaker RR (2002) Genetic control by a metabolite binding mRNA. *Chemistry & biology* 9(9):1043.
- Najafi-Shoushtari SH & Famulok M (2005) Competitive regulation of modular allosteric aptazymes by a small molecule and oligonucleotide effector. *Rna-a Publication of the Rna Society* 11(10):1514-1520.
- Najeh S, Zandi K, Djerroud S, Kharma N & Perreault J (2021) Computer-Aided Design of Active Pseudoknotted Hammerhead Ribozymes. *Methods Mol Biol* 2167:91-111.
- Nakamaye KL & Eckstein F (1994) AUA-cleaving hammerhead ribozymes: attempted selection for improved cleavage. *Biochemistry* 33(5):1271-1277.
- Nakano S, Cerrone AL & Bevilacqua PC (2003) Mechanistic characterization of the HDV genomic ribozyme: classifying the catalytic and structural metal ion sites within a multichannel reaction mechanism. *Biochemistry* 42(10):2982-2994.
- Nawrot B, Antoszczyk S, Maszewska M, Kuwabara T, Warashina M, Taira K & Stec WJ (2003) Efficient inhibition of beta-secretase gene expression in HEK293 cells by tRNAVal-driven and CTE-helicase associated hammerhead ribozymes. *European journal of biochemistry / FEBS* 270(19):3962-3970.
- Neil EE & Bisaccia EK (2019) Nusinersen: A Novel Antisense Oligonucleotide for the Treatment of Spinal Muscular Atrophy. *The journal of pediatric pharmacology and therapeutics : JPPT : the official journal of PPAG* 24(3):194-203.
- Nelson JA & Uhlenbeck OC (2008) Hammerhead redux: does the new structure fit the old biochemical data? *RNA* 14(4):605-615.
- Neveu M, Kim HJ & Benner SA (2013) The "strong" RNA world hypothesis: fifty years old. *Astrobiology* 13(4):391-403.
- Nezlin R (2016) Use of aptamers in immunoassays. *Molecular immunology* 70:149-154.

- Ng EW, Shima DT, Calias P, Cunningham ET, Jr., Guyer DR & Adamis AP (2006) Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nature reviews. Drug discovery* 5(2):123-132.
- Ng L, Li X, Wan TM, Iyer D, Sin RW, Lo OS, Foo DC & Law WL (2021) Investigation of miRNA dysregulation and association with immune cell profile during malignant transformation of colorectal cells. *European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology* 10.1016/j.ejso.2021.09.016.
- Nguyen LA, Wang J & Steitz TA (2017a) Crystal structure of Pistol, a class of self-cleaving ribozyme. *Proceedings of the National Academy of Sciences of the United States of America* 114(5):1021-1026.
- Nguyen VT, Kwon YS & Gu MB (2017b) Aptamer-based environmental biosensors for small molecule contaminants. *Current opinion in biotechnology* 45:15-23.
- Nicolas Kamel NK, Aida Abu-Baker, Guy Rouleau, Alexander Bailey (2020) Evolutionary Design and Experimental Evaluation of Selective Hammerhead Ribozymes. *2020 IEEE 20th International Conference on Bioinformatics and BioEngineering (BIBE)*.
- Nielsen AA, Der BS, Shin J, Vaidyanathan P, Paralanov V, Strychalski EA, Ross D, Densmore D & Voigt CA (2016) Genetic circuit design automation. *Science* 352(6281):aac7341.
- Nishikawa F, Chiba A, Shirai M, Fauzi H, Taira K, Kumar PK & Nishikawa S (1995) In vitro selection analysis of trans-acting HDV ribozyme. *Nucleic acids symposium series* (34):115-116.
- Nishikawa F, Kawakami J, Chiba A, Shirai M, Kumar PK & Nishikawa S (1996) Selection in vitro of trans-acting genomic human hepatitis delta virus (HDV) ribozymes. *European journal of biochemistry / FEBS* 237(3):712-718.
- Nissen P, Hansen J, Ban N, Moore PB & Steitz TA (2000) The structural basis of ribosome activity in peptide bond synthesis. *Science* 289(5481):920-930.
- Nissim L, Wu MR, Pery E, Binder-Nissim A, Suzuki HI, Stupp D, Wehrspaun C, Tabach Y, Sharp PA & Lu TK (2017) Synthetic RNA-Based Immunomodulatory Gene Circuits for Cancer Immunotherapy. *Cell* 171(5):1138-1150 e1115.
- Niu T, Liu Y, Li J, Koffas M, Du G, Alper HS & Liu L (2018) Engineering a Glucosamine-6-phosphate Responsive glmS Ribozyme Switch Enables Dynamic Control of Metabolic Flux in *Bacillus subtilis* for Overproduction of N-Acetylglucosamine. *ACS synthetic biology* 7(10):2423-2435.
- Noller HF, Hoffarth V & Zimniak L (1992) Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* 256(5062):1416-1419.
- Nomura Y, Kumar D & Yokobayashi Y (2012a) Synthetic mammalian riboswitches based on guanine aptazyme. *Chem Commun* 48(57):7215-7217.
- Nomura Y, Kumar D & Yokobayashi Y (2012b) Synthetic mammalian riboswitches based on guanine aptazyme. *Chem Commun (Camb)* 48(57):7215-7217.
- Nomura Y, Zhou L, Miu A & Yokobayashi Y (2013) Controlling mammalian gene expression by allosteric hepatitis delta virus ribozymes. *ACS synthetic biology* 2(12):684-689.
- Norris JS, Hoel B, Voeks D, Maggouta F, Dahm M, Pan W & Clawson G (2000) Design and testing of ribozymes for cancer gene therapy. *Advances in experimental medicine and biology* 465:293-301.
- Nussinov R & Jacobson AB (1980) Fast algorithm for predicting the secondary structure of single-stranded RNA. *Proceedings of the National Academy of Sciences of the United States of America* 77(11):6309-6313.

- Odeh F, Nsairat H, Alshaer W, Ismail MA, Esawi E, Qaqish B, Bawab AA & Ismail SI (2019) Aptamers Chemistry: Chemical Modifications and Conjugation Strategies. *Molecules* 25(1).
- Ogawa A (2009) Aptazyme-based biosensors using a eukaryotic cell-free translation system. *Nucleic Acids Symp Ser (Oxf)* 10.1093/nass/nrp131(53):261-262.
- Ogawa A (2011) Rational design of artificial riboswitches based on ligand-dependent modulation of internal ribosome entry in wheat germ extract and their applications as label-free biosensors. *RNA* 17(3):478-488.
- Ogawa A (2014) Rational design of artificial ON-riboswitches. *Methods Mol Biol* 1111:165-181.
- Ogawa A & Maeda M (2008) An artificial aptazyme-based riboswitch and its cascading system in *E. coli*. *ChemBiochem : a European journal of chemical biology* 9(2):206-209.
- Oliveto S, Mancino M, Manfrini N & Biffo S (2017) Role of microRNAs in translation regulation and cancer. *World journal of biological chemistry* 8(1):45-56.
- Olson KE & Muller UF (2012) An in vivo selection method to optimize trans-splicing ribozymes. *RNA* 18(3):581-589.
- Olzog VJ, Gartner C, Stadler PF, Fallmann J & Weinberg CE (2021) cyPhyRNA-seq: a genome-scale RNA-seq method to detect active self-cleaving ribozymes by capturing RNAs with 2',3' cyclic phosphates and 5' hydroxyl ends. *RNA biology* 18(sup2):818-831.
- Ospina-Villa JD, Lopez-Camarillo C, Castanon-Sanchez CA, Soto-Sanchez J, Ramirez-Moreno E & Marchat LA (2018) Advances on Aptamers against Protozoan Parasites. *Genes* 9(12).
- Ospina-Villa JD, Zamorano-Carrillo A, Castanon-Sanchez CA, Ramirez-Moreno E & Marchat LA (2016) Aptamers as a promising approach for the control of parasitic diseases. *The Brazilian journal of infectious diseases : an official publication of the Brazilian Society of Infectious Diseases* 20(6):610-618.
- Ossowski S, Schwab R & Weigel D (2008) Gene silencing in plants using artificial microRNAs and other small RNAs. *The Plant journal : for cell and molecular biology* 53(4):674-690.
- Oteng EK, Gu W & McKeague M (2020) High-efficiency enrichment enables identification of aptamers to circulating Plasmodium falciparum-infected erythrocytes. *Scientific reports* 10(1):9706.
- Padilla R & Sousa R (1999) Efficient synthesis of nucleic acids heavily modified with non-canonical ribose 2'-groups using a mutant T7 RNA polymerase (RNAP). *Nucleic acids research* 27(6):1561-1563.
- Palade GE (1955) A small particulate component of the cytoplasm. *The Journal of biophysical and biochemical cytology* 1(1):59-68.
- Palazzo AF & Lee ES (2015) Non-coding RNA: what is functional and what is junk? *Frontiers in genetics* 6:2.
- Palchetti I & Mascini M (2008) Nucleic acid biosensors for environmental pollution monitoring. *The Analyst* 133(7):846-854.
- Pan T (1997) Novel and variant ribozymes obtained through in vitro selection. *Current opinion in chemical biology* 1(1):17-25.
- Pang KC, Frith MC & Mattick JS (2006) Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. *Trends in genetics : TIG* 22(1):1-5.
- Park SV, Yang JS, Jo H, Kang B, Oh SS & Jung GY (2019a) Catalytic RNA, ribozyme, and its applications in synthetic biology. *Biotechnology advances* 37(8).
- Park SV, Yang JS, Jo H, Kang B, Oh SS & Jung GY (2019b) Catalytic RNA, ribozyme, and its applications in synthetic biology. *Biotechnology advances* 37(8):107452.
- Pavlou MP, Diamandis EP & Blasutig IM (2013) The long journey of cancer biomarkers from the bench to the clinic. *Clinical chemistry* 59(1):147-157.

- Pehlivan ZS, Torabfam M, Kurt H, Ow-Yang C, Hildebrandt N & Yuce M (2019) Aptamer and nanomaterial based FRET biosensors: a review on recent advances (2014-2019). *Mikrochimica acta* 186(8):563.
- Penchovsky R (2012) Engineering integrated digital circuits with allosteric ribozymes for scaling up molecular computation and diagnostics. *ACS synthetic biology* 1(10):471-482.
- Penchovsky R (2013) Computational design and biosensor applications of small molecule-sensing allosteric ribozymes. *Biomacromolecules* 14(4):1240-1249.
- Penchovsky R (2014) Computational design of allosteric ribozymes as molecular biosensors. *Biotechnology advances* 32(5):1015-1027.
- Penchovsky R & Breaker RR (2005) Computational design and experimental validation of oligonucleotide-sensing allosteric ribozymes. *Nature biotechnology* 23(11):1424-1433.
- Penchovsky R & Kostova GT (2013) Computational selection and experimental validation of allosteric ribozymes that sense a specific sequence of human telomerase reverse transcriptase mRNAs as universal anticancer therapy agents. *Nucleic acid therapeutics* 23(6):408-417.
- Pendergrast PS, Marsh HN, Grate D, Healy JM & Stanton M (2005) Nucleic acid aptamers for target validation and therapeutic applications. *Journal of biomolecular techniques : JBT* 16(3):224-234.
- Penedo JC, Wilson TJ, Jayasena SD, Khvorova A & Lilley DM (2004) Folding of the natural hammerhead ribozyme is enhanced by interaction of auxiliary elements. *RNA* 10(5):880-888.
- Peng Y & Croce CM (2016) The role of MicroRNAs in human cancer. *Signal transduction and targeted therapy* 1:15004.
- Perreault J, Weinberg Z, Roth A, Popescu O, Chartrand P, Ferbeyre G & Breaker RR (2011a) Identification of Hammerhead Ribozymes in All Domains of Life Reveals Novel Structural Variations. *PLoS computational biology* 7(5).
- Perreault J, Weinberg Z, Roth A, Popescu O, Chartrand P, Ferbeyre G & Breaker RR (2011b) Identification of hammerhead ribozymes in all domains of life reveals novel structural variations. *PLoS computational biology* 7(5):e1002031.
- Perreault JaW, Zasha and Roth, Adam and Popescu, Olivia and Chartrand, Pascal and Ferbeyre, Gerardo and Breaker, Ronald R (2011) Identification of hammerhead ribozymes in all domains of life reveals novel structural variations. *PLoS computational biology* 7:e1002031.
- Perrotta AT & Been MD (1991) A pseudoknot-like structure required for efficient self-cleavage of hepatitis delta virus RNA. *Nature* 350(6317):434-436.
- Peselis A & Serganov A (2014) Structure and function of pseudoknots involved in gene expression control. *Wiley interdisciplinary reviews. RNA* 5(6):803-822.
- Petkovic S, Badelt S, Block S, Flamm C, Delcea M, Hofacker I & Muller S (2015) Sequence-controlled RNA self-processing: computational design, biochemical analysis, and visualization by AFM. *RNA* 21(7):1249-1260.
- Piganeau N (2009) In vitro selection of allosteric ribozymes. *Methods Mol Biol* 535:45-57.
- Piganeau N, Thuillier V & Famulok M (2001) In vitro selection of allosteric ribozymes: theory and experimental validation. *Journal of molecular biology* 312(5):1177-1190.
- Pipas JM & McMahon JE (1975) Method for predicting RNA secondary structure. *Proceedings of the National Academy of Sciences of the United States of America* 72(6):2017-2021.
- Pleij CW, Rietveld K & Bosch L (1985) A new principle of RNA folding based on pseudoknotting. *Nucleic acids research* 13(5):1717-1731.

- Porchetta A, Vallee-Belisle A, Plaxco KW & Ricci F (2012) Using distal-site mutations and allosteric inhibition to tune, extend, and narrow the useful dynamic range of aptamer-based sensors. *Journal of the American Chemical Society* 134(51):20601-20604.
- Porter EB, Polaski JT, Morck MM & Batey RT (2017) Recurrent RNA motifs as scaffolds for genetically encodable small-molecule biosensors. *Nature chemical biology* 13(3):295-301.
- Powers T & Noller HF (1991) A functional pseudoknot in 16S ribosomal RNA. *The EMBO journal* 10(8):2203-2214.
- Prasad Y, Smith JB, Gottlieb PA, Bentz J & Dinter-Gottlieb G (1992) Deriving a 67-nucleotide trans-cleaving ribozyme from the hepatitis delta virus antigenomic RNA. *Antisense research and development* 2(4):267-277.
- Prodeus A, Abdul-Wahid A, Fischer NW, Huang EH, Cydzik M & Garipey J (2015) Targeting the PD-1/PD-L1 Immune Evasion Axis With DNA Aptamers as a Novel Therapeutic Strategy for the Treatment of Disseminated Cancers. *Molecular therapy. Nucleic acids* 4:e237.
- Prody GA, Bakos JT, Buzayan JM, Schneider IR & Bruening G (1986) Autolytic processing of dimeric plant virus satellite RNA. *Science* 231(4745):1577-1580.
- Prommana P, Uthaipibull C, Wongsombat C, Kamchonwongpaisan S, Yuthavong Y, Knuepfer E, Holder AA & Shaw PJ (2013) Inducible Knockdown of Plasmodium Gene Expression Using the glmS Ribozyme. *PLoS one* 8(8).
- Prommana PaU, Chairat and Wongsombat, Chayaphat and Kamchonwongpaisan, Sumalee and Yuthavong, Yongyuth and Knuepfer, Ellen and Holder, Anthony A and Shaw, Philip J (2013) Inducible knockdown of Plasmodium gene expression using the glmS ribozyme. *PLoS one* 8:e73783.
- Prykhozhiy SV, Rajan V, Gaston D & Berman JN (2015) CRISPR multitargeter: a web tool to find common and unique CRISPR single guide RNA targets in a set of similar sequences. *PLoS one* 10(3):e0119372.
- Pu M, Chen J, Tao Z, Miao L, Qi X, Wang Y & Ren J (2019) Regulatory network of miRNA on its target: coordination between transcriptional and post-transcriptional regulation of gene expression. *Cellular and molecular life sciences : CMLS* 76(3):441-451.
- Pu Q, Zhou S, Huang X, Yuan Y, Du F, Dong J, Chen G, Cui X & Tang Z (2020) Intracellular Selection of Theophylline-Sensitive Hammerhead Aptazyme. *Molecular therapy. Nucleic acids* 20:400-408.
- Puttaraju M, Perrotta AT & Been MD (1993) A circular trans-acting hepatitis delta virus ribozyme. *Nucleic acids research* 21(18):4253-4258.
- Pyle AM (1993) Ribozymes: a distinct class of metalloenzymes. *Science* 261(5122):709-714.
- Quan K, Li J, Wang J, Xie N, Wei Q, Tang J, Yang X, Wang K & Huang J (2019) Dual-microRNA-controlled double-amplified cascaded logic DNA circuits for accurate discrimination of cell subtypes. *Chem Sci* 10(5):1442-1449.
- Ramesh A & Winkler WC (2014) Metabolite-binding ribozymes. *Biochimica et biophysica acta* 1839(10):989-994.
- Rastogi T, Beattie TL, Olive JE & Collins RA (1996) A long-range pseudoknot is required for activity of the Neurospora VS ribozyme. *The EMBO journal* 15(11):2820-2825.
- Ray P & White RR (2010) Aptamers for Targeted Drug Delivery. *Pharmaceuticals (Basel)* 3(6):1761-1778.
- Rehm C & Hartig JS (2014) In vivo screening for aptazyme-based bacterial riboswitches. *Methods Mol Biol* 1111:237-249.

- Rehm C, Klauser B & Hartig JS (2015) Engineering aptazyme switches for conditional gene expression in mammalian cells utilizing an in vivo screening approach. *Methods Mol Biol* 1316:127-140.
- Ren A, Vusurovic N, Gebetsberger J, Gao P, Juen M, Kreutz C, Micura R & Patel DJ (2016) Pistol ribozyme adopts a pseudoknot fold facilitating site-specific in-line cleavage. *Nature chemical biology* 12(9):702-708.
- Ren J, Rastegari B, Condon A & Hoos HH (2005) HotKnots: heuristic prediction of RNA secondary structures including pseudoknots. *RNA* 11(10):1494-1504.
- Retwitzer MD, Reinharz V, Churkin A, Ponty Y, Waldispuhl J & Barash D (2020) incaRNAfbinv 2.0: a webserver and software with motif control for fragment-based design of RNAs. *Bioinformatics* 36(9):2920-2922.
- Reuter JS & Mathews DH (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC bioinformatics* 11:129.
- Reyon D, Kirkpatrick JR, Sander JD, Zhang F, Voytas DF, Joung JK, Dobbs D & Coffman CR (2011) ZFNGenome: a comprehensive resource for locating zinc finger nuclease target sites in model organisms. *BMC genomics* 12:83.
- Rich A & RajBhandary UL (1976) Transfer RNA: molecular structure, sequence, and properties. *Annual review of biochemistry* 45:805-860.
- Robertson MP & Ellington AD (1999) In vitro selection of an allosteric ribozyme that transduces analytes to amplicons. *Nature biotechnology* 17(1):62-66.
- Robertson MP & Ellington AD (2001) In vitro selection of nucleoprotein enzymes. *Nature biotechnology* 19(7):650-655.
- Robertson MP, Knudsen SM & Ellington AD (2004) In vitro selection of ribozymes dependent on peptides for activity. *RNA* 10(1):114-127.
- Rodrigo G & Jaramillo A (2014) RiboMaker: computational design of conformation-based riboregulation. *Bioinformatics* 30(17):2508-2510.
- Rossi JJ, Cantin EM, Zaia JA, Ladne PA, Chen J, Stephens DA, Sarver N & Chang PS (1990) Ribozymes as therapies for AIDS. *Annals of the New York Academy of Sciences* 616:184-200.
- Rossi JJ & Sarver N (1990) RNA enzymes (ribozymes) as antiviral therapeutic agents. *Trends in biotechnology* 8(7):179-183.
- Roth A & Breaker RR (2004) Selection in vitro of allosteric ribozymes. *Methods Mol Biol* 252:145-164.
- Roth A, Weinberg Z, Chen AG, Kim PB, Ames TD & Breaker RR (2014a) A widespread self-cleaving ribozyme class is revealed by bioinformatics. *Nature chemical biology* 10(1):56-60.
- Roth A, Weinberg Z, Chen AGY, Kim PB, Ames TD & Breaker RR (2014b) A widespread self-cleaving ribozyme class is revealed by bioinformatics. *Nature chemical biology* 10(1):56-U92.
- Roy G, Ananvoranich S & Perreault J-P (1999) Delta ribozyme has the ability to cleave in trans an mRNA. *Nucleic acids research* 27(4):942-948.
- Rozenblum GT, Lopez VG, Vitullo AD & Radrizzani M (2016) Aptamers: current challenges and future prospects. *Expert opinion on drug discovery* 11(2):127-135.
- Ruder WCaL, Ting and Collins, James J (2011) Synthetic biology moving into the clinic. *Science* 333:1248--1252.
- Rueda D & Walter NG (2006) Fluorescent energy transfer readout of an aptazyme-based biosensor. *Methods Mol Biol* 335:289-310.
- Ruffner DE, Stormo GD & Uhlenbeck OC (1990) Sequence requirements of the hammerhead RNA self-cleavage reaction. *Biochemistry* 29(47):10695-10702.

- Russell R (1990) Ribozymes versus HIV. *British dental journal* 169(2):50.
- Sabrina Najeh KZ, Nawwaf Kharma and Jonathan Perreault (2019) Computational design and experimental verification of pseudoknotted ribozymes. *RNA*.
- Sack M, Stifel J, Kreft SG, Deuerling E & Hartig JS (2019) Neomycin-dependent hammerhead ribozymes for the direct control of gene expression in *Saccharomyces cerevisiae*. *Methods* 10.1016/j.ymeth.2019.01.004.
- Sagan SM, Nasheri N, Luebbert C & Pezacki JP (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 CH & Wu GY (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.
- Saksmerprome V, Roychowdhury-Saha M, Jayasena S, Khvorova A & Burke DH (2004) Artificial tertiary motifs stabilize trans-cleaving hammerhead ribozymes under conditions of submillimolar divalent ions and high temperatures. *Rna* 10(12):1916-1924.
- Sambrook J, Fritsch, E.F., and Maniatis, T (1989) *Molecular Cloning: A Laboratory Manual*. Cold Springs Harbor Laboratory Press, Cold Springs Harbor Laboratory, N,Y
- Sanchez-Luque FJ, Stich M, Manrubia S, Briones C & Berzal-Herranz A (2014) Efficient HIV-1 inhibition by a 16 nt-long RNA aptamer designed by combining in vitro selection and in silico optimisation strategies. *Scientific reports* 4:6242.
- Saragliadis A, Krajewski SS, Rehm C, Narberhaus F & Hartig JS (2013) Thermozymes: Synthetic RNA thermometers based on ribozyme activity. *RNA biology* 10(6):1010-1016.
- Sarver N, Cantin EM, Chang PS, Zaia JA, Ladne PA, Stephens DA & Rossi JJ (1990) Ribozymes as potential anti-HIV-1 therapeutic agents. *Science* 247(4947):1222-1225.
- Sato K, Akiyama M & Sakakibara Y (2021) RNA secondary structure prediction using deep learning with thermodynamic integration. *Nature communications* 12(1):941.
- Sato KaK, Yuki and Hamada, Michiaki and Akutsu, Tatsuya and Asai, Kiyoshi (2011) IPknot: fast and accurate prediction of RNA secondary structures with pseudoknots using integer programming. *Bioinformatics* 27:i85--i93.
- Savanur MA, Weinstein-Marom H & Gross G (2021) Implementing Logic Gates for Safer Immunotherapy of Cancer. *Frontiers in immunology* 12:780399.
- Saville BJ & Collins RA (1990) A site-specific self-cleavage reaction performed by a novel RNA in *Neurospora mitochondria*. *Cell* 61(4):685-696.
- Savinov A & Block SM (2018) Self-cleavage of the glmS ribozyme core is controlled by a fragile folding element. *Proceedings of the National Academy of Sciences of the United States of America* 115(47):11976-11981.
- Scanlon KJ, Jiao L, Funato T, Wang W, Tone T, Rossi JJ & Kashani-Sabet M (1991) Ribozyme-mediated cleavage of c-fos mRNA reduces gene expression of DNA synthesis enzymes and metallothionein. *Proceedings of the National Academy of Sciences of the United States of America* 88(23):10591-10595.
- Scarborough RJ & Gagnon A (2015) HIV and Ribozymes. *Advances in experimental medicine and biology* 848:97-116.
- Scarborough RJ, Levesque MV, Perreault JP & Gagnon A (2014) Design and evaluation of clinically relevant SOFA-HDV ribozymes targeting HIV RNA. *Methods Mol Biol* 1103:31-43.
- Scherr M, Grez M, Ganser A & Engels JW (1997) Specific hammerhead ribozyme-mediated cleavage of mutant N-ras mRNA in vitro and ex vivo. Oligoribonucleotides as therapeutic agents. *The Journal of biological chemistry* 272(22):14304-14313.

- Schmidt KS, Borkowski S, Kurreck J, Stephens AW, Bald R, Hecht M, Friebe M, Dinkelborg L & Erdmann VA (2004) Application of locked nucleic acids to improve aptamer in vivo stability and targeting function. *Nucleic acids research* 32(19):5757-5765.
- Schnall-Levin MaC, Leonid and Berger, Bonnie (2008) Inverting the Viterbi algorithm: an abstract framework for structure design. *Proceedings of the 25th international conference on Machine learning* :904--911.
- Schubert S & Kurreck J (2004) Ribozyme- and deoxyribozyme-strategies for medical applications. *Current drug targets* 5(8):667-681.
- Schultz EP, Vasquez EE & Scott WG (2014) Structural and catalytic effects of an invariant purine substitution in the hammerhead ribozyme: implications for the mechanism of acid-base catalysis. *Acta crystallographica. Section D, Biological crystallography* 70(Pt 9):2256-2263.
- Schwab R, Ossowski S, Riester M, Warthmann N & Weigel D (2006) Highly specific gene silencing by artificial microRNAs in Arabidopsis. *The Plant cell* 18(5):1121-1133.
- Sciabola S, Xi H, Cruz D, Cao Q, Lawrence C, Zhang T, Rotstein S, Hughes JD, Caffrey DR & Stanton RV (2021) PFRED: A computational platform for siRNA and antisense oligonucleotides design. *PLoS one* 16(1):e0238753.
- Scott WG (2010) What can the New Hammerhead Ribozyme Structures Teach us About Design? *RNA Technologies and Their Applications*, V.A. Erdmann JB (Édit.)Springer. p 305-324.
- Scott WG, Horan LH & Martick M (2013) The hammerhead ribozyme: structure, catalysis, and gene regulation. *Progress in molecular biology and translational science* 120:1-23.
- Scott WG, Murray JB, Arnold JR, Stoddard BL & Klug A (1996) Capturing the structure of a catalytic RNA intermediate: the hammerhead ribozyme. *Science* 274(5295):2065-2069.
- Seetin MG & Mathews DH (2012) RNA structure prediction: an overview of methods. *Methods Mol Biol* 905:99-122.
- Sepehri Zarandi H, Behbahani M & Mohabatkar H (2020) In Silico Selection of Gp120 ssDNA Aptamer to HIV-1. *SLAS discovery : advancing life sciences R & D* 25(9):1087-1093.
- Serganov A & Nudler E (2013) A decade of riboswitches. *Cell* 152(1-2):17-24.
- Serra MJ & Turner DH (1995) Predicting thermodynamic properties of RNA. *Methods in enzymology* 259:242-261.
- Sewgobind NV, Albers S & Pieters RJ (2021) Functions and Inhibition of Galectin-7, an Emerging Target in Cellular Pathophysiology. *Biomolecules* 11(11).
- Shah SA & Brunger AT (1999) The 1.8 Å crystal structure of a statically disordered 17 base-pair RNA duplex: principles of RNA crystal packing and its effect on nucleic acid structure. *Journal of molecular biology* 285(4):1577-1588.
- Shamas-Din A, Kale J, Leber B & Andrews DW (2013) Mechanisms of action of Bcl-2 family proteins. *Cold Spring Harbor perspectives in biology* 5(4):a008714.
- Shapiro E & Gil B (2008) RNA computing in a living cell. *Science* 322(5900):387-388.
- Shi YZ, Jin L, Feng CJ, Tan YL & Tan ZJ (2018) Predicting 3D structure and stability of RNA pseudoknots in monovalent and divalent ion solutions. *PLoS computational biology* 14(6):e1006222.
- Shivram H, Cress BF, Knott GJ & Doudna JA (2021) Controlling and enhancing CRISPR systems. *Nature chemical biology* 17(1):10-19.
- Shu D, Li H, Shu Y, Xiong G, Carson WE, 3rd, Haque F, Xu R & Guo P (2015) Systemic Delivery of Anti-miRNA for Suppression of Triple Negative Breast Cancer Utilizing RNA Nanotechnology. *ACS nano* 9(10):9731-9740.

- Shum KT & Tanner JA (2008) Differential inhibitory activities and stabilisation of DNA aptamers against the SARS coronavirus helicase. *Chembiochem : a European journal of chemical biology* 9(18):3037-3045.
- Singh V (2014) Recent advances and opportunities in synthetic logic gates engineering in living cells. *Systems and synthetic biology* 8(4):271-282.
- Sivakumar P, Kim S, Kang HC & Shim MS (2019) Targeted siRNA delivery using aptamer-siRNA chimeras and aptamer-conjugated nanoparticles. *Wiley interdisciplinary reviews. Nanomedicine and nanobiotechnology* 11(3):e1543.
- Sokilde R, Persson H, Ehinger A, Pirona AC, Ferno M, Hegardt C, Larsson C, Loman N, Malmberg M, Ryden L, Saal L, Borg A, Vallon-Christerson J & Rovira C (2019) Refinement of breast cancer molecular classification by miRNA expression profiles. *BMC genomics* 20(1):503.
- Soler M, Estevez MC, Alvarez M, Otte MA, Sepulveda B & Lechuga LM (2014) Direct detection of protein biomarkers in human fluids using site-specific antibody immobilization strategies. *Sensors (Basel)* 14(2):2239-2258.
- Song T, Shah S, Bui H, Garg S, Eshra A, Fu D, Yang M, Mokhtar R & Reif J (2019) Programming DNA-Based Biomolecular Reaction Networks on Cancer Cell Membranes. *Journal of the American Chemical Society* 141(42):16539-16543.
- Soukup GA (2006) Core requirements for glmS ribozyme self-cleavage reveal a putative pseudoknot structure. *Nucleic acids research* 34(3):968-975.
- Soukup GA & Breaker RR (1999a) Design of allosteric hammerhead ribozymes activated by ligand-induced structure stabilization. *Structure* 7(7):783-791.
- Soukup GA & Breaker RR (1999b) Engineering precision RNA molecular switches. *Proceedings of the National Academy of Sciences of the United States of America* 96(7):3584-3589.
- Soukup GA & Breaker RR (1999c) Engineering precision RNA molecular switches. *Proceedings of the National Academy of Sciences* 96(7):3584-3589.
- Staple DW & Butcher SE (2005) Pseudoknots: RNA structures with diverse functions. *PLoS biology* 3(6):e213.
- Stifel J, Sporing M & Hartig JS (2019a) Expanding the toolbox of synthetic riboswitches with guanine-dependent aptazymes. *Synth Biol (Oxf)* 4(1):ysy022.
- Stifel J, Sporing M & Hartig JS (2019b) Expanding the toolbox of synthetic riboswitches with guanine-dependent aptazymes. *Syn Biol* 4(1).
- Strobel B, Sporing M, Klein H, Blazevic D, Rust W, Sayols S, Hartig JS & Kreuz S (2020) High-throughput identification of synthetic riboswitches by barcode-free amplicon-sequencing in human cells. *Nature communications* 11(1):714.
- Studnicka GM, Rahn GM, Cummings IW & Salser WA (1978) Computer method for predicting the secondary structure of single-stranded RNA. *Nucleic acids research* 5(9):3365-3387.
- Sudarsan N, Hammond MC, Block KF, Welz R, Barrick JE, Roth A & Breaker RR (2006) Tandem riboswitch architectures exhibit complex gene control functions. *Science* 314(5797):300-304.
- Sullenger BA & Gilboa E (2002) Emerging clinical applications of RNA. *Nature* 418(6894):252-258.
- Sun C, Liu X, Feng K, Jiang J, Shen G & Yu R (2010) An aptazyme-based electrochemical biosensor for the detection of adenosine. *Analytica chimica acta* 669(1-2):87-93.
- Szabat M, Prochota M, Kierzek R, Kierzek E & Mathews DH (2022) A Test and Refinement of Folding Free Energy Nearest Neighbor Parameters for RNA Including N(6)-Methyladenosine. *Journal of molecular biology* 10.1016/j.jmb.2022.167632:167632.
- Tafer H, Ameres SL, Obernosterer G, Gebeshuber CA, Schroeder R, Martinez J & Hofacker IL (2008) The impact of target site accessibility on the design of effective siRNAs. *Nature biotechnology* 26(5):578-583.

- Taneda A (2011) MODENA: a multi-objective RNA inverse folding. *Advances and applications in bioinformatics and chemistry: AABC* 4:1.
- Taneda A (2015) Multi-objective optimization for RNA design with multiple target secondary structures. *BMC bioinformatics* 16:280.
- Tang J & Breaker RR (1997) Rational design of allosteric ribozymes. *Chemistry & biology* 4(6):453-459.
- Tang X, Ren Q, Yang L, Bao Y, Zhong Z, He Y, Liu S, Qi C, Liu B, Wang Y, Sretenovic S, Zhang Y, Zheng X, Zhang T, Qi Y & Zhang Y (2018) Single transcript unit CRISPR 2.0 systems for robust Cas9 and Cas12a mediated plant genome editing. *Plant biotechnology journal* 10.1111/pbi.13068.
- Tanzer A, Hofacker IL & Lorenz R (2019) RNA modifications in structure prediction - Status quo and future challenges. *Methods* 156:32-39.
- Theis CaJ, Stefan and Giegerich, Robert (2010) Prediction of RNA Secondary Structure Including Kissing Hairpin Motifs. *WABI*, Springer. p 52--64.
- Thiel KW & Giangrande PH (2010) Intracellular delivery of RNA-based therapeutics using aptamers. *Therapeutic delivery* 1(6):849-861.
- Thiel WH & Giangrande PH (2016) Analyzing HT-SELEX data with the Galaxy Project tools--A web based bioinformatics platform for biomedical research. *Methods* 97:3-10.
- Thiemann S & Baum LG (2016) Galectins and Immune Responses-Just How Do They Do Those Things They Do? *Annual review of immunology* 34:243-264.
- Thompson KM, Syrett HA, Knudsen SM & Ellington AD (2002) Group I aptazymes as genetic regulatory switches. *BMC biotechnology* 2:21.
- Tinder RF (2000) Background for Digital Design. *Engineering Digital Design*, Academic Press, United States of America, 2 Ed. p 79-130.
- Tinoco I, Jr., Borer PN, Dengler B, Levin MD, Uhlenbeck OC, Crothers DM & Bralla J (1973) Improved estimation of secondary structure in ribonucleic acids. *Nature: New biology* 246(150):40-41.
- Tinoco I, Jr., Uhlenbeck OC & Levine MD (1971) Estimation of secondary structure in ribonucleic acids. *Nature* 230(5293):362-367.
- Toh SY, Citartan M, Gopinath SC & Tang TH (2015) Aptamers as a replacement for antibodies in enzyme-linked immunosorbent assay. *Biosensors & bioelectronics* 64:392-403.
- Townshend B, Kennedy AB, Xiang JS & Smolke CD (2015) High-throughput cellular RNA device engineering. *Nature methods* 12(10):989-994.
- Townshend B, Xiang JS, Manzanarez G, Hayden EJ & Smolke CD (2021a) A multiplexed, automated evolution pipeline enables scalable discovery and characterization of biosensors. *Nature communications* 12(1):1437.
- Townshend RJJ, Eismann S, Watkins AM, Rangan R, Karelina M, Das R & Dror RO (2021b) Geometric deep learning of RNA structure. *Science* 373(6558):1047-1051.
- Tuerk C & Gold L (1990a) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249(4968):505-510.
- Tuerk C & Gold L (1990b) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249(4968):505-510.
- Turner DH (1996) Thermodynamics of base pairing. *Current opinion in structural biology* 6(3):299-304.
- Tuschl T, Sharp PA & Bartel DP (1998) Selection in vitro of novel ribozymes from a partially randomized U2 and U6 snRNA library. *The EMBO journal* 17(9):2637-2650.
- Uhlenbeck OC (1987) A Small Catalytic Oligoribonucleotide. *Nature* 328(6131):596-600.

- Vaish NK, Dong F, Andrews L, Schweppe RE, Ahn NG, Blatt L & Seiwert SD (2002) Monitoring post-translational modification of proteins with allosteric ribozymes. *Nature biotechnology* 20(8):810-815.
- Vanarsa K, Enan S, Patel P, Strachan B, Sam Titus A, Dennis A, Lotan Y & Mohan C (2021) Urine protein biomarkers of bladder cancer arising from 16-plex antibody-based screens. *Oncotarget* 12(8):783-790.
- Varani G & Tinoco I, Jr. (1991) RNA structure and NMR spectroscopy. *Quarterly reviews of biophysics* 24(4):479-532.
- Veedu RN & Wengel J (2010) Locked nucleic acids: promising nucleic acid analogs for therapeutic applications. *Chemistry & biodiversity* 7(3):536-542.
- Vert JP, Foveau N, Lajaunie C & Vandenbrouck Y (2006) An accurate and interpretable model for siRNA efficacy prediction. *BMC bioinformatics* 7:520.
- Viladoms J & Fedor MJ (2012) The glmS ribozyme cofactor is a general acid-base catalyst. *Journal of the American Chemical Society* 134(46):19043-19049.
- Villalonga A, Perez-Calabuig AM & Villalonga R (2020) Electrochemical biosensors based on nucleic acid aptamers. *Analytical and bioanalytical chemistry* 412(1):55-72.
- Vladoiu MC, Labrie M, Letourneau M, Egesborg P, Gagne D, Billard E, Grosset AA, Doucet N, Chatenet D & St-Pierre Y (2015) Design of a peptidic inhibitor that targets the dimer interface of a prototypic galectin. *Oncotarget* 6(38):40970-40980.
- Vladoiu MC, Labrie M & St-Pierre Y (2014) Intracellular galectins in cancer cells: potential new targets for therapy (Review). *International journal of oncology* 44(4):1001-1014.
- Voigt CA (2006) Genetic parts to program bacteria. *Current opinion in biotechnology* 17(5):548-557.
- Vorobyeva M, Vorobjev P & Venyaminova A (2016) Multivalent Aptamers: Versatile Tools for Diagnostic and Therapeutic Applications. *Molecules* 21(12).
- Vusurovic N, Altman RB, Terry DS, Micura R & Blanchard SC (2017) Pseudoknot Formation Seeds the Twister Ribozyme Cleavage Reaction Coordinate. *Journal of the American Chemical Society* 139(24):8186-8193.
- Wallden K, Stenmark P, Nyman T, Flodin S, Graslund S, Loppnau P, Bianchi V & Nordlund P (2007) Crystal structure of human cytosolic 5'-nucleotidase II: insights into allosteric regulation and substrate recognition. *The Journal of biological chemistry* 282(24):17828-17836.
- Walter JG & Stahl F (2020) Aptazymes: Expanding the Specificity of Natural Catalytic Nucleic Acids by Application of In Vitro Selected Oligonucleotides. *Advances in biochemical engineering/biotechnology* 170:107-119.
- Walter NG & Engelke DR (2002) Ribozymes: catalytic RNAs that cut things, make things, and do odd and useful jobs. *Biologist (London)* 49(5):199-203.
- Walton T, DasGupta S, Duzdevich D, Oh SS & Szostak JW (2020) In vitro selection of ribozyme ligases that use prebiotically plausible 2-aminoimidazole-activated substrates. *Proceedings of the National Academy of Sciences of the United States of America* 117(11):5741-5748.
- Wan Q, Liu X & Zu Y (2021) Oligonucleotide aptamers for pathogen detection and infectious disease control. *Theranostics* 11(18):9133-9161.
- Wan Y, Kertesz M, Spitale RC, Segal E & Chang HY (2011) Understanding the transcriptome through RNA structure. *Nature reviews. Genetics* 12(9):641-655.
- Wang CX, Lu YQ, Qi P, Chen LH & Han JX (2010) Efficient inhibition of hepatitis B virus replication by hepatitis delta virus ribozymes delivered by targeting retrovirus. *Virology journal* 7:61.

- Wang K, Chen M & Wu W (2017) Analysis of microRNA (miRNA) expression profiles reveals 11 key biomarkers associated with non-small cell lung cancer. *World journal of surgical oncology* 15(1):175.
- Wang R, Wang S, Xu X, Jiang W & Zhang N (2021a) MNzyme probes mediated DNA logic platform for microRNAs logic detection and cancer cell identification. *Analytica chimica acta* 1149:338213.
- Wang X, Gao J, Zhou B, Xie J, Zhou G & Chen Y (2019) Identification of prognostic markers for hepatocellular carcinoma based on miRNA expression profiles. *Life sciences* 232:116596.
- Wang X, Liu L, Wu H, Wu Z, Tang LJ & Jiang JH (2021b) Programming DNA cascade circuits on live cell membranes for accurate cancer cell recognition and gene silencing. *Chem Commun (Camb)* 57(31):3816-3819.
- Wang YH, Wei KY & Smolke CD (2013) Synthetic biology: advancing the design of diverse genetic systems. *Annual review of chemical and biomolecular engineering* 4:69-102.
- Wang Z, Liu X, Xie X, Deng L, Zheng H, Pan H, Li D, Li L & Zhong C (2021c) ABE8e with Polycistronic tRNA-gRNA Expression Cassette Significantly Improves Adenine Base Editing Efficiency in *Nicotiana benthamiana*. *International journal of molecular sciences* 22(11).
- Webb CH, Riccitelli NJ, Ruminski DJ & Luptak A (2009) Widespread occurrence of self-cleaving ribozymes. *Science* 326(5955):953.
- Wei F, Patel P, Liao W, Chaudhry K, Zhang L, Arellano-Garcia M, Hu S, Elashoff D, Zhou H, Shukla S, Shah F, Ho CM & Wong DT (2009) Electrochemical sensor for multiplex biomarkers detection. *Clinical cancer research : an official journal of the American Association for Cancer Research* 15(13):4446-4452.
- Weinberg CE, Weinberg Z & Hammann C (2019a) Novel ribozymes: discovery, catalytic mechanisms, and the quest to understand biological function. *Nucleic acids research* 47(18):9480-9494.
- Weinberg CE, Weinberg Z & Hammann C (2019b) Novel ribozymes: discovery, catalytic mechanisms, and the quest to understand biological function. *Nucleic acids research* 47(18):9480-9494.
- Weinberg MS, Ely A, Passman M, Mufamadi SM & Arbutnot P (2007) Effective anti-hepatitis B virus hammerhead ribozymes derived from multimeric precursors. *Oligonucleotides* 17(1):104-112.
- Weinberg Z, Kim PB, Chen TH, Li S, Harris KA, Lunse CE & Breaker RR (2015a) New classes of self-cleaving ribozymes revealed by comparative genomics analysis. *Nature chemical biology* 11(8):606-610.
- Weinberg Z, Kim PB, Chen TH, Li SS, Harris KA, Lunse CE & Breaker RR (2015b) New classes of self-cleaving ribozymes revealed by comparative genomics analysis. *Nature chemical biology* 11(8):606-U111.
- Weng DE, Masci PA, Radka SF, Jackson TE, Weiss PA, Ganapathi R, Elson PJ, Capra WB, Parker VP, Lockridge JA, Cowens JW, Usman N & Borden EC (2005) A phase I clinical trial of a ribozyme-based angiogenesis inhibitor targeting vascular endothelial growth factor receptor-1 for patients with refractory solid tumors. *Molecular cancer therapeutics* 4(6):948-955.
- Wieland M, Benz A, Klauser B & Hartig JS (2009) Artificial ribozyme switches containing natural riboswitch aptamer domains. *Angew Chem Int Ed Engl* 48(15):2715-2718.
- Wieland M, Berschneider B, Erlacher MD & Hartig JS (2010) Aptazyme-mediated regulation of 16S ribosomal RNA. *Chemistry & biology* 17(3):236-242.

- Wieland M & Hartig JS (2008) Improved aptazyme design and in vivo screening enable riboswitching in bacteria. *Angew Chem Int Ed Engl* 47(14):2604-2607.
- Wilds CJ & Damha MJ (2000) 2'-Deoxy-2'-fluoro-beta-D-arabinonucleosides and oligonucleotides (2'-F-ANA): synthesis and physicochemical studies. *Nucleic acids research* 28(18):3625-3635.
- Win MN & Smolke CD (2007a) A modular and extensible RNA-based gene-regulatory platform for engineering cellular function. *Proceedings of the National Academy of Sciences of the United States of America* 104(36):14283-14288.
- Win MN & Smolke CD (2007b) A modular and extensible RNA-based gene-regulatory platform for engineering cellular function. *Proceedings of the National Academy of Sciences of the United States of America* 104(36):14283-14288.
- Win MN & Smolke CD (2008) Higher-order cellular information processing with synthetic RNA devices. *Science* 322(5900):456-460.
- Winkler WC, Nahvi A, Roth A, Collins JA & Breaker RR (2004) Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* 428(6980):281-286.
- Wittmann A & Suess B (2011) Selection of tetracycline inducible self-cleaving ribozymes as synthetic devices for gene regulation in yeast. *Molecular bioSystems* 7(8):2419-2427.
- Wolter O & Mayer G (2017) Aptamers as Valuable Molecular Tools in Neurosciences. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 37(10):2517-2523.
- Wong-Staal F, Poeschla EM & Looney DJ (1998) A controlled, Phase 1 clinical trial to evaluate the safety and effects in HIV-1 infected humans of autologous lymphocytes transduced with a ribozyme that cleaves HIV-1 RNA. *Human gene therapy* 9(16):2407-2425.
- Wu JC, Gardner DP, Ozer S, Gutell RR & Ren P (2009) Correlation of RNA secondary structure statistics with thermodynamic stability and applications to folding. *Journal of molecular biology* 391(4):769-783.
- Wu L, Pan M, Chen Y, Huang H, Zhang X, Wang F & Zhou X (2019) The construction of DNAzyme-based logic gates for amplified microRNA detection and cancer recognition. *The Analyst* 144(24):7278-7282.
- Wurmthaler LA, Klausner B & Hartig JS (2018) Highly motif- and organism-dependent effects of naturally occurring hammerhead ribozyme sequences on gene expression. *RNA biology* 15(2):231-241.
- Wurmthaler LA, Sack M, Gense K, Hartig JS & Gamerdinger M (2019) A tetracycline-dependent ribozyme switch allows conditional induction of gene expression in *Caenorhabditis elegans*. *Nature communications* 10(1):491.
- Xiang Y, Dalchau N & Wang B (2018) Scaling up genetic circuit design for cellular computing: advances and prospects. *Natural computing* 17(4):833-853.
- Xie Z, Wroblewska L, Prochazka L, Weiss R & Benenson Y (2011) Multi-input RNAi-based logic circuit for identification of specific cancer cells. *Science* 333(6047):1307-1311.
- Yamada M & Tanaka Y (2020) Structure-activity relationship of pseudoknot-type hammerhead ribozyme reveals key structural elements for enhanced catalytic activity(dagger). *Nucleosides, nucleotides & nucleic acids* 39(1-3):245-257.
- Yamagami R, Kayedkhordeh M, Mathews DH & Bevilacqua PC (2019) Design of highly active double-pseudoknotted ribozymes: a combined computational and experimental study. *Nucleic acids research* 47(1):29-42.
- Yang B, Xiong WY, Hou HJ, Xu Q, Cai XL, Zeng TX & Ha XQ (2019) Exosomal miRNAs as Biomarkers of Cancer: a Meta-Analysis. *Clinical laboratory* 65(5).
- Yang J & Ding S (2021) Chimeric RNA-binding protein-based killing switch targeting hepatocellular carcinoma cells. *Molecular therapy. Nucleic acids* 25:683-695.

- Yang P, Wang J, Pang Q, Zhang F, Wang J, Wang Q & Qi Q (2017) Pathway optimization and key enzyme evolution of N-acetylneuraminic acid biosynthesis using an in vivo aptzyme-based biosensor. *Metabolic engineering* 43(Pt A):21-28.
- Yang Q, Yang F, Dai W, Meng X, Wei W, Cheng Y, Dong J, Lu H & Dong H (2021) DNA Logic Circuits for Multiple Tumor Cells Identification Using Intracellular MicroRNA Molecular Bispecific Recognition. *Advanced healthcare materials* 10(21):e2101130.
- Yerukala Sathipati S & Ho SY (2018) Identifying a miRNA signature for predicting the stage of breast cancer. *Scientific reports* 8(1):16138.
- Yip KW & Reed JC (2008) Bcl-2 family proteins and cancer. *Oncogene* 27(50):6398-6406.
- Yokobayashi Y (2019a) Aptamer-based and aptzyme-based riboswitches in mammalian cells. *Current opinion in chemical biology* 52:72-78.
- Yokobayashi Y (2019b) Aptamer-based and aptzyme-based riboswitches in mammalian cells. *Current opinion in chemical biology* 52:72-78.
- Yokobayashi Y (2020) High-Throughput Analysis and Engineering of Ribozymes and Deoxyribozymes by Sequencing. *Accounts of chemical research* 53(12):2903-2912.
- Yokoyama T, Tsukakoshi K, Yoshida W, Saito T, Teramoto K, Savory N, Abe K & Ikebukuro K (2017) Development of HGF-binding aptamers with the combination of G4 promoter-derived aptamer selection and in silico maturation. *Biotechnology and bioengineering* 114(10):2196-2203.
- Yu L, Cui X, Li H, Lu J, Kang Q & Shen D (2019) A ratiometric electrochemical sensor for multiplex detection of cancer biomarkers using bismuth as an internal reference and metal sulfide nanoparticles as signal tags. *The Analyst* 144(13):4073-4080.
- Yue R, Chen M & Ma N (2020) Dual MicroRNA-Triggered Drug Release System for Combined Chemotherapy and Gene Therapy with Logic Operation. *ACS applied materials & interfaces* 12(29):32493-32502.
- Yue R, Li Z, Wang G, Li J & Ma N (2019) Logic Sensing of MicroRNA in Living Cells Using DNA-Programmed Nanoparticle Network with High Signal Gain. *ACS sensors* 4(1):250-256.
- Zadeh JN, Steenberg CD, Bois JS, Wolfe BR, Pierce MB, Khan AR, Dirks RM & Pierce NA (2011a) NUPACK: Analysis and design of nucleic acid systems. *Journal of computational chemistry* 32(1):170-173.
- Zadeh JN, Wolfe BR & Pierce NA (2011b) Nucleic acid sequence design via efficient ensemble defect optimization. *Journal of computational chemistry* 32(3):439-452.
- Zalatan JG, Lee ME, Almeida R, Gilbert LA, Whitehead EH, La Russa M, Tsai JC, Weissman JS, Dueber JE, Qi LS & Lim WA (2015) Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell* 160(1-2):339-350.
- Zandi K, Butler G & Kharna N (2016) An Adaptive Defect Weighted Sampling Algorithm to Design Pseudoknotted RNA Secondary Structures. *Frontiers in genetics* 7:129.
- Zangar RC, Daly DS & White AM (2006) ELISA microarray technology as a high-throughput system for cancer biomarker validation. *Expert review of proteomics* 3(1):37-44.
- Zhang B & Cech TR (1997) Peptide bond formation by in vitro selected ribozymes. *Nature* 390(6655):96-100.
- Zhang C, Zhao Y, Xu X, Xu R, Li H, Teng X, Du Y, Miao Y, Lin HC & Han D (2020a) Cancer diagnosis with DNA molecular computation. *Nature nanotechnology* 15(8):709-715.
- Zhang H, Zhang L, Mathews DH & Huang L (2020b) LinearPartition: linear-time approximation of RNA folding partition function and base-pairing probabilities. *Bioinformatics* 36(Suppl_1):i258-i267.
- Zhang J, Fu H & Chu X (2021a) Metal-Organic Framework Nanoparticles Power DNAzyme Logic Circuits for Aberrant MicroRNA Imaging. *Analytical chemistry* 93(44):14675-14684.

- Zhang MM, Bahal R, Rasmussen TP, Manautou JE & Zhong XB (2021b) The growth of siRNA-based therapeutics: Updated clinical studies. *Biochemical pharmacology* 189:114432.
- Zhang S, Chen H & Wang J (2019a) Generate TALE/TALEN as Easily and Rapidly as Generating CRISPR. *Molecular therapy. Methods & clinical development* 13:310-320.
- Zhang W, Liu QX, Guo ZH & Lin JS (2018) Practical Application of Aptamer-Based Biosensors in Detection of Low Molecular Weight Pollutants in Water Sources. *Molecules* 23(2).
- Zhang Y, Hong H & Cai W (2011) Tumor-targeted drug delivery with aptamers. *Current medicinal chemistry* 18(27):4185-4194.
- Zhang Y, Lai BS & Juhas M (2019b) Recent Advances in Aptamer Discovery and Applications. *Molecules* 24(5).
- Zhang Y, Wang J, Cheng H, Sun Y, Liu M, Wu Z & Pei R (2017) Conditional control of suicide gene expression in tumor cells with theophylline-responsive ribozyme. *Gene therapy* 24(2):84-91.
- Zhao S, Yu L, Yang S, Tang X, Chang K & Chen M (2021) Boolean logic gate based on DNA strand displacement for biosensing: current and emerging strategies. *Nanoscale horizons* 6(4):298-310.
- Zheng X, Gao S, Wu J & Hu X (2020) Recent Advances in Aptamer-Based Biosensors for Detection of *Pseudomonas aeruginosa*. *Frontiers in microbiology* 11:605229.
- Zhong G, Wang H, Bailey CC, Gao G & Farzan M (2016a) Rational design of aptazyme riboswitches for efficient control of gene expression in mammalian cells. *Elife* 5:e18858.
- Zhong G, Wang H, Bailey CC, Gao G & Farzan M (2016b) Rational design of aptazyme riboswitches for efficient control of gene expression in mammalian cells. *eLife* 5.
- Zhong J, Karberg M & Lambowitz AM (2003) Targeted and random bacterial gene disruption using a group II intron (targetron) vector containing a retrotransposition-activated selectable marker. *Nucleic acids research* 31(6):1656-1664.
- Zhou Q, Liu Y, Shin DS, Silangcruz J, Tuleuova N & Revzin A (2012) Aptamer-containing surfaces for selective capture of CD4 expressing cells. *Langmuir : the ACS journal of surfaces and colloids* 28(34):12544-12549.
- Zhou W, Huang PJ, Ding J & Liu J (2014) Aptamer-based biosensors for biomedical diagnostics. *The Analyst* 139(11):2627-2640.
- Zhu G & Chen X (2018) Aptamer-based targeted therapy. *Advanced drug delivery reviews* 134:65-78.
- Zhu J, Zheng Z, Wang J, Sun J, Wang P, Cheng X, Fu L, Zhang L, Wang Z & Li Z (2014) Different miRNA expression profiles between human breast cancer tumors and serum. *Frontiers in genetics* 5:149.
- Zhuo Z, Yu Y, Wang M, Li J, Zhang Z, Liu J, Wu X, Lu A, Zhang G & Zhang B (2017) Recent Advances in SELEX Technology and Aptamer Applications in Biomedicine. *International journal of molecular sciences* 18(10).
- Zou X, Wu J, Gu J, Shen L & Mao L (2019) Application of Aptamers in Virus Detection and Antiviral Therapy. *Frontiers in microbiology* 10:1462.
- Zuker M (2003a) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research* 31:3406-3415.
- Zuker M (2003b) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research* 31(13):3406-3415.

Annexes

Annexe I

Ribozymes de Ribosoft *in vivo* (article 3)

Suite aux résultats obtenus avec les tests des ribozymes *in vitro* ayant montré une activité de clivage assez élevée, quelques séquences ont été choisies pour vérifier l'activité des ribozymes correspondant dans des cellules exprimant la Gal-7 et la Bcl-2. Le choix de Bcl-2 et Gal-7 est basé sur le fait qu'elles sont impliquées dans la survie des cellules cancéreuses et leurs résistances aux thérapies comme la chimio- et la radiothérapie. Le ciblage de leurs ARNm par dégradation permettra de réduire leurs taux d'expression ou de l'inhiber en réduisant alors leurs effets. Ceci permettra selon notre hypothèse d'avoir moins de survie cellulaire ou une meilleure réponse aux thérapies.

Matériel et méthode

Test des ribozymes dans des cellules exprimant la Gal-7

- **Sélection d'une lignée stable exprimant la Gal-7**

Afin de tester l'activité des ribozymes sur l'ARNm de la Gal-7 dans les cellules, les premiers essais ont été faits avec la lignée cellulaire MDA-MB-468, mais la transfection n'avait pas réussi. Donc la lignée cellulaire HEK293 a été choisie pour générer des transfectants stables exprimant la Gal-7. Pour ce faire, les cellules HEK293 confluentes ont été trypsinées, centrifugées et diluées. Après quantifications, les cellules ont été mises en culture (à 37°C et 5% CO₂) dans une plaque de 12 puits à 120 000 cellules/puits dans du DMEM (*Dulbecco's Modified Eagle Medium*) comprenant 10% (v/v) de sérum fœtal de veau (DMEM complet). Le lendemain les cellules ont été transfectées avec le plasmide pET-22b portant le gène LGALS7 codant la Gal-7 humaine. Comme agent de transfection, la lipofectamine 3000 (invitrogen) a été utilisée et la transfection a été faite selon le protocole du fournisseur.

Après 48 heures, les cellules ont été trypsinées et diluées (1/2) dans un nouveau milieu contenant la puromycine à une concentration finale de 3 µg/ml. Les cellules ont été mises en culture et le milieu a été changé au besoin. Comme témoins, des cellules non transfectées et mises en culture dans du DMEM complet avec ou sans puromycine ont été utilisées.

Après deux semaines, deux à quatre colonies/puits sont apparues dans six puits. Chaque colonie a été repiquée sous microscope et transférée à un puits d'une plaque de 6 puits. Les cellules ont été gardées en culture jusqu'à atteindre 100% de confluence. À confluence, les cellules de chaque puits ont été trypsinées, un petit volume a été utilisé pour le passage de cellule dans une nouvelle plaque de six puits et le reste a été utilisé pour l'analyse de l'expression de la Gal-7 par les cellules.

- **Immunobuvardage western**

Un volume de 1.5 ml de cellules confluentes issues de chaque colonie a été utilisé pour faire l'extraction de protéine. Les cellules ont été centrifugées 15 min à 4°C à 5000 rpm et le surnageant a été jeté. Le tampon de lyse (150 mM NaCl, 1%(v/v) Nonidet P-40, 0.5% (p/v) déoxycholate, 50 mM Tris pH 7.6, 0.1% (p/v) SDS) a été rajouté dans chaque tube et incubé sur glace pendant 30 minutes et vortexé à chaque 5 minutes. Ensuite les lysats ont été centrifugés pendant 15 minutes à 15000 rpm à 4°C et le surnageant de chaque tube a été conservé.

Pour déterminer la concentration des protéines issues de chaque clone un essai de Bradford a été réalisé. Des aliquots de 20 µg de protéines de chaque clone ont été diluées 1 :1 (v/v) avec le tampon d'échantillon (95% tampon Laemmli 2X, 5% 2-mercaptoéthanol) et dénaturés à 95°C pendant 3 min. Les échantillons ont ensuite été centrifugés pendant 1 minute à 15000 rpm et le surnageant récupéré pour l'analyse par SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis 15% (v/v)). La migration sur gel a été faite pendant 1h30 minutes (15 minutes à 100V et 1h15 à 150V). Suite à la migration le gel a été incubé avec une membrane PVDF (Polyvinylidene difluoride) pendant 1h dans le tampon de transfert (39 mM Glycine, 48 mM Tris, 20% méthanol) suivie de la coloration au rouge de ponceau permettant la validation du transfert des protéines. Après le transfert, la membrane a été bloquée dans le tampon de blocage (5% lait dans TBS (Tris-buffered saline)-Tween20 0.1%) pendant une heure. L'incubation a été suivie par un lavage de la membrane quatre fois pendant 15 minutes avec du TBS-Tween 0.1%.

L'anticorps primaire de Gal-7 (anti-Galectine-7 humain-R&D systems) et celui de l'actine (anti-béta actin de la souris, sigma) ont été incubés avec la membrane à 4°C toute la nuit. La membrane a été lavée quatre fois avec le même tampon de lavage avant d'être incubée avec l'anticorps secondaire pendant une heure et lavée encore une fois. 2 ml du réactif

ECL (enhanced chemiluminescence) a été rajoutée à la fin et la révélation des résultats a été faite avec un chemi-doc (Bio-Rad).

- **Préparation des plasmides exprimant les ribozymes**

Des ribozymes montrant une haute activité catalytique ont été choisis et clonés dans deux plasmides différents. Les ribozymes 2, 4 et 8 contre l'ARNm de la galectin7 et le ribozyme 43 contre l'ARNm Bcl-2 ont été clonés chacun dans le plasmide pUC-KE-tRNA-CTE (Nawrot *et al.*, 2003). Ce plasmide contient une cassette d'expression avec un promoteur humain de l'ARN de transfert valine avec une séquence en plus en 3' inhibant la maturation de l'ARNt (séquences verte). Ceci fait que le ribozyme sera transcrit par l'ARN polymérase III et transporté du noyau au cytoplasme par le biais de l'ARNt sans être polyadénylé comme un ARNm. Le CTE (*Constitutive Transport Element*) est un élément de transport viral capable de lier des hélicases. Cet élément est rajouté en 3' de chaque ribozyme pour aider à l'ouverture des ARNm cibles des ribozymes (séquence rose). Chaque ribozyme a été cloné entre les sites BstBI et KpnI.

Une autre construction a été préparée en assemblant deux ribozymes (11 et 14) sur la même cassette d'expression. Le ribozyme 11 a été précédé par la séquence du promoteur de l'ARNt valine et en amont de la séquence du CTE. Le ribozyme 14 avait la séquence du promoteur de l'ARNt glycine en 5' et en 3' une autre séquence CTE (génomme de souris). Cette cassette a été commandée déjà clonée dans un vecteur de chez Twist biosciences. Les séquences des constructions ainsi que les amorces utilisées pour l'ajout des sites de restrictions sont dans **la table S1**

Chaque ribozyme a été amplifié à l'aide des amorces permettant de rajouter les sites de restriction (BstBI en 5' et KpnI en 3'). Une double digestion du plasmide ainsi que des inserts a été effectuée avec les enzymes BstBI et KpnI (NEB) selon les indications du manufacturier. Ceci a été suivi d'une purification sur gels d'agarose 1% et une déphosphorylation du plasmide digéré avec de l'antarctique phosphatase afin d'empêcher sa re-fermeture. Ensuite la T4 ADN ligase (NEB) a été utilisé pour la ligation du plasmide avec chaque ribozyme. Afin de valider le clonage des ribozymes et de produire une quantité suffisante, les plasmides ont été transformés dans *E. coli* DH5 alpha par choc thermique. Ensuite, les plasmides ont été extraits et envoyés au séquençage Sanger (Genome Québec) afin de vérifier la séquence des clonages effectués. Les trois clonages ont fonctionné.

Table S1 Cassette d'expression de ribozymes contre la Gal-7

Nom	Séquence
Cassette ribozyme 2	ACCC TCGAA ACCGGGC ACTACAAAAACCAAC GCAGTGAGCTGATGAGAA CAAACCCGCCCTCGGGCTCGAAACTTGTGGGGGACGT GGTACC CGTA
F Rz 2	ACCC TCGAA ACCGGGC ACTACAAAAACCAAC GCAGTGAGCTGATGAGAA
R Rz 2	TACGGGTACCACGTCCCCACAAGTTTC
Cassette ribozyme 4	ACCC TCGAA ACCGGGC ACTACAAAAACCAAC GCTGTTGAACTGATGAGA ACAAACCCGCTCCTTGCCTCGAAACCACCTCCG GGTACC CGTA
F Rz 4	ACCC TCGAA ACCGGGC ACTACAAAAACCAAC GCTGTTGAACTGATGAGA
R Rz 4	TACGGGTACCCGGAGGTGGTTTCGACGC
Cassette ribozyme 8	ACCC TCGAA ACCGGGC ACTACAAAAACCAAC CGTGGTTAGGGCCACGTT AAATAGGCCCTTAAGCCCTAAGCGCTCCACCAGGCGCACG GGTACC CGTA
F Rz 8	ACCC TCGAA ACCGGGC ACTACAAAAACCAAC CGTGGTTAGGGCCACGTT
R Rz 8	TACGGGTACCCGTGCGCCTGGTGGAGCG
Cassette ribozyme 11+14	GATTCAGGACTAGTCTTTTAGGTCAAAAAGAAGAAGCTTTGTAACCGTTG GTTTCCGTAGTGTAGTGGTTATCACGTTTCGCCTAACACGCGAAAGGTCCC CGG TCGAA ACCGGGC ACTACAAAAACCAAC CGTGGTTAGGGCCACGTTA AATAGGGGGTTAAGCCCTAAGCGGTTGGACATG GGTACC AGACCACCTCC CCTGCGAGCTAAGCTGGACAGCCAATGACGGGTAAGAGAGTGACATTGTT CACTAACCTAAGACAGGAGGGCCGTCAGAGCTACTGCCTAATCCAAAGAC GGGTAAAAGTGATAAAAATGTATCACTCCAACCTAAGACAGGCGCAGCTT CCGAGGGATTTGGATA.TCAACAAAGCACCAGTGGTCTAGTGGTAGAATAG TACCTGCCACGGTACAGACCCGGGTTTCGATTCCCGGCTGGTGCA ACTAC AAAAACCAAC CCCTCGGGCAAATGTGAGCTGATGAGTCGCTGAAATGCCA CGAAACTTGTGGGGGACGT GGTACC CACTGTGGCTTGGCATGCTAGAGA AGTAGTCAATGACGGGTAAGACTCCCTGGGCGTGCACCAACCTAAGACA GGGATCAAACCAATGTTGTTTGTCTCCCCAGGACAGGTAAGGGGCATTGC TGCAGGGGGCAACCTAAGACAGGCATTCTCTCTGCCAATAAGTAAAGACA GGATATC TTTTTTT

Code de couleurs : rose : séquence CTE, bleu : site de restriction et insertion de ribozymes, vert : séquence altérée empêchant la maturation de l'ARNt, brun : séquence du ribozyme.

- **Transfection des cellules**

Pour les transfections, les cellules HEK293 exprimant la Gal-7 ont été utilisées. Dans une plaque de 24 puits, 120 000 cellules/puits ont été mises en culture une nuit avant la transfection. Le lendemain 1 µg de chaque plasmide ayant la séquence d'un ribozyme a été utilisé pour les transfections transitoires. Les transfections ont été faites selon le protocole du kit de la lipofectamine 3000 (Invitrogen-Thermofisher scientific).

- **Extraction d'ARN et RT-qPCR**

Après 48 h de culture, l'ARN total des cellules dans chaque puits a été extrait à l'aide du kit purelink RNA mini kit (Invitrogen-Thermofisher scientific). L'ARN extrait de chaque cellule a été traité avec 1 µl de Dnase I (NEB) pendant 20 minutes à 37°C. Des amorces spécifiques ont été utilisées pour détecter le taux d'expression de chaque ribozyme et la cible Gal-7 (**Table S2**).

Table S2 Amorces de RT-PCR des ribozymes et Gal-7

Nom	Séquence	Tailles amplicons
F Gal-7Q	ATGTCCAACGTCCCCACAAGT	120 nct
R Gal-7 Q	CCCGCACAGCAGGTTTA	
F Rz2 Q	ACCGGGCACTACAAAACCA	-
R Rz2 Q	CGTCCCCACAAGTTTCGAC	76 nct
R Rz4 Q	GGAGGTGGTTTCGACGCAAG	72 nct
R Rz8 Q	GTGCGCCTGGTGGAG	78 nct
F Rz11 Q	CGAAACCGGGCACTACAA	78 nct
R Rz11 Q	CCCATGTCCAACCGCTTA	
F Rz14 Q	TACAGACCCGGGTTTCGATT	106 nct
R Rz14 Q	TCGCATTCAGCGACTCATC	
F Bcl-2 Q	TGCGGCCTCTGTTTGATT	104 nct
R Bcl-2 Q	CACTTGTGGCCAGATAGG	

Pour les réactions de RT-qPCR le kit 'Luna Universal one step RT qPCR' a été utilisé. Les réactions ont été faites pour détecter le taux d'expression des ribozymes ainsi que leurs

cibles. Comme témoin négatif, pour les cellules transfectées avec les ribozymes ciblant la Gal-7, le taux d'expression de Bcl-2 a été vérifié.

Résultat

Nous avons validé la transfection de la Gal-7 en détectant les niveaux d'expression de la Gal-7 chez différents clones de transfectants stables par immunobuvardage. Parmi les six clones, cinq exprimaient notre protéine d'intérêt (**Figure S1**). Le clone 3 a été choisi pour faire les tests de ribozymes, car il semblait exprimer le taux le plus élevé de Gal-7.

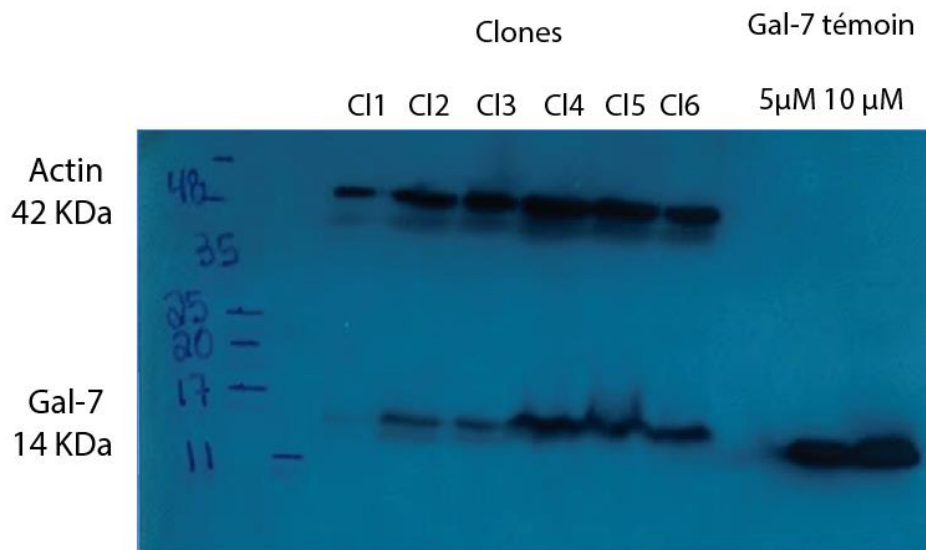


Figure S1 Résultat d'immunobuvardage western de Gal-7

- **Résultat de transfection et RT-PCR**

Les résultats de RT-qPCR ne montraient pas un changement du taux d'expression des ARNm ciblés (Gal-7 et Bcl-2), plus de tests doivent être faits pour valider les résultats. Un immuno-buvardage western donnera un résultat plus visuel du taux d'expression des protéines ciblées.

Annexe II

Librairie de 1000 ribozymes

Matériels et méthodes

- **Préparation de bibliothèques d'ADN de 1000 ribozymes**

Les séquences de 1000 ribozymes ont été commandées de chez GenScript. Afin de les amplifier tous dans la même réaction de PCR, des régions communes aux extrémités 3' et 5' ont été utilisées pour créer des amorces. À l'amorce sens la séquence du promoteur T7 a été rajoutée. Les amorces ainsi que les séquences des miARN ont été commandées de chez IDT (**Table S3**).

Table S3 Séquences miARN et amorces

Nom	Séquence
In1-miARN 18	UAAGGUGCAUCUAGUGCAGAUAG
In2-miARN 155	UUA AUGCUAAUCGUGAUAGGGGUU
In3-miARN 135	UAUGGCUUUUCAUCCUAUGUGA
F librairie Rzs	TAATACGACTCACTATAGGGGGTGC GACTGACG
R librairie Rzs	CCCATAGTAGTGCGA
R+Adaptateur	GCTGACTGTACTGGGCCGAATCGCACTACTATGGG

L'amplification a été faite à l'aide d'une réaction PCR en utilisant l'enzyme ADN Taq polymérase (NEB) en présence de son tampon réaction (1X) avec 2 mM de dNTP. Les tailles des produits ont été vérifiées par électrophorèse sur gel d'agarose 2%. L'ADN a été finalement purifié avec une précipitation à l'éthanol.

- **Préparation des ARN radio-marqués et froids**

Les produits de réactions de clivage ont été utilisés pour la génération de l'ADN qui était envoyé au séquençage. L'ARN radiomarqué a été utilisé comme un marqueur de taille et les réactions des ribozymes ont été faites avec de l'ARN froid.

L'ADN obtenu après la précipitation a été transcrit en présence ou en absence d'UTP³² avec le tampon de transcription 1X, 10 mM de rNTP, sauf pour l'UTP froid qui était à 4 mM pour la transcription chaude, et les enzymes T7 ARN polymérase, la pyrophosphatase

(Roche) et l'inhibiteur de RNase ribolock (Thermofisher). Les réactions ont été incubées pendant 2h30 à 37°C avant d'ajouter 1 µl de Dnase I (NEB) suivie d'une autre incubation de 30 minutes.

L'ARN chaud obtenu a été purifié par une précipitation à l'éthanol et pour l'ARN froid une extraction phénol/chloroforme a été effectuée avant la précipitation. Ceci a été suivi d'une migration sur gel de polyacrylamide dénaturant 10%. Après la migration une plaque phosphore a été exposée au gel pendant 5 min, ensuite la lecture a été faite au Typhoon FLA 9500 (GE healthcare). L'image obtenue a été utilisée pour couper les bandes du gel correspondant aux ribozymes chauds. Quant aux ARN froids, les bandes ont été visualisées sous UV et les bandes correspondantes ont été coupées. L'ARN a été élué du gel dans un tampon de 0.3 M de NaCl suivi d'une précipitation à l'éthanol.

- **Test des ribozymes avec les différentes combinaisons**

Puisque le nombre de miARN total était 'trois', il y a 'huit' combinaisons possibles (**Table S4**). L'ARN froid a été incubé dans le tampon de clivage contenant les miARN de chaque combinaison. L'ARN chaud a été utilisé comme outil de contrôle permettant le suivi de la migration des ARN non marqués. Ce dernier a été incubé soit sans miARN ou en présence des trois combinés. Les réactions ont été incubées à 37°C pendant une heure, suivi par une migration dans un gel de polyacrylamide dénaturant de 6%. Une cassette phosphore a été exposée au gel après migration pendant 30 minutes. La photo du gel obtenue après le scan au Typhoon a été utilisée pour couper toutes les bandes résultantes de la réaction de clivage des ribozymes (coupés et non coupés) (**Figure 7.3**). Les ARN ont été élués pendant une nuit à 4°C et précipité à l'éthanol le lendemain.

Table 0S4 Les combinaisons possibles de miARN

miARN	miARN 18	miARN 155	miARN 135	ET
Combinaison 1	-	-	-	0
Combinaison 2	+	-	-	0
Combinaison 3	-	+	-	0
Combinaison 4	-	-	+	0
Combinaison 5	+	+	-	0
Combinaison 6	+	-	+	0
Combinaison 7	-	+	+	0
Combinaison 8	+	+	+	1

- **Ligation d'adaptateur et RT-PCR**

Les ARN issus d'un clivage de ribozymes ont un phosphate 2'-3' cyclique. Afin de régénérer la région coupée des ribozymes pour pouvoir amplifier les ribozymes résultant du test de clivage un adaptateur a été utilisé. Avant de liguer l'adaptateur, la T4 PNK a été utilisée afin d'enlever le phosphate cyclique des ribozymes coupés. La réaction a été faite en présence du tampon de la T4 PNK (1X) (NEB) et 1 U de ribolock (Thermofisher) pendant une heure à 37°C. L'enzyme a été inactivée à 65°C pendant 5 minutes. Les ARN déphosphorylés ont été incubés avec une séquence adaptatrice comprenant la séquence manquante du ribozyme clivé (**Figure S2**). La réaction de ligation a été faite à l'aide de la T4 ARN ligase et son tampon (1X) (NEB). Avant d'envoyer les ADN à séquencer, en utilisant les mêmes amorces de la première amplification, les ARN issus de la ligation ont été amplifiés par PCR et précipité à l'éthanol.

1- Incubation de ribozymes avec différentes combinaisons de mi-ARN

→ 2- Récupération des ARN du gel

→ 3- Phosphorylation puis ligation d'adaptateur

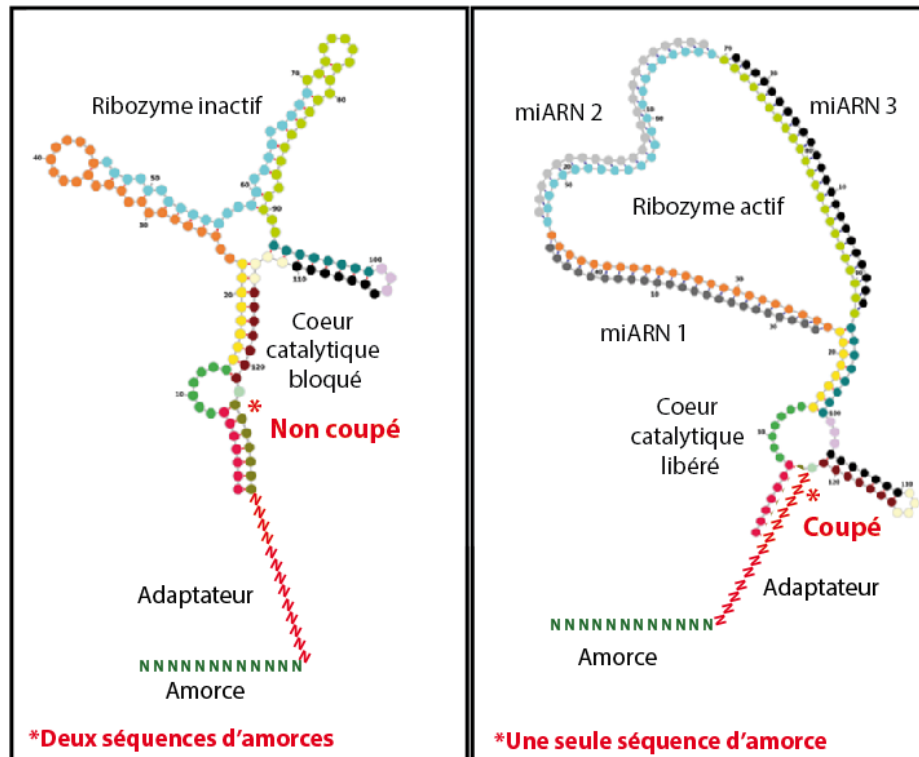


Figure S2 Ligation d'adaptateur après la réaction de clivage.

Suite à la purification de l'ARN après la réaction de clivage, l'adaptateur est rajouté à chaque ribozyme. Pour distinguer les ribozymes actifs de ceux inactifs, l'adaptateur a en 3' la même séquence de l'amorce anti-sens qui correspond à la région coupée. Donc les ribozymes actifs auront un seul exemplaire de l'amorce sens en 3' alors que ceux inactifs auront deux.

Résultats

Les réactions de clivages de ribozymes radioactifs en présence de trois miARN avaient montré une intensité de bandes clivées supérieure à celles des ribozymes sans miARN. (Figure S3)

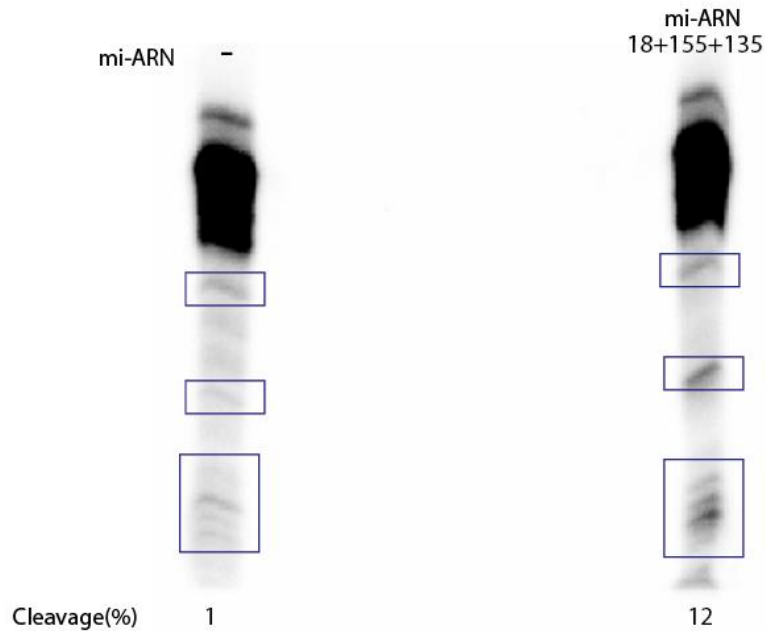


Figure S3 Librairie de ribozymes en absence d'input ou en présence de trois miARN

Ce résultat était encourageant, mais il n'était malheureusement pas possible de confirmer cela avec les résultats de séquençage. Plusieurs problèmes sont survenus avec le séquençage et les milliers de séquences obtenues pour chaque combinaison, dont bon nombre n'avaient pas la bonne taille, ce qui a rendu impossible d'analyser ces résultats et d'en obtenir une conclusion logique. Le séquençage est à refaire mais les résultats ne sont disponibles qu'après quelques mois.

Table S5 Séquence des portes logique AND et NOT-AND

Nom	Séquence
3 AND 1	GGCACUGGGCUGACGAGGCGCUAGUUUUUGUUUGUCUGGGUGCGUUUUUGUCUGG UUCUUAUUAUGAUUAGUGUUGGUUCUCACAUGGGAGUGAAAAGUCGUGGGGCUAG UGUCGAAAUAUUGAAUACGGUUGUACUCAGUG
3 AND 2	GGGUCUGGGCUGAAGAGCGGUUAGUGUUUGUUUGCAUUGGGUGUAUUUUAGUUAG UUUUUGUCGCGAUUGGUGUUGAGUUUCGCAUGGGGAUGGAAGGUUGUAUCCCUAA UUGCGAAAGUUGACGUGCGUUGACUACCCAGAC
3 AND 3	GGGCAGUGGCUGAAGAGCGGUUAGUUUUUGUUUGCAUUAGAUGCAUUUUUGUACAG UUUUUGUUGCGAUUAGUGUUGAUACUCGCGUGGGGGUGGAAGGUUGUGUUUCUAA UCGCGAAACAGAACCACCGUUCUGUACCCAGUC
3 AND 4	GGGCUGUGGCUGAAGAGCGAUUAGGCUCUGUUUGCAUUAGAUGCACUUUAGGUGG UUUUUGUCGCGAUUAGUGUUGAGUAUCGCAUAGGGAUGAAAGGUUAUACACCUGA UUGCGAAACUACGCGAAAGCGUAGUACCAUAGC
3 AND 5	GGGCAGUGGCUGAAGAGCAUUUAGGCACUGUCUGCAUUAGGUGCAUUUUUAUAGG UUUUUGUCGCGAUUAGUGUUGAAUUUCGCAUGGGGAUGGAAGGUCAUAUCACUGA GUGCGAAACUGCUCGACAGGGCAGUACCAUUGC
3 AND 6	GGGCAAAGGCUGACGAGCGGUUAGGUUUUGUCUGCGUUAGAUGCAUCUUAUGUAG UCUUUGUUGCGAUUGGUGUUAGAUUUUCGCAUAGGGAUGGAAGGUUGUGCAUCUGA UCGCGAAACCAUCGCAAGGUUGGUACCUUUGC
3 AND 7	GGAUCGUGGCUGAAGAGUUGCUACGUCCUGUCUGCAUUAGAUGCGUCUUGGUCAG UUUCUGUCAUGGUUAGCGUUGGGUAUCGUAUGGGGGUGGAAAGUCAUGGUUGUAG UAACGAAACAUCCCAUAGGAGUGUACCAUGAU
2 AND 1	GGGGCCUUGCUGAGGAGCCCUUGGGCCCUAUCUGUGCUAGGUGCGUCUUGUUUAA CCCUAUCGUGGUUAGCGUUGACAGCCUGGGGCGAAACGCUUCGAACGAGGUGUA CAAGGCC
2 AND 2	GGGGCCUUGCUGAAGAGGCAGGGGUCUCUAUCUGUGCUAGGUGCAUCUUGUUGAG CCCUUGUCGUGGUUAGCGUUGAUCGCCUCUGCCGAAACAUUUCGAACGAGAUGUA CAAGGCC
2 AND 3	GGGCACCGGCUGAAGAGGUACCUGGGACUAUUUGUGUUAGGUGUGUCUUGGGUGG CUCCUGUUACGGUUGGUGUUGGUGUCAGGUGCCGAAACCACCCACUUGGGUGGUA CCGGUGC
2 AND 4	GGGGCCUUGCUGAAGAGCCAGGGGCCUCUAUCUGUGCUAGGUGCAUCUUGUUUAG CCCUUGUCGUGGUUAGCGUUGAUCGCCUCUGGCGAAACAUUUCGAACGAGAUGUA CAAGGCC
2 AND 5	GGGCAAGGGCUGAGGAGCGGGGCAUUUUUGUUUGUGCUAGGUGUAUCUUGAGGGG CCCUUGUCGCGGUUAGCGUUGAGGUGCCCUUGCGAAAGGAUGCUUCUGCAUCCUA CCCUUGC

2 AND 6	GGGCCAGGGCUGAGGAGCUAGUGUGGAUUAUCUGUGUUGGGUGUGUUUUGUCGAG CUUCUGUUAUGGUUAGCGUUA AAAUACGCUGGGCGAAACGAUCCGCGUGGGUUGUA CCCUGGC
2 AND 7	GGCAGAGUGCUGAUGAGGGGUGUUUUUUAUCUGUGCUGGGUGUAUUUUGCGAAG CCUCUGUCGUGAUUAGCGUUGAGUGGGCACUCGAAUAUCGCGAAGGCGGUGUC CACUCUG
2 NAND 1	GGGACGGGGCUGACGAGGCGUCGAACAUCUCAGCAGGACUAUCUGCGCUGGAUGU ACUUUAGGAGACUCCUGUUGUGGUUGGUAUUGGCCUUCGACGCCGAAACAUCUG CUGAGGAUGUACCCCGUC
2 NAND 2	GGGUAGUGGCUGAAGAGAUGCCGCAGGUCGCCGCAGGGCUAUCUGUAUUAGAUGU ACCUUACGCAACUCCUGUUGUGAUUGGUAUUGGACUGCGGCGUCGAAACGGGCUU UUUAGUCCGUACCACUAC
2 NAND 3	GGGCCUGCGCUGAUGAGUCGUUCCAUGUCGGCGCGGGACUAUCUGUGUUGGAUGU ACCUUAUAAGACUCCUGUUAACGAUUGGUAUUGACCUGGGACGGCGAAAGACGGCU GCCGCCGUCUACGUAGGC
2 NAND 4	GGGCAGGUCCUGAGGAGGGGCCACGAGGCUGCUUCAGCCUGUUUGUCUAGGUGC ACUUUAUCAAUCUCUGUCACGAUUGGUGUAGGGAGUGGCCUCGAAACUGCGAG AUUUCGCGGUAGACUUGU
2 NAND 5	GGGCGUGUGCUGAAGAGUUGUCCACGCCGACGCAGGGCUAUCUGCGUUGGAUGU ACCUAAUAAACUCCUGUUGCGGUUGGUAUUGGCCUGGGACGGCGAAAGACGGCA GCGGCCGUCUACAUACGC
2 NAND 6	GGGGCUGUGCUGAAGAGCUCCACUUGUCCGCCGGAUUGUUUGUGUUGGAUGU ACUUUAUUAGAUCUUUGUUGUGAUUGGCAUUGGAUCGUGGGGGCGAAACUAGAUG GAGAUUCUAGUACAUGGCC
2 NAND 7	GGACAGUGCCUGAAGAGGACGGGCGUUGGCCGGCUGGAUUGUUUGUAUUAGAUGU AUCUUGCCGGAUCCUUGUUGUGGUUAGCGUUGAACAGCCCGUCCGAAACCGGUCC UAAGACCGGUAGUGCUGU

Table 0S6 Séquences des matrices et amorces pour l'amplification des portes logiques

Nom	Séquence
3 AND 1	ACCCGATGGATACAGTCGCATAATACGACTCACTATAGGGGCACTGGGCTG ACGAGCGCTAGTTTTTGTGTTGTCTGGGTGCGTTTTGTCTGGTCTTATT ATGATTAGTGTTGGTTCTCACATGGGAGTGAAAAGTCGTGGGGCTAGTGTC GAAATAATTGAATACGGTTGTACTCAGTG
3 AND 2	ACCCGATGGATACAGTCGCATAATACGACTCACTATAGGGGGTCTGGGCTG AAGAGCGGTTAGTGTTGTTTTGCATTGGGTGTATTTTAGTTAGTTTTTGTCTC GCGATTGGTGTTGAGTTTCGCATGGGGATGGAAGGTTGTATCCCTAATTGC GAAAGTTGACGTGCGTTGACTACCCAGAC
3 AND 3	ACCCGATGGATACAGTCGCATAATACGACTCACTATAGGGGGCAGTGGCTG AAGAGCGGTTAGTTTTTGTGTTGCATTAGATGCATTTTGTACAGTTTTTGT GCGATTAGTGTTGATACTCGCGTGGGGGTGGAAGGTTGTGTTTCTAATCGC GAAACAGAACCACCGTTCGTACTACTGC
3 AND 4	ACCCGATGGATACAGTCGCATAATACGACTCACTATAGGGGGTGTGGCTG AAGAGCGATTAGGCTCTGTTTGCATTAGATGCACTTTAGGTGGTTTTTGTCTC GCGATTAGTGTTGAGTATCGCATAGGGATGAAAGGTTATACACCTGATTGC GAAACTACGCCAAAGCGTAGTACCATAGC
3 AND 5	ACCCGATGGATACAGTCGCATAATACGACTCACTATAGGGGGCAGTGGCTG AAGAGCATTTAGGCACTGTCTGCATTAGGTGCATTTTATAAGGTTTTTGTCTC GCGATTAGTGTTGAATTCGCATGGGGATGGAAGGTCATATCACTGAGTGC GAAACTGCTCGACAGGGCAGTACCATTGC
3 AND 6	ACCCGATGGATACAGTCGCATAATACGACTCACTATAGGGGGCAAAGGCTG ACGAGCGGTTAGGTTTTTGTCTGCGTTAGATGCATCTTATGTAGTCTTTGTT GCGATTGGTGTTAGATTTTCGCATAGGGATGGAAGGTTGTGCATCTGATCGC GAAACCAATCGCAAGGTTGGTACCTTTGC
3 AND 7	ACCCGATGGATACAGTCGCATAATACGACTCACTATAGGGGGATCGTGGCTG AAGAGTTGCTACGTCTGTCTGCATTAGATGCGTCTTGGTCAGTTTCTGTCTC ATGGTTAGCGTTGGGTATCGTATGGGGGTGAAAAGTCATGGTTGTAGTAAC GAAACATTTCCCATAGGAGTGTACCATGAT
2 AND 1 + 2 NAND 5	ACCCGATGGATACAGTCGCATAATACGACTCACTATAGGGGGCCCTTGCTG AGGAGCCCTGGGCCCTATCTGTGCTAGGTGCGTCTTGTTTTAACCCCTATC GTGGTTAGCGTTGACAGCCTGGGGCGAAACGCTTTCGAACGAGGTGTACAAG GCCGTCTACGACAGCTACATATATAATACGACTCACTATAGGGGGCCTGTG CTGAAGAGTTGTTCCACGCCGACGCAGGGCTATCTGCGTTGGATGTACCTT AATAAACTCCTGTTGCGGTTGGTATTGGCCTGGGACGGCGAAAGACGGCAG CGGCCGTCTACATACGC
2 AND 2 + 2 NAND 6	ACCCGATGGATACAGTCGCATAATACGACTCACTATAGGGGGCCCTTGCTG AAGAGGCAGGGGTCTCTATCTGTGCTAGGTGCATCTTGTTGAGCCCCGTGC GTGGTTAGCGTTGATCGCCTCTGCCGAAACATTTTCGAACGAGATGTACAAG GCCGTCTACGACAGCTACATATATAATACGACTCACTATAGGGGGCCTGTG

	CTGAAGAGCTCCCACTTGTCCGCCGCGATTGTTTGTGTTGGATGTACTTT ATTAGATCTTTGTTGTGATTGGCATTGGATCGTGGGGGCGAAACTAGATGG AGATCTAGTACATGGCC
2 AND 3 + 2 NAND 7	ACCCGATGGATACAGTCGCATAAATACGACTCACTATAGGGGGCACCGGCTG AAGAGGTACCTGGGACTATTTGTGTTAGGTGTGTCTTGGGTGGCTCCTGTT ACGGTTGGTGTGGTGTGTCAGGTGCCGAAACCACCCACTTGGGTGGTACCGG TGCGTCTACGACAGCTACATATATAAATACGACTCACTATAGGGGACAGTGC CTGAAGAGGACGGGCGTTGGCGGGCTGGATTGTTTGTATTAGATGTATCTT GCCGGATCCTTGTGTTGGTTAGCGTTGAACAGCCCGTCCGAAACCGTCTCT AAGACCGGTAGTGCTGT
2 AND 4 + 2 AND 5	ACCCGATGGATACAGTCGCATAAATACGACTCACTATAGGGGGGCTTGCTG AAGAGCCAGGGGCTCTATCTGTGCTAGGTGCATCTTGTTTAGCCCTGTCTC GTGGTTAGCGTTGATCGCCTCTGGCGAAACATTTGAAACGAGATGTACAAG GCCGCTACGACAGCTACATATATAAATACGACTCACTATAGGGGGCAAGGG CTGAGGAGCGGGGCATTTTGTGTTGTGCTAGGTGTATCTTGGGGGCCCTT GTCCGGTTAGCGTTGAGGTGCCCTTGCAGAAAGGATGCTTCTGCATCCTAC CCTTGC
2 AND 6 + 2 AND 7	ACCCGATGGATACAGTCGCATAAATACGACTCACTATAGGGGGCCAGGGCTG AGGAGCTAGTGTGGATTATCTGTGTTGGGTGTGTTTTGTGCGAGCTTCTGTT ATGGTTAGCGTTAAAAACGCTGGCGAAACGATCCGCGTGGGTTGTACCCT GGCGTCTACGACAGCTACATATATAAATACGACTCACTATAGGGGCAGAGTG CTGATGAGGGGTGTTTTTTTATCTGTGCTGGGTGTATTTTGCAGAAAGCTCT GTCGTGATTAGCGTTGAGTGGGCACTCCGAAATATCGCGAAGGCGGTGTCC ACTCTG
2 NAND 1 + 2 NAND 2	ACCCGATGGATACAGTCGCATAAATACGACTCACTATAGGGGGACGGGGCTG ACGAGGCGTGAACATCTCAGCAGGACTATCTGCGCTGGATGTACTTTAGG AGACTCCTGTTGTGGTTGGTATTGGCCTTCGACGCCGAAACATTTCTGCTGA GGATGTACCCCGTCTGCTACGACAGCTACATATATAAATACGACTCACTATA GGGGGTAGTGGCTGAAGAGATGCCGCAGGTGCGCCGAGGGCTATCTGTATT AGATGTACCTTACGCAACTCCTGTTGTGATTGGTATTGGACTGCGGCGTGC AAACGGGCTTTTTTAGTCCGTACCACTAC
2 NAND 3 + 2 NAND 4	ACCCGATGGATACAGTCGCATAAATACGACTCACTATAGGGGGCCTGCGCTG ATGAGTCGTTCCATGTCGGCGCGGGACTATCTGTGTTGGATGTACCTTATA AGACTCCTGTTACGATTGGTATTGACCTGGGACGGCGAAAGACGGCTGCCG CCGCTACGTAGGCGTCTACGACAGCTACATATATAAATACGACTCACTATA GGGGGCAGGTCTGAGGAGGGGCCACGAGGCTGCTTCAGCCTGTTTGTGCT AGGTGCACTTTATCAAATCTCTGTACAGATTGGTGTAGGGAGTGGCCTCG AAACTGCGAGATTTGCGGGTAGACTTGT
Adaptateur 1	ACCCGATGGATACAGTCGCA
Adaptateur 2	GTCTACGACAGCTACATATA

R 3 AND 1	CACTGAGTACAACCGTATTCA
R 3 AND 2	GTCTGGGTAGTCAACGCACG
R 3 AND 3	GCAGTGGTACAGAACGGTGG
R 3 AND 4	GCTATGGTACTACGCTTTCG
R 3 AND 5	GCAATGGTACTGCCCTGTCG
R 3 AND 6	GCAAAGGTACCAACCTTGCG
R 3 AND 7	ATCATGGTACACTCCTATGG
R 2 AND 1-2	GGCCTTGTACACCTCGTTCG
R 2 AND 3	GCACCGGTACCACCCAAGTG
R 2 AND 4	GGCCTTGTACATCTCGTTCG
R 2 AND 5	GCAAGGGTAGGATGCAGAAG
R 2 AND 6	GCCAGGGTACAACCCACGCG
R 2 AND 7	CAGAGTGGACACCGCCTTCG
R 2 NAND 1	GACGGGGTACATCCTCAGCA
R 2 NAND 2	GTAGTGGTACGGACTAAAAA
R 2 NAND 3	GCCTACGTAGACGGCGGCAG
R 2 NAND 4	ACAAGTCTACCGCGAAATCT
R 2 NAND 5	GCGTATGTAGACGGCCGCTG
R 2 NAND 6	GGCCATGTACTAGATCTCCA
R 2 NAND 7	ACAGCACTACCGGTCTTAGG

Annexe III

- Matrice utilisé par Ribosoft pour la génération de ribozyme liant un miARN

Ribozyme pour miARN 18

- ✓ Matrice du ribozyme :

ACCGTTGGTTTTCCGTAGTGTAGTGGTTATCACGTTTCGCCTAACACGCGAAAGGTCCCCGGTTCGAAACCGG
...((((((((...((((((...(((...)))...((((...)))...))))))))))

GC-**ACTACAAAAACCAAC**NNNNNNNNCUGAUG**AGA**ACAAACCC**NNNNNNNN**CG**UCGAA****ACNNnnnnnnnnn**

))-)))))...)))))01234567.....((.....89abcdef..))...ghijklmnopqrs
nnCUAUCUGCACUAGAUGCACCUAA

tu.....

- ✓ Matrice de la cible :

nnnnnnnnnnNNGUCGNNNNNNNNNNNNNNNN

utsrqponmlkjihg.76543210fedcba98

Le code des couleurs définit les régions complémentaires sur le ribozyme et sa cible.

La séquence **rouge** est la partie complémentaire à la séquence du miARN 18.

La séquence **brune** est la séquence de l'ARNt qui permettra l'expression et le transport du ribozyme dans le cytoplasme cellulaire.

Table S7 Séquences de ribozymes liant des miARN en 3'

Nom	Séquence
Rz 1 miARN 18	ACCGUUGGUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCG AAAGGUCCCCGGUUCGAAACCGGGCACUACAAAAACCAACUCGUGGCC CUGAUGAGAACAACCCUGCCCCAGCGUCGAAACGAAGAUGAUCUUC UAUCUGCACUAGAUGCACCUUA
Rz 2 miARN 18	ACCGUUGGUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCG AAAGGUCCCCGGUUCGAAACCGGGCACUACAAAAACCAACUCGUGGCC CUGAUGAGAACAACCCUGCCCCAGCGUCGAAACGAAGAUGAUCUUC AUCUGCACUAGAUGCACCUUA
Rz 3 miARN 18	ACCGUUGGUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCG AAAGGUCCCCGGUUCGAAACCGGGCACUACAAAAACCAACUCGUGGCC CUGAUGAGAACAACCCUGCCCCAGCGUCGAAACGAAGAUGAUCUCUA UCUGCACUAGAUGCACCUUA
Rz 1 miARN 155	ACCGUUGGUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCG AAAGGUCCCCGGUUCGAAACCGGGCACUACAAAAACCAACUCGUGGCC CUGAUGAGAACAACCCUGCCCCAGCGUCGAAACGAAGAUAACCCCU AUCACGAUUAGCAUUA
Rz 2 miARN 155	ACCGUUGGUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCG AAAGGUCCCCGGUUCGAAACCGGGCACUACAAAAACCAACUCGUGGCC CUGAUGAGAACAACCCUGCCCCAGCGUCGAAACGAAGUAACCCCUA UCACGAUUAGCAUUA
Rz 3 miARN 155	ACCGUUGGUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCG AAAGGUCCCCGGUUCGAAACCGGGCACUACAAAAACCAACUCGUGGCC CUGAUGAGAACAACCCUGCCCCAGCGUCGAAACGAAGAAACCCCUA CACGAUUAGCAUUA
Rz 1 miARN 135	ACCGUUGGUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCG AAAGGUCCCCGGUUCGAAACCGGGCACUACAAAAACCAACUCGUGGCC CUGAUGAGAACAACCCUGCCCCAGCGUCGAAACGAAGAUGAUCUUC CACAUAGGAAUGAAAAGCCAUA
Rz 2 miARN 135	ACCGUUGGUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCG AAAGGUCCCCGGUUCGAAACCGGGCACUACAAAAACCAACUCGUGGCC CUGAUGAGAACAACCCUGCCCCAGCGUCGAAACGAAGAUGAUCUUC ACAUAGGAAUGAAAAGCCAUA
Rz 3 miARN 135	ACCGUUGGUUCCGUAGUGUAGUGGUUAUCACGUUCGCCUAACACGCG AAAGGUCCCCGGUUCGAAACCGGGCACUACAAAAACCAACUCGUGGCC CUGAUGAGAACAACCCUGCCCCAGCGUCGAAACGAAGAUGAUCUUC CAUAGGAAUGAAAAGCCAUA

Table S8 Séquences des matrices et amorces pour l'amplification de ribozymes liant des miARN en 3'

Nom	Séquence
F1 Rz1 miARN18	TAATACGACTCACTATAGGGACCGTTGGTTTCCGTAGTGTAGTGGTTATCA CGTTCGCCTAA
R1 Rz1 miARN18	CGGTTTCGAACCGGGGACCTTTCGCGTGTTAGGCGAACGTGATAACCACTA CACTACGGA
F2 Rz1 miARN18	CCCCGGTTCGAAACCGGGCACTACAAAACCAACTCGTGGCCCTGATGAGAA CAAACCCGCCCCAGC
R2 Rz1 miARN18	TAAGGTGCATCTAGTGCAGATAGGAAGATCATCTTCGTTTCGACGCTGGGG CAGGGTTTGTCT
F1 Rz2 miARN18	TAATACGACTCACTATAGGGACCGTTGGTTTCCGTAGTGTAGTGGTTATCA CGTTCGCCTAA
R1 Rz2 miARN18	GGTTTCGAACCGGGGACCTTTCGCGTGTTAGGCGAACGTGATAACCACTAC ACTACGGAA
F2 Rz2 miARN18	CCCCGGTTCGAAACCGGGCACTACAAAACCAACTCGTGGCCCTGATGAGA ACAAACCCGCCCCAGC
R2 Rz2 miARN18	TAAGGTGCATCTAGTGCAGATAGAAGATCATCTTCGTTTCGACGCTGGGGC AGGGTTTGTCTC
F1 Rz3 miARN18	TAATACGACTCACTATAGGGACCGTTGGTTTCCGTAGTGTAGTGGTTATCA CGTTCGCCTAA
R1 Rz3 miARN18	GGTTTCGAACCGGGGACCTTTCGCGTGTTAGGCGAACGTGATAACCACTAC ACTACGGAA
F2 Rz3 miARN18	CCCCGGTTCGAAACCGGGCACTACAAAACCAACTCGTGGCCCTGATGAGA ACAAACCCGCCCCAGC
R2 Rz3 miARN18	TAAGGTGCATCTAGTGCAGATAGAGATCATCTTCGTTTCGACGCTGGGGCA GGGTTTGTCT
F1 Rz1 miARN155	TAATACGACTCACTATAGGGACCGTTGGTTTCCGTAGTGTAGTGGTTATCA CGTTCGCCT
R1 Rz1 miARN155	GGTTTCGAACCGGGGACCTTTCGCGTGTTAGGCGAACGTGATAACCACTAC ACTACGGAA
F2 Rz1 miARN155	AACACGCGAAAGGTCCCCGGTTCGAAACCGGGCACTACAAAACCAACTCG TGGCCCTGA
R2 Rz1 miARN155	TTAATGCTAATCGTGATAGGGGTTTCATCTTCGTTTCGACGCTGGGGCAGGG TTTGTCTCATCAGGGCCACGAGTT

F1 Rz2 miARN155	TAATACGACTCACTATAGGGACCGTTGGTTTTCCGTAGTGTAGTGGTTATCA CGTTCGCCT
R1 Rz2 miARN155	GGTTTTCGAACCGGGGACCTTTTCGCGTGTTAGGCGAACGTGATAACCACTAC ACTACGGAA
F2 Rz2 miARN155	CCCCGGTTCGAAACCGGGCACTACAAAAACCAACTCGTGGCCCTGATGAGA ACAAACCCTGCCCA
R2 Rz2 miARN155	TTAATGCTAATCGTGATAGGGGTTATCTTCGTTTTCGACGCTGGGGCAGGGT TTGTTCTCAT
F1 Rz3 miARN155	TAATACGACTCACTATAGGGACCGTTGGTTTTCCGTAGTGTAGTGGTTATCA CGTTCGCCT
R1 Rz3 miARN155	GGTTTTCGAACCGGGGACCTTTTCGCGTGTTAGGCGAACGTGATAACCACTAC ACTACGGAA
F2 Rz3 miARN155	CCCCGGTTCGAAACCGGGCACTACAAAAACCAACTCGTGGCCCTGATGAGA ACAAACCCTGCCCA
R2 Rz3 miARN155	TTAATGCTAATCGTGATAGGGGTTTCTTCGTTTTCGACGCTGGGGCAGGGT TGTTCTCAT
F1 Rz1 miARN135	TTCTAATACGACTCACTATATAATACGACTCACTATAGGGACCGTTGGTTT CCGTAGTGTAGTGGTTATC
R1 Rz1 miARN135	AACCGGGGACCTTTTCGCGTGTTAGGCGAACGTGATAACCACTACACTACGG AAACCAACG
F2 Rz1 miARN135	AACACGCGAAAGGTCCCCGGTTCGAAACCGGGCACTACAAAAACCAACTCG TGGCCCTGA
R2 Rz1 miARN135	TATGGCTTTTTCATTCCTATGTGAGAAGATCATCTTCGTTTTCGACGCTGGGG CAGGGTTTGTTCATCAGGGCCACGAGTT
F1 Rz2 miARN135	TAATACGACTCACTATAGGGACCGTTGGTTTTCCGTAGTGTAGTGGTTATCA CGTTCGCCTAA
R1 Rz2 miARN135	CGGTTTCGAACCGGGGACCTTTTCGCGTGTTAGGCGAACGTGATAACCACTA CACTACGGA
F2 Rz2 miARN135	CCCGGTTTCGAAACCGGGCACTACAAAAACCAACTCGTGGCCCTGATGAGAA CAAACCCTGCCCA
R2 Rz2 miARN135	TATGGCTTTTTCATTCCTATGTGAAAGATCATCTTCGTTTTCGACGCTGGGGC AGGGTTTGTTCCT
F1 Rz3 miARN135	TAATACGACTCACTATAGGGACCGTTGGTTTTCCGTAGTGTAGTGGTTATCA CGTTCGCCTAA
R1 Rz3 miARN135	CGGTTTCGAACCGGGGACCTTTTCGCGTGTTAGGCGAACGTGATAACCACTA CACTACGGA

F2 Rz3 miARN135	CCCGGTTTCGAAACCGGGCACTACAAAAACCAACTCGTGGCCCTGATGAGAA CAAACCCCTGCCCA
R2 Rz3 miARN135	TATGGCTTTTCATTCCATATGTGAAGATCATCTTCGTTTCGACGCTGGGGCA GGGTTTGTCT

Matériels et méthodes des tests *in vivo*

- **Préparation des cassettes d'expression du ribozyme et du miARN**

Suite à la validation de la capacité des ribozymes liés aux séquences de miARN à cliver la cible comme prévu, des cassettes d'expression d'un ribozyme et de différent miARN ont été préparées. Les vecteurs ayant les cassettes d'expression déjà clonés, ont été commandés de chez Twist biosciences.

Les séquences des pri-miARN ont été utilisées pour les constructions, en plus des surplus de séquences, permettant la formation de la structure d'une tige avec des boucles internes, reconnue par Drosha/DGCR8. Cette reconnaissance permet la maturation du pri-miARN en pre-miARN qui sera coupé par Dicer générant ainsi le miARN mature qui rejoindra le complexe RISC (Ely *et al.*, 2009; Han *et al.*, 2006).

La séquence du pri-miARN visée a été clonée en aval de la séquence d'un promoteur CMV (disponible dans le vecteur fourni par Twist bioscience) et suivie de la séquence d'un promoteur d'ARNt qui précède la séquence d'un deuxième miARN. Ceci a été fait pour tester notre capacité d'exprimer plusieurs miARN avec le même promoteur afin d'éviter la co-transfection de plusieurs vecteurs ayant des effets négatifs sur le taux d'expression de différents gènes d'intérêt.

Au final, sur le même vecteur il y avait les pri-miARN-135 et pri-miARN-18 et sur un autre le pri-miARN-18 et le pri-miARN-155. Le deuxième vecteur a été commandé afin de l'utiliser comme un témoin négatif dans la preuve de concept et pour des utilisations futures si des ribozymes liant l'un des deux miARN seront testés. Les cassettes d'expression sont présentées ci-dessous.

Pour la preuve de concept, la séquence d'un seul ribozyme a été choisie pour faire le test d'activité dans les cellules. Pour le ribozyme liant le miARN-135 et ciblant la luciférase, la cassette d'expression est la même que celle utilisée pour les ribozymes ciblant l'ARNm de la Gal-7 avec un promoteur d'ARNt et une séquence CTE à la fin (en rose).

- ✓ Les cassettes d'expression des miARN et ribozymes
- ✓ Cassette d'expression du miARN 135 et miARN 155:

```
GGTTGCTGTTGACAGTGAGCGACCTGTTAATGCTAATCGTGATAGGGGTTTTTGCCTCCTCAACTGA
CTCCTACATATTAGCATTAAACAGTGCCTACTGCCTCGGACTTCATAACAACGAAGAGGGATGGTA
TTGCTCCTGTAACCTCGGAACCTGGAGAGGCACCTCTGCTGTGGCCTATGGCTTTTCATTCCTATGTG
ATTGCTGTCCCAAACCTCATGTAGGGCTAAAAGCCATGGGCTACAGTGAGGGGCGAGCTCCTTTCC
TGTCTGACAGCAGCTTGGCTACCTCCGTCCTGTTCCCTTGTCTTATTTTTTTTT
```

- ✓ Cassette d'expression des miARN 155 et 18 :

```
GGTTGCTGTTGACAGTGAGCGACTGTTCTAAGGTGCATCTAGTGCAGATAGTGAAGTAGATTAGC
ATCTACTGCCCTAAGTGCTCCTTCTGGCAATGCCTACTGCCTCGGACTTCATAACAACGAAGAGGG
ATGGTATTGCTCCTGTAACCTCGGAACCTGGAGAGGCCTGTTAATGCTAATCGTGATAGGGGTTTTTG
CCTCCTCAACTGACTCCTACATATTAGCATTAAACAGTTTCCTGTCTGACAGCAGCTTGGCTACCTCC
GTCCCTGTTCCCTTGTCTTATTTTTTTTT
```

Vecteur : CMV Hygro, Promoteur UbC de chez Twist

Les séquences jaune et verte sont les régions présent du pri-miARN 30 pour le premier miARN et pri-miARN 31 pour le deuxième miARN, reconnues par Drosha/DGCR8 et la queue polyT permettra la terminaison de la transcription.

- ✓ **Cassette d'expression du Rz2-miR-135a :**

```
GATTCAGGACTAGTCTTTTAGGTCAAAAAGAAGAAGCTTTGTAACCGTTGGTTTCCGTAGTGTAG
TGGTTATCACGTTTCGCCTAACACGCGAAAGGTCCCCGGTTCGAAAACCGGGCACTACAAAACCAA
CTCGTGGCCCTGATGAGAACAAACCCTGCCCCAGCGTCGAAACGAAGATGATCTTTCACATAGGA
ATGAAAAGCCATAAGTACCAGACCACCTCCCCTGCGAGCTAAGCTGGACAGCCAATGACGGGTAA
GAGAGTGACATTGTTCACTAACCTAAGACAGGAGGGCCGTCAGAGCTACTGCCTAATCCAAAGAC
GGGTAAAAGTGATAAAAATGTATCACTCCAACCTAAGACAGGCGCAGCTTCCGAGGGATTGAT
ATCTTTTTTT
```

- **Test du ribozyme dans les cellules exprimant la luciférase**

Dans le but d'exprimer la luciférase, plusieurs essais de génération de lignée stable exprimant l'enzyme ont été faites, mais sans réussite. Au final, en absence d'une lignée stable, nous avons choisis de co-transfecter le plasmide exprimant la luciférase avec les autres vecteurs d'intérêt.

Afin de tester l'activité du ribozyme dans les cellules et l'hypothèse de son inhibition à l'aide du complexe RISC en présence du miARN, différentes combinaisons ont été testées.

- La luciférase
- La luciférase + ribozyme contre l'ARNm de la luciférase et liant le miARN 135
- La luciférase + ribozyme non spécifique (contre l'ARNm de la Gal-7)
- La luciférase + ribozyme contre l'ARNm de la luciférase+miARN135
- La luciférase + ribozyme contre l'ARNm de la luciférase+miARN155
- La luciférase + miARN-135

La même concentration (200 ng/μl) a été utilisée pour chaque plasmide pour essayer d'avoir un taux d'expression similaire, bien que ce ne soit pas garanti car l'efficacité de la transfection est très variable. De plus, le fait que deux ou plusieurs gènes ont été exprimés avec le même promoteur même si ce n'est pas sur un seul plasmide, le taux d'expression est certainement variable.

- **Extraction d'ARN et RT-qPCR :**

Les transfections ont été faites en triplicatas, après 48h de transfection, les cellules dans chaque puits ont été lavées au PBS 1X puis l'ARN total a été extrait en utilisant le kit Nucleospin RNA plus de chez Takara Bio.

Pour vérifier la qualité des ARN extrait, le bioanalyzer 2100 de chez Agilent a été utilisé avec le kit 'Agilent RNA 6000 Nano kit' pour détecter l'intégrité des ARN ribosomiaux 28S et 18S.

Les réactions de RT-qPCR ont été faites avec des amorces conçues à l'aide de l'outil 'Primer Quest' de chez IDT où les oligonulécotides ont été commandés (**Table S8**). Le kit 'Luna Universal One step RT-qPCR' a été utilisé et les PCR ont été faites dans une machine de qPCR de chez BioRad.

Table S8 Amorces de RT-PCR des ribozymes et miARN

Nom	Séquence
F luciférase Q	TACCTGACCGCTTGGTTTG
R luciférase Q	GCTCGTAGGCGTAGTGAAAG
F miARN155 Q	CAGTGAGCGACTGTTCTAAGG
R miARN155 Q	CATGCCAGAAGGAGCACTTA
F miARN135 Q	CACTCTGCTGTGGCCTATG
R miARN135 Q	CACTGTAGCCCATGGCTTT
F Rz2 miARN135 Q	CGAAACCGGGCACTACAA
R Rz2 miARN135 Q	ATCATCTTCGTTTCGACGCT
F miARN 18 Q	ACAACGAAGAGGGATGGTATTG
R miARN18 Q	GCTAATATGTAGGAGTCAGTTGGAG

- **Essai de luciférase**

Deux essais de luciférase ont été faits, un avec les cellules vivantes et un avec le lysat cellulaire.

- **Dans les cellules vivantes**

Après 48 h de transfection, les cellules de chaque puits ont été lavées au PBS 1X. Afin de visualiser les noyaux cellulaires une coloration au Hoechst 33342 (Thermofisher scientific) a été effectuée. Les cellules ont été lavées deux fois au PBS, ensuite la luciférine D du kit Pierce™ Firefly Luciferase Glow Assay Kit (thermofisher scientific) a été diluée (1:1000) dans du Opti-MEM. A chaque puits 1 ml de luciférine diluée a été rajouté et les cellules ont été rapidement visualisées sous microscope confocal.

- **Avec le lysat cellulaire**

Les essais ont été selon le protocole du kit Pierce™ Firefly Luciferase Glow Assay Kit (thermofisher scientific). Suite à la lyse cellulaire, 50 µl de la luciférine D diluée dans le tampon d'essai a été mélangé avec 20 µl de lysat cellulaire dans des puits d'une plaque 96 puits. La lecture du résultat a été faite immédiatement avec un lecteur de plaque avec une longueur d'onde de 525 nm.

Résultats

Plusieurs essais ont été faits pour générer une lignée stable exprimant la luciférase, mais aucun n'a réussi. Pour cela nous avons co-transfecté le plasmide ayant le gène de la luciférase avec chaque plasmide des combinaisons mentionnées.

Les résultats de transfection ont été vérifiés avec trois techniques : une RT-qPCR pour chaque gène exprimé, un essai de luciférase après une lyse cellulaire et un essai de luciférase avec des cellules vivantes.

La visualisation de cellules sous microscope confocal après transfection et en présence de la luciférine n'avait pas montré un taux de transfection élevé, sur quelques dizaines de cellules, une ou deux cellules exprimaient la luciférase (montrant une luminescence).

De plus, les résultats de RT-qPCR n'avaient pas montré de différence dans le taux d'expression de la luciférase en présence ou en absence des ribozymes spécifiques ou non spécifiques. Ceci est le même cas pour le ribozyme liant le miARN 135 qui ne montrait pas un changement marquant du taux d'expression dans les différentes conditions.

Les résultats ne sont pas encourageants, mais le fait que tous les plasmides ont été co-transfectés et nos gènes d'intérêt sont exprimés avec le même promoteur (CMV), les conditions de l'expérience n'étaient pas optimales. Une meilleure façon sera d'utiliser des cellules qui expriment déjà la luciférase et le miARN d'intérêt ce qui donnera des résultats probablement plus fiables et valables.