

Université de Sherbrooke

**Investigation de la spécificité nucléotidique de l'hélicase DHX36 lors du déroulement
de structures d'ARN G-quadruplex.**

Par
Julien RAINVILLE SIROIS
Programmes de Biochimie

Mémoire présenté à la Faculté de médecine et des sciences de la santé
en vue de l'obtention du grade de maître ès sciences (M. Sc.)
en Biochimie

Sherbrooke, Québec, Canada
[Octobre, 2016]

Membres du jury d'évaluation
Martin Bisaillon, Biochimie
Guylain Boissonneault, Biochimie
Daniel Lafontaine, Biologie, Sciences, Université de Sherbrooke

RÉSUMÉ

Investigation de la spécificité nucléotidique de l'hélicase DHX36 lors du déroulement de structures d'ARN G-quadruplex.

Par
Julien Rainville Sirois
Programmes de Biochimie

Mémoire présenté à la Faculté de médecine et des sciences de la santé en vue de l'obtention du diplôme de maître ès sciences (M.Sc.) en biochimie, Faculté de médecine et des sciences de la santé, Université de Sherbrooke, Sherbrooke, Québec, Canada, J1H 5N4

La déstabilisation des structures G-quadruplex au niveau des acides nucléiques a des répercussions physiologiques importantes. L'accentuation des connaissances concernant les processus cellulaires associés au métabolisme des structures des G4 est primordiale. Une panoplie d'hélicases à G4 est impliquée dans le métabolisme des structures G4, notamment l'hélicase humaine DHX36. Il a été déterminé au préalable par certains groupes de recherche que l'hélicase DHX36 se lie à son substrat l'ARN G4 et utilise des nucléosides triphosphates afin de catalyser le dépliement de la structure G-quadruplex. Toutefois, l'interaction avec l'ARN G4 a été sommairement caractérisée et la spécificité nucléotidique n'a toujours pas été évaluée. Ainsi, nous avons décidé d'approfondir les connaissances du mécanisme de dépliement de la structure du G4 d'ARN par l'hélicase DHX36. Notamment, en évaluant la thermodynamique de l'interaction entre l'hélicase et l'ARN G4 afin de révéler particulièrement l'efficacité de liaison mais également en évaluant la spécificité nucléotidique de l'hélicase DHX36 afin d'effectuer le dépliement de l'ARN G4. La combinaison des analogues de nucléotides et le modèle structural permettent de révéler les caractéristiques structurales et fonctionnelles de l'interaction entre l'hélicase humaine DHX36 et l'ATP. Nos analyses permettent de constater que l'enzyme DHX36 est en mesure d'utiliser autant l'ATP que GTP afin de dérouler les structures G4 d'ARN ayant, par contre, une spécificité accrue pour la molécule d'ATP.

Mots clés : Hélicase, DHX36, RHAU, ARN G-quadruplex, Analogues, Modélisation.

TABLE DES MATIÈRES

Résumé	ii
Table des matières	iii
Liste des tableaux	v
Liste des figures	vi
Liste des abréviations	vii
Remerciements	xi
Chapitre 1 – Introduction	1
1.1 Les acides nucléiques	1
1.1.1 L'acide désoxyribonucléique (ADN).....	1
1.1.2 L'acide ribonucléique (ARN).....	2
1.1.3 Structures secondaires et tertiaires des acides nucléiques.....	4
1.1.3.1 Structures secondaires et tertiaires de l'ADN.....	4
1.1.3.2 Structures secondaires et tertiaires de l'ARN.....	5
1.2 Les G-quadruplex (G4)	7
1.2.1 Structure.....	7
1.2.2 Localisation.....	8
1.2.3 Fonctions.....	9
1.2.4 Problématique.....	10
1.2.5 Hélicases de G4.....	12
1.2.6 Les hélicases.....	13
1.2.6.1 Familles.....	13
1.2.6.2 Superfamilles 1 et 2.....	14
1.2.6.3 Hélicases DEAD-box et DEAH-box.....	17
1.2.6.4 Structure et mécanisme.....	18
1.2.7 DHX36.....	20
1.2.7.1 Caractéristiques.....	20
1.2.7.2 Structure.....	20
1.2.7.3 Conservation.....	21
1.2.7.4 Fonctions probables.....	22
1.2.7.5 Dépliement des G4 d'ARN.....	23
1.2.7.6 Mécanisme.....	23
1.2.7.7 Liaison de l'ARN G4 et utilisation de l'ATP.....	24

1.3	Hypothèse/problématique	25
1.3.1	Objectifs.....	25
Chapitre 2 – Matériel et méthodes.....		26
2.1	Clonage, expression et purification de DHX36 dans <i>E. coli</i>.....	26
2.2	Clonage, transfection, expression transitoire et purification de DHX36 dans HEK293T.....	27
2.3	Formation du complexe tétramoléculaire d'ARN G4.....	28
2.4	Étiquetage au ³²P de l'extrémité 5' de l'ARN G4.....	29
2.5	Essais de dépliement des structures G4 d'ARN.....	29
2.6	Analogues de nucléotides	30
2.7	Modélisation par homologie.....	30
Chapitre 3 – Résultats.....		31
3.1	Expression, purification et activité de dépliement des G4	31
3.1.1	Système bactérien (<i>E. coli</i>).....	31
3.1.2	Système de culture de cellules humaines (HEK293T).....	32
3.2	Activité enzymatique de l'hélicase DHX36	34
3.3	Modèle d'homologie	34
3.3.1	Interactions entre la protéine DHX36 et l'ATP	37
3.4	Analogues d'ATP afin d'identifier la spécificité nucléotidique	41
3.5	Groupements fonctionnels nécessaires.....	42
3.6	Spécificité nucléotidique des purines ATP/GTP	45
Chapitre 4 – Discussion.....		47
Liste des références		51

LISTE DES TABLEAUX

Tableau 1 Fonctions biologiques des G-quadruplex d'ARN	10
Tableau 2 Propriétés biochimiques des hélicases à G4 et les phénotypes associés à leur dérèglement.....	12
Tableau 3 Niveau d'activité enzymatique selon l'analogue d'ATP	36

LISTE DES FIGURES

Figure 1 Structure de la double hélice d'ADN	1
Figure 2 Dogme central de la biologie moléculaire.....	2
Figure 3 Sucres à 5 carbones	3
Figure 4 Résumé des principales étapes de la production de protéines.....	4
Figure 5 Différentes structures secondaires retrouvées dans l'ADN par appariement de bases canonique de type Watson-Crick.....	5
Figure 6 Différentes structures secondaires et tertiaires retrouvées dans l'ARN par appariement de bases canonique de type Watson-Crick.....	6
Figure 7 Structure moléculaire de G-quadruplex d'ARN.....	7
Figure 8 Différences entre les G-quadruplex d'ADN et d'ARN.....	8
Figure 9 Fonctions des G-quadruplex localisés à travers différentes régions du pré-ARNm	9
Figure 10 Classification des hélicases	14
Figure 11 Les familles des hélicases SF1 et SF2.....	15
Figure 12 Organisation de la séquence et de la structure du noyau hélicase des protéines SF1 et SF2	16
Figure 13 La représentation de la structure d'une hélicase DEAD-box engagé avec l'ARN et l'ATP et le schéma des motifs principaux des hélicase de la superfamille 2 (SF2)	18
Figure 14 Représentation schématique des 1008 acides aminés de la protéine DHX36.....	21
Figure 15 Conservation en acides aminés de la protéine DHX36	22
Figure 16 Modèle du dépliement des G4 dépendant de DHX36.....	24
Figure 17 Purification de l'hélicase de G4 d'ARN humaine DHX36 à partir d'un système bactérien (<i>E. coli</i>).....	32
Figure 18 Purification et activité enzymatique de l'hélicase de G4 d'ARN humaine DHX36 à partir de culture de cellules humaines HEK293T	33
Figure 19 Modèle d'homologie de la protéine DHX36.....	37
Figure 20 Site de liaison de l'ADP	38
Figure 21 Résidus conservés dans la région du motif I.....	40
Figure 22 Analogues d'ATP utilisés afin d'analyser le site actif de l'hélicase DHX36	42

LISTE DES ABRÉVIATIONS

γ -phosphate	Groupement phosphate à l'extrémité du groupe triphosphate d'une molécule d'ATP.
$[\gamma^{32}\text{P}]\text{ATP}$	Molécule d'ATP portant l'isotope ^{32}P du groupement γ -phosphate.
μg	Microgramme [1×10^{-6} gramme].
μl	Microlitre [1×10^{-6} litre].
μM	Micromolaire [1×10^{-6} Molaire].
^{32}P	Isotope radioactif du phosphore.
ADN	Acide désoxyribonucléique.
ADN-B	Forme canonique de la double hélice d'ADN, de pas droit.
ADN G4	Structure G-quadruplex à l'intérieur de l'ADN.
ADN-H	Triple hélice d'ADN dans lequel un brin d'ADN effectue un appariement de type Hoogsteen avec une double hélice de type WC.
ADN-Z	Double hélice d'ADN, de pas gauche.
AMP-PNP	Molécule d'ATP non hydrolysable, l'atome de phosphore du groupement β -phosphate est substitué par un atome d'azote.
Ala	Alanine, acide aminé entrant dans la composition des protéines.
Asp	Aspartate, acide aminé entrant dans la composition des protéines.
Arg	Arginine, acide aminé entrant dans la composition des protéines.
ARN	Acide ribonucléique.
ARN G4	Structures G-quadruplex à l'intérieur de l'ARN.
ARNm	Acide ribonucléique messager, molécule temporaire codant pour une protéine.
ARNdb	Acide ribonucléique double brin.
ARNnc	Acide ribonucléique non-codant.
ARNr	Acide ribonucléique ribosomaux, entre dans la composition des ribosomes.
ARNsb	Acide ribonucléique simple brin (monocaténaire).

ARNt	Acide ribonucléique de transfert, interagit avec les ribosomes et l'ARNm afin de fournir les acides aminés adéquats pour la production de protéines.
ATP γ S	Molécule d'ATP non hydrolysable, l'atome de phosphore du groupement γ -phosphate est substitué par un atome de soufre.
ATPase	Adénosine triphosphatase, classe d'enzyme hydrolysant ou synthétisant une molécule d'ATP.
Ci	Unité de radioactivité, $3,7 \times 10^{10}$ désintégrations par seconde.
C-terminale	Extrémité carboxy-terminale, une des deux extrémités d'une protéine.
DEAD-box	Motif II caractéristique des hélicases de type DEAD, composé de quatre acides aminés spécifiques : Asp-Glu-Ala-Asp.
DEAH-box	Motif II caractéristique des hélicases de type DEAH, composé de quatre acides aminés spécifiques : Asp-Glu-Ala-His.
EDTA	Éthylène Diamine Tétra-Acétique, agent chélateur de cation métallique.
fmol	Femtomole [1×10^{-15} mole].
g	Gramme
G4	Structure G-quadruplex.
Glu	Glutamate, acide aminé entrant dans la composition des protéines.
Gly	Glycine, acide aminé entrant dans la composition des protéines.
His	Histidine, acide aminé entrant dans la composition des protéines.
HRP	<i>Horseradish peroxydase</i> , enzyme catalysant des réactions de réduction.
K ⁺	Cation de potassium.
K _m	Constante de Michaelis-Menten.
L	Litre
Li ⁺	Cation de lithium.
Lys	Lysine, acide aminé entrant dans la composition des protéines.
M	Concentration molaire d'un soluté exprimée en moles par litre.
mg	Milligramme [1×10^{-3} gramme].

Mg ²⁺	Cation de magnésium.
microARN	Courts acides ribonucléiques monocaténaire.
ml	Millilitre [1 x 10 ⁻³ litre].
mM	Millimolaire [1 x 10 ⁻³ Molaire].
mmol	Millimol [1 x 10 ⁻³ mole].
mole	Quantité de matière, 1 mole = 6,023 x 10 ²³ atomes.
Motif Q	Glutamine conservée en amont du motif I des hélicases de type DEAD.
Na ⁺	Cation de sodium.
ng	Nanogramme [1 x 10 ⁻⁹ gramme].
nm	Nanomol [1 x 10 ⁻⁹ mole].
nM	Nanomolaire [1 x 10 ⁻⁹ Molaire].
N-terminale	Extrémité terminale amine, une des deux extrémités d'une protéine.
NTP	Nucléoside triphosphate, au nombre de quatre selon la base azotée: Adénine (A), Guanine (G), Thymine (T) et Uracile (U).
NTPase	Nucléoside triphosphatase, classe d'enzyme hydrolysant ou synthétisant des molécules NTPs.
Phe	Phénylalanine, acide aminé entrant dans la composition des protéines.
Poly-A	Modification post-transcriptionnelle de l'extrémité 3' d'un transcrit d'ARNm, une série de plusieurs centaines de ribonucléotides de type adénosine (A).
Pré-ARNm	Acide ribonucléique monocaténaire immature issue de la transcription d'un gène.
RBPs	<i>RNA binding proteins</i> , protéines ayant la capacité de lier les molécules d'ARN.
RHA	<i>RNA helicase A</i> , hélicase codée par le gène <i>DHX9</i> .
Riboswitch	Segment riborégulateur d'un ARNm liant une petite molécule résultant d'un changement de l'efficacité de traduction d'une protéine.

RIP-chip	<i>RNA immunoprecipitation-chip</i> , technique d'immunoprécipitation spécifique de l'ARN sur puce permettant de détecter les interactions protéines-ARN.
RMN	Résonance magnétique nucléaire.
RSM	<i>RHAU specific motif</i> , motif spécifique à la protéine RHAU (DHX36).
SDS-PAGE	<i>Sodium dodecylsulfate polyacrylamide gel electrophoresis</i> , électrophorèse sur gel de polyacrylamide en présence de dodécylsulfate de sodium.
Ser	Sérine, acide aminé entrant dans la composition des protéines.
Thr	Thréonine, acide aminé entrant dans la composition des protéines.
Trp	Tryptophane, acide aminé entrant dans la composition des protéines.
UTR	<i>Untranslated regions</i> , régions non traduites flaquant la séquence codante d'une molécule d'ARNm.
V_{\max}	Vitesse de l'activité enzymatique maximale.
WC	Appariements de bases canoniques de type Watson-Crick, A et T; G et C.

REMERCIEMENTS

J'aimerais évidemment remercier mon directeur de recherche, Martin Bisailon afin de m'avoir accueilli dans son laboratoire pendant plusieurs années. Je le remercie également pour le savoir qu'il a réussi à me transmettre, non seulement sur le plan scientifique mais également sur le plan personnel. J'aimerais également remercier les membres du laboratoire qui ont su rendre ce parcours unique et agréable. Je remercie notamment l'assistante de recherche Maude Tremblay-Létourneau pour son soutien et son aide précieuse tout au long de ma maîtrise. Merci également aux membres du laboratoire Jean-Pierre Perreault qui ont partagé leur savoir concernant les structures G-quadruplex. Merci également à Louis-Philippe Morency du laboratoire de Rafael Najmanovich qui m'a aidé avec la modélisation. Finalement, je tiens à remercier Evan Booy du laboratoire de Sean McKenna au Manitoba afin d'avoir produit et fourni l'hélicase DHX36.

CHAPITRE 1 – INTRODUCTION

1.1 Les acides nucléiques

1.1.1 L'acide désoxyribonucléique (ADN)

La découverte de la structure canonique de l'acide désoxyribonucléique (ADN) par James Watson et Francis Crick a permis de paver la voie de l'étude de la biologie moléculaire contemporaine. Leurs analyses par rayons X ont révélé que l'ADN adopte une structure hélicoïdale de pas droit. En effet, la double hélice d'ADN est composée de deux brins antiparallèles qui établissent le squelette de désoxyribose-phosphate, ceux-ci sont appariés par l'entremise de liaisons hydrogènes entre les bases azotées ordonnées (Figure 1A). Les unités fondamentales de la séquence primaire d'ADN sont les désoxyribonucléotides enchaînés par des liaisons phosphodiester, ils se distinguent selon la base nucléique associée, elles sont au nombre de quatre : l'adénine (A), la guanine (G), la cytosine (C) et la thymine (T) (Figure 1B). L'agencement entre bases azotées proposé par Watson et Crick correspond à l'interaction entre une purine (A ou G) et une pyrimidine (T ou C). Cette interaction canonique A : T et C : G est dit un appariement de type Watson-Crick (WC) (Figure 1C) (Watson et Crick, 1953).

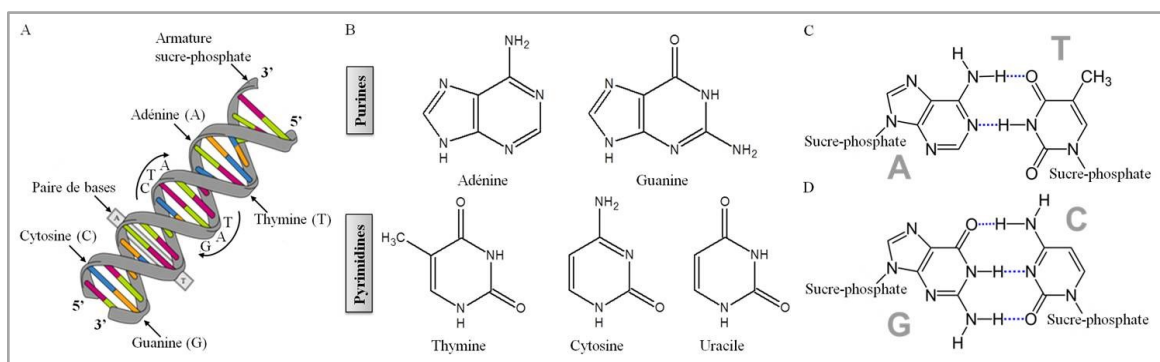


Figure 1. Structure de la double hélice d'ADN.

A. Schématisation d'une molécule d'ADN, les bases azotées sont représentées par le vert (adénine), le bleu (cytosine), l'orange (guanine) et le fuchsia (thymine). L'armature désoxyribose-phosphate est simplifiée en gris. **B.** Bases azotées. **C et D.** Appariement de type Watson-Crick entre l'adénine (A) et la thymine (T) ainsi qu'entre la guanine (G) et la cytosine (C), les liaisons hydrogènes sont représentées par les pointillés.

A: Adapté de https://upload.wikimedia.org/wikipedia/commons/a/a2/DNA_structure_and_bases_color_FR.svg.

C: Adapté de https://upload.wikimedia.org/wikipedia/commons/c/ce/Base_Pair_AT_Hydrogen_Bridge_V.1.svg.

D: Adapté de https://upload.wikimedia.org/wikipedia/commons/5/5c/Base_Pair_GC_Hydrogen_Bridge_V.1.svg.

L'ordre des nucléotides de la séquence primaire permet notamment de coder l'information nécessaire à l'expression génique. Ce processus s'effectue par l'entremise de la synthèse de l'acide ribonucléique (ARN) (Figure 2). Ce biopolymère est l'intermédiaire entre le code génétique (ADN) et ses effecteurs, les protéines. En effet, le dogme fondamental de la biologie moléculaire, représenté par la figure 2, dicte que le code génétique supporté par l'ADN est transcrit temporairement sous forme d'ARN pour que celui-ci soit ensuite traduit en protéines. Conséquemment, la séquence nucléotidique de l'ADN et de l'ARN influencera la séquence primaire en acides aminés de la protéine résultante. L'ordre des nucléotides peut également réguler le processus d'expression génique. Effectivement, il est possible de retrouver des séquences régulatrices en périphérie des gènes qui permettront d'activer ou d'inhiber leur expression. Somme toute, la séquence primaire de l'ADN est primordiale quant à l'expression génique et le phénotype résultant.

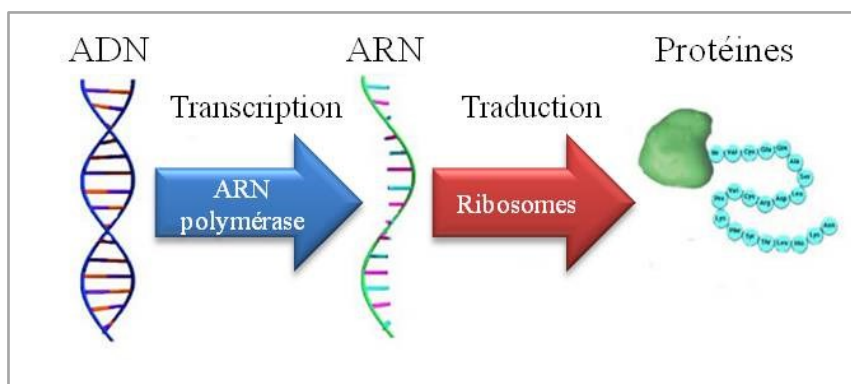


Figure 2. Dogme central de la biologie moléculaire.

Un segment d'ADN est transcrit en ARN par l'ARN polymérase, le transcrit est traduit en protéine par les ribosomes.

Adapté de London Health Sciences Centre, *The central dogma of biology*, http://www.lhsc.on.ca/Patients_Families_Visitors/Genetics/Inherited_Metabolic/Mitochondria/BasicBiology.htm.

1.1.2 L'acide ribonucléique (ARN)

Contrairement à l'ADN, l'ARN est un biopolymère monocaténaire formé de ribonucléotides; le pentose associé est le ribose et l'uracile (U) remplace la thymine (Figure 1B). La présence

du groupement hydroxyle en 2' des ribonucléotides et la nature monocaténaire de l'ARN minimisent la stabilité relative de la molécule. La distinction chimique entre le ribose et le désoxyribose est schématisé dans la figure 3.

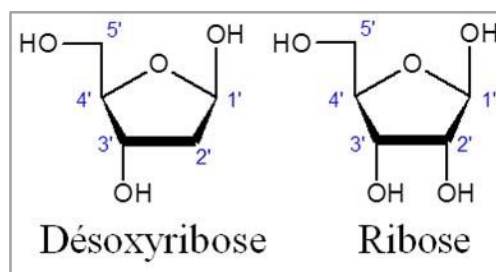


Figure 3. Sucres à 5 carbones.

Sucres qui entrent dans la composition des désoxyribonucléotides et des ribonucléotides.

Le produit de transcription de l'ADN est le pré-ARN messager (pré-ARNm). Le pré-ARNm subira différentes modifications post-transcriptionnelles telles que l'ajout d'une coiffe à l'extrémité 5', l'épissage et l'ajout d'une queue poly-A à l'extrémité 3' (Figure 4). Ces étapes de maturation permettent d'obtenir une molécule d'ARN messager mature (ARNm), celle-ci sera ensuite exportée au cytoplasme afin d'y être traduite en protéines par les ribosomes. Les régions non-traduites (UTR – *Untranslated region*) flanquant la région codante influenceront l'efficacité de la traduction de l'ARNm.

La participation au processus de production de protéines n'est pas la seule fonction d'une molécule d'ARN. Un transcrit d'ARN non-codant (ARNnc) peut également avoir diverses fonctions physiologiques, notamment les ARN de transfert (ARNt) et les ARN ribosomiaux (ARNr). La fonction des ARNnc sera dictée par leur structure secondaire influencée par leur séquence en ribonucléotides. La structure secondaire adoptée par les molécules d'ARN est primordiale quant au processus de transcription, d'épissage, de traduction, de localisation et du renouvellement de l'ARN (Mortimer *et al.* 2014).

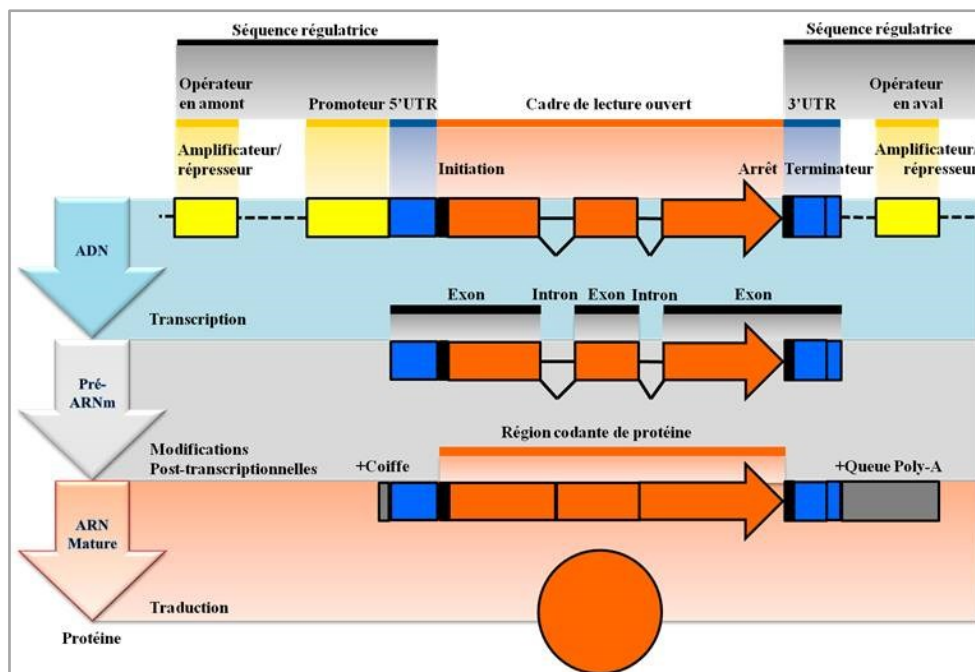


Figure 4. Résumé des principales étapes de la production de protéines.

1.1.3 Structures secondaires et tertiaires des acides nucléiques

1.1.3.1 Structures secondaires et tertiaires de l'ADN

Les structures secondaires formées par l'ARN et l'ADN sont variées, la principale forme adoptée par l'ADN est la double hélice de type-B. Celle-ci est caractérisée par une double hélice de pas droit composée de 10,5 paires de bases WC par tour. À travers les années, différents groupes de recherche ont permis de révéler que la conformation de type-B de la double hélice d'ADN n'était pas la seule forme adoptée par une molécule d'ADN. Ces différentes conformations ont été expérimentalement révélées notamment par diffraction aux rayons X et par résonance magnétique nucléaire (RMN). L'ADN peut se retrouver sous forme de cruciformes, de triple hélice d'ADN-H ainsi que de G-quadruplex (G4) pour ne donner que quelques exemples. En résumé, selon la séquence et les différentes conditions environnantes, l'ADN manifestera une variété de structures particulières (Figure 5). Ces structures d'ADN affectent les processus de réplication de l'ADN, de transcription, des systèmes de recombinaison et de réparation (Svozil *et al.*, 2008).

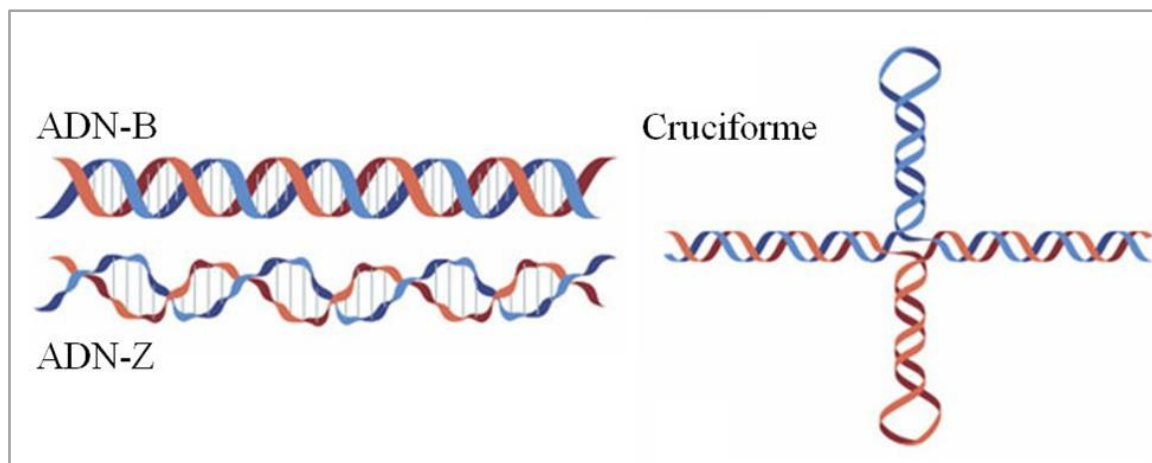


Figure 5. Différentes structures secondaires retrouvées dans l'ADN par appariement de bases canonique de type Watson-Crick.

ADN-Z : Contrairement à l'ADN-B, l'ADN-Z est une conformation double hélice de pas gauche causé favorisé par un surenroulement négatif (Gessner *et al.*, 1989).

Cruciforme : Un surenroulement négatif peut également provoquer la formation de cette structure à partir de l'ADN-B, cette structure ressemble notamment à jonction de type Holliday (Palecek, 1991).

Adapté de Matthew L. Bochman, Katrin Paeschke & Virginia A. Zakian. *DNA secondary structures: stability and function of G-quadruplex structures*. Nature Reviews Genetics, 2012, **13**:770-780. doi:10.1038/nrg3296.

1.1.3.2 Structures secondaires et tertiaires de l'ARN

Contrairement à l'ADN, l'ARN est retrouvé naturellement sous la forme monocaténaire permettant à cette molécule de former plusieurs conformations différentes. Notamment, les structures secondaires les plus simples retrouvées, par appariement de bases de type WC, sont la boucle à épingle à cheveux, la tige-boucle et le pseudo-nœud (Figure 6).

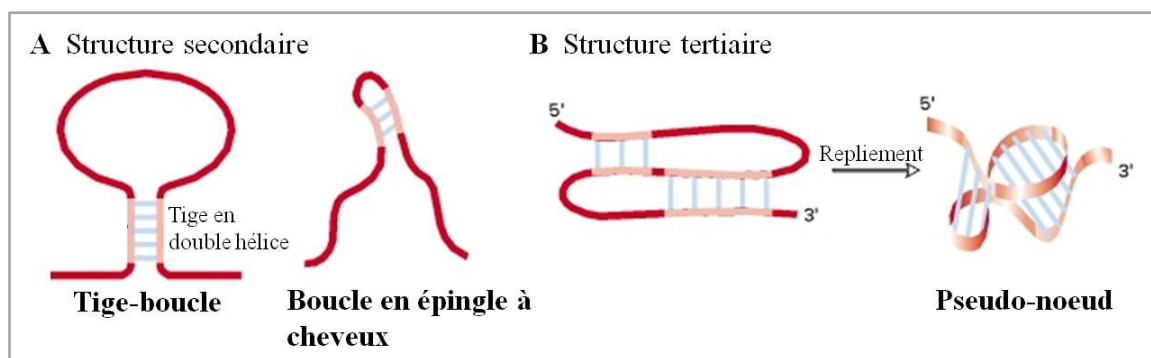


Figure 6. Différentes structures secondaires et tertiaires retrouvées dans l'ARN par appariement de bases canonique de type Watson-Crick.

Adapté de Molecular Cell Biology. 4th edition. Lodish H, Berk A, Zipursky SL, *et al.* New York: W. H. Freeman; 2000.

La présence de structures secondaires et tertiaires dans l'ARN influence la fonction d'une panoplie de types d'ARN, que ce soit des ARNm mais également parmi les ARN non-codants, tels que des ribozymes, des riboswitch, des long ARN non-codants et des microARN. Les structures secondaires de l'ARN ont un rôle parmi toutes les étapes du processus d'expression génique. Ces structures peuvent notamment affecter le processus de traduction en interférant sur la liaison de la sous-unité ribosomale 40S à l'ARNm (Kozak, 1980). Les capacités des structures secondaires d'ARN sont variées, celles-ci peuvent permettre notamment à un ARN d'interagir avec soi-même, avec d'autres ARN, avec certains ligands et avec des protéines liant l'ARN (RBPs - *RNA-binding proteins*). Ces structures secondaires permettent au RBPs d'interagir avec un ARN selon un site spécifique mais sont également en mesure de limiter l'accès à ces sites en modifiant leur structure (Wan *et al.*, 2011). Ainsi, il existe une variété de structures secondaires d'ARN ayant des rôles bien définis parmi le processus d'expression génique. La nature transitoire des structures secondaires d'ARN permet notamment de moduler le rôle de celles-ci et, par extension, moduler par exemple l'efficacité de la traduction (Kozak, 1980). Néanmoins, des travaux de recherche ont permis de découvrir une structure secondaire non canonique ayant un rôle potentiel à travers le processus d'expression génique. Il s'agit des structures G-quadruplex, celles-ci peuvent se retrouver à la fois dans l'ADN et dans l'ARN.

1.2 Les G-quadruplex (G4)

1.2.1 Structure

Les structures G-quadruplex ont été découvertes expérimentalement par le groupe de recherche de Gellert en 1960 (Gellert, Lipsett et Davies, 1962). Ces travaux ont permis de révéler la structure à quatre brins non canonique d'une région d'ADN ou d'ARN riche en guanine en employant la méthode de diffraction des rayons X. Les G-quadruplex sont formés d'un empilement de G-quartets dont l'unité élémentaire est la guanine. Les G-quartets représentent un arrangement cyclique planaire de quatre guanines liées par des interactions de type Hoogsteen (Figure 7A). Un cation monovalent stabilise la structure G-quadruplex en interagissant avec les oxygènes chargés négativement au centre des G-quartets (Figure 7).

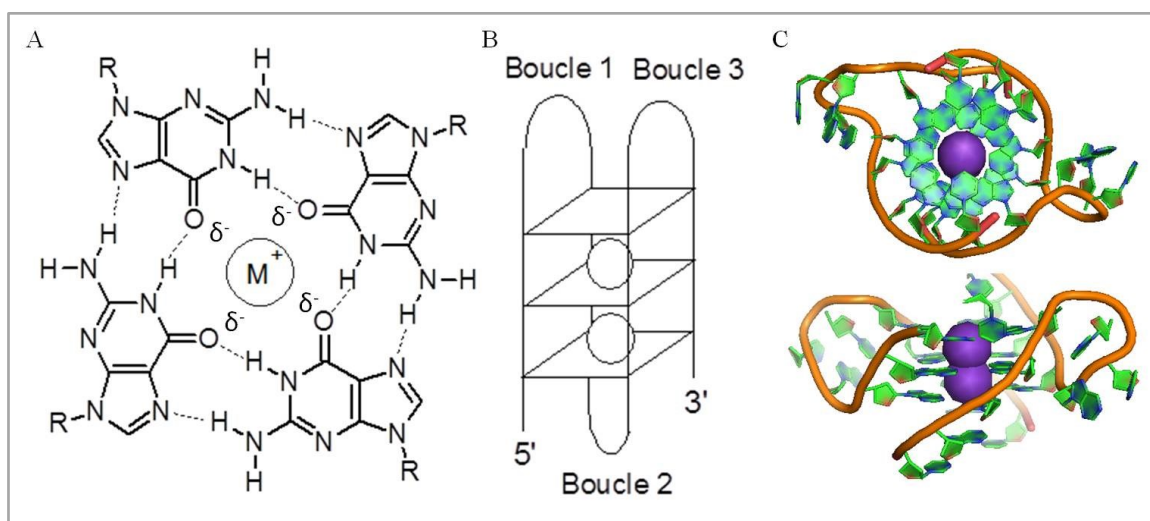


Figure 7. Structure moléculaire de G-quadruplex d'ARN.

A. Représentation schématique d'un G-quartet. Le cation monovalent (M^+) stabilise les charges partielles négatives des oxygènes centraux. **B.** Représentation schématique d'un G-quadruplex intramoléculaire. Les G-quartets sont schématisés par des losanges, les cations monovalents par des sphères et le squelette ribose-phosphate par un trait dont celui-ci est orienté en 5'-3'. **C.** Structure cristalline d'un G-quadruplex d'ARN intermoléculaire (PDB ID – 3IBK). (Collie *et al.*, 2010) Le ribose et les bases azotées sont représentés en vert et le squelette de phosphate en orange. Les ions de potassium (K^+) sont en mauve.

Les cations monovalents stabilisant ces structures sont énumérés en ordre décroissant d'efficacité de stabilisation: $K^+ > Na^+ > Li^+$ (Burge *et al.*, 2006). Ainsi, le potassium est le cation majoritairement retrouvé au sein des structures G4. Les G4 peuvent manifester différentes topologies en fonction de plusieurs facteurs : le cation monovalent (K^+ ou Na^+),

le nombre de molécules d'acide nucléique impliquées dans leur formation (intramoléculaire, bimoléculaire ou tétramoléculaire), l'orientation relative des brins (parallèle ou antiparallèle), le nombre de G-quartets empilés et finalement la séquence nucléotidique (Keniry, 2000; Patel, Phan et Kuryavyi, 2007). Comparativement aux G4 d'ADN, les G4 d'ARN sont thermodynamiquement plus stables, plus compactes et moins hydratés (Joachimi, Benz et Hartig, 2009; Arora et Maiti, 2009). Tel que mentionné au préalable, le ribose possède un groupement hydroxyle en 2', celui-ci permet d'accentuer les interactions intramoléculaires et permettant de stabiliser davantage le G4 d'ARN (Figure 8). Par ailleurs, les G4 d'ARN sont majoritairement retrouvés dans le cytoplasme (Ji *et al.*, 2011).

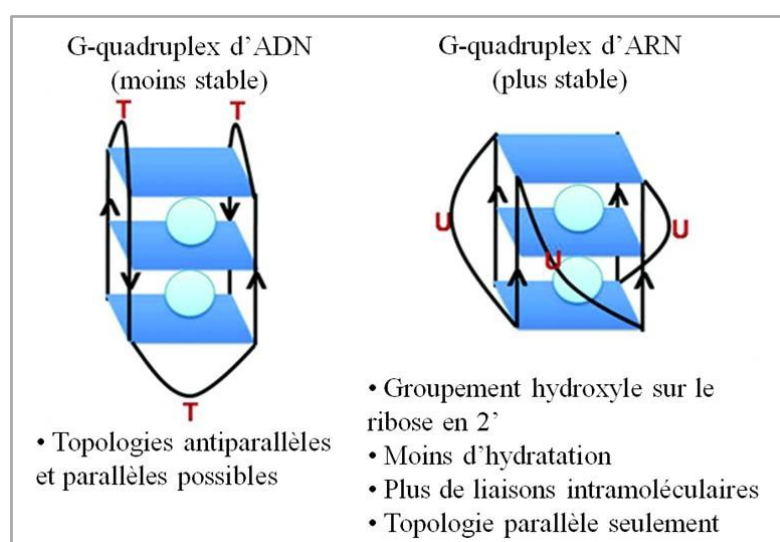


Figure 8. Différences entre les G-quadruplex d'ADN et d'ARN.

Adapté de Prachi Agarwala, Satyaprakash Pandey & Souvik Maiti. *The tale of RNA G-quadruplex*. *Org. Biomol. Chem.*, 2015, **13**, 5570. doi: 10.1039/c4ob02681k.

1.2.2 Localisation

Les études bio-informatiques ont révélé qu'il y avait environ 376 000 séquences ayant le potentiel de former des G-quadruplex d'ADN à l'intérieur du génome humain (Simonsson, 2001). Les structures G4 ne sont pas aléatoirement distribuées, elles ont été principalement retrouvées et analysées au niveau des régions télomériques. Cependant, il est possible de retrouver un enrichissement de G4 dans les régions promotrices de proto-oncogènes, dans les extrémités 5' des gènes (5'UTR, premier exon et premier intron) et finalement dans les régions à haute fréquence de mutation (Simonsson, 2001). La figure 9 résume la localisation des G4 d'ARN au niveau des pré-ARNm incluant les processus desquels les G4 pourraient

être impliqués. Des études bio-informatiques sur l'étendue du génome ont révélé la présence de plus de 4000 motifs ayant le potentiel de former des G4 dans le 5'UTR des ARNm (Huppert *et al.*, 2008).

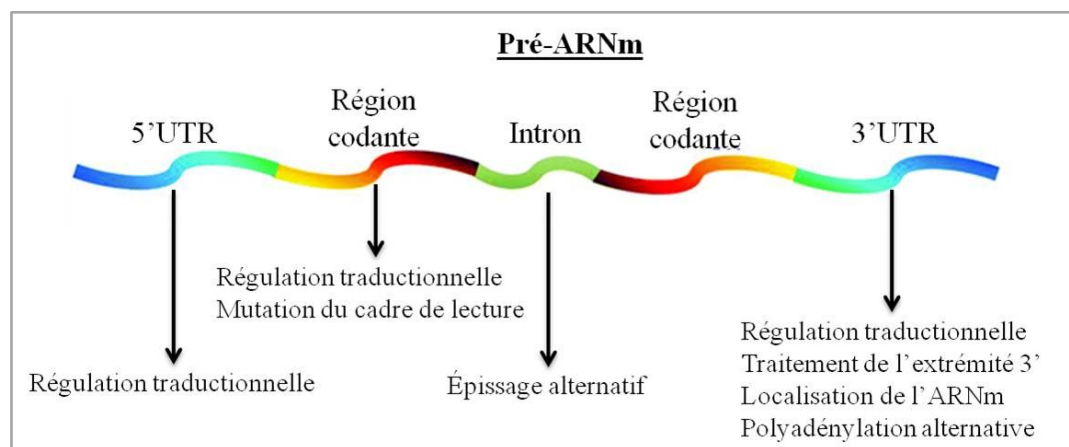


Figure 9. Fonctions des G-quadruplex localisés à travers différentes régions du pré-ARNm.

Adapté de Prachi Agarwala, Satyaprakash Pandey & Souvik Maiti. *The tale of RNA G-quadruplex*. *Org. Biomol. Chem.*, 2015, **13**, 5570. doi: 10.1039/c4ob02681k.

1.2.3 Fonctions

Des évidences suggèrent que les G4 sont impliqués dans le maintien de l'intégrité génomique et dans la régulation des processus de réplication, de transcription et de recombinaison (Simonsson, 2001). Ainsi, le fait que les structures G4 d'ADN soient associées aux régions télomériques, aux oncogènes et, par extension, au cancer, la recherche pour trouver un procédé thérapeutique en fait un sujet chaud dans la communauté scientifique (Wu et Brosh, 2010). Les fonctions des ARN G4 sont variées; la figure 9 résume le rôle de ceux-ci à l'intérieur d'une molécule de pré-ARNm (Agarwala, Pandey et Maiti, 2015). Globalement, plusieurs études révèlent que la présence d'un G4 à l'intérieur du 5'UTR inhibe la traduction de l'ARNm en limitant l'accès des ribosomes, notamment le transcrite du gène humain *NRAS* (Kumari *et al.*, 2007). Cependant, un groupe de recherche a démontré que la présence d'un G4 à l'intérieur du 5'UTR activait la traduction du transcrite du gène *TGFβ2* (Agarwala *et al.*, 2013). Ainsi, la présence d'un G4 dans le 5'UTR permettrait à la fois l'activation et l'inhibition de la traduction des ARNm. De son côté, peu d'études ont permis d'élucider la conséquence de la présence d'un G4 dans le 3'UTR. Néanmoins, la présence d'un G4 dans le 3'UTR du proto-oncogène *PIMI* peut avoir comme conséquence la répression de la

traduction du transcrit; cependant le mécanisme reste inconnu (Arora et Sues, 2011). Les fonctions biologiques des G-quadruplex d'ARN sont résumées dans le tableau 1 (Agarwala, Pandey et Maiti, 2015). Somme toute, le dérèglement de certains processus associés aux structures G4 révèle que celles-ci ont un rôle prépondérant au niveau physiologique.

Tableau 1. Fonctions biologiques des G-quadruplex d'ARN.

Rôle	ARNm	Fonctions possibles	Références
1. Régulation traductionnelle			
A Répression	NRAS	Un G4 dans le 5'UTR de ces ARNm cellulaires réprime la traduction.	Kumari <i>et al.</i> , 2007.
	Bcl-2		Shahid <i>et al.</i> , 2010.
	TRF-2	Un G4 présent dans le cadre de lecture ouvert de cet ARNm provoque un arrêt de l'élongation traductionnelle.	Gomez <i>et al.</i> , 2010.
	Estrogen receptor		Endoh <i>et al.</i> , 2013.
	PIM1	Un G4 dans le 3'UTR de cet ARNm est associé à une répression de la traduction.	Arora et Sues, 2011.
B Activation	TGFβ2	Un G4 dans cet ARNm augmente la traduction.	Agarwala <i>et al.</i> , 2013.
2. Traitement de l'extrémité 3'			
	IGF-II	Ce transcrit subit un clivage spécifique <i>in vivo</i> à un site en amont d'un motif G4.	Christiansen <i>et al.</i> , 1994. - Nielsen et Christiansen, 1995.
	p53	Un G4 en aval du site de clivage du pré-ARNm p53 est crucial pour maintenir l'efficacité du traitement de l'extrémité 3' de p53 suivant une irradiation aux rayons UV.	Decorsiere <i>et al.</i> , 2011.
Polyadénylation alternative	FXR1	Un G4 situé dans le 3'UTR dans cet ARNm augmente l'efficacité polyadénylation alternative générant plusieurs petits transcrits.	Beaudoin et Perreault, 2013.
3. Terminaison de la transcription			
	Control region-conserved sequence block II (CSB II) des mitochondries	Un G4 favorise la terminaison de la transcription et le mécanisme est similaire à la terminaison Rho-indépendante chez les procaryotes.	Wanrooij <i>et al.</i> , 2010.
4. Localisation de l'ARNm			
	CaMKIIa	Un G4 dans le 3'UTR de cet ARNm est un élément de localisation aux neurites qui permet le transport de cet ARNm du soma vers les dendrites.	Subramanian <i>et al.</i> , 2011.
5. Épissage alternatif			
	hTERT	Un G4 affecte l'épissage et l'expression génique de ces ARNm en fournissant des sites de liaisons de protéines nécessaires à l'épissage.	Gomez <i>et al.</i> , 2004.
	FMR1	Un G4 dans l'intron 3 module l'épissage de l'intron 2 ce qui mène à l'expression différentiel de transcrits qui codent pour différentes isoformes de p53.	Didiot <i>et al.</i> , 2008.
	p53		Marcel <i>et al.</i> , 2011.

Adapté de Prachi Agarwala, Satyaprakash Pandey & Souvik Maiti. *The tale of RNA G-quadruplex*. Org. Biomol. Chem., 2015, **13**, 5570. doi: 10.1039/c4ob02681k.

1.2.4 Problématique

Les structures G4 ont été répertoriées à travers une variété de pathologies humaines; une recherche approfondie du dérèglement entourant ces structures au niveau des processus moléculaires et cellulaires s'avère un défi actuel pour la communauté scientifique. L'étude des processus métaboliques associés aux structures G-quadruplex permettrait d'élaborer des

moyens thérapeutiques dans le but de limiter les maladies reliées. Par ailleurs, la diminution de la longueur des télomères représente un facteur de risque du développement du cancer (Greenberg *et al.*, 1999; Lansdorp, 2009). Les télomères sont un complexe nucléoprotéique situés aux extrémités chromosomiques retrouvés parmi les vertébrés. Les répétitions d'ADN télomériques ont pour séquence (TTAGGG)_n sur plusieurs milliers de bases. Ainsi, les études biochimiques ont démontré que les structures G-quadruplex se forment au niveau des télomères *in vivo* (Arora *et al.*, 2008). Si les G4 aux régions télomériques procurent une stabilité structurelle aux télomères, une déstabilisation de ceux-ci pourrait entraîner des défauts fonctionnels des télomères (Agarwala, Pandey et Maiti, 2015). Les maladies humaines associées à la perturbation des télomères comprennent des formes d'anémie, l'hypertension, des maladies coronariennes, une infection chronique au virus de l'immunodéficience humaine (VIH) et une maladie chronique du foie (Lansdorp, 2009; Armanios, 2009).

Par ailleurs, le génome humain contient également des séries de répétitions nucléotidiques en tandem, ces unités sont classifiées en tant que microsatellites et minisatellites. Ces séquences répétées sont sujettes à l'expansion ou bien à la contraction, elles contribuent à l'instabilité génomique et elles se sont avérées être impliquées dans plusieurs maladies humaines, notamment le syndrome du X fragile (Crawford, Acuna et Sherman, 2001). Le syndrome du X fragile correspond à un retard mental héréditaire causé par une répétition trinuécléotidique CGG dans le 5'UTR du gène *FMRI*. Lorsque la quantité de répétitions CGG est supérieure à 200, celles-ci subissent une hyperméthylation causant une répression de l'expression du gène *FMRI* au niveau transcriptionnel. Par contre, il a été démontré que les répétitions de CGG au niveau du gène *FMRI* pouvaient former des structures G-quadruplex stables ayant pour conséquence de nuire à la transcription du gène (Fry et Loeb, 1994; Usdin et Woodford, 1995). Évidemment, l'inhibition de l'expression du gène *FMRI* et par conséquence l'absence de la protéine encodée par le gène, FMRP, entraîne le phénotype aberrant. Il a été révélé que ces séquences répétitives CGG formant des structures d'ADN G4 peuvent être déroulées par des hélicases spécifiques, telles que WRN (Fry et Loeb, 1999) et FANCI (Wu, Shin-ya et Brosh, 2008). La mutation altérant l'activité ou l'expression génique de ces protéines entraîne les désordres génétiques du syndrome de Werner et

l'anémie de Fanconi respectivement. Ainsi, la régulation adéquate des structures G-quadruplex est nécessaire au développement d'un individu.

1.2.5 Hélicases de G4

Afin de réguler la stabilité structurale des G-quadruplex, le génome humain code pour des hélicases dont l'activité catalytique consiste à dérouler les conformations G4 autant au niveau de l'ADN que l'ARN. Le tableau 2 énumère les hélicases de G-quadruplex découvertes ainsi que le phénotype associé au dérèglement de celles-ci. Afin de comprendre davantage le mécanisme de régulation des structures G4 dans la cellule, on peut d'abord s'intéresser aux enzymes responsables de leur déroulement, c'est-à-dire les hélicases.

Tableau 2. Propriétés biochimiques des hélicases à G4 et les phénotypes associés à leur dérèglement.

Protéine	Organisme	ADN G4	ARN G4	Phénotype	Références
WRN	Humain	3' → 5'	ND	Vieillesse prématurée, instabilité génomique.	Fry et Loeb, 1999, Brosh et Bohr, 2007, Mohaghegh <i>et al.</i> , 2001.
BLM	Humain	3' → 5'	ND	Cancer, hausse des échanges en chromatides sœurs, instabilité génomique.	Sun <i>et al.</i> , 1998, Chu et Hickson, 2009.
FANCD1	Humain	5' → 3'	ND	Anémie de Fanconi, cancer du sein, réparation par recombinaison homologe anormale et progression tardive en phase S.	Wu, Shin-ya et Brosh, 2008, Wu, Suhasini et Brosh, 2009, London <i>et al.</i> , 2008.
G4R1/RHAU	Humain	Oui	Oui	Déadénylation et dégradation anormales de l'ARNm.	Vaugh <i>et al.</i> , 2005, Creacy <i>et al.</i> , 2008, Tran <i>et al.</i> , 2004.
Pif1	Humain	ND	ND	Défauts télomériques possible.	Zhang <i>et al.</i> , 2006, Bochman, Sabouri et Zakian, 2010, Ribeyre <i>et al.</i> , 2009,
	Levure	5' → 3'	ND	Instabilité de l'ADN mitochondrial et des régions riches en guanine, télomères allongés et métabolisme du fragment d'Okazaki aberrant.	Budd <i>et al.</i> , 2006, Cheng, Dunaway et Ivessa, 2007, Pike <i>et al.</i> , 2009.
Dna2	Humain	5' → 3'	ND	Ralentissement du cycle cellulaire et division cellulaire aberrante, instabilité de l'ADN génomique et mitochondrial.	Duxin <i>et al.</i> , 2009, Zheng <i>et al.</i> , 2008.
	Levure	5' → 3'	ND	Métabolisme du fragment d'Okazaki aberrant.	Masuda-Sasa <i>et al.</i> , 2008, Kang, Lee et Seo, 2008, Stewart, Campbell et Bambara, 2009.
Sgs2	Levure	3' → 5'	ND	Réparation des cassures doubles brins inadéquate, reprise de l'arrêt des fourches de réplication inadéquate, traitement des événements de recombinaison intermédiaires défectueux et instabilité des télomères.	Huber, Lee et Maizels, 2001, Ashton et Hickson, 2010.

Adapté de Wu, Yuliang & Robert M. Brosh. *G-Quadruplex nucleic Acids and Human Disease*. The FEBS journal. 2010;277(17):3470-3488. doi:10.1111/j.1742-4658.2010.07760.x.

1.2.6 Les hélicases

Les hélicases regroupent les enzymes responsables de la catalyse de la séparation d'acide nucléique double-brin en utilisant l'énergie d'hydrolyse des nucléosides phosphate (NTP). Le produit de réaction résultant est un acide nucléique monocaténaire; celui-ci sera ensuite le substrat de plusieurs réactions cellulaires. Par conséquent, les hélicases sont impliquées dans pratiquement toutes les étapes du métabolisme de l'ADN et de l'ARN notamment la réplication, la réparation de l'ADN, la recombinaison, la transcription, la traduction, les réarrangements de la chromatine, la synthèse des ribosomes, la maturation et l'épissage des pré-ARNm, l'export nucléaire et finalement le retrait des protéines de l'ADN et de l'ARN (Byrd et Raney, 2012).

1.2.6.1 Familles

Gorbalenya et Koonin ont classé les hélicases selon la comparaison de leur séquence en acides aminés *in silico*. Les résultats de leurs travaux se résument à la classification des hélicases parmi 3 superfamilles et 2 plus petites familles d'hélicases putatives. (Gorbalenya et Koonin, 1993) Ils ont déterminé que les protéines réunies à travers chacune des superfamilles 1 et 2 possèdent une série de motifs conservés au nombre de 7. Cependant, à travers les années, certaines études ont révélé qu'il y avait plusieurs déficiences à propos de leur conclusion. Notamment, certaines séquences analysées *in silico* suggérant qu'ils s'agissaient d'hélicases se sont avérées de faux positifs. Singleton *et al.* ont proposé une nouvelle organisation des hélicases qui font l'inclusion de certains motifs qui ont été omis préalablement tels que TxGx, TRG et les motifs Q et 4a (Singleton, Dillingham et Wigley, 2007) (Figure 10). Les hélicases sont donc classées parmi 6 superfamilles (SF1 à SF6) basées sur la conservation de séquence et les éléments structuraux, dont tous les membres des familles possèdent les motifs Walker qui sont également répertoriés dans plusieurs enzymes NTPase (Walker *et al.*, 1982).

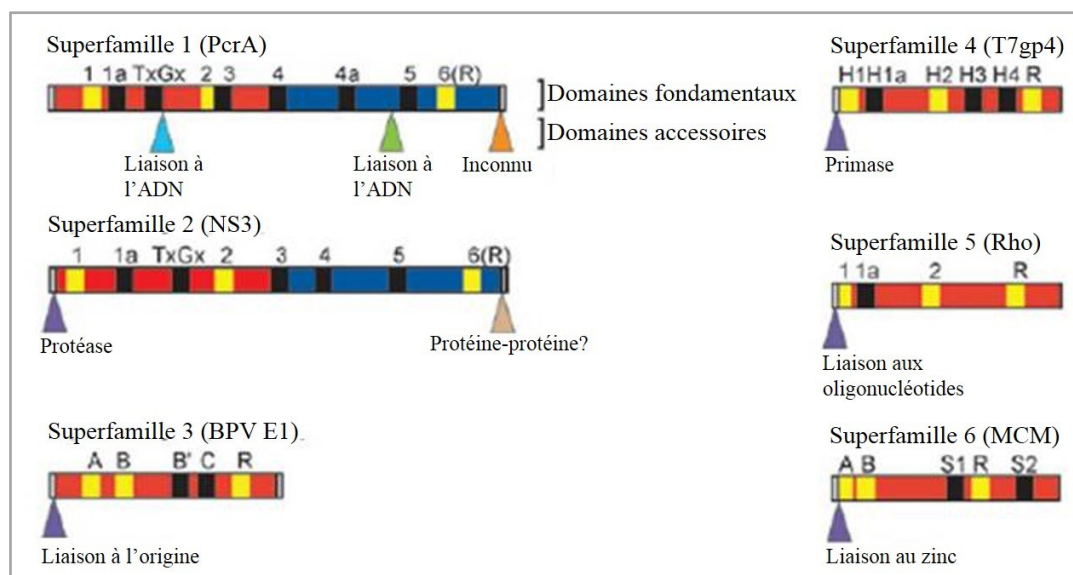


Figure 10. Classification des hélicases.

Le nom d'un membre de chacune des superfamilles est écrit entre parenthèses. Les domaines fondamentaux sont représentés pour chacune des classes d'hélicases. Les motifs en jaune constituent les éléments structuraux retrouvés dans toutes les hélicases. Les domaines accessoires sont spécifiques à l'hélicase représentée pour chacune des familles mais peuvent également être partagés à travers les membres d'une même famille.

Adapté de Martin R. Singleton, Mark S. Dillingham & Dale B. Wigley. *Structure and mechanism of helicases and nucleid acid translocases*. Annual Review of Biochemistry, 2007, **76**: 23-50. doi: 10.1146/annurev.biochem.76.052305.115300.

1.2.6.2 Superfamilles 1 et 2

La caractéristique particulière des hélicases comprises parmi les superfamilles 1 (SF1) et 2 (SF2) est le noyau hélicase conservé réunissant deux domaines similaires qui ressemblent au repliement de la protéine bactérienne de recombinaison RecA (Singleton, Dillingham et Wigley, 2007). Par ailleurs, les protéines de la famille SF2 sont ubiquitaires à travers les espèces. Il y a également plusieurs paralogues encodés par le génome d'une multitude d'organismes (Pyle, 2008; Anantharaman, Koonin et Aravind, 2002; Fairman-Williams, Guenther et Jankovsky, 2007). Fairman-Williams Guenther et Jankovsky ont subdivisé les membres des superfamilles 1 et 2 en alignant les séquences des noyaux hélicases des protéines chez l'humain, la levure à bourgeon, *E. coli* et certaines protéines virales. Leurs résultats leur permettent de grouper les protéines selon leur similarité de séquence (Figure 11).

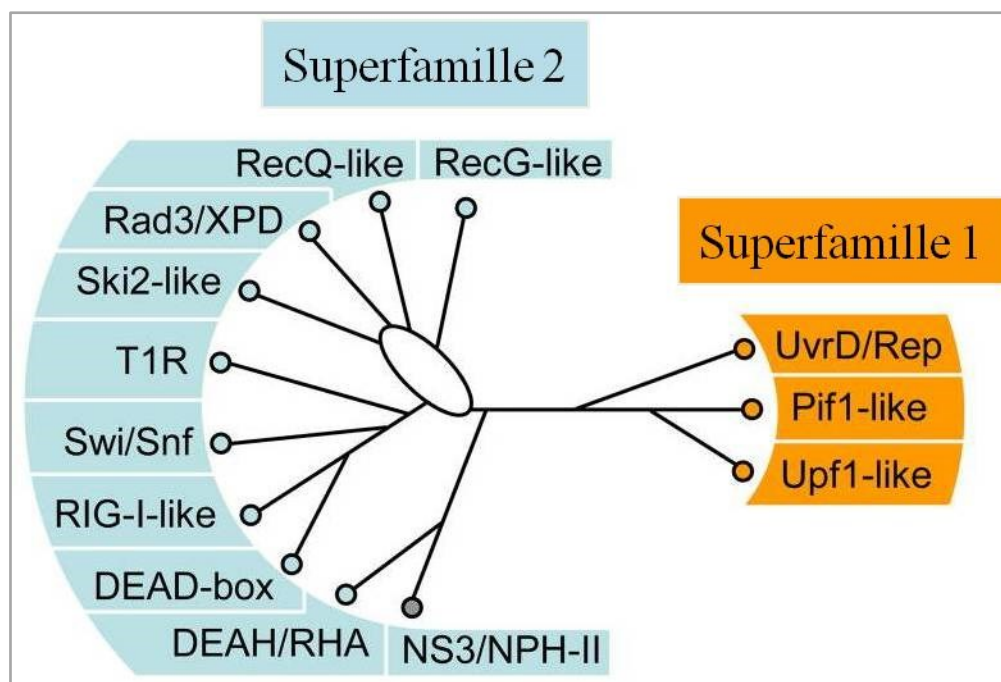


Figure 11. Les familles des hélicases SF1 et SF2.

Cladogramme schématisé montrant les trois familles de SF1 (droite) et les neuf familles et un groupe de SF2 (gauche).

Adapté de Margaret E Fairman-Williams, Ulf-Peter Guenther & Eckhard Jankowsky. *SF1 and SF2 helicases: family matters*. *Curr. Opin Struct Biol*, 2007, **17**:316-324. doi: 10.1016/j.sbi.2007.05.007.

Les caractéristiques du domaine central hélicase partagé par les protéines des superfamilles 1 et 2 résident parmi 12 motifs conservés. Ces 12 motifs sont représentés dans la figure 12. Le plus haut degré de conservation de séquence entre les superfamilles réside majoritairement au niveau des acides aminés qui coordonnent la liaison et l'hydrolyse du triphosphate (motifs I, II et VI). Ces résidus se situent notamment entre les deux domaines hélicase conservés RecA (Leipe *et al.*, 2002). La structure et les motifs de séquence conservés seront davantage examinés à travers les sous-groupes des hélicases DEAD/H-box de la superfamille 2.

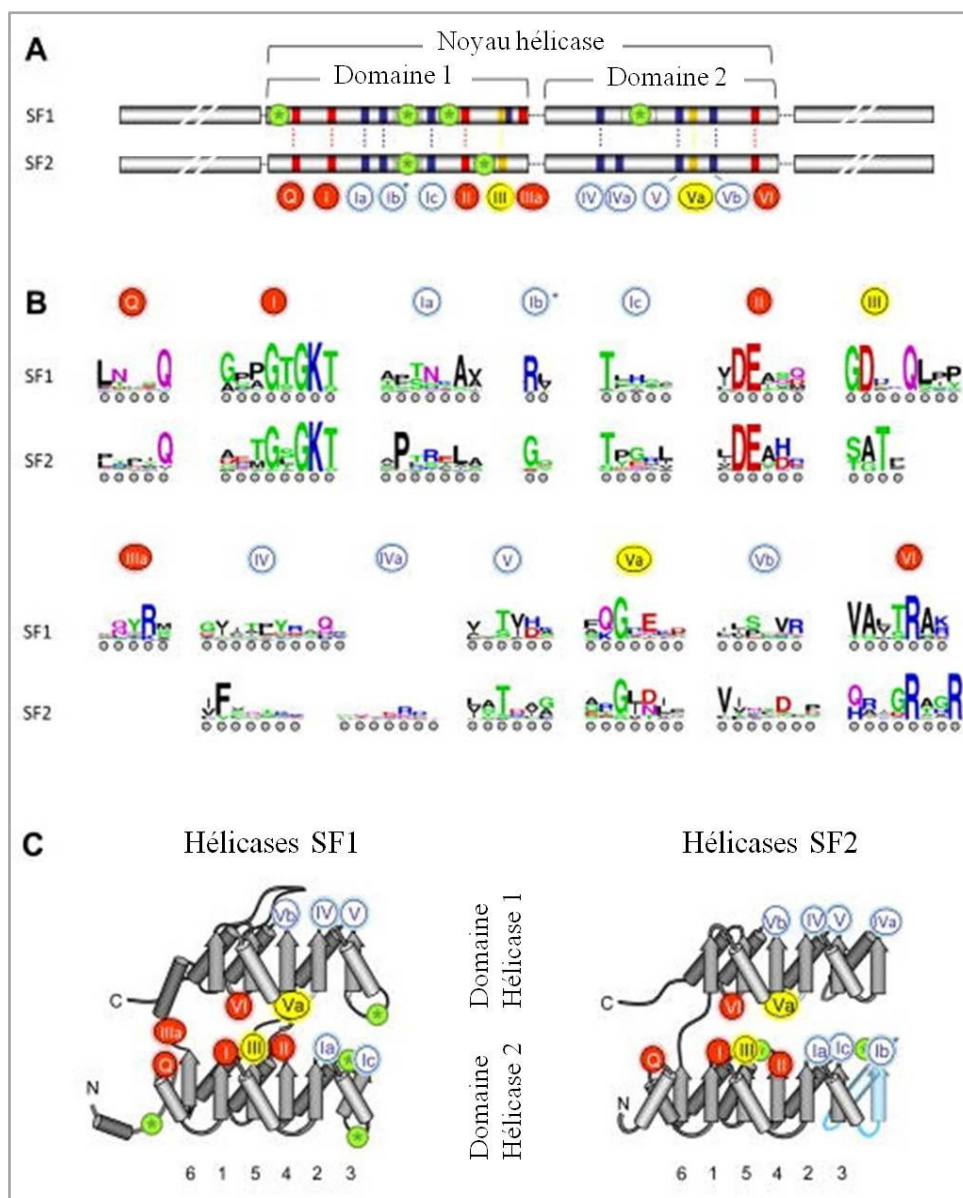


Figure 12. Organisation de la séquence et de la structure du noyau hélicase des protéines SF1 et SF2.

A Organisation de la séquence du noyau hélicase de SF1 et SF2. Les motifs de séquence caractéristiques sont colorés selon leur fonction biochimique prédominante: rouge, liaison et hydrolyse de l'ATP; jaune, coordination entre l'acide nucléique et les sites de liaison des NTP; bleu, liaison à l'acide nucléique. Les astérisques entourés d'un cercle vert marquent les sites d'insertions de domaines additionnels. Les longueurs des blocs et la distance entre les domaines conservés ne sont pas à l'échelle. Les motifs caractéristiques sont issus de l'alignement de toutes les protéines SF1 et SF2 chez l'humain, la levure à bourgeon (*S. cerevisiae*), *E. coli* et certains virus. Les motifs des protéines SF1 et SF2 sont situés aux mêmes positions au niveau du repliement de RecA (partie B). **B** La conservation de séquence à travers les motifs caractéristiques des hélicases. La hauteur des acides aminés reflète le niveau de conservation à cette position, plus la taille de la lettre est grande plus son niveau

de conservation l'est. Les couleurs indiquent les propriétés chimiques de l'acide aminé: vert et mauve – polaire, bleu – basique, rouge – acide et noir – hydrophobe. **C** Position des motifs caractéristiques à travers le repliement des domaines du noyau hélicase similaire à RecA. Les feuillets β sont représentés par des flèches et les hélices α par des cylindres. Les feuillets β du premier domaine similaire à RecA sont numérotés selon leur position au niveau de la structure primaire. La position des motifs caractéristiques est indiquée par des cercles numérotés et colorés comme dans la partie A. Le feuillet β et l'hélice α colorés en bleu à l'extrémité droite dans la représentation SF2 indique l'absence de cette partie à travers plusieurs familles de SF2.

Adapté de Margaret E Fairman-Williams, Ulf-Peter Guenther & Eckhard Jankowsky. *SF1 and SF2 helicases: family matters*. Curr. Opin Struct Biol, 2007, **17**:316-324. doi: 10.1016/j.sbi.2007.05.007.

1.2.6.3 Hélicases DEAD-box et DEAH-box

La plus grande famille, SF2, est formée majoritairement des DEAD-box et des protéines apparentées DEAH-box et hélicases à ARN DExD/H. Ces protéines partagent au minimum 12 motifs de séquence caractéristiques, certains motifs sont conservés par l'ensemble de la famille SF2 alors que certains varient entre les sous-groupes (Linder et Jankowsky, 2011). Notamment, les membres de la famille DEAD et les variations DExD/H sont caractérisés par la présence du motif Asp-Glu-Ala-Asp/His (DEAD/H) au niveau du motif II; responsable de l'activité ATPase de l'hélicase. Les motifs de séquence conservés sont représentés dans la figure 13A à partir de la structure de l'hélicase DEAD-box Vasa chez la drosophile. Il est possible d'observer que les motifs en verts, les motifs Ia, Ib, IV et V, sont responsables des interactions avec l'ARN alors que les motifs en rouges, les motifs I, II, III, VI et Q, sont responsables des interactions avec l'ATP, l'hydrolyse de celui-ci et le transfert d'énergie générée par l'hydrolyse d'ATP afin de permettre l'activité hélicase (Hardwick et Luisi, 2013).

L'activité NTPase de ces hélicases est stimulée par la présence de l'ARN, les hélicases DEAD-box utilisent l'ATP en tant que substrat alors que les protéines DEAH peuvent utiliser différents NTPs. Notamment, par la présence du motif Q chez les DEAD-box, celui-ci est responsable de la coordination avec la base adénine alors que l'absence de celui-ci chez les DEAH-box expliquerait l'utilisation non spécifique de l'ATP à l'intérieur de ce sous-groupe. (Tanner *et al.*, 2003)

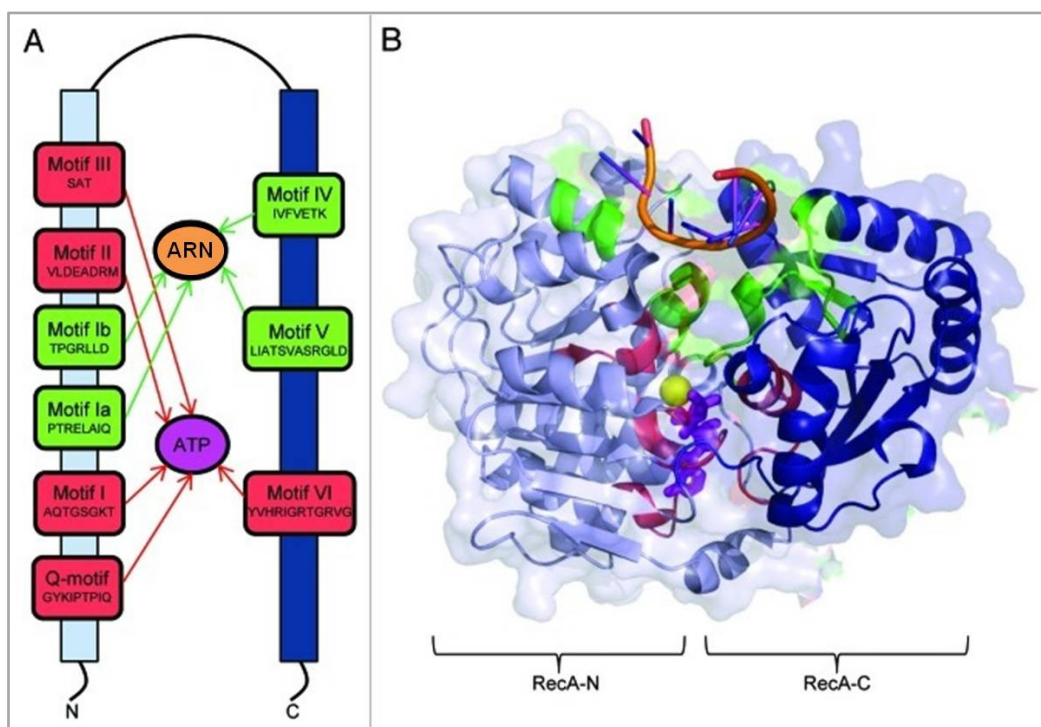


Figure 13. La représentation de la structure d'une hélicase DEAD-box engagé avec l'ARN et l'ATP et le schéma des motifs principaux des hélicases de la superfamille 2 (SF2).

A Une vue schématisée des motifs DEAD-box conservés. Le domaine RecA N-terminale est montré en bleu pâle alors que le domaine RecA C-terminale est en bleu foncé. Les motifs de séquence impliqués dans la liaison à l'ATP/Mg²⁺ sont en rouges, ceux impliqués à l'interaction avec l'ARN sont en verts. **B** La structure cristalline de l'hélicase Vasa chez la drosophile en complexe avec un ARN monocaténaire et de l'ATP non hydrolysable. Les éléments de structures secondaires de l'hélicase sont représentés en style cartoon, la surface protéique est semi-transparente. Les motifs de liaison à l'ARN conservés sont montrés en vert et l'ARN en orange. Les motifs impliqués dans la liaison à l'ATP et l'hydrolyse sont colorés en rouges, l'ATP non hydrolysable est en magenta et l'ion de magnésium est représenté selon une sphère jaune.

Adapté de Steven W. Hardwick & Ben F. Luisi, *Rarely at rest: RNA helicases and their busy contributions to RNA degradation, regulation and quality control*. RNA Biology, 2013, **10**:1, 56-70. doi: 10.4161/rna.22270.

1.2.6.4 Structure et mécanisme

La première vue structurale publiée sur l'organisation des domaines RecA chez les hélicases DEAD-box lors de l'interaction avec l'ARN et l'ATP fut la structure cristalline de la protéine

Vasa de *D. melanogaster* en complexe avec une molécule d'ARNsb et d'ATP non hydrolysable, l'AMP-PNP (Figure 13B) (Sengoku *et al.*, 2006). Cette structure cristalline et les suivantes de protéines DEAD/H-box, en plus de données biophysiques, ont permis de révéler que les deux domaines RecA n'ont pas d'orientation définitive sous la forme apo et ces domaines sont également reliés par un domaine flexible. Cependant, l'interaction avec les molécules d'ARN et d'ATP fait en sorte que les domaines RecA se rejoignent provoquant ainsi la formation d'une clé dans laquelle le nucléotide s'imbrique (le nucléotide est représenté en magenta dans la figure 13B). À l'intérieur, le lien ester du γ -phosphate de l'ATP est disponible pour l'attaque hydrolytique (Sengoku *et al.*, 2006; Del Campo et Lambowitz, 2009; Andersen *et al.*, 2006; Bono *et al.*, 2006; von Moeller, Basquin et Conti, 2009; Fan *et al.*, 2009). Le changement conformationnel provoqué par la liaison de l'ATP génère un site de liaison à l'ARN distal de celui de la liaison à l'ATP, l'ARN lié à cette région pontre les deux domaines RecA (haut de la figure 13B). Les domaines RecA se lient principalement au squelette de sucre-phosphate de l'ARN, les interactions sont orchestrées par un motif structural dont l'origine évolutive proviendrait d'un ancien module de liaison anionique (Milner-White, Pietras et Luisi, 2010). L'activité de déroulement de l'ARN des hélicases DExD/H est normalement restreinte à moins de deux tours complets, ainsi même à leur maximum d'activité elles sembleraient peu efficaces. Il est possible d'observer le tout par l'entremise de la structure cristalline de Vasa qui montre que le site d'interaction avec l'ARN n'est pas extensif. Par ailleurs, la surface de liaison à l'ARN restreinte limiterait la capacité de l'hélicase à s'emparer une région monocaténaire exposé de manière opportuniste. De ce fait, les hélicases ne sembleraient pas préférentiellement dérouler dans une direction donnée, c'est-à-dire $5' \rightarrow 3'$ ou l'inverse (Jarmoskaite et Russell, 2011).

Les interactions entre les domaines RecA et le squelette de l'ARN n'impliquent pas de reconnaissance de bases azotées, il ne s'agit donc pas d'une interaction séquence spécifique. Néanmoins, les différentes hélicases DExD/H-box sont en mesure de reconnaître spécifiquement un ARN particulier et ont des rôles bien définis *in vivo*. Il a été déterminé que les régions N et C-terminales d'une variété d'hélicases de la famille des DEAD/H-box comportaient des séquences leur conférant certaines fonctions spécifiques d'interactions à des structures d'ARN ou bien des protéines régulatrices (Diges et Uhlenbeck, 2001; Tsu,

Kossen et Uhlenbeck, 2001; Kossen, Karginov et Uhlenbeck, 2002; Wang et Guthrie, 1998; Schneider et Schwer, 2001; Ohno et Shimura, 1996). Notamment, l'hélicase DHX36 possède un motif spécifique (RSM) à l'intérieur de la région N-terminale responsable de l'interaction avec les structures G4 (Meier *et al.*, 2013).

1.2.7 DHX36

1.2.7.1 Caractéristiques

La découverte de cette protéine remonte à 1997 par le groupe d'Akman. La protéine résultante de l'expression du gène *DHX36* possède plusieurs noms couramment utilisés: RHAU (*RNA helicase associated with AU-rich elements*), G4R1 (*G4 resolvase 1*) et DHX36 (*DEAH-box protein 36*). DHX36 possède une activité catalytique de dépliement des structures G-quadruplex autant au niveau de l'ADN que l'ARN (Vaugh *et al.*, 2005). Il s'agit de la seule hélicase répertoriée qui utilise les structures G4 d'ARN comme substrat (Creacy, 2008). Son activité nécessite la présence de l'adénosine triphosphate (ATP) et du magnésium (Mg^{2+}); elle est dite ATP et Mg^{2+} -dépendante. DHX36 fait partie de la grande famille d'hélicases SF2 et du sous-groupe DEAH-box. DHX36 est majoritairement retrouvée à l'intérieur du noyau et une séquence dans la partie N-terminale de la protéine est responsable de l'import de celle-ci au noyau (Tran *et al.*, 2004 et Iwamoto *et al.*, 2008). Il a été démontré que *DHX36* est un gène essentiel. En effet, Lai *et al.* ont délété ce gène au niveau de cellules germinales de souris ayant pour conséquence une létalité embryonnaire (Lai *et al.*, 2012). Le rôle physiologique de l'expression de ce gène est méconnu, la nécessité d'accroître les connaissances sur la protéine DHX36 au niveau du métabolisme des structures G4 s'avère importante.

1.2.7.2 Structure

Au niveau structural, DHX36 comporte un domaine central hélicase d'environ 400 acides aminés incluant tous les motifs spécifiques de la famille des hélicases DEAH-box (Figure 14). Le noyau hélicase regroupe notamment les motifs conservés responsables de l'hydrolyse de l'ATP (motifs I, II et VI), de la liaison aux acides nucléiques (motifs Ia, Ib IV et V) et finalement le motif III qui couple l'énergie issue de l'hydrolyse de l'ATP à la séparation des brins nucléiques (Hall et Matson, 1999). La région C-terminale est d'une longueur de quelques 400 acides aminés alors que la région N-terminale est de 200 acides aminés. Le rôle

de portion C-terminale de la protéine reste toujours méconnu. Cependant, la portion N-terminale possède des éléments impliqués dans la reconnaissance spécifique des G4 *in vitro*. En effet, il a été affirmé que la séquence RSM (*RHAU-specific motif*) était impliquée dans la reconnaissance des structures G4. Son abolition diminue la liaison et l'activité de dépliement des G4 de la protéine DHX36 (Lattmann *et al.*, 2010). La structure du motif RSM a par ailleurs été déterminée par RMN et diffraction des rayons X (Meier *et al.*, 2013).

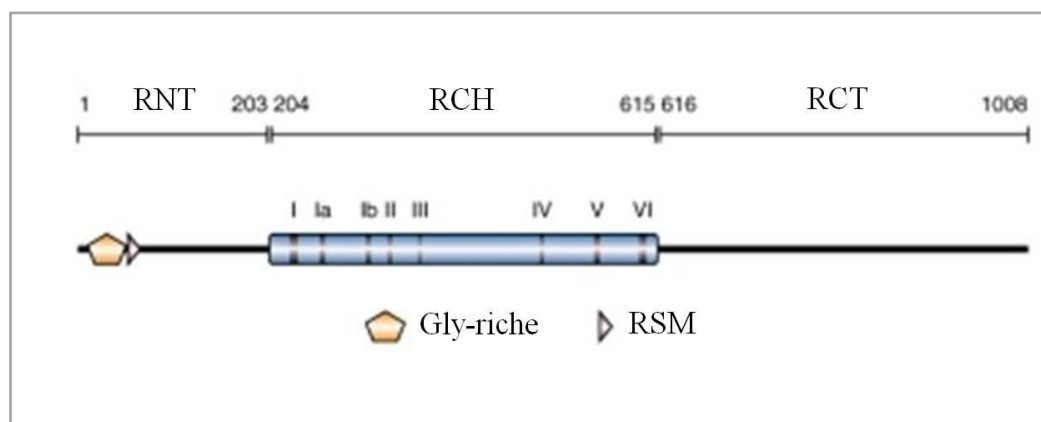


Figure 14. Représentation schématique des 1008 acides aminés de la protéine DHX36.

Les motifs ATPase/hélicase I-VI conservés de la famille DEAD/H-box sont indiqués par des barres à l'intérieur de la région centrale hélicase (RCH). La RCH est flanquée par les régions N-terminale (RNT) et C-terminale (RCT) d'une longueur de 203 et 393 acides aminés respectivement. Les séquences Gly-riche (acides aminés 10-51) et RSM (*RHAU specific motif*) sont indiqués dans la région N-terminale.

Adapté de Simon Lattmann, Banabihari Giri, James P. Vaughn, Steven A. Akman & Yoshikuni Nagamine. *Role of the amino terminal RHAU-specific motif in the recognition and resolution of guanine quadruplex-RNA by the DEAH-box RNA helicase RHAU*. *Nucleic Acids Research*, 2010, Vol. 38, No. 18 6219-6233. doi: 10.1093/nar/gkq372.

1.2.7.3 Conservation

Selon l'analyse de séquence de Lattman *et al.*, DHX36 possède des orthologues à travers pratiquement toutes les espèces, passant des organismes eucaryotes unicellulaires tels que les choanoflagellés vers des organismes pluricellulaires plus complexes comme l'humain (Lattman *et al.*, 2010). Ils ont effectué un alignement de séquence de 8 orthologues qui montrent les régions conservées au cours de l'évolution; notamment les motifs du domaine central hélicase et la région C-terminale. La figure 15 montre les séquences conservées de DHX36.

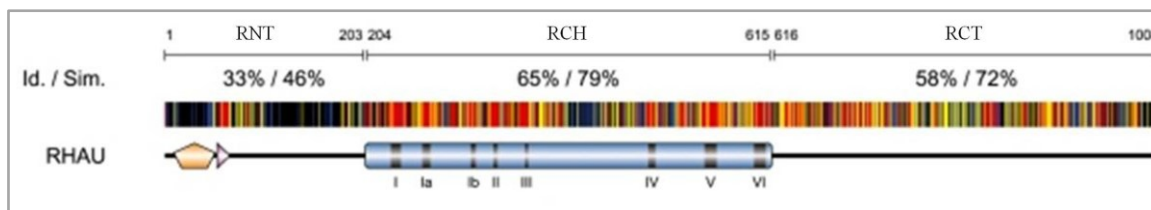


Figure 15. Conservation en acides aminés de la protéine DHX36.

Représentation schématique de la conservation en acides aminés de la protéine humaine DHX36 à travers l'évolution. La protéine DHX36 a été alignée avec huit orthologues de DHX36 par le groupe de Nagamine en utilisant MAFFT (version 6). Chacun des résidus de DHX36 est représenté par un code de couleur qui indique le niveau de conservation parmi les huit séquences orthologues. La similarité est montrée en rouge pour 100%, en jaune pour 99-80% et en bleu pour 79-60%.

Adapté de Simon Lattmann, Banabihari Giri, James P. Vaughn, Steven A. Akman & Yoshikuni Nagamine. *Role of the amino terminal RHAU-specific motif in the recognition and resolution of guanine quadruplex-RNA by the DEAH-box RNA helicase RHAU*. *Nucleic Acids Research*, 2010, Vol. 38, No. 18 6219-6233. doi: 10.1093/nar/gkq372.

1.2.7.4 Fonctions probables

Tout d'abord, DHX36 a été identifiée en tant que protéine apte à lier et dérouler les structures G4 intermoléculaires d'ADN sous forme monocaténaire (Vaughn *et al.*, 2005). Par la suite, il s'est avéré que les structures G4 intramoléculaires issues des promoteurs de deux gènes cibles (*c-Myc* et *Zic1*) sont déroulées par DHX36. Ensuite, il a été révélé que DHX36 possédait également une fonction de dépliement des structures G4 d'ARN (Creacy *et al.*, 2008). Ce faisant, les fonctions potentielles de DHX36 se multiplient. L'invalidation partielle ou *knockdown* de DHX36 a révélé avoir un impact significatif, autant positivement que négativement, sur les niveaux d'ARNm de plusieurs gènes détecté par micro-puces (Iwamoto *et al.*, 2008). Ainsi, DHX36 aurait un rôle prépondérant au niveau de la régulation de l'expression d'ARNm que leur stabilité.

Par ailleurs, DHX36 est connu pour avoir un rôle global au niveau de la régulation de l'expression des ARNm autant d'un point de vue traductionnel que de la stabilité des molécules d'ARNm (Tran, 2004). DHX36 est en mesure d'interagir avec la déadénylase PARN et l'exosome afin d'induire la déadénylation et la dégradation des ARNm dont la région non traduite à l'extrémité 3' est riche en adénine et en uracile.

Le groupe de recherche d'Akman a utilisé la méthode de RIP-chip afin de déterminer les ARN qui interagissent avec DHX36 *in vivo*. Il a été montré qu'un des ARN les plus abondants qui interagit avec DHX36 était l'ARN matrice de la télomérase (TERC) (Lattmann, 2011). La région 5' de TERC possède une séquence formant un G4 intramoléculaire, celle-ci est reconnue et liée par DHX36. Ainsi, DHX36 est associé à l'holoenzyme télomérase par l'entremise de l'ARN TERC. De plus, le groupe de McKenna a renchérit en démontrant l'implication de DHX36 dans le processus du maintien des télomères en effectuant un *knockdown* de DHX36 ayant pour résultat une diminution de la longueur des télomères (Booy *et al.*, 2012).

1.2.7.5 Déplieement des G4 d'ARN

La protéine DHX36 est une des seules hélicases répertoriées jusqu'à présent ayant la fonction de déplieement des ARN G4 (Creacy *et al.*, 2008). Sachant que les fonctions biologiques des ARN G4 regroupent notamment l'activation ou la répression de la traduction de certains transcrits d'oncogènes répertoriés tels que *NRAS* (Kumari *et al.*, 2007), *TGFβ2* (Agarwala *et al.*, 2013) et *YY1* (Huang *et al.*, 2012), l'impact de la stabilité des ARN G4 par DHX36 s'avère un sujet d'étude utile. Puisqu'il s'agit d'une fonction unique à cette hélicase, il est intéressant d'étudier le mécanisme du déplieement des structures G4 d'ARN.

1.2.7.6 Mécanisme

Le groupe de recherche de Balasubramanian, par l'entremise de la stabilité thermique des structures G4, a proposé le mécanisme du déplieement des structures nucléiques G4 par l'hélicase DHX36 présenté par la figure 16. Leurs travaux révèlent que le taux de déplieement de DHX36 est sensible à la stabilité thermique de la structure du G4 et que l'utilisation de l'ATP par l'hélicase est indépendante de la stabilité du G4 (Chen *et al.*, 2015). Ainsi, DHX36 déroule les structures G4 de manière non-processive comme le ferait notamment l'hélicase à DEAD-box eIF4A spécifique pour les ARNdb (Rogers, Richter et Merrick, 1999).

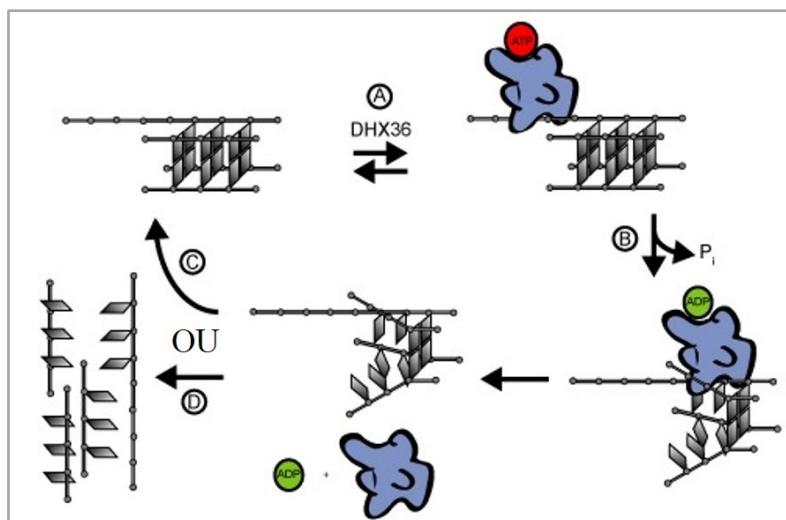


Figure 16. Modèle du dépliement des G4 dépendant de DHX36.

A DHX36 lie la région 3' du G4 indépendamment de la présence d'ATP. **B** L'hydrolyse d'ATP entraîne le dépliement partiel du G4 et la dissociation de DHX36 générant un G4 instable. **C** et **D** Étant donné la non processivité de DHX36, le G4 est seulement partiellement déroulé. Ainsi, le G4 peut se régénérer ou bien se dénaturer complètement selon la stabilité du G4 dans l'état déstabilisé.

Adapté de Michael C. Chen, Pierre Murat, Keren Abecassis, Adrian R. Ferré-D'Amaré & Shankar Balasubramanian. *Insights into the mechanism of G-quadruplex-unwinding DEAH-box helicase*. *Nucleic Acids Research*, 2015, Vol. 43, No. 4 2223-2231. doi: 10.1093/nar/gkv051.

Par conséquent, les fondements du dépliement des structures G4 par l'hélicase DHX36 résident particulièrement sur la liaison du substrat de G4 par DHX36, l'hydrolyse de l'ATP et le transfert de l'énergie libérée afin de déstabiliser la structure d'ARN G4. Dans un premier temps, certaines études ont montré la liaison de la structure d'ARN G4 par DHX36 sans toutefois la caractériser. De plus, aucune étude n'a caractérisé l'utilisation d'ATP par l'hélicase DHX36.

1.2.7.7 Liaison de l'ARN G4 et utilisation de l'ATP

Le groupe de recherche d'Akman a fait la découverte que DHX36 était en mesure de lier l'ARN G4 (Creacy *et al.*, 2008). Ils ont utilisé la méthode de rétention sur gel afin d'identifier et d'analyser l'interaction entre DHX36 et l'ARN G4. Cependant, il s'agissait d'une caractérisation sommaire de cette interaction.

Le groupe d'Akman a également déterminé que l'activité du dépliement de structures G4 par DHX36 était dépendante des NTPs (Vaugh *et al.*, 2005). Cependant, la liaison et l'utilisation des nucléotides n'ont pas été caractérisées.

1.3 Hypothèse/problématique

La déstabilisation des structures G4 au niveau des acides nucléiques a des répercussions physiologiques importantes. L'accentuation des connaissances concernant les processus cellulaires associés au métabolisme des structures des G4 est primordiale. Une panoplie d'hélicases à G4 est impliquée dans le métabolisme des structures G4, notamment l'hélicase humaine DHX36. Il a été déterminé au préalable par certains groupes de recherche que l'hélicase DHX36 se lie à son substrat l'ARN G4 et utilise des NTPs afin de catalyser le dépliement de la structure G-quadruplex. Toutefois, l'interaction avec l'ARN G4 a été sommairement caractérisée et la spécificité nucléotidique n'a toujours pas été évaluée. Ainsi, nous avons décidé d'approfondir les connaissances du mécanisme de dépliement de la structure du G4 d'ARN par l'hélicase DHX36. Notamment, en évaluant la thermodynamique de l'interaction entre l'hélicase et l'ARN G4 afin de révéler particulièrement l'efficacité de liaison mais également en évaluant la spécificité nucléotidique de l'hélicase DHX36 afin d'effectuer le dépliement de l'ARN G4.

1.3.1 Objectifs

Évaluer la thermodynamique de l'interaction entre l'hélicase DHX36 et l'ARN G4.

Utiliser la technique de spectroscopie à fluorescence afin de déterminer les caractéristiques thermodynamiques de l'interaction entre l'hélicase et l'ARN G4.

Évaluer la spécificité nucléotidique de l'hélicase DHX36.

Utiliser une série d'analogues de nucléotide afin d'analyser la nature de la pochette de liaison nucléotidique ainsi que la spécificité de l'hydrolyse de NTP par l'hélicase DHX36. Nous utilisons également la modélisation structurale afin d'étudier la liaison de l'ATP par la protéine en déterminant les résidus requis pour l'interaction avec l'ATP.

CHAPITRE 2 – MATÉRIEL ET MÉTHODES

2.1 Clonage, expression et purification de DHX36 dans *E. coli*.

Un plasmide pour l'expression de l'hélicase de structure G4 humaine DHX36 a été généré en introduisant la séquence entière du gène DHX36 à l'intérieur du plasmide d'expression pET-28a (Novagen). La séquence code pour une protéine ayant 1008 acides aminés. Celle-ci a été clonée entre les sites de clonage NcoI et XhoI du plasmide pET-28a. Une étiquette de 6 résidus d'histidine en tandem a été ajoutée à l'extrémité C-terminale de la protéine et l'expression de DHX36 His₆ est modulée par le promoteur de l'ARN polymérase T7. Le plasmide recombiné résultant (pET28a-DHX36) a été transformé dans *Escherichia coli* BL21 (DE3). Une culture de 3000 ml de *E. coli* BL21(DE3)/pET28a-DHX36 a crû à 37 °C dans le milieu de culture Luria-Bertani (LB) supplémenté avec 30 µg/ml de l'antibiotique kanamycine jusqu'à l'obtention d'une densité optique à 600nm (DO₆₀₀) de 0,5. La culture a ensuite été supplémentée avec de l'IPTG (isopropyl-β-D-thiogalactoside) à une concentration finale de 0,4 mM et 2% (v/v) d'éthanol, l'incubation s'est poursuivie à 18 °C pendant 18h. Les cellules ont été récoltées par centrifugation à 5000 g pendant 5 minutes en utilisant un rotor Sorval SLA-1500. Les procédures subséquentes ont été réalisées à 4 °C. Le culot de bactéries a été resuspendu dans 50 ml de tampon de purification sans Triton [50 mM Tris-HCl, pH 7,5, 150 mM NaCl et 10% (m/v) sucrose]. La lyse cellulaire a été réalisée par l'addition de lysozyme et de Triton X-100 pour des concentrations finales de 50 µg/ml et 0,1% (v/v) respectivement. Les lysats ont été soniqués afin de réduire la viscosité et tout matériel insoluble a été retiré par centrifugation à 13000 rpm pendant 45 minutes en utilisant un rotor Sorval SS-34. L'extrait soluble a été déposé dans une colonne de 5 ml de Ni-NTA (Ni²⁺ - nitroacétate)-agarose (Qiagen) qui avait été préalablement équilibré avec le tampon de purification avec 0,1% (v/v) de Triton X-100. La colonne a été ensuite lavée avec ce même tampon et éluée séquentiellement, par gravité, avec le tampon d'éluion [50 mM Tris-HCl, pH 8,0, 100 mM NaCl et 10% (v/v) glycérol] contenant 50, 100, 200, 500 et 1000 mM d'imidazole. La composition en protéines de chacune des fractions éluées a été analysée par SDS-PAGE coloré au bleu de Coomassie. Un immunobuvardage a été effectué afin de déterminer les fractions parmi lesquelles la protéine recombinante a été éluee. L'immunobuvardage a été réalisé en séparant les protéines de chacune des fractions sur SDS-

PAGE 10% puis transféré sur une membrane de PVDF (*polyvinylidene fluoride membrane*). La membrane a été bloquée et incubée pendant la nuit avec un anticorps primaire de souris 1 :1000 dirigé contre 5 résidus d'histidine (Abcam) à 4°C. Ensuite, la membrane fut incubée pendant 1h en présence d'un anticorps secondaire 1 :5000 de chèvre HRP-anti-souris (GE Healthcare). La révélation a été effectuée par détection au ECL (Perkin-Elmer) puis le tout fut exposé à un appareil détectant la chimiluminescence (GE Healthcare). La protéine recombinante a été récupérée parmi les fractions 200 et 500 mM d'imidazole. Une dialyse s'ensuivit en utilisant le tampon de dialyse [50 mM Tris-HCl, pH 8,0, 50 mM NaCl, 2 mM dithiothreitol, 10% (v/v) glycérol et 0,05% (v/v) Triton X-100]. La préparation purifiée a été conservée à -80°C. La concentration protéique a été déterminée en employant la méthode du colorant Bio-Rad, en utilisant l'albumine sérique de bœuf (BSA) comme standard.

2.2 Clonage, transfection, expression transitoire et purification de DHX36 dans HEK293T.

Le vecteur pCDNA3 (Invitrogen Corporation) contenant l'ADNc de DHX36 avec une étiquette FLAG (DYKDDDDK) en C-terminale a été généré. Le plasmide recombiné résultant (pCDNA3-DHX36) a été transfecté dans des cellules HEK293T. Trois jours avant la transfection, 10 cultures de HEK293T (American Type Culture Collection, Manassas, VA, USA) ont été incubées dans des plaques de 150 mm à 37°C dans un incubateur à 5% CO₂ humidifié afin d'obtenir une confluence de 80% au moment de la transfection. Les cellules ont été cultivées en utilisant du milieu DMEM (*Dubelcco's Modified Eagle's Medium*) (Invitrogen, Burlington, ON, Canada) supplémenté avec 10% FBS (*Fetal Bovine Serum*), 100 U/ml Pénicilline et 100 µg/ml Streptomycine. La veille de la transfection, le milieu de culture a été remplacé par 26 ml de DMEM supplémenté avec 10% FBS sans antibiotique. 40 ml de mélange de transfection a été préparé à partir de 40 ml de DMEM sans sérum dans lequel 600 µg de plasmide recombiné (pCDNA3-DHX36-FLAG) et 600 µl de réactif de transfection Turbofect (Thermo-Fischer Scientific) ont été ajoutés. Le mélange de transfection a été incubé à température pièce pendant 15 minutes. Ensuite, 4 ml du mélange de transfection a été ajouté goutte à goutte aux plaques de culture. Le lendemain les cultures transfectées ont été séparés 1 :2 parmi un total de 20 plaques de culture de 150mm puis incubé 24h afin d'obtenir une confluence de 100%.

La purification de DHX36 s'ensuivit. Les cellules ont été resuspendues avec 10 ml de PBS (*Phosphate buffered saline*) [10 mM Na₂HPO₄, 1,8 mM KH₂PO₄, 137 mM NaCl, 2,7 mM KCl, pH 7,4] froid par plaque. Les cellules ont été culotées par centrifugation à 1000g pendant 5 min à 4°C puis culotées à nouveau dans un seul tube par centrifugation. Le culot total a été resuspendu dans 3 ml de tampon de lyse 1 % [10 mM Na₂HPO₄, 1,8 mM KH₂PO₄, 600 mM NaCl, 2,7 KCl et 1% IGEPAL (NP-40), pH 7,4] puis 27 ml d'un mélange de PBS + 600 mM NaCl et d'inhibiteurs de protéases 1x (Thermo-Fischer Scientific). Le lysat fut par la suite soumis au vortex et soniqué. Une centrifugation s'ensuivit à 20 000g pendant 15 min à 4°C afin de culoter les débris insolubles. L'extrait soluble a été déposé dans une colonne de 2 ml de résine d'affinité Anti-FLAG M2 (Sigma-Aldrich) qui avait été préalablement lavée à trois reprises avec une solution de glycine 0,1 M pH 3,5 puis équilibrée avec le tampon de lyse 0,1% (v/v) de NP-40. La colonne a été ensuite lavée avec 15 ml de ce même tampon et éluée par gravité avec 10 ml du tampon d'éluion du peptide FLAG [0,2 mg/ml peptide FLAG, 50 mM Tris-HCl pH 7,5, 50 mM KCl, 5 mM DTT et 20% glycérol] en récoltant des fractions de 1 ml. Le contenu protéique de chacune des fractions fut dialysé en utilisant le tampon de conservation [50 mM Tris-HCl pH 7,5, 50 mM KCl, 5 mM DTT et 20% glycérol]. La composition en protéines de chacune des fractions éluées est monitorée par des essais de Bradford. Les aliquotes de DHX36-FLAG furent conservés à -80°C. Le contenu protéique de chacune des fractions a été analysé par SDS-PAGE 8% coloré au bleu de Coomassie.

2.3 Formation du complexe tétramoléculaire d'ARN G4

Les étapes de production des complexes d'ARN G4 ont été basées sur les expérimentations de Creacy *et al.*, 2008. Les oligoribonucléotides utilisés ont pour séquence : 5' – AAA AAA AAA AAA AAA GGG GAA AAA AAA AAA A – 3'. Ceux-ci ont été achetés de la compagnie Integrated DNA Technologies. Les oligoribonucléotides ont été suspendus pour obtenir une concentration de 0,5 mM dans une solution de 10mM Tris-HCl, 1 mM EDTA, pH 7,5 sans ribonucléase (RNase). La solution d'oligoribonucléotides a été aliquotée dans des tubes à PCR et incubés dans un thermocycleur (Eppendorf) à 98 °C pendant 10 minutes puis maintenue à 80°C. Les tubes sont immédiatement ouverts afin d'ajouter de l'EDTA, pH

8 afin d'obtenir une concentration finale de 25 mM. Les tubes ont été refermés puis graduellement refroidis afin d'atteindre la température pièce. Les aliquotes ont été combinés et réservés à 4°C pendant 2-3 jours pour finalement être conservés à -20°C.

2.4 Étiquetage au ^{32}P de l'extrémité 5' de l'ARN G4

Les oligonucléotides d'ARN G4 5'-[^{32}P] a été obtenu en incubant 1500 μmol d'oligonucléotides non-étiquetés d'ARN G4 avec 3 unités de la polynucléotide kinase T4 (NEB) et 1,6 μmol de [γ - ^{32}P]ATP (6000 Ci/mmol; Perkin-Elmer) pendant 1 heure à 37°C. Les oligonucléotides d'ARN G4 étiquetés furent directement purifiés en utilisant les colonnes Illustra ProbeQuant G-50 Micro Columns (GE HealthCare) puis conservés à -20°C. Les colonnes ont été équilibrées au préalable avec le tampon fourni par le manufacturier soit le tampon STE [150 mM NaCl, 10 mM Tris-HCl, pH 8,0 et 1 mM EDTA]. La purification sur colonne G-50 a été réalisée telle que décrite par le fournisseur.

2.5 Essais de dépliement des structures G4 d'ARN

La protéine recombinante DHX36 à une concentration de 300 nM (2 μl) fut incubée en présence de 50 μM d'ARN G4 étiqueté en 5' au [^{32}P] dans le tampon K-Res [100 mM KCl, 10 mM NaCl, 3 mM MgCl_2 , 50 mM Tris-acétate, pH 7,8, 70 mM glycine, 0.012%, α -lactalbumine de boeuf, 10% glycérol] (Giri *et al.*, 2011) en présence de 10 à 5000 μM d'ATP pendant 30 minutes à température pièce. Un tampon de chargement non dénaturant contenant du bleu de bromophénol et du xylène cyanol fut ajouté aux produits de réaction d'hélicase. Celles-ci furent ensuite analysées par électrophorèse sur gel polyacrylamide non dénaturant 6% supplémenté au KCl (100 mM). L'électrophorèse a été réalisée à 90 volts pendant 2h45 dans une chambre froide (4 °C) dans un tampon de TBE supplémenté au KCl (100mM). Les gels ont été exposés sur des écrans photostimulables (Amersham Biosciences) puis révélés par l'imageur biomoléculaire Typhoon (Molecular Dynamics). Différents analogues furent utilisés afin d'évaluer l'efficacité de la réaction de dépliement des G4 d'ARN relativement à l'utilisation de l'ATP comme substrat. La concentration d'analogues dans les réactions fut de 50 μM ce qui correspond environ à 4 fois le K_m de la réaction en présence d'ATP. Les réactions d'hélicase furent effectuées 2 fois indépendamment. L'intensité des bandes fut

analysée par le logiciel ImageQuant (GE Healthcare) et l'analyse statistique fut effectuée à l'aide du logiciel GraphPad Prism.

2.6 Analogues de nucléotides

Tous les nucléotides utilisés ont été achetés chez Jena Biosciences (Allemagne) et TriLink Biotechnologies (États-Unis).

2.7 Modélisation par homologie

La prédiction de structure a été obtenue à partir de l'outil de modélisation disponible gratuitement en ligne : I-TASSER (<http://zhanglab.ccmb.med.umich.edu/I-TASSER/>) (Zhang, 2015). La structure cristalline de l'hélicase d'ARN Prp43p de *S. cerevisiae* liée à de l'ADP a notamment été utilisée comme matrice par le logiciel I-TASSER pour modéliser le domaine hélicase de DHX36. Les coordonnées atomiques furent obtenues par le fichier 2XAU de Protein Data Bank, représentant la structure cristalline du complexe entre le domaine hélicase de la protéine Prp43p, du magnésium et de l'ADP. La structure 3D prédite du site actif de l'hélicase DHX36 a été générée en utilisant Pymol.

CHAPITRE 3 – RÉSULTATS

3.1 Expression, purification et activité de dépliement des G4

Dans le but d'analyser l'interaction entre DHX36 et son substrat l'ATP *in vitro*, via son activité catalytique de dépliement des G4, une quantité pure et suffisante était nécessaire. Des études biochimiques antérieures portant sur DHX36 ont montré qu'il était possible d'exprimer la protéine recombinante pleine longueur à partir d'un système bactérien (souche Rosetta 2 de *E. coli*) (Vaughn *et al.*, 2005) et d'un système de culture de cellules humaines (HEK293T) (Booy *et al.*, 2015; Lattmann *et al.*, 2010).

3.1.1 Système bactérien (*E. coli*)

Nous avons réalisé l'expression de l'enzyme étiquetée His₆ en N-terminale dans *E. coli*, celle-ci fut purifiée de l'extrait bactérien soluble par chromatographie d'affinité de résine d'agarose Ni-NTA. La protéine purifiée était localisée principalement dans les fractions éluées d'imidazole de 50 et 100 mM (Figure 17A). La localisation fut confirmée par immunobuvardage en utilisant un anticorps monospécifique anti-His (Figure 17B). Cependant, tel que montré par le gel SDS-PAGE coloré au bleu de Coomassie (Figure 17A), la purification et la quantité de la protéine ne sont pas optimales pour effectuer des tests *in vitro* avec celle-ci. Ainsi, DHX36 fut exprimée et purifiée à partir du système de culture de cellules humaines HEK293T.

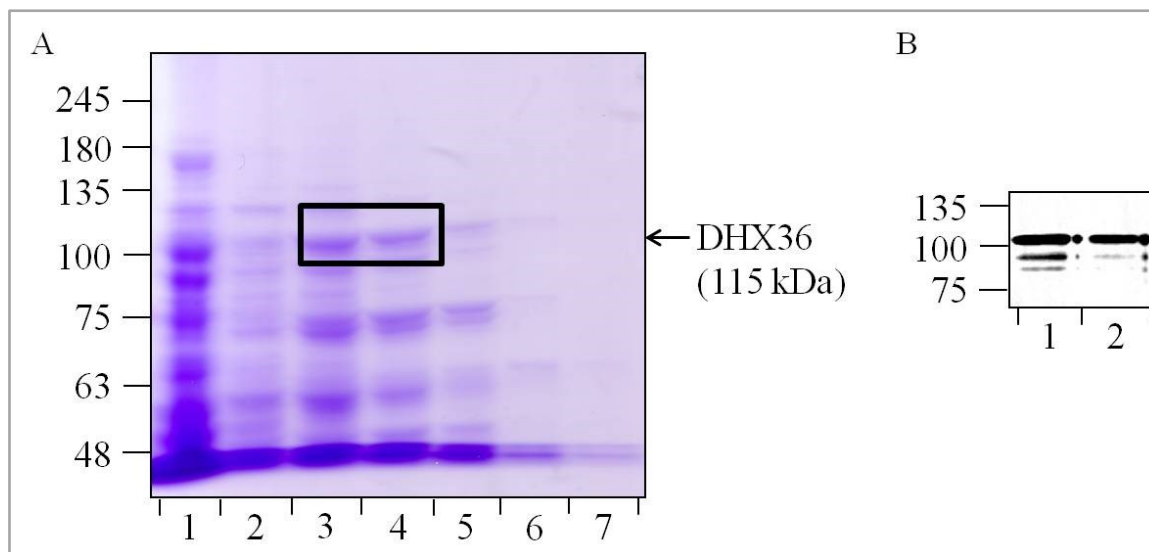


Figure 17. Purification de l'hélicase de G4 d'ARN humaine DHX36 à partir d'un système bactérien (*E. coli*).

A La composition peptidique des différentes fractions de purification analysées par SDS-PAGE. Le gel a été fixé et coloré au Bleu de Coomassie. Piste 1, le *flow-through* de la résine d'agarose Ni-NTA; piste 2, lavage de la résine d'agarose Ni-NTA; piste 3, éluat de 50 mM d'imidazole; piste 4, éluat de 100 mM d'imidazole; piste 5, éluat de 250 mM d'imidazole; piste 6, éluat de 500 mM d'imidazole et piste 7, éluat de 1000 mM d'imidazole. La position et la taille (en kDa) du marqueur de poids moléculaire sont indiquées sur la gauche. **B** La protéine recombinante humaine DHX36 (fractions, piste 1, 50 mM et piste 2, 100 mM d'imidazole) a été analysée par immunobuvardage en utilisant un anticorps monospécifique anti-His.

3.1.2 Système de culture de cellules humaines (HEK293T)

DHX36 étiquetée FLAG en C-terminale dans HEK293T, celle-ci a été purifiée de l'extrait soluble par chromatographie d'affinité de résine d'agarose M2 Anti-FLAG. La protéine purifiée était localisée dans les 9 dernières fractions éluées (Figure 18A). La pureté de la protéine dans chacune des fractions est optimale. Ces 9 fractions furent donc récoltées et combinées, la concentration de la protéine a été évaluée à 178 ng/ μ l. Toutes les expériences subséquentes furent réalisées à partir de cette combinaison de fractions.

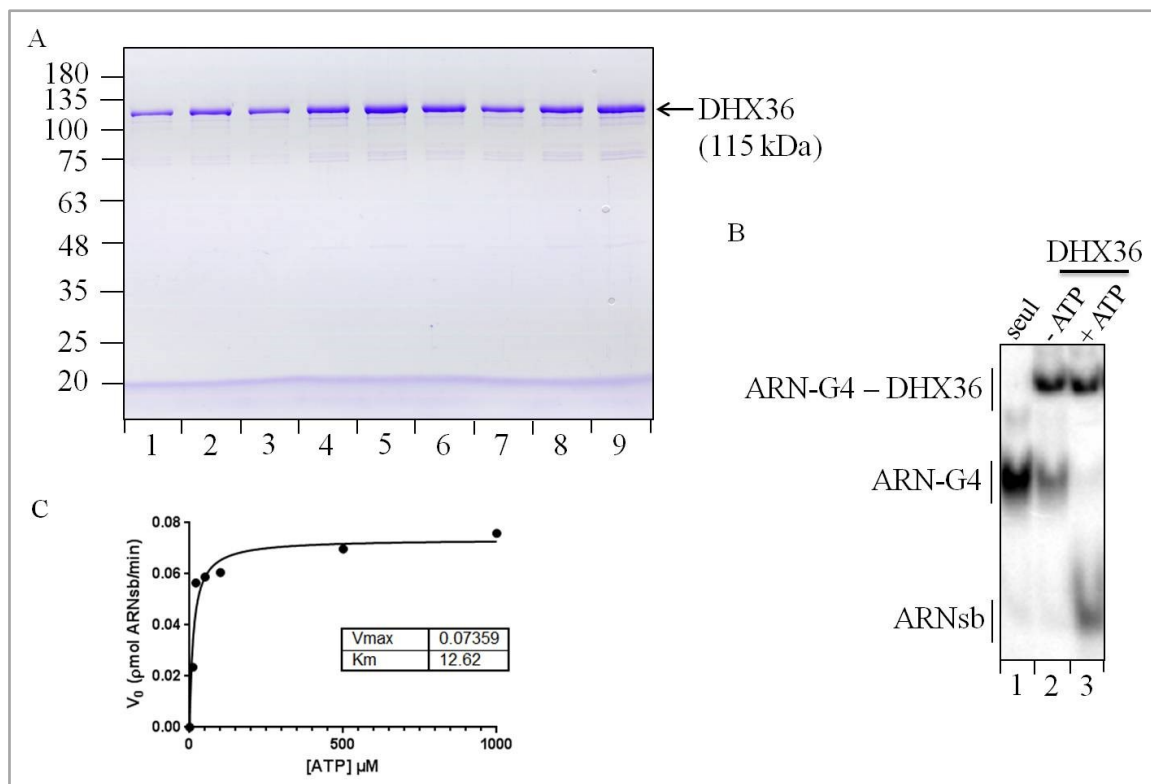


Figure 18. Purification et activité enzymatique de l'hélicase de G4 d'ARN humaine DHX36 à partir de culture de cellules humaines HEK293T.

A La composition peptidique des différentes fractions de purification analysées par SDS-PAGE. Le gel a été fixé et coloré au Bleu de Coomassie. Les 9 pistes correspondent aux 9 dernières fractions sur 10 récupérées suite à l'éluion de l'extrait protéique soluble de cellules HEK293T sur colonne d'agarose M2 anti-FLAG avec une solution contenant 200 $\mu\text{g/ml}$ du peptide FLAG. La position et la taille (en kDa) du marqueur de poids moléculaire sont indiquées sur la gauche. **B** Les essais hélicases de G4 d'ARN furent réalisés sur l'ARN G4 tétramoléculaire. Le mélange réactionnel (10 μl) contenant 50 mM Tris-acétate (pH 7,8), 100 mM KCl, 10 mM NaCl, 3 mM MgCl_2 , 70 mM glycine, 0.012% α -lactalbumine de boeuf, 10% glycérol, 50 μM d'ARN G4 étiqueté en 5' au ^{32}P et avec (pistes 2 et 3) ou sans (piste 1) 300 nM de l'enzyme recombinante DHX36 en présence (piste 2) ou en absence (piste 3) de 1 mM d'ATP pendant 30 minutes à température pièce. Les produits de réaction ont été analysés par électrophorèse sur gel polyacrylamide 6% non dénaturant supplémenté au KCl (100 mM). Un autoradiogramme du gel est présenté, la position du substrat, l'interaction du substrat-enzyme et le produit de réaction sont également présentés. **C** Une quantité croissante d'ATP fut ajoutée en présence d'une quantité fixe de G4 d'ARN et d'enzyme. Les mélanges réactionnels contenant 50 μM d'ARN G4 étiqueté en 5' au ^{32}P , 300 nM d'enzyme et d'une quantité croissante d'ATP de 10 à 5000 μM ont été incubés pendant 30 min à température pièce. Les produits de réaction furent analysés par électrophorèse sur gel polyacrylamide 6% non dénaturant. L'autoradiogramme résultant fut analysé par le logiciel ImageQuant (GE Healthcare).

3.2 Activité enzymatique de l'hélicase DHX36

Nous avons démontré que l'enzyme recombinante humaine DHX36 produite à partir de cellules humaines en culture (HEK293T) était en mesure de dérouler les structures d'ARN G4 tétramoléculaire. La figure 18B montre que l'enzyme purifiée est apte à lier son substrat et de le dérouler en présence d'ATP. En effet, lorsque le substrat d'ARN G4 est mis en présence de l'enzyme (piste 2) il y a retardement sur gel ce qui reflète l'interaction protéine-substrat mais également de l'ARN G4 résiduel puisque celui-ci se retrouve en excès (50 μM) comparativement à la quantité d'enzyme (300 nM). Par contre, lorsque l'ATP est ajouté (piste 3), il y a diminution de la quantité d'ARN G4; substrat de réaction, au profit de l'apparition d'ARN monocaténaire; le produit de réaction. Tel que démontré au préalable (Harrington *et al.*, 1997), l'activité de DHX36 est dépendante de l'hydrolyse de NTP. Nous avons donc examiné la cinétique apparente de l'activité de dépliement des G4 d'ARN en utilisant l'ATP comme cofacteur. Les valeurs apparentes de K_m et V_{max} ont été déterminées en quantifiant le produit de réaction en fonction d'une concentration croissante d'ATP. La quantité de produit d'ARN G4 résolu a été tracée graphiquement selon la concentration d'ATP et les données ont été traitées par l'équation de Michaelis-Menten (Figure 18C). Les valeurs apparentes de K_m et V_{max} obtenues sont 12 μM et 74 fmol d'ARNsb min^{-1} respectivement ce qui est cohérent avec les résultats obtenus d'études précédentes (Vaugh *et al.*, 2005).

3.3 Modèle d'homologie

Étant donné que la structure cristalline de DHX36 n'a toujours pas été élucidée, il est tout de même possible de l'interpréter par l'entremise d'études révélant les structures cristallines de la protéine NS3 des *Flaviviridae* (Luo *et al.*, 2008) et de l'hélicase Prp43p de *S. cerevisiae* (Walbott *et al.*, 2010). Toutefois, la spécificité de ces enzymes pour un NTP particulier, considérant qu'elle soit déterminée par l'interaction entre l'enzyme et la base azotée du substrat NTP, n'a toujours pas été révélée. Nous remarquons dans nos expérimentations que la molécule d'ATP est favorablement utilisée par l'enzyme comparativement au GTP. En effet, l'activité de dépliement de l'ARN-G4 par l'enzyme utilisant l'ATP est environ deux fois plus importante que celle utilisant le GTP (Tableau 3). Nous avons donc généré un

modèle 3D du site actif de DHX36 afin d'évaluer les interactions qui justifient l'utilisation d'ATP le discriminant des autres NTP. Le cristal de l'hélicase apparentée Prp43p de la levure *S. cerevisiae*, possédant 72% de similarité de séquence avec celle de DHX36, liée à l'ADP a été utilisé comme matrice afin de produire le modèle 3D de la protéine DHX36. En effet, la figure 19 montre le modèle 3D obtenu par l'outil de modélisation I-TASSER.

Tableau 3. Niveau d'activité enzymatique selon l'analogue d'ATP.

Nucléotide	Ratio produit/substrat ^a (relative à l'ATP)
ATP	1.00
GTP	0.62 ± 0.05
(A ₂) 2',3'-ddATP	1.02 ± 0.40
(A ₃) 2'-dATP	0.12 ± 0.15
(A ₄) 2'F,2'-dATP	0.24 ± 0.04
(A ₅) ara-ATP	Pas d'activité
(A ₆) 3'-dATP	0.04 ± 0.02
(A ₇) N1-methyl-ATP	0.05 ± 0.03
(A ₈) 2-amino-ATP	0.61 ± 0.49
(A ₉) 2-hydroxy-ATP	0.37 ± 0.35
(A ₁₀) N6-methyl-ATP	0.78 ± 0.75
(A ₁₁) ITP	0.23 ± 0.02
(A ₁₂) 6-mercaptapurine-RTP	0.62 ± 0.06
(A ₁₃) 6-methyl-thio-ITP	0.59 ± 0.55
(A ₁₄) 6-bromo-ATP	Pas d'activité
ATP _γ S	Pas d'activité

Les essais d'activité du dépliement d'ARN-G4 ont été réalisés à deux reprises séparément. Les résultats présentés sont des moyennes des expériences indépendantes effectuées.

^aLes activités spécifiques du dépliement d'ARN-G4 ont été calculées à partir de la formation du produit de réaction comparativement à la présence initiale de substrat en utilisant le logiciel ImageQuant. Les activités ont été normalisées selon l'activité résultante de l'utilisation d'ATP.

La protéine peut être divisée en six domaines formant une structure ressemblant à une pince. Le domaine N-terminale est un domaine spécifique à la protéine DHX36 dans lequel on

retrouve le motif spécifique à DHX36 (RSM), dont il est responsable de l'interaction entre la protéine et son substrat; les G4. Les domaines 2 et 3 correspondent aux deux domaines ubiquitaires des hélicases soient les domaines RecA responsables de la liaison à l'ATP, son hydrolyse et la liaison aux acides nucléiques. La structure de ces domaines est principalement homologue aux hélicases DEAH virales NS3 (Code PDB : 2JLW, Luo *et al.*, 2008). Le motif boucle- β antiparallèle qui sort du plan (mauve dans la figure 19) est retrouvé à l'intérieur du domaine RecA2 entre les motifs V et VI. Cet élément est une particularité seulement conservée à travers les hélicases *DExH-box* puisque celui-ci n'est pas retrouvé chez les hélicases *DEAD-box*. Le domaine 4 a un repliement de type *winged-helix* (WH), celui-ci se presse contre le domaine RecA1 (en orange sur la figure 19). Le domaine 5 est un regroupement de sept hélices qui se lie à travers les deux domaines RecA, il s'agit du domaine « ratchet » (en jaune sur la figure 19) (Büttner *et al.*, 2007). Les domaines WH et ratchet sont annotés à travers les banques de données de domaines à titre de *helicase-associated domain* (HA2), celui-ci est retrouvé à travers toutes les séquences protéiques des hélicases à *DEAH-box* et des *RNA helicase A* (RHA). Le modèle d'homologie basée sur la structure cristalline de Prp43 permet de révéler une homologie structurale du domaine HA2 avec les domaines caractéristiques WH et ratchet des hélicases de la famille Ski2 (Walbott *et al.*, 2010). Ce faisant, DHX36 possède un domaine C-terminale caractéristique des hélicases *DEAH/RHA* qui est distinct des hélicases DEAH virales.

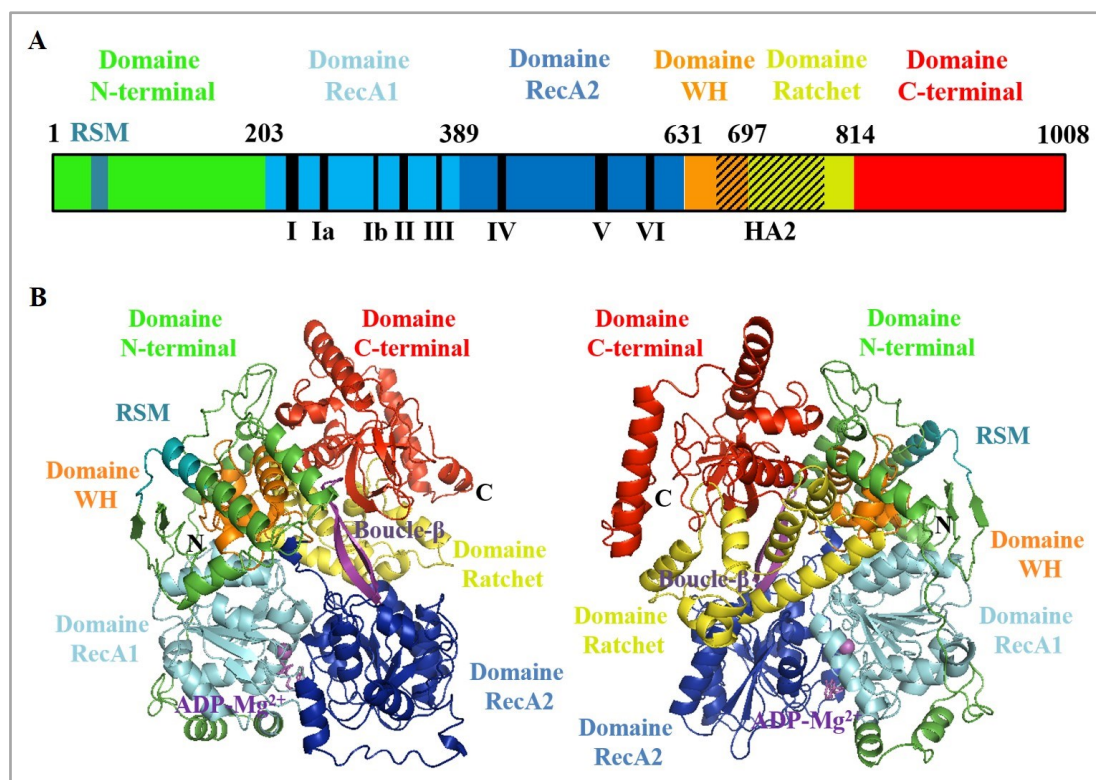


Figure 19. Modèle d'homologie de la protéine DHX36.

A Représentation schématique de l'organisation des domaines de DHX36. Le nom des domaines structuraux et leurs limites sont indiqués au-dessus, les motifs conservés (chiffres romains) et le domaine HA2 (*Helicase associated domain*) se retrouvent en dessous. Le motif spécifique à DHX36 : RSM (*RHAU-specific motif*) localisé dans le domaine N-terminale est indiqué en turquoise. **B** Représentation du modèle de DHX36 en complexe avec une molécule d'ADP et l'ion Mg²⁺ selon deux orientations différentes de 180°. Ce modèle a été généré à partir du logiciel Pymol en utilisant la structure cristalline de l'hélicase Prp43p de *S. cerevisiae* en complexe avec de l'ADP et du magnésium comme matrice (code PDB : 2xau). Les six domaines structuraux de DHX36 sont colorés comme en A, le motif RSM est en turquoise et la boucle-β au niveau du domaine RecA2 est en mauve. La molécule d'ADP est représentée en bâtonnets magenta alors que l'ion de magnésium est montré par une sphère magenta.

3.3.1 Interactions entre la protéine DHX36 et l'ATP

Tel que montré à travers les travaux de la découverte de la structure cristalline de Prp43p, la molécule d'ADP se retrouve à l'intérieur des domaines RecA1 et RecA2, celle-ci est liée à DHX36 par l'entremise d'interactions avec les motifs conservés I, V et VI. (Walbott *et al.*, 2010). Le modèle structural révèle les interactions entre les résidus spécifiques de la protéine et le substrat nucléotidique (Figure 20).

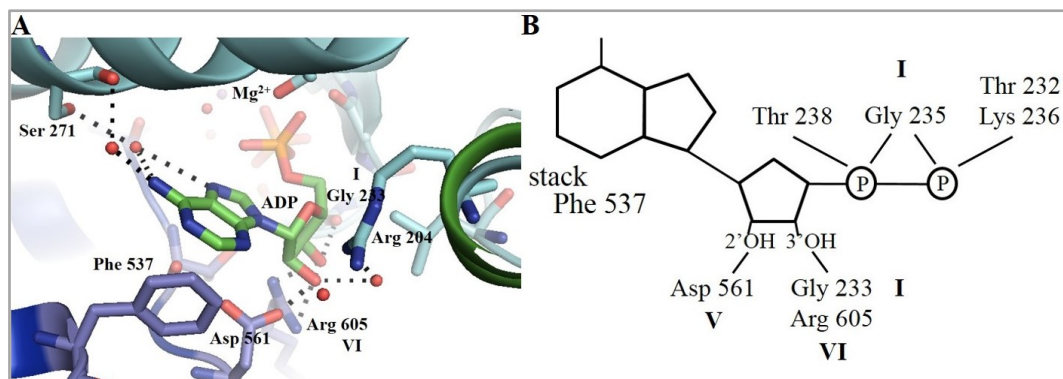


Figure 20. Site de liaison de l'ADP.

A Représentation 3D rapprochée du site de liaison de la base d'adénine entre la Ser-271 et Phe-537 du modèle de l'hélicase DHX36. Les résidus qui interagissent avec l'ADP (en bâtonnets verts) sont montrés sous forme de bâtonnets et étiquetés. Les domaines RecA1 et RecA2 sont colorés tels que dans la figure 19, l'ion de magnésium est représenté par une sphère magenta alors que les molécules d'eau sont représentées par des sphères rouges. **B** Représentation schématisée des interactions directes entre la molécule d'ADP et les résidus de DHX36. Les différents motifs sont inscrits en chiffres romains.

Tel que déjà démontré par les structures cristallines d'hélicases résolues jusqu'à présent, la boucle-P (motif I) de DHX36 est responsable de la liaison aux phosphates de la molécule d'ADP (Figure 20B). En effet, les phosphates α et β sont stabilisés par les interactions avec la chaîne latérale ou principale des résidus du motif I : Thr-232, Gly-235, Lys-236 et Thr-238 (non montrés dans la figure 20A). Les phosphates de l'ADP sont également coordonnés par des molécules d'eau qui interagissent avec les résidus Arg-204, Thr-237 et les résidus issus du motif V : Thr-556, Ser-557 et Thr-559. Quant au 3'OH du ribose, celui-ci interagit directement avec deux groupements amines de Arg-605 (motif VI) et avec le groupement carbonyle de la chaîne principale de Gly-233 (motif I). Le 2'OH du ribose est stabilisé par des interactions avec le groupement carboxylique de Asp-561 (motif V) et avec les groupements amines de Arg-204 par l'entremise d'une molécule d'eau. Le modèle de la protéine DHX36 et celui de l'hélicase Prp43 montre également que l'anneau d'ATP est emprisonné à l'intérieur des deux domaines RecA plutôt que d'être enfoui à l'intérieur du domaine RecA1 (Walbott *et al.*, 2010). En effet, la base adénine est empaquetée à l'intérieur par l'anneau de la phénylalanine F537 du domaine RecA2. Contrairement à l'hélicase DEAH NS3 virale, la base adénine de l'ADP est complètement exposée au solvant sans interaction d'empaquetage (Wu *et al.*, 2005; Luo *et al.*, 2008). Finalement, il y a également interactions

entre les positions N7 et N6 de la base d'adénine avec le groupement carbonyle de la chaîne principale et le groupement hydroxyle de la chaîne latérale du résidu Ser-271 par l'entremise de molécules d'eau. Les interactions entre la molécule d'ADP et les résidus de la protéine DHX36 modélisée sont constantes avec celles retrouvées parmi les hélicases de type DEAH (Walbott *et al.*, 2010). Nous supposons que les interactions révélées par le modèle entre la protéine et l'ADP sont similaires aux interactions entre la protéine et l'ATP. Évidemment, les observations de l'interaction entre la protéine et différents groupements des nucléotides ont été effectuées en l'absence d'ARN. Nécessairement, la présence d'ARN pourrait influencer ces interactions et ainsi les observations seraient légèrement différentes.

Il est possible de répertorier à travers la littérature un élément conservé parmi les hélicases de type DEAD qui serait responsable de l'utilisation spécifique de l'ATP comme source d'énergie pour résoudre des structures secondaires d'acides nucléiques. En effet, ce motif correspond à une glutamine conservée en amont du motif I conservé responsable de l'interaction avec le segment triphosphate des nucléotides. Tanner *et al.* ont nommé cet élément le motif Q (Tanner *et al.*, 2003). Celui-ci est notamment retrouvé dans l'hélicase eIF4A impliquée dans l'initiation de la traduction chez *S. cerevisiae* (Benz *et al.*, 1999) et dans l'hélicase Vasa, essentielle au développement embryonnaire de la drosophile (Sengoku *et al.*, 2006) (Figure 13). La structure cristalline de Vasa (PDB : 2DB3) montre que le motif Q forme des interactions ponts hydrogène avec les positions N6 et N7 de l'ATP (Sengoku *et al.*, 2006). Ainsi, les hélicases de type DEAD sont en mesure de discriminer l'ATP du GTP par l'entremise de ce motif conservé. La reconnaissance non spécifique de la base expliquerait pourquoi DHX36 et d'autres hélicases DEAH seraient considérées comme des NTPases (Warrener *et al.*, 1993) plutôt qu'être exclusivement des ATPases telles que les hélicases de type DEAD (Cordin *et al.*, 2006). Ainsi, il s'agirait d'un mode de liaison typiquement exclusif aux hélicases de la famille des DEAH virales et DEAH/RHA. Tanner *et al.* ont montré que le motif Q se situait précisément à 21 résidus en amont du motif I d'après la séquence consensus des hélicases de type DEAD. Cependant, le modèle structural de la protéine DHX36 révèle qu'il n'y a pas de glutamine en périphérie du site actif de liaison à l'ATP. De plus, parmi la séquence peptidique de DHX36 on ne retrouve pas de glutamine en amont à 21 résidus du motif I (Figure 21). En effet, l'alignement de séquence réalisé entre

différents représentants des hélicases de type DEAH et DEAD a souligné l'absence des résidus conservés du motif Q à travers les hélicases de type DEAH et notamment les orthologues de DHX36.

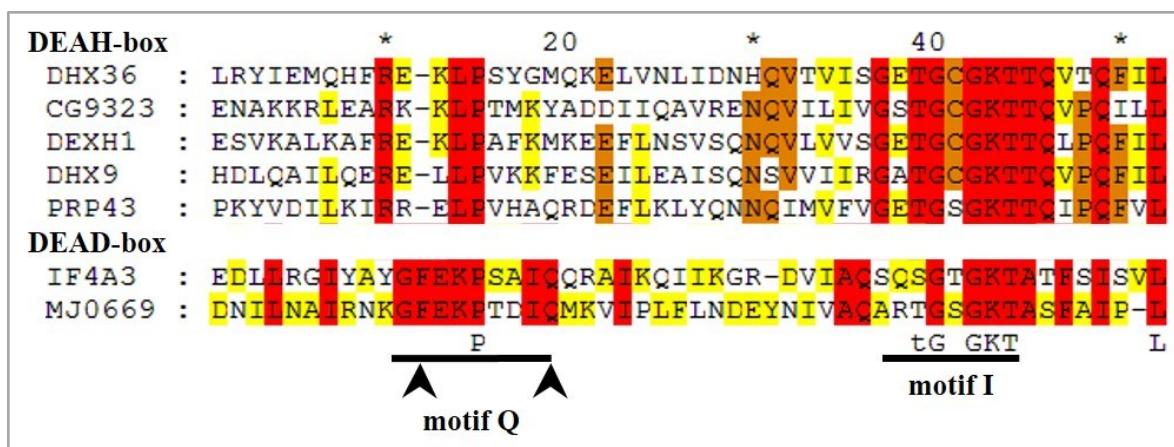


Figure 21. Résidus conservés dans la région du motif I.

Alignement de séquences du motif I et de la région en amont de différentes hélicases de type DEAH et DEAD. L'alignement de séquences multiple a été obtenu à partir du logiciel MAFFT (version 7) (Kato et Toh, 2008). L'analyse de similarité a été obtenue à partir du logiciel GeneDoc (version 2.7) en utilisant la matrice de notation BLOSUM62. La similarité est indiquée parmi les sous-groupes en rouge (100%), en orange (99-80%) et en jaune (79-60%). Les acides aminés identiques parmi toutes les séquences sont indiqués sous l'alignement. Les motifs I et Q sont indiqués. Les flèches indiquent les résidus conservés du motif Q exclusif aux hélicases de type DEAD (une glutamine et une phénylalanine 7 résidus en amont). Les séquences protéiques alignées et leurs (numéros d'accès UniProtKB) sont les suivantes: DHX36, hélicase d'ARN humaine ATP-dépendante (Q9H2U1-1), CG2923, orthologue de DHX36 retrouvé chez la drosophile (Q8SWT2-1), DEXH1, orthologue de DHX36 retrouvé chez *A. thaliana* (F4ILR7-1), DHX9, hélicase d'ARN humaine A ATP-dépendante (Q08211-1), Prp43, facteur d'épissage et hélicase d'ARN ATP-dépendante chez la levure à bourgeon (P53131-1), eIF4A3, hélicase de type DEAD et facteur d'initiation 4A-III humain (P38919-1), MJ0669, hélicase de type DEAD chez l'archaébactérie *Methanococcus jannaschii* (Q58083-1).

Somme toute, nous observons que le motif Q est absent de la séquence protéique de DHX36 et qu'il s'agirait plutôt de Ser-271 qui interagit avec les positions N6 et N7 de l'ATP (Figure 20A). L'absence du motif Q permettrait à l'enzyme d'utiliser différents nucléotides afin de fournir l'énergie nécessaire pour catalyser la réaction du dépliement des G4.

3.4 Analogues d'ATP afin d'identifier la spécificité nucléotidique

Afin de confirmer expérimentalement le modèle d'homologie et d'en apprendre davantage sur les groupements fonctionnels de l'ATP qui permettent à l'hélicase DHX36 de discriminer celui-ci des autres nucléotides, nous avons utilisé une quantité étendue d'analogues d'ATP. Les analogues utilisés possèdent différentes modifications au niveau de la base azotée, du ribose et de la partie triphosphate (Figure 22A). Les modifications sont variées, elles peuvent correspondre à de simples modifications d'un seul groupement fonctionnel à de plus complexes telles que des modifications de plusieurs groupements fonctionnels simultanément pour un analogue donné. L'effet de la substitution de l'ATP au profit d'un des analogues de l'ATP sur l'efficacité de la réaction révélera les groupements fonctionnels qui sont nécessaires pour le dépliement optimal de la réaction. L'activité enzymatique relative a été évaluée selon la quantité de produit pour chacun des analogues substitués et listée dans le tableau 3. La technique utilisée afin de révéler la quantité de produit de réaction obtenu est décrite dans la section matériel et méthodes; la figure 18B illustre le résultat typiquement obtenu. L'activité enzymatique relative des variétés d'analogues s'étend de 0 à 1,02.

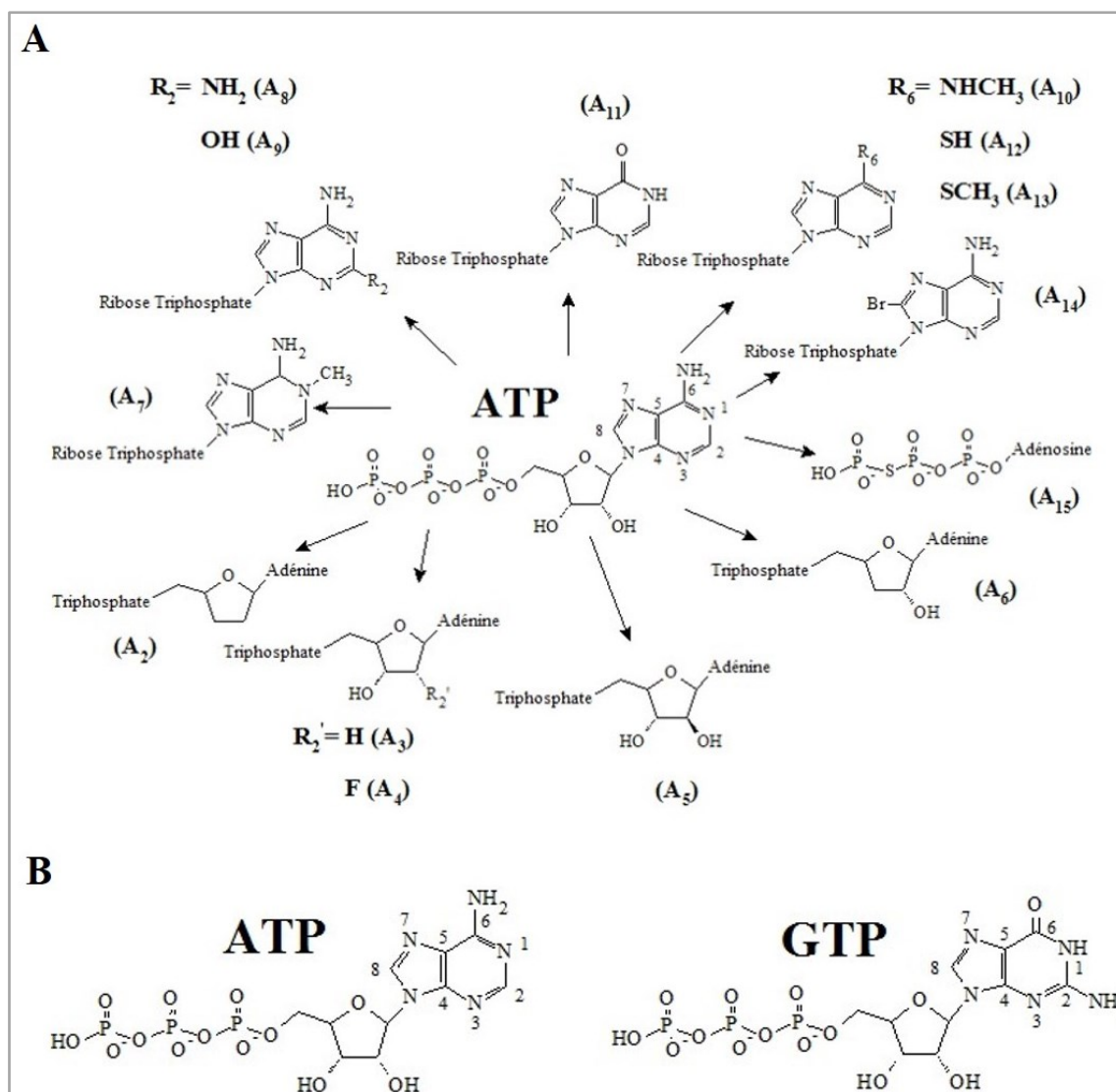


Figure 22. Analogues d'ATP utilisés afin d'analyser le site actif de l'hélicase DHX36.

A Structure des analogues nucléotidiques d'ATP utilisés. Les analogues d'ATP possèdent différentes modifications au niveau du ribose, de la base azotée et du triphosphate. **B** Différences atomiques au niveau de la base azotée entre le GTP et l'ATP, le groupement amine en N6 est remplacé par un groupement oxo dans le GTP, de plus, il y a ajout d'un groupement amine en position 2 dans la molécule de GTP.

3.5 Groupements fonctionnels nécessaires

Nous avons analysé plus de 14 analogues différents et nous avons évalué l'efficacité enzymatique lorsque ceux-ci substituaient l'ATP. La capacité de DHX36 à effectuer la catalyse de la réaction de dépliement du complexe G4 en présence d'analogues d'ATP reflète la flexibilité du site de liaison de l'ATP de DHX36 permettant d'accommoder différents

analogues d'ATP. En effet, les analogues portant des modifications chimiques sur la base azotée de l'ATP (Figure 22A) permettent plus précisément de révéler la flexibilité du site actif ATPase de l'enzyme. Le groupement amine en N6 de l'ATP semble être un groupement prépondérant pour la reconnaissance, la liaison et l'hydrolyse de celui-ci par l'enzyme puisqu'il possède à la fois les capacités de donneur de liaisons hydrogène que d'accepteur. D'autant plus que ce groupement est substitué par un groupement oxo en position N6 (accepteur d'hydrogène) au niveau de la base azotée du GTP, suggérant que la capacité donneur d'hydrogène contribue à la sélectivité de l'enzyme pour l'ATP au détriment du GTP. L'activité enzymatique du dépliement de l'ARN G4 de quatre analogues d'ATP portant des modifications chimiques au niveau de la position N6 (A₁₀, A₁₁, A₁₂, A₁₃) permet de révéler l'incidence de ce groupement sur la sélectivité nucléotidique. Nous avons observé que l'absence d'un donneur de liaisons hydrogène à cette position réduisait la capacité de l'analogue A₁₁ (ITP) à favoriser la réaction hélicase (Tableau 3). Par contre, la présence d'un groupement modérément bon donneur de liaison hydrogène (Zhou *et al.*, 2009) à cette position, soit un groupement thiol (A₁₂, 6-mercaptopurine-RTP), ne diminuerait pas sévèrement l'activité de l'enzyme. Ainsi, nous supposons que la présence d'un groupement fonctionnel qui possède la capacité donneur de liaisons hydrogène en position 6 de l'ATP est critique pour le dépliement adéquat de la réaction. De plus, les analogues A₁₀ (N6-methyl-ATP) et A₁₃ (6-methyl-thio-ITP) ne semblent pas affecter sérieusement l'activité enzymatique. En effet, les groupements N6-methyl et 6-methyl-thio retrouvés sur les analogues A₁₀ et A₁₃ respectivement possèdent tous deux les propriétés accepteur et donneur de liaisons hydrogène. De plus, ces groupements ont la caractéristique d'être plus volumineux que le groupement amine retrouvé dans l'ATP, le fait que l'activité ne soit pas grandement affectée par la présence de ceux-ci suppose que le site de liaison soit suffisamment flexible pour accommoder les analogues d'ATP portant ces modifications chimiques. Si on s'attarde au modèle produit de la protéine DHX36, on observe l'interaction entre la position N6 de la base d'adénine avec le groupement carbonyle de la chaîne principale et le groupement hydroxyle de la chaîne latérale de Ser-271 par l'entremise de molécules d'eau (Figure 20A). Le fait que le résidu Ser-271 de la protéine, qui interagit avec le groupement amine en position 6 de l'ATP, soit négativement polaire (accepteur de liaisons hydrogène) reflète la nécessité de la propriété donneur de liaisons hydrogène du groupement

en position 6 de l'ATP. De plus, la figure 20A montre qu'un groupement volumineux en position 6 des analogues d'ATP (A_{10} et A_{13}) n'affecterait pas l'alignement de l'analogue dans le site de liaison de l'ATP de la protéine puisque cette espace est occupée par le solvant. Ainsi, les résultats obtenus avec les analogues possédant des modifications au niveau de la position 6 de l'ATP révèlent que la propriété donneur de liaisons hydrogène en position 6 est requise pour la catalyse de la réaction du dépliement des G4. De plus, ces résultats permettraient de soutenir le modèle obtenu en préconisant un résidu négativement polaire qui interagit avec la position 6 de l'ATP, ainsi la sérine-271 pourrait être le résidu permettant cette interaction.

L'addition d'un substituant halogéné possédant un grand rayon atomique en position C8 de la base d'adénine (A_{14} , 8-bromo ATP) avait un impact significatif sur l'activité enzymatique. En effet, l'activité enzymatique a été complètement abolie en présence d'un groupement d'une taille volumineuse sur l'analogue d'ATP. Cela expliquerait que l'ATP requiert son inclusion à l'intérieur d'une pochette de liaison dédiée à l'interaction entre la protéine et l'ATP et l'hydrolyse de celui-ci pour effectuer la réaction enzymatique. En effet, la figure 19B du modèle de l'hélicase DHX36 obtenu montre que l'ADP se retrouve empaqueter entre les deux domaines RecA1 et RecA2. Ainsi, l'abolition de l'activité enzymatique en présence de l'analogue possédant un substituant halogéné à grand rayon atomique permet de confirmer que le site de liaison de l'ATP de la protéine DHX36 est une pochette formée entre les deux domaines RecA1 et RecA2. Ainsi, l'intégration de l'ATP à l'intérieur de cette pochette est nécessaire pour l'hydrolyse de celui-ci.

Ensuite, nous avons observé que l'addition d'un groupement amine en position C2 de l'ATP (A_8 , 2-amino-ATP) n'avait pas d'impact significatif sur l'activité de l'enzyme. Par contre, l'ajout d'un groupement hydroxyle à la même position (A_9 , 2-hydroxy ATP) affectait grandement l'activité enzymatique. Cela suggère que le groupement OH établi des interactions défavorables avec des résidus à proximité de la position C2. En regardant le modèle structural de la protéine DHX36, il est possible d'observer que la Phe-537 est l'acide aminé se trouvant en périphérie de la position C2. Le rôle de la phénylalanine au niveau de l'interaction entre la protéine et l'ATP est l'empaquetage de l'anneau de l'ATP dans le site

actif. Ainsi, l'ajout d'un groupement électronégatif en position C2 induirait la répulsion électronique de la chaîne latérale de la phénylalanine qui possède un caractère hydrophobe dépourvue de charge. La phénylalanine en position 537 semble importante au niveau de l'alignement adéquat de l'ATP à l'intérieur de la pochette du site actif ATPase de la protéine DHX36.

Finalement, l'ajout d'un groupement méthyl en position N1 de la base azotée de l'ATP annule l'activité enzymatique de la protéine. L'analogue utilisé portant ces modifications fût le N1-méthyl ATP (A₇) et le résultat de l'activité enzymatique subséquente à son utilisation était pratiquement nul (Tableau 3). L'hypothèse soulevée par ce résultat est que l'ajout du groupement méthyl à cet endroit provoquerait un encombrement stérique au niveau des résidus du site actif limitant ainsi l'hydrolyse de celui-ci. En regardant le modèle structural (Figure 20A) on retrouve encore une fois la Phe-537 à proximité de la position N1 de l'adénine. Ainsi, l'encombrement stérique provoqué par la présence du groupement méthyl limiterait l'efficacité de la phénylalanine à aligner adéquatement l'ATP à l'intérieur du site actif et permettre l'hydrolyse de l'ATP adéquatement.

3.6 Spécificité nucléotidique des purines ATP/GTP

Les essais de dépliement des G4 ont été évalués en présence de nucléotides à pyrimidines (CTP et UTP), il a été révélé que l'enzyme n'est pas en mesure d'utiliser ces nucléotides afin d'exécuter la réaction enzymatique (résultats non présentés). Donc, DHX36 est en mesure de discriminer les nucléotides à purines des nucléotides à pyrimidines. Le tableau 3 montre que l'utilisation du GTP par l'enzyme est possible mais l'efficacité de l'activité enzymatique correspond à la moitié relativement à celle de l'ATP. Ainsi, l'enzyme ne discrimine pas totalement l'ATP du GTP. Par contre, il y a certaines modifications au niveau du GTP qui rendre défavorable l'utilisation de celui-ci à travers la réaction de dépliement des G4 d'ARN. Les distinctions chimiques au niveau de la base azoté entre l'ATP et le GTP sont montrées dans la figure 22B. La substitution du groupement amine en position 6 de l'ATP par un groupement oxo dans la molécule de GTP et l'ajout d'un groupement amine en position 2 sont les modifications qui distinguent le GTP de l'ATP. Les analogues 2-amino ATP (A₈) et l'ITP (A₁₁) sont porteurs de modifications similaires à celles du GTP. Dans un premier temps,

nous avons montré que la substitution du groupement amine par un groupement oxo en position 6 de l'ATP à travers l'analogue A₁₁ diminuait l'efficacité de la réaction. Ensuite, l'ajout du groupement amine en position 2 de l'ATP à travers l'analogue A₈ n'influence pas significativement l'activité enzymatique. Ainsi, le groupement oxo du GTP semblerait être l'élément responsable de la diminution de l'activité enzymatique sans toutefois permettre à l'enzyme de discriminer l'ATP du GTP. Cela serait concordant avec la littérature suggérant que l'absence du motif Q parmi les hélicases de type DEAH limiterait l'enzyme à discriminer l'ATP du GTP. Néanmoins, la présence de ce groupement oxo dans la molécule de GTP, de par sa nature physicochimique d'accepteur de liaisons hydrogène, défavoriserait l'utilisation du GTP par l'enzyme.

CHAPITRE 4 – DISCUSSION

Notre étude permet de soulever les caractéristiques importantes et fonctionnelles des interactions entre l'ATP et l'hélicase DHX36. Elle permet d'expliquer d'un point de vue moléculaire la préférence de la protéine pour la molécule d'ATP afin d'effectuer le dépliement de son substrat l'ARN G4. En effet nous avons réussi à déterminer (i) les groupements fonctionnels de l'ATP qui permettaient sa liaison et son hydrolyse par l'enzyme (ii) les résidus de DHX36 potentiellement impliqués dans la liaison à l'ATP (iii) déterminé l'absence du motif Q permettant à DHX36 d'utiliser également le GTP comme substrat à moindre efficacité que l'ATP.

La combinaison des analogues de nucléotides et le modèle structural permet de révéler les caractéristiques structurales et fonctionnelles de l'interaction entre l'hélicase humaine DHX36 et l'ATP. En effet, en utilisant une panoplie d'analogues d'ATP nous avons réussi à montrer l'importance de la propriété donneur de liaisons hydrogène en position C6 du substrat nucléotidique. Cela a été révélé par la diminution radicale de l'activité enzymatique lors de la substitution de ce groupement par un groupement strictement donneur de liaisons hydrogène, un groupement oxo, avec l'ITP (A₁₁). L'importance fonctionnelle de ce groupement reflète la raison pour laquelle le GTP est moins efficace que l'ATP afin de catalyser la réaction hélicase de DHX36 sur l'ARN G4. Nous avons également révélé que le site actif ATPase de l'enzyme DHX36 représente une pochette de liaison à l'intérieur duquel se déroule la réaction enzymatique d'hydrolyse du substrat nucléotidique. En effet, l'utilisation du 6-bromo ATP, soit l'analogue d'ATP portant un groupement halogéné à grand rayon atomique, révélait qu'il ne pouvait être hydrolysé et ainsi permettre la catalyse du dépliement de l'ARN G4. De plus, nous montrons également que la phénylalanine F537 serait impliquée à l'alignement adéquat du nucléotide à l'intérieur du site actif de DHX36. Finalement, nous avons déterminé que le motif Q, responsable de l'interaction spécifique avec les positions N6 et N7 de l'ATP, était absent de la séquence protéique de DHX36. Ce faisant, l'hélicase DHX36 n'est pas en mesure de discriminer totalement l'ATP du GTP. Ce résultat est cohérent avec les résultats obtenus avec les hélicases de la famille des DEAH virales et DEAH/RHA (Walbott *et al.*, 2010).

Nous avons orienté nos analyses et notre recherche principalement autour des groupements fonctionnels des purines nécessaires à l'interaction et à l'efficacité catalytique de l'hélicase DHX36. Néanmoins, plusieurs analogues portant des modifications au niveau du ribose du substrat nucléotidique ont été utilisés. En effet, le modèle de DHX36 révèle que l'hélicase interagit avec les groupements OH du ribose par l'entremise des résidus Arg-605 et Gly-233 concernant le groupement 3'OH et les résidus Asp-561 Arg-204 pour le groupement 2'OH. Conséquemment avec nos observations à partir du modèle de DHX36, l'utilisation d'analogues portant des modifications au niveau du 3'OH et 2'OH du ribose montre une différence de l'efficacité de l'activité enzymatique. En effet, le retrait, le remplacement ou bien la modification d'isomérisation de l'un ou l'autre des groupements hydroxyles du ribose (A₃, A₄, A₅ et A₆) inhibe l'activité enzymatique de DHX36. Cependant, l'absence des deux groupements hydroxyles du ribose (A₂) n'affecte pas l'activité enzymatique de la protéine. Ainsi, l'absence simultanée des deux groupements n'influencerait pas l'activité de l'hélicase alors que l'absence, le remplacement ou la modification d'isomérisation d'un des deux groupements hydroxyles inhibe l'activité enzymatique. Ainsi, les deux groupements seraient dépendants afin de favoriser l'activité enzymatique de l'hélicase. Puisqu'il y a des résidus spécifiques qui interagissent avec ces groupements, la présence des deux groupements serait nécessaire au bon positionnement du nucléotide à l'intérieur du site actif. De plus, l'absence simultanée des deux groupements permettrait l'hydrolyse du nucléotide sans spécificité. Ainsi, l'hélicase DHX36 interagit spécifiquement avec les groupements hydroxyles du ribose de l'ATP afin de reconnaître la molécule, favoriser son positionnement à l'intérieur du site actif permettant l'hydrolyse du nucléotide.

Par ailleurs, nous n'avons pas réussi à déterminer la thermodynamique de liaison entre DHX36 et son substrat l'ARN G4. La technique privilégiée afin d'évaluer celle-ci fût la spectroscopie à fluorescence (résultats non montrés). Nonobstant le fait que la spectroscopie à fluorescence puisse être une technique permettant d'évaluer une panoplie de paramètres d'interaction entre une protéine et son ligand, celle-ci possède tout de même certaines limitations. En effet, la spectroscopie à fluorescence permet de détecter les changements du microenvironnement des résidus de tryptophanes retrouvés parmi la séquence de la protéine analysée. Ces changements de microenvironnement sont provoqués notamment, dans le

cadre de notre étude, par le changement conformationnel de la protéine suite à l'interaction avec son ligand. Il se peut qu'avec une protéine aussi volumineuse, l'interaction avec l'ARN G4 n'induit suffisamment pas de changement afin de modifier l'environnement des tryptophanes et ainsi révéler quantitativement la liaison. Ainsi, nous nous sommes abstenus d'effectuer des essais de liaison à l'aide du *RNA electromobility shift assay* ou REMSA puisque cette technique est trop limitante concernant les conditions au moment de l'interaction entre la protéine et son ligand. Plusieurs études ont déjà révélé l'affinité de liaison entre la protéine et l'ARN G4 en utilisant le REMSA (Vaugh *et al.*, 2005; Creacy, 2008). Ainsi, il serait plutôt intéressant d'évaluer la thermodynamique de liaison entre DHX36 et l'ARN G4 en utilisant l'ITC (*Isothermal titration calorimetry*). L'utilisation de l'ITC permettrait également de varier les conditions lors de l'interaction entre DHX36 et l'ARN G4 (Feig, 2009). De plus, il aurait été intéressant d'effectuer des mutagénèses dirigées afin de confirmer les résultats obtenus à l'aide des analogues d'ATP et du modèle produit. En effet, en utilisant la mutagénèse dirigée, nous aurions été en mesure de déterminer notamment si la capacité donneur de liaisons hydrogène de la sérine en position 271 permet l'interaction entre l'enzyme et son substrat nucléotidique. Ainsi, en effectuant ces mutations ponctuelles sur chacun des résidus susceptibles d'interagir avec le substrat nucléotidique, il aurait été possible d'affirmer qu'ils sont essentiels à l'interaction entre la protéine et son ligand. Finalement, le modèle produit de la protéine DHX36 n'est évidemment pas parfait, il serait intéressant d'obtenir, ultérieurement, la structure 3D de la protéine DHX36 en utilisant la technique du RMN (Résonance magnétique nucléaire). En effet, cette technique permettrait d'obtenir la structure 3D de la protéine telle qu'elle est plutôt qu'employer un modèle prédictif de la protéine.

Somme toute, nous avons vu que l'hélicase DHX36 était dépendante du substrat nucléotidique, plus spécifiquement les nucléotides à purines, afin de dérouler les structures G4 dans l'ARN. De plus, il a été montré que les G4 d'ARN à l'intérieur d'une molécule d'ARNm avaient une gamme de rôles. Considérant que l'hélicase DHX36 puisse déstabiliser ces structures G4 à l'intérieur de l'ARNm, conséquemment la fonction de ceux-ci serait inhibée. Ainsi, l'activité de DHX36 pourrait potentiellement favoriser la traduction de l'ARNm de proto-oncogènes, notamment *NRAS* (Kumari *et al.*, 2007). La modulation de

l'activité de l'hélicase DHX36 *in vivo* serait un défi intéressant afin de contrôler les répercussions en aval issues de la déstabilisation des G4 d'ARN. En améliorant nos connaissances vis-à-vis la protéine DHX36, nous serions potentiellement en mesure de moduler son activité *in vivo*.

LISTE DES RÉFÉRENCES

- Agarwala, P., Pandey, S. & Maiti, S. (2015) The tale of RNA G-quadruplex. *Org. & Biomol. Chem.*, **13**, 5570–5585.
- Agarwala, P., Pandey, S., Mapa, K. & Maiti, S. (2013) The G-Quadruplex Augments Translation in the 5' Untranslated Region of Transforming Growth Factor β 2. *Biochemistry*, **52**, 1528–1538.
- Anantharaman, V., Koonin, E. V. & Aravind, L. (2002) Comparative genomics and evolution of proteins involved in RNA metabolism. *Nucleic Acids Res.* **30**, 1427–1464.
- Andersen, C. B. F., Ballut, L., Johansen, J.S., Chamieh, H., Nielsen, K.H., Oliveira, C.L.P., Pedersen, J.S., Séraphin, B., Hir, H.L., Andersen, G.R. (2006) Structure of the Exon Junction Core Complex with a Trapped DEAD-Box ATPase Bound to RNA. *Science*, **313**, 1968–1972.
- Armanios, M. (2009) Syndromes of Telomere Shortening. *Annu. Rev. Genomic- Hum. Genet.*, **10**, 45–61.
- Arora, A., Balasubramanian, C., Kumar, N., Agrawal, S., Ojha, R.P., Maiti, S. (2008) Binding of berberine to human telomeric quadruplex – spectroscopic, calorimetric and molecular modeling studies. *FEBS J.*, **275**, 3971–3983.
- Arora, A. & Maiti, S. (2009) Differential Biophysical Behavior of Human Telomeric RNA and DNA Quadruplex. *J. Phys. Chem. B*, **113**, 10515–10520.
- Arora, A. & Suess, B. (2011) An RNA G-quadruplex in the 3' UTR of the proto-oncogene PIM1 represses translation. *RNA Biol.*, **8**, 802–805.
- Ashton, T. M. & Hickson, I. D. (2010) Yeast as a model system to study RecQ helicase function. *DNA Repair*, **9**, 303–314.
- Beaudoin, J.-D. & Perreault, J.-P. (2013) Exploring mRNA 3'-UTR G-quadruplexes: evidence of roles in both alternative polyadenylation and mRNA shortening. *Nucleic Acids Res.*, **41**, 5898–5911.
- Benz, J., Trachsel, H. & Baumann, U. (1999) Crystal structure of the ATPase domain of translation initiation factor 4A from *Saccharomyces cerevisiae*--the prototype of the DEAD box protein family. *Structure*, **7**, 671–679.
- Bochman, M. L., Paeschke, K. & Zakian, V. A. (2012) DNA secondary structures: stability and function of G-quadruplex structures. *Nat. Rev. Genet.*, **13**, 770–780.
- Bochman, M. L., Sabouri, N. & Zakian, V. A. (2010) Unwinding the functions of the Pif1 family helicases. *DNA Repair*, **9**, 237–249.
- Bono, F., Conti, E., Ebert, J. & Lorentzen, E. (2006) The Crystal Structure of the Exon Junction Complex Reveals How It Maintains a Stable Grip on mRNA. *Cell*, **126**, 713–725.
- Booy, E. P., McRae, E. K. S. & McKenna, S. A. (2015) Biochemical Characterization of G4 Quadruplex Telomerase RNA Unwinding by the RNA Helicase RHAU. Marc Boudvillain (ed.), *RNA Remodeling Proteins: Methods and Protocols, Methods in Molecular Biology*, Springer Science & Business Media New York, **1259**, 125–135.
- Booy, E. P., Meier, M., Okun, N., Novakowski, S.K., Xiong, S., Stetefeld, J., McKenna, S.A. (2012) The RNA helicase RHAU (DHX36) unwinds a G4-quadruplex in human telomerase RNA and promotes the formation of the P1 helix template boundary. *Nucleic Acids Res.*, **40**, 4110–4124.
- Bougie, I. & Bisailon, M. (2003) Initial binding of the broad spectrum antiviral nucleoside ribavirin to the hepatitis C virus RNA polymerase. *J. Biol. Chem.*, **278**, 52471–52478.

- Bougie, I., Bisailon, M. & Picard-Jean, F. (2007) Characterization of the DNA- and dNTP-binding activities of the human cytomegalovirus DNA polymerase catalytic subunit UL54. *Biochem. J.*, **407**, 331–341.
- Bougie, I., Charpentier, S. & Bisailon, M. (2003) Characterization of the metal ion binding properties of the hepatitis C virus RNA polymerase. *J. Biol. Chem.*, **278**, 3868–3875.
- Bougie, I., Parent, A. & Bisailon, M. (2004) Thermodynamics of ligand binding by the yeast mRNA-capping enzyme reveals different modes of binding. *Biochem. J.*, **384**, 411–420.
- Brosh, R. M. & Bohr, V. A. (2007) Human premature aging, DNA repair and RecQ helicases. *Nucleic Acids Res.*, **35**, 7527–7544.
- Budd, M. E., Reis, C. C., Smith, S., Myung, K. & Campbell, J. L. (2006) Evidence suggesting that Pif1 helicase functions in DNA replication with the Dna2 helicase/nuclease and DNA polymerase delta. *Mol. Cell. Biol.*, **26**, 2490–2500.
- Kumari, S., Bugaut, A., Huppert, J. L. & Balasubramanian, S. (2007) An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation. *Nat. Chem. Biol.*, **3**.
- Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K. & Neidle, S. (2006) Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.*, **34**, 5402–5415.
- Byrd, A. K. & Raney, K. D. (2012) Superfamily 2 helicases. *Front. Biosci.*, **17**, 2070–2088.
- Büttner, K., Nehring, S. & Hopfner, K.-P. (2007) Structural basis for DNA duplex separation by a superfamily-2 helicase. *Nat. Struct. & Mol. Biol.*, **14**, 647–652.
- Chen, M. C., Murat, P., Abecassis, K., Ferré-D'Amaré, A. R. & Balasubramanian, S. (2015) Insights into the mechanism of a G-quadruplex-unwinding DEAH-box helicase. *Nucleic Acids Res.*, **43**, 2223–2231.
- Cheng, X., Dunaway, S. & Ivessa, A. S. (2007) The role of Pif1p, a DNA helicase in *Saccharomyces cerevisiae*, in maintaining mitochondrial DNA. *Mitochondrion*, **7**, 211–222.
- Christiansen, J., Kofod, M. & Nielsen, F. C. (1994) A guanosine quadruplex and two stable hairpins flank a major cleavage site in insulin-like growth factor II mRNA. *Nucleic Acids Res.*, **22**, 5709–5716.
- Christiansen, J., Nielsen, J., Østergaard, L. & Nielsen, F. C. (1995) Growth-dependent translation of IGF-II mRNA by a rapamycin-sensitive pathway. *Nature*, **377**, 358–362.
- Chu, W. K. & Hickson, I. D. (2009) RecQ helicases: multifunctional genome caretakers. *Nat. Rev. Cancer*, **9**, 644–654.
- Collie, G. W., Haider, S. M., Neidle, S. & Parkinson, G. N. (2010) A crystallographic and modelling study of a human telomeric RNA (TERRA) quadruplex. *Nucleic Acids Res.*, **38**, 5569–5580.
- Cordin, O., Banroques, J., Tanner, N. K. & Linder, P. (2006) The DEAD-box protein family of RNA helicases. *Gene*, **367**, 17–37.
- Crawford, D. C., Acuña, J. M. & Sherman, S. L. (2001) FMR1 and the fragile X syndrome: Human genome epidemiology review. *Genetics in Medicine*, **3**, 359–371.
- Creacy, S. D., Routh, E.D., Iwamoto, F., Nagamine, Y., Akman, S.A., Vaughn, J.P. (2008) G4 resolvase 1 binds both DNA and RNA tetramolecular quadruplex with high affinity and is the major source of tetramolecular quadruplex G4-DNA and G4-RNA resolving activity in HeLa cell lysates. *J. Biol. Chem.*, **283**, 34626–34634.
- Decorsière, A., Cayrel, A., Vagner, S. & Millevoi, S. (2011) Essential role for the interaction between hnRNP H/F and a G quadruplex in maintaining p53 pre-mRNA 3'-end processing and function during DNA damage. *Genes & Dev.*, **25**, 220–225.

- Del Campo, M. & Lambowitz, A. M. (2009) Structure of the Yeast DEAD box protein Mss116p reveals two wedges that crimp RNA. *Mol. Cell*, **35**, 598–609.
- Didiot, M.-C., Tian, Z., Schaeffer, C., Subramanian, M., Mandel, J., Moine, H. (2008) The G-quartet containing FMRP binding site in FMR1 mRNA is a potent exonic splicing enhancer. *Nucleic Acids Res.*, **36**, 4902–4912.
- Diges, C. M. & Uhlenbeck, O. C. (2001) Escherichia coli DbpA is an RNA helicase that requires hairpin 92 of 23S rRNA. *EMBO J.*, **20**, 5503–5512.
- Duxin, J. P., Dao, B., Martinsson, P., Rajala, N., Guittat, L., Campbell, J.L., Spelbrink, J.N., Stewart, S.A. (2009) Human Dna2 is a nuclear and mitochondrial DNA maintenance protein. *Mol. Cell. Biol.*, **29**, 4274–4282.
- Eftink, M. R. & Ghiron, C. A. (1976) Exposure of tryptophanyl residues in proteins. Quantitative determination by fluorescence quenching studies. *Biochemistry*, **15**, 672–680.
- Endoh, T., Kawasaki, Y. & Sugimoto, N. (2013) Stability of RNA quadruplex in open reading frame determines proteolysis of human estrogen receptor α . *Nucleic Acids Res.*, **41**, 6222–6231.
- Fairman-Williams, M. E., Guenther, U.-P. & Jankowsky, E. (2010) SF1 and SF2 helicases: family matters. *Curr. Opin. Struct. Biol.*, **20**, 313–324.
- Fan, J.-S., Cheng, Z., Zhang, J., Noble, C., Zhou, Z., Song, H., Yang, D. (2009) Solution and crystal structures of mRNA exporter Dbp5p and its interaction with nucleotides. *J. Mol. Biol.*, **388**, 1–10.
- Feig, A. L. (2006) Chapter 19 - Studying RNA–RNA and RNA–Protein interactions by Isothermal Titration Calorimetry. *Methods in Enzymology*, **468**.
- Flowers, S., Biswas, E. E. & Biswas, S. B. (2003) Conformational dynamics of DnaB helicase upon DNA and nucleotide binding: analysis by intrinsic tryptophan fluorescence quenching. *Biochemistry*, **42**, 1910–1921.
- Fry, M. & Loeb, L. A. (1994) The fragile X syndrome d(CGG)n nucleotide repeats form a stable tetrahelical structure. *Proc. Natl. Acad. Sci.*, **91**, 4950–4954.
- Fry, M. & Loeb, L. A. (1999) Human Werner Syndrome DNA Helicase Unwinds Tetrahelical Structures of the Fragile X Syndrome Repeat Sequence d(CGG) n. *J. Biol. Chem.*, **274**, 12797–12802.
- Garner, M. M. & Revzin, A. (1981) A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the Escherichia coli lactose operon regulatory system. *Nucleic Acids Res.*, **9**, 3047–3060.
- Gellert, M., Lipsett, M. & Davies, D. (1962) Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. United States Am.*, 2013–2018.
- Gessner, R. V., Frederick, C. A., Quigley, G. J., Rich, A. & Wang, A. H. (1989) The molecular structure of the left-handed Z-DNA double helix at 1.0-Å atomic resolution. Geometry, conformation, and ionic interactions of d(CGCGCG). *J. Biol. Chem.*, **264**, 7921–7935.
- Giri, B., Smaldino, P.J., Thys, R.G., Creacy, S.D., Routh, E.D., Hantgan, R.R., Lattmann, S., Nagamine, Y., Akman, S.A., Vaughn, J.P. (2011) G4 resolvase 1 tightly binds and unwinds unimolecular G4-DNA. *Nucleic Acids Res.*, **39**, 7161–7178.
- Gomez, D. Guédin, A., Mergny, J., Salles, B., Riou, J., Teulade-Fichou, M., Calsou, P.A. (2010) G-quadruplex structure within the 5'-UTR of TRF2 mRNA represses translation in human cells. *Nucleic Acids Res.*, **38**, 7187–7198.

- Gomez, D., Lemarteleur, T., Lacroix, L., Mailliet, P., Mergny, J., Riou, J. (2004) Telomerase downregulation induced by the G-quadruplex ligand 12459 in A549 cells is mediated by hTERT RNA alternative splicing. *Nucleic Acids Res.*, **32**, 371–379.
- Gorbalenya, A. E. & Koonin, E. V. (1993) Helicases: amino acid sequence comparisons and structure-function relationships. *Curr. Opin. Struct. Biol.*, **3**, 419–429.
- Greenberg, R. A., Chin, L., Femino, A., Lee, K. H., Gottlieb, G. J., Singer, R. H., Greider, C. W. & DePinho, R. A. (1999) Short dysfunctional telomeres impair tumorigenesis in the INK4a(Δ 2/3) cancer-prone mouse. *Cell*, **97**, 515–525.
- Hafsi, H., Ortiz-Cuaran, S., Ma, D., Olivier, M., Sagne, C., Bourdon, J., Hainaut, P., Mollereau, B., Dichtel-Danjoy, M., Hall, J., Marcel, V. (2011) Biological functions of p53 isoforms through evolution: lessons from animal and cellular models. *Cell Death & Differ.*, **18**, 1815–1824.
- Hall, M. C. & Matson, S. W. (1999) Helicase motifs: the engine that powers DNA unwinding. *Mol. Microbiol.*, **34**, 867–877.
- Hardwick, S. W. & Luisi, B. F. (2013) Rarely at rest: RNA helicases and their busy contributions to RNA degradation, regulation and quality control. *RNA Biol.*, **10**, 56–70.
- Harrington, C., Lan, Y. & Akman, S. A. (1997) The identification and characterization of a G4-DNA resolvase activity. *J. Biol. Chem.*, **272**, 24631–24636.
- Huang, W., Smaldino, P.J., Zhang, Q., Miller, L.D., Cao, P., Stadelman, K., Wan, M., Giri, B., Lei, M., Nagamine, Y., Vaughn, J.P., Akman, S.A., Sui, G. (2012) Yin Yang 1 contains G-quadruplex structures in its promoter and 5'-UTR and its expression is modulated by G4 resolvase 1. *Nucleic Acids Res.*, **40**, 1033–1049.
- Huber, M. D., Lee, D. C. & Maizels, N. (2001) G4 DNA unwinding by BLM and Sgs1p: substrate specificity and substrate-specific inhibition. *Nucleic Acids Res.*, **30**, 3954–3961.
- Huppert, J. L. (2008) Hunting G-quadruplexes. *Biochimie*, **90**, 1140–1148.
- Iwamoto, F., Stadler, M., Chalupnikova, K., Oakeley, E. & Nagamine, Y. (2008) Transcription-dependent nucleolar cap localization and possible nuclear function of DExH RNA helicase RHAU. *Exp. cell Res.*, **314**, 1378–1391.
- Jarmoskaite, I. & Russell, R. (2011) RNA helicase proteins as chaperones and remodelers. *Annu. Rev. Biochem.*, **83**, 697–725.
- Ji, X., Sun, H., Zhou, H., Xiang, J., Tang, Y., Zhao, C. (2011) Research progress of RNA quadruplex. *Nucleic acid Ther.*, **21**, 185–200.
- Joachimi, A., Benz, A. & Hartig, J. S. (2011) A comparison of DNA and RNA quadruplex structures and stabilities. *Bioorganic & Med. Chem.*, **17**, 6811–6815.
- Kang, Y.-H., Lee, C.-H. & Seo, Y.-S. (2008) Dna2 on the road to Okazaki fragment processing and genome stability in eukaryotes. *Crit. Rev. Biochem. Mol. Biol.*, **45**, 71–96.
- Katoh, K. & Toh, H. (2008) Improved accuracy of multiple ncRNA alignment by incorporating structural information into a MAFFT-based framework. *BMC Bioinformatics*, **9**.
- Keniry, M. A. (2000) Quadruplex structures in nucleic acids. *Biopolymers*, **56**, 123–146.
- Kossen, K., Karginov, F. V. & Uhlenbeck, O. C. (2002) The Carboxy-terminal Domain of the DExDH Protein YxiN is Sufficient to Confer Specificity for 23 S rRNA. *J. Mol. Biol.*, **324**, 625–636.
- Kozak, M. (1980) Influence of mRNA secondary structure on binding and migration of 40S ribosomal subunits. *Cell*, **19**, 79–90.

- Lai, J. C., Ponti, S., Pan, D., Kohler, H., Skoda, R.C., Matthias, P., Nagamine, Y. (2012) The DEAH-box helicase RHAU is an essential gene and critical for mouse hematopoiesis. *Blood*, **119**, 4291–4300.
- Lansdorp, P. M. (2009) Telomeres and disease. *EMBO J.*, **28**, 2532–2540.
- Lattmann, S., Giri, B., Vaughn, J. P., Akman, S. A. & Nagamine, Y. (2010) Role of the amino terminal RHAU-specific motif in the recognition and resolution of guanine quadruplex-RNA by the DEAH-box RNA helicase RHAU. *Nucleic Acids Res.*, **38**, 6219–6233.
- Lattmann, S., Stadler, M. B., Vaughn, J. P., Akman, S. A. & Nagamine, Y. (2011) The DEAH-box RNA helicase RHAU binds an intramolecular RNA G-quadruplex in TERC and associates with telomerase holoenzyme. *Nucleic Acids Res.*, **39**, 9390–9404.
- Leipe, D. D., Wolf, Y. I., Koonin, E. V. & Aravind, L. (2002) Classification and evolution of P-loop GTPases and related ATPases1. *J. Mol. Biol.*, **317**, 41–72.
- Linder, P. & Jankowsky, E. (2011) From unwinding to clamping - the DEAD box RNA helicase family. *Nat. Rev. Mol. cell Biol.*, **12**, 505–516.
- London, T. B. C. Barber, L.J., Mosedale, G., Kelly, G.P., Balasubramanian, S., Hickson, I.D., Boulton, S.J., Hiom, K. (2008) FANCI is a structure-specific DNA helicase associated with the maintenance of genomic G/C tracts. *J. Biol. Chem.*, **283**, 36132–36139.
- Luo, D., Xu, T., Watson, R.P., Scherer-Becker, D., Sampath, A., Jahnke, W., Yeong, S.S., Wang, C.H., Lim, S.P., Strongin, A., Vasudevan, S.G., Lescar, J. (2008) Insights into RNA unwinding and ATP hydrolysis by the flavivirus NS3 protein. *EMBO J.*, **27**, 3209–3219.
- Marcel, V. & Hainaut, P. (2011) p53 isoforms - A conspiracy to kidnap p53 tumor suppressor activity? *Cell. Mol. Life Sci.*, **66**, 391-406.
- Masuda-Sasa, T., Polaczek, P., Peng, X. P., Chen, L. & Campbell, J. L. (2008) Processing of G4 DNA by Dna2 helicase/nuclease and replication protein A (RPA) provides insights into the mechanism of Dna2/RPA substrate recognition. *J. Biol. Chem.*, **283**, 24359–24373.
- Lodish, H., Matsudaira, P., Berk, A., Zipursky, S.L., Darnell, J., Baltimore, D. (2000) *Molecular Cell Biology*. 4th edition. New York: W. H. Freeman.
- Meier, M., Patel, T.R., Booy, E.P., Marushchak, O., Okun, N., Deo, S., Howard, R., McEleney, K., Harding, S.E., Stetefeld, J., McKenna, S.A. (2013) Binding of G-quadruplexes to the N-terminal recognition domain of the RNA helicase associated with AU-rich element (RHAU). *J. Biol. Chem.*, **288**, 35014–35027.
- Milner-White, E. J., Pietras, Z. & Luisi, B. F. (2010) An ancient anion-binding structural module in RNA and DNA helicases. *Proteins: Struct. Funct. Bioinforma.*, **78**, 1900–1908.
- Mohaghegh, P., Karow, J. K., Brosh Jr, R. M., Bohr, V. A. & Hickson, I. D. (2001) The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases. *Nucleic Acids Res.*, **29**, 2843–2849.
- Mortimer, S. A., Kidwell, M. A. & Doudna, J. A. (2014) Insights into RNA structure and function from genome-wide studies. *Nat Rev Genet.*, **15**, 469–479.
- Ohno, M. & Shimura, Y. (1996) A human RNA helicase-like protein, HRH1, facilitates nuclear export of spliced mRNA by releasing the RNA from the spliceosome. *Genes & Dev.*, **10**, 997–1007.
- Paleček, E. (1991) Local Supercoil-Stabilized DNA Structure. *Crit. Rev. Biochem. Mol. Biol.*, **26**, 151–226.

- Pan, J. Y., Sanford, J. C. & Wessling-Resnick, M. (1995) Effect of guanine nucleotide binding on the intrinsic tryptophan fluorescence properties of Rab5. *J. Biol. Chem.*, **270**, 24204–24208.
- Patel, D. J., Phan, A. T. & Kuryavyi, V. (2007) Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res.*, **35**, 7429–7455.
- Pike, J. E., Burgers, P. M. J., Campbell, J. L. & Bambara, R. A. (2009) Pif1 helicase lengthens some Okazaki fragment flaps necessitating Dna2 nuclease/helicase action in the two-nuclease processing pathway. *J. Biol. Chem.*, **284**, 25170–25180.
- Pyle, A. M. (2008) Translocation and Unwinding Mechanisms of RNA and DNA Helicases. *Annu. Rev. Biophys.*, **37**, 317–336.
- Ribeyre, C., Lopes, J., Boulé, J., Piazza, A., Guédin, A., Zakian, V.A., Mergny, J., Nicolas, A. (2009) The yeast Pif1 helicase prevents genomic instability caused by G-quadruplex-forming CEB1 sequences in vivo. *PLoS Genet.*, **5**.
- Rogers, G. W., Richter, N. J. & Merrick, W. C. (1999) Biochemical and kinetic characterization of the RNA helicase activity of eukaryotic initiation factor 4A. *J. Biol. Chem.*, **274**, 12236–12244.
- Schneider, S. & Schwer, B. (2001) Functional Domains of the Yeast Splicing Factor Prp22p. *J. Biol. Chem.*, **276**, 21184–21191.
- Sengoku, T., Nureki, O., Nakamura, A., Kobayashi, S. & Yokoyama, S. (2006) Structural basis for RNA unwinding by the DEAD-box protein Drosophila Vasa. *Cell*, **125**, 287–300.
- Shahid, R., Bugaut, A. & Balasubramanian, S. (2010) The BCL-2 5' Untranslated Region Contains an RNA G-Quadruplex-Forming Motif That Modulates Protein Expression. *Biochemistry*, **49**, 8300–8306.
- Simonsson, T. (2001) G-Quadruplex DNA Structures Variations on a Theme. *Biol. Chem.*, **382**, 621–628.
- Singleton, M. R., Dillingham, M. S. & Wigley, D. B. (2007) Structure and Mechanism of Helicases and Nucleic Acid Translocases. *Annual Review of Biochemistry*, **76**, 23–50.
- Stewart, J. A., Campbell, J. L. & Bambara, R. A. (2009) Significance of the dissociation of Dna2 by flap endonuclease 1 to Okazaki fragment processing in *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **284**, 8283–8291.
- Subramanian, M., Rage, F., Tabet, R., Flatter, E., Mandel, J., Moine, H. (2011) G-quadruplex RNA structure as a signal for neurite mRNA targeting. *EMBO reports*, **12**, 697–704.
- Sun, H., Karow, J. K., Hickson, I. D. & Maizels, N. (1998) The Bloom's Syndrome Helicase Unwinds G4 DNA. *J. Biol. Chem.*, **273**, 27587–27592.
- Svozil, D., Kalina, J., Omelka, M. & Schneider, B. (2008) DNA conformations and their sequence preferences. *Nucleic Acids Res.*, **36**, 3690–3706.
- Tanner, N. K., Cordin, O., Banroques, J., Doère, M. & Linder, P. (2003) The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. *Mol. Cell*, **11**, 127–138.
- Tran, H., Schilling, M., Wirbelauer, C., Hess, D. & Nagamine, Y. (2003) Facilitation of mRNA deadenylation and decay by the exosome-bound, DExH protein RHAU. *Mol. Cell*, **13**, 101–111.
- Tsu, C. A., Kossen, K. & Uhlenbeck, O. C. (2001) The *Escherichia coli* DEAD protein DbpA recognizes a small RNA hairpin in 23S rRNA. *RNA*, **7**, 702–709.

- Usdin, K. & Woodford, K. J. (1995) CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis in vitro. *Nucleic Acids Res.*, **23**, 4202–4209.
- Vaughn, J. P., Creacy, S.D., Routh, E.D., Joyner-Butt, C., Jenkins, G.S., Pauli, S., Nagamine, Y., Akman, S.A. (2005) The DEXH Protein Product of the DHX36 Gene Is the Major Source of Tetramolecular Quadruplex G4-DNA Resolving Activity in HeLa Cell Lysates. *J. Biol. Chem.*, **280**.
- Walbott, H., Mouffok, S., Capeyrou, R., Lebaron, S., Humbert, O., van Tilbeurgh, H., Henry, Y., Leulliot, N. (2010) Prp43p contains a processive helicase structural architecture with a specific regulatory domain. *EMBO J.*, **29**, 2194–2204.
- Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.*, **1**, 945–951.
- Wan, Y., Kertesz, M., Spitale, R. C., Segal, E. & Chang, H. Y. (2011) Understanding the transcriptome through RNA structure. *Nat. Rev. Genet.*, **12**, 641–655.
- Wang, Y. & Guthrie, C. (1998) PRP16, a DEAH-box RNA helicase, is recruited to the spliceosome primarily via its nonconserved N-terminal domain. *RNA*, **4**, 1216–1229.
- Wanrooij, P. H., Gustafsson, C. M., Falkenberg, M., Simonsson, T. & Uhler, J. P. (2010) G-quadruplex structures in RNA stimulate mitochondrial transcription termination and primer formation. *Proc. Natl. Acad. Sci.*, **107**, 16072–16077.
- Warrener, P., Tamura, J. K. & Collett, M. S. (1993) RNA-stimulated NTPase activity associated with yellow fever virus NS3 protein expressed in bacteria. *J. Virol.*, **67**, 989–996.
- Watson, J. D. & Crick, F. H. C. (1953) Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature*, 737–738.
- Wu, J., Bera, A. K., Kuhn, R. J. & Smith, J. L. (2003) Structure of the Flavivirus helicase: implications for catalytic activity, protein interactions, and proteolytic processing. *J. Virol.*, **79**, 10268–10277.
- Wu, Y. & Brosh, R. M. (2010) G-quadruplex nucleic acids and human disease. *FEBS J.*, **277**, 3470–3488.
- Wu, Y., Shin-ya, K. & Brosh, R. M. (2008) FANCD1 Helicase Defective in Fanconi Anemia and Breast Cancer Unwinds G-Quadruplex DNA To Defend Genomic Stability. *Mol. Cell. Biol.*, **28**, 4116–4128.
- Wu, Y., Suhasini, A. N. & Brosh, R. M. (2009) Welcome the family of FANCD1-like helicases to the block of genome stability maintenance proteins. *Cell. Mol. life Sci.*, **66**, 1209–1222.
- Zhang, D.-H., Zhou, B., Huang, Y., Xu, L.-X. & Zhou, J.-Q. (2006) The human Pif1 helicase, a potential Escherichia coli RecD homologue, inhibits telomerase activity. *Nucleic Acids Res.*, **34**, 1393–1404.
- Zhang, Y., Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J. (2015) The I-TASSER Suite: protein structure and function prediction. *Nat. Methods*, **12**, 7–8.
- Zheng, L., Zhou, M., Guo, Z., Lu, H., Qian, L., Dai, H., Qiu, J., Yakubovskaya, E., Bogenhagen, D.F., Demple, B., Shen, B. (2008) Human DNA2 is a mitochondrial nuclease/helicase for efficient processing of DNA replication and repair intermediates. *Mol. Cell*, **32**, 325–336.
- Zhou, P., Tian, F., Lv, F. & Shang, Z. (2009) Geometric characteristics of hydrogen bonds involving sulfur atoms in proteins. *Proteins*, **76**, 151–163.

- Zhou, T. & Rosen, B. P. (1997) Tryptophan fluorescence reports nucleotide-induced conformational changes in a domain of the ArsA ATPase. *J. Biol. Chem.*, **272**, 19731–19737.
- Von Moeller, H., Basquin, C. & Conti, E. (2009) The mRNA export protein DBP5 binds RNA and the cytoplasmic nucleoporin NUP214 in a mutually exclusive manner. *Nat. Struct. & Mol. Biol.*, **16**, 247–254.