Etude du rôle d'un G-quadruplex d'ARN dans l'épissage alternatif du gène abti-apoptique BCL-x et exploitation comme cible thérapeutique pour lutter contre les cancers résistants aux chimiothérapies

Ronan Le Sénéchal

COLLEGE DOCTORAL BRETAGNE SCIENCES DE LA VIE ET DE LA SANTE

Université de Bretagne Occidentale

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE
Liberté Égalité Fraternité

THESE DE DOCTORAT DE

L'UNIVERSITE
DE BRETAGNE OCCIDENTALE

ECOLE DOCTORALE N° 637
Sciences de la Vie et de la Santé
Spécialité : *Biologie cellulaire, Biologie du développement*

Par

Ronan LE SENECHAL

Etude du rôle d'un G-quadruplex d'ARN dans l'épissage alternatif du gène anti-apoptotique *BCL-x* et exploitation comme cible thérapeutique pour lutter contre les cancers résistants aux chimiothérapies

L'épissage alternatif de *BCL-x* est contrôlé par la liaison du facteur d'épissage RBM25 à un G-quadruplex du pré-ARNm de *BCL-x*

Thèse présentée et soutenue à Brest, le 23 novembre 2023
Unité de recherche : UMR 1078

Rapporteurs avant soutenance :

Stefania MILLEVOI	Directrice de Recherche Inserm, CRCT, Toulouse
Laurent POULAIN	Directeur de Recherche CLCC François Baclesse, UMR 1086, UNICAEN, Caen

Composition du jury :

Président :	Laurent CORCOS	Directeur de Recherche Inserm, UMR1078 IBRBS, Brest
Examinateurs :	Stefania MILLEVOI	Directrice de Recherche Inserm, CRCT, Toulouse
	Laurent POULAIN	Directeur de Recherche CLCC F. Baclesse, UMR1086, Caen
	Alexandra MARTINS	Chargée de Recherche Inserm, UMR1245, Rouen
	Rodrigo PRADO MARTINS	Chargé de Recherche INRAe, Centre Val de Loire, Nouzilly
	Marie-Paule TEULADE-FICHOU	Directrice de Recherche CNRS, CMIB, Institut Curie, Orsay
Dir. de thèse :	Marc BLONDEL	Professeur UBO, UMR1078, IBRBS, Brest

Remerciements

La thèse est une aventure qui ne peut pas être résumée à un doctorant et son directeur de thèse. C'est pourquoi je tiens aussi à remercier les acteurs silencieux de cette thèse qui m'ont accompagné, aidé et soutenu pendant cette tranche de ma vie.

En premier lieu, j'aimerais remercier les membres de mon jury de thèse et de mon comité de suivi pour avoir pris le temps d'évaluer mon travail.

Ce travail n'aurait pas pu avoir lieu sans l'aide des différents financeurs : la ligue contre le cancer, l'association Gaëtan Saleün, la fondation de l'avenir et le ministère de l'enseignement supérieur et de la recherche.

Un grand merci aux membres de l'unité 1078 pour leur accueil et pour avoir pris le temps de m'aider quand vous le pouviez. Je voudrais en particulier remercier Emmanuelle Génin pour m'avoir accueilli dans son unité.

En écrivant ces quelques lignes, je prends le temps de me rappeler mes compagnons d'aventure sans qui cette thèse n'aurait pas eu la même saveur :

Les membres YL2 (désormais ASTRE), Alicia, Marc, Nadège, Trang et Gaëlle. Il y a tant de choses pour lequel je vous suis reconnaissant. La façon dont vous m'avez accueilli dans l'équipe, vos connaissances que vous m'avez partagées sans retenu, les sorties (sportives, gourmandes ou arrosées) et les autres nombreux bons moments. Un grand merci à vous tous.

Au reste de l'équipe ASTRE, merci pour votre convivialité pendant les goûters et les repas, et parfois aussi un peu pour la science. Un merci particulier à Enora et Solène, les sportives du midi (ou du dimanche), pour les sessions mémorables d'escalade et de wakeboard !

Je ne peux faire de remerciements sans mentionner Pierre, mon frère de thèse. Je ne m'étendrai pas sur tes remerciements au risque de ne pouvoir m'arrêter alors simplement merci.

Au Professeur Marc Blondel, merci pour ces quatre magnifiques années d'aventures à tes côtés. Tu as été un exemple aussi bien professionnellement qu'humainement. Je sors plus enrichi de cette thèse que je ne l'aurais imaginé. Merci de m'avoir inspiré.

A mes parents et ma sœur, merci pour votre soutien, votre patience et votre affection.

Table des matières :

Table des figures

Liste des abréviations

A

ADN : acide désoxyribonucléique

APAF-1 : apoptotic peptidase activating factor 1

ARN : acide ribonucléique

ATP : adénosine-triphosphate

B

Bcl-2 : B-cell lymphoma 2

Bcl-xL : B-cell lymphoma extra large

Bcl-xS : B-cell lymphoma extra short

BH : bcl-2 Homologie

BID : BH3 interacting-domain death agonist

BIN : Myc box-dependant interaction protein 1

BP : branching point

C

CAD : caspase-activated DNase

cFLIP : cellular FLICE-like inhibitory protein

D

DISC : death inducing signaling complex

DR6 : death receptor 6

E

EMT : epithélio-mésenchimateuse

ESE : exonic splicing enhancer

ESS : exonic splicing silencer

F

FADD : fas associated death domain

FoLDER : footprinting of long 7-deazaguanine-substituted RNAs

G

G4 : G-quadruplex

G4BP : G-quadruplex binding protein

H

hnRNP : heterogeneous nuclear ribonucleoproteins

I

ISE : intronic splicing enhancer

ISS : intronic splicing silencer

L

LAM : leucémies aigus myéloïdes

LB : lymphocyte B

M

Mcl-1 : induced myeloid leukemia cell differentiation protein

N

NCL : nucléoline

NHE III : nuclease hypersensitivity element III

NMD : nonsense-mediated decay

O

OH : groupement hydroxyle

P

PEMME : perméabilisation de la membrane externe mitochondriale

PKC : protein kinase C

PLA : proximity ligation assay

PPT : polypirimidine tract

PTBP1 : protein tyrosine phosphatase 1

PWI : proline tryptophane isoleucine rich

R

RBM : RNA binding motif

RE : arginine glutamate rich

RRM : RNA recognition motif

S

SMN2 : survival of motor neuron 2

snRNP : small nuclear ribonucleoprotein

SR : serine arginine rich

SS : splice site

T

TM : transmembranaire domain

U

UVRAG : UV radiation resistance-associated gene

Introduction

1. L'apoptose

La mort cellulaire est un processus physiologique essentiel aux organismes. Elle permet l'élimination des cellules endommagées, inutiles ou en excès et contribue, avec la synthèse de nouvelles cellules par mitose, au développement puis au maintien de l'homéostasie cellulaire. Ainsi, en moyenne dix milliards de cellules sont produites par jour afin de remplacer les cellules qui meurent quotidiennement (Renehan et al., 2001).

La compréhension de cette balance entre synthèse et élimination des cellules est d'autant plus importante que des déséquilibres de cette balance conduisent à des pathologies. C'est le cas de cancers où le cycle cellulaire et/ou la mort cellulaire ne sont plus à un niveau physiologique. L'un des principaux mécanismes de résistance des cancers consiste à prévenir leur propre mort, les rendant à la fois moins susceptible à la mort cellulaire mais entrainant également une résistance au traitement visant à éliminer ces cellules (Jin et al., 2004). A l'inverse, un excès de mort cellulaire conduit à certaines pathologies, telles que des maladies auto-immunes ou des maladies neurodégénératives (Ethell & Buhler, 2003).

L'apoptose est critique durant le développement du système nerveux et immunitaire. Ces systèmes sont développés à partir d'une surproduction de cellules. Les cellules nerveuses et immunitaires sont ensuite soumises à un contrôle qualité qui va éliminer par apoptose les cellules n'ayant pas de fonction synaptique ou n'ayant pas réussi à produire des antigènes. La syndactylie est l'une des malformations héréditaires les plus communes qui est provoquée par un défaut d'apoptose durant le développement de l'embryon. Cette pathologie caractérisée par la fusion des doigts ou des orteils. Bien que cette pathologie soit multifactorielle elle est toujours caractérisée par une suppression de l'apoptose et de la dégradation de la matrice extracellulaire (Al-Qattan, 2019).

Différents types de mort cellulaire sont retrouvés dans l'organisme. Les deux morts canoniques sont la mort cellulaire programmée et la nécrose. Bien que conduisant tous deux à la mort de la cellule, ces deux mécanismes diffèrent par les conséquences

sur l'environnement des cellules mourantes. La nécrose se produit en réponse à des stress physiologiques externes tels que l'inflammation ou l'hypoxie. Ce mécanisme passif est régulé par des protéines de l'inflammation et conduit à la rupture de la membrane plasmique. Le contenu cellulaire se retrouve alors dans l'espace intercellulaire entraînant une inflammation et des dommages au niveau des tissus. A l'inverse de la nécrose qui impacte fortement son environnement, la mort cellulaire programmée n'a pas d'effet sur celui-ci. Le mécanisme canonique de la mort cellulaire programmée est celui de l'apoptose. Le terme d'apoptose fut choisi en 1972 par Kerr et collaborateurs pour décrire un mécanisme d'élimination des cellules ayant un rôle complémentaire et opposé à la mitose conduisant à des changements moléculaires, biochimiques et morphologiques en réponse à un stimuli interne ou externe (Kerr et al., 1972). L'apoptose est caractérisée par une diminution de la taille de la cellule et une condensation de son noyau et de la chromatine. S'en suit une fragmentation du noyau, le décollement de la membrane plasmique et le bourgeonnement de corps apoptotique à la membrane plasmique. Ces corps apoptotiques sont ensuite pris en charge par les macrophages environnant afin d'être recyclés (Kurosaka et al., 2003). Puisqu'aucun matériel cellulaire n'est déversé dans l'espace intercellulaire, l'apoptose est immunologiquement inerte. Le mécanisme d'apoptose est un mécanisme cellulaire très conservé qui s'active en réponse à différents stimuli. Elle peut être activée par la cellule elle-même suite à la détection de dommages au niveau de senseurs intracellulaires. En résulte l'activation de la voie intrinsèque de l'apoptose. Cette voie se distingue d'une seconde voie, la voie extrinsèque qui résulte le plus souvent, d'une interaction entre une cellule endommagée ou anormale (par exemple infectée ou tumorale) et une cellule du système immunitaire (figure 1). Ces deux voies apoptotiques fonctionnent de façon synergique afin d'assurer le développement harmonieux, le maintien de l'homéostasie et le bon fonctionnement de l'organisme.

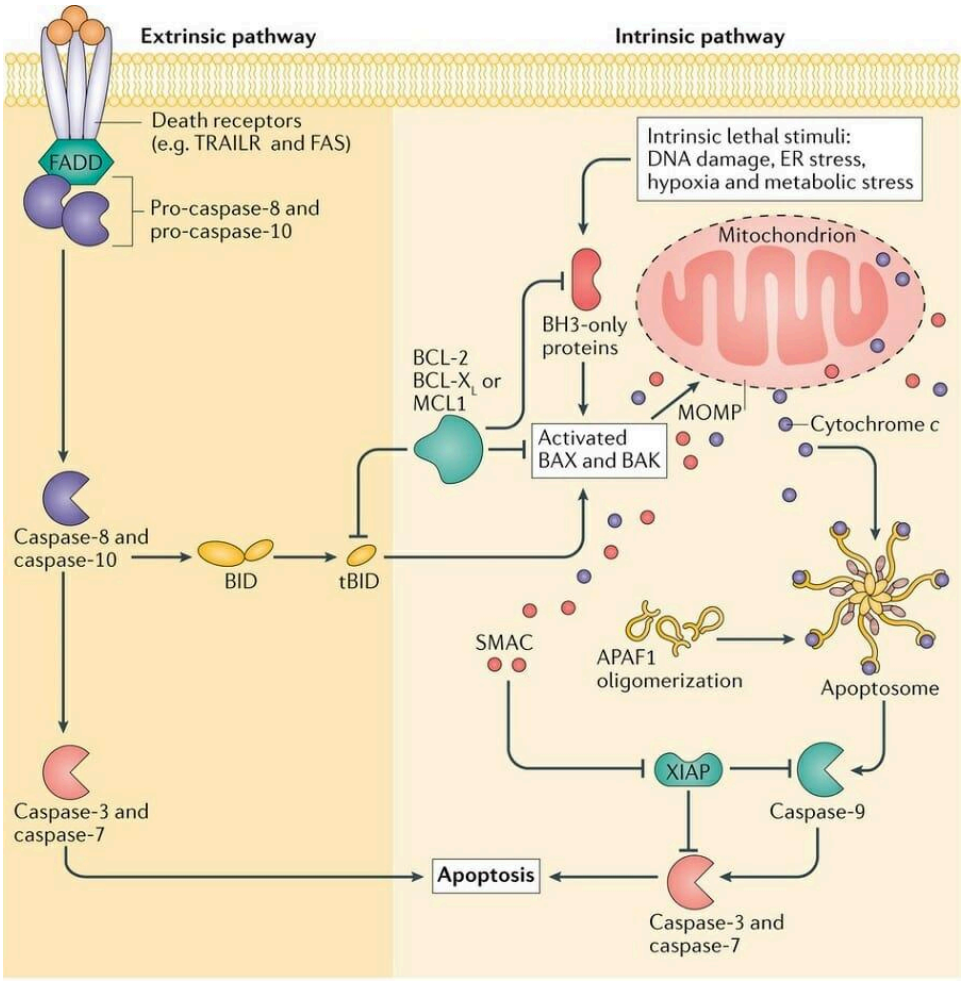


Figure 1 : représentation schématique des voies extrinsèque et intrinsèque de l'apoptose d'après Ichim et Tait 2016.

La voie extrinsèque de l'apoptose est déclenchée par la liaison de ligands de mort (FasL) aux récepteurs de mort (Fas). Après la trimérisation des récepteurs de mort, les protéines FADD, procaspases 8 et 10 sont recrutées pour former le complexe DISC. Ce complexe permet le clivage et donc l'activation des procaspases 8 et 10 qui vont ensuite activer les procaspases 3 et 7 exécutrices de l'apoptose.

La voie intrinsèque de l'apoptose débute par la détection d'un stimulus interne tel que des dégâts à l'ADN, un stress du réticulum endoplasmique, l'hypoxie ou encore un stress métabolique. Ce stimulus induit l'activation des protéines *BH3-only* et l'activation des protéines pro-apoptotique Bax et Bak qui vont former des pores au niveau de la membrane externe de la mitochondrie. Par ces pores, va sortir le cytochrome c, une molécule qui va s'oligomériser avec la protéine APAF1 pour former l'apoptosome. Ce complexe va activer la procaspase 9 et déclencher la cascade d'activation des caspases et l'apoptose de la cellule. Les protéines *BH3-only* sont inhibées par les protéines anti-apoptotique Bcl-2, Bcl-xL et Mcl-1 qui empêchent ainsi la perméabilisation de la membrane mitochondriale. La protéine *BH3-only* Bid, fait le lien entre les deux voies de l'apoptose. Cette protéine *BH3-only* peut être activée lors des voies extrinsèque ou intrinsèque pour aboutir à la PEMME et donc à l'apoptose de la cellule.

a. Voie extrinsèque de l'apoptose

La voie extrinsèque est la voie apoptotique médiée par les récepteurs de mort. Elle est initiée par la liaison de facteurs externes à la cellule, tel que FasL (APO-1), à son récepteur de mort associé, Fas (CD95), localisé à la membrane plasmique des cellules. La liaison de FasL à Fas déclenche la trimérisation des récepteurs à la membrane plasmique et la transduction d'un signal conduisant à l'apoptose (Papoff et al., 1999). Suite à la formation de ce trimère, les protéines FADD (*Fas Associated Death Domain*) sont recrutées du côté cytoplasmique de la cellule. L'association des protéines FADD avec les procaspases 8/10 et les protéines cFLIP conduit à la formation du complexe DISC (*Death Inducing Signaling Complex*) (Kischkel et al., 1995). Ce complexe permet le rapprochement et le clivage des procaspases 8/10 qui vont se dimériser au niveau de leur domaine protéase C-terminal et ainsi s'activer. Ces caspases initiatrices sont essentielles à l'activation de la voie extrinsèque et vont enclencher la cascade d'activation des caspases exécutrices 3 et 7 en les clivant, ce qui conduit à la mort de la cellule par apoptose. Il est important de noter que la protéine cFLIP est un composant particulier du complexe DISC. Cette protéine qui préente une organisation similaire aux procaspases 8/10, à ceci près que son domaine protéase C-terminal est un pseudo-domaine protéase qui est dépourvu d'activité protéolytique. Lorsque cette protéine est fortement exprimée dans le complexe DISC, elle va avoir une fonction anti-apoptotique par un effet dominant-négatif et ainsi inhiber l'activation des procaspases. Au contraire, lorsqu'elle est faiblement exprimée, cFLIP favorise l'activation des procaspases. Ainsi, en fonction de son niveau d'expression, cFLIP peut contrôler l'activation de la voie extrinsèque et le déclenchement de l'apoptose.

Les caspases sont des protéases à cystéine qui clivent les peptides et les protéines spécifiquement après un résidu aspartate. Elles coordonnent et exécutent le processus de mort cellulaire. Les caspases sont d'abord exprimées comme précurseurs sous forme inactive et nécessitent une activité de protéolyse au niveau de leurs résidus aspartates pour les activer. Cette particularité permet de contrôler l'activité pro-apoptotique de ces protéines. Les caspases peuvent s'activer soit par auto-protéolyse ou suite à un clivage par d'autres caspases ou protéases à cystéine (Stennicke et al., 1998). Les caspases sont divisées en deux groupes. Les caspases initiatrices (caspase 2, 8, 9 et 10) et les caspases exécutrices (3, 6 et 7). Les caspases

initiatrices sont responsables de l'initiation de la cascade d'activation, leur activation permet d'induire le clivage des caspases exécutrices qui sont responsables de dégradation protéolytiques dans la cellule au cours de l'apoptose. Les caspases exécutrices vont conduire au clivage d'endonucléases cytoplasmiques capables de dégrader le matériel nucléaire ainsi que d'autres protéases. La caspase 3 est particulièrement importante puisqu'elle active l'endonucléase CAD responsable de la dégradation de l'ADN chromosomique dans le noyau et de la condensation de la chromatine (Sakahira et al., 1998). La caspase 3 est également responsable de la réorganisation du cytosquelette et de la formation de corps apoptotique qui seront ensuite recyclés par les macrophages (Kurosaka et al., 2003).

b. Voie intrinsèque de l'apoptose

En plus de stimuli externes, l'apoptose peut être activée par des stimuli internes qui peuvent être de nature physique ou chimique. Ainsi, un stress du réticulum endoplasmique, des dégâts à l'ADN, ou l'hypoxie sont des stimuli capables d'activer la seconde voie de l'apoptose, la voie intrinsèque. Ces stimuli conduisent à la formation de pores apoptotiques dans la membrane externe de la mitochondrie ainsi qu'une perte du potentiel de la membrane mitochondriale entrainant le relargage de cytochrome c dans l'espace cytoplasmique. L'efflux de cytochrome c dans le cytoplasme conduit à la formation d'un complexe moléculaire : l'apoptosome. Ce complexe est à l'origine de l'initiation de la cascade d'activation des caspases. L'apoptosome est constitué de cytochrome c, d'APAF-1 et de la caspase 9. Dans une cellule saine, la protéine cytoplasmique APAF-1 se trouve sous forme de monomère. Cette protéine agit comme senseur de dégâts cellulaires en détectant la présence de cytochromes c dans le cytoplasme. La détection de cytochrome c induit l'oligomérisation d'APAF-1 pour former l'apoptosome avec la procaspase 9 (Srinivasula et al., 1998). Cette caspase est alors clivée par APAF-1 pour prendre sa forme active et pouvoir déclencher la cascade d'activation des caspases notamment des caspase 3 et 7, responsables de l'apoptose de la cellule.

La régulation de cette la voie intrinsèque est liée à la perméabilisation de la membrane externe de la mitochondrie. Cette étape clef est régulée par les protéines de la famille

Bcl-2 qui sont localisées à la membrane mitochondriale et dont les activités peuvent être anti- ou pro-apoptotique ce qui en fait des protéines régulatrices essentielles de la régulation de l'apoptose.

2. Régulation de la voie intrinsèque de l'apoptose

a. Perméabilisation de la membrane mitochondriale externe

La perméabilisation de la membrane mitochondriale externe (PEMME) est un processus rapide qui en cinq minutes conduit au relargage de cytochromes c dans le cytosol et en 15 minutes à l'activation complète des caspases, aboutissant à l'apoptose de la cellule (Goldstein et al., 2000). La PEMME est régulée par un réseau de protéines appartenant à la famille des protéines Bcl-2. Ces protéines sont cruciales à la régulation de l'apoptose et peuvent être anti- ou pro-apoptotiques. Ces protéines possèdent des domaines d'homologie Bcl-2 (BH) qui facilitent l'interaction entre les protéines de cette famille. Ces protéines sont classées en trois sous-familles en fonction des domaines BH qui composent ces protéines : les protéines multi-domaines anti-apoptotiques, les protéines pro-apoptotiques *BH3-only* et les-protéines exécutrices pro-apoptotiques (figure 2) (Bock & Tait, 2020).

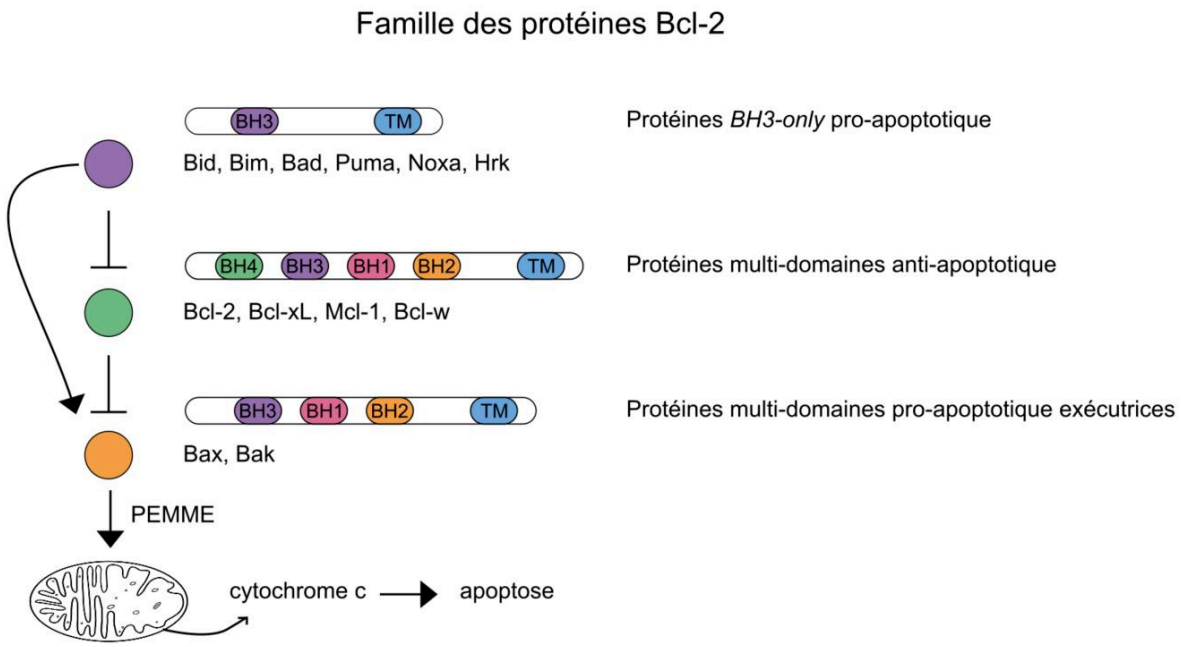


Figure 2 : représentation de la régulation de la perméabilisation de la membrane mitochondriale externe (PEMME) par les protéines de la famille Bcl-2.

L'activation des protéines pro-apoptotiques Bax et Bak est responsable de la PEMME qui conduit au relargage de cytochrome c dans le cytoplasme et ainsi à l'apoptose. Les protéines pro-apoptotiques possèdent les domaines d'homologie Bcl-2 (BH) 1 à 3 et sont régulées par les protéines anti-apoptotique possédant les quatre domaines BH (1-4). Enfin, les protéines *BH3-only* régulent la voie intrinsèque en inhibant l'activité des protéines anti-apoptotiques et en activant Bax et Bak.

Ce dernier groupe comprend les protéines Bax et Bak dont l'activation est directement responsable de la PEMME (Wei et al., 2001). Ces protéines sont responsables de la formation de pores mitochondriaux au niveau de la membrane mitochondriale ce qui en fait des protéines centrales dans la régulation de la voie intrinsèque de l'apoptose. Bien que ces protéines possèdent un rôle redondant, elles possèdent des différences. Bak est principalement localisé à la membrane mitochondriale tandis que Bax est majoritairement cytosolique avec une rétro-translocation constante de la mitochondrie vers le cytosol (Edlich et al., 2011). Suite à des stress, Bax s'accumule à la mitochondrie et est activée par des protéines *BH3-only* activatrices. La liaison des protéines *BH3-only* activatrices à Bax conduit à des changements de conformation au niveau du domaine transmembranaire de Bax permettant son insertion à la membrane mitochondriale (Letai et al., 2002). Les protéines *BH3-only* vont également permettre l'exposition du domaine BH3 de Bax et Bak permettant ainsi leur homodimérisation

(Dewson et al., 2009, 2012). L'insertion dans la membrane mitochondriale externe et l'activation de Bax/Bak par les protéines *BH3-only* présente le premier type d'activité pro-apoptotique des protéines *BH3-only*. Une autre fonction pro-apoptotique de ces protéines est portée par les protéines *BH3-only*, dites sensibilisatrices. Ces protéines sont capables de séquestrer les protéines multi-domaines anti-apoptotiques qui inhibent l'activation de Bax/Bak. Ainsi, ces protéines *BH3-only* vont-elles libérer Bax/Bak de dimères avec les protéines anti-apoptotiques. De sorte, les protéines pro-apoptotiques *BH3-only* sont des protéines essentielles à la régulation de l'apoptose en assistant les protéines Bax et Bak pro-apoptotiques.

Suite à la dimérisation de Bax et Bak, ces protéines vont entrainer la formation de pores lipidiques à la membrane mitochondriale formée par la fusion des feuillets de la membrane externe et interne (Bleicken et al., 2013). Une fois ce processus initié, la PEMME est renforcée par une boucle de rétroaction où les caspases vont également activer BID, une protéine *BH3-only* qui va accentuer le processus de PEMME par l'activation de nouvelles protéines Bax et Bak.

b. Régulation de la perméabilisation de la membrane mitochondriale externe par les protéines de la famille Bcl-2

Les protéines pro-apoptotiques Bax et Bak peuvent être inhibées par les protéines multi-domaines anti-apoptotiques. Ces protéines sont caractérisées par leur activité anti-apoptotique, par la présence de quatre domaines BH ainsi qu'un domaine transmembranaire (TM) nécessaire pour la liaison à la membrane mitochondriale. Cette sous-famille comprend les protéines Bcl-2, Bcl-xL, Bcl-W, BFL1 & Mcl-1. Le ratio entre les protéines anti- et pro-apoptotiques est crucial pour déterminer si la cellule entre en apoptose ou non. Pour décrire les interactions entre les trois sous-familles de protéines Bcl-2, quatre modèles sont proposés.

Dans le modèle d'activation directe, les protéines *BH3-only* sont divisées en deux groupes : les protéines *BH3-only* activatrices, comprenant tBid, Bim et Puma, et les protéines sensibilisatrices, comprenant Bad, Noxa, Bik, Bmf, Hrk et Bnip3. Les protéines *BH3-only* activatrices sont capables de s'associer avec Bax et Bak pour les activer et déclencher l'apoptose (Letai et al., 2002). Ces protéines sont également

reconnues par les protéines anti-apoptotiques qui vont les séquestrer et ainsi prévenir l'activation de Bax et Bak. Les protéines *BH3-only* sensibilisatrices ont la capacité de se lier avec les protéines anti-apoptotiques et, par cette interaction, vont libérer les protéines activatrices séquestrées par les protéines anti-apoptotiques. Dans ce modèle, l'inhibition de l'apoptose serait due à la séquestration des protéines *BH3-only* activatrices par les protéines anti-apoptotiques (H. Kim et al., 2006).

Dans le second modèle, dit de déplacement, Bax et Bak seraient constitutivement actives, mais séquestrées par les protéines anti-apoptotiques. Les protéines *BH3-only* vont libérer Bax et Bak de la liaison avec les protéines anti-apoptotiques pour promouvoir la PEMME. Les protéines *BH3-only* possèdent un spectre d'interaction limité à quelques protéines anti-apoptotique, aussi, pour initier l'apoptose, il faut un ensemble de plusieurs protéines *BH3-only* (L. Chen et al., 2005). Ce modèle permet ainsi d'expliquer la diversité des protéines *BH3-only* malgré une fonction redondante.

Dans le modèle incorporé, la membrane joue un rôle crucial dans l'activation de Bax et Bak et dans la formation de la PEMME. L'interaction de la membrane mitochondriale avec les protéines de la famille Bcl-2 déclenche des changements de conformation des protéines pro- et anti-apoptotique qui vont permettre l'insertion de ces protéines dans la membrane mitochondriale. Dans ce modèle les *BH3-only* activatrices jouent un rôle double dans lequel elles vont promouvoir l'insertion de Bax et Bak dans la membrane ainsi que permettre leur oligomérisation en les libérant des protéines anti-apoptotique (Lovell et al., 2008).

Enfin, Le modèle unifié vise à expliquer le rôle double des protéines *BH3-only* en deux modes. Dans le mode 1, les protéines anti-apoptotique se lient aux *BH3-only* afin de les séquestrer et d'inhiber l'activation des protéines pro-apoptotiques. Dans le mode 2, les protéines anti-apoptotique se lient à Bax et Bak pour les séquestrer et prévenir leur oligomérisation. Dans ce modèle, la séquestration de Bax et Bak est vue comme un mécanisme d'inhibition de l'apoptose plus efficace que le mode 1 (Llambi et al., 2011).

3. Le gène régulateur de l'apoptose *BCL-x*

a. Bcl-xL, une protéine anti-apoptotique majeure

i. Rôle de Bcl-xL dans la régulation de l'apoptose

Le contrôle de l'apoptose par les protéines multi-domaines anti-apoptotiques est un mécanisme essentiel à la survie des cellules et au développement des tissus. La PEMME peut être inhibée par la protéine Bcl-xL. Bcl-xL est capable d'inhiber l'apoptose en empêchant l'activation de Bax. Une fois intégrée à la mitochondrie, Bcl-xL forme un hétérodimère avec Bax, ce qui bloque l'homodimérisation de Bax et donc la formation de pore mitochondriaux. En outre, Bcl-xL régule de manière continue la quantité de protéines Bax à la membrane mitochondriale en la relocalisant depuis la membrane vers le cytosol. Ces deux mécanismes font de Bcl-xL une protéine anti-apoptotique majeur dans la régulation et l'inhibition de la voie intrinsèque de l'apoptose.

La découverte de la famille des protéines Bcl-2 a suivi une analyse de la translocation des chromosomes 14 à 18 dans les cellules B de lymphome montrant une surexpression du gène *BCL-2* leur conférant une résistance à la mort cellulaire (Cleary et al., 1986; Galteland et al., 2005). Cette propriété a permis de définir Bcl-2 comme une protéine anti-apoptotique. Le gène *BCL-X* fut cloné pour la première fois en 1993 (Boise et al., 1993). Ce gène conduit à la synthèse de la protéine anti-apoptotique Bcl-xL qui possède, avec Bcl-2, la séquence la plus longue des protéines de la famille Bcl-2. Bcl-xL contient les quatre domaines BH ainsi qu'un domaine TM. Cette protéine longue de 233 acides aminés partage le plus d'homologie avec les protéines anti-apoptotiques Bcl-W (51%) et Bcl-2 (45%), et ne partage qu'entre 18 et 25% d'homologie avec le reste des protéines de la famille Bcl-2.

La structure 3D de Bcl-xL fut déterminée en 1996 par cristallographie aux rayons X et par résonance magnétique nucléaire (Muchmore et al., 1996). Du fait de la complexité de produire les autres membres de cette famille protéique, l'étude de Bcl-xL a servi de référence pour comprendre le rôle des domaines BH dans le contrôle de l'apoptose et permis de comprendre comment les protéines de cette famille interagissent entre elles. Comme toutes les protéines anti-apoptotique, Bcl-xL possède les quatre domaines d'homologie BH ainsi qu'une région C terminale hydrophobe agissant comme une

ancre transmembranaire. Cette protéine est composée de huit hélices α. Les hélices 5 et 6 forment une épingle à cheveux centrale entourée par les hélices 3 et 4 d'un côté et les hélices 1, 2 et 8 de l'autre. Les domaines BH participent aussi à la formation de la structure tertiaire. Les domaines BH1 et BH3 forment une poche hydrophobe où le domaine BH3 d'autres protéines, notamment Bax, peuvent venir s'insérer. Ce domaine BH3 est donc essentiel pour l'interaction de Bcl-xL avec Bak. Le domaine BH3 est défini par une courte séquence de quatre résidus hydrophobes aux positions h1 à h4. Par la présence de ce domaine de 16 nucléotides de Bcl-xL est ainsi capable de séquestrer le domaine BH3 de Bax.

Bcl-xL est une protéine anti-apoptotique majeur, elle est capable d'inhiber la PEMME en se liant à la protéine *BH3-only* tBid et à la protéine pro-apoptotique Bax intégrée à la membrane mitochondriale (Edlich et al., 2011; Yao et al., 2009). Malgré une structure très similaire entre Bax et Bcl-xL, ces deux protéines ont pourtant une fonction antagoniste sur le déclenchement de l'apoptose (Suzuki et al., 2000). Bax et Bcl-xL fonctionnent de manière similaire au point qu'elles agissent en compétition dans la plupart des étapes conduisant à la PEMME, à la grande différence que Bcl-xL n'est pas capable d'induire la perméabilisation de la membrane externe de la mitochondrie. Par conséquent, Bcl-xL agit comme un dominant négatif de la protéine Bax.

Bax est une protéine cytoplasmique qui s'accumule à la mitochondrie suite à la détection de stress cytotoxique (Wolter et al., 1997). Elle est alors soumise à un changement de conformation lorsqu'elle interagit avec la membrane mitochondriale. Ce changement de conformation permet l'intégration de Bax à la membrane externe et une fois Bax intégrée, la protéine Bax va interagir avec la protéine *BH3-only* tBid intégrée à la membrane. Cette interaction déclenche un nouveau changement de conformation de Bax qui sera alors intégrée à la membrane dans une conformation qui permet son oligomérisation (Czabotar et al., 2013). D'autres molécules Bax seront ensuite activées par tBid. Bax peut aussi autoactiver d'autres protéines Bax cytoplasmiques pour induire leur intégration à la membrane et permettre l'oligomérisation de Bax et la formation de pores mitochondriaux conduisant à la PEMME (Tan et al., 2006).

De manière similaire, Bcl-xL est une protéine cytoplasmique qui, suite à une interaction avec la membrane mitochondriale, va subir un changement de conformation qui permet son insertion dans la membrane. tBid va ensuite déclencher l'intégration de

Bcl-xL ainsi que son activation (García-Sáez et al., 2009). Contrairement à Bax, Bcl-xL n'est pas capable de se dimériser, elle va alors séquestrer les protéines tBid empêchant ainsi l'activation de Bax. En outre, Bcl-xL va se lier avec les protéines Bax intégrées à la membrane pour empêcher leur oligomérisation et bloquer le recrutement et le mécanisme d'autoactivation de Bax. De la sorte, Bcl-xL agit comme un dominant négatif de Bax. Enfin, une fois que Bcl-xL et Bax se sont dimérisés, Bcl-xL est capable de rétro-transloquer Bax depuis la membrane mitochondriale dans le cytoplasme et de stabiliser Bax dans une conformation inactive (Edlich et al., 2011).

ii. Bcl-xL et cancer :

L'importance des protéines de la famille Bcl-2 dans la régulation de l'apoptose fait que ces protéines sont fréquemment retrouvées dérégulées dans des cancers. Bcl-xL est considérée comme l'un des régulateurs majeurs de la voie intrinsèque de l'apoptose. Aussi, Bcl-xL est surexprimée dans de nombreux cancers dans lesquels sa surexpression est associée à une résistance des cellules aux chimiothérapies et aux radiothérapies (Liu et al., 1999; Scherr et al., 2016). Il a été montré que Bcl-xL est surexprimée dans les chondrosarcomes et qu'inhiber l'activité anti-apoptotique de cette protéine conduit à une augmentation de l'apoptose ainsi qu'une sensibilisation des cellules cancéreuses aux chimiothérapies (de Jong et al., 2018). Bcl-xL est donc considérée comme une cible thérapeutique pertinente pour traiter les cancers et augmenter la sensibilité des cellules cancéreuses aux chimiothérapies.

Cette stratégie thérapeutique a donné naissance aux BH3 mimétiques, des molécules capables de se fixer à la poche hydrophobe des protéines anti-apoptotique (Bcl-2, Bcl-xL et Mcl-1) empêchant ainsi l'interaction des protéines anti-apoptotiques avec les protéines pro-apoptotiques (Chin, 2001; Holinger et al., 1999). Cette stratégie permet une augmentation de l'oligomérisation des protéines Bax et Bak et ainsi à une activation de la PEMME. L'ABT-737 est une molécule identifiée en utilisant une approche de prédictions bio-informatique basée sur la capacité du composé à se fixer à la poche hydrophobe de Bcl-2 et de Bcl-xL (E. F. Lee et al., 2007). Son efficacité fut avérée dans des lignées dérivées de lymphome et de leucémie où son utilisation conduit à une augmentation de l'apoptose des cellules cancéreuses. De plus, il est possible d'associer cette molécule avec des chimiothérapies plus classiques afin

d'augmenter les effets apoptotiques (van Delft et al., 2006). L'ABT-737 a ensuite été améliorée afin d'augmenter la bio-accessibilité et la solubilité de la molécule. Ce nouveau composé, nommé Navitoclax (ABT-263), présente également des effets prometteurs pour augmenter l'apoptose des cellules (Skwarska & Konopleva, 2023; Tse et al., 2008). Cependant, il a été montré que ces molécules conduisent à une augmentation de la mort des plaquettes. En diminuant l'activité anti-apoptotique de Bcl-xL, l'activité de Bax et Bak est augmentée et l'utilisation d'ABT-263 est associée de manière dose-dépendante à une diminution de la demi-vie des plaquettes et une thrombopénie. Cet effet secondaire important a considérablement réduit l'utilisation de ces composés en recherche clinique (de Vos et al., 2021). De plus, L'utilisation de BH3 mimétiques ciblant Bcl-2 et Bcl-xL est aussi fréquemment associée à des nouveaux mécanismes de résistance des cancers. Dans ces cancers résistants aux BH3 mimétiques, la protéine anti-apoptotique Mcl-1 est surexprimé et abroge l'effet de sensibilisation des cellules à l'apoptose par les molécules (B. Wang et al., 2014). Toutefois, L'utilisation d'un cotraitement basé sur des molécules visant à inhiber l'activité de Mcl-1, tel que l'EGF-R, en association avec le Navitoclax permet de contourner cette résistance aux BH3 mimétiques en cellule K562 et d'augmenter l'apoptose des cellules cancéreuses (Y.-C. Lee et al., 2020). Bien que le développement de drogues basées sur l'inhibition de Bcl-xL présente des difficultés, cibler Bcl-xL reste une stratégie pertinente dans le traitement de cancers notamment en association avec d'autres thérapies.

iii. Autres fonctions biologiques de Bcl-xL

En plus d'être un important régulateur de la voie intrinsèque de l'apoptose, Bcl-xL est aussi impliquée dans d'autres processus biologiques. Ainsi, Bcl-xL est également un régulateur de l'autophagie, le processus qui permet le recyclage et l'autodégradation de composés intra-cellulaires par les lysosomes. Bcl-xL interagit avec la protéine Beclin-1, le régulateur clef de l'autophagie. Cette protéine possède un domaine BH3 qui lui permet de se fixer dans la poche hydrophobe de Bcl-xL. De sorte, Bcl-xL est capable d'inhiber l'autophagie médié par Beclin-1 en l'isolant de hVps34 et en réduisant l'affinité de Beclin-1 avec UVRAG en stabilisant Beclin-1 sous forme de dimère (F. Zhou et al., 2011).

Bcl-xL est principalement localisée à la mitochondrie et dans le cytoplasme. Elle peut aussi être localisée dans le réticulum endoplasmique où elle régule le transport du calcium. Dans le réticulum endoplasmique, Bcl-xL interagit avec IP3R3 et facilite le transport du calcium depuis le RE jusqu'à la mitochondrie, participant ainsi de façon essentielle à l'homéostasie de calcium ainsi qu'à la production d'énergie de la cellule (Eno et al., 2012).

De par son importance dans la régulation de l'apoptose et dans le métabolisme du calcium, la protéine Bcl-xL est impliquée dans le développement des neurones, du système immunitaire et participe de manière importante à la neuroprotection. Bcl-xL est fortement exprimée dans l'hypothalamus, l'hippocampe et le cortex. Son expression est essentielle du développement des cellules nerveuses jusqu'à l'âge adulte où Bcl-xL contribue à la protection des cellules du système nerveux central. Son rôle neuroprotecteur est lié à ses fonctions anti-apoptotique ainsi qu'à sa capacité à augmenter la production d'ATP en modifiant l'apport de calcium. Ce rôle est essentiel puisque les neurones ont une demande importante en ATP, ainsi Bcl-xL participe directement à la neuroprotection en régulant la production d'énergie de la cellule et en prévenant l'hypoxie et la carence en glucose des neurones (Alavian et al., 2011). De plus, Bcl-xL protège les cellules de la mort en régulant l'expression du Death Receptor 6 (DR6), une protéine impliquée dans la dégradation des axones et l'apoptose des neurones. Dans ces cellules, la délétion de Bcl-xL est associée à une activation de la caspase 6 ainsi qu'à une augmentation de l'expression de DR6 (Park et al., 2015). Bien que le mécanisme exact ne soit pas encore connu. Les auteurs de ces travaux proposent un modèle dans lequel, suite à un stimulus interne tel que l'hypoxie, le récepteur DR6 est activé par une interaction avec un ligand. Ceci conduit à la séquestration de Bcl-xL par des protéines *BH3-only* et donc à l'activation de Bax qui elle-même conduit à la perméabilisation de la mitochondrie et l'activation des caspases. Ce modèle propose également l'existence d'une boucle de rétroaction positive où l'activation des caspases et/ou la séquestration de Bcl-xL conduirait à une augmentation de l'expression de DR6 et à la mort des neurones. Ainsi, Bcl-xL semble protéger les neurones de l'activation de DR6 en empêchant l'oligomérisation de Bax (figure 3).

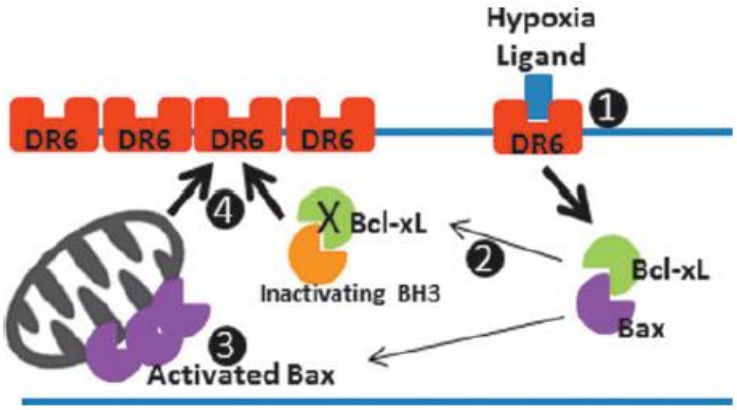


Figure 3 : modèle du rôle de Bcl-xL dans la régulation de la mort des neurones d'après Park *et al* 2014.

Suite à la détection d'un stimulus tel que l'hypoxie, le récepteur DR6 (*Death Receptor 6*) est activé suite à une interaction avec un ligand, ce qui va conduire à la séquestration de Bcl-xL par les protéines *BH3-only*. La protéine pro-apoptotique Bax est alors libérée et activée et va alors pouvoir induire la PEMME et le relargage de cytochrome c dans le cytoplasme. Ceci déclenche l'activation des caspases responsables de l'apoptose et induit des boucles de rétroaction positives qui vont conduire à l'augmentation de l'expression de DR6. Tout ceci aboutit à la mort des neurones.

Pendant le développement du système immunitaire, l'expression de Bcl-xL est variable et permet, en association avec Bcl-2, d'assurer un contrôle de l'apoptose pendant le développement des lymphocytes B (LB). Cette régulation de l'apoptose pendant le développement des LB est essentielle afin d'assurer la production de cellules fonctionnelles et prévenir des maladies auto-immunes. Bcl-xL participe avec Bcl-2 au développement des LB. Leurs profils d'expressions sont opposés pendant les différents stades de développement des LB (figure 4). Ainsi, pendant le stade des cellules Pro-B, Bcl-xL n'est pas exprimée tandis que Bcl-2 est fortement exprimée. Ces différences d'expression, de deux protéines anti-apoptotiques de la même famille, suggèrent des rôles différents dans le développement des LB. De manière intéressante, Bcl-2 et Bcl-xL ne sont pas exprimées pendant la transition de cellule Pré-B en cellules immature B. Ceci suggère que l'absence de deux protéines anti-apoptotiques majeures est nécessaire pour induire la mort des LB non fonctionnel ne passant pas le contrôle qualité (López-Hoyos et al., 1998).

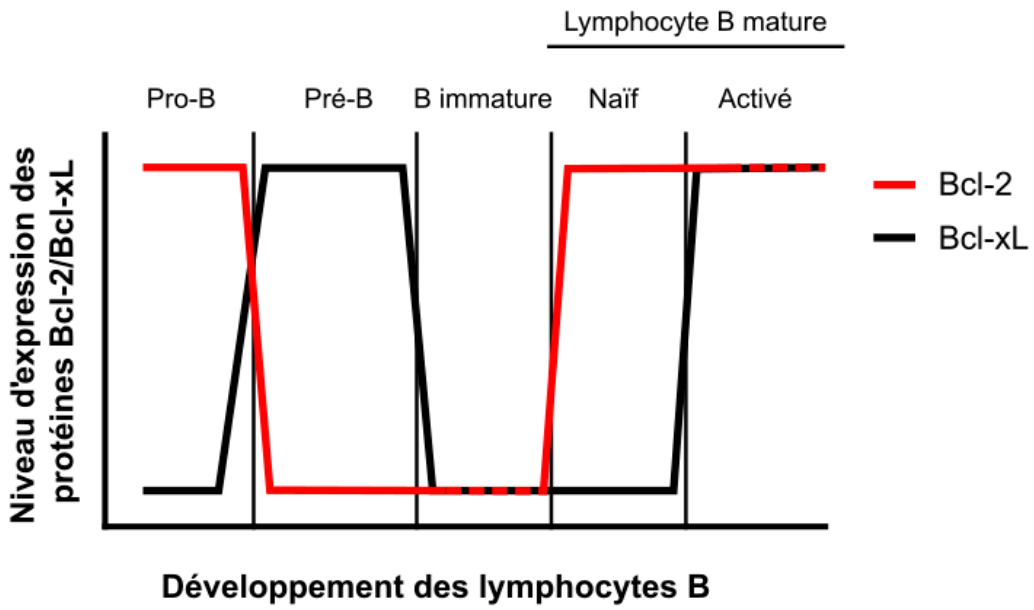


Figure 4 : expression des protéines anti-apoptotique Bcl-xL et Bcl-2 pendant le développement des lymphocytes B (LB). Adapté de Lopez-Hoyos *et al* 1998.

Les protéines Bcl-xL et Bcl-2 sont exprimées différentiellement et différemment pendant le développement des LB. Ainsi, il est suggéré que ces deux protéines ont des fonctions différentes et complémentaires dans la sélection des LB. Pendant le stade de LB immature, Bcl-xL et Bcl-2 ne sont pas exprimées. L'absence de ces deux protéines anti-apoptotiques majeurs est essentielle pour permettre la sélection des LB et ainsi éviter l'apparition de maladies auto-immunes.

b. Bcl-xS, une protéine pro-apoptotique particulière

La régulation de Bcl-xL est donc un processus essentiel qui permet le bon fonctionnement de la cellule. De plus, cette protéine s'avère être une cible thérapeutique pertinente pour certaines pathologies. Aussi, il est important de comprendre les mécanismes de régulation de Bcl-xL. Il est intéressant de noter que, comme plusieurs gènes de la famille des protéines Bcl-2, le gène *BCL-x* est régulé par un évènement d'épissage alternatif (Warren et al., 2019). L'épissage alternatif est un mécanisme qui permet la production de plusieurs ARNm matures à partir d'un seul pre-mRNA. Dans la plupart des cas, les isoformes alternatives conservent leurs fonctions pro- ou anti-apoptotique. Cependant, l'épissage alternatif des gènes *MCL-1* et *BCL-x* conduit à la synthèse de deux protéines aux fonctions antagonistes. Les formes canoniques de ces gènes possèdent des propriétés anti-apoptotiques, tandis que les isoformes synthétisées par épissage alternatif ont des fonctions pro-apoptotiques. L'isoforme alternative du gène *BCL-x* est Bcl-xS. Cette protéine est

produite suite à l'utilisation d'un site alternatif 5' d'épissage dans l'exon 2. Bcl-xS est une protéine de 170 acides aminés dépourvue des domaines BH1 et BH2, mais possède les domaines BH3, BH4 et TM (Boise et al., 1993). Bcl-xS possède des propriétés pro-apoptotiques ce qui en fait une protéine unique en son genre puisqu'il s'agit de la seule protéine pro-apoptotique qui possède le domaine BH4, aussi elle n'appartient pas à la famille des protéines *BH3-only*. Bcl-xS est capable de former des hétérodimères avec les protéines anti-apoptotiques Bcl-xL et Bcl-2 afin d'inhiber leur fonction anti-apoptotique (Lindenboim et al., 2001). Bcl-xS semble aussi être capable de former des hétérodimères avec certaines protéines pro-apoptotiques afin de déclencher l'apoptose de manière similaire à celle de Bid qui est capable d'activer Bax. L'équilibre entre la synthèse de Bcl-xS et de Bcl-xL est donc essentiel à la décision de la cellule à entrer, ou non, en apoptose. Cette balance est un mécanisme essentiel de régulation de l'homéostasie cellulaire et également au cours du développement. Pendant le processus de sélection des cellules T, le ratio Bcl-xS/Bcl-xL est augmenté afin d'éliminer les cellules ne passant pas le contrôle qualité. A l'inverse le ratio Bcl-xS/Bcl-xL est plus bas dans les cellules qui doivent être maintenues dans un état anti-apoptotique. Comprendre et maîtriser ce mécanisme d'épissage alternatif semble donc être une stratégie efficace pour limiter la synthèse de Bcl-xL et contrôler l'apoptose, notamment dans les cellules tumorales.

4. L'épissage des ARN

Les gènes eucaryotes sont pour la majorité interrompus par les introns, des séquences le plus souvent non codantes, qui doivent être éliminées de l'ARN pré-messager afin de produire un ARNm mature. Ainsi, les exons des gènes représentent les portions qui sont conservées et qui codent les protéines. Les exons doivent être joints ensemble afin de former l'ARNm mature. Dans les quelques 20.000 gènes qui composent l'Homme, le nombre moyen d'introns est de 8, et ils sont le plus souvent éliminés du transcrit principal par le mécanisme d'épissage (Roy & Gilbert, 2006). Le clivage des introns se fait au niveau de séquences conservées qualifiées de sites d'épissage. Ces séquences sont trouvées aux extrémités 5' et 3' des introns. Elles servent de point de reconnaissance pour la définition d'introns et d'exons. Le plus souvent, ces séquences

sont conservées et la séquence consensus qui définit l'extrémité 5' des introns est la séquence GU, tandis que la séquence AG est le plus souvent retrouvée à l'extrémité 3' des introns. Ces séquences sont très importantes au point que la substitution au niveau d'une de leurs bases inhibe le plus souvent complètement l'évènement d'épissage. Le point de branchement (BP) est la troisième séquence essentielle pour l'évènement d'épissage. Cette séquence est localisée entre 18 et 40 pb avant la fin 3' de l'intron et est toujours caractérisée par la présence d'une adénine. Enfin, le point de branchement est suivi par un tract polypyrimidines (PPT) de 15 à 20 nucléotides, localisé en amont du site d'épissage 3'.

a. La réaction d'épissage

L'épissage permet l'élimination des introns et la jonction des exons par deux réactions de transestérification qui vont conduire à l'excision des introns du pré-mRNA (Moore & Sharp, 1993). La première réaction de transestérification est l'étape de branchement (*branching* en anglais). Le groupement hydroxyle (OH) de l'adénosine du BP va « attaquer » le groupement phosphate de la guanine localisée dans le 5'ss (figure 5). Cette attaque nucléophilique produit une extrémité 5' clivée et un lasso intronique connecté au BP via une liaison phosphodiester. Dans la seconde étape de transestérification, le groupement OH de la guanine du 5'ss va venir attaquer le phosphate du site 3'. Cette modification conduit à la fusion des exons de manière covalente ainsi qu'à l'excision du lasso intronique qui est éliminé du pré-ARNm. Ces deux réactions sont catalysées par le splicéosome, un complexe macromoléculaire composé de cinq *uridines rich small nuclear ribonucleic acids* (snRNP) et de plus de 200 protéines associées qui vont participer à la formation du splicéosome. Cette machinerie s'assemble puis se dissocie après chaque évènement d'épissage.

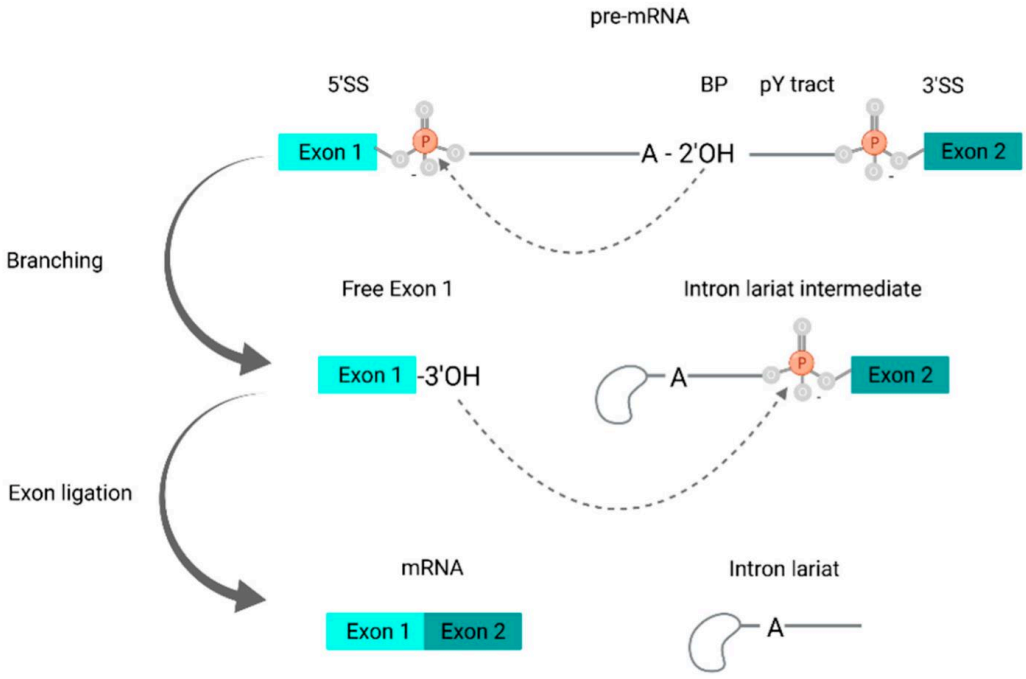


Figure 5 : l'épissage des pré-ARNm est réalisé suite à deux réactions de transestérifications, d'après Jacob *et al* 2022.

L'épissage est réalisé par deux réactions de transestérifications successives. La première réaction se fait pendant l'étape de branchement. Le groupement hydroxyle du point de branchement (BP) va attaquer le groupement phosphate de la guanine localisée dans le site d'épissage 5' (5'ss), ce qui produit une extrémité 5' clivée et conduit à la formation d'un lasso intronique connecté au BP. Dans la seconde réaction de transestérification, le groupement hydroxyle de la guanine du 5'ss attaque le groupement phosphate du site 3'ss ce qui aboutit à la fusion des exons de manière covalente ainsi qu'à l'excision du lasso intronique.

La formation du splicéosome s'effectue de manière séquentielle aux travers de nombreuses interactions entre les snRNPs et différents facteurs d'épissages (figure 6). Dans un premier temps, U1 snRNP est recrutée au niveau du site 5'ss. Les facteurs non snRNP SF1 et U2AF vont se fixer au BP et au PPT respectivement. Ensuite, U2 snRNP s'associe avec le point de branchement. Ces deux premières étapes constituent le pré-splicéosome ou complexe A. L'ajout du tri-snRNP U4/U6.U5 au splicéosome va permettre la formation du complexe pré-catalytique B. Ce tri-snRNP est un complexe préassemblé où U4 et U6 sont associées par appariement de bases complémentaire. La DEAD-Box hélicase Prp28 va alors libérer U1 snRNP du site accepteur 5' pour que U6 vienne prendre sa place. L'hélicase Brr2 va ensuite séparer U4 et U6, ce qui conduit à un changement de conformation de U6 qui va se replier et

s’associer avec U2, aboutissant au complexe B actif composé de U2/U5/U6 qui est catalytiquement actif. Ce complexe actif contient les deux ions métalliques Mg^{2+} nécessaires aux deux réactions de transestérifications (Fedor, 2002). Une fois U1 et U4 libérés du splicéosome, la première étape de transestérification a lieu. Une fois cette première réaction terminée, les snRNP U2 et U5 vont subir un réarrangement pour former le complexe C afin d’effectuer la seconde étape de transestérification. A la suite de cette seconde réaction, les exons sont joints bout à bout tandis que lasso intronique est libéré puis dégradé. Les trois derniers snRNP U2, U5 et U6 sont libérés et tous les composés seront réutilisés pour un nouvel événement d’épissage. Cette réaction constitue le premier splicéosome ou splicéosome « majeur » duquel dépend la majorité des introns. Chez la plupart des eucaryotes, un second splicéosome dit « mineur » prend en charge l’évènement d’épissage dans moins d’1% des introns. Ce second splicéosome est formé par l’assemblage des snRNP U11, U12, U4atac, U5 et U6atac et reconnait généralement des sites d’épissages aux séquences alternative (Sharp & Burge, 1997).

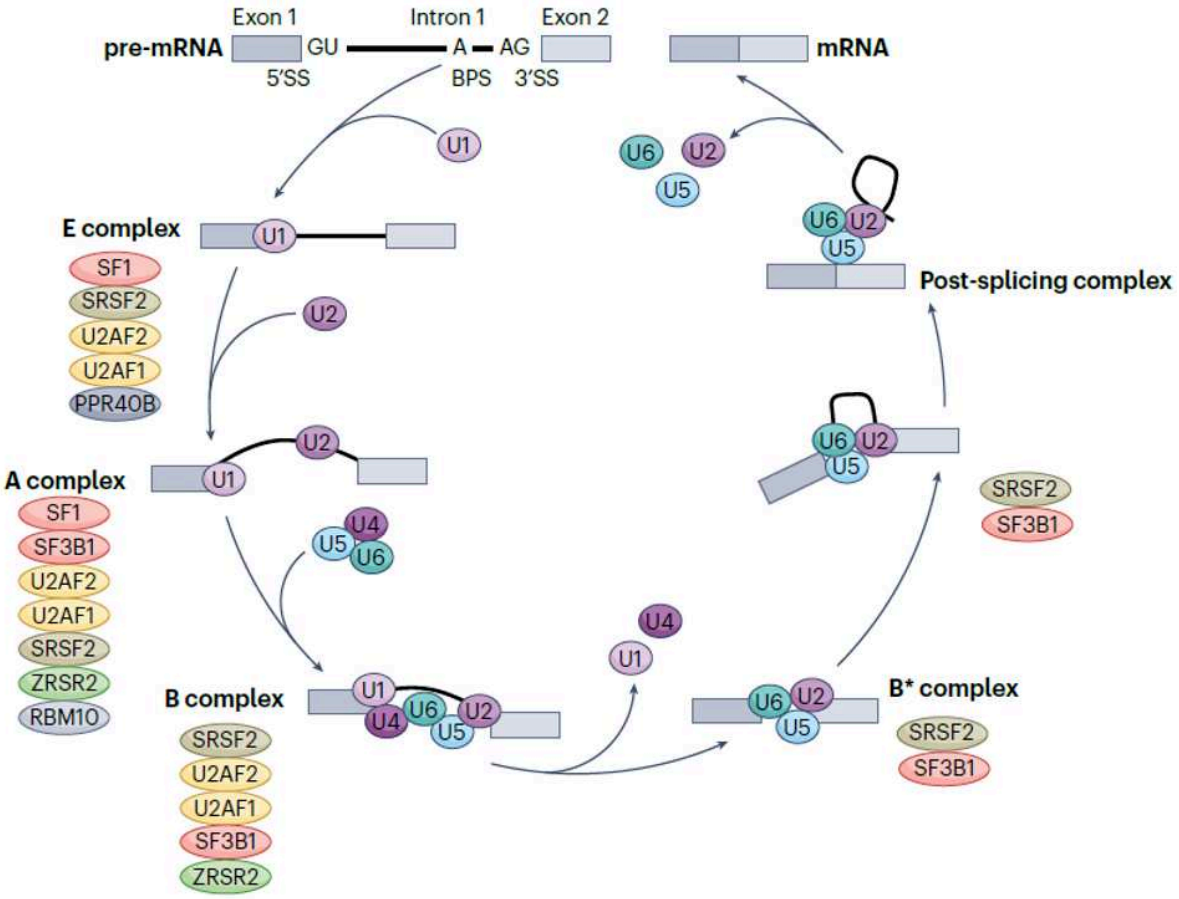


Figure 6 : l'assemblage du splicéosome et l'épissage d'un exon, d'après Bradley et Anczuków 2023.

Le splicéosome, un complexe macromoléculaire, s'assemble de manière séquentielle autour du pré-ARNm qui doit être épissé avec l'aide de snRNP et de facteurs d'épissages. La première étape de la formation du splicéosome est le recrutement de U1 snRNP au site 5' d'épissage, ce qui forme le complexe E. La formation de ce complexe active le recrutement de SF1 au BP et de UA1F2 au niveau du tract de polypyrimidine (PPT), respectivement. U2 snRNP est ensuite recruté au niveau du BP et forme le complexe A. Le recrutement du tri-snRNP U4/U6/U5 forme ensuite le complexe B qui est catalytiquement activé après le changement de conformation de U6. Ce complexe B désormais actif va réaliser la première étape de transestérification et la formation du lasso intronique. Les snRNP U2 et U5 vont ensuite subir un changement de conformation qui va aboutir à la seconde étape de transestérification et la jonction des deux exons. Le splicéosome est ensuite désassemblé et les composants pourront être réutilisés pour réaliser d'autres réactions d'épissage.

b. L'épissage alternatif

Quand le *Human Genome Project* a été initié dans les années 1990, il était estimé que 100.000 gènes seraient identifiés dans l'espèce humaines. Après une décennie, ce nombre a été revu à la baisse et estimé à 26.000 gènes. Aujourd'hui, ce nombre a encore diminué et il est estimé que le génome humain contient environ 20.000 gènes. Ce nombre est relativement bas comparé aux 6.000 gènes qui constituent la levure (*Saccharomyces cerevisiae*), un organisme unicellulaire bien moins complexe que l'Homme. La différence entre les organismes ne réside pas tant dans le nombre de gènes que dans la présence d'introns dans le génome. Chez la levure, seulement 3% des gènes possèdent des introns alors que chez l'homme, 97% des gènes possèdent des introns (Clark et al., 2002). Ces différences sont corrélées avec la diversité du protéome de chaque espèce. La levure exprime 6.050 protéines différentes tandis que ce nombre est d'environ 80.000 dans l'espèce humaine. Ainsi, l'idée d'un mécanisme de base permettant une expansion importante de la diversité protéique chez les eucaryotes a été proposée de manière simultanée par Philip Sharp et Richard Roberts en 1977. Cette idée implique qu'un pré-ARNm puisse conduire à plusieurs ARNm matures qui vont aboutir à des protéines aux structures et aux fonctions différentes. De sorte, la proportion des gènes sujets à un épissage alternatif est corrélée avec le niveau de complexité de l'organisme qui est à son plus haut niveau chez les primates. L'épissage alternatif ne conduit pas qu'à la synthèse de variants mineurs, mais également à la production d'isoformes qui peuvent agir comme des protéines distinctes. Ce mécanisme aboutit ainsi à une augmentation du potentiel biochimique d'un gène.

L'épissage dépend de la sélection des différents sites d'épissage 5' et 3' présent dans un gène, qui sont en compétition entre eux. En effet, ces séquences ne possèdent pas toutes la même capacité à attirer la machinerie d'épissage et, lorsque certaines sont moins souvent reconnues par le splicéosome, l'épissage devient alternatif. Les sites dits « forts » ont généralement une séquence consensus de GU pour le 5'ss et AG pour le site 3'ss. Ces séquences sont plus efficacement reconnues par les composants du splicéosome, ce qui conduit le plus souvent à un épissage constitutif. Un site moins bien reconnu par la machinerie d'épissage est alors qualifié de site « faible » et conduit à un épissage produisant à un ARNm mature différent. L'épissage des ARN est assisté

par des éléments réguleurs -*cis* et des facteurs d'épissage -*trans*. Les éléments -*cis* sont des séquences présentes dans l'ARN-pré messager qui participent à la sélection des sites 5' et 3'. Ces séquences -*cis* peuvent, soit favoriser l'utilisation d'un site d'épissage auquel cas il s'agit d'*ISE* (*Intronic Splicing Enhancer*) et d'*ESE* (*Exonic Splicing Enhancer*), soit inhiber l'utilisation d'un site d'épissage, ce sont les *ISS* (*Intronic Splicing Silencer*) et les *ESS* (*Exonic Splicing Silencer*). Ces séquences spécifiques de l'ARN sont reconnues par les facteurs -*trans* qui sont généralement représentés par les protéines de la famille des SR (Serine Arginine Rich) et de la famille des hnRNP. En général, les protéines SR ont tendance à favoriser la reconnaissance d'exons au travers des *ESE* et des *ISE*, tandis que les hnRNP vont interagir avec les *ESS* et les *ISS* afin d'inhiber la reconnaissance du site d'épissage. Ces deux familles protéiques possèdent donc des rôles opposés et doivent être finement régulées pour assurer l'épissage optimal pour les cellules (Cáceres & Kornblihtt, 2002).

L'épissage alternatif conduit à l'organisation différentielle d'exons d'un transcrit de base de façon à produire des ARNm distincts. Ces combinaisons différentielles peuvent se produire de différentes manières. Soit par saut d'exon dans 40% des événements d'épissage chez les eucaryotes, soit par l'utilisation d'un 3'ss alternatif dans 18,4% des cas, ou encore l'utilisation d'un 5'ss dans 7,9% des cas et enfin la rétention d'intron dans 5% des cas. Il est possible d'observer d'autres formes d'épissage alternatif dans de plus rares cas comme des exons mutuellement exclusifs, des promoteurs alternatifs ou des sites de polyadénylation particulier (figure 7).

En plus de pouvoir produire des ARNm aboutissant à des protéines aux fonctions différentes, l'épissage alternatif peut agir comme un mécanisme on/off des gènes par l'introduction de codons stop prématurés. L'ajout de codons stop prématurés dans un transcrit en fait des cibles pour le mécanisme de dégradation des ARNm non-sens (NMD). La NMD agit comme mécanisme de contrôle qualité des nouveaux ARNm afin de dégrader ceux qui présentent un codon stop prématuré. Ce mécanisme est essentiel pour empêcher la synthèse de protéine tronqués qui pourraient avoir un effet nocif sur l'organisme. Ainsi, couplée à l'épissage alternatif, la NMD permet un contrôle de la régulation de gènes et de l'expression de protéines en diminuant la traduction de la protéine (Lewis et al., 2003).

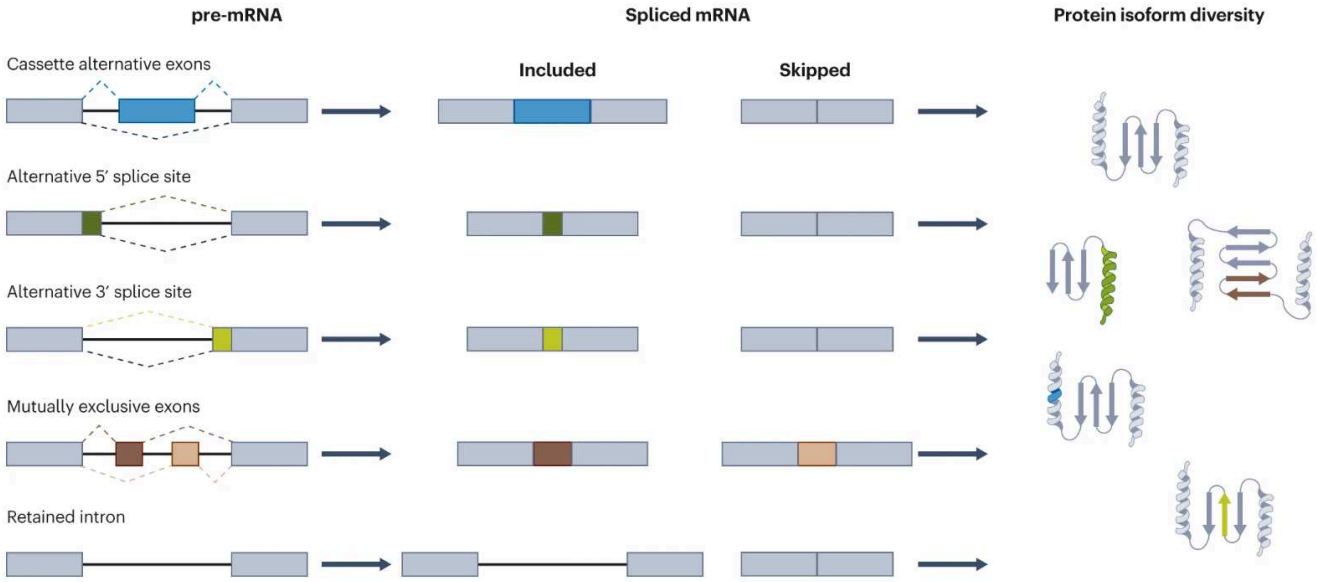


Figure 7 : les différentes sortes d'épissage alternatif, adapté de Bradley et Anczuków 2023.

Différentes sortes d'épissage alternatif sont possibles : épissage alternatif de l'exon cassette, site 5' d'épissage alternatif, site 3' d'épissage alternatif, exons mutuellement exclusif et rétention d'introns. Ces différents épissages conduisent à des ARNm qui seront traduit en des protéines avec des séquences et des fonctions différentes.

5. L'épissage de *BCL-x*

Le gène régulateur de l'apoptose *BCL-x* est l'un des nombreux gènes régulés par un épissage alternatif. Ce gène est constitué de trois exons, la traduction débute au début de l'exon 2 et l'exon 3 contient le codon stop (figure 8). Depuis sa découverte en 1993, il est connu que le gène *BCL-x* est régulé par un évènement d'épissage alternatif au niveau de l'exon 2 qui conduit à la synthèse de deux protéines aux effets antagonistes (Boise et al., 1993). La sélection du site 5' canonique conduit à la synthèse de la protéine anti-apoptotique Bcl-xL qui contient l'exon 2 en entier et donc les quatre domaines d'homologie BH. La protéine alternative pro-apoptotique Bcl-xS est produite suite à la sélection du site 5' alternatif. Cette protéine, plus courte de 63 acides aminés, est codée par un ARNm dépourvu de la partie 3' de l'exon 2, elle ne contient donc pas les domaines BH1 et BH2. De sorte, Bcl-xS est capable de former des hétérodimères avec Bcl-xL ce qui a pour conséquence de libérer les protéines Bax et Bak séquestrées par Bcl-xL et d'initier la PEMME. La sélection site 5' d'épissage de *BCL-x* est donc un

mécanisme essentiel dans la résistance à l'apoptose de la cellule puisque le ratio entre Bcl-xS et Bcl-xL détermine la capacité d'une cellule à entrer, ou non, en apoptose. En outre, il a été montré que la surexpression de Bcl-xS entraine une augmentation de l'apoptose des cellules cancéreuses. Dans certaines formes de diabète du type 1, les cellules ß du pancréas présentent une apoptose excessive à la suite d'un changement dans l'épissage de *BCL-x* en faveur de Bcl-xS. Ainsi, pouvoir moduler l'épissage de *BCL-x* représente une cible thérapeutique pertinente pour traiter différentes pathologies, parmi lesquelles des cancers chimio-résistants ou encore certaines formes de diabète. C'est dans ce but que de nombreuses études ont cherché à comprendre les facteurs capables d'influer sur la sélection du site d'épissage afin de pouvoir mieux comprendre la régulation de ce gène qui est un régulateur majeur de l'apoptose. Aujourd'hui, de nombreux éléments régulateurs ont été montrés comme capables d'influer sur le choix des sites d'épissage. Cela a permis l'identification de facteurs d'épissage, mais également de voies de transduction qui sont impliquées dans l'épissage de *BCL-x* (figure 9) (Dou et al., 2021).

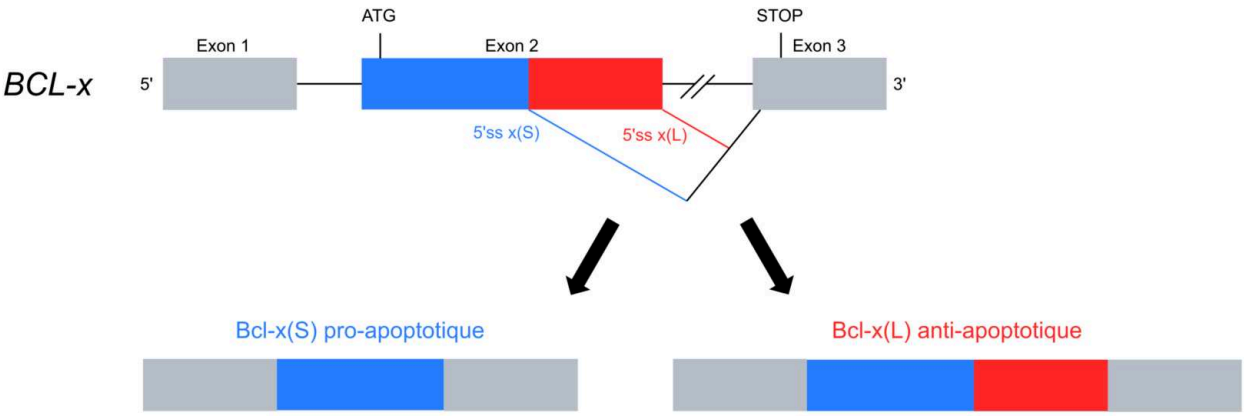


Figure 8 : l'épissage (alternatif) du gène *BCL-x.*

Le gène *BCL-x* est constitué de trois exons. La traduction débute au début de l'exon 2. Deux sites 5' d'épissage sont localisés dans l'exon 2. Lorsque le site 5' d'épissage canonique est sélectionné, la protéine anti-apoptotique Bcl-x(L) est synthétisée. Lorsque le site 5' d'épissage alternatif est sélectionné, l'ARNm mature est amputé de la partie 5' de l'exon 2 et conduit à la synthèse de la protéine pro-apoptotique Bcl-x(S).

a. Facteurs d'épissage influençant l'épissage alternatif de *BCL-x*

L'épissage alternatif de *BCL-x* est régulé par des facteurs d'épissage nucléaires. Ces facteurs peuvent augmenter ou diminuer l'utilisation de l'un ou l'autre des deux sites d'épissage 5'. L'expression de ces facteurs est variable et dépendante de changements cellulaires qui peuvent conduire à des changements dans l'épissage de *BCL-x* et donc dans le ratio Bcl-xS/Bcl-xL.

La plupart des protéines impliquées dans l'épissage de *BCL*-x sont des facteurs d'épissage généraux. Les protéines hnRNPs A1, A2/B1, K, I (PTPB1) et F/H participent à la sélection du site d'épissages 5' (Dominguez et al., 2010; Garneau et al., 2005a; Revil et al., 2009). Le site de liaison de hnRNP I est localisé dans la région *-cis* régulatrice B2G localisé après le site 5' de Bcl-xS. Cet élément B2G sert également de site de liaison pour la protéine hnRNP F/H qui une fois fixée sur cet élément, va augmenter la synthèse de Bcl-xS. Les protéines SR, SRSF1, SRSF2, SRSF3, SRSF7, SRSF9 et SRSF10 sont également impliquées dans la sélection des sites d'épissages 5' (Bielli et al., 2014; Shkreta et al., 2016). Alors que SRSF1, SRSF7 et SRSF9 favorisent la sélection du site 5' canonique, SRSF2, SRSF3 et SRSF10 favorisent elles la sélection du site 5' alternatif conduisant à la synthèse de Bcl-xS.

SRSF9 est capable de se fixer à l'élément régulateur B3 localisé en amont du site 5' d'épissage canonique. Cet élément sert généralement de point de liaison pour le snRNP U1 qui va favoriser l'épissage de Bcl-xS. En se fixant à l'élément B3, SRSF9 est capable de déplacer U1 snRNP et *in fine* conduit à une augmentation de la synthèse de Bcl-xL (Michelle et al., 2012).

En plus des protéines hnRNPs et SR, d'autres facteurs d'épissage influent sur la sélection des sites d'épissage de *BCL-x*, tel que RBM4, RBM10, RBM11, RBM25, TRA2ß, SF3B1 (Inoue et al., 2014; Massiello et al., 2006; Pedrotti et al., 2012; Y. Wang et al., 2014; A. Zhou et al., 2008). La protéine Sam68 s'associe à la protéine hnRNP A1 pour se lier au pré-mRNA de *BCL-x*. La liaison de Sam68 au pré-ARNm conduit à la sélection du site 5' alternatif et à la synthèse de Bcl-xS (Paronetto et al., 2007). Ces facteurs d'épissage sont généralement en compétition pour des éléments de liaison au pré-ARNm ou fonctionnent par association pour aboutir à la sélection d'un des sites d'épissages. Aussi, de nombreuses interactions ont été décrites entre les différents

facteurs d'épissage. La liaison de hnRNP A1 à Sam68 est essentielle pour l'activité d'épissage de Sam68, en absence de la protéine hnRNP A1 ou lorsqu'elle est mutée, la synthèse de Bcl-xS est significativement diminuée. SRSF10 fonctionne en association avec les protéines hnRNP A1, A2 et Sam68 afin d'augmenter la synthèse de Bcl-xS en réponse à des dégât à l'ADN (Cloutier et al., 2018). RBM4 est en compétition avec le même site que SRSF1. Lorsque RBM4 utilise ce site, la synthèse de Bcl-xS est augmentée. hnRNP I augmente la synthèse de Bcl-xS en se liant à un site polypyrimidique localisé entre les deux sites 5'ss (Y. Wang et al., 2014). Suite à cette liaison hnRNP I déplace SRSF1 de son site de liaison ce qui conduit à une diminution de la sélection du site de Bcl-xL. En cellules saines, SRSF10 interagit avec hnRNP K et F/H pour diminuer Bcl-xS. Cependant, suite à des dommages à l'ADN, SRSF10 est déphosphorylée et perd sa capacité d'interagir avec hnRNP F/H. Ainsi, hnRNP F/H est alors capable de se fixer dans la région B2G, un élément riche en G localisé après le site d'épissage alternatif conduisant à la production de Bcl-xS, pour augmenter sa synthèse de en favorisant l'utilisation du site d'épissage 5' alternatif (Shkreta et al., 2016).

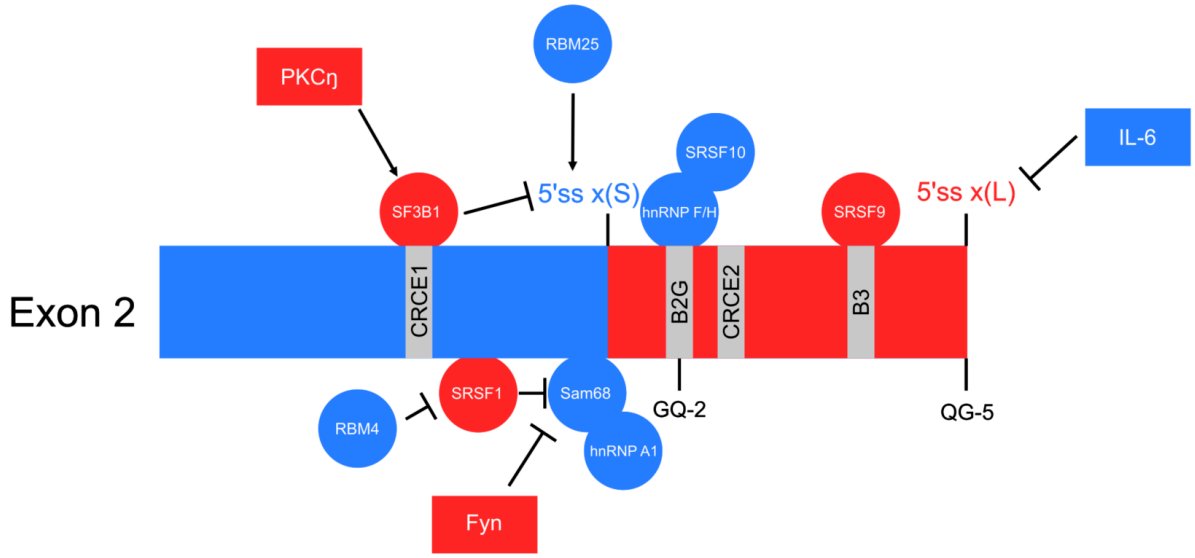


Figure 9 : régulation de l'épissage de *BCL-x* et facteurs connus influençant la sélection du site 5' d'épissage.

Plusieurs facteurs d'épissages influencent la sélection des sites 5' d'épissage. Ces facteurs sont représentés par des ronds rouges lorsqu'ils favorisent la sélection du site 5' canonique d'épissage conduisant à la forme longue Bcl-x(L), ou bleus lorsqu'ils favorisent la sélection de site 5' alternatif de Bcl-x(S). Les voies de signalisation impliquées dans la sélection du site d'épissage sont représentées avec des carrés avec le même code couleur. En gris sont représentées les séquences *cis* régulatrices du pré-ARNm de *BCL-x.* Enfin, deux G-quadruplex (rG4), le GQ-2 et le GQ-5, sont indiqués sur le schéma. Le rG4 GQ-2 est localisé à proximité du site 5' alternatif tandis que le rG4 GQ-5 est situé au niveau du site 5' d'épissage canonique de *BCL-x*.

b. Voie de signalisation influençant l'épissage alternatif de *BCL-x*

La liaison des protéines SR aux éléments régulateur de l'ARNm est régulée par la phosphorylation par des protéines kinases telles que CLK1 et SRPK1. La phosphorylation de SRSF1 par SRPK1 facilite le transport de SRSF1, sa localisation dans les *speckles* et l'interaction avec les pré-ARNm pour réguler l'épissage des ARN cibles (Serrano et al., 2016). Les voies de signalisation sont donc des facteurs influençant l'épissage des gènes. Une dérégulation récurrente des voies de signalisation ainsi qu'un épissage aberrant sont d'ailleurs retrouvés dans tous les cancers.

La protéine kinase C (PKC) est une famille de sérine/thréonine kinases (des kinases impliquées dans la phosphorylation des groupes hydroxyles des chaines latérales de sérines et de thréonines de leurs protéines cibles. L'activité des PKC est associée avec une diminution de l'utilisation du site 5' alternatif de *BCL-x,* mais le mécanisme exact

n'est pas encore connu aujourd'hui. Les PI3K sont une famille de lipide kinases impliquées dans la croissance cellulaire, la prolifération et sont capables de phosphoryler le groupe hydroxyle en position trois des phosphatidylinositols. Ces kinases sont essentielles à l'épissage alternatif de *BCL-x* et sont associées avec une augmentation de la protéine anti-apoptotique Bcl-xL dans les cancers NSCLC (Shultz et al., 2012). La PKC atypique, PKC-η, régule quant à elle l'épissage de *BCL-x* en modifiant l'expression de SF3B1, un des composants du spliceosome.

La tyrosine kinase Fyn appartient à la famille des Src kinases. Elle a un rôle important dans le fonctionnement neuronal ainsi que dans les processus inflammatoires. Fyn peut phosphoryler le facteur d'épissage Sam68 afin de réguler l'épissage de *BCL-x*. La protéine Sam68 a une localisation principalement nucléaire où elle promeut l'utilisation du site 5' alternatif conduisant à la synthèse de la protéine pro-apoptotique Bcl-xS. Lorsque Sam68 est phosphorylée par Fyn, la protéine est relocalisée dans le cytoplasme et perd sa capacité à réguler l'épissage de *BCL-x*. Cette phosphorylation de Sam68 conduit à une augmentation de Bcl-xL en cellule. De plus, Fyn régule l'expression des hnRNP A2 et B. Fyn conduit à une surexpression de ces deux facteurs qui vont favoriser la synthèse de la protéine anti-apoptotique Bcl-xL (Paronetto et al., 2007).

L'IL-6 est une cytokine pro-inflammatoire et une myokine anti-inflammatoire. Il a été observé que le traitement de cellules K562 par l'IL-6 entraine une réduction de Bcl-xL. Cependant, l'effet et les cibles de ces cytokines ne sont pas connus , aussi l'identification des facteurs de d'épissage en aval reste à déterminer pour comprendre l'effet des cytokines sur la sélection du site d'épissage (C. Y. Li et al., 2004).

Enfin, les céramides, des régulateurs importants de la réponse des cellules au stress et aux mécanismes de croissance, sont impliquées dans l'épissage de *BCL-x.* Les céramides sont capables de moduler l'activité des protéines SR en provoquant leur déphosphorylation. Ainsi, l'activité des céramides est associée à une augmentation de Bcl-xS ainsi qu'à une sensibilité accrue aux chimiothérapies. De plus, deux éléments de réponse aux céramides, CRCE1 et CRCE2, ont été identifiés dans l'exon 2 de *BCL-x* et sont nécessaires pour la sélection du site 5' alternatif de *BCL-x* en réponse aux céramides. SF3B1 est nécessaire pour la réponse aux céramides dans l'épissage de *BCL*-x en se fixant au CRCE1 (Massiello et al., 2004, 2006).

c. Structures secondaires influençant l'épissage alternatif de *BCL-x*

Les molécules d'ARN ont la capacité de former des structures secondaires et tertiaires. La formation de ces structures secondaires est un mécanisme essentiel de régulation de nombreux processus des ARN tels que la transcription, l'épissage et la traduction. Les structures secondaires qu'adoptent un pré-ARN messager influent sur la perception des sites d'épissages et des exons de la machinerie d'épissage. Par exemple, le saut de l'exon 7 du gène *SMN2* est corrélé avec la stabilité d'une structure tige-boucle qui séquestre le site 5' d'épissage (Singh et al., 2007). Ces structures secondaires peuvent ainsi empêcher la machinerie d'épissage d'accéder aux sites d'épissage. Au contraire, certaines structures secondaires peuvent être perçues par des protéines et attirer le splicéosome. Au sein du pré-ARN messager de *BCL-x* différentes structures secondaires ont été identifiées et influent sur l'épissage du gène. Ainsi, dans la région régulatrice -*cis* E3b de l'exon 3 a été identifiée une structure tige-boucle de 27 nucléotides. La présence de cette structure augmente l'utilisation du site d'épissage 5' alternatif. Des mutations visant à empêcher la formation de la tige-boucle conduisent à l'inverse à une augmentation de l'utilisation du site d'épissage 5' canonique et donc à une augmentation de Bcl-xL, validant ainsi l'importance de la formation de cette structure dans l'épissage alternatif de *BCL-x* (J. Lee et al., 2012).

La région régulatrice -*cis* B2 est essentielle à la synthèse de Bcl-xS. Lorsqu'elle est manquante ou mutée, la synthèse de Bcl-xS est totalement inhibé (Garneau et al., 2005a). La présence de répétitions de guanine au niveau de l'exon 2 et plus particulièrement au niveau de cette région B2 a abouti à des recherches sur la capacité de l'ARN pré-messager de *BCL-x* à former des G-quadruplex d'ARN (rG4). Par une technique de substitution des guanines en déazaguanines (FoLDER), deux rG4 ont été identifiés dans l'ARNm de *BCL-x* au niveaux des deux sites 5' d'épissages (figure 10) (Weldon et al., 2017).

Les G-quadruplex (G4) sont des structures secondaires non canoniques qui se forment dans les acides nucléiques riches en guanines (partie détaillée dans la section 7 de l'introduction et dans la figure 11). Dans ces structures, quatre guanines sont liées entre elles par des liaisons hydrogènes de type Hoogsteen, pour former un G-quartet, l'unité de base du G4 qui est stabilisé par un cation, généralement le potassium (K^+).

L’empilement d’au moins deux G-quartets permet la formation du G4. Ces structures sont impliquées dans différents processus cellulaires dont la réplication, la transcription, l’épissage et la traduction.

Dans l’épissage, les G4 sont des structures secondaires qui peuvent augmenter ou diminuer l’utilisation de sites d’épissage. L’identification de rG4 au niveau des sites d’épissages 5’ de *BCL-x* pose la question de leur rôle potentiel dans la régulation de l’épissage alternatif de ce gène. En effet, la stabilisation de ces rG4 par des dérivés d’éllipticines ou de quindolines affecte l’épissage de *BCL-x* (Weldon et al., 2018). De plus, le rG4 localisé à proximité du site 5’ d’épissage alternatif, est localisé au niveau du site -*cis* B2 où se fixe hnRNP H/F. Il a été montré que la séquestration de motifs GGG potentiellement capables de former des G4 par hnRNP H/F, empêche la formation de rG4 (Dominguez et al., 2010). La liaison de hnRNP H/F est donc mutuellement exclusive avec la formation de rG4. Le rôle potentiel de rG4 dans l’épissage de *BCL-x* pourrait donc être modulé par des protéines telles que hnRNP H/F. Cependant, au moment où cette thèse a démarré, il n’était pas encore connu par quels moyens ces rG4 pouvaient influencer la sélection des sites d’épissages 5’.

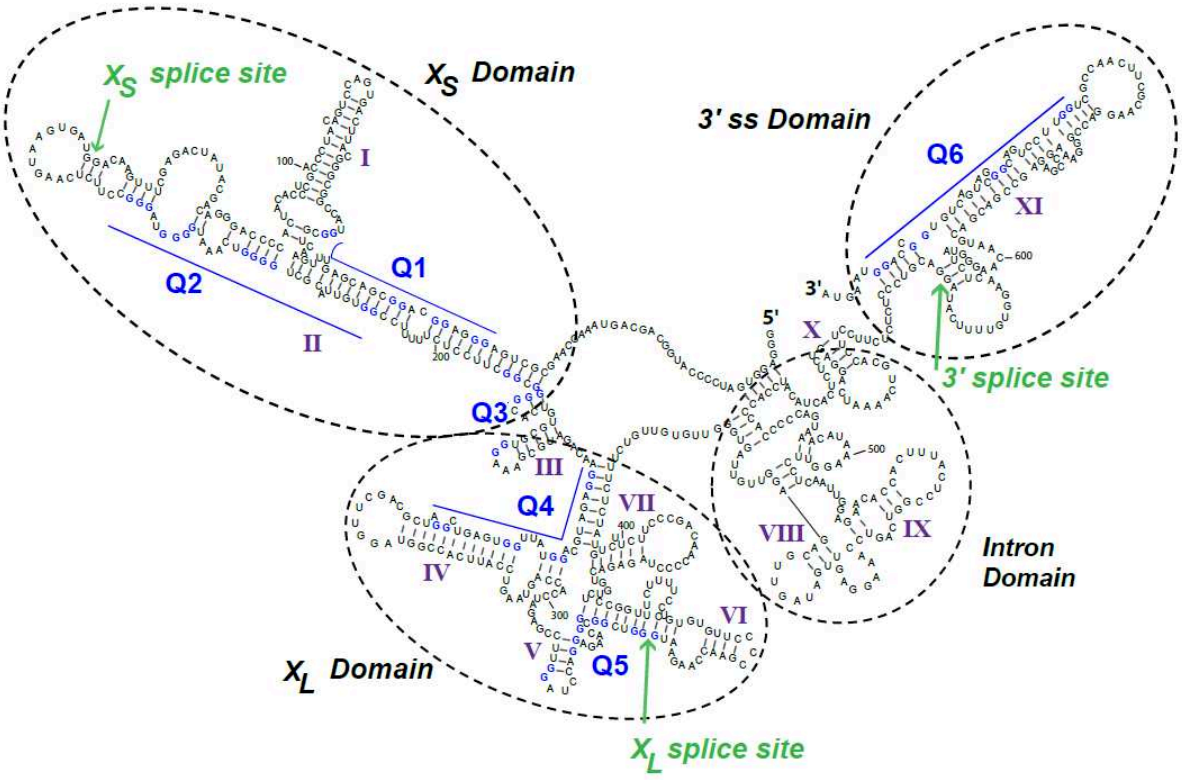


Figure 10 : modèle des structures secondaires présentes dans l'ARNm du minigène Bcl-x-681 d'après Weldon *et al* 2017.

Modèle généré par Mfold avec les contraintes générées par une expérience de footprinting. Les différents sites d'épissages sont indiqués en vert. Les différentes tiges-boucles sont désignés de I à XI. Les séquences des G-quadruplex potentiels prédit par QGRS Mapper sont indiqués de Q1 à Q6 en bleu. Les domaines structurellement indépendants sont entourés et comprend les domaines XS, XL, Intron et 3'SS.

6. Le facteur d'épissage RBM25 :

a. Structure de RBM25

RBM25 est une protéine appartenant à la famille des protéines RBP (RNA Binding Protein) et à la sous-famille des protéines RBM (*RNA Binding Motif*). Les protéines RBM sont des protéines intracellulaires essentielles, impliquées dans de nombreux processus importants pour les ARN tels que l'épissage, le transport, la localisation et la traduction. Cette sous-famille comprend généralement au moins un motif RRM (RNA Recognition Motif) nécessaire pour l'interaction avec les ARN. A ce jour, plus de 50 RBM protéines ont été identifiées.

RBM25, anciennement RED120 chez l'Homme ou Snu71 chez la levure, est une protéine très conservée au cours de l'évolution et est ainsi retrouvée dans de nombreux modèles biologiques tels que les plantes, les levures et les mammifères (Ester & Uetz, 2008; Fortes et al., 2007; X. Li, 2017).

RBM25 a été pour la première fois identifiée en 2003 par la présence de son domaine PWI du côté C-terminal, semblable à celui de la protéine SRm160 (Szymczyna et al., 2003). Ce domaine particulier avait été identifié comme un domaine nécessaire pour l'interaction avec les acides nucléiques et pour le traitement des ARN. Suite à l'identification de ce domaine, RBM25 est classée comme protéine capable de réguler le devenir des ARN. En plus de ce domaine PWI, RBM25 contient un domaine RRM, un domaine riche en arginine et acide glutamique (RE) central, et un domaine riche en proline du côté N-terminal (figure 11).

La localisation de RBM25 est principalement nucléaire avec une localisation subcellulaire au niveau des *speckles* qui est régulée par son domaine RE (A. Zhou et al., 2008). Les *speckles*, sont des compartiments sans membrane du noyau. Ils sont riches en composants du splicéosome et en facteurs d'épissage. Ces zones du noyau servent de zones de stockage ou d'assemblage pour fournir aux sites actifs de transcription les facteurs d'épissage nécessaires à la formation du splicéosome. Ainsi, les protéines localisées dans les *speckles* sont généralement associées avec une fonction dans la régulation de l'épissage. De par sa localisation dans les *speckles* et les domaines la constituant, RBM25 a rapidement été assimilée à une protéine de l'épissage alors que sa fonction exacte n'est toujours pas clairement définie.

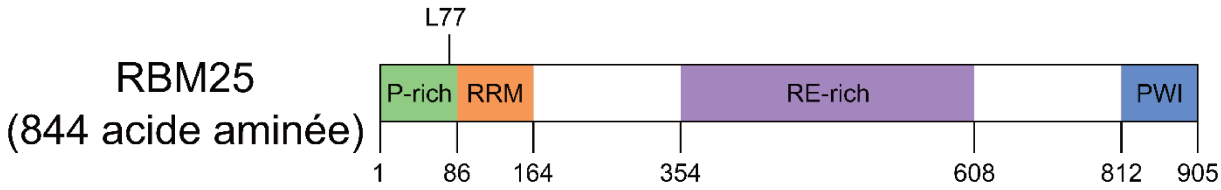


Figure 11 : organisation fonctionnelle de la protéine RBM25

Le facteur d'épissage RBM25 est constitué d'un domaine riche en proline (P-rich) situé à l'extrémité N-terminale avec un domaine de reconnaissance des ARN (RRM) immédiatement adjacent. Le domaine PWI (proline-tryptophan-isoleucine-rich) est localisé à l'extrémité C-terminale et est également un domaine de reconnaissance des acides nucléiques. Un domaine riche en arginine et acide glutamique (RE-rich) est retrouvé au centre de RBM25. La lysine en position 77 est essentielle pour l'interaction avec le facteur d'épissage SRSF2. Lorsque cette lysine est méthylée, RBM25 perd sa capacité à interagir avec SRSF2.

b. Fonctions de RBM25 :

RBM25 est principalement impliquée dans le processus d'épissage. De manière générale, il semble que RBM25 soit un facteur d'épissage permettant l'inclusion d'exon alternatif (Carlson et al., 2017). De façon plus précise, le mécanisme de régulation de l'épissage par RBM25 a été décrit pour différents gènes.

Il a ainsi été montré que RBM25 régule l'épissage alternatif de *BCL-x* en favorisant la sélection du site 5' faible de *BCL-x* conduisant à la forme alternative pro-apoptotique Bcl-xS. La liaison de RBM25 à proximité du site 5' faible d'épissage activerait ce site d'épissage par interaction avec la protéine Luc7A. Cette protéine a été montrée comme capable de se lier directement avec la sous-unité U1 snRNP, qui est nécessaire pour la reconnaissance du site d'épissage à la fois en levure et en cellule humaine. Il semblerait que RBM25 puisse s'associer avec Luc7A par le biais de son domaine RE qui, semblable aux domaines RS, permet les interactions protéine-protéines. De plus, la partie C-terminale de Luc7A est constituée de répétitions RE et RS permettant l'interaction avec RBM25 de manière ARN-dépendante. L'interaction de RBM25 avec le site d'épissage alternatif 5' de *BCL-x* est essentielle à la synthèse de Bcl-xS ainsi, la surexpression de RBM25 est associée à une augmentation de la forme pro-apoptotique Bcl-xS, tandis que la sous-expression est associée avec une augmentation de la forme anti-apoptotique Bcl-xL (A. Zhou et al., 2008). Par cette implication dans la régulation de Bcl-xS, RBM25 est considéré comme un gène

suppresseur de tumeurs. La protéine RBM25 est également impliquée dans l'épissage de BIN1. BIN1 est un gène suppresseur de tumeur produisant deux isoformes. L'isoforme canonique Bin1 est une protéine cytoplasmique capable d'interagir avec une région conservée de c-Myc, un oncogène et facteur important pour la prolifération cellulaire afin d'inhiber son activité. La seconde isoforme est produite suite à l'inclusion de l'exon 12 et conduit à l'ajout d'un domaine à la protéine qui bloque la capacité d'interagir avec c-Myc. En absence d'expression de RBM25, l'expression du variant avec l'exon 12 inclus est fortement augmentée. RBM25 est donc capable de réguler l'exclusion de l'exon 12 dans le pré-ARNm de BIN1 et donc d'inhiber la synthèse des dominants négatif de BIN1 (Ge et al., 2019). La régulation de MYC par RBM25 est donc une des fonctions principales de RBM25 et permet de maintenir indirectement l'activité de MYC dans les cellules normales. L'effet suppresseur de tumeur de RBM25 est amplifié par le rôle de RBM25 dans l'épissage alternatif de *BCL-x*. Ceci conduit à un mécanisme de sécurité pour maintenir la cellule dans un état normal. Ainsi, cette fonction double de RBM25 permet l'élimination des cellules malignes. RBM25 est également impliquée dans l'épissage du gène du canal cardiaque au sodium en se liant au pré-ARNm de ce dernier, et permet la sélection du site cryptique d'épissage conduisant à une version non fonctionnelle de la protéine (Noyes et al., 2017).

RBM25 peut être soumise à des modifications post-traductionnel qui vont impacter sa capacité à interagir avec d'autres protéines. La méthylation au niveau de la lysine 77 de RBM25 semble être essentielle pour l'interaction avec d'autres facteurs. Cette mono-méthylation de la lysine située dans la région riche en proline de RBM25 permet l'interaction avec le domaine RS de SRSF2. Lorsque cette lysine est méthylée, RBM25 perd la capacité d'interagir avec SRSF2. Cette modification permet ainsi une régulation du rôle de RBM25 dans l'épissage en modulant son interaction avec SRSF2, une protéine cruciale dans la définition d'exon (Carlson et al., 2017).

c. Les pathologies associées à RBM25

Des dérèglements dans l'épissage sont associés avec une susceptibilité aux cancers, aux maladies cardiaques, à des maladies neurodégénératives ainsi qu'à des maladies auto-immunes. Aussi, par son rôle essentiel dans le contrôle de l'épissage et de

l'apoptose, RBM25 est un gène dont la dérégulation est associée au développement de cancers ainsi qu'à un mauvais prognostique pour ces derniers. RBM25 est sous-exprimée dans le carcinome hépatocellulaire, le cancer de la prostate et le cancer colorectal (Guo et al., 2021). Dans les cellules tumorales des leucémies aigus myéloïdes (LAM), la sous-expression de RBM25 est associée à une augmentation de la prolifération des cellules. Inversement, maintenir RBM25 à un niveau physiologique pourrait prévenir le développement de leucémies (Bramlett, 2023).

Le gène *RBM25* est une des cibles du facteur de transcription p53 qui favorise l'expression de *RBM25*. Ce facteur de transcription est retrouvé muté ou manquant dans 50 % des cancers, aussi, l'absence de ce facteur important entraine une diminution de l'expression de RBM25. Cette diminution de l'expression de *RBM25* est associée à une évolution vers des métastases dans les cancers de la prostate et les cancers colorectaux. De plus, dans ces cancers, la diminution de RBM25 est associée avec un épissage aberrant du gène AMOTLT1. En l'absence d'épissage par RBM25, AMOTLT1 produit une version circulaire de son ARN, circAMOTLT1, qui est reconnue par les miRNA-193a-5p entrainant la transition épithélio-mésenchimateuse (EMT) et la progression métastatique dans le cas du cancer de la prostate. L'EMT est un mécanisme qui permet la transition de cellules épithéliales vers des cellules avec un phénotype mésenchymateux défini par des marqueurs spécifiques tels que l'E-cadhérine et la vimentine. Ce mécanisme est essentiel pendant le développement embryonnaire et pendant la régénération des tissus à l'âge adulte. L'EMT peut être réactivé dans certains cancers où il participe à la croissance des tumeurs, leur invasion et l'activité des métastases.

De façon surprenante, RBM25 peut également être retrouvée surexprimée dans certains cancers tels que le carcinome hépatocellulaire (Yang et al., 2019). Bien que le mécanisme d'action ne soit pas encore compris, il semble que RBM25 interagisse de façon excessive avec les gènes CDCA5 et INCENP qui sont impliqués dans le contrôle de la division cellulaire.

Enfin, dans le cas d'ischémies cérébrales, la transcription de RBM25 est augmentée par la liaison d'EGR1 au promoteur de RBM25. Cette surexpression de RBM25 dans la cellule neuronale conduit à une augmentation du volume de l'ischémie ainsi qu'à une augmentation de l'apoptose des cellules. Ce mécanisme est renforcé par la surexpression et l'interaction du lncRNA NEAT1 avec EGR1. Lorsque NEAT1 se lie à

EGR1, ce lncRNA empêche l'ubiquitinylation et donc la dégradation de EGR1. L'accumulation de EGR1 facilite alors la transcription, l'expression de RBM25 et, *in fine*, l'apoptose des cellules (Cao et al., 2022).

7. Les G-quadruplex

a. Définition

Dans les années 1960, des chercheurs ont montré que les guanines pouvaient s'assembler entre elles afin de former des structures secondaires particulières, les G-quadruplex (G4) (Gellert et al., 1962). Il fallut attendre les années 1990 pour que le premier G4 d'ADN (dG4) soit identifié dans la séquence des télomères (F. M. Chen, 1992; Y. Wang & Patel, 1993). Depuis, les G4 font l'objet de nombreuses études et ces structures particulières sont maintenant reconnues comme étant impliquées dans la régulation de nombreux processus cellulaires. Les G4 sont qualifiés de structures secondaires non canoniques. En effet, les G4 impliquent des liaisons hydrogène de type Hoogsteen à la différence des interactions canoniques de type Waston-Crick qui impliquent des liaisons hydrogènes éponymes. Les liaisons de types Hoogsteen sont caractérisées par l'utilisation des N7 et C6 pour la formation des liaisons hydrogènes entre guanines. Ce type de liaison permet la formation de duplex, et plus fréquemment de triplex et de quadruplex. La formation de quadruplex est unique à l'association de guanines entre elles ce qui a donné le nom de G-quadruplex (figure 12). Les G4 sont composés de l'empilement de G-quartets, l'unité de base de G4. Ces G-quartets sont des structures planes assemblées par la liaison de quatre résidus de guanine entres elles. Dans les G-quartets, chaque guanine sert à la fois de donneur et d'accepteur de deux liaisons hydrogène de type Hoogsteen. Les G-quartets sont ensuite stabilisés par des cations monovalents tels que le sodium (Na^+) ou le potassium (K^+), ce dernier étant le cation le plus efficace pour stabiliser les G-quartet et les G4 (Hardin et al., 1992). De façon intéressante, le potassium est également l'ion métallique le plus présent dans les cellules de mammifères dans lesquelles sa concentration est de l'ordre de 100 mM (soit la concentration optimale pour stabiliser les G4), ce qui permet la formation des G4 dans les conditions physiologiques.

Les G4 sont des structures polymorphiques qui peuvent adopter de nombreuses conformations. Les liaisons entre les guanines peuvent être intra- ou intermoléculaires. Les G4 intramoléculaires nécessitent la présence de quatre guanines sur la même molécule d'acide nucléique tandis que les G4 intermoléculaires peuvent impliquer entre deux et quatre molécules différentes d'acides nucléiques. Tandis que les G4 intramoléculaires (par exemple dans les ARN) adoptent uniquement une conformation parallèle, dans laquelle les quatre molécules d'acides nucléiques sont orientées dans la même direction, les G4 intermoléculaires peuvent être parallèles ou antiparallèles. La diversité des G4 est accrue par la possibilité de former des boucles qui peuvent être de taille et de séquence extrêmement variables. Ces boucles sont constituées d'acides nucléiques qui séparent les répétitions de G nécessaires à la formation des G-quartet. Ces boucles influent sur la structure et la stabilité des G4.

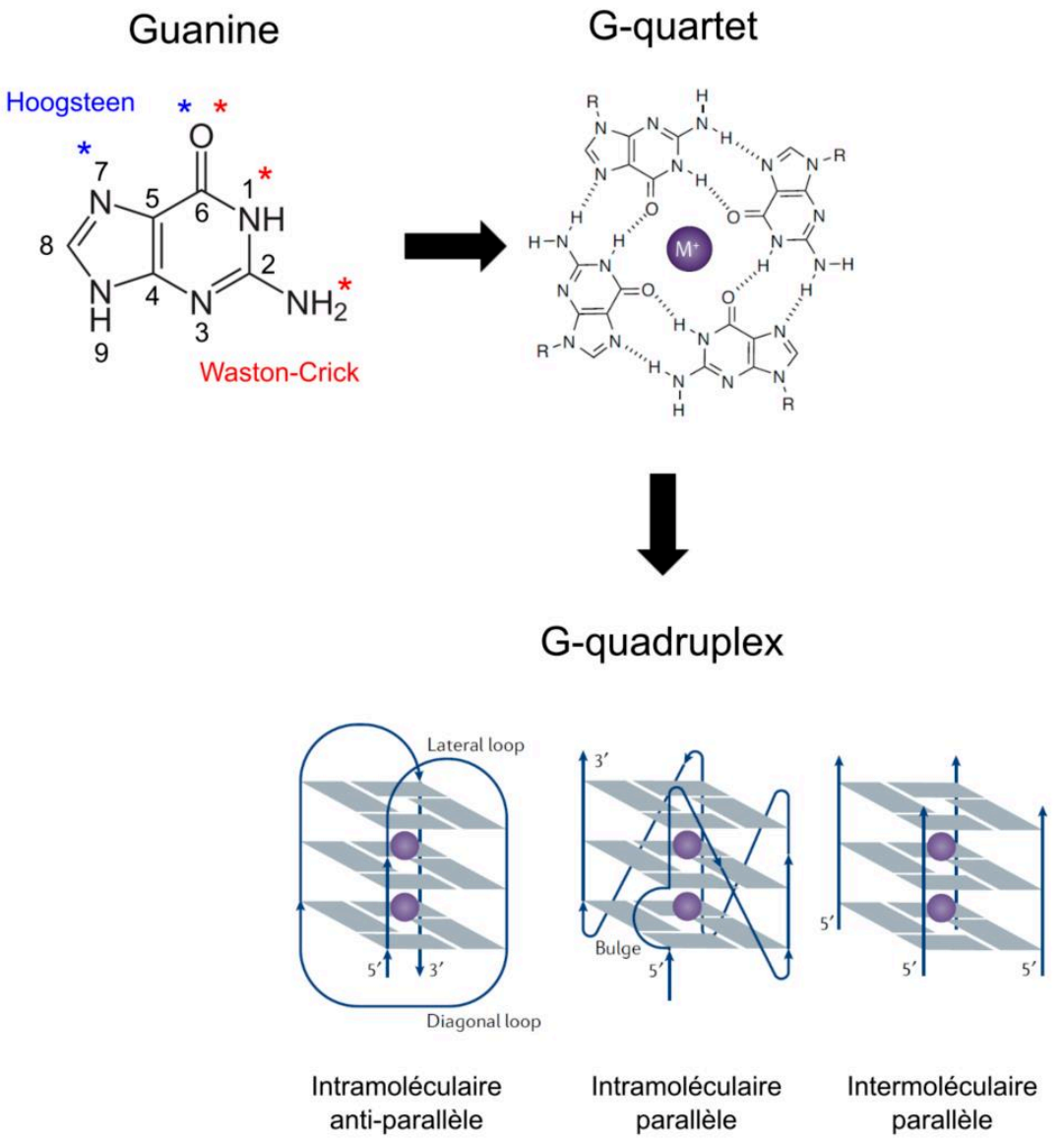


Figure 12 : structure des G-quadruplex, adapté de Vashney *et al* 2020.

Les liaisons Hoogsteen impliquent les positions C6 et N7 de la purine tandis que les liaisons Waston-Crick impliquent les positions N1 et C2 et C6. L'association de quatre guanines entre-elles par des liaisons de type Hoogsteen forme un G-quartet, l'unité de base du G-quadruplex. Le G-quartet est stabilisé par un cation, généralement le potassium (K^+). L'empilement d'au moins deux G-quartet permet la formation d'un G-quadruplex (G4), une structure secondaire capable de se former dans les acides nucléiques riches en guanine. En fonctions du nombre de molécules d'acide nucléique impliqués dans la formation du G4, les G4 peuvent être intra- ou inter-moléculaire. Enfin, en fonction de l'orientation des molécules d'acides nucléiques, les G4 peuvent être parallèles ou anti-parallèles.

b. Outils pour identifier et caractériser les G-quadruplex

Durant ces dernières décennies, la recherche sur les G4 en biologie s'est grandement développée, aboutissant à l'identification du rôle régulateur de nombreux G4 dans différents processus biologiques. En parallèle, différents outils ont été développés pour identifier et caractériser ces structures. Une des premières approches pour l'identification de G4 est généralement l'utilisation de logiciels prédictifs, tels que QGRS Mapper et G4 Hunter (Brázda et al., 2019; Frees et al., 2014). Ces logiciels sont capables de prédire la propension de la séquence analysée à former un G4. Cependant, ces logiciels ne sont pas capables d'identifier tous les G4. De plus, les G4 identifiés ne sont que potentiels. C'est pourquoi cette méthode doit être associée avec d'autres approches pour confirmer la formation d'un G4. Pour cela, de nombreuses techniques biophysiques peuvent être utilisées telles que le dichroïsme circulaire, la spectroscopie RMN ou encore des expériences d'*UV melting* (Luo et al., 2022; Mergny & Lacroix, 2009; Strahan et al., 1998). Ces techniques permettent, à partir de courts oligonucléotides, de déterminer si une séquence peut former un G4. Cette approche biophysique permet d'avoir une forte indication quant à la capacité d'une séquence d'ADN ou d'ARN à former un G4. Cependant, ces techniques sont limitées à une approche *in vitro* et ne permettent donc pas d'affirmer la formation d'un G4 dans un contexte physiologique.

Les ligands de G4, des petites molécules présentant une forte affinité pour les G4, représentent des outils précieux pour l'étude des G4 et peuvent être utilisés dans des tests *in vitro* et *in vivo*. Ces molécules peuvent stabiliser, ou déstabiliser, des G4 et, dans des tests in *vivo*, permettent ainsi de réaliser des études fonctionnelles. Ces ligands de G4 sont notamment utilisés dans des études sur le rôle régulateur potentiel des G4 dans les promoteurs. Stabiliser ces G4 avec des ligands accentue la capacité inhibitrice de la transcription des G4 confirmant le rôle du G4 comme inhibiteur de la transcription (Dickerhoff et al., 2021). Cependant, ces ligands possèdent généralement un spectre large et sont donc capables de stabiliser de nombreux G4. Les effets observés peuvent donc être le résultat de *off-target.* De plus, ces ligands peuvent aussi stabiliser la formation de G4 qui ne se formeraient pas dans des conditions physiologiques. Enfin, certains seraient capables de déstabiliser les G4 (Mitteaux J… Granzhan A et al *JACS* 2021 -référence 72 de ton papier *NAR*).

Le développement d'anticorps présentant de l'affinité pour les G4 a permis l'observation de G4 dans des cellules (Biffi et al., 2013). Cet outil peut être combiné avec d'autres techniques telles que des tests de retards sur gel, des ELISA ou des tests de *Proximity Ligation Assay* (PLA) afin d'étudier des G4. Plus récemment, la technique de *Footprinting Of Long 7-Deazaguanine-substituted RNAs* (FOLDeR) a été développée afin d'identifier la formation de G4 dans des longues séquences d'ARN (Weldon et al., 2017). En remplaçant les guanines par des 7-déazaguanine, il est possible bloquer la formation des G4 sans empêcher la formation d'autres structures secondaires, ainsi en comparant deux oligonucléotides, un modifié et l'autre non, il est possible de déterminer de la capacité d'une séquence à former, ou non, des G4 dans des conditions physiologiques.

c. Les G-quadruplex d'ADN (dG4)

En 1987, un premier dG4 est identifié *in vitro* dans la séquence ADN des télomères (Henderson et al., 1987). Les télomères sont des complexes ADN-protéines localisés à l'extrémité des chromosomes eucaryotes. Ils sont caractérisés par des répétitions en tandem de séquences riches en guanines, généralement GGGTTA. Les télomères sont des éléments essentiels aux chromosomes eucaryotes en les protégeant des mécanismes de réparation aberrants et en prévenant la fusion des chromosomes, ainsi que leur dégradation. A chaque cycle de réplication de l'ADN, environ 50 à 200 pb d'ADN télomérique sont perdues (Makarov et al., 1997). De la sorte, les télomères raccourcissent avec l'âge (le nombre division) des cellules jusqu'à devenir trop courts, ce qui entraîne la sénescence (le vieillissement) cellulaire. Les télomères sont synthétisés par la télomérase, un complexe ribonucléoprotéique avec une activité de rétro-transcription capable d'étendre le télomère. L'activité de cette télomérase est surexprimée dans les cancers où son activité ainsi augmentée permet la maintenance des télomères et une division ininterrompue des cellules (N. W. Kim et al., 1994). La richesse en guanines des séquences télomérique est propice à la formation de dG4. Lorsque des dG4 se forment dans les séquences télomériques, leur structure est capable de bloquer la liaison de la télomérase. Ainsi, la stabilisation des G4 dans les télomères est associée à une diminution de la taille des télomères et donc à la senescence des cellules. Pour cette raison, des stratégies anticancéreuses visant à

stabiliser les G4 sont désormais en développement (Kosiol et al., 2021; Nakanishi & Seimiya, 2020; Oganesian & Bryan, 2007). Elles visent à stabiliser les G4 des télomères, mais également les G4 présents dans les promoteurs d'oncogènes de façon à interférer avec leur expression. En effet, plus de 40% des promoteurs de gènes humains contiendraient au moins un motif capable de former des G4 (Huppert & Balasubramanian, 2007; Varshney et al., 2020). Les G4 présents dans les promoteurs peuvent être associés avec diminution ou une augmentation de la transcription. Dans le cas de l'oncogène *c-MYC,* la transcription est régulée par un élément NHE III (*Nuclease Hypersensitivity Element*) localisé dans le promoteur P1 qui est responsable de 80% de la transcription du gène (González & Hurley, 2010). L'élément NHE III sert de point de reconnaissance pour les facteurs de transcription SP1, CNBP et hnRNP K afin d'activer la transcription de *c-MYC*. Cet élément contient une région riche en G capable de former un G4 qui régule négativement la transcription. En se formant, le G4 bloque la liaison de ces facteurs et entraine ainsi une diminution de la transcription de *c-MYC* (Mathad et al., 2011). A l'inverse, empêcher la formation du G4, par exemple par des mutations substituant les guanines en adénines, conduit à une très forte augmentation de l'activité du promoteur de *c-MYC*. Ce mécanisme d'inhibition de la transcription par les G4 est retrouvé dans d'autres gènes liés aux cancers tels que *BCL2, VEGF, KIT, KRAS* et *RET.* Enfin, dans d'autres cas, les G4 peuvent, au contraire, augmenter la transcription des gènes en servant de plateforme de reconnaissance pour des facteurs de transcription. Ainsi, la déstabilisation du G4 dans le promoteur du gène suppresseur de tumeur *BAP1* par mutagénèse dirigée entraine-t-elle une diminution de sa transcription. Ce G4 régule donc de manière positive l'expression de ce gène (Y. Li et al., 2018).

d. Les G-quadruplex d'ARN (rG4)

Les G4 d'ARN (rG4), bien que structurellement similaires aux G4 d'ADN (dG4), présentent des différences notables par rapport à ceux-ci et sont généralement plus stables que les dG4 (Joachimi et al., 2009). A la différence de l'ADN, l'ARN est simple brin, aussi l'ARN adopte de nombreuses structures secondaires telles que des épingles à cheveux (*hairpins*), des tiges-boucles *(stem-loops),* des *bulges* et des pseudonoeuds *(pseudoknots)*. Dans les ARN, le groupe 2' hydroxyle du sucre ribose

a plusieurs effets sur la stabilité du G4. Ce groupement, permet plus d'interactions dans les rG4 ce qui conduit à une augmentation de la stabilité des rG4. En effet, il a été montré que les rG4 ont une plus grande stabilité thermodynamique et thermique que leur version dG4. De plus, le groupement 2'-OH supplémentaire dans l'anneau ribose organise les liaisons hydrogènes et ainsi favorise une topologie parallèle stable dans les rG4 (Collie et al., 2010; Fay et al., 2017). Intrinsèquement, les rG4 présentent une spécificité des cations stabilisateurs différente de celle des dG4. Ainsi, des études sur les rG4 *TERRA* et *NRAS* ainsi que leur version dG4 ont montrées que si les deux versions de G4 sont stabilisées par le potassium (K^+), seuls les dG4 sont stabilisés par le sodium (Na^+) (Guiset Miserachs et al., 2016). Les rG4 représentent des structures secondaires importantes dans la biologie des ARN et ont été impliquées dans certaines maladies. Ainsi, les rG4 sont impliqués dans l'épissage, l'export, la localisation et la traduction des ARN.

Suite à la transcription, l'ARN nouvellement synthétisé peut s'apparier avec l'ADN complémentaire pour former des hybrides ARN:ADN. Ces structures sont connues sous le nom de *R-loops*, et si deux molécules d'acides nucléiques contiennent des guanines, elles peuvent aussi former des G4 intermoléculaires à base d'ARN et d'ADN (Miglietta et al., 2020). Ces G4 hybrides inhibent la transcription et représentent des éléments *-cis* régulateurs intégrés aux gènes qui sont activés de manière co-transcriptionnelle. De plus, ces G4 sont impliqués dans l'inhibition de la terminaison de la transcription et bloquent par encombrement stérique l'ARN polymérase II.

D'après une étude à grande échelle du génome, plusieurs centaines de milliers de G4 potentiels pourraient se former dans le génome humain (Tu et al., 2021). Les G4 localisés dans le pré-ARNm peuvent impacter l'épissage en modifiant les liaisons des RBP aux sites d'épissages. Plusieurs mécanismes de régulation ont ainsi été observés. Les rG4 peuvent soit bloquer de manière stérique les éléments régulateurs, soit interférer avec le recrutement de certaines protéines et ainsi inhiber l'épissage. A l'inverse, les rG4 peuvent aussi recruter des protéines régulatrices de l'épissage et favoriser le recrutement de la machinerie d'épissage. Des rG4 présents dans les pré-ARNm des gènes *TP53, BCL-2, BCL-x, FMRP1, NRAS*, ont été montrés comme étant impliqués dans l'épissage de ces gènes (Didiot et al., 2008; Dumas et al., 2021; Marcel et al., 2011; Varshney et al., 2020; Weldon et al., 2018). Cependant, peu d'études font état du mécanisme exact par lequel ces rG4 régulent l'épissage de ces gènes.

L'épissage alternatif du gène *CD44* est régulé par un rG4 au niveau de l'un de ses introns. Le gène *CD44* est régulé par un évènement d'épissage alternatif qui aboutit à deux isoformes. La protéine canonique CD44v, est une molécule de surface exprimée dans les cellules épithéliales. La forme alternative, CD44s conduit à un changement de phénotype des cellules de l'état épithélial à un vers un état mésenchymateux (transition épithélio-mésenchymateuse ou *EMT* en anglais). CD44 est régulée pendant l'EMT, un programme de développement qui est anormalement activé pendant la métastase des tumeurs. L'épissage de *CD44* est régulé par un élément cis-régulateur I-8 capable de former un rG4 qui stimule l'épissage alternatif (Huang et al., 2017). Les auteurs de cette étude ont montré que le facteur d'épissage hnRNP F est capable de potentialiser l'épissage alternatif de *CD44*. Sa surexpression est associée à la synthèse de CD44v et une inhibition de l'EMT. Aussi les auteurs de cette étude proposent-ils que hnRNP F viendrait stabiliser le rG4 situé dans l'intron de CD44 afin de favoriser l'épissage alternatif de ce gène. Cependant, cette hypothèse est en désaccord avec autre modèle de régulation des rG4, le mécanisme de liaison-dépliement-blocage (*bind-unfold-lock*) proposé par le groupe de Stéfania Millevoi (Dumas et al., 2021; Herviou et al., 2020). Dans ce modèle, les rG4 peuvent être reconnus par des hélicases d'ARN telles que DHX36, qui vont dérouler le rG4. Une fois le brin d'ARN linéarisé, la protéine de liaison hnRNP F vient se fixer au niveau de la séquence capable de former le G4 afin de la maintenir dans sa conformation non structurée. Dans ce mécanisme, les hélicases apparaissent comme des régulateurs importants de rG4 et ce mécanisme a été observé dans la traduction de certains ARN (Lyu et al., 2021; Millevoi et al., 2012).

L'intérêt de la communauté des scientifiques pour les rG4 est en pleine essor, cependant, les mécanismes de régulation précis par lesquels ces structures secondaires non-canoniques des ARN exercent leur rôle régulateur sont encore mal compris. Aussi, des études plus approfondies sur l'implication des rG4 dans l'épissage sont-elles tout à fait pertinentes, notamment pour pouvoir cibler ces rG4 à des fins thérapeutiques.

e. Les protéines de liaisons aux G-quadruplex (G4BP)

Les G4 sont impliqués dans de nombreux processus physiologiques tels que la régulation de la transcription, de l'épissage et de la traduction. Pour réguler ces processus, les G4 sont souvent associés à des G4BP, des protéines présentant de l'affinité pour eux. En interagissant avec les G4, ces protéines peuvent moduler la stabilité et la conformation des G4. Ces protéines peuvent également reconnaitre les G4 et s'en servir comme plateforme de recrutement pour d'autres protéines régulatrices. Ainsi, les G4BP peuvent être catégorisés en trois groupes, les protéines stabilisatrices des G4, les protéines déstabilisatrices des G4, et enfin les protéines de liaisons aux G4 (Shu et al., 2022).

Dans les télomères, les dG4 qui se forment bloquent par encombrement stérique la télomérase. Les protéines hnRNP A1 et UP1 ont été identifiées comme capables d'interagir avec ces dG4 qui se forment dans les télomères (Ghosh & Singh, 2018). Ces protéines, une fois liées au G4, vont déplier ces structures, ce qui permet à la télomérase de reprendre son activité et ainsi d'allonger les télomères. Des études *in vitro* ont montré que la protéine UP1 présente une très forte affinité pour des séquences linéaires et qu'elle peut linéariser les séquences capables de former des dG4. Cependant, le mécanisme précis par lequel UP1 régule la stabilité des dG4 n'est pas encore connu. Les protéines hnRNP sont en général associées à une capacité de déstabilisation des G4. Ainsi, certaines hnRNP, sont-elles capables de déplier les dG4 qui se forment dans le promoteur des oncogènes *KRAS* et *TRA2ß* (Paramasivam et al., 2009). Les dG4 localisés dans le promoteur de ces gènes agissent comme des inhibiteurs de la transcription. L'expression des protéines hnRNP est associée à un dépliement des G4 des promoteurs ainsi qu'à une reprise de la transcription de ces gènes.

Les hélicases de la famille des *DEAD box* sont des régulateurs importants des G4. Ces protéines très conservées sont impliquées dans de nombreux processus métaboliques tels que l'épissage. Les hélicases sont capables de dérouler aussi bien les G4 d'ADN que d'ARN (Mendoza et al., 2016). Elles sont notamment essentielles pour maintenir l'intégrité du génome en agissant comme des protéines capables de résoudre des structures secondaires telles que les R-loops et les G4. L'hélicase DDX5 est un régulateur transcriptionnel de *c-MYC* dont une des fonctions est de déplier la

séquence capable de former un dG4 dans le promoteur de *c-MYC* (Wu et al., 2019). De façon intéressante, ce mécanisme est conservé chez la levure, où l'orthologue Dbp2 est également capable de déplier aussi bien les G4 d'ADN que d'ARN (Yan et al., 2021). Différents motifs de liaisons aux G4 ont été identifiés dans les G4BP. Un des motifs de liaison aux G4 les mieux décrits aujourd'hui est le motif RGG qui est composé de répétitions du tri-peptide arginine-glycine-glycine. Les arginines sont des acides aminés chargé positivement connus pour établir des interactions électrostatiques fortes avec le squelette phosphodiester des acides nucléiques ainsi que de faciliter les liaisons hydrogènes. Le motif RGG conduit à des régions intrinsèquement désordonnées qui sont impliqués dans des interactions avec des acides nucléiques et également d'autres protéines.

Le domaine RGG est un des domaines les plus conservés au cours de l'évolution et il est retrouvé dans des protéines responsables de la maintenance des télomères, de la transcription ou encore de la traduction de régions d'ARNm contenant des rG4. Outre les hélicases qui possèdent un domaine RGG, la NCL fut la première protéine avec un motif RGG identifiée (Thandapani et al., 2013). Cette protéine est localisée essentiellement dans le nucléole où elle est impliquée dans la maturation des ARN ribosomiques (ARNr) et la synthèse du ribosome. Elle est également impliquée dans la transcription et l'apoptose. La NCL est une G4BP qui interagit avec le dG4 qui se forme dans le promoteur du gène *c-MYC*. La NCL agit comme une protéine stabilisatrice de ce dG4, ce qui entraine une diminution de la transcription de *c-MYC* (González et al., 2009). La NCL est aussi capable d'interagir avec les G4 localisés dans les promoteurs de *BCL2, hTERT, VEGF, RET, PDFGF-1* et de *c-kit* afin de réguler leur expression. Malgré de nombreuses études faisant état de l'interaction entre des G4 et des G4BP, le rôle fonctionnel exact de ces interactions G4/protéines, ainsi que leur effet sur les rG4 eux-mêmes, restent le plus souvent relativement flous, voire indéterminés.

f. Les ligands de G4 comme pistes thérapeutiques potentielles

Avec l'identification de G4 dans les télomères et dans les promoteurs d'oncogènes, les G4 sont rapidement apparus comme des cibles thérapeutiques potentielles pour le

traitement de certains cancers. Ainsi, l'utilisation de ligands de G4 capables de stabiliser ou déstabiliser ces structures représente aujourd'hui un espoir pour traiter les cancers. Plus de 900 ligands de G4 avec des structures variées ont aujourd'hui été développés, tels que les ligands de la famille des porphyrines (TMPyP4), des dérivés de phénanthroline (PhenDC3) ou encore des dérivés de pyridostatine (PDS). Ces petites molécules présentent une spécificité pour les G4, et parfois pour l'un ou l'autre type de conformation du G4. De la sorte, selon que le G4 adopte une structure parallèle ou antiparallèle, la liaison du ligand de G4 peut être différente. L'effet stabilisateur des ligands de G4 dépend aussi du mode de liaison du ligand au G4, certains ligands de G4 établissent des interactions Π sur les quartets terminaux du G4, d'autres avec les boucles, d'autres encore avec les sillons du G4, ou bien encore utilisent une combinaison de ces différents modes. Les G4 étant des structures très dynamiques, pouvoir stabiliser ces structures secondaires dans leur conformation permet d'accentuer leurs effets. Cette propriété des ligands de G4 a permis de développer plusieurs stratégies anticancéreuses basées sur la stabilisation de G4 localisés dans les télomères, dans les promoteurs d'oncogènes ou pour encore pour induire une instabilité génomique.

Les ligands de G4 ont été envisagés dans le traitement de certains cancers, notamment du pancréas. Le traitement de cellules par des ligands de G4 réduit la viabilité des cellules cancéreuses et inhibe la croissance de ces cellules. Le ligand TMPyP4 induit la mort de cellules cancéreuses du pancréas MIA Pa-Ca-2 mais n'a pas d'effet sur les cellules normales. De plus, le ligand MM41 réduit de manière significative la croissance des tumeurs dans des xénogreffes de MIA Pa-Ca-2 (Ohnmacht et al., 2015). Ces ligands agiraient sur la stabilité des G4 présent dans les promoteurs des gènes *KRAS* et *BCL-2* (Pattanayak et al., 2018). De plus, le traitement de ces cellules par des ligands de G4 est associé à une augmentation de l'expression des protéines pro-apoptotiques Bax et p53 ainsi qu'à une augmentation de l'apoptose. Bien que l'utilisation des ligands de G4 semble prometteuse, ces molécules présentent des limites. En effet, les molécules utilisées présentent généralement une capacité de stabilisation de nombreux G4, tant d'ADN que d'ARN ce qui pose question quant à leur spécificité d'action. En conséquence, il est difficile d'identifier la cause exacte des effets observés par les ligands de G4. Ainsi, bien que l'effet des ligands de G4 semblent être prometteuse dans des thérapies anticancéreuses, leur mécanisme

d’action n’est pas encore bien compris et les effets peuvent généralement être attribués à la stabilisation des G4 dans les télomères, de ceux des promoteurs d’oncogènes ou encore à une induction de l’instabilité génomique. Ainsi, développer des molécules capables de lier un G4 spécifique permettrait de développer des thérapies plus précises et contrôlées. Enfin, il semble pertinent d’identifier les protéines capables d’interagir avec les G4. Ces protéines sont essentielles pour comprendre le rôle dans G4 dans les cellules et peuvent elles-mêmes représenter des cibles ou des moyens de réguler des G4 spécifiquement en ciblant la protéine elle-même ou encore son interaction avec le G4.

Objectifs de la thèse

L'identification de deux rG4 qui se forment à proximité étroite des deux sites alternatif 5' d'épissage dans le pré-ARNm de *BCL-x* soulève de nouvelles questions sur la régulation du gène *BCL-x*. *BCL-x* produit par épissage deux isoformes alternatives aux fonctions antagonistes. Bcl-xL, l'isoforme canonique pleine longueur qui est anti-apoptotique et la forme alternative Bcl-xS qui, elle, est pro-apoptotique (Boise et al., 1993). De sorte, l'épissage alternatif de *BCL-x* contrôle la balance entre la fonction anti-apoptotique canonique de *BCL-x* et sa fonction alternative qui est au contraire d'activer l'apoptose. Dans certains cancers, l'augmentation de Bcl-xL est corrélée à une résistance des cellules à l'apoptose et est associée à une sensibilité moindre aux chimiothérapies. A l'inverse, dans certaines formes de diabète de type II tout comme de pathologies cardiaques, l'isoforme alternative pro-apoptotique Bcl-xS est surexprimée, ce qui engendre une apoptose accrue, par exemple des cellules des ilots de Langerhans dans le cas du diabète. Comprendre la régulation du gène *BCL-x* afin d'exploiter la balance Bcl-xL/Bcl-xS comme cible thérapeutique a donc fait l'objet de nombreuses études qui ont notamment permis le développement des BH3 mimétiques, des molécules capables d'inhiber la fonction anti-apoptotique de Bcl-xL (van Delft et al., 2006). Cependant, les effets secondaires importants de ces molécules font que d'autres pistes doivent être explorées. Pouvoir moduler l'épissage alternatif de *BCL-x* représente donc une piste thérapeutique pertinente. De plus, l'identification de rG4 au niveau des sites d'épissages 5' pose la question de leur rôle potentiel dans la régulation de l'épissage alternatif de *BCL-x* (Weldon et al., 2018). Les G4, font l'objet de recherches poussées ces dernières années et leurs implications dans de nombreux processus biologiques en font des cibles intéressantes pour le traitement de certains cancers. L'utilisation de ligands de G4 pour stabiliser ou déstabiliser certains G4 est aujourd'hui envisagée pour traiter certains cancers, c'est le cas par exemple pour la stabilisation de G4 dans les télomères ou dans les promoteurs d'oncogène (Kosiol et al., 2021). Les G4BP sont des partenaires importants des G4, notamment pour la régulation de l'épissage. Identifier les partenaires spécifiques d'un rG4 qui semble impliqué dans la régulation de l'épissage, comme c'est le cas pour le gène *BCL-x*, permet de mieux comprendre le rôle de ces rG4 dans l'épissage ainsi que de s'en servir pour réguler à façon ce processus. Bien que de nombreux facteurs d'épissages

aient été identifiés dans le pré-ARNm de *BCL-x*, aucun n'a pour le moment été identifié comme interagissant avec les rG4 présents dans son pré-ARNm.

La protéine RBM25 est un facteur d'épissage régulant l'épissage alternatif de plusieurs gènes dont *BCL-x*. En effet, RBM25 favorise la production de l'isoforme pro-apoptotique Bcl-xS en augmentant la sélection du site alternatif 5' d'épissage par la machinerie d'épissage (A. Zhou et al., 2008). De plus, selon des données d'une étude globale récemment publiée sur les potentielles protéines de liaisons aux G4, RBM25 serait aussi une G4BP potentielle (Su et al., 2021). Ainsi, l'objectif de départ de cette thèse était d'étudier un nouveau mécanisme potentiel de régulation de l'épissage alternatif du gène *BCL*-x impliquant les rG4 de son pré-ARNm et de rechercher des facteurs d'épissage présentant de l'affinité pour ces derniers, en premier RBM25 étant donné son implication connue dans l'épissage alternatif de *BCL-x* et sa capacité potentielle à lier des rG4. Si ce mécanisme impliquant RBM25 et les rG4 du gène BCL-x existe, le but dans un second temps sera de l'exploiter, et d'utiliser l'interaction entre les rG4 du pré-ARNm de *BCL-x* et RBM25 comme cible thérapeutique potentielle pour identifier des composés, tels que des ligands de G4, capables de moduler l'épissage alternatif de *BCL-x*. Ceci présenterait un intérêt fort pour améliorer l'efficacité de certaines thérapies anticancéreuses en interférant avec la résistance de certains cancers aux chimiothérapies dont l'efficacité est limitée, en particulier dans les cancers associés à une surexpression de la forme Bcl-xL. De façon intéressante, l'identification de composés qui, à l'inverse, favoriseraient la forme anti-apoptotique Bcl-xL présenterait un intérêt pour les pathologies associées à une surexpression de la forme pro-apoptotique Bcl-xS, telles que le diabète de type II.

Pour mener à bien ce projet, il faudra tout d'abord déterminer si la protéine RBM25 présente de l'affinité pour les rG4 de *BCL-x*, puis, si tel est le cas, déterminer l'effet de la stabilisation des G4 sur l'épissage de *BCL-x*. Une fois cette interaction caractérisée, il sera intéressant de déterminer quel domaine de RBM25 serait responsable de sa liaison potentielle aux rG4 puisque cette protéine présente plusieurs domaines de liaisons aux acides nucléiques (RRM, PWI, RE). Enfin, des ligands de G4 pourront être testés, d'abord comme outils, pour caractériser l'interaction entre RBM25 et le G4, puis pour déterminer l'effet de la stabilisation du G4 sur l'épissage de *BCL-x.* De sorte, des ligands de G4 capables de moduler l'épissage de *BCL-x* en faveur de la forme pro-apoptotique représenteraient des molécules anti-cancéreuses candidates

intéressantes et leur effet potentiel sur la viabilité de cellules cancéreuses pourrait alors être testées.

Résultats

Organisation des résultats :

La première partie des résultats décrit le cœur de mes travaux de thèse qui portent sur la régulation de l'épissage du gène anti-apoptotique *BCL-x* par un rG4 localisé dans son pré-ARNm. Dans le cadre de ce travail, nous montrons que ce rG4, localisé à proximité du site 5' d'épissage alternatif, sert de plateforme de reconnaissance pour le facteur d'épissage RBM25. Cette interaction entre le facteur d'épissage et le rG4 de *BCL-x* est essentielle pour la sélection du site 5' alternatif d'épissage et la synthèse de la protéine alternative pro-apoptotique Bcl-xS. Ces résultats ont fait l'objet d'une publication très récemment (29 août 2023) acceptée dans *Nucleic Acids Research*.

La seconde partie décrit ma participation à un autre projet de l'équipe qui n'est pas directement relié à mon sujet de thèse. Elle est constituée de deux articles expérimentaux (publiés tous les deux en 2022 dans *Nucleic Acids Research*) plus un article de synthèse (publié en 2023 dans *Biochimie*). Ces trois articles portent sur l'implication de la nucléoline dans la furtivité au système immunitaire de l'oncovirus d'Epstein-Barr (EBV). Ce virus échappe à la détection par le système immunitaire en limitant la traduction de l'ARNm viral codant sa *genome maintenance protein* (*GMP*) EBNA1, une protéine essentielle à la réplication ainsi qu'à la maintenance de son génome mais également très antigénique, afin de minimiser la production de peptides antigéniques dérivés de cette *GMP*. En effet, il est estimé que la grande majorité des peptides antigéniques proviendraient de sous-produits de la traduction. Ainsi, en limitant la traduction de l'ARNm viral codant EBNA1, le virus assure une production minimale d'EBNA1, à un niveau suffisant pour que cette *GMP* assure sa fonction essentielle, mais suffisamment bas pour minimiser la production de peptides antigéniques, et donc sa détection par le système immunitaire. La *GMP* EBNA1 est donc le talon d'Achille d'EBV et sa capacité à limiter son expression est cruciale pour permette à EBV, qui infecte de façon latente et qui persiste ainsi tout au long de la vie de l'individu infecté, de persister. L'échappement de la *GMP* EBNA1 d'EBV au système immunitaire dépend notamment de son domaine de répétition de glycine et d'alanine (GAr). La nucléoline (NCL) interagit directement avec des G4 qui se forment dans

l'ARNm d'EBNA1 et permet ainsi l'inhibition GAr dépendante de la traduction d'EBNA1.

Partie I : Régulation de l'épissage de *BCL-x* par le rG4 GQ-2 et la protéine RBM25

Présentation de l'article :

L'identification de rG4 dans le pré-ARNm de *BCL-x* suggère un nouveau mode de régulation de l'épissage de ce gène (Weldon et al., 2017, 2018). Celui-ci pourrait représenter une nouvelle cible thérapeutique pour traiter des pathologies où l'expression de *BCL-x* est dérégulée. Le rôle des deux G4 présents dans le pré-ARNm de *BCL-x* sur l'épissage de ce gène n'était pas encore bien compris quand j'ai démarré mon travail de thèse en 2019. De plus, même si de nombreux facteurs d'épissage sont décrits comme impactant l'épissage de *BCL-x*, dont des G4BP, aucun n'est connu pour interagir avec les rG4 de *BCL-x*. D'après une étude globale récente, le facteur d'épissage RBM25 pourrait être une G4BP (Su et al., 2021). Ce facteur est connu pour induire une augmentation de l'isoforme pro-apoptotique Bcl-xS et se lie à proximité du site 5' d'épissage alternatif. Ainsi, le premier objectif de cette thèse était l'identification de protéines, telles que RBM25, interagissant avec l'un des deux rG4 présents dans le pré-ARNm de *BCL-x*, de façon à mieux comprendre la régulation de ce gène par les rG4.

Les résultats de cette étude ont été publiés dans *Nucleic Acids Research* en 2023. Dans cet article, nous avons montré par des tests *in vitro* de *RNA Pulldown* que le facteur d'épissage RBM25 se lie de manière directe et G4 dépendante au rG4 GQ-2. De plus, par des expériences de *Proximity Ligation Assay* (PLA), nous avons montré que la protéine RBM25 interagit avec l'ARNm endogène de *BCL-x*. Ainsi nous avons montré que RBM25 est une nouvelle G4BP qui se lie de manière directe au rG4 GQ-2, le rG4 localisé à proximité du site 5' d'épissage conduisant à la forme pro-apoptotique Bcl-xS.

Afin d'étudier l'effet du rG4 GQ-2 sur l'épissage de *BCL-x*, nous avons utilisé deux minigènes de *BCL-x*, un premier avec la séquence GQ-2 WT et un second muté avec

le minimum de mutation pour prévenir la formation du G4. Ainsi, nous avons montré qu'empêcher la formation du rG4 GQ-2 conduit à une inhibition de la sélection du site 5' d'épissage alternatif. En outre, la surexpression de RBM25 induit une augmentation dose-dépendante de la synthèse de Bcl-xS dans le minigène WT mais n'a aucun effet sur l'épissage du minigène muté. Par des tests de PLA entre la protéine RBM25 et l'ARNm produits par les minigènes, nous avons montré qu'inhiber la formation du rG4 GQ-2 réduit de manière importante l'interaction entre la protéine RBM25 et l'ARN issus des minigènes. Ces résultats confirment ainsi que la protéine RBM25 interagit avec le rG4 GQ-2 présent dans le pré-ARNm et que cette interaction est nécessaire pour la production de Bcl-xS.

La protéine RBM25 comprend différents domaines d'interaction avec les acides nucléiques. Aussi, nous nous sommes demandé quel domaine était nécessaire pour l'interaction avec le rG4 GQ-2. Pour cela, différentes constructions ont été générées, chacune manquant un des domaines d'intérêt de RBM25. Puis, par des tests de *RNA pulldown*, nous avons montré pour la première fois que le domaine RE est nécessaire pour la liaison avec les rG4. Nous avons aussi montré qu'en absence de ce domaine, la protéine RBM25 n'augmente pas la synthèse de la forme pro-apoptotique Bcl-xS.

De par l'importance du rG4 GQ-2 sur l'épissage de *BCL-x*, nous avons déterminé l'effet du composé PhenDC3, un ligand de G4, sur l'épissage de *BCL-x*. Nous avons montré que cette molécule induit une augmentation de la forme Bcl-xS aussi bien dans le modèle du minigène WT que sur le gène *BCL-x* endogène. Nous avons alors émis l'hypothèse que le composé PhenDC3 stabilisait le rG4 GQ-2 sous sa conformation de rG4, augmentant ainsi l'interaction de RBM25 avec l'ARNm de *BCL-x* ce qui conduirait à l'augmentation de la forme pro-apoptotique Bcl-xS. Nous avons testé cette hypothèse par des tests biophysiques et montré que le composé PhenDC3 stabilisait effectivement très significativement le rG4 GQ-2 *in vitro*. Ensuite, par des tests de *PLA*, nous avons montré que le composé PhenDC3 augmentait l'interaction entre la protéine RBM25 et l'ARNm de *BCL-x.* En outre, nous avons montré que l'induction de Bcl-xS par le composé PhenDC3 est corrélée avec une augmentation de l'apoptose.

Nous avons ensuite testé 90 ligands de G4 structurellement proches du composé PhenDC3 afin d'identifier de nouveaux composés capables de moduler l'épissage de *BCL-x*. De ce criblage, deux molécules capables d'augmenter encore plus efficacement que le composé PhenDC3 le ratio de Bcl-xS/Bcl-xL ont été identifiées :

les composés PhenDH8 et PhenDH9. Comme pour le composé PhenDC3, ces deux composés sont capables de stabiliser le rG4 GQ-2 et d'augmenter la synthèse de Bcl-xS au niveau ARN et protéique. Cette augmentation de la forme pro-apoptotique Bcl-xS est d'ailleurs associée à une augmentation de l'apoptose dans des cellules cancéreuses.

Nous avons donc montré non seulement une nouvelle régulation de *BCL-x* par l'interaction entre le facteur d'épissage RBM25 et le rG4 GQ-2, mais également que cette interaction peut être ciblée par des ligands de G4 afin de stabiliser ce rG4 pour augmenter la synthèse de Bcl-xS. Cette approche représente une nouvelle cible thérapeutique pertinente pour des pathologies ou la forme anti-apoptotique Bcl-xL est surexprimée telles que les cancers chimiorésistants.

Résumé graphique :

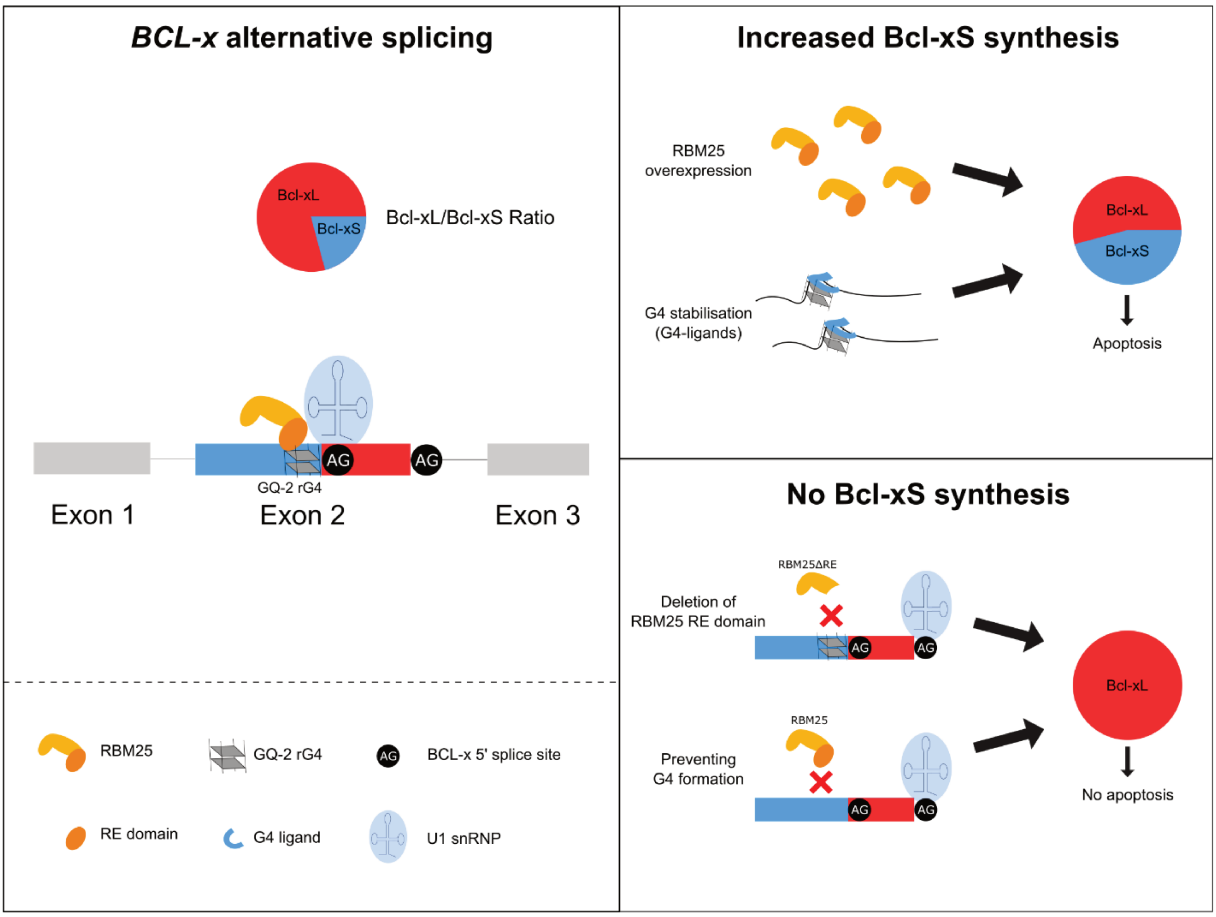


Figure 13 : régulation de l'épissage de *BCL-x* par le rG4 GQ-2 et le facteur d'épissage RBM25.

Dans des cellules saines, la protéine pro-apoptotique Bcl-xS est minoritaire. Cette protéine est produite lorsque le site 5' d'épissage alternatif est sélectionné. La sélection de ce site 5' d'épissage est renforcée par la formation du rG4 GQ-2. Ce rG4 sert de plateforme de reconnaissance pour le facteur d'épissage RBM25. Une fois RBM25 fixé au GQ-2, le snRNP U1 est recruté au niveau du site 5' d'épissage alternatif. La partie C-terminale de l'exon 2 est ensuite coupée par la machinerie d'épissage et se retrouve absente de l'ARNm mature. La surexpression de RBM25 et la stabilisation du rG4 GQ-2 par des ligands de G4 conduisent toutes deux à l'augmentation de la forme pro-apoptotique Bcl-xS ainsi qu'à une augmentation de l'apoptose dans des cellules cancéreuses. A l'inverse, déléter le domaine RE de liaison au G4 de RBM25 ou empêcher la formation du rG4 GQ-2 conduit à une inhibition de la sélection du site 5' alternatif et à l'absence de synthèse de la protéine Bcl-xS.

Nucleic Acids Research, 2023, 1–19
https://doi.org/10.1093/nar/gkad772
RNA and RNA-protein complexes

Alternative splicing of *BCL-x* is controlled by RBM25 binding to a G-quadruplex in *BCL-x* pre-mRNA

Ronan Le Sénéchal [1], Marc Keruzoré [1,†], Alicia Quillévéré [1,†], Nadège Loaëc [1], Van-Trang Dinh [1], Oksana Reznichenko [2], Pedro Guixens-Gallardo [2], Laurent Corcos [1], Marie-Paule Teulade-Fichou [2], Anton Granzhan [2] and Marc Blondel [1,*]

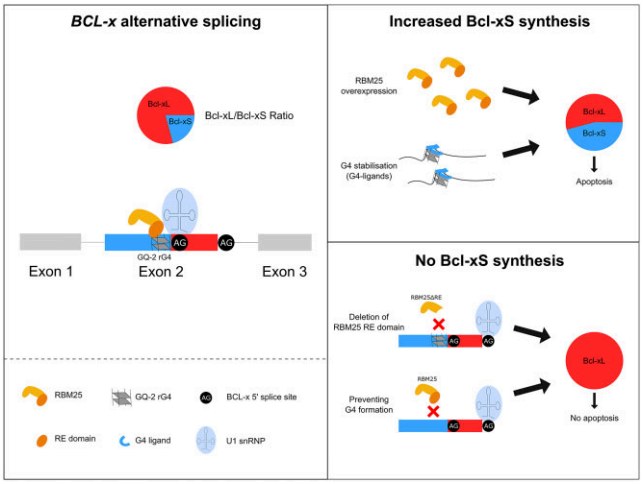
[1]Univ Brest; Inserm UMR1078; Etablissement Français du Sang (EFS) Bretagne; CHRU Brest, Hôpital Morvan, Laboratoire de Génétique Moléculaire, 22 avenue Camille Desmoulins, F-29200 Brest, France
[2]Chemistry and Modelling for the Biology of Cancer (CMBC), CNRS UMR9187, Inserm U1196, Institut Curie, Université Paris Saclay, F-91405 Orsay, France

[*]To whom correspondence should be addressed. Tel: +33 2 98 01 83 88; Email: marc.blondel@univ-brest.fr
[†]Contributed equally.

Abstract

BCL-x is a master regulator of apoptosis whose pre-mRNA is alternatively spliced into either a long (canonical) anti-apoptotic Bcl-xL isoform, or a short (alternative) pro-apoptotic Bcl-xS isoform. The balance between these two antagonistic isoforms is tightly regulated and overexpression of Bcl-xL has been linked to resistance to chemotherapy in several cancers, whereas overexpression of Bcl-xS is associated to some forms of diabetes and cardiac disorders. The splicing factor RBM25 controls alternative splicing of *BCL-x*: its overexpression favours the production of Bcl-xS, whereas its downregulation has the opposite effect. Here we show that RBM25 directly and specifically binds to GQ-2, an RNA G-quadruplex (rG4) of *BCL-x* pre-mRNA that forms at the vicinity of the alternative 5′ splice site leading to the alternative Bcl-xS isoform. This RBM25/rG4 interaction is crucial for the production of Bcl-xS and depends on the RE (arginine-glutamate-rich) motif of RBM25, thus defining a new type of rG4-interacting domain. PhenDC3, a benchmark G4 ligand, enhances the binding of RBM25 to the GQ-2 rG4 of *BCL-x* pre-mRNA, thereby promoting the alternative pro-apoptotic Bcl-xS isoform and triggering apoptosis. Furthermore, the screening of a combinatorial library of 90 putative G4 ligands led to the identification of two original compounds, PhenDH8 and PhenDH9, superior to PhenDC3 in promoting the Bcl-xS isoform and apoptosis. Thus, favouring the interaction between RBM25 and the GQ-2 rG4 of *BCL-x* pre-mRNA represents a relevant intervention point to re-sensitize cancer cells to chemotherapy.

Graphical abstract



Introduction

Alternative splicing (AS), discovered simultaneously with splicing >40 years ago (1–3), is prevalent in multicellular eukaryotes in which it is the main mechanism allowing one gene to produce various proteins, thereby increasing the complexity of the proteome from a limited number of genes. In addition, AS plays key roles in determining tissue- and species-specific differentiation patterns. AS is regulated by multiple co- and post-transcriptional regulatory mechanisms and can

Received: March 22, 2023. **Revised:** August 5, 2023. **Editorial Decision:** August 29, 2023. **Accepted:** September 9, 2023

play a causal role in various hereditary disorders as well as cancers (4,5).

The *BCL-x* gene (aka *BCL2-like 1*, *BCL2L1*) encodes Bcl-x (6), a member of the Bcl-2 (B-cell lymphoma) family of proteins that are major regulators of apoptosis, a highly conserved and essential programmed cell death process that eliminates unwanted cells, including damaged cells, lifelong including during development (7). Bcl-x is a mitochondrial transmembrane protein that regulates the intrinsic apoptosis pathway. Owing to the existence of two alternative 5′ splice sites (5′ss) in its exon 2, the *BCL-x* pre-mRNA is alternatively spliced into either a long (canonical) anti-apoptotic Bcl-xL isoform, or a short (alternative) pro-apoptotic Bcl-xS isoform (6). The ability of Bcl-xL to inhibit apoptosis involves several mechanisms that include inhibition of Bax-induced mitochondrial outer membrane permeabilization (MOMP), a key step in apoptosis execution (8). Besides this canonical function, Bcl-xL has been also shown to remotely control apoptosis by inhibiting IP3R channels and therefore the release of Ca^{2+} from the endoplasmic reticulum (RE) (9,10), the transfer of Ca^{2+} from the RE to the mitochondria being also involved in MOMP (11). Bcl-xS has been proposed to activate apoptosis by inhibiting Bcl-xL via a direct interaction (8). Hence, depending on the balance between Bcl-xL and Bcl-xS, *BCL-x* can either be pro- or anti-apoptotic. As such, the balance between these two antagonistic isoforms is tightly regulated and overexpression of the anti-apoptotic Bcl-xL isoform has been linked to resistance to chemotherapy in several cancers (12,13), whereas overexpression of the pro-apoptotic Bcl-xS isoform is associated to some forms of diabetes and cardiac disorders (14,15). Therefore, pharmacological means to switch the splicing in favour of Bcl-xS may have therapeutic benefits for a number of cancers, in particular those resistant to chemotherapy due to Bcl-xL overexpression. Conversely, favouring Bcl-xL represents a possible intervention point for diseases linked to Bcl-xS overexpression such as some forms of diabetes. In line, splice switching oligonucleotides (SSO) that block the 5′ss leading to the long canonical Bcl-xL isoform have been developed and shown to reduce the overall cell viability of a variety of cancer cell lines (16–19). In addition, inhibitors of SR kinases that phosphorylate serine-arginine-rich (SR) proteins, have been tested (20,21) with the inherent limitation that alterations in the phosphorylation of SR proteins, which play prominent roles in pre-mRNA constitutive and alternative splicing decisions, are likely to have widespread effects on the splicing of many genes.

G-quadruplex (G4) are non-canonical secondary structures that may assemble in guanine-rich (G-rich) DNA or RNA. G4 are formed by the stacking of at least two G-quartets, each representing a planar arrangement of four guanines connected by Hoogsteen hydrogen bonds and stabilized by a central cation, most often K^+. G4 structures within G-rich DNA or RNA have been implicated in gene regulation where they can affect transcription (22), splicing (23–25) or translation (26–32).

G4 have been shown to be involved in *BCL-x* alternative splicing (33,34). Indeed, the *BCL-x* pre-mRNA contains a number of G-rich sequences that can potentially form at least six RNA G4 (rG4) as predicted *in silico*. In addition, thanks to FOLDeR, a new strategy based on footprinting of long 7-deazaguanine-substituted RNA, the formation of two of these rG4 was confirmed *in vitro* under functional conditions: one (GQ-2) lying in close proximity with the 5′ss leading to the short (alternative) pro-apoptotic Bcl-xS isoform (xS 5′ss) while the other (GQ-5) is overlapping the 5′ss leading to the long (canonical) anti-apoptotic Bcl-xL isoform (xL 5′ss) (33). In addition, the ability of GQ-2 to adopt a G4 conformation has been recently confirmed by various biophysical experiments (34,35).

Multiple splice factors and signalling pathways have been implicated in the regulation of *BCL-x* pre-mRNA (alternative) splicing. Most of these proteins are general splice factors and include SRSF1 (36,37), SRSF2 (38), SRSF 3 and 7 (39), SRSF9 (36) and SRSF10 (40). The following hnRNP have also been involved: A1 (37), K (41), F/H (42,43) and PTBP1 (39) as well as other RNA-binding protein: Sam68 (37), SF3B1 (44), TRA2 (39), RBM4 (45), RBM10 (46), RBM11 (47) and RBM25 (48). RBM25 (also known as RED120) belongs to a family of RNA-binding proteins whose members share an arginine-glutamate-rich (RE) central region and a C-terminal proline-tryptophan-isoleucine (PWI) motif. RBM25 also contains an N-terminal RNA recognition motif (RRM). RBM25 associates with various splicing factors that include SRm160/300 and U snRNAs, as well as with spliced RNA (49). Overexpression of RBM25 has been shown to increase the expression of the alternative pro-apoptotic Bcl-xS isoform and to trigger apoptosis whereas its downregulation leads to accumulation of the canonical anti-apoptotic Bcl-xL isoform (48).

Here we show that RBM25 directly and specifically binds in a G4 structure-dependent manner to the GQ-2 rG4 of the *BCL-x* pre-mRNA, which forms in the vicinity of the alternative 5′ splice site that leads to the alternative Bcl-xS isoform (xS 5′ss). We also show that this RBM25/rG4 interaction is crucial for the production of Bcl-xS and depends on the central RE motif of RBM25, thus defining a new type of rG4-interacting domain. This RBM25/rG4 interaction is druggable as PhenDC3, a benchmark G4 ligand, stabilizes GQ-2 rG4 *in vitro* and enhances the binding of RBM25 to the GQ-2 rG4 of the *BCL-x* pre-mRNA, hence promoting the alternative pro-apoptotic Bcl-xS isoform and apoptosis. Therefore, favouring the interaction between RBM25 and the GQ-2 rG4 of *BCL-x* pre-mRNA represents a relevant intervention point to re-sensitize cancer cells to chemotherapy, whereas interfering with this interaction may represent a therapeutic avenue for diseases associated with Bcl-xS overexpression, such as some forms of diabetes. For this reason, we have also screened a library of 90 putative G4 ligands related to PhenDC3 (50) and identified PhenDH8 and PhenDH9, two original G4-ligands, that are more potent than PhenDC3 to promote the alternative pro-apoptotic Bcl-xS isoform and apoptosis.

Material and methods

In vitro characterization of RNA oligonucleotides

RNA oligonucleotides (Eurogentec, HPLC-purified) were dissolved in deionized water at a concentration of 100 μM and further diluted to a concentration of 4 μM with a lithium cacodylate buffer (10 mM, pH 7.2) supplemented with 100 mM KCl, 100 mM LiCl or their mixture to obtain the desired concentration of K^+, in a sample volume of 1 mL. Thermal denaturation spectra (TDS) were recorded as described by Mergny *et al.* (51) using sample temperatures of 20 and 80°C. Ultraviolet (UV) melting profiles were recorded using a heating ramp of 0.2°C min^{-1} and a detection wavelength of 295 nm on a Cary 300 Bio spectrophotometer (Agilent) equipped with

a Peltier-controlled multicell holder, and T_m values were obtained by the first-derivative analysis of melting transitions. FRET melting experiments were performed according to the published procedure (52), using 5′-FAM / 3′-TAMRA labelled GQ2 oligonucleotide (F-GQ2-T, Eurogentec).

Plasmid constructions

All the RBM25 plasmids were generated using standard procedures. The T4 DNA ligase was obtained from Promega and the restriction enzymes were purchased from New England Biolabs. RBM25 was amplified from complementary DNA (cDNA) prepared from extracts of H1299 cells with primers 5′-GCGCGCGGTACCATGTACCCATACGATGTTCCAG ATTACGCTTCTTTTCCACCTCATTTGAATCGC-3′ and 5′-GCGCGCGCGGCCGCTTACTTCACAAGACCAATTT TCTTGGC-3′. The PCR fragment was cloned into KpnI and NotI cloning sites of pcDNA 3.1(+) vector. The generated construction was verified by PCR amplification, restriction enzymes digestion and sequencing.

The Bcl-x 672 minigenes synthesis was carried out by GenScript. On reception, the minigenes were amplified in bacteria. Plasmid extraction was performed using the NucleoBond® Xtra Midi EF kit protocol (Macherey-Nagel). The plasmids were verified by PCR amplification and sequencing.

Cell culture, treatment and transfection

H1299 cells were derived from metastatic lymph node from lung carcinoma. HeLa cells were derived from a cervical carcinoma. A549 cells were derived from a lung carcinoma. H1299, HeLa and A549 cells were purchased from ATCC and cultured in RPMI-1640 supplemented with 10% foetal bovine serum and 2 mM L-glutamine. Transfections were performed using GeneJuice® (Merck) or FuGENE® HD (Promega) transfection reagent according to the manufacturers' protocols. Where indicated, cells were treated with various concentrations of G4 ligands, doxorubicin (Sigma) or with DMSO (compounds vehicle) for 24 or 48 hours.

Protein extraction from mammalian cells

Whole cells at 75–90% confluency were harvested, washed with 1x phosphate buffered saline (PBS) and suspended into lysis buffer (20 mM HEPES pH 7.5, 50 mM β-glycerolphosphate, 1 mM EDTA pH 8, 0.5 mM Na_3VO_4, 100 mM KCl, 10% glycerol, 1% Triton and anti-proteases cocktail (Roche 11697498001)). These cell suspensions were mechanically lysed before centrifugation at 16 000*g* for 20 min at 4°C and protein concentration was determined using a Bradford assay.

Western blotting

Equal protein quantities and volumes of all samples were loaded and run on Bolt or NuPAGE™ 10% Bis-Tris Protein gels (Invitrogen), then transferred onto 0.45 μM nitrocellulose membrane (GE Healthcare). Membranes were blocked in 1x PBS, 0.1% Igepal and 3% bovine serum albumin (BSA), and incubated with the indicated primary antibodies: mouse anti-GAPDH (Abcam ab125247, 1/5000), rat anti-HA (Roche 11867423001, 1/2000), rabbit anti-RBM25 (Merck HPA003025, 1/1000) or rabbit anti-Bcl-xL (CellSignaling no. 2764, 1/500). The membranes were then washed with fresh 1x PBS 0.1% Igepal and incubated with the indicated secondary antibodies conjugated to horseradish peroxidase: rabbit anti-mouse (Dako P0161, 1/3000), swine anti-rabbit (Dako P0217, 1/2000) or goat anti-rat (Millipore AP136P, 1/3000). The membranes were washed afresh and analysed by enhanced chemiluminescence in the following buffer (Tris-base pH 8.5, 12.5 nM coumaric acid, 2.25 nM luminol and 0.15% H_2O_2) using a Vilber-Lourmat Photodocumentation Chemistart 5000 imager. Relative proteins levels for each sample were normalized to GAPDH protein levels using the ImageJ software.

RNA and RNA pulldowns

RNA pulldowns experiments were performed as previously described (27,31). Briefly, the following G-quadruplex forming RNA oligonucleotides 3′-tagged with tetraethylene glycol (TEG)-Biotin were used:

- GQ-1 GGGAGGCAGGCGACGAGUUUGAACUGCG GUACCGGCGGGCA-3′-TEG-Biotin,
- GQ-2 GGGAUGGGGUAAACUGGGGUCGCAUUGU GG-3′-TEG-Biotin,
- GM-2 GAGAUGAGGUAAACUGAGGUCGCAUUGU GG-3′-TEG-Biotin,
- GQ-5 GGAUCCAGGAGAACGGCGGCUGG-3′-TEG-Biotin and
- GM-5 GGAUCCAGAAGAACGGCGGCUGA-3′-TEG-Biotin.

The G-quadruplexes were formed by heating the RNA oligonucleotides at 95°C for 5 min then cooling down to 4°C at a rate of 2°C per minute in folding buffer (10 mM Tris-HCl pH 7.5, 0.1 mM EDTA) in the presence of 100 mM KCl, or in the presence of 100 mM LiCl. To avoid unspecific binding, high affinity streptavidin Sepharose beads (GE Healthcare, 28985799) were incubated in 1 mL of blocking buffer (10 mM Tris-HCl, pH 7.5, 100 mM KCl, 0.1 mM EDTA, 1 mM DTT, 0.01% Triton, 0.1% BSA, 0.02% *S. cerevisiae* tRNAs (Sigma-Aldrich 10109495001)) for 1 h at 4°C on a rotating wheel. An amount of 10 μg of each batch of folded biotinylated RNA oligonucleotides was incubated with 50 μL of solution containing the streptavidin sepharose beads for 90 min at 4°C on a rotating wheel. Next, 500 μg of cell extracts or 200 ng of recombinant RBM25 were treated with 200 U/mL of RNase Inhibitor (NEB, M0307S) for 90 min at 4°C on a rotating wheel. These extracts were then incubated with the RNA oligonucleotides bound to the streptavidin beads for 90 min at room temperature. Beads were then washed five times with lysis buffer and then with lysis buffer with increasing KCl or LiCl concentrations (200 to 800 mM). Protein still bound to beads after the washes were eluted using 2x loading buffer (2x Laemmli Buffer with 5% β-mercaptoethanol) and analysed by western blotting with antibodies raised against RBM25 or HA.

The same protocol was applied for DNA pulldowns, except that G-quadruplex forming DNA oligonucleotides 3′-tagged with TEG-Biotin were used.

Proximity ligation assay adapted to protein–RNA interactions

An adaptation of proximity ligation assay (PLA) to detect protein–RNA interactions (31,53) has been used. Briefly, cells were fixed in 1x PBS, 4% paraformaldehyde for 20 min and permeabilized for 10 min with 0.4% Tri-

ton X-100, 0.05% CHAPS. The Bcl-x-mRNA-digoxigenin probe (5′-CCATTGTCCAAAACACCTGCTCACTCACTG AGTCTCGTCTCTGGAAAAA-3′) at a quantity of 100 ng per well was denatured 5 min at 80°C. The probe hybridization reaction was carried out in 40 μL of hybridization buffer (10% formamide, 2X SSC, 0.2 mg/ml *E. coli* tRNA, 0.2 mg/ml salmon sperm DNA and 2 mg/ml BSA). Fixed cells were washed and blocked in the blocking solution (1x PBS, 3% BSA, 0.1% Saponine) before incubation with the primary antibodies (anti-digoxigenin, Sigma 1/200 and anti-RBM25, Merck 1/4000). The PLA reaction was performed according to the manufacturer's protocol using the Duolink® PLA *in situ* kit, PLA probe anti-rabbit Plus, PLA probe anti-mouse Minus and the FarRed *in situ* detection reagent, all purchased from Sigma. The results were analysed using a Zeiss Axio Imager M2 and the ImageJ software. All PLA experiments were performed at least twice independently and the following controls were implemented: with sense mRNA probe or without primary antibodies. The PLA method with the use of Flag Bcl-x minigene was carried out as follow: H1299 cells were plated and transfected with 400 ng of Flag Bcl-x minigene plasmid 24 h before the paraformaldehyde fixation process described before was performed. The Flag-Bcl-x-mRNA-digoxigenin probe (5′-CCATGGGGATCACC TCCCGCTTGTCGTCATCGTCTTTGTAGTAAAAA-3′) at a quantity of 100 ng per well was denatured 5 min at 80°C prior to the hybridization step. After the hybridization of the probe, an incubation with the primary antibodies (anti-digoxigenin Sigma 1/200 and anti-RBM25 Merck 1/1000) was performed. The remaining steps of the protocol are as described before. The PLA to assess the interaction between RBM25 and G4s used the same protocol (without the hybridization step) except that the anti-G4 (BG4) antibody (Absolute Antibody, 1/200) was used together with the anti-RBM25 antibody (Merck 1/4000).

RNA extraction and semi-quantitative RT–PCR

Total A549 or HeLa cellular RNA were extracted using the RNeasy Plus kit (Qiagen). cDNA synthesis was carried out using 500 ng of RNA and the SuperScript™ IV Reverse Transcriptase (Invitrogen) together with oligo$(dT)_{20}$ primer (Invitrogen). cDNA samples were analysed using semi-quantitative PCR using the MasterMix OneTaq Polymerase (NEB). Bands were quantified using the ImageJ software to determine Bcl-xS percentage.

For Bcl-x 672 minigene splicing analysis, PCR was performed using P1 primer specific to pcDNA 3.1 (+) plasmid (5′-TAGAAGGCACAGTCGAGG-3′) and P2 primer specific to *BCL-x* (5′-GGGAGGTGATCCCCATGGCAG-3′).

For endogenous *BCL-x* splicing analysis, PCR was performed using P2 primer (5′-GGGAGGTGATCCCCATGGCA G-3′) and P3 primer specific to *BCL-x* (5′-GGCCACAGTCAT GCCCGTCAG-3′).

Immunofluorescence

H1299 cells were plated on 13 mm-diameter coverslips in 24-well plates and transfected with HA-RBM25-FL or HA-RBM25ΔRE plasmids. At 24 h post-transfection, cells were fixed with 4% paraformaldehyde in PBS for 20 min and permeabilized with PBS 0.4%, Triton X-100, CHAPS 0.05% for 10 min at room temperature. After incubation in blocking buffer (1x PBS, 3% BSA, 0.1% Saponin) for 30 min at room temperature, samples were incubated with a mouse polyclonal anti-HA antibody (a kind gift from Borek Vojtesek, Masaryk Memorial Cancer Institute, Brno, Czech Republic) at 1/1000 and a rabbit anti-RBM25 antibody (Merck HPA003025) at 1/500 for 2 h at room temperature, followed by 1 h incubation with 1/500 dilutions of the goat anti-mouse immunoglobulin G (IgG) secondary antibody conjugated to Alexa Fluor® 594 and of the goat anti-rabbit immunoglobulin G (IgG) secondary antibody conjugated to Alexa Fluor® 488 (both purchased from Invitrogen). Both antibodies were diluted in blocking buffer. 4',6-Diamidino-2-phenylindole (DAPI) was used for nuclear counterstaining and the images were taken using a Zeiss Axio Imager M2.

Caspase activity assay

About 4000 A549 cells were plated in 0.1 mL of medium per well in 96-well flat-bottom plates. After 18 h, cells were exposed to the indicated compounds at the indicated concentrations or to dimethylsulphoxide (DMSO; vehicle) as a negative control. After 48 h of treatment, the Caspase 3/7 activity was determined using the Caspase Glo® 3/7 Assay (Promega) according to the manufacturer's protocol.

MTT assay

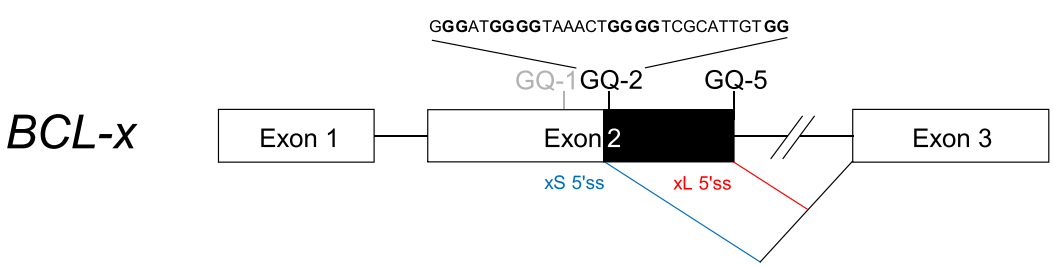
About 4000 A549 cells were plated in 0.1 ml of medium per well in 96-well flat-bottom plates. After 18 h, cells were exposed to the indicated compounds at the indicated concentrations or to DMSO (vehicle) as a negative control. After 48 h of treatment, 10 μL of MTT solution (5 mg/mL in PBS (pH 7.4), CT01-5, Merck Millipore) was added to each well and cells were incubated for 4 h. Next, 100 μL of an isopropanol solution containing 0.1 N HCl and 10% (v/v) Triton X100 was added to each well to dissolve the formazan crystals, and the absorbance at 570 nm was then measured.

Results

RBM25 selectively binds to GQ-2 rG4 in a G4 structure-dependent manner

Whereas exons are enriched in guanines in eukaryotes (54), the presence of G4 seems to have been counter selected (55,56). However, as stated before, the formation of two rG4 from two potential quadruplex-forming sequences (PQS) in the exon 2 of the *BCL-x* gene have been confirmed *in vitro* under functional conditions (33). One of these two rG4, GQ-2, is localized at the close vicinity of the 5′ss leading to the short alternative pro-apoptotic Bcl-xS isoform (xS 5′ss) whereas the other, GQ-5, is overlapping the 5′ss leading to the long canonical anti-apoptotic Bcl-xL isoform (xL 5′ss) (Figure 1A). The sequence of these two PQS is shown in Figure 1B, together with the sequence of GQ-1, a PQS upstream of GQ-2 whose formation was not confirmed *in vitro* under physiological conditions, and of GM-2 and GM-5, two mutated versions of GQ-2 and GQ-5, respectively, in which a minimal number of guanines (3 and 2, respectively) involved in G4 formation were replaced by adenines to prevent G4 formation, as predicted using the G4Killer software that has been developed to design mutations crippling G4 propensity (57). The capacity of GQ-2 fragment to form rG4 structure *in vitro* has recently

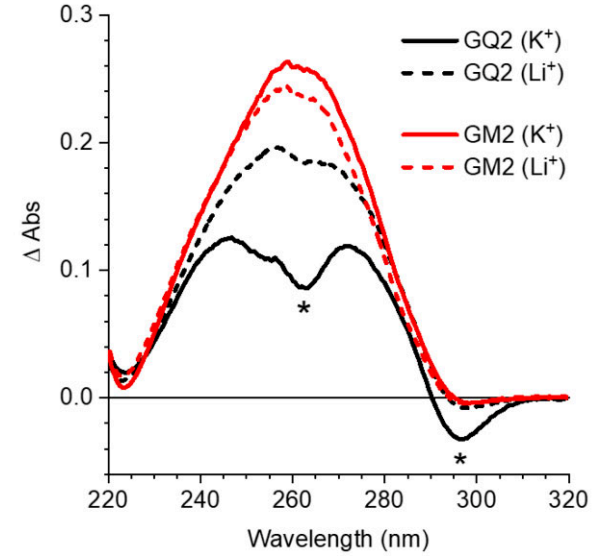
A



B

Name	Length (nt)	Sequence (5'→3')	QGRS G-Score	G4Hunter Score
GQ-1	41	GGG AGG CAG GCG ACG AGU UUG AAC UGC GGU ACC GGC GGG CA	6	N/A
GQ-2	30	GGG AUG GGG UAA ACU GGG GUC GCA UUG UGG	17	1.467
GM-2	30	GAG AUG AGG UAA ACU GAG GUC GCA UUG UGG	0	0
GQ-5	23	UUG GAU CCA GGA GAA CGG CGG CUG GGU A	21	N/A
GM-5	23	UUG GAU CCA GAA GAA CGG CGG CUG GAU A	0	N/A

C



D

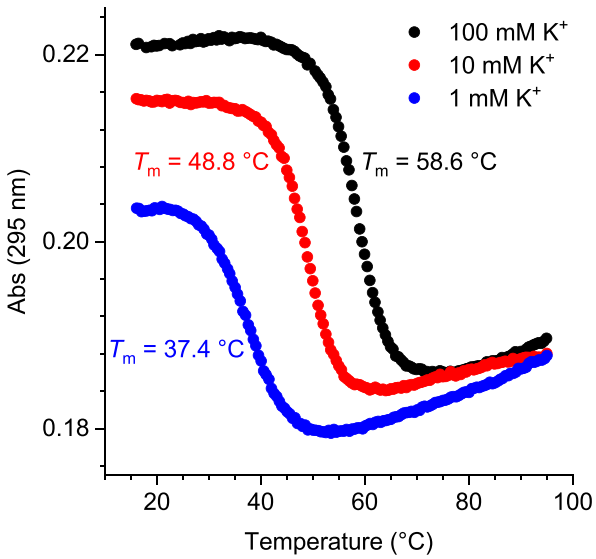


Figure 1. GQ-2 is an rG4 that may form in close proximity to the alternative splice site that leads to the synthesis of Bcl-xS, the short alternative pro-apoptotic isoform of Bcl-x. (**A**) Schematic representation of the *BCL-x* gene. The two alternative 5′ splice sites in the exon 2 are indicated (xS 5′ss is highlighted in blue and xL 5′ss in red) as well as the two putative rG4 (GQ-2 and GQ-5) that may assemble at close proximity to these two splice sites. GQ-1 (written in grey) is a putative quadruplex sequence (PQS), as predicted by G4-finder, which has not been confirmed *in vitro*. (**B**) Sequences of the various RNA oligonucleotides that have been used for RNA pulldown experiments. The various sequences, together with their predicted propensity to form G4 (as determined using two different software: G-score and G4-hunter) are indicated and the guanines potentially important for G4 formation are underlined. (**C**) TDS (ΔAbs = Abs (80°C) − Abs (20°C)) of GQ-2 and GM-2 RNA oligonucleotides (4 μM) in 10 mM lithium cacodylate buffers supplemented with 100 mM KCl (solid lines) or 100 mM LiCl (dashed lines). G4-characteristic peaks are indicated with asterisks. (**D**) UV-melting analysis of GQ-2 (4 μM) in the presence of various concentrations of K^+ as indicated.

been demonstrated by Bhogadia *et al.* through CD, ^{1}H nuclear magnetic resonance and gel electrophoresis experiments (35). Here, we additionally assessed the capacity of GQ-2 and its variants to form G4 structures *in vitro* using TDS and UV melting experiments. In agreement with previously published results and bioinformatics analysis (QGRS Mapper (58) and G4Hunter (59), cf. the corresponding scores in Figure 1B), we found that, in the presence of K^+, but not Li^+, GQ-2 adopted a G4 conformation evidenced by characteristic negative TDS peaks at 295 and 260 nm, whereas GM-2 was unable to do so, even in the presence of K^+ (Figure 1C). The UV-melting profile of GQ-2 featured a negative transition at 295 nm, with melting temperatures being strongly dependent on K^+ concentration (T_m = 37.4, 48.8 and 58.6°C in the presence of 1, 10 or 100 mM KCl, respectively, Figure 1D), which is a strong hallmark of G4 formation (60). To get preliminary insight into the rG4 structure formed by GQ-2, we studied three additional sequence variants in which isolated guanines G22 or/and G27 were replaced by adenines (GQ2-A22, GQ2-A27 and GQ2-A22A27, respectively). Their thermal difference spectra (Supplementary Figure 1) revealed G4-characteristic negative peaks at 295 and 260 nm in the presence of K^+, but not Li^+, giving evidence that these variants were also able to form rG4 structures. In addition, UV-melting analysis (Supplementary Figure 2) indicated that thermal stability of these variants was only marginally affected compared to GQ-2, indicating that guanines G22 and G27 do not participate in formation of the rG4 core. However, in the case of GQ2-A27 and, especially, GQ2-A22A27, the amplitude of G4-specific melting transition was reduced at low K^+ concentration (1 mM), suggesting possible formation of alternative structures in addition to rG4. Altogether, these observations suggest that GQ2 forms a two-layer rG4 structure without bulge.

Next, to identify protein(s) presenting potential affinity for GQ-2 and/or GQ-5 rG4, we performed RNA pulldown assays as previously described (31,61). Briefly, this assay is based on the use of an RNA oligonucleotide containing a sequence that can form G4, 3′-conjugated to biotin through a TEG spacer, which can therefore be precipitated by magnetic sepharose beads conjugated to streptavidin. This way, proteins with affinity for the selected RNA-G4 can be precipitated and detected by western blotting. Interestingly, RBM25 has recently been identified in a comprehensive approach as one of 281 potential G4-related proteins in MM231 cells, thanks to an original method based on a photoactive G4 ligand (62). Hence, in addition to regulate *BCL-x* alternative splicing as presented in the introduction, RBM25 is a potential G4-binding protein. For these reasons we first determined, using RNA pulldown assays, if endogenously expressed RBM25 binds, or not, to one or the other of the following matrices: GQ-2, GM-2, GQ-5 and GM-5 RNA oligonucleotides conjugated to biotin. For this purpose, we used total protein extracts prepared from H1299 cells. As shown in Figure 2A, an efficient binding of endogenous RBM25 was observed only on the GQ-2 matrix. This binding mostly depends on G4 since only a residual binding was observed with the GM-2 sequence unable to adopt a G4 structure, similar to the residual binding observed with GQ-5 and GM-5 matrices.

We then repeated these RNA pulldown experiments using total protein extracts prepared from H1299 cells transfected with a plasmid allowing expression of HA-tagged RBM25 (HA-RBM25-FL) (Figure 2B). As an additional control, we used biotin-conjugated GQ-1 RNA oligonucleotide. We found that exogenously expressed HA-RBM25 efficiently binds to GQ-2 whereas only a residual binding was observed on all the other matrices.

The binding of RBM25 to the GQ-2 rG4 may be indirect and involve other protein(s). To determine whether the specific binding of RBM25 is direct or indirect, we repeated the same RNA pulldown experiments using purified recombinant His-tagged RBM25 produced in bacteria (His-RBM25). We obtained the same result as His-RBM25 efficiently binds to GQ-2, whereas only a residual binding was observed on GM-2 control matrix that does not form G4 (Figure 2C). This result shows that RBM25 directly binds to GQ-2 rG4. We also performed DNA pulldown experiments to determine whether RBM25 is able to bind to DNA G-quadruplexes (dG4). For this purpose, we have used the same protocol than the one for RNA pulldown, except that the following DNA matrices were employed: GQ-2, GM-2, BCL-2, BCL-2mut, MYC and MYCmut DNA oligonucleotides 3′-conjugated to biotin that contain, or not for the mutants, a sequence that can form dG4. As shown in Supplementary Figure 3, RBM25 is able, at least *in vitro*, to bind to these various dG4, compared to their mutated versions that cannot form dG4. Of note, the binding of RBM25 appears less efficient on GQ-2 dG4 than on GQ-2 rG4 suggesting that RBM25 may have more affinity for rG4 than for dG4.

To confirm that the binding on the GQ-2 matrix was G4-dependent, we performed an additional control experiment solely based on the GQ-2 matrix folded in the presence of K^+ (100 mM KCl), which allows G4 formation, or Li^+ (100 mM LiCl), which does not allow efficient G4 formation. As shown in Figure 2D, the binding of recombinant His-RBM25 is more efficient in the condition where G4 formation is favoured (K^+), thereby confirming that the interaction of RBM25 with the GQ-2 rG4 is indeed G4 structure-dependent.

Finally, we performed PLA to assess whether, and where in the cell, endogenously expressed RBM25 interacts with endogenously expressed *BCL-x* mRNA. For this purpose, we used a recently described adaptation of PLA to monitor protein–RNA interactions (53). As shown in Figure 2E, we observed a clear nuclear signal (median value of six dots per cell, most of them being nuclear, as compared to zero dots per cell as a median value for the negative control). This confirms *in cellulo* the interaction between RBM25 protein and *BCL-x* RNA (both partners being expressed at their physiological level) and indicates that this interaction takes place in the nucleus, as expected for interactions between splice proteins and pre-mRNA. In addition, to assess the ability of RBM25 to bind G4s *in cellulo*, we performed an additional PLA using an antibody raised against G4 (BG4, (63)) together with the antibody against RBM25 (Supplementary Figure 4). We observed numerous PLA dots, which are mostly localized in the nucleus (median value of 40 dots per nucleus as compared to one dot per nucleus as a median value for the negative control), clearly showing that RBM25 binds G4s in the cell, mostly in the nucleus.

Taken together, all these results indicate that RBM25 selectively binds in a G4-dependent manner to GQ-2, the rG4 that forms at the vicinity of the 5′ss leading to the short alternative pro-apoptotic Bcl-xS isoform (xS 5′ss).

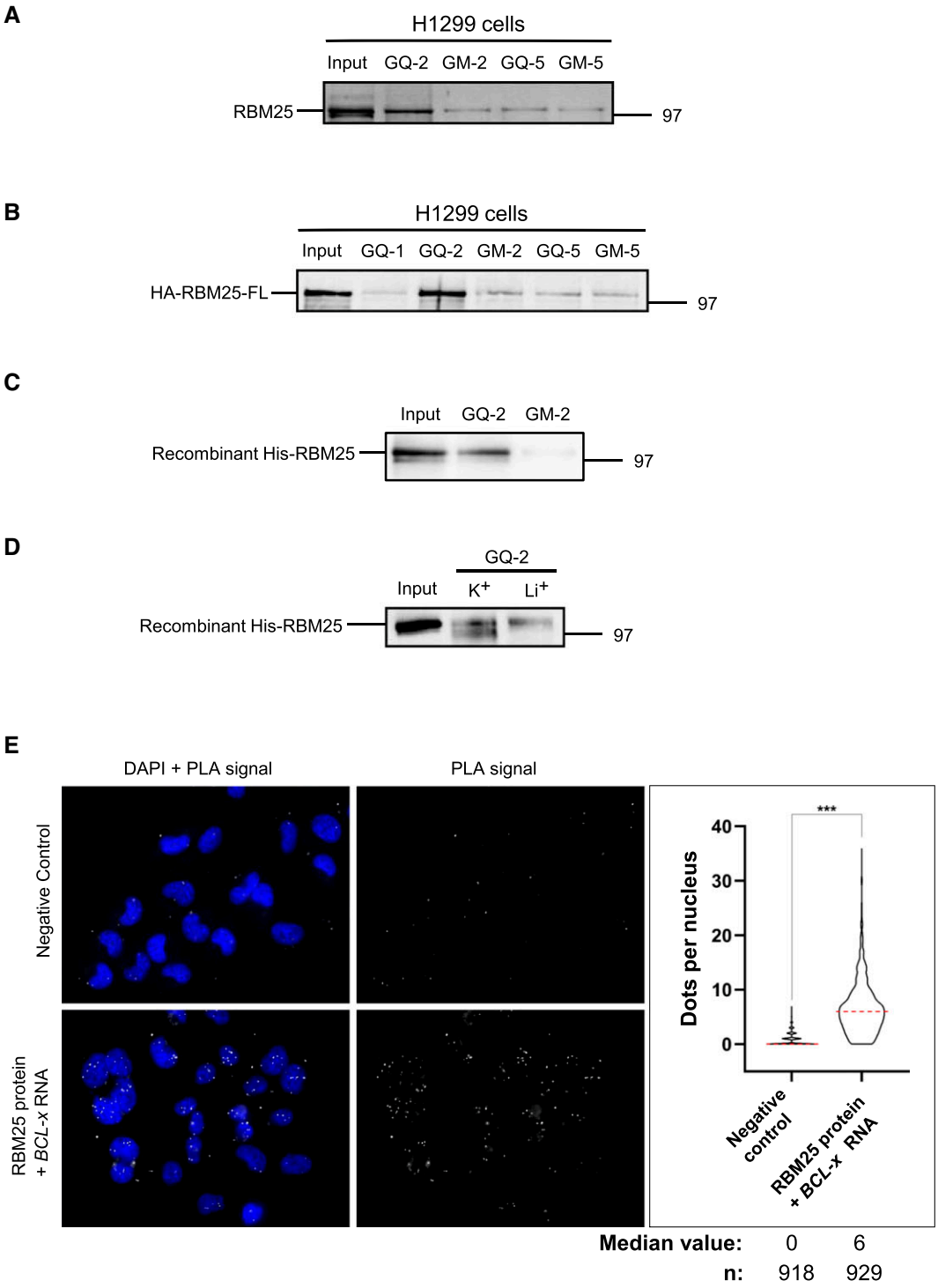


Figure 2. RBM25 selectively binds to GQ-2 rG4 in a G4-dependent manner (**A**) RNA pulldown of extracts from H1299 cells using the indicated RNA oligonucleotide matrices. The proteins still bound after an 800 mM KCl wash were eluted and analysed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS–PAGE) and western blot using an anti-RBM25 antibody to reveal endogenous RBM25. Gel represents $n \geq 3$. (**B**) Same RNA pulldown experiment as in (**A**) except that extracts from H1299 cells transfected by a plasmid allowing expression of HA-tagged RBM25 (HA-RBM25-FL) were used. Exogenous HA-RBM25-FL was revealed using an anti-HA antibody. Gel represents $n \geq 3$. (**C**) The same RNA pulldown experiment as in (**A**) and (**B**) except that a recombinant polyhistidine-tagged RBM25 protein (His-RBM25) was used instead of extracts from H1299 cells. Gel represents $n \geq 3$. (**D**) Control experiment using the GQ-2 matrix folded in the presence of 100 mM KCl (that favours G4 formation) or 100 mM LiCl (that does not favour formation of G4), as indicated, and recombinant His-RBM25. Gel represents $n \geq 3$. (**E**) Adaptation of the PLA to monitor the RBM25 protein-*BCL-x* mRNA interaction performed in H1299 cells natively expressing RBM25 protein and *BCL-x* mRNA. Microscopy images of H1299 cells analysed using a probe specifically hybridizing to the *BCL-x* RNA (lower panels) or not (upper panels, negative controls) as indicated. Nuclei were revealed by DAPI staining and appear in blue. White dots (PLA signals) indicated interaction (close proximity) between RBM25 protein and *BCL-x* RNA. The graph on the right indicates the number of PLA dots per cell in each condition. Data from three biological replicates, at least 250 cells per sample, were analysed by a non-parametric Mann–Whitney's test using the GraphPad Prism 8 Software (***$P < 0.0001$).

GQ-2 rG4 is required for both RBM25 binding and the synthesis of the alternative pro-apoptotic Bcl-xS isoform

We next designed two minigenes to monitor the alternative splicing of *BCL-x* and the effect of suppressing the ability of GQ-2 to form an rG4 structure on synthesis of the alternative product Bcl-xS and on the ability of RBM25 to interact with *BCL-x* RNA. First, based on an already published minigene (34,64,65), we constructed Bcl-x 672 WT (wild-type), a wt minigene that encompasses most of exon 2 (80% including all its 3′ part), part of intron 2 and the beginning of exon 3 (Figure 3A and B). Then, using site-directed mutagenesis we created a mutated version of this *WT* minigene (termed Bcl-x 672 GM-2, Figure 3B) by introducing the same modifications as in GM-2 (Figure 1B) making it unable of G4 formation (Figure 1C). We first assessed, using semi-quantitative RT–PCR, the alternative splicing from these two minigenes transfected in HeLa cells. As shown in Figure 3C, Bcl-x 672 WT led to ∼30% of Bcl-xS and ∼70% of Bcl-xL, in line with previously published results (34). By contrast, the Bcl-x 672 GM-2 minigene exclusively led to the synthesis of the canonical full length Bcl-xL isoform, readily suggesting that GQ-2 rG4 is important for alternative splicing of *BCL-x* and the synthesis of the alternative short pro-apoptotic Bcl-xS isoform.

Next, to further evaluate the validity of our minigene and the role of RBM25 binding on the GQ-2 rG4 in the production of the RNA encoding the alternative Bcl-xS isoform, we tested the effect of RBM25 overexpression on alternative splicing from either Bcl-x 672 WT or Bcl-x 672 GM-2 minigene. For this purpose, we transfected HeLa cells with both a plasmid that allows expression of Bcl-x 672 WT or Bcl-x 672 GM-2 minigene and increasing quantities (1 or 3 μg) of a plasmid allowing ectopic expression of a N-terminally HA-tagged version of RBM25 (HA-RBM25-FL). As shown in Figure 3D, the fraction of Bcl-xS RNA produced from the Bcl-x 672 WT minigene increased from 22 to 42% (with 1 μg of HA-RBM25-FL plasmid) and 60% (with 3 μg of HA-RBM25-FL plasmid), in good correlation with the increase in RBM25 (as assessed by comparing to the GAPDH loading control) and in line with previously published data (34,48). By contrast, irrespective of the level of RBM25, no Bcl-xS RNA was detected from the Bcl-x 672 GM-2 minigene, suggesting that the GQ-2 rG4/RBM25 interaction is required for synthesis of the alternative Bcl-xS isoform. To further test this possibility, we adapted the PLA shown in Figure 2E to specifically assess the interaction between endogenously expressed RBM25 and pre-mRNA expressed from either Bcl-x 672 WT or Bcl-x 672 GM-2 minigene transfected in H1299 cells. GFP was used as a marker to detect transfected cells. As shown in Figure 3E, the nuclei of GFP^+ cells transfected by the Bcl-x 672 WT minigene contained multiple PLA dots (median value of 7.5 dots per cell) whereas the nuclei of the non-transfected GFP^- cells did not. By contrast, the nuclei of GFP^+ cells transfected by the Bcl-x 672 GM-2 minigene contained only a few dots (median value of 1 dot per cell), which in addition were less intense and similar to the dots detected in non-transfected GFP^- cells. Additional controls are shown in Supplementary Figure 5. Altogether, these results further confirm that RBM25 interacts with GQ-2 rG4 present in the *BCL-x* pre-mRNA and that this interaction is necessary for the production of the alternative Bcl-xS isoform.

The RE domain of RBM25 is necessary for binding to GQ-2 rG4 and defines a new rG4-binding motif

We next sought to identify which domain(s) of RBM25 is involved in its ability to interact with the GQ-2 rG4 of *BCL-x* pre-mRNA. As shown in Figure 4A, RBM25 contains three domains that may potentially be involved in binding to nucleic acids: an N-terminal RRM (RNA recognition motif) domain, a central RE (arginine-glutamate-rich) motif and a C-terminal PWI domain. To determine which of these three domains is important for binding of RBM25 to GQ-2 G4, we overexpressed HA-tagged versions of RBM25 deleted, or not, for one of the other of these three domains in H1299 cells (hence they are respectively termed HA-RBM25-ΔRRM, HA-RBM25-ΔPWI, HA-RBM25-ΔRE-rich and HA-RBM25-FL, depicted in Figure 4A) and then tested their ability to interact with GQ-2 G4 using the same RNA pulldown assay used in Figure 2A to C. Of note, the RE domain is composed of two sub-domains: one that is RE-rich, whereas the other is significantly less enriched in RE (termed RE-poor in Figure 4A). To minimize the size of the deletion, we choose to delete only the RE-rich part, hence the name of the construct: HA-RBM25-ΔRE-rich. As shown in Figure 4B, whereas HA-RBM25-FL, HA-RBM25-ΔPWI and HA-RBM25-ΔRRM interact in a G4-dependent manner with GQ-2 G4, only a residual binding barely superior to the residual binding on control GM-2 matrix was observed for HA-RBM25-ΔRE-rich. To confirm this result, we repeated this RNA pulldown experiment using protein extracts from H1299 cells transfected with HA-RBM25-ΔRE-rich and used an anti-RBM25 antibody to detect both endogenous and overexpressed HA-tagged RBM25. As shown in Figure 4C, endogenous RBM25 was clearly enriched on the GQ-2 matrix and only a residual binding was observed on the control GM-2 matrix, whereas only a residual binding of HA-RBM25-ΔRE-rich occurred on the GQ-2 matrix, which was barely superior to the one observed on the control GM-2 matrix. These results indicate that the RE motif is important for the ability of RBM25 to bind GQ-2 rG4 and readily suggests that the RE domain is directly mediating this interaction. To test this possibility, we expressed the HA-tagged RE motif (HA-RE), the RE-rich part only (HA-RE-rich) or HA-RBM25-FL (as a positive control) in H1299 cells and determined their ability to bind to the GQ-2 matrix compared to the control GM-2 matrix. As shown in Figure 4D, HA-RE is able to bind to GQ-2 in a G4-dependent manner, confirming that the RE motif is directly mediating the interaction of RBM25 with GQ-2 rG4. As for the HA-RE-rich construct, perhaps because it is most probably an intrinsically disordered motif, only a weak binding to GQ-2 RNA oligonucleotide was observed. However, even this weak binding was G4-dependent as no significant binding was observed with GM-2 oligonucleotide. Altogether, these results suggest that RE motif represents a new rG4-binding domain.

Next, we determined the effect of overexpressing HA-RBM25ΔRE-rich, as compared to HA-RBM25-FL, on splicing from the Bcl-x 672 WT minigene, using similar experiments as in Figure 3D. We observed that, although expressed at a level similar to HA-RBM25-FL (as determined by comparison to the GAPDH loading control, lower panels of Figure 4E), overexpression of HA-RBM25-ΔRE-rich only had a modest effect on alternative splicing (the proportion of Bcl-xS increasing from 24% in the negative control transfected by an empty plasmid to 27 and 35% in cells transfected with 1 or 3 μg, respectively, of the HA-RBM25-ΔRE-rich plasmid),

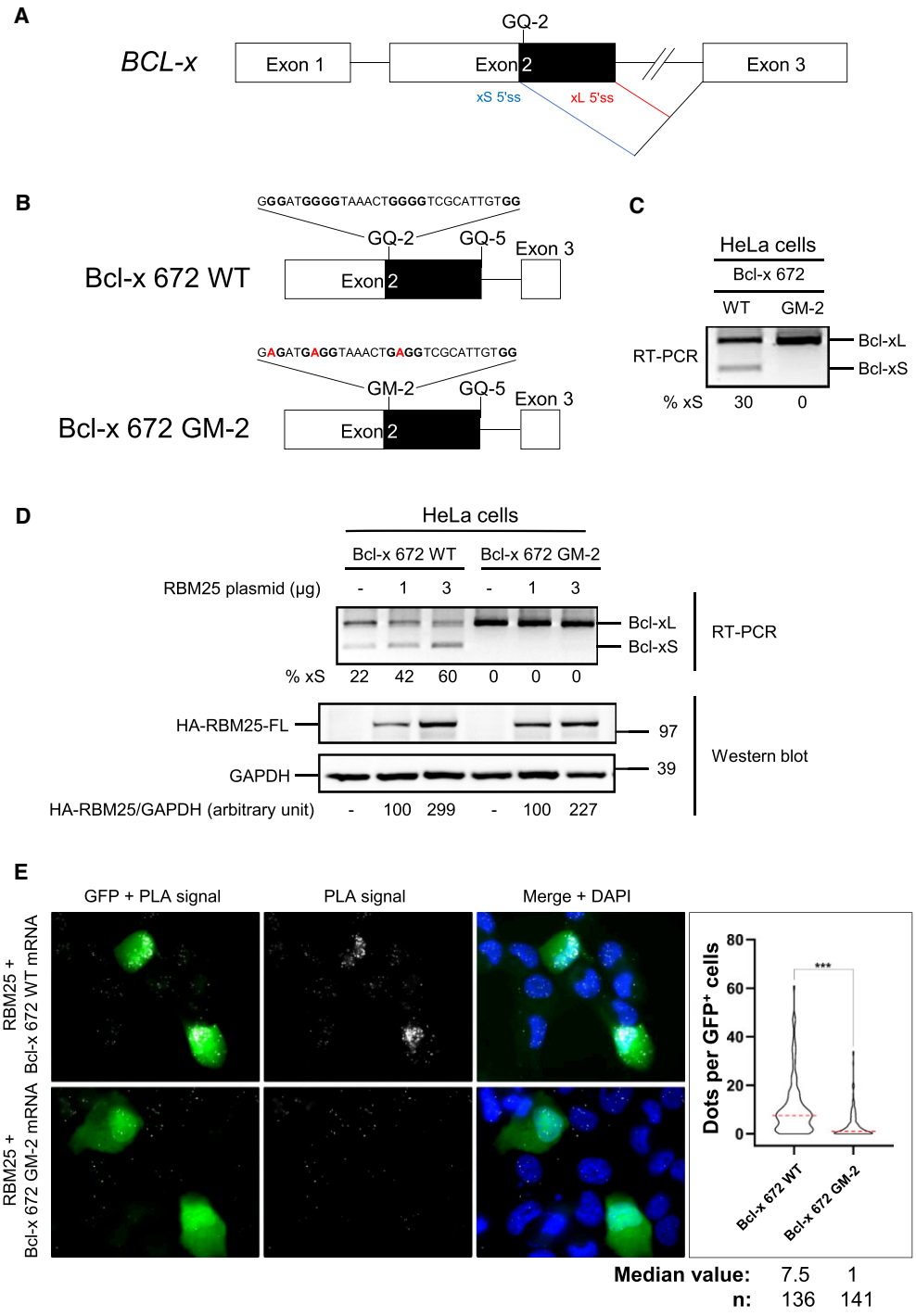


Figure 3. GQ-2 rG4 is required for both RBM25 binding and synthesis of the alternative pro-apoptotic Bcl-xS isoform (**A**) Schematic representation of the *BCL-x* gene with the two alternative 5′ splice sites in the exon 2 indicated (xS 5′ss in blue and xL 5′ss in red) as well as the GQ-2 rG4. (**B**) Schematic representation of the Bcl-x 672 WT (upper panel) and Bcl-x 672 GM-2 (lower panel) minigenes. The only differences between the WT and the GM-2 minigenes are the replacement of three guanines critical for GQ-2 rG4 formation by three adenines (highlighted in red). (**C**) Semi-quantitative RT–PCR experiment from HeLa cells transfected by the WT or GM-2 minigene, as indicated, to determine the relative proportion of the Bcl-xS and Bcl-xL isoforms synthesized from the minigenes. The PCR products were analysed by electrophoresis in a 1% agarose gel, revealed by an ethidium bromide staining and quantified by densitometry. An image of a resulting gel is shown ($n \geq 3$). (**D**) Same semi-quantitative RT–PCR experiment as in (**C**) except that, in addition to the WT or GM-2 minigene, HeLa cells were also transfected by the indicated quantities of a plasmid allowing expression of HA-tagged RBM25 wt (HA-RBM25-FL) or, as a control, by an empty plasmid. The result of the semi-quantitative RT–PCR experiment is shown in the upper panel and western blot analysis of the level of expression of HA-RBM25-FL, as compared to the loading control GAPDH, is shown in the two lower panels ($n \geq 3$). (**E**) Same PLA experiments as in Figure 2E, except that H1299 cells were transfected by a GFP-encoding plasmid that also express the WT or the GM-2 minigene, as indicated. Microscopy images of cells analysed using a probe specifically hybridizing to the *BCL-x* RNA transcribed from the minigenes. GFP signal indicates cells that were efficiently transfected and nuclei were revealed by DAPI staining and appear in blue. White dots (PLA signals) indicated interaction (close proximity) between the RBM25 protein and the *BCL-x* RNA transcribed from the minigene. The graph on the right indicates the number of PLA dots per cells in each condition. Data from two biological replicates, at least 35 GFP^+ cells per sample, were analysed by a non-parametric Mann–Whitney's test using the GraphPad Prism 8 Software (***$P < 0.0001$).

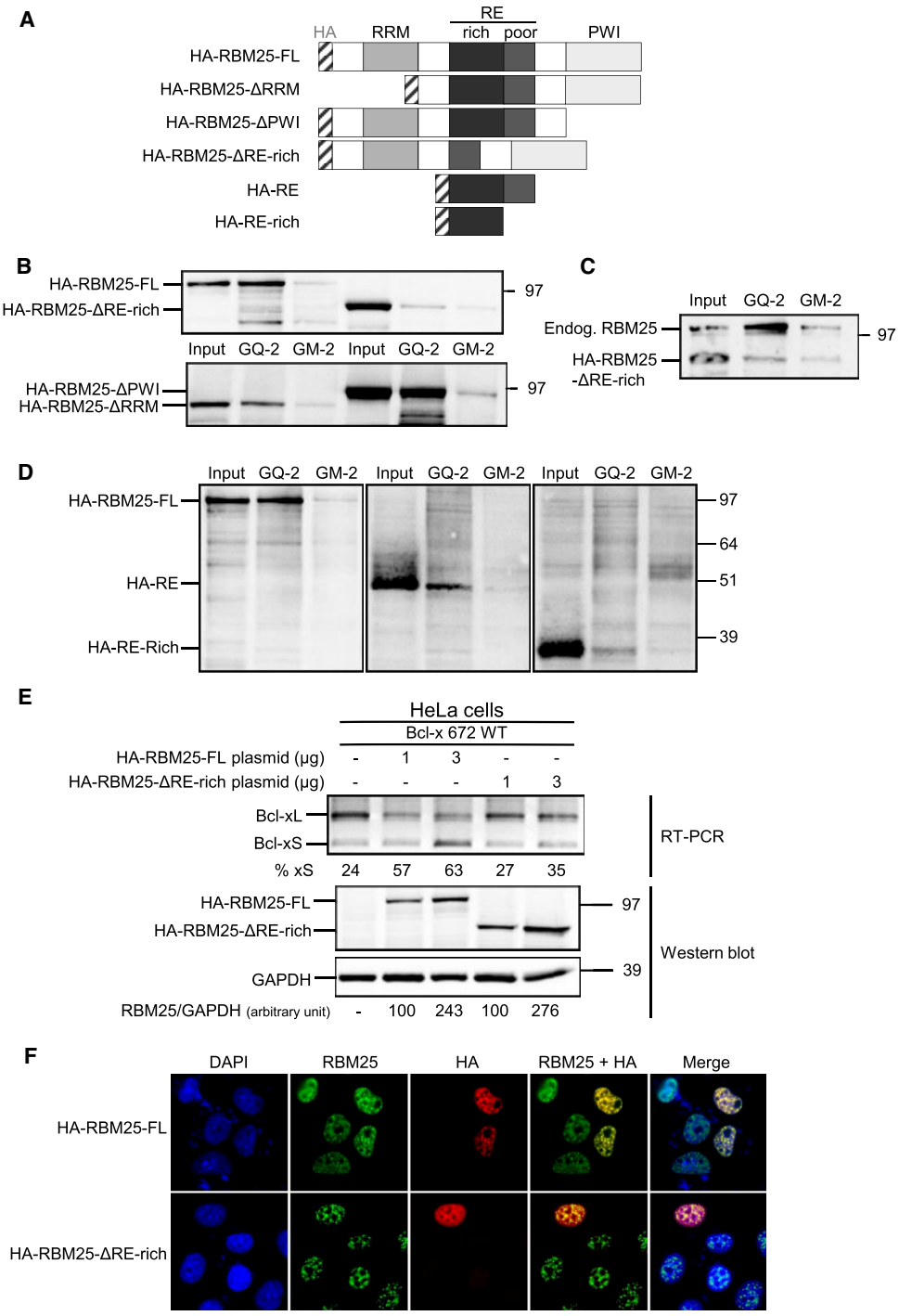


Figure 4. The RE domain of RBM25 is necessary and sufficient for binding to GQ-2 rG4. (**A**) Schematic representation of the RBM25 protein with its three motifs known to potentially interact with nucleic acids (RRM, RE and PWI) and of the various constructions used for the determination of the role of each of these three domains in binding to GQ-2 rG4. All the constructs are N-terminally HA-tagged (HA tag written in grey and represented as a hatched box). (**B**) Same RNA pulldown experiments as in Figure 2B using the GQ-2 or GM-2 matrices as indicated except that, in addition to H1299 cells transfected with an HA-tagged RBM25 (HA-RBM25-FL)-encoding plasmid, H1299 cells transfected with one or the other of the plasmids expressing the various HA-tagged forms of RBM25 deleted for one of its three domains were also used. The proteins still bound after an 800 mM KCl wash were eluted and analysed by SDS–PAGE and western blot using an anti-HA antibody to reveal the exogenously expressed various forms of HA-tagged RBM25. Gels represent $n \geq 3$. (**C**) Same experiment as in (**B**) using only the HA-RBM25-ΔRE-rich plasmid except that an anti-RBM25 antibody was used for the western blot analysis to reveal both exogenously expressed HA-RBM25-ΔRE-rich and endogenously expressed RBM25 that thus serves as an internal positive control. (**D**) Same experiment as in (**B**) except that a plasmid allowing expression of a HA-tagged full length RE motif (HA-RE) or only the RE-rich part of RE motif (HA-RE-rich) was used. (**E**) Same experiment as in Figure 3D except that, in addition to the effect of overexpressing HA-RBM25-FL on alternative splicing from the Bcl-x 672 WT minigene, the effect of overexpressing HA-RBM25-ΔRE-rich was also assessed. The result of the semi-quantitative RT–PCR experiment is shown in the upper panel and the western blot analysis of the level of expression of HA-RBM25-FL or HA-RBM25-ΔRE-rich, as compared to the loading control GAPDH, is shown in the two lower panels. Gels represent $n \geq 3$. (**F**) Immunofluorescence analysis of H1299 cells expressing HA-RBM25-FL or HA-RBM25-ΔRE-rich, as indicated. Microscopy images of fixed cells analysed using anti-RBM25 or anti-HA antibodies were shown as indicated. Merged images were also shown.

which is in line with its limited ability to interact with GQ-2 rG4, as compared to the strong effect of HA-RBM25-FL (for which the proportion of Bcl-xS increased from 24% in the negative control to 57 and 63% in cells transfected with 1 or 3 μg, respectively, of the HA-RBM25-FL plasmid) (Figure 4E, upper panel). We have also determined the effect of over-expressing only HA-RE on splicing from the Bcl-x 672 WT minigene and found it had no effect (Supplementary Figure 6), suggesting that other domains of RBM25 are required for an effect on alternative splicing, most probably by recruiting various splicing factors.

Because the altered ability of HA-RBM25-ΔRE-rich to induce alternative splicing of RBM25 could be due to a change in its intracellular localization, we determined, using immunofluorescence, the localization of both HA-RBM25-FL and HA-RBM25-ΔRE expressed in H1299 cells. As shown in Figure 4F, we found a similar nuclear localization for both HA-RBM25-ΔRE and HA-RBM25-FL, in line with already published results (48). Hence, deleting the RE motif of RBM25, in addition to interfering with its capacity to interact with GQ-2 rG4 in *BCL-x* mRNA, also strongly affects its ability to induce alternative splicing of *BCL*-x. We conclude that interfering with the ability of RBM25 to interact with GQ-2 rG4 by deleting its RE domain has the same effect as interfering with this interaction by preventing rG4 formation through replacement of the three guanines critical for G4 formation in GQ-2 with adenines. Altogether, these results clearly indicate that the RE-mediated interaction between RBM25 and GQ-2 rG4 of *BCL-x* pre-mRNA is critical for the synthesis of the alternative pro-apoptotic Bcl-xS isoform.

Effect of the G4 ligand PhenDC3 on RBM25/*BCL-x* RNA interaction and the synthesis of the alternative Bcl-xS isoform

Given the importance of GQ-2 rG4 on *BCL-x* alternative splicing, we next tested the effect of PhenDC3, a benchmark G4 ligand (66), on the splicing from the Bcl-x 672 WT minigene transfected into HeLa cells. As shown in Figure 5A, treatment with increasing concentration of PhenDC3 (from 0 to 40 μM) led to a dose-dependent increase (from 39 to 71%) in the proportion of the alternative Bcl-xS isoform while having no effect on Bcl-x 672 GM-2 minigene for which no alternative splicing was observed with or without PhenDC3.

To assess whether this effect was also observed for the endogenous *BCL-x* gene, we tested the effect of PhenDC3 on the alternative splicing of *BCL-x* in A549 cells. In this cell line, as in other cancer-derived cell lines, the level of alternative Bcl-xS is barely detectable using RT–PCR and represents only ∼6% of total Bcl-x in this experiment. However, increasing concentrations of PhenDC3 (from 0 to 40 μM) led to a dose-dependent increase in the alternative Bcl-xS RNA that reached ∼15% at 40 μM PhenDC3 (Figure 5B), further confirming that PhenDC3 induces the alternative splicing of *BCL-x* and suggesting that this G4 ligand favours RBM25 binding to GQ-2 rG4 and thereby the production of the alternative pro-apoptotic Bcl-xS isoform. To test this hypothesis, using similar PLA experiments as in Figure 2E, we determined the effect of PhenDC3 on the interaction between endogenous RBM25 and endogenous *BCL*-x RNA in H1299 cells. We observed (Figure 5C) that a treatment with 20 μM PhenDC3 significantly increases the number of PLA dots (median value of 11 dots per cell, most of them being nuclear, as compared to a median value of six dots per cell for the DMSO-treated cells). We concluded that the G4 ligand PhenDC3 favours the binding of RBM25 to GQ-2 rG4 of *BCL-x* RNA, readily explaining its ability to boost alternative splicing of *BCL-x*.

We then tested the effect of PhenDC3 on the stability of GQ-2 RNA G4. For this purpose, we performed UV-melting experiments. In a buffer containing 1 mM K^+, the addition of 1 or 2 molar equivalents of PhenDC3 resulted in an increase of melting temperature of GQ-2 from 37.4 to 71.3°C ($\Delta T_m = 33.9$°C, Figure 5D) with 2 molar equivalents giving evidence of a strong stabilizing effect. In buffers with higher K^+ content (10 or 100 mM) the ligand-induced stabilization prevented the melting of G4 at temperatures below 95°C, prohibiting the precise determination of T_m values. This behaviour is in line with the known capacity of PhenDC3 to strongly stabilize various DNA and RNA G4 (66–69). Hence, it is likely that PhenDC3 also stabilizes GQ-2 rG4 in native *Bcl-x* mRNA, or promotes the formation of an rG4 at the expense of alternative secondary structures, thereby favouring RBM25 binding and consequently production of the alternative pro-apoptotic Bcl-xS isoform.

For this reason, we tested the effect of PhenDC3 on the induction of apoptosis in A549 cells only expressing endogenous Bcl-x. Using the *Caspase Glo® 3/7 Assay* which allows to determine the relative caspase 3/7 activity, and using the BH3-mimetics Navitoclax and ABT-737 (70,71) as positive controls, we found that PhenDC3, in line with its ability to induce Bcl-xS, does induce apoptosis (Figure 5E), as compared to the negative control (DMSO 0.4%; compound vehicle).

Screening of a library of 90 putative G4 ligands leads to the identification of two original and more active modulators of *BCL-x* splicing

The ability of PhenDC3 to enhance the binding of RBM25 to rG4 in *BCL-x* pre-mRNA, and thereby to promote the alternative pro-apoptotic Bcl-xS isoform and apoptosis indicates that favouring the interaction between RBM25 and the GQ-2 rG4 in *BCL-x* pre-mRNA represents a relevant intervention point to re-sensitize cancer cells to chemotherapy. Hence, we screened a combinatorial library of 90 'as-synthesized' cationic bis(acylhydrazones) **Aa-Ji** (Figure 6A and Supplementary Figure 7), structurally related to PhenDC3 and previously validated as putative G4 ligands (50) using a fixed concentration of ∼10 μM and the splicing from the Bcl-x 672 WT minigene transfected into HeLa cells as a readout (Figure 6B and Supplementary Figure 8). From these 90 compounds, two derivatives containing a phenanthroline fragment (chemical structures: Figure 6C) significantly increased the Bcl-xS ratio (**Ef**, or PhenDH8: 53% xS; **Ei**, or PhenDH9: 51% xS). Of note, several structurally close phenanthroline derivatives, such as **Eb** (PhenDH2), were inactive. We then tested pure samples of PhenDH8, PhenDH9, PhenDC3 and PhenDH2 (as inactive control) at various concentrations (from 0 to 40 μM) to rank them in terms of efficacy. We found the following order of potency: PhenDH8 > PhenDH9 > PhenDC3 > PhenDH2, the latter being only weakly active at the highest concentration tested (40 μM), whereas PhenDH8 was already fully active at the lowest concentration (10 μM) (Supplementary Figure 9). Next, we tested the effect of these four compounds on the alternative splicing of the endogenous *BCL-x* gene in A549 cells. As shown in Figure 6D, we obtained similar results than when using the minigene, with the same order of potency,

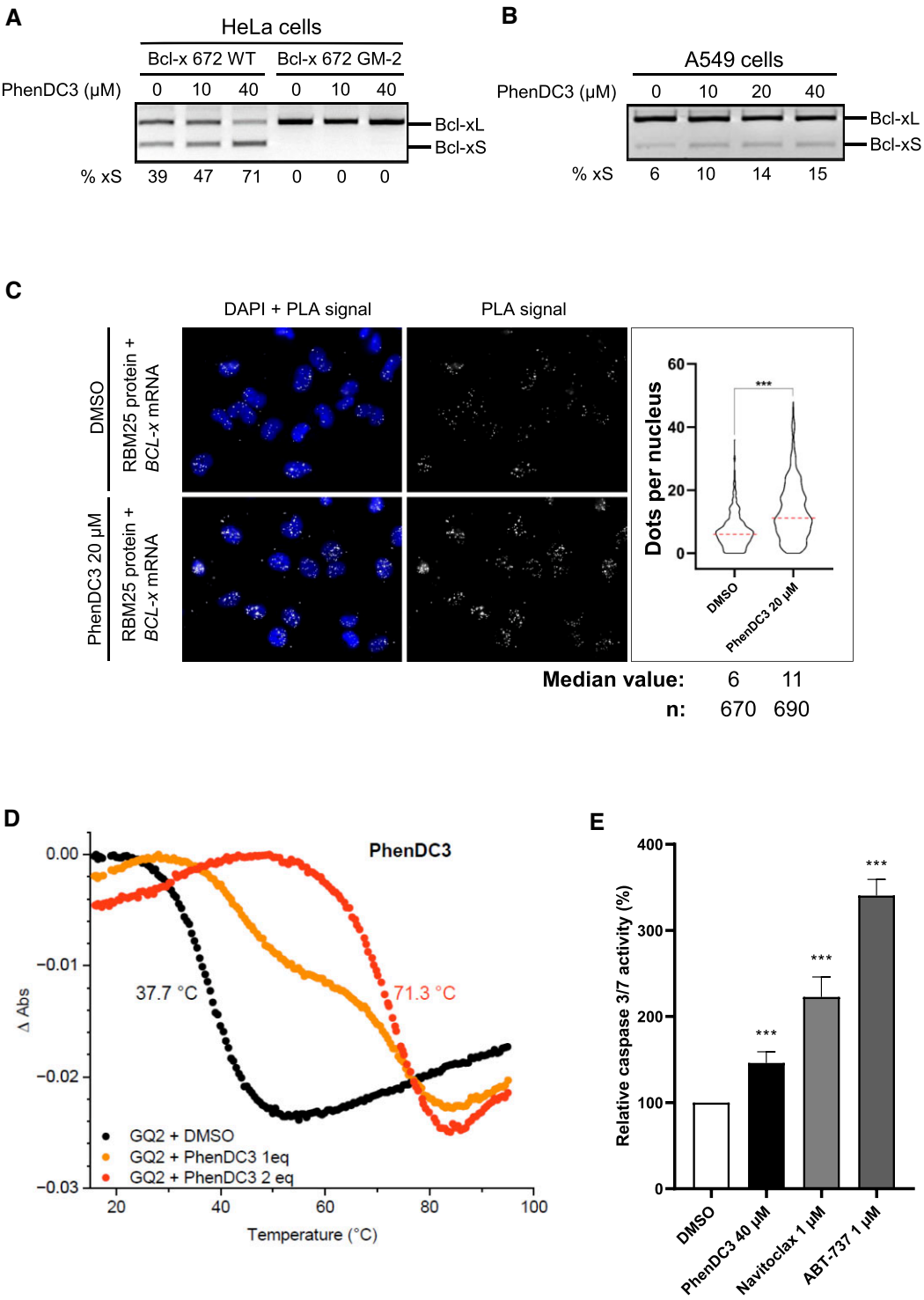


Figure 5. Effect of the PhenDC3 G4 ligand on RBM25/*BCL-x* RNA interaction, synthesis of the alternative pro-apoptotic Bcl-xS isoform and apoptosis. (**A**) Same semi-quantitative RT–PCR experiment as in Figure 3C to determine the relative proportion of the Bcl-xS and Bcl-xL isoforms synthesized from the minigenes, except that HeLa cells transfected by the WT or GM-2 minigene (as indicated) were also treated, or not, by various concentrations of the benchmark G4 ligand PhenDC3. Gel represents $n \geq 3$. (**B**) Same experiment as in (**A**) except that the alternative splicing of the endogenous *BCL-x* gene was analysed in A549 cells. Gel represents $n \geq 3$. (**C**) Same PLA experiment as in Figure 2E except that cells were treated, or not, with 20 μM PhenDC3. Microscopy images of H1299 cells analysed using a probe specifically hybridizing to the *BCL-x* RNA. Nuclei were revealed by DAPI staining and appear in blue. White dots (PLA signals) indicated interaction (close proximity) between endogenous RBM25 protein and endogenous *BCL-x* RNA. The graph on the right indicates the number of PLA dots per cells in each condition. Data from three biological replicates, at least 200 cells per sample, were analysed by a non-parametric Mann–Whitney's test using the GraphPad Prism 8 Software (***$P < 0.0001$). (**D**) Normalized UV-melting curves of GQ-2 (4 μM in lithium cacodylate buffer supplemented with 1 mM KCl and 99 mM LiCl) in the absence (black) or in the presence of 1 molar equivalent (orange) of 2 molar equivalents (red) of PhenDC3. (**E**) Determination, using the *Caspase Glo 3/7 Assay*, of the relative caspase 3/7 activity in A549 cells treated with DMSO (vehicle, negative control), 40 μM PhenDC3 or 1 μM Navitoclax or 1 μM ABT-737 as positive controls. Data from four biological replicates were analysed by ANOVA in conjunction with Tukey's test using GraphPad Prism 8 Software (***$P < 0.0001$).

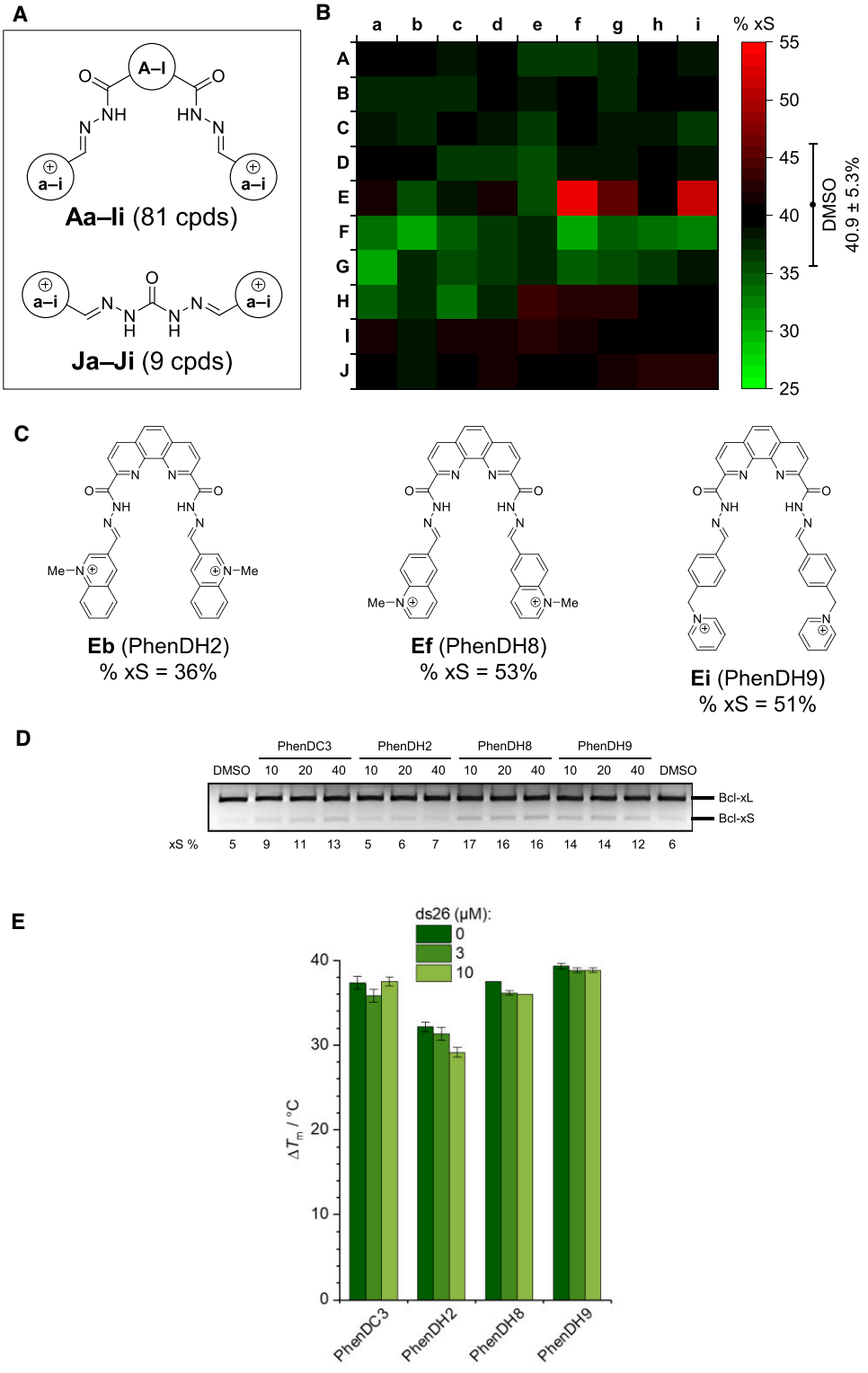


Figure 6. Screening of a library of 90 putative G4 ligands and identification of PhenDH2, 8 and 9. (**A**) Generic structure of the combinatorial library of 90 cationic bis(acylhydrazones) (**Aa**–**Ji**). (**B**) The effect of individual library members (at ∼10 μM) on alternative splicing from the Bcl-x 672 WT minigene, as determined by semi-quantitative RT–PCR experiments on HeLa cells transfected by the Bcl-x 672 WT minigene was determined. The percentage of Bcl-xS obtained for each compound is indicated as a heatmap (% xS). PhenDC3 at the same concentration was used as a positive control. (**C**) Chemical structures of hit compounds **Ef** (PhenDH8), **Ei** (PhenDH9) and the inactive analogue **Eb** (PhenDH2). (**D**) Semi-quantitative RT–PCR experiments to assess alternative splicing of the endogenous *BCL-x* gene in A549 cells submitted to increasing concentrations of PhenDC3, PhenDH8, PhenDH9 or PhenDH2, as indicated. PhenDH8 and 9 are more potent than PhenDC3 to induce the alternative pro-apoptotic Bcl-xS whereas PhenDH2, although structurally close, is inactive at all tested concentrations. Gel represents $n \geq 3$. (**E**) Ligand-induced stabilization of GQ2 observed in FRET-melting experiments, performed with double-labelled oligonucleotide (F-GQ2-T, 0.2 μM) in the presence of the ligands (1.0 μM each) and in the absence or in the presence of double-stranded DNA competitor (ds26: 0, 3 or 10 μM). Experiments were performed in 10 mM lithium cacodylate buffer (pH 7.2) supplemented with 10 mM KCl and 90 mM LiCl.

confirming that the Bcl-x 672 WT minigene is a relevant tool to model *BCL-x* alternative splicing.

To rationalize the results obtained with these compounds, we assessed their capacity to stabilize the rG4 structure formed by GQ-2 by FRET-melting experiments. The results (Figure 6E) indicated that PhenDH8 was as effective as PhenDC3 in stabilizing rG4 ($\Delta T_m = 37.5°C$), PhenDH9 was slightly more effective ($\Delta T_m = 39.3°C$), and PhenDH2 significantly less effective than PhenDC3 ($\Delta T_m = 32.2°C$). Moreover, and in contrast to PhenDC3, PhenDH8 and PhenDH9, the stabilization effect induced by PhenDH2 was further reduced when the experiment was performed in the presence of double-stranded DNA (ds26) as a competitor, indicating that in cellular conditions this ligand has poor capacity to stabilize rG4 in *BCL-x* pre-mRNA. Taken together, these experiments shed light on the relative capacity of the ligands to modulate *BCL-x* splicing.

Next, we tested the effect of a range of lower concentrations of the various G4 ligands on alternative splicing of the endogenous *BCL-x* gene (Figure 7A). This allowed us to estimate the EC_{50} (half-maximum effective concentration) for the three active G4 ligands, thereby confirming their order of potency: PhenDH8 ($EC_{50} \approx 0.2$ μM) > PhenDH9 ($EC_{50} \approx 0.7$ μM) > PhenDC3 ($EC_{50} \approx 2$ μM) (Figure 7B). Then, to determine whether the effect of these G4 ligands on alternative splicing of the endogenous *BCL-x* affects the anti-apoptotic Bcl-xL protein level, we performed western blots using an antibody specific for the Bcl-xL isoform. As shown in Figure 7C, PhenDH8 and PhenDH9 led to a significant decrease in the level of Bcl-xL protein, which is consistent with their effect on the alternative splicing of the *BCL-x* gene. By contrast, PhenDH2 had no effect on Bcl-xL level, in line with its lack of effect on splicing of *BCL-x*.

The ability of PhenDH8, PhenDH9 and PhenDC3 to affect the alternative splicing of *BCL-x* and to promote the alternative pro-apoptotic Bcl-xS isoform being confirmed, we then tested their ability to induce apoptosis, using PhenDH2 and DMSO (the vehicle) as negative controls and Navitoclax and ABT-737, two BH3-mimetics (70,71), as positive controls. For this purpose, we used the same Caspase Glo® 3/7 Assay as in Figure 5E, which allows to determine the relative caspase 3/7 activity. As shown in Figure 7D, we first confirmed that PhenDC3 induces apoptosis. We also observed that PhenDH8 and PhenDH9 were even more active, whereas PhenDH2 was inactive. Importantly, we got the same order of potency than for the ability of the compounds to induce the synthesis of the alternative pro-apoptotic Bcl-xS isoform.

Finally, using the MTT assay, we measured the ability of the compounds to induce cell death (Figure 7E). We obtained the same result than in the two splicing assays and in the Caspase Glo® 3/7 Assay. To determine whether the compounds have the same order of potency than on alternative splicing, we repeated the MTT assays with a range of concentrations for each of the three active G4 ligands (Figure 7F). This allowed us to estimate the GI_{50} (the concentration of the agent that inhibits growth by 50%) for each of these compounds. We found again that PhenDH8 was the most active compound, followed by PhenDH9 and then by PhenDC3.

Taken together, these results indicate that PhenDC3, a benchmark G4-ligand, and its two derivatives PhenDH8 and PhenDH9, induce both the alternative pro-apoptotic Bcl-xS isoform and apoptosis and that the intensity of their effect on apoptosis and cell death is correlated with their effect on alternative splicing.

Discussion

In this paper, using RNA pulldown and PLA adapted for monitoring protein–RNA interactions, we first show that RBM25 binds directly and specifically to the GQ-2 rG4 of *BCL-x* RNA. We also show that this interaction is G4 structure-dependent and involves the central RE motif of RBM25, thereby defining a new G4-interacting domain. In addition, we found that this RBM25/GQ-2 rG4 interaction is necessary for alternative splicing of *BCL-x* and is a limiting factor for the synthesis of the pro-apoptotic Bcl-xS isoform. Indeed, interfering with this interaction, either by preventing G4 formation (GM-2 mutant) or by deleting the RE motif of RBM25, abolishes the synthesis of Bcl-xS, the pro-apoptotic alternatively spliced form of Bcl-x. On the contrary, favouring the RBM25/GQ-2 rG4 interaction by overexpressing RBM25 or by using PhenDC3, a G4 ligand that strongly stabilizes GQ-2 rG4 *in vitro*, promotes the synthesis of Bcl-xS and apoptosis. Importantly, this last result indicates that the RBM25/GQ-2 rG4 interaction represents an intervention point to modulate the balance between Bcl-xL and Bcl-xS, the two antagonistic forms of Bcl-x. This is of particular interest for the various disorders linked to overexpression of one or the other of Bcl-x isoforms, in particular for cancers in which overexpression of the anti-apoptotic Bcl-xL isoform has been associated to resistance to chemotherapy (15). Hence, small molecules that promote the pro-apoptotic alternative Bcl-xS isoform (such as PhenDC3, PhenDH8 and PhenDH9) may represent prototypes of candidate drugs to re-sensitize cancer cells to chemotherapy. Importantly, the fact that PhenDC3, PhenDH8 or PhenDH9 promote both the Bcl-xS isoform and apoptosis is not a proof of a mechanism in which they would induce apoptosis as a consequence of their effect on *BCL-x* alternative splicing as these two effects could be independent. However, the fact that the order of potency of these compounds on apoptosis (PhenDH8 > PhenDH9 > PhenDC3 > PhenDH2) is the same than on induction of the pro-apoptotic Bcl-xS isoform, and that PhenDH2, their close chemical derivative, is inactive on both processes suggest that a causal link exists between induction of Bcl-xS and of apoptosis by our active G4 ligands. Of note, novel molecular tools have recently been developed that rather destabilize G4 (72). According to the results presented here, these new compounds may thus represent therapeutic options to interfere with the production of the alternative pro-apoptotic Bcl-xS isoform whose overexpression has been linked to various disorders such as some forms of diabetes in which it has been associated to exacerbated apoptosis of β cells in the islets of Langerhans (15).

Important questions remain, such as the precise mechanism of action of RBM25 in *BCL-x* alternative splicing. At least two possible models could be envisioned. In the first model, the GQ-2 rG4 may represent a protein-binding platform recognized by the splicing factor RBM25 that itself might promote the selection of the xS 5′ss by the splicing machinery, thereby leading to the synthesis of the alternative pro-apoptotic Bcl-xS isoform. In the second model, the main role of RBM25 in *BCL-x* alternative splicing could be to promote and/or stabilize GQ-2 rG4 formation which would be required for the

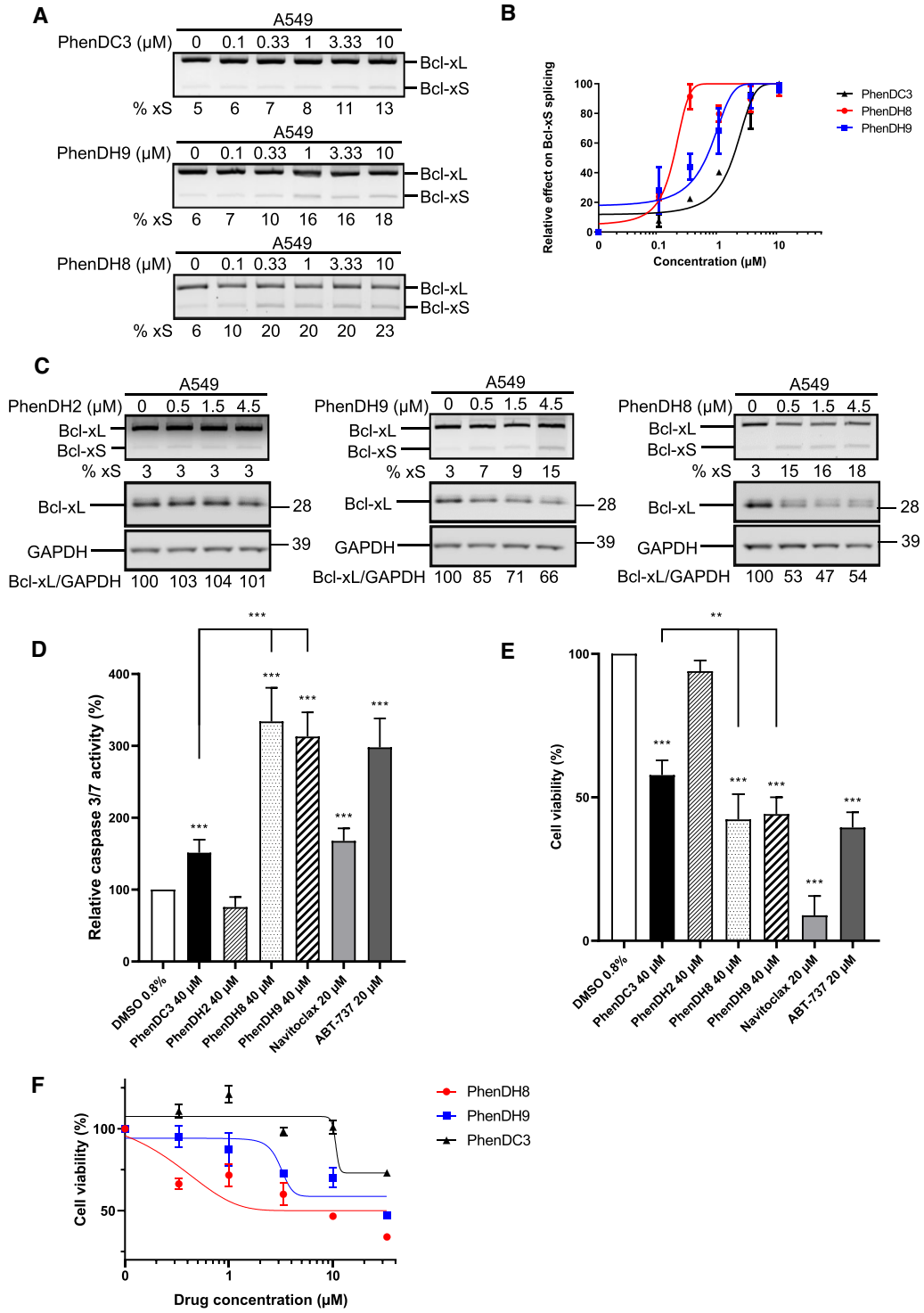


Figure 7. Quantitative analysis of the effect of PhenDC3, PhenDH8 and PhenDH9 on alternative splicing of *BCL-x*, apoptosis and cytotoxicity. (**A**) Semi-quantitative RT–PCR experiments to assess alternative splicing of the endogenous *BCL-x* gene in A549 cells submitted to increasing concentrations of PhenDC3, PhenDH9 or PhenDH8, as indicated. Gels represent $n \geq 3$. (**B**) Semi-logarithmic representation of the data shown in (**A**). (**C**) Western blot analysis of the effect of a range of concentration of PhenDH2, PhenDH9 or PhenDH8 (as indicated) on the protein level of the anti-apoptotic isoform Bcl-xL (middle gels) as compared to GAPDH (loading control, lower gels). The impact of the various G4 ligands on the alternative splicing of the endogenous *BCL-x* gene was determined in the same samples using semi-quantitative RT–PCR experiments (upper gels) as in (**A**). Gels represent $n \geq 3$. (**D**) Effect of PhenDC3, PhenDH8, PhenDH9 or PhenDH2 (all at 40 μM) on apoptosis as determined using the *Caspase Glo 3/7 Assay* as in Figure 5E. DMSO (vehicle) was used as a negative control and Navitoclax and ABT-737 were used as positive controls. Data from three biological replicates were analysed by ANOVA in conjunction with Tukey's test (***$P < 0.0001$). (**E**) Effect of PhenDC3, PhenDH8, PhenDH9 or PhenDH2 (all at 40 μM) on cell viability as determined using the MTT assay. DMSO (vehicle) was used as a negative control and Navitoclax and ABT-737 were used as positive controls. Data from three biological replicates were analysed by ANOVA in conjunction with Tukey's test using GraphPad Prism 8 Software (**$P < 0.001$, ***$P < 0.0001$). (**F**) Semi-logarithmic representation of the data obtained when testing the effect of a range of concentrations of PhenDC3, PhenDH8, PhenDH9 or PhenDH2 on cell viability as determined using the MTT assay as in (**E**).

selection of the xS 5′ss by the splicing machinery. The first hypothesis is supported by the effects of G4 ligands such as PhenDC3, which are likely to increase the lifetime and/or the fraction of GQ-2 rG4 at the expense of alternative secondary RNA structures, thereby enhancing the binding of RBM25. At the same time, in favour of the second hypothesis, recent *in vitro* data (35) showed that the stabilizing effect of a G4 ligand (GQC-05, an ellipticin derivative) on GQ-2 rG4 could be observed only in the presence of nuclear extract. This suggests that the existence of this rG4 *in cellulo* requires the presence of a nuclear factor, which may be an RNA binding protein such as RBM25, that would promote its formation and thereby its recognition and further stabilization by a G4 ligand. Clearly, further investigations are required to decide between the two hypotheses.

Importantly, whatever the precise mechanism of action of RBM25 in promoting the alternative splicing of *BCL-x*, this mechanism represents an intervention point at which to interfere using notably G4 ligands (73,74). Indeed, several modes of action are possible for G4 ligands: they can either stabilize or destabilize G4, or they can prevent binding of a factor on G4, as shown for PhenDC3, PhenDH2 and PyDH2 that have been reported to compete with nucleolin for binding to rG4 that form in the Esptein–Barr virus-encoded EBNA1 mRNA and in the Kaposi's sarcoma-associated herpesvirus-encoded LANA1 mRNA, thereby interfering with immune evasion of these two oncoviruses (30,31,75).

An inherent limitation in the use of G4 ligands may be their potential lack of specificity given that, in addition to the numerous rG4, there are hundreds of thousands potential DNA G4 that may form in the human genome. Importantly, rG4 are more stable than DNA G4, in particular because rG4 predominantly adopt parallel structures, in contrast to DNA G4 that may fold into either parallel, antiparallel of hybrid conformations. In addition, due to the presence of the complementary strand, DNA G4 are in competition with the formation of DNA duplex by canonical Watson–Crick pairing. Hence, G4 ligands may preferentially target rG4. Also, RNA is inherently more flexible than DNA and may adopt a plethora of alternative conformations involving G4 as well as other structural features (76). Thus, G4 in the RNA context are often in a dynamic equilibrium with other structural isoforms, and even low doses of G4 ligand may have a profound impact on this equilibrium, preventing or leading to the recruitment of the corresponding structure-specific binding factors. Another important point is that, beside the G-quartets, which are the constant elements constitutive of G4, there are other structural characteristics, which vary from one sequence to another: the loops and the flanking regions. These elements ensure that each quadruplex is, in principle, unique within the whole genome and, as such, could be exploited as key targets to gain specificity through, for instance, oligonucleotide (or derivatives)-based approaches. Hence it is not excluded that some specificity toward particular G4 may be reached and various methods to assess G4 ligands specificity have been described (77,78). In line, the screening of various libraries of G4 ligands against some particular G4 clearly indicates that selectivity may be reached (34,52,78) and coupling oligonucleotides complementary to the loops or the flanking sequences to G4 ligands can direct them onto some specific G4 (79). Importantly, the specificity issue is also encountered in drug development programmes based on slicing factors or splicing kinases such as CLK or DYRK1A as the developed compounds may, in principle, act on the (alternative) splicing of thousands of genes.

Another important issue, when interfering with the Bcl-xL/Bcl-xS balance, is the risk of potential side effects, for example when promoting Bcl-xS, given its suspected role in various disorders such as some forms of diabetes or of cardiac disorders. Also, promoting Bcl-xL may favour tumour formation. Hence, the targeting of drugs interfering with the Bcl-xL/Bcl-xS balance to specific organs or to tumour cells is an important point to consider in future drug development programmes.

Finally, other means to modify the Bcl-xL/Bcl-xS balance have been developed, such as splice switching oligonucleotides (SSO), splicing kinases inhibitors or BH3-mimetics (15,70,71). Some of these drug candidates have shown promising effects in various preclinical or clinical trials, in particular BH3-mimetics, but their further development has been limited by their side effects, in particular for BH3-mimetics such as ABT-737 (80) and ABT-263 (Navitoclax, (81)). Hence, the possibility to develop the use of G4 ligands such as PhenDH8 or PhenDH9 in co-treatment with reduced quantities of these drug candidates may represent an appealing possibility to limit their side effects.

Data availability

The data underlyingthis article are either available in the article or in its online supplementary materials, or will be made available from the corresponding author upon reasonable request.

Supplementary data

Supplementary Data are available at NAR Online.

Acknowledgements

We thank Rodrigo Prado Martins (INRAe, Nouzilly) for valuable discussions and critical proofreading of the manuscript and Laurent Maillet and Philippe Juin (Inserm Nantes) for valuable discussions and the kind gift of the Bcl-x null cell line.

Funding

This work was supported by the following grant agencies: 'La Ligue contre le cancer CSIRGO' and 'Fondation pour l'Avenir' to M.B., 'Agence Nationale de la Recherche' (ANR-17-CE07-0004-01, to A.G.) and the French Ministry of Higher Education, Research and Innovation (PhD fellowships to R.L.S. and O.R.). V.T.D. is the recipient of a 'La Ligue contre le cancer Bretagne' and 'Région Bretagne' joint fellowship.

Conflict of interest statement

None declared.

References

1. Berget,S.M., Moore,C. and Sharp,P.A. (1977) Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc. Natl Acad. Sci. USA*, **74**, 3171–3175.

2. Chow,L.T., Gelinas,R.E., Broker,T.R. and Roberts,R.J. (1977) An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell*, **12**, 1–8.
3. Chow,L.T., Roberts,J.M., Lewis,J.B. and Broker,T.R. (1977) A map of cytoplasmic RNA transcripts from lytic adenovirus type 2, determined by electron microscopy of RNA:DNA hybrids. *Cell*, **11**, 819–836.
4. Dvinge,H., Kim,E., Abdel-Wahab,O. and Bradley,R.K. (2016) RNA splicing factors as oncoproteins and tumour suppressors. *Nat. Rev. Cancer*, **16**, 413–430.
5. Scotti,M.M. and Swanson,M.S. (2016) RNA mis-splicing in disease. *Nat. Rev. Genet.*, **17**, 19–32.
6. Boise,L.H., Gonzalez-Garcia,M., Postema,C.E., Ding,L., Lindsten,T., Turka,L.A., Mao,X., Nunez,G. and Thompson,C.B. (1993) bcl-x, a bcl-2-related gene that functions as a dominant regulator of apoptotic cell death. *Cell*, **74**, 597–608.
7. Singh,R., Letai,A. and Sarosiek,K. (2019) Regulation of apoptosis in health and disease: the balancing act of BCL-2 family proteins. *Nat. Rev. Mol. Cell Biol.*, **20**, 175–193.
8. Shamas-Din,A., Kale,J., Leber,B. and Andrews,D.W. (2013) Mechanisms of action of Bcl-2 family proteins. *Cold Spring Harb. Perspect. Biol.*, **5**, a008714.
9. Popgeorgiev,N. and Gillet,G. (2022) Bcl-xL and IP3R interaction: intimate relationship with an uncertain outcome. *Cell Calcium*, **101**, 102504.
10. Rosa,N., Ivanova,H., Wagner,L.E. 2nd, Kale,J., La Rovere,R., Welkenhuyzen,K., Louros,N., Karamanou,S., Shabardina,V., Lemmens,I., *et al.* (2022) Bcl-xL acts as an inhibitor of IP(3)R channels, thereby antagonizing $Ca^{(2+)}$-driven apoptosis. *Cell Death Differ.*, **29**, 788–805.
11. Prudent,J., Zunino,R., Sugiura,A., Mattie,S., Shore,G.C. and McBride,H.M. (2015) MAPL SUMOylation of Drp1 stabilizes an ER/mitochondrial platform required for cell death. *Mol. Cell*, **59**, 941–955.
12. Chen,W. and Li,J. (2021) Alternative splicing of BCL-X and implications for treating hematological malignancies. *Oncol. Lett.*, **22**, 670.
13. Dou,Z., Zhao,D., Chen,X., Xu,C., Jin,X., Zhang,X., Wang,Y., Xie,X., Li,Q., Di,C., *et al.* (2021) Aberrant Bcl-x splicing in cancer: from molecular mechanism to therapeutic modulation. *J. Exp. Clin. Cancer Res.*, **40**, 194.
14. Rohrbach,S., Muller-Werdan,U., Werdan,K., Koch,S., Gellerich,N.F. and Holtz,J. (2005) Apoptosis-modulating interaction of the neuregulin/erbB pathway with anthracyclines in regulating Bcl-xS and Bcl-xL in cardiomyocytes. *J. Mol. Cell Cardiol.*, **38**, 485–493.
15. Stevens,M. and Oltean,S. (2019) Modulation of the apoptosis gene Bcl-x function through alternative splicing. *Front. Genet.*, **10**, 804.
16. Bauman,J.A., Li,S.D., Yang,A., Huang,L. and Kole,R. (2010) Anti-tumor activity of splice-switching oligonucleotides. *Nucleic Acids Res.*, **38**, 8348–8356.
17. Li,Z., Li,Q., Han,L., Tian,N., Liang,Q., Li,Y., Zhao,X., Du,C. and Tian,Y. (2016) Pro-apoptotic effects of splice-switching oligonucleotides targeting Bcl-x pre-mRNA in human glioma cell lines. *Oncol. Rep.*, **35**, 1013–1019.
18. Mercatante,D.R. and Kole,R. (2002) Control of alternative splicing by antisense oligonucleotides as a potential chemotherapy: effects on gene expression. *Biochim. Biophys. Acta*, **1587**, 126–132.
19. Mercatante,D.R., Mohler,J.L. and Kole,R. (2002) Cellular response to an antisense-mediated shift of Bcl-x pre-mRNA splicing and antineoplastic agents. *J. Biol. Chem.*, **277**, 49374–49382.
20. Moore,M.J., Wang,Q., Kennedy,C.J. and Silver,P.A. (2010) An alternative splicing network links cell-cycle control to apoptosis. *Cell*, **142**, 625–636.
21. Shkreta,L., Blanchette,M., Toutant,J., Wilhelm,E., Bell,B., Story,B.A., Balachandran,A., Cochrane,A., Cheung,P.K., Harrigan,P.R, *et al.* (2017) Modulation of the splicing regulatory function of SRSF10 by a novel compound that impairs HIV-1 replication. *Nucleic Acids Res.*, **45**, 4051–4067.
22. Siddiqui-Jain,A., Grand,C.L., Bearss,D.J. and Hurley,L.H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl Acad. Sci. USA*, **99**, 11593–11598.
23. Didiot,M.C., Tian,Z., Schaeffer,C., Subramanian,M., Mandel,J.L. and 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.
24. Georgakopoulos-Soares,I., Parada,G.E., Wong,H.Y., Medhi,R., Furlan,G., Munita,R., Miska,E.A., Kwok,C.K. and Hemberg,M. (2022) Alternative splicing modulation by G-quadruplexes. *Nat. Commun.*, **13**, 2404.
25. Marcel,V., Tran,P.L., Sagne,C., Martel-Planche,G., Vaslin,L., Teulade-Fichou,M.P., Hall,J., Mergny,J.L., Hainaut,P. and Van Dyck,E. (2011) G-quadruplex structures in TP53 intron 3: role in alternative splicing and in production of p53 mRNA isoforms. *Carcinogenesis*, **32**, 271–278.
26. Angrand,G., Quillevere,A., Loaec,N., Daskalogianni,C., Granzhan,A., Teulade-Fichou,M.P., Fahraeus,R., Martins,P. and Blondel,M. (2019) Sneaking out for happy hour: yeast-based approaches to explore and modulate immune response and immune evasion. *Genes (Basel)*, **10**, 667.
27. Angrand,G., Quillevere,A., Loaec,N., Dinh,V.T., Le Senechal,R., Chennoufi,R., Duchambon,P., Keruzore,M., Martins,R.P., Teulade-Fichou,M.P, *et al.* (2022) Type I arginine methyltransferases are intervention points to unveil the oncogenic Epstein-Barr virus to the immune system. *Nucleic Acids Res.*, **50**, 11799–11819.
28. Beaudoin,J.D. and Perreault,J.P. (2010) 5'-UTR G-quadruplex structures acting as translational repressors. *Nucleic Acids Res.*, **38**, 7022–7036.
29. Dinh,V.T., Loaec,N., Quillevere,A., Le Senechal,R., Keruzore,M., Martins,R.P., Granzhan,A. and Blondel,M. (2023) The hide-and-seek game of the oncogenic Epstein-Barr virus-encoded EBNA1 protein with the immune system: an RNA G-quadruplex tale. *Biochimie.*, https://doi.org/10.1016/j.biochi.2023.07.010.
30. Lista,M.J., Martins,R.P., Angrand,G., Quillevere,A., Daskalogianni,C., Voisset,C., Teulade-Fichou,M.P., Fahraeus,R. and Blondel,M. (2017) A yeast model for the mechanism of the Epstein-Barr virus immune evasion identifies a new therapeutic target to interfere with the virus stealthiness. *Microb. Cell*, **4**, 305–307.
31. Lista,M.J., Martins,R.P., Billant,O., Contesse,M.A., Findakly,S., Pochard,P., Daskalogianni,C., Beauvineau,C., Guetta,C., Jamin,C., *et al.* (2017) Nucleolin directly mediates Epstein-Barr virus immune evasion through binding to G-quadruplexes of EBNA1 mRNA. *Nat. Commun.*, **8**, 16043.
32. Song,J., Perreault,J.P., Topisirovic,I. and Richard,S. (2016) RNA G-quadruplexes and their potential regulatory roles in translation. *Translation (Austin)*, **4**, e1244031.
33. Weldon,C., Behm-Ansmant,I., Hurley,L.H., Burley,G.A., Branlant,C., Eperon,I.C. and Dominguez,C. (2017) Identification of G-quadruplexes in long functional RNAs using 7-deazaguanine RNA. *Nat. Chem. Biol.*, **13**, 18–20.
34. Weldon,C., Dacanay,J.G., Gokhale,V., Boddupally,P.V.L., Behm-Ansmant,I., Burley,G.A., Branlant,C., Hurley,L.H., Dominguez,C. and Eperon,I.C. (2018) Specific G-quadruplex ligands modulate the alternative splicing of bcl-X. *Nucleic Acids Res.*, **46**, 886–896.
35. Bhogadia,M., Stone,B., Del Villar Guerra,R., Muskett,F.W., Ghosh,S., Taladriz-Sender,A., Burley,G.A., Eperon,I.C., Hudson,A.J. and Dominguez,C. (2022) Biophysical characterisation of the Bcl-x pre-mRNA and binding specificity of the ellipticine derivative GQC-05: implication for alternative splicing regulation. *Front. Mol. Biosci.*, **9**, 943105.
36. Cloutier,P., Toutant,J., Shkreta,L., Goekjian,S., Revil,T. and Chabot,B. (2008) Antagonistic effects of the SRp30c protein and cryptic 5' splice sites on the alternative splicing of the apoptotic regulator bcl-x. *J. Biol. Chem.*, **283**, 21315–21324.

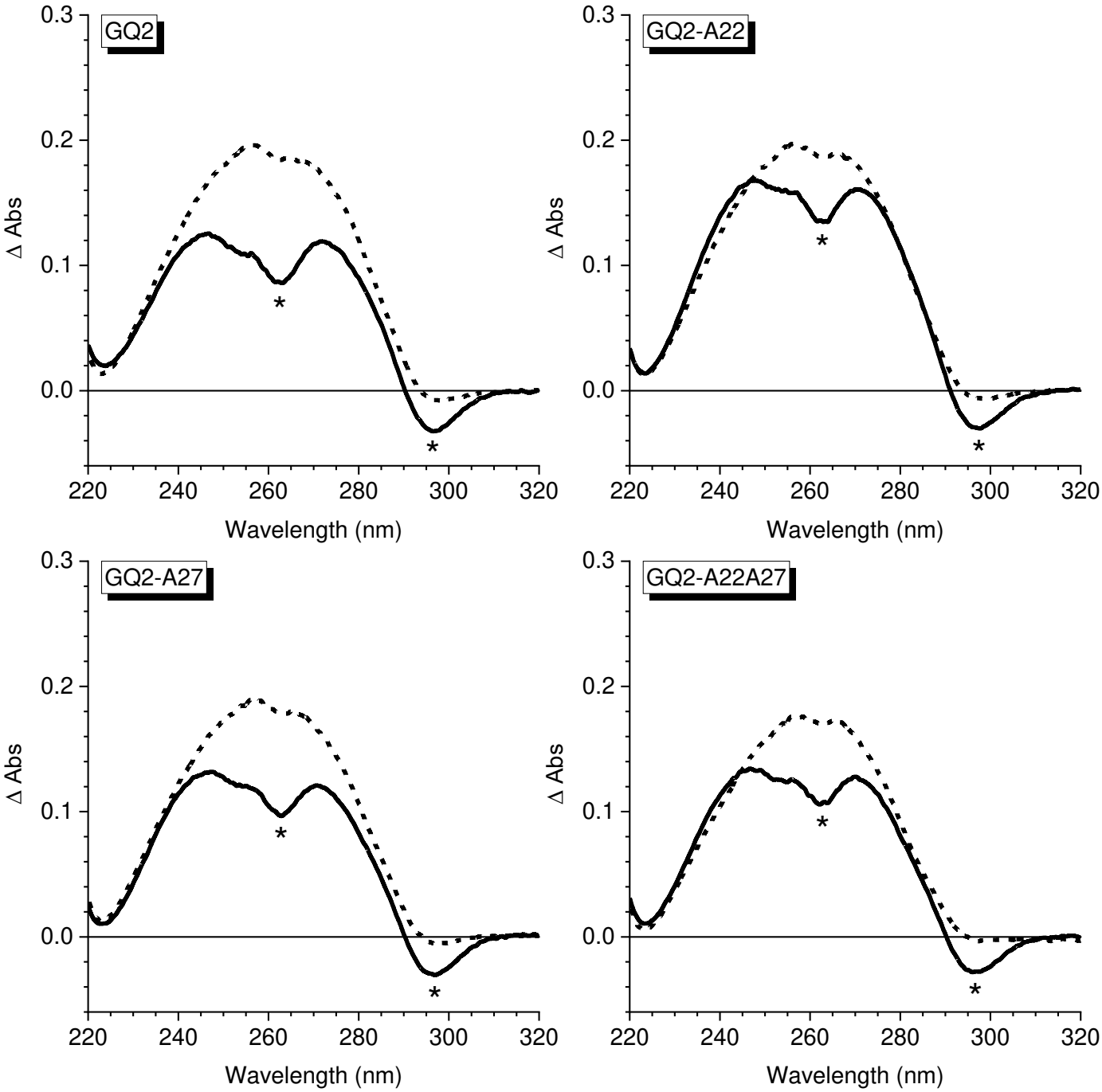
37. Paronetto,M.P., Achsel,T., Massiello,A., Chalfant,C.E. and Sette,C. (2007) The RNA-binding protein Sam68 modulates the alternative splicing of Bcl-x. *J. Cell Biol.*, **176**, 929–939.
38. Merdzhanova,G., Edmond,V., De Seranno,S., Van den Broeck,A., Corcos,L., Brambilla,C., Brambilla,E., Gazzeri,S. and Eymin,B. (2008) E2F1 controls alternative splicing pattern of genes involved in apoptosis through upregulation of the splicing factor SC35. *Cell Death Differ.*, **15**, 1815–1823.
39. Bielli,P., Bordi,M., Di Biasio,V. and Sette,C. (2014) Regulation of BCL-X splicing reveals a role for the polypyrimidine tract binding protein (PTBP1/hnRNP I) in alternative 5' splice site selection. *Nucleic Acids Res.*, **42**, 12070–12081.
40. Shkreta,L., Toutant,J., Durand,M., Manley,J.L. and Chabot,B. (2016) SRSF10 Connects DNA damage to the alternative splicing of transcripts encoding apoptosis, cell-cycle control, and DNA repair factors. *Cell Rep.*, **17**, 1990–2003.
41. Revil,T., Pelletier,J., Toutant,J., Cloutier,A. and Chabot,B. (2009) Heterogeneous nuclear ribonucleoprotein K represses the production of pro-apoptotic bcl-xS splice isoform. *J. Biol. Chem.*, **284**, 21458–21467.
42. Dominguez,C., Fisette,J.F., Chabot,B. and Allain,F.H. (2010) Structural basis of G-tract recognition and encaging by hnRNP F quasi-RRMs. *Nat. Struct. Mol. Biol.*, **17**, 853–861.
43. Garneau,D., Revil,T., Fisette,J.F. and Chabot,B. (2005) Heterogeneous nuclear ribonucleoprotein F/H proteins modulate the alternative splicing of the apoptotic mediator Bcl-x. *J. Biol. Chem.*, **280**, 22641–22650.
44. Massiello,A., Roesser,J.R. and Chalfant,C.E. (2006) SAP155 Binds to ceramide-responsive RNA cis-element 1 and regulates the alternative 5' splice site selection of Bcl-x pre-mRNA. *FASEB J.*, **20**, 1680–1682.
45. Wang,Y., Chen,D., Qian,H., Tsai,Y.S., Shao,S., Liu,Q., Dominguez,D. and Wang,Z. (2014) The splicing factor RBM4 controls apoptosis, proliferation, and migration to suppress tumor progression. *Cancer Cell*, **26**, 374–389.
46. Inoue,A., Yamamoto,N., Kimura,M., Nishio,K., Yamane,H. and Nakajima,K. (2014) RBM10 regulates alternative splicing. *FEBS Lett.*, **588**, 942–947.
47. Pedrotti,S., Busa,R., Compagnucci,C. and Sette,C. (2012) The RNA recognition motif protein RBM11 is a novel tissue-specific splicing regulator. *Nucleic Acids Res.*, **40**, 1021–1032.
48. Zhou,A., Ou,A.C., Cho,A., Benz,E.J. Jr and Huang,S.C. (2008) Novel splicing factor RBM25 modulates Bcl-x pre-mRNA 5' splice site selection. *Mol. Cell. Biol.*, **28**, 5924–5936.
49. Fortes,P., Longman,D., McCracken,S., Ip,J.Y., Poot,R., Mattaj,I.W., Caceres,J.F. and Blencowe,B.J. (2007) Identification and characterization of RED120: a conserved PWI domain protein with links to splicing and 3'-end formation. *FEBS Lett.*, **581**, 3087–3097.
50. Reznichenko,O., Leclercq,D., Pinto,F., J.,M., L.,G. and Granzhan,A. (2023) Optimization of G-quadruplex ligands through a SAR study combining parallel synthesis and screening of cationic bis(acylhydrazones). *Chem. Eur. J.*, **29**, e202202427.
51. Mergny,J.L., Li,J., Lacroix,L., Amrane,S. and Chaires,J.B. (2005) Thermal difference spectra: a specific signature for nucleic acid structures. *Nucleic Acids Res.*, **33**, e138.
52. Reznichenko,O., Quillevere,A., Martins,R.P., Loaec,N., Kang,H., Lista,M.J., Beauvineau,C., Gonzalez-Garcia,J., Guillot,R., Voisset,C., *et al.* (2019) Novel cationic bis(acylhydrazones) as modulators of Epstein-Barr virus immune evasion acting through disruption of interaction between nucleolin and G-quadruplexes of EBNA1 mRNA. *Eur. J. Med. Chem.*, **178**, 13–29.
53. Prado Martins,R., Findakly,S., Daskalogianni,C., Teulade-Fichou,M.P., Blondel,M. and Fahraeus,R. (2018) In cellulo protein-mRNA interaction assay to determine the action of G-quadruplex-binding molecules. *Molecules*, **23**, 3124.
54. Amit,M., Donyo,M., Hollander,D., Goren,A., Kim,E., Gelfman,S., Lev-Maor,G., Burstein,D., Schwartz,S., Postolsky,B., *et al.* (2012) Differential GC content between exons and introns establishes distinct strategies of splice-site recognition. *Cell Rep.*, **1**, 543–556.
55. Guiblet,W.M., DeGiorgio,M., Cheng,X., Chiaromonte,F., Eckert,K.A., Huang,Y.F. and Makova,K.D. (2021) Selection and thermostability suggest G-quadruplexes are novel functional elements of the human genome. *Genome Res.*, **31**, 1136–1149.
56. Maizels,N. and Gray,L.T. (2013) The G4 genome. *PLoS Genet.*, **9**, e1003468.
57. Brazda,V., Kolomaznik,J., Mergny,J.L. and Stastny,J. (2020) G4Killer web application: a tool to design G-quadruplex mutations. *Bioinformatics*, **36**, 3246–3247.
58. Kikin,O., D'Antonio,L. and Bagga,P.S. (2006) QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res.*, **34**, W676–W682.
59. Brazda,V., Kolomaznik,J., Lysek,J., Bartas,M., Fojta,M., Stastny,J. and Mergny,J.L. (2019) G4Hunter web application: a web server for G-quadruplex prediction. *Bioinformatics*, **35**, 3493–3495.
60. Luo,Y., Granzhan,A., Marquevielle,J., Cucchiarini,A., Lacroix,L., Amrane,S., Verga,D. and Mergny,J.L. (2022) Guidelines for G-quadruplexes: I. in vitro characterization. *Biochimie.*, https://doi.org/10.1016/j.biochi.2022.12.019.
61. von Hacht,A., Seifert,O., Menger,M., Schutze,T., Arora,A., Konthur,Z., Neubauer,P., Wagner,A., Weise,C. and Kurreck,J. (2014) Identification and characterization of RNA guanine-quadruplex binding proteins. *Nucleic Acids Res.*, **42**, 6630–6644.
62. Su,H., Xu,J., Chen,Y., Wang,Q., Lu,Z., Chen,Y., Chen,K., Han,S., Fang,Z., Wang,P., *et al.* (2021) Photoactive G-quadruplex ligand identifies multiple G-quadruplex-related proteins with extensive sequence tolerance in the cellular environment. *J. Am. Chem. Soc.*, **143**, 1917–1923.
63. Biffi,G., Tannahill,D., McCafferty,J. and Balasubramanian,S. (2013) Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.*, **5**, 182–186.
64. Massiello,A., Salas,A., Pinkerman,R.L., Roddy,P., Roesser,J.R. and Chalfant,C.E. (2004) Identification of two RNA *cis*-elements that function to regulate the 5' splice site selection of Bcl-x pre-mRNA in response to ceramide. *J. Biol. Chem.*, **279**, 15799–15804.
65. Wilusz,J.E., Devanney,S.C. and Caputi,M. (2005) Chimeric peptide nucleic acid compounds modulate splicing of the Bcl-x gene in vitro and in vivo. *Nucleic Acids Res.*, **33**, 6547–6554.
66. De Cian,A., Delemos,E., Mergny,J.L., Teulade-Fichou,M.P. and Monchaud,D. (2007) Highly efficient G-quadruplex recognition by bisquinolinium compounds. *J. Am. Chem. Soc.*, **129**, 1856–1857.
67. Chung,W.J., Heddi,B., Hamon,F., Teulade-Fichou,M.P. and Phan,A.T. (2014) Solution structure of a G-quadruplex bound to the bisquinolinium compound phen-DC(3). *Angew. Chem. Int. Ed Engl.*, **53**, 999–1002.
68. De Rache,A. and Mergny,J.L. (2015) Assessment of selectivity of G-quadruplex ligands via an optimised FRET melting assay. *Biochimie*, **115**, 194–202.
69. Monchaud,D., Allain,C., Bertrand,H., Smargiasso,N., Rosu,F., Gabelica,V., De Cian,A., Mergny,J.L. and Teulade-Fichou,M.P. (2008) Ligands playing musical chairs with G-quadruplex DNA: a rapid and simple displacement assay for identifying selective G-quadruplex binders. *Biochimie*, **90**, 1207–1223.
70. Wang,J.L., Liu,D., Zhang,Z.J., Shan,S., Han,X., Srinivasula,S.M., Croce,C.M., Alnemri,E.S. and Huang,Z. (2000) Structure-based discovery of an organic compound that binds bcl-2 protein and induces apoptosis of tumor cells. *Proc. Natl Acad. Sci. USA*, **97**, 7124–7129.
71. Warren,C.F.A., Wong-Brown,M.W. and Bowden,N.A. (2019) BCL-2 family isoforms in apoptosis and cancer. *Cell Death. Dis.*, **10**, 177.
72. Mitteaux,J., Lejault,P., Wojciechowski,F., Joubert,A., Boudon,J., Desbois,N., Gros,C.P., Hudson,R.H.E., Boule,J.B., Granzhan,A., *et al.* (2021) Identifying G-quadruplex-DNA-disrupting small molecules. *J. Am. Chem. Soc.*, **143**, 12567–12577.

73. Neidle,S. (2016)Quadruplex nucleic acids as novel therapeutic targets. *J. Med. Chem.*, **59**, 5987–6011.
74. Neidle,S. (2017) Quadruplex nucleic acids as targets for anticancer therapeutics. *Nat. Rev. Chem.*, **1**, 0041.
75. Zheng,A.J., Thermou,A., Guixens Gallardo,P., Malbert-Colas,L., Daskalogianni,C., Vaudiau,N., Brohagen,P., Granzhan,A., Blondel,M., Teulade-Fichou,M.P., *et al.* (2022) The different activities of RNA G-quadruplex structures are controlled by flanking sequences. *Life Sci. Alliance*, **5**, e202101232.
76. Kharel,P., Becker,G., Tsvetkov,V. and Ivanov,P. (2020) Properties and biological impact of RNA G-quadruplexes: from order to turmoil and back. *Nucleic Acids Res.*, **48**, 12534–12555.
77. Halder,K. and Chowdhury,S. (2007) Quadruplex-coupled kinetics distinguishes ligand binding between G4 DNA motifs. *Biochemistry*, **46**, 14762–14770.
78. Scalabrin,M., Nadai,M., Tassinari,M., Lago,S., Doria,F., Frasson,I., Freccero,M. and Richter,S.N. (2021) Selective recognition of a single HIV-1 G-quadruplex by ultrafast small-molecule screening. *Anal. Chem.*, **93**, 15243–15252.
79. Cadoni,E., De Paepe,L., Manicardi,A. and Madder,A. (2021) Beyond small molecules: targeting G-quadruplex structures with oligonucleotides and their analogues. *Nucleic Acids Res.*, **49**, 6638–6659.
80. Oltersdorf,T., Elmore,S.W., Shoemaker,A.R., Armstrong,R.C., Augeri,D.J., Belli,B.A., Bruncko,M., Deckwerth,T.L., Dinges,J., Hajduk,P.J, *et al.* (2005) An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature*, **435**, 677–681.
81. Tse,C., Shoemaker,A.R., Adickes,J., Anderson,M.G., Chen,J., Jin,S., Johnson,E.F., Marsh,K.C., Mitten,M.J., Nimmer,P., *et al.* (2008) ABT-263: a potent and orally bioavailable Bcl-2 family inhibitor. *Cancer Res.*, **68**, 3421–3428.

Received: March 22, 2023. **Revised:** August 5, 2023. **Editorial Decision:** August 29, 2023. **Accepted:** September 9, 2023

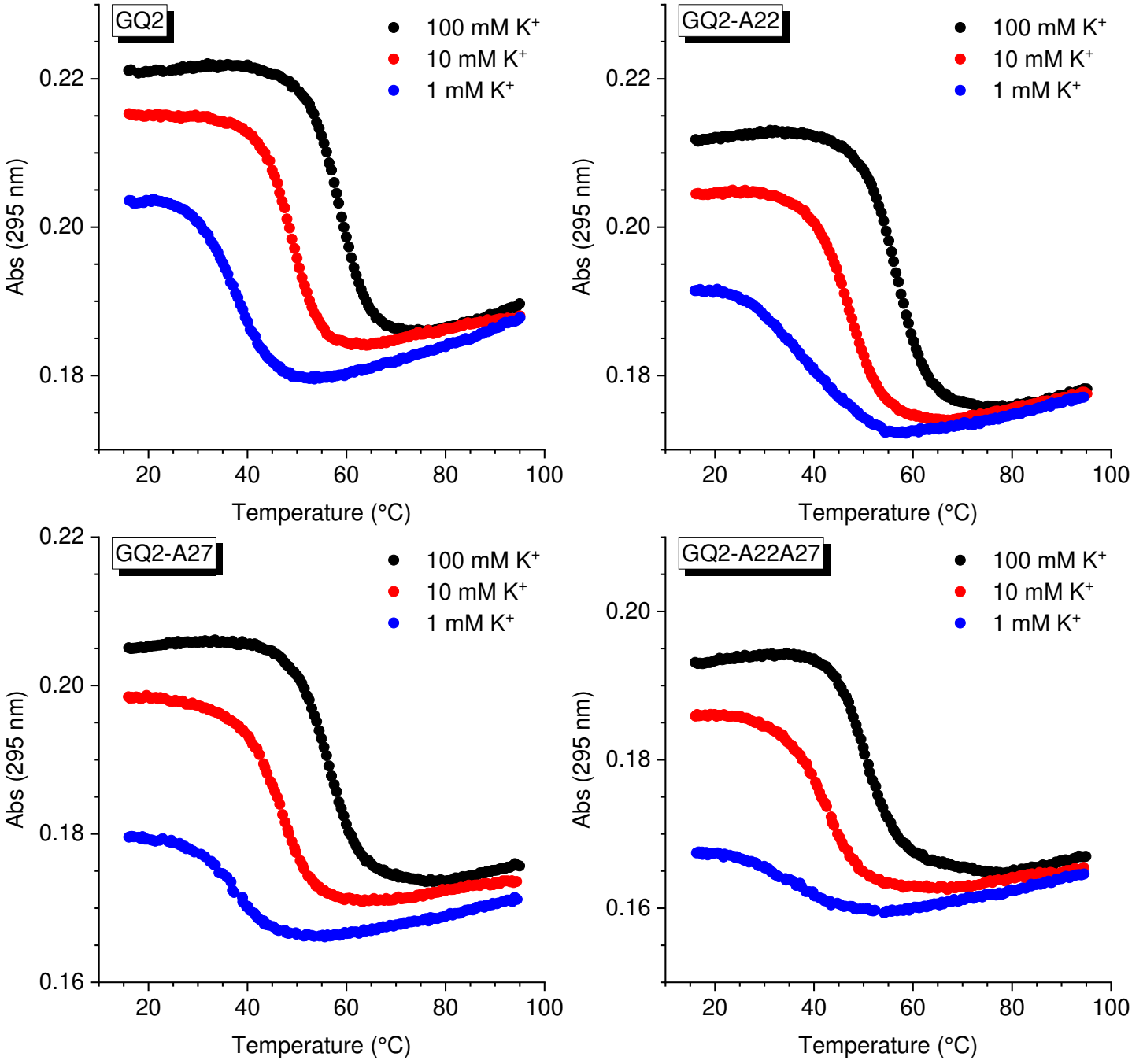
Supplementary Figure 1

Oligonucleotide	Sequence (5′→3′)
GQ2	GGGAUGGGGUAAACUGGGGUCGCAUUGUGG
GM2	GAGAUGAGGUAAACUGAGGUCGCAUUGUGG
GQ2-A22	GGGAUGGGGUAAACUGGGGUCACAUUGUGG
GQ2-A27	GGGAUGGGGUAAACUGGGGUCGCAUUAUGG
GQ2-A22A27	GGGAUGGGGUAAACUGGGGUCACAUUAUGG



Sequences and thermal difference spectra (ΔAbs = $Abs_{80\ °C}$ – $Abs_{20\ °C}$) of GQ-2 and its variants. TDS were obtained with 4 μM oligonucleotides in 10 mM lithium cacodylate buffer (pH 7.2) supplemented with 100 mM KCl (solid lines) or 100 mM LiCl (dashed lines). G4-characteristic peaks are indicated with asterisks.

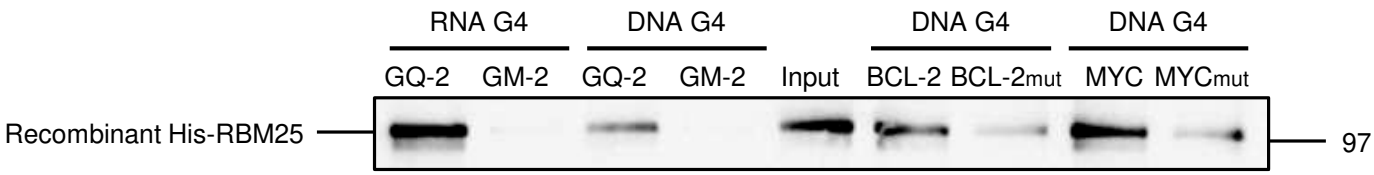
Supplementary Figure 2



Oligo	T_m / °C, at K^+ concentration (mM):		
	1	10	100
GQ2	37.4	48.8	58.6
GQ2-A22	38.5	47.1	56.6
GQ2-A27	36.6	46.9	58.4
GQ2-A22A27	35.6	42.8	50.4

UV-melting analysis of GQ-2 and its variants. UV melting profiles were obtained with 4 μM oligonucleotides in 10 mM lithium cacodylate buffer (pH 7.2) supplemented with various concentrations of KCl (as indicated) and LiCl to a total ionic strength of 110 mM.

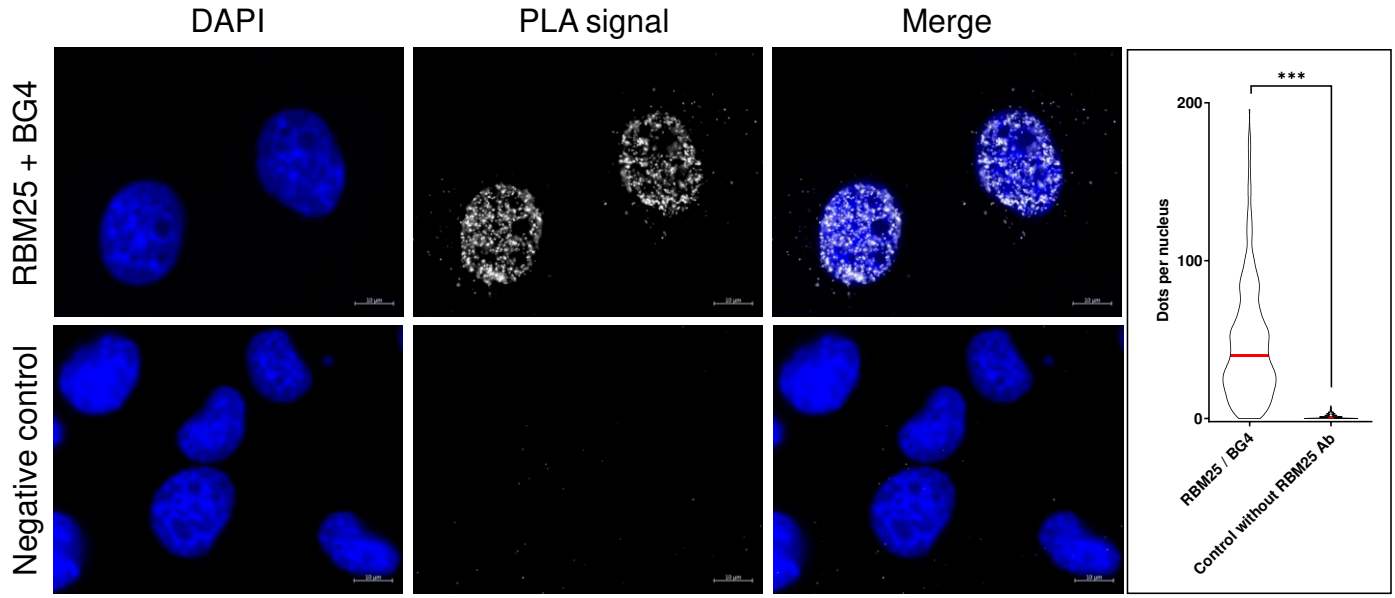
Supplementary Figure 3



RNA or DNA pulldown (as indicated) of a recombinant polyhistidine-tagged RBM25 protein (His-RBM25) using the indicated RNA or DNA oligonucleotides matrices were performed as indicated in **Figure 2**. Recombinant RBM25 protein still bound after an 800 mM KCl wash was eluted and analysed by SDS-PAGE and western blot using and anti-RBM25 antibody. Three different DNA matrices were used: GQ-2 which has the same sequence than the GQ-2 RNA matrix used in **Figure 2** to compare the ability of RBM25 to bind to rG4 and dG4 of the same sequence, BCL-2 and MYC (with their mutated versions, respectively BCL-2mut and MYCmut in which a minimal number of guanines involved in G4 formation were replaced by adenines or thymidines to prevent G4 formation as predicted using the G4Killer software, Brazda V et al, *Bioinformatics* 2020).

The following G-quadruplex forming oligonucleotides (RNA or DNA as indicated) 3'-tagged with TEG-Biotin were used:
GQ-2 (RNA) GGGAUGGGGUAAACUGGGGUCGCAUUGUGG-3'-TEG-Biotin,
GM-2 (RNA) GAGAUGAGGUAAACUGAGGUCGCAUUGUGG-3'-TEG-Biotin,
GQ-2 (DNA) GGGATGGGGTAAACTGGGGTCGCATTGTGG-3'-TEG-Biotin,
GM-2 (DNA) GAGATGAGGTAAACTGAGGTCGCATTGTGG-3'-TEG-Biotin,
BCL-2 (DNA) AGGGGCGGGCGCGGGAGGAAGGGGGCGGGAGCGGGGCTG-3'-TEG-Biotin,
BCL-2mut (DNA) AGGTGCGGGCGCGTGAGGAAGGTGGCGAGAGCGGAGCTG -3'-TEG-Biotin,
MYC (DNA) TGGGGAGGGTGGGGAGGGTGGGGAAGG -3'-TEG-Biotin and
MYCmut (DNA) TGGAGAGAGTGGAGAGAGTGGTGAAGG-3'-TEG-Biotin.

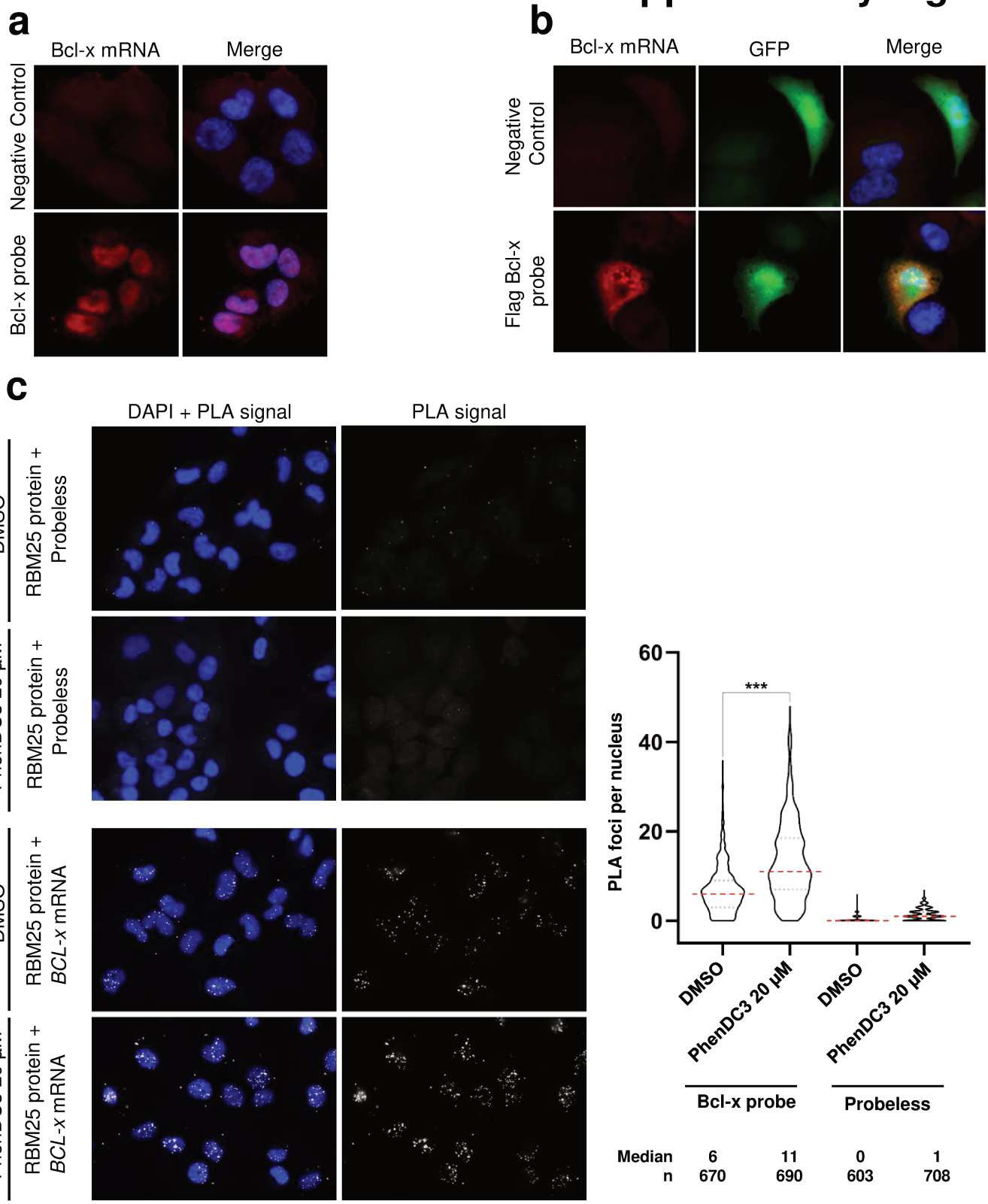
Supplementary Figure 4



Median value:	40	1
n:	543	564

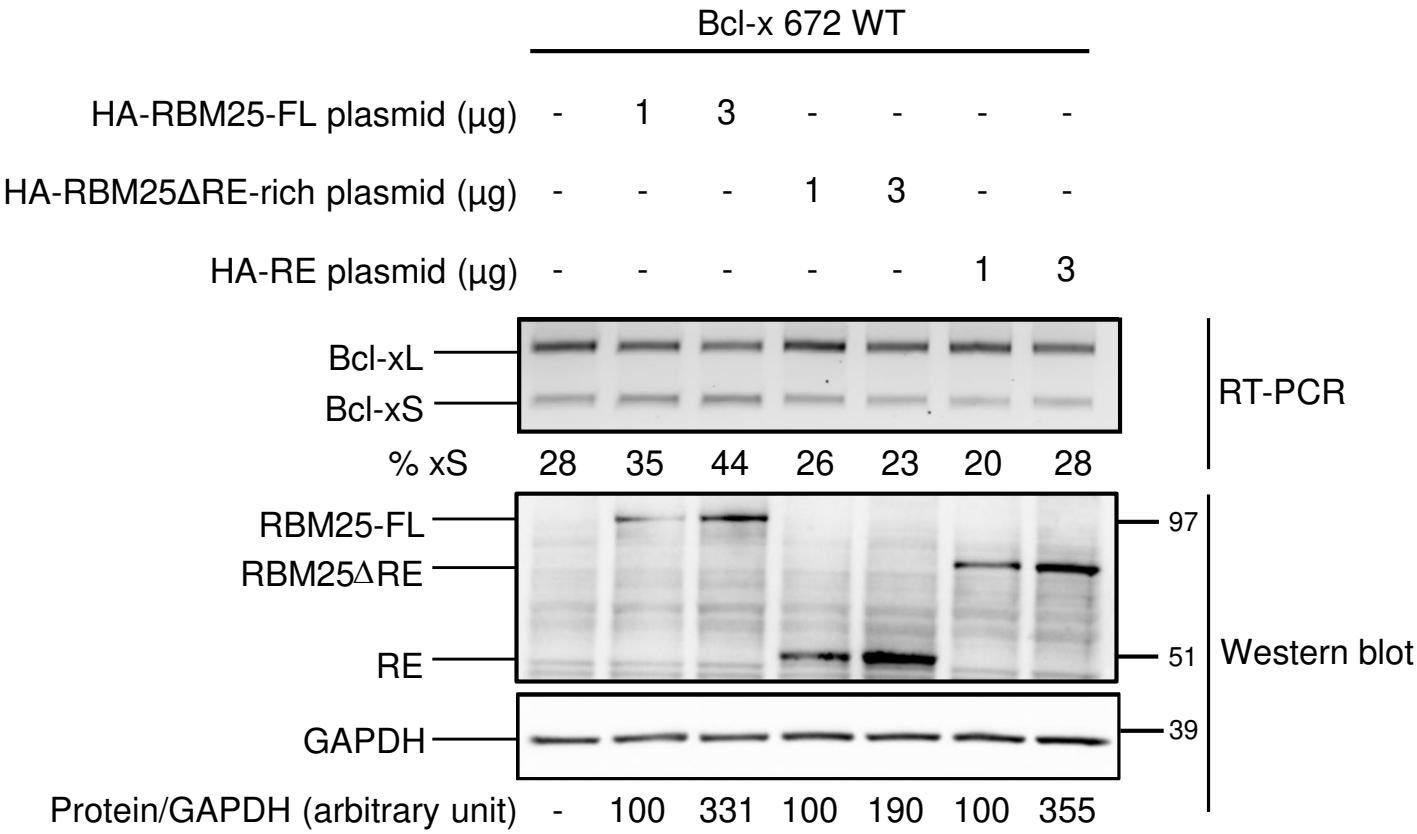
Proximity ligation assay (PLA) performed in H1299 cells to monitor *in cellulo* the interaction between endogenously expressed RBM25 protein and G4 using a BG4 antibody and an anti-RBM25 antibody. Microscopy images of H1299 cells analysed using both BG4 (anti-G4 antibody) and an anti-RBM25 antibody (upper panels) or only the BG4 antibody (lower panels, negative controls). Nuclei were revealed by DAPI staining and appear in blue. White dots (PLA signals) indicate interaction (close proximity) between RBM25 protein and G4 (that can be rG4 or dG4). The graph on the right indicates the number of PLA dots per cells in each condition. Data from three biological replicates, at least 500 cells per sample, were analysed by a non-parametric Mann-Whitney's test using the GraphPad Prism 8 Software (*** $P<0.0001$).

Supplementary Figure 5



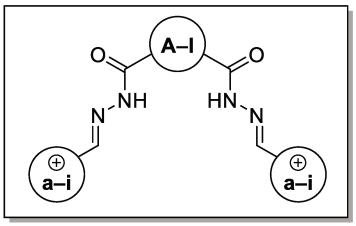
H1299 cells **(a)** and H1299 cells transfected with the minigene Bcl-x 672 WT **(b)** were analyzed by RNA *in situ* hybridization-immunofluorescence (rISH-IF) to verify the specificity of the Bcl-x digoxigenin probe **(a)** and the Flag Bcl-x digoxigenin probe **(b)** and to validate the detection of probe-mRNA complexes. **(c)** Adaptation of the proximity ligation assay (PLA) to monitor the RBM25 protein-*BCL-x* mRNA interaction performed in H1299 cells natively expressing RBM25 protein and *BCL-x* mRNA. Microscopy images of H1299 cells analysed using a probe specifically hybridizing to the *BCL-x* RNA (lower panels) or not (upper panels, negative controls) with or without PhenDC3 treatment as indicated. Nuclei were revealed by DAPI staining and appear in blue. White dots (PLA signals) indicated interaction (close proximity) between RBM25 protein and *BCL-x* RNA. The graph on the right indicates the number of PLA dots per cells in each condition. Data from three biological replicates, at least 250 cells per sample, were analysed by t-test in conjunction with Mann-Whitney's test using GraphPad Prism 8 Software (*** $P<0.0001$).

Supplementary Figure 6

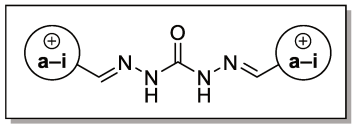


H1299 cells were transfected with the Bcl-x 672 WT minigene and a plasmid allowing expression of RBM25wt (HA-RBM25-FL), or a form of RBM25 deleted for its RE-rich motif (HA-RBM25-ΔRE-rich), or only the RE motif (HA-RE) to assess the effect of overexpressing one or the other of these various forms of RBM25 on alternative splicing from the Bcl-x 672 WT minigene. The result of the semi-quantitative RT-PCR experiment is shown in the upper panel and the western blot analysis of the level of expression of HA-RBM25-FL, HA-RBM25-ΔRE-rich or HA-RE, as compared to the loading control GAPDH, is shown in the two lower panels.

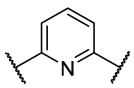
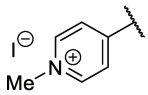
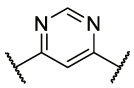
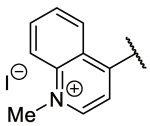
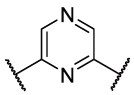
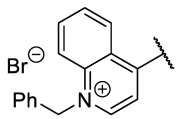
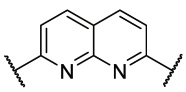
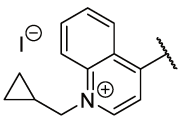
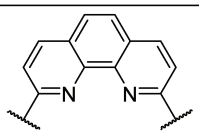
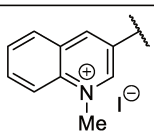
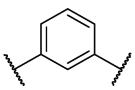
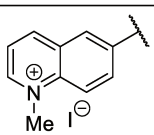
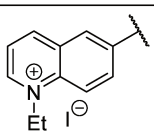
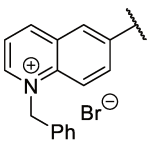
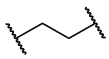
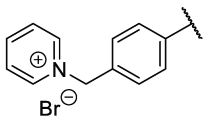
Supplementary Figure 7



Aa–Ii (81 cpds)

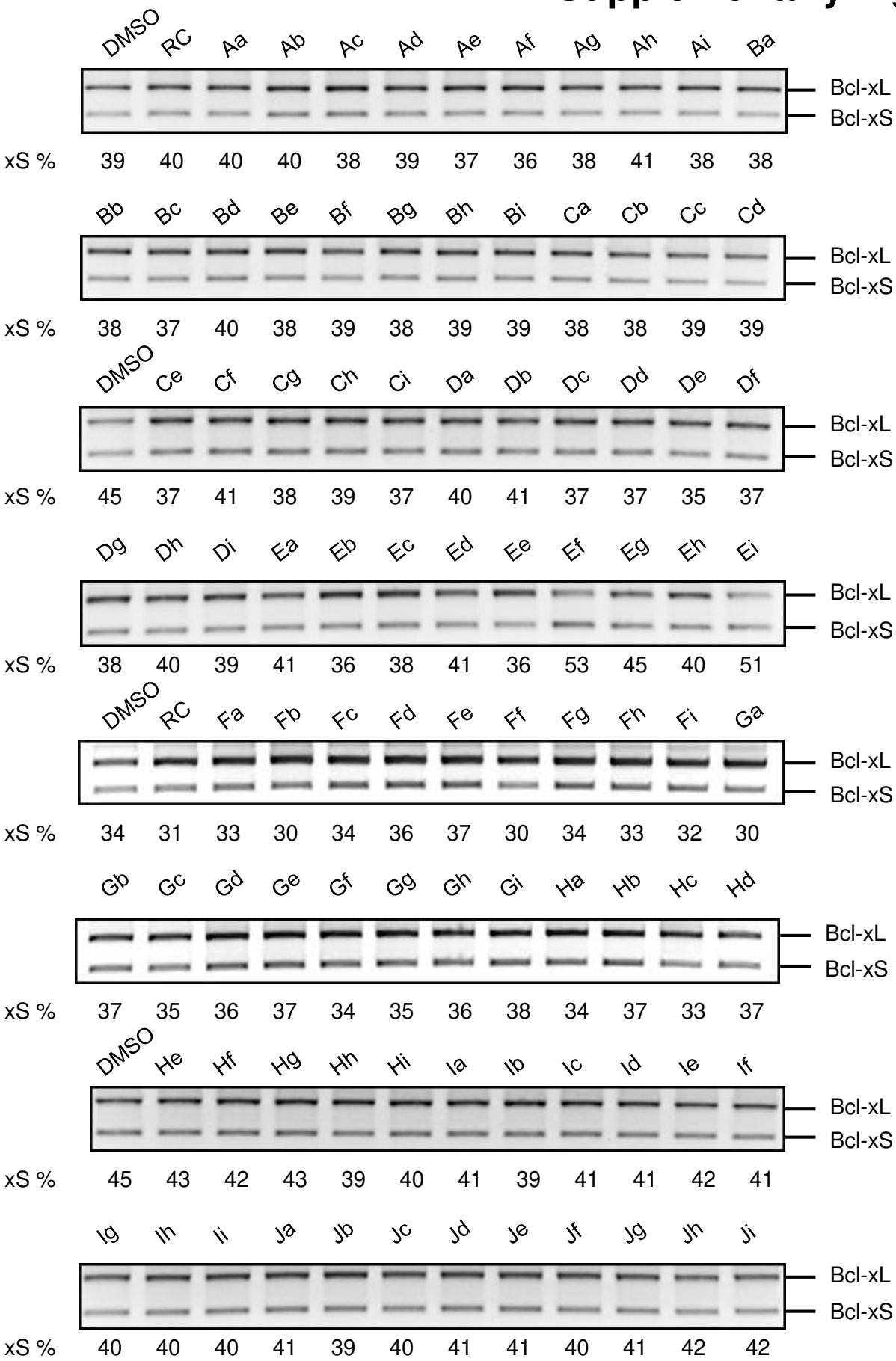


Ja–Ji (9 cpds)

A–I =		a–i =	
A		**a**	
B		**b**	
C		**c**	
D		**d**	
E		**e**	
F		**f**	
G		**g**	
H		**h**	
I		**i**	

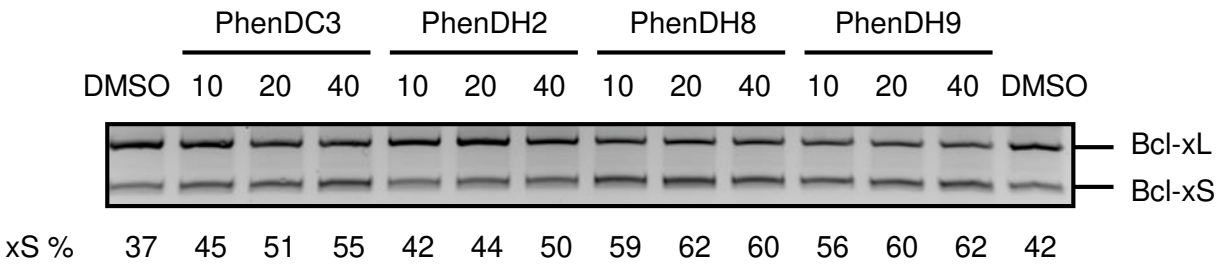
Combinatorial matrix of 90 cationic bis(acylhydrazones) screened as Bcl-x splicing modulators.

Supplementary Figure 8



Screening of a combinatorial library of 90 "as-synthesized" cationic bis(acylhydrazones) structurally related to PhenDC3 and previously validated as putative G4 ligands (Reznichenko O et al *Chemistry* 2023) using a fixed concentration of ~10 µM and the splicing from the Bcl-x 672 WT minigene as a readout. The splicing from the Bcl-x 672 WT minigene was assessed using semi-quantitative RT-PCR experiments. DMSO and RC (Reaction Control, 1M AcOH in DMSO, a catalyst present in every sample) were used as negative controls.

Supplementary Figure 9



Effect of various concentrations of PhenD3, PhenDH2, PhenDH8 or PhenDH9 on the splicing from the Bcl-x 672 WT minigene as assessed using semi-quantitative RT-PCR experiments. DMSO (compounds vehicle) was used as a negative control.

Partie II : Régulation du mécanisme d'échappement d'EBV au système immunitaire par sa protéine EBNA1

Préambule :

En plus du travail principal sur l'étude d'un rG4 présent dans le pré-ARNm de *BCL-x*, cette thèse a fait l'objet de plusieurs collaborations, dont deux ont été valorisées par deux articles tous les deux acceptés pour publication dans *Nucleic Acids Research* en 2022. Ces articles sont présentés ci-après et précédés, pour chacun, par une courte présentation.

Rôle du domaine RGG de la nucléoline dans l'inhibition GAr-dépendante de la traduction de l'ARNm d'EBNA1 du virus d'Epstein-Barr

Le virus d'Epstein-Barr (EBV) est un virus de la famille des *Herpesviridae* et de la sous-famille des *Gammaherpesvirinae*. En 1964, le virus EBV est identifié comme le premier oncovirus humain (Epstein et al., 1964). Il est estimé qu'au moins 1% des cancers dans le monde sont provoqués par le virus EBV. Ce virus est notamment associé aux lymphomes de Burkitt et d'Hodgkin, aux carcinomes nasopharyngés et à certains cancers gastriques. Comme de nombreux virus, EBV a développé des mécanismes d'évasion du système immunitaire. EBV possède une stratégie d'échappement immunitaire basée sur le contrôle de l'expression de sa protéine EBNA1 (Wilson et al., 2018). Cette protéine de maintenance du virus est considérée comme le talon d'Achille d'EBV. En effet, cette protéine est essentielle au virus pour la réplication et la maintenance de son génome, mais EBNA1 est également hautement antigénique. De sorte, des lymphocytes T dirigés contre cette protéine sont présents chez les individus infectés par EBV. Afin de rester invisible au système immunitaire et de limiter la présentation de peptides antigéniques dérivés d'EBNA1, cette protéine inhibe la traduction de son propre ARNm. De sorte, le niveau d'expression d'EBNA1 est maintenu au niveau minimum pour assurer la réplication et la maintenance du génome d'EBV (et donc la latence du virus) et en même temps suffisamment bas pour échapper au système immunitaire.

Ce mécanisme est régulé par le domaine GAr présent dans l'ARNm d'EBNA1 (Levitskaya et al., 1995). Le domaine GAr est composé de répétitions de glycine et d'alanine qui inhibent la traduction de son propre ARNm en *cis*. Ce mécanisme représente donc un point d'intervention pertinent pour dévoiler les cellules infectées par EBV au système immunitaire, en particulier les cellules tumorales des cancers liés à EBV. Plusieurs rG4 ont été identifiés dans la séquence de l'ARNm codant le domaine GAr d'EBNA1 et il a été montré qu'empêcher la formation de ces G4 aboutit à une augmentation de l'expression d'EBNA1 ainsi que de la présentation de peptides antigéniques dérivés d'EBNA1. En outre, la nucléoline (NCL) a été identifiée dans l'équipe comme le premier facteur de la cellule hôte de l'infection nécessaire pour l'inhibition de la traduction GAr-dépendante d'EBNA1 (Lista, Martins, Billant, et al., 2017). La NCL interagit de manière directe avec les rG4 de l'ARNm d'EBNA1 et cette interaction est nécessaire à l'inhibition de la traduction de l'ARNm d'EBNA1. La NCL est composée quatre domaines RRM, un motif RGG C-terminal et un domaine acide localisé du coté N-terminal. Dans le premier article présenté dans cette partie, nous avons cherché à déterminer le ou les domaines de la NCL impliqués dans l'inhibition GAr-dépendante de la traduction.

Afin d'identifier le domaine d'intérêt, nous avons utilisé un modèle de levure qui récapitule les aspects de l'inhibition GAr-dépendante de la traduction (Lista, Martins, Angrand, et al., 2017; Lista, Martins, Billant, et al., 2017; Voisset et al., 2014). Ce modèle est basé sur la fusion d'un domaine GAr de 43 acides aminés (43GAr) avec le gène rapporteur *ADE*2. Quand le niveau d'expression d'Ade2 est optimal, les colonies prennent une coloration blanche. A l'inverse, sans expression d'Ade2, les colonies apparaissent rouges. Enfin, lorsque l'expression est intermédiaire, les colonies prennent une coloration rose dont l'intensité de coloration est proportionnelle au niveau d'expression d'Ade2. La souche 43GAr-Ade2 qui est utilisée dans cette étude conduit à une inhibition partielle de l'expression d'Ade2p et ainsi une coloration rose des colonies. Ce modèle permet ainsi d'identifier des éléments capables d'augmenter ou d'inhiber l'effet de GAr sur la traduction en se basant sur la coloration des colonies. C'est dans ce modèle de levure que nous avons réintroduit des constructions de Nsr1 délétés de l'un ou l'autre de ses différents domaines afin d'identifier le domaine nécessaire pour l'inhibition GAr-dépendante de la traduction. Cette expérience a conduit à l'identification du domaine RGG C-terminal de Nsr1 comme le domaine de

la nucléoline essentiel à l'inhibition GAr-dépendante de la traduction. De plus, par des expériences de *RNA pulldown*, nous avons montré que le domaine RGG de Nsr1 était nécessaire pour son interaction avec les rG4 de l'ARNm d'EBNA1. Ces expériences ont ensuite été reproduites avec la nucléoline humaine et ont confirmé le rôle essentiel du domaine RGG C-terminal de la nucléoline dans la liaison avec les rG4 de l'ARNm d'EBNA1 et donc dans l'inhibition GAr-dépendante de la traduction.

Les motifs RGG sont les principales cibles des protéines arginines méthyltransférases (PRMTs ; notamment des PRMTs de type I) qui vont catalyser l'ajout de groupements méthyles au niveau des atomes d'azote des arginines. Aussi, nous avons regardé si la méthylation des arginines de la NCL était nécessaire pour l'inhibition GAr-dépendante de la traduction. Pour cela nous avons délété *HMT1*, le gène codant la principale PRMT de type I chez la levure qui est responsable de la majorité de la méthylation des arginines chez cet organisme eucaryote modèle. Nous avons alors observé que, en réponse à cette délétion d'*HMT1*, l'inhibition de la traduction d'EBNA1 était significativement diminuée. De plus, la délétion d'*HMT1* réduit également la capacité de Nsr1 d'interagir avec les rG4 d'EBNA1. Ces résultats montrent que le domaine RGG de Nsr1 doit être méthylé afin d'induire l'inhibition GAr-dépendante de la traduction. Ces résultats ont été reproduits en cellules humaines en utilisant des ARNi dirigés contre les principales PRMTs de types I ou des inhibiteurs pharmacologiques de ces enzymes et ont permis d'identifier les PRMTs de type I comme nécessaires pour l'inhibition GAr-dépendante de la traduction EBNA1 définissant ainsi de nouveaux points d'intervention possible pour interférer avec le rôle de la nucléoline dans l'inhibition GAr-dépendante de la traduction, et donc avec l'échappement d'EBV au système immunitaire.

Ma contribution dans cette publication a été de tester si les arginines du motif RGG étaient, ou pas la cible, des PRMT de type I. Pour cela, nous avons tiré parti du fait que la nucléoline de levure Nsr1 n'est pas essentielle pour cet organisme eucaryote modèle. Nous avons donc utilisé une souche délété pour le gène *NSR1* (*nsr1*∆). J'ai ensuite construit, par mutagénèse dirigée, des versions mutées de Nsr1 dans lesquelles chacune des 8 arginines du motif RGG C-terminal ont été remplacées soit par des alanines (qui ne peuvent être méthylées, mutant nsr1-8A), soit par des phénylalanines (qui miment des arginines constitutivement méthylées, mutant nsr1-8F) puis exprimées ces différents mutants de nsr1 dans la souche *nsr1*∆. Nous avons

ainsi montré que les levures exprimant nsr1-8A expriment un niveau élevé d'une protéine de fusion 43GAr-Ade2 de façon similaire à la souche *nsr1*Δ dans laquelle l'inhibition GAr-dépendante de la traduction n'est plus opérante, tandis que ce niveau est plus faible dans les levures exprimant le mutant nsr1-8F. Ces résultats ont ensuite été confirmés avec la nucléoline humaine avec des mutations de ses 10 arginines soit en alanines, soit en phénylalanines. La NCL avec les mutations en phénylalanine (NCL-10F) agit de manière similaire à la protéine WT sur l'inhibition GAr-dépendante de la traduction tandis que la NCL-10A est totalement inactive. En conclusion, ces résultats montrent que la méthylation des arginines du domaine RGG de la NCL est nécessaire pour l'interaction avec les rG4 de l'ARNm d'EBNA1.

D'après ces résultats, deux modèles sont proposés pour expliquer la régulation entre la NCL et les rG4 d'EBNA1. Le premier modèle, dit direct, propose que la méthylation des arginines de la NCL par la PRMT I soit nécessaire pour que la protéine interagisse avec les rG4 d'EBNA1, favorisant ainsi l'inhibition de la traduction et l'échappement au système immunitaire d'EBV. Le deuxième modèle propose un rôle indirect de la méthylation. Dans ce modèle, la NCL serait séquestrée par une protéine partenaire qui empêcherait ainsi l'interaction de la NCL avec les rG4 de l'ARNm d'EBNA1. La méthylation du domaine RGG viendrait dissocier la NCL de sa protéine partenaire et donc la libérer, pour lui permettre d'aller s'associer avec les rG4 de l'ARNm d'EBNA1, ce qui conduirait à l'inhibition de la traduction de l'ARNm d'EBNA1, et donc à l'échappement d'EBV au système immunitaire.

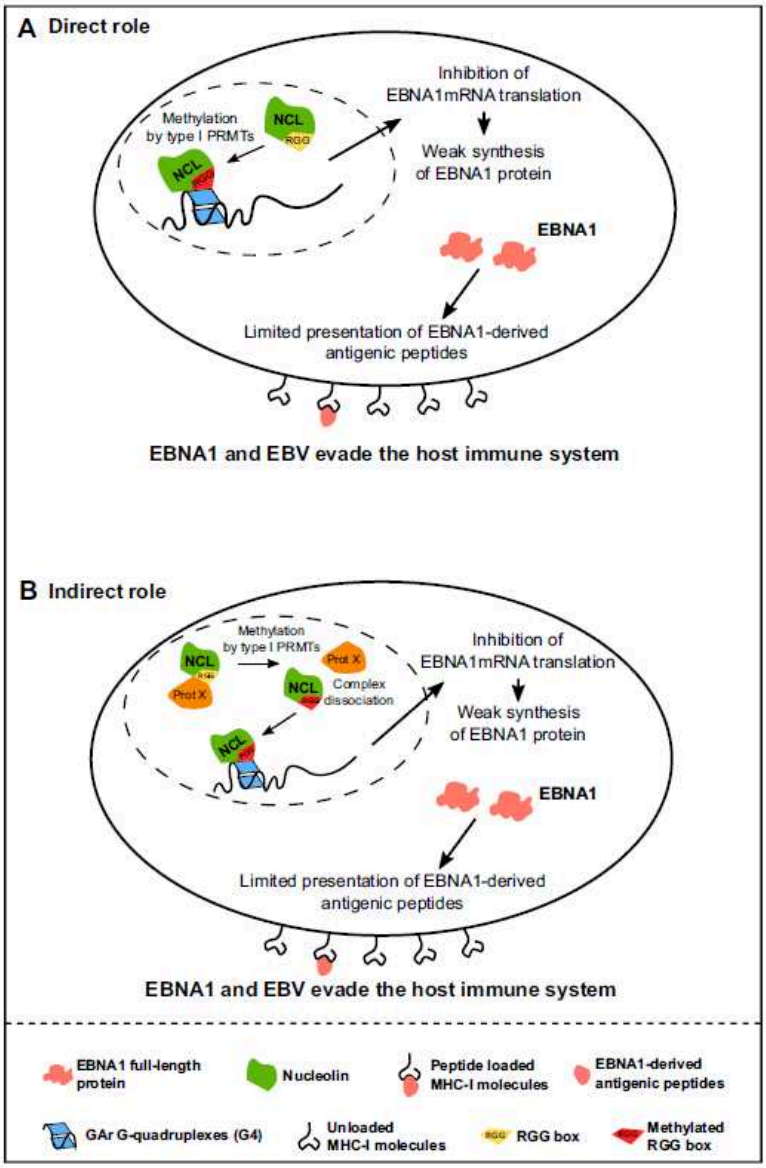


Figure 14 : modèles de l'interaction entre la NCL et les rG4 de l'ARNm d'EBNA1.

L'interaction de la NCL avec les rG4 de l'ARNm d'EBNA1 est nécessaire pour permettre l'inhibition GAr-dépendante de la traduction et ainsi l'échappement au système immunitaire d'EBV. La méthylation des arginines du domaine RGG de la NCL par la PRMT I est essentielle pour l'interaction de la NCL avec les rG4. Aussi, deux modèles sont proposés pour expliquer la régulation par la PRMT I, un modèle direct et un modèle indirect. Dans le modèle direct, la méthylation des arginines du domaine RGG de la NCL favorise directement l'interaction avec les rG4 et l'inhibition de la traduction ce qui conduit à l'échappement au système immunitaire du virus.

Dans le modèle indirect, la NCL et son domaine RGG non méthylé sont séquestrés par une protéine partenaire (protéine X), ce qui empêche la liaison de la NCL avec les rG4. Dans ce cas, il n'y a pas d'inhibition de la traduction. Les arginines du motif RGG doivent être méthylées par la PRMT I pour que la NCL soit libérée de la protéine X pour pouvoir ensuite aller se lier aux rG4 de l'ARNm d'EBNA1 et permettre ainsi l'échappement d'EBV au système immunitaire.

Nucleic Acids Research, 2022 **1**
https://doi.org/10.1093/nar/gkac915

Type I arginine methyltransferases are intervention points to unveil the oncogenic Epstein-Barr virus to the immune system

Gaelle Angrand[1,†], Alicia Quillévéré[1,†], Nadège Loaëc[1,†], Van-Trang Dinh[1], Ronan Le Sénéchal[1], Rahima Chennoufi[2], Patricia Duchambon[2], Marc Keruzoré[1], Rodrigo Prado Martins[3], Marie-Paule Teulade-Fichou[2], Robin Fåhraeus[4,5] and Marc Blondel[1,*]

[1]Institut National de la Santé et de la Recherche Médicale UMR1078; Université de Bretagne Occidentale, Faculté de Médecine et des Sciences de la Santé; Etablissement Français du Sang (EFS) Bretagne; CHRU Brest, Hôpital Morvan, Laboratoire de Génétique Moléculaire, 22 avenue Camille Desmoulins, F-29200 Brest, France, [2]Chemistry and Modelling for the Biology of Cancer, CNRS UMR9187 - Inserm U1196, Institut Curie, Université Paris-Saclay, Orsay, Campus universitaire, Bat. 110, F-91405, France, [3]ISP, INRAE, Université de Tours, UMR1282, Tours, Nouzilly, France, [4]Cibles Thérapeutiques, Institut National de la Santé et de la Recherche Médicale UMR1162, Institut de Génétique Moléculaire, Université Paris 7, Hôpital St. Louis, 27 rue Juliette Dodu, F-75010 Paris, France and [5]RECAMO, Masaryk Memorial Cancer Institute, Zluty kopec 7, 65653 Brno, Czech Republic

Received September 28, 2022; Editorial Decision September 30, 2022; Accepted October 06, 2022

ABSTRACT

The oncogenic Epstein-Barr virus (EBV) evades the immune system but has an Achilles heel: its genome maintenance protein EBNA1. Indeed, EBNA1 is essential for viral genome maintenance but is also highly antigenic. Hence, EBV seemingly evolved a system in which the glycine–alanine repeat (GAr) of EBNA1 limits the translation of its own mRNA to the minimal level to ensure its essential function, thereby, at the same time, minimizing immune recognition. Therefore, defining intervention points at which to interfere with GAr-based inhibition of translation is an important step to trigger an immune response against EBV-carrying cancers. The host protein nucleolin (NCL) plays a critical role in this process via a direct interaction with G-quadruplexes (G4) formed in the GAr-encoding sequence of the viral EBNA1 mRNA. Here we show that the C-terminal arginine–glycine-rich (RGG) motif of NCL is crucial for its role in GAr-based inhibition of translation by mediating interaction of NCL with G4 of EBNA1 mRNA. We also show that this interaction depends on the type I arginine methyltransferase family, notably PRMT1 and PRMT3: drugs or small interfering RNA that target these enzymes prevent efficient binding of NCL on G4 of EBNA1 mRNA and relieve GAr-based inhibition of translation and of antigen presentation. Hence, this work defines type I arginine methyltransferases as therapeutic targets to interfere with EBNA1 and EBV immune evasion.

INTRODUCTION

The Epstein-Barr virus (EBV) is the first oncogenic virus described in human (1–3). EBV is linked to at least 1% of cancers worldwide (4), which include Burkitt's and Hodgkin's lymphomas, nasopharyngeal carcinoma, 10% of gastric cancers and potentially gliomas (5–7). Like the other gamma herpesviruses, EBV is a latent virus that evades the host immune system. However, EBV presents an Achilles heel: its virally encoded genome maintenance protein EBNA1 (Epstein-Barr nuclear antigen 1), which is both essential for the virus (for the maintenance of its genome) and highly antigenic. In addition, infected individuals may contain T cells raised against EBNA1 (8–10). Hence, EBV evolved a mechanism by which EBNA1 self-inhibits the translation of its own mRNA, thereby minimizing the production of EBNA1-derived antigenic peptides (11,12). Although not fully elucidated, this mechanism critically involves the Gly–Ala repeat (GAr) of EBNA1 which is encoded by a guanine-repeat-containing mRNA sequence that was shown to be able to form clusters of up to thirteen

*To whom correspondence should be addressed. Tel: +33 2 98 01 83 88; Email: marc.blondel@univ-brest.fr
†The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.

G-quadruplexes (G4) that have been involved in GAr-based inhibition of translation (13). G4 are non-canonical secondary structures that may assemble in guanine-rich DNA or RNA. G4 are formed by the stacking of at least two G-quartets which consist of a planar arrangement of four guanines connected by Hoogsteen hydrogen bonds and stabilized by a central cation, most often K^+. G4 structures within G-rich DNA or RNA have been implicated in gene regulation where they can affect transcription, splicing or translation (14–19). Recently, the host cell nucleolin protein (NCL in human, Nsr1 in the budding yeast *Saccharomyces cerevisiae*) has been shown to directly mediate EBV immune evasion through binding to G4 of the GAr-encoding sequence of EBNA1 mRNA (20,21). Indeed, the binding of NCL to G4 of EBNA1 mRNA is required for GAr-based inhibition of both translation and antigen presentation. Importantly, this protein–RNA-G4 interaction is a relevant intervention point to interfere with EBNA1 immune evasion, as several G4 ligands, which include the benchmark compound PhenDC3 as well as the new derivatives PyDH2 and PhenDH2, are able to interfere with both this interaction and, consequently, with GAr-based limitation of translation and antigen presentation (21,22). Hence G4 that form in the GAr-encoding sequence of EBNA1 mRNA constitute a recognition platform for the binding of the host cell NCL, a mechanism which is critically involved in EBNA1 immune evasion. However, several important questions remain such as about the precise role of NCL in GAr-based translation inhibition, or the identification of the domain of NCL that is involved in binding of G4 of EBNA1 mRNA.

NCL is a multifunctional nucleolar DNA/RNA-binding protein widely conserved among eukaryotes, whose major role is in RNA metabolism, in particular rRNA maturation (23–25). NCL also plays additional roles in chromatin remodelling, cell cycle control, transcription and apoptosis. NCL is one of the first G4-interacting proteins identified (26). One of its best documented mechanisms of action on G4 is its role as a repressor of transcription through stabilization of DNA-G4 in the c-Myc promoter (27). Later, NCL was shown to recognize DNA-G4 within the long terminal repeat (LTR) promoter of human immunodeficiency virus (HIV), thereby silencing the provirus transcription (28). More recently, in addition to its role in the immune evasion of EBV through binding to G4 of EBNA1 mRNA (21), NCL has been involved in suppressing hepatitis C virus (HCV) replication through binding to the viral core RNA-G4 structure (29). NCL is ubiquitous and is therefore likely to interact at many G4 loci of DNA or RNA. The structural determinants of G4–NCL interactions are not precisely known but, recently, a strong preference of NCL for binding G4-DNA harbouring a long central loop has been reported by several groups (30–32). NCL is composed of three main domains: (i) the N-terminal domain that contains four acidic stretches; (ii) the central region that contains four tandem RNA recognition motifs (RRM1–RRM4); and (iii) the arginine/glycine/glycine (RGG) motif at the C-terminus. Importantly, most nucleolin homologues known to date share the same organization, with some variations in the number of RRMs; for example, the yeast nucleolin Nsr1 contains only two RRMs but possesses the acidic N-terminal domain as well as the C-terminal RGG motif (Figure 1A). RGG motifs represent low complexity sequences which can lead to disordered regions (33) and have been involved in interaction with nucleic acids or various proteins (34).

Methyl groups can be attached to the nitrogen atoms of arginine within polypeptides, a process termed arginine methylation. This post-translational modification is conducted by protein arginine methyltransferases (PRMTs) (35). PRMTs catalyse the transfer of a methyl group from *S*-adenosylmethionine (SAM) to the guanidino nitrogen atoms of arginine, resulting in the formation of methylarginine and *S*-adenosylhomocysteine. Arginine can be methylated in three different ways on its guanidino group: it can be monomethylated, symmetrically dimethylated or asymmetrically dimethylated, each of which has potentially different functional consequences (36). There are nine PRMTs in human that are divided into three groups: type I, type II and type III (37). PRMT7 is the single type III PRMT and only catalyses the formation of monomethylarginine (MMA), whereas both type I and type II PRMTs catalyse the formation of MMA first, and then, from this intermediate, of dimethylarginine (DMA). Starting from MMA, type I PRMTs (PRMT1, PRMT2, PRMT3, PRMT4, PRMT6 and PRMT8) catalyse the formation of asymmetric dimethylarginine (aDMA), whereas type II PRMTs (PRMT5 and PRMT9) catalyse the formation of symmetric dimethylarginine (sDMA). Arginine methylation occurs on a variety of protein sequence motifs, but the RGG motifs are the most commonly reported (38,39) and, in budding yeast, it is estimated that the majority of arginine methylation occurs in RGG motifs (40). This is in line with the fact that RGG is the canonical motif substrate established for Hmt1p, a type I PRMT and the main budding yeast PRMT, which is responsible for 66% of arginine monomethylation and 89% of asymmetric dimethylation of intracellular proteins in *S. cerevisiae* (41).

Here, we have first determined that the C-terminal RGG motif of nucleolin is essential for both its ability to bind G4 of EBNA1 mRNA and its role in GAr-based inhibition of translation. This is true for both Nsr1, the yeast nucleolin, and for NCL, the human nucleolin. We also show that the ability of nucleolin to interact with G4 of EBNA1 mRNA, as well as the GAr-based limitation of both translation and antigen presentation, depend on type I PRMTs, in particular on PRMT1 and PRMT3.

Hence, the RGG motif of the host protein NCL is critically involved in the binding of G4 of the viral EBNA1 mRNA, and this interaction depends on type I arginine methyltransferases that thus represent relevant therapeutic targets to interfere with EBNA1 and EBV immune evasion.

MATERIALS AND METHODS

Yeast strains, genetic manipulation and culture media

All the strains used in this study are derived from the W303a *WT* K699 strain (42): *MAT*a, *leu2-3*, 112 *trp1-1*, *can 1-100*, *ura3*-1, *ade2-1*, *his3-11, 15*.

***Y32**: MAT a, leu2-3, 113 trp1-1, can 1-100, ura3-1, ade2-1:: his5 $^{S.pombe}$, his3-11,15, met15::HA-ade2*

***Y33**: MAT a, leu2-3, 113 trp1-1, can 1-100, ura3-1, ade2-1:: his5 $^{S.pombe}$, his3-11,15, met15::HA-43GAr -ade2*

***Y32 nsr1Δ**: MAT a, leu2-3, 113 trp1-1, can 1-100, ura3-1, ade2-1:: his5 $^{S.pombe}$, his3-11,15, met15::HA-ade2, nsr1::KANMX6*

***Y33 nsr1Δ**: MAT a, leu2-3, 113 trp1-1, can 1-100, ura3-1, ade2-1:: his5 $^{S.pombe}$, his3-11,15, met15::HA-43GAr-ade2, nsr1::KANMX6*

***Y32 nsr1ΔRGG**: MAT a, leu2-3, 113 trp1-1, can 1-100, ura3-1, ade2-1:: his5 $^{S.pombe}$, his3-11,15, met15::HA-ade2, nsr1ΔRGG::KANMX6*

***Y33 nsr1ΔRGG**: MAT a, leu2-3, 113 trp1-1, can 1-100, ura3-1, ade2-1:: his5 $^{S.pombe}$, his3-11,15, met15::HA-43GAr-ade2, nsr1ΔRGG::KANMX6*

***Y32 hmt1Δ**: MAT a, leu2-3, 113 trp1-1, can 1-100, ura3-1, ade2-1:: his5 $^{S.pombe}$, his3-11,15, met15::HA-ade2, hmt1::KANMX6*

***Y33 hmt1Δ**: MAT a, leu2-3, 113 trp1-1, can 1-100, ura3-1, ade2-1:: his5 $^{S.pombe}$, his3-11,15, met15::HA-43GAr-ade2, hmt1::KANMX6*

Yeasts were growth on the following media: YPD (10 g/l yeast extract, 20 g/l peptone, 20 g/l D-glucose) and DO-TRP (6.7 g/l yeast extract without amino acids, 0.74 g of CSM-TRP, 20 g/l D-glucose). For solid media, agar was added at a final concentration of 20 g/l.

Plasmid construction

All the plasmids were generated using standard procedures. The T4 DNA ligase was obtained from Promega and the restriction enzymes were purchased from New England Biolabs. Plasmids maintenance and amplification were carried out in a chemically competent *Escherichia coli* strain. The plasmids pRS414 containing NSR1 and parts of the coding sequence were constructed as follows: the DNA fragments were amplified by the forward primers 5′-CGCGGATCCATGGCTAAGACTACTAAAGTAAA AGGTAAC-3′ or 5′-GCGCGCGGATCCATGTCTTC CAACAAGAAGCAAAAA-3′, and the reverse primers 5′-CCGCTCGAGCGGTTAATCAAATGTTTTCTTT GAACCAG-3′, 5′-CGGCCGCTCGAGTTACTCTTCG TCTTCTTCTTCTTC-3′, 5′-CGGCCGCTCGAGTTAC TTGGCACGATCGTTGTTAC-3′ or 5′-CGGCCGCT CGAGTTAACCATCGTTGTTTGGTCTTGG-3′. The corresponding polymerase chain reaction (PCR) fragments were cloned into BamHI and XhoI cloning sites of the pRS414-ADH vector.

All generated constructions were verified by PCR amplification on clones, restriction enzyme digestion and sequencing.

Yeast protein extraction

A 4 ml aliquot of exponentially growing yeast of 0.7–0.8 $OD_{600\,nm}$ was harvested, washed in 1× TE, pelleted and then suspended into 300 μl of lysis buffer [25 mM Tris–HCl pH 6.8, 10% glycerol, 5% β-mercaptoethanol, 5% sodium dodecylsulphate (SDS), 8 M urea, 0.02% bromophenol blue)

Cell culture and transfection

H1299 cells are derived from metastatic lymph node from lung carcinoma. Mutu-1 cells are derived from an EBV-positive Burkitt lymphoma. Raji cells are from a type III latency EBV-infected Burkitt's lymphoma. B3Z T cells are from a hybridoma expressing a T-cell receptor (TCR) that specifically recognizes ovalbumin (OVA; 257–264: SIINFEKL) in the context of H-2Kb. H1299, Mutu-1 and Raji cells were cultured in RPM1-1640 supplemented with 10% foetal bovine serum and 2 mM L-glutamine. B3Z cells were cultured in the same medium supplemented with 50 μM β-mercaptoethanol. Transfections were performed using GeneJuice® transfection reagent (Merck), or by electroporation using the Amaxa Cell Line Nucleofector® kit V (Lonza) for Raji and Mutu-1 cells, in both cases according to the manufacturer's protocol.

Small interfering RNA (siRNA)-mediated knockdown

A total of 50 000 cells per well were seeded in 6-well plates and transfected the following day with 0.75 μg of EBNA1 or EBNA1ΔGAr expression vectors using Genejuice according to the manufacturer's protocol (Merck). Transfected cells were treated with the indicated concentration of control siRNA or FlexiTube GeneSolution for PRMT 1, 3 and 5 (Qiagen). siRNAs were implemented according to the manufacturer's protocol using HiPerFect transfection reagent (Qiagen). Cells were collected for western blot analysis 48 h after expression vector transfection.

Protein extraction from mammalian cells

Whole cells of 75–90% confluence in 6-well plates were harvested, washed in 1× phosphate-buffered saline (PBS) and suspended in lysis buffer [20 mM HEPES pH 7.5, 50 mM β-glycerolphosphate, 1 mM EDTA pH 8, 0.5 mM Na_3VO_4, 100 mM KCl, 10% glycerol, 1% Triton, anti-protesase cocktail (Roche 11697498001)]. These cell suspensiosn were mechanically lysed before centrifugation at 16 000 *g* for 20 min at 4°C, and the protein concentration was determined using the Bradford assay.

Western blotting

Equal protein quantities and volumes of all samples were loaded and run on Bolt or NuPAGE (polyacrylamide gel electrophoresis) 10% Bis-Tris Protein gels (Invitrogen) then transferred onto a 0.45 μM nitrocellulose membrane (GE Healthcare). Membranes were blocked in 1× PBS, 0.1% Igepal and 3% bovine serum albumin (BSA), and incubated with the indicated primary antibodies: mouse anti-Actin (Abcam ab3280, 1/5000), mouse anti-glyceraldehyde phoshate dehydrogenase (GAPDH; Abcam ab125247, 1/5000), mouse anti-EBNA1 (Cytobarr OTX1-EBNA1, 1/2000), rat anti-haemagglutinin (HA; Roche 11867423001, 1/2000), rabbit anti-NCL (Abcam ab70493, 1/5000), mouse anti-NSR1 (Abcam ab4642, 1/5000), rabbit anti-PRMT1 (Merck 07-404, 1/2000), rabbit anti-PRMT5 (Merck 07-405, 1/2000) and rabbit anti-PRMT3 (Abcam ab191562, 1/10000). The membranes were then washed with fresh 1× PBS, 0.1% Igepal and incubated with the

indicated secondary antibodies, conjugated to horseradish peroxidase: rabbit anti-mouse (Dako P0161, 1/3000), swine anti-rabbit (Dako P0217, 1/2000) and goat anti-rat (Millipore AP136P, 1/3000). The membranes were washed again and analysed by enhanced chemiluminescence in buffer (Tris-base pH 8.5, 12.5 nM coumaric acid, 2.25 nM luminol and 0.15% H_2O_2) using a Vilber-Loumart Photodocumentation Chemistart 5000 imager. All the experiments were repeated at least three times. Relative proteins levels for each sample were normaliazed to GAPDH or Actin protein levels as indicated, using Image J software.

T-cell proliferation assay

A total of 50 000 H1299 cells were seeded in 6-well plates and co-transfected the following day with 0.5 μg of Kb expression vector and 0.25 μg of 235GAr-OVA, OVA or control plasmids using Genejuice according to the manufacturer's protocol (Merck). Forty-eight hours later, the transfected cells were seeded in 24-well plates at a density of 1×10^5 in the presence of 1×10^5 cells of the B3Z T hybridoma in 2 ml final volume. After 48 h, the supernatant was withdrawn for an enzyme-linked immunosorbent assay (ELISA) to evaluate the interleukin-2 (IL2) production in order to estimate T-cell activation and therefore antigen presentation. IL2 levels were measured using the IL2 ELISA MAX™ Standard kit (Biolegend, USA) according to the manufacturer's instructions.

RNA pulldown

Yeast protein extracts. A 50 ml aliquot of 0.8–1.0 $OD_{600\ nm}$ exponentially growing cells was collected and cell pellets were suspended in 2 ml of lysis buffer [25 mM Tris–HCl, pH 7,4, 100 mM NaCl, 100 mM EDTA 0.1% Triton, antiprotease cocktail (Roche, 11697498001)]. After addition of 450–600 μm glass beads (Sigma-Aldrich, G8772), cells were lysed by six cycles of vortexing for 30 s followed by 30 s ice-cooling, shaken for 2 min at 25 Hz in a mixer mill (Retsch MM400) and centrifuged for 3 min at 800 *g* at 4°C. Supernatants were recovered and protein concentrations were determined using the Bradford assay.

Mammalian cell protein extracts. Cells were collected 48 h post-treatment and suspended in 200 μl of lysis buffer [20 mM HEPES, pH 7,5, 50 mM β-glycerophosphate, 1 mM EDTA, pH 8, 0.5 mM Na_3VO_4, 100 mM KCl, 10% glycerol, 1% Triton, antiprotease cocktail (Roche, 11697498001)]. Cells were lysed as mentioned above. Supernatants were recovered and protein concentrations were measured using the Bradford assay.

RNA pulldown experiments. Yeast extracts or mammalian cell extracts were used for pulldown assays with the following G4-forming oligonucleotides: GQ 5′-GGGGCAGGAGCAGGAGGA-3′-Biotin-TEG, GM 5′-GAGGCAGUAGCAGUAGAA-3′-Biotin-TEG and ARPC2 5′-AGCCGGGGGCUGGGCGGGGACCGG GCUUGU-3′Biotin-TEG. G4 were formed by heating the RNA oligonucleotides at 95°C during 5 min then cooling them down to 4°C at a rate of 2°C/min in the presence of 100 mM KCl to favour G4 formation. To avoid unspecific binding, high-affinity streptavidin–Sepharose beads (GE Healthcare, 28985799) were incubated in 1 ml of blocking buffer [10 mM Tris–HCl, pH 7.5, 100 mM KCl, 0.1 mM EDTA, 1 mM dithiothreitol (DTT), 0.01% Triton, 0.1% BSA, 0.02% *S. cerevisiae* tRNAs (Sigma-Aldrich, 10109495001)] for 1 h at 4°C on a rotating wheel. A 10 μg aliquot of each folded biotinylated RNA oligonucleotide was incubated with 50 μl of solution containing the streptavidin–Sepharose beads for 90 min at 4°C on a rotating wheel. Then 500 μg of cell extracts (mammalian or yeast) or 200 ng of recombinant NCL were treated with 200 U/ml of RNase inhibitor (NEB, M0307S) for 90 min at 4°C on a rotating wheel. These extracts or recombinant NCL were incubated with the RNA oligonucleotides bound to the streptavidin beads for 90 min at room temperature. Beads were then washed five times with lysis buffer and lysis buffer with increasing KCl concentrations (200–800 mM). Proteins still bound to beads after the washes were eluted using 2× loading buffer (2× Laemmli buffer with 5% β-mercaptoethanol) and analysed by western blotting against NCL or Nsr1.

Proximity ligation assay (PLA)

Cells were fixed in 1× PBS, 4% paraformaldehyde for 20 min and permeabilized for 10 min with 0.4% Triton X-100, 0.05% CHAPS. The EBNA1–digoxigenin probe mRNA (5′-CTTTCCAAACCACCCTCCTTTTTTGCGC CTGCCTCCATCAAAAA-3′) at 50 ng/well was denaturated for 5 min at 80°C. The probe hybridization reaction was carried out in 40 μl of hybridization buffer (10% formamide, 2× SCC, 0.2 mg/ml *E. coli* tRNA, 0.2 mg/ml salmon sperm DNA and 2 mg/ml BSA). The fixed cells were washed and blocked into the blocking solution (1× PBS, 3% BSA, 0.1% saponin) before incubation with the primary antibodies (anti-digoxigenin, Sigma 1/200 and anti-NCL, Abcam 1/1000). The PLA reaction was performed under the manufacturer's protocol using the Duolink PLA *in situ* kit, PLA probe anti-rabbit plus, PLA probe anti-mouse Minus and the *in situ* detection reagent FarRed, all from Sigma. The results were analysed using a Zeiss Axio Imager M2. All the PLA experiments were performed at least three times independently, and the following controls probes were implemented: without mRNA probe or without primary antibodies.

RNA extraction and quantitative reverse transcription–PCR (RT–qPCR)

Total H1299 or Mutu-1 cellular RNA was extracted using NucleoSpin® RNA Plus (Macherey Nagel). cDNA synthesis was carried out using 1 μg of RNA and M-MuLV Reverse Transcriptase (NEB) together with random primer. cDNA samples were analysed by qPCR using Master Mix PCR Power SYBR™ Green (Applied Biosystems™). The relative abundance of target mRNA was normalized using GAPDH as an endogenous control. Quantification of gene expression was determined using the $2^{-\Delta\Delta CT}$ method. The primers used for PCR were GAPDH-forward, 5′-GAGTCAACGGATTTGGTC

GT-3′; GAPDH-reverse, 5′-CACAAGCTTCCCGTTCTC AG-3′; EBNA1-forward, 5′-GGCAGTGGACCTCAAA GAAGAG-3′; and EBNA1-reverse, 5′- CAATGCAACT TGGACGTTTTTG-3′. All the experiments were performed in duplicate and repeated twice.

MTT assay

About 15 000 Mutu-1 cells were plated at 0.1 ml per well in 96-well, flat-bottom plates and exposed to the indicated compounds at the indicated concentrations or dimethylsulphoxide (DMSO; vehicle). After 48 h, 10 μl of MTT solution [5 mg/ml in PBS (pH 7.4), CT01-5, Merck Millipore] was added to each well and cells were incubated for 4 h. A mixture of isopropanol/0.1 N HCl/10% Triton X-100 (0.1 ml) was added to each well to dissolve the formazan crystals, and the absorbance was then measured at 540 nm.

Immunofluorescence

H1299 cells were plated on 13 mm diameter coverslips in 24-well plates and were transfected with HA-NCL-wt, HA-NCL-R10A or HA-NCL-R10F plasmids. At 24 h post-transfection, cells were fixed with 4% paraformaldehyde in PBS for 20 min and permeabilized with 0.4% PBS, Triton X-100, 0.05% CHAPS for 10 min at room temperature. After incubation in blocking buffer (1× PBS, 3% BSA, 0.1% saponin) for 30 min at room temperature, samples were incubated with a mouse polyclonal anti-HA antibody (a kind gift from Borek Vojtesek, Masaryk Memorial Cancer Institute, Brno, Czech Republic) at 1/1000 for 2 h at room temperature followed by incubation with a 1/500 dilution of the goat anti-mouse immunoglobulin G (IgG) secondary antibody conjugated to Alexa Fluor® 594 (Invitrogen). Both antibodies were diluted in blocking buffer. 4′,6-Diaminidino-2-phenylindole (DAPI) was used for nuclear counterstaining and the images were taken using a Zeiss Axio Imager M2. The experiments were repeated twice.

NCL methylation

NCL recombinant protein was prepared and purified as previously described (32). Methylation of purified NCL recombinant protein was performed in a buffer containing 50 mM Tris pH 7.5, 1 mM EDTA, 200 mM NaCl. NCL (5 μM) was incubated with 50 nM SAM (Sigma-Aldrich, Saint-Quentin-Fallavier) and 240 μg/ml of recombinant human PRMT1 protein (ab89007, Abcam) at 30°C for 3 h. The conditions of the methylation reaction i.e. time, buffer and stoichiometry of the partners (SAM:NCL:PRMT = 10:1:1), were previously determined through preliminary qualitative experiments using [^{3}H]SAM followed by nitrocellulose membrane filtration (Bio-dot). Methylated NCL was purified by using a HiLoad Superdex® column and stored in 50 mM Tris pH 8, 200 mM NaCl buffer at −80°C. Mass spectrometry analysis was performed to identify the number and sites of methylation.

Electrophoretic mobility shift assay

The annealed EBNA1 and EBNA1-2rep were prepared by mixing 5′-radiolabelled [γ-^{32}P]oligonucleotide with non-radiolabelled oligonucleotide (100 nM) in K-100 buffer [10 mM LiCaco2, 100 mM KCl (pH 7.2)], then heating at 95°C for 5 min followed by slow cooling. For protein or peptide binding, the annealed RNAs (final concentration of 20 nM) were incubated with proteins or peptides in the desired concentration range (0–1 μM for proteins and 0–56 nM for peptides) in 1× binding buffer [10 mM Tris–HCl (pH 7.5), 1 mM EDTA, 100 mM KCl, 0,1 mM DTT, 5% glycerol and 0.01 mg/mL BSA] at room temperature for 1 h. The samples were loaded on a 10% or 15% polyacrylamide native gel (37.5:1 acrylamide:bis ratio) for recombinant proteins or peptides, respectively. After electrophoresis, the gels were exposed in a phosphoimager cassette and scanned on a Typhoon Trio Variable mode imager. The quantitative gel was fitted in GraphPad Prism7 with the Hill slope. Representation of means and standard deviation (SD) values are from three independent experiments.

Sequences of peptides and oligonucleotides for EMSAs

RGG, H-EGGFGGRGGGRGGFGGRGGGRGGRGG FGGRGRGGFGGRGGFRGGRGGGGD-OH

RGG-met; H-EGGFGG(ADMA)GGG(ADMA)GG FGG(ADMA)GGG(ADMA)GG(ADMA)GGFG GADMA)G(ADMA)GGFGG(ADMA)GGFADM A)GG(ADMA)GGGGD-OH (where ADMA is asymmetrically dimethylated arginine); EBNA1, GGG-GCA-GGA-GCA-GGA-GGA; EBNA1-2repeat, GGG-GCA-GGA-GCA-GGA-GGA-GGG-GCA-GGA-GCA-GGA-GGA

Oligonucleotides were purchased from Eurogentec (HPLC purification). Peptides were synthesized by Pepscan (HPLC purification).

Statistics analyses

Data were analysed by analysis of variance (ANOVA) in conjunction with Tukey's test using GraphPad Prism 5 Software. Data shown are the mean ± SD. *P <0.05; ***P <0.0001; ns, not significant.

RESULTS

The C-terminal RGG motif of yeast nucleolin Nsr1 is necessary for its role in GAr-based inhibition of translation and for the ability of yeast and human nucleolin to bind G4 of EBNA1 mRNA

As stated above, Nsr1, the budding yeast (*S. cerevisiae*) nucleolin, shares the same domain organization as NCL, the human nucleolin, apart from the fact that Nsr1 contains only two central RRMs, instead of four in NCL, which probably correspond to RRM3 and 4 of NCL (Figure 1A). In addition, GAr-based inhibition of translation is fully operational in budding yeast (43) where it also involves nucleolin (20). This explains the relevance of the yeast-based genetic screen that led to the identification of nucleolin as the first host cell factor critically involved in GAr-based inhibition of translation (21). This study also shows that human

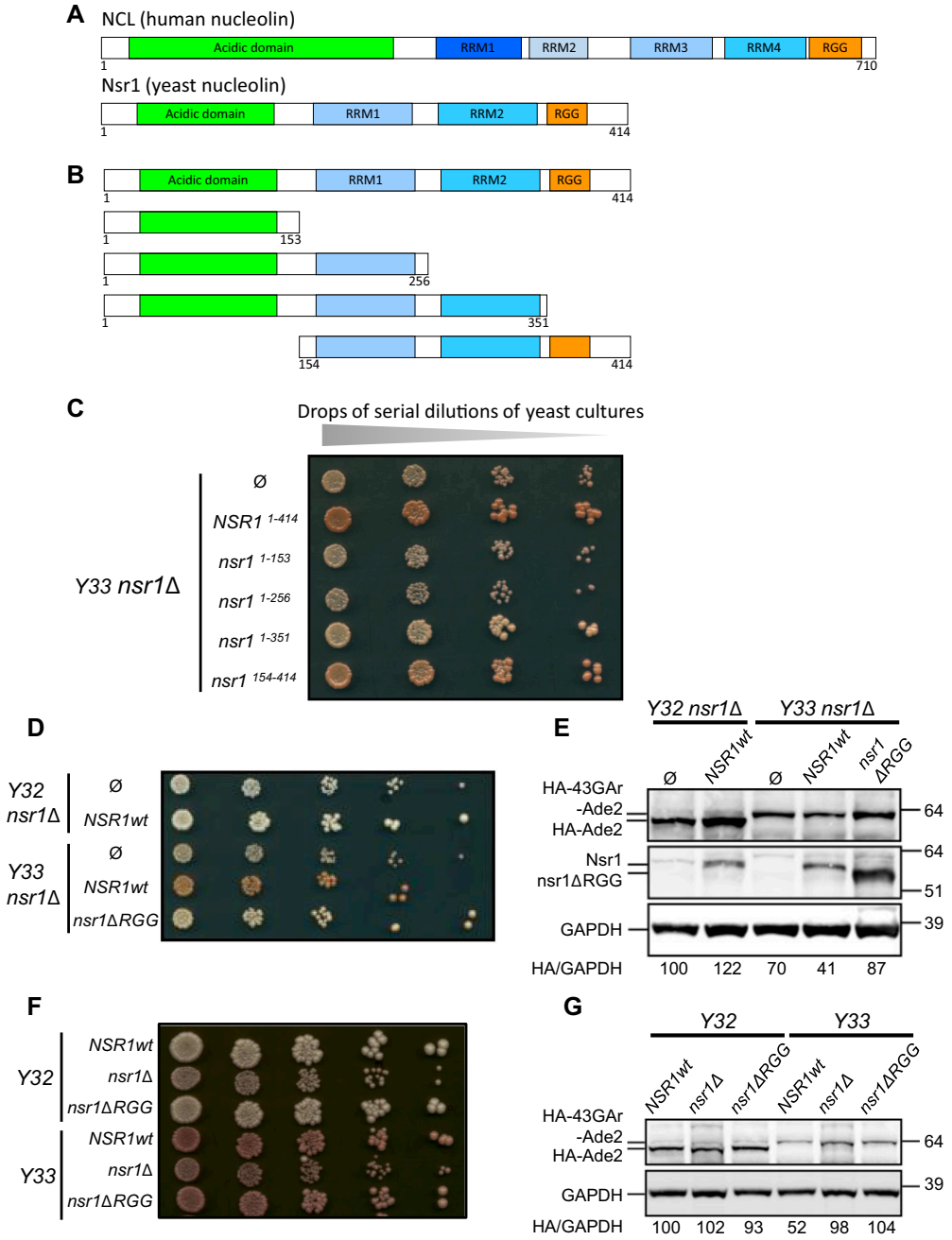


Figure 1. The C-terminal arginine–glycine rich (RGG) motif of yeast nucleolin Nsr1 is necessary for its role in GAr-based inhibition of translation. (**A**) Schematic representation of the various domains of the human (NCL) and yeast (Nsr1) nucleolin. Note that the two proteins share the same organization, in particular central RRMs (four for NCL and only two for Nsr1 which are probably orthologues of RRM3 and 4, hence the colour code) and the C-terminal RGG motif. Both RRM and RGG are known as RNA-binding motifs. (**B**) Schematic representation of the various constructions used for the domain analysis of the yeast nucleolin Nsr1. (**C**) Analysis of the ability of these various constructs to complement the role of Nsr1 in GAr-based inhibition of translation in a *Y33 nsr1Δ* strain, as compared with the wild-type Nsr1. Serial dilutions of the various transformed strains were spotted onto agar-based solid medium. The colour of yeast colonies is used as the readout. Note that all the constructs that lack the C-terminus that contains the RGG motif are unable to complement the deletion of the *NSR1* gene, as indicated by the white colour of the yeast colonies. (**D**) Ability of a form of yeast nucleolin deleted of its C-terminal RGG motif (nsr1ΔRGG) to complement the role of Nsr1 in GAr-based inhibition of translation in a *Y33 nsr1Δ* strain as compared with the wild-type Nsr1. Briefly, a yeast strain that expresses HA-43GAr-Ade2 or, as a control, HA-Ade2 as sole source of Ade2, and deleted for the *NSR1* gene was transformed with either an empty vector (negative control), or a plasmid allowing expression of full-length Nsr1wt (positive control), or of nsr1ΔRGG, as indicated. Serial dilutions of the various transformed strains were spotted onto agar-based solid medium and the colour of the resulting transformed strains was assessed. A full level of expression of HA-Ade2 leads to the formation of white colonies, whereas the reduced level of HA-43GAr-Ade2 due to the GAr-based inhibition of translation leads to pink colonies. (**E**) Western blot analysis of the level of Ade2 or HA-43GAr-Ade2 in the same strains. HA-Ade2/GAPDH or HA-43-GAr-Ade2/GAPDH ratios are indicated below the gels. (**F**) *Y33 nsr1ΔRGG*, a yeast strain in which only the sequence encoding the RGG motif was deleted from the endogenous chromosomal *NSR1* gene, displays the same pink phenotype as *Y33 nsr1Δ* regarding GAr-based inhibition of translation but grows like a *Y33 NSR1wt* strain, indicating that nsr1ΔRGG is functionally expressed. (**G**) Western blot analysis of the level of Ade2 or HA-43GAr-Ade2 in the same strains. HA-Ade2/GAPDH or HA-43-GAr-Ade2/GAPDH ratios are indicated below the gels.

NCL can complement the effect of the loss of yeast Nsr1 on GAr-based inhibition of translation in yeast. Taken together, these results indicate that yeast is a relevant system to model NCL- and GAr-based inhibition of translation. Importantly, and in contrast to NCL in human cells, Nsr1 is not an essential gene, as yeast *nsr1Δ* strains, although growing quite poorly, are viable (21,44). For all these reasons, we decided to use yeast *nsr1Δ* strains to determine the involvement of the various domains of nucleolin in GAr-based inhibition of translation, as this cellular context allows clear-cut situations where the only source of nucleolin is either full-length Nsr1, or parts of it, that were expressed back in yeast. We also made use of the yeast model for GAr-based inhibition of translation which we recently created (43). This model is based on a fusion between a GAr domain of 43 amino acids (43GAr) and the yeast Ade2 reporter protein. Because GAr is able to self-inhibit the translation of its own mRNA in yeast, this leads to a reduction in Ade2 level (43). This can easily be monitored as yeast which express Ade2 at a functional level form white colonies, whereas yeast that do not express Ade2 readily form red colonies, and any intermediate level of Ade2 leads to the formation of pink colonies whose intensity of coloration is inversely proportional to the level of Ade2 expressed. Hence, a yeast strain expressing the *HA-43GAr-ADE2* construct (*Y33* strain) forms pink colonies and exhibits an intermediate level of HA-43GAr-Ade2 protein, whereas a control strain expressing *HA-ADE2* from the same promoter (*Y32* strain) forms white colonies that contain a higher level of HA-Ade2 protein (43). Importantly, the deletion of the *NSR1* gene in the *Y33* strain (hence termed *Y33 nsr1Δ*) led to the formation of white colonies that express a high level of HA-43GAr-Ade2 as, in yeast too, the GAr-based inhibition of translation depends on nucleolin. In contrast, the deletion of the *NSR1* gene had no effect on HA-Ade2 expression in Y32 strains (hence termed *Y32 nsr1Δ*), demonstrating that the effect of nucleolin on protein expression is GAr dependent (21). Starting from the *Y33 nsr1Δ*, we first determined the effect of re-expressing, from a low copy number centromeric plasmid, full-length Nsr1, or various parts of it (Figure 1B), on the GAr-based inhibition of translation, using as readouts the colour of the various yeast strains and the level of expression of 43GAr-Ade2 or Ade2. As shown in Figure 1C and D, as expected, reintroduction of full-length Nsr1 (1–414) led to pink colonies that express a reduced level of HA-43GAr-Ade2, similar to the original *Y33* strain. The same pink colour was obtained when expressing a form of Nsr1 deleted of its N-terminal acidic domain (154–414), indicating that this domain is not required for the involvement of Nsr1 in GAr-based inhibition of translation (Figure 1C). In contrast, the *Y33 nsr1Δ* strain expressing various forms of Nsr1 that all lack the C-terminal end containing the RGG motif [1–351, that lacks only the C-terminal RGG motif (hence termed nsr1ΔRGG in the following); 1–256, that lacks both RRM2 and RGG; and 1–153, that lacks RRM1, RRM2 and RGG] remains white, suggesting that the inhibitory effect of GAr on protein expression was not restored (Figure 1C). This was checked for the *Y32 nsr1Δ* and *Y33 nsr1Δ* expressing nsr1ΔRGG (Figure 1D, E), indicating that the RGG motif is necessary for the crucial role of Nsr1 in GAr-based inhibition of translation. To confirm this result, we deleted only the RGG motif-encoding sequence from the endogenous chromosomal copy of the *NSR1* gene in the *Y33* strain (to yield *Y33 nsr1ΔRGG*), and in the *Y32* strain (to yield *Y32 nsr1ΔRGG*) as a control. As shown in Figure 1F and G, the *Y33 nsr1ΔRGG* strain displays the same phenotype (white colonies expressing a higher level of 43GAr-Ade2) than the *Y33 nsr1Δ* strain regarding GAr-based inhibition of translation. Of note, the *Y33 nsr1ΔRGG* strain grows like the *Y33 NSR1wt* strain, whereas the *Y33 nsr1Δ* strain exhibits a significant growth defect as reported by us and others (21,44). This shows that the RGG motif is necessary for the role of Nsr1 in GAr-based inhibition of translation, but not for its other functions required for an optimal yeast growth, in good agreement with recently published results (45). In contrast, the results shown in Figure 1C suggest that the two RRMs (RRM1 and 2) are necessary for optimal yeast growth of the *Y33 nsr1Δ* strain as only the forms of Nsr1 that lack RRM1 or both RRM1 and 2, in addition to the C-terminal RGG motif, display the same small colony phenotype as the control strain transformed with an empty vector. Importantly, this result also indicates that the defect in GAr-based inhibition of translation of the yeast strains deleted for the *NSR1* gene cannot be attributed to their slow growth phenotype. It also shows that the role of Nsr1 in GAr-based inhibition of translation can be genetically separated from its central role in cell fitness.

Because the C-terminal RGG motif of Nsr1 is required for its role in GAr-based inhibition of translation, and as we previously showed that the direct interaction between human nucleolin NCL and G4 of EBNA1 mRNA is required for the GAr-based limitation of both translation and antigen presentation (21), we then compared the ability of the Nsr1 and nsr1Δrgg proteins to interact with G4 of EBNA1 mRNA. For this purpose, we performed RNA pulldown assays as previously described (21,46). Briefly, this assay is based on the use of an RNA oligonucleotide containing a sequence that can form G4, conjugated to biotin, which can therefore be precipitated by magnetic Sepharose beads conjugated to streptavidin (Figure 2A). In this way, proteins which present affinity for the selected RNA-G4 can be precipitated and detected by western blotting. We first determined if yeast nucleolin Nsr1 does bind to G4 of EBNA1 mRNA, as has been shown previously for human nucleolin NCL (21), and if this binding depends on the presence of its C-terminal RGG motif. For this purpose, lysate from *Y32* or *Y32 rggΔ* cells was applied to the following matrices: streptavidin-coupled beads either together with GQ (containing the most probable G4 of the GAr-encoding sequence of EBNA1 mRNA), GM (negative control: same sequence except that four guanines critical for G4 formation were replaced by adenine or uracil) or ARPC2 (positive control containing a G4 present in *ARPC2* mRNA and that has been shown to bind NCL) RNA oligonucleotides (Figure 2A). As shown in Figure 2B, an efficient binding of Nsr1 to the prototypical G4 found in EBNA1 mRNA was observed, which is fully in line with the already described binding of human nucleolin NCL on the same matrix (21). As for NCL, this binding mostly depends on G4 since only a residual binding was observed on GM control that cannot

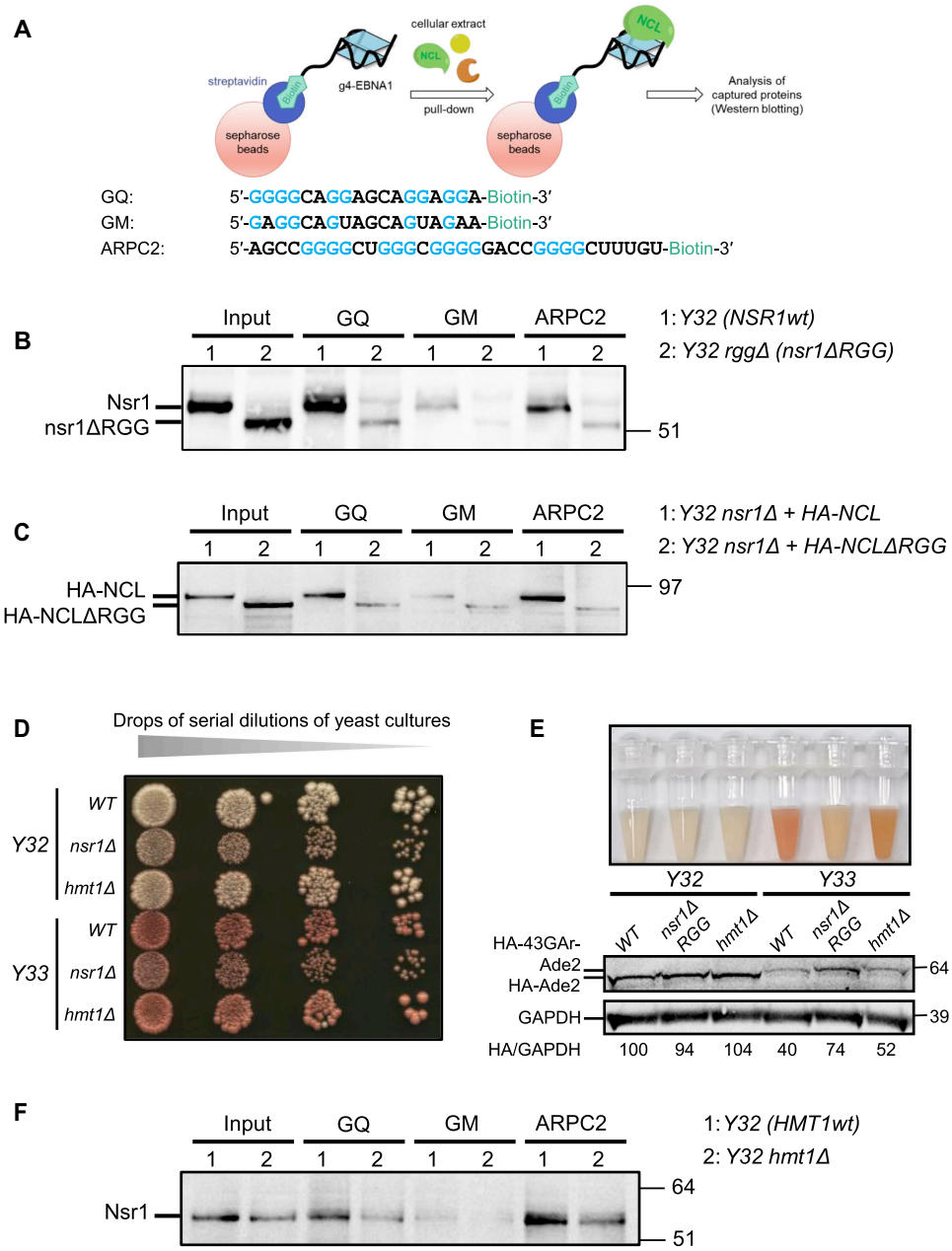


Figure 2. The C-terminal RGG motif of yeast nucleolin Nsr1 is necessary for its binding to G4 of EBNA1 mRNA which also depends on the main yeast type I arginine methyltransferases Hmt1. (**A**) Scheme depicting the principle of the RNA pulldown experiment. The sequence of the various RNA oligonucleotides used is shown and the guanines (G) implicated in the formation G4 are highlighted in blue. GQ: RNA oligonucleotide containing the prototypal repeated sequence which forms the most probable G4 in GAr-encoding sequence of EBNA1 mRNA. GM (negative control): mutated version of GQ in which four guanines were replaced by adenine or uridine to prevent formation of G4. ARPC2 (positive control): RNA oligonucleotide containing the sequence of *ARPC2* mRNA that forms a G4 known to bind NCL (46). (**B**) RNA pulldown of extracts from yeast cells expressing Nsr1wt [1: *Y32* (*NSR1wt*)] or nsr1ΔRGG [2: *Y32 rggΔ* (*nsr1ΔRGG*)] as the sole source of nucleolin. The protein still bound after an 800 mM KCl wash were eluted and analysed by SDS–PAGE and western blot. Like the human nucleolin NCL, the yeast nucleolin Nsr1 binds to G4 of *EBNA1* and *ARPC2* mRNA, and this binding is G4 specific as only residual binding was found on GM beads. In contrast, only residual binding is observed for nsr1ΔRGG. The blot represents $n \geq 3$. (**C**) The same experiment as in (B), except that RNA pulldowns were performed using extracts from *nsr1Δ* yeast cells expressing human HA-tagged wild-type nucleolin, HA-NCL (1: *Y32 nsr1Δ + HA-NCL*) or a version deleted of its C-terminal RGG motif, NCLΔRGG (2: *Y32 nsr1Δ + HA-NCLΔRGG*) as the sole source of nucleolin. In contrast to NCL which binds to G4 of both EBNA1 and ARPC2 in a G4-dependent manner, only residual binding is observed for NCLΔRGG. The blot represents $n \geq 3$. (**D**) A similar experiment to that shown in Figure 1C was performed and indicates that the ability of GAr to limit protein expression is lower in an *hmt1Δ* yeast strain (Y33 *hmt1Δ*) as compared with the *HMT1wt* Y33 strain. (**E**) The same yeast strain as in (D) was grown in synthetic liquid medium up to stationary phase and photographed. In addition, western blot analysis of the level of Ade2 or HA-43GAr-Ade2 was performed. HA-Ade2/GAPDH or HA-43-GAr-Ade2/GAPDH ratios are indicated below the gels. The blot represents $n \geq 3$. This experiment confirms that the ability of GAr to limit protein expression is less in an *hmt1Δ* yeast strain (Y33 *hmt1Δ*) as compared with the *HMT1wtY33* strain. (**F**) The same RNA pulldown experiment as in (B) and (C) was performed except that the following yeast extracts were used: 1: *Y32* (*HMT1wt*) or 2: *Y32 hmt1Δ*. The binding of the yeast nucleolin Nsr1 on G4 of EBNA1 mRNA is significantly decreased when the *HMT1* gene is absent. The deletion of the *HMT1* gene also affects the binding of Nsr1 on ARPC2 mRNA G4 (positive control).

form G4, whereas a significant binding was also observed on ARPC2 matrix (positive control that forms RNA-G4 that binds NCL). Importantly, only a residual binding on GQ, comparable with the binding of Nsr1 on GM, was observed for nsr1Δrgg, whereas Nsr1 and nsr1Δrgg (which were expressed from the *NSR1* endogenous promoter in their normal chromosomal context) were both expressed at a similar level (input, two first lanes). The same result was observed with ARPC2 RNA oligonucleotide, confirming that the binding of yeast nucleolin Nsr1 on RNA-G4 depends on its C-terminal RGG motif.

We then performed the same experiments with human nucleolin NCL. For this purpose, we expressed NCL, or NCLΔRGG, from the constitutive *ADH* promoter in the *Y32 nsr1Δ* strain. We chose to express NCL, or NCLΔRGG, in an *nsr1Δ* yeast strain to ensure that the sole source of nucleolin is the one we expressed. Indeed, in human cells in which NCL is an essential gene whose level of expression cannot be significantly modified, dealing with a mixture of NCL and NCLΔRGG would be inevitable, thereby complicating the analysis. Hence, here again, we made use of yeast genetics and of the fact that nucleolin is not essential in yeast. The *ADH* promoter was chosen because it leads to expression of human nucleolin NCL which is comparable with that of yeast nucleolin Nsr1. As shown in Figure 2C, we obtained the same result with NCL and NCLΔRGG as when using yeast nucleolin Nsr1 and nsr1Δrgg: NCL bound efficiently to the GQ matrix whereas only a residual binding was observed for NCLΔRGG, although it was repeatedly more expressed than NCL in yeast (input, first two lanes). The same result was observed with ARPC2 RNA oligonucleotide, confirming that the binding of human nucleolin NCL on RNA-G4, in particular on G4 of EBNA1 mRNA, also depends on its C-terminal RGG motif.

Hence, we concluded that the binding of both yeast and human nucleolin (Nsr1 and NCL, respectively) on the G4 of the GAr-encoding sequence of EBNA1 mRNA requires their C-terminal RGG motif. The residual binding observed for wild-type Nsr1 and NCL on the GM matrix, or for NCLΔRGG and nsr1Δrgg on the GQ or ARPC2 matrix, is probably due to the RRM domains which possess an intrinsic RNA binding affinity, whereas the RGG motif seems to be crucial for the specific affinity of nucleolin for RNA-G4.

The main yeast type I arginine methyltransferase Hmt1p is necessary for efficient GAr-based inhibition of translation and binding of nucleolin to G4 of EBNA1 mRNA

In several instances, methylation of RGG motifs has been reported to interfere, positively or negatively, with their ability to interact with other proteins as well as with nucleic acids (34). RGG is the canonical substrate motif established for Hmt1p, a type I PRMT which is the main yeast PRMT and responsible for 66% of arginine monomethylation and 89% of asymmetric dimethylation of intracellular proteins in this model eukaryote. Since it is estimated that, in budding yeast, the majority of arginine methylation occurs on RGG motifs, we tested the impact of inactivating the *HMT1* gene on the GAr-based inhibition of translation and on the ability of Nsr1 to interact with G4 of EBNA1 mRNA. For this, we deleted the *HMT1* gene in both *Y32* and *Y33* strains to yield *Y32 hmt1Δ* and *Y33 hmt1Δ* strains, respectively. We then observed the impact of *HMT1* gene deletion on GAr-based inhibition of protein expression using the same assay as in Figure 1C, D and F. As shown in Figure 2D and E, the *Y33 hmt1Δ* strain presents a whiter phenotype than the control *Y33* strain, a phenotype that is comparable, although less pronounced, with that of the *Y33 nsr1Δ* strain, suggesting that the ability of 43GAr to inhibit expression of 43GAr-Ade2 is partially compromised in the *Y33 hmt1Δ* strain. As we observed no difference in colour in the control strains *Y32* and *Y32 hmt1Δ*, we concluded that the partial loss of GAr-based inhibition of protein expression observed in the *Y33 hmt1Δ* strain, as compared with the *Y33* strain, is GAr dependent. In addition, using western blot analysis, we determined the level of Ade2 and 43GAr-Ade2 proteins in *hmt1Δ* and *nsr1ΔRGG* strains, as compared with wild-type controls in both Y32 and Y33 backgrounds (Figure 2E). We found that 43GAr-Ade2 was slightly more expressed in *Y33 hmt1Δ* as compared with *Y33 HMT1wt*, in good agreement with the slightly whiter phenotype observed on plates (Figure 2D) and in liquid culture (Figure 2E, upper panel). This suggests that modulation of the methylation of the RGG motif of Nsr1 may interfere with its ability to interact with G4 of the GAr-encoding sequence of *EBNA1* mRNA. To test this hypothesis, we performed the same RNA pull-down experiment as described in Figure 2B, except that we compared the ability of Nsr1 extracted from *Y32* or *Y32 hmt1Δ* strains to interact with G4 of EBNA1 mRNA, the G4 of *ARPC2* mRNA (positive control) or the GM RNA oligonucleotide (negative control). As shown in Figure 2F, the ability of Nsr1 to interact with both GQ and ARPC2 RNA oligonucleotides is significantly decreased when Nsr1 is extracted from the *Y32 hmt1Δ* strain as compared with the *Y32* strain. Of note, this decrease is not as strong as that observed in Figure 2B, when comparing the binding of Nsr1 with that of nsr1Δrgg. This is in line with the partial loss of GAr-based inhibition of protein expression observed in the *Y33 hmt1Δ* strain and with the fact that Hmt1 is not the sole PRMT in yeast, which implies that the effect on RGG methylation might only be partial in *hmt1Δ* strains.

We concluded that the requirement of the RGG motif of Nsr1 for the role of this protein in GAr-based inhibition of translation and for its ability to interact with G4 of EBNA1 mRNA in yeast both depend on an optimal type I protein arginine methyltransferase (PRMT) activity.

Inhibition of human type I protein arginine methyltransferases impacts GAr-based inhibition of protein expression

We next assessed if inhibition of human PRMTs, in particular type I PRMTs (since Hmt1p is a type I PRMT), also affects GAr-based limitation of translation. For this purpose, we first determined the effect of drug- or siRNA-based inhibition of PRMTs, in particular type I PRMTs, on GAr-based limitation of protein expression. Hence, we tested the effect of 7,7′-carbonylbis(azanediyl)bis(4-hydroxynaphthalene-2-sulphonic acid) (AMI-1) or 5′-

deoxy-5′(methylthio)adenosine (MTA) molecules on the level of EBNA1 or EBNA1ΔGAr in H1299 cells. AMI-1 was the first identified pharmacological compound targeting endogenous methyltransferases and has been isolated on the basis of its ability to inhibit both yeast and human type I protein arginine methyltransferases Hmt1p and PRMT1 (47). It is generally considered that AMI-1 selectively inhibits type I PRMTs (PRMT1, 3, 4 and 6), but it has been suggested that it could also inhibit PRMT5, the main type II PRMT (48). In contrast, MTA is a natural metabolite that is considered to be a general inhibitor of the three types of PRMTs (49), in particular of PRMT5 (50,51). As shown in Figure 3, treatment of H1299 cells transfected with vectors allowing expression of EBNA1 (Figure 3A), or EBNA1ΔGAr (Figure 3B), by AMI-1 or MTA, led to a GAr-dependent increase in EBNA1 expression which was even more pronounced than that induced by PhenDH2 or PyDH2, two G4 ligands which have been optimized for binding to G4 of EBNA1 mRNA as well as for their ability to outcompete NCL for binding on these G4 (22). These results confirmed those obtained in yeast and indicate that type I PRMTs are important for GAr-based inhibition of protein expression.

To further test this hypothesis, we next examined the effect of MS023, a recently isolated PRMT inhibitor that has been shown to specifically inhibit type I PRMTs (52). MS023 gave similar results to AMI-1 (Figure 3C), whereas it had no noticeable effect on EBNA1ΔGAr (Figure 3D). We have also tested the effect of MS023 on GAr-OVA or OVA proteins as it has been shown that GAr is able to limit expression of OVA. As shown in Figure 3E and F, MS023 strongly increased the expression of GAr-OVA whereas it had no noticeable effect on OVA. As a control, we determined the effect of MTA and MS023 on NCL expression, as a decrease in NCL level upon treatment with these PRMTs inhibitors might readily explain their effect on EBNA1 expression. Importantly, neither MTA nor MS023 had an effect on the NCL protein level (Figure 3G). Finally, we also determined that MS023 had no effect on EBNA1 and EBNA1ΔGAr mRNA levels (Figure 3H; Supplementary Figure S1). Altogether, these experiments confirm that inhibition of type I PRMTs specifically interferes with GAr-based inhibition of protein expression.

We then tested the effect of down-regulating various PRMTs using siRNA. We first determined the effect of down-regulating, independently or together, PRMT1 and PRMT3, the two main type I PRMTs. As shown in Figure 4A and C, down-regulation of PRMT1 alone, or of PRMT3 alone, led to a significant and comparable increase in EBNA1 expression, whereas down-regulation of both PRMT1 and 3 led to a slightly stronger effect. All these effects were GAr specific since no significant effect was observed for EBNA1ΔGAr (Figure 4B, D). We also tested the effect of down-regulating PRMT5, the main type II PRMT, and found that it had no significant effect on EBNA1 expression (not shown). In line with this, we found no additive effect when down-regulating PRMT5 together with PRMT1.

Taken together, all these results confirm that inhibition of type I PRMTs (by drugs or siRNA) interferes with GAr-based inhibition of translation.

Inhibition of human type I protein arginine methyltransferases impacts the ability of NCL to interact with G4 of EBNA1 mRNA

As inhibition of type I PRMTs impacts the GAr-based inhibition of protein expression in both yeast and human cells, we next determined, using RNA pulldown experiments and proximity ligation assay (PLA), if inhibition of PRMTs also interferes with the binding of human nucleolin NCL on G4 of EBNA1 mRNA to confirm the results obtained for yeast nucleolin Nsr1 in Figure 2F. We first performed the same pulldown experiments as in Figure 2 except that, instead of yeast cell lysates, we used cell lysates from human H1299 cells treated with MTA or MS023, or with vehicle (DMSO) as a negative control. As shown in Figure 5A, the treatment with MTA or MS023 decreases in a dose-dependent manner the binding of NCL to G4 of EBNA1 mRNA, whereas it had no significant effect on the steady-state level of NCL (Figure 3G). This result indicates that type I PRMT activity is required for efficient binding of NCL on G4 of EBNA1 mRNA. Next, we wanted to verify if the inhibition of PRMTs affects the NCL–GAr G4 interaction *in cellulo*. Hence, we performed a PLA adapted to protein–RNA interaction (53) to assess if MS023 (Figure 5B) interferes with the binding of endogenous NCL on endogenous EBNA1 mRNA in EBV-infected Mutu-1 cells. As previously reported (21,53), we observed PLA dots mostly in the nucleus, or close to it, in Mutu-1 cells treated with vehicle (DMSO), confirming that the NCL–EBNA1 mRNA interaction essentially takes place in the nucleus or in the cytoplasm in the vicinity of the nucleus. In contrast, treatment with 1 μM MS023 led to a significant decrease in the number of PLA dots per cell, reaching a value comparable with that obtained for the negative control without probe (Figure 5B). Similar results were obtained with treatment with 5 μM MTA (Supplementary Figure S2).

Taken together, these results indicate that type I PRMTs are required, both in yeast and in human cells, for the interaction between nucleolin and G4 of EBNA1 mRNA. As this interaction is crucial for GAr-based inhibition of translation, this readily explains why inactivating type I PRMTs interferes with GAr-based inhibition of translation.

Changing arginines of the RGG motif of both Nsr1 and NCL to alanines suppresses GAr-based inhibition of protein expression and interaction of NCL with G4 of EBNA1 mRNA, whereas replacement by the bulky hydrophobic phenylalanines mimicking methylated arginines maintains both these phenomenon

Our results indicate that the binding of nucleolin to G4 of EBNA1 mRNA depends both on its C-terminal RGG motif and on an optimal type I protein arginine methyltransferase (PRMT) activity. The binding of nucleolin to the G4 of EBNA1 mRNA is direct (21). In addition, arginines of RGG motifs are known to be the main substrate of type I PRMT. Therefore, the most likely hypothesis is that type I PRMTs methylate arginines of the C-terminal RGG motif of nucleolin, thereby favouring its interaction with G4 of EBNA1 mRNA. To test this possibility, we again made use of the fact that the yeast nucleolin Nsr1 is not essential and generated a yeast strain that expresses a mutant

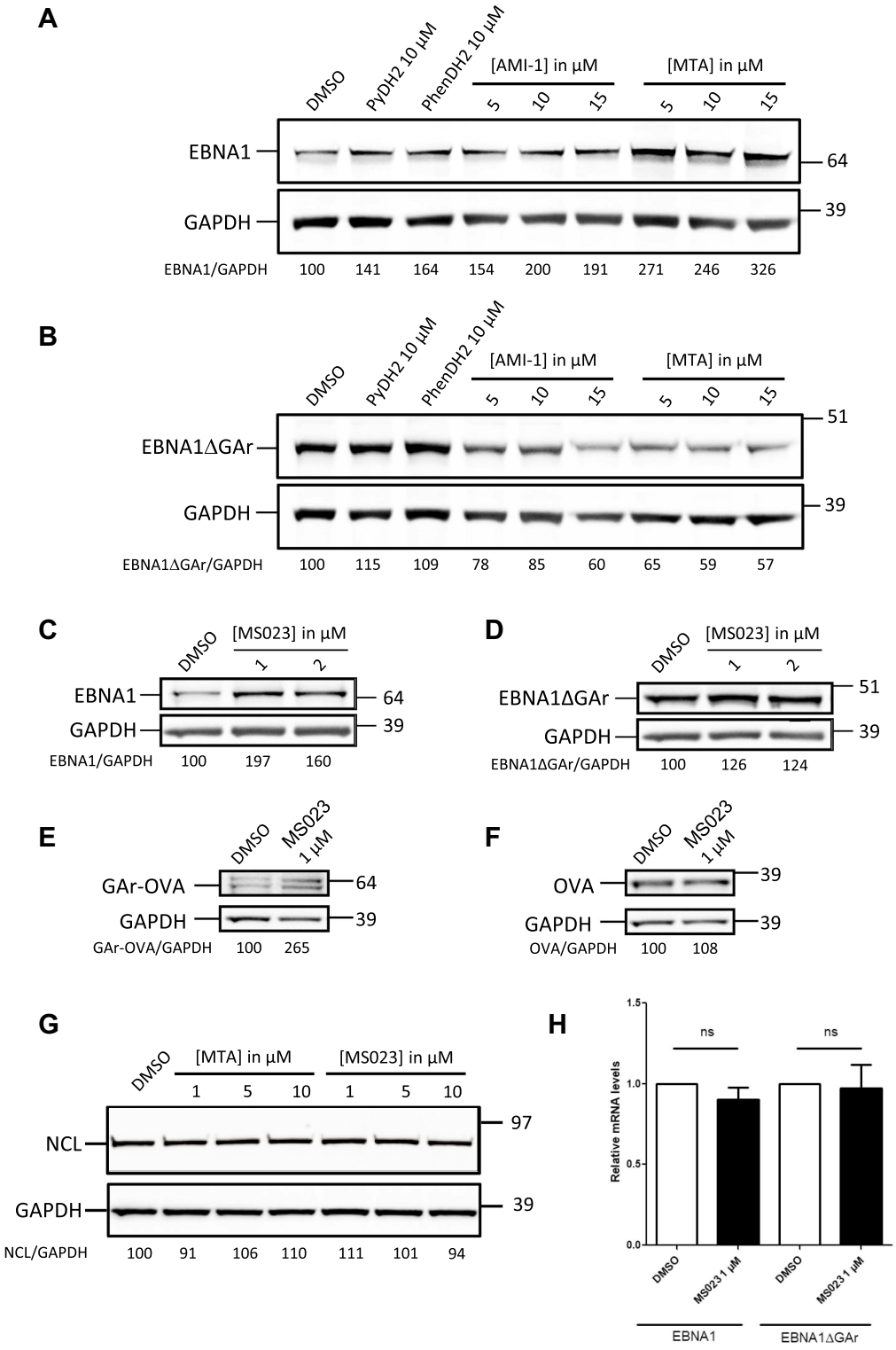


Figure 3. Inhibition of human type I arginine methyltransferases using specific inhibitors impacts GAr-based inhibition translation and has no effect on NCL level. SDS–PAGE and western blot analysis of the level of EBNA1 or EBNA1ΔGAr in response to treatments with various inhibitors of arginine methyltransferases. H1299 cells were transfected with EBNA1- (**A** and **C**) or EBNA1ΔGAr-expressing vectors (**B** and **D**) and treated, or not, with the various drugs, as indicated. GAPDH was used as a loading control. EBNA1/GAPDH or EBNA1ΔGAr/GAPDH ratios are indicated below the gels. The G4 ligands PyDH2 and PhenDH2 were used as positive controls in (A) and (B). SDS–PAGE and western blot analysis of the level of GAr-OVA or OVA in response to treatments with MS023. H1299 cells were transfected with GAr-OVA- (**E**) or OVA-expressing vectors (**F**) and treated, or not, with MS023, as indicated. GAPDH was used as a loading control. GAr-OVA/GAPDH or OVA/GAPDH ratios are indicated below the gels. The NCL steady-state level was determined in H1299 cells treated with various concentrations of MTA or MS023, as indicated (**G**). None of the drugs impacts NCL level. (**H**) Levels of EBNA1 or EBNA1ΔGAr as determined using RT–qPCR.

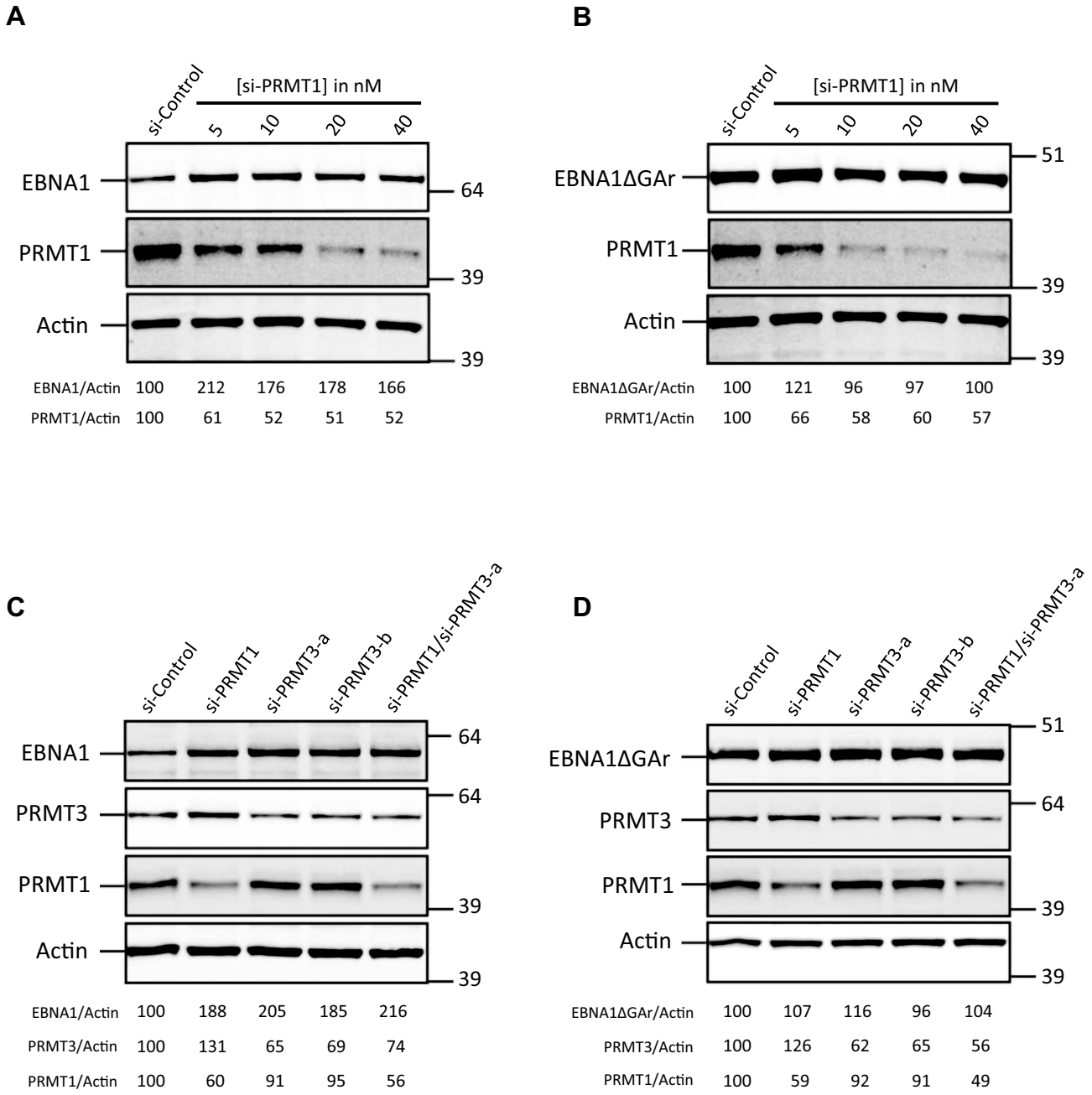


Figure 4. Inhibition of human type I arginine methyltransferases using siRNA impacts GAr-based inhibition of protein expression. SDS–PAGE and western blot analysis of the level of EBNA1 or EBNA1ΔGAr in response to knockdown of various PRMTs. H1299 cells were transfected with EBNA1- (**A** and **C**) or EBNA1ΔGAr-expressing vectors (**B** and **D**) and with control siRNA or siRNA targeting various PRMTs, as indicated. Actin was used as a loading control. EBNA1/Actin or EBNA1ΔGAr/Actin ratios are indicated below the gels. Blots represent $n \geq 3$.

form of Nsr1 in which all eight arginines of its RGG motif were replaced by alanines (nsr1-R8A) as the sole source of nucleolin. As shown in Figure 6A, this strain (*Y33 nsr1Δ + nsr1-R8A*) behaves like a strain that expresses no nucleolin or only nsr1ΔRGG (*Y33 nsr1Δ + nsr1ΔRGG*) as it forms white colonies, in contrast to the control strain that expresses Nsr1 (*Y33 nsr1Δ + NSR1wt*) which forms pink colonies due to the inhibitory effect of 43GAr on translation of its own mRNA. This is consistent with the possibility that the role of type I PRMT in GAr-based inhibition of translation involves methylation of the arginines of the C-terminal RGG motif of nucleolin, which may in turn regulate the ability of this motif to directly interact with the G4 of EBNA1 mRNA.

We then performed the same type of experiment with the human nucleolin NCL. Using site-directed mutagenesis, we replaced the 10 arginines of the C-terminal RGG motif of HA-tagged NCL (HA-NCL) by either alanine (HA-NCL-R10A), lysine (HA-NCL-R10K) to prevent methylation while maintaining the positive charge, or phenylalanine

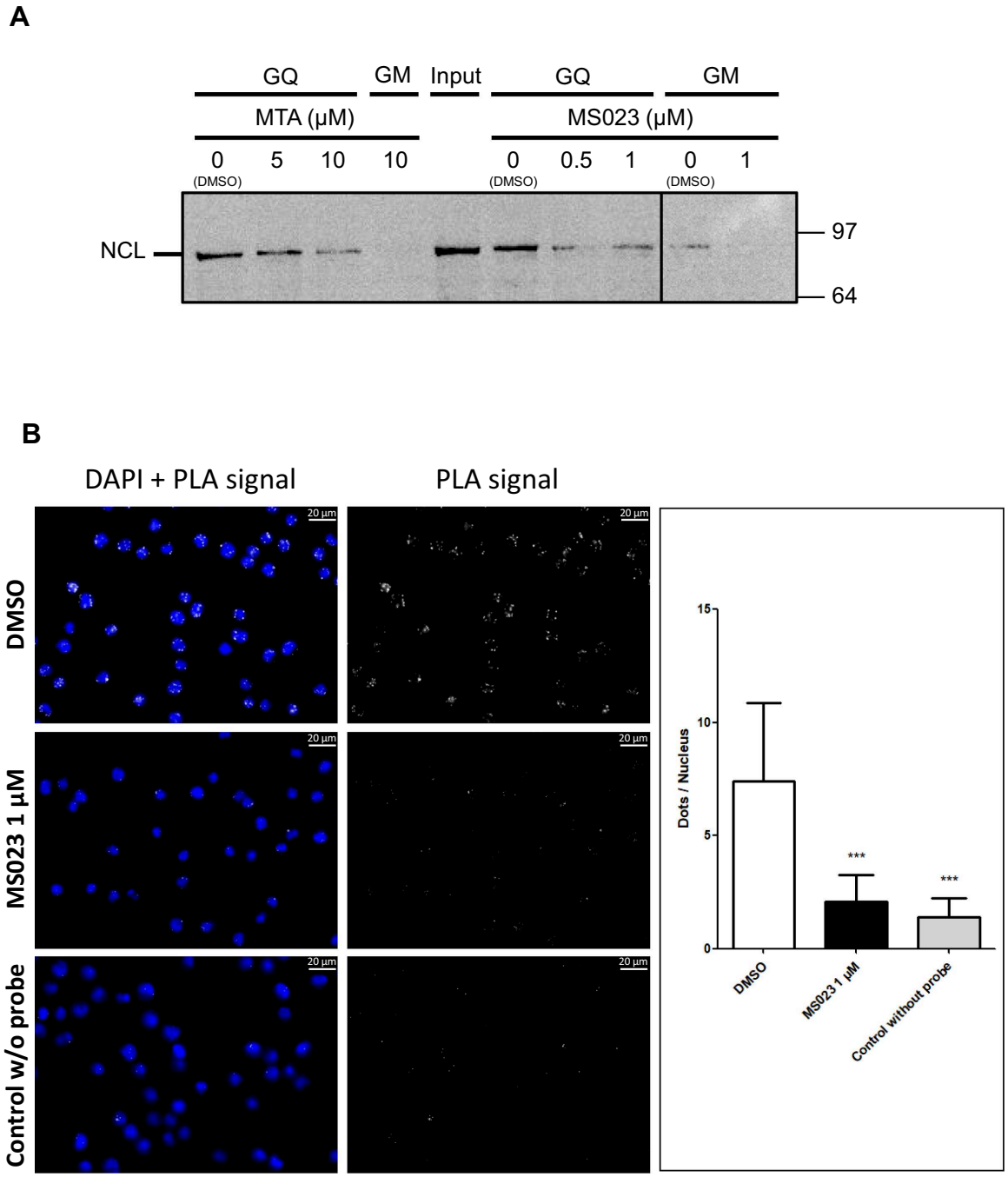


Figure 5. Inhibition of human type I arginine methyltransferases impacts the ability of NCL to interact with G4 of EBNA1 mRNA. (**A**) The same RNA pulldown experiments as in Figure 2 were performed except that extracts from human H1299 cells treated, or not (DMSO control vehicle), with various concentrations of the two PRMT inhibitors MTA or MS023 as indicated were used. Either MTA or MS023 interferes with binding of NCL to G4 of EBNA1 mRNA. (**B**) Adaptation of the PLA to monitor protein–RNA interaction performed in Mutu-1 cells natively expressing EBNA1. Left and middle panels: microscopy images of cells treated with DMSO (compound vehicle, control) or MS023 (1 μM) or negative control cells without probe, as indicated. Nuclei were revealed by DAPI staining and appear in blue; white dots (PLA signals) indicate interaction between NCL and G4 of EBNA1 mRNA. Right panel: number of nuclear PLA signals (dots) per cell in Mutu-1 cells treated with DMSO (control) or with MS023 (1 μM) or in cells of the negative control (without probe). Data from two biological replicates, 200 cells per sample were analysed by ANOVA in conjunction with Tukey's test using GraphPad Prism 5 for Windows (GraphPad Software) (***P <0.0001).

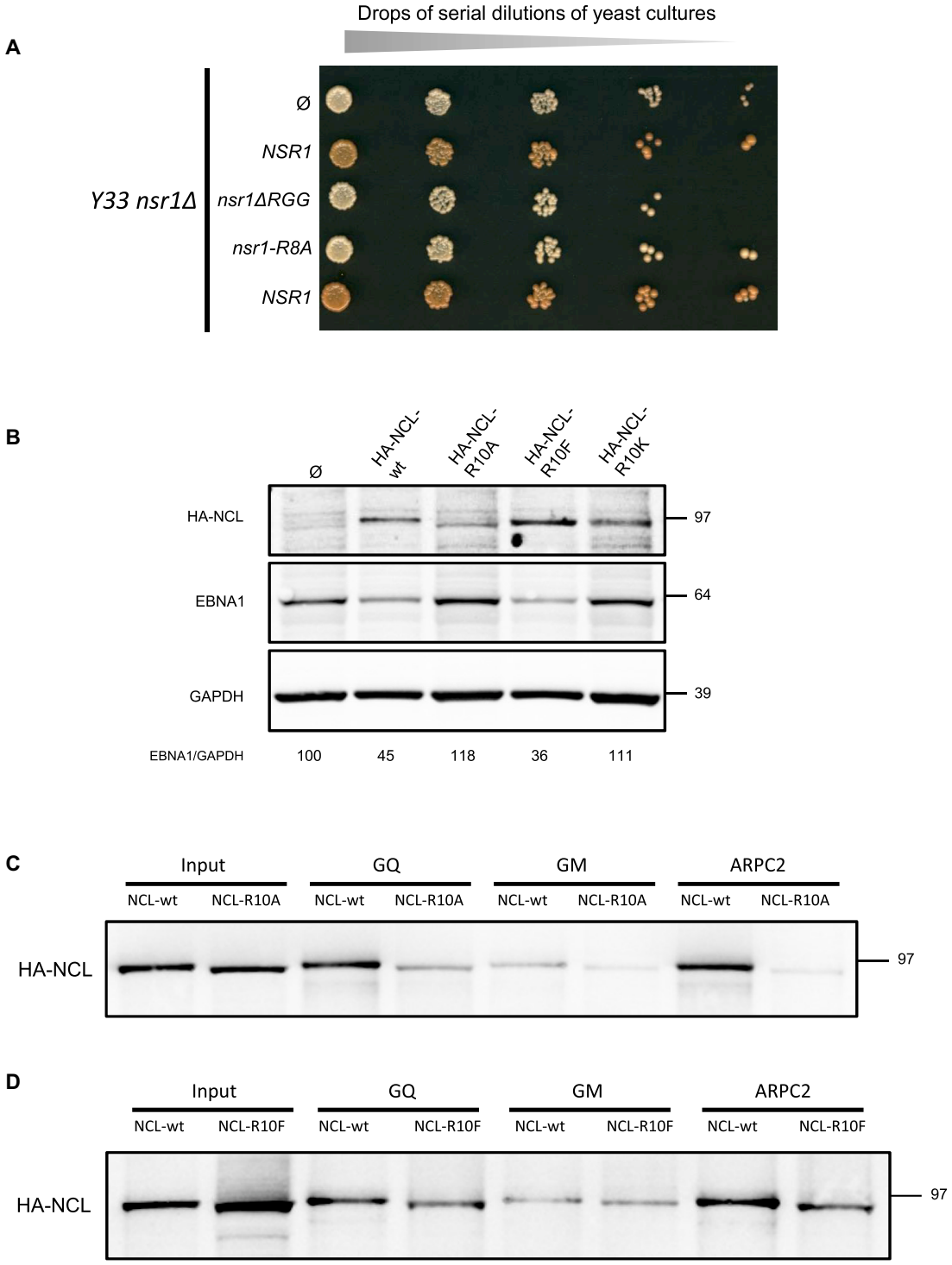


Figure 6. Changing arginines of the RGG motif of both Nsr1 and NCL to alanines suppresses GAr-based inhibition of protein expression and interaction of NCL with G4 of EBNA1 mRNA, whereas changing arginines of NCL to the bulky hydrophobic phenylalanine mimicking methylated arginine maintains both of these phenomena. (**A**) The same experiments as that shown in Figures 1C and 2D were performed and indicate that the eight arginines of the RGG motif of yeast nucleolin Nsr1 are necessary for its function in GAr-based inhibition of protein expression. (**B**) SDS–PAGE and western blot analysis of the level of endogenous EBNA1 in the EBV-infected Raji cell line overexpressing wild-type or mutated versions of HA-tagged nucleolin (HA-NCL-wt, HA-NCL-R10A, HA-NCL-R10F or HA-NCL-R10K as indicated), or not (empty vector). GAPDH was used as a loading control. EBNA1/GAPDH ratios are indicated below the gels. (**C**) The same experiment as in Figure 2, except that RNA pulldowns were performed using extracts from H1299 cells expressing HA-tagged wild-type nucleolin (HA-NCL-wt) or HA-NCL-R10A as indicated. In contrast to HA-NCL-wt which binds to G4 of both EBNA1 and ARPC2 in a G4-dependent manner, only residual binding is observed for HA-NCL-R10A. The blot represents $n \geq 3$. (**D**) The same experiment as in (C), except that RNA pulldowns were performed using extracts from H1299 cells expressing HA-tagged wild-type nucleolin (HA-NCL-wt) or HA-NCL-R10F as indicated. Similarly to HA-NCL-wt, HA-NCL-R10F binds to G4 of both EBNA1 and ARPC2 in a G4-dependent manner, albeit less efficiently.

(HA-NCL-R10F) as a mimic of methylated arginine. Indeed, it is known that methylation of arginine alters charge distribution and increases the volume and π-surface area of the guanidinium moiety, making this residue more hydrophobic and more prone to π-stacking (54,55). Hence it has been shown that phenylalanine may mimic methylated arginine as phenylalanine carries a bulky hydrophobic aromatic moiety (56,57). We overexpressed these various forms of NCL in Raji cells, which are type III latency EBV-infected Burkitt's lymphoma cells, to assess their effect on the endogenous EBNA1 level. Indeed, overexpression of HA-NCL-wt has been shown to lead to a significant decrease in the EBNA1 endogenous level in this cell line, consistent with its role in GAr-based inhibition of translation (21). As shown in Figure 6B, whereas overexpression of wild-type NCL (HA-NCL-wt, second lane) led to a significant decrease in EBNA1 level as compared with GAPDH and with the control with empty plasmid (first lane) as previously observed (21), both HA-NCL-R10A (third lane) and HA-NCL-R10K (fifth lane) had no effect despite being expressed at a similar level as NCL-wt. In contrast, HA-NCL-R10F (fourth lane) led to a decrease in EBNA1 level which is similar to that induced by overexpression of wild-type NCL, suggesting that it is also able to decrease EBNA1 expression. Next we tested the ability of HA-NCL-R10F and HA-NCL-R10A to interact with G4 of EBNA1 mRNA, as compared with HA-NCL-wt. For this, we performed RNA pulldown experiments similar to those described in Figure 2C. In line with the effect of the mutations on EBNA1 expression, we observed that HA-NCL-R10A does not bind to G4 of EBNA1 mRNA (Figure 6C), whereas HA-NCL-R10F does, albeit less efficiently than HA-NCL-wt (Figure 6D). We have also tested the functionality of all the mutants of NCL and Nsr1 in *Y33 nsr1Δ* and found that NCL-R10F and, to a lesser extent, *nsr1-R8F* are able to partially complement the GAr-based inhibition of protein expression in yeast (Supplementary Figure S3). In addition, we determined the cellular localization of HA-NCL-wt, HA-NCL-R10A, HA-NCL-R10F and, as a control, HA-NCLΔRGG, and found that all these forms of NCL are mainly localized in the nucleus (Supplementary Figure S4). All these results are consistent with the possibility that the role of type I PRMT in GAr-based inhibition of translation involves methylation of the arginines of the C-terminal RGG motif of nucleolin.

Inhibition of human type I protein arginine methyltransferases impacts GAr-based inhibition of antigen presentation

Our results indicate that type I PRMTs are critically involved in the interaction between NCL and G4 of the GAr-encoding sequence of EBNA1 mRNA. As the GAr-based limitation of both translation and antigen presentation critically depends on this interaction (21), this suggests that human type I PRMTs are therapeutic targets to unveil the oncogenic EBV to the immune system. To test this hypothesis, we determined if inhibition of type I PRMTs by MS023 has an effect on GAr-restricted major histocompatibility complex (MHC) class I antigen presentation. Indeed, as MS023 led to a GAr-dependent increase in protein expression, it was also expected to stimulate GAr-restricted antigen presentation. For this purpose, we determined the effect of MS023 treatment on the GAr-restricted presentation of the ovalbumin-derived antigenic peptide SIINFEKL ($OVA_{257–264}$) complexed to the murine Kb MHC class I receptor using a previously described T-cell assay (21,58). Once complexed to the murine Kb class I receptor, SIINFEKL is specifically recognized by T cells from the B3Z hybridoma, leading to their activation and therefore to production of IL2. Hence, we measured, by ELISA, IL2 produced by B3Z T-hybridoma cells cultured with H1299 cells co-transfected with both Kb expression vector and 235GAr-OVA or, as a control, OVA-expressing plasmids. As previously reported (22), when mixed with B3Z T cells, 235GAr-OVA-expressing H1299 cells led to a much weaker IL2 production (∼22 pg/ml), as compared with OVA-expressing H1299 cells (∼75 pg/ml) (Figure 7A). This is due to the GAr-based limitation of translation and antigen presentation. Strikingly, treatment of cells transiently transfected with 235GAr-OVA with increasing concentrations of MS023 significantly increased the production of IL2 in a dose-dependent manner, whereas no significant effect was observed for OVA-expressing control cells (Figure 7A). These observations are in line with MS023's effect on GAr-restricted protein expression as MS023 led to a significant increase of GAr-OVA expression, whereas it had no significant effect on OVA expression (Figure 7B). This shows that inhibition of type I PRMTs increases GAr-restricted antigen presentation by MHC-I, suggesting that human type I PRMTs are potential therapeutic targets to unveil the oncogenic EBV to the immune system.

DISCUSSION

In this paper we first show that the C-terminal RGG motif of NCL is crucial for its interaction with G4 of the GAr-encoding sequence of EBNA1 mRNA. This is fully consistent with recent biophysical studies that assess a role for the RGG motif of NCL in the binding to DNA G4 (32,59). This is also in line with the recently described dependence on the C-terminal RGG of Nsr1 for its role in G4-associated genome instability as well as its ability to bind DNA-G4 (45). In line with this, the N-terminal RGG motif of AVEN has been shown to bind RNA-G4 within the coding region of *MLL1* and *MLL4* mRNA, thereby increasing their polysomal association and translation (60). Interestingly, the role of AVEN was also described to be dependent on arginine methylation (60). Of note, this does not mean that one and/or the other of the various RRM domains of NCL are not important for (or do not participate in) the binding of NCL on RNA-G4.

Next, given that RGG motifs are among the main substrates of type I PRMTs, particularly in yeast, this prompted us to determine if these enzymes may modulate the ability of NCL to interact with G4 of EBNA1 mRNA. Hence we observed that the inhibition of type I PRMTs (by specific inhibitors or by siRNA) in human cells does affect: (i) the ability of NCL to interact with G4 of EBNA mRNA and (ii) the GAr-based inhibition of protein expression. This is also true in yeast, as we observed that the deletion of the *HMT1* gene, which encodes the main yeast type I PRMT1 that accounts for the vast majority of arginine methylation (41)

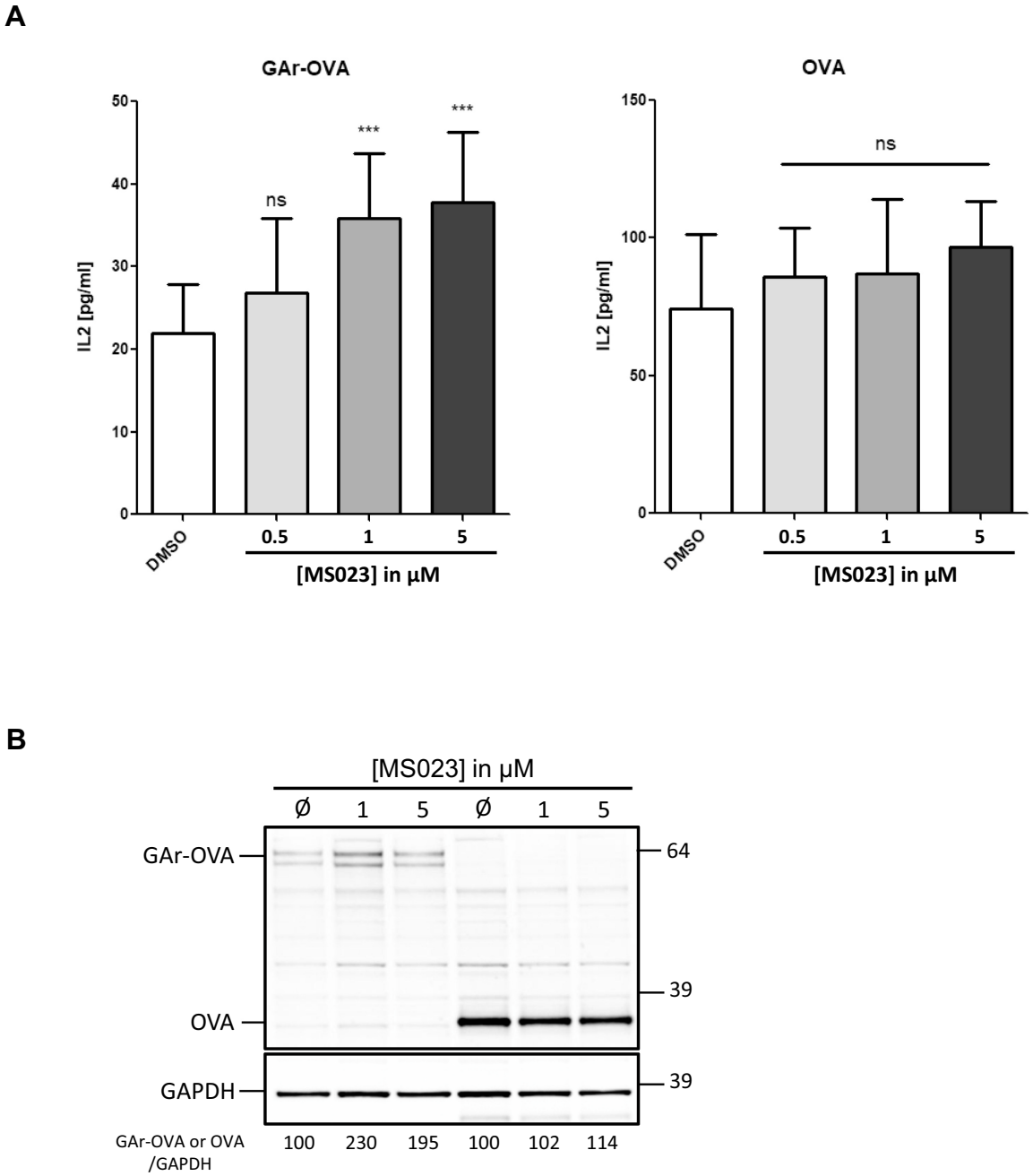


Figure 7. Inhibition of human type I arginine methyltransferases impacts GAr-restricted antigen presentation. (**A**) IL2 concentration (pg/mL) in the supernatant of H1299 cells expressing GAr-OVA (left) or OVA (right) and co-cultured with naive $OVA_{257–264}$ (SIINFEKL peptide)-specific $CD8^+$ cells, were determined following their treatment with DMSO (control), or MS023 at various concentrations as indicated. Data are from three biological replicates (***P <0.0001; ns, not significant). (**B**) SDS–PAGE and western blot analysis of cells used in (A).

and whose main substrates are RGG motifs (40), interferes with the ability of the yeast nucleolin Nsr1 to interact with G4 of EBNA1 mRNA and also with the GAr-based inhibition of protein expression. This led us to examine if inhibition of type I PRMT1 could interfere with GAr-based limitation of antigen presentation, a mechanism at the root of immune evasion of EBNA1, and thus of EBV. Importantly, we found that inhibition of type I PRMTs by MS023, which was recently described as a specific inhibitor of this family of PRMTs (52), suppresses the GAr-based limitation of antigen presentation in a dose-dependent manner. Hence we have identified type I PRMTs as new druggable therapeutic targets to unveil EBNA1 to the immune system. Of note, type I PRMTs may appear as relatively non-specific therapeutic targets as their inhibition could in principle disturb methylation of many proteins, as has been initially considered for drugs inhibiting protein kinases or phosphatases. However, since RGG motifs are the main substrates of these enzymes, and as, in our various assays, type I PRMT inhibitors did not exhibit significant toxicity at concentration

ranges in which they significantly affect both GAr-based translation inhibition and the NCL/G4 of EBNA1 mRNA interaction (Supplementary Figure S5), the lack of specificity may not be an issue.

The precise mechanism by which type I PRMTs control EBNA1 immune evasion remains unclear. However, given that RGG motifs represent a major substrate for type I PRMTs, and since we show here that the RGG motif of NCL is crucial for its interaction with G4 of the GAr-encoding sequence of the EBNA1 mRNA [an interaction which is direct (21)], it is tempting to speculate that methylation of arginines in the C-terminal RGG motif of NCL by type I PRMTs favours this interaction. Fully consistent with this possibility is our observation that replacing the eight arginines of the RGG motif of yeast nucleolin Nsr1 by alanines abolishes its ability to participate in GAr-based inhibition of translation. We obtained the same result when replacing the 10 arginines of the C-terminal RGG motif of NCL by alanines or by lysines. Conversely, the replacement of the 10 arginines of the C-terminal RGG motif of NCL by the bulky hydrophobic constitutive methylated arginine-mimetic phenylalanine restored GAr-based inhibition of protein expression, suggesting that π-stacking interactions of the arginines of the RGG motif are likely to be involved in the ability of NCL to interact with G4 of EBNA1 mRNA and that these interactions might be reinforced by arginine methylation. Indeed, arginines are known to establish preferential interactions with guanine rings through complex interactions involving π-cation, π-π stacking and H-bonding (61). Hence it makes sense that arginine–guanine pairs are determinant for binding to the external guanine quartets of G4 structures which are recognition elements for G4-binding proteins (32,59). In line with this, we observed that inhibition of type I PRMTs (by small molecular weight inhibitors or by siRNA) interferes with the ability of both yeast and human nucleolin to bind to G4 of EBNA1 mRNA, as shown by RNA pulldown experiments and PLAs. Postulating that methylation of the arginines of the RGG motif of nucleolin is crucial for the role of nucleolin in GAr-based inhibition of translation, the question is now to determine if the role of RGG methylation is direct or indirect. In other words and as shown in Figure 8, is the methylation of various arginines directly involved by physically promoting the interaction between the C-terminal RGG motif of NCL and G4 of EBNA1 mRNA, or does it plays its role by preventing the interaction of NCL with other partner(s), thereby releasing NCL which is then free to interact with G4 of EBNA1 mRNA? In favour of the first possibility ('direct' role of RGG methylation) data are available that show that RGG methylation may interfere with its ability to interact with some of its partners (that can be proteins, DNA or RNA) (34). In favour of the second possibility ('indirect' role of RGG methylation) are the recent observations that Scd6, a yeast protein involved in translation inhibition, is able to self-associate via its RGG motif and that this self-association prevents its translation repression activity and is negatively regulated by Hmt1-dependent methylation (33). To test these two possibilities ('direct' or 'indirect' role of RGG methylation), we have performed two types of *in vitro* experiments. We first performed *in vitro* arginine methylation of bacterially produced recombinant NCL (hence being initially non-methylated) and then assessed its ability to bind G4 of EBNA1 mRNA, as compared with non-methylated recombinant NCL. The result of this new experiment is shown in Supplementary Figure S6. We found no significant difference between methylated and non-methylated NCL for binding to the EBNA1 mRNA sequence forming one G4 (18-mer) and a slightly stronger binding of methylated NCL when a two-repeat EBNA1 mRNA sequence forming two G4 (36-mer) was used as a binding partner. This suggests that the effect of arginine methylation on the ability of the C-terminal RGG motif of NCL to bind G4 of EBNA1 mRNA could be indirect, by interfering with the interaction of the RGG motif with unknown partner(s) whose binding may prevent the RGG motif from interacting with G4 of EBNA1 mRNA. The factor interacting with the RGG motif of NCL could be a ribosomal protein as NCL has been reported to interact with several ribosomal proteins through its RGG motif (62). Alternatively, the binding partner could be NCL itself as it has been reported to self-associate (63). However, in this case, this self-interaction rather involves NCL's central RRM domains, and is therefore less likely to prevent its C-terminal RGG motif from interacting with G4 of EBNA1 mRNA. In line with this, using two-hybrid, we were not able to observe self-interaction for either yeast (Nsr1) or human (NCL) nucleolin, even in a *hmt1Δ* strain (G.A. and M.B., unpublished observations).

Of note, one caveat with *in vitro* arginine methylation of NCL is that the extent of methylation is difficult to control. Hence, we obtained a heterogeneous mixture of NCLs harbouring various degrees of methylation (on average eight arginines per protein featuring mono-methylation, or symmetrical or asymmetrical di-methylations, Supplementary Figure S7) which may not reflect the biological situation. To overcome this limitation, we have also performed another *in vitro* experiment based on two NCL RGG peptides—one being asymmetrically di-methylated on the 10 arginines, while the other was non-methylated. We used these two peptides to determine if arginine methylation may have a direct effect on G4 binding. We found that both peptides display a similarly high capacity to bind G4 of EBNA1 mRNA (k_D in the nM range with a slight difference in favour of the methylated peptide, Supplementary Figure S8), which is fully in line with the results of the experiment using the recombinant protein. Taken together, these *in vitro* data suggest that the effect of arginine methylation on G4 binding is most probably indirect as methylation only displays a marginal effect on the *in vitro* direct interaction between G4 and the RGG motif of NCL. Importantly, the identification of the putative binding partner of the NCL RGG motif that sequesters it, thereby preventing its interaction with G4 of EBNA1 mRNA, is an important aspect that has not been solved in the present study. Given the central role of NCL in RNA biology, the identification of this factor should be the subject of future studies. This will be an important step as, beyond its involvement in EBNA1 immune evasion, this mechanism may represent an original and general way to regulate the ability of NCL to interact with RNA.

Apart from NCL, which is the first host factor crucially involved in GAr-based inhibition of translation and of antigenic peptide production, an important question is to deter-

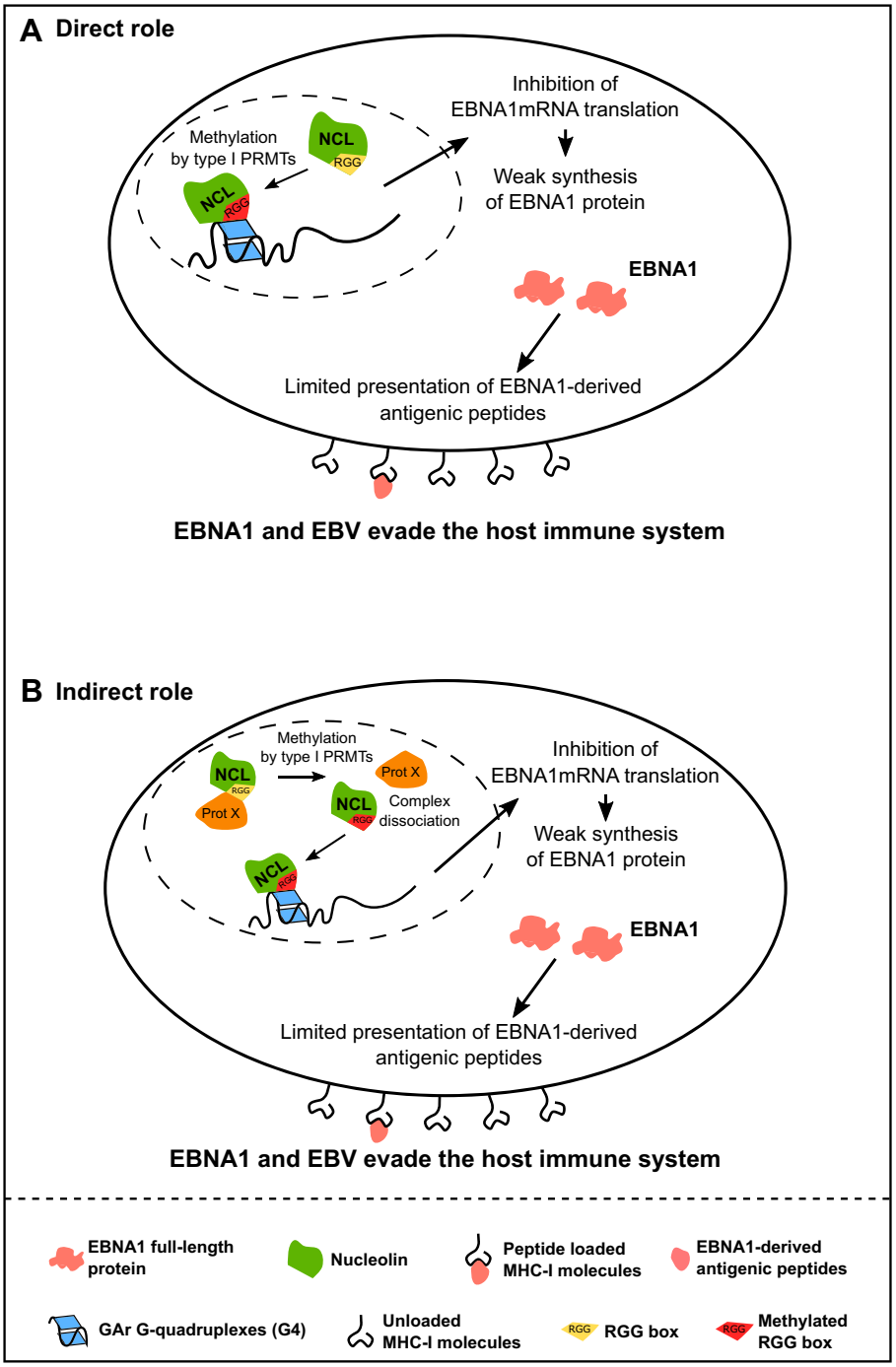


Figure 8. Model. The interaction between NCL and G4 of EBNA1 mRNA is direct and the C-terminal RGG motif of NCL is required for both the ability of NCL to interact with G4 that form in the GAr-encoding sequence of EBNA1 mRNA and for NCL's crucial role in GAr-based inhibition of translation, the molecular mechanism at the root of EBNA1 immune evasion. RGG motifs are the main substrates of type I PRMTs. The interaction between NCL and G4 of EBNA1 mRNA, as well as the GAr-based inhibition of translation and EBNA1 immune evasion, depend on type I PRMTs, in particular PRMT1. Results of the site-directed mutagenesis of the arginines of the RGG motif of NCL experiments suggest that the role of type I PRMTs in EBNA1 immune evasion involves methylation of the arginines of the RGG motif of NCL. Hence, two possible models for the role RGG methylation by type I PRMTs in EBNA1 evasion can be envisaged: direct (**A**) or indirect (**B**). In the direct model (A), the methylation of arginines of the RGG motif of NCL would directly favour the interaction of the latter with G4 of EBNA1 mRNA and therefore the inhibition of its translation, which in turn allows EBNA1 and EBV to evade the immune system. In the indirect model (B), the non-methylated RGG motif of NCL would be sequestered by a protein partner [protein X, which may be nucleolin itself or another protein such as the various ribosomal proteins which have been shown to interact with the RGG motif of NCL (62)], thereby preventing its interaction with G4 of EBNA1 mRNA. In this model, methylation of the arginines of the RGG motif would interfere with its interaction with protein X, thus releasing NCL and allowing its RGG motif to interact with G4 of EBNA1 mRNA. Our *in vitro* assays did not show any significant difference between methylated and unmethylated RGG for the binding of EBNA1 mRNA G4, suggesting the indirect role is the most likely.

mine if other proteins are involved in this process, through binding to G4 of EBNA1 mRNA like NCL. Interestingly, EBNA1 itself contains two RGG motifs: one upstream of GAr and one directly downstream, and these two RGG motifs have been shown to bind RNA-G4 (64), leaving the possibility that EBNA1 may contribute to the limitation of its own expression by binding to G4 of its own mRNA. Importantly, this putative role of EBNA1 could also be controlled by type I PRMTs.

Finally, the deimination of arginine into citrulline, a process called citrullination and which is catalysed by peptidylarginine deiminases, prevents their methylation. Interestingly, it has been reported that the infection by EBV activates several PRMTs, including PRMT1, and inactivates PADI4, the main peptidylarginine deiminase (65). This is fully in line with the results reported here and it is tempting to speculate that the infection by EBV, by activating PRMT1 and inactivating PADI4, favours the methylation of the RGG motif of NCL (and possibly of other RGG motif-containing proteins such as EBNA1), thereby increasing their ability to bind to G4 of EBNA1 mRNA, which in turn leads to inhibition of EBNA1 mRNA translation, thus limiting the production of EBNA1-derived antigenic peptides, ultimately leading to immune evasion of EBV.

The precise and comprehensive role of type I PRMTs in GAr-based inhibition of translation, a molecular mechanism at the root of EBNA1 immune evasion and potentially another role for NCL in RNA biology, clearly deserves deeper investigations and will be the subject of future studies. Importantly, whatever the precise mechanism (see Figure 8), our present work unambiguously defines type I PRMTs as therapeutic targets to unveil EBV-infected cells to the immune system, thereby providing new potential therapeutic avenues to unveil EBV-related cancers to the immune system.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank all group members for valuable discussion, and Drs A. Granzhan and S. Bombard (Institut Curie, CNRS UMR9187-Inserm U1196, Orsay, France) for critical proofreading of the manuscript and for their help in the arginine methylation experiments. This work has also benefited from the facilities and expertise of the SICaPS Proteomic platform of Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Univ. Paris-Sud, Université Paris-Saclay, 91198, Gif-sur-Yvette Cedex, France.

FUNDING

This work was supported by the following grant agencies: 'La Ligue contre le cancer' to R.F., 'La Ligue contre le cancer CSIRGO' to M.B., 'Institut National du cancer (INCa)' [2016–169 to M.B., R.F. and M.P.T.F.] and 'Fondation pour la Recherche Médicale' [DCM20181039571 to M.P.T.F., R.F., M.B. and R.C. (postdoctoral fellowship)], European Regional Development Fund - Project ENOCH [No. CZ.02.1.01/0.0/0.0/16_019/0000868] and by MH CZ - DRO [MMCI, 00209805]. G.A. was the recipient of a joint PhD grant from 'La Ligue contre le cancer Bretagne' and Région Bretagne and of a fourth year from 'Fondation pour la Recherche Médicale' [FRM FDT201904008299], R.L.S. is the recipient of a PhD grant from the French government and V.T.D. of a joint PhD grant from 'La Ligue contre le cancer Bretagne' and Région Bretagne. Funding for open access charge: 'Institut National de Recherche Médicale' (Inserm).

Conflict of interest statement. None declared.

REFERENCES

1. Ozoya,O.O., Sokol,L. and Dalia,S. (2016) EBV-related malignancies, outcomes and novel prevention strategies. *Infect. Disord. Drug Targets*, **16**, 4–21.
2. Thorley-Lawson,D.A. and Allday,M.J. (2008) The curious case of the tumour virus: 50 years of Burkitt's lymphoma. *Nat. Rev. Microbiol.*, **6**, 913–924.
3. Young,L.S. and Rickinson,A.B. (2004) Epstein–Barr virus: 40 years on. *Nat. Rev. Cancer*, **4**, 757–768.
4. Cohen,J.I., Fauci,A.S., Varmus,H. and Nabel,G.J. (2011) Epstein–Barr virus: an important vaccine target for cancer prevention. *Sci. Transl. Med.*, **3**, 107fs107.
5. Akhtar,S., Vranic,S., Cyprian,F.S. and Al Moustafa,A.E. (2018) Epstein–Barr virus in gliomas: cause, association, or artifact?*Front. Oncol.*, **8**, 123.
6. Daskalogianni,C., Pyndiah,S., Apcher,S., Mazars,A., Manoury,B., Ammari,N., Nylander,K., Voisset,C., Blondel,M. and Fahraeus,R. (2015) Epstein–Barr virus-encoded EBNA1 and ZEBRA: targets for therapeutic strategies against EBV-carrying cancers. *J. Pathol.*, **235**, 334–341.
7. Wilson,J.B., Manet,E., Gruffat,H., Busson,P., Blondel,M. and Fahraeus,R. (2018) EBNA1: oncogenic activity, immune evasion and biochemical functions provide targets for novel therapeutic strategies against Epstein–Barr virus-associated cancers. *Cancers (Basel)*, **10**, 109.
8. Blake,N., Lee,S., Redchenko,I., Thomas,W., Steven,N., Leese,A., Steigerwald-Mullen,P., Kurilla,M.G., Frappier,L. and Rickinson,A. (1997) Human CD8+ T cell responses to EBV EBNA1: HLA class I presentation of the (Gly–Ala)-containing protein requires exogenous processing. *Immunity*, **7**, 791–802.
9. Lee,S.P., Brooks,J.M., Al-Jarrah,H., Thomas,W.A., Haigh,T.A., Taylor,G.S., Humme,S., Schepers,A., Hammerschmidt,W., Yates,J.L. *et al.* (2004) CD8 T cell recognition of endogenously expressed Epstein–Barr virus nuclear antigen 1. *J. Exp. Med.*, **199**, 1409–1420.
10. Voo,K.S., Fu,T., Wang,H.Y., Tellam,J., Heslop,H.E., Brenner,M.K., Rooney,C.M. and Wang,R.F. (2004) Evidence for the presentation of major histocompatibility complex class I-restricted Epstein–Barr virus nuclear antigen 1 peptides to CD8+ T lymphocytes. *J. Exp. Med.*, **199**, 459–470.
11. Tellam,J., Connolly,G., Green,K.J., Miles,J.J., Moss,D.J., Burrows,S.R. and Khanna,R. (2004) Endogenous presentation of CD8+ T cell epitopes from Epstein–Barr virus-encoded nuclear antigen 1. *J. Exp. Med.*, **199**, 1421–1431.
12. Yin,Y., Manoury,B. and Fahraeus,R. (2003) Self-inhibition of synthesis and antigen presentation by Epstein–Barr virus-encoded EBNA1. *Science*, **301**, 1371–1374.
13. Murat,P., Zhong,J., Lekieffre,L., Cowieson,N.P., Clancy,J.L., Preiss,T., Balasubramanian,S., Khanna,R. and Tellam,J. (2014) G-quadruplexes regulate Epstein–Barr virus-encoded nuclear antigen 1 mRNA translation. *Nat. Chem. Biol.*, **10**, 358–364.
14. Beaudoin,J.D. and Perreault,J.P. (2010) 5′-UTR G-quadruplex structures acting as translational repressors. *Nucleic Acids Res.*, **38**, 7022–7036.
15. Didiot,M.C., Tian,Z., Schaeffer,C., Subramanian,M., Mandel,J.L. and 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.

16. Gomez,D., Lemarteleur,T., Lacroix,L., Mailliet,P., Mergny,J.L. and Riou,J.F. (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.
17. Marcel,V., Tran,P.L., Sagne,C., Martel-Planche,G., Vaslin,L., Teulade-Fichou,M.P., Hall,J., Mergny,J.L., Hainaut,P. and Van Dyck,E. (2011) G-quadruplex structures in TP53 intron 3: role in alternative splicing and in production of p53 mRNA isoforms. *Carcinogenesis*, **32**, 271–278.
18. Siddiqui-Jain,A., Grand,C.L., Bearss,D.J. and Hurley,L.H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl Acad. Sci. USA*, **99**, 11593–11598.
19. Song,J., Perreault,J.P., Topisirovic,I. and Richard,S. (2016) RNA G-quadruplexes and their potential regulatory roles in translation. *Translation (Austin)*, **4**, e1244031.
20. Lista,M.J., Martins,R.P., Angrand,G., Quillevere,A., Daskalogianni,C., Voisset,C., Teulade-Fichou,M.P., Fahraeus,R. and Blondel,M. (2017) A yeast model for the mechanism of the Epstein–Barr virus immune evasion identifies a new therapeutic target to interfere with the virus stealthiness. *Microb. Cell*, **4**, 305–307.
21. Lista,M.J., Martins,R.P., Billant,O., Contesse,M.A., Findakly,S., Pochard,P., Daskalogianni,C., Beauvineau,C., Guetta,C., Jamin,C. *et al.* (2017) Nucleolin directly mediates Epstein–Barr virus immune evasion through binding to G-quadruplexes of EBNA1 mRNA. *Nat. Commun.*, **8**, 16043.
22. Reznichenko,O., Quillevere,A., Martins,R.P., Loaec,N., Kang,H., Lista,M.J., Beauvineau,C., Gonzalez-Garcia,J., Guillot,R., Voisset,C. *et al.* (2019) Novel cationic bis(acylhydrazones) as modulators of Epstein–Barr virus immune evasion acting through disruption of interaction between nucleolin and G-quadruplexes of EBNA1 mRNA. *Eur. J. Med. Chem.*, **178**, 13–29.
23. Abdelmohsen,K. and Gorospe,M. (2012) RNA-binding protein nucleolin in disease. *RNA Biol*, **9**, 799–808.
24. Ginisty,H., Sicard,H., Roger,B. and Bouvet,P. (1999) Structure and functions of nucleolin. *J. Cell Sci.*, **112**, 761–772.
25. Ugrinova,I., Petrova,M., Chalabi-Dchar,M. and Bouvet,P. (2018) Multifaceted nucleolin protein and its molecular partners in oncogenesis. *Adv. Protein Chem. Struct. Biol.*, **111**, 133–164.
26. Hanakahi,L.A., Sun,H. and Maizels,N. (1999) High affinity interactions of nucleolin with G–G-paired rDNA. *J. Biol. Chem.*, **274**, 15908–15912.
27. Gonzalez,V. and Hurley,L.H. (2010) The C-terminus of nucleolin promotes the formation of the c-MYC G-quadruplex and inhibits c-MYC promoter activity. *Biochemistry*, **49**, 9706–9714.
28. Tosoni,E., Frasson,I., Scalabrin,M., Perrone,R., Butovskaya,E., Nadai,M., Palu,G., Fabris,D. and Richter,S.N. (2015) Nucleolin stabilizes G-quadruplex structures folded by the LTR promoter and silences HIV-1 viral transcription. *Nucleic Acids Res.*, **43**, 8884–8897.
29. Bian,W.X., Xie,Y., Wang,X.N., Xu,G.H., Fu,B.S., Li,S., Long,G., Zhou,X. and Zhang,X.L. (2019) Binding of cellular nucleolin with the viral core RNA G-quadruplex structure suppresses HCV replication. *Nucleic Acids Res.*, **47**, 56–68.
30. Dickerhoff,J., Onel,B., Chen,L., Chen,Y. and Yang,D. (2019) Solution structure of a MYC promoter G-quadruplex with 1:6:1 loop length. *ACS Omega*, **4**, 2533–2539.
31. Lago,S., Tosoni,E., Nadai,M., Palumbo,M. and Richter,S.N. (2017) The cellular protein nucleolin preferentially binds long-looped G-quadruplex nucleic acids. *Biochim. Biophys, Acta*, **1861**, 1371–1381.
32. Saha,A., Duchambon,P., Masson,V., Loew,D., Bombard,S. and Teulade-Fichou,M.P. (2020) Nucleolin discriminates drastically between long-loop and short-loop quadruplexes. *Biochemistry*, **59**, 1261–1272.
33. Poornima,G., Mythili,R., Nag,P., Parbin,S., Verma,P.K., Hussain,T. and Rajyaguru,P.I. (2019) RGG-motif self-association regulates eIF4G-binding translation repressor protein Scd6. *RNA Biol*, **16**, 1215–1227.
34. Thandapani,P., O'Connor,T.R., Bailey,T.L. and Richard,S. (2013) Defining the RGG/RG motif. *Mol. Cell*, **50**, 613–623.
35. Wu,Q., Schapira,M., Arrowsmith,C.H. and Barsyte-Lovejoy,D. (2021) Protein arginine methylation: from enigmatic functions to therapeutic targeting. *Nat. Rev. Drug Discov.*, **20**, 509–530.
36. Bedford,M.T. (2007) Arginine methylation at a glance. *J. Cell Sci.*, **120**, 4243–4246.
37. Blanc,R.S. and Richard,S. (2017) Arginine methylation: the coming of age. *Mol. Cell*, **65**, 8–24.
38. Hamey,J.J., Separovich,R.J. and Wilkins,M.R. (2018) MT-MAMS: protein methyltransferase motif analysis by mass spectrometry. *J. Proteome Res.*, **17**, 3485–3491.
39. Lorton,B.M. and Shechter,D. (2019) Cellular consequences of arginine methylation. *Cell. Mol. Life Sci.*, **76**, 2933–2956.
40. Yagoub,D., Tay,A.P., Chen,Z., Hamey,J.J., Cai,C., Chia,S.Z., Hart-Smith,G. and Wilkins,M.R. (2015) Proteogenomic discovery of a small, novel protein in yeast reveals a strategy for the detection of unannotated short open reading frames. *J. Proteome Res.*, **14**, 5038–5047.
41. Chia,S.Z., Lai,Y.W., Yagoub,D., Lev,S., Hamey,J.J., Pang,C.N.I., Desmarini,D., Chen,Z., Djordjevic,J.T., Erce,M.A. *et al.* (2018) Knockout of the hmt1p arginine methyltransferase in *Saccharomyces cerevisiae* leads to the dysregulation of Phosphate-associated genes and processes. *Mol. Cell. Proteomics*, **17**, 2462–2479.
42. Blondel,M., Bach,S., Bamps,S., Dobbelaere,J., Wiget,P., Longaretti,C., Barral,Y., Meijer,L. and Peter,M. (2005) Degradation of Hof1 by SCF(Grr1) is important for actomyosin contraction during cytokinesis in yeast. *EMBO J.*, **24**, 1440–1452.
43. Voisset,C., Daskalogianni,C., Contesse,M.A., Mazars,A., Arbach,H., Le Cann,M., Soubigou,F., Apcher,S., Fahraeus,R. and Blondel,M. (2014) A yeast-based assay identifies drugs that interfere with immune evasion of the Epstein–Barr virus. *Dis. Model. Mech.*, **7**, 435–444.
44. Lee,W.C., Xue,Z.X. and Melese,T. (1991) The NSR1 gene encodes a protein that specifically binds nuclear localization sequences and has two RNA recognition motifs. *J. Cell Biol.*, **113**, 1–12.
45. Singh,S., Berroyer,A., Kim,M. and Kim,N. (2020) Yeast nucleolin nsr1 impedes replication and elevates genome instability at an actively transcribed guanine-rich G4 DNA-forming sequence. *Genetics*, **216**, 1023–1037.
46. von Hacht,A., Seifert,O., Menger,M., Schutze,T., Arora,A., Konthur,Z., Neubauer,P., Wagner,A., Weise,C. and Kurreck,J. (2014) Identification and characterization of RNA guanine-quadruplex binding proteins. *Nucleic Acids Res.*, **42**, 6630–6644.
47. Cheng,D., Yadav,N., King,R.W., Swanson,M.S., Weinstein,E.J. and Bedford,M.T. (2004) Small molecule regulators of protein arginine methyltransferases. *J. Biol. Chem.*, **279**, 23892–23899.
48. Zhang,B., Dong,S., Zhu,R., Hu,C., Hou,J., Li,Y., Zhao,Q., Shao,X., Bu,Q., Li,H. *et al.* (2015) Targeting protein arginine methyltransferase 5 inhibits colorectal cancer growth by decreasing arginine methylation of eIF4E and FGFR3. *Oncotarget*, **6**, 22799–22811.
49. Borchardt,R.T. (1980) *S*-Adenosyl-L-methionine-dependent macromolecule methyltransferases: potential targets for the design of chemotherapeutic agents. *J. Med. Chem.*, **23**, 347–357.
50. Kryukov,G.V., Wilson,F.H., Ruth,J.R., Paulk,J., Tsherniak,A., Marlow,S.E., Vazquez,F., Weir,B.A., Fitzgerald,M.E., Tanaka,M. *et al.* (2016) MTAP deletion confers enhanced dependency on the PRMT5 arginine methyltransferase in cancer cells. *Science*, **351**, 1214–1218.
51. Mavrakis,K.J., McDonald,E.R., Schlabach,M.R., Billy,E., Hoffman,G.R., deWeck,A., Ruddy,D.A., Venkatesan,K., Yu,J., McAllister,G. *et al.* (2016) Disordered methionine metabolism in MTAP/CDKN2A-deleted cancers leads to dependence on PRMT5. *Science*, **351**, 1208–1213.
52. Eram,M.S., Shen,Y., Szewczyk,M., Wu,H., Senisterra,G., Li,F., Butler,K.V., Kaniskan,H.U., Speed,B.A., Dela Sena,C. *et al.* (2016) A potent, selective, and cell-active inhibitor of human type I protein arginine methyltransferases. *ACS Chem. Biol.*, **11**, 772–781.
53. Prado Martins,R., Findakly,S., Daskalogianni,C., Teulade-Fichou,M.P., Blondel,M. and Fahraeus,R. (2018) In cellulo protein–mRNA interaction assay to determine the action of G-quadruplex-binding molecules. *Molecules*, **23**, 314.
54. Evich,M., Stroeva,E., Zheng,Y.G. and Germann,M.W. (2016) Effect of methylation on the side-chain pKa value of arginine. *Protein Sci.*, **25**, 479–486.
55. Hughes,R.M. and Waters,M.L. (2006) Arginine methylation in a beta-hairpin peptide: implications for Arg–pi interactions, DeltaCp(o), and the cold denatured state. *J. Am. Chem. Soc.*, **128**, 12735–12742.
56. Campbell,M., Chang,P.C., Huerta,S., Izumiya,C., Davis,R., Tepper,C.G., Kim,K.Y., Shevchenko,B., Wang,D.H., Jung,J.U. *et al.* (2012) Protein arginine methyltransferase 1-directed methylation of

Kaposi sarcoma-associated herpesvirus latency-associated nuclear antigen. *J. Biol. Chem.*, **287**, 5806–5818.
57. Mostaqul Huq,M.D., Gupta,P., Tsai,N.P., White,R., Parker,M.G. and Wei,L.N. (2006) Suppression of receptor interacting protein 140 repressive activity by protein arginine methylation. *EMBO J.*, **25**, 5094–5104.
58. Martins,R.P., Malbert-Colas,L., Lista,M.J., Daskalogianni,C., Apcher,S., Pla,M., Findakly,S., Blondel,M. and Fahraeus,R. (2019) Nuclear processing of nascent transcripts determines synthesis of full-length proteins and antigenic peptides. *Nucleic Acids Res.*, **47**, 3086–3100.
59. Masuzawa,T. and Oyoshi,T. (2020) Roles of the RGG domain and RNA recognition motif of nucleolin in G-quadruplex stabilization. *ACS Omega*, **5**, 5202–5208.
60. Thandapani,P., Song,J., Gandin,V., Cai,Y., Rouleau,S.G., Garant,J.M., Boisvert,F.M., Yu,Z., Perreault,J.P., Topisirovic,I. *et al.* (2015) Aven recognition of RNA G-quadruplexes regulates translation of the mixed lineage leukemia protooncogenes. *Elife*, **4**, e06234.
61. Chavali,S.S., Cavender,C.E., Mathews,D.H. and Wedekind,J.E. (2020) Arginine forks are a widespread motif to recognize phosphate backbones and guanine nucleobases in the RNA major groove. *J. Am. Chem. Soc.*, **142**, 19835–19839.
62. Bouvet,P., Diaz,J.J., Kindbeiter,K., Madjar,J.J. and Amalric,F. (1998) Nucleolin interacts with several ribosomal proteins through its RGG domain. *J. Biol. Chem.*, **273**, 19025–19029.
63. Hanakahi,L.A., Bu,Z. and Maizels,N. (2000) The C-terminal domain of nucleolin accelerates nucleic acid annealing. *Biochemistry*, **39**, 15493–15499.
64. Norseen,J., Johnson,F.B. and Lieberman,P.M. (2009) Role for G-quadruplex RNA binding by Epstein–Barr virus nuclear antigen 1 in DNA replication and metaphase chromosome attachment. *J. Virol.*, **83**, 10336–10346.
65. Leonard,S., Gordon,N., Smith,N., Rowe,M., Murray,P.G. and Woodman,C.B. (2012) Arginine methyltransferases are regulated by Epstein–Barr virus in B cells and are differentially expressed in Hodgkin's lymphoma. *Pathogens*, **1**, 52–64.

Le complexe NAC est nécessaire pour le recrutement de la NCL sur l'ARNm d'EBNA1 et pour l'inhibition GAr-dépendante de la traduction

Bien qu'il ait été montré que l'inhibition de la traduction de l'ARNm d'EBNA1 est dépendante de la séquence codant GAr , une étude a montré que des anticorps dirigés contre GAr sont aussi capables d'inactiver son effet inhibiteur de la traduction (Yin et al., 2003). C'est pourquoi, dans le second article présenté, nous avons cherché à savoir si le peptide contenant le domaine GAr est aussi un élément de régulation de la traduction de son ARNm. Notre hypothèse est que ce mécanisme se produirait au moment de la traduction, avant le départ de la protéine du ribosome. Généralement, une fois la traduction terminée, le peptide nouvellement synthétisé rencontre le complexe NAC situé à la fin du ribosome. Ce complexe est composé des protéines NACA et BTF3. NACA est la première protéine chaperonne que rencontre le peptide naissant qui est ensuite exporté et protégé de la protéolyse par BTF3.

Pour tester l'effet du complexe NAC sur la traduction, nous avons inhibé l'expression de NACA avec des siRNA. Ces expériences ont montré qu'inhiber l'expression de NACA augmente l'expression d'EBNA1 sans changer son niveau d'ARNm. De plus, par des tests de présentation antigénique, nous avons aussi montré que l'inhibition de NACA augmente la présentation antigénique des protéines possédant un domaine GAr. En outre, par des expériences de PLA, nous avons montré que la protéine NACA et la protéine EBNA1 interagissent dans le noyau des cellules. En revanche, lorsque le domaine GAr est délété, l'interaction entre les deux protéines est significativement diminuée.

Par une adaptation du protocole de PLA couplé avec un ELISA, nous avons ensuite montré que le peptide contenant le domaine GAr, nouvellement synthétisé interagissait avec NACA. Cependant, lorsque le domaine GAr est absent cette interaction est fortement diminuée. Grâce à l'ajout d'un site de clivage qui sépare le domaine GAr du reste de la protéine, nous avons montré que le peptide GAr libère NACA du ribosome. Ces résultats montrent donc que les peptides contenant GAr nouvellement synthétisés interagissent avec le complexe NAC afin de déplacer la protéine NACA. Cette protéine, désormais libérée de son complexe NAC viendrait alors participer à l'inhibition de la traduction d'EBNA1.

Pour tester cette hypothèse et rechercher un lien entre NAC et le recrutement de la NCL à l'ARNm d'EBNA1, nous avons réalisé des co-immunoprécipitations entre la NCL et des ARN contenant le domaine GAr. En absence d'expression de NACA, la capacité de la NCL à se lier avec des ARN contenant le domaine GAr est fortement diminuée, cependant lorsque ces ARN ne possèdent pas de domaine GAr, l'absence de NACA n'a pas d'effet sur la liaison de la NCL avec ces ARN.

Ma contribution à cet article a été de déterminer si l'interaction entre l'ARNm d'EBNA1 et NACA était directe par des tests de *RNA pulldown*. Nous avons ainsi montré que NACA lie de manière G4 dépendante aux rG4 de l'ARNm d'EBNA1. De plus, l'utilisation du PhenDH2, un ligand de G4 a montré que celui-ci compétait NACA pour la liaison aux rG4 et donc à l'ARNm d'EBNA1. NACA en interagissant avec les rG4 favoriserait la liaison de la NCL à ces rG4 à la manière du mécanisme de liaison-dépliement-blocage (*bind-unfold-lock*) où une première protéine est nécessaire pour la liaison d'une seconde à un G4 (Herviou et al., 2020).

L'expression des protéines contenant le domaine GAr semble nécessaire pour l'interaction entre l'ARNm et la NCL. Pour tester cette hypothèse, nous avons traité des cellules avec des drogues inhibant la traduction. En réponse à un traitement par ces molécules, l'interaction entre l'ARNm contenant le domaine GAr et la NCL est fortement réduite alors qu'elle demeure inchangée en absence du domaine GAr. Des tests supplémentaires, avec des plasmides ne contenant pas de codon d'initiation de la traduction, mais dont le niveau ARNm est normal, montrent qu'en absence de traduction, la NCL interagit moins avec l'ARNm contenant un domaine GAr et dépourvu de codon START. En conclusion, l'ARNm contenant le domaine GAr doit d'abord être traduit afin de libérer la protéine NACA de son complexe. Celle-ci va alors interagir avec les rG4 d'EBNA1 et recruter la NCL au niveau des rG4, ce qui va conduire à l'inhibition de la traduction de l'ARNm d'EBNA1 et permettre ainsi l'évasion d'EBV au système immunitaire.

10110–10122 *Nucleic Acids Research, 2022, Vol. 50, No. 17* *Published online 15 September 2022*
https://doi.org/10.1093/nar/gkac751

The nascent polypeptide-associated complex (NAC) controls translation initiation *in cis* by recruiting nucleolin to the encoding mRNA

Alice J.-L. Zheng[1,†], Aikaterini Thermou[1,2,†], Chrysoula Daskalogianni[1,2], Laurence Malbert-Colas[1], Konstantinos Karakostis[1], Ronan Le Sénéchal[3], Van Trang Dinh[3], Maria C. Tovar Fernandez[1,2], Sébastien Apcher[4], Sa Chen[5], Marc Blondel[3] and Robin Fahraeus[1,5,6,*]

[1]Inserm UMRS 1131, Institut de Génétique Moléculaire, Université de Paris, Hôpital St. Louis, F-75010 Paris, France, [2]ICCVS, University of Gdańsk, Science, ul. Wita Stwosza 63, 80-308 Gdańsk, Poland, [3]Inserm UMR 1078, Université de Bretagne Occidentale (UBO), Etablissement Français du Sang (EFS) Bretagne, CHRU Brest, 29200, Brest, France, [4]Institut Gustave Roussy, Université Paris Sud, Unité 1015 département d'immunologie, 114, rue Edouard Vaillant, 94805 Villejuif, France, [5]Department of Medical Biosciences, Building 6M, Umeå University, 901 85 Umeå, Sweden and [6]RECAMO, Masaryk Memorial Cancer Institute, Zluty kopec 7, 65653 Brno, Czech Republic

Received July 21, 2022; Editorial Decision August 12, 2022; Accepted September 10, 2022

ABSTRACT

Protein aggregates and abnormal proteins are toxic and associated with neurodegenerative diseases. There are several mechanisms to help cells get rid of aggregates but little is known on how cells prevent aggregate-prone proteins from being synthesised. The EBNA1 of the Epstein-Barr virus (EBV) evades the immune system by suppressing its own mRNA translation initiation in order to minimize the production of antigenic peptides for the major histocompatibility (MHC) class I pathway. Here we show that the emerging peptide of the disordered glycine–alanine repeat (GAr) within EBNA1 dislodges the nascent polypeptide-associated complex (NAC) from the ribosome. This results in the recruitment of nucleolin to the GAr-encoding mRNA and suppression of mRNA translation initiation *in cis*. Suppressing NAC alpha (NACA) expression prevents nucleolin from binding to the GAr mRNA and overcomes GAr-mediated translation inhibition. Taken together, these observations suggest that EBNA1 exploits a nascent protein quality control pathway to regulate its own rate of synthesis that is based on sensing the nascent GAr peptide by NAC followed by the recruitment of nucleolin to the GAr-encoding RNA sequence.

INTRODUCTION

The formation and persistence of protein aggregates is associated with numerous neurodegenerative diseases, including Alzheimer's disease with the development of β-amyloid plaques, or in Huntington's disease where the synthesis of polyQ stretches in the huntingtin protein causes the formation of assemblies through weak side-by-side interactions (1). Apart from β-amyloids, aggregates are predominantly composed of misfolded or unfolded proteins (2). Upon acute stress, cells accumulate unfolded proteins that are cleared by different cellular pathways, including autophagy, secretion in exophers, or interactions with heat shock proteins (HSPs). HSPs prevent the formation of aggregates by supporting the proper folding or engaging with misfolded or aggregate-prone proteins for refolding or to target them for degradation (2–4). Ribosome stalling leads to co-translational ubiquitination and degradation of the nascent polypeptide via the ribosome-associated protein quality control (RQC) pathway (5–8) but how the cell senses and prevents the synthesis of aggregate-prone proteins is still relatively unknown.

The nascent polypeptide-associated complex (NAC) is present at the ribosome exit tunnel and is composed of NAC alpha (NACA) and BTF3 (also called NAC beta or NACB) (9,10). NAC is suggested to be the first chaperone the nascent polypeptide encounters and crystallography studies show that the BTF3 N-tail is present within the exit tunnel and escorts the nascent polypeptide (11). It interacts with short nascent polypeptide sequences and protects them from proteolysis and from interacting

*To whom correspondence should be addressed. Tel: +33 172639366; Email: robin.fahraeus@inserm.fr
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

with other nascent polypeptide-associated factors, such as the SRP (signal recognition particle) and the methionine aminopeptidase (12,13). The flexible conformation of NAC allows it to interact with a wide range of substrates and NAC chaperone activity has been shown to prevent aggregate accumulation outside of its ribosome-bound state (14–16).

mRNA translation initiation is a tightly regulated dynamic process governed by multiple signalling pathways and factors to allow a selection of mRNAs to be translated depending on cell type and conditions (17,18). RNA-binding proteins (RBPs) are adaptable regulators that interact with RNA sequences, or with RNA structures, and play a key role in guiding translation (19,20). It has also been suggested that translation drives the formation of RNA structures (21). The binding of nucleolin (NCL) to G-quadruplex (G4) structures in the glycine-alanine repeat (GAr) domain of the Epstein-Barr virus (EBV)-encoded *EBNA1* mRNA forms an essential part in the viral immune evasion strategy by suppressing translation initiation *in cis* in order to minimize the production of antigenic peptides for the major histocompatibility (MHC) class I pathway (22–25). The GAr-encoded peptide forms an aggregate-prone disordered domain consisting of single alanines separated by one, two, or three glycines and varies in length depending on viral strain but can reach over 200 amino acids (Supplementary Table S1). The GAr-encoding mRNA suppresses mRNA translation initiation of every open reading frame to which it is fused and introducing serine residues in every eight amino acids alleviates GAr-mediated inhibition of translation (24,26).

Here, we have used the GAr to show how a nascent disordered peptide acts together with G4 structures within the encoding RNA to control mRNA translation initiation *in cis*. Our results suggest that in order to minimise the synthesis of antigenic peptides, EBNA1 exploits a nascent protein quality control pathway that serves to prevent the synthesis of aggregate-prone proteins and is based on the dislodgment of NAC by the nascent peptide.

MATERIALS AND METHODS

Cell culture

Human carcinoma-derived cell line H1299, EBV carrying Burkitt's lymphoma cell line Raji, the *in vitro* transformed B95.8 and nasopharyngeal carcinoma cell line NPC 666-1 were all cultivated under standard conditions in RPMI 1640 medium containing 10% fetal calf serum, 2 mM L-glutamine, and 100 IU/ml penicillin and streptomycin (Gibco-BRL). NPC 666-1 was a kind gift from Pr. Kwok-Wai Lo from the Chinese University of Hong Kong.

RNA G4 structures and solubility prediction of the polypeptides

RNA G4 structures predictions were realised using the webserver QGRS mapper (27). Solubilities of the polypeptides for *E.coli* expression were predicted using the webserver SoluProt v1.0 (28).

Western blotting

Cells were harvested 40h post-transfection and lysed in the presence of a complete protease inhibitor cocktail (Roche Diagnostics). Total cell extracts were fractionated by SDS-PAGE, transferred to BioTrace®NT nitrocellulose blotting membranes (Pall Corp.). After incubation with the appropriate primary antibodies and peroxidase-conjugated secondary antibodies (Dako), proteins were visualized by using Pierce ECL, WestDura, or West Femto (ThermoFisher) and quantified with the MyECL Imager system (Thermo Scientific) and ImageJ. The following primary antibodies were used: mouse monoclonal anti-EBNA1 antibody (OT1X, Cyto-Barr), mouse monoclonal anti-p53 antibody (DO-1), mouse polyclonal anti-HA antobody (provided by Dr B. Vojtěšek, Masaryk Memorial Cancer Institute, Brno, Czech Republic), rabbit polyclonal anti-NACA antibody (Abcam), anti-Ovalbumin whole serum (Sigma), rabbit polyclonal anti-NCL antibody (Abcam), rabbit polyclonal anti-LC3B antibody (Sigma) and mouse monoclonal anti-actin antibody (Sigma). Relative quantifications of the HRP signals normalised with the corresponding actin bands are mentioned above the band.

RNA extraction, RT-qPCR, and RNA *in vitro* co-IP assay

Cells were washed in cold PBS and total RNA extraction was performed using RNAeasy Mini Kit (Qiagen), following the manufacturer's instructions. cDNA synthesis was carried out using the Moloney murine leukaemia virus M-MLV reverse transcriptase and Oligo(dT)$_{12\text{-}18}$ primer (Life technologies). qPCR was performed using the StepOne real-time PCR system (Applied Biosystems) with Perfecta SYBR Green FastMix (Quanta Biosciences).

In vitro RNA co-IP was carried out as described elsewhere (29). Briefly, 1μg of total RNA extracted from cells was co-incubated under agitation with 100 ng of recombinant NCL (provided by Dr. M.-P. Teulade-Fichou, Institut Curie, Paris, France) in binding buffer (50 mM Tris pH7.5, 150 mM NaCl, 0.02 mg/ml yeast tRNA, 0.2 mg/ml BSA) for 15 min at 37°C. After incubation, NCL-RNA complexes were pulled down at 4°C using G-coated sepharose beads (Sigma) with anti-NCL rabbit polyclonal antibody (Abcam) according to standard conditions and purified using the TRIzol (Life Technologies). Precipitated RNAs were then analysed by RT-qPCR.

MHC class I restricted antigen presentation

T-cell assays using the B3Z SIINFEKL:Kb-specific T cell hybridoma were carried out as described previously (30). Briefly, B3Z SIINFEKL:Kb-specific T cell hybridoma were co-cultured with H1299 cells transfected with both Kb and the reporter construct, or with the empty vector (EV), for 20 h. B3Z CD8+ T cell hybridoma expresses LacZ in response to activation of T cell receptors specific for the ovalbumin's immuno-dominant SIINFEKL peptide in the context of H-2Kb MHC class I molecules.

The cells were harvested and washed 2 times with 1X cold PBS before lysis in 0.2% TritonX-100, 0.5M K_2HPO_4, 0.5M KH_2PO_4 for 5 min on ice. Supernatants from each condition were transferred into 96-well optiplate counting plates (Packard Bioscience, Randburg, SA) and incubated for 1 h at room temperature, protected from light, and tested for

β-galactosidase activity using the Luminescence assay (BD Biosciences Clontech) on a FLUOstar OPTIMA (BMG LABTECH Gmbh, Offenburg, Germany). The results are expressed in luminescence units.

RNA FiSH

H1299 cell were seeded in 24-well plates (2 x 100 cells/ well) and trasient transfections were carried out 24 h later using the Genejuice reagent (Merck Biosciences) according to the manufacturer's protocol. 24 h after transfection the cells were briefly washed with ice-cold PBS, fixed with 4% PFA during 20 minutes at room temperature, and washed again with PBS. Cell were then incubated in 70% ethanol for 4-24 h at 4°C. For rehydration, the cells were incubated in 50% and 30% ethanol and further washed with PBS. Subsequently, cells were permabilised with PBS 0.4% Triton 0.05% CHAPS for 5 minutes at room temperature. Coverslips were incubated overnight in a wet chamber at 37°C in FiSH hybridisation buffer supplemented with 10% dextran sulphate and 100nM of FiSH Stellaris probes targeting Ovalbumin or EBNA1 mRNAs (Biosearch Technologies). Coverslips were washed twice 20 min in FiSH hybridisation buffer and 5 min in FiSH Wash buffer and subsequently stained with DAPI. Images were obtained using Zen software (Zeiss) and the Costes colocalisation factor between nuclei (DAPI channel) and targeted mRNAs (Cy3 channel), called Correlation_Costes_Cy3_DAPI, were obtained using the software CellProfiler (31), and later used for statistical analysis.

Proximity ligation assay (PLA) protein–protein and immunostaining

Cells were cultured, fixed, and permeabilised as described above. Primary antibodies incubation and PLA were carried out using the Naveniflex MR kit (Navinci), following the manufacturer's protocol. The aggregates staining were realised using the Proteostat Aggresome detection kit (Enzo) following the manufacturer's instructions. PLA antibodies: mouse polyclonal anti-HA antibody, rabbit polyclonal anti-NACA antibody (Abcam), rabbit polyclonal anti-HA antibody (Sigma), mouse polyclonal anti-NACA antibody (Abnova) and mouse monoclonal anti-EBNA1 antibody (OT1X, Cyto-Barr). Images were obtained using Zen software and analysed using CellProfiler. Data were processed by taking into account the difference in targeted protein expression, meaning the number of PLA dots were normalised with the corresponding cells integrated immunofluorescence signal measured, and then the mean relative difference between the different conditions was calculated.

Protein co-immunoprecipitation (co-IP)

After centrifugation, cell pellets from Raji or B95.8 cultures were lysed in buffer containing 20 mM Tris (pH 7.5), 150 mM NaCl, 1% NP-40 in the presence of complete protease inhibitor cocktail (Roche). Lysates were immunoprecipitated with anti-EBNA1 goat antibody or goat IgG and protein G-sepharose. The beads were washed with PBS and lysis buffer x4 and boiled in SDS loading buffer. Immunoprecipitates were analysed by SDS/PAGE using 4–12% precast gels (Invitrogen).

Polysome fractionation and PLEA

5–50% (w/vol) linear sucrose gradients were freshly cast on SW41 ultracentrifuge tubes (Beckmann) using the Gradient master (BioComp instruments) following the manufacturer's instructions. Polysome fractions were collected and concentrated to 100 μl using the Millipore concentrating falcon tubes. Protein concentration was measured by Bradford and an equal amount of protein from each sample was used for the PLEA experiment to study the interaction of three molecules. 96-well ELISA plates were incubated with the capture antibody at a dilution of 1:200, o/n, at 4°C. Samples were incubated with the PLA secondary antibodies and the PLA kit (Sigma) was used according to manufacturers' instructions. The fluorescence at 640 nm was measured by the FLUOstar plate reader (excitation at 644 nm and emission at 669 nm). The values were used for the preparation of a graph and for statistical analysis to calculate the mean and the standard deviation using the GraphPad Prism 9 software. Each sample was tested in triplicates.

RNA pulldown assays

For the preparation of whole-cell extracts, confluent H1299 cells were collected after trypsin treatment and washed twice with 1X PBS (Gibco). Cells were suspended in 500 μl of lysis buffer (20 mM Tris–HCl pH 7.5; 200 mM NaCl and 0.1% Igepal) containing 1× protease inhibitor cocktail (Roche). Cell lysis was performed by five series of vortex followed by 10 min incubation on ice, and 3 series of 3s sonication at 20% amplitude. After lysis cells were centrifuged at 4°C for 15 min at 16 000g, and the supernatant was quantified by Bradford. The whole-cell extracts or recombinant GST-NCL (Abnova) and His-tagged NACA (homemade, Umea University) were used for pulldown assays with the following G-quadruplex forming oligonucleotides: GQ-18 5′-GGGGCAGGAGCAGGAGGA-3′Biotin TEG. The negative control for EBNA1 G4 was the GM-18 5′ GAGGCAGUAGCAGUAGAA-3′Biotin TEG oligonucleotide which, according to the QGRS mapper software (Ramapo College New Jersey https://bioinformatics.ramapo.edu/QGRS/credits.php), is unable to form G4 structures. To avoid unspecific binding, high-affinity streptavidin sepharose beads (GE Healthcare) were incubated in 1 ml blocking buffer containing 10 mM Tris–HCl pH 7.5; 100 mM KCl; 0.1 mM EDTA; 1 mM DTT; 0.01% Triton X-100; 0.1% BSA; 0.02% *S. cerevisiae* tRNAs (Sigma), for 1 h at 4 °C on a rotating wheel. An amount of 10 pg of each folded biotinylated RNA oligos was incubated with 50 μl of the solution containing the streptavidin sepharose beads for 90 min at 4°C on a rotating wheel. Five hundred micrograms of cell extract were incubated with the RNA oligonucleotides bound to the streptavidin beads for 90 min at room temperature. Beads were washed with increasing KCl concentration (200–800 mM). Protein still bound to beads after the washes were eluted using 2× SDS loading buffer and analysed by WB against NACA or NCL, as previously described. In the input lane of the WBs was loaded a fraction

of the extract that was incubated with the beads for each condition.

Statistical analysis

Data were analysed by unpaired Mann–Whitney's test or Student's *t*-test on GraphPad Prism 9. On graphs, represented data are the mean and the standard deviation or SEM of a minimum of three independent experiments. Statistical significance of the difference is stated as following: $P > 0.05$ (ns), $P < 0.05$ (*), $P < 0.01$ (**) and $P < 0.001$ (***).

RESULTS

NAC controls the synthesis of GAr-containing polypeptides

The gly-ala repeat (GAr) domain is located in the N-terminal part of the EBNA1 and deleting this domain (EBNAΔGAr) increases mRNA translation and protein expression (Figure 1A and B) (23–26,32). Previous studies have shown that the binding of nucleolin (NCL) to the *EBNA1* mRNA is important for translation suppression *in cis* (22). However, antibodies against the GAr domain of EBNA1 overcome translation suppression *in vitro* (23), showing that the peptide also plays a role in controlling translation initiation. We wanted to know how the mRNA and the encoded peptide both play a role in controlling EBNA1 synthesis. In order for both the RNA and the encoded peptide to control translation initiation *in cis*, we reasoned that the mechanism of action should take place before the protein leaves the translating ribosome. NAC is described as a ribosome-associated chaperone targeting aggregate-prone proteins and we tested the impact of NAC on GAr peptide-mediated protein synthesis control. siRNA-mediated silencing of NAC subunit alpha (NACA) in H1299 cells leads to a decrease in *NACA* mRNA and protein levels and this resulted in an increase of protein aggregates and an increase in LC3 protein and polyQ-fused protein levels, showing that the reduced expression resulted in a functional response (Figure 1C and Supplementary Figures S1A, S1B and S1C) (16,33).This was accompanied by an approximately 4-fold increase in EBNA1 expression without affecting *EBNA1* mRNA levels (Figure 1C and Supplementary Figure S1D). A similar increase in expression following NACA knockdown was also observed when GAr was fused to the N-terminus of p53 (GAr-p53) while only a limited, or no, effect was observed on p53 alone (Figure 1D). Fusing the GAr to ovalbumin (Ova) (GAr-Ova) also resulted in a NACA-dependent expression control (Supplementary Figure S1E). These results show that silencing NACA interferes with GAr-mediated translation inhibition. In addition, the GAr is known to be more efficient in suppressing translation when placed in the 5′ of the coding sequence, as compared to the 3′ (23,26), and NACA silencing had less impact when GAr was fused to the 3′ of the Ova (Ova-GAr), supporting the notion that NACA acts on GAr-mediated translation control (Supplementary Figure S1E). To further test that the effect of NACA on EBNA1 expression is at the level of mRNA translation, we took advantage of the fact that fusing the 5′ untranslated region (UTR) of *c-Myc* to the 5′ UTR of EBNA1 (cMyc_EBNA1) overcomes GAr-mediated translation inhibition (Figure 1E) (26,32,34). When we knocked out NACA we observed no effect on the expression of EBNA1 carrying the *cMyc* sequence in the 5′ UTR (Figure 1F). Similarly, when we fused the *cMyc* sequence to the 5′ UTR of GAr-p53 or GAr-Ova constructs, we did not observe any increase in expression following NACA knockdown, validating that the effect of NACA is indeed on the synthesis of EBNA1 and not on protein stability (Figure 1G and Supplementary Figure S1F). Overexpressing an HA-tagged NACA did not affect EBNA1 expression, indicating that the endogenous NACA levels are sufficient to control EBNA1 synthesis (Figure 1H). To ensure that endogenous EBNA1 expression is controlled by NAC, we also treated the EBV-carrying NPC 666-1 and Raji cells with siRNA against NACA and we observed a similar increase in EBNA1 expression (Figure 1I and Supplementary Figure S1G). The physiological role of the GAr is to minimise the production of EBNA1-derived peptide substrates for the MHC class I pathway and we wanted to know if NACA affects the production of antigenic peptide substrates. The Ova includes an antigenic peptide (SL8; see Figure 1A) that is presented on the murine Kb class I molecules. When we expressed Ova, or GAr-Ova, and exposed the cells to B3Z CD8 + T cells that are specific for SL8, we observed an increase in antigen presentation following NACA siRNA treatment in cells expressing GAr-Ova but not in cells expressing Ova. Silencing SRP (Signal Recognition Particle), another nascent polypeptide-associated factor, had no impact on antigen presentation (Figure 1J). Taken together, these results show that NAC plays a key role in GAr-mediated suppression of mRNA translation initiation and in the viral strategy to evade the immune system.

NAC interacts with the GAr peptide domain

To address how NAC mediates GAr-dependent translation control, we performed Proximity Ligation Assay (PLA) to see if endogenous NACA interacts with EBNA1 or the EBNAΔGAr *in situ*. We observed fewer interactions with the EBNAΔGAr, as compared to EBNA1, despite EBNAΔGAr being expressed at four times higher levels. The interactions were mainly observed in the nuclear compartment (Figure 2A). As a control we used the HA-polyQ-p53 as poly-glutamine repeat (PolyQ) causes aggregates and is known to interact with NAC (Supplementary Figure S2A). In line with previous studies, immunofluorescence assays confirmed a cytoplasmic and nuclear localisation of NACA (Supplementary Figure S2B) (35). After taking into account the difference in expression levels between EBNA1 and EBNAΔGAr (see Figure 1B), we estimated that deleting the GAr significantly decreased the number of interactions between endogenous NACA and EBNA1 peptide (Figure 2B). Overexpressing an HA-NACA construct together with EBNA1, or EBNAΔGAr, showed an approximately 3-fold increase in the number of interactions between NACA and EBNA1 peptide, as compared to NACA and EBNAΔGAr (Figure 2C). Co-immunoprecipitation assays from two EBV-carrying B cell lines (B95.8 and Raji) showed that endogenous EBNA1 protein interacts with NACA (Figure 2D). We also wanted to see if NACA interacts with

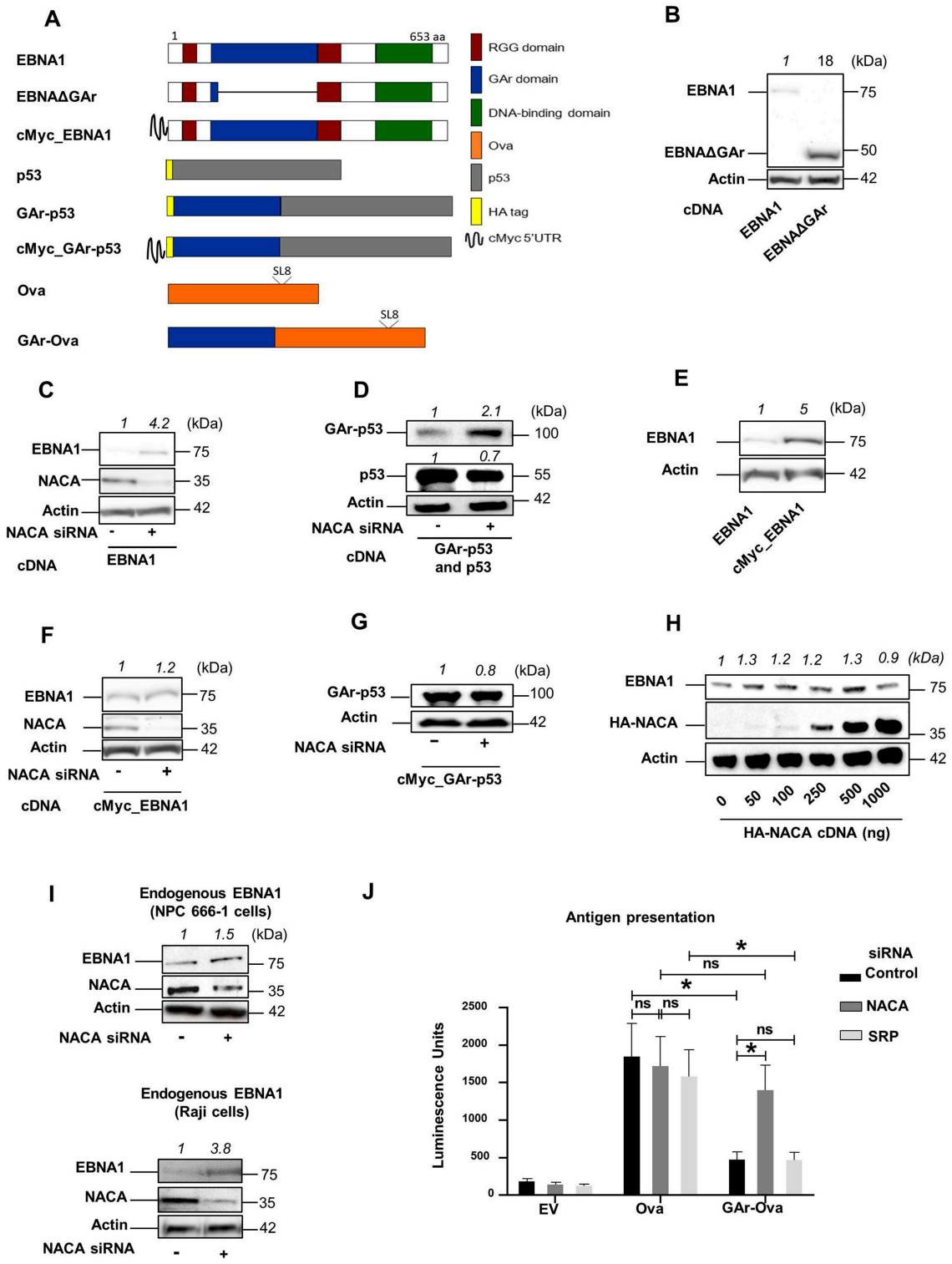


Figure 1. (**A**) Cartoon illustrating the constructs used. Position of Ova-derived SIINFEKL antigenic MHC class I peptide (SL8) is indicated. (**B**) Western Blot (WB) from H1299 cells transfected with EBNA1 cDNA or an EBNA1 that lacks the gly-ala repeat (EBNAΔGAr). (**C**) WB from H1299 cells transfected with EBNA1 and treated with siRNAs control (-) or with siRNA against NACA (NAC subunit alpha) (+). (**D**) WB from H1299 cells transfected with GAr-p53 or p53 constructs and treated with indicated siRNAs. (**E**) WB from H1299 cells transfected with EBNA1 or an EBNA1 construct carrying *cMyc* in its 5′ UTR (cMyc_EBNA1). (**F**) WB from H1299 cells expressing cMyc_EBNA1 and treated with indicated siRNAs. (**G**) WB of H1299 expressing the cMyc_GAr-p53 construct and treated with indicated siRNAs. (**H**) WB of H1299 cells transfected with an EBNA1 construct and increasing amounts of HA-NACA-encoding plasmid. (**I**) WB of endogenous EBNA1 in EBV-carrying NPC 666-1 cells (upper panel) or Raji cells (lower panel) treated with indicated siRNAs. (**J**) Antigen presentation assay performed using B3Z SIINFEKL:Kb-specific T cell hybridoma co-cultured with H1299 cells transfected with Ova, GAr-Ova or empty vector (EV) and treated with siRNAs against NACA or SRP. Quantification of WBs is indicated on the top of each immunoblot. Proteins detected are indicated on the left of the immunoblots. Data represent three independent experiments.

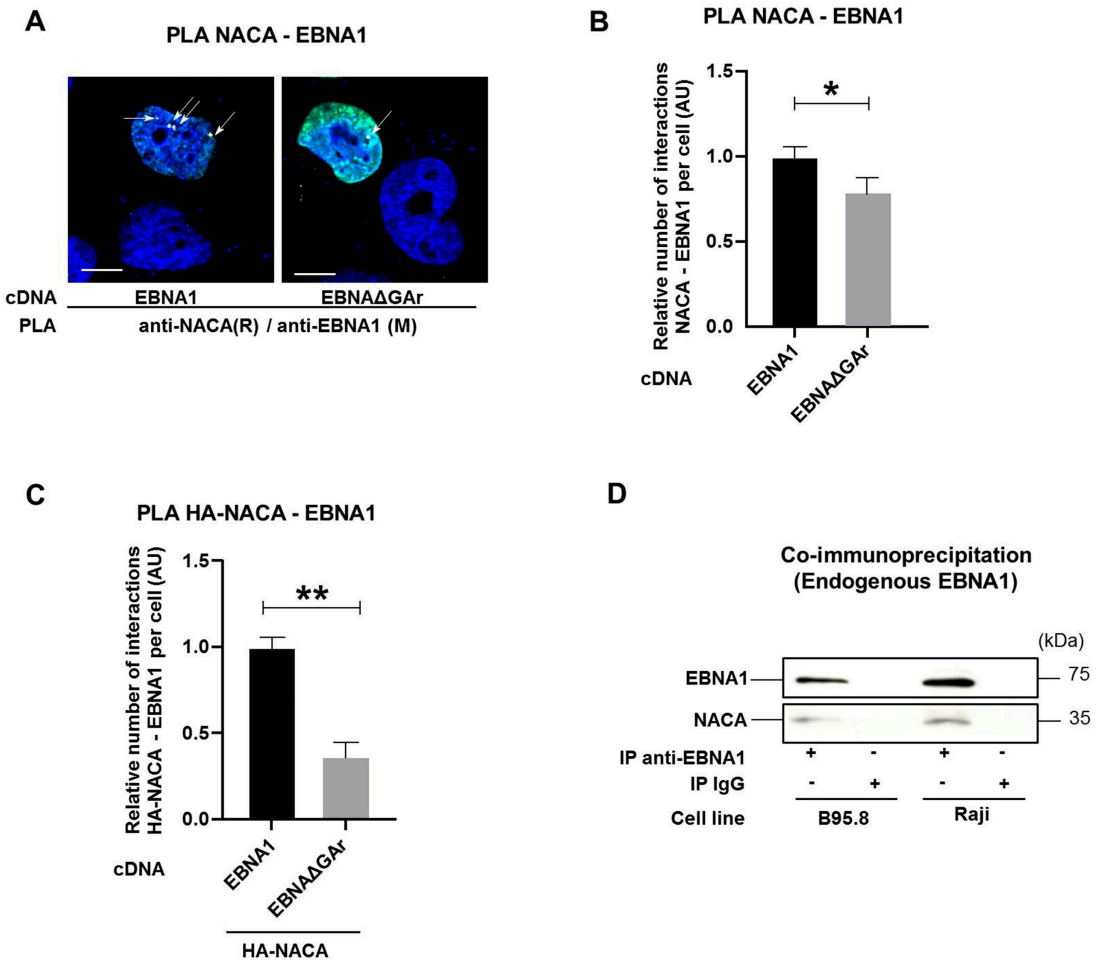


Figure 2. (**A**) Proximity ligation assay (PLA) (white arrows) shows interactions in H1299 cells between NACA and EBNA1 or NACA and EBNAΔGAr. Co-immunofluorescence of respective proteins shown in green. DAPI nuclear staining in blue. Scale bar represents 10 μm. (**B**) Relative number of interactions of indicated reporter proteins with endogenous NACA detected by PLA and normalised with the expression of the corresponding reporter proteins in H1299 cells expressing EBNA1 or EBNAΔGAr. (**C**) Relative number of interactions between an exogenous HA-tagged NACA and EBNA1 or EBNAΔGAr. (**D**) Co-IP using anti-EBNA1 or anti-IgG (negative control) antibodies in EBV-carrying B95.8 and Raji cell lysates. Data represent three independent experiments. A minimum of 50 cells expressing indicated reporter protein was counted for each PLA experiment.

the newly synthesised EBNA1 peptide and we therefore performed a metabolic labelling assay using S35 methionine in order to visualise newly synthesised polypeptides that co-immunoprecipitate with NACA. When comparing control cells (EV) with cells transfected with EBNA1, we observed a band corresponding to the size of EBNA1 in cell lysates immunoprecipitated with NACA antibodies. The same size band was detected in Western Blot (Supplementary Figure S2C). These results indicate that NAC interacts with the newly synthesized GAr peptide.

The nascent GAr peptide dislodges NAC from the ribosome

We next tested if NAC interacts with the nascent EBNA1 on the ribosome, as this would help explain how the interaction between GAr and NAC controls *EBNA1* translation *in cis.* We expressed p53 and GAr-p53 and carried out polysomal fractionation using sucrose gradients on cycloheximide-treated cell lysates (Figure 3A). The isolated polysomes were tested for the presence of *GAr-p53* mRNA and nascent GAr-p53 derived peptides in order to ensure that they were actively translating the GAr-p53 message (Supplementary Figures S3A and S3B). The polysomes were captured on 96-well plates using goat anti-RPL5 sera. We did an adapted PLA ELISA (PLEA) by adding increasing amounts of polysomes followed by *in vitro* PLA using anti-p53 (rabbit) and anti-NACA (mouse) antibodies (Figure 3B). The fusion of the GAr sequence with p53 resulted in an average 3-fold stronger interaction between the nascent GAr-p53 peptide and NACA, supporting the notion that NAC interacts strongly with the nascent GAr at the riboosme (Figure 3C). To test if the high affinity between GAr and NAC could lead to the dissociation of NAC from the ribosome we designed a construct in which a TEV protease cleavage site was inserted between the GAr and the p53 (GAr-TEV-p53). The GAr peptide will, thus, exit the ribosome and encounter NAC before the TEV and the p53 sequence. To ensure that TEV cleaved the GAr-TEV-p53 fusion protein we added TEV enzyme to cell lysates for 60nbsp;min and we could observe that a major part of

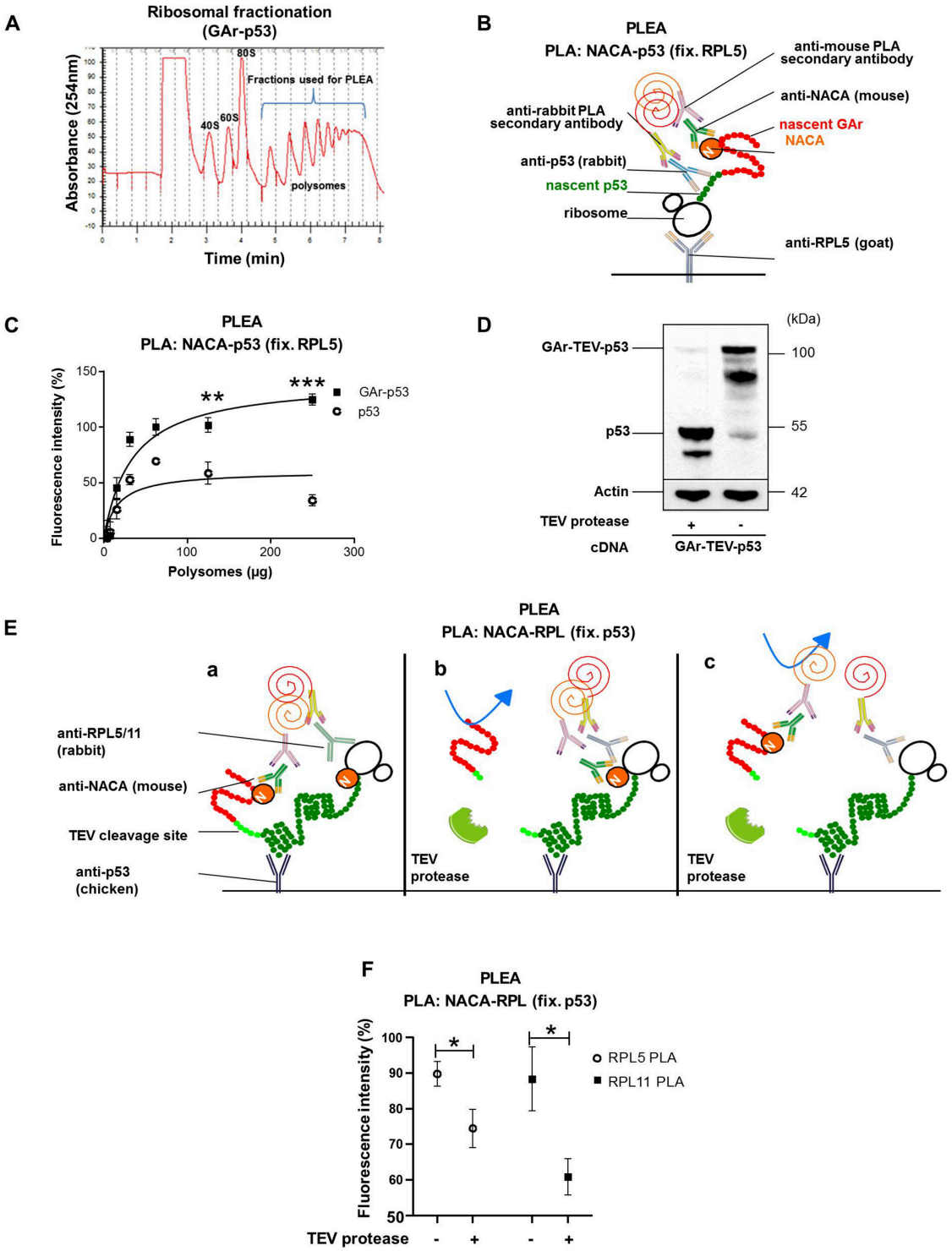


Figure 3. (**A**) Ribosomal fractionation of H1299 cells. Polysomal fractions collected and used for the PLEA experiments are indicated in blue. (**B**) Cartoon illustrating the PLEA experiment. Polysomes were fixed with a goat anti-RPL5 antibody and the PLA was performed on the captured polysomes and their associated complex using anti-NACA (M = mouse) and anti-p53 (R = rabbit) antibodies. (**C**) The graph shows PLEA results where more interactions between p53 and NACA were detected on GAr-p53, as compared to p53, expressing polysomes. The data show normalised fluorescence intensity for GAr-p53 and p53 polysomes. T-tests were performed between GAr-p53 and p53 values for the same amount of polysomes. (**D**) WB of H1299 cells transfected with a construct in which the TEV proteolytic cleavage site was inserted between GAr and p53 (GAr-TEV-p53). The lysates were treated, or not, with TEV protease. (**E**) Cartoons illustrate the PLEA experiment to assess if GAr dislodges NACA from the ribosome. (**F**) Polysomes from H1299 cell lysates expressing GAr-TEV-p53 and treated with TEV protease, or not, were captured using a chicken anti-p53 antibody. PLEA experiments were performed using 125µg of polysomes and PLA was carried out using anti-NACA(M = mouse) with anti-RPL5(R = rabbit) antibodies (white circles) or anti-NACA(M) with anti-RPL11(R) antibodies (black squares). The graph shows the relative amount of NACA bound to GAr-TEV-p53 translating ribosomes before and after treatment with TEV protease. The data represent three independent experiments.

GAr was indeed cleaved from p53 (Figure 3D). We then expressed GAr-TEV-p53 and carried out ribosomal fractionations on TEV-treated and non-TEV-treated lysates (Supplementary Figure S3C). p53-carrying polysomes (125 μg) were isolated and fixed to 96-well plates using chicken anti-p53 antibodies and PLEA was performed against ribosomal proteins and NACA (see Figure 3C). If the GAr did not affect NACA's placement at the ribosome, we would expect that TEV protease treatment would not make any difference on the PLEA signal between NACA and ribosomal proteins (RPL5 or RPL11) (Figure 3E; a and b). However, if the GAr sequence caused NACA to detach from the ribosome, TEV treatment would result in reduced interactions between NACA and RPL5 or RPL11 (Figure 3E; c). The TEV treatment indeed resulted in less PLA signal between NACA - RPL5/RPL11 on GAr-TEV-p53-expressing polysomes, in line with the notion that NACA's recruitment to the GAr sequence causes NACA to dissociate from the ribosome (Figure 3F). These results show that the nascent GAr peptide interacts strongly with NACA on the ribosome and leads to NACA being dislodged from the ribosome, offering a first insight into how NAC affects mRNA translation of GAr-carrying mRNAs *in cis*.

NAC is necessary for the interaction between the *GAr*-encoding mRNA and NCL

The results show that the nascent GAr peptide plays an important role in GAr-mediated mRNA translation control. However, previous works have demonstrated the importance of the *GAr*-encoding mRNA sequence in suppressing mRNA translation (22,29). The effect of the *GAr* mRNA is attributed to NCL binding G-quadruplex (G4) structures in the *GAr*-encoding mRNA and we wanted to know if there is a link between NAC and the recruitment of NCL to the *GAr* mRNA. To test this, we carried out *in vitro* RNA-protein co-immunoprecipitation (RNA-coIP) assays using recombinant NCL together with RNA isolated from cellular extracts, followed by RT-qPCR for the indicated mRNAs. Silencing NACA resulted in a dramatic decrease in the binding of NCL to the *EBNA1* and *GAr-Ova* mRNAs while no significant change was observed for the already weak interaction between NCL and the *EBNAΔGAr* or *Ova* mRNAs (Figure 4A and B). NAC has been described to bind nucleic acids and we next tested if this capacity can be the link between the *GAr*-encoding RNA sequence and the GAr peptide (10). RNA-coIP using recombinant NACA and mRNAs isolated from cell lysates showed that NACA binds to the *EBNA1* mRNA, but not to the *EBNA1ΔGAr* mRNA (Figure 4C). To test if the NACA – *EBNA1* mRNA interaction is direct, we carried out *in vitro* RNA pulldown assays using recombinant NACA protein together with synthetic oligos forming the RNA G4 structure of the GAr (GQ-18) or a mutated version of GQ-18 that does not form a G4 (GM-18) (22). This showed that NACA binds RNA oligos in a G4-dependent manner (Figure 4D). Interestingly, adding recombinant NCL did not prevent NACA interaction, suggesting that they have different interaction sites (Figure 4D). NACA binding to the G4 forming sequence of the GAr (GQ-18) was further confirmed using endogenous NACA protein pulled down from H1299 cell lysates (Figure 4E). Addition of an increasing amount of the G4 ligand PhenDH2 (32) prevented the interaction of endogenous NACA to the GQ-18, further confirming that NACA binds to the G4 RNA structure of *GAr*-encoding mRNA (Figure 4F). Having observed that both NACA and NCL interact with the *GAr* G4 RNA, we next tested if NACA and NCL proteins interact with each other. *In situ* PLA showed that endogenous NACA and NCL interactions were detected in the nuclear compartment and that this is not limited to EBNA1 expressing cells, as it is also detected in cells transfected with the EBNAΔGAr construct or with the empty vector (EV) (Figure 4G and Supplementary Figure S4A and S4B). These results show that NACA in addition to binding the GAr peptide also interacts with *GAr* mRNA sequence and promotes the NCL–*EBNA1* mRNA interaction.

Translation is required for the interaction between *GAr*-encoding mRNAs and NCL.

In line with the scenario that nascent GAr peptide dislodges NAC from the ribosome and this results in the recruitment of NCL to the *GAr* RNA G4 structure, the interaction between NCL and the *GAr* RNA would require *GAr*-encoding mRNAs to be translated. We tested this by treating cells with cycloheximide (CHX) or harringtonine (Harr), two drugs that prevent mRNA translation. As expected, both drugs suppress the expression of GAr-Ova, Ova, and endogenous p21, while NCL and actin with longer half-life were less affected (Figure 5A). Importantly, *in vitro* RNA-coIP showed that the interaction between the *GAr-Ova* mRNA and NCL was reduced by ∼4-fold following CHX or Harr treatment. However, the NCL–*Ova* mRNA interaction was not affected by translation inhibitor treatments (Figure 5B). To further test if translation is needed to allow NCL to bind the *GAr*-encoding mRNA, we designed Ova and GAr-Ova constructs lacking AUG start codons (OvaΔATG and GAr-OvaΔATG) abrogating the expression of respective proteins (Figure 5C) and we observed that NCL had a reduced affinity for the *GAr-OvaΔATG* mRNA, as compared to the *GAr-Ova* mRNA (Figure 5D). The deletion of the AUGs had no significant effect on mRNA levels (Supplementary Figure S5A) or on their respective subcellular localisation (Supplementary Figure S5B and C). Thus, the *GAr*-encoding mRNAs need to be translated for binding NCL, in line with the notion that the interaction between NAC and the nascent GAr peptide is required for the recruitment of NCL to the G4 structure of the *GAr*-encoding mRNA on the EBNA1 polysome.

DISCUSSION

Our data support a *cis*-acting protein-quality control pathway in which the disordered and aggregate-prone nascent gly-ala repeat (GAr) peptide of the EBNA1 dislodges NAC alpha (NACA) from the ribosome. NACA then interacts with the encoding *GAr* RNA sequence and helps recruit nucleolin (NCL) to the G-quadruplex (G4) structure of the *GAr* RNA, which leads to the suppression of mRNA translation initiation *in cis*. This illustrates how the EBV-encoded EBNA1 exploits a *cis*-acting cellular pathway for post-transcriptional gene regulation. The effect of the GAr

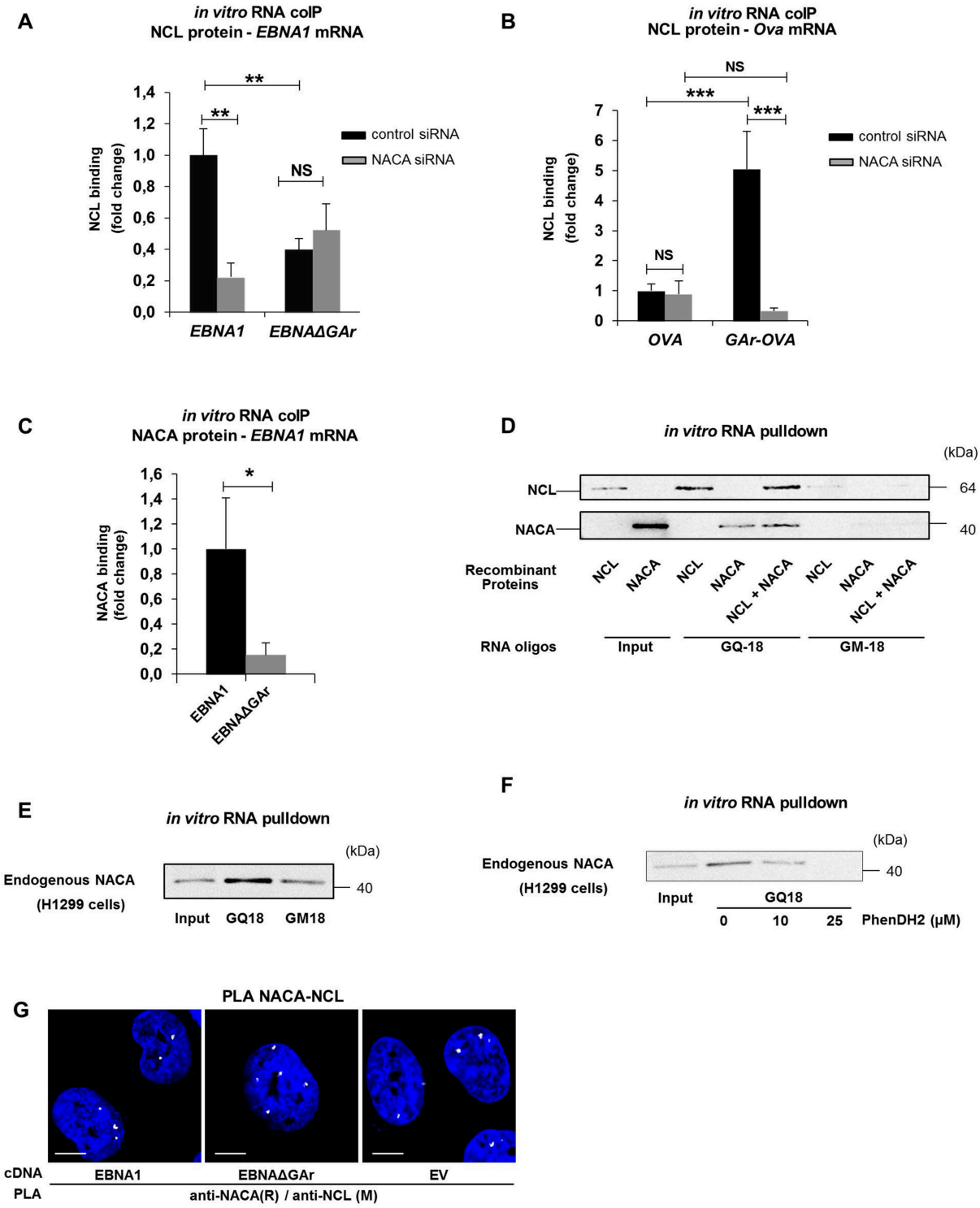


Figure 4. (**A**) *In vitro* RNA coIP experiments performed with recombinant NCL and *EBNA1* or *EBNAΔGAr* mRNA extracted from H1299 cells treated with the indicated siRNAs. The graph shows the fold change in NCL binding. (**B**) *In vitro* RNA coIP performed with recombinant NCL and *Ova* or *GAr-Ova* mRNAs from cells treated with the indicated siRNAs. The graph shows the fold change in NCL binding. (**C**) *In vitro* RNA coIP experiments performed with recombinant NACA and *EBNA1* or *EBNAΔGAr* mRNAs extracted from H1299 transfected cells. The graph shows relative fold change in NACA binding. (**D**) *in vitro* RNA pulldown assay using recombinant NCL and NACA and the indicated RNA oligos. GQ-18 is derived from the G4-forming sequence coding for the GAr. GM-18 is a mutated non-G4-forming version of GQ-18. (**E**) *in vitro* RNA pulldown assay using endogenous NACA from H1299 cells and the indicated RNA oligos. (**F**) *in vitro* RNA pulldown assay using endogenous NACA from H1299 cells, the GQ-18 oligo derived from the G4-forming sequence coding for the GAr and increasing amounts of the PhenDH2 G4 ligand. (**G**) PLA assessing endogenous NACA - NCL interactions in H1299 cells expressing EBNA1 or EBNAΔGAr (EV = empty vector). In blue: nucleus, in white: PLA dots. Scale bar represents 10μm. The data represent a minimum of three independent experiments.

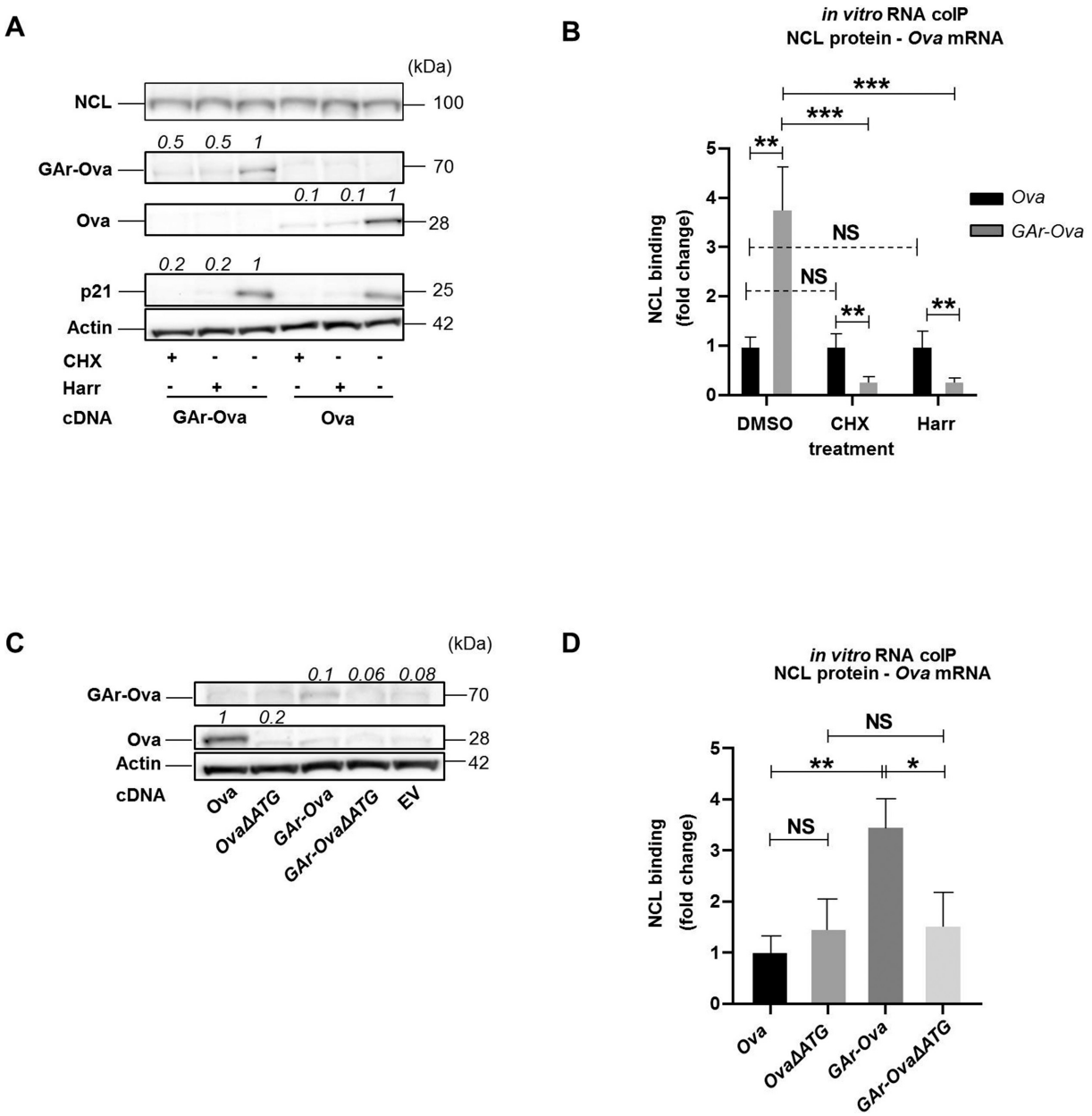


Figure 5. (**A**) WB of lysates from H1299 cells expressing indicated constructs and treated with cycloheximide (CHX), harringtonine (Harr) or DMSO control for 5h. (**B**) *in vitro* RNA coIP experiments performed with recombinant NCL and *Ova* or *GAr-Ova* mRNAs extracted from H1299 transfected cells treated with CHX, Harr or DMSO. The graph shows the relative fold change in NCL binding. (**C**) WB of H1299 transfected cells expressing constructs where the +1 AUG has been deleted (OvaΔATG and GAr-OvaΔATG) . (**D**) *In vitro* RNA coIP performed with recombinant NCL and mRNAs extracted from H1299 cells expressing *Ova,* OvaΔATG, GAr-Ova and GAr-OvaΔATG. The numbers indicate relative protein expression levels. The graph shows the relative fold change in NCL binding with the indicated mRNAs. Data presented are from three independent experiments.

on translation suppression is via initiation (26) and together with the fact that neither the RNA, nor the peptide, are degraded suggests that GAr-mediated translation control is a regulated process. For example, EBNA1 binds its own mRNA via a RGG-rich region and this could allow control of its own synthesis (36). EBNA1 is essential for viral survival but at the same time EBNA1 is immunogenic and this self-regulation of translation serves to minimize the production of antigenic peptides for the MHC class I pathway and at the same time guarantee a functional level of EBNA1 expression. The scenario in which the nascent GAr peptide serves as a trigger to suppress its own synthesis helps explain the observation that the GAr is more potent in suppressing synthesis when fused in the 5′ coding sequence rather than in the 3′ (Supplementary Figure S1E) (32). It is more likely that the presence of a disordered domain in the N-terminus, rather than the C-terminus, activates this pathway, in line with the fact that fusing a leader sequence to the N-terminus of an insoluble protein facilitates expression and solubility. There are few other examples of nascent peptides affecting

protein expression but a more recent study on tubulin illustrates how the nascent peptide affects protein expression via the stability of the encoding mRNA, suggesting that the nascent peptide can activate different pathways.

We show that NACA binds the GAr peptide sequence as well as the encoding mRNA and, thus, providing the link between the GAr peptide and the encoding mRNA in controlling EBNA1 synthesis. *GAr* mRNA forms RNA G4 structures that interact with NCL and this is required for GAr-mediated translation initiation inhibition *in cis*. The observation that NACA binds the *GAr* G4 RNA structure and recruits NCL to the *GAr*-encoding RNA suggests that NACA makes the RNA structure NCL 'friendly'. Previous studies have shown that the G4 RNA structures are highly dynamic *in vivo* with multi-functional properties and it is plausible that NACA affects the G4 structure to suit NCL binding (32). A similar 'bind-unfold-lock' mechanism was proposed for USP1 translational regulation by hnRNP H/F and DHX36 in glioblastoma or for the CNBP-targeted mRNAs (37,38).

An interesting aspect of NAC-mediated recruitment of NCL to the mRNA is where in the cell this event takes place. While NCL has a predominantly nucleolar localisation, NAC is present in the cytoplasm and the nuclear compartment. The PLA data show that the NACA – EBNA1 interaction takes place in the nucleus and it is an interesting possibility that this nascent peptide control pathway takes place during the ribosome-mediated RNA quality control that precedes nonsense-mediated decay.

The mechanism whereby the RNA and the encoded peptide act together via NAC to regulate mRNA translation *in cis* is specific in terms of both the RNA and the peptide sequence. Introducing a single serine in every eight residues of the GAr completely abolishes its translation inhibitory capacity (26) and introducing synonymous mutations also affects the efficacy of translation suppression (23). The GAr-mediated pathway of translation control is conserved in yeast (39). It is, however, plausible that this pathway reflects a specific example of a broader concept of regulating the synthesis of aggregate-prone proteins that includes the nascent peptide and the encoding RNA sequence but not necessarily via the same molecular mechanism.

It is not clear why both the RNA sequence and the encoding peptide should be required to suppress the synthesis of aggregate-prone proteins. One could argue that the sensing of a nascent disordered peptide should be sufficient to trigger suppression of synthesis in order to protect the cell from toxic aggregates. However, GC rich RNA sequences forming G4 structures, or disordered peptide domains in the N-terminus, are not uncommon and the fact that both peptide and RNA need to act together to trigger this pathway ensures that neither RNA sequence alone, nor peptide, will prevent synthesis. This could imply that aggregate-prone proteins that do not fulfil this double criterion would be more likely to be expressed at higher levels and accumulate aggregates. In the case of the aggregate-prone PolyQ of Huntingtin protein (HTT), the encoding RNA does not form G4 structures and it does not suppress translation when fused to the N-terminus of p53 (Supplementary Table S1) (26). It is also noteworthy that GC-rich sequences forming G4 structures encode peptides that can form aggregates, as is illustrated by patients suffering from Fragile X Syndrome. These patients have a (CGG) expansion that promotes alternative translation upstream of the classical *FMR1* ORF via repeat-associated non-AUG (RAN) (40,41). RAN is an alternative initiation mechanism mostly described for non-coding regions and introns and the encoded peptides cause toxic aggregates (42). Another example is from amyotrophic lateral sclerosis (ALS) which is associated with a GC-rich sequences from the (GGGGCC) repeats expansion in the *C9ORF72* gene (43). Both (CGG) and (GGGGCC) repeat expansions form G4 structures (43–45) and the encoded polypeptides are prone to form aggregates (46,47).

DATA AVAILABILITY

All data are available upon request.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the personnel of the Plateforme Technologique de l'Institut de Recherche Saint-Louis (Paris, France) for their help in microscopy. We would also like to thank Dr M.-P. Teulade-Fichou, Dr A. Granzhan, and their teams (Institut Curie, Paris, France) for supplying us with recombinant NCL. We are grateful to Dr Sa Chen and Dr. Mikael Lindberg for the preparation of recombinant His-tagged NACA proteins. The chicken anti-p53 antibodies was generated by Dr Mario Señorale at the Sección Bioquímica, Facultad de Ciencias, Universidad de la República, Iguá 4225, Montevideo, Uruguay.

Author contributions: A.Z., A.T. and C.D. carried out the majority of the work. L.M.-C., A.Q., R.L.S., T.D., M.T.F., G.M., K.K., S.A. contributed. M.B. and R.F. supervised and helped write the MS.

FUNDING

A.Z. was funded by Université de Paris and the Fondation pour la Recherche Médicale [FDT202001010912]; European Regional Development Fund [ENOCH, CZ.02.1.01/0.0/0.0/16_019/0000868]; MH CZ – DRO [MMCI, 00209805]; Cancerforskningsfonden Norr, Cancerfonden [160598]; Vetenskapsradet; International Centre for Cancer Vaccine Science within the International Research Agendas program of the Foundation for Polish Science co-financed by the European Union under the European Regional Development Fund. Funding for open access charge: Department of Medical Biosciences, Building 6M, Umeå University Umeå, Sweden.

Conflict of interest statement. None declared.

REFERENCES

1. Candelise,N., Scaricamazza,S., Salvatori,I., Ferri,A., Valle,C., Manganelli,V., Garofalo,T., Sorice,M. and Misasi,R. (2021) Protein aggregation landscape in neurodegenerative diseases: clinical relevance and future applications. *Int. J. Mol. Sci.*, **22**, 6016.

2. Mogk,A., Bukau,B. and Kampinga,H.H. (2018) Cellular handling of protein aggregates by disaggregation machines. *Mol. Cell*, **69**, 214–226.
3. Goncalves,C.C., Sharon,I., Schmeing,T.M., Ramos,C.H.I. and Young,J.C. (2021) The chaperone HSPB1 prepares protein aggregates for resolubilization by HSP70. *Sci. Rep.*, **11**, 17139.
4. Genest,O., Wickner,S. and Doyle,S.M. (2019) Hsp90 and hsp70 chaperones: collaborators in protein remodeling. *J. Biol. Chem.*, **294**, 2109–2120.
5. Mizuno,M., Ebine,S., Shounai,O., Nakajima,S., Tomomatsu,S., Ikeuchi,K., Matsuo,Y. and Inada,T. (2021) The nascent polypeptide in the 60S subunit determines the Rqc2-dependency of ribosomal quality control. *Nucleic Acids Res.*, **49**, 2102–2113.
6. Joazeiro,C.A.P. (2019) Mechanisms and functions of ribosome-associated protein quality control. *Nat. Rev. Mol. Cell Biol.*, **20**, 368–383.
7. Stein,K.C. and Frydman,J. (2019) The stop-and-go traffic regulating protein biogenesis: how translation kinetics controls proteostasis. *J. Biol. Chem.*, **294**, 2076–2084.
8. Chandrasekaran,V., Juszkiewicz,S., Choi,J., Puglisi,J.D., Brown,A., Shao,S., Ramakrishnan,V. and Hegde,R.S. (2019) Mechanism of ribosome stalling during translation of a poly(A) tail. *Nat. Struct. Mol. Biol.*, **26**, 1132–1140.
9. Deuerling,E., Gamerdinger,M. and Kreft,S.G. (2019) Chaperone interactions at the ribosome. *Cold Spring Harb. Perspect. Biol.*, **11**, a033977.
10. Liu,Y., Hu,Y., Li,X., Niu,L. and Teng,M. (2010) The crystal structure of the human nascent polypeptide-associated complex domain reveals a nucleic acid-binding region on the NACA subunit. *Biochemistry*, **49**, 2890–2896.
11. Gamerdinger,M., Kobayashi,K., Wallisch,A., Kreft,S.G., Sailer,C., Schlomer,R., Sachs,N., Jomaa,A., Stengel,F., Ban,N. *et al.* (2019) Early scanning of nascent polypeptides inside the ribosomal tunnel by NAC. *Mol. Cell*, **75**, 996–1006.
12. del Alamo,M., Hogan,D.J., Pechmann,S., Albanese,V., Brown,P.O. and Frydman,J. (2011) Defining the specificity of cotranslationally acting chaperones by systematic analysis of mRNAs associated with ribosome-nascent chain complexes. *PLoS Biol.*, **9**, e1001100.
13. Wang,S., Sakai,H. and Wiedmann,M. (1995) NAC covers ribosome-associated nascent chains thereby forming a protective environment for regions of nascent chains just emerging from the peptidyl transferase center. *J. Cell Biol.*, **130**, 519–528.
14. Martin,E.M., Jackson,M.P., Gamerdinger,M., Gense,K., Karamonos,T.K., Humes,J.R., Deuerling,E., Ashcroft,A.E. and Radford,S.E. (2018) Conformational flexibility within the nascent polypeptide-associated complex enables its interactions with structurally diverse client proteins. *J. Biol. Chem.*, **293**, 8554–8568.
15. Shen,K., Gamerdinger,M., Chan,R., Gense,K., Martin,E.M., Sachs,N., Knight,P.D., Schlomer,R., Calabrese,A.N., Stewart,K.L. *et al.* (2019) Dual role of ribosome-binding domain of NAC as a potent suppressor of protein aggregation and aging-related proteinopathies. *Mol. Cell*, **74**, 729–741.
16. Kirstein-Miles,J., Scior,A., Deuerling,E. and Morimoto,R.I. (2013) The nascent polypeptide-associated complex is a key regulator of proteostasis. *EMBO J.*, **32**, 1451–1468.
17. Fabbri,L., Chakraborty,A., Robert,C. and Vagner,S. (2021) The plasticity of mRNA translation during cancer progression and therapy resistance. *Nat. Rev. Cancer*, **21**, 558–577.
18. Holcik,M. and Sonenberg,N. (2005) Translational control in stress and apoptosis. *Nat. Rev. Mol. Cell Biol.*, **6**, 318–327.
19. Sysoev,V.O., Fischer,B., Frese,C.K., Gupta,I., Krijgsveld,J., Hentze,M.W., Castello,A. and Ephrussi,A. (2016) Global changes of the RNA-bound proteome during the maternal-to-zygotic transition in drosophila. *Nat. Commun.*, **7**, 12128.
20. Garcia-Moreno,M., Noerenberg,M., Ni,S., Jarvelin,A.I., Gonzalez-Almela,E., Lenz,C.E., Bach-Pages,M., Cox,V., Avolio,R., Davis,T. *et al.* (2019) System-wide profiling of RNA-Binding proteins uncovers key regulators of virus infection. *Mol. Cell*, **74**, 196–211.
21. Beaudoin,J.D., Novoa,E.M., Vejnar,C.E., Yartseva,V., Takacs,C.M., Kellis,M. and Giraldez,A.J. (2018) Analyses of mRNA structure dynamics identify embryonic gene regulatory programs. *Nat. Struct. Mol. Biol.*, **25**, 677–686.
22. Lista,M.J., Martins,R.P., Billant,O., Contesse,M.A., Findakly,S., Pochard,P., Daskalogianni,C., Beauvineau,C., Guetta,C., Jamin,C. *et al.* (2017) Nucleolin directly mediates epstein-barr virus immune evasion through binding to G-quadruplexes of EBNA1 mRNA. *Nat. Commun.*, **8**, 16043.
23. Yin,Y., Manoury,B. and Fahraeus,R. (2003) Self-inhibition of synthesis and antigen presentation by epstein-barr virus-encoded EBNA1. *Science*, **301**, 1371–1374.
24. Apcher,S., Daskalogianni,C., Manoury,B. and Fahraeus,R. (2010) Epstein barr virus-encoded EBNA1 interference with MHC class i antigen presentation reveals a close correlation between mRNA translation initiation and antigen presentation. *PLoS Pathog.*, **6**, e1001151.
25. Murat,P., Zhong,J., Lekieffre,L., Cowieson,N.P., Clancy,J.L., Preiss,T., Balasubramanian,S., Khanna,R. and Tellam,J. (2014) G-quadruplexes regulate epstein-barr virus-encoded nuclear antigen 1 mRNA translation. *Nat. Chem. Biol.*, **10**, 358–364.
26. Apcher,S., Komarova,A., Daskalogianni,C., Yin,Y., Malbert-Colas,L. and Fahraeus,R. (2009) mRNA translation regulation by the gly-ala repeat of epstein-barr virus nuclear antigen 1. *J. Virol.*, **83**, 1289–1298.
27. Kikin,O., D'Antonio,L. and Bagga,P.S. (2006) QGRS mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res.*, **34**, W676–W682.
28. Hon,J., Marusiak,M., Martinek,T., Kunka,A., Zendulka,J., Bednar,D. and Damborsky,J. (2021) SoluProt: prediction of soluble protein expression in escherichia coli. *Bioinformatics*, **37**, 23–28.
29. Martins,R.P., Malbert-Colas,L., Lista,M.J., Daskalogianni,C., Apcher,S., Pla,M., Findakly,S., Blondel,M. and Fahraeus,R. (2019) Nuclear processing of nascent transcripts determines synthesis of full-length proteins and antigenic peptides. *Nucleic Acids Res.*, **47**, 3086–3100.
30. Apcher,S., Daskalogianni,C., Lejeune,F., Manoury,B., Imhoos,G., Heslop,L. and Fahraeus,R. (2011) Major source of antigenic peptides for the MHC class i pathway is produced during the pioneer round of mRNA translation. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 11572–11577.
31. McQuin,C., Goodman,A., Chernyshev,V., Kamentsky,L., Cimini,B.A., Karhohs,K.W., Doan,M., Ding,L., Rafelski,S.M., Thirstrup,D. *et al.* (2018) CellProfiler 3.0: Next-generation image processing for biology. *PLoS Biol.*, **16**, e2005970.
32. Zheng,A.J., Thermou,A., Guixens Gallardo,P., Malbert-Colas,L., Daskalogianni,C., Vaudiau,N., Brohagen,P., Granzhan,A., Blondel,M., Teulade-Fichou,M.P. *et al.* (2022) The different activities of RNA G-quadruplex structures are controlled by flanking sequences. *Life Sci. Alliance*, **5**, e202101232.
33. Guo,B., Huang,J., Wu,W., Feng,D., Wang,X., Chen,Y. and Zhang,H. (2014) The nascent polypeptide-associated complex is essential for autophagic flux. *Autophagy*, **10**, 1738–1748.
34. Tovar Fernandez,M.C., Sroka,E.M., Lavigne,M., Thermou,A., Daskalogianni,C., Manoury,B., Prado Martins,R. and Fahraeus,R. (2022) Substrate-specific presentation of MHC class I-restricted antigens via autophagy pathway. *Cell. Immunol.*, **374**, 104484.
35. Meury,T., Akhouayri,O., Jafarov,T., Mandic,V. and St-Arnaud,R. (2010) Nuclear alpha NAC influences bone matrix mineralization and osteoblast maturation in vivo. *Mol. Cell. Biol.*, **30**, 43–53.
36. Lin,Z., Gasic,I., Chandrasekaran,V., Peters,N., Shao,S., Mitchison,T.J. and Hegde,R.S. (2020) TTC5 mediates autoregulation of tubulin via mRNA degradation. *Science*, **367**, 100–104.
37. Herviou,P., Le Bras,M., Dumas,L., Hieblot,C., Gilhodes,J., Cioci,G., Hugnot,J.P., Ameadan,A., Guillonneau,F., Dassi,E. *et al.* (2020) hnRNP H/F drive RNA G-quadruplex-mediated translation linked to genomic instability and therapy resistance in glioblastoma. *Nat. Commun.*, **11**, 2661.
38. Benhalevy,D., Gupta,S.K., Danan,C.H., Ghosal,S., Sun,H.W., Kazemier,H.G., Paeschke,K., Hafner,M. and Juranek,S.A. (2017) The human CCHC-type zinc finger nucleic acid-binding protein binds G-Rich elements in target mRNA coding sequences and promotes translation. *Cell Rep.*, **18**, 2979–2990.
39. Lista,M.J., Voisset,C., Contesse,M.A., Friocourt,G., Daskalogianni,C., Bihel,F., Fahraeus,R. and Blondel,M. (2015) The long-lasting love affair between the budding yeast saccharomyces cerevisiae and the epstein-barr virus. *Biotechnol. J.*, **10**, 1670–1681.
40. Kearse,M.G. and Wilusz,J.E. (2017) Non-AUG translation: a new start for protein synthesis in eukaryotes. *Genes Dev.*, **31**, 1717–1731.
41. Todd,P.K., Oh,S.Y., Krans,A., He,F., Sellier,C., Frazer,M., Renoux,A.J., Chen,K.C., Scaglione,K.M., Basrur,V. *et al.* (2013)

CGG repeat-associated translation mediates neurodegeneration in fragile x tremor ataxia syndrome. *Neuron*, **78**, 440–455.
42. Cleary,J.D., Pattamatta,A. and Ranum,L.P.W. (2018) Repeat-associated non-ATG (RAN) translation. *J. Biol. Chem.*, **293**, 16127–16141.
43. Fratta,P., Mizielinska,S., Nicoll,A.J., Zloh,M., Fisher,E.M., Parkinson,G. and Isaacs,A.M. (2012) C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Sci. Rep.*, **2**, 1016.
44. Ofer,N., Weisman-Shomer,P., Shklover,J. and Fry,M. (2009) The quadruplex r(CGG)n destabilizing cationic porphyrin TMPyP4 cooperates with hnRNPs to increase the translation efficiency of fragile x premutation mRNA. *Nucleic Acids Res.*, **37**, 2712–2722.
45. Kharel,P., Becker,G., Tsvetkov,V. and Ivanov,P. (2020) Properties and biological impact of RNA G-quadruplexes: from order to turmoil and back. *Nucleic Acids Res.*, **48**, 12534–12555.
46. Ash,P.E., Bieniek,K.F., Gendron,T.F., Caulfield,T., Lin,W.L., Dejesus-Hernandez,M., van Blitterswijk,M.M., Jansen-West,K., Paul,J.W.,3rd, Rademakers,R. *et al.* (2013) Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron*, **77**, 639–646.
47. Mori,K., Arzberger,T., Grasser,F.A., Gijselinck,I., May,S., Rentzsch,K., Weng,S.M., Schludi,M.H., van der Zee,J., Cruts,M. *et al.* (2013) Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol.*, **126**, 881–893.

Intégration de l'ensemble des résultats obtenus sur l'inhibition GAr-dépendante de la traduction de l'ARNm d'EBNA1 du virus d'Epstein-Barr

Nous avons rédigé un article de synthèse présentant l'ensemble de toutes les données obtenues par notre groupe ainsi que par l'ensemble de la communauté travaillant sur EBV. J'ai participé à l'élaboration de cet article présenté ci-dessous.

Biochimie xxx (xxxx) xxx

Contents lists available at ScienceDirect

Biochimie

journal homepage: www.elsevier.com/locate/biochi

The hide-and-seek game of the oncogenic Epstein-Barr virus-encoded EBNA1 protein with the immune system: An RNA G-quadruplex tale

Van-Trang Dinh [a], Nadège Loaëc [a], Alicia Quillévéré [a], Ronan Le Sénéchal [a], Marc Keruzoré [a], Rodrigo Prado Martins [b], Anton Granzhan [c], Marc Blondel [a, *]

[a] Univ Brest; Inserm UMR1078; Etablissement Français Du Sang (EFS) Bretagne; CHRU Brest, Hôpital Morvan, Laboratoire de Génétique Moléculaire, 22 Avenue Camille Desmoulins, F-29200 Brest, France
[b] ISP, INRAE, Univ Tours, UMR1282, Tours, Nouzilly, France
[c] Chemistry and Modelling for the Biology of Cancer (CMBC), CNRS UMR9187, Inserm U1196, Institut Curie, Université Paris Saclay, F-91405 Orsay, France

ARTICLE INFO

Article history:
Received 10 June 2023
Received in revised form
11 July 2023
Accepted 13 July 2023
Available online xxx

Handling Editor: J.L. Mergny

ABSTRACT

The Epstein-Barr virus (EBV) is the first oncogenic virus described in human. EBV infects more than 90% of the human population worldwide, but most EBV infections are asymptomatic. After the primary infection, the virus persists lifelong in the memory B cells of the infected individuals. Under certain conditions the virus can cause several human cancers, that include lymphoproliferative disorders such as Burkitt and Hodgkin lymphomas and non-lymphoid malignancies such as 100% of nasopharyngeal carcinoma and 10% of gastric cancers. Each year, about 200,000 EBV-related cancers emerge, hence accounting for at least 1% of worldwide cancers. Like all gammaherpesviruses, EBV has evolved a strategy to escape the host immune system. This strategy is mainly based on the tight control of the expression of its Epstein-Barr nuclear antigen-1 (EBNA1) protein, the EBV-encoded genome maintenance protein. Indeed, EBNA1 is essential for viral genome replication and maintenance but, at the same time, is also highly antigenic and T cells raised against EBNA1 exist in infected individuals. For this reason, EBNA1 is considered as the Achilles heel of EBV and the virus has seemingly evolved a strategy that employs the binding of nucleolin, a host cell factor, to RNA G-quadruplex (rG4) within EBNA1 mRNA to limit its expression to the minimal level required for function while minimizing immune recognition. This review recapitulates in a historical way the knowledge accumulated on EBNA1 immune evasion and discusses how this rG4-dependent mechanism can be exploited as an intervention point to unveil EBV-related cancers to the immune system.

Contents

* Corresponding author.
E-mail addresses: dinhtrangtnhp@gmail.com (V.-T. Dinh), marc.blondel@univ-brest.fr (M. Blondel).

https://doi.org/10.1016/j.biochi.2023.07.010

Please cite this article as: V.-T. Dinh, N. Loaëc, A. Quillévéré *et al.*, The hide-and-seek game of the oncogenic Epstein-Barr virus-encoded EBNA1 protein with the immune system: An RNA G-quadruplex tale, Biochimie, https://doi.org/10.1016/j.biochi.2023.07.010

1. Introduction

Although first observed at the beginning of the 20th century [1], the Burkitt lymphoma (BL) was first described in 1958 as a prevalent lymphosarcoma affecting young children in sub-Saharan Africa [2]. It was then rapidly realized that the zone of incidence of BL coincided with the one of malaria, leading to the hypothesis that malaria could be responsible, or at least a cofactor for BL. In this context, Anthony Epstein and his student Yvonne Barr made the first observation, thanks to electron microscopy, of the so-called Epstein-Barr virus (EBV) in cultured lymphoblasts from Burkitt lymphoma ([3,4]). A couple of years later, translocation and resulting activation of the *c-MYC* oncogene was detected in BL [5,6]. Altogether these observations have paved the way for the current model that EBV, *c-MYC* and malaria collaborate to cause BL (reviewed in Ref. [7]). Indeed, shortly after its discovery in 1964, the hypothesis that EBV may cause cancers has emerged, notably based on the observation that almost 100% of BL in tropical Africa were latently infected by the virus [8] and carry the viral genome as a clonal extra-chromosomal episome [9]. This way, EBV was the first oncogenic virus described in human and is one of the most efficient method for transforming and immortalizing human B-cells. EBV infects more than 90% of the human population worldwide with geographical disparities. Most EBV infections are asymptomatic even though, when the primary infection occurs during adolescence or young adulthood, it can lead to an acute infection known as infectious mononucleosis. After the primary infection, the virus persists lifelong in the memory B cells of the infected individuals. EBV has been associated to several human cancers, that include lymphoproliferative disorders such as 100% of Burkitt lymphomas and 40% of Hodgkin lymphomas, non-lymphoid malignancies such as 100% of nasopharyngeal carcinoma, 10% of gastric cancers, rare T- and NK-cells lymphomas as well as leiomyosarcoma [10–12] and, potentially, gliomas [13]. EBV is therefore found in a variety of human cancers whose localization mostly overlaps with the known tissue tropism of the virus for B and epithelial cells. According to some estimations, about 200,000 EBV-related cancers emerge per year, hence accounting for at least 1% of cancer worldwide [14]. In addition, a link between primary EBV infection and multiple sclerosis has been suspected for a long time and firmly confirmed recently [15]. For all these reasons, EBV became one of the most studied human viruses so far. As a latent virus, like all herpesviruses, EBV has evolved a strategy to escape the host immune system. This strategy is mainly based on the strict control of the expression of its Epstein-Barr nuclear antigen-1 (EBNA1) protein, the EBV-encoded genome maintenance protein of the virus [16]. Indeed, EBNA1 is considered as the Achilles heel of EBV due to the fact that it is essential for viral genome replication and maintenance but, at the same time, it is also highly antigenic and T cells raised against EBNA1 exist in infected individuals [17–22]. This review aims at recapitulating the knowledge about the EBNA1-based immune evasion of EBV and how the discovered mechanisms can define original therapeutic targets to unveil EBV-related cancers to the immune system.

2. EBNA1 is the genome maintenance protein (GMP) of EBV

EBNA1 is a multifunctional EBV-encoded protein whose primary function is to ensure EBV genome replication and maintenance, hence being considered the genome maintenance protein (GMP) of the virus [23,24]. As such, EBNA1 is the only viral protein expressed in all the latency programs of EBV [25]. In addition, EBNA1 plays a role in transcriptional activation of the EBV latency genes in latent infection [26,27]. EBNA1 is a multi-domain protein (Fig. 1a) that contains, from residue 90 to 328 in its *N*-term part, a glycine-alanine repeat (GAr) that consists of a stretch of single alanine separated by one, two or three glycine(s). This GAr domain is conserved across all EBV strains and plays a critical role in the ability of EBNA1, and thereby of EBV, to evade the immune system during the latent phase of viral infection (described in details later in this review). Importantly, a natural polymorphism in the length of GAr has been observed and the ability of GAr to allow immune evasion of EBNA1 is GAr-length-dependent: the longer GAr, the most efficient the immune evasion is. The GAr domain is flanked on each side by a glycine-arginine-rich (RGG) motif. RGG motifs (aka RGG boxes) are low complexity sequences that contain repetitions of the arginine-glycine-glycine tripeptide. RGG motifs can lead to intrinsically disordered regions [28] and have been involved in interactions with nucleic acids and various proteins [29]. The first RGG motif of EBNA1 is located from residue 33 to 64 whereas the second lies between residues 329 and 377 and both are responsible for EBNA1 association to host cell metaphase chromosomes, hence supporting EBV genome replication and maintenance [25,30,31]. The second RGG motif of EBNA1 is followed by a monopartite nuclear localization signal NLS (KRPRSPSS, residues 379 to 386) and then by a *C*-terminal DNA-binding and dimerization domain (from residue 459 to 609), which is also responsible for the attachment of the viral genome to the cellular chromatin to ensure the viral replication [16,32].-

The function of EBNA1 in genome replication and maintenance also relies on its capacity to bind to multiple sites on the origin of DNA replication (OriP) of EBV genome [33,34]. The EBV OriP contains two *cis*-acting components, a dyad symmetry (DS) element, and a family of repeats (FR) [35,36]. The DS element is a repetition of four EBNA1 binding sites divided in two by a 9-bp sequence, it is responsible for DNA replication and it has also been shown to be sufficient for EBNA1-dependent EBV replication [32,35,37]. The FR element is composed of 20 tandem copies of a 30-bp sequence containing an 18-bp palindromic element on which EBNA1 binds, followed by a 12-bp AT-rich element [38]. This FR element is not directly involved in DNA replication but, the mitotic segregation of the EBV episome requires binding of EBNA1 to its multiple recognition sites in the FR element for ensuring efficient attachment to host cell chromosomes [32,39].

In addition, thanks to a yeast-based two-hybrid screen, EBNA1 was shown to specifically interact with a host cell nuclear protein, the so-called Epstein-Barr binding protein 2 (EBP2). This protein-protein interaction appears to be important for the role of EBNA1 in EBV episome stable segregation through association to host cell

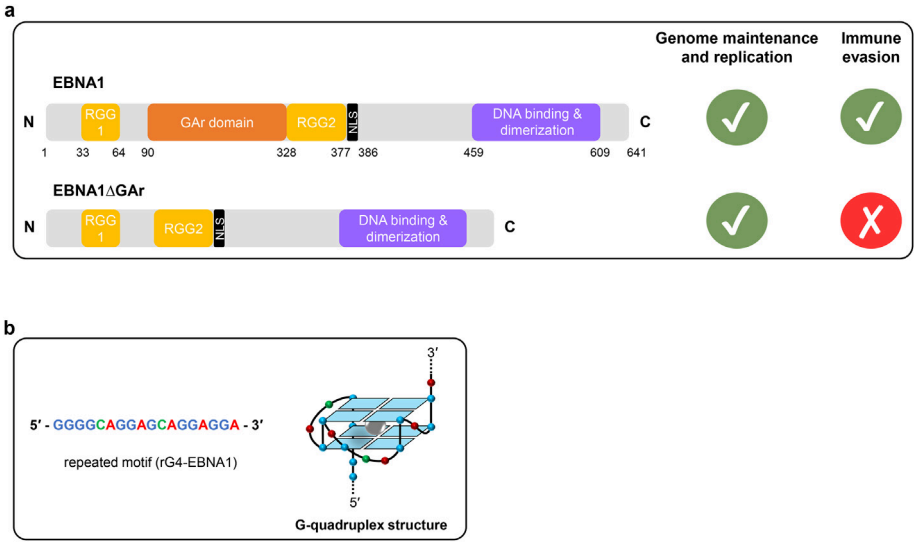


Fig. 1. Schematic representation of the EBV-encoded EBNA1 protein highlighting the essential role of the GAr domain in immune evasion of the virus. **a**. The first 32 amino acids of EBNA1 full length are not associated to a particular function. The two RGG (Arginine-Glycine-Glycine) motifs are highlighted in gold (from amino acids 33 to 64 and 329 to 377), the GAr (Glycine-Alanine repeat) domain is shown in orange (from amino acids 90 to 328), the nuclear localization signal (NLS) is located between amino acids 378 and 386 and colored in black. The DNA-binding and dimerization domain (amino acids 459 to 609) is colored in purple. As EBNA1 full length, EBNA1ΔGAr is able to fulfill its essential functions for EBV genome replication and maintenance. **b**. On the left is depicted the prototypal G-rich motif represented 13 times in the GAr-encoding sequence of EBNA1 mRNA. On the right is represented a model of the parallel two-quartet RNA G4 that this G-rich motif may form.

chromosomes in metaphase [32,40,41]. EBNA1 region that binds to EBP2 has been mapped to the second RGG motif of EBNA1 (residues 329 to 377) whose deletion prevents interaction of EBNA1 with EBP2 and consequently affects EBV genome maintenance [41]. This EBNA1-EBP2 protein-protein interaction has been furthermore reconstituted in the budding yeast *Saccharomyces cerevisiae*. This way, EBP2 was demonstrated to be crucial for the maintenance of an EBV FR element-containing plasmid maintenance [42–44]. EBNA1 is therefore a crucial protein for EBV DNA replication and episome maintenance, particularly in B cells which are actively dividing cells.

As stated above, EBNA1 is also implicated in transcriptional activation of the EBV latency genes in latent infection [26,27]. The 65–83 *N*-terminal EBNA1 residues [45,46] as well as the second RGG motif (residues 329 to 377) [47–49] are required for transcriptional activation. EBNA1 requires these two regions to activate transcription, as deletion of one or the other abolishes its ability to activate transcription [46,47]. Of note, a Δ61–83 EBNA1 mutant is fully active for replication and segregation of the viral genome, indicating that transcriptional activation is separable from the GMP role of EBNA1 [46].

3. EBNA1 is the Achilles heel of EBV and its central GAr domain is crucial for its immune evasion

Like all herpesviruses, EBV is a latent virus that evades the immune system. However, as said above, EBV possesses an Achilles heel: its virally-encoded genome maintenance protein EBNA1 (Epstein-Barr nuclear antigen 1). Indeed, being essential for replication and maintenance of the viral genome, EBNA1 is required for EBV persistence but, at the same time, highly antigenic. In addition, T cells directed against EBNA1 epitopes exist in nearly all EBV-infected individuals [17–22]. Hence, EBV seemingly evolved a mechanism by which EBNA1 limits the presentation of EBNA1-derived antigenic peptides. This ability of EBNA1 to evade the immune system has been rapidly attributed to its GAr domain which prevents EBNA1-derived antigenic peptides to be presented to the MHC class I pathway [50]. Importantly, this *in cis* effect of GAr is autonomous and transferable since, when fused to other proteins such as EBNA4, ovalbumin (OVA), or the herpes simplex virus thymidine-kinase, GAr also interfere with their antigenic presentation [50–53]. Of note, the study by Maria Masucci and collaborators represents a milestone since, not only it demonstrated the crucial role of GAr in the immune evasion of EBNA1, but also represented a proof of principle that interfering with GAr-based immune evasion of EBNA1 may represent an intervention point to unveil EBV-infected cells, in particular tumoral cells from EBV-associated cancers, to the immune system [20,50].

3.1. A historical perspective on the role of GAr in EBNA1 immune evasion: the proteasome era followed by the ribosome era

At the time the role of GAr in immune evasion of EBNA1, and therefore of EBV, was discovered, the main dogma was that most antigenic peptides for the MHC class I pathway comes from 26S proteasome-mediated degradation of full-length proteins (reviewed in Ref. [54]). In this context, several teams have observed that the GAr domain is able to efficiently inhibit *in cis* protein degradation by the 26S proteasome. This has been shown by observing that the fusion of GAr to several short half-life proteins, that include p53 and Iκβα, interferes with their 26S proteasome-dependent degradation [55–57] and this effect has been shown to be GAr-length-dependent [56,58,59]. It was thus quite logical, on this ground, to infer that most of the capacity of GAr to inhibit antigen presentation was due to its ability to inhibit 26S-dependent proteolysis (Fig. 2, **left part**). However, at the same time, some important problems associated to this model of antigenic presentation, in particular regarding the kinetics of antigenic presentation by the MHC class I pathway, have started to emerge [60–62]. In line, as early as in 1996, Jonathan Yewdell proposed an alternative hypothesis in which the most important source of antigenic peptides would mainly originate from the so-called defective ribosomal products (DRiPs) [61,62]. DRiPs correspond to abortive byproducts of the translation process that include defective, misfolded or prematurely terminated polypeptides. The first observation supporting this hypothesis is that most viral proteins are very

stable, suggesting that the degradation of full-length proteins (the so-called "retirees") by the proteasome and the consecutive presentation of derived antigenic peptides by the MHC class I would take long time, from hours to days, which is not consistent with the experimental observations that the kinetic of presentation of antigenic peptides by the MHC class I pathway following viral infection is rather rapid (from a few minutes to a few hours) and may happen even before the synthesis of mature full-length proteins. In addition, there is a strong correlation between the inception of protein expression and epitope presentation for most antigens, which represents a strong indication that antigen presentation is mostly associated to translation and not to degradation of mature full-length proteins (retirees) [60,63]. Indeed, from a number of T cell based-studies, it appears that numerous antigenic peptides from multiple viruses are rapidly produced following infection [64]. The strong kinetic link between viral gene translation and antigenic peptides production has been confirmed at the proteomic level using mass spectrometry, for the vaccinia virus [63] and recently for the influenza A virus (IAV) [65]. Hence, the idea that most viral antigenic peptides come from the degradation of full-length viral proteins did not seem anymore a viable hypothesis. Finally, the ribosome hypothesis was further supported by data showing that the inhibition of *de novo* protein synthesis rapidly reduces the antigenic presentation by MHC class I molecules [66,67]. All these findings suggest that antigenic peptides have a main source from translation itself (Fig. 2, **right part**), in particular from DRiPs [60,68]. Hence, the model in which the ability of the GAr domain of EBNA1 to enable its immune evasion, and therefore immune evasion of EBV, comes from its capacity to inhibit own degradation by the 26S proteasome started to be fragilized. In addition, several studies have found that EBNA1 protein, with or without GAr, is a very stable protein [53,55,69] like most viral proteins, although its stability may be tissue-dependent [69]. Also, the steady-state level of EBNA1 protein was strongly increased in cells expressing a version of EBNA1 deleted for its GAr domain (EBNA1ΔGAr) as compared to cells expressing the wild-type EBNA1 (EBNA1wt), which was not expected when removing a domain that would inhibit proteasomal degradation [53,56]. Finally, when expressing in the budding yeast *Saccharomyces cerevisiae* GAr domains of various lengths fused to destabilized versions of the green fluorescent protein (GFP), harboring *N*-end rule or ubiquitin fusion degradation signals, the length-dependent effect of the GAr domain on protein stability was confirmed. However, the steady-state level of the different GAr-GFP fusions was inversely correlated to their degree of stabilization: the longer the GAr domain, the stronger the stabilizing effect on the reporter protein, but the smaller its steady-state level, which is quite counterintuitive [59]. Altogether, all these indirect data suggested that the ability of GAr to promote immune evasion was not linked to its ability to inhibit the 26S proteasome-mediated proteolysis and, even if GAr is able to inhibit *in cis* proteasomal degradation when fused to unstable reporter proteins, this effect is not operant in EBNA1 which is a stable protein.

3.2. The ability of GAr to interfere with EBNA1 immune evasion is linked to its capacity to interfere in cis with the translation of its own mRNA

In a seminal study published in 2003, the use of *in vitro* expression systems and *in cellulo* metabolic labeling led to the observation that the increase in EBNA1ΔGAr, as compared to EBNA1wt, was due to an increase in EBNA1 synthesis [53]. Not only this striking result explained all the previously described results that did not fit with an effect of GAr on proteasomal degradation, but, more importantly, it was the starting point for other studies that unambiguously demonstrated that GAr does in fact inhibit *in cis* the translation of its own mRNA (EBNA1 mRNA) or of any mRNA to which it is fused, without affecting *in trans* the translation of other mRNA [51,70]. This important demonstration was fully in line with the growing role of translation and of DRIPs in the genesis of antigenic peptides for the MHC class I pathway and readily

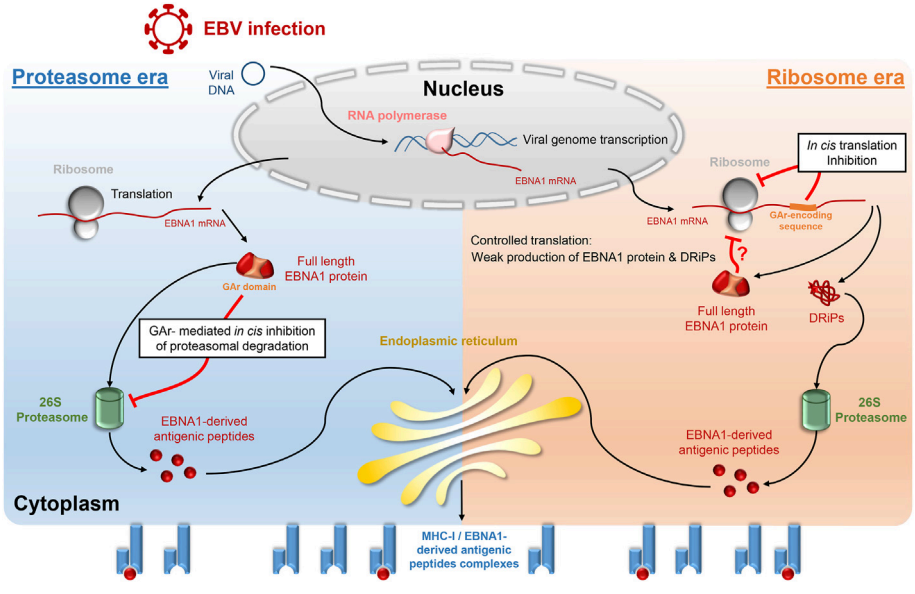


Fig. 2. The proteasome and the ribosome era parallel with the two consecutive models for EBNA1 immune evasion. On the left (**Proteasome era**) is represented the former model in which processing of the full-length EBNA1 protein by the 26S proteasome would be the main source of EBNA1-derived antigenic peptide. On this ground, the ability of GAr to inhibit *in cis* 26S proteasome-mediated proteolysis gave a plausible explanation for its role in immune evasion of EBNA1. On the right (**Ribosome era**) is represented the current model in which most of antigenic peptides would originate from byproducts of the translation (the so-called defective ribosomal products -DRIPs-) of EBNA1 mRNA. Hence, in this model, the ability of GAr to limit *in cis* the translation of its own mRNA fully explains its inhibitory effect on antigen presentation.

explained the effect of GAr on immune evasion of EBNA1, and therefore of EBV [71]. These important results have then been confirmed [69] and converged to the model in which the endogenous presentation of EBNA1-derived antigenic peptides is dependent on the newly synthesized EBNA1 protein rather than on the long-lived stable EBNA1 protein, hence highlighting the inhibitory role of GAr on translation as the main mechanism explaining its inhibitory effect on antigen presentation (**Fig. 2**, **right part**). Importantly, the GAr-encoding sequence does not represent *per se* a hurdle to mRNA translation by the ribosome since the insertion of the *c-Myc* internal ribosome entry site (IRES) in the 5' untranslated transcribed region (UTR) of EBNA1 or GAr-OVA fusion-coding mRNA abolishes GAr-based inhibition of both translation and antigen presentation [51,70]. This result also indicated that the GAr-inhibitory effect on translation requires cap-dependent translation initiation factors.

3.3. Tools to monitor the GAr inhibitory effect on translation: setup of a yeast-based assay amenable to high throughput drug and genetic screenings

Once the GAr inhibitory effect on translation firmly established, it rapidly appeared as a relevant intervention point to unveil EBV-infected cells, in particular tumor cells of cancers linked to EBV, to the immune system, hence the need for robust cell-based systems allowing both drug and genetic screenings. For various reasons, the budding yeast *Saccharomyces cerevisiae* has emerged as a relevant model system: notably, the study by the Masucci group already discussed above and based on the expression in yeast of destabilized GFP fused to GAr domains of various lengths indirectly suggested that the length-dependent inhibitory effect of GAr on translation is operant in yeast [59]. In addition, the budding yeast *S. cerevisiae* is a eukaryote as easy to handle as a prokaryote and is most certainly the model system for which the most molecular tools are available. Finally, for decades, yeast has been harnessed by biomedical research, notably on the Epstein-Barr virus itself (reviewed in Ref. [44]) and on virus immune evasion [72], and widely used for modeling various pathophysiological mechanisms responsible for a large variety of human disorders [73,74]. A number of these yeast-based models have been successfully used for performing large-scale pharmacological screening that aimed at identifying candidate drugs or chemical probes (for examples see Ref. [75] or [76]) as well as genetic screenings that focused on deciphering novel key factors involved in the considered diseases (for examples see Ref. [77] or [78]). Therefore, a yeast model for monitoring the GAr-based inhibition of the translation of its own mRNA *in cis* has been set up [79]. In this model (described in Fig. 3a), the yeast protein Ade2 belonging to adenine biosynthesis pathway was used as a reporter protein since: (i) like EBNA1 it is a stable protein, and (ii) a convenient colorimetric system allows to easily (without any treatment) determine its steady-state level in living yeast cells. Indeed, whereas yeast cells that express Ade2 at a functional level form white colonies, yeast cells that do not express Ade2 (for example cells deleted for the *ADE2* gene -termed *ade2Δ*-) readily form red colonies on rich YPD medium. In addition, any intermediate level of Ade2 leads to pink colonies whose color intensity is inversely proportional to the level of Ade2 expressed. Hence, this simple reporter system allows to easily monitor subtle changes in the level of Ade2. For this reason, and based on the fact that GAr-based inhibition of translation is probably operant in yeast, fusions of GAr domains of various lengths (21, 43 or 235 amino acids) to Ade2 were tested in yeast and led to the formal demonstration that GAr inhibits *in cis* and in a length-dependent manner the translation of its own mRNA in yeast, without any significant effect on mRNA level nor on protein stability (which was expected given that Ade2 is a very stable protein). Hence, the 43GAr-Ade2 fusion leads to pink yeasts, thereby allowing to easily monitor the effect of GAr and serving as a convenient readout for various screening aiming at identifying compounds or genes able to interfere with its *in cis* inhibitory effect on translation: modifiers that counteract the GAr effect on translation would lead to whiter yeast colonies, whereas those exacerbating the GAr effect would lead to redder yeast colonies. This yeast-based assay has first been used for drug screening using a drop test previously developed [75,80] (Fig. 3b, **left part**) and led to the identification of doxorubicin as a compound specifically counteracting the inhibitory effect of GAr on the expression of GAr-Ade2. This effect was GAr-dependent since doxorubicin had no effect on the expression of Ade2. The yeast assay was then definitively validated when doxorubicin was not only shown to stimulate EBNA1 expression in a GAr-dependent manner in human cells, but also to overcome the GAr-dependent restriction of MHC class I antigen presentation [79]. These results also suggested that, should host cell factors involved in GAr-based translation inhibition exist, they would be functionally conserved from yeast to human. This conclusion was the ground for the use of the yeast-based assay to perform a genetic screening (Fig. 3b, **right part**) that aimed at identifying those potential host factors. The same readout (changes in the pink color of the 43GAr-Ade2 yeast strain) was used and the effect of overexpressing every single yeast gene was thus tested and led to the identification of the yeast *NSR1* gene [78,81] which encodes the Nsr1 protein, the yeast ortholog of the human nucleolin (NCL). When overexpressed in yeast, Nsr1 or NCL led to a darker color of the 43GAr-Ade2 yeast strain due to the exacerbation of the GAr-based inhibition of translation, whereas the 43GAr-Ade2 yeast strain deleted for the *NSR1* gene form white colonies that express a higher level of the 43GAr-Ade2 fusion protein. Similar results were obtained when overexpressing or downregulating human NCL in human cells expressing EBNA1 of GAr-OVA fusions [78]. Altogether these results have highlighted nucleolin (NCL) as the first identified host cell factor required for GAr-based inhibition of translation, the mechanism at the basis of EBNA1 immune evasion. The question that came next was the precise role of NCL in GAr-based inhibition of translation.

3.4. Crucial role of RNA G-quadruplex (rG4) and of their interaction with NCL in GAr-based inhibition of translation and of antigen presentation

G-quadruplexes (G4) are non-canonical secondary structures that may fold in guanine-rich DNA or RNA. At the basis of G4 are G-quartets which consist of the planar arrangement of four guanines connected by Hoogsteen hydrogen bonds and stabilized by a central cation, most often potassium (K^+). Then, the stacking of at least two G-quartets forms a G4. G4 can thus represent kind of "knots" in nucleic acids that may interfere with the progression of processive enzymes such as polymerases. They can also constitute binding platforms for a number of factors such as RNA-binding proteins for RNA G-quadruplex (rG4). Hence, G4 may have biological regulatory roles and have notably been shown to affect transcription, splicing or translation [82–86]. Importantly, the central regions of most gamma herpesviruses genome maintenance proteins are highly homologous in terms of nucleotide sequence and are purine enriched [87]. At about the time NCL was identified as the first host factor involved in GAr effect on translation, the ability of the GAr-encoding sequence of EBNA1 mRNA to form a cluster of up to thirteen rG4 was demonstrated and the possibility that these rG4 could impede ribosome progression has been proposed [88]. The prototypal repeated sequence which forms the most probable rG4 in GAr-encoding sequence of EBNA1

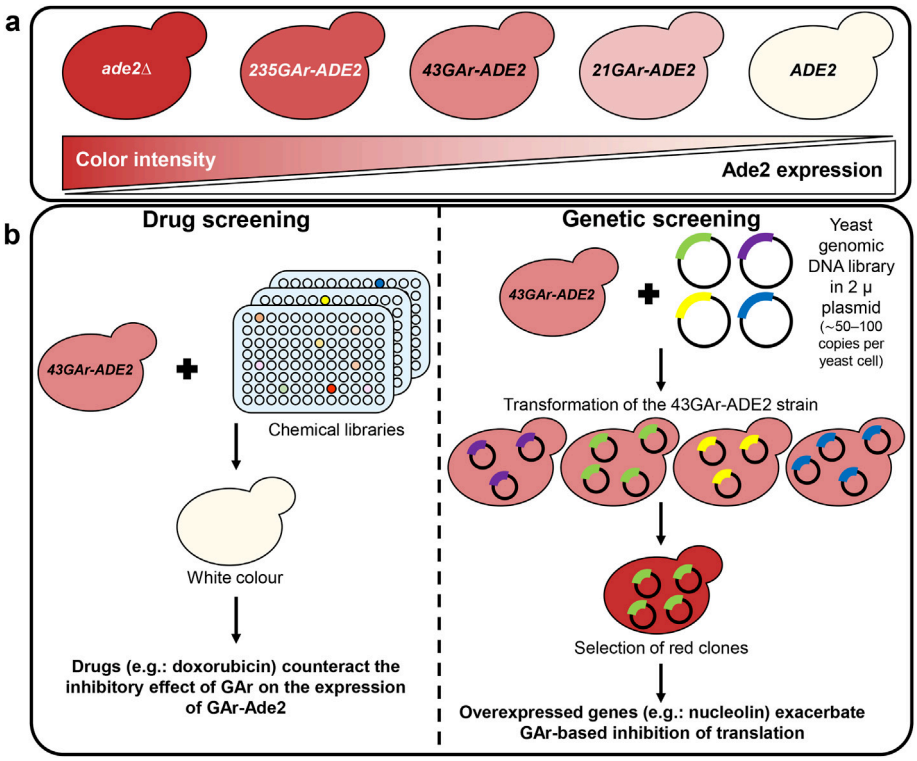


Fig. 3. Rationale of the yeast-based assay for GAr-based inhibition of translation and its use for both drug and genetic screenings. **a**. Rationale of the use of the Ade2 reporter protein in yeast. When Ade2 is expressed at a functional level yeast cells are white. When not expressed (*ade2Δ* cells), yeast cells are red. Any intermediate level of Ade2 leads to pink colonies whose color intensity is inversely proportional to the level of Ade2 expression. This allows to easily monitor the GAr-inhibitory effect on translation in yeast. **b**. Examples of use of the Ade2 reporter system for drug screening (left panel) aiming at identifying compounds that counteract the *in cis* inhibitory effect of GAr on translation, or genetic screening (right panel) to identify host genes involved in GAr-based inhibition of translation.

mRNA is represented in Fig. 1b. In line, in the same study, it was shown that destabilization of these rG4 by antisense oligonucleotides enhances both EBNA1 translation and presentation of EBNA1-derived antigenic peptides. However, as stated above, the fact that driving the translation initiation of a GAr-OVA mRNA by the *c-Myc* internal ribosome entry site (IRES) abolishes GAr-based inhibition of both translation and antigen presentation suggests that the GAr-encoding sequence does not represent *per se* a hurdle to the ribosome progression [51,70]. In addition, at a low concentration (less than 1 μM), EBNA1 mRNA was translated *in vitro* as efficiently as the EBNA1ΔGAr mRNA [10]. In line, GAr-based inhibition of translation is not operant in a yeast *nsr1Δ* strain deleted for nucleolin [78,81] also suggesting that the series of rG4 that may assemble in GAr-encoding RNA are not *per se* able to significantly impede ribosome progression. Importantly, human nucleolin (NCL) has long been known to interact with G-quadruplexes [89,90], hence opening the possibility that NCL interacts with rG4 of the EBNA1 mRNA and that this interaction is crucial for GAr-based inhibition of EBNA1 translation and antigenic peptides presentation. This potential NCL/rG4 of EBNA1 mRNA interaction has been evidenced using both RNA pulldown [78,91] and an adaptation of the proximity ligation assay (PLA) for monitoring RNA/protein interactions [78,92]. Following these various observations, the benchmark G4-ligand PhenDC3 was tested and shown to interfere with NCL binding to rG4 of EBNA1 mRNA, thereby relieving GAr-based inhibition of both translation and antigenic peptide presentation by the MHC class I [78]. This result has motivated the screening of a series of 20 cationic bis(acylhydrazone) derivatives and led to the identification of two original G4-ligands, PyDH2 and PhenDH2, more active and less cytotoxic than PhenDC3 and interfering with EBV immune evasion also by disrupting the interaction between NCL and rG4 of EBNA1 mRNA [93]. Of note, doxorubicin, the compound isolated in the yeast assay that models GAr-based inhibition of translation described above [79] was recently shown to be a G4-ligand [94,95], thus giving a plausible explanation for its ability to interfere with EBNA1 immune evasion. In contrast, pyridostatin (PDS), another benchmark G4-ligand, has been shown to decrease EBNA1 expression in an *in vitro* transcription-translation coupled system [88]. Although this result was not verified *in cellulo* where PDS had no effect on EBNA1 expression [78], which may due to the possibility that PDS affects transcription rather than translation in the *in vitro* transcription-translation coupled system, this suggests at least three different possible mechanisms of action for G4-ligands. Indeed, the binding of G4-ligands on G4 may stabilize them and therefore increase the recruitment of G4-binding proteins, or interfere with the binding of cellular partners, either by direct competition or by remodelling G4 structure. Hence, targeting G4 structures by G4-ligands is not a straightforward process as the precise effect of the tested compounds may be difficult to anticipate. Of note, BRACO19, another G4-ligand, has been described to interfere with EBNA1-dependent stimulation of EBV DNA replication [96]. Importantly, targeting G4 is not unique to EBV as the genome of many viruses contain G4 [97] which may represent intervention points and therapeutic opportunities [98,99]. Of note, NCL has also been found to interact with G4 in the LTR promoter of human immunodeficiency virus-1 (HIV-1), thereby stabilizing them, which in turn silences HIV-1 viral transcription [100]. NCL has also been involved in suppression of hepatitis C virus replication by binding to rG4 of the viral genome

[101]. Another example is the herpes simplex virus-1 (HSV-1) whose genome contains multiple clusters of G4 that can be targeted by G4-ligands for antiviral strategies [102,103].

Importantly, it has been recently observed that insertion of the *c-Myc* IRES to the 5'UTR of GAr-OVA RNA prevents the interaction between NCL and rG4 of the GAr-encoding RNA, thus readily explaining its ability to abolish GAr-based inhibition of both translation and antigen presentation [104]. In addition, this result further confirms the importance of the NCL/rG4 of EBNA1 mRNA interaction for EBNA1 immune evasion.

Finally, the mRNA encoding LANA1, the genome maintenance protein of the Kaposi's sarcoma-associated herpesvirus (KHSV), another oncogenic herpesvirus that also evades the immune system, has been proposed to also form rG4 in a central region presenting sequence similarity with the one encoding GAr, readily suggesting rG4 could represent a common feature that may allow immune evasion of various stealthy oncoviruses [88,105]. The existence and the crucial role of the rG4 of LANA1 mRNA in immune evasion of KHSV have recently been shown [106,107].

3.5. The C-terminal RGG motif of NCL is crucial for its role in GAr-based inhibition of both translation and antigen presentation

The importance of the various domains of NCL protein for GAr-based inhibition of translation has been tested. NCL is composed of three domains: (i) an *N*-terminal domain containing acidic stretches; (ii) a central region consisting of four tandem RNA recognition motifs (RRM1 to RRM4); and (iii) a *C*-terminal RGG motif (Fig. 4a). Nucleolin homologues share the same organization, with some variation in the number of RRMs. Hence, the yeast nucleolin Nsr1 contains only two RRMs (RRM1 & 2) but does possess the acidic *N*-terminal domain as well as the *C*-terminal RGG motif. Importantly, and contrary to NCL in human, Nsr1 is not essential as yeast cells deleted for the *NSR1* gene (*nsr1Δ* strains), although growing rather poorly, are viable [108]. In addition, human nucleolin NCL can complement the effect of the loss of yeast Nsr1 on GAr-based inhibition of translation in yeast [78,81]. For these reasons, an approach combining the use of the yeast model for GAr-based inhibition of translation in which the *NSR1* gene has been deleted (*43GAr-ADE2 nsr1Δ* strain) and the expression of either Nsr1 or NCL, or of parts of them, has been exploited to determine which domain of nucleolin is involved in this process and has identified the *C*-terminal RGG motif of both yeast and human nucleolin. Then, this RGG motif has been shown to interact with the rG4 of EBNA1 mRNA, thus being essential for EBNA1 ability to self-limit its expression [109].

One or two methyl groups can be attached to the nitrogen atoms of arginine, a reaction termed arginine methylation. This post-translational modification is performed by protein arginine methyltransferases (PRMTs) [110]. These enzymes catalyze the transfer of a methyl group from *S*-adenosylmethionine (SAM) to the guanidino nitrogen atoms of arginine. In human there are nine PRMTs that are distributed into three groups: type I, type II and type III which differ in the pattern of arginine methylation [111]. Arginine methylation happens in a variety of protein domains, but the RGG motifs are the most commonly reported [112,113] and, in *S. cerevisiae*, the majority of arginine methylation was observed in RGG motifs [114]. For these reasons, the potential role of PRMTs in GAr-based inhibition of translation and antigen presentation has been tested. Hence, using drug-based or siRNA-mediated inhibition of PRMTs, it has been shown that type I PRMTs are required for the crucial role of nucleolin in GAr-based inhibition of translation and antigen presentation [109]. Next, the role of methylation of the arginines of the RGG motif of both yeast and human nucleolins has been directly evaluated by changing the eight arginines of the RGG motif of Nsr1, or the ten arginines of the RGG motif of NCL, by either alanines (to prevent methylation), lysines (to prevent methylation while maintaining the positive charge), or phenylalanines as mimics of methylated arginines. Indeed, phenylalanine has been shown to mimic methylated arginine as it carries a bulky hydrophobic moiety [115,116]. Changing arginines into alanines or lysines suppress the interaction of NCL and Nsr1 with rG4 of EBNA1 mRNA, and thereby GAr-based inhibition of translation, whereas replacement of arginines by phenylalanines maintains both these phenomena [109]. Altogether, these results indicate that methylation of the arginines of the *C*-terminal RGG motif of nucleolin is necessary for its interaction with rG4 of the GAr-encoding sequence of EBNA1 mRNA, an interaction which is required for GAr-based inhibition of translation and, as a consequence, for EBNA1 immune evasion. Following these observations, two possible models for the role of methylation of the arginines of the *C*-terminal RGG motif of nucleolin could be envisioned: either this methylation directly promotes the interaction of NCL with rG4 of EBNA1 mRNA, or this methylation may interfere with the ability of the RGG motif of nucleolin to interact, and thus to be sequestered, by some of its partners (that can be proteins, DNA or RNA). In favor of this second possibility is the recent observation that Scd6, a yeast protein also acting in translation inhibition, self-associates *via* its RGG motif and that this self-association abolishes its ability to inhibit translation. Importantly, this self-association is prevented by arginine methylation catalyzed by Hmt1, the main yeast type I PRMT which is thus required for Scd6-mediated translation inhibition [28]. *In vitro* experiments based on arginine methylation of bacterially produced recombinant NCL, or on either methylated or non-methylated RGG peptides have shown that methylation of the arginines of the *C*-terminal RGG motif of NCL does not significantly change its ability to interact with rG4 of EBNA1 mRNA [109]. Hence, it seems that the RGG motif of NCL may be trapped by an unknown partner, thereby preventing its interaction with rG4 of EBNA1 mRNA and, upon type I PRMTs-mediated arginine methylation, the RGG motif of NCL is released and thus free to interact with rG4 of EBNA1 mRNA, hence leading to immune evasion of EBNA1 (Fig. 4b, **left part**). Of note, the unknown partner of the RGG motif of NCL could be L3, a ribosomal protein that has been shown to interact with the RGG motif of NCL [117].

Finally, all these results reinforce the role of NCL in the persistence of EBV in its host. Indeed, a role of NCL in EBV episome maintenance and transcription has also been described. This role involves a direct interaction of NCL with the EBNA1 protein [118]. Hence, NCL appears to be involved in both EBV episome maintenance and transcription on the one hand, and the GAr-dependent self-limitation of EBNA1 (the genome maintenance protein of EBV) expression on the other. From the virus point of view, it makes sense to have the same host cell protein regulating these two key aspects of EBV's latency. Indeed, if NCL is weakly expressed in the host of the infection, as a result, the maintenance and transcription of EBV episome will be compromised but, as a direct consequence of the role of NCL in GAr-based limitation of EBNA1 expression, EBNA1 mRNA will be more efficiently translated, which may compensate for its reduced level, thereby allowing the maintenance of EBV genome. On the contrary, if NCL is strongly expressed, then EBV episome will be efficiently maintained and transcribed, hence leading to a high level of EBNA1 mRNA, but then, due to the role of NCL in self-limitation of EBNA1 expression, its translation will be further inhibited, thus limiting the level of EBNA1 protein and thereby its detection by the immune system. Importantly, the role of NCL in EBV episome maintenance and transcription involves its interaction with EBNA1 protein, whereas its role in immune evasion of EBV involves its interaction with EBNA1 mRNA. In addition, both roles of NCL involve different domains: NCL role in

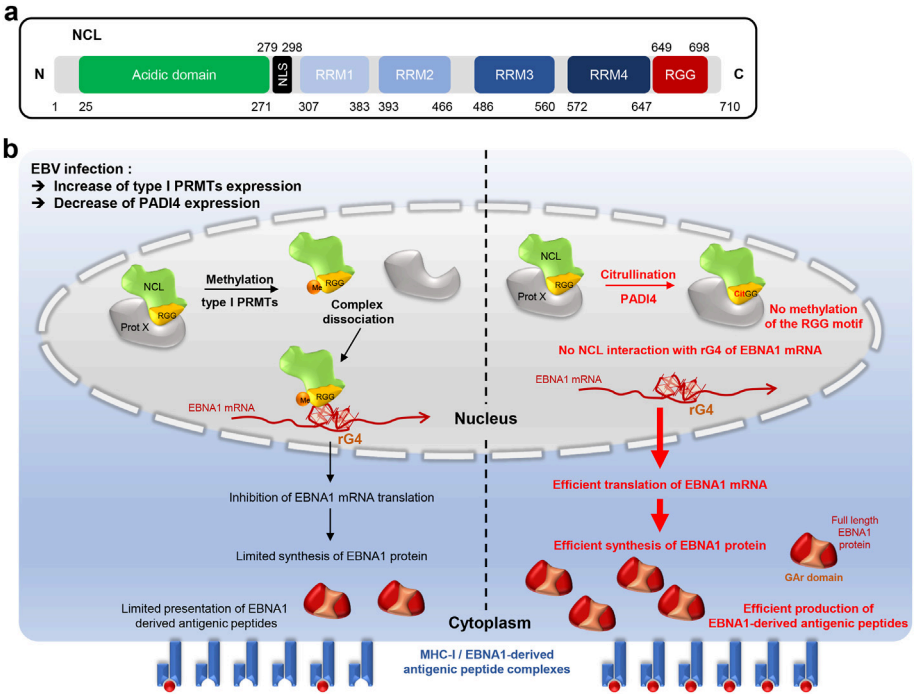


Fig. 4. Putative model for the role of NCL and of the arginine methylation and citrullination of its *C*-terminal RGG motif in EBNA1 immune evasion. **a**. Schematic representation of human nucleolin (NCL). **b**. In this model, NCL is trapped by an unknown factor (Protein X) which interacts with the unmethylated *C*-terminal RGG motif of NCL. Upon type I PRMTs-mediated methylation of the arginines of this RGG motif (left), the NCL/Protein X complex is dissociated. Hence, the RGG motif of NCL becomes available to interact with the RNA G-quadruplexes (rG4) of EBNA1 mRNA, which is required for GAr-based limitation of translation, a phenomenon at the basis of EBNA1 immune evasion. Citrullination (on the right), by preventing arginine methylation, interferes with type I PRMTs-mediated methylation of the *C*-terminal RGG motif of NCL, hence interfering with GAr-based inhibition of translation and presentation of EBNA1-derived antigenic peptides.

EBNA1 immune evasion involves its *C*-terminal RGG motif (which interacts with rG4 that form in the GAr-encoding sequence of EBNA1 mRNA), whereas its role in episome maintenance and transcription involves its central RNA recognition motifs (RRMs, which interact with the *N*-terminal first 100 amino acids of EBNA1 protein) [118]. Therefore, targeting NCL / EBNA1 mRNA interaction should specifically affect EBNA1 immune evasion.

4. Therapeutic perspectives: potential intervention points and therapeutic avenues

Based on the current model of GAr-based inhibition of both EBNA1 expression and antigen presentation, two intervention points may be envisioned to unveil EBV-infected cells, in particular tumor cells of EBV-related cancers, to the immune system: (i) the interaction between rG4 of EBNA1 mRNA and the *C*-terminal RGG motif of NCL, and (ii) type I PRMTs (Fig. 5). Importantly, in both cases proof-of-principles exist as G4 ligands that interfere with the rG4/NCL interaction on the one hand, and drugs targeting type I PRMTs on the other, interfere with the GAr-based inhibition of translation and antigen presentation [78,93,109]. Of note, for both these therapeutic avenues, selectivity may be an issue (as for most of the drugs). Indeed, type I PRMTs may appear as relatively non-specific therapeutic targets as their inhibition should impact on the methylation of many proteins. Importantly, this limit is probably encountered with most of drugs targeting enzymes like, for example, drugs inhibiting protein kinases or phosphatases. Of note, this potential lack of specificity has not prevented the development to the clinics of a number of kinase inhibitors [119], in particular in the field of cancer, and several inhibitors of PRMTs are currently being evaluated in clinical trials. Also, given that RGG motifs are the main substrates of PRMTs and also that, at concentration ranges where they affect both GAr-based translation inhibition and the interaction between NCL and rG4 of EBNA1 mRNA, type I inhibitors did not exhibit significant toxicity, the lack of specificity may not be a hurdle for their potential developments to the clinics [109]. As for G4 ligands, an inherent limitation in their development to the clinics may also be their potential lack of specificity given that, in addition to the numerous potential rG4 [120], there are hundreds of thousands potential DNA G4 (dG4) that may form in the human genome [121,122]. However, G4 in the RNA context are often in a dynamic equilibrium with other structural isoforms, and even low doses of G4 ligands may have a profound impact on this equilibrium, preventing or leading to the recruitment of the corresponding structure-specific binding factors (nucleolin in the case of the rG4 of EBNA1 mRNA). Another important point is that, besides the G-quartets, which are the constant elements constitutive of G4, there are other structural characteristics which vary from one sequence to another: the loops and the flanking regions. These elements ensure that each G4 is, in principle, unique within the whole genome or transcriptome and, as such, could be exploited as key targets to gain specificity through, for instance, the coupling of complementary oligonucleotides [123,124] or minor-groove binders [125] to G4-ligands. Hence, it is not excluded that some specificity toward particular G4, notably rG4, may be reached and various methods to assess G4 ligands specificity have been described [126,127]. Finally, the combined use of G4-ligands and inhibitors of type I PRMTs may also help to solve the specificity issue by potentially allowing the use of much reduced doses of both types of drugs.

The citrullination of arginines of the RGG motif of NCL may also be involved in the GAr-based mechanism of inhibition of both EBNA1 translation and antigen presentation. Indeed, deamination, also called citrullination, is the process of conversion of arginine

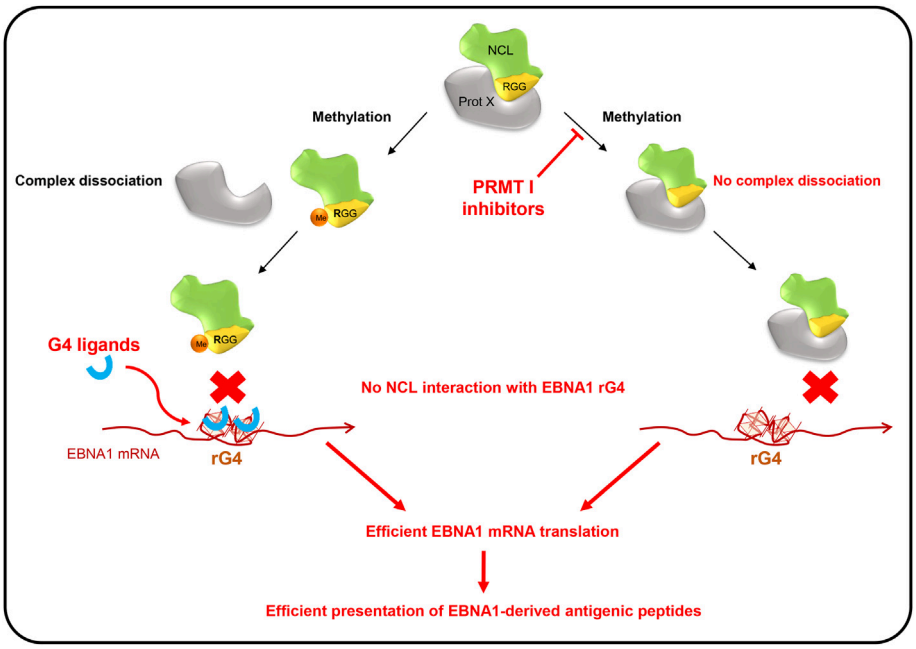


Fig. 5. Possible intervention points to unveil EBNA1 to the immune system. Based on the model presented in Fig. 4, at least to intervention points could be exploited to unveil EBV-infected cells (in particular tumor cells of EBV-related cancers) to the immune system: G4 ligands such as PhenDC3 that interfere with the binding of NCL on rG4 of EBNA1 mRNA (left) and compounds inhibiting type I PRMTs (right). By distinct modes of action, both types of drugs interfere with the binding of NCL on rG4 of EBNA1 mRNA, and thereby with the ability of GAr to self-limit the translation of its own mRNA. This, in turn, relieves the limitation of the production of EBNA1-derived antigenic peptides, ultimately interfering with immune evasion of EBNA1 and thus of EBV.

into citrulline through enzymatic activation of peptidyl arginine deiminases (PADIs), notably PADI4 (aka PAD4), the main peptidyl arginine deiminase in human, thus preventing their methylation by PRMTs [128,129]. Citrullination of RGG motifs was observed in several RGG-containing proteins, including NCL [130]. Interestingly, it has been reported that EBV infection activates several PRMTs, especially PRMT1, the main type I PRMTs, and at the same time inactivates PADI4 [131]. Hence, a possibility exists that EBV infection, by activating PRMT1 as well as inactivating PADI4, favours arginine methylation of the NCL *C*-terminal RGG motif, thus promoting the interaction of this protein with the G4s of EBNA1 mRNA, leading to limitation of its translation, and therefore to the evasion of EBNA1/EBV from the immune system (Fig. 4).

Importantly, in addition to the GAr-encoding mRNA sequence, the GAr peptide could also participate to the inhibition of its own expression. For example, the addition of antibodies directed against the GAr domain can enhance the translation of GAr-carrying mRNAs, which supports the idea that the nascent GAr polypeptides could participate to the translation inhibition of its own mRNA [51,70]. In addition, the nascent GAr peptides have recently been shown to interact with the nascent polypeptide-associated complex alpha (NACA) and displace it from the ribosome [132]. Endogenous NACA, in addition to bind to the GAr peptides, also interacts with GAr mRNA sequence as well as with NCL in the cell nucleus, thus facilitating NCL-EBNA1 mRNA interaction, thereby favoring mRNA translation inhibition *in cis*. Importantly, siRNA-mediated silencing of NACA expression interferes with the interaction between NCL and the rG4 of GAr mRNA and thus with GAr-based inhibition of both translation and antigenic peptides production [132]. Also, as stated above, the EBNA1 protein itself contains two RGG motifs: one upstream and one downstream of the GAr domain, and it has been shown that EBNA1 could bind to its own mRNA through its two RGG motifs [133] and also that these two RGG motifs could bind to rG4 [96]. Moreover, these RGG domains are encoded by G-rich sequences in the EBNA1 mRNA that also have potential to form G-quadruplexes. These observations open two non-mutually exclusive possibilities that would implicate the RGG domains of EBNA1 in the mechanism of regulation of its expression. Firstly, like NCL, the EBNA1 protein could bind to the rG4 of its own mRNA through its two RGG motifs, and thus participate in the inhibition of its translation. The other possibility would be that, like the GAr-encoding sequence, the EBNA1 mRNA sequences encoding the two RGG motifs of EBNA1 would be capable of forming rG4 that may also recruit NCL and/or EBNA1 protein(s) and would therefore participate to the inhibition of EBNA1 translation as well as the production of EBNA1-derived antigenic peptides, leading thus to the immune evasion of EBNA1/EBV. Of note, in both cases, type I PRMTs and PADIs may play a regulatory role. Interestingly, NCL and EBNA1 proteins have been reported to interact [118], hence opening the possibility of a cooperative binding on rG4 of EBNA1 mRNA, and thereby of a cooperative inhibitory effect of these two proteins on EBNA1 mRNA translation. All these possibilities clearly deserve detailed investigations.

Importantly, in addition to its crucial role in EBV genome replication and maintenance and in the virus immune evasion, EBNA1 has also been linked to cell transformation and, if the mechanism underlying its oncogenic activity is still unclear [16], it may involve the *c*-Myc oncogene. Indeed, the EμEBNA1 mice that express EBNA1 in their B-cells thanks to the immunoglobulin heavy chain intronic enhancer (Eμ) develop and succumb to B-cell lymphoma [134]. Importantly, the B cell tumours of these mice are monoclonal and similar to those associated to transgenic *c*-Myc expression (Eμmyc transgenic mice [135,136]). However, contrary to the characteristic translocations observed in the *c-MYC* locus which are linked to the deregulation of its expression, in B cells tumours arising from EμEBNA1 transgenic mice no 5' rearrangements were detected in the murine *c-MYC* locus, whereas the expression of *c*-Myc was also deregulated [134]. Hence, EBNA1 is able to induce overexpression of the *c-MYC* oncogene and, as such,

may participate to the oncogenicity of EBV. Interestingly, recently a direct link between EBNA1 and *C*-Myc expression involving the GAr domain of EBNA1 has been described [137]. In this model, the *in cis* GAr-mediated mRNA translation inhibition induces a translational stress that in turn increases the E2F1 transcription factor expression *via* a phosphoinositide-3-kinase-δ- (PI3Kδ-) dependent signalling. The resulting increase in the level of E2F1 induces expression of the *c-MYC* oncogne. In support of this model are the observations that drug-based or siRNA-mediated specific inhibition of PI3Kδ abrogate EBNA1-dependent induction of both E2F1 and *c*-Myc. In addition, a multi-omics approach has recently revealed that EBNA1 binds directly to enhancer elements in a subset of genes that notably include key reregulators of purine metabolism that are crucial for proliferation of EBV-immortalized B-cells [138]. Hence, in addition to be the GMP of EBV and to play a central role in the ability of the virus to evade the immune system, EBNA1 participate to the oncogenicity of EBV. Interestingly, both oncogenicity and immune evasion of EBV depend on the GAr domain of EBNA1. For this reason, pharmacological means that would result in exacerbating the GAr inhibitory effect on translation, by further decreasing EBNA1 level, may also be of therapeutic value for EBV-related cancers. In line, EBNA1 protein itself has been exploited as an intervention point and various EBNA1 inhibitors have been developed with potential to treat EBV-related cancers [139,140].

Finally, beyond cancer, EBV, and more particularly EBNA1, have been involved in other diseases, notably multiple sclerosis (MS). Indeed, the long-time suspected link between EBV and MS has recently been firmly confirmed in a large-scale study in which the authors analysed serum EBV antibodies from individuals who developed MS among a cohort of more than 10 million young adults in a more than 20 years follow-up, and have reported a 32-fold increase in the risk of developing MS after EBV infection [15]. In parallel, a role for EBNA1 in MS has been proposed, notably *via* molecular mimicry leading to the development of cross-reactive antibodies and T cells [141,142]. In line, a vaccine against EBNA1 appears a relevant therapeutic option to treat EBV-related disorders, in particular EBV-related cancers and MS, and several experimental vaccines have been developed against EBV, including against EBNA1 [143–145]. Notably, vaccination against EBNA1 has been shown efficient in inducing global $CD4^+$ and $CD8^+$ T cell responses able to protect against EBV antigen-expressing lymphomas [144,146].

5. Concluding remark

Beyond the clear interest of all the studies on EBNA1 biological roles and immune evasion for defining intervention points and novel therapeutic avenues to treat EBV-related diseases, in particular EBV-related malignancies, they also represent another illustration that “viruses are the best cell biologists” in that their study quite often leads to seminal discovery in cell biology.

Acknowledgments

This work was supported by the following grant agencies: ‘La Ligue contre le cancer CSIRGO’ and ‘Fondation pour l’Avenir’ to MB, ‘Agence Nationale de la Recherche’ (ANR-17-CE07-0004-01, to AG), French Ministry of Higher Education, Research and Innovation (PhD fellowships to RLS), ‘La Ligue contre le cancer Bretagne’ and ‘Région Bretagne’ (PhD fellowship to VTD), INCa PLBIO (PLBIO16-225) to MB and ‘Fondation pour la Recherche Médicale’ (DCM20181039571) to MB.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.biochi.2023.07.010.

References

[1] J.N. Davies, et al., Cancer in an african community, 1897–1956. An analysis of the records of mengo hospital, kampala, Uganda. I, Br. Med. J. 1 (5378) (1964) 259–264.
[2] D. Burkitt, A sarcoma involving the jaws in African children, Br. J. Surg. 46 (197) (1958) 218–223.
[3] M.A. Epstein, B.G. Achong, Y.M. Barr, Virus particles in cultured lymphoblasts from burkitt's lymphoma, Lancet 1 (7335) (1964) 702–703.
[4] M.A. Epstein, Y.M. Barr, Cultivation in vitro of human lymphoblasts from burkitt's malignant lymphoma, Lancet 1 (7327) (1964) 252–253.
[5] G. Klein, Specific chromosomal translocations and the genesis of B-cell-derived tumors in mice and men, Cell 32 (2) (1983) 311–315.
[6] G. Manolov, Y. Manolova, Marker band in one chromosome 14 from Burkitt lymphomas, Nature 237 (5349) (1972) 33–34.
[7] D.A. Thorley-Lawson, M.J. Allday, The curious case of the tumour virus: 50 years of Burkitt's lymphoma, Nat. Rev. Microbiol. 6 (12) (2008) 913–924.
[8] I. Magrath, The pathogenesis of Burkitt's lymphoma, Adv. Cancer Res. 55 (1990) 133–270.
[9] T. Lindahl, et al., Covalently closed circular duplex DNA of Epstein-Barr virus in a human lymphoid cell line, J. Mol. Biol. 102 (3) (1976) 511–530.
[10] C. Daskalogianni, et al., Epstein-Barr virus-encoded EBNA1 and ZEBRA: targets for therapeutic strategies against EBV-carrying cancers, J. Pathol. 235 (2) (2015) 334–341.
[11] R. Longnecker, F. Neipel, Introduction to the human gamma-herpesviruses, in: A. Arvin, et al. (Eds.), Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis, 2007 (Cambridge).
[12] S.A. Rezk, X. Zhao, L.M. Weiss, Epstein-Barr virus (EBV)-associated lymphoid proliferations, a 2018 update, Hum. Pathol. 79 (2018) 18–41.
[13] S. Akhtar, et al., Epstein-Barr virus in gliomas: cause, association, or artifact? Front. Oncol. 8 (2018) 123.
[14] J.I. Cohen, et al., Epstein-Barr virus: an important vaccine target for cancer prevention, Sci. Transl. Med. 3 (107) (2011) 107fs7.
[15] K. Bjornevik, et al., Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis, Science 375 (6578) (2022) 296–301.
[16] J.B. Wilson, et al., EBNA1: oncogenic activity, immune evasion and biochemical functions provide targets for novel therapeutic strategies against epstein-barr virus- associated cancers, Cancers 10 (4) (2018).
[17] N. Blake, et al., Human CD8+ T cell responses to EBV EBNA1: HLA class I presentation of the (Gly-Ala)-containing protein requires exogenous processing, Immunity 7 (6) (1997) 791–802.
[18] S.P. Lee, et al., CD8 T cell recognition of endogenously expressed epstein-barr virus nuclear antigen 1, J. Exp. Med. 199 (10) (2004) 1409–1420.
[19] A. Leen, et al., Differential immunogenicity of Epstein-Barr virus latent-cycle proteins for human CD4(+) T-helper 1 responses, J. Virol. 75 (18) (2001) 8649–8659.
[20] C. Munz, Epstein-barr virus nuclear antigen 1: from immunologically invisible to a promising T cell target, J. Exp. Med. 199 (10) (2004) 1301–1304.
[21] C. Munz, et al., Human CD4(+) T lymphocytes consistently respond to the latent Epstein-Barr virus nuclear antigen EBNA1, J. Exp. Med. 191 (10) (2000) 1649–1660.
[22] K.S. Voo, et al., Evidence for the presentation of major histocompatibility complex class I-restricted Epstein-Barr virus nuclear antigen 1 peptides to CD8+ T lymphocytes, J. Exp. Med. 199 (4) (2004) 459–470.
[23] N. Blake, Immune evasion by gammaherpesvirus genome maintenance proteins, J. Gen. Virol. 91 (Pt 4) (2010) 829–846.
[24] A. De Leo, A. Calderon, P.M. Lieberman, Control of viral latency by episome maintenance proteins, Trends Microbiol. 28 (2) (2020) 150–162.
[25] M.S. Kang, E. Kieff, Epstein-Barr virus latent genes, Exp. Mol. Med. 47 (1) (2015) e131.
[26] T.A. Gahn, B. Sugden, An EBNA-1-dependent enhancer acts from a distance of 10 kilobase pairs to increase expression of the Epstein-Barr virus LMP gene, J. Virol. 69 (4) (1995) 2633–2636.
[27] B. Sugden, N. Warren, A promoter of Epstein-Barr virus that can function during latent infection can be transactivated by EBNA-1, a viral protein required for viral DNA replication during latent infection, J. Virol. 63 (6) (1989) 2644–2649.
[28] G. Poornima, et al., RGG-motif self-association regulates eIF4G-binding translation repressor protein Scd6, RNA Biol. 16 (9) (2019) 1215–1227.
[29] P. Thandapani, et al., Defining the RGG/RG motif, Mol. Cell. 50 (5) (2013) 613–623.
[30] D. Mackey, B. Sugden, The linking regions of EBNA1 are essential for its support of replication and transcription, Mol. Cell Biol. 19 (5) (1999) 3349–3359.
[31] J. Sears, et al., The amino terminus of Epstein-Barr Virus (EBV) nuclear antigen 1 contains AT hooks that facilitate the replication and partitioning of latent EBV genomes by tethering them to cellular chromosomes, J. Virol. 78

(21) (2004) 11487–11505.
[32] L. Frappier, The Epstein-Barr Virus EBNA1 Protein, Scientifica, Cairo), 2012 438204.
[33] J. Dheekollu, et al., Cell-cycle-dependent EBNA1-DNA crosslinking promotes replication termination at oriP and viral episome maintenance, Cell 184 (3) (2021) 643–654 e13.
[34] Y. Mei, et al., Cryo-EM structure and functional studies of EBNA1 binding to the family of repeats and dyad symmetry elements of epstein-barr virus oriP, J. Virol. 96 (17) (2022) e0094922.
[35] T.A. Gahn, C.L. Schildkraut, The Epstein-Barr virus origin of plasmid replication, oriP, contains both the initiation and termination sites of DNA replication, Cell 58 (3) (1989) 527–535.
[36] D. Reisman, J. Yates, B. Sugden, A putative origin of replication of plasmids derived from Epstein-Barr virus is composed of two cis-acting components, Mol. Cell Biol. 5 (8) (1985) 1822–1832.
[37] D.A. Wysokenski, J.L. Yates, Multiple EBNA1-binding sites are required to form an EBNA1-dependent enhancer and to activate a minimal replicative origin within oriP of Epstein-Barr virus, J. Virol. 63 (6) (1989) 2657–2666.
[38] D.R. Rawlins, et al., Sequence-specific DNA binding of the Epstein-Barr virus nuclear antigen (EBNA-1) to clustered sites in the plasmid maintenance region, Cell 42 (3) (1985) 859–868.
[39] T.L. Hodin, T. Najrana, J.L. Yates, Efficient replication of Epstein-Barr virus-derived plasmids requires tethering by EBNA1 to host chromosomes, J. Virol. 87 (23) (2013) 13020–13028.
[40] V.K. Nayyar, K. Shire, L. Frappier, Mitotic chromosome interactions of Epstein-Barr nuclear antigen 1 (EBNA1) and human EBNA1-binding protein 2 (EBP2), J. Cell Sci. 122 (Pt 23) (2009) 4341–4350.
[41] K. Shire, et al., EBP2, a human protein that interacts with sequences of the Epstein-Barr virus nuclear antigen 1 important for plasmid maintenance, J. Virol. 73 (4) (1999) 2587–2595.
[42] P. Kapoor, L. Frappier, EBNA1 partitions Epstein-Barr virus plasmids in yeast cells by attaching to human EBNA1-binding protein 2 on mitotic chromosomes, J. Virol. 77 (12) (2003) 6946–6956.
[43] P. Kapoor, K. Shire, L. Frappier, Reconstitution of Epstein-Barr virus-based plasmid partitioning in budding yeast, EMBO J. 20 (1–2) (2001) 222–230.
[44] M.J. Lista, et al., The long-lasting love affair between the budding yeast Saccharomyces cerevisiae and the Epstein-Barr virus, Biotechnol. J. 10 (11) (2015) 1670–1681.
[45] G. Kennedy, B. Sugden, EBNA-1, a bifunctional transcriptional activator, Mol. Cell Biol. 23 (19) (2003) 6901–6908.
[46] H. Wu, P. Kapoor, L. Frappier, Separation of the DNA replication, segregation, and transcriptional activation functions of Epstein-Barr nuclear antigen 1, J. Virol. 76 (5) (2002) 2480–2490.
[47] D.F. Ceccarelli, L. Frappier, Functional analyses of the EBNA1 origin DNA binding protein of Epstein-Barr virus, J. Virol. 74 (11) (2000) 4939–4948.
[48] S. Van Scoy, et al., Human p32: a coactivator for Epstein-Barr virus nuclear antigen-1-mediated transcriptional activation and possible role in viral latent cycle DNA replication, Virology 275 (1) (2000) 145–157.
[49] Y. Wang, et al., P32/TAP, a cellular protein that interacts with EBNA-1 of Epstein-Barr virus, Virology 236 (1) (1997) 18–29.
[50] J. Levitskaya, et al., Inhibition of antigen processing by the internal repeat region of the Epstein-Barr virus nuclear antigen-1, Nature 375 (6533) (1995) 685–688.
[51] S. Apcher, et al., Epstein Barr virus-encoded EBNA1 interference with MHC class I antigen presentation reveals a close correlation between mRNA translation initiation and antigen presentation, PLoS Pathog. 6 (10) (2010) e1001151.
[52] M. Ossevoort, et al., Characterization of an immuno 'stealth' derivative of the herpes simplex virus thymidine-kinase gene, Cancer Gene Ther. 13 (6) (2006) 584–591.
[53] Y. Yin, B. Manoury, R. Fahraeus, Self-inhibition of synthesis and antigen presentation by Epstein-Barr virus-encoded EBNA1, Science 301 (5638) (2003) 1371–1374.
[54] K.L. Rock, A.L. Goldberg, Degradation of cell proteins and the generation of MHC class I-presented peptides, Annu. Rev. Immunol. 17 (1999) 739–779.
[55] C. Daskalogianni, et al., Gly-Ala repeats induce position- and substrate-specific regulation of 26 S proteasome-dependent partial processing, J. Biol. Chem. 283 (44) (2008) 30090–30100.
[56] J. Levitskaya, et al., Inhibition of ubiquitin/proteasome-dependent protein degradation by the Gly-Ala repeat domain of the Epstein-Barr virus nuclear antigen 1, Proc. Natl. Acad. Sci. U. S. A. 94 (23) (1997) 12616–12621.
[57] A. Sharipo, et al., A minimal glycine-alanine repeat prevents the interaction of ubiquitinated I kappaB alpha with the proteasome: a new mechanism for selective inhibition of proteolysis, Nat. Med. 4 (8) (1998) 939–944.
[58] N.P. Dantuma, et al., Inhibition of proteasomal degradation by the gly-Ala repeat of Epstein-Barr virus is influenced by the length of the repeat and the strength of the degradation signal, Proc. Natl. Acad. Sci. U. S. A. 97 (15) (2000) 8381–8385.
[59] S. Heessen, et al., Inhibition of ubiquitin/proteasome-dependent proteolysis in Saccharomyces cerevisiae by a Gly-Ala repeat, FEBS Lett. 555 (2) (2003) 397–404.
[60] J.W. Yewdell, DRiPs solidify: progress in understanding endogenous MHC class I antigen processing, Trends Immunol. 32 (11) (2011) 548–558.
[61] J.W. Yewdell, L.C. Anton, J.R. Bennink, Defective ribosomal products (DRiPs): a major source of antigenic peptides for MHC class I molecules? J. Immunol. 157 (5) (1996) 1823–1826.
[62] J.W. Yewdell, C.V. Nicchitta, The DRiP hypothesis decennial: support, controversy, refinement and extension, Trends Immunol. 27 (8) (2006) 368–373.
[63] N.P. Croft, et al., Kinetics of antigen expression and epitope presentation during virus infection, PLoS Pathog. 9 (1) (2013) e1003129.
[64] L.C. Anton, J.W. Yewdell, Translating DRiPs: MHC class I immunosurveillance of pathogens and tumors, J. Leukoc. Biol. 95 (4) (2014) 551–562.
[65] T. Wu, et al., Quantification of epitope abundance reveals the effect of direct and cross-presentation on influenza CTL responses, Nat. Commun. 10 (1) (2019) 2846.
[66] B.P. Dolan, et al., Distinct pathways generate peptides from defective ribosomal products for CD8+ T cell immunosurveillance, J. Immunol. 186 (4) (2011) 2065–2072.
[67] S. Khan, et al., Cutting edge: neosynthesis is required for the presentation of a T cell epitope from a long-lived viral protein, J. Immunol. 167 (9) (2001) 4801–4804.
[68] J.W. Yewdell, J. Holly, DRiPs get molecular, Curr. Opin. Immunol. 64 (2020) 130–136.
[69] J. Tellam, et al., Endogenous presentation of CD8+ T cell epitopes from Epstein-Barr virus-encoded nuclear antigen 1, J. Exp. Med. 199 (10) (2004) 1421–1431.
[70] S. Apcher, et al., mRNA translation regulation by the Gly-Ala repeat of Epstein-Barr virus nuclear antigen 1, J. Virol. 83 (3) (2009) 1289–1298.
[71] R. Fahraeus, Do peptides control their own birth and death? Nat. Rev. Mol. Cell Biol. 6 (3) (2005) 263–267.
[72] G. Angrand, et al., Sneaking out for happy hour: yeast-based approaches to explore and modulate immune response and immune evasion, Genes 10 (9) (2019).
[73] S. Bach, P. Colas, M. Blondel, [Budding yeast, a model and a tool... also for biomedical research], Med. Sci. 36 (5) (2020) 504–514.
[74] A.H. Kachroo, et al., Humanized yeast to model human biology, disease and evolution, Dis Model Mech 15 (6) (2022).
[75] S. Bach, et al., Isolation of drugs active against mammalian prions using a yeast-based screening assay, Nat. Biotechnol. 21 (9) (2003) 1075–1081.
[76] E. Couplan, et al., A yeast-based assay identifies drugs active against human mitochondrial disorders, Proc. Natl. Acad. Sci. U. S. A. 108 (29) (2011) 11989–11994.
[77] F. Giorgini, et al., A genomic screen in yeast implicates kynurenine 3-monooxygenase as a therapeutic target for Huntington disease, Nat. Genet. 37 (5) (2005) 526–531.
[78] M.J. Lista, et al., Nucleolin directly mediates Epstein-Barr virus immune evasion through binding to G-quadruplexes of EBNA1 mRNA, Nat. Commun. 8 (2017) 16043.
[79] C. Voisset, et al., A yeast-based assay identifies drugs that interfere with immune evasion of the Epstein-Barr virus, Dis Model Mech 7 (4) (2014) 435–444.
[80] S. Bach, et al., A yeast-based assay to isolate drugs active against mammalian prions, Methods 39 (1) (2006) 72–77.
[81] M.J. Lista, et al., A yeast model for the mechanism of the Epstein-Barr virus immune evasion identifies a new therapeutic target to interfere with the virus stealthiness, Microb Cell 4 (9) (2017) 305–307.
[82] J.D. Beaudoin, J.P. Perreault, 5'-UTR G-quadruplex structures acting as translational repressors, Nucleic Acids Res. 38 (20) (2010) 7022–7036.
[83] M.C. Didiot, et al., The G-quartet containing FMRP binding site in FMR1 mRNA is a potent exonic splicing enhancer, Nucleic Acids Res. 36 (15) (2008) 4902–4912.
[84] D. Gomez, et al., Telomerase downregulation induced by the G-quadruplex ligand 12459 in A549 cells is mediated by hTERT RNA alternative splicing, Nucleic Acids Res. 32 (1) (2004) 371–379.
[85] V. Marcel, et al., G-quadruplex structures in TP53 intron 3: role in alternative splicing and in production of p53 mRNA isoforms, Carcinogenesis 32 (3) (2011) 271–278.
[86] A. Siddiqui-Jain, et al., Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription, Proc. Natl. Acad. Sci. U. S. A. 99 (18) (2002) 11593–11598.
[87] A.D. Cristillo, et al., Double-stranded RNA as a not-self alarm signal: to evade, most viruses purine-load their RNAs, but some (HTLV-1, Epstein-Barr) pyrimidine-load, J. Theor. Biol. 208 (4) (2001) 475–491.
[88] P. Murat, et al., G-quadruplexes regulate Epstein-Barr virus-encoded nuclear antigen 1 mRNA translation, Nat. Chem. Biol. 10 (5) (2014) 358–364.
[89] L.A. Dempsey, et al., G4 DNA binding by LR1 and its subunits, nucleolin and hnRNP D, A role for G-G pairing in immunoglobulin switch recombination, J. Biol. Chem. 274 (2) (1999) 1066–1071.
[90] L.A. Hanakahi, H. Sun, N. Maizels, High affinity interactions of nucleolin with G-G-paired rDNA, J. Biol. Chem. 274 (22) (1999) 15908–15912.
[91] A. von Hacht, et al., Identification and characterization of RNA guanine-quadruplex binding proteins, Nucleic Acids Res. 42 (10) (2014) 6630–6644.
[92] R. Prado Martins, et al., Cellulo protein-mRNA interaction assay to determine the action of G-quadruplex-binding molecules, Molecules 23 (12) (2018).
[93] O. Reznichenko, et al., Novel cationic bis(acylhydrazones) as modulators of Epstein-Barr virus immune evasion acting through disruption of interaction between nucleolin and G-quadruplexes of EBNA1 mRNA, Eur. J. Med. Chem. 178 (2019) 13–29.
[94] I. Manet, et al., Affinity of the anthracycline antitumor drugs Doxorubicin and

Sabarubicin for human telomeric G-quadruplex structures, Phys. Chem. Chem. Phys. 13 (2) (2011) 540–551.
[95] L. Scaglioni, et al., Nemorubicin and doxorubicin bind the G-quadruplex sequences of the human telomeres and of the c-MYC promoter element Pu22, Biochim. Biophys. Acta 1860 (6) (2016) 1129–1138.
[96] J. Norseen, F.B. Johnson, P.M. Lieberman, Role for G-quadruplex RNA binding by Epstein-Barr virus nuclear antigen 1 in DNA replication and metaphase chromosome attachment, J. Virol. 83 (20) (2009) 10336–10346.
[97] E. Lavezzo, et al., G-quadruplex forming sequences in the genome of all known human viruses: a comprehensive guide, PLoS Comput. Biol. 14 (12) (2018) e1006675.
[98] E. Ruggiero, S.N. Richter, G-quadruplexes and G-quadruplex ligands: targets and tools in antiviral therapy, Nucleic Acids Res. 46 (7) (2018) 3270–3283.
[99] E. Ruggiero, S.N. Richter, Targeting G-quadruplexes to achieve antiviral activity, Bioorg. Med. Chem. Lett. 79 (2023) 129085.
[100] E. Tosoni, et al., Nucleolin stabilizes G-quadruplex structures folded by the LTR promoter and silences HIV-1 viral transcription, Nucleic Acids Res. 43 (18) (2015) 8884–8897.
[101] W.X. Bian, et al., Binding of cellular nucleolin with the viral core RNA G-quadruplex structure suppresses HCV replication, Nucleic Acids Res. 47 (1) (2019) 56–68.
[102] S. Artusi, et al., The Herpes Simplex Virus-1 genome contains multiple clusters of repeated G-quadruplex: implications for the antiviral activity of a G-quadruplex ligand, Antivir. Res. 118 (2015) 123–131.
[103] I. Frasson, et al., Quindoline-derivatives display potent G-quadruplex-mediated antiviral activity against herpes simplex virus 1, Antivir. Res. 208 (2022) 105432.
[104] R.P. Martins, et al., Nuclear processing of nascent transcripts determines synthesis of full-length proteins and antigenic peptides, Nucleic Acids Res. 47 (6) (2019) 3086–3100.
[105] P. Murat, J. Tellam, Effects of messenger RNA structure and other translational control mechanisms on major histocompatibility complex-I mediated antigen presentation, Wiley Interdiscip Rev RNA 6 (2) (2015) 157–171.
[106] P. Dabral, et al., LANA and hnRNP A1 regulate the translation of LANA mRNA through G-quadruplexes, J. Virol. 94 (3) (2020).
[107] A.J. Zheng, et al., The different activities of RNA G-quadruplex structures are controlled by flanking sequences, Life Sci. Alliance 5 (2) (2022).
[108] W.C. Lee, Z.X. Xue, T. Melese, The NSR1 gene encodes a protein that specifically binds nuclear localization sequences and has two RNA recognition motifs, J. Cell Biol. 113 (1) (1991) 1–12.
[109] G. Angrand, et al., Type I arginine methyltransferases are intervention points to unveil the oncogenic Epstein-Barr virus to the immune system, Nucleic Acids Res. 50 (20) (2022) 11799–11819.
[110] Q. Wu, et al., Protein arginine methylation: from enigmatic functions to therapeutic targeting, Nat. Rev. Drug Discov. 20 (7) (2021) 509–530.
[111] R.S. Blanc, S. Richard, Arginine methylation: the coming of age, Mol. Cell. 65 (1) (2017) 8–24.
[112] J.J. Hamey, R.J. Separovich, M.R. Wilkins, M.T.- Mams, Protein methyltransferase motif analysis by mass spectrometry, J. Proteome Res. 17 (10) (2018) 3485–3491.
[113] B.M. Lorton, D. Shechter, Cellular consequences of arginine methylation, Cell. Mol. Life Sci. 76 (15) (2019) 2933–2956.
[114] D. Yagoub, et al., Yeast proteins Gar1p, Nop1p, Npl3p, Nsr1p, and Rps2p are natively methylated and are substrates of the arginine methyltransferase Hmt1p, Proteomics 15 (18) (2015) 3209–3218.
[115] M. Campbell, et al., Protein arginine methyltransferase 1-directed methylation of Kaposi sarcoma-associated herpesvirus latency-associated nuclear antigen, J. Biol. Chem. 287 (8) (2012) 5806–5818.
[116] M.D. Mostaqul Huq, et al., Suppression of receptor interacting protein 140 repressive activity by protein arginine methylation, EMBO J. 25 (21) (2006) 5094–5104.
[117] P. Bouvet, et al., Nucleolin interacts with several ribosomal proteins through its RGG domain, J. Biol. Chem. 273 (30) (1998) 19025–19029.
[118] Y.L. Chen, et al., Nucleolin is important for Epstein-Barr virus nuclear antigen 1-mediated episome binding, maintenance, and transcription, Proc. Natl. Acad. Sci. U. S. A. 111 (1) (2014) 243–248.
[119] P. Cohen, D. Cross, P.A. Janne, Kinase drug discovery 20 years after imatinib: progress and future directions, Nat. Rev. Drug Discov. 20 (7) (2021) 551–569.
[120] A. Vannutelli, et al., Where are G-quadruplexes located in the human transcriptome? NAR Genom Bioinform 2 (2) (2020) lqaa035.
[121] J.L. Huppert, S. Balasubramanian, Prevalence of quadruplexes in the human genome, Nucleic Acids Res. 33 (9) (2005) 2908–2916.
[122] J.L. Huppert, S. Balasubramanian, G-quadruplexes in promoters throughout the human genome, Nucleic Acids Res. 35 (2) (2007) 406–413.
[123] E. Cadoni, et al., Beyond small molecules: targeting G-quadruplex structures with oligonucleotides and their analogues, Nucleic Acids Res. 49 (12) (2021) 6638–6659.
[124] M. Tassinari, et al., Selective targeting of mutually exclusive DNA G-quadruplexes: HIV-1 LTR as paradigmatic model, Nucleic Acids Res. 48 (9) (2020) 4627–4642.
[125] S. Mandal, et al., Submolecular dissection reveals strong and specific binding of polyamide-pyridostatin conjugates to human telomere interface, Nucleic Acids Res. 47 (7) (2019) 3295–3305.
[126] K. Halder, S. Chowdhury, Quadruplex-coupled kinetics distinguishes ligand binding between G4 DNA motifs, Biochemistry 46 (51) (2007) 14762–14770.
[127] M. Scalabrin, et al., Selective recognition of a single HIV-1 G-quadruplex by ultrafast small-molecule screening, Anal. Chem. 93 (46) (2021) 15243–15252.
[128] G.L. Cuthbert, et al., Histone deimination antagonizes arginine methylation, Cell 118 (5) (2004) 545–553.
[129] Y. Wang, et al., Human PAD4 regulates histone arginine methylation levels via demethylimination, Science 306 (5694) (2004) 279–283.
[130] C. Tanikawa, et al., Citrullination of RGG motifs in FET proteins by PAD4 regulates protein aggregation and ALS susceptibility, Cell Rep. 22 (6) (2018) 1473–1483.
[131] S. Leonard, et al., Arginine methyltransferases are regulated by epstein-barr virus in B cells and are differentially expressed in hodgkin's lymphoma, Pathogens 1 (1) (2012) 52–64.
[132] A.J.L. Zheng, et al., The nascent polypeptide-associated complex (NAC) controls translation initiation in cis by recruiting nucleolin to the encoding mRNA, Nucleic Acids Res. 50 (17) (2022) 10110–10122.
[133] J. Norseen, et al., RNA-dependent recruitment of the origin recognition complex, EMBO J. 27 (22) (2008) 3024–3035.
[134] J.B. Wilson, J.L. Bell, A.J. Levine, Expression of Epstein-Barr virus nuclear antigen-1 induces B cell neoplasia in transgenic mice, EMBO J. 15 (12) (1996) 3117–3126.
[135] J.M. Adams, S. Cory, Myc oncogene activation in B and T lymphoid tumours, Proc. R. Soc. Lond. B Biol. Sci. 226 (1242) (1985) 59–72.
[136] J.M. Adams, et al., The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice, Nature 318 (6046) (1985) 533–538.
[137] S.V. Gnanasundram, et al., Pl3Kdelta activates E2F1 synthesis in response to mRNA translation stress, Nat. Commun. 8 (1) (2017) 2103.
[138] R.J. Lamontagne, et al., A multi-omics approach to Epstein-Barr virus immortalization of B-cells reveals EBNA1 chromatin pioneering activities targeting nucleotide metabolism, PLoS Pathog. 17 (1) (2021) e1009208.
[139] T.E. Messick, et al., Structure-based design of small-molecule inhibitors of EBNA1 DNA binding blocks Epstein-Barr virus latent infection and tumor growth, Sci. Transl. Med. 11 (482) (2019).
[140] S.S. Soldan, et al., EBNA1 inhibitors have potent and selective antitumor activity in xenograft models of Epstein-Barr virus-associated gastric cancer, Gastric Cancer 24 (5) (2021) 1076–1088.
[141] T.V. Lanz, et al., Clonally expanded B cells in multiple sclerosis bind EBV EBNA1 and GlialCAM, Nature 603 (7900) (2022) 321–327.
[142] S.S. Soldan, P.M. Lieberman, Epstein-Barr virus and multiple sclerosis, Nat. Rev. Microbiol. 21 (1) (2023) 51–64.
[143] X. Cui, C.M. Snapper, Epstein barr virus: development of vaccines and immune cell therapy for EBV-associated diseases, Front. Immunol. 12 (2021) 734471.
[144] J. Ruhl, C.S. Leung, C. Munz, Vaccination against the epstein-barr virus, Cell. Mol. Life Sci. 77 (21) (2020) 4315–4324.
[145] L. Zhong, et al., Urgency and necessity of Epstein-Barr virus prophylactic vaccines, NPJ Vaccines 7 (1) (2022) 159.
[146] J. Ruhl, et al., Heterologous prime-boost vaccination protects against EBV antigen-expressing lymphomas, J. Clin. Invest. 129 (5) (2019) 2071–2087.

Discussion

L'épissage des ARN est un processus essentiel qui est régulé par de nombreux éléments qui peuvent agir en *cis* ou en *trans*. L'épissage du gène *BCL-x* produit deux isoformes aux fonctions antagonistes, une protéine canonique Bcl-xL qui est anti-apoptotique et une protéine alternative Bcl-xS qui elle est pro-apoptotique. La modification de la balance entre ces deux isoformes est associée à diverses pathologies telles que des cancers, des pathologies cardiaques ou encore certaines formes de diabète de type II. En ce qui concerne l'épissage de *BCL-x*, de nombreux facteurs *cis* et *trans* ont déjà été identifiés et influencent la sélection du site 5' fort ou du site 5' faible d'épissage. Plus récemment, l'identification de deux G-quadruplex d'ARN (rG4) dans le pré-ARNm de *BCL-x* a conduit à l'hypothèse d'un nouveau mode de régulation possible de ce gène par ces rG4. De plus, ces deux rG4 sont localisés dans l'exon 2 à proximité des deux sites 5' d'épissages de *BCL-x* ce qui tendait à conforter cette hypothèse. En effet, si les exons sont enrichis en guanines, la présence de G4 semble y être contre-sélectionnée et ils sont le plus souvent retrouvés dans les introns. La présence de rG4 dans les introns est alors associée à une activité régulatrice sur l'épissage dont l'effet est dépendant de la position du G4 par rapport au site d'épissage. La présence de ces deux rG4 dans l'exon 2 de *BCL-x* suggère que ces G4 ont été conservés au cours de l'évolution, ce qui est cohérent avec un rôle possible dans la régulation de l'épissage de *BCL-x*. Bien que ces rG4 aient déjà été caractérisés, par nos données et par d'autres publications, la question au début de cette thèse était de déterminer s'ils jouaient, ou non, un rôle dans la régulation de l'épissage alternatif de *BCL-x* puis, si tel était le cas, de comprendre leur rôle et comment ces structures secondaires sont régulées. L'identification et l'étude de protéines se liant aux G4 (*G4-binding proteins* -G4BP-) représente un aspect important de la compréhension de la régulation des rG4 et de leur rôle potentiel. Aussi avons-nous cherché à identifier des protéines ayant de l'affinité pour l'un ou l'autre des rG4 de *BCL-x*. En se basant sur les données d'une étude globale qui avait identifié le facteur d'épissage RBM25, déjà connu pour jouer un rôle dans la régulation de l'épissage alternatif de *BCL-x*, comme une G4BP potentielle, nous avons testé si la protéine RBM25 présentait, ou non, de l'affinité pour l'un des deux rG4 du pré-ARNm de *BCL-x*.

La protéine RBM25, une protéine de liaison au G-quadruplex qui régule l'épissage de *BCL-x*

Grâce à l'utilisation d'oligonucléotides capables de former, ou non -du fait de mutations-, les rG4 de *BCL-x* nous avons montré que la protéine RBM25 possède une affinité spécifique pour le rG4 GQ-2. De plus, nous avons montré que la protéine RBM25 est capable de se lier avec le rG4 GQ-2 de manière directe et dépendante de la formation du rG4. Conformément à des données de la littérature, nous avons observé que RBM25 est une protéine qui interagit avec l'ARN de *BCL-x* dans le noyau des cellules par des expériences de PLA. L'adaptation de cette technique pour observer l'interaction entre la protéine RBM25 et l'ARN produit par des minigènes nous a permis de montrer qu'en mutant le GQ-2 pour empêcher la formation du rG4, RBM25 ne se lie plus avec l'ARN de *BCL-x*. Ces données confirment que RBM25 est bien une nouvelle G4BP. Nous montrons également que RBM25 est capable de lier des G4 d'ADN (dG4) en plus des rG4, avec toutefois une préférence pour les rG4 comme le montre l'expérience de *RNA pulldown* entre la protéine RBM25 et le G4 GQ-2 dans sa version rG4 versus sa version dG4 (Figure Supplémentaire 3 du m/s Le Sénéchal R et al *Nucleic Acids Research* 2023).

RBM25, qui est une protéine peu étudiée, est composée de différents domaines de liaison aux acides nucléiques (un domaine RRM, un domaine RE et un domaine PWI). L'identification de la capacité de RBM25 à lier les rG4 soulève la question de savoir par quel domaine RBM25 est capable d'interagir avec ces rG4. Grâce à différentes constructions de RBM25 délétées, ou pas, pour l'un ou l'autre de ces domaines, nous avons identifié par *RNA pulldown* que le domaine RE est nécessaire et suffisant pour la liaison avec le rG4 GQ-2. Le domaine RE représente donc un nouveau domaine de liaison aux rG4. Cependant, bien que ce domaine seul puisse lier le rG4 GQ-2, il n'est pas suffisant pour réguler l'épissage de *BCL-x*. Ceci indique que les autres domaines de RBM25 sont nécessaires, très certainement *via* le recrutement de la machinerie d'épissage.

Dans la littérature, les domaines RE sont aussi qualifiés de domaines BAD (*Basic-Acidic-Dipeptide*) (Bishof et al., 2018). Ces domaines désordonnés de faible complexité sont impliqués dans de nombreux processus tels que la régulation des

ARN, l'assemblage du splicéosome, le repliement des protéines et la formation du ribosome. Il a également été suggéré que ces domaines BAD soient nécessaires pour la localisation nucléaire des protéines. Dans le cas de RBM25, il semble que ce domaine RE soit nécessaire plus précisément pour la localisation de la protéine dans les *nuclear speckles* (figure 4f Le Sénéchal et al, Zhou et al., 2008). De plus, ce domaine est nécessaire pour l'interaction de RBM25 avec la protéine LUC7L3 et le recrutement de U1 snRNP (Bishof et al., 2018). D'après les résultats de cette thèse, nous montrons un nouveau rôle de ce domaine RE, comme nouveau domaine de liaison aux rG4. Il serait dès lors intéressant de rechercher si d'autres protéines avec des domaines BAD sont également des G4BP. D'après l'étude globale de Su qui a servi de base à notre travail sur RBM25, la protéine SRSF3, qui contient un domaine BAD, serait une potentielle G4BP (Su et al., 2021). Il serait donc intéressant de rechercher si SRSF3 présente de l'affinité pour le rG4 GQ-2, ou avec d'autres rG4. En outre, il a déjà été montré que cette protéine se lie dans la région où se situe le GQ-2 (Bielli et al., 2014). Si tel est le cas, il serait également intéressant de savoir dans quelle mesure cette protéine agit en compétition ou en coopération/association à la protéine RBM25.

RBM25 régule donc l'épissage de *BCL-x* en interagissant avec le rG4 GQ-2. Le mécanisme d'action de RBM25 sur l'épissage de *BCL-x* reste pour le moment inconnu. Au moins deux mécanismes peuvent être envisagés pour expliquer le rôle de RBM25 dans la régulation de *BCL-x* et du GQ-2 (figure 15). Dans le premier modèle, le rG4 GQ-2 représenterait une plateforme de reconnaissance pour RBM25. Le facteur d'épissage RBM25 recruterait alors la machinerie d'épissage sur le site 5' alternatif et favoriserait ainsi la synthèse de la protéine pro-apoptotique Bcl-xS. Cette hypothèse est soutenue par la capacité de RBM25 à recruter LUC7L3 et indirectement la snRNP U1, un des composants essentiels à l'assemblage du splicéosome. De plus, ce modèle est renforcé par les résultats que nous avons obtenus avec le composé PhenDC3 puisque celui-ci augmente la stabilité du rG4 GQ-2 et donc la liaison entre la protéine RBM25 et l'ARNm de *BCL-x*. Dans le second modèle, RBM25 agirait directement sur la stabilité et la capacité de la séquence GQ-2 à former un rG4. La formation de ce rG4 serait nécessaire à l'épissage de *BCL-x*. Cette seconde hypothèse est soutenue par des travaux qui ont montré que la stabilisation du rG4 GQ-2 de *BCL*-x par un ligand de G4 (GQC-05) est uniquement observée en présence d'extraits nucléaires

(Bhogadia et al., 2022). Ces résultats suggèrent ainsi que l'existence du rG4 GQ-2 dans les cellules requiert un facteur nucléaire, qui pourrait être RBM25. Récemment, une étude a montré que certains ligands de G4 peuvent agir comme des chaperons pour des rG4 (Lejault et al., 2023). Ces ligands chaperons facilitent et/ou stabilisent la formation de rG4. Ils pourraient ainsi permettre la formation de G4 qui ne se formeraient pas dans les conditions physiologiques. Certaines protéines, comme RBM25, pourraient ainsi agir de la même manière et seraient, par là-même, nécessaires à la formation de certains rG4.

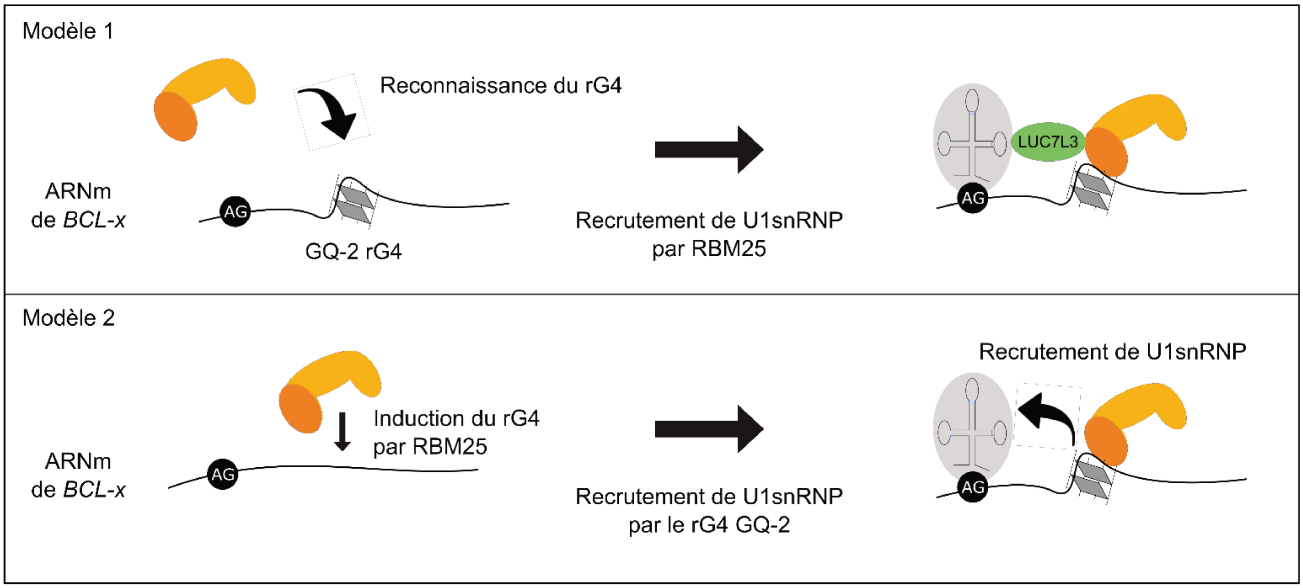


Figure 15 : modèles hypothétiques de la régulation du rG4 GQ-2 par RBM25.

L'interaction de la protéine RBM25 avec le rG4 GQ-2 est nécessaire pour le recrutement du splicéosome sur le site 5' d'épissage alternatif. Deux hypothèses sont proposées pour expliquer le rôle de cette protéine sur le rG4 et le recrutement de U1.

Dans le premier modèle, le rG4 GQ-2 sert de plateforme de reconnaissance qui est détectée par RBM25. Une fois associé à l'ARNm de BCL-x, RBM25 est capable de recruter d'autres facteurs d'épissage tels que LUC7L3, qui vont recruter U1snRNP sur le site 5' faible.

Dans le second modèle, RBM25 se lie sur la séquence du GQ-2 de BCL-x. Cette interaction conduit à la formation et la stabilisation du rG4 qui est nécessaire pour le recrutement du splicéosome sur le site 5' alternatif.

Le rG4 GQ-2 est essentiel pour la sélection du site 5' d'épissage alternatif et la synthèse de Bcl-xS

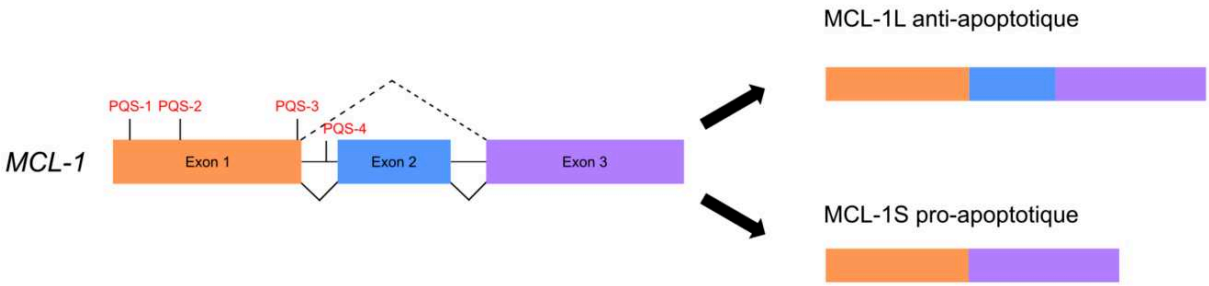
L'identification de la protéine RBM25 comme une nouvelle G4BP capable de se lier au rG4 GQ-2 nous a amenés à étudier le rôle de ce rG4 dans l'épissage de *BCL-x*. Pour cela nous avons synthétisé des minigènes capables, ou non, de former le rG4 GQ-2. Avec ces outils, nous avons montré que la formation du rG4 du GQ-2 est essentielle pour la synthèse de l'isoforme alternative pro-apoptotique Bcl-xS. De plus, en l'absence du rG4 GQ-2, la protéine RBM25 n'est plus capable de se lier à l'ARNm de *BCL-x*. De sorte, nous avons identifié le rG4 GQ-2 comme un *ESE* (*Exon Splicing Enhancer*) qui favorise l'utilisation du site 5' d'épissage alternatif et la synthèse de la forme pro-apoptotique Bcl-xS.

Il est intéressant de remarquer que le rG4 GQ-2 est localisé au niveau de l'élément *cis* régulateur B2, identifié pour la première fois en 2005 par l'équipe de Chabot (Garneau et al., 2005b). Dans cette étude, les auteurs décrivent une séquence de 30 nucléotides dans la région B2, nommée B2G, comme la région active pour l'épissage de B2 et essentielle pour la synthèse de Bcl-xS. Lorsque cette région est mutée au niveau de deux guanines, l'épissage de la forme alternative est complètement aboli. Cette région B2G correspond à la séquence GQ-2 étudiée dans cette thèse. De plus, l'effet de la substitution des guanines sur l'épissage observé dans l'étude de 2005 est similaire aux résultats observés dans notre étude avec GM-2, la version mutante du GQ-2, qui empêche la formation du rG4. Les effets régulateurs de la région B2 observés dans l'article de 2005 et dans le nôtre, semblent donc être le résultat de la formation du rG4 au niveau de GQ-2 qui agirait comme un *ESE* pour le site 5' alternatif de *BCL-x*. Cette hypothèse est renforcée par une autre publication de 2008 qui caractérise une région riche en guanine essentielle à l'épissage de la forme alternative Bcl-xS (Hai et al., 2008). Encore une fois, cette région correspond à celle du GQ-2. En outre, les auteurs de cette étude montrent que le rôle du GQ-2 est variable en fonction de sa localisation par rapport au site d'épissage. Introduite entre deux sites d'épissages de même force, le GQ-2 favoriserait le site d'épissage situé en amont de la séquence. Une hypothèse serait qu'en fonction de la position du G4, il puisse agir comme un *ESE* ou un *ESS* (*Exon Splicing Silencer*) en fonction de sa position par rapport au site d'épissage. Cette hypothèse est déjà vérifiée pour certains facteurs d'épissages dont l'effet sur

l'épissage est dépendant du site de liaison du facteur par rapport au site d'épissage. Ainsi, il a été montré que PTBP1 régule de manière différente l'épissage et peut induire soit l'inclusion d'exon, ou au contraire le saut d'exon, en fonction de la position de son site de liaison par rapport aux sites d'épissage (Xue et al., 2009).

L'identification du rG4 GQ-2 comme un nouvel *ESE* de *BCL-x* ajoute un niveau de régulation supplémentaire à l'épissage de ce gène. Cependant, le rôle du r GQ-5 n'a pas été exploré pendant cette thèse. Il est pourtant vraisemblable que le rG4 qui se forme dans la séquence GQ-5 soit également un élément régulateur de *BCL-x* qui mérite d'être étudié. Ce rG4 est localisé sur le site 5' d'épissage canonique, aussi il est possible que la formation du rG4 entre directement en compétition avec l'assemblage du splicéosome sur ce site, et donc qu'il régule de manière négative la sélection du site 5' canonique.

D'après les résultats présentés dans cette thèse, le rG4 GQ-2 représente un mécanisme essentiel pour la synthèse de la forme alternative pro-apoptotique de *BCL-x*. Il serait désormais intéressant de déterminer si des rG4 régulent la synthèse de formes alternatives d'autres gènes régulateurs de l'apoptose. En effet, de nombreux gènes régulateurs de la voie intrinsèque de l'apoptose produisent des versions alternatives avec une fonction antagoniste à l'isoforme canonique. Ainsi, les caspases 2 et 3 sont capables de synthétiser des protéines alternatives anti-apoptotiques. A l'inverse, et tout comme le gène *BCL-x*, le gène anti-apoptotique *MCL-1* produit une forme alternative pro-apoptotique (figure 16). De plus, d'après le logiciel G4 Hunter, quatre PQS (*Potential Quadruplex-forming Sequence*) pourraient se former dans l'exon 1 et l'intron 1 de *MCL-1* dont deux à proximité des sites d'épissage. L'étude de ce gène et de sa régulation par des potentiels rG4 représente ainsi une piste intéressante pour mieux comprendre à la fois la régulation de l'épissage par les rG4, mais aussi l'implication de ces derniers dans la fonction de gènes régulateurs majeurs de l'apoptose.



PQS de *MCL-1*	Séquence	G4 hunter Score
PQS-1	GTGGGGGGGCCGGCTTGGGGGCCGGCAGCGG	1,237
PQS-2	GGCCCGGCGAGAGATAGGGGGAGGGGAGGCCGGCGCGG	1,089
PQS-3	GACGGGTTGGGGATGGCGTGCAGCG	1,2
PQS-4	GGGTGGGATGTCAATTTCAAGTGGGG	1,308

Figure 16 : épissage alternatif du gène anti-apoptotique *MCL-1*, un régulateur majeur de l'apoptose aux fonctions redondantes avec *BCL-x*

Le gène *MCL-1* est constitué de trois exons. Le transcrit entier code la protéine anti-apoptotique et canonique Mcl-1L. Le variant alternatif Mcl-1S est produit suite au saut de l'exon 2. Ce saut d'exon conduit à un décalage du cadre de lecture et la perte de fonction des domaines localisés dans l'exon 3. La protéine MCL-1S devient alors pro-apoptotique et s'associe à la protéine MCL-1L pour inhiber sa fonction anti-apoptotique. Quatre *PQS* (*Potentiel Quadruplex-forming Sequences*) localisés dans l'exon 1 et dans l'intron 1 sont prédits par le logiciel G4 hunter.

Les ligands de G4 : des outils pour étudier l'épissage de *BCL-x* et des candidats médicaments potentiels pour traiter les cancers chimiorésistants

D'après les résultats présentés dans cette thèse, le rG4 GQ-2 est essentiel pour l'épissage alternatif du gène *BCL-x*. Pour continuer d'étudier ce mécanisme, nous avons étudié l'effet de différents ligands de G4 sur ce système. Ces molécules peuvent stabiliser ou déstabiliser les rG4. Elles peuvent également agir en compétition avec les G4BP qui interagissent avec ces rG4. Pour cette raison, nous avons observé les effets du composé PhenDC3 sur l'interaction de RBM25 avec le rG4 GQ-2. En traitant des cellules par le composé PhenDC3 nous avons observé une augmentation dose-dépendante de la synthèse de la forme alternative Bcl-xS, aussi bien dans le modèle minigène, que sur le gène *BCL-x* endogène. Nous avons ensuite montré par des tests *in vitro*, que le PhenDC3 stabilisait le rG4 GQ-2, puis par des tests de PLA que le composé PhenDC3 augmentait la liaison de la protéine RBM25 avec l'ARNm de *BCL-x*. Ainsi, le composé PhenDC3 stabilise et/ou promeut la formation du rG4 GQ-2 et, en conséquence, favorise la liaison de RBM25 à l'ARNm de *BCL-x*, ce qui, *in fine*, augmente la sélection du site 5' alternatif et la synthèse de la forme pro-apoptotique Bcl-xS.

L'identification de l'interaction RBM25/rG4 GQ-2 représente donc un point d'intervention pour réguler la balance entre Bcl-xL et Bcl-xS. Cette piste est pertinente pour de nombreuses pathologies où la balance entre ces deux isoformes est dérégulée, en particulier dans les cancers où la forme anti-apoptotique Bcl-xL est surexprimée, ce qui est associé à une résistance aux chimiothérapies. Contrôler l'épissage en faveur de la forme alternative pro-apoptotique Bcl-xS représente donc un moyen de resensibiliser des cellules cancéreuses aux chimiothérapies. En outre, nous avons testé l'effet du composé PhenDC3 sur l'apoptose de cellules. Ce test réalisé sur des cellules cancéreuses, montre que le composé PhenDC3 induit une augmentation de l'apoptose qui est corrélée avec l'augmentation de la forme pro-apoptotique Bcl-xS. Ainsi, cibler l'interaction de RBM25 avec le rG4 GQ-2 représente un point d'intervention pertinent pour sensibiliser les cellules cancéreuses aux chimiothérapies. C'est avec cette idée que nous avons recherché des nouveaux ligands de G4 capables de moduler l'épissage de *BCL-x*. Parmi la librairie de 90

molécules structurellement proches du composé PhenDC3, nous avons identifié deux nouveaux composés, PhenDH8 et PhenDH9. L'identification de ces deux molécules est d'autant plus intéressante que ces ligands sont plus efficaces que le composé PhenDC3 pour augmenter la synthèse de Bcl-xS ainsi que pour induire l'apoptose des cellules. Toutefois, bien que nous ayons observé une augmentation de la synthèse de Bcl-xS et de l'apoptose, cela ne constitue pas une preuve que l'induction observée de l'apoptose soit une conséquence directe de l'augmentation de la forme pro-apoptotique Bcl-xS. En effet, de nombreuses études utilisent déjà les ligands de G4 pour induire la mort de cellules cancéreuses. Les ligands de G4 ayant une capacité de se lier à de nombreux G4, aussi bien ADN qu'ARN, les effets peuvent être attribués à la stabilisation des ligands dans les télomères, au niveau de promoteurs d'oncogènes, à l'induction d'instabilité génomique ou même à l'augmentation de la synthèse de protéines pro-apoptotiques. Ainsi, une des principales limites des ligands de G4 est leur manque de spécificité. Développer des ligands de G4 avec une spécificité pour des G4 précis permettrait de s'en servir comme outils. Dans le cas du GQ-2 cela permettrait de confirmer le lien entre l'augmentation de Bcl-xS et l'apoptose. Bien que les G4 sont structurellement similaires entre eux avec une base de G-quartet, de nombreuses caractéristiques rendent chaque G4 unique, telles que les séquences bornant le G4 ou les boucles qu'ils contiennent. Ces caractéristiques uniques peuvent, en théorie, être exploitées pour développer des ligands de G4 spécifiques à un G4 (Scalabrin et al., 2021). En outre, en plus d'être des outils pour étudier spécifiquement certains G4 dans un contexte cellulaire, de tels ligands représenteraient des candidats médicaments pertinents pour lutter contre certains cancers. A l'inverse, des nouveaux composés capables de déstabiliser des G4 ont récemment été développés. Dans le cas de l'épissage du gène *BCL-x*, ils présenteraient un intérêt thérapeutique pour traiter des pathologies, telles que certaines formes de diabète ou de maladie cardiaque, où la forme pro-apoptotique Bcl-xS est surexprimée (Mitteaux et al., 2021).

Pour terminer, il serait également intéressant de tester les composés PhenDC3, PhenDH8 et PhenDH9 en co-traitement avec des molécules anti-cancéreuses dans un premier temps dans un modèle de cellules résistantes aux chimiothérapies pour observer si les effets sont intéressants, puis dans des modèles *in vivo*. En effet, il est possible d'envisager des stratégies basées sur le co-traitement de ligands de G4 avec les BH3 mimétiques. Le mécanisme d'action des BH3 mimétiques est basé sur leur

capacité à inhiber l'interaction entre les protéines pro-apoptotiques et les protéines anti-apoptotiques en interagissant avec les molécules anti-apoptotiques. Les protéines Bax et Bak ne sont alors plus séquestrées et peuvent induire la perméabilisation de la membrane mitochondriale externe et l'apoptose de la cellule. Le travail sur les BH3 mimétiques a conduit au développement de plusieurs composés dont le chef de file est le Navitoclax (ABT-263), un composé présentant une forte affinité pour les poches BH3 des protéines Bcl-2 et Bcl-xL. Les BH3 mimétiques sont capables d'induire la mort cellulaire de cellules cancéreuses seules ou en association avec des molécules anti-cancéreuses plus classiques (doxorubicine, cisplatine) en présentant un effet synergique sur la diminution de la viabilité cellulaire. Cependant, des essais cliniques réalisés avec les BH3 mimétiques ont montré des effets secondaires importants, dont notamment des thrombopénies et l'apparition de résistance à ces molécules.

Ainsi, l'utilisation de ligands de G4 en co-traitement avec des BH3 mimétiques représente une stratégie pertinente car, en modulant le ratio entre les protéines Bcl-xL et Bcl-xS, les ligands de G4 devraient permettre d'utiliser des concentrations plus faibles de BH3 mimétiques afin de minimiser les effets secondaires de ces derniers. Il est aussi envisageable d'adapter cette stratégie avec d'autres molécules anticancéreuses et de réaliser des co-traitements avec des ligands de G4 afin de sensibiliser les cellules à ces chimiothérapies en rétablissant la balance Bcl-xL/Bcl-xS. L'utilisation de ligands de G4 pourrait être d'autant plus pertinente qu'il est possible que ces molécules régulent également l'épissage d'autres gènes régulateurs de l'apoptose tels que *MCL-1*.

Bibliographie :

Alavian, K. N., Li, H., Collis, L., Bonanni, L., Zeng, L., Sacchetti, S., Lazrove, E., Nabili, P., Flaherty, B., Graham, M., Chen, Y., Messerli, S. M., Mariggio, M. A., Rahner, C., McNay, E., Shore, G. C., Smith, P. J. S., Hardwick, J. M., & Jonas, E. A. (2011). Bcl-xL regulates metabolic efficiency of neurons through interaction with the mitochondrial F1FO ATP synthase. *Nature Cell Biology*, *13*(10), 1224-1233. https://doi.org/10.1038/ncb2330

Al-Qattan, M. M. (2019). A Review of the Genetics and Pathogenesis of Syndactyly in Humans and Experimental Animals : A 3-Step Pathway of Pathogenesis. *BioMed Research International*, *2019*, 9652649. https://doi.org/10.1155/2019/9652649

Bhogadia, M., Stone, B., Del Villar Guerra, R., Muskett, F. W., Ghosh, S., Taladriz-Sender, A., Burley, G. A., Eperon, I. C., Hudson, A. J., & Dominguez, C. (2022). Biophysical characterisation of the Bcl-x pre-mRNA and binding specificity of the ellipticine derivative GQC-05 : Implication for alternative splicing regulation. *Frontiers in Molecular Biosciences*, *9*, 943105. https://doi.org/10.3389/fmolb.2022.943105

Bielli, P., Bordi, M., Biasio, V. D., & Sette, C. (2014). Regulation of BCL-X splicing reveals a role for the polypyrimidine tract binding protein (PTBP1/hnRNP I) in alternative 5′ splice site selection. *Nucleic Acids Research*, *42*(19), 12070-12081. https://doi.org/10.1093/nar/gku922

Biffi, G., Tannahill, D., McCafferty, J., & Balasubramanian, S. (2013). Quantitative visualization of DNA G-quadruplex structures in human cells. *Nature Chemistry*, *5*(3), 182-186. https://doi.org/10.1038/nchem.1548

Bishof, I., Dammer, E. B., Duong, D. M., Kundinger, S. R., Gearing, M., Lah, J. J., Levey, A. I., & Seyfried, N. T. (2018). RNA-binding proteins with basic-acidic dipeptide (BAD) domains self-assemble and aggregate in Alzheimer's disease. *The Journal of Biological Chemistry*, *293*(28), 11047-11066. https://doi.org/10.1074/jbc.RA118.001747

Bleicken, S., Landeta, O., Landajuela, A., Basañez, G., & García-Sáez, A. J. (2013). Proapoptotic Bax and Bak proteins form stable protein-permeable pores of tunable size. *The Journal of Biological Chemistry*, *288*(46), 33241-33252. https://doi.org/10.1074/jbc.M113.512087

Bock, F. J., & Tait, S. W. G. (2020). Mitochondria as multifaceted regulators of cell death. *Nature Reviews. Molecular Cell Biology*, *21*(2), 85-100. https://doi.org/10.1038/s41580-019-0173-8

Boise, L. H., González-García, M., Postema, C. E., Ding, L., Lindsten, T., Turka, L. A., Mao, X., Nuñez, G., & Thompson, C. B. (1993). Bcl-x, a bcl-2-related gene that functions as a dominant regulator of apoptotic cell death. *Cell*, *74*(4), 597-608. https://doi.org/10.1016/0092-8674(93)90508-n

Bramlett, C. (2023). *RNA splicing factor Rbm25 underlies heterogeneous preleukemic clonal expansion in mice*. *141*(24).

Brázda, V., Kolomazník, J., Lýsek, J., Bartas, M., Fojta, M., Šťastný, J., & Mergny, J.-L. (2019). G4Hunter web application : A web server for G-quadruplex prediction. *Bioinformatics*, *35*(18), 3493-3495. https://doi.org/10.1093/bioinformatics/btz087

Cáceres, J. F., & Kornblihtt, A. R. (2002). Alternative splicing : Multiple control mechanisms and involvement in human disease. *Trends in Genetics: TIG*, *18*(4), 186-193. https://doi.org/10.1016/s0168-9525(01)02626-9

Cao, J.-W., Tang, Z.-B., Zhao, J.-W., Zhao, J.-K., Yao, J.-L., Sheng, X.-M., Zhao, M.-Q., Duan, Q., Han, B.-C., & Duan, S.-R. (2022). LncRNA nuclear-enriched abundant transcript 1 aggravates cerebral ischemia/reperfusion injury through activating early growth response-1/RNA binding motif protein 25 axis. *Journal of Neurochemistry*, *163*(6), 500-516. https://doi.org/10.1111/jnc.15692

Carlson, S. M., Soulette, C. M., Yang, Z., Elias, J. E., Brooks, A. N., & Gozani, O. (2017). RBM25 is a global splicing factor promoting inclusion of alternatively spliced exons and is itself regulated by lysine mono-methylation. *Journal of Biological Chemistry*, *292*(32), 13381-13390. https://doi.org/10.1074/jbc.M117.784371

Chen, F. M. (1992). Sr2+ facilitates intermolecular G-quadruplex formation of telomeric sequences. *Biochemistry*, *31*(15), 3769-3776. https://doi.org/10.1021/bi00130a006

Chen, L., Willis, S. N., Wei, A., Smith, B. J., Fletcher, J. I., Hinds, M. G., Colman, P. M., Day, C. L., Adams, J. M., & Huang, D. C. S. (2005). Differential targeting of prosurvival Bcl-2 proteins by their BH3-only ligands allows complementary apoptotic function. *Molecular Cell*, *17*(3), 393-403. https://doi.org/10.1016/j.molcel.2004.12.030

Clark, T. A., Sugnet, C. W., & Ares, M. (2002). Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. *Science (New York, N.Y.)*, *296*(5569), 907-910. https://doi.org/10.1126/science.1069415

Cleary, M. L., Smith, S. D., & Sklar, J. (1986). Cloning and structural analysis of cDNAs for bcl-2 and a hybrid bcl-2/immunoglobulin transcript resulting from the t(14;18) translocation. *Cell*, *47*(1), 19-28. https://doi.org/10.1016/0092-8674(86)90362-4

Cloutier, A., Shkreta, L., Toutant, J., Durand, M., Thibault, P., & Chabot, B. (2018). hnRNP A1/A2 and Sam68 collaborate with SRSF10 to control the alternative splicing response to oxaliplatin-mediated DNA damage. *Scientific Reports*, *8*, 2206. https://doi.org/10.1038/s41598-018-20360-x

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 Research*, *38*(16), 5569-5580. https://doi.org/10.1093/nar/gkq259

Czabotar, P. E., Westphal, D., Dewson, G., Ma, S., Hockings, C., Fairlie, W. D., Lee, E. F., Yao, S., Robin, A. Y., Smith, B. J., Huang, D. C. S., Kluck, R. M., Adams, J. M., & Colman, P. M. (2013). Bax crystal structures reveal how BH3 domains activate Bax and nucleate its oligomerization to induce apoptosis. *Cell*, *152*(3), 519-531. https://doi.org/10.1016/j.cell.2012.12.031

de Jong, Y., Monderer, D., Brandinelli, E., Monchanin, M., van den Akker, B. E., van Oosterwijk, J. G., Blay, J. Y., Dutour, A., & Bovée, J. V. M. G. (2018). Bcl-xl as the most promising Bcl-2 family member in targeted treatment of chondrosarcoma. *Oncogenesis*, *7*(9), 74. https://doi.org/10.1038/s41389-018-0084-0

Design and Evolution of a Miniature Bcl-2 Binding Protein—Chin—2001—Angewandte Chemie International Edition—Wiley Online Library. (s. d.). Consulté 29 août 2023, à l'adresse https://onlinelibrary-wiley-com.proxy.insermbiblio.inist.fr/doi/10.1002/1521-3773(20011015)40:20%3C3806::AID-ANIE3806%3E3.0.CO;2-B

de Vos, S., Leonard, J. P., Friedberg, J. W., Zain, J., Dunleavy, K., Humerickhouse, R., Hayslip, J., Pesko, J., & Wilson, W. H. (2021). Safety and efficacy of navitoclax, a BCL-2 and BCL-XL inhibitor, in patients with relapsed or refractory lymphoid malignancies : Results from a phase 2a study. *Leukemia*

& lymphoma, *62*(4), 810-818. https://doi.org/10.1080/10428194.2020.1845332

Dewson, G., Kratina, T., Czabotar, P., Day, C. L., Adams, J. M., & Kluck, R. M. (2009). Bak activation for apoptosis involves oligomerization of dimers via their alpha6 helices. *Molecular Cell*, *36*(4), 696-703. https://doi.org/10.1016/j.molcel.2009.11.008

Dewson, G., Ma, S., Frederick, P., Hockings, C., Tan, I., Kratina, T., & Kluck, R. M. (2012). Bax dimerizes via a symmetric BH3:groove interface during apoptosis. *Cell Death and Differentiation*, *19*(4), 661-670. https://doi.org/10.1038/cdd.2011.138

Dickerhoff, J., Dai, J., & Yang, D. (2021). Structural recognition of the MYC promoter G-quadruplex by a quinoline derivative : Insights into molecular targeting of parallel G-quadruplexes. *Nucleic Acids Research*, *49*(10), 5905-5915. https://doi.org/10.1093/nar/gkab330

Didiot, M.-C., Tian, Z., Schaeffer, C., Subramanian, M., Mandel, J.-L., & Moine, H. (2008). The G-quartet containing FMRP binding site in FMR1 mRNA is a potent exonic splicing enhancer. *Nucleic Acids Research*, *36*(15), 4902-4912. https://doi.org/10.1093/nar/gkn472

Dominguez, C., Fisette, J.-F., Chabot, B., & Allain, F. H.-T. (2010). Structural basis of G-tract recognition and encaging by hnRNP F quasi-RRMs. *Nature Structural & Molecular Biology*, *17*(7), 853-861. https://doi.org/10.1038/nsmb.1814

Dou, Z., Zhao, D., Chen, X., Xu, C., Jin, X., Zhang, X., Wang, Y., Xie, X., Li, Q., Di, C., & Zhang, H. (2021). Aberrant Bcl-x splicing in cancer : From molecular mechanism to therapeutic modulation. *Journal of Experimental & Clinical*

Cancer Research: CR, *40*(1), 194. https://doi.org/10.1186/s13046-021-02001-w

Dumas, L., Herviou, P., Dassi, E., Cammas, A., & Millevoi, S. (2021). G-Quadruplexes in RNA Biology : Recent Advances and Future Directions. *Trends in Biochemical Sciences*, *46*(4), 270-283. https://doi.org/10.1016/j.tibs.2020.11.001

Edlich, F., Banerjee, S., Suzuki, M., Cleland, M. M., Arnoult, D., Wang, C., Neutzner, A., Tjandra, N., & Youle, R. J. (2011). Bcl-x(L) retrotranslocates Bax from the mitochondria into the cytosol. *Cell*, *145*(1), 104-116. https://doi.org/10.1016/j.cell.2011.02.034

Eno, C. O., Eckenrode, E. F., Olberding, K. E., Zhao, G., White, C., & Li, C. (2012). Distinct roles of mitochondria- and ER-localized Bcl-xL in apoptosis resistance and Ca2+ homeostasis. *Molecular Biology of the Cell*, *23*(13), 2605-2618. https://doi.org/10.1091/mbc.E12-02-0090

Epstein, M. A., Achong, B. G., & Barr, Y. M. (1964). VIRUS PARTICLES IN CULTURED LYMPHOBLASTS FROM BURKITT'S LYMPHOMA. *The Lancet*, *283*(7335), 702-703. https://doi.org/10.1016/S0140-6736(64)91524-7

Ester, C., & Uetz, P. (2008). The FF domains of yeast U1 snRNP protein Prp40 mediate interactions with Luc7 and Snu71. *BMC Biochemistry*, *9*, 29. https://doi.org/10.1186/1471-2091-9-29

Ethell, D. W., & Buhler, L. A. (2003). Fas ligand-mediated apoptosis in degenerative disorders of the brain. *Journal of Clinical Immunology*, *23*(6), 439-446. https://doi.org/10.1023/b:joci.0000010420.96419.a8

Fay, M. M., Lyons, S. M., & Ivanov, P. (2017). RNA G-quadruplexes in biology : Principles and molecular mechanisms. *Journal of molecular biology*, *429*(14), 2127-2147. https://doi.org/10.1016/j.jmb.2017.05.017

Fedor, M. J. (2002). The role of metal ions in RNA catalysis. *Current Opinion in Structural Biology*, *12*(3), 289-295. https://doi.org/10.1016/S0959-440X(02)00324-X

Fortes, P., Longman, D., McCracken, S., Ip, J. Y., Poot, R., Mattaj, I. W., Caceres, J. F., & Blencowe, B. J. (2007). Identification and characterization of RED120 : A conserved PWI domain protein with links to splicing and 30-end formation. *FEBS Letters*.

Frees, S., Menendez, C., Crum, M., & Bagga, P. S. (2014). QGRS-Conserve : A computational method for discovering evolutionarily conserved G-quadruplex motifs. *Human Genomics*, *8*, 8. https://doi.org/10.1186/1479-7364-8-8

Galteland, E., Sivertsen, E. A., Svendsrud, D. H., Smedshammer, L., Kresse, S. H., Meza-Zepeda, L. A., Myklebost, O., Suo, Z., Mu, D., Deangelis, P. M., & Stokke, T. (2005). Translocation t(14;18) and gain of chromosome 18/BCL2 : Effects on BCL2 expression and apoptosis in B-cell non-Hodgkin's lymphomas. *Leukemia*, *19*(12), 2313-2323. https://doi.org/10.1038/sj.leu.2403954

García-Sáez, A. J., Ries, J., Orzáez, M., Pérez-Payà, E., & Schwille, P. (2009). Membrane promotes tBID interaction with BCL(XL). *Nature Structural & Molecular Biology*, *16*(11), 1178-1185. https://doi.org/10.1038/nsmb.1671

Garneau, D., Revil, T., Fisette, J.-F., & Chabot, B. (2005a). Heterogeneous nuclear ribonucleoprotein F/H proteins modulate the alternative splicing of the apoptotic mediator Bcl-x. *The Journal of Biological Chemistry*, *280*(24), 22641-22650. https://doi.org/10.1074/jbc.M501070200

Garneau, D., Revil, T., Fisette, J.-F., & Chabot, B. (2005b). Heterogeneous nuclear ribonucleoprotein F/H proteins modulate the alternative splicing of the

apoptotic mediator Bcl-x. *The Journal of Biological Chemistry*, *280*(24), 22641-22650. https://doi.org/10.1074/jbc.M501070200

Ge, Y., Schuster, M. B., Pundhir, S., Rapin, N., Bagger, F. O., Sidiropoulos, N., Hashem, N., & Porse, B. T. (2019). The splicing factor RBM25 controls MYC activity in acute myeloid leukemia. *Nature Communications*, *10*(1), 172. https://doi.org/10.1038/s41467-018-08076-y

Gellert, M., Lipsett, M. N., & Davies, D. R. (1962). Helix formation by guanylic acid. *Proceedings of the National Academy of Sciences of the United States of America*, *48*(12), 2013-2018. https://doi.org/10.1073/pnas.48.12.2013

Ghosh, M., & Singh, M. (2018). RGG-box in hnRNPA1 specifically recognizes the telomere G-quadruplex DNA and enhances the G-quadruplex unfolding ability of UP1 domain. *Nucleic Acids Research*, *46*(19), 10246-10261. https://doi.org/10.1093/nar/gky854

Goldstein, J. C., Waterhouse, N. J., Juin, P., Evan, G. I., & Green, D. R. (2000). The coordinate release of cytochrome c during apoptosis is rapid, complete and kinetically invariant. *Nature Cell Biology*, *2*(3), 156-162. https://doi.org/10.1038/35004029

González, V., Guo, K., Hurley, L., & Sun, D. (2009). Identification and characterization of nucleolin as a c-myc G-quadruplex-binding protein. *The Journal of Biological Chemistry*, *284*(35), 23622-23635. https://doi.org/10.1074/jbc.M109.018028

González, V., & Hurley, L. H. (2010). The c-MYC NHE III(1): Function and regulation. *Annual Review of Pharmacology and Toxicology*, *50*, 111-129. https://doi.org/10.1146/annurev.pharmtox.48.113006.094649

Guiset Miserachs, H., Donghi, D., Börner, R., Johannsen, S., & Sigel, R. K. O. (2016). Distinct differences in metal ion specificity of RNA and DNA G-

quadruplexes. *JBIC Journal of Biological Inorganic Chemistry*, *21*(8), 975-986. https://doi.org/10.1007/s00775-016-1393-4

Guo, A.-J., Wang, F.-J., Ji, Q., Geng, H.-W., Yan, X., Wang, L.-Q., Tie, W.-W., Liu, X.-Y., Thorne, R. F., Liu, G., & Xu, A.-M. (2021). Proteome Analyses Reveal S100A11, S100P, and RBM25 Are Tumor Biomarkers in Colorectal Cancer. *Proteomics Clin. Appl.*

Hai, Y., Cao, W., Liu, G., Hong, S.-P., Elela, S. A., Klinck, R., Chu, J., & Xie, J. (2008). A G-tract element in apoptotic agents-induced alternative splicing. *Nucleic Acids Research*, *36*(10), 3320-3331. https://doi.org/10.1093/nar/gkn207

Hardin, C. C., Watson, T., Corregan, M., & Bailey, C. (1992). Cation-dependent transition between the quadruplex and Watson-Crick hairpin forms of d(CGCG3GCG). *Biochemistry*, *31*(3), 833-841. https://doi.org/10.1021/bi00118a028

Henderson, E., Hardin, C. C., Walk, S. K., Tinoco, I., & Blackburn, E. H. (1987). Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine·guanine base pairs. *Cell*, *51*(6), 899-908. https://doi.org/10.1016/0092-8674(87)90577-0

Herviou, P., Le Bras, M., Dumas, L., Hieblot, C., Gilhodes, J., Cioci, G., Hugnot, J.-P., Ameadan, A., Guillonneau, F., Dassi, E., Cammas, A., & Millevoi, S. (2020). hnRNP H/F drive RNA G-quadruplex-mediated translation linked to genomic instability and therapy resistance in glioblastoma. *Nature Communications*, *11*(1), 2661. https://doi.org/10.1038/s41467-020-16168-x

Holinger, E. P., Chittenden, T., & Lutz, R. J. (1999). Bak BH3 peptides antagonize Bcl-xL function and induce apoptosis through cytochrome c-independent

activation of caspases. *The Journal of Biological Chemistry*, *274*(19), 13298-13304. https://doi.org/10.1074/jbc.274.19.13298

Huang, H., Zhang, J., Harvey, S. E., Hu, X., & Cheng, C. (2017). RNA G-quadruplex secondary structure promotes alternative splicing via the RNA-binding protein hnRNPF. *Genes & Development*, *31*(22), 2296-2309. https://doi.org/10.1101/gad.305862.117

Huppert, J. L., & Balasubramanian, S. (2007). G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Research*, *35*(2), 406-413. https://doi.org/10.1093/nar/gkl1057

Inoue, A., Yamamoto, N., Kimura, M., Nishio, K., Yamane, H., & Nakajima, K. (2014). RBM10 regulates alternative splicing. *FEBS Letters*, *588*(6), 942-947. https://doi.org/10.1016/j.febslet.2014.01.052

Jin, Z., McDonald, E. R., Dicker, D. T., & El-Deiry, W. S. (2004). Deficient tumor necrosis factor-related apoptosis-inducing ligand (TRAIL) death receptor transport to the cell surface in human colon cancer cells selected for resistance to TRAIL-induced apoptosis. *The Journal of Biological Chemistry*, *279*(34), 35829-35839. https://doi.org/10.1074/jbc.M405538200

Joachimi, A., Benz, A., & Hartig, J. S. (2009). A comparison of DNA and RNA quadruplex structures and stabilities. *Bioorganic & Medicinal Chemistry*, *17*(19), 6811-6815. https://doi.org/10.1016/j.bmc.2009.08.043

Kerr, J. F. R., Wyllie, A. H., & Currie, A. R. (1972). Apoptosis : A Basic Biological Phenomenon with Wideranging Implications in Tissue Kinetics. *British Journal of Cancer*, *26*(4), Article 4. https://doi.org/10.1038/bjc.1972.33

Kim, H., Rafiuddin-Shah, M., Tu, H.-C., Jeffers, J. R., Zambetti, G. P., Hsieh, J. J.-D., & Cheng, E. H.-Y. (2006). Hierarchical regulation of mitochondrion-

dependent apoptosis by BCL-2 subfamilies. *Nature Cell Biology*, *8*(12), 1348-1358. https://doi.org/10.1038/ncb1499

Kim, N. W., Piatyszek, M. A., Prowse, K. R., Harley, C. B., West, M. D., Ho, P. L., Coviello, G. M., Wright, W. E., Weinrich, S. L., & Shay, J. W. (1994). Specific association of human telomerase activity with immortal cells and cancer. *Science (New York, N.Y.)*, *266*(5193), 2011-2015. https://doi.org/10.1126/science.7605428

Kischkel, F. C., Hellbardt, S., Behrmann, I., Germer, M., Pawlita, M., Krammer, P. H., & Peter, M. E. (1995). Cytotoxicity-dependent APO-1 (Fas/CD95)-associated proteins form a death-inducing signaling complex (DISC) with the receptor. *The EMBO Journal*, *14*(22), 5579-5588. https://doi.org/10.1002/j.1460-2075.1995.tb00245.x

Kosiol, N., Juranek, S., Brossart, P., Heine, A., & Paeschke, K. (2021). G-quadruplexes : A promising target for cancer therapy. *Molecular Cancer*, *20*(1), 40. https://doi.org/10.1186/s12943-021-01328-4

Kurosaka, K., Takahashi, M., Watanabe, N., & Kobayashi, Y. (2003). Silent cleanup of very early apoptotic cells by macrophages. *Journal of Immunology (Baltimore, Md.: 1950)*, *171*(9), 4672-4679. https://doi.org/10.4049/jimmunol.171.9.4672

Lee, E. F., Czabotar, P. E., Smith, B. J., Deshayes, K., Zobel, K., Colman, P. M., & Fairlie, W. D. (2007). Crystal structure of ABT-737 complexed with Bcl-xL : Implications for selectivity of antagonists of the Bcl-2 family. *Cell Death and Differentiation*, *14*(9), 1711-1713. https://doi.org/10.1038/sj.cdd.4402178

Lee, J., Zhou, J., Zheng, X., Cho, S., Moon, H., Loh, T. J., Jo, K., & Shen, H. (2012). Identification of a novel cis-element that regulates alternative splicing of

Bcl-x pre-mRNA. *Biochemical and Biophysical Research Communications*, *420*(2), 467-472. https://doi.org/10.1016/j.bbrc.2012.03.029

Lee, Y.-C., Wang, L.-J., Huang, C.-H., Chiou, J.-T., Shi, Y.-J., & Chang, L.-S. (2020). Inhibition of EGFR pathway promotes the cytotoxicity of ABT-263 in human leukemia K562 cells by blocking MCL1 upregulation. *Biochemical Pharmacology*, *178*, 114047. https://doi.org/10.1016/j.bcp.2020.114047

Lejault, P., Prudent, L., Terrier, M.-P., & Perreault, J.-P. (2023). Small molecule chaperones facilitate the folding of RNA G-quadruplexes. *Biochimie*, S0300-9084(23)00207-9. https://doi.org/10.1016/j.biochi.2023.08.016

Letai, A., Bassik, M. C., Walensky, L. D., Sorcinelli, M. D., Weiler, S., & Korsmeyer, S. J. (2002). Distinct BH3 domains either sensitize or activate mitochondrial apoptosis, serving as prototype cancer therapeutics. *Cancer Cell*, *2*(3), 183-192. https://doi.org/10.1016/s1535-6108(02)00127-7

Levitskaya, J., Coram, M., Levitsky, V., Imreh, S., Steigerwald-Mullen, P. M., Klein, G., Kurilla, M. G., & Masucci, M. G. (1995). Inhibition of antigen processing by the internal repeat region of the Epstein-Barr virus nuclear antigen-1. *Nature*, *375*(6533), 685-688. https://doi.org/10.1038/375685a0

Lewis, B. P., Green, R. E., & Brenner, S. E. (2003). Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(1), 189-192. https://doi.org/10.1073/pnas.0136770100

Li, C. Y., Chu, J. Y., Yu, J. K., Huang, X. Q., Liu, X. J., Shi, L., Che, Y. C., & Xie, J. Y. (2004). Regulation of alternative splicing of Bcl-x by IL-6, GM-CSF and TPA. *Cell Research*, *14*(6), 473-479. https://doi.org/10.1038/sj.cr.7290250

Li, X. (2017). RBM25 Mediates Abiotic Responses in Plants. *Frontiers in Plant Science*, *8*.

Li, Y., Zhang, X., Gao, Y., Shi, J., Tang, L., & Sui, G. (2018). G-quadruplexes in the BAP1 promoter positively regulate its expression. *Experimental Cell Research*, *369*(1), 147-157. https://doi.org/10.1016/j.yexcr.2018.05.016

Lindenboim, L., Borner, C., & Stein, R. (2001). Bcl-x(S) can form homodimers and heterodimers and its apoptotic activity requires localization of Bcl-x(S) to the mitochondria and its BH3 and loop domains. *Cell Death and Differentiation*, *8*(9), 933-942. https://doi.org/10.1038/sj.cdd.4400888

Lista, M. J., Martins, R. P., Angrand, G., Quillévéré, A., Daskalogianni, C., Voisset, C., Teulade-Fichou, M.-P., Fåhraeus, R., & Blondel, M. (2017). A yeast model for the mechanism of the Epstein-Barr virus immune evasion identifies a new therapeutic target to interfere with the virus stealthiness. *Microbial Cell (Graz, Austria)*, *4*(9), 305-307. https://doi.org/10.15698/mic2017.09.590

Lista, M. J., Martins, R. P., Billant, O., Contesse, M.-A., Findakly, S., Pochard, P., Daskalogianni, C., Beauvineau, C., Guetta, C., Jamin, C., Teulade-Fichou, M.-P., Fåhraeus, R., Voisset, C., & Blondel, M. (2017). Nucleolin directly mediates Epstein-Barr virus immune evasion through binding to G-quadruplexes of EBNA1 mRNA. *Nature Communications*, *8*, 16043. https://doi.org/10.1038/ncomms16043

Liu, R., Page, C., Beidler, D. R., Wicha, M. S., & Núñez, G. (1999). Overexpression of Bcl-xL Promotes Chemotherapy Resistance of Mammary Tumors in a Syngeneic Mouse Model. *The American Journal of Pathology*, *155*(6), 1861-1867.

Llambi, F., Moldoveanu, T., Tait, S. W. G., Bouchier-Hayes, L., Temirov, J., McCormick, L. L., Dillon, C. P., & Green, D. R. (2011). A unified model of

mammalian BCL-2 protein family interactions at the mitochondria. *Molecular Cell*, *44*(4), 517-531. https://doi.org/10.1016/j.molcel.2011.10.001

López-Hoyos, M., Carrió, R., Merino, J., & Merino, R. (1998). Regulation of B cell apoptosis by Bcl-2 and Bcl-XL and its role in the development of autoimmune diseases (Review). *International Journal of Molecular Medicine*, *1*(2), 475-483. https://doi.org/10.3892/ijmm.1.2.475

Lovell, J. F., Billen, L. P., Bindner, S., Shamas-Din, A., Fradin, C., Leber, B., & Andrews, D. W. (2008). Membrane binding by tBid initiates an ordered series of events culminating in membrane permeabilization by Bax. *Cell*, *135*(6), 1074-1084. https://doi.org/10.1016/j.cell.2008.11.010

Luo, Y., Granzhan, A., Marquevielle, J., Cucchiarini, A., Lacroix, L., Amrane, S., Verga, D., & Mergny, J.-L. (2022). Guidelines for G-quadruplexes : I. In vitro characterization. *Biochimie*, S0300-9084(22)00344-3. https://doi.org/10.1016/j.biochi.2022.12.019

Lyu, K., Chow, E. Y.-C., Mou, X., Chan, T.-F., & Kwok, C. K. (2021). RNA G-quadruplexes (rG4s) : Genomics and biological functions. *Nucleic Acids Research*, *49*(10), 5426-5450. https://doi.org/10.1093/nar/gkab187

Makarov, V. L., Hirose, Y., & Langmore, J. P. (1997). Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell*, *88*(5), 657-666. https://doi.org/10.1016/s0092-8674(00)81908-x

Marcel, V., Tran, P. L. T., Sagne, C., Martel-Planche, G., Vaslin, L., Teulade-Fichou, M.-P., Hall, J., Mergny, J.-L., Hainaut, P., & Van Dyck, E. (2011). G-quadruplex structures in TP53 intron 3 : Role in alternative splicing and in production of p53 mRNA isoforms. *Carcinogenesis*, *32*(3), 271-278. https://doi.org/10.1093/carcin/bgq253

Massiello, A., Roesser, J. R., & Chalfant, C. E. (2006). SAP155 Binds to ceramide-responsive RNA cis-element 1 and regulates the alternative 5′ splice site selection of Bcl-x pre-mRNA. *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology*, *20*(10), 1680-1682. https://doi.org/10.1096/fj.05-5021fje

Massiello, A., Salas, A., Pinkerman, R. L., Roddy, P., Roesser, J. R., & Chalfant, C. E. (2004). Identification of two RNA cis-elements that function to regulate the 5′ splice site selection of Bcl-x pre-mRNA in response to ceramide. *The Journal of Biological Chemistry*, *279*(16), 15799-15804. https://doi.org/10.1074/jbc.M313950200

Mathad, R. I., Hatzakis, E., Dai, J., & Yang, D. (2011). c-MYC promoter G-quadruplex formed at the 5′-end of NHE III1 element : Insights into biological relevance and parallel-stranded G-quadruplex stability. *Nucleic Acids Research*, *39*(20), 9023-9033. https://doi.org/10.1093/nar/gkr612

Mendoza, O., Bourdoncle, A., Boulé, J.-B., Brosh, R. M., & Mergny, J.-L. (2016). G-quadruplexes and helicases. *Nucleic Acids Research*, *44*(5), 1989-2006. https://doi.org/10.1093/nar/gkw079

Mergny, J.-L., & Lacroix, L. (2009). UV Melting of G-Quadruplexes. *Current Protocols in Nucleic Acid Chemistry*, *Chapter 17*, 17.1.1-17.1.15. https://doi.org/10.1002/0471142700.nc1701s37

Michelle, L., Cloutier, A., Toutant, J., Shkreta, L., Thibault, P., Durand, M., Garneau, D., Gendron, D., Lapointe, E., Couture, S., Le Hir, H., Klinck, R., Elela, S. A., Prinos, P., & Chabot, B. (2012). Proteins associated with the exon junction complex also control the alternative splicing of apoptotic regulators. *Molecular and Cellular Biology*, *32*(5), 954-967. https://doi.org/10.1128/MCB.06130-11

Miglietta, G., Russo, M., & Capranico, G. (2020). G-quadruplex–R-loop interactions and the mechanism of anticancer G-quadruplex binders. *Nucleic Acids Research*, *48*(21), 11942-11957. https://doi.org/10.1093/nar/gkaa944

Millevoi, S., Moine, H., & Vagner, S. (2012). G-quadruplexes in RNA biology. *Wiley Interdisciplinary Reviews. RNA*, *3*(4), 495-507. https://doi.org/10.1002/wrna.1113

Mitteaux, J., Lejault, P., Wojciechowski, F., Joubert, A., Boudon, J., Desbois, N., Gros, C. P., Hudson, R. H. E., Boulé, J.-B., Granzhan, A., & Monchaud, D. (2021). Identifying G-Quadruplex-DNA-Disrupting Small Molecules. *Journal of the American Chemical Society*, *143*(32), 12567-12577. https://doi.org/10.1021/jacs.1c04426

Moore, M. J., & Sharp, P. A. (1993). Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature*, *365*(6444), 364-368. https://doi.org/10.1038/365364a0

Muchmore, S. W., Sattler, M., Liang, H., Meadows, R. P., Harlan, J. E., Yoon, H. S., Nettesheim, D., Chang, B. S., Thompson, C. B., Wong, S. L., Ng, S. L., & Fesik, S. W. (1996). X-ray and NMR structure of human Bcl-xL, an inhibitor of programmed cell death. *Nature*, *381*(6580), 335-341. https://doi.org/10.1038/381335a0

Nakanishi, C., & Seimiya, H. (2020). G-quadruplex in cancer biology and drug discovery. *Biochemical and Biophysical Research Communications*, *531*(1), 45-50. https://doi.org/10.1016/j.bbrc.2020.03.178

Noyes, A. M., Zhou, A., Gao, G., Gu, L., Day, S., Wasserstrom, J. A., & Dudley, S. C. (2017). Abnormal sodium channel mRNA splicing in hypertrophic cardiomyopathy. *International Journal of Cardiology*.

Oganesian, L., & Bryan, T. M. (2007). Physiological relevance of telomeric G-quadruplex formation : A potential drug target. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *29*(2), 155-165. https://doi.org/10.1002/bies.20523

Ohnmacht, S. A., Marchetti, C., Gunaratnam, M., Besser, R. J., Haider, S. M., Di Vita, G., Lowe, H. L., Mellinas-Gomez, M., Diocou, S., Robson, M., Šponer, J., Islam, B., Pedley, R. B., Hartley, J. A., & Neidle, S. (2015). A G-quadruplex-binding compound showing anti-tumour activity in an in vivo model for pancreatic cancer. *Scientific Reports*, *5*, 11385. https://doi.org/10.1038/srep11385

Papoff, G., Hausler, P., Eramo, A., Pagano, M. G., Di Leve, G., Signore, A., & Ruberti, G. (1999). Identification and characterization of a ligand-independent oligomerization domain in the extracellular region of the CD95 death receptor. *The Journal of Biological Chemistry*, *274*(53), 38241-38250. https://doi.org/10.1074/jbc.274.53.38241

Paramasivam, M., Membrino, A., Cogoi, S., Fukuda, H., Nakagama, H., & Xodo, L. E. (2009). Protein hnRNP A1 and its derivative Up1 unfold quadruplex DNA in the human KRAS promoter : Implications for transcription. *Nucleic Acids Research*, *37*(9), 2841-2853. https://doi.org/10.1093/nar/gkp138

Park, H.-A., Licznerski, P., Alavian, K. N., Shanabrough, M., & Jonas, E. A. (2015). Bcl-xL is necessary for neurite outgrowth in hippocampal neurons. *Antioxidants & Redox Signaling*, *22*(2), 93-108. https://doi.org/10.1089/ars.2013.5570

Paronetto, M. P., Achsel, T., Massiello, A., Chalfant, C. E., & Sette, C. (2007). The RNA-binding protein Sam68 modulates the alternative splicing of Bcl-x. *The*

Journal of Cell Biology, *176*(7), 929-939. https://doi.org/10.1083/jcb.200701005

Pattanayak, R., Barua, A., Das, A., Chatterjee, T., Pathak, A., Choudhury, P., Sen, S., Saha, P., & Bhattacharyya, M. (2018). Porphyrins to restrict progression of pancreatic cancer by stabilizing KRAS G-quadruplex : In silico, in vitro and in vivo validation of anticancer strategy. *European Journal of Pharmaceutical Sciences: Official Journal of the European Federation for Pharmaceutical Sciences*, *125*, 39-53. https://doi.org/10.1016/j.ejps.2018.09.011

Pedrotti, S., Busà, R., Compagnucci, C., & Sette, C. (2012). The RNA recognition motif protein RBM11 is a novel tissue-specific splicing regulator. *Nucleic Acids Research*, *40*(3), 1021-1032. https://doi.org/10.1093/nar/gkr819

Renehan, A. G., Booth, C., & Potten, C. S. (2001). What is apoptosis, and why is it important? *BMJ (Clinical Research Ed.)*, *322*(7301), 1536-1538. https://doi.org/10.1136/bmj.322.7301.1536

Revil, T., Pelletier, J., Toutant, J., Cloutier, A., & Chabot, B. (2009). Heterogeneous nuclear ribonucleoprotein K represses the production of pro-apoptotic Bcl-xS splice isoform. *The Journal of Biological Chemistry*, *284*(32), 21458-21467. https://doi.org/10.1074/jbc.M109.019711

Roy, S. W., & Gilbert, W. (2006). The evolution of spliceosomal introns : Patterns, puzzles and progress. *Nature Reviews. Genetics*, *7*(3), 211-221. https://doi.org/10.1038/nrg1807

Sakahira, H., Enari, M., & Nagata, S. (1998). Cleavage of CAD inhibitor in CAD activation and DNA degradation during apoptosis. *Nature*, *391*(6662), 96-99. https://doi.org/10.1038/34214

Scalabrin, M., Nadai, M., Tassinari, M., Lago, S., Doria, F., Frasson, I., Freccero, M., & Richter, S. N. (2021). Selective Recognition of a Single HIV-1 G-

Quadruplex by Ultrafast Small-Molecule Screening. *Analytical Chemistry*, *93*(46), 15243-15252. https://doi.org/10.1021/acs.analchem.0c04106

Scherr, A.-L., Gdynia, G., Salou, M., Radhakrishnan, P., Duglova, K., Heller, A., Keim, S., Kautz, N., Jassowicz, A., Elssner, C., He, Y.-W., Jaeger, D., Heikenwalder, M., Schneider, M., Weber, A., Roth, W., Schulze-Bergkamen, H., & Koehler, B. C. (2016). Bcl-xL is an oncogenic driver in colorectal cancer. *Cell Death & Disease*, *7*(8), e2342. https://doi.org/10.1038/cddis.2016.233

Serrano, P., Aubol, B. E., Keshwani, M. M., Forli, S., Ma, C.-T., Dutta, S. K., Geralt, M., Wüthrich, K., & Adams, J. A. (2016). Directional Phosphorylation and Nuclear Transport of the Splicing Factor SRSF1 Is Regulated by an RNA Recognition Motif. *Journal of Molecular Biology*, *428*(11), 2430-2445. https://doi.org/10.1016/j.jmb.2016.04.009

Sharp, P. A., & Burge, C. B. (1997). Classification of Introns : U2-Type or U12-Type. *Cell*, *91*(7), 875-879. https://doi.org/10.1016/S0092-8674(00)80479-1

Shkreta, L., Toutant, J., Durand, M., Manley, J. L., & Chabot, B. (2016). SRSF10 Connects DNA Damage to the Alternative Splicing of Transcripts Encoding Apoptosis, Cell-Cycle Control, and DNA Repair Factors. *Cell Reports*, *17*(8), 1990-2003. https://doi.org/10.1016/j.celrep.2016.10.071

Shu, H., Zhang, R., Xiao, K., Yang, J., & Sun, X. (2022). G-Quadruplex-Binding Proteins : Promising Targets for Drug Design. *Biomolecules*, *12*(5), 648. https://doi.org/10.3390/biom12050648

Shultz, J. C., Vu, N., Shultz, M. D., Mba, M.-U. U., Shapiro, B. A., & Chalfant, C. E. (2012). The Proto-oncogene PKCι regulates the alternative splicing of Bcl-x pre-mRNA. *Molecular Cancer Research: MCR*, *10*(5), 660-669. https://doi.org/10.1158/1541-7786.MCR-11-0363

Singh, N. N., Singh, R. N., & Androphy, E. J. (2007). Modulating role of RNA structure in alternative splicing of a critical exon in the spinal muscular atrophy genes. *Nucleic Acids Research*, *35*(2), 371-389. https://doi.org/10.1093/nar/gkl1050

Skwarska, A., & Konopleva, M. (2023). BCL-xL Targeting to Induce Apoptosis and to Eliminate Chemotherapy-Induced Senescent Tumor Cells : From Navitoclax to Platelet-Sparing BCL-xL PROTACs. *Cancer Research*, *83*(21), 3501-3503. https://doi.org/10.1158/0008-5472.CAN-23-2804

Srinivasula, S. M., Ahmad, M., Fernandes-Alnemri, T., & Alnemri, E. S. (1998). Autoactivation of procaspase-9 by Apaf-1-mediated oligomerization. *Molecular Cell*, *1*(7), 949-957. https://doi.org/10.1016/s1097-2765(00)80095-7

Stennicke, H. R., Jürgensmeier, J. M., Shin, H., Deveraux, Q., Wolf, B. B., Yang, X., Zhou, Q., Ellerby, H. M., Ellerby, L. M., Bredesen, D., Green, D. R., Reed, J. C., Froelich, C. J., & Salvesen, G. S. (1998). Pro-caspase-3 is a major physiologic target of caspase-8. *The Journal of Biological Chemistry*, *273*(42), 27084-27090. https://doi.org/10.1074/jbc.273.42.27084

Strahan, G. D., Keniry, M. A., & Shafer, R. H. (1998). NMR structure refinement and dynamics of the K+-[d(G3T4G3)]2 quadruplex via particle mesh Ewald molecular dynamics simulations. *Biophysical Journal*, *75*(2), 968-981. https://doi.org/10.1016/S0006-3495(98)77585-X

Su, H., Xu, J., Chen, Y., Wang, Q., Lu, Z., Chen, Y., Chen, K., Han, S., Fang, Z., Wang, P., Yuan, B.-F., & Zhou, X. (2021). Photoactive G-Quadruplex Ligand Identifies Multiple G-Quadruplex-Related Proteins with Extensive Sequence Tolerance in the Cellular Environment. *Journal of the American Chemical Society*, *143*(4), 1917-1923. https://doi.org/10.1021/jacs.0c10792

Suzuki, M., Youle, R. J., & Tjandra, N. (2000). Structure of Bax : Coregulation of dimer formation and intracellular localization. *Cell*, *103*(4), 645-654. https://doi.org/10.1016/s0092-8674(00)00167-7

Szymczyna, B. R., Bowman, J., McCracken, S., Pineda-Lucena, A., Lu, Y., Cox, B., Lambermon, M., Graveley, B. R., Arrowsmith, C. H., & Blencowe, B. J. (2003). Structure and function of the PWI motif : A novel nucleic acid-binding domain that facilitates pre-mRNA processing. *Genes & Development*, *17*(4), 461-475. https://doi.org/10.1101/gad.1060403

Tan, C., Dlugosz, P. J., Peng, J., Zhang, Z., Lapolla, S. M., Plafker, S. M., Andrews, D. W., & Lin, J. (2006). Auto-activation of the apoptosis protein Bax increases mitochondrial membrane permeability and is inhibited by Bcl-2. *The Journal of Biological Chemistry*, *281*(21), 14764-14775. https://doi.org/10.1074/jbc.M602374200

Thandapani, P., O'Connor, T. R., Bailey, T. L., & Richard, S. (2013). Defining the RGG/RG motif. *Molecular Cell*, *50*(5), 613-623. https://doi.org/10.1016/j.molcel.2013.05.021

Tse, C., Shoemaker, A. R., Adickes, J., Anderson, M. G., Chen, J., Jin, S., Johnson, E. F., Marsh, K. C., Mitten, M. J., Nimmer, P., Roberts, L., Tahir, S. K., Xiao, Y., Yang, X., Zhang, H., Fesik, S., Rosenberg, S. H., & Elmore, S. W. (2008). ABT-263 : A potent and orally bioavailable Bcl-2 family inhibitor. *Cancer Research*, *68*(9), 3421-3428. https://doi.org/10.1158/0008-5472.CAN-07-5836

Tu, J., Duan, M., Liu, W., Lu, N., Zhou, Y., Sun, X., & Lu, Z. (2021). Direct genome-wide identification of G-quadruplex structures by whole-genome resequencing. *Nature Communications*, *12*(1), 6014. https://doi.org/10.1038/s41467-021-26312-w

van Delft, M. F., Wei, A. H., Mason, K. D., Vandenberg, C. J., Chen, L., Czabotar, P. E., Willis, S. N., Scott, C. L., Day, C. L., Cory, S., Adams, J. M., Roberts, A. W., & Huang, D. C. S. (2006). The BH3 mimetic ABT-737 targets selective Bcl-2 proteins and efficiently induces apoptosis via Bak/Bax if Mcl-1 is neutralized. *Cancer Cell*, *10*(5), 389-399. https://doi.org/10.1016/j.ccr.2006.08.027

Varshney, D., Spiegel, J., Zyner, K., Tannahill, D., & Balasubramanian, S. (2020). The regulation and functions of DNA and RNA G-quadruplexes. *Nature Reviews. Molecular Cell Biology*, *21*(8), 459-474. https://doi.org/10.1038/s41580-020-0236-x

Voisset, C., Daskalogianni, C., Contesse, M.-A., Mazars, A., Arbach, H., Le Cann, M., Soubigou, F., Apcher, S., Fåhraeus, R., & Blondel, M. (2014). A yeast-based assay identifies drugs that interfere with immune evasion of the Epstein-Barr virus. *Disease Models & Mechanisms*, *7*(4), 435-444. https://doi.org/10.1242/dmm.014308

Wang, B., Ni, Z., Dai, X., Qin, L., Li, X., Xu, L., Lian, J., & He, F. (2014). The Bcl-2/xL inhibitor ABT-263 increases the stability of Mcl-1 mRNA and protein in hepatocellular carcinoma cells. *Molecular Cancer*, *13*, 98. https://doi.org/10.1186/1476-4598-13-98

Wang, Y., Chen, D., Qian, H., Tsai, Y. S., Shao, S., Liu, Q., Dominguez, D., & Wang, Z. (2014). The splicing factor RBM4 controls apoptosis, proliferation, and migration to suppress tumor progression. *Cancer Cell*, *26*(3), 374-389. https://doi.org/10.1016/j.ccr.2014.07.010

Wang, Y., & Patel, D. J. (1993). Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex. *Structure (London, England: 1993)*, *1*(4), 263-282. https://doi.org/10.1016/0969-2126(93)90015-9

Warren, C. F. A., Wong-Brown, M. W., & Bowden, N. A. (2019). BCL-2 family isoforms in apoptosis and cancer. *Cell Death & Disease*, *10*(3), 177. https://doi.org/10.1038/s41419-019-1407-6

Wei, M. C., Zong, W. X., Cheng, E. H., Lindsten, T., Panoutsakopoulou, V., Ross, A. J., Roth, K. A., MacGregor, G. R., Thompson, C. B., & Korsmeyer, S. J. (2001). Proapoptotic BAX and BAK : A requisite gateway to mitochondrial dysfunction and death. *Science (New York, N.Y.)*, *292*(5517), 727-730. https://doi.org/10.1126/science.1059108

Weldon, C., Behm-Ansmant, I., Hurley, L. H., Burley, G. A., Branlant, C., Eperon, I. C., & Dominguez, C. (2017). Identification of G-quadruplexes in long functional RNAs using 7-deazaguanine RNA. *Nature Chemical Biology*, *13*(1), 18-20. https://doi.org/10.1038/nchembio.2228

Weldon, C., Dacanay, J. G., Gokhale, V., Boddupally, P. V. L., Behm-Ansmant, I., Burley, G. A., Branlant, C., Hurley, L. H., Dominguez, C., & Eperon, I. C. (2018). Specific G-quadruplex ligands modulate the alternative splicing of Bcl-X. *Nucleic Acids Research*, *46*(2), 886-896. https://doi.org/10.1093/nar/gkx1122

Wilson, J. B., Manet, E., Gruffat, H., Busson, P., Blondel, M., & Fahraeus, R. (2018). EBNA1 : Oncogenic Activity, Immune Evasion and Biochemical Functions Provide Targets for Novel Therapeutic Strategies against Epstein-Barr Virus-Associated Cancers. *Cancers*, *10*(4), 109. https://doi.org/10.3390/cancers10040109

Wolter, K. G., Hsu, Y. T., Smith, C. L., Nechushtan, A., Xi, X. G., & Youle, R. J. (1997). Movement of Bax from the cytosol to mitochondria during apoptosis. *The Journal of Cell Biology*, *139*(5), 1281-1292. https://doi.org/10.1083/jcb.139.5.1281

Wu, G., Xing, Z., Tran, E. J., & Yang, D. (2019). DDX5 helicase resolves G-quadruplex and is involved in *MYC* gene transcriptional activation. *Proceedings of the National Academy of Sciences*, *116*(41), 20453-20461. https://doi.org/10.1073/pnas.1909047116

Xue, Y., Zhou, Y., Wu, T., Zhu, T., Ji, X., Kwon, Y.-S., Zhang, C., Yeo, G., Black, D. L., Sun, H., Fu, X.-D., & Zhang, Y. (2009). Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping. *Molecular Cell*, *36*(6), 996-1006. https://doi.org/10.1016/j.molcel.2009.12.003

Yan, K. K.-P., Obi, I., & Sabouri, N. (2021). The RGG domain in the C-terminus of the DEAD box helicases Dbp2 and Ded1 is necessary for G-quadruplex destabilization. *Nucleic Acids Research*, *49*(14), 8339-8354. https://doi.org/10.1093/nar/gkab620

Yang, Z., Qu, C.-B., Zhang, Y., Zhang, W.-F., Wang, D.-D., Gao, C.-C., Ma, L., Chen, J.-S., Liu, K.-L., Zheng, B., Zhang, X.-H., Zhang, M.-L., Wang, X.-L., Wen, J.-K., & Li, W. (2019). Dysregulation of p53-RBM25-mediated circAMOTL1L biogenesis contributes to prostate cancer progression through the circAMOTL1L-miR-193a-5p-Pcdha pathway. *Oncogene*, *38*(14), 2516-2532. https://doi.org/10.1038/s41388-018-0602-8

Yao, Y., Bobkov, A. A., Plesniak, L. A., & Marassi, F. M. (2009). Mapping the interaction of pro-apoptotic tBID with pro-survival BCL-XL. *Biochemistry*, *48*(36), 8704-8711. https://doi.org/10.1021/bi901171n

Yin, Y., Manoury, B., & Fåhraeus, R. (2003). Self-inhibition of synthesis and antigen presentation by Epstein-Barr virus-encoded EBNA1. *Science (New York, N.Y.)*, *301*(5638), 1371-1374. https://doi.org/10.1126/science.1088902

Zhou, A., Ou, A. C., Cho, A., Benz, E. J., & Huang, S.-C. (2008). Novel Splicing Factor RBM25 Modulates Bcl-x Pre-mRNA 5′ Splice Site Selection. *Molecular and Cellular Biology*, *28*(19), 5924-5936. https://doi.org/10.1128/MCB.00560-08

Zhou, F., Yang, Y., & Xing, D. (2011). Bcl-2 and Bcl-xL play important roles in the crosstalk between autophagy and apoptosis. *The FEBS Journal*, *278*(3), 403-413. https://doi.org/10.1111/j.1742-4658.2010.07965.x

COLLEGE SCIENCES
DOCTORAL DE LA VIE
BRETAGNE ET DE LA SANTE

UBO
Université de Bretagne Occidentale

Titre : Etude du rôle d'un G-quadruplex d'ARN dans l'épissage alternatif du gène anti-apoptotique *BCL*-x et exploitation comme cible thérapeutique pour lutter contre les cancers résistants aux chimiothérapies

Mots clés : Epissage alternatif, apoptose, cancers, G-quadruplex, *BCL-x*, RBM25

Résumé : L'apoptose joue un rôle crucial dans le contrôle de l'homéostasie cellulaire au cours du développement, tout au long de la vie comme dans certaines maladies comme les cancers pour lesquels nombre de traitements visent d'ailleurs à promouvoir l'apoptose des cellules tumorales. La fonction principale de la protéine Bcl-x est de réguler la voie intrinsèque de l'apoptose. Le gène *BCL-x* est soumis à un épissage alternatif qui produit deux isoformes aux fonctions antagonistes : la forme longue Bcl-xL anti-apoptotique, canonique, et la forme alternative courte Bcl-xS qui, elle, est pro-apoptotique. Dans certains cancers, la surexpression de Bcl-xL est corrélée à une survie augmentée des cellules et associée à une sensibilité moindre aux chimiothérapies. Récemment, des G-quadruplex (G4) ont été identifiés à proximité étroite des deux sites 5' d'épissages de *BCL-x* conduisant aux formes Bcl-xL et Bcl-xS. Nous avons mis en évidence une interaction spécifique entre le facteur d'épissage RBM25 et le rG4 GQ-2 localisé à proximité du site 5' d'épissage conduisant à l'isoforme pro-apoptotique Bcl-xS. Cette interaction est essentielle à la production de Bcl-xS et peut être ciblée par des ligands de G4 afin de moduler l'épissage de *BCL-x*. Ces résultats mettent en lumière un nouveau mécanisme de régulation de l'épissage alternatif de *BCL-x* et son utilisation potentielle comme cible thérapeutique, notamment pour les cancers résistants aux chimiothérapies.

Title: Study of the role of an RNA G-quadruplex in the alternative splicing of the anti-apoptotic *BCL-x* gene and exploitation as a therapeutic target for cancers resistant to chemotherapy

Keywords: Alternative splicing, apoptosis, cancers, G-quadruplex, *BCL-x*, RBM25

Résumé: Apoptosis is a highly regulated process involved in the control of cellular homeostasis during development, as well as lifelong or in some cancers for which many treatments aim to promote apoptosis of cancer cells. The main function of the Bcl-x protein is to regulate the intrinsic pathway of apoptosis. The *BCL-x* gene undergoes alternative splicing producing two isoforms with opposite functions: the long canonical anti-apoptotic Bcl-xL isoform, and the short alternative pro-apoptotic Bcl-xS isoform. Bcl-xL overexpression is frequently associated with increased cell survival and decreased sensitivity to chemotherapy in cancer. Recently, two G-quadruplex (G4) have been identified in close vicinity to either 5' splice sites of *BCL-x* pre-mRNA and their presence have been linked to alternative splicing. In this context, we have identified a specific interaction between the splicing factor RBM25 and the rG4 GQ-2 lying in close proximity with the splice site leading to the pro-apoptotic isoform Bcl-xS, and have demonstrated that this interaction can be targeted by G4 ligands to modulate *BCL-x* alternative splicing. Hence, we identified an actionable mechanism regulating *BCL-x* alternative splicing with the potential to identify molecules able to relieve chemoresistance of various cancers.