



HAL
open science

How the Epstein-Barr virus-encoded EBNA1 mRNA translation is regulated in cis by its mRNA dynamic structure and its nascent polypeptide

Alice Jia-Li Zheng

► **To cite this version:**

Alice Jia-Li Zheng. How the Epstein-Barr virus-encoded EBNA1 mRNA translation is regulated in cis by its mRNA dynamic structure and its nascent polypeptide. Human health and pathology. Université Paris Cité, 2021. English. NNT : 2021UNIP7132 . tel-03783685

HAL Id: tel-03783685

<https://theses.hal.science/tel-03783685>

Submitted on 22 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université de Paris

Ecole doctorale 561

HOB : Hématologie – Oncogénèse – Biothérapies

INSERM U1131

How the Epstein-Barr virus-encoded *EBNA1* mRNA translation is regulated *in cis* by its mRNA dynamic structure and its nascent polypeptide

Thèse de doctorat en oncogénèse

Par Alice ZHENG

Dirigé par Dr. Robin FAHRAEUS

Présentée et soutenue publiquement le 22 novembre 2021

Membres du jury / Jury members:

Dr. Stéphan VAGNER – Université Paris-Saclay

Président du jury

Dr. Stefania MILLEVOI – Université Toulouse III

Rapporteuse

Dr. Marie-Françoise O'DONOHUE – Université Toulouse III

Rapporteuse

Pr. Marc Blondel – Université de Bretagne Occidentale

Examineur

Dr. Isabelle DUSANTER-FOURT – Université de Paris

Examinatrice

Dr. Robin FAHRAEUS – Université de Paris

Directeur de thèse

“Ne regrettez pas, souvenez-vous.”

Portrait de la jeune fille en feu

Abstract

mRNA translation and protein synthesis are tightly regulated events in the cell. Mechanisms describing these key cellular events involve the mRNA sequence and its structure, the association of RNA-binding protein to the mRNA, as well as the quality of the polypeptide product encoded by the mRNA, assessed notably through ribosome-associated quality control.

In this context, the Epstein-Barr virus *EBNA1* (Epstein-Barr Nuclear Antigen 1) mRNA translation regulation is an interesting example. EBNA1 is known to be an essential protein for the Epstein-Barr virus (EBV) survival in the host cells. Even though EBNA1 is present in every infected cell, its protein levels are remarkably low. As EBNA1 is highly antigenic, it has been suggested that EBNA1 levels in the cells are maintained low enough to escape the immune system of the host, but sufficient to maintain EBV infection. This balance requires a tightly controlled EBNA1 production. Further studies showed that the GAR (glycine-alanine repeats) domain, located in the N-terminal part of EBNA1, triggers an *in cis* mechanism leading to the inhibition of the translation initiation of its own mRNA, without affecting translation of other mRNAs in the cell. Thus, the GAR domain of EBNA1 is a unique tool to study selective mRNA translation control without affecting general protein synthesis.

It was previously shown that RNA G4 (G-quadruplex) structures can be folded in the GAR-encoding mRNA. Numerous studies underlined the importance of these RNA structures in the regulation of *EBNA1* mRNA translation. Furthermore, the cellular factor nucleolin can interact with these RNA G4 structures, competitively with G4 ligands. However, it was also formerly demonstrated that the GAR peptide itself plays a role in controlling *in cis* the translation of *EBNA1* mRNA, rather than just the RNA sequence. The main focus of the study presented here is to shed light on how this translation event and the fate of the corresponding mRNA are regulated *in cis* by the mRNA and the encoded nascent polypeptide.

In line with the fact that RNA G4 structures are highly dynamic, we first show that GAR RNA G4-associated functions, namely mRNA localisation, translation, and ability to bind RNA-binding proteins, are dependent on the molecular context, *i.e.* their position in the mRNA, the RNA structures in their surrounding or the factors binding the mRNA, such as G4 ligands. We next demonstrate that translation of the *EBNA1* mRNA is necessary for nucleolin to bind it, meaning that the translation event modifies some properties of the *EBNA1* mRNA. In parallel, we show that the NACA, a subunit of the NAC chaperone complex, is detached from the ribosome and interacts with

the GAR polypeptide. Interestingly, NACA processing of the GAR differs from the polyQ processing, another highly intrinsically disordered domain. Finally, and unexpectedly, we show that translation initiation factor eIF4A1 is also a key player in the downregulation of the *EBNA1* mRNA translation, affecting the mRNA nucleolin-binding capacity. We propose that eIF4A1 acts as an RNA structure modulator, controlling the formation of ribonucleoprotein complex, here composed of nucleolin and the EBNA1 protein and mRNA, by favouring a NCL-friendly RNA G4 structure.

These results support the idea that both the RNA sequence and structure as well as the corresponding nascent polypeptide are involved in the downregulation of *EBNA1* mRNA translation. However, it does not rule out the possibility that both the RNA structure and the polypeptide sequence trigger their own separated inhibitory pathway.

As viruses use components already present in the cells to maintain themselves, the cellular biology elements brought out here can provide insights on many other pathologies in addition to EBV-associated diseases.

Key words: mRNA translation, EBNA1, RNA G4 structures, nascent polypeptide-associated control, EBV

Résumé

La traduction des ARNm et la production de protéines sont des phénomènes étroitement contrôlés dans la cellule. La séquence de l'ARNm et sa structure, auxquelles sont associées des protéines se liant à l'ARN, ainsi que la qualité de la protéine produite à partir de cet ARNm, évaluée notamment par la voie de contrôle qualité associée aux ribosomes, sont des éléments impliqués dans ce processus majeur pour la cellule.

Dans ce domaine, le contrôle de la production de protéines EBNA1 (Epstein-Barr Nuclear Antigen 1) du virus d'Epstein-Barr est un exemple intéressant. La protéine EBNA1 est essentielle pour la survie du virus dans les cellules hôtes. La protéine EBNA1 est présente dans toutes les cellules infectées en faible quantité. Cette dernière est aussi extrêmement antigénique. Il est aujourd'hui admis que la quantité de protéine EBNA1 présente dans les cellules est suffisante pour assurer le maintien du virus dans la cellule, mais assez basse pour lui permettre d'échapper au système immunitaire de l'hôte. Un contrôle de sa production est nécessaire au maintien de cet équilibre. Des études précédentes ont montré que le domaine GAR (répétitions de glycine et alanine), présent dans la partie N-terminale de la protéine, déclenche un mécanisme conduisant à l'inhibition de l'initiation de la traduction de l'ARNm d'EBNA1 en *cis*, sans affecter la traduction des autres ARNm présents dans la cellule.

Il a été montré précédemment que les structures G4s (G-quadruplex) peuvent être formées dans l'ARNm codant le GAR. De nombreuses études ont montré l'importance de ces structures secondaires de l'ARN dans la régulation de la traduction de l'ARNm d'EBNA1. La nucléoline, un facteur nucléaire, peut se lier aux G4s de l'ARNm du GAR. Cependant, il a aussi été montré que le peptide GAR, et non l'ARNm associé, est nécessaire au contrôle de la traduction de l'ARNm du GAR en *cis*. L'objectif principal de ma thèse est de mieux comprendre le mécanisme déclenché par l'ARNm et le polypeptide naissant conduisant au contrôle de la traduction de l'ARNm d'EBNA1 en *cis*.

En accord avec le fait que les structures G4s de l'ARN sont extrêmement dynamiques, nous avons montré dans un premier temps que les fonctions associées au G4s de l'ARNm du GAR, à savoir la localisation de l'ARNm, sa traduction et sa capacité à se lier à certaines protéines, dépendent du contexte dans lesquelles ces structures se trouvent. Nous montrons ensuite que la traduction de l'ARNm d'EBNA1 est nécessaire à l'interaction nucléoline-ARNm, signifiant que la traduction de l'ARNm induit des changements dans les propriétés de l'ARNm. En parallèle, nous avons étudié le NACA, une sous-unité du complexe chaperon NAC (nascent polypeptide-associated complex). NACA

se détache du ribosome lors de la synthèse du GAR et interagit avec le GAR. Le mécanisme contrôlé par NACA par rapport au domaine GAR est différent de celui réalisé en présence de polyQ, un autre polypeptide intrinsèquement désordonné. Enfin, nous montrons que eIF4A1, un des composants du complexe d'initiation eIF4F, est impliqué dans ce mécanisme d'inhibition de la traduction. Nous proposons que dans ce contexte, eIF4A1 est un modulateur des structures ARN et contrôle la formation de complexe ribonucléoprotéique, composé dans notre cas par la nucléoline et l'ARN et la protéine EBNA1, en favorisant les structures ARN G4 ayant la capacité de se lier à la nucléoline.

Ces résultats indiquent que la séquence et la structure de l'ARNm et le polypeptide naissant correspondant sont impliqués dans l'inhibition de l'initiation de la traduction de l'ARNm d'EBNA1. Cependant, cela n'écarte pas la possibilité que l'ARNm et le polypeptide naissant déclenchent chacun une voie d'inhibition de la synthèse d'EBNA1 distincte l'une de l'autre.

Les virus utilisent des éléments déjà présents dans la cellule pour assurer leur maintien dans la cellule hôte. Ainsi, les principes de biologie cellulaire décrits ici peuvent apporter des indications importantes pour une meilleure compréhension d'autres pathologies en plus de celles liées au virus d'Epstein-Barr.

Mots-clés : traduction des ARN messagers, EBNA1, G-quadruplex (ARN), contrôle des polypeptides naissants, Virus d'Epstein-Barr

Résumé détaillé en français

La traduction des ARNm (ARN messagers) est un évènement central pour les cellules, ce processus permettant la synthèse de protéines. Le modèle communément accepté pour décrire la traduction des ARNm est le suivant chez les organismes eucaryotes : lors d'une première étape dite d'initiation, la sous-unité 40S se lie à différents facteurs d'initiation de la traduction et forme le PIC (*pre-initiation complex* ou complexe de pré-initiation). L'ARNm de son côté est activé par la formation d'un complexe avec d'autres facteurs d'initiation, entraînant le recrutement du PIC. La sous-unité 40S du ribosome scanne alors la partie 5'UTR (5' *Untranslated Region*, la partie non-codante de l'ARNm en 5') jusqu'à ce qu'un codon START, le codon AUG en général, soit repéré. La reconnaissance de ce codon START par la sous-unité 40S du ribosome signifie la fin de la phase d'initiation et le début de la phase d'élongation. Les facteurs d'initiation de la traduction sont libérés de la sous-unité 40S du ribosome, qui est rejoint par la sous-unité 60S et différents facteurs d'élongation. Le ribosome 80S alors formé, aidé par les facteurs d'élongation, synthétise la chaîne polypeptidique, par polymérisation séquentielle des acides aminés correspondants aux codons lus par le ribosome lors de sa progression sur l'ARNm. Lorsque le ribosome arrive à un codon STOP (codon UAG, UAA ou UGA), l'élongation est terminée et la phase de terminaison de la traduction est lancée, au cours de laquelle le ribosome se désolidarise de l'ARNm, et les deux sous-unités 40S et 60S sont séparées et recyclées pour d'autres cycles de traduction. La chaîne polypeptidique formée est relâchée du ribosome, et différentes modifications sont souvent nécessaires par la suite avant d'obtenir *in fine* une protéine fonctionnelle.

Plusieurs mécanismes contrôlent la traduction des ARNm, de manière globale dans la cellule ou de façon spécifique à chaque ARNm. En effet, après des stress ponctuels tel qu'un choc thermique, la traduction des ARNm est globalement réduite, à l'exception de quelques ARNm régulant la réponse cellulaire aux chocs, comme ceux codant HSP70 ou p53. Au contraire, certains stimuli entraînent l'activation globale de la traduction des ARNm, notamment dans des contextes cancéreux où les cellules prolifèrent de façon incontrôlée. Dans ces cas, les voies de prolifération cellulaire liées aux kinases MAPK ou PI3K entraînent la libération et l'activation de facteurs de traduction, dont les facteurs d'initiation eIF4A et eIF4E entre autres. Cependant, l'activation de la traduction des ARNm dans les cellules cancéreuses semble viser certains ARNm en particulier, ceux nécessaires à la croissance et la prolifération cellulaire. Ainsi, la traduction des ARNm, qui obéit au schéma global initiation – élongation – terminaison décrit précédemment, est en fait un évènement contrôlé de manière spécifique pour chaque type d'ARNm. De plus, de nombreux complexes

ribonucléoprotéiques, différents les uns des autres en termes de composition, sont capables d'initier la traduction des ARNm. En outre, des protéines ou ARN ribosomaux spécifiques sont nécessaires à certains types de traduction mais facultatifs pour d'autres, mettant en évidence l'hétérogénéité des ribosomes eux-mêmes. Cette multiplicité des moyens explique en partie le caractère dynamique de la traduction des ARNm, qui permet une adaptation rapide du protéome de la cellule en réponse à différents stimuli. Le contrôle de la traduction dans les neurones est un exemple illustrant l'importance de la spécificité de la traduction des ARNm. Dans ces cellules extrêmement différenciées, les différentes parties de la cellule ont des rôles et des caractéristiques spécifiques, se reflétant dans leurs compositions cellulaires variées en termes d'ARNm et de protéines selon la partie du neurone étudiée. La traduction des ARNm dans les neurones est un phénomène contrôlé dans le temps et l'espace. En effet, certaines protéines sont synthétisées exclusivement dans les extrémités des neurones, en réponse à certains stimuli, pendant une période donnée. Ainsi, la traduction des ARNm regroupe de multiples mécanismes spécifiques à chaque type d'ARNm, avec des complexes traductionnels variés. Ces procédés dynamiques permettent à la cellule de s'adapter à un environnement changeant.

L'initiation de la traduction est l'étape identifiée comme étant la plus régulée. En effet, la disponibilité des différents facteurs d'initiation de la traduction sous leur forme active est la cible de nombreuses voies de régulation cellulaire. De plus, les caractéristiques de l'ARNm lui-même impactent l'efficacité traductionnelle. Les nucléotides composant l'ARNm peuvent être modifiés, notamment par l'ajout de groupements méthyles. Une autre caractéristique importante de l'ARNm pour la régulation de ses fonctions dans la cellule est sa structure. Les molécules d'ARN peuvent adopter des structures secondaires et tertiaires diverses, contrôlées par des liaisons de type Watson-Crick, mais aussi par des interactions non classiques, comme des liaisons hydrogènes de type Hoogsteen. Enfin, différentes protéines, appelées RBP (*RNA-binding proteins*, ou protéines se liant à l'ARN), peuvent interagir avec les ARNm. Les effets des liaisons RBP-RNA sont variables selon les interactions étudiées. Ces trois paramètres, à savoir structure de l'ARNm, modification chimique de certains nucléotides et RBP interagissant avec l'ARNm, déterminent le futur de l'ARNm dans la cellule, et évoluent dans le temps.

Une des fonctions de l'ARNm dépendante de sa structure et de ses interactions avec certains RBPs est la synthèse de protéines. Chez les bactéries, certaines structures dynamiques nommées *riboswitch* sont sensibles aux conditions cellulaires. La conformation de ces *riboswitch* varie selon la présence de certains métabolites, impactant la traduction de l'ARNm en question. Dans certains

ARNm viraux et cellulaires, des IRES (*Internal Ribosome Entry Site*) sont présents dans la partie 5'UTR et permettent le recrutement de complexes traductionnels directement au niveau de ces structures.

Les G-quadruplex (G4s) sont des structures secondaires se formant dans les acides nucléiques, ADN ou ARN, riches en guanine. Quatre bases de guanine peuvent s'associer par des liaisons de Hoogsteen dans un même plan et former un G-quartet ; l'empilement de deux G-quartets au minimum forme un G-quadruplex. L'ensemble est stabilisé par la présence de cations monovalents K^+ ou Na^+ . Ces structures jouent un rôle essentiel pour les ARNm. En effet, les G4s sont impliqués dans la maturation, le contrôle de la localisation, la stabilité et la traduction d'un grand nombre d'ARNm, et dans la formation de complexes ribonucléoprotéiques. Les G4s ARN sont aussi des structures dynamiques dans les cellules. Certains RBPs se lient préférentiellement aux séquences ayant la capacité de former des G4s, comme la protéine nucléaire nucléoline. Parmi ces RBPs, plusieurs groupes peuvent être définis en fonction de leur action sur les G4s : certains stabilisent les structures G4, d'autres empêchent leur formation en se liant à ces régions lorsque l'ARNm est en conformation linéaire, et certaines hélicases sont capables d'ouvrir ces structures. L'hélicase eIF4A, précédemment mentionné pour son rôle dans l'initiation de la traduction, est capable d'ouvrir les structures G4 ARN.

En comparaison avec l'étape d'initiation, l'avancée du ribosome sur l'ARNm lors de l'étape d'élongation est irrégulière. En effet, lors de l'élongation, le ribosome se déplace par triplet de bases et non base par base comme cela semble être le cas pendant l'initiation. Certaines modifications post-traductionnelles sont effectuées sur les polypeptides en cours de formation, telles que le repliement de la chaîne polypeptidique, l'ajout de certains sucres ou autres groupements chimiques, mais aussi le ciblage et le transport du polypeptide naissant dans des compartiments cellulaires spécifiques, tels que le réticulum endoplasmique. Ces phénomènes nécessaires à la formation d'une protéine fonctionnelle à partir de la chaîne polypeptidique requièrent parfois une pause lors de l'élongation. Pourtant, différentes voies de contrôle associées aux ribosomes, la NGD et la NSD (*No-Go* et *No-Stop Decay*, respectivement), contrôlent l'avancée du ribosome sur l'ARNm et déclenchent la dégradation du polypeptide naissant et de l'ARNm associé en cas de pause et collision de ribosomes lors de la traduction. Il est suggéré que les cellules disposent d'un moyen leur permettant de reconnaître les traductions problématiques. Cette hypothèse est confirmée par le fait que toutes les collisions de ribosomes ne déclenchent pas la dégradation de l'ARNm, mais les mécanismes derrière ce triage restent à être éclaircis.

En plus du contrôle de l'avancée du ribosome sur l'ARNm lors de l'élongation, des facteurs associés à certains ribosomes reconnaissent la nature du polypeptide en cours de formation de manière spécifique. Cela est le cas avec la synthèse de tubuline dans les cellules. Il a été montré récemment que la protéine tétrapeptide 5 (TTC5) se lie aux ribosomes synthétisant la tubuline quand cette dernière est présente en excès dans la cellule. Le mécanisme détaillé est toujours inconnu, mais TTC5 reconnaît les polypeptides naissants de tubuline juste à leur sortie du ribosome, entraînant la dégradation des ARNm codant la tubuline lorsque cette dernière est présente en quantité suffisante dans la cellule.

La sortie du tunnel emprunté par les polypeptides en cours de formation est une zone du ribosome où de nombreuses protéines se liant aux polypeptides naissants sont présentes. Le NAC (*nascent polypeptide-associated complex*) est un complexe composé de deux sous-unités NACA et BTF3, aussi respectivement appelées NAC α et NAC β . Ce complexe joue un rôle déterminant pour le ciblage cellulaire des protéines pour le réticulum endoplasmique, en augmentant la spécificité de ce ciblage. De plus, ce complexe agit en tant que chaperon et limite la formation et l'accumulation d'agrégats protéiques. Le NAC est un complexe flexible, pouvant interagir de manière transitoire avec le ribosome à plusieurs points de contact et s'adaptant à différents substrats protéiques. Du fait de sa position, ses propriétés structurelles et ses fonctions associées, le NAC a été suggéré comme étant le premier chaperon entrant en contact avec le polypeptide naissant à la sortie du ribosome, contrôlant les futures interactions de la chaîne polypeptidique avec les autres facteurs se liant aux polypeptides naissants.

Notre modèle d'étude pour ce projet est le contrôle de la traduction de l'ARNm codant la protéine EBNA1 (*Epstein-Barr Nuclear Antigen 1*) du virus d'Epstein-Barr (EBV, *Epstein-Barr virus*). La protéine EBNA1 est en grande partie intrinsèquement désordonnée, ce qui signifie qu'elle n'a pas de structure définie. Elle est nécessaire au maintien d'EBV dans les cellules infectées, en contrôlant la réplication du génome viral et en assurant la répartition des différents épisomes viraux dans les cellules filles lors de la division cellulaire. Cependant cette protéine est aussi hautement antigénique, et les personnes ayant été en contact avec EBV ont, sauf cas particuliers, un système immunitaire capable de repérer et d'éliminer les cellules infectées produisant EBNA1. EBV utilise donc un mécanisme permettant une régulation fine de la production d'EBNA1, assurant sa survie dans les cellules en produisant assez de protéines EBNA1 quand cela est nécessaire, tout en limitant la traduction de l'ARNm en question, échappant ainsi au système immunitaire de l'hôte.

Le domaine GAR (*Glycine-Alanine repeats*), une région d'EBNA1 constituée d'une répétition de glycine et alanine et située dans la moitié N-terminale de la protéine, est particulièrement désordonné. Ce domaine est associé à l'inhibition en *cis* de la traduction de l'ARNm codant EBNA1. Cette inhibition vise l'étape d'initiation de la traduction, et conduit à une limitation de la production de protéines ainsi que de peptides antigéniques dérivant d'EBNA1. Les études précédentes sur le sujet ont montré que l'ARNm codant la protéine EBNA1 forme des structures G4, notamment au niveau de la partie codant le domaine GAR. La protéine nucléoline se lie avec les séquences formant ces structures, et empêcher cette interaction conduit à lever l'inhibition de la traduction de l'ARNm codant EBNA1. D'autres études ont montré que le polypeptide GAR joue aussi un rôle dans le contrôle de la traduction de l'ARNm, car modifier légèrement la séquence en acides aminés, sans changer la capacité de l'ARNm à former des structures G4s, annule aussi l'inhibition de la traduction de l'ARNm codant EBNA1.

Le but de mon projet de thèse est de mieux comprendre comment le polypeptide naissant et l'ARNm contrôle la traduction en *cis* de l'ARNm codant EBNA1.

Nous nous sommes d'abord concentrés sur le fait que, tout comme les structures G4 ARN, les fonctions associées à ces structures sont dynamiques et sensibles au contexte. Ensuite, nous montrons que l'inhibition de la traduction de l'ARNm codant EBNA1 est dépendante de NACA et cette dernière est détachée du ribosome lors de la synthèse du GAR. La capacité de l'ARNm d'EBNA1 à se lier à la nucléoline est conditionnée par la traduction de l'ARNm et la présence de NACA. Enfin, le facteur d'initiation eIF4A est impliqué dans le contrôle de la traduction de cet ARNm, et la capacité de l'ARNm d'EBNA1 à se lier à la nucléoline dépend de la présence de ce facteur.

A partir de ces résultats et des données des études précédentes, nous proposons le modèle suivant pour le contrôle de la traduction de l'ARNm codant la protéine EBNA1. Lors de la synthèse de la protéine EBNA1, le NAC est détaché du ribosome et est lié au domaine GAR et les structures G4s dans l'ARNm sont défaites. Il a été montré précédemment que la protéine EBNA1 se lie à son propre ARNm, via les domaines RGG de la protéine et les séquences pouvant former des G4s dans l'ARNm. Etant donné que les structures ARN G4 sont favorisées d'un point de vue thermodynamique dans des contextes cellulaires, nous proposons que les G4s sont ensuite reformées dans l'ARNm d'EBNA1. Différents types de G4s sont ainsi mis en place, certains pouvant interagir avec la nucléoline et d'autres non. Il est possible qu'eIF4A1 agisse en tant que modulateur des structures G4 et transforment les G4s en G4s pouvant se lier à la nucléoline. Ces G4s pouvant interagir avec la nucléoline semblent aussi être stabilisées par le NAC. Ainsi, la nucléoline peut venir se lier à

l'ensemble. Le complexe ainsi formé, composé de la nucléoline, de la protéine EBNA1 et de l'ARNm correspondant, potentiellement stabilisé par le NAC, empêche la traduction de l'ARNm et donc la production de protéines et de peptides antigéniques dérivant d'EBNA1. Sachant que le complexe EBNA1 - nucléoline - G4 ARN est aussi nécessaire au maintien du génome viral dans les cellules hôtes, ce modèle présente un mécanisme intéressant pour le virus, car un même complexe ribonucléoprotéique assure à la fois le maintien de l'infection et l'échappement au système immunitaire de l'hôte tout en autorégulant sa propre production.

Le modèle proposé suggère que l'ARNm et le polypeptide naissant contrôlent la production d'EBNA1. Une telle coopération est un argument en faveur d'une coévolution de la séquence ARN et du polypeptide. Ce modèle souligne aussi le fait que les interactions ARN – protéine et protéine – protéine sont dynamiques, et que les fonctions des ARNm et des protéines varient selon l'environnement cellulaire.

Les structures G4 ARN sont dynamiques et multifonctionnelles. Une des propriétés remarquables de ces structures, ou du moins des ARN pouvant former ces structures, est leur capacité à se lier à de nombreux partenaires, régulant ainsi la formation et la composition de complexes ribonucléoprotéiques. Il a aussi été proposé que les G4s ARN servent de chaperons pour les régions intrinsèquement désordonnées des protéines, limitant ainsi la formation d'agrégats. Ce mécanisme cellulaire peut être utilisé par EBNA1, qui est une protéine largement désordonnée.

La production de protéines contenant des domaines intrinsèquement désordonnés est une nécessité pour les cellules, car ces domaines jouent un rôle important dans les interactions entre différents facteurs cellulaires. Cependant, ces régions sont aussi thermodynamiquement favorables à la formation d'agrégats, dont certains sont associés au développement de pathologies neurodégénératives. Ainsi, le contrôle de la production d'EBNA1 peut tirer parti d'un mécanisme régulant la formation de protéines intrinsèquement désordonnées, car ces protéines, tout comme EBNA1, sont indispensables mais doivent être régulées quantitativement.

L'implication d'eIF4A dans la régulation de la traduction de l'ARNm codant la protéine EBNA1 souligne le fait que certains facteurs d'initiation ne sont pas seulement impliqués dans l'initiation de la traduction. Dans notre modèle d'étude, eIF4A1 ne serait pas seulement une hélicase à ARN, mais aussi une protéine capable de se lier de manière spécifique et relativement stable aux G4s. Avec la mise en évidence de la diversité des complexes traductionnels, plusieurs hypothèses quant à l'arrivée d'eIF4A1 au niveau de la séquence codante du GAR sont possibles : eIF4A1 pourrait être recruté indépendamment de l'initiation de la traduction, ou ne pas être désolidarisé du

ribosome à la fin de l'initiation et rester ainsi présent dans le complexe traductionnel au cours de l'élongation. Un autre point intéressant est le rôle central de la nucléoline dans les fonctions d'EBNA1. Sachant que la nucléoline et la protéine homologue à EBNA1BP2 (*EBNA1-binding protein 2*) chez *S. cerevisiae* sont impliquées dans la ribogenèse, nous pouvons nous questionner sur l'impact de la protéine EBNA1 sur la synthèse de certains ribosomes, d'autant plus que certains complexes traductionnels, comme celui recruté par l'IRES de *c-Myc*, ne sont pas sensibles à ce mécanisme d'inhibition de la traduction de l'ARNm. Hypothétiquement, EBNA1 pourrait favoriser la synthèse de complexes traductionnels sensibles à son mécanisme d'inhibition de la traduction.

L'étude de la protéine virale EBNA1 nous apporte des connaissances nouvelles en biologie cellulaire. Ainsi, en plus de nous permettre de mieux comprendre et de traiter les pathologies liées au virus d'Epstein-Barr, les informations issues de ces études peuvent aussi nous aider à lutter contre d'autres types de pathologies, des maladies neurodégénératives à certains cancers.

Table of contents

Abstract.....	4
Résumé	6
Résumé détaillé en français.....	8
Table of contents	15
Abbreviations.....	18
INTRODUCTION.....	24
I – mRNA translation in cells	25
A – The mRNA translation process	25
1 – Cap-dependent translation	26
a – Overview of the different steps in the cap-dependent translation model.....	26
b – Co-translational events and processing of the nascent-polypeptide	29
c – Regulation of translation by trans-acting factors.....	31
2 – Heterogeneity of the translational machinery.....	32
B – mRNA translation: a dynamic and plastic process.....	34
1 – Local translation	34
2 – Selective translation in response to external stimuli.....	37
C – Viruses hijack the host mRNA translation machinery.....	39
II – mRNA structure and its influence on its processing	41
A – RNA secondary structures and functions.....	41
1 – RNA modifications impact on functions.....	41
2 – RNA G-quadruplexes	43
a – RNA G4 structures and their predicted distribution in the transcriptome	43
b – RNA G4-associated function in mRNA processing	45
i – mRNA localisation	45
ii – RNP granules formation and composition	45
iii – mRNA maturation (splicing and poly-adenylation)	46
iv – mRNA stability	47
v – mRNA translation	47
RNA G4s in the 5'UTR and 3'UTR and translation initiation	47
RNA G4s in the mRNA coding sequence (CDS) affect elongation and the future protein processing	48
3 – RNA G4s in viruses.....	50
B – RNA-binding proteins.....	51

1 – RNA-binding proteins (RBPs) and RNA binding motif (RBM) recognition.....	51
2 – Nucleolin: a multifunctional RNA-binding protein.....	52
3 – RBP – RNA interactions are dynamic	54
C – Dynamics of RNA structure in the cells.....	56
1 – Are RNA G4s folded in living cells?.....	56
2 – DEAD-box RNA helicase as RNA structure modulator	56
3 – Dynamic RNA structures and functions	59
III – Translation-dependent mRNA surveillance and protein quality control in the cell via ribosome-associated quality control	62
A – Translation-dependent mRNA surveillance by assessing the ribosome progress on the mRNA	62
1 – Nonsense-mediated decay (NMD).....	62
2 – No-go decay (NGD).....	63
3 – Non-stop mRNA decay (NSD).....	64
B – Translation-dependent mRNA surveillance pathways: beyond quality control of the nascent polypeptide	64
1 – Antigenic peptides production.....	64
2 – Cellular stress integration	65
C – Nascent polypeptide assessment during translation.....	66
1 – The nascent polypeptide nature controlling its encoding mRNA processing.....	66
2 – Avoiding aggregates formation.....	67
3 – The nascent-associated complex NAC	70
a – NAC prevents nascent polypeptides mistargeting to specific cellular compartments ...	71
b – NAC as a ribosome-associated chaperone.....	72
c – NAC as a cytosolic chaperone.....	73
IV – The Epstein-Barr Virus	75
A – The first human oncovirus identified.....	75
B – EBV latency programs	76
C – EBNA1.....	78
1 – EBNA1: the genome maintenance protein of EBV.....	78
a – EBNA1 roles in viral genome maintenance	78
b – EBNA1 controls other viral and cellular gene expressions	80
2 – EBNA1 controls its protein level and escape the host immune system	80
a – GAR-associated proteasomal degradation inhibition	81
b – The GAR domain of EBNA1 and translation inhibition	81

3 – EBNA1 is oncogenic.....	83
V – Aim of my Ph.D. project.....	84
RESULTS	85
Paper 1: Dynamic regulation of RNA G4s-associated functions	86
Paper 2: The nascent polypeptide controls mRNA translation via the NAC and changes in its encoding mRNA NCL-binding capacity.....	109
Paper 3: eIF4A1 impacts mRNA translation via remodelling RNP complex composition.....	144
DISCUSSION.....	171
I – Proposed model for EBNA1 mRNA translation regulation	172
II – Controlling intrinsically disordered protein synthesis with the NAC.....	178
III – Multifunctional mRNAs regulated by their G4 structures	180
IV – eIF4A1 as an RNA structure modulator, impacting RNP complexes composition	181
V – Diversity in the translation machinery.....	182
VI – Concluding remarks	183
References	184
Acknowledgements.....	204

Abbreviations

4E-BP	eIF4E - binding protein 1
4EHP	eIF4E - homolog protein (also known as eIF4E2)
Å	ångström
ABCE1	ATP-binding cassette sub-family E member 1
ABCF1	ATP-binding cassette sub-family F member 1
AD	Alzheimer's disease
ADAM10	Disintegrin and metalloproteinase domain-containing protein 10
Akt	RAC-alpha serine/threonine-protein kinase
ALKBH5	RNA demethylase ALKBH5, also called Alpha-ketoglutarate-dependent dioxygenase alkB homolog 5
ALS	Amyotrophic Lateral Sclerosis
APP	Amyloid precursor protein
ARE	AU-rich elements
ASCC	ASC-1 Complex
ATP	Adenosine triphosphate
A β	Amyloid β peptides
BACE1	β -Site amyloid precursor protein (APP) cleaving enzyme 1
BDNF	Brain-derived neurotrophic factor
BHRF1	BamHI fragment H rightward open reading frame 1
BL	Burkitt's Lymphoma
BTF3	Basic transcription factor 3
CaMKII α	Calcium/calmodulin-dependent protein kinase 2 α
Cdc2	Cell division control protein 2
CDK1	Cyclin-dependent kinase 1
CDS	Coding sequence
CJD	Creutzfeldt–Jakob Disease
CK2	Casein kinase 2
CMV	Cytomegalovirus
CpG	Cytosine nucleotide followed by guanine nucleotide in the linear sequence of bases along its 5'-3' direction
cPPT	central polypurine tract

CR	Central repeats
Cue2	Coupling of ubiquitin conjugation to ER degradation protein 2
DEAD-box	Asp-Glu-Ala-Asp box
DEAH-box	Asp-Glu-Ala-His box
DENV	Dengue virus
DIS	Dimer initiation site
DNA	Deoxyribonucleic acid
DRiPs	Defective ribosomal products
E2F (1)	Transcription factor E2F (1)
EBER	EBV-encoded RNA
EBNA	Epstein-Barr Nuclear Antigen
EBNA1	Epstein-Barr Nuclear Antigen 1
EBNA-LP	Epstein-Barr Nuclear Antigen Leader Protein
EBOV	Zaire ebolavirus
EBV	Epstein-Barr virus
EDF1	Endothelial differentiation-related factor 1
eEF	eukaryotic elongation factor
eIF	eukaryotic initiation factor
eIF4A1	eukaryotic initiation factor A1
eIF4B	eukaryotic initiation factor B
eIF4E	eukaryotic initiation factor E
eIF4F	eukaryotic initiation factor F
eIF4G	eukaryotic initiation factor G
eIF4H	eukaryotic initiation factor H
EJC	Exon junction complex
EMT	Epithelial-to-mesenchymal transition
ER	Endoplasmic reticulum
eRF	Eukaryotic translation termination factor
ERK	Extracellular signal-regulated kinase
FBL	rRNA 2'-O-methyltransferase fibrillarin
FMR1	Fragile X mental retardation 1
FMRP	Fragile X mental retardation 1 protein
FTD	Frontotemporal Dementia

FTO	Fat mass and obesity-associated protein (Alpha-ketoglutarate-dependent dioxygenase FTO)
FXS	Fragile X syndrome
G4	G-quadruplex
G4C2	(GGGGCC) nucleotide motif
GAIT	Interferon (IFN)- γ -activated inhibitor of translation
GA _r	Glycine – Alanine repeats
GCN2	General control non-derepressible protein 2
GIGYF2	Grb10-interacting GYF (glycine-tyrosine-phenylalanine) domain protein 2
GMP	Genome maintenance protein
HAV	Hepatitis A virus
HCV	Hepatitis C virus
hEBP2	human EBNA1-binding protein 2
Hel2	Histone E3 ligase 2/ E3 ubiquitin-protein ligase HEL2
hER α	human estrogen receptor α
HIV-1	Human Immunodeficiency Virus 1
HnRNP	Heterogeneous nuclear ribonucleoprotein
HPV	Human Papillomavirus
Hsf1	Heat shock factor 1
HSP	Heat Shock Protein
HSV-1 /2	herpes simplex virus 1 /2
HuR	Hu antigen R
IDR	intrinsically disordered region
IFN- γ	interferon- γ
IM	infectious mononucleosis
IRES	Internal Ribosome Entry Site
ITAF	IRES trans-acting factors
I κ B α	NF- κ B inhibitor α
JNK	c-Jun N-terminal kinase
kb	kilobases
KH protein domain	hnRNP K-homology domain
KSHV	Kaposi's sarcoma-associated herpesvirus
LANA1	Latency associated-nuclear antigen 1

LCL	Lymphoblastoid cell lines
LMP	Latent membrane protein
LTR	Long Terminal Repeat (in HIV-1)
m6A	N6-methyladenosine
m7G	methylguanosine cap
MAP	Mitogen-activated protein
MAPK	Mitogen-activated protein kinase
MDM2	E3 ubiquitin-protein ligase Mdm2
METTL	N(6)-adenine-specific methyltransferase
mGluR5	Metabotropic glutamate receptor 5
MHC	Major Histocompatibility Complex
miRISC	miRNA-induced silencing complexes
miRNA	microRNA
MNK	MAPK-interacting kinase
MREC	Met-Arg-Glu-Cys
MREI	Met-Arg-Glu-Ile
mRNA	messenger RNA
mTOR	mechanistic target of rapamycin
mTORC	mechanistic target of rapamycin complex
MZT	Maternal-to-zygotic transition
NAC	Nascent polypeptide-associated complex
NACA	Nascent polypeptide-associated complex α
NCL	Nucleolin
ncRNA	non-coding RNA
NF- κ B1	Nuclear factor of kappa light polypeptide gene enhancer in B-cells 1
NGD	No-go decay
NLS	Nuclear localisation signal
NMD	Nonsense-mediated decay
NPC	Nasopharyngeal carcinomas
NSD	Non-Stop decay
Nsp	Non-structural protein
ORC	Origin recognition complex
ORF	Open Reading Frame

oriP	EBV episome origin of replication
P54nrb	PSF binding partner
PABP	poly(A)-binding protein / Polyadenylate-binding protein
PDCD4	Programmed cell death protein 4
PD-L1	Programmed cell death 1 ligand 1
PDS	Pyridostatin
PEL	Primary effusion lymphoma
PI3K δ	Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit delta isoform
PIC	Pre-initiation complex
PKC	Protein kinase C
poly(A)-tail	Polyadenylate tail
poly(G)	Poly-guanine
polyQ	Poly-glutamine
PrP	Proteinaceous infectious particle
PRTEs	Pyrimidine- rich translational elements
PSD-95	Postsynaptic density protein 95
PSF	PTB-associated splicing factor
PTC	Premature termination codon
PTPs	Pioneer translation products
qRMM	quasi RRM
r(G4C2)	repeat of (G4C2) motif
RAC	Ribosome-associated complex
RBM	RNA-binding motif
RBP	RNA-binding protein
RecA	Recombinase A
RG/RGG-rich	Arginine-Glycine/Arginine-Glycine-Glycine rich domain
RHD	Rel homology domain
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
RON	Tyrosine kinase receptor RON
RQC	Ribosome-associated quality control
RRM	RNA recognition motif
rRNA	ribosomal RNA

S6K	Ribosomal protein S6 kinase
SAP domain	SAF-A/B, Acinus, and PIAS domain
SAPKs	Stress-activated protein kinases
sAPP- α	Soluble amyloid precursor protein α
SMNDC1	Survival motor neuron domain containing 1
SRP	Signal recognition particle
Ssb	Ribosome-associated molecular chaperone Ssb
Ssz	Ribosome-associated complex subunit SSZ
TDP-43	TAR DNA-binding protein 43
TGF β	Transforming growth factor-beta
TOP	Terminal oligopyrimidine
TP53	Tumor protein p53
tRNA	transfer RNA
TTC5	Tetratricopeptide protein 5
Upf	Up-frameshift
USP7	Ubiquitin carboxyl-terminal hydrolase
UTR	Untranslated Region
XRN1	5'-3' exoribonuclease 1
YB-1	Y-box binding protein 1
YTHDC2	3'-5' RNA helicase YTHDC2
YTHDF1	YTH domain-containing family protein 1
YTHDF2	YTH domain-containing family protein 2
ZAK α	alpha motif- and leucine zipper-containing kinase α / Mitogen-activated protein kinase kinase kinase 20 α
ZIKV	Zika Virus
ZNF598	E3 ubiquitin-protein ligase ZNF598
Zuo1	Zuotin

INTRODUCTION

I – mRNA translation in cells

A protein is a polypeptide formed by the polymerisation of amino acids, which can adopt a specific 3D structure. Proteins are essential for the cells, as they ensure structural and functional roles. Thus, their production via mRNA translation is tightly regulated and involves numerous factors.

A – The mRNA translation process

A key element of the synthesis of protein is the ribosome, which is a large ribonucleotide particle of ~4 MDa in eukaryotes, composed of two subunits, the small (40S in eukaryotes) ribosomal subunit and the large (60S in eukaryotes) ribosomal subunit (1). Both subunits are composed of an assembly of different ribosomal proteins and RNAs (rRNAs). The polymerisation of amino acids implies the formation of peptide bonds, a process realised by the ribosome, in the peptidyl-transferase centre, located in a cleft in the large ribosomal subunit (1). On the small ribosomal subunit, selection of the correct aminoacyl-transfer RNA (tRNA) specified by the codon in the mRNA is enabled by the decoding centre, where codon-anticodon recognition occurs (2).

The second major element of mRNA translation is the mRNA itself. The primary transcript, also called pre-mRNA, undergoes a maturation process before being exported in the cytoplasm. The mature mRNA is spliced and bears a methyl-guanosine cap (m7G) at its 5' extremity, and a poly(A)-tail at its 3' end (3). In most of the case, the coding sequence (CDS) of the mRNA is flanked by untranslated regions (5'UTR and 3'UTR), which play important roles in the mRNA processing in the cell. Additionally, the mRNA molecule is not strictly single-stranded, as RNA structures can be formed through canonical base-pairing or non-canonical interactions (4), thus translation is often associated with RNA unwinding.

1 – Cap-dependent translation

a – Overview of the different steps in the cap-dependent translation model

The mRNA translation process is divided into three steps, called initiation, elongation, and termination (**Fig. A**). In all steps, the ribosome is associated with a wide variety of proteins, modulating its activity.

The first step is initiation. Initiation designs the recruitment of the mRNA to the ribosome and the scanning of the 5'UTR until the start codon AUG. The initiation is the rate-limiting and most regulated step of translation (3).

The m7G cap is recognised by the eIF4F complex, and more precisely by the eIF4E1 protein, the cap-binding protein of this multi-protein complex (3). The other elements present in the eIF4F complex are the eIF4A1 and the scaffold protein eIF4G. Other proteins can also bind to the eIF4F complex, like the eIF4A1 co-factors eIF4B or eIF4H. Indeed, eIF4A1 is a DEAD-box RNA helicase, having an RNA-dependent ATPase and an ATP-dependent RNA unwinding activity. Alone, eIF4A1 is a non-processive helicase, meaning it unwinds only the structure right next to it but does not translocate or unwind further structures on the mRNA (5). eIF4A1 association with eIF4B, or its homolog eIF4H, and eIF4G1 increases its RNA unwinding activity, by allowing it to unwind more stable and longer RNA structures. In addition, eIF4A1 in the eIF4F complex also has a directionality (5–8). The poly(A)-binding protein PABP is also considered as an important player for translation. It is thought that the binding of PABP with eIF4G1 results in a circularisation of the mRNA, activating the mRNA for translation (3).

On the ribosome side, other initiation factors are preparing the small ribosomal subunit for translation. The 43S pre-initiation complex (PIC) is made of the 40S small ribosomal subunit, associated with the initiation factors eIF3, eIF1, eIF1A, eIF5, and the ternary complex, which designs the GTP-bound eIF2 and charged initiator methionyl tRNA. The 43S PIC is joining the eIF4F complex on the mRNA, notably through the interaction of eIF4G1 with eIF3, although other interactions between the different initiation factors might also help. The complex formed by the association of the PIC with eIF4F is called the 48S initiation complex and scans the 5'UTR until it reaches a start codon AUG (2, 3, 9).

When the AUG codon is recognised by the ternary complex via codon-anticodon recognition, the GTP bound on eIF2 is hydrolysed, and a GTP-bound eIF5B is recruited. eIF5B joining triggers the

release of other initiation factors, such as the eIF2-GDP, eIF1, eIF3, and eIF5. Subsequently, the 60S ribosomal subunit joins the 40S, leading to eIF5B GTP hydrolysis and the release of eIF5B and eIF1A. The joined 40S and 60S ribosomal subunits form the functional 80S ribosome and can start the elongation process (2, 3, 9).

During elongation, the amino acid chain is polymerised. Proper codon recognition by the anticodon of a charged tRNA allows the synthesis of a polypeptide, that will form a functional protein after further post-translational processing. In eukaryotes, elongation factors like eEF1A and eEF2 are required, notably for the ribosome translocation on the mRNA (3).

When the ribosome reaches a stop codon (UAA, UAG, or UGA), the termination process begins. In normal conditions, the stop codon is closed to the 3'end, not far from PABP. PABP promotes termination by recruiting eRF1 and eRF3a to the A site of the ribosome, leading to the release of the polypeptide in formation (10). The tRNA associated with the stop codon is released and the 80S ribosome is dissociated into the 40S and 60S ribosomal subunits. Both ribosomal subunits can be recycled via the factor ABCE1 for other translations (3, 10).

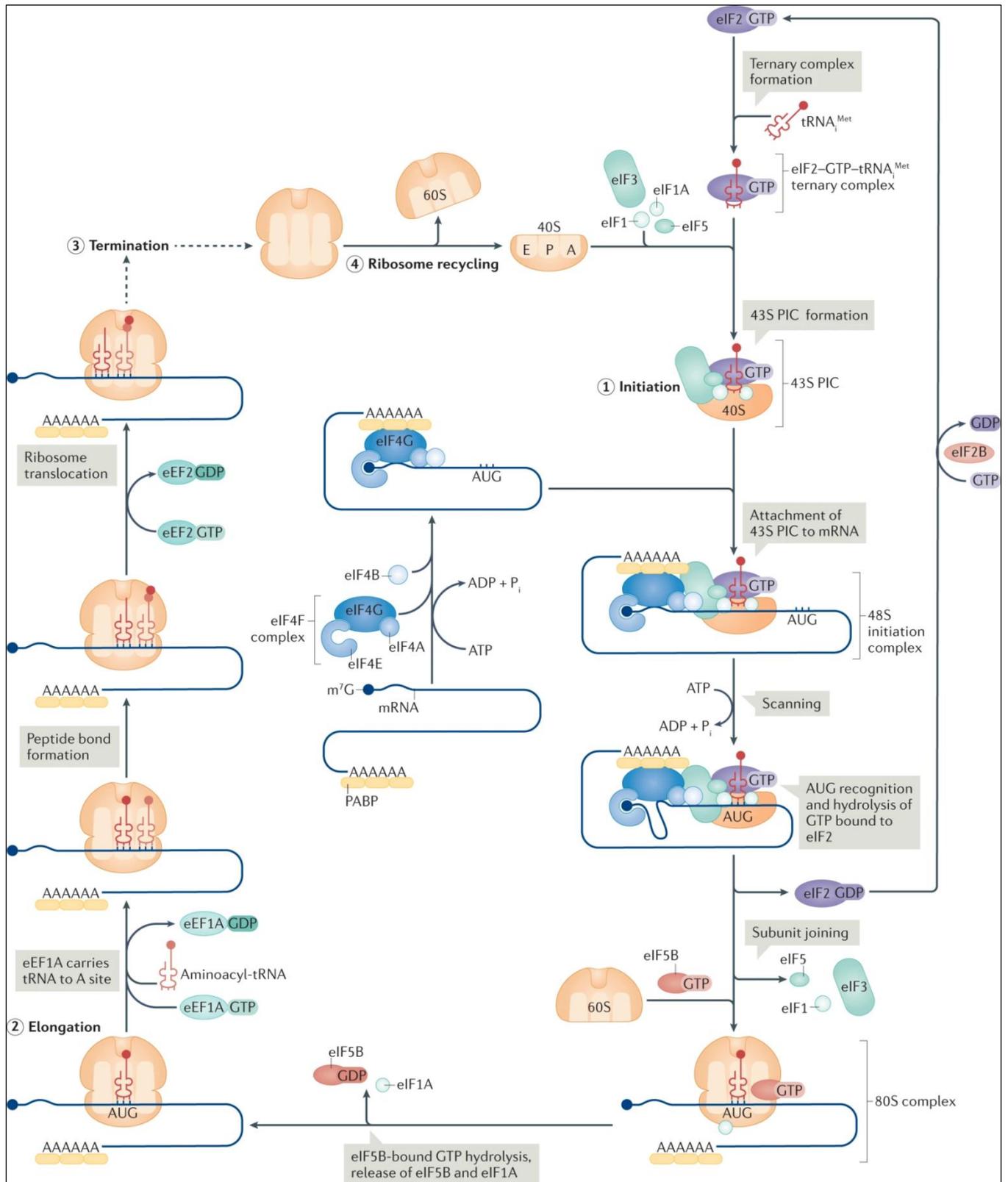


Figure A: The process of mRNA translation (from (3)). The mRNA translation model presented here is the cap and eIF4F-dependent model.

b – Co-translational events and processing of the nascent-polypeptide

Translation, and especially elongation, is not a smooth process. During elongation, the speed of the ribosome can vary, and ribosomes can even pause at some codons (11). Different elements can impact the progression of the ribosome on the ORF. Rare codons clusters, mRNA secondary structures as well as the nascent polypeptide chain electric charges can slow down the elongation process (12). Moreover, pauses can be necessary for the production of the protein, as post-translational modifications can already occur on the nascent chain, co-translationally, like enzymatic processing of the nascent polypeptides, complex formation, proper targeting of the polypeptides to the endoplasmic reticulum (ER), or protein folding (13). The latter will be more detailed in another section.

Near the polypeptide exit site of the ribosome, a factor-binding platform allows different proteins to be in close proximity with the nascent polypeptide (1). Enzymes needed for post-translational modifications of the polypeptides can bind the ribosome on this platform, and thus polypeptide modifications can already occur while the polypeptide is still polymerising. Among these modifications, the excision of the N-terminal methionine of some polypeptides by the methionine aminopeptidase (MAP) requires specific timing, suggesting that it can be impacted by the speed of elongation (14). N-terminal acetylation and glycosylation can also occur co-translationally (11).

Ribosome pausing is also crucial for the formation of protein complexes, whether homomers or heteromers. An example of this co-translationally event is the production of NF- κ B1 (nuclear factor of kappa light polypeptide gene enhancer in B-cells 1), a member of the NF- κ B family of transcription factors. NF- κ B1 is a heterodimer of p50 and p105 subunits, with p50 being a truncated form of p105. p50 contains an RHD (Rel homology domain) in its N-terminal part, an NLS (nuclear localisation signal), and a glycine-arginine rich (RG-rich) region in its C-terminal region. p105 has in addition to p50 an ankyrin repeats domain in the C-terminal part, after the glycine-arginine rich region. It has been shown that dimerisation between RHDs occurs co-translationally between a nascent p105 and a freshly formed or another nascent p105. This dimerisation in the early stage of elongation protects one of the p105 from co-translational cleavage by the proteasome into p50, allowing the formation of functional p50-p105 heterodimer NF- κ B1 (15) (**Fig. B**).

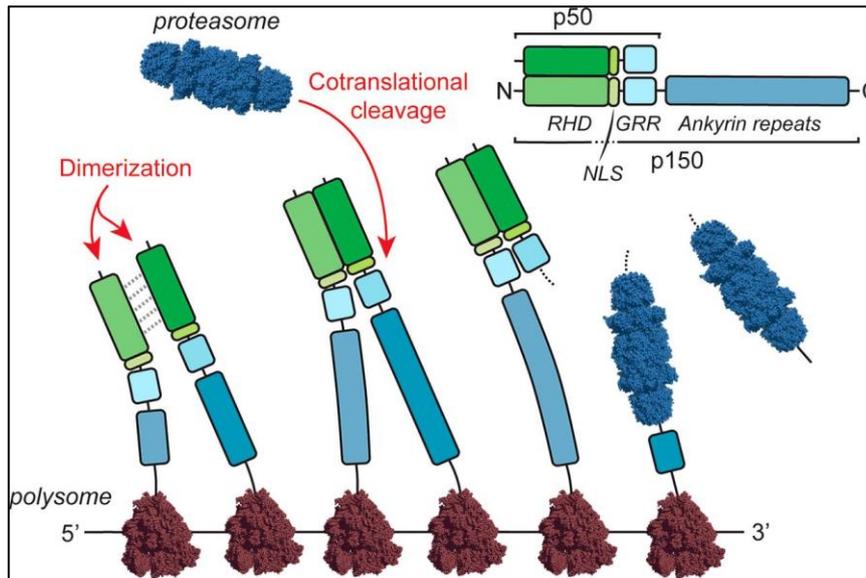


Figure B: Co-translational assembly of the p50-p105 heterodimer (from (15)): The RHD domains (in green) dimerise and the proteasome cleaves one of the p105 ankyrin repeats (in dark blue), co-translationally. It is not sure that dimerisation occurs between two nascent p105, as represented here, or between a nascent p105 and just terminated p105.

Ribosome-bound factors can also induce elongation pauses. It is the case for the ER (Endoplasmic Reticulum)-targeting of nascent chains via SRP (Signal Recognition Particle). Proteins whose maturation requires entrance in the ER exhibits a strong signal sequence for the ER in their N-terminal part. The SRP, a protein binding the ribosome near the polypeptide exit tunnel, recognises this signal selectively, inducing a halt in translation elongation. The SRP is then recognised by SRP receptors on the ER membrane, and interaction between SRP and SRP receptors resume translation by the ribosome, with the nascent polypeptide entering the ER co-translationally (11). In *Caenorhabditis elegans*, yeast, and human, the selectivity of this process depends greatly on the NAC (nascent polypeptide-associated complex), another chaperone present near the exit tunnel of the ribosome (11, 16, 17). NAC decreases nonspecific targeting for ER by binding both the SRP and the target polypeptide and reducing nonspecific SRP interaction with nascent chains not targeted for ER (16).

c – Regulation of translation by trans-acting factors

Many signalling pathways mediate translation, mainly by targeting the initiation step. Most of the time, these pathways either activate translation factors or target some mRNAs via their 5'UTR.

eIF4E1 and eIF4A1 are key elements in translation initiation. Their availability and activity are tightly regulated in the cell, and both are downstream targets of oncogenic pathways (3). Indeed, tumorous cells are proliferating cells, and this activity demands a high production of some specific proteins, linked to cell growth and proliferation.

eIF4E1 is trapped by 4E-binding proteins (4E-BP1 and 2). Upon activation of the PI3k-Akt-mTOR oncogenic pathway, 4E-BP1/2 are phosphorylated. This induces a lower affinity of 4E-BPs for eIF4E1, and thus eIF4E1 is released from 4E-BPs and can form a complex with eIF4G (3). After forming a complex with eIF4G1, eIF4E1 can be further stimulated by phosphorylation on Ser209 by MAPK-interacting kinase 1 (MNK1) and MNK2, downstream of the MAPK pathway (18).

Increasing availability of eIF4A also happens downstream of the PI3k-Akt-mTOR pathway. eIF4A1 is trapped by the programmed cell death 4 protein (PDCD4). mTORC1 activation leads to phosphorylation of ribosomal protein S6 kinase (S6K), which then phosphorylates PDCD4, releasing eIF4A1 (3). In addition to eIF4E1 and eIF4A1, other initiation factors are targeted by oncogenic pathways, increasing the rates of some mRNAs translation (3).

Specific sequence or structure in the mRNA 5'UTR can enhance translation of the mRNA via recruitment of translation complexes. It is the case for mRNA bearing a 5' terminal oligopyrimidine (TOP) sequence, or pyrimidine rich translational elements (PRTes), which mainly encode ribosomal protein and prometastatic mRNAs respectively. Upon mTORC activation, translation of these mRNAs is induced, promoting cell growth (3).

In addition to the canonical translation process, many other translational complexes begin to catch more attention lately. Indeed, the alternative translation events, like the non-AUG translations, seem to play an important role in the cellular metabolism, rather than just being an accident and requires different translational machinery than the one described above (19).

2 – Heterogeneity of the translational machinery

In vitro translation studies showed that by combining the different members of the eIF4G, eIF4A, and eIF4E families (namely eIF4G1 and 2, eIF4A1 and 2, and eIF4E1 and 3), eight different translation initiation complexes can operate translation initiation, showing that different types of eIF4F complex can be functional (20). Additionally, both eIF4A cofactors, eIF4B and H, are competing for the same binding site on eIF4A1, meaning that both cannot be at the same time in the same initiation complex (21).

Recent studies showed that the canonical translational machinery is not needed for the translation of all mRNAs. Studies using eIF4A inhibitors, the RNA helicase component of eIF4F, showed that some mRNAs depend on eIF4A for translation, namely the ones with complex 5'UTR structures upstream the start codon (22, 23, 23–26), meaning that other mRNAs can be translated without eIF4A. eIF4B, eIF4A cofactor, is also indispensable for the translation of some mRNAs with complex structures, independently from eIF4A. Interestingly, eIF4G seems to be needed for the translation of mRNAs with short unstructured 5'UTR and with high closed-loop conformation potentiality (27). Taken together, these studies show that the canonical translational machinery, at least for translation initiation, is actually diverse.

Concerning the cap-binding protein, eIF4E proteins do not have a monopoly. Initiation of translation can also start with eIF3d as the cap-binding protein. eIF3d cap-dependent translation is upregulated during stress, and usually targets mRNAs with a stem-loop in their 5'UTR (3). Among these mRNAs, a lot of them encode cell proliferation regulators, like *c-Jun*. Interestingly, *c-Jun* mRNA also bears an inhibitory RNA element preventing eIF4E recruitment, which further promotes eIF3d interaction with the cap (28).

Another example of cap-dependent but eIF4E-independent translation is the pioneer round of translation which occurs before or during the nuclear export of the mRNA. This translational event has been described as an mRNA quality control, enabling the cell to check if the mRNA is mature for further translation and production of protein. Additional cellular roles for this translational event will be discussed later. mRNAs bound by the nuclear protein CBP80 on their cap undergo a pioneer round of translation, before CBP80-CBP20 replacement by eIF4E (29). This pioneer round of translation requires eIF4G1, PABP, eIF3, eIF4A1, and eIF2 α , meaning that some translational factors can be part of different translational complexes (30, 31). Localisation of the mRNA during its translation is also an important feature, as the repartition of the different translation factors can

impact the efficiency and the way the targeted mRNA is translated. A recent study shows that mRNAs bound to the cap-binding complex (CBC, the heterodimer formed by the association of CPB80 and CBP20) requires a specific translation initiation factor called CTIF (CBC-dependent translation initiation factor), which is tethered to the DEAD-box helicase 19B in the perinuclear region, and this seems to trigger translation of the mRNA while it is exporting from the nucleus (32).

First associated with some viral mRNAs' translation, cap-independent translation initiation also occurs for some cellular mRNAs. In this case, an RNA structure, called IRES (Internal Ribosome Entry Site), located in the 5'UTR of the mRNA, allows initiation of translation directly at this structure, without the need for cap recognition by the translation initiation complex. The recruitment of the small ribosomal subunit is mediated in different ways, depending on the IRES (2). Some IRESs structures, like *HCV* IRES, recruit directly initiation factors without the help of other proteins, although their activity can be enhanced by *trans* factors (33–35). On the other hand, some IRESs require IRES *trans*-acting factors (ITAFs) to be able to recruit the 40S and initiation factors. The *Myc* family of IRESs, to which the *c-Myc* IRES belongs, requires specific ITAFs, namely the PSF (PTB-associated splicing factor), p54nrb (PSF binding partner), GRSF-1 (G-rich RNA sequence binding factor 1), and YB-1 (Y-box binding protein 1), which seems to activate specifically *Myc* IRESs (36). Exact requirements and mechanisms for ITAF-dependent IRES-mediated translation seem to be specific to each IRESs and are for now poorly understood (33).

Another mechanism of cap-independent translation in cells is the translation initiation via N6-methyladenosine (m6A) modifications in the 5'UTR of the mRNA, like in the *HSP70* mRNA (37). The exact mechanism is still unclear, but it was shown that eIF3 can directly bind m6A, promoting 40S recruitment on the mRNA (37). Other reports suggest the involvement of m6A reader for activation of m6A-dependent translation, as ABCF1 or YTHDF1 and 3 (38, 39).

Finally, the ribosomes themselves are heterogeneous. Indeed, the composition of the ribosomes might vary in terms of ribosomal proteins and rRNA present in a specific ribosome. Additionally, post-translational modifications on the ribosomal proteins as well as chemical modifications on the rRNA increase ribosomal heterogeneity (40). Specialised ribosome synthesis can be controlled by regulation of ribosomal proteins and rRNA production, or during the assembly. A major part of ribogenesis occurs in the nucleolus (41). Associated with this diversity in composition, different specialised ribosomes have different functions by targeting a specific subset of mRNAs (42). For example, ribosomes with RPL10A are necessary for efficient translation of a sub-pool of mRNAs, part of them bearing IRESs (42). Furthermore, different studies showed that some

specific ribosomal proteins are needed in precise parts of the organism at key points of the development (43). Ribosomal composition is diverse from both intracellular and intercellular points of view but is also responsive to the cell status. Cells losing p53 expression produce more rRNA 2'-O-methyltransferase fibrillarin (FBL), an enzyme catalysing the 2'-O-methylation on rRNA. This chemical modification on rRNA is associated with increase translation of oncogenes like IGF1R, MYC, and VEGFA (3).

Taken together, all these different examples support the idea that specific translational machinery can target a specific pool of mRNAs, depending on the cellular environment.

B – mRNA translation: a dynamic and plastic process

mRNA translation is an event happening in all cells. However, differentiated cells in a same organism have different translational profiles, depending on the cell function and stage of differentiation. This translational profile is changing over time, according to the change in the environment of the cells. mRNA translation plasticity, *i.e.* the ability of the cells to quickly change their translational profile, is particularly critical when cells are under stress, notably by enabling selective mRNA translation (3). mRNA translation plasticity is also an important phenomenon during cell differentiation and development of an organism, as well as in highly differentiated cells like neurons.

1 – Local translation

Neurons are highly differentiated cells with a particular shape. They have a cell body, also called soma, where the nucleus resides, and two main types of filaments that extrude from the soma called neurites, which can be either dendrites or axons (**Fig. C**). Contact between neurons happens at synapses, which are the junction between the two communicating neurons. At a synapse, the axonal end of the presynaptic neuron sends information to the dendrite terminal end of the postsynaptic neuron, called the spine. Dendrites are extremely branched and can be tens of millimetres long, whereas axons can be up to one meter long (44). Both spine and axon terminals present a very specific proteome, compared to the rest of the cells, and some proteins are highly expressed in the synaptic areas. CaMKII α , glutamate receptors, and PSD-95 have been identified as typical spine

proteins. Maintenance of these synaptic connections also requires the renewal of the specific dendritic or axonic proteome in a time-specific manner (45).

At the spines and the axon terminal, the very specific proteome is partly explained by the movement of proteins from the soma to the extremities of the cells. However, local translation has been visualised in these neurons. Indeed, it has been shown that some mRNAs are carried and stored in the dendrites and axons, where they are translated and lead to local production of specific factors. Illustrating this, *CaMKII α* and *PSD-95* mRNAs localise preferentially in dendrites and spines (46, 47). In the terminal part of axons, *β -catenin* mRNAs are preferentially localised in the newly formed pre-synaptic terminal and are translated there, regulating synaptic vesicle release (48). Another interesting example is the *BDNF* (brain-derived neurotrophic factor) mRNA translation. BDNF transcripts present different splice variants, which are transported in different areas in the neurons, according to the 5'UTR and 3'UTR of these transcripts, and are translated locally, each form of BDNF resulting in the translation of a specific splice variant having a special function in a specific compartment of the neuron (49).

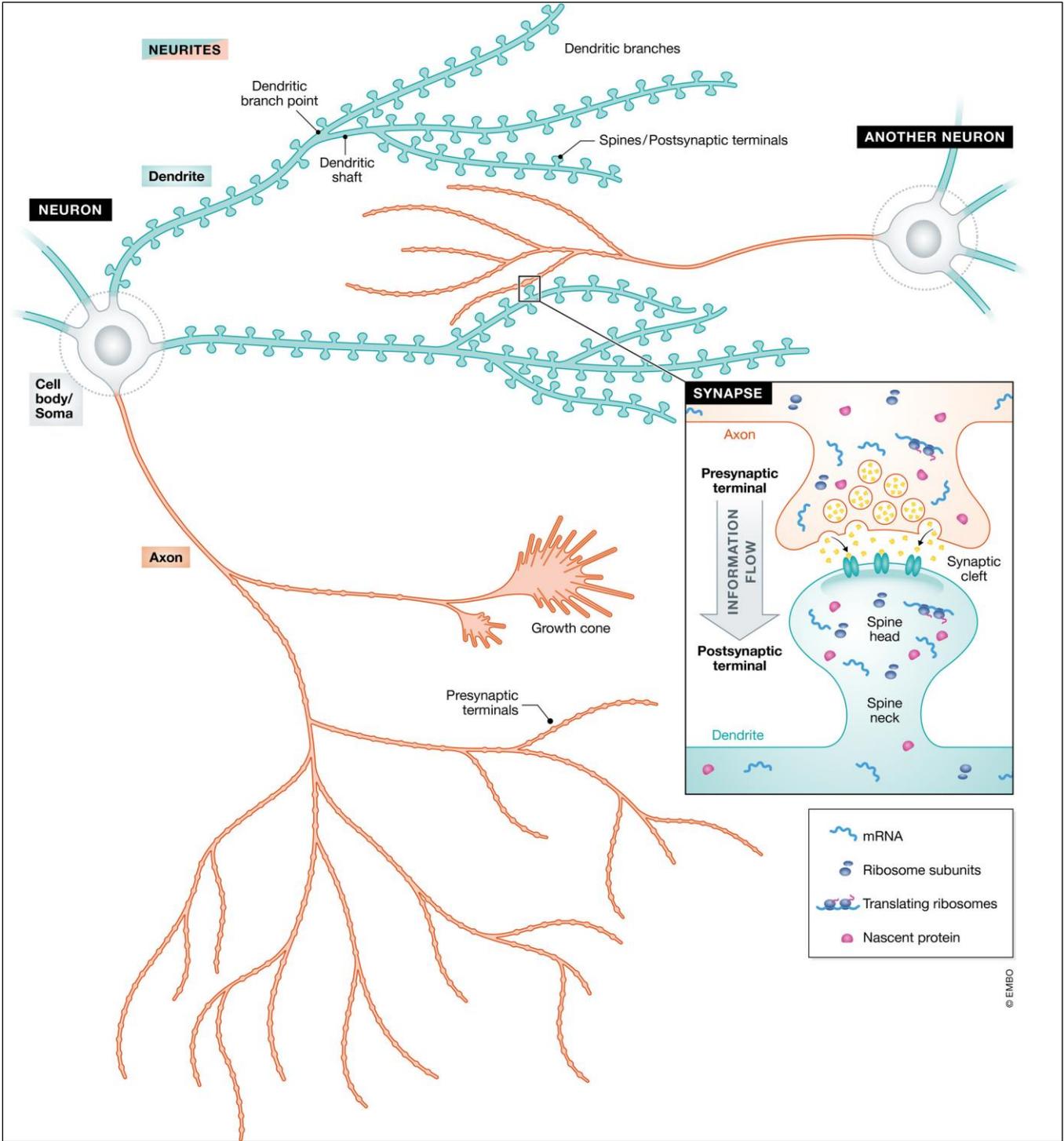


Figure C: Neuron and its structural compartments (from (45)): Neuron with its cell body (in gray) and its neurites (dendrites in blue, axon in red). The inset shows a synapse formed between the presynaptic terminal on the axon of a neuron (in red) and the postsynaptic terminal on the dendrite of another neuron (in blue).

Failure in mRNA localisation in neurons is associated with the development of many neurodegenerative diseases. One of the factors regulating mRNA transport in neurons is the TAR DNA-binding protein 43 (TDP-43), a DNA/RNA-binding protein. TDP-43 proteinopathy is associated with the majority of ALS (Amyotrophic Lateral Sclerosis) cases and nearly 50% of all FTD (Frontotemporal Dementia) cases. In TDP-43 proteinopathy, TDP-43 proteins are trapped in ubiquitinated and hyperphosphorylated aggregates. Both formation of aggregates and loss of TDP-43 functions as an mRNA transporter lead to the dysregulation of the neuron function and the development of ALS or FTD (50).

Another RNA-binding protein important for mRNA transport and postsynaptic translation is FMRP (fragile X mental retardation 1 protein). Loss of FMRP is a marker of Fragile X syndrome (FXS), the most common inherited intellectual disability, which is characterised in the neurons by spatiotemporally aberrant translation of a subset of mRNAs, leading to erroneous production of synaptic proteins, internalisation of receptors, and impaired synaptic plasticity (50). *In vitro*, it has been shown that the RGG-rich and the C-terminal domains of FMRP stop elongation of translation of targeted mRNAs, suggesting that FMRP, in addition to its mRNA transport function, also prevents translation of a subset of mRNAs before reaching their final destination in the neurons (51).

Local translation in the terminal parts of the neurons has been observed even after the separation between the soma and the dendrites, and this local translation is also reactive to stimuli perceived by the spines (52).

2 – Selective translation in response to external stimuli

Neurons are also characterised by their plasticity. Indeed, connections between neurons can evolve over time, with the growth of axons led by growth cones, which are specialised structures at their tips. Growth cones are able to detect and react to growth cues, forming new synapses with the spines of the targeted neurons when they reach their destination (45). An example of selective mRNA translation in response to growth cues is the translation of β -actin mRNA in response to Netrin-1 molecules. Studies in *Xenopus laevis* showed that Netrin-1 triggers the movement of β -actin mRNAs in the growth cone, by moving Vg1RBP granules in which the β -actin mRNAs are trapped. In addition, a gradient of Netrin-1 directs the polarisation of 4E-BPs phosphorylation, leading to a polarised activation of β -actin mRNAs translation and β -actin polymerisation, thus promoting turn and growth of the axonal cone toward Netrin-1 source (53).

Selective mRNA translation in response to stimuli is also important for the reception and integration of information in the spines and dendrites in general. Upon stimulation, the local concentration of mRNAs and ribosomes peaks at the stimulation point in the dendrite, in an area called the source compartment. This results in space and time-restricted translation of specific mRNAs. After a short period of time, during which translation activity is intense in the source compartment, translational activity decreases, and the nascent proteins produced spread gradually around (45).

As we have seen earlier, oncogenic signalling pathways can activate the eIF4F complex for mRNA translation. mTORC1 activation triggers the release of eIF4E and eIF4A from 4E-BP1/2 and PDCD4, respectively (3). However, not all mRNA translations are equally stimulated by these pathways, although a general increase of translation is observed in the early stage of cancer development. eIF4F-targeted mRNAs seem to be the ones bearing long and structured 5'UTR, usually encoding proteins involved in cell growth and proliferation (54). Upon extracellular growth factors, like transforming growth factor-beta (TGF β), eIF4A promotes *STAT1* mRNA translation and subsequent increase of PD-L1 gene expression, enabling cancer cells to bypass the immune checkpoint surveillance (55). TGF β can also activate the MAPK pathway, which triggers a cascade of phosphorylation downstream ERK, leading to activation of eIF4E by MNK1 (56) and promoting *Snail* and *Mmp-3* mRNAs translation, which are key factors in the epithelial-to-mesenchymal transition (EMT) (56). Thus, eIF4F is nowadays more considered as a potential oncogenic complex, stimulating oncogenic protein production, rather than just a general initiation factor.

As tumour cells proliferate at a fast pace, a lack of nutrients and oxygen in their environment is reached, generating genotoxic, metabolic, oxidative, or proteotoxic stress. These stresses induce in some tumour cells a modification in the mRNA translation profile. Global downregulation of mRNA translation is observed at this stage, via the limitation of Ser51 phosphorylation on eIF2 α , which inhibits the ternary complex formation via preventing GDP release and GTP loading on eIF2. Non-AUG and IRES-mediated translation are more common in these conditions, leading to preferential production of pro-survival proteins like BCL2 and MYC, and angiogenic factors like VEGF (3).

mRNA translation can be quickly adapted to the cell-specific requirement and external stimuli. The mechanisms enabling mRNA translation plasticity and specificity can be targeted by viruses, enabling efficient production of viral protein and virus survival and proliferation in the host.

C – Viruses hijack the host mRNA translation machinery

Viruses are host-specific biological elements, composed of a genome, made either of RNA or DNA, and additional proteins which allow its entrance, persistence, and proliferation in the host cells. The vast majority of viruses known currently, if not all, depends on the host translation machinery to replicate their genome and produce viral proteins, and thus their survival, proliferation and infectivity capacity depend on their ability to use the host translation machinery. Viruses hijack the translation machinery by preventing some cellular mRNAs translation and favouring theirs, in a tightly regulated manner. To do so, viruses in infected cells can change the proportion of translation factors in the host cells or modify the functions of some translation factors. Consequently, viruses can maintain their genome in the infected cells during latent infections or produce virions when the lytic cycle is triggered, in both cases promoting their infectivity. Controlling the host mRNAs translation also led to a change in the repartition of energy and translation factors, which are limited resources in the cells, in favour of viral mRNA translation (57).

Adenoviruses limit host immune response by selectively inhibiting host mRNAs translation. Upon infection, eIF4E phosphorylation by eIF4G-associated kinase Mnk1 is inhibited, leading to decrease translation of the *IκBα* mRNA. *IκBα* is a promoter of and blocks NF-κB action. Therefore, by limiting the production of *IκBα*, adenoviruses prevent IFN-γ production and mitigate the host immune response.

RNA structures play an important role in the viral modulation of the host translational cell machinery, notably by recruiting the ribosome on the viral mRNAs, like the *HCV* IRES described earlier. Other viruses also use IRES elements to promote viral mRNA translation. The hepatitis A virus (*HAV*) requires eIF4G and eIF4A binding on the viral mRNA for translation. When eIF4E is available, the whole eIF4F complex bound to the *HAV* IRES has an efficient helicase activity, resulting in efficient viral mRNA translation. In absence of eIF4E, the affinity of the eIF4G-eIF4A complex for *HAV* IRES is reduced, as well as its helicase activity, resulting in low translational efficiency (58).

Some viruses enhance their own mRNAs translations by stopping the host mRNAs translation and attracting translation factors to viral mRNAs, like the poliovirus. In poliovirus-infected cells, the poliovirus 2A protease cleaves eIF4G1 and eIF4G2 translation initiation factors, resulting in a decrease of cap-dependent mRNA translations. As poliovirus mRNAs are uncapped, the viral mRNA translation is not affected. In addition to eIF4G1 and 2 cleavage, poliovirus 2A protease is also essential for the formation and stabilisation of polysomes on viral mRNAs (59). Additionally, cleavage

of eIF4G by poliovirus 2A protease results in a higher affinity of the eIF4G-eIF4A complex for the poliovirus IRES in association with a more efficient helicase activity, leading to a greater viral translation efficiency (58). Another example of general decrease of cellular mRNAs translation in favour of viral mRNAs translation has been observed during SARS-CoV-2 infection. The non-structural protein 1 (Nsp1) of SARS-CoV-2 binds the 40S ribosomal subunit, with its C-terminal region blocking the mRNA channel of the ribosome, thus interfering with host mRNA translation, while the 5'UTR of the genomic viral RNA increases translation efficiency of its own (60).

Viral genomes are optimised regarding the amount of information carried by size-limited nucleic acids molecules. A lot of viral mRNAs are polycistronic, with sometimes the different ORF overlapping each other. Therefore, the expression of different viral genes depends on translational events such as upstream ORF translation, ribosomal read-through and hopping, translation reinitiation after termination, and leaky scanning. Ribosomal frameshifting is also a widely used mechanism by viruses to produce viral proteins. Ribosomal frameshifting occurs during the elongation phase of translation when the translating ribosome stalls at a slippery sequence because of an RNA structure like a pseudoknot or the presence of a stretch of rare codons. Influenza A virus endonuclease PA-X is produced by ribosomal frameshifting (61). PA-X is an important protein for influenza A virus virulence, and its main function is to limit host gene expression, mainly by degrading host mRNAs right after the maturation step (62, 63). Being able to compact different ORF in the same transcript spares energy for recruitment of the translational machinery, but the tight regulation of the translation of the different ORF is necessary to ensure viral proteins are produced in necessary amounts (57).

The translational machinery of the hosts is used by viruses in many ways. IRESs, pseudoknots, and stem-loops are RNA structures playing an important role in viral, but also cellular, mRNA translation regulation. We are going to see now why these structures are essential for the cells.

II – mRNA structure and its influence on its processing

A – RNA secondary structures and functions

1 – RNA modifications impact on functions

Nucleic acids are molecules that can be structured if put in the right conditions, which are often met in the cellular environment. DNA is mostly double-stranded in cells, with a helical structure that is opened during few specific cellular events, like DNA transcription and replication. Contrary to DNA, RNA molecules are mainly single-stranded and free to adopt a wide range of secondary structures, and even tertiary structures, in the cells (3, 4).

RNA can adopt diverse structures through canonical base-pairing interaction, as well as through non-canonical interactions (4, 64). RNA plays a central role in many cellular processes and is tightly regulated via its sequence or its structure. Different types of RNA exist in the cells, mainly categorised on their functional features, but also some physical ones. Non-coding RNAs are mostly predicted to be highly structured, more than the mRNAs. For non-coding RNAs, the structure is usually crucial for their function, such as the tRNAs which adopt a functional tertiary structure, or their processing, like the pre-miRNA which must adopt a stem-loop conformation before cleavage and miRNA formation (65).

In cis elements can control RNA functions such as RNA chemical modifications and structures. These two parameters can evolve rapidly and by doing so help the cells to adapt quickly to fast-changing environments. For example, mRNAs are tightly controlled during their life, from their processing where the pre-mRNA is transformed into a functional mRNA, to its localisation, translation, and stability. The presence of m6A on mRNAs is controlled by two types of enzymes: m6A writers, like METTL3 or 4, which add m6A on the mRNA, and m6A erasers, such as demethylase FTO and ALKBH5. Then, m6A readers bind m6A-transcripts, and influence mRNA processing. m6A modifications affect the splicing of mRNAs, with around 30% of m6A modifications found in introns. Misregulation of these intronic m6A is associated with defects in gene expression (4). The translation of some mRNAs is also mediated by m6A. m6A are rare in the 5'UTR and mostly found in 3'UTR and CDS, near the stop codon. The response to m6A modifications regarding the translation of the mRNA depends on its localisation. For example, 5'UTR m6A can mediate cap-independent translation initiation, with direct interaction between eIF3 and m6A. As m6A writing and erasing is highly dynamic, it allows the control of mRNA translation following the cellular context. For example, in

normal conditions, FTO removes m6A from heat-shock responsive mRNAs like *HSP70*, preventing their cap-independent translation. During heat-shock, YTHDF2 translocates to the nucleus, binding the m6A in the 5'UTR of heat-shock responsive mRNAs, preventing demethylation by FTO and by doing so promoting cap-independent translation, allowing selective translation after heat shock (3, 4).

These RNA modifications can stabilise or disrupt structural elements on the mRNA, and thus control mRNA functions like RBP binding (4). For example, m6A destabilises RNA duplexes, and thus RNA structures. m6A in 3'UTR induce destabilisation of 3'UTR structures, giving the possibility for miRNA to bind the mRNA, by making the miRNA binding site available, leading to mRNA decay (4).

Bioinformatics analyses found that mRNAs were the less structured RNA category, although this depends on the metric used for evaluating structure content in RNA (4, 66). UTR of mRNA are consistently found more structured than the CDS (66), but still structures can be folded in the CDS. Structures on the mRNA have been associated with almost all the steps of the mRNA life, starting from its formation during transcription, to its maturation, cellular transportation, and translation (4).

Riboswitches are tertiary RNA structures on the mRNA, which serve as adaptative regulatory elements following the cell state. These structures can adopt two mutually exclusive conformations. Changes in metabolites concentrations switch the riboswitch structure from a metabolite-bound to a metabolite-free conformation, or *vice versa*, hence modifying gene expression (67).

RNA structures play different roles throughout an mRNA's life. Here, we are going to focus on a type of RNA structure that caught more attention these recent years: the RNA G-quadruplexes (G4s).

2 – RNA G-quadruplexes

a – RNA G4 structures and their predicted distribution in the transcriptome

G-quadruplexes (G4s) are non-canonical secondary structures formed by G-rich nucleic acid molecules (68, 69). These were first noticed in the 1910s by Ivar Bang, who saw that high concentrations of guanylic acid solution have a specific viscosity, and can form a gel-like substance in certain conditions (69, 70). Fibre crystallography data, then supported by other biophysical experiments, showed that G4s consist in a stable four-stranded conformation of DNA or RNA, where G-quartets (also called G-tetrads) are stacked on each other and stabilised thanks to Hoogsten hydrogen bonds and the presence of a monovalent cation at the centre of the G-quartet (68–71). This structure can be intramolecular, where all guanine bases involved are belonging to the same nucleic acid molecule, or intermolecular, where the different strands involved come from different nucleic acid molecules (69) (**Fig. D**). For unimolecular and bimolecular G4s, the bases linking the G-stretches participating in the G4 structure form what is called a loop.

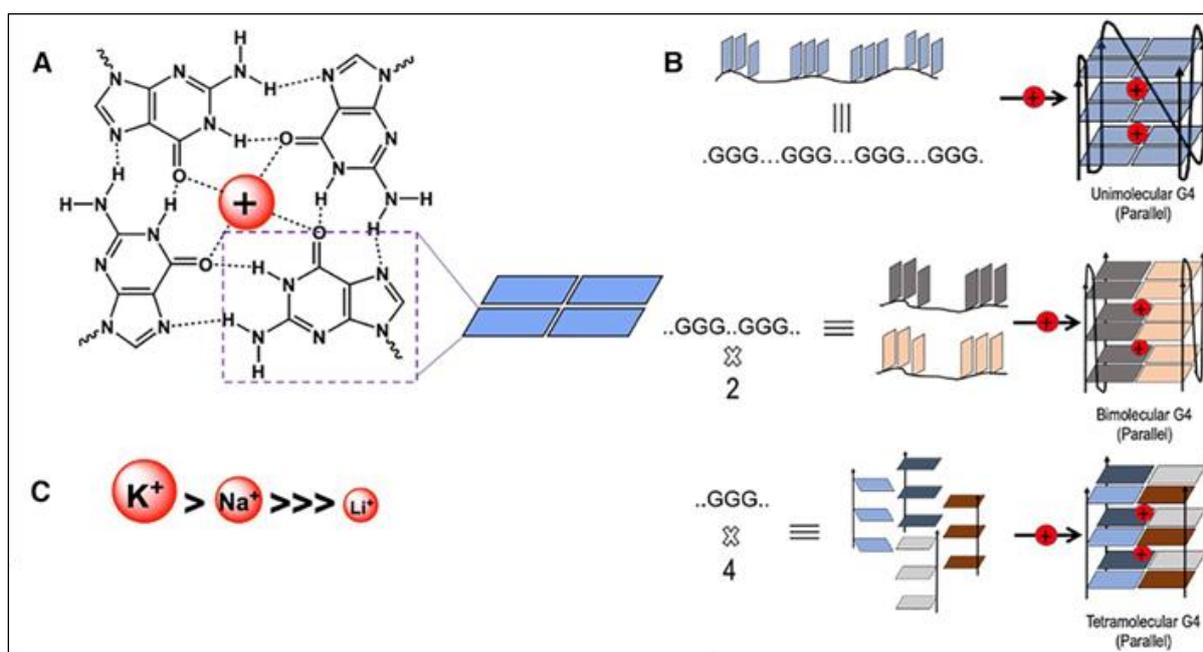


Figure D: RNA G-quadruplexes (G4s) (from (69)): **A:** A G-tetrad (also called G-quartet), a square planar arrangement of 4 guanines stabilised by a cation. Each guanine is then represented as a square in B. **B:** unimolecular (intramolecular) and multimolecular (intermolecular) G4s. Different G-quartets stacking on each other form a G4 structure. The guanine bases forming the G4s can all come from the same molecule (unimolecular G4, top one) or coming from different molecules (bi- and tetra-molecular G4, middle and bottom ones respectively). **C:** G4s stability according to the monovalent ion present in the structure. The K^+ monovalent ion gives more stable G4s than the ones with Na^+ in their centre.

A diverse range of G4s exists, related to the different loop lengths and compositions possible, explaining partially the different thermodynamic stability potential between G4s (68, 70). The nature, and most importantly the size, and the availability of the monovalent cation is a key feature for the stabilisation of the G4 structures, although DNA G4s seem more sensitive to these features than RNA G4s (70). The G4s can have different topologies, named after the orientation of the strands composing the structure (parallel, antiparallel, or mix). Until recently, all RNA G4s known have a parallel conformation, with one exception being the artificial spinach aptamer, which has an anti-parallel conformation (70). In addition, RNA G4s are more compact, less hydrated and have shown greater stability *in vitro* compared to their DNA counterparts (69, 72–75). Although RNA G4s are showing great thermodynamic stability *in vitro*, the different characteristics of an RNA G4 can impact its stability, mainly the loop length and the number of quartets involved in the G4s (72, 76, 77). *In vitro*, RNA G4s are stable in molecular crowding conditions, especially 3 and 4-layer G4s, whereas 2-layer G4s do not seem to be affected by molecular crowding (76), suggesting that this is also an interesting parameter to take into account when considering RNA G4s in the cells.

Using bio-informatics tools to study the transcriptome, potential RNA G4s were identified in up to 6000 genes in humans (69). RNA G4s are mostly predicted and found in the mRNAs UTRs for humans in *in silico* and in high-throughput *in vitro* studies (68–70, 78). The RNA G4s density in the non-coding region of the mRNA is four to five times higher compared to its density in coding regions of mRNAs, and gene ontology analyses predict that G4-forming sequences are mainly found in mRNAs coding proteins involved in RNA processing, regulation of transcription and RNA stability (78). Thus, it was suggested that RNA G4s are an important element for mRNA life, and many studies tend to support this idea.

G4 structures can be targeted by some small chemical molecules, which stabilise them. The specificity of these molecules towards DNA or RNA G4s depends on the molecule properties. Different G4 ligands have been developed and tested, among them TMPyP4, Pyridostatin (PDS), and Bisquinolinium compounds like PhenDC3 or the newly developed PhenDH2 (68, 79). Almost all G4 ligands tested biologically present anti-tumour and anti-viral activities and have been great tools to uncover the RNA G4s functions.

b – RNA G4-associated function in mRNA processing

i – mRNA localisation

As pre-mRNA is synthesised and processed into mRNA in the nucleus, but protein productions mainly happen in the cytoplasm, different mechanisms for the nuclear export of mRNAs have been described.

In specialised cells, like neurons, the transport of mRNA for local translation is critical. It has been shown that 30% of dendritic mRNAs bear G4-forming sequences in their 3'UTR, including *APP*, *BDNF*, *PSD-95*, and *CaMKII α* mRNAs. Mutating these G4-forming sequences causes failure in proper localisation of these mRNAs, with physiological consequences for the plasticity and the overall function of the neuron (69, 80). TDP-43, the mRNA transporter protein which dysfunction is associated with multiple neurodegenerative diseases, recognises its target mRNAs like *PSD-95* and *CaMKII α* mRNAs via their 3'UTR G4 motifs, and transport them to neurites for local translation (81). Therefore, G4s are considered as neurite-targeting signals in neurons (80).

ii – RNP granules formation and composition

In cells, many membrane-less organelles, also called RNP (ribonucleoprotein) granules, are formed by transient RNA-RNA, RNA-protein, and protein-protein interactions condensed in one part of the cell at one moment (71). RNA G4s, with their ability to bind many different partners, as well as their ability to form multimeric RNA structures, are perfect candidates for the formation of RNP granules, or the deregulation of already formed ones. Moreover, the fact that G4 structures can form gel-like substance *in vitro* is in favour of them being at the core of RNP granules, as the same biophysical event is at play, namely liquid-liquid phase separation (69, 71).

In ALS and FTD, expansions of the intronic GGGGCC (G4C2) hexanucleotide repeat in the *C9ORF72* gene can form RNA G4 (82, 83). The expansion of this repeat is directly correlated with formation of r(G4C2)-foci and the development of the disease in patients (69). It has also been shown that some RBPs are trapped in the foci, supposedly via interaction with the RNA G4 structures. HnRNP H (heterogeneous nuclear ribonucleoprotein H) and TDP-43 are two proteins

trapped in these foci, and therefore they are not functional anymore in the cells, leading to pre-mRNA splicing and mRNA localisation misregulation, respectively (81, 84).

iii – mRNA maturation (splicing and poly-adenylation)

Pre-mRNA processing leading to the formation of a mature mRNA involves many modifications of the RNA molecules, such as the addition of an m7G cap in the 5' extremity of the mRNA, polyadenylation at 3' end, and splicing of the mRNA, the latter process possibly giving different mRNAs from a same native transcript. For some mRNAs, the splicing process and the polyadenylation are impacted by RNA G4s.

The polyadenylation of the *TP53* mRNA is sensitive to the cellular stress signal. Under normal conditions, the 3' end processing of *TP53* pre-mRNA is inhibited by G4-forming sequences present downstream of the cleavage site. When cells are stressed by a DNA damage-inducing treatment, the overall 3' end processing in the cells is downregulated, except for a few transcripts like the *TP53* pre-mRNA. It has been proposed that hnRNP H/F, a splicing regulator, recognises G4 structures in the 3'UTR of *TP53* pre-mRNA, and promotes the 3' end processing and polyadenylation of this specific mRNA in response to DNA damage (85). The DEAH-box helicase DHX36 (also known as RHAU or G4R1) is also involved in *TP53* pre-mRNAs polyadenylation, by binding to the 3'UTR G4s (86). G4-stabilising ligands like PhenDC3 prevent DHX36 binding to G4s and the subsequent unwinding, stopping the polyadenylation of the *TP53* pre-mRNA (86).

RNA G4s in pre-mRNAs are also mediators in the splicing process. They have been described either as splicing enhancers, as in the case of the exon-located RNA G4s in the *FMR1* pre-mRNA or the intronic RNA G4s in *TP53* mRNA, or as splicing silencer, like the intronic RNA G4 in human telomerase (*hTERT*) pre-mRNA (69). RNA G4s have also been associated with alternative splicing regulation. To illustrate this, we are going to focus on the β -Site amyloid precursor protein (APP) cleaving enzyme 1 (BACE1), the transmembrane aspartyl protease that catalyses the first cleavage step in the proteolysis of the APP into amyloid β peptides (A β). Human APP can be processed in two different pathways, called non-amyloidogenic and amyloidogenic pathways. In the non-amyloidogenic pathways, α -secretase enzymes like ADAM10 cleave the extracellular domain, producing a short soluble N-terminal fragment and leaving a membrane-bound 83 kDa fragment. In the amyloidogenic pathway, APP is first cleaved by the β -secretase BACE1, which produces a 99 kDa C-terminal moiety. This region is further processed heterogeneously by a γ -secretase, resulting in the

production of A β peptides of sizes ranging from 37 to 49 amino acids (87). Accumulation of neurotoxic A β and their aggregation in plaques are a hallmark of Alzheimer's disease (69). Two isoforms of BACE1 exist, the full-length version being the active one whereas the shorter isoform is inactive in the cleavage of APP. The *BACE1* pre-mRNA have G4-forming sequence in the 5' splice site of its exon 3. hnRNP H recruitment to this part of the pre-mRNA via the RNA G4 sequence favours the long splicing variant and thus the production of full-length BACE1, whereas hnRNP H knockdown leads to decrease in full-length BACE1 and A β production from APP (88).

iv – mRNA stability

mRNAs 3'UTR in the human transcriptome can bear binding sites for miRNA, leading to the translational blockade on this mRNA and its decay via the recruitment of miRISC (miRNA-induced silencing complexes), a ribonucleoprotein complex associated with miRNAs (89). RNA G4 structures in the 3'UTR can prevent miRNA binding on the mRNA, by hiding in the structure the target sequence. It is the case with the *PSD-95* mRNA, which bears two RNA G4-forming sequences in its 3'UTR. One of the RNA G4 structures is extremely stable, whereas the folding of the other one is dynamic, dependent on the K⁺ ions concentration *in vitro*. The dynamic G4 exposes the miR-125a binding site when unfolded (90).

v – mRNA translation

RNA G4s have been broadly studied in the context of mRNA translation. In the UTRs, RNA G4s seem to be translation enhancers or silencers, being either roadblocks for the ribosome on the mRNA or recruiting translation factors, depending on the transcript. Although rarer in the CDS, RNA G4s are also associated with elongation-related events.

RNA G4s in the 5'UTR and 3'UTR and translation initiation

In the 5'UTR, RNA G4s are mainly associated with translation inhibition, by blocking the scanning of the 5'UTR by the small ribosomal subunit (70). RNA G4s can be formed in the 5'UTR of *NRAS* and *KRAS* mRNAs, and in both cases *in vitro* translation is impaired by the RNA G4 structures

(91, 92). In addition, treatment with a G4-stabilising ligand decreases KRAS protein levels in transfected cells (92). Additionally, RNA G4s in the 5'UTR play a role in the development of neurodegenerative diseases, like Alzheimer's disease (AD). The amyloid precursor APP degradation into A β peptides is increasing the risk of A β accumulation and plaques formation. However, APP can also be processed via the non-amyloid pathway, where ADAM10 cleaves APP within the A β domain and releases the neuroprotective N-terminal portion of APP, sAPP- α (50). One unimolecular G4 folds in 5'UTR of *ADAM10* mRNA and inhibits its translation *in vitro*. Mutation on the G4-forming sequence led to an increase of *ADAM10* mRNA translation, as well as high secretion levels of sAPP- α in cells (50). Similarly, expansion of (CGG) n repeats in the 5'UTR of *FMR1* mRNA, which is associated with FMRP repression and FXS development, can fold into RNA G4 structures. RNA G4-mediated mechanism, as well as aberrant CpG islands hypermethylation before the *FMR1* main ORF, are suggested to repress FMRP production (50).

However, for some transcripts, RNA G4s in their 5'UTR increase the translation of the mRNA (70), through recruitment of translation enhancer factors. It is the case for the *RON* mRNA, which encodes the tyrosine kinase receptor RON/MTS1R, a protein associated with tumour dissemination. *RON* mRNAs bear G4-forming sequences in their 5'UTR, which can interact with the splicing regulator hnRNP A1. HnRNP A1 is normally located in the nucleus, but in breast cancer cells from patients with bad prognosis, hnRNP A1 is also found in the cytoplasm, where it binds to *RON* mRNA G4 structures and activates *RON* mRNA translation (93).

In the 3'UTR, RNA G4s have been associated with inhibition of translation. In the *APP* mRNA, RNA G4s have been associated with low translation efficiency, and weak APP production without affecting *APP* mRNA stability (94, 95). Although this can prevent Alzheimer's disease development, as production of A β is reduced, diminution of APP expression also results in impaired learning and memory, with a reduced number of dendritic spines (94).

RNA G4s in the mRNA coding sequence (CDS) affect elongation and the future protein processing

RNA G4s in CDS are often a roadblock for the ribosome, leading to pausing of translation elongation followed either by the dissociation of the 80S ribosome and suppression of translation, or translation resuming, depending on the G4 stability *in vitro* and *in vivo* (96, 97).

RNA G4s are predicted in the coding sequence of the *APP* mRNA, and their formation has been confirmed *in vitro*. These elements are targeted by FMRP in dendrites. FMRP association with *APP* mRNA inhibits translation, probably by blocking ribosomal progression and storing the mRNA in RNP complexes. Upon glutamate receptor mGluR5 activation, FMRP is dislodged from the *APP* mRNA, and APP proteins are produced locally in response to the stimuli (98). A potential actor in this mGluR5-dependent dislodgement of FMRP from *APP* mRNA is hnRNP C, which can bind to the same G4-forming sequence in *APP* mRNA CDS. hnRNP C interaction with *APP* mRNAs has been associated with increase translation of *APP* mRNAs (99). Lack of FMRP results in constitutive and increase production of APP independently from mGluR5 activation (98), illustrating the need for spatiotemporal regulation of translation for the function of neuron cells, which is controlled in this case by RBPs.

RNA G4s in the CDS do not always induce arrest of translation and inhibition of translation, but also elongation pauses. These elongation pauses at a specific sequence cause -1 frameshifting *in vitro* (100) and *in vivo* (101) as well as +1 ribosomal frameshifting. For both types of frameshifting, the ribosome pauses at a slippery sequence positioned 4 to 8 nucleotides away from the G4. The codon readthrough depends on the G4 stability (101). Different G4 ligands have different effects on the ribosomal readthrough: PhenDC3 and berberine increase -1 frameshifting (100), whereas TMPyP4 decreases it (101).

As we have seen earlier, translational pauses are needed for the post-translational modifications of some nascent polypeptides chains. In the case of hER α (human estrogen receptor α) production, the RNA G4 in the ORF induces pause in translation *in vitro*. *In vivo*, the G4-forming sequence is associated with post-translational cleavage of the nascent hER α , a step needed to form functional hER α . This suggests that RNA G4s also play a role in post-translational modifications of the encoded polypeptides by inducing programmed translational pauses at a specific moment of the polypeptide chain polymerisation (102).

3 – RNA G4s in viruses

With the key functions they are playing in cellular mRNAs, RNA G4s are also expected to be important elements for viruses. Indeed, many G4 ligands have an anti-viral effect, although it is difficult to exactly know if this is due to their action on DNA G4s or RNA G4s (68).

Indeed, G4s, either DNA or RNA, are predicted in a lot of viruses, like in HPV (Human Papillomavirus) (103), or in the EP0 promoter of pseudorabies virus. In herpesviruses, RNA G4s are conserved and dense in the regulatory regions and repeat regions of immediate early genes (104, 105).

In HIV-1 (Human Immunodeficiency Virus 1) RNA genome, RNA G4s are identified in the U3 region of the LTR promoter (106), in the Nef CDS (68, 107), in the gag gene (108, 109), and in the cPPT (central polypurine tract) (110). All these regions in the HIV-1 genome can form intramolecular as well as intermolecular RNA G4s. These regions are associated with the dimerisation of the two HIV-1 RNA genomes present in one virion, additionally to the DIS (dimer initiation site) region (111, 112). These RNA G4s are also increasing the recombination rate between the two copies of the HIV-1 genome, as these structures help the RNA-dependent RNA-polymerase to change strand easily. Moreover, stabilising the three RNA G4s in the ORF encoding the Nef protein with TMPyP4 treatment leads to decrease expression of Nef. In line with the fact that Nef is fundamental for HIV-1 replication and infectivity, treatment with G4 ligand decreased HIV-1 infectivity (107).

RNA G4s are also predicted in the Zika virus (ZIKV) genome, where G4-stabilising ligands impair the replication of the genome (113). RNA G4s are characterised in the core gene of the HCV genome, as well as in the 3'UTR of its negative strand. Treating HCV-infected cells with G4-ligand leads to these RNA G4s stabilisation, blocking RNA-dependent RNA synthesis, genome replication, and translation inhibition, resulting in decrease HCV replication in cells (114, 115). In the Ebola Zaire virus (EBOV) L gene, RNA G4s are also forming *in vitro*, and treatment with TMPyP4 led to the inhibition of RNA-dependent synthesis and decrease in the L gene mRNA translation (116).

Among DNA viruses, herpesviruses have a high density of G4s predicted. LANA1, the GMP (genome maintenance protein) of KSHV (Kaposi's Sarcoma Herpes Virus), encoded in its ORF73, is essential for the virus but present at very low levels. The Central Repeat (CR) domains of LANA1 are important for viral immune evasion and inhibition of translation (117). Our team and Dabral *et al.* (118) found that RNA G4s can be formed in these CR domains, and that antigen presentation and full-length protein production inhibition are RNA G4-related. Dabral *et al.* showed that RNA G4s in

the *LANA1* mRNA interact with hnRNP A1, and that TMPyP4 treatment increases antigen presentation (118).

B – RNA-binding proteins

1 – RNA-binding proteins (RBPs) and RNA binding motif (RBM) recognition

RBPs (RNA-binding proteins) are proteins with the capacity to bind RNA. The degree of specificity for the RBP binding on RNA depends on the protein, with some RBPs recognising specific RNA structures or RNA sequences, called RNA binding motif (RBM), while other RBPs binding RNAs widely. With the increasing awareness that RBP-RNA interactions play an important role in the cells, a high-throughput technique called RNA-interactome capture has been developed to study RBP-RNA interactions at the cellular level (119). RNA-binding domain on RBPs can be divided into two groups: the globular domains and the repetitive disordered motifs.

Globular domains on RBPs are structured domains that are found in some families of RBPs. This category includes domains like RRM (RNA recognition motif) domain, KH protein domain (hnRNP K-homology domain), zinc finger domains, and SAP (SAF-A/B, Acinus, and PIAS) domains (119). In some cases, they have been associated with recognition and binding to specific RNA-binding motifs, like the KH protein domain in DDX43 which allows specific RNA binding before unwinding (120), or the RRM and qRRM (quasi RRM) domains in the hnRNP family (69). RNA G4s are recognised by some RNA-binding domains specifically, as hnRNP F qRRM domains which bind to RNA G4 on *Bcl-X* pre-mRNA, controlling alternative splicing of the mRNA (121, 122).

The second group of RNA-binding domains is characterised by short repetitive disordered motifs with low complexity, with a prevalence of G (Glycine), R (Arginine), K (Lysine), and Y (Tyrosine) amino acids, like in RGG, SR, and YGG motifs. hnRNPs, FXS proteins, and DEAD - and DEAH-box helicases are RBPs bearing RGG or RG-rich domains (69). For example, FMRP recognises RNA G4 structures on its target mRNAs, like in the 5'UTR survival motor neuron domain containing 1 (SMNDC1) mRNA or in the coding sequence *APP* mRNA, via its RGG-rich domain (50). Interestingly, these repetitive motifs in disordered regions of RBPs are conserved in non-homologous proteins, suggesting an important role for disordered domains in RNA biology (119).

On the RNA side, evaluating the actual presence of RNA structure *in vivo* is difficult, thus it is hard to conclude about the status of the RNA when identifying an RBP. Nevertheless, Dominguez *et al.* conducted an *in vitro* study of RNA binding motifs (RBM) recognised by 78 human RBPs. Their studies show that both RNA structure and RNA sequence are targeted elements by RBPs. RBMs can also be bipartite, meaning that they are separated into two parts on the RNA sequence, and they underline the fact that flanking regions can also affect RBP binding on RBMs (123).

2 – Nucleolin: a multifunctional RNA-binding protein

The nucleolin (NCL) is a multifunctional phosphoprotein, with multiple phosphorylation sites in its N-terminal part. NCL is mainly localised in the nucleus, more specifically in the nucleolus, the ribosome biogenesis centre of the cell, where it is the most abundant non-ribosomal protein found. NCL can also be found under some circumstances in the cytoplasm (124).

NCL is divided into three main structural domains. The N-terminal domain contains acidic stretches, the central region contains four RRM in animal, two in yeast and plants. The C-terminal domain is glycine-arginine rich, with RGG motifs (124, 125) (**Fig. E**).

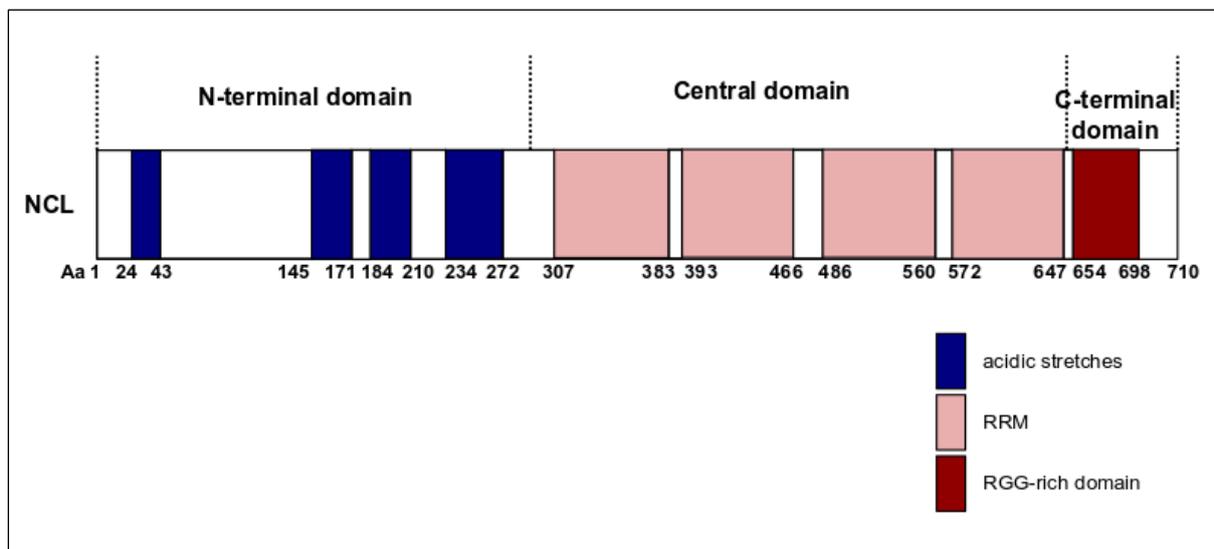


Figure E: Nucleolin protein in humans and its different binding domains

NCL has a DNA helicase activity, a DNA-dependent ATPase activity, and a self-cleaving activity (124, 125). NCL can bind proteins via its RGG-rich domain, as well as specific nucleic acid sequences via both its RRM central and RGG C-terminal domains (124). NCL is involved in pre-rRNA synthesis and maturation. Indeed, NCL represses rRNA production by inhibiting RNA polymerase I activity. NCL also binds the stem-loop structure in pre-rRNA present in the 5'external transcribed spacer and induces the cleavage of the rRNA precursor transcript. As it can bind ribosomal proteins, NCL is involved in the early steps of pre-ribosomal particle assembly and since this protein interacts directly with several ribosomal proteins it may participate to their nuclear import (126, 127). In addition, NCL is also involved in DNA metabolism, by regulating DNA replication, telomere maintenance, and DNA repair. NCL, by binding to histones, is also a chromatin modulator and thus can regulate RNA polymerase II activity and some mRNA synthesis (124, 125). In addition, NCL has a role in pre-miRNA processing (128). Finally, NCL function follows the cell cycle stages, and its N-terminal domain can be phosphorylated by cell cycle control proteins like cell division control protein 2 homolog (Cdc2), casein kinase 2 (CK2), protein kinase C (PKC), and cyclin-dependent kinase 1 (CDK1). Thus, NCL is a key element for cell growth and cell proliferation.

Following this statement, NCL expression is upregulated in cancers, both at the mRNA and protein levels. In this context, NCL has been shown to promote cell proliferation and survival. Upregulation of NCL in cancer cells could contribute to increase ribosome biogenesis, which is crucial for carcinogenesis. Furthermore, NCL is also involved in the regulation of oncogenic transcripts and miRNAs (128).

Indeed, NCL can bind to mRNAs, mainly on AU-rich elements (ARE) in the 3'UTR of some mRNAs, like on *bcl2* mRNA, for which NCL binding is associated with enhanced mRNA stability, but also to G-rich sequences, notably the ones predicted to form RNA G4 structures (125, 129). In one study, NCL seems to promote translation of mRNAs bearing G-rich motifs, regardless of the position of the G-rich motif in the transcript (129). *In vitro*, NCL has been shown to bind both DNA and RNA G4 structures (130–133). For DNA G4s, NCL discriminates G4s and bind preferentially to parallel G4s with 5 to 9-nucleotide length loop rather than the ones with short loops (3 nucleotides). Further experiments show that the RGG C-terminal domain of NCL is responsible for G4 structure recognition, whereas the RRM 2 domain gives the loop-length specificity (132). Until now no work has looked at the potential specific NCL binding towards some categories of RNA G4s.

NCL is also involved in viral processes, from viral entry in the host cells to viral RNA translation and viral replication (128). Cell surface NCL is an element promoting either viral

attachment or entry in the cells for some viruses, like HIV-1 and some influenza viruses (H1N1, H3N2, H5N1, H7N9) (128). In HIV-1 infected cells, NCL stabilises the DNA G4s present in the LTR promoter and blocks HIV-1 viral transcription (134). Upon dengue virus (DENV) infection, NCL interacts with the DENV C protein, a viral protein involved in the structure of the virion as well as in the regulation of the viral genome replication. Treatment targeting NCL decreased DENV infection (128). Similarly, NCL interacts with and stabilises the RNA G4s in the core gene of HCV, and suppresses HCV viral replication and gene expression (135).

3 – RBP – RNA interactions are dynamic

We have seen previously that interactions between RBPs and RNA G4 structures or G4-forming sequences are important for the regulation of many functions of the targeted mRNAs. As much as the processing on the mRNAs is dynamic, RBP-RNA interactions are dynamic and sensitive to external stimuli or cellular cues. Indeed, RBP-RNA interactions in the cell are dependent on the cell type. Riechert *et al.* studied the cardiomyocyte proteome and showed that some RBP-RNA interactions are specific to cardiomyocytes (136). Additionally, by using comparative RNA-interactome capture, they show that the RBP-RNA interactome in cardiac cells evolves during cells differentiation toward cardiomyocytes (136). Other studies using this technique showed that RBP-RNA interactions are dynamic in other cellular contexts. During the maternal-to-zygotic transition (MZT) in drosophila, a process occurring at the earliest stage of embryo development, RBP-RNA interactions are evolving. Indeed, before the MZT, the zygotic genome is not activated and maternal RNAs and proteins are essential for embryo survival and development. Progressively, during the MZT, maternal RNAs and proteins are decayed and replaced by products made following the zygotic genome instructions. Comparative studies between pre and post-MZT show that RBP-RNA interactions go through great modification during the MZT (137). Viral infections also trigger modifications at the cellular level of the RNA-RBPome, as it was shown by Garcia-Moreno *et al.* in the context of sindbis virus infection. In this context, modifications in the RNA-interactome are linked to changes in RNA availability, host cellular RNAs being less available for RBP interactions and viral RNAs being more numerous. Viral infection also activates some RBPs (138).

Availability of a specific RNA structure or sequence is an important feature regulating RBP interaction with RNA. RNA structures can control accessibility to RNA by being the targeted element or by hiding an RNA target sequence in their structure (66). Reversely, RBPs can stabilize, unwind, or

prevent structure formation. Competition between different RBPs for the same target mRNA has been observed, for example in the case of hnRNP C and FMRP binding to *APP* RNA G4s in the coding sequence (98, 99).

Chemical RNA modifications are associated with the regulation of RBP binding: m6A, the most abundant and dynamic RNA modification in eukaryotic cells, is an RBP regulator (139). m6A modifications can alter the mRNA structure and this leads to changes in RNA-RBP interactions. For example, hnRNP C binding to the RNA is promoted by m6A-induced changes in RNA structure. hnRNP C binding to these RNAs promotes pre-mRNA processing and affects the splicing pattern of the targeted mRNAs (140).

Oncogenic signalling pathways also alter RBP-RNA interactions. We have seen it via the activation of eIF4F, where an increase in eIF4E and eIF4A availability also means more interactions of eIF4F with the targeted mRNAs (3). Oncogenic signalling pathways are also characterised by positive auto-feedback regulation, where RBP-RNA interactions are key for activation of oncogenic mRNA translation. For example, downstream of the mTORC2-Akt-HSF1 signalling pathways, the RBP Hu antigen R (HuR) is activated and promotes translation of the mRNA encoding Rictor, a subunit of mTORC2, by binding to its 3' UTR. This leads to increase production and activation of mTORC2 and has been associated with glioblastoma cells invasiveness (3, 141).

Previous work in our lab showed that the TP53 mRNA – MDM2 protein interaction depends on the *TP53* mRNA structure. MDM2, an E3 ubiquitin ligase, can bind the p53 protein, a tumour suppressor, and trigger p53 ubiquitination, leading to p53 degradation. The binding of MDM2 on the *TP53* mRNA is done through the RING domain of MDM2. The MDM2 – *TP53* mRNA interaction inhibits MDM2 ubiquitination activity on p53 protein and promotes *TP53* mRNA translation, leading to increase levels of p53 in the cell. A silent mutation on p53, which affects the RNA structure but not the p53 protein sequence, leads to weaker binding of *TP53* mRNA on MDM2, limiting the anti-tumour activity of p53 (142, 143).

As the interest in RBP-RNA interactions dynamic regulation in the cells is spreading, increasing amounts of data are pointing toward the idea that RNA structures in the cells are dynamic.

C – Dynamics of RNA structure in the cells

1 – Are RNA G4s folded in living cells?

With their high stability *in vitro*, and the optimal free metal monovalent cations concentrations in cells for RNA G4 formation, in addition to the numerous functions they are associated with, RNA G4-forming sequences were assumed to be folded in RNA G4 structures in living cells if they were *in vitro* (70). However, some studies underline the fact that RNA G4s, and RNA structures in general, are not always folded in living cells, and even mainly unfolded in some cases (144).

Since then, numerous studies have aimed to show their presence *in vivo*. Although all techniques present drawbacks in terms of sample preparation, many showed the actual presence of RNA G4s in living cells, using live imaging techniques notably. The development of probes that can be used in living cells, with no treatment or very few sample preparation steps, definitely showed that RNA G4s do fold in living cells (145–147). But these studies also underline the fact the RNA G4s are actually very dynamic structures, meaning they can be folded and unfolded in the cells, implying that a molecular machinery regulates the RNA G4s, and probably RNA structures in general, in the cells.

2 – DEAD-box RNA helicase as RNA structure modulator

With the observation that RNA structures are dynamic, RNA helicases' importance in the cells has grown. RNA helicases are proteins that are able to separate strands in RNA duplexes and thus unwind RNA structures. Among the RNA helicases, a family of proteins presents particular features.

DEAD-box proteins are a family of RNA helicases characterised by the presence of an Asp-Glu-Ala-Asp (DEAD) motif. Members of the DEAD-box RNA helicase family have an RNA-dependent ATPase activity and an ATP-dependent RNA helicase activity, and RNA strand separation is coupled with ATP binding and hydrolysis (148, 149). DEAD-box proteins all have a helicase core which is conserved throughout the family (**Fig. F**). This helicase core is composed of two domains, called domains 1 and 2, which are similar to the bacterial recombination protein recombinase A (RecA). Spread in the helicase core, twelve characteristic sequence motifs are present at conserved

positions, the motif II being the DEAD motif which gives its name to the family. The N and C-terminus domains flanking the helicase core are variable among the DEAD-box helicases, and play a role in binding-partner recognition and the specificity of each DEAD-box helicase activity (148). DEAD-box RNA helicases have been associated with virtually all RNA-associated processes in the cells (148).

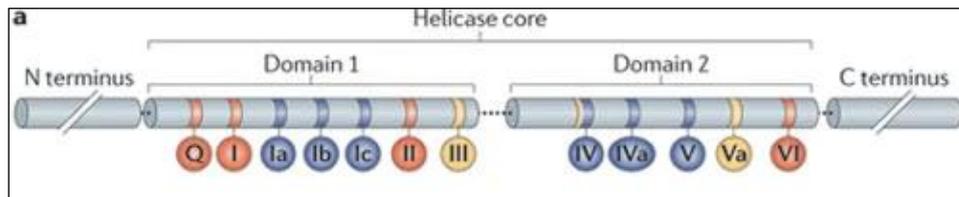


Figure F: Helicase core of the DEAD-box proteins (from (148)). In red: regions involved in ATP binding and hydrolysis, in blue: RNA binding regions, in yellow: communication between ATP binding and RNA-binding site

The functional particularity of this family of RNA helicase is that their RNA unwinding activity is not associated with translocation of the protein, contrary to other nucleic acids helicase. This means that a DEAD-box helicase can unwind locally RNA structure without moving, or with very little translocation (148–150).

eIF4A, the DEAD-box RNA helicase component of the eIF4F translation initiation complex, is the protein from which the DEAD-box RNA helicase family has been characterised (148, 149). eIF4A has three isoforms, encoded by different genes. eIF4A1 and eIF4A2 (also called eIF4AI and eIF4AII respectively) are functionally indistinguishable and are often referred to as eIF4A. Despite sharing 65% homology with eIF4A1, eIF4A3 cannot substitute eIF4A1 *in vitro* and seem to be rather a translation suppressor in *in vitro* translation assays (151). In fact, eIF4A3 and eIF4A1 have different functions in the cells. eIF4A3 is a part of the exon junction complex (EJC) and acts as an ATP-dependent clamp on RNA, allowing the binding of the EJC to the mRNA during splicing and afterward (148).

eIF4A has many functions in the cells, the most described one being its role in translation initiation. As an isolated protein, eIF4A, and most DEAD-box RNA helicase, are non-processive helicases, meaning that alone eIF4A unwinds the structure or part of the structure locally, but does not translocate and unwind completely long structures (5). eIF4F alone is a little more processive than eIF4A, but do not unwind long and complex structure in the 5'UTR of mRNAs neither (5, 7). Despite these observations, in transcriptome-wide studies, eIF4A-dependent mRNA translations are

associated with the presence of (CGG)₄ motifs that can fold into two-layered RNA G4 on the mRNA, and are mainly found in oncogene mRNAs (24). This implies that eIF4A is an RNA helicase unwinding specifically G4s. However, other groups found that eIF4A targets principally canonical secondary structures formed by Watson-Crick base-pairing (26, 152). Nevertheless, both studies confirm the capacity of eIF4A to unwind long and complex 5'UTR structures.

EIF4A alone is a non-processive helicase, but its activity is enhanced by the binding of its co-factors. eIF4B and eIF4H both increase eIF4A1, but not eIF4A3, helicase activity and interact with eIF4A1 via the same region, suggesting that they compete for eIF4A1 binding (5, 7, 21). Both eIF4B and eIF4H can bind mRNA via their RRM domains and prevent RNA re-annealing, giving to eIF4A a directionality while unwinding (153). In presence of ATP, the eIF4A1-eIF4AH complex can repeat rapid cycles of RNA unwinding and re-annealing by transitioning between an “open” and a “closed” conformation, consuming one ATP molecule per unwinding cycle (154, 155). Although unwinding of long and stable RNA duplexes requires ATP hydrolyses, unwinding of an RNA structure with low complexity *in vitro* requires less than one ATP hydrolysis per cycle, and sometimes even no ATP hydrolysis is observed. In these cases, ATP binding to eIF4A might be enough to support eIF4A RNA helicase activity (156).

Several chemical compounds have been identified as eIF4A inhibitors. The flavaglines are a family of natural products extracted from the genus *Aglaia*. Among this family, silvestrol is the compound exhibiting the most potent anti-cancer activity *in vitro* and *in vivo*, by enhancing eIF4A binding to RNA and transforming eIF4A in a clamp stuck on mRNAs, thus inhibiting translation initiation (157). Silvestrol has also anti-viral properties. Silvestrol was shown to inhibit Ebola, Zika, and Chikungunya virus replications (158–160), and to delay and decrease viral protein synthesis for the latter (160). Treating Hepatitis E virus-infected cells with silvestrol diminished the number of viral particles released and mitigate the spread of the virus in cultured cells (161).

Another compound targeting eIF4A is hippuristanol, a polyoxygenated steroid extracted from the Okinawan coral *Isis hippuris*. Hippuristanol is targeting specifically eIF4A1 and 2, but not eIF4A3 or other DEAD-box helicases, and binds to eIF4A C-terminal. Hippuristanol inhibits the eIF4A ATP binding activity by blocking eIF4A in close conformation, and this way decreases eIF4A affinity for RNA (162, 163).

In line with the eIF4A-dependency of some oncogenic mRNAs translation, compounds targeting eIF4A show interesting anti-cancer potential. Treating cell lines derived from primary effusion lymphoma (PEL) caused by Kaposi's sarcoma-associated herpesvirus with hippuristanol induces cell

cycle arrest in the G1 phase and caspase activation, leading to apoptosis (25). Moreover, diverse eIF4A inhibitors alleviate the expression of PD-L1 in melanoma cells, thus blocking cancer cells' immune evasion (55).

3 – Dynamic RNA structures and functions

RNA structures are impacted by the RBPs they are bound to. Those RBPs can either stabilise them or unwind them. In the case of the picornaviruses, IRESs mediate the translation initiation of viral mRNAs. For the Hepatitis A virus and the poliovirus, it was shown that eIF4F binding onto the IRES promotes the RNA re-structuration and unwinding, which further enhances viral mRNA translation (58). In zebrafish embryos, Beaudouin *et al.* show that on a global cellular scale, translation controls RNA structure and that highly translated mRNAs are globally unfolded, possibly due to the multiple unwinding rounds from the different ribosomes advancing on the mRNA (164).

In addition to their interplay with and regulation by RBPs, RNA structures are also sensitive and adapt quickly to environmental conditions and cellular cues. Moreover, RNA structures can compete with each other, at least *in vitro* (78). In bacteria, riboswitches are an example of fast-changing RNA structures according to metabolites concentrations (67). In humans, some environmental-responsive RNA switches are also identified, although their response to environmental cues is also RBP-dependent. The *VEGF* mRNA 3'UTR can adopt two different conformations, depending on the cellular environment. In presence of IFN- γ , the IFN- γ -activated inhibitor of translation complex (GAIT) is active and silences *VEGF* mRNA translation, by constraining the 3'UTR of *VEGF* mRNA in a translation-suppressing conformation. In the condition of hypoxia, translation of *VEGF* mRNA is enhanced in response to a need for angiogenesis. On the cellular level, hypoxia conditions in presence of IFN- γ inhibit the proteasomal degradation of hnRNP L, which binds the *VEGF* mRNA and prevents GAIT binding, therefore stabilising the translation-permissive mRNA conformation. Interestingly, hnRNP L-mediated activation of *VEGF* mRNA translation is dependent on IFN- γ presence (165). Taken together, these studies show that RNA structures depend on the cell type and the cell state, and that RNA can respond quickly to change in the environment, with the help of or impacting RBP interactions.

The dynamic features of RNA structures allow regulation of RNA processing in the cells. In zebrafishes, changes in the RNA structures were observed during the early stages of development and throughout the MZT. Key maternal mRNAs regulate the MZT, and modifications in their 3'UTR

mRNA structure impair MZT and the development of the zebrafish embryos. After multiple mRNA translation rounds, *ccna1* and *nanog* maternal mRNAs lose their structure, and decay-prone elements in the 3'UTR of these mRNAs are available, like the miR-430 binding site on the *nanog* mRNA. Thus, dynamic RNA structure regulates the MZT, and mutation leading to modified RNA structure of key maternal mRNA can impair proper MZT (164). The fact that mRNA structures control the accessibility of RBP-targeted sequences is also observed in human cell lines. Ruijtenberg *et al.* demonstrate that RNA structures can hide Argonaute 2-target sites and that ribosomes promote Argonaute 2 binding on mRNA target.

The impact of the RNA structure on its processing depends also greatly on its stability. RNA G4s are often blockroads to the ribosome advance on the mRNA, and their presence induces ribosomal pauses during either the scanning of the 5'UTR or during elongation when in the CDS. Their ability to block the ribosome advance depends on their stability (96, 97, 101), but also on their localisation in the mRNA. Indeed, while the ribosome advances one nucleotide by one in the 5'UTR, the ribosome translocates three nucleotides by three during elongation. The energy required to unfold and pass the RNA G4 structure in the CDS depends on the distance between the ribosome and the RNA G4s. When unfolding of the RNA G4 can be a two-step process, spanning during two ribosomal translocations, less energy is required to unwind the G4 compared to the situation where the RNA G4 has to be unfolded in one step (166).

Taken together, these examples show that RNA structure can control as much RBP binding as RBPs control RNA structure, and this interplay between RNA structure and RBP binding has functional implications.

In line with the fact that RNA structures are dynamic, it has been observed that DEAD-/DEAH-box RNA helicases bind preferentially folded RNA G4 structures, whereas other RBPs rather bind unfolded RNA G4-forming sequences (64). Work by Herviou *et al.* shows that in glioblastoma cells, DHX36 and hnRNP H/F are both needed to maintain RNA G4 unfolded. They suggest that RNA helicase and RBP works by pair, with specific recognition of some RNA-binding motifs and unwinding of the RNA structure by the helicase followed by RBP binding to the unfolded RNA G4 sequence, preventing reformation of the RNA G4 structure (71). Supporting this idea, cooperation between RNA helicase and an RBP has been observed in the processing of G4-bearing mRNAs. Aven-targeted mRNAs, like *MLL1* and *MLL4* mRNAs, bear RNA G4s in their coding sequence, and Aven binding on the RNA G4-forming sequences promotes these mRNA translations in presence of DHX36 (167). Similarly, DHX36 is also involved in the regulation of *TP53* pre-mRNA maturation via the binding of

the RNA G4s in the 3'UTR of the pre-mRNA, likely in cooperation with hnRNP H/F, although this is not demonstrated yet (85, 86).

RNA helicase can also be recruited via recognition of chemical modification on the mRNA. M6A in structured mRNA coding regions are correlated with ribosome pauses, and prevention of methylation on these transcripts leads to a further decrease of translation. Recruitment of YTHDC2, an m6A reader with RNA helicase activity, allows the ribosome to resume translation after unwinding of the RNA structures by YTHDC2 (168).

Overall, RPB interactions with RNA, RNA chemical modifications, and RNA structures affect each other and control the processing of the mRNA. Strengthening this idea, it was recently showed that selective mRNA translation in VRAF and MEK inhibitor-resistant melanoma cells can be impaired by targeting both eIF4A and m6A modifications (169).

III – Translation-dependent mRNA surveillance and protein quality control in the cell via ribosome-associated quality control

Proteins are essential components of the cells, ensuring diverse functions, from structuring the cell to controlling chemical reactions. Proteostasis can be a matter of life and death for them (170, 171). Moreover, defective mRNA translation products can be harmful to the cells. We saw how protein synthesis can be regulated by the mRNA and are going to see how the translation event itself can affect protein synthesis.

A – Translation-dependent mRNA surveillance by assessing the ribosome progress on the mRNA

1 – Nonsense-mediated decay (NMD)

The nonsense-mediated decay (NMD) is a pathway targeting mRNA harbouring a premature termination codon (PTC), but also mRNA with uORF or long 3'UTR (10). Many factors were identified as part of this pathway, but three factors, called the core NMD factors, are needed: Upf (up-frameshift) 1, 2, and 3b. First thought to happen only during the pioneer round of translation, pieces of evidence showed that already translated mRNAs can also be targeted for NMD (172, 173). Two mechanisms are proposed to describe NMD.

First, the exon junction complex (EJC) model supposes that the PTC is far from the poly(A)-tail, and thus when the ribosome encounters the PTC, the PABP cannot trigger normal termination of translation and most importantly is not able to prevent the recruitment of NMD factors. Instead, the EJC, still present on the mRNA during the pioneer round of translation, will recruit NMD factors. The EJC also recruits termination factors eRF3a and recycling factor ABCE1, leading to the dissociation and recycling of the ribosome (10, 174). The second model, called the faux UTR model, has been proposed in yeast, where the Upf factors directly bind the mRNA, triggering the NMD. In both cases, subsequent recruitment of nuclease involved in mRNA decay, like the exonuclease XRN1, leads to decapping and deadenylation of the mRNA before further degradation (10). Nascent chains are also degraded (175).

2 – No-go decay (NGD)

A recently discovered mRNA decay pathway is the no-go decay (NGD). NGD is triggered by ribosome stalling and collision during translation (176), the minimal unit for triggering NGD being the di-ribosome or disome (177). Many elements can cause ribosome stalling and collision, such as the presence of stable secondary structures on the mRNA, tracts of rare codons, or damaged RNA bases (10).

Ribosome collision is detected by Endothelial differentiation-related factor 1 (EDF1) which stabilises the ubiquitin ligase ZNF598 in human cells (Hel2 in yeast) at the collided ribosomes (178). ZNF598 (or Hel2) ubiquitinates the ribosomal protein eS10, also called RPS10 (RPS3 in yeast) (176, 177). In addition, EDF1 also stabilises the Grb10-interacting GYF (glycine-tyrosine-phenylalanine) domain protein 2 (GIGYF2) at collided ribosomes. GIGYF2 interacts with 4EHP (eIF4E-homologous protein), an inhibitor of translation that competes with eIF4E for binding the mRNA 5'-cap (178), thus inhibiting further rounds of translation on the targeted mRNA, and promoting its degradation by recruiting nucleases. Endonuclease Cue2 has been described to cut the mRNA between the collided ribosome, while the exonuclease XRN1 proceeds by decapping the mRNA (10, 177). Ubiquitination by ZNF598 of the 40S subunit of the collided ribosome promotes recruitment of ABCE1 and ribosome dissociation, which is followed by degradation of the nascent polypeptide attached to the 60S subunit by the RQC (ribosome-associated quality control) complex, mainly via ubiquitination of the nascent product followed by proteasomal degradation (10, 179).

However, as we have seen above, ribosome pausing is a necessary event for co-translational processing of the nascent polypeptide (13). It seems that not all ribosome collision triggers mRNA decay via NGD. Further studies showed that ribosome collision and subsequent ubiquitination by ZNF598 can also lead to ribosome dissociation and resuming of translation on the same mRNA by other ribosomes. In this context, the ASC-1 Complex (ASCC) disassembles collided ribosomes and the nascent chains present on the dissociated ribosomes are targeted by the RQC complex (180). How the choice between mRNA decay or translation resuming is made is still under investigation. A mechanism involving the nascent polypeptide has been recently discovered and will be described in more details further (part III.C.1).

3 – Non-stop mRNA decay (NSD)

The non-stop decay pathway (NSD) targets mRNA without a stop codon. In this case, the ribosome stalls while translating the poly(A)-tail of the mRNA. The mechanism resembles the NGD pathway, with the involvement of Cue2 and ZNF598 and ubiquitination of the 40S ribosomal proteins eS10, uS10, and eS1 in mammals (10).

These three mRNA surveillance pathways based on the ribosome state during elongation are similar from a mechanistic point of view, as all sense an abnormal event during translation elongation and trigger mRNA decay, ribosome dissociation and rescue, and degradation of the nascent chains. In addition to these functions, translation-dependent mRNA surveillance is also associated with other functions at the cellular level.

B – Translation-dependent mRNA surveillance pathways: beyond quality control of the nascent polypeptide

1 – Antigenic peptides production

Antigens are little polypeptides presented at the surface of the cells by MHC Class I or II, the whole complex being recognised by immune cells and playing a major role in the immune system. The first proposed origin for the antigenic peptides is the product of full-length proteins degradation by the proteasome. However, in the last years, numerous studies showed that antigenic peptides also come from alternative mRNA translation events, as they can be encoded in the UTRs of the mRNAs (181), or be the product of non-AUG translation and alternate ORF translation (181, 182). Further studies showed that indeed antigenic peptides and full-length proteins productions are distinguishable events (183–185) and that some ribosomal proteins are involved in antigenic peptides but not in full-length proteins production (186).

DRiPs (defective ribosomal products) were first identified as truncated products of rapid mRNA translation, allowing the production of antigen peptides without the need to produce the corresponding full-length proteins (187). New data are now pointing towards the idea that the nascent chains produced and targeted for decay during NGD could also be used for antigen processing (188). Indeed, a recent report showed that peptides resulting from RQC are presented on MHC I (189).

Another source of antigenic peptides is the PTPs (pioneer translation products). Indeed, previous work in our lab showed that pre-spliced pre-mRNA are translated in the nucleus and generate peptides for MHC I (190). Furthermore, mRNAs bearing PTC, and so targeted by NMD and producing no full-length proteins, are also a source for antigen peptides production.

Taken together, these observations show that mRNA surveillance pathways are involved in antigenic peptides production, meaning that these translational events have a more global effect than just controlling an mRNA molecule.

2 – Cellular stress integration

General cellular stress conditions induce high levels of ribosome collisions. Such events are signals for the cell and trigger coordinate stress responses (10). In yeast, ribosome-associated quality control triggers degradation of the nascent polypeptides but is also detected distinctly from other cellular stress, by the Hsf1 (Heat shock factor 1), the proteotoxic stress detector in yeast (191).

In human cell lines, ribosome collisions induce the recruitment and the auto-phosphorylation of ZAK α at the collision site. This event triggers either the activation of the integrated cell response via GCN2, leading to cell survival, or the ribotoxic stress response after UV treatment or amino acid starvation, by activating SAPKs (stress-activated protein kinases) like p38 and c-Jun N-terminal kinase (JNK) (192).

Therefore, translational events also impact the cell fate, in addition to the mRNA one.

The mRNA surveillance pathways presented here monitor the ribosome progress on the mRNA to assess the quality of the translation. These pathways are tightly balanced and regulated, and trigger either the RQC-mediated degradation of the nascent chain and inhibition of translation, or translation pausing and resuming. We are going to see that the quality of the nascent polypeptide itself is also assessed.

C – Nascent polypeptide assessment during translation

1 – The nascent polypeptide nature controlling its encoding mRNA processing

Tubulins are essential elements in the cytoskeleton. It has been shown that cellular tubulin concentrations are autoregulated, and excess of α or β -tubulin proteins lead to α or β -tubulin-encoding mRNAs degradation. More insight about this auto-regulation loop was obtained recently and offers us an example of mRNA translation regulation triggered by the nature of the polypeptide. Lin *et al.* showed that the tetratricopeptide protein 5 (TTC5) specifically recognises and binds to the N-terminal regulatory motif MREI (Met-Arg-Glu-Ile) of β -tubulin and MREC (Met-Arg-Glu-Cys) of α -tubulin, the three first amino acids being the most critical ones for this interaction. Without these N-terminal regulatory motifs, levels of tubulin are not regulated anymore. In addition, TTC5 binds the ribosome near the polypeptide exit site, on the 28S rRNA backbone and ribosomal protein uL24. Mutations in either the ribosome-interacting domain or the MREI-interacting domain of TTC5 impair its interaction with tubulin-producing ribosomes and the autoregulation of β -tubulin production. Thus, both domains are necessary for TTC5 specific and high-affinity binding to *tubulin* mRNA-translating ribosomes, linking the quality control of the peptide with the regulation of its encoding mRNA translation and stability. TTC5 interaction with the ribosome is strongly regulated in cells, where it is prevented until an excess of $\alpha\beta$ -tubulin is perceived (193).

Another example where the nascent polypeptide influences the mRNA processing is the control or modulation of RQC by the nascent polypeptide. It has been recently shown in yeast that with a minimum of 11 consecutive tryptophan residues in a nascent polypeptide, a Rqc2-independent RQC pathway is triggered. This RQC pathway is still Hel2-dependent. When this occurs while the ribosome is stalled on CGA codon cluster, the expected CAT-tailing of the nascent polypeptide, which is catalysed by Rqc2 and leads to the C-terminal extension of the RQC-targeted nascent polypeptide with alanine and threonine residues, does not occur. Thus, it is thought that the nascent polypeptide present in the ribosome exit tunnel modulates the RQC pathway (194).

As presented earlier, the assessment of the nature of the nascent chain during translation is critical for nascent chains proper ER-targeting. NAC and SRP, both nascent-chain binding proteins present at the exit tunnel of the ribosome, control for ER signals on the polypeptide co-translationally. Thus the nature of the nascent polypeptide influences the polypeptide localisation.

2 – Avoiding aggregates formation

Intrinsic disorder domains in proteins are associated with a lot of functions, linked to their ability to bind a wide variety of partners, which can be proteins or nucleic acids (87, 119). The composition of the sequence rather than the order of the amino acids in the sequence is determinant to judge the intrinsic disorder degree of a protein or a domain. Enrichment of the domain in some amino acids, namely R (arginine), K (lysine), E (glutamic acid), Q (Glutamine), S (Serine), P (Proline), G (Glycine), A (Alanine), increases the disorder degree, as it leads to the formation of hydrophobic and net charge domains. The low complexity of the domain is also a factor favouring disorder in proteins (87). However, disorder domains are also natively unstructured, and their misregulation can lead to the formation of protein aggregates (87).

Protein aggregates are an assembly of disordered, misfolded, or unfolded proteins, or can be made of structured amyloids. Aggregates can contain a single type or multiple types of proteins. These assemblies are dynamic, as aggregates can grow over time, by trapping more proteins. They are mainly described after acute cellular stress, which leads to global protein unfolding, or in the case of disease-causing mutations, which make a protein aggregate-prone or impair the cellular folding machinery (195).

Sensing of aggregates in the cells can trigger different pathways aiming at resolving these structures. A first pathway involves the heat-shock proteins (HSPs) and leads to disaggregation of the assembly, followed by either refolding of the proteins or proteasomal degradation. In this case, small HSPs, which are ATP-independent chaperones, coat the aggregates. The interactions between small HSPs and the protein assemblies are dynamic on the outer shell but stable in the core of the assemblies. This pre-coating by small HSPs is followed by the arrival of HSP70, an ATP-dependent chaperone, on the assembly. HSP70 displaces the small HSPs present on the surface of the aggregate and coat it instead, leading to disaggregation of the assembly with the help of HSP100 (196, 197). The binding of the small HSPs seems determinant for the decision-making regarding the aggregate-prone fate between refolding or degradation (197). Other ways for the cells to deal with aggregates are autophagy clearance or secretion of the aggregates in exophers outside the cells (195). Some reports also suggest that aggregates can be sequestered in membrane-less organelles and asymmetrical cell divisions could then make sure that at least one daughter cell is free of aggregates (195).

Failure to clear aggregates or avoid their formation has been linked to numerous neurodegenerative diseases. Two examples that have been detailed earlier are the A β peptides accumulation resulting in the formation of β -amyloid plaques and the development of Alzheimer's disease, or TDP-43 proteinopathies characterised by hyperphosphorylated and ubiquitinated TDP-43 found in proteinaceous deposits in post-mortem brains of patients affected by ALS and FTD. We can also cite the accumulation of PrP (proteinaceous infectious particle) linked to Creutzfeldt–Jakob Disease (CJD), and the (CAG) expansions producing polyQ (poly-glutamine) stretches associated with many neurodegenerative diseases as Huntington's Disease (87, 195).

In the cells, some mechanisms prevent aggregates by avoiding their formation in the first place, by sensing aggregate-prone polypeptides and assuring their proper folding.

The ribosome itself has a nascent polypeptide chain folding activity. A ribosomal tunnel inside the large subunit connects the peptidyl transferase centre to the polypeptide exit of the ribosome, with a length of 80–100 Å and a diameter of ~10 Å at its narrowest point and ~20 Å at its widest point (1). This tunnel in *E.coli* can accommodate a linear nascent chain of approximately 30–35 amino acids (12). The tunnel wall is mainly composed of rRNAs, giving it an overall electronegative potential (1), and a few proteins. Between the constriction point formed by uL4 and uL22 in the tunnel wall ~30 Å from the PTC (1) and the tunnel exit, early folding events take place, like the formation of α -helical elements, but also the formation of small zinc-finger and β -hairpin motifs as well as compacted folding intermediates (1, 11, 12).

The tunnel exit is also a hotspot for protein folding. It is composed of rRNA and four conserved ribosomal proteins, namely uL22, uL23, uL24, and uL29 (1). Some other ribosomal proteins can also be present at the exit site, depending on the species kingdom: eL19, eL31, eL39 are found in archaea and eukaryotes whereas bL17 and bL32 are found only in bacteria (1). A lot of factors can bind near the tunnel exit, to one of the ribosomal proteins present (11, 12). The protein uL23 is the docking site for many ribosome- and nascent-polypeptides-associated factors (1, 12), which control the nascent chain future processing in the cell. Ribosome-associated chaperones also contribute to the proper folding of the nascent polypeptides. In yeast, the RAC (ribosome-associated complex) cooperates with Ssb and forms a chaperone at the exit of the ribosome (**Fig. G**). Ssb interaction with the ribosome is transient. Ssb recognises short peptide segments enriched in hydrophobic (like valine) and positively charged (such as lysine and arginine) residues at the tunnel exit at an average distance of 35–53 residues from the PTC. The RAC complex, a co-chaperone composed of two subunits Ssz and Zuo1, and bound to the ribosome, enhances Ssb ATPase activity

and its substrate-binding activity. It is suggested that the RAC serves as a sensor of nascent chains and promotes Ssb activity by coordinating Ssb ATP-hydrolysis-dependent substrate binding with the emergence of a specific nascent chain. RAC is also present in humans, where it is composed of two subunits called Hsp70L1 and MPP11, respectively homologous to yeast Ssz and Zuo1, and it is hypothesised that HSP70 replaces the yeast Ssb in humans (1, 11).

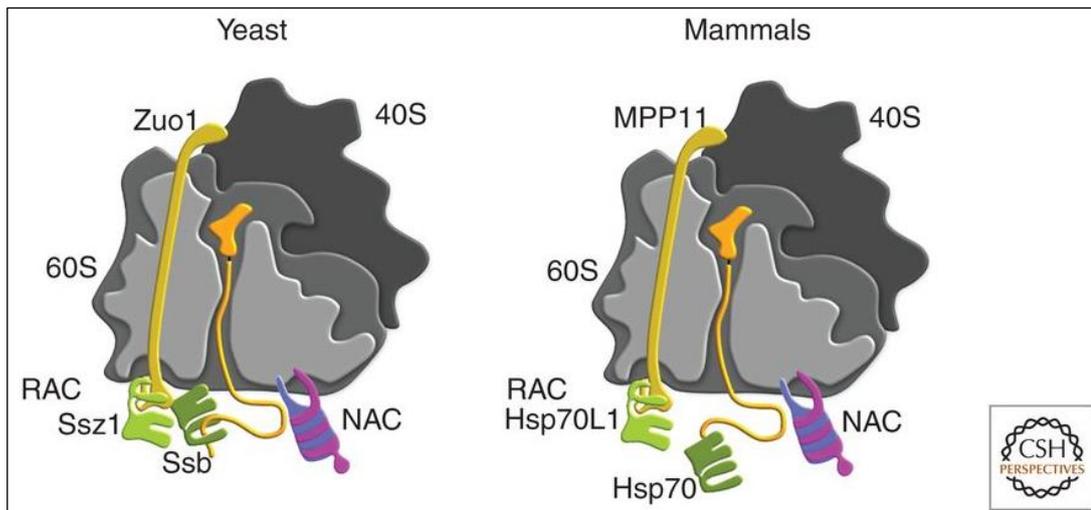


Figure G: Yeast and mammal ribosome-associated chaperones RAC and NAC (from (1)): On yeast and mammals ribosomes, the RAC and NAC complexes are present near the exit polypeptide sites. In both cases, the NAC is composed of two subunits, called NACA and BTF3. The RAC is composed of two subunits, called Zuo1 and Ssz1 in yeast and MPP11 and HSP70L1 in mammals. RAC assists the Ssb chaperone in yeast, and by homology is thought to do the same with HSP70 in mammals.

Another critical ribosome-associated factor that also controlled nascent peptides fate is the nascent polypeptide-associated complex NAC. Indeed, the NAC is associated with proper targeting of the nascent polypeptide to specific cellular compartments, regulation of the nascent polypeptides' interactions with other chaperones, and has a folding and chaperone activity itself.

3 – The nascent-associated complex NAC

The NAC is a heterodimer conserved across archaea, yeast, and mammals and is composed of two subunits, called NACA (also called α NAC or NAC α) and BTF3 (basic transcription factor 3, also called NACB or β NAC) (1, 198). Both subunits contain a homologous NAC domain (**Fig. H**), which dimerises by forming a β -barrel-like structure with a hydrophobic core, the whole forming a stable heterodimer complex (140, 198, 199). NACA can form a homodimer, which is less stable than the NAC heterodimer, and in presence of BTF3, NACA does not form homodimers (198, 200). The C-terminus of NACA contains motifs characteristic of ubiquitin-associated domains (UBAs). Besides these two domains, both NAC subunits are mainly composed of intrinsically disordered regions (1). The NACA subunit has two distinct nucleic acid-binding domains, one binding DNA and one binding nucleic acids in general (199, 200). Interestingly, the nucleic-acid binding region in NACA is covered by a helix region from BTF3 when both associate in the NAC heterodimer. This region is available for RNA or DNA binding when NACA forms a homodimer or is alone (200). Mutations in the nucleic-acid binding region of NACA impairs its nuclear localisation (200). NAC is an essential protein for metazoans, as its deletion in *C.elegans*, *D.melanogaster*, and *M.musculus* are lethal for these organisms, whereas in yeast it does not have any deadly effect (201).

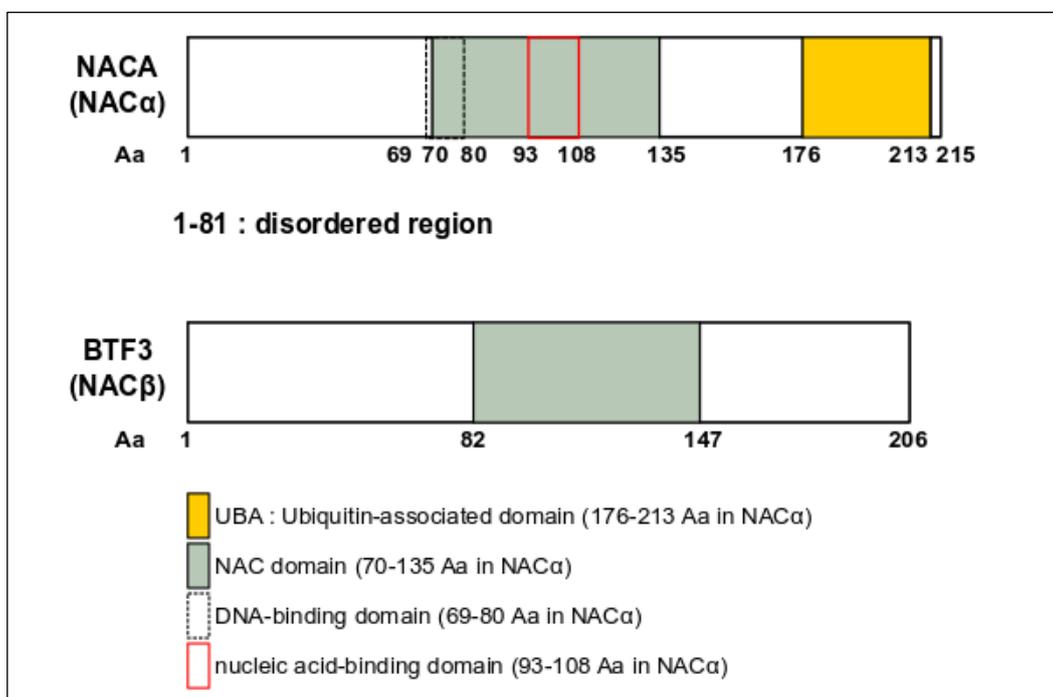


Figure H: NAC subunits NACA and BTF3 and their different functional domains.

Additionally, to its chaperone and ribosome-associated functions that will be described further in detail, both subunits are involved in transcription regulation. Both BTF3 and NACA can be found in the nucleus when they are not associated with each other. BTF3 is a general transcription factor and binds RNA polymerase II, leading to transcription activation (202, 203). NACA forms a homodimer in the nucleus and functions as a transcriptional coactivator. NACA can also bind the TATA box-binding protein in cells (204) and has sequence-specific DNA binding activity (205).

a – NAC prevents nascent polypeptides mistargeting to specific cellular compartments

NAC binds to the ribosome via a conserved motif RRK(X)nKK in the N-terminal region of BTF3. Its docking site is the ribosomal protein uL23, similarly to SRP (12, 206, 207). Many studies showed the importance of NAC for proper and specific ER targeting of nascent polypeptides bearing a signal motif for ER (17, 199, 208–210). The SRP is a complex bound to the ribosome, which recognises ER-targeting element on the nascent polypeptide and allows subsequent ribosome interaction with the SRP receptors on the ER membrane, followed by co-translational translocation of the nascent polypeptide in the ER. However, the SRP binding to SRP receptors is not specific, unless the NAC is present (16, 17). In fact, the NAC does not stimulate SRP binding to the SRP receptor on the ER membrane, but rather prevents the SRP-SRP receptor interaction when the nascent polypeptide does not bear an ER-targeting signal. Crystallography studies show that the N-terminal tail of BTF3 is inserted in the ribosomal polypeptide exit tunnel at the very early stage of translation, and thus can sense the N-terminus of the nascent polypeptide while being very close to the peptidyl transferase centre. As the nascent chain is growing, the N-terminal tail of BTF3 escorts the nascent polypeptide out of the ribosome and then relocates on the ribosomal surface when it gets out of the exit tunnel. This early scanning of the N-terminal part of the nascent polypeptides is needed for proper ER targeting (211). It was also shown that the NAC can change the conformation of the SRP, thus controlling its interaction with SRP receptors. Without the recognition of nascent polypeptide bearing an ER-targeting signal, the NAC keeps the co-bound SRP in a conformational state unfavourable for interaction with SRP receptors. When the NAC senses in the ribosomal tunnel the ER-targeting signal on the nascent polypeptide, it changes the SRP conformation, allowing SRP – SRP receptor interaction on the ER membrane (16). Therefore, the early scanning in the exit tunnel

of the ribosome by the N-tail of BTF3 can regulate SRP conformation and the future location of the nascent polypeptide in the cell.

In addition to preventing ER mistargeting of nascent polypeptides, the NAC also seems to regulate mitochondrial protein transport, by promoting co-translational mitochondrial protein importation in the mitochondria (1). In both cases, the NAC is in contact with the ribosome and can sense the nascent polypeptide, suggesting a possible function of NAC as a chaperone.

b – NAC as a ribosome-associated chaperone

Early studies on NAC have shown that NAC protects short nascent polypeptides from proteolysis and cross-linked with them, regardless of their sequence or their folding state (212), suggesting that the NAC can interact with a wide range of polypeptides. In yeast, NAC is associated with almost all ribosomes and is quantitatively present at a ratio of 1:1 with ribosomes (210, 213). Contrary to the NAC, other ribosome-bound chaperones, like the RAC, are associated with only a fraction of the total ribosome population (213). Finally, the NAC controls SRP and the methionine aminopeptidase interactions with the nascent polypeptide, two events happening early during translation. All three factors can be neighbours on the ribosome at this stage of translation (16, 199, 208, 209, 211, 214). Thus, the NAC has been proposed as being the first chaperone to interact with and sense the nascent polypeptide, controlling its future processing by helping its folding and regulating its interaction with other chaperones. However, apart from the case of SRP and the methionine aminopeptidase, no evidence confirms this idea.

NAC ribosome-associated chaperone function is complementary with the RAC chaperone function, as deletions of both NAC and RAC contribute to a more detrimental phenotype than the deletion of only one of the two (215). Moreover, the pools of NAC and RAC targets do not overlap totally (215), and the RAC binding on the ribosome does not depend on the NAC binding on it (206). Therefore, it has been suggested that both chaperones cooperate when located on the ribosome.

The NAC interaction with the ribosome is complex. Both NACA and BTF3 have a ribosome-binding domain in their N-terminal region, respectively KKAR and RRKKKK (206, 211, 214), but BTF3 interaction seems stronger than NACA interaction with the ribosome (199, 206). The DSD motif present in the N-tail of NACA can inhibit interaction with the ribosome by binding both ribosome-binding domains of the NAC. This auto-inhibition for ribosome binding is relieved by a change in

conformation of the N-tail of NACA, by which the DSD domain binds to ribosomal protein eL19 and releases both ribosome-binding domains (211). Different cross-linking studies show that NAC can bind the ribosome at multiple points: BTF3 can interact with uL23 (206), eL22 (211), eL31 (211, 214, 216), uL22 (211) and NACA can bind uL29 (214), eL19(211), uL22 (216). In all cases, the interactions between the NAC and the ribosome occur next to the exit tunnel, or even in the exit tunnel for uL22. Surprisingly, no studies were able to show all these interactions on the same ribosome – NAC complex, as at best two interactions were detected simultaneously (214, 216).

In addition, the NAC conformation is flexible and can be adapted to different nascent polypeptides (217). Although binding almost all the nascent chains produced by the ribosome, the affinity of the NAC for the nascent chain varies according to the nascent chain sequence (213), and this interaction is in all cases weak (217), suggesting a reversible and dynamic binding of NAC to the nascent polypeptide. Furthermore, when the NAC is cross-linked with the nascent polypeptide, elongation and release of the polypeptide from the ribosome complex is not prevented, and the NAC leaves the ribosome with the polypeptide in this case (212).

Taken together, these observations suggest that the NAC interactions with the ribosome and the nascent chain are dynamic and adaptive, meaning that the NAC could be changing conformation on the ribosome according to the nascent polypeptide nature. Interestingly, the amount of NAC in mammalian cells is greater than the amount of ribosome, implying that either more than one NAC can bind a ribosome or that some free NAC can be found in the cytoplasm (218).

c – NAC as a cytosolic chaperone

In addition to its ribosome-associated functions, the NAC also has a chaperone activity as a free cytosolic complex (219) and can bind both folded and unfolded proteins (217). First, NAC associates to aggregate-prone proteins and prevents aggregates formation independently from the ribosome (201, 217, 219), despite its interaction with polypeptides being relatively weak (217). Secondly, the ribosomal binding domain in the N-tail of BTF3 has a chaperone function, meaning that both functions of this domain can compete (219). Finally, NAC is a sensor and regulator of proteostasis in *C.elegans*. Upon acute stress like heat shock, the NAC localisation is modified from the ribosomes to the unfolded proteins, preventing aggregate formation. In addition, when NAC is localised in protein aggregates rather than bound to ribosomes, the global translation in the cell is slowed down. After heat shock, the NAC is necessary for aggregate clearance, and recruitment of

other chaperones to refold proteins, restoring proteostasis. Thus, changes in NAC localisation in the cells allow monitoring of proteostasis and prevention of further protein synthesis while aggregates are still present (201).

In conclusion, NAC has a tight relation with most nascent polypeptides on the ribosome. However, its chaperone function remains unclear, as it is unknown how the NAC exactly binds the ribosome and the polypeptides and how it transitions on and off of it in cells.

IV – The Epstein-Barr Virus

A – The first human oncovirus identified

The Epstein-Barr virus (EBV) was identified in 1964 in endemic Burkitt's Lymphomas biopsies, using electron microscopy, by Epstein, Barr, and Achong (220). Upon its observation, it was clear that the EBV belongs to the herpesvirus family.

Among the nine species of herpesviruses infecting humans, five of them are widespread: the herpes simplex virus 1 and 2 (HSV-1 and HSV-2), the varicella zoster virus, the Epstein-Barr virus (EBV), and the cytomegalovirus (CMV), causing orolabial and genital herpes, chickenpox and shingles, mononucleosis, and some cancers (68). Herpesviruses have in common a long double-stranded genomic DNA, and the ability to establish latent infection, most of the time asymptomatic or with mild effect on the host. In these healthy carriers, the genome remains as an extrachromosomal episome, replicating and segregating between the daughter cells during cell division, thanks to the viral genome maintenance protein (GMP) (68).

EBV is a good representative of this family: its genome is large, being 172kb-length approximately, and encoding 80 proteins and 46 functional non-coding RNAs. EBV is present in 90 to 95% of the adult population, establishing a life-long latent infection most of the time harmless for the host. In healthy carriers, the persistent infection is in the memory B cell of the host, with around 0.01% of the B cell carrying the EBV genome in this case, although this number varies in function of the individual (220, 221).

Despite being asymptomatic to most of the human population, EBV can cause diseases upon infection. EBV is the cause of infectious mononucleosis (IM), which most of the time does not have long time repercussions on the host health as far as we know. EBV is also associated with the development of specific malignancies, both epithelial and lymphoblastic. It is estimated that 1.5% of cancer cases worldwide are linked to EBV (221). Many EBV-associated cancers are derived from the B cells, which are the preferred cellular hosts of EBV: Burkitt's lymphomas, 30% of Hodgkin lymphomas, and most immunoblastic lymphomas in immunocompromised patients like post-transplant lymphoproliferative disease, X-linked lymphoproliferative disease type 1, HIV and EBV-associated B cell lymphomas, and diffuse large B cell lymphomas. Surprisingly, EBV is also found in some NK and T cells-derived lymphomas. EBV is also strongly associated with some epithelial malignancies, mainly undifferentiated (also called anaplastic) nasopharyngeal carcinomas (NPC) and

10% gastric carcinomas (220, 221). Additionally, recent epidemiological studies suggest that some EBV strains might be associated with multiple sclerosis development (222). In all tumour cells, EBV is present, expressing one of its latency programs (220), but until this day no treatment targeting specifically EBV is widely available (223).

EBV oncogenic activity is supported by the observation that EBV can transform human B lymphocytes into immortalised lymphoblastoid cell lines (LCL) (220, 221). All the cells in LCL carry around 10 copies of the EBV genome, and the copy number is maintained in dividing cells (221).

Geographically, EBV-related cancers are more common in some areas: while the so-called endemic Burkitt's lymphoma is in the vast majority of cases diagnosed in the African population, NPC is the most frequent cancer diagnosed for males in South China and is also relatively frequently diagnosed in the Eskimos and African populations. This specific repartition does not seem to be linked to any specific EBV strains (221). It is thought that both genetic and environmental factors play a role in the development of EBV-related cancers, but no clear factors have been identified yet.

B – EBV latency programs

EBV latent infection in healthy individuals persists in a few memory B cells of the host. How the infection reaches these B cells is unknown, and two hypotheses could explain the infection of memory B cells (**Fig.1**): one favours a direct infection of memory B cells by EBV, whereas the other, called the germinal centre model, favours first an infection of naïve B cells and then the survival of EBV in these B cells while they undergo selection in the germinal centre. According to the last model, EBV infecting naïve B cells express latency program III, also called the growth program, where all EBNAs (-1, -2, -3A, -3B, -3C and -LP) are expressed as well as the LMPs (latent membrane proteins, LMP1, -2A and -2B) and BHRF1 (224). This latency program III promotes infected B cell proliferation until these naïve B cells reach the germinal centre, where normal naïve B cells undergo selection and maturation according to antigens captured and presented by antigen-presenting immune cells in the body. In the germinal centre, EBV switches to latency program II, still expressing EBNA1 and LMP1 and LMP2 proteins which mimic respectively the host CD40 and B cell receptor. This allows EBV-infected B cells to survive the germinal centre triage. At the exit of the germinal centre, EBV is downregulating its expression and is present in resting memory B cell either in latency program 0, where only the EBV episome is detected, or in latency program I during cell division. During latency

program I, only EBNA1 is expressed and ensures the maintenance of the EBV infection by proper EBV episome replication and segregation between daughter cells (220, 221).

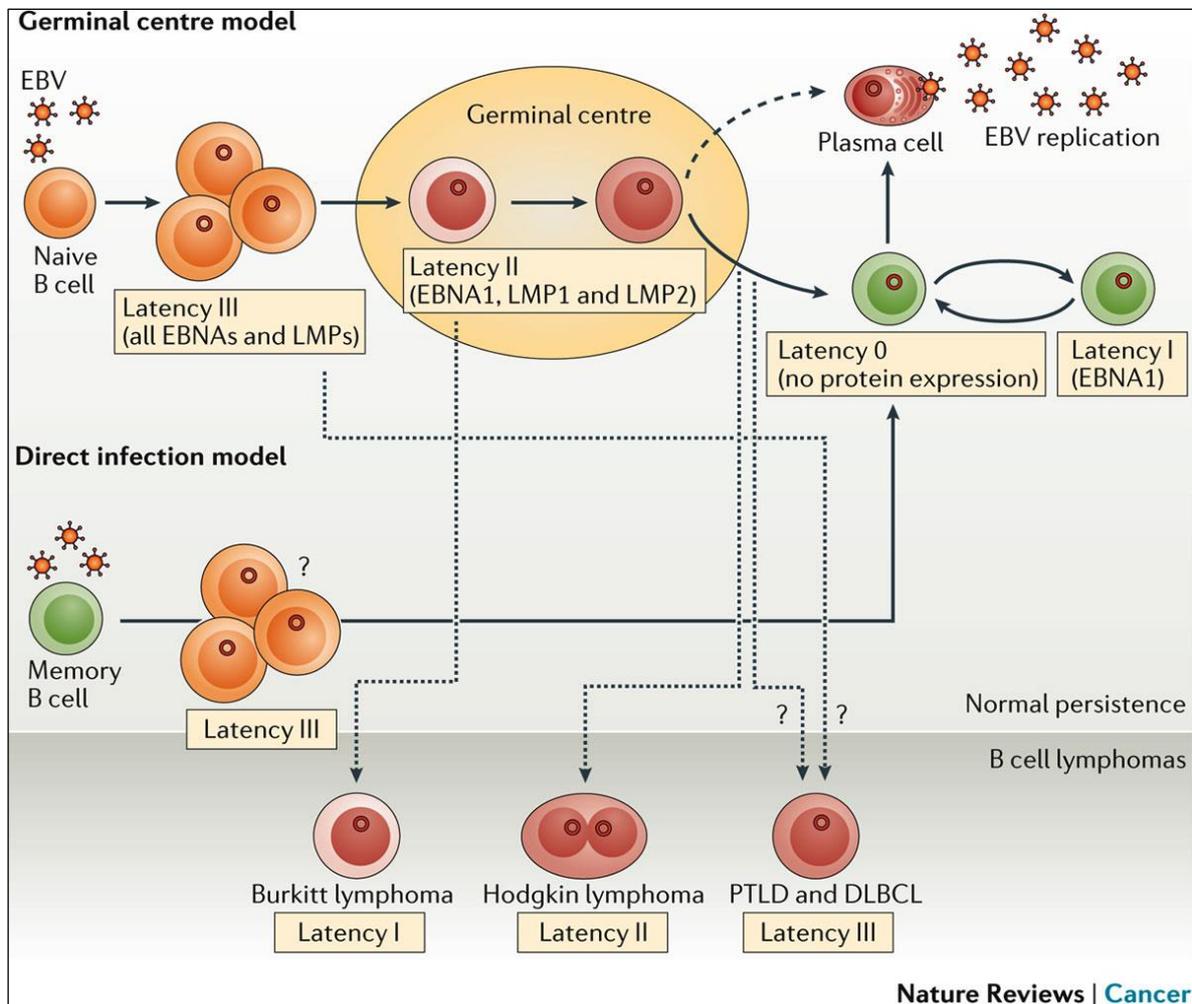


Figure 1: the two EBV infection models (from (220)). On the top, the germinal centre model, where EBV first infect naïve B cells and survive in them while they go through selection in the germinal centre and evolve into mature B cells. In the middle, the direct infection model, where EBV directly infects memory B cells. At the bottom: EBV-associated B-cell derived lymphomas. Latency programs and protein expressed are indicated below each step.

The switch between the different latency programs is controlled by the activation of different promoters. The LMPs have their own promoters, whereas three different promoters control the EBNAs expression: activation of the promoter Cp lead to the production of a primary transcript carrying all EBNAs as well as BHRF1 encoding sequences, which is then further processed to give mRNAs encoding one viral protein; the activation of the Wp promoter leads to the same

protein profile expression as Cp, except for EBNA2 which is deleted. Finally, the Qp promoter leads to the expression of only the EBNA1 protein (220, 221).

The lytic cycle of EBV can be triggered by diverse stimuli (220). In the human host, EBV switches to its lytic cycle when EBV-infected memory B cells are activated by antigen-presenting cells. When these B cells are differentiated into plasma cells, EBV virions are produced. It is also worth noting that many functional non-coding RNAs play an important role in the EBV life cycle (220).

In EBV-infected cancer cells, the virus expresses one of the three latency programs (I, II, or III). The latency program expressed is thought to depend on the B cell maturation stage from which the malignant cells come (220). As EBNA1 is present in all latency programs, it is an interesting target for the development of treatment against EBV.

C – EBNA1

1 – EBNA1: the genome maintenance protein of EBV

The EBNA1 protein is necessary for EBV genome maintenance, as it allows proper episome replication and segregation. Associated with this maintenance function, EBNA1 is also a general key regulator of EBV gene expression.

a – EBNA1 roles in viral genome maintenance

EBNA1 is the genome maintenance protein (GMP) of EBV (**Fig. J**). Structurally, EBNA1 dimerisation and DNA-binding domain is located in its C-terminus. This DNA-binding domain recognises with high affinity the oriP (origin of replication) site on the EBV genome, and dimerisation of EBNA1 occurs when it is binding the EBV episome (221). The N-terminal region of EBNA1, until the DNA-binding domain, is mainly disordered and composed of a glycine – alanine repeats (GAR) domain and RGG-rich domains (221, 225). Each RGG-rich domain of EBNA1 contributes additively to EBNA1 RNA-binding capacity to a wide range of RNA oligos, and especially to poly(G) oligos *in vitro*. In Akata cells, EBNA1 binds via its RGG domains EBV-encoded non-coding RNA *EBER* (226).

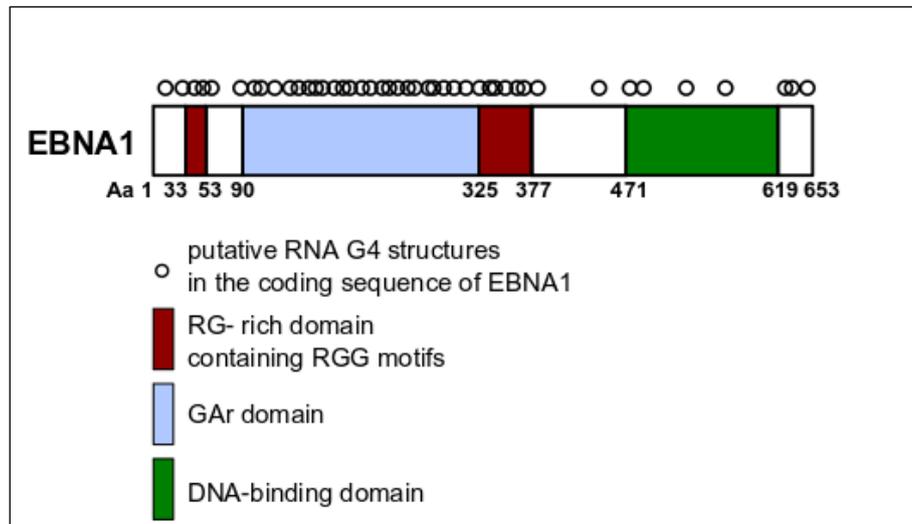


Figure J: The EBV-encoded EBNA1 protein. Additionally to the regions indicated, EBNA1 interacts with USP7 via amino acids 436 to 450. Apart from the DNA-binding domain, the rest of the protein is predicted to be disordered (225).

Recruitment of the origin recognition complex (ORC) to the EBV episome OriP site is mediated by EBNA1 RGG motifs and is dependent on the presence of G-rich RNAs (227). Additionally, in the Raji cells, which are latently infected by EBV, the EBNA1 protein binds its own mRNA (227), which is G-rich and form RNA G4 structures *in vitro* (228) (**Fig. J**). Treatment with the G4-targeting drugs BRACO-19 impairs EBNA1-mediated viral episome replication and proper episomes segregation in daughter cells during cell division, by inhibiting EBNA1 attachment to metaphase chromosomes (229). Thus, it has been suggested that during viral replication, EBNA1 DNA-binding domain binds the OriP DNA sequence and that its RGG-rich domains bind RNA G4s with which ORC is interacting, forming a stable RNP complex (227). Furthermore, the N-terminus of EBNA1 permits attachment of the EBV episome to the host chromosome during cell division in an RNA G4-forming sequence-dependent manner (229) and this also requires human EBNA1 binding protein 2 (hEBP2) (230). Chromosomal attachment of EBV episomes allows proper repartition of the viral episomes between daughter cells and avoid EBV episome degradation as erroneously replicated DNA (221).

The cellular factor NCL is important for EBNA1 function as EBV's GMP. Downregulation of NCL level in cells reduces EBNA1 recruitment to the oriP site and EBNA1-mediated viral replication. In LCL, NCL depletion is associated with impairment in EBV genome maintenance and reduces cell growth and proliferation. Oppositely, NCL overexpression increases EBNA1 binding to OriP. Chen *et al.* also showed that EBNA1-mediated OriP activation is dependent on NCL ATP-binding activity

(231). On the EBNA1 side, the first 100 N-terminal amino acids are necessary for its interaction with NCL. Interestingly, they showed that NCL RNA-binding domain 2 (RRM 2) is critical for both ATP and EBNA1 binding by observing that mutation K429A in the NCL RNA-binding domain decreased EBNA1-related maintenance of the viral episome. However, EBNA1 and NCL interactions are still observed after RNase and DNase treatment, although the authors did not exclude the possibility that RNA or DNA strengthens this interaction (231).

b – EBNA1 controls other viral and cellular gene expressions

In addition to its role in viral genome maintenance, EBNA1 also regulates the transcription of other viral genes. EBNA1 binding to the OriP site promotes LMP promoters (232) as well as the Cp promoter activation (233), leading to activation of latent gene expression. NCL is also involved in the transcriptional activity of EBNA1 (231).

EBNA1 is also a transcription enhancer of cellular genes, like *nox2* or recombinases RAG-1 and RAG-2, all associated with the development of EBV-related malignancies (234).

2 – EBNA1 controls its protein level and escape the host immune system

In healthy EBV carriers, CD4+ and CD8+ T cells against EBNA1 are present and can be activated when put in contact with EBNA1-derived antigens (235–239). Moreover, activated CD4+ T cells limit EBV-associated B-cell proliferation (240) and some attempts using adoptive *ex vivo* transfer of EBNA1-specific T-cell gave promising results but are still limited (241). Despite the presence of immune cells targeting EBNA1, a protein expressed in all malignant EBV-infected cells, EBV escapes the host immune system, by controlling the levels of EBNA1 proteins and antigens, and thus avoiding recognition by the host immune system (242, 243).

a – GAr-associated proteasomal degradation inhibition

The GAr domain of EBNA1 first caught attention for its ability to reduce antigen presentation (243). As the first antigen processing pathway discovered involves proteasomal degradation of full-length proteins, further efforts were put in understanding the effect of the GAr on proteasomal degradation. It was shown that the GAr domain inhibits proteasomal degradation, but not its ubiquitination (244). However, other studies showed that the GAr on its own does not explain the EBNA1 long stability in cells (245). The GAr affects proteasomal degradation in a substrate and position-dependent manner, mostly link to its position in the protein related to other unfolded regions: if the GAr is near other unfolded regions, 26S proteasomal degradation of the protein is inhibited, whereas when the GAr is far from other unfolded regions, the resulting protein is degraded (245). Hence, the GAr is a specific regulator of the 26S proteasome (245).

b – The GAr domain of EBNA1 and translation inhibition

Even though the GAr is preventing its proteasomal degradation in the context of EBNA1, this does not explain entirely why EBV-infected cells produce undetectable levels of EBNA1-derived antigens. Indeed, the team previously showed that inhibiting proteasomal degradation weakly affects GAr-derived antigens production compared to inhibiting translation (183). Since antigens can have a different origin than proteasomal degradation of full-length proteins and be produced by alternative translation events like DriPS and PTPs (181, 187, 190, 235, 246), further studies focused on the impact of the GAr on translation. It was shown that the GAr also inhibits its own mRNA translation via an *in cis* mechanism, hence limiting the production of full-length proteins and antigenic peptides (183, 247). Both GAr-mediated inhibition of proteasomal degradation and inhibition of translation are two distinguishable mechanisms (183). Additionally, the GAr-derived antigens are mainly coming from alternative translation events and antigen presentation depends more on the rate of mRNA translation than on the stability of EBNA1 (235, 248).

The GAr-mediated inhibition of translation is dependent on the GAr peptide length and, like for the inhibition of proteasomal degradation, on its position in the protein, as putting the GAr in the N-terminus of the ovalbumin protein inhibited protein and antigenic peptides production more than fusing the GAr to the C-terminus (183, 247). *In vitro* translation assays also show that this mechanism is dependent on the dose of GAr-encoding mRNA used (247). Pulse-chase experiments

and analyses of ribosomal profiles suggest that the inhibition targets the initiation step of translation. To confirm this, different IRES structures were put in the 5'UTR of GAR-coding sequences, and the *c-Myc* IRES could override the GAR-related inhibition of translation, indicating that some, but not all, translation initiation mechanisms are targeted by the GAR pathway (247).

The GAR-encoding sequence is G-rich, and a first report suggested that GAR-encoding mRNAs can form RNA G4 structures (249), which was confirmed *in vitro* (131, 228). Destabilising the RNA G4s by mutating the sequence cancelled the inhibition of antigen presentation induced by the GAR and resulted in the T cell activation in mice (235). Similarly, using oligos which destabilise the RNA G4s *in vitro* enhances mRNA translation and antigen presentation in cells (228). Treatment with PDS, a G4-stabilising ligand, further inhibited GAR mRNA translation, production of full-length proteins, and antigen presentation (228). However, PhenDC3 and its derivatives like PhenDH2 and PyDH2, which are also G4 ligands, increase full-length protein and antigenic peptides production from GAR-encoding mRNAs (79, 131). The cellular factor NCL is directly binding to the G4-forming sequence in the GAR-encoding mRNA (131) and this interaction is prevented by treating the cells with PhenDC3 or PhenDH2, but not with PDS (131, 250). Despite additional studies, how NCL regulates GAR-encoding mRNAs translation is still unclear (131, 184).

Even though the mRNA sequence and structure are important for GAR-mediated inhibition of translation, a second body of work suggests that the peptide sequence also plays an important role. When using a short 30 amino acid-long GAR, which is enough to trigger GAR-mediated inhibition of translation, the team showed that silent mutations on the RNA sequence do not affect GAR-bearing mRNA translation, whereas small mutations in the amino acid sequence can cancel the GAR-mediated inhibition of translation (247). It should be noted that all the mutations made in this study did not affect the predicted RNA G4 structures (247).

3 – EBNA1 is oncogenic

EBV-associated Burkitt's Lymphoma (BL) pathogenesis is essentially due to the translocation of the *c-Myc* gene in the loci of one of the three immunoglobulin genes (234). In these cells, it has been shown that EBNA1 induction of *nox2* gene leads to increase formation of reactive oxygen species, causing DNA damage and genomic instability (251, 252). EBNA1 also induces the expression of recombinase RAG-1 and RAG-2, which can contribute to the *c-myc/Ig* translocation (253, 254). Thus, EBNA1 can trigger oncogenesis in multiple ways.

Cell growth in healthy cells is tightly regulated, and the uncontrolled cell proliferation characteristic of cancer is due to abnormalities in the autoregulation of cell growth signals. In addition to triggering the development of BL, EBNA1 is also the factor preventing apoptosis of these cells (255). Interaction of EBNA1 with the ubiquitin-specific protease USP7 lowers p53 protein levels and therefore increases cell survival by preventing apoptosis (256). Parallel studies showed that GAR-mediated translational stress triggered by *EBNA1* mRNA translation activates PI3K δ (257). This stabilises MDM2 and promotes MDM2's binding on the *E2F1* mRNA, enhancing its translation (258). MDM2 binding to *E2F1* mRNA is mutually exclusive with its binding to *TP53* mRNA, and Gnanasundram *et al.* showed that MDM2 stimulates exclusively either p53 or E2F1 synthesis, depending on the cellular conditions (258). Consequently to its increased levels in the cells, E2F1 activates transcriptionally cyclins and c-Myc expression, leading to increase ribosome biogenesis and sustained cell proliferation (257). Thus, by preventing apoptosis, *EBNA1* mRNA translation also maintains cell proliferation signals.

V – Aim of my Ph.D. project

As both the GAR polypeptide and the mRNA sequence and structure play a role in the GAR-dependent translation inhibition, the goal of my Ph.D. project is to determine if this mechanism links the nascent polypeptide quality control with the mRNA structure, resulting in the regulation *in cis* of the mRNA translation.

The first part is dedicated to better understand the G4s dynamics impact on GAR-related inhibition of full-length proteins and antigenic peptides production, as well as its effect on mRNA localisation and NCL-binding capacity.

Then, we focused on the interplay between the nascent polypeptide GAR and the NAC and show that this translational event is linked to changes in the GAR-encoding mRNA, notably on the mRNA capacity to bind NCL.

Finally, we tried to understand how the NAC-related quality control of the GAR polypeptide can affect the mRNA NCL-binding capacity, by investigating eIF4A1 involvement in the GAR-mediated translation inhibition.

RESULTS

Paper 1: Dynamic regulation of RNA G4s-associated functions

Research Article



The different activities of RNA G-quadruplex structures are controlled by flanking sequences

Alice J-L Zheng^{1,*}, Aikaterini Thermou^{1,4,*}, Pedro Guixens Gallardo^{6,7}, Laurence Malbert-Colas¹, Chrysoula Daskalogianni^{1,4}, Nathan Vaudiau¹, Petter Brohagen¹, Anton Granzhan^{6,7}, Marc Blondel⁸, Marie-Paule Teulade-Fichou^{6,7}, Rodrigo Prado Martins⁵, Robin Fahraeus^{1,2,3,4,*}

The role of G-quadruplex (G4) RNA structures is multifaceted and controversial. Here, we have used as a model the EBV-encoded EBNA1 and the Kaposi's sarcoma-associated herpesvirus (KSHV)-encoded LANA1 mRNAs. We have compared the G4s in these two messages in terms of nucleolin binding, nuclear mRNA retention, and mRNA translation inhibition and their effects on immune evasion. The G4s in the EBNA1 message are clustered in one repeat sequence and the G4 ligand PhenDH2 prevents all G4-associated activities. The RNA G4s in the LANA1 message take part in similar multiple mRNA functions but are spread throughout the message. The different G4 activities depend on flanking coding and non-coding sequences and, interestingly, can be separated individually. Together, the results illustrate the multifunctional, dynamic and context-dependent nature of G4 RNAs and highlight the possibility to develop ligands targeting specific RNA G4 functions. The data also suggest a common multifunctional repertoire of viral G4 RNA activities for immune evasion.

DOI 10.26508/lsa.202101232 | Received 11 September 2021 | Revised 8 November 2021 | Accepted 8 November 2021 | Published online 16 November 2021

Introduction

RNA-protein interactions are key regulators of the selective processing of RNAs which includes RNA splicing, localisation, translation, and stability, and are therefore involved in various cellular processes (Lewis et al, 2017; Duss et al, 2019; Rodgers & Woodson, 2019) including embryonic development (Beaudoin et al, 2018), neuronal activity (Lin et al, 2020), and oncogenesis (Ceci et al, 2021), to mention some. RNA secondary structures serve as binding platforms for the RNA-binding proteins, determine the assembly of ribonucleoprotein complexes and consequently affect gene expression (Lewis et al, 2017; Beltran et al, 2019; Duss et al, 2019; Rodgers &

Woodson, 2019; Sanchez de Groot et al, 2019). This is well illustrated by riboswitches and internal ribosome entry sites (IRESs) that control prokaryotic and eukaryotic/viral gene expression, respectively (Jackson et al, 2010; Serganov & Nudler, 2013). These regulatory elements are usually located in the 5' UTRs but RNA structures within the coding sequences can also mediate translation. However, little is yet known about the dynamics and regulation of RNA structures in vivo.

G-quadruplexes (G4) are secondary structures found in both RNA and DNA that are formed by the stacking of at least two G-quartets, which are planar arrangements of four guanines connected by Hoogsteen hydrogen bonds (Fay et al, 2017; Reznichenko et al, 2019). G4 conformation diversity depends on directionality (parallel, antiparallel, and hybrid), on the number of stacked G-quartets and on the length and sequence of the loops connecting the different strands, affecting stability and binding capacity to specific factors (Harris & Merrick, 2015; Tosoni et al, 2015; Fay et al, 2017). RNA G4 structures are linked to translation suppression when present in 5' UTR of the mRNA (Beaudoin & Perreault, 2010; Endoh & Sugimoto, 2016; Herdy et al, 2018) and to ribosome stalling when placed in the ORF (Endoh and Sugimoto, 2013, 2016; Endoh et al, 2013a, 2013b). Despite being associated with important functions, the actual formation, stability, and activity of RNA G4s in vivo is still poorly understood and controversial and most of the predicted G4-forming sequences in RNAs do not form stable G4 structures in eukaryotes (Guo & Bartel, 2016). On the other hand, G4 structures in vitro using shorter RNA sequences show high thermostability (Arora & Maiti, 2009).

RNA G4s are implied across various virus families, such as *Flaviviridae* (Fleming et al, 2016; Wang et al, 2016a; Jaubert et al, 2018), *Herpesviridae* (Murat et al, 2014), *Filoviridae* (Wang et al, 2016b), *Paramyxoviridae* (Majee et al, 2020), *Retroviridae* (Ruggiero et al, 2019), and *Coronaviridae* (Bezzi et al, 2021; Ji et al, 2021; Lavigne et al, 2021; Zhao et al, 2021), making them potential targets for drug development. In HIV-1, RNA G4s play an important role in the packaging of the virions (Lyonnais et al, 2003) and are part of the dimerisation process of the

¹Inserm UMR51131, Institut de Génétique Moléculaire, Université Paris 7, Hôpital St. Louis, Paris, France ²RECAMO, Masaryk Memorial Cancer Institute, Brno, Czech Republic ³Department of Medical Biosciences, Umeå University, Umeå, Sweden ⁴ICCVS, University of Gdańsk, Science, Gdańsk, Poland ⁵ISP, INRAE, Université de Tours, UMR1282, Tours, France ⁶CNRS UMR9187, INSERM U1196, Institut Curie, PSL Research University, Orsay, France ⁷CNRS UMR9187, INSERM U1196, Université Paris Sud, Université Paris-Saclay, Orsay, France ⁸Inserm UMR1078, Université de Bretagne Occidentale (UBO), Etablissement Français du Sang (EFS) Bretagne, CHRU Brest, Brest, France

Correspondence: robin.fahraeus@inserm.fr

*Alice J-L Zheng, Aikaterini Thermou, and Robin Fahraeus contributed equally to this work

two copies of the HIV-1 RNA genome, facilitating recombination during reverse transcription (Marquet et al, 1991, 1994; Piekna-Przybylska et al, 2013; Métifiot et al, 2014). In HCV, RNA G4s are implicated in the translation and replication of the genomic RNA (Wang et al, 2016a; Jaubert et al, 2018). EBNA1 and LANA1 are essential genome maintenance proteins of the EBV and the Kaposi's Sarcoma-associated herpesvirus (KSHV), respectively (Lan et al, 2004; Münz, 2015). Both are highly antigenic and to allow the viruses to escape immune surveillance the translation of their respective messenger RNAs is kept at the minimum to prevent the production of antigenic peptide substrates for the MHC-I pathway (Yin, 2003; Kwun et al, 2007). The synthesis of antigenic peptides for the MHC-I pathway is via non-canonical translation that is distinct from the canonical translation which generates full-length proteins, in line with the notion that full-length proteins are a poor substrate for antigenic peptides (Cardinaud et al, 2004; Starck et al, 2008; Apcher et al, 2011; Wei et al, 2019; Yewdell et al, 2019). A glycine-alanine repeat (GAR) domain in EBNA1 is encoded by a guanine-rich sequence that forms G4 structures and binds nucleolin (NCL), inhibiting canonical and non-canonical translation initiation *in cis* to support viral immune evasion but how G4 structures support these different activities is not known (Yin, 2003; Apcher et al, 2009; Lista et al, 2017b; Martins et al, 2019). More recently, it was shown that the mRNA translation stress caused by the GAR domain activates the E2F1 oncogene and stimulates c-myc and ribosomal biogenesis (Gnanasundram et al, 2017). Hence, by suppressing its own synthesis, EBNA1 evades the immune system and promotes cell proliferation. The LANA1 of the KSHV also uses a *cis*-acting mechanism to evade immune surveillance (Murat et al, 2014) and as for EBNA1, both RNA G4 structures and repeat domains of LANA1 plays an important role in translation regulation for both full-length proteins and antigenic peptides production (Kwun et al, 2007; Apcher et al, 2009; Murat et al, 2014; Dabral et al, 2019). However, two main differences are worth being noticed when comparing these two functionally similar mRNAs and proteins, whereas the G4s in EBNA1 are located in the GAR-encoding domain, the G4s of LANA1 are spread throughout the message (Murat et al, 2014; Dabral et al, 2019). In addition, the peptide repeats of LANA1 and EBNA1 are different, although both are encoded by mRNA sequences predicted to form G4s.

We have used the properties of the EBNA1 and the LANA1 mRNAs to study the multifunctional aspects of G4 RNA structures. Each message has different G4 structures encoding different peptide sequences but share functional similarities in suppressing mRNA translation *in cis*, interacting with NCL and preventing mRNA export. The functions of their respective G4s can be separated by alterations in the flanking sequences and together with the observation that RNA G4s can reform in cells, these results illustrate how multifunctional and dynamic G4 structures act together to support viral immune evasion strategies. This sheds new light on some of the controversies surrounding the stability of G4 RNA structures and illustrates the importance of the context in which they are placed.

Results

G4 ligands PhenDC3 and PhenDH2 bind LANA1 mRNA G4 structures *in vitro*

Using the QGRS mapper tool, and according to the work of Dabral et al (2019), the LANA1 message can form G4 structure at multiple

points on the mRNA, but mostly in the region encoding the Central Repeat Domain of LANA1 (Fig 1A). To confirm the binding of the G4 ligands we selected, namely, PhenDC3 and PhenDH2, on LANA1 RNA G4 structures, two quadruplex-prone G-rich short RNA sequences (LANA13 and LANA16 with QGRS scores of 21 and 18, respectively, Fig 1B) from the LANA1 Central Repeat (CR) domains, which have been already identified for their potential to form multiple quadruplex structures (Dabral et al, 2019), have been selected. First, to experimentally verify the RNA capacity to fold into quadruplex, their circular dichroism (CD) spectra were recorded (Fig 1B). In both cases, a positive peak at 265 nm together with a broad shoulder around 300 nm was observed. The CD spectra indicated the co-existence of different conformations in equilibrium, where the parallel conformation was the predominant. The RNA sequences exhibited a negative transition at 295 nm in the UV melting experiments which is characteristic of G4 unfolding (Fig 1C) and allowed determining T_m values of 65–68°C and 45–48°C for LANA13 and LANA16, respectively. In presence of PhenDC3 or PhenDH2, the CD intensity is slightly decreased indicating that interaction is occurring between the G4 conformation and the ligands (Fig 1B). Interestingly, when PhenDC3 was added before the annealing step, the shoulder at 300 nm was not observed (Fig S1), thereby suggesting that the ligand shifts the equilibrium towards the parallel conformation as has been recently reported for other G4 sequences (Aznauryan et al, 2021). Finally, we monitored the effects of PhenDH2 and PhenDC3 on the RNA sequences by CD melting. PhenDH2 and PhenDC3 induced on LANA16 a very strong stabilization ($\Delta T_m = 40^\circ\text{C}$ for PhenDH2 and $>47^\circ\text{C}$ with incomplete melting T_m $> 95^\circ\text{C}$ for PhenDC3) (Fig 1D). No complete melting of the G4 of LANA13 could be achieved in the presence of any ligand, indicating a very strong stabilization effect. Taken together, these data support the formation of G4 structures within the LANA1 CR domains and the possibility of targeting them with high affinity ligands such as PhenDC3 and PhenDH2.

G4 ligands affect LANA1 mRNA nuclear export, translation, and nucleolin binding

The glycine-alanine repeat (GAR) domain of the EBV-encoded EBNA1 mediates a unique *cis*-acting suppression of its own synthesis via G4 RNA structures to minimise the production of EBNA1-derived antigenic peptides for the MHC-1 pathway (Yin, 2003; Lista et al, 2017b). Similarly, the LANA1 of the KSHV also uses a *cis*-acting mechanism to evade the immune system via its CR domains (Kwun et al, 2007). Whereas the GAR-encoding message consists of GC-rich sequences encoding single alanine residues separated by two or three glycines, the CR domains of LANA1 are also GC rich, but each repeat region encodes for a different peptide motif (Kwun et al, 2007).

To know if LANA1 and EBNA1 are using similar molecular mechanisms involving RNA G4 structures to control mRNA translation, we first treated KSHV-carrying lymphoma B cells (BCP-1) and the EBV-carrying B-cell line B95.8 with 2 μM of the G4 ligands PhenDC3 or PhenDH2 for 24 h (Reznichenko et al, 2019). This resulted in an increase in the expression of respective proteins (Fig 2A). For EBNA1, this is linked to prevention of NCL binding to the RNA G4s by the G4 ligands (Lista et al, 2017a, 2017b). Thus, this result

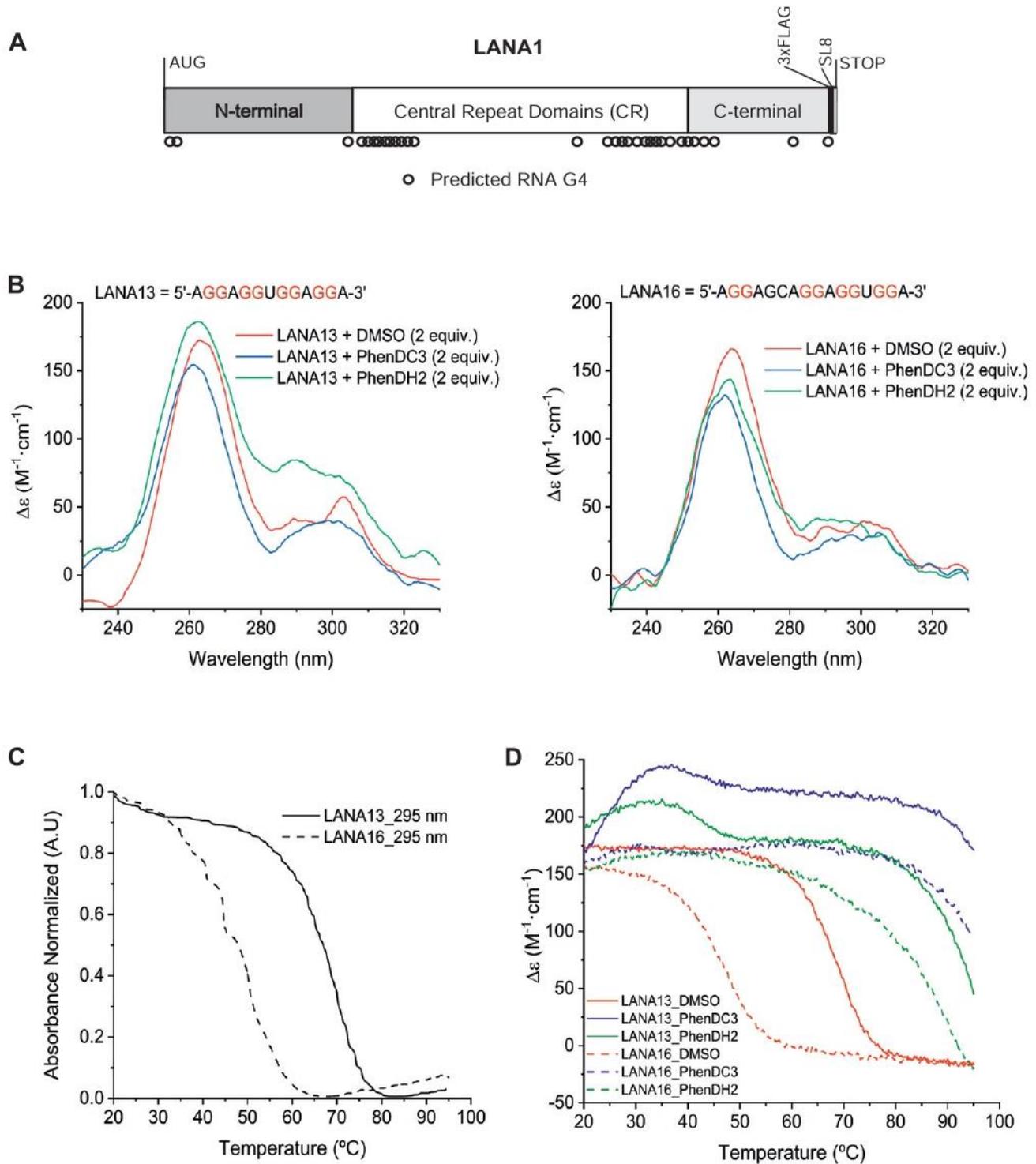


Figure 1. LANA1 central domains encoding mRNA form G4 structures in vitro.

(A) cDNA construct encoding LANA1 with a FLAG-tag and the SL8 MHC-1 antigenic peptide sequence in its C-terminus. Circles underneath represent the 32 predicted RNA G4 structures. (B) Circular dichroism spectra of RNA fragments ($5 \mu\text{M}$) of LANA1 CR domain in the presence of PhenDC3 (blue), PhenDH2 (green) or without ligand (red). Left, RNA-13mer (LANA13); right, RNA 16-mer (LANA16). (C) UV melting experiment at 295 nm of LANA13 (straight line) or LANA16 (dashed line) RNA sequences. (D) Circular dichroism melting experiment recorded at 262 nm of LANA13 (straight line) or LANA16 (dashed line) in the presence of PhenDC3 (blue), PhenDH2 (green), or without ligand (red).

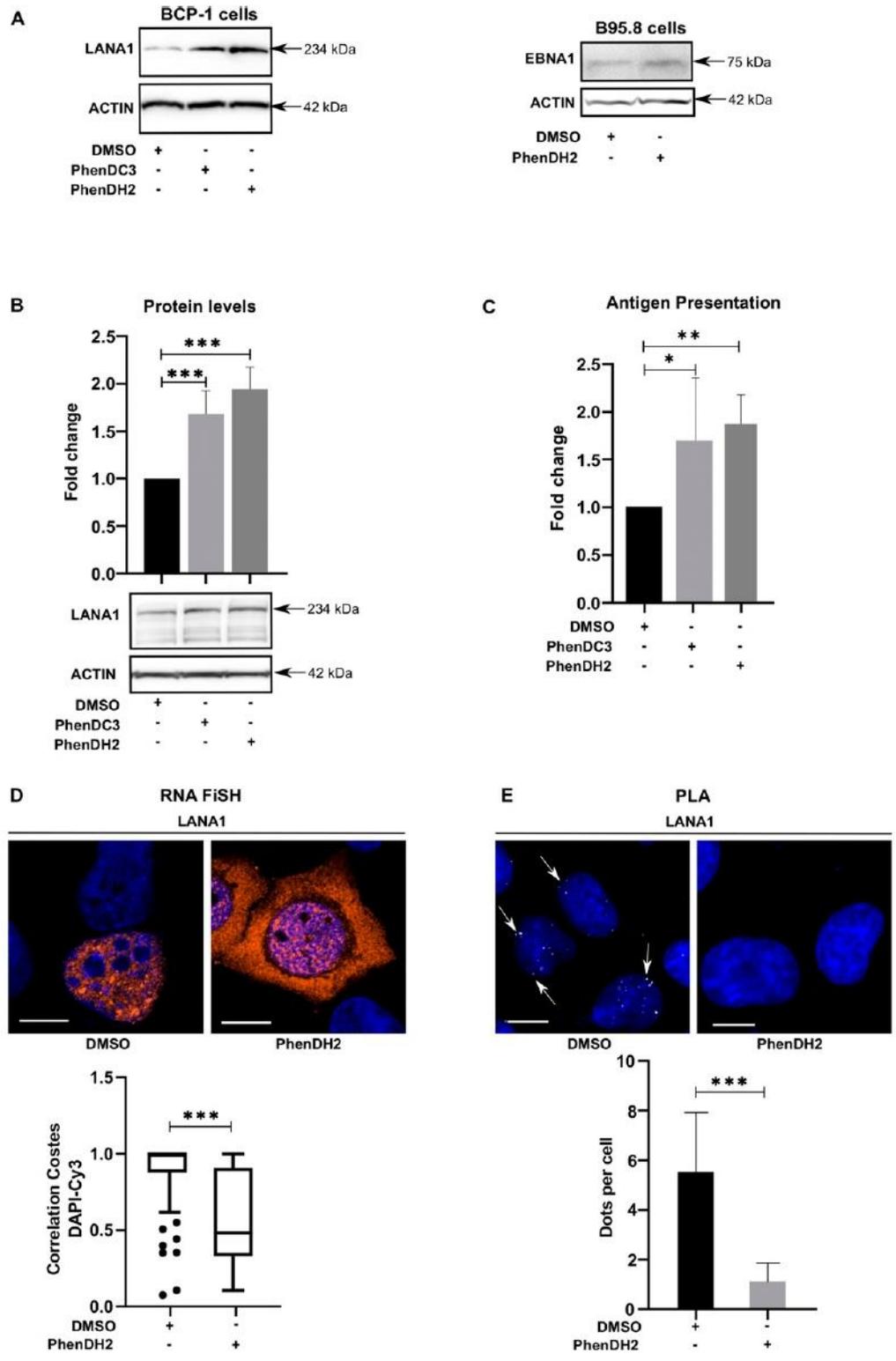


Figure 2. G4 RNA structures control the translation and processing of the LANA1 mRNA.

(A) Western blots (WB) show the expression of LANA1 in KSHV-infected BCP-1 cells (left) or EBNA1 in EBV-infected B95-8 cells (right) after treatment with 2 μ M of G4 ligands PhenDC3, PhenDH2, or DMSO for 24 h. WB show one of at least three similar experiments. (B) WB (below) shows the expression of LANA1 from cDNA construct presented in Fig 1A after indicated treatments. The graph above shows expression relative to actin from three independent experiments. (C) Relative amount of SL8 antigen peptide produced from the LANA1 cDNA presented in Fig 1A under similar conditions in H1299 cells expressing the murine MHC-I (Kb). The levels were estimated by measuring IL2 release from OT1 CD8⁺ T cells. (D) LANA1 mRNA localisation (RNA FISH) in H1299 cells transfected with the LANA1 construct presented in Fig 1A after PhenDH2 or DMSO treatment. The boxplot graph below shows Costes Correlation factor for cytoplasmic versus nuclear localisation. (E) The proximity ligation assay shows the interaction between the LANA1 mRNA and nucleolin (NCL) (white dots) after PhenDH2 treatment. The graph below shows the average interactions per cell from three independent experiments. Scale bar corresponds to 10 μ m.

shows that G4 structures in the LANA1 message, like in the EBNA1 message, are involved in translation suppression.

We next expressed a FLAG-tagged LANA1 construct (Fig 1A) in human H1299 carcinoma-derived cells and we observed a similar

increase in protein levels after PhenDC3 and PhenDH2 treatment (Fig 2B), showing that the G4-mediated translation control is not dependent on cell line or on the presence of the whole viral genome, as for EBNA1 (Lista et al, 2017b). Unlike PhenDC3 and

PhenDH2, another G4 ligand, the pyridostatin (PDS), is not able to interfere with the inhibition of LANA1 expression (Fig S2A), similarly to what was shown with EBNA1 (Lista et al, 2017b). We also inserted an eight amino acid-MHC-I antigenic peptide sequence (SL8) derived from the chicken ovalbumin gene (Ova) in the C-terminus of LANA1. The SL8 peptide is specifically recognised by CD8⁺ T cells (OT-1) when presented on the murine MHC-I (Kb) (Martins et al, 2019). After transient expression of LANA1 together with Kb in H1299 cells, we observed a twofold activation of CD8⁺ T cells after G4 ligand treatment, meaning that more SL8 antigenic peptide substrates were synthesised (Fig 2C), whereas the PDS is not able to interfere with the production of antigenic peptides for MCH-I pathway (Fig S2B). As PhenDH2 is a more potent G4 ligand when it comes to EBNA1 mRNA compared with PhenDC3 (Reznichenko et al, 2019), we focused our investigation on PhenDH2 treatment. By using RNA FISH, we observed that the LANA1 mRNA is located in the nucleus and that treatment with PhenDH2 promoted a cytoplasmic location (Fig 2D). Furthermore, using the proximity ligation assay (PLA) we showed that PhenDH2 interferes with the LANA1 mRNA-nucleolin (NCL) interaction in situ (Fig 2E).

These results show that the EBV and KSHV have evolved a common strategy to exploit *cis*-acting G4 mRNA structures in the EBNA1 and LANA1 messages to prevent the production of antigenic peptide substrates for the MHC-1 pathway.

Deleting the LANA1 CR domain overcomes nuclear retention and stimulates synthesis of antigenic peptide substrates but does not affect the synthesis of full-length proteins or the interaction with nucleolin

It has been suggested that the CR domains of the LANA1 account for mRNA translation control (Kwun et al, 2007) and we wanted to know if G4 activity in this region mediates the same functions as the EBNA1-encoded GAR sequence. When we deleted the CR domains (LANA1ΔCR), we observed no, or little, change in LANA1 protein expression, whereas we observed an average 60% increase in antigen presentation, indicating that different G4-forming domains in LANA1 control translation of antigenic peptide substrates and full-length proteins (Fig 3A–C). RNA FISH analysis detected most LANA1ΔCR messages in the cytoplasm (Fig 3D). The PLA showed NCL-LANA1ΔCR mRNA interactions in the nucleus, even though the number of interactions were reduced when compared with full-length LANA1 (Fig 3E). It should be pointed out that NCL is nuclear and preventing nuclear retention of the mRNAs could be sufficient to explain the differences in number of interactions.

Thus, G4 structures of the CR domain of the LANA1 message mediate nuclear retention and control of antigenic peptide synthesis but not full-length protein synthesis and NCL interaction.

G4s outside the CR domains control LANA1 protein expression and nucleolin binding

The above results indicate that different LANA1 mRNA processing activities can be mediated by individual G4 structures throughout the LANA1 message. To test if NCL binding and the control of full-length protein synthesis of the LANA1ΔCR construct is mediated by G4s predicted outside the CR domain, we treated cells expressing

LANA1ΔCR with PhenDH2. This resulted in an average 50% increase in protein levels (Fig 4A) and antigen presentation (Fig 4B). The treatment did not significantly affect LANA1ΔCR mRNA localisation (Fig 4C) but prevented furthermore the interaction with NCL (Fig 4D).

This confirms that the *cis*-mediated regulation of LANA1 mRNA translation is more complex as compared with the EBNA1 and involves G4s from different regions of the coding sequence and not just from the CR domains.

The position of the G4s within the coding sequence controls its activity

It has been suggested that the interaction between NCL and G4s of EBNA1 mRNA is sufficient for suppressing synthesis of antigenic peptides and full-length proteins (Lista et al, 2017b), and it was surprising to see that deletion of the CR differentiates these two activities in the LANA1 mRNA. Therefore, we were interested to know if the G4 activities of the EBNA1 message can also be separated. To test this, we used constructs in which the GAR was fused to the 5' or the 3' end of the chicken ovalbumin (Ova) ORF, resulting in fusion proteins with the GAR in the N terminus or the C terminus (GAR-Ova and Ova-GAR, respectively) (Fig 5A) (Yin, 2003). Both messages retained their nuclear localisation but the expression of the GAR-Ova fusion protein was considerable less, as compared with the Ova-GAR, whereas the RNA levels followed an opposite trend (Figs 5B and C and S3). When we isolated these RNAs from transiently transfected cells and compared their binding capacity to recombinant NCL, we observed that the GAR-Ova mRNA had a higher affinity to NCL compared with the Ova-GAR mRNA (Fig 5D).

Hence, the location of the GAR G4s within the coding sequence of the message determines NCL binding and mRNA translation inhibitory activity but not nuclear retention. Of note, these results support the notion that NCL plays no role in G4-dependent mRNA export.

Flanking UTR sequences affect LANA1 and EBNA1 G4s differently

The GAR-Ova and the Ova-GAR reporter constructs showed that the localisation of the G4-containing translational inhibitory sequence within the coding sequence is important for some, but not all, G4-related activities. We next wanted to know if changes in the non-coding sequences can also affect G4 functions. We introduced RNA structures in the form of the *c-myc* and HCV IRESs in the 5' UTR of the GAR-Ova construct. We could confirm that the *c-myc*-GAR-Ova construct and *c-myc*-Ova are equally efficiently translated (Apcher et al, 2009). The fusion of the HCV IRES (HCV-GAR-Ova) instead further suppressed expression of GAR-Ova as well as Ova alone (HCV-Ova) (Fig 6A) (Martins et al, 2019). When the two IRESs were individually fused to the LANA1 message, they both stimulated synthesis of full-length proteins (Fig 6B). Interestingly, although only the HCV IRES suppressed synthesis of GAR-Ova and Ova proteins, the presence of either IRES in the 5' UTRs resulted in an increase in the synthesis of antigenic peptide substrates from GAR-Ova and LANA1 messages (Fig 6C and D). Hence, although the HCV IRES suppressed expression of full-length GAR-Ova and Ova, it stimulated synthesis of antigenic peptide substrates from both GAR-Ova and LANA1 messages. Both IRESs also prevented NCL

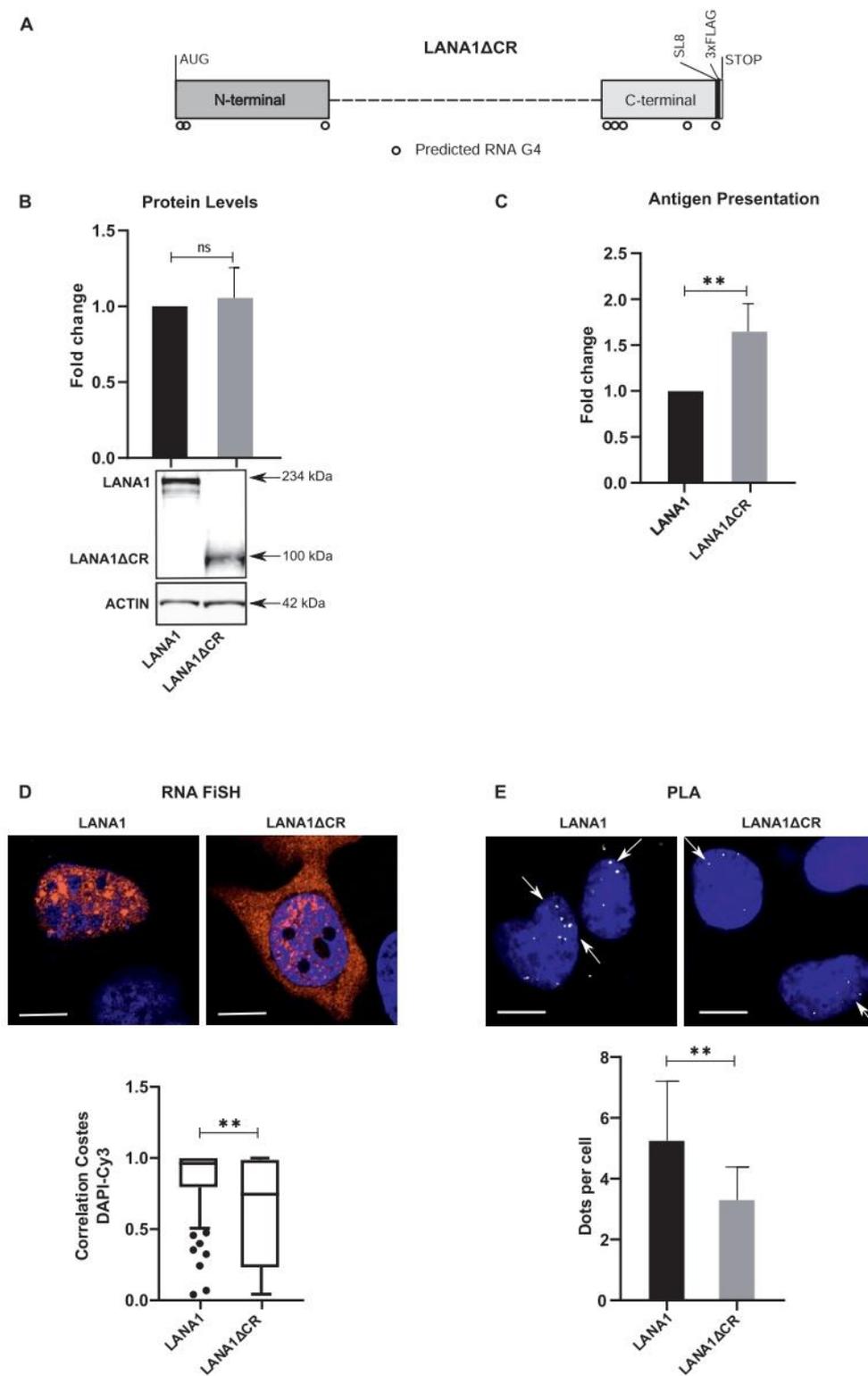


Figure 3. Central repeat (CR) domains of the *LANA1* mRNA control antigen presentation, nuclear retention but not canonical translation for production of full-length protein or nucleolin interaction.

(A) cDNA construct encoding *LANA1* without CR domains (*LANA1*ΔCR) and with a FLAG Tag and the SL8 in the C terminus. Circles underneath the rectangles represent the eight predicted RNA G4 structures. **(B)** WB of H1299 cells transfected with the *LANA1* or *LANA1*ΔCR constructs. Graph shows relative protein levels, adjusted to corresponding actin control. WB shows one of at least three similar experiments. **(C)** Relative antigen presentation levels estimated by measuring IL2 release by OT1 cells after incubation with H1299 cells expressing murine MHC-I (Kb). **(D)** *LANA1* mRNA localisation (RNA FISH) with boxplot graph below representing three independent experiments. **(E)** The interaction with NCL and the *LANA1* mRNA (PLA NCL-*LANA1* mRNA) in H1299 cells transfected with indicated cDNA constructs. The graph below shows the average of three independent experiments. Scale bar corresponds to 10 μm.

binding to GAR-encoding and *LANA1* mRNAs, but only the *HCV* IRES promoted mRNA nuclear export (Figs 6E and S4–S6). Fusing the *c-myc* and the *HCV* structures to the 3' UTR of *LANA1* had little effects on G4 activities (Fig S7A and B).

This demonstrates that the different functions of the G4s within the messages can be separated by altering the 5' UTR. These data further underline that the synthesis of full-length proteins and antigenic peptides are derived from two independent mRNA translation events.

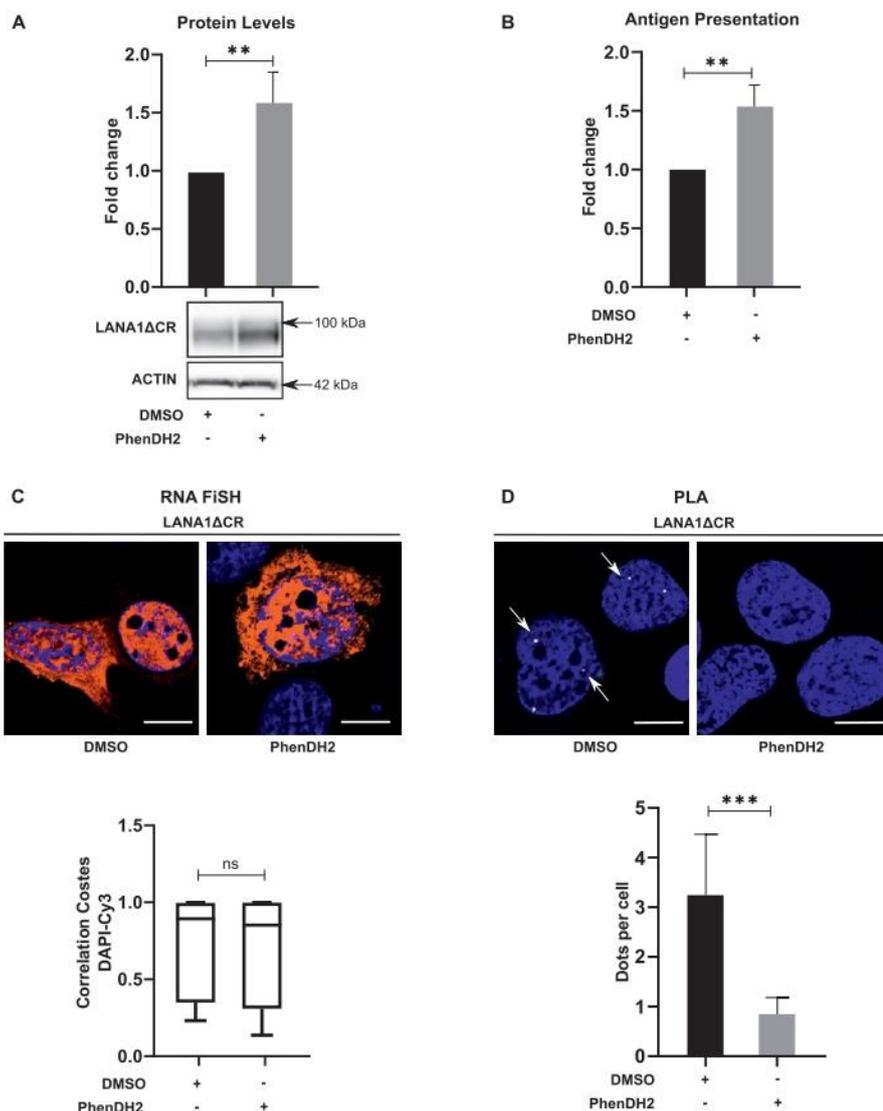


Figure 4. G4 structures outside the central repeat (CR) of *LANA1* control NCL binding and canonical translation.

(A) WB from H1299 cells expressing *LANA1ΔCR* treated with the G4 ligands PhendH2 or DMSO. Graph shows relative protein levels adjusted with the corresponding actin levels. WB shows one of at least three similar experiments. (B) Relative antigen presentation estimated by measuring IL2 release by OT1 cells after incubation with H1299 cells expressing murine MHC-I (Kb) and *LANA1ΔCR* and treated with PhendH2 or DMSO. (C) *LANA1ΔCR* mRNA localisation (RNA FISH). (D) Proximity ligation assay showing the RNA-NCL interaction (proximity ligation assay NCL-*LANA1* mRNA) in H1299 cells transfected with *LANA1ΔCR* construct after PhendH2 or DMSO treatment. The graphs show the average of three independent experiments. Scale bar represents 10 μ m.

Discussion

We have compared four different activities related to G4 structures within the coding sequences of the KSHV- and EBV-encoded *LANA1* and *EBNA1*, respectively. These proteins play similar key roles in maintaining viral genome integrity and both are expressed in all virus-infected cells, including cancerous ones, making their respective expression interesting therapeutic targets for virus-associated diseases (Ballestas, 1999; Ballestas & Kaye, 2001; Ciuffo et al, 2001; Kwun et al, 2007; Münz, 2015). Both viruses need to ensure that the production of peptide substrates for the MHC-I pathway is kept at a minimum to evade the immune system (Kwun et al, 2007; Murat et al, 2014). To achieve this, they have evolved *cis*-acting mechanisms to suppress translation of the *LANA1* and *EBNA1* messages. Even though the peptide sequences mediating this effect are not similar (Kwun et al, 2007), the concept to control mRNA processing and translation to achieve immune evasion is. This suggests that controlling their own rate of translation might be a

common strategy whereby, at least latent, viral messages evade the immune system. In agreement with this hypothesis is the observation that viruses causing persistent infections in *Metazoa* hosts are significantly enriched with putative G-quadruplex forming sequences (PQS), whereas viruses causing acute infections are significantly depleted in PQS (Bohálová et al, 2021).

Treatment with the G4 ligand PhendH2 resulted in an increase in protein expression, antigen presentation, nuclear export, and loss of NCL binding in the context of *EBNA1* and *LANA1*, indicating that these four events are indeed all G4 dependent. The G4s of *EBNA1* mRNA are mostly clustered within the gly-ala repeat GAr-encoding sequence (Murat et al, 2014), whereas the G4 activities of *LANA1* are spread throughout the message. The G4s in the CRs of *LANA1* are determinant for nuclear mRNA retention and suppression of antigenic peptide substrates, whereas G4s outside the CR domains are important for canonical translation control and nucleolin interaction. Both CR and non-CR-related G4s play roles in antigen production. The fact that the CR domains play an important role in

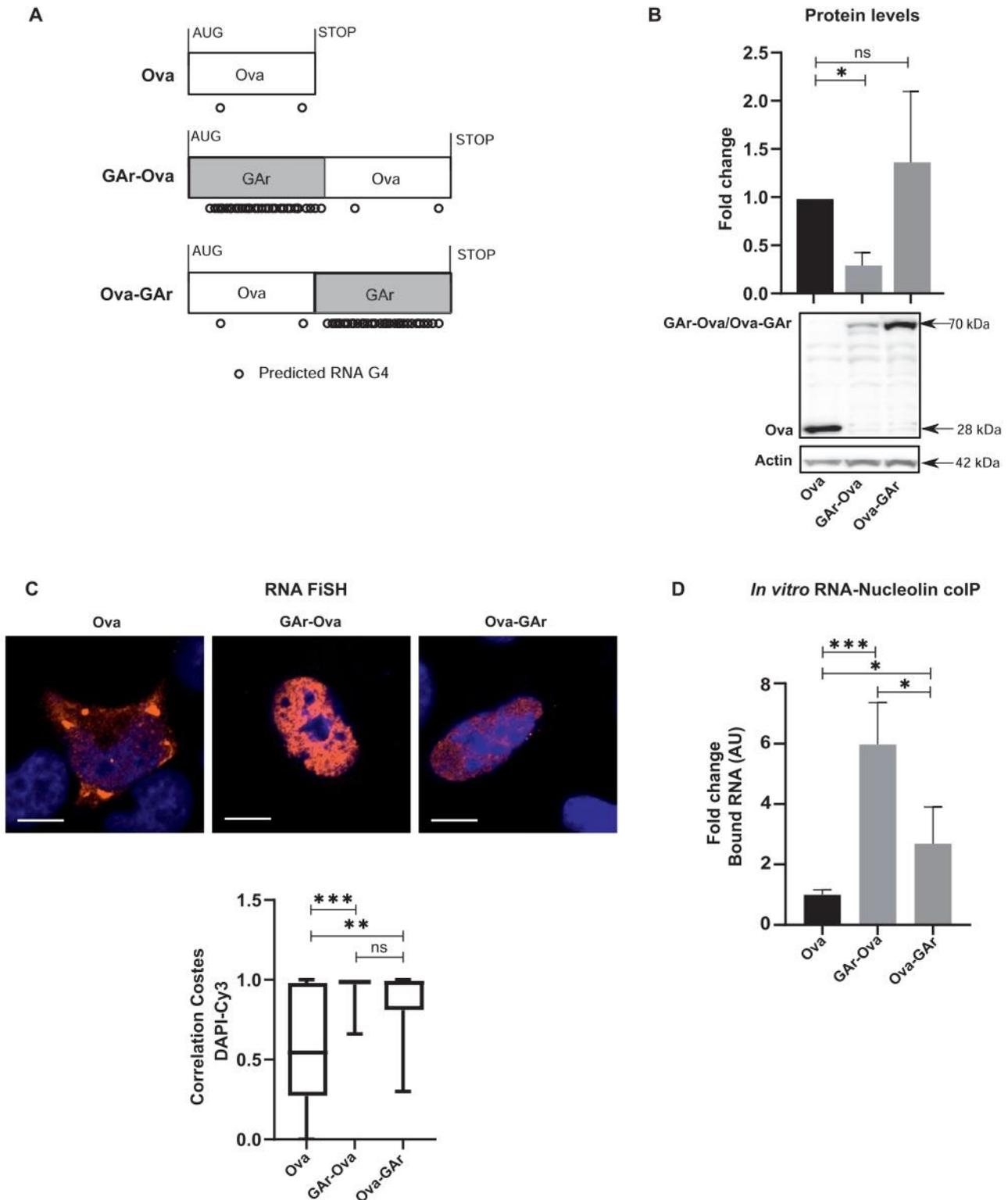


Figure 5. The position of G4s within the coding sequence affects their respective activities.

(A) Illustrations of the cDNA constructs to which GAR was fused to the N terminus (GAR-Ova) or C terminus (Ova-GAR) of chicken ovalbumin (Ova). Circles underneath represent the predicted RNA G4 structures for each construct (2 for Ova, 34 for GAR-Ova, and 33 for Ova-GAR). (B) Western blot of H1299 cells after transfection with indicated cDNA constructs. Graph shows relative expression adjusted with the corresponding actin levels. WB shows one of at least three similar experiments. (C) Ova-encoding mRNA localisation (RNA FISH) in H1299 cells transfected with indicated constructs. Scale bar represents 10 μ m. (D) Relative binding of Ova fusion mRNAs to recombinant NCL protein measured by *in vitro* RNA co-IP assay. mRNAs were extracted from H1299 cells transfected with indicated constructs. The graphs show the average of three independent experiments.

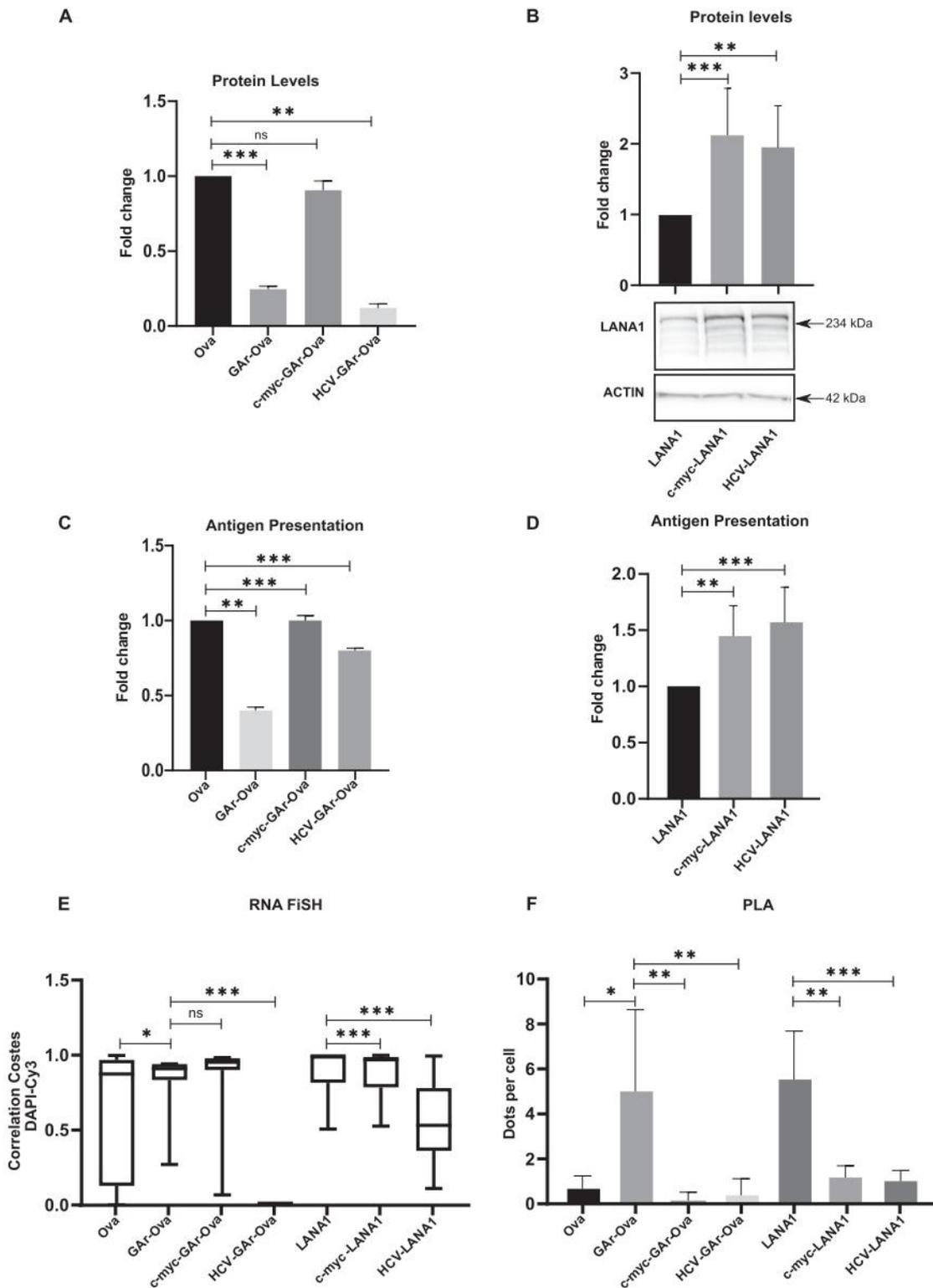


Figure 6. Altering the UTRs differentiates the activities of the GAR-encoded and LANA1 G4s.

Structured RNA sequences from the 5' of the *c-myc* and the *HCV* were inserted in the 5' of indicated constructs. **(A)** WB of H1299 cell lysates after transfection with indicated Ova fusion constructs. GAPDH was used as loading control. **(B)** WB of H1299 cell lysates after transfection with indicated LANA1 fusion constructs. **(C)** Antigen presentation levels estimated by measuring IL2 release by OT1 cells after incubation with H1299 cells transfected with GAR fusion constructs together with cDNA encoding murine MHC-I (Kb). **(D)** Antigen presentation levels derived from LANA1 fusion constructs. **(E)** Graph from RNA FISH observations shows the median subcellular colocalisation of indicated mRNAs with the nucleus. **(F)** Graph shows proximity ligation assay data about the interaction between NCL and indicated mRNAs from transfected constructs in H1299 cells. The graphs show the values of three independent experiments.

controlling the synthesis of antigenic peptides is in line with previous observations (Kwun et al, 2007). However, the results presented here do not support the CR domains being the sole responsible for controlling synthesis of antigenic peptides and full-length proteins from the *LANA1* mRNA. Nevertheless, we did not analyse the role of the individual *LANA1* repeat domains and together with the fact that flanking sequences play a role in determining G4 activities this can help in explaining this discrepancy.

An important conclusion from this study is that different G4 activities within an mRNA can be differentiated by altering flanking sequences, illustrating the dynamic nature of G4 RNA structures. For example, moving the GAr sequence from the 5' towards the 3' end of the coding sequence reduces its translation inhibitory capacity and NCL binding without affecting nuclear retention. Introducing the structured RNAs of the *HCV* or the *c-myc* IRESs affect *LANA1* and GAr mRNA translation and localisation differently. The *HCV* IRES further suppressed synthesis of GAr-carrying mRNAs, whereas it instead stimulated *LANA1* expression and released the nuclear retention of both mRNAs. The *c-myc* IRES, on the other hand, does not affect RNA localisation but induces translation of both messages (Martins et al, 2019). Also G4-mediated regulation of canonical translation producing full-length proteins and non-canonical translation linked to antigenic peptides synthesis could be separated by introducing the *HCV* IRES in the 5' of the GAr.

Previous works show that RNA G4 structures can fold, unfold and refold *in cellulo* (Biffi et al, 2014; Chen et al, 2018; Yang et al, 2018). The *EBNA1* message does not contain introns but placing the GAr in the context of a gene prevents NCL access *in cellulo*. However, placing the GAr in this context does not affect the G4 structure and RNAs isolated from cells interact with recombinant NCL (Martins et al, 2019). The interaction between the GAr derived from a spliced construct and NCL was observed in the cytoplasm when the nuclear localisation signal of NCL was removed (Martins et al, 2019). As the spliced message is subjected to the pioneer round of translation, the GAr G4s would have had to refold to allow NCL interaction in the cytoplasm. These observations indicate that the function of G4s in the regulation of mRNA processing is strongly influenced by events taking place during the mRNA maturation process and that they can act both in the nucleus as well as in the cytoplasm. In this context, G4s multifunctionality might be affected by any change impacting the nature of the ribonucleoparticles. In fact, several RNA binding proteins are reported to be involved in the folding and unfolding of G4s RNA for a specific regulation event (Wolfe et al, 2014; Herviou et al, 2020; Vannutelli et al, 2020).

The role, and actual *in vivo* existence, of G4 RNA structures is controversial (Guo & Bartel, 2016; Waldron et al, 2018; Dumas et al, 2021). Our data support a model in which G4 structures form dynamic multifunctional units, showing similarity to the concept of intrinsic disordered proteins. The allosteric conformation of intrinsic disordered proteins is regulated by post-translational modifications and/or ligand binding that alter their interactomes and their functions (Habchi et al, 2014; Borkosky et al, 2017). The RNA G4s could follow a similar scheme, in which ligand binding, or RNA modifications, could affect the structure of the G4s and alter their activity. In the examples described here, the activities of the *LANA1* and *EBNA1* G4s can be disrupted by chemical ligands, by moving their position or by changing flanking sequences.

It is conceivable that the dynamic multifunctional roles of these herpes viral RNA G4s are not unique but reflect a more common nature of regulated RNA structures within coding sequences. In the case of the *p53* mRNA, a structured region in the 5' of the coding sequence is folded during genotoxic stress so that it binds to MDM2 and stimulates *p53* synthesis. The RNA folding is prevented by a single cancer-derived synonymous mutation, illustrating how single nucleotides can affect the folding of an mRNA and its interaction with regulatory proteins (Candeias et al, 2008; Naski et al, 2009). In line with this, modifications of mRNAs can result in structural changes and, consequently, affect their functions and interactions with cellular factors (Lewis et al, 2017). For example, the binding of hnRNPC to cryptic U-tract sequences can be made available via *N6*-methylation on adenosine residues (Liu et al, 2015).

Treatment with chemical G4 ligands PhenDC3 or PhenDH2 disrupts GAr and *LANA1* G4s functions. However, the PDS G4 ligand stimulates the inhibition of *EBNA1* synthesis *in vitro* (Murat et al, 2014) but not *in vivo* and it does not affect NCL binding (Lista et al, 2017b; Martins et al, 2018). Such differences are also observed for *LANA1*. Dabral et al (2019) report that the G4-stabilising ligand TMPyP4 reduces the translation of *LANA1* mRNA, but we show here that PhenDC3 or PhenDH2 treatments increase *LANA1* protein expression and the presentation of *LANA1*-derived antigens. It is, thus, likely that different G4 ligands affect RNA structures differently by being able to either stabilise or destabilise them and/or to interfere with their ability to represent binding platforms for various RNA-binding factors. Considering the dynamic function of the herpes viral G4s, this suggests that G4 structures and activities can be specifically targeted for therapeutic intervention and that screening of G4-targeting compounds should be carried out in an *in vivo* setting with an intact message to obtain a specific activity.

Materials and Methods

Annealing of RNA sequences for CD spectroscopy

The G-rich RNA oligonucleotide (ON) sequences *LANA13* (5'-AGGAGGUGGAGGA), genomic loci 2563, G-score 21, and *LANA16* (5'-AGGAGCAGGAGGUGGA, genomic loci 2635, G-score 18) (Dabral et al, 2019) were purchased in Eurogentec as lyophilized solid after HPLC separation. They were reconstituted in DEPC water, and their concentration calculated by measuring the UV absorption at 260 nm on a Hitachi 2900 spectrometer. RNA ON buffered solutions (5 μ M ON; Li Cacodylate 10 mM, KCl 100 mM) were annealed in a thermo-block by heating 2 min at 95°C and let it cool until reaching room temperature.

CD footprint

The CD of the previously annealed ON sequences (*LANA13* and *LANA16*) were measured alone or in the presence of 2 M equivalents of PhenDC3 or PhenDH2 in a Jasco J-1500 CD spectrometer using a quartz cuvette of 0.5-cm path length. The parameters were four accumulations, CD scale 200 mdeg/0.1 dOD, scanning speed 100 nm/min. The spectra shown were mathematically processed using

the Savitzky–Golay method with 20 points of window (del Villar-Guerra et al, 2017).

CD melting

The CD melting of the ON sequences containing (LANA13 and LANA16), alone or in the presence of ligand, were monitored using a Jasco J-1500 CD spectrometer equipped with external temperature water-circuit control. The CD was recorded at 262 and 295 nm every 0.2°C change, when heating the samples from 20°C to 95°C at a speed of 0.2°C/min.

UV melting

The UV melting of the ON sequences containing (LANA13 and LANA16) in absence of ligand were monitored using a Spectro UV Cary-300 equipped with external temperature water-circuit control. The UV was recorded at 295 nm when heating the samples from 20°C to 95°C at a speed of 0.2°C/min.

Expression plasmids

Putative G4 RNA sequences were identified in the LANA1-encoding and Ova-encoding mRNAs using the QGRS Mapper: <https://bioinformatics.ramapo.edu/QGRS/index.php> (Kikin et al, 2006).

All expression plasmids were constructed using the pcDNA3 vector and amplified in the *Escherichia coli* DH5 α strain. The constructs named Ova, GAR-Ova and Ova-GAR encode the following proteins, respectively: Ovalbumin without the first 50 amino acids, the same protein with the GAR domain from EBNA1 fused to its N terminus or C terminus (Yin, 2003).

The pA3F-LANA1 plasmid carrying the Flag-tagged ORF73 of KSHV in pcDNA3 vector, which encodes wild-type LANA1 protein, was a kind gift from Pr. Blondel (University of Brest). LANA1-SL8 construct was created by inserting the SL8 (SIINFEKL peptide from chicken ovalbumin) coding sequence after the C-terminal domain of LANA1. LANA1 Δ CR-SL8 construct was generated using the QuikChange Site-Directed Mutagenesis System (Stratagene) and the construct LANA1-SL8 as template.

The plasmids c-myc_IRES_Ova, c-myc_IRES_GAR-Ova, HCV_IRES_Ova, and HCV_IRES_GAR-Ova encode Ovalbumin or a fusion of the GAR domain to the N terminus of ovalbumin, with either the c-myc IRES or the HCV IRES in their 5' UTR (Apcher et al, 2009). HCV and c-myc IRES were amplified from these constructions and fused to the 5' UTR or 3' UTR of LANA1-SL8 plasmid to produce the 5' c-myc-IRES-LANA1, 3' c-myc-IRES-LANA1, 5' HCV-IRES-LANA1, and 3' HCV-IRES-LANA1 constructs. The list of primers, cloning method, and resulting plasmids is provided in Table S1.

Cell culture

The human lung carcinoma cell line H1299 and the EBV-infected B cell line B95.8 were cultured in RPMI 1640 medium supplemented with 10% FBS, 2 mM L-glutamine, 100 U/ml penicillin and 100 μ g/ml streptomycin. KSHV-infected BCP-1 cell line was cultured in RPMI medium containing 10% FBS, 4 mM L-glutamine, 100 U/ml penicillin,

100 μ g/ml streptomycin, and 5 mM Hepes. All cells were cultured at 37°C in a humidified 5% CO₂ incubator.

Transfection and drug treatment

H1299 cells were seeded in six-well plates (7×10^4 cells/well) and transient transfections were carried out 20–24 h later using the Genejuice reagent (Merck Bioscience) according to the manufacturer's protocol. For cell treatments with PhenDC3, PhenDH2, and PDS (Lista et al, 2017b; Reznichenko et al, 2019), cells were incubated with 2 μ M of drug 24 h after transfection. Drug stock solutions were prepared in DMSO (Euromedex).

Western Blotting

40 h after transfection, cells were resuspended in lysis buffer (20 mM Hepes KOH, pH 7.5, 50 mM β -glycero-phosphate, 1 mM EDTA, pH 8, 1 mM EGTA, pH 8, 0.5 mM Na₃VO₄, 100 mM KCl, 10% Glycerol, and 1% Triton X-100) supplemented with complete protease inhibitor cocktail (Roche) and protein concentration was measured using a Bradford assay. Samples were separated by electrophoresis in Bolt Bis-Tris Plus gels 4–12% (Invitrogen/Thermo Fisher Scientific), and then transferred onto nitrocellulose membrane (BioTrace NT, Pall Life Science). Membranes were blocked for 1 h in 5% non-fat dry milk in Tris-buffered saline, pH 7.6, containing 0.1% Tween-20. Proteins were probed by overnight incubation at 4°C with the following antibodies: anti-Ovalbumin whole serum (C6534; Sigma-Aldrich), anti- β -actin mouse monoclonal antibody (clone AC-15; Sigma-Aldrich), anti-NCL polyclonal rabbit antibody (ref 22758; Abcam), anti-Flag mouse monoclonal antibody (F3165; Sigma-Aldrich), anti-EBNA1 mouse monoclonal antibody OT1X (Cyto-Barr BV), and anti-LANA1 (KSHV ORF73) mouse monoclonal antibody (D325-3; MBL International). Membranes were then incubated with appropriate HRP-conjugated secondary antibodies (Dako) and detection of immune complexes was performed using Pierce ECL, WestDura, or West Femto (Thermo Fisher Scientific) and myECL Imager (Thermo Fisher Scientific). Protein bands were quantified by densitometry analysis in Fiji/ImageJ (Schneider et al, 2012), using actin bands for normalisation.

Predicted molecular weight for the full-length LANA1 is 135 kD; however, as it was previously described, the LANA1 full-length protein migrate at 234 kD in SDS-PAGE (Rainbow et al, 1997). In addition, the LANA1 Δ CR product is predicted to have a molecular weight of 61 kD, but migration with SDS-PAGE resulted in an observed band of 100 kD.

RNA FISH

H1299 cells were seeded in 24-well plates (2×10^4 cells/well) and transient transfections were carried out 20–24 h later using the Genejuice reagent (Merck Bioscience) according to the manufacturer's protocol. 24 h after transfection or after drug treatment, the cells were briefly washed with ice-cold PBS, fixed with 4% PFA during 20 min at room temperature, and washed again with PBS. Cells were then incubated in 70% ethanol for 4–24 h at 4°C, after intermediate dehydrating steps using 30% and 50% ethanol. For rehydration, the cells were incubated in 50% and 30% ethanol and

further washed with PBS. Subsequently, cells were permeabilised with PBS 0.4% Triton 0.05% CHAPS for 5 min at room temperature. After three PBS washes, samples were pre-treated with two subsequent incubations in FISH wash buffer (10% formamide and 2X SSC in ddH₂O) for 10 min and then in FISH hybridisation buffer (10% formamide, 2X SSC, 2 mg/ml BSA, 0.2 mg/ml *E. coli* tRNA, 0.2 mg/ml sheared salmon sperm DNA) for 30 min at room temperature. Coverslips were then incubated overnight in a wet chamber at 37°C in FISH hybridisation buffer supplemented with 10% dextran sulphate and 100 nM of FISH Stellaris probes targeting *Ovalbumin* or *LANA1* 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. Samples were mounted using Dako mounting solution (Dako) and observed with a LSM 800 (Zeiss) confocal laser microscope. Images were obtained using Zen software (Zeiss) and the Costes colocalisation factor between nuclei (DAPI channel) and targeted mRNAs (Cy3 channel), called Correlation_Costes_DAPI_Cy3, were obtained using the software CellProfiler (McQuin et al, 2018), and later used for statistical analysis.

Proximity ligation assay (RNA-protein)

Cells were cultured, fixed, permeabilised, and submitted to a pre-hybridisation step as described above. For target-mRNAs hybridisations, DNA probes were denatured for 5 min at 80°C and diluted in hybridisation buffer at a concentration of 50 ng of probes per sample. The probes 5'-GCAGCAGACTACACCTCCACACTCACC-biotin 3' and 5'-CTGCTCATTGATTTCGCATGTGCTGCATGGACAGCTTGAAAAA-digoxigenin 3' were designed to target *LANA1* and *Ovalbumin* mRNAs, respectively. Samples were overnight incubated with the denatured DNA probes in a wet chamber at 37°C. Samples were then washed in hybridisation buffer, FISH wash buffer, and PBS and saturated with PBS 3% BSA 0.1% saponin for 1 h at room temperature. Primary antibodies were diluted in the saturation solution, kept for 30 min at room temperature and then overnight incubated with tested samples at 4°C in a wet chamber. The following antibodies were used to perform the PLA: anti-Digoxigenin mouse monoclonal antibody (clone DI-22, D8156; Sigma-Aldrich), anti-Biotin mouse monoclonal (clone BN-34, B7653; Sigma-Aldrich), and anti-NCL rabbit polyclonal antibody (ab22758; Abcam). Afterwards, PLA was performed using Duolink PLA kit and plus/minus probes (Sigma-Aldrich) following the manufacturer's instructions. Coverslips were finally mounted using Dako mounting buffer (Dako) after nuclear staining with DAPI and observed using an LSM 800 confocal laser microscope. Images were obtained using the Zen software and analysed using Fiji/ImageJ.

Antigen presentation

Naïve OVA257-264-specific CD8⁺ T-cells were isolated from the peripheral and mesenteric lymph nodes of OT-I-mice using the CD8⁺-isolation kit (Miltenyi Biotec). Subsequently, 4×10^5 CD8⁺ T-cells were co-cultured with 10^5 H1299 cells previously transfected with the indicated constructs and a mouse Kb expression vector (a kind gift from C Watts, University of Dundee) and treated or not with the tested drugs. Cells were cultured in RPMI 1640 (Thermo Fisher Scientific) supplemented with 2 mM L-glutamine, 100 U/ml streptomycin and 100 U/ml penicillin (Invitrogen), 10% FBS, 0.05 mM

2-mercaptoethanol, and 5 mM Hepes (Sigma-Aldrich) for 3 d at 37°C with 5% CO₂. Supernatants were collected from the co-cultures and IL-2 levels were measured by ELISA using ELISA MAX Standard kit (BioLegend) following the manufacturer's protocol. Signals were measured using a FLUOstar Optima (BMG Labtech) and data were analysed using the software Optima Control v2.20R2.

RNA extraction, RT-qPCR, and RNA co-IP assay

Transfected cells were washed in cold PBS and total RNA extraction was performed using RNeasy 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)₁₂₋₁₈ primer (Life technologies). qPCR was performed using the StepOne real-time PCR system (Applied Biosystems) with Perfecta SYBR Green FastMix (Quanta Biosciences) and the following primers (custom primers by Invitrogen): Actin Forward 5'-TCACCCACACTGTGCCATCTACGA-3', Actin Reverse 5'-TGAGGTAGT-CAGTCAGTCCCG-3', *LANA1* Forward: 5'-ATCTCCTGCATTGCCACC-CACGC-3', *LANA1* Reverse: 5'-TCTCAGGCTACGCAGGGTAGACG-3', *Ova* Forward 5'-GCAAACCTGTGCAGATGATG-3', *Ova* Reverse 5'-CTGCTCAAGGCCTGAGACTT-3'. In vitro RNA co-IP was carried out as described elsewhere (Candeias et al, 2008). Briefly, 1 µg of total RNA extracted from cells was co-incubated under agitation with 100 ng of recombinant NCL (provided by Dr. Teulade-Fichou) in the binding buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 0.02 mg/ml yeast tRNA, and 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-Aldrich) with an anti-NCL rabbit polyclonal antibodies (ab22758; Abcam) according to standard conditions and purified using the TRIzol (Life Technologies). Precipitated RNAs were then analysed by RT-qPCR.

Statistical analysis

Data were analysed by unpaired Mann-Whitney's test or *t* test on GraphPad Prism 9. For Western blot, PLA dot counts, and qPCR analyses, represented data are the mean and the SD of a minimum of three independent experiments. For RNA co-IP analyses, represented data are the mean and the standard error of a minimum of three independent experiments. For FISH analysis, data represented on graphs are the median value, 25% and 75% quartiles delimiting the interquartile range, the maximum and minimum values and the outliers. In all cases, $P > 0.1234$ (ns), $P < 0.0332$ (*), $P < 0.0021$ (**), and $P < 0.0002$ (***).

Supplementary Information

Supplementary Information is available at <https://doi.org/10.26508/lsa.202101232>.

Acknowledgements

We thank the staff of the Plateforme Technologique de l'Institut de Recherche Saint-Louis for their help with microscopy. FUNDING: AJ-L Zheng

was funded by Université de Paris and the Fondation pour la Recherche Médicale (grant number FDT202001010912). This work was partially supported by European Regional Development Fund (ENOC, CZ.02.1.01/0.0/0.0/16_019/0000868), MH CZ-DRO (MMCI, 00209805), Cancerforskningsfonden Norr, Cancerfonden (160598), Vetenskapsradet, and by the 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.

Author Contributions

AJ-L Zheng: investigation.
 A Thermou: investigation.
 P Guixens-Gallardo: investigation.
 L Malbert-Colas: investigation.
 C Daskalogianni: investigation.
 N Vaudiau: investigation.
 P Brohagen: investigation.
 A Granzhan: investigation.
 M Blondel: supervision.
 M-P Teulade-Fichou: supervision.
 RP Martins: supervision.
 R Fahrenaus: conceptualization.

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

References

- Apcher S, Daskalogianni C, Lejeune F, Manoury B, Imhoos G, Heslop L, Fahrenaus 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. doi:10.1073/pnas.1104104108
- Apcher S, Komarova A, Daskalogianni C, Yin Y, Malbert-Colas L, Fahrenaus R (2009) mRNA translation regulation by the Gly-Ala repeat of Epstein-Barr virus nuclear antigen 1. *J Virol* 83: 1289–1298. doi:10.1128/JVI.01369-08
- Arora A, Maiti S (2009) Differential biophysical behavior of human telomeric RNA and DNA quadruplex. *J Phys Chem B* 113: 10515–10520. doi:10.1021/jp810638n
- Aznauryan M, Noer SL, Pedersen CW, Mergny J, Teulade-Fichou M, Birkedal V (2021) Ligand binding to dynamically populated G-quadruplex DNA. *ChemBioChem* 22: 1811–1817. doi:10.1002/cbic.202000792
- Ballestas ME (1999) Efficient persistence of extrachromosomal KSHV DNA mediated by latency-associated nuclear antigen. *Science* 284: 641–644. doi:10.1126/science.284.5414.641
- Ballestas ME, Kaye KM (2001) Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen 1 mediates episome persistence through cis-acting terminal repeat (TR) sequence and specifically binds TR DNA. *J Virol* 75: 3250–3258. doi:10.1128/JVI.75.7.3250-3258.2001
- Beaudoin J-D, Novoa EM, Vejnar CE, Yartseva V, Takacs CM, Kellis M, Giraldez AJ (2018) Analyses of mRNA structure dynamics identify embryonic gene regulatory programs. *Nat Struct Mol Biol* 25: 677–686. doi:10.1038/s41594-018-0091-z
- Beaudoin J-D, Perreault J-P (2010) 5'-UTR G-quadruplex structures acting as translational repressors. *Nucleic Acids Res* 38: 7022–7036. doi:10.1093/nar/gkq557
- Beltran M, Tavares M, Justin N, Khandelwal G, Ambrose J, Foster BM, Worlock KB, Tvardovskiy A, Kunzelmann S, Herrero J, et al (2019) G-tract RNA removes Polycomb repressive complex 2 from genes. *Nat Struct Mol Biol* 26: 899–909. doi:10.1038/s41594-019-0293-z
- Bezzi G, Piga EJ, Binolfi A, Armas P (2021) CNBP binds and unfolds in vitro G-quadruplexes formed in the SARS-CoV-2 positive and negative genome strands. *Int J Mol Sci* 22: 2614. doi:10.3390/ijms22052614
- Biffi G, Di Antonio M, Tannahill D, Balasubramanian S (2014) Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells. *Nat Chem* 6: 75–80. doi:10.1038/nchem.1805
- Boh lova N, Cantara A, Bartas M, Kaura P, Štastny J, Pecinka P, Fojta M, Mergny J-L, Brazda V (2021) Analyses of viral genomes for G-quadruplex forming sequences reveal their correlation with the type of infection. *Biochimie* 186: 13–27. doi:10.1016/j.biochi.2021.03.017
- Borkosky SS, Camporeale G, Chemes LB, Risso M, Noval MG, Sanchez IE, Alonso LG, de Prat Gay G (2017) Hidden structural codes in protein intrinsic disorder. *Biochemistry* 56: 5560–5569. doi:10.1021/acs.biochem.7b00721
- Candeias MM, Malbert-Colas L, Powell DJ, Daskalogianni C, Maslon MM, Naski N, Bourougaa K, Calvo F, Fahrenaus R (2008) P53 mRNA controls p53 activity by managing Mdm2 functions. *Nat Cell Biol* 10: 1098–1105. doi:10.1038/ncb1770
- Cardinaud S, Moris A, Fevrier M, Rohrlich P-S, Weiss L, Langlade-Demoyen P, Lemonnier FA, Schwartz O, Habel A (2004) Identification of cryptic MHC I-restricted epitopes encoded by HIV-1 alternative reading frames. *J Exp Med* 199: 1053–1063. doi:10.1084/jem.20031869
- Ceci M, Fazi F, Romano N (2021) The role of RNA-binding and ribosomal proteins as specific RNA translation regulators in cellular differentiation and carcinogenesis. *Biochim Biophys Acta Mol Basis Dis* 1867: 166046. doi:10.1016/j.bbadis.2020.166046
- Chen X-C, Chen S-B, Dai J, Yuan J-H, Ou T-M, Huang Z-S, Tan J-H (2018) Tracking the dynamic folding and unfolding of RNA G-quadruplexes in live cells. *Angew Chem Int Ed* 57: 4702–4706. doi:10.1002/anie.201801999
- Ciufo DM, Cannon JS, Poole LJ, Wu FY, Murray P, Ambinder RF, Hayward GS (2001) Spindle cell conversion by Kaposi's sarcoma-associated herpesvirus: Formation of colonies and plaques with mixed lytic and latent gene expression in infected primary dermal microvascular endothelial cell cultures. *J VIROL* 75: 5614–5626. doi:10.1128/JVI.75.12.5614-5626.2001
- Dabral P, Babu J, Zareie A, Verma SC (2019) LANA and hnRNP A1 regulate the translation of LANA mRNA through G-quadruplexes. *J Virol* 94: e01508–e01519. doi:10.1128/JVI.01508-19
- del Villar-Guerra R, Gray RD, Chaires JB (2017) Characterization of quadruplex DNA structure by circular dichroism. *Curr Protoc Nucleic Acid Chem* 68: 17.8.1–17.8.16. doi:10.1002/cpsc.23
- Dumas L, Herviou P, Dassi E, Cammas A, Millevoi S (2021) G-quadruplexes in RNA biology: Recent advances and future directions. *Trends Biochem Sci* 46: 270–283. doi:10.1016/j.tibs.2020.11.001
- Duss O, Stepanyuk GA, Puglisi JD, Williamson JR (2019) Transient protein-RNA interactions guide nascent ribosomal RNA folding. *Cell* 179: 1357–1369.e16. doi:10.1016/j.cell.2019.10.035
- Endoh T, Kawasaki Y, Sugimoto N (2013a) Stability of RNA quadruplex in open reading frame determines proteolysis of human estrogen receptor α . *Nucleic Acids Res* 41: 6222–6231. doi:10.1093/nar/gkt286
- Endoh T, Kawasaki Y, Sugimoto N (2013b) Suppression of gene expression by G-quadruplexes in open reading frames depends on G-quadruplex stability. *Angew Chem Int Ed* 52: 5522–5526. doi:10.1002/anie.201300058
- Endoh T, Sugimoto N (2016) Mechanical insights into ribosomal progression overcoming RNA G-quadruplex from periodical translation suppression in cells. *Sci Rep* 6: 22719. doi:10.1038/srep22719

- Endoh T, Sugimoto N (2013) Unusual -1 ribosomal frameshift caused by stable RNA G-quadruplex in open reading frame. *Anal Chem* 85: 11435–11439. doi:10.1021/ac402497x
- Fay MM, Lyons SM, Ivanov P (2017) RNA G-quadruplexes in biology: Principles and molecular mechanisms. *J Mol Biol* 429: 2127–2147. doi:10.1016/j.jmb.2017.05.017
- Fleming AM, Ding Y, Alenko A, Burrows CJ (2016) Zika virus genomic RNA possesses conserved G-quadruplexes characteristic of the Flaviviridae family. *ACS Infect Dis* 2: 674–681. doi:10.1021/acsfeddis.6b00109
- Gnanasundram SV, Pyndiah S, Daskalogianni C, Armfield K, Nylander K, Wilson JB, Fähræus R (2017) PI3K δ activates E2F1 synthesis in response to mRNA translation stress. *Nat Commun* 8: 2103. doi:10.1038/s41467-017-02282-w
- Guo JU, Bartel DP (2016) RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. *Science* 353: aaf5371. doi:10.1126/science.aaf5371
- Habchi J, Tompa P, Longhi S, Uversky VN (2014) Introducing protein intrinsic disorder. *Chem Rev* 114: 6561–6588. doi:10.1021/cr400514h
- Harris LM, Merrick CJ (2015) G-quadruplexes in pathogens: A common route to virulence control? *PLoS Pathog* 11: e1004562. doi:10.1371/journal.ppat.1004562
- Herdy B, Mayer C, Varshney D, Marsico G, Murat P, Taylor C, D'Santos C, Tannahill D, Balasubramanian S (2018) Analysis of NRAS RNA G-quadruplex binding proteins reveals DDX3X as a novel interactor of cellular G-quadruplex containing transcripts. *Nucleic Acids Res* 46: 11592–11604. doi:10.1093/nar/gky861
- 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. doi:10.1038/s41467-020-16168-x
- Jackson RJ, Hellen CUT, Pestova TV (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* 11: 113–127. doi:10.1038/nrm2838
- Jaubert C, Bedrat A, Bartolucci L, Di Primo C, Ventura M, Mergny J-L, Amrane S, Andreola M-L (2018) RNA synthesis is modulated by G-quadruplex formation in Hepatitis C virus negative RNA strand. *Sci Rep* 8: 8120. doi:10.1038/s41598-018-26582-3
- Ji D, Juhas M, Tsang CM, Kwok CK, Li Y, Zhang Y (2021) Discovery of G-quadruplex-forming sequences in SARS-CoV-2. *Brief Bioinform* 22: 1150–1160. doi:10.1093/bib/bbaa114
- Kikin O, D'Antonio L, Bagga PS (2006) QGRS mapper: A web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res* 34: W676–W682. doi:10.1093/nar/gkl253
- Kwun HJ, da Silva SR, Shah IM, Blake N, Moore PS, Chang Y (2007) Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen 1 mimics epstein-barr virus EBNA1 immune evasion through central repeat domain effects on protein processing. *J Virol* 81: 8225–8235. doi:10.1128/JVI.00411-07
- Lan K, Kuppers DA, Verma SC, Robertson ES (2004) Kaposi's sarcoma-associated herpesvirus-encoded latency-associated nuclear antigen inhibits lytic replication by targeting rta: A potential mechanism for virus-mediated control of latency. *J Virol* 78: 6585–6594. doi:10.1128/JVI.78.12.6585-6594.2004
- Lavigne M, Helynck O, Rigolet P, Boudria-Souilah R, Nowakowski M, Baron B, Guittat L, Beauvineau C, Petres S, Granzhan A, et al (2021) SARS-CoV-2 Nsp3 unique domain SUD interacts with guanine quadruplexes and G4-ligands inhibit this interaction. *Nucleic Acids Res* 49: 7695–7712. doi:10.1093/nar/gkab571
- Lewis CJT, Pan T, Kalsotra A (2017) RNA modifications and structures cooperate to guide RNA–protein interactions. *Nat Rev Mol Cell Biol* 18: 202–210. doi:10.1038/nrm.2016.163
- Lin JQ, van Tartwijk FW, Holt CE (2020) Axonal mRNA translation in neurological disorders. *RNA Biol* 18: 936–961. doi:10.1080/15476286.2020.1822638
- Lista MJ, Martins RP, Angrand G, Quillévère A, Daskalogianni C, Voisset C, Teulade-Fichou M-P, Fähræus R, Blondel M (2017a) 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. doi:10.15698/mic2017.09.590
- Lista MJ, Martins RP, Billant O, Contesse M-A, Findakly S, Pochard P, Daskalogianni C, Beauvineau C, Guetta C, Jamin C, et al (2017b) Nucleolin directly mediates Epstein-Barr virus immune evasion through binding to G-quadruplexes of EBNA1 mRNA. *Nat Commun* 8: 16043. doi:10.1038/ncomms16043
- Liu N, Dai Q, Zheng G, He C, Parisien M, Pan T (2015) N6-methyladenosine-dependent RNA structural switches regulate RNA–protein interactions. *Nature* 518: 560–564. doi:10.1038/nature14234
- Lyonnais S, Gorelick RJ, Mergny J-L, Le Cam E, Mirambeau G (2003) G-quartets direct assembly of HIV-1 nucleocapsid protein along single-stranded DNA. *Nucleic Acids Res* 31: 5754–5763. doi:10.1093/nar/gkg716
- Majee P, Kumar Mishra S, Pandya N, Shankar U, Pasadi S, Muniyappa K, Nayak D, Kumar A (2020) Identification and characterization of two conserved G-quadruplex forming motifs in the Nipah virus genome and their interaction with G-quadruplex specific ligands. *Sci Rep* 10: 1477. doi:10.1038/s41598-020-58406-8
- Marquet R, Baudin F, Gabus C, Darlix J-L, Mougel M, Ehresmann C, Ehresmann B (1991) Dimerization of human immunodeficiency virus (type 1) RNA: Stimulation by cations and possible mechanism. *Nucleic Acids Res* 19: 2349–2357. doi:10.1093/nar/19.9.2349
- Marquet R, Paillart J-C, Skripkin E, Ehresmann C, Ehresmann B (1994) Dimerization of human immunodeficiency virus type 1 RNA involves sequences located upstream of the splice donor site. *Nucleic Acids Res* 22: 145–151. doi:10.1093/nar/22.2.145
- Martins RP, Findakly S, Daskalogianni C, Teulade-Fichou M-P, Blondel M, Fähræus R (2018) In cellulo protein-mRNA interaction assay to determine the action of G-quadruplex-binding molecules. *Molecules* 23: 3124. doi:10.3390/molecules23123124
- Martins RP, Malbert-Colas L, Lista MJ, Daskalogianni C, Apcher S, Pla M, Findakly S, Blondel M, Fähræus R (2019) Nuclear processing of nascent transcripts determines synthesis of full-length proteins and antigenic peptides. *Nucleic Acids Res* 47: 3086–3100. doi:10.1093/nar/gky1296
- McQuin C, Goodman A, Chernyshev V, Kametsky L, Cimini BA, Karhohs KW, Doan M, Ding L, Rafelski SM, Thirstrup D, et al (2018) CellProfiler 3.0: Next-generation image processing for biology. *PLoS Biol* 16: e2005970. doi:10.1371/journal.pbio.2005970
- Métifiot M, Amrane S, Litvak S, Andreola M-L (2014) G-quadruplexes in viruses: Function and potential therapeutic applications. *Nucleic Acids Res* 42: 12352–12366. doi:10.1093/nar/gku999
- Münz C (ed) (2015) Epstein barr virus volume 2: One herpes virus: Many diseases. In *Current Topics in Microbiology and Immunology*. Cham: Springer International Publishing. doi:10.1007/978-3-319-22834-1
- Murat P, Zhong J, Lekieffre L, Cowieson NP, Clancy JL, Preiss T, Balasubramanian S, Khanna R, Tellam J (2014) G-quadruplexes regulate Epstein-Barr virus-encoded nuclear antigen 1 mRNA translation. *Nat Chem Biol* 10: 358–364. doi:10.1038/nchembio.1479
- Naski N, Gajjar M, Bourougaa K, Malbert-Colas L, Fähræus R, Candéias MM (2009) The p53 mRNA-Mdm2 interaction. *Cell Cycle* 8: 31–34. doi:10.4161/cc.8.1.7326
- Piekna-Przybylska D, Sharma G, Bambara RA (2013) Mechanism of HIV-1 RNA dimerization in the central region of the genome and significance for viral evolution. *J Biol Chem* 288: 24140–24150. doi:10.1074/jbc.M113.477265

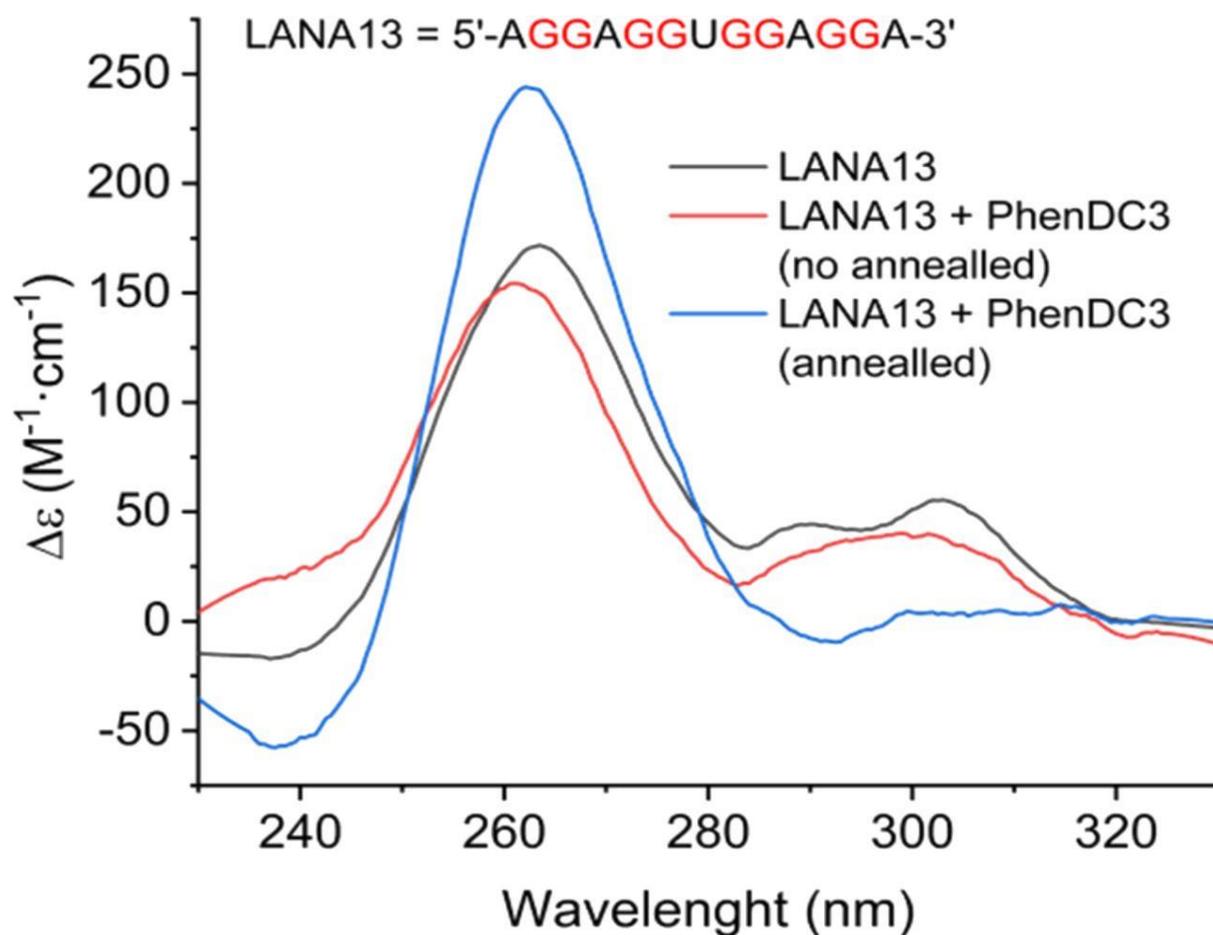
- Rainbow L, Platt GM, Simpson GR, Sarid R, Gao SJ, Stoiber H, Herrington CS, Moore PS, Schulz TF (1997) The 222- to 234-kilodalton latent nuclear protein (LNA) of Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) is encoded by orf73 and is a component of the latency-associated nuclear antigen. *J Virol* 71: 5915–5921. doi:10.1128/JVI.71.8.5915-5921.1997
- Reznichenko O, Quillévère A, Martins RP, Loaëc N, Kang H, Lista MJ, Beauvineau C, González-García 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. doi:10.1016/j.ejmech.2019.05.042
- Rodgers ML, Woodson SA (2019) Transcription increases the cooperativity of ribonucleoprotein assembly. *Cell* 179: 1370–1381.e12. doi:10.1016/j.cell.2019.11.007
- Ruggiero E, Tassinari M, Perrone R, Nadai M, Richter SN (2019) Stable and conserved G-quadruplexes in the long terminal repeat promoter of retroviruses. *ACS Infect Dis* 5: 1150–1159. doi:10.1021/acsinfecdis.9b00011
- Sanchez de Groot N, Armaos A, Graña-Montes R, Alriquet M, Calloni G, Vabulas RM, Tartaglia GG (2019) RNA structure drives interaction with proteins. *Nat Commun* 10: 3246. doi:10.1038/s41467-019-10923-5
- Schneider CA, Rasband WS, Eliceiri KW (2012) NIH image to ImageJ: 25 years of image analysis. *Nat Methods* 9: 671–675. doi:10.1038/nmeth.2089
- Serganov A, Nudler E (2013) A decade of riboswitches. *Cell* 152: 17–24. doi:10.1016/j.cell.2012.12.024
- Starck SR, Ow Y, Jiang V, Tokuyama M, Rivera M, Qi X, Roberts RW, Shastri N (2008) A distinct translation initiation mechanism generates cryptic peptides for immune surveillance. *PLoS One* 3: e3460. doi:10.1371/journal.pone.0003460
- Tosoni E, Frasson I, Scalabrin M, Perrone R, Butovskaya E, Nadai M, Palù G, Fabris D, Richter SN (2015) Nucleolin stabilizes G-quadruplex structures folded by the LTR promoter and silences HIV-1 viral transcription. *Nucleic Acids Res* 43: 8884–8897. doi:10.1093/nar/gkv897
- Vannutelli A, Belhamiti S, Garant J-M, Ouangraoua A, Perreault J-P (2020) Where are G-quadruplexes located in the human transcriptome? *NAR Genomics Bioinforma* 2: lqaa035. doi:10.1093/nargab/lqaa035
- Waldron JA, Raza F, Le Quesne J (2018) eIF4A alleviates the translational repression mediated by classical secondary structures more than by G-quadruplexes. *Nucleic Acids Res* 46: 3075–3087. doi:10.1093/nar/gky108
- Wang S-R, Min Y-Q, Wang J-Q, Liu C-X, Fu B-S, Wu F, Wu L-Y, Qiao Z-X, Song Y-Y, Xu G-H, et al (2016a) A highly conserved G-rich consensus sequence in hepatitis C virus core gene represents a new anti-hepatitis C target. *Sci Adv* 2: e1501535. doi:10.1126/sciadv.1501535
- Wang S-R, Zhang Q-Y, Wang J-Q, Ge X-Y, Song Y-Y, Wang Y-F, Li X-D, Fu B-S, Xu G-H, Shu B, et al (2016b) Chemical targeting of a G-quadruplex RNA in the ebola virus L gene. *Cell Chem. Biol* 23: 1113–1122. doi:10.1016/j.chembiol.2016.07.019
- Wei J, Kishton RJ, Angel M, Conn CS, Dalla-Venezia N, Marcel V, Vincent A, Catez F, Ferré S, Ayadi L, et al (2019) Ribosomal proteins regulate MHC class I peptide generation for immunosurveillance. *Mol Cell* 73: 1162–1173.e5. doi:10.1016/j.molcel.2018.12.020
- Wolfe AL, Singh K, Zhong Y, Drewe P, Rajasekhar VK, Sanghvi VR, Mavrakis KJ, Jiang M, Roderick JE, Van der Meulen J, et al (2014) RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer. *Nature* 513: 65–70. doi:10.1038/nature13485
- Yang SY, Lejault P, Chevrier S, Boidot R, Robertson AG, Wong JMY, Monchaud D (2018) Transcriptome-wide identification of transient RNA G-quadruplexes in human cells. *Nat Commun* 9: 4730. doi:10.1038/s41467-018-07224-8
- Yewdell JW, Dersh D, Fähræus R (2019) Peptide channeling: The key to MHC class I immunosurveillance? *Trends Cell Biol* 29: 929–939. doi:10.1016/j.tcb.2019.09.004
- Yin Y (2003) Self-inhibition of synthesis and antigen presentation by Epstein-Barr virus-encoded EBNA1. *Science* 301: 1371–1374. doi:10.1126/science.1088902
- Zhao C, Qin G, Niu J, Wang Z, Wang C, Ren J, Qu X (2021) Targeting RNA G-quadruplex in SARS-CoV-2: A promising therapeutic target for COVID-19? *Angew Chem Int Ed* 60: 432–438. doi:10.1002/anie.202011419



License: This article is available under a Creative Commons License (Attribution 4.0 International, as described at <https://creativecommons.org/licenses/by/4.0/>).

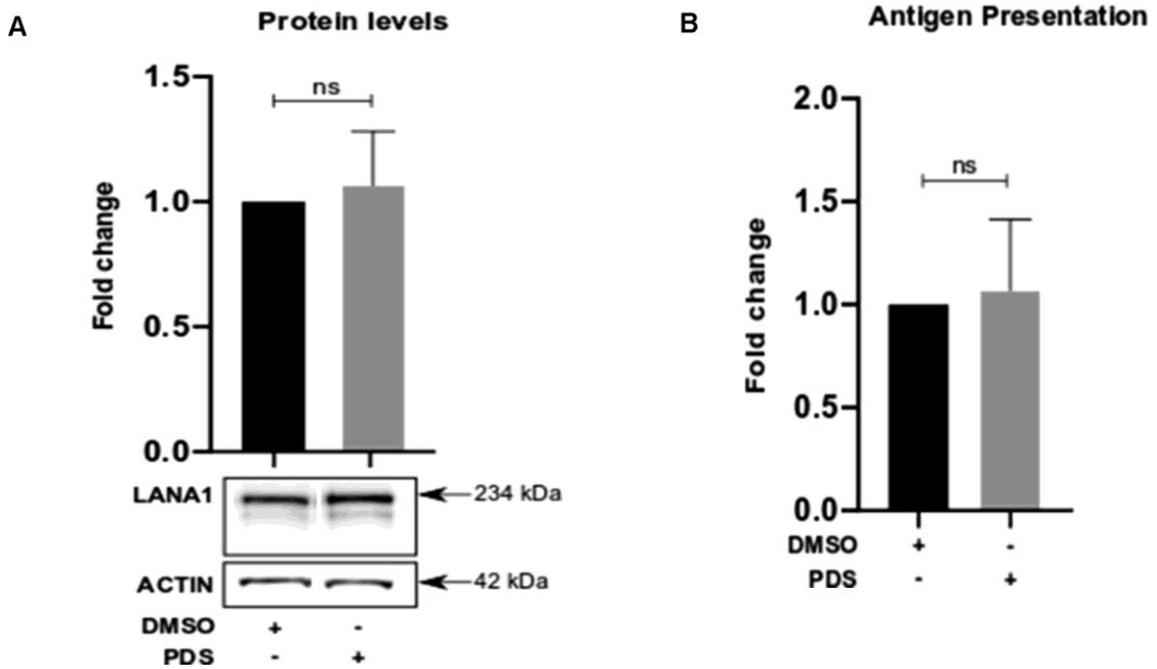
Generated Construct	Template	Cloning method	Sequence (5'-3')
LANA1-SL8	LANA1 - pcDNA3	Side Directed Mutagenesis	CTGGGGACTCTCCACAGGAAATGACAAGTATAATCAACTTTGAAAACTG GGGATATCCATCACACTGGCGGCCGCT AGCGGCCGCCAGTGTGATGGATATCCCCAGTTTTTCAAAGTTGATTATAC TTGTCATTTCTGTGGAGAGTCCCCAG
LANA1-N-terminal	LANA1-SL8	PCR, restriction ligation	GCATAAGCTTATGGCGCCCCGGGAATGCGC CGCGGAATTCCTGACTTTCCTTGCTAATCTCGTTGT
LANA1-C-terminal- SL8	LANA1-SL8	PCR, restriction ligation	GCGCGAATTCATCTTGACGGGTCGTCATCCGAGGACG CGCGGGGCCCTTACTTGTGTCATCGTCTTTGTAGTCC
LANA1ΔCR	LANA1-N-terminal, LANA1-C-terminal- SL8	restriction ligation	-
5' c-myc-IRES-LANA1	c-myc-Ova (41)	PCR, restriction ligation	GCGCAAGCTTAACTCGCTGTAGTAATCCAGCG GCGCGGATCCTGGGTTTACTCTTCCCG
3' c-myc-IRES-LANA1	c-myc-Ova (41)	PCR, restriction ligation	GCGCTCTAGATCCCCTGTGAGGAACTACTGT CGTAGGGCCCCGATGCACGGTCTACGAGACCT
5' HCV-IRES-LANA1	HCV-Ova (41)	PCR, restriction ligation	GCGCAAGCTT TCCCCTGTGAGGAACTACTGT CGTAGGATCCAGGTCTCGTAGACCGTGCATC
3' HCV-IRES-LANA1	HCV-Ova (41)	PCR, restriction ligation	GCGCTCTAGAAACTCGCTGTAGTAATCCAGCG GCGCGGGCCCTGGGTTTACTCTTCCCG

Supplementary Table 1. Primers and constructs used for cloning



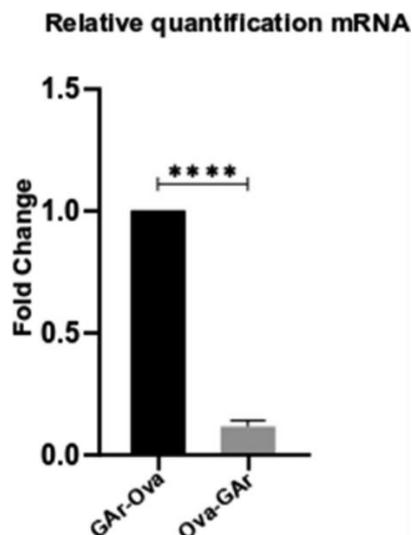
Supp. Fig. S1: Effects of PhenDC3 ligand on the circular dichroism spectra of the RNA sequence LANA13.

Effects of PhenDC3 ligand on the circular dichroism spectra of the RNA sequence LANA13. RNA (black), in the presence of ligand prior annealing (blue), or after annealing (red). Relates to Fig 1.



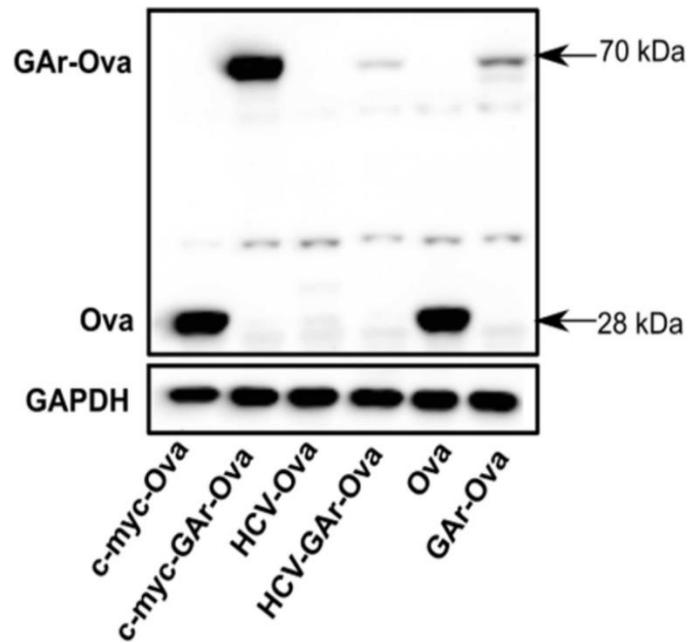
Supp. Fig. S2: Effects of the G4 RNA PDS ligand.

Effects of the G4 RNA PDS ligand. **(A)** Western blots (WB) show the expression of LANA1 from cDNA construct presented in (Fig 1A) the following treatment with 2 μ M of G4 ligands PDS for 24 h. WB show one of at least three similar experiments. **(B)** Relative amount of SL8 antigen peptide produced from the LANA1 cDNA presented in (Fig 1A) under similar conditions in H1299 cells expressing the murine MHC-I (Kb). The levels were estimated by measuring IL2 release from OT1 CD8⁺ T cells. Relates to Fig 2B and C.



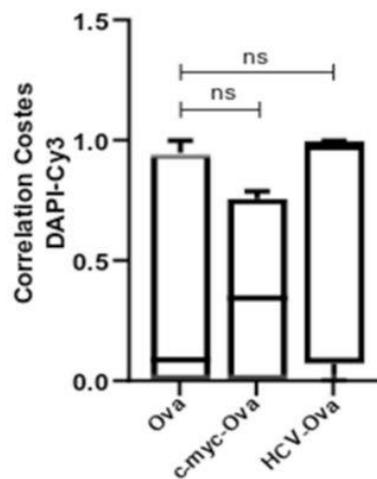
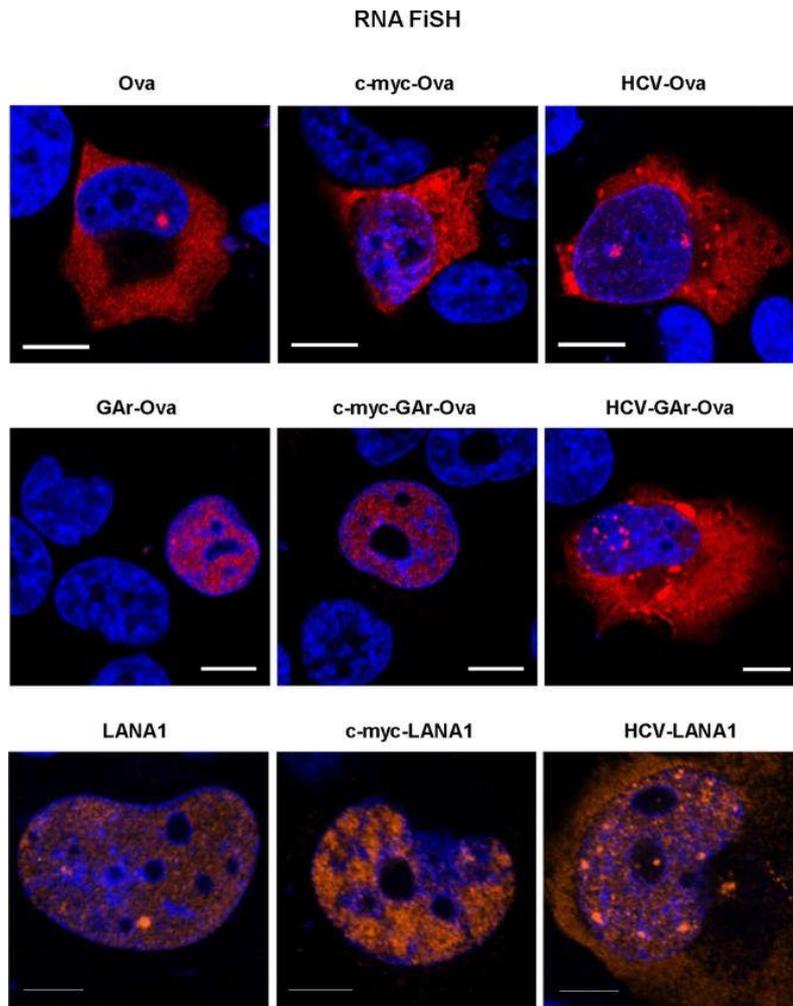
Supp. Fig. S3: mRNAs encoding GAR fused to Ova relative levels 24 h after transfection of H1299 cells with the indicated constructs.

mRNAs encoding GAR fused to Ova relative levels 24 h after transfection of H1299 cells with the indicated constructs. Relates to Fig 5.



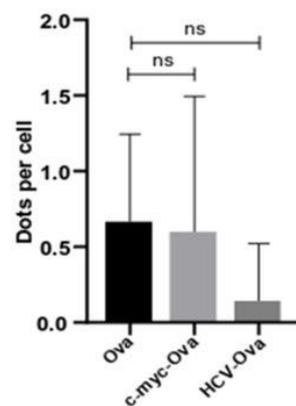
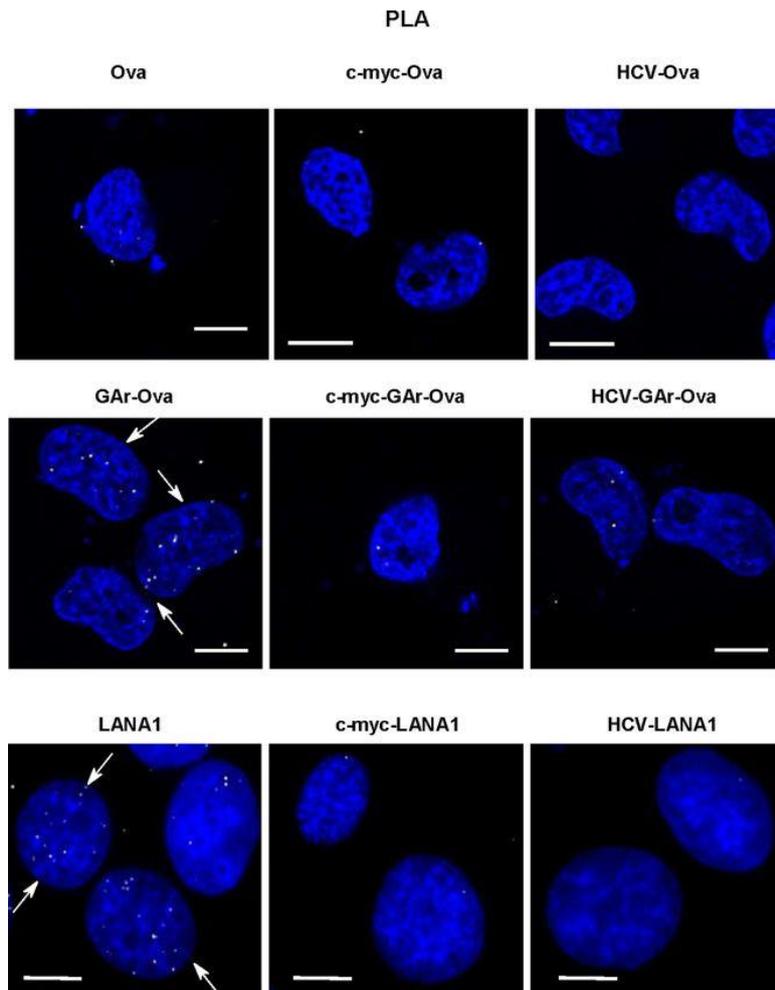
Supp. Fig. S4: Western blot with cell lysate from H1299 cells transfected with Ova/GAr-Ova constructs bearing c-myc or HCV IRES in the 5'UTR.

Western blot with cell lysate from H1299 cells transfected with Ova/GAr-Ova constructs bearing c-myc or HCV IRES in the 5'UTR. Relates to Fig 6.



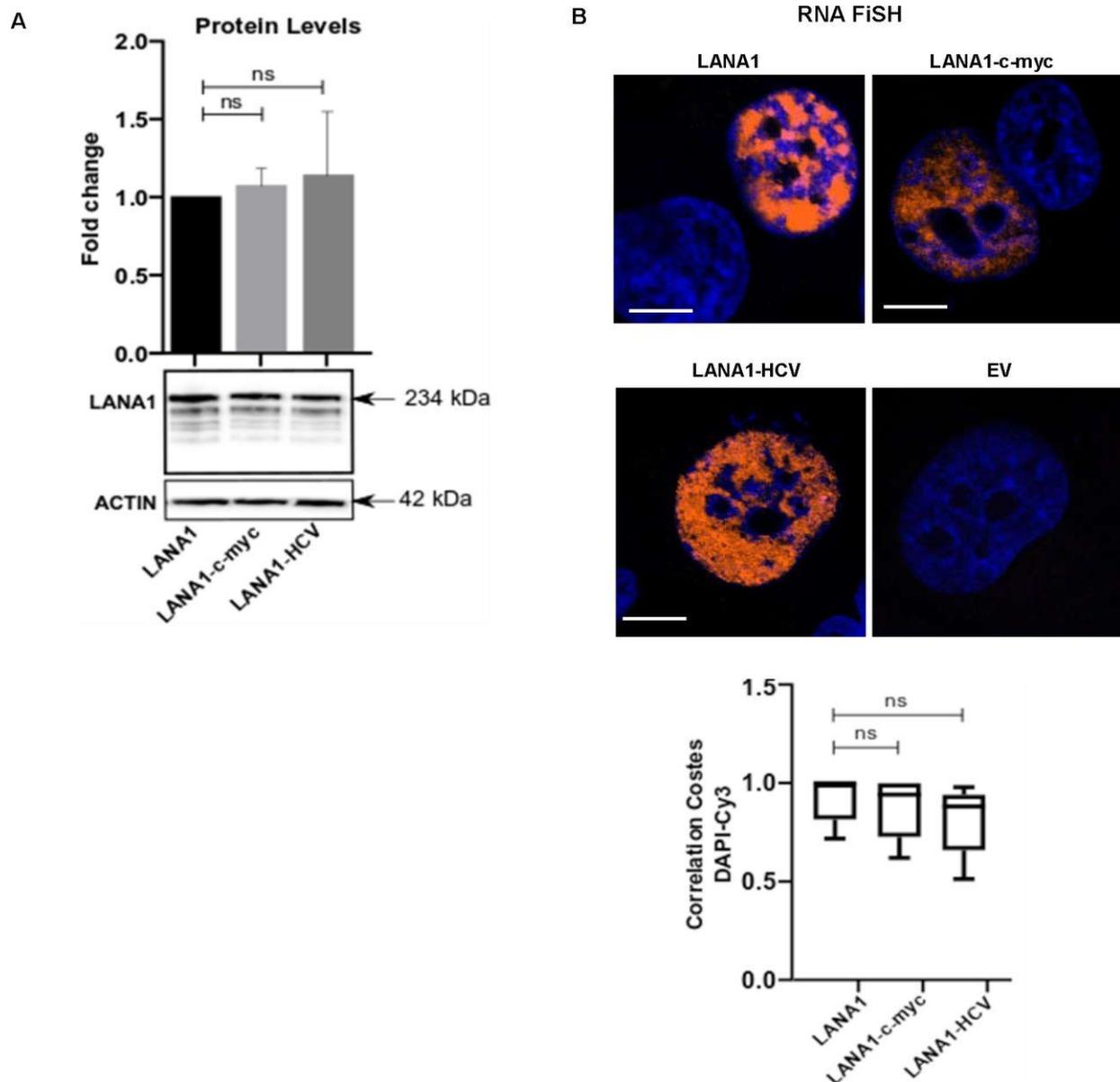
Supp. Fig. S5: Ova (upper panels) and LANA1 (lower panels) mRNA localisation (RNA FISH) in H1299 cells transfected with the indicated constructs (scale bar: 10 μ m).

Ova (upper panels) and LANA1 (lower panels) mRNA localisation (RNA FISH) in H1299 cells transfected with the indicated constructs (scale bar: 10 μ m). The graph below shows Costes Correlation factor between nucleus and LANA1 mRNA staining in H1299 cells transfected with the indicated cDNA. Relates to Fig 6E.



Supp. Fig. S6: Proximity ligation assay for screening nucleolin-Ova mRNA (upper panels) and nucleolin-LANA1 mRNA (lower panels) in H1299 cells transfected with the indicated constructs (scale bar: 10 μ m).

Proximity ligation assay for screening nucleolin-Ova mRNA (upper panels) and nucleolin-LANA1 mRNA (lower panels) in H1299 cells transfected with the indicated constructs (scale bar: 10 μ m). The graph below shows the number of dots per cells resulting from the proximity ligation assay between NCL protein and Ova mRNA. Relates to Fig 6F.



Supp. Fig. S7: Altering the 3' UTR does not affect the activities of the LANA1 G4s.

Altering the 3' UTR does not affect the activities of the *LANA1* G4s. **(A)** Western blot with cell lysate from H1299 cells transfected with LANA1 FL constructs bearing c-myc or HCV IRES in the 3' UTR. The graph above shows expression relative to actin from three independent experiments. **(B)** LANA1 mRNA localisation (RNA FiSH) in H1299 cells transfected with cDNA constructs as in (A) (scale bar: 10 μ m). The graph below shows Costes Correlation factor between nucleus and LANA1 mRNA staining . Relates to Fig 6.

Paper 2: The nascent polypeptide controls mRNA translation via the NAC and changes in its encoding mRNA NCL-binding capacity

The ribosome-associated NAC chaperone mediates mRNA translation regulation *in cis* via the nascent peptide and the encoding RNA sequence.

Alice J.-L. Zheng¹, Chrysoula Daskalogianni¹, Laurence Malbert-Colas¹, Konstantinos Karakostis¹, Mario Señorale², Aikaterini Thermou^{1,3}, Ronan Le Sénéchal^{4,8}, Van Trang Dinh^{4,8}, Maria C. Tovar Fernandez^{1,3}, Sébastien Apcher⁵, Sa Chen⁶, Marc Blondel⁴ and Robin Fahraeus^{1,6,7*}

¹Inserm UMRs 1131, Institut de Génétique Moléculaire, Université de Paris, Hôpital St. Louis, F-75010 Paris, France

²Sección Bioquímica, Facultad de Ciencias, Universidad de la República, Iguá 4225, Montevideo, Uruguay

³ICCVS, University of Gdańsk, Science, ul. Wita Stwosza 63, 80-308 Gdańsk, Poland

⁴Inserm UMR 1078, Université de Bretagne Occidentale (UBO), Etablissement Français du Sang (EFS) Bretagne, CHRU Brest, 29200, Brest, France

⁵Institut Gustave Roussy, Université Paris Sud, Unité 1015 département d'immunologie, 114, rue Edouard Vaillant, 94805 Villejuif, France

⁶Department of Medical Biosciences, Building 6M, Umeå University, 901 85 Umeå, Sweden

⁷RECAMO, Masaryk Memorial Cancer Institute, Zlutý kopec 7, 65653 Brno, Czech Republic.

⁸These authors contributed equally

*To whom correspondence should be addressed

Running title: Relationships between quality control of the nascent polypeptide, mRNA-RBP interactions, and mRNA translation control.

Keywords: nascent polypeptide quality control, G-quadruplex RNA-binding protein, NAC, aggregate-prone proteins, EBNA1

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 evades the immune system by suppressing its own synthesis. Here we show how the emerging peptide of a disordered glycine–alanine repeat (GAR) within EBNA1 dislodges the nascent polypeptide-associated complex (NAC) from the ribosome and recruits nucleolin to the G-quadruplex (G4) structures in the GAR-encoding mRNA, resulting in the suppression of mRNA translation initiation *in cis*. Downregulating NACA overcomes translation inhibition and prevents nucleolin from binding to the GAR mRNA. 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 the sensing of the nascent GAR peptide by NAC followed by the recruitment of nucleolin to the encoding RNA sequence.

INTRODUCTION

The formation and persistence of aggregates in cells are associated with numerous neurodegenerative diseases, including Alzheimer's disease with the development of β -amyloid plaques, or in Huntington's disease where CAG expansions and the consequent 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 of proteins, but also by engaging with misfolded or aggregate-prone proteins for refolding or for targeting them to degradation (2–4). But how the cell senses and prevents the synthesis of aggregate-prone proteins is still relatively unknown.

A network of chaperones engages with the nascent polypeptides (5). The nascent polypeptide-associated complex (NAC) is present at the ribosome exit tunnel on the large ribosomal subunit. NAC is composed of two subunits named NACA and BTF3 (also called NACB) (6–8), interacts with short nascent polypeptides and protects them from proteolysis (9, 10). NAC is suggested to be the first chaperone the nascent polypeptide encounters (9, 11, 12). Crystallography studies show that the BTF3 N-tail is present in the exit tunnel and escorts the nascent polypeptide (10). NAC controls the co-translational interaction of the nascent chains with other nascent polypeptide-associated factors (5), such as the SRP (signal recognition particle) and the methionine aminopeptidase (13). The NAC has a flexible conformation and can interact with a wide range of substrates (14) and recent studies show that NAC chaperone activity can prevent aggregates accumulation outside of its ribosome-bound state (15, 16).

mRNA translation is a tightly regulated and dynamic process regulated by multiple signalling pathways (17). This process is selective as some, but not all, pools of mRNAs are actively translated, depending on cell type and status (17, 18). RNA-binding proteins (RBPs) are adaptable regulators of mRNA translation that play a vital role in most physiological processes (18–20). RBPs interact with RNA sequences or with RNA structures to guide translation, but it has also been suggested that translation drives the formation of RNA structures (21).

The ribosome plays an important role in the quality control of both the mRNA and the encoded peptide. Faulty mRNAs are prevented from maturation and are targeted for degradation via different pathways such as nonsense-mediated decay, no-go decay, and non-stop-decay (22). Ribosome stalling leads to co-translational ubiquitination and degradation of the nascent polypeptide via the ribosome-associated protein quality control (RQC) pathway or the resuming of this process, but little is known about how the decision between degradation and resuming is made (22–26).

EBNA1 is the genome maintenance protein of the Epstein-Barr virus that latently infects over 90% of the world population (27). EBNA1 is required for viral genome replication and is indispensable

for EBV survival and expressed in all virus-carrying cells (27). Despite the presence of a primed immune system against it (28–32), EBNA1 escapes the host immune surveillance thanks to a *cis*-acting mechanism whereby an aggregate-prone glycine-alanine repeat (GAR) domain of EBNA1 suppresses its own synthesis *in cis* thus minimising the production of antigenic peptides for the major histocompatibility (MHC) class I pathway (33–35). The GAR-encoding mRNA forms G-quadruplex (G4) structures that interact with nucleolin (NCL) which inhibits mRNA translation initiation *in cis* of every open reading frame to which it is fused (36, 37). The GAR is a disordered domain consisting of single alanines separated by one, two, or three glycines and is predicted to be insoluble in *E.coli* (**Table S1**). Introducing serine residues in every eight amino acids of the GAR or disrupting the G4 structures alleviates GAR-mediated inhibition of translation (35).

Here we have used the GAR to show how a nascent disordered peptide acts together with G4 structures within the encoding RNA sequence to control its mRNA translation initiation *in cis*. Our results suggest a model whereby EBNA1 exploits a nascent protein quality control pathway to minimise the synthesis of antigenic peptides in order to evade the immune system.

MATERIAL AND METHODS

Cell culture

Human carcinoma-derived cell line H1299 and EBV carrying Burkitt's lymphoma cell lines Raji, B95.8 and nasopharyngeal carcinoma cell line NPC-6661 were 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-6661 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 (38). Solubilities of the polypeptides for *E.coli* expression were predicted using the webserver SoluProt v1.0 (39).

Western Blotting

Cells were harvested 40h post-transfection and lysed in the presence of a complete protease inhibitors cocktail (Roche Diagnostics). The whole procedure is detailed in supplementary Material and Methods. The following primary antibodies were used: anti-EBNA1 mouse monoclonal antibody (OT1X, Cyto-Barr), anti-p53 mouse monoclonal antibody (DO-1), anti-HA mAb (provided by Dr. B. Vojtěšek, Masaryk Memorial Cancer Institute, Brno, Czech Republic), anti-NACA rabbit polyclonal antibody (Abcam), anti-Ovalbumin whole serum (Sigma), anti-NCL polyclonal rabbit antibody (Abcam), anti-NACA mAb (Abnova), anti-LC3B rabbit polyclonal antibody (Sigma) and anti-actin mAb (Sigma). Relative quantifications of the HRP signals normalised with the corresponding actin bands are mentioned above the band. The condition used as a reference for the relative quantification is put in *italic*.

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

Cells were washed in cold PBS and total RNA extraction was performed using RNeasy 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)₁₂₋₁₈ primer (Life technologies). qPCR was performed using the StepOne real-time PCR system (Applied Biosystems) with Perfecta SYBR Green FastMix (Quanta Biosciences).

MHC class I restricted antigen presentation

T-cell assays using the B3Z SIINFEKL:Kb-specific T cell hybridoma were carried out as described previously (40). 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 20h. 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 K₂HPO₄, 0.5M KH₂PO₄ 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 hour 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 unit.

RNA FISH

Coverslips covered with cells 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 20min 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 (41), and later used for statistical analysis.

Proximity Ligation Assay (PLA) protein-protein and immunostaining

Cells were cultured, fixed, and permeabilised as described in supplementary Material and Methods. Primary antibodies incubation and PLA were carried out using the Naveniflex MR kit (Navinci), following the manufacturer's protocol. The aggregates staining was realised using the Proteostat Aggresome detection kit (Enzo) following the manufacturer's instructions. PLA primary antibodies: anti-HA mouse monoclonal antibody (provided by Dr. B. Vojtěšek, Masaryk Memorial Cancer Institute, Brno, Czech Republic), anti-NACA rabbit polyclonal antibody (Abcam), anti-HA rabbit polyclonal antibody (Sigma), anti-NACA monoclonal mouse antibody (Abnova), and anti-EBNA1 mouse monoclonal antibody (OT1X, Cyto-Barr). Images were obtained using Zen software and analysed using CellProfiler (41). 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 20mM Tris (pH 7.5), 150 mM NaCl, 1% NP-40 in the presence of complete protease inhibitor cocktail (Roche, Mannheim, Germany). 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% pre-cast gels (Invitrogen).

Polysome fractionation and PLEA

5-50% wt/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 μ l 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

H1299 cells were collected and washed twice with 1X PBS (Gibco) before lysis was performed. 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 G-quadruplex forming RNA oligonucleotides bound to the streptavidin beads. In the input lane of the western blots was loaded a quantity of extract which corresponds to the quantity 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

NACA controls the synthesis of GAR-containing polypeptides

Previous studies have shown that the binding of nucleolin (NCL) to the *EBNA1* mRNA is important for translation suppression but the addition of antibodies to the gly-ala repeat (GAR) domain of EBNA1 overcomes translation suppression showing that the peptide also plays a role (34, 36, 37). We wanted to know how the mRNA and the encoded peptide act together to control EBNA1 synthesis. The GAR domain is located in the N-terminal part of the EBNA1 and deleting this domain (*EBNA1* Δ GAR) increases mRNA translation and protein expression (**Figs. 1A and 1B**) (34, 35). NAC is described as a ribosome-associated chaperone targeting aggregate-prone proteins and we tested the impact of NAC on GAR-bearing protein synthesis. siRNA-mediated silencing of NACA in H1299 cells leads to a decrease in *NACA* mRNA and protein levels that 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 (**Fig. 1C and Supp. Fig. S1A, S1B and S1C**) (16, 42). This was accompanied by an approximately 4-fold increase in EBNA1 expression without affecting *EBNA1* mRNA levels (**Figs. 1C and D**). A similar increase in expression following NACA knockdown was observed for the GAR fused to p53 (GAR-p53) or GAR fused to ovalbumin (Ova) (GAR-Ova) (**Fig. 1E and Supp. Fig. S1D**). This shows that silencing NACA interferes with GAR-mediated translation inhibition. The GAR is more efficient in suppressing translation when placed in the 5' of the coding sequence, as compared to the 3' (34, 35), 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, rather than protein stability (**Supp. Fig. S1D**). We also treated NPC666-1 cells expressing endogenous EBNA1 with siRNA against NACA and we observed a similar increase, even though not as strong as in transfected cells (**Fig. 1F**). The physiological role of the GAR is to minimise the production of EBNA1-derived peptide substrates for the MHC class I pathway (33) and we wanted to know if NACA also controls the presentation of EBNA1 antigen. The Ova includes an antigenic peptide (SL8) 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 transfected with GAR-Ova but not with Ova. Silencing SRP, another nascent polypeptide-associated factor, did not impact antigen presentation (**Fig. 1G**).

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 *cMyc* to the 5' UTR of GAR-carrying constructs overcomes GAR-mediated translation inhibition (**Fig. 1H**) (35). When we knocked out NACA we observed no effect on the expression of EBNA1 carrying *cMyc* IRES in the 5' UTR, validating that the effect of NACA is indeed on the synthesis of EBNA1 and not on potential EBNA1 aggregates stability (**Fig. 1I**). Similarly, when we fused the *cMyc* 5'UTR to the 5' UTR of GAR-p53 or GAR-Ova constructs, we did not observe any increase in expression following NACA knockdown (**Fig. 1J and Supp. Fig. S1E**). Silencing of NACA on constructs bearing HCV IRES also did not affect expression for both EBNA1 and GAR-Ova reporters (**Fig. 1H, 1I and Supp. Fig. S1E**). We also wanted to see if overexpressing NACA would affect EBNA1 expression. However, we observed no effect following co-expression of an HA-tagged NACA construct, indicating that the endogenous NACA levels are sufficient to control EBNA1 synthesis (**Fig. 1K**). Taken together, these results show that NACA controls translation initiation of mRNAs which bear the GAR sequence.

NACA interacts with the GAR domain

To address how NACA specifically interferes with GAR-mediated translation control, we performed Proximity Ligation Assay (PLA) between endogenous NACA and the EBNA1, EBNA Δ GAR, HA-p53, or HA-polyQ-p53 substrates (**Fig. 2A and Supp. Fig S2A**). We observed interactions between NACA and its different substrates mainly in the nucleus (**Fig. 2A and Supp. Fig S2A**). Immunofluorescence assays confirmed a cytoplasmic and nuclear localisation of NACA (**Supp. Fig S2B**) in line with previous studies (6, 43, 44).

The poly-glutamine repeat (PolyQ) is known to cause aggregates and adding an HA-tagged polyQ to p53 instead increased the relative number of PLA interactions between HA and NACA by nearly two-fold (**Fig. 2B, left and right graphs**). After taking into account the difference in expression between EBNA1 and EBNA Δ GAR, we observed that deleting the GAR decreased on average the number of interactions between endogenous NACA and EBNA1 by 30%. Similarly, overexpressing an HA-NACA construct together with EBNA1, or EBNA Δ GAR, showed an approximately 3-fold increase in the number of interactions between HA and EBNA1, as compared to HA and EBNA Δ GAR (**Fig. 2C**). Co-immunoprecipitation assays from two EBV-carrying B cell lines (Raji and B95.8) showed that endogenous EBNA1 interacts with NACA (**Fig. 2D**). These results show that NACA interacts with the two aggregate-prone repeat sequences, GAR and PolyQ. Only the GAR, and not the PolyQ, interferes with synthesis (35).

The nascent GAR dislodges the NACA from the ribosome

Having confirmed the interaction of the GAR with the alpha subunit of the ribosome-associated NAC complex (NACA), we next tested if NACA interacts with the nascent EBNA1 on the ribosome as this would help to explain how the interaction between the nascent GAR and NACA controls *EBNA1* translation *in cis*. We expressed p53 and GAR-p53 and carried out polysomal fractionation using sucrose gradients on cycloheximide-treated cell lysates (**Fig. 3A**). The isolated polysomes were captured to 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 (**Fig. 3B**). The fusion of the GAR sequence with p53 resulted in a stronger interaction between the nascent peptide and NACA, indicating more NACA associated with nascent GAR, despite GAR suppressing protein synthesis (**Fig. 3C**). We next tested if the high affinity of NACA for the GAR results in NACA becoming dislodged from the ribosome and for that we designed a construct in which a TEV protease cleavage site was inserted between the GAR and the p53 (GAR-TEV-p53). Following 60 minutes of TEV enzyme added to cell lysates, a major part of GAR was cleaved from p53 (**Fig. 3D**). We then expressed GAR-TEV-p53 and carried out polysomal gradients on TEV-treated and non-TEV-treated lysates (**Supp. Fig. S3A**). P53-carrying polysomes were fixed to 96-well plates using chicken anti-p53 antibodies and PLEA was performed using 125µg of polysomes as this was enough to accurately see a difference in PLA signals (see **Fig. 3C**). If NACA remained attached to the ribosome, TEV protease treatment would not make any difference on the PLA signal between NACA and RPL5, or NACA and RPL11 (**Fig. 3E; a and b**), whereas if the GAR sequence detached NACA from the ribosome, the detached NACA would wash away with the GAR peptide, resulting in a reduced NACA - RPL5/RPL11 PLA signal (**Fig. 3E; c**). The TEV treatment indeed resulted in less PLA signal between NACA - RPL5/RPL11 on the GAR-TEV-p53-expressing polysomes, in line with the notion that the GAR sequence detaches NACA from the ribosome (**Fig. 3F**). These results show that the nascent GAR dislodges NACA from the ribosome and offer a first insight into how NACA affects mRNA translation of GAR-carrying mRNAs *in cis*.

NACA is necessary for the GAR-encoding mRNA interaction with NCL

The results so far indicate that the nascent GAR peptide plays an important role in GAR-mediated mRNA translation control *in cis* by interacting with the NACA and dissociating it from the ribosome. However, previous works have demonstrated the importance of the GAR-encoding mRNA sequence in suppressing mRNA translation (37, 45). The effect of the mRNA is attributed to NCL binding a G-quadruplex (G4) structure in the GAR-encoding mRNA and we wanted to know if there is a link between dissociating NACA from the ribosome and the recruitment of NCL to the RNA. 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 of respective mRNAs. Silencing NACA resulted in a dramatic decrease in NCL binding to the *EBNA1* and *GAR-Ova* mRNAs but had little effect on the interaction with *Ova* mRNA (**Figs. 4A and 4B**). Silencing NACA does not affect *EBNA1* mRNA levels (see **Fig. 1D**). NACA has been described to bind nucleic acids (6) and using RNA-coIP and recombinant NACA and *EBNA1* mRNA from cell

lysates we observed that NACA binds to the *EBNA1* mRNA in a *GAr*-dependent fashion (**Fig. 4C**). To test if the NACA – *EBNA1* mRNA interaction is direct, we carried out RNA pulldown assays using recombinant NACA protein together with synthetic oligos mimicking the RNA G4 forming sequences from the *GAr* (GQ-18), a mutated version of GQ-18 that does not form a G4 (GM-18) or a well-known RNA G4-forming sequence (ARPC2). This showed that NACA binds RNA oligos in a G4-dependent manner (**Fig. 4D**). Adding recombinant NCL did not prevent NACA interaction (**Fig. 4D**) and *in situ* PLA showed that endogenous NCL and NACA interact in the nuclear compartment. The NACA - NCL interaction is detected in the nucleus and 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) (**Fig. 4E and Supp. Fig. S4A**). We next tested if NACA affects the interaction between the *EBNA1* mRNA and NCL. The interaction between recombinant NCL and the *EBNA1* mRNA derived from cell lysates was enhanced 1.6-fold in the presence of recombinant NACA. Importantly, the suppression of the NCL – *EBNA1* mRNA interaction observed following siRNA-mediated NACA silencing (see **Fig. 4A**) was almost completely reversed if recombinant NACA was added (**Fig. 4F**). Taken together, these results show that NACA binds directly to the *GAr* sequence and allows NCL to interact with the *EBNA1* mRNA.

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

In line with the scenario that the nascent *GAr* peptide dislodges NACA from the ribosome and that NACA recruits NCL to the *GAr* RNA G4 structure, the interaction between NCL and the *GAr* RNA would require the *GAr*-encoding mRNAs to be translated. We tested this hypothesis by treating cells with cycloheximide (CHX) or harringtonine (Harr), two drugs that stop mRNA translation. As expected, both drugs prevent the synthesis of *GAr*-Ova, Ova, and endogenous p21 (**Fig. 5A**, lower panel). Importantly, *in vitro* RNA-coIP showed that the interaction between the *GAr*-Ova mRNA and NCL was reduced approximately by 4-fold following CHX or Harr treatment. NCL has a 4-fold less affinity to the Ova mRNA as compared to *GAr*-Ova, and the NCL – Ova mRNA interaction was not affected by translation inhibitor treatments (**Fig. 5A**, upper panel), while for both constructs the mRNA levels were higher when treated with translation inhibitors compared to DMSO (**Supp. Fig. 5A**). To further test if the translation is needed to allow NCL to bind the *GAr*-encoding mRNA, we designed an Ova and *GAr*-Ova constructs lacking AUG start codons (*OvaΔATG* and *GAr*-*OvaΔATG* respectively) (**Fig. 5B**) abrogating the expression of respective proteins (**Fig. 5C**, lower panel) and we observed that NCL does not interact with the *OvaΔATG* and *GAr*-*OvaΔATG* mRNAs (**Fig. 5C**, higher panel). The deletion of the AUGs had no significant effect on mRNA levels (**Supp. Fig. S5B**) and RNA FiSH experiments showed that deleting the start codon from Ova and *GAr*-Ova mRNAs had no major impact on their respective subcellular localisation (**Fig. 5D** and **Supp. Fig. S5C**). Thus, *GAr*-encoding mRNAs need to be translated before binding NCL, further supporting the notion that the interaction between NACA and the nascent *GAr* is required for the recruitment of NCL to the G4 structure of the *GAr*-encoding mRNA.

DISCUSSION

Our data support the existence of a *cis*-acting protein-quality control pathway in which aggregate-prone nascent peptides dislodge NAC from the ribosome whereby NACA recruits nucleolin (NCL) to G-quadruplex structures in the encoding RNA sequence to suppress mRNA translation. Neither the mRNA nor the encoded peptide is targeted for degradation, illustrating a novel nascent peptide quality control pathway for post-transcriptional gene regulation. The NAC chaperone has a flexible conformation allowing it to interact with different substrates and its affinity for short polypeptides depends on the sequence of the nascent chain (11, 14). NAC contacts the ribosome at different points but it is not known if NAC binds to all its ribosomal partners at once (10, 13, 46, 47). Insertion of single serines in every eight residues of the GAR completely abolishes its translation inhibitory capacity (35). This, together with the observation that NACA has a high affinity for the GAR, supports the notion that the nature of the non-polar gly-ala repeat and the chaperone activity of NAC cause GAR to dislodge NACA from the ribosome. A different example of a nascent peptide affecting protein expression control comes from tubulin where the encoding mRNA is targeted for degradation to regulate tubulin expression levels (48).

The recruitment of NCL to the GAR-encoding RNA requires NACA, suggesting that NACA makes the G4 RNA structure accessible for NCL. Previous observations have shown that NCL does not have access to a GAR-carrying message that is subject to splicing, whereas the same mRNA interacts with NCL *in vitro* (45). The *EBNA1* message is, however, not spliced and it is unlikely that NACA would play a role in a splicing-dependent NCL – GAR mRNA interaction but it suggests that NCL friendly GAR-encoding G4 structures are influenced by different cellular mechanisms. Previous studies have shown that G4 RNA structures *in vivo* are highly dynamic with multi-functional properties and it is likely that NACA affects the G4 structure to suit NCL binding (49). 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 (50, 51).

In line with the notion that the virus is exploiting a *cis*-acting pathway that senses aggregate-prone proteins to minimise the synthesis of EBNA1-derived antigenic peptide substrates, we have previously reported that this pathway is conserved in yeast (52). The effect on translation suppression is via initiation and together with the fact that neither RNA nor peptide is degraded suggests that GAR-mediated translation control can be regulated. EBNA1 binds its own mRNA via its RGG-rich region (53) and it is, thus, possible that EBNA1 protein can feedback to control its own synthesis. This hypothesis needs to be tested but it is known that EBNA1 feeds back to control its own transcription (54). Another potential effect of the GAR RNA sequence comes from more recent studies suggesting that RNA G4, or RNA structures in general, can act as selective protein chaperones (55, 56) and, thus, help prevent aggregate formation. The essential role of the RNA sequence in this nascent peptide quality control pathway implies that aggregate-prone proteins that are not encoded by an RNA sequence that engages this pathway are expressed without control and leads to aggregate

formation in the cell. For example, the high affinity between NACA and the PolyQ of Huntingtin protein (HTT) does not result in *cis*-mediated translation suppression (**Table S1**) (35).

Patients suffering from Fragile X Syndrome have a (CGG) expansion that promotes alternative translation upstream of the classical *FMR1* ORF via repeat-associated non-AUG (RAN) mechanism (57, 58). Another example of GC-rich sequences affecting translation comes from the (GGGGCC) repeats expansion in the *C9ORF72* gene in the context of amyotrophic lateral sclerosis (ALS) (59). Both (CGG) and (GGGGCC) repeat expansions form G4 structures (60–63) and the encoded polypeptides produced are prone to form aggregates (64). Despite similarities in terms of GC-rich expansions between RAN translation and GAR-mediated translation control, there are important differences. RAN is an alternative initiation mechanism mostly described for non-coding regions and introns and the encoded peptides cause toxic aggregates. In the case of the GC repeat of the GAR, it forms part of a translation regulatory pathway that acts on initiation, requires translation of the coding sequence, and prevents aggregate formation.

Another aspect of NAC-mediated recruitment of NCL to the mRNA is where in the cell this event takes place. The PLA data show that NACA – GAR, NACA – polyQ, and NACA – NCL interactions take place in the nucleus. NACA is present in the nucleus and the cytoplasm while NCL is nuclear. It is interesting, and speculative, to consider that this nascent peptide control pathway is taking place in the nucleus that governs how the mRNA is processed and translated in the cytoplasm. If so, one would expect the synthesis of peptides in the nucleus with a high turnover rate. It has indeed been reported that a nuclear co-transcriptional-translation event produces peptides with a high turnover rate (65, 66). Nuclear translation has been reported by several groups but remains controversial (67–72) as the physiological relevance has remained unclear and it is, thus, possible that a peptide quality control scanning pathway can offer one explanation for these different observations.

ACKNOWLEDGEMENT

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.

AUTHORS CONTRIBUTIONS

A.Z. and C.D. carried out the majority of the work. L.M-C., A.T., 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.

DECLARATION OF INTERESTS

The authors declare no competing interests.

FUNDING

A.Z. was funded by Université de Paris and the Fondation pour la Recherche Médicale [grant number FDT202001010912]. This work was partially supported by 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, and by the 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.

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. Gonçalves,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. Deuerling,E., Gamerdinger,M. and Kreft,S.G. (2019) Chaperone Interactions at the Ribosome. *Cold Spring Harb. Perspect. Biol.*, **11**, a033977.
6. 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.

7. Wang,L., Zhang,W., Wang,L., Zhang,X.C., Li,X. and Rao,Z. (2010) Crystal structures of NAC domains of human nascent polypeptide-associated complex (NAC) and its α NAC subunit. *Protein Cell*, **1**, 406–416.
8. Beatrix,B., Sakai,H. and Wiedmann,M. (2000) The α and β Subunit of the Nascent Polypeptide-associated Complex Have Distinct Functions. *J. Biol. Chem.*, **275**, 37838–37845.
9. 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.
10. Gamerdinger,M., Kobayashi,K., Wallisch,A., Kreft,S.G., Sailer,C., Schlömer,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.e8.
11. Raue,U., Oellerer,S. and Rospert,S. (2007) Association of Protein Biogenesis Factors at the Yeast Ribosomal Tunnel Exit Is Affected by the Translational Status and Nascent Polypeptide Sequence. *J. Biol. Chem.*, **282**, 7809–7816.
12. Alamo,M. del, 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. Nyathi,Y. and Pool,M.R. (2015) Analysis of the interplay of protein biogenesis factors at the ribosome exit site reveals new role for NAC. *J. Cell Biol.*, **210**, 287–301.
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., Schlömer,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.e7.
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. Riechert,E., Kmietczyk,V., Stein,F., Schwarzl,T., Sekaran,T., Jürgensen,L., Kamuf-Schenk,V., Varma,E., Hofmann,C., Rettel,M., *et al.* (2021) Identification of dynamic RNA-binding proteins uncovers a Cpeb4-controlled regulatory cascade during pathological cell growth of cardiomyocytes. *Cell Rep.*, **35**, 109100.
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., Järvelin,A.I., González-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.e11.
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. Inada,T. (2020) Quality controls induced by aberrant translation. *Nucleic Acids Res.*, **48**, 1084–1096.
23. Joazeiro,C.A.P. (2019) Mechanisms and functions of ribosome-associated protein quality control. *Nat. Rev. Mol. Cell Biol.*, **20**, 368–383.
24. 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.
25. 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.
26. Chandrasekaran,V., Juszkievicz,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.
27. Young,L.S., Yap,L.F. and Murray,P.G. (2016) Epstein–Barr virus: more than 50 years old and still providing surprises. *Nat. Rev. Cancer*, **16**, 789–802.
28. 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.
29. Münz,C., Bickham,K.L., Subklewe,M., Tsang,M.L., Chahroudi,A., Kurilla,M.G., Zhang,D., O’Donnell,M. and Steinman,R.M. Human CD4 α T Lymphocytes Consistently Respond to the Latent Epstein-Barr Virus Nuclear Antigen EBNA.
30. Long,H.M., Haigh,T.A., Gudgeon,N.H., Leen,A.M., Tsang,C.-W., Brooks,J., Landais,E., Houssaint,E., Lee,S.P., Rickinson,A.B., *et al.* (2005) CD4 $^{+}$ T-Cell Responses to Epstein-Barr Virus (EBV) Latent-Cycle Antigens and the Recognition of EBV-Transformed Lymphoblastoid Cell Lines. *J. Virol.*, **79**, 4896–4907.
31. Paludan,C., Bickham,K., Nikiforow,S., Tsang,M.L., Goodman,K., Hanekom,W.A., Fonteneau,J.-F., Stevanović,S. and Münz,C. (2002) Epstein-Barr Nuclear Antigen 1-Specific CD4 $^{+}$ Th1 Cells Kill Burkitt’s Lymphoma Cells. *J. Immunol.*, **169**, 1593–1603.
32. Blake,N., Lee,S., Redchenko,I., Thomas,W., Steven,N., Leese,A., Steigerwald-Mullen,P., Kurilla,M.G., Frappier,L. and Rickinson,A. Human CD8 α T Cell Responses to EBV EBNA1: HLA Class I Presentation of the (Gly-Ala)–Containing Protein Requires Exogenous Processing.
33. Levitskaya,J., Coram,M., Levitsky,V., Imreh,S., Steigerwald-Mullen,P.M., Klein,G., Kurilla,M.G. and Masucci,M.G. (1995) Inhibition of antigen processing by the internal repeat region of the Epstein–Barr virus nuclear antigen-1. *Nature*, **375**, 685–688.

34. Yin,Y., Manoury,B. and Fåhraeus,R. (2003) Self-Inhibition of Synthesis and Antigen Presentation by Epstein-Barr Virus–Encoded EBNA. **301**, 5.
35. Apcher,S., Komarova,A., Daskalogianni,C., Yin,Y., Malbert-Colas,L. and Fåhraeus,R. (2009) mRNA Translation Regulation by the Gly-Ala Repeat of Epstein-Barr Virus Nuclear Antigen 1. *J. Virol.*, **83**, 1289–1298.
36. 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.
37. 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.
38. 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.
39. 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.
40. Apcher,S., Daskalogianni,C., Manoury,B. and Fåhraeus,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.
41. McQuin,C., Goodman,A., Chernyshev,V., Kametsky,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.
42. 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.
43. Quélo,I., Gauthier,C., Hannigan,G.E., Dedhar,S. and St-Arnaud,R. (2004) Integrin-linked Kinase Regulates the Nuclear Entry of the c-Jun Coactivator α -NAC and Its Coactivation Potency. *J. Biol. Chem.*, **279**, 43893–43899.
44. Meury,T., Akhouayri,O., Jafarov,T., Mandic,V. and St-Arnaud,R. (2010) Nuclear α NAC Influences Bone Matrix Mineralization and Osteoblast Maturation *In Vivo*. *Mol. Cell. Biol.*, **30**, 43–53.
45. Martins,R.P., Malbert-Colas,L., Lista,M.J., Daskalogianni,C., Apcher,S., Pla,M., Findakly,S., Blondel,M. and Fåhraeus,R. (2019) Nuclear processing of nascent transcripts determines synthesis of full-length proteins and antigenic peptides. *Nucleic Acids Res.*, **47**, 3086–3100.
46. Wegrzyn,R.D., Hofmann,D., Merz,F., Nikolay,R., Rauch,T., Graf,C. and Deuerling,E. (2006) A Conserved Motif Is Prerequisite for the Interaction of NAC with Ribosomal Protein L23 and Nascent Chains. *J. Biol. Chem.*, **281**, 2847–2857.
47. Pech,M., Spreter,T., Beckmann,R. and Beatrix,B. (2010) Dual Binding Mode of the Nascent Polypeptide-associated Complex Reveals a Novel Universal Adapter Site on the Ribosome. *J. Biol. Chem.*, **285**, 19679–19687.
48. 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.

49. Zheng,A.J.-L., 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.
50. 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.
51. Benhalevy,D., Gupta,S.K., Danan,C.H., Ghosal,S., Sun,H.-W., Kazemier,H.G., Paeschke,K., Hafner,M. and Juraneck,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.
52. Lista,M.J., Voisset,C., Contesse,M., Friocourt,G., Daskalogianni,C., Bihel,F., Fähræus,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.
53. Norseen,J., Thomae,A., Sridharan,V., Aiyar,A., Schepers,A. and Lieberman,P.M. (2008) RNA-dependent recruitment of the origin recognition complex. *EMBO J.*, **27**, 3024–3035.
54. Young,L.S. and Murray,P.G. (2003) Epstein-Barr virus and oncogenesis: from latent genes to tumours. *Oncogene*, **22**, 5108–5121.
55. Begeman,A., Son,A., Litberg,T.J., Wroblewski,T.H., Gehring,T., Huizar Cabral,V., Bourne,J., Xuan,Z. and Horowitz,S. (2020) G-Quadruplexes act as sequence-dependent protein chaperones. *EMBO Rep.*, **21**.
56. Aarum,J., Cabrera,C.P., Jones,T.A., Rajendran,S., Adiutori,R., Giovannoni,G., Barnes,M.R., Malaspina,A. and Sheer,D. (2020) Enzymatic degradation of RNA causes widespread protein aggregation in cell and tissue lysates. *EMBO Rep.*, **21**.
57. Kearse,M.G. and Wilusz,J.E. (2017) Non-AUG translation: a new start for protein synthesis in eukaryotes. *Genes Dev.*, **31**, 1717–1731.
58. Todd,P.K., Oh,S.Y., Krans,A., He,F., Sellier,C., Frazer,M., Renoux,A.J., Chen,K., Scaglione,K.M., Basrur,V., *et al.* (2013) CGG Repeat-Associated Translation Mediates Neurodegeneration in Fragile X Tremor Ataxia Syndrome. *Neuron*, **78**, 440–455.
59. Cleary,J.D., Pattamatta,A. and Ranum,L.P.W. (2018) Repeat-associated non-ATG (RAN) translation. *J. Biol. Chem.*, **293**, 16127–16141.
60. 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.
61. Fratta,P., Mizielińska,S., Nicoll,A.J., Zloh,M., Fisher,E.M.C., 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.
62. Reddy,K., Zamiri,B., Stanley,S.Y.R., Macgregor,R.B. and Pearson,C.E. (2013) The Disease-associated r(GGGGCC) Repeat from the C9orf72 Gene Forms Tract Length-dependent Uni- and Multimolecular RNA G-quadruplex Structures. *J. Biol. Chem.*, **288**, 9860–9866.

63. 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.
64. Tabet,R., Schaeffer,L., Freyermuth,F., Jambeau,M., Workman,M., Lee,C.-Z., Lin,C.-C., Jiang,J., Jansen-West,K., Abou-Hamdan,H., *et al.* (2018) CUG initiation and frameshifting enable production of dipeptide repeat proteins from ALS/FTD C9ORF72 transcripts. *Nat. Commun.*, **9**, 152.
65. Baboo,S., Bhushan,B., Jiang,H., Grovenor,C.R.M., Pierre,P., Davis,B.G. and Cook,P.R. (2014) Most Human Proteins Made in Both Nucleus and Cytoplasm Turn Over within Minutes. *PLoS ONE*, **9**, e99346.
66. Iborra,F.J., Jackson,D.A. and Cook,P.R. (2001) Coupled Transcription and Translation Within Nuclei of Mammalian Cells. *Science*, **293**, 1139–1142.
67. Ryu,I. and Kim,Y.K. (2017) Translation initiation mediated by nuclear cap-binding protein complex. *BMB Rep.*, **50**, 186–193.
68. Brogna,S., Sato,T.-A. and Rosbash,M. (2002) Ribosome Components Are Associated with Sites of Transcription. *Mol. Cell*, **10**, 93–104.
69. Al-Jubran,K., Wen,J., Abdullahi,A., Roy Chaudhury,S., Li,M., Ramanathan,P., Matina,A., De,S., Piechocki,K., Rugjee,K.N., *et al.* (2013) Visualization of the joining of ribosomal subunits reveals the presence of 80S ribosomes in the nucleus. *RNA*, **19**, 1669–1683.
70. Bühler,M., Wilkinson,M.F. and Mühlemann,O. (2002) Intranuclear degradation of nonsense codon-containing mRNA. *EMBO Rep.*, **3**, 646–651.
71. David,A., Dolan,B.P., Hickman,H.D., Knowlton,J.J., Clavarino,G., Pierre,P., Bennink,J.R. and Yewdell,J.W. (2012) Nuclear translation visualized by ribosome-bound nascent chain puromycylation. *J. Cell Biol.*, **197**, 45–57.
72. Iborra,F.J., Escargueil,A.E., Kwek,K.Y., Akoulitchev,A. and Cook,P.R. (2004) Molecular cross-talk between the transcription, translation, and nonsense-mediated decay machineries. *J. Cell Sci.*, **117**, 899–906.

FIGURE LEGENDS

Figure 1. NACA controls the translation of GAR-containing mRNAs.

A. Cartoon illustrating constructs. **B.** Western Blot (WB) from H1299 cells expressing EBNA1 or an EBNA1 that lacks the gly-ala repeats domain (EBNA Δ GA_r). **C.** WB from H1299 cells expressing EBNA1 and treated with siRNAs control (-) or against NACA (+). **D.** Relative *EBNA1* mRNA levels as determined by RT-qPCR from H1299 cells expressing EBNA1 and treated with indicated siRNAs. **E.** WB showing GAR-p53 or p53 expression following treatment with indicated siRNAs. **F.** WB of endogenous EBNA1 in EBV-carrying NPC-6661 cells treated with indicated siRNAs. **G.** Antigen presentation assay performed using B3Z SIINFEKL:Kb-specific T cell hybridoma co-cultured with transfected H1299 cells and treated with siRNAs against NACA or SRP. **H.** WB from H1299 cells transfected with an EBNA1 construct carrying the *cMyc* 5'UTR and *HCV* IRES in the 5' UTR (*cMyc_EBNA1* and *HCV_EBNA1* respectively). **I.** WB from H1299 cells expressing indicated constructs and treated with siRNAs. **J.** WB of H1299 expressing the *cMyc_GAr-p53* construct and treated with indicated siRNAs. **K.** WB of H1299 cells transfected with EBNA1 construct and increasing amounts of HA-NACA-encoding plasmids. Data represent three independent experiments.

Figure 2. NACA interacts with the GAR polypeptide

A. Proximity ligation assay (PLA) (white arrows) shows interactions in H1299 cells between NACA and EBNA1 or NACA and EBNA Δ GA_r. Co-immunofluorescence of respective proteins in green. DAPI nuclear staining in blue. Scale bar represents 10 μ m. **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 EBNA Δ GA_r (left). The right graph shows relative PLA signals between endogenous NACA and HA tag in cells expressing HA-tagged p53 and an HA-tagged p53 carrying the PolyQ repeat in the N-terminus. **C.** Relative number of interactions between and exogenous HA-tagged NACA and EBNA1 or EBNA Δ GA_r. **D.** co-IP using anti-EBNA1 or anti-IgG antibodies in EBV-carrying Raji and B95.8 cell lysates. Data represent three independent experiments. A minimum of 50 cells expressing indicated reporter protein was counted for the PLAs.

Figure 3. The nascent GAR dislodges NACA from the ribosome.

A. Ribosomal profile 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) with anti-RPL5(R) 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

Figure 4. The interaction between the GAR mRNA and nucleolin (NCL) is NACA-dependent

A. *In vitro* RNA coIP experiments performed with recombinant NCL and *EBNA1* mRNA extracted from H1299 transfected cells treated with indicated siRNA. 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 indicated siRNA. **C.** *In vitro* RNA coIP experiments performed with recombinant NACA and *EBNA1* or *EBNAΔGAR* mRNA extracted from H1299 transfected cells. The graph shows relative fold change in NACA binding. **D.** RNA pulldown assay using H1299 cell lysates and indicated RNA oligos. GQ-18 is derived from G4-forming sequence coding for the GAR. GM-18 is a mutated non-G4-forming version of GQ-18. **E.** PLA assessing endogenous NACA - NCL interactions in H1299 cells expressing *EBNA1* or *EBNAΔGAR*. EV is for empty vector. In blue: nucleus, in white: PLA dots. Scale bar represents 10µm. **F:** *In vitro* RNA coIP experiments performed with recombinant NCL and NACA proteins and *EBNA1* mRNA extracted from H1299 cells lysates treated, or not, with siRNA against NACA. The graph shows the fold change in NCL binding. NACA promotes NCL binding to the *EBNA1* mRNA. The data represent a minimum of three independent experiments.

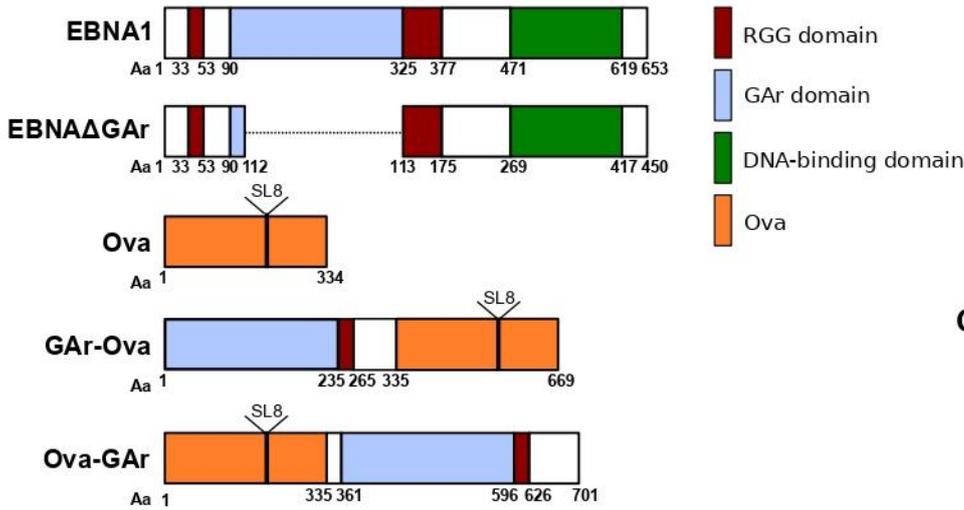
Figure 5. The NCL-GAR mRNA interaction requires translation

A. Top panel shows *in vitro* RNA coIP experiments performed with recombinant NCL and *Ova* and *GAR-Ova* mRNAs extracted from H1299 transfected cells treated with cycloheximide (CHX), harringtonine (Harr) or DMSO control for 5h. The graph shows relative fold change in NCL binding. Bottom panel: WB of H1299 transfected cells treated with CHX, Harr or DMSO. The numbers indicate relative expression levels. **B.** Cartoons illustrating constructs in which the start codons for the *Ova* and the *GAR-Ova* coding sequences were deleted. **C:** Top panel shows *in vitro* RNA coIP performed with recombinant NCL and mRNAs extracted from H1299 cells expressing *Ova*, *GAR-Ova*, *OvaΔATG*

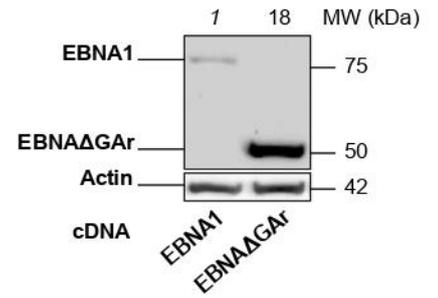
and *GAr-OvaΔATG*. The graph shows relative fold change in NCL binding. Bottom panel shows WB of H1299 transfected expressing indicated constructs or empty vector (EV). **D.** Co-localisation of the *Ova*, *GAr-Ova*, *OvaΔATG* and *GAr-OvaΔnoATG* mRNAs with the nucleus using the mean Costes correlation factor for each condition (n=19 for *GAr-Ova*, 31 for *GAr-OvaΔATG*, 27 for *Ova*, 47 for *OvaΔATG*). A correlation factor of 1 indicates both signals perfectly colocalise. Data presented are from three independent experiments.

Figure 1

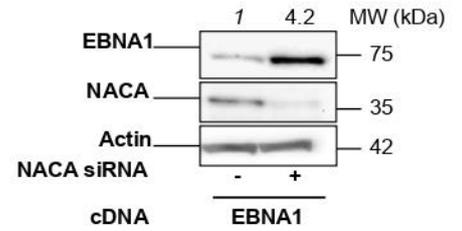
A



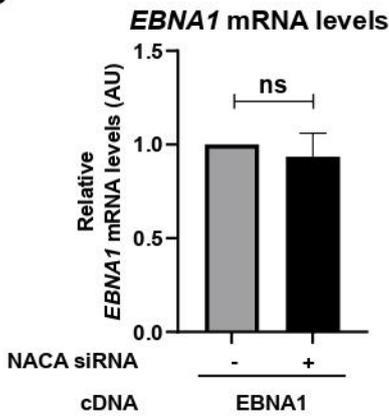
B



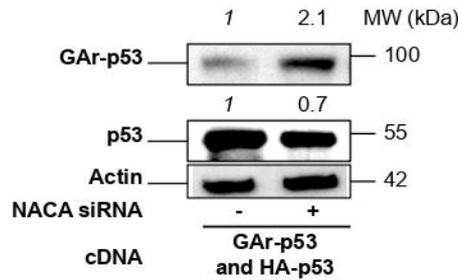
C



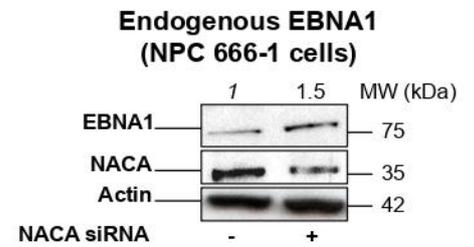
D



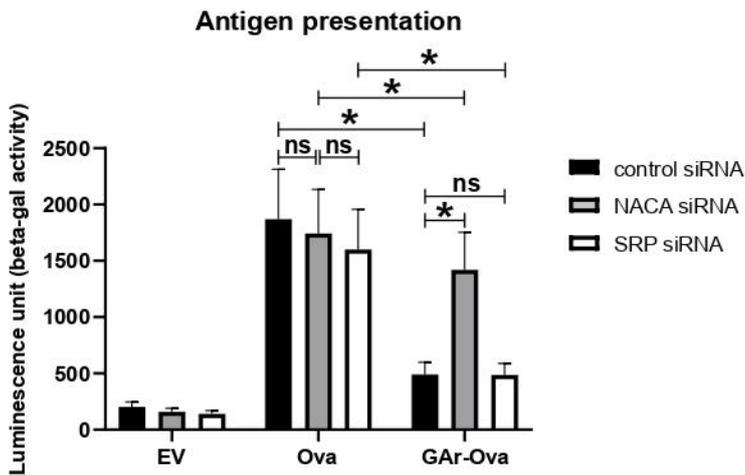
E



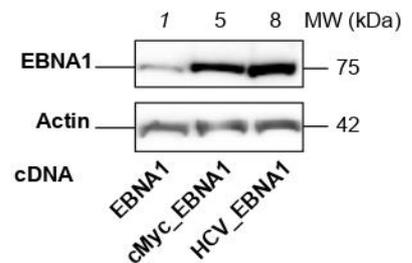
F



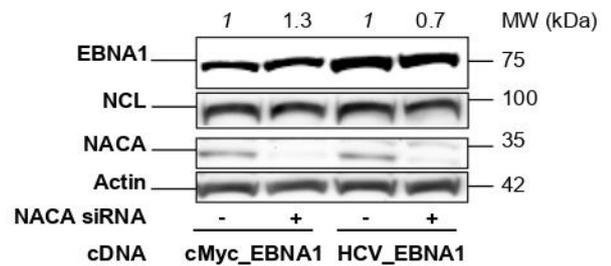
G



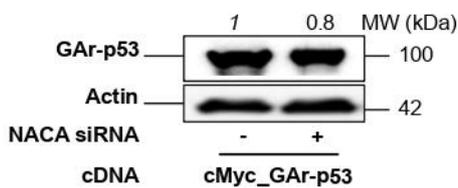
H



I



J



K

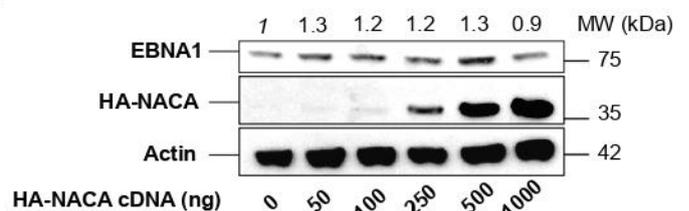
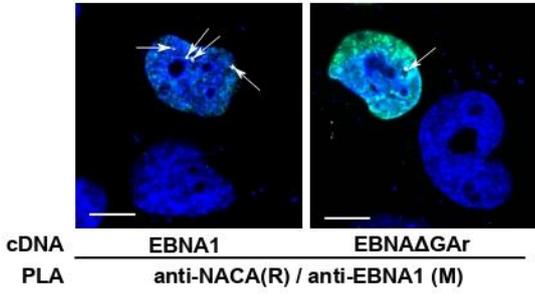
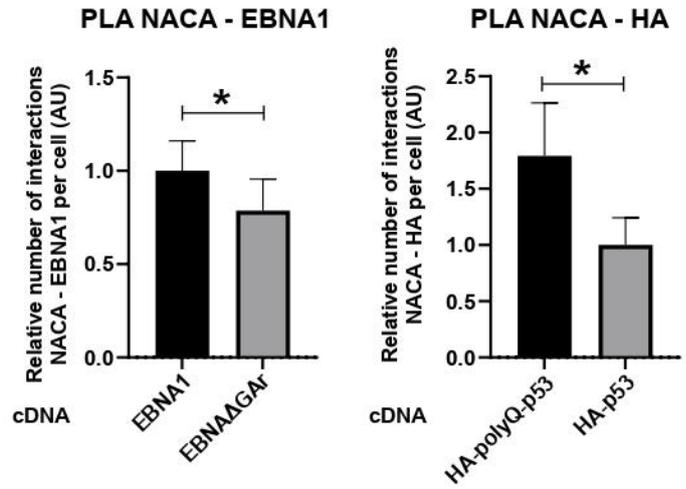


Figure 2

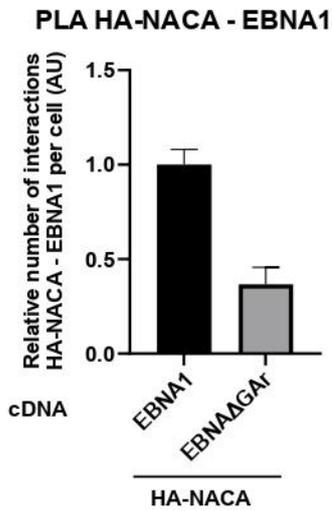
A



B



C



D

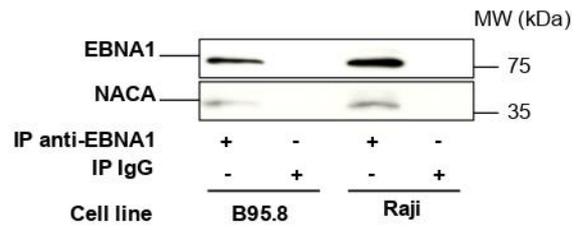


Figure 3

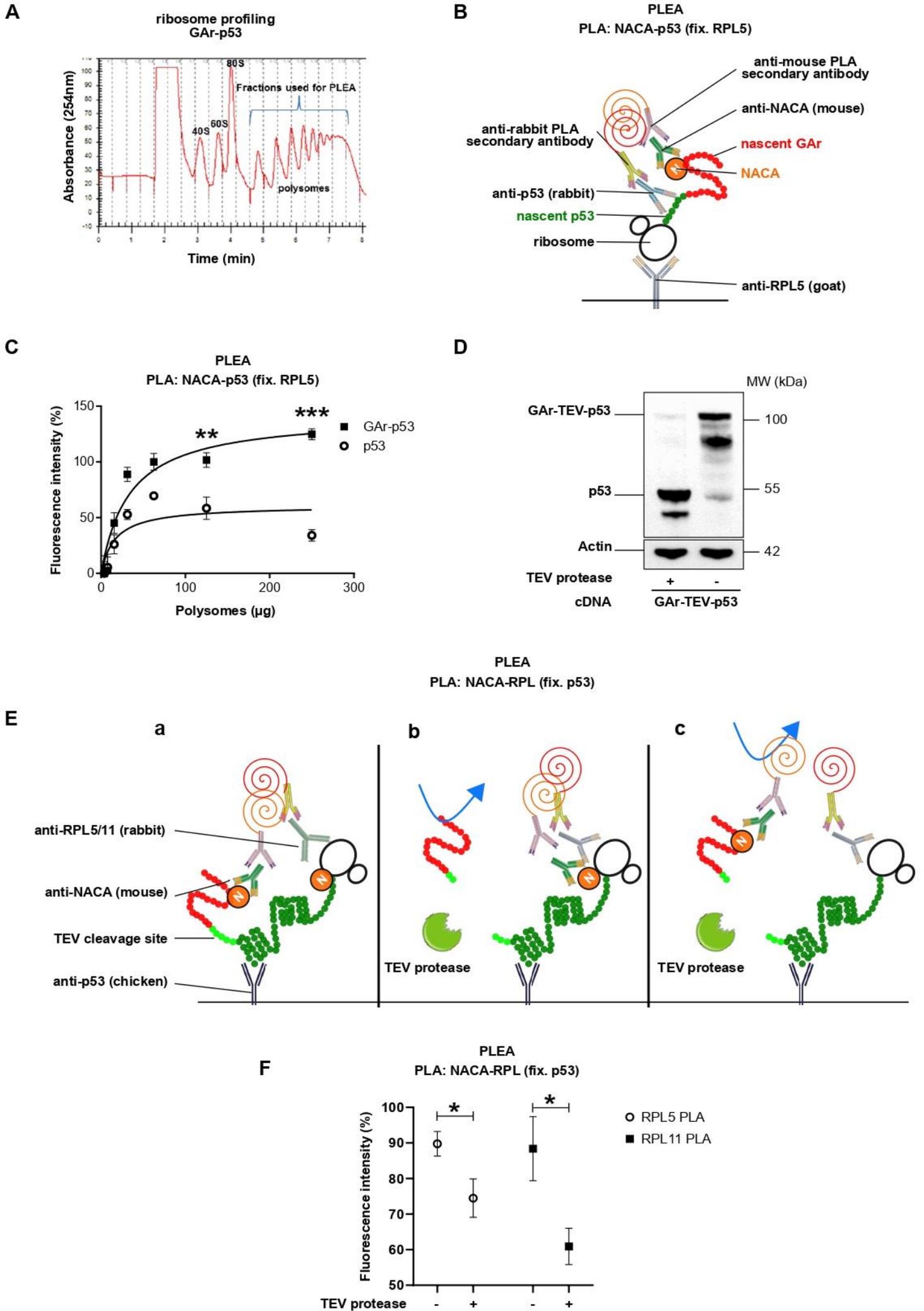
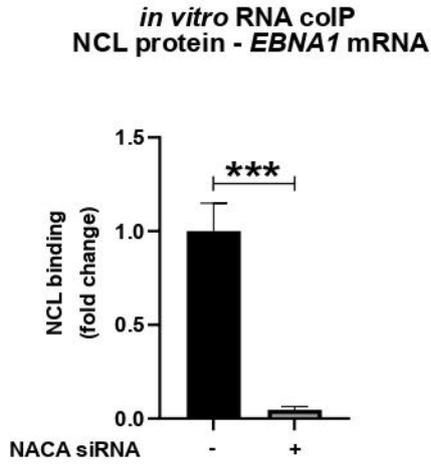
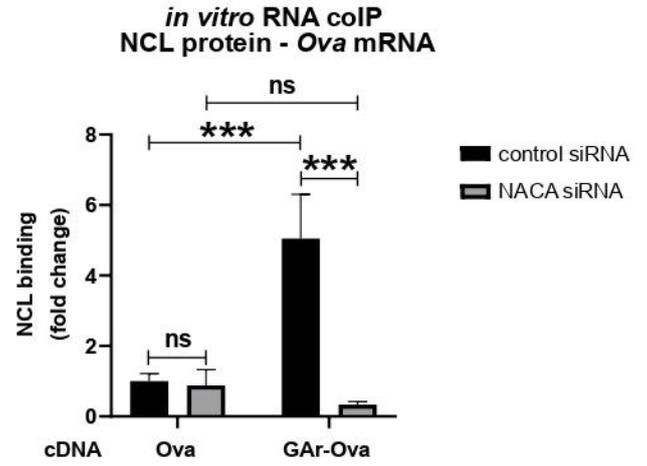


Figure 4

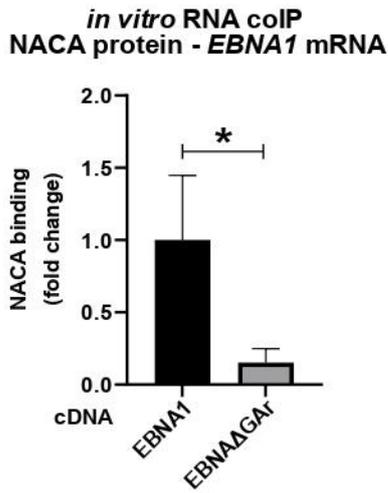
A



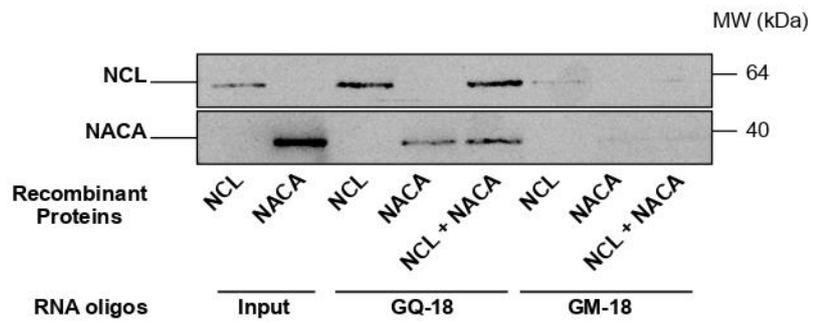
B



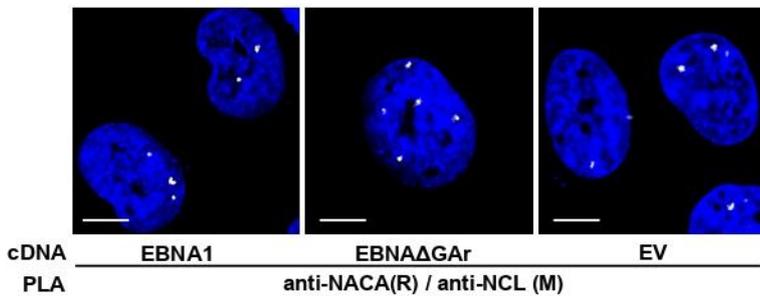
C



D



E



F

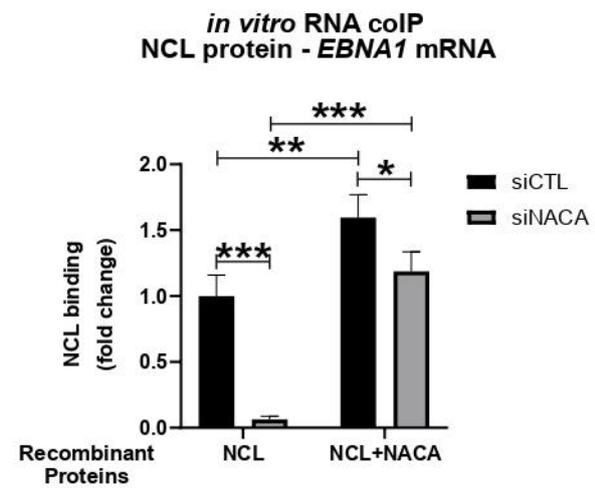
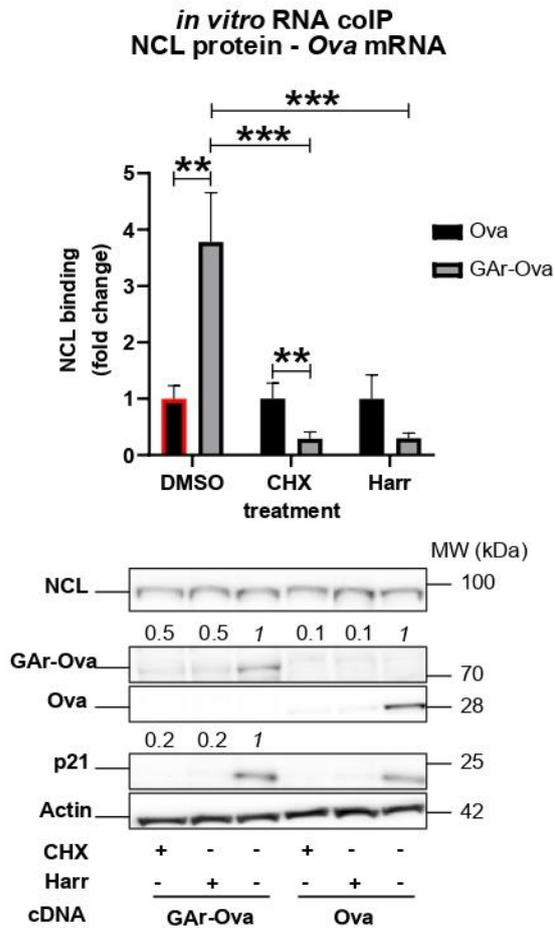
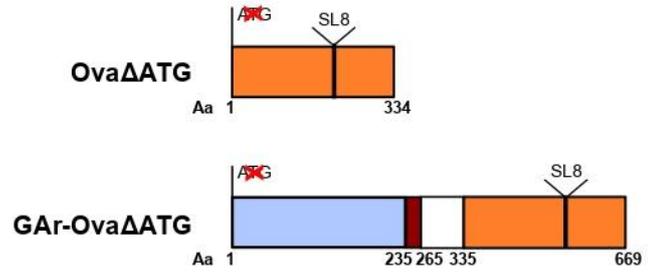


Figure 5

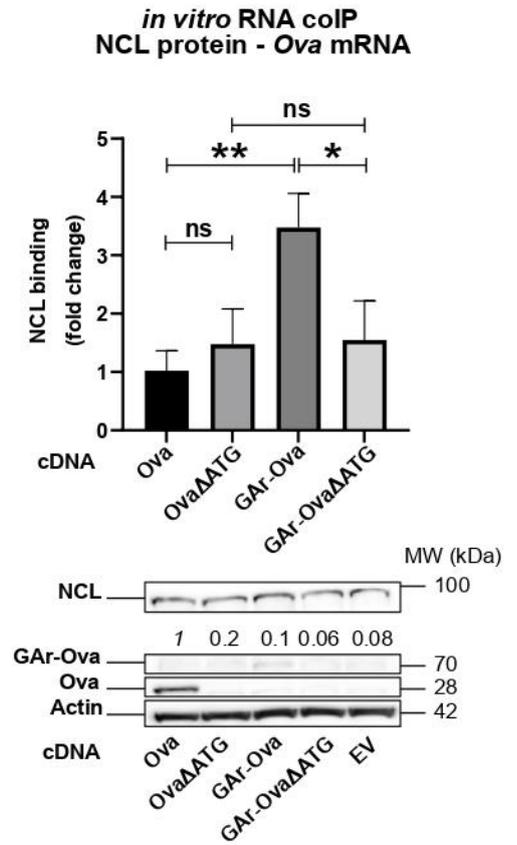
A



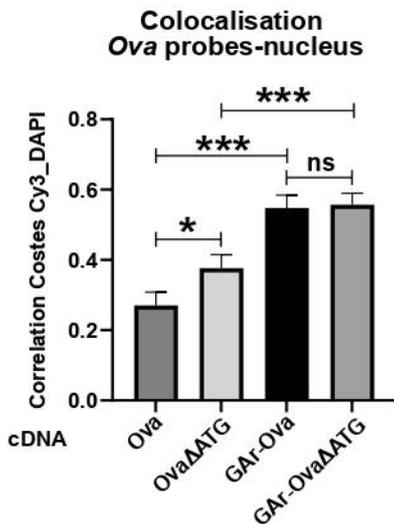
B



C



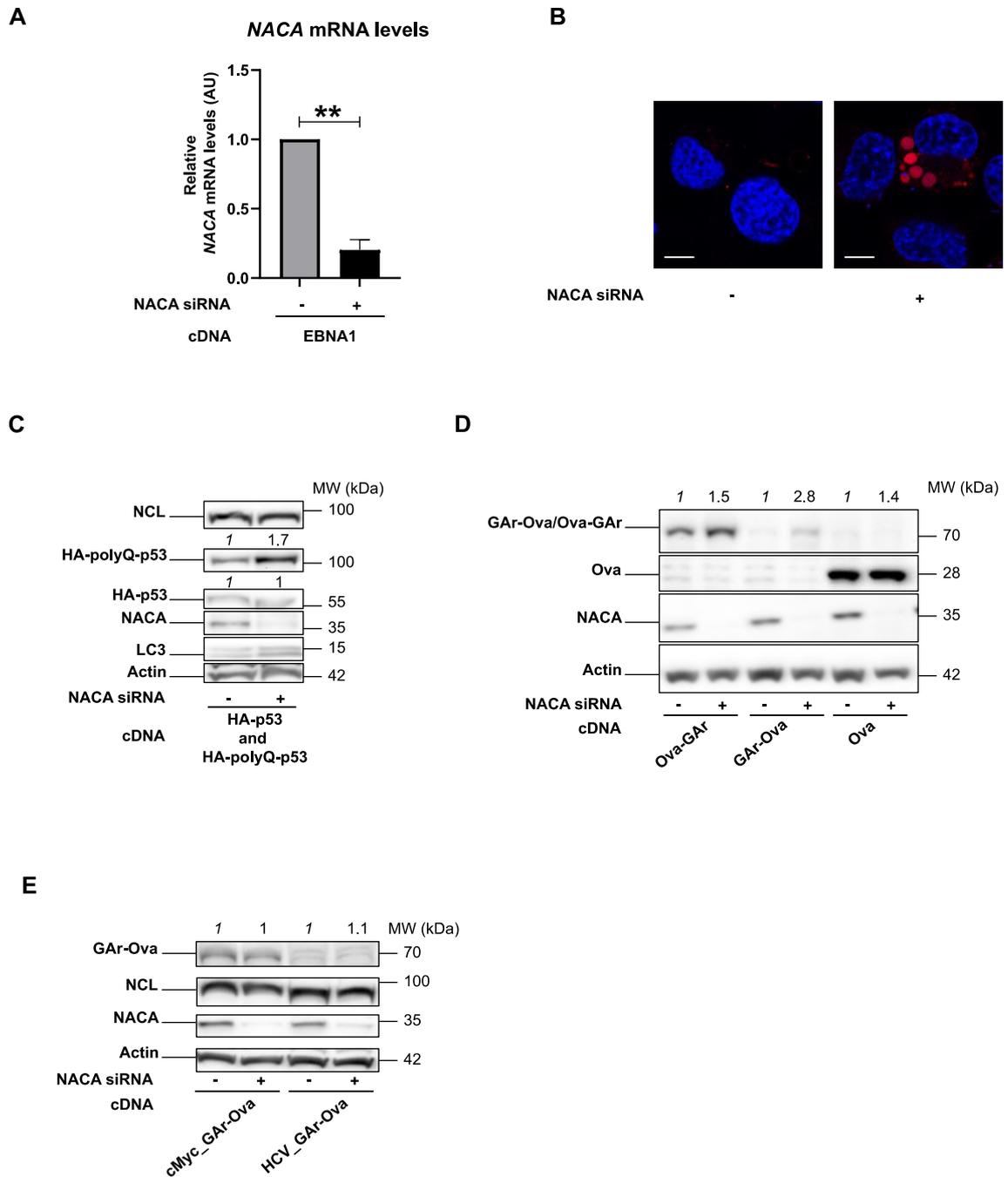
D



Peptide sequence	Length (nucleotides)	Predicted RNA G4s (no overlaps)	Length (amino acids)	Predicted solubility
GAr	708	27	236	0.382
polyQ	381	0	127	0.703
Half GAr	354	14	117	0.357

Table S1: Predicted solubility of the GAr and polyQ polypeptides. Predicted RNA G4 structures in the coding sequence of the mRNAs are identified using QGRS Mapper (71). Predicted solubilities for protein production in *E.coli* are calculated using the SoluProt v1.0 Webserver (72). For predicted solubility, a score below 0.5 indicates insoluble expression in *E.coli*.

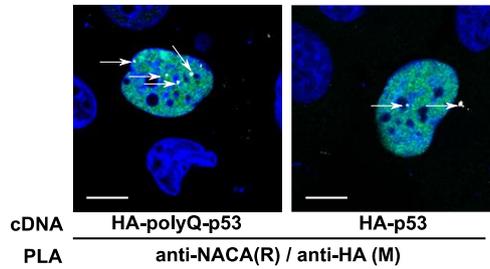
Supp. Figure S1



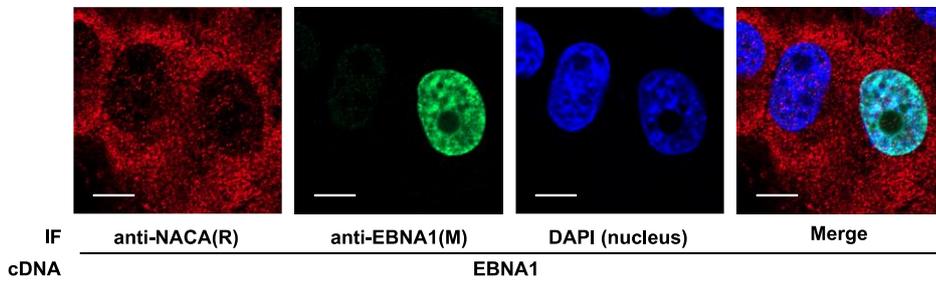
Supp. Fig. S1. A. Relative *NACA* mRNA levels determined by RT-qPCR from H1299 cells transfected with the EBNA1 construct, treated with siRNA control or siRNA against *NACA*. Graph obtained with data from three independent experiments. **B.** Microscopy images obtained with H1299 cells treated with siRNA control or targeting *NACA*, using the proteostat kit. Aggregates are coloured in red, nucleus in blue. Scale bar represents 10 μ m. **C.** Western Blot from H1299 cells co-transfected with HA-polyQ-p53 and HA-p53-encoding plasmids, treated with siRNAs control or targeting *NACA*. **D.** Western Blot from H1299 cells transfected with the indicated constructs, treated with siRNA control or targeting *NACA*. **E.** Western Blot from H1299 cells transfected with the cMyc 5'UTR or HCV IRES fused to the 5' UTR of GAr-Ova and treated with siRNA control or targeting *NACA*. Data relate to Figure 1.

Supp. Figure S2

A



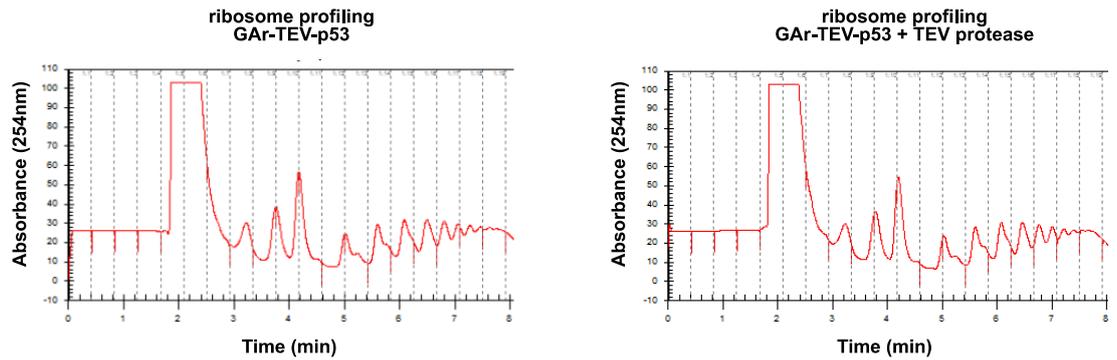
B



Supp. Fig. S2. A. PLA assessing NACA-HA interactions in H1299 cells transfected with HA-polyQ-p53 or HA-p53 constructs, with co-immuno-staining of the reporter proteins. In blue: nucleus, in green: immuno-staining of the reporter proteins. White arrows and dots indicate PLA dots. Scale bar represents 10 μ m. **B.** Immuno-staining of H1299 cells expressing EBNA1 using antibodies against NACA (in red) or EBNA1 (in green). Nuclear staining (DAPI) in blue. Scale bar represents 10 μ m. Data related to Figure 2.

Supp. Figure S3

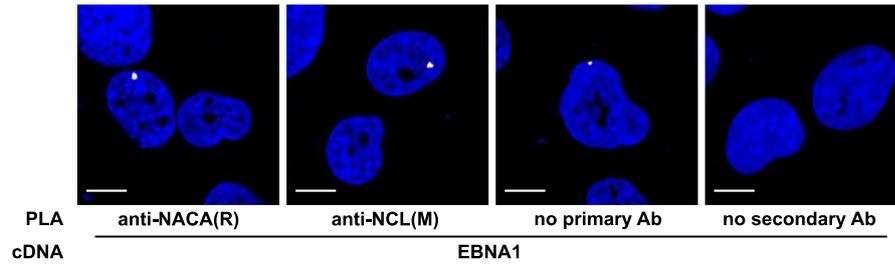
A



Supp. Fig. S3. A. Ribosomal profile of H1299 cells transfected with GAR-TEV-p53 construct and treated, or not, with TEV protease, starting with the free RNA pool at the left and with the polysomes on the right. Related to Figure 3F.

Supp. Figure S4

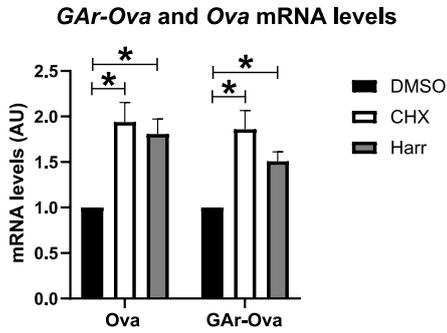
A



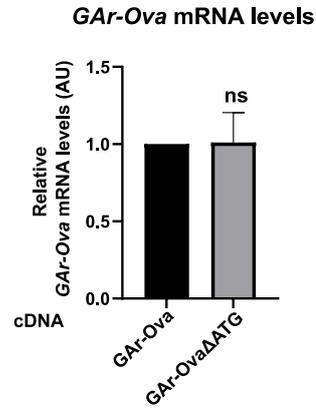
Supp. Fig. S4. A. Control conditions for the NACA - NCL PLA, where only the anti-NAC (1st image) or anti-NCL (2nd image) antibodies were used, or no primary antibody (3rd image) or no secondary antibody (4th image) was used. The figures show the occasional unspecific PLA signal. Related to Figure 4E.

Supp. Figure S5

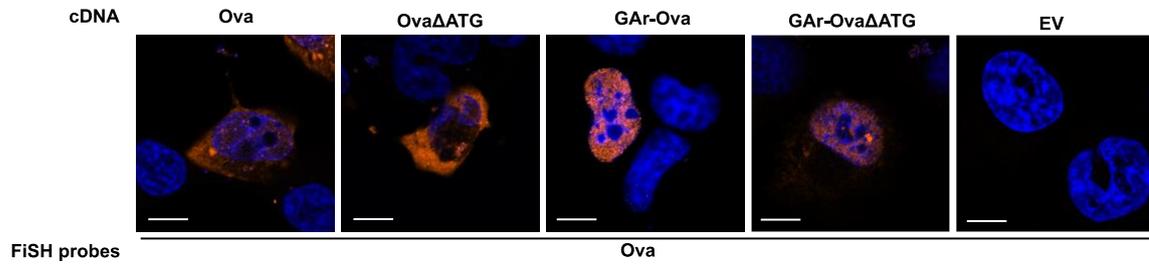
A



B



C



Supp. Fig. S5. A. Relative mRNA levels of *Ova* and *GAr-Ova* mRNAs as determined by RT-qPCR from H1299 cells. Graph shows data from three independent experiments. **B.** Relative mRNA levels of *GAr-Ova* and *GAr-OvaΔATG* (lacks the 1st AUG) mRNAs as determined by RT-qPCR from H1299 cells. Graph shows data from three independent experiments. **C.** RNA FISH using *Ova* primers shows the subcellular localisation of *Ova*, *OvaΔATG*, *GAr-Ova*, *GAr-OvaΔATG* and empty vector (EV). Related to Figure 5.

Cell culture, transfection, drug treatment, and siRNA

H1299 and NPC-6661 cells were seeded in six-well plates at a density of $1.2-1.5 \times 10^5$ cells/well. The following day the cells were co-transfected with a maximum of 1 μ g of total expression plasmid along with 3 μ l of Genejuice according to the manufacturer's protocol (Merck Biosciences).

NACA (GeneSolution siRNA) and negative control siRNAs (AllStars Negative control), called here siCTL, were purchased from Qiagen. Silencing was performed using JetPRIME according to the manufacturer's protocol (Polyplus Transfection). Inhibition of protein synthesis for WB and RNA coIP was performed by treating cells with 10 μ g/ml cycloheximide (EMD Chemicals) or 0.2 μ g/ml harringtonine (LKT Laboratories) for 5 h.

Plasmids construction

All plasmids were generated using standard procedures in DH5 α or TOP10 bacteria strains and cloned in the pCDNA3.1 or EGFP vectors. Restriction enzymes, T4 DNA ligase, and calf intestinal alkaline phosphatase were obtained from New England Biolabs. Purified synthetic oligonucleotides were obtained either from Eurogentec or Invitrogen/ThermoFisher. The EBNA1, EBNA Δ GA, p53, GAR-p53, Ovalbumin, GAR-Ovalbumin were described previously (34, 35, 69, 70).

cMyc-GAR-p53 was generated by amplification of full-length human *c-Myc* 5'UTR by PCR, using the 5' sense primer 5'-**CGGATCC**ACTAGAACTCGCTGTAGTAATTC-3' (restriction sites are in bold) and 3' antisense primer 5'-TCC**GGATCC**CGGGAGGCTGCTGG-3' and cloned into the 5' UTR of GAR-p53. The plasmids cMyc_Ova, cMyc_GAR-Ova, HCV_Ova, and HCV_GAR-Ova encode Ovalbumin or a fusion of the GAR domain to the N-terminus of Ovalbumin, with either the *c-Myc* 5'UTR or the *HCV* IRES in their 5'UTR. *HCV* IRES and *c-Myc* 5'UTR were amplified from these constructions and fused to the 5'UTR of EBNA1 and coding sequence to produce the c-Myc_EBNA1 and HCV_EBNA1 constructs, using for both the sense primer 5'-AATAAGCTTCCACTGCTTACTGGCTTATCG-3', and reverse primers 5'-TAAAAGCTTCGGCCGTTACTAGTGGATCC-3' and 5'-TAAAAGCTTGGAGGGGCAAACAACAGATG-3' for *c-Myc* 5'UTR and *HCV* IRES, respectively. NAC was generated by amplification of human NAC from HEK 293T cell total cDNA by PCR using the 5' sense primer 5'-GCGCG**GGATCC**ATGCCGGCGAAGCCACAGAAACC-3' and the 3' antisense primer 5'-GCGCG**GAATTC**TTACATTGTTAATTCATAATCGC-3'. The amplification was then cloned in the pCDNA3.1 vector. The HA-NAC construct was generated by cloning the NAC construct in the HindIII-EcoRI sites of a pCDNA3.1-HA tag vector. The GAR-TEV-p53 construct was generated by annealing oligonucleotides 5'-P-AATTCGAGAACCTGTACTTTTCAGGGCGGATTCG-3' and 5'-P-AATTCGAATCCGCCCTGAAAGTACAGTTCTCG-3' and then ligating them in the EcoRI site of the GAR-p53 construct. For the Ova Δ ATG and GAR-Ova Δ ATG constructs, site-directed mutagenesis was performed using the QuikChange™ Site-Directed Mutagenesis System (Stratagene) with the Ova and Ova-GAR constructs as templates, followed by digestion ligation in the GAR-Ova or Ova plasmids.

Western Blotting

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.

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

The following primers (custom primers by Invitrogen) were used: Actin Forward 5' - TCACCCACACTGTGCCCATCTACGA- 3', Actin Reverse 5' - TGAGGTAGTCAGTCAGGTCCCG - 3', Ova Forward 5'-GCAAACCTGTGCAGATGATG-3', Ova Reverse 5'- CTGCTCAAGGCCTGAGACTT-3', EBNA1 Forward 5'-GGCAGTGGACCTCAAAGAAGAG-3', EBNA1 Reverse 5'-

CAATGCAACTTGGACGTTTTTG-3', NACA Forward 5'- GGCTGAGACAGGGTCTGG-3', NACA Reverse 5'- GACATAGCCTTCCGTGCC-3'.

In vitro RNA co-IP was carried out as described elsewhere (45). 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 (50mM Tris pH7.5, 150mM NaCl, 0.02mg/mL yeast tRNA, 0.2mg/mL BSA) for 15min 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 antibodies (Abcam, ab22758) according to standard conditions and purified using the TRIzol (Life Technologies). Precipitated RNAs were then analysed by RT-qPCR.

RNA FISH

H1299 cells were seeded in 24-well plates (2x10⁴cells/well) and transient transfections were carried out 20-24h later using the Genejuice reagent (Merck Bioscience) according to the manufacturer's protocol. 24h after transfection or after drug treatment, cells were briefly washed with ice-cold PBS, fixed with 4% PFA for 20 min at room temperature, and washed again with PBS. Cells were then incubated in 70% Ethanol for 4-24 h at 4°C, after intermediate dehydrating steps using 30% and 50% Ethanol. For rehydration, cells were incubated in 50% and 30% Ethanol and further washed with PBS. Subsequently, cells were permeabilised with PBS 0.4% Triton 0.05% CHAPS for 5 min at room temperature. After three PBS washes, samples were pre-treated with two subsequent incubations in FiSH wash buffer (10% formamide and 2X SSC in ddH₂O) for 10 min and then in FiSH hybridisation buffer (10% formamide, 2X SSC, 2mg/mL BSA, 0.2mg/mL *E.coli* tRNA, 0.2mg/mL sheared salmon sperm DNA) for 30min at room temperature. Coverslips were then 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 20min in FiSH hybridisation buffer and 5 min in FiSH Wash buffer and subsequently stained with DAPI. Samples were mounted using Dako mounting solution (Dako) and observed with the LSM 800 (Zeiss) confocal laser microscope. 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 (82), and later used for statistical analysis.

Polysome fractionation and PLEA

5-50% wt/vol linear sucrose gradients were freshly cast on SW41 ultracentrifuge tubes (Beckmann) using the Gradient master (BioComp instruments) following the manufacturer's instructions. 48h post-transfection, H1299 cells (with 80% confluency) were treated with Cycloheximide 100 µg/ml for 5 minutes at 37°C and then washed twice with 1x PBS (Dulbecco modified PBS, GIBCO) containing cycloheximide 100µg/ml. Cells were then scrapped and lysed with polysome lysis buffer (KCl 100 mM, HEPES-50 mM, 5 mM MgCl₂, 0.1% NP-40, 1 mM DTT, Cycloheximide 100 µg/ml, pH 7.4) using narrow gauge syringes. Lysates were then treated with AcProtease (Invitrogen) for 1h at 16°C and spun down at 5000 rpm for 10 minutes at 4 °C. Lysates were then loaded on a sucrose gradient and centrifuged at 36000 rpm for 2h at 4°C in an SW41 rotor. Samples were then fractionated using a Foxy R1 fraction collector (Teledyne ISCO) at 0.5 min intervals. Polysome fractions were collected and concentrated down 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.

The PLEA technique is used to study the interaction of three molecules. The capture antibody is used to capture the complex in the well plate and a set of two additional antibodies is used for the amplification of the PLA signal. As a capture antibody here we used either a chicken anti-p53 antibody (home-made, against full-length p53wt, Biosan (Uruguay) and Prof. Mónica Marín, Biochemistry-Molecular Biology, Facultad de Ciencias, Universidad de la República, Montevideo, Uruguay) or a goat anti-RPL5 antibody (Santa Cruz). Antibodies used for the PLA amplification were a mouse anti-NAC antibody (Abnova), a rabbit anti-p53 antibody (CM-1), a rabbit

anti-RPL5 antibody (Invitrogen), and a rabbit anti-RPL11 antibody (Santa Cruz). 96-well ELISA plates were incubated with the capture antibody at a dilution of 1:200, o/n, at 4 °C. Then the plates were blocked with 5% BSA in PBS, o/n, at 4°C, and were incubated with serial dilutions of the polysomal fractions in 3% BSA in PBST, o/n at 4°C. The plates were then washed six times with PBST (0.1 M tween in PBS) and a set of primary antibodies developed in rabbit and mouse was used at a dilution of 1:200, for 2 h, at RT. After washing six times with PBST, samples were incubated with the PLA secondary antibodies and the PLA kit (Sigma) was used, following the manufacturers' instructions, with modifications. Samples were incubated with the set of the detection secondary PLA antibodies (anti-mouse and anti-rabbit) for 1h at 37 °C. Excess of antibodies was washed out six times with PBST and samples were incubated with 0.5 µl ligase (PLA Far RED kit, Sigma) for 30min at 37 °C and washed six times with PBST. Samples were incubated with 0.25 µl Φ29 DNA polymerase in amplification buffer (PLA Far RED kit, Sigma) for 2 h, at 37 °C, and washed 6 times with PBST. Each sample was tested in triplicates. 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.

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 1X protease inhibitor cocktail (Roche). Cell lysis was performed by five series of vortex followed by 10 min incubation on ice, and three series of 3 s 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, ARPC2- 5' AGCCGGGGCUGGGCGGGGACCGGGCUUGU-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, 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. 500 µg 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). Proteins still bound to beads after the washes were eluted using 2X SDS loading buffer and analysed by western blot against NACA or NCL, as previously described. In the input lane of the western blots was loaded a quantity of extract which corresponds to the quantity that was incubated with the beads for each condition.

Paper 3: eIF4A1 impacts mRNA
translation via remodelling RNP
complex composition

eIF4A1 as a negative regulator of *EBNA1* mRNA translation via RNA G4 structures

Alice J.-L. Zheng¹, Van Trang Dinh^{2,5}, Ronan Le Sénéchal^{2,5}, Laurence Malbert-Colas^{1,5}, Alicia Quillévéré^{2,5}, Chrysoula Daskalogianni^{1,3}, Marc Blondel² and Robin Fahraeus^{1,3,4*}

¹Inserm UMR51131, Institut de Génétique Moléculaire, Université Paris 7, Hôpital St. Louis, F-75010 Paris, France

²Inserm UMR1078, Université de Bretagne Occidentale (UBO), Etablissement Français du Sang (EFS) Bretagne, CHRU Brest, 29200, Brest, France

³ICCVS, University of Gdańsk, Science, ul. Wita Stwosza 63, 80-308 Gdańsk, Poland

⁴Department of Medical Biosciences, Building 6M, Umeå University, 901 85 Umeå, Sweden

⁵These authors contributed equally

*To whom correspondence should be addressed

Running title: eIF4A1 controls RNP complex

Keywords: DEAD-box RNA helicase, EBNA1, RBP, RNA structure, mRNA translation

ABSTRACT

RNA-RBP (RNA-binding protein) interactions control many aspects of an mRNA life, including its maturation, stability, and translation. The resulting RNP (ribonucleoprotein) complexes are dynamic, and their formation and stability depend on many cellular factors. RNA helicases, and particularly DEAD-box RNA helicases, were proposed to regulate RNP complex formation, by either modulating the RNA structures or by directly interacting with the RBPs. Here, we show that eIF4A1, a DEAD-box RNA helicase mostly known for being a component of the translation initiation eIF4F complex, has a role in the translation regulation of the mRNA by controlling RNP formation. The glycine-alanine repeats (GAR) domain of EBV-encoded EBNA1 protein is encoded by an RNA sequence forming RNA G4 structures *in vitro*, and the binding of the nuclear factor nucleolin (NCL) on the G4-forming sequence of the mRNA in the cells leads to the inhibition of the mRNA translation. Silencing eIF4A1 or treating the cells with hippuristanol suppresses this translation inhibition, whereas silvestrol treatment enhances this inhibition. Furthermore, NCL binding on GAR-encoding mRNAs is suppressed by eIF4A1 silencing or the use of eIF4A1 inhibitors. We also show that eIF4A1 binds specifically to RNA G4 structures derived from the GAR coding sequence and that its action on mRNA translation is supported by NACA but does not rely on it, and *vice versa*. Taken together, these results suggest a central role of RNA G4 structures dynamics in the formation of RNP, and we suggest that eIF4A1, by binding and modulating the G4 structures, controls the RNP and *EBNA1* mRNA translation.

INTRODUCTION

RNP (ribonucleoprotein) complexes play an essential role in the cell, and the interactions between RNA and RBP (RNA-binding proteins) are context-dependent. Indeed, RNA-RBP interactions are dynamic, modelling RNP existence and composition, and can adapt to the cell status (1–3). These interactions have an important influence on mRNA translation. For example, MDM2 binding to *TP53* mRNA promotes *TP53* mRNA translation (4, 5), and the classical initiation factor eIF4F is a complex of proteins binding to mRNA, enhancing its translation (6).

Chemical modifications and RNA structure can guide the binding of RBPs on the mRNA (7). For example, N⁶-methyladenosine (m6A) modifications in the 5'UTR can mediate cap-independent translation initiation, with direct interaction between eIF3 and m6A (6, 7). mRNA structure has been widely studied in the context of translation initiation. IRESs (Internal Ribosome Entry Sites) are RNA structures located in the 5'UTR of a cellular or viral mRNA, promoting translation either by recruiting directly the translational machinery or doing so with the help of ITAFs (IRES *trans*-acting factors) (8). The importance of the RNA structure in RBP interaction and translation regulation was demonstrated with different examples, among which is the translation regulation of the *TP53* mRNA via MDM2 interaction. As stated earlier, MDM2 binding to *TP53* mRNA promotes *TP53* mRNA translation, but silence mutations that affect the RNA structure but not the polypeptide sequence lead to weaker interaction of *TP53* mRNA with MDM2, thus limiting p53 synthesis (4, 5).

RNA G4 (G-quadruplex) structures are non-canonical secondary structures formed by G-rich nucleic acid molecules, consisting in a stable four-stranded conformation of DNA or RNA, where G-quartets (also called G-tetrads) are stacked on each other and stabilised thanks to Hoogsteen hydrogen bonds and the presence of a monovalent cation (9–12). Numerous functions in RNA processing have been associated with these structures, like mRNA maturation, localisation, and translation. They have been proposed to serve as roadblocks when localised in the 5'UTR or the CDS of the mRNA, inducing translational pause leading to abortion of the translation or ribosomal frameshifting (13–17). In the case of *NRAS* and *KRAS* mRNAs, both coding for oncogenic factors, RNA G4-forming sequence in their 5'UTR are associated with translation inhibition (18, 19). However in some cases, RNA G4-forming sequences have also been connected with mRNA translation promotion, as observed in the case of the *RON* mRNA where hnRNP A1 binding to G4-forming sequence in the 5'UTR of the mRNA activates its translation (20), or in Aven-targeted mRNAs like *MLL1* and *MLL4* mRNAs (21).

DEAD-box RNA helicases are a family of proteins with an ATP-dependent RNA helicase and RNA-dependent ATPase activity. They have been associated with virtually all RNA processing steps. DEAD-box proteins all have a helicase core which is conserved throughout the family, composed of two domains (called domains 1 and 2), which are similar to the bacterial recombination protein recombinase A (RecA). Spread in the helicase core, twelve characteristic sequence motifs are present at conserved positions, the motif II being the DEAD motif which gives its name to the family. The N and C-terminus domains flanking the helicase core are variable among the DEAD-box helicases, and play a role in binding partner recognition and the specificity of each DEAD-box helicase activity (22, 23).

The classic model of translation initiation requires the presence of the eIF4F complex on the mRNA (8). The eIF4F complex is composed of eIF4E, a cap-binding protein, eIF4G, a scaffolding protein, and eIF4A (either eIF4A1 or eIF4A2), which is a DEAD-box RNA helicase (8). Initiation is a highly regulated step of translation, and many drugs target it. Silvestrol and hippuristanol both inhibit eIF4A activity, the first transforming eIF4A in a clamp on the mRNA (24), and the latter by inhibiting eIF4A ATP-binding activity, decreasing this way eIF4A affinity for RNA (25, 26). Using these drugs, different studies suggest that eIF4A is required for translation of mRNAs with long and highly structured 5'UTR or with RNA G4-forming sequence in their 5'UTR (27–29). This indicates that eIF4A is required for unwinding structures during the scanning step of initiation, supporting the small ribosomal subunit progress on the mRNA. However, without its co-activator eIF4B and eIF4H, eIF4A is a non-processive helicase, meaning it can unwind locally the structure alone, but would not translocate and unwind further structures on the mRNA without being in association with other initiation factors (30–33).

In this study, we use the *EBNA1* mRNA translation regulation as our model. EBNA1 is the genome maintenance protein of the Epstein-Barr virus, which infections are in the majority of the case latent (34). As EBNA1 is required for viral genome replication and proper segregation in daughter cells during cell division, its presence, at least during cell division, is indispensable for EBV survival (35). However, EBNA1 is antigenic, and individuals who are carrying this virus have CD8+ and CD4+ T cells targeting EBNA1-derived antigens (36–40). Despite the presence of a primed immune system against it, EBV escapes the host immune surveillance, thanks to low levels of both EBNA1 proteins and antigens in infected cells (41, 42). Thus, EBNA1 is using a mechanism that tightly regulates its production, enabling the virus to reach a high enough level of proteins when needed for genome maintenance but keeping it low, escaping this way the host immune system. EBNA1 protein contains a glycine-alanine repeats domain, responsible for the inhibition of its own mRNA translation

initiation *in cis* (43, 44). Previous studies show that the RNA G4s predicted in the GAR-encoding mRNA are folding into G4s *in vitro* and are necessary for the inhibition pathway (45, 46). The cellular factor nucleolin (NCL) binds to these structures and treating cells with G4 ligands, like PhenDC3 and PhenDH2, prevents NCL interaction with the GAR-encoding mRNA, leading to an increase in both EBNA1 full-length proteins and antigenic peptides production (46). In addition, previous studies show that the EBNA1 protein binds both its mRNA and NCL, thus forming an RNP complex which is required for viral genome replication and proper segregation (47–49).

Here we show that eIF4A1 is involved in the GAR-mediated inhibition of mRNA translation *in cis* and that it also controls the mRNA capacity to bind NCL. In addition, eIF4A1 can bind directly to RNA G4s, in a specific manner and competes against the G4 ligand PhenDH2 for RNA G4 binding *in vitro*. We also provide first evidence that NACA and eIF4A1 acts similarly on *EBNA1* mRNA NCL-binding capacity but do not rely on each other to inhibit translation. Finally, we show that RNA structures, and particularly the G4s, impact NACA interaction with the GAR polypeptide of EBNA1. We propose a model in which eIF4A1 modifies RNA G4 structures in the GAR-encoding sequences, therefore allowing the GAR-mediated inhibition of translation via NCL binding.

RESULTS

eIF4A1 inhibits EBNA1 mRNA translation

eIF4A has been proposed to unwind RNA G4 structures with two G-quartets (27), and as the majority of predicted RNA G4s in the GAR-encoding mRNA of EBNA1 are with two G-quartets (45), we checked if eIF4A could be involved in the GAR-mediated inhibition of translation initiation, in which RNA G4 structures play a central role (45, 46). eIF4A1 silencing increases by almost 7-fold the steady-state level of EBNA1 protein detected by Western Blot, whereas eIF4A2 silencing only has a mild effect on EBNA1 expression and eIF4A3 silencing leads to a further decrease in EBNA1 detection (**Fig. 1A**). The effectiveness of our siRNAs against eIF4A1 was further confirmed by checking mRNA levels (**Fig. 1B, left panel**). Additionally, silencing eIF4A1 did not have any effect on *EBNA1* mRNA levels, meaning eIF4A1 impacts EBNA1 protein production at the translational level (**Fig. 1B, right panel**). Treating the cells with hippuristanol, a molecule preventing eIF4A1 binding to RNA, relieved *EBNA1* mRNA translation inhibition in a dose-dependent and GAR-dependent manner (**Fig. 1C**). Treating the cells with silvestrol, a chemical promoting eIF4A1 clamping onto RNA, further inhibited *EBNA1* mRNA translation in a dose-dependent way, but not in a GAR-dependent fashion (**Fig. 1D**).

However, of note, both molecules increase *EBNA1* mRNA levels (**Supp. Fig. S1A**). Taken together, these experiments show that eIF4A1 is inhibiting *EBNA1* mRNA translation.

We wanted to see if other translation initiation factors could have an impact on *EBNA1* mRNA translation regulation. We observed that silencing either eIF4G1 or eIF4G2 relieves at certain concentrations mRNA translation inhibition (**Supp. Fig. S1B and S1C**).

eIF4A1 control GAR-encoding mRNAs binding capacity to NCL and binds specifically to RNA G4s derived from the GAR-encoding sequence

The GAR-mediated inhibition of translation initiation involves the interaction between the RBP NCL and RNA G4 structures in the GAR-encoding mRNA. As eIF4A1 can unwind RNA G4s, we wanted to know if it can have an impact on the binding capacity of the mRNA towards NCL. To test this, we extracted mRNAs from cells treated with siRNA against eIF4A1 or eIF4A3 and performed an RNA coIP with recombinant NCL. We observed that eIF4A1 silencing decreases by three-fold the NCL binding onto *EBNA1* mRNA (**Fig. 2A**). Treating cells with 50nM hippuristanol before mRNA extraction reduces by two the mRNA capacity to bind recombinant NCL (**Fig. 2B**). Moreover, treating cells with 50nM silvestrol also induces a reduction of mRNA binding capacity to NCL, although milder than with hippuristanol (**Fig. 2C**). These different experiments suggest that eIF4A1 is regulating the *EBNA1* mRNA binding capacity to NCL.

NCL is binding directly to *EBNA1* mRNA G4 structures (46). To see if eIF4A1 does too, we performed RNA pulldown experiments where H1299 cell lysates are put in contact with different RNA oligos. GQ-18 oligos are derived from the GAR-encoding sequence and are predicted to fold into RNA G4 structures, whereas GM-18 is its mutated version without the predicted RNA G4 folding capacity (46). eIF4A1 binds to the RNA G4-forming oligo GQ-18, but not to the mutated version GM-18. Furthermore, this interaction is competed by the G4 ligand PhenDH2 (**Fig. 2D**). Although eIF4A1 is mainly described as a cytoplasmic factor, it can also be detected in the nucleus (**Supp. Fig. S2A**). Taken together, this indicates that eIF4A1 binding to the RNA is G4-dependent, at least in the context of the GAR.

NACA and eIF4A1 collaborate in the EBNA1 mRNA translation inhibition pathway, but do not rely on each other

It was previously shown that NACA binds the GAR polypeptide, an event that is required for NCL binding on the GAR-encoding mRNA and translation inhibition. In addition, adding IRES or treating the cells with G4 ligands affect the G4-associated functions of the GAR-encoding mRNA, like

NCL binding and mRNA localisation and translation. All of this indicates a key role of RNA G4 structures in the GAR-mediated inhibition of translation. As eIF4A1 has been described as an RNA G4 structure modulator, we wanted to see if silencing both NACA and eIF4A could have an impact on EBNA1 expression. As expected, silencing either eIF4A1 or NACA increased EBNA1 steady-state levels in H1299 transfected cells, but silencing both eIF4A1 and NACA together further increased EBNA1 expression in cells (**Fig. 3A**). This supports an idea where eIF4A1 and NACA both are negative regulators of *EBNA1* mRNA translation which can collaborate in a same pathway, but also have an impact on EBNA1 expression independently.

RNA G4 structures stability impacts NACA-EBNA1 protein interaction

We wanted to further test the impact of RNA G4 structures on RNP complex formation. For this, we first tested how affecting RNA G4s can affect NACA-EBNA1 interactions. Both *c-Myc* IRES and *HCV* IRES-mediated translation are not affected by the GAR-mediated inhibition of translation initiation. By performing PLA to detect EBNA1-NACA interactions in cells, we observed that *HCV* IRES, but not *c-Myc* IRES, decreases the number of interactions (**Fig. 3B**). In parallel, using the FISH technique, we saw that mRNA encoding mRNA without IRES or with the *c-Myc* IRES in the 5'UTR are nuclear, whereas the ones with the *HCV* IRES in the 5'UTR are also cytoplasmic (**Supp. Fig. S3A**). We wonder if the presence of mRNA is required for the stabilisation of the NACA-EBNA1 interactions, as it is for the NCL-EBNA1 interactions (49). To further test this hypothesis, we treated the transfected cells with G4-stabilising ligands and performed PLA. PhenDC3 and PhenDH2 had different effects regarding their impact on NACA-EBNA1 interactions. While PhenDC3 increases by almost two-fold the number of detected NACA-EBNA1 interactions, treating with PhenDH2 did not influence it (**Fig. 3C**). To see how specific this effect is, we repeat the experiment with the polyQ polypeptide, which is known to interact with NAC (50). G4 ligands did not have any effect on the polyQ interaction with NACA (**Supp. Fig. S3B**).

We also started to study the impact of eIF4A1 on RNP formation. Treating the cells with eIF4A inhibitors did not have any effect on NCL-EBNA1 interactions, but this conclusion will need to be further assessed (**Supp. Fig. S3C**).

DISCUSSION

eIF4A1, the RNA helicase of the eIF4F complex, has mainly been studied for its role as an initiation factor, but other cellular functions might be attributed to it (22, 23). In the context of

EBNA1, we found that it is part of the GAR-mediated translation inhibition of mRNA *in cis*. eIF4A1-dependent translation mainly takes care of mRNAs with long and structured 5'UTR (27, 28). However, it is a non-processive helicase and needs either eIF4B or eIF4H to be efficient in the eIF4F complex (31, 33). The data presented here does not exclude eIF4A1 to be an effector of the inhibition pathway triggered by the GAR at the initiation step, and therefore determining if the eIF4A1 co-factors are critical for the inhibition of *EBNA1* mRNA translation could help to sort this question. eIF4G1 and eIF4G2 silencing also cancel *EBNA1* mRNA translation inhibition, but as these factors have been identified to partner with eIF4A1 in other context than as part of a translation initiation complex, it remains still to be determined if eIF4A1, eIF4G1 and eIF4G2 are active solely during the translation initiation phase in this case (56). Moreover, our data also supports an idea where eIF4A1 and NACA are both negative regulators of *EBNA1* mRNA translation which can collaborate in a same pathway but do not depend on each other to inhibit *EBNA1* mRNA translation. In all cases, the RNA structures and especially RNA G4s are key to this inhibitory pathway.

eIF4A1 separates strands in RNA duplexes and this way unwinds the mRNA or modify its structure. Local unwinding or modulation of RNA structure does not require a full ATP hydrolysis cycle, and in some cases, ATP binding alone is enough for eIF4A1 helicase activity *in vitro* (51). We showed here that NCL binding on G4-forming sequence in the GAR-encoding mRNA requires the presence of eIF4A1 in the cells, and that eIF4A1 binds directly and specifically RNA G4 structures *in vitro*. Hence, we hypothesised that eIF4A1 could either binds the mRNA and is subsequently joined by NCL, or that eIF4A1 modulates specifically the RNA G4 structures, allowing NCL interaction with the NCL-friendly G4. As it was shown that NCL binds directly to RNA oligos *in vitro*, we favour the modulating structure hypothesis (46). This hypothesis can seem at odds with the RNA pulldown results obtained by Lista *et al.* (46), but we cannot exclude the possibility that, by artificially folding the RNA oligos into RNA G4 structures, some oligos might stay unfolded and others fold into different type of G4s. Another interesting detail is that EBNA1 protein co-immunoprecipitates with both NCL and *EBNA1* mRNA (47–49). The RNP complex formed by the EBNA1 protein, NCL, and a G4 forming RNA, which could be *EBNA1* mRNA, is required for the proper replication of the viral episome (47–49). Furthermore, the EBNA1 protein interaction with NCL and efficient viral replication requires the ATP-binding activity of NCL (49). If NCL and eIF4A1 compete for the same binding site on the mRNA, one can speculate that NCL, by bringing ATP, gives enough energy for eIF4A1 to modify many RNA G4 structures and binds subsequently to the mRNA while interacting with the EBNA1 protein, forming a stable RNP complex. This model remains highly speculative, and more studies are needed to better understand this mechanism. A first step would be to confirm that

eIF4A1 is actually modifying locally RNA structure. Nevertheless, supporting this idea, adding an IRES structure on the mRNA or treating the cells with G4 ligands, both impacting the G4 structure-associated functions of *EBNA1* mRNA, modify the number of EBNA1-NACA interactions detected by PLA.

The fact that eIF4A1 unwinds or modulates mRNA in a non-processive manner and without any directionality when it is not associated with other initiation factors suggests that eIF4A1 could have another function in the cells additionally to its role as an initiation factor. Some studies support the idea that DEAD-box helicases are local RNA structure modulators, and can displace proteins from mRNAs and thus remodel RNP complexes (22). Two DEAD/H-box helicases, DED1 and NPH-II, are able *in vitro* to displace RBP from mRNA in an ATP hydrolysis-dependent manner (52). Additionally, the authors show that different DEAD/H-box helicases are selective regarding the targeted RNP (52). DEAD-box helicases Ded1p and Mss116p can unwind RNA duplexes locally, even if there is no RNA single-strand before their arrival and with no translocation associated with unwinding, contrary to other families of nucleic acids helicases (53, 54). In yeast, the DEAD-box RNA helicase Dbp5 controls mRNA export by displacing Nab2 from mRNAs *in vitro*, in an ADP binding-dependent fashion but independently from ATP hydrolysis (55). Finally, DEAD-box ATPases have been shown to control membrane-less organelles formation and composition, in cooperation with eIF4G MIF4G domains (56), and have been proposed to avoid RNA aggregate formation by being an ATP-dependent RNA chaperone (57).

MATERIAL AND METHODS

Cell culture, transfection, drug treatment, and siRNA

Human carcinoma-derived H1299 cells were 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).

H1299 cells were seeded in six-well plates at a density of $0.8-1.2 \times 10^5$ cells/well. The following day the cells were cotransfected with a maximum of 1 μ g of total expression plasmid along with 3 μ l of Genejuice according to the manufacturer's protocol (Merck Biosciences).

eIF4A1, eIF4A2 and eIF4A3 siRNAs (GeneSolution siRNA) and negative control siRNAs (AllStars Negative control) were purchased from Qiagen. Silencing was performed using JetPRIME according to the manufacturer's protocol (Polyplus Transfection).

When indicated, cells were treated for 24 hours with 50 nM of silvestrol (MedChemExpress) or hippuristanol (kind gift from Pr. J. Tanaka, University of the Ryukyus, Okinawa, Japan) unless other concentrations are mentioned. Treatments with G4 ligands were done during the last 36-40h of cell culture, with 2 μ M of either PhenDC3 or PhenDH2 (both provided by Dr. A. Granzhan, Institut Curie, Paris, France).

Plasmids construction

Plasmids used for transfections were described elsewhere (43, 44). Briefly, the coding sequence of EBNA1, EBNA Δ GA_r (EBNA1 protein without the GA_r domain), Ova (ovalbumin without the first 50 amino acids), GA_r-Ova (GA_r fused in the N-terminus of ovalbumin without the first 50 amino acids), and HA-polyQ-p53 (HA-tagged protein composed of a poly-glutamine stretch of 127 amino acids fused in the N-terminus of p53) are cloned in the PcDNA3 vector. C-Myc_EBNA1 and HCV_EBNA1 constructs, where the *c-Myc* and *HCV* IRES respectively are put in the 5'UTR of the EBNA1 coding sequence, were obtained by restriction ligation.

Electrophoresis and Western Blotting

Cells were harvested 40h post-transfection and lysed in lysis buffer (20mM HEPES KOH pH 7.5, 50mM b-glycerophosphate, 1mM EDTA pH 8.0, 1mM EGTA pH 8.0, 0.5mM Na₃VO₄, 100mM KCl, 10% glycerol and 1% Triton X-100) supplemented with complete protease inhibitor cocktail (Roche Diagnostics). Protein concentrations were measured using a Bradford assay. Total cell extracts were fractionated by SDS-PAGE, transferred to BioTrace[®]NT nitrocellulose blotting membranes (Pall Corp.), and probed with anti-EBNA1 mouse monoclonal antibody (OT1X, Cyto-Barr), anti-p53 mouse monoclonal antibody (DO-1), anti-HA mouse monoclonal antibody (provided by Dr. B. Vojtěšek, Masaryk Memorial Cancer Institute, Brno, Czech Republic), anti-Ovalbumin whole serum (Sigma), anti-NCL polyclonal rabbit antibody (Abcam), anti-NACA rabbit polyclonal antibody (Abcam), anti-eIF4A1 rabbit polyclonal antibody (CST), anti-eIF4A2 mouse polyclonal antibody (Abcam), anti-eIF4A3 rabbit polyclonal antibody (Abcam), and anti-actin mouse monoclonal antibody (Sigma). After incubation with the appropriate peroxidase-conjugated secondary antibodies (Dako), proteins were visualised by using Pierce ECL, WestDura, or West Femto (ThermoFisher) and quantified with the MyECL Imager system (Thermo Scientific) and ImageJ. Relative quantifications of the HRP signals normalised with the corresponding actin bands are mentioned above the corresponding band.

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

Transfected 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)₁₂₋₁₈ primer (Life technologies). qPCR was performed using the StepOne real-time PCR system (Applied Biosystems) with Perfecta SYBR Green FastMix (Quanta Biosciences) and the following primers (custom primers by Invitrogen): Actin Forward 5' - TCACCCACACTGTGCCCATCTACGA- 3', Actin Reverse 5' - TGAGGTAGTCAGTCAGGTCCCG - 3', Ova Forward 5'-GCAAACCTGTGCAGATGATG-3', Ova Reverse 5'-CTGCTCAAGGCCTGAGACTT-3', EBNA1 Forward 5'-GGCAGTGGACCTCAAAGAAGAG-3', EBNA1 Reverse 5'-CAATGCAACTTGGACGTTTTTG-3', eIF4A1 Forward 5'- GGCATCTACGCGTATGGTTT-3', eIF4A1 Reverse 5'- CCTAGTGCCATGACCACCTT-3'. *In vitro* RNA co-IP was carried out as described elsewhere (4). Briefly, 1 μ g 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 (50mM Tris pH7.5, 150mM NaCl, 0.02mg/mL yeast tRNA, 0.2mg/mL BSA) for 15min 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 antibodies (Abcam) according to standard conditions and purified using the TRIzol (Life Technologies). Precipitated RNAs were then analysed by RT-qPCR.

RNA pulldown

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 1X protease inhibitor cocktail (Roche). Cell lysis was performed by five series of vortex followed by 10 min incubation on ice, and three series of 3 s 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) were used for pulldown assays with the following G-quadruplex forming oligonucleotides: GQ- 5'-GGGGCAGGAGCAGGAGGA-3'Biotin TEG, ARPC2- 5' AGCCGGGGGCUGGGCGGGACCGGGCUUGU-3'Biotin TEG. The negative control for EBNA1 G4 was the GM- 5' GAGGCAGUAGCAGUAGAA-3'Biotin TEG oligonucleotide which, according to the GQRS mapper software, 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 µg 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 2X SDS loading buffer and analysed by western blotting against NACA or eIF4A1, as previously described. In the input lane of the western blots was loaded a quantity of extract which corresponds to the quantity that was incubated with the beads for each condition.

RNA FISH

H1299 cells were seeded in 24-well plates (2x10⁴ cells/well) and transient transfections were carried out 20-24h later using the Genejuice reagent (Merck Bioscience) according to the manufacturer's protocol. 24h after transfection or after drug treatment, cells were briefly washed with ice-cold PBS, fixed with 4% PFA for 20 min at room temperature, and washed again with PBS. Cells were then incubated in 70% Ethanol for 4-24 h at 4°C, after intermediate dehydrating steps using 30% and 50% Ethanol. For rehydration, cells were incubated in 50% and 30% Ethanol and further washed with PBS. Subsequently, cells were permeabilised with PBS 0.4% Triton 0.05% CHAPS for 5 min at room temperature. After three PBS washes, samples were pre-treated with two subsequent incubations in FiSH wash buffer (10% formamide and 2X SSC in ddH₂O) for 10 min and then in FiSH hybridisation buffer (10% formamide, 2X SSC, 2mg/mL BSA, 0.2mg/mL *E.coli* tRNA, 0.2mg/mL sheared salmon sperm DNA) for 30min at room temperature. Coverslips were then 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 20min in FiSH hybridisation buffer and 5 min in FiSH Wash buffer and subsequently stained with DAPI. Samples were mounted using Dako mounting solution (Dako) and observed with the LSM 800 (Zeiss) confocal laser microscope. 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 (58), and later used for statistical analysis.

Proximity Ligation Assay (PLA) protein-protein and immunostaining

Cells were cultured, fixed, permeabilised, and submitted to a pre-hybridisation step as described above. Then, primary antibodies incubation and PLA were carried out using the Naveniflex MR kit (Navinci), following the manufacturer's protocol. The following antibodies were used to perform the Proximity Ligation Assay (PLA): anti-HA mouse monoclonal antibody (provided by Dr. B. Vojtěšek), anti-NACA rabbit polyclonal (Abcam), anti-HA rabbit polyclonal antibody (Sigma), anti-NACA mouse monoclonal antibody (Abnova), anti-NCL polyclonal rabbit antibody (Abcam) and anti-EBNA1 mouse monoclonal antibody (OT1X, Cyto-Barr). For immunostaining, cells were incubated with an anti-mouse goat antibody coupled with Alexa Fluor 488 (ThermoFisher), after primary antibodies incubation and before PLA. Coverslips were finally mounted using Dako mounting buffer (Dako) after nuclear staining with DAPI and observed using the LSM 800 confocal laser microscope. Images were obtained using the Zen software and analysed using CellProfiler (58). Data were then processed to take into account the difference in targeted protein amounts as well as in antibody quality: 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.

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 significances of the difference are stated as following: $p > 0.05$ (ns), $p < 0.05$ (*), $p < 0.01$ (**), and $p < 0.001$ (***)).

ACKNOWLEDGEMENT

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 PhenDC3, PhenDH2, and recombinant NCL. We thank Pr. Junichi Tanaka (University of the Ryukyus, Okinawa, Japan) for kindly giving us hippuristanol.

AUTHORS CONTRIBUTIONS

A.Z., R.S., T.D., L. M.-C., A.Q., and C.D. performed the experiments. A.Z. and R.F. wrote the manuscript, with the help of C.D. and M.B.. M.B. and R.F. supervised the work.

DECLARATION OF INTERESTS

The authors declare no competing interests.

FUNDING

A.Z. was funded by Université de Paris and the Fondation pour la Recherche Médicale [grant number FDT202001010912]. This work was partially supported by 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, and by the 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.

REFERENCES

1. 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.
2. Riechert, E., Kmietczyk, V., Stein, F., Schwarzl, T., Sekaran, T., Jürgensen, L., Kamuf-Schenk, V., Varma, E., Hofmann, C., Rettel, M., *et al.* (2021) Identification of dynamic RNA-binding proteins uncovers a Cpeb4-controlled regulatory cascade during pathological cell growth of cardiomyocytes. *Cell Rep.*, **35**, 109100.
3. Garcia-Moreno, M., Noerenberg, M., Ni, S., Järvelin, A.I., González-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.e11.
4. Candeias, M.M., Malbert-Colas, L., Powell, D.J., Daskalogianni, C., Maslon, M.M., Naski, N., Bourougaa, K., Calvo, F. and Fåhraeus, R. (2008) p53 mRNA controls p53 activity by managing Mdm2 functions. *Nat. Cell Biol.*, **10**, 1098–1105.
5. Naski, N., Gajjar, M., Bourougaa, K., Malbert-Colas, L., Fåhraeus, R. and Candeias, M.M. (2009) The p53 mRNA-Mdm2 interaction. *Cell Cycle*, **8**, 31–34.
6. 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.
7. Lewis, C.J.T., Pan, T. and Kalsotra, A. (2017) RNA modifications and structures cooperate to guide RNA–protein interactions. *Nat. Rev. Mol. Cell Biol.*, **18**, 202–210.
8. Jackson, R.J., Hellen, C.U.T. and Pestova, T.V. (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell Biol.*, **11**, 113–127.
9. 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.
10. Fay, M.M., Lyons, S.M. and Ivanov, P. (2017) RNA G-Quadruplexes in Biology: Principles and Molecular Mechanisms. *J. Mol. Biol.*, **429**, 2127–2147.
11. Dumas, L., Herviou, P., Dassi, E., Cammas, A. and Millevoi, S. (2021) G-Quadruplexes in RNA Biology: Recent Advances and Future Directions. *Trends Biochem. Sci.*, **46**, 270–283.
12. Ruggiero, E. and Richter, S.N. (2018) G-quadruplexes and G-quadruplex ligands: targets and tools in antiviral therapy. *Nucleic Acids Res.*, **46**, 3270–3283.
13. Endoh, T. and Sugimoto, N. (2013) Unusual –1 Ribosomal Frameshift Caused by Stable RNA G-Quadruplex in Open Reading Frame. *Anal. Chem.*, **85**, 11435–11439.

14. Endoh,T. and Sugimoto,N. (2016) Mechanical insights into ribosomal progression overcoming RNA G-quadruplex from periodical translation suppression in cells. *Sci. Rep.*, **6**, 22719.
15. Endoh,T., Kawasaki,Y. and Sugimoto,N. (2013) Translational halt during elongation caused by G-quadruplex formed by mRNA. *Methods*, **64**, 73–78.
16. Yu,C.-H., Teulade-Fichou,M.-P. and Olsthoorn,R.C.L. (2014) Stimulation of ribosomal frameshifting by RNA G-quadruplex structures. *Nucleic Acids Res.*, **42**, 1887–1892.
17. Beaudoin,J.-D. and Perreault,J.-P. (2010) 5'-UTR G-quadruplex structures acting as translational repressors. *Nucleic Acids Res.*, **38**, 7022–7036.
18. Kumari,S., Bugaut,A., Huppert,J.L. and Balasubramanian,S. (2007) An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation. *Nat. Chem. Biol.*, **3**, 218–221.
19. Miglietta,G., Cogoi,S., Marinello,J., Capranico,G., Tikhomirov,A.S., Shchekotikhin,A. and Xodo,L.E. (2017) RNA G-Quadruplexes in Kirsten Ras (KRAS) Oncogene as Targets for Small Molecules Inhibiting Translation. *J. Med. Chem.*, **60**, 9448–9461.
20. Cammas,A., Lacroix-Triki,M., Pierredon,S., Le Bras,M., Iacovoni,J.S., Teulade-Fichou,M.-P., Favre,G., Roché,H., Filleron,T., Millevoi,S., *et al.* (2016) hnRNP A1-mediated translational regulation of the G quadruplex-containing RON receptor tyrosine kinase mRNA linked to tumor progression. *Oncotarget*, **7**, 16793–16805.
21. 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.
22. Linder,P. and Jankowsky,E. (2011) From unwinding to clamping — the DEAD box RNA helicase family. *Nat. Rev. Mol. Cell Biol.*, **12**, 505–516.
23. Jarmoskaite,I. and Russell,R. (2011) DEAD-box proteins as RNA helicases and chaperones: DEAD-box proteins. *Wiley Interdiscip. Rev. RNA*, **2**, 135–152.
24. Cencic,R., Carrier,M., Galicia-Vázquez,G., Bordeleau,M.-E., Sukarieh,R., Bourdeau,A., Brem,B., Teodoro,J.G., Greger,H., Tremblay,M.L., *et al.* (2009) Antitumor Activity and Mechanism of Action of the Cyclopenta[b]benzofuran, Silvestrol. *PLoS ONE*, **4**, e5223.
25. Lindqvist,L., Oberer,M., Reibarkh,M., Cencic,R., Bordeleau,M.-E., Vogt,E., Marintchev,A., Tanaka,J., Fagotto,F., Altmann,M., *et al.* (2008) Selective Pharmacological Targeting of a DEAD Box RNA Helicase. *PLoS ONE*, **3**, e1583.
26. Bordeleau,M.-E., Mori,A., Oberer,M., Lindqvist,L., Chard,L.S., Higa,T., Belsham,G.J., Wagner,G., Tanaka,J. and Pelletier,J. (2006) Functional characterization of IRESes by an inhibitor of the RNA helicase eIF4A. *Nat. Chem. Biol.*, **2**, 213–220.
27. Wolfe,A.L., Singh,K., Zhong,Y., Drewe,P., Rajasekhar,V.K., Sanghvi,V.R., Mavrakis,K.J., Jiang,M., Roderick,J.E., Van der Meulen,J., *et al.* (2014) RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer. *Nature*, **513**, 65–70.

28. Waldron, J.A., Raza, F. and Le Quesne, J. (2018) eIF4A alleviates the translational repression mediated by classical secondary structures more than by G-quadruplexes. *Nucleic Acids Res.*, **46**, 3075–3087.
29. Waldron, J.A., Tack, D.C., Ritchey, L.E., Gillen, S.L., Wilczynska, A., Turro, E., Bevilacqua, P.C., Assmann, S.M., Bushell, M. and Le Quesne, J. (2019) mRNA structural elements immediately upstream of the start codon dictate dependence upon eIF4A helicase activity. *Genome Biol.*, **20**, 300.
30. Rogers, G.W., Richter, N.J. and Merrick, W.C. (1999) Biochemical and Kinetic Characterization of the RNA Helicase Activity of Eukaryotic Initiation Factor 4A. *J. Biol. Chem.*, **274**, 12236–12244.
31. Rogers, G.W., Richter, N.J., Lima, W.F. and Merrick, W.C. (2001) Modulation of the Helicase Activity of eIF4A by eIF4B, eIF4H, and eIF4F. *J. Biol. Chem.*, **276**, 30914–30922.
32. Rozovsky, N., Butterworth, A.C. and Moore, M.J. (2008) Interactions between eIF4A and its accessory factors eIF4B and eIF4H. *RNA*, **14**, 2136–2148.
33. Parsyan, A., Svitkin, Y., Shahbazian, D., Gkogkas, C., Lasko, P., Merrick, W.C. and Sonenberg, N. (2011) mRNA helicases: the tacticians of translational control. *Nat. Rev. Mol. Cell Biol.*, **12**, 235–245.
34. Young, L.S., Yap, L.F. and Murray, P.G. (2016) Epstein–Barr virus: more than 50 years old and still providing surprises. *Nat. Rev. Cancer*, **16**, 789–802.
35. Farrell, P.J. (2018) Epstein–Barr Virus and Cancer.
36. 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.
37. Münz, C., Bickham, K.L., Subklewe, M., Tsang, M.L., Chahroudi, A., Kurilla, M.G., Zhang, D., O’Donnell, M. and Steinman, R.M. Human CD4⁺ T Lymphocytes Consistently Respond to the Latent Epstein-Barr Virus Nuclear Antigen EBNA.
38. Long, H.M., Haigh, T.A., Gudgeon, N.H., Leen, A.M., Tsang, C.-W., Brooks, J., Landais, E., Houssaint, E., Lee, S.P., Rickinson, A.B., *et al.* (2005) CD4⁺ T-Cell Responses to Epstein-Barr Virus (EBV) Latent-Cycle Antigens and the Recognition of EBV-Transformed Lymphoblastoid Cell Lines. *J. Virol.*, **79**, 4896–4907.
39. Paludan, C., Bickham, K., Nikiforow, S., Tsang, M.L., Goodman, K., Hanekom, W.A., Fonteneau, J.-F., Stevanović, S. and Münz, C. (2002) Epstein-Barr Nuclear Antigen 1-Specific CD4⁺ Th1 Cells Kill Burkitt’s Lymphoma Cells. *J. Immunol.*, **169**, 1593–1603.
40. Blake, N., Lee, S., Redchenko, I., Thomas, W., Steven, N., Leese, A., Steigerwald-Mullen, P., Kurilla, M.G., Frappier, L. and Rickinson, A. Human CD8⁺ T Cell Responses to EBV EBNA1: HLA Class I Presentation of the (Gly-Ala)–Containing Protein Requires Exogenous Processing.
41. Levitskaya, J., Coram, M., Levitsky, V., Imreh, S., Steigerwald-Mullen, P.M., Klein, G., Kurilla, M.G. and Masucci, M.G. (1995) Inhibition of antigen processing by the internal repeat region of the Epstein–Barr virus nuclear antigen-1. *Nature*, **375**, 685–688.

42. Mautner,J., Pich,D., Nimmerjahn,F., Milosevic,S., Adhikary,D., Christoph,H., Witter,K., Bornkamm,G.W., Hammerschmidt,W. and Behrends,U. (2004) Epstein-Barr virus nuclear antigen 1 evades direct immune recognition by CD4+ T helper cells. *Eur. J. Immunol.*, **34**, 2500–2509.
43. Yin,Y., Manoury,B. and Fåhraeus,R. (2003) Self-Inhibition of Synthesis and Antigen Presentation by Epstein-Barr Virus-Encoded EBNA. **301**, 5.
44. Apcher,S., Komarova,A., Daskalogianni,C., Yin,Y., Malbert-Colas,L. and Fåhraeus,R. (2009) mRNA Translation Regulation by the Gly-Ala Repeat of Epstein-Barr Virus Nuclear Antigen 1. *J. Virol.*, **83**, 1289–1298.
45. 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.
46. 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.
47. Norseen,J., Thomae,A., Sridharan,V., Aiyar,A., Schepers,A. and Lieberman,P.M. (2008) RNA-dependent recruitment of the origin recognition complex. *EMBO J.*, **27**, 3024–3035.
48. 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.
49. Chen,Y.-L., Liu,C.-D., Cheng,C.-P., Zhao,B., Hsu,H.-J., Shen,C.-L., Chiu,S.-J., Kieff,E. and Peng,C. -w. (2014) Nucleolin is important for Epstein-Barr virus nuclear antigen 1-mediated episome binding, maintenance, and transcription. *Proc. Natl. Acad. Sci.*, **111**, 243–248.
50. 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.
51. Chen,Y., Potratz,J.P., Tijerina,P., Del Campo,M., Lambowitz,A.M. and Russell,R. (2008) DEAD-box proteins can completely separate an RNA duplex using a single ATP. *Proc. Natl. Acad. Sci.*, **105**, 20203–20208.
52. Fairman,M.E. (2004) Protein Displacement by DExH/D ‘RNA Helicases’ Without Duplex Unwinding. *Science*, **304**, 730–734.
53. Yang,Q., Del Campo,M., Lambowitz,A.M. and Jankowsky,E. (2007) DEAD-Box Proteins Unwind Duplexes by Local Strand Separation. *Mol. Cell*, **28**, 253–263.
54. Yang,Q. and Jankowsky,E. (2006) The DEAD-box protein Ded1 unwinds RNA duplexes by a mode distinct from translocating helicases. *Nat. Struct. Mol. Biol.*, **13**, 981–986.
55. Tran,E.J., Zhou,Y., Corbett,A.H. and Wente,S.R. (2007) The DEAD-Box Protein Dbp5 Controls mRNA Export by Triggering Specific RNA:Protein Remodeling Events. *Mol. Cell*, **28**, 850–859.

56. Hondele,M., Sachdev,R., Heinrich,S., Wang,J., Vallotton,P., Fontoura,B.M.A. and Weis,K. (2019) DEAD-box ATPases are global regulators of phase-separated organelles. *Nature*, **573**, 144–148.
57. Tauber,D., Tauber,G., Khong,A., Van Treeck,B., Pelletier,J. and Parker,R. (2020) Modulation of RNA Condensation by the DEAD-Box Protein eIF4A. *Cell*, **180**, 411-426.e16.
58. McQuin,C., Goodman,A., Chernyshev,V., Kametsky,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.

FIGURES LEGENDS

Figure 1: eIF4A1 inhibits *EBNA1* mRNA translation

A: Western Blot from H1299 cells transfected with the *EBNA1*-encoding construct, treated or not with siRNA against eIF4A1 (left panel), eIF4A2 (middle panel), or eIF4A3 (right panel). Western Blots are representative of at least three independent experiments. **B:** *eIF4A1* (left panel) and *EBNA1* (right panel) mRNAs levels in H1299 cells transfected with *EBNA1* construct, treated or not with siRNA against eIF4A1. Graphs were obtained with data from at least three independent experiments. **C:** Western Blot from H1299 cells transfected with the *EBNA1* or *EBNA1* Δ *Gar*-encoding construct, treated or not with hippuristanol. Western Blot is representative of at least three independent experiments. **D:** Western Blot from H1299 cells transfected with the *EBNA1* or *EBNA1* Δ *Gar*-encoding construct, treated or not with silvestrol. Western Blot is representative of at least three independent experiments.

Figure 2: eIF4A1 binds directly the G4 structures derived from the *EBNA1* mRNA and regulates NCL binding onto the mRNA

A: *in vitro* RNA coIP experiments performed with recombinant NCL and *EBNA1* mRNA extracted from H1299 transfected cells treated with siRNA against eIF4A1, eIF4A3 or control siRNAs. The graph shows relative fold change in NCL binding, using the control siRNA-treated cells as reference. **B:** *in vitro* RNA coIP experiments performed with recombinant NCL and *EBNA1* mRNA extracted from H1299 transfected cells treated with DMSO or hippuristanol. The graph shows relative fold change in NCL binding, using the control DMSO-treated cells as reference. **C:** *in vitro* RNA coIP experiments performed with recombinant NCL and *EBNA1* mRNA extracted from H1299 transfected cells treated with DMSO or silvestrol. The graph shows relative fold change in NCL binding, using the control DMSO-treated cells as reference. **D:** eIF4A1 binds to *Gar* G4s. RNA-pulldown experiment with RNA oligos and H1299 cell extracts, in presence of PhenDH2. The GQ-18 oligo is derived from the G4-forming sequence of the *Gar* mRNA, the GM-18 oligos is the equivalent of the GQ-18 with mutations impairing G4 formation, ARPC2 oligo is derived from the G4-forming sequence of *ARPC2*.

Figure 3: eIF4A1 and NACA inhibit *EBNA1* synthesis in a collaborative but independent manner

A: Western Blot from H1299 cells transfected with the *EBNA1*-encoding construct, treated or not with siRNA against NACA and/or eIF4A1. Western Blots are representative of at least three independent experiments. **B:** Left panel: Relative number of *EBNA1*-NACA interactions in H1299 transfected cells detected by PLA, normalised with the immunostaining signal intensity of the

corresponding reporter proteins. (n=63 cells for EBNA1, n=85 for cMyc_EBNA1, n=89 for HCV_EBNA1). Right panel: PLA assessing NACA-EBNA1 interactions in H1299 cells transfected with the indicated constructs, with immuno-co-staining of EBNA1. In blue: nucleus, in green: immuno-staining of the reporter proteins, in white: PLA dots. Scale bar represents 10µm. **C:** Left panel: Relative number of EBNA1-NACA interactions detected by PLA in H1299 transfected cells and treated with either DMSO, PhenDC3, or PhenDH2, normalised with the immunostaining signal intensity of the corresponding reporter proteins (n=41 cells for DMSO, n=14 for PhenDC3, n=58 for PhenDH2). Right panel: PLA assessing NACA-EBNA1 interactions in H1299 cells transfected with the indicated constructs and treated with either DMSO, PhenDC3, or PhenDH2, with immuno-co-staining of EBNA1. In blue: nucleus, in green: immuno-staining of the reporter proteins, in white: PLA dots. Scale bar represents 10µm.

Supplementary Figure S1

A: *EBNA1* mRNAs levels in H1299 cells transfected with EBNA1 construct, treated or not with siRNA against eIF4A1. Graphs were obtained with data from at least three independent experiments. **B:** Western Blot from H1299 cells transfected with the EBNA1-encoding construct, treated with increasing amount of siRNA against eIF4G1 (quantity indicated represent number of pmol of siRNA added in a 6-well plate well). Western Blots are representative of at least three independent experiments. **C:** Western Blot from H1299 cells transfected with the EBNA1-encoding construct, treated with increasing amount of siRNA against eIF4G2 (quantity indicated represent number of pmol of siRNA added in a 6-well plate well). Western Blots are representative of at least three independent experiments.

Supplementary Figure S2

A: Immunofluorescence against eIF4A1 and EBNA1 in H1299 cells transfected with the indicated construct. In blue: nucleus, in green: immuno-staining of EBNA1, in red: immunostaining of eIF4A1. Scale bar represents 10µm.

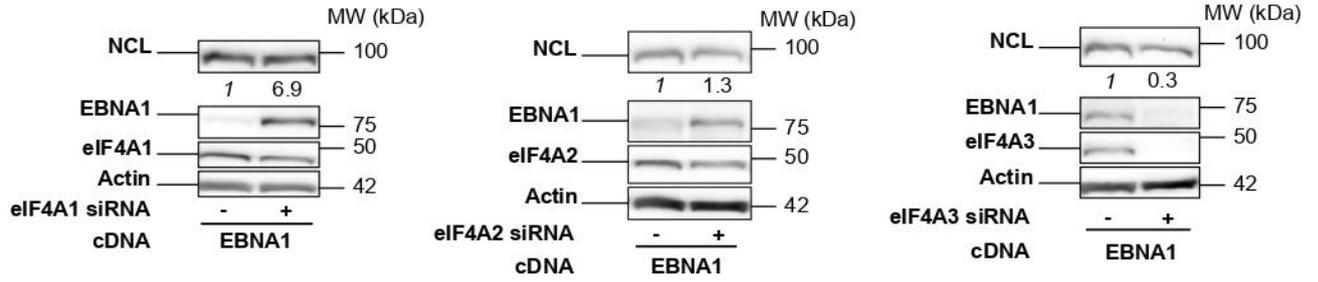
Supplementary Figure S3

A: Colocalisation of the *EBNA1*, *cMyc_EBNA1*, and *HCV_EBNA1* mRNAs with the nucleus. Left panel: graph showing the mean Costes correlation factor for each condition (n=17 for EBNA1, 20 for cMyc_EBNA1, 18 for HCV_EBNA1). A correlation factor of 1 means that both signals perfectly colocalise. Right panel: microscopy images obtained with H1299 cells transfected with the corresponding constructs or with the PcDNA3 empty vector (EV). In blue: nucleus (DAPI), in red:

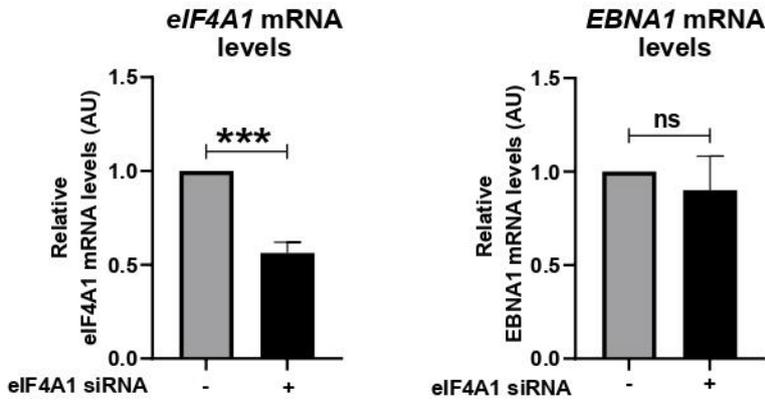
mRNA (FISH probes). Scale bar represents 10 μ m. **B:** Left panel: Relative number of HA-NACA interactions detected by PLA in H1299 transfected cells and treated with either DMSO, PhenDC3, or PhenDH2, normalised with the immunostaining signal intensity of the corresponding reporter proteins (n= 44 cells for DMSO, n=12 for PhenDC3, n=39 for PhenDH2). Right panel: PLA assessing NACA-HA interactions in H1299 cells transfected with the indicated constructs and treated with either DMSO, PhenDC3, or PhenDH2, with immuno-co-staining of HA-polyQ-p53. In blue: nucleus, in green: immuno-staining of the reporter proteins, in white: PLA dots. Scale bar represents 10 μ m. **C:** Left panel: Relative number of EBNA1-NCL interactions detected by PLA in H1299 transfected cells and treated with either DMSO, hippuristanol, or silvestrol, normalised with the immunostaining signal intensity of the corresponding reporter protein (n= 67 cells for DMSO, n=87 for hippuristanol, n=35 for silvestrol). Right panel: PLA assessing EBNA1-NCL interactions in H1299 cells transfected with the indicated construct and treated with either DMSO, hippuristanol, or silvestrol, with immuno-co-staining of EBNA1. In blue: nucleus, in green: immuno-staining of the reporter protein, in red: PLA dots. Scale bar represents 10 μ m.

Figure 1

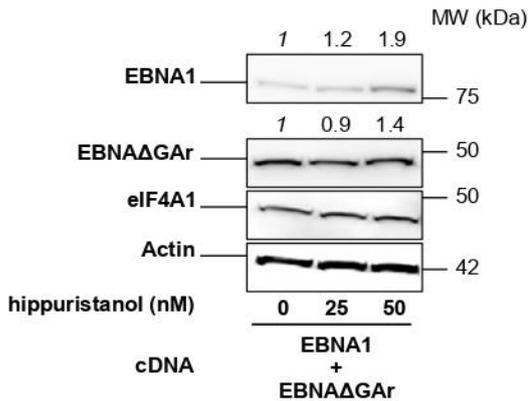
A



B



C



D

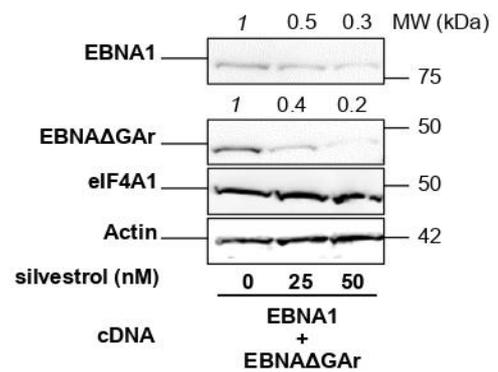
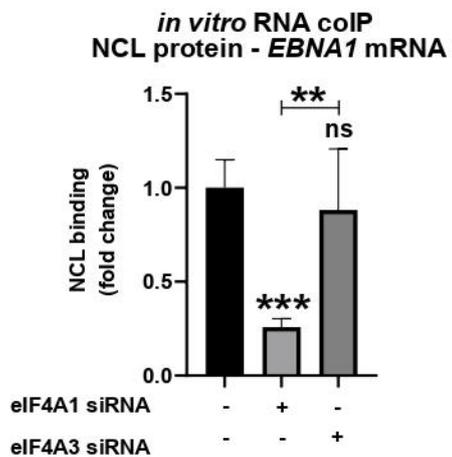
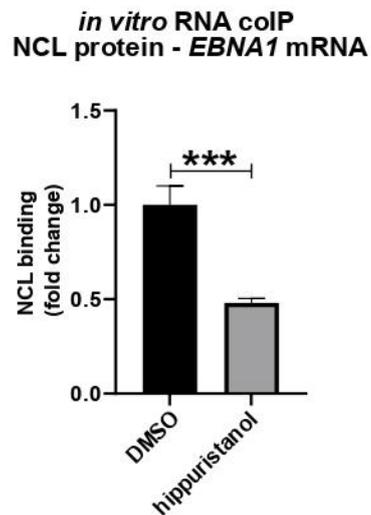


Figure 2

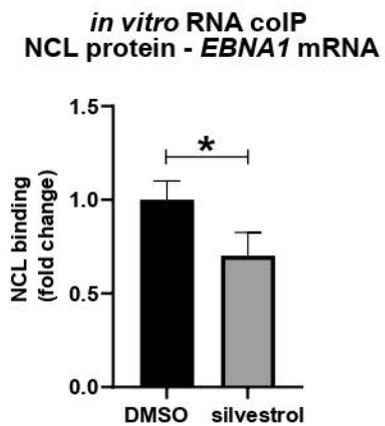
A



B



C



D

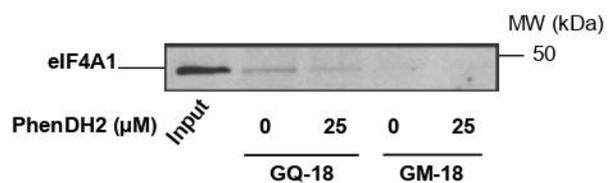
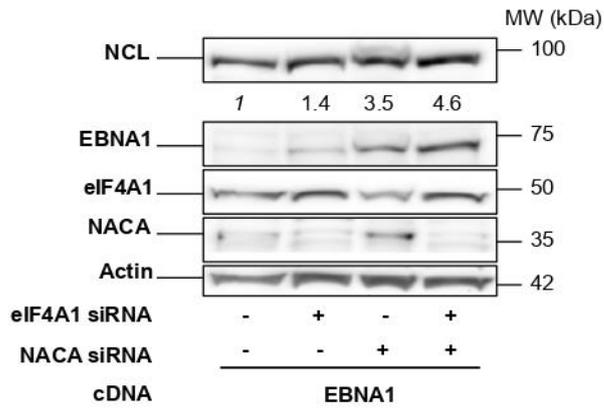
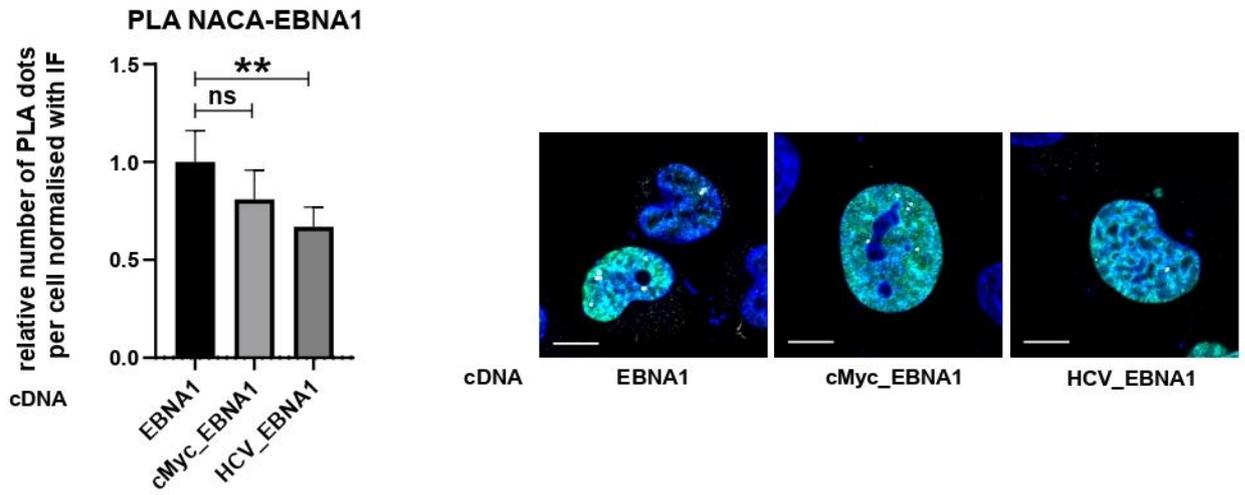


Figure 3

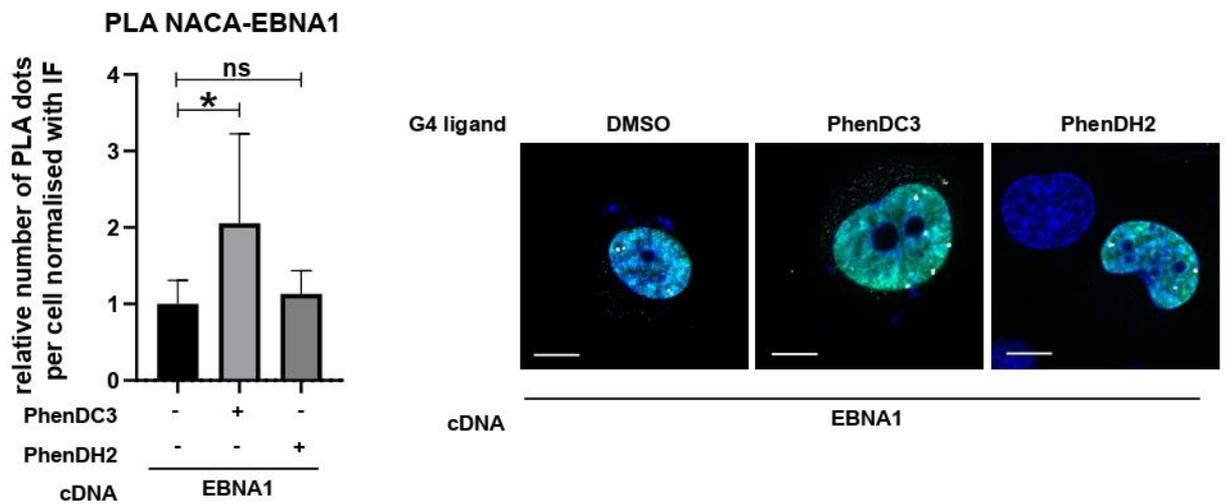
A



B

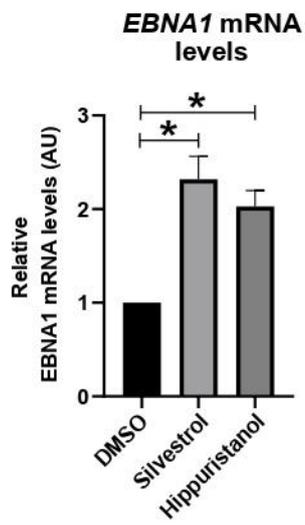


C

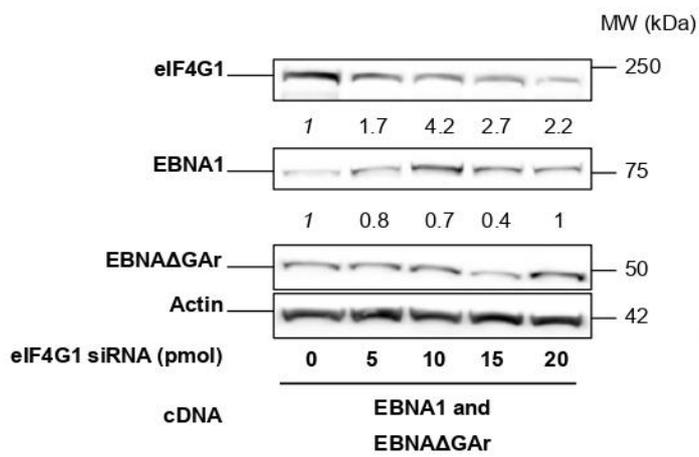


Supp. Figure S1

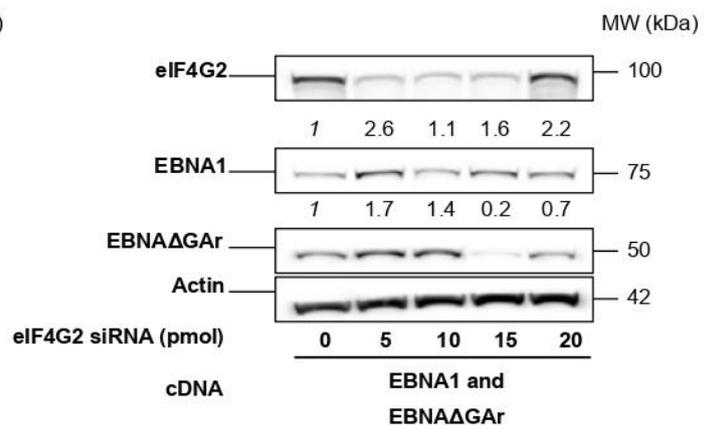
A



B

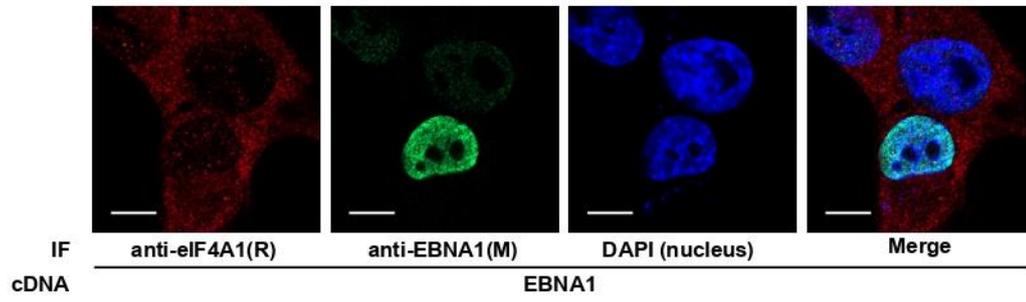


C

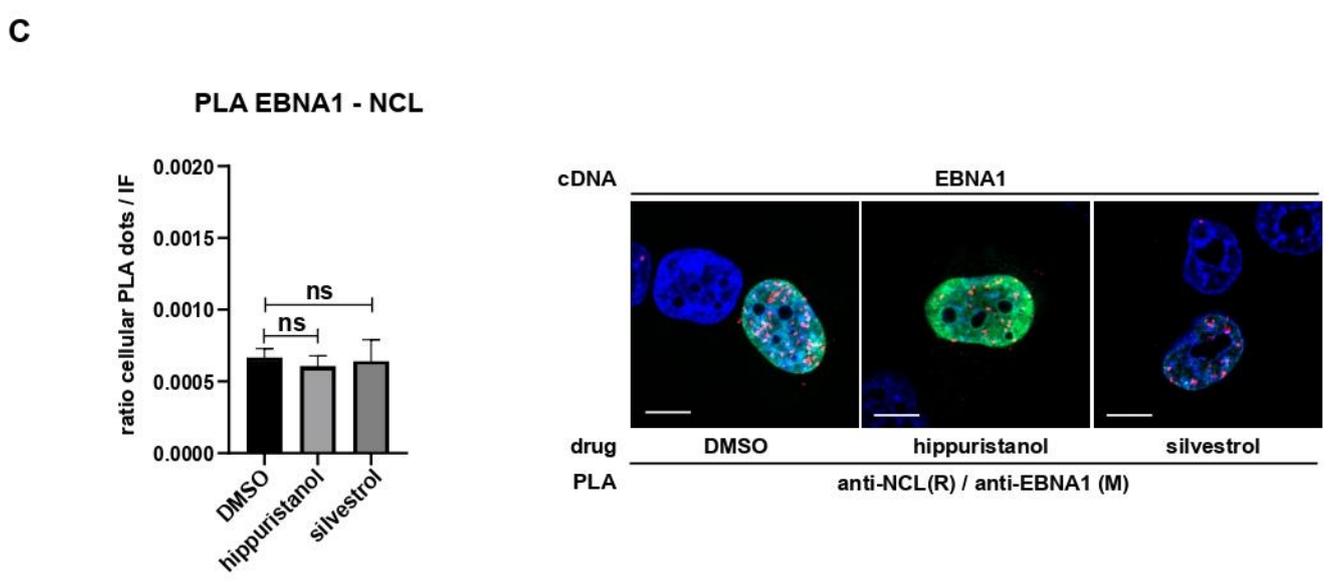
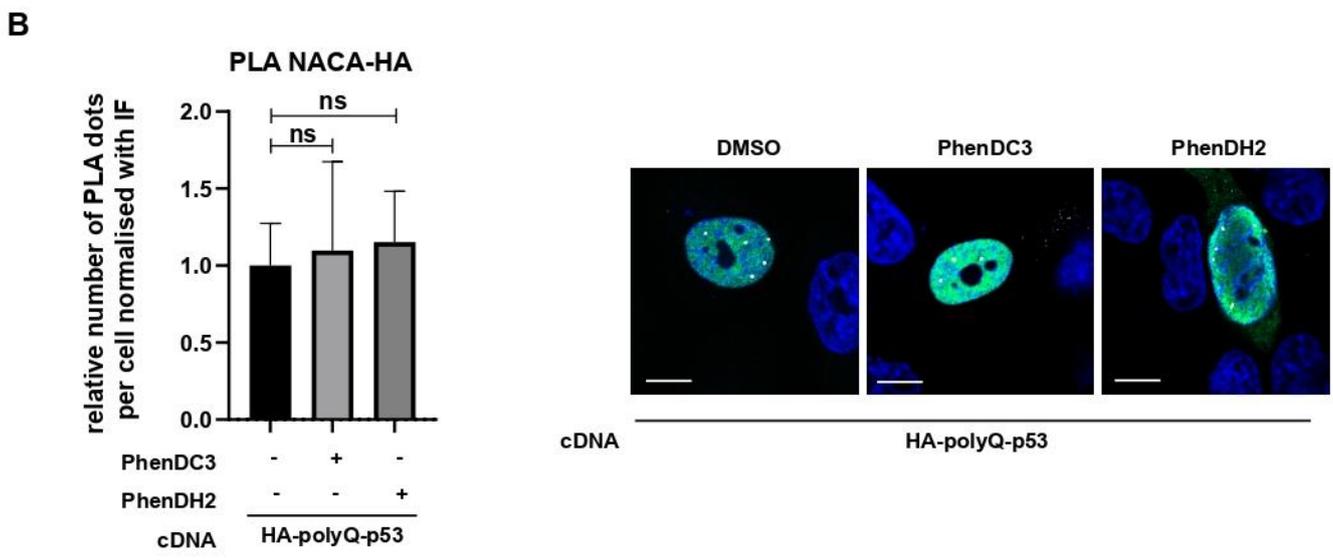
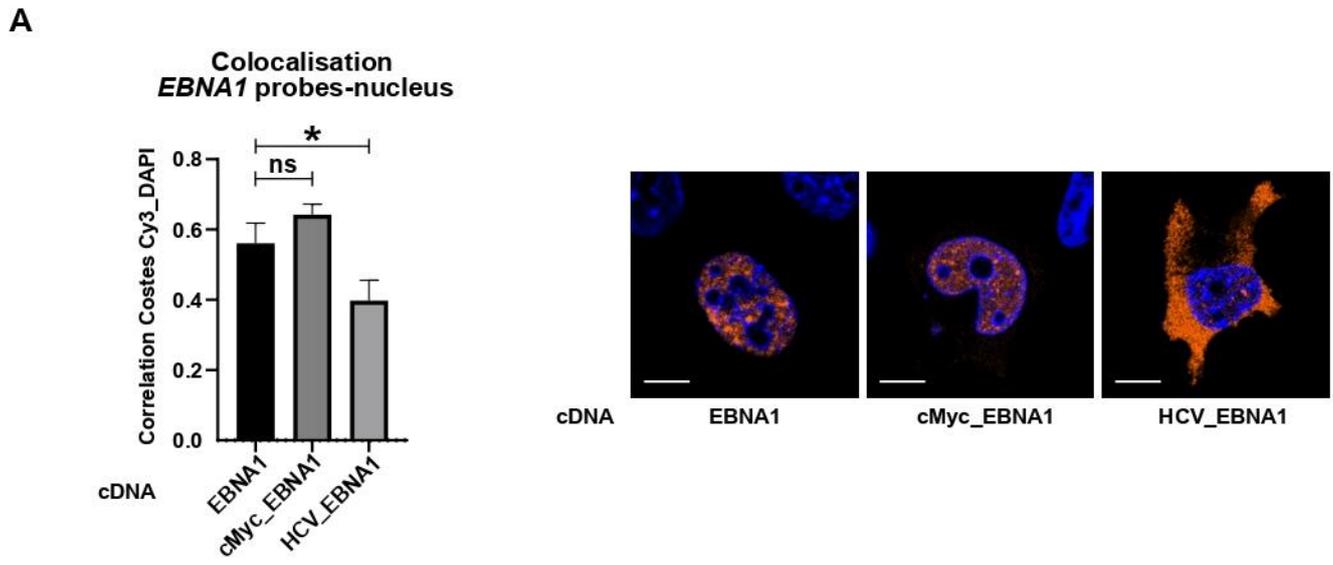


Supp. Figure S2

A



Supp. Figure S3



DISCUSSION

I – Proposed model for EBNA1 mRNA translation regulation

EBNA1 is an essential protein for the EBV but is highly antigenic and most healthy carriers have immune cells able to recognise it. With all these constraints, EBV maintenance in the host cells depends on a tight regulation of EBNA1 production, since enough EBNA1 must be present during cell division to ensure viral episome replication and proper segregation, while EBNA1 levels have to be kept at a minimum to avoid recognition by the immune cells. The GAR domain in EBNA1 N-terminal half is responsible for the inhibition of *EBNA1* mRNA translation *in cis*, hence inhibiting full-length proteins as well as antigenic peptides production.

From previous works, we know that the GAR polypeptide plays a crucial role in the GAR-mediated translation inhibition and that modifying the polypeptide sequence cancels this pathway (247). On the mRNA side, the GAR-encoding mRNA can fold into RNA G4 structures *in vitro* (228), and mutations impairing the formation of these structures relieved the GAR-mediated inhibition of translation (228). Moreover, NCL can bind directly and specifically to RNA G4 in the *GAR* coding sequence, and this interaction is prevented by some G4 ligands like PhenDC3 and its derivatives (79, 131, 250). Preventing NCL binding to the GAR-encoding mRNAs suppresses the GAR-related inhibition of translation (79, 131). In addition, the EBNA1 protein can bind to its own mRNA in a G4-dependent manner as well as to NCL (227, 229, 231). EBNA1 interaction with NCL depends on NCL binding to ATP (231). The whole RNP complex formed by NCL, EBNA1 protein, and G4-forming RNAs, which can be mRNAs, is required for replication of the viral genome and proper segregation of the EBV episomes between the daughter cells (227, 229, 231). RNA G4s are also needed for EBNA1 protein – mRNA interaction and genome replication (227, 229). Finally, the GAR polypeptide inhibits its proteasomal degradation when it is near other unfolded regions in the polypeptide (244, 245).

What our work shows is that RNA G4-associated functions are changing following different parameters, in line with the fact that RNA G4 structures are dynamic (Paper I). Additionally, we show that the NAC interacts with intrinsically disordered domains, like the GAR or the polyQ stretches. In presence of the GAR, the NAC is detached from the translating ribosomes, and this event is required for the *EBNA1* mRNA to gain its ability to bind NCL. Silencing NAC also leads to cellular *EBNA1* mRNA losing its ability to interact with recombinant NCL *in vitro*. Interestingly, recombinant NAC alone can restore *EBNA1* mRNA NCL-binding activity, even if mRNAs are extracted from cells where NAC has been silenced (Paper II). Finally, we show that eIF4A1 is involved in the GAR-mediated translation inhibition and is also needed for *EBNA1* mRNA to be able to bind NCL. Furthermore, we show that eIF4A1 can also interact directly with the RNA G4 oligos derived from the GAR-coding sequence. It

seems that NACA and eIF4A1 are both having similar activity, *i.e.* inhibiting *EBNA1* mRNA translation by controlling its interaction with NCL via RNA G4 structure. Thus, they can both be part of the same pathway and support each other activities, but they do not rely on each other and can act independently (Paper III).

By taking together all these pieces of information, our model is the following (**Fig. K**): when the *EBNA1* mRNA is translated, the NAC is the sensor of the nascent polypeptide, and decide which pathway the nascent polypeptide will follow. We propose that the presence of RNA G4-forming sequence in the mRNA, maybe in combination with additional specific sequence embedded in these G4-forming sequences, and/or the capacity of the protein to bind its own mRNA are determinant elements in the sorting process. In our proposed scenario, the translation of *EBNA1* mRNA leads to the unfolding of RNA G4 structures present in its coding sequence. RGG domains from the EBNA1 protein bind to unfolded RNA G4-forming sequence either in GAR or nearby RGG-encoding sequences. This interaction prevents some RNA G4s to be folded again. However, as the concentration of putative RNA G4s is high in the RGG- and GAR-encoding sequences, and since RNA G4s are thermodynamically favoured in cellular conditions, we propose that some RNA G4s are folded again in the GAR-encoding sequence. These RNA G4s could either be NCL or non-NCL-friendly, the latest being changed into NCL-friendly G4s by the action of eIF4A1. NCL-friendly RNA G4s can be further stabilised by the NAC present nearby as it is binding the GAR peptide. The action of both eIF4A1 and NAC supports NCL binding onto *EBNA1* mRNA. This way, the mRNA is trapped in an RNP complex with both EBNA1 and NCL, preventing further rounds of translation. This could be particularly interesting from the virus point of view, as this RNP complex is also needed for replication and proper segregation of the viral episome during cell division. Like this, EBNA1 ensures a feedback loop monitoring the presence of EBNA1 proteins and adapting its production. When the levels of EBNA1 are sufficient to ensure proper viral replication and segregation, the *EBNA1* mRNAs are trapped in this NCL-EBNA1 stable RNP complex, whereas when levels of EBNA1 are too low for ensuring viral genome maintenance, free to be translated *EBNA1* mRNAs are available, since EBNA1 protein is missing and the whole RNP complex is destabilised.

eIF4A1 role in this GAR-mediated inhibition of translation is still unclear and remains to be further investigated. A first step would be to see if inhibiting eIF4A1 results in any change in NACA interaction with the GAR or NCL interaction with EBNA1 protein or mRNA in the cells. Furthermore, we do not know if any of these three cellular factors interact with each other directly, if they bind together simultaneously or follow a precise timing, and if NAC and eIF4A1 are released from the complex when NCL is binding.

Nevertheless, we propose that eIF4A1 is involved in the GAR-mediated inhibition of translation as a modulator of RNA complex. eIF4A1 could be recruited on the RNA G4 of *EBNA1* mRNA, clamp on it and modify the RNA G4 into NCL-friendly RNA G4. As NCL binding to ATP is required for NCL interaction with EBNA1, we hypothesise that the ATP brought by NCL is used by eIF4A to completely modify the RNA G4 structures and leave the mRNA, letting the mRNA free to be bound by NCL. However, this idea about eIF4A1 role as an RNP modeler needs more investigation and the data we currently have does not exclude the potential role of eIF4A1 as an effector receiving the final signal and preventing further rounds of translation. It has also been shown *in vitro* that eIF4A can control mRNA recruitment by the initiation complex, in a helicase-independent manner. Using ATP binding and hydrolysis, eIF4A can control the eIF3j affinity to the PIC. As mRNA and eIF3j binding to the PIC are mutually exclusive, eIF4A by controlling eIF3j affinity for the PIC controls mRNA recruitment (259).

The GAR-mediated translation inhibition is likely a process acting very early on in the life of the mRNA. Our team formerly showed that putting the GAR in the context of a gene, *i.e.* in a primary transcript that will be spliced before translation, cancels the inhibition pathway. However, with the same construct where the GAR-encoding pre-mRNA also contains introns, inhibiting the splicing machinery restored the GAR-mediated inhibition pathway (184). With this information, we propose that RNA G4 formation, although thermodynamically favoured in the cells, is slower than the recruitment of the splicing complex, an event that can happen co-transcriptionally (260). During this short period of time where the mRNA is unfolded, the EBNA1-NCL RNP complex could bind the GAR-encoding sequence, if enough EBNA1 proteins are available, and in competition with the splicing complex when introns are also present. Thus, this would prevent even the primary round of translation to happen. In line with this, the NCL-EBNA1 RNP complex is participating in viral replication and viral transcription regulation (231), meaning it is spatially in proximity with the nascent transcript.

In our work, we only considered the EBNA1 protein, but we have to keep in mind that other viral elements are also present in the infected cells, like the EBV-encoded miRNAs. A recent study showed that EBV-encoded miRNAs target MHC Class I antigen processing and presentation pathways. These miRNAs downregulate gene expression of different antigen processing and presenting factors, as well as EBNA1 production (261). Knowing that *in vitro* EBNA1 interacts with *EBER*, a group of EBV-encoded non-coding RNA (ncRNA), one could wonder if some viral ncRNAs are also part of the NCL-EBNA1 RNP complex and strengthen the inhibition of *EBNA1* mRNA translation.

EBNA1 is an essential protein for EBV but is also highly disordered and prone to form aggregates. In addition to its translation inhibitory role, the RNA G4s in the GAR-encoding mRNA could also serve as a chaperone for the protein, as it has been recently suggested (262, 263), hence avoiding loss of EBNA1 functions by preventing its aggregation. NAC could also recruit other chaperones in the process. Wang *et al.* showed that targeting HSP70 molecular chaperone function impairs EBV replication and carcinogenicity in cells, but in a GAR-independent way (264).

EBNA1 is an oncogenic protein, and one of the pathways activated by the translational stress triggered by the GAR leads to increasing E2F1 production. Interestingly, E2Fs were shown to regulate *EBNA1* transcription during the cell cycle via the Qp promoter (233, 265). One could then wonder if EBNA1 induce tumourigenesis, or if it is maintaining and amplifying carcinogenic stress under certain circumstances. Indeed, to a certain extent, some oncogenic pathways auto-maintain themselves via a positive auto-feedback loop, like in the mTORC2 oncogenic pathway. In glioblastoma, mTORC2 activates a positive auto-feedback loop, where the production of one of its subunits, Rictor, is promoted (141).

Interestingly, EBNA1 interacts with NCL and NAC, one being clearly related to ribosomal biogenesis and the latter likely to participate in it. The ribosomal protein RPL4 is also a part of the EBNA1-NCL RNP, and its expression is promoted in EBV-infected cells (266). As we begin to grasp the importance of ribosomal heterogeneity in the cells, could it be that EBNA1 interferes in the ribosome synthesis not just in a quantitative, but also in a qualitative manner? This hypothesis still needs to be investigated, but this is in line with the idea that EBNA1 targets a specific translational pathway, sensitive to the GAR-mediated regulation. Even if no data confirm this idea, the fact that some IRES-mediated translations are not affected by the GAR is supporting it.

It is likely that different mechanisms, depending on the context, are activated in EBNA1-expressing cells. It is also possible that different viral-mediated pathways are active simultaneously in one EBV-infected cell. Here our data suggest a mechanism in which both RNA and peptide play an important role, but each viral element could also trigger their own inhibitory pathway. Considering all the information provided by the numerous studies on EBNA1, we can consider that EBNA1, instead of being EBV Achille's heel, is rather its best and most subtle arrow.

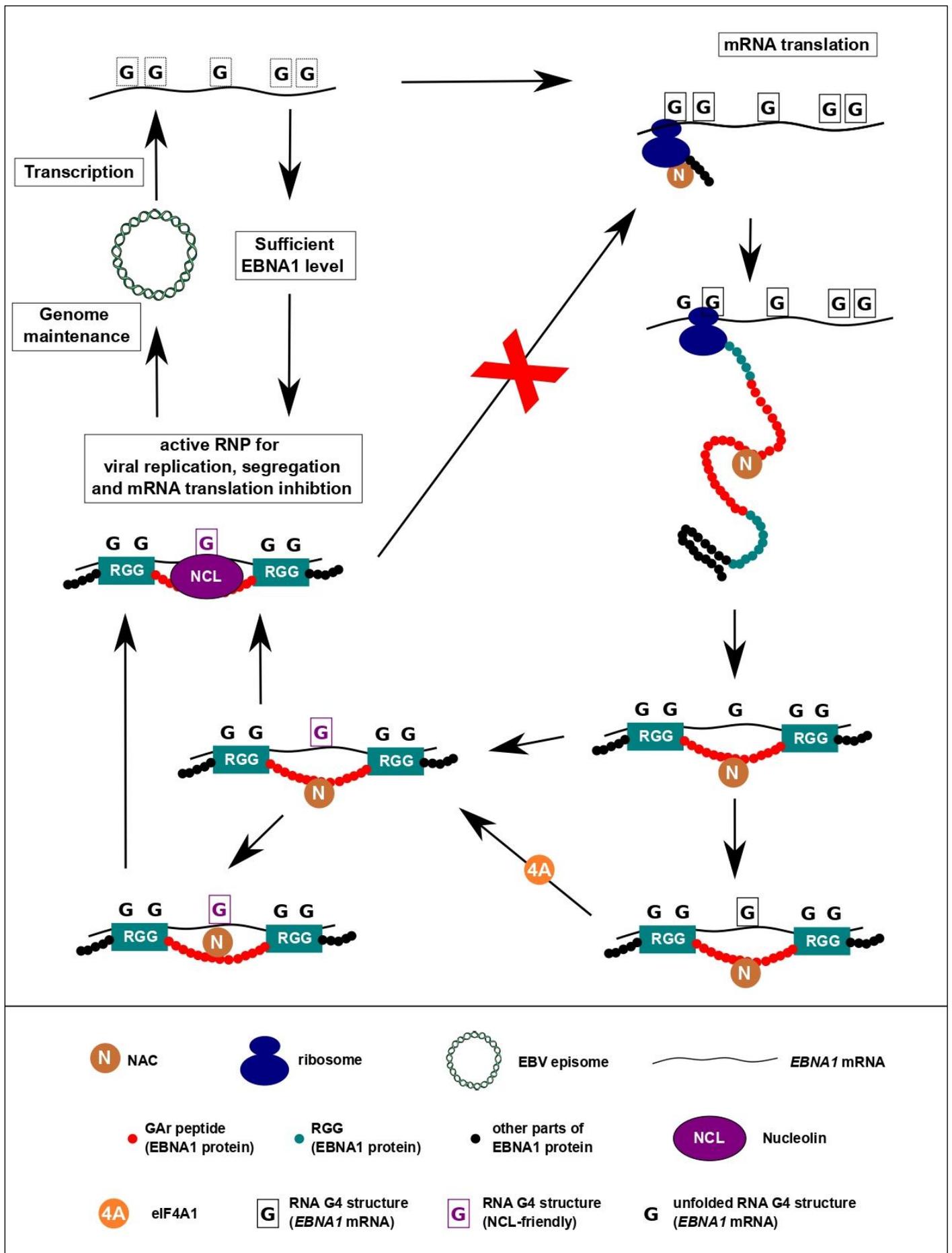


Figure K: Proposed model for EBNA1 mRNA translation regulation.

Figure K: Proposed model for *EBNA1* mRNA translation regulation (legend): When the GAR polypeptide is produced, the NAC is detached from the ribosome. Subsequently or co-translationally, the EBNA1 protein binds its own mRNA on RNA G4-forming sequences, which could be either in the GAR- or in the nearby RGG-encoding sequences. Following this event, RNA G4s in the GAR-encoding mRNA are folded. It is still unclear how this happens, but G4s are thermodynamically favoured in cellular environments. These newly formed RNA G4s could be NCL-friendly or not. We hypothesised that eIF4A1 turns non-NCL-friendly RNA G4s into NCL-friendly RNA G4s, and that NAC stabilises them. NCL can then bind to this NCL-friendly G4 structures. Newly synthesised *EBNA1* mRNA are trapped in the NCL – EBNA1 RNP complex if EBNA1 protein levels are high enough.

II – Controlling intrinsically disordered protein synthesis with the NAC

Our model suggests that a pathway exists to protect intrinsically disordered protein like EBNA1 from aggregation in the cells. The NAC, which is sorting the different nascent chains and decides their future interactions, could be the factor sorting the intrinsically disordered proteins and targeting them to a specific translation regulation pathway in the cells. In the context of the GAR, the protein is produced at sufficient levels to fulfil its function, but at low levels to be manageable for the cells to avoid their aggregation. Interestingly, this seems to be linked to the presence of RNA G4s in the mRNA and/or the ability of the protein to binds its own mRNA. Thus, RNA G4s could serve as chaperones for intrinsically disordered proteins and by doing so, be involved in both their translation regulation and aggregation prevention.

The NAC is a heterodimer, and our studies do not give any information about BTF3, the second subunit of NAC. BTF3 is needed for the interaction of NAC with the ribosome, and the ribosome-binding domain of BTF3 also has a chaperone function when NAC is unbound to ribosomes. Contrary to NACA, BTF3 has not been shown to interact with any nucleic acids. Its transcription factor activity is linked to its stable interaction with RNA polymerase II, which enhances transcription initiation. As NACA, BTF3 also senses nascent polypeptides during their synthesis. Moreover, BTF3 covers the NACA nucleic acid-binding domain. We hypothesised that during GAR synthesis, the NACA is detached from BTF3 by the GAR, and thus the DNA/RNA-binding domain of NACA would be available for binding to the RNA. However, because NACA binding on RNA is not sequence-specific and knowing that NAC interactions with the ribosome are transient and weak, we now favour the idea that the whole NAC is detached from the ribosome, although none of our data can answer to this question. Furthermore, it is also possible that other NAC binds the GAR, in addition to the one detached from the translating ribosome, as the NAC on its own is a chaperone (219).

Surprisingly, NACA RNA-binding activity, although not RNA G4 specific, is suppressed in presence of PhenDH2 during RNA pulldown experiments. Although the specificity of PhenDH2 towards RNA G4s among RNA oligos is well defined, the activity of this component on proteins has not been tested.

IDRs (intrinsically disordered regions) are common and necessary in cells since they are key elements in cellular factors interactions (267), and aggregates themselves have been suggested to have a functional role in healthy cells, as a means to store mRNAs and proteins (268). In addition,

IDR structures and functions depend on their binding partners. Therefore, the regulation of protein with IDR in the cells is a challenge, as, like for EBNA1, enough proteins must be produced to ensure their functions, but the production should also be low enough to avoid improper interactions with other factors or uncontrolled aggregate formation. Interestingly, half of RNA binding domains are in RBP-encoding mRNAs, suggesting an auto-regulation of these factors production (269). In this context, the regulation pathway dependent on the NACA and RNA G4 structures could be a way cells found to regulate the synthesis of essential but also dangerous intrinsically disordered proteins.

We can then wonder if some latent viruses like human herpesviruses, which tightly control viral proteostasis, do not slightly impair the cell proteostasis, and by doing so break the proteome balance of the cells. In neurons, which are highly differentiated cells, aggregate cannot be diluted by cell division (195). The debate on the association of herpesviruses with neurodegenerative diseases is still ongoing, and no clear evidence has been found yet (270–274).

III – Multifunctional mRNAs regulated by their G4 structures

RNA G4 structures are present in the human genome and transcriptome, and as the DNA serves as the template for RNA synthesis, G4-forming sequences must be at specific points. It is particularly clear in viruses, for which the genetic information is encoded in rather limited genome lengthwise. In HIV-1, the LTR promoter can form G4 structures, both in the proviral DNA and in the genomic RNA. In the DNA, the formation of the G4 structures in the LTR promoter U3 region prevents Sp1 binding, and thus regulates the viral gene expression (275). In the RNA version of the LTR promoter, RNA G4s can also fold and promote recombination of the two genomic strands of HIV-1 (106, 110). With EBV, the G4-ligand berberine can impact EBNA1 expression both on a transcriptional and a translational level (276). Hence, isolating and studying RNA G4 structures effects in the cells can be complicated.

Nevertheless, RNA G4s have a function in the cells and interact with different cellular factors. The dynamic nature of G4s allows a single RNA molecule to adapt its function to the cellular context. In our case, we propose that RNA G4s first serve as a chaperone for the nascent aggregate-prone chains, and at the same time allow the formation of an RNP between the *EBNA1* mRNA and the protein. This event triggers modification of the structure, hypothetically by the NAC or eIF4A1, and allows NCL binding. Thus, one RNA molecule structure is adaptable, and the different conformations taken by the RNA permit it to fulfill different functions in the cells, following the cellular cues.

This idea that RNA G4 structures are adaptable to different stimuli could have implications in other cellular contexts. As RNA G4 structures are sensitive to K⁺ concentrations, the electric stimuli in neurons coupled with ions movement could impact the folding and unfolding of these structures. As the G4s structures are present in around 30% of neurite-targeted mRNAs, local variations in K⁺ ions concentrations in the neurons can impact the RNA G4 structures, and the mRNAs can dynamically respond to the external stimuli send to the neuron cells, by moving to a specific location in the neuron and being translated at a specific time, depending on the RBP binding to the folded or unfolded G4 (80).

IV – eIF4A1 as an RNA structure modulator, impacting RNP complexes composition

eIF4A1 is mostly known and studied for its role as a translation initiation factor. However, as a member of the DEAD-box RNA helicase family, eIF4A1 RNA helicase activity is non-processive. Its helicase activity depends greatly on its binding partners (6, 8). Furthermore, eIF4A1 and other DEAD-box proteins can unwind locally structures (148, 277). Other helicases from the DEAD or DEAH-box family have been shown to remodel RNP complexes (278–280). DEAD-box proteins control membrane-less organelles formation and composition, by controlling the RNAs flux in and out (281), or by acting as an ATP-dependent RNA chaperone, avoiding RNA aggregation by controlling the RNA structures (282).

Although we do not have strong evidence here, we suggest that eIF4A specific RNA-binding activity (148) and its ability to modify locally RNA structure, rather than its unwinding activity, is involved in the GAR-mediated pathway. However, our data do not exclude the possibility that eIF4A1 acts as an effector of this inhibition pathway at the initiation step.

Inhibiting eIF4A using chemical compounds leads to the decrease of the translation of a pool of mRNAs (24, 152). However, what is the common feature between the mRNAs translated in an eIF4A-dependent manner is still unclear, as one group found that eIF4A-dependent translation is associated with the presence of RNA G4s in the 5'UTR (24), whereas another group demonstrated that RNA structures in general, and not specifically RNA G4s, cause this eIF4A dependence for translation (152). As eIF4A activity greatly depends on its co-factors, we can speculate that depending on the status of the cell, eIF4A would preferentially unwind different pools of mRNAs, with different effects on the mRNA future processing.

V – Diversity in the translation machinery

The GAR inhibition of translation initiation is efficient most of the time, but in the *cMyc* IRES-mediated translation, the GAR inhibition mechanism is not visible. This supports the idea that the translational process relies on diverse translational machinery. We could then wonder if EBNA1 affects ribogenesis and production of translation factors, favouring the ones targeted by the GAR-mediated pathway.

In line with this idea, EBNA1 interacts with factors that have been associated with ribogenesis regulation. NCL is involved in many steps of ribosomal biogenesis (128). RNA G4 structures interact with different ribosomal proteins (133). In yeast, NAC and its co-chaperone Ssb knockdown impaired mainly ribogenesis (215), and the yeast homolog of the hEBP2 is required for pre-rRNA processing and ribosomal subunit assembly (283, 284). Although this is only speculative, this idea could be tested nowadays by using comparative proteomics techniques.

Other viruses target specific ribosomal proteins during infections. Some viruses rely on specific ribosomal proteins for viral translation, while others induce post-translational modifications of ribosomal proteins (285). Until this day, the mechanisms underlying these events are still unclear.

VI – Concluding remarks

This project revolves around proteins and mRNAs which functions in the cell vary depending on their binding partners and the cellular circumstances. The case studied here is one example of protein moonlighting, *i.e.* the fact that one polypeptide chain has more than one function in a cell. This adds another layer of complexity in living cells and depends greatly on the structures of and chemical modifications on both proteins and RNA molecules (286, 287). Hopefully, high-throughput techniques enable us to grasp a more global view of the cells nowadays and perhaps will help us to clarify the relationships between proteins and RNAs. Additionally, the mechanism studied here show that both the mRNA and the encoded polypeptide trigger pathways regulating specifically mRNA translation. This cooperation is not unique to *EBNA1* mRNA, and we hypothesise that for some multifunctional proteins, both the mRNA and the proteins co-evolved, in a way that tightly control the different functions of the protein.

Recently, some moonlighting or multifunctional proteins show that mRNA and protein biology in the cells are tightly linked. For example, Upf1 which is mainly associated to its mRNA surveillance role in NMD pathways, also plays a role in protein quality control, notably by regulating the process targeting NMD-derived protein products (288–290). Further studies are needed to better understand how the mRNA surveillance and translation pathways communicate with protein quality control pathways, and how spread this mechanism is in living beings, from a cellular to an evolutionary scale.

Viruses are small biological elements with a genome of a relatively limited size carrying all the information needed for its replication and proliferation. Thus, it is not a surprise that with such constraints, viral RNAs and proteins have varied functions depending on the cellular context, and studying them can help us better understand the different cellular pathways.

References

1. Deuerling,E., Gamedinger,M. and Kreft,S.G. (2019) Chaperone Interactions at the Ribosome. *Cold Spring Harb. Perspect. Biol.*, **11**, a033977.
2. Jackson,R.J., Hellen,C.U.T. and Pestova,T.V. (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell Biol.*, **11**, 113–127.
3. 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.
4. Lewis,C.J.T., Pan,T. and Kalsotra,A. (2017) RNA modifications and structures cooperate to guide RNA–protein interactions. *Nat. Rev. Mol. Cell Biol.*, **18**, 202–210.
5. Rogers,G.W., Richter,N.J. and Merrick,W.C. (1999) Biochemical and Kinetic Characterization of the RNA Helicase Activity of Eukaryotic Initiation Factor 4A. *J. Biol. Chem.*, **274**, 12236–12244.
6. Özeş,A.R., Feoktistova,K., Avanzino,B.C. and Fraser,C.S. (2011) Duplex Unwinding and ATPase Activities of the DEAD-Box Helicase eIF4A Are Coupled by eIF4G and eIF4B. *J. Mol. Biol.*, **412**, 674–687.
7. Rogers,G.W., Richter,N.J., Lima,W.F. and Merrick,W.C. (2001) Modulation of the Helicase Activity of eIF4A by eIF4B, eIF4H, and eIF4F. *J. Biol. Chem.*, **276**, 30914–30922.
8. Garcia-Garcia,C., Frieda,K.L., Feoktistova,K., Fraser,C.S. and Block,S.M. (2015) Factor-dependent processivity in human eIF4A DEAD-box helicase. *Science*, **348**, 1486–1488.
9. Aitken,C.E. and Lorsch,J.R. (2012) A mechanistic overview of translation initiation in eukaryotes. *Nat. Struct. Mol. Biol.*, **19**, 568–576.
10. Morris,C., Cluet,D. and Ricci,E.P. (2021) Ribosome dynamics and mRNA turnover, a complex relationship under constant cellular scrutiny. *WIREs RNA*, 10.1002/wrna.1658.
11. Kramer,G., Shiber,A. and Bukau,B. (2019) Mechanisms of Cotranslational Maturation of Newly Synthesized Proteins. *Annu. Rev. Biochem.*, **88**, 337–364.
12. Gloge,F., Becker,A.H., Kramer,G. and Bukau,B. (2014) Co-translational mechanisms of protein maturation. *Curr. Opin. Struct. Biol.*, **24**, 24–33.
13. Collart,M.A. and Weiss,B. (2020) Ribosome pausing, a dangerous necessity for co-translational events. *Nucleic Acids Res.*, **48**, 1043–1055.
14. Yang,C.-I., Hsieh,H.-H. and Shan,S. (2019) Timing and specificity of cotranslational nascent protein modification in bacteria. *Proc. Natl. Acad. Sci.*, **116**, 23050–23060.
15. Wells,J.N., Bergendahl,L.T. and Marsh,J.A. (2015) Co-translational assembly of protein complexes. *Biochem. Soc. Trans.*, **43**, 1221–1226.
16. Hsieh,H.-H., Lee,J.H., Chandrasekar,S. and Shan,S. (2020) A ribosome-associated chaperone enables substrate triage in a cotranslational protein targeting complex. *Nat. Commun.*, **11**, 5840.
17. Gamedinger,M., Hanebuth,M.A., Frickey,T. and Deuerling,E. (2015) The principle of antagonism ensures protein targeting specificity at the endoplasmic reticulum. *Science*, **348**, 201–207.

18. Ueda,T., Watanabe-Fukunaga,R., Fukuyama,H., Nagata,S. and Fukunaga,R. (2004) Mnk2 and Mnk1 Are Essential for Constitutive and Inducible Phosphorylation of Eukaryotic Initiation Factor 4E but Not for Cell Growth or Development. *Mol. Cell. Biol.*, **24**, 6539–6549.
19. Kearsse,M.G. and Wilusz,J.E. (2017) Non-AUG translation: a new start for protein synthesis in eukaryotes. *Genes Dev.*, **31**, 1717–1731.
20. Robert,F., Cencic,R., Cai,R., Schmeing,T.M. and Pelletier,J. (2020) RNA-tethering assay and eIF4G:eIF4A obligate dimer design uncovers multiple eIF4F functional complexes. *Nucleic Acids Res.*, **48**, 8562–8575.
21. Rozovsky,N., Butterworth,A.C. and Moore,M.J. (2008) Interactions between eIF4A1 and its accessory factors eIF4B and eIF4H. *RNA*, **14**, 2136–2148.
22. Rubio,C.A., Weisburd,B., Holderfield,M., Arias,C., Fang,E., DeRisi,J.L. and Fanidi,A. (2014) Transcriptome-wide characterization of the eIF4A signature highlights plasticity in translation regulation.
23. Waldron,J.A., Tack,D.C., Ritchey,L.E., Gillen,S.L., Wilczynska,A., Turro,E., Bevilacqua,P.C., Assmann,S.M., Bushell,M. and Le Quesne,J. (2019) mRNA structural elements immediately upstream of the start codon dictate dependence upon eIF4A helicase activity. *Genome Biol.*, **20**, 300.
24. Wolfe,A.L., Singh,K., Zhong,Y., Drewe,P., Rajasekhar,V.K., Sanghvi,V.R., Mavrakis,K.J., Jiang,M., Roderick,J.E., Van der Meulen,J., *et al.* (2014) RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer. *Nature*, **513**, 65–70.
25. Ishikawa,C., Tanaka,J., Katano,H., Senba,M. and Mori,N. (2013) Hippuristanol Reduces the Viability of Primary Effusion Lymphoma Cells both in Vitro and in Vivo. *Mar. Drugs*, **11**, 3410–3424.
26. Steinberger,J., Shen,L., J. Kiniry,S., Naineni,S.K., Cencic,R., Amiri,M., Aboushawareb,S.A.E., Chu,J., Maïga,R.I., Yachnin,B.J., *et al.* (2020) Identification and characterization of hippuristanol-resistant mutants reveals eIF4A1 dependencies within mRNA 5' leader regions. *Nucleic Acids Res.*, **48**, 9521–9537.
27. Sen,N.D., Zhou,F., Harris,M.S., Ingolia,N.T. and Hinnebusch,A.G. (2016) eIF4B stimulates translation of long mRNAs with structured 5' UTRs and low closed-loop potential but weak dependence on eIF4G. *Proc. Natl. Acad. Sci.*, **113**, 10464–10472.
28. Lee,A.S.Y., Kranzusch,P.J., Doudna,J.A. and Cate,J.H.D. (2016) eIF3d is an mRNA cap-binding protein that is required for specialized translation initiation. *Nature*, **536**, 96–99.
29. Ishigaki,Y., Li,X., Serin,G. and Maquat,L.E. Evidence for a Pioneer Round of mRNA Translation: mRNAs Subject to Nonsense-Mediated Decay in Mammalian Cells Are Bound by CBP80 and CBP20.
30. Lejeune,F., Ranganathan,A.C. and Maquat,L.E. (2004) eIF4G is required for the pioneer round of translation in mammalian cells. *Nat. Struct. Mol. Biol.*, **11**, 992–1000.
31. Chiu,S.-Y. (2004) The pioneer translation initiation complex is functionally distinct from but structurally overlaps with the steady-state translation initiation complex. *Genes Dev.*, **18**, 745–754.
32. Park,Y., Park,J., Hwang,H.J., Kim,L., Jeong,K., Song,H.K., Rufener,S.C., Mühlemann,O. and Kim,Y.K. (2021) Translation mediated by the nuclear cap-binding complex is confined to the perinuclear region via a CTIF–DDX19B interaction. *Nucleic Acids Res.*, **49**, 8261–8276.

33. Yang,Y. and Wang,Z. (2019) IRES-mediated cap-independent translation, a path leading to hidden proteome. *J. Mol. Cell Biol.*, **11**, 911–919.
34. Kwan,T. and Thompson,S.R. (2019) Noncanonical Translation Initiation in Eukaryotes. *Cold Spring Harb. Perspect. Biol.*, **11**, a032672.
35. Jaafar,Z.A. and Kieft,J.S. (2019) Viral RNA structure-based strategies to manipulate translation. *Nat. Rev. Microbiol.*, **17**, 110–123.
36. Cobbold,L.C., Spriggs,K.A., Haines,S.J., Dobbyn,H.C., Hayes,C., de Moor,C.H., Lilley,K.S., Bushell,M. and Willis,A.E. (2008) Identification of Internal Ribosome Entry Segment (IRES)- *trans* -Acting Factors for the Myc Family of IRESs. *Mol. Cell. Biol.*, **28**, 40–49.
37. Meyer,K.D., Patil,D.P., Zhou,J., Zinoviev,A., Skabkin,M.A., Elemento,O., Pestova,T.V., Qian,S.-B. and Jaffrey,S.R. (2015) 5' UTR m6A Promotes Cap-Independent Translation. *Cell*, **163**, 999–1010.
38. Coots,R.A., Liu,X.-M., Mao,Y., Dong,L., Zhou,J., Wan,J., Zhang,X. and Qian,S.-B. (2017) m6A Facilitates eIF4F-Independent mRNA Translation. *Mol. Cell*, **68**, 504-514.e7.
39. Jiang,X., Liu,B., Nie,Z., Duan,L., Xiong,Q., Jin,Z., Yang,C. and Chen,Y. (2021) The role of m6A modification in the biological functions and diseases. *Signal Transduct. Target. Ther.*, **6**, 74.
40. Genuth,N.R. and Barna,M. (2018) Heterogeneity and specialized functions of translation machinery: from genes to organisms. *Nat. Rev. Genet.*, **19**, 431–452.
41. Xue,S. and Barna,M. (2012) Specialized ribosomes: a new frontier in gene regulation and organismal biology. *Nat. Rev. Mol. Cell Biol.*, **13**, 355–369.
42. Shi,Z., Fujii,K., Kovary,K.M., Genuth,N.R., Röst,H.L., Teruel,M.N. and Barna,M. (2017) Heterogeneous Ribosomes Preferentially Translate Distinct Subpools of mRNAs Genome-wide. *Mol. Cell*, **67**, 71-83.e7.
43. Shi,Z. and Barna,M. (2015) Translating the Genome in Time and Space: Specialized Ribosomes, RNA Regulons, and RNA-Binding Proteins. *Annu. Rev. Cell Dev. Biol.*, **31**, 31–54.
44. Holt,C.E., Martin,K.C. and Schuman,E.M. (2019) Local translation in neurons: visualization and function. *Nat. Struct. Mol. Biol.*, **26**, 557–566.
45. Rangaraju,V., tom Dieck,S. and Schuman,E.M. (2017) Local translation in neuronal compartments: how local is local? *EMBO Rep.*, **18**, 693–711.
46. Ifrim,M.F., Williams,K.R. and Bassell,G.J. (2015) Single-Molecule Imaging of PSD-95 mRNA Translation in Dendrites and Its Dysregulation in a Mouse Model of Fragile X Syndrome. *J. Neurosci.*, **35**, 7116–7130.
47. Mayford,M., Baranes,D., Podsypanina,K. and Kandel,E.R. (1996) The 3'-untranslated region of CaMKII is a cis-acting signal for the localization and translation of mRNA in dendrites. *Proc. Natl. Acad. Sci.*, **93**, 13250–13255.
48. Taylor,A.M., Wu,J., Tai,H.-C. and Schuman,E.M. (2013) Axonal Translation of -Catenin Regulates Synaptic Vesicle Dynamics. *J. Neurosci.*, **33**, 5584–5589.

49. Baj,G., Leone,E., Chao,M.V. and Tongiorgi,E. (2011) Spatial segregation of BDNF transcripts enables BDNF to differentially shape distinct dendritic compartments. *Proc. Natl. Acad. Sci.*, **108**, 16813–16818.
50. Wang,E., Thombre,R., Shah,Y., Latanich,R. and Wang,J. (2021) G-Quadruplexes as pathogenic drivers in neurodegenerative disorders. *Nucleic Acids Res.*, **49**, 4816–4830.
51. Athar,Y.M. and Joseph,S. (2020) The Human Fragile X Mental Retardation Protein Inhibits the Elongation Step of Translation through Its RGG and C-Terminal Domains. *Biochemistry*, **59**, 3813–3822.
52. Martin,K.C., Casadio,A., Zhu,H., E,Y., Rose,J.C., Chen,M., Bailey,C.H. and Kandel,E.R. (1997) Synapse-Specific, Long-Term Facilitation of Aplysia Sensory to Motor Synapses: A Function for Local Protein Synthesis in Memory Storage. *Cell*, **91**, 927–938.
53. Leung,K.-M., van Horck,F.P., Lin,A.C., Allison,R., Standart,N. and Holt,C.E. (2006) Asymmetrical β -actin mRNA translation in growth cones mediates attractive turning to netrin-1. *Nat. Neurosci.*, **9**, 1247–1256.
54. De Benedetti,A. and Graff,J.R. (2004) eIF-4E expression and its role in malignancies and metastases. *Oncogene*, **23**, 3189–3199.
55. Cerezo,M., Guemiri,R., Druillennec,S., Girault,I., Malka-Mahieu,H., Shen,S., Allard,D., Martineau,S., Welsch,C., Agoussi,S., *et al.* (2018) Translational control of tumor immune escape via the eIF4F–STAT1–PD-L1 axis in melanoma. *Nat. Med.*, **24**, 1877–1886.
56. Robichaud,N., del Rincon,S.V., Huor,B., Alain,T., Petruccelli,L.A., Hearnden,J., Goncalves,C., Grotegut,S., Spruck,C.H., Furic,L., *et al.* (2015) Phosphorylation of eIF4E promotes EMT and metastasis via translational control of SNAIL and MMP-3. *Oncogene*, **34**, 2032–2042.
57. Jan,E., Mohr,I. and Walsh,D. (2016) A Cap-to-Tail Guide to mRNA Translation Strategies in Virus-Infected Cells. *Annu. Rev. Virol.*, **3**, 283–307.
58. Avanzino,B.C., Fuchs,G. and Fraser,C.S. (2017) Cellular cap-binding protein, eIF4E, promotes picornavirus genome restructuring and translation. *Proc. Natl. Acad. Sci.*, **114**, 9611–9616.
59. Kempf,B.J. and Barton,D.J. (2008) Poliovirus 2A^{Pro} Increases Viral mRNA and Polysome Stability Coordinately in Time with Cleavage of eIF4G. *J. Virol.*, **82**, 5847–5859.
60. Schubert,K., Karousis,E.D., Jomaa,A., Scaiola,A., Echeverria,B., Gurzeler,L.-A., Leibundgut,M., Thiel,V., Mühlemann,O. and Ban,N. (2020) SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation. *Nat. Struct. Mol. Biol.*, **27**, 959–966.
61. Jagger,B.W., Wise,H.M., Kash,J.C., Walters,K.-A., Wills,N.M., Xiao,Y.-L., Dunfee,R.L., Schwartzman,L.M., Ozinsky,A., Bell,G.L., *et al.* (2012) An Overlapping Protein-Coding Region in Influenza A Virus Segment 3 Modulates the Host Response. *Science*, **337**, 199–204.
62. Gaucherand,L., Porter,B.K., Levene,R.E., Price,E.L., Schmalig,S.K., Rycroft,C.H., Kevorkian,Y., McCormick,C., Khapersky,D.A. and Gaglia,M.M. (2019) The Influenza A Virus Endoribonuclease PA-X Usurps Host mRNA Processing Machinery to Limit Host Gene Expression. *Cell Rep.*, **27**, 776–792.e7.
63. Levene,R.E., Shrestha,S.D. and Gaglia,M.M. (2021) The Influenza A Virus Host Shutoff Factor PA-X Is Rapidly Turned Over in a Strain-Specific Manner. *J. Virol.*, **95**.

64. 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.
65. Mirihana Arachchilage,G., Kharel,P., Reid,J. and Basu,S. (2018) Targeting of G-Quadruplex Harboring Pre-miRNA 92b by LNA Rescues PTEN Expression in NSCL Cancer Cells. *ACS Chem. Biol.*, **13**, 909–914.
66. Sanchez de Groot,N., Armaos,A., Graña-Montes,R., Alriquet,M., Calloni,G., Vabulas,R.M. and Tartaglia,G.G. (2019) RNA structure drives interaction with proteins. *Nat. Commun.*, **10**, 3246.
67. Serganov,A. and Patel,D.J. (2012) Molecular recognition and function of riboswitches. *Curr. Opin. Struct. Biol.*, **22**, 279–286.
68. Ruggiero,E. and Richter,S.N. (2018) G-quadruplexes and G-quadruplex ligands: targets and tools in antiviral therapy. *Nucleic Acids Res.*, **46**, 3270–3283.
69. 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.
70. Fay,M.M., Lyons,S.M. and Ivanov,P. (2017) RNA G-Quadruplexes in Biology: Principles and Molecular Mechanisms. *J. Mol. Biol.*, **429**, 2127–2147.
71. Dumas,L., Herviou,P., Dassi,E., Cammas,A. and Millevoi,S. (2021) G-Quadruplexes in RNA Biology: Recent Advances and Future Directions. *Trends Biochem. Sci.*, **46**, 270–283.
72. Zhang,D.-H., Fujimoto,T., Saxena,S., Yu,H.-Q., Miyoshi,D. and Sugimoto,N. (2010) Monomorphic RNA G-Quadruplex and Polymorphic DNA G-Quadruplex Structures Responding to Cellular Environmental Factors. *Biochemistry*, **49**, 4554–4563.
73. Arora,A. and Maiti,S. (2009) Differential Biophysical Behavior of Human Telomeric RNA and DNA Quadruplex. *J. Phys. Chem. B*, **113**, 10515–10520.
74. Joachimi,A., Benz,A. and Hartig,J.S. (2009) A comparison of DNA and RNA quadruplex structures and stabilities. *Bioorg. Med. Chem.*, **17**, 6811–6815.
75. Zaccaria,F. and Fonseca Guerra,C. (2018) RNA versus DNA G-Quadruplex: The Origin of Increased Stability. *Chem. - Eur. J.*, **24**, 16315–16322.
76. Matsumoto,S., Tateishi-Karimata,H., Takahashi,S., Ohyama,T. and Sugimoto,N. (2020) Effect of Molecular Crowding on the Stability of RNA G-Quadruplexes with Various Numbers of Quartets and Lengths of Loops. *Biochemistry*, **59**, 2640–2649.
77. Zhang,A.Y.Q., Bugaut,A. and Balasubramanian,S. (2011) A Sequence-Independent Analysis of the Loop Length Dependence of Intramolecular RNA G-Quadruplex Stability and Topology. *Biochemistry*, **50**, 7251–7258.
78. Kwok,C.K., Marsico,G., Sahakyan,A.B., Chambers,V.S. and Balasubramanian,S. (2016) rG4-seq reveals widespread formation of G-quadruplex structures in the human transcriptome. *Nat. Methods*, **13**, 841–844.
79. Reznichenko,O., Quillévéré,A., Martins,R.P., Loaëc,N., Kang,H., Lista,M.J., Beauvineau,C., González-García,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.
80. Subramanian,M., Rage,F., Tabet,R., Flatter,E., Mandel,J. and Moine,H. (2011) G–quadruplex RNA structure as a signal for neurite mRNA targeting. *EMBO Rep.*, **12**, 697–704.
81. Ishiguro,A., Kimura,N., Watanabe,Y., Watanabe,S. and Ishihama,A. (2016) TDP-43 binds and transports G-quadruplex-containing mRNAs into neurites for local translation. *Genes Cells*, **21**, 466–481.
82. Fratta,P., Mizielińska,S., Nicoll,A.J., Zloh,M., Fisher,E.M.C., 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.
83. Reddy,K., Zamiri,B., Stanley,S.Y.R., Macgregor,R.B. and Pearson,C.E. (2013) The Disease-associated r(GGGGCC) Repeat from the C9orf72 Gene Forms Tract Length-dependent Uni- and Multimolecular RNA G-quadruplex Structures. *J. Biol. Chem.*, **288**, 9860–9866.
84. Conlon,E.G., Lu,L., Sharma,A., Yamazaki,T., Tang,T., Shneider,N.A. and Manley,J.L. (2016) The C9ORF72 GGGGCC expansion forms RNA G-quadruplex inclusions and sequesters hnRNP H to disrupt splicing in ALS brains. *eLife*, **5**, e17820.
85. Decorsiere,A., Cayrel,A., Vagner,S. and Millevoi,S. (2011) Essential role for the interaction between hnRNP H/F and a G quadruplex in maintaining p53 pre-mRNA 3'-end processing and function during DNA damage. *Genes Dev.*, **25**, 220–225.
86. Newman,M., Sfaxi,R., Saha,A., Monchaud,D., Teulade-Fichou,M.-P. and Vagner,S. (2017) The G-Quadruplex-Specific RNA Helicase DHX36 Regulates p53 Pre-mRNA 3'-End Processing Following UV-Induced DNA Damage. *J. Mol. Biol.*, **429**, 3121–3131.
87. 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.
88. Fiset,J.-F., Montagna,D.R., Mihailescu,M.-R. and Wolfe,M.S. (2012) A G-Rich element forms a G-quadruplex and regulates BACE1 mRNA alternative splicing: G-rich element regulates BACE1 alternative splicing. *J. Neurochem.*, **121**, 763–773.
89. Fabian,M.R., Sonenberg,N. and Filipowicz,W. (2010) Regulation of mRNA Translation and Stability by microRNAs. *Annu. Rev. Biochem.*, **79**, 351–379.
90. Stefanovic,S., Bassell,G.J. and Mihailescu,M.R. (2015) G quadruplex RNA structures in PSD-95 mRNA: potential regulators of miR-125a seed binding site accessibility. *RNA*, **21**, 48–60.
91. Kumari,S., Bugaut,A., Huppert,J.L. and Balasubramanian,S. (2007) An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation. *Nat. Chem. Biol.*, **3**, 218–221.
92. Miglietta,G., Cogoi,S., Marinello,J., Capranico,G., Tikhomirov,A.S., Shchekotikhin,A. and Xodo,L.E. (2017) RNA G-Quadruplexes in Kirsten Ras (KRAS) Oncogene as Targets for Small Molecules Inhibiting Translation. *J. Med. Chem.*, **60**, 9448–9461.
93. Cammas,A., Lacroix-Triki,M., Pierredon,S., Le Bras,M., Iacovoni,J.S., Teulade-Fichou,M.-P., Favre,G., Roché,H., Filleron,T., Millevoi,S., *et al.* (2016) hnRNP A1-mediated translational regulation

of the G quadruplex-containing RON receptor tyrosine kinase mRNA linked to tumor progression. *Oncotarget*, **7**, 16793–16805.

94. Crenshaw,E., Leung,B.P., Kwok,C.K., Sharoni,M., Olson,K., Sebastian,N.P., Ansaloni,S., Schweitzer-Stenner,R., Akins,M.R., Bevilacqua,P.C., *et al.* (2015) Amyloid Precursor Protein Translation Is Regulated by a 3'UTR Guanine Quadruplex. *PLOS ONE*, **10**, e0143160.

95. Lyu,K., Chen,S.-B., Chan,C.-Y., Tan,J.-H. and Kwok,C.K. (2019) Structural analysis and cellular visualization of APP RNA G-quadruplex. *Chem. Sci.*, **10**, 11095–11102.

96. Endoh,T., Kawasaki,Y. and Sugimoto,N. (2013) Suppression of Gene Expression by G-Quadruplexes in Open Reading Frames Depends on G-Quadruplex Stability. *Angew. Chem. Int. Ed.*, **52**, 5522–5526.

97. Endoh,T., Kawasaki,Y. and Sugimoto,N. (2013) Translational halt during elongation caused by G-quadruplex formed by mRNA. *Methods*, **64**, 73–78.

98. Westmark,C.J. and Malter,J.S. (2007) FMRP Mediates mGluR5-Dependent Translation of Amyloid Precursor Protein. *PLoS Biol.*, **5**, e52.

99. Lee,E.K., Kim,H.H., Kuwano,Y., Abdelmohsen,K., Srikantan,S., Subaran,S.S., Gleichmann,M., Mughal,M.R., Martindale,J.L., Yang,X., *et al.* (2010) hnRNP C promotes APP translation by competing with FMRP for APP mRNA recruitment to P bodies. *Nat. Struct. Mol. Biol.*, **17**, 732–739.

100. Endoh,T. and Sugimoto,N. (2013) Unusual –1 Ribosomal Frameshift Caused by Stable RNA G-Quadruplex in Open Reading Frame. *Anal. Chem.*, **85**, 11435–11439.

101. Yu,C.-H., Teulade-Fichou,M.-P. and Olsthoorn,R.C.L. (2014) Stimulation of ribosomal frameshifting by RNA G-quadruplex structures. *Nucleic Acids Res.*, **42**, 1887–1892.

102. Endoh,T., Kawasaki,Y. and Sugimoto,N. (2013) Stability of RNA quadruplex in open reading frame determines proteolysis of human estrogen receptor α . *Nucleic Acids Res.*, **41**, 6222–6231.

103. Tlučková,K., Marušič,M., Tóthová,P., Bauer,L., Šket,P., Plavec,J. and Viglasky,V. (2013) Human Papillomavirus G-Quadruplexes. *Biochemistry*, **52**, 7207–7216.

104. Biswas,B., Kandpal,M., Jauhari,U.K. and Vivekanandan,P. (2016) Genome-wide analysis of G-quadruplexes in herpesvirus genomes. *BMC Genomics*, **17**, 949.

105. Lavezzo,E., Berselli,M., Frasson,I., Perrone,R., Palù,G., Brazzale,A.R., Richter,S.N. and Toppo,S. (2018) G-quadruplex forming sequences in the genome of all known human viruses: A comprehensive guide. *PLOS Comput. Biol.*, **14**, e1006675.

106. Piekna-Przybylska,D., Sullivan,M.A., Sharma,G. and Bambara,R.A. (2014) U3 Region in the HIV-1 Genome Adopts a G-Quadruplex Structure in Its RNA and DNA Sequence. *Biochemistry*, **53**, 2581–2593.

107. Perrone,R., Nadai,M., Poe,J.A., Frasson,I., Palumbo,M., Palù,G., Smithgall,T.E. and Richter,S.N. (2013) Formation of a Unique Cluster of G-Quadruplex Structures in the HIV-1 nef Coding Region: Implications for Antiviral Activity. *PLoS ONE*, **8**, e73121.

108. Sundquist,W.I. and Heaphy,S. (1993) Evidence for interstrand quadruplex formation in the dimerization of human immunodeficiency virus 1 genomic RNA. *Proc. Natl. Acad. Sci.*, **90**, 3393–3397.
109. Shen,W., Gao,L., Balakrishnan,M. and Bambara,R.A. (2009) A Recombination Hot Spot in HIV-1 Contains Guanosine Runs That Can Form a G-quartet Structure and Promote Strand Transfer in Vitro. *J. Biol. Chem.*, **284**, 33883–33893.
110. Piekna-Przybylska,D., Sharma,G. and Bambara,R.A. (2013) Mechanism of HIV-1 RNA Dimerization in the Central Region of the Genome and Significance for Viral Evolution. *J. Biol. Chem.*, **288**, 24140–24150.
111. Marquet,R., Baudin,F., Gabus,C., Darlix,J.-L., Mougél,M., Ehresmann,C. and Ehresmann,B. (1991) Dimerization of human immunodeficiency virus (type 1) RNA: stimulation by cations and possible mechanism. *Nucleic Acids Res.*, **19**, 2349–2357.
112. Marquet,R., Paillart,J.-C., Skripkin,E., Ehresmann,C. and Ehresmann,B. (1994) Dimerization of human immunodeficiency virus type 1 RNA involves sequences located upstream of the splice donor site. *Nucleic Acids Res.*, **22**, 145–151.
113. Fleming,A.M., Ding,Y., Alenko,A. and Burrows,C.J. (2016) Zika Virus Genomic RNA Possesses Conserved G-Quadruplexes Characteristic of the Flaviviridae Family. *ACS Infect. Dis.*, **2**, 674–681.
114. Wang,S.-R., Min,Y.-Q., Wang,J.-Q., Liu,C.-X., Fu,B.-S., Wu,F., Wu,L.-Y., Qiao,Z.-X., Song,Y.-Y., Xu,G.-H., *et al.* (2016) A highly conserved G-rich consensus sequence in hepatitis C virus core gene represents a new anti-hepatitis C target. *Sci. Adv.*, **2**, e1501535.
115. Jaubert,C., Bedrat,A., Bartolucci,L., Di Primo,C., Ventura,M., Mergny,J.-L., Amrane,S. and Andreola,M.-L. (2018) RNA synthesis is modulated by G-quadruplex formation in Hepatitis C virus negative RNA strand. *Sci. Rep.*, **8**, 8120.
116. Wang,S.-R., Zhang,Q.-Y., Wang,J.-Q., Ge,X.-Y., Song,Y.-Y., Wang,Y.-F., Li,X.-D., Fu,B.-S., Xu,G.-H., Shu,B., *et al.* (2016) Chemical Targeting of a G-Quadruplex RNA in the Ebola Virus L Gene. *Cell Chem. Biol.*, **23**, 1113–1122.
117. Kwun,H.J., da Silva,S.R., Shah,I.M., Blake,N., Moore,P.S. and Chang,Y. (2007) Kaposi's Sarcoma-Associated Herpesvirus Latency-Associated Nuclear Antigen 1 Mimics Epstein-Barr Virus EBNA1 Immune Evasion through Central Repeat Domain Effects on Protein Processing. *J. Virol.*, **81**, 8225–8235.
118. Dabral,P., Babu,J., Zareie,A. and Verma,S.C. (2020) LANA and hnRNP A1 Regulate the Translation of LANA mRNA through G-Quadruplexes. *J. Virol.*, **94**.
119. Castello,A., Fischer,B., Eichelbaum,K., Horos,R., Beckmann,B.M., Strein,C., Davey,N.E., Humphreys,D.T., Preiss,T., Steinmetz,L.M., *et al.* (2012) Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. *Cell*, **149**, 1393–1406.
120. Yadav,M., Singh,R.S., Hogan,D., Vidhyasagar,V., Yang,S., Chung,I.Y.W., Kusalik,A., Dmitriev,O.Y., Cygler,M. and Wu,Y. (2021) The KH domain facilitates the substrate specificity and unwinding processivity of DDX43 helicase. *J. Biol. Chem.*, **296**, 100085.

121. Dominguez,C. (2006) NMR structure of the three quasi RNA recognition motifs (qRRMs) of human hnRNP F and interaction studies with Bcl-x G-tract RNA: a novel mode of RNA recognition. *Nucleic Acids Res.*, **34**, 3634–3645.
122. Samatanga,B., Dominguez,C., Jelesarov,I. and Allain,F.H.-T. (2013) The high kinetic stability of a G-quadruplex limits hnRNP F qRRM3 binding to G-tract RNA. *Nucleic Acids Res.*, **41**, 2505–2516.
123. Dominguez,D., Freese,P., Alexis,M.S., Su,A., Hochman,M., Palden,T., Bazile,C., Lambert,N.J., Van Nostrand,E.L., Pratt,G.A., *et al.* (2018) Sequence, Structure, and Context Preferences of Human RNA Binding Proteins. *Mol. Cell*, **70**, 854-867.e9.
124. Tajrishi,M.M., Tuteja,R. and Tuteja,N. (2011) Nucleolin: The most abundant multifunctional phosphoprotein of nucleolus. *Commun. Integr. Biol.*, **4**, 267–275.
125. Abdelmohsen,K. and Gorospe,M. (2012) RNA-binding protein nucleolin in disease. *RNA Biol.*, **9**, 799–808.
126. 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.
127. Ginisty,H. (1998) Nucleolin functions in the first step of ribosomal RNA processing. *EMBO J.*, **17**, 1476–1486.
128. Jia,W., Yao,Z., Zhao,J., Guan,Q. and Gao,L. (2017) New perspectives of physiological and pathological functions of nucleolin (NCL). *Life Sci.*, **186**, 1–10.
129. Abdelmohsen,K., Tominaga,K., Lee,E.K., Srikantan,S., Kang,M.-J., Kim,M.M., Selimyan,R., Martindale,J.L., Yang,X., Carrier,F., *et al.* (2011) Enhanced translation by Nucleolin via G-rich elements in coding and non-coding regions of target mRNAs. *Nucleic Acids Res.*, **39**, 8513–8530.
130. Figueiredo,J., Miranda,A., Lopes-Nunes,J., Carvalho,J., Alexandre,D., Valente,S., Mergny,J.-L. and Cruz,C. (2021) Targeting nucleolin by RNA G-quadruplex-forming motif. *Biochem. Pharmacol.*, **189**, 114418.
131. 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.
132. 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.
133. Hacht,A. von, Seifert,O., Menger,M., Schütze,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.
134. Tosoni,E., Frasson,I., Scalabrin,M., Perrone,R., Butovskaya,E., Nadai,M., Palù,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.
135. 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.

136. Riechert,E., Kmietczyk,V., Stein,F., Schwarzl,T., Sekaran,T., Jürgensen,L., Kamuf-Schenk,V., Varma,E., Hofmann,C., Rettel,M., *et al.* (2021) Identification of dynamic RNA-binding proteins uncovers a Cpeb4-controlled regulatory cascade during pathological cell growth of cardiomyocytes. *Cell Rep.*, **35**, 109100.
137. 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.
138. Garcia-Moreno,M., Noerenberg,M., Ni,S., Järvelin,A.I., González-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.e11.
139. Perez-Perri,J.I., Rogell,B., Schwarzl,T., Stein,F., Zhou,Y., Rettel,M., Brosig,A. and Hentze,M.W. (2018) Discovery of RNA-binding proteins and characterization of their dynamic responses by enhanced RNA interactome capture. *Nat. Commun.*, **9**, 4408.
140. Liu,N., Dai,Q., Zheng,G., He,C., Parisien,M. and Pan,T. (2015) N6-methyladenosine-dependent RNA structural switches regulate RNA–protein interactions. *Nature*, **518**, 560–564.
141. Holmes,B., Benavides-Serrato,A., Freeman,R.S., Landon,K.A., Bashir,T., Nishimura,R.N. and Gera,J. (2018) mTORC2/AKT/HSF1/HuR constitute a feed-forward loop regulating Rictor expression and tumor growth in glioblastoma. *Oncogene*, **37**, 732–743.
142. Candeias,M.M., Malbert-Colas,L., Powell,D.J., Daskalogianni,C., Maslon,M.M., Naski,N., Bourougaa,K., Calvo,F. and Fåhræus,R. (2008) p53 mRNA controls p53 activity by managing Mdm2 functions. *Nat. Cell Biol.*, **10**, 1098–1105.
143. Naski,N., Gajjar,M., Bourougaa,K., Malbert-Colas,L., Fåhræus,R. and Candeias,M.M. (2009) The p53 mRNA-Mdm2 interaction. *Cell Cycle*, **8**, 31–34.
144. Guo,J.U. and Bartel,D.P. (2016) RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. *Science*, **353**, aaf5371–aaf5371.
145. Laguerre,A., Hukezalie,K., Winckler,P., Katranji,F., Chanteloup,G., Pirrotta,M., Perrier-Cornet,J.-M., Wong,J.M.Y. and Monchaud,D. (2015) Visualization of RNA-Quadruplexes in Live Cells. *J. Am. Chem. Soc.*, **137**, 8521–8525.
146. Laguerre,A., Wong,J.M.Y. and Monchaud,D. (2016) Direct visualization of both DNA and RNA quadruplexes in human cells via an uncommon spectroscopic method. *Sci. Rep.*, **6**, 32141.
147. Xu,S., Li,Q., Xiang,J., Yang,Q., Sun,H., Guan,A., Wang,L., Liu,Y., Yu,L., Shi,Y., *et al.* (2015) Directly lighting up RNA G-quadruplexes from test tubes to living human cells. *Nucleic Acids Res.*, 10.1093/nar/gkv1040.
148. Linder,P. and Jankowsky,E. (2011) From unwinding to clamping — the DEAD box RNA helicase family. *Nat. Rev. Mol. Cell Biol.*, **12**, 505–516.
149. Jarmoskaite,I. and Russell,R. (2011) DEAD-box proteins as RNA helicases and chaperones: DEAD-box proteins. *Wiley Interdiscip. Rev. RNA*, **2**, 135–152.

150. Gilman,B., Tijerina,P. and Russell,R. (2017) Distinct RNA-unwinding mechanisms of DEAD-box and DEAH-box RNA helicase proteins in remodeling structured RNAs and RNPs. *Biochem. Soc. Trans.*, **45**, 1313–1321.
151. Li,Q., Imataka,H., Morino,S., Rogers,G.W., Richter-Cook,N.J., Merrick,W.C. and Sonenberg,N. (1999) Eukaryotic Translation Initiation Factor 4AIII (eIF4AIII) Is Functionally Distinct from eIF4AI and eIF4AII. *Mol. Cell. Biol.*, **19**, 7336–7346.
152. Waldron,J.A., Raza,F. and Le Quesne,J. (2018) eIF4A alleviates the translational repression mediated by classical secondary structures more than by G-quadruplexes. *Nucleic Acids Res.*, **46**, 3075–3087.
153. Parsyan,A., Svitkin,Y., Shahbazian,D., Gkogkas,C., Lasko,P., Merrick,W.C. and Sonenberg,N. (2011) mRNA helicases: the tacticians of translational control. *Nat. Rev. Mol. Cell Biol.*, **12**, 235–245.
154. Sun,Y., Atas,E., Lindqvist,L., Sonenberg,N., Pelletier,J. and Meller,A. (2012) The eukaryotic initiation factor eIF4H facilitates loop-binding, repetitive RNA unwinding by the eIF4A DEAD-box helicase. *Nucleic Acids Res.*, **40**, 6199–6207.
155. Sun,Y., Atas,E., Lindqvist,L.M., Sonenberg,N., Pelletier,J. and Meller,A. (2014) Single-Molecule Kinetics of the Eukaryotic Initiation Factor 4AI upon RNA Unwinding. *Structure*, **22**, 941–948.
156. Chen,Y., Potratz,J.P., Tijerina,P., Del Campo,M., Lambowitz,A.M. and Russell,R. (2008) DEAD-box proteins can completely separate an RNA duplex using a single ATP. *Proc. Natl. Acad. Sci.*, **105**, 20203–20208.
157. Cencic,R., Carrier,M., Galicia-Vázquez,G., Bordeleau,M.-E., Sukarieh,R., Bourdeau,A., Brem,B., Teodoro,J.G., Greger,H., Tremblay,M.L., *et al.* (2009) Antitumor Activity and Mechanism of Action of the Cyclopenta[b]benzofuran, Silvestrol. *PLoS ONE*, **4**, e5223.
158. Biedenkopf,N., Lange-Grünweller,K., Schulte,F.W., Weißer,A., Müller,C., Becker,D., Becker,S., Hartmann,R.K. and Grünweller,A. (2017) The natural compound silvestrol is a potent inhibitor of Ebola virus replication. *Antiviral Res.*, **137**, 76–81.
159. Elgner,F., Sabino,C., Basic,M., Ploen,D., Grünweller,A. and Hildt,E. (2018) Inhibition of Zika Virus Replication by Silvestrol. *Viruses*, **10**, 149.
160. Henss,L., Scholz,T., Grünweller,A. and Schnierle,B. (2018) Silvestrol Inhibits Chikungunya Virus Replication. *Viruses*, **10**, 592.
161. Glitscher,M., Himmelsbach,K., Woytinek,K., Johne,R., Reuter,A., Spiric,J., Schwaben,L., Grünweller,A. and Hildt,E. (2018) Inhibition of Hepatitis E Virus Spread by the Natural Compound Silvestrol. *Viruses*, **10**, 301.
162. Lindqvist,L., Oberer,M., Reibarkh,M., Cencic,R., Bordeleau,M.-E., Vogt,E., Marintchev,A., Tanaka,J., Fagotto,F., Altmann,M., *et al.* (2008) Selective Pharmacological Targeting of a DEAD Box RNA Helicase. *PLoS ONE*, **3**, e1583.
163. Bordeleau,M.-E., Mori,A., Oberer,M., Lindqvist,L., Chard,L.S., Higa,T., Belsham,G.J., Wagner,G., Tanaka,J. and Pelletier,J. (2006) Functional characterization of IRESes by an inhibitor of the RNA helicase eIF4A. *Nat. Chem. Biol.*, **2**, 213–220.

164. 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.
165. Ray,P.S., Jia,J., Yao,P., Majumder,M., Hatzoglou,M. and Fox,P.L. (2009) A stress-responsive RNA switch regulates VEGFA expression. *Nature*, **457**, 915–919.
166. Endoh,T. and Sugimoto,N. (2016) Mechanical insights into ribosomal progression overcoming RNA G-quadruplex from periodical translation suppression in cells. *Sci. Rep.*, **6**, 22719.
167. 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.
168. Mao,Y., Dong,L., Liu,X.-M., Guo,J., Ma,H., Shen,B. and Qian,S.-B. (2019) m6A in mRNA coding regions promotes translation via the RNA helicase-containing YTHDC2. *Nat. Commun.*, **10**, 5332.
169. Shen,S., Faouzi,S., Bastide,A., Martineau,S., Malka-Mahieu,H., Fu,Y., Sun,X., Mateus,C., Routier,E., Roy,S., *et al.* (2019) An epitranscriptomic mechanism underlies selective mRNA translation remodelling in melanoma persister cells. *Nat. Commun.*, **10**, 5713.
170. Höhfeld,J., Benzing,T., Bloch,W., Fürst,D.O., Gehlert,S., Hesse,M., Hoffmann,B., Hoppe,T., Huesgen,P.F., Köhn,M., *et al.* (2021) Maintaining proteostasis under mechanical stress. *EMBO Rep.*, **22**.
171. Giandomenico,S.L., Alvarez-Castelao,B. and Schuman,E.M. (2021) Proteostatic regulation in neuronal compartments. *Trends Neurosci.*, 10.1016/j.tins.2021.08.002.
172. Durand,S. and Lykke-Andersen,J. (2013) Nonsense-mediated mRNA decay occurs during eIF4F-dependent translation in human cells. *Nat. Struct. Mol. Biol.*, **20**, 702–709.
173. Rufener,S.C. and Mühlemann,O. (2013) eIF4E-bound mRNPs are substrates for nonsense-mediated mRNA decay in mammalian cells. *Nat. Struct. Mol. Biol.*, **20**, 710–717.
174. Singh,G., Rebbapragada,I. and Lykke-Andersen,J. (2008) A Competition between Stimulators and Antagonists of Upf Complex Recruitment Governs Human Nonsense-Mediated mRNA Decay. *PLoS Biol.*, **6**, e111.
175. Inada,T. (2017) The Ribosome as a Platform for mRNA and Nascent Polypeptide Quality Control. *Trends Biochem. Sci.*, **42**, 5–15.
176. Simms,C.L. Ribosome Collision Is Critical for Quality Control during No-Go Decay.
177. Juszkiwicz,S., Chandrasekaran,V., Lin,Z., Kraatz,S., Ramakrishnan,V. and Hegde,R.S. (2018) ZNF598 Is a Quality Control Sensor of Collided Ribosomes. *Mol. Cell*, **72**, 469-481.e7.
178. Juszkiwicz,S., Slodkowitz,G., Lin,Z., Freire-Pritchett,P., Peak-Chew,S.-Y. and Hegde,R.S. (2020) Ribosome collisions trigger cis-acting feedback inhibition of translation initiation. *eLife*, **9**, e60038.
179. Joazeiro,C.A.P. (2019) Mechanisms and functions of ribosome-associated protein quality control. *Nat. Rev. Mol. Cell Biol.*, **20**, 368–383.
180. Juszkiwicz,S., Speldewinde,S.H., Wan,L., Svejstrup,J.Q. and Hegde,R.S. (2020) The ASC-1 Complex Disassembles Collided Ribosomes. *Mol. Cell*, **79**, 603-614.e8.

181. Schwab,S.R. (2003) Constitutive Display of Cryptic Translation Products by MHC Class I Molecules. *Science*, **301**, 1367–1371.
182. Starck,S.R. and Shastri,N. (2011) Non-conventional sources of peptides presented by MHC class I. *Cell. Mol. Life Sci.*, **68**, 1471–1479.
183. Yin,Y., Manoury,B. and Fåhraeus,R. (2003) Self-Inhibition of Synthesis and Antigen Presentation by Epstein-Barr Virus–Encoded EBNA. **301**, 5.
184. Martins,R.P., Malbert-Colas,L., Lista,M.J., Daskalogianni,C., Apcher,S., Pla,M., Findakly,S., Blondel,M. and Fåhraeus,R. (2019) Nuclear processing of nascent transcripts determines synthesis of full-length proteins and antigenic peptides. *Nucleic Acids Res.*, **47**, 3086–3100.
185. Dolan,B.P., Li,L., Veltri,C.A., Ireland,C.M., Bennink,J.R. and Yewdell,J.W. (2011) Distinct Pathways Generate Peptides from Defective Ribosomal Products for CD8⁺ T Cell Immunosurveillance. *J. Immunol.*, **186**, 2065–2072.
186. Wei,J., Kishton,R.J., Angel,M., Conn,C.S., Dalla-Venezia,N., Marcel,V., Vincent,A., Catez,F., Ferré,S., Ayadi,L., *et al.* (2019) Ribosomal Proteins Regulate MHC Class I Peptide Generation for Immunosurveillance. *Mol. Cell*, **73**, 1162-1173.e5.
187. Cardinaud,S., Starck,S.R., Chandra,P. and Shastri,N. (2010) The Synthesis of Truncated Polypeptides for Immune Surveillance and Viral Evasion. *PLoS ONE*, **5**, 12.
188. Anton,L.C. and Yewdell,J.W. (2014) Translating DRiPs: MHC class I immunosurveillance of pathogens and tumors. *J. Leukoc. Biol.*, **95**, 551–562.
189. Trentini,D.B., Pecoraro,M., Tiwary,S., Cox,J., Mann,M., Hipp,M.S. and Hartl,F.U. (2020) Role for ribosome-associated quality control in sampling proteins for MHC class I-mediated antigen presentation. *Proc. Natl. Acad. Sci.*, **117**, 4099–4108.
190. Apcher,S., Millot,G., Daskalogianni,C., Scherl,A., Manoury,B. and Fahraeus,R. (2013) Translation of pre-spliced RNAs in the nuclear compartment generates peptides for the MHC class I pathway. *Proc. Natl. Acad. Sci.*, **110**, 17951–17956.
191. Brandman,O., Stewart-Ornstein,J., Wong,D., Larson,A., Williams,C.C., Li,G.-W., Zhou,S., King,D., Shen,P.S., Weibezahn,J., *et al.* (2012) A Ribosome-Bound Quality Control Complex Triggers Degradation of Nascent Peptides and Signals Translation Stress. *Cell*, **151**, 1042–1054.
192. Wu,C.C.-C., Peterson,A., Zinshteyn,B., Regot,S. and Green,R. (2020) Ribosome Collisions Trigger General Stress Responses to Regulate Cell Fate. *Cell*, **182**, 404-416.e14.
193. 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.
194. 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.
195. Mogk,A., Bukau,B. and Kampinga,H.H. (2018) Cellular Handling of Protein Aggregates by Disaggregation Machines. *Mol. Cell*, **69**, 214–226.

196. Źwirowski,S., Kłosowska,A., Obuchowski,I., Nillegoda,N.B., Piróg,A., Ziętkiewicz,S., Bukau,B., Mogk,A. and Liberek,K. (2017) Hsp70 displaces small heat shock proteins from aggregates to initiate protein refolding. *EMBO J.*, **36**, 783–796.
197. Gonçalves,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.
198. Wang,L., Zhang,W., Wang,L., Zhang,X.C., Li,X. and Rao,Z. (2010) Crystal structures of NAC domains of human nascent polypeptide-associated complex (NAC) and its α NAC subunit. *Protein Cell*, **1**, 406–416.
199. Beatrix,B., Sakai,H. and Wiedmann,M. (2000) The α and β Subunit of the Nascent Polypeptide-associated Complex Have Distinct Functions. *J. Biol. Chem.*, **275**, 37838–37845.
200. 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.
201. 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.
202. Zheng,X.-M., Moncollin,V., Egly,J.-M. and Chambon,P. (1987) A general transcription factor forms a stable complex with RNA polymerase B (II). *Cell*, **50**, 361–368.
203. Zheng,X.M., Black,D., Chambon,P. and Egly,J.M. (1990) Sequencing and expression of complementary DNA for the general transcription factor BTF3. *Nature*, **344**, 556–559.
204. Yotov,W.V., Moreau,A. and St-Arnaud,R. (1998) The Alpha Chain of the Nascent Polypeptide-Associated Complex Functions as a Transcriptional Coactivator. *Mol. Cell. Biol.*, **18**, 1303–1311.
205. Akhouayri,O., Quélo,I. and St-Arnaud,R. (2005) Sequence-Specific DNA Binding by the α NAC Coactivator Is Required for Potentiation of c-Jun-Dependent Transcription of the Osteocalcin Gene. *Mol. Cell. Biol.*, **25**, 3452–3460.
206. Wegrzyn,R.D., Hofmann,D., Merz,F., Nikolay,R., Rauch,T., Graf,C. and Deuerling,E. (2006) A Conserved Motif Is Prerequisite for the Interaction of NAC with Ribosomal Protein L23 and Nascent Chains. *J. Biol. Chem.*, **281**, 2847–2857.
207. Raden,D. and Gilmore,R. (1998) Signal Recognition Particle-dependent Targeting of Ribosomes to the Rough Endoplasmic Reticulum in the Absence and Presence of the Nascent Polypeptide-associated Complex. *Mol. Biol. Cell*, **9**, 117–130.
208. Luring,B., Sakai,H., Kreibich,G. and Wiedmann,M. (1995) Nascent polypeptide-associated complex protein prevents mistargeting of nascent chains to the endoplasmic reticulum. *Proc. Natl. Acad. Sci.*, **92**, 5411–5415.
209. Luring,B., Kreibich,G. and Weidmann,M. (1995) The intrinsic ability of ribosomes to bind to endoplasmic reticulum membranes is regulated by signal recognition particle and nascent-polypeptide-associated complex. *Proc. Natl. Acad. Sci.*, **92**, 9435–9439.
210. Alamo,M. del, 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.

211. Gamerdinger, M., Kobayashi, K., Wallisch, A., Kreft, S.G., Sailer, C., Schlömer, 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.e8.
212. 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.
213. Raue, U., Oellerer, S. and Rospert, S. (2007) Association of Protein Biogenesis Factors at the Yeast Ribosomal Tunnel Exit Is Affected by the Translational Status and Nascent Polypeptide Sequence. *J. Biol. Chem.*, **282**, 7809–7816.
214. Nyathi, Y. and Pool, M.R. (2015) Analysis of the interplay of protein biogenesis factors at the ribosome exit site reveals new role for NAC. *J. Cell Biol.*, **210**, 287–301.
215. Koplin, A., Preissler, S., Ilina, Y., Koch, M., Scior, A., Erhardt, M. and Deuerling, E. (2010) A dual function for chaperones SSB–RAC and the NAC nascent polypeptide–associated complex on ribosomes. *J. Cell Biol.*, **189**, 57–68.
216. Pech, M., Spreter, T., Beckmann, R. and Beatrix, B. (2010) Dual Binding Mode of the Nascent Polypeptide-associated Complex Reveals a Novel Universal Adapter Site on the Ribosome. *J. Biol. Chem.*, **285**, 19679–19687.
217. 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.
218. Rospert, S., Dubaquié, Y. and Gautschi, M. (2002) Nascent-polypeptide-associated complex. *Cell. Mol. Life Sci.*, **59**, 1632–1639.
219. Shen, K., Gamerdinger, M., Chan, R., Gense, K., Martin, E.M., Sachs, N., Knight, P.D., Schlömer, 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.e7.
220. Young, L.S., Yap, L.F. and Murray, P.G. (2016) Epstein–Barr virus: more than 50 years old and still providing surprises. *Nat. Rev. Cancer*, **16**, 789–802.
221. Farrell, P.J. (2018) Epstein–Barr Virus and Cancer.
222. Varvatsi, D., Richter, J., Tryfonos, C., Pantzaris, M. and Christodoulou, C. (2021) Association of Epstein-Barr virus latently expressed genes with multiple sclerosis. *Mult. Scler. Relat. Disord.*, **52**, 103008.
223. Rühl, J., Citterio, C., Engelmann, C., Haigh, T., Dzionek, A., Dreyer, J., Khanna, R., Taylor, G.S., Wilson, J.B., Leung, C.S., *et al.* (2019) Heterologous prime-boost vaccination protects against EBV antigen–expressing lymphomas. *J. Clin. Invest.*, **129**, 2071–2087.
224. Rowe, M., Lear, A.L., Croom-Carter, D., Davies, A.H. and Rickinson, A.B. (1992) Three pathways of Epstein-Barr virus gene activation from EBNA1-positive latency in B lymphocytes. *J. Virol.*, **66**, 122–131.

225. The UniProt Consortium, Bateman,A., Martin,M.-J., Orchard,S., Magrane,M., Agivetova,R., Ahmad,S., Alpi,E., Bowler-Barnett,E.H., Britto,R., *et al.* (2021) UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.*, **49**, D480–D489.
226. Lu,C.-C., Wu,C.-W., Chang,S.C., Chen,T.-Y., Hu,C.-R., Yeh,M.-Y., Chen,J.-Y. and Chen,M.-R. (2004) Epstein–Barr virus nuclear antigen 1 is a DNA-binding protein with strong RNA-binding activity. *J. Gen. Virol.*, **85**, 2755–2765.
227. Norseen,J., Thomae,A., Sridharan,V., Aiyar,A., Schepers,A. and Lieberman,P.M. (2008) RNA-dependent recruitment of the origin recognition complex. *EMBO J.*, **27**, 3024–3035.
228. 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.
229. 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.
230. Kapoor,P., Lavoie,B.D. and Frappier,L. (2005) EBP2 Plays a Key Role in Epstein-Barr Virus Mitotic Segregation and Is Regulated by Aurora Family Kinases. *Mol. Cell. Biol.*, **25**, 4934–4945.
231. Chen,Y.-L., Liu,C.-D., Cheng,C.-P., Zhao,B., Hsu,H.-J., Shen,C.-L., Chiu,S.-J., Kieff,E. and Peng,C. -w. (2014) Nucleolin is important for Epstein-Barr virus nuclear antigen 1-mediated episome binding, maintenance, and transcription. *Proc. Natl. Acad. Sci.*, **111**, 243–248.
232. Gahn,T.A. and Sugden,B. (1995) 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**, 2633–2636.
233. Boreström,C., Forsman,A., Rüetschi,U. and Rymo,L. (2012) E2F1, ARID3A/Bright and Oct-2 factors bind to the Epstein–Barr virus C promoter, EBNA1 and oriP, participating in long-distance promoter–enhancer interactions. *J. Gen. Virol.*, **93**, 1065–1075.
234. Schulz,T.F. and Cordes,S. (2009) Is the Epstein-Barr virus EBNA-1 protein an oncogen? *Proc. Natl. Acad. Sci.*, **106**, 2091–2092.
235. 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.
236. Münz,C., Bickham,K.L., Subklewe,M., Tsang,M.L., Chahroudi,A., Kurilla,M.G., Zhang,D., O’Donnell,M. and Steinman,R.M. Human CD4 ϵ T Lymphocytes Consistently Respond to the Latent Epstein-Barr Virus Nuclear Antigen EBNA.
237. Long,H.M., Haigh,T.A., Gudgeon,N.H., Leen,A.M., Tsang,C.-W., Brooks,J., Landais,E., Houssaint,E., Lee,S.P., Rickinson,A.B., *et al.* (2005) CD4⁺ T-Cell Responses to Epstein-Barr Virus (EBV) Latent-Cycle Antigens and the Recognition of EBV-Transformed Lymphoblastoid Cell Lines. *J. Virol.*, **79**, 4896–4907.
238. Paludan,C., Bickham,K., Nikiforow,S., Tsang,M.L., Goodman,K., Hanekom,W.A., Fonteneau,J.-F., Stevanović,S. and Münz,C. (2002) Epstein-Barr Nuclear Antigen 1-Specific CD4⁺ Th1 Cells Kill Burkitt’s Lymphoma Cells. *J. Immunol.*, **169**, 1593–1603.

239. Blake,N., Lee,S., Redchenko,I., Thomas,W., Steven,N., Leese,A., Steigerwald-Mullen,P., Kurilla,M.G., Frappier,L. and Rickinson,A. Human CD8⁺ T Cell Responses to EBV EBNA1: HLA Class I Presentation of the (Gly-Ala)_n-Containing Protein Requires Exogenous Processing.
240. Nikiforow,S., Bottomly,K., Miller,G. and Münz,C. (2003) Cytolytic CD4⁺ T-Cell Clones Reactive to EBNA1 Inhibit Epstein-Barr Virus-Induced B-Cell Proliferation. *J. Virol.*, **77**, 12088–12104.
241. Icheva,V., Kayser,S., Wolff,D., Tuve,S., Kyzirakos,C., Bethge,W., Greil,J., Albert,M.H., Schwinger,W., Nathrath,M., *et al.* (2013) Adoptive Transfer of Epstein-Barr Virus (EBV) Nuclear Antigen 1–Specific T Cells As Treatment for EBV Reactivation and Lymphoproliferative Disorders After Allogeneic Stem-Cell Transplantation. *J. Clin. Oncol.*, **31**, 39–48.
242. Mautner,J., Pich,D., Nimmerjahn,F., Milosevic,S., Adhikary,D., Christoph,H., Witter,K., Bornkamm,G.W., Hammerschmidt,W. and Behrends,U. (2004) Epstein-Barr virus nuclear antigen 1 evades direct immune recognition by CD4⁺ T helper cells. *Eur. J. Immunol.*, **34**, 2500–2509.
243. Levitskaya,J., Coram,M., Levitsky,V., Imreh,S., Steigerwald-Mullen,P.M., Klein,G., Kurilla,M.G. and Masucci,M.G. (1995) Inhibition of antigen processing by the internal repeat region of the Epstein-Barr virus nuclear antigen-1. *Nature*, **375**, 685–688.
244. Levitskaya,J., Sharipo,A., Leonchiks,A., Ciechanover,A. and Masucci,M.G. (1997) 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.*, **94**, 12616–12621.
245. Daskalogianni,C., Apcher,S., Candeias,M.M., Naski,N., Calvo,F. and Fåhræus,R. (2008) Gly-Ala Repeats Induce Position- and Substrate-specific Regulation of 26 S Proteasome-dependent Partial Processing. *J. Biol. Chem.*, **283**, 30090–30100.
246. 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.*, **108**, 11572–11577.
247. Apcher,S., Komarova,A., Daskalogianni,C., Yin,Y., Malbert-Colas,L. and Fåhræus,R. (2009) mRNA Translation Regulation by the Gly-Ala Repeat of Epstein-Barr Virus Nuclear Antigen 1. *J. Virol.*, **83**, 1289–1298.
248. Apcher,S., Daskalogianni,C., Manoury,B. and Fåhræus,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.
249. Tellam,J.T., Zhong,J., Lekieffre,L., Bhat,P., Martinez,M., Croft,N.P., Kaplan,W., Tellam,R.L. and Khanna,R. (2014) mRNA Structural Constraints on EBNA1 Synthesis Impact on In Vivo Antigen Presentation and Early Priming of CD8⁺ T Cells. *PLoS Pathog.*, **10**, e1004423.
250. Prado Martins,R., Findakly,S., Daskalogianni,C., Teulade-Fichou,M.-P., Blondel,M. and Fåhræus,R. (2018) In Cellulo Protein-mRNA Interaction Assay to Determine the Action of G-Quadruplex-Binding Molecules. *Molecules*, **23**, 3124.
251. Gruhne,B., Sompallae,R., Marescotti,D., Kamranvar,S.A., Gastaldello,S. and Masucci,M.G. (2009) The Epstein-Barr virus nuclear antigen-1 promotes genomic instability via induction of reactive oxygen species. *Proc. Natl. Acad. Sci.*, **106**, 2313–2318.

252. Kamranvar,S.A., Gruhne,B., Szeles,A. and Masucci,M.G. (2007) Epstein–Barr virus promotes genomic instability in Burkitt’s lymphoma. *Oncogene*, **26**, 5115–5123.
253. Kuhn-Hallek,I., Sage,D.R., Stein,L., Groelle,H. and Fingerroth,J.D. Expression of Recombination Activating Genes (RAG-1 and RAG-2) in Epstein-Barr Virus-Bearing B Cells.
254. Srinivas,S.K. and Sixbey,J.W. (1995) Epstein-Barr virus induction of recombinase-activating genes RAG1 and RAG2. *J. Virol.*, **69**, 8155–8158.
255. Kennedy,G., Komano,J. and Sugden,B. (2003) Epstein-Barr virus provides a survival factor to Burkitt’s lymphomas. *Proc. Natl. Acad. Sci.*, **100**, 14269–14274.
256. Saridakis,V., Sheng,Y., Sarkari,F., Holowaty,M.N., Shire,K., Nguyen,T., Zhang,R.G., Liao,J., Lee,W., Edwards,A.M., *et al.* (2005) Structure of the p53 Binding Domain of HAUSP/USP7 Bound to Epstein-Barr Nuclear Antigen 1. *Mol. Cell*, **18**, 25–36.
257. Gnanasundram,S.V., Pyndiah,S., Daskalogianni,C., Armfield,K., Nylander,K., Wilson,J.B. and Fåhræus,R. (2017) PI3K δ activates E2F1 synthesis in response to mRNA translation stress. *Nat. Commun.*, **8**, 2103.
258. Gnanasundram,S.V., Malbert-Colas,L., Chen,S., Fusée,L., Daskalogianni,C., Muller,P., Salomao,N. and Fåhræus,R. (2020) MDM2’s dual mRNA binding domains co-ordinate its oncogenic and tumour suppressor activities. *Nucleic Acids Res.*, **48**, 6775–6787.
259. Sokabe,M. and Fraser,C.S. A helicase-independent activity of eIF4A in promoting mRNA recruitment to the human ribosome. *Comput. Biol.*
260. Payán-Bravo,L., Fontalva,S., Peñate,X., Cases,I., Guerrero-Martínez,J.A., Pareja-Sánchez,Y., Odriozola-Gil,Y., Lara,E., Jimeno-González,S., Suñé,C., *et al.* (2021) Human prefoldin modulates co-transcriptional pre-mRNA splicing. *Nucleic Acids Res.*, **49**, 6267–6280.
261. Albanese,M., Tagawa,T., Bouvet,M., Maliqi,L., Lutter,D., Hoser,J., Hastreiter,M., Hayes,M., Sugden,B., Martin,L., *et al.* (2016) Epstein–Barr virus microRNAs reduce immune surveillance by virus-specific CD8⁺ T cells. *Proc. Natl. Acad. Sci.*, **113**, E6467–E6475.
262. Aarum,J., Cabrera,C.P., Jones,T.A., Rajendran,S., Adiutori,R., Giovannoni,G., Barnes,M.R., Malaspina,A. and Sheer,D. (2020) Enzymatic degradation of RNA causes widespread protein aggregation in cell and tissue lysates. *EMBO Rep.*, **21**.
263. Begeman,A., Son,A., Litberg,T.J., Wroblewski,T.H., Gehring,T., Huizar Cabral,V., Bourne,J., Xuan,Z. and Horowitz,S. (2020) G-Quadruplexes act as sequence-dependent protein chaperones. *EMBO Rep.*, **21**.
264. Wang,H., Bu,L., Wang,C., Zhang,Y., Zhou,H., Zhang,X., Guo,W., Long,C., Guo,D. and Sun,X. (2018) The Hsp70 inhibitor 2-phenylethanesulfonamide inhibits replication and carcinogenicity of Epstein–Barr virus by inhibiting the molecular chaperone function of Hsp70. *Cell Death Dis.*, **9**, 734.
265. Davenport,M.G. and Pagano,J.S. (1999) Expression of EBNA-1 mRNA Is Regulated by Cell Cycle during Epstein-Barr Virus Type I Latency. *J. Virol.*, **73**, 3154–3161.
266. Shen,C.-L., Liu,C.-D., You,R.-I., Ching,Y.-H., Liang,J., Ke,L., Chen,Y.-L., Chen,H.-C., Hsu,H.-J., Liou,J.-W., *et al.* (2016) Ribosome Protein L4 is essential for Epstein–Barr Virus Nuclear Antigen 1 function. *Proc. Natl. Acad. Sci.*, **113**, 2229–2234.

267. Zarin,T., Strome,B., Nguyen Ba,A.N., Alberti,S., Forman-Kay,J.D. and Moses,A.M. (2019) Proteome-wide signatures of function in highly diverged intrinsically disordered regions. *eLife*, **8**, e46883.
268. Fassler,J.S., Skuodas,S., Weeks,D.L. and Phillips,B.T. (2021) Protein Aggregation and Disaggregation in Cells and Development. *J. Mol. Biol.*, **433**, 167215.
269. Castello,A., Fischer,B., Frese,C.K., Horos,R., Alleaume,A.-M., Foehr,S., Curk,T., Krijgsveld,J. and Hentze,M.W. (2016) Comprehensive Identification of RNA-Binding Domains in Human Cells. *Mol. Cell*, **63**, 696–710.
270. Allnutt,M.A., Johnson,K., Bennett,D.A., Connor,S.M., Troncoso,J.C., Pletnikova,O., Albert,M.S., Resnick,S.M., Scholz,S.W., De Jager,P.L., *et al.* (2020) Human Herpesvirus 6 Detection in Alzheimer’s Disease Cases and Controls across Multiple Cohorts. *Neuron*, **105**, 1027-1035.e2.
271. Danics,K., Forrest,S.L., Kapas,I., Erber,I., Schmid,S., Törő,K., Majtenyi,K. and Kovacs,G.G. (2021) Neurodegenerative proteinopathies associated with neuroinfections. *J. Neural Transm.*, 10.1007/s00702-021-02371-7.
272. Woulfe,J.M., Gray,M.T., Gray,D.A., Munoz,D.G. and Middeldorp,J.M. (2014) Hypothesis: A role for EBV-induced molecular mimicry in Parkinson’s disease. *Parkinsonism Relat. Disord.*, **20**, 685–694.
273. Jakimovski,D., Weinstock-Guttman,B., Ramanathan,M., Dwyer,M.G. and Zivadinov,R. (2020) Infections, Vaccines and Autoimmunity: A Multiple Sclerosis Perspective. *Vaccines*, **8**, 50.
274. Cairns,D.M., Rouleau,N., Parker,R.N., Walsh,K.G., Gehrke,L. and Kaplan,D.L. (2020) A 3D human brain-like tissue model of herpes-induced Alzheimer’s disease. *Sci. Adv.*, **6**, eaay8828.
275. Perrone,R., Nadai,M., Frasson,I., Poe,J.A., Butovskaya,E., Smithgall,T.E., Palumbo,M., Palù,G. and Richter,S.N. (2013) A Dynamic G-Quadruplex Region Regulates the HIV-1 Long Terminal Repeat Promoter. *J. Med. Chem.*, **56**, 6521–6530.
276. Wang,C., Wang,H., Zhang,Y., Guo,W., Long,C., Wang,J., Liu,L. and Sun,X. (2017) Berberine inhibits the proliferation of human nasopharyngeal carcinoma cells via an Epstein-Barr virus nuclear antigen 1-dependent mechanism. *Oncol. Rep.*, **37**, 2109–2120.
277. Yang,Q., Del Campo,M., Lambowitz,A.M. and Jankowsky,E. (2007) DEAD-Box Proteins Unwind Duplexes by Local Strand Separation. *Mol. Cell*, **28**, 253–263.
278. Fairman,M.E. (2004) Protein Displacement by DExH/D ‘RNA Helicases’ Without Duplex Unwinding. *Science*, **304**, 730–734.
279. Jankowsky,E. and Bowers,H. (2006) Remodeling of ribonucleoprotein complexes with DExH/D RNA helicases. *Nucleic Acids Res.*, **34**, 4181–4188.
280. Tran,E.J., Zhou,Y., Corbett,A.H. and Wentz,S.R. (2007) The DEAD-Box Protein Dbp5 Controls mRNA Export by Triggering Specific RNA:Protein Remodeling Events. *Mol. Cell*, **28**, 850–859.
281. Hondele,M., Sachdev,R., Heinrich,S., Wang,J., Vallotton,P., Fontoura,B.M.A. and Weis,K. (2019) DEAD-box ATPases are global regulators of phase-separated organelles. *Nature*, **573**, 144–148.
282. Tauber,D., Tauber,G., Khong,A., Van Treeck,B., Pelletier,J. and Parker,R. (2020) Modulation of RNA Condensation by the DEAD-Box Protein eIF4A. *Cell*, **180**, 411-426.e16.

283. Tsujii,R., Miyoshi,K., Tsuno,A., Matsui,Y., Toh-e,A., Miyakawa,T. and Mizuta,K. (2000) Ebp2p, yeast homologue of a human protein that interacts with Epstein-Barr virus Nuclear Antigen 1, is required for pre-rRNA processing and ribosomal subunit assembly: Role of yeast EBP2 in ribosome biogenesis. *Genes Cells*, **5**, 543–553.
284. Huber,M.D., Dworet,J.H., Shire,K., Frappier,L. and McAlear,M.A. (2000) The Budding Yeast Homolog of the Human EBNA1-binding Protein 2 (Ebp2p) Is an Essential Nucleolar Protein Required for Pre-rRNA Processing. *J. Biol. Chem.*, **275**, 28764–28773.
285. Miller,C.M., Selvam,S. and Fuchs,G. (2021) Fatal attraction: The roles of ribosomal proteins in the viral life cycle. *WIREs RNA*, **12**.
286. Balcerak,A., Trebinska-Stryjewska,A., Konopinski,R., Wakula,M. and Grzybowska,E.A. (2019) RNA–protein interactions: disorder, moonlighting and junk contribute to eukaryotic complexity. *Open Biol.*, **9**, 190096.
287. Jeffery,C.J. (2018) Protein moonlighting: what is it, and why is it important? *Philos. Trans. R. Soc. B Biol. Sci.*, **373**, 20160523.
288. Hwang,H.J., Park,Y. and Kim,Y.K. (2021) UPF1: From mRNA Surveillance to Protein Quality Control. *Biomedicines*, **9**, 995.
289. Feng,Q., Jagannathan,S. and Bradley,R.K. (2017) The RNA Surveillance Factor UPF1 Represses Myogenesis via Its E3 Ubiquitin Ligase Activity. *Mol. Cell*, **67**, 239-251.e6.
290. Park,Y., Park,J., Hwang,H.J., Kim,B., Jeong,K., Chang,J., Lee,J.-B. and Kim,Y.K. (2020) Nonsense-mediated mRNA decay factor UPF1 promotes aggresome formation. *Nat. Commun.*, **11**, 3106.

Acknowledgements

These last four years have not been a smooth journey, and I would like to acknowledge the people who have shared a bit of it and helped me to get through it.

I would like to thank the members of the jury for accepting to be part of it, and for the great discussion we had during the defense. I am grateful to Dr. Stéphan Vagner for being the president of my jury and for his advice during some PhD committees. I also would like to thank Dr. Stefania Millevoi and Dr. Marie-Françoise O'Donohue for being the rapporteures and for their useful comments regarding the manuscript and the project in general. I warmly thank Pr. Marc Blondel and Dr. Isabelle Dusanter-Fourt for their advice regarding this project over these last four years.

I am grateful to Robin, for giving me the opportunity to join his lab and for coping with my up and downs during these years. I also thank him for his attention to my project, despite his busy schedule, and for letting me gain independence and explore the ideas I thought interesting. I also thank him for all the philosophical discussions we had and promise I will try to expand my cinematographic culture one day.

I would like to take the time to thank all the people who worked on EBNA1 and its GAR domain before I stepped in. I am aware that I arrived at just the right time on the EBNA1 subject. After four years, I know how tricky the GAR and EBNA1 are, and I wish good luck to the people who will pursue the studies on EBNA1 and its GAR domain.

I want to thank all the members of the lab. As you know I am not super comfortable with this kind of stuff and would like to get rid of it as easily as just saying bye. But I really enjoyed spending time with you all, so I will make the effort ;)

I would like to thank Chrys and Laurence, for welcoming me to the lab, giving me advice, for the nice vibes they bring to the lab, and for the raclettes of course. Also, without them, I would barely have half of the results presented in this manuscript, so really a big thank you to both of them.

I would like to thank Ewa, Katerina, Maria, and Norman. If you ever read this line and wonder how I choose the order, I just took your arrival order in the lab ^^ . Thank you, Ewa, for these evenings where we were wondering what we were doing and debating about all kinds of stuff, and for the pizzas we shared! Thank you, Maria, for reassuring me when I thought I would never make it, and for sharing a culture I had no clues about. Merci Norman pour ton positivisme, malgré

les aléas de la vie de labo (et ils ont été bien nombreux ^^). And a big thank you to Katerina, who helped me a lot with my experiments, but also for being patient with me, and for sharing her story and bringing positive vibes all along this project. I really enjoyed spending time with you all and wish you good luck with your Ph.D. projects. If I could make it, you can for sure ;)

Merci aussi à Irène, pour avoir été ma voisine de bureau pendant quasi la moitié de ma thèse, pour avoir partagé plein de bons plans sur la vie parisienne, et pour s'occuper du labo en général. Merci aussi à Marie-Christine, pour m'avoir aidée à m'y retrouver (autant que possible) dans le labyrinthe de l'administration française.

Thank you also to the people that were in the lab and left before the end of my thesis. Merci Avesta pour m'avoir montrée comment le labo fonctionnait. Thank you Justine for sharing your crazy stories. Thank you Siva and Rodrigo for giving me great advice over these years and being available. Thank you Konstantinos for helping me understand the PLEA, and for welcoming us in Greece. Thank you Agata, although we did not spend a lot of time together, it was nice to meet you and spend some time with you. Et merci Mathilde pour tous les fous rires autour d'un petit verre de gin et pour toutes les discussions (plus ou moins philosophiques certes) qu'on a pu avoir. (A cause de toi, mon objectif est de finir en 208 pages.) I hope that everything is going and will go well for you all.

J'aimerais aussi remercier Nathan et Diane, mes premiers stagiaires, avec qui j'ai beaucoup appris. Je pense que j'ai eu beaucoup de chance d'avoir pu travailler avec vous en tant que "maître de stage" débutant.

J'aimerais aussi remercier le personnel de la plateforme technologique de l'IRSL, et en particulier Niclas et Christelle, pour leur disponibilité et leur écoute attentive, et pour toute l'aide qu'ils m'ont apportée en microscopie.

Je remercie tous nos collaborateurs pour m'avoir aidé dans ce projet. Tout d'abord les Bretons, avec Alicia, Trang et Ronan. I would also like to thank Dr. Teulade-Fichou, Dr. Granzhan and their team in Paris for supplying us in reagents, and Sa and Mikael for the proteins from Umea.

Je remercie la FRM pour m'avoir financée lors de cette dernière année, qui a été bien nécessaire.

Une petite pensée pour mon ancienne équipe : Bruno, Nabil, Ahmed, Alison, Léa, Louise, Sofia, Amélie, Amélie (oui x2), Laure, Mickaël, Océane, Marie, Isabelle, Alain et Florence. C'est avec vous que j'ai vraiment fait mes débuts en sciences, et vous pouvez constater le résultat ;)

Un très très grand merci à mes amis “hors labo”. Le groupe de l’agro, avec Lina, Hortense, Manon de la team paname, mais aussi Saki (encore merci d’avoir accepté de relire quelques pages), Gaïa, Zoé, Clara, St Carolus, Zélie, Florie et Laure. J’ai pu découvrir différentes parties de la France grâce à vous et aussi voyager à travers l’Europe. Bien évidemment, je n’oublierai pas nos multiples expériences culinaires, très très souvent réussies 😊. Un grand merci pour votre écoute et pour m’avoir fait sortir du labo. J’espère qu’on arrivera à continuer à se voir de temps en temps toutes ensemble malgré notre éparpillement géographique.

Le petit groupe de prépa, avec Elisabeth, Sarah, Marie, et Mélanie. J’espère qu’on arrivera à se revoir toutes ensemble dans pas trop longtemps.

Merci à ceux qui me connaissent depuis plus de 10 ans (et oui, ça commence à faire ^^) : Laetitia, Audrey P, Audrey B, Léa, Angèle, Kémilia, Radoen, Lucas, Alice, Camille mais aussi Héloïse et Lucie, pour tous ces moments partagés depuis le lycée. Merci aussi au petit groupe de terminale : Ophélie, Ludovic et Manon, déjà pour avoir répondu à mes messages malgré toutes ces années, et pour ces retrouvailles.

I would like to thank Nick, Roos, Rianne, Mieke, and Zwier for welcoming me when I was in the Netherlands. Soon enough I hope I will be able to say it in dutch to you!

Un très très grand merci à ma famille. Tout d’abord à mes parents, qui quoiqu’il arrive m’ont toujours soutenue. Merci à Anne et Xavier aussi pour avoir pris des nouvelles de temps en temps malgré nos vies bien occupées maintenant. J’espère qu’on arrivera à se faire des week-ends un peu plus souvent. Merci à Wo Yonggang et nounou, pour s’être occupés de nous toutes ces années (et ça ne devait pas être super facile). I would like to thank also my family in China, although we do not know each other very well, they always welcome me whenever I went there.

Last, but not least, a big thank you to Martijn, for supporting me and listening to me all these years, and giving me the strength to pursue this Ph.D. Without you, I would not even have applied to this project and would have probably given up so many times. Even though I am not sure you would be happy to hear that, this Ph.D. project is also a bit yours. I was incredibly lucky to meet you, and I am happy that we shared our little journey through life since then. Now let’s see what kind of crazy adventures life will bring us.