Étude de l'export nucléaire des ARNm du virus de la mosaïque du chou-fleur (CaMV) chez Arabidopsis thaliana

Julie Kubina

▶ **To cite this version:**

Julie Kubina. Étude de l'export nucléaire des ARNm du virus de la mosaïque du chou-fleur (CaMV) chez Arabidopsis thaliana. Biologie végétale. Université de Strasbourg, 2019. Français. NNT : 2019STRAJ052 . tel-03270818

HAL Id: tel-03270818
https://theses.hal.science/tel-03270818

Submitted on 25 Jun 2021

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 STRASBOURG

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ

Institut de Biologie Moléculaire des Plantes du CNRS – UPR 2357

THÈSE

présentée par :

Julie KUBINA

soutenue le : **20 septembre 2019**

pour obtenir le grade de : **Docteur de l'Université de Strasbourg**

Mention : Sciences de la Vie et de la Santé

Discipline : Aspects Moléculaire et Cellulaire de la Biologie

Étude de l'export nucléaire des ARNm du virus de la mosaïque du chou-fleur (CaMV) chez *Arabidopsis thaliana*

THÈSE dirigée par :

Mme DIMITROVA Maria	Professeur, Université de Strasbourg IBMP CNRS

RAPPORTEURS :

M. JACQUOT Emmanuel	Directeur de recherche, CIRAD INRA (Montpellier)
M. OHLMANN Théophile	Directeur de recherche, CIRI INSERM (Lyon)

EXAMINATEURS :

M. GILMER David	Professeur, Université de Strasbourg IBMP CNRS
Mme UZEST Marilyne	Chargée de recherche, CIRAD INRA (Montpellier)

Remerciements

Je tiens tout d'abord à remercier Maria Dimitrova, ma directrice de thèse, pour avoir accepté de diriger mes travaux de thèse. Maria, Merci, de m'avoir permis de continuer le travail engagé par Clément sur l'export des ARN de CaMV, pour m'avoir fait confiance, portée et guidée durant toutes ces années. Merci d'avoir fait preuve de patience et de compréhension face à ma sensibilité, merci pour ton écoute, ton soutien en toutes situations et surtout tes encouragements ! Beaucoup d'ingrédients qui m'ont permis de m'épanouir dans ce laboratoire au contexte difficile et qui ont été essentiels dans mon éveil scientifique.

J'aimerais également remercier Angèle Geldreich, pour m'avoir prise sous son aile au laboratoire et accepté de partager, avec patience, ses connaissances et son savoir-faire. Angèle, Merci, pour ton immense gentillesse et ton soutien sans faille, tes conseils avisés, ton enthousiasme inébranlable et nos conversations très enrichissantes, sur le CaMV mais pas que.

Un grand merci à Mario Keller, mon ancien maître de stage et chef (d'équipe), de m'avoir donné ma chance et introduite dans le laboratoire, d'avoir cru en moi et portée jusqu'à la MRT. Mario, Merci, d'avoir allumé en moi cette flamme pour ce *Caulimovirus* si particulier. Merci de m'avoir guidée et corrigée durant la rédaction de ce manuscrit de thèse, pour ta grande patience, ton écoute et tes encouragements au cours de ces trois derniers mois.

Je remercie aussi les membres de mon jury, Madame Uzest Marilyne et Messieurs Jacquot Emmanuel, Ohlmann Théophile et Gilmer David, pour avoir accepté de juger mon travail.

Merci à mes anciens profs, particulièrement Salah, Corinne et David. Merci d'avoir partagé avec moi vos connaissances et votre passion pour la viro.

Une pensée pour mes petits pioupious, Vincent, Laetitia, Myriam, Audrey, Charlotte, Manu et Jon, j'espère ne pas vous avoir trop traumatisés !

Merci aux copains « des sciences » de master et du labo, spécialement Isabelle, Élodie, Antoine, Natalia, Amelia, Xue, Valentin, Marion, Thibaut, Lucie et Yihan ; ainsi qu'aux collègues de la cantine, Rozenn, Thierry, Franck et Jean-Étienne, pour nos discussions enjouées de la pause déjeuner et nos franches rigolades.

Merci aux copines de la salle de rédac' pour avoir sué avec moi ces dernières semaines, Magdalène, Caroline, Chloé et Lucie.

Merci aussi à tous les autres membres de l'IBMP et en particulier les membres de mon équipe, pour votre aide, vos conseils et votre sympathie.

Enfin, un immense merci à mes parents, à ma sœur et mon beau-frère, ma marraine et mon amoureux, Maryse, Martial, Sophie, Théo, Maryline et Julien, et mes amies, les « filles ». Merci pour m'avoir toujours soutenue et permis d'aller aussi loin.

Pour finir, une pensée pour mon papi, à qui je dédie ce manuscrit.

Avant-propos

Les travaux qui vous seront présentés dans ce manuscrit, sont consacrés à l'étude de l'export nucléaire des ARN messagers et génomique du virus de la mosaïque du chou-fleur (CaMV). Il s'agit d'une nouvelle thématique de recherche développée au laboratoire afin de comprendre cette étape nucléaire cruciale pour le cycle de réplication viral. Cette étude a également pour objet de contribuer à l'élucidation du mécanisme de l'export nucléaire des ARNm cellulaires des plantes.

Le manuscrit de thèse commence par une introduction générale présentant de manière globale, les mécanismes régissant l'export nucléaire des protéines et des ARN chez la levure, les métazoaires et les virus des animaux. Un accent tout particulier est mis sur l'export nucléaire des ARN messagers de ces organismes mais l'export des protéines et des ARN non codants y sont aussi présentés. Afin d'introduire notre projet de recherche, les avancées récentes sur l'export nucléaire des ARNm chez les plantes sont aussi décrites et comparées avec le mécanisme d'export chez les autres organismes, Le travail bibliographique réalisé au cours de la rédaction de cette introduction générale sera valorisé par la rédaction d'une revue qui sera proposée au journal Virologie, à l'issue de cette thèse.

Dans la seconde partie, vous sera amené notre projet de recherche et présenté un aperçu de la biologie du CaMV. Cette partie a pour but de vous fournir les informations essentielles et nécessaires, à la bonne compréhension de notre étude.

La partie consacrée aux résultats vous montrera dans un premier temps comment ont été créés et/ou mis au point les outils biologiques et cellulaires nécessaires à cette étude. Par la suite, les principaux résultats obtenus au cours de la thèse vous seront présentés sous forme d'un article, qui sera très prochainement soumis à la revue *Nucleic Acids Research* et démontrant l'usurpation par CaMV de la voie majeure d'export nucléaire des ARNm cellulaires pour transporter ses ARN 35S. Enfin, nous terminerons cette partie avec des résultats supplémentaires suggérant également l'utilisation par le CaMV d'une autre voie d'export des ARN, la voie mineure CRM1.

Cette partie sera suivie par une discussion/conclusion où sera mis en perspective l'ensemble des résultats obtenus au cours de ces travaux.

Pour terminer, vous trouverez à la fin du manuscrit, en annexes, le chapitre dédié au « Matériels et méthodes » et, deux articles consacrés, pour le premier (très prochainement soumis), à un protocole mis au point pour la transfection de protoplastes et pour le deuxième (publié), à la formation des viroplasmes denses de CaMV, une autre thématique de recherche du laboratoire, à laquelle j'ai participée au tout début de ma thèse.

Sommaire

Sommaire

Introduction générale **1**

1. Mécanisme général de l'import-export nucléo-cytoplasmique 2
   1.1. Les pores nucléaires 2
   1.2. Translocation des macromolécules à travers les pores nucléaires 4
   1.3. L'export nucléaire des protéines 6
   1.4. Export des ARN non-codants 7
      1.4.1. Le transport des précurseurs des sous-unités ribosomiques 60S et 40S 8
      1.4.2. L'export des ARNt 9
      1.4.3. L'export des miARN 10
      1.4.4. L'export des snARN 10
      1.4.5. L'export des lncARN 12
2. L'export nucléaire des ARN messagers cellulaires 13
   2.1. La voie canonique d'export nucléaire des ARNm 14
      2.1.1. Les récepteurs de l'export nucléaire des ARNm 15
      2.1.2. Les protéines adaptatrices des récepteurs d'export nucléaire 16
      2.1.3. Le complexe TREX-1 17
      2.1.4. Couplage entre la biosynthèse et l'export nucléaire des ARNm 19
      2.1.5. Ancrage des mRNP au pore nucléaire et translocation 21
      2.1.6. Contrôle qualité des ARNm 23
   2.2. Les voies alternatives d'export nucléaire des ARNm 25
      2.2.1. Export des ARNm par la voie ALREX 25
      2.2.2. Export nucléaire des ARNm des histones 26
      2.2.3. Export des ARNm cellulaires par la voie de CRM1 26
3. L'export nucléaire des ARNm chez les plantes 28
   3.1. Le complexe TREX-1 d'Arabidopsis thaliana 29
   3.2. Les protéines adaptatrices d'Arabidop*sis* pour l'export des ARNm 31
   3.3. État de l'art de l'export nucléaire des ARNm chez les plantes 31
4. L'export nucléaire des ARN viraux 33
   4.1. L'export des ARN rétroviraux 34
      4.1.1. L'export chez le virus de l'immunodéficience humaine, un rétrovirus complexe 34
      4.1.2. Le virus Mason-Pfizer du singe, l'export chez les rétrovirus simples 38
      4.1.3. À chacun sa stratégie 39
   4.2. L'export nucléaire des ARNm des virus à ADN 43
      4.2.1. L'export complexe des ARNm des adénovirus 43
      4.2.2. L'export chez les Herpesviridae des ARNm issus de gènes sans introns 46
      4.2.3. Le virus de l'hépatite B, l'export nucléaire chez un pararétrovirus des animaux 47
   4.3. Le virus de la grippe, un virus à ARN qui effectue une partie de son cycle dans le noyau 51

Projet de recherche **59**

1. Ouverture sur le sujet d'étude : L'export des ARN viraux chez les plantes 59
2. Le virus de la mosaïque du chou-fleur 61
   2.1. Plantes hôtes et symptômes 61
   2.2. Le génome du CaMV et ses transcrits 62
   2.3. Fonctions des protéines virales 64
   2.4. Cycle de réplication du CaMV 68
3. Projet de thèse 73

Résultats **77**

1. Les outils biologiques pour l'étude de l'export nucléaire des transcrits du CaMV 77
   1.1. Sélection d'*Arabidopsis thaliana* mutants n'exprimant pas certains facteurs d'export 77
   1.2. Mise au point du système d'étude cellulaire 81

1.2.1. Préparation et transfection des protoplastes d'Arabidopsis 81
1.2.2. Constructions plasmidiques 85
1.2.3. Analyse de l'expression des ARN du CaMV 87
2. Résultats présentés en vue de publication 91
3. Résultats supplémentaires 151
3.1. La voie de l'exportine 1 participe à l'export nucléaire des transcrits 35S du CaMV 151

Discussion et perspectives **159**

Matériel et méthodes **171**

Matériel 171
1. Les plantes 171
1.1. *Arabidopsis thaliana* 171
1.1.1. *Plantes sauvages* 171
1.1.2. Mutants d'insertion 171
1.1.3. Le mutant *CA-rop2* 172
1.2. *Brassica napus* 172
2. Isolats viraux 172
2.1. CaMV 172
2.2. TuMV-GFP 173
3. Bactéries 173
3.1. Souches bactériennes 173
3.2. Milieux de cultures et antibiotiques 173
3.1.1. Milieux de culture bactérienne 173
4. Plasmides 174
4.1. Plasmide pCaMV-GFP 174
4.2. Plasmides du type Litmus28i 175
4.3. Plasmides du type Litmus35S-*Xba*I-tNos-35S-mRFP 175
4.4. Plasmide du type pCK-eGFP 177
4.5. Plasmides du type pGEX-6-P1 178
4.6. Plasmides du type pGAD-T7 et pGBK-T7 178
5. Amorces 178
6. Anticorps 180

Méthodes 181
1. Infections virales 181
1.1. Semis de graines en terre 181
1.2. Inoculation mécanique des plantes 181
1.3. Analyse de l'ADN par PCR 182
1.4. Analyse des protéines par western blot 182
1.4.1. Préparation des échantillons protéiques 182
1.4.2. Électrophorèse sur gel de SDS-polyacrylamide (SDS-PAGE) 183
1.4.3. Coloration au bleu de Coomassie™ 183
1.4.4. Transfert des protéines sur membrane 183
1.4.5. Immunodétection des protéines (western blot) 184
2. Préparation et transfection des protoplastes et fractionnement subcellulaire 185
2.1. Protoplastes de mésophylle d'*A. thaliana* 185
2.2. Transfection des protoplastes d'*A. thaliana* 186
2.3. Fractionnement subcellulaire 187
2.4. Échantillons protéiques de protoplastes 187
2.5. Extraction des ARN 187
3. Quantification des ARN 35S par RT-qRT-PCR 187
3.1. DNase 188
3.2. Rétrotranscription 188

3.3. Quantification par PCR en temps réel 188
3.4. Détection des ARN 35S compétiteurs par RT-PCR 189
4. Analyses des interactions entre protéines 189
4.1. Analyse des interactions protéiques *in vitro* 189
4.1.1. Expression et marquage des protéines à la méthionine ^{35}S 189
4.1.2. Expression de la protéine de fusion GST-MOS11 et de GST 189
4.1.3. Capture des protéines en « batch » 190
4.1.4. Analyses par GST *pull-down* 190
4.2. Analyse de la localisation intracellulaire des protéines de fusion à l'eGFP ou la mRFP par microscopie confocale 191
5. Techniques de biologie moléculaire 191
5.1. Transformation des bactéries compétentes 191
5.1.1. Transformation des bactéries chimio-compétentes 191
5.1.2. Transformation des bactéries électro-compétentes 191
5.2. Purification de l'ADN plasmidique produit dans *E. coli* 192
5.3. Amplification de séquences d'ADN par PCR 192
5.3.1. Amplification de séquences d'ADN par PCR de routine 192
5.3.2. Amplification de séquences d'ADN par PCR dans le cadre du clonage 193
5.4. Électrophorèse en gel d'agarose 193
5.5. Extraction sur gel et purification de l'ADN 193
5.6. Les techniques de clonage 194
5.6.1. Clonage par restriction 194
5.6.2. Clonage par la technique du Gibson assembly™ 194
5.7. Génotypage 195

Annexes **199**

Annexe 1 199

Annexe 2 217

Bibliographie **249**

Abréviations

Abréviations

CaMV : virus de la mosaique du chou-fleur
VIH : virus de l'immunodéficience humaine
HTLV : virus T-lymphotropique humain
MPMV : virus Mason-Pfizer du singe
RSV : virus du sarcome de Rous
PFV : *Prototype foamy virus*
MLV : virus de la leucémie murine
AdV : adénovirus
HSV : virus herpès simplex
KSHV : virus herpès associé au sarcome de Kaposi
EBV : virus d'Epstein-Barr
HBV : virus de l'hépatite B
IAV : virus de la grippe A
RRE : *Rev response element*
CTE : *constitutive transport element*
PRE : *post-transcriptional regulatory element*
P6/TAV : P6/*transactivator viroplasmin*
P5_PR : domaine protéase de P5
P5_RT : domaine transcriptase inverse de P5
P5_RH : domaine RNase H de P5
NPC : *nuclear pore complex*
Nups : *nucleoporins*
NTR : *nuclear transport receptor*
TAP : *tip-associated protein*
TREX : *transcription-export*
ALY : *ally of AML-1 and LEF-1*
CIP29 : *cytokine-induced protein*
MOS11 : *modifier of SNC1, 11*
UAP56 : *U2AF65 associated protein*
UIF : *UAP56-interacting factor*
CRM1 : *chromosome region maintenance 1*
XPO1 : *exportin 1*
NES : *nuclear export signal*
NLS : *nuclear localization signal*

mRNP : *messenger ribonucleoprotein*
CBC : *cap binding complex*
EJC : *exon junction complex*
SR : *serine arginine rich*
CA-rop2 : *constitutively active Rho-related GTPase of plants*

Introduction générale

Introduction générale

L'export nucléaire chez les eucaryotes et leurs virus

Une des différences majeures entre procaryotes et eucaryotes est l'existence d'une enveloppe nucléaire qui sépare le matériel génétique du reste de la cellule. Malgré cette séparation physique, le noyau reste connecté au cytoplasme grâce à d'innombrables pores nucléaires ou NPC – *nuclear pore complexes* – qui fusionnent les membranes nucléaires interne et externe pour former des canaux au travers de l'enveloppe nucléaire (Callan et al., 1949). Le cloisonnement de l'ADN nucléaire a conduit chez les eucaryotes, au découplage de la transcription et de la traduction et, à l'élaboration d'un système de transport nucléo-cytoplasmique bidirectionnel actif – on parle d'import et d'export – d'ARN, de protéines et de complexes ribonucléoprotéiques. Le transport de ces macromolécules au travers des pores nucléaires est très sélectif et finement régulé, spatialement et temporellement, grâce à des transporteurs spécifiques (Köhler et Hurt, 2007; Chatterjee et al., 2018). La régulation de ce transit actif est cruciale pour le métabolisme et la croissance cellulaires (Hayama et al., 2017) comme le démontre en cas de dysfonctionnement, l'apparition de pathologies diverses chez les vertébrés, comme des maladies neurodégénératives et des cancers (Fu et al., 2018b) et chez les plantes, des défauts de croissance et de développement (Ehrnsberger et al., 2019a).

Le franchissement de l'enveloppe nucléaire a constitué un véritable challenge pour les virus à ADN, hormis les *Poxvirus* et les *Megavirus* dont le cycle infectieux est cytoplasmique, (Knipe et Howley, 2013), les rétrovirus et certains autres virus à ARN dont les *Orthomyxovirus,* pour pouvoir infecter les eucaryotes (Cullen, 2003; Le Sage et Mouland, 2013). En effet, les virus à ADN ont besoin d'accéder au noyau pour répliquer leur génome et pour exprimer pleinement leurs gènes de protéines alors que les rétrovirus doivent passer, après rétrotranscription de leur ARN génomique, par une phase obligatoire d'intégration de leur ADN dans les chromosomes de l'hôte. La capacité de ces virus à se reproduire chez ces hôtes et à perturber le métabolisme cellulaire démontre qu'ils ont parfaitement réussi, au cours de l'évolution, à déployer des stratégies efficaces pour détourner le système de transport nucléo-cytoplasmique bidirectionnel de la cellule en leur faveur. En effet, il s'avère que les constituants viraux sont dotés des éléments structuraux adéquats ou les acquièrent pendant l'infection comme chez les *Orthomyxovirus*, pour être reconnus et pris en charge par les machineries cellulaires impliquées dans l'import et l'export nucléaires (York et Fodor, 2013; Okamura et al., 2015). Par ailleurs, les virus s'accaparent de ce système de transport non seulement à leurs propres fins mais aussi pour perturber le passage des composants cellulaires et contrecarrer ainsi les stratégies antivirales de leurs hôtes (Kuss et al., 2013). L'appropriation du transport nucléo-cytoplasmique des eucaryotes explique pourquoi les virus des animaux servent de modèles pour élucider les mécanismes de

l'import-export des protéines et des ARN cellulaires (Knipe et Howley, 2013; Yarbrough et al., 2014).

1. Mécanisme général de l'import-export nucléo-cytoplasmique

Pour des raisons de lisibilité, nous utiliserons principalement la nomenclature des protéines nucléaires des pores et de l'export de l'Homme lorsque les propriétés de ces dernières sont conservées parmi les eucaryotes. Les nomenclatures utilisées chez la levure et les plantes ne seront utilisées que lorsque des aspects spécifiques à ces organismes seront abordés.

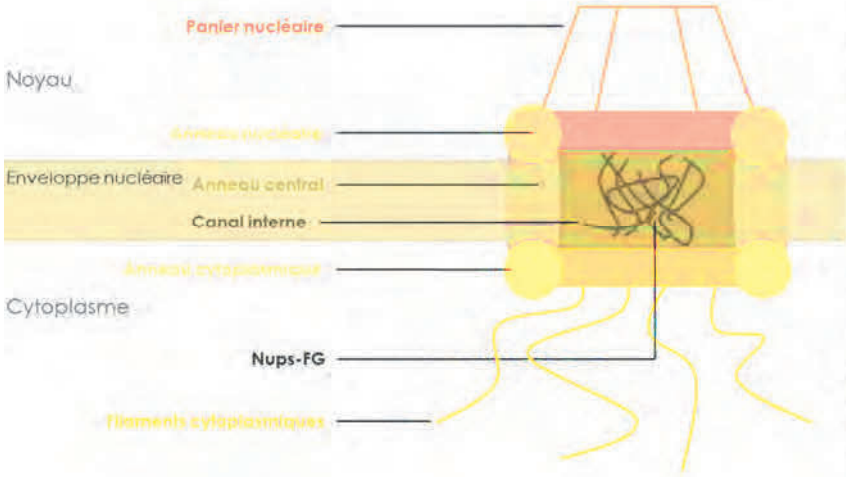
1.1. Les pores nucléaires

Le pore nucléaire est le plus grand complexe protéique eucaryotique, sa masse moléculaire, gigantesque, atteignant les 125 MDa chez l'Homme (Reichelt et al., 1990). Il possède une symétrie rotationnelle d'ordre 8, en d'autres termes, le NPC est constitué de 8 unités protéiques organisées autour d'un axe central. Ces unités ont une face nucléaire associée à une structure en forme de panier et une face cytoplasmique d'où émergent 8 filaments cytoplasmiques. Ces derniers ne sont pas des filaments à proprement dit, comme l'actine, mais des protéines avec des domaines non ordonnés très flexibles. L'intérieur du pore nucléaire consiste en un tunnel central d'un diamètre de 40 nm et d'une longueur de 80 nm chez l'Homme, au travers duquel passent les ions, les métabolites et les macromolécules (Adam, 2001; Kabachinski et Schwartz, 2015).

Malgré leur grande taille, les NPC ne comprennent qu'une trentaine de protéines différentes appelées nucléoporines, en abrégé Nups (le chiffre associé correspond à la masse moléculaire), dont la structure globale est relativement bien conservée chez les différents organismes (Maul et al., 1972; Alber et al., 2007; Grossman et al., 2012). Les Nups sont présentes au sein du NPC en plusieurs exemplaires à raison d'un multiple de 8 et ce, jusqu'à 64 fois. Elles sont classées sur le plan fonctionnel en trois catégories : les Nups transmembranaires (ou POMs) qui ancrent le pore dans l'enveloppe nucléaire (Neumann et al., 2010), les Nups d'échafaudage qui ont une fonction architecturale et les Nups-FG impliquées directement dans le trafic nucléo-cytoplasmique (Beck et Hurt, 2017) (**figure I-1**). Le complexe Y ou NUP107 chez l'Homme (Kelley et al., 2015) est une des principales Nups d'échafaudage du pore. Il est à l'origine des anneaux cytoplasmique et nucléaire et, avec l'aide du complexe NUP93, il forme l'anneau central interne (Vollmer et Antonin, 2014). Les Nups-FG sont des protéines non structurales dont la particularité est de contenir jusqu'à une cinquantaine de motifs répétés de 4 acides aminés (FXFG et GLFG) riches en phénylalanines (F) et glycines (G) d'où leur dénomination. La majorité des Nups-FG décorent le canal central du NPC et jouent un rôle primordial dans le transport actif des macromolécules (Terry et Wente, 2009). Toutefois, des Nups-FG sont aussi détectées dans les parties cytoplasmique et nucléaire des pores et

participent à leur assemblage grâce à la présence de motifs d'interaction protéine-protéine. Des modifications post-traductionnelles de Nups (N- et O- glycosylation, phosphorylation, ubiquitylation, sumoylation) seraient impliquées dans la liaison entre les NPC et le cytosquelette ainsi que dans la ségrégation des NPC lors de la mitose (Beck et Hurt, 2017).

De nombreuses protéines solubles sont également associées aux NPC et jouent un rôle essentiel dans le transit nucléo-cytoplasmique actif des molécules. Les mieux caractérisées d'entre-elles sont les importines α et β (Li et Thomas, 1989; Görlich et al., 1994), la GTPase Ran (Melchior et al., 1993; Moore et Blobel, 1993), le facteur de transport nucléaire 2 NTF2 (Smith et al., 1998a), l'exportine 1, plus connue sous le nom de CRM1 - *chromosome region maintenance 1* - (Fukuda et al., 1997) et l'exportine 2 (Kutay et al., 1997) aussi connue sous le nom de CAS - *cell apoptosis susceptibility* – .



Nups nucléaires

Levure	Homme	Plantes
Nup1	-	NUP1
Nup2	NUP50	NUP50
Nup60	NUP153	NUP136
Mlp1	TPR	NUA
Mlp2	-	-

Complexe Y

Levure	Homme	Plantes
Nup133	NUP133	NUP133
Nup84	NUP107	NUP107
Nup85	NUP75	NUP75
Nup120	NUP160	NUP160
Nup145C	NUP96	NUP96
Sec13	SEC13	SEC13
seh1	SEH1	SEH1
-	NUP37	-
-	NUP43	NUP43

Complexe de l'anneau central

Levure	Homme	Plantes
Nic96	NUP93	NUP93
Nup53/59	NUP35	NUP35
Nup192	NUP205	NUP205
Nup188	NUP188	NUP188
Nup170/157	NUP155	NUP155

Nups transmembranaires

Levure	Homme	Plantes
Ndc1	NDC1	NDC1
Pom152	GP210	GP210
Pom34	POM121	-
Pom33	-	-

Nups-FG

Levure	Homme	Plantes
Nsp1	NUP62	NUP62
Nup49	NUP58	NUP58
Nup57	NUP54	NUP54
Nup145N	NUP98	NUP98
Nup100	-	-
Nup116	-	-

Nups cytoplasmiques

Levure	Homme	Plantes
Nup159	NUP214	NUP214
Nup42	CG1	CG1
Nup82	NUP88	NUP88
Gle2	RAE1	RAE1
Gle1	GLE1	GLE1
-	NUP358	-
-	ALADIN	ALADIN
-	ELYS	ELYS

Figure I-1 | Le pore nucléaire. Le pore nucléaire ou NPC, enchâssé dans l'enveloppe nucléaire, est composé d'environ trente nucléoporines ou Nups différentes, présentes en multiples exemplaires. On distingue du noyau vers le cytoplasme, le panier nucléaire suivi de l'anneau nucléaire, l'anneau central dont le canal interne est tapissé de Nups-FG et l'anneau cytoplasmique paré de 8 filaments. La localisation et le nom des différentes Nups structurant le pore nucléaire sont donnés pour la levure, l'Homme et les plantes. D'après Grossman et al., 2012; Tamura et Hara-Nishimura, 2013; Beck et Hurt, 2017.

1.2. Translocation des macromolécules à travers les pores nucléaires

Avant de voir le mécanisme du transport nucléo-cytoplasmique à travers les NPC, notons que ces derniers interviennent également dans l'organisation spatiale de complexes localisés à la périphérie interne du noyau comme les complexes TREX – *transcription-export* -, impliqués dans l'assemblage et la surveillance des mRNP avant leur export dans le cytosol ou encore, dans l'organisation spatiale de la chromatine (Beck et Hurt, 2017).

Les pores nucléaires n'entravent pas le passage des ions, des métabolites et des petites molécules de moins de 40 kDa. Ceux-ci diffusent passivement dans les deux sens entre noyau et cytoplasme (Kabachinski et Schwartz, 2015; Knockenhauer et Schwartz, 2016). En revanche, les NPC constituent une barrière sélective pour les ARN (quelle que soit leur taille), les protéines et les complexes de plus de 40 kDa, ces macromolécules ne pouvant les traverser qu'avec l'assistance d'un ou plusieurs facteurs de transport appelés NTR - *nuclear transport receptor* -. Cette barrière est matérialisée par un réseau dense et dynamique formé par les motifs FG d'une dizaine de Nups-FG qui décorent le canal interne du NPC. Diverses études ont montré que les motifs FG répétés du complexe NUP62 (NUP62, NUP58 et NUP54) et de NUP98, sont localisés à la surface interne du canal interne (Allen et al., 2001; Frey et al., 2006). Les NTR interagissent transitoirement avec une faible affinité avec les motifs FG ce qui leurs permet de traverser le pore nucléaire, seuls ou associés à leur cargo (Schmidt et Görlich, 2016). Selon le modèle le plus accepté, les NTR ouvrent transitoirement et localement les mailles du réseau de Nups-FG pour atteindre le cytoplasme (Ribbeck et Görlich, 2001). Peu importe le NTR ou la nature des cargos (protéines, ARNt, sous-unités ribosomiques, petits ARN, ARNm ou encore virus), la translocation à travers les NPC s'effectue toujours selon ce même principe. Ce transit effectué par les NTR est très efficace vu qu'on a déterminé que le passage des complexes NTR-protéines à travers les pores nucléaires ne prend en moyenne que 5 millisecondes (Sakiyama et al., 2017), en d'autres termes, environ 200 protéines traversent chaque pore nucléaire par seconde (Grünwald et al., 2011). Le trafic entre les deux compartiments cellulaires est d'autant plus dense que le nombre de pores par noyau peut varier de 50 NPC pour le noyau des cellules de *Nicotiana tabacum* BY-2, jusqu'à 5000 pour celui des cellules humaines (Chatterjee et al., 2018).

Les cargos sont généralement reconnus par les NTR grâce à des domaines de transport nucléaire appelés NLS - *nuclear localisation signal* - pour l'import nucléaire, ou NES – *nuclear export signal* - pour l'export nucléaire. La plupart de ces NTR appartient à la famille des karyophérines β qui effectuent la navette nucléo-cytoplasmique. L'association entre les karyophérines et les cargos est modulée par une petite GTPase, Ran, qui est présente sous deux formes, Ran-GTP et Ran-GDP. Un taux élevé de Ran-GTP est maintenu dans le noyau grâce à un facteur d'échange de nucléotides lié à la chromatine, la protéine Ran-GEF – *Ran-guanine nucleotide exhange factor* - encore connue sous le nom de RCC1 – *regulator of chromosome condensation 1*- qui promeut la substitution du GDP par le GTP. Quant à la forme Ran-GDP, elle

prédomine dans le cytoplasme car le GTP associé à Ran est rapidement hydrolysé après l'export, par l'activité GTPasique de Ran. Cette hydrolyse est stimulée par le facteur Ran-GAP – *Ran-GTP-activating protein* – et, par RanBP1 – *Ran-GTP binding protein 1* – dont le rôle est de sensibiliser la GTPase à l'action de Ran-GAP. Cette dernière est associée pour une grande partie, aux filaments cytoplasmiques du NPC *via* une interaction indirecte avec NUP358 encore appelée RanBP2 – *Ran-GTP binding protein 2* –. C'est ce gradient de Ran-GTP entre noyau et cytoplasme qui dicte la directionnalité du transport nucléo-cytoplasmique dépendant des karyophérines (**figure I-2**)(Fried et Kutay, 2003).

En résumé, une exportine ne peut se lier à son cargo qu'en association avec Ran-GTP alors qu'à l'inverse, dans le mécanisme d'import, la liaison de Ran-GTP, dans le noyau, au complexe d'import provoque sa dissociation, libérant ainsi le cargo.

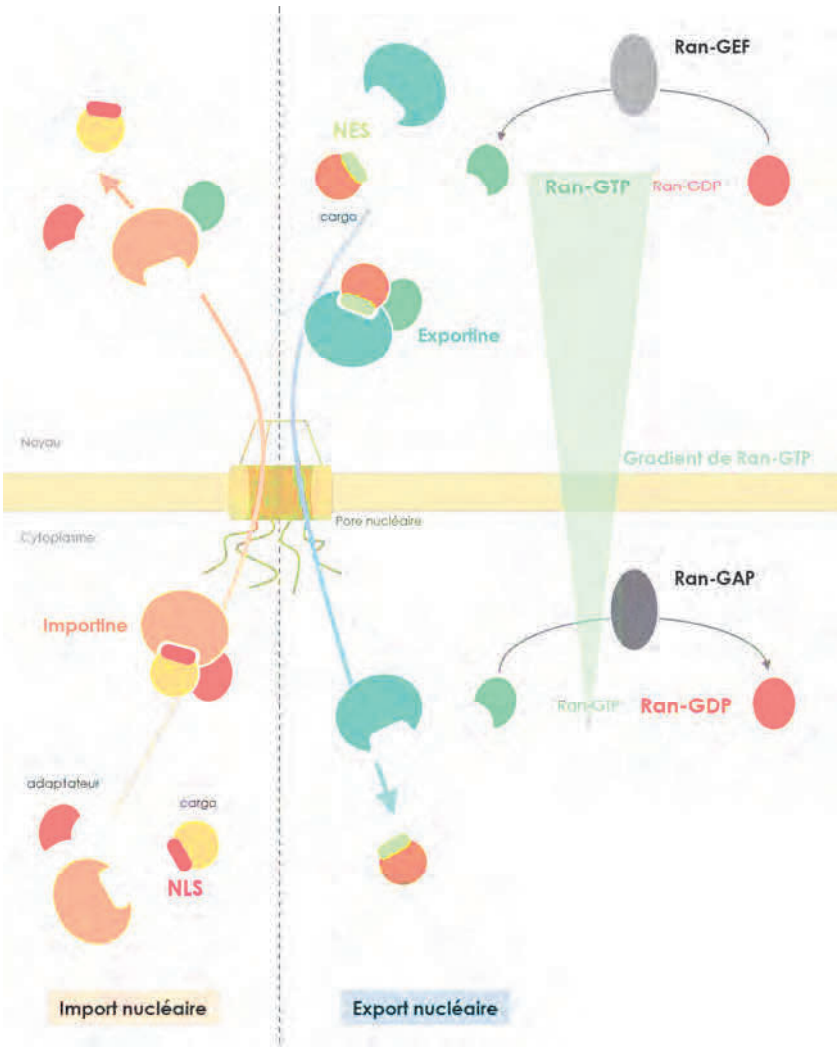


Figure I-2 | Représentation schématique du transport nucléo-cytoplasmique Ran-dépendant. Dans ce modèle, les importines, comme l'importine β et son adaptateur l'importine α, reconnaissent les cargos cytoplasmiques grâce à leur NLS. Elles les transportent jusqu'au noyau puis à travers les pores nucléaires en interagissant avec les motifs répétés des Nups-FG du canal central. Dans le noyau, la liaison à la GTPase Ran-GTP permet la dissociation du complexe. La forte concentration intranucléaire en Ran-GTP est assurée par le facteur Ran-GEF. Les exportines, comme CRM1, liées à Ran-GTP reconnaissent les cargos nucléaires grâce à leur NES. Ces complexes ternaires traversent activement les NPC et rejoignent le cytoplasme où l'hydrolyse du GTP en GDP stimulée par Ran-GAP, entraîne le relarguage du cargo. D'après Grossman et al., 2012; Beck et Hurt, 2017.

1.3. L'export nucléaire des protéines

Toute protéine nucléaire portant un NES est reconnue et exportée par l'exportine 1 (XPO1), plus connue sous le nom de CRM1, un membre de la famille des karyophérines β. C'est le facteur d'export par excellence des protéines et aussi celui qui est le mieux caractérisé (Ishizawa et al., 2015).

Le signal d'export nucléaire (NES) est un domaine protéique riche en résidus hydrophobes [Φ] (leucine, valine, isoleucine, phénylalanine ou méthionine) de séquence consensus Φ_1-$X_{(2-3)}$-Φ_2-$X_{(2-3)}$-Φ_3-X-Φ_4 où les positions X peuvent être occupées par n'importe lequel des 20 acides aminés (Cautain et al., 2015; Fu et al., 2018a). Les premiers NES ont été identifiés, il a plus de 24 ans chez la protéine Rev du virus de l'immunodéficience humaine de type 1 (VIH-1) et l'inhibiteur PKI de la protéine kinase A cellulaire (Fischer et al., 1995; Wen et al., 1995). Depuis, des NES ont été caractérisés sur des centaines de protéines de fonctions et structures différentes, par analyse des séquences puis confirmés par mutagenèse couplée à l'emploi de la leptomycine B, un inhibiteur spécifique de l'export des protéines (Mathew et Ghildyal, 2017). Il s'est avéré que les NES ne s'alignent pas toujours à la séquence canonique décrite ci-dessus (Fu et al., 2013; Xu et al., 2012, 2015; Kırlı et al., 2015; Fu et al., 2018a). L'exposition des NES est parfois induite par des changements conformationnels des protéines suite à des modifications post-traductionnelles telles que la phosphorylation, l'ubiquitylation, la sumoylation et l'acétylation, ou suite à la liaison de co-facteurs, permettant ainsi la reconnaissance des protéines à exporter par CRM1 (Turner et al., 2012).

Le gène *XPO1,* codant pour CRM1, a été identifié pour la première fois chez la levure *Schizosaccharomyces pombe,* comme étant impliqué dans le contrôle et le maintien de la structure des chromosomes chez cette levure (Adachi et Yanagida, 1989) expliquant pourquoi la protéine correspondante a été dénommée CRM1. Il faudra attendre plus de 8 ans avant qu'elle ne soit caractérisée et identifiée comme le principal NTR pour l'export des protéines à NES (Ossareh-Nazari et al., 1997). Cette karyophérine β possède des domaines HEAT - *Huntington, Elongation factor 3, protein phosphatase PP2A, protein kinase TOR* - qui lui permettent de se lier à Ran-GTP. Le sillon formé par les répétitions HEAT 11 et 12 renferme la poche hydrophobe qui interagit avec les résidus hydrophobes des NES (Lindsay et al., 2001). Cette poche est la cible de la leptomycine B qui, en s'y fixant de manière covalente à un résidu cystéine, inhibe non seulement l'export nucléaire des cargos (Kudo et al., 1999; Sun et al., 2013) mais également, l'import nucléaire de CRM1 (Rahmani et Dean, 2017).

CRM1 ne se lie que faiblement à Ran-GTP ou au NES des protéines mais son affinité pour ces deux substrats est 500 à 1000 fois plus élevée lorsqu'elle s'y fixe simultanément et forme ainsi un complexe ternaire suffisamment stable pour permettre l'export du cargo vers le cytoplasme (Dong et al., 2009a, 2009b). La reconnaissance du cargo par CRM1 est stimulée par la protéine RanBP3 – *Ran binding protein 3* – dont la fonction serait de maintenir une forte concentration de

Ran-GTP à proximité de CRM1 et d'augmenter l'affinité de CRM1 pour le NES (Lindsay et al., 2001). À la sortie du canal central du NPC, le complexe d'export est dissocié aux abords des filaments suite à l'hydrolyse du GTP en GDP par la GTPase Ran. À la fin de l'export, CRM1 est re-importée dans le noyau pour prendre en charge un nouveau cargo.

CRM1 exporte également des protéines dépourvues de NES mais dans ce cas, elle fait intervenir un adaptateur qui lui permet de charger le cargo ; citons par exemple, la protéine kinase LKB1 et son adaptateur STRADα – *Ste20-related kinase adapter protein alpha* – (Dorfman et Macara, 2008).

Les métazoaires disposent de voies d'export de protéines alternatives à celle médiée par CRM1. Certaines d'entre elles sont spécialisées dans le transport nucléo-cytoplasmique d'un cargo unique comme CAS/exportine 2, un membre de la famille des karyophérines β, qui exporte seulement l'importine α (Oka et Yoneda, 2018; Kutay et al., 1997). D'autres NTR exportent plusieurs protéines différentes, comme l'exportine 7 (XPO7) qui est impliquée dans la translocation nucléo-cytoplasmique de diverses protéines (Aksu et al., 2018; Mingot et al., 2004). Quant à la calréticuline, elle exporte le récepteur aux glucocorticoïdes, de manière calcium dépendante, et le récepteur α1 aux hormones thyroïdiennes (Grespin et al., 2008; Holaska et al., 2001). Enfin, signalons que des NTR de vertébrés sont doués à la fois d'activités d'export et d'import comme l'exportine XPO4 qui exporte le facteur de traduction eIF5A et importe des facteurs de transcription dont SOX2 (Gontan et al., 2009; Lipowsky et al., 2000; Aksu et al., 2016).

1.4. Export des ARN non-codants

Comme pour les protéines, les pores nucléaires sont le siège d'un flux permanent des ARN codants, autrement dit, les ARN messagers (ARNm) et, de nombreux ARN non-codants, de taille, de structure et de fonction différentes (Cech et Steitz, 2014). Parmi ces multiples ARN, nous décrirons les mécanismes de l'export nucléaire, des ARN ribosomiques (ARNr), des ARN de transfert (ARNt) et des microARN (miARN) impliqués dans la traduction et/ou sa régulation, les petits ARN nucléaires (snARN) du spliceosome, le complexe engagé dans l'épissage des précurseurs des ARNm d'origine nucléaire et enfin, les longs ARN non-codants (lncARN) qui jouent un rôle dans des processus cellulaires très divers aussi bien dans le noyau que dans le cytoplasme : remodelage de la chromatine, échafaudage de complexes, titration de protéines fixant l'ARN, modulation des voies de signalisation etc. (Noh et al., 2018; Ulitsky et Bartel, 2013). Les mécanismes impliqués dans l'export nucléaire des ARNm seront détaillés dans une partie qui leur sera entièrement dédiée.

Le transport nucléo-cytoplasmique des ARN non-codants est un processus actif bien que leur masse moléculaire, hormis celle des ARNr et des lncARN, soit nettement inférieure à celle permettant théoriquement leur diffusion passive à travers les NPC.

1.4.1. Le transport des précurseurs des sous-unités ribosomiques 60S et 40S

Les ribosomes sont composés chez l'Homme, de 80 protéines et de 4 ARNr : la grande sous-unité 60S comprend les ARNr 28S, 5,8S et 5S et, 47 protéines et la petite sous-unité 40S, l'ARNr 18S et 33 protéines. Les deux sous-unités ribosomiques sont synthétisées et assemblées en plusieurs étapes, dans le nucléole sous la forme de gros complexes ribonucléoprotéiques (RNP) d'environ 2 MDa, les précurseurs des sous-unités 60S (pré-60S) et 40S (pré-40S). Ces précurseurs sont exportés dans le cytoplasme où est alors achevée leur maturation (Nerurkar et al., 2015). L'export des complexes pré-60S et pré-40S requiert pour leur translocation à travers le NPC, l'intervention de nombreux facteurs, certains sont communs, d'autres spécifiques à chacun des deux précurseurs ribosomiques (Gerhardy et al., 2014; Nerurkar et al., 2015).

Ainsi, le transport nucléo-cytoplasmique des complexes pré-60S et pré-40S implique le récepteur CRM1 et les composants du système Ran décrits précédemment, l'hétérodimère TAP-p15 ou Mex67/Mtr2 chez *Saccharomyces cerevisiae* (Cole et Scarcelli, 2006; Yao et al., 2007) et, la protéine d'assemblage RRP12 (Nerurkar et al., 2015).

L'export de pré-60S nécessite un adaptateur, la protéine NMD3 qui permet de recruter spécifiquement CRM1 au niveau du précurseur ribosomique. Le rôle de NMD3 s'explique par le fait qu'elle possède deux NES alors que les protéines ribosomiques en sont dépourvues. Suite à l'export, CRM1 est dissociée de son adaptateur NMD3 par hydrolyse de Ran-GTP, puis NMD3 est relarguée de la sous-unité ribosomique 60S par une GTPase cytoplasmique (Lsg1 chez la levure) et remplacée par la protéine ribosomique RPL10 (Hedges et al., 2005). Des récepteurs non essentiels intervenant dans l'export de pré-60S ont également été identifiés, ARX1, BUD20 et ECM1(Hung et al., 2008; Yao et al., 2008; Sengupta et al., 2010). Il ressort de ces études qu'il faut plutôt un *quorum* qu'un jeu complet de récepteurs pour assurer une translocation nucléo-cytoplasmique efficace des pré-60S (**figure 1-3 d.**). En effet, vu la taille du complexe ribosomique, sa prise en charge par un récepteur unique s'avère insuffisant pour sa translocation, le rôle des récepteurs étant à la fois, d'interagir avec les domaines FG des Nups-FG et de masquer les charges à la surface de pré-60S et faciliter ainsi le transit à travers le canal hydrophobe du pore.

L'export nucléaire du complexe pré-40S n'est pas aussi bien caractérisé que celui de pré-60S (Hurt et al., 1999; Stage-Zimmermann et al., 2000; Faza et al., 2012). Contrairement à ce dernier, l'export de pré-40S n'implique pas NMD3 mais dépendrait d'au moins deux autres protéines adaptatrices, porteuses de NES, LTV1 et RIO2 qui interagissent avec CRM1. Il a été montré chez la levure que le recrutement de CRM1 par Rio2 est facilité par l'assistance de Slx9 (Fischer et al., 2015), une protéine qui intervient dans le clivage du pré-ARNr et permet ainsi la séparation des précurseurs pré-66S et pré-43S lors de la biogenèse du ribosome (Bax et al., 2006;

Chatterjee et al., 2018). Il a été également montré que le transport de pré-40S par CRM1 nécessite l'intervention toute particulière de la protéine ribosomique RPS15 (Rouquette et al., 2005).

1.4.2. L'export des ARNt

Le transport nucléo-cytoplasmique des ARNt est bidirectionnel, contrairement à celui des ARNr. En effet, après un premier export nucléaire, ces ARNt sont ré-importés dans le noyau, un processus appelé import rétrograde, pour achever la modification de leurs nucléotides et pour un contrôle qualité, puis finalement ré-exportés dans le cytosol. Initialement identifié chez la levure (Kramer et Hopper, 2013), ce passage répété des ARNt est conservé chez de nombreux organismes y compris chez l'Homme (Chatterjee et al., 2018).

Les ARNt sont transcrits par l'ARN polymérase III sous la forme d'un précurseur qui subit dans le noyau plusieurs étapes de maturation : élimination des extrémités 5' et 3' par des endo- et exonucléases dont la RNase P, addition en 3' de la séquence CCA qui est indispensable pour l'aminoacylation des ARNt, épissage si besoin et modifications des nucléotides (Lund et Dahlberg, 1998; Cook et al., 2009; Hopper, 2013).

Les études réalisées chez différents organismes démontrent que les machineries d'export sélectionnent préférentiellement les ARNt matures sauf chez la levure où les pré-ARNt contenant un intron sont d'abord exportés vers le cytoplasme où ils sont épissés par une endonucléase localisée au niveau de la membrane mitochondriale externe (Kramer et Hopper, 2013). Après épissage, ces ARNt sont importés dans le noyau où ils subiront d'autres modifications post-transcriptionelles puis re-exportés dans le cytoplasme pour intervenir dans la traduction (Yoshihisa et al., 2007; Ohira et Suzuki, 2011). Les exportines se fixent directement sur les ARNt au niveau des extrémités 3' et 5' parfaitement finies et, des bras D et T de la structure tertiaire, gage de qualité des ARNt avant leur export du noyau.

L'export des ARNt est Ran-GTP-dépendant (Okamura et al., 2015) et effectué par des importines β, l'exportine Los1 chez la levure, l'exportine T (XPOT ou XPO3) chez l'Homme et PAUSED chez Arabidopsis (Arts et al., 1998; Hellmuth et al., 1998; Hunter et al., 2003) (**figure I-3 a.**). Toutefois, d'autres voies d'export des ARNt existent chez la levure et les métazoaires comme le montrent l'étude du mutant Δ*LOS1* de *S. cerevisiae* qui reste viable, le fait qu'aucun orthologue de Los1 n'a été identifié chez la drosophile ou encore, que l'inhibition de l'exportine T par des anticorps ne bloque pas totalement l'export des ARNt chez les vertébrés (Arts et al., 1998; Sarkar et Hopper, 1998; Lippai et al., 2000). Chez la levure, l'export des ARNt matures requiert l'exportine Msn5 qui ne permet l'export que des ARNt parfaitement finis alors que chez les vertébrés et les plantes, ses orthologues, l'exportine 5 (XPO5) et HASTY ne semblent jouer qu'un rôle mineur dans le transport de ces molécules (Takano et al., 2005; Shibata et al., 2006; Huang et Hopper, 2015). Enfin, les exportines TAP-p15 et Mex67/Mtr2 peuvent également être

impliquées dans l'export des ARNt ainsi que CRM1 (Wu et al., 2015; Chatterjee et al., 2017). Il a été montré par exemple que p15 stimule l'export des ARNt dans des cellules HeLa et que la surexpression de l'hétérodimère Mex67/Mtr2 chez la levure, permet de compenser le phénotype des cellules Δ*LOS1* (Chatterjee et al., 2017; Ossareh-Nazari et al., 2000).

1.4.3. L'export des miARN

Les microARN (miARN) sont des petits ARN non codants d'environ 22 nucléotides (nt), impliqués dans une large palette de processus cellulaires tels que le développement, la différenciation cellulaire, l'apoptose ou encore l'immunité anti-virale *via* le mécanisme de *RNA silencing* (Bartel, 2004; Valencia-Sanchez et al., 2006; Leung, 2015).

Leur biogenèse fait intervenir plusieurs étapes de clivages enzymatiques qui ont lieu à la fois dans le noyau et le cytoplasme. Les microARN primaires (pri-miARN) sont synthétisés dans le noyau par les ARN polymérases II ou III à partir de leurs propres gènes ou excisés des introns des transcrits messagers primaires. Les pri-miRNA sont coiffés en 5' et polyadénylés en 3' et renferment une grande tige-boucle qui contient le futur miARN (Cai et al., 2004). Ils sont d'abord clivés dans le noyau, à la base de la tige-boucle par la RNase de type III DROSHA assistée par le co-facteur DGCR8, pour générer un précurseur d'environ 65 pb, le pré-miARN. Ce dernier est exporté vers le cytoplasme par l'exportine 5 (XPO5) associée à Ran-GTP qui reconnaît la base du pré-miRNA et plus précisément, son extrémité 3' qui est plus longue, en général de 2 nucléotides que son extrémité 5' à laquelle elle est appariée (Lund et al., 2004; Okada et al., 2009; Ha et Kim, 2014). Ce ne sont donc pas les miARN matures qui sont exportés mais leurs précurseurs (**figure I-3 b.**).

XPO5 est l'exportine canonique des miARN ; elle fait également partie de la famille des karyophérine β. Ses orthologues chez la levure et les plantes sont respectivement Msn5 et HASTY (Shibata et al., 2006). Comme vu ci-dessus, ces exportines peuvent aussi exporter les ARNt voire dans le cas de XPO5, le facteur de traduction eEF1A à condition que celui-ci soit associé à un ARNt aminoacylé (Calado et al., 2002). Une fois dans le cytoplasme, l'hydrolyse du GTP en GDP permet le relarguage des pré-miRNA et leur clivage par la RNase de type III DICER générant ainsi un duplex miARN de 22 nt. L'un des deux brins de ce duplexe est pris en charge par les protéines ARGONAUTE (AGO1-4, chez l'Homme) pour former le complexe RISC - *RNA-induced silencing complex* - (Kim et al., 2016) qui dirige l'appariement du miARN à l'ARNm cible ce qui conduit à la répression traductionnelle ou à la dégradation de cet ARNm.

1.4.4. L'export des snARN

L'export nucléaire des *small nuclear RNAs* (snARN) est une étape primordiale dans la biogenèse du spliceosome chez les eucaryotes. Ces petits ARN non codants dits ARN U sont transcrits par l'ARN polymérase II, à l'exception du snARN U6 produit par l'ARN polymérase III,

sous forme de précurseurs ou pré-snARN. Ces derniers sont coiffés en 5' mais ne possèdent pas de queue poly A en 3' car cette extrémité est clivée co-transcriptionnellement, comme pour les ARNm, par le complexe CPSF - *cleavage and polyadenylation specificity factor* - chez les plantes et INT1, son paralogue chez les métazoaires. Une exonucléase 3'-5' finit la maturation de l'extrémité 3' dans le cytoplasme (Medlin et al., 2003; Ohtani, 2017). Pour cela, les snARN empruntent la voie de l'exportine CRM1 (Fischer et al., 1995; Fornerod et al., 1997) pour atteindre le cytoplasme où ils formeront, en s'associant avec des protéines, les complexes snRNP fonctionnels avant leur retour dans le noyau.

Après leur transcription/maturation, les snARN sont reconnus par la protéine adaptatrice PHAX phosphorylée qui se fixe sur le complexe CBC - *cap binding complex* - associé à leur coiffe ; ce complexe est composé des protéines NCBP1 et NCBP2 (Cbp80 et Cbp20 chez la levure)(Görlich et al., 1996). PHAX possède un NES ce qui permet à CRM1 de transporter de manière Ran-GTP dépendante, les snARN vers le cytoplasme (Izaurralde et al., 1995; Ohno et al., 2000) (**figure I-3 c.**). L'hydrolyse du GTP par Ran et la déphosphorylation de PHAX permettent la dislocation du complexe et le relarguage des snARN dans le cytosol.

Dans le cytoplasme, le complexe protéique SMN – *survival motor neurone* –, facilite l'assemblage des 7 protéines Sm (*hommage à Stéphanie Smith ayant un lupus érythémateux disséminé et chez qui on a trouvé des anticorps dirigés contre ces protéines*) sur un site de liaison qui est conservé chez tous les snARN. La fixation des protéines Sm induit la triméthylation de la coiffe (2,2',7mG) des snARN et la maturation de leur extrémité 3'(Kiss, 2004; Lührmann, 1990). La coiffe triméthylée (TMG) et les protéines Sm fournissent le signal pour la localisation nucléaire des snRNP. Ces derniers sont importés dans le noyau grâce à trois facteurs : la snurportine-1 (SPN1) qui reconnaît le TMG des snARN, l'importine β qui s'associe avec SPN1 et le complexe SMN qui joue le rôle d'adaptateur entre les protéines Sm et cette importine (Patel et Bellini, 2008). Dans le noyau, les snRNP relargués formeront avec des facteurs protéiques, un spliceosome fonctionnel pour l'épissage des pré-ARNm (Yong et al., 2004).

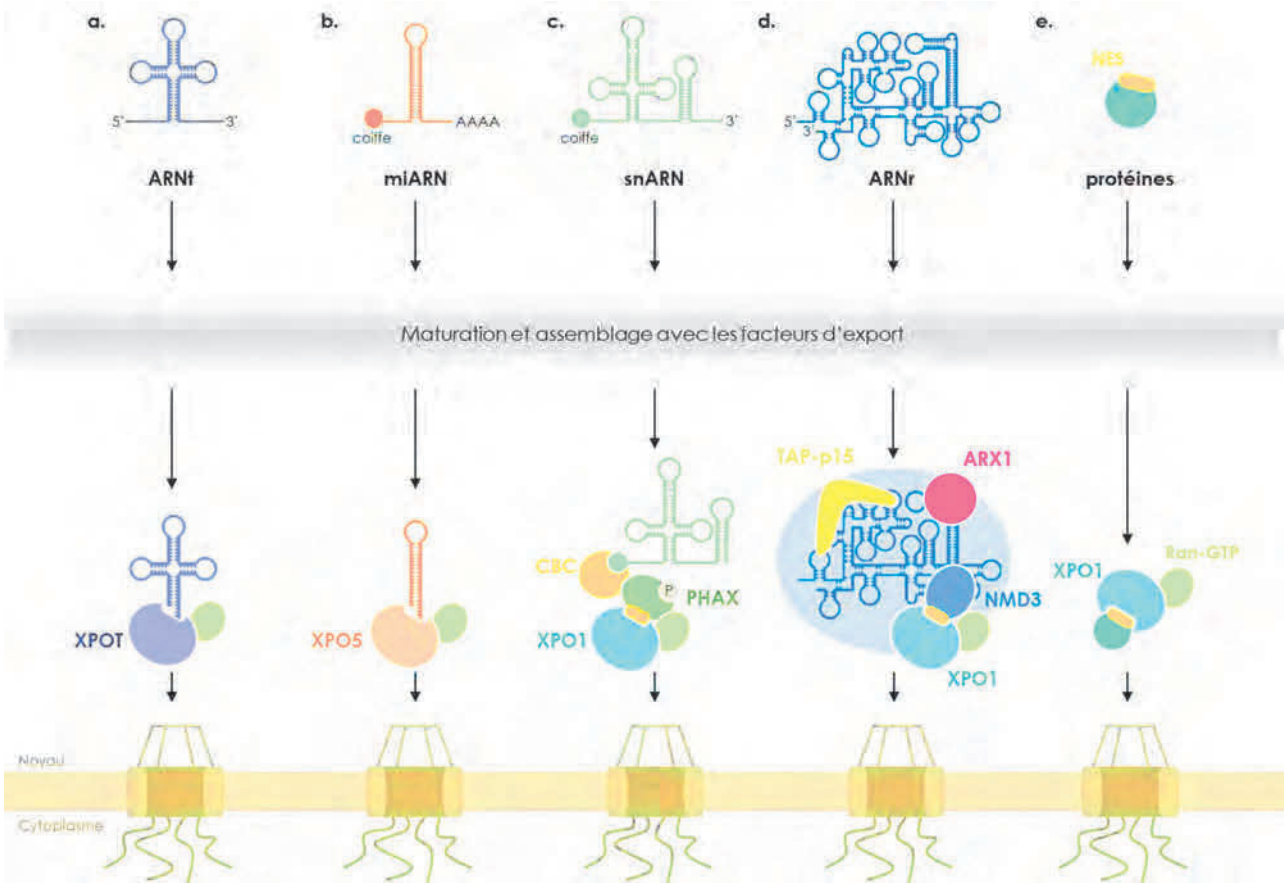


Figure I-3 | Export des ARNt, des miARN, des snARN, des ARNr et des protéines, sous le contrôle du gradient de Ran-GTP. Chacune des voies d'export présentées ci-dessus sont représentées de façon schématique. Les ARNt (**a**) reconnus et exportés par l'exportine t (XPOT). L'extrémité 3' des pré-miARN (**b**) est reconnue par l'exportine 5 (XPO5) qui les transporte vers le cytoplasme. La protéine adaptatrice PHAX interagit avec le CBC fixé sur la coiffe des snARN (**c**) et permet leur export par XPO1 qui reconnait le NES de PHAX. Le précurseur de la sous-unité ribosomique 60S (**d**) est exporté par la voie de XPO1/CRM1, grâce à la protéine adaptatrice NMD3 et en présence d'autres facteurs d'export comme ARX1 et TAP-p15, qui sont nécessaires pour un transport efficace. L'export des protéines à NES (**e**) se fait par la voie de XPO1/CRM1. Modifiée d'après Köhler et Hurt, 2007.

1.4.5. L'export des lncARN

Les lncARN sont des ARN non codants coiffés et polyadénylés d'une taille minimale de 200 nt (Cech et Steitz, 2014) présents dans le noyau et le cytoplasme. Ces ARN ont la capacité d'interagir avec d'autres ARN et/ou avec des protéines se liant à l'ARN ce qui explique l'étendue de leurs fonctions au sein de la cellule (voir **page 7**)(Noh et al., 2018; Ulitsky et Bartel, 2013).

Les lncARN cytoplasmiques utilisent la même voie d'export que certains ARNm à savoir l'exportine CRM1 assistée par l'antigène humain R (HuR) (Dong et al., 2017). Cette protéine adaptatrice recrute CRM1 au niveau des extrémités 5' et 3' du lncARN. Il s'est avéré que les lncARN qui sont exportés, se distinguent des lncARN à résidence nucléaire, par l'absence des séquences, généralement riches en C, impliquées dans la rétention nucléaire des lncARN (Palazzo et Lee, 2018).

2. L'export nucléaire des ARN messagers cellulaires

La biogenèse et l'export nucléaire des ARNm sont des processus intimement liés au cours desquels les ARNm font l'objet d'un contrôle permanent de leur qualité de messager avant leur translocation vers le cytoplasme. Lorsque des erreurs apparaissent lors de la transcription et/ou de la maturation, il y a généralement arrêt de la synthèse des ARNm puis dégradation. Ce système de contrôle évite l'export des ARNm aberrants dont la traduction produirait des protéines non conformes qui seraient délétères pour le métabolisme de la cellule.

Les gènes de protéines sont exprimés sous la forme de précurseurs immatures, les pré-ARNm, qui subiront pendant leur transcription, des modifications structurales et de conformation pour que les ARNm soient traduisibles et exportables vers le cytoplasme : coiffage, épissage et polyadénylation. Des centaines de protéines, souvent associées sous forme de complexes, interviennent dans ces processus et plusieurs d'entre elles restent liées aux ARNm pour former un complexe ribonucléoprotéique (mRNP). Ces protéines ont un rôle très important car elles permettent le recrutement des transporteurs impliqués dans l'export nucléaire et/ou le remodelage des mRNP.

Avant de voir le couplage entre la formation des mRNP et leur export nucléaire, il nous parait judicieux de rappeler très brièvement les principales étapes de la biogenèse des ARNm cellulaires matures.

Les gènes de protéines sont transcrits, au sein de la chromatine activée par la modification des histones, par l'ARN polymérase II dont le domaine C-terminal (CTD) de la grande sous-unité sert de plateforme aux différents complexes impliqués dans le coiffage, l'épissage, la polyadénylation et l'export nucléaire des ARNm. Parmi ces complexes, TREX-1 et TREX-2 jouent un rôle fondamental dans le couplage entre la transcription/maturation de l'ARNm et son export.

- La coiffe (7mGpppG) est formée à l'extrémité 5' du pré-ARNm dès que les premiers nucléotides émergent de l'ARN polymérase puis elle est associée au complexe CBC – *cap binding complex* –, qui protège le pré-ARNm naissant de la dégradation. Ce complexe est également impliqué dans les autres étapes de la maturation ainsi que dans l'export nucléaire du mRNP (Sen et al., 2019).
- L'épissage du pré-ARNm débute dès qu'apparaît le premier site d'épissage, suite à l'assemblage du *spliceosome* et grâce à l'intervention de plus d'une centaine de protéines notamment les protéines SR. L'achèvement de chaque événement d'épissage est marqué par le dépôt d'un complexe protéique appelé EJC – *exon junction complex* – en amont des jonctions exon-exon. Ce complexe et certaines protéines impliquées dans l'épissage, contribuent à rendre le mRNP compétent pour l'export (Woodward et al., 2017).

• La polyadénylation de l'ARNm est catalysée par des complexes appelés CPSF – *cleavage and polyadenylation specificity factor* –, chez l'Homme, et CPF chez la levure. L'extrémité 3' non traduite est clivée à un site bien défini, puis une queue polyA est ajoutée par la polyA polymérase (Casañal et al., 2017; Stewart, 2019).

Suite à ces évènements, les ARNm matures se retrouvent compactés et empaquetés sous la forme de mRNP comprenant diverses protéines dont des hnRNP, des protéines SR et des protéines adaptatrices associées aux transporteurs impliqués dans leur transit nucléo-cytoplasmique. Après un dernier contrôle qualité des mRNP, lors de leur ancrage aux pores nucléaires, pour détecter les ARNm défectueux et un remodelage, les mRNP franchissent les NPC puis ils sont disloqués dans le cytoplasme pour libérer l'ARNm mature (**figure I-4**).

2.1. La voie canonique d'export nucléaire des ARNm

Comme nous l'avons vu, les classes d'ARN très structurés tels que les ARNt et miARN, présentent chacune des motifs d'identité communs qui sont reconnus par une exportine spécifique.

A l'inverse, les ARNm cellulaires constituent une population d'ARN très hétérogène d'un point de vue longueur, séquence et structure secondaire et ils ne renferment pas d'éléments spécifiques directement reconnaissables par les exportines. Pour cette raison, ces ARN sont exportés grâce à une stratégie qui consiste à recruter au cours de leur transcription et leur maturation, des protéines liant l'ARN et jouant le rôle d'adaptateurs reconnus par les récepteurs d'export. D'autres protéines dont des hnRNP se fixeront également sur l'ARNm pour contribuer à son remodelage. Cette stratégie qui conduit à la formation de mRNP compétents pour l'export, est en fait coordonnée par le complexe TREX-1 qui est conservé de la levure à l'Homme en passant par la drosophile et les plantes, reflet de son importance physiologique (**figure I-4**)(Strässer et al., 2002; Rehwinkel et al., 2004; Masuda et al., 2005; Ehrnsberger et al., 2019a).

Avant de détailler cette stratégie, voyons d'abord la nature des récepteurs de l'export et des protéines adaptatrices et, la composition du complexe TREX-1.

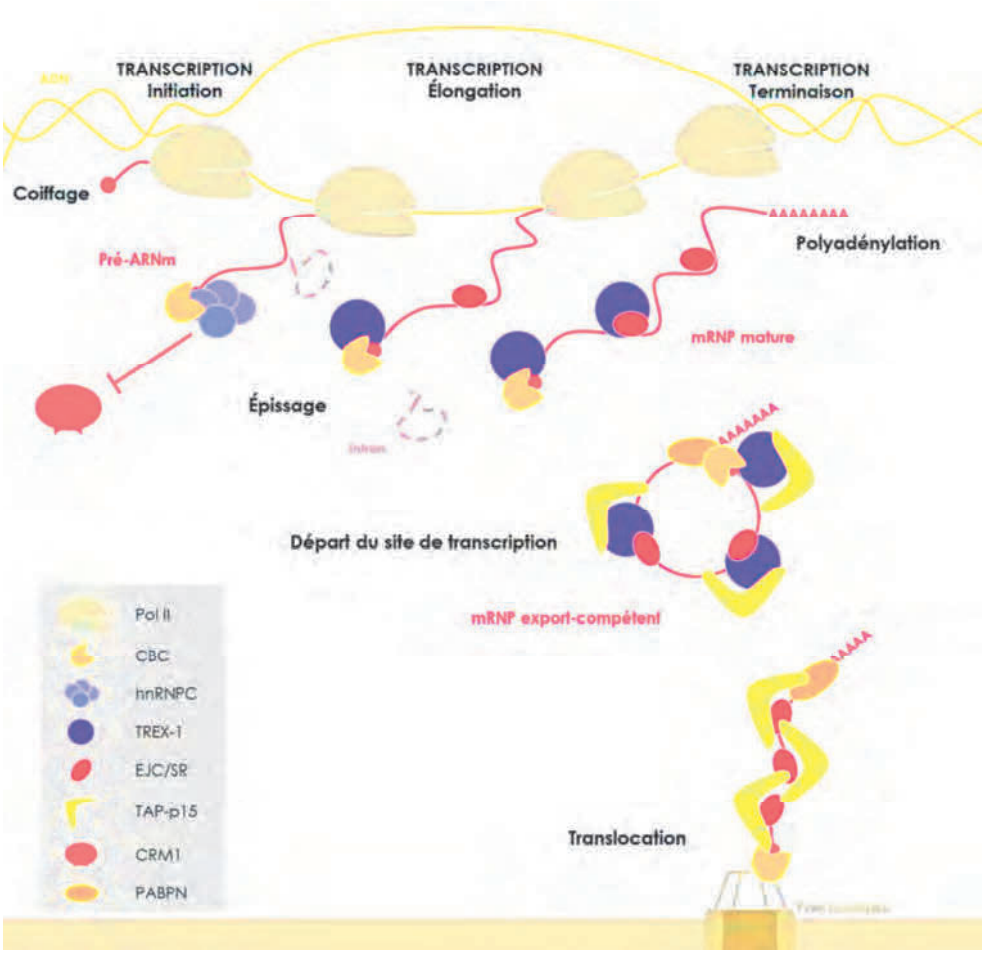


Figure I-4 | La transcription, la maturation et l'export nucléaire des ARNm sont des processus couplés. Pendant la transcription par l'ARN polymérase II (Pol II), les facteurs et les complexes requis pour le coiffage, l'épissage et la polyadénylation interagissent successivement avec le pré-ARNm, formant ainsi un mRNP. La coiffe ajoutée à l'extrémité 5' de l'ARN au début de la transcription est directement reconnue par le CBC. L'estimation de la taille du transcrit en cours de synthèse (>300nt) par la hnRNPC permet de détourner l'ARN de la voie CRM1 et l'orienter vers la voie de TAP-p15. Subséquemment, des protéines du complexe TREX-1 par interaction avec le CBC et les facteurs d'épissage (EJC/SR) sont recrutés sur le messager. Durant sa biogenèse, la structure et la composition protéique du mRNP changent ce qui permet le recrutement des récepteurs TAP-p15 qui procèdent alors à l'export de l'ARNm mature au travers du pore nucléaire. D'après Katahira, 2015.

2.1.1. Les récepteurs de l'export nucléaire des ARNm

Les NTR responsables de l'export nucléaire des ARNm sont conservés de la levure à l'Homme mais demeurent inconnus chez les plantes. Il s'agit de l'hétérodimère TAP-p15 chez les métazoaires formé de l'exportine TAP – *tip associated protein* – et de son co-facteur p15 (encore appelées NXF1-NXT1) et Mex67/Mtr2 chez la levure (Segref et al., 1997; Grüter et al., 1998; Herold et al., 2000). Ces deux récepteurs n'appartiennent pas à la famille des karyophérines β et fonctionnent indépendamment du gradient de Ran-GTP mais requièrent à la place de l'ATP, pour l'orientation et le remodelage des mRNP, au cours de leur transit nucléo-cytoplasmique (Stewart, 2007).

L'exportine TAP possède 5 domaines différents : un domaine N-terminal RBD – *arginine-rich RNA-binding domain* –, un domaine RRM – *RNA recognition motif* –, un domaine LRR – *leucine rich repeat* –, un domaine *NTF2-Like*, qui lui permet de se lier à p15 et enfin, un domaine C-terminal UBA – *ubiquitin associated* – (Viphakone et al., 2012). Ses domaines UBA et *NTF2-like* lui permettent d'interagir avec les motifs Phé-Gly répétés des Nups-FG lors du franchissement du NPC et les domaines RBD, RRM et LRR, de se lier à l'ARNm (Valkov et al., 2012; Bonnet et Palancade, 2014). L'hétérodimérisation de TAP avec son co-facteur p15 est essentielle pour maintenir le domaine *NTF2-like* de TAP dans une structure fonctionnelle (Fribourg et Conti, 2003). TAP-p15 s'associe aux ARN avec peu d'affinité, malgré la présence de domaines RBD et RRM, et de ce fait, requiert l'assistance de protéines dites adaptatrices pour qu'il puisse interagir fortement avec ses cargos et promouvoir leur export.

Un autre récepteur intervient également dans le transport nucléo-cytoplasmique de certaines populations d'ARNm, la protéine CRM1 qui, rappelons-le, a pour fonction principale d'exporter les protéines, les complexes ribosomiques et les snARN.

2.1.2. Les protéines adaptatrices des récepteurs d'export nucléaire

La protéine de liaison aux ARN, ALY/REF - *ally of AML-1 and LEF-1* - (Yar1 chez la levure) est la première des protéines adaptatrices impliquées dans le recrutement des exportines TAP-p15 sur les ARNm à avoir été identifiée. Ses extrémités N- et C-terminales sont responsables *via* des motifs UBM – *UAP56-binding domain* – de son interaction avec l'ARN hélicase UAP56 - *U2AF65 associated protein* -, alors que sa partie centrale contient un motif RRM – *RNA recognition motif* – flanqué de régions riches en arginines qui sont impliquées dans la liaison d'ALY/REF aux ARN. Ces dernières permettent à la protéine ALY/REF de faire le pont entre les ARNm et l'hétérodimère TAP-p15 et subséquemment, de participer à leur compaction. Ces régions riches en arginines constituent aussi les sites principaux d'interaction avec TAP-p15 ; leurs interactions avec ces récepteurs induiraient le transfert de l'ARN de ALY/REF sur TAP-p15. (Hautbergue et al., 2009; Stutz et al., 2000). La protéine ALY/REF, comme l'ARN hélicase UAP56 quittent le mRNP avant ou au cours de sa translocation nucléocytoplasmique (Lund et al., 2004).

ALY/REF est impliquée dans l'export nucléaire à la fois des ARNm épissés et des ARNm dépourvus d'introns mais le nombre de messagers concernés chez l'Homme est moins important qu'attendu (Smyk et al., 2006), contrairement à la levure où la protéine Yra1, orthologue de ALY/REF, est engagée dans l'export de la très grande majorité des ARNm (Kurshakova et al., 2016). Ce contraste s'explique par la redondance fonctionnelle chez les métazoaires, entre ALY/REF et d'autres protéines, telles que UIF - *UAP56-interacting factor* -, LUZP4 - *leucine zipper protein 4* - ou hTHOC5 du complexe TREX-1 et les protéines SR – *serine arginine-rich* - impliquées dans l'épissage (Gatfield et Izaurralde, 2002; Viphakone et al., 2012; Heath et al., 2016). Ainsi, un *knockdown* simultané de ALY/REF et UIF conduit à un blocage très important de l'export des

ARNm chez les mammifères (Hautbergue et al., 2009). Enfin, signalons le cas de la protéine HuD qui sert d'adaptateur pour l'exportine TAP-p15, dans les cellules neuronales où elle est exprimée exclusivement (Saito et al., 2004)

2.1.3. Le complexe TREX-1

Le complexe TREX-1 joue le rôle de coordinateur entre la transcription/maturation et l'export nucléaire des mRNP (Cautain et al., 2015). Ses composants sont impliqués fonctionnellement dans chaque étape de la synthèse de l'ARNm mature comme le démontre son association physique avec les différents complexes impliqués dans ce processus (Heath et al., 2016)(**figure I-5**). Ce complexe permet de par son action, le dépôt de multiples protéines qui provoquent le repliement de l'ARN. Par ailleurs, il empêche au sein de la chromatine active, la formation d'hybrides entre le transcrit en élongation et le brin non-codant du gène – les *R-loops* – et éviter ainsi des recombinaisons transcription-dépendantes (Katahira, 2012). Chez la levure, TREX-1 est recruté sur les transcrits au cours de la phase d'élongation de la transcription (Strässer et al., 2002; Zenklusen et al., 2002) alors que chez l'Homme, il s'associe aux complexes EJC déposés sur les transcrits et ce, dès l'épissage du premier intron. Le complexe TREX-1 humain peut se lier directement à l'ARNm contrairement à celui de la levure (Huertas et al., 2006).

Le complexe TREX-1 (**tableau I-1**) est formé de plusieurs protéines dont la composition peut varier non seulement entre espèces mais également en fonction du type cellulaire (Viphakone et al., 2012; Katahira, 2015; Heath et al., 2016). Il comprend un complexe protéique appelé THO, résistant aux sels, auquel sont associés divers facteurs dont la protéine adaptatrice ALY/REF, anciennement appelé THOC4 (Yra1 chez la levure), l'ARN hélicase ATP dépendante UAP56 (Sub2), et CIP29 (Tho1) – *cytokine-induced protein* – (Sugiura et al., 2007; Dufu et al., 2010). La protéine CIP29 forme un complexe trimérique ATP-dépendant avec ALY/REF et UAP56 dont elle stimule les activités ATPase et hélicase (Sugiura et al., 2007; Dufu et al., 2010). Chez les mammifères, Sub2 existe sous deux formes, UAP56 et URH49, également appelées respectivement DDX39B/A. On y trouve aussi, associées au complexe THO, d'autres protéines telles que hCHOP, UIF et LUZP4.

Le complexe THO est composé chez les métazoaires de THOC1, THOC2, THOC3, THOC5, THOC6 et THOC7 et chez la levure de Hpr1, Tho2, Thp2, Mft1 et Tex1 (Katahira, 2012; Heath et al., 2016). THOC2 joue un rôle d'échafaudage et de ce fait, est essentielle pour la formation et la fonction du complexe TREX-1.

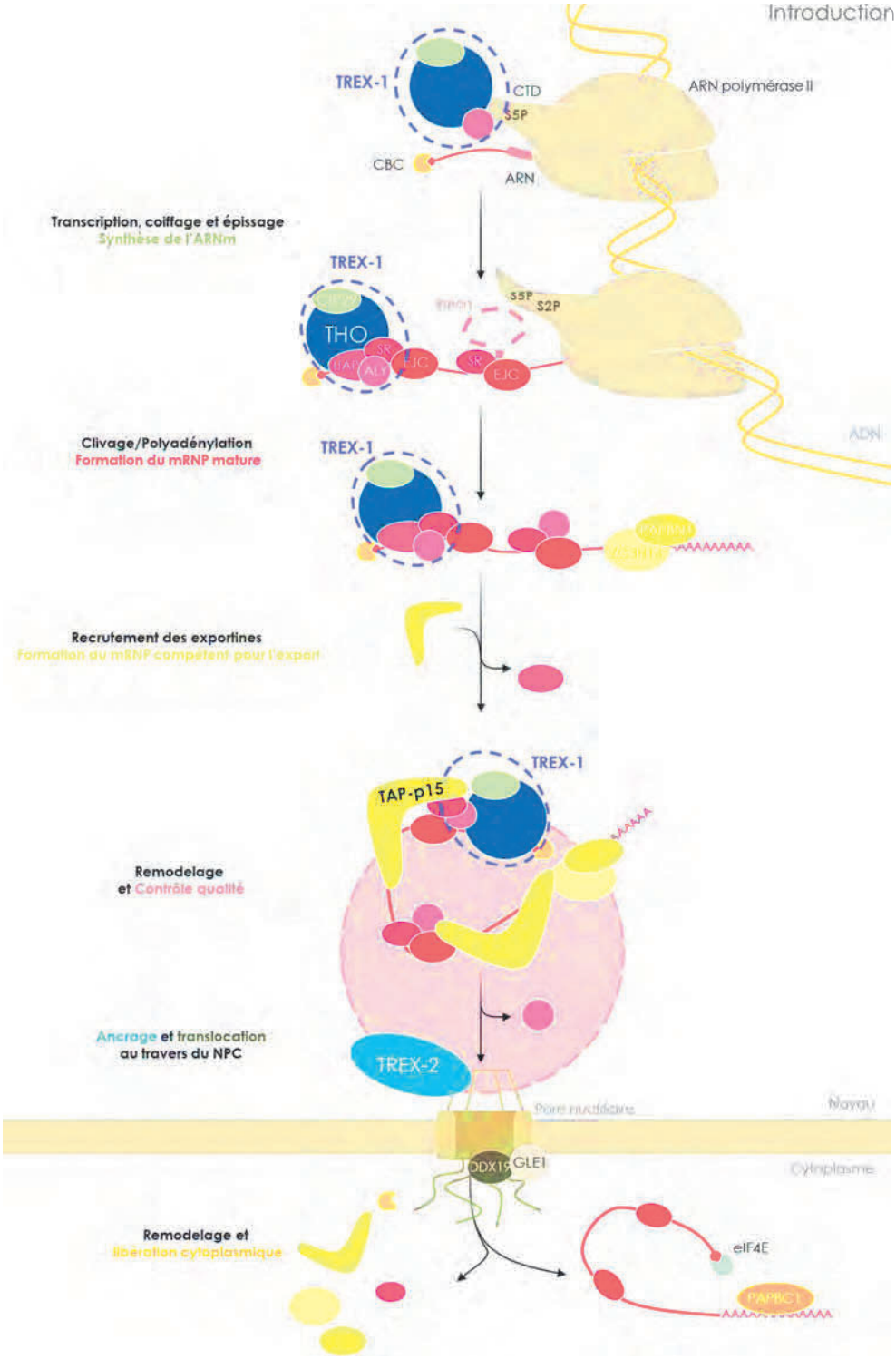


Figure I-5 | Aperçu de la formation et de l'export nucléaire des mRNP dépendant de la voie de TREX-1/TAP-p15. Lorsque les gènes de protéines sont exprimés, le complexe TREX-1 est recruté co-transcriptionnellement en interagissant avec le domaine CTD phosphorylé de l'ARN polymérase II. Chaque étape de maturation du pré-ARNm (coiffage, épissage et polyadénylation) permet l'assemblage de TREX-1 (THO, CIP29, UAP56 et ALY) et le dépôt de nombreux facteurs impliqués dans la formation du mRNP mature (CBC, EJC et protéines SR, PABPN1 ou encore ZC3H14). ALY, grâce à UAP56 et l'hydrolyse de l'ATP, permet alors le recrutement des récepteurs d'export TAP-p15. Aux abords du NPC, l'ARNm est remodelé et ancré *via* le complexe TREX-2 au pore nucléaire où un dernier contrôle qualité est effectué avant son passage au travers de l'enveloppe nucléaire. Du coté cytoplasmique, les facteurs d'export sont relâchés par l'action conjointe de Dbp5 et Gle1 qui libèrent l'ARNm dans le cytosol. D'après Heath et al., 2016.

Complexes	Nom Humain	Noms alternatifs	Orthologues de levure	Fonctions cellulaires
THO	THOC1	hHPR1	Hpr1	Régulateur dans l'apoptose
	THOC2		Tho2	Protéine d'échafaudage
	THOC3	TEX1	Tex1	
	THOC5			Co-adaptateurs d'export
	THOC6			
	THOC7			
TREX-1	UAP56	DDX39B	Sub2	Facteur d'épissage, associé aux EJC
	URH49	DDX39A		Paralogue d'UAP56
	ALY/REF	ALY, REF, THOC4	Yra1	Adaptateur, associé aux EJC
	UIF			Adaptateur, associé aux EJC
	LUZP4			Adaptateur, associé aux EJC
	CHTOP			Adaptateur, associé aux EJC
	CIP29	SARNP	Tho1	Co-adaptateur d'export
Associées	TAP	NXF1	Mex67	NTR
	p15	NXT1	Mtr2	NTR, requis pour stabiliser TAP
	NCBP1	CBP80	Cbp80, Cbc1	CBC : élongation, export et stabilité du transcrit
	NCBP3			Export de transcrit en cas d'infection virale
	ZC3H18			NEXT
	ARS2			Maturation des pri-miARN et des ARNm en 3'

Tableau I-1 | Composition du complexe TREX-1. D'après Heath et al., 2016.

2.1.4. Couplage entre la biosynthèse et l'export nucléaire des ARNm

Le recrutement du complexe TREX-1 sur les pré-ARNm naissants, est dépendant du domaine C-terminal (CTD) de la grande sous-unité de l'ARN polymérase II et plus particulièrement, de l'heptapeptide très conservé et répété $Y_1S_2P_3T_4S_5P_6S_7$. En effet, cette séquence répétée, de 26 fois chez la levure jusqu'à 52 fois chez l'Homme, sert de plateforme flexible pour l'arrimage du complexe TREX-1 et de nombreux facteurs et complexes impliqués dans le coiffage, l'épissage et la polyadénylation (Katahira, 2015). Leur association avec l'ARN polymérase II est déterminée par l'état de phosphorylation de l'heptapeptide (Bentley, 2005; Hsin et al., 2014). Ainsi, chez *S. cerevisiae*, THO interagit spécifiquement avec le CTD phosphorylé en positions S_2 et S_5 et comme sa phosphorylation devient plus intense lors de la phase d'élongation, il y a recrutement de plus en plus de complexes TREX-1 tout le long du pré-ARNm (**figure I-5**)(Meinel et al., 2013).

• **Coiffage du pré-ARNm.** Le complexe TREX-1 est mobilisé à l'extrémité 5' du transcrit émergeant de l'ARN polymérase II, dès la formation de la coiffe et l'association de celle-ci avec le complexe CBC. Ce dernier est un hétérodimère constitué de deux sous-unités, NCBP1 et NCBP2, respectivement Cbp80 et Cbp20 chez *S. cerevisiae* (Gonatopoulos-Pournatzis et Cowling, 2014). Cette première étape de la maturation du pré-ARNm est déterminante pour son orientation vers la voie TAP-p15 (Mex67/Mtr2) car dès que l'ARN dépasse une longueur de 300 nt, il s'enroule autour d'un tétramère formé des protéines hnRNPC1/C2 recrutées par le complexe CBC et empêche la protéine PHAX, adaptateur de l'exportine CRM1, de s'associer au CBC et de diriger le transcrit vers cette voie d'export (Daneholt, 2001; McCloskey et al., 2012).

Le recrutement de TREX-1 se fait par des interactions directes entre NCBP1 et ALY/REF et le complexe THO (**figure I-5**). Il s'avère que TREX-1 interagit également avec la protéine NCBP3 qui se substitue à NCBP2, préférentiellement dans des conditions de stress, par exemple une infection virale. L'implication de NCBP3 dans l'export nucléaire de ARNm est démontrée par l'accumulation nucléaire de ces messagers dans des cellules *knockdown* pour cette protéine (Gebhardt et al., 2015). Il est probable que TREX-1 régule aussi l'association entre le CBC et le complexe NEXT – *nuclear exosome targeting* – impliqué dans le recrutement de l'exosome ce qui suggère que l'assemblage de TREX-1 sur la coiffe constitue à ce stade une étape décisionnaire concernant la poursuite de la synthèse des ARNm ou leur dégradation.

• **Épissage du pré-ARNm.** Le recrutement de TREX-1 est également dépendant de l'épissage du pré-ARNm chez l'Homme, ceci étant corroboré par le fait que l'ARN hélicase UAP56, associée à THO, est aussi impliquée, en combinaison avec le facteur d'épissage U2AF2 dans l'assemblage du spliceosome (Fleckner et al., 1997).

Ainsi, en plus d'être recruté au niveau de l'extrémité 5' proximale grâce à la présence de la coiffe et du CBC, chez les métazoaires, TREX-1 s'associe au complexe EJC qui est déposé à 24 nt en amont de la première jonction exon-exon (**figure I-5**)(Cheng et al., 2006). En effet, des composants de TREX-1 sont co-purifiées avec le complexe EJC et il a été montré que ALY/REF possède un motif (WxHD) qui facilite son interaction avec la protéine eIF4AIII, constitutive du complexe EJC (Gromadzka et al., 2016). Les complexes CBC et EJC servent donc tous les deux de plateformes pour le recrutement de TREX-1 et plus particulièrement celui de l'adaptateur ALY/REF qui à son tour, incorporera les exportines TAP-p15 dans le mRNP. Les complexes TREX-1 sont-ils aussi déposés à l'intérieur des ARNm chez les métazoaires, au niveau des différents EJC ? Cette question n'est pas encore résolue mais il ne semble pas illogique que ce soit le cas vu le nombre important d'introns présents dans les pré-ARNm. Par ailleurs, le recrutement de multiples complexes TREX-1 sur l'ARNm est également observé chez *S. cerevisiae* mais l'incorporation de ces complexes n'est pas liée à l'épissage, les gènes de protéines renfermant rarement des introns chez cette levure, mais à l'élongation de la transcription.

Les protéines SR connues pour leur rôle dans l'épissage des pré-ARNm (Huang et Steitz, 2005), sont également impliquées dans l'export nucléaire des mRNP où elles servent d'adaptateurs pour TAP-p15 (Huang et al., 2003). Après épissage, plusieurs d'entre elles - SRP20 et 9G8 chez l'Homme et, Gbp2 et Hrb1 chez la levure - restent sur l'ARN et recrutent les exportines qui se lient directement à leurs motifs SR à condition que ces derniers soient amenés à un état hypophosphorylé après l'épissage. Après export du mRNP, les protéines SR sont réimportées dans le noyau où elles sont hyperphosphorylées, ce qui leur permet d'être à nouveau recrutées par la machinerie d'épissage (**figure I-5**)(Hurt et al., 2004; Cáceres et al., 1998). Ainsi, la protéine SR NPL3 recrutée au niveau du transcrit *via* l'ARN polymérase II, est hyperphosphorylée progressivement par la caséine kinase CK2 pendant la phase d'élongation,

puis après épissage, elle est déphosphorylée par la phosphatase GCL7 lui permettant de demeurer sur l'ARNm et de recruter les exportines TAP-p15. Après l'export, NPL3 est de nouveau phosphorylée par la protéine kinase SKY1 ce qui conduit à libérer l'ARNm et le transporteur et, au ré-import de la protéine NPL3 (Gilbert et Guthrie, 2004). En d'autres mots, le cycle de phosphorylation et de déphosphorylation des protéines SR, entre noyau et cytoplasme, oriente leur fonction et la directionnalité de l'export. Signalons aussi le cas de la protéine SR SRSF3 qui se fixe sur le dernier exon d'ARNm ayant une région 3' UTR – *untranslated region* – particulièrement longue et qui promeut leur export (Heath et al., 2016).

• **Polyadénylation de l'extrémité 3' du pré-ARNm.** La synthèse de la queue polyA en 3' des transcrits constitue la dernière étape de maturation des ARNm et une nouvelle opportunité pour recruter l'adaptateur ALY/REF chez l'Homme et Yar1, chez la levure, selon des mécanismes similaires. Cette étape, indispensable pour libérer l'ARNm de son gène, est assurée respectivement par les complexes CPSF et CPF (Casañal et al., 2017).

Le clivage dans la région 3' UTR est réalisée par les protéines, PCF11 et CLP1. PCF11 est associé, tôt au cours de la transcription, au CTD phosphorylé de l'ARN polymérase II et recrute ALY/REF à l'extrémité 3' du gène. Toutefois, l'interaction subséquente de CLP1 avec PCF11 provoque la dissociation d'ALY/REF, un processus qui est stimulé par l'ARN hélicase UAP56. Suite au clivage de l'extrémité 3', la poly A polymérase (PAP1) catalyse l'addition de résidus A. La longueur de la queue polyA, est régulée par Nab2 chez *la* levure et par les facteurs PABPN1 et ZC3H14, chez l'Homme. Ces protéines induisent le départ de PAP1 en s'agrégeant sur la queue poly A en formation (Kühn et al., 2017). Le relargage de PAP1 induit un changement conformationnel au sein des complexes CPF/CPSF qui est transmis au complexe TREX-1 sans doute *via* le facteur PCF11 (Johnson et al., 2009b). L'adaptateur ALY/REF est alors sumoylé ce qui activerait UAP56 et son départ du complexe permettant *in fine* à ALY/REF de recruter les exportines TAP-p15 et rendant le mRNP encore plus compétent pour l'export.

L'implication du complexe TREX-1 dans l'étape de polyadenylation des ARNm est démontrée par l'interaction de THOC5 avec des facteurs du complexe CPSF (Katahira et al., 2013) et l'étude de mutants Tho et Sub2 chez la levure. Ces mutants provoquent un arrêt de la formation des mRNP et la rétention de facteurs de polyadénylation au niveau de la chromatine – on parle de chromatine lourde - (Rougemaille et al., 2008). Il apparaît donc clairement que TREX-1 joue un rôle clé dans la maturation en 3' et le relargage concomitant des mRNP de la chromatine.

2.1.5. Ancrage des mRNP au pore nucléaire et translocation

Les mRNP ainsi formés comprennent de nombreuses protéines dont plusieurs exportines TAP-p15 recrutées grâce aux adaptateurs, ce qui facilitera leur translocation à travers le canal intérieur du pore nucléaire.

Les mRNP relargués de leurs gènes se déplacent par diffusion et de manière aléatoire à travers l'espace interchromatinien jusqu'au pore nucléaire. Le trajet des mRNP peut durer quelques minutes voire presque une heure, en fonction de la taille du mRNP, de l'encombrement intranucléaire et de la position du gène de protéine correspondant au sein de la chromatine, par rapport aux NPC (Björk et Wieslander, 2017). Une fois arrivés à la périphérie nucléaire, ces mRNP matures s'accumulent à la face nucléaire des pores où ils sont ancrés au panier nucléaire par le complexe TREX-2 (**figure I-5**)(Grünwald et Singer, 2010; Wickramasinghe et al., 2010).

TREX-2 est un complexe multi-protéique conservé de la levure à l'Homme en passant par la drosophile où il prend le nom d'AMEX – *anchoring and mRNA export* – (Kurshakova et al., 2007; Lu et al., 2010). Outre son rôle dans l'export, il prend part à d'autres mécanismes comme la transcription, la réplication et la stabilité du génome (Jani et al., 2014). Il est composé chez l'Homme, des protéines GANP, PCID2, ENY2, CETN2/3 et DSS1 et respectivement chez la levure, des protéines Sac3, Thp1, Sus1, Cdc31 et Sem1 (Jani et al., 2009; Schubert et Köhler, 2016). Chez Arabidopsis, la protéine ENY2 (Sus1) n'entre pas dans la composition du complexe TREX-2 (voir **§3**). La protéine GANP (Sac3) est la clé de voûte du complexe TREX-2 car elle organise les interactions entre les différentes protéines (Kurshakova et al., 2016).

TREX-2 interagit grâce à Sus1 (ENY2) avec le complexe SAGA – *Spt-Ada-Gcn5-Acetyltransférase* - qui joue un rôle central dans l'acétylation et la déubiquitylation des histones de la chromatine et donc, dans l'activation de la transcription (García-Oliver et al., 2012). En fait, la protéine Sus1 est également constitutive de ce deuxième complexe. Les études réalisées sur Sus1 montrent qu'elle a un rôle important dans le couplage entre la transcription et l'export des mRNP. Elle intervient à différents stades de la formation des mRNP dont l'élongation de la transcription voire chez la drosophile, dans la maturation de l'extrémité 3' des ARNm (Huertas et al., 2006). Son implication dans l'élongation de la transcription est démontrée entre autres, par sa co-purification avec l'ARN polymérase II phosphorylée en S_2 et S_5 du CTD. Quant à son rôle dans l'export, Sus1 interagit avec l'exportine Mex67/Mtr2 et son adaptateur Yra1 et elle cible le complexe TREX-2 vers les NPC (García-Oliver et al., 2012; Valkov et al., 2012). La protéine Sus1 permet aussi de retenir, avec la protéine Thp1 du complexe TREX-1, le mRNP au niveau de la chromatine, en agissant avec les facteurs impliqués dans le contrôle qualité du mRNP (Pascual-García et Rodríguez-Navarro, 2009). L'association des complexes TREX-2 et SAGA permet de positionner les gènes de protéines activés à proximité des NPC de sorte que les mRNP puissent y accéder rapidement, dès leur détachement de la chromatine. Ce processus corrobore la théorie dite du « *gene gating* » (Blobel, 1985) suivant laquelle tout gène est lié physiquement au pore nucléaire.

Le complexe TREX-2 est impliqué grâce notamment aux protéines ENY2 et GANP, dans l'ancrage du mRNP au NPC : il interagit avec les nucléoporines NUP153 et TRP (Nup1 et Mlp1 chez la levure) du panier nucléaire qui s'ouvre et permet aux mRNP d'entrer à l'intérieur puis d'accéder au canal central du NPC. Les mRNP interagissent grâce à TAP-p15 avec NUP98, une Nup-FG située à l'entrée du pore nucléaire, puis avec d'autres Nups-FG dont NUP62, le long du canal interne. La translocation des mRNP au travers du canal central est un processus extrêmement rapide, estimé à 20 ms dans le cas de l'ARNm de l'actine (Björk et Wieslander, 2017).

Après translocation, les mRNP sont maintenus près du canal central du NPC par le biais d'interactions entre leurs NTR et les Nups-FG, pour subir un remodelage grâce à DDX19 (Dbp5 chez la levure), une ARN hélicase ATP-dépendante ; ce remodelage empêche le ré-import des mRNP (**figure I-5**)(Fernandez-Martinez et al., 2016). DDX19 est localisée avantageusement du côté cytoplasmique grâce à son association avec NUP214 (Nup159 chez la levure) des filaments du NPC, à proximité du canal central. DDX19 est en réalité, déjà recrutée par les mRNP au cours de leur transcription mais elle n'agit pas à ce stade car son activité hélicase est trop faible. Cette dernière est stimulée après export, par la protéine GLE1 qui est localisée sur les filaments du pore nucléaire. Chez la levure, cette stimulation est optimisée en présence de la molécule de signalisation inositol hexakisphosphate (IP6)(Alcázar-Román et al., 2006; Lin et al., 2018). DDX19 permet donc de démanteler les mRNP, libérant l'ARNm mature des facteurs d'export et de la plupart des protéines associées aux messagers (**figure I-5**). Le facteur RanBP2 (Nup358) associé aux filaments permet ensuite le recyclage des facteurs d'export vers le noyau (Bernad et al., 2004). Les ribosomes débarrasseront les ARNm du reste des protéines/complexes du mRNP, lors de la traduction.

2.1.6. Contrôle qualité des ARNm

La complexité de la transcription/maturation des ARNm rend les erreurs inévitables au cours de la biogenèse des ARNm. La production et l'export nucléaire des ARNm défectifs ou aberrants peuvent avoir des effets délétères sur le métabolisme cellulaire et *in extenso* sur la physiologie de l'organisme. Les transcrits aberrants peuvent également saturer les machineries impliquées dans la maturation des transcrits et interférer avec la production d'ARNm matures ou s'apparier à leur gène et former des hybrides ARN:ADN qui induiraient des doubles cassures au niveau de l'ADN.

Le contrôle de la qualité des ARNm dans le noyau, au cours de leur synthèse et avant de franchir le pore nucléaire, est donc indispensable pour éviter aux ARNm aberrants d'accéder au cytoplasme. Ce système de contrôle permanent des ARN, à la recherche d'erreurs possibles – on parle aussi de *surveillance* des ARN – dépend essentiellement chez des organismes aussi

divers que la levure, les plantes et l'Homme, de l'exosome, un complexe enzymatique nucléaire portant une activité endonucléase et 3' exonucléase et qui agit sur tous les ARN cellulaires.

• **L'exosome.** Il est impliqué à la fois dans la dégradation nécessaire de l'extrémité 3' des ARN pour leur maturation et dans celle des transcrits aberrants comme des pré-ARNm présentant un épissage inadéquat, une polyadenylation trop courte ou un empaquetage incorrect, en d'autres termes, des transcrits dont les défauts les rendent incompétents pour l'export et leur expression.

L'activité de l'exosome dépend de différents co-facteurs qui ont pour fonction de l'aiguiller sur les transcrits et de permettre à l'exosome de faire la distinction entre la maturation et la dégradation. En effet, plusieurs de ces facteurs sont impliqués dans la maturation des pré-ARN ce qui suggère que l'exosome est déjà recruté sur les transcrits naissants pour pouvoir agir dès qu'une erreur est détectée lors de la transcription et/ou de la maturation. Ces modulateurs sont chez la levure, les complexes TRAMP (Trf4/5, Air1/2 et Mtr4 *polydenylaton complex*) et NNS (Nrd1, Nab3 et Sen1) qui interviennent respectivement dans la phase d'élongation et de terminaison de la transcription (Bresson et Tollervey, 2018; Hilleren et al., 2001; Singh et al., 2018). Grâce à ces complexes, les transcrits considérés comme incorrects sont dégradés et éliminés du pool des ARNm au sein du noyau.

• **Le complexe Rai1/Rat1.** Les transcrits qui n'ont pas été coiffés ou dont la coiffe est défectueuse comme l'absence de méthylation (Kiledjian et al., 2012) sont quant à eux, dégradés par l'hétérodimère Rai1/Rat1 chez la levure et DOM3Z/XRN2 chez l'Homme. L'enzyme Rai1 enlève la coiffe non méthylée grâce à son activité pyrophosphorylase générant une extrémité 5' phosphate tandis que la 5' exonucléase Rat1 dégrade l'ARN à partir de son extrémité 5' (Jiao et al., 2010; Tutucci et Stutz, 2011).

• **Les Nups Mlp1 et Mlp2.** Un dernier contrôle qualité des ARNm est effectué au niveau des pores nucléaires avant de les laisser franchir le canal central. Ce système de surveillance implique deux nucléoporines du panier nucléaire, Mlp1 (Trp chez l'Homme) et Mlp2 qui jouent le rôle de « vigiles » du NPC sans être pour autant associées à une machinerie de dégradation (Sommer et Nehrbass, 2005). Leur rôle est double, empêcher l'export des ARNm considérés comme impropres en les retenant dans le noyau et accélérer l'export des ARNm matures en permettant leur ancrage aux NPC (Bonnet et al., 2015).

Les Nups Mlp1 et 2 interagissent avec le facteur Nab2 (PABPN1 chez l'Homme) constitutif des mRNP matures et compactés ce qui leur permet de recruter directement les mRNP à l'entrée du pore nucléaire. La Nups Mlp2 – son homologue n'a pas été trouvé chez les métazoaires - interagit également avec l'exportine Yra1 et aide ainsi à l'ancrage des mRNP au NPC (Vinciguerra et al., 2005). Ces interactions garantissent le repérage des ARNm présentant une queue polyA aberrante, non compacte et qui n'est donc pas associée à TAP-p15

(Vinciguerra et al., 2005). Mlp1 est aussi capable de reconnaitre les ARNm qui présentent un défaut d'épissage ou retenant des introns, en interagissant avec les facteurs d'épissage SF1 et BBP ; ces dernières ont pour fonction de faciliter la reconnaissance des séquences à épisser par le spliceosome (Galy et al., 2004). Ainsi, si ces deux facteurs sont présents, Mlp1 retient l'ARN incorrectement épissé dans le noyau. Plus généralement, les protéines SR associés aux ARNm déterminent si ces derniers sont retenus dans le noyau et voués à la dégradation ou s'ils sont autorisés à franchir le canal interne.

• **Le NMD.** Dans le cytoplasme, les ARNm sont soumis à une dernière inspection grâce au système de surveillance appelé *Nonsense-mediated decay* ou NMD qui permet de détecter les transcrits présentant des codons STOP prématurés ou PTC – *premature termination codons* – (Fasken et Corbett, 2005; Lejeune, 2017).

2.2. Les voies alternatives d'export nucléaire des ARNm

Les ARNm cellulaires sont dans leur grande majorité exportés efficacement par la voie de l'exportine TAP-p15 suivant le mécanisme décrit précédemment. Toutefois, certains ARNm qui ne sont pas soumis à l'épissage ou qui ne possèdent pas toutes les structures communes aux ARNm cellulaires, requièrent des facteurs spécifiques et supplémentaires pour leur export. Les ARNm qui ne répondent pas aux critères de sélection de la voie TAP-p15 ou qui sont exprimés en conditions de stress ou dans certains tissus spécifiques, sont exportés quant à eux par le transporteur CRM1. Cette voie d'export sera d'ailleurs exploitée par certains rétrovirus pour amener leurs ARN épissés et non épissés dans le cytoplasme.

2.2.1. Export des ARNm par la voie ALREX

Les ARNm issus de gènes dépourvus d'introns, ne peuvent pas emprunter la voie classique d'export des ARNm car celle-ci est intimement liée à la machinerie d'épissage. Ils utilisent pour leur export, une autre voie, appelée ALREX – *alternative RNA export* – (**figure I-6 b.**)(Müller-McNicoll et Neugebauer, 2013). Cette dernière requiert aussi l'intervention de protéines du complexe TREX-1 telles que ALY/REF et UAP56 mais leur recrutement sur ces ARNm se fait différemment et implique essentiellement le complexe CBC fixé à l'extrémité 5' coiffée des transcrits (Lei et al., 2011; Katahira, 2012). Il a été montré par exemple, que l'export de l'ARNm de la protéine HSP70 - *Heat-shock protein of 70 kDa* - codée par un gène dépourvu d'introns, est particulièrement dépendant de la protéine THOC5 qui sert de co-adapteur pour l'export. En effet, THOC5 interagit avec TAP-p15 sur un site non chevauchant avec celui de l'adaptateur ALY/REF, de sorte que cette exportine peut fixer ces deux protéines simultanément (Katahira et al., 2009).

2.2.2. Export nucléaire des ARNm des histones

Les histones sont codées par une famille de gènes dépourvus d'introns et dits réplication-dépendants car leur expression est restreinte à la phase S du cycle cellulaire ce qui leur permet de compacter l'ADN néosynthétisé. Les ARNm des histones, synthétisés par l'ARN polymérase II, sont coiffés mais ils ne possèdent pas de queue polyA à leur extrémité 3'. Cette dernière se termine à la place par une séquence de 26 nt très conservée, contenant une tige-boucle (SL) de 16 nt qui participe à leur maturation et ensuite à leur traduction (Marzluff et al., 2008; Sullivan et al., 2009). Cette structure est le site de liaison de la protéine SLBP– *stem-loop-binding protein* – qui interagit également avec le CBC associé à la coiffe ce qui permet de circulariser l'ARNm et de faciliter ainsi l'expression de la protéine. Huang et Steitz (2005) ont montré que la région codante de l'ARNm des histones constitue un site de liaison pour la protéine SR 9G8 et qu'elle promeut l'export nucléaire d'autres ARNm dans les oocytes de *Xenopus laevis* ce qui suggère que cette protéine intervient comme adaptateur pour TAP-p15 (**figure I-6 c.**). SLBP s'avère également toute aussi importante, du moins chez les mammifères, puisque son absence (*knockdown*) conduit à la rétention de ces messagers dans le noyau (**figure I-6 c.**)(Sullivan et al., 2009).

2.2.3. Export des ARNm cellulaires par la voie de CRM1

Certains ARNm n'empruntent pas la voie majeure de TAP-p15 mais la voie de CRM1/XPO1, comme par exemple le messager de l'interféron α en conditions de stress (Kimura et al., 2004).

L'export des ARNm par CRM1 requiert l'aide de protéines adaptatrices contenant un NES et pouvant s'associer à l'ARNm, CRM1 étant incapable d'interagir directement avec l'ARN (**figure I-6 d.**). Nous avons vu que les petits snARN, coiffés et polyadénylés étaient exportés par CRM1 grâce à la protéine adaptatrice PHAX associée au CBC et que cette stratégie est exclue pour les ARNm en raison de leur taille (Ohno, 2012) car dès que les transcrits dépassent une longueur de 300 nt, ils sont recouverts par des hétérotétramères hnRNPC1/C2 inhibant ainsi la fixation de PHAX sur le CBC (**figure I-4**) (McCloskey et al., 2012; Ohno, 2012). De ce fait, l'export des ARNm par CRM1 implique d'autres protéines adaptatrices, les protéines RBP - *RNA binding protein* – qui sont déjà liées à l'ARNm à des sites spécifiques, ou qui s'y fixent après leur interaction avec CRM1.

Parmi les ARNm exportés par CRM1, il y a des messagers de proto-oncogènes, de cytokines et de lymphokines qui se caractérisent par la présence de séquences riches en A et U (élément ARE) dans leur extrémité 3' UTR. Ces séquences sont reconnues par la protéine HuR – *Human antigen R* – (Gallouzi et Steitz, 2001; Wickramasinghe et Laskey, 2015). Cette protéine quasi ubiquitaire permet, en association avec ses ligands, les protéines PP32 (ANP32A) et APRIL

(ANP32B) – *acidic protein rich in leucin* –, le recrutement de l'exportine CRM1 sur ces ARNm (Fries et al., 2007). HuR contient un NES et un domaine de liaison aux éléments ARE alors que PP32 et APRIL possèdent un NES apparenté appelé *rev-like protein rich in leucines* (**figure I-6 c.**)(Brennan et al., 2000). Toutefois, il a été montré dans le cas spécifique l'interféron α humain que son messager délété de son extrémité 3' UTR était aussi exporté par l'exportine CRM1 ce qui indique que cette dernière peut se dispenser des éléments ARE et/ou de la protéine HuR (Kimura et al., 2004).

CRM1 joue également un rôle important dans l'export dit eIF4E-dépendant de certains ARNm codant des protéines impliquées entre autres dans la prolifération et la survie cellulaire et, la formation des métastases (Culjkovic et al., 2006; Wickramasinghe et Laskey, 2015). Le facteur d'initiation de la traduction eIF4E est présent en grandes quantités dans le noyau où il promeut l'export nucléaire d'ARNm qui se caractérisent par l'existence d'une structure secondaire de 50 nt spécifique au niveau de leur 3' UTR. Cet élément est appelé *eIF4E-sensitivity element* ou 4E-SE (Culjkovic et al., 2006). eIF4E reconnaît l'élément 4E-SE grâce à la protéine adaptatrice riche en leucines LRPPRC pour - *leucine rich pentatricopeptide repeat containing* - (Topisirovic et al., 2009). Cette dernière extrait, par compétition, le facteur eIF4E des corps PML (*promyelocytic leukemia*) et le relocalise sur les ARNm. LRPPRC interagit directement à la fois avec cet élément et eIF4E et, avec CRM1 permettant ainsi la translocation des ARNm à travers les NPC. Cette voie d'export nucléaire particulière implique des facteurs communs avec la voie classique Tap-p15, comme les ARN hélicases UAP56 et DDX3 et encore hnRNPA1. En revanche, le CBC n'intervient pas dans cette voie (Topisirovic et al., 2009).

Chez l'Homme, NXF3, une protéine apparentée à TAP, sert aussi de protéine adaptatrice à CRM1 pour l'export de certains ARNm, notamment dans les cellules de Sertoli des testicules où elle est exprimée préférentiellement. En effet, NXF3 interagit avec les ARNm mais elle ne possède pas, contrairement à TAP, le domaine C-terminal critique pour l'export et la liaison aux Nups. Par contre elle renferme une séquence riche en leucines, rappelant le NES de la protéine Rev du VIH-1 (Fischer et al., 1995) ce qui lui permet d'interagir avec CRM1 (Yang et al., 2001; Björk et Wieslander, 2017).

Un cas particulier est observé avec les pré-ARNm, non épissés, des protéines adaptatrices Yra1 de levure et ALY des métazoaires, ceux-ci étant acheminés vers le cytoplasme conjointement par les exportines Mex67 (TAP) et CRM1, pour y être dégradés. Chez la levure, lorsque Yra1 est présente en quantité suffisante dans le noyau, elle s'associe à son pré-ARNm durant la phase précoce de sa transcription et inhibe son épissage (Dong et al., 2007). Yra1 liée à son pré-ARNm recrute alors Mex67 qui interagit grâce à son NES avec CRM1 qui exporte alors l'ARN non épissé vers le cytosol pour sa dégradation. Ce mécanisme d'export atypique permettrait donc à l'adaptateur Yra1 d'autoréguler sa propre synthèse (Strässer et al., 2000; Kshirsagar et Parker, 2004; Dong et al., 2007).

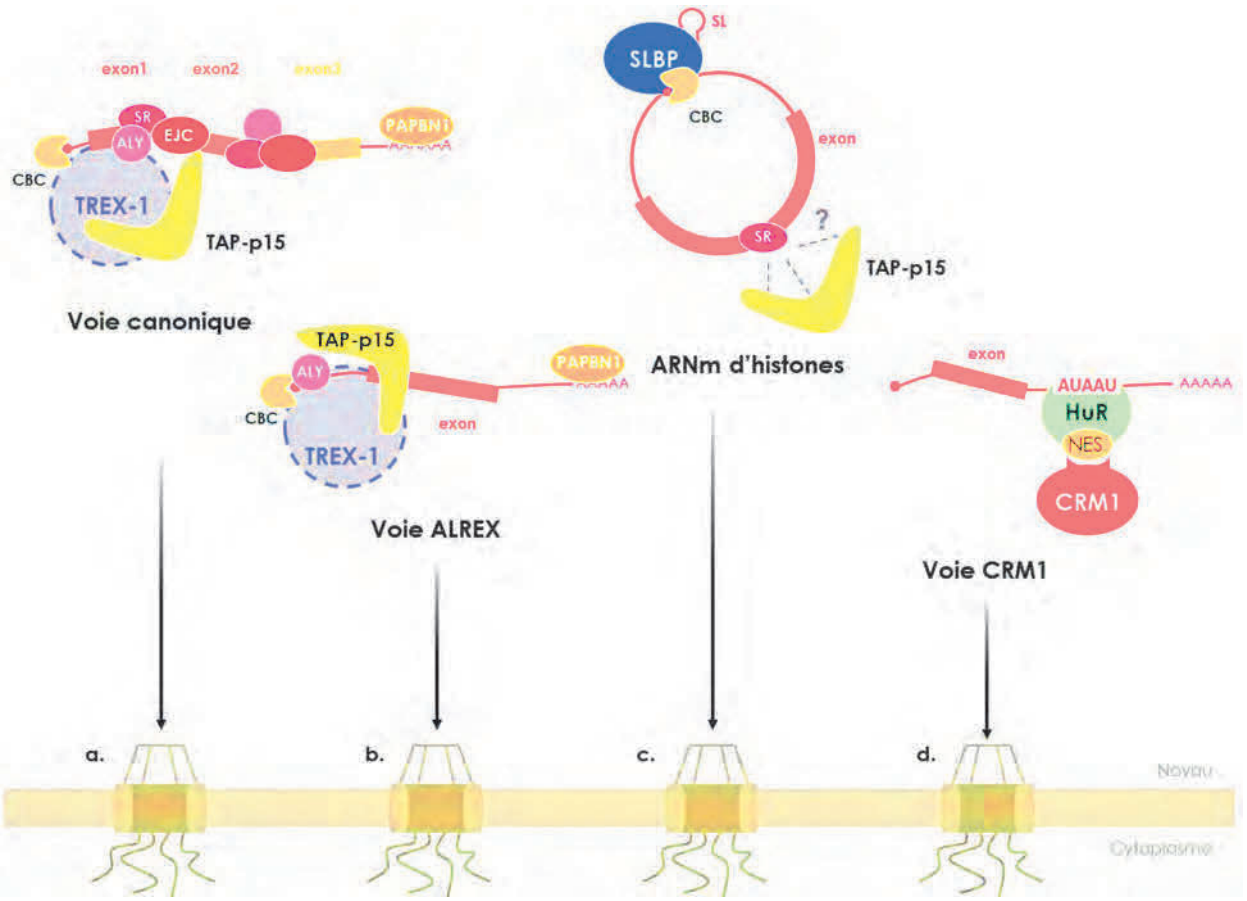


Figure I-6 | Les différentes voies d'export des ARNm cellulaires. Les ARNm matures et épissés sont normalement exportés par la voie canonique des exportines TAP-p15, recrutées sur l'ARNm grâce au complexe TREX-1 et des protéines adaptatrices (**a**). Les ARNm dépourvus d'introns ne peuvent pas utiliser les facteurs d'épissage pour recruter leurs exportines. Pour cela, ils utilisent la voie alternative d'export ALREX qui implique au moins les protéines ALY et THOC5 du complexe TREX-1 recrutées sur l'ARNm à son extrémité 5' grâce au CBC (**b**). Les messagers d'histones dépourvus aussi d'introns, utilisent pour leur export, la protéine SLBP qui est recrutée au niveau de leur extrémité 5' *via* le CBC (ces transcrits sont coiffés) et en 3' *via* une tige-boucle SL et, des protéines SR (**c**). Enfin, dans certaines conditions, notamment de stress, certains messagers sont exportés par la voie de l'exportine CRM1. Pour ce faire, ils utilisent des protéines adaptatrices possédant un NES reconnu par CRM1. C'est l'exemple des ARNm aux éléments ARE (AU–rich element), exportés grâce à la protéine adaptatrice HuR (**d**).

3. L'export nucléaire des ARNm chez les plantes

Contrairement à l'Homme et la levure, le mécanisme détaillé de l'export nucléaire des ARNm est méconnu chez les plantes, principalement du fait que les exportines impliquées dans ce transport n'ont pas encore été identifiées à ce jour (Merkle, 2011; Gaouar et Germain, 2013; Ehrnsberger et al., 2019a). Certes, le transporteur XPO1 qui exporte exceptionnellement certains ARNm chez les métazoaires et chez des virus des animaux, existe chez d'*Arabidopsis thaliana* mais son engagement dans ce type de transport n'est pas prouvé. Néanmoins, de réelles avancées ont été faites sur l'identification et la caractérisation des protéines des complexes TREX-1 et -2 et sur les nucléoporines d'Arabidopsis (Meier et al., 2017; Ehrnsberger et al., 2019a).

Malgré les difficultés d'interprétation de certains résultats obtenus avec les plantes d'Arabidopsis mutées car le complexe TREX-1 est aussi impliqué dans la biogenèse des petits ARN régulateurs (siARN) de l'expression des gènes, on peut esquisser en se référant aux connaissances acquises chez l'Homme et la levure, un schéma de l'export des mRNP chez les plantes.

3.1. Le complexe TREX-1 d'Arabidopsis thaliana

Le complexe TREX-1 d'Arabidopsis ressemble fortement à celui de l'Homme. Le complexe central THO est composé des protéines HPR11 (THOC1), THO2 (THOC2), TEX1 (THOC3), THO5 (THOC5), THO6 (THOC6) et THO7 (THOC7) (**tableau I-2**)(Yelina et al., 2010a; Ehrnsberger et al., 2019a). Les gènes THO5 et THO7 sont dupliqués de sorte qu'il existe deux isoformes, THO5A/5B et, THO7A/7B dont chacune d'entre elles, peut intégrer alternativement le complexe THO. Des protéines sont aussi associées à THO pour former TREX-1 : l'ARN hélicase ATP-dépendante UAP56 - il existe également deux isoformes UAP56A/B (respectivement UAP56/DDX39B et UHR49/DDX39A) -, l'adaptateur ALY pour lequel il existe quatre orthologues (ALY1 à 4) et MOS11, l'orthologue de CIP29 (Kammel et al., 2013; Sørensen et al., 2017). MOS11 fût d'abord associée au système de défense contre les pathogènes de plantes car elle est capable de supprimer le phénotype induit par le mutant *snc1* chez Arabidopsis, une propriété qui lui a d'ailleurs valu son nom - *modifier of snc1* - (Germain et al., 2010). La protéine UAP56 interagit avec MOS11 et ALY, ce qui indique qu'elle recrute ces deux protéines lors de l'assemblage du complexe TREX-1 (Kammel et al., 2013; Sørensen et al., 2017; Ehrnsberger et al., 2019a).

	Arabidopsis	Gènes	Homme	Levure
Complexe THO	HPR1(THO1)	AT5G09860	THOC1	Hpr1
	THO2	AT1G24706	THOC2	Tho2
	TEX1 (THO3)	AT5G56130	THOC3	Tex1
	THO5A	AT5G42920	THOC5	
	THO5B	AT1G45233		
	THO6 (DWA1)	AT2G19430	THOC6	
	THO7A	AT3G02950	THOC7	
	THO7B	AT5G16790		
Complexe TREX-1	UAP56A	AT5G11170	UAP56	Sub2
	UAP56B	AT5G11200	UHR49	
	ALY1	AT5G59950		Yra1
	ALY2	AT5G02530		
	ALY3	AT1G66260		
	ALY4	AT5G37720		
	UIEF1	AT4G10970	UIF	
	UIEF2	AT4G23910		
	MOS11	AT5G02770	CIP29	Tho1
			CHTOP	
			LUZP4	
Complexe TREX-2	SAC3A	AT2G39340	GANP	Sac3
	SAC3B	AT3G06290		
	SAC3C	AT3G54380		
	CEN1	AT3G50360	CETN2	Cdc31
	CEN2	AT4G37010	CETN3	
	ENY2	AT2G19560	ENY2	Sus1
	DSS1	AT5G45010	DSS1	Sem1
	THP1	AT3G27100		Pcid2
NTR			NXF1	Mex67
			NXT1	Mtr2
Facteurs cyto.	LOS4	AT3G5110	DDX19	Dbp5
	GLE1	AT1G13120	GLE1	Gle1

Tableau I-2 | Les facteurs d'export des ARNm chez *Arabidopsis thaliana*. D'après (Ehrnsberger et al., 2019a)

Des expériences d'hybridation *in situ* ciblant les ARNm polyadénylés, grâce à une sonde fluorescente oligo dT, ont été réalisées sur plusieurs mutants d'Arabidopsis n'exprimant plus les différents composants du complexe TREX-1 afin de déterminer leur rôle dans l'export nucléaire des ARNm. Les plants d'Arabidopsis mutés dans le gène *MOS11* montrent une accumulation des ARNm dans le noyau suggérant très fortement que MOS11 est impliquée dans leur export nucléaire (Germain et al., 2010). Cette accumulation des ARNm est exacerbée chez le double mutant d'Arabidopsis *tex1-4/mos11-2* (Sørensen et al., 2017) alors qu'une mutation simple dans le gène *TEX1* n'a strictement aucun effet, ce qui indique que la protéine TEX1 est aussi impliquée d'une certaine manière dans l'export nucléaire chez Arabidopsis. La fonction exacte de MOS11 n'est pas connue mais il est vraisemblable qu'elle forme, par analogie avec son homologue humain CIP29, un complexe trimérique avec UAP56 et ALY, et qu'elle stimule l'activé hélicase de la protéine UAP56. Enfin, la protéine HPR1 – *hyper recombination 1* - du complexe THO est également un facteur d'export nucléaire car la mutation de son gène conduit à la rétention d'ARNm dans le noyau (Pan et al., 2012; Sørensen et al., 2017; Pfaff et al., 2018).

3.2. Les protéines adaptatrices d'Arabidop*sis* pour l'export des ARNm

Les quatre protéines ALY d'Arabidopsis sont exprimées de manière ubiquitaire dans les tissus végétatifs mais d'une façon différenciée dans les tissus reproducteurs comme le montre l'absence d'ALY4 dans le pollen. Elles sont toutes présentes dans le nucléoplasme et ALY2 est également dans le nucléole (Pfaff et al., 2018).

L'étude de mutants *aly* montre que les simples et doubles mutants présentent un phénotype sauvage alors que le quadruple mutant, *4xaly*, est très affecté et se caractérise par des plantes dont la rosette, la hampe florale et la racine primaire sont réduites (Pfaff et al., 2018). De plus, la morphologie florale de ce mutant est modifiée et les ovules anormalement développés, ce qui affecte la production des graines. Des hybridations *in situ* réalisées sur les cellules du quadruple mutant montrent clairement une accumulation nucléaire des ARNm confirmant le rôle primordial des protéines ALY dans l'export nucléaire (Pfaff et al., 2018). Le blocage de l'export des ARNm n'étant pas total, cela signifie que d'autres protéines adaptatrices interviennent dans ce processus, à l'image de ce qui se passe chez les métazoaires.

Des analyses *in silico* ont permis la détection de deux protéines similaires aux facteurs UIF qui jouent chez l'Homme, le même rôle que ALY (Heath et al., 2016). Ces deux protéines d'Arabidopsis, UIEF1/2, sont exprimées comme ALY, dans les tissus végétatifs et interagissent avec l'ARN hélicase UAP56. Dans le cas d'UIEF1, il a été montré qu'elle interagit avec l'ARN malgré l'absence de motifs de liaison à l'ARN connu et que son extrémité C-terminale contribue à cette interaction (Ehrnsberger et al., 2019a). Les plantes double mutantes *2xuief* montrent une rétention nucléaire des ARNm qui devient très accentuée sans être complète pour autant, chez le sextuple mutant *4xaly 2xuief* ce qui démontre d'une part, que les facteurs ALY et UIEF agissent de concert dans le transport nucléo-cytoplasmique des ARNm et d'autre part, que d'autres adaptateurs participent à l'export nucléaire chez les plantes (Ehrnsberger et al., 2019a, 2019b). Les protéines SR impliquées dans l'épissage feraient d'excellentes candidates (Rausin et al., 2010) d'autant plus que de nombreux pré-ARNm renferment des introns, notamment chez Arabidopsis (Lorković et al., 2000; Reddy, 2007). Des études faites sur la protéine SR RSZp22 (orthologue de SR 9G8), mutée au niveau de son site de liaison aux ARN, semblent conforter cette hypothèse (Rausin et al., 2010).

3.3. État de l'art de l'export nucléaire des ARNm chez les plantes

Arabidopsis se singularise des mammifères et de la levure, par la diversité de plusieurs constituants de la machinerie d'export nucléaire des ARNm. En effet, comme nous venons de le voir, son génome code quatre protéines adaptatrices ALY, deux protéines UIEF ainsi que respectivement, deux protéines THO5, THO7 et UAP56 ce qui conduit à un assemblage de

différentes versions du complexe TREX-1. Cette diversité suggère que les complexes TREX-1 pourraient être impliqués de par leur composition, en fonction de la présence des différentes ALY ou UIEF, dans l'export nucléaire de certaines classes d'ARNm et/ou dans des tissus particuliers. Cette hypothèse est d'autant plus probable vu l'expression différenciée des protéines ALY dans les tissus végétatifs et reproducteurs. Il est possible enfin que sa composition varie dans des conditions de stress telle qu'une infection virale. Une étude récente de Choudury et al. (2019) semble corroborer cette hypothèse car ces auteurs ont montré que la protéine ALY1 joue un rôle central dans l'export d'ARNm codant certaines protéines, par exemple AGO6, impliquées dans la voie du RdDM – *RNA-directed DNA methylation* –.

Le recrutement des adaptateurs et des transporteurs au niveau des ARNm, au cours de la transcription par l'ARN polymérase II et de la maturation des transcrits, reste énigmatique. En effet, l'analyse protéomique des partenaires d'interaction de l'ARN polymérase II, a permis d'identifier de nombreux facteurs impliqués dans la maturation des pré-ARNm (épissage, polyadénylation), mais aucun des acteurs de l'export nucléaire identifiés à ce jour, peut-être tout simplement, parce que ceux-ci interagissent faiblement avec la polymérase (Antosz et al., 2017).

Evidemment, le maillon manquant essentiel pour comprendre le mécanisme de l'export nucléaire des ARNm chez les plantes est l'identification de leurs exportines. Aucune protéine présentant des homologies avec TAP-p15 des métazoaires n'a été détectée jusqu'à maintenant, malgré l'analyse approfondie des banques de données. Arabidopsis possède deux gènes *XPO1* codant deux orthologues, XPO1a et XPO1b, de l'exportine CRM1 (Haasen et al., 1999; Blanvillain et al., 2008). Leur implication dans l'export nucléaire des protéines pourvues d'un NES a été clairement démontrée mais malheureusement aucune étude n'a été entreprise sur les mutants simples d'Arabidopsis *xpo1a-3* et *xpo1b-1*, - le double mutant étant létal -, pour déterminer si l'absence d'une des deux exportines conduisait à une rétention nucléaire des ARNm. Par conséquent, on ne sait pas si ces transporteurs pourraient contribuer à l'export d'ARNm comme décrit chez l'Homme (Blanvillain et al., 2008). Certes, une étude a bien montré que des ARNm non coiffés introduits dans des protoplastes de *Nicotiana plumbaginifolia* sont retenus dans le noyau après traitement avec la leptomycine B, l'inhibiteur de CRM1, conduisant les auteurs à conclure qu'il existe bel et bien chez les plantes un système d'export des ARNm NES-dépendant similaire à celui décrit pour CRM1(Stuger et Forreiter, 2004). Toutefois, pour l'instant, il n'a pu être établi clairement, sans ambiguité, que l'exportine CRM1 intervenait dans l'export des ARNm matures chez les végétaux.

Enfin, le complexe TREX-2 d'Arabidopsis, impliqué dans l'ancrage du mRNP au pore nucléaire, présente également quelques différences avec ses homologues de la levure et des vertébrés (**Tableau I-2**). Cette plante possède 3 orthologues de la protéine Sac3 mais seules

SAC3a et SAC3b semblent faire partie du complexe TREX-2. Des analyses structurales montrent que Sac3 est dépourvue du domaine central impliqué dans l'interaction avec la protéine ENY2 (Sus1) chez les autres organismes, ce qui explique pourquoi cette dernière n'a pas été détectée dans le complexe TREX-2. Pourtant, ENY2 (Sus1) est bien exprimée chez Arabidopsis et fait partie du complexe SAGA (Pfaff et al., 2018; Sørensen et al., 2017). Actuellement, peu d'information est disponible sur les mécanismes sous-jacents impliquant TREX-2 dans l'export nucléaire des ARNm, à part, l'existence d'interactions entre THP1 et SAC3 et, avec des Nups et le fait, que la déficience en THP1 provoque une nette accumulation nucléaire des ARNm arguant pour son rôle direct dans l'export nucléaire. En revanche, une déficience en SAC3 n'a strictement aucune répercussion sur l'export chez les plantes mutées, contrairement à la levure et aux vertébrés.

Les résultats de toutes ces études suggèrent ainsi que le mécanisme et la régulation de l'export nucléaire des ARNm chez les plantes diffèrent de ceux de la levure et des métazoaires. La complexité de TREX-1 et l'existence d'une population d'ARNm hétérogène constituée de transcrits partiellement ou totalement épissés, laissent supposer que le système d'export nucléaire est moins stricte chez les plantes.

4. L'export nucléaire des ARN viraux

Les virus sont des parasites intracellulaires obligatoires qui détournent et/ou modifient à leur propre profit, les machineries cellulaires de leurs hôtes pour effectuer leur cycle de réplication. Les systèmes d'export nucléaire des protéines et des ARN n'échappent pas à cette règle.

Pratiquement tous les virus à ADN, les rétrovirus et quelques rares virus à ARN de polarité négative passent par une étape nucléaire au cours de leur cycle infectieux. Par conséquent, ils doivent exporter leurs ARN dans le cytoplasme où s'effectuent la synthèse de leurs protéines et la production des nouvelles particules virales. Or de nombreux transcrits viraux exportés contiennent encore des introns, à l'inverse des ARNm des animaux, car ils sont souvent partiellement épissés ou pas du tout, afin que les virus puissent exprimer l'ensemble de leurs protéines. De ce fait, les virus ont du adopter des stratégies d'une part, pour échapper à l'épissage intégral et au système de contrôle-qualité et d'autre part, pour usurper la machinerie d'export et la forcer à exporter leurs ARN. Pour illustrer notre propos, nous avons sélectionné quelques virus qui sont remarquables par l'originalité de leurs stratégies d'export de leurs ARN. Ainsi, les rétrovirus possèdent au niveau de leur ARN génomique des structures secondaires reconnues par des protéines adaptatrices virales ou cellulaires et selon les virus, ils utilisent la voie CRM1 ou la voie TAP-p15. Concernant les virus à ADN, nous présentons les adénovirus et les herpesvirus dont l'export de leurs ARNm requiert des protéines virales et, le virus de l'hépatite B

où une structure secondaire de l'ARN est importante pour l'export des transcrits. Nous finirons avec le virus de la grippe A dont le cycle infectieux comprend une étape nucléaire ; ses ARN génomiques de polarité négative sont exportés par la voie CRM1 grâce à des protéines virales.

4.1. L'export des ARN rétroviraux

L'export des ARN des *Retroviridae* est étudié depuis de nombreuses années. Les rétrovirus sont des virus enveloppés renfermant deux copies d'un génome à ARN de polarité positive, coiffé et polyadénylé. Ils synthétisent, en début de cycle, par transcription inverse de leur génome à ARN, un ADNc qui est importé dans le noyau de la cellule hôte, puis intégré dans les chromosomes ; il prend alors le nom d'ADN proviral. Chez la majorité des rétrovirus, l'import nucléaire de l'ADN proviral est dépendant de la division cellulaire mais s'effectue chez les rétrovirus du genre *Lentivirus*, à travers les pores nucléaires, en dehors de la mitose (Le Sage et Mouland, 2013). Après intégration, l'ADN proviral est exprimé par l'ARN polymérase II cellulaire en ARN génomique et ARN épissés qui sont ensuite exportés dans le cytoplasme où s'effectuent toutes les autres étapes du cycle infectieux.

Nous traiterons d'abord de l'export nucléaire chez un rétrovirus complexe, le virus de l'immunodéficience humaine, et chez un rétrovirus simple, le virus Mason-Pfizer de singe dont l'étude a largement contribué à mieux comprendre les mécanismes impliqués dans le transport nucléo-cytoplasmique des ARN chez les métazoaires et, à identifier le dimère d'exportine TAP-p15. Si l'export des ARN messagers et génomique de ces deux virus fait souvent cas d'école, il s'avère que certains rétrovirus utilisent d'autres stratégies pour exporter leurs ARN partiellement ou non épissés. Pour exemples, nous aborderons brièvement l'export des ARN chez le virus du sarcome de Rous, les virus *foamy* (spumeux) et chez le virus de la leucémie murine.

4.1.1. L'export chez le virus de l'immunodéficience humaine, un rétrovirus complexe

Les virus de l'immunodéficience humaine de type 1 (VIH-1/HIV-1) ou de type 2 (VIH-2/HIV-2), à l'origine de la maladie du SIDA pour – *syndrome d'immunodéficience acquise* – (Rojas-Araya et al., 2015), font parti de la famille des *Retroviridae* et du genre des *Lentivirus*. En 2017, environ 36,7 millions de personnes étaient infectées par le VIH-1 ou VIH-2 dans le monde et 940 000 personnes sont décédées de maladies liées au SIDA selon ONUSIDA (Vih.org, 2018).

La transcription de l'ADN proviral du VIH-1 est effectuée par l'ARN polymérase II cellulaire qui en raison de sa faible processivité, ne produit que peu de transcrits complets (9 kb), coiffés et polyadénylés. Toutefois, ceux-ci sont suffisants pour générer après épissage, l'ARNm de la protéine virale Tat qui stimule ensuite l'ARN polymérase II (Gatignol, 2007) de sorte que cette dernière synthétise alors de nombreuses copies d'ARN génomique coiffé et polyadénylé (Karn et Stoltzfus, 2012) codant les protéines structurale et enzymatique Gag et Gag-pol. Grâce aux multiples sites d'épissages (au moins 4 sites donneurs et 8 sites accepteurs) qu'ils arborent, ces

longs ARN viraux subissent un épissage alternatif à l'origine des différents ARNm partiellement épissés (4kb) ou totalement épissés (2kb) (**figure I-7**)(Emerman et Malim, 1998; Tazi et al., 2010).

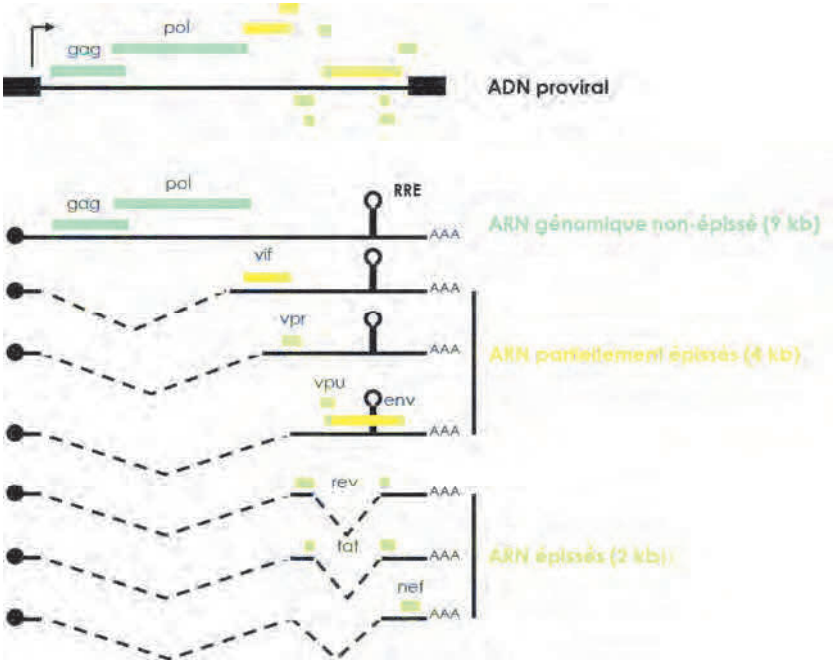


Figure I-7 | Génome et transcrits du VIH-1. L'ADN proviral du VIH-1 est représenté par une ligne noire avec deux rectangles noirs aux extrémiés, représentant ses LTR – *long terminal repeat* -. Cet ADN contient 9 ORFs – on parle aussi de gènes pour *Gag, Pol* et *Env* -. Au moins 70 transcrits différents sont produits grâce à l'utilisation d'au moins 4 sites donneurs et 8 sites accepteurs d'épissage (non représentés). Les ARN complètement épissés du VIH-1 codent les protéines Rev, Tat et Nef. Vif, Vpr, Vpu et Env sont exprimées à partir d'ARN partiellement épissés. L'ARN non épissé code les protéines Gag et Pol, il s'agit de la forme génomique infectieuse encapsidée dans les virions. Tous les ARN possédant des introns portent la tige-boucle RRE présente dans le gène *env* et responsable de leur export. La coiffe en 5' est représentée par un rond noir, la queue polyA par AAA et les introns par des traits en pointillés. D'après Mahiet et Swanson, 2016.

Les premiers ARN exportés correspondent aux transcrits des protéines virales Tat, Rev et Nef, des ARNm complètement épissés (2 kb)(Purcell et Martin, 1993). En tout point ressemblant aux messagers cellulaires, ces ARN viraux utilisent la voie canonique TAP-p15 impliquée dans l'export nucléaire des ARNm épissés (Cullen, 2003; Mahiet et Swanson, 2016) (**figure I-11**). En revanche, les ARN partiellement épissés (4 kb) codant Env/Vpu, Vif, Vpr et Tat de même que l'ARN génomique complet non épissé (9 kb) n'empruntent pas cette voie mais usent d'une stratégie bien particulière impliquant l'exportine CRM1 (**figure I-10**)(Neville et al., 1997). Normalement, les ARNm retenant des introns sont considérés comme aberrants par les machineries en charge du contrôle qualité et retenus dans le noyau jusqu'à leur épissage complet ou dégradés (Cao et Liu, 2007). Pour être exportés, les ARN non épissés du VIH-1 requiert une petite protéine adaptatrice d'origine virale appelée Rev – *regulator of expression of virion proteins* – (Emerman et al., 1989).

Rev est une phophoprotéine produite précocement dans le cycle de réplication du VIH-1, à partir d'un ARNm parfaitement épissé (Watts et al., 2018). Son ARNm est exporté par TAP-p15, traduit dans le cytoplasme puis Rev est acheminée dans le noyau par l'importine β qui reconnait un domaine riche en arginines situé à son extrémité N-terminale et lui servant de NLS (Groom et al., 2009). Ce domaine est flanqué de part et d'autre, par deux séquences qui permettent à la protéine d'oligomériser (DiMattia et al., 2010). Dans le noyau, ce même domaine permet à Rev (Hauber et al., 1988; Cochrane et al., 1989) d'interagir avec l'ARN et de reconnaitre une structure secondaire appelée RRE – *rev response element* – contenue, au niveau du gène *env,* dans tous les transcrits du VIH-1 retenant des introns (Cochrane et al., 1990; Watts et al., 2009, 2018).

L'élément RRE, de 351 nt, est une structure secondaire complexe formée de plusieurs tige-boucles et de *bulges* (Fang et al., 2013) où le bras IIB est le site initial de fixation et d'oligomérisation de la protéine Rev (Rausch et Le Grice, 2015). Un deuxième bras, IA, constitue aussi un site où Rev se lie avec une grande affinité (Fang et al., 2013)(**figure I-8**). Cependant, la manière dont ces deux sites sont reconnus par Rev reste encore floue. Quatre à 13 copies de la protéine Rev peuvent être retrouvées sur la structure RRE (Fang et al., 2013). Selon le modèle le plus récent, des dimères de Rev préformés s'associeraient à chacun des deux bras pour former un tétramère (DiMattia et al., 2016; Watts et al., 2018). Cette dimérisation permettrait à la protéine de masquer son NES, situé à son extrémité C-terminale, avant d'être associée à l'ARN et ainsi, de prévenir son export sans son cargo et de favoriser son accumulation dans le noyau (Behrens et al., 2017). Le démasquage du NES indispensable à l'export pourrait être gouverné par la phosphorylation de Rev, comme cela a été démontré notamment pour le facteur cellulaire NFAT1(Zhu et McKeon, 1999; Hogan et al., 2003; Behrens et al., 2017).

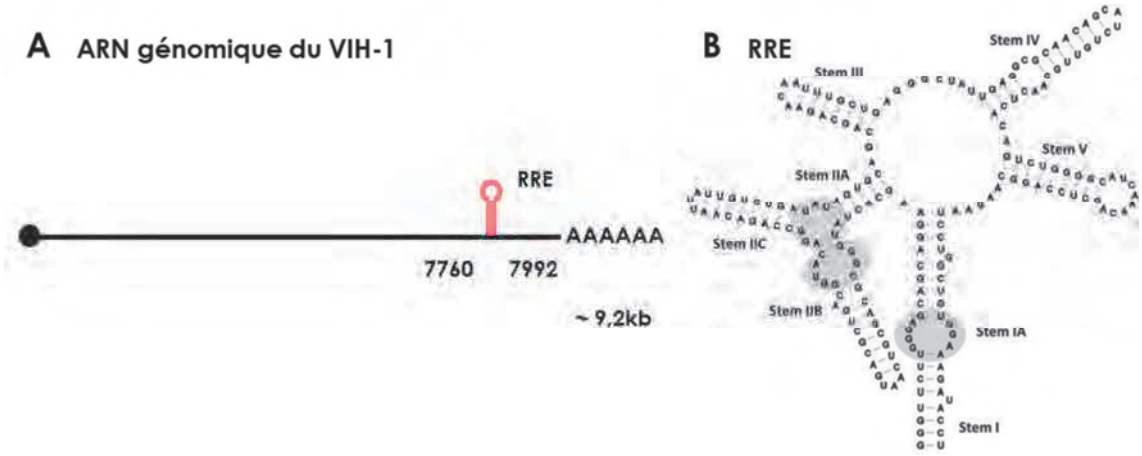


Figure I-8 | Structure secondaire du RRE porté par les ARN non épissés de VIH-1. (**A**) Localisation (nt 7760-7992) de la structure secondaire RRE – *rev response element* – (rouge) sur l'ARN génomique coiffé et polyadénylé du VIH-1 (9,2 kb). (**B**) Structure en 5 branches, adoptée par RRE à son extrémité. La base de la tige (Stem I) a été tronquée. La numérotation des différents éléments structuraux est indiquée. Les sites de liaison de Rev sont indiqués par des ronds gris. D'après Rausch et Le Grice, 2015.

Une fois fixées, les protéines Rev recrutent au moins deux exportines CRM1 (Booth et al., 2014) *via* leur NES riche en leucines (Malim et al., 1991; Fischer et al., 1995). Une étude récente montre clairement que Rev possède des affinités différentes pour les exportines CRM1 de l'Homme et de la souris. Toutefois, cette différence d'affinité ne serait pas liée à son NES mais plutôt à un autre domaine, inconnu, de la protéine (Yue et al., 2018). Ainsi, l'ensemble de ses domaines permet à Rev de faire le pont entre les ARN viraux *via* leur RRE et l'exportine CRM1 (Booth et al., 2014; Fornerod et al., 1997). Rev interfère également avec la protéine NCBP1 du CBC des transcrits viraux coiffés ce qui lui permet de bloquer leur association avec l'exportine TAP-p15 (Taniguchi et al., 2014). En plus de CRM1 et de NCBP1, Rev interagit aussi avec pléthores de facteurs d'export tel que eIF5A, hRIP, DDX3, DDX1 ou encore SAM68 (Suhasini et Reddy, 2009) ce qui force le système d'export cellulaire à sélectionner de préférence les mRNP viraux. Plusieurs études montrent que ces derniers interagissent notamment avec les nucléoporines NUP214, NUP153, NUP98 et NUP62 grâce à Rev et ses différents co-facteurs (Hofmann et al., 2001; Monette et al., 2011). Signalons qu'au-delà de ses fonctions dans l'export, Rev est aussi impliquée dans d'autres étapes du cycle du VIH-1, la polyadénylation, la stabilisation, la traduction et l'encapsidation de ARN génomique du virus (Groom et al., 2009; Rojas-Araya et al., 2015).

D'autres rétrovirus complexes qui recrutent CRM1 pour exporter leurs ARN partiellement ou non épissés, expriment aussi une protéine adaptatrice. Il s'agit des protéines Rex chez le virus T-lymphotropique humain de type 1 (HTLV-1) (Bogerd et al., 1998), Rej pour le rétrovirus jaagsiekte du mouton (JRSV) (Nitta et al., 2009), Rem chez le virus de la tumeur mammaire de la souris (MMTV) (Müllner et al., 2008) et K-Rev qui est exprimée par les rétrovirus endogènes humains (HERV-Ks) (Yang et al., 1999).

- **Le virus T-lymphotrophique humain 1.** Le HTLV-1 appartient au genre *Deltaretrovirus*. Il est à l'origine d'une leucémie à cellules T chez l'Homme mais il peut aussi provoquer chez les patients infectés une paraparésie spastique tropicale ou une uvéite (Mathew et Ghildyal, 2017). Tout comme le VIH-1, son ADN proviral est transcrit par une ARN polymérase II activée par un facteur de transcription viral, la protéine Tax, en quelques copies complètes d'ARN génomique qui subiront ou non, un épissage alternatif. Tax permet ainsi, en début du cycle infectieux, la production d'ARN complètement épissés codant la protéine Rex impliquée dans l'export des ARNm partiellement et non épissé dans la phase tardive du cycle viral (Nakano et Watanabe, 2016). Comme Rev, Rex reconnait une structure secondaire sur les ARN du HTLV-1 appelée RxRE – *Rex responsive element* – qui permet à ces ARN d'emprunter la voie de l'exportine CRM1. Cependant, contrairement aux transcrits du VIH-1, RxRE est présent sur tous les ARN du HTLV-1 y compris ses isoformes parfaitement épissées car il est localisé au niveau du 3' LTR – *long terminal repeat* - (Baydoun et al., 2008), mais l'efficacité de reconnaissance des différents ARN par Rex est variable. On pense que la structure RxRE intervient aussi dans l'étape de polyadénylation

des transcrits viraux (Ahmed et al., 1991) ce qui expliquerait pourquoi le virus a conservé cette structure sur tous ses transcrits et à cette position.

4.1.2. Le virus Mason-Pfizer du singe, l'export chez les rétrovirus simples

Le virus Mason-Pfizer du singe (MPMV) est un rétrovirus dit simple c'est-à-dire que son génome ne porte que les gènes *gag, pol* et *env*. Contrairement aux *Lentivirus* et autres rétrovirus complexes, les rétrovirus simples ne codent pas des protéines auxiliaires et plus particulièrement, des protéines de type Rev. Pourtant, comme tous les rétrovirus, ils doivent exporter au cours de leur cycle, des ARN épissés ou non étant donné que les gènes *gag* et *pol* sont exprimés à partir de l'ARN génomique et le gène *env*, à partir d'un ARN épissé (Okamura et al., 2015). Le MPMV comme d'autres rétrovirus simples, fait appel à une stratégie différente de celle des rétrovirus complexes. Ces rétrovirus possèdent au niveau de la région 3' UTR de leurs ARN, un élément en *cis* appelé CTE – *constitutive transport element* – suffisant à lui seul pour recruter les exportines TAP-p15 et permettre l'export de ces ARN sans l'intervention d'un quelconque autre facteur viral (**figure I-10**)(Bray et al., 1994; Ernst et al., 1997a). Cette observation a conduit à la découverte du rôle de l'hétérodimère TAP-p15 dans l'export nucléaire des ARNm cellulaires chez les métazoaires, grâce à une approche biochimique (Grüter et al., 1998) et à des expériences de compétition menées dans des oocytes de *Xenopus*. En effet, Pasquinelli et al. (1997) ont constaté que l'export nucléaire des ARNm cellulaires était bloqué suite à la micro-injection dans le noyau des oocytes, de grandes quantités de petits ARN contenant le CTE du MPMV et que deux protéines interagissaient avec cette structure.

Le CTE est une structure secondaire en forme d'épingle à cheveux avec une boucle apicale de 9 nt et de 2 boucles internes de 16 nt (**figure I-9**)(Ernst et al., 1997b). Les domaines RBD - *RNA binding domain* - et LRR - *Leucin-rich repeats* - de l'extrémité N-terminale de l'exportine TAP interagissent physiquement avec les boucles internes du CTE et permettent ainsi l'export direct de ces ARN sans qu'il ait eu épissage (Liker et al., 2000). Li et al. (2006) ont montré qu'un CTE du même type était présent dans le pré-ARNm de Tap, au niveau de son dixième intron. La traduction de l'ARNm incomplètement épissé donne naissance à une forme tronquée de l'exportine dont la fonction est inconnue pour le moment (Okamura et al., 2015). Ce résultat suggère que la présence du CTE permettrait à certains ARNm cellulaires spécifiques non épissés d'échapper au système de contrôle qualité du noyau et d'être exportés vers le cytoplasme ce qui pose la question de savoir lesquels des deux, des virus ou de la cellule, a développé cette stratégie.

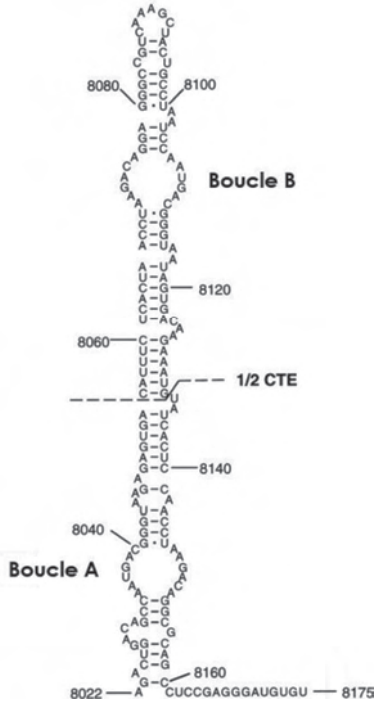


Figure I-9 | Structure secondaire du CTE du MPMV. Structure prédite par Ernst et al., (1997) de l'élément CTE – *constitutive transport element* – porté par les ARN de MPMV. Il est admis que le CTE se replie en une structure hélicoïdale comprenant deux boucles de même taille, nommées respectivement A et B. Elles servent de sites spécifiques de liaison aux protéines. La structure couramment appelée « ½ CTE » correspond à la partie supérieure du CTE ; elle fixe TAP. La numérotation se réfère au génome complet du MPMV. D'après Kang et Cullen, 1999.

4.1.3. À chacun sa stratégie

• **Le virus du sarcome de Rous exporte son ARN génomique grâce à sa protéine Gag.** Le virus du sarcome de Rous (RSV) est un rétrovirus aviaire du genre *Alpharétrovirus* qui se distingue des autres rétrovirus simples, par la présence d'un gène additionnel, l'oncogène *src* en aval de *env*. Les gènes *src* et *env* sont tous deux exprimés à partir d'ARN épissés. Le RSV exporte ces deux ARNm épissés et son ARN génomique d'une manière similaire à MPMV, en utilisant la voie de l'exportine TAP-p15 qui est recrutée sur les ARN viraux *via* deux séquences répétées appelée DR – *direct repeat* – (**figure I-10**)(LeBlanc et al., 2007). Néanmoins, certaines données suggèrent que ce virus utiliserait également une seconde voie qui impliquerait sa protéine Gag, pour exporter son ARN génomique.

Gag est capable de faire la navette entre le noyau et le cytoplasme. Elle est importée dans le noyau grâce à deux NLS présents l'un dans le domaine NC – nucléocapside – et l'autre dans le domaine MA - matrice -. Ces deux NLS interagissent indépendamment, celui de NC avec les importines α/β (Butterfield-Gerson et al., 2006) et l'autre avec Mtr10p (encore appelé importine 11 ou transportine SR) et Kap120p (Gudleski et al., 2010; Jouvenet et al., 2011). Elle possède également un NES qui lui permet d'être exportée du noyau par CRM1 (Scheifele et al., 2005). Gag se lie à l'ARN génomique du RSV en interagissant avec son site d'encapsidation ψ et de dimérisation qui est constitué d'une structure à deux épingles à cheveux (Miyazaki et al., 2010). Elle procède ensuite à son encapsidation en oligomérisant sur ce site ce qui conduit à un changement de conformation exposant son NES (D'Souza et Summers, 2005; Scheifele et al., 2007). Ainsi, en interagissant avec CRM1 *via* son NES et l'ARN génomique, Gag promeut l'export nucléaire du génome du RSV et ce, indépendamment des deux ARN viraux épissés. Le

complexe Gag-génome n'est pas réimporté dans le noyau car les domaines MA et NC de Gag, sans doute engagés dans une plus forte interaction avec l'ARN, ne peuvent plus interagir avec les importines (Jouvenet et al., 2011). La navette nucléo-cytoplasmique effectuée par Gag est donc primordiale pour l'export nucléaire et l'encapsidation du génome viral (Garbitt-Hirst et al., 2009; Parent, 2011).

• **Les virus foamy ou pourquoi ne pas utiliser une protéine adaptatrice d'origine cellulaire pour recruter CRM1 ?** Les *Foamy virus* (FV) ou virus spumeux font parti du genre des *Spumavirus*, dont le membre type est le *Prototype foamy virus* (PFV) encore connu sous le nom d'HFV – *Human foamy virus* – (Hamann et Lindemann, 2016). Contrairement à la majorité des rétrovirus, le génome des FV possède, en plus du promoteur classique présent dans le LTR, un promoteur interne (IP) (Löchelt, 2003), à partir duquel sont produits des ARN subgénomiques épissés codant deux protéines auxiliaires, l'activateur transcriptionnel Tas et Bet, une protéine de fonction mal définie (Lindemann et Rethwilm, 2011). La transcription à partir du promoteur de la région LTR est strictement Tas-dépendante et permet la synthèse d'ARN complets coiffés et polyadénylés qui serviront de génome et de messager pour la protéine Gag et, après épissage simple ou multiple, d'ARNm pour les protéines Env, Pol et Tas. Cette dernière est donc exprimée à partir de transcrits issus de deux promoteurs différents (Lindemann et Rethwilm, 2011).

L'export des ARN partiellement et non épissés de FV est dépendant comme pour le VIH-1 de la voie de l'exportine CRM1, mais à l'inverse des autres rétrovirus complexes, les *Spumavirus* ne codent aucune protéine virale auxiliaire impliquée dans ce processus (Bodem et al., 2011). L'export de l'ARN génomique du FV dépend en fait d'une protéine adaptatrice cellulaire riche en leucines, la protéine HuR qui se lie à l'ARN viral sur une structure encore non identifiée et de deux co-facteurs, les protéines ANP32A et ANP32B qui établissent un pont entre HuR et CRM1 (**figure I-10**)(Bodem et al., 2011; Hamann et Lindemann, 2016; Moschall et al., 2017).

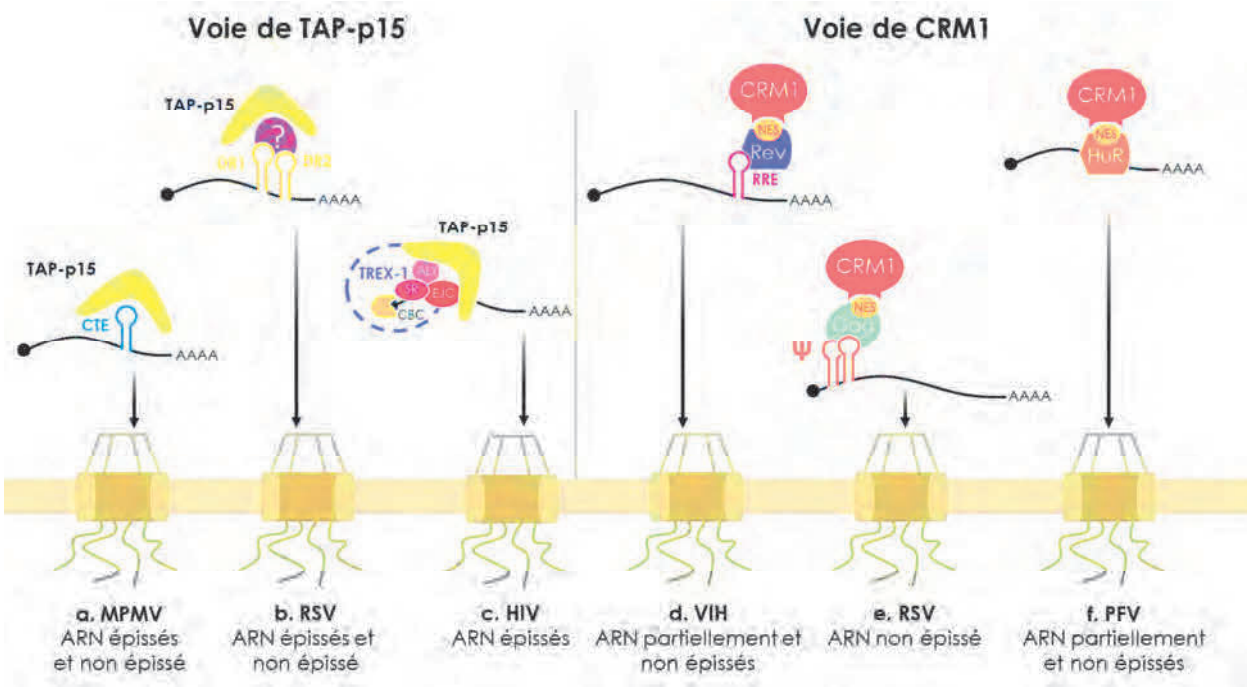


Figure I-10 | Représentation schématique des différentes stratégies employées par les rétrovirus pour exporter leurs ARN. Les ARN rétroviraux possèdent dans leur grande majorité des structures secondaires reconnues par la machinerie d'export, le CTE des ARN du MPMV reconnu directement par TAP (**a**), ou des tige-boucles (DR) sur les ARN du RSV reconnues par TAP grâce à une protéine adaptatrice encore inconnue (?)(**b**). Le VIH utilise les deux voies d'export, TAP-p15 et CRM1, la première pour l'export de ses ARNm parfaitement épissés (**c**) et la deuxième, pour celui des ARN partiellement ou non épissés grâce à Rev, une protéine virale à NES, qui se fixe sur la structure secondaire RRE des ARN (**d**). Le RSV utilise la protéine virale Gag pour l'export de son ARN génomique par la voie CRM1. Gag se lie au signal d'encapsidation ψ, situé à l'extrémité 5' de l'ARN génomique (**e**). PFV ne code pas de protéines adaptatrices mais utilise pour exporter ses ARN partiellement et non épissé, la protéine cellulaire HuR. La structure secondaire reconnue par cette dernière n'est toujours pas identifiée (**f**).

• **Le virus de la leucémie murine.** Le MLV, membre type du genre *Gammaretrovirus*, est l'un des premiers rétrovirus à avoir été découvert, il y a plus de 70 ans (Pessel-Vivares et al., 2015a). Le MLV est également un rétrovirus simple et à ce titre, comme le MPMV, il ne code pas de protéines auxiliaires pour réguler l'épissage et l'export de ses transcrits et il utilise uniquement la voie des récepteurs TAP-p15 pour exporter ses ARN (Pessel-Vivares et al., 2014). Cependant la stratégie adoptée par le MLV est quelque peu différente de celle des autres rétrovirus simples car elle implique non seulement plusieurs facteurs d'export mais aussi de multiples séquences de l'ARN viral (Jouvenet et al., 2011; Bartels et Luban, 2014; Pessel-Vivares et al., 2014; Sakuma et al., 2014).

Sakuma et al., (2014) ont montré que l'export nucléaire des ARN du MLV dépend des protéines UAP56, THOC5 et THOC7 du complexe TREX-1 mais pas de ses autres composants dont ALY/REF et CIP29. De plus, seule UAP56 parait primordiale pour l'export des isoformes épissées de cet ARN alors que les protéines THOC5 et THOC7 semblent être impliquées spécifiquement dans l'export de l'ARN génomique (Sakuma et al., 2014). L'export de ce dernier dépend par ailleurs du facteur d'épissage SRp20 qui participe au recrutement de l'exportine sur le transcrit (Bartels et Luban, 2014). Ces résultats suggèrent que l'ensemble des ARN du MLV emprunte la même

voie d'export, celle de TAP-p15, mais que leur export est régulé différemment en fonction de leur épissage et qu'elle implique des facteurs spécifiques.

Concernant les séquences en *cis* de l'ARN du MLV, cinq éléments structuraux répartis le long du génome semblent être impliqués dans l'export des ARN viraux alors qu'un seul l'est chez le MPMV : la tige-boucle RSL dans la séquence R, l'élément ψ, la structure régulatrice PTE – *post transcriptional element* - contenant un CTE (γ-CTE) et, l'élément CAE – *cytoplasmic accumulaion element* – ces deux derniers étant localisés dans le gène *pol* (Pessel-Vivares et al., 2015a). La boucle de γ-CTE contient la séquence conservée AAGACA qui interagit directement avec les facteurs d'export TAP-p15 (Grüter et al., 1998). Les structures ψ (Mougel et al., 1996; Basyuk et al., 2005; Smagulova et al., 2005) et PTE y compris γ-CTE (Bartels et Luban, 2014) sont uniquement présents dans les ARN codant Gag et Gag-Pol et réguleraient l'export de l'ARN viral non épissé alors que les structures RSL (Cupelli et al., 1998) et CAE (Sakuma et al., 2014) sont dans toutes les isoformes de l'ARN viral et permettraient leur export. Enfin, des études *in vitro* montrent qu'une petite séquence, dans le gène *Pol,* codant l'intégrase et la région U3 du MLV, sont capables de promouvoir l'export de l'ARN non épissé et ce, indépendamment de leurs positions sur l'ARN (**figure I-11**) (Hoshi et al., 2002; Volkova et al., 2014).

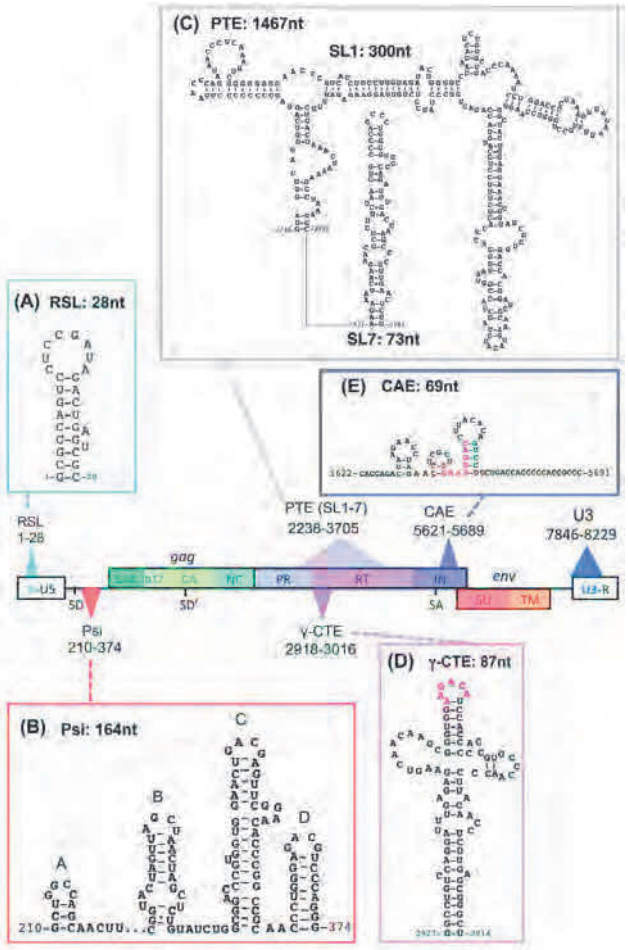


Figure I-11 | Structures secondaires de l'ARN de MLV impliquées dans l'export. De 5' vers 3' : la tige-boucle de 28 nt RSL – *R stem-loop* – (**A**), le signal d'encapsidation ψ de 164 nt organisé en 4 tige-boucles (**B**), L'élément d'export PTE – *post-transcriptional export* – de 1467 nt dont la structure complexe renferme 7 tige-boucles (SL 1 à 7) (**C**) et aussi γ-CTE (**D**) et la séquence AAGACA (en rouge) retrouvée aussi chez MPMV. L'élément CAE de 69 nt arborant 3 tige-boucles distinctes et sa séquence core en rouge (**E**). En 3' de l'ARNg du MLV, la région U3, certainement requise pour l'export de cet ARN non épissé. D'après Pessel-Vivares, 2015.

4.2. L'export nucléaire des ARNm des virus à ADN

Exceptés les *Poxviridae* et les *Megaviridae* dont le cycle infectieux est strictement cytoplasmique (Knipe et Howley, 2013; Claverie et Abergel, 2016, 2018), tous les virus à ADN répliquent et transcrivent leur génome dans le noyau de leurs cellules hôtes. De ce fait, ces virus doivent faire exporter leurs ARNm vers le cytoplasme pour leur traduction afin de produire de nouveaux virions. Parmi les nombreuses familles de virus à ADN infectant les eucaryotes, nous décrirons l'export nucléaire des ARN des *Adénoviridae* et des *Herpesviridae*, pour lesquels les mécanismes d'export sont les mieux connus. L'export chez d'autres virus comme les *Polyomaviridae* (Acheson, 1981; Huang et Carmichael, 2009; Carmichael, 2016) ou les *Papillomaviridae* (Graham, 2010) pour lesquels cette étape est bien moins décrite, ne sera pas abordé ici. Finalement, nous traiterons plus particulièrement de l'export des ARN du virus de l'hépatite B, un *Hepadnaviridae*, qui appartient au super-groupe des pararétrovirus. Ces derniers sont intéressants car ils doivent exporter du noyau non seulement leurs ARNm mais également un ARN légèrement plus long que le génome, appelé ARN pré-génomique, qui servira dans le cytoplasme de matrice à la transcriptase inverse virale pour synthétiser l'ADN viral double brin.

4.2.1. L'export complexe des ARNm des adénovirus

Les adénovirus (AdV) doivent leur nom aux végétations adénoïdes à partir desquels ils ont été isolés pour la première fois, il y a plus de 65 ans (Knipe et Howley, 2013). Cette famille virale est divisée en 5 genres dont celui des *Mastadenovirus* qui comprend plus de 60 types d'adénovirus humains différents répartis en 7 espèces classées de A à G (Benkö, 2008; Leppard, 2014). Ces virus sont responsables de maladies respiratoires, gastro-intestinales, urogénitales ou encore oculaires (Meng et al., 2019). Nos connaissances portent principalement sur les adénovirus de type 2 (Ad2) et 5 (Ad5) de l'espèce C.

Les adénovirus possèdent une capside icosaédrique non enveloppée, avec des fibres polypeptidiques très caractéristiques, qui renferme une molécule d'ADN double-bin linéaire dont la longueur varie de 30 à 36 kpb chez les *Mastadenovirus* (Davison et al., 2003).

Le génome viral est transcrit dans le noyau par l'ARN polymérase II cellulaire, à l'exception des gènes *va* – *virus-associated* – qui le sont par l'ARN polymérase III. L'expression de l'ADN viral est finement régulée temporellement, avec au total plus de 50 protéines synthétisées au cours du cycle. On distingue ainsi les unités transcriptionnelles immédiatement précoce E1A, précoces E2A, E2B, E3 et E4, intermédiaires E1B, IX, MLTU (L4) et IVa2 et, tardives MLTU (L1 à L5) et UXP, soit 11 promoteurs reconnus par l'ARN polymérase II (**figure I-12**)(Leppard, 2014). L'expression de ces unités est à l'origine d'une multitude d'ARNm, grâce à la mise en

œuvre d'un épissage alternatif intensif et l'utilisation de différents sites de polyadénylation. Seules les unités IX et UXP donnent naissance à un seul produit (**figure I-12**).

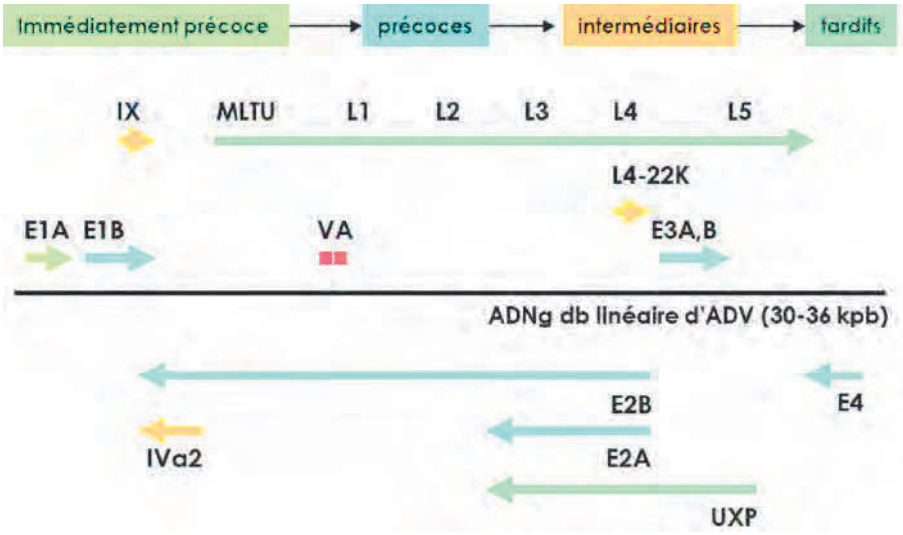


Figure I-12 | Les unités transcriptionnelles de l'adénovirus (type 5). L'ADN génomique est représenté par une ligne noire. Les différents gènes exprimés par l'ARN polymérase II sont représentés par des flèches dont la couleur indique à quelle phase l'expression commence. Ainsi on distingue les gènes immédiatement précoce (vert clair), précoces (bleu), intermédiaires (jaune) et tardifs (vert foncés). Les ARN *va – virus associated* - exprimés par l'ARN polymérase III sont représentés par deux rectangles rouges. D'après Leppard, 2014.

L'épissage alternatif génère évidemment des ARNm viraux renfermant des séquences introniques (Dobner et Kzhyshkowska, 2001) ce qui les fait apparaître au système de contrôle qualité comme des pré-ARNm immatures. Les adénovirus ont donc établi comme les rétrovirus, une stratégie pour éviter la rétention nucléaire de ces ARNm. Cette stratégie implique les protéines virales E4 ORF6 et E1B 55K qui favorisent avec E4 ORF3, l'export nucléaire des ARNm viraux tardifs (Halbert et al., 1985; Gonzalez et Flint, 2002) par la voie de l'exportine TAP-p15 (Dobner et Kzhyshkowska, 2001; Yatherajam et al., 2011) tout en inhibant celui des ARNm cellulaires (Blanchette et al., 2008). Les protéines E4 ORF6 et E1B 55K sont majoritairement à localisation nucléaire bien qu'elles puissent faire la navette entre le noyau et le cytoplasme.

Après son entrée dans le noyau, le génome adénoviral s'associe aux corps nucléaires PML – *promyelocytic leukemia* – qui sont impliquées notamment dans la défense antivirale et dans le contrôle post-transcriptionnel des ARN cellulaires (Lallemand-Breitenbach et de Thé, 2018). La réplication et transcription virales ont lieu à leur périphérie, comme le prouve la détection de l'ADN et des ARN du virus (Bridge et Pettersson, 1996). Au fur et à mesure que le cycle viral progresse, les protéines précoces E4 ORF3 et E1B 55K s'associent à ces structures et entraînent leur réorganisation (Carvalho et al., 1995; Leppard et Everett, 1999). L'accumulation des transcrits viraux tardifs conduit vers la fin du cycle, à la formation de structures appelées granules interchromatiniens (*nuclear speckles* en anglais) qualifiés « d'élargis » à cause de leur grande taille (Carvalho et al., 1995). Ces derniers sont enrichis en facteurs d'épissage et

constituent le lieu de stockage des transcrits tardifs épissés compétents pour l'export (Aspegren et al., 1998). À ce stade, E1B 55K est relocalisée à la périphérie de ces granules (Ornelles et Shenk, 1991) grâce à E4 ORF6 (Rubenwolf et al., 1997) qui rentre en compétition avec E4 ORF3 (Leppard et Everett, 1999). Toute cette réorganisation a pour effet d'inhiber la maturation et l'export des ARNm cellulaires tout en favorisant celui des ARNm viraux bien que renfermant des introns, dans leur grande majorité (Dobner et Kzhyshkowska, 2001; Flint et Gonzalez, 2003). Il est admis que des protéines du système de surveillance et de la machinerie d'export, seraient ciblées par le complexe viral E1B 55K/E4 ORF6 ce qui permettrait aux ARNm adénoviraux d'échapper à la dégradation et d'être exportés vers le cytosol (Leppard, 1993). Ce complexe viral pourrait en fait induire la dégradation par le protéasome, de certains facteurs indispensables pour l'export des messagers cellulaires mais pas pour celui des ARN viraux, en les polyubiquitinylant (Woo et Berk, 2007). En effet, E1B 55K et E4 ORF6 forment avec les élongines B et C, la culline 5 et le facteur RBX1, une E3 ubiquitine ligase dont on sait que l'activité est essentielle à l'export des ARNm viraux tardifs (Blanchette et al., 2008).

Les adénovirus utilisent la voie de l'exportine TAP-p15 pour exporter leurs ARNm tardifs (**figure I-15 a.**)(Flint et al., 2005; Yatherajam et al., 2011) alors que dans un premier temps, on pensait que ces ARNm étaient exportés par la voie CRM1, notamment à cause du NES porté par E1B 55K (Dobbelstein et al., 1997; Rabino et al., 2000; Carter et al., 2003). En effet, l'inhibition par ARN interférence de l'exportine TAP ou l'emploi d'un dominant négatif de cette protéine entrainent une inhibition significative de l'export de ces ARN (Yatherajam et al., 2011). Des études suggèrent que la protéine E1B 55K pourrait être impliquée dans l'export des ARNm tardifs bien qu'elle soit incapable d'interagir directement avec l'exportine TAP et d'autres composants de la machinerie d'export. Cette hypothèse a été émise suite à la découverte d'une interaction entre E1B 55K et une hnRNP, appellée à cet effet, E1B-AP5 – *E1B 55K associated protein* -, (Gabler et al., 1998). Cette protéine cellulaire interagit avec l'ARN (Dobner et Kzhyshkowska, 2001) et *via* son extrémité N-terminale, avec l'exportine TAP (Bachi et al., 2000). Elle partage aussi, au niveau de sa région centrale, une forte homologie de séquence avec le domaine de liaison au GTP de la petite GTPase Ran (Gabler et al., 1998). De plus, E1B-AP5 colocalise avec E1B 55K en périphérie des centres de réplication virale (Yatherajam et al., 2011) et sa surexpression stimule l'export des ARNm tardifs (Gabler et al., 1998). Enfin, des mutations dans E1B 55K pour abolir son interaction avec E1B-AP5, entraînent une inhibition de l'export des ARNm tardifs viraux et/ou une relocalisation de la protéine virale en dehors des sites de réplication (Yatherajam et al., 2011).

4.2.2. L'export chez les Herpesviridae des ARNm issus de gènes sans introns

La famille des *Herpesviridae* est constituée par un groupe important de virus causant une très grande variété de maladies à une large gamme d'hôtes, dont l'Homme pour neuf d'entre eux (Gruffat et al., 2019; Knipe et Howley, 2013). Tous les virions de cette famille présentent la même organisation : un ADN double-brin linéaire, long de 124 à 295 kpb, enfermé dans une capside icosaédrique recouverte d'un tégument et d'une enveloppe lipidique dans laquelle sont enchâssées des glycoprotéines (Pires de Mello et al., 2016). Les herpesvirus sont remarquables par leur capacité à effectuer un cycle infectieux lytique ou à entrer en latence et d'être ultérieurement réactivé lors d'un stress. L'infection lytique ou l'entrée en latence des herpesvirus est dictée par les protéines immédiatement précoces ICP – *infected cell polypeptide* - et les modifications des histones associées à l'ADN viral, ce dernier étant converti en chromatine dès son entrée dans le noyau de la cellule hôte (Nevels et al., 2011).

Malgré la grande taille du génome de ces virus, comparée à celle d'autres virus à ADN comme celui des *Adenovirus*, il contient très peu de séquences promotrices et régulatrices, et la majorité des gènes de protéines sont dépourvus d'introns : 75 % des gènes chez le *Kaposi sarcoma-associated herpesvirus* (KSHV) et jusqu'à 98 % pour l'*Herpes simplex virus* (HSV) (Vogt et Bohne, 2016).

Diverses études montrent que la protéine immédiatement précoce ICP27 des HSV-1 et HSV-2, joue un rôle majeur dans l'export nucléaire des ARNm viraux et qu'elle est nécessaire à l'expression des gènes précoces (β) et tardifs (γ). De plus, ICP27 perturbe fortement le métabolisme cellulaire en interférant avec l'épissage des transcrits cellulaires et en bloquant l'import nucléaire des protéines de l'hôte en interagissant avec les nucléoporines Nup-FG du complexe NUP62 des NPC (Koffa et al., 2001; Sandri-Goldin, 2011; Malik et al., 2012). Cette protéine multifonctionnelle est exprimée à partir d'un ARNm épissé – son gène contient un intron – qui est exporté par la voie canonique TAP-p15 (Vogt et Bohne, 2016). L'importance fonctionnelle de cette protéine pour les *Herpesviridae* est corroborée par le fait que des protéines similaires sont trouvées chez tous les herpesvirus étudiés à ce jour (Tunnicliffe et al., 2015, 2018), comme par exemple, ORF57 du KSHV (Majerciak et Zheng, 2015) ou encore Mta/EB2 du virus d'Epstein-Barr (EBV), tous deux appartenant au genre *Gammaherpesvirus* (Sergeant et al., 2008).

ICP27 et ses homologues sont capables d'interagir directement avec les transcrits viraux et la protéine adaptatrice ALY/REF du complexe TREX-1 ce qui leur permet de recruter les exportines TAP-p15 sur les ARN viraux et de promouvoir leur export (**figure I-15 b.**)(Johnson et al., 2009a; Tian et al., 2013; Tunnicliffe et al., 2018). Les domaines fonctionnels de ces protéines sont : une boite RGG impliquée dans la liaison aux ARN (Mears et Rice, 1996), un NLS pour leur importation nucléaire (Sandri-Goldin, 1998), un domaine d'interaction avec la protéine adaptatrice ALY/REF (Tunnicliffe et al., 2018) et, un domaine globulaire, appelé IHD – *ICP27-*

homology domain – responsable de sa dimérisation (Patel et al., 2015), à son extrémité C-terminale.

La manière dont les protéines de type ICP27 reconnaissent et interagissent avec les ARNm viraux reste vague. Pour l'ICP27 des HSV, les séquences reconnues sont vraisemblablement riches en GC, non structurées et flexibles (**figure I-15 b.**)(Corbin-Lickfett et al., 2009) alors que l'ORF57 du KSHV pourrait reconnaitre des structures secondaires appelées ORE – *ORF-57 response element* – formées de deux tige-boucles (**figure I-15 c.**). Cet élément structural a été notamment mis en évidence chez le KSHV, en 5' de l'ARNm codant ORF59 et de l'ARN PAN, un long ARN non codant mais polyadénylé (Massimelli et al., 2011, 2015; Vogt et Bohne, 2016). Cependant l'interaction de ORF57 avec les ARN viraux étant trop faible (Majerciak et al., 2011), elle doit être renforcée par l'action conjointe de protéines cellulaires liant l'ARN. De nombreuses protéines impliquées dans l'export des ARNm cellulaires sont des candidates potentielles étant donné qu'ORF57 est capable d'interagir avec diverses protéines associées au complexe TREX-1 : UIF (Jackson et al., 2011), CHTOP, CIP29 (Schumann et al., 2013), RBM15 ou encore OTT3 (Majerciak et al., 2011). Concernant ICP27, elle peut interagir directement avec TAP (Johnson et Sandri-Goldin, 2009; Juillard et al., 2009; Hernandez et Sandri-Goldin, 2010), le CTD de l'ARN polymérase II (Dai-Ju et al., 2006) et avec des facteurs d'épissage (Bryant et al., 2001; Sciabica et al., 2003; Escudero-Paunetto et al., 2010).

La protéine ICP27 et ses homologues ont donc un rôle majeur dans le cycle de réplication des *Herpesviridae,* grâce aux interactions qu'elles engagent avec les ARN viraux et les protéines cellulaires. Elles permettent aux ARNm viraux d'être exportés par la voie classique TAP-p15 et ALY/REF, bien que non épissés, leurs gènes étant dépourvus d'introns. En interagissant, avec d'autres protéines virales, avec les nucléoporines, ICP27 provoque sans doute une réorganisation des NPC perturbant ainsi le transport nucléo-cytoplasmique des ARNm cellulaires (Moroianu et al., 1997). Si tel est le cas, il n'est pas exclu que le contrôle qualité des ARNm viraux soit moins contraignant au niveau des NPC, facilitant ainsi leur export vers le cytoplasme.

4.2.3. Le virus de l'hépatite B, l'export nucléaire chez un pararétrovirus des animaux

Le virus de l'hépatite B (HBV) est le plus petit des virus à ADN avec un génome de 3,2 kpb (Chen et al., 2005). C'est le membre type de la famille des *Hepadnaviridae*, une famille de virus au tropisme strictement hépatique et responsable de cirrhoses et cancers du foie chez l'Homme et les animaux. Il fait partie du super-groupe des pararétrovirus qui se distinguent des rétrovirus par le fait qu'ils possèdent un génome à ADN, que l'intégration de cet ADN dans les chromosomes de l'hôte n'est pas obligatoire pour le cycle infectieux et, que leur génome est répliqué en fin de cycle, dans le cytoplasme, par transcription inverse d'un ARN dit pré-génomique (ARNpg).

En raison de la taille réduite du génome, les 4 gènes majeurs du HBV se chevauchent et nécessitent des mécanismes originaux pour être exprimés. Tous les transcrits viraux sont synthétisés, par l'ARN polymérase II cellulaire, à partir de l'ADN circulaire clos covalemment (ADNccc), coiffés et polyadénylés (**figure I-13**). Les sites d'initiation de la transcription sont différents mais les 4 gènes partagent tous le même signal de terminaison de la transcription et de polyadénylation (Lee et al., 2008b). On trouve dans la cellule infectée, 5 transcrits majeurs. Deux de ces transcrits (3,5 kb) sont légèrement plus longs que le génome viral, les ARN messagers pré-core (ARNpc) et pré-génomique (ARNpg). L'ARNpc code la protéine précurseur pré-core qui est clivée post-traductionnellement générant ainsi la protéine HBeAg qui est sécrétée dans le milieu extérieur (Chen et al., 2005). L'ARNpg est polycistronique, il code pour les protéines de capside ou core et la transcriptase inverse (Pol) et sert de matrice pour la synthèse de l'ADN viral. Enfin, il y a les ARNm spécifiques des antigènes de surface HBsAg L (grand) de 2,4 kb (ARNpreS) et, HBsAg M (moyen) et HBsAg S (petit) de 2,1 kb (ARNpreS/S) et l'ARNm du facteur de transcription X de 0,7 kb (ARNX) (**figure I-13**)(Mitra et al., 2018). Environ 30 % des transcrits pc et pg sont épissés et certains ARNm épissés codent pour une protéine appelée HBSP pour – *HBV splice-generated protein* – dont le rôle est inconnu. Ces ARN épissés sont également exportés dans le cytoplasme où ils sont rétrotranscrits et encapsidés pour donner naissance à des particules virales défectueuses (Lee et al., 2008a; Sommer et Heise, 2008; Chen et al., 2014).

Figure I-13 | Génome et transcrits du HBV. Le génome à ADN circulaire de HBV (3,2 kpb), représenté ici sous une forme linéaire, est la matrice de 5 transcrits viraux : l'ARNpc, l'ARNpg, l'ARNpreS, l'ARNpreS/S et l'ARNX. Ces ARN codent 7 protéines virales : la protéine précore, la protéine de capside ou core, la polymérase (P), les grand (preS1), moyen (preS2) et petit antigènes de surface (S) et la protéine X. L'élément PTE – *post-transcriptional element* – est indiqué par l'encadré gris (nt 1151-1805). Le dédoublement du site de polyadénylation (nt 1919, flèches rouges), s'explique par la réprésentation linéaire du génome. Le point 0 sur l'ADN génomique est indiqué par la double flèche noire et la numérotation suit celle de ce dernier. D'après Sommer et Heise, 2008.

Tous les ARNm du HBV portent en *cis* une structure appelée PRE – *post-transcriptional regulatory element* – à leur extrémité 3' (**figure I-14**). Cette structure est subdivisée en deux éléments d'export SEP1 et SEP2, ce dernier contenant deux sous-éléments appelés PREα et PREβ qui adoptent chacun une structure en tige-boucle, respectivement SLα (absente chez l'ARNX) et SLβ (Smith et al., 1998b; Chi et al., 2014; Lim et Brown, 2016). La structure PRE peut être divisée en deux parties fonctionnellement distinctes. La première moitié de PRE (1151-1433) joue un rôle central dans l'export des ARN et serait un élément régulateur de l'épissage. La deuxième moitié en revanche ne serait impliquée que dans l'export des transcrits et recouvre notamment la séquence PRE-III qui est située entre les éléments SEP1 et SEP2 **(figure I-14)**(Visootsat et al., 2015).

PRE stimule l'expression des transcrits du HBV et promeut l'export des ARN subgénomiques non épissés du virus. Il a été démontré que PRE est crucial pour l'export de l'ARN preS/S (Sommer et Heise, 2008) mais qu'il n'est pas impliqué dans celui des ARNpc et ARNpg complets qui sont exportés grâce à un autre élément (**figure I-15-d**)(Lim et Brown, 2016). PRE peut se substituer partiellement à l'élément RRE du VIH-1 dans l'export d'un ARNm rapporteur qui est totalement dépendant du système RRE-Rev et de CRM1(Huang et Liang, 1993). Pourtant, on sait depuis de nombreuses années, grâce à des expériences réalisées avec la leptomycine B, que l'export des transcrits non épissés du HBV est indépendant de l'exportine CRM1 (Zang et Yen, 1999) et fonctionne indépendamment de toute protéine virale (Chen et al., 2014). Parallèlement, des analyses *in silico* ont montré que PRE ne partageait aucune similarité de séquence avec l'élément structural CTE des rétrovirus simples ce qui suggère qu'il est incapable d'interagir, en tout cas, directement avec l'exportine TAP-p15 (Chen et al., 2014). En résumé, le HBV emploierait des mécanismes tout à fait singuliers, différents de ceux des rétrovirus, pour exporter ses transcrits.

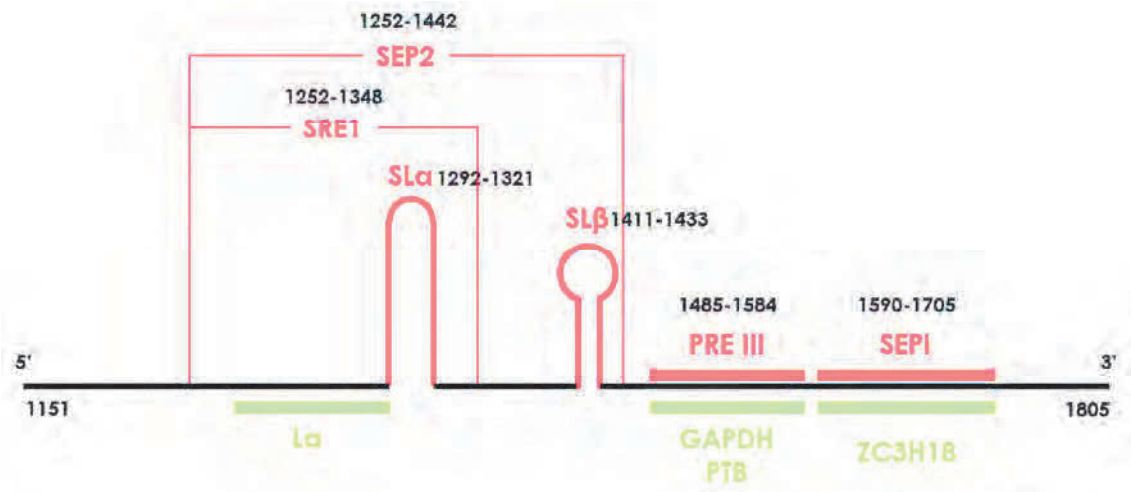


Figure I-14 | L'élément PRE des ARN de HBV. Carte rassemblant l'ensemble des éléments fonctionnels et structuraux connus de PRE (nt 1151-1805). PRE contient deux éléments majeurs d'export SEP1 (nt 1590-1705) et SEP2 (nt 1239-1442). SEP1 constitue le domaine de liaison à ZC3H18. SEP2 contient SRE-1 (nt 1252-1348), le site de liaison à la protéine La (nt 1275-1291) et, les tige-boucles SLα (nt 1292-1321) et SLβ (1411-1433). PRE peut également interagir avec les protéines PTB et GAPDH qui reconnaissent son élément PRE III (1485-1584). D'après Visoostat, 2015 ; Lim et Brown, 2016.

La protéine de capside ou core du HBV est capable de faire la navette entre le noyau et le cytoplasme. Elle possède à son extrémité C-terminale quatre NES qui sont riches en arginines, contrairement au NES canonique, et capables d'interagir directement avec l'exportine TAP-p15 (Li et al., 2010a). Core reconnait aussi certaines protéines du complexe TREX-1 (Yang et al., 2014). La protéine de capside qui interagit avec l'ARNpg (Gerelsaikhan et al., 1996; Gallucci et Kann, 2017), jouerait un rôle d'adaptateur pour l'exportine TAP-p15 (**figure I-15 e.**). Cependant, l'inhibition par ARN interférence des protéines ALY et DDX39 (UAP56) n'a pas d'effet sur l'export de l'ARNpg (Yang et al., 2014), mais ceci pourrait s'expliquer par la redondance des protéines cellulaires ayant les mêmes fonctions.

Concernant l'implication des protéines cellulaires dans l'export des transcrits du HBV, il a été montré que PRE interagit *via* son élément SEP1, avec la protéine cellulaire ZC3H18 qui permet de recruter le complexe TREX et des facteurs d'export comme ARS2 (Chi et al., 2014), La protéine La se lie *in vitro* à l'extrémité 5' PRE et les protéines GAPDH et PTB, à l'élément PRE-III (**figure I-14**) (Zang et al., 2001). L'interaction entre PRE-III et PTB est impliquée dans l'export du transcrit preS/S (Li et al., 2010b). Le rôle de la séquence Pre-III dans l'export nucléaire avait déjà été démontré par Huang et al., (1996) avec l'export du transcrit S dans lequel l'élément PRE avait été substitué par six copies de PRE-III.

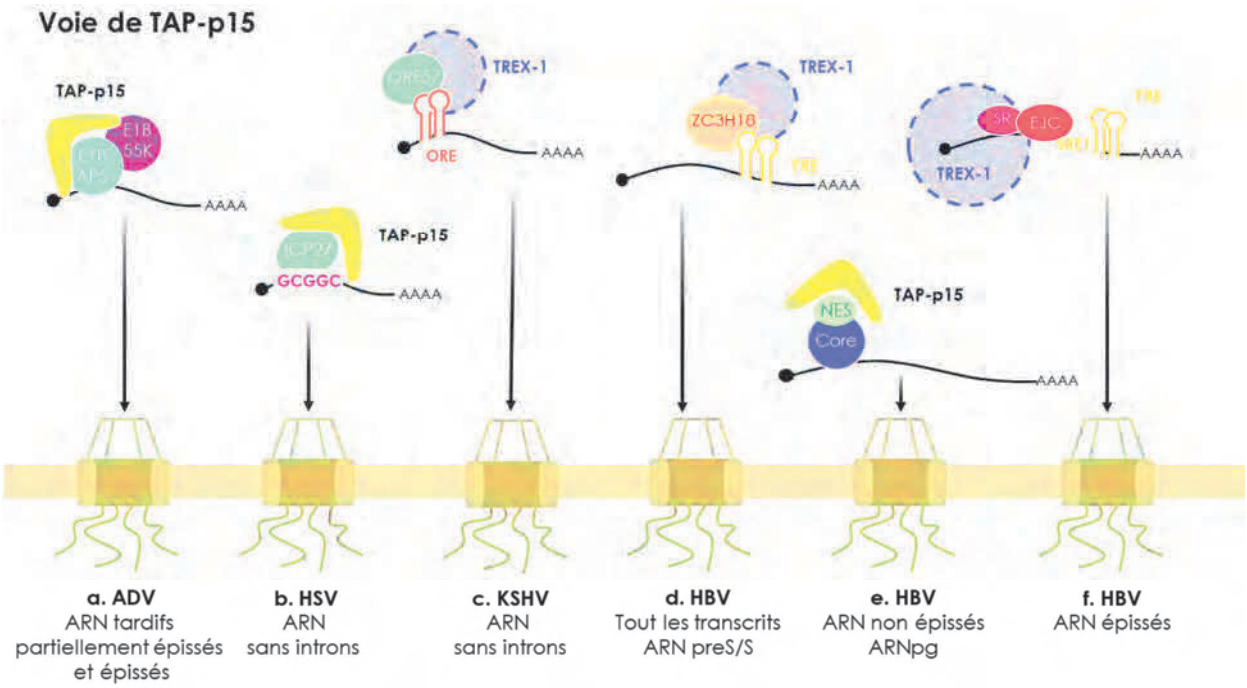


Figure I-15 | Stratégies d'export employées par les virus à ADN. Les virus à ADN détournent la voie de l'exportine TAP-p15 pour exporter leurs ARNm grâce à des protéines virales. Les adénovirus utilisent la protéine virale E1B 55K qui reconnait les ARN tardifs et recrute l'exportine TAP, probablement via l'hnRNP E1B-AP5 (**a**) Chez les herpesvirus, il s'agit de la protéine ICP27 chez le HSV (**b**) et ORF57 chez le KSHV (**c**). Les transcrits d'HBV portent une structure secondaire (PRE) reconnue par différentes protéines d'export dont la protéine ZC3H18 associée au complexe TREX-1 (**d**). L'export de l'ARN pré-génomique (pg) non épissé du HBV se ferait par l'intermédiaire de sa protéine de capside core dont les NES atypiques, riches en arginines, sont reconnus par TAP (**e**). Les transcrits épissés générés par le HBV à partir de l'ARNpg et recouverts de facteurs d'épissage, empruntent certainement la voie de l'exportine TAP-p15 (**f**).

Signalons que PRE est aussi impliqué dans l'épissage de l'ARN pré-génomique (Mitra et al., 2018; Su et al., 1989). Il possède l'élément régulateur d'épissage SRE-1 qui fonctionnellement peut être substitué par l'élément ESE, un *enhancer* présent chez les ARN rétroviraux. Cette observation suggère que la structure PRE pourrait promouvoir indirectement l'export des formes épissées de l'ARNpg grâce au dépôt sur ces transcrits, de facteurs d'épissage permettant le recrutement de Tap-p15 (Heise et al., 2006; Lee et al., 2008a; Sommer et Heise, 2008) (**figure I-15 f.**). Cet épissage ne concerne que 30% des transcrits car ce processus est apparemment contrecarré par la structure SLα qui serait reconnue par un voire plusieurs facteurs cellulaires empêchant l'assemblage *du spliceosome* et encourageant l'export (Hass et al., 2005; Lim et Brown, 2016). L'un de ces facteurs pourrait être la protéine TARDBP, découverte grâce à son interaction avec l'élément TAR dans l'ADN proviral du VIH-1 (Ou et al., 1995). En effet, Makokha et al., (2019) ont montré que TARDBP active non seulement la transcription du génome du HBV *via* son interaction avec le promoteur core, mais qu'elle inhibe également l'épissage des ARNpc et pg en se liant à leur extrémité 3' (Duriez et al., 2017; Makokha et al., 2019).

4.3. Le virus de la grippe, un virus à ARN qui effectue une partie de son cycle dans le noyau

Le virus de la grippe A ou virus influenza A (IAV) est l'un des très rares virus à ARN dont le cycle infectieux comprend une étape nucléaire et qui doivent exporter leurs ARN génomiques et ARNm dans le cytoplasme pour générer des particules virales.

Contrairement à la grande majorité des virus à ARN, le génome du virus de la grippe A (IAV) de la famille des *Orthomyxoviridae* est transcrit et répliqué dans le noyau des cellules hôtes (York et Fodor, 2013). Ce virus se présente sous la forme d'une particule enveloppée renfermant 8 molécules d'ARN génomique simple brin de polarité négative (ARN(-)) codant 10 protéines virales essentielles et 7 protéines virales auxiliaires. Chaque fragment est associé aux nucléoprotéines virales (NP) et à un complexe ARN polymérase ARN-dépendant (RdRp), composé des protéines PB1 et PB2 – *polymerase basic1 et* 2 – et PA – *polymerase acid* –, le tout formant un ensemble de 8 ribonucléoprotéines virales (RNPv) (Knipe et Howley, 2013).

Le virus de la grippe effectue à la fois la réplication de son génome multipartite et la synthèse de ses ARNm dans le noyau ce qui paraît surprenant à première vue pour un virus à ARN. Toutefois, l'étude approfondie de cette étape nucléaire montre que celle-ci constitue en fait, un véritable avantage évolutif pour l'IAV car elle lui permet d'accéder non seulement aux différentes machineries nucléaires de son hôte mais également à d'abondantes quantités d'ARN pré-messagers qu'il exploite pour la synthèse de ses propres ARNm (Amorim et Digard, 2006). Certains de ses ARNm font également l'objet d'un épissage alternatif par le spliceosome (Shih et al., 1998). Par ailleurs, il échappe plus facilement au sein du noyau, au système immunitaire inné cellulaire (Rehwinkel et Reis e Sousa, 2010).

Cette stratégie nucléaire a une contrepartie : IAV doit détourner à son profit le système d'export nucléaire et à cet effet, il a du adapter certaines de ses protéines pour l'export nucléaire de ses ARNm épissés ou non et, les ARN génomiques (**figure I-17**)(Fodor, 2013).

Le complexe RdRp viral étant démuni d'activité de coiffage, il va prélever l'extrémité 5' coiffée (environ 10-15nt) des ARNm cellulaires en voie de synthèse, appariée à l'extrémité 3' des ARN viraux matriciels puis après clivage, se servir de cette extrémité d'origine cellulaire comme amorce pour synthétiser les ARNm viraux. Ce mécanisme appelé vol de coiffe - *cap snaching* - a également pour effet de rendre les messagers cellulaires impropres à la traduction (De Vlugt et al., 2018). Les transcrits codant respectivement la protéine de matrice M1 et la protéine non structurale NS1, qui contrecarre l'action de l'interféron (Hale et al., 2008), peuvent subir un épissage alternatif pour générer des ARNm codant pour trois autres protéines virales. L'ARNm épissé issu du transcrit M permet la synthèse de la protéine M2, une viroporine à protons (Lamb et Choppin, 1981) alors que ceux, obtenus à partir du transcrit NS code respectivement pour la protéine NS2 encore appelée NEP et impliquée dans l'export nucléaire des génomes viraux (ARNv) (Lamb et Lai, 1980) et, pour la protéine NS3 chez certains isolats d'Influenza A (Selman et al., 2012). L'épissage des transcrits M et NS réalisé par le spliceosome (Lamb et Horvath, 1991; Dubois et al., 2014) fait intervenir des éléments en *cis* de ces ARN (Dubois et al., 2014) et des protéines virales, au moins NS1 pour le transcrit M (Robb et Fodor, 2012).

La proximité physique de l'ARN polymérase virale et de l'ARN polymérase II cellulaire lors du processus de *cap snatching* pendant la synthèse des ARNm viraux (Engelhardt et al., 2005) constitue un excellent tremplin pour que les transcrits viraux en cours de synthèse accèdent aux protéines impliquées dans la voie d'export canonique TAP-p15. L'importance de l'ARN polymérase II dans l'export nucléaire des transcrits d'IAV est démontré par la rétention nucléaire des ARNm HA, M1 et NS1 viraux lorsqu'on utilise la drogue DRB – 5,6-dicholoro-1-β-D-ribofuranosyl-benzimidazole - qui inhibe l'ARN polymérase II en phase d'élongation (Amorim et al., 2007). Quant à l'implication de l'exportine TAP-p15 dans l'export d'ARNm grippaux, elle s'est vérifiée par la co-immunoprécipitation de ces ARN avec la protéine TAP (Wang et al., 2008). Toutefois, le traitement des cellules avec des siARN – *small interfering RNA* – avant l'infection par IAV, pour empêcher l'expression de TAP, de ALY/REF ou de UAP56, indique que la dépendance vis-à-vis de TAP-p15 est variable selon les ARN viraux, les plus affectés par ce traitement étant les ARNm viraux HA, M1 et M2 (Read et Digard, 2010).

Le virus de la grippe doit exporter quatre catégories différentes d'ARN viraux : les ARN génomiques dépourvus de coiffe et de queue polyA, les ARNm dépourvus d'introns, les ARNm contenant un ou des introns et, leurs formes épissées (**figure I-17**). À l'heure actuelle, malgré un faisceau de données expérimentales, les mécanismes régissant l'export de l'ensemble des ARN de ce virus restent encore mal compris (York et Fodor, 2013)

• **L'export des ARNm d'IAV.** Il est vraisemblable que l'export des ARNm épissés M et NS s'effectue selon le même mécanisme que celui des ARNm cellulaires (Shih et Krug, 1996; York et Fodor, 2013). Le dépôt des complexes EJC lors de l'épissage permettrait à ces ARN de recruter l'exportine TAP-p15.

Les ARNm non épissés (M et NS) et les ARNm ne contenant pas d'intron (PB1, PB2, PA, HA, NP et NA), pourraient être exportés par la voie alternative d'export ou ALREX qui est dépendante du complexe CBC associé à la coiffe. Ce complexe fixé sur la coiffe des transcrits viraux recruterait directement l'adaptateur ALY/REF, des sous-unités du complexe THO et l'hélicase UAP56 puis les transcrits seraient exportés par TAP-p15 (Bier et al., 2011). L'autre alternative pour l'export des ARNm viraux dépourvus d'introns est le recrutement de TAP-p15 par les protéines SR, 9G8 et SRp20, sachant que ces dernières s'associent et participent à l'export nucléaire d'ARNm sans introns chez les métazoaires (Huang et al., 2003). Enfin, le recrutement de l'exportine TAP-p15 pourrait s'effectuer directement *via* des éléments structuraux en cis des ARNm du IAV comme décrit pour le rétrovirus MPMV (Morita et al., 2013). Cette hypothèse est d'ailleurs étayée par la co-immunoprécipitation de complexes TAP-ARNm viraux (Wang et al., 2008) mais les structures en *cis* pouvant être reconnues par l'exportine n'ont pas encore été identifiées.

Deux protéines du IAV participent directement à l'export de certains ARN, les protéines NS1 et NS2 (NEP). La protéine NS1 interagit avec l'exportine TAP (Wang et al., 2008; Zhang et al., 2019) et l'ARN hélicase UAP56 (Chiba et al., 2018) et, grâce à son domaine N-terminal, avec les ARNm viraux (Hale et al., 2008; Schneider et Wolff, 2009) suggérant fortement que NS1 joue le rôle de protéine adaptatrice. Par ailleurs, Pereira et al. (2017 et 2018) ont montré que NS1 est impliquée dans l'export des ARNm issus du fragment génomique viral 7, codant pour les protéines M1 et M2. Cette activité dépend du domaine « effecteur » C-terminal de NS1 et de sa capacité à interagir avec des facteurs cellulaires des voies de signalisation de PI3K ou RIG-I mais pour l'instant il n'y a pas de résultats probants pour établir un lien direct entre ces deux voies de signalisation et l'export nucléaire. NS1 interagit aussi avec une protéine cellulaire désignée NS1-BP qui est impliquée dans l'épissage du transcrit M et dans son export (Tsai et al., 2013). NS1-BP interagit avec l'adaptateur ALY/REF, l'ARN polymérase II et plusieurs facteurs d'épissage (hnRNP K, PTBP1 et SART1), ce qui suggère qu'elle est impliquée avec NS1, dans la formation du complexe d'export des ARNm M1 et/ou M2 (Zhang et al., 2018).

• **L'export des ARNv (-) d'IAV.** Le transport des nouveaux génomes viraux (ARNv) (O'Neill et al., 1998), du noyau vers le cytoplasme où a lieu la morphogenèse virale, s'effectue par la voie CRM1 (Watanabe et al., 2001; Iwatsuki-Horimoto et al., 2004). Il requiert la protéine virale NEP pour – *nuclear export protein* – (Paterson et Fodor, 2012), appelée originellement NS2 avant la découverte de son rôle dans l'export nucléaire (Lamb et Choppin, 1981). Elle joue le rôle de protéine adaptatrice entre l'exportine CRM1 et la protéine M1 associée aux complexes RNPv.

Chaque complexe est formé d'un ARNv recouvert de protéines NP et associé à la RpRd (PB1, PB2, PA).

La région C-terminale de NEP renferme deux hélices α qui lui confèrent un caractère amphipathique lui permettant de reconnaitre le NLS de la protéine de matrice M1 (Yasuda et al., 1993) qui interagit au sein des complexes RNPv avec les protéines NP (**figure I-16/17**)(Baudin et al., 2001; Paterson et Fodor, 2012). La région N-terminale de NEP, sensible aux protéases comparée à la partie C-terminale (Akarsu et al., 2003; Iwatsuki-Horimoto et al., 2004) renferme deux NES, l'un, NES1, inhabituel car il est riche en méthionines et l'autre, NES2, canonique, tous deux impliqués dans l'export nucléaire des RNPv par CRM1(la Cour et al., 2003; Paterson et Fodor, 2012; Huang et al., 2013). Dans les faits, c'est NES2 qui permet à NEP de recruter l'exportine sur les RNPv (Huang et al., 2013)(**(figure I-17).** Quant au rôle exact de NES1, il n'est pas compris car il a plutôt une action antagoniste voire inhibitrice sur CRM1 vu que sa délétion potentialise l'interaction de NEP avec CRM1 (Neumann et al., 2000; Huang et al., 2013). Ce modèle d'export des RNPv a été affiné récemment pour attribuer également un rôle dans ce transport à la RpRd suite à la mise en évidence d'une interaction entre NEP et les sous-unités PB1 et PB2 des RNPv (Brunotte et al., 2014). Cette interaction de NEP avec la RpRd dans les RNPv faciliterait l'interaction entre les protéines M1 et NP (**figure I-16**)(Lakdawala et al., 2016).

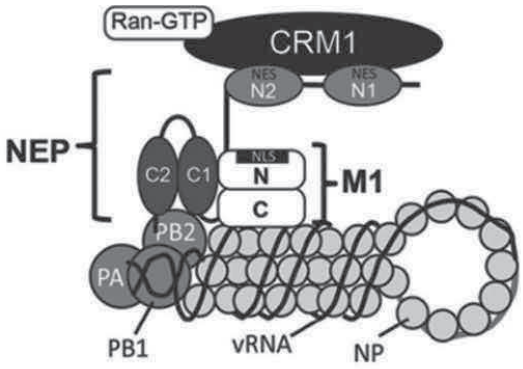


Figure I-16 | Représentation schématique des RNPv compétentes pour l'export. L'exportine CRM1 permet l'export des RNPv en liant le domaine N-terminal de NEP qui contient ses NES, ainsi que la petite GTPase Ran. L'extrémité C-terminale de NEP est engagée dans des interactions à la fois avec la protéine M1 (liée aux NP) et la polymérase virale. Ainsi, NES joue le rôle de protéine adaptatrice et fait le pont entre l'exportine et les ARN génomiques viraux, permettant ainsi leur export. D'après Brunotte et al., 2014.

Il est possible que l'export des RNPv puisse se faire aussi indépendamment de la protéine NEP uniquement grâce à la protéine de nucléocapside NP (Elton et al., 2001) car celle-ci contient plusieurs NES dont au moins un, NES3, interagit avec CRM1 (Yu et al., 2012). Il a été démontré récemment, en utilisant le composé ZBMD-1 qui cible directement le NES3, que l'export des NP et des RNPv était bloqué et que la réplication du IAV était inhibée *in vivo* et *in vitro* (Huang et al., 2018). Pour finir, M1 pourrait aussi intervenir dans l'export des nouvelles RNP virales car elle aussi contient un NES reconnu par CRM1 et sa surexpression permet de surmonter le blocage de ce transport, induit par une inhibition de l'activité de NEP (Bui et al., 2013).

Toutes ces données montrent que le virus de la grippe dispose de plusieurs alternatives pour exporter ses nouveaux génomes viraux, une redondance fonctionnelle qui évite sans doute au virus d'être bloqué en fin de cycle, au moment de la morphogenèse virale.

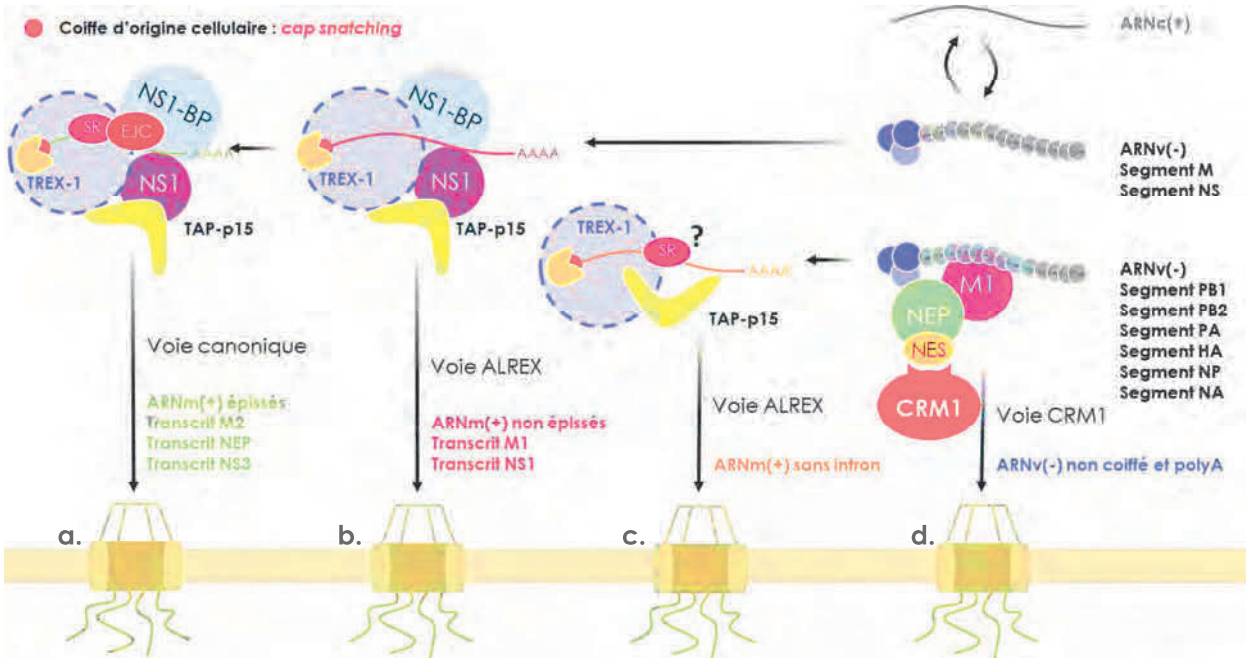


Figure I-17 | Export des ARN(+) et (-) du virus de la grippe A. Durant son cycle, l'IAV doit exporter 4 types d'ARN différents. Des ARN épissés, ressemblant aux ARNm cellulaires, qui empruntent certainement la voie canonique d'export TAP-p15 (**a**). Les formes non épissées de ces ARN utiliseraient la même voie d'export en recrutant les exportines par la voie ALREX grâce à leur extrémité 5' coiffée (**b**). La protéine virale NS1 et son facteur cellulaire associé NS1-BP sont impliqués dans l'export les transcrits M en interagissant avec ces ARN et TAP. IAV exporterait ses ARNm sans introns par la voie TAP-p15, grâce au complexe CBC associé à la coiffe et à des protéines SR (**c**). Les ARN génomiques (ARNv(-)) sont exportés par l'exportine CRM1 qui est recrutée par la protéine adaptatrice virale NEP. Cette dernière est recrutée sur les génomes par le complexe polymérase en 5' de ces ARN et la protéine M1 associée aux protéines NP des RNPv (**d**).

Projet de recherche

Projet de recherche

Étude de l'export nucléaire des ARNm du virus de la mosaïque du chou-fleur (CaMV) chez *Arabidopsis thaliana*

1. Ouverture sur le sujet d'étude : L'export des ARN viraux chez les plantes

Depuis sa découverte, il y a environ 40 ans, le CaMV a toujours suscité un certain intérêt bien qu'il ne soit pas un pathogène très important pour les plantes maraîchères, l'impact économique des viroses attribuées au CaMV étant minimes comparées à celles occasionnées par d'autres phytovirus. Toutefois, les découvertes faites en biologie moléculaire grâce au CaMV, comme le mécanisme du saut du ribosome, et sans doute la rareté des virus à ADN chez les plantes, lui ont valu d'être classé parmi les dix phytovirus les plus importants en virologie végétale (Scholthof et al., 2011). Différents volets de la pathogenèse du CaMV ont fait l'objet d'une recherche intensive par plusieurs équipes, mais peu d'études ont été entreprises pour comprendre l'étape nucléaire du cycle infectieux de ce virus, exception faite de la transcription et de l'épissage, ou chez d'autres membres de la famille des *Caulimoviridae*. Le peu d'intérêt pour cette étape du cycle de réplication du CaMV est d'autant plus surprenante qu'elle est déterminante pour l'issue de l'infection virale.

Jusqu'à cette étude, aucune information n'était disponible sur l'export nucléaire des transcrits du CaMV, ni sur l'ARN 19S, un messager qui ressemble en tout point aux ARNm cellulaires, ni sur l'ARN 35S qui est exporté dans son entier pour permettre l'expression de l'ensemble des protéines virales et la réplication de l'ADN et, sous la forme d'une population hétérogène d'ARN épissés. En effet, plus de 70% des ARN 35S sont soumis à un épissage alternatif intensif (Kiss-László et al., 1995; Bouton et al., 2015a) dont l'objectif serait d'empêcher, en début du cycle viral, la synthèse de la protéine P2 qui s'avère être toxique pour le CaMV et/ou de réguler temporellement l'expression des protéines virales. L'analyse moléculaire des ARN 35S exportés dans le cytoplasme a révélé l'existence de nombreux sites donneurs et accepteurs d'épissage de sorte que les ARN 35S épissés ou non, renferment toujours des introns, quel que soit le schéma d'épissage (Bouton et al., 2015a).

L'épissage alternatif est également intense chez les plantes où il est particulièrement impliqué dans la régulation de l'expression des protéines en réponse à des changements environnementaux et à des stress (Staiger et Brown, 2013). De ce fait, la rétention d'introns dans les ARNm matures est, à l'inverse des animaux, un phénomène très répandu chez les plantes, notamment chez *Arabidopsis thaliana* où plus de 61% des gènes de protéines contiennent des introns (Chaudhary et al., 2019). L'export d'ARNm partiellement épissés des plantes et du CaMV

indique que des mécanismes particuliers sont mis en œuvre chez les plantes, d'une part pour empêcher le *spliceosome* d'épisser totalement les transcrits et d'autre part, pour permettre aux ARNm renfermant des introns d'échapper au système de contrôle qualité et d'être exportés vers le cytoplasme.

Le mécanisme d'export nucléaire des ARNm est totalement méconnu chez les végétaux bien que différentes études aient permis de caractériser, deux homologues de CRM1 (XPO1a et b), les complexes TREX-1 et TREX-2, des facteurs d'export dont des protéines adaptatrices de même que des nucléoporines (voir **§3**). L'implication de certains de ces facteurs protéiques dans l'export des ARNm cellulaires a été démontrée grâce à l'étude de mutants d'Arabidopsis (Ehrnsberger et al., 2019a). En revanche, les transporteurs voués spécifiquement à cet export sont inconnus, aucun homologue de l'exportine TAP n'a été trouvé chez les plantes, ce qui suggère que celles-ci utilisent sans doute des exportines et/ou des voies d'export différentes de celles décrites chez la levure et les métazoaires. Quant aux exportines CRM1 de plantes, impliquées dans l'export des protéines, comme chez tous les organismes, aucune donnée expérimentale ne permet de leur attribuer un rôle quelconque dans l'export des ARNm.

Ainsi, l'étude de l'export nucléaire des ARN 35S et de l'ARN 19S du CaMV chez Arabidopsis est donc d'autant plus d'actualité qu'elle permettrait non seulement de comprendre cette étape du cycle infectieux de ce virus mais également de contribuer de manière plus générale à l'élucidation du mécanisme d'export des ARNm cellulaires chez les végétaux.

Evidemment, tous les phytovirus à ADN, les *Caulimoviridae*, les *Geminiviridae* et les *Nanoviridae* voire les *Nucléorhabdoviridae*, des virus à ARN effectuant une partie de leur cycle infectieux dans le noyau, auraient pu servir de modèles d'étude de l'export nucléaire des ARNm chez les plantes et plus particulièrement, le RTBV - *rice tungro bacilliform virus* - et le FMV - *figwort mosaic virus* - deux *Caulimoviridae* et, le MSV – *maize streak virus* - un *Geminiviridae,* qui produisent des ARN épissés et non épissés. En revanche, les viroïdes comme le PSTVd – *Potato spindle tuber viroid* – qui se répliquent dans le noyau sont totalement inadaptés pour une telle étude bien que l'implication de CRM1 dans leur export ait été évoquée dans une publication (Eiras et al., 2011) car il s'agit de petits ARN simple brin circulaires non codants sans aucune ressemblance avec les ARNm cellulaires.

Le choix du CaMV comme modèle d'étude pour investiguer l'export nucléaire des ARNm chez les phytovirus et plus généralement chez les plantes, s'est en fait imposé pour deux raisons essentielles :

- Le CaMV est depuis des années, notre modèle d'étude au laboratoire et par conséquent, nous avons une bonne connaissance de la biologie de ce virus. Nous disposons déjà d'une collection importante de clones recombinants qui permettent l'expression *in vitro* et/ou *in vivo* de plusieurs protéines virales et plus particulièrement, l'expression de la protéine

P6/TAV et de nombreuses formes mutées. Nous avons aussi des vecteurs de transcription recombinants pour produire des transcrits 35S complet et mutés/tronqués ainsi qu'un vecteur viral permettant la réplication du CaMV dans les plantes entières et en protoplastes.

- L'isolat de CaMV Cabb B-JI que nous avons utilisé, a notamment pour plante hôte, *Arabidopsis thaliana*, la plante modèle par excellence en biologie végétale. Nos connaissances sur les facteurs et les complexes impliqués dans l'export nucléaire chez les plantes, proviennent presqu'exclusivement des travaux réalisés sur Arabidopsis. Par ailleurs, plusieurs collections de plantes d'Arabidopsis mutés sont disponibles de même que des banques d'ADNc. Enfin, nous avons également créé au laboratoire des plantes d'Arabidopsis transgéniques exprimant P6/TAV et diverses versions mutées.

Avant d'exposer notre approche expérimentale et les résultats obtenus au cours de ma thèse, nous allons faire une description brève du CaMV et de son cycle de réplication. Des informations plus précises sont données dans des revues récentes (Bouton et al., 2015b; Schoelz et Leisner, 2017; Leisner et Schoelz, 2018; Pooggin et Ryabova, 2018).

2. Le virus de la mosaïque du chou-fleur

Le virus de la mosaïque du chou-fleur ou *Cauliflower mosaic virus* (CaMV), est le chef de file du genre des *Caulimovirus* et de la famille des *Caulimoviridae* (Bouton et al., 2015b). Seule famille de virus à génome ADN double brin (db) chez les plantes, c'est aussi le premier phytovirus à ADN identifié et dont le génome a été entièrement séquencé (Franck et al., 1980) et pour lequel la réplication de son ADN par transcription inverse d'un ARN pré-génomique a été démontrée (Pfeiffer et Hohn, 1983), une caractéristique majeure du super-groupe des pararétrovirus.

2.1. Plantes hôtes et symptômes

Le CaMV infecte la famille des Brassicacées (choux, navet, colza, *A.thaliana*…) et certaines solanacées (datura…), grâce à son vecteur, le puceron (Webster et al., 2018). La plupart des plantes infectées par le CaMV vont développer des symptômes dont la sévérité dépendra de l'isolat viral, de l'espèce et du cultivar végétal mais aussi des conditions environnementales ambiantes. Ainsi, chez *A. thaliana* Col-0, l'isolat virulent Cabb B-JI induit 10 à 14 jours après inoculation du virus ou de l'ADN viral et en conditions de lumière et de température fixes, une chlorose sous forme de mosaïque, un éclaircissement des nervures et un gaufrage des feuilles infectées. Généralement, un nanisme de la plante est aussi observé avec une réduction atteignant le quart de la taille d'une plante saine (**figure P-1**).

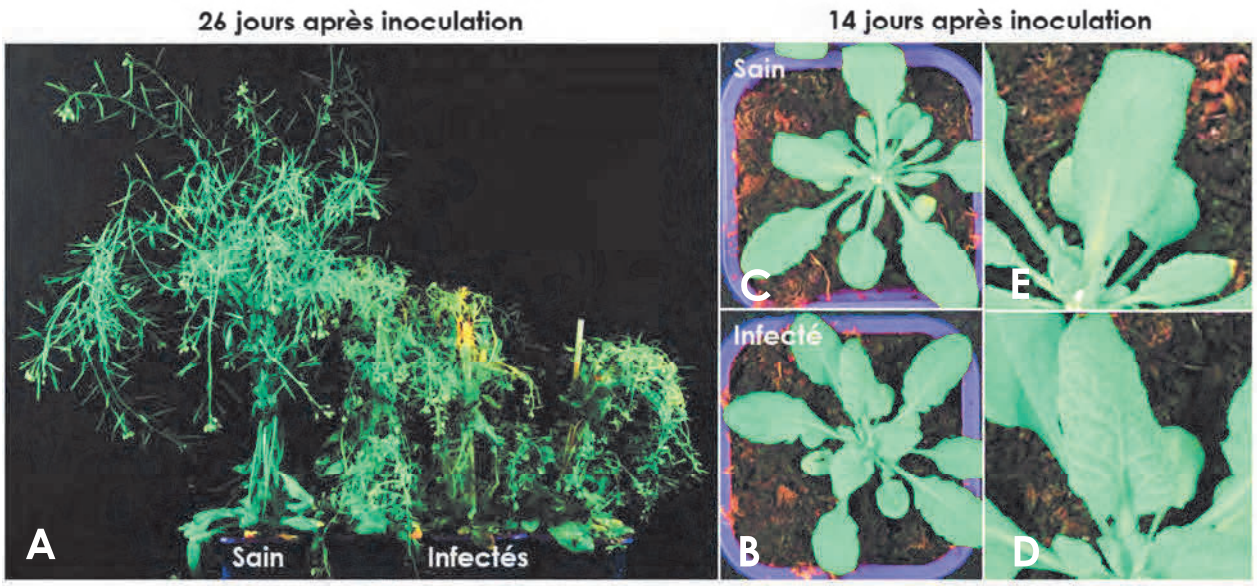


Figure P-1 | Symptômes exprimés par des plants d'Arabidopsis infectés par l'isolat de CaMV Cabb B-JI. **A** : nanisme des plantes infectées en comparaison avec la plante saine. **B** : mosaïque et éclaircissement des nervures observées sur les feuilles d'une plante infectée par le CaMV comparée à une plante saine (**C**). **D** et **E** sont des agrandissements respectivement d'une feuille de plantes infectée et saine.

2.2. Le génome du CaMV et ses transcrits

Le CaMV est un virus non enveloppé dont la capside icosaédrique (T=7) renferme une molécule d'ADN double-brin circulaire relâchée d'environ 8 kpb (**figure P-2**). Cette forme du génome viral est due à la présence de plusieurs interruptions de séquences, dont le nombre et les positions varient selon les isolats viraux. Elles sont générées au cours de la transcription inverse de l'ARN pré-génomique pour synthétiser l'ADN génomique et forment une structure en triple-brin caractéristique où l'extrémité 5' du brin chevauchant commence par une séquence ARN (Bouton et al., 2015b). Le génome de CaMV porte 7 ORFs (cistrons), tous présents sur son brin (-), et deux régions intergéniques de 700 et 150 pb, contenant respectivement les promoteurs de la transcription 35S et 19S (Franck et al., 1980). Les ORFs du CaMV se chevauchent ou sont seulement séparés de quelques nucléotides. Seul l'ORF codant la protéine P6 est isolé et situé entre les deux régions promotrices de l'ADN.

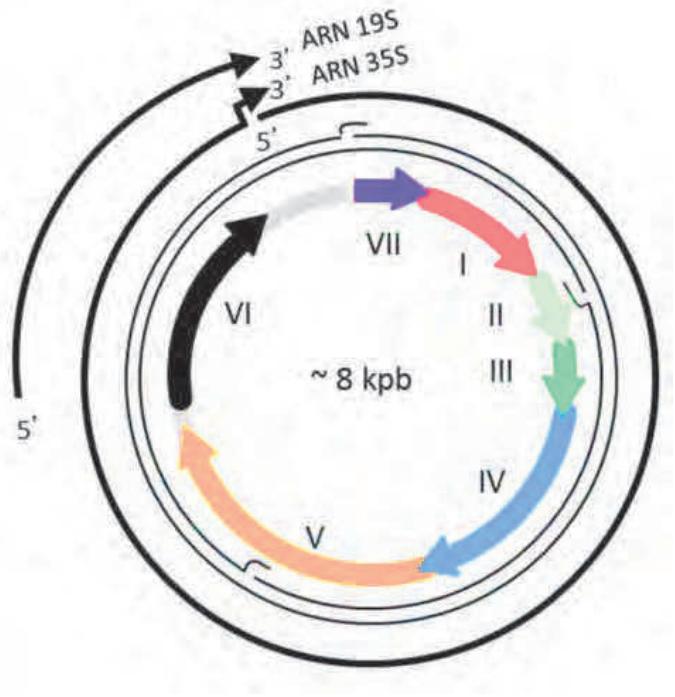


Figure P-2 | Organisation génétique de l'ADN du CaMV. Le génome ADN db de CaMV isolat Cabb B-JI est représenté avec ses 3 interruptions de séquence par des lignes noires fines. Les différents ORF (I à VII) sont indiqués par les flèches de couleur et les grande et petite région intergéniques sont en gris. Les transcrits majeurs du virus 19S et 35S sont représenté par les lignes noires épaisses. D'après Bouton et al., 2015b.

• **L'ADN de CaMV** est transcrit par l'ARN polymérase II cellulaire, très certainement contrôlée par le facteur de transcription P-TEFb, en deux transcrits majeurs, les ARN 19S et 35S, à partir de leurs promoteurs du même nom (**figure P-2**). Le promoteur 35S est un promoteur fort et constitutif, ce qui lui a valu célébrité et sa très large utilisation en ingénierie et biologie fondamentale (Benfey et Chua, 1990; Bhullar et al., 2007). Il contient notamment les boîtes TATA et CAAT, des éléments *enhancer* de type $GTGGX_3G$, ainsi que des domaines spécifiques gouvernant l'expression des ORFs dans différents tissus et à différents stades de développement de la plante. Les deux transcrits du CaMV sont coiffés et polyadénylés. Un ARN 8S est également produit à partir du promoteur 35S ; il sert de leurre à la machinerie de l'ARN interférence de la cellule (Blevins et al., 2011).

• **L'ARN 19S** (1,5 kb), qualifié d'ARN subgénomique puisque sa séquence est colinéaire à celle de l'extrémité 3' de l'ARN 35S, ne renferme que l'ORF VI et partage avec l'ARN 35S le même site de terminaison de la transcription et de polyadénylation (**figure P-2**).

• **L'ARN 35S** est un peu plus long que le génome (8,2 kb). Il s'agit d'un ARN polycistronique, avec une longue région *leader* (600 nt), qui sert à la fois de messager pour la production de toutes les protéines virales et d'intermédiaire à la transcriptase inverse virale, P5, pour la synthèse des nouveaux génomes viraux. C'est pour cette raison qu'on parle d'ARN pré-génomique (**figure P-2**). Sa taille légèrement plus grande que celle de l'ADN génomique est due à la présence de séquences répétées à ses extrémités, très importantes lors de l'étape de

rétrotranscription. Celles-ci sont générées par l'ARN polymérase II lors de la transcription du génome car le signal de terminaison de la transcription et de polyadénylation est situé en aval, à 180 nt du site d'initiation de la transcription. Le site de terminaison et polyadénylation est ignoré par la polymérase lors de son premier passage et par conséquent, cette dernière transcrit deux fois la même séquence. Le signal de polyadénylation est composé de la séquence caractéristique AAUAAA qui détermine le clivage du transcrit et sa polyadénylation. Des séquences répétées UUUGUA situées aux abords de ce signal agissent en *cis* pour stimuler l'arrêt de la transcription et la maturation en 3' des ARN 35S et 19S (Sanfaçon et al., 1991).

Après leur transcription, 70% des ARN 35S subissent un épissage alternatif alors que 30% de ces molécules échappent à la machinerie d'épissage et sont exportées vers le cytoplasme tel quel (Bouton et al., 2015a). Cet épissage s'opère par l'utilisation alternative de quatre sites donneurs, nommés D1 à D4, localisés dans l'ORF I et la région 5'UTR de l'ARN 35S et, un site accepteur A unique présent dans l'ORF II (Kiss-László et al., 1995). On pense que le rôle principal de cet épissage serait de limiter, en début de cycle, l'expression de la protéine P2 qui serait toxique à ce stade pour le virus (Froissart et al., 2004). En effet, Bouton et al., montrèrent en 2015 que l'épissage alternatif de l'ARN 35S était maintenu coûte que coûte par le virus, grâce à l'utilisation de sites d'épissage cryptiques en cas d'inactivation des sites majeurs ce qui témoigne de l'importance fonctionnelle de l'épissage pour le cycle biologique du CaMV.

2.3. Fonctions des protéines virales

Le CaMV possède 7 ORFs, mais seules 6 protéines, P1 à P6, sont produites à partir des ORF I à VI portés par l'ARN 35S polycistronique et en plus, pour la protéine P6 à partir de l'ARN monocistronique 19S. L'ORF VII peut être exprimé dans des systèmes d'expression hétérologues (Wurch et al., 1990; Lutz et al., 2012) mais sa protéine n'a jamais été détectée en contexte infectieux (Bouton et al., 2015b; Schoelz et Leisner, 2017).

- **P1, P2 et P3** sont impliquées dans le transport du CaMV. La protéine P1 (40 kDa), aussi appelée protéine de mouvement ou MP, forme des tubules au niveau des plasmodesmes qui permettent aux virions (Sánchez-Navarro et al., 2010) de passer de cellule-à-cellule afin d'atteindre le phloème et infecter systémiquement la plante hôte (Perbal et al., 1993). Pour remplir cette fonction, P1 interagit avec des protéines cellulaires (PMI7, µ-AP, PDLP1 et PME) (Chen et al., 2000; Huang et al., 2001; Sánchez-Navarro et al., 2010) et avec les protéines virales P6 et P3 (Lutz et al., 2012). P3 (15 kDa) aussi appelée VAP - *virus associated protein* - ou protéine mineure de capside, est associée aux virions *via* son extrémité C-terminale tandis que sa partie N-terminale interagit avec P1 (Sánchez-Navarro et al., 2010). En plus de son rôle dans le mouvement, P3 est aussi indispensable à la transmission du virus par le puceron (Leh et al., 1999; Khelifa et al., 2007). En effet, elle fait le pont entre le facteur assistant de transmission du virus, P2 ou FAT, et la capside virale (Leh et al., 2001). En fin d'infection, P2 forme les viroplasmes clairs

(Hoh et al., 2010), corps d'inclusion encore appelés corps de transmission car ils renferment les deux facteurs de transmission, P2 et P3, et des particules virales néoformées, en attente de transmission par le puceron (Martinière et al., 2009). P2 (18 kDa) interagit avec son extrémité N-terminale avec son récepteur, la styline 1, présente au niveau de l'acrostyle, situé à la pointe du stylet du puceron où se confondent les canaux alimentaire et salivaire (Webster et al., 2018). L'extrémité C-terminale de P2 interagit avec l'extrémité N-terminale de P3 : cet édifice permet de maintenir la nucléocapside virale au niveau de l'acrostyle de l'insecte.

• **P4** (57 kDa) (**figure P-3**) ou CP est le précurseur des protéines de capside et l'équivalent de Gag chez les rétrovirus. Ce précurseur est clivé en plusieurs étapes, concomitamment avec la formation de la capside. Ses extrémités très acides sont maturées par la protéase à acide aspartique codée par la partie 5' de l'ORF V et, certainement, par des protéases cellulaires, engendrant plusieurs polypeptides de 44, 39 et 37 kDa (Torruella et al., 1989; Champagne et Leclerc, 2006; Hohn et Rothnie, 2013). Après clivages, les polypeptides possèdent tous en C-terminal, un domaine basique et un motif en doigt de zinc, $CX_2CX_4HX_4C$, également retrouvé chez toutes les protéines de nucléocapside rétrovirales exceptés les *Spumavirus* (Chapdelaine et Hohn, 1998). Cette région permet aux polypeptides d'interagir directement et spécifiquement avec une séquence riche en purines, nommée ψ, présente dans la région *leader* de l'ARN 35S (**figure P-4**)(Guerra-Peraza et al., 2000). Ce motif est d'une importance primordiale pour le virus car il participe à l'encapsidation des nouveaux génomes viraux lors de l'étape de rétrotranscription ; sa délétion ou toute mutation empêchant son interaction avec l'ARN 35S, abolissent le cycle viral (Guerra-Peraza et al., 2000). Les protéines de capside possèdent aussi à leur extrémité N-terminale, un NLS qui n'est exposé que lors de l'assemblage de la capside au moment des clivages de P4 (Leclerc et al., 1999; Champagne et Leclerc, 2006). Cette stratégie permet au CaMV de s'assurer que P4 ne soit pas importée dans le noyau avant son assemblage, et que seuls les virions matures néoformés ou transmis par le puceron vecteur, atteignent le noyau (Champagne et Leclerc, 2006). Ainsi, les particules de CaMV matures interagissent avec l'importine α et vont s'amarrer aux NPC (Karsies et al., 2002) puis libèrent l'ADN qui pénètre dans nucléoplasme (Bouton et al., 2015b). Champagne et al., (2007) ont montré que le précurseur P4 et l'isoforme de 44 kDa de P4 sont phosphorylés au niveau de leurs extrémités par la caséine kinase II cellulaire et suggèrent que cette phosphorylation régule le ciblage nucléaire des particules virales, à la manière de ce qui se passe chez le HBV où la phosphorylation des protéines de capside permet l'exposition de leur NLS (Kann et al., 1999).

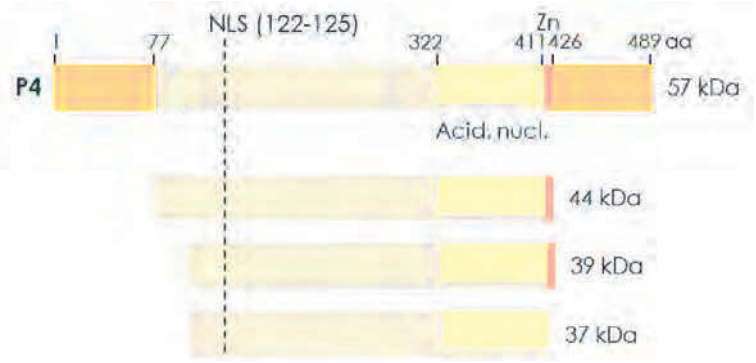


Figure P-3 | La protéine P4 de CaMV et ses produits de maturation. Le précurseur P4 des protéines de la capside (en haut) est clivé au niveau de ses extrémitées N- et C-terminales acides pour donner naissance aux différents produits p44, p39 et p37 (en bas) et qui composent la capside virale. Les domaines acides (en orange) et basique (jaune), le NLS, ainsi que le domaine en doigt de zinc (Zn) et le site d'interaction avec les acides nucléiques vous sont représentés. D'après Champagne et al., 2007.

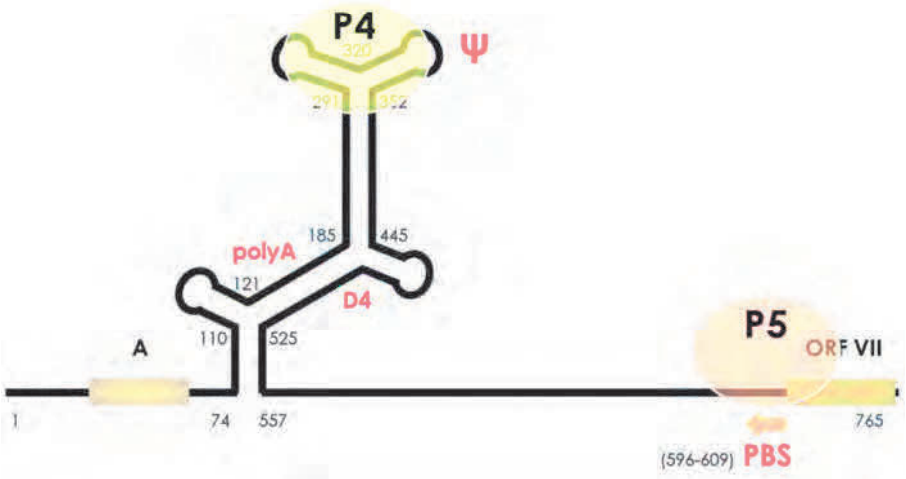


Figure P-4 | La région *leader* de l'ARN 35S de CaMV. Représentation schématique de la structure secondaire de la région *leader* (nt 1-765) obtenue grâce au logiciel MFLOD (Pooggin et al., 2001) à partir du plasmide pCaMV-GFP. La numérotation est celle de l'ARN 35S. La position du sORF A, du site de polyadénylation (nt 170-177), du signal d'encapsidation ψ reconnu par la protéine de capside P4, du site d'épissage D4 (nt 479-482), du PBS – *primer binding site* – reconnu par l'ARNt^{Met}i et par P5, et de l'ORF VII sont représentés en partant de la gauche.

• **P5** (78 kDa) (**figure P-5**) est une protéine multifonctionnelle qui joue un rôle prépondérant dans la réplication du génome viral ; c'est l'équivalent de la protéine Pol des rétrovirus. Le domaine N-terminale de P5 constitue la protéase à acide aspartique évoquée ci-dessus tandis que la région C-terminale constitue la transcriptase inverse ; cette dernière possède également une activité RNase H (RH). La protéase se libère de P5 par un clivage en *cis* (Torruella et al., 1989). L'association de ces trois activités enzymatiques PR-RT-RH est caractéristique des pararétrovirus et des rétrovirus. Toutefois, à l'inverse des rétrovirus, Pol n'a pas d'activité intégrase chez les *Caulimoviridae* et les pararétrovirus en général, car leur ADN génomique reste sous forme d'épisome, l'intégration de l'ADN viral dans le génome hôte n'étant pas une étape obligatoire pour le cycle infectieux. Les seuls cas d'intégration sont observés chez les pararétrovirus endogènes (EPRV) où l'ADN viral est intégré par recombinaison illégitime (Chabannes et Iskra-Caruana, 2013). P5 interagit avec l'ARN 35S, au niveau de sa région *leader*,

lorsqu'elle initie la rétrotranscription et lie l'ARNt^{Met} initiateur qui se fixe à l'élément PBS – *primer binding site* – de cette région (**figure P-4**). P5 est une des protéines exprimées précocement durant le cycle de CaMV (Kobayashi et al., 1998) mais son activité de transcriptase inverse ne semble pas très active à ce moment-là (Khelifa et al., 2010). C'est l'expression de P4, plus tardivement et sa maturation qui semblent activer la polymérase ce qui suggère l'existence d'un couplage entre la rétrotranscription et l'assemblage de la capside (Marsh et Guilfoyle, 1987; Himmelbach et al., 1996).

Figure P-5 | La protéine P5 du CaMV et ses différents domaines. P5 est une polyprotéine de 78 kDa qui possède trois activités enzymatiques différentes, une activité protéase à acide aspartique (PR, Asp46), une activité ADN polymérase ARN-dépendante (RT) et une activité RNase H (RH). La protéase libérée par auto-clivage intervient également dans la maturation de P4. Les domaines catalytiques des différentes enzymes vous sont indiqués en rouge. D'après Monsion, 2008.

• **P6** (62 kDa) (**figure P-6**), aussi appelée TAV – *transactivator viroplasmin* –, est une protéine très abondante qui est essentielle dans la pathogenèse du CaMV. En effet, elle joue de multiples rôles tout au long du cycle viral : dans le déterminisme du spectre d'hôte (Schoelz et al., 1986), l'expression des symptômes viraux (Yu et al., 2003), la suppression de l'ARN interférence (Love et al., 2007; Haas et al., 2008; Shivaprasad et al., 2008; Leisner et Schoelz, 2018), la morphogénèse virale (Himmelbach et al., 1996) et naturellement, son rôle le plus connu, la transactivation traductionnelle (Ryabova et al., 2006). En effet, TAV permet l'expression de toutes les protéines du CaMV à partir de l'ARN 35S grâce au mécanisme dit de réinitiation de la traduction, en maintenant les ribosomes sous une forme compétente tout le long du messager (voir **§2.2.4.**). Ses nombreuses fonctions font qu'on peut la considérer comme étant la clé de voûte du cycle infectieux de CaMV (Bouton et al., 2015b). P6 est capable d'interagir avec toutes les autres protéines du virus de même qu'avec au moins 10 facteurs d'hôte, dont les propriétés et les fonctions sont très différentes (Schoelz et Leisner, 2017). Cette protéine est une navette nucléo-cytoplasmique ; elle arbore deux NLS qui lui permettent d'emprunter la voie de l'importine α pour aller au noyau (Haas et al., 2008) et un NES riche en leucines qui est reconnu par l'exportine CRM1 et qu'elle utilise pour rejoindre le cytoplasme (Haas et al., 2005). P6 forme les usines virales de CaMV aussi appelés viroplasmes denses car ceux-ci apparaissent opaques au microscope électronique (Fujisawa et al., 1967; Martelli et Russo, 1977). La réplication de l'ADN génomique viral et la morphogenèse des nouveaux virions ont lieu dans ces corps d'inclusion (Bak et al., 2014) tandis que la synthèse des protéines virales s'effectue à leur périphérie (Martelli et Russo, 1977; Haas et al., 2002). P6 forme les viroplasmes denses, en absence de toute autre protéine virale, grâce à des interactions intermoléculaires

impliquant plusieurs domaines de P6, notamment l'extrémité N-terminale où le motif EKI joue un rôle prépondérant (Geldreich et al., 2017). En effet, lorsque ce dernier est substitué par 3 alanines, plus aucun viroplasme n'est formé dans la cellule infectée – *cette découverte à laquelle j'ai participé, a fait l'objet d'une publication que vous trouverez en Annexe 2* –.

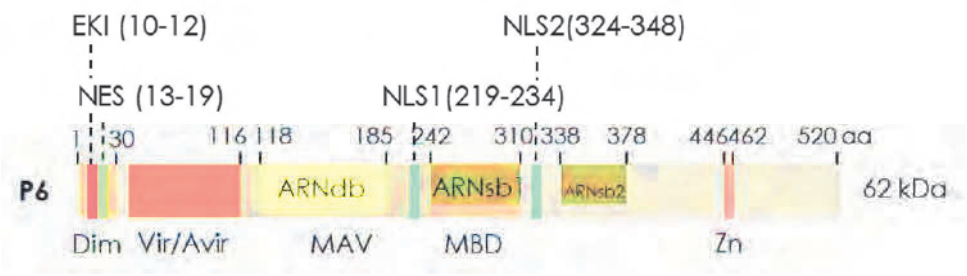


Figure P-6 | Organisation des domaines fonctionnels de la protéine P6 du CaMV. P6 possède plusieurs domaines fonctionnels dont ceux impliqués dans sa dimérisation Dim, dans la virulence/avirulence Vir/Avir, dans la transactivation traductionnel de l'ARN 35S MAV – mini TAV – et MBD – multiple binding domain – qui reconnaissent un nombre important de co-facteurs cellulaires. En plus on retrouve plusieurs motifs dont EKI (formation des viroplasmes denses), NES (reconnu par CRM1) et deux NLS. P6 possède également un activité de liaison aux ARN double brin (ARNdb) et deux motifs de liaison aux ARN simple brin (ARNsb1 et ARNsb2). En C-terminal, P6 contient un motif en doigt de zinc (Zn). D'après Bouton et al., 2015b.

2.4. Cycle de réplication du CaMV

Suite à l'introduction du CaMV dans la cellule hôte, par le puceron ou *via* les plasmodesmes à partir d'une cellule infectée, la nucléocapside est acheminée jusqu'au noyau grâce aux NLS portés par les protéines de capside et la voie de l'importine α. Le génome est ensuite libéré, probablement sous forme de complexe avec quelques monomères de protéines de capside (Schoelz et Leisner, 2017). En raison de sa taille (50 à 54 nm de diamètre) il est peu vraisemblable que la particule virale entière passe au travers des pores nucléaires, elle n'y a d'ailleurs jamais été détectée (Champagne et Leclerc, 2006). Les mécanismes impliqués dans ces premières étapes du cycle viral sont peu comprises à l'heure actuelle (**figure P-6**)(Merkle et al., 1996; Karsies et al., 2002).

Dans le noyau, l'ADN db circulaire relâché est réparé au niveau de ses interruptions de séquence par des enzymes nucléaires, générant une molécule d'ADN db circulaire covalemment fermée, puis associé aux histones cellulaires pour former un minichromosome (**figure P-6**)(Olszewski et al., 1982). Reconnu par l'ARN polymérase II, le génome est alors transcrit en deux ARN coiffés et polyadénylés, l'ARN 19S et l'ARN 35S qui est alors épissé ou non. Les ARN viraux sont ensuite exportés vers le cytoplasme selon un mécanisme inconnu (**figure P-6**).

Dans le cytoplasme, l'ARN 19S est traduit selon un mécanisme de traduction coiffe-dépendante tout à fait classique qui donne naissance à la protéine P6 (**figure P-6**). L'expression de l'ARN 35S s'effectue grâce à des stratégies d'expression non-canoniques très bien

caractérisées et impliquant largement la protéine P6 (Pooggin et Ryabova, 2018). Ces stratégies sont nécessaires en raison de la structure particulière de l'ARN 35S. En effet, s'il est coiffé en 5' et polyadénylé en 3', l'ARN 35S de CaMV est polycistronique, une caractéristique procaryotique. Il possède 7 ORF successifs impossibles à traduire pour une machinerie ribosomale eucaryote classique. De plus, son extrémité 5' non traduite est très longue (600 nt environ), jalonnée de 8 petits ORF (sORF) de 3 à 30 codons et se présente sous la forme d'une grande tige-boucle multibranchée. Juste en amont de cette structure secondaire, on trouve le premier sORF A et en aval, l'ORF VII. Cette région *leader* (**figure P-3**) est présente chez tous les membres de la famille des *Caulimoviridae* (Pooggin et al., 1999; Hohn et al., 2001). Sa conservation parmi tous les genres de cette famille indique, nonobstant la complexité de cette structure qui complique la traduction, que le *leader* est d'une importance capitale pour ces virus, – certainement dans la réplication et l'encapsidation du génome - et par conséquent, ces derniers ont élaboré une stratégie de traduction, le saut du ribosome, pour contourner cet obstacle et exprimer l'ensemble de leurs protéines (**figure P-7**)(Pooggin et Ryabova, 2018). Seule exception, le *Cestrum yellow leaf curling virus* du genre *Soymovirus,* où la traduction de l'ARN 35S s'effectue par un mécanisme (inconnu) distinct du saut du ribosome à cause de la position inhabituelle de la tige-boucle et de l'absence en 5' de cette structure du sORFA, indispensable pour le saut (Stavolone et al., 2003).

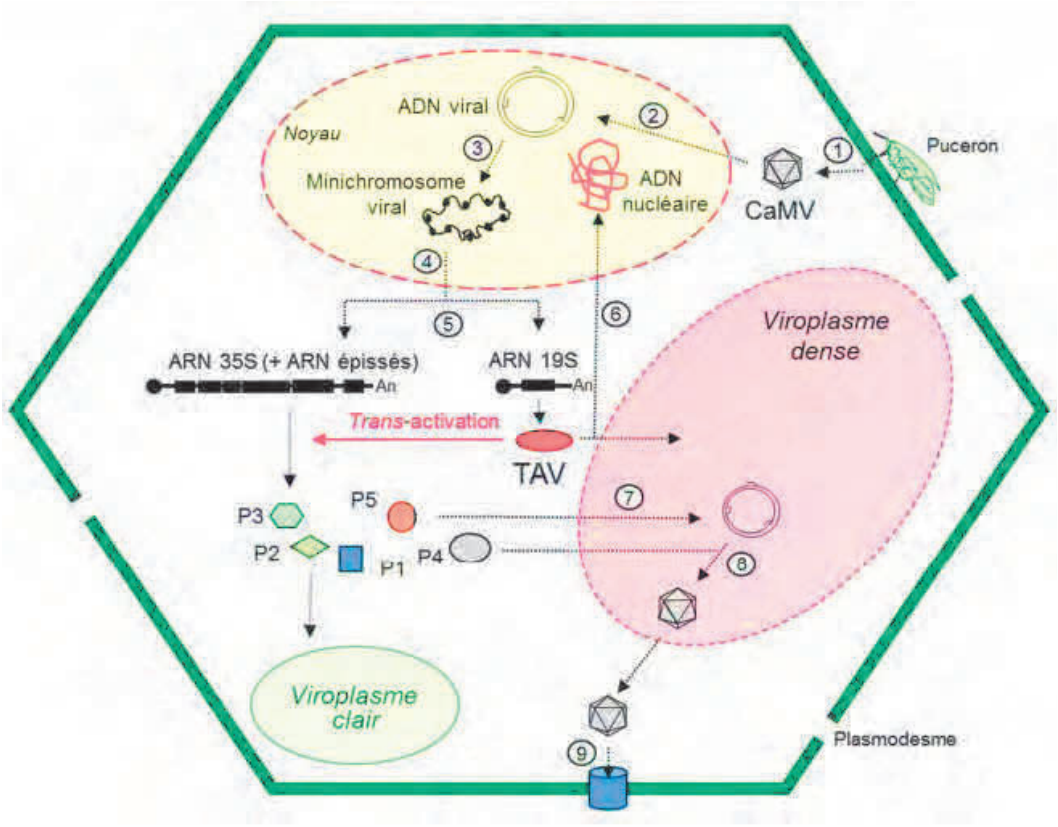


Figure P-6 | Cycle infectieux du CaMV. 1 : Transmission du virus par le puceron (*Myzus persicae*). 2 : Entrée de l'ADN viral dans le noyau. 3 : Réparation de l'ADN viral et formation du minichromosome. 4 : Transcription de l'ADN par l'ARN polymérase II cellulaire et épissage alternatif. 5 : Export nucléaire de l'ARN 19S qui est traduit en P6/TAV et, de l'ARN 35S et des ARN épissés qui sont exprimés par réinitiation de la traduction grâce à P6/TAV pour donner naissance à un jeu complet de protéines (P1, P2, P3, P4, P5 et P6/TAV). La protéine P2 s'assemble pour former un viroplasme clair. 6 : Formation des viroplasmes denses et import nucléaire de P6/TAV, nécessaire à sa fonction de suppresseur du *RNA silencing*. 7 : Réplication du génome viral par rétrotranscription de l'ARN 35S. 8 : Morphogenèse des virions dans les viroplasmes denses. 9 : Mouvement intercellulaire des virions à travers les tubules formés au niveau des plasmodesmes. 10 : Réinfection du noyau par les virions nouvellement formés. D'après (Bouton et al., 2015b).

Dans cette stratégie, les ribosomes sont recrutés sur l'ARN 35S de façon coiffe-dépendante puis traduisent le premier sORF, le sORF A composé seulement de 3 codons. La taille de cet sORF est primordiale dans le mécanisme du saut du ribosome car les ribosomes n'ont pas le temps de perdre leurs facteurs d'initiation de la traduction lorsqu'ils arrivent à son codon STOP. Comme le codon d'initiation de l'ORF VII se retrouve à proximité grâce à la structure secondaire du *leader*, les ribosomes peuvent alors immédiatement réinitier la traduction de l'ORF VII et court-circuiter ainsi la grande tige-boucle de l'ARN 35S (Pooggin et Ryabova, 2018). Ce mécanisme s'opère également chez d'autres virus comme le HBV (Sen et al., 2004; Cao et Tavis, 2011) et le PFV (Schepetilnikov et al., 2009) et pour différents ARNm cellulaires (Jackson et al., 2010). L'intervention de P6 n'est pas indispensable mais sa présence permet de potentialiser la réinitiation de la traduction en assurant probablement le maintien du facteur d'initiation eIF3 sur les ribosomes après lecture du sORF A.

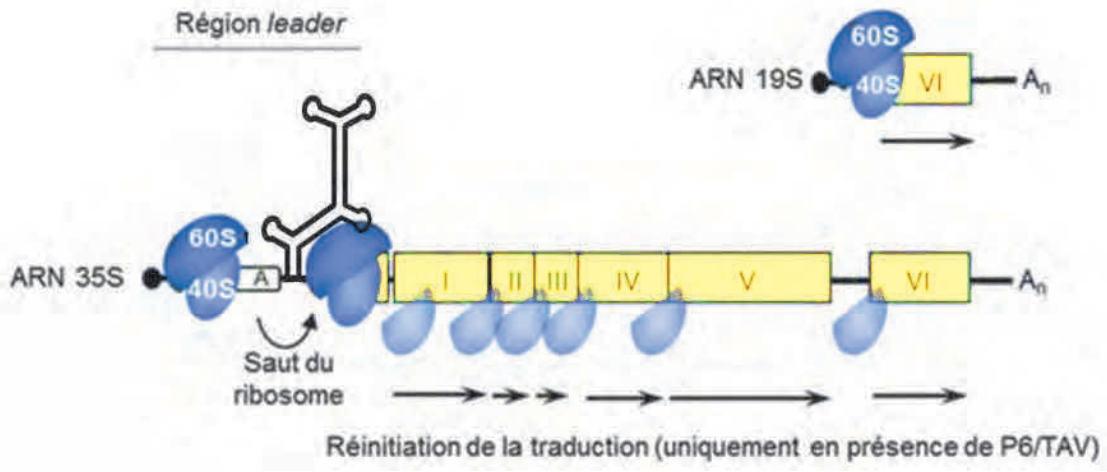


Figure P-7 | Représentation schématique de l'organisation structurale des ARN 35S et 19S, et des deux stratégies de traduction non-canoniques mises en œuvre pour exprimer l'ARN 35S, le saut du ribosome et la réinitiation de la traduction. La boite A représente le sORF A (3 codons), dont la traduction complète est requise pour qu'il y ait saut du ribosome. La structure secondaire de la région *leader* renferme les autres sORF (non représentés). L'ORF VII est caché par le ribosome 80S. Modifiée d'après Bouton et al., 2015b.

P6 est en revanche absolument nécessaire au mécanisme de réinitiation de la traduction des longs ORF (I à VI) de l'ARN 35S (**figure P-7**). Contrairement au mécanisme de réinitiation entre le sORF A et l'ORF VII, la réinitiation après traduction d'un ORF long nécessite le maintien des facteurs d'initiation de la traduction sur les ribosomes au cours de la phase d'élongation de la traduction. Ce rôle est rempli par la protéine P6 qui interagit avec différents composants de la machinerie traductionnelle cellulaire. P6 se lie directement à la grande sous-unité 60S en s'associant avec les protéines ribosomales L13, L18 et L24 (Leh et al., 2000; Park et al., 2001; Bureau et al., 2004). En plus, elle interagit avec la sous-unité g du facteur eIF3, la protéine cellulaire RISP – *reinitiation stimulating protein* – (Thiébeauld et al., 2009) et la protéine kinase TOR – *target of rapamycin* – (Schepetilnikov et al., 2011) qu'elle active (Schepetilnikov et Ryabova, 2014). Ce réseau très complexe agencé par P6 et la phosphorylation de plusieurs de ces facteurs par TOR facilitent la traduction de tous les ORF de l'ARN 35S, hormis apparemment l'ORF VII, en maintenant artificiellement les deux sous-unités ribosomiques sur cet ARN polycistronique (**figure P-8**).

L'étude de la cinétique d'expression des protéines du CaMV en protoplastes de navet montre qu'elles ne sont pas toutes exprimées en même temps, durant le cycle infectieux. Ainsi les premières protéines virales à s'accumuler sont P6 qui est principalement exprimée à partir de son propre messager, l'ARN 19S, P5 et P1, puis P3, et finalement P4 et P2 (Kobayashi et al., 1998; Khelifa et al., 2010). L'épissage de l'ARN 35S qui place les longs ORFs plus ou moins près de l'extrémité 5' coiffée de cet ARN permettrait à CaMV d'exprimer avec plus ou moins d'efficacité ses protéines par le complexe réinitiateur (Schoelz et Leisner, 2017).

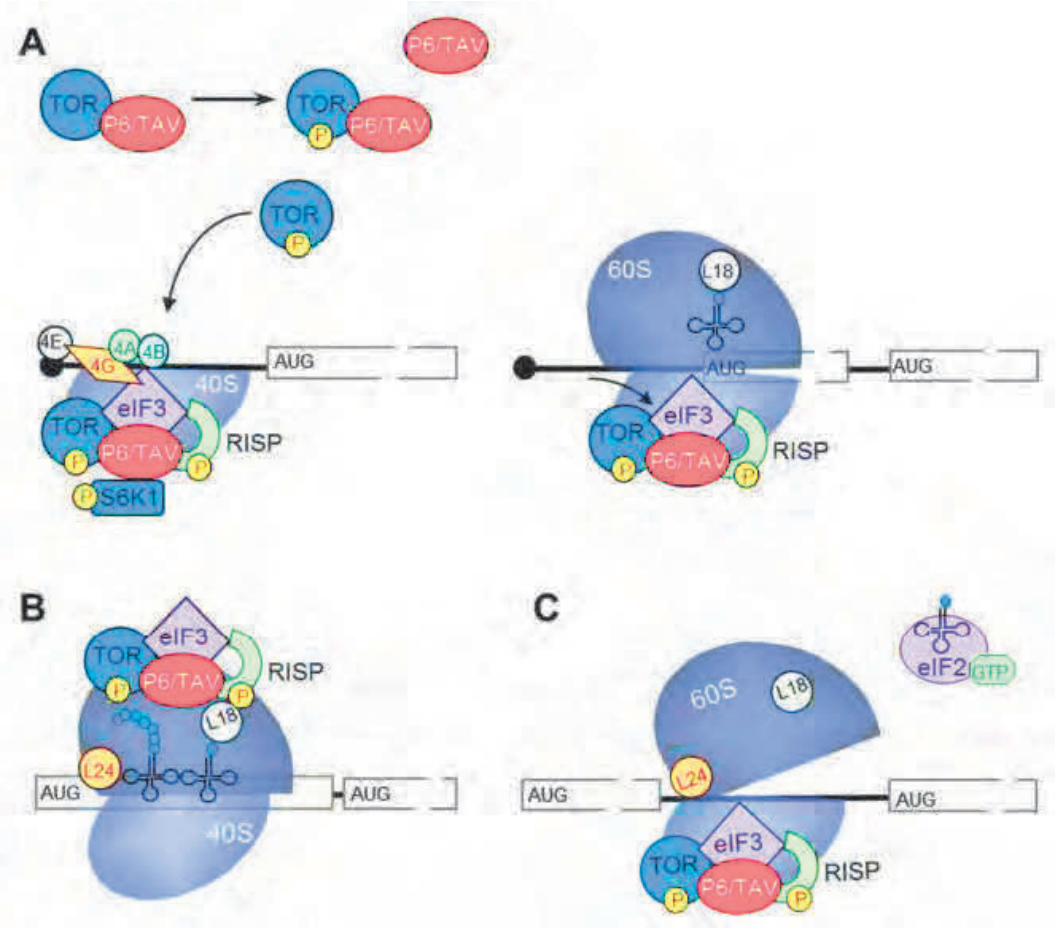


Figure P-8 | Modèle hypothétique du mécanisme de la réinitiation de la traduction des ARNm polycistroniques, grâce à la protéine P6/TAV. (**A**) La protéine kinase TOR activée par P6/TAV se fixe sur le complexe de préinitiation (PIC 43S), vraisemblablement en même temps que eIF3 et RISP et ce, avant la jonction de la sous unité 60S du ribosome. TOR phosphoryle alors ses substrats RISP et S6K1, qui est relarguée de la sous-unité 40S. Le mécanisme impliqué dans la phosphorylation de TOR est pour l'instant non élucidé. (**B**) Au cours de l'élongation, le complexe eIF3-P6/TAV-RISP-TOR est transloqué sur la sous-unité 60S puis, (**C**) avant la terminaison de la traduction de l'ORF, est ramené au niveau de la sous-unité 40S afin de la rendre compétente pour une nouvelle initiation de la traduction. D'après Bouton et al., 2015b.

En fin de cycle, l'ARN 35S sert aussi d'intermédiaire matriciel pour la transcriptase inverse (P5) pour la synthèse nouveaux ADN génomiques viraux (**figure P-6**). Bien que découvert, il y a de nombreuses années (Pfeiffer et Hohn, 1983) le mécanisme exact de réplication du génome de CaMV n'a pas encore été entièrement élucidé. Nos connaissances à ce sujet viennent des travaux menés par Hull et al. (1987), qui ont montré que l'ARNt^{Met} initiateur cellulaire sert d'amorce, et de celles acquises chez les rétrovirus (Hu et Hughes, 2012). Il est fort probable que la rétrotranscription soit couplée à l'étape d'encapsidation de l'ADN génomique (Thomas et al., 1985; Marsh et Guilfoyle, 1987), comme c'est le cas chez les *Hepadnaviridae* (Bruss, 2007) car la protéine de capside P4 interagit avec le sommet de la tige-boucle – site ψ - de la région *leader* de l'ARN 35S (Guerra-Peraza et al., 2000) ce qui suggère que la capside commence à s'assembler dès que démarre la synthèse des nouveaux génomes viraux. L'incapacité de P4 et ses isoformes à former des pseudo-virions en système d'expression hétérologue (Chapdelaine et

Hohn, 1998), contrairement à d'autres virus, serait aussi un argument en faveur de ce couplage, l'assemblage nécessitant la présence du génome viral. Toutefois, il n'est pas exclu que la formation de la capside requiert l'intervention d'un ou d'autre(s) facteur(s) du CaMV comme P6 ou P3 (Hohn et Rothnie, 2013).

Les nouveaux viraux sont stockés principalement dans les usines virales à savoir les viroplasmes denses formés par P6 et accessoirement dans les viroplasmes clairs (**figure P-6**), en attente d'être prélevés par le puceron pour être transmis à d'autres plantes hôtes (Martinière et al., 2009; Bak et al., 2013). Le mécanisme exact permettant aux virions de passer des usines virales jusqu'aux tubules formés par la protéine de mouvement P1 au sein des plasmodesmes, n'est pas connu (Rodriguez et al., 2014; Schoelz et Leisner, 2017). Les viroplasmes denses seraient directement impliqués dans l'acheminement des particules virales, le long des filaments d'actine jusqu'aux plasmodesmes (Schoelz et Leisner, 2017). P6 pourrait être engagée dans ce processus (Martinière et al., 2013) étant donné qu'elle interagit avec le cytosquelette d'actine (Harries et al., 2009), les protéines de capside P3 et P4, la protéine de mouvement P1 ainsi qu'avec la protéine cellulaire AtSRC2 – *Arabidopsis thaliana soybean response to cold* –, une protéine qui co-localise avec PDLP1 – plasmodesmata located protein -. Les virions traversent les tubules des plasmodesmes par le biais d'interactions entre la protéine P3 associée à la capside et la protéine P1 et envahissent les cellules adjacentes pour démarrer un nouveau cycle infecteux (Stavolone et al., 2005).

3. Projet de thèse

L'étude de l'export nucléaire des ARN du CaMV a été abordée au cours de ma thèse, en considérant trois questions qui nous paraissent essentielles :

- Quelle voie est empruntée par le CaMV pour exporter ses ARN 35S ? S'agit-il de la même voie pour les ARN 35S épissés ou non ? La voie XPO1 (CRM1) est–elle impliquée dans ce transport, comme chez certains rétrovirus ?

- L'export est–il régi par les facteurs d'export d'origine cellulaire et/ou virale ?

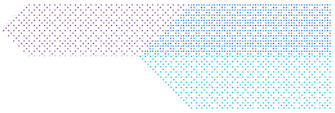
- Quelles sont les séquences en *cis* voire les structures secondaires qui gouvernent leur export ?

Afin d'apporter une réponse à ces différentes questions, nous avons mis en œuvre des expériences *in planta* sur des mutants d'Arabidopsis n'exprimant pas certains facteurs d'export et quantifié l'export nucléaire des ARN 35S du CaMV *ex vivo* dans des protoplastes préparés à partir de ces plantes mutantes ou encore, testé l'effet de la leptomycine B, l'inhibiteur de CRM1, sur le transport des transcrits viraux.

Nous avons également transfecté des protoplastes d'Arabidopsis avec des génomes mutés de CaMV n'exprimant plus toutes ou certaines protéines virales puis mesuré l'effet de l'absence de ces protéines sur le transport des ARN du virus. L'interaction des protéines virales avec des facteurs d'export a également été testée dans des expériences de GST *pull down* pour identifier les facteurs du CaMV pouvant être impliqués dans l'export nucléaire. Enfin, des études de microscopie à épifluorescence ont été entreprises pour suivre le comportement intracellulaire des protéines virales pouvant potentiellement jouer un rôle dans ce processus.

Pour finir, des expériences de compétition dans des protoplastes d'Arabidopsis ont été réalisées pour cartographier les régions de l'ARN 35S importantes pour l'export.

Résultats



Résultats

À mon arrivée dans l'équipe, la thématique sur l'export nucléaire des ARN de CaMV venait de débuter et quasiment aucun des outils nécessaires à cette étude n'était alors disponible au laboratoire. Clément Bouton qui avait initié cette dernière au cours de sa thèse (Bouton, 2014) se heurtait à un obstacle majeur : la difficulté de travailler avec des protoplastes de plantes. En effet, si l'étude de l'export nucléaire des ARNm chez les métazoaires a été largement facilitée grâce à la disponibilité et l'utilisation routinière de très nombreuses lignées cellulaires animales variées et facilement transfectables, malheureusement, chez les plantes, nous manquons cruellement d'un système cellulaire équivalent. Pourtant, ce genre d'outil biologique est plus qu'indispensable pour la réalisation d'un tel projet et représente la façon la plus simple d'aborder et d'isoler un mécanisme moléculaire tel que l'export. Ce qui explique en partie aussi pourquoi cette thématique est si peu étudiée et par conséquent si mal caractérisée pour les plantes et leurs virus. Aussi, avons-nous dû dans un premier temps consacrer beaucoup de temps et d'énergie, à la fois, à sélectionner le matériel végétal dont nous avions besoin mais surtout, à mettre au point le protocole expérimental adéquat afin d'obtenir des protoplastes performants c'est-à-dire pouvant être transfectés par différents plasmides à des taux suffisamment élevés et de façon reproductible pour détecter les ARN de CaMV et déterminer comment se passe leur export.

Ci-dessous, nous présentons d'abord les résultats des travaux expérimentaux entrepris pour nous doter de ces outils moléculaires et cellulaires. Puis nous présenterons sous forme d'un article à soumettre prochainement, les résultats essentiels obtenus sur l'export des ARN 35S du CaMV. Nous terminerons par des résultats additionnels tout aussi intéressants.

1. Les outils biologiques pour l'étude de l'export nucléaire des transcrits du CaMV

1.1. Sélection d'*Arabidopsis thaliana* mutants n'exprimant pas certains facteurs d'export

Avant et durant ma thèse, des orthologues des protéines du complexe TREX-1, ALY (Pfaff et al., 2018), MOS11 (Germain et al., 2010) et TEX1 (Yelina et al., 2010b) ou encore de l'exportine CRM1, XPO1a et XPO1b (Blanvillain et al., 2008; Merkle, 2011) ont été identifiés et décrits comme impliquées dans l'export chez Arabidopsis (Haasen et al., 1999; Stuger et Forreiter, 2004; Ehrnsberger et al., 2019a). Par chance, parmi les collections de mutants de cette plante générées au début des années 2000 (Samson et al., 2002; Sessions et al., 2002; Alonso et al., 2003; Rosso et al., 2003; Woody et al., 2007) se trouvaient des mutants n'exprimant pas ces protéines, et ce, grâce à l'insertion (aléatoire) d'un ADN de transfert (ADN-T) d'*Agrobacterium tumefaciens* modifié et de séquence connue dans leurs gènes (**figure R-1**).

Nous nous sommes procurés des graines de ces plantes mutantes, soit auprès de l'ABRC™ (*Arabidopsis Biological Ressources Center*) de l'Université de l'état de l'Ohio aux Etats-Unis pour les mutants *mos11-2*, *tex1-4*, *xpo1a-3* et *xpo1b-1* soit auprès du Pr. Klaus Grasser de l'Université de Regensburg en Allemagne pour les mutants de ALY (*aly1 aly2*, *aly3 aly4* et *4xaly*) (Pfaff et al., 2018) et le double mutant de TEX1 et MOS11 (*tex1-4/mos11-2*) (Sørensen et al., 2017), dans le but de déterminer si l'une ou plusieurs de ces protéines joue(nt) un rôle dans l'export des ARN du CaMV et donc quelle(s) voie(s) de TREX-1 ou CRM1 est (sont) utilisée(s) par le virus. Nous parlerons par défaut de voie de TREX-1, l'exportine associée à cette voie n'ayant pas encore été identifiée chez les plantes. Pour cela, des expériences d'infection de plantes par le CaMV ont été réalisées afin d'observer si l'absence de ces protéines cellulaires pouvaient entraîner un retard, voire une résistance de la plante au virus ou éventuellement entraîner une modification ou une atténuation des symptômes viraux. Parallèlement, nous avons également mesuré directement l'effet de ces mutations sur l'export des ARN 35S en protoplastes.

Avant de commencer ces expériences, il était cependant indispensable de confirmer le caractère homozygote de ces mutants. Pour cela, un génotypage a été réalisé sur des extraits de feuilles. Cette méthode repose sur la réalisation de 3 réactions de PCR qui permettent de déterminer si l'ADN-T de d'Agrobacterium s'est bien inséré dans les deux allèles du gène d'intérêt et dont la méthode est décrite dans la partie **Méthodes** en **5.7** (**figure R-1**).

Nous avons ainsi pu isoler plusieurs lignées homozygotes mutées au niveau des 2 allèles des gènes *XPO1a* (*xpo1a-3*) et *XPO1b* (*xpo1b-1*) pour la voie de CRM1 mais aussi *MOS11* (*mos11-2*), *TEX1* (*tex1-4*) et des différents gènes *ALY* (*aly1 aly2*, *aly3 aly4* et *4xaly*) pour la voie de TREX-1. Les résultats du génotypage de ces plantes ainsi que les cartes génétiques indiquant l'insertion des ADN-T dans ces différents gènes sont indiqués dans la **figure R-1**. À noter que le génotype des graines d'*A. thaliana* double mutant *tex1-4/mos11-2* fournies par notre collaborateur le Prof. Klaus Grasser, avait déjà été établi par son équipe.

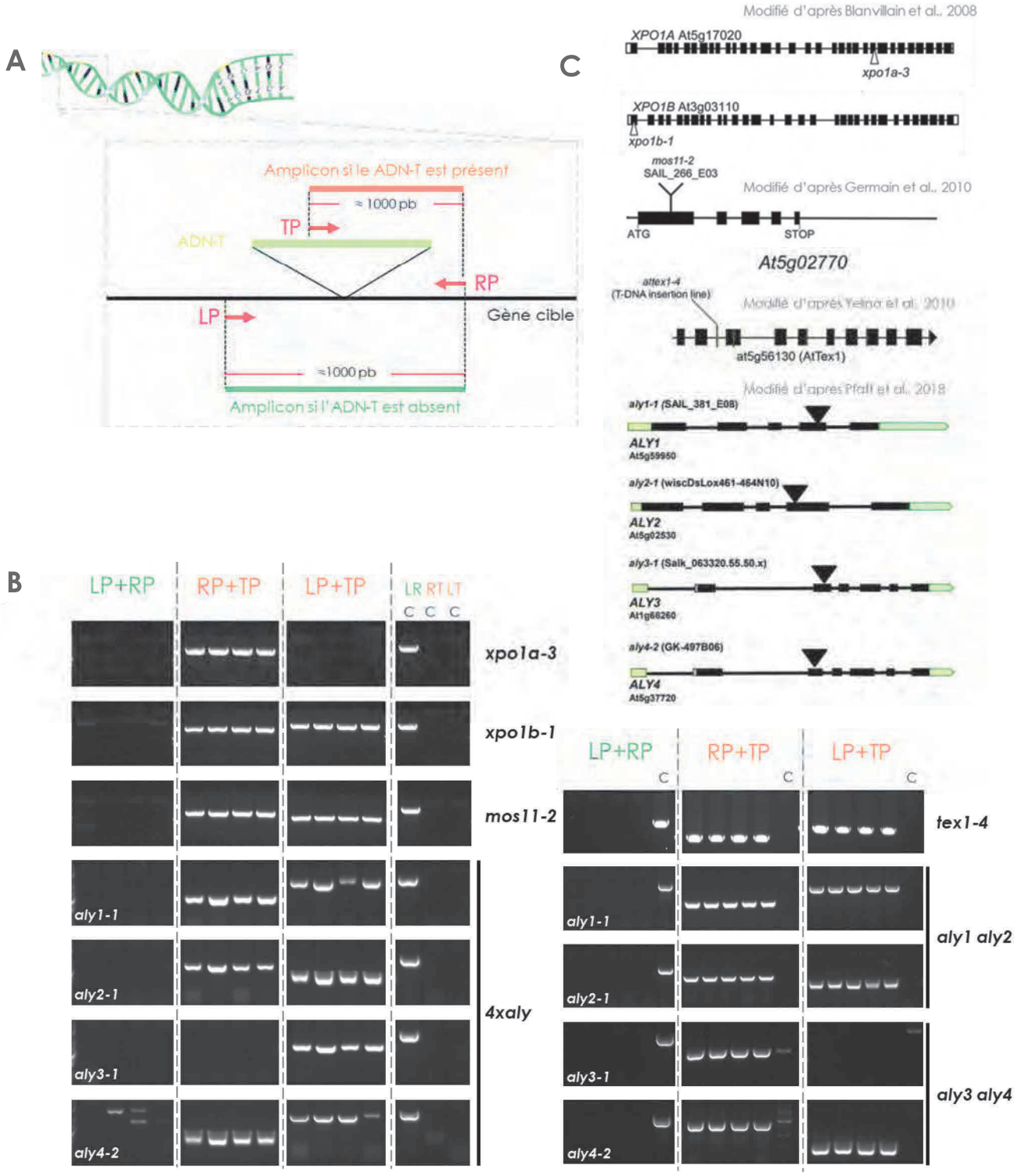


Figure R-1 | Génotypage par PCR des mutants *xpo1a-3*, *xpo1b-1*, *mos11-2*, *tex1-4*, *aly1 aly2*, *aly3 aly4* et *4xaly*. Les amplicons résultent de réactions de PCR sur la région d'insertion de l'ADN-T dans le gène à muter. Les amorces RP, LP et TP situées, respectivement, de part et d'autre de la région d'insertion et dans l'ADN-T ont été utilisées (**A**). Les produits d'amplification d'environ 1 kbp ont été séparés sur gel d'agarose 1% et révélés sous UV grâce à l'addition dans le gel de bromure d'éthidium (**B**) Une représentation schématique des différents gènes indique la zone d'insertion des ADN-T (**C**). C = *Arabidopsis thaliana* Col-0 sauvage.

Les mutants d'Arabidopsis ont été inoculés avec un extrait brut de feuilles d'Arabidopsis infectées par le CaMV, isolat Cabb B-JI, afin de suivre l'infection virale et la symptomatologie. La **figure R-2** montre à la fois le phénotype de ces plantes mutantes non inoculées, qui est

identique à celui des plantes sauvages (WT) excepté pour les mutants *tex1-4/mos11-2* et *4xaly* comme cela a été décrit respectivement par Sørensen et al. (2017) et Pfaff et al. (2018) et, les symptômes développés par ces plantes lorsqu'elles sont infectées par le virus.

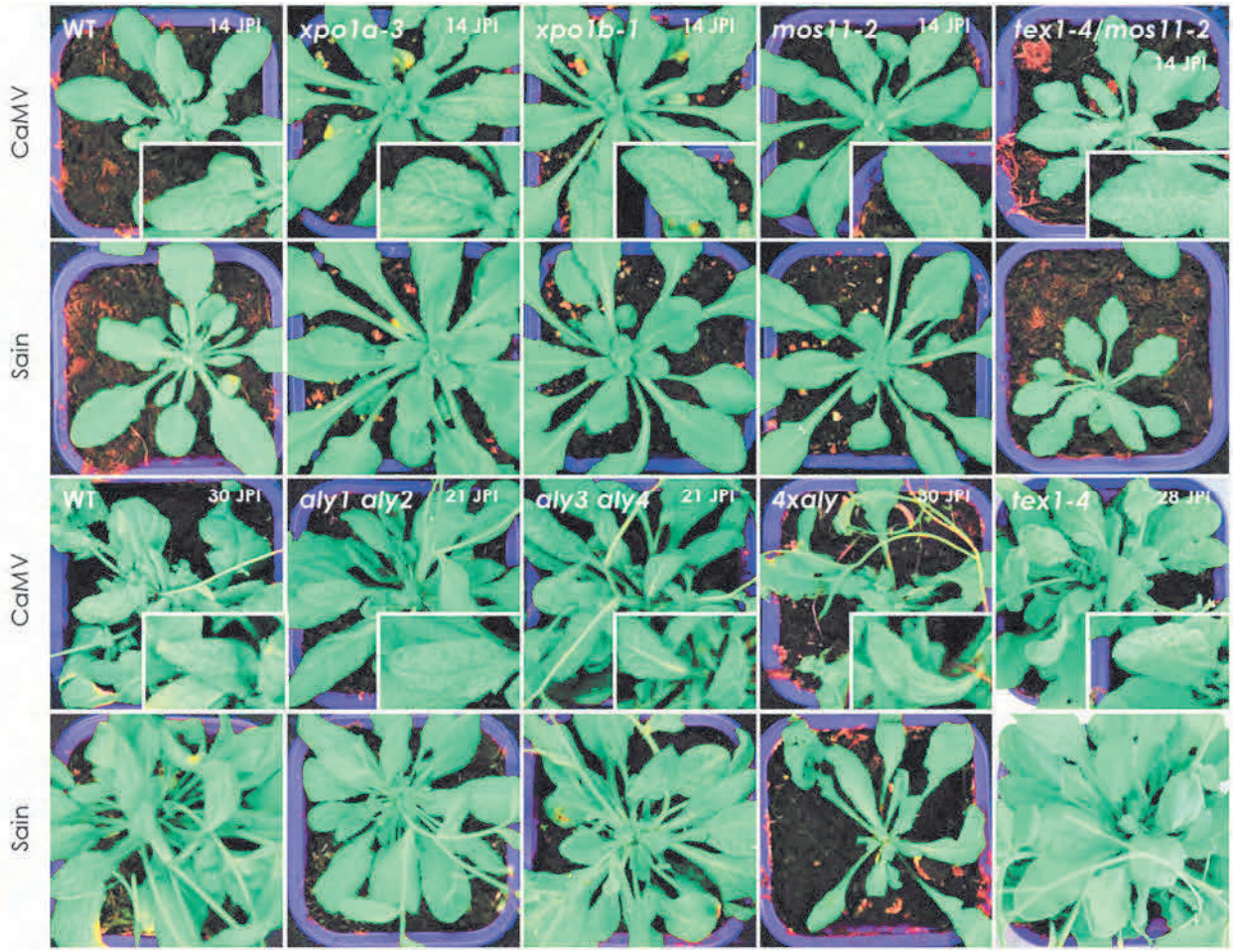


Figure R-2 | Symptômes observés sur les feuilles d'*A. thaliana* Col-0 et ses mutants infectées par le CaMV. Les plantes sauvage Col-0 (WT) et mutantes pour les gènes *XPO1a* (AT5G17020), *XPO1b* (AT3G03110), *MOS11* (AT5G02770), *TEX1* (AT5G56130) et *ALY(1-4)* (AT5G59950, AT5G02530, AT1G66260, AT5G37720), ont été inoculées au stade 4 feuilles de la rosette en frottant 3 de leurs feuilles avec un extrait brut de feuilles d'Arabidopsis infectées par l'isolat virulent de CaMV Cabb B-JI. Les symptômes viraux caractéristiques (chlorose, éclaircissement des nervures et/ou gaufrage) sont observés sur les plantes au minimum 10 jours après leur inoculation (JPI). Les stades auxquels les photos ont été prises sont donnés à côté du nom de chaque plante. Un agrandissement (en bas à droite des cases) montre en détail, les symptômes foliaires des plantes infectées.

Si un retard ou une résistance partielle ont été observés pour certains mutants (*mos11-2*, *tex1-4/mos11-2*, *aly1 aly2*, *aly3 aly4* et *4xaly*), comme le montre les résultats présentés dans l'article ci-dessous, aucune différence dans la sévérité des symptômes caractéristiques de CaMV (mosaïque, éclaircissement des nervures, gaufrage ou encore nanisme) n'a été observée pour aucune de ces plantes à partir du début de la virose. Les mutants simples *tex1-4*, *xpo1a-3* et *xpo1b-1* n'ont quant à eux pas montré de retard, ni de résistance partielle au cours des différentes expériences d'infections que nous avons réalisées (résultats non présentés). Vraisemblablement, l'absence de ces protéines peut être compensée, pour XPO1a et XPO1b

par la présence de l'autre isoforme de l'exportine 1 et pour TEX1 par les autres protéines du complexe multi-protéique TREX-1, comme cela avait été démontré pour les ARNm cellulaires (Sørensen et al., 2017). Par ailleurs, TEX1 étant également un facteur d'épissage, nous avons vérifié par RT-PCR le profil des ARN 35S épissés produit dans le mutant *tex-1-4*, en ciblant la région épissée des ARN 35S (amorces : **35S_(1-21)_FW** et **35S_(2280-2304)_REV**, voir **§ Matériel tableau MM-6**). Même en l'absence du facteur d'épissage, le profil obtenu est strictement identique à celui observé chez les plantes sauvages infectées, suggérant que dans ce mutant, l'épissage des ARN 35S, tout comme l'export des ARN viraux, ne sont pas affectés (résultats non présentés).

1.2. Mise au point du système d'étude cellulaire

1.2.1. Préparation et transfection des protoplastes d'Arabidopsis

Les protoplastes (**figure R-3**) sont des cellules végétales débarrassées expérimentalement de la paroi pecto-cellulosique par digestion enzymatique. Pour les obtenir et les maintenir en vie, ceux-ci sont préparés et maintenus dans un milieu iso-osmotique qui empêche les cellules d'éclater. Comme elles n'ont plus de paroi, ces cellules peuvent recevoir du matériel génétique étranger, ADN ou ARN ou même des virus. Historiquement, la façon d'obtenir des protoplastes était de placer un morceau de tissus végétal stérilisé dans un milieu légèrement hypertonique (additionné de saccharose ou de mannitol à 0,5 M) et de le dilacérer à l'aide d'un scalpel (Cocking, 1960). D'autres techniques permettent de préparer les protoplastes à partir de cals ou encore de cellules maintenues en suspension par agitation d'une culture liquide (Davey et al., 2005). Après plusieurs heures de digestion par la pectinase pour séparer les cellules, et la cellulase pour digérer leur paroi, et sous agitation, les protoplastes s'échappent du tissu digéré et sont récupérés après plusieurs lavages. Comme pour les bactéries, ces cellules peuvent être transformées de plusieurs manières en exerçant sur la membrane plasmique soit une force chimique, soit une force physique ou les deux (Davey et al., 2005). Néanmoins, les protoplastes restent des cellules très fragiles en raison de l'absence de leur paroi et doivent donc être manipulés très précautionneusement, ce qui rend cette étape particulièrement difficile et laborieuse.

Notre protocole s'inspire largement du procédé publié par Yoo et al., en 2007 qui est basé sur la transfection de protoplastes grâce à une solution de PEG et de calcium, une façon devenue classique pour transformer ces cellules végétales. Nous y avons néanmoins apporté quelques modifications afin de l'améliorer et de l'adapter à notre étude, car les protoplastes devaient notamment être transfectés par un plasmide viral (pCaMV-GFP) de très grande taille, 14 kpb. En effet s'il est plutôt aisé d'introduire de l'ARN (Rao, 2007) ou de petits plasmides

(Schepetilnikov et al., 2011), la transfection d'un plasmide de cette taille est inhabituelle et ardue.

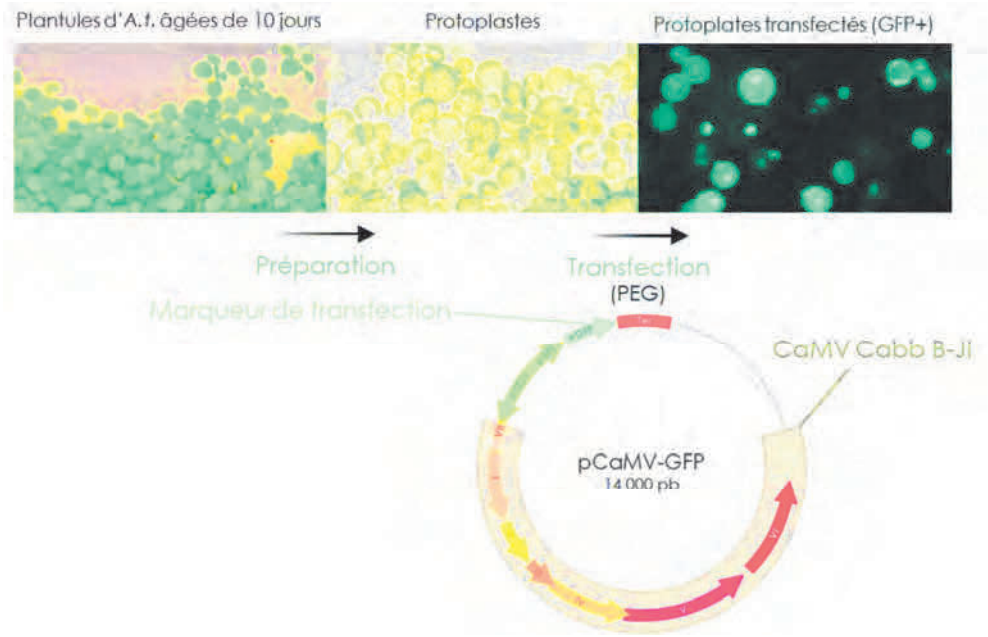


Figure R-3 | Représentation schématique de notre modèle cellulaire de protoplastes. Des protoplastes de mésophylle sont préparés à partir de plantules d'*Arabidopsis thaliana* âgées de 10 jours puis transfectés avec le plasmide viral pCaMV-GFP (Khelifa et al., 2010). Vingt heures après transfection, l'expression du marqueur de transfection eGFP est déterminée sous microscope à épi-fluorescence.

Après avoir essayé d'utiliser des protoplastes préparés à partir de différents matériaux végétaux (feuilles adultes ou encore suspension cellulaire), notre système actuel est basé sur la préparation de protoplastes à partir de plantules d'*A. thaliana* âgées de 7 à 12 jours et cultivées *in vitro* (**figure R-3**). Ce choix permet de gommer légèrement l'effet que semblent avoir les saisons sur ces cellules et qui affecte grandement la reproductibilité des expériences. Même *in vitro*, une sorte d'horloge saisonnière semble persister (Davey et al., 2005) chez les protoplastes plus performants, c'est-à-dire facilement transfectables et de façon reproductible, de l'hiver au printemps, c'est-à-dire de fin novembre à début mai (Keskitalo, 2001).

Nous avons également observé qu'augmenter le temps d'incubation des plantules dans le milieu de digestion améliorait spécialement la capacité des protoplastes à recevoir l'ADN exogène. Il est probable qu'en permettant aux enzymes d'enlever beaucoup plus de paroi, mais sans trop fragiliser les cellules, nous facilitons l'entrée des molécules d'ADN plus grandes. Ainsi le temps de digestion a été augmenté de 4 à 14h. Il fallût néanmoins adapter ce temps en fonction du type de plantules (sauvage ou mutant) utilisé dans notre étude. Alors qu'une digestion longue s'est avéré nécessaire pour transfecter efficacement des protoplastes sauvages ou le mutant *mos11-2* (dont les feuilles sont particulièrement épaisses), pour le quadruple mutant *4xaly*, très affecté par l'absence des protéines ALY, (Pfaff et al., 2018), cette digestion fût ramenée à 7h maximum sous peine d'obtenir des protoplastes trop fragiles et qui

se lysaient après le temps d'incubation nécessaire à l'expression des gènes viraux. Parallèlement, l'utilisation d'un ADN entraîneur, ici le plasmide bactérien Litmus28i, a aussi nettement amélioré l'efficacité de transfection et sa reproductibilité.

Enfin, le recours à un mutant d'Arabidopsis appelé *CA-rop2* nous a également très fortement aidé. Ce mutant exprime de façon constitutive la forme active de la petite GTPase cytoplasmique ROP2 pour – *Rho-related GTPase from plants* – (Li et al., 2001). Les protoplastes préparés à partir de ces plantules se sont avérés être bien plus facilement transfectables et ce, de façon reproductible, que ceux obtenus à partir des Arabidopsis sauvages. Il est vraisemblable que des réarrangements du cytosquelette d'actine induit par cette mutation dans ces cellules, contribue à la qualité des protoplastes. En effet les Rho GTPases sont des commutateurs moléculaires très conservés au sein des eucaryotes qui contrôlent l'organisation du cytosquelette d'actine (Fu et al., 2002; Li et Yang, 2000). Des remaniements du cytosquelette au sein de ces cellules pourraient donc expliquer pourquoi ces protoplastes semblent acquérir plus facilement l'ADN exogène que les protoplastes sauvages, même si cela reste à confirmer. Si à notre connaissance il n'a jamais été démontré que les GTPases ROP pouvaient influencer l'export nucléaire des ARNm, il nous fallait toutefois vérifier que l'activation constitutive de ROP2 n'est pas d'effet sur l'export des ARN viraux. En effet nous avons montré que ROP2 une fois active était capable d'activer elle-même TOR –*target of rapamycin* – (Schepetilnikov et al., 2017), un complexe kinase également contrôlé par le CaMV *via* sa protéine P6 et qui permet au virus de stimuler la réinitiation de la traduction de son ARN 35S (Schepetilnikov et al., 2011).

Aucune différence significative, ni dans l'accumulation des ARNm viraux, ni dans leur export n'a été observée entre les protoplastes sauvages et ces protoplastes mutants (**figure R-4**).

On notera également que les valeurs de Ct mesurées en RT-qPCR pour les 4 ARNm cellulaires (*TIP41*, *GAPDH*, *ACT2* et *SAND*), ciblés pour normaliser nos résultats et testés à la fois dans les protoplastes sauvage d'Arabidopsis Col-0 et le mutant *CA-rop2* lors de la mise au point de ce protocole, ne varient pas de manière significative (résultats non présentés).

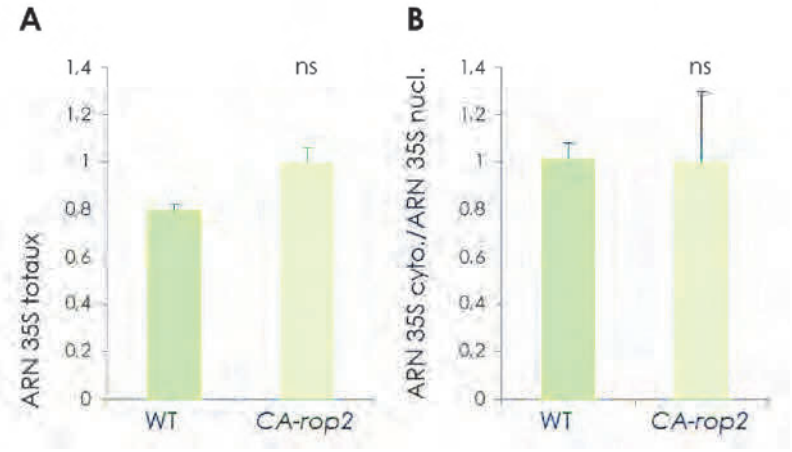


Figure R-4 | Analyse du comportement des protoplastes mutants *CA-ROP2* par rapport aux protoplastes sauvages transfectés avec pCaMV-GFP. Quantification relative par RT-qPCR des ARN 35S (épissés et non épissé ; les ARN ont été ciblé au niveau de l'ORF V) de CaMV dans les fractions cytoplasmiques, nucléaires et totales des protoplastes sauvages (WT) ou mutants *CA-rop2* et transfectés avec pCaMV-GFP. Les valeurs correspondant aux ARN 35S ont été normalisées par rapport à celles obtenues pour les ARN messagers contrôles cellulaires GAPDH et de transfection eGFP (ΔCt). Afin de déterminer si la mutation a un effet sur l'export des ARN viraux, les ratios ARN 35S cytoplasmiques/nucléaires sont calculés dans les deux conditions. Les valeurs obtenues avec les protoplastes sauvage (WT) sont fixées à 1 puis celles des protoplastes *CA-rop2* calculées en fonction (ΔΔCt). Les barres d'erreurs indiquent les écart-types observés entre les différents échantillons. Trois expériences indépendantes ont été réalisées, dont chacune en triplicat, eux même analysés en 3 réplicats techniques. La significativité a été vérifiée par test de Student (ns : non significatif ; p>0,05).

Enfin, les cinétiques d'infection de ces deux types de plantes par le CaMV, suivies par détection de l'ADN et des protéines virales dans les feuilles systémiques, sont similaires (**figure R-5**).

Ainsi, nous avons utilisé pour la préparation de protoplastes, en plus des plantules d'Arabidopsis sauvage Col-0 (difficiles à transfecter de façon reproductible), des plantules *CA-rop2* (protoplastes plus faciles à transfecter).

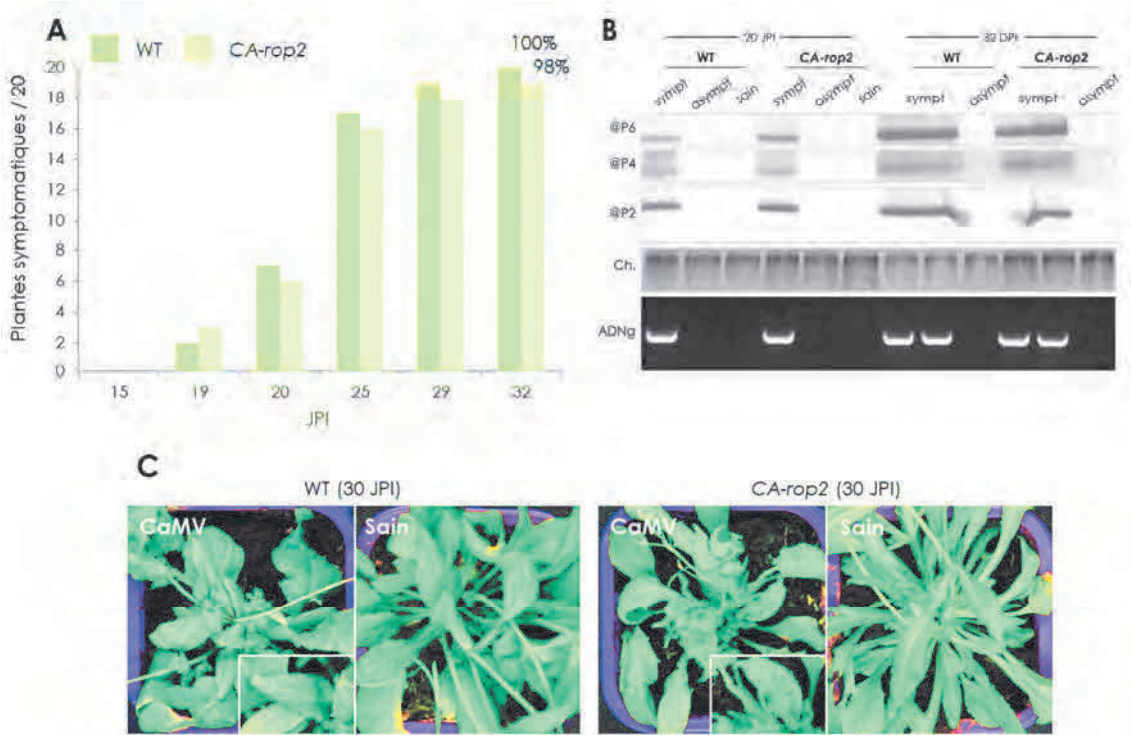


Figure R-5 | Suivi de l'infection des plants d'*A.thaliana* Col-0 sauvage (WT) ou mutants *CA-rop2* infectées par CaMV. L'apparition des symptômes caractéristiques (chlorose, éclaircissement des nervures et/ou gaufrage) de CaMV a été suivie durant plusieurs jours post-inoculation (JPI) sur 20 plants d'Arabidopsis sauvages (WT) ou mutants *CA-rop2* (**A**). En parallèle, les protéines et l'ADN obtenus à partir d'échantillons de feuilles symptomatiques (sympt), asymptomatiques (asympt) ou saines prélevés sur les plantes sauvages (WT) ou mutantes (*CA-rop2*) ont été analysés (**B**). Les protéines ont été séparées par électrophorèse sur gel de polyacrylamide 15% en conditions dénaturantes, puis révélées par coloration au bleu de Coomassie™ (Ch.) ou transférées sur membrane. Les protéines virales P6 (@P6), P4 (@P4) ou P2 (@P2) ont été recherchées dans les échantillons par western blot, à l'aide d'anticorps polyclonaux de lapins. L'ADN de CaMV (ADNg) a été recherché par PCR. Les amplicons obtenus résultent de réactions de PCR dans la région codant pour la protéine P5 de CaMV. Les produits d'amplification ont été séparés sur gel d'agarose 1% et révélés sous UV grâce à l'addition de bromure d'éthidium dans le gel. Les photos des plants sauvages (WT) ou mutants (*CA-rop2*) infectés par CaMV ou sains ont été prises à 30 JPI.

Le protocole mis au point pour ce travail fait l'objet d'une publication en collaboration avec l'équipe du Dr Laurence Maréchal-Drouard à l'IBMP. Accepté, il sera inclus dans un manuscrit consacré à l'étiquetage des ARN et soumis à la rentrée, à Springer par son éditeur, Dr Manfred Heinlein. Une copie de cet article est fournie en **Annexe 1**.

1.2.2. Constructions plasmidiques

Le plasmide viral utilisé dans le système de protoplastes, pCaMV-GFP (**Matériels 4.1**, **figure R-6**), nous a été généreusement fourni par le Dr Martin Drucker, actuellement à l'INRA de Colmar.

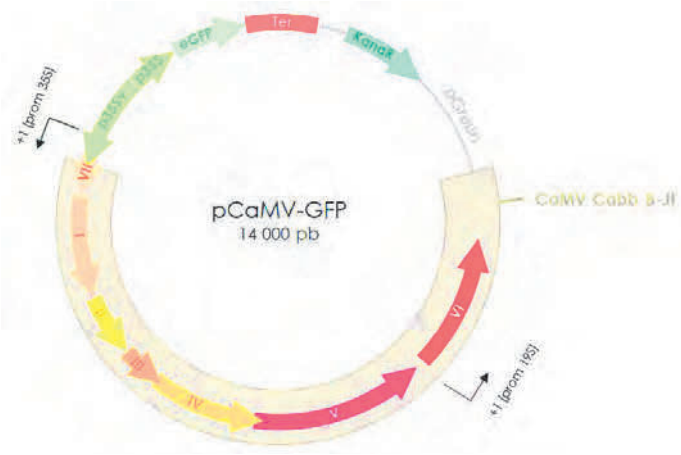


Figure R-6 | Carte du vecteur pCaMV-GFP. Le vecteur viral pCaMV-GFP correspond à un plasmide pGreen modifié portant 1,1 copie du génome de l'isolat Cabb B-JI de CaMV sous le contrôle du promoteur 35S viral complet (p35Sv). Il permet l'expression en cellule végétale des ARN 35S et 19S et le déroulement du cycle viral entier. pCaMV-GFP possède également la séquence codant l'eGFP sous le contrôle par une version tronquée du promoteur 35S (p35S).

Ce vecteur contient 1,1 copie du génome du CaMV de l'isolat Cabb B-JI, placée sous le contrôle du promoteur 35S et qui permet l'expression par l'ARN polymérase II cellulaire de l'ARN 35S complet. Pour information, les plasmides conventionnels employés lors des expériences d'inoculation aux plantes pour produire le CaMV ne sont pas adaptés à la production de l'ARN 35S complet dans ce genre de système d'expression transitoire. Pour pouvoir exprimer cet ARN dans son entier dans les protoplastes, ces vecteurs doivent posséder une duplication de séquence permettant la transcription des répétions présentes aux deux extrémités de l'ARN 35S.

Le promoteur viral interne 19S permet quant à lui la synthèse de l'ARN 19S qui code la protéine P6 indispensable pour l'expression de l'ensemble des protéines virales. L'expression du jeu complet des protéines du CaMV en système protoplastes, a été décrite par Khelifa et al. (2010).

Enfin, en plus du génome de CaMV, ce plasmide contient également, cloné en sens inverse des éléments viraux et placé sous sa propre cassette d'expression, le gène de la protéine fluorescente eGFP pour – *enhanced green fluorescent protein* – utilisé ici comme marqueur de transfection.

Afin de déterminer quelle(s) région(s) de l'ARN est (sont) impliquée(s) dans son export nucléaire, grâce à des expériences de compétitions, mais aussi pour confirmer l'implication de nos différentes protéines candidates dans ce processus, par des expériences de complémentation, nous avions également besoin de mettre au point un vecteur facilement modifiable nous permettant d'exprimer ces différents facteurs/éléments viraux dans notre système.

Pour cela nous avons exploité un plasmide créé par Clément Bouton au cours de son doctorat, le vecteur Litmus35S-*Xba*I-tNos (3,9 kpb) (Bouton, 2014; non publié) que nous avons modifié grâce à la technique de clonage du Gibson assembly™, afin de lui faire exprimer la protéine fluorescente rouge mRFP utilisée comme deuxième marqueur de transfection. Le plasmide obtenu, appelé Litmus35S-*Xba*I-tNos-35S-mRFP (**Matériels 4.3**, **figure R-7**), porte une cassette de clonage composée du promoteur 35S et du terminateur Nos qui peut être linéarisé en plein milieu de cette dernière grâce au site de restriction unique *Xba*I. Brièvement, la technique développée par le Dr Daniel Gibson (Gibson et al., 2009, 2010) permet de cloner facilement et rapidement dans des vecteurs linéarisés différents fragments obtenus par PCR et partageant des extrémités 3' et 5' chevauchantes et complémentaires avec les extrémités du vecteur ouvert (**Méthodes 5.6.2.**). Grâce à cette technique nous avons pu générer plus de 20 vecteurs recombinants différents qui sont présentés dans le **tableau MM-6**, partie **Matériel**.

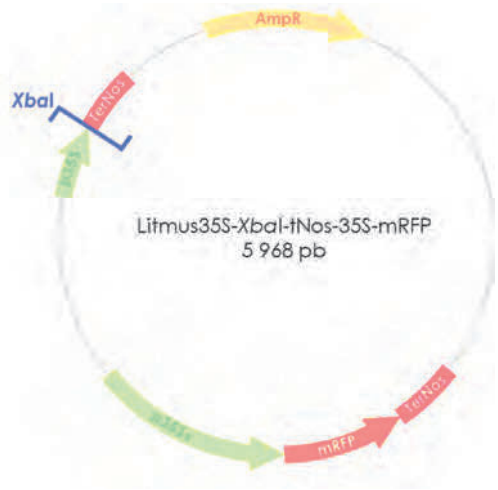


Figure R-7 | Carte du vecteur Litmus35S-*XbaI*-tNos-35S-mRFP. Le plasmide Litmus35S-*Xba*I-tNos-35S-mRFP correspond à plasmide Litmus28i modifié dans lequel a été insérée sous contrôle du promoteur 35S complet de CaMV (p35Sv) la séquence codant la mRFP – *monomeric red fluorescent protein* - suivie du terminateur Nos (TerNos). La mRFP exprimée dans les protoplastes sert de marqueur de transfection du plasmide. Une seconde cassette d'expression a été créée de part et d'autre du site de restriction *XbaI* encadré par le promoteur 35S (p35S) et un terminateur Nos (TerNos). Le site unique *XbaI* a principalement été utilisé pour linéariser le vecteur afin de réaliser nos différents clonages par la technique du Gibson assembly™. Dans notre étude, le vecteur Litmus35S-*Xba*I-tNos-35S-mRFP sert de plateforme d'expression en protoplastes d'Arabidopsis.4

1.2.3. Analyse de l'expression des ARN du CaMV

Vingt heures après transfection, les protoplastes sont observés sous microscope à épi-fluorecence afin d'estimer l'efficacité de transfection et d'expression en recherchant l'eGFP et la mRFP, le cas échéant. En améliorant notre protocole nous avons pu atteindre des efficacités de transfection généralement de l'ordre de 40 %. L'aspect de ces protoplastes, observé au microscope à différents grossissements, est présenté dans la **figure R-8**.

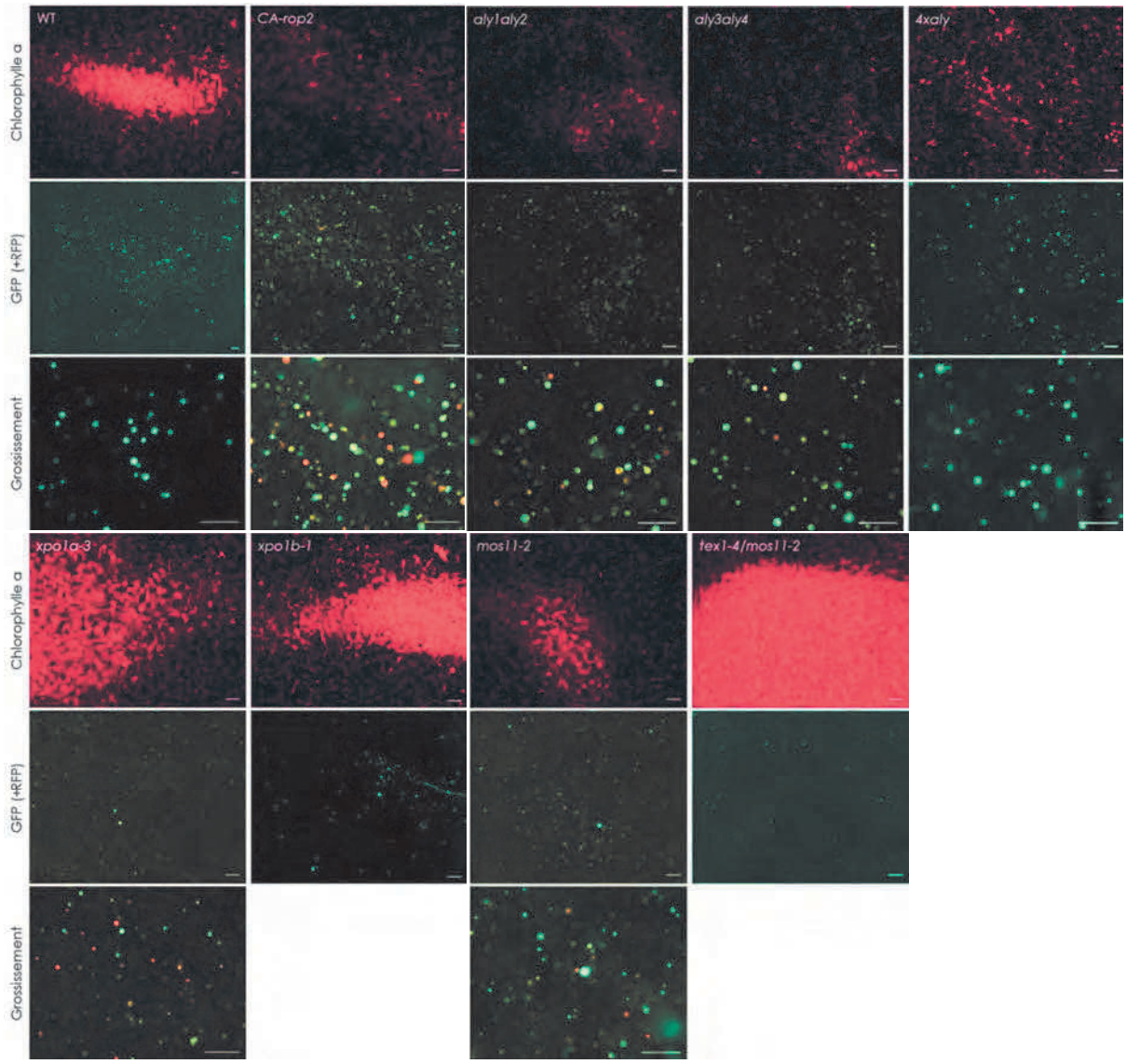


Figure R-8 | Observation au microscope à épi-fluorescence des protoplastes d'*A. thaliana* transfectés avec le plasmide Litmus35S-*Xba*I-tNos-35S-mRFP et/ou le vecteur viral pCaMV-GFP. Les protoplastes de mésophylle d'*A. thaliana* Col-0 sauvage (WT) ou mutants ont été transfectés avec pCaMV-GFP (vert) et Litmus35S-*Xba*I-tNos-35S-mRFP (orange) ou le plasmide entraîneur Litmus28i (Ø). Les protéines eGFP (λ excitation = 488 nm) et mRFP (λ excitation = 561 nm) ont été observées 20 h après transfection au microscope à épi-fluorescence. L'efficacité de la transfection est estimée en observant le nombre de protoplastes oranges et/ou verts et les protoplastes totaux visibles par excitation de la chlorophylle a (rouge)(λ excitation = 650 nm). Les images ont été prises aux grossissements X40 et X112. Les barres d'échelles correspondent à 200 µm.

Des différences d'efficacité de transfection ont été observées entre les protoplastes d'Arabidopsis sauvage et ceux préparés à partir de plantules mutantes parce que ces derniers étaient soit plus fragiles soit plus difficiles à transfecter, Il est important de souligner que lors de la mise au point du système nous avons aussi essayé de transfecter les protoplastes avec des quantités croissantes de pCaMV-GFP et qu'utiliser 10 µg de ce plasmide en présence de 15 µg de plasmide « entraineur » était la combinaison qui assurait la meilleure efficacité de transfection. Pour pallier aux différences de transfection observées, les valeurs de Ct mesurées

par RT-qPCR ont aussi été normalisées en ciblant le messager de l'eGFP, le gène rapporteur porté par le plasmide viral. Enfin, dans nos expériences de compétition et de complémentation, nous avons transfecté les protoplastes avec un excès du plasmide Litmus35S-*Xba*I-tNos-35S-mRFP par rapport à pCaMV-GFP à savoir 2/1 pour les expériences de complémentation et 8/1 pour les expériences de compétition (**Méthodes 2.2.**). Utiliser une plus grande quantité du plasmide Litmus modifié (qui exprime la mRFP) et de taille inférieure (entre 6-8 kbp, **Matériel**, **tableau MM-4**) à celle du plasmide viral (qui exprime l'eGFP) pCaMV-GFP (14 kbp) devrait assurer sa co-transfection optimale et une expression importante des séquences virales qu'il apporte. Dans toutes les expériences de co-transfection, les valeurs de Ct mesurées par RT-qPCR ont été normalisées non seulement par rapport aux ARNm cellulaires de la GAPDH et ceux du marqueur de transfection eGFP, mais aussi par rapport aux ARNm de la mRFP.

Pour analyser l'export nucléaire des transcrits viraux, les protoplastes ont été soumis à un fractionnement subcellulaire afin de séparer les noyaux de la fraction cytosolique, en les incubant en présence d'un tampon hypotonique (NaCl 0,01 M) additionné de NP-40, un détergent non ionique et non dénaturant qui permet de déstabiliser la membrane plasmique sans abimer l'enveloppe nucléaire (Bouton, 2014)(**Méthodes 2.3**).

Après centrifugation, le surnageant constitue la fraction cytoplasmique et le culot contient les noyaux. La pureté de ces fractions a été vérifiée par western blot en ciblant d'une part une protéine strictement nucléaire l'histone H3 (17 kDa) et d'autre part une protéine cytosolique, l'UDP-glucose pyrophosphorylase 1 (52 kDa). La très faible intensité du signal avec des anticorps anti-UGPase 1, dans la fraction nucléaire suggère que la fraction nucléaire est signifiactivement enrichie alors que l'absence de signal avec les anti-H3 dans la fraction cytosolique indique que cette dernière est apparemment exempte de toute contamination par des extraits nucléaires. La présence de traces de la grande sous-unité de la ribulose 1,5, bisphosphate carboxylase/oxygénase (RuBisCO) dans la fraction nucléaire, comme le montre la coloration du gel à l'Instant Blue™, indique une probable contamination des noyaux par des chloroplastes (**figure R-9**).

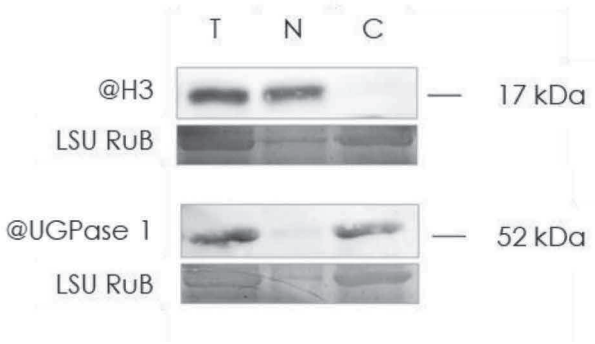


Figure R-9 | Analyse de la pureté des fractions cytoplasmique, nucléaire et totale de protoplastes d'*A. thaliana* par immunodétection de l'histone H3, et de l'UDP-glucose pyrophosphorylase A. Les protéines des fractions totale (T), nucléaire (N) et cytoplasmique (C) de protoplastes non transfectés ont été séparées par électrophorèse sur gel de polyacrylamide 15 %, en conditions dénaturantes, puis révélées à l'Instant Blue™ ou transférées sur membrane. L'UDP-glucose pyrophosphorylase 1 (@UGPase 1), une protéine à localisation exclusivement cytoplasmique et l'histone H3 nucléaire (@H3) ont été recherchées dans les fractions par western blot à l'aide d'anticorps polyclonaux. La coloration du gel montre la grande sous-unité de la RuBisCO (LSU RuB)(55 kDa). La masse moléculaire (en kDa) des protéines est indiquée à droite des figures.

Comme mentionné plus haut, la transfection des protoplastes d'Arabidopsis s'étant faite en présence de très grandes quantités de plasmide recombinant (au moins 10 µg), il nous a fallu éliminer entièrement l'ADN plasmidique contenant le génome du CaMV, pour que ce dernier n'interfère pas, par la suite, dans l'analyse quantitative des transcrits viraux. De ce fait, après extraction au Trizol™, les fractions ARN ont été soumises à un traitement à la TurboDNase™ (Ambion™) dont l'activité six fois supérieure à celle de la DNase I, permet d'éliminer toute trace d'ADN dont la présence donnerait un signal faussement positif en qPCR. La TurboDNase™ est inactivée non pas par chauffage, mais grâce à un réactif, fourni par le producteur et agissant à température ambiante, ce qui préserve les ARN de la dégradation, ces derniers étant sensibles à des températures élevées en présence de cations divalents. L'élimination complète de l'ADN plasmidique est vérifiée par PCR.

Le contenu en ARN 35S des fractions cytosolique et nucléaire est analysé par RT-qPCR en utilisant, pour synthétiser les ADNc, la transcriptase inverse SuperScript IV™ (Invitrogen™). Cette version améliorée de la RT du MLV est d'une excellente processivité ce qui nous a permis de détecter les ARN viraux à l'inverse des autres RT disponibles sur le marché. Dans la réaction de PCR quantitative en temps réel, nous avons utilisé comme chimie le SYBR Green I, avec, dans la plupart de nos expériences, un couple d'amorces qui s'hybrident spécifiquement dans l'ORF V de l'ARN 35S afin de détecter l'ensemble des isoformes de cet ARN, épissés et non épissé, car seule la région de l'ARN 35S située en amont de l'ORF III fait l'objet d'un épissage (Bouton et al., 2015a). Dans les expériences de complémentation par la protéine P5, nous avons utilisé des amorces ciblant la séquence de l'ORF IV. Pour normaliser nos résultats, nous avons également cherché à déterminer, parmi plusieurs ARNm cellulaires (*TIP41* (AT4G34270.1), *GAPDH* (AT3G04120.1), *ACT2* (AT3G18780.2) ou encore *SAND* (AT2G28390), lequel était le plus stable, à la fois dans les fractions cytoplasmiques et nucléaires et, totales ; le messager cellulaire le plus

adapté a été celui de la GAPDH. Pour rappel, les messagers codant l'eGFP (pour pCaMV-GFP) et la mRFP (pour Litmus35S-*Xba*I-tNos-35S-mRFP) sont utilisés pour normaliser le taux de transfection entre les différents échantillons de protoplastes transfectés avec ces différents plasmides.

Toutes les expériences qui seront présentées ci-dessous ont, sauf mention contraire, été réalisées au moins trois fois de manière indépendante et ont inclus au moins une fois des protoplastes d'*Arabidopsis thaliana* Col0 WT, comme témoin.

Les résultats de la quantification relative des ARNm viraux par la PCR en temps réel sont exprimés sous forme d'un rapport, C/N, des ARN 35S des fractions cytoplasmiques (C) par rapport aux ARN des fractions nucléaires (N). Les valeurs obtenues dans les différentes conditions sont comparées au contrôle Arabidopsis WT ou *CA-rop2* pour lequel ce rapport est fixé à 1. Ainsi, on peut détecter par exemple l'accumulation d'ARN 35S dans le noyau suite à l'utilisation d'un inhibiteur comme la leptomycine B ou dans des protoplastes n'exprimant pas un facteur d'export comme par exemple, MOS11.

Ce système facilement adaptable nous a permis l'étude de l'implication dans l'export de l'ARN 35S, des protéines qui interviennent en *trans*, en transfectant des protoplastes mutants n'exprimant pas certains facteurs cellulaires ou en utilisant des plasmides pCaMV-GFP mutés n'exprimant pas l'une ou l'autre protéine virale. Enfin, ce système convient également pour déterminer quels éléments en *cis* de l'ARN 35S pourraient intervenir dans l'export nucléaire grâce à la mise en place d'expériences de compétition consistant à co-exprimer avec les transcrits viraux des séquences d'ARN correspondant à différentes régions de l'ARN 35S.

2. Résultats présentés en vue de publication

Cet article sera très prochainement soumis au journal *Nucleic Acid Research*.

Nuclear export of plant pararetrovirus mRNAs involves the TREX complex, two viral proteins and the 5' highly structured leader region

Julie Kubina[co], Angèle Geldreich[co], Jón Pol Gales, Clément Bouton, Lyubov Ryabova, Klaus Grasser, Mario Keller and Maria Dimitrova

Résumé

Les protéines TEX1, MOS11 et ALY du complexe TREX-1 d'Arabidopsis sont utilisées par le CaMV pour exporter ses ARN 35S

Comme décrit précédemment aucune protéine orthologue ou homologue fonctionnelle à l'exportine TAP et son co-facteur p15, qui représentent la voie majeure d'export

nucléaire des ARNm chez la levure et les métazoaires, n'a été identifiée chez les plantes à l'heure actuelle. En revanche, les orthologues de nombreuses protéines du complexe TREX-1 dont MOS11 (orthologue de CIP29; Germain et al., 2010), TEX1 (orthologue de THO3/TEX1; Sørensen et al., 2017) et ALY (orthologue de ALY/REF; Pfaff et al., 2018) sont connues et décrites comme intervenant dans l'export des ARNm d'*A. thaliana* (Ehrnsberger et al., 2019a). Les ARN 35S du CaMV étant coiffés, polyadénylés et pour certains épissés, nous avons voulu vérifier s'ils empruntent la voie TREX-1 pour rejoindre le cytoplasme.

Pour cela, nous avons inoculé 30 plantes d'Arabidopsis mutées dans les gènes codant les protéines MOS11 (*mos11-2*), MOS11 et TEX1 (*mos11 tex1*) ou ALY (*aly1 aly2*, *aly3 aly4* et *4xaly4*) ainsi que 30 plantes sauvages, avec un extrait brut préparé à partir de feuilles d'Arabidopsis infectées par CaMV Cabb B-JI, puis suivi l'apparition des symptômes viraux sur les plantes au cours du temps.

Chez les plantes *mos11-2,* le prodrome montre un retard de 5 jours en moyenne, comparé aux plantes sauvages. Plus de 20 jours après inoculation, la quasi-totalité des plantes Col-0 sont symptomatiques alors que seules 60 % des plantes *mos11-2* finiront par développer une infection. Des résultats similaires ont été obtenus avec les plantes *mos11 tex1* dont la moitié seulement est infectée à chaque fin de cinétique ; suggérant une importance toute particulière de ces éléments du complexe TREX-1 pour le cycle infectieux de CaMV.

Si les doubles mutants *aly1 aly2* et *aly3 aly4* ne montrent pas de retard dans l'apparition des symptômes, seules 80 % d'entre elles finissent par être infectées en fin de cinétique, lorsque la quasi totalité (97%) des plantes sauvages le sont. Une résistance à l'infection virale encore plus prononcée est obtenue avec les mutants n'exprimant aucun des orthologues de ALY (*4xaly*), avec seulement 10 à 20 % de plantes montrant des symptômes 3 à 4 semaines après inoculation. Ces résultats démontrent que chacun de ces facteurs d'export et spécialement l'adaptateur ALY, revêtent d'une importance toute particulière pour le bon déroulement du cycle de réplication du CaMV, dont l'infection virale semble considérablement compromise lorsque ceux-ci sont absents.

Afin de vérifier que l'effet observé ne soit pas dû à une inhibition plus générale de l'export des ARNm cellulaires, ces plantes mutantes ont également été infectées avec un virus au cycle infectieux exclusivement cytoplasmique, le TuMV - *Turnip mosaic virus* -. Nos résultats montrent clairement que la réplication de ce dernier n'est affectée dans aucune des plantes mutantes, suggérant que l'effet observé avec CaMV soit majoritairement due à une inhibition de l'export de ses ARN messagers plutôt que des messagers cellulaires.

Nous avons poursuivi nos investigations avec des protoplastes préparés à partir des plantules d'Arabidopsis mutants qui n'expriment pas ces protéines (*aly1 aly2*, *aly3 aly4*, *4xaly*, *mos11-2* et *tex1-4*/*mos11-2*). Les protoplastes ont été transfectés avec le vecteur pCaMV-GFP. Nous avons mesuré par RT-qPCR, une diminution de plus de 60 % de l'export des ARN 35S dans le

mutant *mos11-2* par rapport aux protoplastes contrôles et une inhibition encore plus importante (70 %) avec les protoplastes n'exprimant ni MOS11 ni TEX1 (*mos11 tex1*). Enfin, une accumulation nucléaire d'environ 80 % des ARN viraux est observée avec les mutants de ALY, *aly1 aly2* et spécialement *aly3 aly4* et, *4xaly* où le phénomène est le plus prononcé. Ces résultats soulignent, une fois de plus, l'importance de ces protéines nucléaires pour le virus et confirment que le complexe TREX-1 contribue de manière essentielle à l'export des ARN 35S de CaMV.

Les protéines virales P4 et P5 sont importantes pour l'export nucléaire des ARN 35S

Nous l'avons vu, il est commun que les virus utilisent une de leurs propres protéines pour exporter leurs ARNm vers le cytoplasme, et, en particulier, ceux qui sont partiellement ou non épissés. Plusieurs protéines de CaMV pourraient jouer ce rôle au regard de leurs différentes propriétés. D'abord P6, qui est capable de lier à la fois les ARN simple et double-brin et de faire la navette entre le noyau et le cytoplasme, grâce à ses deux NLS et un NES. Ensuite la protéine de capside P4, qui possède un NLS mais celui-ci rappelons-le, n'est exposé qu'après l'assemblage de la capside et sa maturation ce qui lui permet d'être acheminée au niveau des pores nucléaires. En protoplastes, les formes clivées de P4 ont été également localisées dans le noyau (Karsies et al., 2002), ce qui suggère qu'elles peuvent y être transportées dès la déstabilisation des particules virales au début du cycle. Enfin, P4 possède un domaine en doigt de zinc à son extrémité C-terminale qui lui permet d'interagir avec l'élément ψ présent dans la région 5'UTR des ARN 35S probablement lors de l'encapsidation. Nous avons également envisagé que la transcriptase inverse voire son précurseur, la protéine P5, puisse être impliquée dans l'export de l'ARN 35S. Comme sa fonction l'indique, P5 est aussi capable de lier l'ARN 35S pré-génomique (Pietrzak et Hohn, 1987) et ce, au niveau de sa région *leader*. Même si une localisation nucléaire lui était inconnue jusqu'à cette étude, une analyse *in sillico* de sa séquence par deux logiciels différents a également montré l'existence d'un NLS bipartite putatif au sein de la protéine.

Dans le but de déterminer si l'une ou plusieurs protéines de CaMV sont impliquées dans l'export de ses ARN ou si ce mécanisme était purement cellulaire, nous avons introduit, dans le plasmide viral pCaMV-GFP, 2 ou 1 codon(s) STOP, respectivement, en aval des codons AUG initiateurs des protéines P6 et/ou P5 afin d'abolir leur expression. En empêchant l'expression de P6, plus aucune protéine virale ne peut être synthétisée car l'une des multiples fonctions de cette protéine est d'activer l'expression des 6 ORF de l'ARN 35S polycistronique de CaMV, par le mécanisme de réinitiation de la traduction (Pooggin et Ryabova, 2018).

Nous avons montré qu'en absence de P6, de P5 ou des deux protéines à la fois, l'export nucléaire des ARN 35S est fortement diminué (de 70 %), en protoplastes. La rétention partielle dans le noyau, des ARN viraux en absence de la protéine P6 permet de conclure que celle-ci

et/ou l'une des autres protéines virales contribue de quelque manière à l'export de ces ARN. En tout cas, l'une de ces protéines s'est avérée être P5, étant donné que l'export est fortement inhibé quand elle n'est pas exprimée. Des expériences de complémentation, toujours en protoplastes, nous ont permis de confirmer son rôle dans ce processus, puisque l'export nucléaire des ARN viraux est rétabli lorsqu'elle est apportée en *trans*. Mais uniquement sous sa forme entière, car l'expression en *trans* de ses formes matures (protéase et rétrotranscriptase/RNaseH) ne complémente pas le facteur d'export viral manquant. Ce résultat est remarquable car il suggère que P5 a une localisation nucléaire alors qu'elle est considérée depuis longtemps comme une protéine strictement cytoplasmique. En effet, la transcriptase inverse et la protéase issues du clivage en *cis* de P5 agissent dans le cytosol, respectivement, pour synthétiser l'ADN viral et pour produire les protéines de la capside par de multiples clivages de P4.

Par la suite, des expériences de GST-p*ull down,* nous ont permis de démontrer que P5 est en fait capable, tout comme la protéine de capside virale P4, d'interagir directement avec la protéine nucléaire MOS11. Une étude plus détaillée indique que c'est avant tout le domaine RT de P5 qui permet son interaction avec ce facteur de l'export.

L'utilisation de l'étiquette fluorescente eGFP fusionnée aux protéines P4 et P5, nous a finalement permis de visualiser, pour la toute première fois, que P5 mais aussi P4, sous leurs formes non-maturées et fusionnées à l'eGFP, ont une double localisation, cytoplasmique et nucléaire, dans les protoplastes. Enfin, l'analyse de leur localisation intracellulaire en parallèle de celle du facteur d'export MOS11, strictement nucléaire et exprimée en fusion avec la mRFP, a fini par montrer une co-localisation nucléaire de ces 3 protéines dans les protoplastes d'Arabidopsis, suggérant qu'une interaction entre MOS11 et ses partenaires viraux pourrait avoir lieu *in vivo* et dans le noyau.

La région leader de l'ARN 35S est un élément agissant en *cis* qui est impliqué dans l'export

Dans le but d'identifier les séquences en *cis*, impliquées dans l'export des ARN de CaMV, nous avons utilisé notre système d'étude en protoplastes pour réaliser des expériences de compétition entre l'ARN 35S et des régions définies de cet ARN. Des protoplastes d'Arabidopsis ont été co-transfectés avec le plasmide viral pCaMV-GFP et un large excès du plasmide Litmus35S-*Xba*I-tNos-35S-mRFP dans lequel nous avons cloné différents ADNc codant pour des ARN plus courts (de 289 à 865 nt) et correspondant à différentes séquences couvrant les 3 000 premiers nucléotides de l'ARN 35S pré-génomique ainsi que sa région 3' UTR (**Matériel**, **tableau MM-4**).

Lorsque le vecteur pCaMV-GFP a été transfecté seul, les ARNm viraux ont été exportés vers le cytoplasme. En revanche, lors de la co-expression d'un ARN plus court, de 765 nucléotides, et correspondant à la région *leader* de l'ARN 35S et au début de l'ORF VII, nous

avons observé une forte inhibition de l'export nucléaire des ARN 35S. Ceci suggère que cette région pourrait contenir des éléments en *cis* impliqués dans l'export puisqu'elle entre en compétition avec l'ARN pré-génomique et inhibe partiellement l'export de celui-ci.

Pour rappel, le *leader* de l'ARN 35S est très structuré et forme une grande tige-boucle qui pourrait être reconnue par des facteurs d'export à la manière de la tige-boucle CTE des rétrovirus simples. Partant de cette hypothèse, la suite de nos expériences a donc consisté à générer 8 variantes plus courtes de cette structure afin de déterminer si celle-ci était importante pour sa reconnaissance par la machinerie d'export. Pour cela, nous avons disséqué la région *leader* et réalisé des expériences de compétition avec ces 8 nouveaux fragments. Nos résultats montrent clairement que seuls les fragments contenant la tige-boucle entière du *leader* sont capables d'inhiber l'export des ARN 35S. Ainsi, nos résultats montrent non seulement que la région *leader* de l'ARN 35S de CaMV constitue au moins un des éléments en *cis* nécessaires pour l'export nucléaire de ses ARN 35S mais suggère aussi que la structure secondaire de sa tige-boucle semble particulièrement importante dans cette fonction.

La protéine ALY1 interagit avec l'ARN 35S de CaMV

Finalement, afin de déterminer si la région *leader* de l'ARN 35S pouvait effectivement être reconnue et lier les facteurs d'export liant l'ARN du complexe TREX-1 nous avons réalisé des expériences de gel retard *in vitro* dans lesquels l'interaction de la protéine ALY1 avec l'ARN leader mais aussi avec un autre fragment, correspondant à une région plus en aval de l'ARN 35S et vraisemblablement non impliquée dans l'export, a été testée. Nos résultats montrent qu'ALY1 est capable d'interagir de la même manière avec ces deux ARN, sans les reconnaitre de manière spécifique.

Nuclear export of plant pararetrovirus mRNAs involves the TREX complex, two viral proteins and the 5' highly structured leader region

Julie Kubina[1], Angèle Geldreich[1], Jón Pol Gales[1], Clément Bouton[1], Lyubov Ryabova[1], Klaus Grasser[2], Mario Keller[1] and Maria Dimitrova[1*]

[1] Institut de biologie moléculaire des plantes, CNRS, Université de Strasbourg, Strasbourg, France
[2] Cell Biology & Plant Biochemistry, Biochemistry Centre, University of Regensburg, D-93053 Regensburg, Germany

* To whom correspondence should be addressed. Tel: +33 (0)3 67 15 53 72; Fax: +33 (0) 3 67 15 53 00; Email: maria.dimitrova@ibmp-cnrs.unistra.fr

The authors wish it to be known that, in their opinion, the first 2 authors should be regarded as joint First Authors.

Present Address: Clément Bouton, Mucosal Infection and Immunity, Faculty of Medicine, Imperial College London, London SW7 2AZ, UK

ABSTRACT

In eukaryotes, the major nuclear export pathway of mature mRNAs is represented by the dimeric receptor TAP/p15 that is recruited on mRNAs *via* the multisubunit TREX complex composed of the THO core and different export adaptors. Viruses replicating in nuclei adopt different strategies to hijack the export factors and achieve cytoplasmic translation of their mRNAs. In plants, the nuclear export remains poorly described: the export receptors are unknown albeit the TREX complex resembles the mammalian one, with a conserved hexameric THO core, associated with ALY and UIEF proteins, UAP56 and MOS11, the orthologue of mammalian CIP29. The nuclear export of viral mRNAs has never described in plants. We investigated and characterized for the first time the export of mRNAs of plant pararetrovirus CaMV in Arabidopsis and demonstrated that it is inhibited in plants deficient in ALY, MOS11 and TEX1 expression, which makes these plants partially resistant to CaMV infection. Two viral proteins, coat protein P4 and reverse transcriptase P5, are important for nuclear export, and interact and colocalise in nuclei with export factor MOS11. The long and highly structured 5' leader region of 35S RNAs was identified as a *cis*-acting export element interacting with ALY1 *in vitro*.

INTRODUCTION

Nuclear mRNA export plays an important role in gene expression and has been largely studied and well documented since many years in yeast and metazoa. To be efficiently exported transcripts must undergo several maturation steps including capping (1), splicing (2) and 3′ end formation (3). During this co-transcriptional maturation that, except capping, takes place in the nuclear perispeckles (2) maturation and export factors are recruited on the transcripts thus forming large messenger ribonucleoproteins (mRNPs). These factors then allow mRNPs docking onto the nuclear basket of the nuclear pore complex (NPC), transiting the central channel and release from the cytoplasmic fibrils (4–6). The major transporter for mRNPs is the protein heterodimer TAP/p15 (also known as NXF1/NTF1 in metazoa and Mex67/Mtr2 in yeast). Even if TAP is able to directly bind single stem-loop structures named CTE (Constitutive Transport Element) (7, 8), in the case of spliced mRNAs it is rather recruited through adaptor proteins then through direct sequence specific RNA binding (5, 9). A key player in mRNP biogenesis, maturation and TAP/p15 recruitment is the TRanscription and EXport (TREX-1) complex, conserved across a wide range of organisms (5). TREX-1 consists of a stable hexameric core called THO (10) and some important additional proteins such as the adaptors ALY (11) and its functional homologue UIF (12), UAP56 (13) and CIP29 (14). ALY (and its associated THO complex) is rapidly recruited on the nascent mRNAs by the cap-binding complex (CBC) (15–17), which associates with the 5’ cap. During splicing the exon junction complexes (EJCs) are deposited on mRNAs and additional ALY copies are the then loaded upstream of the exon-exon junctions (18) or transferred from the CBC (17). DEAD box RNA helicase UAP56, also considered as a component of EJC (18), associates with both ALY and CIP29 in an ATP-dependent manner (19). Finally, the TREX adaptor proteins, and mostly ALY, together with the THO sub-complex, act as a binding platform for export receptor TAP (and its associated p15 protein) (20), which directs the export competent mRNPs to the nuclear pore (21). UAP56 and ALY are ultimately released before or during translocation through the NPC (22).

Even if splicing in very important for the nuclear export of mature mRNAs, intronless or intron retaining mRNAs are also able to recruit TREX components and/or adaptors in order to further load the export transporters TAP/p15. In these cases the association between CBC and TREX is of big importance (23) but probably also the ability of the scanning spliceosomes to deposit, at lower efficiency, EJC and mRNA export factors on single-exon transcripts (17). Intronless mRNAs, however, likely use different strategies to attract TAP/p15: histone mRNAs bind SR adaptor proteins that directly recruit TAP/p15 (24–26), whereas other intronless cellular mRNAs harbour conserved cytoplasmic accumulation regions (CARs) that interact with TREX components (27, 28).

The nuclear export of mRNAs still remains poorly characterized in plants where only the composition of the Arabidopsis TREX complex is well described (29). With some minor differences, the hexameric THO sub-complex resembles the metazoa one (30). These differences might however be important and allow plants to adapt the export of different RNA populations and to adapt, more generally, to the changing environmental conditions. The differences might also be explained by the fact that in Arabidopsis the THO sub-complex is also involved in mRNA processing and the biogenesis of small RNAs (29). Four adaptor ALY proteins are ubiquitously expressed in vegetative cells from four different genes in Arabidopsis (31). They localize to the nuclei, interact with UAP56 and bind to single and double stranded RNA (32). Single and double knockout *aly* mutants do not exhibit any significant phenotypes whereas the quadruple *4xaly* is affected in growth, flowering and seed production, and displays nuclear accumulation of mRNAs suggesting an inefficient nuclear export (32). Two UIF-like proteins have also been described in Aradidopsis: they bind to RNA and to UAP56, and cooperate with ALY1-4 to mediate efficient nucleocytosolic mRNA transport (33). The DEAD-box RNA helicase UAP56 is encoded by two adjacent genes in Arabidopsis and the resulting identical proteins interact with single and double stranded RNA, and with ALY and MOS11 (the Arabidopsis orthologue of CIP29) proteins (34). MOS11 is also a nuclear protein and the knockout *mos11* mutant is characterized by nuclear accumulation of mRNAs (35). Very surprisingly, orthologues of the main mRNP transporter TAP which is loaded *via* its interactions with adaptor proteins such as ALY, do not exist in the Arabidopsis genome, suggesting that plants may use different exportins to direct and transport mRNPs through the nuclear pore (29).

Like all living organisms, yeast, plants and metazoa are exposed and susceptible to multiple viral infections. Many DNA or RNA viruses with nuclear replication steps in their life cycles, export viral mRNAs by hijacking the TAP/p15 pathway. Herpesviruses, whose majority of mRNAs is intronless, have evolved very efficient mechanisms for their maturation and export: they use a conserved multifunctional protein (ICP27 for herpes simplex virus 1 (HSV1) or its homologue ORF57 for herpes virus saimiri and Kaposi sarcoma herpesvirus (KSHV), for example) which interacts both with viral transcripts and cellular adaptors like ALY, UIF and CIP29, thus recruiting the whole TREX complex and enhancing viral mRNA accumulation and export (36–42). Simple retroviruses, like Mason-Pfizer monkey virus (MPMV) use highly structured CTEs (constitutive transport element) on their mRNAs that directly bind TAP with high affinity (43). Hepatitis B (pararetro)virus (HBV) mRNAs also harbour a structured region named posttranscriptional regulatory element (PRE) necessary for TREX recruitment on intronless viral mRNAs and efficient expression of viral proteins (44).

The paraphyllic group *Pararetrovirus*, which members replicate their circular double-stranded DNA genomes by reverse transcription, include not only the vertebrate *Hepadnaviridae* family but also plant *Caulimoviridae*. *Caulimoviridae* together with *Geminiviridae*, *Nanoviridae* (both with circular single-stranded DNA genomes) and *Nucleorhabdoviruses* (with single-stranded negative–sense RNA genomes) make synthesise viral mRNAs in the nuclei of the infected plant cells. These mRNAs are then necessarily exported into the cytoplasm in order to be translated and allow the achievement of the viral replication cycles. If the general mechanism of the nuclear export of mRNAs is poorly characterized in plants, in our knowledge nothing is known about how phytoviruses manage to export viral mRNAs out of the cell nuclei. In this context, we investigated the nuclear export of mRNAs of *Cauliflower mosaic virus* (CaMV), the type member of the family *Caulimoviridae*.

The replication cycle of CaMV includes two main steps, in the nucleus and the cytoplasm: (a) Following entry into the plant cell and disassembly of the capsid proteins on the nuclear envelope (45), the viral dsDNA is imported into the nucleus, where it associates with histones to form minichromosomes that are used as templates for transcription by the host DNA-dependent RNA polymerase II (Pol II) producing a capped and polyadenylated polycistronic pregenomic 35S RNA (pgRNA), comprising the entire genome and encoding six main proteins (P1 to P6), and a monocistronic subgenomic 19S RNA encoding a single protein, P6 (46). Even if splicing rarely occurs within plant viruses, about 70% of total 35S RNA of CaMV undergoes alternative splicing that generates four spliced isoforms and involves four splice donor sites and a single splice acceptor site (47). Alternative splicing however appeared to be a conserved and complex phenomenon since upon inactivation of splice sites by mutagenesis it is kept functional by activation of numerous cryptic splice donor and acceptor sites, emphasising the key role of this process in CaMV biology (48). (b) All these different mRNA isoforms are then exported to the cytoplasm, where their translation and reverse transcription of pgRNA take place. The six viral proteins are: P1, the movement protein; P2, the aphid transmission factor; P3, the capsid-associated protein; P4, the coat protein; the polyprotein P5, homologous to the Pol gene product of retroviruses, that harbours a C-terminal reverse transcriptase/RNaseH and an N-terminal aspartic proteinase that is released by self-cleavage (49), and the multifunctional P6 protein that allows, among others, transactivation of translation of the polycistronic 35S RNAs (50), and formation of numerous electron-dense inclusion bodies also known as viral factories (46). The newly synthesized dsDNA is packaged into the virions to move from cell to cell and to be transmitted from plant to plant.

In this study, we describe, for the first time, the nuclear export of mRNAs of a plant virus, the *Cauliflower mosaic virus*. Our data demonstrate that CaMV exports its 35S polycistronic

mRNAs using the TREX export complex, and that adaptor protein ALY, MOS11 and the THO component TEX1 are required for efficient 35S mRNAs export in Arabidopsis. We also show that viral capsid protein P4 and reverse transcriptase P5 are important for export and that they interact each with MOS11 *in vitro*, and co-localize each with this export factor in cell nuclei. We also provide evidences that the highly structured 5' untranslated region (5' UTR) of the 35S RNAs, also called the leader region, is a *cis*-acting export element, which can load ALY proteins *in vitro*.

MATERIAL AND METHODS

Plant material and growth

For virus infection (CaMV or TuMV) Arabidopsis (*Arabidopsis thaliana*) Col-0 was grown on soil in a growth chamber, with 12h photoperiod at 22°C and 12h dark at 18°C, while for protoplasts preparation, plants were grown on MS medium (51). In the latter case, after sowing, seeds were stratified in darkness for at least 24h at 4°C prior to incubation in the plant growth chamber under long-day conditions (16h photoperiod at 22°C, 8h dark at 18°C) for 10 ± 2 days.

CA-rop2 mutant was described in (52). Seeds of *mos11-2* T-DNA insertion line (SAIL_266_E03)(35) were obtained from the Arabidopsis Stock Center (ABRC) and characterized by PCR-based genotyping to distinguish between plants being heterozygous or homozygous for the T-DNA insertion. PCR were performed on genomic DNA isolated from leaves, with ABRC recommended primers specific for the T-DNA insertion and the target gene *mos11* (data not shown). Arabidopsis double mutant *tex1 mos11* was described in (53) while double mutants *aly1 aly2* and *aly3 aly4*, and quadruple mutant *4xaly* were characterised in (32).

Viruses and host inoculation

All CaMV experiments were performed with the Cabb B-JI isolate (John Innes Institute, Norwich, GB), by mechanical inoculation on 4-leaf stage rosette of infectious Arabidopsis sap. TuMV-GFP (UKI isolate)(54) infectious sap from *Brassica napus* cv. Drakkar, was provided by Manfred Heinlein and inoculated in the same manner.

Plasmids

pCaMV-GFP, which contains an infectious genome of CaMV under the control of the 35S promoter and a reporter GFP in a separate expression cassette for monitoring transfection, was described in (55) and provided by Martin Drucker. pCaMV-P5_STOP, pCaMV-P6_STOP and pCaMV-P5/P6_STOP were generated by exchanging pCaMV-GFP *Kpn2*I/*Mlu*I fragment (spanning the P4/P5/P6 coding sequences) with the same fragment from pBS-BJI (55), in which P5, P6 or both coding sequences were mutagenized as follows. ORF V harbours, in its first 51 nt, 3 ATG codons in the same reading frame (positions 1, 2 and 17). To prevent the production of P5, ATG_{17} was inactivated (ATG ->TAG). ORF VI also harbours, in its first 45 nt, 3 ATG codons (position 1, 9 and 15). To prevent the production of P6, 2 stop codons were inserted: one between the 2nd and 3rd ATG (K_{12} AAA -> TAA), and the other after the 3rd ATG (E_{17} GAA -> TAA).

Litmus-mRFP-35S was generated in 2 steps as follows. 1. In the *Xba*I (Klenow blunt ended)/*Avr*II sites of Litmus28i (56) was cloned a 1,108 bp PCR fragment (containing 35S promoter (835 bp) followed by a unique *Xba*I site in a 20 bp stretch, and *NOS* terminator (253 bp), amplified with modifications from pGWB14 (57) to give Litmus-35S-*Xba*I-Nos. 2. For monitoring transfection, in the *BamH*I/*EcoR*I linearized Litmus-35S-*Xba*I-Nos were cloned by Gibson assembly and according to the manufacturers instructions (NEBuilder® HiFi DNA Assembly Master Mix) 2 amplicons: a 1,766 bp fragment, containing the mRFP coding sequence under the control of 35S promoter, amplified from pH7WGR2 (58)(provided by Khalid Amari, IBMP, Strasbourg, France), and a 258 bp fragment containing the *NOS* terminator sequence, amplified from pMDC32 (59)(provided by Corinne Keichinger, IBMP, Strasbourg, France) to give Litmus-mRFP-35S.

Litmus-mRFP-35S-cmycx4-P5 and –P5_STOP were generated by introducing, by Gibson assembly in *Xba*I linearized Litmus-mRFP-35S 2 fragments: the P5 or P5_STOP sequence (2,043 bp) amplified from pCaMV or pCaMV-P5_STOP, respectively, and a 186 bp c-mycx4 amplified from pGWB17 (57)(provided by Esther Lechner, IBMP, Strasbourg, France).

Litmus-mRFP-35S-P5_PR and -P5_RTRH were generated by introducing, by Gibson assembly in *Xba*I linearized Litmus-mRFP-35S, PCR fragments amplified from pCaMV-GFP and corresponding, respectively, to the PR domain sequence of P5 (aa 1-201, nt 1-603) or the RTRH domain (aa 202-680, nt 604-2,040).

The 13 vectors Litmus-mRFP-35S-(L-VII), -(VII-I), -(I-II), -(II-IV), -(3'UTR), -(SL), -(SL-Left), -(SL_Right), -(SL_Bottom),-(SL_Top), -(L_5'), -(L_3') and -(SS_3') were generated by introducing, by Gibson assembly in *Xba*I linearized Litmus-mRFP-35S, 12 PCR fragments amplified from pCaMV-GFP and corresponding to different sections of 35S RNA or its 5' leader region: (L-VII): nt 1-765; (VII-I): nt 766-1,439; (I-II): nt 1,440-2,304; (II-IV): nt 2,305-3,000; (3' UTR): nt 7,928-8,216; (SL): nt 75-556; (SL_Left): nt 75-324; (SL_Right): nt 325-556; (SL_Bottom): nt 75-185 and 445-556; (SL_Top): nt 186-444; (L_5'): nt 1-186; (L_3'): nt 187-608 and (SS_3'): nt 557-765.

Litmus28i-(L-VII) and –(I-II), used for *in vitro* T7 transcription, were generated by cloning PCR amplified DNAs corresponding to 35S RNA fragments L-VII and I-II, described above, in *BssH*II/*Pst*I linearized Litmus28i.

The 5 vectors pGAD-T7-P5_FL, -P5_PR, -P5_RT, -P5_RTRH, -P5_RH were obtained by introducing, by Gibson assembly in *BamH*I linearized pGAD-T7, 5 PCR fragments amplified from pCaMV-GFP and corresponding to different domains of P5 protein: PR and RTRH were described above, P5_FL: aa 1-680 (nt 1-2,040), RT corresponding to P5 aa 202-547 (nt 604-1,641) and RH to P5 aa 548-680 (nt 1,642-2,040).

pGAD-T7-P3 and –P4, were obtained by cloning P3 and P4 coding sequences, PCR amplified from pCaMV-GFP in, respectively, *EcoR*I/*Xho*I and *Sma*I/*BamH*I restriction sites of pGADT7.

All resulting pGAD-T7 plasmids were used for *in vitro* T7 transcription, translation and radiolabelling in rabbit reticulocytes lysates.

pGEX-6P1-MOS11 was generated by insertion of a PCR fragment harbouring MOS11 coding sequence and amplified from pENTR223-G12556 (provided from ABRC) in *Sal*I/*Not*I digested pGEX-6P1 (60), in frame with GST.

pET24b-ALY1-GB1 was described in (32) and pET-GGcHGB1-eYFP was kindly provided by Nicolas Baumberger (IBMP, Strasbourg, France). Both vectors allow the expression of ALY1 or eYFP, respectively, in fusion with GB1 tag for protein solubilisation in *E. coli*, and His_6 for purification by Ni^{2+}-chelating chromatography. pET-GGcHGB1-eYFP was obtained by Golden Gate cloning (GGc).

Litmus-35S-MOS11-mRFP was obtained by cloning, in frame with mRFP, MOS11 coding sequence, PCR amplified as described above, in Litmus-35S-mRFP, digested by *Sal*I/*Not*I (2 sites located between the 35S promoter and the mRFP coding sequence).

pCKeGFP-P4 and –P5 were generated by introducing P4 and P5 coding sequences, amplified from pCaMV-GFP, in frame with eGFP, in, respectively, *Sma*I/*Avr*II and *BsrG*I/*Avr*II digested pCKeGFP, in which eGFP is under the control of a double enhancer 35S promoter.

Analysis of CaMV proteins and genomic DNA in infected Arabidopsis plants

CaMV protein and DNA extraction was performed as previously described by (61). Briefly, total proteins were extracted from 2 symptomatic or asymptomatic 1 cm leaf discs, which were frozen in liquid nitrogen, ground twice for 15 sec in a Precellys homogenizer at 6,000 rpm, suspended in 100 µL 8 M urea and ground again for 30 sec. Proteins were then fractionated by SDS-PAGE and CaMV P2, P4 and P6 proteins detected by western blots with polyclonal rabbit antisera at a dilution of 1:10 000, alkaline

phosphatase-conjugated goat anti-rabbit IgG antibodies (Biosys) at a dilution of 1:5 000 and the colorimetric substrate BCIP/NBT detection system (Promega Corporation). Anti-P2 antibodies were kindly provided by Stéphane Blanc (INRA, Montpellier, France) and anti-P4 (62) and anti-P6 (63) polyclonal antisera were previously obtained and tested in our laboratory.

For DNA analysis, 2 discs from different leaves from the same symptomatic, asymptomatic or mock-inoculated plants were ground as described above. DNA was denatured in 200 µL 0.5 N NaOH, ground again, and 5 µL were added to 45 µL 100 mM Tris-HCl pH 8 for neutralization. PCRs were performed on 1 µL of the treated crude extract with couples of specific oligonucleotides to amplify a 775 bp sequence from CaMV DNA (PCR_CaMV_P5 O9a_FW: AGGGGAGGAGGTTATCAGAAG and PCR_CaMV_P5 O9b_REV: TCTAGACCAAGGAAGTTTATCTTC) and the actin gene *act2* used as an internal control (not shown).

Arabidopsis protoplast isolation and transfection

Protoplasts were isolated as described by (64) and adapted by Bouton (48). Briefly, 10 days-old *in vitro* grown Arabidopsis seedlings were finely chopped and incubated over night, at 26°C under gentle shaking (50 rpm) in 25 mL enzyme solution (20 mM MES pH 7.5 – 0.4 M mannitol – 20 mM KCl – 10 mM $CaCl_2$ - 1.5% cellulase R10 - 0.4% macerozyme R10). The digestion medium was then filtrated to eliminate leaf debris, washed in 20 mL 2 mM MES pH 5.7 - 154 mM NaCl – 125 mM $CaCl_2$ – 5 mM KCl solution and gently centrifuged to get rid of enzymes, and protoplasts were finally suspended in 4 mM MES pH 5.7 – 0.4 M mannitol – 15 mM $MgCl_2$ medium, counted and adjusted at 10^6 protoplasts/mL.

Transfection was performed as described by (64) with modifications: 300 µL protoplasts were mixed with 300 µL 2 mM MES pH 5.7 – 0.2 M mannitol - 30% PEG4000 – 100 µM $CaCl_2$ - DNA transfection solution (10 µg expression plasmid and 10 µg carrier plasmid) and incubated during 15 min at room temperature. Transfection was stopped with 1 mL 2 mM MES pH 5.7 - 154 mM NaCl – 125 mM $CaCl_2$ – 5 mM KCl solution; protoplasts were centrifuged, suspended in 1 mL 4 mM MES pH 5.7 – 0.2 mannitol – 20 mM KCl and incubated 20h in the dark at 23°C. The ratio of transfected fluorescent protoplasts, expressing eGFP and/or mRFP, was estimated on Axio Zoom V.16 microscope (Zeiss) at 10-fold magnification and excitation at 488 nM (eGFP), 561 nm (mRFP) and 660 nm (chlorophyll).

Subcellular fractionation of transfected Arabidopsis protoplasts

The transfected protoplasts were centrifuged 5 min at 100 x g, cell pellets were suspended in 100 µL hypotonic buffer (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 0.3% NP-40) and centrifuged again 10 min at 5,000 x g. 90 µL of supernatants (cytoplasmic fractions) were harvested and pellets (nuclear fractions) washed 3 times in hypotonic buffer and finally suspended in 90 µL hypotonic buffer. Total fractions are prepared by suspending the protoplasts in 100 µL hypotonic buffer. Purity of the fractions was determined by western blot detection of nuclear histone H3 and cytoplasmic UDP-glucose pyrophosphorylase 1 (UGPase) (data not shown).

RNA extraction and analysis by real time RT-PCR

RNAs were extracted from total, cytoplasmic and nuclear fractions with TRI-Reagent (Sigma), applied on Phasemaker columns (Invitrogen), according to the manufacturers instructions, and finally suspended in RNase-free water. RNAs were further treated with TURBO DNase (Invitrogen) for 30 min at 37°C and analysed by PCR to monitor digestion efficiency. cDNAs were synthesised using the same amounts of RNAs for all samples with SuperScript IV and in presence of $Oligo(dT)_{20}$ Primers and Random Hexamers, according to the manufacturers instructions (Invitrogen).

CaMV 35S RNAs in total, cytoplasmic and nuclear fractions were quantified through their respective cDNAs, diluted to 1/20, with LightCycler™ 480 SYBR Green I Master (Roche Life Science),

and 35S RNA specific primers, targeting ether the P5 coding sequence (qPCR_CaMV_P5_FW: TATAGCCCAATGGATCGTGA and qPCR_CaMV_P5_FW: CTTCTCGGCTTCATTGTTGA), producing a 126 bp amplicon (nt 4387-4493 of 35S RNA)), or the P4 coding sequence, (during the P5 complementation assays), (qPCR_CaMV_P4_FW: CGGTTCTGGTATAATCTG and qPCR_CaMV_P4_REV: TCTTCGGTAGTTGTCATA), producing a 161 bp amplicon (nt 2831-2991 of 35S RNA)). Normalisation was initially tested with a set of 4 mRNAs of cellular housekeeping genes (*TIP41* (AT4G34270), *GAPDH* (AT3G04120.1), *ACTINE 2* (AT3G18780) and *SAND* (AT2G28390)) and *GAPDH* was chosen as the most stable in all fractions. Besides *GAPDH*, Ct values were also normalized to those of *GFP* and/or *mRFP* as indicators of the transfection efficiencies.

Protein purification

Using pET24b plasmids, His_6-ALY1-GB1 and His_6-GB1-eYFP were expressed in *E. coli* BL21 (DE3) and purified as described by (34) and (32) in 2 steps: Ni^{2+}-chelating chromatography (*via* His_6 tag, in 50 mM Tris-HCl, pH 7.5 - 1 M NaCl - 10% (v/v) glycerol - 10 mM 2-mercaptoethanol - 10mM imidazole, on HisTrap™ FastFlow columns (Sigma-Aldrich), with protein elution at 350 mM imidazole and without NaCl). Eluted proteins were further purified by FPLC ion-exchange chromatography using a Resource Q column (GE Healthcare), eluted with a 0 to 1 M NaCl gradient and passed through PD-10 columns (GE Healthcare) to change to storage buffer (50 mM Tris-HCl, pH 8 - 100 mM NaCl - 10 mM 2-mercaptoethanol - 10% (v/v) glycerol - 0.5% (v/v) Triton X-100 - 50 mM L-Arg - 50 mM L-Glu). Finally, proteins were characterized by SDS-PAGE and Coomassie Blue staining, and their concentrations were determined by Bradford colorimetric assay.

GST pull-down assays

Using plasmids pGEX-6P1, GST and GST-MOS11 were expressed in *E. coli* BL21 (DE3) and affinity captured on Glutathione Sepharose 4B (GE Healthcare) as described by (65). Briefly, after 4h IPTG induction at 20°C, 500 mL bacterial cultures were harvested, suspended in 30 mL 50 mM Tris-HCl pH 7.5 – 1 M NaCl – 2 mM DTT – 0.1 mM EDTA – 0.5% NP-40 – Protease Inhibitors (Roche) - 10U RQ1 RNase-free DNase (Promega Corporation), and lysed by sonication. Soluble protein fractions, obtained after centrifugation, were incubated over night with 500 µL of 50% slurry of glutathione sepharose beads and washed 3 times with the same buffer at 150 mM NaCl. Protein capture on beads was analysed by SDS-PAGE and Coomassie Blue staining.

P3, P4, P5 and its truncated versions were *in vitro* transcribed from the T7 promoter present in pGADT7 and translated in the presence of 10 µCi [^{35}S-methionone] in TnT T7 Coupled Reticulocyte Lysate System (Promega Corporation), according to the manufacturers instructions.

For *in vitro* binding analyses, 50 µl of GST or GST-MOS11 proteins on glutathione- Sepharose 4B beads (50% slurry) were incubated on a rotating device with 10 µl of each *in vitro* synthesized CaMV protein in a final volume of 200 µl. The beads were then washed three times, pelleted at 500 x *g* for 5 min, and boiled in SDS-PAGE sample buffer. GST and GST-MOS11 bound and unbound fractions were analysed by SDS-PAGE and autoradiography.

EMSA

L-VII and I-II RNAs were *in vitro* produced by T7 RNA Polymerase (Thermo Scientific) from 1 µg *BamH*I linearized Litmus28i, according to the manufacturers instructions and in presence of 20 µCi [α-^{32}P]-UTP. Radiolabelled RNAs were purified by acidic phenol/chloroform treatment and isopropanol precipitation, and suspended in 50 µl RNase-free water with 20 U Ribolock RNase Inhibitor (Thermo Scientific).

2 µl (≈ 5,000 cpm) radiolabeled RNAs were incubated for 60 min at 25°C with increasing amounts (0 - 400 pmole) of purified His_6-ALY1-GB1 or His_6-GB1-eYFP, in a final volume of 20 µl protein storage buffer. 4 µl non-denaturing loading buffer 6x (30% (v/v) glycerol – 0.3% bromophenol blue – 0.3% xylene cyanol) were then added and protein-RNA complexes were analysed by non-denaturing 45 min gel electrophoresis at 100 V (TBE 0.5x, pH 10.5 - 4% (v/v) acrylamide-bisacrylamide (37.5:1) - 4% glycerol (v/v)), with a 30 min pre-equilibration run at 100 V in TBE 0.5 (pH 10.5), and subsequent autoradiography.

Fluorescence localisation analysis

Fluorescent Arabidopsis protoplasts, transfected with expression plasmids coding eGFP and/or mRFP or proteins fused to them, were observed at 63-fold magnification between a poly-lysine slide and cover slip, with a silicone spacer in-between, with a Zeiss LSM780 confocal microscope (Jena, Germany). eGFP and mRFP were viewed by excitation, respectively, at 488 nm and 561 nm, with an argon laser using an appropriate emission filter to collect the green or red signal from the optical section.

RESULTS

Assessment of nuclear export of CaMV 35S RNAs in Arabidopsis WT and *CA-rop2* protoplasts

In order to measure the 35S RNA nuclear export we set up a protoplast transfection assay. Protoplasts were prepared from WT Arabidopsis seedlings and transfected with the 14 kbp pCaMV-GFP plasmid (55), which contains an infectious CaMV genome under the control of the 35S promoter and a reporter GFP in a separate expression cassette for monitoring transfection efficiency. It was previously shown that upon transfection, viral protein P6 that is early expressed from 19S RNA, independently of, and before the other viral proteins (whose expression depends on the transactivation activity of P6), is first detected at 15 h post-transfection (hpt) (55). P2 protein is detected about 18 hpt, and capsid protein P4 and the newly synthetized and encapsidated viral DNA, the presence of which indicates that reverse transcriptase P5 is also synthesised and active, are detected 21 hpt, colocalising with P6 in inclusions that correspond to viral factories (55). Taken together, these observations suggest that the time needed for one round of replication of CaMV in transfected protoplasts is about 21 h (55).

The transfection efficiency of the 14 kbp viral genome encoding plasmid in protoplasts prepared from WT Arabidopsis seedlings was however not always high enough to allow an accurate quantification of CaMV RNAs by real-time RT-PCR. On the contrary, protoplasts prepared from Arabidopsis *CA-rop2* seedlings were systematically and reproducibly transfected with high efficiency with pCaMV-GFP. *CA-rop2* plants express the constitutively active GTPase ROP2 (Rho-related GTPase from plants), closely associated with plasma membrane (66) and shown to control multiple developmental processes such as seed dormancy, shoot apical dominance, lateral roots initiation etc. (52). To our knowledge, it has never been shown that ROP GTPases might influence the nuclear export of mRNAs. However, in human HeLa cells, inhibition of the PI3 kinase/AKT pathway was shown to either increase the nuclear export of a subset of poly(A) spliced or intronless mRNAs, or having no effect on the export of some other candidate mRNAs (67). It is yet well established, that PI3K/AKT signalling regulates mammalian TOR (target-of-rapamycin) kinase complex 1 (mTORC1) by activating a GTPase, Rheb, which interacts with, and activates mTOR (68). On the other hand, we have previously shown that Arabidopsis active ROP2 GTPase directly associates with, and upregulates TOR, which then promotes reinitiation after uORF translation (69). Arabidopsis TOR can also be recruited and activated by CaMV P6 protein, and this culminates in a strong stimulation of translation reinitiation of CaMV 35S polycistronic RNA (70).

To rule out the possibility that the active ROP2 GTPase might somehow influence the nuclear export of viral mRNAs, we compared CaMV 35S RNA nuclear export in WT and *CA-rop2* protoplasts. Twenty hpt, RNAs were isolated from total, cytoplasmic and nuclear fractions and quantified by RT-qPCR using primers targeting the P5 coding region present on all 35S RNA isoforms. As shown, there is no significant difference neither in the accumulation of total 35S RNAs (**Figure 1A**), nor in their nucleo-cytoplasmic partitioning (**Figure 1B**) in WT and *CA-rop2* protoplasts, and these findings suggest that the nuclear export of mRNAs (viral and cellular, since the Ct values of the set of common housekeeping gene mRNAs used for normalisation also remained unchanged) is not affected in *CA-rop2* protoplasts.

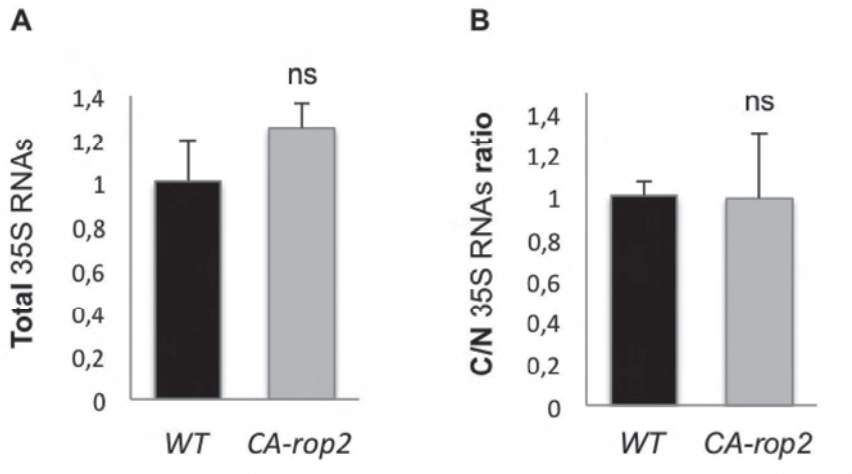
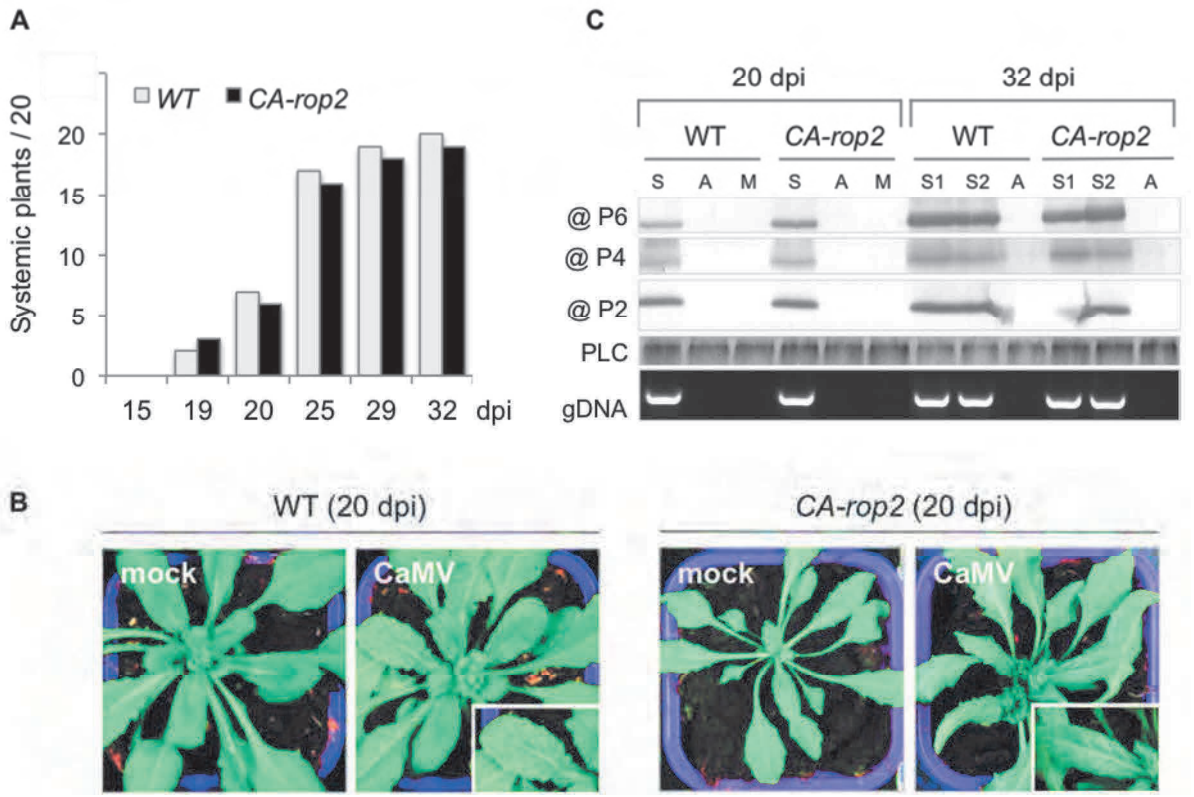


Figure 1. Total 35S RNA accumulation and nucleo-cytoplasmic partitioning are similar in WT and *CA-rop2* protoplasts. Leaf mesophyll protoplasts were prepared from 10 days-old WT and *CA-rop2* Arabidopsis seedlings and transfected with 10 µg pCaMV-GFP. Twenty hpt, the relative levels of 35S mRNAs were quantified in total (**A**), or nuclear and cytoplasmic fractions (**B**) by P5-specific real-time RT-PCR and normalised to the cellular housekeeping gene GAPDH and to GFP (as transfection indicator). Three independent experiments were performed in triplicate, each with 3 technical replicates. The results are expressed as the mean fold change ± SEM of total 35S RNAs (**A**), or of the 35S RNA cytoplasmic-to-nuclear ratios (C/N) (**B**). Significance was tested with a Student's unpaired 2-samples t-test (ns, not significant, $p>0.05$).

In order to investigate whether not only CaMV mRNAs nuclear export but also infection by CaMV occurs similarly in WT and *CA-rop2* plants, we inoculated mechanically with CaMV 20 Arabidopsis plants (WT or *CA-rop2*). As shown in **Supplementary Figure 1**, there was no delay in the infection kinetics, the 2 types of plants becoming systemically infected from day 19 (2 WT

and 3 *CA-rop2* symptomatic Arabidopsis*)* up to day 32 (20 WT and 19 *CA-rop2* symptomatic plants) (**SF1A**), and displaying the same symptoms (typical vein-clearing and leaf-crinkling also visible on the characteristic pointed and serrated *CA-rop2* leaves) (**SF1B**) and the same accumulation of viral proteins and genomic DNA (**SF1B**). We can notice that viral proteins and DNA were detected only in symptomatic (S) but not in asymptomatic (A) or mock-inoculated plants. These CaMV infection results were somehow surprising when considering the fact that in a context of a constitutively active ROP2 GTPase, TOR kinase, which is also activated (69), is expected to stimulate the reinitiation of translation of the 35S polycistronic RNAs and thus to promote the viral life cycle in plants. According to our results, this is clearly not the case, probably because, in order to transactivate CaMV translation, TOR needs to interact with viral protein P6, which is synthesised by canonical translation from its own 19S RNA. As we can see (**SF1C, top panel**) P6 is not overexpressed in *CA-rop2* plants compared to WT, and might represent the limiting factor. Moreover, the main population of P6 is involved in homoassociation, which allows the formation of the electron-dense aggregates used as viral factories (71) whereas only a minor and soluble sub-population of P6 acts in translation transactivation (72).



Supplementary figure 1. Wild type and *CA-rop2* Arabidopsis display the same susceptibility to CaMV infection. Twenty WT or mutant plants were mechanically inoculated with CaMV crude extract and followed for symptoms appearance (**A, B**) and viral proteins and DNA accumulation (**C**). The results shown are representative of 3 independent experiments. (**B**) Symptoms expressed, or not, by WT and *CA-rop2*, CaMV- or mock-inoculated plants 20 dpi, and enlargement of a systemic leaf (on the right bottom). (**C**) 20 and 32 dpi, viral proteins P6 (**top panel**), P4 (**2nd panel**) and P2 (**3d panel**) were revealed by western blot of crude protein extracts from 1 or 2 symptomatic (S), 1 asymptomatic (A) and 1 mock-inoculated (M) WT or *CA-rop2* plant. Viral DNA was detected by PCR amplification of a 775 bp fragment from the P5 encoding sequence on total DNA extracted from different leaves of the same symptomatic, asymptomatic and mock-inoculated plants used above (**bottom panel**). PLC: protein loading control; dpi: days post infection.

Taking in consideration the results described above, we decided us to use as controls in the nuclear export experiments, both WT (often only poorly transfected) and *CA-rop2* (usually transfected with a better efficiency) Arabidopsis protoplasts. All independent experiments that are presented below, were, unless mentioned otherwise, performed at least three times and included at least once WT Arabidopsis as control.

Arabidopsis TREX components TEX1, MOS11 and ALYs are required for nuclear export of CaMV 35S RNAs

The Arabidopsis TREX export complex is composed of the hexameric THO sub-complex (TEX1 being one of its components) associated with the adaptor proteins MOS11 and ALY1-4 (29). Compared to the WT control, homozygous *mos11*, *mos11 tex1* and *4xaly* mutants display significant nuclear accumulation of cellular poly(A) RNAs, which is indicative of mRNA export defects (32, 35, 53). We therefore analysed the nuclear export of CaMV 35S RNAs in transfected protoplasts prepared from the above-described Arabidopsis mutants, as well as from two additional ALY double-mutants, *aly1 aly2* and *aly3 aly4* (32), in comparison with WT and *CA-rop2* controls. To this end, transfected WT and mutant protoplasts were subjected to subcellular fractionation 20 hpt. The levels of cytoplasmic and nuclear 35S RNAs were the determined by real-time RT-PCR (**Figure 2A, 2C**). A major difficulty was the low transfection efficiency of almost all TREX mutant protoplasts, compared to *CA-rop2* and WT, and the fact that, in the initial experiments, the cytoplasmic-to-nuclear ratio of the mRNAs of five different commonly used housekeeping gene that we tested for normalisation, as well as the one of the transfection indicator GFP, were significantly changed in the quadruple *4xaly* mutant. To overcome this problem, we used improved quantification for measuring the nuclear and cytoplasmic fractions of the 35S RNAs by normalisation to the mean value of all Cts obtained, respectively, for nuclear or cytoplasmic GAPDH and GFP mRNAs in WT, *CA-rop2* and the 4

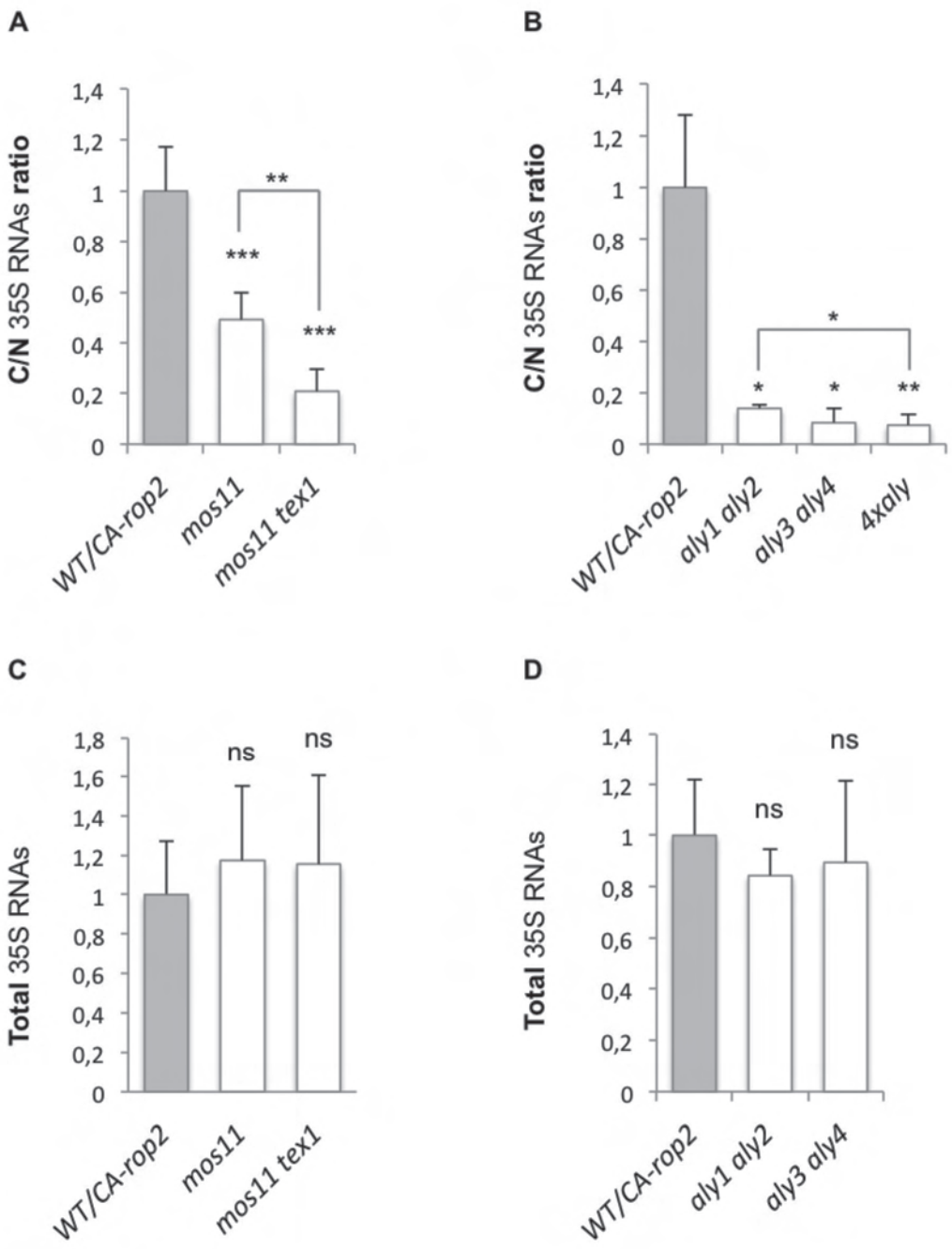


Figure 2. Nuclear export of 35S RNAs is inhibited in TREX export mutants. Leaf mesophyll protoplasts were prepared from 10 days-old WT, *CA-rop2, mos11*, *mos11 tex1, aly1 aly2*, *aly3 aly4* and *4xaly* Arabidopsis seedlings and transfected with 10 µg pCaMV-GFP. Twenty hpt, the cytoplasmic-to-nuclear ratios (**A, B**) or the total levels (**C, D**) of 35S mRNAs were determined by P5-specific real-time RT-PCR, as described in Figure 2. Ct values for *4xaly* where normalised to the mean Ct calculated for cytoplasmic or nuclear GAPDH and GFP in the other mutant and WT protoplasts. Three independent experiments were performed in triplicate. The results are expressed as the mean fold change ± SEM. Significance was tested with a Student's unpaired 2-samples t-test (ns, not significant, $p>0.05$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$).

other non-affected mutants. In MOS11-deficient cells, the cytoplasmic-to-nuclear ratio was reduced by half indicating a significant nuclear accumulation of 35S RNAs and hence an inhibition of the nuclear export (**Figure 2A**). This inhibition was significantly increased in the double mutant *mos11 tex1* and also in ALYs double mutants *aly1 aly2* and *aly3 aly4*, and especially in the quadruple *4xaly* (**Figure 2B**). Surprisingly, *aly3 aly4* and *4xaly* displayed almost the same levels of retention of nuclear 35S RNAs which is more important than in *aly1 aly2* suggesting that ALY3 and 4 might be more important for the export of viral mRNAs.

We also investigated whether the accumulation of total 35S RNAs was affected in the five export mutants. Unfortunately, we did not get enough transfected *4xaly* protoplasts to measure the total 35S RNAs. As we can see in **Figure 2C, 2D,** even if the viral RNA levels appear increased in *mos11* and *mos11/tex1*, these differences are not statistically significant indicating that the TREX export mutants have if any, only very limited effect on primary viral transcription.

Taken together, our data clearly establish the essential contribution of the TREX export complex to the nuclear export of CaMV 35S RNAs.

TREX components-deficient plants are partially resistant to CaMV infection

To test whether the export inhibition of CaMV 35S RNAs, observed in TREX components mutant Arabidopsis, affects the viral life cycle, we studied susceptibility to viral infection of the same TREX-deficient plants. Appearance of symptoms and virus replication kinetics (**Figure 3** and **Supplementary Figure 2**) in 20 or 30 mechanically inoculated WT or mutant plants, were compared in at least three independent experiments. It should be noticed that, even if the general course of CaMV infection was very similar among the different experiments, depending on the plants physiology and the particular infectious sap used, systemic symptoms first appeared at different points in time for WT/*CA-rop2* Arabidopsis: 14 – 17 – 18 or 19 dpi (**Figure 3** and **Supplementary Figure 1**). According to Sørensen *et al.* (53) and Pfaff *et al.* (32) we noticed that double-mutant *mos11 tex1* and especially quadruple *4xaly* were phenotypically severely affected, with reduced size of leaves and rosette diameter and elongated hypocotyls at the seedling stage (**Figure 3B, 3E**).

In this experiment, the first systemic WT Arabidopsis appeared 14 dpi whereas *mos11* plants displayed a 5 days delay and only 10% became systemic 19 dpi, when already 60% of the WT were symptomatic (**Figure 3A**). 23 dpi the WT/*mos11* ratio was 90%/25%, 95%/50% 30 dpi and reached 100%/60% 39 dpi. In all 3 independent infection experiments *mos11* plants always became systemic later then WT and only around 60% of the inoculated plants displayed

symptoms and were infected. As shown in **Supplementary Figure 2A**, viral proteins and genomic DNA were detected only in symptomatic but not in asymptomatic or mock-inoculated WT and *mos11* plants, and the same observation was made for the other 4 mutants. Almost the same observations were made with the double-mutant *mos11 tex1* (**Figure 3B**) in 4 independent experiments. The kinetics of infection presented in Figure 3B is somehow slightly different since WT plants became systemic later than in the experiment described above (17 dpi) ant the first infected *mos11 tex1* were visible the day after (18 dpi) but to a significantly lesser extent (5% vs 35% for the WT). 31 dpi 75% of WT Arabidopsis were systemic and this ratio slightly increased to 80% 39 dpi, when only 45% of the double-mutant were infected by CaMV. The end ratio (39 dpi) of systemic *mos11 tex1* varied between 45% and 55% in the different experiments and was fairly similar to that of *mos11* indicating that even if 35S RNAs nuclear export is less efficient in the double-mutant, the viral life cycle doesn't seem to be more affected in *mos11 tex1* than in *mos11*. Thus, our results show that when the nuclear export of viral mRNAs is twice less efficient in the absence of MOS11 and/or TEX1 (Figure 3B), the replication cycle of CaMV is blocked in half of the inoculated plants suggesting the importance of these components of the TREX export complex for the nuclear export of viral mRNAs and for the outcome of CaMV infection.

The situation was surprisingly different with the ALY adaptor protein mutants (**Figure 3C, 3D, 3E**). Compared to WT, double-mutants *aly1 aly2* (**Figure 3C**) and *aly3 aly4* (**Figure 3D**) did not show any delay in symptoms appearance in any of the 3 independent experiments, but remained reproducibly less infected at each time point, with an end ratio (39 dpi) of about 80% compared to the 97% of systemic WT plants. Infection by CaMV was however severely hampered in the quadruple mutant *4xaly* (**Figure 3E**) with roughly 17% of systemically infected plants 39 dpi versus 93% for the WT. Among all TREX-deficient mutants, *4xaly* was the one in which exhibited the most important inhibition of the nuclear export of CAMV mRNAs (**Figure 2B**) and it therefore not surprising that this significant nuclear retention of viral mRNAs blocks the viral life cycle in the large majority of the inoculated plants.

Considering the fact that the nuclear export of 35S RNAs was also strongly inhibited in *aly1 aly2* and *aly3 aly4* double-mutants, even if to a lesser extent than in *4xaly* (**Figure 2B**), one might expect that the 2 double-mutants would be also less susceptible to CaMV infection. Our results however show that this is definitely not the case thus suggesting that the inhibition of viral infection in *4xaly* might be due not only to the very poor export of viral mRNAs out of the nucleus but also to a general nuclear retention of cellular mRNAs (as observed for the housekeeping genes) some of which might encode important host factors.

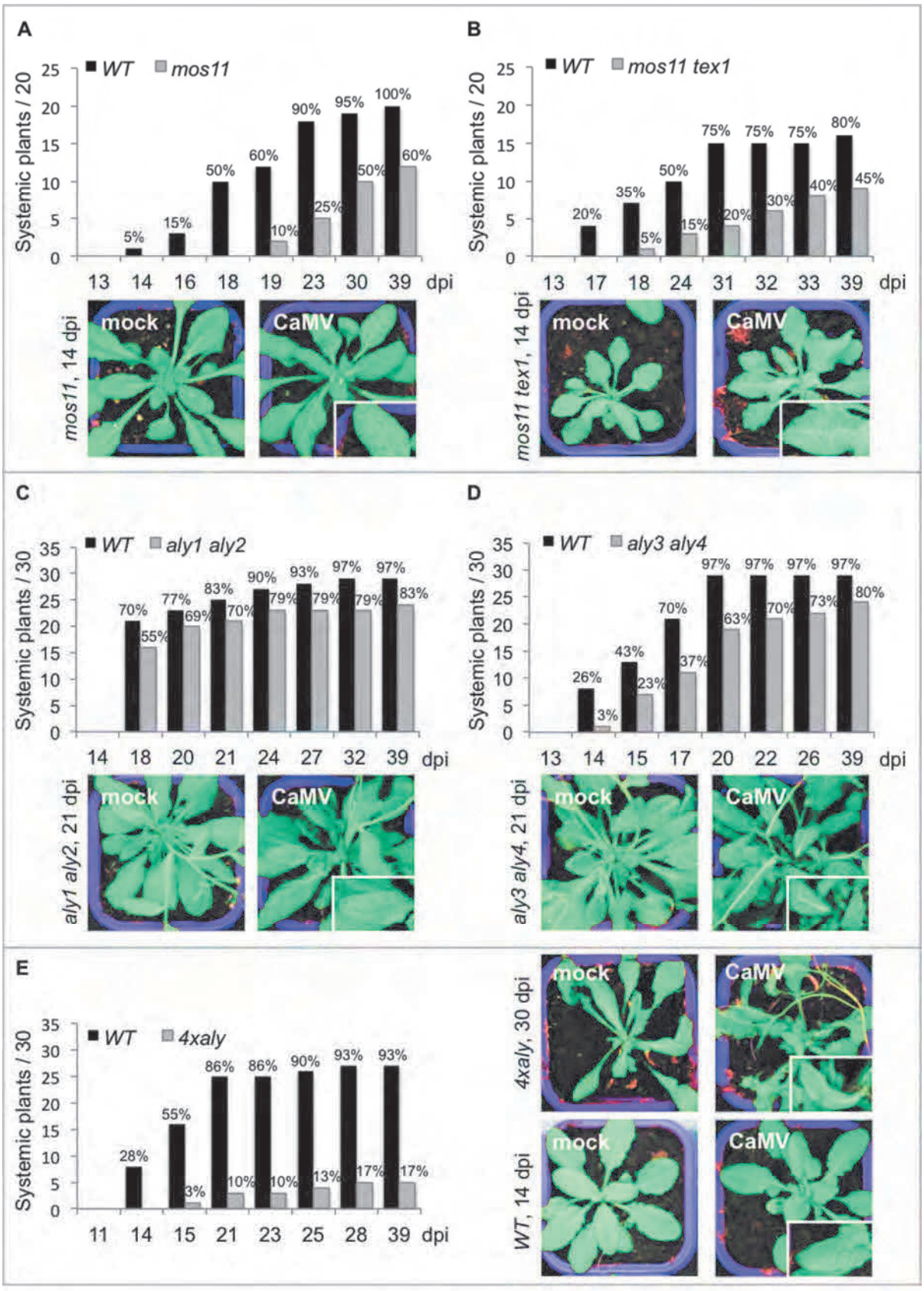
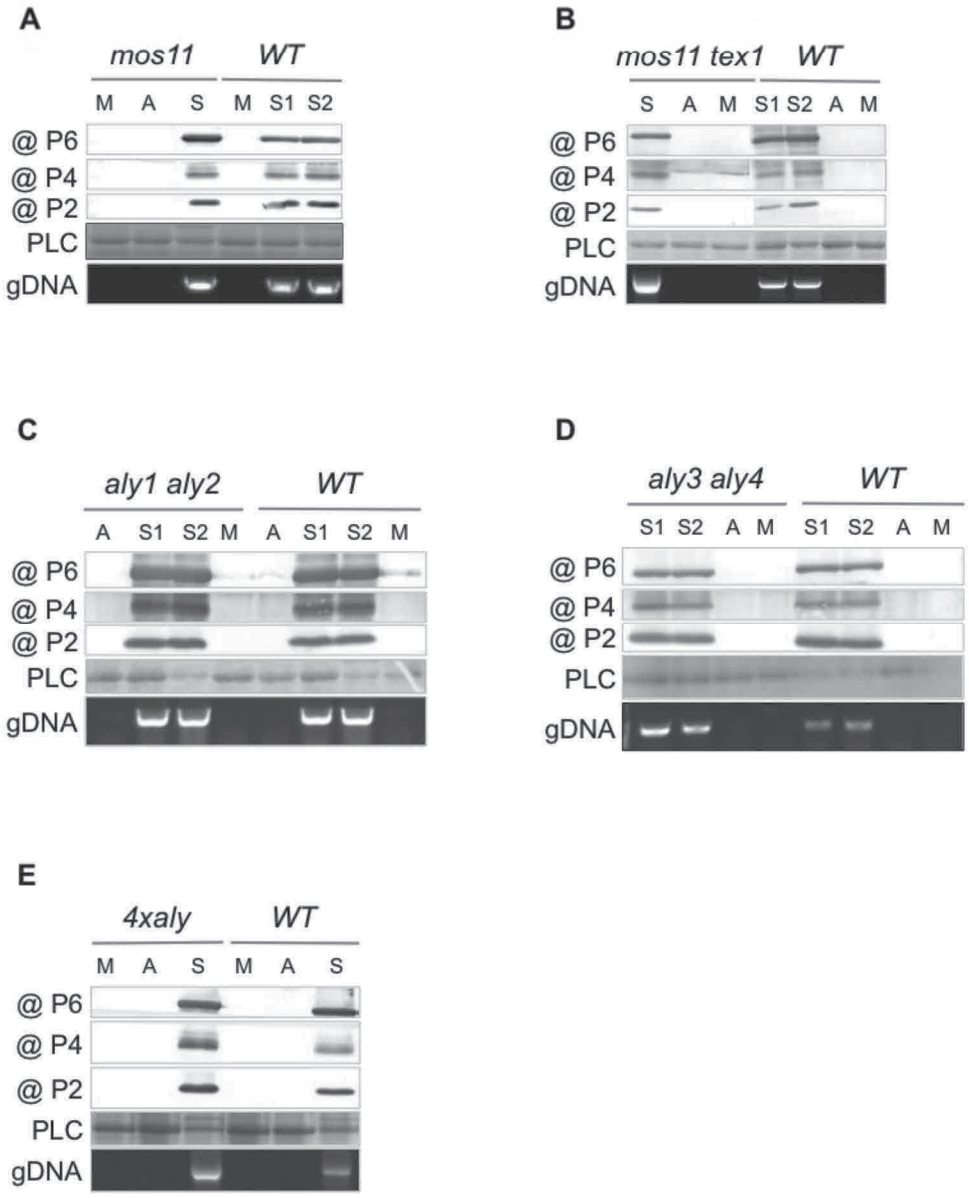


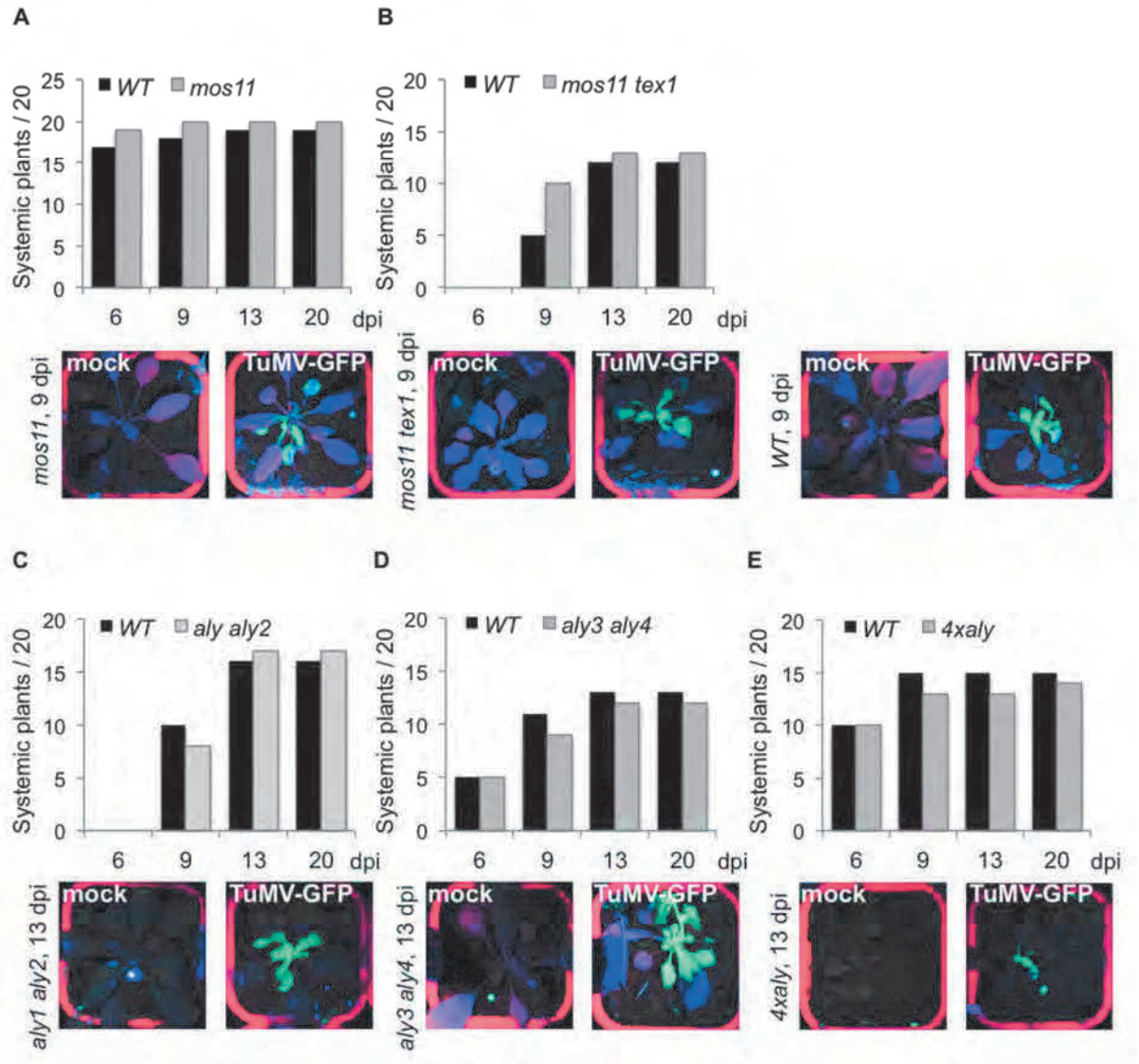
Figure 3. TREX-deficient plants are partially resistant to CaMV. 20 (**A, B**) or 30 (**C, D, E**) WT, *mos11* (**A**), *mos11 tex1* (**B**), *aly1 aly2* (**C**), *aly3 aly4* (**D**) and *4xaly* (**E**) plants were mechanically inoculated with CaMV crude extract and followed for symptoms appearance. The results are representative of 3 (*mos11*,

aly3 aly4 and *4xaly*) or 4 (*mos11 tex1, aly1 aly2*) independent experiments. The percentages of infected WT and mutant plants at each time point are indicated. Representative mock- or CaMV-inoculated systemic plants with enlargement of a symptomatic leaf (on the right bottom) are shown below (*mos11*, *mos11 tex1, aly1 aly2* and *aly3 aly4*) or beside (*4xaly*, WT) the corresponding histograms.



Supplementary figure 2. CaMV proteins expression and genomic DNA replication in inoculated WT and TREX mutant Arabidopsis. 24 dpi viral proteins P6 (**top panels**), P4 (**2nd panels**) and P2 (**3rd panels**) were revealed by western blot of crude protein extracts from 1 or 2 symptomatic (S, S1, S2), asymptomatic (A) and mock-inoculated (M) WT, *mos11* (**A**), *mos11 tex1* (**B**), *aly1 aly2* (**C**), *aly3 aly4* (**D**) and *4xaly* (**E**) plants. Viral DNA was detected by PCR amplification of a 775 bp fragment from the P5 encoding sequence on total DNA extracted from different leaves of the same symptomatic, asymptomatic and mock-inoculated plants used above (**bottom panels**). PLC: protein loading control.

In order to confirm that the partial resistance of TREX mutant Arabidopsis to CaMV infection was mainly due to the more or less severely hampered nuclear export of viral mRNAs, we examined the susceptibility of the same TREX complex mutants to TuMV (Turnip mosaic virus), a *Potyvirus* with single stranded positive-sense RNA genome and strictly cytoplasmic replication cycle (73). In the experiments we used a recombinant TuMV encoding a GFP tag (TuMV-GFP)(54) the expression of which is easily followed under UV light and reflects viral replication. As shown in **Supplementary Figure 3,** most TuMV-GFP infections were detected 6 dpi (**SF 3A, 3D, 3E**), indifferently on WT and mutant plants, except those on *mos11 tex1* (**SF 3B**) and *aly1 aly2* **(SF 3C**) and their respective WT controls, for which systemy became visible 9 dpi. TuMV-GFP infection expanded more rapidly than this of CaMV since themaximal number of systemic plants was already reached 13 dpi (**SF 3**).



Supplementary Figure 3. WT and TREX-deficient plants display the same susceptibility to TuMV-GFP. 20 WT, *mos11* (**A**), *mos11 tex1* (**B**), *aly1 aly2* (**C**), *aly3 aly4* (**D**) and *4xaly* (**E**) plants were mechanically inoculated with TuMV-GFP crude extract produced on *Brassica napus* cv. Drakkar, and followed for symptoms appearance and the expression of GFP under UV light. The results are representative of at least 3 independent experiments. Representative mock- or TuMV-GFP inoculated and GFP expressing, or not, plants, are shown below the corresponding histograms.

There were also no significant differences in the ratios of infected plants between WT and the 5 assessed simple, double and quadruple mutants, the presented end ratios (20 dpi) varying by ±1 plant (**SF 3**). Our results clearly show that TuMV-GFP replication cycle was not affected in any of the TREX mutants we tested, which is not really surprising knowing that all viral RNAs are synthesised in the cytoplasm and don't require nuclear export. On the other hand, these observations also suggest that even in the most severely hampered export mutant, *4xaly*, the cellular mRNAs encoding host factors essential for TuMV continue to be somehow exported out of the nucleus, possibly *via* the ALYs analogues UIFs (33), thus allowing the viral replication to occur in an unchanged manner. We can, of course, not exclude the possibility that in the case of CaMV different cellular factors are strictly needed and that their mRNAs remain retained in the nuclei, in the same way as the viral ones, and that the lack of both the viral and some host proteins inhibits the infection spread. It appears however very likely that the main reason for the incapacity of CaMV to infect most *4xaly* plants (**Figure 3E**) is the strong inhibition of the nuclear export of its 35S RNAs.

Having shown that the nuclear export of CaMV 35S RNAs involves the TREX complex and especially its components TEX1, MOS11 and ALYs, and that their absence reduces more or less strongly the viral infection, we then decided to investigate whether these export factors are directly recruited on viral RNAs or whether they necessitate the participation of some viral proteins.

CaMV 35S RNAs nuclear export depends at least on the expression of viral protein P5

As mentioned previously, mammalian viruses, like Herpesviruses, hijack the cellular mRNA export pathway by the expression of viral-specific export adaptors that serve to recruit TREX and the export receptors TAP/p15 to viral mRNA. The most striking example is KSHV ORF57 protein, which interacts with several TREX components including ALY (39), UIF (41) and CIP29 (42).

In order to explore if CaMV proteins are involved in the nuclear export of viral mRNAs, we mutagenized the CaMV genome-encoding plasmid (pCaMV-GFP) by creating 2 STOP codons in the beginning of the P6 coding sequence (each of them downstream of the AUG, at positions 9 and 15, respectively). Without expression of P6, synthesised from the monocistronic 19S RNA, translation reinitiaton on the polycistronic 35S RNA is completely abolished and no viral protein can then be expressed (74–77). We also generated 2 additional mutant genomes, the double-mutant pCaMV-P5/P6_STOP, in which the AUG codon at position 17 in the P5 coding sequence was also changed to a STOP codon, and the single-mutant pCaMV-P5_STOP, in which the expression of P5 only is abolished.

These plasmids were then transfected into WT/*CA-rop2* protoplasts and we quantified 20 hpt the total, cytoplasmic and nuclear 35S RNAs in order to determine the export ratios (**Figure 4A**). In the absence of expression of viral proteins, as evidenced by the detection of capsid protein P4 (**Figure 4C**) only with pCaMV-P5_STOP (**lane 4**) but not with pCaMV-P6_STOP (**lane 2**) or pCaMV-P5/P6_STOP (**lane 3**), the nuclear export was strongly inhibited (pCaMV-P6_STOP, **bar 2**, and pCaMV-P5/P6_STOP, **bar 3**). Exactly the same result was obtained when only one of the viral proteins, protease/reverse transcriptase P5, was not present (pCaMV-P5_STOP, **bar 4**) meaning that P5, at least, might act as a viral-specific adaptor in the nuclear export of 35S RNA. We next assessed, by cotransfection experiments (**Figure 4A**), whether a P5 protein expressed *in trans* might restore the nuclear export, and demonstrated that only a myc-tagged uncleaved full-length protein (myc-P5_FL, **bar 6,** and **lane 6** in **C** (78 kDa)) but neither the Nt protease domain (P5_PR, **bar 7**) nor the reverse transcriptase / RNaseH domain (P5_RTRH, **bar 8**) is able to complement the missing viral export factor. At the same time, as shown in **Figure 5B**, the total amounts of 35S RNAs, determined by RT-qPCR with P4-specific primers which anneal only on 35S RNAs but not on the cotransfected complementing P5-encoding plasmids, were not significantly different in the 8 transfection and co-transfection conditions. This observation suggests that the inhibition or the complementation of the nuclear export of 35S RNAs, respectively, in absence or in presence of P5, is not due to a different level of RNA accumulation.

To further study the role of P5 in the export of viral RNAs we investigated its capacity to interact with some of the components of the TREX complex.

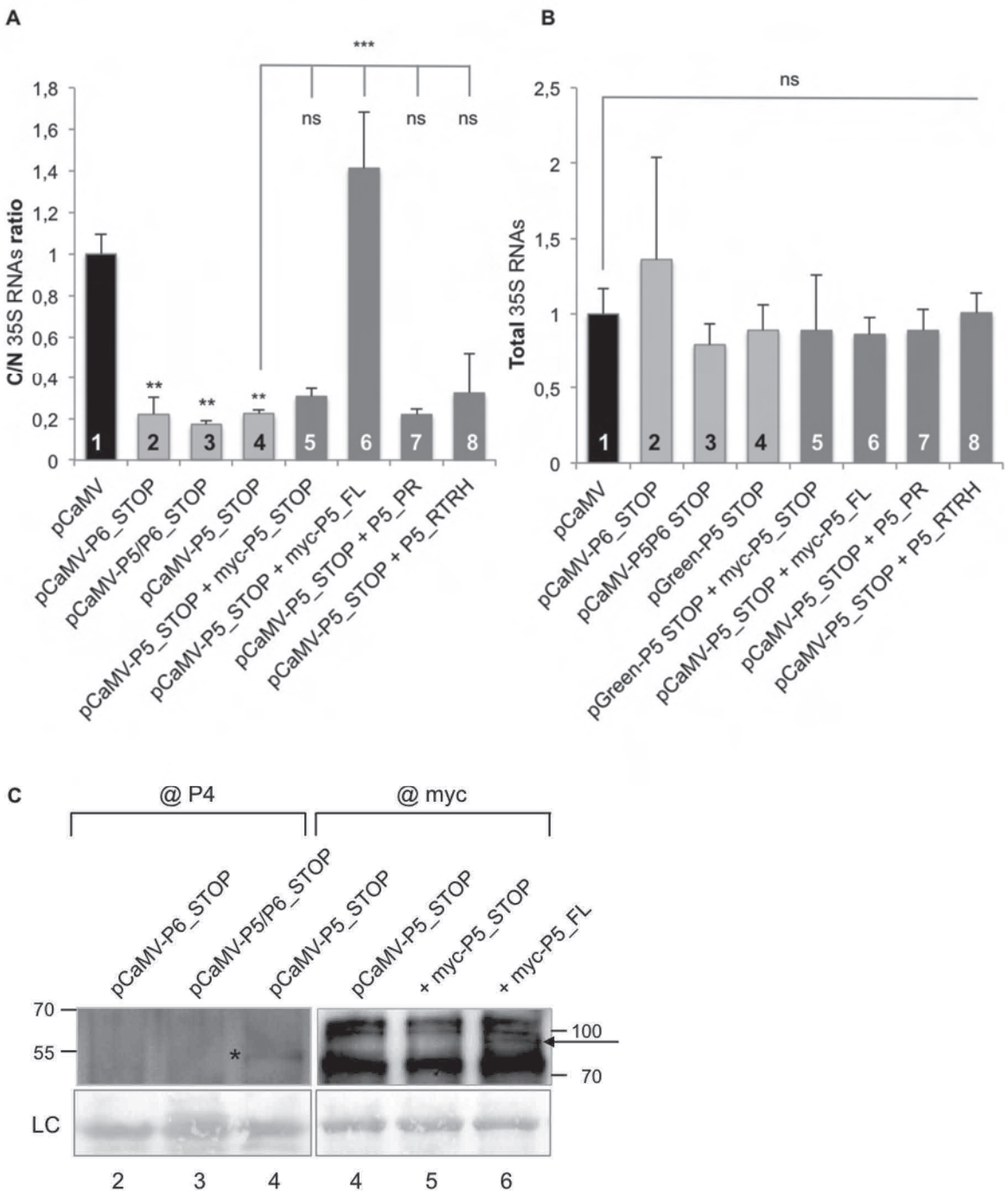


Figure 4. Nuclear export of 35S RNAs is inhibited in the absence of viral protein P5. Leaf mesophyll protoplasts were prepared from 10 days-old WT/*CA-rop2* Arabidopsis seedlings and transfected with 10 µg pCaMV-GFP (**1**), pCaMV-P6_STOP (**2**), pCaMV-P5/P6_STOP (**3**) or pCaMV-P5_STOP (**4 - 8**). In the 3 STOP plasmids (in light grey) one (for P5) or two (for P6) STOP codons were introduced after the START AUG of viral proteins P5, P6 or both. In **5 to 8** (dark grey) pCaMV-P5_STOP was co-transfected with a plasmid supplementing *in trans* with myc-P5_STOP (**5**), full-length myc-P5 (myc-P5_FL, **6**), P5 protease domain (amino acids (aa) 1-201, P5_PR, **7**) or P5 reverse transcriptase/RNaseH domain (aa 202-680, P5_RTRH, **8**). Twenty hpt, the cytoplasmic-to-nuclear ratios (**A**) or the total levels (**B**) of 35S mRNAs were determined by P4-specific real-time RT-PCR and normalised to the cellular housekeeping gene GAPDH, to GFP (transfection indicator for pCaMV-P5_STOP) and to mRFP (transfection indicator

for the P5-expressing plasmids). Three independent experiments were performed in triplicate. The results are expressed as the mean fold change ± SEM of the 35S RNA cytoplasmic-to-nuclear ratios (C/N) (**A**) or of total 35S RNAs (**B**). Significance was tested with a Student's unpaired 2-samples t-test (ns, not significant, p>0.05, ** p<0.01, *** p<0.001). (**C**) Detection of P4 (pointed by the asterisk) and myc-P5 (pointed by the arrow) 20 hpt by western blot with, respectively, anti-P4 or anti-myc antibodies in protoplasts transfected with pCaMV-P6_STOP (**2**), pCaMV-P5/P6_STOP (**3**), pCaMV-P5_STOP (**4**), pCaMV-P5_STOP + myc-P5_STOP (**5**), pCaMV-P5_STOP + myc-P5_FL (**6**), corresponding, respectively, to bars 2, 3, 4, 5 and 6 in (A). LC: loading control. Positions of molecular size standards, in kilodaltons, are shown on the left and on the right.

TREX component MOS11 interacts with CaMV proteins P4 and P5

During the viral replication cycle P5 is synthesized as an inactive polyprotein precursor of 78 kDa which progressively accumulates in the viral factories where it is activated by self-cleavage by its Nt protease domain to yield a 20 - 22 kDa aspartic protease (PR) and a 56 - 58 kDa reverse transcriptase / RNaseH (RTRH) (49, 55, 78). In this context, we produced by *in vitro* translation (**Figure 5A**) radiolabelled full-length P5 (FL) and, according to Torruella *et al.* (49), its different functional domains: protease (PR), reverse transcriptase (RT), RNaseH (RH) and reverse transcriptase / RNaseH together (RR). We also included viral proteins P3 (15 kDa) and P4 (56 kDa), and firefly Luciferase (Luc, ≈ 70 kDa) as control. In this *in vitro* translation conditions, P5_FL is obviously synthesised as an uncleaved polyprotein precursor, since only an about 78 kDa product is detected upon autoradiography (**Figure 5A, P5_FL**) and none of the 20 kDa and 58 kDa potential cleavage products. Regarding the P4 capsid protein, this one is also synthesised during viral infection as a 56 kDa precursor (pre-CP) which is further partially processed during morphogenesis in the viral factories by the virus-encoded protease (PR) into 4 subspecies: p42, p39, p37 and p35, all of them lacking the Nt and/or Ct domains of pre-CP and which are produced to relative amounts depending on the virus preparation (45, 79–81). As we can see on **Figure 5A**, *in vitro* synthesised P4 comigrates with 78 kDa P5_FL and probably corresponds to incompletely denatured and fast-migrating dimers of the full-length precursor (79, 80) since it is produced in the absence of P5 or any other cellular protease susceptible to cleave it. We cannot exclude the possibility that this band represents dimers of shorter P4 proteins, synthesised either by initiation on an internal AUG codon, or by premature translation arrest. The dimers were revealed with specific @ P4 polyclonal antibodies, but not in the P3 and P5 translation reactions (**Figure 5A, bottom**). Several bands of weak intensity are observed below the major translation products and might correspond to degradation or incompletely synthesised polypeptides, or to some proteins endogenously produced in the rabbit reticulocyte lysates used for the transcription/translation assays.

To determine whether the *in vitro* produced viral proteins or domains interact with components of the TREX complex, we performed GST pull-down assays (**Figure 5C - E**) with beads-bound GST or GST-MOS11 (**Figure 5B**). The viral and control proteins were mostly present in the GST unbound fraction (**Figure 5C**) except for the RT and RH domains of P5 for which a faint signal was also detected bound to GST (**Figure 5E**) and this observations suggests that there was almost no unspecific retention with GST alone. Luciferase control and P3 were both found exclusively in the GST-MOS11 unbound fraction meaning that they do not interact at all with MOS11, whereas P5 PR and RH domains were almost equally bound and unbound (**Figure 5D, 5E**) indicating a probably weak association in our assay conditions (150 mM NaCl). The situation was however different for P4, full-length P5 and P5_RR: these proteins were reproducibly and significantly (data not shown) enriched in the GST-MOS11 bound fraction, which demonstrates that they indeed interact with MOS11 (**Figure 5E**).

Surprisingly, in the GST-MOS11 bound fraction, we found 2 forms of "full-length" P5: the about 78 kDa translation product that corresponds to the unique P5_FL input (**Figure 5A, FL),** and a second polypeptide migrating slightly higher than RT and likely comigrating with the RR doublet (a second band of very weak intensity was also visible on **Figure 5A** (**RR**) where the electrophoresis conditions were not exactly the same). This observation was confirmed in all 4 GST pull-down assays we performed (varying the stringency conditions, data not shown) and suggests that the interaction of full-length P5 with MOS11 might activate P5 Nt protease domain, which then partially cleaves itself but is not retained on the GST- MOS11 beads unlike the 2nd generated cleavage product, reverse transcriptase / RNaseH (RTRH) which is still able to interact with MOS11. The cleaved PR peptide is however not recovered in the GST-MOS11 unbound fraction, either because the analysed amounts are not sufficient to reveal its presence, or because, interacting weakly with MOS11 (as suggested by the behaviour of PR alone), and in competition with FL and RTRH, it was released during the extensive washing steps. The migration differences observed for this putative cleaved RTRH and the recombinant that we produced *in vitro* (RR) may also be due to the fact that RR is slightly larger than the natural RTRH. In fact, we delimited the different functional domains of P5 according to Torruella *et al.* (49) who proposed that the protease cleavage site resides between P5 amino acids 184 and 200 and that active PR is usually detected as a 22 and 20 kDa doublet. For this reason we decided to generate a 201 aa PR (22 kDa), a 346 aa RT (38 kDa) and a 408 aa RTRH domain (53 kDa). It is although possible that the activated upon interaction with MOS11 PR generates a larger PR polypeptide and a smaller RTRH, that migrates differently. An additional, in our opinion less likely possibility, might be the capture of bacterial protease(s) by the beads or by beads-bound GST-MOS11, which then exclusively cleaves full-length P5 but neither its different functional domains nor the other viral and control proteins.

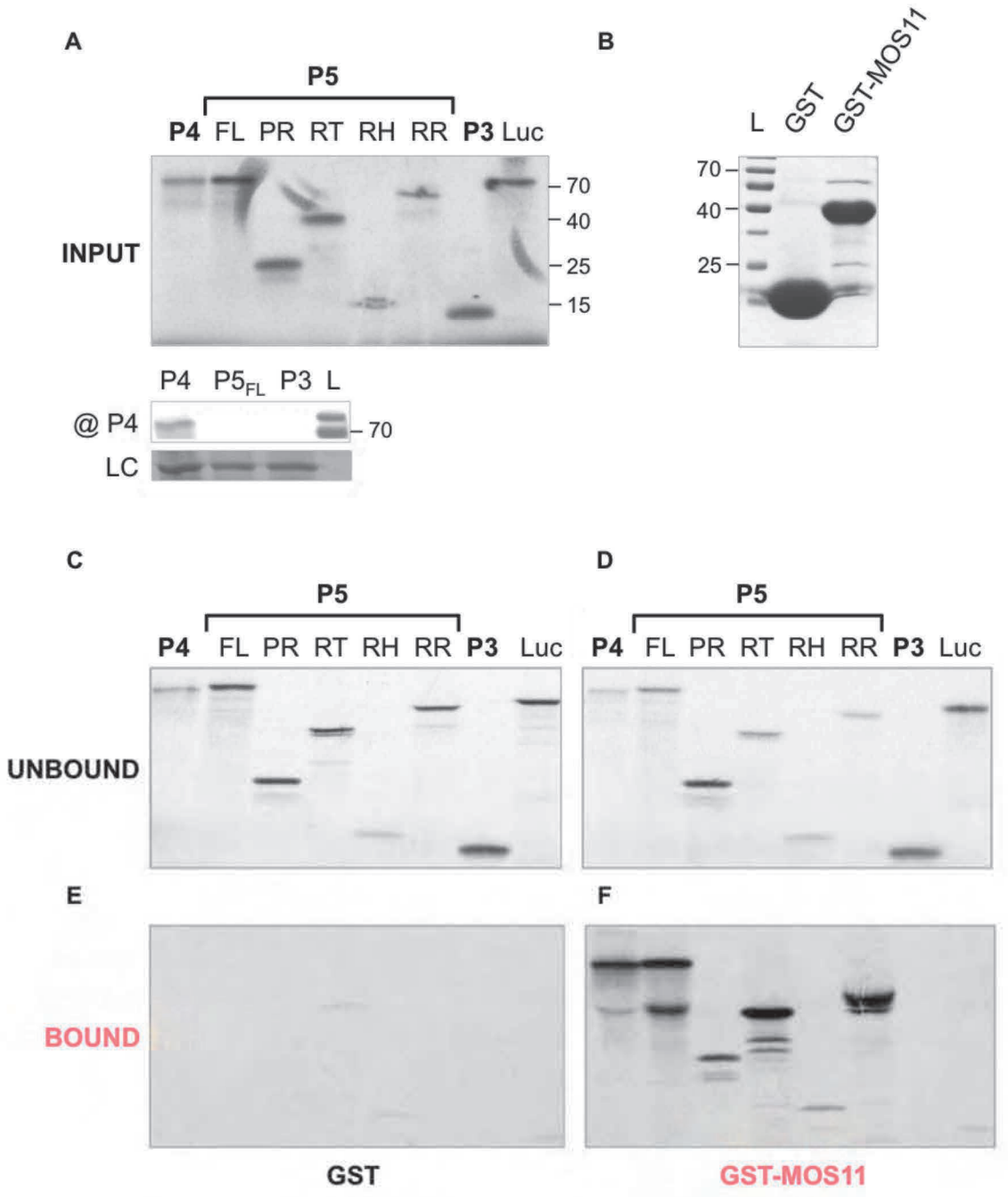


Figure 5. Viral proteins P4 and P5 interact with MOS11 *in vitro*. *In vitro* translated and [^{35}S] methionine-labelled viral proteins P3, P4 and P5 (FL: full-length, PR: protease domain (aa 1-200), RT: reverse transcriptase domain (aa 201-547), RH: RNaseH domain (aa 548-680), RR: reverse transcriptase / RNaseH domains (aa 201-608)) and firefly luciferase (Luc) as control (**A,** autoradiography) were incubated with bacterially expressed and glutathione beads-bound GST or GST-MOS11 (**B,** Coomassie blue staining). The inputs of P4, $P5_{FL}$ and P3 were subjected to a western blot with specific @ P4 antibodies (**A, bottom**). The beads were washed and the GST- and GST-MOS11 unbound (**C, D**) and bound (**E, F**) fractions were analysed by SDS-PAGE followed by autoradiography. The results are representative of 4 independent experiments. The inputs in (A) and (B) represent 1/9 of the amounts used in the GST pull-down assay (C-E). 1/10 of the total unbound fractions were analysed by SDS-PAGE (C, E). LC: loading control. Positions of molecular size standards, in kilodaltons, are shown on the left and on the right in (A) and (B).

Thus, our *in vitro* interaction results provide evidences that 2 viral proteins interact with export factor MOS11: P4 (in a dimeric form) and P5, *via* its RT domain. Moreover, the P5-MOS11 interaction likely induces the proteolytical cleavage of P5 generating a fragment that comigrates with RT / RTRH domains. These observations were somehow surprising and the biological relevance of the *in vitro* interactions needed absolutely to be addressed. Regarding P4, it is well documented that CaMV capsid proteins possess, close to their N-termini, a nuclear localization signal (NLS), which is exposed on the surface of mature virions. However, the acidic N-terminus of the capsid precursor prevents nuclear targeting and only P4 proteins deleted from their 77 Nt residues can be detected in the nuclei (45). MOS11 displaying an exclusively nuclear but not nucleolar localisation in Arabidopsis (35, 53), one might imagine that the interaction between P4 and MOS11, observed *in vitro*, might also take place *in vivo*. There was however a critical point concerning the apparent inability of full-length P4 to localize into the nucleus (45) that necessitated further investigations. Regarding P5, the situation was even more unclear since among all viral proteins, P5 definitely remains the less characterized and a putative nuclear localisation has never been clearly shown for it, even if Pfeiffer *et al.* (82) likely detected CaMV DNA polymerase activity in infected nuclear extracts.

For these reasons we decided to re-examine the subcellular localisation of viral proteins P4 and P5 in Arabidopsis protoplasts, in fusion with a fluorescent eGFP tag, and to look for a possible nuclear colocalisation with MOS11.

GFP-tagged CaMV proteins P4 and P5 colocalise each with MOS11-mRFP in cell nuclei

We addressed the question of the subcellular localisation of P4 and P5 by transfecting Arabidopsis protoplasts with plasmids expressing eGFP-P4 or eGFP-P5 under the control of a 35S promoter, and also performed co-transfections with a MOS11-mRFP coding vector (**Figure 6**). It is important to notice that eGFP-P4 was expressed in the absence of P5 and persisted in these conditions as a full-length fusion protein of about 83 kDa detected by western blot with polyclonal @ P4 antibodies (**Figure 6C, left**).

Because of its small size, eGFP (27 kDa) can enter nuclear pores, so that eGFP not fused to other polypeptides is found in both cytoplasm and nucleus in plant cells (83, 84)(**Figure 6B, panels 1**). As illustrated on **Figure 6A**, eGFP-P4 displayed 3 main distribution patterns in transfected protoplasts: 1. Either forming numerous cytoplasmic only aggregates of different sizes (**Figure 6A, line 4**), as already described by Karsies *et al.* (45) for uncleaved P4 precursor, or 2. Forming one or several cytoplasmic aggregates but also slightly accumulating in the nucleus (**6A, line 1**), or 3. Forming cytoplasmic aggregates and more or less strongly

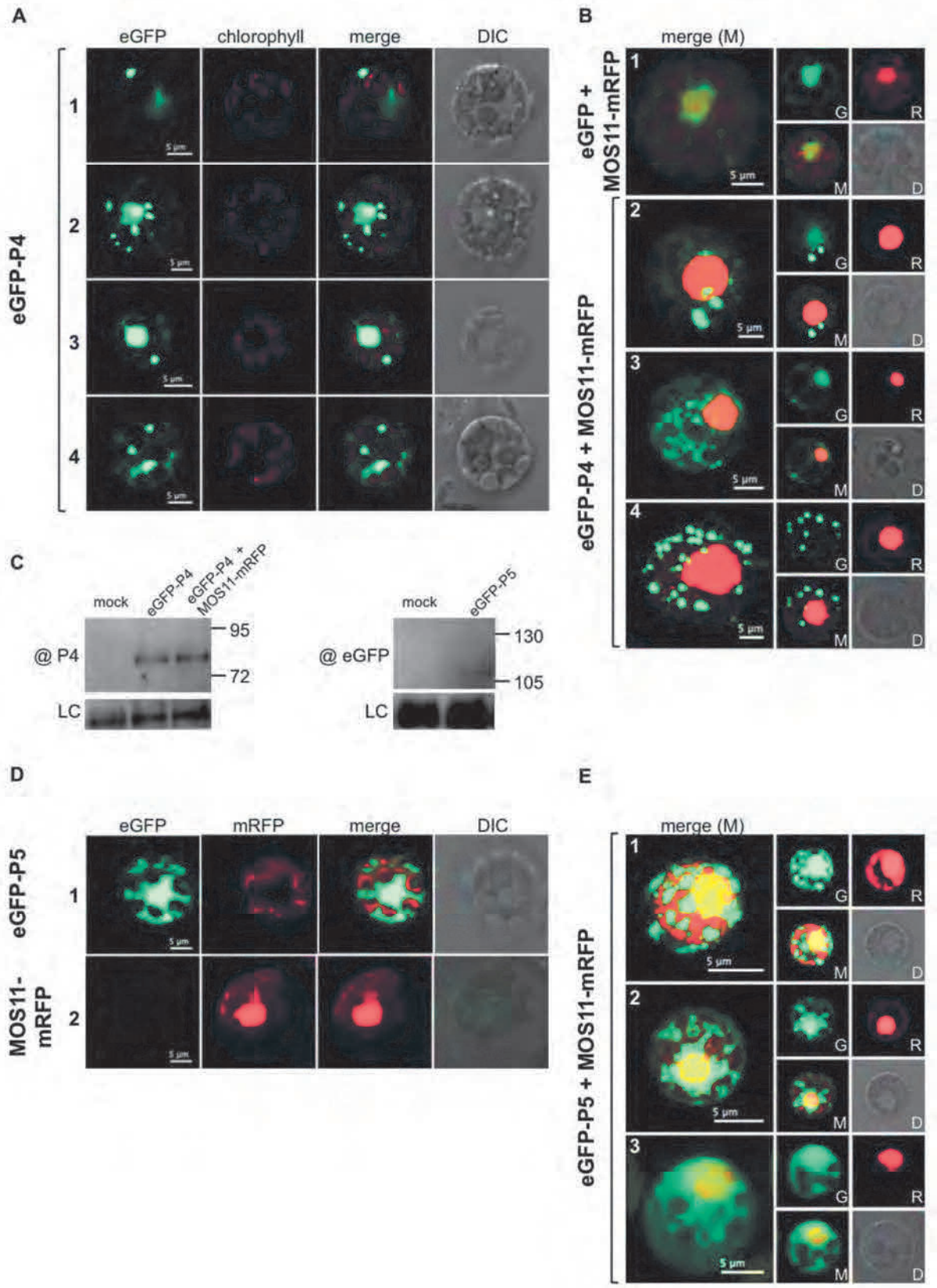


Figure 6. eGFP-P4 and eGFP-P5 colocalise each with nuclear mRFP-MOS11. Leaf mesophyll protoplasts were prepared from 10 days-old WT/*CA-rop2* Arabidopsis seedlings and transfected (or cotransfected) with 10 µg (**A**) pCK-eGFP-P4, (**B**) and (**C, left**) pCK-eGFP + Litmus-35S-MOS11-mRFP or pCK-eGFP-P4 + Litmus-35S-MOS11-mRFP, (**D**) pCK-eGFP-P5 (**line 1**) or Litmus-35S-MOS11-mRFP

(**line 2**), and (**E**) or pCK-eGFP-P5 + Litmus-35S- MOS11-mRFP. Observations of eGFP and mRFP fusion proteins were made 20 hpt by LSCM. The LSCM settings and acquisition conditions of the images were identical in all panels. In (B) and (E) panels 1 to 4, and 1 to 3, respectively are projections while all other images show single sections. G: eGFP; R: mRFP; M: merge; D: DIC (differential interference contrast images that indicate the localization of chloroplasts and nuclei). Fusion proteins eGFP-P4 and eGFP-P5 were also detected in transfected protoplasts by western blot (**C)** with anti-P4**, (left**) and anti-eGFP antibodies (**right**), respectively. Positions of molecular size standards, in kilodaltons, are shown on the right.

accumulating in the nucleus (**6A, lines 3 and 4**). In coexpression conditions with strictly nuclear MOS11-mRFP (**6B**, **panels 1; 6D, line 2**) the subpopulations of eGFP-P4 forming only cytoplasmic aggregates showed no redistribution and did not colocalise with MOS11-mRFP (**Figure 6B**, **panels 4**), whereas those accumulating more or less significantly in the nuclei clearly merged with MOS11-mRFP without loosing the cytoplasmic aggregates (**6B**, **panels 2 and 3**). These observations suggest that full-length P4, expressed in fusion with eGFP likely exists in 2 forms, an aggregates forming cytoplasmic subpopulation, which correlates with its primary role of self-associating capsid protein, and a 2nd, "soluble" subpopulation which shows a more diffuse localisation pattern and is able to enter the cell nucleus. We can, of course, not exclude the possibility that, despite the presence of the 77 Nt acidic amino acids, which are supposed to mask the close NLS (45), the Nt eGFP fusion stabilises and induces a conformational change in P4 that is responsible for the NLS exposure and the subsequent nuclear targeting. Nevertheless, our results demonstrate, for the first time to our knowledge that a subpopulation of eGFP-tagged full-length P4 localises in the nucleus where it might interact with the nuclear export factor MOS11, as it did *in vitro*.

Studying the subcellular localization of P5 was also very exciting since there is only sparse information about this viral protein. eGFP-P5 was revealed in transfected protoplasts by western blot, with anti-eGFP antibodies, as a fusion protein of about 105 kDa (**Figure 6C, right**). Expressed in fusion with eGFP, P5 exhibited a clear nucleo-cytoplasmic distribution with several cytoplasmic aggregates (**Figure 6D**, **line 1**). In coexpression conditions nuclear eGFP-P5, perfectly merged with MOS11-mRFP (**Figure 6E**, **panels 1, 2, 3**) suggesting that the interaction between the 2 proteins might occur in the nucleus.

However, unlike P4, the presence of a NLS has never been described for P5. Scanning P5 sequence with different tools for predicting general eukaryote or specific plant protein targeting to nuclei and NLSs (cNLS Mapper (85), Localizer (86), we found that a motif in the RT domain (I**K**PS**K**SPHMAPAFLVNNEAE**KRRGKKR**MVV), between amino acids 276 and 305, defines a bipartite NLS potentially responsible for a typical nucleo-cytoplasmic localisation of the harbouring protein, exactly as the one we observed for eGFP-P5 (**Figure 6C**).

Having shown that P4 and P5, 2 viral proteins that display a nucleo-cytoplasmic localisation and interact with export adaptor MOS11, we addressed the question whether there is a *cis*-acting element on the 35S RNAs, which might preferentially recruit the TREX complex on viral RNAs.

The leader region of CaMV 35S RNAs is a *cis*-acting element in nuclear export

To look for a potential *cis*-acting element on 35S RNA we performed competition assays for the nuclear export machinery between CaMV 35S RNAs and shorter RNA fragments (289 to 865 nt) corresponding to 35S 5' part (nt 1 – 3,000) or to its 3' UTR (**Figure 7A**).

For this purpose, Arabidopsis protoplasts were cotransfected with pCaMV-GFP and a 8-fold molar excess of Litmus-mRFP-35S plasmid encoding one of the five short RNA fragments: RNA L-VII corresponding to nt 1-765 of 35S RNA and encompassing the 5' UTR, or leader, and the first 157 nt of ORF VII; RNA VII-I (nt 766-1,439) overlapping ORF VII and I; RNA I-II nt (1,440-2,304) overlapping ORF I and II; RNA II-IV (nt 2,305-3,000) overlapping ORF II, III and IV; and 3' UTR corresponding to the terminal nt 7,928 – 8,216 (**Figure 7A**). Total RNAs were next isolated from nuclear and cytoplasmic, or from total fractions, and 35S RNAs were quantified by real-time RT-PCR (**Figure 7B, 7C**).

While the coexpression of RNAs VII-I, I-II, II-IV and the 3' UTR did not influence significantly the nuclear export of 35S RNAs, in presence of L-VII fragment there was a strong nuclear retention of 35S viral RNAs (**Figure 7B**) suggesting a possible hijacking of the export machinery by the overexpressed RNA fragment, mainly corresponding to the 35S leader region, and that was detected by RT-PCR in total RNA fraction of transfected protoplasts, in the same way as the 4 other RNA fragments (**Figure 7D, 7E**). The detection of RNA I-II was however problematic since 4 different bands of almost the same intensity were visible, among which the uppermost corresponds to full-length I-II fragment (**Figure 7E**). This observation, together with the less important signal for L-VII and 3'UTR might be explained by secondary structures making the RNAs less accessible during the cDNA synthesis or, alternately, by increased degradation. Total 35S RNAs accumulation was not affected by the coexpression of any of the RNA fragments (**Figure 7C**) meaning either that transcription efficiency was not influenced or that 35S RNAs were stabilized and less degraded in these conditions.

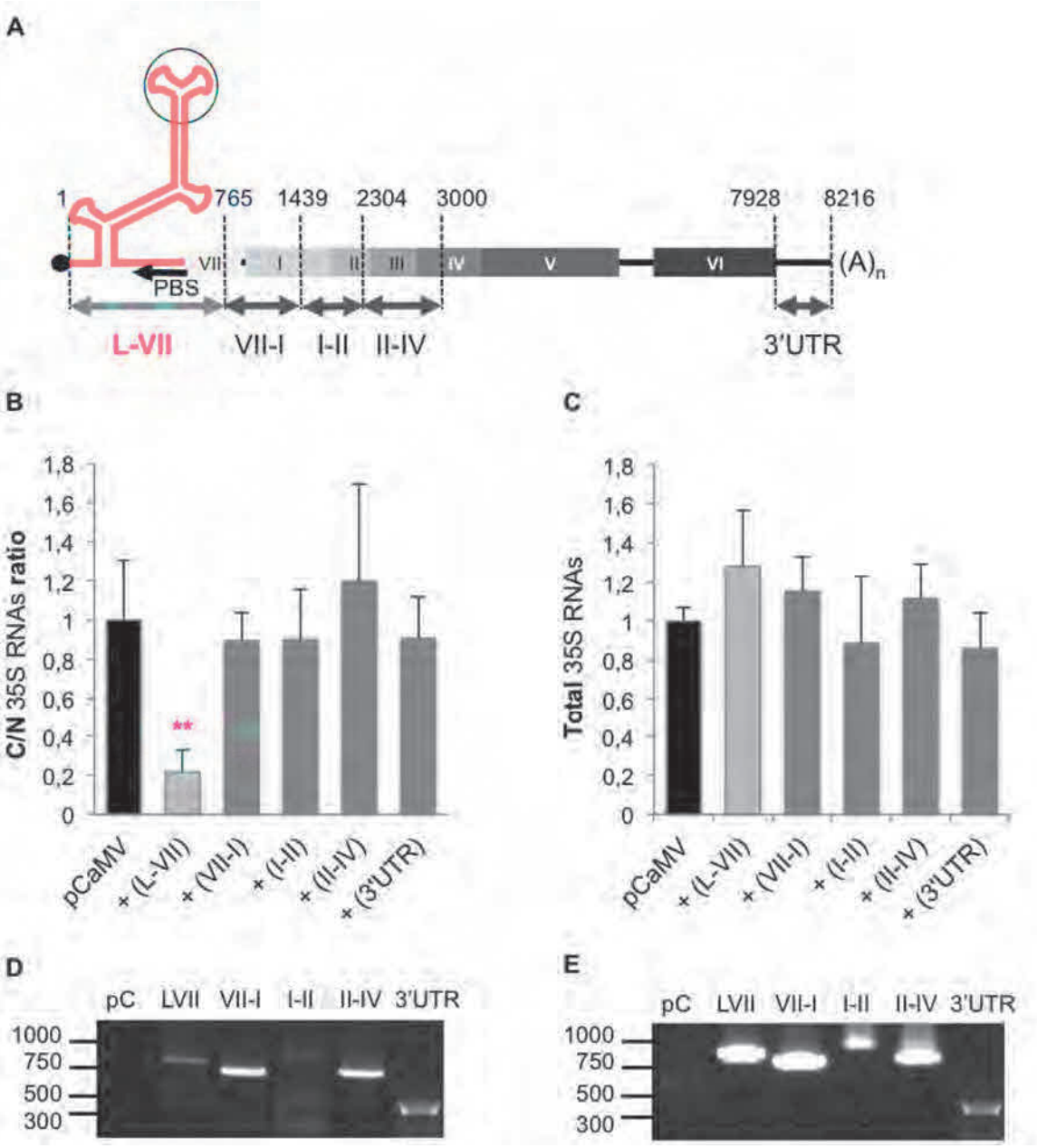


Figure 7. Co-expression of 5'UTR inhibits the nuclear export of 35S RNAs. (A) Structure of the capped and polyadenylated 35S RNA: the 5' UTR (or leader (**L**) in red), 608 nt long, folds into a stable multibranched stem-loop (**SL**), and precedes the 7 ORFs (VII, I, II, III, IV, V and VI) and the 3' UTR (nt 7,928-8,216). The 5 RNA fragments used in the competition assays and overlapping, respectively, the leader and ORF VII (**L-VII**), ORF I and II (**I-II**), ORF II, III and IV (**II-IV**) and the **3' UTR**, are shown with double-headed arrows. The position of the primer ($tRNA^{Meti}$) binding site (PBS, nt 596-609) is indicated by a black arrow. The coat protein P4 binding site is shown with a black circle. (**B**) to (**E**) Leaf mesophyll protoplasts were prepared from 10 days-old WT/*CA-rop2* Arabidopsis seedlings and transfected with 10 µg pCaMV-GFP (black), or co-transfected with pCaMV-GFP and a 10-fold molar excess of Litmus-mRFP-35S-(L-VII) (light grey), Litmus-mRFP-35S-(I-II), Litmus-mRFP-35S-(II-IV) or Litmus-mRFP-35S-(3'UTR) (dark grey). Twenty hpt, the cytoplasmic-to-nuclear ratios (**B**) or the total levels (**C**) of 35S mRNAs were determined by P5-specific real-time RT-PCR and normalised to the cellular housekeeping gene GAPDH, to GFP (transfection indicator for pCaMV-GFP) and to mRFP (transfection indicator for the RNA fragments expressing plasmids). Three independent experiments were performed in triplicate. The results are expressed as the mean fold change ± SEM. Significance was tested with a Student's unpaired 2-

samples t-test (** p<0.01). (**D**) Detection of the 5 competition RNA fragments 20 hpt by RT-PCR in total RNA fraction, with primers specific to their common short 5' and 3' sequences: (L-VII): 765 bp, (VII-I): 674 bp, (I-II): 865 bp, (II-III): 696 bp and (3' UTR): 289 bp. (**E**) PCR positive controls, directly amplified with the same primers from the respective Litmus-mRFP-35S plasmids. Positions of molecular size standards, in bp, are shown on the left.

The L-VII RNA fragment that we demonstated to presumabaly titrate the nuclear export machinery, is composed (**Figure 7A**), at its 5' end, of a 74 nt long unstructured sequence harbouring a short ORF (sORF A, nt 57-68) which is essential for ribosome shunting during translation (72) and which terminates in front of the stable helical section of a large stem-loop secondary structure of 482 nt (87) that contains 5 additional sORFs (for Cabb B-JI isolate), the poly(A) signal and the first splicing donor (SD) site located, respectively, at the ascending and descending arms of a middle section of the stem-loop structure (88), and the 35S pregenomic RNA packaging signal exposed on an uppermost section of the stem-loop structure (nt 288-355) and which interacts with P4 coat protein (89). The 209 nt following the stem-loop and preceding ORF VII are likely unstructured and end with the $tRNA^{Meti}$ primer binding site (nt 596-609) where P5 reverse transcriptase is loaded (90).

In order to further investigate whether some sections of the L-VII fragment are in particular involved in the nuclear export, we splitted the sequence in 8 different sections (**Figure 8C, upper part**): the stem-loop (SL), its left (SL_Left) and right (SL_right) arms, the bottom (SL_bottom) and top (SL_top) sections, the 5' (L_5') and 3' ends (L_3') of the leader, and the single-stranded sequence preceding ORF VII (SS_3'). The respective plasmids were then cotransfected, together with pCaMV-GFP, in protoplasts and the nuclear export of 35S RNAs was analysed as previously described (**Figure 8A**). As we can see, the important nuclear retention of 35S RNAs was confirmed in this assay and, even if to a lesser extent, a significant inhibition effect was also observed with the coexpression of the stem-loop.

These results suggest that 2 elements are likely recognised by the export machinery, the highly structured SL and an additional element located at sequences upstream or downstream of it. Disrupting the SL structure in separated arms (SL_Left with the poly(A) signal, and SL_Right with the SD site) or expressing only parts of it (SL_Bottom, SL_Top with the packaging signal, L_5', L_3' and SS-3' both containing the PBS) completely abolished the competition for the export proteins between these short RNA fragments and the 35S RNAs that were then efficiently exported into the cytoplasm (**Figure 8A**). Total 35S RNAs accumulation was again not influenced by the expression of these additional RNA fragments (**Figure 8B**) that could be easily detected in transfected protoplasts by RT-PCR (**Figure 8C, 8D**).

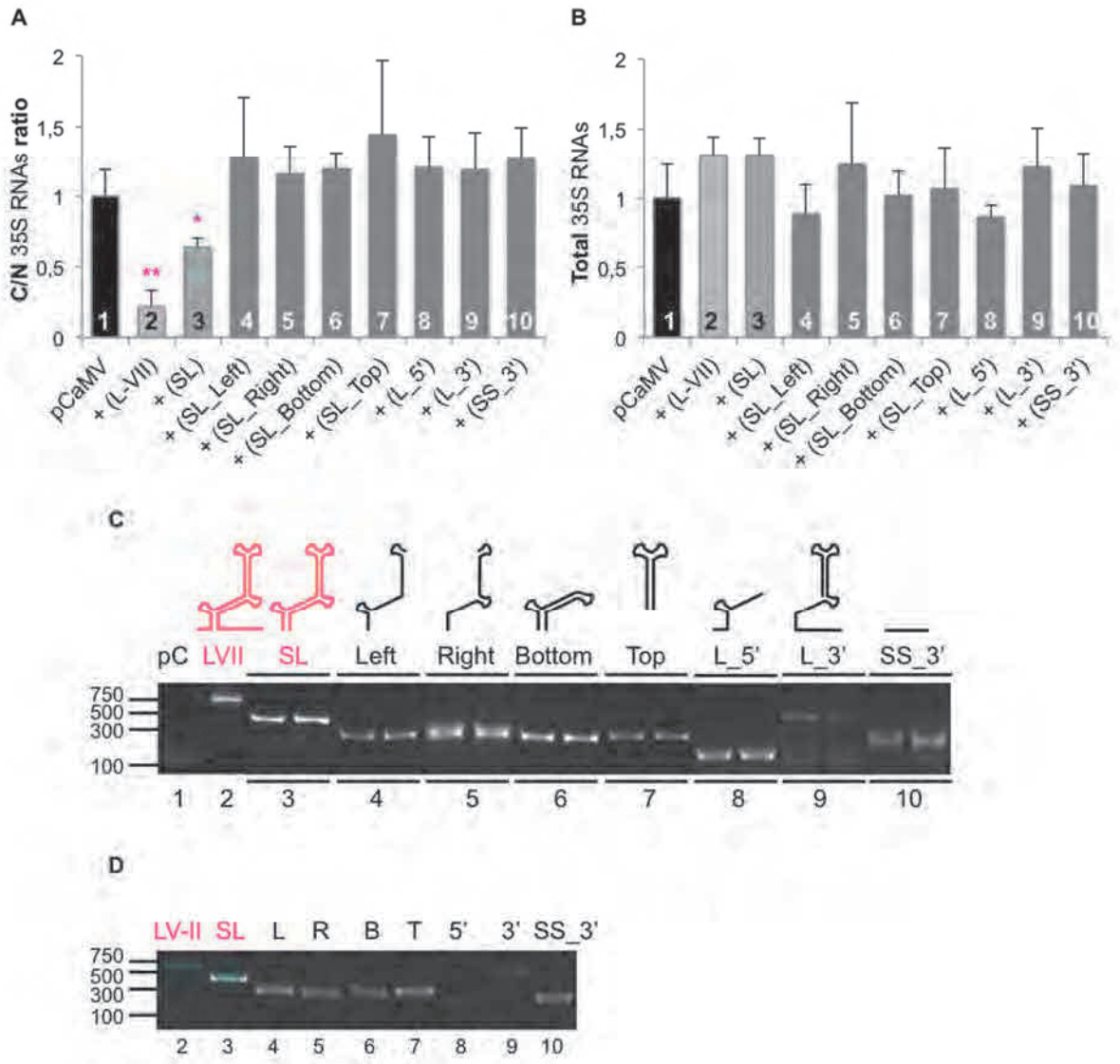


Figure 8. Co-expression of 5'UTR or its stem-loop inhibits the nuclear export of 35S RNA. Leaf mesophyll protoplasts were prepared from 10 days-old WT/*CA-rop2* Arabidopsis seedlings and transfected with 10 µg pCaMV-GFP (**1**, black), or cotransfected, with pCaMV-GFP + Litmus-mRFP-35S encoding (**2**) the RNA fragment L-VII (nt 1-765), (**3**) the stem-loop (SL, nt 75-556), (**4**) the SL left arm (SL_Left, nt 75-324), (**5**) the SL right arm (SL_Right, nt 325-556), (**6**) the SL bottom section (SL_Bottom, nt 75-185 and 445-556), (**7**) the SL top section (SL_Top, nt 186-444), (**8**) the leader 5' end (L_5', nt 1-186), (**9**) the leader 3' end (L_3', nt 187-608), or (**10**) the 3' unstructured sequence following SL (SS_3', nt 557-765). Twenty hpt, the cytoplasmic-to-nuclear ratios (**A**) or the total levels (**B**) of 35S mRNAs were determined by P5-specific real-time RT-PCR and normalised to the cellular housekeeping gene GAPDH, to GFP (transfection indicator for pCaMV-GFP) and to mRFP (transfection indicator for Litmus-mRFP-35S). Three independent experiments were performed in triplicate. The results are expressed as the mean fold change ± SEM. Significance was tested with a Student's unpaired 2-samples t-test (*p<0.05, ** p<0.01). (**C**) Detection of the 9 competition RNA fragments (in duplicate) 20 hpt by RT-PCR with primers specific to their common, vector-derived 5' and 3' sequences: (**2**, L-VII): 765 bp, (**3**, SL): 482 bp, (**4**, SL_Left (L)): 250 bp, (**5**, SL_Right (R)): 232 bp, (**6**, SL_Bottom (B)): 223 bp, (**7**, SL_Top (T)): 259 bp, (**8**, L_5'): 186 bp, (**9**, L_3'): 422 bp and (**10**, SS_3): 209 pb. (**D**) PCR positive controls, directly amplified with the same primers from the respective Litmus-mRFP-35S plasmids. Positions of molecular size standards, in bp, are shown on the left.

Taken together, our results monitor that the 5' UTR of 35S RNA and its stable stem-loop secondary structure in particular, is a *cis*-acting element that is likely recognized by the export proteins. Therefore, we investigated whether this region can effectively load some RNA-binding proteins of the TREX complex, like ALY1 (32) or MOS11 (53).

Export adaptor ALY1 interacts with CaMV 35S RNA *in vitro*

To test the putative RNA – protein interactions, we performed *in vitro* binding EMSA assays with *in vitro* produced and radiolabeled RNA L-VII (765 nt) and I-II (865 nt) (**Figure 9A**) and bacterially expressed and affinity purified ALY1 and eYFP, as control (**Figure 9B**). Both proteins were His_6-tagged at their Nt and expressed in fusion with GB1 polypeptide to increase protein solubility in *E. coli* (32). We showed previously that RNA L-VII (and its stable and long stem-loop) outcompete 35S RNAs for nuclear export factors, while RNA I-II is likely not involved in export (**Figure 8**) and, according to mfold predictions (data not shown), it does not fold into a stable secondary structure.

In presence of 400 pmole control protein His_6-GB1-eYFP the electrophoretic mobility of neither RNA L-VII nor RNA I-II was slowed down (**Figure 9C**) demonstrating that this protein, harbouring the same tags as ALY1, did not interact with any of the 35S RNA fragments. On the contrary, increasing amounts of His_6-ALY1-GB1 (0 to 400 pmole) progressively slowed down RNA L-VII migration (**Figure 9D**): compared to free RNA L-VII (0 pmole ALY1), the band-shift appeared in presence of 100 pmole ALY1 and increased at 200, 300 and 400 pmole protein suggesting that ALY1 interacts with RNA L-VII and that multiple copies of the protein are progressively loaded on the 765 nt long fragment. RNA L-VII being composed of 5' and 3' unstructured sequences (74 and 209 nt, respectively), and the 482 nt stem-loop, it is likely that ALY1 might interact with both structured and unstructured RNA sections thus forming larger and larger ribonucleoproteins, the migration of which is progressively shifted. This observation is in accordance with the results of Pfaff *et al.* (32) who demonstrated that Arabidopsis ALY1 behaves as a typical RNA binding protein which interacts *via* its central RNA recognition motif (RRM) with both single-stranded (ss) and double-stranded (ds) RNA, and even with ssDNA, albeit exhibiting a preference for interaction with ssRNA.

Consistent with these findings, it was not surprising that we observed almost the same gel-shift of RNA I-II in presence of increasing amounts (0 to 400 pmole) of ALY1 (**Figure 9E**), starting, as for L-VII, at 100 pmole protein. This RNA, albeit not involved in 35S RNA nuclear export, interacted with export adaptor ALY1 as did *cis*-acting fragment L-VII. One might imagine that *in vivo*, and in the context of CaMV infection, ALY1 (and the TREX complex) would be

preferentially recruited on 35S 5' UTR by the adjacent cap binding complex (CBC) and/or EJCs but probably also *via* the interactions of MOS11 with capsid protein P4, which binds the uppermost part of the stem-loop, and reverse transcriptase P5, loaded on the primer binding site downstream of the stem-loop.

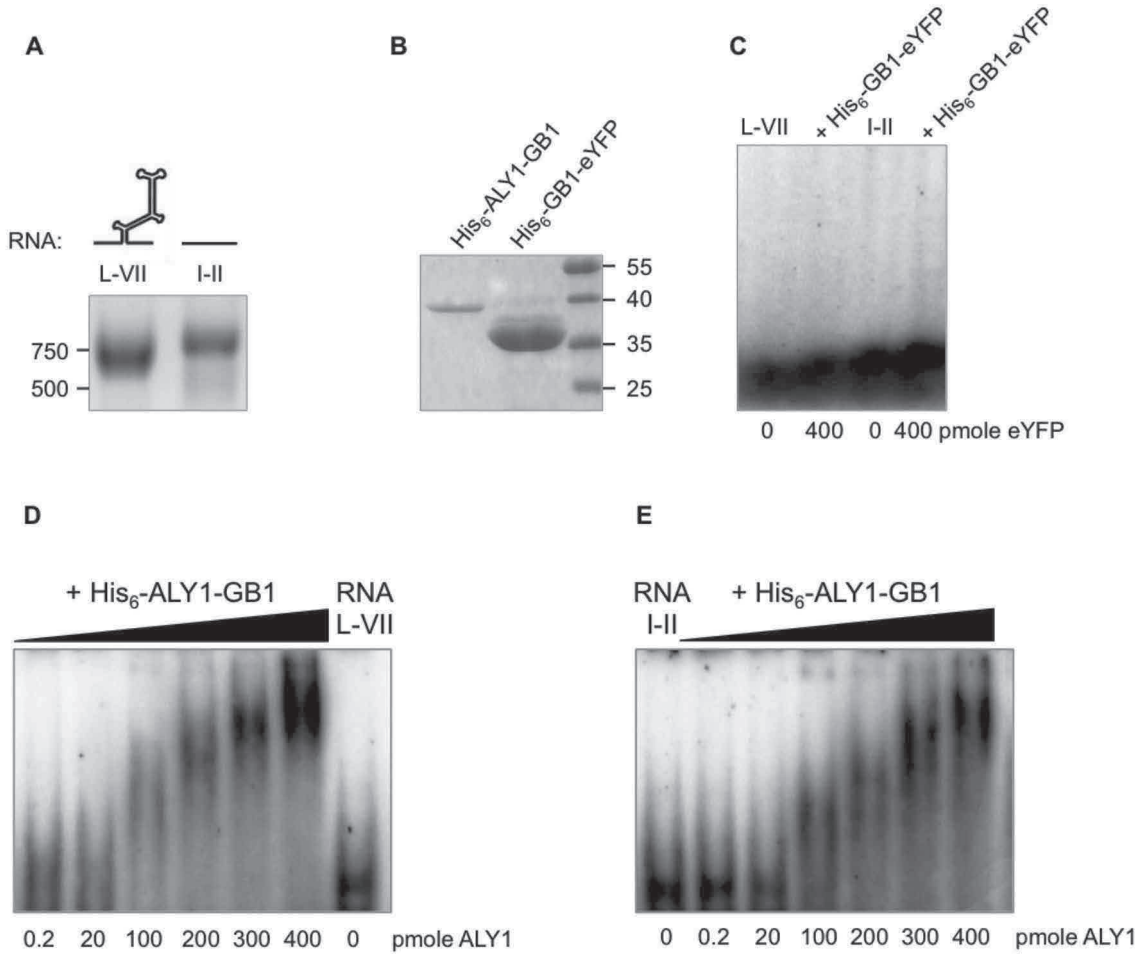


Figure 9. CaMV 35S RNA can load ALY1 *in vitro*. (A) 35S RNA fragments L-VII and I-II were *in vitro* synthesised by T7 RNA polymerase, labelled with [α-^{32}P]-UTP, and analysed by denaturing agarose electrophoresis. (**B**) ALY1 and eYFP, as control, were expressed in *E. coli* as His_6 - GB1 fusion proteins, purified by two-step chromatography, and examined by SDS-PAGE. (**C**) EMSA with ≈ 5,000 cpm L-VII or I-II RNA and 400 pmole purified His_6-GB1-eYFP, analysed by non-denaturing electrophoresis and autoradiography. (**D**) EMSA with ≈ 5,000 cpm RNA L-VII or I-II (**E**) and increasing amounts (0 to 400 pmole) of purified His_6-ALY1-GB1, analysed by non-denaturing electrophoresis and autoradiography. Positions of molecular size standards, in nt (A) or kDa (B) are shown on the left and on the right, respectively.

DISCUSSION

Host cellular machineries play indispensable roles in virus life cycles. Among them is the nuclear export machinery that allows mRNAs to reach the cytoplasm to be translated. In Arabidopsis, the main export receptor TAP that are conserved from yeast, through Drosophila and Xenopus, to humans (91) do not exist despite the presence of the as well conserved TREX adaptor complex, the main role of which is to recognise and interact with cargo mRNAs before recruiting the exportins.

In this study, we investigated and described for the first time the mechanisms used by a plant virus, CaMV, to export its spliced and unspliced 35S mRNAs out of the nucleus. We demonstrated that this export is mediated by the TREX complex and involves the THO subunit and splicing factor TEX1 (53), and adaptor proteins MOS11 (35, 53) and ALYs (32). In mutant protoplasts, deficient for the expression of one or more of the export factors, 35S RNAs were retained in the nuclei and this export inhibition was responsible for a partial resistance to CaMV infection of the respective mutant Arabidopsis.

It has been already described that among the 6 THO components (HPR1, THO2, TEX1, THO5, 6, 7) TEX1 and HPR1 are required for efficient export of cellular mRNAs in Arabidopsis (53, 92, 93). In our hands however, CaMV 35S RNA nuclear export was not affected in single-mutant *tex1*, *hpr1* or *tho6* protoplasts and the respective mutant plants displayed the same susceptibility to CaMV (data not shown). These results suggest that the composition and build-up of the THO core is likely less important for the export of viral mRNAs as it is for the cellular, and that there are certainly differences in the TREX-dependent export of these different mRNA populations.

Export adaptor MOS11 revealed as a truly important one for 35S RNAs export since in the absence of this protein the nuclear export was reduced by almost 50%. This observation is in line with the important nuclear retention of cellular mRNAs described for *mos11* plants (53), and, as for cellular mRNAs, the export inhibition of 35S was significantly increased in double-mutant *mos11 tex1*. One might imagine that TEX1 enhances the effect of MOS11 and is important for CaMV because of its role in splicing: TEX1 interacts with several splicing factors, colocalises in a speckled pattern in nucleoplasm with SR proteins, and its absence in *tex1* and *mos11 tex1* mutants is responsible for quantitative changes in alternatively spliced transcript variants (53). According to these findings, it was not excluded that in the absence of TEX1, CaMV 35S splicing was modified leading to the accumulation of certain isoforms that are more retained in nuclei of *mos11 tex1* protoplasts. However, when we examined, by RT-PCR, the splicing profile of 35S RNA in tex1 single- and double-mutants, we did not observe any

significant differences since all 4 spliced and the unspliced isoform were present at levels, comparable to those in WT protoplasts (data not shown) suggesting that the potentiating effect of TEX1 on export is not linked to splicing and requires further clarification.

TEX1, and probably indirectly MOS11, were also shown to be involved in biogenesis of tasiRNA (trans-acting RNAs) TAS1 and TAS3 since reduced levels of both tasiRNAs, as well as of miRNA miR173, which initiates TAS1 and TAS2 biogenesis, were measured in *tex1* and *tex1 mos11* mutants (53). We can wonder whether this tasiRNA pathway deregulation is not indirectly and partially responsible for the effects observed on CaMV mRNAs export in those mutant plants. However, according to recent data, tasiRNAs generated from TAS1 and TAS2 loci mainly regulate the expression of pentatricopeptide mRNAs, while those from TAS3 and TAS4 control mRNAs encoding auxin-response and MYB transcription factors, respectively (94–98) but not nuclear export factors. Moreover, it is known that CaMV, and especially its P6 protein, are responsible for the same effect as the one observed in *tex1* and *mos11 tex1* plants: altered biogenesis of the 4 known tasiRNA families (TAS1-4), with overaccumulation of long unprocessed precursors (99).

The proteins that appeared to be the most important for CaMV 35S RNA nuclear export were definitely ALY1-4, since the export was strongly inhibited in double-mutants *aly1 aly2* and *aly3 aly4*, and especially in the quadruple-mutant *4xaly*, by more than 90%. In line with this observation, 75% of *4xaly* plants displayed resistance to CaMV infection.

ALYs are all ubiquitously expressed in vegetative tissues but display, by pairs, a different sub-nuclear localisation, ALY3 and ALY4 being much more strongly enriched in nucleoli than ALY1 and ALY2 (32). Considering the important viral mRNA export inhibition in the 2 ALY double-mutants, it is surprising that these plants did not exhibit a more pronounced resistance to CaMV. Our results suggest that *in planta* there is certainly a functional compensation of the lacking ALYs by the 2 other still expressed paralogues. We can also wonder that, despite the huge and indisputable importance of ALYs for viral and cellular mRNA export, in general, *4xaly* plants are severely affected but viable and not totally resistant to CaMV infection while yeast cells lacking Yra1, the otrthologue of ALY, are not viable (100, 101). These findings can however be explained with the recently demonstrated existence, in Arabidopsis, of 2 additional ALY functional homologues, namely the UAP56 interacting export factors UIEF1 and 2, that can also bind RNA (33). Double-mutant *2xuief* displays only modest growth effects and is mildly affected in nuclear mRNA accumulation but sextuple *4xaly 2xuief*, without markedly exacerbating *4xaly* growth and reproductive defects, significantly enhances the nuclear mRNA accumulation of *4xaly* plants (33). It would be therefore of great interest to investigate 35S RNA nuclear export in this newly generated mutant and its susceptibility to CaMV infection, for both

of which we can expect an almost total inhibition regarding the results we obtained in *4xaly* plants.

It was also recently reported that, as in mammals (21) some Arabidopsis export factors could selectively participate in the export of particular subsets of mRNAs. Thus, ALY1 has been described as being involved in the export of mRNAs encoding proteins from the RNA-directed DNA methylation pathway (RdDM), including AGO6, since *aly1* mutants display a strong nuclear accumulation of these mRNAs and deficiencies in RdDM (102). It is also known that CaMV genomic DNA that persists in the nuclei of infected plants is methylated (103) and that this methylation may lead to transcriptional gene silencing (TGS) mediated by the high amounts of DCL3-produced 24 nt siRNAs (104, 105). In line with these findings, the inhibition of RdDM machinery in *aly1* mutants should not have a negative effect on CaMV replication and/or mRNA export since it should lead to reduction of TGS in the infected plants. However, we did not detect any increase in total 35S RNA accumulation neither in double-mutant *aly1 aly2* nor in quadruple *4xaly* suggesting that either the potential RdDM inhibition by *aly1* did not influence 35S RNA levels, or that this effect is somehow attenuated in double- and quadruple-mutants.

In this study we did not analyse the importance of the last major TREX components, the 2 UAP56 helicases, the importance of which in nuclear export of mRNAs in Arabidopsis was clearly demonstrated: both UAP56a and UAP56b bind RNA, and interact with and recruit into the TREX complex MOS11, ALYs and UIEFs (33, 34, 53). To our knowledge, double-mutant *uap56a uap56b* were not described, while in both single-mutants the intact gene compensates for the lacking protein of the mutated gene and the mutants display WT phenotypes (34). Due to the close proximity of the two Arabidopsis *UAP56* genes on chromosome 5, it is unlikely to obtain double mutants that do not express UAP56 by crossing of the single mutants, it would be therefore important to generate double-mutants by CRISPR/Cas9 technology (106). According to the multiple but still viable mutants obtained for ALYs and UIEFs, we might expect viable *uap56a uap56b*, even if depletion of UAP56 in mammals, Drosophila, Caenorhabditis and yeast was shown to severely affect mRNA export and cell growth/survival (107–111).

Considering the important diversification of several components of the Arabidopsis mRNA export pathway, another interesting issue that should be raised is to determine whether there is an “optimal“ TREX complex for the nuclear export CaMV 35S RNAs, constituted of particular THO components (rather THO5a or 5b, THO7a or 7b) or export factors and adaptors (UAP56a or b; ALY1, 2, 3 or 4, UIEF1 or 2) with specific functions in the export of different subpopulations of mRNAs.

The involvement of the TREX complex in the nuclear export of viral mRNAs is not surprising since TREX is part of the major export pathway of transporters TAP/p15 that are hijacked by many viruses in mammals, like simple and complex retroviruses (112–116) and DNA viruses like adenoviruses (117, 118), herpesviruses (40, 119, 120) and HBV (121–123). Among all these viruses many of them use a specific and, in general, multifunctional viral protein that interacts with both viral mRNAs one or several export factors and thus recruits the export machinery. Herpesviruses ICP27/ORF57 proteins, for example, can associate with TAP transporter (124), ALY (39), UIF (41) and CIP29 (42).

Therefore, it was not surprising when we discovered that in the case of CaMV, at least 1 viral protein, protease/reverse transcriptase P5, is important for viral mRNA export, as demonstrated in protoplasts transfected with a mutated pCaMV-GFP, which does not allow the expression of, either any viral protein, or of P5. In these protoplasts 35S RNA were massively retained in nuclei and this export deficiency was efficiently rescued *in trans* by full-length uncleaved P5 but not by its individual functional PR or RTRH domains.

To our knowledge, involvement of reverse transcriptases in the nuclear export of viral mRNAs has never been described before, neither for retroviruses nor for pararetrovirus HBV. Involvement, however, of viral RNA-dependent RNA-polymerases was already demonstrated for influenza A virus (IAV), the heterotrimeric polymerase of which recruits cellular helicase DDX19 (also known as DBP5) on viral mRNPs to enhance their nuclear export (125). It was also shown very recently that the same polymerase, and especially its PB1 and PB2 subunits, recruit cellular RNA helicase eIF4A3, which is part of the EJC, to promote viral mRNA splicing and nuclear export of the spliced mRNAs (126).

The role of P5 in nuclear export can be easily understood when considering several previous findings. As genome encapsidation probably occurs during reverse transcription of 35S unspliced (pregenomic) RNA (127, 128), P5 is present inside the virions (129), probably at, at least 3 copies, needed to perform the negative-strand synthesis and the discontinuous synthesis of the positive-strand DNA (90, 130). After capsid dissociation on nuclear pores, some capsid and P5 proteins are very likely imported into the nucleus, in association with the viral genome. Pfeiffer *et al.* (82) also described an increase of a DNA polymerase of about 75 kDa, present in infected turnip nuclear extracts. During viral life cycle P5, together with P6, which is independently expressed from its own 19S RNA, and P1, synthesised from unspliced 35S RNAs, accumulate very early in turnip protoplasts (131) and this observation is quite surprising when considering the fact that the main functions of P5 (genome replication and capsid maturation) take place late in the viral cycle, and suggests other potentially early roles for P5. For this reason, we investigated whether P5, being important for viral mRNA export, is able to

interact with some export factors and demonstrated, by GST pull-down assays, that *in vitro* P5, and especially its central RT domain, associates with MOS11. Moreover, this interaction is very likely responsible for some conformational changes in P5 that might activate its Nt aspartic proteinase domain, as in the GST-bound fraction we found not only full-length P5 but also a protein, co-migrating with the RTRH domain, which is naturally produced after P5 self-cleavage operated, probably *in cis*, by the PR domain (49). Similar enzymatic activation upon interaction with MOS11 was already described in mammals where CIP29, the orthologue of MOS11, stimulates the helicase activity of UAP56 (132); as is doing ALY that stimulates the ATPase activity of this same helicase (133). This activation of P5 protease domain by MOS11 needs definitely to be further studied and characterised but it is possible that it might take place during the nuclear export of 35S RNA, in association with the TREX complex, the presence of P5 facilitating not only the RNA export but also guiding the exported cytoplasmic viral mRNAs to the viral factories, made by P6, where reverse transcription by active RT and encapsidation can start. Finally, we do not believe that the involvement of P5 in viral mRNA nuclear export might be dependent on its reverse transcriptase activity, as in unprocessed P5, needed for export, RT seems not to be active (78). Moreover, our analysis were always performed ≈ 20 hpt, according to Khelifa *et al.* (55) who demonstrated that in transfected protoplasts newly synthesised and encapsidated viral cytoplasmic DNA (i. e. new generation of virions that might re-infect the nuclei) are not detected before 21 hpt. Our hypothesis can be further verified by mutating the RT conserved catalytic site (Y_{401}-V-D-D_{404}) and then measuring viral mRNA nuclear export in protoplasts.

In the same GST pull-down experiments we found that a second viral protein, P4, likely unprocessed and as a dimer, also interacts *in vitro* with MOS11. This interaction suggests that P4 is also involved in the TREX-mediated viral mRNA export, and the absence of not only P5 but also of P4 in protoplasts, transfected with pCaMV-GFP-P6_STOP, might explain the strong nuclear retention of 35S RNA. In this case, we can of course not completely exclude the possibility that one or more of the 4 other viral proteins might also be important for export but this hypothesis seems less likely. In mammals, the involvement of HBV Core protein in the nuclear export of some viral mRNAs by direct loading, not of CIP29 but of transporter TAP/p15, has been described (134).

Proteolytically matured (by P5 proteinase) coat proteins, lacking their N- and C-termini, can enter nuclei *via* their Nt NLS that becomes accessible only upon cleavage (45). However, using eGFP fusion proteins, we visualised, for the first time to our knowledge, that not only full-length P4 protein but also uncleaved P5 display a clear nucleo-cytoplasmic distribution, with typical cytoplasmic aggregates coherent with the cytoplasmic functions of both proteins, and also perfectly colocalised with MOS11-mRFP in the cell nuclei. This nuclear colocalisation

suggests that the *in vitro* observed physical interaction between each of the viral proteins and the export factor might occur in the nuclei in order to promote the nuclear export of viral mRNAs. We are currently working to confirm these interactions in protoplasts, by co-immunoprecipitation (co-IP) and BiFC (bimolecular fluorescence complementation) experiments. Another point that is worth to be raised is to confirm the importance of the P4 already described NLS as well as of the new one we identified in P5 RT domain, by mutating each of them and then investigating the nuclear export of viral 35S RNA in protoplasts. We are also aiming to monitor whether MOS11 is able to interact in the same time with P4 and P5 to form a trimeric complex, like it is doing with UAP56 and ALY (19), and even a tetrameric one, in which UAP56 might also bind MOS11.

Unlike P5, which accumulates early in infected protoplasts, P4, together with P2, is considered as a late accumulating protein (55, 131). This finding suggests that, on the contrary of P5 that might continuously participate in viral mRNA export, P4 would rather be involved in late export events.

We finally looked for the presence of *cis*-acting elements in the 35S RNAs that might be important for the loading of the export machinery, and clearly demonstrated, by complementation assays in protoplasts, that an RNA fragment corresponding to the 5' leader region and the first 1/3 of ORFVII (L-VII) efficiently outcompetes full-length viral RNAs for the export factors. The overexpression of the highly structured stem-loop of the leader (87) that harbours several important elements like 5 regulatory sORFs, the conditional poly(A) site, the coat protein P4 binding site and the first splice donor site, SD4, and that is bypassed during translation by ribosome shunting (72), also inhibited viral RNA nuclear export but to a lesser extent. This observation suggests that the short 5' and/or the longer 3' regions that flank the stem-loop are also important for export. In line with our previous results on the role of P5, we can expect that this additional *cis*-acting element is the $tRNA^{Meti}$ primer-binding site that precedes ORFVII and where P5 is loaded on 35S RNA (90). The sequences downstream of ORFVII that we also tested in the competition assays because they are concerned by 35S alternative splicing (3 additional major SD sites are present in ORFI and the unique major acceptor site is in ORFII (47, 48) did not influence 35S RNA nuclear export. The additional experiments we performed with different sections of the stem-loop or of its flanking 5' and 3' regions confirmed the importance of the secondary structure for nuclear export. These findings however raise the question about the export of the shortest 35S spliced isoform, generated by the use of SD4 in the ascending arm of the stem-loop, and which lacks the secondary structure and the downstream sequences that precede the acceptor site in ORFII, including the PBS. Among all spliced and the unspliced 35S isoform, this one is the less represented (only 10%) (47) and we can imagine that it would also be the less exported, by a "minimal" recruitment of

TREX complexes, as on cellular mRNAs, *via* the CBC and EJCs but without the participation of viral proteins and *cis*-acting elements.

In the competition experiments we set up, there was a potential risk for the short, capped and polyadenylated 35S RNA fragments, and especially for the 5' leader region (L-VII) and its stem-loop (SL), to be recognised by the plant silencing machinery and diced by DCLs to produce viral small RNAs (vsRNAs) that might further target spliced and unspliced "full-length" 35S RNAs. However, in all experiments the competing RNA fragments were always detected by RT-PCR, some of them, like L-VII and I-II, less efficiently, as already discussed, and the accumulation of total 35S RNAs detected by real-time RT-PCR targeting ORFV, was not significantly changed in presence of those RNA fragments. These observations can be explained with the existence, besides RNA silencing suppressor P6 (99, 135, 136) of a second, unusual RNA-based decoy mechanism (104, 105). According to these authors, when Pol II-driven viral transcription starts in nuclei at 35S promoter on still unrepaired and gapped CaMV genomic DNA, there is an abrupt termination of transcription on the Met-tRNA gap, which results in production of an aberrant 8S RNA lacking poly(A) tail (137). This 8S RNA matches the 35S RNA leader sequence, and is predicted to form a viroid-like secondary structure, which is converted to dsRNA by, probably, Pol II. 8S RNA-derived dsRNAs serve as a decoy to engage all four DCLs in massive production of 21, 22 and 24 nt vsRNAs. These leader-region vsRNAs compete with vsRNAs derived from other regions for AGO proteins. Resulting antisense vsRNAs potentially guide AGO-meditated silencing of 35S RNAs, but the structured nature of the leader likely hinders vsRNA pairing and 35S RNA degradation (104, 105). In line with these findings 8S-derived dsRNAs probably protect from degradation not only 35S RNA but also the competing short RNA fragments. On the other hand, there in no risk that 8S RNAs, which match the 5' leader region, enter in competition with 35S RNAs for the nuclear export machinery, as they are mostly converted into dsRNAs.

Finally, we also showed that RNA-binding adaptor ALY1 (32) can interact, likely without preferences, with 35S RNA 5' leader (L-VII) and with a downstream coding sequence (I-II) that is not involved in export and is devoid of stable secondary structures. These RNA-protein binding assays can be further analysed by competition experiments with both (L-VII) and (I-II) **RNA** fragments, but we are mostly focusing on characterising MOS11 capacity to interact with 35S RNA leader, as this 2nd RNA-binding protein was described to associate with more affinity with dsRNA than with ssRNA (53). We believe that the TREX complex can be preferentially enriched on 35S 5' secondary structure by the interactions existing between MOS11, P4, loaded on its binding site on the top of the stem-loop, and/or P5, recruited on the downstream located PBS, in presence, or not, of tRNA^{Meti} primer. In addition, it has been reported that, like human ALY (138), ALY1 exhibits ~3-fold higher affinity for 5-methycytosine-(m^{5}C)-modified RNA

when compared to unmodified control RNA (32). As methylation of RNA bases, such as (m^5C) and N^6-methyladenosine (m^6A), is the most prevalent mRNA modification found in eukaryotes (and in Arabidopsis, in particular (139, 140)), and as CaMV mRNAs are, like their cellular counterpart, Pol II- derived, it is very likely that they also harbour epitranscriptomic marks that might facilitate ALYs loading.

In mammals, the secondary structure of cellular mRNAs does not appear to be often involved in the direct recruitment of nuclear export factors. A well-described example is the mRNA of transporter TAP itself, which contains a structural *cis*-acting element in an alternatively retained intron that recruits by direct interaction the TAP/p15 dimer (141). Inversely, recognition sites for the export machinery have been described for several viral RNAs. These are more or less complex secondary structures, such as the HBV post-transcriptional element PRE (123), the KSHV ORF57-responsive element ORE (142), the CTE of MPMV (143) or RSV DR (direct repeat) elements (114). These *cis*-acting structures represent a composite set of encapsidation and dimerization signals and are therefore often multifunctional. They are recognized by viral proteins or host adaptor proteins and are used for exportins recruitment.

According to the results presented in this study, we can propose the following model for CaMV 35S mRNAs nuclear export. When viral DNA is released from the virion into the nucleoplasm, the gaps, left by RT during the previous replication cycle, are repaired by host DNA repair enzymes to create covalently closed molecules. This closed viral DNA is then transcribed by Pol II into pregenomic polycistronic 35S RNA, 70% of which is alternatively spliced, and into subgenomic 19S RNA, which is the mRNA for the viral protein P6. Unspliced 35S and 3 out of the 4 spliced isoforms contain the highly structured 5' leader whereas the last isoform and 19S RNA do not. During the first round of viral transcription only few copies of coat protein P4 and reverse transcriptase P5 are present in the nucleus. For this reason, a first, minimal round of nuclear export takes place, during which TREX complexes are loaded on viral mRNAs mainly by CBCs and EJCs, probably more efficiently on the leader-containing mRNA species than for the other that can be considered as regular cellular mRNAs. In the cytoplasm, these mRNAs are then translated and viral proteins progressively accumulate. P5 polyprotein first, and P4 coat precursor later, enter the nuclei *via* their NLSs, where they bind their respective recognition sites on the leader-containing 35S RNAs and, interacting with MOS11, promote the recruitment of TREX and enhance the nuclear export of these viral RNAs.

Besides the export model, we propose, an important question however still remains open and needs to be further examined: is the TREX complex export pathway the only one used by CaMV in Arabidopsis or are there other export possibilities, as this is the case for HIV-1 that exports its multi-spliced mRNA *via* TREX and transporter TAP/p15 but also hijacks, later in

its life cycle, another transporter, CRM1, to export its unspliced and partially spliced mRNAs? Indeed, two orthologues of CRM1, exportins 1a and 1b (XPO1a and 1b), have been identified in Arabidopsis where they mediate, as in yeast and metazoa, the nuclear export of proteins, ribosomal subunits and small nuclear RNAs (144, 145) and probably also of mRNAs.

AVAILABILITY

All datasets generated for this study are included in the manuscript and/or the **Supplementary Files**.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

ACKNOWLEDGEMENT

We gratefully acknowledge Martin Drucker (INRA, Colmar, France) for kindly providing pCaMV-GFP, Manfred Heinlein and Nicolas Pitzalis (IBMP, Strasbourg, France) for TuMV, Nicolas Baumberger (IBMP, Strasbourg, France) for precious help with protein purification, Corinne Keichinger (INRA, Colmar, France), Khalid Amari (JKI, Braunschweig, Germany) and Esther Lechner (IBMP, Strasbourg, France), for providing plasmids, Stéphane Blanc (INRA, Montpellier, France), Anne-Marie Duchêne and David Gilmer (IBMP, Strasbourg, France) for providing antibodies, and Marie Meister (Institut de Zoologie, Strasbourg, France) for critical reading of the manuscript.

FUNDING

This work was supported by Centre National de la Recherche Scientifique and Université de Strasbourg. Julie Kubina and Clément Bouton PhD were supported by fellowships from Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation and from Université de Strasbourg (JK).

CONFLICT OF INTEREST

The authors have declared that no conflict of interest exists.

REFERENCES

1. Ramanathan,A., Robb,G.B. and Chan,S.-H. (2016) mRNA capping: biological functions and applications. *Nucleic Acids Res.*, **44**, 7511–7526.

2. Girard,C., Will,C.L., Peng,J., Makarov,E.M., Kastner,B., Lemm,I., Urlaub,H., Hartmuth,K. and Lührmann,R. (2012) Post-transcriptional spliceosomes are retained in nuclear speckles until splicing completion. *Nat. Commun.*, **3**, 994.

3. Stewart,M. (2019) Polyadenylation and nuclear export of mRNAs. *J. Biol. Chem.*, **294**, 2977–2987.

4. Delaleau,M. and Borden,K.L.B. (2015) Multiple Export Mechanisms for mRNAs. *Cells*, **4**, 452–473.

5. Heath,C.G., Viphakone,N. and Wilson,S.A. (2016) The role of TREX in gene expression and disease. *Biochem. J.*, **473**, 2911–2935.

6. Williams,T., Ngo,L.H. and Wickramasinghe,V.O. (2018) Nuclear export of RNA: Different sizes, shapes and functions. *Semin. Cell Dev. Biol.*, **75**, 70–77.

7. Kang,Y., Bogerd,H.P., Yang,J. and Cullen,B.R. (1999) Analysis of the RNA binding specificity of the human tap protein, a constitutive transport element-specific nuclear RNA export factor. *Virology*, **262**, 200–209.

8. Teplova,M., Wohlbold,L., Khin,N.W., Izaurralde,E. and Patel,D.J. (2011) Structure-function studies of nucleocytoplasmic transport of retroviral genomic RNA by mRNA export factor TAP. *Nat. Struct. Mol. Biol.*, **18**, 990–998.

9. Björk,P. and Wieslander,L. (2017) Integration of mRNP formation and export. *Cell. Mol. Life Sci.*, **74**, 2875–2897.

10. Masuda,S., Das,R., Cheng,H., Hurt,E., Dorman,N. and Reed,R. (2005) Recruitment of the human TREX complex to mRNA during splicing. *Genes Dev.*, **19**, 1512–1517.

11. Hautbergue,G.M., Hung,M.-L., Golovanov,A.P., Lian,L.-Y. and Wilson,S.A. (2008) Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP. *Proc. Natl. Acad. Sci. U. S. A.*, **105**, 5154–5159.

12. Hautbergue,G.M., Hung,M.-L., Walsh,M.J., Snijders,A.P.L., Chang,C.-T., Jones,R., Ponting,C.P., Dickman,M.J. and Wilson,S.A. (2009) UIF, a New mRNA export adaptor that works together with REF/ALY, requires FACT for recruitment to mRNA. *Curr. Biol. CB*, **19**, 1918–1924.

13. Shen,J., Zhang,L. and Zhao,R. (2007) Biochemical characterization of the ATPase and helicase activity of UAP56, an essential pre-mRNA splicing and mRNA export factor. *J. Biol. Chem.*, **282**, 22544–22550.

14. Fukuda,S., Wu,D.W., Stark,K. and Pelus,L.M. (2002) Cloning and characterization of a proliferation-associated cytokine-inducible protein, CIP29. *Biochem. Biophys. Res. Commun.*, **292**, 593–600.

15. Cheng,H., Dufu,K., Lee,C.-S., Hsu,J.L., Dias,A. and Reed,R. (2006) Human mRNA Export Machinery Recruited to the 5′ End of mRNA. *Cell*, **127**, 1389–1400.

16. Chi,B., Wang,Q., Wu,G., Tan,M., Wang,L., Shi,M., Chang,X. and Cheng,H. (2013) Aly and THO are required for assembly of the human TREX complex and association of TREX components with the spliced mRNA. *Nucleic Acids Res.*, **41**, 1294–1306.

17. Viphakone,N., Sudbery,I., Griffith,L., Heath,C.G., Sims,D. and Wilson,S.A. (2019) Co-transcriptional Loading of RNA Export Factors Shapes the Human Transcriptome. *Mol. Cell*, **75**, 310-323.e8.

18. Le Hir,H., Gatfield,D., Izaurralde,E. and Moore,M.J. (2001) The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J.*, **20**, 4987–4997.

19. Dufu,K., Livingstone,M.J., Seebacher,J., Gygi,S.P., Wilson,S.A. and Reed,R. (2010) ATP is required for interactions between UAP56 and two conserved mRNA export proteins, Aly and CIP29, to assemble the TREX complex. *Genes Dev.*, **24**, 2043–2053.

20. Viphakone,N., Hautbergue,G.M., Walsh,M., Chang,C.-T., Holland,A., Folco,E.G., Reed,R. and Wilson,S.A. (2012) TREX exposes the RNA-binding domain of Nxf1 to enable mRNA export. *Nat. Commun.*, **3**, 1006.

21. Wickramasinghe,V.O. and Laskey,R.A. (2015) Control of mammalian gene expression by selective mRNA export. *Nat. Rev. Mol. Cell Biol.*, **16**, 431–442.

22. Kiesler,E., Miralles,F. and Visa,N. (2002) HEL/UAP56 binds cotranscriptionally to the Balbiani ring pre-mRNA in an intron-independent manner and accompanies the BR mRNP to the nuclear pore. *Curr. Biol. CB*, **12**, 859–862.

23. Nojima,T., Hirose,T., Kimura,H. and Hagiwara,M. (2007) The Interaction between Cap-binding Complex and RNA Export Factor Is Required for Intronless mRNA Export. *J. Biol. Chem.*, **282**, 15645–15651.

24. Huang,Y. and Steitz,J.A. (2001) Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. *Mol. Cell*, **7**, 899–905.

25. Hargous,Y., Hautbergue,G.M., Tintaru,A.M., Skrisovska,L., Golovanov,A.P., Stevenin,J., Lian,L.-Y., Wilson,S.A. and Allain,F.H.-T. (2006) Molecular basis of RNA recognition and TAP binding by the SR proteins SRp20 and 9G8. *EMBO J.*, **25**, 5126–5137.

26. Tintaru,A.M., Hautbergue,G.M., Hounslow,A.M., Hung,M.-L., Lian,L.-Y., Craven,C.J. and Wilson,S.A. (2007) Structural and functional analysis of RNA and TAP binding to SF2/ASF. *EMBO Rep.*, **8**, 756–762.

27. Lei,H., Dias,A.P. and Reed,R. (2011) Export and stability of naturally intronless mRNAs require specific coding region sequences and the TREX mRNA export complex. *Proc. Natl. Acad. Sci. U. S. A.*, **108**, 17985–17990.

28. Lei,H., Zhai,B., Yin,S., Gygi,S. and Reed,R. (2013) Evidence that a consensus element found in naturally intronless mRNAs promotes mRNA export. *Nucleic Acids Res.*, **41**, 2517–2525.

29. Ehrnsberger,H.F., Grasser,M. and Grasser,K.D. (2019) Nucleocytosolic mRNA transport in plants: export factors and their influence on growth and development. *J. Exp. Bot.*, 10.1093/jxb/erz173.

30. Yelina,N.E., Smith,L.M., Jones,A.M.E., Patel,K., Kelly,K.A. and Baulcombe,D.C. (2010) Putative Arabidopsis THO/TREX mRNA export complex is involved in transgene and endogenous siRNA biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 13948–13953.

31. Uhrig,J.F., Canto,T., Marshall,D. and MacFarlane,S.A. (2004) Relocalization of nuclear ALY proteins to the cytoplasm by the tomato bushy stunt virus P19 pathogenicity protein. *Plant Physiol.*, **135**, 2411–2423.

32. Pfaff,C., Ehrnsberger,H.F., Flores-Tornero,M., Sørensen,B.B., Schubert,T., Längst,G., Griesenbeck,J., Sprunck,S., Grasser,M. and Grasser,K.D. (2018) ALY RNA-Binding Proteins Are Required for Nucleocytosolic mRNA Transport and Modulate Plant Growth and Development1[OPEN]. *Plant Physiol.*, **177**, 226–240.

33. Ehrnsberger,H.F., Pfaff,C., Hachani,I., Flores-Tornero,M., Sørensen,B.B., Längst,G., Sprunck,S., Grasser,M. and Grasser,K.D. (2019) The UAP56-Interacting Export Factors UIEF1 and UIEF2 Function in mRNA Export. *Plant Physiol.*, **179**, 1525–1536.

34. Kammel,C., Thomaier,M., Sørensen,B.B., Schubert,T., Längst,G., Grasser,M. and Grasser,K.D. (2013) Arabidopsis DEAD-box RNA helicase UAP56 interacts with both RNA and DNA as well as with mRNA export factors. *PloS One*, **8**, e60644.

35. Germain,H., Qu,N., Cheng,Y.T., Lee,E., Huang,Y., Dong,O.X., Gannon,P., Huang,S., Ding,P., Li,Y., *et al.* (2010) MOS11: A New Component in the mRNA Export Pathway. *PLoS Genet.*, **6**.

36. Koffa,M.D., Clements,J.B., Izaurralde,E., Wadd,S., Wilson,S.A., Mattaj,I.W. and Kuersten,S. (2001) Herpes simplex virus ICP27 protein provides viral mRNAs with access to the cellular mRNA export pathway. *EMBO J.*, **20**, 5769–5778.

37. Malik,P., Blackbourn,D.J. and Clements,J.B. (2004) The evolutionarily conserved Kaposi's sarcoma-associated herpesvirus ORF57 protein interacts with REF protein and acts as an RNA export factor. *J. Biol. Chem.*, **279**, 33001–33011.

38. Boyne,J.R., Colgan,K.J. and Whitehouse,A. (2008) Herpesvirus saimiri ORF57: a post-transcriptional regulatory protein. *Front. Biosci. J. Virtual Libr.*, **13**, 2928–2938.

39. Tunnicliffe,R.B., Hautbergue,G.M., Kalra,P., Jackson,B.R., Whitehouse,A., Wilson,S.A. and Golovanov,A.P. (2011) Structural basis for the recognition of cellular mRNA export factor REF by herpes viral proteins HSV-1 ICP27 and HVS ORF57. *PLoS Pathog.*, **7**, e1001244.

40. Tunnicliffe,R.B., Tian,X., Storer,J., Sandri-Goldin,R.M. and Golovanov,A.P. (2018) Overlapping motifs on the herpes viral proteins ICP27 and ORF57 mediate interactions with the mRNA export adaptors ALYREF and UIF. *Sci. Rep.*, **8**, 15005.

41. Jackson,B.R., Boyne,J.R., Noerenberg,M., Taylor,A., Hautbergue,G.M., Walsh,M.J., Wheat,R., Blackbourn,D.J., Wilson,S.A. and Whitehouse,A. (2011) An interaction between KSHV ORF57 and UIF provides mRNA-adaptor redundancy in herpesvirus intronless mRNA export. *PLoS Pathog.*, **7**, e1002138.

42. Schumann,S., Baquero-Perez,B. and Whitehouse,A. (2016) Interactions between KSHV ORF57 and the novel human TREX proteins, CHTOP and CIP29. *J. Gen. Virol.*, **97**, 1904–1910.

43. Grüter,P., Tabernero,C., von Kobbe,C., Schmitt,C., Saavedra,C., Bachi,A., Wilm,M., Felber,B.K. and Izaurralde,E. (1998) TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. *Mol. Cell*, **1**, 649–659.

44. Chi,B., Wang,K., Du,Y., Gui,B., Chang,X., Wang,L., Fan,J., Chen,S., Wu,X., Li,G., *et al.* (2014) A Sub-Element in PRE enhances nuclear export of intronless mRNAs by recruiting the TREX complex via ZC3H18. *Nucleic Acids Res.*, **42**, 7305–7318.

45. Karsies,A., Merkle,T., Szurek,B., Bonas,U., Hohn,T. and Leclerc,D. (2002) Regulated nuclear targeting of cauliflower mosaic virus. *J. Gen. Virol.*, **83**, 1783–1790.

46. Haas,M., Geldreich,A., Bureau,M., Dupuis,L., Leh,V., Vetter,G., Kobayashi,K., Hohn,T., Ryabova,L., Yot,P., *et al.* (2005) The open reading frame VI product of Cauliflower mosaic virus is a nucleocytoplasmic protein: its N terminus mediates its nuclear export and formation of electron-dense viroplasms. *Plant Cell*, **17**, 927–943.

47. Kiss-László,Z., Blanc,S. and Hohn,T. (1995) Splicing of cauliflower mosaic virus 35S RNA is essential for viral infectivity. *EMBO J.*, **14**, 3552–3562.

48. Bouton,C., Geldreich,A., Ramel,L., Ryabova,L.A., Dimitrova,M. and Keller,M. (2015) Cauliflower mosaic virus Transcriptome Reveals a Complex Alternative Splicing Pattern. *PloS One*, **10**, e0132665.

49. Torruella,M., Gordon,K. and Hohn,T. (1989) Cauliflower mosaic virus produces an aspartic proteinase to cleave its polyproteins. *EMBO J.*, **8**, 2819–2825.

50. Ryabova,L., Park,H.-S. and Hohn,T. (2004) Control of translation reinitiation on the cauliflower mosaic virus (CaMV) polycistronic RNA. *Biochem. Soc. Trans.*, **32**, 592–596.

51. Murashige,T. and Skoog,F. (1962) A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiol. Plant.*, **15**, 473–497.

52. Li,H., Shen,J.J., Zheng,Z.L., Lin,Y. and Yang,Z. (2001) The Rop GTPase switch controls multiple developmental processes in Arabidopsis. *Plant Physiol.*, **126**, 670–684.

53. Sørensen,B.B., Ehrnsberger,H.F., Esposito,S., Pfab,A., Bruckmann,A., Hauptmann,J., Meister,G., Merkl,R., Schubert,T., Längst,G., *et al.* (2017) The Arabidopsis THO/TREX component TEX1 functionally interacts with MOS11 and modulates mRNA export and alternative splicing events. *Plant Mol. Biol.*, **93**, 283–298.

54. Lellis,A.D., Kasschau,K.D., Whitham,S.A. and Carrington,J.C. (2002) Loss-of-susceptibility mutants of Arabidopsis thaliana reveal an essential role for eIF(iso)4E during potyvirus infection. *Curr. Biol. CB*, **12**, 1046–1051.

55. Khelifa,M., Massé,D., Blanc,S. and Drucker,M. (2010) Evaluation of the minimal replication time of Cauliflower mosaic virus in different hosts. *Virology*, **396**, 238–245.

56. Evans,P.D., Cook,S.N., Riggs,P.D. and Noren,C.J. (1995) LITMUS: multipurpose cloning vectors with a novel system for bidirectional in vitro transcription. *BioTechniques*, **19**, 130–135.

57. Nakagawa,T., Kurose,T., Hino,T., Tanaka,K., Kawamukai,M., Niwa,Y., Toyooka,K., Matsuoka,K., Jinbo,T. and Kimura,T. (2007) Development of series of gateway binary vectors, pGWBs, for

realizing efficient construction of fusion genes for plant transformation. *J. Biosci. Bioeng.*, **104**, 34–41.

58. Karimi,M., Inzé,D. and Depicker,A. (2002) GATEWAY vectors for Agrobacterium-mediated plant transformation. *Trends Plant Sci.*, **7**, 193–195.

59. Curtis,M.D. and Grossniklaus,U. (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol.*, **133**, 462–469.

60. Smith,D. and Corcoran,L. (1990) Expression and purification of glutathione-S-transferase fusion proteins. In *Current protocols in molecular biology*. Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K, New York, NY: John Wiley, pp. 16.7.1–16.7.8.

61. Geldreich,A., Haas,G., Kubina,J., Bouton,C., Tanguy,M., Erhardt,M., Keller,M., Ryabova,L. and Dimitrova,M. (2017) Formation of large viroplasms and virulence of Cauliflower mosaic virus in turnip plants depend on the N-terminal EKI sequence of viral protein TAV. *PLOS ONE*, **12**, e0189062.

62. Mesnard,J.M., Mougeot,J.L., Geldreich,A. and Lebeurier,G. (1993) Characterization of different electrophoretic forms of cauliflower mosaic virus virions (strain Cabb-S). *Biochimie*, **75**, 645–649.

63. Xiong,C., Muller,S., Lebeurier,G. and Hirth,L. (1982) Identification by immunoprecipitation of cauliflower mosaic virus in vitro major translation product with a specific serum against viroplasm protein. *EMBO J.*, **1**, 971–976.

64. Yoo,S.-D., Cho,Y.-H. and Sheen,J. (2007) *Arabidopsis* mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat. Protoc.*, **2**, 1565–1572.

65. Dimitrova,M., Imbert,I., Kieny,M.P. and Schuster,C. (2003) Protein-Protein Interactions between Hepatitis C Virus Nonstructural Proteins. *J. Virol.*, **77**, 5401–5414.

66. Sorek,N., Gutman,O., Bar,E., Abu-Abied,M., Feng,X., Running,M.P., Lewinsohn,E., Ori,N., Sadot,E., Henis,Y.I., *et al.* (2011) Differential effects of prenylation and s-acylation on type I and II ROPS membrane interaction and function. *Plant Physiol.*, **155**, 706–720.

67. Quaresma,A.J.C., Sievert,R. and Nickerson,J.A. (2013) Regulation of mRNA export by the PI3 kinase/AKT signal transduction pathway. *Mol. Biol. Cell*, **24**, 1208–1221.

68. Dibble,C.C. and Cantley,L.C. (2015) Regulation of mTORC1 by PI3K signaling. *Trends Cell Biol.*, **25**, 545–555.

69. Schepetilnikov,M., Makarian,J., Srour,O., Geldreich,A., Yang,Z., Chicher,J., Hammann,P. and Ryabova,L.A. (2017) GTPase ROP2 binds and promotes activation of target of rapamycin, TOR, in response to auxin. *EMBO J.*, **36**, 886–903.

70. Schepetilnikov,M., Kobayashi,K., Geldreich,A., Caranta,C., Robaglia,C., Keller,M. and Ryabova,L.A. (2011) Viral factor TAV recruits TOR/S6K1 signalling to activate reinitiation after long ORF translation. *EMBO J.*, **30**, 1343–1356.

71. Schoelz,J.E., Angel,C.A., Nelson,R.S. and Leisner,S.M. (2016) A model for intracellular movement of Cauliflower mosaic virus: the concept of the mobile virion factory. *J. Exp. Bot.*, **67**, 2039–2048.

72. Pooggin,M.M. and Ryabova,L.A. (2018) Ribosome Shunting, Polycistronic Translation, and Evasion of Antiviral Defenses in Plant Pararetroviruses and Beyond. *Front. Microbiol.*, **9**.

73. Revers,F. and García,J.A. (2015) Molecular biology of potyviruses. *Adv. Virus Res.*, **92**, 101–199.

74. Bonneville,J.M., Sanfaçon,H., Fütterer,J. and Hohn,T. (1989) Posttranscriptional trans-activation in cauliflower mosaic virus. *Cell*, **59**, 1135–1143.

75. Gowda,S., Wu,F.C., Scholthof,H.B. and Shepherd,R.J. (1989) Gene VI of figwort mosaic virus (caulimovirus group) functions in posttranscriptional expression of genes on the full-length RNA transcript. *Proc. Natl. Acad. Sci. U. S. A.*, **86**, 9203–9207.

76. Fütterer,J., Gordon,K., Sanfaçon,H., Bonneville,J.M. and Hohn,T. (1990) Positive and negative control of translation by the leader sequence of cauliflower mosaic virus pregenomic 35S RNA. *EMBO J.*, **9**, 1697–1707.

77. Scholthof,H.B., Gowda,S., Wu,F.C. and Shepherd,R.J. (1992) The full-length transcript of a caulimovirus is a polycistronic mRNA whose genes are trans activated by the product of gene VI. *J. Virol.*, **66**, 3131–3139.

78. Takatsuji,H., Yamauchi,H., Watanabe,S., Kato,H. and Ikeda,J.E. (1992) Cauliflower mosaic virus reverse transcriptase. Activation by proteolytic processing and functional alteration by terminal deletion. *J. Biol. Chem.*, **267**, 11579–11585.

79. Ani,R.A., Pfeiffer,P., Lebeurier,G. and Hirth,L. (1979) The structure of cauliflower mosaic virus I. pH-Induced structural changes. *Virology*, **93**, 175–187.

80. Martinez-Izquierdo,J. and Hohn,T. (1987) Cauliflower mosaic virus coat protein is phosphorylated in vitro by a virion-associated protein kinase. *Proc. Natl. Acad. Sci. U. S. A.*, **84**, 1824–1828.

81. Karsies,A., Hohn,T. and Leclerc,D. (2001) Degradation signals within both terminal domains of the cauliflower mosaic virus capsid protein precursor. *Plant J.*, **27**, 335–343.

82. Pfeiffer,P., Laquel,P. and Hohn,T. (1984) Cauliflower Mosaic Virus replication complexes: characterization of the associated enzymes and of the polarity of the DNA synthesized in vitro. *Plant Mol. Biol.*, **3**, 261–270.

83. Haseloff,J. and Siemering,K. (1998) The uses of green fluorescent protein in plants. In *Green fluorescent protein: properties, applications and protocols*. Chalfie M, Kain S, New York: Wiley-Liss, Inc., pp. 191–220.

84. Hanson,M.R. and Köhler,R.H. (2001) GFP imaging: methodology and application to investigate cellular compartmentation in plants. *J. Exp. Bot.*, **52**, 529–539.

85. Kosugi,S., Hasebe,M., Tomita,M. and Yanagawa,H. (2009) Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 10171–10176.

86. Sperschneider,J., Catanzariti,A.-M., DeBoer,K., Petre,B., Gardiner,D.M., Singh,K.B., Dodds,P.N. and Taylor,J.M. (2017) LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. *Sci. Rep.*, **7**, 44598.

87. Hemmings-Mieszczak,M., Steger,G. and Hohn,T. (1997) Alternative structures of the cauliflower mosaic virus 35 S RNA leader: implications for viral expression and replication. *J. Mol. Biol.*, **267**, 1075–1088.

88. Hohn,T. and Rothnie,H. (2013) Plant pararetroviruses: replication and expression. *Curr. Opin. Virol.*, **3**, 621–628.

89. Guerra-Peraza,O., de Tapia,M., Hohn,T. and Hemmings-Mieszczak,M. (2000) Interaction of the cauliflower mosaic virus coat protein with the pregenomic RNA leader. *J. Virol.*, **74**, 2067–2072.

90. Pfeiffer,P. and Hohn,T. (1983) Involvement of reverse transcription in the replication of cauliflower mosaic virus: a detailed model and test of some aspects. *Cell*, **33**, 781–789.

91. Köhler,A. and Hurt,E. (2007) Exporting RNA from the nucleus to the cytoplasm. *Nat. Rev. Mol. Cell Biol.*, **8**, 761–773.

92. Pan,H., Liu,S. and Tang,D. (2012) HPR1, a component of the THO/TREX complex, plays an important role in disease resistance and senescence in Arabidopsis. *Plant J. Cell Mol. Biol.*, **69**, 831–843.

93. Xu,C., Zhou,X. and Wen,C.-K. (2015) HYPER RECOMBINATION1 of the THO/TREX complex plays a role in controlling transcription of the REVERSION-TO-ETHYLENE SENSITIVITY1 gene in Arabidopsis. *PLoS Genet.*, **11**, e1004956.

94. Hohn,T. and Vazquez,F. (2011) RNA silencing pathways of plants: silencing and its suppression by plant DNA viruses. *Biochim. Biophys. Acta*, **1809**, 588–600.

95. Allen,E., Xie,Z., Gustafson,A.M. and Carrington,J.C. (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*, **121**, 207–221.

96. Yoshikawa,M., Peragine,A., Park,M.Y. and Poethig,R.S. (2005) A pathway for the biogenesis of trans-acting siRNAs in Arabidopsis. *Genes Dev.*, **19**, 2164–2175.

97. Rajeswaran,R. and Pooggin,M.M. (2012) RDR6-mediated synthesis of complementary RNA is terminated by miRNA stably bound to template RNA. *Nucleic Acids Res.*, **40**, 594–599.

98. Rajeswaran,R., Aregger,M., Zvereva,A.S., Borah,B.K., Gubaeva,E.G. and Pooggin,M.M. (2012) Sequencing of RDR6-dependent double-stranded RNAs reveals novel features of plant siRNA biogenesis. *Nucleic Acids Res.*, **40**, 6241–6254.

99. Shivaprasad,P.V., Rajeswaran,R., Blevins,T., Schoelz,J., Meins,F., Hohn,T. and Pooggin,M.M. (2008) The CaMV transactivator/viroplasmin interferes with RDR6-dependent trans-acting and secondary siRNA pathways in Arabidopsis. *Nucleic Acids Res.*, **36**, 5896–5909.

100. Strässer,K. and Hurt,E. (2000) Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. *EMBO J.*, **19**, 410–420.

101. Zenklusen,D., Vinciguerra,P., Strahm,Y. and Stutz,F. (2001) The yeast hnRNP-Like proteins Yra1p and Yra2p participate in mRNA export through interaction with Mex67p. *Mol. Cell. Biol.*, **21**, 4219–4232.

102. Choudury,S.G., Shahid,S., Cuerda-Gil,D., Panda,K., Cullen,A., Ashraf,Q., Sigman,M.J., McCue,A.D. and Slotkin,R.K. (2019) The RNA Export Factor ALY1 Enables Genome-Wide RNA-Directed DNA Methylation. *Plant Cell*, **31**, 759–774.

103. Tang,W. and Leisner,S. (1998) Methylation of nonintegrated multiple copy DNA in plants. *Biochem. Biophys. Res. Commun.*, **245**, 403–406.

104. Blevins,T., Rajeswaran,R., Aregger,M., Borah,B.K., Schepetilnikov,M., Baerlocher,L., Farinelli,L., Meins,F., Hohn,T. and Pooggin,M.M. (2011) Massive production of small RNAs from a non-coding region of Cauliflower mosaic virus in plant defense and viral counter-defense. *Nucleic Acids Res.*, **39**, 5003–5014.

105. Hohn,T. (2015) RNA based viral silencing suppression in plant pararetroviruses. *Front. Plant Sci.*, **6**, 398.

106. Miki,D., Zhang,W., Zeng,W., Feng,Z. and Zhu,J.-K. (2018) CRISPR/Cas9-mediated gene targeting in Arabidopsis using sequential transformation. *Nat. Commun.*, **9**, 1967.

107. Gatfield,D., Le Hir,H., Schmitt,C., Braun,I.C., Köcher,T., Wilm,M. and Izaurralde,E. (2001) The DExH/D box protein HEL/UAP56 is essential for mRNA nuclear export in Drosophila. *Curr. Biol. CB*, **11**, 1716–1721.

108. MacMorris,M., Brocker,C. and Blumenthal,T. (2003) UAP56 levels affect viability and mRNA export in Caenorhabditis elegans. *RNA N. Y. N*, **9**, 847–857.

109. Strässer,K. and Hurt,E. (2001) Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p. *Nature*, **413**, 648–652.

110. Kapadia,F., Pryor,A., Chang,T.-H. and Johnson,L.F. (2006) Nuclear localization of poly(A)+ mRNA following siRNA reduction of expression of the mammalian RNA helicases UAP56 and URH49. *Gene*, **384**, 37–44.

111. Kistler,A.L. and Guthrie,C. (2001) Deletion of MUD2, the yeast homolog of U2AF65, can bypass the requirement for sub2, an essential spliceosomal ATPase. *Genes Dev.*, **15**, 42–49.

112. Kang,Y. and Cullen,B.R. (1999) The human Tap protein is a nuclear mRNA export factor that contains novel RNA-binding and nucleocytoplasmic transport sequences. *Genes Dev.*, **13**, 1126–1139.

113. Cullen,B.R. (2003) Nuclear RNA export. *J. Cell Sci.*, **116**, 587–597.

114. LeBlanc,J.J., Uddowla,S., Abraham,B., Clatterbuck,S. and Beemon,K.L. (2007) Tap and Dbp5, but not Gag, are involved in DR-mediated nuclear export of unspliced Rous sarcoma virus RNA. *Virology*, **363**, 376–386.

115. Sakuma,T., Davila,J.I., Malcolm,J.A., Kocher,J.-P.A., Tonne,J.M. and Ikeda,Y. (2014) Murine leukemia virus uses NXF1 for nuclear export of spliced and unspliced viral transcripts. *J. Virol.*, **88**, 4069–4082.

116. Pessel-Vivares,L., Houzet,L., Lainé,S. and Mougel,M. (2015) Insights into the nuclear export of murine leukemia virus intron-containing RNA. *RNA Biol.*, **12**, 942–949.

117. Flint,S.J., Huang,W., Goodhouse,J. and Kyin,S. (2005) A peptide inhibitor of exportin1 blocks shuttling of the adenoviral E1B 55 kDa protein but not export of viral late mRNAs. *Virology*, **337**, 7–17.

118. Yatherajam,G., Huang,W. and Flint,S.J. (2011) Export of adenoviral late mRNA from the nucleus requires the Nxf1/Tap export receptor. *J. Virol.*, **85**, 1429–1438.

119. Johnson,L.A., Li,L. and Sandri-Goldin,R.M. (2009) The cellular RNA export receptor TAP/NXF1 is required for ICP27-mediated export of herpes simplex virus 1 RNA, but the TREX complex adaptor protein Aly/REF appears to be dispensable. *J. Virol.*, **83**, 6335–6346.

120. Tian,X., Devi-Rao,G., Golovanov,A.P. and Sandri-Goldin,R.M. (2013) The interaction of the cellular export adaptor protein Aly/REF with ICP27 contributes to the efficiency of herpes simplex virus 1 mRNA export. *J. Virol.*, **87**, 7210–7217.

121. Heise,T., Sommer,G., Reumann,K., Meyer,I., Will,H. and Schaal,H. (2006) The hepatitis B virus PRE contains a splicing regulatory element. *Nucleic Acids Res.*, **34**, 353–363.

122. Lee,G.H., Wasser,S. and Lim,S.G. (2008) Hepatitis B pregenomic RNA splicing—The products, the regulatory mechanisms and its biological significance. *Virus Res.*, **136**, 1–7.

123. Sommer,G. and Heise,T. (2008) Posttranscriptional control of HBV gene expression. *Front. Biosci. J. Virtual Libr.*, **13**, 5533–5547.

124. Hernandez,F.P. and Sandri-Goldin,R.M. (2010) Head-to-tail intramolecular interaction of herpes simplex virus type 1 regulatory protein ICP27 is important for its interaction with cellular mRNA export receptor TAP/NXF1. *mBio*, **1**.

125. Diot,C., Fournier,G., Dos Santos,M., Magnus,J., Komarova,A., van der Werf,S., Munier,S. and Naffakh,N. (2016) Influenza A Virus Polymerase Recruits the RNA Helicase DDX19 to Promote the Nuclear Export of Viral mRNAs. *Sci. Rep.*, **6**, 33763.

126. Ren,X., Yu,Y., Li,H., Huang,J., Zhou,A., Liu,S., Hu,P., Li,B., Qi,W. and Liao,M. (2019) Avian Influenza A Virus Polymerase Recruits Cellular RNA Helicase eIF4A3 to Promote Viral mRNA Splicing and Spliced mRNA Nuclear Export. *Front. Microbiol.*, **10**, 1625.

127. Marsh,L.E. and Guilfoyle,T.J. (1987) Cauliflower mosaic virus replication intermediates are encapsidated into virion-like particles. *Virology*, **161**, 129–137.

128. Himmelbach,A., Chapdelaine,Y. and Hohn,T. (1996) Interaction between cauliflower mosaic virus inclusion body protein and capsid protein: implications for viral assembly. *Virology*, **217**, 147–157.

129. Menissier,J., Laquel,P., Lebeurier,G. and Hirth,L. (1984) A DNA polymerase activity is associated with Cauliflower Mosaic Virus. *Nucleic Acids Res.*, **12**, 8769–8778.

130. Haas,M., Bureau,M., Geldreich,A., Yot,P. and Keller,M. (2002) Cauliflower mosaic virus: still in the news. *Mol. Plant Pathol.*, **3**, 419–429.

131. Kobayashi,K., Nakayashiki,H., Tsuge,S., Mise,K. and Furusawa,I. (1998) Accumulation kinetics of viral gene products in cauliflower mosaic virus-infected turnip protoplasts. *Microbiol. Immunol.*, **42**, 65–69.

132. Sugiura,T., Sakurai,K. and Nagano,Y. (2007) Intracellular characterization of DDX39, a novel growth-associated RNA helicase. *Exp. Cell Res.*, **313**, 782–790.

133. Taniguchi,I. and Ohno,M. (2008) ATP-dependent recruitment of export factor Aly/REF onto intronless mRNAs by RNA helicase UAP56. *Mol. Cell. Biol.*, **28**, 601–608.

134. Li,H.-C., Huang,E.-Y., Su,P.-Y., Wu,S.-Y., Yang,C.-C., Lin,Y.-S., Chang,W.-C. and Shih,C. (2010) Nuclear export and import of human hepatitis B virus capsid protein and particles. *PLoS Pathog.*, **6**, e1001162.

135. Love,A.J., Laird,J., Holt,J., Hamilton,A.J., Sadanandom,A. and Milner,J.J. (2007) Cauliflower mosaic virus protein P6 is a suppressor of RNA silencing. *J. Gen. Virol.*, **88**, 3439–3444.

136. Haas,G., Azevedo,J., Moissiard,G., Geldreich,A., Himber,C., Bureau,M., Fukuhara,T., Keller,M. and Voinnet,O. (2008) Nuclear import of CaMV P6 is required for infection and suppression of the RNA silencing factor DRB4. *EMBO J.*, **27**, 2102–2112.

137. Guilley,H., Dudley,R.K., Jonard,G., Balàzs,E. and Richards,K.E. (1982) Transcription of Cauliflower mosaic virus DNA: detection of promoter sequences, and characterization of transcripts. *Cell*, **30**, 763–773.

138. Yang,X., Yang,Y., Sun,B.-F., Chen,Y.-S., Xu,J.-W., Lai,W.-Y., Li,A., Wang,X., Bhattarai,D.P., Xiao,W., *et al.* (2017) 5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m5C reader. *Cell Res.*, **27**, 606–625.

139. David,R., Burgess,A., Parker,B., Li,J., Pulsford,K., Sibbritt,T., Preiss,T. and Searle,I.R. (2017) Transcriptome-Wide Mapping of RNA 5-Methylcytosine in Arabidopsis mRNAs and Noncoding RNAs. *Plant Cell*, **29**, 445–460.

140. Hu,J., Manduzio,S. and Kang,H. (2019) Epitranscriptomic RNA Methylation in Plant Development and Abiotic Stress Responses. *Front. Plant Sci.*, **10**, 500.

141. Li,Y., Bor,Y.-C., Misawa,Y., Xue,Y., Rekosh,D. and Hammarskjöld,M.-L. (2006) An intron with a constitutive transport element is retained in a Tap messenger RNA. *Nature*, **443**, 234–237.

142. Vogt,C. and Bohne,J. (2016) The KSHV RNA regulator ORF57: target specificity and its role in the viral life cycle. *Wiley Interdiscip. Rev. RNA*, **7**, 173–185.

143. Ernst,R.K., Bray,M., Rekosh,D. and Hammarskjöld,M.L. (1997) A structured retroviral RNA element that mediates nucleocytoplasmic export of intron-containing RNA. *Mol. Cell. Biol.*, **17**, 135–144.

144. Haasen,D., Köhler,C., Neuhaus,G. and Merkle,T. (1999) Nuclear export of proteins in plants: AtXPO1 is the export receptor for leucine-rich nuclear export signals in Arabidopsis thaliana. *Plant J. Cell Mol. Biol.*, **20**, 695–705.

145. Blanvillain,R., Boavida,L.C., McCormick,S. and Ow,D.W. (2008) EXPORTIN1 Genes Are Essential for Development and Function of the Gametophytes in Arabidopsis thaliana. *Genetics*, **180**, 1493–1500.

3. Résultats supplémentaires

3.1. La voie de l'exportine 1 participe à l'export nucléaire des transcrits 35S du CaMV

Nos résultats démontrent que la voie majeure d'export des ARNm cellulaires régulée par le complexe TREX-1 est empruntée par le CaMV pour acheminer ses ARN 35S épissés ou non, dans le cytoplasme. Cependant, il n'est pas exclu, que le CaMV utilise aussi l'autre voie d'export des ARNm chez les métazoaires, la voie de l'exportine 1 (CRM1), puisque comme nous l'avons vu l'export nucléaire des ARN du CaMV n'est pas complètement altéré chez les mutants du complexe TREX-1. De plus, nous savons que la voie CRM1 est également employée pour l'export de certains ARNm chez les métazoaires et, exploitée par des rétrovirus complexes des animaux comme le VIH-1 et le PFV - *Prototype Foamy Virus* – pour transporter leurs ARN partiellement et non épissés. L'utilisation de la voie CRM1 par la PFV est d'autant plus intéressante dans le cadre de notre étude, que ce virus partage de nombreuses propriétés structurales et fonctionnelles avec le CaMV (Bodem et al., 2011). Par conséquent, nous avons réalisé différentes expériences pour déterminer si CRM1, l'exportine XPO1 chez Arabidopsis, est aussi impliquée dans l'export nucléaire des ARN 35S du CaMV.

Les exportines XPO1a et XPO1b d'*A. thaliana* (Blanvillain et al., 2008) sont impliquées dans l'export des protéines (Merkle, 2011) et probablement dans celui des ARN mais aucune preuve concrète n'a été apportée quant à l'export des ARNm matures, coiffés et polyadénylés, par ces transporteurs (Haasen et al., 1999; Stuger et Forreiter, 2004).

La première approche a été naturellement de tester dans notre système de protoplastes, l'effet de la leptomycine B (LMB)(Mathew et Ghildyal, 2017) sur l'export des ARN viraux, un métabolite produit par *Streptomyces spp.* qui inhibe spécifiquement l'exportine CRM1 en se liant de manière covalente et irréversible au domaine impliqué dans la liaison avec le NES des protéines cargos.

Des protoplastes d'*A. thaliana* Col-0 et *CA-rop2* ont été transfectés avec le plasmide viral pCaMV-GFP puis incubés, 20 h après leur transfection, avec la LMB à une concentration finale de 100 nM, pendant 1 h. Des expériences contrôles ont été réalisées en parallèle, en incubant les protoplastes transfectés avec du méthanol, solvant dans lequel est conservé l'inhibiteur, afin de déterminer si la LMB et/ou le méthanol perturbent l'expression des transcrits viraux.

Des analyses par qRT-PCR effectuées sur les fractions totales de protoplastes montrent clairement que les quantités d'ARN 35S produites sont pratiquement identiques dans les protoplastes contrôles et ceux traités à la LMB, qu'ils viennent des plants d'Arabidopsis sauvage ou mutants *CA-rop2* et, que l'emploi de cet inhibiteur n'a aucun incidence sur l'étape de transcription du génome du CaMV (**figure R-10**).

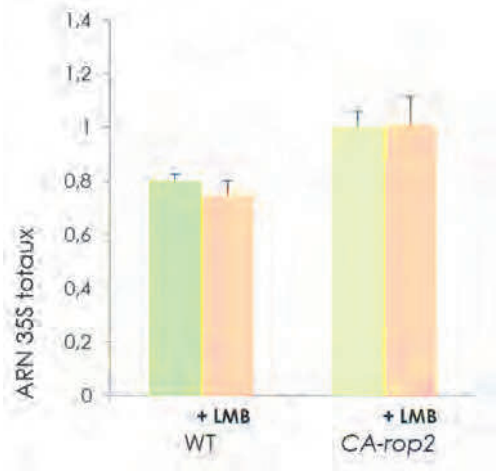


Figure R-10 | Analyse des quantités d'ARN 35S total dans les échantillons de protoplastes sauvages et mutants CA-rop2 traités ou non avec la leptomycine B. Quantification relative par RT-qPCR des ARN 35S épissés et non épissé du CaMV dans les fractions totales des protoplastes sauvages (WT) et mutants *CA-rop2* transfectés avec pCaMV-GFP et traités (LMB) ou non avec l'inhibiteur. Les valeurs correspondant aux ARN 35S ont été normalisées par rapport à celles obtenues pour les ARN messagers contrôles cellulaires GAPDH et de transfection eGFP (ΔCt). Les rapports ARN 35S cytoplasmiques/nucléaires ont ensuite été calculés. Les valeurs de ces rapports sans l'inhibiteur sont fixées à 1 et celles obtenues en sa présence sont calculées en fonction de celles-ci (ΔΔCt). Les barres d'erreurs indiquent les écart-types observés entre les différents échantillons.

L'efficacité d'inhibition de la voie XPO1 par la leptomycine B a été déterminée en observant le comportement de la protéine de fusion eGFP-A (42 kDa) dont le domaine A renferme le NES de la protéine P6. Haas et al., (2005) ont montré que l'export nucléaire de cette protéine est inhibé dans des cellules de tabac BY-2 traitées avec la LMB. Les fractions totales, cytoplasmiques et nucléaires préparées à partir des protoplastes traitées (1 et 4 h) ou non à la LMB ont été analysées par western blot à l'aide d'anticorps dirigés contre la GFP et la protéine P6 (**figure R-11**).

Comme le montre la révélation de la protéine eGFP-A avec les anticorps anti-GFP et anti-P6, la protéine de fusion est retenue dans le noyau lorsque les protoplastes sont traités à la LMB, suggérant que l'export CRM1 dépendant de la protéine de fusion est inhibé et démontrant ainsi que la LMB a efficacement inhibé l'exportine **(figure R-11 N).** Très peu de différence est observée dans la rétention nucléaire de eGFP-A si le traitement est prolongé jusqu'à 4 h. La présence en faible quantité de eGFP-A dans le noyau **(figure R-11 N**, **piste M)** s'explique par le fait que cette protéine entre par diffusion dans le noyau même si elle ne possède pas de NLS (Haas et al., 2005).

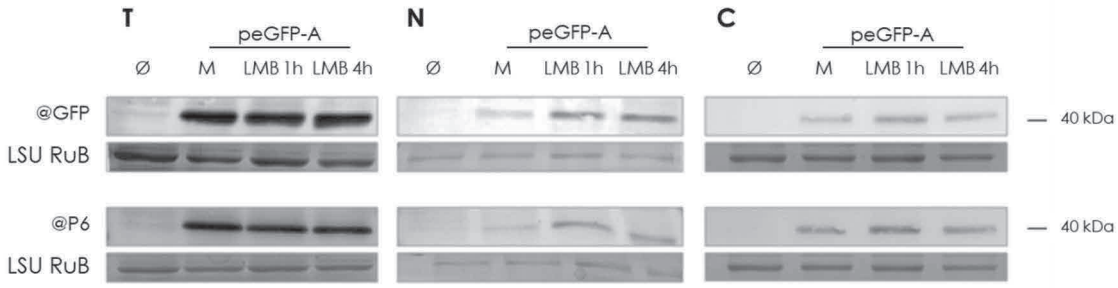


Figure R-11 | Analyse, par immunodétection de la protéine de fusion eGFP-A, de l'inhibition de la voie XPO1 par la leptomycine B (LMB). Les protéines des fractions totales (**T**), nucléaires (**N**) et cytoplasmiques (**C**) des protoplastes transfectés ou non (Ø) avec le plasmide peGFP-A, permettant l'expression de la protéine de fusion eGFP-A, ont été séparées par électrophorèse sur gel de polyacrylamide 15 %, en conditions dénaturantes, puis révélées par coloration au bleu de Coomassie™ ou transférées sur membrane. Les protoplastes ont été traités pendant 1 (LMB 1 h) ou 4 (LMB 4 h) heures à la leptomycine B (100 nM) ou non (M). La protéine a été recherchée dans les fractions par western blot, à l'aide d'anticorps polyclonaux dirigés soit contre l'eGFP (@GFP) ou le domaine A de P6 (@P6). La masse moléculaire (en kDa) des protéines est indiquée à droite des figures. M : méthanol, solvant de la leptomycine B. La coloration du gel montre la grande sous-unité de la RuBisCO (LSU RuB)(55 kDa).

Suite à ces analyses préliminaires indispensables, nous avons étudié l'importance de XPO1 dans l'export des ARN 35S du CaMV, en déterminant, après RT-qPCR, le rapport entre les ARN 35S épissés et non épissés cytoplasmiques et nucléaires de protoplastes transfectés par le plasmide viral pCaMV-GFP et traités ou non à la LMB. Nos résultats montrent que ce rapport est diminué pratiquement de la même manière (à environ 40 %) dans les protoplastes d'Arabidopsis sauvage et mutants *CA-rop2*, lorsqu'ils sont traités à la leptomycine B. Ces résultats démontrent une importante rétention des transcrits 35S du CaMV dans les noyaux des deux types de protoplastes lorsque l'exportine 1 est inhibée par la LMB, suggérant que XPO1 est impliquée dans leur export (**figure R-12 A**). Par ailleurs, ils confirment aussi que l'activation constitutive de ROP2 ne modifie pas le processus d'export des ARN viraux, même lorsqu'il est bloqué pour certains d'entre eux.

Pour confirmer ces résultats, les mêmes expériences ont été réalisées avec des protoplastes d'Arabidopsis n'exprimant pas la protéine XPO1a (*xpo1a-3*), le double mutant *xpo1a-3/xpo1b-1* étant létal (Blanvillain et al., 2008). Nous constatons qu'une forte diminution de ce rapport est obtenue dans les protoplastes mutants par rapport aux protoplastes sauvages. En d'autres termes, une rétention nucléaire très forte des ARN 35S en absence de XPO1a. Cette rétention est exacerbée avec le traitement à la leptomycine B, confirmant ainsi l'implication de la voie CRM1/XPO1 dans l'export des ARN du CaMV chez *A. thaliana* (**figure R-12 B**).

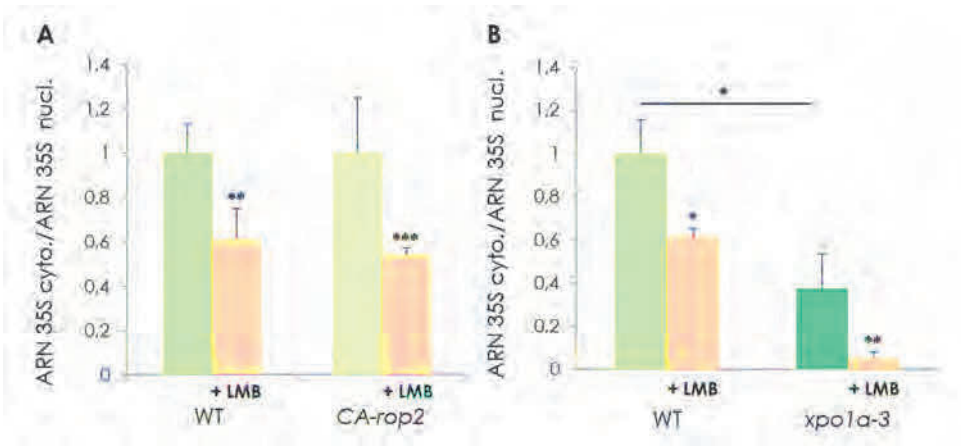


Figure R-12 | Analyse des ARN 35S dans les fractions de protoplastes sauvages et mutants *CA-rop2* et *xpo1a-3* traités ou non avec la leptomycine B (LMB). Quantification relative par RT-qPCR des ARN 35S épissés et non épissés de CaMV dans les fractions cytoplasmiques et nucléaires des protoplastes sauvages (WT) et mutants *CA-rop2* (**A**) ou *xpo1a-3* (**b**) transfectés avec pCaMV-GFP et traités (LMB) ou non avec l'inhibiteur. Les valeurs correspondant aux ARN 35S ont été normalisées par rapport à celles obtenues pour les ARN messagers contrôles cellulaire GAPDH et de transfection eGFP (ΔCt). Les rapports ARN 35S cytoplasmiques/nucléaires ont ensuite été calculés. Les valeurs de ces rapports sans l'inhibiteur sont fixées à 1 et celles obtenues en sa présence sont calculées en fonction de celles-ci (ΔΔCt). Les données des différents expériences ont été vérifiées statistiquement par un test de Student (* p<0,05, ** p<0,01 et ***p<0,001). Les barres d'erreurs indiquent les écart-types observés entre les différents échantillons.

Finalement, des expériences réalisées avec des protoplastes d'Arabidopsis mutants *4xaly*, incubés ou non avec la LMB, montrent une nette diminution de l'export des ARN 35S du CaMV en présence de l'inhibiteur, le rapport des ARN35S cytoplasmiques et nucléaires étant divisé quasiment de moitié (**figure R-13**).

En conclusion, les deux voies d'export des ARNm cellulaires, la voie majeure impliquant TREX-1 et la voie mineure de XPO1/CRM1 sont utilisées par le CaMV pour exporter les différents isoformes de l'ARN 35S, épissées et non épissée chez *A. thaliana*.

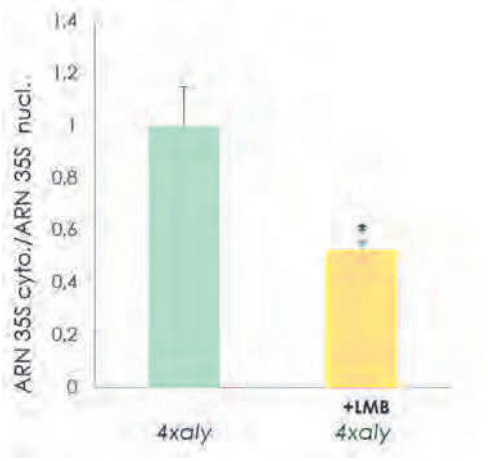


Figure R-13 | Analyse des ARN 35S dans les fractions de protoplastes mutants *4xaly* traités ou non avec la leptomycine B (LMB). Quantification relative par RT-qPCR des ARN 35S épissés et non épissés de CaMV dans les fractions cytoplasmiques et nucléaires des protoplastes mutants 4xaly transfectés avec pCaMV-GFP et traités (LMB) ou non avec l'inhibiteur. Les valeurs correspondant aux ARN 35S ont été normalisées par rapport à celles obtenues pour les ARN messagers contrôles cellulaire GAPDH et de transfection eGFP (ΔCt). Les rapports ARN 35S cytoplasmiques/nucléaires ont ensuite été calculés. Les valeurs de ces rapports sans l'inhibiteur sont fixées à 1 et celles obtenues en sa présence sont calculées en fonction de celles-ci (ΔΔCt). Les données des différents expériences ont été vérifiées statistiquement par un test de Student (* $p<0,05$,). Les barres d'erreurs indiquent les écart-types observés entre les différents échantillons.

Discussion et perspectives

Discussion et perspectives

Le CaMV effectue une partie de son cycle infectieux dans le noyau où sont produits l'ARN 35S polycistronique et des versions épissées, l'ARN monocistronique 19S et un ARN 8S non codant. Seuls 30% des ARN 35S restent intacts, les 70% restants subissant un épissage alternatif dans la région 5' codante de sorte que tous ces ARN 35S constituent une population très hétérogène de messagers (Kiss-László et al., 1995; Bouton et al., 2015a). Les ARN 35S complet et épissés sont exportés dans le cytoplasme où ils sont traduits et rétrotranscrits (Hirochika et al., 1985).

Les résultats que nous avons présentés dans cette thèse, permettent d'apporter des éléments de réponse concernant l'export nucléaire de l'ARN 35S du CaMV et de ses formes épissées, une étape du cycle viral qui est encore totalement méconnue chez ce virus et plus généralement, chez tous les phytovirus à ADN.

Chez les métazoaires et la levure, l'export des ARNm implique le complexe TREX-1 pour rendre les ARNm compétents pour leur export par la voie du dimère d'exportine TAP-p15 (Heath et al., 2016). Ce transporteur n'a pas été caractérisé chez les plantes malgré des analyses approfondies, contrairement au complexe TREX-1 dont la composition en protéines a été établie chez Arabidopsis (Ehrnsberger et al., 2019a). Le complexe TREX-1 montre chez les plantes, une certaine plasticité grâce à l'existence de plusieurs isoformes pour certains composants dont les protéines adaptatrices ALY et UIEF qui sont requises pour recruter les transporteurs et pour lesquels il existe respectivement 4 et 2 isoformes (Pfaff et al., 2018; Ehrnsberger et al., 2019b).

Des mutants simples ou multiples d'Arabidopsis qui n'expriment plus ALY, MOS11 et/ou TEX1 ou déficients en certaines isoformes de ALY, inoculés par un isolat virulent de CaMV, Cabb B-JI, montrent une résistance partielle à l'infection virale à l'exception du double mutant *aly1 aly2* où la cinétique d'infection est pratiquement identique à celle de la plante sauvage. Les doubles mutations *aly3 aly4* ont un effet plus prononcées sur l'infection virale vu que le taux maximal (80%) de plants infectés systémiquement n'est atteint que près de trois semaines après celui (97%) obtenu avec les plantes sauvages. L'effet des mutations des 4 gènes *aly* est dramatique, les plants d'Arabidopsis (*4xaly*) étant difficilement infectés par le CaMV. Seuls 17% d'entre eux, montrent les symptômes d'une infection systémique, une mosaïque foliaire et un éclaircissement des nervures qui s'additionnent aux changements phénotypiques de la plante, occasionnés par l'absence des 4 protéines ALY.

Parallèlement à ces tests *in planta*, nous avons montré en quantifiant par qRT-PCR, les ARN 35S nucléaires et cytoplasmiques, sans faire de distinction entre formes épissées et non épissé, dans des protoplastes préparés à partir de plants mutants et, transfectés avec un

vecteur viral (pCaMV-GFP), que l'export nucléaire des ARN 35S est impacté de façon drastique chez les doubles mutants *aly1 aly2* et *aly3 aly4* et, encore plus sévèrement chez le quadruple mutant *4xaly* comparé à l'export dans les protoplastes de plantes sauvages (et du mutant d'Arabidopsis *CA-rop2*), alors que la quantité d'ARN 35S totale produite chez tous ces protoplastes est équivalente. Des tests au préalable ont montré que l'export des ARN 35S viraux est identique chez les Arabidopsis sauvages et le mutant *CA-rop2,* choisi car la transfection des protoplastes par le vecteur pCaMV-GFP est plus efficace qu'avec les protoplastes sauvages. La rétention nucléaire des ARN 35S chez le double mutant *aly1 aly2* est surprenante, car le CaMV infecte cette plante mutée et la plante sauvage suivant une même cinétique ce qui laissait supposer que le taux d'export des ARN viraux était maintenu chez ce mutant grâce aux deux autres isoformes d'ALY (3 et 4), vu la redondance fonctionnelle de ces protéines. En effet, Pfaff et al., (2018) ont montré que les simples (*aly1-4*) et doubles mutants (*aly1 aly2* et *aly3 aly4*) d'Arabidopsis ont un phénotype sauvage et que l'export des ARNm cellulaires n'est pas défectueux, l'absence d'une ou deux isoformes ALY chez ces mutants étant compensée par les autres. Par contre, le quadruple mutant *4xaly* se caractérise par une forte accumulation des ARNm dans le noyau et par des modifications morphologiques prononcées de la plante (Pfaff et al., 2018). L'incapacité du CaMV à infecter les mutants *4xaly* d'Arabidopsis est donc certainement due à l'effet combiné de l'accumulation nucléaire des ARN 35S et, des ARNm cellulaires codant sans doute pour des facteurs indispensables à la réplication virale.

Collectivement, ces résultats montrent que la rétention nucléaire des ARN 35S du CaMV est liée à la déficience en protéines ALY et que l'ensemble des 4 isoformes de ALY est requis pour un export efficace des transcrits viraux. Ils suggèrent aussi que d'autres protéines adaptatrices probablement les protéines UIEF (1 et 2) (Ehrnsberger et al., 2019b) et des protéines SR, contribuent à l'export des ARN du CaMV mais avec très peu d'efficacité ; le taux d'export étant faible chez les double et quadruple mutants *aly* d'Arabidopsis. L'étude du sextuple mutant *4xaly 2xuief* permettrait de confirmer l'implication des protéines UIEF dans l'export des ARN viraux. Toutefois, malgré la forte inhibition de l'export nucléaire chez les mutants d'Arabidopsis, les transcrits 35S transportés dans le cytoplasme restent en quantité suffisante pour que le CaMV puisse se répliquer et infecter systémiquement les plantes. Les protéines virales, dans les feuilles d'Arabidopsis infectées systémiquement par le CaMV, sont d'ailleurs présentes en quantités équivalentes chez les plantes mutées et sauvage.

Les protéines adaptatrices d'Arabidopsis sélectionnent-elles indifféremment tous les ARNm ou sont-elles spécialisées dans l'export de certains messagers particuliers, comme le suggère l'existence de plusieurs isoformes chez les plantes ? Des études récentes indiquent que tel pourrait être effectivement le cas puisque qu'ALY1 contribue chez Arabidopsis à l'export d'une classe d'ARNm particulière (Choudury et al., 2019; Ehrnsberger et al., 2019a). L'analyse des ARN 35S cytoplasmiques chez des mutants d'Arabidopsis infectés par le CaMV, permettrait de déterminer si les protéines ALY sont impliquées indifféremment dans l'export de tous les

transcrits 35S ou si certaines d'entre elles sont plus adaptées à l'export de l'ARN 35S complet ou des formes épissées.

Des expériences similaires réalisées avec des mutants *mos11* et/ou *tex1* d'Arabidopsis confirment le rôle du complexe TREX-1 dans l'export des transcrits du CaMV. L'absence de la protéine MOS11, l'orthologue de CIP29, conduit à une résistance partielle des plantes mutées à l'infection virale et à une accumulation nucléaire des ARN viraux, mais moins importante que celle observe avec les mutants *aly*. L'inhibition de l'export des ARN 35S est amplifiée par l'action combinée des mutations dans les gènes *MOS11* et *TEX1 ;* ce dernier code pour une protéine du complexe THO de TREX-1 et joue un rôle dans l'épissage (Ehrnsberger et al., 2019a). Il est étonnant que l'absence des protéines MOS11 et de TEX1 n'ait pas plus d'impact sur l'export des ARN vu leurs rôles au sein du complexe TREX-1, plus particulièrement, MOS11 qui est impliquée dans l'activation de l'ARN hélicase UAP56. En effet, à l'inverse de ALY et UIEF, il n'existe pas d'isoformes pour MOS11 et TEX1 permettant de complémenter ces mutants, ces deux protéines étant codées chez Arabidopsis par un gène unique.

L'ensemble de ces résultats montre clairement que les ARN 35S du CaMV empruntent la voie d'export TREX-1 appelée ainsi par défaut, l'exportine associée à cette voie étant inconnue. Toutefois, l'utilisation à la fois de la voie TREX-1 et celle de CRM1 (XPO1) ne sont pas forcément mutuellement exclusives comme le montrent les études menées chez plusieurs virus des animaux. Ainsi, le VIH-1 utilise à la fois la voie du dimère d'exportine TAP-p15 pour exporter ses ARN épissés et la voie de CRM1 pour exporter ses ARN partiellement et non épissés (Rojas-Araya et al., 2015), ou encore le RSV dont l'ARN génomique semble emprunter les deux voies d'export TAP et CRM1 pour rejoindre le cytoplasme (LeBlanc et al., 2007; Jouvenet et al., 2011). Par ailleurs, le virus foamy, un rétrovirus qui s'apparente, à bien des égards, plus au CaMV qu'aux autres rétrovirus, utilise la voie CRM1 pour l'export de son ARN génomique (Bodem et al., 2011). Enfin, l'exportine CRM1, bien qu'impliquée dans l'export des protéines et des ARN non codants, est aussi sollicitée pour transporter des ARNm cellulaires aux structures particulières et dans certaines conditions physiologiques, chez les métazoaires (Brennan et al., 2000; Topisirovic et al., 2009).

L'étude de l'export nucléaire des ARN du CaMV dans des protoplastes d'Arabidopsis sauvage ou mutants *4xaly,* incubés en présence de leptomycine B, montre que la voie XPO1 contribue de manière non négligeable au transport nucléo-cytoplasmique des transcrits viraux. En effet, cet inhibiteur, spécifique de XPO1, bloque partiellement leur export ; seuls 60% des ARN 35S parvenant à gagner le cytoplasme en sa présence. Ce résultat est confirmé grâce à des expériences similaires réalisées avec des plants d'arabidopsis mutés dans l'un des deux gènes *xpo1* - le double mutant est létal –. La quantité d'ARN 35S cytoplasmique est fortement réduite chez ces plantes mutées et encore plus, en présence de leptomycine B.

Les virus des mammifères usurpent généralement les voies d'export nucléaires grâce à des protéines virales spécifiques qui servent d'adaptateurs pour recruter le complexe TREX-1 et l'exportine TAP-p15 au niveau de leurs ARN génomiques et/ou messagers. Ainsi, les protéines ICP27 de HSV et ORF57 du KSHV, interagissent avec le récepteur TAP et différents composants de TREX-1 dont les protéines adaptatrices ALY (Tunnicliffe et al., 2011) ou UIF (Jackson et al., 2011) ou encore CIP29 (Schumann et al., 2016).

Nous avons donc investigué l'implication éventuelle de certaines protéines du CaMV dans l'export nucléaire de l'ARN 35S et de ses formes épissées. L'étude de la cinétique d'apparition des protéines du CaMV au cours du cycle infectieux montre que P5, est l'une des premières protéines, avec P6 et P1, à être détectée dans les protoplastes de navet infectés, P4, le précurseur des protéines de capside n'apparaissant que tardivement dans le cycle viral (Kobayashi et al., 1998; Khelifa et al., 2010). La présence de la protéine de mouvement P1 en début du cycle, signifie que l'export de l'ARN 35S non épissé s'opère dès les premières heures de l'infection, P1 ne pouvant être exprimée qu'à partir de cet ARN grâce à P6 qui est elle, produite à partir de son propre messager, l'ARN 19S (Kobayashi et al., 1998; Khelifa et al., 2010).

Nous nous sommes intéressés d'abord à P5, le précurseur de la transcriptase inverse et de la protéase à acide aspartique, pour les deux raisons suivantes. En premier, la présence très précoce de P5 au cours de l'infection suggère fortement qu'elle a d'autres fonctions que la transcription inverse, la synthèse de l'ADN viral ne s'effectuant que tardivement dans le cycle infectieux, lorsque la protéine P4 est produite, la rétrotranscription et l'encapsidation étant vraisemblablement couplées chez le CaMV, comme chez les rétrovirus. Deuxièmement, la transcriptase inverse est présente dans les virions transmis par le puceron vecteur ou transportées dans les cellules saines *via* les plasmodesmes. Ces particules virales atteignent les pores nucléaires grâce aux protéines de la capside pour y libérer l'ADN viral qui emporte sans doute avec lui, la protéine P5. Pfeiffer et al., (1984) ont montré que la protéine complète P5 (78 kDa) est associée aux complexes de réplication présents dans le cytoplasme. P5 est vraisemblablement présente au minimum en trois exemplaires par génome, celui-ci étant synthétisé de manière discontinue, un pour produire le brin d'ADN (-) et deux autres pour les deux brins (+). Ces molécules de P5 déversées dans le noyau, pourraient ainsi intervenir dans l'export nucléaire, avec les facteurs d'export cellulaires dont le complexe TREX-1, dès le début du cycle de réplication du CaMV, l'export étant par la suite promu par les protéines P5 nouvellement synthétisées.

L'implication de P5 dans l'export des transcrits viraux a été explorée en protoplastes à l'aide de différents vecteurs viraux mutagénéisés au préalable pour empêcher l'expression des protéines P5 et/ou P6, cette dernière étant indispensable pour l'expression de l'ensemble des protéines virales, P5 comprise, à partir de l'ARN 35S. L'analyse du rapport entre les ARN 35S cytoplasmiques et nucléaires montre à l'évidence que l'absence de P5 conduit à une forte

rétention de ces ARN dans le noyau et par conséquent, que la protéine P5 est un acteur de l'export nucléaire des transcrits viraux. L'implication de P5 dans ce processus a été confirmée par le rétablissement de l'export des ARN viraux lors d'essais de complémentation, en apportant en *trans,* une copie complète de P5, en co-transfectant les protoplastes, avec le plasmide pCaMV-GFP muté dans l'ORF V, pour empêcher l'expression de la protéine, et un vecteur exprimant l'ORF V sauvage. Des expériences de complémentation additionnelles faites avec des fragments polypeptidiques correspondant à la protéinase (N-terminal) et à la transcriptase inverse (C-terminal) montrent qu'aucun des deux polypeptides n'est capable de restaurer l'export nucléaire, ce qui suggère que la protéine P5 complète est requise pour le transport nucléo-cytoplasmique des ARN viraux. Ces résultats semblent étayer notre hypothèse initiale, attribuant un rôle de facteur d'export au précursseur P5. Ils suggèrent aussi que des molécules de P5 doivent être préservées du clivage en *cis* par la protéase à acide aspartique pour conserver sa conformation tridimensionnelle qui est vraisemblablement primordiale pour intervenir dans l'export, notamment pour son interaction avec l'ARN 35S et/ou le complexe TREX-1.

Des expériences de GST-*pull down* ont été effectuées avec la protéine P5 et ses différents domaines, produits *in vitro* afin de voir s'ils interagissent avec des composants du complexe TREX-1 et plus particulièrement, dans un premier temps, avec la protéine MOS11. Des expériences similaires ont été réalisées en parallèle avec la protéine P4, celle-ci pouvant aussi constituer un facteur d'export à l'instar de la protéine core du HBV (Li et al., 2010a) et, avec la protéine P3 du CaMV et la luciférase comme témoins négatifs.

Nos résultats montrent que la protéine P5 complète interagit avec MOS11 et que la transcriptase inverse – RT-RNase H (RH) – est tout particulièrement impliquée dans cette interaction. Le domaine protéase de P5 ne semble pas interagir, du moins fortement avec MOS11, car elle est retrouvée en quantité équivalente dans les fractions liées et non liées de la GST-MOS11. L'activité protéase de la protéine semble d'ailleurs, et de façon surprenante, être stimulée lors de l'interaction de P5 avec MOS11 comme l'atteste la présence à la fois, de P5 complète et d'un polypeptide qui a la même migration électrophorétique que la transcriptase inverse (55 kDa). La stimulation de la protéase par MOS11 n'est pas sans rappeler celle de l'ARN hélicase UAP56, bien qu'il ne s'agisse pas de la même activité enzymatique, au sein du complexe trimérique formé entre MOS11, ALY et UAP56. Il n'est pas exclu que le clivage en *cis* de la protéine P5 intervienne aussi au moment de l'export des ARN du CaMV.

La protéine P4 interagit également avec MOS11 ce qui suggère qu'elle est également impliquée dans l'export TREX-1-dépendant des ARN du CaMV. Les protéines de capside peuvent entrer dans le noyau grâce à leur NLS situé à l'extrémité N-terminale, il n'est donc pas exclu qu'elles le fassent de manière concomitante avec P5 et l'ADN viral au moment de l'ancrage du virion au pore nucléaire et de la dislocation partielle de la capside au début du cycle infectieux. Par contre, son implication dans l'export au cours du cycle était initialement

plus discutable car selon Karsies et al., (2002), la protéine P4 complète est incapable d'aller au noyau, son NLS n'étant exposé qu'après assemblage et clivage de P4. En d'autres termes, en se basant sur les travaux de ces auteurs, si la protéine P4 non maturée accède au noyau, c'est sans doute sous forme de cargo avec une protéine cellulaire ou virale reconnue par les importines. Par ailleurs, P4 ne pourra contribuer à l'export nucléaire que tardivement parce qu'elle n'est produite que vers la 20ème heure après le début de l'infection (Khelifa et al., 2010) à l'inverse de P5 qui est synthétisée tout le long de l'infection virale et qui peut exercer son rôle de facteur d'export de manière ininterrompue.

L'implication des protéines P5 et P4 dans l'export suppose qu'il existe deux populations de chacune de ces protéines, l'une accédant au noyau pour l'export et l'autre présente dans les viroplasmes denses où elle réplique et encapside les nouveaux génomes viraux transcrit à partir de l'ARN 35S entier. Ainsi, il est raisonnable d'imaginer que P5 (et P4), après avoir capturé avec les facteurs d'export cellulaires des ARN 35S dans le noyau, les achemine(nt) directement vers les viroplasmes denses pour la biogenèse des nouveaux virions.

Nous avons donc suivi au microscope confocal, en protoplastes d'Arabidopsis, le comportement intracellulaire de la protéine P5 fusionnée à l'eGFP. Il s'est avéré que la protéine eGFP-P5 diffuse à la fois dans le cytoplasme et dans le noyau où sa localisation se superpose avec celle de la protéine MOS11 lorsque celle-ci est co-exprimée avec eGFP-P5, sous une forme fusionnée à la mRFP. C'est la première fois, que la protéine P5 est détectée à l'intérieur du compartiment nucléaire. L'analyse à l'aide de logiciels différents, de la séquence en aminoacides de P5 montre que cette dernière renferme dans le domaine RT, un NLS bipartite riche en résidus basiques caractéristique des NLS des protéines de plantes (Hicks et Raikhel, 1993) et par conséquent, suggère que P5 est importée dans le noyau, sans l'assistance d'une autre protéine, hormis ses importines. Des expériences de mutagenèse dirigée contre ce NLS doivent être menées pour démontrer l'implication de ce signal dans l'import nucléaire de P5 et confirmer notre hypothèse.

L'investigation de la localisation intracellulaire de la protéine de fusion eGFP-P4 en protoplastes d'Arabidopsis montre que P4 forme de multiples agrégats cytoplasmiques, de tailles diverses, comme décrit dans la littérature, mais qu'elle est aussi présente dans le noyau où elle co-localise, comme P5, avec la protéine MOS11-mRFP. La co-localisation *in situ,* des protéines P4 et P5 avec le facteur MOS11, corrobore l'interaction physique observée *in vitro* entre MOS11 et les deux protéines virales. À notre connaissance, l'implication d'une transcriptase inverse dans l'export nucléaire des transcrits viraux n'a jamais été décrite ni chez les rétrovirus ni chez le HBV. Par contre, il a été montré que la protéine Core du HBV sert aussi de protéine adaptatrice pour l'exportine TAP-p15 ou encore de Gag chez le rétrovirus RSV pour l'export par CRM1(Li et al., 2010a; Jouvenet et al., 2011).

Les protéines P4 et/ou P5 sont-elles aussi impliquées dans l'export des transcrits viraux par les exportines XPO1 ? Nous ne pouvons pas l'exclure à ce stade des investigations mais il nous

semble de prime abord que la protéine virale P6 soit plus adaptée pour s'investir dans cette voie d'export. En effet, P6 possède les propriétés intrinsèques requises pour jouer le rôle d'adaptateur pour ces exportines. Elle fait la navette nucléo-cytoplasmique grâce à des NLS et un NES, riche en leucines, reconnue par XPO1 et elle interagit à la fois avec les ARN simple et double brin. P6 est exprimée précocement à partir de son propre ARNm, l'ARN 19S, ce qui lui permettrait de s'impliquer dans l'export des transcrits 35S dès le début du cycle infectieux.

L'expérience en protoplastes avec le vecteur viral muté au niveau de l'ORF VI montre bien une rétention nucléaire importante des ARN 35S, mais comme nous l'avons démontré par la suite, cette rétention est imputable à l'absence de P5. Pour savoir si la protéine P6 est engagée dans l'export, Il est envisagé de tester un vecteur pCaMV-GFP dont les ORF IV et V sont mutés afin de s'affranchir à la fois des protéines P5 et P4 (pCaMV-GFP-P5/P4_STOP) et de déterminer le taux d'export des transcrits 35S en présence de P6 originelle (ORF VI sauvage). Ces expériences, combinées à des essais de complémentation, où différentes combinaison seront testées : apporter *en trans* P5 et/ou P4 et/ou diverses variantes de P6, mutée dans ses NLS ou son NES, nous permettront certainement d'étayer nos hypothèses.

L'export nucléaire des ARNm par la voie TAP-p15, est dépendante des étapes de maturation du pré-ARNm (coiffage, épissage et polyadénylation) et du complexe TREX-1 pour la liaison des protéines adaptatrices aux ARNm et le recrutement subséquent de l'exportine (Heath et al., 2016). La structure secondaire des ARNm cellulaires ne semble pas intervenir souvent dans le recrutement de ces facteurs d'export sur les messagers cellulaires, à l'inverse des ARN viraux qui disposent à cet effet des sites de reconnaissance pour ces facteurs. Il s'agit de structures secondaires plus ou moins complexes comme les structures PRE du HBV (Sommer et Heise, 2008) et ORE du KHSV (Vogt et Bohne, 2016) ou encore les éléments CTE du MPMV (Ernst et al., 1997a) et DR du RSV (LeBlanc et al., 2007). Ces structures en *cis* constituent un ensemble composite de signaux d'encapsidation et de dimérisation et sont donc souvent multifonctionnelles. Elles sont reconnues par des protéines virales ou les protéines adaptatrices de l'hôte et servent au recrutement des exportines.

La région *leader* de l'ARN 35S du CaMV renferme une structure secondaire en tige-boucle multi-branchée (Hemmings-Mieszczak et al., 1997; Pooggin et Ryabova, 2018) qui renferme, le site d'encapsidation reconnue par le précurseur P4. Elle renferme aussi dans sa partie distale, le site de fixation de l'ARNt^{Met} initiateur et de la transcriptase inverse, le *primer binding site*. La comparaison de la séquence *leader* avec les éléments en *cis* impliqués dans l'export chez les rétrovirus et le HBV n'a révélé aucune homologie structurale, indicative d'une implication potentielle de cette région dans l'export de l'ARN 35S. Des expériences de compétition ont donc été entreprises en protoplastes, avec des fragments d'ARN de la région *leader*, le 3' UTR, celui-ci étant impliqué dans l'export chez le MLV par exemple (Pessel-Vivares

et al., 2015b) et, des fragments de la séquence codante correspondant aux ORF VII, I, II et III. Les séquences entre l'ORF VII et l'ORF II et, entre l'ORF I et l'ORF II sont éliminées de l'ARN 35S, lors de l'épissage alternatif en fonction du choix du site donneur (un site donneur dans le leader et 3, dans l'ORF I), le site accepteur de l'épissage dans l'ORF II étant unique (Kiss-László et al., 1995; Bouton et al., 2015a), de sorte que cette région pourrait contenir des éléments en *cis* permettant à la machinerie d'export de sélectionner spécifiquement l'ARN 35S complet.

Les résultats montrent que le fragment comprenant la région *leader* et l'extrémité 5' de l'ORF VII, entre en compétition avec l'ARN 35S pour la machinerie d'export, la quantité d'ARN 35S cytoplasmique représentant environ 20 % de celle trouvée en absence de compétiteur. Les autres fragments n'ont strictement aucun effet sur l'export de l'ARN 35S indiquant a *priori* que la région codante proximale (ORF VII à ORF III) ne renferme pas de signal spécifique pour l'export du transcrit 35S non épissé. La tige-boucle en soi de la région *leader* a également un effet inhibiteur sur l'export mais nettement moins prononcée que la région complète, plus de la moitié des ARN 35S pouvant gagner le cytoplasme comparé au témoin, ce qui indique que les séquences 5' et 3' flanquantes contribuent à l'efficacité de l'export nucléaire, sans doute en stabilisant la tige-boucle. En effet, la dissection de cette dernière en plusieurs fragments et leur utilisation dans les tests de compétition, révèle sans équivoque l'importance de l'intégrité de la tige-boucle et de sa structure secondaire. Des expériences de mutagenèse dirigée permettront d'identifier de manière plus précise les sous-structures voire les séquences importantes pour l'export. Il apparait donc que l'export des ARN 35S et ses versions épissées soit dictée par la tige-boucle du *leader,* celle-ci étant commune à tous ces ARN sauf l'ARN épissé D1-A qui en est dépourvu, le site donneur d'épissage D1 étant localisé dans le segment montant de la structure (Kiss-László et al., 1995). Cet ARN 35S épissé, débarrassé de toutes les séquences introniques, pourrait être exporté, comme d'ailleurs l'ARN 19S qui est dépourvu d'intron, par le même mécanisme que celui mis en œuvre par l'hôte pour exporter ses propres ARNm.

Il s'avère que les plantes infectées par le CaMV renferment également un petit ARN 8S résultant de la transcription de la région *leader*, probablement à partir de la forme non réparée, et qui, théoriquement, peut entrer en compétition avec les ARN 35S pour la machinerie d'export. Toutefois, ceci n'est pas le cas, parce que cet ARN est converti en un ARN double brin déstabilisant ainsi la tige-boucle du *leader* et fait en réalité office de leurre pour la machinerie du *RNA silencing* (Blevins et al., 2011; Hohn, 2015).

Des tests de retard sur gel réalisés en présence de la protéine ALY1 exprimée *in vitro*, avec la région *leader* et le fragment I-II, comprenant les extrémités 3' et 5' respectivement des ORF I et II, montrent que ALY1 ralentit de plus en plus la migration des deux fragments d'ARN avec le nombre molécules d'ALY associées à ces fragments. L'interaction entre la région *leader* et ALY1 corrobore nos résultats montrant l'implication de la tige-boucle dans l'export, à l'inverse de l'interaction avec le fragment I-II. L'interaction de ALY1 avec ces deux régions de l'ARN 35S s'explique par le fait que la protéine ALY1 d'Arabidopsis a une nette préférence pour les ARN

simple brin mais elle peut aussi se lier aux ARN double brin (Pfaff et al., 2018). Il serait intéressant de voir par mutagénèse dirigée si ALY1 interagit à proximité du site d'encapsidation où se fixe la protéine de capside P4 et/ou au niveau du site de liaison de P5. En effet, ces deux protéines virales interagissent avec MOS11, il est donc envisageable qu'elles recrutent le complexe TREX-1 *via* l'interaction MOS11-P4 ou P5 au niveau de leurs sites d'interaction et qu'elles permettent ainsi aux protéines adaptatrices associées à TREX-1, en l'occurrence ALY, d'interagir avec l'ARN 35S. Cette interaction impliquant P4 et P5 n'est pas conflictuelle avec le recrutement de complexes TREX-1 au niveau du *cap binding complex* (CBC) à l'extrémité 5' coiffée des transcrits 35S et des complexes EJC présents en amont de la jonction exon-exon des ARN 35S épissés. Ce recrutement basal et non optimal de quelques complexes TREX-1 sur les molécules d'ARN 35S en absence de protéines virales pourrait d'ailleurs expliquer pourquoi en protoplastes, le cycle viral démarre alors que seul un plasmide viral, pCaMV-GFP, est transfecté.

La partie 5' de l'ARN 35S, et plus particulièrement la région *leader,* semble donc être une région stratégique pour le recrutement des complexes TREX-1, grâce aux protéines P4 et P5, pour l'export des transcrits 35S du CaMV ; il n'est pas exclu que le recrutement des exportines soit confiné à cette seule région. Quant à l'ARN 19S, son export se ferait rapide, sans l'assistance de protéines virales, car son expression est indispensable pour la synthèse des autres protéines virales et la suppression du *RNA silencing*.

Des investigations supplémentaires sont évidemment indispensables pour parfaire nos connaissances sur le mécanisme d'export des transcrits du CaMV et, de comprendre comment ces transcrits viraux arrivent à se soustraire au système de surveillance de la plante.

Matériel et méthodes

Matériel et méthodes

Matériel

1. Les plantes

1.1. *Arabidopsis thaliana*

1.1.1. *Plantes sauvages*

Notre modèle d'étude, *Arabidopsis thaliana* encore connu sous le nom d'arabette des dames, fait partie de la famille des *Brassicaceae*, hôtes naturels de l'isolat Cabb B-JI du virus de la mosaïque du chou-fleur (CaMV) utilisé au laboratoire. Nous avons choisi *A. thaliana* d'écotype Columbia (Col-0) pour les tests d'infection.

1.1.2. Mutants d'insertion

Les mutants d'insertions d'ADN-T d'*A. thaliana* (Col-0 ou Col-3) utilisés dans cette étude sont listés dans le **tableau MM-1** ci-après.

Plusieurs collections de mutants d'insertion d'*A. thaliana* ont été générées à la fin des années 90 et au début des années 2000 et sont disponibles dans les banques d'Arabidopsis (Alonso et al., 2003). Ces mutants sont générés par insertion aléatoire d'un ADN de transfert (ADN-T) d'*Agrobacterium tumefaciens* modifié et de séquence connue dans le génome d'*A. thaliana* par transformation (O'Malley et al., 2015).

Les graines des mutants de l'exportine 1 (XPO1), de MOS11 et de TEX1 ont été commandées auprès de l'ABRC™ - *Arabidopsis Biological Ressources Center* – de l'Université d'État de l'Ohio aux États-Unis et les graines des mutants de ALY (*aly1 aly2*, *aly3 aly4* et *4xaly*) et du double mutant *tex1-4/mos11-2* ont été obtenues auprès du Pr. Klaus Grasser de l'Université de Regensburg en Allemagne (Pfaff et al., 2018; Sørensen et al., 2017).

Lignées mutantes	Locus	Fond	Annotations	Première utilisation
xpo1a-3 SALK_078639	AT5G17020.1	Col-0	XPO1a	Blanvillain *et al.*, 2008
xpo1b-1 SALK_088267	AT3G03110.1	Col-0	XPO1b	Blanvillain *et al.*, 2008
mos11-2 SAIL_266 E03	AT5G02770.1	Col-3	MOS11	Germain *et al.*, 2010
tex1-4 SALK_100012	AT5G56130.1	Col-0	TEX1, THO3	Jauvion *et al.*, 2010; Yelina *et al.*, 2010
aly1-1 SAIL_381_E08	AT5G59950.5	Col-3	ALY1	Pfaff *et al.*, 2018
aly2-1 wiscDsLox461-464N10	AT5G02530.1	Col-0	ALY2	Pfaff *et al.*, 2018
aly3-1 SALK_063320.55.50.x	AT1G66260.1	Col-0	ALY3	Pfaff *et al.*, 2018
aly4-2 GK-497B06-019698	AT5G37720.1	Col-0	ALY4	Pfaff et al., 2018

Tableau MM-1 | Mutants d'insertion d'ADN-T d'*A. thaliana*.

1.1.3. Le mutant *CA-rop2*

Le mutant *CA-rop2* est un mutant d'*A. thaliana* exprimant une forme constitutivement active, c'est-à-dire liée au GTP, de la *Rho-related GTPase from plants* (ROP2) (Li et al., 2001). Les protoplastes préparés à partir des plantules de ce mutant présentent la particularité d'être plus facilement transfectables.

1.2. *Brassica napus*

Des plants de colza (*Brassica napus*) d'écotype Drakkar ont été utilisés pour produire les extraits bruts de TuMV-GFP utilisés dans les tests d'infection. Les graines de ce cultivar ont été obtenues auprès du Dr Manfred Heinlein (IBMP, Strasbourg).

2. Isolats viraux

2.1. CaMV

L'isolat de CaMV utilisé au cours de notre étude est celui de Cabb B-JI du *John Innes Institute* de Norwich en Angleterre. Les plantes infectées par cet isolat viral présentent des symptômes sévères, caractéristiques de CaMV, à savoir chloroses, gaufrage des feuilles, éclaircissement des nervures et nanisme.

2.2. TuMV-GFP

Le virus de la mosaïque du navet (TuMV - *Turnip Mosaic Virus -)* utilisé dans les expériences d'infection de plantes correspond à l'isolat UKI du TuMV modifiée, dans laquelle a été introduite la séquence codante de la GFP – *green fluorescent protein* – entre les cistrons de P1 et de HCpro de la polyprotéine virale (Lellis et al., 2002; Dunoyer et al., 2004). L'expression de la GFP dans les plantes infectées par ce virus chimère permet de suivre l'évolution de l'infection sous UV.

3. Bactéries

3.1. Souches bactériennes

Les souches bactériennes utilisées dans cette étude sont listées dans le **tableau MM-2** ci-après.

Les plasmides ont été amplifiés dans *Escherichia coli* DH5α (préparation d'ADN plasmidique), *XL1-Blue* (clonage par restriction) et *NEB™5α* (clonage par Gibson Assembly™) (Bullock et al., 1987; Grant et al., 1990).

Escherichia coli BL21 (DE3) a été utilisée pour la production des protéines de fusion. Cette souche, lysogène pour le phage λ DE3, est employée pour l'expression des protéines hétérologues car sa déficience en protéase Lon et OmpT limite la dégradation des protéines qui sont exprimées. Elle renferme dans son chromosome le gène *lacI* et le gène de l'ARN polymérase du phage T7, ce dernier étant placé sous le contrôle du promoteur *lacUV5* inductible par l'IPTG (isopropyl-β-D-thiogalactopyranoside) (Studier et al., 1990).

Souches	Génotypes/caractéristiques	Origines
Escherichia coli		
DH5α	F– endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20 φ80lacZΔM15 Δ(lacZYA-argF)U169 hsdR17(rK–mK+) λ–	Life Technologies™
XL1-blue	endA1 gyrA96(nalR) thi-1 recA1 relA1 lac glnV44 F'[::Tn10 proAB+ lacIq Δ(lacZ)M15] hsdR17(rK- mK+)	Stratagene™
NEB™5α	Dérivé de DH5α optimisé par New England Biolabs. Cette souche est résistante au phage T1 et est endA-	New England Biolabs™
BL21 (DE3)	*E. coli* str. B F– ompT gal dcm lon hsdSB(rB–mB–) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB+]K-12(λS)	Novagen™

Tableau MM-2 | Souches bactériennes utilisées dans cette étude

3.2. Milieux de cultures et antibiotiques

3.1.1. Milieux de culture bactérienne

Les bactéries ont été cultivées dans le milieu nutritif *Lysogenic Browth* (LB), complémenté des antibiotiques appropriés si nécessaire et conservé à température ambiante.

Les concentrations et les modes de dissolution des antibiotiques sont indiqués dans le tableau **MM-3**. Leur stockage s'est fait à -20°C.

Les milieux solides sont préparés par addition de 15 g/L d'agar avant autoclavage et conservés à 4°C après avoir été coulés.

Milieu Lysogenic Brewth

Tryptone	1 %
Extrait de levure	0,5 %
NaCl	0,5 %

Antibiotiques	Concentration stock (mg/mL)	Concentration finale (µg/mL)	Préparé dans
Ampicilline (sel sodique)	100	100	EtOH 50 %
Kanamycine (sel disulfate)	50	50	H_2O

Tableau MM-3 | **Antibiotiques utilisés dans cette étude**

4. Plasmides

4.1. Plasmide pCaMV-GFP

Le plasmide pCaMV-GFP nous a été fourni par le Dr. Martin Drucker (INRA, Colmar). Il s'agit d'un plasmide recombinant (14 kpb) dérivant du vecteur viral pGreen-35S-B-JI dans lequel a été introduit le gène codant la protéine eGFP – *enhanced green fluorescent protein* – sous le contrôle de sa propre cassette d'expression. Le vecteur viral pGreen-35S-B-JI (environ 12 kpb) contient 1,1 copie du génome du CaMV isolat Cabb B-JI sous le contrôle du promoteur 35S ce qui permet à l'ARN polymérase II cellulaire de synthétiser directement à partir du plasmide une copie complète de l'ARN 35S et l'ARN 19S (Khelifa et al., 2010) (**figure MM-1**). Le vecteur pCaMV-GFP porte un gène de résistance à la kanamycine.

Les plasmides pCaMV-P5$_{STOP}$-GFP, pCaMV-P6$_{STOP}$-GFP et pCaMV-P5P6$_{STOP}$-GFP ont été obtenus par échange de fragments, grâce à l'utilisation d'enzyme de restriction, entre ce plasmide viral et des versions modifiées par mutagenèse dirigée du plasmide pBS-BJI (Delseny et Hull, 1983). Dans ces versions, un codon STOP a été introduit en aval de l'AUG initiateur (Met$_i$) de l'ORF V et deux codons STOP dans l'ORF VI afin d'empêcher l'expression des protéines P5 et P6.

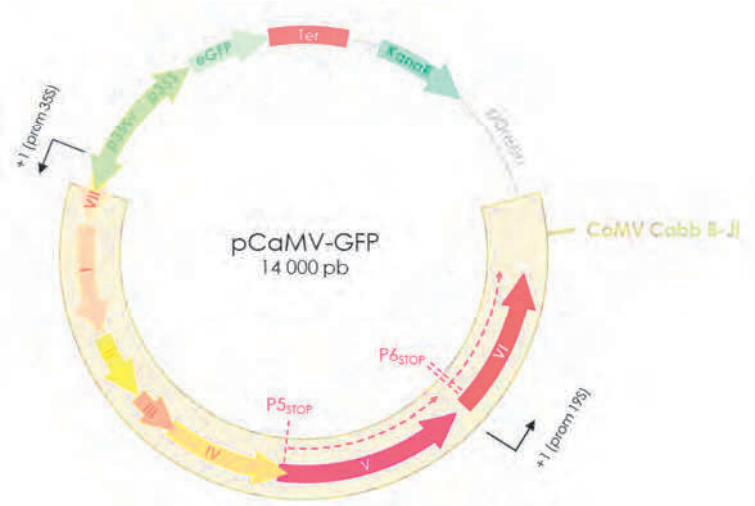


Figure MM-1 | Carte du vecteur pCaMV-GFP et ses mutants. Le vecteur viral pCaMV-GFP correspond à un plasmide pGreen modifié portant 1,1 copie du génome de l'isolat Cabb B-JI de CaMV sous le contrôle du promoteur 35S viral complet (p35Sv). Il permet l'expression en cellule végétale des ARN 35S et 19S et le déroulement du cycle viral entier. pCaMV-GFP possède également la séquence codant l'eGFP sous le contrôle d'une version tronquée du promoteur 35S (p35S). Les emplacements des codons STOP [TAG] à la place du codon [ATG] (n°17) dans l'ORF V et [TAA] à la place des codons [GAA] (n°17) et [AAA] (n°12) dans l'ORF VI sont schématiquement représentés en rouge.

4.2. Plasmides du type Litmus28i

Le plasmide Litmus28i (New Englang BioLabs™) est un vecteur d'expression de 2,8 kpb portant l'origine de réplication de pUC19 et le gène de la β-lactamase conférant une résistance à l'ampicilline. La casette de clonage est flanquée de part et d'autre du promoteur de l'ARN polymérase du bactériophage T7 permettant la transcription *in vitro* de l'ADNc inséré quelle que soit son orientation.

4.3. Plasmides du type Litmus35S-*Xba*I-tNos-35S-mRFP

Le plasmide Litmus35S-*Xba*I-tNos (3,9 kpb) a été obtenu par insertion de la cassette de clonage du plasmide pGWB14, contenant un promoteur 35S et un terminateur Nos (Nakagawa et al., 2007) dans le plasmide Litmus28i par *Avr*II et *Xba*I. Cette cassette de clonage a ensuite été modifiée pour ne garder que le site unique de restriction *Xba*I entre le promoteur 35S et le terminateur Nos (Clément Bouton, 2014, non publié) (**figure MM-2**).

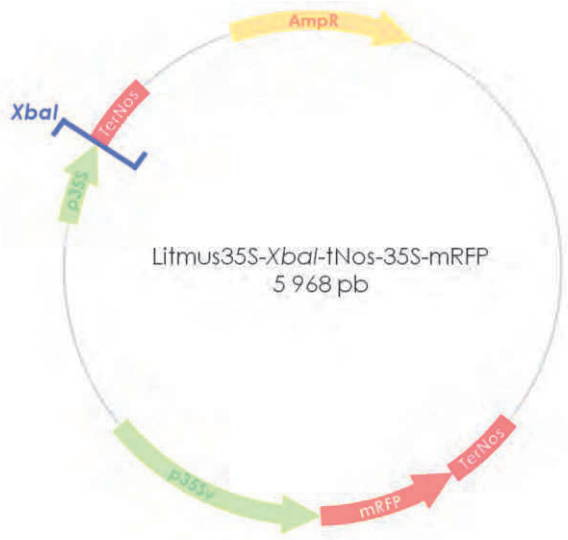


Figure MM-2 | Carte du vecteur Litmus35S-*XbaI*-tNos-35S-mRFP. Le plasmide Litmus35S-*XbaI*-tNos-35S-mRFP correspond à plasmide Litmus28i modifié dans lequel a été inséré sous contrôle du promoteur 35S complet de CaMV (p35Sv) la séquence codante de la mRFP – *monomeric Red Fluorescent Protein* - suivie du terminateur Nos (TerNos). La mRFP exprimée dans les protoplastes sert de marqueur de transfection du plasmide. Une seconde cassette d'expression a été créée de part et d'autre du site de restriction *Xba*I encadré par le promoteur 35S (p35S) et un terminateur Nos (TerNos). Le site unique *Xba*I a principalement été utilisé pour linéariser le vecteur afin de réaliser nos différents clonages par la technique du Gibson assembly™. Dans notre étude, le vecteur Litmus35S-*XbaI*-tNos-35S-mRFP sert de plateforme d'expression en protoplastes d'Arabidopsis.

Les séquences codant la mRFP sous le contrôle du promoteur 35S et du terminateur Nos, amplifiées respectivement à partir du plasmide pH7GWR et pMDC32 ont ensuite été clonées dans les sites *EcoR*I et *BamH*I du plasmide Litmus35S-*Xba*I-tNos par la technique du Gibson assembly™. Nous obtenons ainsi une plateforme d'expression sous promoteur 35S qui code également pour la mRFP, témoin de transfection. Les séquences codantes, résumées dans le **tableau MM-4** ci-après, ont ensuite été clonées dans le plasmide Litmus35S-*Xba*I-tNos-35S-mRFP (5,9 kpb) par la technique du Gibson assembly™. Concernant le plasmide Litmus35S-MOS11-mRFP, la séquence codante de MOS11 amplifiée à partir du vecteur HALO_SFI_21-H02 (ABRC) a été clonée en fusion avec la mRFP entre les sites de restriction *Sal*I et *Not*I du plasmide.

Inserts	Taille plasmides (kbp)	Séquences en nt. sur l'ARN 35S	Taille amplicon (nt) ou Poids moléculaire (kDa)	Origine
L-VII	6,7 kpb	1 à 763	765 nt	pCaMV-GFP
SL	6,4 kpb	75 à 556	482 nt	pCaMV-GFP
SL_Right	6,2 kpb	326 à 556	232 nt	pCaMV-GFP
SL_Left	6,2 kpb	75 à 325	250 nt	pCaMV-GFP
SL_Top	6,2 kpb	186 à 444	259 nt	pCaMV-GFP
SL_Bottom	6,2 kpb	75 à 185 + 445 à 556	223 nt	pCaMV-GFP
SS_3'	6,2 kpb	557 à 765	209 nt	pCaMV-GFP
L_5'	6,4 kpb	187 à 608	422 nt	pCaMV-GFP
L_3'	6,7 kpb	383 à 1106	724 nt	pCaMV-GFP
VII-I	6,6 kpb	764 à 1434	670 nt	pCaMV-GFP
I-II	6,8 kpb	1435 à 2299	866 nt	pCaMV-GFP
II-IV	6,7 kpb	2300 à 3000	701 nt	pCaMV-GFP
3'UTR	6,3 kpb	7928 à 8030 + 1 à 186	289 nt	pCaMV-GFP
P5	8 kpb	4221 à 6263	78 kDa	pCaMV-GFP
P5$_{STOP}$	8 kpb	/	/	pCaMV-P5$_{STOP}$-GFP
5myc-P5*	8 kpb	/	79 kDa	pGWB18
5myc-P5$_{STOP}$*	8 kpb	/	/	pGWB18
P5-PR	6,5 kpb	4221 à 4823	22 kDa	pCaMV-GFP
P5-RT	7 kpb	4824 à 5861	38 kDa	pCaMV-GFP
P5-RH	6,4 kpb	5862 à 6263	15 kDa	pCaMV-GFP
P5-RTRH	7,4 kpb	4824 à 6263	53 kDa	pCaMV-GFP
eGFP	6,7 kpb	/	32,7 kDa	pCK-eGFP
eGFP-P5	8,7 kpb	/	105 kDa	pCK-eGFP-P5
eGFP-P4	8,1 kpb	/	84 kDa	pCK-eGFP-P4
MOS11	6,6 kpb	/	48,2 kDa	HALO_SFI_21-H02

Tableau MM-4 | Inserts clonés dans le vecteur d'expression Litmus35S-*Xba*I-tNos-35S-mRFP

*Les plasmides Litmus35S-5Xmyc-P5-tNos-35S-mRFP et Litmus35S-5Xmyc-P5$_{STOP}$-tNos-35S-mRFP ont été obtenus par introduction non orientée dans le site *Pac*I d'un amplicon PCR généré à partir du plasmide pGWB18, contenant les répétitions myc, gracieusement fourni par le Dr Ester Lechner (IBMP).

4.4. Plasmide du type pCK-eGFP

Le plasmide pCK-eGFP (environ 4,5 kpb) est un vecteur d'expression transitoire comprenant l'ORF de l'eGFP (Clontech™) placé sous le contrôle du promoteur 35S du CaMV dont la séquence *enhancer* a été dédoublée pour stimuler la transcription du gène d'intérêt. Le gène de la β-lactamase conférant une résistance à l'ampicilline permet de sélectionner les bactéries transformées.

Afin d'obtenir les protéines de fusion eGFP-P5 et eGFP-P4, leurs ORF, amplifiés à partir du plasmide pCaMV-GFP, ont été clonés en 3'-terminal et en phase avec l'ORF de l'eGFP par restriction entre les sites *Xba*I et *Bsr*GI*.

*Les fragments PCR amplifiés à partir de pCaMV-GFP ont été digérés par les enzymes de restrictions *Avr*II et *Sma*I pour P4 (la digestion par cette dernière génère une extrémité franche), ou *Bsr*GI pour P5. Dans le cas de P4, après digestion du vecteur pCK-eGFP par *Bsr*GI, les extrémités sortantes créées par l'enzyme ont été comblées par l'ADN polymérase I (Klenow, ThermoScientific™) afin de créer des extrémités franches. Le vecteur a ensuite été digéré par l'enzyme *Xba*I générant des extrémités cohésives compatibles avec celles créées par *Avr*II.

4.5. Plasmides du type pGEX-6-P1

Ce vecteur d'expression (environ 4,9 kpb) (GE Healthcare™) permet la fusion en phase de l'extrémité 5' de la séquence codante d'intérêt avec celle de la GST – *Gluthathion-S-transférase* –. L'expression des protéines de fusion est inductible par l'IPTG grâce à la présence d'un promoteur T7 et de séquences régulatrices dérivées de l'opéron lactose. Le vecteur pGEX-6-P1 porte un gène de résistance à l'ampicilline. Pour fusionner la séquence codante de la GST à celle de la protéine d'export MOS11, la séquence codante de cette dernière a été amplifiée à partir du clone HALO_SFI_21-H02 (ABRC™) puis clonée entre les sites de restriction uniques *BamH*I et *Sal*I.

4.6. Plasmides du type pGAD-T7 et pGBK-T7

Les vecteurs d'expression pGAD-T7 (8 kpb) et pGBK-T7. (7,3 kpb) (Clontech™) ont été utilisés dans cette étude pour l'expression par transcription/traduction *in vitro* des protéines d'intérêt grâce à la présence du promoteur du bactériophage T7 présent en aval de la séquence des domaines AD et BD mais en amont de la cassette de clonage. Ces vecteurs portent respectivement les gènes de résistance à l'ampicilline ou à la kanamycine.

Les séquences codantes des protéines virales P3, P4, P5 et des différents domaines de cette dernière, amplifiés à partir de pCaMV-GFP, ont été clonées par *Sma*I et *BamH*I dans pGAD-T7 et *Sma*I et *Not*I dans pGBK-T7.

Les amorces utilisées pour l'ensemble des clonages de ces mutants, sont décrites ci-après dans le paragraphe **5. Amorces.** Toutes les constructions ont été vérifiées par séquençage.

5. Amorces

Les amorces utilisées dans cette étude ont toutes été commandées auprès de la société IDT – *Integrated DNA Technologies* –.

Amorces	Séquences (5'-3')
Amorces de clonage	
GA_Lit_RFP_FW	GGGAAATTCTTACGTTTCTCGTTCAGCTTTTTTGTAC
GA_Lit_RFP_REV	CATGGGAAGCTTCGTGGATCCGACTAGAGCCAAGCTGATCTC
GA_Lit_Nos_FW	CCGTACGTCTCGAGGAATTCAGTAACATAGATGACACCGCGC
GA_Lit_Nos_REV	GAAACGTAAGAATTTCCCCGATCGTTCAAACATTTG
GA_Lit_L-VII_FW	ATTTTGGAGAGAACACGTCTAGACTGAAATCACCAGTCTCTC
GA_Lit_L-VII_REV	TCCCTAGGGTGTTATATCTAGACGATGTTGTTCTGTAACC
GA_Lit_VII-I_FW	ATTTGGAGAGAACACGTCTAGAGAATGCGTTTACGCCAAC
GA_Lit_VII-I_REV	TCCCTAGGGTGTTATATCTAGAATTGATTCTATCATCGATTAAAGC
GA_Lit_I-II_FW	ATTTGGAGAGAACACGTCTAGAAGAGATTGCCTTCTCGGTG
GA_Lit_I-II_REV	TCCCTAGGGTGTTATATCTAGACTTTTAATTCTAGTATTTTGAGATTCTAAAAG
GA_Lit_II-IV_FW	ATTTGGAGAGAACACGTCTAGAAAAAGCAATTCAATCCTTAG
GA_Lit_II-IV_REV	TCCCTAGGGTGTTATATCTAGAGAGATCGAGTCTTCGGTAG
GA_Lit_L_5'_FW	TTCATTTGGAGAGAACACGTCTAGAGATGTGACATCTCCACTGAC
GA_Lit_L_5'_REV	CGATCCCTAGGGTGTTATATCTAGAGAATTAGAAATTTTATTGATAGAAGTATTTTAC
GA_Lit_SS_3'_REV	CGATCCCTAGGGTGTTATATCGATGTTGTTCTGTAACCTTAC

GA_Lit_ SS_3'_FW	TTCATTTGGAGAGAACACGTCTAGAAATTTAAGAAATAATCCGCATAAG
GA_Lit_L_3'_REV	CGATCCCTAGGGTGTTATATGGTTTTAGAGATATTCTTTAATTGATC
GA_Lit_L_3'_FW	TTCATTTGGAGAGAACACGTCTAGAGGTTTAAATATGATGAAGTGG
GA_Lit_SL_FW	TTCATTTGGAGAGAACACGTCTAGAGATAAGGGAAATTAGGGTTC
GA_Lit_SL_REV	CGATCCCTAGGGTGTTATATGATAAGGGGTTATTAGGGG
GA_Lit_SL_Top_FW	CGATCCCTAGGGTGTTATATCCTGCAACCTTATCCTTC
GA_Lit_SL_Top_REV	TTCATTTGGAGAGAACACGTCTAGACCTAAAACCAAAATCCAGTAC
GA_Lit_SL_Bottom_FW	AATTTCTAATTCCCTGTGCAAGGTAAGAAG
GA_Lit_SL_Bottom_REV	TTGCACAGGGAATTAGAAATTTTATTGATAGAAGTATTTTAC
GA_Lit_SL_Left_REV	CGATCCCTAGGGTGTTATATTCCTGCCTAAGCGGTACTTC
GA_Lit_SL_Right_FW	TTCATTTGGAGAGAACCGTCTAGACTAAAACCAAAATCCAGTACTAAAATC ?
GA_Lit_SL-Right_REV	CGATCCCTAGGGTGTTATATGGCTCTGATACCAATTTTTAAG
GA_Lit_SL-Right_325-556	TTCATTTGGAGAGAACACGTCTAGAGGCCGTTAGGGAAAAGATG
GA_Lit_P5_FW	TTCATTTGGAGAGAACACGTCTAGAATGATGGATCATCTACTTCTG
GA_Lit_P5 _REV	CGATCCCTAGGGTGTTATATCTAGATTAGGAATTAACCTTATTGAATTC
GA_Lit_P5_PR_FW	TTCATTTGGAGAGAACACGTCTAGAATGATGGATCATCTACTTCTG
GA_Lit_P5 _PR_REV	CGATCCCTAGGGTGTTATATCTAGATTATTCTTCTGATAACCTCCTC
GA_Lit_P5_RT_FW	TTCATTTGGAGAGAACACGTCTAGAATGAAACTCTTCATCACTCAAC
GA_Lit_P5 _RT_REV	CGATCCCTAGGGTGTTATATCTAGATTACAGCTTCTCTTCTGGTAAG
GA_Lit_P5_RH_FW	TTCATTTGGAGAGAACACGTCTAGAATGATCATCGAGACCGAC
GA_Lit_P5 _RH_REV	CGATCCCTAGGGTGTTATATCTAGATTAGGAATTAACCTTATTGAATTC
GA_Lit_P5_RTRH_FW	TTCATTTGGAGAGAACACGTCTAGAATGAAACTCTTCATCACTCAAC
GA_Lit_P5 _RTRH_REV	CGATCCCTAGGGTGTTATATCTAGATTAGGAATTAACCTTATTGAATTC
GA_Lit_P5_c-myc_FW	CGCTGTTATCATGATGGATCATCTACTTCTG
GA_Lit_P5 _c-myc_REV	GATCCATCATGATAACAGCGCTACCGTTC
Lit_SalI_MOS11_FW	ATGGCGACCAACGGAGAGAAG
Lit_MOS11_NotI_REV	AGGCAGCGCTTCCTGATACAG
pGEX_ SalI_MOS11_FW	AGTCGTCGACATGGCGACCAACGGAGAGAAG
pGEX_MOS11_NotI_REV	AGTCGCGGCCGCAGGCAGCGCTTCCTGATACAG
pGAD/BK_EcoRI_P3_FW	ATCGGAATTCATGGCTAATCTTAATCAAATCCAAAAAG
pGAD _P3_XhoI_REV	ATCGCTCGAGCTACTAAAATTGATTCGGCCATCC
pGBK_P3_SalI_REV	ATCGGTCGACCTACTAAAATTGAT TCGGCCATCC
pGAD/BK_SmaI_P4_FW	ATCGCCCGGGAATGGCCGAATCAATTTTAGAC
pGAD_P4_BamHI_REV	ATCGGGATCCTCAGTCTGAGTCTGAGTCTT
pGBK_P4_NotI_REV	ATCGGCGGCCGCTCAGTCTGAGTCTGAGTCTTCAG
pGAD/BK_SmaI_P5_FW	ATCGCCCGGGAATGATGGATCATCTACTTCTG
pGAD_P5_BamHI_REV	ATCGGGATCCTTAGGAATTAACCTTATTGAATTCTCTTGAAAGGAAG
pGBK_P5_NotI_REV	ATCGGCGGCCGCTTAGGAATTAACCTTATTGAATTC
pGAD_PR_P5_BamHI_REV	ATCGGGATCCTTATTCTTCTGATAACCTCCTCCCC
pGBK_PR_P5_NotI_REV	ATCGGCGGCCGCTTATTCTTCTGATAACC
pGAD/BK_SmaI_RT_P5_FW	ATCGCCCGGGAATGAAACTCTTCATCACTC
pGAD_RT_P5_BamHI_REV	ATCGGGATCCTTACAGCTTCTCTTCTGGTAAGG
pGBK_RT_P5_NotI_REV	ATCGGCGGCCGCTTACAGCTTCTCTTCTGGTAAG
pGAD/BK_SmaI_RNaseH_P5_FW	ATCGCCCGGGAATGATCATCGAGACCGACG
GA_pGAD_RTRH_P5_FW	GAGGCCAGTGAATTCCACCCGGGTAAACTCTTCATCACTCAACAAAG
GA_pGAD_RTRH_P5_REV	TCCCGTATCGATGCCCACCCTTAGGAATTAACCTTATTGAATTC
GA_pGBK_RTRH_P5_FW	GCCATGGAGGCCGAATTCCCGGGGAAACTCTTCATCACTCAACAAAG
GA_pGBK_RTRH_P5_REV	CTGCAGGTCGACGGATCCCCTTAGGAATTAACCTTATTGAATTC
pCKeGFP_SmaI_P4_FW	CACCCGGGATGGCCGAATCAATTTTAGAC
pCKeGFP_P4_AvrII_REV	GATCCTAGGTCAGTCTGAGTCTGAGTCTTCAGA
pCKeGFP_BsrGI_P5_FW	CATGTACAAGATGATGGATCATCTACTTCTGAAGA
pCKeGFP_P5_AvrII_REV	GATCCTAGGTTAGGAATTAACCTTATTGAATTCTCTTGAAAGG
Amorces de génotypage	
Atxpo1a-3_RP	TGGGCAGAAATCATAGGACAG
Atxpo1a-3_LP	ATCAAGGCAGGGAAACAAAAC
Atxpo1b-1_RP	CCTCAAAAAGTGTTGGAAATTTG
Atxpo1b-1_LP	TGTTCAACGGGTAATGCATTC
Atmos11-2_RP	GGCTATTTAGATCAGACCCGG
Atmos11-2_LP	CATCTGCTGAAGCCTCTTTTG
Atxpo1a-3_RP	TGGGCAGAAATCATAGGACAG
Atxpo1a-3_LP	ATCAAGGCAGGGAAACAAAAC
Atxpo1b-1_RP	CCTCAAAAAGTGTTGGAAATTTG
Atxpo1b-1_LP	TGTTCAACGGGTAATGCATTC
Atmos11-2_RP	GGCTATTTAGATCAGACCCGG
Atmos11-2_LP	CATCTGCTGAAGCCTCTTTTG
Attex1-4_RP	TACCGGAATGACTAAACGCTG
Attex1-4_LP	AACAGCTACATGTGTCCCGTC
Ataly1-1_RP	ACCCTAATCGGAAATCAACCCGATC
Ataly1-1_LP	TTTGAATTCTTAGTTTGTCTCCATATCTCCAGAATGG
Ataly2-1_RP	GTGGCTCGTCTATCGAGACTG
Ataly2-1_LP	TCCATTGCCTCTTTGTGGTAC

Ataly3-1_RP	AAAGCCTGAAAATGTGTGTGG
Ataly3-1_LP	GTCTCCCAGAAGGACCTGAAC
Ataly4-2_RP	CTCTTGCTGTGAATGCTCGTCCAT
Ataly4-2_LP	CATAACGCTCTACCTCCCCAATCTC
LBa1_TP	TGGTTCACGTAGTGGGCCATCG
LB3_TP	TAGCATCTGAATTTCATAACCAATCTCGATACAC
Gabi1_TP	GGGCTACACTGAATTGGTAGGCTC
TAG1_TP	GGACGTAACATAAGGGACTGAC
SAIL_TP	GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC
WiscDsLox_TP	AACGTCCGCAATGTGTTATTAAGTTGTC
SALK_TP	GTTGCCCGTCTCACTGGTGA
GABI_TP	ATATTGACCATCATACTCATTGC
Amorces de qPCR	
qPCR_GAPDH_FW	TTGTGACAACAGGTCAAGCA
qPCR_GAPDH_REV	AAACTTGTCGCTCAATGCAATC
qPCR_GFP_FW	GAAGCGCGATCACATGGT
qPCR_GFP_REV	CCATGCCGAGAGTGATCC
qPCR_mRFP_FW	GACGGCGAGTTCATCTAC
qPCR_mRFP_REV	GGTCTTCTTCTGCATTACG
qPCR_CaMV_P5_FW	TATAGCCCAATGGATCGTGA
qPCR_CaMV_P5_REV	CTTCTCGGCTTCATTGTTGA
qPCR_CaMV_P4_FW	CGGTTCTGGTATAATCTG
qPCR_CaMV_P4_REV	TCTTCGGTAGTTGTCATA
Amorces de PCR	
35S_(1-21)_FW	CAAGCGCGCTCTAGACTGAAATCACCAGTCTCTCTC
35S_(2280-2304)_REV	CAACTGCAGTCTAGACTTTTAATTCTAGTATTTT
35S_(4801-4821)_FW	CAAGCGCGCTCTAGAAGGGGAGGAGGTTATCAGAAG
35S_(5552-5575)_REV	CAACTGCAGTCTAGACCAAGGAAGTTTATCTTC
Compétiteur_FW	TGAACGATCCCTAGGGTGTTATAT
Compétiteur_REV	TCATTTGGAGAGAACACGTCTAGA

Tableau MM-6 | Séquences des amorces.

6. Anticorps

Les anticorps utilisés dans cette étude sont listés dans le tableau ci-après. L'incubation a été réalisée dans le tampon PBS Tween™ lait aux dilutions indiquées.

Tampon PBS Tween™ lait (de saturation)	
NaCl	140 mM
KCl	3 mM
KH_2PO_4	1,5 mM
Na_2HPO_4	8 mM
Tween™-20	0,1 %
Lait écrémé (poudre)	5 %

Anticorps	Description	Dilution	Référence/Source
Primaires			
anti-P6/HisTAV	Anticorps polyclonaux de lapin dirigés contre la protéine P6/TAV de CaMV	1 : 10 000	Dr Thomas Hohn (FMI, Bâle)
anti-CaMV	Anticorps polyclonaux de lapin dirigés contre la capside de CaMV	1 : 10 000	Dr Pierre Yot (IBMP)
anti-P2	Anticorps polyclonaux de lapin dirigés contre la protéine P2 de CaMV	1 : 10 000	Dr Stéphane Blanc (INRA, Montpellier)
anti-GFP	Anticrops polyclonaux de lapin dirigés contre l'eGFP	1 : 10 000	Dr David Gilmer (IBMP) (Link et al., 2005)
anti-H3	Anticorps polyclonaux de lapin dirigés contre l'histone H3	1 : 10 000	Dr. Alexandre Berr (IBMP)
anti-UGPase	Anticorps polyclonaux de lapin dirigés contre l'UDP-Glucose pyrophosphorylase	1 : 10 000	Pr. Anne-Marie Duchène (IBMP)
anti-c-Myc	Anticorps polyclonaux de lapin dirigés contre l'étiquette c-myc	1 : 500	c-Myc (A-14) : sc-789 Santa Cruz biotechnology™
anti-c-Myc	Anticorps monoclonaux de souris dirigés contre l'étiquette c-myc	1 : 1 000	c-Myc (9E10) : sc-40 Santa Cruz biotechnology™
Secondaires			
anti-lapin	Anticorps polyclonaux de chèvre dirigés contre les IgG de lapin	1 : 5 000	Bethyl™ A120-101AP (phosphatase alcaline) Invitrogen™ G21234 (Peroxydase)
anti-souris	Anticorps polyclonaux de chèvre dirigés contre les IgG de souris	1 : 10 000	Invitrogen™ G21040 (Peroxydase)

Tableau MM-7 | **Collection d'anticorps pour notre étude**

Méthodes

1. Infections virales

1.1. Semis de graines en terre

Avant semis, toutes les graines sont placées une nuit à -20°C afin de limiter leur infestation par les thysanoptères. Les graines d'Arabidopsis ou de Brassica sont semées en terre et mises à pousser pendant 2 semaines. Les plantules sont ensuite repiquées et cultivées pendant 1 à 2 semaines avant d'être inoculées au stade de 4 petites feuilles (12h de lumière à 22°C, 12h d'obscurité à 18°C).

1.2. Inoculation mécanique des plantes

Des plants de colza (*Brassica napus* cv. Drakkar) au stade quatre feuilles, sont inoculés mécaniquement, à l'aide d'un abrasif (célite), avec un extrait brut de feuilles de colza infectées par TuMV-GFP fourni gracieusement par le Dr Manfred Heinlein (IBMP, Strasbourg) et préparé dans du tampon potassium phosphate 0,5 M pH 7 dilué 50X (2 g de feuilles réduits en poudre fine dans de l'azote liquide et repris dans un volume final d'environ 5 mL).

Tampon potassium phosphate 0,5 M pH 7

$K_2HPO_4, 3H_2O$	1 M
KH_2PO_4	1 M

Après inoculation, les plantes sont cultivées en conditions de jour long (16h de lumière à 22°C, 8h d'obscurité à 18°C).

Des plants d'*A. thaliana* (Col-0) sauvage et mutants (*mos11-2*, *tex1-4*, *tex1-4/mos11-2*, *xpo1a-3*, *xpo1b-1*, *aly1 aly2*, *aly3 aly4*, *4xaly*) au stade quatre feuilles de la rosette sont inoculés mécaniquement, à l'aide d'un abrasif (célite), et par frottement de 3 feuilles par plante avec un extrait brut de plantes virosées préparé à partir d'Arabidopsis infectées par l'isolat de CaMV Cabb B-JI ou de feuilles de colza infectées par le virus chimère TuMV-GFP.

Après inoculation, les plantes sont cultivées dans des conditions de photopériode jour/nuit égale (12h de lumière à 22°C, 12h d'obscurité à 18°C).

1.3. Analyse de l'ADN par PCR

La présence de l'ADN génomique de CaMV dans les expériences d'infection de plantes a été déterminée par PCR selon le protocole décrit ci-après (**Méthodes 5.3.1.**).

L'ADN total est extrait à partir de disques de feuilles présentant des symptômes d'infection virale ou asymptomatiques, de 1 cm de diamètre, congelés dans l'azote liquide et conservés à -80°C. Les disques sont broyés au Precellys™ (2 x 15 sec avec 30 sec de pause) repris dans 100 µL de NaOH à 0,5 N et à nouveau broyés. 5 µL de broyat sont ensuite dilués 10 fois dans du tampon Tris-HCl 0,1 M pH 8, vortexé et congelé à - 80°C. La PCR est réalisée sur 1 µL d'échantillon.

1.4. Analyse des protéines par western blot

1.4.1. Préparation des échantillons protéiques

Les protéines sont extraites à partir de disques de feuilles présentant des symptômes d'infection virale ou asymptomatiques de 1 cm de diamètre conservés à -80°C. Les disques sont broyés au Precellys™ comme décrit auparavant, repris dans 100 µL d'urée 8 M et à nouveau broyés. Les broyats sont ensuite repris dans tampon PAGE, vortexés, chauffés 5 min à 95°C puis congelés à - 80°C.

Tampon PAGE (concentré 2 fois)

Tris-HCl 1M, pH6,8	125 mM
SDS	5 %
Glycérol	20 %
β-mercaptoéthanol	9,25 %
H_2O	qsp 20 mL
Bleu de bromophénol	quelques µg

1.4.2. Électrophorèse sur gel de SDS-polyacrylamide (SDS-PAGE)

Les protéines sont séparées par électrophorèse dans un gel SDS-PAGE composé d'un gel de concentration 5 % et d'un gel de séparation 15 %.

Gel de concentration		**Gel de séparation**	
Acrylamide/Bisacrylamide	5 %	Acrylamide/Bisacrylamide	15 %
Tris-HCl (pH6,8)	250 mM	Tris-HCl (pH 8,8)	750 mM
SDS	0,2 %	SDS	0,2 %

L'électrophorèse s'effectue dans le tampon de migration suivant.

Tampon de migration	
Tris-base	25 mM
Glycine	190 mM
SDS	0,1 %

Les échantillons protéiques sont chargés sur le gel de concentration et la migration est réalisée par application d'un courant électrique constant de 100-150 V. L'échelle de poids moléculaire PageRuler™ Prestained Protein Ladder (Thermo Scientific™) a été utilisée. Les protéines sont ensuite visualisées soit par coloration au bleu de Coomassie™ ou par analyse western blot.

1.4.3. Coloration au bleu de Coomassie™

Le gel de séparation est coloré grâce à une solution de coloration au bleu de Coomassie™ une nuit sous agitation.

Solution de coloration	
Brilliant Blue R-250	0,25 %
Éthanol	40 %
Acide acétique	10 %

Le gel est ensuite décoloré avec une solution de décoloration sous agitation plusieurs heures à l'aide de bains successifs.

Solution de décoloration	
Éthanol	15 %
Acide acétique	15 %

1.4.4. Transfert des protéines sur membrane

Les protéines séparées par SDS-PAGE sont transferées électrophorétiquement sur une membrane de nitrocellulose Immobilion™ (Millipore™). Pour cela, le gel de protéines et la membrane accolée sont pris en sandwich entre 2 couches de papier Whatman (3 mm d'épaisseur) imprégné de tampon de transfert.

Tampon de transfert

Tris base	30 mM
Glycine	230 mM
Éthanol	20 %

Le sandwich est ensuite inséré dans une cellule de transfert électrophorétique Criterion Blotter™ de BioRad™ et le transfert est effectué sous un courant électrique constant de 100 V pendant 1 h.

1.4.5. Immunodétection des protéines (western blot)

Afin de détecter les protéines transférées sur membranes de nitrocellulose, ces dernières sont incubées 1 h sous agitation dans un tampon de saturation afin de prévenir toute liaison aspécifique des anticorps sur les membranes dans les étapes suivantes.

Tampon de saturation

NaCl	140 mM
KCl	3 mM
KH_2PO_4	1,5 mM
Na_2HPO_4	8 mM
Tween™-20	0,1 %
Lait écrémé (poudre)	5 %

Les membranes sont alors mises en présence des anticorps primaires préparés dans du tampon de saturation toute une nuit à 4°C et sous agitation. Les membranes sont ensuite lavées 3 fois au moins 15 minutes dans un tampon PBS Tween™ puis incubées avec les anticorps secondaires correspondants préparés dans du tampon de saturation, 2 h sous agitation à température ambiante. Les anticorps secondaires non fixés sont éliminés par 3 lavages successifs d'au moins 15 minutes. Les dilutions auxquelles les anticorps primaires et secondaires sont utilisés sont indiquées dans le **tableau MM-7.** (**Matériel 6.**).

Lorsque les anticorps secondaires sont couplés à la peroxydase du raifort (HRP), les protéines reconnues sont visualisées grâce à la chimiluminescence renforcée (ECL). Les membranes sont exposées sous film X-ray pour autoradiographie (Fujifilm™) qui sont développés à l'aide d'un processeur automatique de film radiologique (Agfa™).

Lorsque les anticorps secondaires sont couplés à la phosphatase alcaline, les membranes sont équilibrées durant 30 minutes dans une solution de diéthanolamine à 0,1 M et à pH 9,6 puis les complexes immuns sont révélés en présence de NBT (chlorure de bleu de 4-nitrotétrazolium) 100 µg/mL, de BCIP (5-bromo-4-chloro-3-indolyl phosphate) 50 µg/mL et de $MgCl_2$ 3 mM.

2. Préparation et transfection des protoplastes et fractionnement subcellulaire

2.1. Protoplastes de mésophylle d'*A. thaliana*

Les graines d'*A. thaliana* Col-0 sauvages et mutantes sont stérilisées par un lavage de 20 minutes sous agitation dans de l'éthanol 70 %, javel 10 %. Elles sont ensuite sédimentées pendant 30 secondes à l'aide d'une centrifugeuse de paillasse puis lavées à l'éthanol absolu. Les graines sont ensuite étalées sur une boite de milieu de culture *in vitro* MS-agar après avoir été séchées sous la hotte.

Milieu MS-agar

Milieu Murashige-Skoog Référence M5519, Sigma™	4,4 g/L
Sucrose	30 g/L
MES-KOH pH 5,7	0,5 g/L
Agar	7 g/L

Les graines sont alors vernalisées pendant 24 h, au moins, à 4°C.

Les plantules d'Arabidopsis sont cultivées dans des conditions de jours longs (16 h de lumière à 21°C, 8 h d'obscurité à 17°C) pendant 7 à 12 jours. Elles sont ensuite récoltées, dilacérées finement puis mises à digérer dans 25 mL de milieu de digestion.

Milieu de digestion

MES-KOH, pH 5,7	20 mM
Mannitol	0,4 M
KCl	20 mM
Cellulase R10 Yakult Pharmaceutical™	1,5 %
Macérozyme R10 Yakult Pharmaceutical™	0,4 %
► Chauffer 10 min à 55°C	
$CaCl_2$	10 mM
Serum d'albumine bovine acétylée	0,1%

La digestion se fait sous agitation à 50 rpm pendant 4h-14h à 26°C. Les débris non-digérés sont séparés par filtration sur une couche de Miracloth (Millipore™) et les résidus d'enzymes sont éliminés par trois étapes de lavage successives lors desquelles les protoplastes sont sédimentés par centrifugation (2 min à 100 x g, dans une centrifugeuse à rotor libre et sans frein) et mis en suspension dans 20 mL de tampon de lavage, puis 20 mL de Mannitol 0,5 M et enfin 20 mL de tampon de suspension.

Tampon de lavage

MES-KOH, pH 5,7	2 mM
NaCl	154 mM
$CaCl_2$	125 mM
KCl	5 mM

Tampon de suspension

MES-KOH, pH 5,7	4 mM
Mannitol	0,4 M
$MgCl_2$	15 mM

Les protoplastes sont finalement comptés à l'aide d'une cellule de numération Fuchs-Rosenthal et mis dans du tampon de suspension à une concentration de 10^6 protoplastes par mL.

2.2. Transfection des protoplastes d'*A. thaliana*

Dans un tube de 2 mL à bout rond, 300 µL de protoplastes préparés à une concentration de 10^6 protoplastes/mL additionés de 300 µL d'un tampon de transfection préparé extemporanément sont mélangés vigoureusement à l'ADN plasmidique. Les différentes quantités d'ADN plasmidique utilisées en fonction du type d'expériences sont indiquées dans le **tableau MM-8** ci-après.

Tampon de transfection

Mannitol	0,2 M
MES-KOH, pH5.7	2 mM
CaCl2	100 µM
Polyéthylène glycol 4000 Référence 1546569, Sigma™	30 %

Plasmides	Quantités (µg)
Expériences standards	
pCaMV-GFP	10 µg
Litmus28i	15 µg
Expériences de complémentation	
pCaMV-GFP	10 µg
Litmus28i	5 µg
Litmus35S-*Xbal*-tNos-35S-mRFP	10 µg
Expériences de compétition	
pCaMV-GFP	10 µg
Litmus35S-*Xbal*-tNos-35S-mRFP	40 µg
Expériences de localisation intracellulaire	
pCaMV-GFP	10 µg
Litmus35S-*Xbal*-tNos-35S-mRFP	10 µg
pCK-eGFP	10 µg

Tableau MM-8 | Quantités d'ADN plasmidique en fonction du type d'expériences.

Les protoplastes sont incubés 15 min à température ambiante et la transfection est stoppée par l'ajout d'un mL de tampon de lavage. Le mélange est homogénéisé précautionneusement par inversion du tube puis les protoplastes sont sédimentés par centrifugation (5 min à 100 x g, sans frein). Les cellules sont ensuite resuspendues dans 1 mL de tampon d'incubation et incubées dans une plaque 12 puits pendant 20 h à 23°C et à l'obscurité.

Tampon d'incubation

Mannitol	0,2 M
MES-KOH, pH5,7	4 mM
KCl	20 mM

Les protoplastes sont observés au microscope Axio Zoom (Zeiss™) au grossissement X40 afin d'estimer, au travers de la fluorescence de l'eGFP et de la mRFP, l'efficacité de transfection. Les protéines eGFP et mRFP sont visualisées respectivement aux longueurs d'onde de 488nm et 561 nm et les cellules par excitation de la chlorophylle a à 650 nm.

Dans le cas d'un traitement à leptomycine B (Sigma™), celle-ci est ajoutée 20 h après transfection à une concentration finale de 100 nM pendant 1 h.

2.3. Fractionnement subcellulaire

Les protoplastes sont récoltés dans un micro-tube de 1,5 mL, sédimentés (5 min à 100 x g, sans frein) et le surnageant, correspondant au milieu d'incubation est retiré par pipetage. Cent µL de tampon hypotonique sont ajoutés, les protoplastes sont resuspendus délicatement par aspiration refoulement puis incubés 5 minutes à température ambiante.

Tampon hypotonique

Tris-HCl, pH 7,5	10 mM
NaCl	10 mM
NP-40	0,3 %

Le mélange est ensuite centrifugé 10 min à 5000 rpm et 90 µL de surnageant sont récupérés puis placés dans un nouveau tube, c'est la fraction cytoplasmique. Le reste du surnageant est retiré et le culot correspond à la fraction nucléaire.

Pour les fractions totales, une fois sédimentés, les protoplastes sont récupérés dans 20 µL de tampon hypotonique et délicatement resuspendus.

2.4. Échantillons protéiques de protoplastes

Les fractions cytoplasmiques, nucléaires et totales sont aussitôt additionnées de tampon PAGE (**Méthodes 1.4.1.**, qsp 30 µL), chauffées à 95°C 5 min puis congelées à -80°C en attendant d'être analysées.

2.5. Extraction des ARN

Immédiatement après le fractionnement, les fractions sont reprises dans 300 µL de TRI-Reagent (Sigma™) pour les fractions cytoplasmiques et 200 µL pour les fractions nucléaires et totales, auxquelles sont ajoutés 100 µL d'eau stérile dépourvue de nucléases et vortexés. Les ARN sont ensuite transférés dans des tubes Phasemaker™ (Invitrogen™) extraits selon les instructions du fabricant, puis solubilisés dans de l'eau stérile garantie sans nucléases avant d'être stockés à -80°C.

3. Quantification des ARN 35S par RT-qRT-PCR

Les ARN cytoplasmiques, nucléaires et totaux ainsi isolés sont par la suite traités à la DNase et validés par PCR (**Méthodes 5.3.1.**) afin de vérifier que tout ADN plasmidique contaminant a été éliminé. Les ARN sont ensuite rétrotranscrits pour synthétiser les ADN complémentaires. Les quantités relatives d'ARN 35S des différentes fractions sont ensuite mesurées par qRT-PCR.

3.1. DNase

Le kit TURBO DNA-free™ d'Ambion™ a été employé pour réaliser les traitements DNase en suivant les instructions du fabricant. À chaque expérience, les quantités d'ARN cytoplasmiques, nucléaires et totaux sont normalisées de manière à utiliser la même quantité pour un même type de fractions.

Mix réactionnel

ARN	jusqu'à 1 µg
TURBO DNase (2 unités)	1 µL
Tampon TURBO DNase 10X	2 µL
H_2O stérile (dépourvue de nucléases)	qsp 20 µL
▶ 30 minutes à 37°C	
Réactif d'inactivation	2 µL
▶ 5 minutes à température ambiante	

Les échantillons sont centrifugés (5 min à 6 000 rpm) puis 18 µL de surnageant sont récupérés et stockés à – 80°C.

3.2. Rétrotranscription

Les ADNc sont synthétisés par l'ADN polymérase ARN dépendante SuperScript IV™ (Invitrogen™) en présence d'amorces de type oligod$(T)_{20}$ et d'hexanucléotides aléatoires selon les instructions fournies par le fabricant. Un exemple de mix réactionnel et de conditions standards, vous sont indiqués ci-après. Les échantillons sont stockés à -80°C.

Mix réactionnel

oligod$(T)_{20}$	2,5 µM
Random hexamers	2,5 µM
dNTP*	0,5 mM
H_2O stérile	qsp 13 µL

* Le mix de dNTP contient 10 mM de dATP, dCTP, dGTP et dTTP.

ARN	8 µL (jusqu'à 1 µg)
▶ 5 min à 65°C puis 1 min sur glace	
DTT	5 mM
RiboLock™ (Thermo Scientific™)	40 U
Tampon SSIV™ 5X	4 µL
SuperScript IV™	200 U
▶ 10 min à 28°C, 10 min à 53°C puis 10 min à 80°C	

3.3. Quantification par PCR en temps réel

La PCR quantitative en temps réel est effectuée dans le thermocycleur LightCycler™ 480 (Roche™), avec un mélange enzymatique contenant l'ADN polymerase et le SYBR Green I (LightCycler™ 480 SYBR Green I Master, Roche™), sur des ADNc dilués 20 fois. Les différentes amorces utilisées sont indiquées dans le **tableau MM-6**. Les amorces (500 nM final) ciblant l'ARN 35S s'apparient soit dans la région de l'ORF V (**qPCR_CaMV_P5_FW et REV, tableau MM-6**) ou de l'ORF IV (**qPCR_CaMV_P4_FW et REV, tableau MM-6**) de manière à détecter l'ensemble des

isoformes (sans distinction des formes épissés et non-épissé) de l'ARN 35S. Les messagers codant la GAPDH (AT3G041204), l'eGFP (codé par pCaMV-GFP) et la mRFP (codé par Litmus35S-*XbaI*-tNos-35S-mRFP) ont servi comme références. La température d'hybridation est de 60°C.

3.4. Détection des ARN 35S compétiteurs par RT-PCR

L'expression des fragments d'ARN 35S dans nos expériences de compétition en protoplastes a été vérifiée par RT-PCR grâce à des amorces (**compétiteur_FW et REV**, **tableau MM-6**) s'appariant au niveau des extrémités 5' et 3' communes à l'ensemble de ces ARN, selon le protocole décrit ci-dessous (**Méthodes 5.3.1.**). Les amplicons obtenus sont ensuite séparés sur gel d'agarose (**Méthodes 5.4.).**

4. Analyses des interactions entre protéines

4.1. Analyse des interactions protéiques *in vitro*

4.1.1. Expression et marquage des protéines à la méthionine ^{35}S

Les protéines virales P3, P4, P5 et les différents domaines de cette dernière ont été synthétisés *in vitro* grâce au système de transcription couplée à la traduction dans des lysats de réticulocytes de lapin (TNT™ T7 Coupled Reticulocyte Lysate System) de Promega™, en suivant les instructions du fabricant. Les ARNm codant les protéines d'intérêt sont synthétisés par l'ARN polymérase du phage T7 à partir des plasmides pGBK-T7 ou pGAD-T7 (**Matériel 4.6**). Ces messagers sont ensuite traduits dans les réticulocytes de lapin en présence de 10 µCi de méthionine ^{35}S, dans le mélange réactionnel type suivant.

Mix réactionnel

Lysat de réticulocytes de lapin TNT™	25µL
Tampon TNT™	2 µL
ARN polymerase T7 TNT™	1 µL
Mélange d'acides aminés – Méth.	0,02 mM
Méthionine ^{35}S	10 µCi
RNasin™	40 U
ADN plasmidique	1 µg
H_2O stérile dépourvue de nucléases	qsp 50 µL

▶ 1h30 à 30°C
Les échantillons sont stockés à -20°C.

4.1.2. Expression de la protéine de fusion GST-MOS11 et de GST

Les vecteurs codant la GST et la protéine de fusion GST-MOS11 sont exprimés dans la souche *BL21 (DE3)* d'*E. coli*. Les bactéries transformées (**Méthodes 5.1.1.**) sont mises en préculture de nuit à 37°C et 250 rpm dans 50 mL de milieu LB additionné d'ampicilline (100 µg/mL final). 500 mL de milieu LB additionné du même antibiotique sont ensemencés par 10 mL

de préculture (dilution au 50e), puis mis sous agitation à 37°C et 250 rpm 1,5 à 2 h jusqu'à ce que la culture bactérienne atteigne sa phase exponentielle de croissance (absorbance à 600 nm de 0,6 environ). L'expression des protéines recombinantes est induite par addition d'IPTG (isopropyl-β-D-thiogalactopyranoside) à une concentration finale de 0,2 mM et les bactéries sont cultivées pendant 4 h à 20°C et 250 rpm. Les culots bactériens, obtenus après centrifugation des cultures à 7 000 x g 15 min à 4°C, sont congelés à -20°C.

4.1.3. Capture des protéines en « batch »

La technique du « batch », réalisée dans des tubes de 50 mL, permet de capturer, par des étapes de centrifugations et lavages de billes couplées à un ligand, des protéines fusionnées à une étiquette reconnue par ce ligand. La protéine de fusion GST-MOS11 ainsi que la GST seule, ont été capturées par des billes de sépharose 4B couplées au glutathion et vendues par GE Healthcare™. Toutes les étapes ont été réalisées à 4°C.

Les culots bactériens sont repris dans 30 mL de tampon NET.

Tampon NET

NaCl	100 mM
EDTA	0,1 mM
Tris-HCl, pH 10,5	50 mM
Nonidet™ P-40	0,5 %
DTT	2 mM
Inhibiteurs de protéases Roche™	1 tablette/50 mL

Les bactéries sont lysées par une sonication douce, par pulses, durant 5 cycles de 2 min (2 sec pulse, 2 sec pause, amplitude 40 %) dans la glace, puis centrifugées à 17 000 x g pendant 30 min, afin de sédimenter les protéines insolubles. Le surnageant contenant les protéines de fusion solubles est passé à travers un filtre de 0,45 µM de diamètre de pore, pour enlever les éventuels agrégats. Le filtrat est ensuite incubé en présence de 1 mL de billes de glutathion sépharose 4B (à 50 %, lavées et reprises dans du tampon NET à pH 7,5) toute une nuit sous agitation à 4°C. Les billes sont enfin sédimentées (1000 x g, 5 min, 4°C) et lavées 3 fois avec 500 µL de tampon NET (pH 7,5).

4.1.4. Analyses par GST *pull-down*

Dans un micro-tube de 1,5 mL 50 µL de billes de gluthation sépharose ayant fixé les protéines de fusion GST sont mis en présence de 20 µL du mélange protéique radioactif pendant 2 h à température ambiante et sous agitation dans un volume final de 250 µL. Le surnageant est récupéré et les billes sont lavées 3 fois comme précédemment. Du tampon PAGE est ajouté au surnageant (fraction non-liée) et au culot de billes (fraction liée) qui sont tous les deux congelés à -20°C avant analyse par électrophorèse en gel SDS-PAGE et révélation autoradiographique (**Méthodes 1.4.2.**).

4.2. Analyse de la localisation intracellulaire des protéines de fusion à l'eGFP ou la mRFP par microscopie confocale

La localisation intracellulaire des protéines fusionnées à l'eGFP ou la mRFP est analysée, après transfection des protoplastes d'Arabidopsis, par observation au microscope confocal à balayage laser ZEISS LSM 780.

Les cellules, entre lame polylysinée et lamelle, sont observées à l'aide d'un objectif 63X à immersion à huile. Un intercalaire en silicone est positionné entre lame et lamelle pour éviter l'écrasement des protoplastes.

L'excitation de l'eGFP s'effectue à 488 nm à l'aide d'un laser argon, dont le trajet optique est dévié à partir de la source par un miroir dichroïque, les longueurs d'ondes inférieures ou égales à 488 nm. L'excitation de la mRFP est quant à elle réalisée à 561 nm avec un laser DPSS 561-10.

La fluorescence émise par l'eGFP excitée (respectivement par la mRFP), est mesurée par un filtre d'émission calibré pour sélectionner les longueurs d'ondes comprises entre 490 et 579 nm (582 et 754 nm pour la mRFP) puis recueillie par un photomultiplicateur.

Des images en contraste interférentiel différentiel (DIC) sont obtenues simultanément grâce à l'activation d'un second photomultiplicateur.

5. Techniques de biologie moléculaire

5.1. Transformation des bactéries compétentes

5.1.1. Transformation des bactéries chimio-compétentes

50 ng d'ADN plasmidique ou la totalité du mix réactionnel de ligation est mélangé à 40 µL de bactéries chimio-compétentes, le mélange est ensuite refroidie 30 min sur glace. Afin de faciliter l'entrée de l'ADN dans les bactéries compétentes les cellules sont soumises à un choc thermique 1 min à 42°C puis directement remises sur glace 5 min. 500 µL de milieu sont rajoutés et le mélange réactionnel est incubé 30 min à 1 h à 37°C pour permettre la régénération de la paroi bactérienne. La totalité des bactéries est ensuite étalée sur milieu LB-agar complémenté du ou des antibiotique(s) approprié(s). Les boites sont incubées toute une nuit à 37°C.

5.1.2. Transformation des bactéries électro-compétentes

Deux µL du mix réactionnel de ligation est mélangé à 40 µL de bactéries électro-compétentes. Le mélange est transféré dans une cuve de 0,2 cm de section et soumis à une décharge électrique de 280 V/cm, sous une capacitance de 25 µF et une résistance de 200 Ohm. Après addition de 0,5 mL de milieu LB, les bactéries sont incubées pendant 30 min à 37°C pour permettre la régénération de leur paroi. Elles sont ensuite étalées à raison de 200 µL sur

milieu LB gélosé complémenté avec le ou les antibiotique(s) approprié(s) afin de sélectionner les bactéries transformées, et incubées une nuit à 37°C.

5.2. Purification de l'ADN plasmidique produit dans *E. coli*

Les plasmides utilisés dans cette étude sont listés dans la partie **Matériel 4. Plasmides**. Les ADN plasmidiques ayant servi de matrices pour les PCR, les transformations bactériennes ou les transfections des protoplastes ont été purifiés grâce aux kits de purification sur colonnes et de centrifugation NucleoSpin™ Plasmid EasyPure et NucleoBond™ Xtra Midi (Macherey-Nagel™) en suivant les instructions du fabricant. Les ADN plasmidiques une fois purifiés sont conservés à -20°C dans de l'eau.

5.3. Amplification de séquences d'ADN par PCR

5.3.1. Amplification de séquences d'ADN par PCR de routine

L'ADN polymérase GoTaq™ Flexi de Promega™ a été utilisée pour les réactions de PCR de détection standard avec des amorces ciblant la région 5' de l'ARN 35S (**35S_(1-21)_FW** et **35S_(2280-2304)_REV**) ou l'ORF V (**35S_(4801-4821)_FW** et **35S-(5552-5575)_REV**, **tableau MM-6**). Tous les produits de PCR sont analysés par électrophorèse sur gel d'agarose 1 % (**Méthodes 5.4.**). Un exemple de mix réactionnel et de conditions de PCR standard est indiqué ci-après.

Mix réactionnel

Matrice ADN	50-100 ng
Tampon GoTaq™ Flexi 5X	10 µL
$MgCl_2$	2 mM
dNTP*	0,2 mM
amorce FW	0,5 µM
amorce REV	0,5 µM
H_2O stérile	qsp 50 µL
ADN polymérase GoTaq™ Flexi	0,25 µL

* Le mix de dNTP contient 10 mM de dATP, dCTP, dGTP et dTTP.

Conditions

1.	Dénaturation initiale 95°C	2 min		
2.	Dénaturation	95°C	30 sec	⎫ répéter 30X
3.	Hybridation	45-60°C*	30 sec	⎪
4.	Polymérisation	72°C	1kb/min**	⎪
5.	Extension finale	72°C	5 min	⎭
6.	Pause	10°C	∞	

* La température d'hybridation des amorces dépend du Tm de chaque amorce utilisée et correspond généralement au Tm-5°C

** La taille de l'amplicon définit le temps de polymérisation

5.3.2. Amplification de séquences d'ADN par PCR dans le cadre du clonage

L'ADN polymérase Phusion™ High-Fidelity de Thermo Scientific™ a été utilisée pour la synthèse des amplicons destinés au clonage. La liste des amorces utilisées à cet effet est présentée dans le **tableau MM-6**. Les produits de PCR sont purifiés sur gel d'agarose après électrophorèse (**Méthodes 5.4.**). Un exemple de mix réactionnel et de conditions de PCR type est indiqué ci-après.

Mix réactionnel

Matrice ADN	50-100 ng
2X Phusion™ Master Mix with HF Buffer	10 µL
amorce FW	0,3 µM
amorce REV	0,3 µM
H_2O stérile	qsp 20 µL

Conditions

7.	Dénaturation initiale	98°C	1 min	
8.	Dénaturation	98°C	30 sec	⎫
9.	Hybridation	50-72°C*	30 sec	⎬ répéter 30X
10.	Polymérisation	72°C	2 kb/min**	⎭
11.	Extension finale	72°C	10 min	
12.	Pause	10°C	∞	

* La température d'hybridation des amorces dépend du Tm de chaque amorce utilisée, elle est calculée grâce au Tm Calculator de Thermo Scientific™ disponible en ligne.

** La taille de l'amplicon définit le temps de polymérisation

5.4. Électrophorèse en gel d'agarose

Les échantillons ADN sont analysés par électrophorèse en gel d'agarose préparé dans du tampon TBE. Il contient du bromure d'éthidium à une concentration finale de 0,5 µg/mL. Le gel est ensuite coulé dans un support en plastique transparent aux UV placé dans un moule BioRad Mini-Sub Cell GT™. Après solidification le gel est submergé dans du tampon TBE.

Tampon TBE

Tris	100 mM
Acide borique	90 mM
EDTA	2,5 mM

Les échantillons ADN sont préparés par ajout de tampon de charge (6X Blue/Orange Loading Dye de Promega™), excepté pour les échantillons déjà contenus dans du tampon vert GoTaq™ Flexi 5X, Phire Plant Direct PCR Master Mix 2X ou FastDigest™ Green buffer 10X.

La migration des échantillons est réalisée par application d'un voltage constant de 80 à 120 V pendant 30 à 60 min. Les échelles de tailles 1 kb DNA Ladder de Promega™ et O'GeneRuler Express™ DNA Ladder de Thermo Scientific™ ont été utilisées.

5.5. Extraction sur gel et purification de l'ADN

Les fragments d'ADN et plasmides séparés sur gel d'agarose peuvent être purifiés en vue de leur clonage. Pour cela la bande d'agarose contenant l'ADN est excisée du gel à l'aide d'un scalpel et transféré dans un micro-tube de 1,5 mL. L'ADN est ensuite extrait de la matrice

grâce à l'utilisation du kit NucleoSpin™ Gel and PCR Clean-up (Macherey-Nagel™) en suivant les instructions du fabricant.

5.6. Les techniques de clonage

5.6.1. Clonage par restriction

L'ADN plasmidique est digéré par une ou deux endonucléases de restriction choisies. Deux enzymes sont utilisées simultanément si elles possèdent une activité enzymatique maximale dans le même tampon. Dans le cas contraire, les digestions sont réalisées de manière successive après purification (**Méthodes 5.2.**) et vérification sur gel d'agarose (**Méthodes 5.4.**). Classiquement la digestion est réalisée dans un volume final de 50 µL sur 2-3 µg d'ADN plasmidique (ou fragments ADN amplifié par PCR) pendant 2 à 4 h à 37°C (sauf pour *SmaI*, à 25°C) en utilisant 20 U d'enzymes en suivant les instructions du fabricant (Promega™).

En fin de digestion, les extrémités 5' du vecteur linéarisé sont déphosphorylées en présence de 2 U de phosphatase alcaline d'intestin de veau, pour éviter sa recircularisation lors de la réaction de ligation.

L'ADN digéré est ensuite purifié grâce au kit NucleoSpin™ Gel and PCR Clean-up de Macherey-Nagel™ et repris dans de l'eau.

L'ADN plasmidique linéarisé et les fragments PCR sont mélangés avec l'ADN ligase du phage T4 (Promega™) en respectant un ratio molaire de 5 pour 1. Un exemple de mix réactionnel classique est donné ci-dessous.

Mix réactionnel

Fragment ADN	5 : 1
ADN plasmidique	1 : 5
Tampon de ligation X10	1 µL
ADN ligase T4	1 µL
H_2O stérile	qsp 10 µL

Le milieu de ligation est incubé 1 h à température ambiante avant transformation dans les bactéries compétentes *E. coli XL1 blue* ou *DH5α* (**Matériels 3.1.**). Les transformants sont sélectionnés sur milieu LB + antibiotique et l'ADN plasmidique purifié des colonies résultantes est vérifié soit par digestion avec les enzymes de restrictions appropriés ou par PCR sur colonie. Les produits de digestion ou de PCR sont soumis par électrophorèse sur gel d'agarose avec les contrôles appropriés.

5.6.2. Clonage par la technique du Gibson assembly™

Dans la technique du Gibson assembly™, les amplicons de PCR sont générés de manière à porter des extrémités 5' et 3' chevauchantes et complémentaires à la fois entre elles et avec les extrémités du plasmide receveur linéarisé.

Cette technique, développée par le Dr. Daniel Gibson, permet l'assemblage de plusieurs fragments d'ADN grâce à l'action de trois enzymes successives (Gibson et al., 2009, 2010). Une exonucléase qui crée des extrémités 3' sortantes facilitant l'hybridation des fragments complémentaires, une ADN polymérase qui permet de combler les régions restées en simple-brin entre les fragments appariés, et une ADN ligase permettant la formation des liaisons phospho-diester entre les fragments assemblés.

Dans cette étude nous avons utilisé le kit NEBuilder™ HiFi DNA assembly, comprenant également les bactéries chimio-compétentes *E. coli* NEB5α, (New England BioLabs™)

Un exemple de mélange réactionnel et de transformation bactérienne est donné ci-après.

Mix réactionnel

Vecteur linéarisé	0,02 pmoles
Fragment(s)	0,5 pmoles
NEBuilder™ HiFi DNA Assembly Master Mix	10 µL
H_2O sterile	qsp 20 µL

► 30 min à 50°C

Les échantillons sont stockés à -20°C.

Transformation

Mix réactionnel	2 µL
E. coli NEB5α	50 µL

► 30 min sur glace

► 1 min à 42°C

► 2 min sur glace

+ 300 µL de milieu SOC

► 30 min à 37°C

Etaler 100 µL sur milieu LB-agar + antibiotique

Les boites sont incubées à 37°C toute une nuit.

Milieu de culture SOC

Peptone végétale	2 %
Extrait de levure	0,5 %
NaCl	10 mM
KCl	2,5 mM
$MgCl_2$	10 mM
$MgSO_4$	10 mM
Glucose	20 mM

5.7. Génotypage

Le génotypage des différents mutants d'insertion d'*A. thaliana* a été réalisé à l'aide du kit Phire Plant Direct PCR Master Mix vendu par Thermo Scientific™.

Brièvement 1-2 mm de feuilles d'Arabidopsis âgées de 7 à 15 jours après repiquage sont prélevés à l'aide d'un micro-ciseau de Noyes et écrasés à l'aide d'un cône de P200 dans du tampon de dilution.

La PCR est réalisée à l'aide de 3 amorces différentes. Deux amorces sont spécifiques du gène d'intérêt. Elles s'hybrident de part et d'autre (LP – *Left Primer* – et RP – *Right Primer* –) de la zone d'insertion de l'ADN-T d'*Agrobactérium tumefaciens* et servent à amplifier l'allèle sauvage du gène d'intérêt. La troisième amorce (TP – *T-DNA primer* –) est spécifique de l'ADN-T d'*A.*

tumefaciens utilisé (Lba1 pour les lignées SALK (Alonso et al., 2003), LB3 pour les lignées SAIL (Sessions et al., 2002), Gabi1 pour les lignées GABI-Kat (Rosso et al., 2003), TAG1 pour les lignées INRA-Versailles (Samson et al., 2002) et WiscDsLox pour les lignées WiscDsLox (Woody et al., 2007). Elle sert à amplifier, en combinaison avec l'une ou l'autre des amorces LP ou RP, en fonction du sens d'insertion de l'ADN-T, l'allèle mutant du gène d'intérêt. La liste des amorces utilisées pour les réactions de génotypage est donnée dans le **tableau MM-6**.

Les amorces sont conçues afin d'être distantes de 1000 pb et la réaction de polymérisation est paramétrée pour permettre la synthèse d'un amplicon d'au maximum 1000 pb. Ainsi, seul un allèle sauvage permet la synthèse d'un fragment PCR avec le couple d'amorces LP et RP, et un allèle muté la synthèse d'un fragment PCR avec un couple d'amorces LP ou RP et de l'amorce spécifique de l'ADN-T, TP. L'électrophorèse en gel d'agarose (**Méthodes 5.4.**) permet en fonction du profil électrophorétique obtenu, de déterminer si la plante génotypée est homozygote avec deux allèles sauvages, homozygote avec deux allèles mutés ou hétérozygote (**figure MM-3.**).

Un exemple de mix réactionnel et de conditions de PCR type vous sont indiqués ci-après.

Mix réactionnel

Broyat	0,5 µL
2X Phire Plant Direct PCR Master Mix	10 µL
amore FW	0,5 µM
amorce REV	0,5 µM
H_2O stérile	qsp 20 µL

Conditions

1.	Dénaturation initiale	98°C	5 min	
2.	Dénaturation	98°C	10 sec	
3.	Hybridation	60°C	10 sec	répéter 25X
4.	Polymérisation	72°C	20 sec	
5.	Extension finale	72°C	1 min	
6.	Pause	10°C	∞	

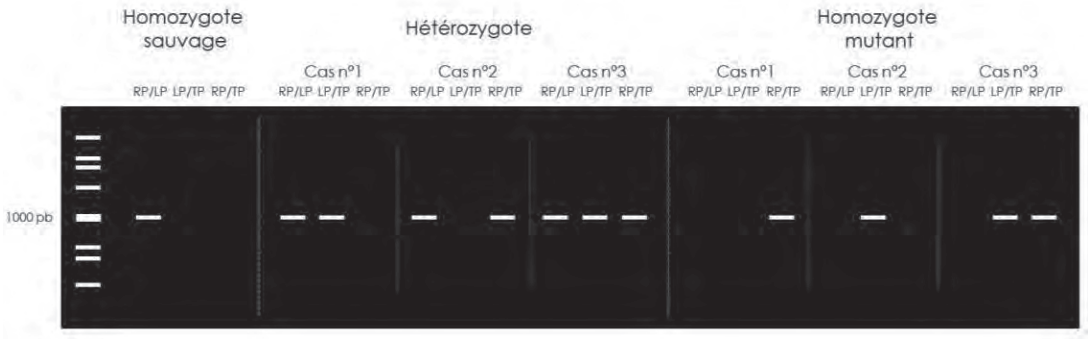
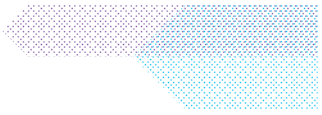


Figure MM-3 | Différents profils théoriques pouvant être obtenus après électrophorèse sur gel d'agarose lors d'un génotypage. Les amplicons résultent de réactions de PCR au niveau de la région d'insertion de l'ADN-T dans le gène à muter. Les amorces RP, LP et TP situées, respectivement, de part et d'autres de la région d'insertion et dans l'ADN-T sont utilisées afin d'établir le profil génétique de chaque plante. Les produits d'amplification sont séparés sur gel d'agarose 1 % et révélés sous UV grâce au bromure d'éthidium. Lorsqu'un amplicon est synthétisé avec les amorces RP et LP, la plante possède au moins un allèle sauvage puisque la présence du T-DNA rendrait les 2 amorces trop distantes pour permettre la synthèse d'ADN double-brin. A contrario, la présence d'un amplicon lorsque l'amorce TP est utilisée indique l'insertion de l'ADN-T dans un des allèles du gène d'intérêt.

Annexes



Annexes

Annexe 1

Vous trouverez ci-dessous l'article rédigé en collaboration avec l'équipe du Dr Laurence Maréchal-Drouard, avec laquelle nous avons partagé notre savoir-faire et notre protocole pour la transfection de protoplastes.

Cet article fait partie d'un manuscrit consacré à l'étiquetage des ARN qui sera soumis à la rentrée, à Springer par son éditeur le Dr Manfred Heinlein.

Ce manuscrit a pour vocation de partager avec un public novice des techniques spécifiques sous forme de protocoles ; en utilisant des termes simples, permettant à tout un chacun de comprendre la méthode sans connaissances préalables. Il se distingue en particulier, par le partage avec le futur utilisateur d'un grand nombre de conseils pratiques (sous forme de notes) qui lui permettront de maitriser plus facilement et rapidement la méthode.

Transfection of small non-coding RNAs into *Arabidopsis thaliana* protoplasts

Stéphanie Lalande, Marjorie Chery, Elodie Ubrig, Guillaume Hummel, **Julie Kubina**, Angèle Geldreich and Laurence Drouard.

In Methods in Molecular Biology : RNA tagging, M. Heinlein (ed), in press.

Transfection of small non-coding RNAs into *Arabidopsis thaliana* protoplasts

Stéphanie Lalande, Marjorie Chery, Elodie Ubrig, Guillaume Hummel, Julie Kubina, Angèle Geldreich, and Laurence Drouard

Institut de biologie moléculaire des plantes-CNRS, Université de Strasbourg, 12 rue du Général Zimmer, F-67084 Strasbourg, France

Running head: Plant protoplasts-based system to transfect RNA

Abstract
Polyethylene glycol transfection of plant protoplasts represents an efficient method to incorporate foreign DNA and study transient gene expression. Here, we describe an optimized protocol to deliver small non-coding RNAs into *Arabidopsis thaliana* protoplasts. An example of application is provided by demonstrating the incorporation of a 20 nt long small non-coding RNA deriving from the 5' extremity of an *A. thaliana* cytosolic alanine tRNA into freshly isolated protoplasts.

Key words Transfection, small non-coding RNAs, tRNA-derived fragments, protoplasts, northern blot

1. Introduction

Among the various classes of small non-coding RNAs (sncRNAs), tRNA-derived fragments (tRFs) are now recognized as important regulators of numerous biological processes [1–4]. In plants, tRFs represent a large and dynamic repertoire of small RNA fragments [5] but their molecular functions are still poorly understood [2]. Among the different ways to decipher the molecular processes in which tRFs are involved, the use of *in vivo* approaches is rather limited. Indeed, the production of specific tRFs in transgenic plants remains difficult to achieve. Facing this bottleneck, cell transfection represents an interesting option. This technique, mainly developed in the animal field to study the biology of miRNAs or siRNAs, is based on the use of lipofectamine, a reagent that facilitates efficient delivery of small RNA molecules into cells (see for instance [6,7]). In plants, to deliver nucleic acids, researchers have chosen to work with plant protoplasts (*i.e.* cells without their cell wall). Nevertheless, chemicals such as lipofectamine appeared toxic and inefficient [8] and thus are rarely utilized. Rather, three other types of methods have been developed to transfect macromolecules: electroporation (*e.g.* [9]), microinjection (*e.g.* [10]) and polyethylene glycol (PEG)-mediated transfection (*e.g.* [11], the latter method being the most widely used). Such protocols are often employed to deliver DNA into protoplasts for transient gene expression studies, but the incorporation of RNAs into plant protoplasts has also been developed. For instance, viral RNAs, siRNAs and tRNAs were electroporated into protoplasts [12,13] and a PEG method allowed viral RNA to be delivered [14]. Here, we have adapted and optimized an efficient DNA transfection method described in details by Yoo *et a*l [11] to incorporate sncRNAs into *Arabidopsis thaliana* protoplasts. As an example of application, we describe here the transfection of an *A. thaliana* 20 nt long tRF alanine (tRF Ala20) [5] into protoplasts isolated either from an *A. thaliana* cell culture or from young seedlings. To validate the PEG transfection protocol, incorporation of tRF Ala20 was

verified by northern blot analysis of the small RNAs population extracted from RNase-treated transfected protoplasts. Therefore, we propose this efficient RNA transfection method as a new tool to study the biology of sncRNAs. For example, tagged (*e.g.* fluorophores, biotin) sncRNAs could be applied for *in vivo* live imaging or interactome studies.

2. Materials

2.1 Plant material

1. Culture of individualized cells obtained from the Landsberg *A. thaliana* ecotype.
2. *A. thaliana* Columbia (Col-0) ecotype seeds.

2.2 Nucleic acids

1. 100 µM synthetic oligoribonucleotide of interest dissolved in diethyl pyrocarbonate (DEPC)-treated water at a 100 µM concentration.
2. Any plasmid of small size (i.e. 3-4 kb).
3. 100 µM oligodeoxynucleotide with a sequence complementary to the sequence of the small RNA of interest.

2.3 Solutions

All solutions should be prepared with RNase-free materials and reagents, and with RNase-free, sterile and ultrapure water. Solutions are autoclaved for 15 min at 120°C (except otherwise stated). Procedures for preparing classical solutions can be found in Sambrook and Russell [15].

1. Cell culture medium: 4.41 g/L MS256, 30 g/L sucrose, 5.4 µM α-naphtaleneacetic acid (NAA)), 0.23 µM kinetin, pH 5.6. Weigh 4.41 g of MS256, 30 g of sucrose. Transfer to a glass beaker. Add about 500 mL of water. Mix well and adjust pH to 5.6 with 1 N KOH. Add 100 µL of 10 mg/mL NAA and 50 µL of 1 mg/mL kinetin. Transfer to a cylinder and adjust volume to 1 L with water. Autoclave and store at room temperature.
2. Seed sterilization solution: 0.1% (v/v) Tween-20, 70% (v/v) ethanol. Under a sterile hood, filter solution with a 0.20 µm filter in a sterile bottle. Store at room temperature.
3. Seedling culture medium: 2.21g/L MS222, 0.5g/L 2-(*N*-morpholin)ethanesulfonic acid (MES), 10 g/L sucrose, 0.68% agar, pH 5.7. Weigh 2.21 g of MS222, 0.5 g of MES monohydrate and 10 g of sucrose. Transfer to a glass beaker. Add about 500 mL of water. Mix well and adjust pH to 5.7 with 1 N KOH. Transfer to a cylinder and adjust volume to 1 L with water. Transfer to a bottle and add 6.8 g of agar. Autoclave and store at room temperature.
4. 2 M KCL. Store at room temperature.
5. 1 M $MgCl_2$. Store at room temperature.
6. 1 M $CaCl_2$. Store at room temperature.
7. 5 M NaCl. Store at room temperature.
8. 0.5 M Mannitol, pH 5.7: Weigh 227.75 g of D-mannitol and 1.75 g of MES monohydrate. Transfer to a glass beaker. Add about 1.5 L of water. Mix and adjust

pH to 5.7 with 10 N KOH. Transfer to a cylinder and adjust volume to 2.5 L with water. Autoclave and store at 4°C.

9. 0.5 M MES, pH 5.7: Weigh 53.3 g of MES monohydrate and transfer to a glass beaker. Add about 400 mL of water. Mix and adjust pH to 5.7 with 10 N KOH. Transfer to a glass cylinder and adjust volume to 500 mL with water. Autoclave and store at room temperature.
10. MM buffer: 0.5 M Mannitol, 4 mM MES, pH 5.7. Pipette 8 mL of 0.5 M MES, pH 5.7, and transfer to a 1 L glass cylinder. Add 0.5 M Mannitol, pH 5.7 to a final volume of 1 L and mix well. Autoclave and store at 4°C.
11. MMG buffer: 0.5 M Mannitol, 4 mM MES, 15 mM $MgCl_2$, pH 5.7. Pipette 7.5 mL of 1 M $MgCl_2$ and transfer to a 1 L glass cylinder. Add MM buffer to a final volume of 500 mL and mix well. Autoclave and store at 4°C.
12. WI buffer: 0.5 M Mannitol, 4 mM MES, 20 mM KCl, pH 5.7. Pipette 5 mL of 2 M KCl and transfer to a 1 L glass cylinder. Add MM buffer to a final volume of 500 mL and mix well. Autoclave and store at 4 °C.
13. W5 buffer : 2 mM MES , 154 mM NaCl, 125 mM $CaCl_2$, 5 mM KCl, pH 5.7. Mix in a glass cylinder 62.5 mL of 1 M $CaCl_2$, 15.4 mL of 5 M NaCl, 2 mL of 0.5 M MES pH 5.7 and 1.25 mL of 2 M KCl. Add water to a final volume of 500 mL and mix well. Autoclave and store at 4 °C.
14. TriReagent™ (Sigma-Aldrich) or TRIzol™ Reagent (Ambion). Store at 4 °C.
15. Chloroform. Store at room temperature.
16. Isopropanol. Store at room temperature.
17. 100 % Ethanol. Store at room temperature.
18. “cold” ethanol: 50 mL 100 % ethanol cooled to -20 °C .
19. 95 % Ethanol: Mix 47.5 mL of 100 % ethanol with 2.5 mL of water. Mix well and store at -20°C.
20. Diethyl pyrocarbonate (DEPC)-treated water: DEPC is toxic, prepare the solution under a chemical hood. In a bottle, add 0.001 % volume of DEPC to 100 mL water. DEPC has limited solubility in water as evidenced by the appearance of globules, thus mix well until globules disappear. Incubate for 16 hours, and autoclave. Make aliquots of 2 to 10 mL, and store at -20 °C.
21. 1 M Tris-HCl, pH 7.5: Weigh 12.11 g of Tris and transfer to a glass beaker. Add about 70 mL of water, mix and adjust pH to 7.5 with 1 N HCl (*see* **Note 1**). Transfer to a glass cylinder and adjust volume to 100 mL with water. Autoclave and store at room temperature.
22. 5 M LiCl solution: 5M LiCl, 50 mM Tris-HCl, pH 7.5. Weigh 21.2 g of LiCl and transfer to a glass beaker. Add 5 mL of 1 M Tris-HCl, pH 7.5, and water to a volume of 90 mL. Mix thoroughly until LiCl is completely dissolved. Transfer to a glass cylinder and adjust volume to 100 mL with water. Autoclave and store at room temperature.
23. 1 M Sodium acetate, pH 4.5. Weigh 8.2 g of sodium acetate and transfer to a glass beaker. Add water to a volume of 70 mL and mix until sodium acetate is completely dissolved. Adjust pH to 4.5 with 99% pure acetic acid. Transfer to a glass cylinder and adjust volume with water to 100 mL. Autoclave and store at room temperature.
24. 0.5 M EDTA, pH 8.0: Weigh 29.2 g of ethylenediaminetetraacetic acid (EDTA) and transfer to a glass beaker. Add 150 mL of water. Mix well and add NaOH pellets

until EDTA is completely dissolved (*see* **Note 2**). Adjust pH to 8.0 with 1 N NaOH. Transfer to a glass cylinder and adjust volume to 200 mL with water. Autoclave and store at room temperature.

25. 2x Loading buffer: 95 % (v/v) formamide, 20 mM EDTA, pH 8.0, 0.05 % (w/v), bromophenol blue, 0.05 % (w/v) xylene cyanol. Weigh 0.025 g of bromophenol blue and 0.025 g of xylene cyanol and transfer to a 50 mL tube. Add 2 mL of 0.5 M EDTA, pH 8.0, and 0.5 mL of water. Mix by vortexing and add 47.5 mL of formamide. Mix well and store at room temperature.
26. 10x TBE: 900 mM Tris, 20 mM EDTA, 900 mM boric acid. Weigh 109 g of Tris and 55.6 g of boric acid and transfer to a glass beaker. Add 40 mL of 0.5 M EDTA, pH 8.0, and water to a volume of about 900 mL, and mix. Transfer to a glass cylinder and adjust volume to 1 L with water. Autoclave and store at room temperature (*see* **Note 3**).
27. 1x TBE: 90 mM Tris, 2 mM EDTA, 90 mM boric acid. Mix 100 mL of 10x TBE with 900 mL of water in a glass cylinder. Mix well and store at room temperature.
28. 0.5x TBE: 45 mM Tris, 1 mM EDTA, 45 mM boric acid. Mix 50 mL of 10x TBE with 950 mL of water in a glass cylinder. Mix well and store at room temperature.
29. 15 % denaturing acrylamide gel: 15 % Acrylamide/Bisacrylamide 19/1, 7 M Urea, 1X TBE. Weigh 84 g of urea and transfer to a glass beaker. Add 75 mL of 40% acrylamide/bisacrylamide (19:1) and 20 mL of 10x TBE. Add water to a volume of 190 mL and mix thoroughly until the urea is completely dissolved. Transfer to a cylinder and adjust volume to 200 mL with water. Mix well and store at room temperature in a bottle wrapped in aluminum foil to avoid light exposure.
30. 10 % (w/v) Ammonium persulfate (APS). Store at 4°C.
31. *N,N,N,N'*-Tetramethyl-ethylenediamine (TEMED). Store at 4°C.
32. 1 % Ethidium bromide. Store at 4°C.
33. G-50 resin: Pour about 1 mL of Sephadex G-50 resin powder in a 14 mL round bottom tube. Add about 10 mL of water and mix well. Autoclave and store at 4°C.
34. Perfect Hyb™ Plus hybridization buffer (Sigma-Aldrich).
35. 20x Saline sodium citrate (SSC): 3 M NaCl, 300 mM trisodium citrate, pH 7.0. Weigh 175.3 g of NaCl and 88.2 g of sodium citrate dihydrate and transfer to a glass beaker. Add about 500 mL of water, and mix. Adjust pH to 7.0 with 99 % pure acetic acid. Transfer to a glass cylinder and adjust volume to 1 L with water. Autoclave and store at room temperature.
36. 2x SSC: 0.3 M NaCl, 30 mM trisodium citrate, pH 7.0. Mix 50 mL of 20x SSC with 450 mL of water in a glass cylinder. Mix well and store at room temperature.
37. 10 % (w/v) Sodium dodecyl sulfate (SDS) (*see* **Note 4**). Autoclave and store at room temperature.
38. 2x SSC, 0.1 % SDS: Transfer 50 mL of 20x SSC into a 500 mL cylinder. Add 400 mL of water and then 5 mL of 10 % SDS. Adjust volume to 500 mL by gently adding water to avoid foaming. Mix and store at room temperature (*see* **Note 5**).

2.4 Enzymes and other reagents

1. Cellulase R10 (Yakult Pharmaceutical Industry Co., Ltd).
2. Macerozyme R10 (Yakult Pharmaceutical Industry Co., Ltd).
3. 10 % acetylated BSA.

4. PEG-4000.
5. RNases A/T1 mix: 2 mg/mL RNase A, 5000 units/mL RNase T1.
6. 20 mg/mL Glycogen.
7. 10 units/µL polynucleotide kinase.
8. 10x polynucleotide kinase buffer.
9. 10 µCi/µL (2500 Ci/mmole) radiolabelled [γ^{32}P]-adenosine triphosphate (ATP).

2.5 Equipment and miscellaneous material

1. Erlenmeyers of 250 mL
2. Water bath.
3. Centrifuge allowing deceleration without brake, with rotors compatible for 2 mL, 14 mL and 50 mL tubes.
4. 2 mL round bottom tubes.
5. 14 mL round bottom tubes.
6. 50 mL tubes.
7. Wide mouth pipettes (tips of at least a 2 mm diameter).
8. Micropipette tips cut at their extremity to enlarge the opening (2 mm diameter).
9. Incubator.
10. Petri dishes (100 mm diameter; 20 mm height).
11. Square Petri dish (120 mm x 120 mm).
12. Scalpel blade.
13. 12 wells plate (3,65 cm^2/well) for cell culture.
14. Micropore gas-permeable tape.
15. Orbital shaker.
16. Tube rotator.
17. Vacuum chamber.
18. Miracloth.
19. Inverted bright field microscope.
20. Fuchs-Rosenthal cell.
21. Mini PROTEAN™ TetraHandcast System (Biorad).
22. Mini PROTEAN™ TetraCell System (Biorad).
23. 15-well comb.
24. Power supply Power Pac basic (Biorad).
25. UV transilluminator.
26. Trans-Blot™ Plus cell (Biorad).
27. Hybond™ N+ nylon membrane.
28. Whatmann 3 MM paper.
29. UV stratalinker.
30. 1 mL syringe.
31. Sephadex G-50.
32. Fiberglass.
33. Hybridization tubes.
34. Hybridization oven.
35. Phosphorimager (with plate) or X-ray developing unit (and X-ray film).

3. Methods

The workflow includes the isolation of protoplasts from either cell cultures or seedlings of *A. thaliana*, transfection of tRFs, extraction of total sncRNA from transfected protoplasts, and a northern blot experiment to confirm the uptake of tRFs into the cells.

Protoplasts must be manipulated very carefully to avoid lysis. Always use wide mouth pipettes with a large hole. Tips for micropipettes must be cut at their extremity to expand the opening (*see* **Note 6**). Never vortex the protoplasts and always centrifuge them without brake (*see* **Note 7**). When removing the supernatant after a centrifugation, always let a minimal volume of culture medium to cover protoplasts with liquid and avoid excessive changes of osmotic pressure. Do not directly pour the solution into the tube with protoplasts but rather always allow the solution to flow along the tube wall (*see* **Note 8**).

3.1 A. thaliana cell culture

3.1.1 Cell culture maintenance

1. Under a sterile hood, pour 100 mL of cell culture medium into a 250 mL Erlenmeyer.
2. Gently stir to homogenize the distribution of cells in a 7 days-old *A. thaliana* cell culture. Delicately transfer 4 mL of the culture into the new medium using a pipette.
3. Incubate for 7 days at 24°C under constant light on an orbital shaker at 130 rpm.
4. Repeat steps 1 to 3 weekly to maintain the cell culture.

3.1.2 Cell culture for protoplast isolation

1. To 100 mL of cell culture medium, add 4 mL of a 7 days-old *A. thaliana* cell culture as described above.
2. Incubate 10 - 11 days at 24°C under constant light condition on an orbital shaker at 130 rpm.

3.2 A. thaliana seedlings

3.2.1 Arabidopsis seed sterilization

All steps requiring opening of a tube need be performed under a sterile hood.

1. Prepare around 50 µL of *A. thaliana* Col-0 seeds in a 1.5 mL microtube (*see* **Note 9**).
2. Add 1 mL of seed sterilization solution. Mix to resuspend the seeds.
3. Gently mix for 15 minutes on a tube rotator.
4. Centrifuge at 3000 x g for 10 seconds to pellet the seeds.
5. Carefully remove the supernatant.
6. Add 1 mL of 100 % ethanol and mix to resuspend the seeds.

7. Gently mix on a tube rotator for 2 minutes.
8. Centrifuge at 3000 x g for 10 seconds to pellet the seeds.
9. Remove as much supernatant as possible.
10. Repeat steps 6 to 9 one more time.
11. Gently hit the tube with a finger to spread seeds, open the tube, and allow the seeds to dry for 1 hour.

3.2.2 Growing Arabidopsis seedlings

1. Heat seedling growth medium in a microwave to melt it. Under a sterile hood, pour 50 mL of the medium into a square Petri dish (120 mm x 120 mm) and let the dish open until the medium is solidified.
2. Scatter the sterilized seeds onto the solidified medium.
3. Seal the dish with micropore gas-permeable tape.
4. Allow the seeds to vernalize by keeping the dish at 4°C for 2 to 5 days.
5. Let seedlings grow for 7 days under long day conditions (8h dark at 17°C, 16h light at 21°C).

3.3 Isolation of protoplasts

3.3.1 Preparation of digestion buffer

Digestion buffer must be prepared just before use. The following protocol is given for 50 mL digestion buffer.

1. Pre-heat a water bath to 55°C.
2. Weigh 0.75 g of cellulase and 0.2 g of macerozyme and transfer to a 50 mL tube.
3. Add 40 mL of 0.5 M mannitol, pH 5.7, 2.5 mL of 0.4 M MES and 500 µL of 2 M KCl. Vortex and heat at 55°C until enzymes are completely dissolved.
4. Add 500 µL of 1 M $CaCl_2$ and 500 µL of acetylated BSA. Adjust volume to 50 mL with water and gently mix.

3.3.2 Removal of cell walls from cultured cells

25 mL of digestion buffer per 10 mL of cell culture are required.

1. Transfer 10 mL of a 10-11 days-old cell culture (Fig. 1A) into a 50 mL tube using a pipette.
2. Centrifuge the cells for 1 minute at 100 x g without brake. Remove most of supernatant (keep a small volume of medium on the top) by pipetting and gently tapping the tubes to resuspend cells.
3. Wash cells by adding 30 mL of 0.5 M mannitol, pH 5.7. Centrifuge for 1 minute at 100 x g without brake. As above, discard supernatant and tap the tubes to resuspend cells.
4. Add 25 mL of digestion buffer and pour the mixture into a Petri dish. Cover with aluminum foil and incubate for 4 hours at 26 °C at 50 rpm (orbital shaker).

3.3.3 Removal of cell walls from seedlings

25 mL of digestion buffer per 200 mg of young seedling leaves are required.

1. Place 2 mL of digestion buffer into a Petri dish.
2. From a 7 days-old culture of seedlings (Fig. 1B), cut seedlings with scissors to remove the roots. Weigh 200 mg of the remaining young leaves and transfer them to the Petri dish.
3. Carefully lacerate leaves with a scalpel blade and cover them with 23 mL of digestion buffer. Lacerated leaves must never be in contact with air and must not dry.
4. Incubate in a vacuum chamber at 0.01 bar until bubbles appear. Wait for 1 minute to allow buffer infiltration into tissues, and slowly break vacuum.
5. Cover Petri dish with aluminum foil, and incubate at 26°C overnight at 50 rpm (orbital shaker).

3.3.4 Purification of protoplasts

1. After cell wall digestion (using cells from either cell culture or seedlings), place a funnel with a single layer of miracloth on an Erlenmeyer. Wet the miracloth with 10 mL of MM buffer.
2. Add 10 mL of MM buffer to the protoplast digestion mixture (*see* **Note 8**). Break cell aggregates by pipetting 2-3 mL of the mixture with a 10 mL pipette and by creating bubbles by pipetting air. Pour the cell culture back into Petri dish, and repeat this operation several times, to allow the release of the majority of protoplasts. Disappearance of aggregates can be followed by observing a drop with an inverted microscope.
3. Filter protoplasts through the miracloth layer and transfer 10 mL aliquots of the filtrate into 14 mL round bottom tubes. This step removes the undigested cell walls and tissues.
4. Centrifuge for 3 minutes at 100 x g without brake. Remove supernatant and resuspend the protoplasts by gentle swirling.
5. Add 5 mL of W5 buffer per tube. Gather the content of two tubes into one. Centrifuge 1 min at 100 x g without brake.
6. Remove supernatant and softly resuspend protoplasts.
7. Repeat steps 5 and 6 until all protoplasts are gathered into two tubes.
8. Gently resuspend protoplasts by inversion of the tubes and add 10 mL of MMG buffer per tube. Centrifuge for 1 minute at 100 x g without brake. Remove supernatant and softly resuspend protoplasts.
9. Add 2 mL of MMG, gather tubes and incubate for at least 15 minutes on ice.
10. Add 10 mL of MMG. Centrifuge for 1 minute at 100 x g. Remove supernatant and softly resuspend protoplasts in 1 or 2 mL of MMG.

3.3.5 Determination of protoplast concentration

1. Clean a Fuchs-Rosenthal cell with 70 % ethanol and fix the cover glass with water.
2. Put a drop of protoplasts suspension between the slide and the cover glass.

3. Count the number of protoplasts under an inverted microscope, and determine the concentration (*see* **Note 10**). Protoplasts prepared either from cell culture or from seedlings are shown in Fig. 1.

3.4 Protoplast transfection

1 Prepare transfection buffer (*see* **Note 11**). Weigh 3 g of PEG-4000 and transfer to a 50 mL tube. Add 1.5 mL of water, 4 mL of 0.5 M mannitol, pH 5.7, and 1 mL of 1 M $CaCl_2$. Vortex until PEG-4000 is completely dissolved.
2 Prepare 2 mL round bottom microtubes each with at least 200 000 protoplasts at a final concentration of 2 million protoplasts/mL of MMG.
3 Add 1 µL of 100 µM small RNA and 15 µg of plasmid carrier DNA to each tube.
4 Gently mix by slowly inverting the tubes. Do not pipette.
5 Add 1 volume of transfection buffer and homogenize the protoplast suspension by inverting the tubes 2-3 times. Do not pipette.
6 Incubate for 15-30 minutes at room temperature.

3.5 Washing of transfected protoplasts

1. Add 2 volumes of W5 and homogenize the protoplast suspension by gently inverting the tubes.
2. Centrifuge for 1 minute at 100 x g without brake. Remove supernatant and gently resuspend protoplasts.
3. Repeat steps 1 and 2 three times.
4. Add 2 volumes of WI and homogenize the protoplast suspension by gently inverting the tubes.
5. Transfer the suspension to a 12 wells plate and close with micropore tape. Wrap the plate in aluminum foil and incubate at 25°C for the time required for preparing your experiment (*e.g.* 20 minutes are sufficient for northern blot analysis) (*see* **Note 12**).

3.6 RNase treatment

RNase treatment allows demonstrating that the small RNA is not just bound to the surface of protoplasts but incorporated into *A. thaliana* protoplasts.

1. Delicately pour 100 µL of the transfected protoplasts into a 2 mL round bottom tube. Add 1 µL of the RNases A/T1 mix. Incubate for 10 minutes at room temperature.
2. Add 1 mL of WI buffer. Centrifuge for 1 minute at 100 x g without brake. Remove supernatant by pipetting and gently resuspend the protoplasts.
3. Repeat step 2. After the last centrifugation, add 100 µL of WI to the protoplast pellet.

3.7 Small RNA extraction

Due to the toxicity of TriReagent™, steps 1 to 5 must be performed under a chemical hood.

1. Add 300 µL of TriReagent™ to the protoplast pellet. Vortex for 10 minutes.
2. Centrifuge for 10 minutes at 12000 x g, at 4°C. Transfer aqueous phase to a new tube.
3. Add 120 µL of chloroform and vortex for 10 minutes.
4. Centrifuge for 15 minutes at 12000 x g, at 4°C. Transfer aqueous phase to a new tube.
5. Add 240 µL of isopropanol. Delicately homogenize the solution by inverting the tube. Incubate for 5 minutes at room temperature, then for 10 minutes at 4°C.
6. Centrifuge for 15 minutes at 12000 x g, at 4°C. Remove supernatant.
7. Wash pellet by adding 175 µL of 95 % ethanol.
8. Centrifuge for 15 minutes at 12000 x g, at 4°C. Carefully remove supernatant and let the pellet dry at room temperature for 10 minutes to ensure that no ethanol remains in the tube.
9. Resuspend the pellet in 200 µL of water. Add 133 µL of LiCl solution. Incubate for 16 hours at 4°C (*see* **Note 13**).
10. Centrifuge for 30 minutes at 16000 x g, at 4°C. Transfer supernatant to a new tube.
11. Add 30 µL of 1 M sodium acetate, pH 4.5, 0.5 µL of glycogen, and 2.5 volumes of cold ethanol. Precipitate the nucleic acids for at least 1 hour at -20°C (*see* **Note 14**).
12. Centrifuge for 30 minutes at 16000 x g, at 4°C. Remove supernatant, let the pellet dry, then, dissolve the nucleic acids in 5 µL of DEPC-treated water.

3.8 RNA fractionation by polyacrylamide gel electrophoresis

1. Assemble plates in the casting frame and fix the frame on the casting stand.
2. Mix 7 mL of 15 % acrylamide, 49 µL of APS and 2.45 µL of TEMED. Quickly pour the solution into the space between the plates. Insert the 15-well comb and wait until the polymerization of the gel is completed.
3. To the samples (5 µL) add 5 µL of 2x loading buffer.
4. Load samples onto the gel. Run the electrophoresis at 150 V in 1x TBE until the bromophenol blue band reaches the bottom of the gel.
5. Prepare a container with water and add ethidium bromide to a final concentration of 0.5 µg/L.
6. Separate the gel plates with a spatula. Carefully take off the gel and transfer it to the ethidium bromide solution. Incubate for 5-10 minutes at room temperature with gentle agitation. Image the RNA profile with a UV transilluminator.

3.9 Electrophoretic blotting procedure

1. Cut 4 Whatman sheets to the size of the plates, and a Hybond N+ membrane square to a size slightly larger than the gel.
2. Soak a sponge from the electroblot apparatus with 0.5x TBE and place it in the opened electroblotting cassette. Wet two sheets of Whatman paper with 0.5x TBE and place them onto the sponge. Soak the Hybond N+ membrane with 0.5x TBE and place it onto the top of the Whatman sheets. Rinse the gel in 0.5x TBE and

transfer the gel onto the nylon membrane. Remove any bubbles (*see* **Note 15**). Soak two more sheets of Whatman paper and place them on top of the membrane. Wet the second sponge with 0.5x TBE and place it on top. Clamp the cassette closed and place the assembled system into the chamber. Pay attention to the cassette orientation in the chamber to allow transfer of RNAs from the gel onto the membrane. Fill the chamber with 0.5x TBE.

3. Carry out the transfer of the RNAs from the gel to the membrane at 300 mA for 80 minutes at 4°C.
4. Place two sheets of Whatmann paper in a container and soak them with 2x SSC. Discard excessive buffer. Disassemble the electroblotting system, remove the nylon membrane and place it onto the sheets of Whatmann paper with the surface carrying the RNA facing up. Incubate for 10 minutes at room temperature. Place the container into a Stratalinker to crosslink the RNA to the membrane by applying two UV exposures of 120000 µJ/cm^2.

3.10 Labelling of the oligodeoxynucleotide probe

All steps requiring manipulation of radioactivity must be carried out in an appropriate room. The safety rules of your institution must be strictly followed.

1. Preheat a water bath at 37°C.
2. Mix 4.5 µL of water, 1 µL of 100 µM oligodeoxynucleotide complementary to the small RNA of interest, 1 µL of 10x T4 polynucleotide kinase (PNK) buffer, 1 µL of PNK and 2.5 µL of [γ^{32}P]-ATP. Incubate for 30 minutes at 37°C.
3. Remove unincorporated [γ^{32}P]-ATP with a Sephadex G-50 spin column. To prepare a G-50 column, add 2 mm of fiberglass into the bottom of a 1 mL syringe. Place the syringe into a 14 mL tube. Load swelling watered G-50 resin into the syringe and allow the water to drop out (but do not let the resin dry). Add resin until the syringe is filled. Centrifuge for 45 seconds at 2000 x g. Place the syringe in a new 14 mL tube.
4. Add 50 µL of water to the PNK reaction mixture and load it onto the top of the G-50 column. Centrifuge for 45 seconds at 2000 x g to elute the radioactive probe. Discard the syringe in the radioactive waste disposal container

3.11 Northern Hybridization

1. Wet the membrane with PerfectHyb^{TM} Plus solution and place it into a cylindrical glass hybridization tube. Make sure that the membrane is well attached to the tube wall.
2. Add 4 mL of PerfectHyb^{TM} Plus hybridization solution into the tube.
3. Add 940 µL of PerfectHyb^{TM} Plus to the radiolabeled probe, and transfer it into the hybridization tube.
4. Place the tube into the hybridization oven, and incubate with gentle rotation overnight at 42°C.
5. Discard the probe into the radioactive waste disposal container or keep it at -20 °C if you wish to reuse the probe another time (the half-time of ^{32}P is 14 days).

6. Pour 10 mL of 2x SSC into the tube and wash the membrane for 10 minutes at 42°C with gentle rotation.
7. Discard the 2x SSC buffer into the waste container dedicated to radioactivity. Repeat step 6.
8. Discard 2x SSC as above. Add 10 mL of 2x SSC, 0.1 % SDS into the tube and wash the membrane for 30 minutes at 42°C with gentle rotation.
9. Discard the washing solution. Place the membrane on a sheet of Whatman paper to remove excess liquid. Then place the membrane between two sheets of plastic wrap.
10. Expose the radioactive signal of membrane to a phosphorimager screen or to X-ray film. Image the screen using a phosphorimager or develop the X-ray film (*see* **Note 16**).
11. As an illustration, the autoradiography of a northern blot showing the successful transfection of sncRNA into *A. thaliana* protoplasts is presented Fig. 2. In this example, an oligoribonucleotide (5'-ACCATCTGAGCTACATCCCC-3') corresponding to a previously identified tRF Ala of 20 nt (tRF Ala20) derived from the 5' extremity of an *A. thaliana* cytosolic alanine tRNA [5] has been incorporated by the protoplasts. The oligoribonucleotide transfection was further analyzed by a northern blot experiment as described above. Control experiments were performed. In the absence of the Ala20 oligoribonucleotide (control with water), or in the absence of PEG, no signal corresponding to the tRF was observed. After RNase treatment, a signal corresponding to the tRF Ala20 was still visible in the transfected protoplasts, thus demonstrating successful incorporation into protoplasts. This oligoribonucleotide transfection method could be widely used to identify the interactome of a biotinylated tRF or the subcellular localization of a fluorescently tagged tRF.

4. Notes

1. The addition of HCl increases the temperature of the solution and distorts pH measurement. Always wait for temperature cooling before measuring pH.
2. EDTA can only be dissolved at a basic pH. Add NaOH pellets and wait a few minutes. Then add additional pellets one by one until the EDTA powder is dissolved.
3. Boric acid precipitates after long storage (around 1 month). Thus, preparing a large volume of 10x TBE is not advised.
4. Heating is required to dissolve SDS.
5. SDS can precipitate if the room temperature decreases below 20°C. Heat the solution to dissolve SDS before use, if needed.
6. Pipetting with normal tips exposes protoplasts to an excessive osmotic pressure change and can smash them. Better use pipettes with wide mouth tips, and cut micropipette tips with scissors to enlarge the opening hole. Avoid pipetting and rather pour protoplast solutions to the new containers if possible.
7. Applying braking combined with centrifuge force can smash the cells.
8. Add the buffer slowly along the wall of the tube or the Petri dish to avoid excess of pressure and mechanical force, which could smash the protoplasts.
9. Quality of the seeds is important. Better to use freshly harvested seeds (less than one year old). If you use old seeds, sterilization is less efficient. In the case of old seeds or

mutant seeds, we have used use 80 % ethanol, 4 % Tween™ 20 (rather than 70% ethanol, 0.1% Tween™ 20) as sterilization solution. Please note that times must be respected to avoid damage to the seeds.

10. Intact protoplasts are perfectly round and well isolated (see Fig. 1). Their size can vary, but both large and small ones can be counted if they are not damaged. The number of protoplasts prepared from 10 mL of cell culture is usually around 2 x 10^6 whereas 200 mg of seedlings will give about 1 x 10^6 protoplasts.
11. The transfection buffer must be freshly prepared the day of transfection.
12. Protoplasts can be kept in the 12-wells plate for a maximum of 30 hours. The longer the protoplasts are kept in the plate, the more of them will die.
13. LiCl causes the precipitation of large RNAs (*i.e.*> 200 nt in size). After centrifugation, the supernatant contains tRNAs and small non-coding RNAs (including tRFs).
14. One hour is usually sufficient for precipitation if ethanol kept at -20°C is used. Otherwise, precipitation can also be performed overnight.
15. To avoid bubbles, the membrane should be wet before placing the gel on it. Bubbles can be chased by smoothly pressing a pipette on the gel.
16. To increase signal with X-ray exposure, use a reflective screen placed on the film and incubate the cassette at -80°C. Under these conditions, a strong signal can be obtained after 6 hours of exposure. At room temperature, an exposure time of 24-48 hours is recommended.

References

1 Shen Y, Yu X, Zhu L, Li T, Yan Z, Guo J (2018) Transfer RNA-derived fragments and tRNA halves: biogenesis, biological functions and their roles in diseases. J Mol Med 96: 1167 -1176

2 Megel C, Morelle G, Lalande S, Duchêne A-M, Small I, Maréchal-Drouard L (2015) Surveillance and cleavage of eukaryotic tRNAs. Int J Mol Sci 16: 1873-1893

3 Schimmel P (2018) RNA Processing and Modifications: The emerging complexity of the tRNA world: Mammalian tRNAs beyond protein synthesis. Nat Rev Mol Cell Biol 19: 45–58

4 Cristodero M, Polacek N (2017) The multifaceted regulatory potential of tRNA-derived fragments. Non-coding RNA Investig 1: 7

5 Cognat V, Morelle G, Megel C, Lalande S, Molinier J, Vincent T, Small I, Duchêne A-M, Maréchal-Drouard L (2017) The nuclear and organellar tRNA-derived RNA fragment population in *Arabidopsis thaliana* is highly dynamic. Nucleic Acids Res 45: 3460–3472

6 Haas G, Cetin S, Messmer M, Chane-Woon-Ming B, Terenzi O, Chicher J, Kuhn L, Hammann P, Pfeffer S (2016) Identification of factors involved in target RNA-directed microRNA degradation. Nucleic Acids Res 44: 2873 - 2887

7 Han Z, Ge X, Tan J, Chen F, Gao H, Lei P, Zhang J (2015) Establishment of lipofection protocol for efficient miR-21 transfection into cortical neurons in vitro. DNA Cell Biol 34: 703–709

8 Spörlein B, Koop H-U (1991) Lipofectin: direct gene transfer to higher plants using cationic liposomes. Theor Appl Genet 83: 1 -5

9 Miao Y, Jiang L (2007) Transient expression of fluorescent fusion proteins in protoplasts of suspension cultured cells. Nat Protoc 2: 2348–2353

10 Masani MYA, Noll GA, Parveez GKA, Sambanthamurthi R, Prüfer D (2014) Efficient transformation of oil palm protoplasts by PEG-mediated transfection and DNA microinjection. PLoS One 9: e96831

11 Yoo S-D, Cho Y-H, Sheen J (2007) Arabidopsis mesophyll protoplasts: a versatile cell system

for transient gene expression analysis. Nat Protoc 2: 1565 -1572
12 Qi Y, Zhong X, Itaya A, Ding B (2004) Dissecting RNA silencing in protoplasts uncovers novel effects of viral suppressors on the silencing pathway at the cellular level. Nucleic Acids Res 32: e179
13 Wintz H, Dietrich A (1996) Electroporation of small RNAs into plant protoplasts: mitochondrial uptake of transfer RNAs. Biochem Biophys Res Commun 223: 204 -210
14 McCormack JC, Simon AE (2006) Callus cultures of Arabidopsis. Curr Protoc Microbiol 00: 16D.1.1-16D.1.9
15 Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual (C. S. Harbor, ed.) Cold Spring Harbor Laboratory Press, NY

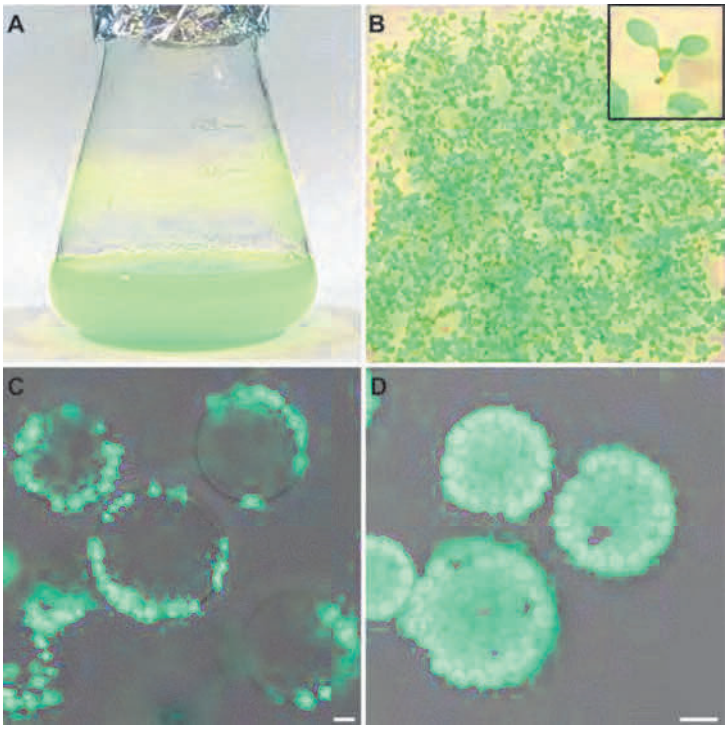


Fig. 1: Protoplast preparation from *A. thaliana* cell culture and seedlings
(**A**) 11 days-old cell culture of *A. thaliana* ecotype Landsberg. (**B**) 7 days-old seedlings of the *A. thaliana* Col-0 ecotype. Enlargement of a seedling is shown in the top right edge (**C**) Visualization by confocal microscopy of isolated protoplasts prepared from cell culture. (**D**) Visualization by confocal microscopy of isolated protoplasts from seedlings. Chlorophyll autofluorescence (green) reveals chloroplasts and their integrity (Zeiss LSM780 confocal microscope, 633 nm laser). Scale bars: 10 µm.

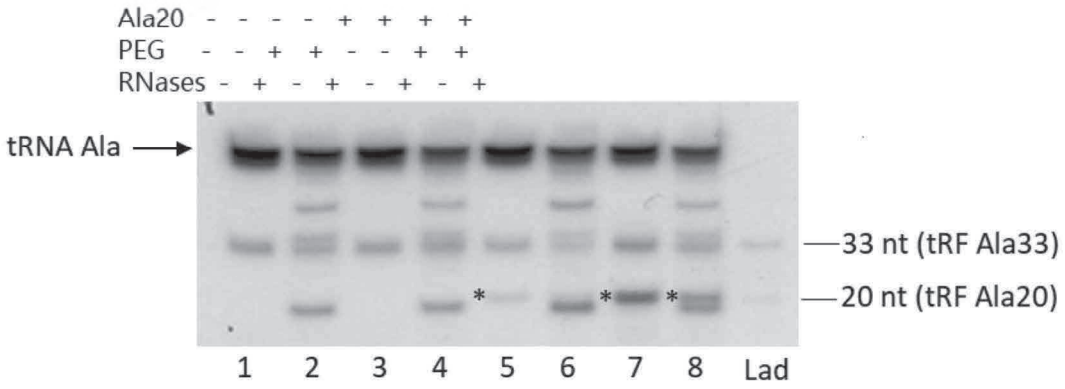


Fig. 2: Analysis by northern blot of the transfection of tRF Ala20 into *A. thaliana* protoplasts
Autoradiography of the northern blot experiment performed with RNA samples extracted from *A. thaliana* protoplasts transfected with an oligoribonucleotide (tRF Ala20) corresponding to the cytosolic *A. thaliana* tRNA Alanine (UGC) of 20 nt in size. The same oligoribonucleotide was used as a ladder (Lad), together with an oligoribonucleotide corresponding to tRF Ala33 (33 nt of the 5' extremity of $tRNA^{Ala}$ (UGC)).
Protoplasts prepared from an *A. thaliana* cell culture were transfected in the absence (-) or in the presence (+) of tRF Ala20 as well as in the absence (-) or in the presence (+) of PEG. Transfected protoplasts were treated (+) or not treated (-) with a mix of RNase A and RNase T1 (RNases). Extracted RNAs were probed with a ^{32}P- radiolabelled oligodeoxynucleotide probe complementary to the 5' extremity of $tRNA^{Ala}$ (UGC), (5'-GGGGATGTAGCTCAGATGGT-3').
In all samples, signals corresponding to the endogenous alanine tRNA and the endogenous tRF Ala33 are seen. RNA samples from protoplasts treated with RNases exhibit additionnal bands, which is likely due to protoplasts lysis and accessibility of alanine tRNA to nucleases (lanes 2, 4, 6 and 8). Lane 5 with RNA samples of protoplasts treated for transfection in the absence of PEG shows a band corresponding to tRF Ala20. This signal is not present upon treatment of the protoplasts with RNases (lane 6). By contrast, this signal is higher when protoplasts were treated for transfection in the presence of PEG (lane 7) and this signal is still present when such protoplasts were post-treated with RNases (lane 8), thus demonstrating the successful transfection and internalization of tRF Ala20 into *A. thaliana* protoplasts.

Annexe 2

En plus de l'étude sur l'export nucléaire des ARN du CaMV, j'ai participé au tout début de ma thèse, aux travaux réalisés par Angèle Geldreich *et al.* et initiés précédemment par Gabrielle Haas, sur une toute autre thématique du laboratoire, la formation des viroplasmes denses du CaMV. Ces travaux ont conduit à l'article suivant, publié dans PLoS One, fin 2017.

Formation of large viroplasms and virulence of Cauliflower mosaic virus in turnip plants depend on the N-terminal EKI sequence of viral protein TAV.

Geldreich A.[co], Haas G.[co], **Kubina J.**, Bouton C., Tanguy M., Erhardt M., Keller M., Ryabova L., Dimitrova M.

PLoS One. 2017 Dec 18;12(12):e0189062. doi: 10.1371/journal.pone.0189062. eCollection 2017.

Résumé

Les viroplasmes sont des structures subcellulaires dépourvues de membrane, formées par de nombreux virus au cours de leur cycle de réplication. Il s'agit de corps d'inclusion où les virus concentrent leurs protéines, leurs acides nucléiques et parfois des facteurs de l'hôte ce qui leur permet de se répliquer avec efficacité, voire de se protéger contre les systèmes antiviraux cellulaires (Netherton et Wileman, 2011; Castro et al., 2013; Schmid et al., 2014).

Chez le CaMV, les viroplasmes dits denses sont formés par la protéine multifonctionnelle P6 qui est exprimée très abondamment au cours du cycle viral. C'est pour cette raison que P6 est aussi connue sous le nom de TAV pour – *trans-activator viroplasmine* – (Bonneville et al., 1989) faisant référence à la fois à sa fonction d'activateur traductionnel et de composant principal des viroplasmes. Ces derniers apparaissent denses aux électrons sous microscope à contraste interférentiel (Fujisawa et al., 1967; Martelli et Russo, 1977) permettant ainsi de les distinguer des viroplasmes dits clairs qui sont constitués majoritairement de P2, les corps de transmission des particules virales par les pucerons. P6 possède des domaines qui lui permettent de s'auto-assembler (Leisner et Schoelz, 2018) en absence de toute autre protéine virale pour former de petits agrégats qui fusionnent ensuite pour donner naissance à ces corps d'inclusion cytoplasmiques si caractéristiques de ce virus. Les viroplasmes denses occupent une place primordiale durant le cycle de réplication du CaMV puisque ce dernier y réplique son ADN génomique et, forme et stocke les nouveaux virions qui seront acheminés vers les plasmodesmes pour infecter les cellules adjacentes ou vers les viroplasmes clairs pour être transmis par les insectes vecteurs (Schoelz et Leisner, 2017). De ce fait, les viroplasmes denses ont toujours été

considérés comme l'usine virale par excellence du CaMV, indispensable pour son pouvoir infectieux.

Une première étude réalisée au laboratoire a montré qu'un mutant particulier de P6 (ou TAV) de l'isolat virulent CaMV Cabb B-JI, appelé TAVm3, et dont les 3 résidus E, K et I en position 11, 12 et 13 ont été mutés en alanines, était incapable de former de grands viroplasmes denses en cellules de tabac BY2 et chez Arabidopsis (Haas et al., 2005, 2008). TAVm3 est seulement capable de former quelques agrégats de petite taille et disséminés dans la cellule. Nous avons montré par la suite que deux autres domaines de P6/TAV sont impliqués dans la formation des viroplasmes, la région MAV et un domaine C-terminal comprenant un doigt de zinc.

L'inaptitude de TAVm3 à former de grands viroplasmes n'a curieusement aucune réelle incidence sur l'infectiosité du CaMV alors qu'il était admis depuis longtemps qu'ils constituaient le site stratégique de son cycle infectieux (Haas et al., 2008). Cette propriété de TAVm3 nous a donc incité à investiguer l'importance fonctionnelle des viroplasmes denses dans le cycle du CaMV en suivant sa symptomatologie qui témoigne de la virulence du virus et, en analysant les différentes étapes clés du cycle à savoir la synthèse des protéines, la réplication de l'ADN viral et la production de virions, chez des plants de navet infectés par un CaMV muté exprimant TAVm3.

L'infection des plants de navet par le CaMV mutant se manifeste par des symptômes systémiques doux qui rendent les plantes difficilement distinguables des plantes saines. À l'inverse, les plantes infectées par le virus sauvage montrent une mosaïque, un gaufrage foliaire et un nanisme prononcés indiquant que le motif EKI est impliqué dans l'expression des symptômes. Par contre, la cinétique d'infection systémique, suivie par détection de l'ADN viral, est similaire chez les plantes inoculées avec le CaMV sauvage et le virus muté.

L'observation au microscope électronique à transmission, des cellules de navet infectées, confirme les résultats de l'expression transitoire du mutant TAVm3 en cellules BY-2 et en protoplastes d'Arabidopsis à savoir l'absence de grands viroplasmes denses. Certaines cellules renferment des petits agrégats (<1 μm^2) d'aspect irrégulier et contenant de rares particules virales. Etonnement, ces cellules sont également dépourvues de viroplasme clair (les cellules infectées ne renferment qu'un seul corps de transmission). Tout l'inverse des cellules de plants de navet infectés par le CaMV sauvage arborant des viroplasmes denses de grande taille et chargés d'innombrables virions ainsi qu'un viroplasme clair.

Une étude plus approfondie du virus mutant a permis de révéler, par le biais de marquages immunofluorescents, que la protéine P2 est en fait incapable de former des viroplasmes clairs bien qu'elle soit produite en grande quantité, indiquant que la mutation du motif EKI de TAV a une répercussion sur l'assemblage de P2 en viroplasmes. La purification des

virions des plantes infectées respectivement par le CaMV sauvage et le virus muté montre que ce dernier produit au moins 50 fois moins de nouveaux virions que le virus sauvage, confirmant ainsi nos observations en microscopie électronique. Et ce, malgré que les plantes infectées par le CaMV-TAVm3 contiennent des quantités appréciables de protéine de capside P4 (environ 40 % par rapport au virus sauvage) et sous sa forme parfaitement maturée, démontrant ainsi que l'assemblage des capsides s'effectue normalement dans les cellules infectées par le CaMV muté. Par contre, la quantité d'ADN viral est nettement réduite dans ces cellules ce qui indique que l'absence de viroplasmes denses affecte fortement l'étape de rétrotranscription.

En résumé, la formation des grands viroplasmes denses et la virulence du CaMV sont dépendantes du motif EKI situé dans la région N-terminale de la protéine TAV. Toutefois, l'absence de ces viroplasmes n'est pas préjudiciable au CaMV *in planta*, étant donné que la faible production de virions est suffisante pour infecter systémiquement les plantes. En revanche, ce défaut des viroplasmes denses pourrait porter préjudice au passage du virus de plante-à-plante car il entraîne l'absence des viroplasmes clairs, considérés comme indispensables pour la transmission du virus par son vecteur le puceron.

RESEARCH ARTICLE

Formation of large viroplasms and virulence of *Cauliflower mosaic virus* in turnip plants depend on the N-terminal EKI sequence of viral protein TAV

Angèle Geldreich[☯], Gabrielle Haas[¤a☯], Julie Kubina, Clément Bouton[¤b], Mélanie Tanguy[¤c], Mathieu Erhardt, Mario Keller, Lyubov Ryabova, Maria Dimitrova*

Institut de Biologie Moléculaire des Plantes, CNRS UPR2357, Université de Strasbourg, Strasbourg, France

☯ These authors contributed equally to this work.
¤a Current address: Institut de Biologie Moléculaire et Cellulaire, CNRS UPR9022, Université de Strasbourg, Strasbourg, France
¤b Current address: Rothamsted Research, Harpenden, United Kingdom
¤c Current address: NIBSC, South Mimms, Potters Bar, United Kingdom
* m.dimitrova@unistra.fr

Check for updates

OPEN ACCESS

Citation: Geldreich A, Haas G, Kubina J, Bouton C, Tanguy M, Erhardt M, et al. (2017) Formation of large viroplasms and virulence of *Cauliflower mosaic virus* in turnip plants depend on the N-terminal EKI sequence of viral protein TAV. PLoS ONE 12(12): e0189062. https://doi.org/10.1371/journal.pone.0189062

Editor: Mikhail M. Pooggin, University of Basel, SWITZERLAND

Received: July 18, 2017

Accepted: November 17, 2017

Published: December 18, 2017

Copyright: © 2017 Geldreich et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by Centre National de la Recherche Scientifique and Université de Strasbourg. Gabrielle Haas, Clément Bouton and Julie Kubina PhD were supported by fellowships from Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation. The funders had no role in study design, data collection

Abstract

Cauliflower mosaic virus (CaMV) TAV protein (TransActivator/Viroplasmin) plays a pivotal role during the infection cycle since it activates translation reinitiation of viral polycistronic RNAs and suppresses RNA silencing. It is also the major component of cytoplasmic electron-dense inclusion bodies (EDIBs) called viroplasms that are particularly evident in cells infected by the virulent CaMV Cabb B-JI isolate. These EDIBs are considered as virion factories, vehicles for CaMV intracellular movement and reservoirs for CaMV transmission by aphids. In this study, focused on different TAV mutants *in vivo*, we demonstrate that three physically separated domains collectively participate to the formation of large EDIBs: the N-terminal EKI motif, a sequence of the MAV domain involved in translation reinitiation and a C-terminal region encompassing the zinc finger. Surprisingly, EKI mutant TAVm3, corresponding to a substitution of the EKI motif at amino acids 11–13 by three alanines (AAA), which completely abolished the formation of large viroplasms, was not lethal for CaMV but highly reduced its virulence without affecting the rate of systemic infection. Expression of TAVm3 in a viral context led to formation of small irregularly shaped inclusion bodies, mild symptoms and low levels of viral DNA and particles accumulation, despite the production of significant amounts of mature capsid proteins. Unexpectedly, for CaMV-TAVm3 the formation of viral P2-containing electron-light inclusion body (ELIB), which is essential for CaMV aphid transmission, was also altered, thus suggesting an indirect role of the EKI tripeptide in CaMV plant-to-plant propagation. This important functional contribution of the EKI motif in CaMV biology can explain the strict conservation of this motif in the TAV sequences of all CaMV isolates.

and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Virus-infected cells often contain cytoplasmic and/or nuclear inclusion bodies mainly composed of viral proteins and called viroplasms. Viroplasms are considered for many viruses as viral factories providing a physical scaffold to concentrate viral components and host factors within specific sites, and thereby increasing the efficiency of genome replication and/or assembly of viral particles [1]. In cells infected by some animal viruses, viroplasms result from the assembly of small aggregates that are transported by dynein along microtubules to the microtubule organization centre at the periphery of the nucleus, where they recruit cellular proteins and mitochondria [2–4]. These viroplasms resemble aggresomes that naturally occur in cells to reduce the toxicity of misfolded proteins and make them susceptible to proteolysis by the proteasome and/or by autophagy [2]. These inclusion bodies appear during the early steps of viral infection, dramatically alter the cell ultrastructure and differ from aggregates of overproduced viral proteins that accumulate late in infection. Most animal and plant RNA viruses modify membranes usually deriving from the secretory pathway, endoplasmic reticulum and Golgi apparatus, or from organelles such as mitochondria or peroxisomes, where they anchor their replication complexes and establish virus assembly sites [1]. RNA viruses belonging to the *Reoviridae* family also form viroplasms in which virions are produced, both in animals [5] and plants [6]. Viroplasms are also observed in cells infected by plant DNA viruses such as *Cauliflower mosaic virus* of the *Caulimoviridae* family [7–9].

Cauliflower mosaic virus (CaMV) is a plant pararetrovirus and the type member of the *Caulimovirus* genus [10,11]. Its circular double-stranded DNA genome (~8 kbp) is replicated through reverse transcription of the pre-genomic 35S RNA. The six proteins (P1 to P6) encoded by the CaMV genome are expressed from the polycistronic 35S RNA, whereas P6, also called TAV (TransActivator/Viroplasmin), is primarily synthesized from the monocistronic 19S RNA. TAV (62 kDa) is a nucleocytoplasmic shuttling protein essential for CaMV infectivity [9,12]. It is the major determinant of host specificity [13,14] and expression of symptoms [15,16] but it also triggers a hypersensitive response in *Nicotiana* plant species depending on the CaMV strains [17]. TAV interferes with plant defence mechanisms by inhibiting signalling responses to salicylic acid [18] and by suppressing RNA silencing possibly *via* the interaction with DRB4, the cofactor of DCL4 [12,19]. TAV also promotes the reinitiation of translation along the polycistronic 35S RNA thanks to its multiple interactions with the translation machinery [20], with a specific plant protein called RISP [21], and TOR protein kinase [22], thus enabling the synthesis of the full set of viral proteins [11,20]. In addition, TAV is required for an efficient replication of the CaMV genome, but the underlying mechanism is still unclear [23]. TAV might be involved in the assembly and the transport of CaMV particles, since it interacts with the capsid protein (P4), with CHUP1, a plant protein responsible for moving chloroplasts on actin microfilaments, and with plasmodesmal proteins [24–27]. Finally, TAV is also the major component of membrane-free amorphous electron-dense inclusion bodies/viroplasms (EDIBs) [8,9,28], easily distinguishable by electron microscopy in the cytoplasm of CaMV-infected cells from electron-lucent inclusion bodies/viroplasms (ELIBs), which are mainly constituted of P2 protein and are used as platforms for plant-to-plant transmission of CaMV by aphids [29]. EDIBs result from the self-association of TAV molecules without the assistance of any other viral protein [9,28]. Small TAV aggregates move along the actin filaments network, but they appear stationary on microtubules [26]. It has been proposed that EDIBs could be considered as mobile viral factories, which, during the late events of infection, dock and deliver virions to the tubules formed by movement protein P1 in plasmodesmata [26,30].

EDIBs used to be considered as containers of overexpressed non-functional TAV protein, until it was shown that they are the sites of CaMV DNA replication, virus assembly and storage of newly formed viral particles. The translation of CaMV transcripts is supposed to occur within or close to EDIBs, since the latter are surrounded by numerous ribosomes. Heterologous proteins (*i.e.* interferon) expressed from recombinant CaMV DNA have been detected in proximity of EDIBs [31]. These observations explain why it is generally accepted that EDIBs are CaMV factories and, consequently, why they are considered to be mandatory for CaMV infectious cycle.

CaMV-infected cells contain many EDIBs with areas ranging from 0.3 to 10 μm^2, depending on both the virus strain and the host plant [32]. Virulent CaMV Cabb B-JI strain triggers in *Brassicaceae* the formation of unusually large perinuclear EDIBs (8 to 10 μm^2) that arise from the gradual assembly of small TAV-aggregates (0.1 to 1 μm^2) during the course of infection. On the contrary, other CaMV strains such as D/H or CM1841 only form small viroplasms at the final stage of infection [32]. The reason for these size differences is to date not fully understood. We have shown previously that the formation of large EDIBs is impaired in cultured tobacco BY2 cells [9] upon mutation of the EKI tripeptide (amino acids 11–13) of TAV, to AAA (mutant called TAVm3), and that this mutation did not abolish CaMV infectivity in *Arabidopsis thaliana* [12]. These observations suggested that large EDIBs might be dispensable for the virus. However, we could not totally exclude that EDIBs can still assemble in a viral context in host plants. Recently, it was shown that mutations within a 35 amino acids region (subdomain D3b) in the central part of TAV halved the size of EDIBs and CaMV infectivity [33].

Here we demonstrate that two other independent domains of TAV are required for the formation of EDIBs, in addition to the EKI motif and the recently described subdomain D3b [33]. We also show that the EKI mutation impairs the formation of large EDIBs in turnip plants, and thereby the virulence of the Cabb B-JI isolate, but does not affect CaMV systemic infection. Finally, we demonstrate that the absence of large EDIBs affects CaMV protein synthesis and capsid assembly, and dramatically reduces CaMV DNA replication efficiency and the subsequent production of infectious particles.

Materials and methods

Virus and host plants

All experiments were performed with the CaMV Cabb B-JI isolate. Three week-old turnip plants (*Brassica rapa*, cv. Tokyo) were mechanically inoculated with *Sal*I-linearized pMD324 or pGH plasmids coding for wild-type TAV or the mutated version TAVm3, respectively [12]. Inoculations were performed using, per plant, 10 μg of linearized CaMV DNA diluted in 20 μl sterile water. Ten turnip plants were inoculated with wild type or mutated viral DNA per assay. Infected leaves were harvested 21 days post-inoculation (dpi), ground in liquid nitrogen, and stored at -80°C. The results reported in this study were exclusively obtained from systemically infected tissues. Transgenic *Arabidopsis thaliana* (ecotype Columbia) expressing CaMV TAV or TAVm3 were described in [12].

Plasmid constructions

The pCK-EGFP vector was used to construct the recombinant plasmids encoding the fusion proteins, consisting of enhanced green fluorescent protein (EGFP) and wild-type or mutant CaMV TAV. Full-length CaMV ORF VI and 3' truncated ORF VI sequences were obtained by PCR performed on pETKaKS.6 recombinant plasmid containing the complete ORF VI of Cabb B-JI isolate [34] with appropriate primers. Amplification products were cloned in *Nco*I

restriction site of the pCK-EGFP vector and verified by PCR to be in fusion and in frame with the 3' end of the EGFP encoding sequence. Deletions and point mutations were introduced in the recombinant plasmid encoding EGFP:TAV by site-directed mutagenesis using specific primers, as previously described [9]. CaMV ORF VI sequence was also amplified using two primers carrying at their 5' ends *Bsr*GI and *Xba*I sites, respectively, and cloned in these restriction sites of the pmRFP vector, in fusion with the sequence encoding the red fluorescent protein (mRFP). Forward and reverse primers used to generate ORF VI and its mutants are listed in S1 Table. Error-free recombinant plasmids were identified by DNA sequencing.

Purification of CaMV particles

CaMV particles were extracted from infected turnip leaves ground in the presence of 1M urea to disrupt viroplasms and purified on a 10–40% sucrose gradient as previously described [35]. The concentration of purified virus was determined by spectrophotometry (one OD at 260 nm corresponding to 7 mg/ml of virions) [35].

Analysis of CaMV proteins in infected turnip plants

Total proteins from systemically infected turnip leaves were extracted in Laemmli buffer 2X concentrated, containing 8 M urea, and fractionated by sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE) 12%. After migration, proteins were electroblotted on PVDF membrane (Immobilon-P, Millipore) that was then stained with Coomassie blue in order to detect ribulose 1,5, bisphosphate carboxylase/oxygenase large subunit (RbcL) used as loading control. CaMV TAV, P4, P3, P2 and P1 proteins were detected using polyclonal antisera at a dilution of 1:10,000, alkaline phosphatase-conjugated goat anti-rabbit IgG antibodies (Biosys) at a dilution of 1:5,000 and the colorimetric substrate BCIP (5-Bromo- 4-Chloro-3-Indolyl Phosphate) / NBT (Nitroblue Tetrazolium) detection system (Promega)). Anti-P1 and anti-P2 antibodies were kindly provided by Andy Maule (John Innes Centre, Norwich, England) and Stéphane Blanc (INRA, Montpellier, France), respectively, and anti-P3 [36], anti-P4 [37] and anti-TAV [38] polyclonal antisera were previously obtained and tested in our laboratory. The amounts of viral proteins were quantified using ImageJ software (http://rsb.info.nih.gov/ij/) based on the intensity of the signal obtained for each of the 5 viral proteins relative to the RbcL control protein. The means for each viral protein from wild-type CaMV and CaMV-TAVm3-infected plants were compared using a Student's 2-samples t-test and those for wild-type CaMV were set at 100%.

26S Proteasome inhibition in transgenic *Arabidopsis* seedlings

Eight-days-old transgenic *A. thaliana* seedlings, expressing CaMV TAV protein [12], grown on solid agar plates containing Murashige and Skoog (MS) medium, were carefully removed and washed in 10 mM MES (2-(N-morpholino)ethanesulfonic acid)) pH 5.7. Eighty seedlings were transferred into 7 mL of MES buffer containing 100 μM MG132 in dimethylsulphoxid (DMSO) (Selleckchem) or, for the mock treatment, DMSO only, infiltrated with these solutions under vacuum for 10 minutes and then incubated under gentle agitation at 21˚C. Ten seedlings were collected at the indicated time points, frozen in liquid nitrogen and ground in a Precellys® homogenizer at 6,000 rpm for 7 seconds. Proteins of the crude extracts were quantified with amido black and finally, analysed by western blotting using rabbit polyclonal antibodies against CaMV TAV at a dilution of 1:10,000 and RGA DELLA (PhytoAB), provided by Patrick Achard (IBMP, Strasbourg, France), at a dilution of 1:1,000, secondary goat anti-rabbit antibodies conjugated to HRP (Horse Radish Peroxydase) (Thermo Fischer Scientific, and

luminol-based enhanced chemiluminescence substrate (Lumi-Light^{Plus} Western Blotting Substrate, Roche).

Analysis of viral DNA from CaMV-infected plants by semi-quantitative (RT)-PCR

At 21 dpi, two discs (diameter 0.5 cm) sampled from CaMV or CaMV-TAVm3-infected turnip leaves were ground in liquid nitrogen in a Precellys® homogenizer at 6,000 rpm twice for 20 seconds. After DNA denaturation in 200 μL 0.5 N NaOH per plant crude extract, 5 μL were added to 45 μL 0.1 M Tris-HCl pH 8 for neutralization. PCRs were performed on 3 μL of the treated crude extract with couples of specific oligonucleotides to amplify sequences from CaMV DNA and the actin gene *act2* used as an internal control (S1 Table). PCRs were run on a T gradient thermocycler Biometra using the Go Taq®flexi DNA polymerase (Promega). The PCR products obtained after 22 to 30 cycles of amplification were analysed on a 1% agarose gel and revealed by ethidium bromide staining.

Detection of CaMV 19S and 35S RNAs, and virus-derived small RNAs

Total RNA from infected plant leaves was extracted using the RNeasy plant mini kit (Qiagen) and followed by a DNase I (Promega) treatment to eliminate residual DNA. RNA was precipitated with isopropanol and dissolved in 50% formamide. Northern blot analyses of low and high molecular weight RNA were performed with 10 and 5 μg of total RNA, respectively, as previously described [12]. CaMV 35S and 19S RNAs were detected using as probes DNA oligonucleotides end-labelled with [γ-^{32}P] ATP (3,000 Ci/mmole) complementary to ORF VI and ORF II, respectively. Virus-derived small RNAs (vsRNAs) were detected with [α-^{32}P] dCTP (800 Ci/mmole) radiolabelled probes resulting from random priming reactions on pMD324, which contains the CaMV full-length genome [12].

Biolistic experiments

CaMV TAV protein and the mutated versions fused to EGFP, and TAV fused to mRFP, were transiently expressed in BY-2 tobacco suspension cells (*Nicotiana tabacum* cv Bright Yellow 2). Cells were subcultured weekly and harvested 3 days after medium renewal for biolistic transfection. Cells were filtered onto Whatman disks. Particles preparation and bombardment assays were performed as follows: 2 mg of 1.1 μm tungsten particles (Bio-Rad, Hercules, CA) were immersed in 1 mL of absolute ethanol for 20 min. Dried particles were then successively mixed with 10 μg of recombinant plasmid DNA (pCK-EGFP-TAV or derivatives, or pmRFP-TAV) supplemented with 18% glycerol, 750 mM $CaCl_2$, and 90 mM spermidine in a final volume of 30 μL. The firing distance was 11 cm and the helium pressure 7 bars. After bombardment, cells were transferred to 0.8% agar MS media plates and incubated in the dark at 28˚C. Transfected BY-2 cells were collected under HBO binoculars (excitation/emission wavelength 488/505 to 545 nm) 20 h after bombardment and before further treatment and/or observations with a Zeiss LSM700 confocal microscope (Jena, Germany).

Immunofluorescence analysis

Tobacco BY-2 cells, transfected with constructions encoding TAV and its mutated versions fused to EGFP or mRFP, were observed between slide and cover slip with laser scanning confocal microscope (LSCM). EGFP and mRFP were detected after excitation at 488 nm and 568 nm with argon and HeNe laser, respectively, and using an appropriate emission filter to collect the signals from the optical section. Cells were observed 16 h and 24 h after bombardment.

For immunofluorescence localization studies, protoplasts were prepared from CaMV-infected turnip plants as described in [39] and fixed for 15 min under gentle shaking in protoplast medium containing 1% glutaraldehyde. Thereafter, they were washed three times with the protoplast medium, once with the medium diluted volume to volume with phosphate-buffered saline (PBS) and then again with PBS. A sample of protoplasts was mounted on poly-L-Lys–coated cover slip, allowed to settle for 1 h at room temperature, and then treated overnight at 4˚C in a 0.1% sodium borohydride solution. Protoplasts were incubated for 1 h in a blocking solution (5% acetylated bovine serum albumin [BSAc, Aurion, Wageningen, The Netherlands]), 5% normal goat serum, and 0.1% cold water fish skin gelatine prepared in PBS) and then overnight, with polyclonal anti-TAV, anti-P4 or anti-P2 antibodies, at a dilution of 1:500. After six washes with PBS containing 0.1% BSAc, protoplasts were treated with goat or mouse anti-rabbit IgG coupled to Alexa 488 or Alexa 568 respectively (Molecular Probes, Eugene, OR) at a dilution of 1:300 for 12 h. After several washes in PBS containing 0.1% BSAc, protoplasts were examined by LSCM.

Electron microscopy of CaMV-infected turnip cells

Infected leaves, cut into even strips (0.1 x 1 cm), were fixed overnight in 4% glutaraldehyde with 10% picric acid, then successively stained with 2% uranyl acetate for 2 h each, and with 0.1% (v/v) osmium tetroxide in 150 mM phosphate buffer pH 7.2. Samples were dehydrated through an ethanol series and infiltrated with EPON812 medium-grade resin (Polysciences). Polymerization was performed for 72 h at 60˚C. Ultrathin sections (90 nm) were cut using an ultracut E microtom (Reichert) and collected on grids coated with formvar (Electron Microscopy Sciences). Finally, samples were visualized with a Hitachi H-600 transmission electron microscope operating at 75 kV and viroplasms surface was measured with ImageJ software.

Results

Formation of large electron-dense inclusion bodies (EDIBs) involves three physically independent regions of TAV

Several functional domains of TAV were mapped (Fig 1): domain A, divided into A1 and A2 subdomains, and involved in host specificity, symptomatology and self-assembly; MAV (Mini-TAV) domain and RNA binding domain 1 (RNA1), also known as MBD (Multiple Binding Domain), are both required for reinitiation of translation and bind double-stranded RNA and RNA-DNA hybrids, respectively; RNA binding domain 2 (RNA2) which binds single-stranded RNA, and, finally, a zinc finger (Zn) domain. Previous studies performed in cultured tobacco BY-2 cells, non-host for CaMV, indicated that the EKI tripeptide located in TAV subdomain A1 of the Cabb B-JI isolate plays an essential role in the formation of EDIBs [9].

To further investigate the mechanism involved in viroplasm formation, we co-transfected tobacco BY-2 cells with recombinant plasmids encoding full-length TAV fused at its N-terminus to mRFP (mRFP-TAV), and two truncated TAV proteins, fused at their N-termini to EGFP (EGFP-A and EGFP-TAVΔA), to verify by competition assays if, in addition to domain A, other TAV sequences are involved in the formation of EDIBs (Fig 2A and 2B). Having already shown in a previous study that the fusion of EGFP did not modify the capacity of TAV to form EDIBs in BY-2 cells [9], we observed the cells using LSCM, 16 h and 24 h after transfection. mRFP-TAV protein assembled into large IBs when expressed alone (Fig 2B, panel 1), thus demonstrating that, as for EGFP, the fusion of mRFP at the N-terminus of TAV did not hinder TAV self-association. When mRFP-TAV was co-expressed with EGFP-TAV in BY-2 cells, the two TAV fusion proteins co-localized, as evidenced by the yellow fluorescent foci

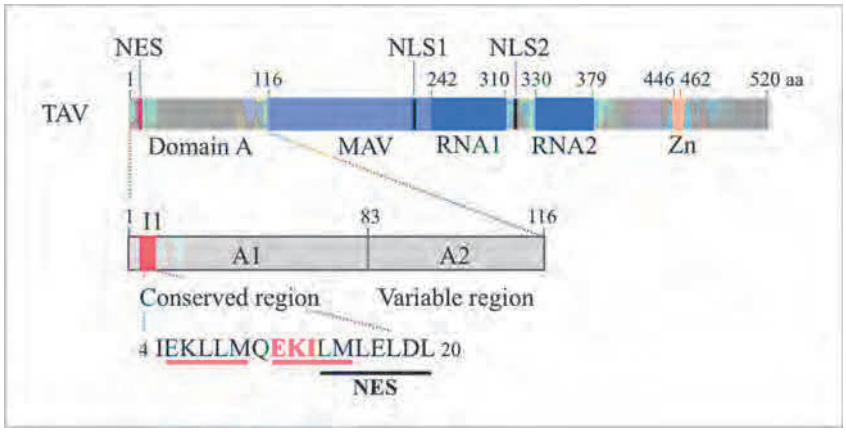


Fig 1. Schematic representation of CaMV TAV protein functional domains. Numbers above the boxes indicate the amino acids and delimitate the functional domains of TAV: domain A (grey box, amino acids 1–116); MAV (light blue box, amino acids 117–242); RNA1 (dark blue box, amino acids 243–310); RNA2 domain (second dark blue box, amino acids 330–379), and the zinc finger (orange vertical bar, amino acids 446–462). Positions of the single NES and the two NLSs are indicated with red and black vertical bars, respectively [9,12]. Below is detailed domain A with its subdomains, the conserved A1 region (amino acids 1–83) and the variable A2 region (amino acids 84–116) [9]. The invariant sequence I1 is represented with a dark red box and its sequence is given with the EKI motif in red letters; the motif is substituted by three alanines (AAA) in TAVm3 mutant. The tandem repeated sequence EKI/LLM and the NES are underlined in red and black, respectively.

https://doi.org/10.1371/journal.pone.0189062.g001

observed in the merge image (Fig 2B, panels 5–7). As previously reported [9], EGFP-A diffused in the cytoplasm (Fig 2B, panel 2) and EGFP-TAVΔA was retained in the nucleus (Fig 2B, panel 3). Their cellular localization is due to the presence of a NES (in domain A) and two NLSs (in the downstream region) that are involved in TAV nucleo-cytoplasmic shuttling (Fig 1). When mRFP-TAV was co-expressed in tobacco cells with EGFP-A (Fig 2B, panels 8 to 10; S1A Fig) or EGFP-TAVΔA (Fig 2B, panels 11–13; S1B Fig) or both (Fig 2B, panels 14–16; S1C Fig), it did not form any large IBs but diffused essentially in the cytoplasm, despite the nuclear accumulation of EGFP-TAVΔA. Few fluorescent foci were detected in some BY2-cells (Fig 2B, panels 11–16; S1B and S1C Fig) indicating that EGFP-A and/or EGFP-TAVΔA did not totally prevent mRFP-TAV self-assembly. The latter was probably overexpressed in these cells compared to EGFP-TAV deletion mutants, as evidenced by the weak green fluorescence of the foci (Fig 2B, panels 11 and 14; panels 1 in S1B and S1C Fig) and their orange colour in the merge images (Fig 2B, panels 13 and 16; panels 4 in S1B and S1C Fig). The capacity of EGFP-TAVΔA to interfere with mRFP-TAV self-association was surprising, since TAVΔA did not bind full-length TAV *in vitro*, in far western assays, contrary to domain A [9]. No TAV bodies were formed even when EGFP-A and EGFP-TAVΔA were expressed together in BY-2 cells in order to reconstitute *in trans* a complete TAV protein (Fig 2B, panel 4). Taken together, these results indicate that both TAV mutants out-competed the interactions between full-length TAV molecules, thus impairing the formation of EDIBs. This reinforced our hypothesis that one or several sequences of TAV, located downstream of domain A, are required for an efficient formation of IBs.

To identify these sequences, we analysed in tobacco BY-2 cells, the behaviour of EGFP-fused TAV mutants with C-terminal deletions (TAVΔC1: residues 1–242, TAVΔC2: residues 1–313 and TAVΔC3: residues 1–400) or functional domain deletions (TAVΔMAV, TAVΔRNA1, TAVΔRNA2 and TAVΔZn) (Fig 1, Fig 2C and 2D). For comparison, we also ectopically expressed EGFP-TAV (Fig 2D, panel 1) and EGFP-TAVm3 (Fig 2D, panel 2) in

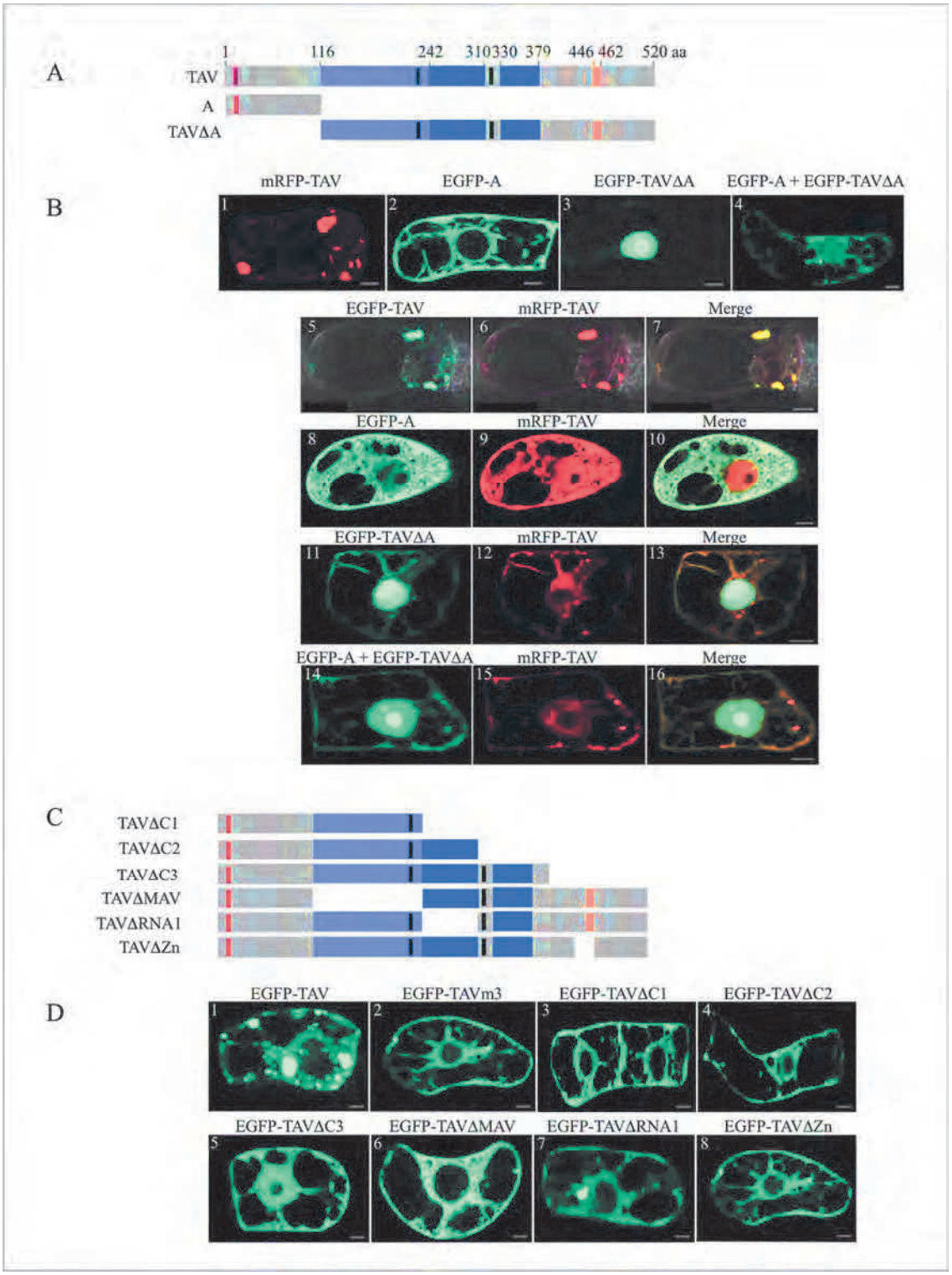


Fig 2. Characterization of CaMV TAV regions involved in the formation of inclusion bodies. (A, C) Schematic representation of TAV and TAV deletion mutants: A, TAVΔA, TAVΔC1, TAVΔC2, TAVΔC3, TAVΔMAV, TAVΔRNA1 and TAVΔZn. Numbers above the diagram of full-length TAV indicate the amino acids. The functional domains of TAV are depicted as in Fig 1. (B) Transient expression of mRFP-TAV (panel 1), EGFP-A (panel 2), EGFP-TAVΔA (panel 3), and EGFP-A and

EGFP-TAVΔA together (panel 4) in tobacco BY-2 cells, and competition assays performed in BY-2 cells co-transfected with plasmids encoding mRFP-TAV and EGFP-TAV (panels 5–7), EGFP-A (panels 8–10), EGFP-TAVΔA (panels 11–13) and both EGFP-A and EGFP-TAVΔA (panels 14–16), respectively. (D) Identification of TAV domains involved in the formation of inclusion bodies by transient expression of EGFP-TAV deletion mutants in BY-2 cells. Observations of TAV and TAV mutants fused to EGFP (B, D) or mRFP (B) were made 16 h after transfection, by LSCM. The LSCM settings and acquisition conditions of the images (single sections) were identical in all (B) and (D) panels. Scale bars: 10 μm.

https://doi.org/10.1371/journal.pone.0189062.g002

BY-2 cells. None of the C-terminally truncated TAV mutants that all possess the domain A generated inclusion bodies (Fig 2D, panels 3–5), confirming that additional domains are required for TAV self-assembly. The partial localization of EGFP-TAVΔC3 in the nucleoplasm, despite possessing both nuclear import and export signals, may be due to an unusual conformation of this TAV mutant, hiding the NES. Two regions of TAV involved in the formation of inclusion bodies could be unambiguously identified when we tested internally deleted TAV mutants in tobacco BY-2 cells: the MAV domain, where the relevant sequence could be restricted to positions 218 to 242, and the zinc finger-encompassing domain (positions 413–462). Their deletion abolished the formation of TAV bodies as evidenced by the diffused green fluorescence in the cytoplasm (Fig 2D, panels 6 and 8), whereas the deletion of RNA1 binding domain did not (Fig 2D, panel 7). The behaviour of EGFP-TAVΔRNA2 was uncertain since it formed small aggregates or showed a diffuse cytoplasmic localization depending on the assay (data not shown).

In conclusion, our results indicate that, in addition to the EKI motif, two other physically separated regions of TAV, MAV and the domain containing the zinc finger, are required for EDIB formation. They likely act collectively, since the deletion of each of these sequences is sufficient to impair this process.

TAVm3 is unable to form large EDIBs in turnip plants in a viral or non-viral context

In previous studies, we noticed that the CaMV Cabb B-JI isolate encoding TAVm3 is as infectious in *A. thaliana* (ecotype Columbia) [12] as in turnip plants (unpublished data), suggesting that either TAVm3 forms EDIBs in a viral context or, on the contrary, that EDIBs are dispensable for CaMV infection.

To verify these hypotheses, we first expressed EGFP-TAV and EGFP-TAVm3 in turnip and in *A. thaliana* (ecotype Columbia) in a non-viral context. As in tobacco BY-2 cells (Fig 2), epidermal cells from turnip leaves bombarded with the plasmid coding for EGFP-TAV displayed green fluorescent aggregates of different sizes, among which were large TAV inclusion bodies (8–10 μm^2) (Fig 3A, panel 1). On the other hand, EGFP-TAVm3 showed a diffused distribution in the cytosol of turnip cells transfected with the corresponding recombinant plasmid (Fig 3A, panel 2). Small fluorescent aggregates ($< 1\ \mu m^2$) were visible in some turnip cells, albeit to a lesser extent than with EGFP-TAV. In transient expression assays performed on *A. thaliana* protoplasts, EGFP-TAV and EGFP-TAVm3 fusion proteins both accumulated at the same levels (S2 Fig) and showed the same behaviour as in turnip and tobacco BY-2 cells (Fig 3B, panels 1 and 2) thus suggesting that TAV's EKI mutation (and not a lower accumulation of EGFP-TAVm3) was responsible for the absence of large EDIBs. To also ensure that TAV and TAVm3 aggregation capacities were not influenced by their fusion to EGFP in these plants, we took advantage of transgenic *A. thaliana* lines expressing TAV and TAVm3 at comparable levels [12]. Immunofluorescence experiments performed on protoplasts isolated from these transgenic plants revealed that TAV and TAVm3 behaved identically as when they were fused to EGFP (Fig 3C, panels 1 and 2), thus demonstrating that the inability of TAVm3 to self-assemble into large EDIBs was related to the EKI mutation *per se*.

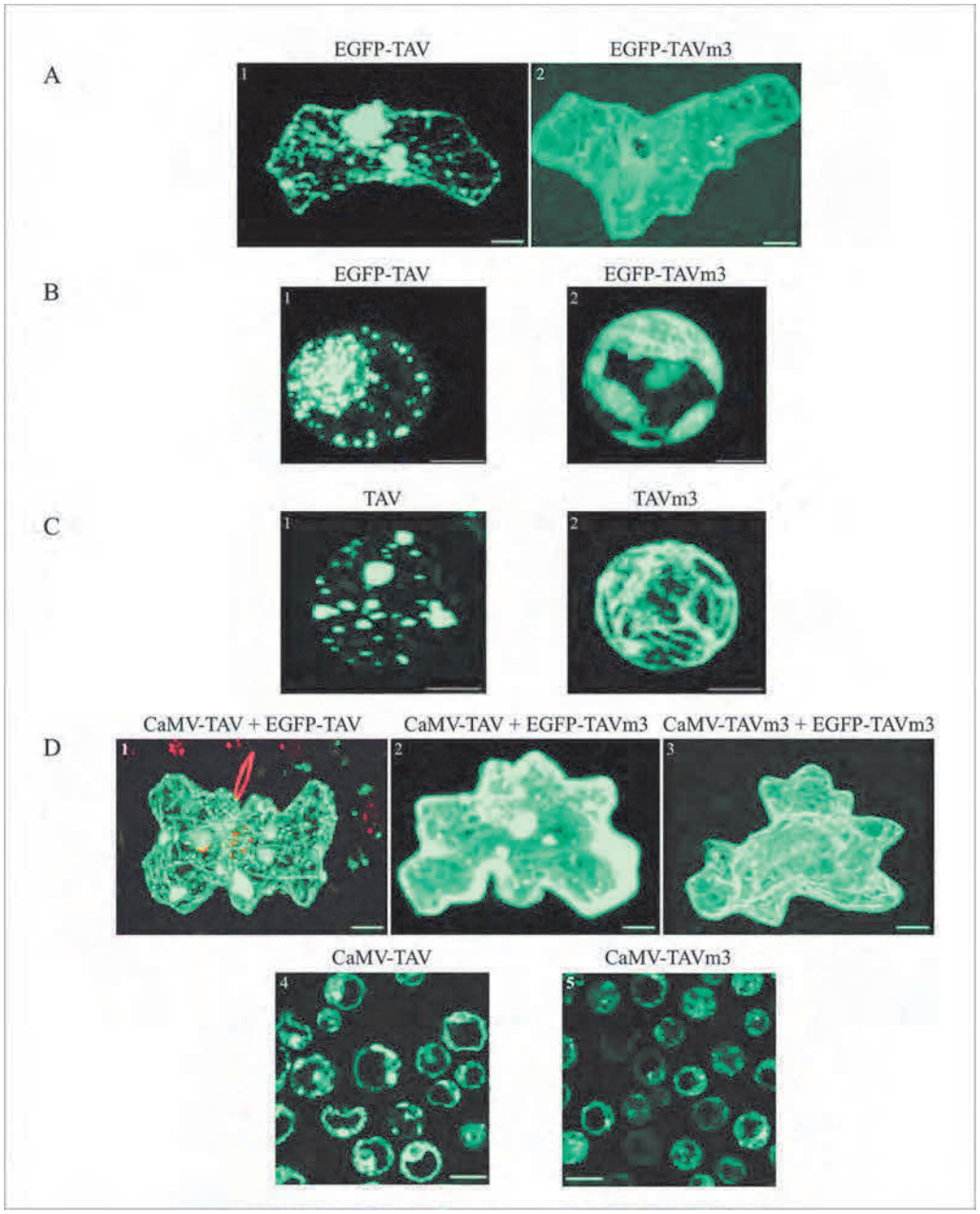


Fig 3. CaMV TAV and TAVm3 proteins expressed in turnip plants and *A. thaliana* in a viral or non-viral context. (A and B) Transient expression of EGFP-TAV and EGFP-TAVm3 in turnip epidermal cells bombarded with the corresponding pCK-EGFP recombinant plasmids (A, panels 1 and 2) and in *A. thaliana* (ecotype Columbia) protoplasts transfected by PEG with same constructions (B, panels 1 and 2). Cells were observed 16 h after transfection, by LSCM. (C) Immunolocalisation of TAV and TAVm3 (panels 1 and 2) in protoplasts obtained from transgenic *A. thaliana* (ecotype Columbia) lines expressing wild-type or TAVm3 [12]. Protoplasts were incubated with polyclonal anti-TAV antibodies [38] and with anti-rabbit IgG secondary antibodies conjugated to green fluorochrome Alexa 488. (D) Detection of TAV proteins by EGFP fluorescence, 21 dpi, in turnip epidermal cells from systemic leaves infected with wild-type CaMV (panels 1 and 2) or CaMV-TAVm3 (panel 3) through interactions with ectopic EGFP-TAV (panel 1) and EGFP-TAVm3 (panels 2 and 3) expressed upon bombardment with the corresponding plasmids. Immunofluorescence detection of TAV and TAVm3 in protoplasts isolated from CaMV-TAV and CaMV-TAVm3-infected turnip leaves (panels 4 and 5). The LSCM settings and acquisition conditions of the images were identical in all panels. Images in (A), (B), (C) and (D) (panels 1–3) are projections; those in (D) (panels 4–5) show single sections. Scale bars (A and D): 10 μm, (B and C): 5 μm.

https://doi.org/10.1371/journal.pone.0189062.g003

Next, turnip plants were mechanically inoculated with linear CaMV DNA, obtained by *Sal*I digestion of viral vectors [12], and encoding wild-type TAV (CaMV-TAV) or TAVm3 (CaMV-TAVm3), in order to determine the behaviour of TAV proteins in a viral context. Systemically infected turnip leaves were analysed to ensure that most of the inclusions were fully developed [32]. CaMV-TAV and CaMV-TAVm3-infected leaves were bombarded with plasmids encoding EGFP-TAVm3 or EGFP-TAV used as control, to detect TAV proteins produced in the course of CaMV replication cycle through interactions with their fluorescent counterpart (Fig 3D, panels 1–3). Ectopic EGFP-TAVm3 revealed in epidermal turnip cells infected with wild-type CaMV that TAV formed numerous cytosolic aggregates, highly variable in size (1–10 μm^2), which were either free or appeared to be associated with the cytoskeleton (Fig 3D, panel 2), as reported in a previous study [26]. The fluorescent aggregates resulted from the interaction between EGFP-TAVm3 and TAV inclusions formed during the viral cycle, since EGFP-TAVm3 alone is unable to self-assemble in turnip cells (Fig 3A, panel 2). Ectopic EGFP-TAV in turnip cells infected with wild-type CaMV labelled large cytoplasmic aggregates and decorated the cytoskeleton network (Fig 3D, panel 1). On the other hand, EGFP-TAVm3 diffused in the cytoplasm or revealed small aggregates ($< 1\ \mu m^2$) (Fig 3D, panel 3), indicating that TAVm3 was unable, as in a non-viral context, to self-assemble into large IBs during an authentic viral infection. The observation of a population of protoplasts isolated from CaMV-TAVm3-infected systemic turnip leaves (Fig 3D, panel 5) and treated with anti-TAV antibodies and secondary antibodies coupled to Alexa 488, revealed that in fact only small cytosolic TAVm3 inclusion bodies were present in protoplasts, often along the plasma membrane, while those infected with wild-type CaMV systematically contained large EDIBs (Fig 3D, panel 4). Sequence analysis showed that CaMV progeny DNA maintained the EKI mutation in the tested plants after several serial passages of CaMV-TAVm3, indicating that EKI is not essential for CaMV infectivity, while the opposite was predicted since it is highly conserved among all CaMV isolates.

Taken together, our results show that the TAVm3 mutant, in *Brassicaceae*, is also unable to self-assemble into large IBs in a viral context, strongly suggesting that large EDIBs are dispensable for systemic CaMV infection of host plants.

TAV's EKI motif is important for CaMV Cabb B-JI virulence on turnip plants

The finding that the CaMV-TAVm3 mutant systemically infected *Brassicaceae* prompted us to investigate, in turnip, the biological relevance of EKI for CaMV pathogenesis.

Turnip plants inoculated with linearized CaMV-TAVm3 DNA were systemically infected 21 dpi, without any delay compared to plants inoculated with wild-type CaMV DNA. Sequence analysis of ORF VI after several rounds of virus replication or successive inoculations (> 10) of host plants with an infectious sap, showed that CaMV did not restore the EKI motif. Turnip plants infected with wild-type CaMV displayed severe symptoms (mosaic, vein clearing and stunting of plant and leaves) (Fig 4A, panels 1 and 2) whereas CaMV-TAVm3-infected plants showed very mild symptoms (faint chlorosis, no plant stunting) (Fig 4A, panels 3 and 4) and were thus hardly distinguishable from mock-inoculated plants (Fig 4A, panels 5 and 6). CaMV presence in systemic turnip leaves was confirmed by detection of viral DNA by PCR and viral proteins by western blot (data not shown). Disease symptoms became no more pronounced over time with CaMV-TAVm3, in contrast to those elicited by the wild-type virus.

Electron-microscope observations of cellular ultrastructure were made on CaMV-infected systemic turnip leaves exhibiting a disease phenotype, to precisely assess the morphology of

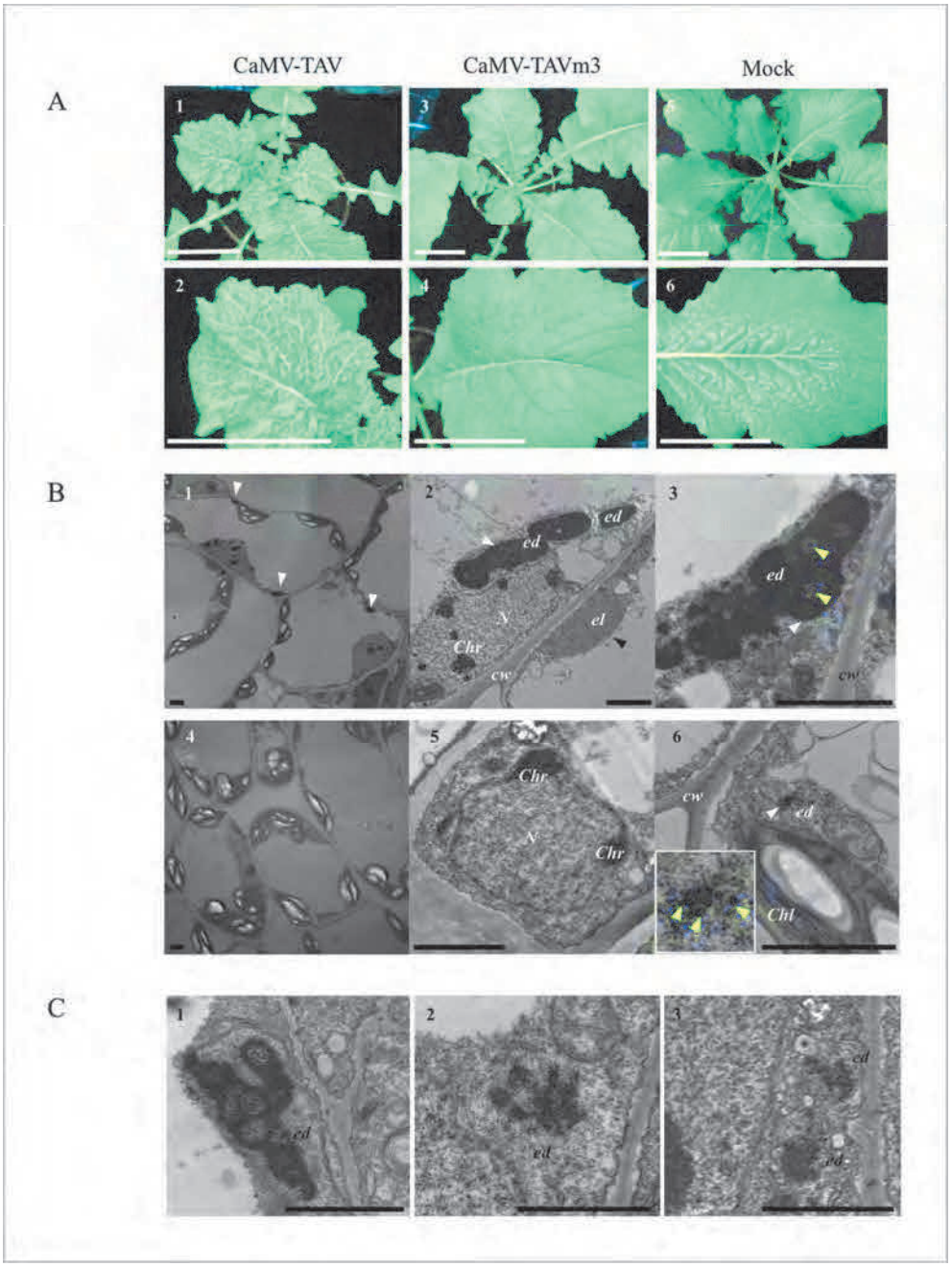


Fig 4. Phenotype of turnip and *A. thaliana* plants infected with wild-type CaMV or CaMV-TAVm3 mutant. (A) Symptoms expressed by turnip plants and systemic leaves infected with CaMV-TAV (panels 1 and 2) or CaMV-TAVm3 (panels 3 and 4) 21 dpi. A mock-inoculated turnip plant and leaf are shown (panels 5 and 6). Scale bars: 6 cm. (B) Photographs of turnip cells infected with CaMV (panels 1–3) or CaMV-TAVm3 (panels 4–6) observed by electron microscopy. Electron-dense viroplasms/aggregates (**ed**) and electron-lucent viroplasms (**el**) are pointed by white and black arrowheads, respectively. Panel 6: enlargement of the electron-dense body (**ed**). Yellow arrowheads indicate some CaMV particles. (C) Electron microscopy photographs of *Arabidopsis* cells from wild-type CaMV-infected plants (panel 1), CaMV-TAVm3-infected plants (panel 2) and a transgenic line expressing TAVm3 (panel 3). **Chl**: chloroplast; **Chr**: chromatin; **cw**: cell wall; **N**: nucleus. Scale bars (B-C): 2 μm.

https://doi.org/10.1371/journal.pone.0189062.g004

TAVm3-bodies. Photographs encompassing approximately 300 cells from four independent infection assays, were taken, all showing very similar results. Fig 4B shows 6 typical representative images. Large perinuclear EDIBs with vacuolar-like areas full of virions and a single ELIB, formed by P2 [40], were found in turnip cells infected with wild-type CaMV (Fig 4B, panels 1–3). Cells from turnip plants infected with CaMV-TAVm3 never contained such EDIBs, but in some cells we observed electron-dense areas certainly corresponding to the small inclusion bodies observed by LSCM (Fig 4B, panels 4–6). They also contained much less virions than wild-type CaMV-infected cells. Magnification of the electron-dense areas in CaMV-TAVm3-infected cells showed that they are irregularly shaped and granular compared to authentic EDIBs, suggesting that they may result from clumping of virus particles. Morphologically similar TAVm3 bodies were also observed in Arabidopsis plants infected with CaMV-TAVm3, as well as in transgenic Arabidopsis lines expressing TAVm3 (Fig 4C). Unexpectedly, CaMV-TAVm3-infected plants did not contain any ELIB, while this type of IB is always formed in CaMV-infected cells. To confirm this, protoplasts from turnip plants infected with CaMV or CaMV-TAVm3, were analysed by EGFP autofluorescence and anti-P2 immunofluorescence to detect both EDIBs and ELIB. CaMV-TAV- or CaMV-TAVm3-infected protoplasts were first transfected with the plasmid encoding EGFP-TAVm3 to reveal EDIBs (Fig 5A, panel 1, and 5B, panel 1). Then, protoplasts were fixed and treated with anti-P2 antibodies and Alexa 568-coupled secondary antibodies to detect ELIBs (Fig 5A, panel 2, and 5B, panel 2). Large EDIBs and a single large ELIB were clearly visible in protoplasts infected with wild-type CaMV (Fig 5A, panel 3; S3 Fig). By contrast, the P2 protein was dispersed throughout the cell, forming small aggregates ($< 1\ \mu m^2$) physically separated from those generated by TAVm3 (Fig 5B, panel 3; S4 Fig), thus confirming that ELIB was not formed when the EKI motif was mutated, despite the fact that P2 was expressed at a significant level (see below), and suggesting that native TAV and/or EDIBs are mandatory for the formation of the ELIB which serves as platform for aphid-transmission of CaMV [29].

In conclusion, mutating EKI dramatically affects the virulence of the CaMV Cabb B-JI isolate on turnip plants, as evidenced by (i) the mild phenotype exhibited by infected plants, and (ii) the few CaMV particles produced in the course of infection. Neither large EDIBs nor ELIB were formed in turnip plants infected with the CaMV-TAVm3 mutant.

Large EDIBs are dispensable for CaMV infectivity but enhance the production of virus progeny

Electron microscopy observations revealed that turnip plants infected with CaMV-TAVm3 contained less virus particles than those infected with wild-type CaMV. Most CaMV particles in plants infected with CaMV-TAVm3 were essentially clustered in electron-dense areas, while virions were concentrated in the matrix of EDIBs in plants infected with wild-type CaMV (Fig 4); it was estimated that EDIBs contain 95% of the virus particles [41]. We could also indirectly confirm this observation by immunodetection of P4, the capsid proteins precursor (pre-CP), since wild-type CaMV-infected protoplasts displayed a cytosolic green fluorescence superimposed by a large green oval-shaped area corresponding to a perinuclear EDIB (Fig 6A, panel 1) while protoplasts infected with CaMV-TAVm3 exhibited fluorescence restricted to small areas (Fig 6A, panel 2). To definitely confirm that CaMV-TAVm3 mutant was produced in lower amounts than wild-type CaMV, virus particles from infected turnip plants 21 dpi were purified on a sucrose density gradient and quantified by spectrophotometry as shown in Fig 6B. Similar amounts of infected plant crude extracts, evidenced by the detection of the almost same quantities of RbcL (55 kDa) (Fig 6C, lanes 1 and 4), were fractionated on sucrose gradients. Capsid proteins (42, 39 and 37 kDa) of purified wild-type viral particles

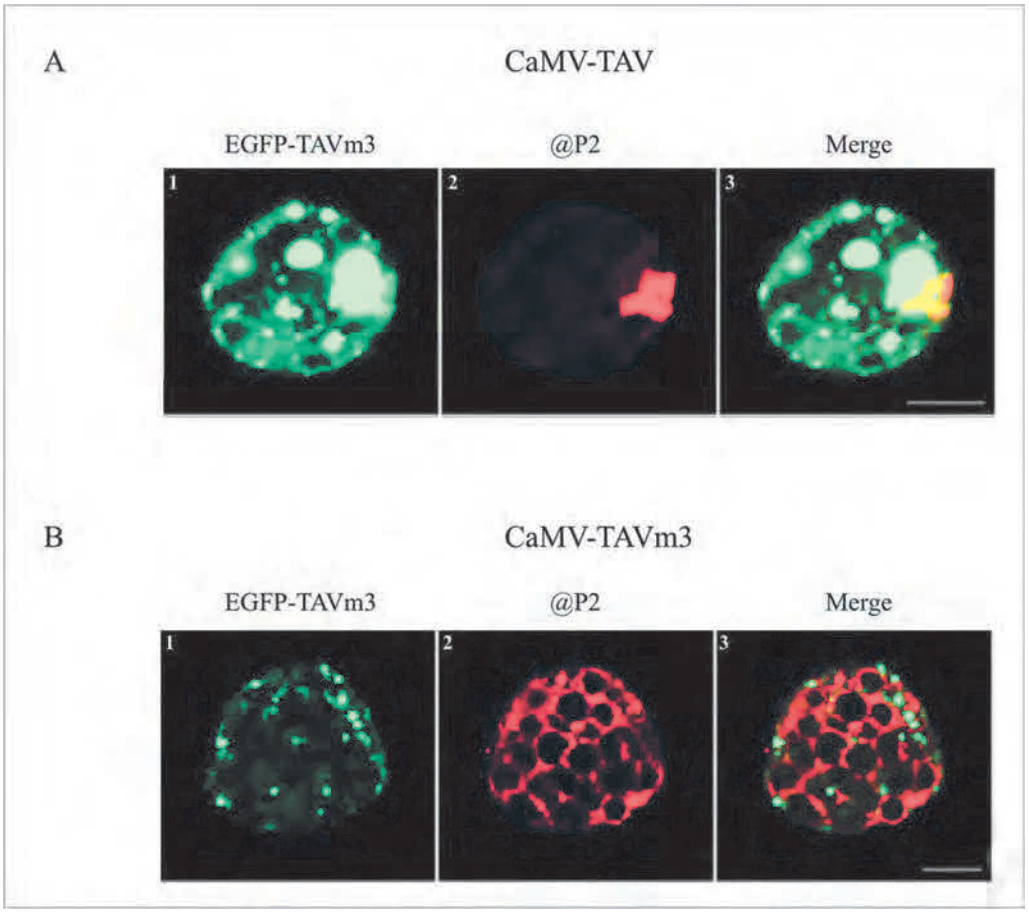


Fig 5. Immunolocalisation of electron-lucent inclusion bodies in protoplasts prepared from turnip plants systemically infected with wild-type CaMV or CaMV-TAVm3. Electron-dense inclusion bodies (EDIBs) were visualized in protoplasts infected with wild-type CaMV (A) or CaMV-TAVm3 (B) by transient expression of EGFP-TAVm3 (A, panel 1 and B, panel 1). Electron-lucent inclusion body/aggregates (ELIB) generated by CaMV P2 protein were observed by LSCM, after treatment of the infected and transfected protoplasts with polyclonal anti-P2 protein and secondary Alexa 568-conjugated antibodies (A, panel 2 and B, panel 2). Merge images showing the cellular distribution of EDIBs and ELIB, and TAVm3 and P2 bodies are presented in A, panel 3, and B, panel 3. All panels show single sections. Scale bars: 5 μm.

https://doi.org/10.1371/journal.pone.0189062.g005

were easily visualized after SDS-PAGE fractionation by Coomassie blue staining (Fig 6C, lane 2) and with anti-P4 antibodies (Fig 6C, panel 3), in contrast to purified CaMV-TAVm3 particles, for which larger aliquots were needed for their detection by staining (Fig 6C, lane 5) or immunodetection (Fig 6C, panel 6). Virus quantification indicated that plants infected with wild-type CaMV contained at least 50 times more particles than those infected with CaMV-TAVm3, *i.e.* 0.43 mg versus 0.008 mg of virions per gram of fresh turnip leaves.

The low number of CaMV-TAVm3 virus particles prompted us to compare the viral protein levels in plants infected with wild-type CaMV and CaMV-TAVm3 to determine whether their synthesis was affected in presence of TAVm3. Western blots were performed on total proteins extracted from leaf discs of infected turnip plants 21 dpi. The results were reproducible independently of the infection assay, as shown in Fig 7A. Overall, CaMV-TAVm3-infected turnip plants contained significantly lower amounts of viral proteins compared to plants infected by the wild-type virus (Fig 7A), and in particular TAVm3, whose quantity was

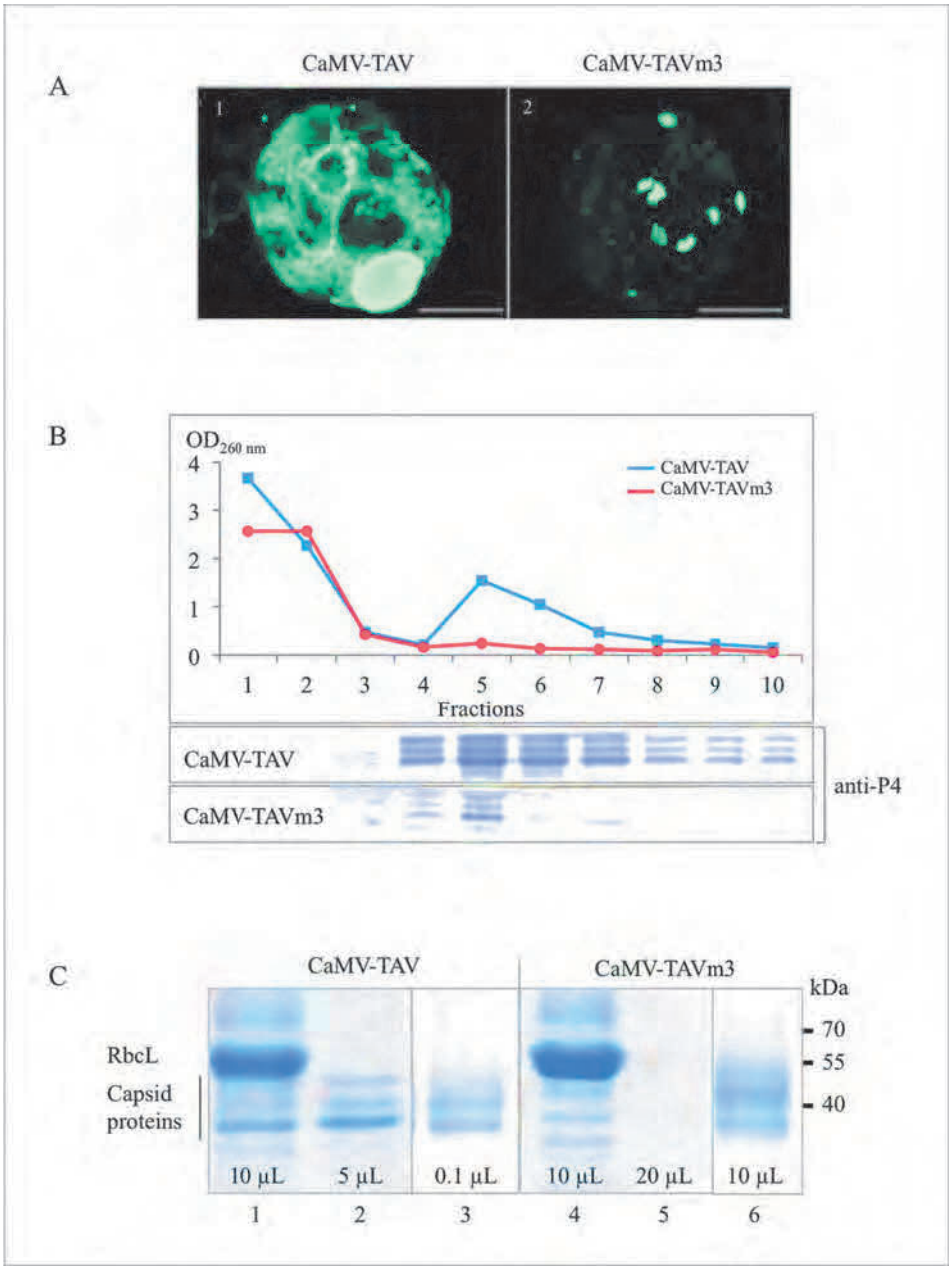


Fig 6. Purification of wild-type CaMV and CaMV-TAVm3 particles from infected turnip plants. (A) Immunolocalisation of CaMV and CaMV-TAVm3 capsids in protoplasts isolated from infected plants using polyclonal anti-P4 (pre-CP) antibodies and Alexa 488-coupled secondary IgG. Protoplasts were observed by LSCM. Images show single sections. (B) Purification of CaMV particles on sucrose density gradient: wild-type CaMV (blue line and squares) and CaMV-TAVm3 (red line and circles) were detected by measurement of OD_{260nm} of each 500 µL-fraction. Fractions 3–10 were tested for the presence of capsid proteins by immunodetection using anti-P4 antibodies. Peak fractions 4–8 corresponding to purified CaMV or CaMV-TAVm3, were pooled, concentrated and analyzed. (C) Analysis of CaMV-TAV (lanes 1 to 3) and CaMV-TAVm3 (lanes 4 to 6) particles (purified simultaneously and under strictly identical conditions) by 12% SDS-PAGE fractionation. Proteins in crude extracts before the gradient purification (aliquots of 10 µL) were revealed by Coomassie blue staining (lanes 1 and 4). Purified viruses from peak fraction 5 of the gradient (aliquots of 5 µL and 20 µL, respectively) were also analyzed by staining (lanes 2 and 5). Purified viruses from the pooled and concentrated peak fractions (aliquots of 0.1 µL and 10 µL, respectively) were immunodetected (lanes 3 and 6) using anti-P4 antibodies. RbcL and CaMV capsid proteins, and molecular markers are indicated at the left and right of panel C, respectively.

https://doi.org/10.1371/journal.pone.0189062.g006

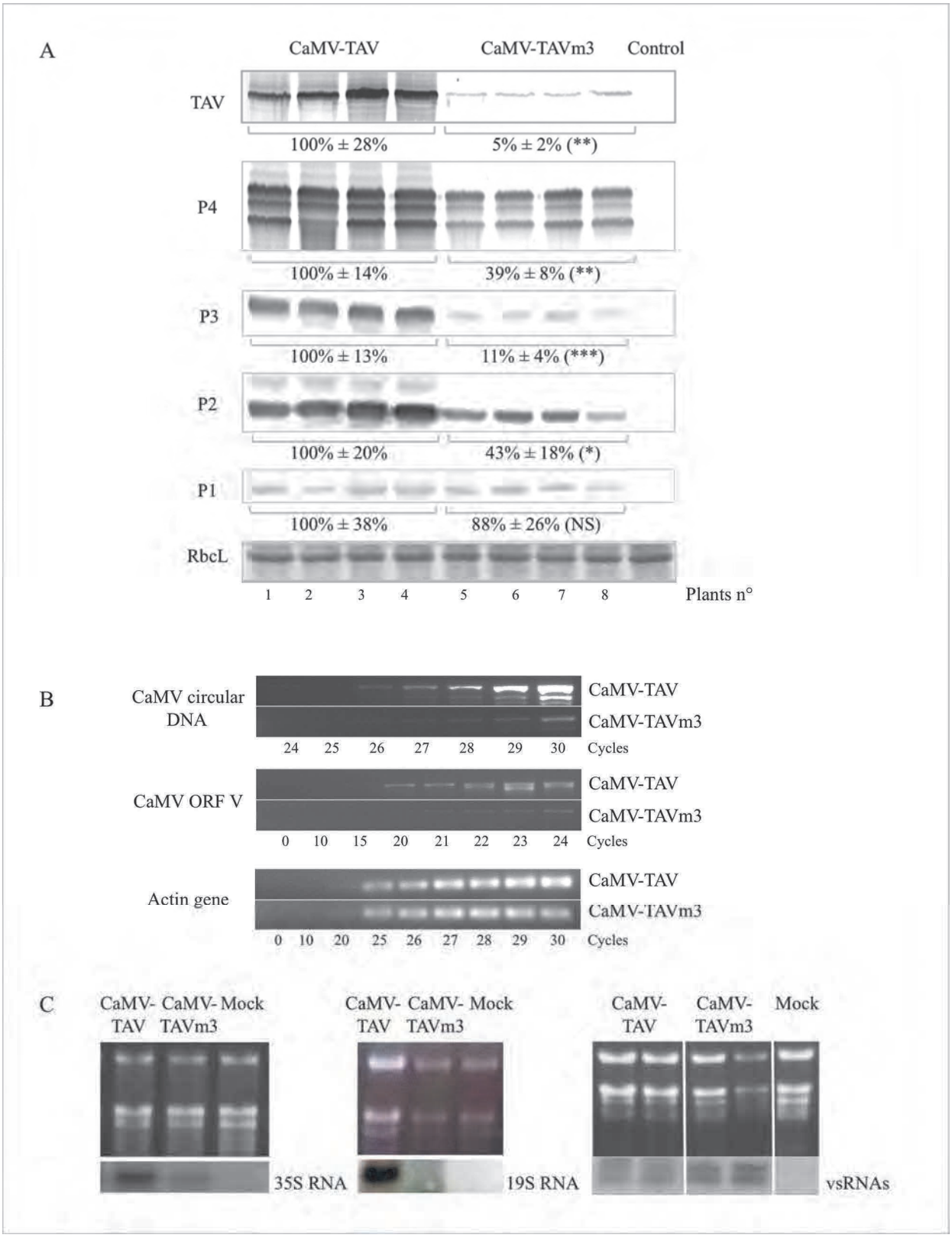

A
CaMV-TAV
CaMV-TAVm3
Control
TAV
100% ± 28%
5% ± 2% (**)
P4
100% ± 14%
39% ± 8% (**)
P3
100% ± 13%
11% ± 4% (***)
P2
100% ± 20%
43% ± 18% (*)
P1
100% ± 38%
88% ± 26% (NS)
RbcL
1 2 3 4 5 6 7 8
Plants n°
B
CaMV circular DNA
CaMV-TAV
CaMV-TAVm3
24 25 26 27 28 29 30 Cycles
CaMV ORF V
CaMV-TAV
CaMV-TAVm3
0 10 15 20 21 22 23 24 Cycles
Actin gene
CaMV-TAV
CaMV-TAVm3
0 10 20 25 26 27 28 29 30 Cycles
C
CaMV-TAV
CaMV-TAVm3
Mock
35S RNA
CaMV-TAV
CaMV-TAVm3
Mock
19S RNA
CaMV-TAV
CaMV-TAVm3
Mock
vsRNAs

Fig 7. Analysis of CaMV proteins, DNA, 35S and 19S RNAs, and vsRNAs produced in turnip plants infected with wild-type CaMV or CaMV-TAVm3. (A) Immunodetection by western blot of CaMV's TAV and TAVm3, P4 and the derived processed capsid proteins, P3, P2 and P1 proteins, using specific polyclonal antibodies and secondary antibodies coupled to alkaline phosphatase. Four independent infected plants were tested for CaMV-TAV (lanes 1–4) and CaMV-TAVm3 (lanes 5–8). Mock-inoculated plants were used as control. The loading control is RubisCO large subunit (RbcL) revealed by Coomassie blue staining. The bands corresponding to each of the 5 viral proteins were quantified as indicated in "Materials and Methods" and their corresponding values indicated, those of the viral proteins expressed in wild-type CaMV-infected plants being set at 100% (NS (not significant): p>0,5; * p<0.01; ** p<0.001; *** p<0.0001) (B) Semi-quantitative PCR performed on CaMV DNA from crude extracts of CaMV-TAV and CaMV-TAVm3-infected turnip plants using two couples of appropriate primers, which permit to only amplify a DNA sequence on the circular genome, and ORF V, respectively. The internal control was the actin gene *act2* amplified in the same crude extracts, by PCR using specific primers (S1 Table). The number of PCR cycles is indicated. (C) Northern blot performed on fractionated total RNA from plants infected by CaMV-TAV and CaMV-TAVm3, to detect the 35S and 19S RNAs (left and middle panels) and viral-derived small RNAs (vsRNAs) (right panel) using ^{32}P-labeled oligonucleotides. RNAs from mock-inoculated plants were also probed.

https://doi.org/10.1371/journal.pone.0189062.g007

estimated to represent only 5% of that of wild-type TAV. Treatment or not of seedlings of transgenic *Arabidopsis* expressing TAVm3 with the 26S proteasome inhibitor MG132 showed that the steady state level of TAVm3 remained unchanged throughout 24 hours, indicating that the low amounts of TAVm3 in CaMV-infected plants were not due to an instability of TAVm3 but rather to its reduced expression. The efficiency of the MG132 treatment was evidenced by the accumulation of the cellular protein RGA-DELLA in the presence of MG132, while this protein was barely detectable in mock-treated *Arabidopsis* seedlings (S5 Fig). Surprisingly, capsid proteins were present in plants infected with CaMV-TAVm3 at 39% of the levels of P4 detected in wild-type CaMV-infected plants, while we expected to detect minute amounts since CaMV-TAVm3 infected plants contained much less virus particles than wild-type CaMV-infected plants. We also noticed that the protein composition of the capsids (42, 37 and 35 kDa) was identical for CaMV and CaMV-TAVm3, indicating that the capsids were properly assembled and the pre-CP properly processed by the viral protease [42,43] in the cells infected by the mutant virus. For an unexplained reason, P3, another structural component of CaMV capsids, was found in very low amounts (11%) in CaMV-TAVm3-infected plants. The P2 protein was, on the contrary, produced in relatively high amounts (43%) in CaMV-TAV-m3-infected turnip plants (Fig 7A), thus indicating that its inability to form ELIBs was due to events independent of its expression level. Finally, the movement protein P1 was expressed at almost similar levels (100% - 88%) in both CaMV- and CaMV-TAVm3-infected turnip plants partly explaining why the systemic propagation of infection occurred at the same rate for CaMV and CaMV-TAVm3.

These results suggest that the EKI mutation in TAV and/or EDIBs has a strong effect on the expression of TAVm3 protein itself but also of P3, and a rather weak or moderate effect on the expression of P1, P2 and P4, and on capsid assembly. However, despite the fact that the expression and processing of the capsid protein precursor remains almost unaffected, TAVm3 inability to form large EDIBs seems to strongly reduce the yield of CaMV progeny. Therefore, we favour the hypothesis that this poor yield in progeny is due to the instability of the CaMV particles.

Replication of CaMV DNA by reverse transcription is strongly reduced in plants infected by CaMV-TAVm3 mutant

The discrepancy between the low yield of CaMV-TAVm3 progeny and the relatively high levels of capsid proteins in infected turnip plants strongly suggested that capsids might disassemble because they do not contain viral DNA. To verify this hypothesis, we performed semi-quantitative PCR that indicates DNA replication efficiency, on crude turnip extracts using a set of primers which anneal to the sequence that encodes the 5' part of the 35S RNA leader region (forward primer), and to the ORF II encoding sequence (reverse primer), respectively,

thus allowing the amplification of the circular viral genome [10] (S1 Table). PCR product analysis on ethidium bromide-stained agarose gel showed that three additional amplification cycles were required for the extract from CaMV-TAVm3-infected plants in order to obtain the same amounts of viral DNA than these amplified from extracts of wild-type CaMV-infected turnip plants (Fig 7B), while the actin gene *act2* used as an internal control was amplified following the same kinetic in both types of infected plants (Fig 7B). The use of another set of primers amplifying ORF V (S1 Table) led to the same conclusion (Fig 7B). Densitometry analyses, performed with the ImageJ software, indicate that plants infected with CaMV-TAVm3 only contained about 12% of the viral DNA amounts synthesized in turnip plants infected with wild-type CaMV, and thus, that viral DNA replication was strongly impaired in the presence of TAVm3.

Northern blots performed on total RNA using specific probes to detect CaMV transcripts showed that turnip plants infected with CaMV-TAVm3 contained approximately half the amount of 35S RNA compared to those infected with wild-type CaMV (Fig 7C, left panel). The 19S RNA coding for TAV was undetectable in CaMV-TAVm3-infected turnip plants, as opposed to those infected with CaMV-TAV (Fig 7C, middle panel). This suggests that the EKI mutation indirectly affected the transcription rate of the CaMV genome and/or the stability of viral RNAs. Analysis of virus-derived small RNAs (vsRNAs) revealed that plants infected with mutant CaMV contained slightly higher levels than in the context of an infection with the wild-type virus. Consequently, 35S and 19S RNAs were likely more susceptible to RNA silencing, despite the fact that the RNA silencing suppressor activity was not impaired in TAVm3 [12,44] (Fig 7C, right panel).

Taken together, our results confirmed that large electron-dense viroplasms are sites of intensive CaMV DNA replication by reverse transcription [45,46], where the 35S RNA is protected from degradation and capsid assembly is enhanced. However, in stark contrast with previous observations, our data suggest that large EDIBs are dispensable for the production of infectious CaMV particles, but we cannot rule out the possibility of the requirement for small TAV aggregates in the infection process.

Discussion

The hallmark of CaMV infection is the formation of EDIBs [7], which are considered as virus factories. The size of these inclusion bodies is extremely variable and depends on the virus isolate, the largest being observed with the virulent Cabb B-JI isolate [32]. Previously, we showed that mutating the EKI sequence located at the N-terminus of the TAV protein (mutant TAVm3), the major component of EDIBs impaired this process in a non-viral context. In this paper, we describe the continuation of our investigation on EDIB formation and the analyses of their biological relevance for CaMV Cabb B-JI pathogenesis, by studying the TAVm3 CaMV mutant.

Testing the behaviour of TAV mutants in tobacco BY-2 cells revealed that three independent sequences of TAV are involved in the formation of inclusion bodies: the N-terminal EKI motif (amino acids 11–13), the C-terminal region of the MAV domain (amino acids 218–242) and a sequence encompassing the zinc finger (amino acids 413–462). Deletion of one of these sequences is sufficient to abolish the formation of large EDIBs, indicating that they participate collectively in this process. They roughly correspond to three of the four domains (D1, D2, D3 and D4) of the TAV protein of CaMV CM1841 isolate shown to interact with full-length TAV in yeast two-hybrid analyses: D1 (amino acids 1–110), D2 (amino acids 156–253) and D4 (amino acids 414–520) [47]. Mutation of the TAV EKI motif appeared to have a similar effect to deletion or point mutations of subdomain D3b (amino acids 309–343) [33] since both,

Cabb B-JI TAVm3 and CM1841 TAV mutants, form small EDIBs in a non-viral context compared to the wild-type TAV proteins. Overlapping of domain D3b and RNA binding domain 2 (amino acids 330 to 379) [33] might partly explain the equivocal behaviour of CaMV Cabb-JI TAVΔRNA2 mutant in forming EDIBs or not in tobacco cells, depending on the assay.

In this study, we demonstrated that mutation of the EKI motif impairs the formation of large perinuclear EDIBs in turnip cells in a viral context, in Arabidopsis protoplasts expressing TAVm3 ectopically, and in protoplasts from TAVm3 transgenic Arabidopsis plants. Small TAVm3 inclusion bodies were formed in some cells, including CaMV-TAVm3-infected cells, whose formation is probably mediated by the two other interacting domains. This also suggests that the EKI mutation was not rescued by the upstream similar EKL tripeptide (Fig 1) in triggering TAVm3 self-assembly and the subsequent coalescence of small bodies to form large EDIBs.

The EKI sequence is conserved in all CaMV isolates (*i.e.* Cabb B-JI, D/H, D4, CM1841. . .) and several *Caulimoviruses*, such as *Dahlia mosaic virus* (DaMV) [48] and *Figwort mosaic virus* [49]. The TAV N-terminal region (domain A or D1) strongly interacts with the full-length protein *in vitro* [9] and *in vivo*, when tested in yeast two-hybrid assays [47]. This interaction is drastically reduced *in vitro* when EKI is mutated (data not shown), and *in vivo* when the glutamic acid and leucine at positions 312 and 316 of subdomain D3b are mutated in full-length TAV [33]. We believe that the EKI sequence interacts with subdomain D3b - it is conserved in many CaMV isolates, including Cabb B-JI—and that this potentiates the interactions between TAV molecules and/or TAV aggregates and, finally, the formation of large EDIBs. This hypothesis is supported by the observation that TAVm3 inclusion bodies never fused together to become large EDIBs, as is observed in wild-type CaMV-infected plants. Whether the EKI motif mediates the movement of TAV inclusion bodies along the cytoskeleton to reach the perinuclear region where the nucleation presumably occurs [2,26,50] remains an open question. Other experiments and a three-dimensional structure of the TAV protein are clearly needed to fully elucidate the mechanism leading to large EDIB formation in CaMV-infected cells, i.e. to determine whether sumoylation of TAV—two consensus sequences ΨKxD/E are present in TAV (M. Bureau, personal communication)–is required, as was shown to be the case for the generation of rotavirus viroplasms [51]. Deeper knowledge of this mechanism will certainly help to explain why EDIBs produced by some CaMV isolates remain small and/or do not fuse into larger viroplasms [32].

The CaMV Cabb B-JI isolate expressing TAVm3 systemically infects turnip plants but loses its virulence, as evidenced by the mild leaf symptoms (faint chlorosis) and the normal development of infected plants. Transgenic *Arabidopsis* lines (ecotype Columbia) encoding TAVm3 [12] also displayed a mild phenotype, resembling healthy plants, whereas lines expressing wild-type TAV showed severe leaf chlorosis and stunting, indicating that the EKI motif is an essential determinant in symptom expression and severity. Similarly to CaMV-TAVm3, CM1841 isolate harbouring mutations in the D3b domain also induced milder symptoms in turnip plants compared to those infected by the wild-type virus [33]. Studies performed with chimeric CaMV isolates, obtained by combining mild and severe CaMV isolates, have already demonstrated that the 5' part of ORF VI, encoding TAV, is responsible for disease severity [16,52,53]. The domain responsible for symptomatology, at least in *Arabidopsis*, was precisely mapped to the distal region of domain D1 (amino acids 40–110) [44], thus excluding the EKI motif. In fact, since the C-terminal part of D1 is also involved in the RNA silencing suppression activity of TAV [44], this activity may contribute to the expression of symptoms in CaMV-infected plants by modifying the expression pattern of cellular genes as observed in transgenic *Arabidopsis* plants expressing TAV [54,55]. Concerning the N-terminal part of domain D1 (amino acids 1–20) that contains the EKI motif, it is required for the suppression

of salicylic acid response gene expression [44]. Whether EKI plays a role in this process was not investigated in this study.

In summary, it appears that the EKI sequence is indirectly involved in symptom expression, probably by mediating the formation of large inclusion bodies, since none of TAV functions required for CaMV infectivity, including translation transactivation [11], nuclear import and export [12] and suppression of RNA silencing [44], is impaired when the EKI motif is mutated. Indeed, in turnip plants infected with the CaMV mutant, TAVm3 diffuses in the cytoplasm or forms few small aggregates, never triggering the formation of large EDIBs. Electron microscopy observations revealed that TAVm3 inclusion bodies are irregularly shaped and likely correspond to clusters of CaMV particles rather than EDIBs. This is in stark contrast with the sharp contour and uniform matrix of the small and large EDIBs contained in wild-type CaMV-infected turnip cells. Because viroplasms increase in size as synthesis of their components progresses [1,56], we thought that the inability of TAVm3 to self-assemble into typical EDIBs could be due to its low accumulation in infected turnip plants and/or to the host environment, as shown for the CaMV D4 isolate. In fact, the D4 wild-type TAV protein forms small EDIBs in turnip plants (1 μm^2) but EDIBs dramatically increase in size (14 μm^2) in *Datura*, the natural host of this isolate [32]. However, neither the abundance of TAVm3 nor the host context influence the behaviour of TAVm3, since large EDIBs resembling those found in turnip plants infected with wild-type CaMV were never observed in CaMV-TAVm3-infected *Arabidopsis* plants and in transgenic *Arabidopsis* lines, despite abundant synthesis of the TAV mutant [12].

A totally unexpected finding, made while analysing turnip mesophyll cells infected by CaMV-TAVm3 by electron microscopy, was the absence of electron-lucent inclusion bodies (ELIBs). Further studies using fluorescent microscopy revealed that the P2 protein, the major component of ELIBs, formed small aggregates scattered across the cytoplasm despite being present in high amounts, instead of forming a single large ELIB as in wild-type CaMV-infected cells, suggesting that TAV and/or EDIBs influence the behaviour of P2. The interaction between TAV and P2 [57] may induce a conformational change that triggers P2 self-assembly to form ELIB and/or allows its transport along the cytoskeleton network toward the nucleation site [50]. Consequently, the transmission of CaMV by aphids should be strongly affected in plants infected with CaMV-TAVm3, since neither EDIBs (the reservoirs of CaMV particles) nor ELIBs (the platforms for transmission) are formed in plants infected with the CaMV mutant [41,58].

The CaMV-TAVm3 mutant virus was able to systemically infect turnip plants, showing no delay compared to wild-type CaMV. Therefore, the mutation in TAVm3 did not perturb the cell-to-cell and the long distance movement of CaMV particles, in contrast with the mutation within domain D3 that affects TAV self-assembly and the rate of CaMV systemic infection [33]. A recent model proposed that TAV is involved in the intracellular movement of CaMV factories, since EDIBs were observed near plasmodesmata and TAV interacts with CaMV movement protein P1, PDLP1 (Plasmodesmatal-Located Protein 1) and AtSRC2 (Soybean Response to Cold) [25], two cellular proteins localized next to the tubules formed by P1 [30]. Decoration of the cytoskeleton with EGFP-TAVm3 in transient assays and the detection of small TAVm3 bodies along the plasma membrane of infected turnip protoplasts suggests that TAVm3 may still fulfil this hypothetic role.

Unexpectedly, CaMV-TAVm3 produced at least 50 times less virus particles than wild-type CaMV. However, infected turnip plants contained large amounts of processed capsid proteins, suggesting that this yield might be due to the scarce TAVm3 inclusions, since EDIBs are thought to be physical supports for capsid assembly [24]. Very recently, Hafrén *et al.* (2017) proposed that another function of EDIBs is to protect CaMV particles and capsid proteins

against degradation by the NRB1-dependent autophagy mechanism [59]. Our data rather suggest that capsids were properly assembled, as the precursor P4 (pre-CP) is processed most likely after self-assembly [42,60]. CaMV proteins P1 and P2—P5 could not be tested in the absence of specific antibodies—were also produced at significant levels, indicating that the extremely low amount of TAVm3 was sufficient to efficiently activate reinitiation of 35S RNA translation. This also suggests that translation of CaMV transcripts occurred in the absence of typical EDIBs at alternative subcellular compartments, such as the endoplasmic reticulum or the microtubule network [21]. Consequently, we hypothesize that most TAV produced in the course of a wild-type CaMV infection accumulates to form EDIBs, while only a small soluble population of TAV is involved in translation reinitation and other functions such as suppression of RNA silencing. TAV, in its aggregated form, might act as a nucleation site for capsid assembly and encapsidation of viral DNA, and in particular as a chaperonin and/or scaffolding protein, since it physically interacts with P4 [24]. Champagne *et al*., (2004) found pre-CP and viral particles in small TAV bodies, while large EDIBs only contain mature CaMV particles. They proposed a model in which CaMV assembly starts soon after viral DNA synthesis, in small bodies in close proximity to the plasma membrane where pre-CP accumulates. After cleavage of the capsid precursor by the viral protease, the small bodies fuse to generate large EDIBs that consequently contain only mature CaMV particles [61]. The presence of CaMV particles in small TAVm3 bodies suggests that the latter are sites of virus assembly and processing of the pre-CP. However, the discrepancy between the amounts of processed capsid proteins and purified virus particles from CaMV-TAVm3-infected plants strongly suggested that most capsids were empty and/or disassembled in the absence of viral DNA and/or P3 which forms a network around the capsomers [62]. Our data clearly show that the level of CaMV DNA was drastically reduced in CaMV-TAVm3-infected turnip plants, since it represented only approximately 12% of the viral DNA level present in plant infected with wild-type CaMV. Processing of pre-CP by the viral protease indicates that precursor P5 was correctly cleaved *in cis* by the protease located at its N-terminus, thereby activating the reverse transcriptase [43]. Therefore, viral DNA synthesis in turnip plants infected with CaMV-TAVm3 was less active, most likely because large EDIBs, which are the physical supports for DNA replication complexes [45,46], are not formed in the presence of TAVm3. Reduced levels of viral DNA were also observed in turnip plants infected with a CaMV CM1841 isolate which codes for a TAV D3b mutant unable to trigger the formation of correctly sized EDIBs [33] but, by contrast to TAVm3, the subdomain D3b overlaps TAV region required for CaMV DNA synthesis [23]. The reduced DNA level may also be explained, at least partly, by the fact that CaMV-TAV-m3-infected plants contain lower amounts of CaMV pregenomic 35S RNA, probably as a result of degradation, as evidenced by the detection of slightly higher amounts of vsRNAs compared to plants infected with wild-type CaMV [63,64]. As the RNA silencing suppressor activity of TAVm3 is unaffected [12] this suggests that the 35S and 19S RNAs were not physically protected from the RNA silencing machinery. Moreover, transcription of viral DNA yielded much less 35S and 19S RNAs since the nucleus was probably less re-infected by infectious particles and, consequently, a smaller amount of proteins was synthesized compared to an infection with wild-type CaMV.

In summary, we found that the mutation of the EKI motif within the N-terminal self-assembly domain of the TAV protein drastically affects the formation of proper EDIBs, and in particular of large EDIBs, and indirectly impacts the formation of CaMV transmission bodies (ELIB). This is harmful for CaMV DNA replication by reverse transcription and, subsequently, for an efficient production of infectious CaMV progeny, contributing to the loss of CaMV Cabb B-JI virulence on *Brassicaceae*. By contrast, in our experiments the EKI mutation did not hinder virus propagation within the plant, capsid morphogenesis or translation of viral 35S

RNA, indicating that CaMV probably adapts by usurping subcellular compartments, as is the case for other *Caulimoviridae* members, such as *Petuviruses or Badnaviruses*, that do not express TAV homologues. Nevertheless, the functional contribution of the EKI motif in the establishment of an efficient CaMV infectious cycle certainly explains why it is conserved in the TAV protein of all CaMV isolates and in some *Caulimovirus* species.

Supporting information

S1 Table. Oligonucleotides used as PCR primers to generate TAV-encoding cDNA and its derivatives (A) and for semi-quantitative RT-PCR (B). PCR products encoding TAV and TAV mutants were cloned into pmRFP and pCK-EGFP vectors. The restriction sites, used for cloning, at the 5' end of the primers are not indicated. Specific couples of primers were used for semi-quantitative RT-PCR to amplify a sequence of CaMV circular DNA and ORF V, and the *act-2* gene used as reference, respectively. Forward (+) and reverse (–) primers.
(TIF)

S1 Fig. Enlargement of Fig 2B (panels 8–16) showing i) transient co-expression in BY-2 cells of: EGFP-A (A, panel 1), EGFP-TAVΔA (B, panel 1) or both (C, panel 1), and mRFP-TAV (A-C, panels 2); ii) DIC (differential interference contrast)-images (A-C, panels 3), and iii) merged images of panels 1–3 (A-C, panels 4). Scale bars: 10 μm.
(TIF)

S2 Fig. Analysis of EGFP-TAV and EGFP-TAVm3 expression in *Arabidopsis* protoplasts. Protoplasts, prepared from 10 days-old *Arabidopsis* seedlings, were PEG400 (Sigma-Aldrich)-transfected with 15 μg carrier plasmid DNA and 10 μg of either pCK-EGFP-TAV or pCK-EGFP-TAVm3. Ectopically expressed EGFP-TAV and EGFP-TAVm3 were immunodetected by western blot 24h post-transfection in the whole protoplasts lysates with polyclonal rabbit antibodies against TAV (@-TAV) or EGFP, kindly provided by D. Gilmer (IBMP, Strasbourg, France) (@-EGFP), HRP (Horse Radish Peroxydase)-conjugated secondary goat anti-rabbit antibodies (Thermo Fischer Scientific), and luminol-based enhanced chemiluminescence substrate (Lumi-Light^{Plus} Western Blotting Substrate, Roche). The loading control (LC) is RubisCO large subunit (RbcL) revealed by Coomassie blue staining.
(TIF)

S3 Fig. A Z-series stack can be played through the protoplast imaged in Fig 5A. The protoplast, observed by LSCM, was infected with wild-type CaMV and also transiently expressed EGFP-TAVm3. The P2 viral protein was immunodetected with specific anti-P2 antibodies and Alexa 568-conjugated secondary antibodies.
(AVI)

S4 Fig. A Z-series stack can be played through the protoplast imaged in Fig 5B. The protoplast, observed by LSCM, was infected with CaMV-TAVm3 and also transiently expressed EGFP-TAVm3. The P2 viral protein was immunodetected with specific anti-P2 antibodies and Alexa 568-conjugated secondary antibodies.
(AVI)

S5 Fig. Analysis of the stability of CaMV TAV protein in transgenic *Arabidopsis* seedlings. Eight-days-old transgenic *A. thaliana* seedlings expressing CaMV TAV [12] were incubated in MES buffer containing (+ MG132) or not (-MG132) 26S proteasome inhibitor MG132, for several hours at 21˚C. At each time point (in hours), 10 seedlings were collected, ground, and proteins were analysed by western blot using anti-TAV and anti-RGA DELLA polyclonal antibodies and secondary antibodies coupled to alkaline phosphatase. Protein loading was

controlled after transfer, by Ponceau S staining of the membrane (control). (-) corresponds to proteins in non-treated *Arabidopsis* seedlings before starting the proteasome inhibition experiment.
(TIF)

Acknowledgments

We gratefully acknowledge Andy Maule (John Innes Centre, Norwich, England), Stéphane Blanc (INRA, Montpellier, France), Patrick Achard and David Gilmer (IBMP, Strasbourg, France) for kindly providing antibodies, and Marco Incarbone (IBMP, Strasbourg, France) for critical reading of the manuscript.

Author Contributions

Conceptualization: Angèle Geldreich, Gabrielle Haas, Mario Keller, Maria Dimitrova.

Formal analysis: Angèle Geldreich, Gabrielle Haas, Mathieu Erhardt, Mario Keller, Lyubov Ryabova, Maria Dimitrova.

Investigation: Angèle Geldreich, Gabrielle Haas, Julie Kubina, Clément Bouton, Mélanie Tanguy, Mathieu Erhardt, Mario Keller, Maria Dimitrova.

Methodology: Angèle Geldreich, Gabrielle Haas, Clément Bouton, Mathieu Erhardt, Mario Keller, Maria Dimitrova.

Project administration: Mario Keller, Maria Dimitrova.

Supervision: Mario Keller, Maria Dimitrova.

Validation: Angèle Geldreich, Gabrielle Haas, Mario Keller, Maria Dimitrova.

Writing – original draft: Mario Keller, Maria Dimitrova.

Writing – review & editing: Mario Keller, Maria Dimitrova.

References

1. Netherton CL, Wileman T. Virus factories, double membrane vesicles and viroplasm generated in animal cells. Curr Opin Virol. 2011; 1(5): 381–387. https://doi.org/10.1016/j.coviro.2011.09.008 PMID: 22440839
2. Kopito RR. Aggresomes, inclusion bodies and protein aggregation. Trends Cell Biol. 2000; 10(12): 524–530. PMID: 11121744
3. Wileman T. Aggresomes and pericentriolar sites of virus assembly: cellular defense or viral design? Annu Rev Microbiol. 2007; 61: 149–167. https://doi.org/10.1146/annurev.micro.57.030502.090836 PMID: 17896875
4. Johnston JA, Illing ME, Kopito RR. Cytoplasmic dynein/dynactin mediates the assembly of aggresomes. Cell Motil Cytoskeleton. 2002; 53: 26–38. https://doi.org/10.1002/cm.10057 PMID: 12211113
5. Patton JT, Silvestri LS, Tortorici MA, Vasquez-Del Carpio R, Taraporewala ZF. Rotavirus genome replication and morphogenesis: role of the viroplasm. Curr Top Microbiol Immunol. 2006; 309: 169–187. PMID: 16909900
6. Zhang C, Liu Y, Liu L, Lou Z, Zhang H, Miao H, et al. Rice black streaked dwarf virus P9-1, an alpha-helical protein, self-interacts and forms viroplasms in vivo. J Gen Virol. 2008; 89(Pt 7): 1770–1776. https://doi.org/10.1099/vir.0.2008/000109-0 PMID: 18559948
7. Fujisawa I, Matsui C. Electron microscopic autoradiography of X-bodies associated with caulimovirus group infection. Uirusu. 1972; 22(3): 136–138. PMID: 4499990
8. Giband M, Stoeckel ME, Lebeurier G. Use of the immuno-gold technique for in situ localization of cauliflower mosaic virus (CaMV) particles and the major protein of the inclusion bodies. J Virol Methods. 1984; 9(4): 277–281. PMID: 6397480

9. Haas M, Geldreich A, Bureau M, Dupuis L, Leh V, Vetter G, et al. The open reading frame VI product of Cauliflower mosaic virus is a nucleocytoplasmic protein: its N terminus mediates its nuclear export and formation of electron-dense viroplasms. Plant Cell. 2005; 17(3): 927–943. https://doi.org/10.1105/tpc.104.029017 PMID: 15746075
10. Haas M, Bureau M, Geldreich A, Yot P, Keller M. Cauliflower mosaic virus: still in the news. Mol Plant Pathol. 2002; 3(6): 419–429. https://doi.org/10.1046/j.1364-3703.2002.00136.x PMID: 20569349
11. Hohn T, Rothnie H. Plant pararetroviruses: replication and expression. Curr Opin Virol. 2013; 3(6): 621–628. https://doi.org/10.1016/j.coviro.2013.08.013 PMID: 24063990
12. Haas G, Azevedo J, Moissiard G, Geldreich A, Himber C, Bureau M, et al. Nuclear import of CaMV P6 is required for infection and suppression of the RNA silencing factor DRB4. EMBO J. 2008; 27(15): 2102–2112. https://doi.org/10.1038/emboj.2008.129 PMID: 18615098
13. Schoelz JE, Shepherd RJ. Host range control of cauliflower mosaic virus. Virology 1988; 162(1): 30–37. PMID: 3341113
14. Schoelz JE, Shepherd RJ, Daubert S. Region VI of cauliflower mosaic virus encodes a host range determinant. Mol Cell Biol. 1986; 6(7): 2632–2637. PMID: 3785205
15. Leisner SM, Howell SH. Symptom variation in different Arabidopsis thaliana ecotypes produced by Cauliflower mosaic virus. Phytopathology. 1992; 82: 1042–1046.
16. Yu W, Murfett J, Schoelz JE. Differential induction of symptoms in Arabidopsis by P6 of Cauliflower mosaic virus. Mol Plant Microbe Interact. 2003; 16(1): 35–42. https://doi.org/10.1094/MPMI.2003.16.1.35 PMID: 12580280
17. Király L, Cole AB, Bourque JE, Schoelz JE. Systemic cell death is elicited by the interaction of a single gene in Nicotiana clevelandii and gene VI of Cauliflower mosaic virus. Mol Plant Microbe Interact. 1999; 12(10): 919–925.
18. Love AJ, Geri C, Laird J, Carr C, Yun BW, Loake GJ, et al. Cauliflower mosaic virus protein P6 inhibits signaling responses to salicylic acid and regulates innate immunity. PLoS One. 2012; 7(10): e47535. https://doi.org/10.1371/journal.pone.0047535 PMID: 23071821
19. Love AJ, Laird J, Holt J, Hamilton AJ, Sadanandom A, Milner JJ. Cauliflower mosaic virus protein P6 is a suppressor of RNA silencing. J Gen Virol. 2007; 88(Pt12): 3439–3444.
20. Ryabova LA, Pooggin MM, Hohn T. Translation reinitiation and leaky scanning in plant viruses. Virus Res. 2006; 119(1): 52–62. https://doi.org/10.1016/j.virusres.2005.10.017 PMID: 16325949
21. Thiebeauld O, Schepetilnikov M, Park HS, Geldreich A, Kobayashi K, Keller M, et al. A new plant protein interacts with eIF3 and 60S to enhance virus-activated translation re-initiation. EMBO J. 2009; 28(20): 3171–3184. https://doi.org/10.1038/emboj.2009.256 PMID: 19745810
22. Schepetilnikov M, Kobayashi K, Geldreich A, Caranta C, Robaglia C, Keller M, et al. Viral factor TAV recruits TOR/S6K1 signalling to activate reinitiation after long ORF translation. EMBO J. 2011; 30(7): 1343–1356. https://doi.org/10.1038/emboj.2011.39 PMID: 21343906
23. Kobayashi K, Hohn T. Dissection of cauliflower mosaic virus transactivator/viroplasmin reveals distinct essential functions in basic virus replication. J Virol. 2003; 77(15): 8577–8583. https://doi.org/10.1128/JVI.77.15.8577-8583.2003 PMID: 12857928
24. Himmelbach A, Chapdelaine Y, Hohn T. Interaction between cauliflower mosaic virus inclusion body protein and capsid protein: implications for viral assembly. Virology. 1996; 217(1): 147–157. https://doi.org/10.1006/viro.1996.0102 PMID: 8599199
25. Rodriguez A, Angel CA, Lutz L, Leisner SM, Nelson RS, Schoelz JE. Association of the P6 protein of Cauliflower mosaic virus with plasmodesmata and plasmodesmal proteins. Plant Physiol. 2014; 166 (3): 1345–1358. https://doi.org/10.1104/pp.114.249250 PMID: 25239023
26. Harries PA, Palanichelvam K, Yu W, Schoelz JE, Nelson RS. The cauliflower mosaic virus protein P6 forms motile inclusions that traffic along actin microfilaments and stabilize microtubules. Plant Physiol. 2009; 149(2): 1005–1016. https://doi.org/10.1104/pp.108.131755 PMID: 19028879
27. Angel CA, Lutz L, Yang X, Rodriguez A, Adair A, Zhang Y, et al. The P6 protein of Cauliflower mosaic virus interacts with CHUP1, a plant protein which moves chloroplasts on actin microfilaments. Virology 2013; 443: 363–374. https://doi.org/10.1016/j.virol.2013.05.028 PMID: 23769239
28. Cecchini E, Gong Z, Geri C, Covey SN, Milner JJ. Transgenic Arabidopsis lines expressing gene VI from Cauliflower mosaic virus variants exhibit a range of symptom-like phenotypes and accumulate inclusion bodies. Mol Plant Microbe Interact. 1997; 10: 1094–1101. https://doi.org/10.1094/MPMI.1997.10.9.1094 PMID: 9390424
29. Khelifa M, Journou S, Krishnan K, Gargani D, Esperandieu P, Blanc S, et al. Electron-lucent inclusion bodies are structures specialized for aphid transmission of cauliflower mosaic virus. J Gen Virol. 2007; 88(Pt 10): 2872–2880. https://doi.org/10.1099/vir.0.83009-0 PMID: 17872542

30. Schoelz JE, Angel CA, Nelson RS, Leisner SM. A model for intracellular movement of Cauliflower mosaic virus: the concept of the mobile virion factory. J Exp Bot. 2016; 67(7): 2039–2048. https://doi.org/10.1093/jxb/erv520 PMID: 26687180

31. De Zoeten GA, Penswick JR, Horisberger MA, Ahl P, Schultze M, Hohn T. The expression, localization, and effect of a human interferon in plants. Virology. 1989; 172(1): 213–222. PMID: 2773316

32. Shalla TA, Shepherd RJ, Petersen LJ. Comparative cytology of nine isolates of cauliflower mosaic virus. Virology. 1980; 102(2): 381–388. PMID: 18631647

33. Lutz L, Okenka G, Schoelz J, Leisner S. Mutations within A 35 amino acid region of P6 influence self-association, inclusion body formation, and Caulimovirus infectivity. Virology. 2015; 476: 26–36. https://doi.org/10.1016/j.virol.2014.11.017 PMID: 25506670

34. Bureau M, Leh V, Haas M, Geldreich A, Ryabova L, Yot P, et al. P6 protein of Cauliflower mosaic virus, a translation reinitiator, interacts with ribosomal protein L13 from Arabidopsis thaliana. J Gen Virol. 2004; 85(Pt 12): 3765–3775. https://doi.org/10.1099/vir.0.80242-0 PMID: 15557250

35. Hull R, Shepherd RJ, Harvey JD. Cauliflower mosaic virus: an improved purification procedure and some properties of the virus particles J Gen Virol. 1976; 31: 93–100.

36. Xiong C, Lebeurier G, Hirth L. Detection in vivo of a new gene product (gene III) of cauliflower mosaic virus. Proc Natl Acad Sci U S A. 1984; 81(21): 6608–6612. PMID: 16593524

37. Mesnard JM, Mougeot JL, Geldreich A, Lebeurier G. Characterization of different electrophoretic forms of cauliflower mosaic virus virions (strain Cabb-S). Biochimie. 1993; 75(8): 645–649. PMID: 8286436

38. Xiong C, Muller S, Lebeurier G, Hirth L. Identification by immunoprecipitation of cauliflower mosaic virus in vitro major translation product with a specific serum against viroplasm protein. EMBO J. 1982; 1(8): 971–976. PMID: 16453427

39. Bouton C, Geldreich A, Ramel L, Ryabova LA, Dimitrova M, Keller M. Cauliflower mosaic virus transcriptome reveals a complex alternative splicing pattern. PLoS One. 2015; 10(7): e0132665. https://doi.org/10.1371/journal.pone.0132665 PMID: 26162084

40. Espinoza AM, Medina V, Hull R, Markham PG. Cauliflower mosaic virus gene II product forms distinct inclusion bodies in infected plant cells. Virology. 1991; 185(1): 337–344. PMID: 1656590

41. Drucker M, Froissart R, Hebrard E, Uzest M, Ravallec M, Esperandieu P, et al. Intracellular distribution of viral gene products regulates a complex mechanism of cauliflower mosaic virus acquisition by its aphid vector. Proc Natl Acad Sci U S A. 2002; 99(4): 2422–2427. https://doi.org/10.1073/pnas.042587799 PMID: 11842201

42. Karsies A, Hohn T, Leclerc D. Degradation signals within both terminal domains of the cauliflower mosaic virus capsid protein precursor. Plant J. 2001; 27(4): 335–343. PMID: 11532179

43. Torruella M, Gordon K, Hohn T. Cauliflower mosaic virus produces an aspartic proteinase to cleave its polyproteins. EMBO J. 1989; 8(10): 2819–2825. PMID: 2684630

44. Laird J, McInally C, Carr C, Doddiah S, Yates G, Chrysanthou E, et al. Identification of the domains of cauliflower mosaic virus protein P6 responsible for suppression of RNA silencing and salicylic acid signalling. J Gen Virol. 2013; 94(Pt 12): 2777–2789. https://doi.org/10.1099/vir.0.057729-0 PMID: 24088344

45. Mazzolini L, Bonneville JM, Volovitch M, Magazin M, Yot P. Strand-specific viral DNA synthesis in purified viroplasms isolated from turnip leaves infected with cauliflower mosaic virus. Virology. 1985; 145 (2): 293–303. PMID: 18640551

46. Thomas CM, Hull R, Bryant JA, Maule AJ. Isolation of a fraction from Cauliflower mosaic virus-infected protoplasts which is active in the synthesis of (+) and (-) strand viral DNA and reverse transcription of primed RNA templates. Nucleic Acids Res. 1985; 13(12): 4557–4576. PMID: 2409536

47. Li Y, Leisner SM. Multiple domains within the Cauliflower mosaic virus gene VI product interact with the full-length protein. Mol Plant Microbe Interact. 2002; 15(10): 1050–1057. https://doi.org/10.1094/MPMI.2002.15.10.1050 PMID: 12437303

48. Raikhy G, Krause C, Leisner S. The Dahlia mosaic virus gene VI product N-terminal region is involved in self-association. Virus Res. 2011; 159(1): 69–72. https://doi.org/10.1016/j.virusres.2011.04.026 PMID: 21571015

49. Richins RD, Scholthof HB, Shepherd RJ. Sequence of figwort mosaic virus DNA (caulimovirus group). Nucleic Acids Res. 1987; 15(20): 8451–8466. PMID: 3671088

50. Martiniere A, Gargani D, Uzest M, Lautredou N, Blanc S, Drucker M. A role for plant microtubules in the formation of transmission-specific inclusion bodies of Cauliflower mosaic virus. Plant J. 2009; 58(1): 135–146. https://doi.org/10.1111/j.1365-313X.2008.03768.x PMID: 19077170

51. Campagna M, Marcos-Villar L, Arnoldi F, de la Cruz-Herrera CF, Gallego P, Gonzalez-Santamaria J, et al. Rotavirus viroplasm proteins interact with the cellular SUMOylation system: implications for

viroplasm-like structure formation. J Virol. 2013; 87(2): 807–817. https://doi.org/10.1128/JVI.01578-12 PMID: 23115286

52. Daubert SD, Schoelz JE, Debao L, Shepherd RJ. Expression of disease symptoms in cauliflower mosaic virus genomic hybrids. J Mol Appl Genet. 1984; 2(6): 537–547. PMID: 6530602

53. Anderson EJ, Qui SG, Schoelz JE. Genetic analysis of determinants of disease severity and virus concentration in cauliflower mosaic virus. Virology. 1991; 181(2): 647–655. PMID: 2014640

54. Geri C, Cecchini E, Giannakou ME, Covey SN, Milner JJ. Altered patterns of gene expression in Arabidopsis elicited by cauliflower mosaic virus (CaMV) infection and by a CaMV gene VI transgene. Mol Plant Microbe Interact. 1999; 12(5): 377–384. https://doi.org/10.1094/MPMI.1999.12.5.377 PMID: 10226370

55. Shivaprasad PV, Rajeswaran R, Blevins T, Schoelz JE, Meins F Jr, Hohn T, et al. The CaMV transactivator/viroplasmin interferes with RDR6-dependent trans-acting and secondary siRNA pathways in Arabidopsis. Nucleic Acids Res. 2008; 36(18): 5896–5909. https://doi.org/10.1093/nar/gkn590 PMID: 18801846

56. Carreno-Torres JJ, Gutierrez M, Arias CF, Lopez S, Isa P. Characterization of viroplasm formation during the early stages of rotavirus infection. Virol J. 2010; 7: 350–361. https://doi.org/10.1186/1743-422X-7-350 PMID: 21114853

57. Lutz L, Raikhy G, Leisner SM. Cauliflower mosaic virus major inclusion body protein interacts with the aphid transmission factor, the virion-associated protein, and gene VII product. Virus Res. 2012; 170(1–2): 150–153. https://doi.org/10.1016/j.virusres.2012.08.017 PMID: 22982205

58. Bak A, Gargani D, Macia JL, Malouvet E, Vernerey MS, Blanc S, et al. Virus factories of cauliflower mosaic virus are virion reservoirs that engage actively in vector transmission. J Virol. 2013; 87(22): 12207–12215. https://doi.org/10.1128/JVI.01883-13 PMID: 24006440

59. Hafren A, Macia JL, Love AJ, Milner JJ, Drucker M, Hofius D. Selective autophagy limits cauliflower mosaic virus infection by NBR1-mediated targeting of viral capsid protein and particles. Proc Natl Acad Sci U S A. 2017; 114(10): 2026–2035.

60. Karsies A, Merkle T, Szurek B, Bonas U, Hohn T, Leclerc D. Regulated nuclear targeting of cauliflower mosaic virus. J Gen Virol. 2002; 83(Pt7): 1783–1790.

61. Champagne J, Benhamou N, Leclerc D. Localization of the N-terminal domain of cauliflower mosaic virus coat protein precursor. Virology. 2004; 324(2): 257–262. https://doi.org/10.1016/j.virol.2004.04.014 PMID: 15207613

62. Plisson C, Uzest M, Drucker M, Froissart R, Dumas C, Conway J et al. Structure of the mature P3-virus particle complex of cauliflower mosaic virus revealed by cryo-electron microscopy. J Mol Biol. 2005; 346(1): 267–277. https://doi.org/10.1016/j.jmb.2004.11.052 PMID: 15663943

63. Blevins T, Rajeswaran R, Aregger M, Borah BK, Schepetilnikov M, Baerlocher L, et al. Massive production of small RNAs from a non-coding region of Cauliflower mosaic virus in plant defense and viral counter-defense. Nucleic Acids Res. 2011; 39(12): 5003–5014. https://doi.org/10.1093/nar/gkr119 PMID: 21378120

64. Hohn T. RNA based viral silencing suppression in plant pararetroviruses. Front Plant Sci. 2015; 6: 398. https://doi.org/10.3389/fpls.2015.00398 PMID: 26113850

Bibliographie

Bibliographie

Acheson, N.H. (1981). Efficiency of processing of viral RNA during the early and late phases of productive infection by polyoma virus. J. Virol. *37*, 628–635.

Adachi, Y., and Yanagida, M. (1989). Higher order chromosome structure is affected by cold-sensitive mutations in a Schizosaccharomyces pombe gene crm1+ which encodes a 115-kD protein preferentially localized in the nucleus and its periphery. J. Cell Biol. *108*, 1195–1207.

Adam, S.A. (2001). The nuclear pore complex. Genome Biol. *2*, reviews0007.1-reviews0007.6.

Ahmed, Y.F., Gilmartin, G.M., Hanly, S.M., Nevins, J.R., and Greene, W.C. (1991). The HTLV-I Rex response element mediates a novel form of mRNA polyadenylation. Cell *64*, 727–737.

Akarsu, H., Burmeister, W.P., Petosa, C., Petit, I., Müller, C.W., Ruigrok, R.W.H., and Baudin, F. (2003). Crystal structure of the M1 protein-binding domain of the influenza A virus nuclear export protein (NEP/NS2). EMBO J. *22*, 4646–4655.

Aksu, M., Trakhanov, S., and Görlich, D. (2016). Structure of the exportin Xpo4 in complex with RanGTP and the hypusine-containing translation factor eIF5A. Nat. Commun. *7*.

Aksu, M., Pleiner, T., Karaca, S., Kappert, C., Dehne, H.-J., Seibel, K., Urlaub, H., Bohnsack, M.T., and Görlich, D. (2018). Xpo7 is a broad-spectrum exportin and a nuclear import receptor. J. Cell Biol. *217*, 2329–2340.

Alber, F., Dokudovskaya, S., Veenhoff, L.M., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B.T., et al. (2007). The molecular architecture of the nuclear pore complex. Nature *450*, 695.

Alcázar-Román, A.R., Tran, E.J., Guo, S., and Wente, S.R. (2006). Inositol hexakisphosphate and Gle1 activate the DEAD-box protein Dbp5 for nuclear mRNA export. Nat. Cell Biol. *8*, 711–716.

Allen, N.P., Huang, L., Burlingame, A., and Rexach, M. (2001). Proteomic analysis of nucleoporin interacting proteins. J. Biol. Chem. *276*, 29268–29274.

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H.M., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., et al. (2003). Genome-wide Insertional mutagenesis of Arabidopsis thaliana. Science *301*, 653–657.

Amorim, M.J., and Digard, P. (2006). Influenza A virus and the cell nucleus. Vaccine *24*, 6651–6655.

Amorim, M.-J., Read, E.K., Dalton, R.M., Medcalf, L., and Digard, P. (2007). Nuclear export of influenza A virus mRNAs requires ongoing RNA polymerase II activity. Traffic Cph. Den. *8*, 1–11.

Antosz, W., Pfab, A., Ehrnsberger, H.F., Holzinger, P., Köllen, K., Mortensen, S.A., Bruckmann, A., Schubert, T., Längst, G., Griesenbeck, J., et al. (2017). The Composition of the Arabidopsis RNA Polymerase II Transcript Elongation Complex Reveals the Interplay between Elongation and mRNA Processing Factors. Plant Cell *29*, 854–870.

Arts, G.J., Fornerod, M., and Mattaj, I.W. (1998). Identification of a nuclear export receptor for tRNA. Curr. Biol. CB *8*, 305–314.

Aspegren, A., Rabino, C., and Bridge, E. (1998). Organization of splicing factors in adenovirus-infected cells reflects changes in gene expression during the early to late phase transition. Exp. Cell Res. *245*, 203–213.

Bachi, A., Braun, I.C., Rodrigues, J.P., Panté, N., Ribbeck, K., von Kobbe, C., Kutay, U., Wilm, M., Görlich, D., Carmo-Fonseca, M., et al. (2000). The C-terminal domain of TAP interacts with the nuclear pore complex and promotes export of specific CTE-bearing RNA substrates. RNA *6*, 136–158.

Bak, A., Blanc, S., Gargani, D., Martinière, A., and Drucker, M. (2014). Multiples fonctions des usines virales : l'exemple du virus de la mosaïque du chou-fleur (*Cauliflower mosaic virus*). Virologie *18*, 201–210.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281–297.

Bartels, H., and Luban, J. (2014). Gammaretroviral pol sequences act in cis to direct polysome loading and NXF1/NXT-dependent protein production by gag-encoded RNA. Retrovirology *11*, 73.

Basyuk, E., Boulon, S., Skou Pedersen, F., Bertrand, E., and Vestergaard Rasmussen, S. (2005). The packaging signal of MLV is an integrated module that mediates intracellular transport of genomic RNAs. J. Mol. Biol. *354*, 330–339.

Baudin, F., Petit, I., Weissenhorn, W., and Ruigrok, R.W. (2001). In vitro dissection of the membrane and RNP binding activities of influenza virus M1 protein. Virology *281*, 102–108.

Bax, R., Raué, H.A., and Vos, J.C. (2006). Slx9p facilitates efficient ITS1 processing of pre-rRNA in Saccharomyces cerevisiae. RNA *12*, 2005–2013.

Baydoun, H.H., Bellon, M., and Nicot, C. (2008). HTLV-1 Yin and Yang: Rex and p30 master regulators of viral mRNA trafficking. AIDS Rev. *10*, 195–204.

Beck, M., and Hurt, E. (2017). The nuclear pore complex: understanding its function through structural insight. Nat. Rev. Mol. Cell Biol. *18*, 73–89.

Behrens, R.T., Aligeti, M., Pocock, G.M., Higgins, C.A., and Sherer, N.M. (2017). Nuclear Export Signal Masking Regulates HIV-1 Rev Trafficking and Viral RNA Nuclear Export. J. Virol. *91*.

Benfey, P.N., and Chua, N.H. (1990). The Cauliflower Mosaic Virus 35S Promoter: Combinatorial Regulation of Transcription in Plants. Science *250*, 959–966.

Benkő, M. (2008). Adenoviruses: Pathogenesis. In Encyclopedia of Virology (Third Edition), B.W.J. Mahy, and M.H.V. Van Regenmortel, eds. (Oxford: Academic Press), pp. 24–29.

Bentley, D.L. (2005). Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. Curr. Opin. Cell Biol. *17*, 251–256.

Bernad, R., van der Velde, H., Fornerod, M., and Pickersgill, H. (2004). Nup358/RanBP2 attaches to the nuclear pore complex via association with Nup88 and Nup214/CAN and plays a supporting role in CRM1-mediated nuclear protein export. Mol. Cell. Biol. *24*, 2373–2384.

Bhullar, S., Datta, S., Advani, S., Chakravarthy, S., Gautam, T., Pental, D., and Burma, P.K. (2007). Functional analysis of cauliflower mosaic virus 35S promoter: re-evaluation of the role of subdomains B5, B4 and B2 in promoter activity. Plant Biotechnol. J. *5*, 696–708.

Bier, K., York, A., and Fodor, E. (2011). Cellular cap-binding proteins associate with influenza virus mRNAs. J. Gen. Virol. *92*, 1627–1634.

Björk, P., and Wieslander, L. (2017). Integration of mRNP formation and export. Cell. Mol. Life Sci. *74*, 2875–2897.

Blanchette, P., Kindsmüller, K., Groitl, P., Dallaire, F., Speiseder, T., Branton, P.E., and Dobner, T. (2008). Control of mRNA Export by Adenovirus E4orf6 and E1B55K Proteins during Productive Infection Requires E4orf6 Ubiquitin Ligase Activity. J. Virol. *82*, 2642–2651.

Blanvillain, R., Boavida, L.C., McCormick, S., and Ow, D.W. (2008). EXPORTIN1 Genes Are Essential for Development and Function of the Gametophytes in Arabidopsis thaliana. Genetics *180*, 1493–1500.

Blevins, T., Rajeswaran, R., Aregger, M., Borah, B.K., Schepetilnikov, M., Baerlocher, L., Farinelli, L., Meins, F., Hohn, T., and Pooggin, M.M. (2011). Massive production of small RNAs from a non-coding region of Cauliflower mosaic virus in plant defense and viral counter-defense. Nucleic Acids Res. *39*, 5003–5014.

Blobel, G. (1985). Gene gating: a hypothesis. Proc. Natl. Acad. Sci. U. S. A. *82*, 8527–8529.

Bodem, J., Schied, T., Gabriel, R., Rammling, M., and Rethwilm, A. (2011). Foamy Virus Nuclear RNA Export Is Distinct from That of Other Retroviruses. J. Virol. *85*, 2333–2341.

Bogerd, H.P., Echarri, A., Ross, T.M., and Cullen, B.R. (1998). Inhibition of human immunodeficiency virus Rev and human T-cell leukemia virus Rex function, but not Mason-Pfizer monkey virus constitutive transport element activity, by a mutant human nucleoporin targeted to Crm1. J. Virol. *72*, 8627–8635.

Bonnet, A., and Palancade, B. (2014). Regulation of mRNA trafficking by nuclear pore complexes. Genes *5*, 767–791.

Bonnet, A., Bretes, H., and Palancade, B. (2015). Nuclear pore components affect distinct stages of intron-containing gene expression. Nucleic Acids Res. *43*, 4249–4261.

Bonneville, J.M., Sanfaçon, H., Fütterer, J., and Hohn, T. (1989). Posttranscriptional trans-activation in cauliflower mosaic virus. Cell *59*, 1135–1143.

Booth, D.S., Cheng, Y., and Frankel, A.D. (2014). The export receptor Crm1 forms a dimer to promote nuclear export of HIV RNA. ELife *3*, e04121.

Bouton, C. (2014). Etude de l'épissage alternatif de l'ARN 35S du virus de la mosaïque du chou-fleur (CaMV) et de l'export nucléaire des ARN viraux. thesis. Strasbourg.

Bouton, C., Geldreich, A., Ramel, L., Ryabova, L.A., Dimitrova, M., and Keller, M. (2015a). Cauliflower mosaic virus Transcriptome Reveals a Complex Alternative Splicing Pattern. PloS One *10*, e0132665.

Bouton, C., Geldreich, A., Ryabova, L., Keller, M., and Dimitrova, M. (2015b). Le virus de la mosaïque du chou-fleur (CaMV), encore et toujours. Virologie *19*, 119–139.

Bray, M., Prasad, S., Dubay, J.W., Hunter, E., Jeang, K.T., Rekosh, D., and Hammarskjöld, M.L. (1994). A small element from the Mason-Pfizer monkey virus genome makes human immunodeficiency virus type 1 expression and replication Rev-independent. Proc. Natl. Acad. Sci. U. S. A. *91*, 1256–1260.

Brennan, C.M., Gallouzi, I.E., and Steitz, J.A. (2000). Protein ligands to HuR modulate its interaction with target mRNAs in vivo. J. Cell Biol. *151*, 1–14.

Bresson, S., and Tollervey, D. (2018). Surveillance-ready transcription: nuclear RNA decay as a default fate. Open Biol. *8*.

Bridge, E., and Pettersson, U. (1996). Nuclear organization of adenovirus RNA biogenesis. Exp. Cell Res. *229*, 233–239.

Brunotte, L., Flies, J., Bolte, H., Reuther, P., Vreede, F., and Schwemmle, M. (2014). The nuclear export protein of H5N1 influenza A viruses recruits Matrix 1 (M1) protein to the viral ribonucleoprotein to mediate nuclear export. J. Biol. Chem. *289*, 20067–20077.

Bruss, V. (2007). Hepatitis B virus morphogenesis. World J. Gastroenterol. *13*, 65–73.

Bryant, H.E., Wadd, S.E., Lamond, A.I., Silverstein, S.J., and Clements, J.B. (2001). Herpes simplex virus IE63 (ICP27) protein interacts with spliceosome-associated protein 145 and inhibits splicing prior to the first catalytic step. J. Virol. *75*, 4376–4385.

Bui, K.H., von Appen, A., DiGuilio, A.L., Ori, A., Sparks, L., Mackmull, M.-T., Bock, T., Hagen, W., Andrés-Pons, A., Glavy, J.S., et al. (2013). Integrated structural analysis of the human nuclear pore complex scaffold. Cell *155*, 1233–1243.

Bullock, W.O., Fernandez, J.M., and Short, J.M. (1987). XL1-Blue: A high efficiency plasmid transforming recA E. coli strain with beta-galactosidase selection.

Bureau, M., Leh, V., Haas, M., Geldreich, A., Ryabova, L., Yot, P., and Keller, M. (2004). P6 protein of Cauliflower mosaic virus, a translation reinitiator, interacts with ribosomal protein L13 from Arabidopsis thaliana. J. Gen. Virol. *85*, 3765–3775.

Butterfield-Gerson, K.L., Scheifele, L.Z., Ryan, E.P., Hopper, A.K., and Parent, L.J. (2006). Importin-beta family members mediate alpharetrovirus gag nuclear entry via interactions with matrix and nucleocapsid. J. Virol. *80*, 1798–1806.

Cáceres, J.F., Screaton, G.R., and Krainer, A.R. (1998). A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm. Genes Dev. *12*, 55–66.

Cai, X., Hagedorn, C.H., and Cullen, B.R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. RNA N. Y. N *10*, 1957–1966.

Calado, A., Treichel, N., Müller, E.-C., Otto, A., and Kutay, U. (2002). Exportin-5-mediated nuclear export of eukaryotic elongation factor 1A and tRNA. EMBO J. *21*, 6216–6224.

Callan, H.G., Randall, J.T., and Tomlin, S.G. (1949). An electron microscope study of the nuclear membrane. Nature *163*, 280.

Cao, F., and Tavis, J.E. (2011). RNA Elements Directing Translation of the Duck Hepatitis B Virus Polymerase via Ribosomal Shunting ▿ . J. Virol. *85*, 6343–6352.

Cao, Y., and Liu, X. (2007). [HIV-1 Rev and related inhibitors]. Yao Xue Xue Bao *42*, 347–351.

Carmichael, G.G. (2016). Gene Regulation and Quality Control in Murine Polyomavirus Infection. Viruses *8*.

Carter, C.C., Izadpanah, R., and Bridge, E. (2003). Evaluating the role of CRM1-mediated export for adenovirus gene expression. Virology *315*, 224–233.

Carvalho, T., Seeler, J.S., Ohman, K., Jordan, P., Pettersson, U., Akusjärvi, G., Carmo-Fonseca, M., and Dejean, A. (1995). Targeting of adenovirus E1A and E4-ORF3 proteins to nuclear matrix-associated PML bodies. J. Cell Biol. *131*, 45–56.

Casañal, A., Kumar, A., Hill, C.H., Easter, A.D., Emsley, P., Degliesposti, G., Gordiyenko, Y., Santhanam, B., Wolf, J., Wiederhold, K., et al. (2017). Architecture of eukaryotic mRNA 3'-end processing machinery. Science *358*, 1056–1059.

Castro, I.F. de, Volonté, L., and Risco, C. (2013). Virus factories: biogenesis and structural design. Cell. Microbiol. *15*, 24–34.

Cautain, B., Hill, R., de Pedro, N., and Link, W. (2015). Components and regulation of nuclear transport processes. FEBS J. *282*, 445–462.

Cech, T.R., and Steitz, J.A. (2014). The noncoding RNA revolution-trashing old rules to forge new ones. Cell *157*, 77–94.

Chabannes, M., and Iskra-Caruana, M.-L. (2013). Endogenous pararetroviruses--a reservoir of virus infection in plants. Curr. Opin. Virol. *3*, 615–620.

Champagne, J., and Leclerc, D. (2006). Nuclear targeting of the cauliflower mosaic virus (CaMV) genome. Can. J. Bot. *84*, 565–571.

Champagne, J., Laliberté-Gagné, M.-E., and Leclerc, D. (2007). Phosphorylation of the termini of Cauliflower mosaic virus precapsid protein is important for productive infection. Mol. Plant-Microbe Interact. MPMI *20*, 648–658.

Chapdelaine, Y., and Hohn, T. (1998). The cauliflower mosaic virus capsid protein: assembly and nucleic acid binding in vitro. Virus Genes *17*, 139–150.

Chatterjee, K., Majumder, S., Wan, Y., Shah, V., Wu, J., Huang, H.-Y., and Hopper, A.K. (2017). Sharing the load: Mex67-Mtr2 cofunctions with Los1 in primary tRNA nuclear export. Genes Dev. *31*, 2186–2198.

Chatterjee, K., Nostramo, R.T., Wan, Y., and Hopper, A.K. (2018). tRNA dynamics between the nucleus, cytoplasm and mitochondrial surface: Location, location, location. Biochim. Biophys. Acta *1861*, 373–386.

Chaudhary, S., Khokhar, W., Jabre, I., Reddy, A.S.N., Byrne, L.J., Wilson, C.M., and Syed, N.H. (2019). Alternative Splicing and Protein Diversity: Plants Versus Animals. Front. Plant Sci. *10*, 708.

Chen, A., Kao, Y.F., and Brown, C.M. (2005). Translation of the first upstream ORF in the hepatitis B virus pregenomic RNA modulates translation at the core and polymerase initiation codons. Nucleic Acids Res. *33*, 1169–1181.

Chen, A., Panjaworayan T-Thienprasert, N., and Brown, C.M. (2014). Prospects for inhibiting the post-transcriptional regulation of gene expression in hepatitis B virus. World J. Gastroenterol. *20*, 7993–8004.

Chen, M.H., Sheng, J., Hind, G., Handa, A.K., and Citovsky, V. (2000). Interaction between the tobacco mosaic virus movement protein and host cell pectin methylesterases is required for viral cell-to-cell movement. EMBO J. *19*, 913–920.

Cheng, H., Dufu, K., Lee, C.-S., Hsu, J.L., Dias, A., and Reed, R. (2006). Human mRNA Export Machinery Recruited to the 5′ End of mRNA. Cell *127*, 1389–1400.

Chi, B., Wang, K., Du, Y., Gui, B., Chang, X., Wang, L., Fan, J., Chen, S., Wu, X., Li, G., et al. (2014). A Sub-Element in PRE enhances nuclear export of intronless mRNAs by recruiting the TREX complex via ZC3H18. Nucleic Acids Res. *42*, 7305–7318.

Chiba, S., Hill-Batorski, L., Neumann, G., and Kawaoka, Y. (2018). The Cellular DExD/H-Box RNA Helicase UAP56 Co-localizes With the Influenza A Virus NS1 Protein. Front. Microbiol. 9.

Choudury, S.G., Shahid, S., Cuerda-Gil, D., Panda, K., Cullen, A., Ashraf, Q., Sigman, M.J., McCue, A.D., and Slotkin, R.K. (2019). The RNA Export Factor ALY1 Enables Genome-Wide RNA-Directed DNA Methylation. Plant Cell *31*, 759–774.

Claverie, J.-M., and Abergel, C. (2016). [Giant viruses: update, enigmas, controversies and perspectives]. Med. Sci. MS *32*, 1087–1096.

Claverie, J.-M., and Abergel, C. (2018). Mimiviridae: An Expanding Family of Highly Diverse Large dsDNA Viruses Infecting a Wide Phylogenetic Range of Aquatic Eukaryotes. Viruses *10*.

Cochrane, A.W., Golub, E., Volsky, D., Ruben, S., and Rosen, C.A. (1989). Functional significance of phosphorylation to the human immunodeficiency virus Rev protein. J. Virol. *63*, 4438–4440.

Cochrane, A.W., Chen, C.H., and Rosen, C.A. (1990). Specific interaction of the human immunodeficiency virus Rev protein with a structured region in the env mRNA. Proc. Natl. Acad. Sci. U. S. A. *87*, 1198–1202.

Cocking, E.C. (1960). A Method for the Isolation of Plant Protoplasts and Vacuoles. Nature *187*, 962.

Cole, C.N., and Scarcelli, J.J. (2006). Transport of messenger RNA from the nucleus to the cytoplasm. Curr. Opin. Cell Biol. *18*, 299–306.

Cook, A.G., Fukuhara, N., Jinek, M., and Conti, E. (2009). Structures of the tRNA export factor in the nuclear and cytosolic states. Nature *461*, 60–65.

Corbin-Lickfett, K.A., Chen, I.-H.B., Cocco, M.J., and Sandri-Goldin, R.M. (2009). The HSV-1 ICP27 RGG box specifically binds flexible, GC-rich sequences but not G-quartet structures. Nucleic Acids Res. *37*, 7290–7301.

la Cour, T., Gupta, R., Rapacki, K., Skriver, K., Poulsen, F.M., and Brunak, S. (2003). NESbase version 1.0: a database of nuclear export signals. Nucleic Acids Res. *31*, 393–396.

Culjkovic, B., Topisirovic, I., Skrabanek, L., Ruiz-Gutierrez, M., and Borden, K.L.B. (2006). eIF4E is a central node of an RNA regulon that governs cellular proliferation. J. Cell Biol. *175*, 415–426.

Cullen, B.R. (2003). Nuclear mRNA export: insights from virology. Trends Biochem. Sci. *28*, 419–424.

Cupelli, L., Okenquist, S.A., Trubetskoy, A., and Lenz, J. (1998). The secondary structure of the R region of a murine leukemia virus is important for stimulation of long terminal repeat-driven gene expression. J. Virol. *72*, 7807–7814.

Dai-Ju, J.Q., Li, L., Johnson, L.A., and Sandri-Goldin, R.M. (2006). ICP27 interacts with the C-terminal domain of RNA polymerase II and facilitates its recruitment to herpes simplex virus 1 transcription sites, where it undergoes proteasomal degradation during infection. J. Virol. *80*, 3567–3581.

Daneholt, B. (2001). Assembly and transport of a premessenger RNP particle. Proc. Natl. Acad. Sci. U. S. A. *98*, 7012–7017.

Davey, M.R., Anthony, P., Power, J.B., and Lowe, K.C. (2005). Plant protoplasts: status and biotechnological perspectives. Biotechnol. Adv. *23*, 131–171.

Davison, A.J., Benko, M., and Harrach, B. (2003). Genetic content and evolution of adenoviruses. J. Gen. Virol. *84*, 2895–2908.

De Vlugt, C., Sikora, D., and Pelchat, M. (2018). Insight into Influenza: A Virus Cap-Snatching. Viruses *10*.

Delseny, M., and Hull, R. (1983). Isolation and characterization of faithful and altered clones of the genomes of cauliflower mosaic virus isolates Cabb B-JI, CM4-184, and Bari I. Plasmid *9*, 31–41.

DiMattia, M.A., Watts, N.R., Stahl, S.J., Rader, C., Wingfield, P.T., Stuart, D.I., Steven, A.C., and Grimes, J.M. (2010). Implications of the HIV-1 Rev dimer structure at 3.2 A resolution for multimeric binding to the Rev response element. Proc. Natl. Acad. Sci. U. S. A. *107*, 5810–5814.

DiMattia, M.A., Watts, N.R., Cheng, N., Huang, R., Heymann, J.B., Grimes, J.M., Wingfield, P.T., Stuart, D.I., and Steven, A.C. (2016). The Structure of HIV-1 Rev Filaments Suggests a Bilateral Model for Rev-RRE Assembly. Structure *24*, 1068–1080.

Dobbelstein, M., Roth, J., Kimberly, W.T., Levine, A.J., and Shenk, T. (1997). Nuclear export of the E1B 55-kDa and E4 34-kDa adenoviral oncoproteins mediated by a rev-like signal sequence. EMBO J. *16*, 4276–4284.

Dobner, T., and Kzhyshkowska, J. (2001). Nuclear Export of Adenovirus RNA. In Nuclear Export of Viral RNAs, J. Hauber, and P.K. Vogt, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 25–54.

Dong, S., Li, C., Zenklusen, D., Singer, R.H., Jacobson, A., and He, F. (2007). YRA1 autoregulation requires nuclear export and cytoplasmic Edc3p-mediated degradation of its pre-mRNA. Mol. Cell *25*, 559–573.

Dong, X., Biswas, A., and Chook, Y.M. (2009a). Structural basis for assembly and disassembly of the CRM1 nuclear export complex. Nat. Struct. Mol. Biol. *16*, 558–560.

Dong, X., Biswas, A., Süel, K.E., Jackson, L.K., Martinez, R., Gu, H., and Chook, Y.M. (2009b). Structural basis for leucine-rich nuclear export signal recognition by CRM1. Nature *458*, 1136–1141.

Dong, Y., Yoshitomi, T., Hu, J.-F., and Cui, J. (2017). Long noncoding RNAs coordinate functions between mitochondria and the nucleus. Epigenetics Chromatin *10*, 41.

Dorfman, J., and Macara, I.G. (2008). STRADalpha regulates LKB1 localization by blocking access to importin-alpha, and by association with Crm1 and exportin-7. Mol. Biol. Cell *19*, 1614–1626.

D'Souza, V., and Summers, M.F. (2005). How retroviruses select their genomes. Nat. Rev. Microbiol. *3*, 643–655.

Dubois, J., Terrier, O., and Rosa-Calatrava, M. (2014). Influenza Viruses and mRNA Splicing: Doing More with Less. MBio *5*.

Dufu, K., Livingstone, M.J., Seebacher, J., Gygi, S.P., Wilson, S.A., and Reed, R. (2010). ATP is required for interactions between UAP56 and two conserved mRNA export proteins, Aly and CIP29, to assemble the TREX complex. Genes Dev. *24*, 2043–2053.

Dunoyer, P., Thomas, C., Harrison, S., Revers, F., and Maule, A. (2004). A cysteine-rich plant protein potentiates Potyvirus movement through an interaction with the virus genome-linked protein VPg. J. Virol. *78*, 2301–2309.

Duriez, M., Mandouri, Y., Lekbaby, B., Wang, H., Schnuriger, A., Redelsperger, F., Guerrera, C.I., Lefevre, M., Fauveau, V., Ahodantin, J., et al. (2017). Alternative splicing of hepatitis B virus: A novel virus/host interaction altering liver immunity. J. Hepatol. *67*, 687–699.

Ehrnsberger, H.F., Grasser, M., and Grasser, K.D. (2019a). Nucleocytosolic mRNA transport in plants: export factors and their influence on growth and development. J. Exp. Bot.

Ehrnsberger, H.F., Pfaff, C., Hachani, I., Flores-Tornero, M., Sørensen, B.B., Längst, G., Sprunck, S., Grasser, M., and Grasser, K.D. (2019b). The UAP56-Interacting Export Factors UIEF1 and UIEF2 Function in mRNA Export. Plant Physiol. *179*, 1525–1536.

Eiras, M., Nohales, M.A., Kitajima, E.W., Flores, R., and Daròs, J.A. (2011). Ribosomal protein L5 and transcription factor IIIA from Arabidopsis thaliana bind in vitro specifically Potato spindle tuber viroid RNA. Arch. Virol. *156*, 529–533.

Elton, D., Simpson-Holley, M., Archer, K., Medcalf, L., Hallam, R., McCauley, J., and Digard, P. (2001). Interaction of the influenza virus nucleoprotein with the cellular CRM1-mediated nuclear export pathway. J. Virol. *75*, 408–419.

Emerman, M., and Malim, M.H. (1998). HIV-1 regulatory/accessory genes: keys to unraveling viral and host cell biology. Science *280*, 1880–1884.

Emerman, M., Vazeux, R., and Peden, K. (1989). The rev gene product of the human immunodeficiency virus affects envelope-specific RNA localization. Cell *57*, 1155–1165.

Engelhardt, O.G., Smith, M., and Fodor, E. (2005). Association of the influenza A virus RNA-dependent RNA polymerase with cellular RNA polymerase II. J. Virol. *79*, 5812–5818.

Ernst, R.K., Bray, M., Rekosh, D., and Hammarskjöld, M.L. (1997a). A structured retroviral RNA element that mediates nucleocytoplasmic export of intron-containing RNA. Mol. Cell. Biol. *17*, 135–144.

Ernst, R.K., Bray, M., Rekosh, D., and Hammarskjold, M.L. (1997b). Secondary structure and mutational analysis of the Mason-Pfizer monkey virus RNA constitutive transport element. RNA N. Y. N *3*, 210–222.

Escudero-Paunetto, L., Li, L., Hernandez, F.P., and Sandri-Goldin, R.M. (2010). SR proteins SRp20 and 9G8 contribute to efficient export of herpes simplex virus 1 mRNAs. Virology *401*, 155–164.

Fang, X., Wang, J., O'Carroll, I.P., Mitchell, M., Zuo, X., Wang, Y., Yu, P., Liu, Y., Rausch, J.W., Dyba, M.A., et al. (2013). An Unusual Topological Structure of the HIV-1 Rev Response Element. Cell *155*, 594–605.

Fasken, M.B., and Corbett, A.H. (2005). Process or perish: quality control in mRNA biogenesis. Nat. Struct. Mol. Biol. *12*, 482–488.

Faza, M.B., Chang, Y., Occhipinti, L., Kemmler, S., and Panse, V.G. (2012). Role of Mex67-Mtr2 in the nuclear export of 40S pre-ribosomes. PLoS Genet. *8*, e1002915.

Fernandez-Martinez, J., Kim, S.J., Shi, Y., Upla, P., Pellarin, R., Gagnon, M., Chemmama, I.E., Wang, J., Nudelman, I., Zhang, W., et al. (2016). Structure and Function of the Nuclear Pore Complex Cytoplasmic mRNA Export Platform. Cell *167*, 1215-1228.e25.

Fischer, U., Huber, J., Boelens, W.C., Mattaj, I.W., and Lührmann, R. (1995). The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. Cell *82*, 475–483.

Fischer, U., Schäuble, N., Schütz, S.C., Altvater, M., Chang, Y., Faza, M.B., and Panse, V.G. (2015). A non-canonical mechanism for Crm1-export cargo complex assembly. In ELife, p.

Fleckner, J., Zhang, M., Valcárcel, J., and Green, M.R. (1997). U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction. Genes Dev. *11*, 1864–1872.

Flint, S.J., and Gonzalez, R.A. (2003). Regulation of mRNA production by the adenoviral E1B 55-kDa and E4 Orf6 proteins. Curr. Top. Microbiol. Immunol. *272*, 287–330.

Flint, S.J., Huang, W., Goodhouse, J., and Kyin, S. (2005). A peptide inhibitor of exportin1 blocks shuttling of the adenoviral E1B 55 kDa protein but not export of viral late mRNAs. Virology *337*, 7–17.

Fodor, E. (2013). The RNA polymerase of influenza a virus: mechanisms of viral transcription and replication. Acta Virol. *57*, 113–122.

Fornerod, M., Ohno, M., Yoshida, M., and Mattaj, I.W. (1997). CRM1 is an export receptor for leucine-rich nuclear export signals. Cell *90*, 1051–1060.

Franck, A., Guilley, H., Jonard, G., Richards, K., and Hirth, L. (1980). Nucleotide sequence of cauliflower mosaic virus DNA. Cell *21*, 285–294.

Frey, S., Richter, R.P., and Görlich, D. (2006). FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. Science *314*, 815–817.

Fribourg, S., and Conti, E. (2003). Structural similarity in the absence of sequence homology of the messenger RNA export factors Mtr2 and p15. EMBO Rep. *4*, 699–703.

Fried, H., and Kutay, U. (2003). Nucleocytoplasmic transport: taking an inventory. Cell. Mol. Life Sci. CMLS *60*, 1659–1688.

Fries, B., Heukeshoven, J., Hauber, I., Grüttner, C., Stocking, C., Kehlenbach, R.H., Hauber, J., and Chemnitz, J. (2007). Analysis of nucleocytoplasmic trafficking of the HuR ligand APRIL and its influence on CD83 expression. J. Biol. Chem. *282*, 4504–4515.

Froissart, R., Uzest, M., Ruiz-Ferrer, V., Drucker, M., Hébrard, E., Hohn, T., and Blanc, S. (2004). Splicing of Cauliflower mosaic virus 35S RNA serves to downregulate a toxic gene product. J. Gen. Virol. *85*, 2719–2726.

Fu, S.-C., Huang, H.-C., Horton, P., and Juan, H.-F. (2013). ValidNESs: a database of validated leucine-rich nuclear export signals. Nucleic Acids Res. *41*, D338-343.

Fu, S.-C., Fung, H.Y.J., Cağatay, T., Baumhardt, J., and Chook, Y.M. (2018a). Correlation of CRM1-NES affinity with nuclear export activity. Mol. Biol. Cell *29*, 2037–2044.

Fu, X., Liang, C., Li, F., Wang, L., Wu, X., Lu, A., Xiao, G., and Zhang, G. (2018b). The Rules and Functions of Nucleocytoplasmic Shuttling Proteins. Int. J. Mol. Sci. *19*.

Fu, Y., Li, H., and Yang, Z. (2002). The ROP2 GTPase Controls the Formation of Cortical Fine F-Actin and the Early Phase of Directional Cell Expansion during Arabidopsis Organogenesis. Plant Cell *14*, 777–794.

Fujisawa, I., Rubio-Huertos, M., Matsui, C., and Yamaguchi, A. (1967). Intracellular appearance of cauliflower mosaic virus particles. Phytopathology *57*, 1130–1132.

Fukuda, M., Asano, S., Nakamura, T., Adachi, M., Yoshida, M., Yanagida, M., and Nishida, E. (1997). CRM1 is responsible for intracellular transport mediated by the nuclear export signal. Nature *390*, 308–311.

Gabler, S., Schütt, H., Groitl, P., Wolf, H., Shenk, T., and Dobner, T. (1998). E1B 55-kilodalton-associated protein: a cellular protein with RNA-binding activity implicated in nucleocytoplasmic transport of adenovirus and cellular mRNAs. J. Virol. *72*, 7960–7971.

Gallouzi, I.E., and Steitz, J.A. (2001). Delineation of mRNA export pathways by the use of cell-permeable peptides. Science *294*, 1895–1901.

Gallucci, L., and Kann, M. (2017). Nuclear Import of Hepatitis B Virus Capsids and Genome. Viruses 9.

Galy, V., Gadal, O., Fromont-Racine, M., Romano, A., Jacquier, A., and Nehrbass, U. (2004). Nuclear retention of unspliced mRNAs in yeast is mediated by perinuclear Mlp1. Cell *116*, 63–73.

Gaouar, O., and Germain, H. (2013). mRNA export: threading the needle. Front. Plant Sci. *4*, 59.

Garbitt-Hirst, R., Kenney, S.P., and Parent, L.J. (2009). Genetic evidence for a connection between Rous sarcoma virus gag nuclear trafficking and genomic RNA packaging. J. Virol. *83*, 6790–6797.

García-Oliver, E., García-Molinero, V., and Rodríguez-Navarro, S. (2012). mRNA export and gene expression: the SAGA-TREX-2 connection. Biochim. Biophys. Acta *1819*, 555–565.

Gatfield, D., and Izaurralde, E. (2002). REF1/Aly and the additional exon junction complex proteins are dispensable for nuclear mRNA export. J. Cell Biol. *159*, 579–588.

Gatignol, A. (2007). Transcription of HIV: Tat and Cellular Chromatin. In Advances in Pharmacology, (Academic Press), pp. 137–159.

Gebhardt, A., Habjan, M., Benda, C., Meiler, A., Haas, D.A., Hein, M.Y., Mann, A., Mann, M., Habermann, B., and Pichlmair, A. (2015). mRNA export through an additional cap-binding complex consisting of NCBP1 and NCBP3. Nat. Commun. *6*, 8192.

Geldreich, A., Haas, G., Kubina, J., Bouton, C., Tanguy, M., Erhardt, M., Keller, M., Ryabova, L., and Dimitrova, M. (2017). Formation of large viroplasms and virulence of Cauliflower mosaic virus in turnip plants depend on the N-terminal EKI sequence of viral protein TAV. PLOS ONE *12*, e0189062.

Gerelsaikhan, T., Tavis, J.E., and Bruss, V. (1996). Hepatitis B virus nucleocapsid envelopment does not occur without genomic DNA synthesis. J. Virol. *70*, 4269–4274.

Gerhardy, S., Menet, A.M., Peña, C., Petkowski, J.J., and Panse, V.G. (2014). Assembly and nuclear export of pre-ribosomal particles in budding yeast. Chromosoma *123*, 327–344.

Germain, H., Qu, N., Cheng, Y.T., Lee, E., Huang, Y., Dong, O.X., Gannon, P., Huang, S., Ding, P., Li, Y., et al. (2010). MOS11: A New Component in the mRNA Export Pathway. PLoS Genet. *6*.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345.

Gibson, D.G., Smith, H.O., Hutchison Iii, C.A., Venter, J.C., and Merryman, C. (2010). Chemical synthesis of the mouse mitochondrial genome. Nat. Methods *7*, 901–903.

Gilbert, W., and Guthrie, C. (2004). The Glc7p nuclear phosphatase promotes mRNA export by facilitating association of Mex67p with mRNA. Mol. Cell *13*, 201–212.

Gonatopoulos-Pournatzis, T., and Cowling, V.H. (2014). Cap-binding complex (CBC). Biochem. J. *457*, 231–242.

Gontan, C., Güttler, T., Engelen, E., Demmers, J., Fornerod, M., Grosveld, F.G., Tibboel, D., Görlich, D., Poot, R.A., and Rottier, R.J. (2009). Exportin 4 mediates a novel nuclear import pathway for Sox family transcription factors. J. Cell Biol. *185*, 27–34.

Gonzalez, R.A., and Flint, S.J. (2002). Effects of mutations in the adenoviral E1B 55-kilodalton protein coding sequence on viral late mRNA metabolism. J. Virol. *76*, 4507–4519.

Görlich, D., Prehn, S., Laskey, R.A., and Hartmann, E. (1994). Isolation of a protein that is essential for the first step of nuclear protein import. Cell *79*, 767–778.

Görlich, D., Kraft, R., Kostka, S., Vogel, F., Hartmann, E., Laskey, R.A., Mattaj, I.W., and Izaurralde, E. (1996). Importin provides a link between nuclear protein import and U snRNA export. Cell *87*, 21–32.

Graham, S.V. (2010). Human papillomavirus: gene expression, regulation and prospects for novel diagnostic methods and antiviral therapies. Future Microbiol. *5*, 1493–1506.

Grant, S.G., Jessee, J., Bloom, F.R., and Hanahan, D. (1990). Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants. Proc. Natl. Acad. Sci. U. S. A. *87*, 4645–4649.

Grespin, M.E., Bonamy, G.M.C., Roggero, V.R., Cameron, N.G., Adam, L.E., Atchison, A.P., Fratto, V.M., and Allison, L.A. (2008). Thyroid Hormone Receptor α1 Follows a Cooperative CRM1/Calreticulin-mediated Nuclear Export Pathway. J. Biol. Chem. *283*, 25576–25588.

Gromadzka, A.M., Steckelberg, A.-L., Singh, K.K., Hofmann, K., and Gehring, N.H. (2016). A short conserved motif in ALYREF directs cap- and EJC-dependent assembly of export complexes on spliced mRNAs. Nucleic Acids Res. *44*, 2348–2361.

Groom, H.C.T., Anderson, E.C., and Lever, A.M.L. (2009). Rev: beyond nuclear export. J. Gen. Virol. *90*, 1303–1318.

Grossman, E., Medalia, O., and Zwerger, M. (2012). Functional architecture of the nuclear pore complex. Annu. Rev. Biophys. *41*, 557–584.

Gruffat, H., Mure, F., and Manet, E. (2019). SRSF3: entre épissage, export et dégradation des ARNm du virus d'Epstein-Barr. médecine/sciences *35*, 103–105.

Grünwald, D., and Singer, R.H. (2010). In vivo imaging of labelled endogenous β-actin mRNA during nucleocytoplasmic transport. Nature *467*, 604–607.

Grünwald, D., Singer, R.H., and Rout, M. (2011). Nuclear export dynamics of RNA–protein complexes. Nature *475*, 333–341.

Grüter, P., Tabernero, C., von Kobbe, C., Schmitt, C., Saavedra, C., Bachi, A., Wilm, M., Felber, B.K., and Izaurralde, E. (1998). TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. Mol. Cell *1*, 649–659.

Gudleski, N., Flanagan, J.M., Ryan, E.P., Bewley, M.C., and Parent, L.J. (2010). Directionality of nucleocytoplasmic transport of the retroviral gag protein depends on sequential binding of karyopherins and viral RNA. Proc. Natl. Acad. Sci. U. S. A. *107*, 9358–9363.

Guerra-Peraza, O., de Tapia, M., Hohn, T., and Hemmings-Mieszczak, M. (2000). Interaction of the cauliflower mosaic virus coat protein with the pregenomic RNA leader. J. Virol. *74*, 2067–2072.

Ha, M., and Kim, V.N. (2014). Regulation of microRNA biogenesis. Nat. Rev. Mol. Cell Biol. *15*, 509–524.

Haas, G., Azevedo, J., Moissiard, G., Geldreich, A., Himber, C., Bureau, M., Fukuhara, T., Keller, M., and Voinnet, O. (2008). Nuclear import of CaMV P6 is required for infection and suppression of the RNA silencing factor DRB4. EMBO J. *27*, 2102–2112.

Haas, M., Bureau, M., Geldreich, A., Yot, P., and Keller, M. (2002). Cauliflower mosaic virus: still in the news. Mol. Plant Pathol. *3*, 419–429.

Haas, M., Geldreich, A., Bureau, M., Dupuis, L., Leh, V., Vetter, G., Kobayashi, K., Hohn, T., Ryabova, L., Yot, P., et al. (2005). The open reading frame VI product of Cauliflower mosaic virus is a nucleocytoplasmic protein: its N terminus mediates its nuclear export and formation of electron-dense viroplasms. Plant Cell *17*, 927–943.

Haasen, D., Köhler, C., Neuhaus, G., and Merkle, T. (1999). Nuclear export of proteins in plants: AtXPO1 is the export receptor for leucine-rich nuclear export signals in Arabidopsis thaliana. Plant J. Cell Mol. Biol. *20*, 695–705.

Halbert, D.N., Cutt, J.R., and Shenk, T. (1985). Adenovirus early region 4 encodes functions required for efficient DNA replication, late gene expression, and host cell shutoff. J. Virol. *56*, 250–257.

Hale, B.G., Randall, R.E., Ortín, J., and Jackson, D. (2008). The multifunctional NS1 protein of influenza A viruses. J. Gen. Virol. *89*, 2359–2376.

Hamann, M.V., and Lindemann, D. (2016). Foamy Virus Protein-Nucleic Acid Interactions during Particle Morphogenesis. Viruses *8*.

Harries, P.A., Palanichelvam, K., Yu, W., Schoelz, J.E., and Nelson, R.S. (2009). The cauliflower mosaic virus protein P6 forms motile inclusions that traffic along actin microfilaments and stabilize microtubules. Plant Physiol. *149*, 1005–1016.

Hass, M., Hannoun, C., Kalinina, T., Sommer, G., Manegold, C., and Günther, S. (2005). Functional analysis of hepatitis B virus reactivating in hepatitis B surface antigen-negative individuals. Hepatol. Baltim. Md *42*, 93–103.

Hauber, J., Bouvier, M., Malim, M.H., and Cullen, B.R. (1988). Phosphorylation of the rev gene product of human immunodeficiency virus type 1. J. Virol. *62*, 4801–4804.

Hautbergue, G.M., Hung, M.-L., Walsh, M.J., Snijders, A.P.L., Chang, C.-T., Jones, R., Ponting, C.P., Dickman, M.J., and Wilson, S.A. (2009). UIF, a New mRNA export adaptor that works together with REF/ALY, requires FACT for recruitment to mRNA. Curr. Biol. CB *19*, 1918–1924.

Hayama, R., Rout, M.P., and Fernandez-Martinez, J. (2017). The nuclear pore complex core scaffold and permeability barrier: variations of a common theme. Curr. Opin. Cell Biol. *46*, 110–118.

Heath, C.G., Viphakone, N., and Wilson, S.A. (2016). The role of TREX in gene expression and disease. Biochem. J. *473*, 2911–2935.

Hedges, J., West, M., and Johnson, A.W. (2005). Release of the export adapter, Nmd3p, from the 60S ribosomal subunit requires Rpl10p and the cytoplasmic GTPase Lsg1p. EMBO J. *24*, 567–579.

Heise, T., Sommer, G., Reumann, K., Meyer, I., Will, H., and Schaal, H. (2006). The hepatitis B virus PRE contains a splicing regulatory element. Nucleic Acids Res. *34*, 353–363.

Hellmuth, K., Lau, D.M., Bischoff, F.R., Künzler, M., Hurt, E., and Simos, G. (1998). Yeast Los1p has properties of an exportin-like nucleocytoplasmic transport factor for tRNA. Mol. Cell. Biol. *18*, 6374–6386.

Hemmings-Mieszczak, M., Steger, G., and Hohn, T. (1997). Alternative structures of the cauliflower mosaic virus 35 S RNA leader: implications for viral expression and replication. J. Mol. Biol. *267*, 1075–1088.

Hernandez, F.P., and Sandri-Goldin, R.M. (2010). Head-to-tail intramolecular interaction of herpes simplex virus type 1 regulatory protein ICP27 is important for its interaction with cellular mRNA export receptor TAP/NXF1. MBio *1*.

Herold, A., Suyama, M., Rodrigues, J.P., Braun, I.C., Kutay, U., Carmo-Fonseca, M., Bork, P., and Izaurralde, E. (2000). TAP (NXF1) Belongs to a Multigene Family of Putative RNA Export Factors with a Conserved Modular Architecture. Mol. Cell. Biol. *20*, 8996–9008.

Hicks, G.R., and Raikhel, N.V. (1993). Specific binding of nuclear localization sequences to plant nuclei. Plant Cell *5*, 983–994.

Hilleren, P., McCarthy, T., Rosbash, M., Parker, R., and Jensen, T.H. (2001). Quality control of mRNA 3'-end processing is linked to the nuclear exosome. Nature *413*, 538–542.

Himmelbach, A., Chapdelaine, Y., and Hohn, T. (1996). Interaction between cauliflower mosaic virus inclusion body protein and capsid protein: implications for viral assembly. Virology *217*, 147–157.

Hirochika, H., Takatsuji, H., Ubasawa, A., and Ikeda, J.E. (1985). Site-specific deletion in cauliflower mosaic virus DNA: possible involvement of RNA splicing and reverse transcription. EMBO J. *4*, 1673–1680.

Hofmann, W., Reichart, B., Ewald, A., Müller, E., Schmitt, I., Stauber, R.H., Lottspeich, F., Jockusch, B.M., Scheer, U., Hauber, J., et al. (2001). Cofactor requirements for nuclear export of Rev response element (RRE)- and constitutive transport element (CTE)-containing retroviral RNAs. An unexpected role for actin. J. Cell Biol. *152*, 895–910.

Hogan, P.G., Chen, L., Nardone, J., and Rao, A. (2003). Transcriptional regulation by calcium, calcineurin, and NFAT. Genes Dev. *17*, 2205–2232.

Hoh, F., Uzest, M., Drucker, M., Plisson-Chastang, C., Bron, P., Blanc, S., and Dumas, C. (2010). Structural insights into the molecular mechanisms of cauliflower mosaic virus transmission by its insect vector. J. Virol. *84*, 4706–4713.

Hohn, T. (2015). RNA based viral silencing suppression in plant pararetroviruses. Front. Plant Sci. *6*, 398.

Hohn, T., and Rothnie, H. (2013). Plant pararetroviruses: replication and expression. Curr. Opin. Virol. *3*, 621–628.

Hohn, T., Park, H.S., Guerra-Peraza, O., Stavolone, L., Pooggin, M.M., Kobayashi, K., and Ryabova, L.A. (2001). Shunting and controlled reinitiation: the encounter of cauliflower mosaic virus with the translational machinery. Cold Spring Harb. Symp. Quant. Biol. *66*, 269–276.

Holaska, J.M., Black, B.E., Love, D.C., Hanover, J.A., Leszyk, J., and Paschal, B.M. (2001). Calreticulin Is a Receptor for Nuclear Export. J. Cell Biol. *152*, 127–140.

Hopper, A.K. (2013). Transfer RNA post-transcriptional processing, turnover, and subcellular dynamics in the yeast Saccharomyces cerevisiae. Genetics *194*, 43–67.

Hoshi, S., Odawara, T., Oshima, M., Kitamura, Y., Takizawa, H., and Yoshikura, H. (2002). cis-Elements involved in expression of unspliced RNA in Moloney murine leukemia virus. Biochem. Biophys. Res. Commun. *290*, 1139–1144.

Hsin, J.-P., Xiang, K., and Manley, J.L. (2014). Function and control of RNA polymerase II C-terminal domain phosphorylation in vertebrate transcription and RNA processing. Mol. Cell. Biol. *34*, 2488–2498.

Hu, W.-S., and Hughes, S.H. (2012). HIV-1 reverse transcription. Cold Spring Harb. Perspect. Med. *2*.

Huang, H.-Y., and Hopper, A.K. (2015). In vivo biochemical analyses reveal distinct roles of β-importins and eEF1A in tRNA subcellular traffic. Genes Dev. *29*, 772–783.

Huang, J., and Liang, T.J. (1993). A novel hepatitis B virus (HBV) genetic element with Rev response element-like properties that is essential for expression of HBV gene products. Mol. Cell. Biol. *13*, 7476–7486.

Huang, Y., and Carmichael, G.G. (2009). RNA processing in the polyoma virus life cycle. Front. Biosci. Landmark Ed. *14*, 4968–4977.

Huang, Y., and Steitz, J.A. (2005). SRprises along a messenger's journey. Mol. Cell *17*, 613–615.

Huang, F., Chen, J., Zhang, J., Tan, L., Lu, G., Luo, Y., Pan, T., Liang, J., Li, Q., Luo, B., et al. (2018). Identification of a novel compound targeting the nuclear export of influenza A virus nucleoprotein. J. Cell. Mol. Med. *22*, 1826–1839.

Huang, S., Chen, J., Chen, Q., Wang, H., Yao, Y., Chen, J., and Chen, Z. (2013). A second CRM1-dependent nuclear export signal in the influenza A virus NS2 protein contributes to the nuclear export of viral ribonucleoproteins. J. Virol. *87*, 767–778.

Huang, Y., Gattoni, R., Stévenin, J., and Steitz, J.A. (2003). SR splicing factors serve as adapter proteins for TAP-dependent mRNA export. Mol. Cell *11*, 837–843.

Huang, Z., Andrianov, V.M., Han, Y., and Howell, S.H. (2001). Identification of arabidopsis proteins that interact with the cauliflower mosaic virus (CaMV) movement protein. Plant Mol. Biol. *47*, 663–675.

Huang, Z.M., Zang, W.Q., and Yen, T.S. (1996). Cellular proteins that bind to the hepatitis B virus posttranscriptional regulatory element. Virology *217*, 573–581.

Huertas, P., García-Rubio, M.L., Wellinger, R.E., Luna, R., and Aguilera, A. (2006). An hpr1 Point Mutation That Impairs Transcription and mRNP Biogenesis without Increasing Recombination. Mol. Cell. Biol. *26*, 7451–7465.

Hull, R., Covey, S.N., and Maule, A.J. (1987). Structure and replication of caulimovirus genomes. J. Cell Sci. Suppl. *7*, 213–229.

Hung, N.-J., Lo, K.-Y., Patel, S.S., Helmke, K., and Johnson, A.W. (2008). Arx1 Is a Nuclear Export Receptor for the 60S Ribosomal Subunit in Yeast. Mol. Biol. Cell *19*, 735–744.

Hunter, C.A., Aukerman, M.J., Sun, H., Fokina, M., and Poethig, R.S. (2003). PAUSED encodes the Arabidopsis exportin-t ortholog. Plant Physiol. *132*, 2135–2143.

Hurt, E., Hannus, S., Schmelzl, B., Lau, D., Tollervey, D., and Simos, G. (1999). A novel in vivo assay reveals inhibition of ribosomal nuclear export in ran-cycle and nucleoporin mutants. J. Cell Biol. *144*, 389–401.

Hurt, E., Luo, M.-J., Röther, S., Reed, R., and Strässer, K. (2004). Cotranscriptional recruitment of the serine-arginine-rich (SR)-like proteins Gbp2 and Hrb1 to nascent mRNA via the TREX complex. Proc. Natl. Acad. Sci. U. S. A. *101*, 1858–1862.

Ishizawa, J., Kojima, K., Hail, N., Tabe, Y., and Andreeff, M. (2015). Expression, function, and targeting of the nuclear exporter chromosome region maintenance 1 (CRM1) protein. Pharmacol. Ther. *153*, 25–35.

Iwatsuki-Horimoto, K., Horimoto, T., Fujii, Y., and Kawaoka, Y. (2004). Generation of influenza A virus NS2 (NEP) mutants with an altered nuclear export signal sequence. J. Virol. *78*, 10149–10155.

Izaurralde, E., Lewis, J., Gamberi, C., Jarmolowski, A., McGuigan, C., and Mattaj, I.W. (1995). A cap-binding protein complex mediating U snRNA export. Nature *376*, 709–712.

Jackson, B.R., Boyne, J.R., Noerenberg, M., Taylor, A., Hautbergue, G.M., Walsh, M.J., Wheat, R., Blackbourn, D.J., Wilson, S.A., and Whitehouse, A. (2011). An interaction between KSHV ORF57 and UIF provides mRNA-adaptor redundancy in herpesvirus intronless mRNA export. PLoS Pathog. *7*, e1002138.

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.

Jani, D., Lutz, S., Marshall, N.J., Fischer, T., Köhler, A., Ellisdon, A.M., Hurt, E., and Stewart, M. (2009). Sus1, Cdc31, and the Sac3 CID Region Form a Conserved Interaction Platform that Promotes Nuclear Pore Association and mRNA Export. Mol. Cell *33*, 727–737.

Jani, D., Valkov, E., and Stewart, M. (2014). Structural basis for binding the TREX2 complex to nuclear pores, GAL1 localisation and mRNA export. Nucleic Acids Res. *42*, 6686–6697.

Jauvion, V., Elmayan, T., and Vaucheret, H. (2010). The Conserved RNA Trafficking Proteins HPR1 and TEX1 Are Involved in the Production of Endogenous and Exogenous Small Interfering RNA in Arabidopsis[C][W]. Plant Cell *22*, 2697–2709.

Jiao, X., Xiang, S., Oh, C., Martin, C.E., Tong, L., and Kiledjian, M. (2010). Identification of a quality-control mechanism for mRNA 5'-end capping. Nature *467*, 608–611.

Johnson, L.A., and Sandri-Goldin, R.M. (2009). Efficient nuclear export of herpes simplex virus 1 transcripts requires both RNA binding by ICP27 and ICP27 interaction with TAP/NXF1. J. Virol. *83*, 1184–1192.

Johnson, L.A., Li, L., and Sandri-Goldin, R.M. (2009a). The cellular RNA export receptor TAP/NXF1 is required for ICP27-mediated export of herpes simplex virus 1 RNA, but the TREX complex adaptor protein Aly/REF appears to be dispensable. J. Virol. *83*, 6335–6346.

Johnson, S.A., Cubberley, G., and Bentley, D.L. (2009b). Cotranscriptional recruitment of the mRNA export factor Yra1 by direct interaction with the 3' end processing factor Pcf11. Mol. Cell *33*, 215–226.

Jouvenet, N., Lainé, S., Vivares, L.P., and Mougel, M. (2011). Cell biology of retroviral RNA packaging. RNA Biol. *8*, 572–580.

Juillard, F., Hiriart, E., Sergeant, N., Vingtdeux-Didier, V., Drobecq, H., Sergeant, A., Manet, E., and Gruffat, H. (2009). Epstein-Barr virus protein EB2 contains an N-terminal transferable nuclear export signal that promotes nucleocytoplasmic export by directly binding TAP/NXF1. J. Virol. *83*, 12759–12768.

Kabachinski, G., and Schwartz, T.U. (2015). The nuclear pore complex – structure and function at a glance. J. Cell Sci. *128*, 423–429.

Kammel, C., Thomaier, M., Sørensen, B.B., Schubert, T., Längst, G., Grasser, M., and Grasser, K.D. (2013). Arabidopsis DEAD-box RNA helicase UAP56 interacts with both RNA and DNA as well as with mRNA export factors. PloS One *8*, e60644.

Kang, Y., and Cullen, B.R. (1999). The human Tap protein is a nuclear mRNA export factor that contains novel RNA-binding and nucleocytoplasmic transport sequences. Genes Dev. *13*, 1126–1139.

Kann, M., Sodeik, B., Vlachou, A., Gerlich, W.H., and Helenius, A. (1999). Phosphorylation-dependent binding of hepatitis B virus core particles to the nuclear pore complex. J. Cell Biol. *145*, 45–55.

Karn, J., and Stoltzfus, C.M. (2012). Transcriptional and posttranscriptional regulation of HIV-1 gene expression. Cold Spring Harb. Perspect. Med. *2*, a006916.

Karsies, A., Merkle, T., Szurek, B., Bonas, U., Hohn, T., and Leclerc, D. (2002). Regulated nuclear targeting of cauliflower mosaic virus. J. Gen. Virol. *83*, 1783–1790.

Katahira, J. (2012). mRNA export and the TREX complex. Biochim. Biophys. Acta *1819*, 507–513.

Katahira, J. (2015). Nuclear export of messenger RNA. Genes *6*, 163–184.

Katahira, J., Inoue, H., Hurt, E., and Yoneda, Y. (2009). Adaptor Aly and co-adaptor Thoc5 function in the Tap-p15-mediated nuclear export of HSP70 mRNA. EMBO J. *28*, 556–567.

Katahira, J., Okuzaki, D., Inoue, H., Yoneda, Y., Maehara, K., and Ohkawa, Y. (2013). Human TREX component Thoc5 affects alternative polyadenylation site choice by recruiting mammalian cleavage factor I. Nucleic Acids Res. *41*, 7060–7072.

Kelley, K., Knockenhauer, K.E., Kabachinski, G., and Schwartz, T.U. (2015). Atomic structure of the Y complex of the nuclear pore. Nat. Struct. Mol. Biol. *22*, 425–431.

Keskitalo, M. (2001). Can protoplast production from in vitro cultured shoots of Tanacetum vary during the season? Agric. Food Sci. *10*, 145–151.

Khelifa, M., Journou, S., Krishnan, K., Gargani, D., Espérandieu, P., Blanc, S., and Drucker, M. (2007). Electron-lucent inclusion bodies are structures specialized for aphid transmission of cauliflower mosaic virus. J. Gen. Virol. *88*, 2872–2880.

Khelifa, M., Massé, D., Blanc, S., and Drucker, M. (2010). Evaluation of the minimal replication time of Cauliflower mosaic virus in different hosts. Virology *396*, 238–245.

Kiledjian, M., Zhou, M., and Jiao, X. (2012). Normal and Aberrantly Capped mRNA Decapping. The Enzymes *31*, 165–180.

Kim, Y.-K., Kim, B., and Kim, V.N. (2016). Re-evaluation of the roles of DROSHA, Exportin 5, and DICER in microRNA biogenesis. Proc. Natl. Acad. Sci. *113*, E1881–E1889.

Kimura, T., Hashimoto, I., Nagase, T., and Fujisawa, J.-I. (2004). CRM1-dependent, but not ARE-mediated, nuclear export of IFN-alpha1 mRNA. J. Cell Sci. *117*, 2259–2270.

Kiss, T. (2004). Biogenesis of small nuclear RNPs. J. Cell Sci. *117*, 5949–5951.

Kiss-László, Z., Blanc, S., and Hohn, T. (1995). Splicing of cauliflower mosaic virus 35S RNA is essential for viral infectivity. EMBO J. *14*, 3552–3562.

Kırlı, K., Karaca, S., Dehne, H.J., Samwer, M., Pan, K.T., Lenz, C., Urlaub, H., and Görlich, D. (2015). A deep proteomics perspective on CRM1-mediated nuclear export and nucleocytoplasmic partitioning. ELife *4*.

Knipe, D., and Howley, P. (2013). Fields virology: Sixth edition.

Knockenhauer, K.E., and Schwartz, T.U. (2016). The Nuclear Pore Complex as a Flexible and Dynamic Gate. Cell *164*, 1162–1171.

Kobayashi, K., Nakayashiki, H., Tsuge, S., Mise, K., and Furusawa, I. (1998). Accumulation kinetics of viral gene products in cauliflower mosaic virus-infected turnip protoplasts. Microbiol. Immunol. *42*, 65–69.

Koffa, M.D., Clements, J.B., Izaurralde, E., Wadd, S., Wilson, S.A., Mattaj, I.W., and Kuersten, S. (2001). Herpes simplex virus ICP27 protein provides viral mRNAs with access to the cellular mRNA export pathway. EMBO J. *20*, 5769–5778.

Köhler, A., and Hurt, E. (2007). Exporting RNA from the nucleus to the cytoplasm. Nat. Rev. Mol. Cell Biol. *8*, 761–773.

Kramer, E.B., and Hopper, A.K. (2013). Retrograde transfer RNA nuclear import provides a new level of tRNA quality control in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U. S. A. *110*, 21042–21047.

Kshirsagar, M., and Parker, R. (2004). Identification of Edc3p as an enhancer of mRNA decapping in Saccharomyces cerevisiae. Genetics *166*, 729–739.

Kudo, N., Matsumori, N., Taoka, H., Fujiwara, D., Schreiner, E.P., Wolff, B., Yoshida, M., and Horinouchi, S. (1999). Leptomycin B inactivates CRM1/exportin 1 by covalent modification at a cysteine residue in the central conserved region. Proc. Natl. Acad. Sci. U. S. A. 96, 9112–9117.

Kühn, U., Buschmann, J., and Wahle, E. (2017). The nuclear poly(A) binding protein of mammals, but not of fission yeast, participates in mRNA polyadenylation. RNA N. Y. N *23*, 473–482.

Kurshakova, M.M., Krasnov, A.N., Kopytova, D.V., Shidlovskii, Y.V., Nikolenko, J.V., Nabirochkina, E.N., Spehner, D., Schultz, P., Tora, L., and Georgieva, S.G. (2007). SAGA and a novel Drosophila export complex anchor efficient transcription and mRNA export to NPC. EMBO J. 26, 4956–4965.

Kurshakova, M.M., Georgieva, S.G., and Kopytova, D.V. (2016). [Protein complexes coordinating mRNA export from the nucleus into the cytoplasm]. Mol. Biol. (Mosk.) *50*, 723–729.

Kuss, S.K., Mata, M.A., Zhang, L., and Fontoura, B.M.A. (2013). Nuclear Imprisonment: Viral Strategies to Arrest Host mRNA Nuclear Export. Viruses *5*, 1824–1849.

Kutay, U., Bischoff, F.R., Kostka, S., Kraft, R., and Görlich, D. (1997). Export of importin alpha from the nucleus is mediated by a specific nuclear transport factor. Cell *90*, 1061–1071.

Lakdawala, S.S., Fodor, E., and Subbarao, K. (2016). Moving On Out: Transport and Packaging of Influenza Viral RNA into Virions. Annu. Rev. Virol. *3*, 411–427.

Lallemand-Breitenbach, V., and de Thé, H. (2018). PML nuclear bodies: from architecture to function. Curr. Opin. Cell Biol. *52*, 154–161.

Lamb, R.A., and Choppin, P.W. (1981). Identification of a second protein (M2) encoded by RNA segment 7 of influenza virus. Virology *112*, 729–737.

Lamb, R.A., and Horvath, C.M. (1991). Diversity of coding strategies in influenza viruses. Trends Genet. TIG *7*, 261–266.

Lamb, R.A., and Lai, C.J. (1980). Sequence of interrupted and uninterrupted mRNAs and cloned DNA coding for the two overlapping nonstructural proteins of influenza virus. Cell *21*, 475–485.

Le Sage, V., and Mouland, A.J. (2013). Viral Subversion of the Nuclear Pore Complex. Viruses *5*, 2019–2042.

LeBlanc, J.J., Uddowla, S., Abraham, B., Clatterbuck, S., and Beemon, K.L. (2007). Tap and Dbp5, but not Gag, are involved in DR-mediated nuclear export of unspliced Rous sarcoma virus RNA. Virology *363*, 376–386.

Leclerc, D., Chapdelaine, Y., and Hohn, T. (1999). Nuclear targeting of the cauliflower mosaic virus coat protein. J. Virol. *73*, 553–560.

Lee, G.H., Wasser, S., and Lim, S.G. (2008a). Hepatitis B pregenomic RNA splicing—The products, the regulatory mechanisms and its biological significance. Virus Res. *136*, 1–7.

Lee, H.-J., Lee, J., Shin, M.-K., and Ryu, W.-S. (2008b). Polyadenylation is dispensable for encapsidation and reverse transcription of hepatitis B viral pregenomic RNA. Mol. Cells *25*, 545–552.

Leh, V., Jacquot, E., Geldreich, A., Hermann, T., Leclerc, D., Cerutti, M., Yot, P., Keller, M., and Blanc, S. (1999). Aphid transmission of cauliflower mosaic virus requires the viral PIII protein. EMBO J. *18*, 7077–7085.

Leh, V., Yot, P., and Keller, M. (2000). The cauliflower mosaic virus translational transactivator interacts with the 60S ribosomal subunit protein L18 of Arabidopsis thaliana. Virology 266, 1–7.

Leh, V., Jacquot, E., Geldreich, A., Haas, M., Blanc, S., Keller, M., and Yot, P. (2001). Interaction between the Open Reading Frame III Product and the Coat Protein Is Required for Transmission of Cauliflower Mosaic Virus by Aphids. J. Virol. *75*, 100–106.

Lei, H., Dias, A.P., and Reed, R. (2011). Export and stability of naturally intronless mRNAs require specific coding region sequences and the TREX mRNA export complex. Proc. Natl. Acad. Sci. U. S. A. *108*, 17985–17990.

Leisner, S.M., and Schoelz, J.E. (2018). Joining the Crowd: Integrating Plant Virus Proteins into the Larger World of Pathogen Effectors. Annu. Rev. Phytopathol. *56*, 89–110.

Lejeune, F. (2017). Nonsense-mediated mRNA decay at the crossroads of many cellular pathways. BMB Rep. *50*, 175–185.

Lellis, A.D., Kasschau, K.D., Whitham, S.A., and Carrington, J.C. (2002). Loss-of-susceptibility mutants of Arabidopsis thaliana reveal an essential role for eIF(iso)4E during potyvirus infection. Curr. Biol. CB *12*, 1046–1051.

Leppard, K. (2014). Adenoviruses: Molecular Biology. Ref. Module Biomed. Sci.

Leppard, K.N. (1993). Selective effects on adenovirus late gene expression of deleting the E1b 55K protein. J. Gen. Virol. *74 (Pt 4)*, 575–582.

Leppard, K.N., and Everett, R.D. (1999). The adenovirus type 5 E1b 55K and E4 Orf3 proteins associate in infected cells and affect ND10 components. J. Gen. Virol. *80 (Pt 4)*, 997–1008.

Leung, A.K.L. (2015). The Whereabouts of microRNA Actions: Cytoplasm and Beyond. Trends Cell Biol. *25*, 601–610.

Li, H., and Yang, Z. (2000). RHO Gtpases and the Actin Cytoskeleton. In Actin: A Dynamic Framework for Multiple Plant Cell Functions, C.J. Staiger, F. Baluska, D. Volkmann, and P. Barlow, Eds (Dordrecht, The Netherlands: Kluwer Academic Press), pp. 301–321.

Li, R.H., and Thomas, J.O. (1989). Identification of a human protein that interacts with nuclear localization signals. J. Cell Biol. *109*, 2623–2632.

Li, H., Shen, J.J., Zheng, Z.L., Lin, Y., and Yang, Z. (2001). The Rop GTPase switch controls multiple developmental processes in Arabidopsis. Plant Physiol. *126*, 670–684.

Li, H.-C., Huang, E.-Y., Su, P.-Y., Wu, S.-Y., Yang, C.-C., Lin, Y.-S., Chang, W.-C., and Shih, C. (2010a). Nuclear export and import of human hepatitis B virus capsid protein and particles. PLoS Pathog. 6, e1001162.

Li, J., Lin, S., Chen, Q., Peng, L., Zhai, J., Liu, Y., and Yuan, Z. (2010b). Inhibition of hepatitis B virus replication by MyD88 involves accelerated degradation of pregenomic RNA and nuclear retention of pre-S/S RNAs. J. Virol. *84*, 6387–6399.

Li, Y., Bor, Y.-C., Misawa, Y., Xue, Y., Rekosh, D., and Hammarskjöld, M.-L. (2006). An intron with a constitutive transport element is retained in a Tap messenger RNA. Nature *443*, 234–237.

Liker, E., Fernandez, E., Izaurralde, E., and Conti, E. (2000). The structure of the mRNA export factor TAP reveals a cis arrangement of a non-canonical RNP domain and an LRR domain. EMBO J. *19*, 5587–5598.

Lim, C.S., and Brown, C.M. (2016). Hepatitis B virus nuclear export elements: RNA stem-loop α and β, key parts of the HBV post-transcriptional regulatory element. RNA Biol. *13*, 743–747.

Lin, D.H., Correia, A.R., Cai, S.W., Huber, F.M., Jette, C.A., and Hoelz, A. (2018). Structural and functional analysis of mRNA export regulation by the nuclear pore complex. Nat. Commun. *9*, 2319.

Lindemann, D., and Rethwilm, A. (2011). Foamy Virus Biology and Its Application for Vector Development. Viruses *3*, 561–585.

Lindsay, M.E., Holaska, J.M., Welch, K., Paschal, B.M., and Macara, I.G. (2001). Ran-Binding Protein 3 Is a Cofactor for Crm1-Mediated Nuclear Protein Export. J. Cell Biol. *153*, 1391–1402.

Link, D., Schmidlin, L., Schirmer, A., Klein, E., Erhardt, M., Geldreich, A., Lemaire, O., and Gilmer, D. (2005). Functional characterization of the Beet necrotic yellow vein virus RNA-5-encoded p26 protein: evidence for structural pathogenicity determinants. J. Gen. Virol. *86*, 2115–2125.

Lipowsky, G., Bischoff, F.R., Schwarzmaier, P., Kraft, R., Kostka, S., Hartmann, E., Kutay, U., and Görlich, D. (2000). Exportin 4: a mediator of a novel nuclear export pathway in higher eukaryotes. EMBO J. *19*, 4362–4371.

Lippai, M., Tirián, L., Boros, I., Mihály, J., Erdélyi, M., Belecz, I., Máthé, E., Pósfai, J., Nagy, A., Udvardy, A., et al. (2000). The Ketel gene encodes a Drosophila homologue of importin-beta. Genetics *156*, 1889–1900.

Löchelt, M. (2003). Foamy virus transactivation and gene expression. Curr. Top. Microbiol. Immunol. *277*, 27–61.

Lorković, Z.J., Wieczorek Kirk, D.A., Lambermon, M.H., and Filipowicz, W. (2000). Pre-mRNA splicing in higher plants. Trends Plant Sci. *5*, 160–167.

Love, A.J., Laird, J., Holt, J., Hamilton, A.J., Sadanandom, A., and Milner, J.J. (2007). Cauliflower mosaic virus protein P6 is a suppressor of RNA silencing. J. Gen. Virol. *88*, 3439–3444.

Lu, Q., Tang, X., Tian, G., Wang, F., Liu, K., Nguyen, V., Kohalmi, S.E., Keller, W.A., Tsang, E.W.T., Harada, J.J., et al. (2010). Arabidopsis homolog of the yeast TREX-2 mRNA export complex: components and anchoring nucleoporin. Plant J. Cell Mol. Biol. *61*, 259–270.

Lührmann, R. (1990). Functions of U-snRNPs. Mol. Biol. Rep. *14*, 183–192.

Lund, E., and Dahlberg, J.E. (1998). Proofreading and aminoacylation of tRNAs before export from the nucleus. Science *282*, 2082–2085.

Lund, E., Güttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. (2004). Nuclear Export of MicroRNA Precursors. Science *303*, 95–98.

Lutz, L., Raikhy, G., and Leisner, S.M. (2012). Cauliflower mosaic virus major inclusion body protein interacts with the aphid transmission factor, the virion-associated protein, and gene VII product. Virus Res. *170*, 150–153.

Mahiet, C., and Swanson, C.M. (2016). Control of HIV-1 gene expression by SR proteins. Biochem. Soc. Trans. *44*, 1417–1425.

Majerciak, V., and Zheng, Z.-M. (2015). KSHV ORF57, a protein of many faces. Viruses *7*, 604–633.

Majerciak, V., Uranishi, H., Kruhlak, M., Pilkington, G.R., Massimelli, M.J., Bear, J., Pavlakis, G.N., Felber, B.K., and Zheng, Z.-M. (2011). Kaposi's sarcoma-associated herpesvirus ORF57 interacts with cellular RNA export cofactors RBM15 and OTT3 to promote expression of viral ORF59. J. Virol. *85*, 1528–1540.

Makokha, G.N., Abe-Chayama, H., Chowdhury, S., Hayes, C.N., Tsuge, M., Yoshima, T., Ishida, Y., Zhang, Y., Uchida, T., Tateno, C., et al. (2019). Regulation of the Hepatitis B virus replication and gene expression by the multi-functional protein TARDBP. Sci. Rep. *9*, 8462.

Malik, P., Tabarraei, A., Kehlenbach, R.H., Korfali, N., Iwasawa, R., Graham, S.V., and Schirmer, E.C. (2012). Herpes simplex virus ICP27 protein directly interacts with the nuclear pore complex through Nup62, inhibiting host nucleocytoplasmic transport pathways. J. Biol. Chem. *287*, 12277–12292.

Malim, M.H., McCarn, D.F., Tiley, L.S., and Cullen, B.R. (1991). Mutational definition of the human immunodeficiency virus type 1 Rev activation domain. J. Virol. *65*, 4248–4254.

Marsh, L.E., and Guilfoyle, T.J. (1987). Cauliflower mosaic virus replication intermediates are encapsidated into virion-like particles. Virology *161*, 129–137.

Martelli, G.P., and Russo, M. (1977). Plant virus inclusion bodies. Adv. Virus Res. *21*, 175–266.

Martinière, A., Gargani, D., Uzest, M., Lautredou, N., Blanc, S., and Drucker, M. (2009). A role for plant microtubules in the formation of transmission-specific inclusion bodies of Cauliflower mosaic virus. Plant J. Cell Mol. Biol. *58*, 135–146.

Martinière, A., Bak, A., Macia, J.-L., Lautredou, N., Gargani, D., Doumayrou, J., Garzo, E., Moreno, A., Fereres, A., Blanc, S., et al. (2013). A virus responds instantly to the presence of the vector on the host and forms transmission morphs. ELife *2*.

Marzluff, W.F., Wagner, E.J., and Duronio, R.J. (2008). Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. Nat. Rev. Genet. *9*, 843–854.

Massimelli, M.J., Kang, J.-G., Majerciak, V., Le, S.-Y., Liewehr, D.J., Steinberg, S.M., and Zheng, Z.-M. (2011). Stability of a long noncoding viral RNA depends on a 9-nt core element at the RNA 5' end to interact with viral ORF57 and cellular PABPC1. Int. J. Biol. Sci. *7*, 1145–1160.

Massimelli, M.J., Majerciak, V., Kang, J.-G., Liewehr, D.J., Steinberg, S.M., and Zheng, Z.-M. (2015). Multiple regions of Kaposi's sarcoma-associated herpesvirus ORF59 RNA are required for its expression mediated by viral ORF57 and cellular RBM15. Viruses *7*, 496–510.

Masuda, S., Das, R., Cheng, H., Hurt, E., Dorman, N., and Reed, R. (2005). Recruitment of the human TREX complex to mRNA during splicing. Genes Dev. *19*, 1512–1517.

Mathew, C., and Ghildyal, R. (2017). CRM1 Inhibitors for Antiviral Therapy. Front. Microbiol. *8*, 1171.

Maul, G.G., Maul, H.M., Scogna, J.E., Lieberman, M.W., Stein, G.S., Hsu, B.Y., and Borun, T.W. (1972). Time sequence of nuclear pore formation in phytohemagglutinin-stimulated lymphocytes and in HeLa cells during the cell cycle. J. Cell Biol. *55*, 433–447.

McCloskey, A., Taniguchi, I., Shinmyozu, K., and Ohno, M. (2012). hnRNP C tetramer measures RNA length to classify RNA polymerase II transcripts for export. Science *335*, 1643–1646.

Mears, W.E., and Rice, S.A. (1996). The RGG box motif of the herpes simplex virus ICP27 protein mediates an RNA-binding activity and determines in vivo methylation. J. Virol. *70*, 7445–7453.

Medlin, J.E., Uguen, P., Taylor, A., Bentley, D.L., and Murphy, S. (2003). The C-terminal domain of pol II and a DRB-sensitive kinase are required for 3' processing of U2 snRNA. EMBO J. *22*, 925–934.

Meier, I., Richards, E.J., and Evans, D.E. (2017). Cell Biology of the Plant Nucleus. Annu. Rev. Plant Biol. *68*, 139–172.

Meinel, D.M., Burkert-Kautzsch, C., Kieser, A., O'Duibhir, E., Siebert, M., Mayer, A., Cramer, P., Söding, J., Holstege, F.C.P., and Sträßer, K. (2013). Recruitment of TREX to the transcription machinery by its direct binding to the phospho-CTD of RNA polymerase II. PLoS Genet. *9*, e1003914.

Melchior, F., Paschal, B., Evans, J., and Gerace, L. (1993). Inhibition of nuclear protein import by nonhydrolyzable analogues of GTP and identification of the small GTPase Ran/TC4 as an essential transport factor. J. Cell Biol. *123*, 1649–1659.

Meng, W., Wang, X.-J., and Wang, H.-C.R. (2019). Targeting nuclear proteins for control of viral replication. Crit. Rev. Microbiol. 1–19.

Merkle, T. (2011). Nucleo-cytoplasmic transport of proteins and RNA in plants. Plant Cell Rep. *30*, 153–176.

Merkle, T., Leclerc, D., Marshallsay, C., and Nagy, F. (1996). A plant in vitro system for the nuclear import of proteins. Plant J. Cell Mol. Biol. *10*, 1177–1186.

Mingot, J.-M., Bohnsack, M.T., Jäkle, U., and Görlich, D. (2004). Exportin 7 defines a novel general nuclear export pathway. EMBO J. *23*, 3227–3236.

Mitra, B., Thapa, R.J., Guo, H., and Block, T.M. (2018). Host functions used by hepatitis B virus to complete its life cycle: Implications for developing host-targeting agents to treat chronic hepatitis B. Antiviral Res. *158*, 185–198.

Miyazaki, Y., Irobalieva, R.N., Tolbert, B.S., Smalls-Mantey, A., Iyalla, K., Loeliger, K., D'Souza, V., Khant, H., Schmid, M.F., Garcia, E.L., et al. (2010). Structure of a conserved retroviral RNA packaging element by NMR spectroscopy and cryo-electron tomography. J. Mol. Biol. *404*, 751–772.

Monette, A., Panté, N., and Mouland, A.J. (2011). HIV-1 remodels the nuclear pore complex. J. Cell Biol. *193*, 619–631.

Monsion, B. (2008). FLUCTUATIONS DEMOGRAPHIQUES AU COURS DU CYCLE DE VIE DU CaMV (Cauliflower mosaic virus). Estimation de la taille efficace des populations virales lors de la colonisation des feuilles de la plante hôte, évaluation de la multiplicité d'infection cellulaire au sein de ces feuilles, et estimation de la taille des goulots d'étranglement lors de sa transmission d'hôte à hôte par vecteur. thesis. Montpellier 2.

Moore, M.S., and Blobel, G. (1993). The GTP-binding protein Ran/TC4 is required for protein import into the nucleus. Nature *365*, 661–663.

Morita, M., Kuba, K., Ichikawa, A., Nakayama, M., Katahira, J., Iwamoto, R., Watanebe, T., Sakabe, S., Daidoji, T., Nakamura, S., et al. (2013). The lipid mediator protectin D1 inhibits influenza virus replication and improves severe influenza. Cell *153*, 112–125.

Moroianu, J., Blobel, G., and Radu, A. (1997). RanGTP-mediated nuclear export of karyopherin alpha involves its interaction with the nucleoporin Nup153. Proc. Natl. Acad. Sci. U. S. A. *94*, 9699–9704.

Moschall, R., Denk, S., Erkelenz, S., Schenk, C., Schaal, H., and Bodem, J. (2017). A purine-rich element in foamy virus pol regulates env splicing and gag/pol expression. Retrovirology *14*, 10.

Mougel, M., Zhang, Y., and Barklis, E. (1996). cis-active structural motifs involved in specific encapsidation of Moloney murine leukemia virus RNA. J. Virol. *70*, 5043–5050.

Müller-McNicoll, M., and Neugebauer, K.M. (2013). How cells get the message: dynamic assembly and function of mRNA-protein complexes. Nat. Rev. Genet. *14*, 275–287.

Müllner, M., Salmons, B., Günzburg, W.H., and Indik, S. (2008). Identification of the Rem-responsive element of mouse mammary tumor virus. Nucleic Acids Res. *36*, 6284–6294.

Nakagawa, T., Kurose, T., Hino, T., Tanaka, K., Kawamukai, M., Niwa, Y., Toyooka, K., Matsuoka, K., Jinbo, T., and Kimura, T. (2007). Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. J. Biosci. Bioeng. *104*, 34–41.

Nakano, K., and Watanabe, T. (2016). HTLV-1 Rex Tunes the Cellular Environment Favorable for Viral Replication. Viruses *8*, 58.

Nerurkar, P., Altvater, M., Gerhardy, S., Schütz, S., Fischer, U., Weirich, C., and Panse, V.G. (2015). Eukaryotic Ribosome Assembly and Nuclear Export. Int. Rev. Cell Mol. Biol. *319*, 107–140.

Netherton, C.L., and Wileman, T. (2011). Virus factories, double membrane vesicles and viroplasm generated in animal cells. Curr. Opin. Virol. *1*, 381–387.

Neumann, G., Hughes, M.T., and Kawaoka, Y. (2000). Influenza A virus NS2 protein mediates vRNP nuclear export through NES-independent interaction with hCRM1. EMBO J. *19*, 6751–6758.

Neumann, N., Lundin, D., and Poole, A.M. (2010). Comparative Genomic Evidence for a Complete Nuclear Pore Complex in the Last Eukaryotic Common Ancestor. PLoS ONE *5*.

Nevels, M., Nitzsche, A., and Paulus, C. (2011). How to control an infectious bead string: nucleosome-based regulation and targeting of herpesvirus chromatin. Rev. Med. Virol. *21*, 154–180.

Neville, M., Stutz, F., Lee, L., Davis, L.I., and Rosbash, M. (1997). The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export. Curr. Biol. CB *7*, 767–775.

Nitta, T., Hofacre, A., Hull, S., and Fan, H. (2009). Identification and Mutational Analysis of a Rej Response Element in Jaagsiekte Sheep Retrovirus RNA. J. Virol. *83*, 12499–12511.

Noh, J.H., Kim, K.M., McClusky, W.G., Abdelmohsen, K., and Gorospe, M. (2018). Cytoplasmic functions of long noncoding RNAs. Wiley Interdiscip. Rev. RNA *9*, e1471.

Ohira, T., and Suzuki, T. (2011). Retrograde nuclear import of tRNA precursors is required for modified base biogenesis in yeast. Proc. Natl. Acad. Sci. U. S. A. *108*, 10502–10507.

Ohno, M. (2012). Size matters in RNA export. RNA Biol. *9*, 1413–1417.

Ohno, M., Segref, A., Bachi, A., Wilm, M., and Mattaj, I.W. (2000). PHAX, a Mediator of U snRNA Nuclear Export Whose Activity Is Regulated by Phosphorylation. Cell *101*, 187–198.

Ohtani, M. (2017). Plant snRNP Biogenesis: A Perspective from the Nucleolus and Cajal Bodies. Front. Plant Sci. *8*, 2184.

Oka, M., and Yoneda, Y. (2018). Importin α: functions as a nuclear transport factor and beyond. Proc. Jpn. Acad. Ser. B Phys. Biol. Sci. *94*, 259–274.

Okada, C., Yamashita, E., Lee, S.J., Shibata, S., Katahira, J., Nakagawa, A., Yoneda, Y., and Tsukihara, T. (2009). A High-Resolution Structure of the Pre-microRNA Nuclear Export Machinery. Science *326*, 1275–1279.

Okamura, M., Inose, H., and Masuda, S. (2015). RNA Export through the NPC in Eukaryotes. Genes *6*, 124–149.

Olszewski, N., Hagen, G., and Guilfoyle, T.J. (1982). A transcriptionally active, covalently closed minichromosome of cauliflower mosaic virus DNA isolated from infected turnip leaves. Cell *29*, 395–402.

O'Malley, R.C., Barragan, C.C., and Ecker, J.R. (2015). A User's Guide to the Arabidopsis T-DNA Insertional Mutant Collections. Methods Mol. Biol. Clifton NJ *1284*, 323–342.

O'Neill, R.E., Talon, J., and Palese, P. (1998). The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins. EMBO J. *17*, 288–296.

Ornelles, D.A., and Shenk, T. (1991). Localization of the adenovirus early region 1B 55-kilodalton protein during lytic infection: association with nuclear viral inclusions requires the early region 4 34-kilodalton protein. J. Virol. *65*, 424–429.

Ossareh-Nazari, B., Bachelerie, F., and Dargemont, C. (1997). Evidence for a role of CRM1 in signal-mediated nuclear protein export. Science *278*, 141–144.

Ossareh-Nazari, B., Maison, C., Black, B.E., Lévesque, L., Paschal, B.M., and Dargemont, C. (2000). RanGTP-binding protein NXT1 facilitates nuclear export of different classes of RNA in vitro. Mol. Cell. Biol. *20*, 4562–4571.

Ou, S.H., Wu, F., Harrich, D., García-Martínez, L.F., and Gaynor, R.B. (1995). Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. J. Virol. *69*, 3584–3596.

Palazzo, A.F., and Lee, E.S. (2018). Sequence Determinants for Nuclear Retention and Cytoplasmic Export of mRNAs and lncRNAs. Front. Genet. *9*, 440.

Pan, H., Liu, S., and Tang, D. (2012). HPR1, a component of the THO/TREX complex, plays an important role in disease resistance and senescence in Arabidopsis. Plant J. Cell Mol. Biol. *69*, 831–843.

Parent, L.J. (2011). New insights into the nuclear localization of retroviral Gag proteins. Nucl. Austin Tex *2*, 92–97.

Park, H.S., Himmelbach, A., Browning, K.S., Hohn, T., and Ryabova, L.A. (2001). A plant viral "reinitiation" factor interacts with the host translational machinery. Cell *106*, 723–733.

Pascual-García, P., and Rodríguez-Navarro, S. (2009). A tale of coupling, Sus1 function in transcription and mRNA export. RNA Biol. *6*, 141–144.

Pasquinelli, A.E., Ernst, R.K., Lund, E., Grimm, C., Zapp, M.L., Rekosh, D., Hammarskjöld, M.L., and Dahlberg, J.E. (1997). The constitutive transport element (CTE) of Mason-Pfizer monkey virus (MPMV) accesses a cellular mRNA export pathway. EMBO J. *16*, 7500–7510.

Patel, S.B., and Bellini, M. (2008). The assembly of a spliceosomal small nuclear ribonucleoprotein particle. Nucleic Acids Res. *36*, 6482–6493.

Patel, V., Dahlroth, S.-L., Rajakannan, V., Ho, H.T., Cornvik, T., and Nordlund, P. (2015). Structure of the C-Terminal Domain of the Multifunctional ICP27 Protein from Herpes Simplex Virus 1. J. Virol. *89*, 8828–8839.

Paterson, D., and Fodor, E. (2012). Emerging roles for the influenza A virus nuclear export protein (NEP). PLoS Pathog. *8*, e1003019.

Perbal, M.C., Thomas, C.L., and Maule, A.J. (1993). Cauliflower mosaic virus gene I product (P1) forms tubular structures which extend from the surface of infected protoplasts. Virology *195*, 281–285.

Pereira, C.F., Read, E.K.C., Wise, H.M., Amorim, M.J., and Digard, P. (2017). Influenza A Virus NS1 Protein Promotes Efficient Nuclear Export of Unspliced Viral M1 mRNA. J. Virol. *91*.

Pereira, C.F., Wise, H.M., Kurian, D., Pinto, R.M., Amorim, M.J., Gill, A.C., and Digard, P. (2018). Effects of mutations in the effector domain of influenza A virus NS1 protein. BMC Res. Notes *11*.

Pessel-Vivares, L., Ferrer, M., Lainé, S., and Mougel, M. (2014). MLV requires Tap/NXF1-dependent pathway to export its unspliced RNA to the cytoplasm and to express both spliced and unspliced RNAs. Retrovirology *11*, 21.

Pessel-Vivares, L., Houzet, L., Lainé, S., and Mougel, M. (2015a). Insights into the nuclear export of murine leukemia virus intron-containing RNA. RNA Biol. *12*, 942–949.

Pessel-Vivares, L., Houzet, L., Lainé, S., and Mougel, M. (2015b). Insights into the nuclear export of murine leukemia virus intron-containing RNA. RNA Biol. *12*, 942–949.

Pfaff, C., Ehrnsberger, H.F., Flores-Tornero, M., Sørensen, B.B., Schubert, T., Längst, G., Griesenbeck, J., Sprunck, S., Grasser, M., and Grasser, K.D. (2018). ALY RNA-Binding Proteins Are Required for Nucleocytosolic mRNA Transport and Modulate Plant Growth and Development1[OPEN]. Plant Physiol. *177*, 226–240.

Pfeiffer, P., and Hohn, T. (1983). Involvement of reverse transcription in the replication of cauliflower mosaic virus: a detailed model and test of some aspects. Cell *33*, 781–789.

Pfeiffer, P., Laquel, P., and Hohn, T. (1984). Cauliflower Mosaic Virus replication complexes: characterization of the associated enzymes and of the polarity of the DNA synthesized in vitro. Plant Mol. Biol. *3*, 261–270.

Pietrzak, M., and Hohn, T. (1987). Translation products of cauliflower mosaic virus ORF V, the coding region corresponding to the retrovirus pol gene. Virus Genes *1*, 83–96.

Pires de Mello, C.P., Bloom, D.C., and Paixão, I.C. (2016). Herpes simplex virus type-1: replication, latency, reactivation and its antiviral targets. Antivir. Ther. *21*, 277–286.

Pooggin, M.M., and Ryabova, L.A. (2018). Ribosome Shunting, Polycistronic Translation, and Evasion of Antiviral Defenses in Plant Pararetroviruses and Beyond. Front. Microbiol. 9.

Pooggin, M.M., Fütterer, J., Skryabin, K.G., and Hohn, T. (1999). A short open reading frame terminating in front of a stable hairpin is the conserved feature in pregenomic RNA leaders of plant pararetroviruses. J. Gen. Virol. *80 (Pt 8)*, 2217–2228.

Pooggin, M.M., Fütterer, J., Skryabin, K.G., and Hohn, T. (2001). Ribosome shunt is essential for infectivity of cauliflower mosaic virus. Proc. Natl. Acad. Sci. U. S. A. *98*, 886–891.

Purcell, D.F., and Martin, M.A. (1993). Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. J. Virol. *67*, 6365–6378.

Rabino, C., Aspegren, A., Corbin-Lickfett, K., and Bridge, E. (2000). Adenovirus Late Gene Expression Does Not Require a Rev-Like Nuclear RNA Export Pathway. J. Virol. *74*, 6684–6688.

Rahmani, K., and Dean, D.A. (2017). Leptomycin B alters the subcellular distribution of CRM1 (Exportin 1). Biochem. Biophys. Res. Commun. *488*, 253–258.

Rao, A.L.N. (2007). Preparation and inoculation of mesophyll protoplasts from monocotyledenous and dicotyledenous hosts. Curr. Protoc. Microbiol. *Chapter 16*, Unit 16D.2.

Rausch, J.W., and Le Grice, S.F.J. (2015). HIV Rev Assembly on the Rev Response Element (RRE): A Structural Perspective. Viruses *7*, 3053–3075.

Rausin, G., Tillemans, V., Stankovic, N., Hanikenne, M., and Motte, P. (2010). Dynamic nucleocytoplasmic shuttling of an Arabidopsis SR splicing factor: role of the RNA-binding domains. Plant Physiol. *153*, 273–284.

Read, E.K.C., and Digard, P. (2010). Individual influenza A virus mRNAs show differential dependence on cellular NXF1/TAP for their nuclear export. J. Gen. Virol. *91*, 1290–1301.

Reddy, A.S.N. (2007). Alternative splicing of pre-messenger RNAs in plants in the genomic era. Annu. Rev. Plant Biol. *58*, 267–294.

Rehwinkel, J., and Reis e Sousa, C. (2010). RIGorous detection: exposing virus through RNA sensing. Science *327*, 284–286.

Rehwinkel, J., Herold, A., Gari, K., Köcher, T., Rode, M., Ciccarelli, F.L., Wilm, M., and Izaurralde, E. (2004). Genome-wide analysis of mRNAs regulated by the THO complex in Drosophila melanogaster. Nat. Struct. Mol. Biol. *11*, 558–566.

Reichelt, R., Holzenburg, A., Buhle, E.L., Jarnik, M., Engel, A., and Aebi, U. (1990). Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components. J. Cell Biol. *110*, 883–894.

Ribbeck, K., and Görlich, D. (2001). Kinetic analysis of translocation through nuclear pore complexes. EMBO J. *20*, 1320–1330.

Robb, N.C., and Fodor, E. (2012). The accumulation of influenza A virus segment 7 spliced mRNAs is regulated by the NS1 protein. J. Gen. Virol. *93*, 113–118.

Rodriguez, A., Angel, C.A., Lutz, L., Leisner, S.M., Nelson, R.S., and Schoelz, J.E. (2014). Association of the P6 protein of Cauliflower mosaic virus with plasmodesmata and plasmodesmal proteins. Plant Physiol. *166*, 1345–1358.

Rojas-Araya, B., Ohlmann, T., and Soto-Rifo, R. (2015). Translational Control of the HIV Unspliced Genomic RNA. Viruses *7*, 4326–4351.

Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K., and Weisshaar, B. (2003). An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. Plant Mol. Biol. *53*, 247–259.

Rougemaille, M., Dieppois, G., Kisseleva-Romanova, E., Gudipati, R.K., Lemoine, S., Blugeon, C., Boulay, J., Jensen, T.H., Stutz, F., Devaux, F., et al. (2008). THO/Sub2p functions to coordinate 3'-end processing with gene-nuclear pore association. Cell *135*, 308–321.

Rouquette, J., Choesmel, V., and Gleizes, P.-E. (2005). Nuclear export and cytoplasmic processing of precursors to the 40S ribosomal subunits in mammalian cells. EMBO J. *24*, 2862–2872.

Rubenwolf, S., Schütt, H., Nevels, M., Wolf, H., and Dobner, T. (1997). Structural analysis of the adenovirus type 5 E1B 55-kilodalton-E4orf6 protein complex. J. Virol. *71*, 1115–1123.

Ryabova, L.A., Pooggin, M.M., and Hohn, T. (2006). Translation reinitiation and leaky scanning in plant viruses. Virus Res. *119*, 52–62.

Saito, K., Fujiwara, T., Katahira, J., Inoue, K., and Sakamoto, H. (2004). TAP/NXF1, the primary mRNA export receptor, specifically interacts with a neuronal RNA-binding protein HuD. Biochem. Biophys. Res. Commun. *321*, 291–297.

Sakiyama, Y., Panatala, R., and Lim, R.Y.H. (2017). Structural dynamics of the nuclear pore complex. Semin. Cell Dev. Biol. *68*, 27–33.

Sakuma, T., Davila, J.I., Malcolm, J.A., Kocher, J.-P.A., Tonne, J.M., and Ikeda, Y. (2014). Murine leukemia virus uses NXF1 for nuclear export of spliced and unspliced viral transcripts. J. Virol. *88*, 4069–4082.

Samson, F., Brunaud, V., Balzergue, S., Dubreucq, B., Lepiniec, L., Pelletier, G., Caboche, M., and Lecharny, A. (2002). FLAGdb/FST: a database of mapped flanking insertion sites (FSTs) of Arabidopsis thaliana T-DNA transformants. Nucleic Acids Res. *30*, 94–97.

Sánchez-Navarro, J., Fajardo, T., Zicca, S., Pallás, V., and Stavolone, L. (2010). Caulimoviridae tubule-guided transport is dictated by movement protein properties. J. Virol. *84*, 4109–4112.

Sandri-Goldin, R.M. (1998). ICP27 mediates HSV RNA export by shuttling through a leucine-rich nuclear export signal and binding viral intronless RNAs through an RGG motif. Genes Dev. *12*, 868–879.

Sandri-Goldin, R.M. (2011). The many roles of the highly interactive HSV protein ICP27, a key regulator of infection. Future Microbiol. *6*, 1261–1277.

Sanfaçon, H., Brodmann, P., and Hohn, T. (1991). A dissection of the cauliflower mosaic virus polyadenylation signal. Genes Dev. *5*, 141–149.

Sarkar, S., and Hopper, A.K. (1998). tRNA Nuclear Export in Saccharomyces cerevisiae: In Situ Hybridization Analysis. Mol. Biol. Cell *9*, 3041–3055.

Scheifele, L.Z., Ryan, E.P., and Parent, L.J. (2005). Detailed mapping of the nuclear export signal in the Rous sarcoma virus Gag protein. J. Virol. *79*, 8732–8741.

Scheifele, L.Z., Kenney, S.P., Cairns, T.M., Craven, R.C., and Parent, L.J. (2007). Overlapping roles of the Rous sarcoma virus Gag p10 domain in nuclear export and virion core morphology. J. Virol. *81*, 10718–10728.

Schepetilnikov, M., and Ryabova, L. (2014). Chapter 17 - Cauliflower mosaic virus (CaMV) upregulates translation reinitiation of its pregenomic polycistronic 35S RNA via interaction with the cell's translation machinery. In Plant Virus–Host Interaction, R.K. Gaur, T. Hohn, and P. Sharma, eds. (Boston: Academic Press), pp. 325–343.

Schepetilnikov, M., Schott, G., Katsarou, K., Thiébeauld, O., Keller, M., and Ryabova, L.A. (2009). Molecular dissection of the prototype foamy virus (PFV) RNA 5'-UTR identifies essential elements of a ribosomal shunt. Nucleic Acids Res. *37*, 5838–5847.

Schepetilnikov, M., Kobayashi, K., Geldreich, A., Caranta, C., Robaglia, C., Keller, M., and Ryabova, L.A. (2011). Viral factor TAV recruits TOR/S6K1 signalling to activate reinitiation after long ORF translation. EMBO J. *30*, 1343–1356.

Schepetilnikov, M., Makarian, J., Srour, O., Geldreich, A., Yang, Z., Chicher, J., Hammann, P., and Ryabova, L.A. (2017). GTPase ROP2 binds and promotes activation of target of rapamycin, TOR, in response to auxin. EMBO J. *36*, 886–903.

Schmid, M., Speiseder, T., Dobner, T., and Gonzalez, R.A. (2014). DNA virus replication compartments. J. Virol. *88*, 1404–1420.

Schmidt, H.B., and Görlich, D. (2016). Transport Selectivity of Nuclear Pores, Phase Separation, and Membraneless Organelles. Trends Biochem. Sci. *41*, 46–61.

Schneider, J., and Wolff, T. (2009). Nuclear functions of the influenza A and B viruses NS1 proteins: Do they play a role in viral mRNA export? Vaccine *27*, 6312–6316.

Schoelz, J.E., and Leisner, S. (2017). Setting Up Shop: The Formation and Function of the Viral Factories of Cauliflower mosaic virus. Front. Plant Sci. *8*, 1832.

Schoelz, J., Shepherd, R.J., and Daubert, S. (1986). Region VI of cauliflower mosaic virus encodes a host range determinant. Mol. Cell. Biol. *6*, 2632–2637.

Scholthof, K.-B.G., Adkins, S., Czosnek, H., Palukaitis, P., Jacquot, E., Hohn, T., Hohn, B., Saunders, K., Candresse, T., Ahlquist, P., et al. (2011). Top 10 plant viruses in molecular plant pathology. Mol. Plant Pathol. *12*, 938–954.

Schubert, T., and Köhler, A. (2016). Mediator and TREX-2: Emerging links between transcription initiation and mRNA export. Nucleus *7*, 126–131.

Schumann, S., Jackson, B.R., Baquero-Perez, B., and Whitehouse, A. (2013). Kaposi's Sarcoma-Associated Herpesvirus ORF57 Protein: Exploiting All Stages of Viral mRNA Processing. Viruses *5*, 1901–1923.

Schumann, S., Baquero-Perez, B., and Whitehouse, A. (2016). Interactions between KSHV ORF57 and the novel human TREX proteins, CHTOP and CIP29. J. Gen. Virol. *97*, 1904–1910.

Sciabica, K.S., Dai, Q.J., and Sandri-Goldin, R.M. (2003). ICP27 interacts with SRPK1 to mediate HSV splicing inhibition by altering SR protein phosphorylation. EMBO J. *22*, 1608–1619.

Segref, A., Sharma, K., Doye, V., Hellwig, A., Huber, J., Lührmann, R., and Hurt, E. (1997). Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)+ RNA and nuclear pores. EMBO J. *16*, 3256–3271.

Selman, M., Dankar, S.K., Forbes, N.E., Jia, J.-J., and Brown, E.G. (2012). Adaptive mutation in influenza A virus non-structural gene is linked to host switching and induces a novel protein by alternative splicing. Emerg. Microbes Infect. *1*, 1–10.

Sen, N., Cao, F., and Tavis, J.E. (2004). Translation of Duck Hepatitis B Virus Reverse Transcriptase by Ribosomal Shunting. J. Virol. *78*, 11751–11757.

Sen, R., Barman, P., Kaja, A., Ferdoush, J., Lahudkar, S., Roy, A., and Bhaumik, S.R. (2019). Distinct Functions of the Cap-Binding Complex in Stimulation of Nuclear mRNA Export. Mol. Cell. Biol. *39*.

Sengupta, J., Bussiere, C., Pallesen, J., West, M., Johnson, A.W., and Frank, J. (2010). Characterization of the nuclear export adaptor protein Nmd3 in association with the 60S ribosomal subunit. J. Cell Biol. *189*, 1079–1086.

Sergeant, A., Gruffat, H., and Manet, E. (2008). The Epstein-Barr virus (EBV) protein EB is an mRNA export factor essential for virus production. Front. Biosci. J. Virtual Libr. *13*, 3798–3813.

Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D., Dietrich, B., Ho, P., Bacwaden, J., Ko, C., et al. (2002). A High-Throughput Arabidopsis Reverse Genetics System. Plant Cell *14*, 2985–2994.

Shibata, S., Sasaki, M., Miki, T., Shimamoto, A., Furuichi, Y., Katahira, J., and Yoneda, Y. (2006). Exportin-5 orthologues are functionally divergent among species. Nucleic Acids Res. *34*, 4711–4721.

Shih, S.R., and Krug, R.M. (1996). Novel exploitation of a nuclear function by influenza virus: the cellular SF2/ASF splicing factor controls the amount of the essential viral M2 ion channel protein in infected cells. EMBO J. *15*, 5415–5427.

Shih, S.R., Suen, P.C., Chen, Y.S., and Chang, S.C. (1998). A novel spliced transcript of influenza A/WSN/33 virus. Virus Genes *17*, 179–183.

Shivaprasad, P.V., Rajeswaran, R., Blevins, T., Schoelz, J., Meins, F., Hohn, T., and Pooggin, M.M. (2008). The CaMV transactivator/viroplasmin interferes with RDR6-dependent trans-acting and secondary siRNA pathways in Arabidopsis. Nucleic Acids Res. *36*, 5896–5909.

Singh, P., Saha, U., Paira, S., and Das, B. (2018). Nuclear mRNA Surveillance Mechanisms: Function and Links to Human Disease. J. Mol. Biol. *430*, 1993–2013.

Smagulova, F., Maurel, S., Morichaud, Z., Devaux, C., Mougel, M., and Houzet, L. (2005). The highly structured encapsidation signal of MuLV RNA is involved in the nuclear export of its unspliced RNA. J. Mol. Biol. *354*, 1118–1128.

Smith, A., Brownawell, A., and Macara, I.G. (1998a). Nuclear import of Ran is mediated by the transport factor NTF2. Curr. Biol. CB *8*, 1403–1406.

Smith, G.J., Donello, J.E., Lück, R., Steger, G., and Hope, T.J. (1998b). The hepatitis B virus post-transcriptional regulatory element contains two conserved RNA stem-loops which are required for function. Nucleic Acids Res. *26*, 4818–4827.

Smyk, A., Szuminska, M., Uniewicz, K.A., Graves, L.M., and Kozlowski, P. (2006). Human enhancer of rudimentary is a molecular partner of PDIP46/SKAR, a protein interacting with DNA polymerase delta and S6K1 and regulating cell growth. FEBS J. *273*, 4728–4741.

Sommer, G., and Heise, T. (2008). Posttranscriptional control of HBV gene expression. Front. Biosci. J. Virtual Libr. *13*, 5533–5547.

Sommer, P., and Nehrbass, U. (2005). Quality control of messenger ribonucleoprotein particles in the nucleus and at the pore. Curr. Opin. Cell Biol. *17*, 294–301.

Sørensen, B.B., Ehrnsberger, H.F., Esposito, S., Pfab, A., Bruckmann, A., Hauptmann, J., Meister, G., Merkl, R., Schubert, T., Längst, G., et al. (2017). The Arabidopsis THO/TREX component TEX1 functionally interacts with MOS11 and modulates mRNA export and alternative splicing events. Plant Mol. Biol. *93*, 283–298.

Stage-Zimmermann, T., Schmidt, U., and Silver, P.A. (2000). Factors affecting nuclear export of the 60S ribosomal subunit in vivo. Mol. Biol. Cell *11*, 3777–3789.

Staiger, D., and Brown, J.W.S. (2013). Alternative splicing at the intersection of biological timing, development, and stress responses. Plant Cell *25*, 3640–3656.

Stavolone, L., Villani, M.E., Leclerc, D., and Hohn, T. (2005). A coiled-coil interaction mediates cauliflower mosaic virus cell-to-cell movement. Proc. Natl. Acad. Sci. U. S. A. *102*, 6219–6224.

Stewart, M. (2007). Ratcheting mRNA out of the nucleus. Mol. Cell *25*, 327–330.

Stewart, M. (2019). Polyadenylation and nuclear export of mRNAs. J. Biol. Chem. *294*, 2977–2987.

Strässer, K., Bassler, J., and Hurt, E. (2000). Binding of the Mex67p/Mtr2p heterodimer to FXFG, GLFG, and FG repeat nucleoporins is essential for nuclear mRNA export. J. Cell Biol. *150*, 695–706.

Strässer, K., Masuda, S., Mason, P., Pfannstiel, J., Oppizzi, M., Rodriguez-Navarro, S., Rondón, A.G., Aguilera, A., Struhl, K., Reed, R., et al. (2002). TREX is a conserved complex coupling transcription with messenger RNA export. Nature *417*, 304–308.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. *185*, 60–89.

Stuger, R., and Forreiter, C. (2004). Uncapped mRNA introduced into tobacco protoplasts can be imported into the nucleus and is trapped by leptomycin B. Plant Cell Rep. *23*, 99–103.

Stutz, F., Bachi, A., Doerks, T., Braun, I.C., Séraphin, B., Wilm, M., Bork, P., and Izaurralde, E. (2000). REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. RNA N. Y. N *6*, 638–650.

Su, T.S., Lai, C.J., Huang, J.L., Lin, L.H., Yauk, Y.K., Chang, C.M., Lo, S.J., and Han, S.H. (1989). Hepatitis B virus transcript produced by RNA splicing. J. Virol. *63*, 4011–4018.

Sugiura, T., Sakurai, K., and Nagano, Y. (2007). Intracellular characterization of DDX39, a novel growth-associated RNA helicase. Exp. Cell Res. *313*, 782–790.

Suhasini, M., and Reddy, T.R. (2009). Cellular proteins and HIV-1 Rev function. Curr. HIV Res. *7*, 91–100.

Sullivan, K.D., Mullen, T.E., Marzluff, W.F., and Wagner, E.J. (2009). Knockdown of SLBP results in nuclear retention of histone mRNA. RNA N. Y. N *15*, 459–472.

Sun, Q., Carrasco, Y.P., Hu, Y., Guo, X., Mirzaei, H., Macmillan, J., and Chook, Y.M. (2013). Nuclear export inhibition through covalent conjugation and hydrolysis of Leptomycin B by CRM1. Proc. Natl. Acad. Sci. U. S. A. *110*, 1303–1308.

Takano, A., Endo, T., and Yoshihisa, T. (2005). tRNA actively shuttles between the nucleus and cytosol in yeast. Science *309*, 140–142.

Tamura, K., and Hara-Nishimura, I. (2013). The molecular architecture of the plant nuclear pore complex. J. Exp. Bot. *64*, 823–832.

Taniguchi, I., Mabuchi, N., and Ohno, M. (2014). HIV-1 Rev protein specifies the viral RNA export pathway by suppressing TAP/NXF1 recruitment. Nucleic Acids Res. *42*, 6645–6658.

Tazi, J., Bakkour, N., Marchand, V., Ayadi, L., Aboufirassi, A., and Branlant, C. (2010). Alternative splicing: regulation of HIV-1 multiplication as a target for therapeutic action. FEBS J. *277*, 867–876.

Terry, L.J., and Wente, S.R. (2009). Flexible gates: dynamic topologies and functions for FG nucleoporins in nucleocytoplasmic transport. Eukaryot. Cell *8*, 1814–1827.

Thiébeauld, O., Schepetilnikov, M., Park, H.-S., Geldreich, A., Kobayashi, K., Keller, M., Hohn, T., and Ryabova, L.A. (2009). A new plant protein interacts with eIF3 and 60S to enhance virus-activated translation re-initiation. EMBO J. *28*, 3171–3184.

Thomas, C.M., Hull, R., Bryant, J.A., and Maule, A.J. (1985). Isolation of a fraction from cauliflower mosaic virus-infected protoplasts which is active in the synthesis of (+) and (-) strand viral DNA and reverse transcription of primed RNA templates. Nucleic Acids Res. *13*, 4557–4576.

Tian, X., Devi-Rao, G., Golovanov, A.P., and Sandri-Goldin, R.M. (2013). The interaction of the cellular export adaptor protein Aly/REF with ICP27 contributes to the efficiency of herpes simplex virus 1 mRNA export. J. Virol. *87*, 7210–7217.

Topisirovic, I., Siddiqui, N., Lapointe, V.L., Trost, M., Thibault, P., Bangeranye, C., Piñol-Roma, S., and Borden, K.L.B. (2009). Molecular dissection of the eukaryotic initiation factor 4E (eIF4E) export-competent RNP. EMBO J. *28*, 1087–1098.

Torruella, M., Gordon, K., and Hohn, T. (1989). Cauliflower mosaic virus produces an aspartic proteinase to cleave its polyproteins. EMBO J. *8*, 2819–2825.

Tsai, P.-L., Chiou, N.-T., Kuss, S., García-Sastre, A., Lynch, K.W., and Fontoura, B.M.A. (2013). Cellular RNA binding proteins NS1-BP and hnRNP K regulate influenza A virus RNA splicing. PLoS Pathog. *9*, e1003460.

Tunnicliffe, R.B., Hautbergue, G.M., Kalra, P., Jackson, B.R., Whitehouse, A., Wilson, S.A., and Golovanov, A.P. (2011). Structural basis for the recognition of cellular mRNA export factor REF by herpes viral proteins HSV-1 ICP27 and HVS ORF57. PLoS Pathog. *7*, e1001244.

Tunnicliffe, R.B., Schacht, M., Levy, C., Jowitt, T.A., Sandri-Goldin, R.M., and Golovanov, A.P. (2015). The structure of the folded domain from the signature multifunctional protein ICP27 from herpes simplex virus-1 reveals an intertwined dimer. Sci. Rep. *5*, 11234.

Tunnicliffe, R.B., Tian, X., Storer, J., Sandri-Goldin, R.M., and Golovanov, A.P. (2018). Overlapping motifs on the herpes viral proteins ICP27 and ORF57 mediate interactions with the mRNA export adaptors ALYREF and UIF. Sci. Rep. *8*, 15005.

Turner, J.G., Dawson, J., and Sullivan, D.M. (2012). Nuclear export of proteins and drug resistance in cancer. Biochem. Pharmacol. *83*, 1021–1032.

Tutucci, E., and Stutz, F. (2011). Keeping mRNPs in check during assembly and nuclear export. Nat. Rev. Mol. Cell Biol. *12*, 377–384.

Ulitsky, I., and Bartel, D.P. (2013). lincRNAs: genomics, evolution, and mechanisms. Cell *154*, 26–46.

Valencia-Sanchez, M.A., Liu, J., Hannon, G.J., and Parker, R. (2006). Control of translation and mRNA degradation by miRNAs and siRNAs. Genes Dev. *20*, 515–524.

Valkov, E., Dean, J.C., Jani, D., Kuhlmann, S.I., and Stewart, M. (2012). Structural basis for the assembly and disassembly of mRNA nuclear export complexes. Biochim. Biophys. Acta *1819*, 578–592.

Vih.org (2018). Le VIH en 2019: Les clefs pour comprendre.

Vinciguerra, P., Iglesias, N., Camblong, J., Zenklusen, D., and Stutz, F. (2005). Perinuclear Mlp proteins downregulate gene expression in response to a defect in mRNA export. EMBO J. *24*, 813–823.

Viphakone, N., Hautbergue, G.M., Walsh, M., Chang, C.-T., Holland, A., Folco, E.G., Reed, R., and Wilson, S.A. (2012). TREX exposes the RNA-binding domain of Nxf1 to enable mRNA export. Nat. Commun. *3*, 1006.

Visootsat, A., Payungporn, S., and T-Thienprasert, N.P. (2015). A conserved RNA structural element within the hepatitis B virus post-transcriptional regulatory element enhance nuclear export of intronless transcripts and repress the splicing mechanism. Mol. Biol. Rep. *42*, 1603–1614.

Vogt, C., and Bohne, J. (2016). The KSHV RNA regulator ORF57: target specificity and its role in the viral life cycle. Wiley Interdiscip. Rev. RNA *7*, 173–185.

Volkova, N.A., Fomina, E.G., Smolnikova, V.V., Zinovieva, N.A., and Fomin, I.K. (2014). The U3 region of Moloney murine leukemia virus contains position-independent cis-acting sequences involved in the nuclear export of full-length viral transcripts. J. Biol. Chem. *289*, 20158–20169.

Vollmer, B., and Antonin, W. (2014). The diverse roles of the Nup93/Nic96 complex proteins - structural scaffolds of the nuclear pore complex with additional cellular functions. Biol. Chem. *395*, 515–528.

Wang, W., Cui, Z.-Q., Han, H., Zhang, Z.-P., Wei, H.-P., Zhou, Y.-F., Chen, Z., and Zhang, X.-E. (2008). Imaging and characterizing influenza A virus mRNA transport in living cells. Nucleic Acids Res. *36*, 4913.

Watanabe, K., Takizawa, N., Katoh, M., Hoshida, K., Kobayashi, N., and Nagata, K. (2001). Inhibition of nuclear export of ribonucleoprotein complexes of influenza virus by leptomycin B. Virus Res. *77*, 31–42.

Watts, J.M., Dang, K.K., Gorelick, R.J., Leonard, C.W., Bess, J.W., Swanstrom, R., Burch, C.L., and Weeks, K.M. (2009). Architecture and secondary structure of an entire HIV-1 RNA genome. Nature *460*, 711–716.

Watts, N.R., Eren, E., Zhuang, X., Wang, Y.-X., Steven, A.C., and Wingfield, P.T. (2018). A new HIV-1 Rev structure optimizes interaction with target RNA (RRE) for nuclear export. J. Struct. Biol. *203*, 102–108.

Webster, C.G., Pichon, E., van Munster, M., Monsion, B., Deshoux, M., Gargani, D., Calevro, F., Jimenez, J., Moreno, A., Krenz, B., et al. (2018). Identification of Plant Virus Receptor Candidates in the Stylets of Their Aphid Vectors. J. Virol. 92.

Wen, W., Meinkoth, J.L., Tsien, R.Y., and Taylor, S.S. (1995). Identification of a signal for rapid export of proteins from the nucleus. Cell *82*, 463–473.

Wickramasinghe, V.O., and Laskey, R.A. (2015). Control of mammalian gene expression by selective mRNA export. Nat. Rev. Mol. Cell Biol. *16*, 431–442.

Wickramasinghe, V.O., Stewart, M., and Laskey, R.A. (2010). GANP enhances the efficiency of mRNA nuclear export in mammalian cells. Nucl. Austin Tex *1*, 393–396.

Woo, J.L., and Berk, A.J. (2007). Adenovirus ubiquitin-protein ligase stimulates viral late mRNA nuclear export. J. Virol. *81*, 575–587.

Woodward, L.A., Mabin, J.W., Gangras, P., and Singh, G. (2017). The exon junction complex: a lifelong guardian of mRNA fate. Wiley Interdiscip. Rev. RNA *8*, e1411.

Woody, S.T., Austin-Phillips, S., Amasino, R.M., and Krysan, P.J. (2007). The WiscDsLox T-DNA collection: an arabidopsis community resource generated by using an improved high-throughput T-DNA sequencing pipeline. J. Plant Res. *120*, 157–165.

Wu, J., Bao, A., Chatterjee, K., Wan, Y., and Hopper, A.K. (2015). Genome-wide screen uncovers novel pathways for tRNA processing and nuclear-cytoplasmic dynamics. Genes Dev. *29*, 2633–2644.

Wurch, T., Kirchherr, D., Mesnard, J.M., and Lebeurier, G. (1990). The cauliflower mosaic virus open reading frame VII product can be expressed in Saccharomyces cerevisiae but is not detected in infected plants. J. Virol. *64*, 2594–2598.

Xu, D., Grishin, N.V., and Chook, Y.M. (2012). NESdb: a database of NES-containing CRM1 cargoes. Mol. Biol. Cell *23*, 3673–3676.

Xu, D., Marquis, K., Pei, J., Fu, S.-C., Cağatay, T., Grishin, N.V., and Chook, Y.M. (2015). LocNES: a computational tool for locating classical NESs in CRM1 cargo proteins. Bioinforma. Oxf. Engl. *31*, 1357–1365.

Yang, C.-C., Huang, E.-Y., Li, H.-C., Su, P.-Y., and Shih, C. (2014). Nuclear Export of Human Hepatitis B Virus Core Protein and Pregenomic RNA Depends on the Cellular NXF1-p15 Machinery. PLoS ONE *9*.

Yang, J., Bogerd, H.P., Peng, S., Wiegand, H., Truant, R., and Cullen, B.R. (1999). An ancient family of human endogenous retroviruses encodes a functional homolog of the HIV-1 Rev protein. Proc. Natl. Acad. Sci. U. S. A. *96*, 13404–13408.

Yang, J., Bogerd, H.P., Wang, P.J., Page, D.C., and Cullen, B.R. (2001). Two closely related human nuclear export factors utilize entirely distinct export pathways. Mol. Cell *8*, 397–406.

Yao, W., Roser, D., Köhler, A., Bradatsch, B., Bassler, J., and Hurt, E. (2007). Nuclear export of ribosomal 60S subunits by the general mRNA export receptor Mex67-Mtr2. Mol. Cell *26*, 51–62.

Yao, W., Lutzmann, M., and Hurt, E. (2008). A versatile interaction platform on the Mex67-Mtr2 receptor creates an overlap between mRNA and ribosome export. EMBO J. *27*, 6–16.

Yarbrough, M.L., Mata, M.A., Sakthivel, R., and Fontoura, B.M.A. (2014). Viral Subversion of Nucleocytoplasmic Trafficking. Traffic Cph. Den. *15*, 127–140.

Yasuda, J., Nakada, S., Kato, A., Toyoda, T., and Ishihama, A. (1993). Molecular assembly of influenza virus: association of the NS2 protein with virion matrix. Virology *196*, 249–255.

Yatherajam, G., Huang, W., and Flint, S.J. (2011). Export of adenoviral late mRNA from the nucleus requires the Nxf1/Tap export receptor. J. Virol. *85*, 1429–1438.

Yelina, N.E., Smith, L.M., Jones, A.M.E., Patel, K., Kelly, K.A., and Baulcombe, D.C. (2010a). Putative Arabidopsis THO/TREX mRNA export complex is involved in transgene and endogenous siRNA biosynthesis. Proc. Natl. Acad. Sci. U. S. A. *107*, 13948–13953.

Yong, J., Wan, L., and Dreyfuss, G. (2004). Why do cells need an assembly machine for RNA-protein complexes? Trends Cell Biol. *14*, 226–232.

Yoo, S.-D., Cho, Y.-H., and Sheen, J. (2007). *Arabidopsis* mesophyll protoplasts: a versatile cell system for transient gene expression analysis. Nat. Protoc. *2*, 1565–1572.

York, A., and Fodor, E. (2013). Biogenesis, assembly, and export of viral messenger ribonucleoproteins in the influenza A virus infected cell. RNA Biol. *10*, 1274–1282.

Yoshihisa, T., Ohshima, C., Yunoki-Esaki, K., and Endo, T. (2007). Cytoplasmic splicing of tRNA in Saccharomyces cerevisiae. Genes Cells Devoted Mol. Cell. Mech. *12*, 285–297.

Yu, M., Liu, X., Cao, S., Zhao, Z., Zhang, K., Xie, Q., Chen, C., Gao, S., Bi, Y., Sun, L., et al. (2012). Identification and characterization of three novel nuclear export signals in the influenza A virus nucleoprotein. J. Virol. *86*, 4970–4980.

Yu, W., Murfett, J., and Schoelz, J.E. (2003). Differential induction of symptoms in Arabidopsis by P6 of Cauliflower mosaic virus. Mol. Plant-Microbe Interact. MPMI *16*, 35–42.

Yue, Y., Coskun, A.K., Jawanda, N., Auer, J., and Sutton, R.E. (2018). Differential interaction between human and murine Crm1 and lentiviral Rev proteins. Virology *513*, 1–10.

Zang, W.Q., and Yen, T.S. (1999). Distinct export pathway utilized by the hepatitis B virus posttranscriptional regulatory element. Virology *259*, 299–304.

Zang, W.Q., Li, B., Huang, P.Y., Lai, M.M., and Yen, T.S. (2001). Role of polypyrimidine tract binding protein in the function of the hepatitis B virus posttranscriptional regulatory element. J. Virol. *75*, 10779–10786.

Zenklusen, D., Vinciguerra, P., Wyss, J.-C., and Stutz, F. (2002). Stable mRNP formation and export require cotranscriptional recruitment of the mRNA export factors Yra1p and Sub2p by Hpr1p. Mol. Cell. Biol. *22*, 8241–8253.

Zhang, K., Shang, G., Padavannil, A., Wang, J., Sakthivel, R., Chen, X., Kim, M., Thompson, M.G., García-Sastre, A., Lynch, K.W., et al. (2018). Structural–functional interactions of NS1-BP protein with the splicing and mRNA export machineries for viral and host gene expression. Proc. Natl. Acad. Sci. *115*, E12218–E12227.

Zhang, K., Xie, Y., Muñoz-Moreno, R., Wang, J., Zhang, L., Esparza, M., García-Sastre, A., Fontoura, B.M.A., and Ren, Y. (2019). Structural basis for influenza virus NS1 protein block of mRNA nuclear export. Nat. Microbiol.

Zhu, J., and McKeon, F. (1999). NF-AT activation requires suppression of Crm1-dependent export by calcineurin. Nature *398*, 256.

Julie KUBINA

Étude de l'export nucléaire des ARNm du CaMV chez *Arabidopsis thaliana*

RÉSUMÉ. L'export nucléaire des ARNm cellulaires et viraux dans les cellules animales est largement étudié depuis de nombreuses années. La majorité des ARNm matures emprunte la voie de TAP-p15 et des complexes TREX, mais une minorité d'entre eux est exportée par CRM1 (XPO1) grâce à des protéines adaptatrices possédant un NES. En revanche, chez les plantes, ce processus fondamental reste encore peu décrit voir même inconnu pour les phytovirus à génome ADN, comme le CaMV. En tant que pararétrovirus, le CaMV doit exporter au cours de son cycle ses ARNm 19S et 35S. Notre objectif a été de clarifier cette étape primordiale mais jamais étudiée du cycle du CaMV et de l'utiliser comme modèle pour décrypter les mécanismes d'export nucléaire des ARNm chez la plante modèle et hôte de CaMV *Arabidopsis thaliana*. Nos résultats des tests d'infection par le CaMV de plants mutants d'*A. thaliana* n'exprimant pas certaines protéines du complexe TREX-1, et de quantification de l'export des ARNm viraux en protoplastes d'Arabidopsis transfectés montrent que trois facteurs de ce complexe, MOS11, TEX1 et ALY sont importants pour le CaMV qui emprunte cette voie pour exporter ses ARN 35S. Nous avons également montré que l'export des transcrits du CaMV est inhibé dans des protoplastes mutants n'exprimant pas l'exportine XPO1a ainsi qu'en présence de la leptomycine B, un inhibiteur très efficace de CRM1, suggérant que les deux voies d'export, la voie majeure associée au complexe TREX-1 et la voie mineure de CRM1/XPO1 sont utilisées par le virus pour exporter ses ARN. En étudiant des mutants de CaMV n'exprimant plus soit aucune protéine virale soit plus sa transcriptase inverse P5, nous avons observé une forte inhibition de l'export des ARN 35S, démontrant l'importance d'au moins P5 dans ce processus. En suivant la localisation subcellulaire dans des protoplastes de la protéine P5 fusionnée à l'eGFP, et de P4 le précurseur de la protéine de capside du virus, nous avons montré que ces deux protéines présentent une distribution nucléo-cytoplasmique et co-localisent dans le noyau avec la protéine de fusion mRFP-MOS11. Ces deux protéines sont également capables d'interagir *in vitro* avec le facteur d'export MOS11, comme le montrent nos expériences de GST *pull-down*. Finalement, nous avons également montré, avec des expériences de compétition en système de protoplastes, que la région *leader* de l'ARN 35S du CaMV, et en particulier sa tige-boucle, est un des élément agissant en *cis*, capable de recruter la machinerie d'export nucléaire des ARNm .

Mots-clés : Export nucléaire, ARNm, TAP-p15, TREX, CRM1, virus de la mosaïque du chou-fleur, *Arabidopsis thaliana*.

ABSTRACT. Nuclear export of cellular and viral mRNAs in animal cells is largely studied since many years. Most of the mature mRNAs use TAP-p15 and TREX complexes while a minority of transcripts reaches the cytoplasm via CRM1 (XPO1) exportine and NES-containing adaptor proteins. In plants, however, this fundamental process is poorly known and even unknown for DNA phytoviruses, as CaMV. As a pararetrovirus, CaMV has to export in the cytoplasm his 19S and 35S mRNAs. Our aim was to clarify this important but never studied step of CaMVs life cycle and use the virus as protoptype to decode the mechanisms of mRNAs nuclear export in the model plant and CaMV host *Arabidopsis thaliana*. By challenging *A. thaliana* TREX-1 mutant plants with CaMV and by monitoring the nuclear export of CaMV mRNAs in Arabidopsis transfected protoplasts, we showed that MOS11, TEX1 and ALY export factors, in particular, and the TREX-1 complex, in general, are used by CaMV 35S mRNAs to reach the cytoplasm. We also demonstrated that the nuclear export of 35S mRNAs is strongly inhibited in *xpo1a* mutant protoplasts or after leptomycin B inhibition of exportin CRM1, suggesting that both the main TREX complex pathway and the one of exportine CRM1 are used by the virus to export its 35S mRNAs. By studying CaMV mutagenised genomes that do not express either any of CaMV proteins or its reverse transcriptase P5, we observed a strong decrease in 35S mRNAs export, showing the importance of P5 in this process. Moreover, following the subcellular localization in protoplasts of eGFP tagged P5 and P4 proteins (P4 being the precursor of CaMV coat protein) we also showed that these two viral proteins display a nucleo-cytoplasmic distribution and co-localize in nuclei with fusion protein mRFP-MOS11. These two proteins also interact with export factor MOS11 *in vitro*, as shown by GST pull-down assays. Finally, by performing competition experiments in protoplasts, we showed that CaMV 35S RNA leader region and especially its highly structured stem-loop is a *cis*-acting element that recruits the mRNAs nuclear export machinery.

Keywords: Nuclear export, mRNA, Tap-p15, TREX, CRM1, *Cauliflower Mosaic Virus*, *Arabidopsis thaliana*.