Rôle et mode d'action de l'UTP : RNA Uridylyltransférase URT1 dans l'uridylation et la dégradation des ARNm chez Aradopsis thaliana

Emilie Ferrier

▶ **To cite this version:**

Emilie Ferrier. Rôle et mode d'action de l'UTP : RNA Uridylyltransférase URT1 dans l'uridylation et la dégradation des ARNm chez Aradopsis thaliana. Génomique, Transcriptomique et Protéomique [q-bio.GN]. Université de Strasbourg, 2013. Français. NNT : 2013STRAJ053 . tel-01124203

HAL Id: tel-01124203
https://theses.hal.science/tel-01124203

Submitted on 6 Mar 2015

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
Ecole Doctorale des Sciences de la Vie et de la Santé
Institut de Biologie Moléculaire des Plantes – CNRS – UPR2357

THESE
De Emilie FERRIER

ROLE ET MODE D'ACTION DE L'UTP:RNA URIDYLYLTRANSFERASE URT1 DANS L'URIDYLATION ET LA DEGRADATION DES ARNM CHEZ ARABIDOPSIS THALIANA

Pour obtenir le grade de
DOCTEUR
Spécialité : Aspects moléculaires et cellulaires de la biologie

Thèse soutenue le 29 Novembre 2013

Directeur de thèse	Dr. Dominique GAGLIARDI	Université de Strasbourg
Rapporteur externe	Dr. Dominique WEIL	Université Pierre et Marie Curie
Rapporteur externe	Dr. Julio SAEZ VASQUEZ	Université de Perpignan
Rapporteur interne	Dr. Sébastien PFEFFER	Université de Strasbourg

ROLE ET MODE D'ACTION DE L'UTP:RNA URIDYLYLTRANSFERASE URT1 DANS L'URIDYLATION ET LA DEGRADATION DES ARNM CHEZ *ARABIDOPSIS THALIANA*

Résumé :

La dégradation des ARN est un mécanisme essentiel à la régulation de l'expression des génomes. L'importance de l'uridylation dans les mécanismes de dégradation des ARN commence juste à être appréciée. Cette thèse présente l'étude de l'*UTP :RNA Uridylyltransferase 1* (URT1) et de son rôle dans la dégradation des ARN chez *Arabidopsis thaliana*. L'étude des propriétés catalytiques de URT1 montre que cette uridylyltransférase est intrinsèquement spécifique des UTP et distributive pour les premières uridines ajoutées. URT1 est responsable *in vivo* de l'uridylation des ARNm après une étape de déadénylation, protégeant leur extrémité 3' et polarisant la dégradation de 5' en 3'. URT1 est localisée dans le cytosol au niveau des granules de stress et des *processing bodies.* Le mécanisme d'adressage de URT1 dans les *processing bodies* implique une partie de la région N terminale prédite comme intrinsèquement désorganisée, alors que le domaine nucléotidyltransférase C terminal semble suffisant pour permettre l'adressage de URT1 au niveau des *processing bodies* et granules de stress en réponse à un stress thermique.

Ces travaux de thèse ont permis de mieux comprendre les mécanismes et les rôles de l'uridylation dans la dégradation des ARNm chez Arabidopsis. Ils ouvrent des perspectives dans l'étude d'autres fonctions de l'uridylation comme l'inhibition de la traduction.

ROLE ET MODE OF ACTION THE UTP:RNA URIDYLYLTRANSFERASE URT1 IN MRNA'S URIDYLATION AND DEGRADATION IN *ARABIDOPSIS THALIANA*

Abstract :

RNA degradation is an essential mechanism for the regulation of genome expression. The importance of uridylation for RNA degradation is just emerging. This thesis presents the study of URT1 (UTP :RNA Uridylyltransferase 1) and its role in RNA degradation in *Arabidopsis thaliana*. URT1 is an uridylyltransferase intrinsically and strictly specific for UTP and is distributive for the first nucleotides added. URT1 uridylates mRNA *in vivo* after a deadenylation step. This uridylation protects mRNA's 3' end from further attacks and polarise degradation in the 5' to 3' direction. This protection of 3' ends by uridylation and its conferred polarity of 5' to 3' degradation are also detected in polysomes. Uridylation is therefore likely important in case of co-translational degradation of mRNAs. A region in URT1's N terminal region predicted to be intrinsically disorganised is required for addressing URT1 to processing bodies. However, following heat shock, the nucleotidyltransferase domain present in the C terminal region of URT1 is sufficient to address URT1 to both *processing bodies* and stress granules,

This work contributes to a better understanding of the mechanisms and roles of uridylation in RNA degradation in *Arabidopsis thaliana*. These results also open perspectives for studying other functions of uridylation such as translation inhibition.

Remerciement

Nombreux sont ceux que je souhaiterai remercier mais je ne ferai pas une liste de noms qui serait longue et certainement ennuyeuse pour les lecteurs. J'irai donc à l'essentiel.

Je remercie tout d'abord mon directeur de thèse, Dominique Gagliardi pour m'avoir permis de réaliser ce travail, m'avoir apporté son expérience et son savoir intarissable sur la dégradation des ARN. J'ai beaucoup appris à ses cotés.

Je souhaite aussi remercier tout particulièrement les membres de mon jury, Sébastien Pfeffer, Julio Saez Vasquez et Dominique Weil pour avoir accepté d'évaluer mon travail de thèse.

D'autre part, je remercie les anciens et les actuels membres du laboratoire « 306 », tout particulièrement Hélène Zuber, François Sement et Heike Lange, pour les nombreuses discussions scientifiques ou non, leur aide et soutien, leurs conseils avisés et leur bonne humeur qui ont rendus ces trois ans beaucoup plus agréable. Je laisserai une mention spéciale pour Hélène, devenue une amie, qui m'a fait découvrir entre autres sa passion pour les statistiques et la bioinformatique.

Merci aux anciens et actuels membres du département « mito », appelé actuellement « Régulation et Coordination de l'expression des Génomes » et les membres de l'IBMP pour leur disponibilité, leur aide, leur soutien et leurs conseils.

Je n'oublie pas non plus tous les membres des plateformes, Marc Bergdoll de la plateforme de biologie structurale, Jérôme Muterrer et Mathieu Erhardt de la plateforme Microscopie et Imagerie, Nicolas Baumberger de la plateforme de production et purification des protéines, l'équipe de protéomique, Philippe Hammann, Lauriane Kuhn et Johana Chicher ; Malek Alioua de la plateforme de séquençage d'ADN et l'équipe de production de plantes de l'IBMP qui ont toujours été à l'écoute, de bon conseils et qui ont rendus ce travail possible.

Je tiens aussi à remercier les membres des équipes des Dr. Jean-Marc Deragon, du Laboratoire Génome et Développement des Plantes de Perpignan et Anton Meinhart du Max Planck Institute for Medical Reasearch de Heidelberg en Allemagne pour leur collaboration à ce projet.

Enfin, je remercie la personne qui a partagé ma vie durant ces trois ans, ma famille et mon entourage pour leur soutien au quotidien.

Abréviations

3’dATP 3’ désoxyAdénineTriPhosphate
3D trois dimension
A ATP
Å angström
ADN Acide DésoxyRiboNucléique ou *DNA*
ADN-T ADN du plasmide Ti
AMP Adénine MonoPhosphate
ARE AU-*Rich destabilizing Elements*
ARN Acide RiboNucléique ou *RNA*
ARNg ARN guides ou *gRNA*
ARNm ARN messagers
ARNnc ARN non codant
ARNr ARN ribosomiques
ARNt ARN de transfert
ATP Adénine TriPhosphate
AUB AU *Binding protein*
C CTP
CaMV *Cauliflower Mosaic Virus*
CDS *Coding DNA Sequence*
CFP Cyan Fluorescent Protein
Ci curi
Col0 Columbia-0
CTP Cytidine TriPhosphate
Da Dalton
DCL Dicer-*like*
DIC *Differential Interference Contrast*
DTT DiThioThreitol
e-YFP enhenced Yellow Fluorescent Protein
EDTA Ethylène Diamine Tétra-Acétate de sodium
eIF4 eucaryotic translation Initiation T-Factor
g intensité de la force artificielle, force centrifuge
G GTP
GFP Green Fluorescent Protein
GST Gluthation S-Transférase
GTP Guanidine TriPhosphate
His Histidine
IPTG IsoPropyl β-D-1 ThioGalactopyranoside
LB Luria Broth
Ler Landsberg

mA milliAmpère
MES 2-(N-Morpholino)EthaneSulfonic acid
miARN microARN
min minutes
MOPS 3-(N-MorphOlino)PropaneSulfonic
MS Murashige et Skoog
n nano
ncPAP *non-canonical* PolyA Polymérase
NGD *No-Go decay*
Ni—NTA Nickel-NitrilioTriacetic Acid
NMD *Non-sens-Mediated Decay*
Nos Nopaline synthase
nt nucléotides
OCS OCtopine Synthase
ORF *Open Reading Frame*
p/v poids/volume
PABP *PolyA Binding Protein*
PAP PolyA Polymérase
piARN piwi ARN
PTGS *Post-Transcriptional Gene Silencing*
PUP PolyU Polymérase
RFP Red Fluorescent Protein
RISC RNA-*Induced Silencing Complex*
RNase RiboNucléase
RNP RiboNucléoProtéique
rNTP riboNucléotideTriPhosphate
SDS Sodium Dodécyl Sulfate
SDS-PAGE Sodium Dodécyl Sulfate *PolyAcrylamide Gel Electrophoresis*
sec secondes
sARN *small* ARN
siARN *small interfering* ARN
TERRA *Telomeric Repeat-containing RNA*
Ti *Tumor-inducing*
U UTP
U5P Uridine 5'Phosphate
UMP Uridine MonoPhosphate
UTP Uridine TriPhosphate
V volt
v/v volume/volume
α alpha
β béta
Δ délétion
µ micro

Φ phi
Ψ psi

Sommaire

Introduction 1
La dégradation de l'ARN 3
A - Les facteurs impliqués dans la dégradation 4
1. Les exoribonucléases agissant en 3' de l'ARN 4
a) Les déadénylases 4
1. Le complexe PAN 5
2. Le complexe Ccr4-Pop2/Caf1-Not 5
3. PARN 6
4. La régulation de la déadénylation 6
5. Les déadénylases chez A. thaliana 6
b) Le complexe de l'exosome 7
1. Les facteurs associés à l'exosome dans le noyau 8
2. Facteurs associés à l'exosome dans le cytoplasme 9
3. L'exosome chez A. thaliana 9
c) Dis3L2/SOV 10
2. Les ribonucléases agissant en 5' de l'ARN 11
a) Le complexe d'élimination de la coiffe 11
1. Les facteurs de régulation du complexe d'élimination de la coiffe 11
2. Le complexe d'élimination de la coiffe chez A. thaliana 12
b) Les exonucléases 5'-3' 12
1. XRN1/PACMAN 13
2. La XRN nucléaire XRN2/RAT1 14
3. Les endonucléases 16
4. Les protéines modifiant et/ou marquant les ARN précédant leur dégradation 16
a) Le complexe Lsm1-7 17
b) Les ncPAP 17
1. Les ncPAP responsable de la polyadénylation 19
2. Les autres ncPAP 20
3. Les ncPAP chez Arabidopsis thaliana 20
B—Les rôles de l'uridylation dans le métabolisme des ARN 21
1. Le petit ARN nucléaire U6 snARN 22
2. Le précurseur du microARN let-7 (pré-let-7) 22
3. Les petits ARN (microARN, siARN et piARN) 23
a) La biogenèse des petits ARN (microARN, siARN et piARN) 23
1. Les piARN spécifiques des animaux 24
2. Les siARN et les microARN chez les plantes 24
b) La modification en 3' des petits ARN 24

L'uridylation des petits ARN 25
4. Les fragments d'ARNm clivés par le complexe RISC 27
a) Le mode d'action des petits ARN 27
b) La dégradation des sous-produits du clivage par RISC de l'ARNm 27
5. Les ARN guides impliqués dans l'édition des ARNm chez les trypanosomes 28
a) L'édition des ARNm 28
b) Les petits ARN guides uridylés 28
6. Les ARNm 29
a) Les ARNm polyadénylés 29
b) Les ARNm oligoadénylés 29
c) Les ARNm non adénylés 30
7. Conclusion 31
C—Les complexes ribonucléoprotéiques cytoplasmiques dans la dégradation et/ou le stockage des ARNm 31
1. Les polysomes 31
2. Les granules cytoplasmiques 33
a) Les points communs et les spécificités des granules de stress et des processing bodies 34
b) Les processing bodies 35
1. La composition des processing bodies 35
2. La formation des processing bodies 36
3. Le rôle putatif des processing bodies 38
c) Les granules de stress 40
1. La composition des granules de stress 40
2. La formation des granules de stress 41
3. Le rôle des granules de stress 43
3. La régulation de la traduction des ARNm et leur stockage dans les processing bodies et/ou les granules de stress lors d'une infection virale 45
4. La relation entre les polysomes, les granules de stress et les processing bodies 46
Objectifs de thèse 48

Chapitre 1 50
Résumé 51
1. URT1 est une nucléotidyltransférase composée de deux régions structuralement différentes 53
2. La région N terminale de URT1 est conservée chez les Angiospermes 53
a) La région N terminale de URT1 est prédite intrinsèquement désorganisée 53
b) Des zones de liaison à d'autres protéines sont prédites au niveau de la région N terminale de URT1 54
c) La région N terminale est conservée chez les Angiospermes 55

3. Le domaine C terminal de URT1 est similaire à Cid1 58
4. Conclusion/Ouverture 58

Chapitre 2 60
Résumé 61
A - URT1 est responsable de l'uridylation des ARNm oligoadénylés et prévient la formation de transcrits tronqués en 3' chez Arabidopsis thaliana 63
1. Article publié dans Nucleic Acids Research 63
2. Informations complémentaires concernant les ARNm uridylés 85
a) Analyse des extensions uridylées des ARNm 85
b) Analyse des extensions uridylées des ARNm associés aux polysomes 85
3. Conclusion et ouverture 87
B- URT1 n'est pas intrinsèquement spécifique envers son substrat ARN 88
1. Modélisation de la structure du domaine nucléotidyltransférase de URT1 avec un transfert des ligands Mg2+ et UTP dans le site actif 88
a) Modélisation du domaine nucléotidyltransférase de URT1 à partir de Cid1 par SWISS-MODEL 89
b) Validation du modèle généré par SWISS-MODEL et transfert des ligands (UTP, magnésium et ARN) 89
1. Diagramme de Ramachandran 89
2. QMEANscore 90
c) Transfert des ligands au modèle généré 91
Validation du transfert de ligands UTP, magnésium et U5P 91
2. Description du site actif du domaine nucléotidyltransférase de URT1 92
3. URT1 est une uridylyltransférase intrinsèquement distributive qui est régulée par l'ion divalent présent lors de la réaction 93
a) URT1 est une uridylyltransférase très spécifique des UTP 93
b) Le magnésium est essentiel à l'activité de URT1 95
4. URT1 est intrinsèquement distributive 95
5. Les régions de liaisons à l'ARN de Cid1 ne sont pas conservées chez URT1 96
6. URT1 uridyle in vitro tous les ARN ayant une extrémité 3'OH libre 96
7. URT1 est active sous forme de monomère 97
8. Conclusion/Ouverture 98

Chapitre 3 100
Résumé 101
1. La région N terminale de URT1 impacte la localisation de URT1 103
2. Une partie de la région N terminale de URT1 est essentielle à sa localisation au niveau des processing bodies 104

3. Le domaine nucléotidyltransférase de URT1 semble être suffisant pour la localisation de URT1 au niveau des processing bodies après un stress thermique 106
4. Le domaine nucléotidyltransférase de URT1 serait aussi suffisant pour la localisation de URT1 au niveau des granules de stress 107
5. Conclusion et ouverture 108

Discussion 110
A. URT1 est une uridylyltransférase intrinsèquement spécifique pour l'UTP 111
B. L'uridylation des ARNm oligoadénylés par URT1 chez A. thaliana 112
1. La reconnaissance des substrats ARN par URT1 112
2. La reconnaissance des ARN oligoadénylés par URT1 113
3. L'implication de la structure de la région N terminale de URT1 dans la reconnaissance de substrat ARN 114
4. Le rôle potentiel des motifs conservés de URT1 dans la reconnaissance de substrats ARN 115
5. URT1 peut éventuellement uridyler d'autres ARN que les ARNm 116
C. La localisation de URT1 au niveau des foci cytoplasmiques 116
D. Le rôle de l'uridylation des ARNm déadénylés 118
1. La polarisation de la dégradation dans le sens 5'-3' des ARNm oligoadénylés ayant une à trois uridines en 3' 119
a) La protection en 3' des ARNm oligoadénylés par URT1 119
b) L'uridylation marque les ARNm pour l'élimination de la coiffe 120
2. L'absence de modification de la demi-vie des ARN dans le mutant urt1-1 121
3. La présence d'oligouridine en 3' des ARNm 122
a) La dégradation des ARNm oligouridylés 122
b) Une seconde uridylyltransférase peut uridyler les ARNm 123
1. La reconnaissance des substrats ARN de la seconde uridylyltransférase uridylant les ARNm 123
2. Les candidats 124
4. Le rôle de l'uridylation des ARNm déadénylés associés aux polysomes 125
La dégradation dans le sens 3'-5' associée aux polysomes ? 126
5. Le rôle de URT1 au niveau des foci cytoplasmiques tels que les processing bodies et les granules de stress 127
a) Les processing bodies 127
b) Les granules de stress 128
Conclusion générale 129

Matériels et méthodes 131
A - Matériels 132
1. Matériel végétal 132

a) Arabidopsis thaliana 132
b) Nicotiana benthamiana 132
2. Souches bactériennes 132
a) Top10 132
b) SurvivalTM 132
c) BL21 DE3 132
d) Agrobacterium tumefaciens GV3101::pMP90 133
3. Plasmides 133
a) Plasmides donneurs 133
b) Plasmides d'expression 133
4. Anticorps 134
a) Anticorps primaires 134
1. Anticorps anti-GFP 135
2. Anticorps anti-c-myc 135
3. Anticorps anti-HA 135
4. Anticorps anti-URT1 135
b) Anticorps secondaires 135
5. Les amorces 135
6. Logiciels de bioinformatiques utilisés pour l'analyse de la séquence de URT1 136
7. Les numéros d'accession des homologues de URT1 138
B - Méthodes 139
1. Modélisation du domaine nucléotidyltransférase de URT1 à partir de la résolution atomique (4FH5) de Cid1 par SWISS-MODEL 139
2. Clonage des gènes 139
3. Mutagenèse dirigée 139
4. Purification des PCR 140
5. Extraction d'ADN plasmidique 140
6. Séquençage d'ADN 140
7. Transformation bactérienne 140
8. Transformation d'Arabidopsis 140
9. Production des protéines recombinantes 141
10. Purification de protéines recombinantes 141
a) Purification par chromatographie d'affinité 141
b) Purification des protéines recombinantes ayant une étiquette histidine par l'agarose Ni-NTA 142
c) Purification des protéines recombinantes ayant une étiquette GST par l'agarose GST Sépharose 142
11. Marquage radioactif d'un oligoribonucléotide par la polynucléotide kinase 143
12. Activité catalytique in vitro de la protéine recombinante URT1 143
13. Précipitation des ARN au TCA 144
14. Gel de polyacrylamide pour ARN 144
15. Chromatographie d'exclusion de taille 144

16. Agro-infiltration de feuilles de Nicotiana benthamiana 144
17. Immuno-précipitation 145
18. Extraction protéique 145
a) Extraction protéique d'une culture bactérienne 145
b) Extraction protéique de feuilles de N. benthamiana 146
19. La résolution des protéines selon leur taille par gel SDS-PAGE 146
20. Coloration des protéines à l'argent 146
21. Coloration des protéines au bleu de coomassie sur gel polyacrylamide 147
22. Transfert des protéines sur une membrane 147
23. Western blot 147
24. Coloration des protéines sur membrane 148
25. Co-localisation des protéines exprimées de façon transitoire 148
26. Microscopie 148
27. Quantification de la colocalisation 148
a) JACoP 148
b) FOCI 149

ANNEXES 150
Annexe 1.1 Séquence protéique de URT1 151
Annexe 1.2 Alignement des homologues de URT1 chez les Angiospermes 152
Annexe 1.3 Prédiction de la désorganisation des homologues d'URT1 chez les Angiospermes par IUPred et ANCHOR 162
Annexe 2.1 Diagramme de Ramachandran de la modélisation de URT1 détaillée (complément de la Figure 2B.1) 169
Annexe D1 Alignement des ncPAP d'A. thaliana par MUSCLE 170

Bibliographie 181

Figures et tableaux

Introduction

Partie A

Figure IA.1 : Les voies principales de la dégradation des ARN
Figure IA.2 : Organisation du complexe de déadénylation Ccr4-Caf1/Pop2-Not chez la levure
Figure IA.3 : La régulation de la déadénylation
Figure 1A4 : Architecture de la RNase PH, des PNPases et de l'exosome
Figure IA.5 : Les cofacteurs de l'exosome chez *S. cerevisiae* et l'homme selon la localisation de l'exosome
Figure IA.6 : Les cofacteurs de l'exosome chez *A. thaliana* selon la localisation de l'exosome

Figure IA.7 : Les facteurs interagissant avec le complexe d'élimination de la coiffe pour la dégradation 5'-3' chez *S. cerevisiae*
Figure IA.8 : Les facteurs du complexe d'élimination de la coiffe chez *A. thaliana*
Figure IA.9: Les mécanismes de clivage des ARNm par les endonucléases
Figure IA.10 : L'implication des ncPAP dans la dégradation des ARN
Tableau IA.1 : Les ncPAP chez *S. cerevisiae*, *S. pombe* et les mammifères
Partie B
Figure IB.1 : L'uridylation des précurseurs des microARN *let-7* a deux rôles antagonistes selon la quantité d'uridines ajoutées en 3'
Figure IB.2 : La voie de biogenèse des piARN
Figure IB.3 : La voie de biogenèse et le fonctionnement des microRNA et des siARN chez les plantes
Figure IB.4 : L'édition des ARNm chez les trypanosomes par la délétion ou l'insertion d'uridines
Figure IB.5 : La biogenèse des ARN guide et des ARN guide-*like* chez les trypanosomes
Partie C
Figure IC.1 : La régulation de la traduction des ARNm
Tableau IC.1 : Les protéines composant les *processing bodies* et les granules de stress chez les mammifères
Figure IC.2 : Les facteurs régulant la formation et la dissociation des *processing bodies*
Tableau IC.2 : Les domaines responsables de la localisation des protéines au niveau des *processing bodies*
Figure IC.3 : Les facteurs régulant la formation et la dissociation des granules de stress
Figure IC.4 : Assemblage du complexe d'initiation de la traduction
Figure IC.5 : DYRK3 régule la stabilité des granules de stress et l'expression de l'ARNm de TORC1
Tableau IC.3 : Les domaines nécessaires à la localisation des protéines au niveau des granules de stress
Figure IC.3 : Les facteurs régulant la formation et la dissociation des granules de stress
Figure IC.4 : Assemblage du complexe d'initiation de la traduction
Figure IC.5 : DYRK3 régule la stabilité des granules de stress et l'expression de l'ARNm de TORC1
Figure IC.6 : Le « cycle de vie » des granules de stress
Figure IC.7 : L'assemblement des granules de stress et des *processing bodies* et leur interférence par les virus chez les mammifères.
Figure IC.8 : Les composants des *processing bodies* et des granules de stress peuvent être détournés par les virus pour leur réplication
Figure IC.9 : Le cycle dynamique de l'ARNm entre les polysomes, les granules de stress et les *processing bodies*

Chapitre 1

Figure 1.1 : URT1 possède deux domaines homologues à Cid1
Figure 1.2 : Prédiction des structures secondaires de URT1
Figure 1.3: Représentation schématique des régions de faibles complexités de URT1
Tableau 1.1 : La région N terminale de URT1 est appauvrie en acides aminés promouvant l'ordre et principalement enrichie en acides aminés promouvant la désorganisation
Figure 1.4 : La région N terminale de URT1 est prédite intrinsèquement désorganisée
Figure 1.5 : La région N terminale de URT1 contient des zones prédites d'interaction protéine-protéine
Figure 1.6: Les monocotylédones possèdent deux homologues de URT1 (URT1A et URT1B)
Figure 1.7: La région N terminale de URT1 contient des régions et des motifs conservées chez les Angiospermes

Chapitre 2

Partie A

Figure 2A.1 : Le nombre d'uridines ajoutées en 3' est plus important lorsque le nombre d'adénines est faible pour les séquences uridylées issues de l'analyse des ARNm dans le mutant *xrn4-3*
Figure 2A.2 : Le nombre d'uridines ajoutées en 3' est plus important lorsque le nombre d'adénines est faible pour les ARNm associés aux polysomes dans le mutant *xrn4-3*

Partie B

Figure 2B.1 : Représentation du diagramme de Ramachandran de la structure 3D modélisée du domaine nucléotidyltransférase de URT1
Figure 2B.2 : Validation statistique du modèle de URT1 et mise en évidence des erreurs prédites par résidus
Figure 2B.3 : Alignement 3D du modèle du domaine nucléotidyltransférase de URT1 et des trois autres uridylyltransférases résolues au niveau atomique
Figure 2B.4 : Transposition des ligands UTP, magnésium et U5P dans le site actif de URT1 modélisé
Figure 2B.5 : Représentation de la structure 3D modélisée du domaine nucléotidyltransférase de URT1
Figure 2B.6 : Les acides aminés importants pour la spécificité de l'UTP chez les autres uridylyltransférases sont conservés chez URT1
Figure 2B.7 : L'activité catalytique de URT1 est régulée par l'ajout de magnésium
Figure 2B.8: URT1 est intrinsèquement distributive au niveau des premières uridines ajoutées
Figure 2B.9 : Les régions de liaison à l'ARN de Cid1 ne sont pas conservées chez URT1

Figure 2B.10 : Représentation de la surface du domaine nucléotidyltransférase de URT1 selon les propriétés chimiques des acides aminées
Figure 2B.11: URT1 est capable d'uridyler *in vitro* les ARN ayant une extrémité 3' libre et non-modifiée
Figure 2B.12 : URT1 ne possède pas d'équivalent à l'arginine nécessaire à la spécificité du substrat ARN de *Tb*TUT4 et *Tb*RET2
Figure 2B.13: URT1 est active sous forme de monomère

Chapitre 3

Figure 3.1 : Les délétions de URT1 utilisées dans les expériences
Figure 3.2 : L'expression de URT1 et de ses délétions avec une étiquette YFP fusionnée en N et en C terminal est similaire
Figure 3.3: La localisation de URT1 et de ses délétions
Figure 3.4: La localisation de URT1 et Ntrf ne dépend pas de l'activité nucléotidyltransférase
Figure 3.5: La co-expression de DCP1 ne change pas la localisation des différentes délétions de URT1
Figure 3.6: Une partie de la région N terminale de URT1 est essentielle à la localisation de URT1 au niveau des *processing bodies*
Figure 3.7 : La localisation de URT1 et de Ntrf après un stress thermique
Figure 3.8 : Ntrf semble suffisant à la localisation de URT1 au niveau des *processing bodies* après un stress thermique
Figure 3.9 : La localisation des marqueurs des granules de stress après un stress thermique
Figure 3.10 : Ntrf serait suffisant à la localisation de URT1 au niveau des granules de stress avec le marqueur des granules de stress RBP47
Figure 3.11 : Ntrf est suffisant à la localisation de URT1 au niveau des granules de stress avec le marqueur PAB2

Discussion

Figure D1 : Outils obtenus lors de ma thèse pour la recherche de protéines partenaires de URT1
Figure D2 : PAB2 inhibe l'activité de URT1
Figure D3 : La région en N terminal de URT1 est essentielle à l'activité optimale de URT1
Figure D4 : L'étiquette YFP en N terminal de URT1 provoque du *silencing* de URT1
Figure D5 : Les anticorps dirigés contre URT1 ne reconnaissent pas le domaine nucléotidyltransférase
Figure D6 : Les ARNm du double mutant *urt1-1/xrn4-3* sont uridylés par une autre uridylyltransférase que URT1
Figure D7 : Arbre phylogénique des ncPAP d'*A. thaliana*
Figure D8 : Le mécanisme d'uridylation des ARNm oligoadénylés par URT1 chez *A. thaliana*

Figure D9 : Les rôles hypothétiques de l'uridylation des ARNm déadénylés par URT1 chez *A. thaliana* dans « le cycle de l'ARNm »

Introduction :
L’uridylation et la dégradation des ARN

La dégradation des ARN est une composante essentielle de la régulation de l'expression génétique. La dégradation des ARN participe à réguler leur quantité dans la cellule et à éliminer les ARN en « fin de cycle». Elle est également essentielle pour l'élimination des transcrits aberrants et des sous-produits de la maturation des ARN issus de la biogenèse d'ARN codants ou non-codants à partir de transcrits primaires. La dégradation des ARN évite ainsi leur accumulation dans la cellule qui peut avoir un effet toxique et, de manière générale, elle permet le recyclage des ribonucléotides.

Dans une première partie de l'introduction, je détaille les principaux acteurs de la dégradation de 5' en 3' ou de 3' en 5' de l'ARN et brièvement les endonucléases qui clivent au sein de la séquence de l'ARN. Je présente aussi certains facteurs impliqués dans la dégradation tels que la famille des poly(A) polymérases non canoniques (ncPAP pour *non-canonical PolyA Polymerase*) ou le complexe Lsm1-7. Dans la suite de l'introduction, je détaille les substrats ARN uridylés connus ainsi que le nombre limité d'uridylyltransférases caractérisées à ce jour. L'uridylation des ARN est un mécanisme impliqué dans la régulation de la stabilité des ARN, la maturation des précurseurs et la dégradation des ARN dans le sens 5'-3' et/ou 3'-5'. Dans une troisième partie de l'introduction, je présente les rôles connus ou supposés d'aggrégats de complexes RiboNucléoProtéiques (RNP) cytoplasmiques tels que les *processing bodies* et les granules de stress dans la dégradation et/ou le stockage des ARNm. En effet, afin d'optimiser la régulation de l'expression génétique, les ARNm non traduits peuvent être stockés dans des foci cytoplasmiques lors de stress et être triés afin d'être dirigés vers la dégradation ou de retourner à la traduction au niveau des polysomes. L'interconnexion entre les foci cytoplasmiques et les polysomes est mis en évidence afin de présenter un « cycle dynamique du métabolisme d'un ARNm ».

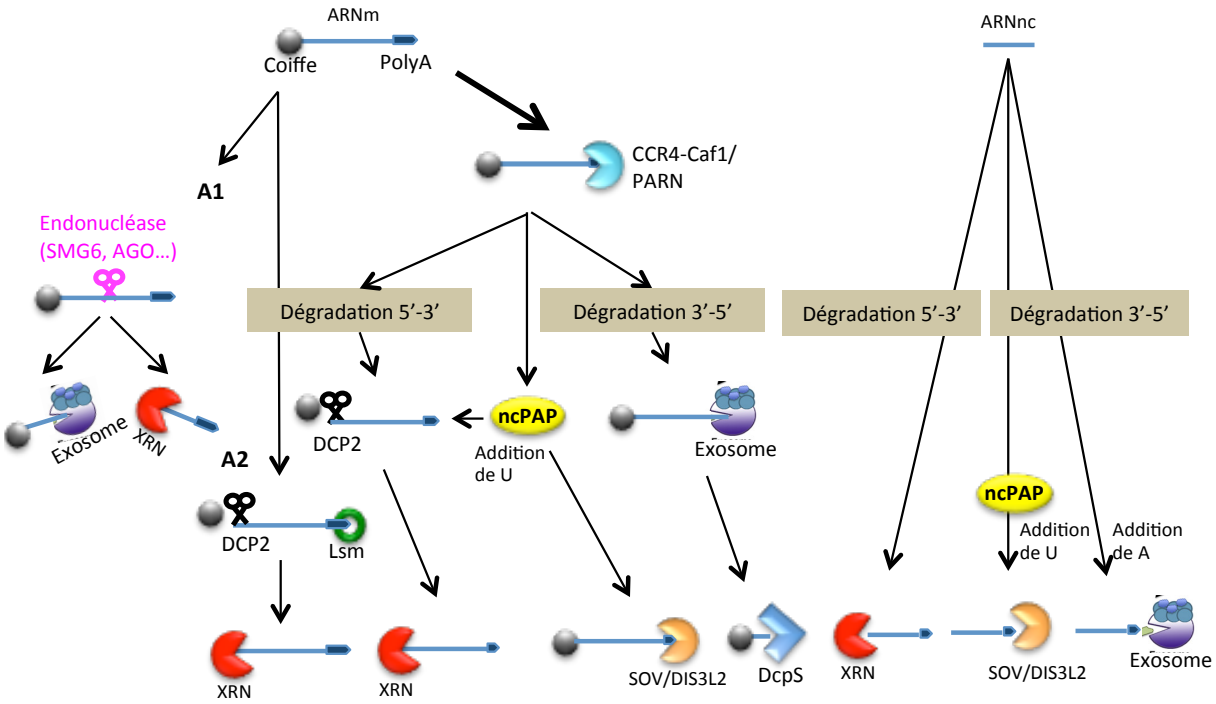


Figure IA.1 : Les voies principales de la dégradation des ARN

Il existe plusieurs facteurs impliqués dans la dégradation : les exoribonucléases progressant de 5' en 3' ou de 3' en 5' ou les endonucléases clivant l'ARN au sein de la séquence de l'ARN. Après dégradation de l'ARNm dans le sens 3'-5', il reste un fragment d'ARN avec la coiffe en 5' (m7GpppN). Elle est libérée et hydrolysée par la pyrophosphatase DCPS/DCS1 de la famille de HIT en m7GMP et laissant une extrémité 5' monophosphate à l'ARN (van Dijk et al., 2003). La dégradation fait aussi intervenir des protéines de liaison à l'ARN ou des PAP non canoniques. A1 Le clivage endonucléolytique de l'ARNm. Il intervient après un contrôle qualité dans la voie du *Nonsense Mediated Decay*, la voie du *No-Go decay* ou encore après la reconnaissance de l'ARN cible par les petits ARN. A2 La dégradation des ARNm indépendamment de la déadénylation. Elle s'effectue dans le sens 5'-3' pour le *Non-Stop Decay* ou après uridylation. D'après (Chang et al., 2013; Lubas et al., 2013; Malecki et al., 2013; Nagarajan et al., 2013; Parker, 2012; Rissland and Norbury, 2009; Shen and Goodman, 2004).

La dégradation de l'ARN

Les plantes, étant statiques, ont besoin de s'adapter continuellement aux modifications de leur environnement. Les stimuli extérieurs et les différentes phases développementales induisent des modifications appropriées de leur expression génétique, notamment via la régulation de la transcription des gènes, la traduction des ARNm et la dégradation des ARN codants ou non-codants.

De manière générale, il existe deux types de dégradation des ARN dans la cellule : la dégradation impliquée dans le *turnover* des ARN et la dégradation intervenant après un contrôle qualité, lorsqu'un ARN est détecté comme étant aberrant (Figure IA.1). La dégradation ou *turnover* des ARN concerne tous les ARNm et les ARN non-codants (ncARN) « en fin de cycle» et peut être régulée en réponse à des signaux développementaux ou environnementaux. Les mécanismes de régulation de la dégradation des ARN sont multiples et peuvent, par exemple, faire intervenir la reconnaissance des ARN cibles par des protéines régulatrices, des petits ARN ou impliquer l'adénylation ou l'uridylation des ARN. La dégradation des ARN est également essentielle à l'élimination des sous-produits issus de la biogenèse des ARN fonctionnels. En effet, lors de la biogenèse des ARN, plusieurs fragments d'ARN dérivés de la maturation des précurseurs des ARN sont produits comme par exemple les introns des ARNm, les fragments d'ARN excisés en 3' ou en 5' des ARN de transfert (ARNt), des microARN ou des ARN ribosomiques (ARNr). La dégradation des ARN peut aussi survenir après un contrôle de la qualité des ARN afin d'éliminer les ARN aberrants, potentiellement toxiques. En effet, lors de la biogenèse des ARN, plusieurs erreurs peuvent survenir lors de leur transcription et de leur maturation, formant ainsi des ARN défectueux. La dégradation des ARN représente donc un mécanisme essentiel pour le contrôle quantitatif et qualitatif de l'expression des gènes (Belostotsky et al., 2012; Eberhardt et al., 2007; Xu et al., 2006).

La dégradation des ARN implique trois types d'activités ribonucléasiques : les exoribonucléases qui progressent de 5' en 3', les exoribonucléases qui progressent de 3' en 5' et les endoribonucléases qui clivent l'ARN au sein de sa séquence. Les mécanismes de dégradation des ARN diffèrent selon la voie de dégradation, les ARN cibles, leur localisation cellulaire et les partenaires protéiques recrutés (Figure IA.1).

A - Les facteurs impliqués dans la dégradation

Nos connaissances sur les mécanismes de dégradation des ARN chez les eucaryotes reposent amplement sur des études effectuées chez la levure *Saccharomyces cerevisiae.* Cet organisme a été et continue à être un organisme modèle exceptionnel pour décrypter des mécanismes fondamentaux de la dégradation des ARN. Il est important de préciser que les mécanismes de dégradation des ARN chez cette levure ne sont pas aussi complexes que chez d'autres organismes (Parker, 2012). Par exemple, les voies de dégradation *via* les petits ARN sont absentes chez *S. cerevisiae*. De même, cette levure possède seulement deux protéines de la famille des polyA polymérases non canoniques impliquées dans la dégradation des ARN et aucune uridylation des ARN n'a été mise en évidence. Cependant, les principaux acteurs de la dégradation en 3', en 5' ou au sein de l'ARN sont très conservés chez les différents organismes de la levure à l'homme. Pour chacun des facteurs principaux impliqués dans la dégradation des ARN, je présente leurs rôles et leurs modes d'action, leurs homologues chez l'homme et les plantes, ainsi que les principaux cofacteurs régulant leur activité. La nomenclature des ARN et des protéines de chacun des organismes est respectée.

2. Les exoribonucléases agissant en 3' de l'ARN

a) Les déadénylases

La déadénylation constitue la première étape de la dégradation de la plupart des ARNm. Le raccourcissement de la queue polyadénylée ne permet plus la liaison des protéines de liaison aux queues polyA (PABP) qui protègent la queue polyA et permettent la circularisation des ARNm par leur interaction avec le complexe de liaison de la coiffe (Tomek and Wollenhaupt, 2012). En effet, il est estimé que les PABP ont besoin de 10 à 12 A chez les plantes et les champignons ou de 5 à 25 A chez les mammifères pour se lier. Le désengagement des PABP de l'ARNm rend son extrémité 3' accessible aux exoribonucléases dans le sens 3'-5', principalement l'exosome ou Dis3L2, lorsque les ARNm sont oligouridylés (Lykke-Andersen et al., 2009; Malecki et al., 2013; Parker, 2012), ou permet la liaison du complexe Lsm1-7 en 3' (partie A4 de l'introduction) (Chowdhury et al., 2007). La liaison du complexe Lsm1-7 provoque le recrutement du complexe d'élimination de la coiffe dirigeant ainsi la dégradation dans le sens 5'-3' (Introduction partie A2) (Song and Kiledjian, 2007).

Deux principaux complexes de déadénylation sont présents chez *S. cerevisiae* et les mammifères : le complexe Pab1-dépendant polyA nucléase (PAN)

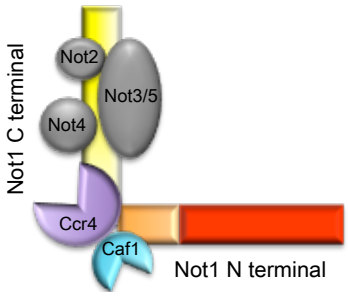


Figure IA.2 : Organisation du complexe de déadénylation Ccr4-Caf1/Pop2-Not chez la levure

Le domaine N terminal de Not1 est en rouge, le domaine C terminal en jaune et en orange la région des acides aminés 754 à 1000, nécessaire à l'interaction de Caf1. Not2, Not3, Not4 et Not5 se lient à la partie C terminale de Not1. Caf40 et Caf130 ne sont pas indiquées car leur région d'interaction n'est pas connue. D'après (Basquin et al., 2012).

et le complexe Ccr4-Pop2/Caf1-Not (*Carbon Catabolite Repressor* 4 et *CCR4 Associated Factor* 1) (Tucker et al., 2001). La déadénylation s'effectue en deux temps. Premièrement, le complexe PAN réduit la taille de la queue polyA d'environ 200 et 90A à 110 et 70A, respectivement chez les mammifères et la levure. La suite de la déadénylation est réalisée par le complexe Ccr4-Pop2/Caf1-Not mais également par PARN (Poly(A) spécifique RiboNucléase) qui est absent chez *S. cerevisiae*.

1. Le complexe PAN

Le complexe PAN est composé de Pan2 et Pan3. Pan2 appartient à la famille des RNase D et est la sous-unité catalytique du complexe PAN (Boeck et al., 1996). Pan3 est, quant à elle, la sous-unité de régulation du complexe (Brown et al., 1996). Ce complexe est activé par la protéine majeure de liaison aux queues polyA Pab1 (Boeck et al., 1996). L'action du complexe PAN est très rapide. En effet, très peu de queues polyA supérieures à 70A sont retrouvées chez la levure. Par contre, chez le mutant ΔPan2 les tailles des queux polyA varient selon les gènes mais elles peuvent atteindre environ 90A (Brown and Sachs, 1998). Dans le contexte du mutant nul ΔCcr4, Pan2 est capable de poursuivre la déadénylation à la place de Ccr4 mais elle agit plus lentement que Ccr4. L'activité de Pan2 en absence de Ccr4 s'arrête lorsque les queues polyA sont réduites entre 20 à 25 résidus. Sachant que Pan2 est activé par Pab1, cet arrêt de la déadénylation suggère une absence de Pab1 liée à la queue polyA à une taille inférieure à 20-25 adénines chez la levure (Tucker et al., 2001).

2. Le complexe Ccr4-Pop2/Caf1-Not

Lorsque les queues polyA des ARNm atteignent une taille de 110 et 70A chez les mammifères et la levure, le complexe Ccr4-Pop2/Caf1-Not peut intervenir. Le complexe Ccr4-Pop2/Caf1-Not est très conservé de la levure à l'homme et est impliqué dans de nombreux processus tels que l'initiation et le contrôle de l'élongation de la traduction, la déadénylation et la dégradation des ARNm (Chen and Shyu, 2011; Collart and Panasenko, 2012; Denis and Chen, 2003; Zhang et al., 2010b). Le complexe Ccr4-Pop2/Caf1-Not est formé de 7 protéines chez la levure Ccr4, Pop2, Not1, Not2 à 5, Caf40 et Caf130 (Denis and Chen, 2003). Ccr4 et Pop2/Caf1 sont deux 3'-5' exonucléases. Ccr4 fait partie de la famille des nucléases ExoIII et est la nucléase majeure actrice du complexe (Tucker et al., 2001, 2002). En effet, l'activité de la seconde nucléase Pop2/Caf1, faisant partie de la famille des RNaseD, n'est pas requise *in vitro* (Thore et al., 2003; Tucker et al., 2002). Not1 est

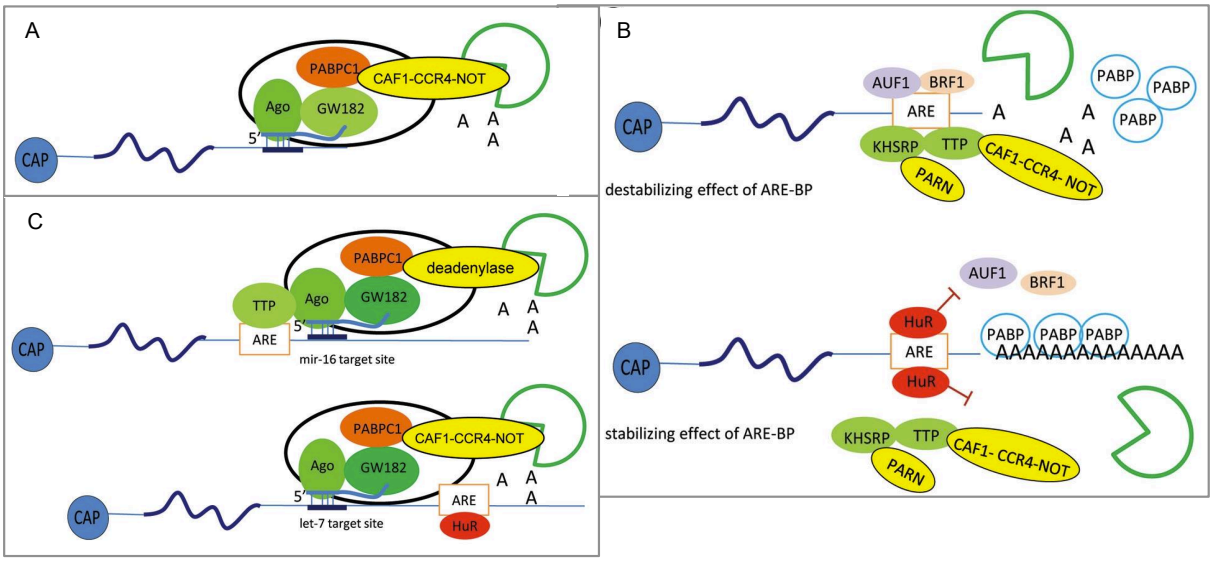


Figure IA.3 : La régulation de la déadénylation

A La déadénylation induite par les microARN chez les mammifères. Le complexe RISC, contenant Ago, GW182, PABPC1 et CCR4-CAF1-NOT ou PAN, délivre le microRNA à l'ARNm cible et induit la déadénylation de l'ARNm, permettant ainsi sa dégradation. **B** La déadénylation des ARNm possédant des séquences ARE. Les protéines de liaisons aux séquences ARE (BP) induisent une déstabilisation et/ou une stabilisation de l'ARNm. AUF1, TTP, BRF1 et KHSRP, recrutent les déadénylases, PARN et CCR4-CAF1-NOT et induisent la déadénylation et la dégradation de l'ARNm. A l'inverse, d'autres BP, comme HuR, stabilisent les ARNm en bloquant la liaison d'autres BP impliquées dans la déstabilisation. Cette compétition stabilise l'association des PABP et prévient le recrutement des déadénylases. **C** La déadénylation induite par les séquences ARE et les microARN. La coopération des BP, des microARN, des déadénylases et de l'exosome est essentielle pour la régulation de la stabilité des ARNm. Le recrutement des BP, HuR ou TTP aux séquences ARE cible de *let-7* ou du complexe RISC contenant le microARN-16, s'effectue au site le plus près de la séquence ARE. D'après (Zhang et al., 2010).

la plateforme de liaison des deux nucléases à l'ARN qui interagit au niveau de sa région C terminale avec Not2, Not3/5 et Not4, formant le module Not (Figure IA.2) (Basquin et al., 2012). Le rôle des protéines Caf40 et Caf130 n'est pas déterminé pour l'instant, ni leur région d'interaction avec Not1.

3. PARN

PARN est une déadénylase présente chez de nombreux eucaryotes mais absente chez *S. cerevisiae*. Chez les mammifères, PARN est processif et interagit avec la coiffe et la queue polyA des ARNm (Balatsos et al., 2012; Virtanen et al., 2013). PARN est impliqué dans la maturation des H/ACA snoRNA chez les mammifères. En effet, ils sont premièrement polyadénylés par PAPD5 puis PARN dégrade la queue polyA ainsi que les 5 à 8 nucléotides de l'extrémité 3' du snoARN H/ACA afin de le rendre fonctionnel (Berndt et al., 2012).

4. La régulation de la déadénylation

La déadénylation des ARN est principalement régulée par Pab1, les microARN et d'autres facteurs de liaison à l'ARNm comme les protéines de liaison aux séquences riches en A/U (ARE) qui sont activées en cas de stress (Zhang et al., 2010b). En effet, en condition de stress, les protéines de liaison aux séquences ARE comme TTP, BRF1 sont activées et répriment la traduction des ARNm en recrutant le complexe Ccr4-Pop2/Caf1-Not et PARN. D'autre part, les microARN, peuvent induire la déadénylation des ARNm cibles en recrutant Ccr4-Pop2/Caf1-Not ou le complexe PAN *via* le complexe RISC (Figure IA.3) (Zhang et al., 2010b).

Dans certains cas, l'étape de déadénylation peut être un processus réversible. En effet, les ARN déadénylés peuvent être à nouveau polyadénylés dans le cytoplasme comme par exemple chez les oocytes des mammifères et les cellules des neurones (Udagawa et al., 2012; Villalba et al., 2011; Weill et al., 2012). Ces mécanismes antagonistes de déadénylation et d'adénylation créent une régulation de la traduction par une répression de celle-ci lors de la déadénylation et à l'inverse, une activation de la traduction lors de la polyadénylation des ARNm (Weill et al., 2012; Zhang et al., 2010b).

5. Les déadénylases chez *A. thaliana*

Les complexes PAN et CCR4-CAF1-NOT, ainsi que PARN, sont présents chez *Arabidopsis thaliana* et certaines de ces protéines sont codées par de petites familles multigéniques. *A. thaliana*, possède ainsi neuf homologues de Pan2, douze homologues de la nucléase CCR4 et quatorze homologues de la nucléase CAF1 d'après la base de donnée de Phytozome (Goodstein et al., 2011). Les deux

homologues de CAF1, AtCAF1a et AtCAF1b, ont un rôle complémentaire et sont induits par différents stress environnementaux et lors d'un stress biotique par *Pseudomonas syringae* (Belostotsky and Sieburth, 2009; Liang et al., 2009; Walley et al., 2010). PARN est essentielle chez *A. thaliana* et les plantes supérieures car une mutation nulle est létale (Chiba et al., 2004; Reverdatto et al., 2004). PARN est nécessaire à la dégradation des ARNm en réponse aux stress par l'acide abscissique, salin et osmotique (Nishimura et al., 2005). D'après un article très récent, PARN semble aussi localisée dans les mitochondries et dégrade les extensions poly(A) ajoutée par une poly(A) polymérase non canonique (Hirayama et al., 2013).

Après déadénylation, la taille de la queue polyA des ARNm atteint environ 15 A chez les champignons et les plantes et de 15 à 25 A chez les mammifères. Ces ARN oligoadénylés ne sont plus protégés en 3' par les PABP. Leurs extrémités 3' libres permettent la liaison du complexe Lsm1-7 en 3' de l'ARN. Ils deviennent ainsi des substrats du complexe de l'élimination de la coiffe dirigeant la dégradation dans le sens 5'-3'. Ils peuvent aussi, en absence du complexe en 3', être dégradés dans le sens 3'-5', par l'exosome ou Dis3L2 (Lykke-Andersen et al., 2009; Malecki et al., 2013; Parker, 2012).

b) Le complexe de l'exosome

L'exosome est un complexe protéique impliqué dans la dégradation de 3' en 5' des ARN et dont la structure générale est conservée de la levure à l'homme (Figure IA.4) (Büttner et al., 2006; Mitchell et al., 1997). Un complexe similaire est également retrouvé chez les Archaea (Lykke-Andersen et al., 2009). L'exosome des eucaryotes et celui des Archaea sont structurés en forme d'anneau, une organisation très similaire à celle observée pour la RNase PH de *E. coli* et les polynucléotide phosphorylases (PNPases) bactérienne, chloroplastique et mitochondriale (Lykke-Andersen et al., 2011). La RNase PH est un homohéxamère tandis que l'anneau des PNPases est formé par l'association de trois sous-unités comprenant chacune deux domaines de type RNase PH, dont un seul est actif. Chez les eucaryotes, l'exosome est composé d'un cœur auquel s'associe des co-facteurs. Le cœur de l'exosome est composé de 9 sous-unités : Csl4, Rrp4, Rrp40, Rrp41, Rrp42, Rrp43, Rrp45, Rrp46 et Mtr3. Ce cœur peut être divisé en deux parties. La première est formée des six protéines Rrp41, Rrp42, Rrp43, Rrp45, Rrp46 et Mtr3 qui forment l'anneau (Figure IA.4) (Mitchell et al., 1997). La deuxième partie du cœur de l'exosome forme une coiffe composée de Csl4, Rrp4 et Rrp40, qui sont des protéines de liaison à l'ARN

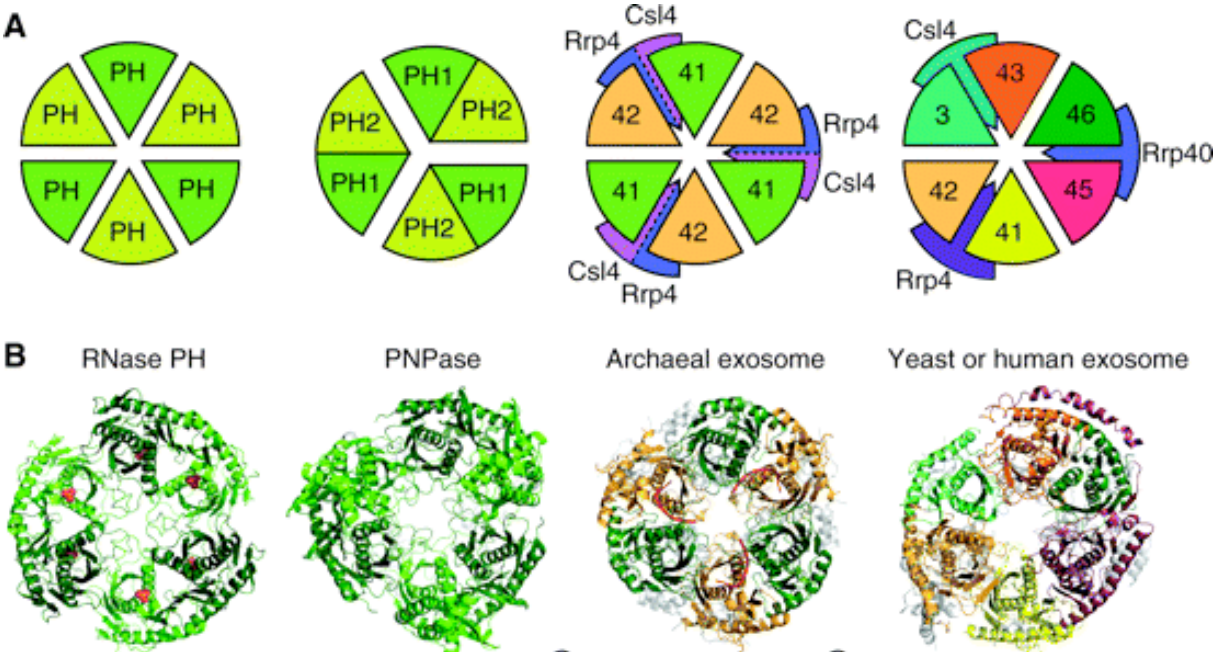


Figure 1A4 : Architecture de la RNase PH, des PNPases et de l'exosome

A Représentation schématique de la RNase PH, des PNPases bactériennes, de l'exosome chez les archées et de l'exosome chez l'homme et la levure. **B** Structure des sous-unités de la RNase PH, de la PNPase et du cœur de l'exosome. Les couleurs utilisées en A et B sont les mêmes. D'après (Lykke-Andersen et al., 2009).

possédant un domaine S1. Rrp4 et Rrp40 possèdent aussi un domaine analogue à la protéine K (K-homologue ou KH). La coiffe est appelée la plateforme S1 de l'hexamère (Figure IA.4). Chez la levure et l'homme, aucune des six sous-unités formant l'anneau de l'exosome ne possède d'activité ribonucléasique, et ce, malgré leur homologie de séquence évidente à la RNase PH. En effet, des mutations dans des acides aminés critiques pour l'activité RNasique sont présentes pour chacune des six sous-unités de l'anneau de l'exosome. L'activité RNasique de l'exosome est apportée par l'association avec des RNases, Rrp6 et Rrp44/Dis3 (Lykke-Andersen et al., 2009, 2011; Tomecki et al., 2010). D'autres co-facteurs de l'exosome incluent des hélicase ARN, des protéines de liaison à l'ARN et des poly(A) polymérases non-canoniques et sont impliqués dans l'activation de l'exosome et la reconnaissance des substrats ARN (Lange and Gagliardi, 2010; Lange et al., 2011; Lykke-Andersen et al., 2011). Chez les eucaryotes, les cofacteurs associés avec le cœur de l'exosome sont différents selon la localisation intracellulaire de l'exosome, l'exosome étant présent dans le cytoplasme et le noyau de toute cellule eucaryote (Figure IA.5) (Butler and Mitchell, 2011; Lubas et al., 2011; Lykke-Andersen et al., 2011; Raijmakers et al., 2004).

1. Les facteurs associés à l'exosome dans le noyau

L'exosome nucléaire s'associe avec les facteurs Rrp44 (ou Dis3), Rrp47, Rrp6, Mpp6 (M-*phase phosphoprotein* 6) ainsi que le complexe TRAMP (Trf4-Air1-Mtr4 Polyadénylation). La sous-unité Rrp44/Dis3, de la famille des RNase R et RNase II, possède une activité 3' exonucléase hydrolytique processive (Lykke-Andersen et al., 2011). Rrp44/Dis3 possède un domaine Pilt N terminus (PIN), très conservé, ayant une activité endonucléase. Il existe deux paralogues de Rrp44 chez l'homme, appelés hDIS3 et hDIS3L interagissant avec l'exosome localisés respectivement dans le nucléole ou le cytoplasme (Lykke-Andersen et al., 2011; Staals et al., 2010; Tomecki et al., 2010). Une autre protéine de cette famille des RNase R, hDis3L2, quant à elle n'interagit pas avec l'exosome (Astuti et al., 2012; Lubas et al., 2013). Rrp6 possède une activité hydrolytique distributive et appartient à la famille des RNase D. Rrp6 et Rrp44/Dis3 interagissent au dessus et dessous, respectivement de la plateforme PH. Mpp6 est impliqué dans la maturation des ARN 5,8S et la dégradation des *cryptic unstable transcripts* (CUTs), *unstable transcripts* (UNTs), *promoter upstream transcripts* (PROMPTs) et des ARN aberrants (Milligan et al., 2008). D'autre part, l'exosome s'associe dans le noyau avec le complexe Nrd1-Nab3-Sen1 pour la terminaison de la transcription des snARN, snoARN (snoARN) et la dégradation des CUTs (Arigo et al., 2006; Creamer et al., 2011;

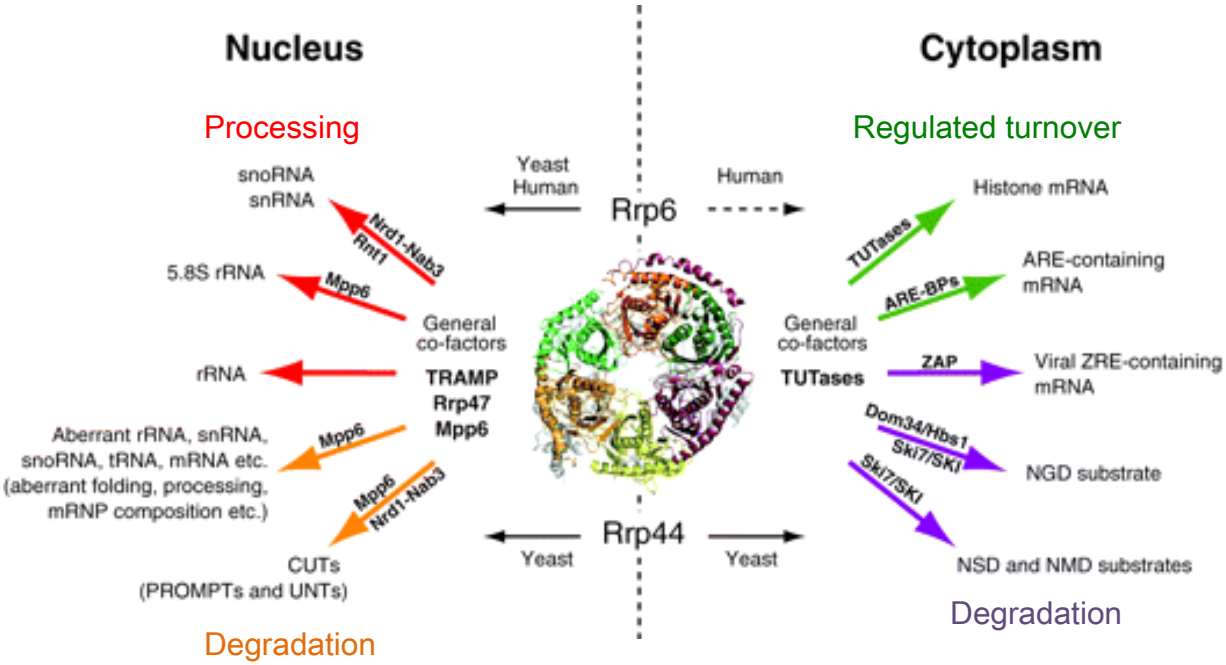


Figure IA.5 : Les cofacteurs de l'exosome chez *S. cerevisiae* et l'homme selon la localisation de l'exosome

Les cofacteurs de l'exosome chez *S. cevevisiae* et l'homme dépendent de la localisation subcellulaire de l'exosome et ciblent différents ARN. D'après (Lykke-Andersen et al., 2009).

Rasmussen and Culbertson, 1998; Steinmetz et al., 2001; Thiebaut et al., 2008; Ursic et al., 1997). Enfin, le complexe TRAMP contient l'hélicase ARN Mtr4 associée à une protéine de liaison à l'ARN (Air1 ou Air2) et à une ncPAP (Trf4 ou Trf5) (Introduction partie A4). Le complexe TRAMP marque les ARN par l'ajout d'adénosines en 3' des ARN *via* les ncPAP Trf4 ou Trf5 pour la dégradation par l'exosome. Le complexe TRAMP stimule une dégradation processive *in vitro via* l'ATPase hélicase Mtr4 qui permet de garder l'ARN déstructuré (Butler and Mitchell, 2011). Le complexe Ccr4-Not a aussi été décrit comme pouvant interagir avec le complexe TRAMP et l'exosome nucléaire mais la fonction de cette association reste à être déterminée (Azzouz et al., 2009).

Dans le noyau, l'exosome peut donc s'associer au complexe TRAMP, à Rrp44, Rrp47, Rrp6, Mpp6 et au complexe Nrd1-Nab3-Sen1 afin de dégrader ou maturer les ARN ribosomiques (ARNr), les petits ARN nucléaires, les petits ARN nucléolaire, les pré-ARN non épissés et les transcrits aberrant (Butler and Mitchell, 2011; Büttner et al., 2006; Lykke-Andersen et al., 2011).

2. Facteurs associés à l'exosome dans le cytoplasme

Dans le cytoplasme, l'exosome dégrade les ARNm ayant des séquences ARE, les ARNm viraux, les ARNm « en fin de cycle» et les ARNm défectueux en s'associant avec différents cofacteurs (Figure IA.5) (Lykke-Andersen et al., 2011). Il s'associe au complexe Ski (Ski2, Ski3, Ski8) par l'intermédiaire de Ski7, une protéine de liaison contenant un domaine GTPase (Araki et al., 2001). Ski7 se lie à la protéine de liaison à l'ARN de la coiffe de l'exosome Csl4. Le complexe Ski est essentiel pour activer la dégradation par l'exosome d'ARN cytosoliques. Il permet notamment la dégradation des ARNm ne contenant pas de codon stop par la voie du NSD (*Non-Stop Decay*) et la dégradation de longs ARN non codants XRN1-*insensitive Unstable-Transcripts* (XUTs) (Badis et al., 2004; Geisler et al., 2012; Muhlrad and Parker, 2005). L'exosome dégrade aussi, associé avec le complexe Ski, les ARNm ayant un codon stop précoce par la voie du NMD (*Non-sens Mediated Decay*) et les ARNm tronqués ou ayant une structure secondaire forte dans la séquence codante par la voie de dégradation du NGD (*No-Go Decay*) en s'associant aux protéines de liaison à l'ARNm Dom34 et Hbs1 (Introduction partie A3) (Garneau et al., 2007; van Hoof et al., 2002).

3. L'exosome chez *A. thaliana*

L'organisation structurale du cœur de l'exosome est conservée chez les plantes (Lange and Gagliardi, 2010). Cependant, différemment de la levure et de

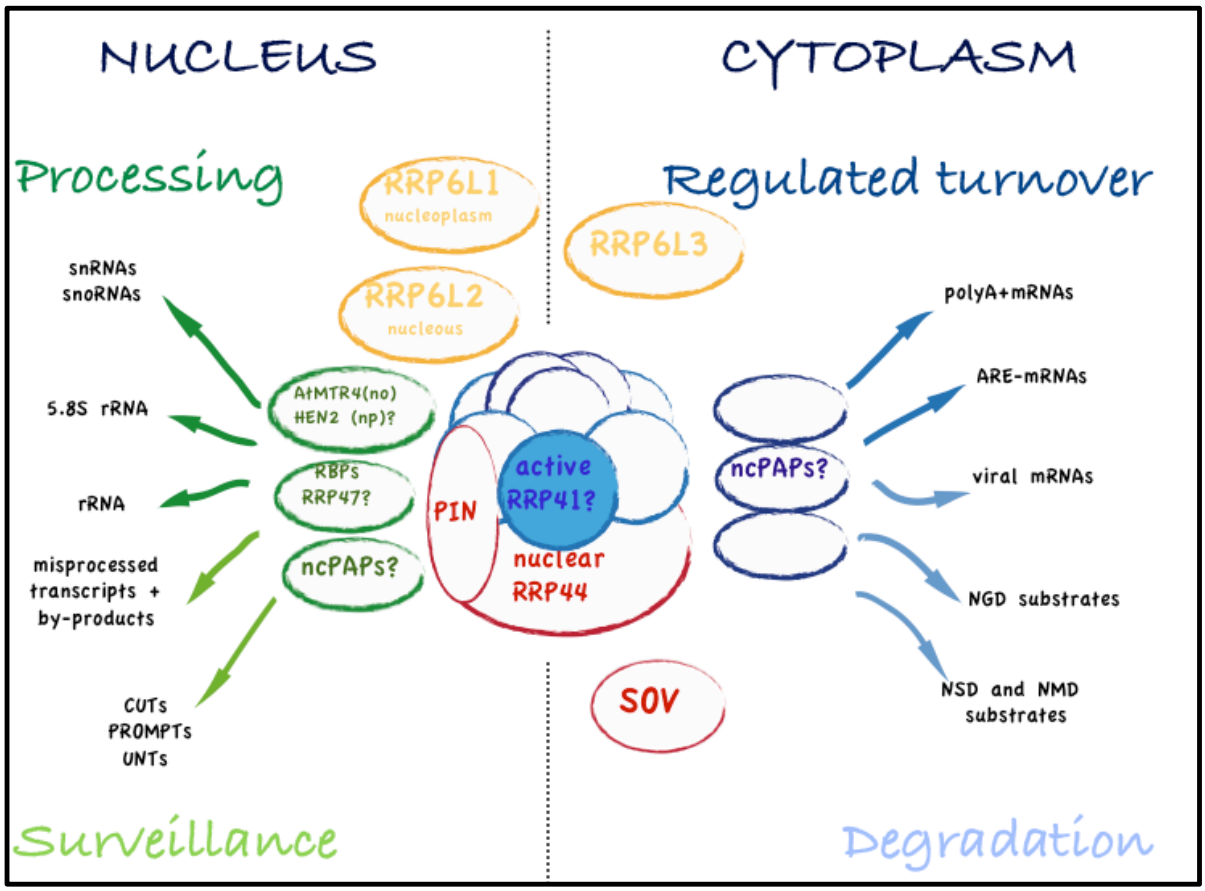


Figure IA.6 : Les cofacteurs de l'exosome chez *A. thaliana* selon la localisation de l'exosome

Les cofacteurs de l'exosome chez *A. thaliana* varient selon sa localisation subcellulaire et les ARN cibles. D'après (Chekanova et al., 2000; Lange et al., 2008, 2011).

l'homme, chez *A. thaliana*, la sous-unité RRP41 conserve les résidus nécessaires à l'activité phosphorolytique (Chekanova et al., 2000). D'autre part, les plantes possèdent deux paralogues de RRP44, appelées AtRRP44 et SOV (*Suppresor Of Varicose*). SOV, n'est pas associée à l'exosome et est l'homologue de Dis3L2 chez l'homme. De plus, chez Arabidopsis, il existe trois homologues du cofacteur Rrp6 : RRP6L1, RRP6L2 et RRP6L3 qui ont toutes trois une localisation différente : RRP6L1 est située dans le nucléoplasme, RRP6L2 dans le nucléole et RRP6L3 dans le cytoplasme (Chekanova et al., 2007). *A. thaliana* possède deux hélicases HEN2 et MTR4, homologues à Mtr4. MTR4 est localisée au niveau du nucléole et HEN2 dans le nucléoplasme (Figure IA.6) (Lange et al., 2008). MTR4 et RRP6L2 coopèrent pour effectuer la maturation et la dégradation des ARNr (Lange et al., 2011). Le complexe Ski est aussi conservé chez les plantes et l'homologue de Ski8, VIP3 (*Vernalization Independant 3*) est associé avec le complexe SKI lors de la dégradation des ARNm (Dorcey et al., 2012; Lange and Gagliardi, 2010). Les ncPAP (Introduction partie A4) peuvent aussi être des cofacteurs de la dégradation des ARN par l'exosome dans le noyau et, peut être aussi dans le cytoplasme (Lange et al., 2009).

c) Dis3L2/SOV

Une nouvelle voie de dégradation 3'-5' indépendante de l'exosome a été mise à jour récemment. Comme décrit précédemment, la sous-unité de l'exosome Rrp44/Dis3, chez l'homme et les plantes, possède plusieurs paralogues. SOV, l'homologue chez les plantes de DIS3L2, est un membre de la famille de Rrp44/Dis3 conservant le domaine RNase II. Cependant, SOV/DIS3L2 ne contient pas le domaine PIN nécessaire à l'interaction avec le cœur de l'exosome. L'écotype sauvage Col0 d'*A. thaliana* semble être le seul écotype à posséder une version de SOV inactive. En effet, il possède une arginine en position 705 au lieu d'une proline chez les autres écotypes (Zhang et al., 2010a). DIS3L2 et son homologue chez la levure, *S. pombe*, dégradent les ARN dans le sens 3'-5' indépendamment de l'exosome (Lubas et al., 2013; Malecki et al., 2013; Zhang et al., 2010a). De manière intéressante, DIS3L2 dégradent préférentiellement les ARN uridylés (Chang et al., 2013; Lubas et al., 2013; Malecki et al., 2013). En effet, Dis3l2 a une affinité pour les ARN ayant des oligouridines en 3' *in vitro* (Chang et al., 2013; Malecki et al., 2013). La découverte de DIS3L2/SOV met à jour une nouvelle voie de dégradation alternative aux deux principales voies de dégradation des ARN effectuées par l'exoribonucléase 5'-3' XRN1/4 et l'exosome.

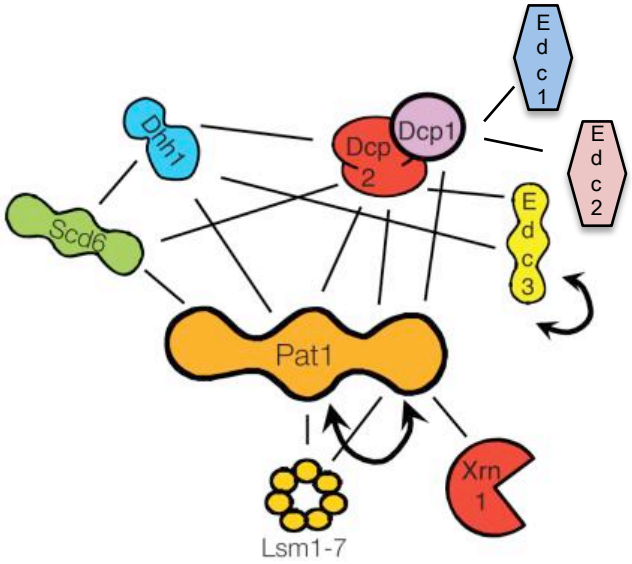


Figure IA.7 : Les facteurs interagissant avec le complexe d'élimination de la coiffe pour la dégradation 5'-3' chez *S. cerevisiae*

Les interactions entre les activateurs de l'élimination de la coiffe (Dhh1, Scd6, Edc1, Edc2, Edc3 et Lsm1-7), la plateforme de recrutement Pat1, le complexe d'élimination de la coiffe Dcp1-Dcp2 et l'exoribonucléase 5'-3' XRN1 sont indiquées. D'après (Borja et al., 2011; Nissan et al., 2010).

3. Les ribonucléases agissant en 5' de l'ARN

a) Le complexe d'élimination de la coiffe

Après déadénylation, lorsque la queue polyA de l'ARNm n'est plus recouverte de PABP, le complexe Lsm1-7 peut se lier à la queue oligoadénylée et recruter le complexe d'élimination de la coiffe *via* Pat1 (Figure IA.7). L'élimination de la coiffe s'effectue par Dcp2 qui est conservée chez les eucaryotes ou par hNudt16 chez l'homme (Dunckley and Parker, 1999; Lu et al., 2011; Song et al., 2010).

Le complexe d'élimination de la coiffe est formé de deux protéines Dcp1 et Dcp2. Dcp2, membre de la famille Nudix/MutT, qui possèdent une activité pyrophosphatase et une zone de liaison à l'ARNm (Steiger et al., 2003). Dcp2 clive la coiffe (7mGpppN) qui protège l'extrémité 5' de l'ARN et libère donc l'ARNm avec une extrémité 5' monophosphate et 7mGDP. Dcp1, de la famille Ervh1/WH1, est la sous-unité régulant Dcp2 par liaison avec des activateurs *via* le domaine EVH1 (Borja et al., 2011). L'élimination de la coiffe de l'ARN en 5' permet l'accessibilité de l'extrémité 5' monophosphate de l'ARN aux exonucléases du sens 5'-3' (Stevens, 2001).

1. Les facteurs de régulation du complexe d'élimination de la coiffe

Les protéines de liaisons à l'ARN Edc1, Edc2, Edc3 et Edc6, pour *Enhancer of Decapping,* Sdc6 et Dhh1 se lient à Dcp2 et l'activent (Borja et al., 2011; Nissan et al., 2010). Chez les mammifères et l'homme, la protéine au domaine WD40, Helds/Ge-1 est un activateur essentiel pour l'élimination de la coiffe (Yu et al., 2005). Dhh1 est l'hélicase de la famille DEAD box qui permet la déstructuration de l'ARN (Nissan et al., 2010). Pat1, quand à elle, sert de plateforme en se liant au complexe Lsm1-7 en 3' de l'ARNm et au complexe d'élimination de la coiffe qui agit en 5' de l'ARNm (Figure IA.7). Edc1 et Edc2 se lient au domaine EVH1 de Dcp1, qui est un domaine de reconnaissance des régions riche en proline, *via* leur région riche en proline située en C terminale (Borja et al., 2011). Sdc6 (reliée à Edc3), réprime eIF4G du complexe de liaison de la coiffe et interagit avec Dcp2, Pat1 ainsi que Dhh1 (Nissan et al., 2010). En effet, afin que le complexe d'élimination de la coiffe puisse accéder à la coiffe, il faut les facteurs d'initiation de la traduction eIF4E-eIF4G n'y soient plus liés. En effet, ceux-ci circularisent l'ARNm en interagissant avec la coiffe et les PABP liées aux queues polyA. Pab1 inhibe l'élimination de la coiffe tout comme le facteur d'initiation eIF4E et certaines autres protéines (Caponigro and Parker, 1995; Schwartz and Parker, 2000). Par exemple, Khd1 se lie à l'ARNm de Mtl1 et au facteur d'initiation de la traduction eIF4E-eIF4G, inhibant l'élimination de la

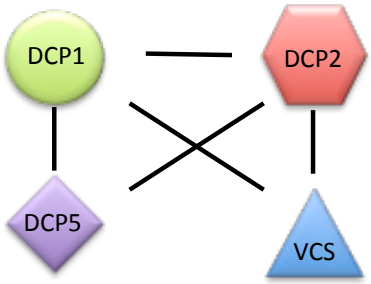


Figure IA.8 : Les facteurs du complexe d'élimination de la coiffe chez *A. thaliana*

Les interactions entre DCP1 et DCP2 sont indiquées, et DCP5, comme VCS interagit avec DCP1 et DCP2. DCP1 peut interagir ave lui même. D'après (Iwasaki et al., 2007; Xu and Chua, 2009; Xu et al., 2006).

coiffe et affectant aussi la localisation de l'ARNm (Mauchi et al., 2010; Paquin et al., 2007).

L'étape de déadénylation est principalement nécessaire à l'élimination de la coiffe. Cependant, elle peut aussi intervenir, pour certains ARNm, sans être précédée d'une étape de déadénylation. Par exemple, les ARNm n'ayant pas de codon stop sont traduits, ce qui élimine les PABP de la queue polyA et permet l'élimination de la coiffe (Garneau et al., 2007). D'autre part, l'élimination de la coiffe pourrait être un processus réversible chez les mammifères, nécessaire au rétablissement des cellules après un stress oxydatif (Otsuka et al., 2009). Ceci suggère que les cellules sont capables de stocker les ARNm après élimination de leur coiffe et de les remobiliser dans la traduction. Ce stockage peut être effectué au niveau des *processing bodies* et/ou des granules de stress (Introduction partie C).

2. Le complexe d'élimination de la coiffe chez *A. thaliana*

Le complexe d'élimination de la coiffe chez *A. thaliana*, contient DCP1-DCP2-VCS/EDC4 et DCP5 (Iwasaki et al., 2007; Xu and Chua, 2009; Xu et al., 2006). Comme chez la levure et l'homme, seule DCP2 possède une activité pyrophosphatase. VCS/EDC4, homologue à Helbs/Ge-1, DCP1 et DCP5 sont des cofacteurs. DCP1 interagit avec lui-même et DCP2, et DCP5, comme VCS interagit avec DCP1 et DCP2 (Figure IA.8) (Iwasaki et al., 2007; Xu and Chua, 2009; Xu et al., 2006). Chacune de ces protéines est essentielle au développement post-embryonnaire, à l'exception de VCS qui est seulement essentielle dans l'écotype Col0 qui possède une version mutée de SOV. En effet, en absence d'un de ces facteurs d'élimination de la coiffe, les plantes meurent au stade cotylédon (Iwasaki et al., 2007; Xu and Chua, 2009; Xu et al., 2006).

L'élimination de la coiffe est régulée par des stimuli environnementaux. Un stress hydrique active la phosphorylation de DCP1 par la kinase MPK6. DCP1 phosphorylée s'associe préférentiellement à DCP5 qui est nécessaire pour l'activité d'élimination de la coiffe *in vivo* (Iwasaki et al., 2007; Xu and Chua, 2009, 2012). Cette voie d'élimination de la coiffe MPK6-DCP1-DCP5 est donc essentielle à l'adaptation d'*A. thaliana* à un stress hydrique.

b) Les exonucléases 5'-3'

Une fois que la coiffe est éliminée en 5' de l'ARN, celui-ci ne possède plus qu'un monophosphate en 5' et il devient un substrat de dégradation pour les exoribonucléases 5'-3' (Stevens, 2001). Il existe plusieurs exoribonucléases 5'-3' qui sont appelées XRN. Elles sont notamment impliquées dans la dégradation des ARNm, le *silencing*, la maturation des ARNr et la terminaison de la transcription

(Nagarajan et al., 2013). Chez les mammifères et la levure, XRN1/PACMAN est localisée dans le cytoplasme et XRN2/RAT1 dans le noyau. XRN2/RAT1 est très similaire à XRN1/PACMAN et leurs sites actifs sont conservés. Chez Arabidopsis, les XRNs nucléaires sont XRN2 et XRN3. XRN4 est aussi un homologue de XRN2/RAT1 mais elle est localisée dans le cytoplasme et assure la même fonction que XRN1/PACMAN (Chiba and Green, 2009).

1. XRN1/PACMAN

XRN1/PACMAN est l'exonucléase 5'-3' du cytoplasme conservée de la levure à l'homme (Larimer et al., 1992). Elle dégrade les ARN incluant les ARNm, les ARN non codants, les microARN et les ARNm ayant un codon stop précoce, substrats du NMD (Nagarajan et al., 2013). XRN1/PACMAN est essentiel au développement des organismes. Par exemple, chez l'homme, XRN1/PACMAN est essentielle pour la formation des os tandis que chez les mouches, elle impacte la fertilité, la formation des ailes et de la pupe (Grima et al., 2008; Laurence et al., 2012; Lin et al., 2008; Newbury and Woollard, 2004; Zabolotskaya et al., 2008). XRN1/PACMAN est aussi impliquée dans la résistance au stress salin et le développement (Nagarajan et al., 2013; Todeschini et al., 2006). Chez la levure, dans un mutant Δhal2, lors d'un stress salin, des 3'-phosphoadenosine 5'-phosphate (pAp) s'accumulent, inhibant les exoribonucléases XRN1/PACMAN et XRN2/RAT1 (Todeschini et al., 2006). XRN2/RAT1 est impliquée dans la biogenèse des ARN structuraux nucléaires mais peut être substituée par XRN1/PACMAN. En effet, dans un mutant XRN2/RAT1, une partie de XRN1/PACMAN est relocalisée dans le noyau (Nagarajan et al., 2013). Il a été montré que XRN1/PACMAN et XRN2/RAT1 sont interchangeables. En effet, XRN1/PACMAN peut compenser la perte de XRN2/RAT1 et XRN2/RAT1 peut substituer XRN1/PACMAN lorsqu'elle est localisée dans le cytoplasme (Johnson, 1997).

• Les cofacteurs de XRN1/PACMAN

XRN1/PACMAN s'associe avec DCP1-DCP2, les activateurs du complexe de l'élimination de la coiffe ainsi que le complexe Lsm1-7 *via* son interaction avec l'activateur de l'élimination de la coiffe, Pat1 (Figure IA.7) (Bouveret et al., 2000; Nissan et al., 2010; Tharun et al., 2000). D'autre part, XRN1/PACMAN s'associe aussi avec le complexe UPF (UPF1-UPF2-UPF3a) qui contient l'ARN hélicase UPF1 intervenant dans la voie de dégradation du NMD (Lejeune et al., 2003). Chez la levure, XRN1/PACMAN interagit avec la protéine dégradant la coiffe m7GpppN en m7Gp, Dcs1 (van Dijk et al., 2003; Liu and Kiledjian, 2005; Sinturel et al., 2012).

Tristetraproline (TTP) est co-purifiée avec XRN1/PACMAN, DCP2 et l'exosome, permettant ainsi la dégradation des transcrits ayant des séquences ARE (Lykke-Andersen and Wagner, 2005; Stoecklin et al., 2006).

- L'exoribonucléase 5'-3' du cytoplasme chez *A. thaliana*

Comme expliqué précédemment, chez *A. thaliana,* l'exoribonucléase 5'-3' XRN4 est un homologue de XRN2 mais elle effectue la fonction de XRN1/PACMAN au niveau du cytoplasme (Kastenmayer and Green, 2000). En effet, XRN4 est l'exoribonucléase 5'-3' responsable de la dégradation des ARN cytoplasmiques. Elle est impliquée dans la voie de dégradation des ARNm ayant un codon stop précoce, substrats du NMD même si elle n'est pas essentielle au NMD (Mérai et al., 2012). XRN4 est aussi impliquée dans la dégradation des ARNm cibles des long siARN (lsiARN) qui ne requièrent pas d'endoclivage de l'ARNm. Ces lsiARN font entre 30 à 40 nucléotides et sont induits lors de l'infection des plantes par *Pseudomonas syringae* ou d'autres stress biotiques (Katiyar-Agarwal et al., 2007). Le mutant *xrn4*-3 n'a pas de phénotype sévère mais les plantes sont insensibles à l'éthylène qui est une hormone connue pour influencer la réponse au stress (Olmedo et al., 2006; Potuschak et al., 2006). XRN4 est également un suppresseur de la voie du PTGS (*Post-Transcriptional Gene Silencing*) (Gazzani et al., 2004). Par exemple, le mutant *xrn4-3* est très résistant au CaMV car l'absence de XRN4 favorise le *silencing* des ARNm viraux par le PTGS (Gy et al., 2007).

2. La XRN nucléaire XRN2/RAT1

XRN2/RAT1 est l'exonucléase 5'-3' localisée dans le noyau. XRN2/RAT1 est très similaire à XRN1/PACMAN et son mécanisme est semblable à celui de XRN1/PACMAN. En effet, elle agit après élimination de la coiffe et est recrutée par le complexe Lsm2-8. XRN2/RAT1 joue un rôle dans le rognage en 5' des ARNr, des snoARN, des microARN, la dégradation des ARNt matures et lors de la terminaison de la transcription par l'ARN Polymérase I et II (Chernyakov et al., 2008; Fatica and Tollervey, 2002; Proudfoot, 2011; Whipple et al., 2011). XRN2/RAT1 est aussi impliquée dans la maintenance des télomères chez l'homme et la levure. En effet, XRN2/RAT1 promeut l'élongation des télomères en dégradant les *Telomeric Repeat-containing RNA* (TERRA) (Luke and Lingner, 2009; Luke et al., 2008). Tout comme XRN1/PACMAN, XRN2/RAT1 est essentielle à la viabilité des organismes, de la levure à la mouche (Chiba and Green, 2009; Nagarajan et al., 2013).

- Les cofacteurs des XRN nucléaires

XRN2/RAT1 est impliquée dans la surveillance de la coiffe dans le noyau lors de stress environnementaux et de stress hydrique (Estavillo et al., 2011; Hirsch et al., 2011; Jiao et al., 2010). XRN2/RAT1 forme un complexe stable avec RAI1 afin de l'activer. RAI1 a une activité pyrophosphohydrolase qui permet d'éliminer la coiffe des ARN possédant une coiffe aberrante. Cependant RAI1 n'est pas conservée chez les eucaryotes (Nagarajan et al., 2013). XRN2/RAT1 s'associe aussi avec DCP1a/DCP2 et le facteur de terminaison TTF2 qui se localise près des sites de début de transcription, coïncidant avec les ARN Polymérase II en pause. Edc3, Dcp1-Dcp2 et XRN2/RAT1 interagissent donc en complexe pour la terminaison de la transcription des transcrits de l'ARN polymérase II dans le noyau (Brannan et al., 2012).

- Les XRN nucléaires d'*A. thaliana* XRN2 et XRN3

A. thaliana possède deux exoribonucléases 5'-3' nucléaires : XRN2 et XRN3, respectivement situées dans le nucléoplasme et le nucléole. XRN3 est essentielle pour le développement des plantes car en absence de XRN3, les plantes ne peuvent pas dépasser le stade d'embryon (Gy et al., 2007). XRN2 et XRN3 dégradent les sous produits de la maturation des ARNr 45S. XRN2 dégrade les ARNr polyadénylés et les pré-ARNr mais elle n'est pas essentielle pour la dégradation des ARNm (Lange et al., 2011; Zakrzewska-Placzek et al., 2010). De plus, elle est essentielle à la première étape de la maturation des pré-ARNr 45S (Lange et al., 2011; Zakrzewska-Placzek et al., 2010). Celle-ci consiste en une coupure endonucléolytique des pré-ARNr au site P par le complexe U3 snoRNP (Sáez-Vasquez et al., 2004; Samaha et al., 2010). XRN2 est aussi impliquée dans la dégradation des petits ARN chez *A. thaliana* et l'algue verte (Cerutti and Ibrahim, 2011). D'autre part, XRN2 et XRN3 influencent le PTGS comme XRN1/XRN4, due à une déficience de la dégradation des ARN issus de transgènes à l'origine de la formation de siARN (Gy et al., 2007; Kastenmayer and Green, 2000). Chez *A. thaliana,* l'absence de l'un des suppresseurs du PTGS, la nucléotide/phosphatase FIERY1 (FRY1), qui est l'une des six orthologues de Hal2 chez la levure, restaure le PTGS dans un mutant *ago1* hypomorphique. En effet, FRY1 réprime XRN2, XRN3 et XRN4, induisant le *silencing* des ARN (Gy et al., 2007). XRN2 et XRN3 régulent négativement des gènes de résistance au stress hydrique. En effet, la délétion des XRN nucléaires rend les plantes plus résistantes au stress hydrique (Estavillo et al., 2011; Hirsch et al., 2011).

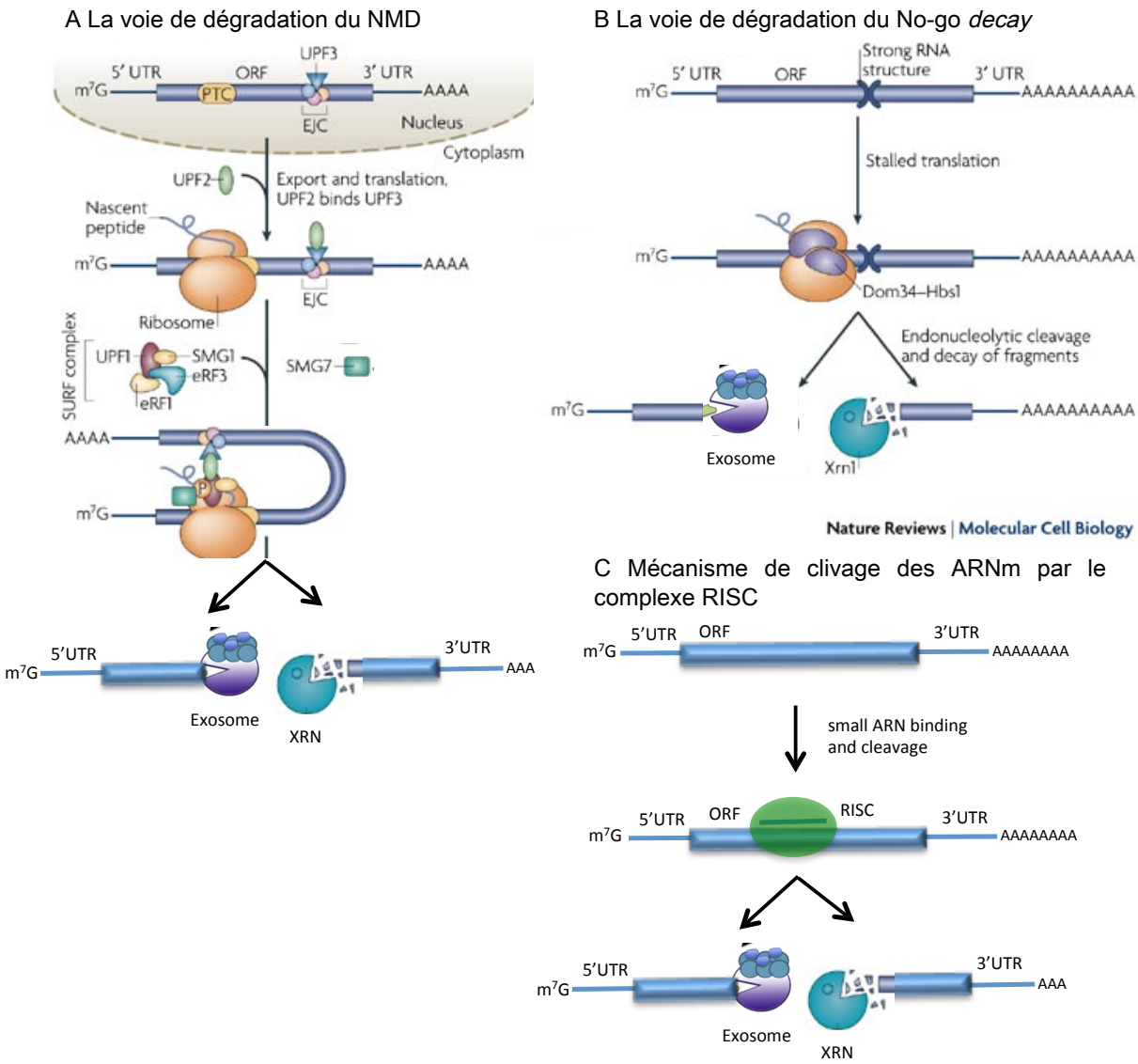


Figure IA.9: Les mécanismes de clivage des ARNm par les endonucléases

A La voie de dégradation du NMD (*non-sense-mediated decay*). Le complexe de jonction des exons (EJC) contenant UPF3 associé avec le transcrit et l'ARNm sont exportés dans le cytoplasme, où se lie UPF2. Les ribosomes associés aux ARNm traduits sont bloqués au niveau du codon stop précoce (PTC). La liaison du complexe SURF aux ribosomes (SMG1, UPF1, eRF1 et eRF3). UPF1 se lie à UPF2 qui se lie à EJC par le PTC. La phosphorylation de UPF1 par SMG1 induit la dissociation de eRF1 et eRF3 et la liaison de SMG7. **B** La voie de dégradation NGD (*No-Go decay*). Les ribosomes sont bloqués dans l'*open reading frame* (ORF) par une structure secondaire forte ou un transcrit tronqué. Les protéines Dom34 et Hbs1 se lient aux ARNm près du ribosome bloqué et initie le clivage endonucléolytique près du site de blocage, permettant le relargage du ribosome et générant deux fragments d'ARNm. **C** Lors qu'un petit ARN est complémentaire à l'ARNm cible, il recrute le complexe RISC et l'ARNm est clivé par AGO. Les fragments d'ARNm issus du clivage endonucléolytiques sont ensuite dégradés par l'exosome ou l'exoribonucléase 5'-3'. D'après (Garneau et al., 2007; Pfeffer et al., 2004; Voinnet, 2009).

4. Les endonucléases

La dégradation des ARN par une endonucléase peut être effectué après un contrôle qualité des ARNm ou suite à la reconnaissance par des petits ARN complémentaires. En effet, lorsque les petits ARN sont parfaitement complémentaires à leur ARN cibles, ils induisent un clivage par une protéine AGO (Introduction partie B4a) (Figure IA.9).

Il existe deux voies de dégradation endonucléolytique des ARNm aberrants après un contrôle qualité, la voie du NMD (*non-sense-mediated decay*) et la voie du NGD (*No-Go decay*). La voie du NMD dégrade les ARNm aberrants ayant un codon stop précoce. La voie du NGD élimine les ARNm qui ne permettent pas aux ribosomes de réaliser correctement l'élongation de la traduction dû à une structure secondaire forte dans la séquence des ARNm ou lorsque les ARNm sont tronqués. Lors de la voie de dégradation du NMD, l'endoclivage est réalisé par l'endonucléase SMG6 (Figure IA.9A). Dans la voie de dégradation NGD (*No-Go decay*), les protéines Dom34 et Hbs1 se lient aux ARNm près du ribosome bloqué et initie le clivage endonucléolytique près du site de blocage (Figure IA.9B) (Garneau et al., 2007). Dans le cas de la dégradation des ARNm après leur reconnaissance par de petits ARN parfaitement complémentaires, l'endoclivage des ARNm s'effectue par AGO (Figure IA.9C) (Voinnet, 2009). Ce clivage peut être effectué après reconnaissance par des microARN (Introduction partie B3c et B4a) ou des petits ARN interférents, issus de transcrits viraux (Voinnet, 2008).

Les fragments d'ARNm clivés par des endonucléases sont ensuite dégradés par l'exonucléases 5'-3' XRN1/XRN4 pour le fragment 3' et dans le sens 3'-5' par l'exosome, principalement, pour le fragment 5'. Cependant, les fragments 5' des ARNm clivés par le complexe RISC, chez les mammifères et les plantes, peuvent également être uridylés par des ncPAP avant d'être dégradés dans le sens 5'-3' (Introduction partie B3b).

5. Les protéines modifiant et/ou marquant les ARN précédant leur dégradation

Afin de faciliter la dégradation des ARN, ceux-ci peuvent être modifiés et/ou marqués en 3' par différentes protéines. Le complexe Lsm1-7 marque les ARN pour la dégradation dans le sens 5'-3'. Les protéines de la famille des ncPAP peuvent, quant à elles, polyadényler ou uridyler les ARN en 3' afin de recruter l'exosome ou DIS3L2/SOV pour la dégradation dans le sens 3'-5' ou augmenter l'affinité du complexe Lsm1-7 par un ajout d'uridine, polarisant la dégradation dans le sens 5'-3'.

a) Le complexe Lsm1-7

Le complexe Lsm1-7 (sm *like*) est formé de sept protéines de liaison à l'ARN possédant le domaine Sm avec les motifs SmI et SmII. Elles s'assemblent en formant un anneau (Hermann et al., 1995). Le complexe Lsm1-7 est conservé de la levure à l'homme et il existe deux variantes Lsm1-7 et Lsm2-8. Le complexe Lsm 2-8 est localisé au niveau du noyau et est impliqué dans l'épissage des introns avec le petit ARN nucléaire snARN U6 (Spiller et al., 2007). Le complexe Lsm1-7, quant à lui, se situe dans le cytoplasme et est impliqué dans la dégradation des ARN (Golisz et al., 2013). Le complexe Lsm1-7 se lie à l'extrémité 3' de l'ARN *in vivo* grâce à la structure du domaine Sm (Wilusz and Wilusz, 2013). Le complexe Lsm1-7 se lie préférentiellement aux ARNm déadénylés. En effet, le complexe Lsm1-7 est intrinsèquement capable de distinguer les ARN ayant une queue polyA de ceux étant oligoadénylés (Chowdhury et al., 2007). Par ailleurs, l'affinité du complexe Lsm1-7 est augmentée lorsque les ARN oligoadénylés sont uridylés (Song and Kiledjian, 2007).

La liaison du complexe Lsm1-7 à l'extrémité 3' de l'ARN protège des attaques exoribonucléique en 3' et stimule l'élimination de la coiffe par le recrutement du complexe d'élimination de la coiffe dcp1-dcp2 et de ses activateurs *via* Pat1 (Figure IA.7) (Nissan et al., 2010; Tharun, 2009). En effet, Pat1 est la plateforme d'inhibition de la traduction et d'activation de l'élimination de la coiffe. Le complexe Lsm1-7 et Pat1 s'accumulent tout les deux dans les *processing bodies* (Buchan et al., 2008; Franks and Lykke-Andersen, 2008; Reijns et al., 2008). Les *processing bodies* sont des foci cytoplasmiques stockant les ARNm non traduits vraisemblablement destinés à être dégradés (Parker and Sheth, 2007; Sheth and Parker, 2003) (Introduction partie C2b).

b) Les ncPAP

La superfamille des ADN polymérase-β (Pol β) contient la famille des PAP canoniques contenant un domaine nucléotidyltransférase, un domaine PAP/25A associé caractéristique et un domaine RRM. Les PAP canoniques sont responsables de l'ajout d'une queue polyA aux transcrits de l'ARN polymérase II à l'exception des ARNm des histones chez les mammifères. La queue polyA des ARNm est requise pour leur stabilisation, leur export du noyau au cytoplasme et pour l'efficacité de leur traduction (Colgan and Manley, 1997). Les PAP non canoniques appartiennent également à la famille des nucléotidyltransférases et régulent la maturation, la traduction et la dégradation de nombreux ARN dans le noyau, le cytoplasme, les chloroplastes ou les mitochondries (Wang et al., 2000).

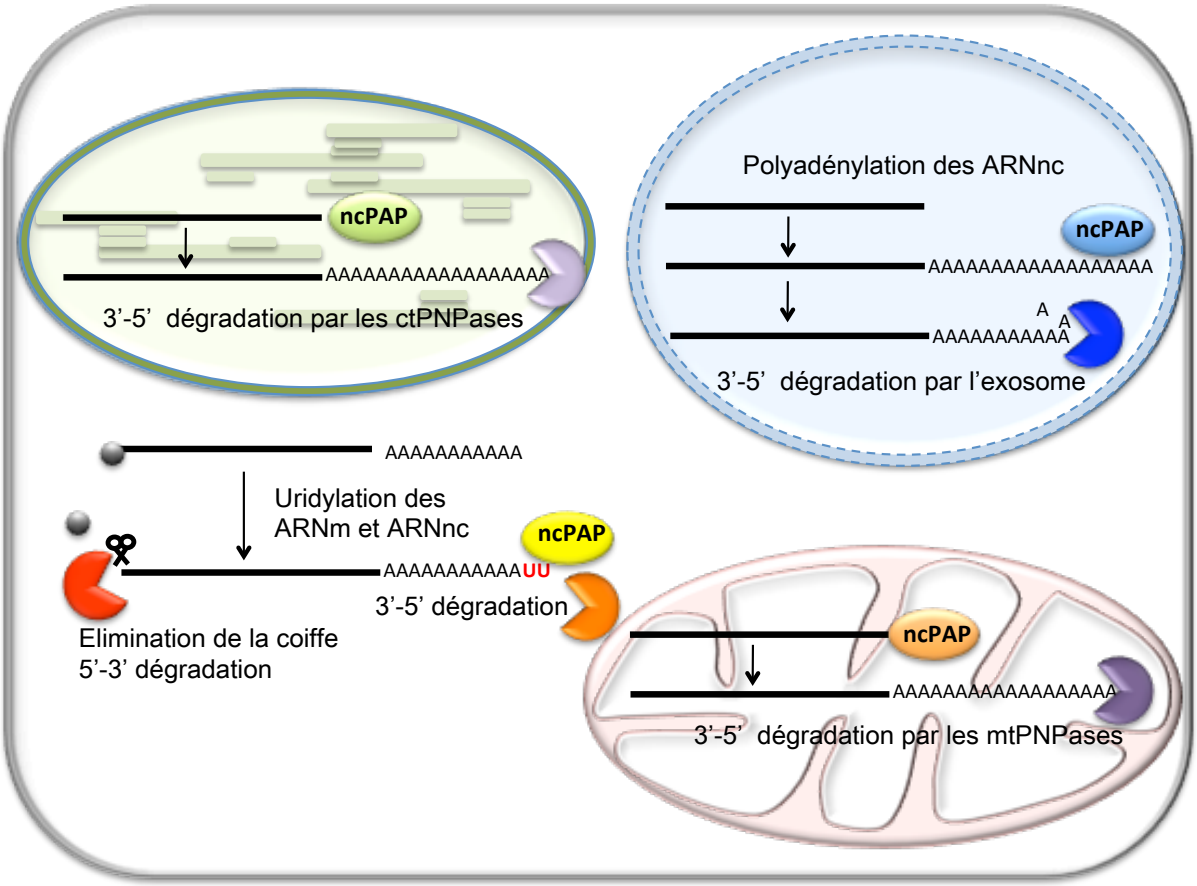


Figure IA.10 : L'implication des ncPAP dans la dégradation des ARN

Chez les plantes, les ARN peuvent être polyadénylés par les ncPAP avant dégradation dans le noyau, les mitochondries et les chloroplastes. Les ARNm et les ARN non codant peuvent être uridylés dans le cytoplasme avant leur dégradation (Chekanova et al., 2000; Heo et al., 2012; Lange et al., 2009; Malecki et al., 2013; Mullen and Marzluff, 2008; Ren et al., 2012; Rissland and Norbury, 2008; Zhao et al., 2012).

Tableau IA.1 : Les ncPAP chez *S. cerevisiae*, *S. pombe* et les mammifères

	S. cerevisiae	*S. pombe*	mammifères	activité	localisation
Trf4 et Trf5-like	Trf4 Trf5	Cid12 Cid14	PAPD5 POLS	oligaoadénylation	noyau
PAP	pas d'équivalent	inconnue	PAPD1	oligaoadénylation	mito-chondrie
GLD2-*like*	pas d'équivalent	Cid11 Cid13	GLD-2	polyadénylation, monoadénylation	noyau
U6TUTase	pas d'équivalent	pas d'équivalent	U6 TUTase	oligouridylation, polyadénylation	noyau
Cid1-*like*	pas d'équivalent	Cid1	ZCCHC6 ZCCHC11 (CCHC-type zinc-finger, Inactive nucleotidyl transferase domain)	monouridylation polyuridylation	cytoplasme
autre	pas d'équivalent	Cid16	pas d'équivalent	inconnue	inconnue

D'après (Halic and Moazed, 2010; Motamedi et al., 2004; Norbury, 2013; Win et al., 2006).

Les ncPAP sont conservées de la levure à l'homme. Cependant, le nombre de ncPAP présent chez les organismes reste très variable. En effet, il existe seulement deux ncPAP chez *S. cerevisiae*, six chez *S. pombe*, sept chez l'homme et quatorze chez *Arabidopsis thaliana* (Martin and Keller, 2007; Lange Et al., 2009). Les six ncPAPs de *S. pombe* sont Cid1, Cid11, Cid12, Cid13, Cid14 et Cid16. Elles forment les ncPAP de la famille de Cid1 (Stevenson and Norbury, 2006). Les ncPAP conservent toutes les mêmes caractéristiques. Elles incorporent un ribonucléoside monophosphate à l'extrémité 3' d'un ARN. Elles sont composées de deux domaines : un domaine nucléotidyltransférase et un domaine PAP/25A-associé, contenant le domaine de reconnaissance des nucléotides. La structure générale des ncPAP et leur site catalytique est conservé. La structure des ncPAPs contient un motif hélice-tour de la famille Pol β contenant le site catalytique DxD ou DxE (Martin and Keller, 2007). Contrairement aux PAP canoniques, les ncPAP ne possèdent principalement pas de motifs de liaison à l'ARN, à l'exception de RBM21/Star PAP/l'U6 TUTase et PAPD5 qui contient une région de liaison à l'ARN riche en acides aminés basiques (Rammelt et al., 2011). A l'inverse des PAP canoniques, les ncPAP semblent agir majoritairement en complexe avec des protéines de liaison à l'ARN.

L'activité des ncPAP peut être différente *in vivo* et *in vitro*. Par exemple, Cid1 *in vitro* est processive même si elle ajoute seulement une à trois uridines *in vivo* (Rissland et al., 2007). La spécificité nucléotidique des ncPAP peut aussi être apportée par un cofacteur. En effet, RBM21/Star PAP, l'U6 TUTase possède une spécificité différente PAP ou PolyU Polymérase (PUP), respectivement selon son interaction ou non avec le cofacteur CPSF (*Cleavage and Polyadenylation Specific Factor*) (Kaufmann et al., 2004; Laishram and Anderson, 2010; Li et al., 2013). Cependant, les ncPAP peuvent aussi avoir une spécificité stricte *in vivo* et accepter plus ou moins d'autres ribonucléotides *in vitro*. Afin de comprendre la spécificité des ncPAP, des données de structures de la famille des ncPAP au niveau atomique ont été obtenues et certains acides aminés responsables de la spécificité nucléotidiques pour les uridines ont été mise en évidence (Chapitre2) (Deng et al., 2005; Lunde et al., 2012; Munoz-Tello et al., 2012; Stagno et al., 2007, 2010; Yates et al., 2012).

Les ncPAP possèdent une spécificité nucléotidique différente et leur localisation varie. Elles peuvent être localisée dans le cytoplasme et les différents organites (le noyau, les chloroplastes et les mitochondries) (Figure IA.10 et Tableau IA.1).

1. Les ncPAP responsable de la polyadénylation

Les deux ncPAP de la levure sont Trf4 et Trf5 qui sont des polyA polymérases déstabilisatrices nucléaires. Elles peuvent toutes deux faire partie du complexe TRAMP qui active la dégradation des ARN par l'exosome (Introduction partie A1b). Il existe une certaine redondance de fonction entre Trf4 et Trf5 mais elles ciblent aussi différents substrats ARN. Trf4 est indispensable *in vivo* et joue un rôle dans le maintien de la longueur des télomères (San Paolo et al., 2009). Trf4 et Trf5 ont des orthologues chez *S pombe* (Cid14 et Cid12) et chez l'homme (hTRF4-2/PAPD5). Chez l'homme, un équivalent au complexe TRAMP a été identifié dans le nucléole lors de l'immunoprécipitation de l'hélicase hMTR4 (Lubas et al., 2011). Ce complexe contient donc hMTR4 et la protéine de liaison à l'ARN ZCCHC7 et la ncPAP hTRF4-2/PAPD5. Chez *S. pombe*, Cid14 fait partie du complexe TRAMP-*like* et induit la dégradation des ARN ayant une demi-vie courte ou les ARN aberrants en interagissant avec l'exosome nucléaire. Cid14 s'associe avec Air1 et forme donc le complexe TRAMP-*like* avec Mtr4-Cid14 et Air1. Le complexe TRAMP-*like* est impliqué dans la maturation des ARNr, l'assemblement des pré-ribosomes ainsi que la formation de l'hétérochromatine au niveau des télomères grâce à la méthylation des histones (Keller et al., 2010; Win et al., 2006a). *S. pombe* possède une autre ncPAP de la famille de TRF4/5 : Cid12 qui est aussi impliquée dans la formation de l'hétérochromatine. Elle forme le *RNA directed RNA Polymerase Complex* en dirigeant la méthylation des histones grâce au niveau des centromères (Halic and Moazed, 2010; Motamedi et al., 2004; Win et al., 2006b).

Cid11 et Cid13, sont deux polyA polymérases localisées au niveau du noyau. Cid11 joue un rôle dans la formation de l'hétérochromatine, la maturation des siARN, le *silencing* et la stabilité des ARNm (Win et al., 2006b). Cid13 est, quant à elle, une polyA polymérase ciblant d'une façon spécifique l'ARNm *suc22* qui code une sous unité de la ribonucléotide réductase. Cid13 ajoute des adénines aux ARN en réponse à l'hydroxyurée au niveau du cytoplasme et du noyau (Read, 2002; Saitoh et al., 2002). GLD2/PAPD4 est responsable de la polyadénylation cytoplasmique d'ARNm, permettant ainsi une activation de la traduction (Villalba et al., 2011). Cid11 et Cid13 sont toutes les deux des ncPAP GLD2-*like* (Kwak et al., 2004).

Il existe d'autres ncPAP étant des polyA polymérases, localisées dans les compartiments cellulaires tels que les mitochondries ou les chloroplastes. Comme, par exemple, la PAP mitochondriale humaine PAPD1 dont le rôle dans le métabolisme des ARN mitochondriaux reste à être complètement défini (Chang and Tong, 2012).

2. Les autres ncPAP

S. pombe possède deux autres ncPAP, Cid1 et Cid16, que celle précédemment présentées et possédant une activité uridylyltransférase. Cid1 est la ncPAP de *S. pombe* la mieux caractérisée. Elle est localisée dans le cytoplasme, son activité catalytique a été étudiée en détails *in vitro* et *in vivo* et sa structure est résolue au niveau atomique (Lunde et al., 2012; Munoz-Tello et al., 2012; Rissland and Norbury, 2009; Rissland et al., 2007; Wang et al., 2000; Yates et al., 2012). Cid1 est impliquée dans l'uridylation des ARNm et induit leur dégradation dans le sens 5'-3' (Introduction partie B6) (Rissland and Norbury, 2009). La résolution de la structure de Cid1 au niveau atomique a mis en évidence les acides aminés en partie responsable de sa spécificité pour les uridines (D330, R339 et H336) (Lunde et al., 2012; Munoz-Tello et al., 2012; Yates et al., 2012). Cid1 est conservée chez l'homme (ZCCHC6/TUT7 et ZCCHC11/TUT4), *Xenopus leavis*, *C. elegans* et *A. thaliana* (AT2G45620) (Kwak and Wickens, 2007; Lapointe and Wickens, 2013). Le rôle et le mécanisme d'action de ces uridylyltransférases homologues à Cid1 sont décrits plus précisément dans la partie B de l'introduction.

Les caractérisations de Cid16 de *S. pombe* ne sont pas quant à elles, pas encore publiées, mais son homologie avec Cid1 suggère qu'elle est également une PolyU Polymérase, ce qui a été confirmé expérimentalement (communication personnelle du Dr Norbury).

3. Les ncPAP chez *Arabidopsis thaliana*

Le génome d'*A. thaliana* code quatorze ncPAP (Lange et al., 2009; Ren et al., 2012; Zhao et al., 2012a). La localisation intracellulaire de ces ncPAP ont été, pour la plupart, précédemment déterminées dans l'équipe (Thèse de François Sement 2012). Les plantes possèdent un homologue de la CCA nucléotidyltransférase, codée par le gène At1g22660 et une polyA polymérase, codée par le gène At1g28090 (Zimmer et al., 2009). Pour l'instant, seulement deux autres ncPAP ont été étudiées et elles sont toutes deux des uridylyltransférases. HESO1, codée par le gène At2g39740 est principalement responsable de l'uridylation des microARN (Zhao et al., 2012a, 2012b). La seconde ncPAP, codée par le gène At2g45620 est une uridylyltransférase homologue de Cid1 (Kwak and Wickens, 2007) que nous avons appelée URT1 (*UTP :RNA uridylyltransferase 1*). Elle est étudiée dans ce manuscrit.

Les ncPAP de la famille de Cid1 de *S. pombe* sont principalement impliquées dans la déstabilisation des ARN par l'ajout d'uridines comme Cid1 (Introduction partie B6) ou d'une queue polyA comme Cid11 et Cid13 (homologues à GLD2 des

mammifères) et Cid12 et Cid14 (homologues à Trf4 et Trf5 de *S. cerevisiae*). Il existe aussi des protéines de la famille des ncPAP qui ont perdu leur activité comme NF45 et NF90 qui régulent de l'expression génétique en agissant au niveau de la transcription et de la traduction des ARNm (Wolkowicz and Cook, 2012).

Certaines protéines de la famille des ncPAP sont impliquées dans l'ajout de bases pyrimidines CUCU chez *A. nidulans,* ayant un rôle dans la dégradation des ARN (Morozov et al., 2010, 2012a). A l'inverse, la ncPAP Star-PAP/U6 TUTase chez l'homme est impliquée dans la stabilisation et la maturation du petit ARN nucléaire snARN U6 par l'ajout d'uridines en 3' (Trippe et al., 1998). D'autres ncPAP peuvent être impliquées dans la maturation des ARN et leur dégradation (Berndt et al., 2012; Chang et al., 2013; Heo et al., 2012; Jones et al., 2009; Lubas et al., 2013). Par exemple, la ATP(CTP):ARNt nucléotidyltransférase ou CCA nucléotidyltransférase est impliquée dans la maturation ou la réparation des extrémités 3' des ARNt par un ajout de CCA, générant un site d'aminoacylation. Cependant, elle est aussi responsables de l'ajout de motifs CCACCA, qui induit la dégradation des ARNt *in vivo* (Wilusz et al., 2011).

Les ncPAP sont donc principalement impliquées dans les voies de dégradation soit par l'ajout d'une queue polyA soit par l'ajout de quelques résidus pyrimidines (C et/ou U). L'uridylation des ARN est un processus conservé de *S. pombe* à l'homme qui cible différents ARN et qui peut avoir différents rôles.

B—Les rôles de l'uridylation dans le métabolisme des ARN

L'uridylation émerge comme un processus important du métabolisme des ARN (Scott and Norbury, 2013; Yates et al., 2013). Dans cette seconde partie de l'introduction, je présente les ARN connus pour être uridylés tels que le petit ARN nucléaire U6 (snARN U6), les précurseurs des microARN, les petits ARN (piARN, microARN, siARN), les fragments des ARNm clivés par le complexe RISC (*RNA induced silencing complex*), les ARN guides (ARNg) impliqués dans l'édition des ARNm chez les trypanosomes et les ARNm. Pour chacun des cas, je présente les ncPAP responsables de l'uridylation, si elles sont connues, et le rôle de l'uridylation de ces ARN. En effet, l'uridylation des ARN peut être responsable de leur dégradation dans le sens 5'-3' ou 3'-5' ou avoir un rôle de stabilisation et de promotion de la maturation des ARN.

1. Le petit ARN nucléaire U6 snARN

Chez les eucaryotes, les introns des précurseurs des ARNm sont épissés par le spliceosome. Le spliceosome est un complexe composé de petits ARN nucléaires snARN (U2, U5 et le duplexe U4-U6) et de protéines associées (Grabowski and Sharp, 1986). Le snARN U6 est le snARN le plus conservé de la levure à l'homme, il n'est pas transcrit par la polymérase II comme la plupart des snARN mais par la polymérase III (Moenne et al., 1990). Ce transcrit non codant de 100 bases participe à la formation du site actif du spliceosome par la formation d'un duplexe avec le snARN U4. L'extrémité 3' du snARN U6 est composée de quatre uridines codées par le génome signalant la fin de la transcription (Paule and White, 2000). Les quatre uridines encodées par le génome stabilisent le précurseur du snARN U6 *via* la liaison avec la protéine La à l'extrémité 5'. La protéine La empêche l'accès de l'extrémité 5' à des exoribonucléases 5'-3' (Xue et al., 2000).

Le snARN U6 est particulier car il subit des modifications en 3' par l'ajout ou l'élimination d'uridines. En effet, le snARN U6 est le premier ARN découvert chez l'homme capable d'être uridylé (Booth and Pugh, 1997). Lors de la maturation du snARN U6, la queue polyU peut être allongée jusqu'à 20 uridines par la Star-PAP/U6 TUTase (Trippe et al., 2006) ou réduite jusqu'à cinq uridines par la 3'-5' exoribonucléase Mpn1. Mpn1 est aussi responsable de la cyclisation 2'-3' phosphate de la dernière uridine *in vitro* (Shchepachev et al., 2012). Celle-ci permet la dissociation de la protéine La et le recrutement du complexe nucléaire Lsm2-8 à l'extrémité 3' du snARN U6. Le complexe Lsm2-8 permet la rétention nucléaire du snARN U6, sa stabilisation et la formation du duplexe snARN U4-U6 (Spiller et al., 2007). En absence de Mpn1, l'extrémité 3' du snARN U6 n'est pas maturée après transcription et sa demi-vie est plus faible (Shchepachev et al., 2012). L'uridylation du snARN U6 est donc importante pour sa maturation et sa stabilité. Chez *S. cerevisiae* et la drosophile, les snARN U6 n'ayant pas de cyclisation 2'-3' phosphate sont polyadénylés par Trf4 et ils sont ensuite dégradés par la sous-unité Rrp6 de l'exosome. Il semble que ce mécanisme soit conservé chez *S. pombe* et l'homme, même si en absence de Mpn1 les snARN polyadénylés ne sont pas accumulés (Schneider et al., 2012; Shchepachev and Azzalin, 2013).

2. Le précurseur du microARN *let-7* (pré-*let-7*)

Chez les animaux, les pri-microARN sont tout d'abord clivés par Drosha pour générer un pré-microARN. Les pré-microARN contiennent une structure tige-boucle qui est reconnue par le RISC-*loading complex*. L'endoribonucléase DICER de la

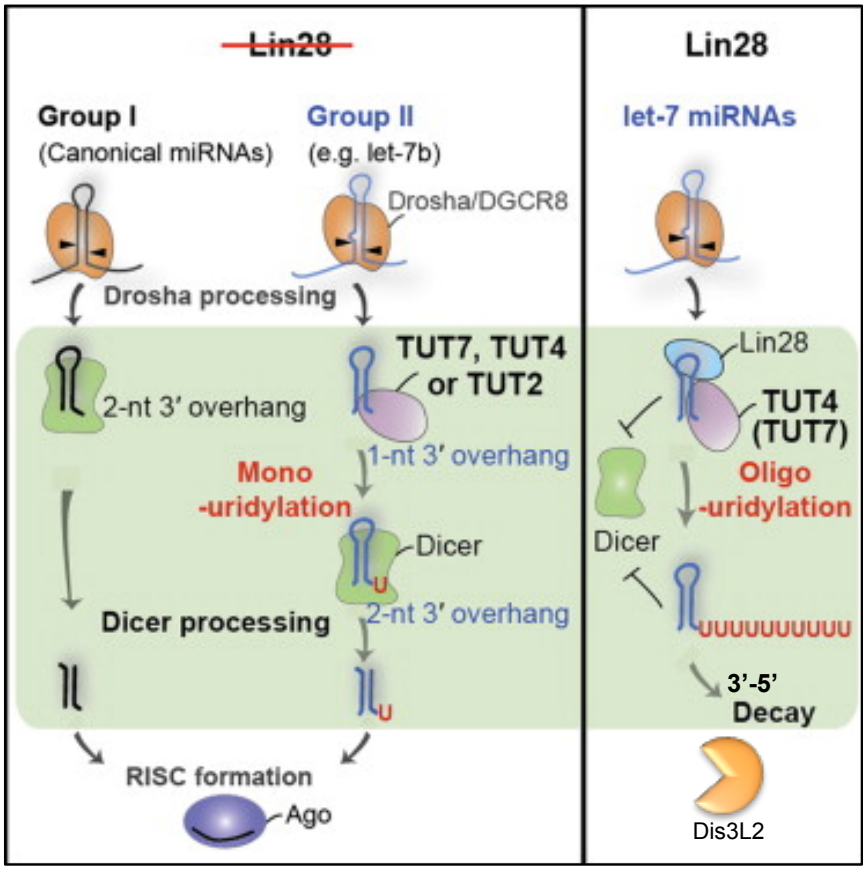


Figure IB.1 : L'uridylation des précurseurs des microARN *let-7* a deux rôles antagonistes selon la quantité d'uridines ajoutées en 3'

Les précurseurs du microARN *let-7* peuvent être monouridylés en 3' par ZCCHC6/TUT7, ZCCHC11/TUT4 ou encore la TUT2/PAPD4/GLD2 (*in vitro*), promouvant leur maturation. A l'inverse, une oligouridylation par les uridylyltransférases interactant avec Lin28, empêche leur maturation et induit leur dégradation dans le sens 3'-5' par l'exoribonucléase Dis3L2. D'après (Chang et al., 2013; Heo et al., 2012).

famille des RNases III, appartenant au RISC-*loading complex* clive l'ARN double brin donnant ainsi naissance aux microARN (Treiber et al., 2012).

Les précurseurs des microARN de la famille de *let-7* peuvent avoir une extremité 3' ayant un nucléotide libre. Les terminales uridylyltransférases ZCCHC6/TUT7, ZCCHC11/TUT4 ou encore la TUT2/PAPD4/GLD2 (*in vitro*), ajoutent une seule uridine au précurseur du microARN *let-7* (pré-*let-7*) (Figure IB.1). En effet, ZCCHC11/TUT4 seule interagit très rapidement avec les pré-*let-7* et n'ajoute qu'une seule uridine en 3'. Cette monouridylation créé une extrémité 3' de deux nucléotides libres, promouvant la maturation de *let-7* après sa reconnaissance par le complexe RISC-*loading complex* (Heo et al., 2012). A l'inverse, en présence de Lin28 qui est une protéine de liaison à l'ARN spécifique de la famille des microARN *let-7*, l'interaction des uridylyltransférases avec le pré-*let-7* est stabilisée. Chez les mammifères, ZCCHC11/TUT4 ou ZCCHC6/TUT7 (Hagan et al., 2009; Thornton et al., 2012) ou PUP-2, son orthologue chez *C. elegans* (Lehrbach et al., 2009) ajoutent donc plusieurs uridines successives. L'oligouridylation en 3' du pré-*let-7* empêche leur reconnaissance par le complexe RISC-*loading complex*, inhibant ainsi leur maturation et à l'inverse, induisant leur dégradation (Heo et al., 2012; Jones et al., 2009, 2012). La dégradation des pré-*let-7* oligouridylés est effectuée, chez la souris, par l'exonucléase 3'-5' Dis3l2. En effet, en absence de Dis3l2, les pré-*let-7* sont stablisés et Dis3l2 dégrade préférentiellement les ARN uridylés *in vitro* (Chang et al., 2013; Malecki et al., 2013) (Figure IB.1).

L'étude de l'uridylation des précurseurs des microARN *let-7* révèle deux rôles antagonistes. En effet, l'ajout d'une uridine en 3' des pré-*let-7* du groupe II promeut leur maturation et, à l'inverse, leur oligouridylation, en présence de Lin28, empêche leur maturation et induit leur dégradation dans le sens 3'-5' par Dis3l2 (Figure IB.1).

3. Les petits ARN (microARN, siARN et piARN)

a) La biogenèse des petits ARN (microARN, siARN et piARN)

Les petits ARN jouent un rôle important dans la régulation de l'expression génétique. Chez les mammifères, les petits ARN, de 21 à 32 nucléotides, peuvent être classés en trois grands groupes: les microARN, les siARN (*small interfering RNA*) et les piARN (*Piwi-interacting RNA*). Ce dernier groupe est spécifique des animaux et n'est donc pas présent chez les plantes. Tous les petits ARN peuvent être uridylés. Je présente, tout d'abord, la voie de biogenèse des piARN chez les animaux et des siARN et microARN chez les plantes et leurs conditions d'uridylation.

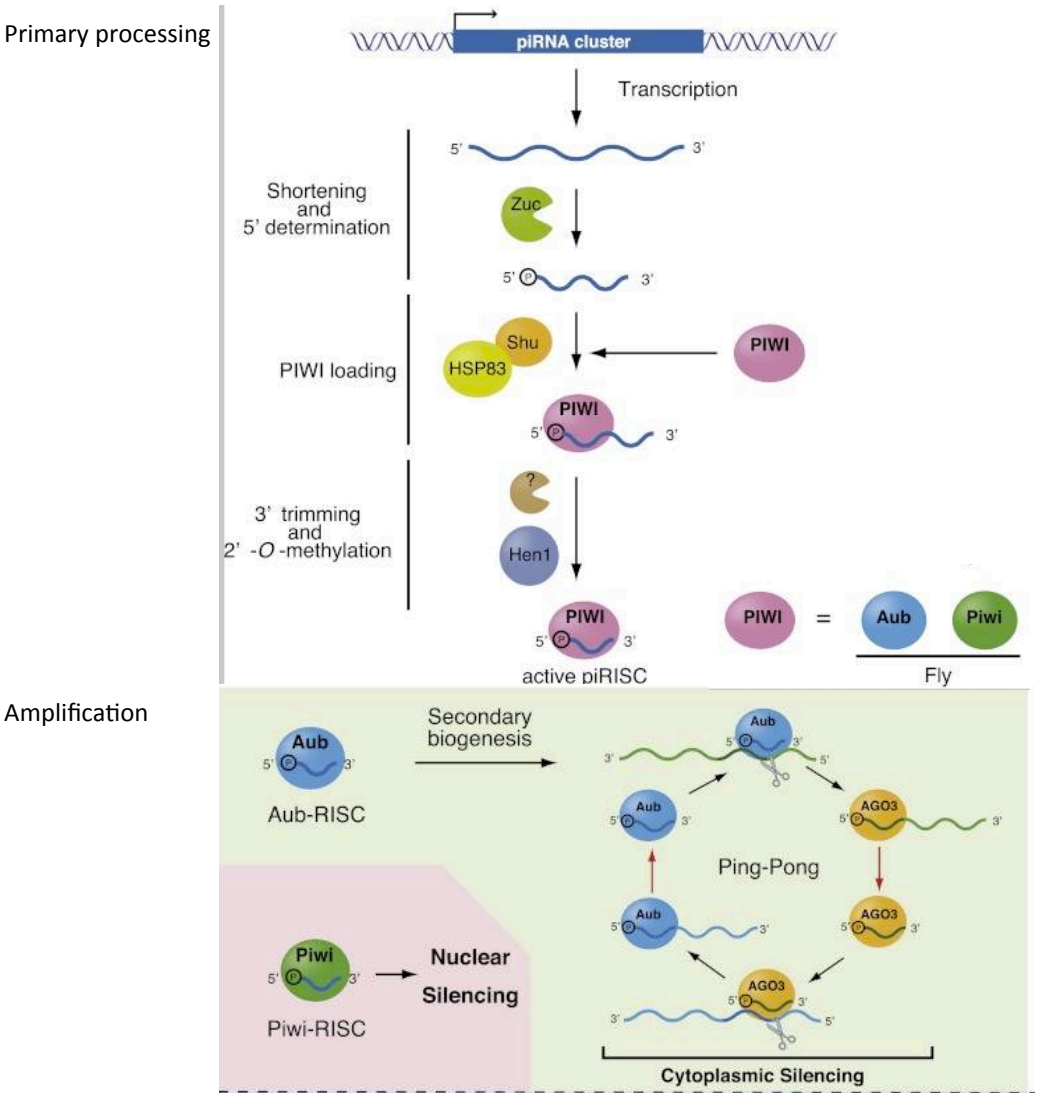


Figure IB.2 : La voie de biogenèse des piARN

La voie de biogenèse des piARN chez la mouche est effectuée par une voie de maturation primaire à partir d'un ARN simple brin indépendant et des enzymes de la famille RNase III, suivit d'un cycle d'amplification. La partie supérieure représente la maturation primaire modèle chez la mouche. Les transcrit primaires des piARN clusters sont raccourcis en petits ARN piARN-*like*. Zucchini pourrait être l'endonucléase formant l'extrémité 5′ des piARN, l'homologue de Zucchini (MitoPLD) semble être la nucléase produisant les transcrits primaires piARN. Les petits ARN piRNA-*like* sont pris en charge par PIWI (en rose) et rognés en 3′ par une nucléase jusqu'à la taille d'un piARN mature. Hsp83/Shu semble être impliqué dans la prise en charge par Piwi. Ensuite, les piARN sont 2′-O-methylés par HEN1/Pimet. Chez la mouche, les PIWI s'associant avec les piARN primaires sont Piwi (vert) et Aub (bleu). Les protéines Piwi associées avec les piARN sont relocalisées dans le noyau et ne semblent pas contribuer à la boucle d'amplification. Aub s'associe avec les piARN et le cycle Ping-Pong est effectuée par clivage des transcrits. Ago3 (orange) prend en charge les piARN secondaires produisant les piARN qui s'associent avec Aub. Les transcrits des transposons sont clivés lors du cycle d'amplification, réprimant les transposons. D'après (Ishizu et al., 2012).

1. Les piARN spécifiques des animaux

Les piARN sont spécifiques des animaux et sont exprimés dans les cellules germinales et l'embryon précoce. Ils répriment l'expression des transcrits en les clivant et en méthylant les loci génomiques leur correspondants (Huntzinger and Izaurralde, 2011; Parent et al., 2012). Les piARN sont assez uniques. En effet, ils font de 24 à 32 nucléotides et sont générés à partir d'éléments intergéniques répétés du génome, enrichies en transposons, appelés piARN clusters (Brennecke et al., 2007). La plupart des piARN sont dérivés de transposons et répriment leur expression (Figure IB.2). La biogenèse des piARN est effectuée par une voie de maturation primaire à partir d'un ARN simple brin indépendamment des enzymes de la famille RNase III, suivie par un cycle d'amplification (Figure IB.2) (Ishizu et al., 2012).

2. Les siARN et les microARN chez les plantes

Chez les plantes, les microARN et les siARN sont impliqués dans la répression génétique et la résistance aux pathogènes. Les microARN sont impliqués dans la dégradation des ARNm lorsqu'ils sont parfaitement complémentaires à leur ARNm cible, en majorité chez les plantes, ou dans la répression de la traduction (Voinnet, 2009). Ce dernier cas de figure est majoritaire chez les animaux, cependant, les microARN parfaitement complémentaires à leurs ARN cibles peuvent être aussi présents chez les mammifères, ou encore être encodés par des virus (Pfeffer et al., 2004; Voinnet, 2009; Yekta et al., 2004). Les siARN, eux, ciblent les ARN viraux et les ARN aberrants. La voie de biogenèse des petits ARN (microARN et siARN) présente des similitudes. Ils ont une taille de 21 à 24 nucléotides et dérivent d'une région double brin d'un ARN précurseur qui est clivé en un ou plusieurs sARN (*small* ARN) duplexe par l'une des 4 enzymes *dicer-like* chez les plantes (Baulcombe, 2004). La plupart des microARN sont issus d'un simple duplexe clivé à partir d'une structure tige-boucle par DCL1 (Figure IB.3A) (Voinnet, 2009). Les siARN sont produits par DCL2, DCL3 et/ou DCL4. Les précurseurs des siARN sont des ARN doubles brins issus d'une transcription inverse ou d'éléments répétés-inversés, de la réplication virale ou de la transcription par la *RNA dependant RNA polymerase* (RDR). Les ARN doubles brins sont ensuite clivés de façon séquentielle en siARN de 21, 22 et 24 nucléotides respectivement par DCL4, DCL2 et DCL3 (Figure IB.3B) (Borsani et al., 2005; Kasschau et al., 2007; Khraiwesh et al., 2012).

b) La modification en 3' des petits ARN

Les microARN et les siARN chez les plantes et les piARN des animaux sont ensuite méthylés en 3' par une enzyme très conservée : HEN1 (*Hua Enhancer 1*),

découverte chez *Arabidopsis thaliana* ou par une de ces homologues comme HENN1 chez *C. elegans* (Bagijn et al., 2012), DmHen1 ou Pimet chez la Drosophile (Horwich et al., 2007; Saito et al., 2007) ou Hen1 chez le poisson zèbre (Kamminga et al., 2010). HEN1 et ses homologues sont responsables de la méthylation en 2O' du dernier ribose des petits ARN qui est important pour leur stabilisation (Li et al., 2005). Chez *A. thaliana*, il existe une compétition entre les microARN et les siARN pour être méthylés par HEN1. Chez *A. thaliana*, l'ARN polymérase IV et l'ARN dépendante ARN polymérase II sont toutes deux essentielles pour la biogenèse des siARN de 24 nucléotides. La mutation de ces deux protéines restaure la méthylation des microARN dans un mutant de perte de fonction partielle de la méthylation des microARN (*hen1-2*). Les siARN ont donc une influence négative sur la méthylation des microARN par HEN1 (Yu et al., 2010).

Le séquençage à haut débit a permis de mettre en évidence des modifications post-transcriptionnelle des petits ARN en 3'. Ces modifications sont conservées et sont principalement soit une monoadénylation soit une monouridylation (Choi et al., 2012; Ruby et al., 2006; Wyman et al., 2011; Zhai and Meyers, 2013). Les modifications en 3′ des microARN varient selon les organismes et impliquent de nombreuses enzymes. Une étude montre que chez les mammifères environ 50% des microARN sont adénylés et environ 25% sont monouridylés (Chiang et al., 2010). A l'inverse, chez *C. elegans* 40% sont monouridylés et 25% monoadénylés. Chez l'homme PAPD4, PAPD5, ZCCHC11, MTPAP, ZCCHC6 et Star PAP/U6 TUTase/TUT1 sont impliqués dans ces modifications en 3' (Burroughs et al., 2010; Jones et al., 2009; Katoh et al., 2009; Wyman et al., 2011). Cette addition de nucléotide en 3' des petits ARN permet de réguler leur stabilité ou d'influencer leur capacité de liaison à l'ARN cible. Par exemple, il a été mis en évidence que l'ajout d'une adénine en 3' des microARN par PAPD4 stabilise leur liaison à leur ARNm cible (Burroughs et al., 2010).

L'uridylation des petits ARN

Les siARN et les microARN chez les plantes ainsi que les piARN chez les animaux peuvent être uridylés en 3'. L'uridylation des petits ARN et son implication dans la dégradation dans le sens 3'-5' a été mis en évidence chez *A. thaliana* grâce à la délétion de HEN1 qui est responsable de la méthylation en 2O' du dernier ribose des petits ARN. Dans un mutant de délétion *hen1-1*, les petits ARN sont oligouridylés et dégradés dans le sens 3'-5'. L'uridylation des microARN non méthylés provoque leur dégradation dans le sens 3'-5' par un mécanisme inconnu (Li et al., 2005). Il est possible que l'uridylation induit une dégradation très processive dans le sens 3'-5'.

Deux publications récentes montre que chez *A. thaliana*, la ncPAP nucléo-cytoplasmique HESO1 (HEn1 SuppressOr 1) est responsable partiellement de l'uridylation des petits ARN lorsqu'ils ne sont pas méthylés en 2O' (Ren et al., 2012; Zhao et al., 2012a).

Chez les algues vertes, *Chlamydomonas reinhardtii*, MUT68 uridyle les microARN et les siARN non méthylés en 2O' par un homologue de HEN1. L'uridylation de ces petits ARN induit leur dégradation par le cofacteur de l'exosome RRP6 humain *in vitro* (Ibrahim et al., 2010). Chez *C. elegans*, les siARN naturellement non méthylés impliqués dans la répression des ARNm CSR-1 sont uridylés par CDE-1. Leur uridylation est nécessaire à leur dégradation car en absence de CDE-1, ils s'accumulent (van Wolfswinkel et al., 2009). Les mécanismes de méthylation qui protège les petits ARN et d'uridylation qui induit leur dégradation sont conservés chez les autres animaux, incluant les mammifères (Horwich et al., 2007; Kamminga et al., 2010; Saito et al., 2007). Cependant, l'enzyme responsable de l'uridylation des piARN chez les animaux n'est pas à ce jour identifiée. Chez l'homme, ZCCHC11/TUT4, ZCCHC6/TUT7, qui sont aussi responsables de l'uridylation des précurseurs des microARN de la famille de *let-7*, peuvent avec Star-PAP/U6 TUTase, uridyler les microARN (Hagan et al., 2009; Jones et al., 2009; Thornton et al., 2012; Wyman et al., 2011).

Le mécanisme de méthylation/uridylation des petits ARN est conservé de l'algue verte à l'homme. Il existe une redondance d'activité (au moins partielle) des ncPAP responsables de l'uridylation chez l'homme, *A. thaliana*, *C. reinhardtii* et *C. elegans* (partie 2b de la Discussion) (Ibrahim et al., 2010; Li et al., 2005; van Wolfswinkel et al., 2009; Wyman et al., 2011). Le mécanisme de méthylation/uridylation des petits ARN peut avoir un rôle de contrôle qualité. En effet, l'uridylation des petits ARN empêcherait leur liaison correcte à AGO (Hung and Stumph, 2011). De plus, elle peut permettre la dégradation des petits ARN ayant perdu leur méthylation en 2O'. Cependant, dans une étude plus récente, il est mis en évidence que l'uridylation intervient après que les microARN soient tronqués en 3', lorsqu'ils sont pris en charge par AGO1 (Zhai et al., 2013). En effet, le séquençage à haut débit des microARN du mutant *hen1* chez *A. thaliana*, le riz et le maïs, met en évidence un mécanisme d'uridylation des microARN tronqués qui est conservé entre les espèces mais les microARN ciblés varient d'une espèce à l'autre. Sachant que l'uridylation s'effectue lorsque les microARN sont en association avec AGO1 et que les mêmes observations sont faites chez les plantes sauvage (seulement à une proportion plus faible), il est possible que l'uridylation des microARN ait un rôle dans la dégradation des microARN durant ou après l'assemblement du complexe RISC

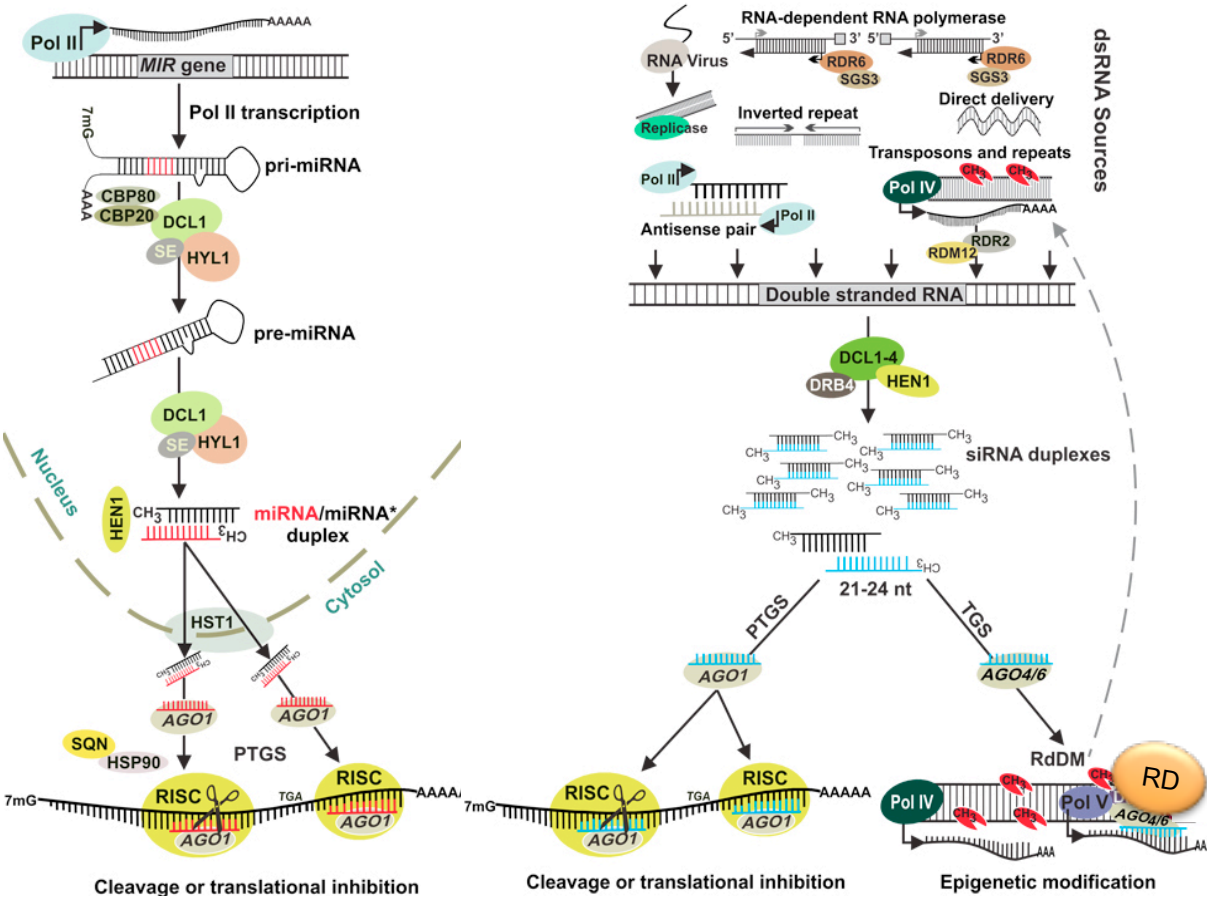


Figure IB.3 : La voie de biogenèse et le fonctionnement des microRNA et des siARN chez les plantes

A La voie de biogenèse des microRNA chez les plantes. Les précurseurs des micoRNA sont des transcrits de l'ARN polymérase II (Pol II) issus de région non codantes. DAWDLE stabilise leur conversion en structure tige-boucle dans des complexes ribonucléoprotéiques nucléaires (D-bodies). Cette étape nécessite l'action et l'interaction du domaine en doigt de zinc de SERRATE (SE), la protéine de liaison à l'ARN double brin HYPONASTIC LEAVES1 (HYL1), Dicer-*like* 1 (DCL1) et le complexe de liaison à la coiffe nucléaire (CBP80 et 20). Les pré-microRNAs, ou les microRNAs produits par DCL1, sont exportés dans le cytoplasme par l'exportine 5 orthologue de HASTY (HST1) et d'autres facteurs inconnus. Les ARN duplexes issus des pré-microRNAs (microRNA/microRNA*, respectivement le brin guide et le brin dégradé) sont méthylés par HEN1, les protégeant de la dégradation par des exonucléases. Les microRNA guides sont pris en charge par le complexe RISC contenant AGO1 afin d'effectuer la réaction de répression de la traduction et le clivage de l'ARNm cible. AGO1 est activé par SQUINT et la HSP90. **B** La voie de biogenèse des siARN chez les plantes. Les sources de l'ARN doubles brins précurseurs des siARN varient et sont clivés par l'une des trois protéines (DCL2, DCL3 et/ou DCL4) assistées par une protéine de liaison à l'ARN double brin. HEN1 stabilise les siARN et sélectionne le brin guide du duplexe siARN dans le complexe RISC contenant AGO pour la répression de la traduction ou le clivage endonucléolytique. Ces siARN peuvent guider AGO4 or AGO6 afin d'intervenir dans la voie de la *RNA-directed DNA Methylation* (RdDM) impliquant Pol IV et Pol V. La méthylation des cytosines à des sites spécifiques implique différents facteurs de remodelage de l'ADN (RD). D'après (Khraiwesh et al., 2012; Voinnet, 2009).

dans les plantes sauvage (Zhai et al., 2013). Tous les facteurs de la dégradation des microARN uridylés après qu'ils soient tronqués ou non ne sont pas encore connus chez les plantes ni chez les animaux.

4. Les fragments d'ARNm clivés par le complexe RISC

a) Le mode d'action des petits ARN

Les petits ARN régulent la quantité des ARNm cibles par une répression de leur traduction ou par une dégradation grâce à un clivage endonucléolytique (Introduction partie A3). Chez les plantes, les microARN sont parfaitement complémentaires à leur ARNm cible, permettant le clivage endonucléolytique de l'ARNm par une protéine Argonaute faisant partie du complexe RISC. Chez les animaux, les microARN ne sont principalement pas parfaitement complémentaires à leur cible, ce mésappariement central exclu la possibilité de clivage de l'ARNm. Les microARN des mammifères permettent donc principalement la répression de la traduction de leur ARNm cible. Cependant, certains microARN, encodés par des virus ou non peuvent permettre le clivage de leur ARNm cible (Pfeffer et al., 2004; Yekta et al., 2004). Chez les plantes, la répression de l'expression des ARNm sans clivage est effectuée principalement par les siARN ainsi que certains microARN. Il existe donc une balance entre la répression de la traduction et le clivage des ARNm selon le microARN se liant sur l'ARNm. Cette balance permet ainsi une grande flexibilité et versatilité de la fonction des microARN. En effet, le clivage des ARNm est irréversible tandis que la répression de leur traduction permet une régulation réversible, très importante lors d'une réponse au stress (Brodersen et al., 2008).

b) La dégradation des sous-produits du clivage par RISC de l'ARNm

Chez les plantes, le fragment 3', sous-produit du clivage de l'ARNm par le complexe RISC possède une extrémité 5' phosphate et est un substrat de l'exonucléase XRN4 chez *A. thaliana* (Souret et al., 2004). Le fragment 5' quant à lui est uridylé par une à sept uridines à son extrémité 3' par une ou plusieurs uridylyltransférase(s) inconnue(s) (Shen and Goodman, 2004). Cette uridylation est un signal de dégradation dans le sens 5'-3', et permettrait une dégradation du fragment 5'. Ce processus d'uridylation du fragment 5' de l'ARNm clivé par le complexe RISC, est conservé chez la souris et les cellules humaines infectées par le virus de Epstein-Barr (Shen and Goodman, 2004). Les ARNm clivés par le complexe RISC après reconnaissance par les siARN peuvent être polyadénylés par MUT68 chez *C. reinhardtii* induisant leur dégradation dans le sens 3'-5' (Ibrahim et al., 2006).

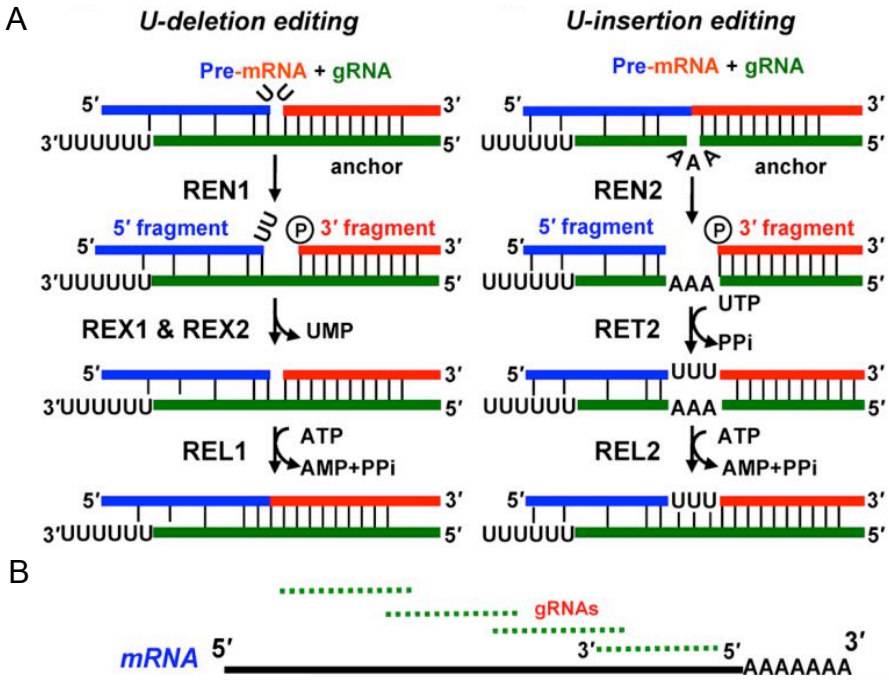


Figure IB.4 : L'édition des ARNm chez les trypanosomes par la délétion ou l'insertion d'uridines

A Les mécanismes de délétion d'uridines (à gauche) et d'insertion d'uridines (à droite). Les trypanosomes possèdent deux complexes d'édition des ARN, un spécifique de la délétion d'uridines et l'autre de l'insertion d'uridines. Les deux complexes contiennent une endonucléase et une ligase et respectivement, deux exonucléases (REX1 et REX2) pour la délétion d'uridines, et l'uridylyltransférase RET2 pour l'insertion d'uridines. L'ARN guide forme une région d'ancrage (*anchor*) de 5 à 15nt double-brin, par la liaison de l'extrémité 5′ de l'ARNg à l'ARNm pré-édité. Le clivage endonucléolytique initial de l'ARNm est localisé au niveau des premiers nucléotides non-appariés en amont de la région d'ancrage. La réaction d'édition commence par un clivage endonucléolytique de l'ARNm par une RNase III-*like*, en amont de la séquence d'ancrage double brin, laissant une extrémité 3'OH et 5' phosphate. Ensuite, les exonucléase 3'-5' éliminent les résidus uridines non appariés du fragment 5' de l'ARNm ou RET2 ajoute des uridines selon le nombre de purine de l'ARN guide (gRNA pour *guide RNA*). Les fragments d'ARNm sont ensuite reliés par une ligase. La modification de la séquence de l'ARNm peut créer un nouveau site de complémentarité avec un autre ARNg, permettant ainsi une hiérarchisation de l'édition dans le sens 3'-5' de l'ARNm. **B** Polarisation de l'édition de l'ARNm dans le sens 3'-5'. La polarité 3′-5′ de l'édition de l'ARNm est déterminée par la hiérarchisation de la liaison de l'ARNg dans la région 3′ de l'ARNm. L'édition créée souvent des sites de liaison pour l'ARNg suivant et ainsi de suite. D'après (Aphasizhev and Aphasizheva, 2011a, 2011b).

La dégradation des ARNm clivés par le complexe RISC n'est donc pas totalement caractérisée.

5. Les ARN guides impliqués dans l'édition des ARNm chez les trypanosomes

a) L'édition des ARNm

L'édition des ARN est un terme utilisé pour nommer un ensemble de mécanismes modifiant la séquence de l'ARN par rapport à la séquence initiale génomique. L'édition des ARN comprend l'insertion, la délétion et la modification de nucléotides. Grâce à l'édition, le même gène donne naissance à des ARNm différents. Chez les plantes, l'édition correspond à une modification de C en U pour de nombreux ARN dans les mitochondries et certains transcrits des chloroplastes. Cependant, le mécanisme d'édition le mieux caractérisé est l'insertion et la délétion d'uridines au niveau des ARNm mitochondriaux chez les trypanosomes, cette édition nécessitant des petits ARN guides polyuridylés (ARNg) (Figure IB.4). Le génome mitochondrial des trypanosomes est composé d'un réseau d'ADN circulaire maxi-cercle, de 25kb, homologues des ADN mitochondriaux des autres eucaryotes, et de mini-cercles, de 1kb, codant pour des petits ARN guides. Les petits ARN guides contrôlent l'ajout ou la délétion d'uridines lors de l'édition des ARN (Figure IB.4).

b) Les petits ARN guides uridylés

Les petits ARN guides sont des ARN de 50 à 60 nucléotides, composés d'une séquence d'ancrage en 5' qui s'apparie aux transcrits pré-édités, au centre, d'une séquence complémentaire à la région éditée qui détermine le nombre d'uridines à ajouter ou à supprimer et, en 3', une queue polyuridine. Lors de la biogenèse des ARNg, le précurseur des ARNg et un brin anti-sens s'apparient. Ils sont ensuite clivés par une endonucléase et l'uridylyltransférase RET1 ajoute 15 à 20 uridines en 3' au brin sens et au brin anti-sens (Figure IB.5) (Aphasizhev and Aphasizheva, 2011a, 2011b; Aphasizheva and Aphasizhev, 2010; McManus et al., 2000). Le rôle de cette queue polyuridine n'est pas déterminé. Même s'il est estimé qu'elle stabiliserait le complexe ARNm/ARNg, des études montrent qu'elle n'est pas nécessaire pour l'édition *in vitro* ni pour la stabilité *in vivo* du complexe ARNm/ARNg (Aphasizheva and Aphasizhev, 2010; Seiwert et al., 1996). Il est possible que cette queue polyuridine soit un signal de dégradation pour le brin sens comme pour le brin anti-sens (Figure IB.5) (Aphasizhev and Aphasizheva, 2011b).

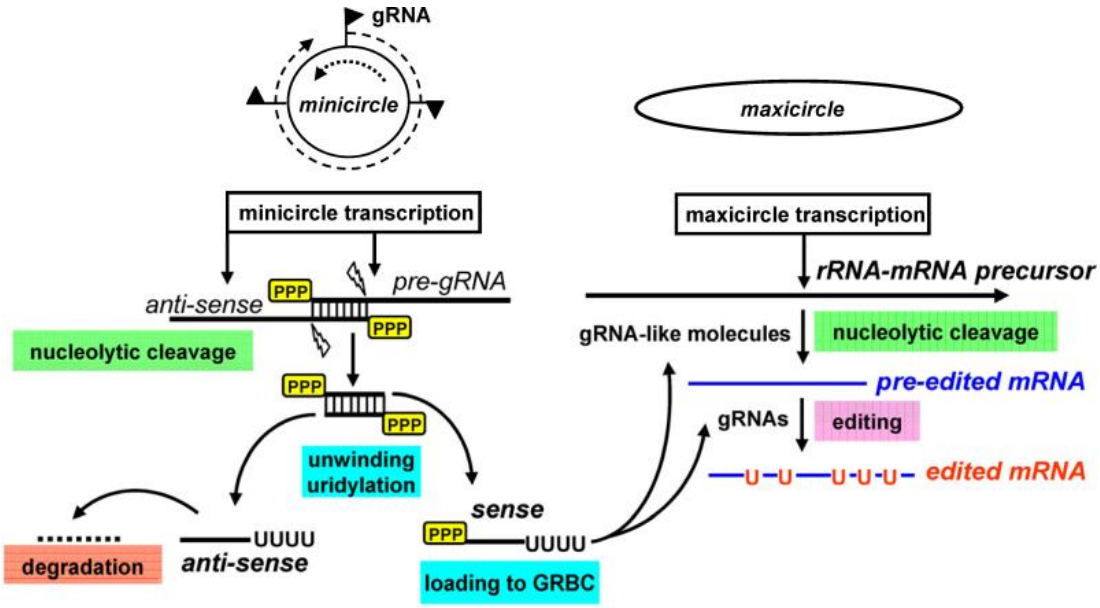


Figure IB.5 : La biogenèse des ARN guide et des ARN guide-*like* chez les trypanosomes

La biogenèse des ARNg et ARNg-*like* chez les trypanosomes. Ils sont respectivement impliqués dans l'édition de l'ARN et les maxicercles précurseurs de leur biogénèse. Le substrat ARN double brin nécessaire à l'endonucléase peut être généré par un appariement des transcrits synthétisés à partir du brin opposé. L'uridylation semble être effectuée après coupure du duplexe ARNg à la taille nécessaire. Il est possible que l'uridylyltransférase RET1 ajoute 15 à 20 uridines au complexe de liaison à l'ARNg ou *gRNA* pour (*guide RNA*) et aux ARNg-*like*. La première étape de l'édition est le clivage de l'ARNm dirigé par l'ARNg. Le clivage du précurseur multicistronique est dirigé par une action en trans de petites molécules d'ARN liées au *gRNA binding complex* (GRBC). Les endonucléases de l'édition sont de la famille des RNases III, suggérant que les membres de la même famille participent dans le clivage des ARNm dirigés par les petits ARN. D'après (Aphasizhev and Aphasizheva, 2011a).

6. Les ARNm

a) Les ARNm polyadénylés

L'uridylation des ARNm polyadénylés a été découverte chez *S. pombe*. La PolyU Polymérase Cid1 est capable d'ajouter une à trois uridines à l'extrémité 3' des ARNm, indépendamment de la taille de la queue polyA des ARNm. L'uridylation est impliquée dans la voie de dégradation des ARNm car la délétion de Cid1 augmente la demi-vie de l'ARNm d'un gène (Rissland and Norbury, 2009). L'uridylation des ARNm chez *S. pombe* induit l'élimination de leur coiffe, permettant ainsi la dégradation dans le sens 5'-3' (Rissland and Norbury, 2009). L'uridylation et la déadénylation jouent un rôle complémentaire dans la dégradation. En effet, l'accumulation de transcrits uridylés est observée dans un contexte d'inactivation du complexe de déadénylation Ccr4 tandis que dans un mutant inactif de Cid1, les queues polyA des transcrits sont plus courtes. Chez les mammifères, l'ajout de cinq uridines à l'extrémité 3' d'un ARN stimule l'élimination de la coiffe *in vitro* (Song and Kiledjian, 2007). En effet, l'ajout de six uridines successives ou plus augmente l'affinité de liaison du complexe Lsm1-7 à l'extrémité 3' de l'ARN (Chowdhury et al., 2007; Song and Kiledjian, 2007).

D'autre part, l'inactivation de l'exonucléase 3'-5' Dis3L2 chez *S. pombe* induit une accumulation des ARNm oligouridylés ayant jusqu'à sept uridines successives. Dis3L2 dégrade préférentiellement les ARN uridylés *in vitro* et sa vitesse de dégradation est plus importante si l'ARN possède quatre uridines en 3' plutôt que deux (Malecki et al., 2013). L'ajout d'oligouridine en 3' des ARNm est donc un signal pour la dégradation dans le sens 3'-5' des ARNm par Dis3L2. L'uridylation des ARNm chez *S. pombe* pourrait avoir un rôle d'activation de la dégradation dans le sens 5'-3' ou dans l'autre 3'-5' selon la quantité d'uridines ajoutées en 3' (partie 4 de la Discussion).

b) Les ARNm oligoadénylés

Contrairement à *S. pombe, A. nidulans*, présente une modification en 3' des ARNm, par un ajout de CUCU, après une étape de déadénylation. En effet, la ncPAP CutA (CUCU nucléotidyltransférase) est capable d'ajouter une queue hétéropolymérique CUCU aux ARNm oligoadénylés n'ayant plus que 15 adénines environ en 3'. Comme chez *S. pombe,* la modification en 3' est impliquée dans la dégradation des ARNm et marque les ARNm décoiffés (Morozov et al., 2010). En effet, lors de la délétion de CutA la demi-vie des ARNm augmente et les *processing bodies*, qui sont des foci cytoplasmiques ayant un rôle putatif dans la dégradation (Introduction partie C2b), se désassemblent (Morozov et al., 2010). Cependant, dans

le mutant ΔCutA, une faible uridylation est toujours observée. *A. nidulans* ne possède qu'une autre ncPAP : CutB. Dans le mutant ΔCutB la demi-vie des ARNm augmente et, comme dans le mutant ΔCutA, les ARNm sont toujours sujets à l'ajout de U et de C en 3'. Cependant, en absence des deux ncPAP ΔCutAΔCutB, les ARNm n'ont plus de modification en 3' (Morozov et al., 2012b). Cette absence totale de modification en 3' du double mutant comparé aux simples indique que ces deux enzymes agissent ensemble. Cependant, lors de la délétion des deux enzymes responsables de la modification en 3' des ARNm oligoadénylés, la vitesse de dégradation n'est pas modifiée comparée à un simple mutant (Morozov et al., 2012b). La modification en 3' des ARNm par des bases pyrimidines aurait donc un rôle de signal non essentiel pour la dégradation dans le sens 5'-3' et un rôle de contrôle qualité des ARNm afin de dégrader les transcrits aberrants et de libérer les ribosomes engagés (Morozov et al., 2012b). Cette dernière étude présente une conservation des modifications en 3' des ARNm chez *A. thaliana*, par l'ajout de bases pyrimidines C et U à l'ARNm *CCR2* lorsque les queues polyadénylées sont de 14 adénines (Morozov et al., 2012b). Ces observations sont en cohérences avec nos résultats (Chapitre 2).

c) Les ARNm non adénylés

Chez l'homme, les ARNm des histones ne sont pas stabilisés par une polyadénylation mais par une structure tige-boucle en 3' très conservée. Cette structure tige-boucle est aussi importante pour la maturation des ARNm des histones, leur export ainsi que leur traduction (Dominski and Marzluff, 1999; Dominski et al., 2003; Sànchez and Marzluff, 2002). La structure tige-boucle en 3' forme un complexe avec la protéine de liaison au structure tige-boucle SLBP/HBP (*Stem-Loop Binding Protein*/*Hairpin-Binding protein*). SLBP/HBP est nécessaire à la biogenèse des pré-ARNm des histones, à la localisation des ARNm des histones dans le cytoplasme et à leur traduction (Erkmann et al., 2005; Marzluff, 2005; Sànchez and Marzluff, 2002). La première étape de la dégradation des ARNm des histones est effectuée par une uridylation en 3' par ZCCHC11/TUT4 (Schmidt et al., 2010). L'ajout d'uridine en 3' permet la liaison du complexe Lsm1-7 et induit l'élimination de la coiffe ainsi qu'une dégradation simultanée dans le sens 5'-3' et 3'-5' (Mullen and Marzluff, 2008). En effet, Lsm1 recrute l'exoribonucléase 3'hExo/ERI1 3'-5', spécifique des ARN simple brin qui dégrade les ARNm des histones après l'ouverture préalable de la structure tige-boucle par l'hélicase Upf1, recrutée par SLBP/HBP (Hoefig et al., 2013; Schmidt et al., 2010). 3'hExo/ERI1 est conservée et impliquée dans la dégradation des ARN, la biogenèse et la maturation de l'ARN ribosomique 5,8S (Tan et al., 2013; Yang et al., 2006b).

7. Conclusion

L'uridylation des ARN est un nouveau mécanisme conservé de la levure *S. pombe* à l'homme, effectuée par des ncPAP. Le rôle de l'uridylation des ARN varie selon les ARN ciblés et peut au sein d'un même ARN, jouer des rôles opposés selon la quantité d'uridines ajoutées en 3'. L'uridylation est principalement impliquée dans la dégradation et l'ajout d'une à trois uridines est responsable de la dégradation dans le sens 5'-3', tandis que l'ajout d'oligouridine est responsable de la dégradation dans le sens 3'-5'. Cependant, l'uridylation des ARN peut être aussi importante pour la stabilité des ARN, la maturation de leurs précurseurs et l'arrêt de la traduction des ARNm.

C—Les complexes ribonucléoprotéiques cytoplasmiques dans la dégradation et/ou le stockage des ARNm

Les *processing bodies* et les granules de stress sont deux structures formées par l'agrégation d'ARNm non-traduits et pourraient être impliquées dans la dégradation et/ou le stockage des ARNm, bien que leurs fonctions respectives ne soient pas totalement comprises à l'heure actuelle. L'association d'un nombre croissant de facteurs de dégradation au niveau des polysomes suggère que les ARNm en cours de traduction peuvent aussi être soumis à la dégradation, révélant ainsi l'existence du phénomène de dégradation co-traductionnelle chez les eucaryotes.

1. Les polysomes

Les ARNm sont stabilisés par une coiffe en 5' à laquelle se lie des protéines et une queue polyA en 3' qui est, elle, recouverte de PABP. Les protéines de liaison de la coiffe et les PABP interagissent ensemble afin de cycliser les ARNm et faciliter l'initiation de la traduction. La traduction des ARNm est régulée par la taille de leur queue polyA, leur capacité d'association à des petits ARN et/ou des protéines de liaison aux ARN comme par exemple celles qui se lient aux séquences riches en adénine et uridine de la 3'UTR (ARE pour AU-*rich destabilizing elements*) (Introduction partie A1a et Figure IA.3). En effet, les protéines de liaison aux séquences ARE (AUB) se lient aux séquences ARE et peuvent soit stabiliser soit déstabiliser les ARNm. Par exemple, HuR inhibe les déadénylases et stabilise les ARNm (Introduction partie A1a). Cependant, des AUB comme BRF1 et TTP, recrutent les activités de déadénylation PARN et Ccr4-Pop2/Caf1-Not et induisent la

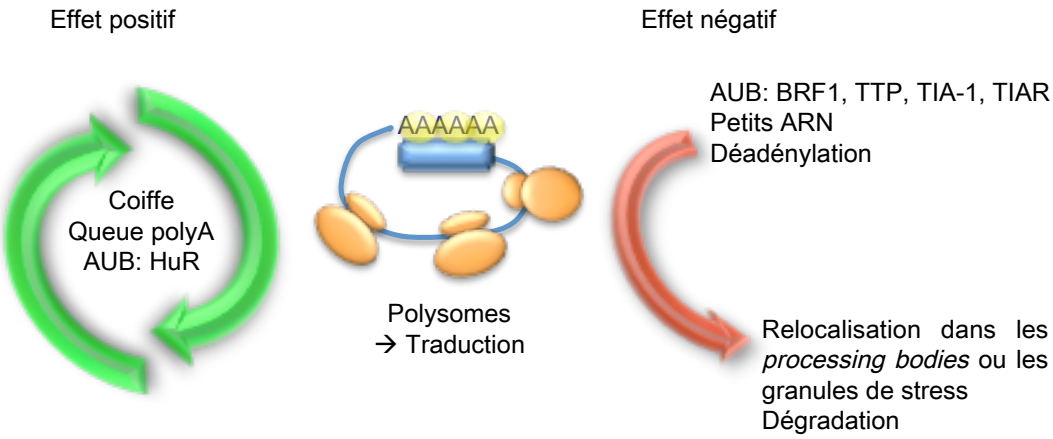


Figure IC.1 : La régulation de la traduction des ARNm

Les facteurs ayant un effet sur la traduction des ARNm au niveau des polysomes sont indiqués, en vert les facteurs activateurs de la traduction et en rouge les facteurs déstabilisants les polysomes et inhibiteurs de la traduction. AUB sont les protéines de liaison aux ARNm ayant des séquences ARE. Elles peuvent être stabilisantes comme HUR ou déstabilisantes en recrutant les complexes de déadénylation (partie A1 de l'introduction). Le rectangle bleu représente le complexe d'initiation de la traduction, la queue polyA de l'ARNm est recouverte de PABP.

dégradation de l'ARNm cible (Introduction partie A1a). D'autres AUB comme TIA-1, TIAR, régulent la traduction des ARNm (Zhang et al., 2010b). D'autre part, les petits ARN ciblent leur ARNm complémentaire et leur traduction peut être réprimée ou les ARNm peuvent être dégradés après un clivage endonucléolytique par le complexe RISC en coopérant ou non avec les AUB (Introduction partie A1a et Figure IA.3). Le raccourcissement de la taille de la queue polyA des ARNm est habituellement la première étape de la dégradation des ARNm (Zhang et al., 2010b). La régulation de la traduction des ARNm et de leur dégradation intervenant au niveau des polysomes permet la translocation des ARNm dans des foci cytoplasmiques (*processing bodies* et les granules de stress) ou leur dégradation (Figure IC.1).

Jusqu'à récemment, il existait un dogme selon lequel les ARNm devaient être dissociés des polysomes avant d'être dégradés. En effet, chez la levure, l'étape de déadénylation des ARNm induit une perte de leurs facteurs de traduction (Pab1, eIF4E et eIF4G) et ces facteurs sont remplacés par des facteurs de dégradation tels que le complexe Lsm1-7 et le complexe d'élimination de la coiffe (Dcp1 et Dcp2) (Tharun and Parker, 2001). Pourtant, chez la levure et l'homme, des facteurs de dégradation tels que XRN1/PACMAN, Dcp1 et Dcp2, et des facteurs de la voie de la dégradation des ARNm ayant un codon stop précoce par la voie du NMD tels que UPF1, UPF2 et UPF3 co-sédimentent avec les polysomes (Mangus and Jacobson, 1999; Wang et al., 2002).

La co-sédimentation des facteurs de dégradation avec les polysomes indique que les ARNm n'ont pas toujours besoin d'être relocalisés afin d'initier leur dégradation. Cette idée est d'autant plus renforcée que, chez la levure, l'élimination de la coiffe s'effectue par l'association de l'ARN hélicase Dhh1/RCK/p54 DEAD-box au niveau des polysomes. L'activateur du complexe d'élimination de la coiffe Dhh1/RCK/p54 réprime les facteurs d'initiation eIF4 et eIF3b et ralentit la translocation des ribosomes sur l'ARNm, provoquant ainsi leur accumulation sur l'ARNm (Sweet et al., 2012). Par la suite, l'hélicase Dhh1/RCK/p54 active Dcp2 qui élimine la coiffe des ARNm. Une étude récente montre que les ARNm décoiffés sont associés au niveau des polysomes. Cette première et très importante étude met ainsi en évidence l'existence de la dégradation co-traductionnelle polarisée dans le sens 5'-3' (Hu et al., 2009). Selon Hu *et. al.*, la voie de dégradation co-traductionnelle dans le sens 5'-3' aurait évoluée pour permettre la translocation complète du dernier ribosome sur l'ARNm avant sa dégradation, évitant ainsi la traduction d'ARNm tronqués (Hu et al., 2009). Les facteurs de la voie de dégradation du NMD (UPF1, UPF2 et UPF3) co-sédimentent avec les polysomes chez *S. cerevisiae* (Mangus and Jacobson, 1999) et les substrats ARN intermédiaires de la dégradation des ARNm

ayant un codon stop précoce de la voie du NMD dans le sens 5'-3' sont associés au niveau des polysomes (Hu et al., 2010). D'après ces résultats, la dégradation des ARNm peut être co-traductionnelle dans le sens 5'-3' pour les ARNm « normaux » et les ARNm ayant un codon stop précoce.

Ce mécanisme de dégradation co-traductionnelle dans le sens 5'-3' observée chez la levure, l'homme et semble être aussi effectif chez *A. thaliana*. En effet, plusieurs facteurs de dégradation dans le sens 5'-3' (tel que XRN4) sont retrouvés au niveau des polysomes (communication personnelle des Dr Deragon, Antonelli et Merret du laboratoire Génome et Développement des Plantes à Perpignan).

Une étude très récente chez l'homme, implique la voie de dégradation dans le sens 3'-5' au niveau des polysomes. L'exoribonucléase 3'-5', hDIS3L2, indépendante de l'exosome, qui, comme hXRN1 co-purifie avec les monosomes et les polysomes en présence d'ARN (Lubas et al., 2013). Cette récente implication de la voie de dégradation dans le sens 3'-5' au niveau des polysomes va à l'encontre de la première hypothèse du rôle de la dégradation co-traductionnelle dans le sens 5'-3' qui empêcherait la traduction de protéines tronquées. Cependant, même si hDIS3L2 co-sédimente avec les polysomes, pour l'instant, aucun ARNm, intermédiaire de la dégradation dans le sens 3'-5' n'a été mis en évidence associé aux polysomes (partie D3 de la Discussion).

2. Les granules cytoplasmiques

Après arrêt de la traduction des ARNm, ils peuvent être relocalisés dans des particules ou foci cytoplasmiques. Il existe plusieurs types de foci dans la cellule situés dans le noyau ou le cytoplasme, visibles au microscope (Anderson and Kedersha, 2006). Ces derniers incluent les *processing bodies* et les granules de stress cytoplasmiques. Ils sont formés par l'agrégation contrôlée de particules ribonucléoprotéiques (RNPs). La composition des deux granules cytoplasmiques (*processing bodies* et granules de stress) est plus connue chez les mammifères et la levure que chez les plantes. Les *processing bodies* et granules de stress partagent un groupe de protéines communes mais ils sont caractérisés par un groupe de protéine spécifique à chacun d'eux (Buchan and Parker, 2009). Les *processing bodies* sont principalement enrichis en facteurs impliqués dans la dégradation des ARN tandis que les granules de stress comprennent des facteurs d'initiation de la traduction. Je présente les points communs de ces deux foci cytoplasmiques, les protéines spécifiques les composants, les facteurs importants pour leur formation et leurs rôles (putatifs) chez les mammifères et la levure. Je présente aussi les données

Tableau IC.1 : Les protéines composant les *processing bodies* et les granules de stress chez les mammifères

	Processing bodies	Granules de stress
protéines communes	Les protéines de répression de la traduction : TTP, BRF1, CPEB, AGO, Lin28 (liaison à ARN) ; Hsp90 ; Mex3b ; FAST (*Fas-Activated Ser/Thr phosphoprotein*) ; PCBP2 (*Poly rC Binding Protein* 2) ; APOBEC le facteur de l'édition ; Rap55, l'activateur de l'élimination de la coiffe, XRN1, l'hélicase Rck/p54/Dhh1, eIF4E, initiant la traduction	
protéines spécifiques des différents types de foci cytoplasmiques	Les facteurs de dégradation : Dcp1, Dcp2, Lsm1-7, Pat1, Edc1-3, GW182, Ccr4, Pop2, Pan2, Hedls, Upf1-3, SMG5, SMG7 ; eIF4E-T et Mex3A, réprimant la traduction	Pabp, la sous-unité 40S du ribosome, les protéines régulant la traduction TIA-1/R et HuR ; les facteurs d'initiation de la traduction eIF4G et eIF4A ; les protéines de liaisons à l'ARN FMRP, FXR1/2, Mex67, Staufen et ZBP1 (*β-actin zipcode-binding protein 1*), impliquée dans le transport des ARNm ; les protéines de modification post-traductionnelle, la kinase RSK2, RACK1 (*Receptor for Activated C Kinase 1*), TRAF2, nécessaire à l'activation de kinase GSBP (*glycosylation site-binding protein*)

D'après (Buchan and Parker, 2009; Matsumoto et al., 2011; Sawarkar et al., 2012; Thomas et al., 2011a; Wilczynska et al., 2005).

connues chez *A. thaliana*. Dans une autre partie, je précise l'interaction existante entre les polysomes, les *processing bodies* et les granules de stress et j'intégre l'importance des *processing bodies* et des granules de stress lors d'une infection virale.

a) Les points communs et les spécificités des granules de stress et des *processing bodies*

Les *processing bodies* et les granules de stress sont tous deux des foci cytoplasmiques dynamiques formés suite à un arrêt de la traduction des ARNm. Les microtubules, sont essentiels à leurs mouvements dans le cytoplasme, leur association et leur dissociation. Les *processing bodies* existent de manière constitutive dans la cellule et leur taille et leur nombre sont modifiés lors de certains stress. A l'inverse, les granules de stress sont nouvellement formés en réponse à un stress et sont ensuite dissociés lorsque le stress est levé. Par ailleurs, les structures microscopiques de ces deux foci cytoplasmiques sont différentes. En effet, les *processing bodies* sont compacts et fibrillaires tandis que les granules de stress sont plus lâches et fibro-granulaires (Souquere et al., 2009).

Les microtubules sont essentiels à la formation des *processing bodies* et des granules de stress car la délétion d'un des moteurs moléculaires composants des microtubules tels que la dynéine ou la kinésine empêche leur formation en réponse à un stress (Loschi et al., 2009). Chez les mammifères, lors d'un stress, la tristetrapolin (TTP), une protéine de liaison aux séquences ARE et/ou son paralogue BRF1 sont déphosphorylés afin de permettre leur appariement avec les séquences ARE des ARNm. Leur association à l'ARNm permet l'arrêt de leur traduction. L'ARNm non-traduit est soit dégradé soit relocalisé au niveau des *processing bodies* et/ou des granules de stress (Franks and Lykke-Andersen, 2007; Kedersha et al., 2005). Ces deux granules cytoplasmiques contiennent principalement des protéines de liaison à l'ARN telles que les protéines de liaison aux séquences ARE, TTP et BRF1, ainsi que d'autres facteurs de répression de la traduction tels que CPEB et AGO (Tableau IC.1). Cependant, les deux granules cytoplasmiques contiennent aussi des protéines impliquées dans la dégradation 5'-3' telles que l'exoribonucléase 5'-3' XRN1/PACMAN et l'hélicase Dhh1/RCK/p54, même si les facteurs de la dégradation 5'-3' sont principalement présents dans les *processing bodies* (Tableau IC.1) (Buchan and Parker, 2009; Thomas et al., 2011).

Chez les plantes, l'exoribonucléase 5'-3' XRN4 est présente au niveau des *processing bodies* et est absente des granules de stress (Weber et al., 2008). A l'inverse, certaines protéines de liaison à l'ARN, homologues de TTP des

mammifères, possédant un domaine en doigt de zinc en tandem CCCH (TZF) telles que AtTZF1, AtTZF4, AtTZF5, AtTZF6, sont aussi localisées dans les granules de stress et les *processing bodies* en réponse à des stress hormonaux par le méthyl jasmonate, l'acide abscissique, l'acide gibbérellique ou un stress lumineux (Bogamuwa and Jang, 2013; Pomeranz et al., 2010). Cependant, peu de données sur les foci cytoplasmiques des plantes sont connues à ce jour.

b) Les *processing bodies*

1. La composition des *processing bodies*

Les ARNm non traduits peuvent s'assembler sous forme de complexes ribonucléoprotéiques formant des foci cytoplasmiques tels que les *processing bodies* appelés P-*bodies* ou GW-*bodies* chez les animaux, même si les GW-bodies contenant GW182 et AGO2 semblent être distincts des P-*bodies* (Gibbings et al., 2009). Les *processing bodies* sont des structures dynamiques, qui se déplacent dans la cellule grâce aux microtubules et qui sont constitutivement présents dans la cellule. Cependant, leur nombre et leur taille varient selon l'efficacité de la traduction des ARNm. Les *processing bodies* peuvent être en contact avec les organelles telles que le réticulum endoplasmique, les multivésicules, qui sont des intermédiaires entre la surface de la cellule et le lysosome, et les mitochondries (Weil and Hollien, 2013). En effet, plusieurs composants des *processing bodies* peuvent interagir au niveau de la surface des mitochondries (Huang et al., 2011).

Les *processing bodies* contiennent principalement des facteurs de la dégradation dans le sens 5'-3' comme le complexe Lsm1-7-Pat1, le complexe d'élimination de la coiffe et ses activateurs tels que Dhh1/RCK/p54, SDC6/Rap55/DCP5, Edc1-3, Dcp1 et Dcp2. Le complexe de déadénylation Ccr4-Caf1/Pop2-Not ainsi que Pan2 peuvent également être retrouvés au niveau des *processing bodies*, bien qu'ils soient majoritairement localisés de manière diffuse dans le cytosol. Les *processing bodies* contiennent aussi des protéines impliquées dans la voie de dégradation des ARNm aberrants contenant des codons stop précoces par la voie du NMD telles que Upf1p, Upf2p, Upf3p, SMG5 et SMG7, des facteurs de répression de la traduction tel que eIF4E-T (Buchan and Parker, 2009; Parker and Sheth, 2007). La machinerie de dégradation dans le sens 3'-5' n'est pas présente chez les animaux et la levure *S. cerevisiae*. Par exemple, la sous-unité Ski7 qui permet la liaison du complexe Ski à l'exosome dans le cytoplasme n'est pas associée aux *processing bodies* (Sheth and Parker, 2003).

Les principaux facteurs des *processing bodies* sont conservés entre les différents organismes. En effet, chez *A. thaliana,* les *processing bodies* comportent

comme chez la levure et l'homme, des facteurs de dégradation dans le sens 5'-3' tels que le complexe d'élimination de la coiffe, DCP1, DCP2, VCS, DCP5, l'activateur de l'élimination de la coiffe Dhh1 et l'exoribonucléase 5'-3' XRN4 (Weber et al., 2008; Xu and Chua, 2009, 2011). Les *processing bodies* chez *A. thaliana* contiennent aussi les facteurs de la voie de dégradation du NMD tels que UPF1 à 3 et SMG7 (Mérai et al., 2012). Différemment de chez les mammifères, chez les plantes, eIF4E est seulement localisée dans les *processing bodies* et non dans les *processing bodies* et les granules de stress (Weber et al., 2008). GW182 n'existe pas chez les plantes mais les *processing bodies* contiennent une hélicase putative, SDE3, impliquée dans la réponse antivirale qui possède une répétition GW (Garcia et al., 2012; Xu and Chua, 2011). D'autres protéines sont associés aux *processing bodies* tels que le suppresseur viral de *silencing* CMV2b et AGO1 qui est impliqué dans la répression des ARNm par complémentarité avec des petits ARN (Pomeranz et al., 2010; Xu and Chua, 2011). L'exoribonucléase 3'-5' SOV de l'écotype Landsberg, qui est localisé au niveau de granules cytoplasmiques, est associée aux *processing bodies* (communication de Zhang et Sieburth du *Department of Biology*, *University of Utah*, Salt Lake City, Utah) (Zhang et al., 2010a).

2. La formation des *processing bodies*

Les *processing bodies* sont constitutivement présents dans les cellules et ils sont dynamiques. En effet, leur nombre et leur taille peuvent être modifiés selon la traduction des ARNm. Les ARNm sont essentiels à la formation des *processing bodies*. En effet, lorsque des cellules perméables sont traitées avec de la RNase, les *processing bodies* se dispersent (Eulalio et al., 2007a). Cet effet est aussi observé lorsque les ARNm sont bloqués au niveau des polysomes, par l'action du cycloheximide (Weber et al., 2008). Cependant, lorsqu'un ARNm est très fortement exprimé et qu'il n'est donc plus piégé au niveau des polysomes, les *processing bodies* peuvent se reformer (Franks and Lykke-Andersen, 2007). Lorsque la traduction des ARNm est inhibée, les *processing bodies* se forment et, à l'inverse, lorsque l'élongation de la traduction est inhibée, bloquant les ARNm au niveau des polysomes, les *processing bodies* se dissocient (Figure IC.2). Chez les mammifères, l'hélicase Dhh1/Rck/p54, l'activateur de l'élimination de la coiffe SDC6/Rap55/DCP5, l'inhibiteur de traduction CPEB1, eIF4E-T et la déadénylase Ccr4 sont aussi essentielles à l'assemblement des *processing bodies*, une délétion de l'une d'entre elles cause la disparition des *processing bodies* (Andrei et al., 2005; Ernoult-Lange et al., 2012; Serman et al., 2007; Yang et al., 2006a). La formation des *processing bodies* s'effectue étape par étape. En effet, tout d'abord, l'hélicase Dhh1/Rck/p54

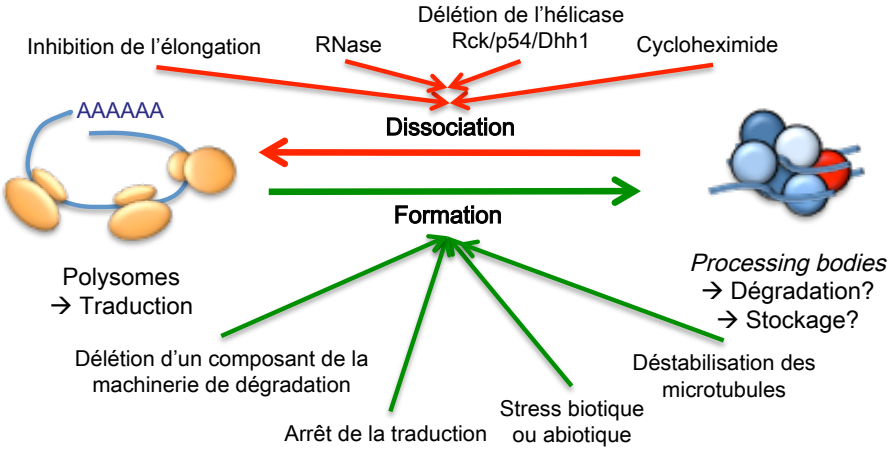


Figure IC.2 : Les facteurs régulant la formation et la dissociation des *processing bodies*

Tableau IC.2 : Les domaines responsables de la localisation des protéines au niveau des *processing bodies*

Domaines	Protéines	Références
Yjef-N	Edc3	(Decker et al., 2007; Reijns et al., 2008; Tritschler et al., 2007)
Lsm	Edc3, DCP1	
riche en Q	GW182	(Eulalio et al., 2009)
riche en GW	GW182	
riche en Q/N	Lsm4, EDC4, CCR4, Caf1/Pop2, Dhh1/RCK/p54*, TIA-1 et TIAR	(Buchan et al., 2008; Gilks et al., 2004; Reijns et al., 2008; Thomas et al., 2011a)

* seulement chez les mammifères et l'homme

relâcherait les structures des ARNm afin de permettre leur entrée dans les *processing bodies* (Ernoult-Lange et al., 2012). Puis Pat1 intervient et induit l'assemblement des *processing bodies* et l'élimination de la coiffe des ARNm par son association avec DCP2. Le facteur de répression de la traduction, CPEB1, quant à lui, serait nécessaire à leur maintenance (Ozgur and Stoecklin, 2013; Serman et al., 2007). GW182 et Ge-1/ Hedls/VCS qui ne sont pas présentes chez *S. cerevisiae*, sont essentielles à la formation des *processing bodies* chez les mammifères (Anantharaman and Aravind, 2004; Yang et al., 2004; Yu et al., 2005). Certains domaines de protéines localisées dans les *processing bodies* sont essentiels pour la formation des agrégats cytoplasmiques (Tableau IC.2). Ces domaines sont le domaine Yjef-N, Lsm, les domaines riches en GW, en Q et en Q/N, ces deux derniers ressemblent à un domaine des prions (Tableau IC.2) (Reijns et al., 2008; Thomas et al., 2011; Tritschler et al., 2007). Le domaine riche en Q/N de Lsm4 n'est pas conservé chez les eucaryotes. Chez les métazoaires, il est remplacé par un domaine riche en RG (Reijns et al., 2008). Ce dernier est responsable de la localisation de plusieurs protéines au niveau des granules de stress (Tableau IC.3) (Krüger et al., 2013; De Leeuw et al., 2007; Marko et al., 2012; Tourrière et al., 2003; Yang et al., 2006a).

Beaucoup des composants des *processing bodies* contiennent des régions riches en Q/N. Chez la levure, lorsque la sous-unité CCT3 de la chaperonne CCT/TRiC est mutée, la quantité de protéines riche en Q/N varie dramatiquement et le nombre de *processing bodies* augmente. CCT3 contient un site de liaison aux protéines riches en Q/N, qui semblent être correctement repliées par la sous-unité CCT3 de la chaperonne CCT/TRiC (Nadler-Holly et al., 2012). La taille et la quantité de *processing bodies* varient selon le stress. En effet, l'inhibition de l'initiation de la traduction et la répression de la traduction des ARNm augmente le nombre de *processing bodies*. Par ailleurs, leur composition peut varier selon le type de stress (Thomas et al., 2011). Par exemple chez *S. cerevisiae*, la formation des *processing bodies* en condition acide requière Scd6//Rap55/DCP5. L'acide lactique, l'acide chlorique et l'acide acétique induisent la formation de *processing bodies* mais pas la formation de granules de stress. Le stress acide réduit l'initiation de la traduction plus faiblement que lorsque le stress est effectué par une carence en glucose. Cependant, l'acide lactique avec un pH supérieur à 3, l'acide propénoïque ou l'acide acétique n'induit pas la formation de *processing bodies*. Il existe un seuil d'induction des *processing bodies* à pH faible (d'environ 2,5) (Iwaki and Izawa, 2012). Lors d'une déstabilisation des microtubules par le bénomyl, inhibant la polymérisation des β-tubulines et/ou par le letrunciline-A, inhibant la polymérisation de l'actine, la formation

des *processing bodies* est stimulée (Sweet et al., 2007). Tout comme, lorsque la capacité d'action des facteurs de la dégradation est ralentie par une délétion de l'un de ces facteurs de la dégradation dans le sens 5'-3', le nombre des *processing bodies* augmente, comme lorsque l'exoribonucléase 3'-5' hDIS3L2, est absente (Figure IC.2) (Chen and Shyu, 2013; Coller and Parker, 2005; Cougot et al., 2004; Eulalio et al., 2007b; Franks and Lykke-Andersen, 2007; Lubas et al., 2013; Sheth and Parker, 2003; Weber et al., 2008).

Comme chez les mammifères et la levure, chez les plantes, les facteurs DCP5, DCP1 et Dhh1 sont nécessaires à la formation des *processing bodies*. Ceux-ci répriment la traduction des ARNm. Ensuite l'élimination de la coiffe peut être effectuée par DCP2 en complexe avec VCS. Cependant, l'élimination de la coiffe par DCP2 n'est pas nécessaire à la formation des *processing bodies* car en absence de DCP2, ils sont toujours présents (Xu and Chua, 2009). Les ARNm ne sont donc pas toujours dégradés dans les *processing bodies* qui contiennent la machinerie de la dégradation dans le sens 5'-3'. Cette observation appuie un rôle putatif de stockage des *processing bodies*.

3. Le rôle putatif des *processing bodies*

Le rôle des *processing bodies* n'est pas totalement résolu. Ils comportent principalement des protéines impliquées dans la dégradation des ARNm, la répression de l'initiation et l'élongation de la traduction. Ils semblent donc être un lieu de stockage ou de dégradation des ARNm mais ce rôle hypothétique reste à être formellement démontré. Les *processing bodies* contiennent des facteurs de la dégradation et des protéines de liaisons à l'ARN impliquées dans la voie de dégradation des ARNm ayant un codon stop précoce par la voie du NMD et la répression de l'expression des ARNm par l'action des microRNA (Kedersha et al., 2005; Teixeira et al., 2005). L'augmentation du nombre des *processing bodies* lorsque la dégradation est ralentie, renforce leur rôle putatif dans la dégradation des ARNm « normaux » ou aberrants et dans la répression de l'expression des ARNm par l'action des petits ARN et des ARNm ayant des séquences ARE (Franks and Lykke-Andersen, 2007; Kedersha et al., 2005; Weil and Hollien, 2013). Cependant, les *processing bodies* ne sont pas essentiels pour la voie de dégradation par la voie du NMD ni pour l'activité du des siARN ni pour la dégradation des ARNm. En effet, chez une souche de *S. cerevisiae*, ne formant pas de *processing bodies*, la vitesse de dégradation de deux transcrits rapporteurs n'est pas affectée (Decker et al., 2007). De plus, chez la mouche et l'homme, la dégradation des ARNm par la voie de

dégradation du NMD et la répression de la traduction par les microARN ou les siARN peuvent être découplées de la formation de *processing bodies* (Chu and Rana, 2006; Eulalio et al., 2007a; Huntzinger and Izaurralde, 2011; Serman et al., 2007). Par ailleurs, chez *S. cerevisiae*, un nombre limité d'ARNm peuvent entrer dans les *processing bodies* et être réengagés dans la traduction lors du rétablissement d'un stress (Arribere et al., 2011). Donc le rôle des *processing bodies* ne se limite pas à la dégradation des ARNm. Par ailleurs, les ARNm semblent pouvoir transiter des *processing bodies* aux granules de stress et/ou aux polysomes (Balagopal and Parker, 2009; Brengues et al., 2005). Les *processing bodies* pourraient donc être impliqués dans la dégradation, le stockage et le tri des ARNm.

Chez les plantes, les composants des *processing bodies* sont importants pour le développement. Comme présenté dans la partie A2a de l'introduction, la délétion des facteurs impliqués dans l'élimination de la coiffe DCP1, DCP2, DCP5 et VCS sont létales au stade post-embryonnaire (Xu et al., 2006), comme la délétion du facteur impliqué dans la voie de dégradation du NMD UPF1. Le phénotype le plus fort est observé lors d'une délétion de LSM4 (Xu and Chua, 2011). En effet, les embryons ne se développent pas au delà du stade globulaire. Il semble donc que les facteurs localisés dans les *processing bodies* chez les plantes soient importants pour les transitions entre les différents stades de développement. Cependant, le rôle des *processing bodies* n'est pas direct dans ces observations, même s'il est envisageable. Par exemple, les graines sèches contiennent une grande quantité de DCP1 et DCP5. Les *processing bodies* permettraient l'arrêt de la traduction des ARNm codant pour les protéines de stockage de la graine sans les dégrader car DCP2 n'est pas présent à ce stade (Xu and Chua, 2011). Chez les plantes, les *processing bodies* sont aussi le lieu de dégradation des transcrits ayant un codon stop prématuré (Mérai et al., 2012). Ils semblent aussi être impliqués dans la lutte contre les virus comme peut l'indiquer la présence du suppresseur viral CMV2b (Thomas et al., 2011). Par ailleurs, les *processing bodies* pourraient aussi avoir un rôle dans l'arrêt de la traduction par l'action des microARN et des siARN sur leur ARNm complémentaire par VCS (Brodersen et al., 2008). Cependant, les *siRNA bodies* accumulant RDR6 et SGS3 sont responsables de la biogenèse des tasi-ARN et sont distincts des *processing bodies* même s'ils partagent certaines protéines (Jouannet et al., 2012; Kumakura et al., 2009; Moreno et al., 2013). Ils contiennent AGO7, RDR6, SGS3, essentielles à la défense de la plante contre les virus par la voie du PTGS et VP6, nécessaire à la réplication des virus. Les *siRNA bodies* semble être un point de contrôle de la réplication du virus et de la défense de l'hôte.

En effet, après un arrêt de la traduction induit pas un stress, ils recueillent des marqueurs de granules de stress (Jouannet et al., 2012).

c) Les granules de stress

1. La composition des granules de stress

Les granules de stress sont des granules cytoplasmiques comprenant des ARNm non traduits et sont formés en réponse à un grand nombre de stress comme un stress oxydatif, thermique, viral ou une exposition aux UV (Buchan and Parker, 2009; Kedersha et al., 1999; Thomas et al., 2011). Le cœur des granules de stress est composé des facteurs d'inhibition de la traduction TIA-1 et TIAR, impliquées dans l'épissage, qui permettent d'assembler le complexe d'initiation de la traduction déficient eIF2-eIF5. TIA-1 et TIAR sont nucléaires mais elles sont relocalisées dans le cytoplasme au niveau des granules de stress en cas de stress (Anderson and Kedersha, 2002; Kedersha et al., 2002). Les granules de stress contiennent des ARNm polyadénylés, le complexe d'initiation de la traduction eIF4G, eIF4E, eIF4A, eIF4B, eIF2, eIF3, les PABP, et la sous unité 40S du ribosome (Balagopal and Parker, 2009; Buchan and Parker, 2009; Thomas et al., 2011). La composition des granules de stress varie selon la nature du stress induisant leur formation (Thomas et al., 2011). Par exemple, chez *S. cerevisiae*, les granules de stress formés lors d'un choc thermique contiennent eIF3 et eIF3 n'est pas présent dans les granules de stress lorsqu'ils sont formés en réponse à une absence de glucose (Buchan et al., 2008; Grousl et al., 2009). Les granules de stress comportent principalement des protéines de liaison à l'ARNm étant des facteurs d'inhibition de la traduction et contrairement aux *processing bodies*, ils ne comportent que XRN1/PACMAN de la machinerie de la dégradation (Tableau IC.1). Cependant, d'une façon particulière, en absence de l'hélicase Pck/p54/Dhh1, lorsque les *processing bodies* ne sont pas formés, les granules de stress peuvent contenir DCP1 lors d'un stress à l'arsénite ou lors d'une surexpression de CPEB1 (Serman et al., 2007; Wilczynska et al., 2005). Pourtant DCP1 est normalement localisé au niveau des *processing bodies.*

Tout comme pour les *processing bodies*, les principaux composants des granules de stress sont conservés entre les organismes. Chez *A. thaliana*, les granules de stress contiennent, comme chez la levure et les mammifères, des ARNm polyadénylés, les PABP et le facteur de l'initiation de la traduction eIFiso4E, eIF2α, eIF3 et eIF4E (Weber et al., 2008). Ils contiennent aussi les protéines d'inhibition de la traduction de la famille TIA-1 (RBP47 et UBP1) lors d'un choc thermique (Weber et al., 2008).

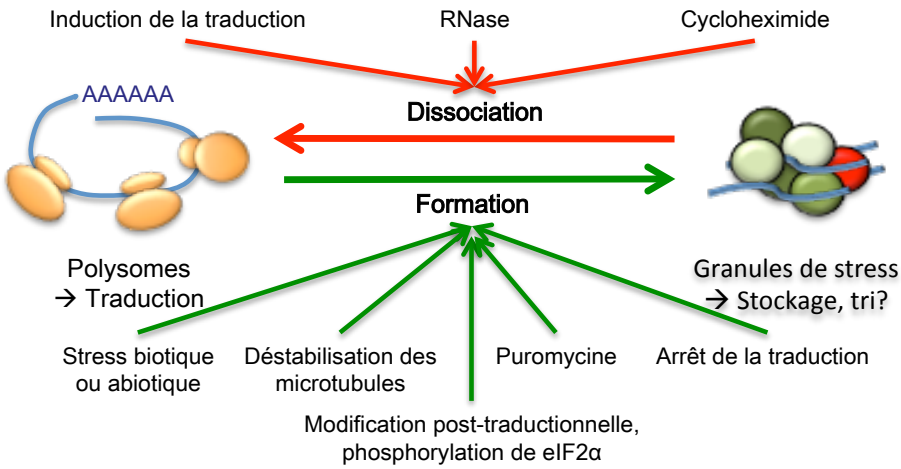


Figure IC.3 : Les facteurs régulant la formation et la dissociation des granules de stress

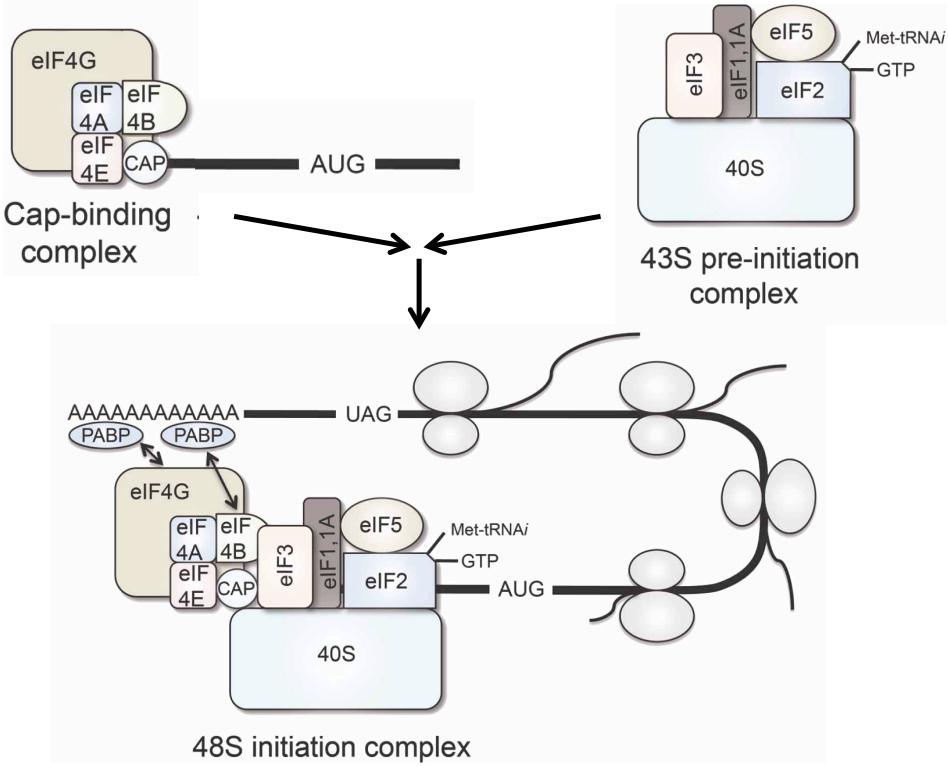


Figure IC.4 : Assemblage du complexe d'initiation de la traduction

eIF4E se lie à la coiffe, le complexe eIF4F, contenant eIF4G et eIF4A. eIF4B se lie à eIF4A et au complexe de liaison à la coiffe. La formation du complexe de pré-initiation 43S débute par l'assemblement du complexe ternaire (eIF2-GTP-ARNt^{met}), suivie par l'interaction avec eIF3, eIF1, eIF1A, et la sous-unité ribosomale 40S. L'interaction entre le complexe de liaison de la coiffe et le complexe de pré-initiation 43S recrute la sous-unité ribosomique 40S à l'extrémité 5′ de l'ARNm pour former de complexe d'initiation de la traduction 48S. La circularisation de l'ARNm est effectuée par l'interaction de eIF4G et eIF4B aux PABP. La traduction est initiée et, une fois que le complexe d'initiation 48S contient le codon d'initiation, eIF5 hydrolyse le complexe ternaire associé au GTP. Le recrutement de la sous-unité 60S du ribosome est suivit par l'initiation de la traduction. Les interactions entre eIF3c et eIF5, comme entre eIF3 et eIF2 dans le complexe d'initiation 43S ne sont pas indiquées dans ce schéma. D'après (Van Der Kelen et al., 2009).

2. La formation des granules de stress

Comme les *processing bodies*, les granules de stress sont des complexes dynamiques. Cependant, différemment des *processing bodies*, les granules de stress ne sont pas constitutivement présents dans la cellule mais ils sont nouvellement formés suite à un stress inhibant la traduction des ARNm. Le cœur des granules de stress est composé de TIA-1 et TIAR. Le stress réprime la traduction des ARNm avec des séquences ARE et des régions riches en cytidines par la liaison de TIAR et TIA-1 au niveau de ces régions de l'ARNm (Kim et al., 2007; Silanes et al., 2005). La présence d'ARNm est essentielle à la formation des granules de stress, comme pour les *processing bodies*. En bloquant les ARNm au niveau des polysomes, la formation des granules de stress est inhibée et à l'inverse, la dissociation les ribosomes des ARNm par la puromycine stimule leur formation (Figure IC.3) (Buchan et al., 2008). La structure des complexes RNP dans les granules de stress n'est pas connue. Il est possible que les ARNm forment un complexe d'initiation non canonique 48S* quand la voie normale d'initiation est inhibée ou que les ARNm et les protéines associées forment un complexe mais ne s'assemblent pas en complexe 48S. Il est également possible que les granules de stress contiennent un mélange entre des complexes 48S canoniques et d'autres pour lesquels l'association de l'ARNm avec le complexe 43S n'est pas effectuée (Figure IC.4) (Buchan and Parker, 2009; Muench et al., 2012). L'assemblage et le désassemblage des granules de stress sont, comme les *processing bodies*, dépendants des microtubules. En effet, une déficience du transport par une délétion de la chaine lourde ou légère de la kinésine ralentit leur dissolution tandis qu'une délétion de la dynéine et de la kinésine empêche leur formation (Loschi et al., 2009).

La formation des granules de stress est induite par plusieurs facteurs tels qu'une modification post-traductionnelle des protéines par phosphorylation, glycosylation ou encore par méthylation. Ces modifications post-traductionnelles permettent de réguler les interactions et/ou la fonction des protéines modifiées (Buchan and Parker, 2009; Thomas et al., 2011). La formation des granules de stress peut être dépendante de la phosphorylation de eIF2α à la suite de certains stress. En effet, la phosphorylation de eIF2α empêche la formation de complexe d'initiation de la traduction eIF2-GTP-ARNt^{met}, inhibant l'initiation de la traduction. Cette absence du complexe d'initiation de la traduction permet à TIA-1 et TIAR d'assembler le complexe d'initiation de la traduction déficient eIF2-eIF5 (Kedersha et al., 2002). La voie d'induction des granules de stress par phosphorylation de eIF2α est effective, par exemple, lors d'un choc thermique chez *S. cerevisiae* ou par le peroxyde d'hydrogène chez les mammifères (Dang et al., 2006; Emara et al., 2012;

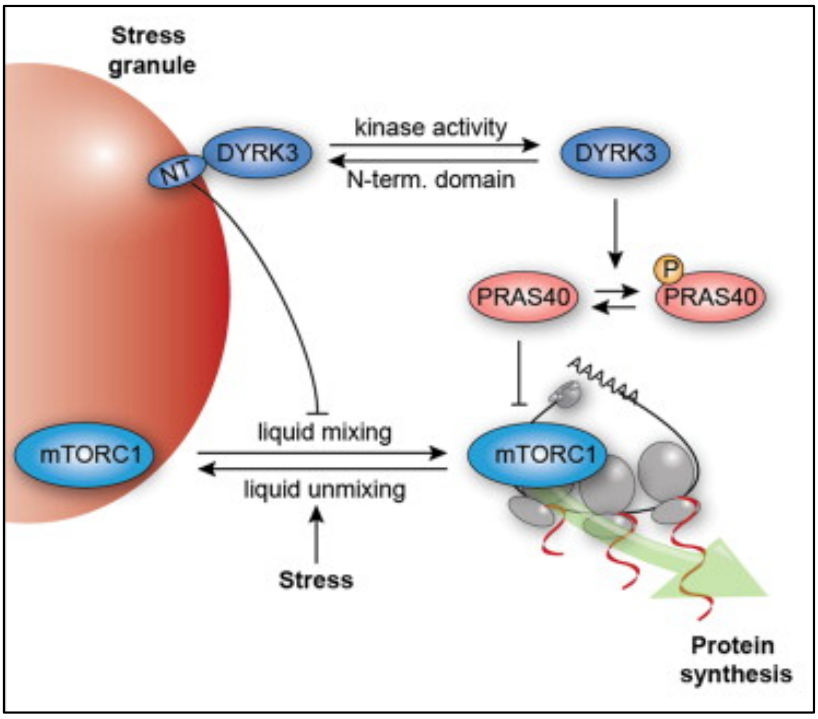


Figure IC.5 : DYRK3 régule la stabilité des granules de stress et l'expression de l'ARNm de TORC1

L'activation de DYRK3 dissocie les granules de stress et l'ARNm de TORC1 est traduit. En effet, DYRK3 phosphoryle l'inhibiteur de l'ARNm de TORC1 PRAS40, empêchant ainsi sa liaison à ARNm TORC1. Le domaine de faible complexité de DYRK3 est capable d'induire la formation de granules de stress lorsque l'activité kinase est inhibée. L'inhibition de l'activité kinase de DYRK3 influe sur la phosphorylation de PRAS40. PRAS40 inhibe donc la traduction de TORC1 et les granules de stress apparaissent, stockant l'ARNm TORC1. D'après (Wippich et al., 2013).

NT représente le domaine de faible complexité N terminal de DYRK3, P représente la phosphorylation de PRAS40, le flèches montre une activation et à l'inverse une flèche plate une inhibition.

Groušl et al., 2009; Thomas et al., 2011). Elle peut aussi permettre de réguler la taille des granules de stress chez la souris (Reineke et al., 2012). Cependant il existe des voies de formation des granules de stress indépendantes de la phosphorylation de eIF2α. En effet, l'inhibition du recrutement des ribosomes pour l'initiation de la traduction effectuée par une inhibition de l'activité de l'hélicase eIF4 induit la formation des granules de stress (Mazroui et al., 2006). Il existe d'autres voies de formation des granules de stress indépendamment de la phosphorylation de eIF2α (Thomas et al., 2011). L'une d'entre elle est effectuée suite à la phosphorylation par la phosphatidylinositol kinase DYRK3 (Wippich et al., 2013). DYRK3 est une kinase qui régule la stabilité des granules de stress et l'expression de l'ARNm de TORC1. En effet, DYRK3 phosphoryle l'inhibiteur de traduction de l'ARNm de TORC1, PRAS40, empêchant ainsi sa liaison à l'ARNm de TORC1. L'inhibition de l'activité kinase de DYRK3 influe la phosphorylation des protéines PRAS40 et inhibe la traduction de TORC1 et les granules de stress se forment, stockant les ARNm de TORC1 non transcrits (Figure IC.5). De plus, le mutant DYRK3 inactif ne peut se dissocier des granules de stress du à son domaine de faible complexité en N terminal (Wippich et al., 2013).

Comme pour les *processing bodies*, plusieurs domaines sont nécessaires à leur assemblement des protéines localisées dans les granules de stress (Tableau IC.3). Ces domaines sont principalement des domaines de liaison à l'ARN tels que les domaines ROQ, RGG, les domaines en doigt de zinc CCCH, des régions non structurées de faible complexité (Athanasopoulos et al., 2010; Chalupníková et al., 2008; Gareau et al., 2013; Goulet et al., 2008; Krüger et al., 2013; De Leeuw et al., 2007; Marko et al., 2012; Murata et al., 2005; Thomas et al., 2011; Tourrière et al., 2003; Wippich et al., 2013; Yang et al., 2006a). Une mutation ciblée du domaine de liaison à l'ARN prévient l'accumulation de CPEB1 au niveau des granules de stress, ce qui peut surement être étendu aux autres protéines possédant un domaine de liaison à l'ARN nécessaire à leur localisation au niveau des granules de stress (Wilczynska et al., 2005). Des régions riches en Q et en Q/N sont aussi nécessaires à la localisation de plusieurs protéines au niveau des granules de stress (Baguet et al., 2007; Colombrita et al., 2009; Gareau et al., 2013; Johnson et al., 2008; Solomon et al., 2007; Thomas et al., 2011; Vessey et al., 2006). Plusieurs protéines possèdent plusieurs fois le même domaine qui permet leur localisation dans les granules de stress comme pour TIS11 qui contient deux domaines doigts de zinc CCCH, IGF2BP1 et IGFBP2 (*insulin-like growth factor*-2) qui ont quatre domaines KH ou encore Vgl1 qui contient 14 domaines KH (Murata et al., 2005; Wächter et al., 2013; Wen et al., 2010). Certaines protéines possèdent des domaines différents

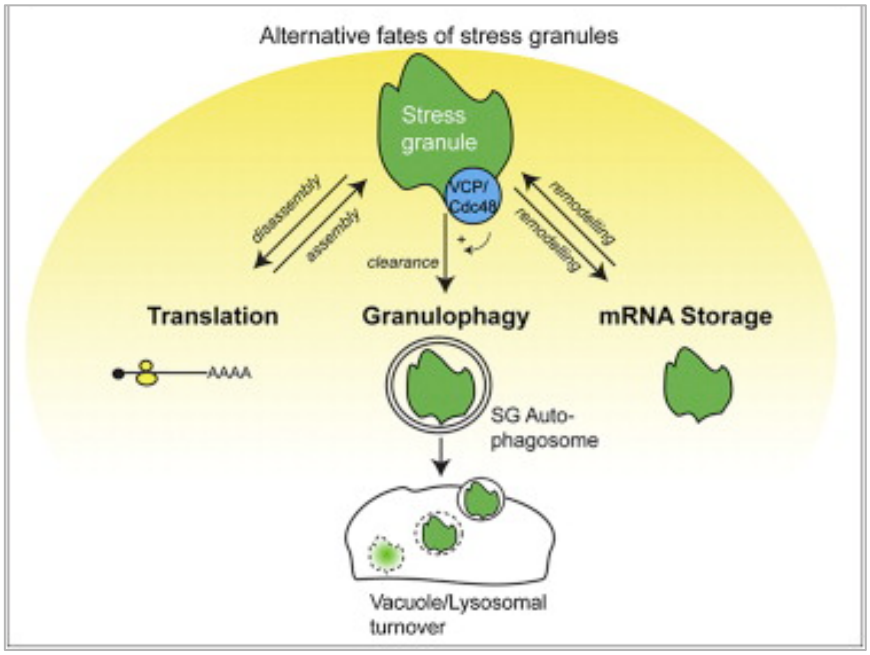


Figure IC.6 : Le « cycle de vie » des granules de stress

Les granules de stress s'assemblent lors d'un arrêt de la traduction et stockent les ARNm non-traduits ou, ils peuvent se désassembler et permettre la reprise de la traduction des ARNm. Lorsque l'AAA-ATPase VCP/Cdc48 s'associe aux granules de stress, les granules de stress sont désassemblés et ils peuvent être dégradés par autophagie. D'après (Buchan et al., 2013).

nécessaires à la localisation de la protéines dans les granules de stress qui ont des rôles complémentaires comme FMRP chez la mouche, qui possède les domaines KH, RGG et une région riche en Q/N (Gareau et al., 2013). D'autre part, ils peuvent avoir une complémentarité du rôle de ces différents domaines ou régions. Par exemple, chez les trypanosomes, la région LSM de SDC6/Rap55 est nécessaire à la formation de granules de stress tandis que le domaine RGG détermine le type et le nombre de granules de stress (Krüger et al., 2013).

Chez la levure, les granules de stress sont éliminés par autophagie. L'AAA-ATPase Cdc48/VCP, promouvant l'autophagie, sert à désassembler les granules de stress et à les diriger vers la granulophagie pour les éliminer. Plusieurs marqueurs des granules de stress sont retrouvés dans les vacuoles et seulement le marqueur des *processing bodies* Edc3. Il existe donc un remodelage des granules cytoplasmiques et un transfert des ARNm des *processing bodies* aux granules de stress avant leur élimination par autophagie. Lorsque la dégradation des ARNm est inhibée, l'autophagie augmente (Figure IC.6) (Buchan et al., 2013).

Chez les plantes, les granules de stress peuvent être formés en cas de choc thermique, lors d'un stress à l'arsénite, lumineux, ou hormonal tels que le méthyl jasmonate, l'acide abscissique, l'acide gibbérellique (Bogamuwa and Jang, 2013; Pomeranz et al., 2010; Weber et al., 2008). UBP1 et RBP47, possèdent un domaine semblable au prion riche en Q et trois domaines de liaison à l'ARN RRM. Ces derniers semblent responsables de l'agrégation de UBP1 et RBP47 au niveau des granules de stress (Weber et al., 2008).

3. Le rôle des granules de stress

Le rôle putatif des granules de stress est principalement le stockage des ARNm non-traduits lors d'un stress. En effet, ils sont nouvellement formés en cas de stress lors d'une inhibition de la traduction et ils peuvent se dissocier après reprise de la traduction. De plus, les principaux composant des granules de stress sont des facteurs d'initiation de la traduction, des protéines impliquées dans les modifications post-traductionnelle et des facteurs d'inhibition de la traduction, renforçant le rôle putatif de stockage des granules de stress. Cependant, leur formation n'est pas requise pour un ralentissement de la traduction (Buchan and Parker, 2009). Donc leur rôle ne se limite pas au stockage. En effet, les ARNm sont très peu de temps dans les granules de stress, environ une minute, alors que les granules de stress sont désassemblés après deux à trois heures. Les ARNm sont donc localisés beaucoup plus longtemps dans le cytoplasme que dans les granules de stress, ce qui va à l'encontre d'un rôle de stockage des granules de stress. Les ARNm

Tableau IC.3 : Les domaines nécessaires à la localisation des protéines au niveau des granules de stress

Domaines	Protéines	Références
riche en Q	MNL51, Pumilio 2, Caprin, TDP43	(Baguet et al., 2007; Colombrita et al., 2009; Johnson et al., 2008; Solomon et al., 2007; Vessey et al., 2006)
riche en Q/N	TIAR, TIA1, FMRP Dhh1/RCK/p54*	(Gareau et al., 2013; Thomas et al., 2011b)
Lsm	SDC6/Rap55	(Krüger et al., 2013)
ROQ, domaine en doigt de zinc RING-1 et CCCH	E3 ubiquitine ligase	(Athanasopoulos et al., 2010)
faible complexité	DYRK3	(Wippich et al., 2013)
RGG, domaine de liaison à l'ARN RRM et une région riche en RG	G3BP, SDC6/Rap55, CIRP, TAF15, FMRP	(Gareau et al., 2013; Krüger et al., 2013; De Leeuw et al., 2007; Marko et al., 2012; Tourrière et al., 2003; Yang et al., 2006)
KH (hnRNP-K Homologue)	IGF2BP1 et IGFBP2 (4 chacun), Vgl1, (14 dont 7 classiques), FMRP	(Gareau et al., 2013; Wächter et al., 2013; Wen et al., 2010)
en doigts de zinc CCCH	TIS11 (2)	(Murata et al., 2005)
c-Jun N terminal kinase (JNK)	MAPKBP1	(Cohen-Katsenelson et al., 2013)
de liaison à l'ARN	L ' hélicase RHAU, TDRD3, CPEB1	(Chalupníková et al., 2008; Goulet et al., 2008; Wilczynska et al., 2005)
Tudor	TDRD3	(Goulet et al., 2008)

* seulement chez les mammifères et l'homme

Lorsque les protéines contiennent plusieurs fois le même domaine, le nombre est indiqué entre parenthèses.

transitent juste par les granules de stress (Mollet et al., 2008). Les granules de stress comportent aussi des protéines reconnues pour être des régulateurs positifs (HUR) ou négatifs (G3BP, Rpp20, TTP, CPEB, Pumilio, SMG et Staufen) de la stabilité des ARNm. L'accumulation des facteurs de régulations négatives des ARNm soutient la possibilité que les granules de stress n'aient pas seulement une fonction d'entreposage mais aussi de dégradation. Cependant, la stabilisation des ARNm durant un stress n'est pas non plus altérée en absence de granules de stress (Buchan et al., 2008). Pourtant, en condition de stress à l'arsénite des facteurs de dégradation peuvent s'associer aux granules de stress. L'expression de CPEB1 induit la formation des granules de stress qui recrutent ensuite les facteurs des *processing bodies* (Wilczynska et al., 2005). Les granules de stress n'ont donc pas qu'un rôle de stockage mais aussi de triage, déterminant la conservation ou la dégradation des ARNm (Anderson and Kedersha, 2006, 2008). Il est aussi possible que les granules de stress soient impliqués dans un contrôle qualité des ARNm non traduits, par exemple en ajoutant une modification post-traductionnelle sur les protéines les recouvrant afin qu'ils restent intacts jusqu'à la reprise de la traduction (Mollet et al., 2008). Donc les rôles des granules de stress semblent complexes et ne sont pas encore complètement élucidés.

Les granules de stress sont principalement considérés comme un lieu de transition entre les polysomes et les *processing bodies* après une accumulation des transcrits non traduits en conséquence d'un stress (Balagopal and Parker, 2009). On estime que les ARNm destinés à retourner à la traduction, sont dirigés vers les polysomes et ceux destinés à être dégradés, vers les *processing bodies*. Plusieurs faits renforcent cette hypothèse. En effet, les granules de stress et les *processing bodies* partagent un groupe de protéines communes, ils peuvent être juxtaposés, suggérant que les deux types de granules cytoplasmiques peuvent interagir et que les ARNm peuvent transiter de l'un à l'autre (Anderson and Kedersha, 2008; Balagopal and Parker, 2009; Buchan and Parker, 2009; Buchan et al., 2008; Kedersha et al., 2005; Olszewska et al., 2012). Chez les plantes, les rôles putatifs de tri et de stockage des ARNm des granules de stress lors d'un stress semble être conservés (Muench et al., 2012; Nakaminami et al., 2012).

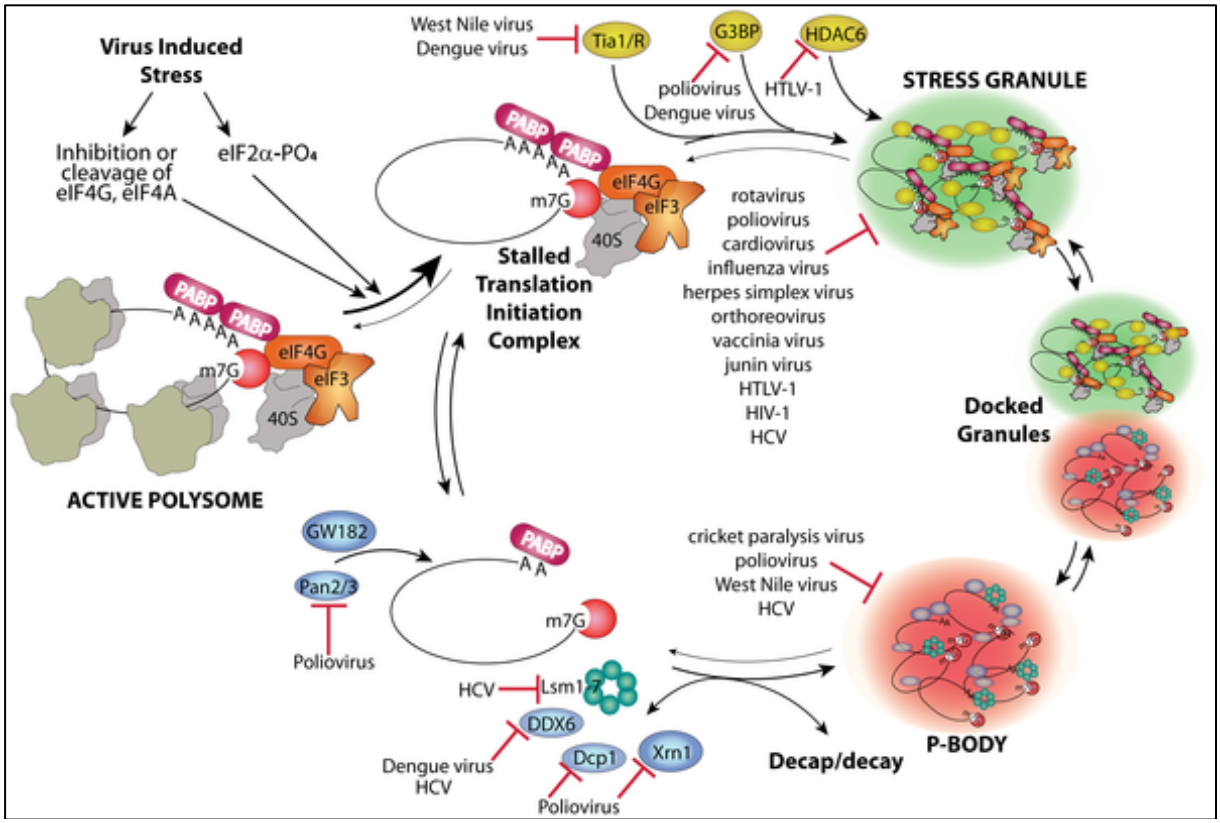


Figure IC.7 : L'assemblement des granules de stress et des *processing bodies* et leur interférence par les virus chez les mammifères.

L'infection par des virus induit un stress, réduisant la traduction des ARNm de l'hôte par l'activation des eIF2 kinases ou d'autres protéines impliquées dans les modifications post-traductionnelles, provoquant leur accumulation dans des foci cytoplasmiques tels que les *processing bodies* (P-body) et les granules de stress. Les points ou les protéines spécifiques cibles des virus qui activent (flèche) ou inhibent (flèche plate rouge) l'assemblement des foci cytoplasmiques sont indiqués. D'après (Lloyd, 2012).

3. La régulation de la traduction des ARNm et leur stockage dans les *processing bodies* et/ou les granules de stress lors d'une infection virale

Comme décrit précédemment, les ARNm non traduits peuvent s'accumuler dans des foci cytoplasmiques tels que les *processing bodies* et les granules de stress. L'infection par certains virus induit un stress, inhibant la traduction des ARNm et provoquant leur accumulation au niveau des *processing bodies* et des granules de stress. Les foci cytoplasmiques sont donc induits par les virus et ils ont un rôle antiviral en accumulant les ARN viraux et en inhibant la traduction des protéines des virus. Cependant, les virus peuvent aussi réprimer la formation des foci cytoplasmiques à un certain point de l'infection en interférant dans leur assemblage (Figure IC.7). De plus, ils peuvent utiliser les facteurs des foci cytoplasmiques pour leur réplication (Figure IC.8) (Beckham and Parker, 2008; Lloyd, 2012, 2013; Reineke and Lloyd, 2013; Valiente-Echeverría et al., 2012).

Les virus peuvent interférer dans l'assemblage des composants des *processing bodies* et les utiliser pour leur réplication (Figure IC.8A). Par exemple, Lsm1-7 est nécessaire à l'induction de la traduction des transcrits viraux et induit la formation des *processing bodies*. D'autre part, le brome mosaïque virus (BMV) des plantes est localisé dans les *processing bodies* afin de détourner leurs machinerie pour sa réplication et possiblement pour d'autres virus à ARN à brin positif (Beckham et al., 2007). Le processus de détournement des facteurs des *processing bodies* par les virus est courant. Les virus peuvent ainsi former des agrégats et/ou de nouveaux foci viraux nécessaires à leur réplication (Figure IC.8A) (Lloyd, 2013; Reineke and Lloyd, 2013).

Les virus peuvent aussi interférer dans l'assemblage des composants des granules de stress afin de former des foci viraux comportant des protéines issues des granules de stress (Figure IC.8B). Quelques exemples existent où les granules de stress et les virus coexistent lorsque le génome du virus est en cours d'expression. Cependant, lorsque les virus ont exprimés leurs gènes, les granules de stress sont absents et/ou ils ont été détournés pour former de nouveaux foci viraux symbiotiques avec le processus de réplication des virus. Les granules de stress ont donc un rôle antiviral. Par exemple, la PKR/ EIF2AK2 (*eukaryotic translation initiation factor 2-alpha kinase 2*) qui est un senseur central des stress induit la formation des granules de stress et est importante pour la lutte antivirale (Figure IC.8B) (Lloyd, 2012, 2013; Reineke and Lloyd, 2013).

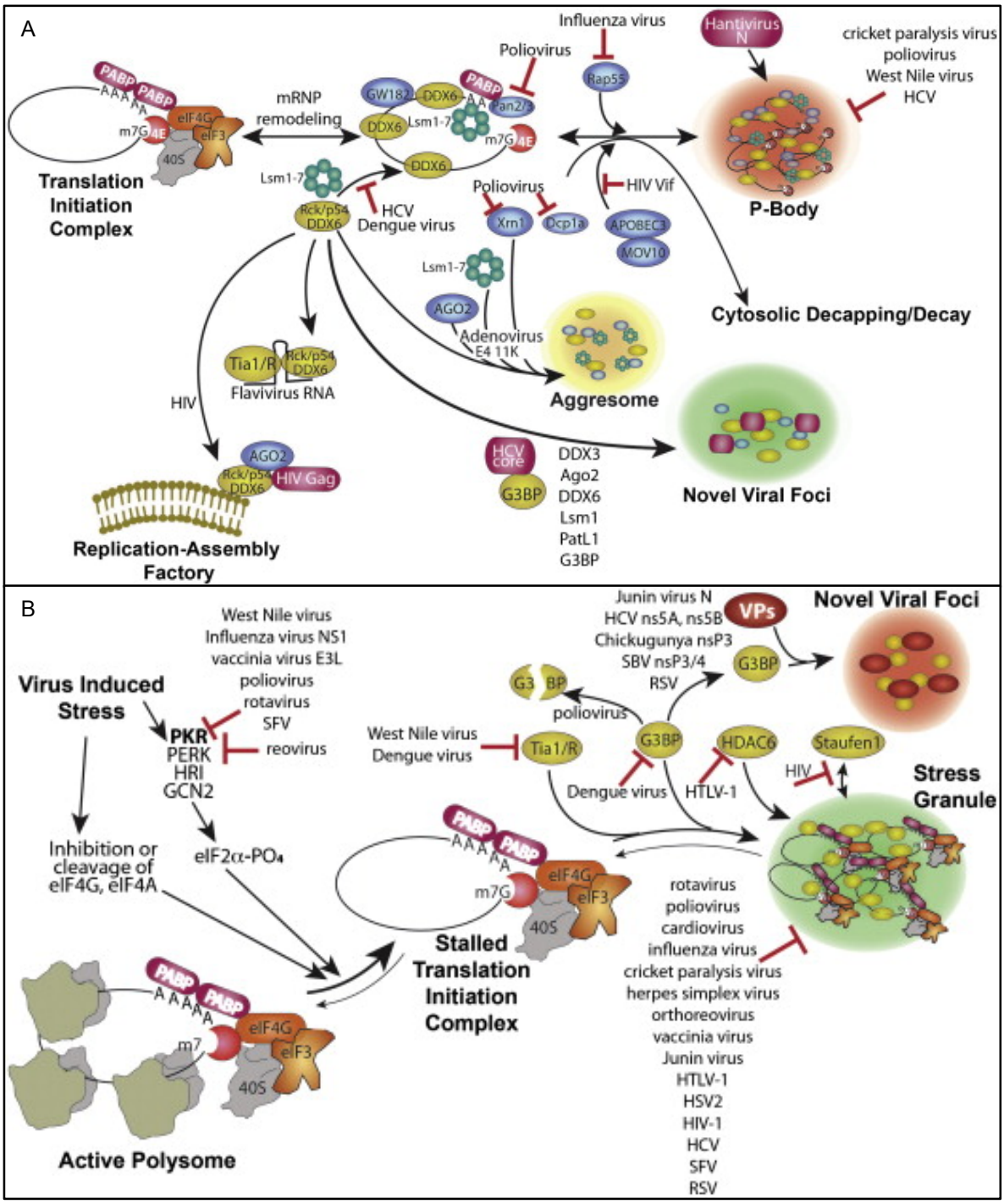


Figure IC.8 : Les composants des *processing bodies* et des granules de stress peuvent être détournés par les virus pour leur réplication

A L'assemblement des *processing bodies* et le détournement de leurs composants par les virus. Pour HCV, les nouveaux foci viraux contiennent des composants des *processing bodies* et quelques composants des granules de stress incluant G3BP. **B** L'assemblement des granules de stress et le détournement de leurs composants par des virus. Plusieurs protéines virales (VPs) interagissent en complexe avec G3BP1, certaines sont localisées dans de nouveaux foci viraux. Beaucoup de virus contrôlent l'activation de PKR, ici seuls quelques exemples sont indiqués. Les points où les protéines spécifiques cibles des virus qui activent (flèche) ou inhibent (flèche plate rouge) ou détournent l'assemblement des foci cytoplasmiques sont indiqués. D'après (Reineke and Lloyd, 2013).

4. La relation entre les polysomes, les granules de stress et les *processing bodies*

Le contrôle de la vitesse de la dégradation et de la transcription des ARNm est important pour la régulation de l'expression génétique chez les eucaryotes. Il est estimé que le taux de traduction et de dégradation des ARNm est influencé par un équilibre entre les polysomes et les foci cytoplasmiques tels que les *processing bodies* et les granules de stress.

Les *processing bodies* et les granules de stress partagent, chez les mammifères, un groupe de protéines communes tels que l'hélicase Dhh1/RCK/p54 et des protéines de répression de la traduction telles que AGO, TTP, BRF1 (Buchan and Parker, 2009). Les deux granules cytoplasmiques comportent aussi des protéines possédant des régions de faible complexité qui sont responsables de leur agrégation au niveau des *processing bodies* et des granules de stress telles que les région riches en Q/N (Gareau et al., 2013; Kato et al., 2012; Thomas et al., 2011; Weber and Brangwynne, 2012) (Tableaux 1C.2 et 1C.3). Les facteurs des granules de stress et des *processing bodies* peuvent naviguer entre les deux foci cytoplasmiques. Ils peuvent être juxtaposés allant jusqu'à fusionner lors d'une surexpression de CPEB1 (Kedersha et al., 2005; Wilczynska et al., 2005). Chez la levure, en absence de glucose, les *processing bodies* se forment en premier et ensuite les granules de stress. Sachant que les deux granules cytoplasmiques contiennent les mêmes ARNm, ils peuvent passer des *processing bodies* aux granules de stress (Buchan et al., 2008). A l'inverse, lors d'un stress thermique les ARNm sont stockés dans les granules de stress, qui se forment puis ils sont transférés dans les *processing bodies* (Kedersha et al., 2005). L'association des granules de stress et des *processing bodies* apparaît au début de la formation des granules de stress (Mollet et al., 2008; Wilczynska et al., 2005). Donc, soit les *processing bodies* permettent l'assemblement des granules de stress, même si ces derniers peuvent se former sans les *processing bodies*, soit, ces deux types de foci cytoplasmiques ont la même origine, même s'ils deviennent deux foci bien distincts (Mollet et al., 2008; Souquere et al., 2009).

Il existe donc une relation entre les polysomes, les granules de stress et les *processing bodies* formant ainsi un cycle dynamiques des ARNm (Figure IC.9). Les ARNm sont traduits dans les polysomes puis, lorsque la traduction est inhibée par un stress, les ARNm non traduits peuvent s'accumuler dans les *processing bodies* et/ou dans les granules de stress où ils peuvent être stockés et triés pour un retour vers la traduction au niveau des polysomes et/ou vers la dégradation. Les ARNm peuvent

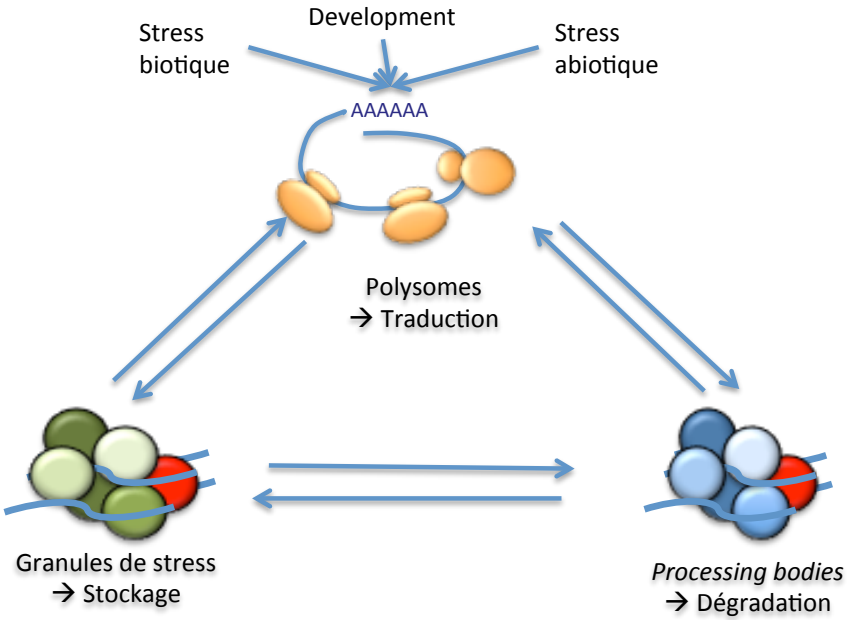


Figure IC.9 : Le cycle dynamique de l'ARNm entre les polysomes, les granules de stress et les *processing bodies*

Lors d'un arrêt de la traduction par un stress, les ARNm non traduits peuvent être stockés dans les granules de stress et/ou les *processing bodies.* Les ARNm peuvent être transférés des *processing bodies* aux granules de stress et *vice versa* et ils peuvent être réengagés dans la traduction.

être transférés d'un type de foci cytoplasmique à l'autre lorsqu'ils sont proches et être relocalisés dans les polysomes lorsque la traduction est reprise. Il existe donc une compétition entre ces trois structures pour les ARNm (Figure IC.9).

Objectifs de thèse

Notre équipe s'intéresse aux mécanismes de la dégradation des ARN chez les plantes et plus particulièrement au rôle de la polyadénylation et de l'uridylation dans de la dégradation des ARN. L'ajout d'uridines et d'adénosines en 3' des ARN est réalisé par des nucléotidyltransférases de la même famille, les polyA polymérases non-canoniques (ncPAP). Le génome d'*A. thaliana* code pour quinze ncPAP dont les localisations intracellulaires ont, pour la plupart, été précédemment déterminées dans l'équipe (Thèse de François Sement 2012). L'une de ces ncPAP, codée par le gène At2g45620, est une nucléotidyltransférase appartenant à la sous-famille des PolyU Polymérases (Kwak and Wickens, 2007). Nous l'avons nommée URT1 pour UTP:RNA uridylyltransférase 1. Précédemment à mon arrivée sur le projet, URT1 a été localisée dans le cytoplasme, associée à des foci cytoplasmiques, les granules de stress et les *processing bodies*. Ces deux types de foci cytoplasmiques sont impliqués dans la dégradation et/ou la régulation de la traduction des ARNm en réponse aux stress. Sachant que l'uridylation et ses mécanismes sont peu connus chez les plantes, mes objectifs de thèse étaient de caractériser URT1 et son activité catalytique, d'identifier ses substrats ARN et l'impact de l'uridylation de URT1 dans la plante, et de comprendre ses mécanismes d'adressage au niveau des granules de stress et des *processing bodies*.

Dans une première partie de ce manuscrit, je présente les caractéristiques de la séquence primaire de URT1 et de ses homologues chez les Angiospermes. Grâce à cette étude, j'ai découvert des motifs et des régions préservés chez les Angiospermes. Leur conservation au cours de la spéciation des Angiospermes suggère une conservation de leur fonction biologique. Par exemple, ils peuvent être impliqués dans la reconnaissance de substrat ARN de URT1, *via* une interaction protéine-protéine ou protéine-ARN ou encore être responsables de la localisation de URT1. Dans une seconde partie de manuscrit, j'ai déterminé les caractéristiques catalytiques intrinsèques *in vitro* de URT1 et ses spécificités de substrats (nucléotide et ARN). La prédiction de structure 3D de URT1 a été réalisée par homologie avec la structure d'autres uridylyltransférases et a contribué à mieux comprendre certaines particularités catalytiques de URT1. J'ai aussi participé à la découverte des substrats ARN de URT1 *in vivo* et du rôle de l'uridylation chez *A. thaliana*. Les données catalytiques de URT1 aussi permis de comprendre certaines particularités de l'uridylation des ARNm déadénylés *in vivo* par URT1. La découverte des substrats

ARN *in vivo* de URT1 a contribué à la découverte d'un nouveau rôle de protection en 3' de l'uridylation des ARNm chez *A. thaliana.* Dans une troisième et dernière partie, j'ai étudié les mécanismes d'adressage de URT1 au niveau des foci cytoplasmiques accumulant des transcrits non-traduits. Dans ce but, j'ai étudié l'impact des régions et/ou motifs conservés chez les Angiospermes sur la localisation de URT1. La région enrichie en P/Q/N/H, de la région N terminale de URT1, est essentielle à sa localisation au niveau des *processing bodies* et son domaine nucléotidyltransférase, quant à lui, semble suffisant pour sa localisation dans les *processing bodies* et les granules de stress après un stress thermique.

Ces travaux de thèse ont permis de mieux comprendre les mécanismes et les rôles de l'uridylation dans la dégradation des ARNm chez Arabidopsis. Ils ouvrent des perspectives dans l'étude d'autres fonctions de l'uridylation comme l'inhibition de la traduction.

Chapitre 1 :
URT1 est une nucléotidyltransférase conservée chez les Angiospermes possédant deux régions structuralement différentes

Résumé

Contexte scientifique

URT1 est une nucléotidyltransférase appartenant à la famille des Poly(A) Polymérases non-canoniques et plus particulièrement à la sous-famille des Poly(U) Polymérases (Kwak and Wickens, 2007). URT1 est une protéine d'une taille de 86 kDa (764 acides aminés) comprenant un domaine nucléotidyltransférase homologue à Cid1 de *Schizosaccharomyces pombe*. Cid1 est une uridylyltransférase impliquée dans l'uridylation des ARNm (Rissland and Norbury, 2009).

Questions biologiques

Quelles sont les particularités de la séquence de URT1 ?

Les régions caractéristiques de URT1 sont-elles conservées entre ses homologues chez les Angiospermes ?

Stratégie d'étude

Afin de caractériser URT1, j'ai réalisé une étude bioinformatique en trois étapes :

- une recherche approfondie des caractéristiques de la séquence primaire et secondaire
- une recherche de la préservation des caractéristiques mises en évidence chez URT1 lors de la première étape chez les Angiospermes
- une recherche de nouveaux motifs conservés entre les Angiospermes.

Principaux résultats

La région N terminale de URT1 est prédite intrinsèquement désorganisée. Sa taille est préservée chez les Angiospermes et elle comporte plusieurs motifs et régions conservés : des motifs prédits comme étant des zones de liaison à des partenaires protéiques et des régions enrichies en Q, en P/Q/N/H ou en E pouvant influencer la localisation de URT1.

Le domaine C terminal de URT1 contient un domaine nucléotidyltransférase et un domaine PAP/25A associé. Il est prédit pour avoir une structure globulaire et est similaire à 31% à Cid1 qui est connu pour uridyler les ARN messagers (Rissland and Norbury, 2009).

Conclusion/Ouverture

La présence de la région N terminale de URT1 et la préservation de motifs et de régions enrichies chez les plantes à fleurs suggèrent une conservation de leur fonction biologique lors de l'évolution, soit dans la localisation de URT1, soit dans la reconnaissance de partenaires protéiques ou de substrats ARN.

Les régions riches en Q et en Q/N se sont révélées responsables de la localisation de protéines au niveau des *processing bodies* et/ou des granules de stress (Introduction partie C et Chapitre 3) (Reijns et al., 2008; Thomas et al., 2011). D'autres parts, les domaines prédits comme intrinsèquement désorganisés sont connus pour être impliqués dans des interactions protéine-protéine et/ou protéine-ARN (Uversky and Dunker, 2013). Ces deux hypothèses de préservation du rôle des motifs et des régions conservées dans la localisation de URT1 et de sa liaison à un partenaire sont étudiées dans les chapitres suivants 2 et 3.

Le domaine nucléotidyltransférase de URT1 est similaire à 31% de Cid1. L'activité catalytique de Cid1 a été caractérisée et sa structure 3D est résolue au niveau atomique (Lunde et al., 2012; Munoz-Tello et al., 2012; Yates et al., 2012). La famille des ribonucléotidyltransférases est connue pour avoir un site actif très conservé (Introduction partie A4b) (Martin and Keller, 2007). Le domaine nucléotidyltransférase de URT1 a donc été modélisé par homologie de structure 3D à Cid1. Cette modélisation m'a permis d'étudier la conformation du site actif et ainsi de relier les propriétés biochimiques de URT1 (spécificité de substrat ARN et nucléotidique) à sa structure (Chapitre 2).

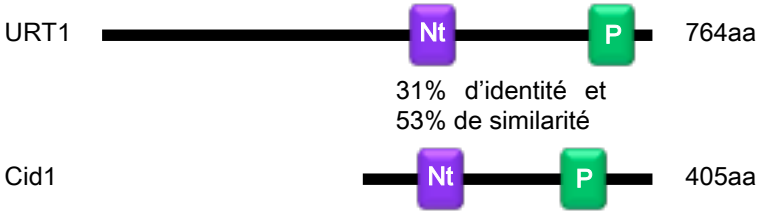


Figure 1.1 : URT1 possède deux domaines homologues à Cid1

Identification des domaines par InterProScan (Quevillon et al., 2005; Zdobnov and Apweiler, 2001), le domaine nucléotidyltransférase (Nt) : IPR002934 est représenté en violet (461-514) et le domaine PAP/25A associé (P) : IPR002058, en vert (656-714). Echelle 1cm = 100 acides aminés (aa).

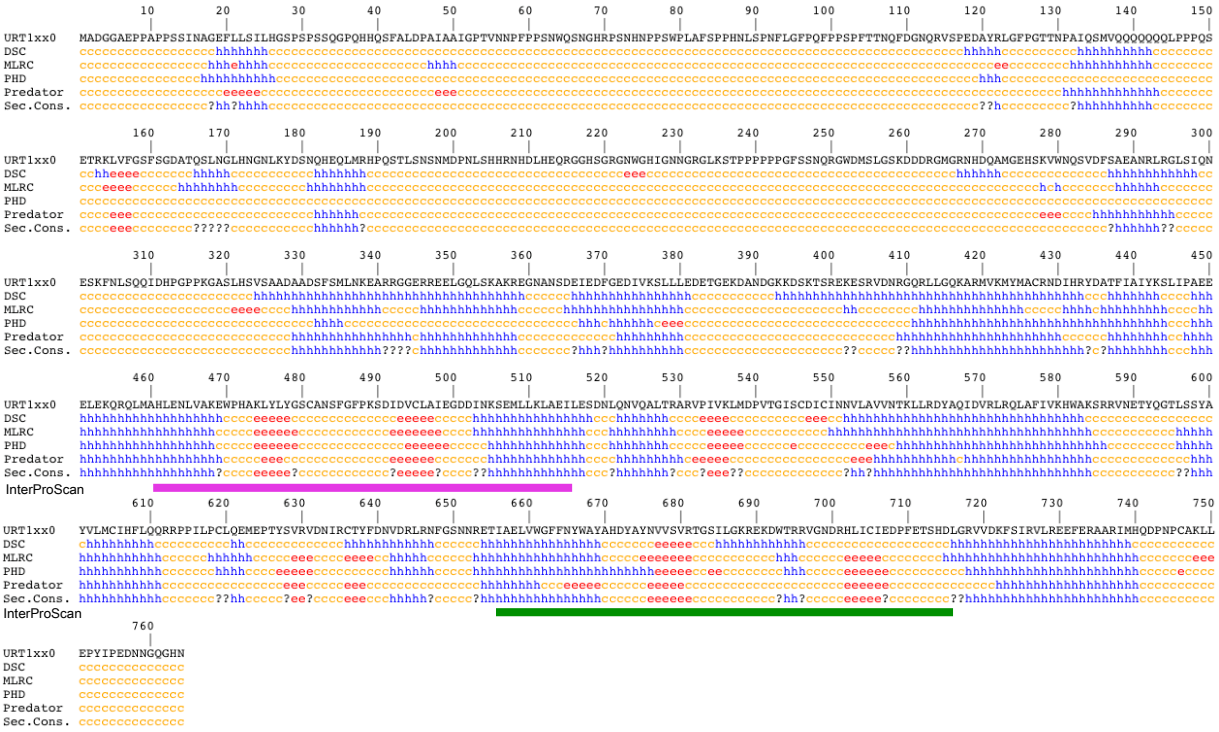


Figure 1.2 : Prédiction des structures secondaires de URT1

Prédiction des structures secondaires effectuée par différent logiciels : MLRC (Guermeur et al., 1999), DSC (King and Sternberg, 1996), PHD (Rost and Sander, 1993), PREDATOR (Frishman and Argos, 1996) grâce au site NPS@ («*Network Protein Sequence Analysis*» http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_seccons.html) (Combet et al., 2000). Les hélices α sont représentées en bleu, les feuillets β en rouge et les boucles en orange. En violet et en vert en dessous sont indiqués respectivement le domaine nucléotidyltransférase (461-514) et le domaine PAP/25A associé (656-714).

Figure 1.3: Représentation schématique des régions de faibles complexités de URT1

Détection des régions de faible complexité effectuée par le logiciel CAST (Promponas et al., 2000) à partir du site Annie v1.2, 2010 (http://annie.bii.a-star.edu.sg/annie/home.do) (Ooi et al., 2009). La région enrichie en Q est représentée en violet (des acides aminés 134 à 154), enrichie en N en bleue (des acides aminés 169 à 233), enrichie en P en vert (des acides aminés 238 à 243) et enrichie en E en rouge (des acides aminés 337 à 412).

1. URT1 est une nucléotidyltransférase composée de deux régions structuralement différentes

URT1 est codée par le locus atg45620 (séquence Annexe 1.1). La protéine URT1 possède un domaine nucléotidyltransférase (des acides aminés 461 à 514) et un domaine PAP/25A associé (des acides aminés 656 à 714) identifiés par InterProScan (Figure 1.1) (Quevillon et al., 2005; Zdobnov and Apweiler, 2001). La présence de ces domaines prédit que URT1 a une activité nucléotidyltransférase (GO : 0016779) (Ashburner et al., 2000; Carbon et al., 2009). Cette protéine de 86kDa peut être divisée en deux régions de taille similaire d'après les prédictions de structure secondaire (Figure 1.2). La première région située en N terminal, correspondant aux 419 premiers acides aminés (46kDa), contient très peu de prédiction de structure secondaire. La seconde région de URT1, située en C terminal, des acides aminés 420 à 764 (40kDa) est prédite contenant des structures secondaires. Ce dernier domaine est homologue à Cid1 avec 31% d'identité et 53% de similarité (Figure 1.1). Les régions N et C terminales de URT1 sont décrites plus précisément dans les paragraphes suivants.

2. La région N terminale de URT1 est conservée chez les Angiospermes

a) La région N terminale de URT1 est prédite intrinsèquement désorganisée

Plusieurs logiciels de prédiction de structure secondaire prédisent très peu de régions organisées en hélices α ou feuillets β dans la région N terminale de URT1 (Figure 1.2) (Combet et al., 2000; Frishman and Argos, 1996; Guermeur et al., 1999; King and Sternberg, 1996; Rost and Sander, 1993). Cette région possède plusieurs régions de faible complexité déterminées par l'algorithme CAST (Annie v1.2 © 2010, http://annie.bii.a-star.edu.sg/annie/home.do) (Kreil and Ouzounis, 2003; Ooi et al., 2009; Promponas et al., 2000). Ces régions sont enrichies en glutamine (Q) (des acides aminés 134 à 154), en asparagine (N) (des acides aminés 169 à 233), en proline (P) (des acides aminés 238 à 243) ou en acide glutamique (E) (des acides aminés 337 à 412) (Figure 1.3). Plusieurs acides aminés sont considérés comme promouvant la désorganisation (A, R, G, Q, S, P, E et K) et d'autres, à l'inverse, promouvant l'organisation (T, Y, F, I, L, V, C et N) (Campen et al., 2008; Williams et al., 2001). Le logiciel Composition Profiler (http://www.cprofiler.org/cgi-bin/profiler.cgi) permet l'analyse de la composition de la région N terminale de URT1 en détectant l'enrichissement ou l'appauvrissement de chaque acide aminé par comparaison au protéome d'*Arabidopsis thaliana* provenant de la base de donnée de TAIR

Tableau 1.1 : La région N terminale de URT1 est appauvrie en acides aminés promouvant l'ordre et principalement enrichie en acides aminés promouvant la désorganisation

Acides aminés	Lettre	Nombre/419	Pourcentage
Ala	(A)	25	6.0%
Arg	(R)	24	5.7%
Asn	(N)	34	8.1%
Asp	(D)	25	6.0%
Cys	(C)	0	0.0%
Gln	(Q)	30	7.2%
Glu	(E)	22	5.3%
Gly	(G)	41	9.8%
His	(H)	19	4.5%
Ile	(I)	10	1.4%
Leu	(L)	31	7.4%
Lys	(K)	17	4.1%
Met	(M)	8	1.9%
Phe	(F)	17	4.1%
Pro	(P)	40	9.5%
Ser	(S)	49	11.7%
Thr	(T)	11	1.6%
Trp	(W)	5	1.2%
Tyr	(Y)	2	0.5%
Val	(V)	9	1.1%

La composition en acides aminés de la région N terminale (des acides aminés 1 à 419) de URT1 est comparée au protéome d'*Arabidopsis thaliana* de TAIR (ftp://ftp.arabidopsis.org/home/tair/Proteins/TAIR10_protein_lists/) grâce au logiciel Composition Profiler (http://www.cprofiler.org/cgi-bin/profiler.cgi) (Vacic et al., 2007). Les acides aminés enrichis de URT1 sont mis en évidence en vert (p< 0,005) et ceux appauvris en rouge (p< 0,05).

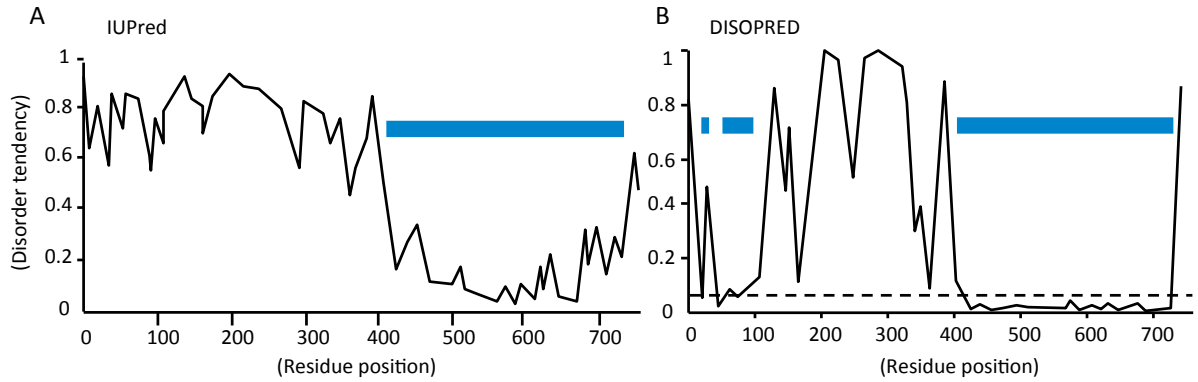


Figure 1.4 : La région N terminale de URT1 est prédite intrinsèquement désorganisée

A Prédiction de la désorganisation intrinsèque de la séquence primaire de URT1 par IUPred (http://iupred.enzim.hu) (Dosztanyi et al., 2005). **B** Prédiction de la désorganisation de URT1 par DISOPRED (http://bioinf.cs.ucl.ac.uk/disopred/disopred.html) (Buchan et al., 2010). Le principe de ces deux méthodes sont expliqués plus amplement dans le texte. La ligne discontinue met en évidence le seuil d'organisation avec DISOPRED et les barres bleues représentent les zones prédites structurées.

(ftp://ftp.arabidopsis.org/home/tair/Proteins/TAIR10_protein_lists/) (Vacic et al., 2007). La région N terminale de URT1 est donc enrichie en acides aminés N, Q, G, H, et P (Tableau 1.1). Ces acides aminés sont principalement considérés promouvant la désorganisation G, Q et P, représentant respectivement 9,7%, 7,1% et 9,5%. Un maximum de 15% de proline est atteint dans les 250 premiers acides aminés. La région N terminale est, de plus, appauvrie en acides aminés promouvant l'organisation tels que T (2,6%), Y (0,5%), I (2,4%), V (2,1%) et C (0%) (Tableau 1.1).

Les 411 premiers acides aminés de la région N terminale de URT1 sont prédits comme étant intrinsèquement désorganisée par le logiciel de prédiction de désorganisation structurale intrinsèque IUPred (http://iupred.enzim.hu) qui s'appuie sur les propriétés biophysiques des acides aminés de la séquence protéique (Dosztanyi et al., 2005) (Figure 1.4A). IUPred prédit, en plus de la région N terminale, une désorganisation des derniers acides aminés de URT1 (des acides aminés 743 à 764).

Dans un second temps, j'ai utilisé un second logiciel de prédiction de la désorganisation structurale intrinsèque des protéines : DISOPRED. DISOPRED est basé sur les résidus non identifiés sur les cartes de densité issues de l'analyse de plus de 750 protéines cristallisées dont la structure a été définie à haute résolution par diffraction de rayons X (http://bioinf.cs.ucl.ac.uk/disopred/disopred.html) (Buchan et al., 2010). Il confirme la désorganisation structurale intrinsèque de la région N terminale, principalement des acides aminés 100 à 410 ainsi que la fin de la séquence de URT1, des acides aminés 754 à 764, déjà observée par IUPred (Figure 1.4B). DISOPRED met aussi en évidence plusieurs prédictions de régions structurées des acides aminés 16 à 26 et 42 à 96. Celles-ci sont très intéressantes car beaucoup de régions intrinsèquement désorganisées sont connues pour être impliquées dans des interactions protéine-protéine ou protéine-ARN par des segments ayant la capacité de s'ordonner lors de ces interactions (Dunker et al., 2001; Uversky, 2013a, 2013b; Uversky and Dunker, 2013).

b) Des zones de liaison à d'autres protéines sont prédites au niveau de la région N terminale de URT1

Les régions désorganisées peuvent former des liaisons avec des protéines et/ou avec des ARN (Dunker et al., 2001; Uversky, 2013a, 2013b; Uversky and Dunker, 2013). Afin de déterminer des zones potentielles de liaison à l'ARN qui sont chargées positivement, j'ai utilisé une méthode d'analyse statistique SAPS (*Statistical Analysis of Protein Sequences*) (Brendel et al., 1992). Cependant aucune région chargée positivement n'est mise en évidence (Annie v1.2 © 2010,

Figure 1.5 : La région N terminale de URT1 contient des zones prédites d'interaction protéine-protéine

Les prédictions des zones d'interaction protéine-protéine dans les régions désorganisées par le logiciel ANCHOR (http://anchor.enzim.hu/) (Dosztányi et al., 2009; Mészáros et al., 2009) sont représentées en bleu. ANCHOR utilise les données prédiction de désorganisation d'IUPred présentées dans la Figure 1.4.

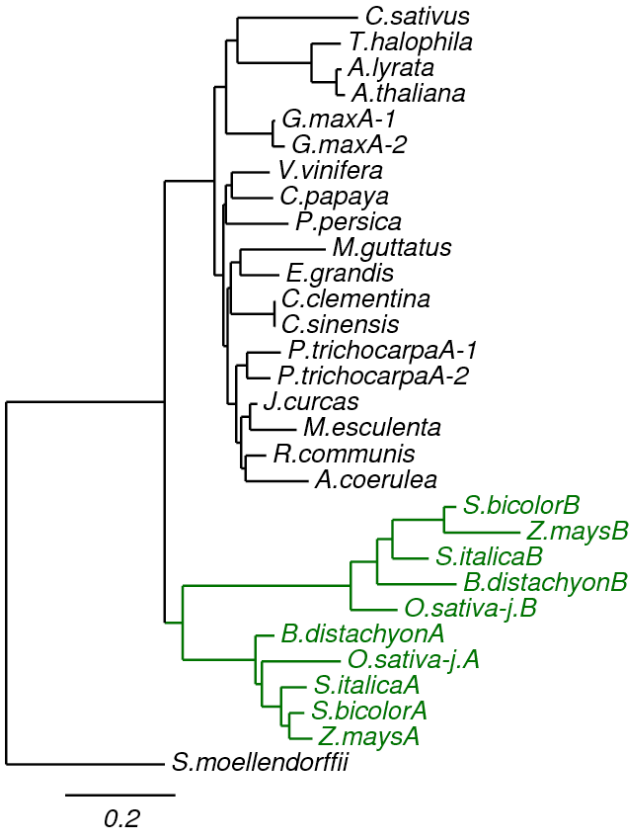


Figure 1.6: Les monocotylédones possèdent deux homologues de URT1 (URT1A et URT1B)

L'arbre phylogénique des homologues de URT1 chez les Angiospermes est construit à partir d'un alignement effectué par MUSCLE (Edgar, 2004), suivi d'une élimination des régions divergentes par GBLOCKS (Castresana, 2000; Talavera and Castresana, 2007) et la construction de l'arbre phylogénétique par PHYML (http://www.phylogeny.fr) (Dereeper et al., 2008). Les dicotylédones sont représentées en noir et les monocotylédones en vert et l'homologue de URT1 chez *Selaginella moellendorffii*, de la famille des Trachéophytes créé la racine de l'arbre phylogénique.

http://annie.bii.a-star.edu.sg/annie/home.do) (Ooi et al., 2009). Au contraire, la région N terminale comporte une région composée d'une charge mixte de 17 charges positives et 21 charges négatives, des acides aminés 338 à 409. Cette région est située au niveau de la région enrichie en E (Figure 1.3). Elle contient aussi une proportion importante des acides aminés K, R et D avec 11% de chaque. Seul E est enrichit selon le logiciel Composition Profiler ($p<0.0001$). Une longue région non chargée est aussi présente des acides aminés 8 à 109. Cette dernière région est riche en sérine (14,7%) et en proline (22,5%). Elle contient les deux régions prédites structurées par DISOPRED des acides aminés 16 à 26 et 42 à 92 (Figure 1.4).

Afin de prédire des zones d'interaction protéine-protéine, j'ai utilisé un logiciel complémentaire d'IUPred nommé ANCHOR (http://anchor.enzim.hu). ANCHOR se base sur la prédiction de la transition d'une longue région désorganisée, issue des données d'IUPred, à des segments organisés. Afin de prédire des zones d'interaction protéine-protéine, les segments ne doivent pas être capables de former d'interaction sur eux-mêmes et, au contraire, ils doivent être stabilisés par une interaction avec une protéine globulaire (Dosztányi et al., 2009; Mészáros et al., 2009). Plusieurs zones d'interaction protéine-protéine sont prédites au niveau de la région N terminale de URT1 (Figure 1.5). Deux zones d'interaction protéine-protéine découvertes par le logiciel de prédiction de la désorganisation intrinsèque DISOPRED correspondent aux segments organisés de la région N terminale des acides aminés 16 à 26 et 42 à 96 (Figure 1.4). Le logiciel ANCHOR prédit beaucoup d'autres zones d'interaction protéine-protéine. Cependant, en compilant les données obtenues par DISOPRED et ANCHOR, ces deux segments des acides aminés 16 à 26 et 42 à 92, semblent être les zones d'interaction protéine-protéine les plus plausibles. En effet, ces deux régions correspondent à deux motifs conservés chez les Angiospermes (paragraphe suivant et Discussion partie B).

c) La région N terminale est conservée chez les Angiospermes

Des homologues de Cid1 sont présents chez plusieurs organismes en dehors du règne végétal (Kwak and Wickens, 2007). Cependant aucun d'entre eux ne possède une région N terminale semblable à URT1. Afin de déterminer si les caractéristiques de la région N terminale de URT1 sont conservées chez les plantes, j'ai analysé les homologues de URT1 après leur identification grâce à Phytozome (http://www.phytozome.net) (Goodstein et al., 2011). Phytozome est une base de donnée regroupant les données génomiques végétales. Elle permet ainsi de déterminer les familles d'homologues intra ou inter espèce(s). Des homologues de URT1 sont présents chez toutes les Angiospermes (alignement présenté en Annexe

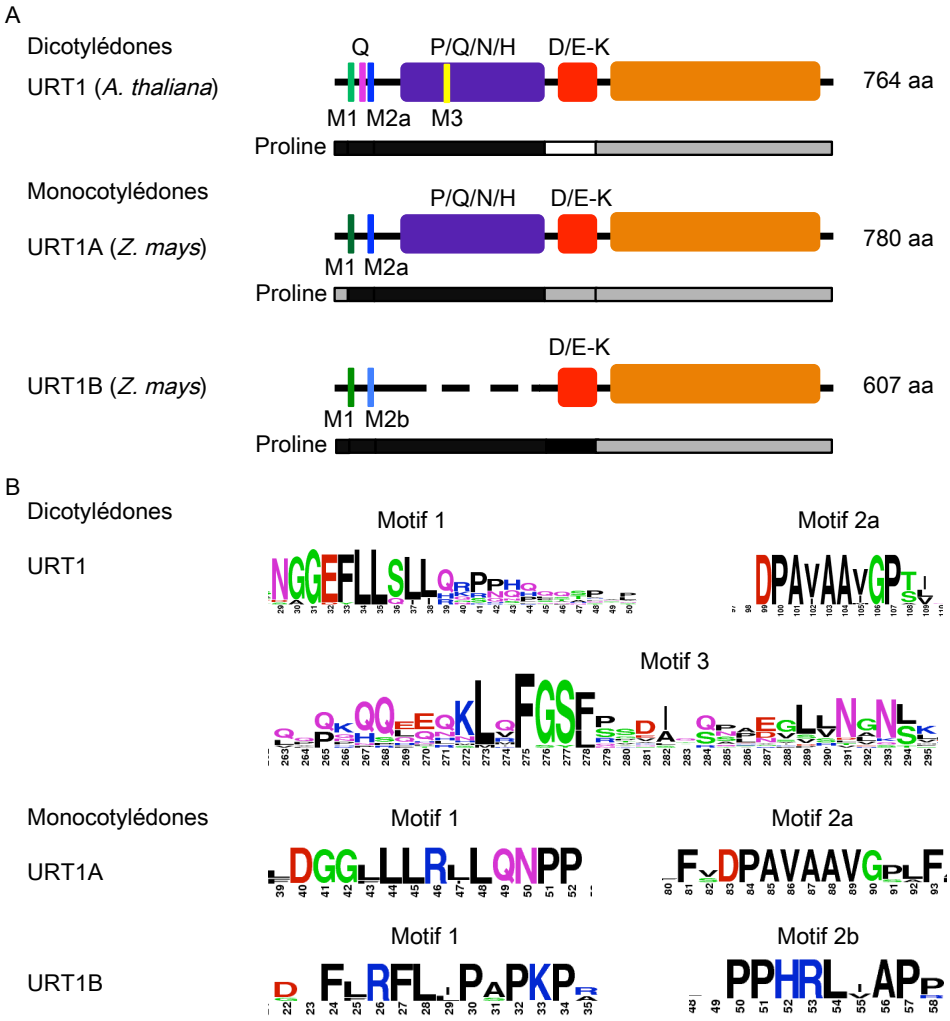


Figure 1.7: La région N terminale de URT1 contient des régions et des motifs conservées chez les Angiospermes

A Les motifs et les régions conservés chez les Angiospermes selon les trois différents clades : dicotylédones et URT1A et URT1B des monocotylédones (de *Z. mays*) représentés sur la séquence de URT1. Les motifs M1 leucine sont représenté en vert, M2a et M2b sont représentés en bleu et le motif M3 présent seulement chez les dicotylédones (des acides aminés 154 à 160 chez *A. thaliana*) en jaune. La région riche en Q est en rose des acides aminés 26 à 45 chez *A. thaliana*, celle enrichie en P/Q/N/H (des acides aminés 166 à 309 chez *A. thaliana*) en violet et celle riche en E en rouge (des acides aminés 339 à 403 chez *A. thaliana*). Le domaine nucléotidyltransférase est coloré en orange. La barre en dessous représente la proportion de proline. En noire est représenté les zones enrichie en proline ($p<0,05$), les zones grise n'étant pas statistiquement différent et en blanc la zone appauvrie en proline ($p<0,05$) à l'aide de Composition profiler en comparaison avec le protéome *d'A. thaliana* ou de *Z. mays,* respectivement issus de la base de donnée TAIR (ftp://ftp.arabidopsis.org/home/tair/Proteins/TAIR10_protein_lists/) et de Phytozome (ZmB73_5a_WGS_translations.fasta) (Goodstein et al., 2011; Vacic et al., 2007). **B** Représentation des motifs conservés en vert et jaune dans A par Weblogo (http://weblogo.berkeley.edu/logo.cgi) (Crooks et al., 2004; Schneider and Stephens, 1990) effectués à partir de l'alignement des homologues de URT1 des Angiospermes selon les trois différents clades : dicotylédones et URT1A et URT1B des monocotylédones.

1.2). Il existe d'autres homologues de URT1 qui gardent une région intrinsèquement désorganisée comme par exemple chez le champignon filamenteux *Aspergillus nidulans* (Morozov et al., 2010) ou chez la plante de la famille des Trachéophytes *Selaginella moellendorffii* (voir Annexe 1.3). Cependant, ces homologues ne possèdent pas de motifs aussi conservés que ceux retrouvés chez les Angiospermes. Une utilisation des homologues de URT1 de la famille des Angiospermes est donc privilégiée dans notre analyse. Leur alignement met en avant des motifs conservés pouvant être responsables d'une interaction avec un partenaire ou de régions importantes pour la fonction biologique de URT1 préservés au cours de l'évolution. En effet, les zones d'interaction avec un partenaire divergent moins que les autres régions car les zones d'interaction avec un partenaire possèdent une pression de sélection plus forte que les autres (Mészáros et al., 2009, 2011, 2012; Schlessinger et al., 2011, 2011; Uversky and Dunker, 2013).

Afin d'analyser l'évolution des homologues de URT1, j'ai effectué un arbre phylogénétique en éliminant les régions divergentes par GBLOCKS (http://www.phylogeny.fr) (Castresana, 2000; Talavera and Castresana, 2007) (Figure 1.6). Les dicotylédones possèdent généralement une seule copie de URT1, exceptée pour *Glycine max* et *Populus trichocarpa* qui en ont deux, due à une récente duplication génomique (Schmutz et al., 2010; Tuskan et al., 2006). Les monocotylédones, quant à elles, possèdent principalement deux homologues de URT1 : URT1A et URT1B, indiquant une duplication antérieure à la spéciation chez les monocotylédones. Le clade A des monocotylédones est plus proche génétiquement des dicotylédones que le clade B. De plus, les protéines du clade A sont généralement de taille similaire à URT1 d'*A. thaliana* avec une taille minimale et maximale de 780 acides aminés chez *Zea mays* et une taille maximale de 828 acides aminés chez *Oryza sativa japonica*. Les protéines du clade de URT1B, quant à elles, sont d'une taille inférieure à *A. thaliana*, de 584 à 647 acides aminés respectivement chez *Setaria italica* et *Sorghum bicolor*.

La région N terminale des homologues de URT1 chez les Angiospermes conserve des caractéristiques intéressantes. En effet, la taille de cette région N terminale est conservée, environ 400 à 500 acides aminés. La région N terminale est prédite intrinsèquement désorganisée avec des zones de liaison à d'autres protéines (Prédiction par IUPred et ANCHOR en Annexe 1.3). Même si la séquence primaire de la région N terminale est divergente entre les Angiospermes, deux motifs d'une dizaine d'acides aminés sont très conservés, nommés ici M1 et M2 et des régions particulières enrichies en proline, en Q, en P/Q/N/H et en E, appelée E/D-K (Figure 1.7).

Les motifs M1 et M2 conservés sont situés au début de la région N terminale et sont flanqués de régions riches en proline. La séquence des protéines URT1 débute par une région (de 12 à 40 acides aminés) riche en proline et en glycine. Elle précède le premier motif conservé nommé M1 correspondant à un motif leucine. La séquence consensus est NGGEFLL(S/Q)(L/I)(L/I) chez les dicotylédones, située au niveau des acides aminés 16 à 25 chez *A. thaliana*, DGG(L/F)LLR(L/I)L pour URT1A des monocotylédones et DxF(L/F)RFL(I/L) pour URT1B (Figure 1.7, Alignement en Annexe 1.1). Entre le motif M1 et M2, se situe une autre région riche en proline et glutamine, située des acides aminés 26 à 45 chez *A. thaliana*. Cette région riche en glutamine est conservée chez les dicotylédones, allant jusqu'à 12 Q successifs chez *Manihot esculenta*. La séquence des homologues de URT1 est suivie par le motif M2a, DPA(V/I)AA(V/I)GP qui est conservé chez les dicotylédones et URT1A des monocotylédones. Chez URT1B des monocotylédones, ce motif est remplacé par la séquence consensus M2b PPHRL(V/I)AP. Chez les Angiospermes, les régions suivantes sont riches en proline. Chez les dicotylédones, une région riche en proline et asparagine est présente des acides aminés 54 à 120, respectivement de 22 et 15% chez *A. thaliana*. La région suivante des acides aminés 120 à 153 est enrichie en Q/P, respectivement de 27 et 15% chez *A. thaliana*. Cette région contient 7Q successives de 138 à 144 chez URT1 de *A. thaliana*. La région riche en Q/N (des acides aminés 54 à 153) est conservée chez les dicotylédones et les monocotylédones du clade A. Cependant, chez les monocotylédones du clade B, celle-ci est absente. Chez les dicotylédones, une autre séquence consensus, le motif M3, est présente, KLxFGS(F/L), des acides aminés 154 à 161 chez *A. thaliana*. Puis, les séquences des homologues de URT1 chez les Angiospermes possèdent une région riche en asparagine et en histidine des acides aminés 166 à 230 chez *A. thaliana*, respectivement de 17 et 14%. Celle-ci peut être étendue jusqu'aux acides aminés 309 avec un taux de Q/N de 20%. La région des acides aminés 54 à 309 de URT1 chez *A. thaliana*, forme la région enrichie en P/Q/N/H ($p<0,001$ d'après Composition Profiler) (voir en complément le chapitre 3). Cette région est poursuivie par la région, enrichie en acide glutamique précédemment mise en évidence lors de l'analyse de la séquence primaire de URT1 de *A. thaliana* (Chapitre 2a). Elle est conservée chez les Angiospermes, 20% de E chez *A. thaliana* des acides aminés 339 à 403. Par exemple, URT1A de *Z. mays* possède 16% de E, des acides aminés 415 à 473 ($p<0,0001$ d'après Composition Profiler). La suite de la région N terminale des Angiospermes ne possède plus de motifs conservés et est suivie par le domaine nucléotidyltransférase. Tous les motifs et les régions conservées chez les Angiospermes sont représentés dans la Figure 1.7.

Cette analyse montre que la séquence primaire de la région N terminale de URT1 a rapidement évolué durant la spéciation des plantes. Cependant, elle préserve certaines caractéristiques comme la propriété d'être intrinsèquement désorganisée, des motifs et des régions préservées chez les plantes à fleurs. Ceux-ci pourraient refléter une conservation de leur fonction biologique lors de l'évolution, soit dans la localisation de la protéine soit dans l'interaction avec un partenaire (Chapitre 3 et Discussion partie B). Les zones d'interactions prédites par ANCHOR et les deux régions prédites organisées par DISOPRED des acides aminés 16 à 26 et 42 à 96 chez URT1, correspondent aux motifs M1 et M2 (Figure 1.4). Les régions d'interactions avec un partenaire évoluent beaucoup plus lentement que les autres régions qui ont une pression sélective moins importante et donc ont tendance à diverger plus rapidement (Discussion partie B).

3. Le domaine C terminal de URT1 est similaire à Cid1

Le domaine C terminal de URT1, des acides aminés 420 à 764, possède deux domaines connus : le domaine nucléotidyltransférase et le domaine PAP/25A associé (Figure 1.1). Ces deux domaines sont suffisants pour prédire que URT1 appartient à la famille des Poly(A) Polymérase de la famille de CID (E-value de $4{,}4e^{-118}$) par PANTHER HMM (*Protein ANalysis THrough Evolutionary Relationships*) (http://www.pantherdb.org/) (Mi et al., 2010, 2013). Le domaine C terminal de URT1 est un homologue de Cid1 avec 31% d'identité et 53% de similarité (Figure 1.1), Cid1 étant connu pour uridyler les ARN messagers chez *S. pombe* (Rissland and Norbury, 2009). Le domaine C terminal de URT1 est prédit pour avoir des structures secondaires par IUPred et DISOPRED (Figure 1.4) et est prédit globulaire par GlobPlot2 des acides aminés 323 à 762 (http://globplot.embl.de/) (Linding, 2003).

4. Conclusion/Ouverture

URT1 est une protéine conservée chez les Angiospermes. La préservation de la région N terminale de URT1, intrinsèquement désorganisée, de ses motifs (M1, M2 et M3) et de ses régions (enrichies en Q, en P/Q/N/H et en E) chez les plantes à fleurs est surement responsable de la conservation de leurs fonctions biologiques lors de l'évolution, soit dans la localisation de la protéine soit dans l'interaction avec un partenaire (Chapitre 2 et 3 et Discussion partie B). En effet, l'importance des régions riches en Q et/ou en Q/N dans la localisation des protéines au niveau des *processing bodies* et des granules de stress a été précédemment mis en évidence (Reijns et al., 2008; Thomas et al., 2011). D'autre part, les motifs conservés sont

souvent impliqués dans des interactions protéine-protéine et/ou protéine-ARN, tout comme les régions riches en proline intrinsèquement désorganisée (Discussion partie B) (Dosztányi et al., 2009; Mészáros et al., 2009, 2012; Schlessinger et al., 2011; Uversky, 2013b; Uversky and Dunker, 2013).

Le domaine C terminal est un homologue de Cid1 avec 31% d'identité et 53% de similarité. La structure tridimensionnelle (3D) de Cid1 a été récemment résolue au niveau atomique (Lunde et al., 2012; Munoz-Tello et al., 2012; Yates et al., 2012). Cid1 possède une structure 3D typique de la famille des ribonucléotidyltransférases (Martin and Keller, 2007). La structure 3D du domaine nucléotidyltransférase de URT1 peut être modélisé par homologie de structure avec Cid1. La modélisation du site actif de URT1 et sa comparaison à d'autres protéines modèles, uridylyltransférase de la famille des ncPAP a servi à mieux comprendre et expliquer certaines des propriétés catalytiques de URT1 présentées dans le chapitre 2. Cependant puisqu'il est intéressant de résoudre la structure atomique du domaine nucléotidyltransférase de URT1, une collaboration a été commencée avec le laboratoire du Dr Meinhart du Max-Planck-Institute of Medical research d'Heidelberg en Allemagne.

Chapitre 2 :
Activité catalytique de URT1, une nucléotidyltransférase responsable de l'uridylation des ARNm oligoadénylés chez *Arabidopsis thaliana*

Résumé

Contexte scientifique

URT1 a été précédemment identifiée comme une Poly(U) Polymérases (PUP) appartenant à la famille de Cid1 (Kwak and Wickens, 2007). Cid1 est connue pour uridyler les ARN messagers et induire leur dégradation dans le sens 5'-3' (Rissland and Norbury, 2009). Les expériences permettant de déterminer URT1 comme une PUP ont été réalisées par micro-injection de l'ARNm de URT1 fusionnée à la séquence codant pour la protéine de la coque du phage MS2 dans des oocytes de *Xenopus laevis*. Cette fusion permet de diriger l'action de URT1 sur un ARN cible contenant une structure tige-boucle reconnue par la protéine virale (Bernardi and Spahr, 1972). Aucune étude catalytique complémentaire *in vitro* ou *in vivo* n'a été effectuée par Kwak et Wickens. Le type de substrats ARN de URT1 n'est donc pas identifié dans ces expériences car l'action de URT1 est dirigée sur un ARN cible et testée en système hétérologue.

Questions biologiques

Quels sont les propriétés catalytiques de URT1 (spécificité de substrat ARN et nucléotidique) ?

Quel est le rôle de l'uridylation des ARN *in vivo* effectuée par URT1 ?

Stratégie d'étude

Afin de mettre en évidence les caractéristiques de l'activité catalytique de URT1, plusieurs stratégies ont été utilisées :

- une étude de l'activité catalytique de la protéine recombinante URT1 in vitro (spécificité nucléotide et ARN),
- une recherche des substrats ARN in vivo grâce à des mutants d'insertion de T-DNA, abolissant l'expression de URT1
- une modélisation du domaine nucléotidyltransférase de URT1 par homologie de structure 3D avec Cid1 afin de mieux comprendre les propriétés catalytiques de URT1

Principaux résultats et ouverture

URT1 est une UTP:RNA uridylyltransférase intrinsèquement spécifique des UTP et distributive au niveau des premiers nucléotides ajoutés.

In vivo URT1 ajoute une à trois uridines aux ARNm ayant subi une étape de déadénylation. Nos résultats montrent que l'uridylation protège l'extrémité 3' des ARNm prévenant ainsi la formation d'ARNm tronqués.

Les transcrits uridylés sont des substrats de la dégradation dans le sens 5'-3' par XRN4. L'uridylation par URT1 des ARNm chez *A. thaliana* contribue à établir la polarité de dégradation de 5' en 3', établissant un nouveau rôle pour l'uridylation des ARNm. Cette polarité de dégradation dans le sens 5'-3' est importante en cas de dégradation co-traductionnelle car l'effet protecteur de l'uridylation est observé au niveau des polysomes.

In vitro, URT1 peut uridyler les ARN ayant le dernier nucléotide libre et non modifié en 2'OH car ce nucléotide apparaît crucial à la stabilisation de l'ARN dans le site actif. URT1 peut uridyler *in vitro* tous les substrats ARN testés, y compris les ARN ribosomiques et les ARN de transfert. La spécificité de URT1 pour les ARNm oligoadénylés *in vivo* pourrait être conférée par un partenaire protéique (Discussion partie B).

D'autre part, dans l'article publié, URT1 est localisée au niveau des granules ribonucléoprotéiques, les *processing bodies* et les granules de stress. Ces deux foci cytoplasmiques sont impliquées dans la régulation de la traduction et/ou de la dégradation des ARN en réponse aux stress (Introduction partie C et Discussion partie D5). Les mécanismes d'adressage de URT1 au niveau de ces structures sont étudiés plus précisément dans le chapitre 3.

A - URT1 est responsable de l'uridylation des ARNm oligoadénylés et prévient la formation de transcrits tronqués en 3' chez *Arabidopsis thaliana*

1. Article publié dans Nucleic Acids Research

Les résultats décrivant le rôle de URT1 dans la protection des extrémités 3' des ARNm déadénylés sont présentés sous forme d'un article publié dans Nucleic Acids Research. Dans cet article, où je figure en deuxième auteur, j'ai caractérisé les propriétés catalytiques de URT1 et sa spécificité pour l'UTP *in vitro*. D'autre part, nous avons mis en évidence que URT1 appartient à la même voie de dégradation que l'exonucléase 5'-3' cytoplasmique XRN4. En effet, les substrats ARNm uridylés sont des substrats de la dégradation effectuée par XRN4. Sachant que le mutant nul *xrn4-3* est insensible à l'éthylène, j'ai comparé le phénotype de *xrn4-3* avec le mutant nul *urt1-1* afin d'observer si URT1 est également un facteur de la réponse à l'éthylène.

Nucleic Acids Research Advance Access published June 6, 2013

Nucleic Acids Research, 2013, 1–13
doi:10.1093/nar/gkt465

Uridylation prevents 3′ trimming of oligoadenylated mRNAs

François Michaël Sement[1], Emilie Ferrier[1], Hélène Zuber[1], Rémy Merret[2], Malek Alioua[1], Jean-Marc Deragon[2], Cécile Bousquet-Antonelli[2], Heike Lange[1] and Dominique Gagliardi[1,*]

[1]Institut de Biologie Moléculaire des Plantes, Centre National de la Recherche Scientifique (CNRS), Université de Strasbourg, 12 rue du général Zimmer, 67084 Strasbourg Cedex, France and [2]Laboratoire Génome et Développement des Plantes, Centre National de la Recherche Scientifique (CNRS), Université de Perpignan Via Domitia, 58 avenue P. Alduy, 66860 Perpignan Cedex, France

Received January 31, 2013; Revised April 15, 2013; Accepted May 5, 2013

ABSTRACT

Degradation of mRNAs is usually initiated by deadenylation, the shortening of long poly(A) tails to oligo(A) tails of 12–15 As. Deadenylation leads to decapping and to subsequent 5′ to 3′ degradation by XRN proteins, or alternatively 3′ to 5′ degradation by the exosome. Decapping can also be induced by uridylation as shown for the non-polyadenylated histone mRNAs in humans and for several mRNAs in *Schizosaccharomyces pombe* and *Aspergillus nidulans*. Here we report a novel role for uridylation in preventing 3′ trimming of oligoadenylated mRNAs in Arabidopsis. We show that oligo(A)-tailed mRNAs are uridylated by the cytosolic UTP:RNA uridylyltransferase URT1 and that URT1 has no major impact on mRNA degradation rates. However, in absence of uridylation, oligo(A) tails are trimmed, indicating that uridylation protects oligoadenylated mRNAs from 3′ ribonucleolytic attacks. This conclusion is further supported by an increase in 3′ truncated transcripts detected in *urt1* mutants. We propose that preventing 3′ trimming of oligo(A)-tailed mRNAs by uridylation participates in establishing the 5′ to 3′ directionality of mRNA degradation. Importantly, uridylation prevents 3′ shortening of mRNAs associated with polysomes, suggesting that a key biological function of uridylation is to confer 5′ to 3′ polarity in case of co-translational mRNA decay.

INTRODUCTION

Uridylation participates in the control of RNA stability in various eukaryotes, from *Schizosaccharomyces pombe* to nematodes, plants or man. The broad spectrum of RNAs subjected to uridylation includes U6 snRNA, mRNAs, RNA-induced silencing complex (RISC)-cleaved fragments, small RNAs and their precursors (1–5). Uridylation has diverse effects on the fate of non-coding RNAs. For instance, oligouridylation destabilizes human let-7 microRNA (miRNA) precursors in human embryonic cells and cancer cells, whereas monouridylation of let-7 miRNA precursors favours let-7 miRNA maturation by Dicer in different cell types (6). Uridylation can therefore promote or prevent let-7 miRNA production depending on the cellular context. In addition to miRNA precursors, mature small RNAs can also be uridylated, with different consequences on their stability. In mammals, uridylation can abrogate miRNA activity without affecting miRNA stability (7,8). However, uridylation can also trigger the degradation of miRNA and siRNA as shown in *Caenorhabditis elegans*, zebrafish, *Chlamydomonas reinhardtii* or Arabidopsis (9–13). In Arabidopsis, the 3′ nucleotide of siRNAs and miRNAs is methylated by the methyltransferase HEN1 (14). Methylation stabilizes small RNA and prevents their uridylation by the terminal uridylyltransferase HESO1 (12–15). Mutations in HESO1 partially stabilize small RNAs in a *hen1* background, revealing that methylation and uridylation have opposing effects on small RNA stability in Arabidopsis (12,13).

Other RNA targets of terminal uridylyltransferases include RISC-cleaved transcripts and mRNAs, and for both RNA types, uridylation enhances decapping followed by 5′ to 3′ degradation (16–22). For instance, oligouridylation of histone mRNAs in humans was shown to favour the binding of the Lsm1-7 complex, which in turn promotes decapping by Dcp2 and subsequent 5′ to 3′ degradation by Xrn1 (19). Uridylation of histone mRNAs triggers also 3′ to 5′ degradation by the exosome (19) and by the Eri1 exoribonuclease, which is recruited on binding of the Lsm1 complex to the uridylated 3′ terminal stem-loop (16). To date, the

*To whom correspondence should be addressed. Tel: +33 3 67 15 53 66; Fax: +33 3 88 61 44 42; Email: dominique.gagliardi@ibmp-cnrs.unistra.fr

© The Author(s) 2013. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from http://nar.oxfordjournals.org/ by guest on July 3, 2013

demonstration of uridylation-mediated mRNA degradation is restricted to non-polyadenylated histone mRNAs in mammals. However, uridylation of several mRNAs by the nucleotidyl transferase Cid1 in *S. pombe* triggers decapping and subsequent 5′ to 3′ degradation, revealing that uridylation can be an integral step of bulk mRNA decay (20). Interestingly, uridylation-induced decapping is independent of deadenylation in *S. pombe* (20). By contrast, deadenylation precedes addition of U/C-rich sequences to *gdhA* mRNAs in *Aspergillus nidulans*, which also triggers decapping (21). Deadenylation before 3′ U/C tagging is not required for all mRNA types in *A. nidulans* because it can be bypassed for transcripts containing premature stop codons that are substrates of the non-sense–mediated decay pathway (22). Collectively, these data reveal an emerging role of uridylation in controlling the stability of coding and non-coding RNAs in eukaryotes (4).

Two lines of evidence argue in favour of mRNA uridylation playing a role in mRNA metabolism in Arabidopsis. First, the uridylation of *CCR2* mRNAs was recently reported (22). However, the biological significance of this post-transcriptional modification remains to be determined. Second, the Arabidopsis Cid1-related protein encoded by At2g45620 catalyses the addition of uridines when artificially tethered to a non-adenylated reporter mRNA ectopically expressed in Xenopus oocytes (23). Here, we describe the function of UTP:RNA uridylyltransferase 1 (URT1), the At2g45620 gene product. We show that URT1 targets oligoadenylated mRNAs *in vivo* and has little impact on mRNA degradation rates. More importantly, in *urt1* mutants lacking mRNA uridylation, we observed the 3′ trimming of oligo(A)-tailed transcripts for all mRNAs tested, indicating that uridylation prevents 3′ to 5′ ribonucleolytic attacks. Importantly, uridylation can prevent the 3′ trimming of mRNAs still engaged on polysomes. We propose that mRNA uridylation participates in establishing the 5′ to 3′ polarity of mRNA degradation, which could be crucial in case of co-translational decay.

MATERIALS AND METHODS

Materials

Arabidopsis thaliana plants, Columbia ecotype (Col-0), were grown on soil with 16 h light/8 h darkness cycles. The characterization of *urt1-1* (Salk_087647C) and *urt1-2* (WISCDSLOXHS208_08D) T-DNA insertion lines is shown in Supplementary Figure S1. *xrn4-3* (SALK_014209) has been described previously (24). Sequences of primers are shown in Supplementary Table S1.

Subcellular localization of URT1

The coding sequence of *URT1* (At2g45620), *DCP1* (At1g08370) and *RBP47* (At1g19130) was fused 5′ to GFP or RFP sequences using the Gateway® destination vectors pB7FWG2 or pH7RWG2, respectively. Biolistic transformation of tobacco BY2 cells and analysis by confocal microscopy were performed by standard methods.

Expression and purification of recombinant proteins

Recombinant His-tagged GST-*URT1* or the catalytically inactive mutant GST-URT1$^{D491/3A}$ were produced in BL21(DE3) cells grown at 17°C. Cells were disrupted by sonication in 20 mM MOPS, pH 7.5, 250 mM KCl, 15% (v/v) glycerol, 1 mM DTT, 0.1% (v/v) Tween 20 in presence of protease inhibitors (Roche). Recombinant proteins were purified by classical Immobilized Metal Affinity Chromatography on Ni-NTA resin followed by glutathion affinity chromatography. Purified proteins were dialysed against 20 mM MOPS, pH 7.5, 100 mM NaCl, 15% (v/v) glycerol, 0.1% (v/v) Tween 20. Aliquots were snap frozen in liquid nitrogen and stored at −80°C.

Activity assays

URT1 and Cid1 activities were compared by using identical amounts of arbitrary units defined by measuring trichloroacetic acid-precipitable incorporation of [α-^{32}P]-UTP into tRNAs and reaction conditions based on (25). *In vitro* assays shown in Figure 1A contained 100 nM of GST-URT1 or GST-URT1$^{D491/3A}$ or 0.1 unit/µl of Cid1 (NEB), 25 mM potassium phosphate, pH 7, 5 mM $MgCl_2$, 5 mM DTT, 10% (v/v) glycerol, 0.1 µg/µl bovine serum albumin (BSA), 0.1 mM UTP, GTP, CTP or ATP and 5 nM of a 21-nucleotide synthetic RNA 5′-end labelled by T4 Polynucleotide Kinase (NEB) and [γ-^{32}P]-ATP. For experiments shown in Figure 1B, the reaction buffer was 20 mM MOPS, pH 7.5, 100 mM NaCl, 15% (v/v) glycerol, 0.1% (v/v) Tween 20, 1 mM $MgCl_2$, 0.5 µg/µl BSA, 1 mM UTP. Reaction products were separated by denaturing 17% (w/v) polyacrylamide gel electrophoresis before autoradiography. To determine the composition of the tails synthesized by URT1 in the presence of the four ribonucleotides, the latter reaction conditions were used except that the concentration for each ribonucleotide was 1 mM. Nucleotide composition was determined by 3′ Rapid Amplification of cDNA ends (RACE) polymerase chain reaction (PCR), cloning and sequencing as detailed below.

3′ RACE and circular reverse transcriptase-polymerase chain reaction (cRT-PCR) analyses

The 3′ RACE PCR protocol used to detect uridylated mRNAs is detailed in (26). Briefly, 5 µg of total RNA was purified with Nucleospin® RNA plant columns (Macherey Nagel), dephosphorylated with calf intestinal phosphatase (CIP, NEB) before ligation to the RNA primer (5′P-CUAG AUGAGACCGUCGACAUGAAUUC-3′NH_2) with T4 RNA ligase (Fermentas). This sequence was used as an anchor to initiate cDNA synthesis by SuperScript® III reverse transcriptase (Invitrogen) and the anchor rv1 primer (Supplementary Table S1). Two nested PCR amplification of 30 cycles were performed using anchor rv1 and rv2 primers in combination with gene-specific sense primers (Supplementary Table S1).

Circular reverse transcriptase (cRT)-PCR analysis was used to characterize the 5′ and 3′ extremities of capped and

Downloaded from http://nar.oxfordjournals.org/ by guest on July 3, 2013

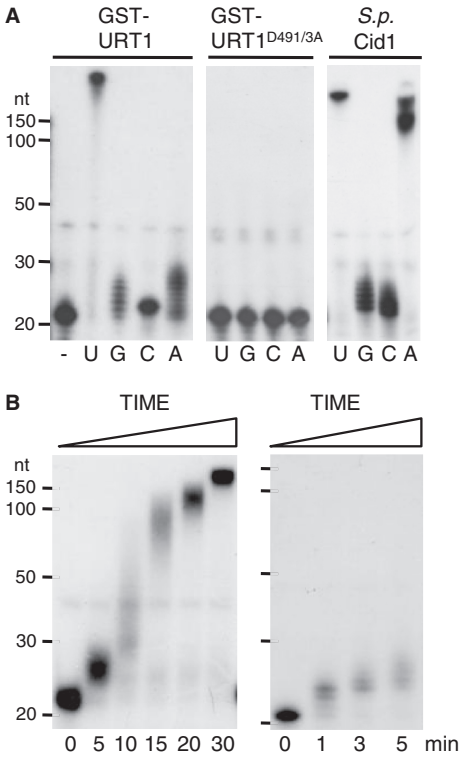


Figure 1. Catalytic properties of URT1. (**A**) URT1 is a UTP:RNA uridylyltransferase. GST-URT1 or the catalytically inactive mutant GST-URT1$^{D491/3A}$ were incubated for 30 min with a 21-nucleotide 5′ [^{32}P]-labelled synthetic RNA without (−) or in the presence of each ribonucleotide as indicated. Reaction products were separated on 17% denaturing polyacrylamide gels before autoradiography. The migration of size standards (150, 100, 50, 30 and 20 nt) is indicated on the left. For comparison, the catalytic activity of *S. pombe* Cid1 is shown. (**B**) URT1's activity is distributive for the first added nucleotides as shown by time-course analysis. Incubation time is indicated in minutes at the bottom of panels.

uncapped *LOM1* transcripts. To detect capped transcripts, 5 µg of total RNA was first treated with 5 units of CIP (NEB) to dephosphorylate uncapped RNAs and then with 2.5 units of tobacco acid pyrophosphatase (TAP) (Epicentre) to generate 5′ phosphorylated transcripts from capped mRNAs. The 5′ phosphorylated transcripts were then circularized by 20 units of T4 RNA ligase at 18°C for 16 h in a final volume of 200 µl. Uncapped transcripts were detected following mRNA circularization from total RNA that was neither treated with CIP nor with TAP. Negative controls were treated only with CIP because neither dephosphorylated nor capped mRNAs can be circularized by T4 RNA ligase.

3′ RACE and cRT PCR products were cloned in pGEM®-T Easy (Promega) before sequence analysis.

Polysome purification

Polysome purification was performed according to (27) except that 400 mg of tissue powder was loaded on 11 ml of 15–60% (w/v) sucrose gradients. A puromycin treatment abolished the typical UV polysome profiles and, as expected, dramatically reduced the RNA content in heavy fractions of the gradients, preventing reliable 3′ RACE analysis of puromycin controls.

Additional methods

Northern blot analyses, mRNA half-life measurements, microarray analysis and statistical analysis of microarray data are described in Supplementary Methods.

RESULTS

Nucleotide specificity of URT1

To characterize the nucleotidyltransferase activity of URT1, we produced recombinant proteins corresponding either to the wild-type protein or to a mutated version in which two conserved aspartate residues essential for catalysis of nucleotidyltransferases (D491 and D493 in URT1) were mutated to alanines. The nucleotidyltransferase activity of both proteins was tested using a 5′ [^{32}P]-labelled oligoribonucleotide substrate and conditions previously used to describe Cid1's activity (25). No activity was observed for the mutant protein URT1$^{D491/3A}$, whereas URT1 showed a robust uridylation activity (Figure 1A). Under these assay conditions, URT1 synthesizes tails of >200 Us, while only 1–9 of the three other ribonucleotides are appended to the RNA substrate in 30 min. To further test that URT1 preferentially incorporates uridines, URT1 was incubated with the unlabelled oligoribonucleotide substrate and the four ribonucleotides ATP, GTP, CTP and UTP. The composition of tails synthesized by URT1 was then determined by sequence analysis of 12 cloned reaction products. URT1-synthesized tails contained 97.8% (947/967 nucleotides) uridines, confirming the strong uridine selectivity of URT1. In line with these *in vitro* results, the residues involved in uridine selectivity in Cid1 (D330, E333, H336) (28–30) are conserved in URT1 (D708, E711, H714). Interestingly, under the *in vitro* conditions used, URT1 better discriminates UTP from ATP as compared with Cid1 (Figure 1A).

Testing the stringency of URT1's nucleotide specificity *in vitro* entailed the use of reaction conditions that maximize activity. As a result, large U-tails are polymerized, which unlikely reflect the *in vivo* situation. In fact, a restricted number of Us (1–10) are added to almost all RNA substrate molecules on short incubation times, indicating that the activity of URT1 is distributive for the first added nucleotides (Figure 1B). Following the addition of ~10 uridines, URT1's activity seems to become more processive. This switch from distributivity to processivity is not due to the short size of the initial oligoribonucleotide substrate (21 nt) because it was also observed for two RNA substrates of 31 nt (Supplementary Figure S2). The distributivity for the first added nucleotides appears therefore an inherent feature of URT1.

Taken together, these data show that Arabidopsis URT1 is a UTP:RNA uridylyltransferase with a marked preference for uridine polymerization and a distributive activity for the first added nucleotides. In line with this conclusion, the *in vivo* RNA substrates of URT1 are

Downloaded from http://nar.oxfordjournals.org/ by guest on July 3, 2013

modified by the addition of a limited number of uridines as detailed below.

URT1 uridylates oligoadenylated mRNAs

Expressing fluorescent URT1 fusion proteins revealed a diffuse cytosolic pattern with foci (Figure 2A). These foci were observed in all cells examined. URT1-GFP and URT1-RFP partially co-localized with known markers of processing bodies (P-Bodies) and stress granules, DCP1-RFP or RBP47-GFP, respectively (31,32), suggesting that URT1 can associate with cytosolic ribonucleoprotein (RNP) granules (Figure 2A). These results also indicate that URT1 targets cytosolic RNAs. To identify such targets, we investigated mRNA uridylation by nested 3′ RACE PCR experiments using gene-specific forward primers and reverse primers complementary to a 3′ ligated RNA anchor sequence as detailed in (26). Initial tests on mRNAs selected on different criteria such as short half-lives or misregulation in *urt1* mutants revealed that any mRNA investigated is a substrate of URT1. Detailed features of mRNA uridylation are presented for five mRNAs: *LOM1* (At2g45160), *DREB2C* (At2g40340), *BAM3* (At4g20270), At2g21560 and At5g48250. Of note, *DREB2C* and *BAM3* are among the shortest-lived mRNAs in Arabidopsis (33) and At2g21560 and At5g48250 were identified among genes misregulated in *urt1* (Supplementary Table S2). mRNA uridylation was first assessed in WT plants. Cloning and sequence analysis of 3′ RACE products revealed that 14% (6/43) of *LOM1*, 43% (26/60) of *DREB2C*, 36% (23/64) of *BAM3*, 57% (33/58) of At2g21560 and 18% (19/105) of At5g48250 clones corresponded to uridylated mRNAs (Figure 2B). Uridine extensions were short in WT, with 1 or 2 uridines detected for 92% (98/107) of the clones corresponding to uridylated mRNAs (Figure 2C). Overall, uridylated transcripts had significant shorter poly(A) tails than non-uridylated ones (Figure 2D). The average size of uridylated oligo(A) tails for the 5 mRNAs analysed in WT was 14 (n = 106). This size is similar to the length recently reported for oligo(A) tails of uridylated *CCR2* mRNAs (22) and is in the range for the known oligo(A) size of deadenylated mRNAs in eukaryotes. Taken together, these data indicate that oligoadenylated mRNAs can be uridylated in Arabidopsis.

Next, we analysed the uridylation status of the five candidate mRNAs in *urt1* mutants. Interestingly, we did not detect uridylated transcripts for *LOM1* (0/55), At2g21560 (0/55) and At5g48250 (0/103) and only 2/72 *DREB2C* clones and 1/107 *BAM3* clones corresponded to uridylated transcripts in *urt1* (Figure 2B). URT1 is therefore the nucleotidyltransferase activity that is mainly responsible for uridylating mRNAs in Arabidopsis.

mRNA uridylation was also determined in a *xrn4* mutant because the Arabidopsis 5′ to 3′ cytosolic exoribonuclease XRN4 (34) is a reasonable candidate for degrading uridylated mRNAs by analogy with data reported for humans and *S. pombe* (19,20). As in WT, uridylated mRNAs were detected in *xrn4* for the five candidate mRNAs (Figure 2B), and uridylated mRNAs had significant shorter oligo(A) tails than non-uridylated ones in *xrn4* (Figure 2D). However, both composition and length of the U tails differ between WT and *xrn4* (Figure 2C). Whereas most U tails of WT mRNAs are 1-2 Us in length, the majority of oligo(U) tails (68/122) in *xrn4* were at least 3 nucleotide long and could reach up to 11 nucleotides (Figure 2C). In addition, the oligo(U)-rich tails in *xrn4* included up to 10% As (37/381), whereas the short tails in WT were almost exclusively composed of Us (180/181 nucleotides). The fact that both length and composition of U-tails are modified in absence of XRN4 suggests that uridylated mRNAs are substrates of XRN4.

Collectively, these data support the idea that URT1 uridylates mRNAs, that a deadenylation step precedes uridylation in Arabidopsis and that uridylated mRNAs are substrates of XRN4.

Uridylation marks uncapped, 5′ shortened *LOM1* transcripts

Uridylation was reported to enhance decapping of mRNAs in *S. pombe* and *A. nidulans* and of non-polyadenylated histone mRNAs in humans (19–21). We therefore compared the proportion of uridylated transcripts between capped and uncapped transcripts by cRT-PCR using *LOM1* as a model substrate. Capped *LOM1* transcripts were detected in WT, *urt1* and *xrn4*, whereas uncapped *LOM1* mRNAs accumulated to detectable levels only in *xrn4* (Figure 3A). This is in line with the idea that uncapped transcripts are degraded by XRN4. The 3′ extremities of both capped and uncapped transcripts in *xrn4* mapped between positions +2074 to +2102 (Figure 3B). For the uncapped samples, 42 out of 70 clones were uridylated and only 2 out of 70 were not. For the rest of the clones (26/70), sequence analysis of cRT-PCR products cannot discriminate between uridines that could be either genomically encoded at the 5′ or added at the 3′ extremities. We favour the latter possibility because >90% of the uridylated *LOM1* mRNAs had shortened poly(A) tails and half of them correspond to 5′ truncated transcripts (see below). Therefore up to 98% of the clones in the uncapped sample correspond to uridylated transcripts. By contrast, only 1 clone out of 73 corresponding to capped transcripts was uridylated. In addition, 95% (69/73) of 5′ extremities of capped transcripts map at or within 27-nt downstream of the known *LOM1* 5′-ends. By contrast, the 5′ extremities of 49% (34/70) of uncapped transcripts map within the coding sequence, indicating that a 5′ shortening occurs even in absence of the known cytosolic 5′ to 3′ exoribonuclease XRN4 (Figure 3B). Whether these 5′ shortened transcripts in *xrn4* are generated by an endoribonuclease activity or by a cytosolic 5′ to 3′ exoribonuclease distinct from XRN4 is presently unknown. Interestingly, the uncapped, 5′ shortened transcripts that are detected in *xrn4* mutants share the same 3′ extremities as capped transcripts. This indicates that even in *xrn4* where 5′ to 3′ RNA degradation is compromised, uridylated *LOM1* transcripts are still prone to 5′ to 3′ degradation rather than to 3′ attacks.

The cRT-PCR experiments also confirmed that uridylated *LOM1* transcripts have a decreased poly(A) size as compared with non-uridylated capped transcripts

Downloaded from http://nar.oxfordjournals.org/ by guest on July 3, 2013

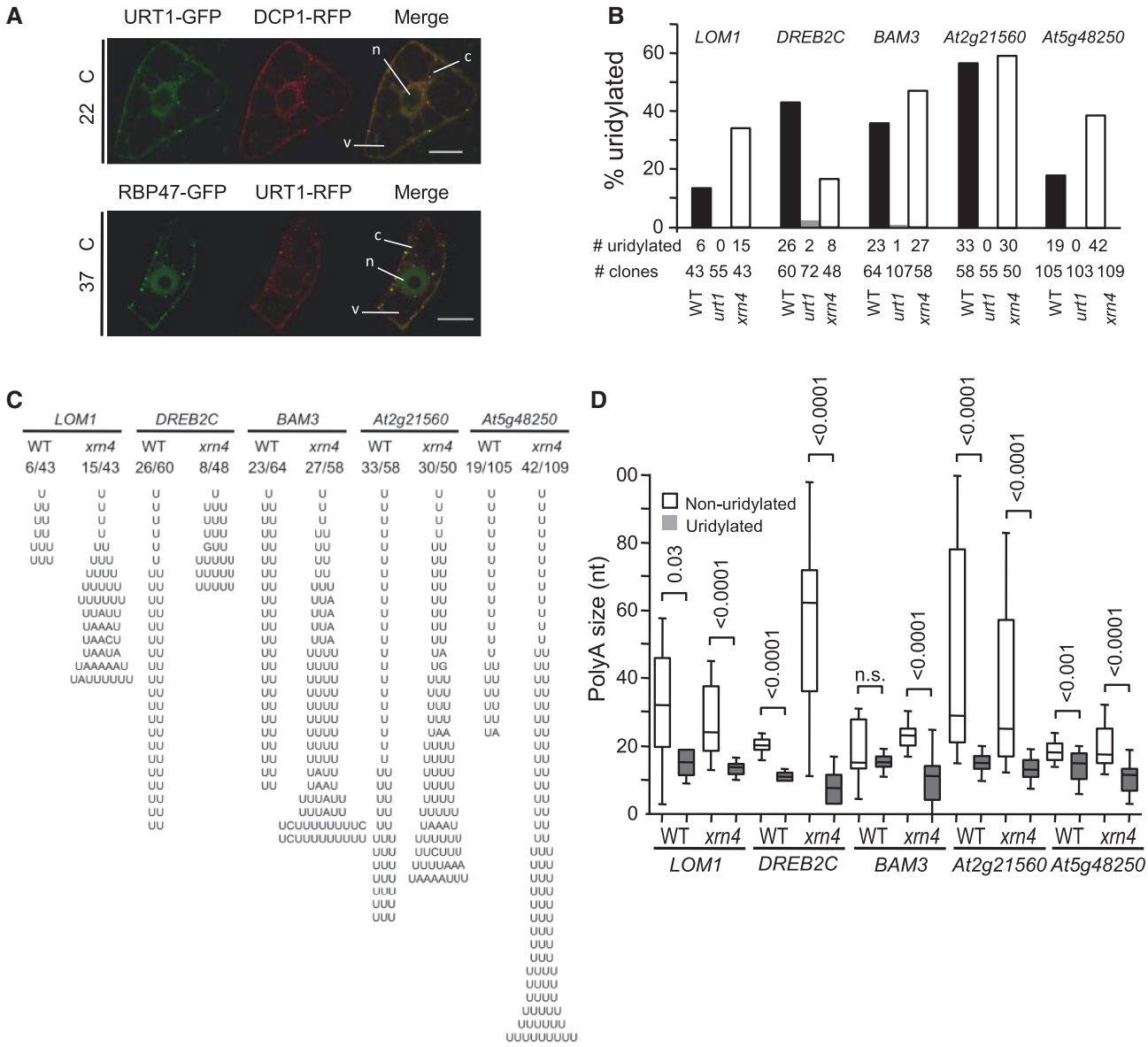


Figure 2. Deadenylation precedes uridylation. (**A**) URT1 is a cytosolic protein. URT1 fused to GFP (upper panel) or RFP (lower panel) was transiently expressed in tobacco BY-2 cells. Markers of P-bodies and stress granules were DCP1-RFP and RBP47-GFP, respectively. Co-localization of URT1-GFP with DCP1-RFP and of URT1-RFP with RBP47-GFP was observed for 27 of 35 and 39 of 50 foci, respectively, as measured from two independent samples. Cells shown in the lower panel were incubated 30 min at 37°C before observation. n, nucleus; v, vacuole; c, cytosol. Scale bars are 30 µm (top) and 15 µm (bottom). (**B**) URT1 uridylates mRNAs. Percentage of uridylated mRNAs determined by 3′ RACE PCR in WT, *urt1* and *xrn4* for *LOM1* (At2g45160), *DREB2C* (At2g40340), *BAM3* (At4g20270), At2g21560 and At5g48250 (n = 2). Number of clones (# clones) and number of clones corresponding to uridylated mRNAs (# uridylated) are indicated for each genotype. (**C**) Composition and length of U-tails detected at the 3′-ends of *LOM1*, *DREB2C*, *BAM3*, At2g21560 and At5g48250 mRNAs in WT and *xrn4*. The number of clones corresponding to uridylated mRNAs on the total number of clones is indicated. (**D**) Deadenylation precedes uridylation. Box plot analysis of poly(A) tail length distribution for non-uridylated (white) and uridylated (grey) *LOM1*, *DREB2C*, *BAM3*, At2g21560 and At5g48250 mRNAs in WT and *xrn4* as indicated. The second and third quartiles are shown by the box edges, the median is indicated by a horizontal bar, whiskers show the range of data except far-outliers. *P*-values (Mann–Whitney test) are indicated for each pair, except when non-significant (n.s.).

(Figure 3C). In *xrn4*, the *LOM1* poly(A) tail distribution had a median of 50 As for capped transcripts and 11 As for uridylated uncapped transcripts ($P < 0.0001$, Mann–Whitney test). The prevalence of uridylated transcripts in the uncapped fraction, the 5′ shortening of uridylated transcripts and the fact that uridylated transcripts have short poly(A) tails suggest that mRNA uridylation in Arabidopsis is linked to decapping and 5′ to 3′ degradation, as previously reported in humans, *S. pombe* and *A. nidulans*. However, it is difficult at present to estimate whether uridylation facilitates decapping in Arabidopsis because uridylation marks oligoadenylated mRNAs, and deadenylation alone can promote decapping. Importantly, none of the phenotypes linked with compromised

Downloaded from http://nar.oxfordjournals.org/ by guest on July 3, 2013

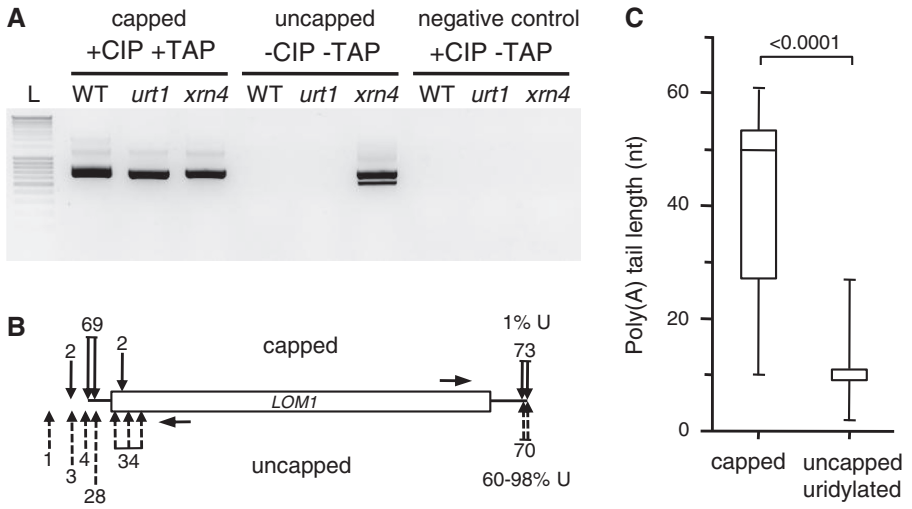


Figure 3. Uridylation marks 5′ shortened uncapped *LOM1* mRNAs. (**A**) Detection of capped and uncapped *LOM1* mRNAs by cRT-PCR analysis following combinations of CIP and TAP treatments as indicated. Capped mRNAs were detected following successive treatments with CIP and TAP. Both CIP and TAP treatments were omitted to detect uncapped transcripts. Negative controls were treated with CIP only (see details in 'Materials and Methods' section). PCR products were run alongside a DNA ladder (L) with 100-bp size increments. The negative image of an ethidium bromide–stained agarose gel is shown. (**B**) Most uncapped *LOM1* mRNAs are uridylated. Positions of 5′- and 3′-ends of capped (black arrows above diagram) and uncapped (dashed arrows below diagram) *LOM1* mRNAs in *xrn4*. Seventy-three and 70 clones were analysed for capped and uncapped mRNAs, respectively. The numbers of clones is indicated for each position. Horizontal arrows indicate the position of primers used for cRT-PCR. The percentage of clones corresponding to uridylated mRNAs is shown at the 3′-ends. (**C**) Uncapped uridylated *LOM1* mRNAs have shorter poly(A) tails in *xrn4* as compared with capped transcripts. Box plot analysis as in Figure 2.

decapping or 5′ to 3′ RNA degradation in Arabidopsis were observed in *urt1* mutants, suggesting that decapping is not adversely compromised in absence of URT1. For instance, *urt1* mutants do not show the ethylene-insensitive phenotype of *xrn4* seedlings (35,36) (Supplementary Figure S3) and we did not observe abnormalities in embryo-to-seedling transition, a developmental step strongly affected by the loss of decapping activity (31,37).

Uridylation has no major impact on mRNA degradation rates

The genome-wide impact of URT1 on mRNA steady-state levels was determined by microarray analysis. One hundred sixty-four genes showed significant different expression levels between WT and *urt1* (65 up, 99 down) (see Supplementary Table S2 for a list of misregulated genes and gene ontology analysis, and Supplementary Figure S4A for the validation of three misregulated genes in *urt1*). The number of differentially expressed genes is comparable with those observed for *xrn4* (38,39). However, there is no significant overlap between *urt1* and *xrn4* data sets because only one gene (At5g49280) is misregulated both in *urt1* and *xrn4*. To test a potential impact of URT1 on mRNA degradation rates, half-lives of five short-lived transcripts (i.e. half-lives ranging between 20 and 40 min) were determined following inhibition of transcription by cordycepin or actinomycin D. The degradation rates of the five mRNAs were not significantly affected regardless of their respective up-regulated, down-regulated or unchanged expression level in *urt1* (Supplementary Figure S4B and C). Therefore, the variation of steady-state levels of these mRNAs in *urt1* is rather due to indirect effects or multifactorial causes that remain to be identified. Although we cannot exclude that the mRNA half-lives determined following transcription inhibition by cordycepin or actinomycin D could not faithfully reproduce the actual *in vivo* half-lives, these data argue that URT1 has no major impact on mRNA degradation rates.

Oligoadenylated mRNAs are trimmed in *urt1* mutants

A new role for mRNA uridylation in Arabidopsis was determined following the observation of a slight but systematic reduction in the size of the main 3′ RACE PCR products for 16 of the most up- or down-regulated mRNAs in *urt1* as well as for the short-lived transcripts *DREB2C* and *BAM3* or any other mRNAs investigated (Figures 4A and B, Supplementary Figure S1C and S5B). This size shift in *urt1* is not only detected for all mRNAs tested but is also observed for the different developmental stages analysed: rosette leaves, seedlings and flowers (Figures 4A and B, Supplementary Figure S5B) showing the robustness of this molecular phenotype in *urt1* mutants.

To understand the nature of the size shift, we cloned size-selected amplification products corresponding to the main band for *DREB2C* and *BAM3*. The distribution of poly(A) sites was similar between WT, *urt1* and *xrn4*. However, the cloned poly(A) tails were shorter in *urt1* (Mann–Whitney test $P < 0.001$) (Figure 4C). Enhanced deadenylation is therefore responsible for the size shift observed by 3′ RACE PCR in *urt1*. Because 3′ RACE PCR preferentially detects shorter over long poly(A) tails, we checked whether only oligo(A) tails rather than full-length poly(A) tails were shortened in *urt1* mutants, a rationale hypothesis because URT1's RNA substrates

Downloaded from http://nar.oxfordjournals.org/ by guest on July 3, 2013

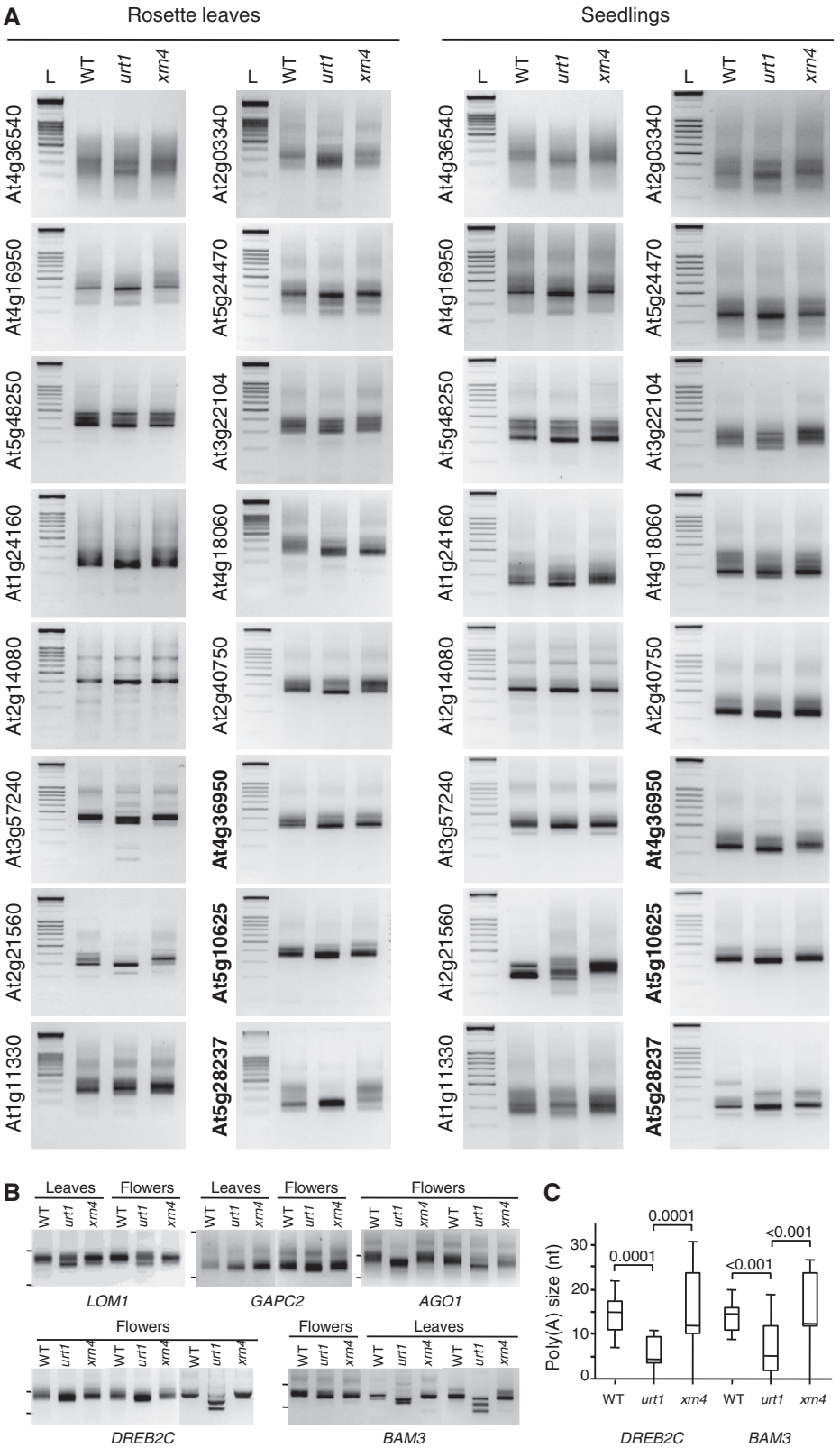


Figure 4. Uridylation protects from mRNA 3′ shortening. (**A**) URT1 prevents 3′ shortening of mRNAs irrespective of their misregulation in *urt1*. Thirteen down- and 3 up-regulated transcripts (regular and bold AGI numbers, respectively) in *urt1* were analysed by 3′ RACE analysis in WT, *urt1* and *xrn4* at two developmental stages, rosette leaves and seedlings. All panels correspond to negative images of ethidium bromide–stained 2% agarose gels. 3′ RACE PCR products were run alongside a DNA ladder (L) with 100-bp size increments. Complex patterns reflect multiple polyadenylation sites frequently observed for Arabidopsis mRNAs. (**B**) 3′ shortening in *urt1* detected by 3′ RACE PCR for 5 mRNAs. Genotypes, tissues and gene identities are indicated for all panels corresponding to negative images of ethidium bromide–stained 2% agarose gels. The migration of 0.5 and 0.2 kb size standards are shown on the left. (**C**) Increased deadenylation of *DREB2C* and *BAM3* mRNAs in *urt1*. Box plot analysis as in Figure 2.

Downloaded from http://nar.oxfordjournals.org/ by guest on July 3, 2013

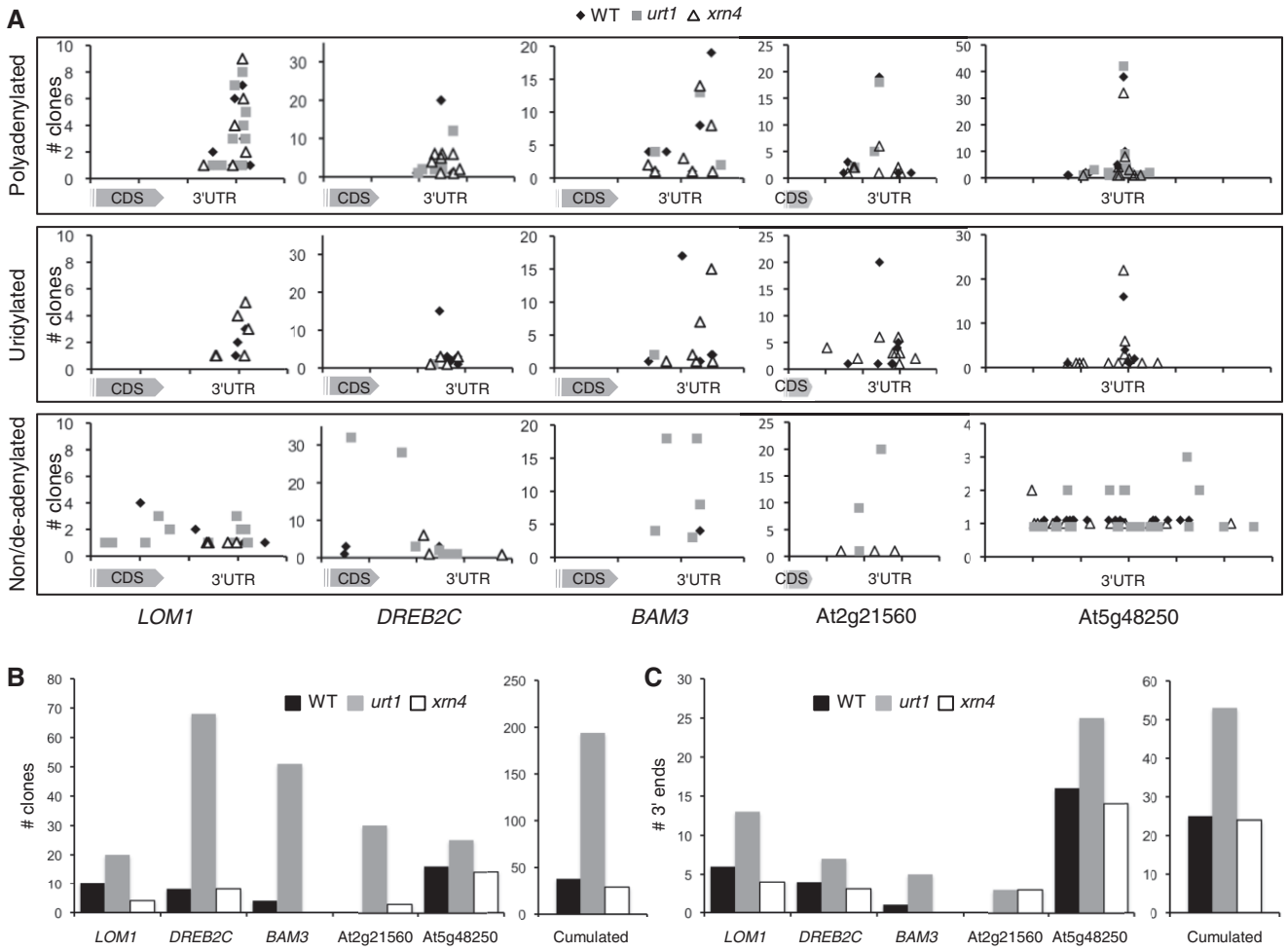


Figure 5. 3′ truncated transcripts accumulate in *urt1*. (**A**) Number of clones (n = 141 for *LOM1*, n = 180 for *DREB2C*, n = 229 for *BAM3*, n = 163 for At2g21560, n = 314 for At5g48250) from at least two independent experiments and corresponding to polyadenylated, uridylated, non/deadenylated (from 0 to 10 As) mRNAs were plotted against the 3′ region of each respective mRNA as indicated. One interval on abscissas is 100 nt. CDS, coding sequence. (**B**) Individual and cumulated numbers of non/deadenylated clones for *LOM1*, *DREB2C*, *BAM3*, At2g21560 and At5g48250 in WT, *urt1* and *xrn4*. (**C**) Individual and cumulated numbers of distinct 3′-ends of non/deadenylated clones for *LOM1*, *DREB2C*, *BAM3*, At2g21560 and At5g48250 in WT, *urt1* and *xrn4*.

correspond to oligoadenylated mRNAs. To this end, we chose the At5g08040 mRNAs which, unlike *DREB2C* and *BAM3*, are expressed to sufficient levels to be detected by northern analysis and are well resolved on acrylamide gels owing to their short size of 371 nt [excluding poly(A) tails]. A northern analysis was performed using RNA samples from WT, *urt1* and *xrn4* flowers that were either treated with RNase H and oligo(dT) to remove poly(A) tails, or not treated with RNase H. This analysis revealed that similar long poly(A) tails of a maximal size of ~100 nt are observed for At5g08040 mRNAs in the three genotypes (Supplementary Figure S5A). The slight size shift for RNA species with shorter tails is more difficult to observe on northern blots but is convincingly detected by 3′ RACE PCR in flowers, as well as in seedlings (Supplementary Figure S5B).

A second molecular phenotype was observed in *urt1*. As shown for several experiments for *BAM3* and *DREB2C*, additional 3′ RACE PCR products with reduced sizes were occasionally amplified in *urt1* mutants (Figure 4B). To understand the nature of these heterogeneous patterns, whole PCR reactions from two to three independent experiments were cloned. The corresponding sequence analysis is presented for *LOM1*, *DREB2C*, *BAM3*, At2g21560 and At5g48250 (Figure 5). Cloned 3′ RACE PCR products were split into three categories: (i) polyadenylated transcripts, (ii) uridylated transcripts and (iii) non-polyadenylated or deadenylated transcripts. Transcripts were considered deadenylated when the poly(A) tail length was <10, a size inferior to the minimal poly(A) tail length required for the binding of a poly(A) binding protein (40). Plotting the clone numbers against the position of 3′-ends on the respective mRNAs shows that multiple polyadenylation sites are detected for the five genes (Figure 5A). This observation is coherent with the intrinsic 3′ heterogeneity described for mRNA 3′-ends in Arabidopsis (41). The overall distribution of polyadenylation sites appeared unchanged in WT, *urt1* and *xrn4* (Figure 5A). Of note, the distribution of uridylated ends in WT and *xrn4* coincides perfectly with

Downloaded from http://nar.oxfordjournals.org/ by guest on July 3, 2013

the positions of polyadenylation sites (Figure 5A). In fact, all uridylated mRNAs detected in the course of this study in WT correspond to adenylated mRNAs [with a strong bias toward oligo(A) tails]. Therefore, although we do not formally exclude the possible uridylation of truncated and/or non-polyadenylated mRNAs in WT, our data indicate that oligoadenylated mRNAs are the main substrates of URT1. Importantly, the number and the diversity of clones corresponding to deadenylated or non-polyadenylated (from 0 to 10 As) transcripts were increased in *urt1* as compared with WT and *xrn4* (Figure 5A). Because clones having identical 3′-ends could correspond either to sister amplification or reflect the amplification of preferential degradation intermediates, we considered both the number of clones and the number of distinct 3′-ends detected for non/de-adenylated mRNAs. Overall, both the total number of clones and the diversity of the 3′-ends were similar in WT and *xrn4* but markedly increased in *urt1* (Figures 5B and C, respectively). For *LOM1* and *DREB2C*, the 3′ extremities of deadenylated or non-polyadenylated transcripts covered the whole range of the sequence investigated by 3′ RACE PCR with some 3′-ends mapping even within the CDS (Figure 5A). This analysis revealed that the occasional shorter PCR products detected by 3′ RACE correspond to 3′ truncated transcripts and that the occurrence of 3′ truncated transcripts is increased in *urt1*.

Taken together, these results show that the major phenotype observed in *urt1* mutants corresponds to a 3′ shortening of oligoadenylated mRNAs. Because this shortening was systematically observed and uridylation was detected for any mRNA investigated, we propose that mRNA uridylation is a frequent process in Arabidopsis and that uridylation prevents complete deadenylation of mRNAs. This is further supported by the increased occurrence of 3′ truncated transcripts in *urt1* mutants that likely results from 3′ exoribonucleolytic attacks in the mRNA body once the mRNA has been fully deadenylated (see 'Discussion' section).

Uridylation can prevent 3′ trimming of polysomal mRNAs

The occurrence of co-translational mRNA decay was reported in *S. cerevisiae* (42) and recently, uridylation was shown to tag polysomal mRNAs in another fungus, *A. nidulans* (22). In this species, uridylation tags oligoadenylated transcripts in a similar way as for Arabidopsis mRNAs (22). We therefore tested whether uridylation could prevent the occurrence of 3′ shortened mRNAs associated with polysomes. Polysomes were isolated from seedlings, and typical UV profiles were observed for WT, *urt1* and *xrn4* samples, indicating no major effects of the *urt1* mutation on polysome assembly (Figure 6A). RNA extracted from polysomal fractions was used to determine 3′ RACE PCR profiles for four transcripts, and the characteristic 3′ shortening was observed in *urt1* (Figure 6B). Two experiments are shown for *BAM3* and At5g48250 to show that the 3′ shortening is systematically observed in *urt1* even though the PCR profiles may vary slightly. In addition, shorter amplicons corresponding to 3′ truncated transcripts were occasionally detected in *urt1* polysomal fractions as shown for *BAM3*. Cloning and sequence analysis of *BAM3* and At5g48250 samples revealed the presence of uridylated mRNAs in polysomal fractions in WT and *xrn4* and confirmed that both the number of clones corresponding to non- or de-adenylated transcripts and the number of distinct 3′-ends are increased in *urt1* (Figure 7). These results indicate that uridylation can prevent 3′ trimming of mRNAs associated with polysomes.

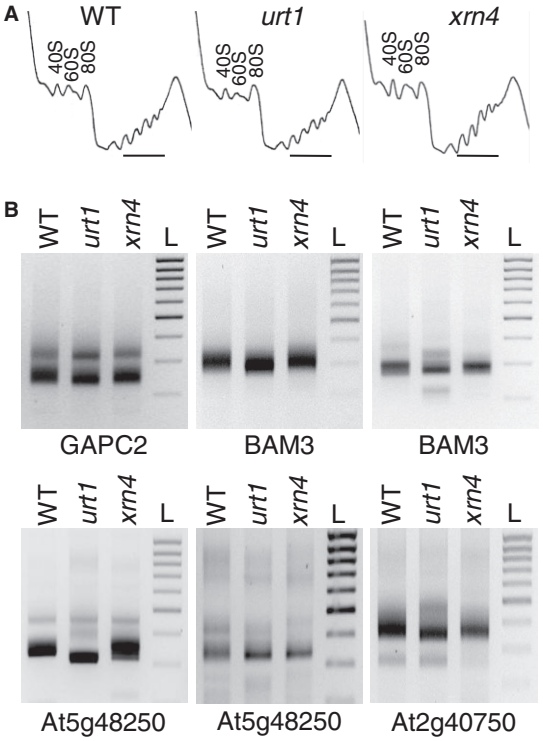


Figure 6. Uridylated mRNAs are present in polysomal fractions. (**A**) Polysome profiles in WT, *urt1* and *xrn4*. The position of the 40S, 60S and 80S peaks is indicated on each panel. Polysomal fractions analysed by 3′ RACE PCR are marked by a horizontal line. (**B**) 3′ shortening of mRNAs in *urt1* is detected in polysomal fractions by 3′ RACE PCR. Genotypes and identity of mRNAs are indicated on all panels corresponding to negative images of ethidium bromide–stained agarose gels. The migration of a DNA ladder (L) (1 to 0.2 kb, 0.1 kb size increment) is shown on the right of each panel. Two experiments are shown for *BAM3* and At5g48250 to show that the PCR profiles may vary but the 3′ shortening is systematically observed in *urt1*.

DISCUSSION

Uridylation of mRNAs was recently reported to promote 5′ to 3′ degradation through decapping (19,20). Here we show that in addition, uridylation prevents the formation of 3′ trimmed mRNAs in Arabidopsis, revealing a novel role for uridylation. We propose that uridylation favours the 5′ to 3′ directionality of mRNA decay, likely by protecting the 3′-ends of oligoadenylated mRNAs that are committed for 5′ to 3′ degradation as discussed below. 5′ to 3′ polarity is crucial in case of co-translational decay (42) and, interestingly, uridylation limits the occurrence of 3′ shortened mRNAs engaged in polysomes, highlighting one role for mRNA uridylation in Arabidopsis.

Downloaded from http://nar.oxfordjournals.org/ by guest on July 3, 2013

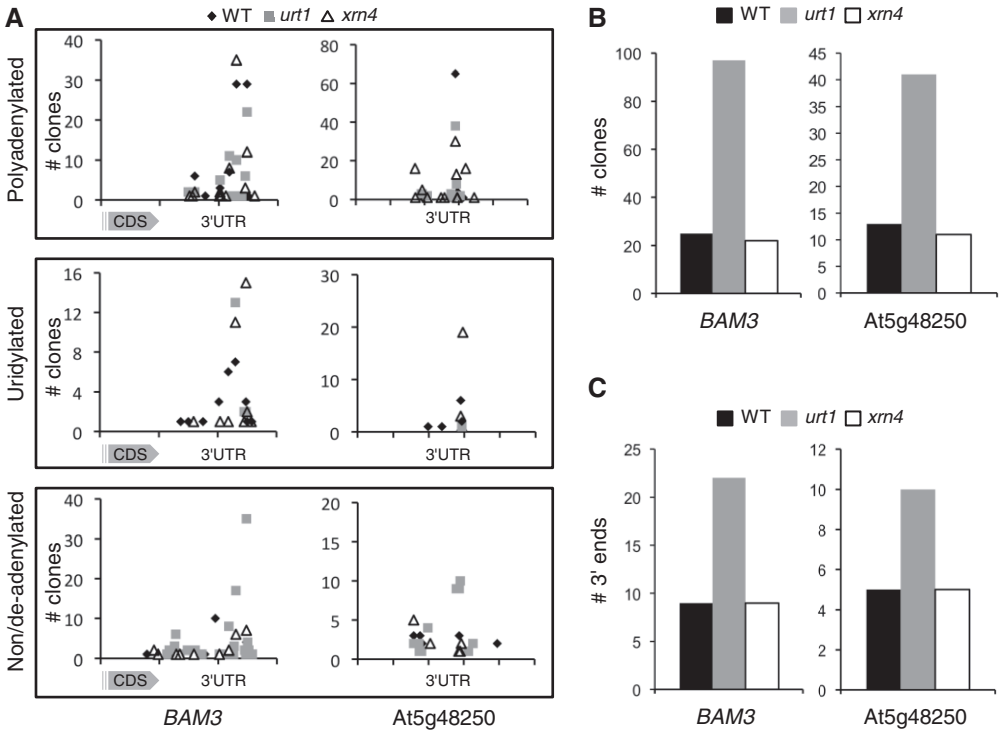


Figure 7. 3′ truncated transcripts accumulate in *urt1* polysomal fractions. (**A**) Number of clones (n = 423 for *BAM3*, n = 326 for *At5g48250*) corresponding to polyadenylated, uridylated, non/deadenylated (from 0 to 10 As) mRNAs plotted against position alongside the 3′ region of each respective mRNA as indicated. One interval on abscissas is 100 nt. CDS, coding sequence. (**B**) Numbers of non/deadenylated clones for *BAM3* and At5g48250 in WT, *urt1* and *xrn4*. (**C**) Numbers of distinct 3′-ends of non/deadenylated clones for *BAM3* and At5g48250 in WT, *urt1* and *xrn4*.

Our results indicate that uridylation marks oligoadenylated mRNAs in Arabidopsis. The exact proportion of a given mRNA to be uridylated remains to be determined because the PCR-based technique used to detect uridylated mRNAs favours the amplification of shorter versus longer poly(A) tails (Supplementary Figure S5) and uridylated transcripts have shorter tails than non-uridylated ones. The ratio of clones corresponding to uridylated mRNAs is therefore likely overestimated. Despite this technical limitation, the 3′ RACE PCR strategy has the advantages of being gene-specific, sensitive and allows the detection of a single uridine present at the 3′-end of a poly(A) tail. This method is therefore well suited for detecting uridylated RNAs and comparing their occurrence between genotypes. To date, uridylation was systematically detected for any mRNA investigated and is therefore likely to be widespread in Arabidopsis.

The absence of a significant impact of URT1 on mRNA degradation rates is coherent with *urt1* mutants displaying none of the phenotypes linked with compromised decapping or 5′ to 3′ RNA degradation. A likely interpretation of these observations is that the mRNAs that are marked by uridylation are already committed to degradation by the deadenylation step. Deadenylation is a rate-limiting step in mRNA degradation and is sufficient to trigger decapping. Both of these features could explain why mRNA degradation is not adversely compromised in *urt1* mutants.

The most important finding of this study is that uridylation is required to limit the occurrence of 3′ trimmed and truncated mRNAs. Three alternative interpretations could explain this observation. First, the 5′ to 3′ RNA degradation pathway could be impaired in *urt1*, thereby promoting 3′ to 5′ attacks of oligoadenylated mRNAs. Second, uridylation could be required to trigger the degradation of 3′ trimmed and truncated transcripts. Third, uridylation of oligoadenylated mRNAs could prevent 3′ trimming by protecting their 3′-ends. We do not favour the two first hypotheses for the following reasons. Several lines of evidence support the idea that Arabidopsis uridylated mRNAs are substrates of the 5′ to 3′ pathway as previously reported for humans and *S. pombe* (19,20): uridylation marks decapped mRNAs as shown for *LOM1*, these transcripts are 5′ shortened (Figure 3) and both the composition and length of U-tails are modified in *xrn4* (Figure 2C). However, we never observed 3′ trimming in *xrn4* for all mRNAs tested (Figures 4A and B, Figure 6, Supplementary Figure S1C and S5B), and truncated transcripts do not accumulate in *xrn4* as compared with WT (Figures 5B and C and 7B and C). Both observations indicate that 3′ trimmed and truncated transcripts are not merely generated by favouring 3′ to 5′ attacks when 5′ to 3′ degradation is impaired.

Importantly, we never detected 3′ truncated transcripts that were uridylated, neither in WT nor in *xrn4*, even though the PCR strategy used here is sensitive enough to detect even low amounts of truncated transcripts in WT and *xrn4* (Figures 5 and 7). Although we cannot rule out that truncated transcripts could be uridylated

Downloaded from http://nar.oxfordjournals.org/ by guest on July 3, 2013

by URT1 at low levels, we have no indication that URT1 is required for the degradation of truncated transcripts. In addition, all uridylated mRNAs detected in WT and all but 2 clones in *xrn4* (387 of 389 clones when combining data presented in Figures 3, 5 and 7) correspond to adenylated mRNAs and only 2 of 389 clones are deadenylated mRNAs in *xrn4*. This observation strongly supports the idea that the main mRNA substrates of URT1 are mRNAs that are still adenylated, albeit with short poly(A) tails.

By contrast, several lines of evidence support the idea that uridylation could prevent 3′ to 5′ attacks of mRNAs that have been deadenylated to ~10–14 As. As described above, the main substrates of URT1 are mRNAs with short poly(A) tails. In addition, when we characterized uncapped uridylated *LOM1* transcripts in *xrn4*, we observed that these RNAs are 5′ shortened but showed no sign of 3′ trimming (Figure 3). This indicates that uridylated mRNAs are prone to 5′ rather than 3′ attacks, even in a *xrn4* genetic background. A similar observation was made in wild-type plants for uridylated RISC-cleaved transcripts that are 5′ but not 3′ shortened (17). Uridylation also prevents 3′ to 5′ exoribonucleolytic decay in mammalian cell extracts, in addition to promote decapping (18). Both studies concluded that uridylation favours the 5′ to 3′ directionality of RNA decay and supports the idea that uridylation could stabilize 3′-ends (17,18). In fact, U-tracts that terminate the transcription of RNA polymerase III precursors are well known to prevent 3′ to 5′ degradation by promoting the binding of the La protein and the Lsm2-8 heptameric complex in the nucleus (43). Interestingly, a link between uridylation and binding of the cytosolic Lsm1-7 complex also exists. The Lsm1-7 complex recognizes preferentially oligo(A) versus poly(A) tails, but the binding of the Lsm1-7 complex is enhanced by uridylation (18). In line with this, the Lsm1-7 complex is an integral component of the uridylation-mediated decapping pathway (19,20). The current model of this pathway is that uridylation enhances binding of the Lsm1-7 complex (19,20), which recruits the decapping complex via Pat1 (44–48). Importantly, in the context of our study, the Lsm1-7 complex also protects the 3′-ends of oligoadenylated mRNAs from ribonucleolytic attacks (49,50). In yeast *lsm* mutants, mRNAs are trimmed by 10–20 nucleotides at their 3′-ends, a phenotype similar to the one observed for *urt1* mutants. Although the trimming reaction itself is not performed by the exosome, 3′ trimming enhances exosome-mediated 3′ to 5′ decay (50,51). The Lsm1-7 complex could therefore play a dual role in RNA decay, by promoting 5′ to 3′ degradation but also by playing a direct role in inhibiting 3′ ribonucleolytic attacks. Both functions of the Lsm1-7 complex are likely conserved in Arabidopsis. However, the presence of the Lsm1-7 complex is not sufficient to ensure protection of the 3′-ends of oligoadenylated mRNAs in Arabidopsis *urt1* mutants, and the uridylation activity of URT1 is mandatory. This observation implies that preventing the formation of 3′ trimmed and truncated transcripts in Arabidopsis requires an active process, uridylation, highlighting its biological significance.

Degradation is commonly preceded by mRNAs exiting translation (45,48) but mRNA degradation can also occur on polysomes in budding yeast (42). This finding led to the hypothesis that the 5′ to 3′ directionality of RNA decay evolved to allow the last translocating ribosome to complete translation. Our data indicate that uridylation definitely participates in the establishment of 5′ to 3′ polarity of RNA degradation in Arabidopsis by preventing the formation of 3′ trimmed and truncated transcripts. Detecting uridylated mRNAs in polysomal fractions strongly suggests that protection by uridylation operates on polysomes. It is unclear at present what proportion of uridylated mRNAs is associated to polysomes. However, even if this proportion is low, protection of mRNA 3′-end by uridylation could be advantageous to limit the occurrence of truncated transcripts in polysomes. Further work is also required to identify the ribonuclease(s) involved in nibbling the unprotected oligoadenylated mRNAs in *urt1*. In yeast *lsm* mutants, the nibbling enzyme is also unknown but the exosome was shown to degrade 3′ trimmed mRNAs (50). The 3′ truncated transcripts in *urt1* mutants could also become substrate of the 3′ to 5′ degradation pathway, either of the exosome or other cytosolic 3′ to 5′ exoribonucleolytic activities such as SUPPRESSOR OF VARICOSE, a member of the RNase II family associated with RNP granules or RRP6L3, a cytosolic RRP6-like exoribonuclease (52,53). Irrespective of the exact fate of 3′ truncated mRNAs in *urt1* mutants, the most important consequence of our data is that the preventive action of URT1 ensures that such aberrant transcripts are not, or to a much lesser extent, produced in WT plants. Whereas mechanisms of RNA quality control have evolved to detect and eliminate defective transcripts, uridylation is rather part of a preventive process that prevents the occurrence of 3′ trimmed and truncated transcripts.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–2, Supplementary Figures 1–5, Supplementary Methods and Supplementary References [54–61].

ACKNOWLEDGEMENTS

The authors are grateful to Benjamin Stupfler and Barbara Class for technical help.

FUNDING

Centre National de la Recherche Scientifique (CNRS, France); Agence Nationale de la Recherche (ANR, France) [ANR-2010-BLAN-1707 to D.G. and J.M.D.], [ANR-2010-LABX-36 to D.G.] in the frame of the *programme d'investissements d'avenir*]. Funding for open access charge: Centre National de la Recherche Scientifique (CNRS), France.

Conflict of interest statement. None declared.

Downloaded from http://nar.oxfordjournals.org/ by guest on July 3, 2013

REFERENCES

1. Trippe,R., Sandrock,B. and Benecke,B.J. (1998) A highly specific terminal uridylyl transferase modifies the 3′-end of U6 small nuclear RNA. *Nucleic Acids Res.*, **26**, 3119–3126.
2. Wilusz,C.J. and Wilusz,J. (2008) New ways to meet your (3′) end oligouridylation as a step on the path to destruction. *Genes Dev.*, **22**, 1–7.
3. Marzluff,W. (2009) A new way to initiate mRNA degradation. *Nat. Struct. Mol. Biol.*, **16**, 613–614.
4. Norbury,C.J. (2010) 3′ Uridylation and the regulation of RNA function in the cytoplasm. *Biochem. Soc. Trans.*, **38**, 1150–1153.
5. Morozov,I.Y. and Caddick,M.X. (2012) Cytoplasmic mRNA 3′ tagging in eukaryotes: does it spell the end? *Biochem. Soc. Trans.*, **40**, 810–814.
6. Heo,I., Ha,M., Lim,J., Yoon,M.J., Park,J.E., Kwon,S.C., Chang,H. and Kim,V.N. (2012) Mono-uridylation of pre-MicroRNA as a key step in the biogenesis of group II let-7 MicroRNAs. *Cell*, **151**, 521–532.
7. Jones,M.R., Quinton,L.J., Blahna,M.T., Neilson,J.R., Fu,S., Ivanov,A.R., Wolf,D.A. and Mizgerd,J.P. (2009) Zcchc11-dependent uridylation of microRNA directs cytokine expression. *Nat. Cell Biol*, **11**, 1157–1163.
8. Jones,M.R., Blahna,M.T., Kozlowski,E., Matsuura,K.Y., Ferrari,J.D., Morris,S.A., Powers,J.T., Daley,G.Q., Quinton,L.J. and Mizgerd,J.P. (2012) Zcchc11 Uridylates Mature miRNAs to Enhance Neonatal IGF-1 Expression, Growth, and Survival. *PLoS Genet.*, **8**, e1003105.
9. van Wolfswinkel,J.C., Claycomb,J.M., Batista,P.J., Mello,C.C., Berezikov,E. and Ketting,R.F. (2009) CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs. *Cell*, **139**, 135–148.
10. Ibrahim,F., Rymarquis,L.A., Kim,E.J., Becker,J., Balassa,E., Green,P.J. and Cerutti,H. (2010) Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in *Chlamydomonas*. *Proc. Natl Acad. Sci. USA*, **107**, 3906–3911.
11. Kamminga,L.M., Luteijn,M.J., den Broeder,M.J., Redl,S., Kaaij,L.J., Roovers,E.F., Ladurner,P., Berezikov,E. and Ketting,R.F. (2010) Hen1 is required for oocyte development and piRNA stability in zebrafish. *EMBO J.*, **29**, 3688–3700.
12. Ren,G., Chen,X. and Yu,B. (2012) Uridylation of miRNAs by HEN1 SUPPRESSOR1 in *Arabidopsis*. *Curr. Biol.*, **22**, 695–700.
13. Zhao,Y., Yu,Y., Zhai,J., Ramachandran,V., Dinh,T.T., Meyers,B.C., Mo,B. and Chen,X. (2012) The arabidopsis nucleotidyl transferase HESO1 uridylates unmethylated small RNAs to trigger their degradation. *Curr. Biol.*, **22**, 689–694.
14. Yu,B., Yang,Z., Li,J., Minakhina,S., Yang,M., Padgett,R.W., Steward,R. and Chen,X. (2005) Methylation as a crucial step in plant microRNA biogenesis. *Science*, **307**, 932–935.
15. Li,J., Yang,Z., Yu,B., Liu,J. and Chen,X. (2005) Methylation protects miRNAs and siRNAs from a 3′-end uridylation activity in *Arabidopsis*. *Curr. Biol.*, **15**, 1501–1507.
16. Hoefig,K.P., Rath,N., Heinz,G.A., Wolf,C., Dameris,J., Schepers,A., Kremmer,E., Ansel,K.M. and Heissmeyer,V. (2012) Eri1 degrades the stem-loop of oligouridylated histone mRNAs to induce replication-dependent decay. *Nat. Struct. Mol. Biol.*, **20**, 73–81.
17. Shen,B. and Goodman,H.M. (2004) Uridine addition after microRNA-directed cleavage. *Science*, **306**, 997.
18. Song,M.G. and Kiledjian,M. (2007) 3′ Terminal oligo U-tract-mediated stimulation of decapping. *RNA*, **13**, 2356–2365.
19. Mullen,T.E. and Marzluff,W.F. (2008) Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5′ to 3′ and 3′ to 5′. *Genes Dev.*, **22**, 50–65.
20. Rissland,O.S. and Norbury,C.J. (2009) Decapping is preceded by 3′ uridylation in a novel pathway of bulk mRNA turnover. *Nat. Struct. Mol. Biol.*, **16**, 616–623.
21. Morozov,I.Y., Jones,M.G., Razak,A.A., Rigden,D.J. and Caddick,M.X. (2010) CUCU modification of mRNA promotes decapping and transcript degradation in *Aspergillus nidulans*. *Mol. Cell. Biol.*, **30**, 460–469.
22. Morozov,I.Y., Jones,M.G., Gould,P.D., Crome,V., Wilson,J.B., Hall,A.J., Rigden,D.J. and Caddick,M.X. (2012) mRNA 3′ tagging is induced by nonsense-mediated decay and promotes ribosome dissociation. *Mol. Cell. Biol.*, **32**, 2585–2595.
23. Kwak,J.E. and Wickens,M. (2007) A family of poly(U) polymerases. *RNA*, **13**, 860–867.
24. Gazzani,S., Lawrenson,T., Woodward,C., Headon,D. and Sablowski,R. (2004) A link between mRNA turnover and RNA interference in *Arabidopsis*. *Science*, **306**, 1046–1048.
25. Rissland,O.S., Mikulasova,A. and Norbury,C.J. (2007) Efficient RNA polyuridylation by noncanonical poly(A) polymerases. *Mol. Cell. Biol.*, **27**, 3612–3624.
26. Sement,F.M. and Gagliardi,D. (2013) Polydenylation: Methods and Protocols. In: Rorbach,J. (ed). *Methods In Molecular Biology*. Humana Press, Springer Science, in press.
27. Mustroph,A., Juntawong,P. and Bailey-Serres,J. (2009) Isolation of plant polysomal mRNA by differential centrifugation and ribosome immunopurification methods. *Methods Mol. Biol.*, **553**, 109–126.
28. Munoz-Tello,P., Gabus,C. and Thore,S. (2012) Functional implications from the Cid1 poly(U) polymerase crystal structure. *Structure*, **20**, 977–986.
29. Lunde,B.M., Magler,I. and Meinhart,A. (2012) Crystal structures of the Cid1 poly (U) polymerase reveal the mechanism for UTP selectivity. *Nucleic Acids Res.*, **40**, 9815–9824.
30. Yates,L.A., Fleurdepine,S., Rissland,O.S., De Colibus,L., Harlos,K., Norbury,C.J. and Gilbert,R.J. (2012) Structural basis for the activity of a cytoplasmic RNA terminal uridylyl transferase. *Nat. Struct. Mol. Biol.*, **19**, 782–787.
31. Xu,J., Yang,J.Y., Niu,Q.W. and Chua,N.H. (2006) Arabidopsis DCP2, DCP1, and VARICOSE form a decapping complex required for postembryonic development. *Plant Cell*, **18**, 3386–3398.
32. Weber,C., Nover,L. and Fauth,M. (2008) Plant stress granules and mRNA processing bodies are distinct from heat stress granules. *Plant J.*, **56**, 517–530.
33. Narsai,R., Howell,K.A., Millar,A.H., O'Toole,N., Small,I. and Whelan,J. (2007) Genome-wide analysis of mRNA decay rates and their determinants in *Arabidopsis thaliana*. *Plant Cell*, **19**, 3418–3436.
34. Kastenmayer,J.P. and Green,P.J. (2000) Novel features of the XRN-family in *Arabidopsis*: evidence that AtXRN4, one of several orthologs of nuclear Xrn2p/Rat1p, functions in the cytoplasm. *Proc. Natl Acad. Sci. USA*, **97**, 13985–13990.
35. Potuschak,T., Vansiri,A., Binder,B.M., Lechner,E., Vierstra,R.D. and Genschik,P. (2006) The exoribonuclease XRN4 is a component of the ethylene response pathway in *Arabidopsis*. *Plant Cell*, **18**, 3047–3057.
36. Olmedo,G., Guo,H., Gregory,B.D., Nourizadeh,S.D., Aguilar-Henonin,L., Li,H., An,F., Guzman,P. and Ecker,J.R. (2006) ETHYLENE-INSENSITIVE5 encodes a 5′–>3′ exoribonuclease required for regulation of the EIN3-targeting F-box proteins EBF1/2. *Proc. Natl Acad. Sci. USA*, **103**, 13286–13293.
37. Goeres,D.C., Van Norman,J.M., Zhang,W., Fauver,N.A., Spencer,M.L. and Sieburth,L.E. (2007) Components of the *Arabidopsis* mRNA decapping complex are required for early seedling development. *Plant Cell*, **19**, 1549–1564.
38. Rymarquis,L.A., Souret,F.F. and Green,P.J. (2011) Evidence that XRN4, an *Arabidopsis* homolog of exoribonuclease XRN1, preferentially impacts transcripts with certain sequences or in particular functional categories. *RNA*, **17**, 501–511.
39. Gregory,B.D., O'Malley,R.C., Lister,R., Urich,M.A., Tonti-Filippini,J., Chen,H., Millar,A.H. and Ecker,J.R. (2008) A link between RNA metabolism and silencing affecting Arabidopsis development. *Dev. Cell*, **14**, 854–866.
40. Sachs,A.B., Davis,R.W. and Kornberg,R.D. (1987) A single domain of yeast poly(A)-binding protein is necessary and sufficient for RNA binding and cell viability. *Mol. Cell. Biol.*, **7**, 3268–3276.
41. Sherstnev,A., Duc,C., Cole,C., Zacharaki,V., Hornyik,C., Ozsolak,F., Milos,P.M., Barton,G.J. and Simpson,G.G. (2012) Direct sequencing of *Arabidopsis thaliana* RNA reveals patterns of cleavage and polyadenylation. *Nat. Struct. Mol .Biol.*, **19**, 845–852.

Downloaded from http://nar.oxfordjournals.org/ by guest on July 3, 2013

42. Hu,W., Sweet,T.J., Chamnongpol,S., Baker,K.E. and Coller,J. (2009) Co-translational mRNA decay in *Saccharomyces cerevisiae*. *Nature*, **461**, 225–229.
43. Beggs,J.D. (2005) Lsm proteins and RNA processing. *Biochem. Soc. Trans.*, **33**, 433–438.
44. Bouveret,E., Rigaut,G., Shevchenko,A., Wilm,M. and Seraphin,B. (2000) A Sm-like protein complex that participates in mRNA degradation. *EMBO J.*, **19**, 1661–1671.
45. Tharun,S. (2009) Roles of eukaryotic Lsm proteins in the regulation of mRNA function. *Int. Rev. Cell. Mol. Biol.*, **272**, 149–189.
46. Haas,G., Braun,J.E., Igreja,C., Tritschler,F., Nishihara,T. and Izaurralde,E. (2010) HPat provides a link between deadenylation and decapping in metazoa. *J. Cell Biol.*, **189**, 289–302.
47. Ozgur,S., Chekulaeva,M. and Stoecklin,G. (2010) Human Pat1b connects deadenylation with mRNA decapping and controls the assembly of processing bodies. *Mol. Cell. Biol.*, **30**, 4308–4323.
48. Coller,J. and Parker,R. (2005) General translational repression by activators of mRNA decapping. *Cell*, **122**, 875–886.
49. Boeck,R., Lapeyre,B., Brown,C.E. and Sachs,A.B. (1998) Capped mRNA degradation intermediates accumulate in the yeast spb8-2 mutant. *Mol. Cell. Biol.*, **18**, 5062–5072.
50. He,W. and Parker,R. (2001) The yeast cytoplasmic LsmI/Pat1p complex protects mRNA 3′ termini from partial degradation. *Genetics*, **158**, 1445–1455.
51. Tharun,S. (2009) Lsm1-7-Pat1 complex: a link between 3′ and 5′-ends in mRNA decay? *RNA Biol.*, **6**, 228–232.
52. Lange,H., Holec,S., Cognat,V., Pieuchot,L., Le Ret,M., Canaday,J. and Gagliardi,D. (2008) Degradation of a polyadenylated rRNA maturation by-product involves one of the three RRP6-like proteins in Arabidopsis thaliana. *Mol. Cell. Biol.*, **28**, 3038–3044.
53. Zhang,W., Murphy,C. and Sieburth,L.E. (2010) Conserved RNaseII domain protein functions in cytoplasmic mRNA decay and suppresses Arabidopsis decapping mutant phenotypes. *Proc. Natl Acad. Sci. USA*, **107**, 15981–15985.
54. Boyes,D.C., Zayed,A.M., Ascenzi,R., McCaskill,A.J., Hoffman,N.E., Davis,K.R. and Gorlach,J. (2001) Growth stage-based phenotypic analysis of *Arabidopsis*: a model for high throughput functional genomics in plants. *Plant Cell*, **13**, 1499–1510.
55. Lurin,C., Andres,C., Aubourg,S., Bellaoui,M., Bitton,F., Bruyere,C., Caboche,M., Debast,C., Gualberto,J., Hoffmann,B. *et al.* (2004) Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell*, **16**, 2089–2103.
56. Yang,Y.H., Dudoit,S., Luu,P., Lin,D.M., Peng,V., Ngai,J. and Speed,T.P. (2002) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.*, **30**, e15.
57. Smyth,G.K. (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.*, **3**, Article3.
58. Storey,J.D. and Tibshirani,R. (2003) Statistical significance for genomewide studies. *Proc. Natl Acad. Sci. USA*, **100**, 9440–9445.
59. Provart,N. and Zhu,T. (2003) A Browser-based functional classification superviewer for *Arabidopsis* genomics. *Curr. Comput. Mol. Biol.*, **2003**, 271–272.
60. Maere,S., Heymans,K. and Kuiper,M. (2005) BiNGO: a cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics*, **21**, 3448–3449.
61. Gagnot,S., Tamby,J.P., Martin-Magniette,M.L., Bitton,F., Taconnat,L., Balzergue,S., Aubourg,S., Renou,J.P., Lecharny,A. and Brunaud,V. (2008) CATdb: a public access to *Arabidopsis* transcriptome data from the URGV-CATMA platform. *Nucleic Acids Res.*, **36**, D986–D990.

Downloaded from http://nar.oxfordjournals.org/ by guest on July 3, 2013

SUPPLEMENTARY DATA

SUPPLEMENTARY METHODS

mRNA half-life measurements

Half-lives of mRNAs were measured as described in (53) except that diced rosette leaves were used as plant material and that either cordycepin or actinomycin D was used at a final concentration of 1mM and 200µM, respectively. *BAM3*, *MURE*, *BPS1*, *BME3* and *ERF1* mRNA levels were quantified at 0, 15, 30, 45 and 60 min by qRT-PCR analysis. cDNA was synthesized from 1-5 µg of total RNA using SuperScript® III reverse transcriptase (Invitrogen) using 37.5 pmol of oligo$(dT)_{20}$ and 5 ng of random hexamer (Fermentas) per 20 µl reaction according to the manufacturer's instructions. Three technical triplicates and two to four biological replicates were analyzed in a LightCycler® 480 Real-Time PCR System (Roche). Each qPCR reaction contained 1x LightCycler® 480 SYBR Green I Master Mix (Roche), 5 pmol of each primer, 0.2-0.5 µL of cDNA in a volume of 10 µL. *ACT2*, *TIP41* and *EXP* were used as reference mRNAs.

Northern blot analyses

URT1 mRNAs were detected using 20 µg of total RNA extracted from flowers, fractionated using 1.5% agarose 5% formaldehyde gels and transferred to a Hybond-N+ membrane (Amersham) (Supplementary Figure S1). The URT1 probe consisted of a DIG-labelled PCR product corresponding to nucleotides 161:840 of URT1 coding sequence. Hybridization and detection were performed using DIG system kits from Roche according to the manufacturer's recommendations. A similar procedure was followed for northern blot analyses presented in Supplementary Figures S4-5 except that [α-^{32}P]-dCTP random-labeled probes were used. Final stringent washes were performed at 68°C with 0.1x SSC, 0.1%(w/v) SDS.

Microarray analysis

Microarray analysis was carried out at the Unité de Recherche en Génomique Végétale (Evry, France), using the CATMAv6.1 array based on Roche-NimbleGen technology. A single high density CATMAv6.1 microarray slide contains twelve chambers, each containing 135,000 primers representing all the *Arabidopsis thaliana* genes: 30,834 probes corresponding to TAIRv8 annotation (including 476 probes of mitochondrial and chloroplast genes) + 1289 probes corresponding to EUGENE software predictions. The array also includes 5,352 probes corresponding to repeat elements, 658 probes for miRNA/MIR, 342 probes for other RNAs (rRNA, tRNA, snRNA, soRNA) and 36 controls. Each long primer is triplicated in each chamber for robust analysis. Three independent biological replicates were produced. For each biological repetition, rosette leaves were collected by pooling rosette leaves from five plants (growth stage 6.5) (54). Total RNA was extracted using TRI-reagent (MRC). For each comparison, one technical replicate with fluorochrome reversal was performed for each biological replicate (*i.e.* six hybridizations per comparison). The labeling of cRNAs with Cy3-dUTP or Cy5-dUTP (Perkin-Elmer-NEN Life Science Products) and the hybridization to the slides were performed as described in (55). Two micron scanning was performed with InnoScan900

scanner (Innopsys^{R}, Carbonne, FRANCE) and raw data were extracted using Mapix^{R} software (Innopsys^{R}, Carbonne, FRANCE).

Statistical analysis of microarray data

Experiments were designed with the statistics group of the Unité de Recherche en Génomique Végétale (Evry, France). For each array, the raw data comprised the logarithm of median feature pixel intensity at wavelengths 635 nm (red) and 532 nm (green). A global intensity-dependent normalization using the loess procedure (56) was performed for each array to correct the dye bias. The differential analysis is based on the log-ratios averaging over the duplicate probes and over the technical replicates. Hence, the number of available data for each gene equals the number of biological replicates and are used to calculate the moderated t-test (57). Under the null hypothesis, no evidence that the specific variances vary between probes is highlighted by Limma and consequently the moderated t-statistic is assumed to follow a standard normal distribution. To control the false discovery rate, adjusted p-values found using the optimized FDR approach of (58) are calculated. We considered as being differentially expressed the probes with an adjusted p-value ≤ 0.05. Analysis was done with the R software (http://www.R-project.org). The function SqueezeVar of the library Limma was used to smooth the specific variances by computing empirical Bayes posterior means. The library Kerfdr was used to calculate the adjusted p-values.

Differentially expressed genes were assigned to a MapMan category using the Arabidopsis Classification Superviewer at http://bar.utoronto.ca/ntools/cgi-bin/ntools_classification_superviewer.cgi (59) (Supplementary Table S2). In addition, a Gene Ontology analysis was performed using the BiNGO plugin (60) to identify GO categories that are significantly overrepresented in the sets of down and up-regulated genes (Supplementary Table S2). Microarray data were deposited at Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/), accession no. GSE38902) and at CATdb (http://urgv.evry.inra.fr/CATdb/; Project: RS11-10_urt1) (61).

SUPPLEMENTARY REFERENCES

53. Zhang, W., Murphy, C. and Sieburth, L.E. (2010) Conserved RNaseII domain protein functions in cytoplasmic mRNA decay and suppresses Arabidopsis decapping mutant phenotypes. *Proc Natl Acad Sci U S A*, **107**, 15981-15985.
54. Boyes, D.C., Zayed, A.M., Ascenzi, R., McCaskill, A.J., Hoffman, N.E., Davis, K.R. and Gorlach, J. (2001) Growth stage-based phenotypic analysis of Arabidopsis: a model for high throughput functional genomics in plants. *Plant Cell*, **13**, 1499-1510.
55. Lurin, C., Andres, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyere, C., Caboche, M., Debast, C., Gualberto, J., Hoffmann, B. *et al.* (2004) Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell*, **16**, 2089-2103.
56. Yang, Y.H., Dudoit, S., Luu, P., Lin, D.M., Peng, V., Ngai, J. and Speed, T.P. (2002) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res*, **30**, e15.
57. Smyth, G.K. (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*, **3**, Article3.
58. Storey, J.D. and Tibshirani, R. (2003) Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A*, **100**, 9440-9445.
59. Provart, N. and Zhu, T. (2003) A Browser-based Functional Classification SuperViewer for Arabidopsis Genomics. *Currents in Computational Molecular Biology*, **2003**, 271-272.
60. Maere, S., Heymans, K. and Kuiper, M. (2005) BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics*, **21**, 3448-3449.
61. Gagnot, S., Tamby, J.P., Martin-Magniette, M.L., Bitton, F., Taconnat, L., Balzergue, S., Aubourg, S., Renou, J.P., Lecharny, A. and Brunaud, V. (2008) CATdb: a public access to Arabidopsis transcriptome data from the URGV-CATMA platform. *Nucleic Acids Res*, **36**, D986-990.

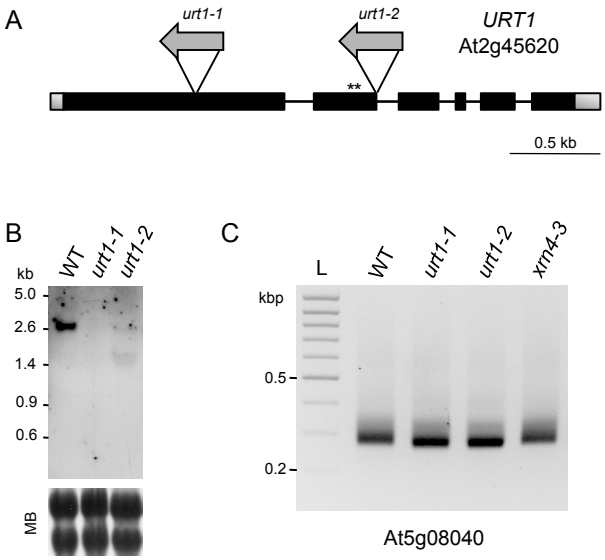


Figure S1. Characterization of *urt1-1* and *urt1-2* mutants. (**A**) Schematic representation of *URT1* gene. 5' and 3' UTRs are indicated by grey boxes, exons by black boxes and introns by solid lines. T-DNA insertions in *urt1-1* and *urt1-2* mutants were confirmed by PCR amplification and sequencing. T-DNA position and orientation are show by grey arrows, the left border is represented by the arrowhead. Positions of the mutations introduced to produce the catalytically-inactive URT1$^{D491/3A}$ mutant protein are shown by asterisks. (**B**) Both *urt1* T-DNA insertions affect *URT1* mRNA expression. Northern analysis was performed using 20 µg of total RNA and a DIG-labelled probe corresponding to the 5' part of the gene. Equal loading is shown by methylene blue (MB) staining of the membrane. (**C**) The mRNA 3' shortening detailed in the text is observed in both *urt1-1* and *urt1-2* as illustrated for At5g08040 transcripts. Genotypes are indicated on top of the panel corresponding to a negative image of an ethidium bromide stained 2% agarose gel. The migration of a DNA ladder (L) (1 to 0.2 kbp, 0.1 kbp size increment) is shown on the left. A slight reduction in the size of the main amplification product due to enhanced deadenylation (see main text) is observed in both *urt1-1* and *urt1-2* mutants as compared to WT and *xrn4*. The sequencing data presented throughout the text were obtained in *urt1*-1.

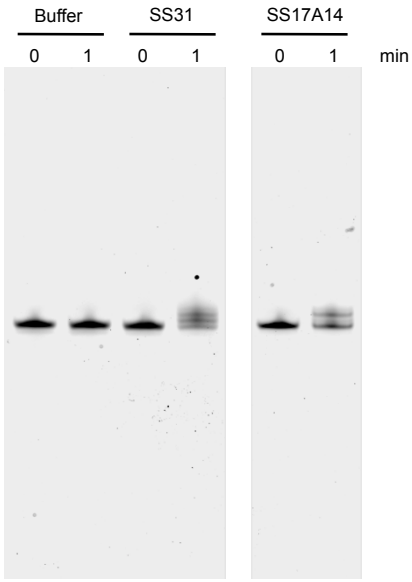


Figure S2. Distributive activity of URT1 on 31 nucleotide RNA substrates. GST-URT1 was incubated for 1 min with 5'-labeled RNA substrates corresponding either to a 31 nucleotide single-stranded synthetic RNA with no oligo(A) tail (SS31) or with a 31 nucleotide synthetic RNA containing 14 As (SS17A14). The sequences of SS31 and SS17A14 are CCCCACCACCAUCACUUCACCACCAUCACUU and CCCCACCACCAUCACUUAAAAAAAAAAAAAA, respectively. Reaction products were separated on 17% denaturing polyacrylamide gels. Incubation time is indicated in minutes at the top of panels. For both SS31 and SS17A14, URT1's activity is distributive for the first added nucleotides.

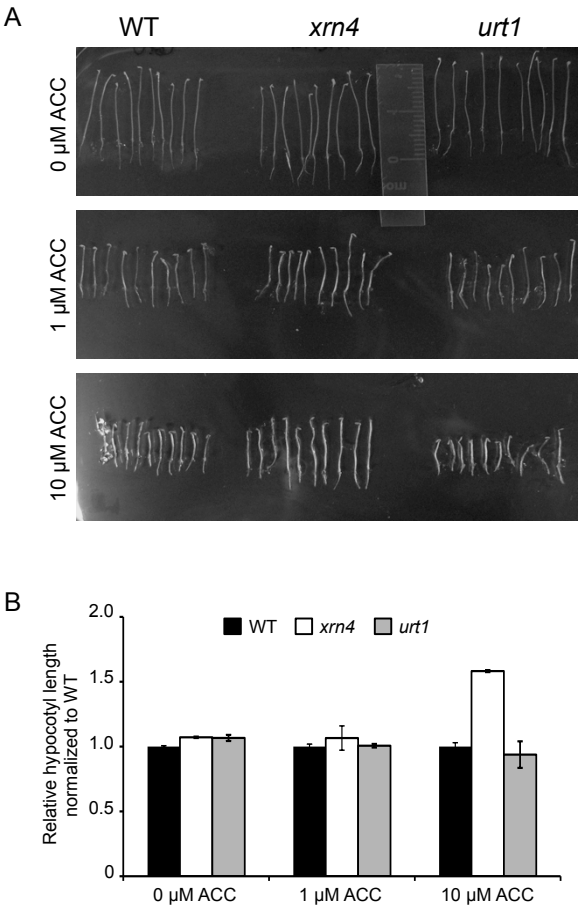


Figure S3. The ethylene-insensitive phenotype of *xrn4* is not observed in *urt1*. (**A**) Phenotypes of 3 day-old etiolated seedlings of WT, *xrn4* and *urt1* grown in absence or presence of 1 or 10 µM of the ethylene precursor 1-aminocyclopropane-1-carboxylic acid (ACC) as described in (35). (**B**) Hypocotyl length measurements of 3 day-old etiolated WT, *xrn4* and *urt1* seedlings germinated in presence of the indicated ACC concentrations. For all genotypes, average hypocotyl lengths were normalized as compared to WT and data are presented as means ± SEM, n=2.

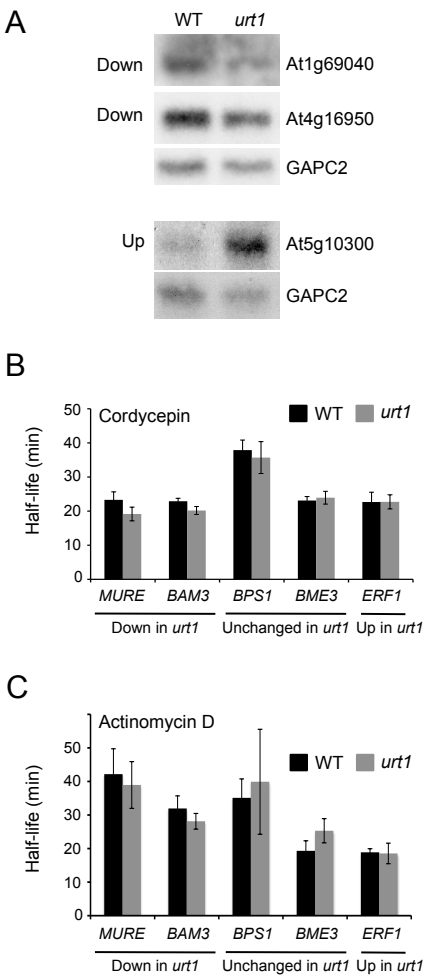


Figure S4. URT1 has no significant impact on degradation rates of mRNAs.
(**A**) Validation of microarray data for selected mRNAs. Northern blot analysis is shown for three mRNAs that have been selected amongst the most misregulated genes in *urt1* (see Supplementary Table S2 for microarray results) and whose expression is detectable in rosette leaves by northern analysis. *GAPC2* transcripts were used as a loading control. (**B**) URT1 has no significant impact on mRNA half-lives regardless of respective misregulation of steady-state levels in *urt1*. *MURE* (At1g63680) and *BAM3* mRNAs are down-regulated, BPS1 (At1g01550) and BME3 (At3g54810) mRNA levels are unchanged and ERF1 (At4g17500) is upregulated in *urt1* (detailed in Supplementary Table S2 except for the low-expressed *BAM3* transcripts analyzed by qPCR). Half-lives (min) were determined in rosette leaves following cordycepin inhibition of transcription and qRT-PCR analysis. Data are presented as means ± SD (n=4). (**C**) Same legend as B except that transcription was inhibited by actinomycin D. Data are presented as means ± SD (*BPS1* n=2, all other genes n=3).

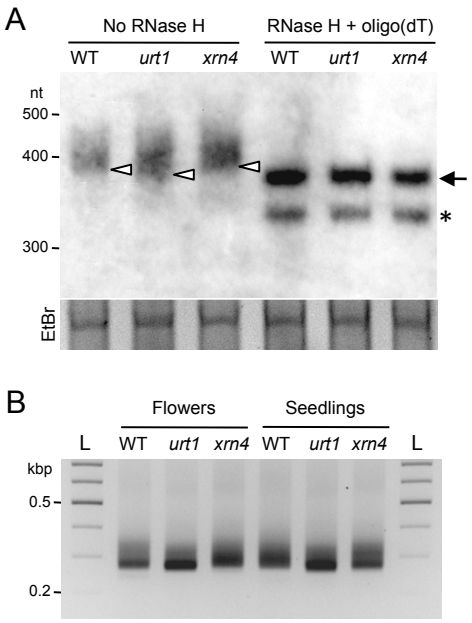


Figure S5. Poly(A) tails and 3' shortening of At5g08040 transcripts in *urt1*. (**A**) 6% acrylamide gel northern analysis of At5g08040 transcripts using total RNA extracted from flowers. RNA were either not treated with RNase H or incubated with RNase H and oligo(dT) to remove poly(A) tails as indicated on the top of the panel. The white arrowheads indicate the bottom of the smears corresponding to adenylated mRNAs whereas the migration of completely deadenylated mRNAs observed upon digestion with RNase H in the presence of oligo(dT) is indicated by a black arrow. The asterisk marks an artefactual band probably due to a partial cleavage by RNase H in a A-rich region within the mRNA sequence. The position of RNA standards run alongside the samples is indicated on the left. A portion of the 6% acrylamide gel stained with ethidium bromide (EtBr) is shown for loading and migration control. (**B**) 3' RACE analysis showing the 3' shortening of At5g08040 transcripts in *urt1*. Genotypes and tissues are indicated on top of panels corresponding to negative images of ethidium bromide stained 2% agarose gels. 3' RACE PCR products were run alongside a DNA ladder (L) with 100 bp size increments.

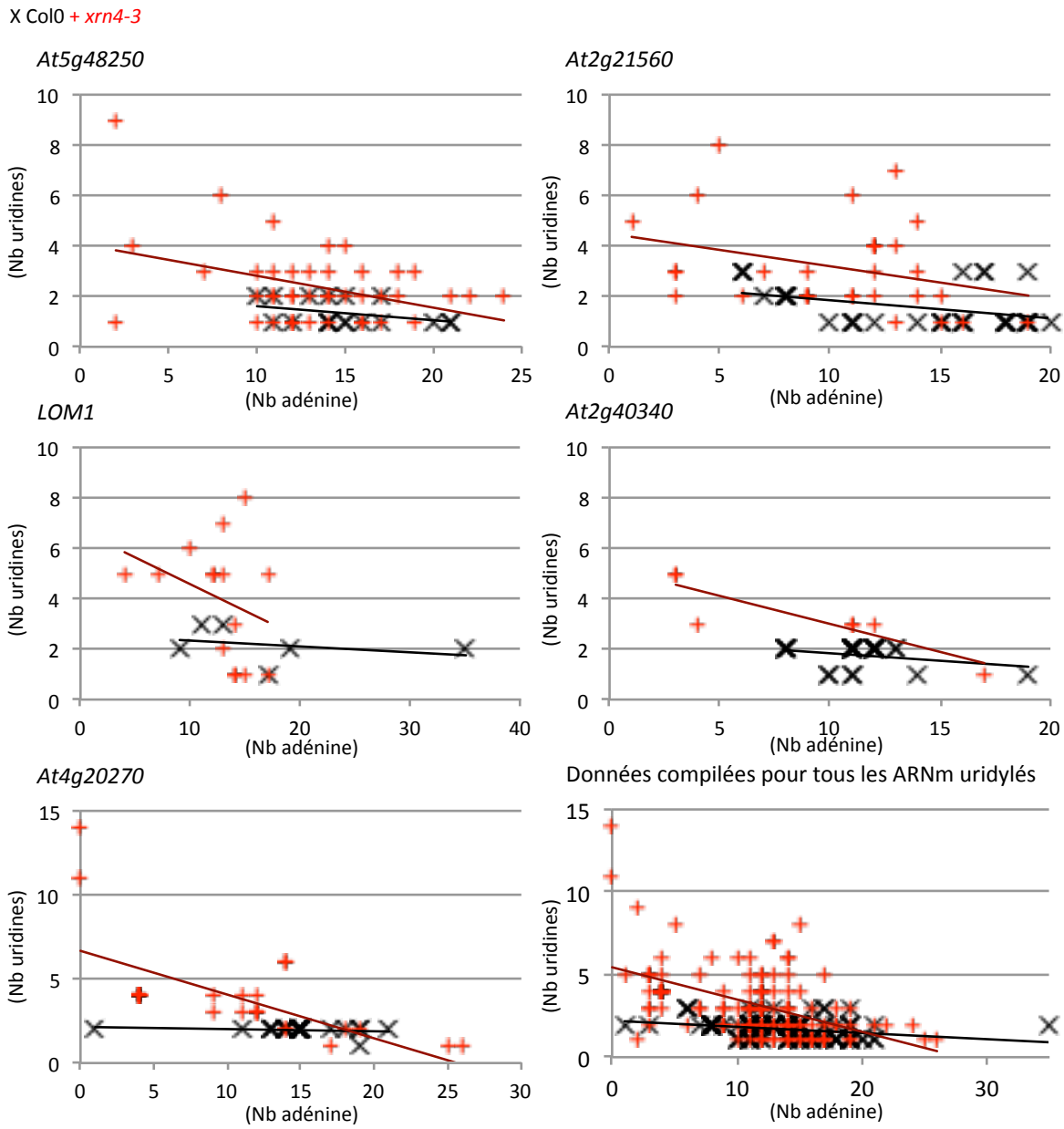


Figure 2A.1 : Le nombre d'uridines ajoutées en 3' est plus important lorsque le nombre d'adénines est faible pour les séquences uridylées issues de l'analyse des ARNm dans le mutant *xrn4-3*

Le nombre d'uridines en fonction du nombre d'adénines est mise en évidence par un graphique pour chacune des séquences uridylés pour les cinq ARNm étudiés dans l'article ainsi qu'une compilation de ces cinq ARNm pour l'écotype Col0 (n=122) et le mutant *xrn4-3* (n=107).

2. Informations complémentaires concernant les ARNm uridylés

Dans cette partie, j'apporte des informations supplémentaires sur les modifications en 3' des ARNm totaux et des ARNm associés aux polysomes par rapport aux informations données dans l'article. L'information importante donnée dans l'article est que les ARNm uridylés par URT1 ont en moyenne 15 adénines et qu'ils sont détectables au niveau des polysomes. Par ailleurs, les ARNm uridylés ont une à trois uridines chez Col0 et une quantité plus importante dans xrn4-3, allant jusqu'à 14 uridines. Les informations complémentaires données dans le cas des ARNm associés aux polysomes sont préliminaires en raison du trop faible nombre de séquences analysées mais révèlent de nouvelles hypothèses qu'il faudra confirmer.

a) Analyse des extensions uridylées des ARNm

Les ARNm uridylés de *xrn4-3* possèdent statistiquement une quantité d'uridines plus grande ajoutée en 3' (en moyenne 2,6 pour n=107) que les ARNm uridylés chez Col0 (en moyenne 1,6 pour n=122) (Test de Mann Whitney $p<0,001$). J'ai mis en relation la quantité d'uridines ajoutée en 3' avec le nombre d'adénines présentes en 3' des ARNm pour chacun des ARNm étudiés et une compilation des cinq ARNm (Figure 2A.1). Pour chacun des cinq ARNm étudiés, les ARNm uridylés ayant le plus d'uridines ajoutées en 3' (au maximum 14) correspondent majoritairement à des ARNm n'ayant pas ou très peu d'adénines en 3'. Il existe donc des séquences uridylées présentes dans le mutant *xrn4-3* qui ont une quantité d'adénines précédent l'uridylation très faible (au minimum 0) qui ne sont faiblement représentées chez Col0, à l'exception d'une séquence d'ARNm uridylée n'ayant qu'une adénine. Les ARNm ayant plus de trois uridines en 3' peuvent être un signal de dégradation pour la voie 3'-5'. En effet, la voie de dégradation du sens 5'-3' est principalement effectuée par l'exoribonucléase 5'-3' XRN4. Lorsque la voie de dégradation 5'-3' n'est pas suffisante (comme dans le mutant *xrn4-3*), elle peut être compensée par la voie de dégradation 3'-5' (Discussion partie D3).

b) Analyse des extensions uridylées des ARNm associés aux polysomes

Afin de vérifier si les observations faites avec les séquences uridylées des ARN totaux sont aussi valables avec les ARNm associés aux polysomes, j'ai aussi analysé les séquences des ARNm *BAM3* et *At5g48250* associés aux polysomes

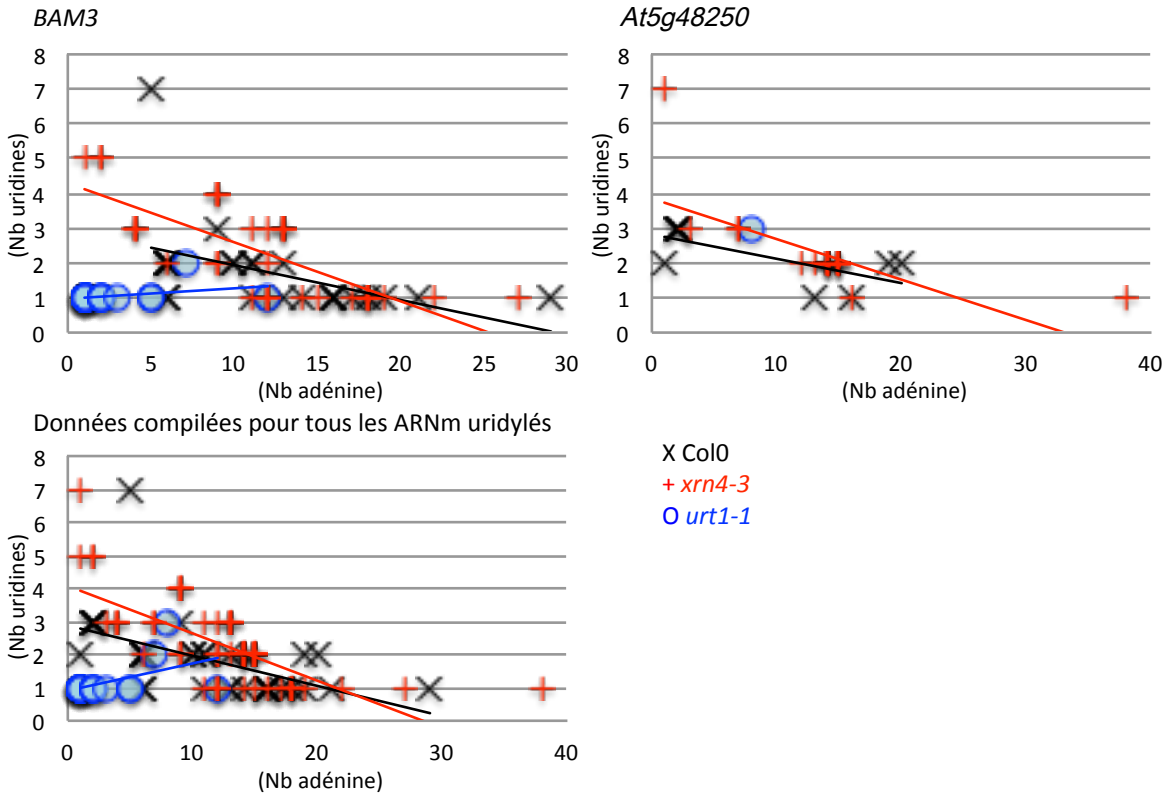


Figure 2A.2 : Le nombre d'uridines ajoutées en 3' est plus important lorsque le nombre d'adénines est faible pour les ARNm associés aux polysomes dans le mutant *xrn4-3*

Le nombre d'uridines en fonction du nombre d'adénines est mise en évidence par un graphique pour chacune des séquences uridylés pour les deux ARNm associés aux polysomes étudiés dans l'article ainsi qu'une compilation de ces deux ARNm pour l'écotype Col0 (n=33 séquences uridylées/224 séquences totales), le mutant *xrn4-3* (n=56 séquences uridylées/245 séquences totales) et *urt1-1* (n= 17 séquences uridylées/279 séquences totales).

présentés dans l'article en reliant le nombre d'uridines ajoutées en 3' des ARNm selon la quantité d'adénine présente en 3'.

Dans l'écotype sauvage Col0, 20% des ARNm polysomaux de *BAM3* et 6% de *At5g48250* séquencés sont uridylés. La majorité des modifications en 3' consiste en l'ajout d'une à trois uridines en 3' lorsque l'ARNm ne possède plus que 9 adénines en moyenne (11 pour *BAM3* et 5 pour *At5g48250*) (Figure 2A.2). Pour *BAM3*, une séquence (1/24) contient sept uridines successives lorsque l'ARNm possède seulement cinq adénines.

Chez le mutant *urt1-1*, en absence de URT1, le taux d'uridylation passe de 20% à 11% et de 9% à 1%, respectivement pour *BAM3* et *At5g48250*. Le taux d'uridylation de *BAM3* est surement surestimé car la complexité des séquences uridylées est faible, par exemple 9 séquences de *BAM3* contiennent une uridine lorsque l'ARNm n'a qu'une adénine en 3'. L'uridylation résiduelle dans *urt1-1* n'est pas effectuée par URT1 donc une autre uridylyltransférase est aussi capable d'uridyler les ARNm, même si URT1 est l'uridylyltransférase majoritaire dans l'écotype Col0. L'uridylation résiduelle chez *urt1-1* montre qu'il existe une secodne uridylyltransférase capable d'uridyler les ARNm ayant une spécificité de substrat différente de URT1. En effet, les modifications apparaissent lorsque les ARNm ont une queue polyA quasi inexistante, plus que 3 adénines en moyenne en 3' (Figure 2A.2). En effet, pour l'ARNm *BAM3*, l'uridylation prend place lorsque les ARNm n'ont plus que 2 adénines en moyenne et 8 adénines pour *At5g48250* (une seule séquence).

Le mutant *xrn4-3* possède un taux d'uridylation plus élevé que l'écotype sauvage Col0, respectivement 26% et 20% pour *BAM3* et 19% et 9% pour *At5g48250*. De plus, les queues uridylées observées chez *xrn4-3* sont plus grandes que Col0, allant jusqu'à huit uridines successives La taille moyenne des queues polyA des ARNm uridylés de *BAM3* ne sont pas significativement différentes entre l'écotype sauvage Col0 et le mutant *xrn4-3* (Test de Mann-Whitney p=0,98),

respectivement de 9 et 10 adénines. Cependant, les tailles des queux polyA des ARNm uridylés de Col0 et *xrn4-3* sont différentes du mutant *urt1-1* (3 adénines). Cependant, pour l'ARNm *At5g48250*, les tailles des queues polyA uridylées ne sont pas statistiquement significatives entre les différents génotypes. Ce qui peut être expliqué par un nombre de séquences uridylées trop faible. En effet, dans cette étude, une seule séquence est uridylée dans le mutant *urt1-1*.

3. Conclusion et ouverture

D'après l'analyse des séquences uridylés selon la quantité d'uridines ajoutées, les ARNm ayant un nombre d'uridines plus important en 3' possèdent un nombre d'adénine en 3' plus faible (Figure 2A.2). D'une manière intéressante, la tendance observée pour les séquences uridylées de *xrn4-3* des ARNm totaux étudiés reste valable pour les ARNm associés aux polysomes. Cependant, pour *BAM3*, dans l'écotype Col0, une séquence possède sept uridines lorsque l'ARNm ne possède plus que cinq adénines. Donc le phénomène observé chez *xrn4-3* pourrait aussi intervenir dans l'écotype Col0 mais à une fréquence beaucoup plus faible (Discussion partie D3). Le rôle de ces oligouridines n'est pas encore découvert.

D'autre part, il serait intéressant de pouvoir comparer l'uridylation globale des ARNm, selon leur association ou non aux polysomes et étudier leur capacité de traduction.

B- URT1 n'est pas intrinsèquement spécifique envers son substrat ARN

J'ai effectué plusieurs expériences complémentaires de l'étude catalytique de URT1 présentée dans la partie A afin de mieux comprendre le mécanisme catalytique de URT1. J'ai effectué une modélisation de la structure 3D du domaine nucléotidyltransférase de URT1 à partir de la structure résolue au niveau atomique de Cid1 et j'ai transféré les ligands d'intérêts (UTP, magnésium et ARN) dans le site actif de la modélisation. Cette modélisation de structure a permis de mieux comprendre certaines des caractéristiques catalytiques observées *in vitro*. Dans les paragraphes suivants, je présente en détail les étapes de la vérification de la fiabilité de la modélisation que j'ai réalisée.

1. Modélisation de la structure du domaine nucléotidyltransférase de URT1 avec un transfert des ligands Mg^{2+} et UTP dans le site actif

Les domaines nucléotidyltransférase de URT1 et Cid1 sont identiques à 31% et similaires à 53%. La structure 3D de Cid1 a récemment été résolue au niveau atomique (Lunde et al., 2012; Munoz-Tello et al., 2012; Yates et al., 2012). Quant une protéine ou seulement un de ses domaines est homologue à une protéine dont la structure a été précédemment résolue, une modélisation de structure 3D peut être effectuée par homologie. La structure de Cid1 a été résolue en présence de différents nucléotides (UTP, GTP, CTP, 3'dATP et ATP) et d'ions, i.e. magnésium ou calcium, donnant plusieurs structures 3D de Cid1. Une structure de Cid1 contenant un substrat ARN sera prochainement publiée (communication personnelle du Dr Thore).

La famille des ribonucléotidyltransférases est connue pour avoir un site actif très conservé (Martin and Keller, 2007). Plusieurs uridylyltransférases de cette famille, respectivement responsable de l'uridylation des ARN dans le cytoplasme ou de l'édition dans les mitochondries chez *Trypanosoma brucei* (*Tb*TUT4 et *Tb*RET2) ont été résolues en présence d'uridine-5'-monophosphate (U5P), afin de mimer le début du substrat ARN. Ces structures ont été résolues aussi en présence d'UTP et de magnésium afin de représenter un état pré-catalytique référencées PDB ID : 2Q0F pour *Tb*TUT4 (Stagno et al., 2007) et PDB ID : 2B56 pour *Tb*RET2 (Deng et al., 2005).

Suite à nos observations sur l'homologie de URT1 avec Cid1, la conservation de structure 3D des ribonucléotidyltransférases et l'étude catalytique de URT1, j'ai

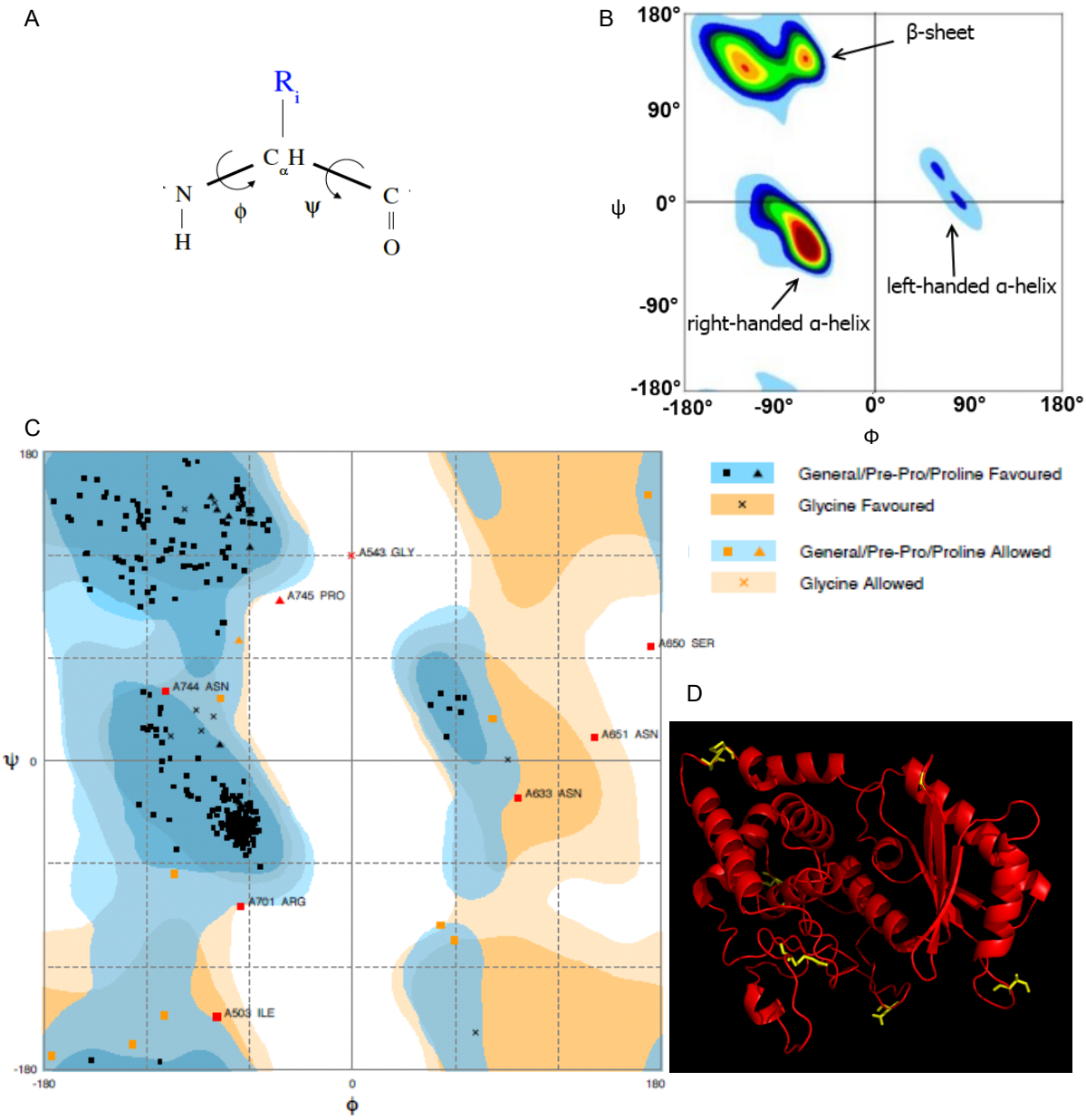


Figure 2B.1 : Représentation du diagramme de Ramachandran de la structure 3D modélisée du domaine nucléotidyltransférase de URT1

A Les angles de torsion phi et psi (Φ, ψ) d'un acide aminé sont représentés, Ri est la chaine latérale. **B** Représentation graphique de la conformation stérique du squelette peptidique d'une protéine, les structures secondaires principalement observées sont indiquées. **C** Diagramme de Ramachandran du domaine nucléotidyltransférase de URT1 modélisé à partir de Cid1 (PDB ID : 4FH5) (Yates et al., 2012) avec les régions favorables (en bleu ou orange pour les glycines), les régions autorisées (en bleu clair ou orange pale pour les glycines) et les régions non autorisées. Ce diagramme est effectué à l'aide du logiciel RAMPAGE (http://mordred.bioc.cam.ac.uk) (voir détails en Annexe 3.1). **D** Représentions par PyMOL des acides aminés situés dans une région non autorisée (en jaune) : G543, S650, N651, N633, R701, N744 et P745, localisée sur la structure 3D du domaine nucléotidyltransférase modélisée de URT1 (en rouge).

privilégié une modélisation par homologie de structure du domaine nucléotidyltransférase de URT1 à partir de Cid1 résolue en présence d'UTP et de magnésium. Par la suite, j'ai transféré les ligands UTP et magnésium à partir de la structure de Cid1 et le substrat ARN à partir de *Tb*TUT4 qui est plus proche de URT1 que *Tb*RET2. J'ai étudié la spécificité de substrat ARN de URT1 par comparaison avec Cid1, *Tb*TUT4 et *Tb*RET2. D'une façon intéressante, Cid1 n'a pas de spécificité intrinsèque de substrat ARN, tandis que *Tb*TUT4 et *Tb*RET2 sont intrinsèquement spécifique d'un substrat ARN différent (Deng et al., 2005; Lunde et al., 2012; Munoz-Tello et al., 2012; Stagno et al., 2007; Yates et al., 2012). En effet, *Tb*TUT4 est spécifique d'un ARN ayant en 3' une uridine, contrairement à *Tb*RET2 qui est spécifique d'un ARN ayant une adénine en 3' (Deng et al., 2005; Stagno et al., 2007). La comparaison des structures de Cid1, *Tb*TUT4 et *Tb*RET2 permettront de poser une hypothèse sur la spécificité de substrat ARN de URT1.

a) Modélisation du domaine nucléotidyltransférase de URT1 à partir de Cid1 par SWISS-MODEL

J'ai utilisé le mode automatique du serveur de SWISS-MODEL pour modéliser la structure du domaine nucléotidyltransférase (http://swissmodel.expasy.org/) (Kiefer et al., 2009; Schwede et al., 2003) (voir matériel et méthodes) à partir de la structure atomique modèle de Cid1 référencée dans la Protein Data Bank (Berman et al., 2000) PDB ID : 4FH5 des acides aminés de Cid1 de 33 à 377 (Lunde et al., 2012). La modélisation obtenue du domaine nucléotidyltransférase de URT1 comprend les résidus 433 à 752 avec une énergie finale de -14745 KJ/mol.

b) Validation du modèle généré par SWISS-MODEL et transfert des ligands (UTP, magnésium et ARN)

J'ai vérifié la fiabilité du modèle généré par SWISS-MODEL par une évaluation structurale et une analyse stéréochimique grâce au Z-score, au QMEANscore (Benkert et al., 2008) et au diagramme de Ramachandran (Ramachandran, 1963; RAMACHANDRAN et al., 1963). Puis, dans un second temps, j'ai vérifié la validité du transfert des ligands dans le site actif.

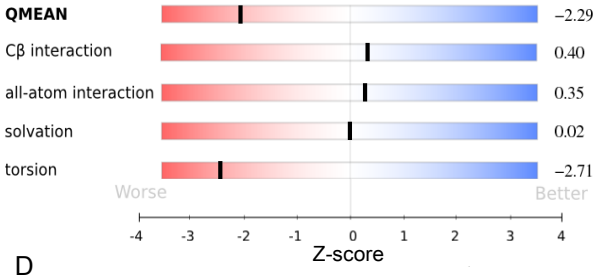
1. Diagramme de Ramachandran

Les acides aminés sont formés d'un carbone α entouré d'un groupe carboxyle et d'un groupe amine, ainsi qu'une chaîne latérale (Ri). Pour chaque acide aminé la chaîne principale possède deux angles dièdres flexibles ϕ et ψ (Figure 2B.1A). Ces deux angles sont restreints géométriquement. Leur valeur rend compte de la possibilité géométrique des angles de torsion du squelette peptidique de chaque

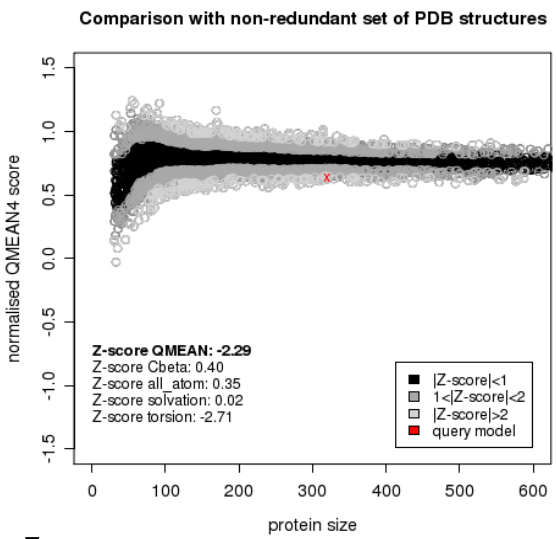
A

	Raw score	Z-score
Interaction des Cβ	-159,20	0,40
Interaction des atomes deux à deux	-9975,33	0,35
Solvatation	-32,75	0,02
Angles de torsion	-35,61	-2,71
QMEAN4 score	0,634	-2,29

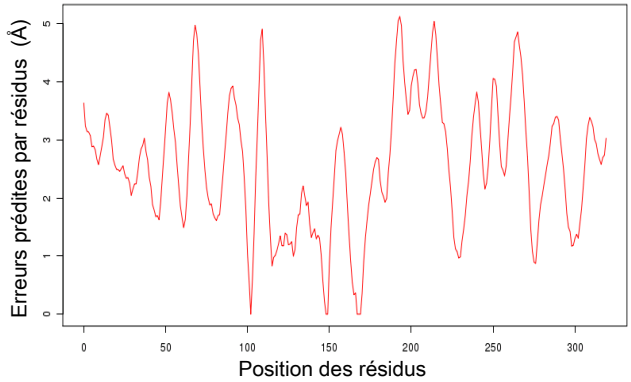
B



C



D



E

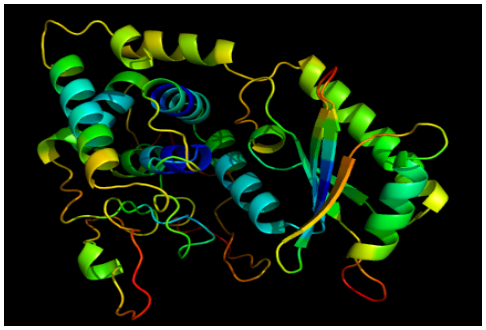


Figure 2B.2 : Validation statistique du modèle de URT1 et mise en évidence des erreurs prédites par résidus

A Détails du score de la modélisation du domaine nucléotidyltransférase de URT1 effectuée par SWISS-MODEL (Kiefer et al., 2009; Schwede et al., 2003) à partir de Cid1 (PDB ID : 4FH5) (Yates et al., 2012). **B** Représentation graphique du Z-score, plus le score est proche de 0 plus le modèle est considéré comme fiable, la barre noire représente la valeur du modèle généré sur une échelle de couleur allant de rouge (non fiable) à bleu (fiable). **C** Comparaison des QMEAN4 scores (Benkert et al., 2008) des protéines cristallisées de la même taille (319 acides aminés) avec celui obtenu par le modèle généré. **D** Erreurs structurales prédites par résidu de la structure effectuée par SWISS-MODEL (http://swissmodel.expasy.org/) (Kiefer et al., 2009; Schwede et al., 2003). **E** Représentation 3D du modèle coloré selon la fiabilité des résidus par PyMol, en allant de bleu pour les résidus fiables à rouge pour les non fiables.

acide aminé car toutes les valeurs de ces angles ne sont pas stériquement possibles. Une étude des combinaisons admissibles de ces angles a permis la réalisation d'un diagramme, appelé le diagramme de Ramachandran (RAMACHANDRAN et al., 1963). C'est une représentation graphique permettant d'analyser la conformation stérique du squelette peptidique d'une protéine, les angles ϕ en abscisse et ψ en ordonné (compris entre -180° et +180°). Le diagramme de Ramachandran comporte trois zones favorables, les deux premières correspondant aux structures secondaires principalement observées (feuillets β dans la partie supérieure gauche et hélices α, plus au centre). La troisième zone représente les conformations en hélice gauche (ψ>0) (Figure 2B.1B). Sachant que les glycines sont plus libres car elles ne possèdent pas de chaîne latérale, elles peuvent se situer en dehors des régions favorables du diagramme.

Le diagramme de Ramachandran de la modélisation de la structure du domaine nucléotidyltransférase est effectué à l'aide du logiciel RAMPAGE (http://mordred.bioc.cam.ac.uk). Il représente les angles de torsion des acides aminés et met en évidence graphiquement ceux qui sont situés dans une région favorable (300/319 soit 94%), une région autorisée (10/319 soit 3%) et une région non autorisée (8/319 soit 2,5%). Ces 8 derniers acides aminés sont I503, G543, S650, N651, N633, R701, N744 et P745 (Figure 2B.1C). Je les ai représentés au niveau la structure 3D modélisée du domaine nucléotidyltransférase de URT1. Ils sont localisés plus à l'extérieur de la structure 3D, au niveau des boucles et non du site actif (Figure 2B.1D). Donc ils ne gêneront pas au transfert des ligands ni à l'étude du site actif. De plus, ils sont conservés chez Cid1 (G152, S262, N164, R323, I368 et P369), à l'exception de N633 qui correspond au G247 de Cid1. D'après cette étude la modélisation de structure du domaine nucléotidyltransférase de URT1 est suffisamment fiable.

2. QMEANscore

Dans un second temps, j'ai effectué une autre étude statistique pour valider l'utilisation du modèle généré à l'aide du QMEANscore4 qui est un score global reflétant la fiabilité de la structure. Ce score est calculé grâce aux Z-scores de quatre paramètres : l'énergie d'interaction entre les carbones β, l'énergie des atomes deux à deux, l'énergie de solvatation ainsi que l'énergie de torsion des angles. Le Z-score est l'erreur quadratique obtenue par comparaison avec les valeurs expérimentales issues de la résolution atomique par rayon X de protéines de tailles similaires à celle du modèle généré (Benkert et al., 2008).

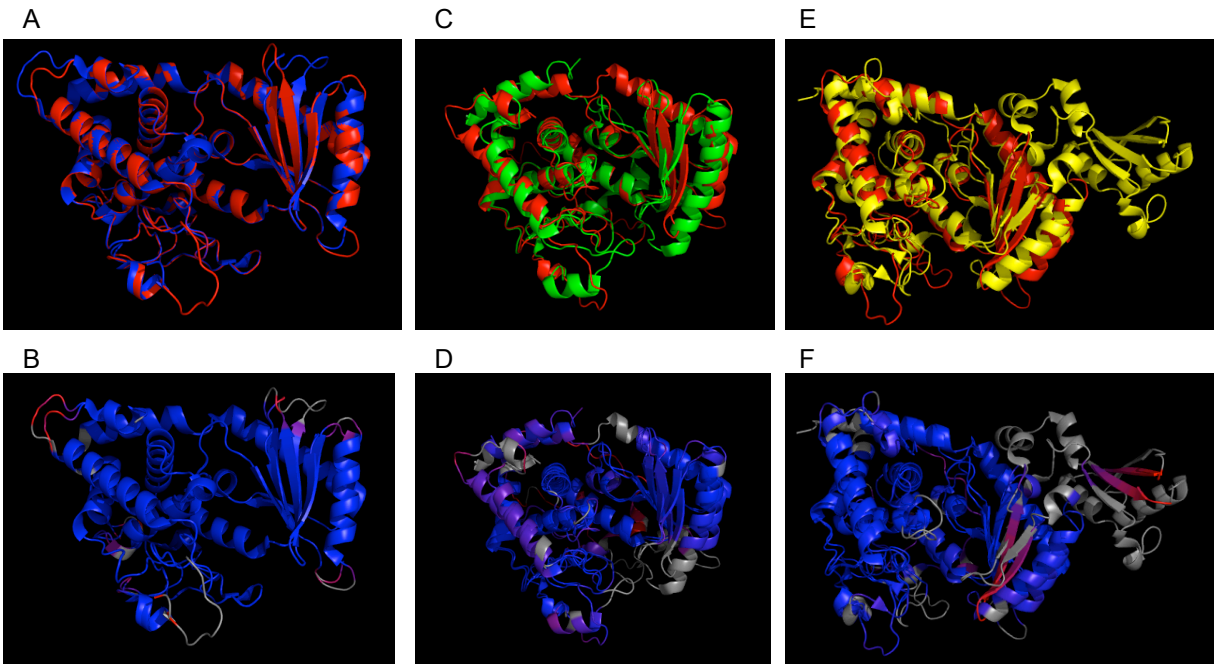


Figure 2B.3 : Alignement 3D du modèle du domaine nucléotidyltransférase de URT1 et des trois autres uridylyltransférases résolues au niveau atomique

A à F Représentation de Cid1, *Tb*TUT4 et *Tb*RET2 alignés deux à deux avec URT1 par PyMol. URT1 et Cid1 (**A** et **B**), URT1 et *Tb*TUT4 (**C** et **D**), URT1 et *Tb*RET2 (**E** et **F**). **A** URT1 est représentée en rouge, Cid1 (PDB ID :4FH5) (Lunde et al., 2012) en bleu, **C** *Tb*TUT4 (PDB ID : 2Q0F) (Stagno et al., 2007) en vert et **E** *Tb*RET2 (PDB ID :2B56) (Deng et al., 2005) en jaune. **B**, **D** et **F**: Les structures sont colorées selon leurs RSMD calculés par PyMol, en bleu pour les zones les plus proches en allant vers les zones les plus éloignées en rouge, les zones non utilisées pour l'alignement sont en gris.

Le score global (QMEANscore4) de la structure modélisée du domaine nucléotidyltransférase de URT1 est de 0,63. Cette valeur est proche de celui de la structure 4FH5 (0,81). Après comparaison avec les valeurs expérimentales issues de la PDB, ce modèle se situe dans la moitié inférieure de la courbe, équivalent à un Z-score inférieur à -2 (-2,29) (Figure 2B.2A à C). Ce score indique une fiabilité suffisante du modèle généré. En effet, on estime que la limite de fiabilité se situe à une valeur inférieure à -3. De plus, les erreurs prédites par cette méthode sont localisées au niveau des boucles externes (Figure 2B.2D et E) (Kiefer et al., 2009; Schwede et al., 2003).

c) Transfert des ligands au modèle généré

Afin de transférer les différents ligands dans le site actif de la structure du domaine nucléotidyltransférase modélisée (UTP, ARN et magnésium), il faut que les protéines et leurs structures respectent différents critères. Il faut que les acides aminés importants pour l'interaction avec les ligands soient conservés, que la distance entre les structures respectives soit suffisamment proche (RSMD<2 Å) (*Root Mean Squared Deviation*) et qu'il n'y ait pas de conflit entre la structure modélisée et le ligand transféré (clash stérique). Le transfert de ces ligands est d'autant plus fiable si des données expérimentales confirment leur interaction dans le site actif de la protéine modélisée.

Validation du transfert de ligands UTP, magnésium et U5P

Afin de valider le transfert des ligands (UTP, U5P et magnésium) j'ai tout d'abord vérifié que la distance entre les structures est suffisamment proche (RMSD<2 Å) (Figure 2B.3) et que les acides aminés sont conservés entre URT1 et les uridylyltransférases résolues en présence des ligands d'intérêt (Cid1, *Tb*TUT4 et *Tb*RET2) (Figure 2B.4). Pour évaluer les distances entre les protéines, j'ai procédé à un alignement des structures 3D deux à deux avec URT1 puis j'ai utilisé le plugin RMSD de PyMOL pour évaluer la distance entre les carbones α. Ce plugin permet aussi de colorer les structures selon la distance entre leurs carbones α dans PyMOL (Figure 2B.3B, D et F).

Les ribonucléotidyltransférases ont une structure 3D très similaire au niveau du site actif. Les acides aminés permettant la coordination de l'ion magnésium sont conservés DxD (D101 et D103 pour Cid1, D491 et D493 pour URT1, D66 et D68 pour *Tb*TUT4 et D106 et D108 pour *Tb*RET2) (Figure 2B.4A). URT1 et Cid1 ont une distance minimale entre leurs carbones α de 0,01 Å au niveau du site actif. Les zones ayant un RMSD plus important (au maximum 6 Å), correspondent aux boucles. Le RMSD moyen entre URT1 et Cid1 reste faible (0,27 Å). D'après ces

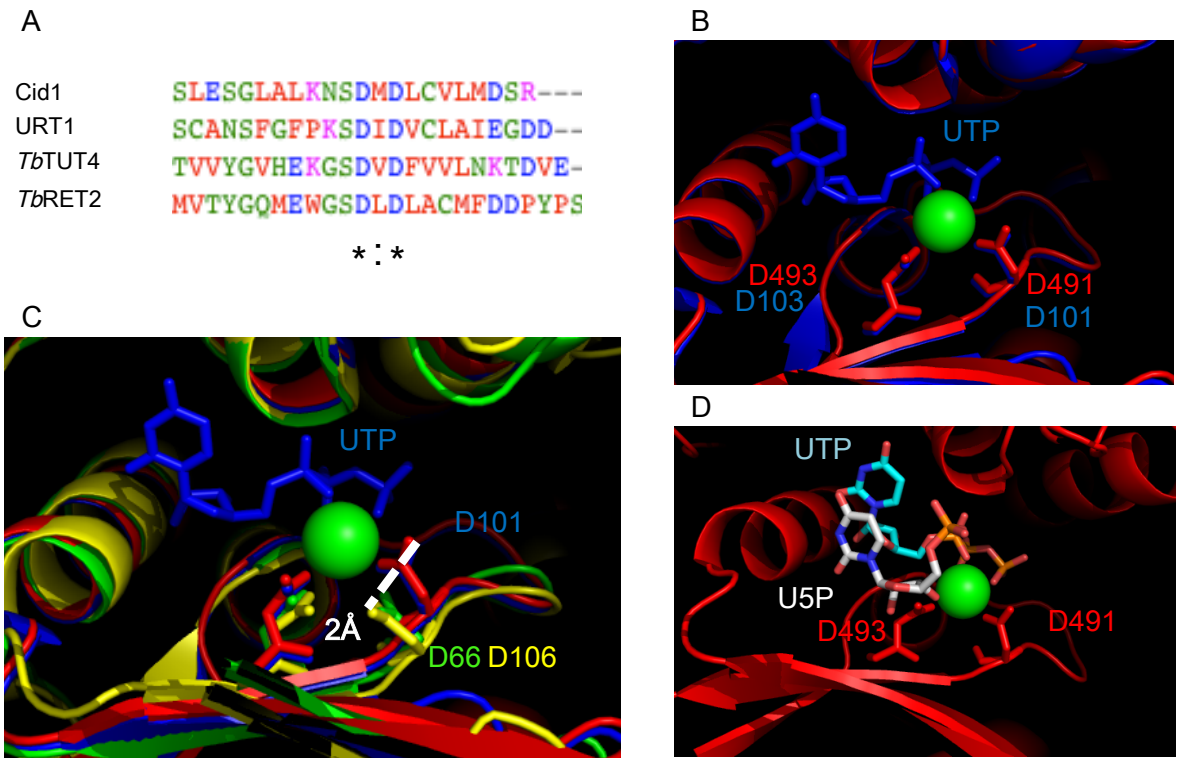


Figure 2B.4 : Transposition des ligands UTP, magnésium et U5P dans le site actif de URT1 modélisé

A Les aspartates DxD coordonnant le magnésium sont conservées chez les uridylyltransférases. **B** et **C**, **B** Représentation des deux aspartates impliquées dans la coordination de l'ion magnésium en présence d'UTP chez URT1 et Cid1 (PDB ID : 4FH5) (Lunde et al., 2012) après alignement des structures, **C** avec *Tb*TUT4 (PDB ID : 2Q0F) (Stagno et al., 2007) et *Tb*RET2 (PDB ID : 2B56) (Deng et al., 2005) ont été ajoutées. Comme pour la figure 3.3, URT1 est colorée en rouge, Cid1 en bleu, l'ion magnésium est représenté par une sphère verte (PDB ID : 4FH5), *Tb*TUT4 est coloré en vert, *Tb*RET2 en jaune, l'UTP de Cid1 (PDB ID : 4FH5) en bleu. **D** Le site actif de URT1 est représenté avec les différents ligands. Les carbones de l'UTP sont en bleu clair, ceux de l'U5P, mimant le substrat ARN, en blanc.

données, le magnésium et l'UTP de la structure 4FH5 sont transposables dans la modélisation de la structure 3D du domaine nucléotidyltransférase de URT1 (Figure 2B.4B).

Afin de valider le transfert du substrat ARN dans le site actif du domaine nucléotidyltransférase de URT1 modélisé, j'ai aussi évalué les distances minimales entre les carbones α grâce au RMSD en comparant les structures de *Tb*TUT4 ou *Tb*RET2 par rapport au domaine nucléotidyltransférase de URT1 modélisé. Les distances minimales du RMSD de ces deux protéines par rapport à URT1 sont très faibles (respectivement 0,24 Å et 0,36 Å) (Figure 2B.3D et F). Cependant les distances maximales sont beaucoup plus élevées (21,79 Å pour *Tb*TUT4, 45,44 Å pour *Tb*RET2). En effet, les structures 3D diffèrent plus entre elles. La distance moyenne est donc élevée (3,67 Å pour *Tb*TUT4 et 5,20 Å pour *Tb*RET2). J'ai tout de même transféré le substrat ARN de *Tb*TUT4 dans le site actif du domaine nucléotidyltransférase de URT1 car les sites actifs sont très proches : le motif DxD est conservé et la distance des Cα des aspartates est au maximum de 2 Å (D491 et D66 *Tb*TUT4 ou D106 *Tb*RET2) (Figure 2B.4C). La structure 3D modélisée du domaine nucléotidyltransférase de URT1 contient donc tous les ligands nécessaires à son étude (Figure 2B.4D). Ces informations ont été utilisées pour interpréter les résultats des tests d'activité de URT1 présentés dans l'article (Chapitre 2 partie A) et ci-dessous.

2. Description du site actif du domaine nucléotidyltransférase de URT1

Le domaine nucléotidyltransférase de URT1 modélisé est similaire aux ribonucléotidyltransférases. Les ribonucléotidyltransférases ont elles-mêmes une homologie avec le domaine catalytique de la famille des DNA polymérase β. Les protéines de cette famille sont constituées d'un domaine central et d'un domaine catalytique (Figure 2B.5). Le domaine catalytique est composé de cinq brins de feuillets β soutenus par deux hélices α, formant ainsi un sandwich à deux étages αβ. Le feuillet β2, contient le site catalytique (DxD). DxD est indispensable à la coordination de l'ion magnésium (Martin and Keller, 2007) (Chapitre 2 partie A et Figure 2B.6B).

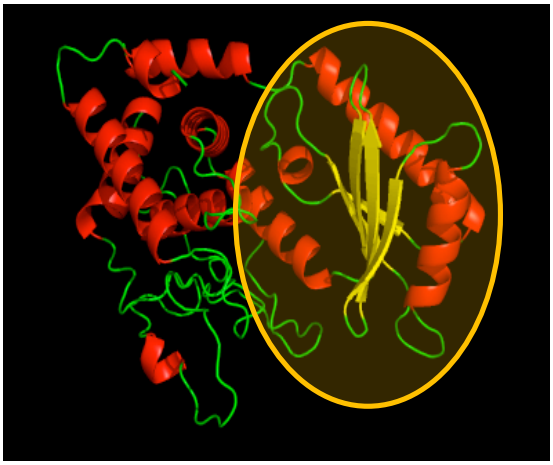
A  B

Figure 2B.5 : Représentation de la structure 3D modélisée du domaine nucléotidyltransférase de URT1

A Modèle de la structure 3D du domaine nucléotidyltransférase de URT1 d'après la structure de Cid1 (PDB ID : 4FH5) (Lunde et al., 2012) coloré selon la structure secondaire, les feuillets β sont représentés en jaunes, les hélices α en rouge et les boucles en vert. Le domaine catalytique est entouré en jaune ; **B** Coloration de la modélisation selon la partie N terminale en rouge en utilisant les couleurs de l'arc en ciel jusqu'à la partie C terminale en bleu.

3. URT1 est une uridylyltransférase intrinsèquement distributive qui est régulée par l'ion divalent présent lors de la réaction

La spécificité nucléotidique de URT1 a été testée *in vitro*, comme présenté dans l'article en partie A (Figure 2B.6-B). Afin de produire URT1, j'ai dû tout d'abord optimiser le protocole d'expression de la protéine recombinante chez *Escherichia coli*. URT1 est stable seulement si elle possède une étiquette en N terminal. En effet, avec une étiquette en C terminal, la protéine est dégradée. Ces observations tendent à confirmer la désorganisation prédite du domaine N terminal de URT1 (Chapitre 1) car il est facilement accessible à des protéases. Même après avoir optimisé la production en testant plusieurs souches bactériennes, la production de URT1 recombinante reste très faible. Cette observation est surement due à sa taille importante (86kD) et sa région N terminale de 46kD intrinsèquement désorganisée d'après les prédictions, rendant URT1 très instable (Chapitre 1).

Afin de tester l'activité catalytique *in vitro* de URT1, la protéine recombinante a dû être purifiée soluble en conditions natives. Pour d'optimiser sa purification et sa solubilité, j'ai induit son expression à 17°C. A cette température, la production de URT1 s'effectue plus lentement qu'à 37°C et permet ainsi une meilleure solubilisation. En effet, cette méthode d'induction améliore le repliement de URT1, empêchant son accumulation dans les corps d'inclusion. Durant le test d'activité catalytique de URT1, j'incube URT1 en présence d'un ARN puis je révèle la taille des ARN après incubation pour observer les nucléotides ajoutés en 3' par URT1. Pour éviter des contaminations par des RNases de *E coli*, j'ai dû effectuer deux purifications successives de URT1. URT1 est exprimé en présence de deux étiquettes en N terminal : 6 histidines et la glutathione S-transférase (GST) et fait une taille totale d'environ 120kDa. Après expression et purification (Matériels et Méthodes), 6his-GST-URT1 est pure mais sa concentration finale reste faible (environ 50ng/µL) (Figure 2B.6A). Celle-ci n'est suffisante pour cristalliser URT1. Cependant, la quantité de URT1 est suffisante pour effectuer les tests d'activité catalytique car la quantité nécessaire lors des tests est faible (100ng par réaction).

a) URT1 est une uridylyltransférase très spécifique des UTP

La modélisation du site actif met en évidence une première sélection des nucléotides par le type de base azotée présente dans le site actif. Seules les bases pyrimidiques qui sont structuralement plus petites que les bases puriques, peuvent se placer de façon optimale. Les phosphates de l'UTP sont impliqués dans la

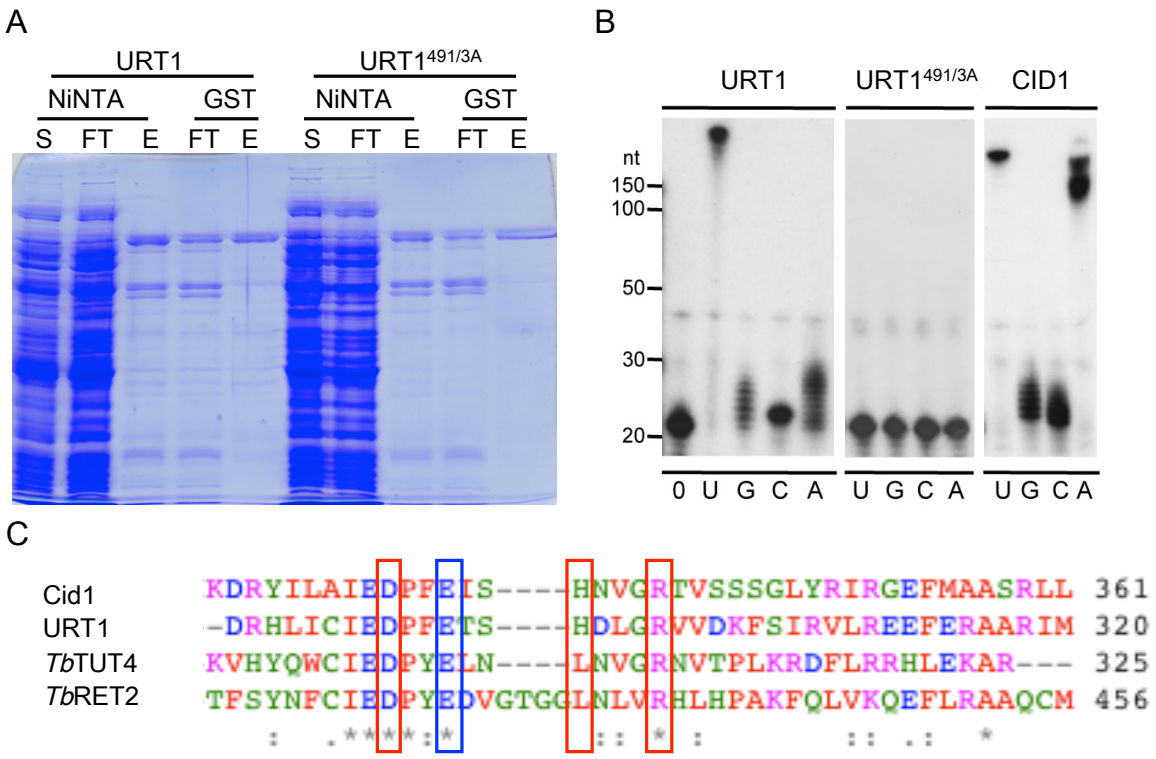


Figure 2B.6 : Les acides aminés importants pour la spécificité de l'UTP chez les autres uridylyltransférases sont conservés chez URT1

A Double purification successive de URT1 et son équivalent inactif URT1$^{491/2A}$, S: Surnageant, FT: Flowthrough, E: Elution. **B** Comparaison de la spécificité nucléotidique de URT1, URT1$^{D491/2A}$ et de Cid1. Les protéines ont été incubées en présence d'UTP (U), rGTP (G), rCTP (C), rATP, ou sans nucléotides (0). **C** Alignement des uridylyltransférases réalisé par ClustalW, les acides aminés conservés et impliqués dans la spécificité de l'UTP sont encadrés en rouge et sont : D (Cid1:D330, URT1:D708, *Tb*TUT4:D297, *Tb*RET2:D421), R (Cid1:R440, URT1:R718, *Tb*TUT4:R307, *Tb*RET2:R435) et H chez Cid1 (H336) et URT1 (H714). E nécessaire à la liaison à l'ARN est encadré en bleu : Cid1:E333, URT1:E711, *Tb*TUT4:E300, *Tb*RET2:E424.

coordination des ions magnésium permettant ainsi la stabilisation de la base dans le site actif. Le Y478, correspondant au F88 de Cid1 serait en contact avec le ribose de l'uridine. Le triphosphate serait sélectionnée par S490, S598 et K580, K584 correspondant au S90, S211 et K193, K197 de Cid1. Les interactions des phosphates et des acides aminés de URT1 maintiendraient comme pour Cid1, l'interaction entre l'UTP et le T599 (T212 chez Cid1) servant à la sélectivité entre ribose et désoxyribose (Lunde et al., 2012; Munoz-Tello et al., 2012; Yates et al., 2012).

La sélectivité de l'UTP est intrinsèque du site actif de URT1. En effet, URT1 conserve plusieurs acides aminés précédemment impliqués dans la sélectivité pour l'UTP chez d'autres uridylyltransférases, tout d'abord un aspartate (*Tb*RET2 :D421, *Tb*TUT4 : D297, Cid1 : D330 et chez URT1 D708) et une arginine (*Tb*RET2 : R435, *Tb*TUT4 : R307, Cid1 : R339 et chez URT1 : R718) (Figure 2B.6-C). L'azote de l'arginine effectue une liaison ionique avec l'oxygène de l'UTP afin de permettre une différenciation avec l'autre base pyrimidique CTP qui possède, elle, une extrémité basique et ne permet pas la liaison ionique (Stagno et al., 2007).

La résolution de la structure de Cid1 a permis de mettre en évidence un autre acide aminé important pour discriminer l'UTP de l'ATP : l'histidine, H336. En effet comme présenté dans la partie A et Figure 2B.6B, Cid1 possède une forte activité en présence d'ATP. En mutant cette H336, la spécificité de Cid1 pour l'UTP est diminuée et l'incorporation d'ATP est favorisée (Lunde et al., 2012; Munoz-Tello et al., 2012; Yates et al., 2012). Il n'existe pas d'équivalent à cette histidine chez *Tb*TUT4 ni chez *Tb*RET2. Cependant, un équivalent a été récemment mis en évidence chez une autre uridylyltransférase XTUT7 chez *Xenopus leavis* (Lapointe and Wickens, 2013). URT1 conserve un équivalent de cette histidine (H714) (Figure 2B.6C). Cependant, dans les mêmes conditions, URT1 est capable d'ajouter 200 UTP tandis qu'elle peut ajouté seulement 9 ATP. Différemment, Cid1 est capable d'ajouter la même quantité d'UTP et d'ATP. URT1 a donc une capacité beaucoup plus importante que Cid1 à discriminer l'UTP de l'ATP (Figure 2B.6B). Cette discrimination entre l'UTP et l'ATP de URT1 doit impliquer d'autres acides aminés importants que ceux qui ont été précédemment mis en évidence. La résolution atomique du domaine nucléotidyltransférase de URT1 devrait permettre de les identifier. La cristallisation du domaine nucléotidyltransférase est en cours en collaboration avec l'équipe du Dr Meinhart, du Max Planck *Institut for Medical Research* (Heidelberg, Allemagne).

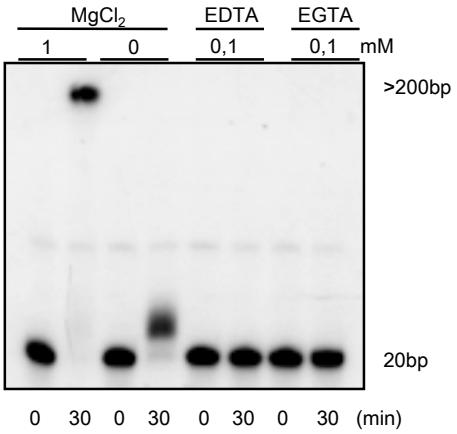


Figure 2B.7 : L'activité catalytique de URT1 est régulée par l'ajout de magnésium

Le magnésium est essentiel pour l'activité de URT1, l'EDTA et l'EGTA chélatent les ions divalents. La réaction est effectuée avec ou sans ajout de magnésium (1mM). Les produits de réaction ont été résolus sur un gel dénaturant de polyacrylamide à 17% après 30 minutes de réaction à température ambiante.

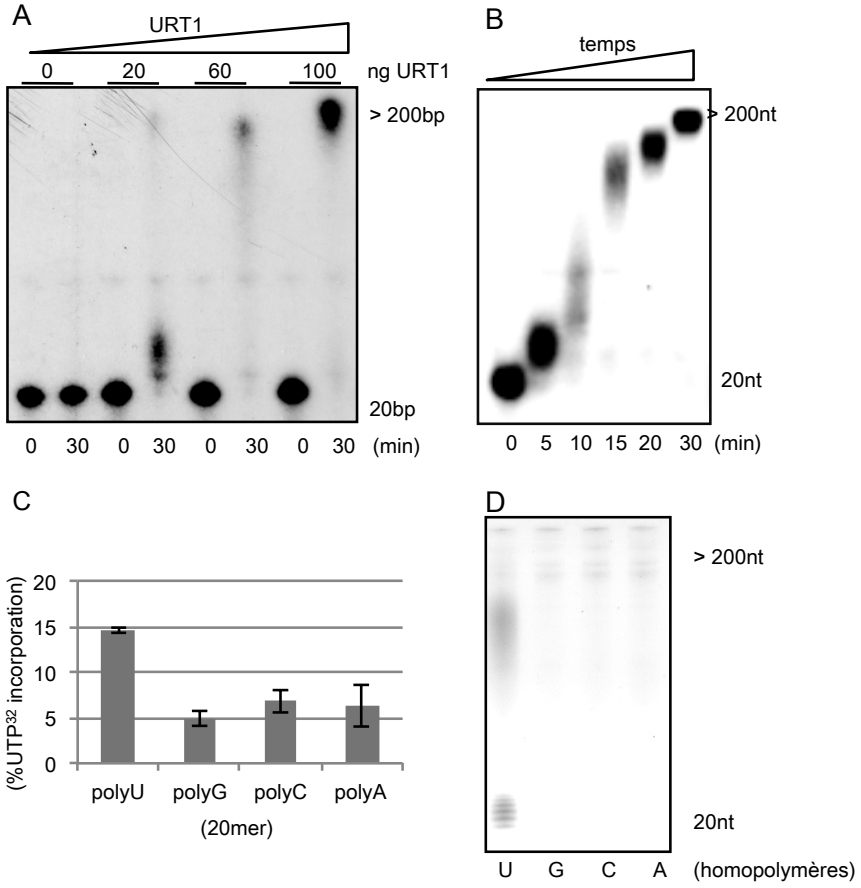


Figure 2B.8: URT1 est intrinsèquement distributive au niveau des premières uridines ajoutées

A L'activité de URT1 dépend du rapport URT1/ARN. **B** Courbe d'activité de URT1 en présence d'uridines au cours du temps. **C** Pourcentage d'incorporation de [α-^{32}P]-UTP par URT1 en présence de l'un des homopolymères de vingt U, rG, rC ou rA après élimination des nucléotides non incorporés par précipitation au TCA. **D** La résolution de la réaction C sur gel de polyacrylamide.

Pour A et B, le temps d'incubation est indiqué en bas des gels de polyacrylamide de 17%. Pour C et D, le temps d'incubation est de 30 minutes.

b) Le magnésium est essentiel à l'activité de URT1

L'utilisation d'ion magnésium est essentielle pour l'activité de URT1. En effet, l'activité de URT1 est inhibée si la coordination du magnésium est prévenue par la mutation des DxD du site actif en AxA ou que les ions divalents présents dans la réaction sont chélatés lors de la réaction enzymatique par la présence d'EDTA ou d'EGTA, (Figure 2B.6B et Figure 2B.7A). De plus, sans ajout externe de magnésium lors de la réaction, URT1 est faiblement active. URT1 doit être liée à un ion divalent pendant la synthèse chez *E. coli*, ion qui co-purifie avec la protéine. Quand la réaction s'effectue avec un ajout de magnésium (1mM), des centaines d'UTP sont ajoutées à l'ARN (Figure 2B.7). L'activité de URT1 est donc régulée par la quantité de magnésium présente lors de la réaction.

4. URT1 est intrinsèquement distributive

Afin de tester la spécificité nucléotidique, la réaction est effectuée en présence d'une quantité importante de magnésium. Mais l'intensité de la réaction est aussi régulée par la quantité de protéine présente, donc par le ratio protéine/ARN et par le temps d'incubation de la réaction (Figure 2B.8A et B).

Après une incubation de quelques minutes, pratiquement tous les substrats ARN ont un à 4 UTP rajouté(s). URT1 est donc distributive (voir partie A). Après ajouts d'environ 10 UTP, URT1 devient processive (Figure 2B.8B). Cette processivité n'est pas due à la taille du substrat ARN comme décrit dans les données supplémentaires accompagnant notre article sur URT1 mais peut être expliquée par la préférence de URT1 pour les homopolymères d'une vingtaine d'uridines. En effet, quant URT1 est incubée en présence de l'un des homopolymères de vingt ribonucléotides : U, rG, rC ou rA et de [α-^{32}P]-UTP radioactif et que les [α-^{32}P]-UTP non incorporés sont éliminés, la quantité de radioactivité est trois fois plus importante en présence de l'homopolymère d'uridines, indiquant une préférence de URT1 pour les homopolymères d'uridines (Figure 2B.8C). Après résolution de la réaction sur un gel d'acrylamide, une incorporation d'une à trois UTP est mise en évidence donc URT1 est aussi distributive pour les premiers UTP ajoutés avant de devenir processive (Figure 2B.8D). Cette distributivité est donc intrinsèque à la protéine URT1.

A l'inverse de URT1, Cid1 est processive et non distributive et Cid1 est capable de se lier à l'ARN (Yates et al., 2012). La conservation des régions responsables de l'interaction de l'ARN à la surface de Cid1 est comparée avec URT1.

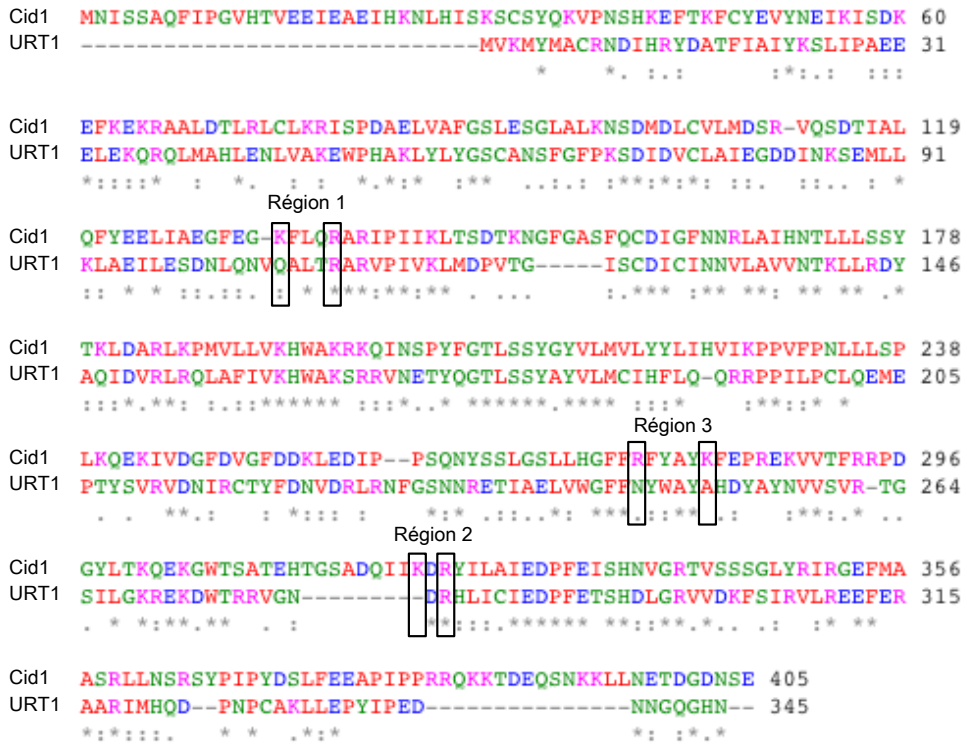


Figure 2B.9 : Les régions de liaison à l'ARN de Cid1 ne sont pas conservées chez URT1

Alignement du domaine nucléotidyltransférase de URT1 et de Cid1 effectué par ClustalW. Les 3 régions impliquées dans la liaison à l'ARN (Yates et al., 2012) sont encadrées en noir.

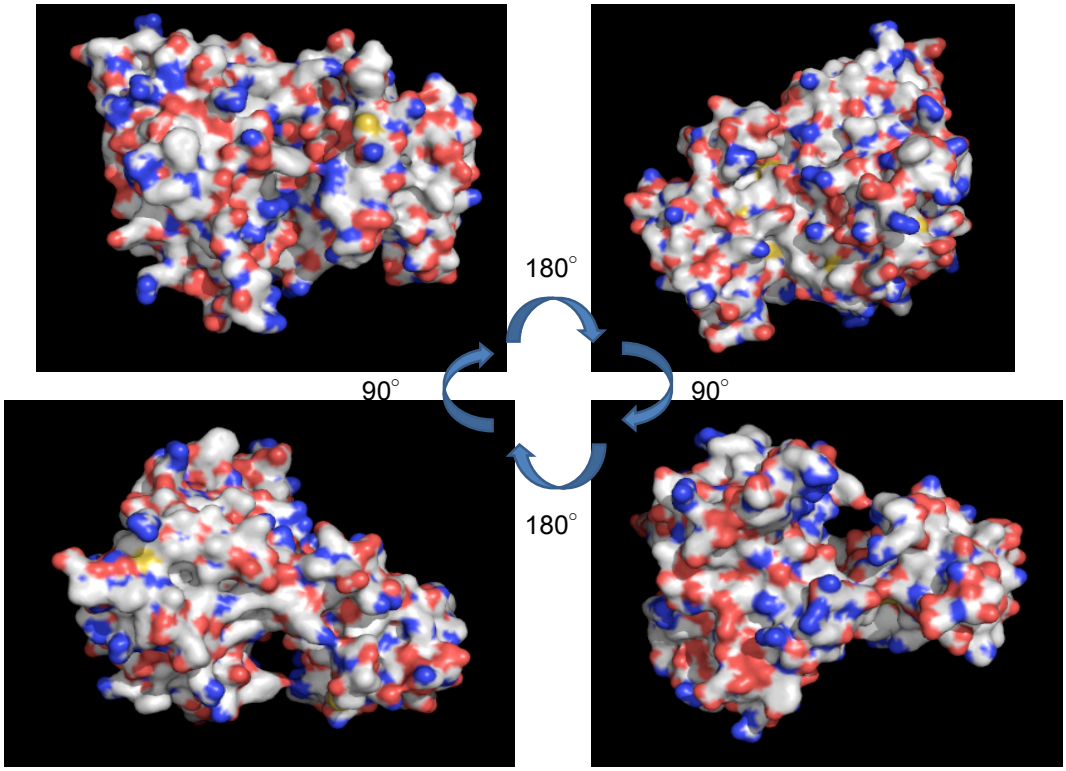


Figure 2B.10 : Représentation de la surface du domaine nucléotidyltransférase de URT1 selon les propriétés chimiques des acides aminées

La surface de URT1 a été prédite en utilisant PyMol et est colorée selon les propriétés chimiques des acides aminés : acides en rouge et basiques en bleu.

5. Les régions de liaisons à l'ARN de Cid1 ne sont pas conservées chez URT1

Cid1 se lie à un ARN d'au moins 13 nucléotides. Cette liaison serait due au couple arginine et lysine de trois régions différentes, la région 1 (R137 et K133), la région 2 (R323 et K321) ainsi que la région 3 (R277 et K282). La liaison à l'ARN dépend d'une part de la région 1 et 2 et de la région 3 (Yates et al., 2012).

Après recherche de la conservation de ces acides aminés chez URT1, j'ai pu mettre en évidence que les lysines des trois régions ne sont pas préservées. Les arginines de la région 1 et 2 sont conservées, et correspondent respectivement à R531 et R703 (Figure 2B.9). D'autre part, il n'existe pas de région chargée positivement sur la surface de la structure modélisée du domaine nucléotidyltransférase de URT1 (Figure 2B.10). Cette divergence entre Cid1 et URT1 pourrait expliquer la distributivité de URT1 dans les conditions utilisées et les résultats négatifs obtenus lorsque j'ai testé l'interaction de URT1 avec de l'ARN, dans les mêmes conditions que celles de l'étude catalytique par la technologie de résonnance plasmonique de surface (SPR) de Biacore™ ni par des tests de co-précipitation. Il semblerait que URT1 ne se lie pas à l'ARN de façon stable.

6. URT1 uridyle *in vitro* tous les ARN ayant une extrémité 3'OH libre

Les expériences précédentes ont montré que URT1 est capable d'uridyler les ARN d'une vingtaine de nucléotides, avec une préférence pour les homopolymères d'uridines. Afin de tester la spécificité de substrats ARN, j'ai incubé URT1 en présence d'ARN total et d'ARN structurés tels que les ARN de transferts (Figure 2B.11A). URT1 est capable d'uridyler *in vitro* tous les ARN testés. Cependant, il faut que l'ARN ait son extrémité 3' libre. En effet, lorsque URT1 est incubée en présence d'un ARN formant une structure tige-boucle incorporant le dernier nucléotide dans cette structure (Figure 2B.11B) (prédiction effectuée par mfold (Zuker, 2003)), URT1 n'est pas capable d'uridyler ce substrat. Si cet ARN est préalablement dénaturé à 70°C durant 5 minutes, URT1 peut faiblement l'uridyler. Cette activité de URT1 reste minimale car l'ARN est restructuré très rapidement et redevient inaccessible à URT1. Cette observation est renforcée par l'utilisation de URT1 d'un autre ARN qui n'incorpore pas le dernier ribonucléotide en 3' dans sa structure tige-boucle (Figure 2B.11B, C et D).

Pour que URT1 puisse uridyler un ARN, il faut que son ribonucléotide en 3' soit libre et que l'extrémité 2'OH ne soit pas modifiée. *In vivo*, il a été montré que les

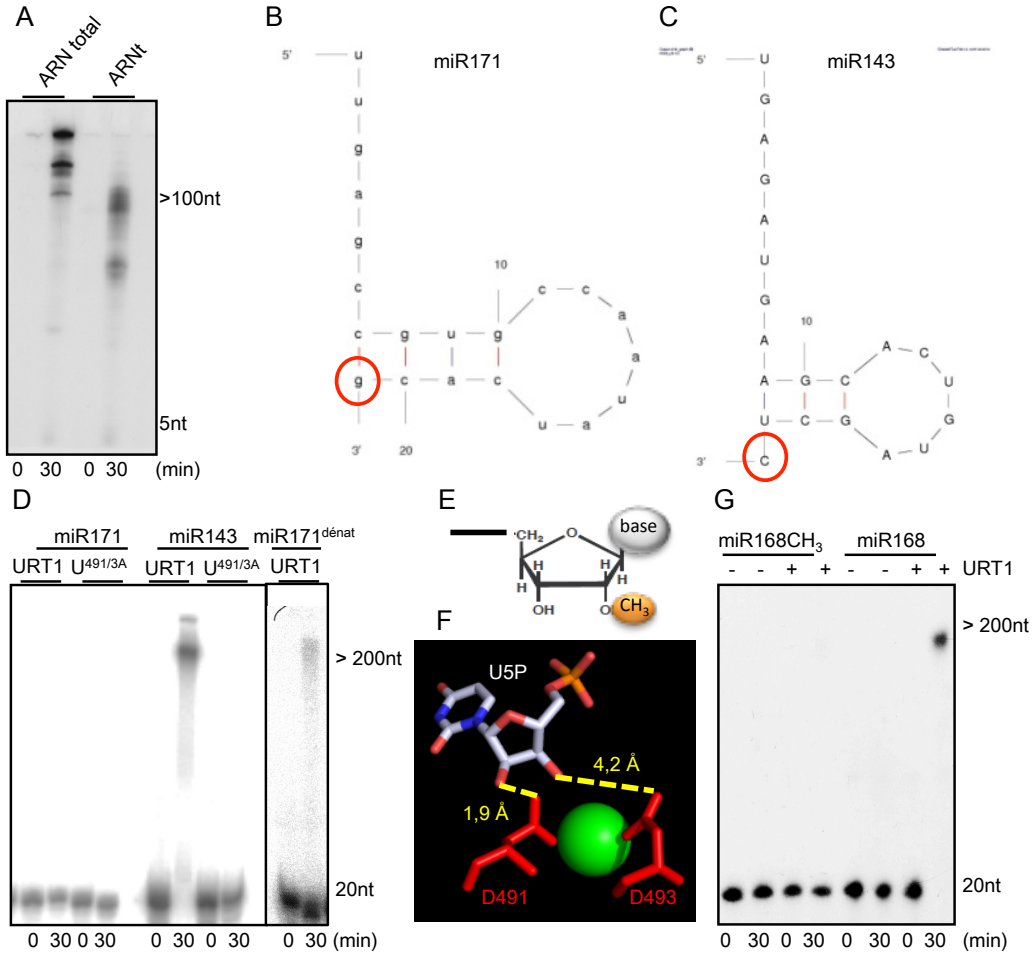


Figure 2B.11: URT1 est capable d'uridyler *in vitro* les ARN ayant une extrémité 3' libre et non-modifiée

A URT1 est capable d'uridyler l'ARN total ou les ARN de transfert (tRNA). **B** et **C** La structure secondaire de deux différents ARN ayant une structure secondaire en épingle à cheveux ayant soit le dernier ribonucléotide inclus (miR171) en **B** ou libre (miR143) en **C** est prédite par mfold à 25°C (http://mfold.rna.albany.edu/) (Zuker, 2003). La dernière base en 3' est entourée en rouge. **D** Activité en présence des ARN structurés en B et C, miR171^{dénat} correspond à l'ARN miR171 préalablement dénaturé 5 minutes à 70°C avant ajout de URT1 pour effectuer la réaction. **E** Représentation du substrat ARN méthylé en 2'OH de la dernière base. **F** Les distances entre le 2'OH et l'aspartate D493 de URT1 sont mises en évidence par PyMOL. **G** L'activité de URT1 en présence d'un ARN méthylé en 2'OH (miR168CH$_3$) ou non (miR168).

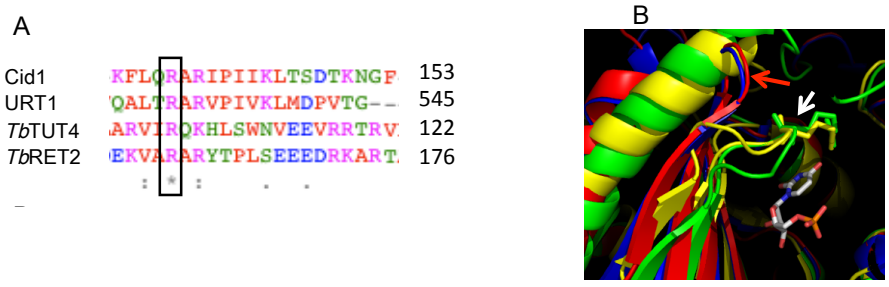


Figure 2B.12 : URT1 ne possède pas d'équivalent à l'arginine nécessaire à la spécificité du substrat ARN de *Tb*TUT4 et *Tb*RET2

A Alignement des uridylyltransférases est effectué par ClustalW. **B** Représentation de la boucle de la structure 3D des uridylyltransférases contenant l'arginine nécessaire à la spécificité du substrat ARN, la flèche rouge montre la boucle avec Cid1 (bleu) et URT1 (rouge) et en blanc la boucle avec R121 et R144 de *Tb*TUT4 (vert) et *Tb*RET2 (jaune) responsable de la spécificité du substrat ARN.

microARN sont méthylés en 2'OH par HEN1 (Li et al., 2005) (Figure 2B.11E). Après transposition de l'uridine-5-phosphate, mimant le substrat ARN, dans le site actif de URT1, les oxygènes du substrat ARN semblent importants pour les liaisons dans le site actif. En effet, la cristallisation de *Tb*TUT4 montre bien l'importance du 2'OH du U5P dans la stabilisation de l'ARN (Stagno et al., 2007). Celui-ci créé une liaison hydrogène avec les aspartates catalytiques conservés chez URT1 : D491 (1,5 Å) et D493 (4,2 Å). Cette liaison est importante pour positionner le 3'OH à bonne distance (1,9 Å) de D491 et orienter ainsi correctement le phosphate catalytique de la molécule d'UTP. Le 3'OH est lui aussi impliqué dans la coordination du magnésium (5,8 Å), tandis que le 2'OH est impliqué dans la coordination avec l'aspartate D493 (Figure 2B.11F). Une modification de ce 2'OH par une méthylation empêche stériquement le dernier ribonucléotide en 3' de l'ARN de se placer correctement dans le site actif. La réaction est donc inhibée en présence d'un ARN méthylé en 2'OH (Figure 2B.11G).

La spécificité ARN de *Tb*TUT4 s'effectue par une liaison avec l'arginine R121 en formant une liaison ionique avec un résidu glutamate : E300 et E424 chez *Tb*TUT4 et *Tb*RET2 (Stagno et al., 2007). Un équivalent est conservé chez Cid1 E333, et sa mutation réduit l'activité de la protéine (Lunde et al., 2012; Munoz-Tello et al., 2012; Yates et al., 2012). Ce résidu est aussi conservé chez URT1 : E711. Un autre acide aminé a été découvert chez Cid1 interagissant avec de façon aspécifique avec l'ARN N165 (communication personnelle du Dr Thore), URT1 possède un équivalent (N554).

L'arginine responsable de la sélectivité du dernier nucléotide du substrat ARN chez *Tb*TUT4 et *Tb*RET2 (réciproquement R121 et R144) possède un équivalent chez Cid1 (R139) et URT1 (R531) (Figure 2B.12A). Cependant la boucle les contenant est beaucoup plus courte et les arginines de Cid1 et URT1 ne sont pas situées aussi près du substrat ARN que pour *Tb*TUT4 et *Tb*RET2 (Figure 2B.12B). La distance plus importante des arginines de Cid1 et URT1 permettrait d'expliquer leur absence de spécificité intrinsèque de substrat ARN *in vitro*. La spécificité de URT1 *in vivo* est très forte pour les ARNm oligoadénylés, au contraire des résultats observés *in vitro*. L'hypothèse la plus plausible est que la spécificité de substrat ARN *in vivo* est apportée par un partenaire protéique (Discussion partie B).

7. URT1 est active sous forme de monomère

Certaines protéines de la famille des ribonucléotidyltransférases ont besoin de former des dimères afin d'être actives (Bai et al., 2011). J'ai donc testé si la protéine

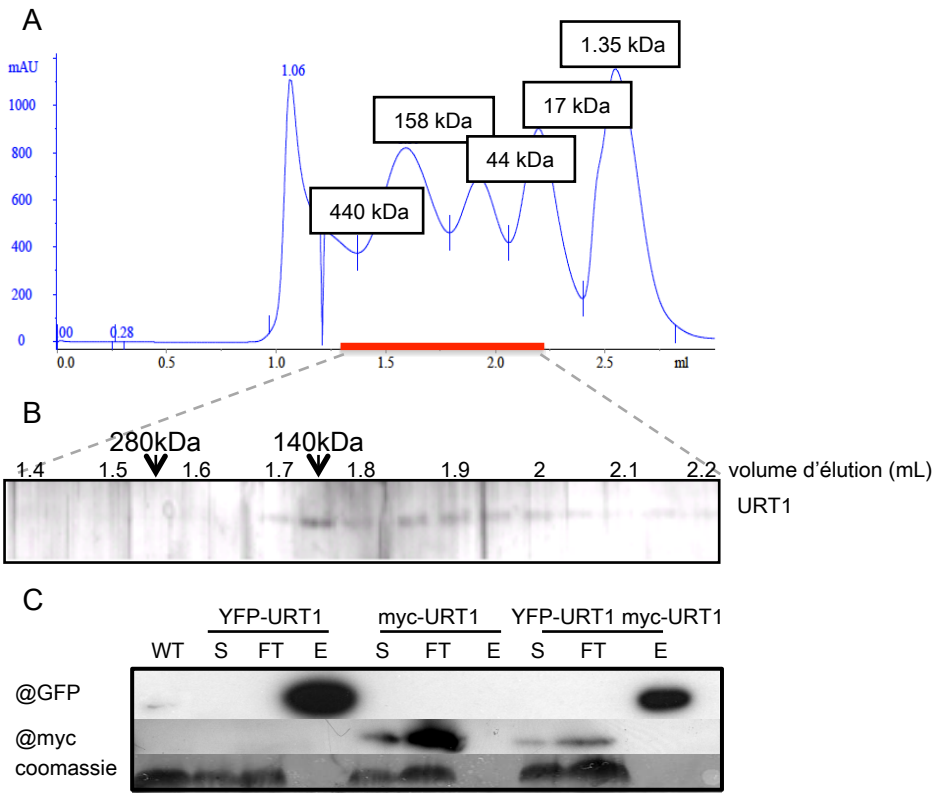


Figure 2B.13: URT1 est active sous forme de monomère

A Chromatographie d'exclusion de taille effectuée avec différents marqueurs de tailles. **B** Coloration à l'argent des élutions de la chromatographie d'exclusion de taille effectuée avec URT1. **C** Représentation des westerns blot effectués à partir de la co-immunoprécipitation de URT1 après expression transitoire des protéines myc-URT1 et YFP-URT1 dans des feuilles de *N. benthamiana*. S: Surnageant, FT: Flowthrough et E: Elution.

recombinante URT1 forme des dimères dans les conditions d'études de l'activité catalytique effectuées *in vitro* par différentes méthodes.

Tout d'abord, j'ai effectué une chromatographie d'exclusion de taille avec la protéine recombinante 6his-GST-URT1 dans les conditions que j'ai utilisées pour tester son activité catalytique. La protéine est éluée à un volume d'élution correspondant à une taille d'environ 140kDa, taille proche de la taille de URT1 et de ses étiquettes 6His-GST (environ 120kDa), indiquant qu'il n'y a pas de dimérisation. Cependant, URT1 est aussi éluée avec des tailles apparentes inférieures. Il semble que URT1 interagit avec la phase utilisée lors de la chromatographie (Figure 2B.13A). Afin de vérifier l'absence de dimérisation de URT1, j'ai utilisé deux autres méthodes en utilisant URT1 recombinante avec deux étiquettes différentes (6His-GST-URT1 et 6His-MBP-URT1). J'ai utilisé la méthode de résonance plasmonique de surface (SPR) de Biacore™ et la co-précipitation *in vitro*. Ces deux techniques n'ont pas permis de mettre en évidence une homodimérisation de URT1. La protéine recombinante URT1 est donc vraisemblablement active sous forme de monomère dans les conditions d'activité *in vitro*.

Afin de vérifier cette hypothèse *in vivo*, j'ai effectué un test d'immunoprécipitation après co-expression transitoire dans les feuilles de *N. benthamiana* de URT1 avec deux étiquettes différentes (myc et HA) situées en N terminal (Matériels et Méthodes). Ce test n'a pas non plus mis en évidence une dimérisation de URT1. URT1 semble donc être aussi active sous forme de monomère *in vivo* (Figure 2B.13B).

8. Conclusion/Ouverture

In vitro, URT1 est une uridylyltransférase intrinsèquement spécifique des UTP. Son étude catalytique montre une activité intrinsèquement distributive au niveau des premiers nucléotides ajoutés. Ce qui explique que *in vivo*, URT1 ajoute une à trois uridines aux ARNm après déadénylation. Leur uridylation par URT1 protège leur extrémité 3', prévenant ainsi la formation d'ARNm tronqués. Ces résultats indiquent que l'uridylation contribue à établir la polarité de dégradation de 5' en 3' des ARNm chez *A. thaliana*. Ces résultats établissent un nouveau rôle de protection de l'uridylation des ARNm. La polarité de la dégradation dans le sens 5'-3' est importante en cas de dégradation co-traductionnelle. En effet, l'effet protecteur de l'uridylation est observé au niveau des polysomes.

J'ai effectué des expériences complémentaires de l'activité catalytique de URT1 *in vitro* présentées dans la partie A et une modélisation du domaine

nucléotidyltransférase de URT1 afin de mieux comprendre son fonctionnement. Contrairement à son homologue Cid1, URT1 est très spécifique des UTP et peut discriminer intrinsèquement entre l'ATP et l'UTP. Pourtant les acides aminés connus pour être responsables de la spécificité pour l'UTP chez Cid1 sont conservés chez URT1. Une résolution atomique du domaine nucléotidyltransférase de URT1 devrait mettre à jour de nouveaux acides aminés qui permettent à URT1 de discriminer l'ATP de l'UTP (collaboration avec le laboratoire du Dr Meinhart) et expliquer sa capacité accrue à différencier l'UTP de l'ATP. La spécificité de URT1 *in vivo* pour les ARNm déadénylés n'est pas observée *in vitro* et pourrait être apportée par un partenaire protéique (Discussion partie B). En effet, *in* vitro, URT1 peut uridyler tous les substrats ARN testés ayant leur extrémité 3' de l'ARN libre tels que les ARN ribosomiques et les ARN de transfert. Il faut cependant la dernière base libre des ARN ne soit pas modifiée en 2'OH. La protéine principalement responsable de l'uridylation des microARN qui peuvent être uridylés lorsqu'ils ne sont pas modifiés en 2'OH, chez *A. thaliana*, a été récemment mise en évidence : HESO1 (Ren et al., 2012; Zhao et al., 2012). Cependant, l'uridylation des microARN chez le mutant nul *heso1-1* ne disparaît pas. URT1 pourrait avoir un rôle complémentaire de HESO1. Cette hypothèse est actuellement testée dans le laboratoire (Discussion partie B).

D'autre part, URT1 est localisée au niveau des *processing bodies* et des granules de stress (Chapitre1 partie A). Ces foci cytoplasmiques sont impliquées dans la régulation de la traduction et/ou de la dégradation des ARNm en réponse aux stress (Introduction partie C). Les mécanismes d'adressage de URT1 au niveau de ces foci sont étudiés plus précisément dans le chapitre 3.

Chapitre 3 :
L'adressage de URT1 au niveau des *processing bodies* dépend de la région N terminale de URT1

Résumé

Contexte scientifique

URT1 est une nucléotidyltransférase responsable de l'uridylation des ARNm déadénylés, protégeant ainsi leur extrémité 3' et polarisant la dégradation dans le sens 5'-3' (Chapitre 2). URT1 peut être divisée en deux régions de taille similaire : la région N terminale, prédite intrinsèquement désorganisée et la région C terminale, structurée qui correspond au domaine nucléotidyltransférase (Chapitre 1). La conservation de la région N terminale de URT1, de ces motifs et de ces régions enrichies en Q, en P/Q/N/H et en E chez les plantes à fleurs suggèrent une préservation de leur fonction biologique lors de l'évolution, soit dans la localisation de URT1, soit dans la reconnaissance de partenaires protéiques ou de substrats ARN. Sachant que URT1 est localisée au niveau des *processing bodies* et des granules de stress (Chapitre 2) et que les régions riches en Q et en Q/N se sont révélées responsables de la localisation d'autres protéines au niveau des *processing bodies* et/ou des granules de stress (Introduction partie C) (Reijns et al., 2008; Thomas et al., 2011; Weber et al., 2008), nous avons étudié l'adressage de URT1 au niveau de ces foci cytoplasmiques.

Questions biologiques

Les motifs conservés et les régions enrichies en Q et en P/Q/N/H chez les Angiospermes sont-ils impliqués dans la localisation de URT1 ?

Stratégie d'étude

Afin d'étudier l'impact des motifs et régions conservées, sur la localisation de URT1, j'ai utilisé des versions de URT1 tronquées, sans les régions et les motifs conservés, précédemment identifiés (Chapitre 1). Ces protéines fusionnées à une étiquette fluorescente sont exprimées de façon transitoire dans des feuilles de tabac *N. benthamiana,* faute de ne pouvoir obtenir de plantes transformées stables dû à une mise en place de *silencing* avec l'étiquette YFP (Discussion partie C), et j'ai observé leur localisation au microscope confocal.

Afin d'étudier le mécanisme d'adressage de URT1, j'ai quantifié la co-localisation de URT1 et des différentes versions de URT1 tronquées avec des marqueurs des *processing bodies* et des granules de stress grâce :

- au plugin JACoP d'imageJ

- à une méthode automatique FOCI, que j'ai mise en place avec la collaboration de la plateforme de microscopie de l'IBMP.

Principaux résultats et ouverture

Le plugin FOCI permet de sélectionner les spots pour l'étude de la co-localisation, éliminant certains artéfacts et permettant ainsi une meilleure quantification de la co-localisation de deux protéines.

Une partie de la région N terminale est essentielle pour l'adressage de URT1 au niveau des *processing bodies*. Par contre, la région N terminale de URT1 ne semblerait pas avoir d'impact sur la localisation de URT1 au niveau des *processing bodies* après un stress thermique. Des résultats préliminaires indiquent que celle-ci n'aurait pas d'impact non plus sur la localisation de URT1 au niveau des granules de stress en cas de stress thermique.

Le rôle de URT1 au niveau de ces foci cytoplasmiques n'est pas encore résolu mais les ARNm uridylés par URT1 semblent être des substrats de la dégradation 5'-3' (Chapitre 2 partie A), pouvant expliquer sa localisation au niveau des *processing bodies.* Cependant, URT1 pourrait aussi inhiber la traduction des ARNm déadénylés comme l'ajout de pyrimidines en 3' comme suggéré chez *A. nidulans* ou chez *Xenopus laevis* (Lapointe and Wickens, 2013; Morozov et al., 2012).

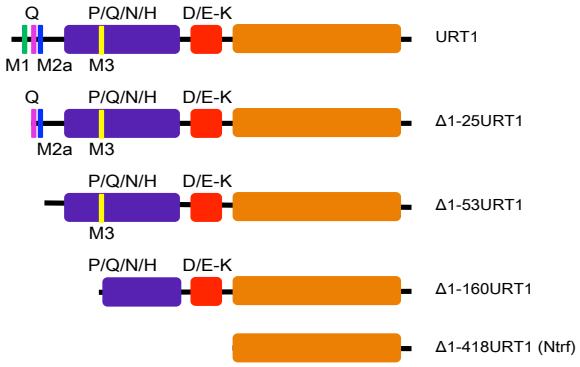


Figure 3.1 : Les délétions de URT1 utilisées dans les expériences

Les motifs et les régions conservées chez les homologues de URT1 des dicotylédones sont représentés sur la séquence de URT1 *A. thaliana*. Le motif M1 leucine est représenté en vert, M2a en bleu et le motif M3 en jaune. La région riche en Q est représentée en rose, celle riche en P/Q/N/H en violet et celle riche en E en rouge (d'après Composition Profiler), appelée D/E-K. Le domaine nucléotidyltransférase est représenté en orange.

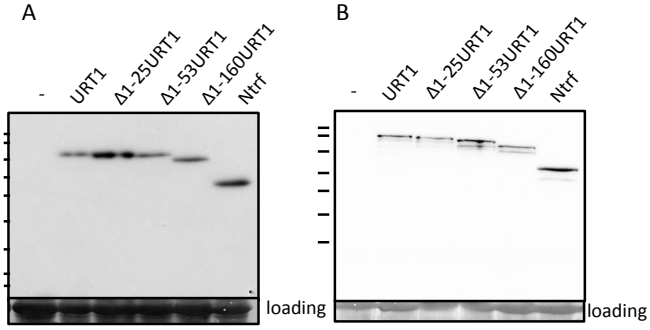


Figure 3.2 : L'expression de URT1 et de ses délétions avec une étiquette YFP fusionnée en N et en C terminale est similaire

A L'expression de URT1 et de ses délétions avec une étiquette YFP en N terminale.
B L'expression de URT1 et de ses délétions avec une étiquette YFP en C terminale.
Les westerns blot sont réalisés sur un disque foliaire de 1cm² de feuille de *N. benthamiana* 2 à 3 jours après agro-infiltration. Les anticorps anti-GFP (1/5000e) couplés à la HRP sont utilisés durant une heure d'incubation.

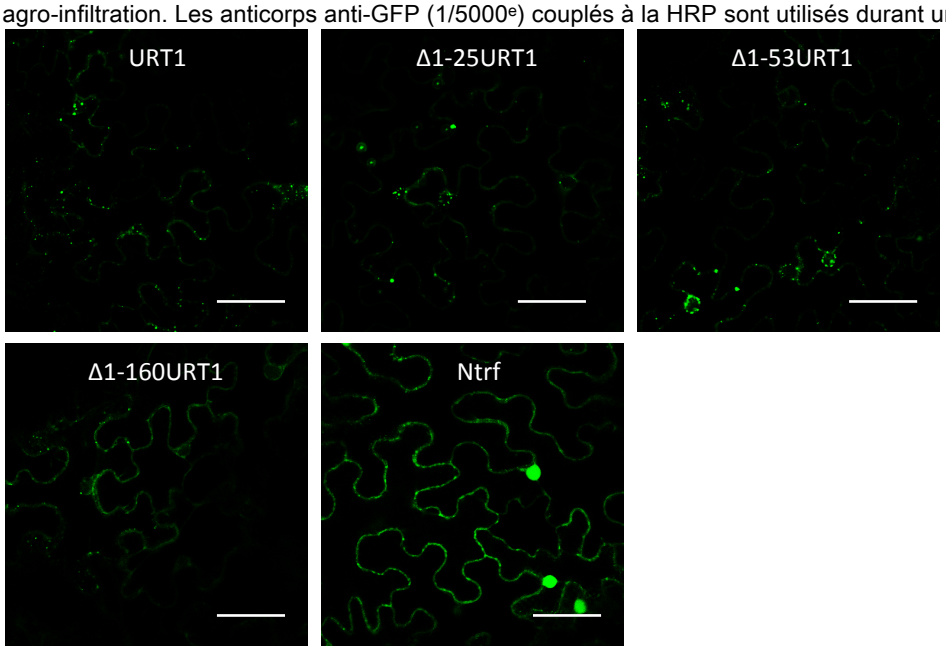


Figure 3.3: La localisation de URT1 et de ses délétions

Les protéines sont exprimées de façon transitoire dans des feuilles de *N. benthamiana* et observées au microscope confocal lsm780 (Zeiss) 2 à 3 jours après agro-infiltration. La barre blanche représente 50 µm.

1. La région N terminale de URT1 impacte la localisation de URT1

La conservation des motifs et des régions au niveau de la région N terminale de URT1 chez les Angiospermes, mise en évidence dans le chapitre 1, suggère qu'ils sont importants pour la fonction biologique de URT1. Sachant que URT1 est localisée au niveau des *processing bodies* et des granules de stress (Chapitre 2 partie A) (Sement et al., 2013), j'ai testé l'impact des motifs et régions conservées de URT1 dans sa localisation. Dans ce but, j'ai exprimé de façon transitoire les différentes constructions de URT1 avec une étiquette fluorescente YFP fusionnée en N ou en C terminal dans des feuilles de *N. benthamina*. La localisation de URT1 est comparée à celles des délétions n'ayant plus la région riche en G/P et le motif M1 (Δ 1-53URT1), manquant également le motif M2 (Δ1-53URT1), manquant en plus une partie de la région P/Q/N/H et le motif M3 (Δ1-160URT1) et ne contenant que le domaine nucléotidyltransférase seul (Ntrf), (Figure 3.1). L'expression des protéines est mise en évidence par un western blot et a confirmé la présence d'une bande à la taille attendue de chacune des délétions, montrant une expression similaire des différentes protéines (Figure 3.2).

Les localisations des délétions de URT1 fusionnées à la YFP aux extrémités N ou C terminales sont similaires. URT1 est localisée dans le cytoplasme avec des foci cytoplasmiques semblable à la localisation précédemment mise en évidence (Chapitre 2 partie A) (Figure 3.3). Différemment, Ntrf apparait plus diffuse dans le cytoplasme et est aussi détectée dans le noyau, même si le domaine nucléotidyltransférase ne possède pas de prédiction de séquence de localisation nucléaire. Cette localisation nucléaire est très probablement due à la diffusion du domaine nucléotidyltransférase car sa taille (au total de 67kDa) est proche de la limite de taille d'un pore nucléaire. Je ne considère pas cette localisation nucléaire comme étant biologiquement significative car URT1 n'a jamais été détectée dans le noyau. Les localisations de URT1 et Ntrf ne dépendent pas de l'activité nucléotidyltransférase car les mutants d'activité URT1$^{D491/493A}$ et Ntrf $^{D491/493A}$ présentent la même localisation que URT1 et Ntrf, cytoplasmique avec des foci cytoplasmiques pour URT1 et nucléo-cytoplasmique plus diffus dans le cytoplasme pour Ntrf (Figure 3.3 et 3.4).

Les motifs conservés M1, M2 et M3 n'ont pas d'impact sur la localisation cytoplasmique de URT1 mais ils pourraient avoir un impact sur la localisation de URT1 au niveau des foci cytoplasmiques (Figure 4.3). En effet, Δ1-160URT1 semble avoir un comportement intermédiaire de URT1 et Ntrf en présentant plus de foci

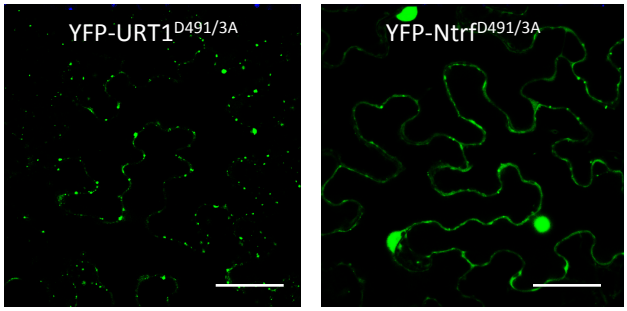


Figure 3.4: La localisation de URT1 et Ntrf ne dépend pas de l'activité nucléotidyltransférase

Les protéines sont exprimées de façon transitoire dans des feuilles de *N. benthamiana* et observées au microscope confocal lsm780 (Zeiss) 2 à 3 jours après agro-infiltration. Les protéines avec une étiquette YFP sont pseudo-colorées en vert. La barre blanche représente 50 µm.

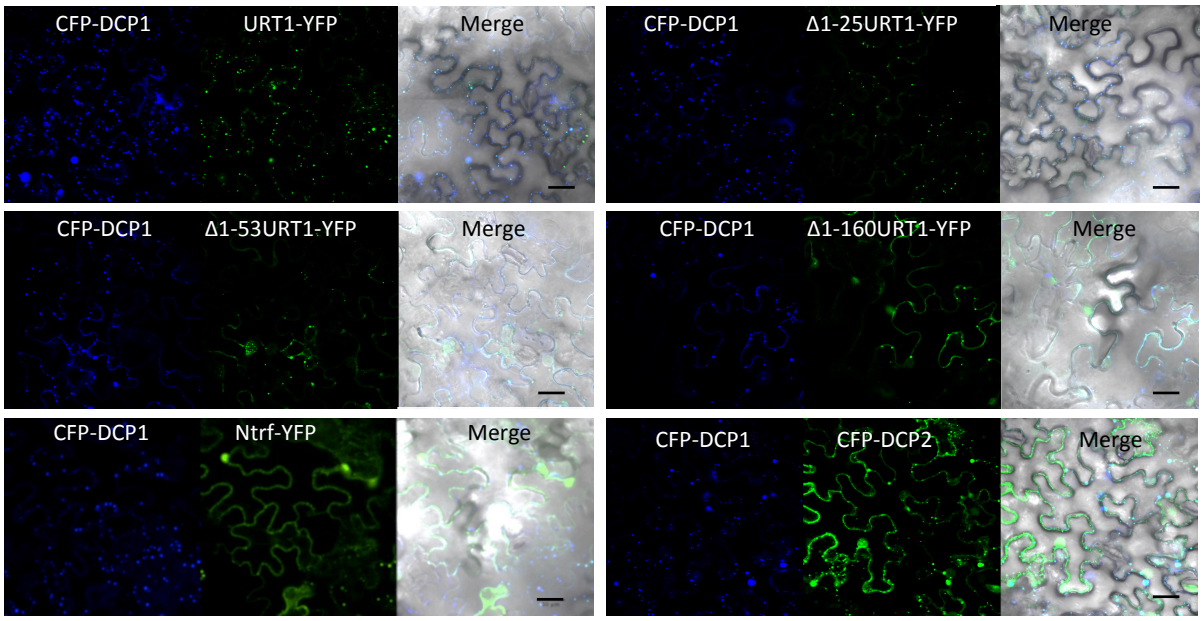


Figure 3.5: La co-expression de DCP1 ne change pas la localisation des différentes délétions de URT1

Image de la co-localisation de URT1 et de ses délétions avec le marqueur des *processing bodies* DCP1

Les protéines sont exprimées de façon transitoire dans des feuilles de *N. benthamiana* et observées au microscope confocal lsm780 (Zeiss) 2 à 3 jours après agro-infiltration. Les protéines avec une étiquette YFP sont pseudo-colorées en vert et celles avec une étiquette CFP en bleu. La barre noire représente 50 µm.

cytoplasmiques que Ntrf mais en étant plus diffuse dans le cytoplasme que URT1. Sachant que URT1 co-localise avec des marqueurs des *processing bodies* et des granules de stress, j'ai étudié l'impact de la région N terminale dans l'adressage de URT1 au niveau de ces foci cytoplasmiques avec les constructions précédemment présentées.

2. Une partie de la région N terminale de URT1 est essentielle à sa localisation au niveau des *processing bodies*

Afin d'étudier l'impact de la région N terminale de URT1 dans son adressage au niveau des *processing bodies*, j'ai co-exprimé les délétions de URT1 avec un marqueur des *processing bodies* (CFP-DCP1) (Weber et al., 2008; Xu and Chua, 2011; Xu et al., 2006). Par la suite j'ai quantifié la co-localisation des délétions de URT1 fusionnée à l'étiquette YFP avec CFP-DCP1, grâce au plugin JACoP (Just Another Colocalization Plugin) du logiciel imageJ (Bolte and Cordelières, 2006). Afin de quantifier la co-localisation plus précisément et de palier les inconvénients du plugin JACoP, j'ai aussi utilisé un second plugin, FOCI, créé en collaboration avec Jérome Mutterer de la plateforme Microscopie et Imagerie de l'IBMP.

La localisation des différentes délétions de URT1 ne changent pas lors de la co-expression avec DCP1 (Figure 3.5). Cependant, comme il a été déjà observé dans de précédentes publications, DCP1 est localisée au niveau des *processing bodies* d'une taille inférieure à 1,88m ou 6 pixels mais DCP1 peut aussi former de plus gros foci qui ne sont pas considérés comme des *processing bodies*. Les images de co-localisation montrent que certains des foci YFP et CFP co-localisent (Figure 3.5). Afin de quantifier cette co-localisation des protéines, les images microscopiques sont analysées automatiquement par le plugin JACoP (Bolte and Cordelières, 2006). Une série d'une dizaine d'images successives est donc prise au hasard une fois que le plan confocal est réglé afin de ne pas biaiser la quantification de la co-localisation. Cette expérience est réalisée trois fois avec URT1 et ses délétions fusionnées avec une étiquette YFP en N terminal ou en C terminal (un total de 454 images). D'autre part, sachant que JACoP utilise une méthode basée sur la corrélation de l'intensité, les lasers sont ajustés et gardés constant lors d'une même expérience. Le coefficient de Pearson R est estimé par image après une sélection d'une zone pertinente d'expression des protéines (Figure 3.6A et B). La valeur de R théorique varie de 1 pour une co-localisation parfaite, -1 pour une exclusion mutuelle et des valeurs

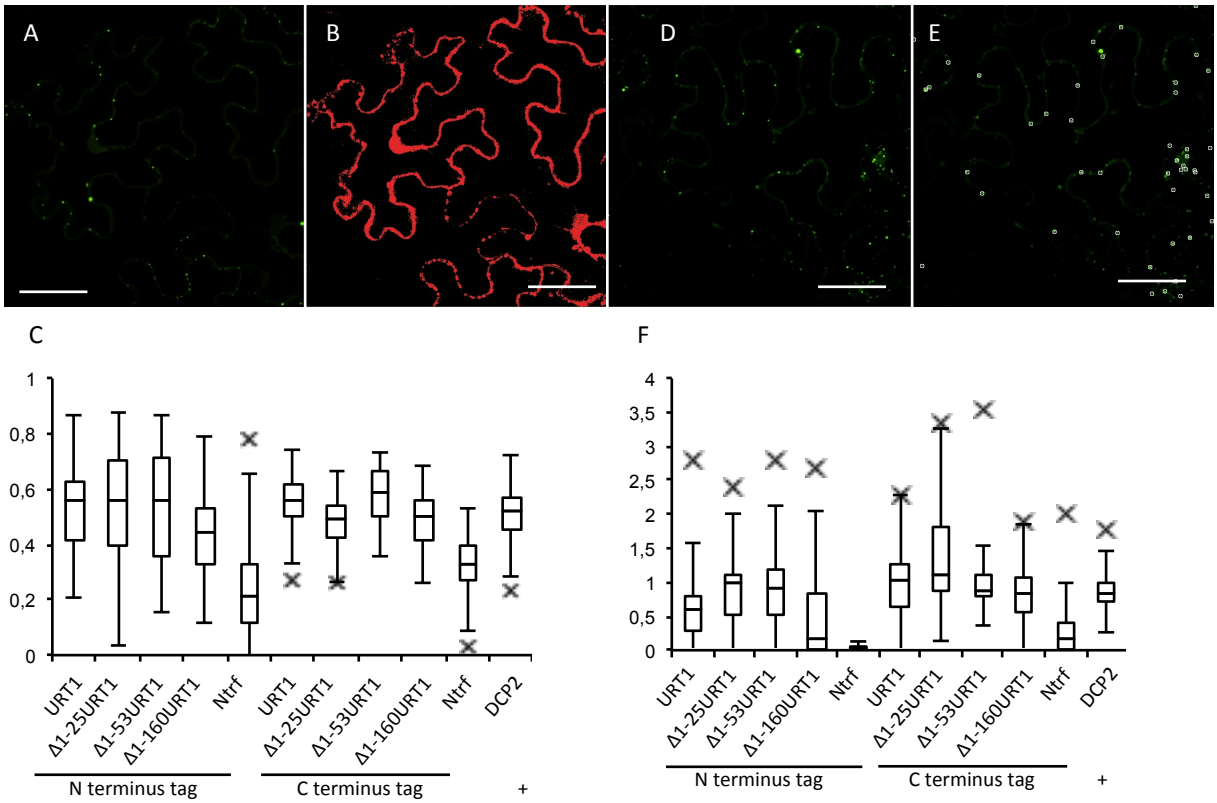


Figure 3.6: Une partie de la région N terminale de URT1 est essentielle à la localisation de URT1 au niveau des *processing bodies*

A Exemple d'image confocal. **B** Sélection de la zone d'étude du plugin JACoP du logiciel ImageJ en rouge à partir de l'image A. **C** Boxplot de la quantification de la co-localisation de URT1 et de ses délétions avec DCP1 par le plugin JACoP du logiciel ImageJ (un total de 454 images). **D** Exemple d'image confocal. **E** Exemple de la sélection de la zone d'étude du plugin FOCI du logiciel ImageJ, créé en collaboration avec la plateforme Bio-image de l'IBMP (Jérome Mutterer) à partir de l'image D. Les cercles blancs représentent les point pris en comptent pour l'analyse. **F** Boxplot de la quantification de la co-localisation de URT1 et de ses délétions avec DCP1 par le plugin FOCI du logiciel ImageJ. Les crois représentes les valeurs extrêmes supérieures ou inférieures.

Les protéines sont co-exprimées avec CFP-DCP1 de façon transitoire dans des feuilles de *N. benthamiana* et observées au microscope confocal lsm780 (Zeiss) 2 à 3 jours après agro-infiltration. Les protéines avec une étiquette YFP sont pseudo-colorées en vert. La barre blanche représente 50 µm.

proches de 0 pour une distribution aléatoire de la fluorescence (Adler and Pagakis, 2003).

Les protéines ayant une étiquette YFP en N et en C terminal sont comparées à un contrôle positif, DCP2. DCP2 et DCP1 forment le complexe d'élimination de la coiffe (Introduction partie A3) et sont tous deux des marqueurs des *processing bodies* chez les plantes (Weber et al., 2008; Xu et al., 2006). Des coefficients de Pearson similaires sont obtenus pour YFP-DCP2, YFP-URT1 et URT1-YFP, indiquant une co-localisation similaire avec CFP-DCP1 (Figure 3.6C). Ces résultats confirment la localisation de URT1 au niveau des *processing bodies* (Chapitre 2 partie A) (Sement et al., 2013). Aucune différence significative n'est observée entre les protéines ∆1-25, ∆1-53, ∆1-160URT1 fusionnées avec une étiquette N ou C terminale et URT1 ou le contrôle positif DCP2 (test de Mann Whitney). Ces résultats indiquent que ni le motif M1, M2 ou M3 ne sont essentiels pour l'association de URT1 au *processing bodies*. A l'inverse, le coefficient de Pearson de YFP-Ntrf et Ntrf-YFP est significativement différent de URT1 et du contrôle positif DCP2 (P value de 2 $10e^{-6}$).

Les résultats obtenus par le plugin JACoP suggèrent que la région 161 à 418 de URT1 est essentielle à l'association de URT1 au niveau des *processing bodies*. Cependant, JACoP possède plusieurs inconvénients. En effet, le coefficient de Pearson est déterminé d'après tous les pixels de l'image et n'est donc pas limité aux objets d'intérêt. De plus, il ne prend pas en compte le nombre de foci dans les images. Ensuite, il n'est pas possible d'exclure que la localisation nucléo-cytoplasmique de Ntrf soit responsable d'une portion de la baisse de la co-localisation avec DCP1. Afin de palier ce problème et vérifier ces résultats, j'ai développé, en collaboration avec la plateforme Microscopie et Imagerie de l'IBMP (Jérome Mutterer), un protocole utilisant aussi ImageJ permettant la détection et l'analyse de l'intensité des spots, que nous avons appelé FOCI. Cette méthode permet une sélection automatique d'un nombre ajustable de spots dans une image selon leur intensité. Elle permet d'exclure les spots n'étant pas biologiques qui peuvent être créé par DCP1 et de palier l'influence de la localisation nucléo-cytoplasmique de Ntrf (Figure 3.6E). De plus, l'automatisation de la détection de spots est plus fiable et plus reproductible qu'une détection manuelle. Les foci sont sélectionnés selon un canal de fluorescence (protéine 1) et sont automatiquement comparés au second canal (protéine 2). Les images précédemment analysées par le plugin JACoP sont donc ré-analysées par FOCI (Figure 4.6F). En utilisant cette méthode plus restrictive, des résultats similaires à la méthode JACoP sont observés. En effet, la co-localisation de URT1 avec DCP1 est similaire à celle du contrôle

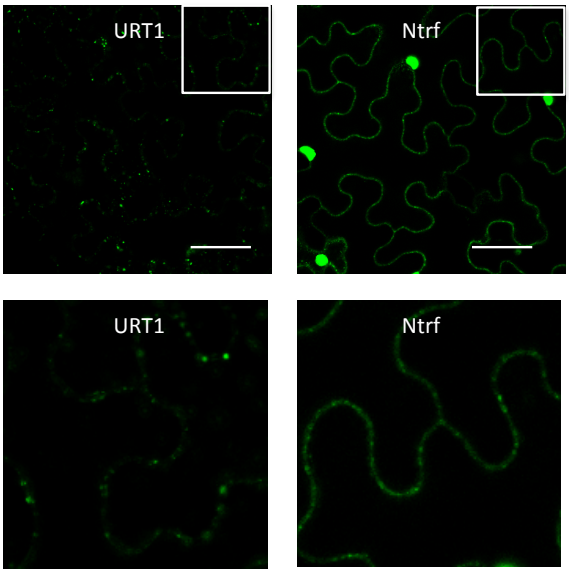


Figure 3.7 : La localisation de URT1 et de Ntrf après un stress thermique

Les images du panel du bas représentent un grossissement de 3 fois la partie encadrée de l'image au dessus d'elle. Les protéines sont exprimées de façon transitoire dans des feuilles de *N. benthamiana* et observées au microscope confocal lsm780 (Zeiss) après un traitement de 20 minutes à 39°C et 20 minutes de rétablissement à température ambiante, 2 à 3 jours après agro-infiltration. La barre blanche représente 50 µm.

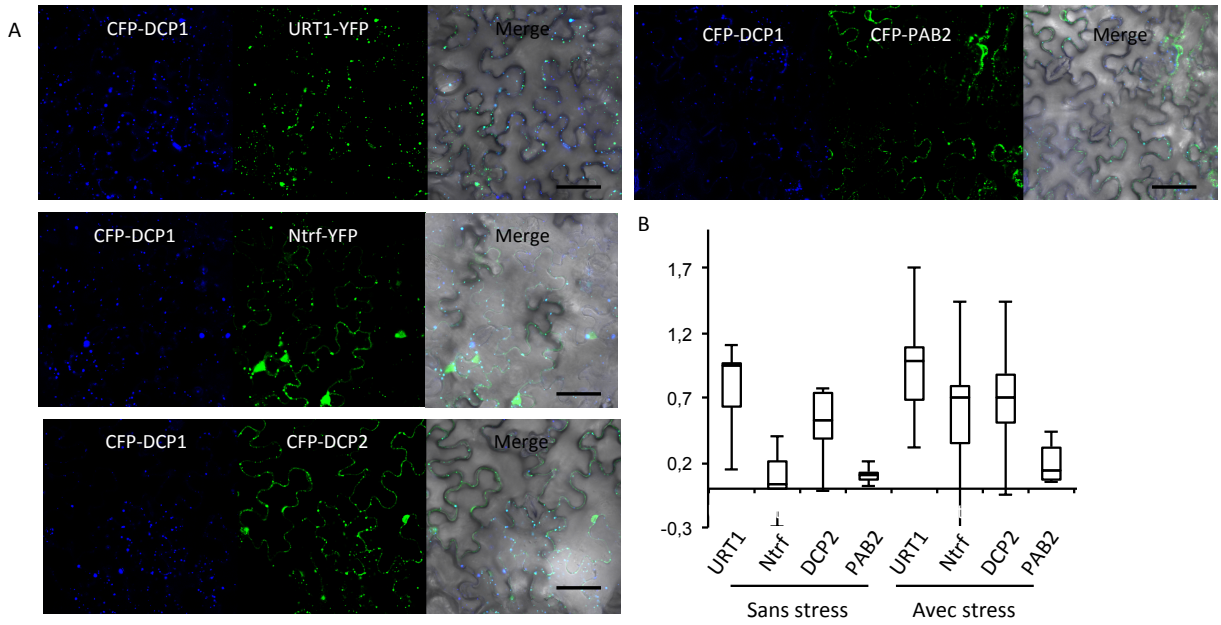


Figure 3.8 : Ntrf semble suffisant à la localisation de URT1 au niveau des *processing bodies* après un stress thermique

A Images de URT1 et de Ntrf co-exprimés avec CFP-DCP1 après un stress thermique

B Boxplot de la quantification de la colocalisation de URT1 et Ntrf avec DCP1 après un stress thermique analysé par le plugin FOCI (au total 55 images de URT1 et Ntrf avec stress).

Les protéines sont co-exprimées avec CFP-DCP1 de façon transitoire dans des feuilles de *N. benthamiana* et observées au microscope confocal lsm780 (Zeiss) après un traitement de 20 minutes à 39°C et 20 minutes de rétablissement à température ambiante, 2 à 3 jours après agro-infiltration. Les protéines avec une étiquette YFP sont pseudo-colorées en vert et celle avec une étiquette CFP en bleue. La barre noire représente 50 µm.

positif DCP2 et la différence entre URT1 et Ntrf est toujours présente ($p<0,01$). Seule l'utilisation de l'étiquette en N terminal pour Δ1-160UT1 montre une quantification de la co-localisation intermédiaire entre Δ 1-53URT1 et Ntrf (Figure 4.6F). Cette délétion de URT1 élimine une partie de la région P/Q/N/H de URT1 et le motif M3. La baisse de la co-localisation de Δ1-160URT1 avec DCP1 peut être due à une délétion d'une partie de la région P/Q/N/H et/ou au manque d'accessibilité de celle-ci par l'étiquette YFP qui a une taille de 26kDa et est située en N terminal (Figure 3.6F). En effet, cette baisse de la co-localisation de Δ1-160URT1 avec DCP1 n'est pas observée si l'étiquette YFP se situe en C terminal (Figure 3.6F). Sachant que l'influence de la partie N terminale de la protéine sur la localisation est étudiée, les résultats avec l'étiquette en C terminal de URT1 sont considérés plus fiables.

Les données obtenues par l'analyse des images avec FOCI confirment les données obtenues par le plugin JACoP. En effet, les acides aminés 1 à 160 de la région N terminale n'influencent pas la localisation de URT1 au niveau des *processing bodies* tandis que la région des acides aminés 161 à 418, contenant la région P/Q/N/H et la région riche en E est, elle, essentielle à l'adressage de URT1 aux *processing bodies*. J'ai tendance à favoriser l'implication du domaine P/Q/N/H dans la localisation de URT1 au niveau des *processing bodies* sachant que les régions riches en Q/N sont responsables de la localisation de plusieurs protéines dans les *processing bodies* (Buchan et al., 2008; Gilks et al., 2004; Reijns et al., 2008; Thomas et al., 2011).

3. Le domaine nucléotidyltransférase de URT1 semble être suffisant pour la localisation de URT1 au niveau des *processing bodies* après un stress thermique

Les *processing bodies* sont constitutivement formés dans la cellule mais ils sont plus nombreux en cas de stress (Buchan and Parker, 2009; Kedersha et al., 2005; Teixeira et al., 2005; Xu and Chua, 2011). Le mécanisme d'adressage des protéines en cas de stress au niveau des *processing bodies* peut être différent. Afin d'étudier l'adressage de URT1 dans les *processing bodies* en cas de stress, j'ai quantifié la co-localisation de URT1 et de Ntrf avec une étiquette YFP en C terminal avec DCP1 après un stress thermique effectué par un traitement de 20 minutes à 39°C suivit de 20 minutes de rétablissement à température ambiante. Trois expériences indépendantes ont été effectuées (avec un total de 55 images). En cas de stress, la localisation de Ntf change. En effet, Ntrf forme un nombre plus important de foci cytoplasmiques (Figure 3.7). Le nombre de ces foci cytoplasmiques n'est pas

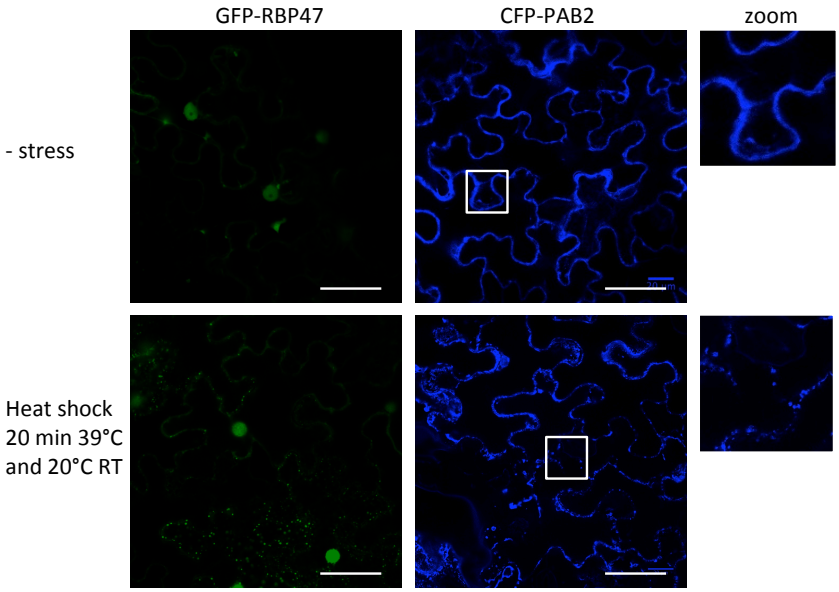


Figure 3.9 : La localisation des marqueurs des granules de stress après un stress thermique

Les protéines sont exprimées de façon transitoire dans des feuilles de *N. benthamiana* et observées au microscope confocal lsm780 (Zeiss) 2 à 3 jours après agro-infiltration. La protéine CFP-PAB2 est pseudo-colorée en bleu et GFP-RBP47 en vert. La barre blanche représente 50 µm.

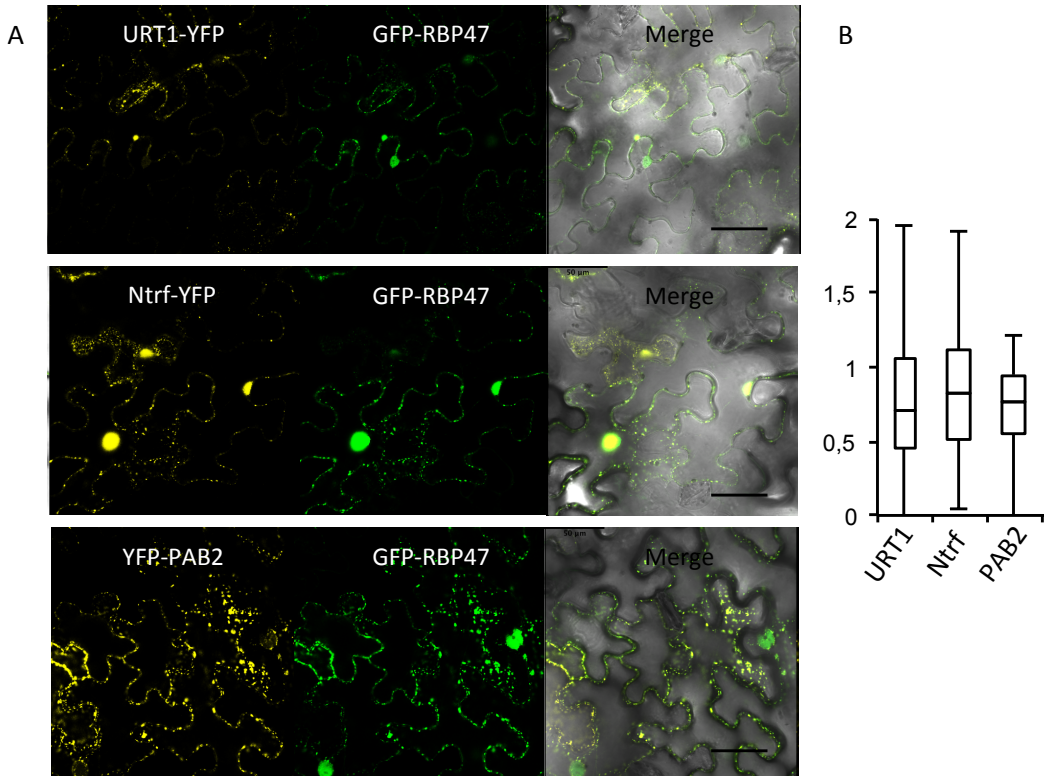


Figure 3.10 : Ntrf serait suffisant à la localisation de URT1 au niveau des granules de stress avec le marqueur des granules de stress RBP47

A Images de URT1 et de Ntrf co-exprimés avec GFP-RBP47 après un stress thermique. **B** Boxplot de la quantification de la colocalisation de URT1 et Ntrf avec RBP47 après un stress thermique analysé par le plugin FOCI (au total 54 images de URT1 et Ntrf).

Les protéines sont co-exprimées avec GFP-RBP47 de façon transitoire dans des feuilles de *N. benthamiana* et observées au microscope confocal lsm780 (Zeiss) après un traitement de 20 minutes à 39°C et 20 minutes de rétablissement à température ambiante, 2 à 3 jours après agro-infiltration. Les protéines avec une étiquette GFP sont pseudo-colorées en vert et celle avec une étiquette YFP en jaune. La barre noire représente 50 µm.

évident à quantifier. En effet, la taille des cellules varie et le contour des cellules n'est pas facile à déterminer, de plus les cellules n'ont pas exactement le même plan dans une image.

D'après la quantification de la co-localisation de URT1 et Ntrf avec DCP1 par FOCI, la co-localisation de Ntrf et de DCP1 est similaire à celle de DCP2 (Figure 3.8). Donc en cas de stress, le domaine nucléotidyltransférase semble être suffisant à la localisation de URT1 au niveau des *processing bodies*. Ces résultats sont bien sûr à prendre avec les précautions de rigueur car aucun contrôle cycloheximide n'a pu être effectué. Afin de confirmer ces résultats et de suivre la formation des *processing bodies* en cas de stress thermique, il faudrait être capable d'obtenir des plantes transformées de façon stable avec la construction YFP-URT1 et YFP-Ntrf (Discussion partie C).

4. Le domaine nucléotidyltransférase de URT1 serait aussi suffisant pour la localisation de URT1 au niveau des granules de stress

URT1 co-localise avec des marqueurs de granules de stress (Chapitre 2 partie A) (Sement et al., 2013). Afin d'étudier l'adressage de URT1 aux granules de stress, j'ai co-exprimé URT1 et Ntrf en présence d'un marqueur des granules de stress GFP-RBP47 ou CFP-PAB2 (Weber et al., 2008). Les disques foliaires sont donc préalablement stressés 20 minutes à 39°C et observés au microscope confocal après 20 minutes de rétablissement à température ambiante. J'ai effectué une quantification de la co-localisation de URT1 et de ses délétions avec l'un des marqueur des granules de stress, en comparaison d'un contrôle positif, PAB2 pour le marqueur RBP47 (Weber et al., 2008) ou LARP1 pour le marqueur PAB2 (communication personnelle des Dr. Bousquet-Antonelli, Merret et Deragon du Laboratoire Génome et Développement des Plantes à Perpignan). RBP47, l'homologue de TIA1/TIAR, est localisée au niveau du noyau et est exportée au niveau des granules de stress après un stress thermique (Figure 3.9). PAB2, quant à elle, est localisée sous forme diffuse dans le cytoplasme et est relocalisée au niveau des granules de stress après un stress thermique (Figure 3.9).

Sachant que RBP47 reste partiellement localisée dans le noyau, j'ai utilisé la méthode FOCI afin de quantifier automatiquement la co-localisation des images obtenues. Celle-ci montre que URT1 co-localise avec RBP47 et PAB2, conformément avec les résultats précédemment obtenus (Figure 3.10 et 3.11) (Chapitre 2 partie A) (Sement et al., 2013). Mais la quantification de la co-localisation

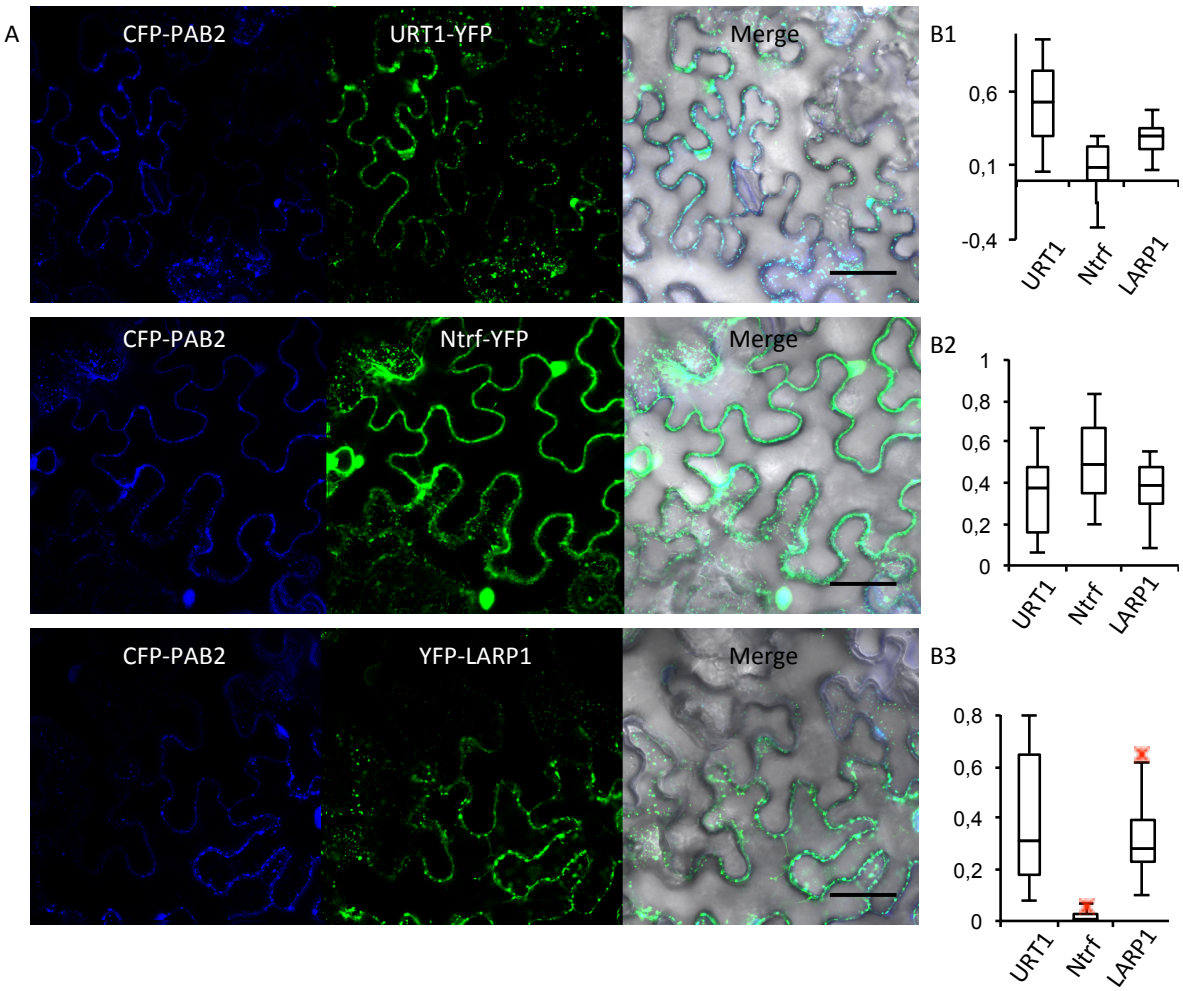


Figure 3.11 : Ntrf est suffisant à la localisation de URT1 au niveau des granules de stress avec le marqueur PAB2

A Images de URT1 et de Ntrf co-exprimés avec CFP-PAB2 après un stress thermique. **B** Boxplot de la quantification de la co-localisation de URT1 et Ntrf avec PAB2 après un stress thermique analysé par le plugin FOCI, B1, B2 et B3 correspondent respectivement à la série 1, 2 et 3. Les croix rouges représentent les valeurs extrêmes supérieures (au total 43 images de URT1 et Ntrf).

Les protéines sont co-exprimées avec CFP-PAB2 de façon transitoire dans des feuilles de *N. benthamiana* et observées au microscope confocal lsm780 (Zeiss) après un traitement de 20 minutes à 39°C et 20 minutes de rétablissement à température ambiante, 2 à 3 jours après agro-infiltration. Les protéines avec une étiquette YFP sont pseudo-colorées en vert et celle avec une étiquette CFP en bleu. La barre noire représente 50 µm.

de Ntrf avec le marqueur des granules de stress varie selon le marqueur utilisé. En effet, la quantification de la co-localisation de Nrtf n'est pas homogène entre les différentes expériences utilisant le marqueur PAB2 (séries 1 à 3).

En analysant les données avec le marqueur RBP47 (un total de 53 images de URT1 et Ntrf issues de trois expériences indépendantes), URT1 et Ntrf co-localisent d'une façon similaire avec le contrôle positif PAB2 (Figure 3.10). Le domaine nucléotidyltransférase serait suffisant à la localisation de URT1 dans les granules de stress en cas de stress thermique, comme pour la localisation de URT1 aux *processing bodies*. En effet, les foci de URT1 et des marqueurs de stress peuvent colocaliser. Cependant, en utilisant le marqueur PAB2, la quantification de la co-localisation de Ntrf n'est pas homogène d'après le plugin FOCI car les images semblent être visuellement similaires entre les expériences (Figure 3.11A). Deux répétitions montrent une différence significative de la co-localisation de Ntrf avec PAB2 ($p<0,01$) tandis qu'une expérience montre une co-localisation similaire de URT1, Ntrf et le contrôle positif LARP1 avec PAB2 (un total de 43 images de URT1 et Ntrf) (Figure 3.11B). La différence de co-localisation de URT1 et de Ntrf avec PAB2 n'est pas visible à l'œil nu à partir des images, seule l'analyse avec le plugin FOCI permet de la mettre en évidence (Figure 3.11).

5. Conclusion et ouverture

Les motifs conservés M1, M2 et M3 chez les homologues de URT1 chez les Angiospermes ne sont pas impliqués dans la localisation de la protéine. Par ailleurs, la localisation de URT1 ne dépend pas non plus de son activité uridylyltransférase. Dans les expériences de co-localisation de URT1 avec les marqueurs des *processing bodies* et les granules de stress, nous avons tout d'abord utilisé la méthode JACoP pour quantifier la co-localisation. D'autre part, nous avons mis en place une nouvelle méthode de détection des spots, appelée FOCI. Elle permet de sélectionner les spots pour l'étude de la co-localisation, éliminant certains artéfacts et permettant une meilleure quantification de la co-localisation de deux protéines. Cette méthode est automatique, rapide, facilement modulable et transposable à d'autres expériences.

D'après ces expériences, une partie de la région N terminale, des acides aminés 161 à 418 et plus probablement la région riche en P/Q/N/H, est essentielle pour l'adressage de URT1 au niveau des *processing bodies*. Différemment, la région N terminale de URT1 ne semble pas avoir d'impact sur la localisation de URT1 au niveau des *processing bodies* ni des granules de stress en cas de stress thermique.

Cependant, l'implication du domaine nucléotidyltransférase de URT1 dans la localisation de URT1 au niveau des granules de stress reste à vérifier. En effet, puisque certains résultats sont contradictoires, plusieurs répétitions de l'expérience avec le marqueur PAB2 nécessitent d'être effectuées. De plus, il est essentiel d'effectuer un contrôle en inhibant la formation des granules de stress, par du cycloheximide par exemple, afin de vérifier que les foci cytoplasmiques observés après un stress thermique ne sont pas des agrégats de protéines ne contenant pas d'ARNm (Discussion partie C).

Le rôle de URT1 au niveau de ces foci cytoplasmiques n'est pas encore résolu mais URT1 est impliquée dans la dégradation des ARNm dans le sens 5'-3' (Chapitre 2 partie A), pouvant expliquer sa localisation au niveau des *processing bodies.* Cependant, URT1 pourrait aussi inhiber la traduction des ARNm déadénylés comme suggéré pour l'ajout de CUCU aux ARNm ayant un codon stop précoce chez *A. nidulans* ou pour l'ajout d'uridines aux ARNm par XTUT7 chez *Xenopus laevis* (Lapointe and Wickens, 2013; Morozov et al., 2012). Il est possible que URT1 inhibe la traduction des ARNm directement ou *via* la liaison d'une protéine de liaison aux ARNm uridylés comme par exemple les protéines La-*related* qui se lient aux ARNm uridylés comme LARP1 chez *C. elegans* (Nykamp et al., 2008). Par ailleurs, il est possible que l'uridylation des ARN soit un processus rapide et réversible en réponse au stress et que les ARNm puissent retourner à la traduction (Discussion partie D5b).

Les motifs M1, M2 et M3 de URT1, conservés chez les homologues des Angiospermes, ne sont pas impliqués dans la localisation de la protéine et ne semble pas non plus impliqués dans l'interaction protéine-ARN. En effet, pour l'instant, aucune interaction stable de URT1 avec de l'ARN n'a été mise en évidence (Chapitre 2 partie B). Cependant, ils pourraient être impliqués dans des interactions protéine-protéine comme le prédit le logiciel ANCHOR et la prédiction de structures secondaires en hélice de ces motifs (Chapitre 1 et Discussion partie B) (Dunker et al., 2001; Uversky, 2013a, 2013b; Uversky and Dunker, 2013)

Discussion

A. URT1 est une uridylyltransférase intrinsèquement spécifique pour l'UTP

La spécificité nucléotidique de plusieurs uridylyltransférases a été précédemment étudiée et leur structure 3D a été résolue au niveau atomique. La capacité de discrimination intrinsèque envers les différents ribonucléotides varient selon les uridylyltransférases (Aphasizhev and Aphasizheva, 2007, 2008a; Lapointe and Wickens, 2013; Rissland and Norbury, 2008; Stagno et al., 2007). Cette capacité de discriminer les différents nucléotides pour les ncPAP est comparée dans les mêmes conditions d'activité en présence de l'un des quatre ribonucléotides. Par exemple, *Tb*RET1 des trypanosomes n'incorpore aucun ATP dans le même laps de temps qu'elle peut ajouter des dizaines d'UTP (Aphasizhev and Aphasizheva, 2007). Cid1 de *S. pombe* peut incorporer autant d'ATP que d'UTP (Rissland and Norbury, 2008). Différemment, la spécificité peut changer en présence d'un cofacteur comme pour Star-PAP/U6 TUTase, spécifique des UTP qui peut devenir une PAP lorsqu'elle interagit avec le cofacteur CPSF (Minasaki and Eckmann, 2012).

URT1 est une uridylyltransférase intrinsèquement spécifique pour les UTP, ayant une capacité accrue de discrimination entre l'UTP de l'ATP par rapport à Cid1. URT1 conserve les mêmes acides aminés (D708, E333, H714) précédemment impliqués dans la spécificité pour l'UTP chez d'autres uridylyltransférases (Aphasizhev and Aphasizheva, 2008b; Deng et al., 2005; Lapointe and Wickens, 2013; Lunde et al., 2012; Munoz-Tello et al., 2012; Ringpis et al., 2010; Stagno et al., 2007, 2010; Yates et al., 2012). URT1 a une donc capacité de discriminer l'UTP des quatre ribonucléotides semblable aux uridylyltransférases des trypanosomes qui pourtant ne possède pas d'équivalent à l'histidine mise en évidence chez Cid1 (Lunde et al., 2012; Munoz-Tello et al., 2012; Yates et al., 2012). Il doit donc exister d'autres acides aminés impliqués dans la spécificité nucléotidique chez URT1, permettant de discriminer l'UTP de l'ATP. URT1 semble être un très bon modèle pour l'étude de la spécificité des uridylyltransférases pour les UTP et sa structure n'a pas encore été résolue au niveau atomique (Chapitre 2).

La résolution au niveau atomique de URT1 permettra de mettre en évidence les acides aminés impliqués dans des liaisons avec les substrats ARN et les nucléotides. Cependant, la résolution au niveau atomique d'une protéine est un travail laborieux, variant d'une protéine à l'autre et URT1 contient une région N terminale intrinsèquement désorganisée qui risque d'être très flexible, instable, rendant surement la résolution de la protéine entière quasi impossible. Afin de palier ce problème, une collaboration est en cours afin d'effectuer la résolution atomique du

domaine nucléotidyltransférase seul avec le laboratoire du Dr Meinhart du Max-Planck-Institute of Medical research d'Heidelberg en Allemagne. Celle-ci permettra de découvrir de nouveaux acides aminés important pour la sélection nucléotidique pour l'UTP et les acides aminés essentiels à la discrimination entre l'UTP et l'ATP pourront être mis en évidence en effectuant une comparaison de la structure atomique de URT1 avec celle des autres uridylyltransférases.

B. L'uridylation des ARNm oligoadénylés par URT1 chez *A. thaliana*

1. La reconnaissance des substrats ARN par URT1

Les ncPAP peuvent contenir des domaines de liaison à l'ARN tel que Star-PAP/U6 TUTase, PAPD5 ou XTUT7 (Lapointe and Wickens, 2013; Rammelt et al., 2011). D'autres uridylyltransférases *Tb*RET2 et *Tb*TUT4 des trypanosomes possèdent une spécificité intrinsèque pour leur substrat ARN qui est apportée par la conformation de leur site actif (Deng et al., 2005; Stagno et al., 2007). Cependant, pour la plupart des ncPAP, la spécificité de substrat ARN provient d'une interaction avec des protéines de liaison à l'ARN (Hagan et al., 2009).

URT1 uridyle *in vitro* les ARN lorsque leur extrémité 3' est libre, non engagée dans une structure tige-boucle. De plus, il faut que ce dernier nucléotide en 3' ne soit pas non plus modifié en 2O', comme par exemple une méthylation. En effet, le 2O' de l'ARN est vraisemblablement nécessaire à la stabilité de l'ARN dans le site actif de URT1. Lorsque le substrat ARN possède une extrémité 3' libre, non méthylée en 2'O, URT1, peut uridyler *in vitro* tous les ARN testés comme les ARNm, les ARN ribosomiques et les ARN de transfert. URT1 n'uridyle pas non plus spécifiquement *in vitro* les ARN comportant plusieurs adénines en 3' (Chapitre 2), même si, *in vivo*, URT1 est très spécifique des ARNm oligoadénylés. En effet, URT1 n'est pas essentielle pour l'uridylation des microARN ni des fragments d'ARN clivés par le complexe RISC (Thèse de François Sement et résultats obtenus par Hélène Zuber). La spécificité de URT1 pour les ARN oligoadénylés n'est donc pas intrinsèquement liée à la conformation de son site actif.

Jusqu'à présent, aucune interaction stable entre URT1 et l'ARN n'a été mise en évidence par différentes méthodes (Biacore™, co-précipitation). Celle-ci est très probablement inexistante car URT1 est distributive et celle-ci peut indiquer une interaction faible avec l'ARN. L'interaction de URT1 et de son substrat ARN, *in vivo*, est très probablement apportée par un partenaire protéique, apportant la spécificité

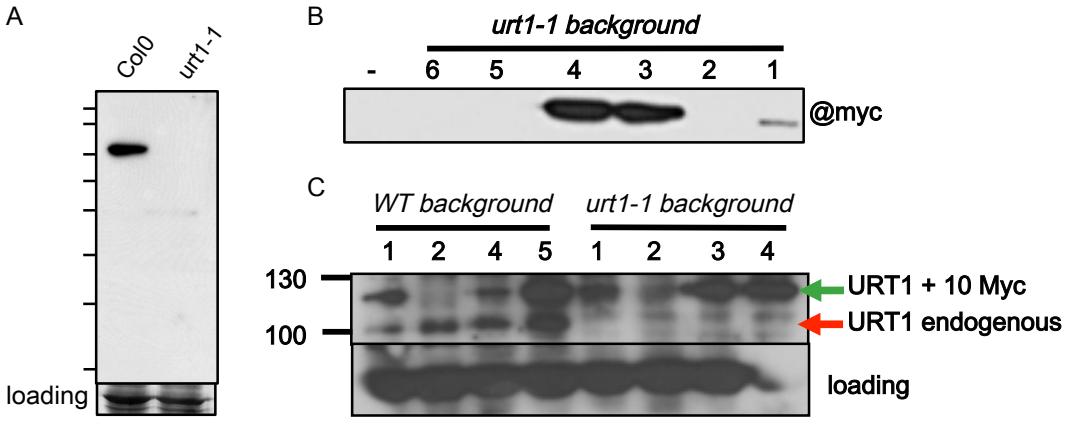


Figure D1 : Outils obtenus lors de ma thèse pour la recherche de protéines partenaires de URT1

A WB utilisant les anticorps dirigés contre URT1 (1/20 000e) sur des fleurs Col0 ou *urt1-1*. **B** Les plantes transgéniques myc-URT1. WB avec les anticorps dirigés contre l'étiquette c-myc sur les plantes transgéniques exprimant URT1 avec une étiquette c-myc en N terminal dans le mutant *urt1-1*. **C** WB avec l'anticorps dirigé contre URT1 sur les plantes transgéniques exprimant URT1 avec une étiquette c-myc en N terminal dans un fond génétique WT ou *urt1-1*. Ce WB sert à quantifier l'expression de URT1 dans les plantes transgéniques exprimant c-myc-URT1 par rapport à l'expression endogène de URT1 (B et C réalisés par Hélène Zuber).

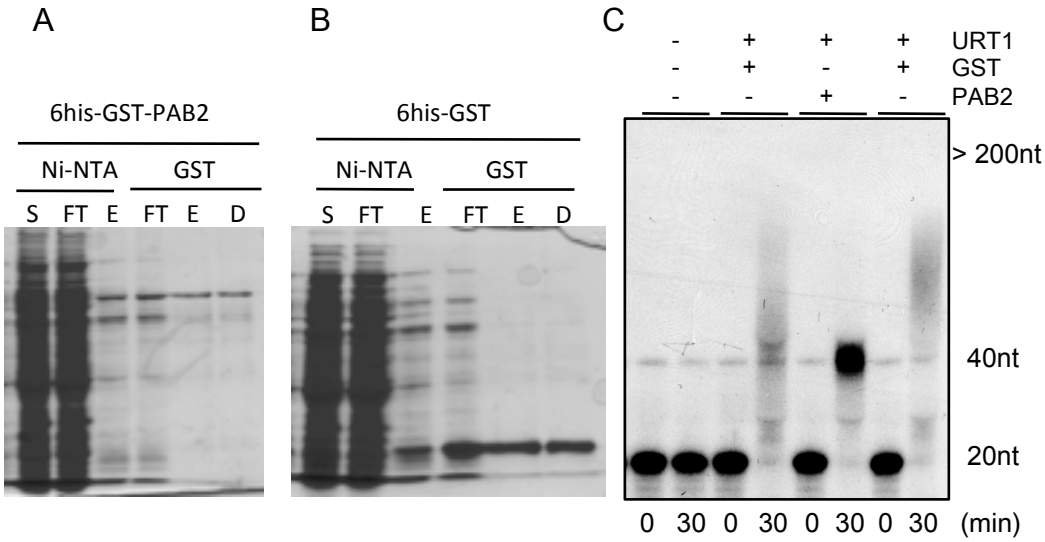


Figure D2 : PAB2 inhibe l'activité de URT1

A Double purification de 6his-GST-PAB2. **B** Double purification de 6his-GST *via* l'étiquette 6histidines avec NiNTA et GST avec la GST Sépharose Fast Flow. **C** PAB2 inhibe l'activité de URT1. L'activité de URT1 est effectuée en présence de 6his-GST-PAB2 ou de 6his-GST (1/5). S, surnageant, FT, *flow-through*, E, élution et D, échantillon dialysé durant 4h.

URT1 est capable d'ajouter un nombre d'UTP plus importante en présence de GST qu'en présence de PAB2. En effet, en présence de PAB2, URT1 semble ajouter au maximum 20 UTP.

de substrat ARN. Cependant, à l'heure actuelle, aucun partenaire protéique n'a été encore mis à jour. Pourtant, durant ma thèse, une recherche d'interactants par un crible double hybride a été sous-traitée à la compagnie Hybrigenics mais il n'a pas révélé de partenaire en raison d'un trop fort taux d'autoactivation des gènes rapporteurs par URT1. Cette autoactivation est probablement due à la région désorganisée qui pourrait provoquer des interactions non-spécifiques avec un grand nombre de protéines. D'autre part, j'ai effectué des tests de co-immunoprécipitation à partir de plantes exprimant une version étiquetée de URT1 après transformation stable ou transitoire de plantes mais aucun interactant de URT1 n'a été mis en évidence après analyse des échantillons en spectrométrie de masse (Figure D1). Afin d'améliorer la stabilité des interactions protéine-protéine, il serait possible d'effectuer préalablement une étape de *crosslink.*

2. La reconnaissance des ARN oligoadénylés par URT1

La reconnaissance des ARNm oligodénylés par URT1 pourrait être effectuée par l'interaction de URT1 et d'une PABP. Les ARNm oligoadénylés uridylés par URT1 possèdent en moyenne 14 adénines et, il est estimé que 10 à 12 A sont suffisantes à la liaison de PABP. Contrairement à la levure et les animaux, chez Arabidopsis, il existe huit PABP ayant une expression différentielle au cours du développement. Les PAB3, PAB5 et PAB7 sont exprimées dans le pollen et les PABP de la classe II : PAB2, PAB4 et PAB8 sont les plus exprimées dans la plante (Belostotsky, 2003; Honys and Twell, 2004). Seul un triple mutant *pab2/pab4/pab8* n'est pas viable. Il est possible, chez Arabidopsis, que la queue polyA ne soit pas recouverte d'un seul type de PABP mais plutôt d'un mélange d'entre les trois PABP de la classe II. Seule la PABP de la classe II PAB2 complémente la délétion de pabp1 chez la levure (Palanivelu et al., 2000). J'ai testé l'hypothèse de l'interaction de URT1 et PAB2. Cependant, je n'ai pas mis en évidence d'interaction lors de tests de co-purification par expression transitoire dans des feuilles de tabac ou par l'utilisation des protéines recombinantes ni lors de tests utilisant la technologie de la résonance plasmonique avec les protéines recombinantes (Biacore™). Lors de tests de l'activité catalytique de URT1 en présence de la protéine recombinante PAB2, PAB2 inhibe l'activité de URT1 (Figure D2). En effet, en présence de PAB2, URT1 est capable d'ajouter seulement un maximum de 20 uridines. Tandis que dans le même laps de temps en absence de PAB2, elle est capable d'en ajouter beaucoup plus. Il est possible que PAB2 ne soit pas la première PABP associée à la queue polyA mais plutôt lorsque la queue polyA est plus longue. On peut donc imaginer que

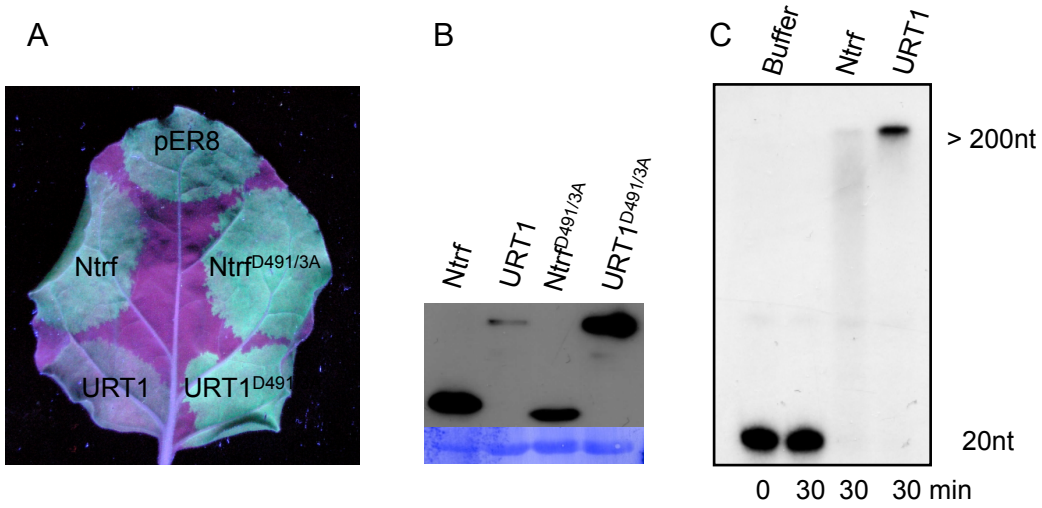


Figure D3 : La région en N terminal de URT1 est essentielle à l'activité optimale de URT1

A Test fonctionnel de URT1 en présence d'un gène rapporteur, la GFP. La GFP est co-agro-infiltrée dans une feuille de tabac, en présence du suppresseur de *silencing* P19 et de soit URT1, actif ou inactif, soit son domaine nucléotidyltransférase, actif ou inactif ou de pER8. pER8 est un vecteur vide, il sert de contrôle de l'expression de la GFP dans une feuille de tabac 4,5 jours après agro-infiltration. L'expression de la GFP est plus faible en présence de URT1. Cette baisse de la fluorescence en présence de URT1 comparée à pER8 est due à l'uridylation car en présence du mutant inactif URT1$^{D491/3A}$, la GFP est exprimée au même niveau que le contrôle pER8. L'impact de l'uridylation sur la faible expression de la GFP est étudié plus amplement dans la Thèse de François Sement. En présence du domaine nucléotidyltransférase de URT1 (Ntrf), la GFP est autant exprimée que dans le contrôle pER8. La région N terminale de URT1 est donc essentielle au phénotype de URT1 (GFP faiblement exprimée). **B** L'expression de la GFP dans le patch Ntrf n'est pas due à une faible expression du domaine nucléotidyltransférase de URT1. **C** *In vitro*, domaine nucléotidyltransférase de URT1 est toujours capable d'uridyler les ARN.

la première PABP associées aux queues polyA soit PAB4 ou PAB8 et qu'ensuite seulement intervient PAB2. PAB2 empêcherait URT1 d'uridyler les ARNm polyadénylés. Il est possible que URT1 interagisse avec PAB4 ou PAB8, permettant ainsi la reconnaissance des ARNm oligoadénylés. Cependant, il est aussi possible que URT1 uridyle les ARNm ayant en moyenne 14 adénines car les PAB2 ne sont plus liées aux ARNm oligoadénylés.

3. L'implication de la structure de la région N terminale de URT1 dans la reconnaissance de substrat ARN

La moitié de la protéine URT1 est prédite intrinsèquement désorganisée. Les régions intrinsèquement désorganisées des protéines ont été longtemps négligées car seulement la partie structurée des protéines semblait importante pour leur fonction. Il est maintenant évident que ces régions jouent également des rôles fondamentaux. Les régions intrinsèquement désorganisées sont impliquées dans la régulation de la fonction des protéines, les interactions protéine-protéine et/ou protéine-ARN et la localisation cellulaire (Mészáros et al., 2011; Tompa, 2005; Uversky, 2013; Uversky and Dunker, 2013).

Une expérience a été précédemment mise en place dans le laboratoire, montrant que lors d'une co-expression d'un gène rapporteur, la GFP, en présence de URT1 active est faiblement exprimée (Thèse de François Sement de 2012). Lorsque ce même test est effectué seulement en présence du domaine nucléotidyltransférase de URT1, le phénotype n'est plus observé (Figure D3). La région N terminale de URT1 est donc essentielle pour diminuer l'expression du gène rapporteur, la GFP. Une de nos hypothèses de travail est que la région N terminale de URT1 peut être impliquée dans la reconnaissance du substrat ARN *via* une protéine partenaire. De manière intéressante, beaucoup de protéines impliquées dans la voie de dégradation des ARN possèdent une ou des régions intrinsèquement désorganisées et ces régions sont impliquées dans des interactions protéines-protéines (Espinosa Angarica et al., 2013; Kalmar et al., 2012; Sun et al., 2013). L'interaction entre URT1 et son partenaire putatif ne semble cependant pas être stable puisque nous n'avons pu l'identifier par co-purification avec URT1. Une deuxième hypothèse pourrait également expliquer que l'expression de la GFP ne soit pas diminuée quand la GFP est co-exprimée avec le domaine nucléotidyltransférase seul. En effet, il est possible que l'activité du domaine nucléotidyltransférase de URT1 soit inférieure à celle de URT1 et qu'elle soit insuffisante pour permettre l'inhibition de l'expression de la GFP (Figure D3C). Cette hypothèse n'est pas à écarter mais il est important de remarquer

que le niveau d'expression du domaine nucléotidyltransférase est très supérieur à celle de URT1 (Figure D3B). Ces expériences restent préliminaires mais ouvrent des perspectives intéressantes. L'analyse moléculaire (niveaux d'expression, d'uridylation, de la coiffe…) des transcrits GFP dans les différentes conditions devrait permettre de mieux comprendre l'implication de la région N terminale dans la diminution de l'expression du gène rapporteur et de tester l'effet sur les gènes endogènes.

4. Le rôle potentiel des motifs conservés de URT1 dans la reconnaissance de substrats ARN

Les zones d'interaction protéine-protéine et protéine-ARN sont conservées entre les différentes espèces car elles sont soumises à une forte pression de sélection. En effet, les régions d'interactions entre les protéines et leurs partenaires doivent co-évoluer pour préserver leur interaction (Mészáros et al., 2011; Uversky and Dunker, 2013). Cependant, les régions prédites intrinsèquement désorganisées peuvent être aussi impliquées dans des interactions protéines-protéines (Dosztányi et al., 2009; Mészáros et al., 2009).

URT1 conserve deux motifs M1 et M2 chez les Angiospermes. Ils sont entourés de deux régions prédites intrinsèquement désorganisées (Chapitre 1 et 3). La conservation des motifs M1 et M2 chez les Angiospermes, alors que la séquence primaire de la région N terminale d'URT1 évolue rapidement, reflète très probablement une préservation de leur fonction biologique telle que la localisation ou l'interaction protéine-protéine ou protéine-ARN de URT1. Cependant, les motifs M1 et M2 ne sont pas impliqués dans la localisation de URT1 (Chapitre 3) et actuellement, aucune liaison stable n'a pu être mise en évidence entre URT1 et de l'ARN. Donc la fonction la plus probable des motifs M1 et M2 est l'interaction protéine-protéine même si à ce jour nous n'avons pu identifier de protéines partenaires. Les données de prédictions de structures secondaires et d'interaction protéine-protéine soutiennent cette hypothèse. Le partenaire de URT1 s'associant à ces motifs M1 et M2 pourrait apporter la spécificité de substrat ARN de URT1 pour les ARNm oligoadénylés observée *in vivo*.

Sachant que le mutant *urt1-1* accumule des transcrits tronqués, il peut être transformé avec la construction de URT1 en absence des deux motifs M1 et M2 conservés afin de tester s'ils ont un rôle dans la reconnaissance de substrat ARN. En effet, si les motifs M1 et M2 sont impliqués dans la reconnaissance de substrat, cette construction ne devrait pas permettre de complémenter le phénotype du mutant *urt1-*

1. D'autre part, un crible double hybride pourrait être effectué avec ces deux motifs afin de rechercher des partenaires. En effet, cette région risque d'être moins autoactivante vis à vis des gènes rapporteurs que la protéine entière URT1.

5. URT1 peut éventuellement uridyler d'autres ARN que les ARNm

Il existe très souvent des redondances d'activité des ncPAP dans la modification post-transcriptionnelle des ARN comme chez l'homme, les mammifères, les algues vertes et *A. thaliana* (Ibrahim et al., 2010; Lapointe and Wickens, 2013; Ren et al., 2012; Wyman et al., 2011; Zhao et al., 2012). Par exemple, chez *A. thaliana*, HESO1 est principalement responsable de l'uridylation des microARN dans le mutant *hen1-1*. Cependant, dans le mutant *heso1-1*, l'uridylation des microARN n'est pas abolie, révélant l'existence d'une autre uridylyltransférase impliquée dans leur uridylation (Ren et al., 2012; Zhao et al., 2012).

Nos résultats montrent que URT1 est essentielle pour l'uridylation des ARNm déadénylés mais qu'elle n'est pas essentielle pour l'uridylation des microARN ni des fragments d'ARNm clivés par le complexe RISC (Chapitre 2, Thèse de François Sement en 2012 et résultats obtenus dans le laboratoire). Cependant, sachant que URT1 est localisée au niveau des *processing bodies,* il est possible que URT1 intervienne secondairement dans ces mécanismes d'uridylation. En effet, les fragments d'ARNm clivés par le complexe RISC sont localisés au niveau des *processing bodies* (Pomeranz et al., 2010). L'uridylyltransférase principalement impliquée dans l'uridylation des fragments d'ARN clivés par le complexe RISC est HESO1 (résultat non-publié du laboratoire). De manière intéressante HESO1 et URT1 co-localisent dans les foci cytoplasmiques (Thèse de François Sement en 2012). Il n'est donc pas exclu que URT1 puisse être impliquée aussi dans l'uridylation des microARN et/ou des fragments d'ARNm clivés par le complexe RISC. Cette hypothèse est actuellement testée dans le laboratoire par l'étude du triple mutant nul *urt1-1*/*hen1-1*/*heso1-1*.

C. La localisation de URT1 au niveau des foci cytoplasmiques

Les transcrits non traduits s'accumulent en foci cytoplasmiques tels que les *processing bodies* et les granules de stress. Ces deux types de foci cytoplasmiques contiennent des protéines communes ou spécifiques des deux types de granules (Buchan and Parker, 2009; Kedersha et al., 2005; Thomas et al., 2011). Les régions

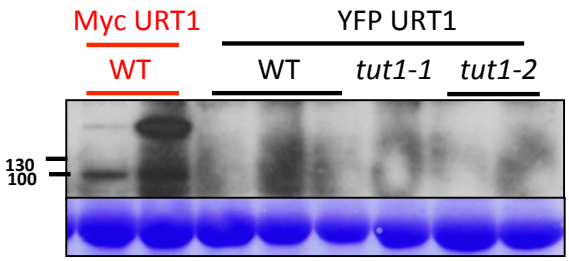


Figure D4 : L'étiquette YFP en N terminal de URT1 provoque du *silencing* de URT1

Le Western blot a été effectué avec les anticorps dirigés contre URT1 (réalisé par Hélène Zuber) avec des plantes transformées 35S-YFP-URT1 dans différents génotypes, Col0 et deux mutant de urt1 (*urt1-1* et *urt1-2*). Les plantes contrôles 35S-myc-URT1 dans Col0 montre qu'en utilisant l'étiquette YFP, du *silencing* est mis en place.

enrichies en Q et en Q/N, les domaines prions-*like*, sont responsables de la localisation de protéines dans les *processing bodies* mais aussi dans les granules de stress (Reijns et al., 2008; Thomas et al., 2011). URT1 contient des régions enrichies en Q et en P/Q/N/H au niveau de la région N terminale intrinsèquement désorganisée, conservée chez les Angiospermes (Chapitre 1 et 3). Seulement la région riche en P/Q/N/H et la région riche en E sont essentielles à la localisation de URT1 dans les *processing bodies.* Cependant, la région N terminale n'est responsable de la localisation de URT1 au niveau des *processing bodies* en cas de stress thermique. L'adressage de URT1 au niveau des granules de stress n'est pas encore résolu. Cependant, la partie N terminale ne semble pas être essentielle. De ce fait, Il existe deux mécanismes différents d'adressage de URT1 au niveau des foci cytoplasmiques en cas de stress ou non.

Des résultats préliminaires montrent que l'activité uridylyltransférase n'intervient pas dans cette localisation. Cependant, ces expériences sont à répéter et confirmer par l'utilisation d'un contrôle inhibant la formation des foci cytoplasmiques. Par exemple, il faudrait utiliser un contrôle comme la cycloheximide permettant d'inhiber la formation de foci cytoplasmiques tels que les *processing bodies* et les granules de stress afin de vérifier que les foci observés après un stress thermique ne sont pas des agrégats protéiques.

L'impact de la région N terminale sur la localisation de URT1 est étudié par l'expression transitoire des différentes constructions de la protéine URT1 en présence de marqueurs de *processing bodies* ou de granules de stress. Afin de confirmer la formation des *processing bodies* et des granules de stress, il faudrait des plantes transgéniques pour effectuer un contrôle avec la cycloheximide qui inhibe leur formation. Cependant, ce test nécessite des plantes transformées de façon stable exprimant les différentes constructions de URT1 étiquetée avec la GFP avec ou sans sa région N terminale. L'obtention de plantes transgéniques exprimant URT1 a été très problématique. En effet, nous avons tout d'abord utilisé l'ADNc, avec différentes étiquettes en N ou en C terminal et nous avons dû cribler plus d'une centaine de plantes avec ces différentes constructions sans succès avec URT1 ou URT1$^{D491/3A}$. Nous avons réussi à obtenir des plantes transgéniques exprimant URT1 avec une construction contenant les introns et le terminateur endogène de URT1 et une étiquette c-myc en N terminal. Cependant, l'expression de URT1 dépend aussi de l'étiquette ajoutée en N terminal. En effet, l'obtention des plantes transgéniques avec une étiquette en YFP n'est pas possible car du *silencing* est systématiquement mis en place (Figure D4).

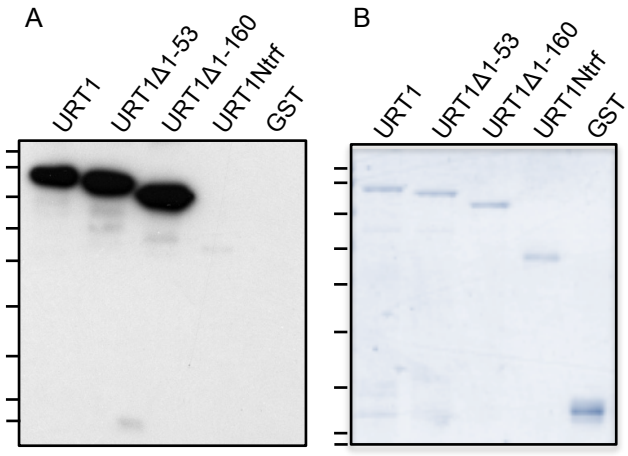


Figure D5 : Les anticorps dirigés contre URT1 ne reconnaissent pas le domaine nucléotidyltransférase

Les protéines recombinantes des constructions de URT1 sont exprimées chez *E. coli* avec deux étiquettes en N terminal, 6 histidines et GST et sont purifiées successivement contre l'étiquette 6histidines par le NiNTA et contre l'étiquette GST par la GST Sépharose FAst Flow.

URT1Ntrf est le domaine nucléotidyltransférase de URT1 et GST correspond au contrôle de spécificité 6 histidines et la GST.

Afin de confirmer la localisation de URT1 observée dans des expériences d'expression transitoire, il serait possible d'effectuer des expériences d'immunolocalisation avec des plantes sauvages. L'utilisation de nos anticorps dirigés contre URT1 s'est malheureusement révélée impossible car ils s'associent de manière non-spécifique à des granules cytoplasmiques dans notre contrôle négatif correspondant au mutant *urt1-1.* De plus, ces anticorps reconnaissent seulement la région N terminale de URT1 bien que la protéine entière ait été utilisée pour immuniser les lapins. Ces anticorps ne permettent donc pas de détecter le domaine nucléotidyltransférase de URT1 et ne peuvent pas permettre d'étudier les mécanismes d'adressage de URT1 au niveau des foci cytoplasmiques (Figure 5.5). L'obtention récente de plantes exprimant URT1 avec l'étiquette myc va permettre de tester la détection d'URT1 avec des anticorps anti-myc dans des expériences d'immunolocalisation.

D. Le rôle de l'uridylation des ARNm déadénylés

Il a été récemment montré chez l'homme que l'uridylation des précurseurs des microARN de la famille de *let-7* a deux rôles antagonistes selon la quantité d'uridines ajoutées en 3'. En effet, l'ajout d'une seule uridine permet la maturation du précurseur tandis que l'ajout d'oligouridine induit sa dégradation dans les sens 3'-5' par l'exonucléase Dis3L2 (Chang et al., 2013; Heo et al., 2012). Par ailleurs, la polarisation de la dégradation par l'uridylation peut être antagoniste selon la quantité d'uridines en 3' des ARNm. En effet, lors d'une délétion de l'exoribonucléase 3'-5' Dis3L2 chez *S. pombe*, les ARNm oligouridylés sont accumulés (Malecki et al., 2013). L'ajout d'oligouridine serait un signal pour la dégradation dans le sens 3'-5' par Dis3L2 tandis que l'ajout d'une à trois uridines, indépendamment de la déadénylation induit l'élimination de la coiffe et la dégradation dans le sens 5'-3' (Rissland and Norbury, 2009). L'induction de la dégradation dans le sens 5'-3' par l'ajout d'une uridine est conservée pour les ARNm des histones non polyadénylés chez l'homme et les ARNm polyadénylés chez *A. nidulans*, précédés ou non par une étape de déadénylation (Morozov et al., 2010a; Mullen and Marzluff, 2008). Chez *A. nidulans*, il existe une voie d'élimination de la coiffe indépendante de la déadénylation dans les mutants de déadénylation Δcaf1, Δccr4 (Morozov et al., 2010a). Celle-ci peut correspondre à la modification en 3' des ARNm contenant un codon stop précoce qui sont les substrats de la voie de dégradation du NMD (Morozov et al., 2012). L'ajout de bases pyrimidiques est effectué après déadénylation, lorsque les ARNm ne possèdent plus que 15 adénines sauf pour les

ARNm ayant un codon stop précoce, dégradés par la voie du NMD (Morozov et al., 2010a). Dans une seconde étude, cette équipe a montré que l'ARNm *CCR2* chez *A. thaliana* est aussi modifié en 3' par des bases pyrimidines lorsque les ARNm comportent 14 adénines (Morozov et al., 2012). La modification en 3' des ARNm oligoadénylés (15 ou 14 A) respectivement observée chez les ARNm de *A. nidulans* et l'ARNm *CCR2* chez *A. thaliana* (Morozov et al., 2010a, 2012) sont en cohérence avec nos expériences qui montrent que les ARNm sont uridylés lorsque les queues polyA sont réduites à 14-15 A (Chapitre 2). Cependant, dans l'étude de l'ARNm *CCR2*, la différence de la taille des queues polyA modifiées ou non en 3' n'est pas significative. Tandis qu'à l'inverse, nos résultats indiquent clairement que l'uridylation des ARNm intervient lorsque les queues polyA sont réduites à 14 adénines. De plus, la protection de l'uridylation des ARNm en 3' a été observée sur tous les gènes testés.

1. La polarisation de la dégradation dans le sens 5'-3' des ARNm oligoadénylés ayant une à trois uridines en 3'

a) La protection en 3' des ARNm oligoadénylés par URT1

Nos résultats montrent une accumulation des ARNm tronqués dans un mutant de délétion *urt1-1* qui n'est pas observée dans Col0. L'uridylation en 3' des ARNm par URT1 les protège en 3' d'un « rognage ». URT1 est intrinsèquement distributive ce qui permet l'accessibilité de l'extrémité 3' uridylée des ARNm à d'autres facteurs protéiques. Certains d'entre eux peuvent être impliqués dans la protection en 3' des ARNm. Les partenaires de la voie de la protection en 3' des ARNm par l'uridylation de URT1 restent à être identifiés.

L'extrémité 3' des ARNm déadénylés est protégée par le complexe Pat1 et Lsm1-7. Le complexe Lsm1-7 a une plus grande affinité pour les ARN oligoadénylés uridylés que pour les ARN oligoadénylés, du moins dans un système *in vitro* à base d'extraits protéiques de cellules humaines (Song and Kiledjian, 2007; Tharun and Parker, 2001). Le complexe Lsm1-7 pourrait protéger l'extrémité en 3' des ARNm uridylés chez Arabidopsis. Cependant, dans un mutant *lsm1* la protection en 3' est toujours présente (données obtenues dans le laboratoire). De plus, le complexe Lsm1-7 est présent dans le mutant *urt1-1* et la protection en 3' n'est pas observée. La protection des ARNm est vraiment due à l'uridylation de URT1 qui recrute une protéine ou un complexe autre que le complexe Lsm1-7.

La protéine La se lie aux snARN U6 uridylés et les protège en 3' (Xue et al., 2000). Certaines protéines ayant des motifs La, les *La-related protein* (LARP)

peuvent se lier aux ARN uridylés (Bayfield et al., 2010). LARP1, LARP6 et LARP7 semblent avoir conservés leur liaison pour les uridines mais LARP7 se lie surement aux transcrits de la polymérase III comme les snARN du noyau (Bousquet-Antonelli and Deragon, 2009). Chez *A. thaliana,* LARP6 possède trois homologues. LARP6a est localisée dans le noyau (Merret et al., 2013). LARP6b est localisée dans le cytoplasme et les granules de stress. Elle est exprimée dans tous les tissus de la plante, sauf le pollen et se lie aux polyA (Merret et al., 2013). LARP6c se lie aux ARN uridylés *in vitro* et est localisée dans le noyau et le cytoplasme et dans des RNP granules, étant très probablement des granules de stress. LARP6c pourrait être un candidat de la protection des ARNm uridylés en 3' mais elle pourrait inhiber la traduction sans être spécifique de l'extrémité 3' (Merret et al., 2013). LARP6c pourrait, comme son homologue LARP6 chez l'homme, moduler la stabilité et/ou la traduction des ARNm par liaison aux 5'UTR (Cai et al., 2010; Merret et al., 2013). Cependant, son expression est limitée au pollen (Merret et al., 2013). Donc LARP6c ne semble pas le candidat responsable de la protection des ARNm uridylés en 3'. LARP1 est la candidate la plus plausible de la protection des ARNm uridylés en 3'. Chez *C. elegans*, en absence de LARP1, le niveau standard des ARNm est augmenté. Elle peut être due à une baisse de la dégradation ou une non inhibition de la traduction des ARNm. Sachant que LARP1 est localisée au niveau des *processing bodies,* les auteurs favorisent l'hypothèse de la baisse de la dégradation en absence de LARP1 (Nykamp et al., 2008). LARP1 aurait donc plutôt un rôle dans la dégradation des ARNm. Cependant, chez la mouche, LARP1 se lie aux PABP et LARP1 pourrait avoir un rôle d'inhibition de la traduction (Burrows et al., 2010). Le rôle de LARP1 n'est pas encore publié chez les plantes. LARP1 est localisée au niveau des granules de stress. Afin de déterminer si LARP1 est responsable de la protection en 3' des ARNm, il faut quantifier les ARNm tronqués dans un mutant nul *larp1*. Pour découvrir les partenaires protégeant l'extrémité 3' des ARNm oligoadénylés uridylés d'une façon globale et non par une approche gène candidat, il serait possible d'effectuer une co-purification à partir d'un extrait protéique de plantes et d'ARNm oligoadénylés possédant une à trois uridine(s).

b) L'uridylation marque les ARNm pour l'élimination de la coiffe

L'uridylation chez *S. pombe*, induit l'élimination de la coiffe, et la dégradation dans le sens 5'-3' (Rissland and Norbury, 2009). Chez *A. nidulans*, l'ajout de CUCU est effectué après une étape optionnelle de déadénylation. Même s'il est estimé que la modification en 3' des ARNm induit l'élimination de la coiffe, chez *A. nidulans*, la réduction de la taille de la queue polyA à 15 adénines est, elle-même, suffisante pour

l'induire (Morozov et al., 2010a). Chez *A. thaliana*, la modification en 3' des ARNm marque les transcrits sans coiffe. En effet, 29% des ARNm *CCR2* modifiés en 3' n'ont plus de coiffe (Morozov et al., 2012). Donc l'uridylation ou l'ajout de CUCU en 3' des ARNm induit l'élimination de la coiffe et/ou marque les transcrits pour la voie de dégradation dans le sens 5'-3'.

Dans nos expériences, tout comme les ARNm de *CCR2*, de 60 à 98% des ARNm de *LOM1* uridylés n'ont pas de coiffe, indiquant comme chez *A. nidulans*, un marquage des ARNm sans coiffe. Cependant, le mutant de délétion *urt1-1* ne montre par d'impact sur le taux de dégradation des ARNm. De plus, nos expériences n'ont pas permis de mettre en évidence que l'uridylation des ARNm induit l'élimination de la coiffe. En effet, comme chez *A. nidulans,* chez *A. thaliana*, l'étape de déadénylation induit en elle-même l'élimination de la coiffe. Afin de relier l'uridylation et l'élimination de la coiffe, il faudrait que des ARNm polyadénylés soient uridylés. Cette étude serait très difficile car *A. thaliana* comporte plusieurs homologues des déadénylases, CCR4, Caf1 et PARN, même si la localisation cellulaire de cette dernière est controversée (Introduction partie A1a).

2. L'absence de modification de la demi-vie des ARN dans le mutant *urt1-1*

Il a été rapporté chez *A. thaliana* que 25% des ARNm de *CCR2* sont uridylés lorsque le taux de *CCR2* baisse, contre 3% lorsque la quantité de transcrits dans la cellule augmente (Morozov et al., 2012). Ce résultat laisse à penser que le taux d'uridylation de *CCR2* est corrélé à son instabilité. Cependant, dans nos expériences, la délétion de URT1 n'entraine pas de modification de la vitesse de dégradation globale des ARNm ni de phénotype développemental sévère comme ceux observés lors de la délétion de DCP1 ou DCP2 (Xu et al., 2006). Il est possible que URT1 n'ait pas une grande influence sur la demi-vie des ARNm car URT1 uridyle les transcrits déadénylés et la déadénylation seule peut induire l'élimination de la coiffe. Cependant, le mutant de l'exoribonucléase 5'-3' *xrn4-3* chez les plantes ne présente pas non plus de phénotype sévère et XRN4 est impliquée dans la voie de réponse à l'éthylène (Potuschak et al., 2006). Cette absence de phénotype sévère de la délétion de XRN4 est sûrement due à une compensation des activités ribonucléolytiques du sens 5'-3' avec celles du sens 3'-5'. L'absence de phénotype de *urt1-1* indique que l'uridylation est donc une étape facultative marquant les ARNm pour la dégradation dans le sens 5'-3', facilitant cette polarisation de la dégradation mais sans en influencer la vitesse de dégradation.

3. La présence d'oligouridine en 3' des ARNm

Dans un mutant de l'exoribonucléase du sens 5'-3' *xrn4-3*, la quantité d'uridines ajoutées en 3' s'allonge jusqu'à quatorze uridines. Cette modification de la quantité d'uridines en 3', dans le mutant *xrn4-3* comparée à l'écotype sauvage Col0, montre que les ARNm uridylés sont des substrats de XRN4. L'allongement de la quantité d'uridine en 3' peut indiquer que les oligouridines (supérieures à 3 uridines) en 3' sont un signal de la dégradation dans le sens 3'-5'. En effet, les ARN oligouridylés en 3' comme les précurseurs des microARN, les microARN chez l'homme et les ARNm chez *S. pombe* sont dégradés par la 3'-5' exoribonucléase Dis3L2 (Chang et al., 2013; Lubas et al., 2013; Malecki et al., 2013). Par ailleurs, *in vitro*, Dis3L2 a une affinité plus forte pour les ARN avec quatre uridines que ceux n'en possédant seulement deux en 3' (Malecki et al., 2013).

a) La dégradation des ARNm oligouridylés

Afin d'étudier les facteurs responsable de la dégradation des ARNm oligouridylés, il est possible de procéder à une délétion de protéine candidate en absence ou non de XRN4 et observer si les ARNm oligouridylés s'accumulent. L'exosome est l'un des facteurs principal de la dégradation des ARN dans le sens 3'-5' (Introduction partie A2). Le cytosol d'*A.thaliana* contient plusieurs exoribonucléases dans le sens 3'-5' comme la sous-unité de l'exosome RRP41, le cofacteur possible de l'exosome Rrp6L3 et SOV qui est indépendante de l'exosome (Chekanova et al., 2000; Lange et al., 2008; Zhang et al., 2010).

SOV semble être la meilleure candidate pour la dégradation dans le sens 3'-5' des ARNm oligouridylés. En effet, SOV est l'homologue de Dis3L2 qui est responsable de la dégradation des ARN oligouridylés dans le sens 3'-5' chez *S. pombe* et l'homme (Chang et al., 2013; Lubas et al., 2013; Malecki et al., 2013). Arabidopsis possède plusieurs écotypes et seul l'écotype sauvage Col0 semble posséder une version de SOV inactive. En effet, chez l'écotype Col0, SOV contiennent une arginine en position 705 chez Col0 ou une proline chez les autres écotypes d'Arabidopsis. Si SOV dégrade les ARNm oligouridylés, la présence de SOV inactive dans l'écotype Col0 permettrait ainsi l'accumulation des ARNm oligouridylés en absence de XRN4. Nous testons actuellement l'implication de SOV dans la dégradation des transcrits oligouridylés dans le laboratoire *in vitro* et *in vivo*. Pour cela plusieurs approches sont utilisées en parallèle. Les versions de SOV de Col0 et d'un autre écotype sont exprimées sous forme de protéines recombinantes et leurs activités sont testées *in vitro*. D'autre part, la version active de SOV d'un autre écotype que Col0 peut être exprimée dans les plantes Col0 *xrn4-3*, et il suffit de

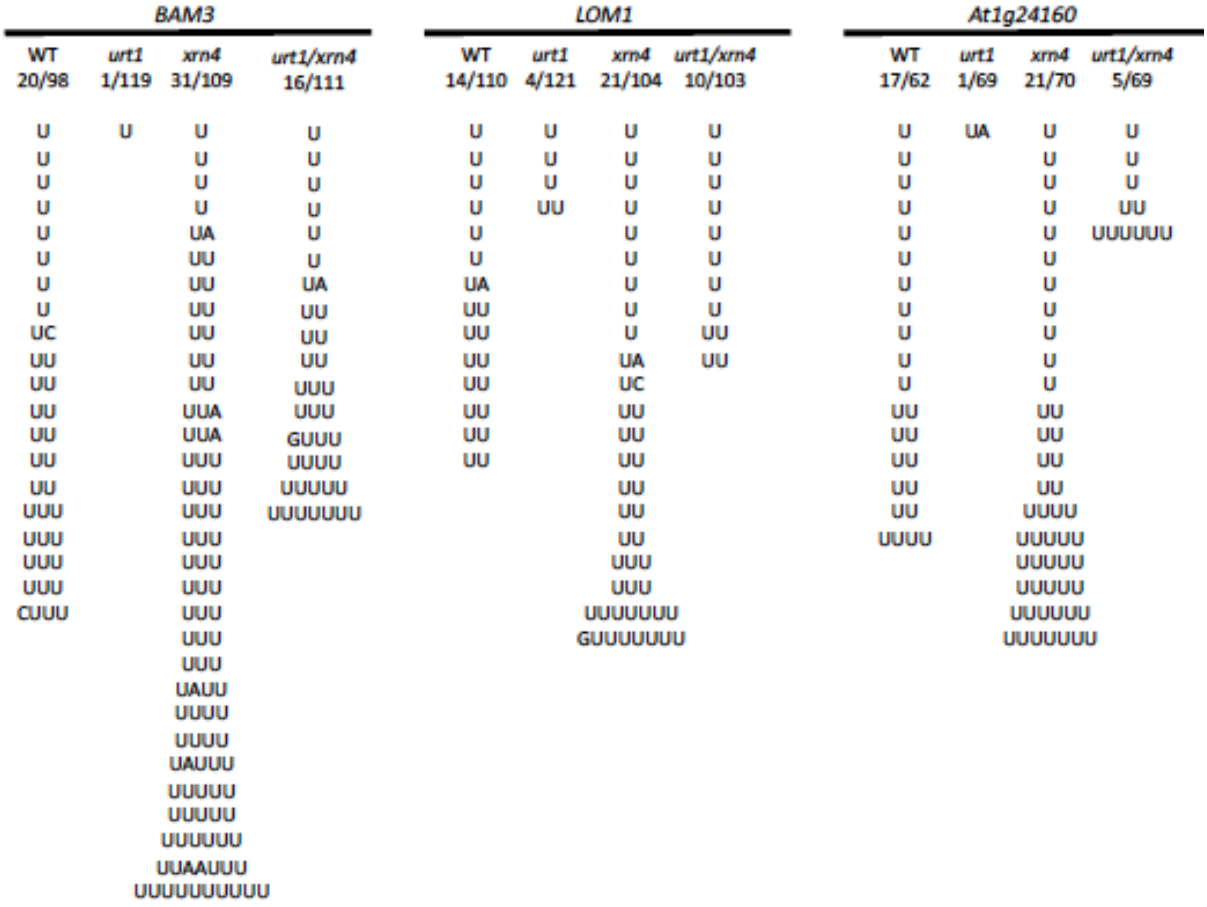


Figure D6 : Les ARNm du double mutant *urt1-1/xrn4-3* sont uridylés par une autre uridylyltransférase que URT1

La composition des modifications en 3' est présentée pour les ARNm *BAM3*, *LOM1* et *At1g24160*, issus de deux répliquas biologiques. Le rapport du nombre de séquence uridylés sur le nombre de clones total est indiqué pour chaque génotype. Les extensions de *xrn4-3* sont significativement plus grande que celles provenant de WT (test de Mann-Whitney $p<0,01$)

vérifier si l'accumulation des ARNm oligouridylés est toujours présente. Il est aussi possible de muter *xrn4-3* dans un autre écotype d'Arabidopsis que Col0 et de comparer l'accumulation des ARNm oligouridylés en absence de XRN4 avec celle observée dans l'écotype Col0.

Chez *A. thaliana*, la sous-unité de l'exosome RRP41 possède tous les résidus catalytiques nécessaires à la phosphorolyse et pourrait avoir une activité exoribonucléasique. Sachant que l'exosome est essentiel à la viabilité des plantes, il est possible d'effectuer des mutations au niveau des résidus catalytiques ou des mutants de *silencing* sous un promoteur inductible de *RRP41* afin de vérifier l'implication de RRP41 dans la dégradation des ARNm uridylés. *A. thaliana* possède trois homologues au cofacteur RRP6, RRP6-*like*, qui ont tous une localisation différente. Sachant que URT1 est localisée au niveau du cytoplasme, RRP6L3 est le candidat favorisé de la dégradation des ARNm déadénylés uridylés car il est le seul des homologues de RRP6 à être localisé dans le cytoplasme. Afin d'étudier cette hypothèse, il est possible d'utiliser un mutant nul *rrp6l3* et d'observer l'accumulation des ARNm uridylés.

b) Une seconde uridylyltransférase peut uridyler les ARNm

1. La reconnaissance des substrats ARN de la seconde uridylyltransférase uridylant les ARNm

Dans le mutant *urt1-1*, il existe une uridylation des ARNm très résiduelle. Il existe donc une autre uridylyltransférase capable d'uridyler les ARNm déadénylés. Cette uridylation intervient lorsque les ARNm ne possèdent plus que 3 à 4 adénines en 3'. Dans le mutant *xrn4-3*, les ARNm ayant un nombre d'uridines en 3' le plus important (appelés oligouridines) possèdent une quantité d'adénines en 3' faible. Sachant que l'ajout d'oligouridines en 3' des ARNm est retrouvée dans le mutant *urt1-1/xrn4-3* (Figure D6), il n'est pas effectué par URT1. De plus, dans ce mutant, l'uridylation des ARNm intervient lorsque les ARNm n'ont plus qu'en moyenne 4 adénines en 3' (31 séquences), semblable à l'uridylation résiduelle chez *urt1-1* (Figure D6). Le mécanisme de reconnaissance des substrats ARN de cette deuxième uridylyltransférase capable d'uridyler les ARNm déadénylés est donc différent de URT1 qui uridyle les ARNm ayant environ 14 adénines en 3'.

D'une façon intéressante, la quantité d'uridines en 3' des ARNm dans le double mutant *urt1-1*/*xrn4-3* (31 séquences) tend à être significativement réduite comparée au simple mutant *xrn4-3* ($p<0,05$) (73 séquences) mais elle n'est pas supérieure à Col0 ($p>0,05$) (51 séquences). Les ARNm oligouridylés ne sont pas ajoutées par URT1. Cependant, l'allongement global de la quantité d'uridines en 3'

Figure D7 : Arbre phylogénique des ncPAP d'*A. thaliana*

L'alignement des ncPAP est effectué par MUSCLE et la construction de l'arbre par ClustalW2 (*neighbour joining* algorithme) phylogeny (http://www.ebi.ac.uk/Tools/msa).

des ARNm chez *xrn4-3* pourrait être due à une synergie d'une autre uridylyltransférase et de URT1. Afin de vérifier si les deux uridylyltransférases sont nécessaires à l'allongement du nombre d'uridines en 3', il faudrait comparer la quantité des uridines ajoutées aux ARNm en absence de XRN4 et de la deuxième uridylyltransférase.

Donc l'ajout d'oligouridines aux ARNm ayant peu d'adénines en 3' est effectué par la seconde uridylyltransférase capable d'uridyler les ARNm. D'autre part, dans le mutant *xrn4-3*, URT1 uridyle les ARNm en premier, probablement *via* un partenaire spécifique des ARNm n'ayant plus que 14 adénines ou par l'absence d'inhibition par la PAB2. Ensuite, la quantité d'uridines en 3' peut être augmentée par URT1 et/ou une autre uridylyltransférase, lorsque la dégradation dans le sens 5'-3' est affectée.

2. Les candidats

Il existe quinze ncPAP chez *A. thaliana*. Actuellement, seules l'orthologue cytoplasmique de MUT68 HESO1, la CCA nucléotidyltransférase, une polyA polymérase et URT1 sont caractérisées (Kwak and Wickens, 2007; Ren et al., 2012; Sement et al., 2013; Zhao et al., 2012; Zimmer et al., 2009). Après une étude phylogénique, URT1, HESO1 et les protéines codées par les locus At3g45750 et At3g45760 forment un groupe de quatre protéines (Figure D7). HESO1, et les protéines codées par les locus At3g45750 et At3g45760 sont donc les trois meilleurs candidats pour être la seconde uridylyltransférase détectée dans nos études. Cependant, la protéine codée par le locus At3g45750 a été précédemment localisée dans le laboratoire dans le noyau après une expression transitoire de la protéine dans des cellules BY2 (Thèse de François Sement). La protéine codée par le locus At3g45760, dont la localisation est encore inconnue, pourrait être un bon candidat pour être la seconde uridylyltransférase responsable de l'uridylation de ARNm. En effet, elle conserve l'aspartate et l'arginine nécessaires à la spécificité des nc PAP pour les UTP (Alignement Annexe D1). Cependant, la plupart des ncPAP contiennent un équivalent de ces deux acides aminés et seule URT1 contient un équivalent de l'histidine de Cid1 et XTUT7, responsable de la discrimination de l'UTP à l'ATP (Alignement Annexe D1) (Lapointe and Wickens, 2013; Lunde et al., 2012; Munoz-Tello et al., 2012; Yates et al., 2012). Seules les protéines codées par les locus At5g53770 et At3g56320 ne contiennent pas d'équivalent à l'arginine nécessaire à la spécificité nucléotidique pour les UTP. Toutes les autres ncPAP pourraient être potentiellement des uridylyltransférases. Une recherche de l'uridylyltransférase impliquée dans l'uridylation des ARNm est débutée par l'approche du candidat HESO1. En effet, HESO1 est le candidat le plus

probablement responsable de l'uridylation des ARNm. HESO1 est une uridylyltransférase localisée dans des granules cytoplasmiques et elle co-localise avec URT1 (Thèse de François Sement) (Ren et al., 2012; Zhao et al., 2012). Un mutant nul *heso1-1*/*urt1-1* est en cours d'étude dans le laboratoire.

4. Le rôle de l'uridylation des ARNm déadénylés associés aux polysomes

Dans les travaux effectués chez *S. cerevisiae*, il a été émis l'hypothèse que la polarité de dégradation de 5' en 3' aurait évolué afin d'empêcher la traduction d'ARNm tronqués en cas de dégradation co-traductionnelle (Hu et al., 2009). La même équipe a découvert que les ARN intermédiaires de la dégradation 5'-3' du NMD sont associés aux polysomes (Hu et al., 2010). Chez *A. nidulans*, dans le cas de la dégradation par la voie du NMD des ARNm ayant un codon stop précoce, leur modification CUCU en 3' permettrait leur dissociation des ribosomes (Morozov et al., 2012). Cette hypothèse est renforcée par l'observation d'une baisse de la traduction des ARNm uridylés par XTUT7 (Lapointe and Wickens, 2013). Il y a donc deux rôles hypothétiques et non mutuellement exclusifs de l'uridylation des ARNm. Les ARNm uridylés peuvent être dégradés co-traductionnellement dans le sens 5'-3' et l'uridylation peut être un signal d'inhibition de la traduction par une dissociation des ARNm des ribosomes. Ces deux rôles peuvent être transposables dans le cas de l'uridylation par URT1.

D'après nos résultats, les ARNm déadénylés uridylés en 3' par URT1 peuvent être détectés au niveau des polysomes. L'uridylation par URT1, protège les ARNm en 3' et polarise la dégradation dans le sens 5'-3' et pourrait être donc importante lors de la dégradation co-traductionnelle. Cependant, la capacité de traduction de ces ARNm uridylés associés aux polysomes n'est pas encore établie. Dans le mutant *urt1-1* une uridylation secondaire des ARNm est présente. Elle semble être plus importante lorsque les ARNm déadénylés sont associés au niveau des polysomes, que lors de l'étude des ARNm totaux. D'après nos résultats, le gène BAM3 dans le mutant *urt1-1* sont respectivement uridylés à 11% (19/178) au niveau des polysomes contre 1% (1/101) au niveau des ARNm totaux. Bien évidemment, il faut prendre ces résultats avec quelques précautions de rigueur car seulement deux gènes ont été étudiés et l'étape de PCR n'est pas quantitative car elle apporte un biais en amplifiant de façon préférentielle les fragments les plus courts. De plus, la faible variabilité des séquences uridylés pourrait être due à une amplification des mêmes fragments ARNm (Chapitre 2).

La seconde uridylation observée chez le mutant *urt1-1* intervient lorsque les ARNm ne possèdent plus que trois ou quatre adénines en moyenne (pour 20 séquences). L'uridylation des ARNm n'est jamais observée dans le corps de l'ARNm. La deuxième uridylyltransférase uridylant les ARNm pourrait aussi empêcher la traduction d'ARNm tronqués dans la CDS, polarisant la dégradation co-traductionnelle dans le sens 5'-3'. Cependant, il est possible que les ARNm tronqués uridylés ne soient pas détectés car la dégradation du corps de l'ARN est vraisemblablement très processive tandis que la déadénylation est plus lente.

Afin de tester l'importance de l'uridylation au niveau des polysomes et peut-être confirmer ces observations, il faudrait traiter en parallèle des échantillons avec ou non une séparation préalable des ARNm associés aux polysomes des ARNm non-associés et utiliser une méthode d'analyse globale de l'uridylation. Cette méthode est actuellement mise en place au laboratoire. La polarité de la dégradation co-traductionnelle dans le sens 5'-3' peut être une l'évolution de la cellule afin d'empêcher la traduction d'ARNm tronquées. Cependant, la traduction des ARNm uridylés associés au niveau des polysomes n'est pas mise en évidence. Pour cela, il serait intéressant de relier l'étude globale de l'uridylation des ARNm au translatome, c'est à dire aux ARNm étant associés aux polysomes (Juntawong and Bailey-Serres, 2012; Mustroph et al., 2009; Zanetti et al., 2005).

La dégradation dans le sens 3'-5' associée aux polysomes ?

Lorsque la dégradation 5'-3' est réduite comme chez le mutant *xrn4-3*, la quantité d'uridines ajoutée en 3' des ARNm après déadénylation augmente. Cette modification du nombre d'uridines en 3' des ARNm pourrait être un signal pour la dégradation dans le sens 3'-5' et ce, même si les ARNm sont associés au niveau des polysomes (Chapitre 2). L'oligouridylation pourrait être responsable de la dégradation 3'-5' des ARNm associés aux polysomes. En effet, l'ajout d'oligouridines a récemment été impliqué dans la dégradation dans le sens 3'-5' par Dis3L2 qui est localisée au niveau des monosomes et des polysomes chez l'homme (Chang et al., 2013; Lubas et al., 2013; Malecki et al., 2013).

L'uridylation pourrait donc être un signal de dégradation dans le sens 5'-3' mais aussi 3'-5'. Pour détecter des transcrits tronqués dans la CDS de l'ARNm et associés aux polysomes, il faudrait employer une méthode d'analyse globale avec une très grande profondeur de séquençage. Il est en effet possible que ces transcrits tronqués soient présents à de très faibles niveaux car la dégradation dans le corps de l'ARNm est vraisemblablement plus rapide que la déadénylation. Il est aussi possible que la dégradation 3'-5' ne soit pas présente au niveau des polysomes. En

effet, la dégradation dans le sens 3'-5' associée aux polysomes n'a pas de sens logique car elle risque de produire des ARNm tronqués et de bloquer les ribosomes sur les fragments d'ARNm, ayant un effet toxique pour la cellule. Donc l'hypothèse de la dégradation co-traductionnelle 3'-5' n'est pas préférée. Comme décrit précédemment, l'uridylation pourrait aussi être un signal d'arrêt de la traduction éventuellement par liaison d'une protéine, protégeant les extrémités 3' uridylés qui pourrait inhiber l'initiation de la traduction (Burrows et al., 2010; Lapointe and Wickens, 2013; Morozov et al., 2012; Nykamp et al., 2008).

5. Le rôle de URT1 au niveau des foci cytoplasmiques tels que les *processing bodies* et les granules de stress

La protéine URT1 est localisée au niveau des *processing bodies* et des granules de stress après son expression transitoire dans des feuilles de tabac et des cellules BY-2 (Chapitre 2 partie A et Chapitre 3). Les *processing bodies* et les granules de stress contiennent un groupe de protéines communes étant des facteurs d'inhibition de la traduction. Il est possible que l'uridylation des ARNm par URT1 soit un signal d'inhibition de la traduction. En effet, même si l'uridylation des ARNm est présente au niveau des polysomes chez Arabidopsis, aucune donnée n'est connue au niveau de la traduction de ces ARNm.

a) Les *processing bodies*

Le rôle putatif des *processing bodies* est principalement la dégradation des ARNm non traduits car ils contiennent des facteurs impliqués dans la dégradation dans le sens 5'-3' comme le complexe Lsm1-7 qui se lie aux ARNm déadénylés (uridylés ou non), le complexe d'élimination de la coiffe et l'exoribonucléase 5'-3' (Buchan and Parker, 2009; Thomas et al., 2011; Weber et al., 2008). Chez *A. nidulans*, l'activateur de l'élimination de la coiffe Edc3 et la protéine cutA, responsable de l'ajout de CUCU en 3' des ARNm, sont nécessaires à la formation des *processing bodies*, et ils serviraient à l'optimisation de la dégradation des transcrits (Morozov et al., 2010b). De plus, les ARNm uridylés sont des substrats de XRN4 qui est, elle-même, localisée dans les *processing bodies* (Weber et al., 2008). La localisation de URT1 au niveau des *processing bodies* est donc tout à fait cohérente avec ces précédentes observations. L'uridylation des ARNm par URT1 intervient lorsque les ARN ont une queue polyA réduite à 14 adénines, les protégeant en 3' et polarisant la dégradation dans le sens 5'-3'. Cependant, l'uridylation n'est pas essentielle à la dégradation des ARNm et elle pourrait être aussi un signal d'inhibition de la traduction.

Il semble que la localisation de URT1 au niveau des *processing bodies* soit importante en cas de stress car le nombre de *processing bodies* augmente en cas de stress. De plus, mes résultats indiquent que le domaine nucléotidyltransférase devient suffisant à la localisation de URT1 dans les *processing bodies*. Le mécanisme d'adressage de URT1 dans les *processing bodies* pourrait donc être différent en cas de stress.

b) Les granules de stress

Les granules de stress sont a priori impliqués dans le stockage et le tri des ARNm non traduits dirigés ensuite vers la dégradation ou redirigés vers la traduction (Anderson and Kedersha, 2008; Buchan and Parker, 2009; Thomas et al., 2011; Weber et al., 2008). Le rôle de URT1 au niveau des granules de stress n'est pas encore compris, mais l'uridylation pourrait avoir un rôle dans l'inhibition de la traduction. L'uridylation peut aussi être responsable du marquage des ARNm destinés à être dégradés dans le sens 5'-3' qui seraient redirigés vers les *processing bodies* par exemple. D'autre part, l'uridylation des ARNm par URT1 peut être un mécanisme de réponse rapide à un stress qui pourrait être réversible, mais aucune donnée n'est connue à ce jour sur ce sujet. Il est possible que les ARNm uridylés retournent à la traduction, après élimination ou non des uridines en 3'. Les mécanismes d'uridylation sont en cours d'étude dans le laboratoire. En effet, l'implication de l'uridylation par URT1 en réponse au stress est actuellement étudiée mais le mutant nul *urt1-1* ne possède pas de phénotype particulier tandis que le double mutant *xnr4-3/urt1-1* est très sensible à des stress biotique et abiotique. En effet, il résiste moins aux infections biotiques (*Pseudomonas syringae*) et sa germination est ralentie en conditions salines. Lorsque la méthode d'analyse globale de l'uridylation du transcriptome sera mise en place dans le laboratoire, elle pourra permettre de mieux comprendre le mécanisme de l'uridylation des ARNm par URT1 lors de différents stress.

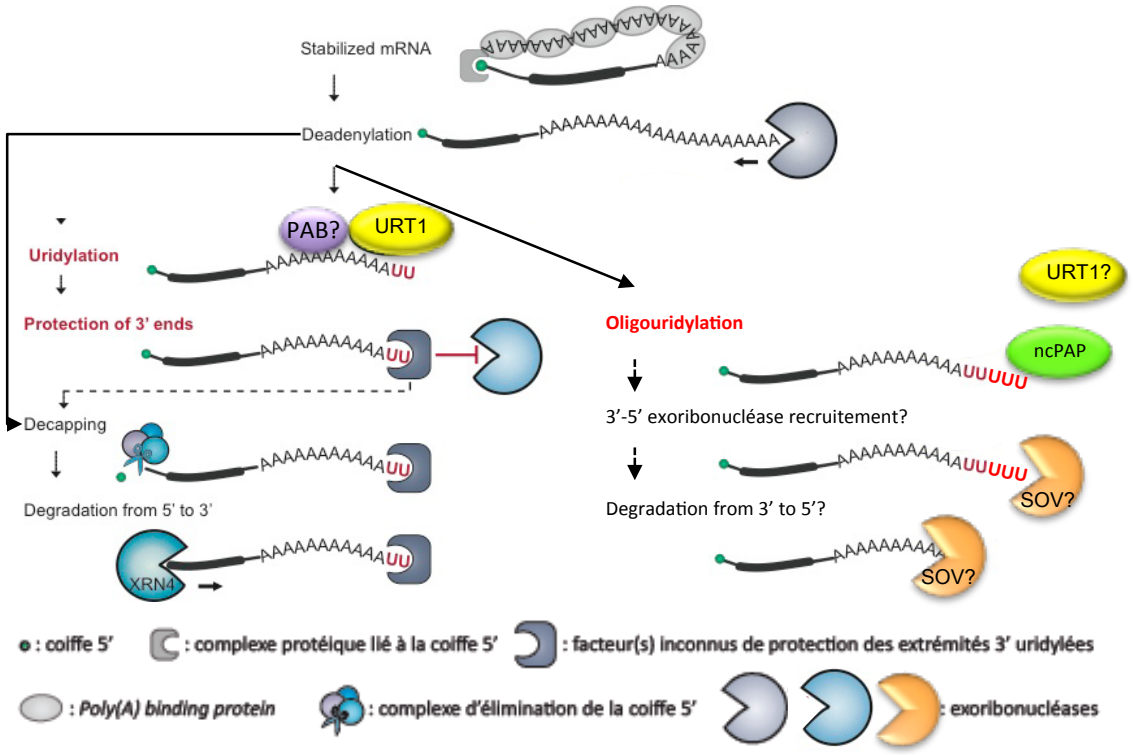


Figure D8 : Le mécanisme d'uridylation des ARNm oligoadénylés par URT1 chez *A. thaliana*

Les étapes non prouvées sont représentées par des flèches en pointillée, les facteurs candidats potentiels intervenant sont suivis d'un point d'interrogation

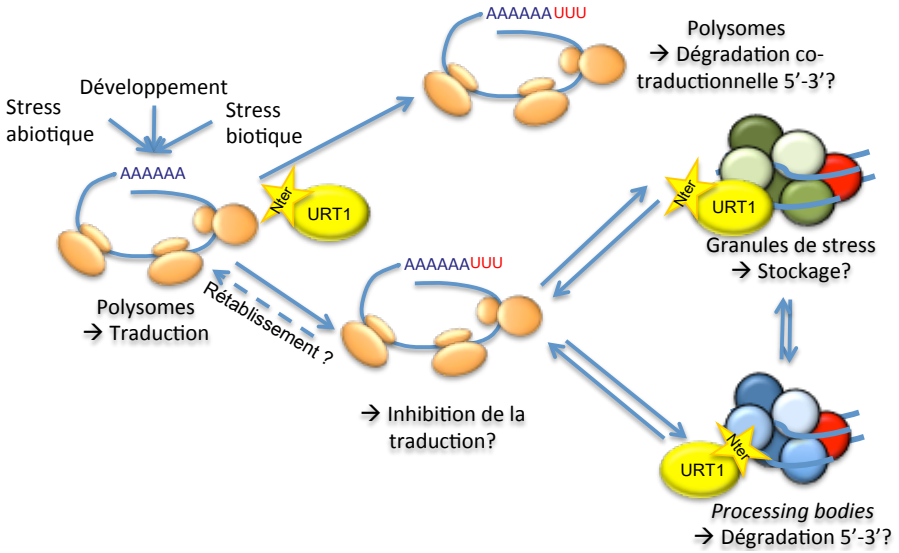


Figure D9 : Les rôles hypothétiques de l'uridylation des ARNm déadénylés par URT1 chez *A. thaliana* dans « le cycle de l'ARNm »

Les ARNm associés aux polysomes peuvent être uridylés. L'uridylation a plusieurs rôles hypothétiques : le marquage des ARNm pour la dégradation 5'-3' co-traductionnelle ou non, ou une inhibition de la traduction. Celle-ci permettrait de stocker les ARNm non transcrits au niveau des foci cytoplasmiques. La région N terminale de URT1 est essentielle à la localisation de URT1 au niveau des *processing bodies* mais pas pour sa localisation au niveau des granules de stress. Il est possible que les ARNm puissent être à nouveau traduits au niveau des polysomes soit après élimination des uridines en 3' ou non. Donc l'uridylation pourrait être un processus rapide et réversible de réponse au stress.

Conclusion générale

Ce travail de thèse a permis de découvrir une nouvelle uridylyltransférase, intrinsèquement spécifique des UTP : URT1. *In vivo,* URT1 est essentielle à l'uridylation des ARNm oligoadénylés bien que URT1 ne présente pas de spécificité intrinsèque pour les ARNm déadénylés *in vitro*. La spécificité de substrat ARN de URT1 est surement apportée par un partenaire protéique qui pourrait se lier à URT1 au niveau des motifs conservés chez les Angiospermes.

L'ajout d'une à trois uridines par URT1 aux ARNm déadénylés protège leur extrémité 3', polarisant leur dégradation dans le sens 5'-3'. Cette polarité est importante en cas de dégradation co-traductionnelle car l'uridylation des ARNm et la protection en 3' qui en découle sont retrouvées au niveau d'ARNm engagés dans les polysomes. Cependant, la dégradation des ARNm n'est pas modifiée en absence de URT1, l'uridylation n'est donc pas une étape requise à la dégradation des ARNm. Lors d'une baisse de la dégradation dans le sens 5'-3', le nombre d'uridines en 3' des ARNm déadénylés augmente, l'ajout d'oligouridines en 3' s'effectue par une autre uridylyltransférase capable d'uridyler les ARNm. Ces ARNm oligouridylés deviendraient des substrats de la voie de dégradation dans le sens 3'-5', éventuellement dégradés par SOV (Figure D8).

URT1 est localisée au niveau des granules de stress et des *processing bodies* après son expression transitoire dans des cellules de tabac BY-2 et des feuilles de tabac. J'ai pu mettre en évidence le mécanisme d'adressage de URT1 au niveau de ces foci cytoplasmiques. Les acides aminés 161 à 418 de la région N terminale de URT1, prédite comme intrinsèquement désorganisée, est essentielle pour sa localisation au niveau des *processing bodies.* Mais elle n'est pas essentielle pour sa localisation au niveau des *processing bodies* et des granules de stress après un stress thermique. La présence de URT1 dans les *processing bodies* est cohérente avec le fait que les ARNm uridylés sont des substrats de XRN4 (Sement et al. 2013). En revanche, le rôle de URT1 au niveau des granules de stress n'est pas encore résolu. L'uridylation des ARNm par URT1 pourrait être impliquée dans l'inhibition de la traduction des ARNm avant le tri des ARNm vers un retour à la traduction au niveau des polysomes ou à la dégradation 5'-3' dans les *processing bodies*. L'uridylation pourrait être éventuellement réversible ou encore, les ARNm uridylés pourrait retourner aux polysomes après polyadénylation avec ou sans étape

d’élimination des uridines afin de répondre très rapidement à un stimulus externe (Figure D9).

Dans l’ensemble, mes travaux de thèse ont participé à mieux comprendre les mécanismes de l’uridylation des ARNm chez *Arabidopsis thaliana* et ouvrent des perspectives pour étudier les rôles potentiels de l’uridylation dans la dégradation et/ou l’inhibition de traduction des ARNm.

Matériels et méthodes

A - Matériels

1. Matériel végétal

a) *Arabidopsis thaliana*

Les plantes utilisées lors des expériences, sauvages ou mutantes, sont des plantes d'*Arabidopsis thaliana* de l'écotype Columbia-0, appelé Col0. Les plantes ont été cultivées deux semaines en condition de jours courts avec une alternance de 8 heures de luminosité et de 16 heures d'obscurité. Elles sont ensuite cultivées à 22°C en condition de jours longs avec une alternance de 16 heures de luminosité et de 8 heures d'obscurité.

b) *Nicotiana benthamiana*

Les plantes de *Nicotiana benthamiana* utilisées pour l'expression transitoire de protéines par agro-infiltrations des feuilles ont cinq à six semaines. Elles sont cultivées à 22°C en condition de jours longs avec une alternance de 16 heures de luminosité et de 8 heures d'obscurité. Les mêmes conditions de culture sont utilisées après l'agro-transformation des feuilles durant 2 jours.

2. Souches bactériennes

a) Top10

La souche bactérienne Top10 est une souche d'*E. coli* (Invitrogen®) utilisée pour propager les plasmides Gateway®. TOP10 présentent une mutation pour endA1, inactivant l'activité endonucléase intracellulaire, et pour recA éliminant les recombinaisons homologues. Son génotype est F- mcrA (mrr-hsdRMS-mcrBC) 80lacZM15 lacX74 recA1 ara139 (ara-leu)7697 galU galK rpsL (Str^R) endA1 nupG.

b) Survival™

La souche bactérienne *ccdB* Survival™ issues de la souche d'*Escherichia coli* Top10 (Invitrogen®) est utilisée pour propager les plasmides de la technologie Gateway®, non recombinés, possédant le gène de la gyrase *ccdB*. Son génotype est F-mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araΔ139 Δ(ara-leu)7697 galU galK rpsL (StrR) endA1 nupG fhuA::IS2.

c) BL21 DE3

La souche bactérienne d'*E. coli* BL21 DE3 (NEB®) contient le gène codant pour l'ARN polymérase T7 sous le contrôle du promoteur lacUV5. Son génotype est fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5.

d) Agrobacterium tumefaciens GV3101::pMP90

Agrobacterium tumefaciens GV3101 ::pMP90 est la souche bactérienne issue de C58 et résistante à la rifampicine (50ng/ µL). Elle est utilisée pour transformer de façon transitoire les feuilles de *N. benthamiana*. *A. tumefasciens* GV3101 possède un plasmide Ti (*Tumor-inducing*) pMP90 (pTiC58 Δ T-DNA) qui contient un gène de résistance à la gentamicine (30ng/ µ L). Le plasmide Ti est nécessaire à la transformation des cellules végétales. *A. tumefaciens* se développe de façon optimale à 28°C.

3. Plasmides

a) Plasmides donneurs

Les plasmides donneurs utilisés sont tous issus de la technologie Gateway®. Ces plasmides donneurs possèdent le gène *ccdB* codant pour un inhibiteur de la gyrase, empêchant des souches sensibles d'*E. coli* comme les Top10 de se multiplier si la recombinaison n'a pas eu lieu. En effet, lors de la recombinaison du gène d'intérêt, le gène *ccdB* du plasmide est excisé. Les plasmides donneurs possèdent les sites de recombinaison attP1 et attP2 nécessaires à la recombinaison d'un fragment d'ADN comportant des sites de recombinaison attB1 et attB2. La recombinaison effectuée par une BP Clonase permet la création de nouveaux sites de recombinaison, attL1 et attL2, nécessaires à l'insertion d'un fragment d'ADN d'intérêt dans un plasmide d'expression.

Nom du plasmide	Fournisseur	Antibiotique de sélection
pDONR™207	Invitrogen	Gentamycine
pDONR™Zéo	Invitrogen	Zéocine

b) Plasmides d'expression

Les plasmides utilisés sont tous issus de la technologie Gateway®. Ces plasmides d'expression possèdent le gène *ccdB* codant pour un inhibiteur de la gyrase. Ce gène empêche des souches sensibles d'*E. coli* comme les Top10 de se multiplier. En effet, lors de la recombinaison d'un fragment d'ADN, le gène *ccdB* du plasmide est excisé. Les plasmides d'expression possèdent les sites de recombinaison attR1 et attR2 nécessaires à la recombinaison d'un fragment d'ADN d'intérêt précédemment inséré dans un plasmide d'entrée. Les plasmides d'expression et leurs étiquettes sont référencés dans le tableau suivant.

Nom du plasmide	Etiquette	Localisation	Promo-teur	Terminateur	Antibiotique de sélection	Particularité
pHGWA	6his	N terminale	Phage T7	Phage T7	100 ng/ µ L Ampicilline	Production de protéines recombinantes
pHGGWA	6his-GST	N terminale	Phage T7	Phage T7	100 ng/ µ L Ampicilline	
pHMGWA	6his-MBP	N terminale	Phage T7	Phage T7	100 ng/ µ L Ampicilline	
pEarleyGate 104	YFP	N terminale	35S CaMV	OCS	50 ng/ µ L Kanamycine	Expression de protéines *in planta*
pEarleyGate 101	YFP-HA	C terminale	35S CaMV	OCS	50 ng/ µ L Kanamycine	
pH7RWG2	RFP	C terminale	35S CaMV	35S CaMV	100 ng/ µ L Spectinomycine	
pB7WGC2	CFP	N terminale	35S CaMV	35S CaMV	100 ng/ µ L Spectinomycine	
pB7FWG2	GFP	C terminale	35S CaMV	35S CaMV	100 ng/ µ L Spectinomycine	
pER8					30 ng/ µ L Gentamycine	Vecteur vide
pGWB617	4 c-myc	N terminale	35S CaMV	Nos	100 ng/ µ L Spectinomycine	Expression de protéines *in planta*
pGWB618	3HA	N terminale	35S CaMV	Nos	100 ng/ µ L Spectinomycine	

Les plasmides d'expression contiennent en 5' des sites de recombinaison un promoteur 35S du virus de la mosaïque du chou fleur (35S CaMV) ou du Phage T7. Ils possèdent un terminateur, soit de la nopaline synthase (Nos), soit de l'octopine synthase gène (OCS) soit du Phage T7. Les plasmides de la série pGWA possèdent un promoteur inductible au lactose. Les plasmides d'expression possèdent différentes étiquettes localisées en 5' ou en 3' de la séquence du gène d'intérêt, respectivement pour une fusion de l'étiquette en N ou en C terminal de la protéine d'intérêt. Les plasmides pGWB ont un gène de résistance au basta exprimé lors de la transformation stable de plante.

4. Anticorps

a) Anticorps primaires

Les anticorps primaires sont dirigés directement contre une protéine ou l'étiquette fusionnée à la protéine d'intérêt. Ils sont incubés une heure à des concentrations spécifiques à chacun des anticorps primaires dilué dans une solution de PBS-Tween 20 (PBS1x, 0,2% (v/v) Tween 20) et 2% (p/v) de lait.

1. Anticorps anti-GFP

L'anticorps dirigé contre la GFP (1/5000e) est directement couplé à la HRP (Invitrogen®) et ne nécessite pas l'utilisation d'anticorps secondaires.

2. Anticorps anti-c-myc

Des anticorps primaires monoclonaux commerciaux sont utilisés pour mettre en évidence les étiquettes c-myc fusionnée à la protéine d'intérêt, l'anticorps est dilué au 1/3000e (Invitrogen®).

3. Anticorps anti-HA

Des anticorps primaires monoclonaux commerciaux sont utilisés pour mettre en évidence l'étiquette HA fusionnée à la protéine d'intérêt, l'anticorps est dilué au 1/5000e (Roche®).

4. Anticorps anti-URT1

Les anticorps primaires dirigés contre la protéine URT1 sont utilisés au 1/5000e. Ils ont été obtenus par l'immunisation de lapins effectuée par l'injection de 600mg de protéine recombinante URT1, avec une étiquette six histidines fusionnée en N terminal, purifiée par une chromatographie d'affinité grâce à la colonne HisTrap™ (GE).

b) Anticorps secondaires

Les anticorps secondaires sont des anticorps de chèvre dirigés contre les immunoglobulines de lapin ou de souris couplés à la péroxydase. Les anticorps secondaires sont dirigés contre la souris pour les anticorps monoclonaux anti-c-myc ou contre le lapin pour les anticorps polyclonaux tels les anticorps anti-URT1 et anti-HA. Ils sont incubés 30 minutes dilués au 1/10 000e dans une solution de PBS-Tween 20 (PBS1x, 0,2% (v/v) Tween 20) et 2% (p/v) de lait.

5. Les amorces

Les amorces sont issues des fournisseurs Sigma® et EUROFINS®. Les différentes amorces utilisées et leurs utilisations sont référencées dans le tableau suivant.

Gène	Nom de l'amorce	Séquence	Utilisation
at2g45620 (URT1)	EF306-TUT1A	ATGCACGGAAGTCCTTCGCCGTCG	amplification
	EF306-TUT1B	ATGACCGTCAATAATCCTTTTCC	amplification
	EF306-TUT1E	ATGTCGGGAGATGCTACTCAAAGCC	amplification
	EF306-TUT1Aattb	ggggacaagtttgtacaaaaaagcaggctATGCACGGAAGTCCTTCGCCGTCG	clonage fusion C terminal
	EF306-TUT1Battb	ggggacaagtttgtacaaaaaagcaggctATGACCGTCAATAATCCTTTTCC	clonage fusion C terminal
	EF306-TUT1Eattb	ggggacaagtttgtacaaaaaagcaggctATGACCGTCAATAATCCTTTTCC	clonage fusion C terminal
	306G_TUT1.1337fw	caccATGGTTAAAATGTATATGGCATGT	clonage fusion C terminal
	306G_TUT1.1337attB1fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGGTTAAAATGTATATGGCATGT	clonage fusion C terminal
	306_ncPAP4.2349rv	GTTGTGGCCTTGTCCATTATTAT	clonage fusion C terminal
	nPAP4.2349attB2rv	GGGGACCACTTTGTACAAGAAAGCTGGGTAGTTGTGGCCTTGTCCATTATTAT	clonage fusion C terminal
	EF306-TUT1AattbN	ggggacaagtttgtacaaaaaagcaggctTGCACGGAAGTCCTTCGCCGTCG	clonage fusion N terminal
	EF306-TUT1BattbN	ggggacaagtttgtacaaaaaagcaggctTGACCGTCAATAATCCTTTTCC	clonage fusion N terminal
	EF306-TUT1EattbN	ggggacaagtttgtacaaaaaagcaggctTGTCGGGAGATGCTACTCAAAGCC	clonage fusion N terminal
	TUT1.stopincrv	TCAGTTGTGGCCTTGTCC	clonage fusion N terminal
	TUT1.stopincattB2	GGGGACCACTTTGTACAAGAAAGCTGGGTTCAGTTGTGGCCTTGTCC	clonage fusion N terminal
	TUT1.1338attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTGGTTAAAATGTATATGGCATGT	clonage fusion N terminal
	TUT1**mut.fw	AAGAGCGCCATCGCCGTT	mutagénèse dirigée
	TUT1**mut.rv	AACGGCGATGGCGCTCTT	mutagénèse dirigée

6. Logiciels de bioinformatiques utilisés pour l'analyse de la séquence de URT1

Les logiciels bioinformatiques utilisés pour l'analyse de la séquence de URT1 et/ou de ses homologues sont présentés dans le tableau suivant.

Rôle	Nom	Site Web	Référence
Détection de domaines connus	InterProScan	http://www.ebi.ac.uk/Tools/pfa/iprscan/	(Quevillon et al., 2005; Zdobnov and Apweiler, 2001)
Prédiction de la fonction moléculaire	GO	http://www.geneontology.org/	(Ashburner et al., 2000; Carbon et al., 2009)
Classification et prédiction de la fonction moléculaire	PANTHER HMM	http://www.pantherdb.org/	(Mi et al., 2010, 2013)
Prédiction des structures secondaires	NPS@	http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_seccons.html	(Combet et al., 2000)
	MLRC	http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_seccons.html	(Guermeur et al., 1999)
	DSC	http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_seccons.html	(King and Sternberg, 1996)
	PHD	http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_seccons.html	(Rost and Sander, 1993)
	PREDATOR	http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_seccons.html	(Frishman and Argos, 1996)
Prédiction de domaine globulaire	GlobPlot2	http://globplot.embl.de/	(Linding, 2003)
Détection des régions de faible complexité	CAST	http://annie.bii.a-star.edu.sg/annie/home.do	(Kreil and Ouzounis, 2003; Ooi et al., 2009; Promponas et al., 2000)
Analyse de l'enrichissement ou l'appauvrissement des acides aminés d'une protéine en comparaison à une référence	Composition Profiler	http://www.cprofiler.org/cgi-bin/profiler.cgi	(Vacic et al., 2007)
Prédiction de la désorganisation structurale	IUPred	http://iupred.enzim.hu	(Dosztanyi et al., 2005)
	DISOPRED	http://bioinf.cs.ucl.ac.uk/disopred/disopred.html	(Buchan et al., 2010)
Prédiction des zones de liaison protéique	ANCHOR	http://anchor.enzim.hu	(Dosztányi et al., 2009; Mészáros et al., 2009)
Méthode d'analyse statistique	SAPS	http://annie.bii.a-star.edu.sg/annie/home.do	(Brendel et al., 1992)
Identification des homologues de URT1	Phytozome	http://www.phytozome.net	(Goodstein et al., 2011)

Construction des arbres phylogénétiques en éliminant les régions divergentes	GBLOCKS et PHYML	http://www.phylogeny.fr	(Castresana, 2000; Talavera and Castresana, 2007)
Représentation des motifs conservés	Weblogo	http://weblogo.berkeley.edu/logo.cgi	(Crooks et al., 2004; Schneider and Stephens, 1990)

7. Les numéros d'accession des homologues de URT1

Les numéros d'accession sont: *Aquilegia coerulea* (Aquca_013_00489), *Arabidopsis lyrata* (483698), *Brachypodium distachyon* (A, Bradi3g22907.1 and B, Bradi3g02160), *Carica papaya* (evm.TU.supercontig_18.161), *Citrus sinensis* (orange1.1g005380m), *Cucumis sativus* (Cucsa.362940.1), *Citrus clementine* (Ciclev10023615m), *Eucalyptus grandis* (Eucgr.D02430), *Glycine max* (A1 Glyma07g09011.1, A2 Glyma03g02290.1), *Manihot esculenta* (cassava4.1_002182m), *Mimulus guttatus* (Migut.B01666), *Oryza sativa Japonica* (A, LOC_Os10g11034.1, B LOC_Os02g02970.1), *Populus trichocarpa* (A1 Potri.002G151000.1, A2 Potri.014G073300.1), *Prunus persica* (ppa002004m), *Ricinus communis* (30174.t000321), *Selaginella moellendorffii* (100054), Setaria italica (A Si034285m, B Si019265m), *Sorghum bicolor* v2.1 (A Sobic.001G261900.1, B Sobic.004G019700.1), *Thellungiella halophile* (Thhalv10001324m), *Vitis vinifera* (GSVIVT01027570001), *Zea mays* (A GRMZM2G001191_T01, B GRMZM2G065097_T01). Une séquence provient de NCBI: *Jatropha curcas* (dbj|BAJ53142.1).

B - Méthodes

Les différentes techniques de base présentées dans la publication ne sont pas expliquées dans cette partie de description des méthodes qui se veut un complément à l'article présenté.

1. Modélisation du domaine nucléotidyltransférase de URT1 à partir de la résolution atomique (4FH5) de Cid1 par SWISS-MODEL

Afin de modéliser le domaine nucléotidyltransférase de URT1, j'ai utilisé le mode automatique du serveur SWISS-MODEL. SWISS-MODEL sélectionne une structure 3D dans la base de donnée PDB par homologie de séquences, ici Cid1. Cid1 possède plusieurs structures 3D résolues au niveau atomique répertoriées dans la base de donnée, avec les différents nucléotides (UTP, CTP, GTP, 3'dATP) et avec des cofacteurs (magnésium ou calcium) (Lunde et al., 2012; Munoz-Tello et al., 2012; Yates et al., 2012). Ces différentes résolutions au niveau atomique de la structure 3D de Cid1 induisent plusieurs modélisations de structures 3D de URT1 possibles. Elles suivent toutes le mêmes processus de modélisation parallèlement afin de les comparer entre elles et de sélectionner le modèle 3D le plus fiable. Le serveur SWISS-MODEL procède tout d'abord à l'alignement des deux protéines de URT1 et de Cid1 (Hang and Miller, 1991), suivi par une étape d'amélioration du score par l'optimisation des gaps. La modélisation de la structure 3D est ensuite construite, premièrement par l'assemblement de la chaîne carbone principale, puis par l'ajout des chaînes latérales. L'énergie du modèle est par la suite optimisée au niveau des chaînes latérales et des boucles.

2. Clonage des gènes

Les gènes d'intérêt sont préalablement amplifiés par PCR et ayant des sites attB1 et attB2 ajoutés à leurs séquences sont clonés par la technologie de Gateway® qui utilise deux réactions de recombinaisons successives par la BP Clonase II et la LR Clonase II. Les réactions sont effectuées dans les conditions préconisées par Invitrogen®.

3. Mutagenèse dirigée

Afin d'introduire les mutations ponctuelles D491/493A, deux amplifications successives sont effectuées. Deux fragments du gène URT1 sont amplifiés en tête-bêche grâce aux amorces contenant les mutations. Puis les amplicons sont purifiés

afin d'éliminer les amorces non intégrées et une seconde PCR est effectuée avec ces deux fragments dilués en présence des amorces du gène URT1 permettant d'amplifier le gène en entier.

4. Purification des PCR

Les produits PCR sont purifiés avant d'être clonés dans les plasmides d'entrée grâce au kit NucleoBond® Plasmid DNA Purification (Macherey-Nagel) selon les conditions standard d'utilisation.

5. Extraction d'ADN plasmidique

Les extractions d'ADN plasmidique sont effectuées par le kit NucleoSpin Plasmid® Quikpure (Macherey-Nagel) selon les conditions standard d'utilisation.

6. Séquençage d'ADN

Les réactions de séquençage sont réalisées par le service de séquençage de l'IBMP (géré par M. Alioua) basé sur la méthode de Sanger. Les acides nucléiques sont séparés par électrophorèse capillaire (Applied Biosystems 3100).

7. Transformation bactérienne

Les transformations bactériennes d'*E. coli* ou *A. tumefaciens* sont effectuées par un choc thermique par un protocole classique. 50µL de bactéries thermocompétentes sont mises en présence de 30 à 100ng de plasmides puis incubées pendant 30 minutes dans la glace. Ce mélange va subir un choc thermique. *E. coli* thermocompétentes sont incubées 30 secondes à 42°C et *A. tumefaciens* cinq minutes à 37°C avant d'être replacé sur la glace durant deux minutes. 1mL de LB (5 g/L d'extrait de levure, 10 g/L de bacto-tryptone, 5 g/L de NaCl pH 7,2) est ajouté au mélange et incubé soit à 37°C durant une heure soit à 28°C pendant deux heures, respectivement pour *E. coli* et *A. tumefaciens*. 200L sont ensuite étalés sur boite de Pétri contenant du LB agar en présence d'antibiotique spécifique du plasmide transformé afin de sélectionner les colonies transformées.

8. Transformation d'Arabidopsis

Une pré-culture de 2 jours de *A. tumefaciens* contenant le plasmide d'expression d'intérêt est utilisée pour ensemencer une culture de 150mL de LB, avec les antibiotiques nécessaires, durant 20h à 28°C sous agitation. Les bactéries sont sédimentées par centrifugation 10min à 3000g à température ambiante. Le culot

bactérien est re-suspendu dans un milieu d'agro-infiltration (1/2 MS (MO255 avec vitamines et MES (Duchefa)), 5% (p/v) sucrose, 0,5mL/L Silvet L-77, 20µM acétosyringone). Les hampes florales, ayant deux semaines après floraison, à transformer sont incubées 30 secondes dans la solution d'agro-infiltration puis les plantes sont préalablement mises à l'obscurité pendant 48h avant d'être remises dans les conditions de cultures usuelles.

9. Production des protéines recombinantes

Une souche d'*E. coli* BL21 transformée avec le plasmide d'expression d'intérêt est tout d'abord cultivée une nuit à 37°C sur une boite de Pétri contenant du milieu LB avec 1% d'agar (p/v) et l'antibiotique spécifique du plasmide d'expression. Une colonie est utilisée pour ensemencer une pré-culture liquide de LB, toujours à 37°C, en présence d'antibiotique et sous agitation. Cette pré-culture est diluée au 1/1000^{e} pour obtenir une culture bactérienne en phase exponentielle de croissance après 16h à une absorbance à 600nm comprise entre 0,5 et 0,6. Lors d'une production classique de protéine, l'induction est effectuée 5h à 37°C en présence de 1mM d'IPTG. Si les protéines doivent être extraites solubles et en condition native l'induction est effectuée à basse température. La croissance est préalablement arrêtée par une incubation de 30 minutes à 4°C puis, l'induction de la production des protéines recombinantes est effectuée durant 2 à 4 heures à 17°C. Après l'induction, les culots bactériens sont sédimentés 15min à 3000g et congelés instantanément dans de l'azote liquide et stockés à -80°C.

10. Purification de protéines recombinantes

a) Purification par chromatographie d'affinité

Les culots protéiques sont re-suspendus dans un tampon de lyse en condition dénaturante (50 mM Tris pH8, 150mM NaCl, 8M urée, 10mM β–mercaptoéthanol, 1% Triton-X100 (v/v)), puis lysés par sonication, 2 minutes par cycles de 2 secondes à 80% par le Vibra-cell™ 75115 (Bioblock Scientific). Le lysat est ensuite centrifugé pour éliminer les débris cellulaires 30 minutes à 20 000g à 4°C.

La purification est effectuée avec la colonne His Trap Crude FF 1 mL (GE) avec un flux de 0,5mL/min et une élution par un gradient de 5-100% du tampon d'élution (50 mM Tris pH8, 150mM NaCl, 8M urée, 500mM imidazole) avec un volume final de 25mL.

Les fractions d'intérêt on été regroupées et concentrées par un filtre d'AMICON MWCO 30kDa dans un volume final de 1,5 mL puis le tampon est

échangé par la colonne HiTrap *desalting* contre 20mM Tris et 6M urée à pH 8. 600µg de protéines sont utilisées pour l'immunisation des lapins pour obtenir des anticorps anti-URT1.

b) Purification des protéines recombinantes ayant une étiquette histidine par l'agarose Ni-NTA

Un culot de culture d'*E. coli* BL21 induite de 200mL durant 2 à 4 heures est repris dans 3mL de tampon de lyse bactérienne (20mM MOPS pH 7,5 ; 250mM NaCl ; 15% glycérol (v/v) ; 0,2mM DTT ; 20mM imidazole ; 0,1% Tween 20 (v/v)) en présence d'inhibiteur de protéase (*complete*-EDTA Roche). La lyse bactérienne est effectuée par une sonication de 30 secondes effectuée deux fois avec 15 secondes de pause à 40% par le Viber-cell™ (Sonics&Materials).

Ce lysat bactérien est ensuite centrifugé 10 minutes à 16 000g afin de se débarrasser des débris cellulaires. Il est incubé 30 minutes à 4°C avec 100µL de phase de Nickel agarose (Ni-NTA) préalablement incubée dans le tampon de lyse. Par la suite, les protéines ne se liant pas sur la colonne Ni-NTA sont éliminées par 3 lavages successifs de 1mL de tampon de lavage (20mM MOPS pH 7,5; 150mM NaCl ; 15% glycérol (v/v) ; 20mM imidazole ; 0,1% Tween 20 (v/v)). L'ajout de 20mM d'imidazole, qui est un compétiteur de l'histidine pour le nickel, permet d'éliminer les éventuelles liaisons de protéines aspécifiques. La protéine recombinante possédant l'étiquette 6 histidines est ensuite éluée grâce à 500µL de tampon d'élution (20mM MOPS pH 7,5 ; 15% glycérol (v/v) ; 300mM imidazole; 0,1% Tween 20 (v/v)), après une incubation de 2 minutes afin d'augmenter la concentration de la protéine recombinante éluée.

c) Purification des protéines recombinantes ayant une étiquette GST par l'agarose GST Sépharose

L'éluât de la purification par l'agarose Ni-NTA est par la suite incubé avec 100µL de GST Sepharose 4 Fast Flow (GE Healthcare®) durant une heure à 4°C. La phase est préalablement incubée dans le tampon d'élution de la purification précédente (20mM MOPS pH 7,5 ; 15% glycérol (v/v) ; 300mM imidazole; 0,1% Tween 20 (v/v)). Comme pour la purification à l'aide de l'agarose Ni-NTA, les protéines ne possédant pas d'étiquette GST sont éliminées par 3 lavages successifs par 1mL de tampon de lavage 2 (20mM MOPS pH 7,5; 150mM NaCl ; 15% glycérol (v/v) ; 0,1% Tween 20 (v/v)). La protéine d'intérêt est ensuite éluée en présence d'un compétiteur (le glutathion). Comme lors de la purification à l'aide de l'agarose Ni-NTA, les billes sont re-suspendues avec 500µL de tampon d'élution (20mM MOPS pH 7,5; 150mM NaCl ; 15% glycérol (v/v) ; 0,1% Tween 20 (v/v), 10mM glutathion pH

8) puis les protéines recombinantes sont éluées après 2 minutes d'incubation afin d'augmenter la concentration de protéine recombinante obtenue.

Une dialyse est ensuite effectuée durant 4h contre 1L tampon de dialyse (20mM MOPS pH 7,5; 150mM NaCl ; 15% glycérol (v/v) ; 0,1% Tween 20 (v/v)) afin d'éliminer le glutathion qui inhibe l'activité de URT1. Par la suite les protéines recombinantes sont aliquotées et congelées de façon instantanée dans de l'azote liquide.

11. Marquage radioactif d'un oligoribonucléotide par la polynucléotide kinase

L'ARN synthétique est préalablement dénaturé 5 minutes à 70°C puis afin de le marquer radioactivement, il est incubé entre 15 à 30 minutes à 37°C en présence de 25µCi de [γ-^{32}P]-ATP et de 1U de T4 Polynucléotide Kinase dans du tampon TNK 1x. Les nucléotides [γ-^{32}P]-ATP non incorporés sont éliminés par une chromatographie d'exclusion sur colonne de Sephadex® G-50 (Sigma-Aldrich®) par une minute de centrifugation à 1 000g.

12. Activité catalytique *in vitro* de la protéine recombinante URT1

Les tests d'activité de URT1 *in vitro* sont effectués avec 100nM de protéine recombinante 6histidines-GST-URT1 ou 6histidines-GST-URT1$^{D491/3A}$

Les tampons de réactions sont soit:

-25mM potassium phosphate pH 7, 5mM $MgCl_2$, 5mM DTT, 10 % glycérol (v/v), 0,1µg/µl BSA, 0,1mM de UTP, rGTP, rCTP ou rATP

-soit 20mM MOPS pH 7,5, 100mM NaCl, 15% glycérol (v/v), 0.1% Tween 20 (v/v), 1mM $MgCl_2$, 0,5µg/µl BSA, 1mM UTP ou 0,5µCi de [α-^{32}P]-UTP.

En présence de 5nM d'ARN synthétique de 21 nucléotides marqués à l'extrémité 5' par la T4 Polynucléotide Kinase et [γ-^{32}P]-ATP ou de 500ng d'ARN non marqués. Les réactions effectuées avec d'autres protéines ajoutées contiennent 20nM de 6his-GST ou de 6his-GST-PAB2 (ratio de 1/5ème comparé à 6histidines-GST-URT1).

Les réactions sont ensuite quantifiées après précipitation des ARN au TCA et/ou résolues sur gel polyacrylamide pour les ARN.

13. Précipitation des ARN au TCA

Les [α-^{32}P]-UTP non incorporés aux ARN par URT1 lors des test d'activité catalytique *in vitro* sont éliminés par une précipitation des ARN au TCA. Les réactions d'activités de URT1 sont déposées sur un papier Whatman (1µL) puis séchées avant d'être précipités par trois incubations successives de 15 minutes dans une solution de 10% (v/v) de TCA. Les ARN sont ensuite lavés par deux incubations successives de cinq à dix minutes dans une solution d'éthanol 100%. La radioactivité incorporée aux ARN est quantifiée par le scintillateur Beckman Coulter LS 6500.

14. Gel de polyacrylamide pour ARN

Les ARN sont résolus sur gel de polyacrylamide de 17% (4M urée ; 1xTBE ; 17% polyacrylamide 19/1 (acrylamide/bisacrylamide)) selon leurs tailles lors d'une migration dans du TBE 1x (90mM Tris, 2,5mM EDTA, 90mM borate) durant une heure sous un ampérage de 25mA. Les réactions d'activité catalytique de URT1 sont résolues en présence de Decade™ Marker Ambion® (Invitrogen) préalablement marqué radioactivement.

15. Chromatographie d'exclusion de taille

La chromatographie d'exclusion de taille est effectuée en utilisant une colonne de Sephadex S200 (5/150) (GE) équilibrée avec le tampon suivant : 20 mM MOPS pH 7,5, 150mM NaCl, 15% glycérol (v/v). 50µL de protéine recombinante 6histidines-GST-URT1 purifiée à une concentration de 50ng/µL est injecté sur la colonne à une vitesse de 0,2mL/minutes avec un volume total de 3mL et un volume mort de 1mL. Des échantillons de 50 µL d'élution sont récoltés. La calibration de la taille est effectuée par des marqueurs de taille de 440 (ferritine), 158 (aldolase), 44 (ovalbumine) et 17 kDa (RNase A) dans des conditions identiques. Les échantillons sont analysés par coloration à l'argent.

16. Agro-infiltration de feuilles de *Nicotiana benthamiana*

Les feuilles de plantes de 5 à 6 semaines de *Nicotiana benthamiana* sont agro-infiltrées par *A. tumefaciens*. Les *A. tumefasciens* GV3101, pMP90 contenant les plasmides d'intérêts utilisés pour les agro-infiltrations sont préalablement cultivées sur une boite de Pétri contenant du milieu LB avec les antibiotiques appropriés à la sélection des plasmides durant 2 jours à 28°C. Ensuite, une colonie est transférée dans un milieu liquide LB de 2mL contenant toujours les antibiotiques appropriés à 28°C durant 2 jours. Cette pré-culture est diluée au 1/1000^{e} dans du LB

liquide en présence des antibiotiques et de 20mM d'acétosyringone durant 16 heures. Ces cultures sont culottées à température ambiante et re-suspendues dans une solution d'agro-infiltration (10mM de MgCl2 et 20mM d'acétosyringone). Une solution d'agro-infiltration ayant une absorbance à 600nm de 0,8 est utilisée pour effectuer l'agro-infiltration des feuilles de *N. benthamiana*. Les solutions d'agro-infiltration des différents plasmides sont mélangées à quantité égale en présence d'un inhibiteur de *silencing* (P19) et/ou d'un marqueur des foci cytoplasmiques. L'agro-infiltration est réalisée coté abaxial de la feuille à l'aide d'une seringue. Les plantes agro-infiltrées sont ensuite mises en cultures dans les mêmes conditions précédant l'agro-infiltration.

Après deux jours, les feuilles sont récoltées puis congelées instantanément dans de l'azote liquide et stockées à -80°C pour l'extraction protéique ou des cercles de feuilles sont montés sur des lames afin d'être observés au microscope confocal (lsm780 Zeiss).

17. Immuno-précipitation

Une feuille de *N. benthamiana* précédemment agro-infiltrée est broyée dans 3mL de tampon de lyse du kit d'isolation µMACS™ de Miltenyi Biotec en présence d'inhibiteur de protéase (*complete*-EDTA Roche). Les débris cellulaires sont éliminés par une centrifugation de 10 minutes à 16 000g. Les immuno-précipitations sont ensuite effectuées selon les conditions préconisées par le kit d'isolation µMACS™ anti-GFP ou anti-c-myc de Miltenyi Biotec, avec une incubation de 15 minutes à 4°C de l'extrait protéique avec les anticorps. Différemment des conditions standard, les lavages sont effectués avec le tampon de lyse. Cependant, l'élution est réalisée dans les conditions standard.

18. Extraction protéique

a) Extraction protéique d'une culture bactérienne

Un volume de la culture bactérienne est centrifugé, puis le surnageant est éliminé. Les protéines de la culture bactérienne sont extraites par une re-suspension des bactéries dans du tampon de lyse dénaturant (80mM Tris-HCl pH 6,8, 4M urée, 0,1M DTT, 2% SDS (p/v), 10% glycérol (v/v)). Les débris cellulaires sont éliminés par une centrifugation de cinq minutes à 16 000g. L'extrait protéique est ensuite incubé cinq minutes à 95°C.

b) Extraction protéique de feuilles de N. benthamiana

Des cercles de feuilles de diametre de 1mm de *N. benthamiana* sont broyés dans 200µL tampon d'extraction (80mM Tris-HCl pH 6,8 ; 4M urée ; 0,1M DTT ; 2% SDS (p/v), 10% glycérol (v/v)). Les débris cellulaires sont éliminés par une centrifugation de cinq minutes à 16 000g. L'extrait protéique est ensuite incubé cinq minutes à 95°C.

19. La résolution des protéines selon leur taille par gel SDS-PAGE

Les protéines sont incubées dans un tampon (Tris-HCl pH 6,8 ; 2% SDS (p/v) ; 0,1% de bleu de bromophénol (p/v) ; 3% β-mercaptoéthanol (v/v) ; 10% glycérol (v/v)). Le SDS charge négativement les protéines. Elles sont ensuite séparées durant une heure (25mA à 220V) selon leurs tailles lors de la migration dans un gel de polyacrylamide dénaturant composé d'un gel de concentration (125mM Tris-HCl pH 6,8 ; 4% polyacrylamide (v/v) ; 0,1% SDS (p/v)) et d'un gel de séparation (375mM Tris-HCl pH 8,8 ; 10% polyacrylamide (v/v) ; 0,1% SDS (p/v)). Les protéines les plus petites migrent plus vite que les protéines ayant une taille plus importante vers le pôle positif (cathode). Un marqueur de taille protéique est utilisé comme référence : le Pageruler™ (Fermentas).

Les protéines des gels peuvent être colorées à l'argent ou au bleu de coomassie ou être transférée sur une membrane pour une analyse en western blot.

20. Coloration des protéines à l'argent

Le matériel d'électrophorèse est préalablement incubé 10 minutes dans une solution de 1M de NaOH afin d'éliminer les contaminations par la créatine. Les gels de polyacrylamides SDS-PAGE utilisés contiennent 0,1% de sodium thiosulfate. Après une résolution des protéines selon leurs tailles, elles sont fixées 40 à 60 minutes par une solution (50% méthanol (v/v), 10% TCA (p/v), 2% $CuCl_2$ (p/v)). Le gel est ensuite lavé 30 minutes dans la solution de lavage 1 (10% éthanol (v/v), 5% acide acétique (v/v)). Le gel est ensuite incubé trois minutes dans une solution de 0,01% de $KMnO_4$ (p/v) puis relavé 10 minutes dans la solution de lavage 1 puis 10 minutes dans une solution de lavage 2 (10% éthanol (v/v)) et 10 minutes dans de l'eau milli-Q® (Millipore) afin d'éliminer l'éthanol. Les protéines sont ensuite mises en présence de 0,2% d'$AgNO_3$ (p/v) durant 15 minutes. L'excès d'argent est retiré par une incubation de 20 secondes dans de l'eau milli-Q® (Millipore). Le gel est ensuite mis en condition alcaline une minute dans 10% de K_2CO_3 (p/v) puis la révélation est

effectuée avec une solution de révélation (0,02% formaldéhyde (v/v), 10% de K_2CO_3 (p/v)). Lorsque la révélation est suffisante, la réaction est arrêtée par une incubation de cinq minutes dans une solution stop (1% acide acétique (v/v)).

21. Coloration des protéines au bleu de coomassie sur gel polyacrylamide

Les protéines sont mises en évidence par coloration de gel polyacrylamide au bleu de coomassie (0,25% bleu de coomassie R250 (p/v) ; 7% acide acétique (v/v) ; 50% méthanol (v/v)). Le gel est décoloré dans 10% acide acétique (v/v) ; 40% méthanol (v/v).

22. Transfert des protéines sur une membrane

Les protéines sont transférées sur une membrane d'Immobilion-P® (Millipore) par un transfert liquide (48mM Tris, 39mM glycine, 15% méthanol (v/v)). Le méthanol est utilisé pour précipiter les protéines sur la membrane. Le transfert est effectué à 250mA durant 45 minutes.

23. Western blot

Le western blot est une méthode permettant la révélation d'une protéine d'intérêt dans un échantillon grâce à des anticorps dirigés contre celle-ci ou contre une étiquette fusionnée à celle-ci.

Le Western blot consiste en 3 étapes : une saturation de la membrane (20mM Tris-HCl pH 7,4 ; 150mM NaCl ; 0,2% Tween-20 (v/v) ; 5% lait écrémé (p/v)), une incubation avec l'anticorps primaire qui se fixe à l'antigène. Les excès d'anticorps sont éliminés par trois lavage de cinq minutes successives dans du tampon de lavage (20mM Tris-HCl pH 7,4 ; 150mM NaCl ; 0,2% Tween-20 (v/v)). Ensuite les anticorps primaires non couplés à la HRP peuvent être reconnues par des anticorps secondaires couplés à la HRP. Les anticorps en excès sont aussi éliminés par trois lavages successifs de cinq minutes dans le tampon de lavage avant l'étape de révélation.

Cette étape de révélation est effectuée grâce au kit Lumi Light^{Plus} (Roche). Une réaction de chimioluminescence est catalysée par la péroxydase et produit une émission de photons là où les anticorps sont localisés. Le signal des photons est révélé par une exposition à des films auto-radiographiques.

24. Coloration des protéines sur membrane

Les protéines présentes sur les membranes sont colorées par un tampon contenant du bleu de coomassie (0,25% bleu de coomassie R250 (p/v) ; 50% méthanol (v/v) ; 7% acide acétique (v/v)). La membrane est décolorée en présence de méthanol 40% et le bleu de coomassie est éliminé avec du méthanol 100%.

25. Co-localisation des protéines exprimées de façon transitoire

Une expression transitoire des protéines est effectuée dans des feuilles de *Nicotiana benthamiana* par agro-infiltration de plantes de 4 à 5 semaines contenant les délétions de URT1 en présence ou non d'un marqueur des granules de stress (PAB2 ou RBP47) ou des *processing bodies* (DCP1 ou XRN4) couplé respectivement à la CFP, la GFP à la RFP. Afin d'optimiser le signal, les couples GFP-YFP, YFP-CFP, RFP-YFP sont utilisés et les feuilles de tabac sont observées deux jours après agro-infiltration.

26. Microscopie

Les protéines fluorescentes sont observées grâce au microscope confocal (lsm780 Zeiss) 40x, 1,4 NA. Les images sont prises par captures séquentielles des signaux fluorescents aux absorbances de 519 à 621nm pour l'YFP et de 454 à 516nm pour la détection de la CFP ou de 493 à 516nm pour la GFP et 519 à 521nm pour l'YFP ou 593 à 643 nm pour la RFP et 519 à 580nm pour l'YFP et en lumière blanche DIC (*Differential Interference Contrast*). Toutes les images sont prises avec un gain similaire.

Les photographies sont prises par le microscope confocal (lsm780 Zeiss) et analysées par le logiciel ImageJ.

27. Quantification de la colocalisation

a) JACoP

La quantification de la colocalisation des délétions de URT1 aux marqueurs de granules de stress ou de *processing bodies* est effectuée par l'utilisation du plugin JACoP du logiciel d'imageJ (Bolte and Cordelières, 2006). Celui-ci détermine de façon automatique un seuil de bruit de fond pour chaque image et calcule le coefficient de corrélation R de Pearson pour la zone sélectionnée.

Afin de palier les inconvénients de la méthode de JACoP, la quantification de la colocalisation est aussi effectuée automatiquement grâce à un plugin appelé FOCI que nous avons mis en place avec la plateforme de microscopie de l'IBMP.

b) FOCI

La sélection des spots est effectuée par une méthode en deux étapes. Premièrement, un bruit de fond est évalué afin de déterminer un maximum de spot ajustable (ici 30) à partir du canal d'intérêt (ici YFP). Ceci permet l'analyse automatique des images de plusieurs expériences. Les spots sont ensuite filtrés selon leurs tailles afin de ne retenir que les spots inférieurs à 1,88 µm (6 pixels) de diamètre. Ensuite, les spots sélectionnés sont ensuite analysés et la corrélation entre l'intensité des fluorescences YFP et CFP/GFP est calculée.

Cette méthode permet ainsi d'éliminer les spots les moins significatifs et ceux qui risquent d'interférer dans l'évaluation de la colocalisation. De plus, elle évite aussi une mauvaise estimation de la colocalisation de deux protéines lorsque celle-ci n'ont pas la même localisation (ex : lorsque l'une est nucléo-cytoplasmique et l'autre cytoplasmique) ce qui pourrait baisser le coefficient de colocalisation. D'autre part, elle élimine les foci « non physiologique » lorsqu'une protéine forme des gros agrégats lors d'une expression transitoire. Ces agrégats peuvent avoir un signal qui fuit entre les différents canaux.

Annexes

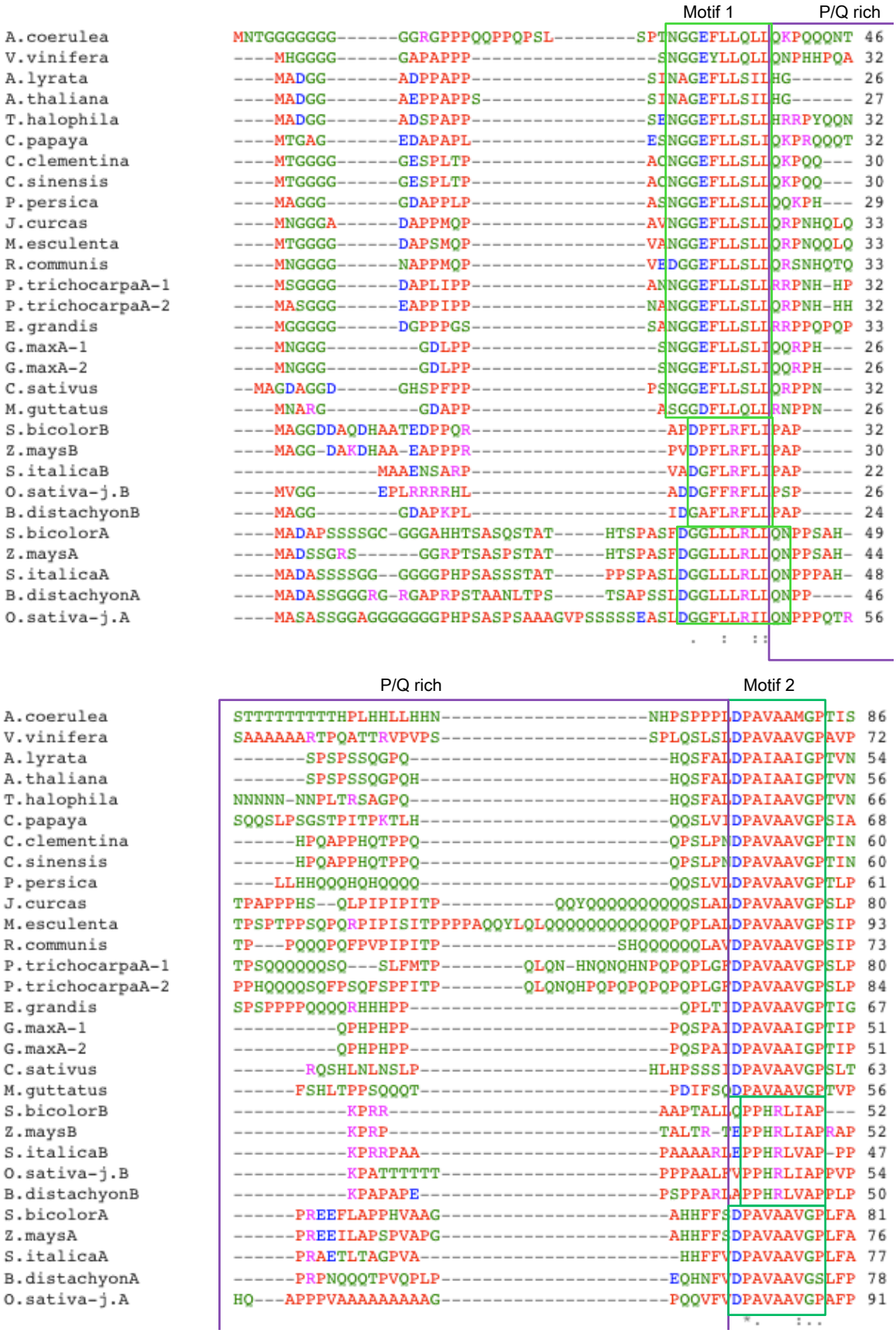
Annexe 1.1

Séquence protéique de URT1

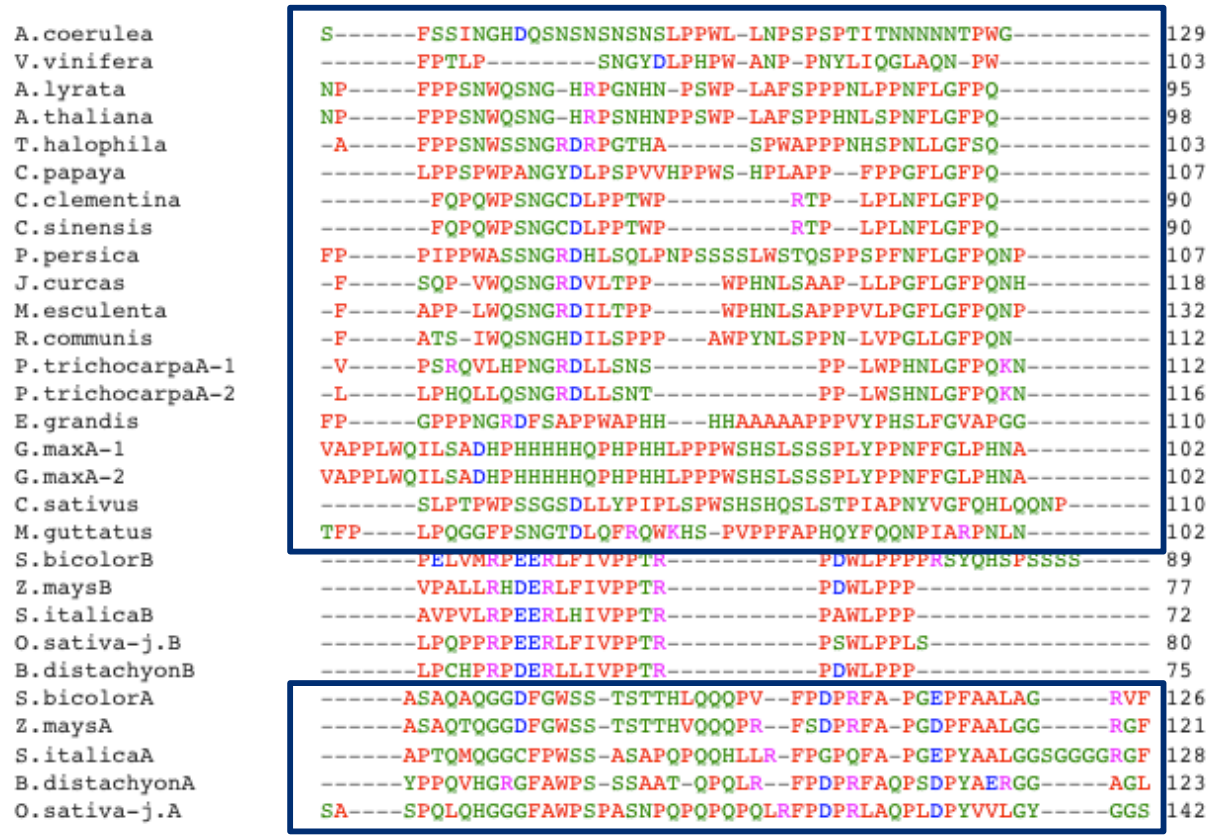
MADGGAEPPAPPSSINAGEFLLSILHGSPSPSSQGPQHHQSFALDPAIAAIGP
TVNNPFPPSNWQSNGHRPSNHNPPSWPLAFSPPHNLSPNFLGFPQFPPSPFTTNQ
FDGNQRVSPEDAYRLGFPGTTNPAIQSMVQQQQQQQLPPPQSETRKLVFGSFSG
DATQSLNGLHNGNLKYDSNQHEQLMRHPQSTLSNSNMDPNLSHHRNHDLHEQRG
GHSGRGNWGHIGNNGRGLKSTPPPPPPGFSSNQRGWDMSLGSKDDDRGMGRN
HDQAMGEHSKVWNQSVDFSAEANRLRGLSIQNESKFNLSQQIDHPGPPKGASLHS
VSAADAADSFSMLNKEARRGGERREELGQLSKAKREGNANSDEIEDFGEDIVKSLL
LEDETGEKDANDGKKDSKTSREKESRVDNRGQRLLGQKARMVKMYMACRNDIHR
YDATFIAIYKSLIPAEEELEKQRQLMAHLENLVAKEWPHAKLYLYGSCANSFGFPKS
DIDVCLAIEGDDINKSEMLLKLAEILESDNLQNVQALTRARVPIVKLMDPVTGISCDICI
NNVLAVVNTKLLRDYAQIDVRLRQLAFIVKHWAKSRRVNETYQGTLSSYAYVLMCIH
FLQQRRPPILPCLQEMEPTYSVRVDNIRCTYFDNVDRLRNFGSNNRETIAELVWGF
FNYWAYAHDYAYNVVSVRTGSILGKREKDWTRRVGNDRHLICIEDPFETSHDLGRV
VDKFSIRVLREEFERAARIMHQDPNPCAKLLEPYIPEDNNGQGHN

Annexe 1.2

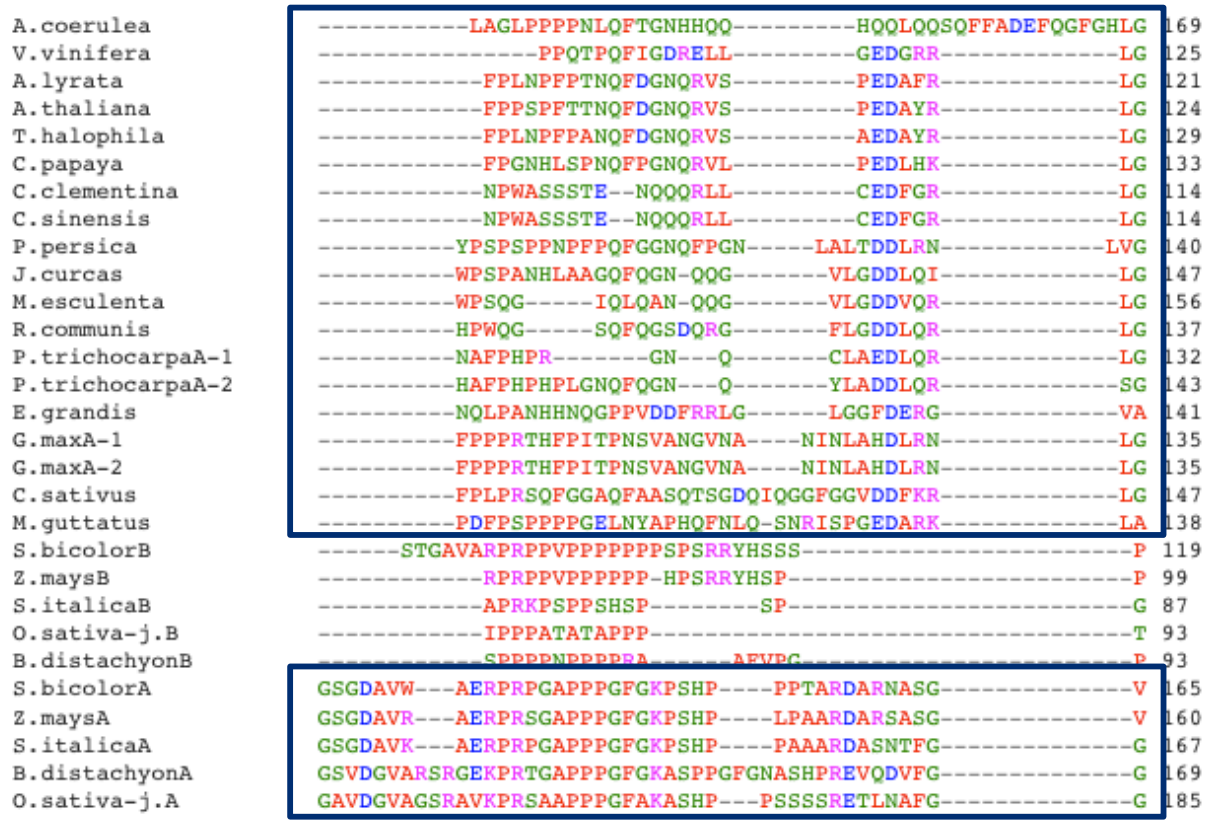
Alignement des homologues de URT1 chez les Angiospermes

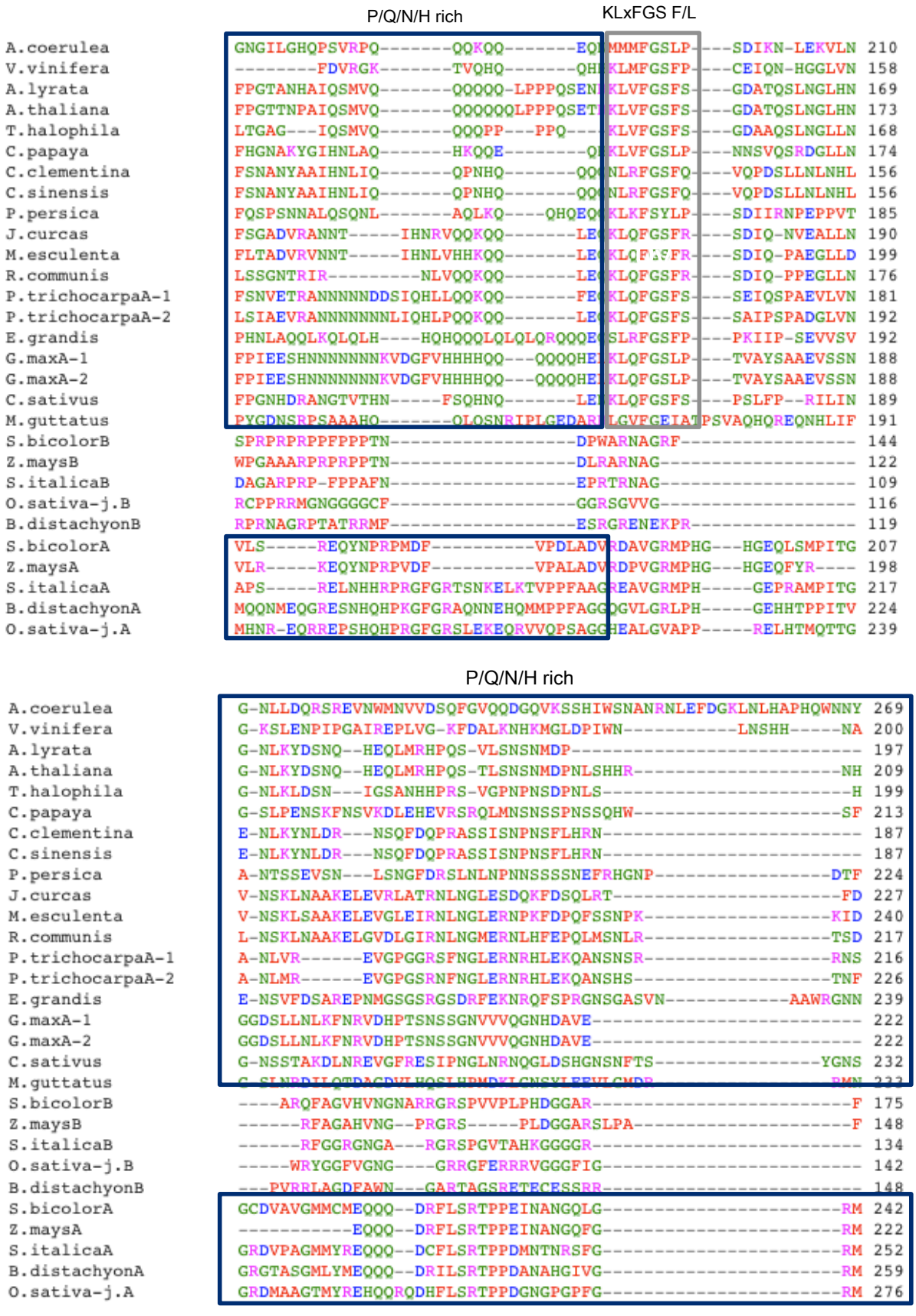

Motif 1
P/Q rich
A.coerulea MNTGGGGGGG------GGRGPPPQQPPQPSL--------SPTNGGEFLLQLLQKPQQQNT 46
V.vinifera ----MHGGGG------GAPAPPP----------------SNGGEYLLQLLQNPHHPQA 32
A.lyrata ----MADGG-------ADPPAPP----------------SINAGEFLLSILHG------ 26
A.thaliana ----MADGG-------AEPPAPPS---------------SINAGEFLLSILHG------ 27
T.halophila ----MADGG-------ADSPAPP----------------SENGGEFLLSLLHRRPYQQN 32
C.papaya ----MTGAG-------EDAPAPL----------------ESNGGEFLLSLIQKPRQQQT 32
C.clementina ----MTGGGG------GESPLTP----------------ACNGGEFLLSLLQKPQQ--- 30
C.sinensis ----MTGGGG------GESPLTP----------------ACNGGEFLLSLLQKPQQ--- 30
P.persica ----MAGGG-------GDAPPLP----------------ASNGGEFLLSLLQQKPH--- 29
J.curcas ----MNGGGA------DAPPMQP----------------AVNGGEFLLSLLQRPNHQLQ 33
M.esculenta ----MTGGGG------DAPSMQP----------------VANGGEFLLSLLQRPNQQLQ 33
R.communis ----MNGGGG------NAPPMQP----------------VEDGGEFLLSLLQRSNHQTQ 33
P.trichocarpaA-1 ----MSGGGG------DAPLIPP----------------ANNGGEFLLSLLRRPNH-HP 32
P.trichocarpaA-2 ----MASGGG------EAPPIPP----------------NANGGEFLLSLLQRPNH-HH 32
E.grandis ----MGGGGG------DGPPPGS----------------SANGGEFLLSLLRRPPQPQP 33
G.maxA-1 ----MNGGG---------GDLPP----------------SNGGEFLLSLIQQRPH--- 26
G.maxA-2 ----MNGGG---------GDLPP----------------SNGGEFLLSLIQQRPH--- 26
C.sativus --MAGDAGGD------GHSPFPP----------------PSNGGEFLLSLLQRPPN--- 32
M.guttatus ----MNARG---------GDAPP----------------ASGGDFLLQLLRNPPN--- 26
S.bicolorB ----MAGGDDAQDHAATEDPPQR----------------APDPFLRFLIPAP----- 32
Z.maysB ----MAGG-DAKDHAA-EAPPPR----------------PVDPFLRFLIPAP----- 30
S.italicaB -------------MAAENSARP----------------VADGFLRFLIPAP----- 22
O.sativa-j.B ----MVGG------EPLRRRRHL----------------ADDGFFRFLIPSP----- 26
B.distachyonB ----MAGG--------GDAPKPL----------------IDGAFLRFLIPAP----- 24
S.bicolorA ----MADAPSSSSGC-GGGAHHTSASQSTAT-----HTSPASFDGGLLLRLIQNPPSAH- 49
Z.maysA ----MADSSGRS------GGRPTSASPSTAT-----HTSPASFDGGLLLRLIQNPPSAH- 44
S.italicaA ----MADASSSSGG--GGGGPHPSASSSTAT-----PPSPASLDGGLLLRLIQNPPPAH- 48
B.distachyonA ----MADASSGGGRG-RGAPRPSTAANLTPS-----TSAPSSLDGGLLLRLIQNPP---- 46
O.sativa-j.A ----MASASSGGAGGGGGGGPHPSASPSAAAGVPSSSSSEASLDGGFLLRILQNPPPQTR 56
P/Q rich
Motif 2
A.coerulea STTTTTTTTTHPLHHLLHHN----------NHPSPPPLDPAVAAMGPTIS 86
V.vinifera SAAAAAARTPQATTRVPVPS----------SPLQSLSLDPAVAAVGPAVP 72
A.lyrata -------SPSPSSQGPQ----------HQSFALDPAIAAIGPTVN 54
A.thaliana -------SPSPSSQGPQH----------HQSFALDPAIAAIGPTVN 56
T.halophila NNNNN-NNPLTRSAGPQ----------HQSFALDPAIAAVGPTVN 66
C.papaya SQQSLPSGSTPITPKTLH----------QQSLVLDPAVAAVGPSIA 68
C.clementina ------HPQAPPHQTPPQ----------QPSLPNDPAVAAVGPTIN 60
C.sinensis ------HPQAPPHQTPPQ----------QPSLPNDPAVAAVGPTIN 60
P.persica ----LLHHQQQHQHQQQQ----------QQSLVLDPAVAAVGPTLP 61
J.curcas TPAPPPHS--QLPIPIPITP----------QQYQQQQQQQQQQSLALDPAVAAVGPSLP 80
M.esculenta TPSPTPPSQPQRPIPISITPPPPAQQYLQLQQQQQQQQQQQQPQPLALDPAVAAVGPSIP 93
R.communis TP---PQQQPQFPVPIPITP----------SHQQQQQQLAVDPAVAAVGPSIP 73
P.trichocarpaA-1 TPSQQQQQQSQ---SLFMTP--------QLQN-HNQNQHNPQPQPLGFDPAVAAVGPSLP 80
P.trichocarpaA-2 PPHQQQQSQFPSQFSPFITP--------QLQNQHPQPQPQPQPQPLGFDPAVAAVGPSLP 84
E.grandis SPSPPPPQQQQRHHHPP----------QPLTIDPAVAAVGPTIG 67
G.maxA-1 ----------QPHPHPP----------PQSPAIDPAVAAIGPTIP 51
G.maxA-2 ----------QPHPHPP----------PQSPAIDPAVAAIGPTIP 51
C.sativus -------RQSHLNLNSLP----------HLHPSSSIDPAVAAVGPSLT 63
M.guttatus ------FSHLTPPSQQQT----------PDIFSQDPAVAAVGPTVP 56
S.bicolorB ----------KPRR----------AAPTALLQPPHRLIAP--- 52
Z.maysB ----------KPRP----------TALTR-TEPPHRLIAPRAP 52
S.italicaB ----------KPRRPAA----------PAAAARLEPPHRLVAP-PP 47
O.sativa-j.B ----------KPATTTTTT----------PPPAALPVPPHRLIAPPVP 54
B.distachyonB ----------KPAPAPE----------PSPPARLAPPHRLVAPPLP 50
S.bicolorA ------PREEFLAPPHVAAG----------AHHFFSDPAVAAVGPLFA 81
Z.maysA ------PREEILAPSPVAPG----------AHHFFSDPAVAAVGPLFA 76
S.italicaA ------PRAETLTAGPVA----------HHFFVDPAVAAVGPLFA 77
B.distachyonA ------PRPNQQQTPVQPLP----------EQHNFVDPAVAAVGSLFP 78
O.sativa-j.A HQ---APPPVAAAAAAAAAG----------PQQVFVDPAVAAVGPAFP 91

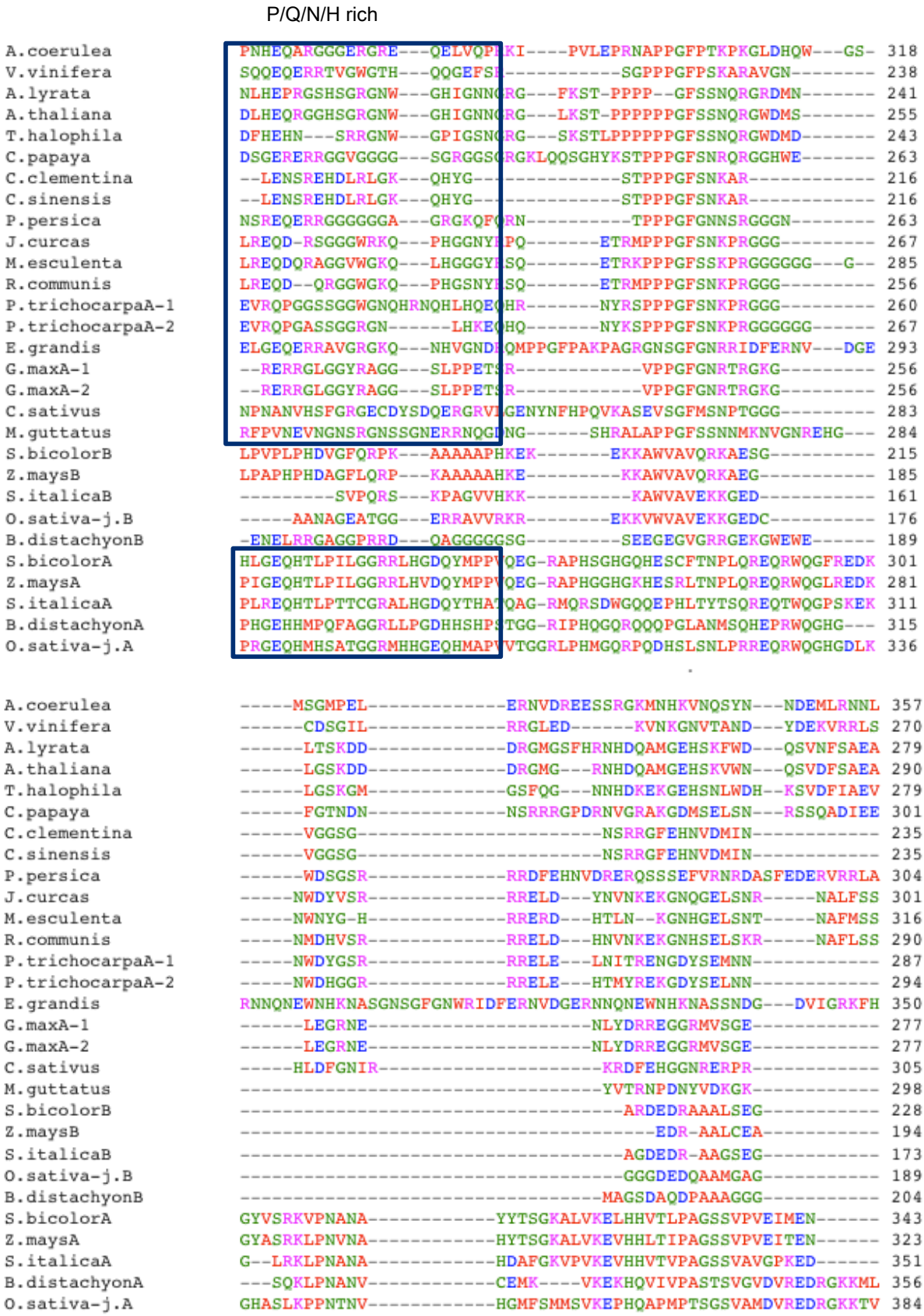
P/Q/N/H rich

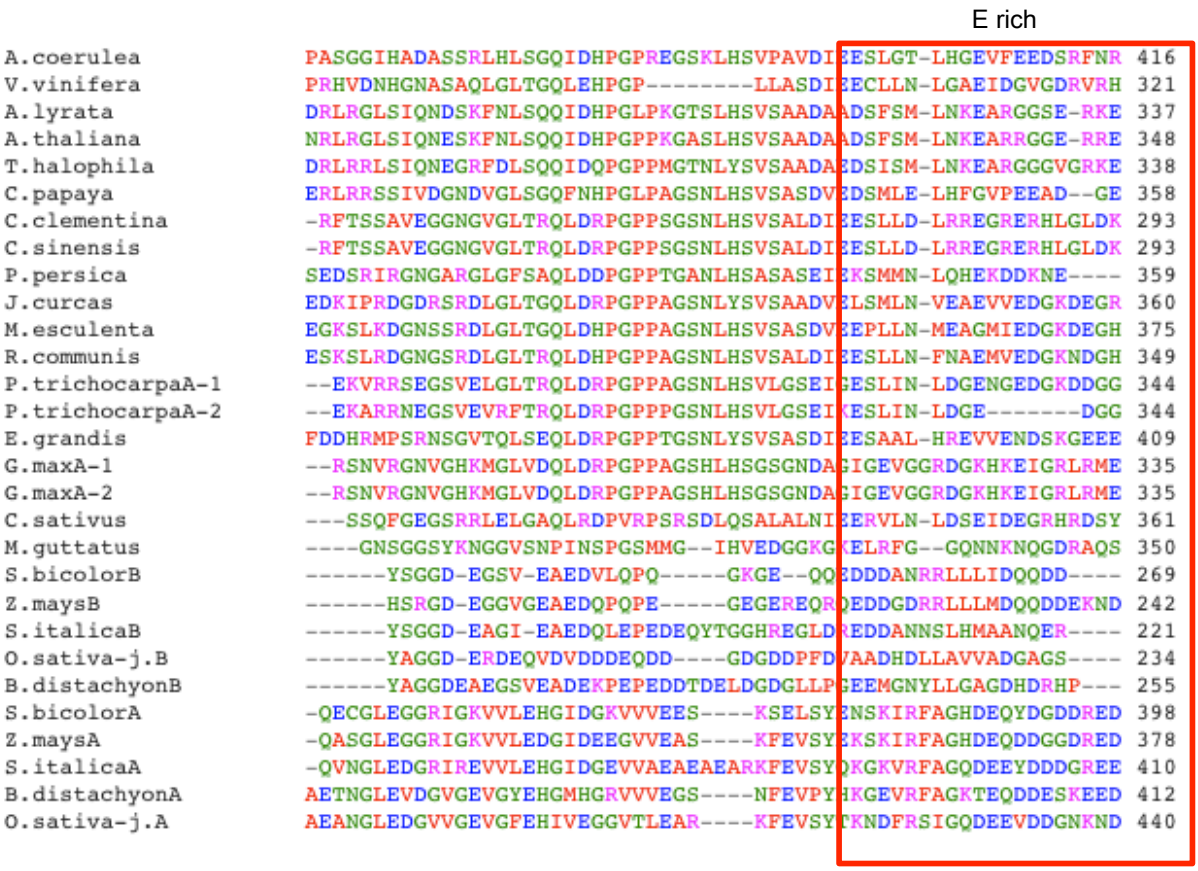


P/Q/N/H rich

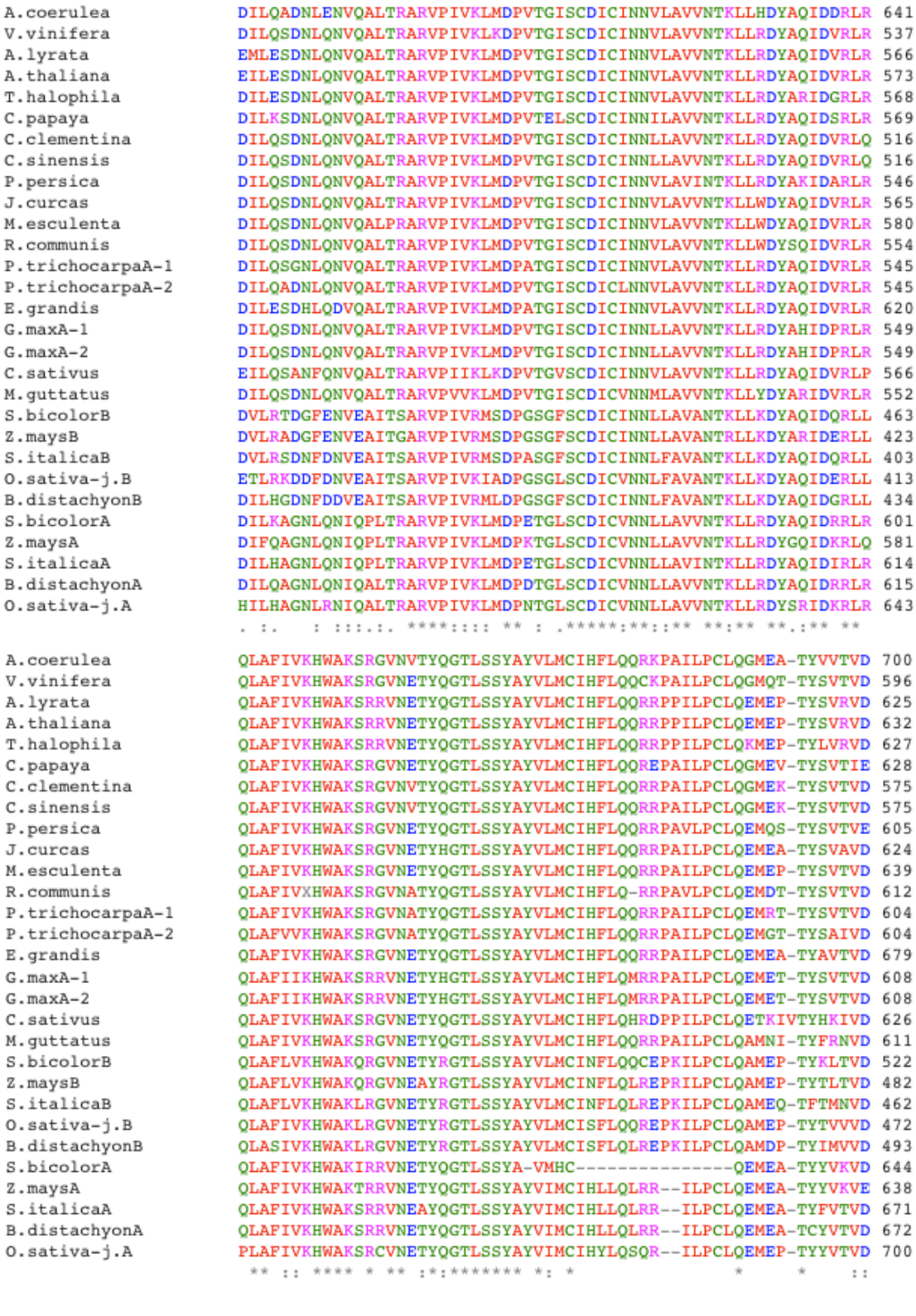



P/Q/N/H rich
KLxFGS F/L
A.coerulea
V.vinifera
A.lyrata
A.thaliana
T.halophila
C.papaya
C.clementina
C.sinensis
P.persica
J.curcas
M.esculenta
R.communis
P.trichocarpaA-1
P.trichocarpaA-2
E.grandis
G.maxA-1
G.maxA-2
C.sativus
M.guttatus
S.bicolorB
Z.maysB
S.italicaB
O.sativa-j.B
B.distachyonB
S.bicolorA
Z.maysA
S.italicaA
B.distachyonA
O.sativa-j.A
P/Q/N/H rich
A.coerulea
V.vinifera
A.lyrata
A.thaliana
T.halophila
C.papaya
C.clementina
C.sinensis
P.persica
J.curcas
M.esculenta
R.communis
P.trichocarpaA-1
P.trichocarpaA-2
E.grandis
G.maxA-1
G.maxA-2
C.sativus
M.guttatus
S.bicolorB
Z.maysB
S.italicaB
O.sativa-j.B
B.distachyonB
S.bicolorA
Z.maysA
S.italicaA
B.distachyonA
O.sativa-j.A


P/Q/N/H rich
A.coerulea PNHEQARGGGERGRE---QELVQPIKI----PVLEPRNAPPGFPTKPKGLDHQW---GS- 318
V.vinifera SQQEQERRTVGWGTH---QQGEFSI----------SGPPPGFPSKARAVGN-------- 238
A.lyrata NLHEPRGSHSGRGNW---GHIGNNGRG---FKST-PPPP--GFSSNQRGRDMN------- 241
A.thaliana DLHEQRGGHSGRGNW---GHIGNNGRG---LKST-PPPPPPGFSSNQRGWDMS------- 255
T.halophila DFHEHN---SRRGNW---GPIGSNGRG---SKSTLPPPPPPGFSSNQRGWDMD------- 243
C.papaya DSGERERRGGVGGGG---SGRGGSGRGKLQQSGHYKSTPPPGFSNRQRGGHWE------- 263
C.clementina --LENSREHDLRLGK---QHYG----------STPPPGFSNKAR---------- 216
C.sinensis --LENSREHDLRLGK---QHYG----------STPPPGFSNKAR---------- 216
P.persica NSREQERRGGGGGGA---GRGKQFQRN---------TPPPGFGNNSRGGGN-------- 263
J.curcas LREQD-RSGGGWRKQ---PHGGNYIPQ-------ETRMPPPGFSNKPRGGG--------- 267
M.esculenta LREQDQRAGGVWGKQ---LHGGGYISQ-------ETRKPPPGFSSKPRGGGGGG---G-- 285
R.communis LREQD--QRGGWGKQ---PHGSNYISQ-------ETRMPPPGFSNKPRGGG--------- 256
P.trichocarpaA-1 EVRQPGGSSGGWGNQHRNQHLHQEQHR-------NYRSPPPGFSNKPRGGG--------- 260
P.trichocarpaA-2 EVRQPGASSGGRGN------LHKEQHQ-------NYKSPPPGFSNKPRGGGGGG------ 267
E.grandis ELGEQERRAVGRGKQ---NHVGNDIQMPPGFPAKPAGRGNSGFGNRRIDFERNV---DGE 293
G.maxA-1 --RERRGLGGYRAGG---SLPPETSR----------VPPGFGNRTRGKG--------- 256
G.maxA-2 --RERRGLGGYRAGG---SLPPETSR----------VPPGFGNRTRGKG--------- 256
C.sativus NPNANVHSFGRGECDYSDQERGRVIGENYNFHPQVKASEVSGFMSNPTGGG--------- 283
M.guttatus RFPVNEVNGNSRGNSSGNERRNQGDNG------SHRALAPPGFSSNNMKNVGNREHG--- 284
S.bicolorB LPVPLPHDVGFQRPK---AAAAAPHKEK-------EKKAWVAVQRKAESG--------- 215
Z.maysB LPAPHPHDAGFLQRP---KAAAAAHKE---------KKAWVAVQRKAEG---------- 185
S.italicaB ---------SVPQRS---KPAGVVHKK----------KAWVAVEKKGED---------- 161
O.sativa-j.B -----AANAGEATGG---ERRAVVRKR--------EKKVWVAVEKKGEDC--------- 176
B.distachyonB -ENELRRGAGGPRRD---QAGGGGGSG---------SEEGEGVGRRGEKGWEWE------ 189
S.bicolorA HLGEQHTLPILGGRRLHGDQYMPPVQEG-RAPHSGHGQHESCFTNPLQREQRWQGFREDK 301
Z.maysA PIGEQHTLPILGGRRLHVDQYMPPVQEG-RAPHGGHGKHESRLTNPLQREQRWQGLREDK 281
S.italicaA PLREQHTLPTTCGRALHGDQYTHATQAG-RMQRSDWGQQEPHLTYTSQREQTWQGPSKEK 311
B.distachyonA PHGEHHMPQFAGGRLLPGDHHSHPSTGG-RIPHQGQRQQQPGLANMSQHEPRWQGHG--- 315
O.sativa-j.A PRGEQHMHSATGGRMHHGEQHMAPVVTGGRLPHMGQRPQDHSLSNLPRREQRWQGHGDLK 336
A.coerulea -----MSGMPEL----------ERNVDREESSRGKMNHKVNQSYN---NDEMLRNNL 357
V.vinifera ------CDSGIL----------RRGLED------KVNKGNVTAND---YDEKVRRLS 270
A.lyrata ------LTSKDD----------DRGMGSFHRNHDQAMGEHSKFWD---QSVNFSAEA 279
A.thaliana ------LGSKDD----------DRGMG---RNHDQAMGEHSKVWN---QSVDFSAEA 290
T.halophila ------LGSKGM----------GSFQG---NNHDKEKGEHSNLWDH--KSVDFIAEV 279
C.papaya ------FGTNDN----------NSRRRGPDRNVGRAKGDMSELSN---RSSQADIEE 301
C.clementina ------VGGSG----------NSRRGFEHNVDMIN---------- 235
C.sinensis ------VGGSG----------NSRRGFEHNVDMIN---------- 235
P.persica ------WDSGSR----------RRDFEHNVDRERQSSSEFVRNRDASFEDERVRRLA 304
J.curcas -----NWDYVSR----------RRELD---YNVNKEKGNQGELSNR-----NALFSS 301
M.esculenta -----NWNYG-H----------RRERD---HTLN--KGNHGELSNT-----NAFMSS 316
R.communis -----NMDHVSR----------RRELD---HNVNKEKGNHSELSKR-----NAFLSS 290
P.trichocarpaA-1 -----NWDYGSR----------RRELE---LNITRENGDYSEMNN---------- 287
P.trichocarpaA-2 -----NWDHGGR----------RRELE---HTMYREKGDYSELNN---------- 294
E.grandis RNNQNEWNHKNASGNSGFGNWRIDFERNVDGERNNQNEWNHKNASSNDG---DVIGRKFH 350
G.maxA-1 ------LEGRNE----------NLYDRREGGRMVSGE---------- 277
G.maxA-2 ------LEGRNE----------NLYDRREGGRMVSGE---------- 277
C.sativus -----HLDFGNIR----------KRDFEHGGNRERPR---------- 305
M.guttatus ----------YVTRNPDNYVDKGK---------- 298
S.bicolorB ----------ARDEDRAAALSEG---------- 228
Z.maysB ----------EDR-AALCEA---------- 194
S.italicaB ----------AGDEDR-AAGSEG---------- 173
O.sativa-j.B ----------GGGDEDQAAMGAG---------- 189
B.distachyonB ----------MAGSDAQDPAAAGGG---------- 204
S.bicolorA GYVSRKVPNANA----------YYTSGKALVKELHHVTLPAGSSVPVEIMEN------ 343
Z.maysA GYASRKLPNVNA----------HYTSGKALVKEVHHLTIPAGSSVPVEITEN------ 323
S.italicaA G--LRKLPNANA----------HDAFGKVPVKEVHHVTVPAGSSVAVGPKED------ 351
B.distachyonA ---SQKLPNANV----------CEMK----VKEKHQVIVPASTSVGVDVREDRGKKML 356
O.sativa-j.A GHASLKPPNTNV----------HGMFSMMSVKEPHQAPMPTSGSVAMDVREDRGKKTV 384


E rich
A.coerulea PASGGIHADASSRLHLSGQIDHPGPREGSKLHSVPAVDIEESLGT-LHGEVFEEDSRFNR 416
V.vinifera PRHVDNHGNASAQLGLTGQLEHPGP--------LLASDIEECLLN-LGAEIDGVGDRVRH 321
A.lyrata DRLRGLSIQNDSKFNLSQQIDHPGLPKGTSLHSVSAADAADSFSM-LNKEARGGSE-RKE 337
A.thaliana NRLRGLSIQNESKFNLSQQIDHPGPPKGASLHSVSAADAADSFSM-LNKEARRGGE-RRE 348
T.halophila DRLRRLSIQNEGRFDLSQQIDQPGPPMGTNLYSVSAADAEDSISM-LNKEARGGGVGRKE 338
C.papaya ERLRRSSIVDGNDVGLSGQFNHPGLPAGSNLHSVSASDVEDSMLE-LHFGVPEEAD--GE 358
C.clementina -RFTSSAVEGGNGVGLTRQLDRPGPPSGSNLHSVSALDIEESLLD-LRREGRERHLGLDK 293
C.sinensis -RFTSSAVEGGNGVGLTRQLDRPGPPSGSNLHSVSALDIEESLLD-LRREGRERHLGLDK 293
P.persica SEDSRIRGNGARGLGFSAQLDDPGPPTGANLHSASASEIEKSMMN-LQHEKDDKNE---- 359
J.curcas EDKIPRDGDRSRDLGLTGQLDRPGPPAGSNLYSVSAADVELSMLN-VEAEVVEDGKDEGR 360
M.esculenta EGKSLKDGNSSRDLGLTGQLDHPGPPAGSNLHSVSASDVEEPLLN-MEAGMIEDGKDEGH 375
R.communis ESKSLRDGNGSRDLGLTRQLDHPGPPAGSNLHSVSALDIEESLLN-FNAEMVEDGKNDGH 349
P.trichocarpaA-1 --EKVRRSEGSVELGLTRQLDRPGPPAGSNLHSVLGSEIGESLIN-LDGENGEDGKDDGG 344
P.trichocarpaA-2 --EKARRNEGSVEVRFTRQLDRPGPPPGSNLHSVLGSEIKESLIN-LDGE-------DGG 344
E.grandis FDDHRMPSRNSGVTQLSEQLDRPGPPTGSNLYSVSASDIEESAAL-HREVVENDSKGEEE 409
G.maxA-1 --RSNVRGNVGHKMGLVDQLDRPGPPAGSHLHSGSGNDAGIGEVGGRDGKHKEIGRLRME 335
G.maxA-2 --RSNVRGNVGHKMGLVDQLDRPGPPAGSHLHSGSGNDAGIGEVGGRDGKHKEIGRLRME 335
C.sativus ---SSQFGEGSRRLELGAQLRDPVRPSRSDLQSALALNIEERVLN-LDSEIDEGRHRDSY 361
M.guttatus ----GNSGGSYKNGGVSNPINSPGSMMG--IHVEDGGKGKELRFG--GQNNKNQGDRAQS 350
S.bicolorB ------YSGGD-EGSV-EAEDVLQPQ-----GKGE--QQEDDDANRRLLLIDQQDD---- 269
Z.maysB ------HSRGD-EGGVGEAEDQPQPE-----GEGEREQRQEDDGDRRLLLMDQQDDEKND 242
S.italicaB ------YSGGD-EAGI-EAEDQLEPEDEQYTGGHREGLDREDDANNSLHMAANQER---- 221
O.sativa-j.B ------YAGGD-ERDEQVDVDDDEQDD----GDGDDPFDVAADHDLLAVVADGAGS---- 234
B.distachyonB ------YAGGDEAEGSVEADEKPEPEDDTDELDGDGLLPGEEMGNYLLGAGDHDRHP--- 255
S.bicolorA -QECGLEGGRIGKVVLEHGIDGKVVVEES----KSELSYENSKIRFAGHDEQYDGDDRED 398
Z.maysA -QASGLEGGRIGKVVLEDGIDEEGVVEAS----KFEVSYEKSKIRFAGHDEQDDGGDRED 378
S.italicaA -QVNGLEDGRIREVVLEHGIDGEVVAEAEAEARKFEVSYQKGKVRFAGQDEEYDDDGREE 410
B.distachyonA AETNGLEVDGVGEVGYEHGMHGRVVVEGS----NFEVPYHKGEVRFAGKTEQDDESKEED 412
O.sativa-j.A AEANGLEDGVVGEVGFEHIVEGGVTLEAR----KFEVSYTKNDFRSIGQDEEVDDGNKND 440
E rich
A.coerulea NENKSKSDASGPNYQRSGR-------GWQDDLSEQLMNSVRLSNG-LDKSSTSHYDN--- 465
V.vinifera QKQ---------GMRREGQ-------GNLDDLSEEMTGSLVLEDGSQDKNDTNQHHN--- 362
A.lyrata ELGRL------SKGKREGNANSGPVDDEIEDFGEDIVKSLLLEDETGEKDAKDGKKD--S 389
A.thaliana ELGQL------SKAKREGNANS----DEIEDFGEDIVKSLLLEDETGEKDANDGKKD--S 396
T.halophila ELGQF------SKGKREGNGECGPGDDDIEGFGEDIVESLLLEDETDDKNAKDGKNN--S 390
C.papaya GYERE------ERLRRE----------GEELVDSLLLEGETE---GRDEKRQ--H 392
C.clementina RRENG------PGYSQGG--------DDMDDFGEDLVDSLLPDDESELKNDTHERNDKKH 339
C.sinensis RRENG------PGYSQGG--------DDMDDFGEDLVDSLLPDDESELKNDTHERNDKKH 339
P.persica ----------EDDKNEAKQHHN--- 371
J.curcas ELDEA----------GEELVDSLLLEGESDGKNDKKQNRH--- 390
M.esculenta ELDDV----------GEELVDSLLLEGESDGKNDRKQNRH--- 405
R.communis DLDDV----------GEELADTLLLEGESEGKNDNKQNRH--- 379
P.trichocarpaA-1 ELDDL----------GEELVDSLLLNGQSEGKKDKKQS----- 372
P.trichocarpaA-2 LLDDL----------GEELMDSLLLEGESDGKKDKKQS----- 372
E.grandis GEGAAVKR----------AGKLGANESDD-LLEDELDEKNERRQQRS--- 445
G.maxA-1 GVPESGGG----------GADVDVLGEQLADSLLVKDESDDRTNLRQRRR--- 375
G.maxA-2 GVPESGGG----------GADVDVLGEQLADSLLVKDESDDRTNLRQRRR--- 375
C.sativus QGHDS----------QELDNIGEQLADSLLLEDEPDEKSDSKFIRR--- 397
M.guttatus KMNGI----------EDQMGSLGIEEESGETSDKKKNPH--- 379
S.bicolorB ----------DVESSLPTDGTLDCSGRAG 288
Z.maysB D----------DVKRSLPTDG---------- 253
S.italicaB ----------SDGGGGGDERIS------ 233
O.sativa-j.B ----------EKPMEQLG---------- 242
B.distachyonB ----------GDDGGEETT---------- 264
S.bicolorA ----------AIIEQMTQNLVIDGSGDAKGVVLEKTI--- 425
Z.maysA ----------AIIEQMAQNLLIDGNGDAKGVVLEKPI--- 405
S.italicaA G----------SIIEQMTESLMIDGNVEAKGVVLQKTI--- 438
B.distachyonA ----------AIIEQFMDTVVIEGNGEANNVVVQSSG--- 439
O.sativa-j.A D----------ATIEQLMETLVIDDNGEAKSTVVQING--- 468

E rich
A.coerulea --SYDKNFRPDMSRELRPPGQRMRNHKR-VVDCRSDIDRMNAPFLAIYESLVPTEEEKSK 522
V.vinifera --SRNRDFRSDT-RGQRMLSQRVRNLKR-HMECRRDIGTLNFRFLSIYESLIPEEEEKAK 418
A.lyrata KTSREKDSRMDN-RGQRLLGQKARMVKM-YMACRNDIHRYDASFIAVYKSLIPAEEELEK 447
A.thaliana KTSREKESRVDN-RGQRLLGQKARMVKM-YMACRNDIHRYDATFIAIYKSLIPAEEELEK 454
T.halophila RTSREKESRMDT-RGQRLLRQSSRIHRWRYMACRYDIHMYDAPFIAVYESLIPAEEELEK 449
C.papaya RNFREKESRVDS-RGQRLLSQRLRNLRR-HIVCRRDIDSLNAHFLSIYESLMPLEEEKAK 450
C.clementina RNSRDKEIRSDN-RGKRLLSQRMRNLKW-QIECRADIGRLNAPFLAIYESLIPAEEEKAK 397
C.sinensis RNSRDKEIRSDN-RGKRLLSQRMRNLKW-QIECRADIGRLNAPFLAIYESLIPAEEEKAK 397
P.persica --SREKDSRSDN-RGQHLLSQRMRIFKS-QMQCRFDIDRLNAPFLAIYDSLIPTEEEKAK 427
J.curcas --SREKESRSDN-RGQRTLSQRMRMLKR-QMECRRDIDRLNAPFLAIYESLVPPEEEKAK 446
M.esculenta --SREKESRLDN-RGQRMLSQRMRILKR-QMECRRDIDRLNASFIEIYESLVLPQEEKAK 461
R.communis --SRDKESRSDN-RGQQILSQRMRMLKR-QMECRRDIDRLNVSFLAIYESLIPPEEEKSK 435
P.trichocarpaA-1 ----NKESRSDN-RGKKILSQRMRMLKK-QTQCCLDIDRLNAAFLAIYESLIPPEEEKMK 426
P.trichocarpaA-2 ----SKESRSDS-RGHNILSQRMRMLKR-QMQCRLDIDRLNAAFLAIYESLVPPEEETAK 426
E.grandis --QREKDFRSDS-RGQRLLTQRMRNFKS-SMECRWDIDRLNVPFLAIYDSLIPPEEEKAK 501
G.maxA-1 ----EKDVRLSDSRGQQIMSQRGRMYRR-QMMCRRDIDVFNVPFLAIYGSLIPPEEEKLK 430
G.maxA-2 ----EKDVRLSDSRGQQIMSQRGRMYRR-QMMCRRDIDVFNVPFLAIYGSLIPPEEEKLK 430
C.sativus ----EKDCRGNR-----LLTHRERIARK-HIHCRGDIDMLTIPLLRIYESLIPPEEEKEK 447
M.guttatus ----DKEYRSDQ-RGQWIMGQRMRHVKM-QTACRKDIDRFNSQFLTVFESLIPADEERVK 433
S.bicolorB GGDRPSEVTPGP---DQVPRLQGRMRRP-EAECRHDIDTFTPGLLALYESLKPSEEHRSK 344
Z.maysB -----TQATPPP---NRVPRLQGRTRGS-PAERRRDMDAFAPGFLAVYESLKPSGEQRSK 304
S.italicaB -----EQLVSQS---NQAPQLSGRMRRW-QVECRPEIDTFTPGLLALYESLKPSEEHKSK 284
O.sativa-j.B -----SPPDQPP---PPPPRQRVGTRRW-RVERRHDIDAFTPGLLSLYESLNPSEEHKAK 293
B.distachyonB -----ENLVSES---NGLLRLRVKMRRG-WIECRHDIDSFAPDLLSLYESLKPSEEHKSK 315
S.bicolorA --LRSKDFRSDFSRGHHVSSQRIRFQRR-NRPCRYDIDRFTPNFLSIFESLMPSEEEIAK 482
Z.maysA --LRSKDFRSDFSRGHHVSSQRIRFQRR-NRPCRYDIDRFTPNFLSIFESLVPSEEEIAN 462
S.italicaA --TRSKDFRSDFSRGHHVSSQRIRFQRR-NRPCRYDIDRFAPSFLSIFESLVPSQDEIAK 495
B.distachyonA --SRSKDFRSDFSRGHHVSSQRVRFQRR-IRACRYDIDQFSPSFLSIFESLVPSEEEIAK 496
O.sativa-j.A --SRSKNFRSD-SRGKNVSSQSVRFQRR-IRPCRYDIDQFTPSFMSIFESLVPSDEEISK 524
QKQLLTSLEKLVNKEWPTAQLYLYGSCANSFGVSKSDIDVCLALEEPDIN-KSEILLKLA 581
V.vinifera QKQLLTLLEKLVSKEWPKAQLFLYGSCANSFGVSKSDIDVCLAIDDADIN-KSEFLLKLA 477
A.lyrata QRQLMAHLENLVAKEWPHAKLYLYGSCANSFGFPKSDIDVCLAIEGDDIN-KSEMLLKLA 506
A.thaliana QRQLMAHLENLVAKEWPHAKLYLYGSCANSFGFPKSDIDVCLAIEGDDIN-KSEMLLKLA 513
T.halophila QKQLMARLEHLVGKEWPHAKLYLYGSCANSFGFPKSDIDVCLAIEDDDIN-KSEMLLKLA 508
C.papaya QKQLLKLLEKLVREEWPEARLFLYGSCANSFGVSKSDIDVCLAIEDTGTK-KSEILLKLA 509
C.clementina QKKLLTLLEKLVCKEWPDARLYLYGSCANSFGVSKSDIDVCLAINDSEIN-KSEVLLKLA 456
C.sinensis QKKLLTLLEKLVCKEWPDARLYLYGSCANSFGVSKSDIDVCLAINDSEIN-KSEVLLKLA 456
P.persica QNQLFTLLETLITKEWPEAQLYVYGSCGNSFGVSKSDIDLCLAIDVADDN-KSEILLRLA 486
J.curcas QKQLLSLLEKLVNKEWPQARLYLYGSCANSFGVLKSDIDVCLAIQNADIN-KSEVLLKLA 505
M.esculenta QKQLLAVLEKLVNREWPQAKLYLYGSCANSFGVLKSDIDVCLAIEDANIN-KSEVLLKLA 520
R.communis QKQLLTLLEKLVNKEWPEARLYLYGSCANSFGVRKSDIDVCLAIQDADIN-KSEVLLKLA 494
P.trichocarpaA-1 QELFLMSLEKLVNKEWPEARLYLYGSGANSFGVSKSDIDVCLAIEDAEIN-KSEVLLKLA 485
P.trichocarpaA-2 QKQFFMLLEKLVSKEWPEARLYLYGSCANSFGVSKSDIDVCLTIEDAEIK-KSEVLLKLA 485
E.grandis QKQLLTMLQKLVSKEWPEARLYLYGSCANSFGVCKSDIDVCLAIQDANID-KSEVILKLA 560
G.maxA-1 QKKLVALLEKLVSKEWPTAKLYLYGSCANSFGVSKSDIDVCLAIEEADME-KSKIIMKLA 489
G.maxA-2 QKKLVALLEKLVSKEWPTAKLYLYGSCANSFGVSKSDIDVCLAIEEADME-KSKIIMKLA 489
C.sativus QRQLLISLEKLVVNEWPHAHLFLFGSCANSFGVSNSDVDVCLVLRDADID-KSEILLKLA 506
M.guttatus QKQLLTVLEKLVAKEWPDARLYLYGSCANSFGFSKSDLDVCLAIELGNNE-KSEVVLKLA 492
S.bicolorB QKQLVDSLAKSVSKEWPNAQMHLYGSCANSFGTSHSDVDVCLEMETGTQD-AIEVLVKLA 403
Z.maysB QEQLVDSLARSVRKECPHAQMHLYGSCANSFGTSHSDVDVCLEMETGAES-AVEVLVRLA 363
S.italicaB QNQLVDSLTKSVSKEWPDAQLHLYGSCANSFGTSHSDVDVCLEMETGTES-TAEALLKLA 343
O.sativa-j.B QRQLIESLTNSVSKEWPNAQLHLYGSCANSFGNSHSDVDVCLQIDTAAEENIAELLLALA 353
B.distachyonB QTQLIDSLAKSLSKEWPNARLHLYGSCANSFGTSHSDVDVCLEIEIGTES-TVEILQRLA 374
S.bicolorA QKQLLTSLSRLINMEWPNAKLYLYGSCANSFGFSNSDIDLCLSIDDKEMS-KVDIILKLA 541
Z.maysA QKQLLTSLSRLINTEWPNSKLYLYGSCANSFGFSNSDIDLCLSIDNKEMS-KVDIILKLA 521
S.italicaA QKQLLTSLSRIINKEWPNSKLYLYGSCANSFGFSNSDIDLCLSIDDKEMS-KVDIILKLA 554
B.distachyonA QNQLLIALSRLITMEWPNSKLYLYGSCANSFGFSNSDIDLCLSIDNNEMS-KVDIILKLA 555
O.sativa-j.A QKQLLATLSRLINKEWPNSKLYLYGSCANSFGFSNSDIDLCLSIDEKEMS-KVDIILKLA 583


A.coerulea DILQADNLENVQALTRARVPIVKLMDPVTGISCDICINNVLAVVNTKLLHDYAQIDDRLR 641
V.vinifera DILQSDNLQNVQALTRARVPIVKLKDPVTGISCDICINNVLAVVNTKLLRDYAQIDVRLR 537
A.lyrata EMLESDNLQNVQALTRARVPIVKLMDPVTGISCDICINNVLAVVNTKLLRDYAQIDVRLR 566
A.thaliana EILESDNLQNVQALTRARVPIVKLMDPVTGISCDICINNVLAVVNTKLLRDYAQIDVRLR 573
T.halophila DILESDNLQNVQALTRARVPIVKLMDPVTGISCDICINNVLAVVNTKLLRDYARIDGRLR 568
C.papaya DILKSDNLQNVQALTRARVPIVKLMDPVTELSCDICINNILAVVNTKLLRDYAQIDSRLR 569
C.clementina DILQSDNLQNVQALTRARVPIVKLMDPVTGISCDICINNLLAVVNTKLLRDYAQIDVRLQ 516
C.sinensis DILQSDNLQNVQALTRARVPIVKLMDPVTGISCDICINNLLAVVNTKLLRDYAQIDVRLQ 516
P.persica DILQSDNLQNVQALTRARVPIVKLMDPVTGISCDICINNVLAVINTKLLRDYAKIDARLR 546
J.curcas DILQSDNLQNVQALTRARVPIVKLMDPVTGISCDICINNVLAVVNTKLLWDYAQIDVRLR 565
M.esculenta DILQSDNLQNVQALPRARVPIVKLMDPVTGISCDICINNVLAVVNTKLLWDYAQIDVRLR 580
R.communis DILQSDNLQNVQALTRARVPIVKLMDPVTGISCDICINNVLAVVNTKLLWDYSQIDVRLR 554
P.trichocarpaA-1 DILQSGNLQNVQALTRARVPIVKLMDPATGISCDICINNVLAVVNTKLLRDYAQIDVRLR 545
P.trichocarpaA-2 DILQADNLQNVQALTRARVPIVKLMDPVTGISCDICLNNVLAVVNTKLLRDYAQIDVRLR 545
E.grandis DILESDHLQDVQALTRARVPIVKLMDPATGISCDICINNVLAVVNTKLLRDYAQIDVRLR 620
G.maxA-1 DILQSDNLQNVQALTRARVPIVKLMDPVTGISCDICINNLLAVVNTKLLRDYAHIDPRLR 549
G.maxA-2 DILQSDNLQNVQALTRARVPIVKLMDPVTGISCDICINNLLAVVNTKLLRDYAHIDPRLR 549
C.sativus EILQSANFQNVQALTRARVPIIKLKDPVTGVSCDICINNVLAVVNTKLLRDYAQIDVRLP 566
M.guttatus DILQSDNLQNVQALTRARVPVVKLMDPVTGISCDICVNNMLAVVNTKLLYDYARIDVRLR 552
S.bicolorB DVLRTDGFENVEAITSARVPIVRMSDPGSGFSCDICINNLLAVANTKLLKDYAQIDQRLL 463
Z.maysB DVLRADGFENVEAITGARVPIVRMSDPGSGFSCDICINNLLAVANTRLLKDYARIDERLL 423
S.italicaB DVLRSDNFDNVEAITSARVPIVRMSDPASGFSCDICINNLFAVANTKLLKDYAQIDQRLL 403
O.sativa-j.B ETLRKDDFDNVEAITSARVPIVKIADPGSGLSCDICVNNLFAVANTKLLKDYAQIDERLL 413
B.distachyonB DILHGDNFDDVEAITSARVPIVRMLDPGSGFSCDICINNLFAVANTKLLKDYAQIDGRLL 434
S.bicolorA DILKAGNLQNIQPLTRARVPIVKLMDPETGLSCDICVNNLLAVVNTKLLRDYAQIDRRLR 601
Z.maysA DIFQAGNLQNIQPLTRARVPIVKLMDPKTGLSCDICVNNLLAVVNTKLLRDYGQIDKRLQ 581
S.italicaA DILHAGNLQNIQPLTRARVPIVKLMDPETGLSCDICVNNLLAVINTKLLRDYAQIDIRLR 614
B.distachyonA DILQAGNLQNIQALTRARVPIVKLMDPDTGLSCDICVNNLLAVVNTKLLRDYAQIDRRLR 615
O.sativa-j.A HILHAGNLRNIQALTRARVPIVKLMDPNTGLSCDICVNNLLAVVNTKLLRDYSRIDKRLR 643
QLAFIVKHWAKSRGVNVTYQGTLSSYAYVLMCIHFLQQRKPAILPCLQGMEA-TYVVTVD 700
QLAFIVKHWAKSRGVNETYQGTLSSYAYVLMCIHFLQQCKPAILPCLQGMQT-TYSVTVD 596
QLAFIVKHWAKSRRVNETYQGTLSSYAYVLMCIHFLQQRRPPILPCLQEMEP-TYSVRVD 625
QLAFIVKHWAKSRRVNETYQGTLSSYAYVLMCIHFLQQRRPPILPCLQEMEP-TYSVRVD 632
QLAFIVKHWAKSRRVNETYQGTLSSYAYVLMCIHFLQQRRPPILPCLQKMEP-TYLVRVD 627
QLAFIVKHWAKSRGVNETYQGTLSSYAYVLMCIHFLQQREPAILPCLQGMEV-TYSVTIE 628
QLAFIVKHWAKSRGVNVTYQGTLSSYAYVLMCIHFLQQRRPAILPCLQGMEK-TYSVTVD 575
QLAFIVKHWAKSRGVNVTYQGTLSSYAYVLMCIHFLQQRRPAILPCLQGMEK-TYSVTVD 575
QLAFIVKHWAKSRGVNETYQGTLSSYAYVLMCIHFLQQRRPAVLPCLQEMQS-TYSVTVE 605
QLAFIVKHWAKSRGVNETYHGTLSSYAYVLMCIHFLQQRRPAILPCLQEMEA-TYSVAVD 624
QLAFIVKHWAKSRGVNETYQGTLSSYAYVLMCIHFLQQRRPAILPCLQEMEP-TYSVTVD 639
QLAFIVXHWAKSRGVNATYQGTLSSYAYVLMCIHFLQ-RRPAVLPCLQEMDT-TYSVTVD 612
QLAFIVKHWAKSRGVNATYQGTLSSYAYVLMCIHFLQQRRPAILPCLQEMRT-TYSVTVD 604
QLAFVVKHWAKSRGVNATYQGTLSSYAYVLMCIHFLQQRRPAILPCLQEMGT-TYSAIVD 604
QLAFIVKHWAKSRGVNETYQGTLSSYAYVLMCIHFLQQRRPAILPCLQEMEA-TYAVTVD 679
QLAFIIKHWAKSRRVNETYHGTLSSYAYVLMCIHFLQMRRPAILPCLQEMET-TYSVTVD 608
QLAFIIKHWAKSRRVNETYHGTLSSYAYVLMCIHFLQMRRPAILPCLQEMET-TYSVTVD 608
QLAFIVKHWAKSRGVNETYQGTLSSYAYVLMCIHFLQHRDPPILPCLQETKIVTYHKIVD 626
QLAFIVKHWAKSRGVNETYQGTLSSYAYVLMCIHFLQQRRPAILPCLQAMNI-TYFRNVD 611
QLAFLVKHWAKQRGVNETYRGTLSSYAYVLMCINFLQQCEPKILPCLQAMEP-TYKLTVD 522
QLAFLVKHWAKQRGVNEAYRGTLSSYAYVLMCINFLQLREPRILPCLQAMEP-TYTLTVD 482
QLAFLVKHWAKLRGVNETYRGTLSSYAYVLMCINFLQLREPKILPCLQAMEQ-TFTMNVD 462
QLAFIVKHWAKLRGVNETYRGTLSSYAYVLMCISFLQQREPKILPCLQAMEP-TYTVVVD 472
QLASIVKHWAKLRGVNETYRGTLSSYAYVLMCISFLQLREPKILPCLQAMDP-TYIMVVD 493
QLAFIVKHWAKIRRVNETYQGTLSSYA-VMHC----------QEMEA-TYYVKVD 644
QLAFIVKHWAKTRRVNETYQGTLSSYAYVIMCIHLLQLRR--ILPCLQEMEA-TYYVKVE 638
QLAFIVKHWAKSRRVNEAYQGTLSSYAYVIMCIHLLQLRR--ILPCLQEMEA-TYFVTVD 671
QLAFIVKHWAKSRRVNETYQGTLSSYAYVIMCIHLLQLRR--ILPCLQEMEA-TCYVTVD 672
PLAFIVKHWAKSRCVNETYQGTLSSYAYVIMCIHYLQSQR--ILPCLQEMEP-TYYVTVD 700

```
A.coerulea          DIECAYFDQVQHLRDFGARNKERIAALVWAFFDYWAYRHDYANGVISVRTGSIISKRAKD 760
V.vinifera          DIQCAFFDQVERLRHFGSHNKESIAQLVWAFFNYWAYHHDYANDVISIRTGSIISKREKD 656
A.lyrata            NIRCAYFDNVDRLRNFGSSNRETIAELVWGFFNYWAYAHDYAYNVVSVRTGSILGKREKD 685
A.thaliana          NIRCTYFDNVDRLRNFGSNNRETIAELVWGFFNYWAYAHDYAYNVVSVRTGSILGKREKD 692
T.halophila         NIRCAYFDNVETLG--RSSNKETIAELVWGFFSYWAYGHDYANTVVSVRTGSILGKREKN 685
C.papaya            DVECAFFDQVEKLQNFGSHNKETIAQLVWGFFNYWAYCHDYANSVISVRTGSIISKREKD 688
C.clementina        DIECAYFDQVDKLHGFGSRNKESIGRLVWAFFNYWAYGHDYASNVISVRTGSTISKQEKD 635
C.sinensis          DIECAYFDQVDKLHGFGSRNKESIGRLVWAFFNYWAYGHDYASNVISVRTGSTISKQEKD 635
P.persica           NIECAFFDQVDKLRDFGSHNREPIAKLVWAFFNYWAYGHDYANSVISVRTGSILSKREKD 665
J.curcas            DIQCAYFDQVEKLRGFGSRNKETIAQLVWAFFNYWAYRHDYANAVISIRTGSIISKREKD 684
M.esculenta         DIKCAYFDRVEKLQGFGSRNREPIAQLVWAFFNYWAYRHDYVNAVVSVRTGSIISKQEKD 699
R.communis          DIECAYFDQVEKLQGLGSRNKETIAQLVWAFFNYWAYRHDYANDVISVRKGTIISKWEKD 672
P.trichocarpaA-1    DIQCAYFDQVEKLRGFGSRNKETIARLVWAFFNYWAYGHDYANAVISVRTGSILSKHEKE 664
P.trichocarpaA-2    DIRCAYFDQVEKLRGFGSRNKETIAQLVWAFFNYWAYRHDYANGVISVRTGSIISKREKD 664
E.grandis           NIECAFFDKVEKLRDFGSRNRETIAQLVWAFFHYWAYCHDYTNTVISVRTGSXXXKRTKD 739
G.maxA-1            DIHCAYFDQVEKLSDFGRHNKESIAQLVRGFFHYWAYCHDYANTVISVRTGSIISKREKD 668
G.maxA-2            DIHCAYFDQVEKLSDFGRHNKESIAQLVRGFFHYWAYCHDYANTVISVRTGSIISKREKD 668
C.sativus           NIECAYFDQVEKLKTFGSDNKESVARLVWGFFHYWAYCHDYANTVVSVRTKNTVSKRAKD 686
M.guttatus          NVECAYFDQVEKLGNFGSRNGESIAHLVWAFFHYWAYCHDYATDVVSVRTGSTLSKRTKD 671
S.bicolorB          GTECAYFDKVDQLQGFGADNKASVAELLWGFFHYWASQHHYKRDVISVRLGKTISKQEKR 582
Z.maysB             GTECAYFDRVDQLQGFGAGNKASIGELLWGFFHYWASQHRYKRDVISVRLGKTIRKQEKG 542
S.italicaB          GTECAYFDEVHHLQDFGADNKESIAELLWAFFHYWAFDHDYRRDVISVRMGNTISKQEKN 522
O.sativa-j.B        GTECAYFDQVDQLKDFGAENKESIAELLWAFFHYWAFHHDYRNDVISVRMGNTISKQEKN 532
B.distachyonB       DTKCTYFDDIHQLHDFGAENKESIAELLWAFFHYWAFQHDYRKDVISIRMGKIISKKEKD 553
S.bicolorA          ENNCAYFDQVEKLNNYGAHNRDTISRLLWAFFHYWAYEHDYTRDVISIRTGRIISKERKD 704
Z.maysA             EINCAYFDQVDKLNNYGAHNRDTVSRLLWSFFHYWAYEHDYTRDVISIRTGRIISKERKD 698
S.italicaA          ESNCAYFDQVDKLNNYGAHSRDTVSRLLWAFFHYWAYEHDYTRDVISIRTGRIISKERKD 731
B.distachyonA       DNHCAYFDQVDKLNNYGAHNKETISSLLWAFFHYWAYQHDYTKDVISIRTGRIISKHMKD 732
O.sativa-j.A        NNICAYFDQVDKLNGFGAQCKDTLSRLLWGFFRYWAYAHNYTKDVISIRTGRTISKNMKD 760
                       *:*** :. *          :. *: .** ***  * *   *:*:*      *  *

A.coerulea          WTRRIGNDRHLICIEDPFEISHDLGRVVDKYSIKTLREEFERAANIMHLDPDPCVTLFEP 820
V.vinifera          WTRRKGNDRHLICIEDPFEISHDLGRVVDKFSIKVLREEFERAAYIMQYDPDPCVTLFEP 716
A.lyrata            WTRRVGNDRHLICIEDPFETSHDLGRVVDKFSIRVLREEFERAARIMHQEPNPCAKLFEP 745
A.thaliana          WTRRVGNDRHLICIEDPFETSHDLGRVVDKFSIRVLREEFERAARIMHQDPNPCAKLLEP 752
T.halophila         WTRRVGNDRHLICIEDPFETSHDLGRVVDKFSIRVLREEFERAAEVMHEDPNPSAKLFEP 745
C.papaya            WTRRIGNDRHLICIEDPFEISHDLGRVVDKFSIKVLREEFERAAEIMQYNPNPCAALFEP 748
C.clementina        WTRRIGNDRHLICIEDPFEVTHDLGRVVDKFSIKVLREEFERAAEIMQHDPNPCVKLFEP 695
C.sinensis          WTRRIGNDRHLICIEDPFEVTHDLGRVVDKFSIKVLREEFERAAEIMQHDPNPCVKLFEP 695
P.persica           WTRRIGNDRHLICIEDPFEISHDLGRVVDKYSIKVLREEFERAAGIMQYDANPCVKLFEP 725
J.curcas            WTRRIGNDRHLICIEDPFEISHDLGRVVDKYSIKVLREEFERAADIMQYDPNPCVTLFEP 744
M.esculenta         WTRRVGNDRHLICIEDPFETSHDLGRVVDKHTIKVLREEFERAADILQYDPNPCVTLFEP 759
R.communis          WTRRIGNDRHLICIEDPFEISHDLGRVVDKYSIKVLREEFERAANIMQYDPNPCVKLFEP 732
P.trichocarpaA-1    WTRRIGNDRHLICIEDPFEISHDLGRVVDKFSIKVLREEFERAADILQYDPNPCATLFEP 724
P.trichocarpaA-2    WTRRIGNDRHLICIEDPFEISHDLGRVVDKFSIKVLREEFERAADIMQYDPNPCVTLFEP 724
E.grandis           WTRRIGNDRHLICIEDPFDVSHDLGRVVDKFSIKVLREEFERAAEIMQYDPNPCVTLFKP 799
G.maxA-1            WTRRIGNDRHLICIEDPFEISHDLGRVVDKHSIKVLREEFERAAEIMQNDPNPCIKLFEP 728
G.maxA-2            WTRRIGNDRHLICIEDPFEISHDLGRVVDKHSIKVLREEFERAAEIMQNDPNPCIKLFEP 728
C.sativus           WTRRIGKDRHLICIEDPFETSHDLGRVVDKYSIKVLREEFERAATILQTYPNPCEKLFEP 746
M.guttatus          WTRRVGNDRHLICIEDPFEVSHDLGRVVDKYSIRVLREEFERAAEIMQYDPNPCVTLFHP 731
S.bicolorB          WTTRVGNDRHLVCIEDPFEVSHDLGRVVDRQTIRILREEMERAAVVLQHDDDPCVTLFEP 642
Z.maysB             WTTRVGGDRHLMCIEDPFETGHDLGRVVDRQTIWIIREEMERAATVLQNDDDPCATLFEP 602
S.italicaB          WTTRVGNDRHLMCIEDPFETGHDLGRVVDRQTLRILREEFERAAGVLQYNDDPCVTLFEP 582
O.sativa-j.B        WTTRVGNDRHLICIEDPFETSHDLGRVVDRQTIRVLREEFERAATILQYDDDPCVALFEP 592
B.distachyonB       WTTRVGNDRHLMCIEDPFEISHDLGRVVDRQTIRIIIEEFERAADVLQHDNDPCVTLFKP 613
S.bicolorA          WTRRVGNDRHLICIEDPFEISHDLGRVVDKFSIKILREEFERAANILQFDPNPSMTLFEP 764
Z.maysA             WTRRVGNDRHLICIEDPFEISHDLGRVVDKFTIKILREEFERAANILQFDPNPSVTLFEP 758
S.italicaA          WTRRVGNDRHLICIEDPFETSHDLGRVVDKFTIKILREEFERAANILQFDPNPSVTLFEP 791
B.distachyonA       WTRRVGNDRHLICIEDPFETSHDLGRVVDKFSIKILREEFERAANILQYDPNPSVTLFEP 792
O.sativa-j.A        WTRRIGNDRHLICIEDPFETSHDLGRVVDNRSIWALREEFERAAEILHLDPNPSITLFEP 820
                    ** * * ****:******:  ********. ::  : **:**** :::   :*.  *:.*
```

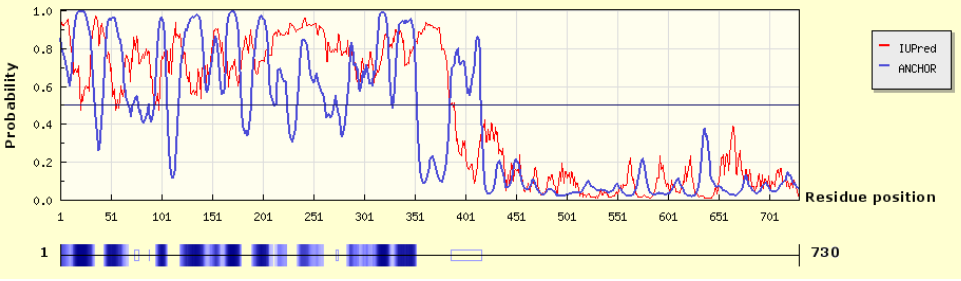
```
A.coerulea              Y-VPN------------------ 824
V.vinifera              Y-VPA------------------ 720
A.lyrata                Y-IPGDNNGQGNN---------- 757
A.thaliana              Y-IPEDNNGQGHN---------- 764
T.halophila             Y-LPGDNNGQAHN---------- 757
C.papaya                Y-IPG------------------ 752
C.clementina            Y-VPG------------------ 699
C.sinensis              Y-VPG------------------ 699
P.persica               Y-VPGV----------------- 730
J.curcas                Y-VPS------------------ 748
M.esculenta             Y-VP------------------- 762
R.communis              Y-VPS------------------ 736
P.trichocarpaA-1        Y-IPS------------------ 728
P.trichocarpaA-2        YNTPS----------------- 729
E.grandis               YSAS------------------ 803
G.maxA-1                Y-VPS------------------ 732
G.maxA-2                Y-VPS------------------ 732
C.sativus               F-VPS------------------ 750
M.guttatus              Y-VPN------------------ 735
S.bicolorB              YGCEN----------------- 647
Z.maysB                 YGYDL----------------- 607
S.italicaB              YD-------------------- 584
O.sativa-j.B            YDYES----------------- 597
B.distachyonB           YNYGT----------------- 618
S.bicolorA              YVPPL-LPNLLQEETANAASN-- 784
Z.maysA                 YVPPP-SPSLLQEETANAAEIEL 780
S.italicaA              YVPPP-LPTLMQEETVNGAEVEL 813
B.distachyonA           YVPPPPSTSMEREGTASTAGVVL 815
O.sativa-j.A            YVPSEAET--------------- 828
                        :
```

Alignment des homologues de URT1 chez les Angiospermes par ClustalW2

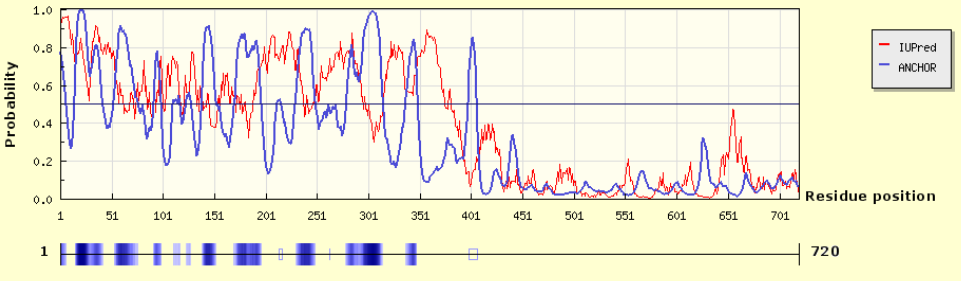
Annexe 1.3

Prédiction de la désorganisation des homologues d'URT1 chez les Angiospermes par IUPred et ANCHOR

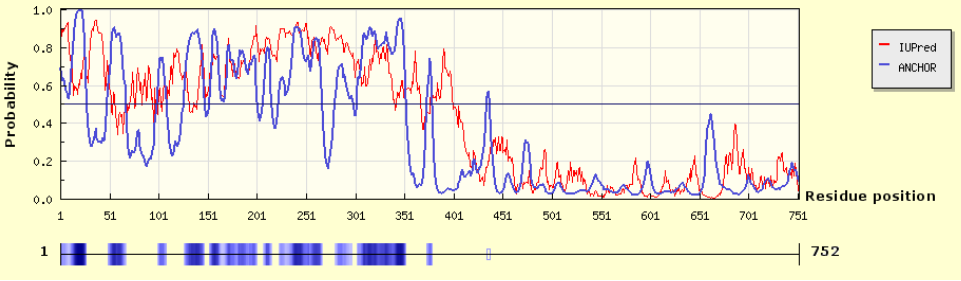
Prunus persica

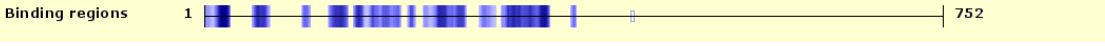


Vitis vinifera

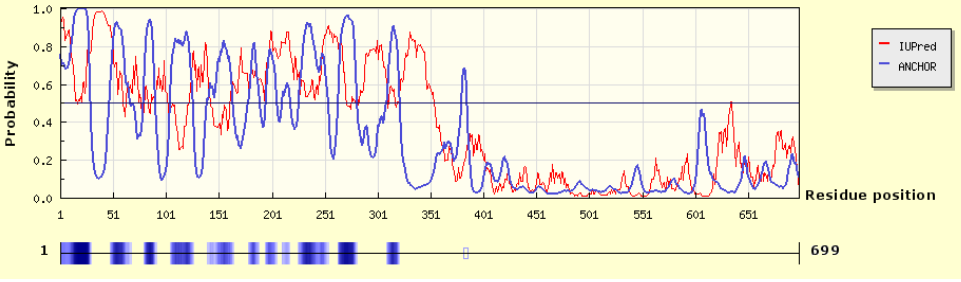


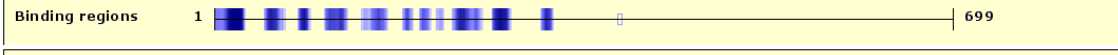
Carica papaya



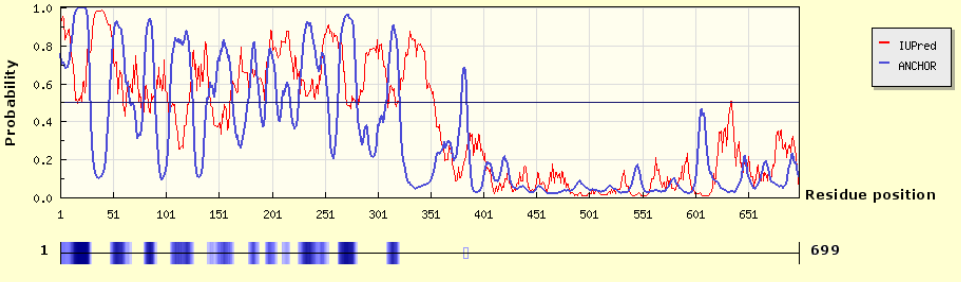


Citrus sinensis

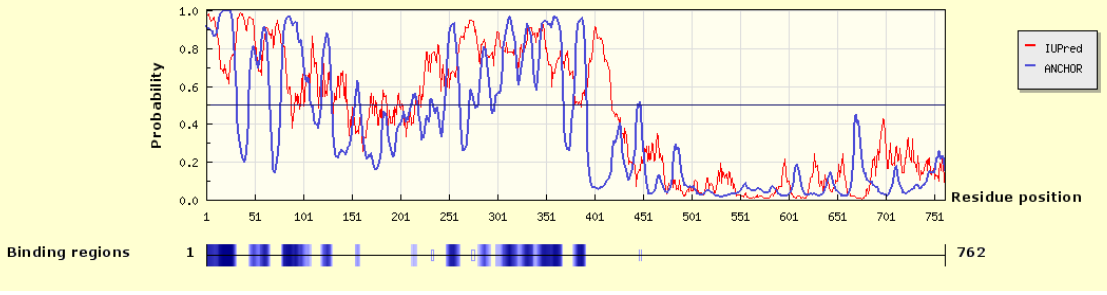




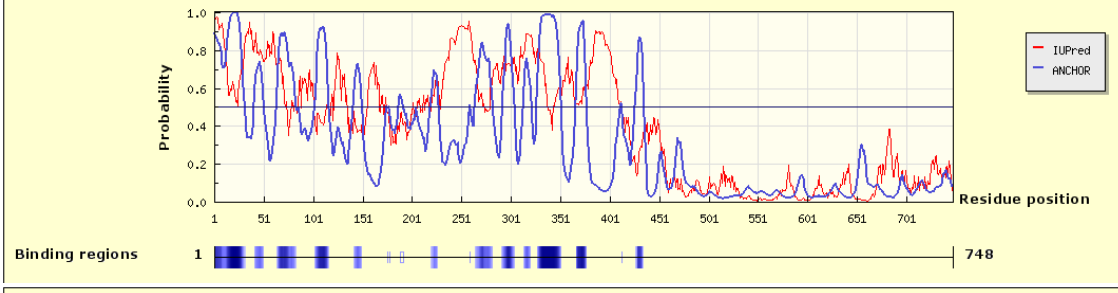
Citrus clementina



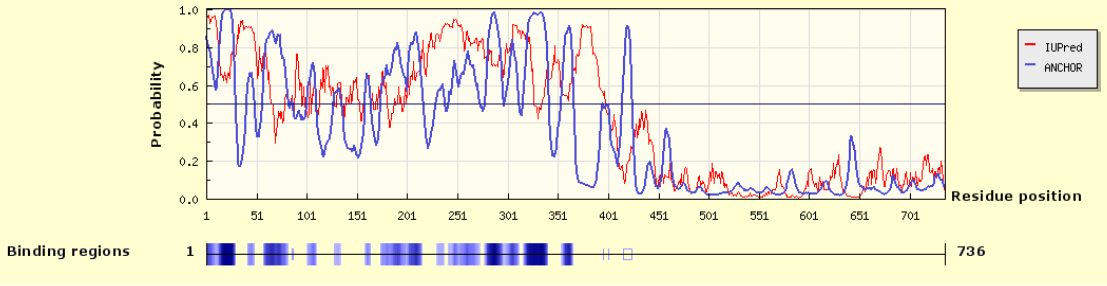
Manihot esculenta



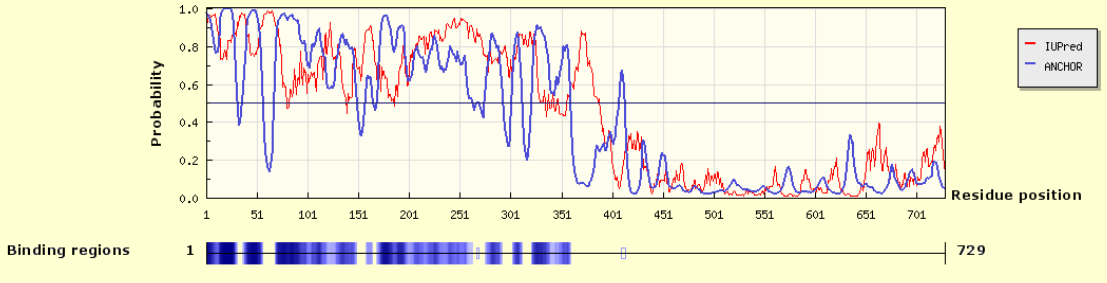
Jatropha curcas



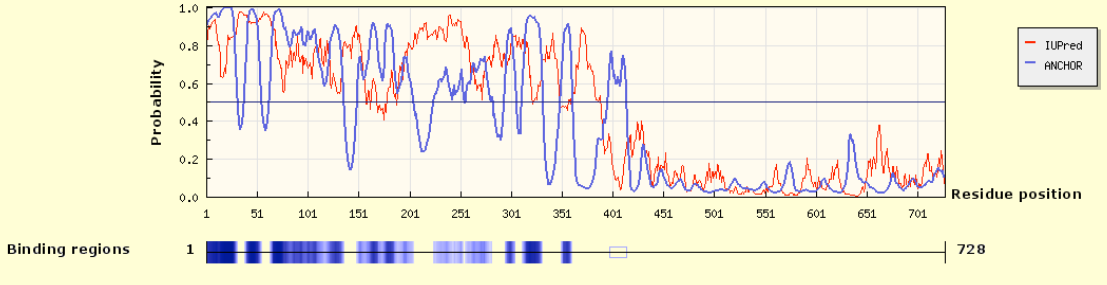
Ricinus communis



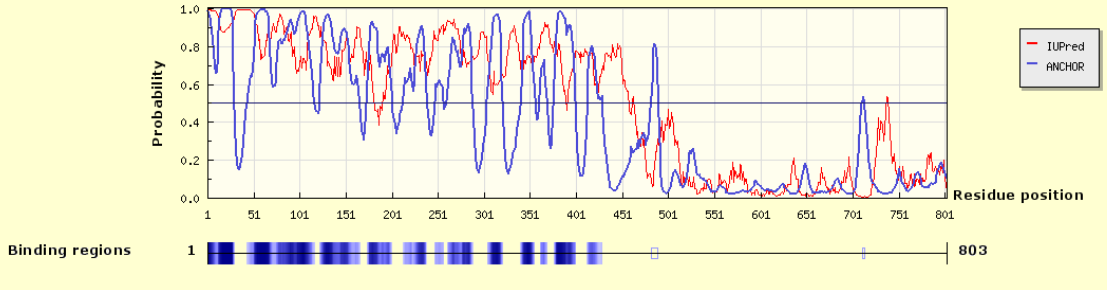
Populus trichocarpa A-2



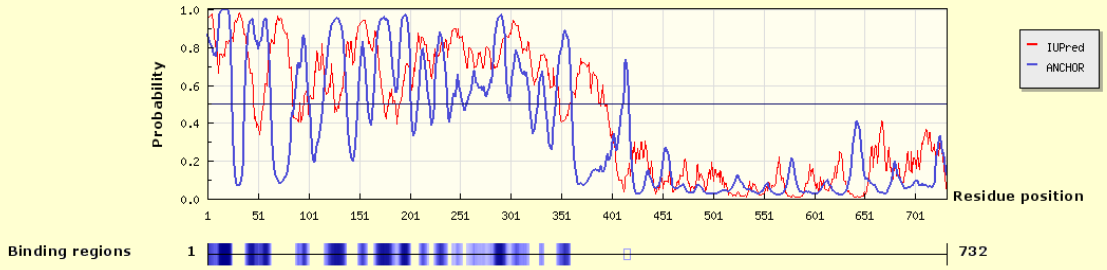
Populus trichocarpa A-1



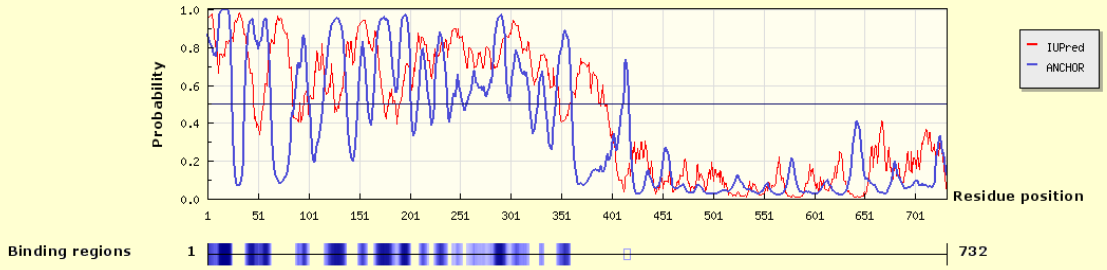
Eucalyptus grandis



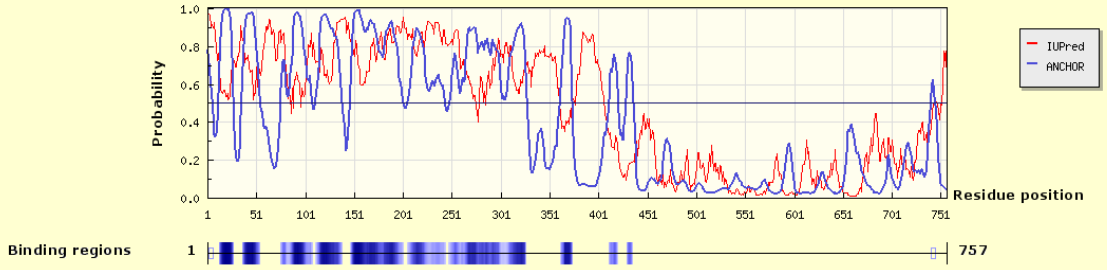
Glycine max A-2



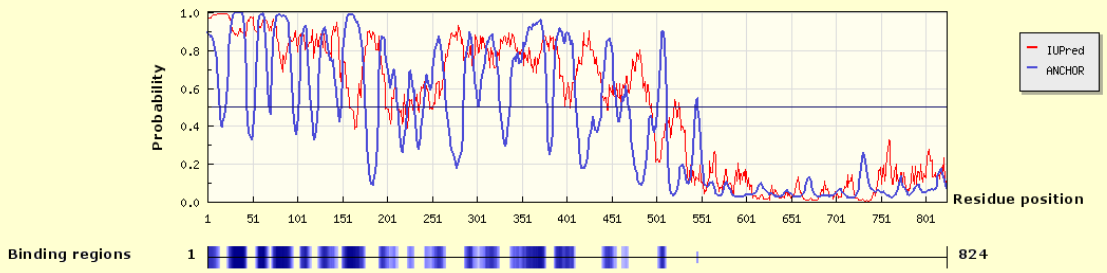
Glycine max A-1



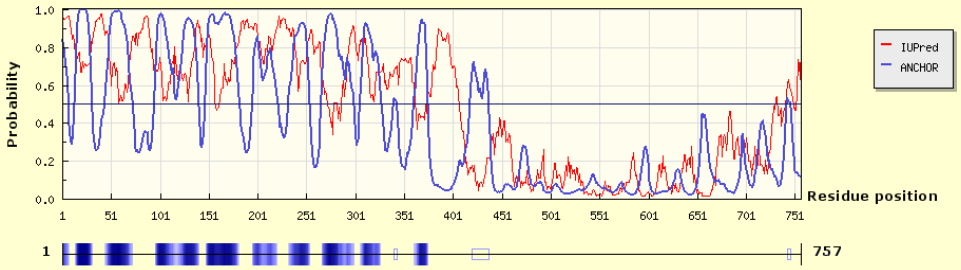
Arabidopsis lyrata



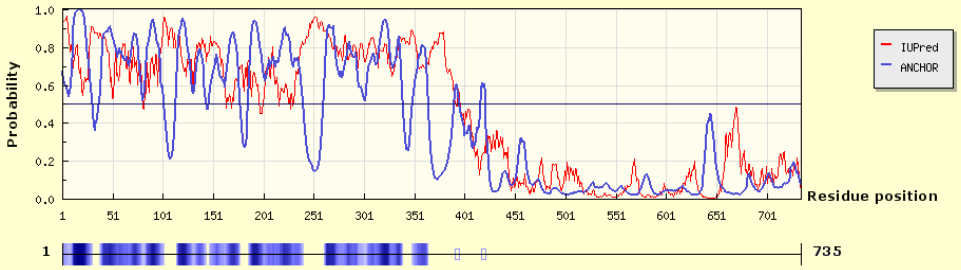
Aquilegia coerulea



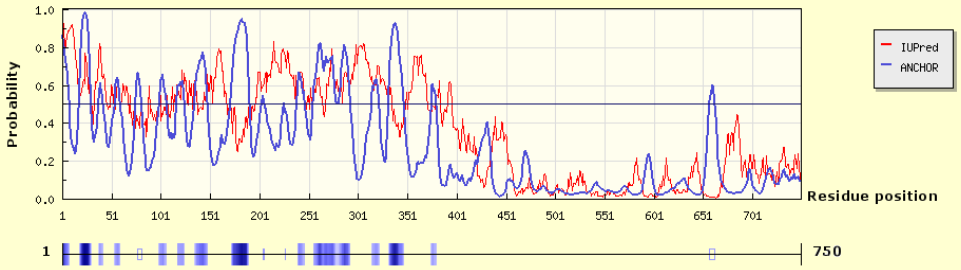
Thellungiella halophila



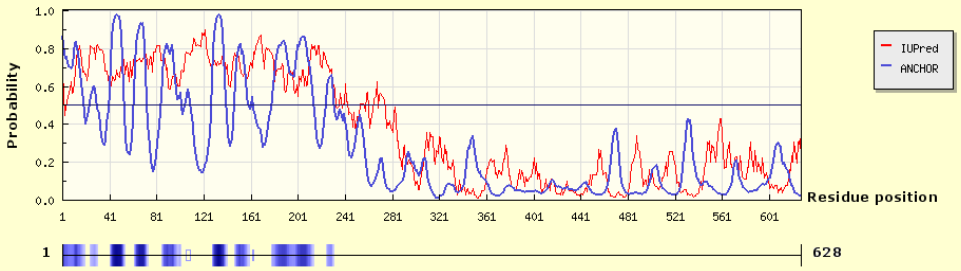
Mimulus guttatus



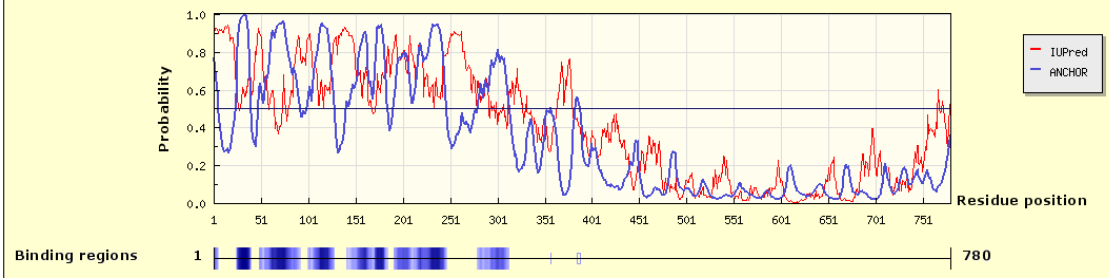
Cucumis sativus



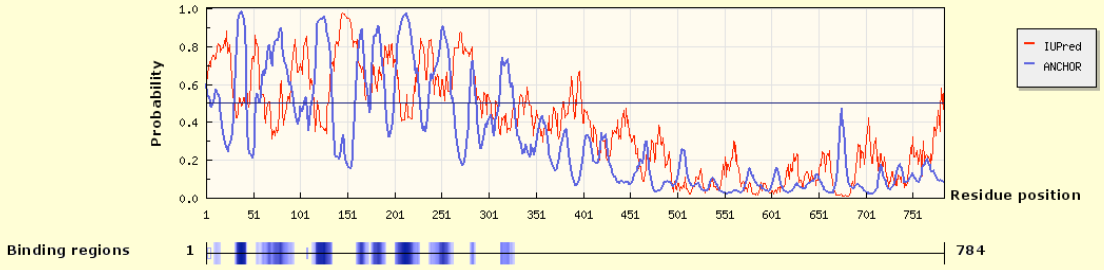
Selaginella moellendorffii



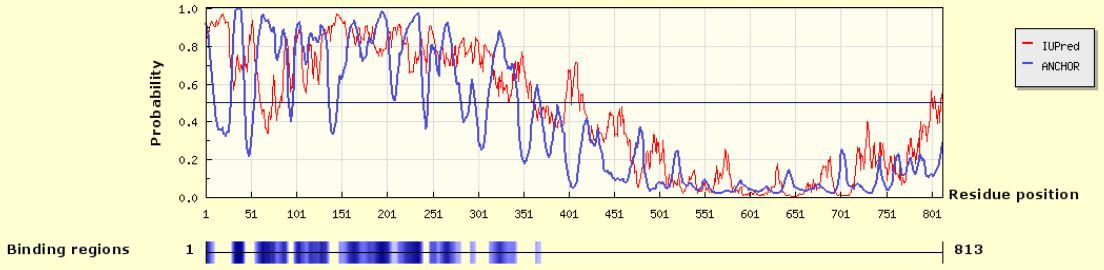
Zea mays A



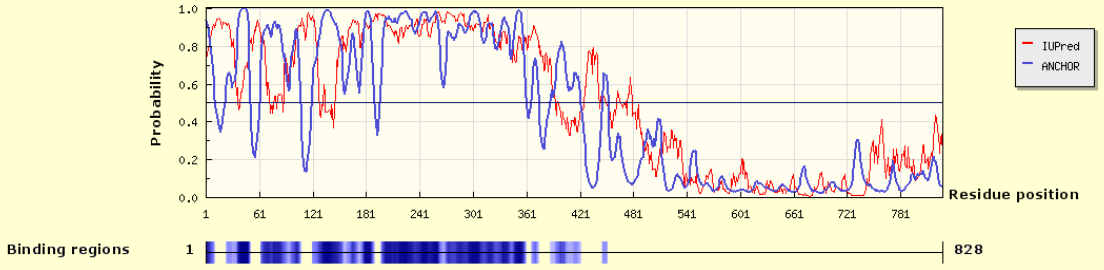
Sorghum bicolor A



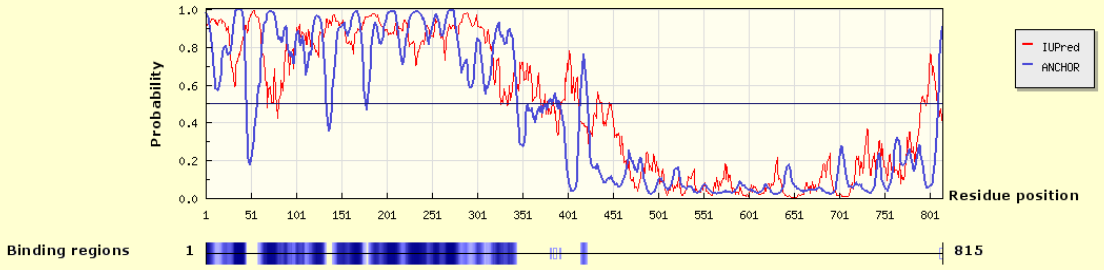
Setaria italica A



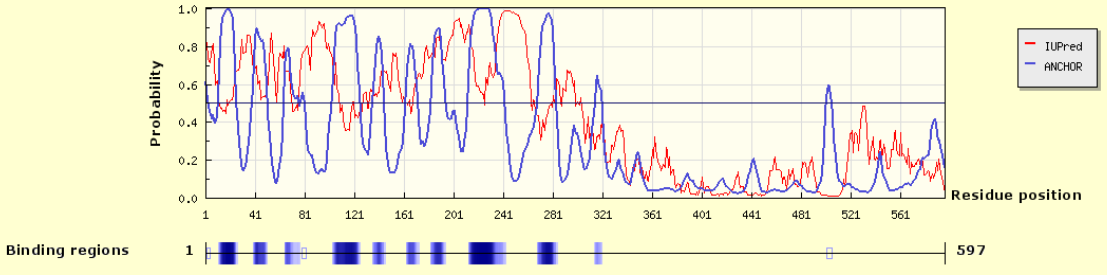
Oryza sativa j. A

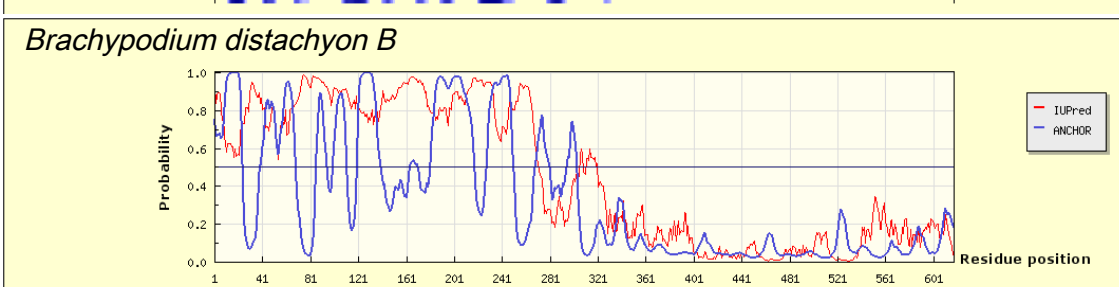


Brachypodium distachyon A



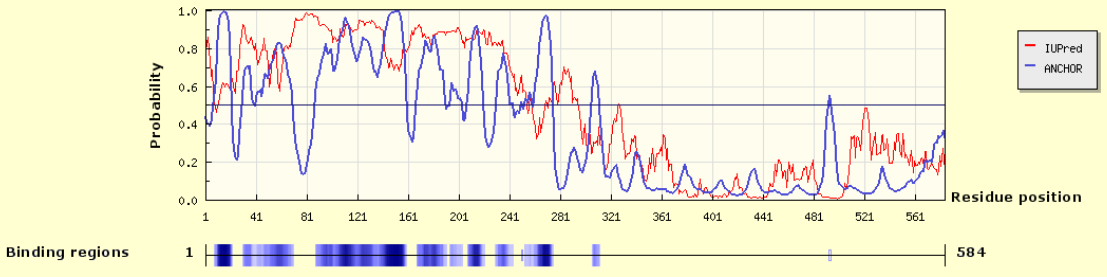
Oryza sativa j. B

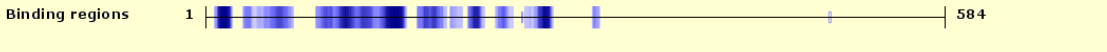




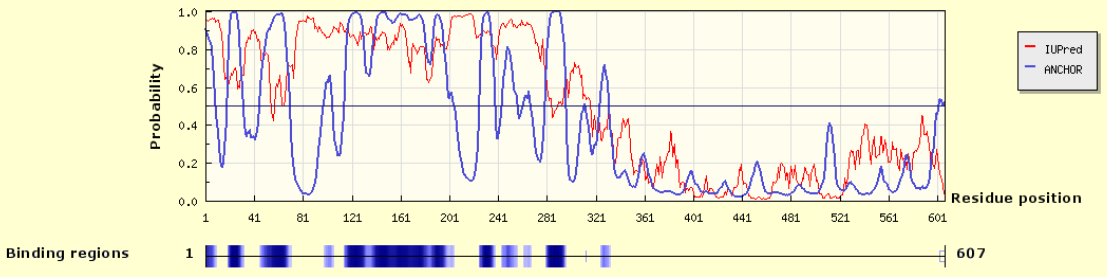
Brachypodium distachyon B

Setaria italica B

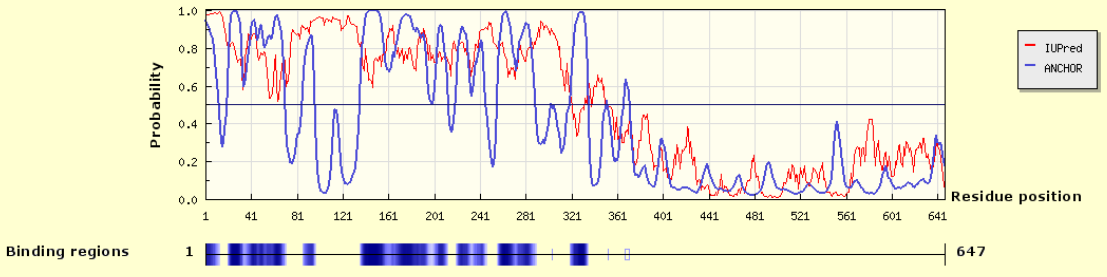


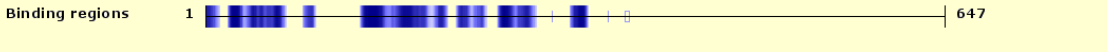


Zea mays B



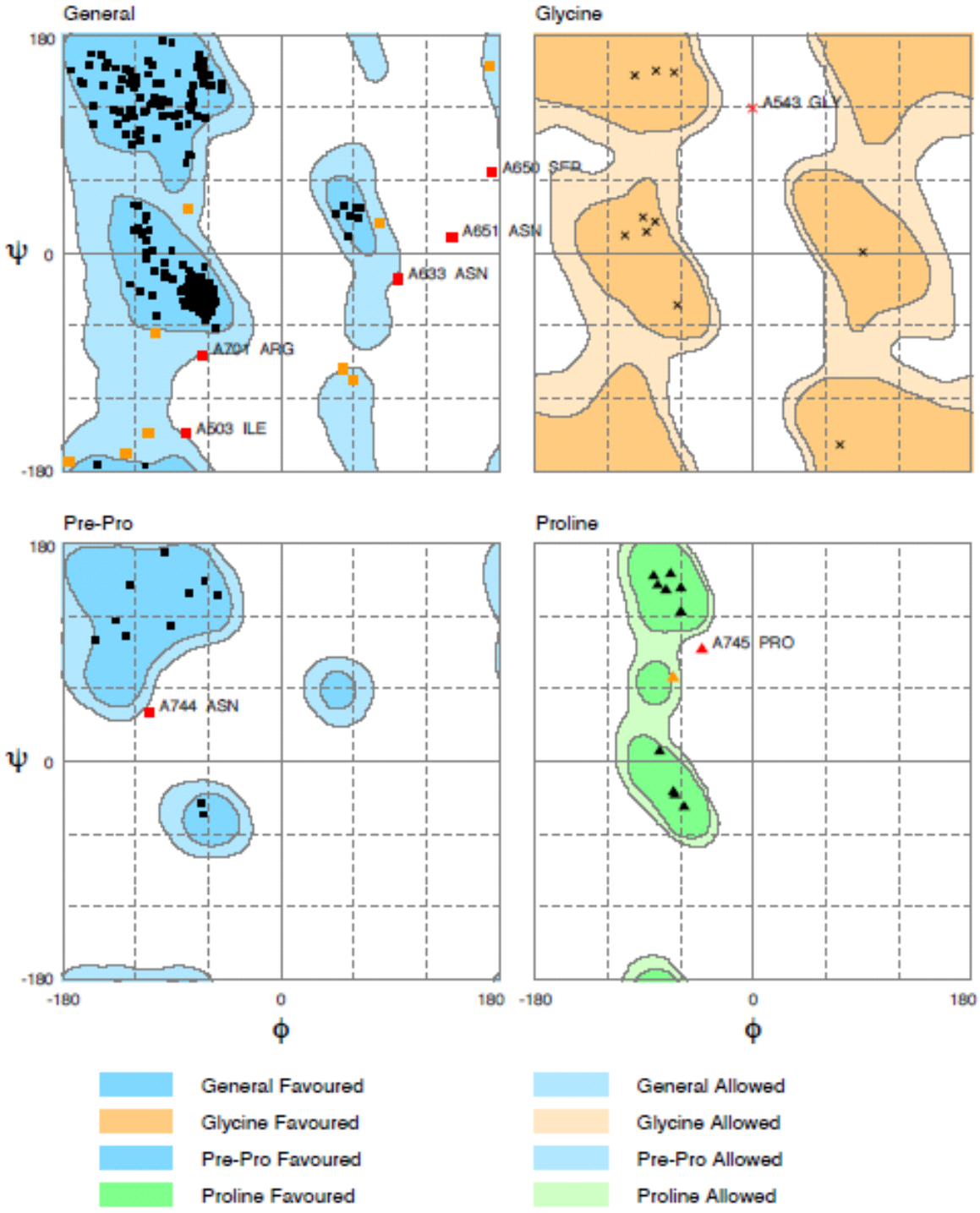
Sorghum bicolor B





Annexe 2.1

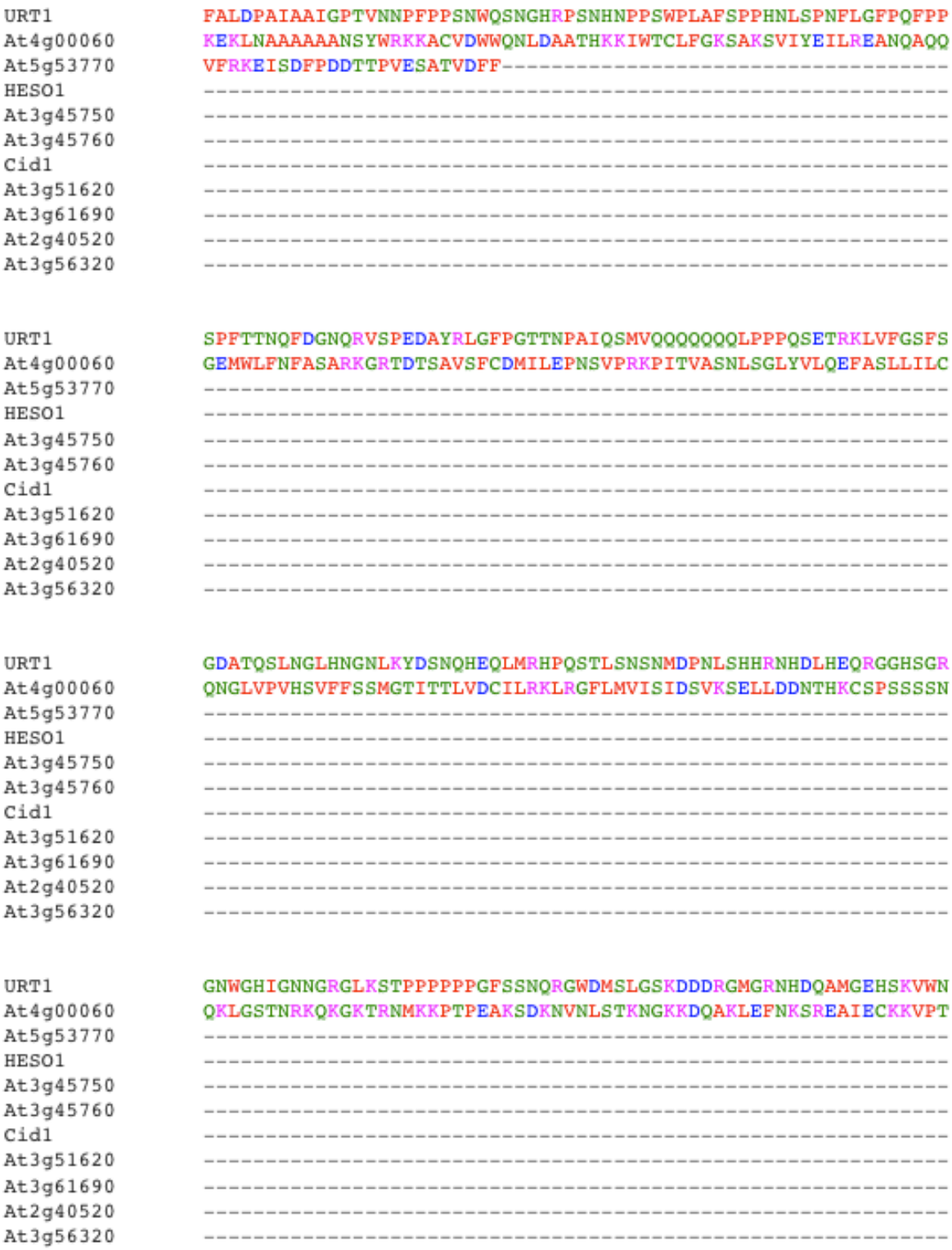
Diagramme de Ramachandran de la modélisation de URT1 détaillée (complément de la Figure 2B.1)

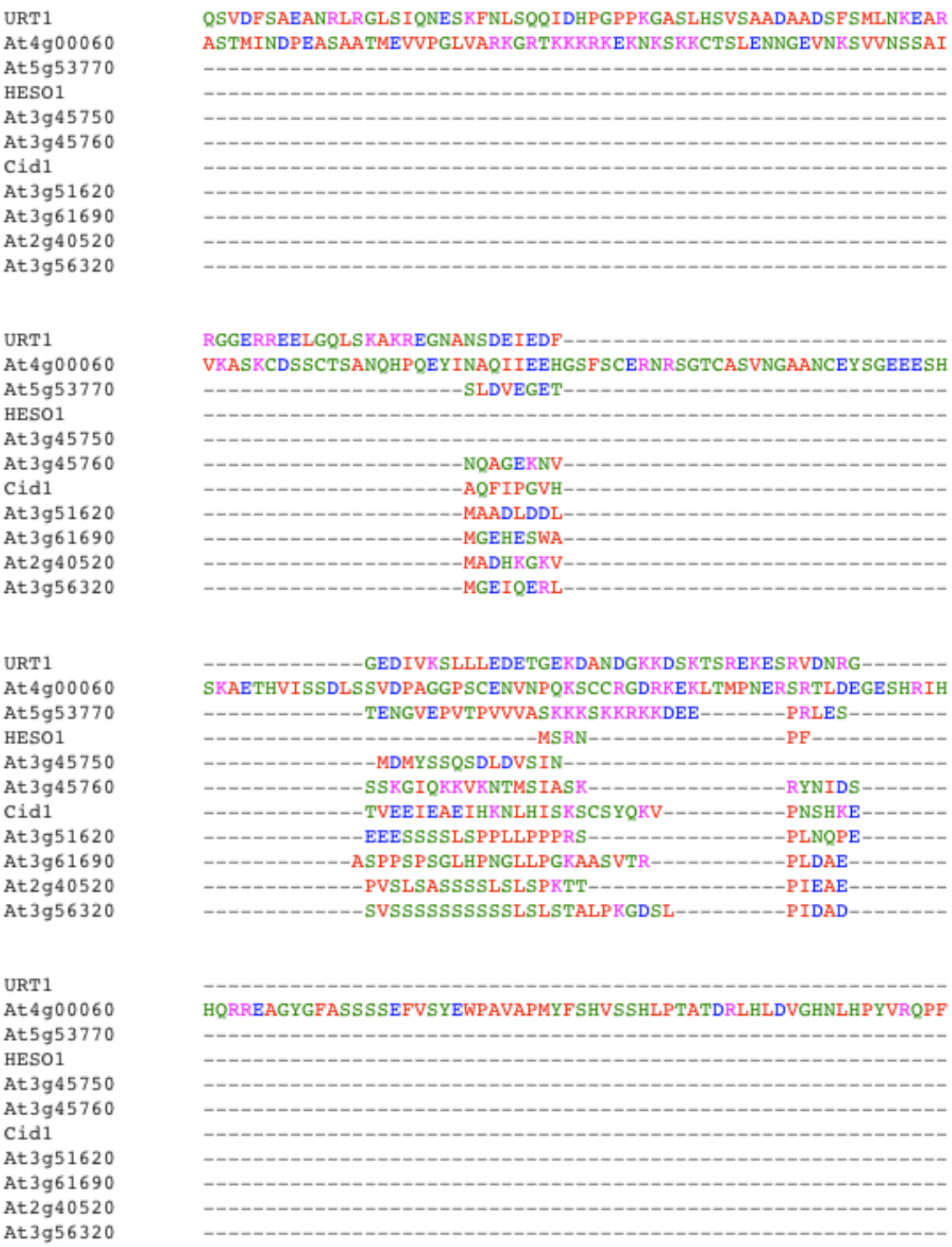


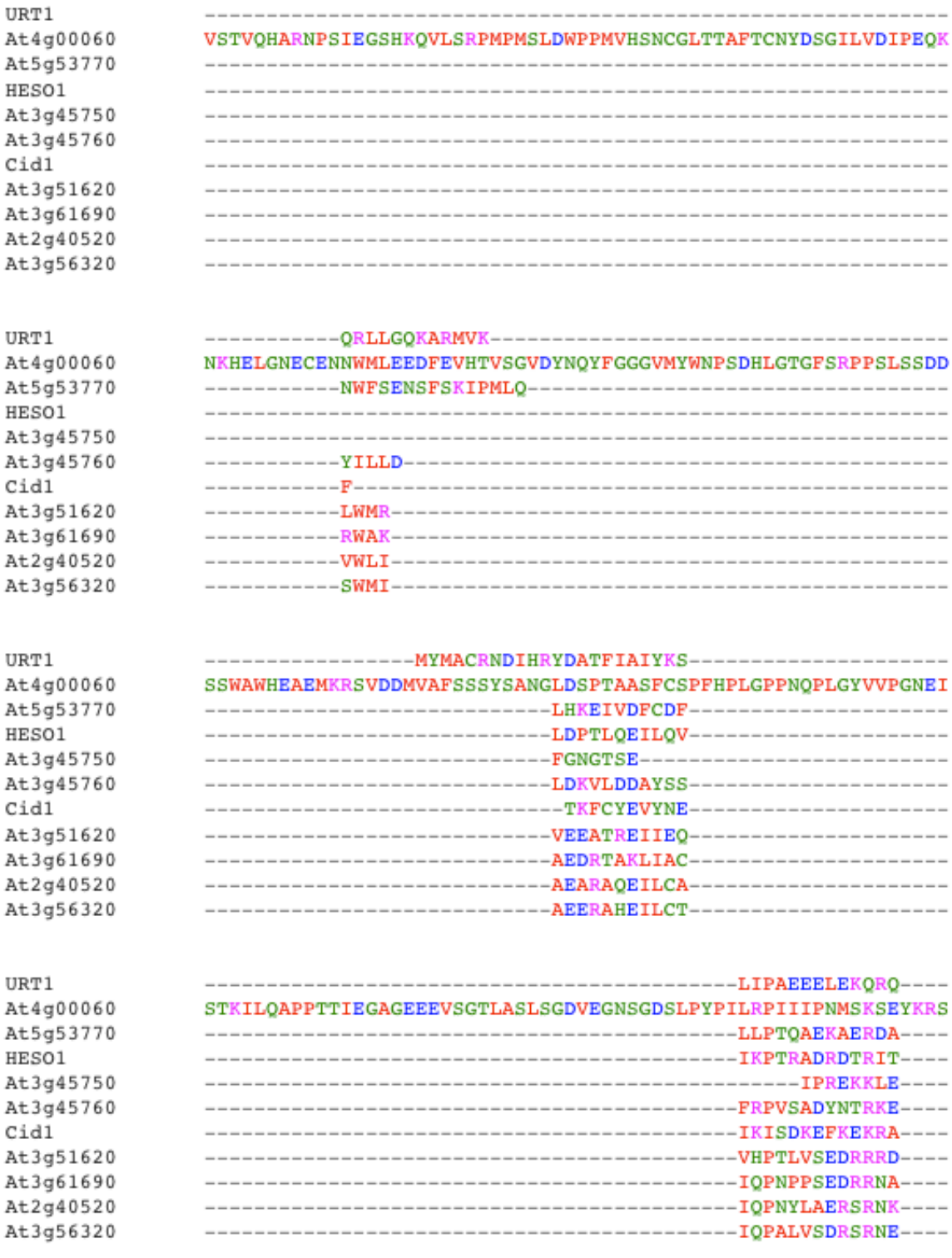
Annexe D1

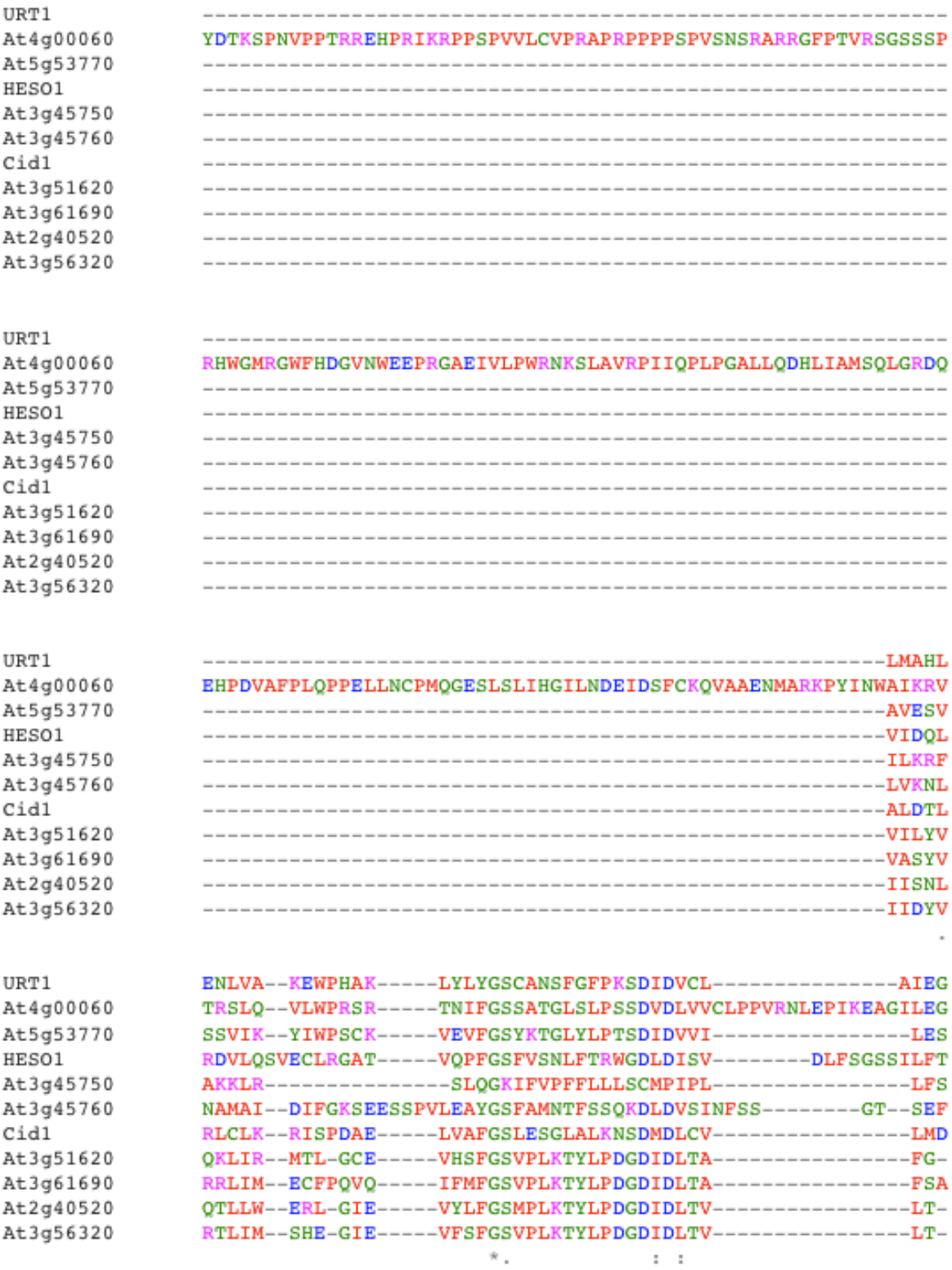
Alignement des ncPAP d'*A. thaliana* par MUSCLE

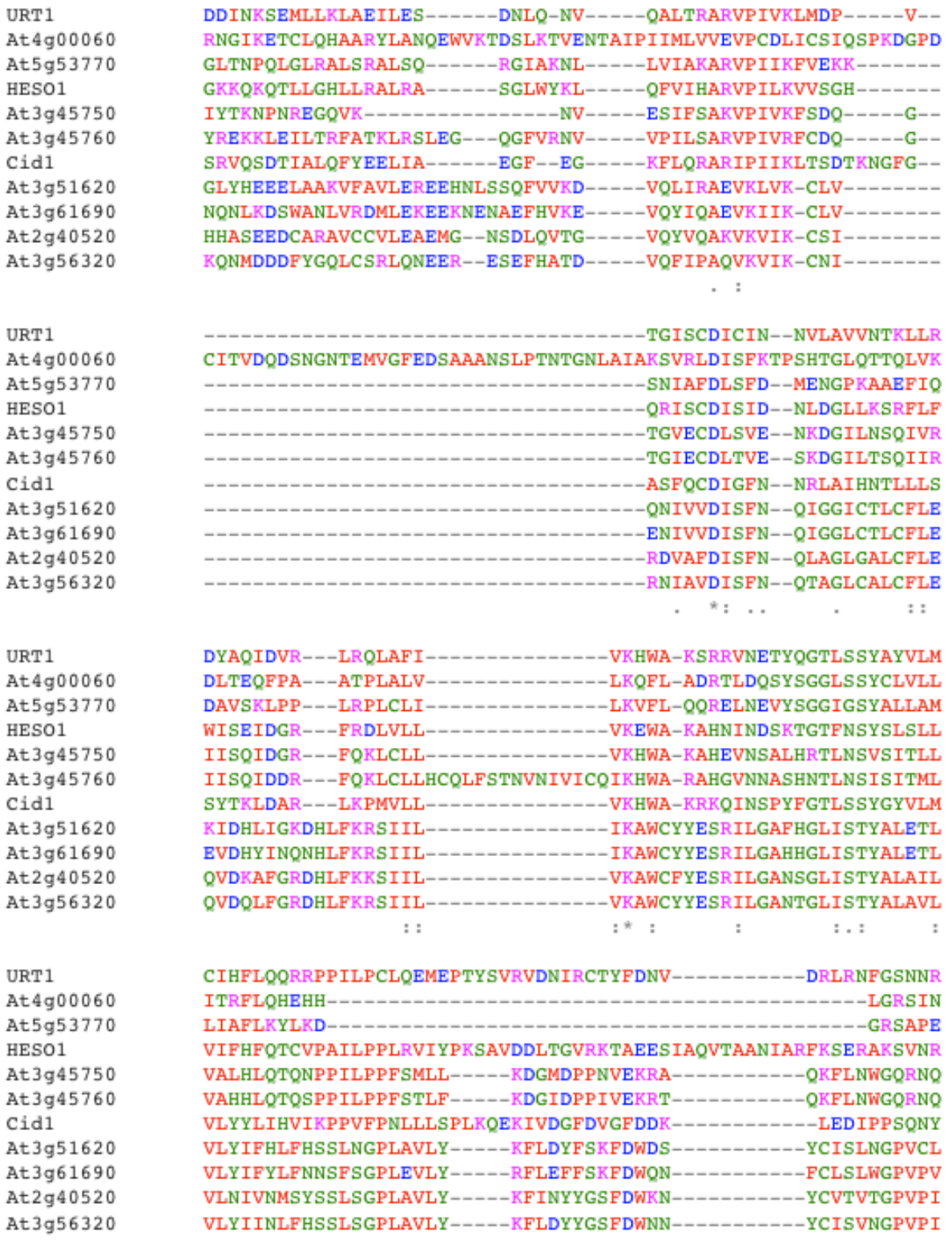
```
URT1          ------------------------------------------------------------
At4g00060     MENRHLPSKFFSSASSMAQNQLIDSLTSHISLYHSHSSSSSMANTIPNPRSAILRWFSSL
At5g53770     ------------------------------------------------------------
HESO1         ------------------------------------------------------------
At3g45750     ------------------------------------------------------------
At3g45760     ------------------------------------------------------------
Cid1          ------------------------------------------------------------
At3g51620     ------------------------------------------------------------
At3g61690     ------------------------------------------------------------
At2g40520     ------------------------------------------------------------
At3g56320     ------------------------------------------------------------

URT1          ------------------------------------------------------------
At4g00060     SVHQRLSHLTVVDPKFVQILLQMLGYIRTKGPCSFIILPDLPSSSDLPSLCFKKSRGLIS
At5g53770     ------------------------------------------------------------
HESO1         ------------------------------------------------------------
At3g45750     ------------------------------------------------------------
At3g45760     ------------------------------------------------------------
Cid1          ------------------------------------------------------------
At3g51620     ------------------------------------------------------------
At3g61690     ------------------------------------------------------------
At2g40520     ------------------------------------------------------------
At3g56320     ------------------------------------------------------------

URT1          ------------------------------------------------------------
At4g00060     RVSESNESERFVFDSTRLFGSGEGERAQDCSCSVNSLDSVVMAEEFLTNVDRFVETMDVL
At5g53770     ------------------------------------------------------------
HESO1         ------------------------------------------------------------
At3g45750     ------------------------------------------------------------
At3g45760     ------------------------------------------------------------
Cid1          ------------------------------------------------------------
At3g51620     ------------------------------------------------------------
At3g61690     ------------------------------------------------------------
At2g40520     ------------------------------------------------------------
At3g56320     ------------------------------------------------------------

URT1          ------------------MADGGAEPPAPPSSINAGEFLLSILHGSPSPSSQGPQHHQS
At4g00060     SDGAFLRGEESDLGSNWVELEWLKAKGYYSMEAFVANRLEVSMRLAWLNTNSGKRRGIKL
At5g53770     ------------------MEDSDAKTEAPAFVYDTLPPLSFSDSNQSPPPTHEESHQYS
HESO1         ------------------------------------------------------------
At3g45750     ------------------------------------------------------------
At3g45760     ------------------MDFRK-------------------------------------
Cid1          ------------------MNISS-------------------------------------
At3g51620     ------------------------------------------------------------
At3g61690     ------------------------------------------------------------
At2g40520     ------------------------------------------------------------
At3g56320     ------------------------------------------------------------
```

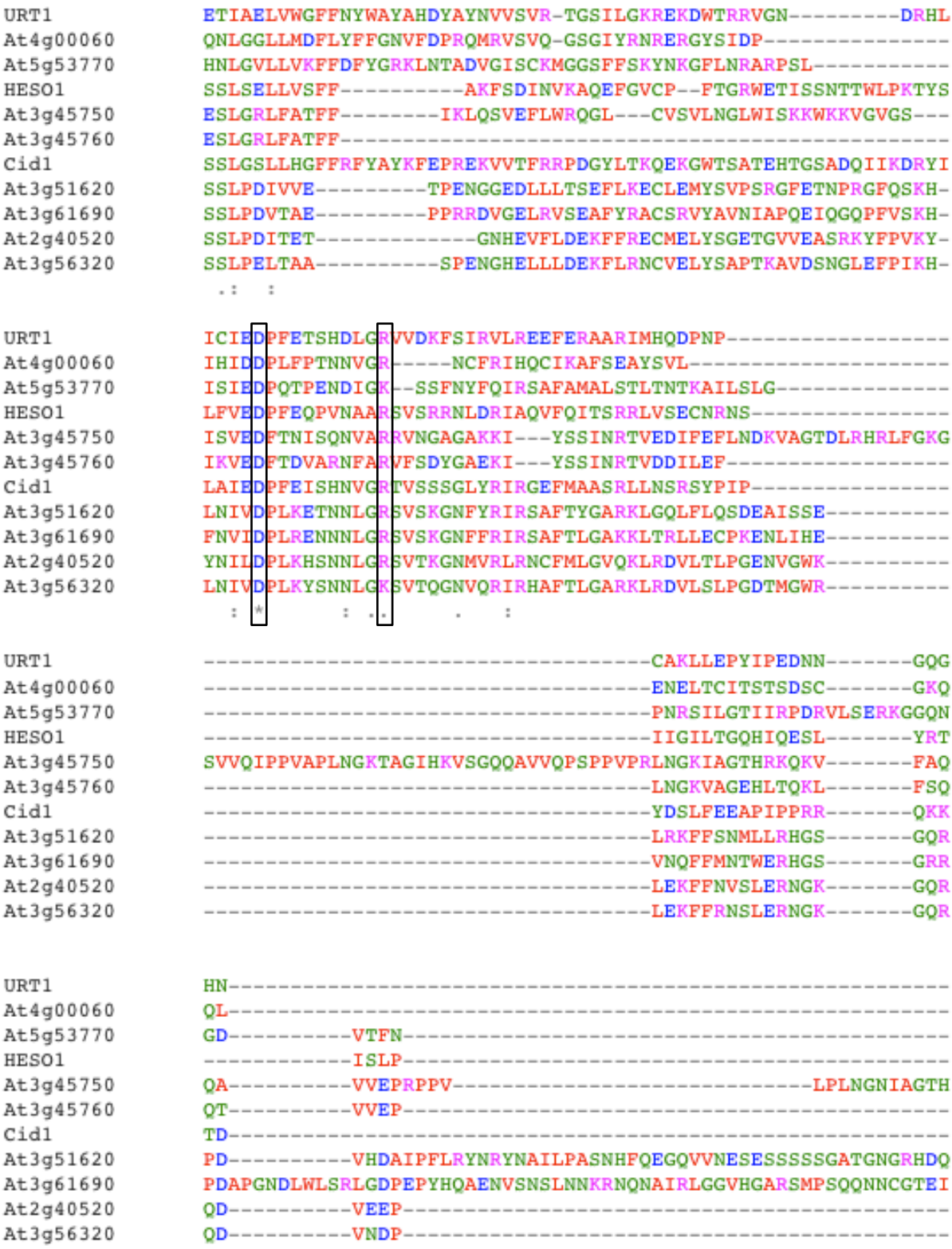

URT1 FALDPAIAAIGPTVNNPFPPSNWQSNGHRPSNHNPPSWPLAFSPPHNLSPNFLGFPQFPP
At4g00060 KEKLNAAAAAANSYWRKKACVDWWQNLDAATHKKIWTCLFGKSAKSVIYEILREANQAQQ
At5g53770 VFRKEISDFPDDTTPVESATVDFF
HESO1
At3g45750
At3g45760
Cid1
At3g51620
At3g61690
At2g40520
At3g56320
URT1 SPFTTNQFDGNQRVSPEDAYRLGFPGTTNPAIQSMVQQQQQQQLPPPQSETRKLVFGSFS
At4g00060 GEMWLFNFASARKGRTDTSAVSFCDMILEPNSVPRKPITVASNLSGLYVLQEFASLLILC
At5g53770
HESO1
At3g45750
At3g45760
Cid1
At3g51620
At3g61690
At2g40520
At3g56320
URT1 GDATQSLNGLHNGNLKYDSNQHEQLMRHPQSTLSNSNMDPNLSHHRNHDLHEQRGGHSGR
At4g00060 QNGLVPVHSVFFSSMGTITTLVDCILRKLRGFLMVISIDSVKSELLDDNTHKCSPSSSSN
At5g53770
HESO1
At3g45750
At3g45760
Cid1
At3g51620
At3g61690
At2g40520
At3g56320
URT1 GNWGHIGNNGRGLKSTPPPPPPGFSSNQRGWDMSLGSKDDDRGMGRNHDQAMGEHSKVWN
At4g00060 QKLGSTNRKQKGKTRNMKKPTPEAKSDKNVNLSTKNGKKDQAKLEFNKSREAIECKKVPT
At5g53770
HESO1
At3g45750
At3g45760
Cid1
At3g51620
At3g61690
At2g40520
At3g56320

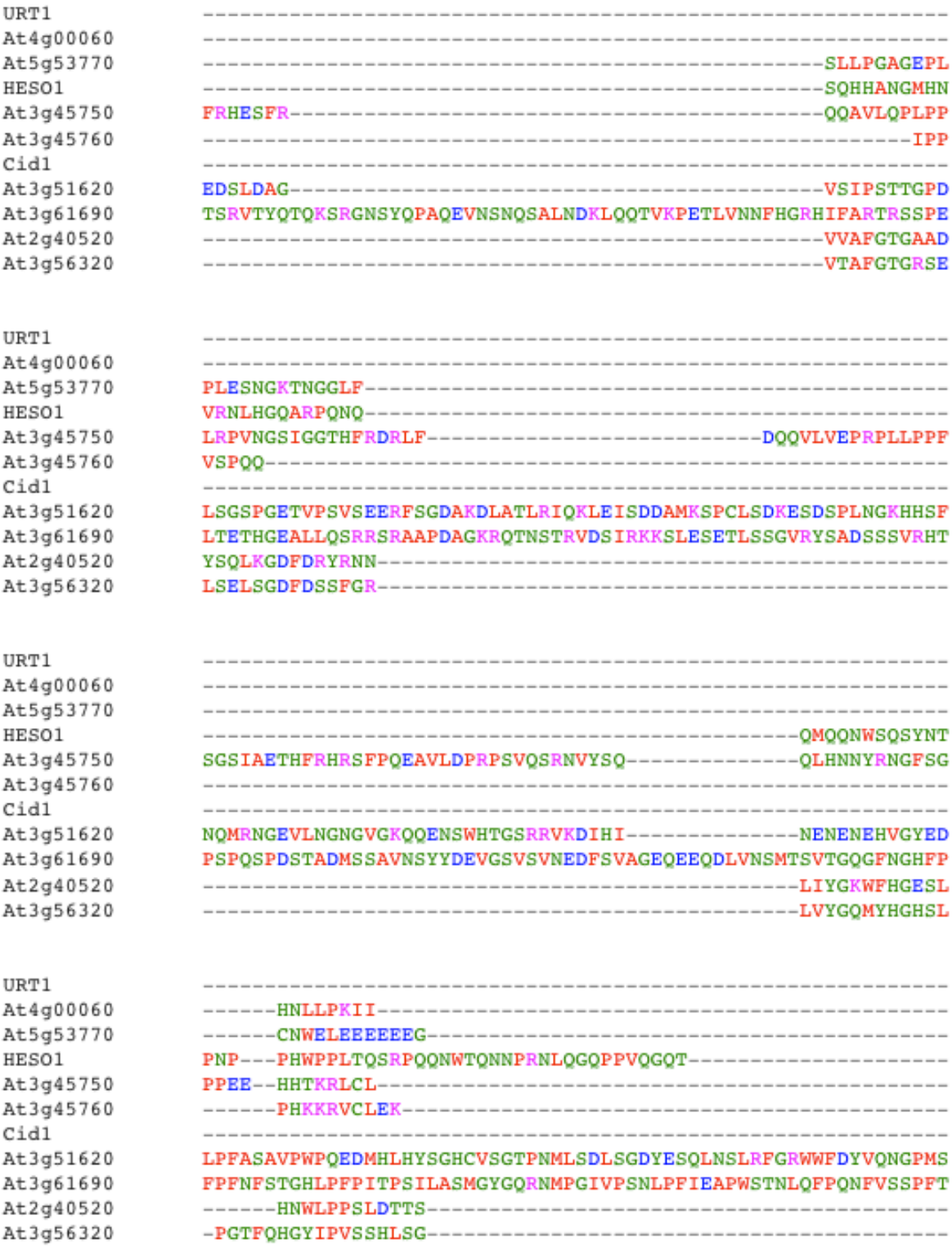

URT1 QSVDFSAEANRLRGLSIQNESKFNLSQQIDHPGPPKGASLHSVSAADAADSFSMLNKEAR
At4g00060 ASTMINDPEASAATMEVVPGLVARKGRTKKKRKEKNKSKKCTSLENNGEVNKSVVNSSAI
At5g53770
HESO1
At3g45750
At3g45760
Cid1
At3g51620
At3g61690
At2g40520
At3g56320
URT1 RGGERREELGQLSKAKREGNANSDEIEDF
At4g00060 VKASKCDSSCTSANQHPQEYINAQIIEEHGSFSCERNRSGTCASVNGAANCEYSGEEESH
At5g53770 SLDVEGET
HESO1
At3g45750
At3g45760 NQAGEKNV
Cid1 AQFIPGVH
At3g51620 MAADLDDL
At3g61690 MGEHESWA
At2g40520 MADHKGKV
At3g56320 MGEIQERL
URT1 GEDIVKSLLLEDETGEKDANDGKKDSKTSREKESRVDNRG
At4g00060 SKAETHVISSDLSSVDPAGGPSCENVNPQKSCCRGDRKEKLTMPNERSRTLDEGESHRIH
At5g53770 TENGVEPVTPVVVASKKKSKKRKKDEE PRLES
HESO1 MSRN PF
At3g45750 MDMYSSQSDLDVSIN
At3g45760 SSKGIQKKVKNTMSIASK RYNIDS
Cid1 TVEEIEAEIHKNLHISKSCSYQKV PNSHKE
At3g51620 EEESSSSLSPPLLPPPRS PLNQPE
At3g61690 ASPPSPSGLHPNGLLPGKAASVTR PLDAE
At2g40520 PVSLSASSSSLSLSPKTT PIEAE
At3g56320 SVSSSSSSSSSSLSLSTALPKGDSL PIDAD
URT1
At4g00060 HQRREAGYGFASSSSEFVSYEWPAVAPMYFSHVSSHLPTATDRLHLDVGHNLHPYVRQPF
At5g53770
HESO1
At3g45750
At3g45760
Cid1
At3g51620
At3g61690
At2g40520
At3g56320

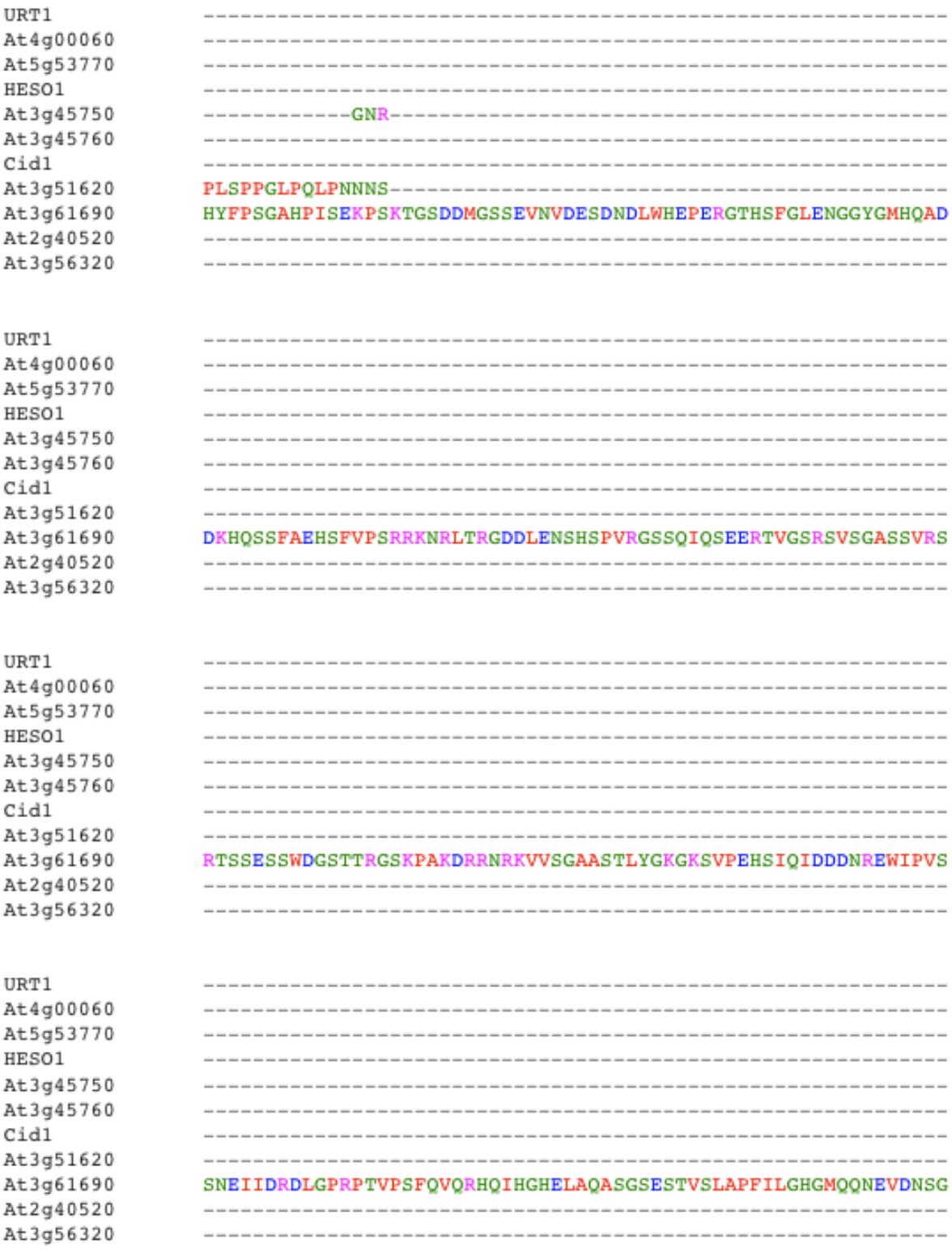

URT1
At4g00060
At5g53770
HESO1
At3g45750
At3g45760
Cid1
At3g51620
At3g61690
At2g40520
At3g56320

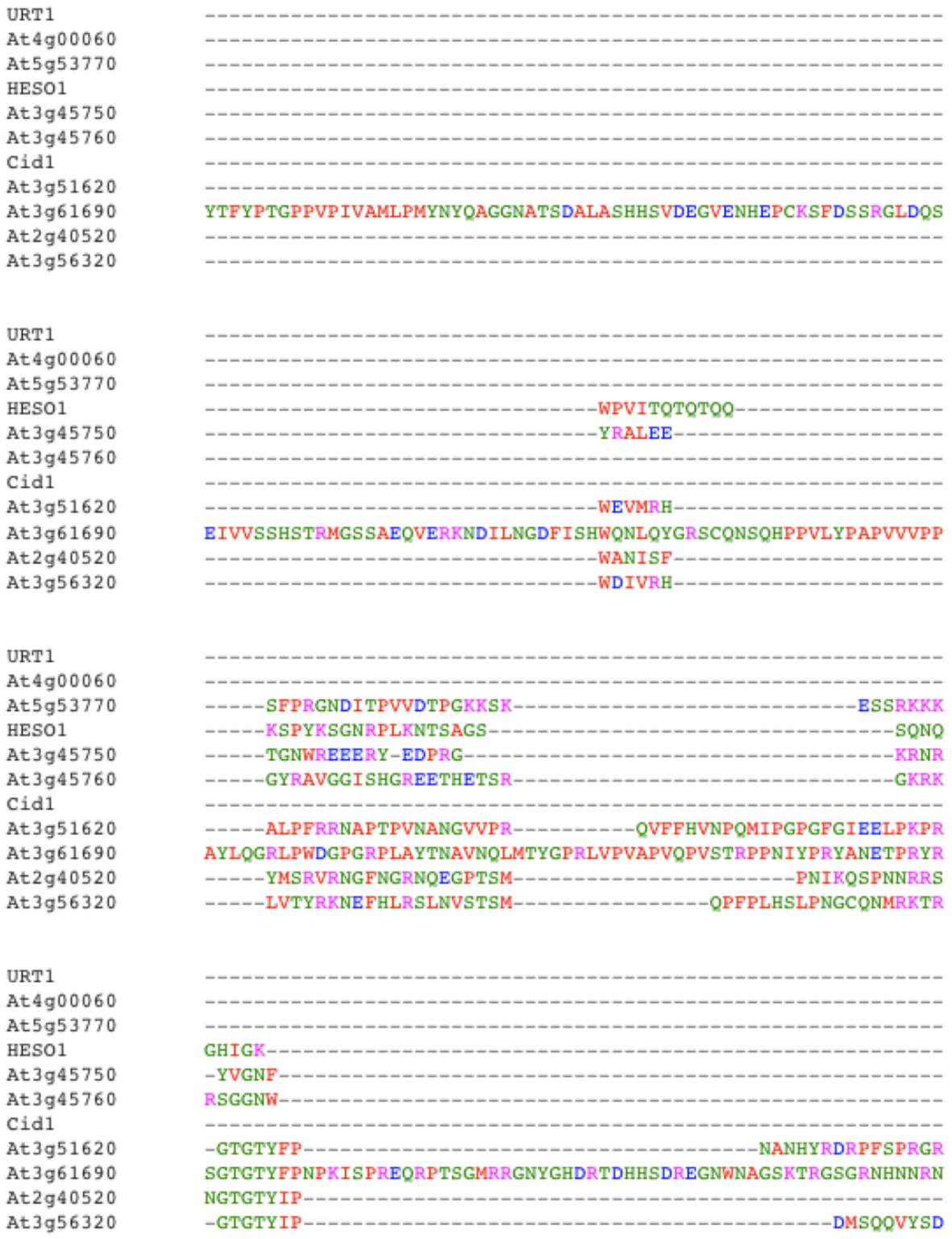

URT1
At4g00060 YDTKSPNVPPTRREHPRIKRPPSPVVLCVPRAPRPPPPSPVSNSRARRGFPTVRSGSSSP
At5g53770
HESO1
At3g45750
At3g45760
Cid1
At3g51620
At3g61690
At2g40520
At3g56320
URT1
At4g00060 RHWGMRGWFHDGVNWEEPRGAEIVLPWRNKSLAVRPIIQPLPGALLQDHLIAMSQLGRDQ
At5g53770
HESO1
At3g45750
At3g45760
Cid1
At3g51620
At3g61690
At2g40520
At3g56320
URT1 LMAHL
At4g00060 EHPDVAFPLQPPELLNCPMQGESLSLIHGILNDEIDSFCKQVAAENMARKPYINWAIKRV
At5g53770 AVESV
HESO1 VIDQL
At3g45750 ILKRF
At3g45760 LVKNL
Cid1 ALDTL
At3g51620 VILYV
At3g61690 VASYV
At2g40520 IISNL
At3g56320 IIDYV
URT1 ENLVA--KEWPHAK-----LYLYGSCANSFGFPKSDIDVCL---------------AIEG
At4g00060 TRSLQ--VLWPRSR-----TNIFGSSATGLSLPSSDVDLVVCLPPVRNLEPIKEAGILEG
At5g53770 SSVIK--YIWPSCK-----VEVFGSYKTGLYLPTSDIDVVI----------------LES
HESO1 RDVLQSVECLRGAT-----VQPFGSFVSNLFTRWGDLDISV--------DLFSGSSILFT
At3g45750 AKKLR----------------SLQGKIFVPFFLLLSCMPIPL----------------LFS
At3g45760 NAMAI--DIFGKSEESSPVLEAYGSFAMNTFSSQKDLDVSINFSS--------GT--SEF
Cid1 RLCLK--RISPDAE-----LVAFGSLESGLALKNSDMDLCV----------------LMD
At3g51620 QKLIR--MTL-GCE-----VHSFGSVPLKTYLPDGDIDLTA----------------FG-
At3g61690 RRLIM--ECFPQVQ-----IFMFGSVPLKTYLPDGDIDLTA----------------FSA
At2g40520 QTLLW--ERL-GIE-----VYLFGSMPLKTYLPDGDIDLTV----------------LT-
At3g56320 RTLIM--SHE-GIE-----VFSFGSVPLKTYLPDGDIDLTV----------------LT-

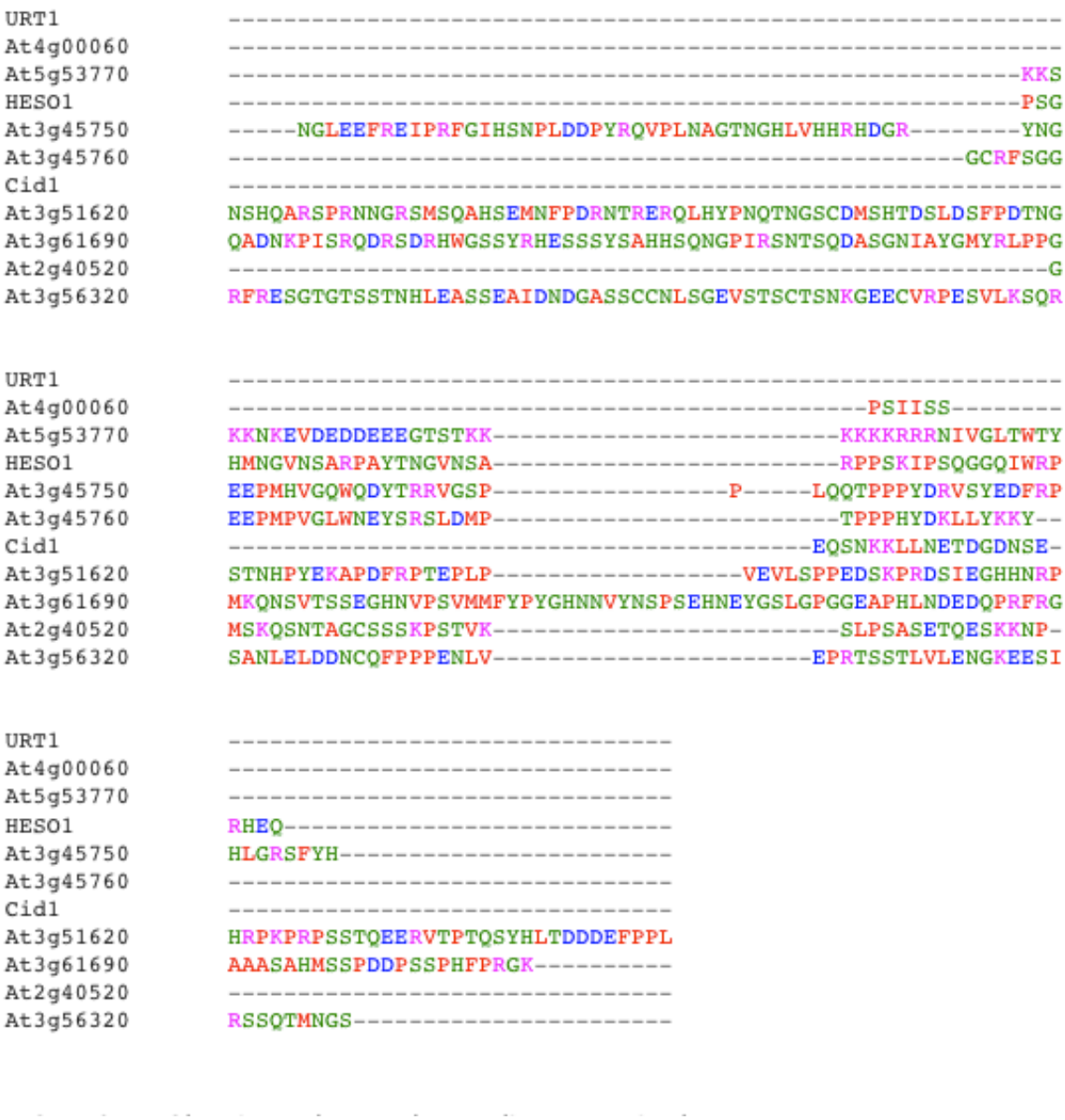

URT1 DDINKSEMLLKLAEILES------DNLQ-NV-----QALTRARVPIVKLMDP-----V--
At4g00060 RNGIKETCLQHAARYLANQEWVKTDSLKTVENTAIPIIMLVVEVPCDLICSIQSPKDGPD
At5g53770 GLTNPQLGLRALSRALSQ------RGIAKNL-----LVIAKARVPIIKFVEKK-------
HESO1 GKKQKQTLLGHLLRALRA------SGLWYKL-----QFVIHARVPILKVVSGH-------
At3g45750 IYTKNPNREGQVK----------------NV-----ESIFSAKVPIVKFSDQ-----G--
At3g45760 YREKKLEILTRFATKLRSLEG---QGFVRNV-----VPILSARVPIVRFCDQ-----G--
Cid1 SRVQSDTIALQFYEELIA------EGF--EG-----KFLQRARIPIIKLTSDTKNGFG--
At3g51620 GLYHEEELAAKVFAVLEREEHNLSSQFVVKD-----VQLIRAEVKLVK-CLV--------
At3g61690 NQNLKDSWANLVRDMLEKEEKNENAEFHVKE-----VQYIQAEVKIIK-CLV--------
At2g40520 HHASEEDCARAVCCVLEAEMG--NSDLQVTG-----VQYVQAKVKVIK-CSI--------
At3g56320 KQNMDDDFYGQLCSRLQNEER--ESEFHATD-----VQFIPAQVKVIK-CNI--------
URT1 ----------TGISCDICIN--NVLAVVNTKLLR
At4g00060 CITVDQDSNGNTEMVGFEDSAAANSLPTNTGNLAIAKSVRLDISFKTPSHTGLQTTQLVK
At5g53770 ----------SNIAFDLSFD--MENGPKAAEFIQ
HESO1 ----------QRISCDISID--NLDGLLKSRFLF
At3g45750 ----------TGVECDLSVE--NKDGILNSQIVR
At3g45760 ----------TGIECDLTVE--SKDGILTSQIIR
Cid1 ----------ASFQCDIGFN--NRLAIHNTLLLS
At3g51620 ----------QNIVVDISFN--QIGGICTLCFLE
At3g61690 ----------ENIVVDISFN--QIGGLCTLCFLE
At2g40520 ----------RDVAFDISFN--QLAGLGALCFLE
At3g56320 ----------RNIAVDISFN--QTAGLCALCFLE
URT1 DYAQIDVR---LRQLAFI----------VKHWA-KSRRVNETYQGTLSSYAYVLM
At4g00060 DLTEQFPA---ATPLALV----------LKQFL-ADRTLDQSYSGGLSSYCLVLL
At5g53770 DAVSKLPP---LRPLCLI----------LKVFL-QQRELNEVYSGGIGSYALLAM
HESO1 WISEIDGR---FRDLVLL----------VKEWA-KAHNINDSKTGTFNSYSLSLL
At3g45750 IISQIDGR---FQKLCLL----------VKHWA-KAHEVNSALHRTLNSVSITLL
At3g45760 IISQIDDR---FQKLCLLHCQLFSTNVNIVICQIKHWA-RAHGVNNASHNTLNSISITML
Cid1 SYTKLDAR---LKPMVLL----------VKHWA-KRKQINSPYFGTLSSYGYVLM
At3g51620 KIDHLIGKDHLFKRSIIL----------IKAWCYYESRILGAFHGLISTYALETL
At3g61690 EVDHYINQNHLFKRSIIL----------IKAWCYYESRILGAHHGLISTYALETL
At2g40520 QVDKAFGRDHLFKKSIIL----------VKAWCFYESRILGANSGLISTYALAIL
At3g56320 QVDQLFGRDHLFKRSIIL----------VKAWCYYESRILGANTGLISTYALAVL
URT1 CIHFLQQRRPPILPCLQEMEPTYSVRVDNIRCTYFDNV----------DRLRNFGSNNR
At4g00060 ITRFLQHEHH----------LGRSIN
At5g53770 LIAFLKYLKD----------GRSAPE
HESO1 VIFHFQTCVPAILPPLRVIYPKSAVDDLTGVRKTAEESIAQVTAANIARFKSERAKSVNR
At3g45750 VALHLQTQNPPILPPFSMLL-----KDGMDPPNVEKRA----------QKFLNWGQRNQ
At3g45760 VAHHLQTQSPPILPPFSTLF-----KDGIDPPIVEKRT----------QKFLNWGQRNQ
Cid1 VLYYLIHVIKPPVFPNLLLSPLKQEKIVDGFDVGFDDK----------LEDIPPSQNY
At3g51620 VLYIFHLFHSSLNGPLAVLY-----KFLDYFSKFDWDS----------YCISLNGPVCL
At3g61690 VLYIFYLFNNSFSGPLEVLY-----RFLEFFSKFDWQN----------FCLSLWGPVPV
At2g40520 VLNIVNMSYSSLSGPLAVLY-----KFINYYGSFDWKN----------YCVTVTGPVPI
At3g56320 VLYIINLFHSSLSGPLAVLY-----KFLDYYGSFDWNN----------YCISVNGPVPI


URT1 ETIAELVWGFFNYWAYAHDYAYNVVSVR-TGSILGKREKDWTRRVGN---------DRHL
At4g00060 QNLGGLLMDFLYFFGNVFDPRQMRVSVQ-GSGIYRNRERGYSIDP----------
At5g53770 HNLGVLLVKFFDFYGRKLNTADVGISCKMGGSFFSKYNKGFLNRARPSL----------
HESO1 SSLSELLVSFF----------AKFSDINVKAQEFGVCP--FTGRWETISSNTTWLPKTYS
At3g45750 ESLGRLFATFF--------IKLQSVEFLWRQGL---CVSVLNGLWISKKWKKVGVGS---
At3g45760 ESLGRLFATFF----------
Cid1 SSLGSLLHGFFRFYAYKFEPREKVVTFRRPDGYLTKQEKGWTSATEHTGSADQIIKDRYI
At3g51620 SSLPDIVVE---------TPENGGEDLLLTSEFLKECLEMYSVPSRGFETNPRGFQSKH-
At3g61690 SSLPDVTAE---------PPRRDVGELRVSEAFYRACSRVYAVNIAPQEIQGQPFVSKH-
At2g40520 SSLPDITET----------GNHEVFLDEKFFRECMELYSGETGVVEASRKYFPVKY-
At3g56320 SSLPELTAA----------SPENGHELLLDEKFLRNCVELYSAPTKAVDSNGLEFPIKH-
URT1 ICIEDPFETSHDLGRVVDKFSIRVLREEFERAARIMHQDPNP----------
At4g00060 IHIDDPLFPTNNVGR-----NCFRIHQCIKAFSEAYSVL----------
At5g53770 ISIEDPQTPENDIGK--SSFNYFQIRSAFAMALSTLTNTKAILSLG----------
HESO1 LFVEDPFEQPVNAARSVSRRNLDRIAQVFQITSRRLVSECNRNS----------
At3g45750 ISVEDFTNISQNVARRVNGAGAKKI---YSSINRTVEDIFEFLNDKVAGTDLRHRLFGKG
At3g45760 IKVEDFTDVARNFARVFSDYGAEKI---YSSINRTVDDILEF----------
Cid1 LAIEDPFEISHNVGRTVSSSGLYRIRGEFMAASRLLNSRSYPIP----------
At3g51620 LNIVDPLKETNNLGRSVSKGNFYRIRSAFTYGARKLGQLFLQSDEAISSE----------
At3g61690 FNVIDPLRENNNLGRSVSKGNFFRIRSAFTLGAKKLTRLLECPKENLIHE----------
At2g40520 YNILDPLKHSNNLGRSVTKGNMVRLRNCFMLGVQKLRDVLTLPGENVGWK----------
At3g56320 LNIVDPLKYSNNLGKSVTQGNVQRIRHAFTLGARKLRDVLSLPGDTMGWR----------
URT1 ----------CAKLLEPYIPEDNN-------GQG
At4g00060 ----------ENELTCITSTSDSC-------GKQ
At5g53770 ----------PNRSILGTIIRPDRVLSERKGGQN
HESO1 ----------IIGILTGQHIQESL-------YRT
At3g45750 SVVQIPPVAPLNGKTAGIHKVSGQQAVVQPSPPVPRLNGKIAGTHRKQKV-------FAQ
At3g45760 ----------LNGKVAGEHLTQKL-------FSQ
Cid1 ----------YDSLFEEAPIPPRR-------QKK
At3g51620 ----------LRKFFSNMLLRHGS-------GQR
At3g61690 ----------VNQFFMNTWERHGS-------GRR
At2g40520 ----------LEKFFNVSLERNGK-------GQR
At3g56320 ----------LEKFFRNSLERNGK-------GQR
URT1 HN----------
At4g00060 QL----------
At5g53770 GD----------VTFN----------
HESO1 ----------ISLP----------
At3g45750 QA----------VVEPRPPV----------LPLNGNIAGTH
At3g45760 QT----------VVEP----------
Cid1 TD----------
At3g51620 PD----------VHDAIPFLRYNRYNAILPASNHFQEGQVVNESESSSSSGATGNGRHDQ
At3g61690 PDAPGNDLWLSRLGDPEPYHQAENVSNSLNNKRNQNAIRLGGVHGARSMPSQQNNCGTEI
At2g40520 QD----------VEEP----------
At3g56320 QD----------VNDP----------


URT1
At4g00060
At5g53770
HESO1
At3g45750
At3g45760
Cid1
At3g51620
At3g61690
At2g40520
At3g56320


URT1
At4g00060
At5g53770
HESO1
At3g45750
At3g45760
Cid1
At3g51620
At3g61690
At2g40520
At3g56320
GNR
PLSPPGLPQLPNNNS
HYFPSGAHPISEKPSKTGSDDMGSSEVNVDESDNDLWHEPERGTHSFGLENGGYGMHQAD
DKHQSSFAEHSFVPSRRKNRLTRGDDLENSHSPVRGSSQIQSEERTVGSRSVSGASSVRS
RTSSESSWDGSTTRGSKPAKDRRNRKVVSGAASTLYGKGKSVPEHSIQIDDDNREWIPVS
SNEIIDRDLGPRPTVPSFQVQRHQIHGHELAQASGSESTVSLAPFILGHGMQQNEVDNSG


URT1
At4g00060
At5g53770
HESO1
At3g45750
At3g45760
Cid1
At3g51620
At3g61690
At2g40520
At3g56320


URT1
At4g00060
At5g53770
HESO1
At3g45750
At3g45760
Cid1
At3g51620
At3g61690
At2g40520
At3g56320

Bibliographie

Adler, J., and Pagakis, S.N. (2003). Reducing image distortions due to temperature-related microscope stage drift. J. Microsc. *210*, 131–137.

Anantharaman, V., and Aravind, L. (2004). Novel conserved domains in proteins with predicted roles in eukaryotic cell-cycle regulation, decapping and RNA stability. BMC Genomics *5*, 45.

Anderson, P., and Kedersha, N. (2002). Stressful initiations. J. Cell Sci. *115*, 3227–3234.

Anderson, P., and Kedersha, N. (2006). RNA granules. J. Cell Biol. *172*, 803–808.

Anderson, P., and Kedersha, N. (2008). Stress granules: the Tao of RNA triage. Trends Biochem. Sci. *33*, 141–150.

Andrei, M.A., Ingelfinger, D., Heintzmann, R., Achsel, T., Rivera-Pomar, R., and Lührmann, R. (2005). A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. RNA New York N *11*, 717–727.

Aphasizhev, R., and Aphasizheva, I. (2007). RNA Editing Uridylyltransferases of Trypanosomatids. In Methods in Enzymology, Jonatha M. Gott, ed. (Academic Press), pp. 55–73.

Aphasizhev, R., and Aphasizheva, I. (2008). Terminal RNA uridylyltransferases of trypanosomes. Biochim. Biophys. Acta *1779*, 270–280.

Aphasizhev, R., and Aphasizheva, I. (2011a). Mitochondrial RNA processing in trypanosomes. Res. Microbiol. *162*, 655–663.

Aphasizhev, R., and Aphasizheva, I. (2011b). Uridine Insertion/Deletion Editing In Trypanosomes: A Playground for RNA-guided Information Transfer. Wiley Interdiscip. Rev. RNA *2*, 669–685.

Aphasizheva, I., and Aphasizhev, R. (2010). RET1-catalyzed uridylylation shapes the mitochondrial transcriptome in Trypanosoma brucei. Mol. Cell. Biol. *30*, 1555–1567.

Araki, Y., Takahashi, S., Kobayashi, T., Kajiho, H., Hoshino, S., and Katada, T. (2001). Ski7p G protein interacts with the exosome and the Ski complex for 3'-to-5' mRNA decay in yeast. EMBO J. *20*, 4684–4693.

Arigo, J.T., Eyler, D.E., Carroll, K.L., and Corden, J.L. (2006). Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. Mol. Cell *23*, 841–851.

Arribere, J.A., Doudna, J.A., and Gilbert, W.V. (2011). Reconsidering Movement of Eukaryotic mRNAs between Polysomes and P Bodies. Mol. Cell *44*, 745–758.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. *25*, 25–29.

Astuti, D., Morris, M.R., Cooper, W.N., Staals, R.H.J., Wake, N.C., Fews, G.A., Gill, H., Gentle, D., Shuib, S., Ricketts, C.J., et al. (2012). Germline mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility. Nat. Genet. *44*, 277–284.

Athanasopoulos, V., Barker, A., Yu, D., Tan, A.H.-M., Srivastava, M., Contreras, N., Wang, J., Lam, K.-P., Brown, S.H.J., Goodnow, C.C., et al. (2010). The ROQUIN family of proteins localizes to stress granules via the ROQ domain and binds target mRNAs. FEBS J. *277*, 2109–2127.

Azzouz, N., Panasenko, O.O., Colau, G., and Collart, M.A. (2009). The CCR4-NOT complex physically and functionally interacts with TRAMP and the nuclear exosome. PloS One *4*, e6760.

Badis, G., Saveanu, C., Fromont-Racine, M., and Jacquier, A. (2004). Targeted mRNA degradation by deadenylation-independent decapping. Mol. Cell *15*, 5–15.

Bagijn, M.P., Goldstein, L.D., Sapetschnig, A., Weick, E.-M., Bouasker, S., Lehrbach, N.J., Simard, M.J., and Miska, E.A. (2012). Function, targets, and evolution of Caenorhabditis elegans piRNAs. Science *337*, 574–578.

Baguet, A., Degot, S., Cougot, N., Bertrand, E., Chenard, M.-P., Wendling, C., Kessler, P., Le Hir, H., Rio, M.-C., and Tomasetto, C. (2007). The exon-junction-complex-component metastatic lymph node 51 functions in stress-granule assembly. J. Cell Sci. *120*, 2774–2784.

Balagopal, V., and Parker, R. (2009). Polysomes, P bodies and stress granules: states and fates of eukaryotic mRNAs. Curr. Opin. Cell Biol. *21*, 403–408.

Balatsos, N.A.A., Maragozidis, P., Anastasakis, D., and Stathopoulos, C. (2012). Modulation of poly(A)-specific ribonuclease (PARN): current knowledge and perspectives. Curr. Med. Chem. *19*, 4838–4849.

Basquin, J., Roudko, V.V., Rode, M., Basquin, C., Séraphin, B., and Conti, E. (2012). Architecture of the Nuclease Module of the Yeast Ccr4-Not Complex: the Not1-Caf1-Ccr4 Interaction. Mol. Cell *48*, 207–218.

Baulcombe, D. (2004). RNA silencing in plants. Nature *431*, 356–363.

Bayfield, M.A., Yang, R., and Maraia, R.J. (2010). Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). Biochim. Biophys. Acta *1799*, 365–378.

Beckham, C.J., and Parker, R. (2008). P Bodies, Stress Granules, and Viral Life Cycles. Cell Host Microbe *3*, 206–212.

Beckham, C.J., Light, H.R., Nissan, T.A., Ahlquist, P., Parker, R., and Noueiry, A. (2007). Interactions between Brome Mosaic Virus RNAs and Cytoplasmic Processing Bodies. J. Virol. *81*, 9759–9768.

Belostotsky, D.A. (2003). Unexpected complexity of poly(A)-binding protein gene families in flowering plants: three conserved lineages that are at least 200 million years old and possible auto- and cross-regulation. Genetics *163*, 311–319.

Belostotsky, D.A., and Sieburth, L.E. (2009). Kill the messenger: mRNA decay and plant development. Curr. Opin. Plant Biol. *12*, 96–102.

Belostotsky, R., Frishberg, Y., and Entelis, N. (2012). Human mitochondrial tRNA quality control in health and disease: a channelling mechanism? RNA Biol. *9*, 33–39.

Benkert, P., Tosatto, S.C.E., and Schomburg, D. (2008). QMEAN: A comprehensive scoring function for model quality assessment. Proteins *71*, 261–277.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res. *28*, 235–242.

Berndt, H., Harnisch, C., Rammelt, C., Stöhr, N., Zirkel, A., Dohm, J.C., Himmelbauer, H., Tavanez, J.-P., Hüttelmaier, S., and Wahle, E. (2012). Maturation of mammalian H/ACA box snoRNAs: PAPD5-dependent adenylation and PARN-dependent trimming. RNA New York N *18*, 958–972.

Bernstein, P., and Ross, J. (1989). Poly(A), poly(A) binding protein and the regulation of mRNA stability. Trends Biochem. Sci. *14*, 373–377.

Boeck, R., Tarun, S., Jr, Rieger, M., Deardorff, J.A., Müller-Auer, S., and Sachs, A.B. (1996). The yeast Pan2 protein is required for poly(A)-binding protein-stimulated poly(A)-nuclease activity. J. Biol. Chem. *271*, 432–438.

Bogamuwa, S., and Jang, J.-C. (2013). The Arabidopsis tandem CCCH zinc finger proteins AtTZF4, 5 and 6 are involved in light-, abscisic acid- and gibberellic acid-mediated regulation of seed germination. Plant Cell Environ.

Bolte, S., and Cordelières, F.P. (2006). A guided tour into subcellular colocalization analysis in light microscopy. J. Microsc. *224*, 213–232.

Booth, B.L., Jr, and Pugh, B.F. (1997). Identification and characterization of a nuclease specific for the 3' end of the U6 small nuclear RNA. J. Biol. Chem. *272*, 984–991.

Borja, M.S., Piotukh, K., Freund, C., and Gross, J.D. (2011). Dcp1 links coactivators of mRNA decapping to Dcp2 by proline recognition. RNA New York N *17*, 278–290.

Borsani, O., Zhu, J., Verslues, P.E., Sunkar, R., and Zhu, J.-K. (2005). Endogenous siRNAs Derived from a Pair of Natural cis-Antisense Transcripts Regulate Salt Tolerance in Arabidopsis. Cell *123*, 1279–1291.

Bousquet-Antonelli, C., and Deragon, J.-M. (2009). A comprehensive analysis of the La-motif protein superfamily. RNA *15*, 750–764.

Bouveret, E., Rigaut, G., Shevchenko, A., Wilm, M., and Séraphin, B. (2000). A Sm-like protein complex that participates in mRNA degradation. EMBO J. *19*, 1661–1671.

Brannan, K., Kim, H., Erickson, B., Glover-Cutter, K., Kim, S., Fong, N., Kiemele, L., Hansen, K., Davis, R., Lykke-Andersen, J., et al. (2012). mRNA decapping factors and the exonuclease Xrn2 function in widespread premature termination of RNA polymerase II transcription. Mol. Cell *46*, 311–324.

Brendel, V., Bucher, P., Nourbakhsh, I.R., Blaisdell, B.E., and Karlin, S. (1992). Methods and algorithms for statistical analysis of protein sequences. Proc. Natl. Acad. Sci. U. S. A. *89*, 2002–2006.

Brengues, M., Teixeira, D., and Parker, R. (2005). Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies. Science *310*, 486–489.

Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and

Hannon, G.J. (2007). Discrete Small RNA-Generating Loci as Master Regulators of Transposon Activity in Drosophila. Cell *128*, 1089–1103.

Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y.Y., Sieburth, L., and Voinnet, O. (2008). Widespread Translational Inhibition by Plant miRNAs and siRNAs. Science *320*, 1185–1190.

Brown, C.E., and Sachs, A.B. (1998). Poly(A) tail length control in Saccharomyces cerevisiae occurs by message-specific deadenylation. Mol. Cell. Biol. *18*, 6548–6559.

Brown, C.E., Tarun, S.Z., Jr, Boeck, R., and Sachs, A.B. (1996). PAN3 encodes a subunit of the Pab1p-dependent poly(A) nuclease in Saccharomyces cerevisiae. Mol. Cell. Biol. *16*, 5744–5753.

Buchan, J.R., and Parker, R. (2009). Eukaryotic Stress Granules: The Ins and Outs of Translation. Mol. Cell *36*, 932–941.

Buchan, D.W.A., Ward, S.M., Lobley, A.E., Nugent, T.C.O., Bryson, K., and Jones, D.T. (2010). Protein annotation and modelling servers at University College London. Nucleic Acids Res. *38*, W563–W568.

Buchan, J.R., Muhlrad, D., and Parker, R. (2008). P bodies promote stress granule assembly in Saccharomyces cerevisiae. J. Cell Biol. *183*, 441–455.

Buchan, J.R., Kolaitis, R.-M., Taylor, J.P., and Parker, R. (2013). Eukaryotic Stress Granules Are Cleared by Autophagy and Cdc48/VCP Function. Cell *153*, 1461–1474.

Burroughs, A.M., Ando, Y., Hoon, M.J.L. de, Tomaru, Y., Nishibu, T., Ukekawa, R., Funakoshi, T., Kurokawa, T., Suzuki, H., Hayashizaki, Y., et al. (2010). A comprehensive survey of 3′ animal miRNA modification events and a possible role for 3′ adenylation in modulating miRNA targeting effectiveness. Genome Res. *20*, 1398–1410.

Burrows, C., Abd Latip, N., Lam, S.-J., Carpenter, L., Sawicka, K., Tzolovsky, G., Gabra, H., Bushell, M., Glover, D.M., Willis, A.E., et al. (2010). The RNA binding protein Larp1 regulates cell division, apoptosis and cell migration. Nucleic Acids Res. *38*, 5542–5553.

Butler, J.S., and Mitchell, P. (2011). Rrp6, rrp47 and cofactors of the nuclear exosome. Adv. Exp. Med. Biol. *702*, 91–104.

Büttner, K., Wenig, K., and Hopfner, K.-P. (2006). The exosome: a macromolecular cage for controlled RNA degradation. Mol. Microbiol. *61*, 1372–1379.

Cai, L., Fritz, D., Stefanovic, L., and Stefanovic, B. (2010). BINDING OF LARP6 TO THE CONSERVED 5' STEM-LOOP REGULATES TRANSLATION OF mRNAs ENCODING TYPE I COLLAGEN. J. Mol. Biol. *395*, 309–326.

Campen, A., Williams, R.M., Brown, C.J., Meng, J., Uversky, V.N., and Dunker, A.K. (2008). TOP-IDP-scale: a new amino acid scale measuring propensity for intrinsic disorder. Protein Pept. Lett. *15*, 956–963.

Caponigro, G., and Parker, R. (1995). Multiple functions for the poly(A)-binding protein in mRNA decapping and deadenylation in yeast. Genes Dev. *9*, 2421–2432.

Carbon, S., Ireland, A., Mungall, C.J., Shu, S., Marshall, B., and Lewis, S. (2009).

AmiGO: online access to ontology and annotation data. Bioinformatics *25*, 288–289.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. *17*, 540–552.

Cerutti, H., and Ibrahim, F. (2011). Turnover of Mature miRNAs and siRNAs in Plants and Algae. Adv. Exp. Med. Biol. *700*, 124–139.

Chalupníková, K., Lattmann, S., Selak, N., Iwamoto, F., Fujiki, Y., and Nagamine, Y. (2008). Recruitment of the RNA helicase RHAU to stress granules via a unique RNA-binding domain. J. Biol. Chem. *283*, 35186–35198.

Chang, J.H., and Tong, L. (2012). Mitochondrial poly(A) polymerase and polyadenylation. Biochim. Biophys. Acta *1819*, 992–997.

Chang, H.-M., Triboulet, R., Thornton, J.E., and Gregory, R.I. (2013). A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28-let-7 pathway. Nature *497*, 244–248.

Chekanova, J.A., Shaw, R.J., Wills, M.A., and Belostotsky, D.A. (2000). Poly(A) tail-dependent exonuclease AtRrp41p from Arabidopsis thaliana rescues 5.8 S rRNA processing and mRNA decay defects of the yeast ski6 mutant and is found in an exosome-sized complex in plant and yeast cells. J. Biol. Chem. *275*, 33158–33166.

Chekanova, J.A., Gregory, B.D., Reverdatto, S.V., Chen, H., Kumar, R., Hooker, T., Yazaki, J., Li, P., Skiba, N., Peng, Q., et al. (2007). Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the Arabidopsis transcriptome. Cell *131*, 1340–1353.

Chen, C.-Y.A., and Shyu, A.-B. (2011). Mechanisms of deadenylation-dependent decay. Wiley Interdiscip. Rev. RNA *2*, 167–183.

Chen, C.-Y.A., and Shyu, A.-B. (2013). Deadenylation and P-bodies. Adv. Exp. Med. Biol. *768*, 183–195.

Chernyakov, I., Whipple, J.M., Kotelawala, L., Grayhack, E.J., and Phizicky, E.M. (2008). Degradation of several hypomodified mature tRNA species in Saccharomyces cerevisiae is mediated by Met22 and the 5?-3? exonucleases Rat1 and Xrn1. Genes Dev. *22*, 1369–1380.

Chiang, H.R., Schoenfeld, L.W., Ruby, J.G., Auyeung, V.C., Spies, N., Baek, D., Johnston, W.K., Russ, C., Luo, S., Babiarz, J.E., et al. (2010). Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. Genes Dev. *24*, 992–1009.

Chiba, Y., and Green, P.J. (2009). mRNA Degradation Machinery in Plants. J. Plant Biol. *52*, 114–124.

Chiba, Y., Johnson, M.A., Lidder, P., Vogel, J.T., van Erp, H., and Green, P.J. (2004). AtPARN is an essential poly(A) ribonuclease in Arabidopsis. Gene *328*, 95–102.

Choi, Y.S., Patena, W., Leavitt, A.D., and McManus, M.T. (2012). Widespread RNA 3′-end oligouridylation in mammals. RNA *18*, 394–401.

Chowdhury, A., Mukhopadhyay, J., and Tharun, S. (2007). The decapping activator Lsm1p-7p-Pat1p complex has the intrinsic ability to distinguish between oligoadenylated and

polyadenylated RNAs. RNA New York N *13*, 998–1016.

Chu, C., and Rana, T.M. (2006). Translation Repression in Human Cells by MicroRNA-Induced Gene Silencing Requires RCK/p54. PLoS Biol. *4*.

Cohen-Katsenelson, K., Wasserman, T., Darlyuk-Saadon, I., Rabner, A., Glaser, F., and Aronheim, A. (2013). Identification and analysis of a novel dimerization domain shared by various members of c-Jun N-terminal kinase (JNK) scaffold proteins. J. Biol. Chem. *288*, 7294–7304.

Colgan, D.F., and Manley, J.L. (1997). Mechanism and regulation of mRNA polyadenylation. Genes Dev. *11*, 2755–2766.

Collart, M.A., and Panasenko, O.O. (2012). The Ccr4--not complex. Gene *492*, 42–53.

Coller, J., and Parker, R. (2005). General Translational Repression by Activators of mRNA Decapping. Cell *122*, 875–886.

Colombrita, C., Zennaro, E., Fallini, C., Weber, M., Sommacal, A., Buratti, E., Silani, V., and Ratti, A. (2009). TDP-43 is recruited to stress granules in conditions of oxidative insult. J. Neurochem. *111*, 1051–1061.

Combet, C., Blanchet, C., Geourjon, C., and Deléage, G. (2000). NPS@: network protein sequence analysis. Trends Biochem. Sci. *25*, 147–150.

Cougot, N., Babajko, S., and Séraphin, B. (2004). Cytoplasmic foci are sites of mRNA decay in human cells. J. Cell Biol. *165*, 31–40.

Creamer, T.J., Darby, M.M., Jamonnak, N., Schaughency, P., Hao, H., Wheelan, S.J., and Corden, J.L. (2011). Transcriptome-Wide Binding Sites for Components of the Saccharomyces cerevisiae Non-Poly(A) Termination Pathway: Nrd1, Nab3, and Sen1. PLoS Genet. *7*.

Crooks, G.E., Hon, G., Chandonia, J.-M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. Genome Res. *14*, 1188–1190.

Dang, Y., Kedersha, N., Low, W.-K., Romo, D., Gorospe, M., Kaufman, R., Anderson, P., and Liu, J.O. (2006). Eukaryotic initiation factor 2alpha-independent pathway of stress granule induction by the natural product pateamine A. J. Biol. Chem. *281*, 32870–32878.

Decker, C.J., Teixeira, D., and Parker, R. (2007). Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in Saccharomyces cerevisiae. J. Cell Biol. *179*, 437–449.

Deng, J., Ernst, N.L., Turley, S., Stuart, K.D., and Hol, W.G. (2005). Structural basis for UTP specificity of RNA editing TUTases from Trypanosoma brucei. EMBO J. *24*, 4007–4017.

Denis, C.L., and Chen, J. (2003). The CCR4-NOT complex plays diverse roles in mRNA metabolism. Prog. Nucleic Acid Res. Mol. Biol. *73*, 221–250.

Van Der Kelen, K., Beyaert, R., Inzé, D., and De Veylder, L. (2009). Translational control of eukaryotic gene expression. Crit. Rev. Biochem. Mol. Biol. *44*, 143–168.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.-

F., Guindon, S., Lefort, V., Lescot, M., et al. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. *36*, W465–469.

Van Dijk, E., Le Hir, H., and Seraphin, B. (2003a). DcpS can act in the 5?-3? mRNA decay pathway in addition to the 3?-5? pathway. Proc. Natl. Acad. Sci. U. S. A. *100*, 12081–12086.

Van Dijk, E., Le Hir, H., and Séraphin, B. (2003b). DcpS can act in the 5'-3' mRNA decay pathway in addition to the 3'-5' pathway. Proc. Natl. Acad. Sci. U. S. A. *100*, 12081–12086.

Dolken, L., Perot, J., Cognat, V., Alioua, A., John, M., Soutschek, J., Ruzsics, Z., Koszinowski, U., Voinnet, O., and Pfeffer, S. (2007). Mouse Cytomegalovirus MicroRNAs Dominate the Cellular Small RNA Profile during Lytic Infection and Show Features of Posttranscriptional Regulation. J. Virol. *81*, 13771–13782.

Dominski, Z., and Marzluff, W.F. (1999). Formation of the 3' end of histone mRNA. Gene *239*, 1–14.

Dominski, Z., Yang, X., Kaygun, H., Dadlez, M., and Marzluff, W.F. (2003). A 3′ Exonuclease that Specifically Interacts with the 3′ End of Histone mRNA. Mol. Cell *12*, 295–305.

Dorcey, E., Rodriguez-Villalon, A., Salinas, P., Santuari, L., Pradervand, S., Harshman, K., and Hardtke, C.S. (2012). Context-dependent dual role of SKI8 homologs in mRNA synthesis and turnover. PLoS Genet. *8*, e1002652.

Dosztanyi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005). IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics *21*, 3433–3434.

Dosztányi, Z., Mészáros, B., and Simon, I. (2009). ANCHOR: web server for predicting protein binding regions in disordered proteins. Bioinformatics *25*, 2745–2746.

Dunckley, T., and Parker, R. (1999). The DCP2 protein is required for mRNA decapping in Saccharomyces cerevisiae and contains a functional MutT motif. EMBO J. *18*, 5411–5422.

Dunker, A.K., Lawson, J.D., Brown, C.J., Williams, R.M., Romero, P., Oh, J.S., Oldfield, C.J., Campen, A.M., Ratliff, C.M., Hipps, K.W., et al. (2001). Intrinsically disordered protein. J. Mol. Graph. Model. *19*, 26–59.

Eberhardt, W., Doller, A., Akool, E.-S., and Pfeilschifter, J. (2007). Modulation of mRNA stability as a novel therapeutic approach. Pharmacol. Ther. *114*, 56–73.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797.

Emara, M.M., Fujimura, K., Sciaranghella, D., Ivanova, V., Ivanov, P., and Anderson, P. (2012). Hydrogen peroxide induces stress granule formation independent of eIF2α phosphorylation. Biochem. Biophys. Res. Commun. *423*, 763–769.

Erkmann, J.A., Wagner, E.J., Dong, J., Zhang, Y., Kutay, U., and Marzluff, W.F. (2005). Nuclear import of the stem-loop binding protein and localization during the cell cycle.

Mol. Biol. Cell *16*, 2960–2971.

Ernoult-Lange, M., Baconnais, S., Harper, M., Minshall, N., Souquere, S., Boudier, T., Bénard, M., Andrey, P., Pierron, G., Kress, M., et al. (2012). Multiple binding of repressed mRNAs by the P-body protein Rck/p54. RNA New York N *18*, 1702–1715.

Espinosa Angarica, V., Ventura, S., and Sancho, J. (2013). Discovering putative prion sequences in complete proteomes using probabilistic representations of Q/N-rich domains. BMC Genomics *14*, 316.

Estavillo, G.M., Crisp, P.A., Pornsiriwong, W., Wirtz, M., Collinge, D., Carrie, C., Giraud, E., Whelan, J., David, P., Javot, H., et al. (2011). Evidence for a SAL1-PAP chloroplast retrograde pathway that functions in drought and high light signaling in Arabidopsis. Plant Cell *23*, 3992–4012.

Eulalio, A., Behm-Ansmant, I., Schweizer, D., and Izaurralde, E. (2007a). P-Body Formation Is a Consequence, Not the Cause, of RNA-Mediated Gene Silencing. Mol. Cell. Biol. *27*, 3970–3981.

Eulalio, A., Behm-Ansmant, I., and Izaurralde, E. (2007b). P bodies: at the crossroads of post-transcriptional pathways. Nat. Rev. Mol. Cell Biol. *8*, 9–22.

Eulalio, A., Tritschler, F., and Izaurralde, E. (2009). The GW182 protein family in animal cells: New insights into domains required for miRNA-mediated gene silencing. RNA *15*, 1433–1442.

Fatica, A., and Tollervey, D. (2002). Making ribosomes. Curr. Opin. Cell Biol. *14*, 313–318.

Franks, T.M., and Lykke-Andersen, J. (2007). TTP and BRF proteins nucleate processing body formation to silence mRNAs with AU-rich elements. Genes Dev. *21*, 719–735.

Franks, T.M., and Lykke-Andersen, J. (2008). The control of mRNA decapping and P-body formation. Mol. Cell *32*, 605–615.

Frishman, D., and Argos, P. (1996). Incorporation of non-local interactions in protein secondary structure prediction from the amino acid sequence. Protein Eng. *9*, 133–142.

Garcia, D., Garcia, S., Pontier, D., Marchais, A., Renou, J.P., Lagrange, T., and Voinnet, O. (2012). Ago hook and RNA helicase motifs underpin dual roles for SDE3 in antiviral defense and silencing of nonconserved intergenic regions. Mol. Cell *48*, 109–120.

Gareau, C., Houssin, E., Martel, D., Coudert, L., Mellaoui, S., Huot, M.-E., Laprise, P., and Mazroui, R. (2013). Characterization of fragile X mental retardation protein recruitment and dynamics in Drosophila stress granules. PloS One *8*, e55342.

Garneau, N.L., Wilusz, J., and Wilusz, C.J. (2007). The highways and byways of mRNA decay. Nat. Rev. Mol. Cell Biol. *8*, 113–126.

Gazzani, S., Lawrenson, T., Woodward, C., Headon, D., and Sablowski, R. (2004). A link between mRNA turnover and RNA interference in Arabidopsis. Science *306*, 1046–1048.

Geisler, S., Lojek, L., Khalil, A.M., Baker, K.E., and Coller, J. (2012). Decapping of long noncoding RNAs regulates inducible genes. Mol. Cell *45*, 279–291.

Gibbings, D.J., Ciaudo, C., Erhardt, M., and Voinnet, O. (2009). Multivesicular bodies associate with components of miRNA effector complexes and modulate miRNA activity. Nat. Cell Biol. *11*, 1143–1149.

Gilks, N., Kedersha, N., Ayodele, M., Shen, L., Stoecklin, G., Dember, L.M., and Anderson, P. (2004). Stress granule assembly is mediated by prion-like aggregation of TIA-1. Mol. Biol. Cell *15*, 5383–5398.

Golisz, A., Sikorski, P.J., Kruszka, K., and Kufel, J. (2013). Arabidopsis thaliana LSM proteins function in mRNA splicing and degradation. Nucleic Acids Res.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., et al. (2011). Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. *40*, D1178–D1186.

Goulet, I., Boisvenue, S., Mokas, S., Mazroui, R., and Côté, J. (2008). TDRD3, a novel Tudor domain-containing protein, localizes to cytoplasmic stress granules. Hum. Mol. Genet. *17*, 3055–3074.

Grabowski, P.J., and Sharp, P.A. (1986). Affinity chromatography of splicing complexes: U2, U5, and U4+U6 small nuclear ribonucleoprotein particles in the spliceosome. Science.

Grima, D.P., Sullivan, M., Zabolotskaya, M.V., Browne, C., Seago, J., Wan, K.C., Okada, Y., and Newbury, S.F. (2008). The 5'-3' exoribonuclease pacman is required for epithelial sheet sealing in Drosophila and genetically interacts with the phosphatase puckered. Biol. Cell Auspices Eur. Cell Biol. Organ. *100*, 687–701.

Groušl, T., Ivanov, P., Frydlová, I., Vašicová, P., Janda, F., Vojtová, J., Malínská, K., Malcová, I., Nováková, L., Janošková, D., et al. (2009). Robust heat shock induces eIF2α-phosphorylation-independent assembly of stress granules containing eIF3 and 40S ribosomal subunits in budding yeast, Saccharomyces cerevisiae. J. Cell Sci. *122*, 2078–2088.

Guermeur, Y., Geourjon, C., Gallinari, P., and Del√©age, G. (1999). Improved performance in protein secondary structure prediction by inhomogeneous score combination. Bioinformatics *15*, 413–421.

Gy, I., Gasciolli, V., Lauressergues, D., Morel, J.-B., Gombert, J., Proux, F., Proux, C., Vaucheret, H., and Mallory, A.C. (2007). Arabidopsis FIERY1, XRN2, and XRN3 are endogenous RNA silencing suppressors. Plant Cell *19*, 3451–3461.

Hagan, J.P., Piskounova, E., and Gregory, R.I. (2009). Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. Nat. Struct. Mol. Biol. *16*, 1021–1025.

Halic, M., and Moazed, D. (2010). Dicer-Independent Primal RNAs Trigger RNAi and Heterochromatin Formation. Cell *140*, 504–516.

Heo, I., Ha, M., Lim, J., Yoon, M.-J., Park, J.-E., Kwon, S.C., Chang, H., and Kim, V.N. (2012). Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. Cell *151*, 521–532.

Hermann, H., Fabrizio, P., Raker, V.A., Foulaki, K., Hornig, H., Brahms, H., and Lührmann, R. (1995). snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions. EMBO J. *14*, 2076–2088.

Hirayama, T., Matsuura, T., Ushiyama, S., Narusaka, M., Kurihara, Y., Yasuda, M., Ohtani, M., Seki, M., Demura, T., Nakashita, H., et al. (2013). A poly(A)-specific ribonuclease directly regulates the poly(A) status of mitochondrial mRNA in Arabidopsis. Nat. Commun. *4*.

Hirsch, J., Misson, J., Crisp, P.A., David, P., Bayle, V., Estavillo, G.M., Javot, H., Chiarenza, S., Mallory, A.C., Maizel, A., et al. (2011). A novel fry1 allele reveals the existence of a mutant phenotype unrelated to 5'->3' exoribonuclease (XRN) activities in Arabidopsis thaliana roots. PloS One *6*, e16724.

Hoefig, K.P., Rath, N., Heinz, G.A., Wolf, C., Dameris, J., Schepers, A., Kremmer, E., Ansel, K.M., and Heissmeyer, V. (2013). Eri1 degrades the stem-loop of oligouridylated histone mRNAs to induce replication-dependent decay. Nat. Struct. Mol. Biol. *20*, 73–81.

Honys, D., and Twell, D. (2004). Transcriptome analysis of haploid male gametophyte development in Arabidopsis. Genome Biol. *5*, R85.

Van Hoof, A., Frischmeyer, P.A., Dietz, H.C., and Parker, R. (2002). Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. Science *295*, 2262–2264.

Horwich, M.D., Li, C., Matranga, C., Vagin, V., Farley, G., Wang, P., and Zamore, P.D. (2007). The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. Curr. Biol. CB *17*, 1265–1272.

Hu, W., Sweet, T.J., Chamnongpol, S., Baker, K.E., and Coller, J. (2009). Co-translational mRNA decay in Saccharomyces cerevisiae. Nature *461*, 225–229.

Hu, W., Petzold, C., Coller, J., and Baker, K.E. (2010). Nonsense-mediated mRNA decapping occurs on polyribosomes in Saccharomyces cerevisiae. Nat. Struct. Mol. Biol. *17*, 244–247.

Huang, L., Mollet, S., Souquere, S., Le Roy, F., Ernoult-Lange, M., Pierron, G., Dautry, F., and Weil, D. (2011). Mitochondria associate with P-bodies and modulate microRNA-mediated RNA interference. J. Biol. Chem. *286*, 24219–24230.

Hung, K.-H., and Stumph, W.E. (2011). Regulation of snRNA gene expression by the Drosophila melanogaster small nuclear RNA activating protein complex (DmSNAPc). Crit. Rev. Biochem. Mol. Biol. *46*, 11–26.

Huntzinger, E., and Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nat. Rev. Genet. *12*, 99–110.

Ibrahim, F., Rohr, J., Jeong, W.-J., Hesson, J., and Cerutti, H. (2006). Untemplated oligoadenylation promotes degradation of RISC-cleaved transcripts. Science *314*, 1893.

Ibrahim, F., Rymarquis, L.A., Kim, E.-J., Becker, J., Balassa, E., Green, P.J., and Cerutti, H. (2010). Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in Chlamydomonas. Proc. Natl. Acad. Sci. U. S. A. *107*, 3906–3911.

Ishizu, H., Siomi, H., and Siomi, M.C. (2012). Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. Genes Dev. *26*, 2361–2373.

Iwaki, A., and Izawa, S. (2012). Acidic stress induces the formation of P-bodies, but not stress granules, with mild attenuation of bulk translation in Saccharomyces cerevisiae. Biochem. J. *446*, 225–233.

Iwasaki, S., Takeda, A., Motose, H., and Watanabe, Y. (2007). Characterization of Arabidopsis decapping proteins AtDCP1 and AtDCP2, which are essential for post-embryonic development. FEBS Lett. *581*, 2455–2459.

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, A.W. (1997). Rat1p and Xrn1p are functionally interchangeable exoribonucleases that are restricted to and required in the nucleus and cytoplasm, respectively. Mol. Cell. Biol. *17*, 6122–6130.

Johnson, B.S., McCaffery, J.M., Lindquist, S., and Gitler, A.D. (2008). A yeast TDP-43 proteinopathy model: Exploring the molecular determinants of TDP-43 aggregation and cellular toxicity. Proc. Natl. Acad. Sci. U. S. A. *105*, 6439–6444.

Jones, M.R., Quinton, L.J., Blahna, M.T., Neilson, J.R., Fu, S., Ivanov, A.R., Wolf, D.A., and Mizgerd, J.P. (2009). Zcchc11-dependent uridylation of microRNA directs cytokine expression. Nat. Cell Biol. *11*, 1157–1163.

Jones, M.R., Blahna, M.T., Kozlowski, E., Matsuura, K.Y., Ferrari, J.D., Morris, S.A., Powers, J.T., Daley, G.Q., Quinton, L.J., and Mizgerd, J.P. (2012). Zcchc11 Uridylates Mature miRNAs to Enhance Neonatal IGF-1 Expression, Growth, and Survival. PLoS Genet *8*, e1003105.

Jouannet, V., Moreno, A.B., Elmayan, T., Vaucheret, H., Crespi, M.D., and Maizel, A. (2012). Cytoplasmic Arabidopsis AGO7 accumulates in membrane-associated siRNA bodies and is required for ta-siRNA biogenesis. EMBO J. *31*, 1704–1713.

Juntawong, P., and Bailey-Serres, J. (2012). Dynamic Light Regulation of Translation Status in Arabidopsis thaliana. Front. Plant Sci. *3*, 66.

Kalmar, L., Acs, V., Silhavy, D., and Tompa, P. (2012). Long-range interactions in nonsense-mediated mRNA decay are mediated by intrinsically disordered protein regions. J. Mol. Biol. *424*, 125–131.

Kamminga, L.M., Luteijn, M.J., den Broeder, M.J., Redl, S., Kaaij, L.J.T., Roovers, E.F., Ladurner, P., Berezikov, E., and Ketting, R.F. (2010). Hen1 is required for oocyte development and piRNA stability in zebrafish. EMBO J. *29*, 3688–3700.

Kasschau, K.D., Fahlgren, N., Chapman, E.J., Sullivan, C.M., Cumbie, J.S., Givan, S.A., and Carrington, J.C. (2007). Genome-Wide Profiling and Analysis of Arabidopsis siRNAs. PLoS Biol *5*, e57.

Kastenmayer, J.P., and Green, P.J. (2000). Novel features of the XRN-family in Arabidopsis: evidence that AtXRN4, one of several orthologs of nuclear Xrn2p/Rat1p,

functions in the cytoplasm. Proc. Natl. Acad. Sci. U. S. A. *97*, 13985–13990.

Katiyar-Agarwal, S., Gao, S., Vivian-Smith, A., and Jin, H. (2007). A novel class of bacteria-induced small RNAs in Arabidopsis. Genes Dev. *21*, 3123–3134.

Kato, M., Han, T.W., Xie, S., Shi, K., Du, X., Wu, L.C., Mirzaei, H., Goldsmith, E.J., Longgood, J., Pei, J., et al. (2012). Cell-free Formation of RNA Granules: Low Complexity Sequence Domains Form Dynamic Fibers within Hydrogels. Cell *149*, 753–767.

Katoh, T., Sakaguchi, Y., Miyauchi, K., Suzuki, T., Kashiwabara, S., Baba, T., and Suzuki, T. (2009). Selective stabilization of mammalian microRNAs by 3′ adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. Genes Dev. *23*, 433–438.

Kaufmann, I., Martin, G., Friedlein, A., Langen, H., and Keller, W. (2004). Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. EMBO J. *23*, 616–626.

Kedersha, N., Chen, S., Gilks, N., Li, W., Miller, I.J., Stahl, J., and Anderson, P. (2002). Evidence That Ternary Complex (eIF2-GTP-tRNAi Met)–Deficient Preinitiation Complexes Are Core Constituents of Mammalian Stress Granules. Mol. Biol. Cell *13*, 195–210.

Kedersha, N., Stoecklin, G., Ayodele, M., Yacono, P., Lykke-Andersen, J., Fritzler, M.J., Scheuner, D., Kaufman, R.J., Golan, D.E., and Anderson, P. (2005a). Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. J. Cell Biol. *169*, 871–884.

Kedersha, N., Stoecklin, G., Ayodele, M., Yacono, P., Lykke-Andersen, J., Fritzler, M.J., Scheuner, D., Kaufman, R.J., Golan, D.E., and Anderson, P. (2005b). Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. J. Cell Biol. *169*, 871–884.

Kedersha, N.L., Gupta, M., Li, W., Miller, I., and Anderson, P. (1999). RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2 alpha to the assembly of mammalian stress granules. J. Cell Biol. *147*, 1431–1442.

Keller, C., Woolcock, K., Hess, D., and Buhler, M. (2010). Proteomic and functional analysis of the noncanonical poly(A) polymerase Cid14. RNA *16*, 1124–1129.

Khraiwesh, B., Zhu, J.-K., and Zhu, J. (2012). Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. Biochim. Biophys. Acta *1819*, 137–148.

Kiefer, F., Arnold, K., Künzli, M., Bordoli, L., and Schwede, T. (2009). The SWISS-MODEL Repository and associated resources. Nucleic Acids Res. *37*, D387–D392.

Kim, H.S., Kuwano, Y., Zhan, M., Pullmann, R., Mazan-Mamczarz, K., Li, H., Kedersha, N., Anderson, P., Wilce, M.C.J., Gorospe, M., et al. (2007). Elucidation of a C-Rich Signature Motif in Target mRNAs of RNA-Binding Protein TIAR. Mol. Cell. Biol. *27*, 6806–6817.

King, R.D., and Sternberg, M.J. (1996). Identification and application of the concepts important for accurate and reliable protein secondary structure prediction. Protein Sci. Publ. Protein Soc. *5*, 2298–2310.

Kreil, D.P., and Ouzounis, C.A. (2003). Comparison of sequence masking algorithms and the detection of biased protein sequence regions. Bioinforma. Oxf. Engl. *19*, 1672–1681.

Krüger, T., Hofweber, M., and Kramer, S. (2013). SCD6 induces RNP granule formation in trypanosomes in a translation-independent manner, regulated by its LSM and RGG domains. Mol. Biol. Cell.

Kumakura, N., Takeda, A., Fujioka, Y., Motose, H., Takano, R., and Watanabe, Y. (2009). SGS3 and RDR6 interact and colocalize in cytoplasmic SGS3/RDR6-bodies. FEBS Lett. *583*, 1261–1266.

Kwak, J.E., and Wickens, M. (2007). A family of poly(U) polymerases. RNA *13*, 860–867.

Kwak, J.E., Wang, L., Ballantyne, S., Kimble, J., and Wickens, M. (2004). Mammalian GLD-2 homologs are poly(A) polymerases. Proc. Natl. Acad. Sci. U. S. A. *101*, 4407–4412.

Laishram, R.S., and Anderson, R.A. (2010). The poly A polymerase Star-PAP controls 3'-end cleavage by promoting CPSF interaction and specificity toward the pre-mRNA. EMBO J. *29*, 4132–4145.

Lange, H., and Gagliardi, D. (2010). The Exosome and 3′–5′ RNA Degradation in Plants. In RNA Exosome, T.H. Jensen, ed. (Springer US), pp. 50–62.

Lange, H., Holec, S., Cognat, V., Pieuchot, L., Le Ret, M., Canaday, J., and Gagliardi, D. (2008). Degradation of a Polyadenylated rRNA Maturation By-Product Involves One of the Three RRP6-Like Proteins in Arabidopsis thaliana. Mol. Cell. Biol. *28*, 3038–3044.

Lange, H., Sement, F.M., Canaday, J., and Gagliardi, D. (2009). Polyadenylation-assisted RNA degradation processes in plants. Trends Plant Sci. *14*, 497–504.

Lange, H., Sement, F.M., and Gagliardi, D. (2011). MTR4, a putative RNA helicase and exosome co-factor, is required for proper rRNA biogenesis and development in Arabidopsis thaliana. Plant J. Cell Mol. Biol. *68*, 51–63.

Lapointe, C.P., and Wickens, M. (2013). The nucleic acid binding domain and translational repression activity of a Xenopus terminal uridylyl transferase. J. Biol. Chem.

Larimer, F.W., Hsu, C.L., Maupin, M.K., and Stevens, A. (1992). Characterization of the XRN1 gene encoding a 5'-->3' exoribonuclease: sequence data and analysis of disparate protein and mRNA levels of gene-disrupted yeast cells. Gene *120*, 51–57.

Laurence, N., Epelman, M., Markowitz, R.I., Jaimes, C., Jaramillo, D., and Chauvin, N.A. (2012). Osteoid osteomas: a pain in the night diagnosis. Pediatr. Radiol. *42*, 1490–1501; quiz 1540–1542.

De Leeuw, F., Zhang, T., Wauquier, C., Huez, G., Kruys, V., and Gueydan, C. (2007). The cold-inducible RNA-binding protein migrates from the nucleus to cytoplasmic stress granules by a methylation-dependent mechanism and acts as a translational repressor. Exp. Cell Res. *313*, 4130–4144.

Lehrbach, N.J., Armisen, J., Lightfoot, H.L., Murfitt, K.J., Bugaut, A., Balasubramanian, S., and Miska, E.A. (2009). LIN-28 and the poly(U) polymerase PUP-2 regulate let-7 microRNA processing in Caenorhabditis elegans. Nat. Struct. Mol. Biol. *16*,

1016–1020.

Lejeune, F., Li, X., and Maquat, L.E. (2003). Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. Mol. Cell *12*, 675–687.

Li, J., Yang, Z., Yu, B., Liu, J., and Chen, X. (2005). Methylation Protects miRNAs and siRNAs from a 3′-End Uridylation Activity in Arabidopsis. Curr. Biol. *15*, 1501–1507.

Li, W., Laishram, R.S., and Anderson, R.A. (2013). The novel poly(A) polymerase Star-PAP is a signal-regulated switch at the 3′-end of mRNAs. Adv. Biol. Regul. *53*, 64–76.

Liang, W., Li, C., Liu, F., Jiang, H., Li, S., Sun, J., Wu, X., and Li, C. (2009). The Arabidopsis homologs of CCR4-associated factor 1 show mRNA deadenylation activity and play a role in plant defence responses. Cell Res. *19*, 307–316.

Lin, M.-D., Jiao, X., Grima, D., Newbury, S.F., Kiledjian, M., and Chou, T.-B. (2008). Drosophila processing bodies in oogenesis. Dev. Biol. *322*, 276–288.

Linding, R. (2003). GlobPlot: exploring protein sequences for globularity and disorder. Nucleic Acids Res. *31*, 3701–3708.

Liu, H., and Kiledjian, M. (2005). Scavenger Decapping Activity Facilitates 5′ to 3′ mRNA Decay. Mol. Cell. Biol. *25*, 9764–9772.

Lloyd, R.E. (2012). How do viruses interact with stress-associated RNA granules? PLoS Pathog. *8*, e1002741.

Lloyd, R.E. (2013). Regulation of stress granules and P-bodies during RNA virus infection. Wiley Interdiscip. Rev. RNA *4*, 317–331.

Loschi, M., Leishman, C.C., Berardone, N., and Boccaccio, G.L. (2009). Dynein and kinesin regulate stress-granule and P-body dynamics. J. Cell Sci. *122*, 3973–3982.

Lu, G., Zhang, J., Li, Y., Li, Z., Zhang, N., Xu, X., Wang, T., Guan, Z., Gao, G.F., and Yan, J. (2011). hNUDT16: a universal decapping enzyme for small nucleolar RNA and cytoplasmic mRNA. Protein Cell *2*, 64–73.

Lubas, M., Christensen, M.S., Kristiansen, M.S., Domanski, M., Falkenby, L.G., Lykke-Andersen, S., Andersen, J.S., Dziembowski, A., and Jensen, T.H. (2011). Interaction profiling identifies the human nuclear exosome targeting complex. Mol. Cell *43*, 624–637.

Lubas, M., Damgaard, C.K., Tomecki, R., Cysewski, D., Jensen, T.H., and Dziembowski, A. (2013). Exonuclease hDIS3L2 specifies an exosome-independent 3′-5′ degradation pathway of human cytoplasmic mRNA. EMBO J. *advance online publication*.

Luke, B., and Lingner, J. (2009). TERRA: telomeric repeat-containing RNA. EMBO J. *28*, 2503–2510.

Luke, B., Panza, A., Redon, S., Iglesias, N., Li, Z., and Lingner, J. (2008). The Rat1p 5’ to 3’ exonuclease degrades telomeric repeat-containing RNA and promotes telomere elongation in Saccharomyces cerevisiae. Mol. Cell *32*, 465–477.

Lunde, B.M., Magler, I., and Meinhart, A. (2012). Crystal structures of the Cid1 poly (U) polymerase reveal the mechanism for UTP selectivity. Nucleic Acids Res. *40*, 9815–9824.

Lykke-Andersen, J., and Wagner, E. (2005). Recruitment and activation of mRNA decay enzymes by two ARE-mediated decay activation domains in the proteins TTP and BRF-1. Genes Dev. *19*, 351–361.

Lykke-Andersen, S., Brodersen, D.E., and Jensen, T.H. (2009). Origins and activities of the eukaryotic exosome. J. Cell Sci. *122*, 1487–1494.

Lykke-Andersen, S., Tomecki, R., Jensen, T.H., and Dziembowski, A. (2011). The eukaryotic RNA exosome: same scaffold but variable catalytic subunits. RNA Biol. *8*, 61–66.

Malecki, M., Viegas, S.C., Carneiro, T., Golik, P., Dressaire, C., Ferreira, M.G., and Arraiano, C.M. (2013). The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. EMBO J.

Mangus, D.A., and Jacobson, A. (1999). Linking mRNA turnover and translation: assessing the polyribosomal association of mRNA decay factors and degradative intermediates. Methods San Diego Calif *17*, 28–37.

Marcinowski, L., Tanguy, M., Krmpotic, A., Rädle, B., Lisnić, V.J., Tuddenham, L., Chane-Woon-Ming, B., Ruzsics, Z., Erhard, F., Benkartek, C., et al. (2012). Degradation of Cellular miR-27 by a Novel, Highly Abundant Viral Transcript Is Important for Efficient Virus Replication In Vivo. PLoS Pathog *8*, e1002510.

Marko, M., Vlassis, A., Guialis, A., and Leichter, M. (2012). Domains involved in TAF15 subcellular localisation: dependence on cell type and ongoing transcription. Gene *506*, 331–338.

Martin, G., and Keller, W. (2007). RNA-specific ribonucleotidyl transferases. RNA *13*, 1834–1849.

Marzluff, W.F. (2005). Metazoan replication-dependent histone mRNAs: a distinct set of RNA polymerase II transcripts. Curr. Opin. Cell Biol. *17*, 274–280.

Matsumoto, K., Minami, M., Shinozaki, F., Suzuki, Y., Abe, K., Zenno, S., Matsumoto, S., and Minami, Y. (2011). Hsp90 is involved in the formation of P-bodies and stress granules. Biochem. Biophys. Res. Commun. *407*, 720–724.

Mauchi, N., Ohtake, Y., and Irie, K. (2010). Stability control of MTL1 mRNA by the RNA-binding protein Khd1p in yeast. Cell Struct. Funct. *35*, 95–105.

Mazroui, R., Sukarieh, R., Bordeleau, M.-E., Kaufman, R.J., Northcote, P., Tanaka, J., Gallouzi, I., and Pelletier, J. (2006). Inhibition of Ribosome Recruitment Induces Stress Granule Formation Independently of Eukaryotic Initiation Factor 2α Phosphorylation. Mol. Biol. Cell *17*, 4212–4219.

McManus, M.T., Adler, B.K., Pollard, V.W., and Hajduk, S.L. (2000). Trypanosoma brucei Guide RNA Poly(U) Tail Formation Is Stabilized by Cognate mRNA. Mol. Cell. Biol. *20*, 883–891.

Mérai, Z., Benkovics, A.H., Nyikó, T., Debreczeny, M., Hiripi, L., Kerényi, Z., Kondorosi, E., and Silhavy, D. (2012). The late steps of plant nonsense-mediated mRNA decay. Plant J. Cell Mol. Biol.

Merret, R., Martino, L., Bousquet-Antonelli, C., Fneich, S., Descombin, J., Billey, E.,

Conte, M.R., and Deragon, J.-M. (2013). The association of a La module with the PABP-interacting motif PAM2 is a recurrent evolutionary process that led to the neofunctionalization of La-related proteins. RNA *19*, 36–50.

Mészáros, B., Simon, I., and Dosztányi, Z. (2009). Prediction of Protein Binding Regions in Disordered Proteins. PLoS Comput Biol *5*, e1000376.

Mészáros, B., Simon, I., and Dosztányi, Z. (2011). The expanding view of protein-protein interactions: complexes involving intrinsically disordered proteins. Phys. Biol. *8*, 035003.

Mészáros, B., Dosztányi, Z., and Simon, I. (2012). Disordered binding regions and linear motifs--bridging the gap between two models of molecular recognition. PloS One *7*, e46829.

Mi, H., Dong, Q., Muruganujan, A., Gaudet, P., Lewis, S., and Thomas, P.D. (2010). PANTHER version 7: improved phylogenetic trees, orthologs and collaboration with the Gene Ontology Consortium. Nucleic Acids Res. *38*, D204–D210.

Mi, H., Muruganujan, A., and Thomas, P.D. (2013). PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic Acids Res. *41*, D377–386.

Milligan, L., Decourty, L., Saveanu, C., Rappsilber, J., Ceulemans, H., Jacquier, A., and Tollervey, D. (2008). A yeast exosome cofactor, Mpp6, functions in RNA surveillance and in the degradation of noncoding RNA transcripts. Mol. Cell. Biol. *28*, 5446–5457.

Minasaki, R., and Eckmann, C.R. (2012). Subcellular specialization of multifaceted 3'end modifying nucleotidyltransferases. Curr. Opin. Cell Biol. *24*, 314–322.

Moenne, A., Camier, S., Anderson, G., Margottin, F., Beggs, J., and Sentenac, A. (1990). The U6 gene of Saccharomyces cerevisiae is transcribed by RNA polymerase C (III) in vivo and in vitro. EMBO J. *9*, 271.

Mollet, S., Cougot, N., Wilczynska, A., Dautry, F., Kress, M., Bertrand, E., and Weil, D. (2008). Translationally Repressed mRNA Transiently Cycles through Stress Granules during Stress. Mol. Biol. Cell *19*, 4469–4479.

Moreno, A.B., Martínez de Alba, A.E., Bardou, F., Crespi, M.D., Vaucheret, H., Maizel, A., and Mallory, A.C. (2013). Cytoplasmic and nuclear quality control and turnover of single-stranded RNA modulate post-transcriptional gene silencing in plants. Nucleic Acids Res. *41*, 4699–4708.

Morozov, I.Y., Jones, M.G., Razak, A.A., Rigden, D.J., and Caddick, M.X. (2010a). CUCU modification of mRNA promotes decapping and transcript degradation in Aspergillus nidulans. Mol. Cell. Biol. *30*, 460–469.

Morozov, I.Y., Jones, M.G., Spiller, D.G., Rigden, D.J., Dattenböck, C., Novotny, R., Strauss, J., and Caddick, M.X. (2010b). Distinct roles for Caf1, Ccr4, Edc3 and CutA in the co-ordination of transcript deadenylation, decapping and P-body formation in Aspergillus nidulans. Mol. Microbiol. *76*, 503–516.

Morozov, I.Y., Jones, M.G., Gould, P.D., Crome, V., Wilson, J.B., Hall, A.J.W.,

Rigden, D.J., and Caddick, M.X. (2012a). mRNA 3′ Tagging Is Induced by Nonsense-Mediated Decay and Promotes Ribosome Dissociation. Mol. Cell. Biol. *32*, 2585–2595.

Morozov, I.Y., Jones, M.G., Gould, P.D., Crome, V., Wilson, J.B., Hall, A.J.W., Rigden, D.J., and Caddick, M.X. (2012b). mRNA 3′ Tagging Is Induced by Nonsense-Mediated Decay and Promotes Ribosome Dissociation. Mol. Cell. Biol. *32*, 2585–2595.

Motamedi, M.R., Verdel, A., Colmenares, S.U., Gerber, S.A., Gygi, S.P., and Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. Cell *119*, 789–802.

Muench, D.G., Zhang, C., and Dahodwala, M. (2012). Control of cytoplasmic translation in plants. Wiley Interdiscip. Rev. RNA *3*, 178–194.

Muhlrad, D., and Parker, R. (2005). The yeast EDC1 mRNA undergoes deadenylation-independent decapping stimulated by Not2p, Not4p, and Not5p. EMBO J. *24*, 1033–1045.

Mullen, T.E., and Marzluff, W.F. (2008). Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. Genes Dev. *22*, 50–65.

Munoz-Tello, P., Gabus, C., and Thore, S. (2012). Functional implications from the Cid1 poly(U) polymerase crystal structure. Struct. Lond. Engl. 1993 *20*, 977–986.

Murata, T., Morita, N., Hikita, K., Kiuchi, K., Kiuchi, K., and Kaneda, N. (2005). Recruitment of mRNA-destabilizing protein TIS11 to stress granules is mediated by its zinc finger domain. Exp. Cell Res. *303*, 287–299.

Mustroph, A., Zanetti, M.E., Jang, C.J.H., Holtan, H.E., Repetti, P.P., Galbraith, D.W., Girke, T., and Bailey-Serres, J. (2009). Profiling translatomes of discrete cell populations resolves altered cellular priorities during hypoxia in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. *106*, 18843–18848.

Nadler-Holly, M., Breker, M., Gruber, R., Azia, A., Gymrek, M., Eisenstein, M., Willison, K.R., Schuldiner, M., and Horovitz, A. (2012). Interactions of subunit CCT3 in the yeast chaperonin CCT/TRiC with Q/N-rich proteins revealed by high-throughput microscopy analysis. Proc. Natl. Acad. Sci. *109*, 18833–18838.

Nagarajan, V.K., Jones, C.I., Newbury, S.F., and Green, P.J. (2013). XRN 5'→3' exoribonucleases: Structure, mechanisms and functions. Biochim. Biophys. Acta *1829*, 590–603.

Nakaminami, K., Matsui, A., Shinozaki, K., and Seki, M. (2012). RNA regulation in plant abiotic stress responses. Biochim. Biophys. Acta *1819*, 149–153.

Newbury, S., and Woollard, A. (2004). The 5'-3' exoribonuclease xrn-1 is essential for ventral epithelial enclosure during C. elegans embryogenesis. RNA New York N *10*, 59–65.

Nishimura, N., Kitahata, N., Seki, M., Narusaka, Y., Narusaka, M., Kuromori, T., Asami, T., Shinozaki, K., and Hirayama, T. (2005). Analysis of ABA hypersensitive germination2 revealed the pivotal functions of PARN in stress response in Arabidopsis. Plant J. Cell Mol. Biol. *44*, 972–984.

Nissan, T., Rajyaguru, P., She, M., Song, H., and Parker, R. (2010). Decapping activators in Saccharomyces cerevisiae act by multiple mechanisms. Mol. Cell *39*, 773–783.

Nykamp, K., Lee, M.-H., and Kimble, J. (2008). C. elegans La-related protein, LARP-1, localizes to germline P bodies and attenuates Ras-MAPK signaling during oogenesis. RNA *14*, 1378–1389.

Olmedo, G., Guo, H., Gregory, B.D., Nourizadeh, S.D., Aguilar-Henonin, L., Li, H., An, F., Guzman, P., and Ecker, J.R. (2006). ETHYLENE-INSENSITIVE5 encodes a 5'-->3' exoribonuclease required for regulation of the EIN3-targeting F-box proteins EBF1/2. Proc. Natl. Acad. Sci. U. S. A. *103*, 13286–13293.

Olszewska, M., Bujarski, J.J., and Kurpisz, M. (2012). P-bodies and their functions during mRNA cell cycle: Mini-review. Cell Biochem. Funct. *30*, 177–182.

Ooi, H.S., Kwo, C.Y., Wildpaner, M., Sirota, F.L., Eisenhaber, B., Maurer-Stroh, S., Wong, W.C., Schleiffer, A., Eisenhaber, F., and Schneider, G. (2009). ANNIE: integrated de novo protein sequence annotation. Nucleic Acids Res.

Otsuka, Y., Kedersha, N.L., and Schoenberg, D.R. (2009). Identification of a cytoplasmic complex that adds a cap onto 5'-monophosphate RNA. Mol. Cell. Biol. *29*, 2155–2167.

Ozgur, S., and Stoecklin, G. (2013). Role of Rck-Pat1b binding in assembly of processing-bodies. RNA Biol. *10*.

Palanivelu, R., Belostotsky, D.A., and Meagher, R.B. (2000). Arabidopsis thaliana poly (A) binding protein 2 (PAB2) functions in yeast translational and mRNA decay processes. Plant J. Cell Mol. Biol. *22*, 187–198.

Paquin, N., Ménade, M., Poirier, G., Donato, D., Drouet, E., and Chartrand, P. (2007). Local activation of yeast ASH1 mRNA translation through phosphorylation of Khd1p by the casein kinase Yck1p. Mol. Cell *26*, 795–809.

Parent, J.-S., Martínez de Alba, A.E., and Vaucheret, H. (2012). The origin and effect of small RNA signaling in plants. Front. Plant Sci. *3*, 179.

Parker, R. (2012). RNA degradation in Saccharomyces cerevisae. Genetics *191*, 671–702.

Parker, R., and Sheth, U. (2007). P Bodies and the Control of mRNA Translation and Degradation. Mol. Cell *25*, 635–646.

Paule, M.R., and White, R.J. (2000). Survey and summary: transcription by RNA polymerases I and III. Nucleic Acids Res. *28*, 1283–1298.

Pfeffer, S., Zavolan, M., Grässer, F.A., Chien, M., Russo, J.J., Ju, J., John, B., Enright, A.J., Marks, D., Sander, C., et al. (2004). Identification of virus-encoded microRNAs. Science *304*, 734–736.

Pfeffer, S., Sewer, A., Lagos-Quintana, M., Sheridan, R., Sander, C., Grässer, F.A., van Dyk, L.F., Ho, C.K., Shuman, S., Chien, M., et al. (2005). Identification of microRNAs of the herpesvirus family. Nat. Methods *2*, 269–276.

Pomeranz, M.C., Hah, C., Lin, P.-C., Kang, S.G., Finer, J.J., Blackshear, P.J., and

Jang, J.-C. (2010). The Arabidopsis tandem zinc finger protein AtTZF1 traffics between the nucleus and cytoplasmic foci and binds both DNA and RNA. Plant Physiol. *152*, 151–165.

Potuschak, T., Vansiri, A., Binder, B.M., Lechner, E., Vierstra, R.D., and Genschik, P. (2006). The exoribonuclease XRN4 is a component of the ethylene response pathway in Arabidopsis. Plant Cell *18*, 3047–3057.

Promponas, V.J., Enright, A.J., Tsoka, S., Kreil, D.P., Leroy, C., Hamodrakas, S., Sander, C., and Ouzounis, C.A. (2000). CAST: an iterative algorithm for the complexity analysis of sequence tracts. Complexity analysis of sequence tracts. Bioinforma. Oxf. Engl. *16*, 915–922.

Proudfoot, N.J. (2011). Ending the message: poly(A) signals then and now. Genes Dev. *25*, 1770–1782.

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., and Lopez, R. (2005). InterProScan: protein domains identifier. Nucleic Acids Res. *33*, W116–W120.

Raijmakers, R., Schilders, G., and Pruijn, G.J.M. (2004). The exosome, a molecular machine for controlled RNA degradation in both nucleus and cytoplasm. Eur. J. Cell Biol. *83*, 175–183.

Ramachandran, G.N. (1963). Protein Structure and Crystallography. Science *141*, 288–291.

RAMACHANDRAN, G.N., RAMAKRISHNAN, C., and SASISEKHARAN, V. (1963). Stereochemistry of polypeptide chain configurations. J. Mol. Biol. *7*, 95–99.

Rammelt, C., Bilen, B., Zavolan, M., and Keller, W. (2011). PAPD5, a noncanonical poly(A) polymerase with an unusual RNA-binding motif. RNA *17*, 1737–1746.

Rasmussen, T.P., and Culbertson, M.R. (1998). The putative nucleic acid helicase Sen1p is required for formation and stability of termini and for maximal rates of synthesis and levels of accumulation of small nucleolar RNAs in Saccharomyces cerevisiae. Mol. Cell. Biol. *18*, 6885–6896.

Read, R.L. (2002). Cytoplasmic poly(A) polymerases mediate cellular responses to S phase arrest. Proc. Natl. Acad. Sci. *99*, 12079–12084.

Reijns, M.A.M., Alexander, R.D., Spiller, M.P., and Beggs, J.D. (2008). A role for Q/N-rich aggregation-prone regions in P-body localization. J. Cell Sci. *121*, 2463–2472.

Reineke, L.C., and Lloyd, R.E. (2013). Diversion of stress granules and P-bodies during viral infection. Virology *436*, 255–267.

Reineke, L.C., Dougherty, J.D., Pierre, P., and Lloyd, R.E. (2012). Large G3BP-induced granules trigger eIF2α phosphorylation. Mol. Biol. Cell *23*, 3499–3510.

Ren, G., Chen, X., and Yu, B. (2012). Uridylation of miRNAs by hen1 suppressor1 in Arabidopsis. Curr. Biol. CB *22*, 695–700.

Reverdatto, S.V., Dutko, J.A., Chekanova, J.A., Hamilton, D.A., and Belostotsky, D.A. (2004). mRNA deadenylation by PARN is essential for embryogenesis in higher plants. RNA New York N *10*, 1200–1214.

Ringpis, G.-E., Stagno, J., and Aphasizhev, R. (2010). Mechanism of U-Insertion RNA Editing in Trypanosome Mitochondria: Characterization of RET2 Functional Domains by Mutational Analysis. J. Mol. Biol. *399*, 696–706.

Rissland, O.S., and Norbury, C.J. (2008). The Cid1 poly (U) polymerase. Biochim. Biophys. Acta BBA-Gene Regul. Mech. *1779*, 286–294.

Rissland, O.S., and Norbury, C.J. (2009). Decapping is preceded by 3' uridylation in a novel pathway of bulk mRNA turnover. Nat. Struct. Mol. Biol. *16*, 616–623.

Rissland, O.S., Mikulasova, A., and Norbury, C.J. (2007). Efficient RNA polyuridylation by noncanonical poly(A) polymerases. Mol. Cell. Biol. *27*, 3612–3624.

Rost, B., and Sander, C. (1993). Prediction of protein secondary structure at better than 70% accuracy. J. Mol. Biol. *232*, 584–599.

Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H., and Bartel, D.P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. Cell *127*, 1193–1207.

Sáez-Vasquez, J., Caparros-Ruiz, D., Barneche, F., and Echeverría, M. (2004). A plant snoRNP complex containing snoRNAs, fibrillarin, and nucleolin-like proteins is competent for both rRNA gene binding and pre-rRNA processing in vitro. Mol. Cell. Biol. *24*, 7284–7297.

Saito, K., Sakaguchi, Y., Suzuki, T., Suzuki, T., Siomi, H., and Siomi, M.C. (2007). Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi- interacting RNAs at their 3' ends. Genes Dev. *21*, 1603–1608.

Saitoh, S., Chabes, A., McDonald, W.H., Thelander, L., Yates, J.R., and Russell, P. (2002). Cid13 is a cytoplasmic poly(A) polymerase that regulates ribonucleotide reductase mRNA. Cell *109*, 563–573.

Samaha, H., Delorme, V., Pontvianne, F., Cooke, R., Delalande, F., Van Dorsselaer, A., Echeverria, M., and Sáez-Vásquez, J. (2010). Identification of protein factors and U3 snoRNAs from a Brassica oleracea RNP complex involved in the processing of pre-rRNA. Plant J. Cell Mol. Biol. *61*, 383–398.

San Paolo, S., Vanacova, S., Schenk, L., Scherrer, T., Blank, D., Keller, W., and Gerber, A.P. (2009). Distinct roles of non-canonical poly(A) polymerases in RNA metabolism. PLoS Genet. *5*, e1000555.

Sànchez, R., and Marzluff, W.F. (2002). The stem-loop binding protein is required for efficient translation of histone mRNA in vivo and in vitro. Mol. Cell. Biol. *22*, 7093–7104.

Sawarkar, R., Sievers, C., and Paro, R. (2012). Hsp90 Globally Targets Paused RNA Polymerase to Regulate Gene Expression in Response to Environmental Stimuli. Cell *149*, 807–818.

Schlessinger, A., Schaefer, C., Vicedo, E., Schmidberger, M., Punta, M., and Rost, B. (2011). Protein disorder — a breakthrough invention of evolution? Curr. Opin. Struct. Biol. *21*, 412–418.

Schmidt, M.-J., West, S., and Norbury, C.J. (2010). The human cytoplasmic RNA

terminal U-transferase ZCCHC11 targets histone mRNAs for degradation. RNA *17*, 39–44.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J., Cheng, J., et al. (2010). Genome sequence of the palaeopolyploid soybean. Nature *463*, 178–183.

Schneider, T.D., and Stephens, R.M. (1990). Sequence logos: a new way to display consensus sequences. Nucleic Acids Res. *18*, 6097–6100.

Schneider, C., Kudla, G., Wlotzka, W., Tuck, A., and Tollervey, D. (2012). Transcriptome-wide Analysis of Exosome Targets. Mol. Cell *48*, 422–433.

Schwartz, D.C., and Parker, R. (2000). mRNA decapping in yeast requires dissociation of the cap binding protein, eukaryotic translation initiation factor 4E. Mol. Cell. Biol. *20*, 7933–7942.

Schwede, T., Kopp, J., Guex, N., and Peitsch, M.C. (2003). SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. *31*, 3381–3385.

Scott, D.D., and Norbury, C.J. (2013). RNA decay via 3′ uridylation. Biochim. Biophys. Acta BBA - Gene Regul. Mech. *1829*, 654–665.

Seitz, H., Ghildiyal, M., and Zamore, P.D. (2008). Argonaute loading improves the 5' precision of both MicroRNAs and their miRNA* strands in flies. Curr. Biol. CB *18*, 147–151.

Seiwert, S.D., Heidmann, S., and Stuart, K. (1996). Direct visualization of uridylate deletion in vitro suggests a mechanism for kinetoplastid RNA editing. Cell *84*, 831–841.

Sement, F.M., Ferrier, E., Zuber, H., Merret, R., Alioua, M., Deragon, J.-M., Bousquet-Antonelli, C., Lange, H., and Gagliardi, D. (2013). Uridylation prevents 3' trimming of oligoadenylated mRNAs. Nucleic Acids Res.

Serman, A., Le Roy, F., Aigueperse, C., Kress, M., Dautry, F., and Weil, D. (2007). GW body disassembly triggered by siRNAs independently of their silencing activity. Nucleic Acids Res. *35*, 4715–4727.

Shchepachev, V., and Azzalin, C.M. (2013). The Mpn1 RNA exonuclease: Cellular functions and implication in disease. FEBS Lett.

Shchepachev, V., Wischnewski, H., Missiaglia, E., Soneson, C., and Azzalin, C.M. (2012). Mpn1, Mutated in Poikiloderma with Neutropenia Protein 1, Is a Conserved 3′-to-5′ RNA Exonuclease Processing U6 Small Nuclear RNA. Cell Reports *2*, 855–865.

Shen, B., and Goodman, H.M. (2004). Uridine addition after microRNA-directed cleavage. Science *306*, 997.

Sheth, U., and Parker, R. (2003). Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. Science *300*, 805–808.

Siddique, M., Gernhard, S., von Koskull-Doring, P., Vierling, E., and Scharf, K.-D. (2008). The plant sHSP superfamily: five new members in Arabidopsis thaliana with unexpected properties. Cell Stress Chaperones *13*, 183–197.

Silanes, I.L. de, Galbán, S., Martindale, J.L., Yang, X., Mazan-Mamczarz, K., Indig, F.E., Falco, G., Zhan, M., and Gorospe, M. (2005). Identification and Functional Outcome of mRNAs Associated with RNA-Binding Protein TIA-1. Mol. Cell. Biol. *25*, 9520–9531.

Sinturel, F., Bréchemier-Baey, D., Kiledjian, M., Condon, C., and Bénard, L. (2012). Activation of 5'-3' exoribonuclease Xrn1 by cofactor Dcs1 is essential for mitochondrial function in yeast. Proc. Natl. Acad. Sci. U. S. A. *109*, 8264–8269.

Solomon, S., Xu, Y., Wang, B., David, M.D., Schubert, P., Kennedy, D., and Schrader, J.W. (2007). Distinct structural features of caprin-1 mediate its interaction with G3BP-1 and its induction of phosphorylation of eukaryotic translation initiation factor 2alpha, entry to cytoplasmic stress granules, and selective interaction with a subset of mRNAs. Mol. Cell. Biol. *27*, 2324–2342.

Song, M.-G., and Kiledjian, M. (2007). 3' Terminal oligo U-tract-mediated stimulation of decapping. RNA New York N *13*, 2356–2365.

Song, M.-G., Li, Y., and Kiledjian, M. (2010). Multiple mRNA decapping enzymes in mammalian cells. Mol. Cell *40*, 423–432.

Souquere, S., Mollet, S., Kress, M., Dautry, F., Pierron, G., and Weil, D. (2009). Unravelling the ultrastructure of stress granules and associated P-bodies in human cells. J. Cell Sci. *122*, 3619–3626.

Souret, F.F., Kastenmayer, J.P., and Green, P.J. (2004). AtXRN4 degrades mRNA in Arabidopsis and its substrates include selected miRNA targets. Mol. Cell *15*, 173–183.

Spiller, M.P., Boon, K.-L., Reijns, M.A.M., and Beggs, J.D. (2007). The Lsm2-8 complex determines nuclear localization of the spliceosomal U6 snRNA. Nucleic Acids Res. *35*, 923–929.

Staals, R.H.J., Bronkhorst, A.W., Schilders, G., Slomovic, S., Schuster, G., Heck, A.J.R., Raijmakers, R., and Pruijn, G.J.M. (2010). Dis3-like 1: a novel exoribonuclease associated with the human exosome. EMBO J. *29*, 2358–2367.

Stagno, J., Aphasizheva, I., Aphasizhev, R., and Luecke, H. (2007). Dual role of the RNA substrate in selectivity and catalysis by terminal uridylyl transferases. Proc. Natl. Acad. Sci. *104*, 14634.

Stagno, J., Aphasizheva, I., Bruystens, J., Luecke, H., and Aphasizhev, R. (2010). Structure of the mitochondrial editosome-like complex associated TUTase 1 reveals divergent mechanisms of UTP selection and domain organization. J. Mol. Biol. *399*, 464–475.

Steiger, M., Carr-Schmid, A., Schwartz, D.C., Kiledjian, M., and Parker, R. (2003). Analysis of recombinant yeast decapping enzyme. RNA New York N *9*, 231–238.

Steinmetz, E.J., Conrad, N.K., Brow, D.A., and Corden, J.L. (2001). RNA-binding protein Nrd1 directs poly(A)-independent 3′-end formation of RNA polymerase II transcripts. Nature *413*, 327–331.

Stevens, A. (2001). 5'-exoribonuclease 1: Xrn1. Methods Enzymol. *342*, 251–259.

Stevenson, A.L., and Norbury, C.J. (2006). The Cid1 family of non-canonical poly(A) polymerases. Yeast Chichester Engl. *23*, 991–1000.

Stoecklin, G., Mayo, T., and Anderson, P. (2006). ARE-mRNA degradation requires the 5?-3? decay pathway. EMBO Rep. *7*, 72–77.

Sun, X., Rikkerink, E.H.A., Jones, W.T., and Uversky, V.N. (2013). Multifarious Roles

of Intrinsic Disorder in Proteins Illustrate Its Broad Impact on Plant Biology. Plant Cell Online *25*, 38–55.

Sweet, T., Kovalak, C., and Coller, J. (2012). The DEAD-box protein Dhh1 promotes decapping by slowing ribosome movement. PLoS Biol. *10*, e1001342.

Sweet, T.J., Boyer, B., Hu, W., Baker, K.E., and Coller, J. (2007). Microtubule disruption stimulates P-body formation. RNA *13*, 493–502.

Talavera, G., and Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst. Biol. *56*, 564–577.

Tan, D., Marzluff, W.F., Dominski, Z., and Tong, L. (2013). Structure of histone mRNA stem-loop, human stem-loop binding protein, and 3'hExo ternary complex. Science *339*, 318–321.

Teixeira, D., Sheth, U., Valencia-Sanchez, M.A., Brengues, M., and Parker, R. (2005). Processing bodies require RNA for assembly and contain nontranslating mRNAs. RNA New York N *11*, 371–382.

Tharun, S. (2009). Lsm1-7-Pat1 complex: a link between 3' and 5'-ends in mRNA decay? RNA Biol. *6*, 228–232.

Tharun, S., and Parker, R. (2001). Targeting an mRNA for Decapping. Mol. Cell *8*, 1075–1083.

Tharun, S., He, W., Mayes, A.E., Lennertz, P., Beggs, J.D., and Parker, R. (2000). Yeast Sm-like proteins function in mRNA decapping and decay. Nature *404*, 515–518.

Thiebaut, M., Colin, J., Neil, H., Jacquier, A., Séraphin, B., Lacroute, F., and Libri, D. (2008). Futile Cycle of Transcription Initiation and Termination Modulates the Response to Nucleotide Shortage in S. cerevisiae. Mol. Cell *31*, 671–682.

Thomas, M.G., Loschi, M., Desbats, M.A., and Boccaccio, G.L. (2011a). RNA granules: The good, the bad and the ugly. Cell. Signal. *23*, 324–334.

Thomas, M.G., Loschi, M., Desbats, M.A., and Boccaccio, G.L. (2011b). RNA granules: The good, the bad and the ugly. Cell. Signal. *23*, 324–334.

Thore, S., Mauxion, F., Séraphin, B., and Suck, D. (2003). X-ray structure and activity of the yeast Pop2 protein: a nuclease subunit of the mRNA deadenylase complex. EMBO Rep. *4*, 1150–1155.

Thornton, J.E., Chang, H.-M., Piskounova, E., and Gregory, R.I. (2012). Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). RNA New York N *18*, 1875–1885.

Todeschini, A.-L., Condon, C., and Bénard, L. (2006). Sodium-induced GCN4 expression controls the accumulation of the 5' to 3' RNA degradation inhibitor, 3'-phosphoadenosine 5'-phosphate. J. Biol. Chem. *281*, 3276–3282.

Tomecki, R., Kristiansen, M.S., Lykke-Andersen, S., Chlebowski, A., Larsen, K.M., Szczesny, R.J., Drazkowska, K., Pastula, A., Andersen, J.S., Stepien, P.P., et al. (2010). The human core exosome interacts with differentially localized processive RNases: hDIS3 and

hDIS3L. EMBO J. *29*, 2342–2357.

Tomek, W., and Wollenhaupt, K. (2012). The “closed loop model” in controlling mRNA translation during development. Anim. Reprod. Sci. *134*, 2–8.

Tompa, P. (2005). The interplay between structure and function in intrinsically unstructured proteins. FEBS Lett. *579*, 3346–3354.

Tourrière, H., Chebli, K., Zekri, L., Courselaud, B., Blanchard, J.M., Bertrand, E., and Tazi, J. (2003). The RasGAP-associated endoribonuclease G3BP assembles stress granules. J. Cell Biol. *160*, 823–831.

Treiber, T., Treiber, N., and Meister, G. (2012). Regulation of microRNA biogenesis and function. Thromb. Haemost. *107*, 605–610.

Trippe, R., Sandrock, B., and Benecke, B.J. (1998). A highly specific terminal uridylyl transferase modifies the 3’-end of U6 small nuclear RNA. Nucleic Acids Res. *26*, 3119–3126.

Trippe, R., Guschina, E., Hossbach, M., Urlaub, H., Lührmann, R., and Benecke, B.-J. (2006). Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. RNA *12*, 1494–1504.

Tritschler, F., Eulalio, A., Truffault, V., Hartmann, M.D., Helms, S., Schmidt, S., Coles, M., Izaurralde, E., and Weichenrieder, O. (2007). A divergent Sm fold in EDC3 proteins mediates DCP1 binding and P-body targeting. Mol. Cell. Biol. *27*, 8600–8611.

Tucker, M., Valencia-Sanchez, M.A., Staples, R.R., Chen, J., Denis, C.L., and Parker, R. (2001). The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in Saccharomyces cerevisiae. Cell *104*, 377–386.

Tucker, M., Staples, R.R., Valencia-Sanchez, M.A., Muhlrad, D., and Parker, R. (2002). Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in Saccharomyces cerevisiae. EMBO J. *21*, 1427–1436.

Tuskan, G.A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., et al. (2006). The genome of black cottonwood, Populus trichocarpa (Torr. & Gray). Science *313*, 1596–1604.

Udagawa, T., Swanger, S.A., Takeuchi, K., Kim, J.H., Nalavadi, V., Shin, J., Lorenz, L.J., Zukin, R.S., Bassell, G.J., and Richter, J.D. (2012). Bidirectional control of mRNA translation and synaptic plasticity by the cytoplasmic polyadenylation complex. Mol. Cell *47*, 253–266.

Ursic, D., Himmel, K.L., Gurley, K.A., Webb, F., and Culbertson, M.R. (1997). The yeast SEN1 gene is required for the processing of diverse RNA classes. Nucleic Acids Res. *25*, 4778–4785.

Uversky, V.N. (2013a). Unusual biophysics of intrinsically disordered proteins. Biochim. Biophys. Acta BBA - Proteins Proteomics *1834*, 932–951.

Uversky, V.N. (2013b). A decade and a half of protein intrinsic disorder: Biology still waits for physics. Protein Sci. *22*, 693–724.

Uversky, V.N. (2013c). A decade and a half of protein intrinsic disorder: Biology still waits for physics. Protein Sci. *22*, 693–724.

Uversky, V.N., and Dunker, A.K. (2013). The case for intrinsically disordered proteins playing contributory roles in molecular recognition without a stable 3D structure. F1000 Biol. Reports *5*.

Vacic, V., Uversky, V.N., Dunker, A.K., and Lonardi, S. (2007). Composition Profiler: a tool for discovery and visualization of amino acid composition differences. BMC Bioinformatics *8*, 211.

Valiente-Echeverría, F., Melnychuk, L., and Mouland, A.J. (2012). Viral modulation of stress granules. Virus Res. *169*, 430–437.

Vessey, J.P., Vaccani, A., Xie, Y., Dahm, R., Karra, D., Kiebler, M.A., and Macchi, P. (2006). Dendritic localization of the translational repressor Pumilio 2 and its contribution to dendritic stress granules. J. Neurosci. Off. J. Soc. Neurosci. *26*, 6496–6508.

Villalba, A., Coll, O., and Gebauer, F. (2011). Cytoplasmic polyadenylation and translational control. Curr. Opin. Genet. Dev. *21*, 452–457.

Virtanen, A., Henriksson, N., Nilsson, P., and Nissbeck, M. (2013). Poly(A)-specific ribonuclease (PARN): an allosterically regulated, processive and mRNA cap-interacting deadenylase. Crit. Rev. Biochem. Mol. Biol. *48*, 192–209.

Voinnet, O. (2008). Post-transcriptional RNA silencing in plant-microbe interactions: a touch of robustness and versatility. Curr. Opin. Plant Biol. *11*, 464–470.

Voinnet, O. (2009). Origin, Biogenesis, and Activity of Plant MicroRNAs. Cell *136*, 669–687.

Wächter, K., Köhn, M., Stöhr, N., and Hüttelmaier, S. (2013). Subcellular localization and RNP formation of IGF2BPs (IGF2 mRNA binding proteins) is modulated by distinct RNA-binding domains. Biol. Chem.

Walley, J.W., Kelley, D.R., Nestorova, G., Hirschberg, D.L., and Dehesh, K. (2010). Arabidopsis deadenylases AtCAF1a and AtCAF1b play overlapping and distinct roles in mediating environmental stress responses. Plant Physiol. *152*, 866–875.

Wang, S.W., Toda, T., MacCallum, R., Harris, A.L., and Norbury, C. (2000). Cid1, a fission yeast protein required for S-M checkpoint control when DNA polymerase delta or epsilon is inactivated. Mol. Cell. Biol. *20*, 3234–3244.

Wang, Z., Jiao, X., Carr-Schmid, A., and Kiledjian, M. (2002). The hDcp2 protein is a mammalian mRNA decapping enzyme. Proc. Natl. Acad. Sci. U. S. A. *99*, 12663–12668.

Weber, S.C., and Brangwynne, C.P. (2012). Getting RNA and protein in phase. Cell *149*, 1188–1191.

Weber, C., Nover, L., and Fauth, M. (2008). Plant stress granules and mRNA processing bodies are distinct from heat stress granules. Plant J. *56*, 517–530.

Weil, D., and Hollien, J. (2013). Cytoplasmic organelles on the road to mRNA decay. Biochim. Biophys. Acta *1829*, 725–731.

Weill, L., Belloc, E., Bava, F.-A., and Méndez, R. (2012). Translational control by changes in poly(A) tail length: recycling mRNAs. Nat. Struct. Mol. Biol. *19*, 577–585.

Wen, W.-L., Stevenson, A.L., Wang, C.-Y., Chen, H.-J., Kearsey, S.E., Norbury, C.J.,

Watt, S., Bähler, J., and Wang, S.-W. (2010). Vgl1, a multi-KH domain protein, is a novel component of the fission yeast stress granules required for cell survival under thermal stress. Nucleic Acids Res. *38*, 6555–6566.

Whipple, J.M., Lane, E.A., Chernyakov, I., D'Silva, S., and Phizicky, E.M. (2011). The yeast rapid tRNA decay pathway primarily monitors the structural integrity of the acceptor and T-stems of mature tRNA. Genes Dev. *25*, 1173–1184.

Wilczynska, A., Aigueperse, C., Kress, M., Dautry, F., and Weil, D. (2005). The translational regulator CPEB1 provides a link between dcp1 bodies and stress granules. J. Cell Sci. *118*, 981–992.

Williams, R.M., Obradovi, Z., Mathura, V., Braun, W., Garner, E.C., Young, J., Takayama, S., Brown, C.J., and Dunker, A.K. (2001). The protein non-folding problem: amino acid determinants of intrinsic order and disorder. Pac. Symp. Biocomput. Pac. Symp. Biocomput. 89–100.

Wilusz, C.J., and Wilusz, J. (2013). Lsm proteins and Hfq: Life at the 3' end. RNA Biol. *10*, 592–601.

Wilusz, J.E., Whipple, J.M., Phizicky, E.M., and Sharp, P.A. (2011). tRNAs marked with CCACCA are targeted for degradation. Science *334*, 817–821.

Win, T.Z., Draper, S., Read, R.L., Pearce, J., Norbury, C.J., and Wang, S.-W. (2006a). Requirement of Fission Yeast Cid14 in Polyadenylation of rRNAs. Mol. Cell. Biol. *26*, 1710–1721.

Win, T.Z., Stevenson, A.L., and Wang, S.-W. (2006b). Fission yeast Cid12 has dual functions in chromosome segregation and checkpoint control. Mol. Cell. Biol. *26*, 4435–4447.

Wippich, F., Bodenmiller, B., Trajkovska, M.G., Wanka, S., Aebersold, R., and Pelkmans, L. (2013). Dual specificity kinase DYRK3 couples stress granule condensation/dissolution to mTORC1 signaling. Cell *152*, 791–805.

Van Wolfswinkel, J.C., Claycomb, J.M., Batista, P.J., Mello, C.C., Berezikov, E., and Ketting, R.F. (2009). CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs. Cell *139*, 135–148.

Wolkowicz, U.M., and Cook, A.G. (2012). NF45 dimerizes with NF90, Zfr and SPNR via a conserved domain that has a nucleotidyltransferase fold. Nucleic Acids Res. *40*, 9356–9368.

Wyman, S.K., Knouf, E.C., Parkin, R.K., Fritz, B.R., Lin, D.W., Dennis, L.M., Krouse, M.A., Webster, P.J., and Tewari, M. (2011). Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity. Genome Res. *21*, 1450–1461.

Xu, J., and Chua, N.-H. (2009). Arabidopsis decapping 5 is required for mRNA decapping, P-body formation, and translational repression during postembryonic development. Plant Cell *21*, 3270–3279.

Xu, J., and Chua, N.-H. (2011a). Processing bodies and plant development. Curr. Opin. Plant Biol. *14*, 88–93.

Xu, J., and Chua, N.-H. (2011b). Processing bodies and plant development. Curr. Opin. Plant Biol. *14*, 88–93.

Xu, J., and Chua, N.-H. (2012). Dehydration stress activates Arabidopsis MPK6 to signal DCP1 phosphorylation. EMBO J. *31*, 1975–1984.

Xu, J., Yang, J.-Y., Niu, Q.-W., and Chua, N.-H. (2006). Arabidopsis DCP2, DCP1, and VARICOSE Form a Decapping Complex Required for Postembryonic Development. PLANT CELL ONLINE *18*, 3386–3398.

Xue, D., Rubinson, D.A., Pannone, B.K., Yoo, C.J., and Wolin, S.L. (2000). U snRNP assembly in yeast involves the La protein. EMBO J. *19*, 1650–1660.

Yang, W.-H., Yu, J.H., Gulick, T., Bloch, K.D., and Bloch, D.B. (2006a). RNA-associated protein 55 (RAP55) localizes to mRNA processing bodies and stress granules. RNA New York N *12*, 547–554.

Yang, X., Purdy, M., Marzluff, W.F., and Dominski, Z. (2006b). Characterization of 3'hExo, a 3' exonuclease specifically interacting with the 3' end of histone mRNA. J. Biol. Chem. *281*, 30447–30454.

Yang, Z., Jakymiw, A., Wood, M.R., Eystathioy, T., Rubin, R.L., Fritzler, M.J., and Chan, E.K.L. (2004). GW182 is critical for the stability of GW bodies expressed during the cell cycle and cell proliferation. J. Cell Sci. *117*, 5567–5578.

Yates, L.A., Fleurdépine, S., Rissland, O.S., De Colibus, L., Harlos, K., Norbury, C.J., and Gilbert, R.J.C. (2012). Structural basis for the activity of a cytoplasmic RNA terminal uridylyl transferase. Nat. Struct. Mol. Biol. *19*, 782–787.

Yates, L.A., Norbury, C.J., and Gilbert, R.J.C. (2013). The Long and Short of MicroRNA. Cell *153*, 516–519.

Yekta, S., Shih, I.-H., and Bartel, D.P. (2004). MicroRNA-directed cleavage of HOXB8 mRNA. Science *304*, 594–596.

Yu, B., Bi, L., Zhai, J., Agarwal, M., Li, S., Wu, Q., Ding, S.-W., Meyers, B.C., Vaucheret, H., and Chen, X. (2010). siRNAs compete with miRNAs for methylation by HEN1 in Arabidopsis. Nucleic Acids Res. *38*, 5844–5850.

Yu, J.H., Yang, W.-H., Gulick, T., Bloch, K.D., and Bloch, D.B. (2005). Ge-1 is a central component of the mammalian cytoplasmic mRNA processing body. RNA New York N *11*, 1795–1802.

Zabolotskaya, M.V., Grima, D.P., Lin, M.-D., Chou, T.-B., and Newbury, S.F. (2008). The 5'-3' exoribonuclease Pacman is required for normal male fertility and is dynamically localized in cytoplasmic particles in Drosophila testis cells. Biochem. J. *416*, 327–335.

Zakrzewska-Placzek, M., Souret, F.F., Sobczyk, G.J., Green, P.J., and Kufel, J. (2010). Arabidopsis thaliana XRN2 is required for primary cleavage in the pre-ribosomal RNA. Nucleic Acids Res. *38*, 4487–4502.

Zanetti, M.E., Chang, I.-F., Gong, F., Galbraith, D.W., and Bailey-Serres, J. (2005). Immunopurification of polyribosomal complexes of Arabidopsis for global analysis of gene expression. Plant Physiol. *138*, 624–635.

Zdobnov, E.M., and Apweiler, R. (2001). InterProScan - an integration platform for the signature-recognition methods in InterPro. Bioinformatics *17*, 847–848.

Zhai, J., and Meyers, B.C. (2013). Deep Sequencing from hen1 Mutants to Identify Small RNA 3' Modifications. Cold Spring Harb. Symp. Quant. Biol.

Zhai, J., Zhao, Y., Simon, S.A., Huang, S., Petsch, K., Arikit, S., Pillay, M., Ji, L., Xie, M., Cao, X., et al. (2013). Plant MicroRNAs Display Differential 3′ Truncation and Tailing Modifications That Are ARGONAUTE1 Dependent and Conserved Across Species. Plant Cell Online.

Zhang, W., Murphy, C., and Sieburth, L.E. (2010a). Conserved RNaseII domain protein functions in cytoplasmic mRNA decay and suppresses Arabidopsis decapping mutant phenotypes. Proc. Natl. Acad. Sci. U. S. A. *107*, 15981–15985.

Zhang, X., Virtanen, A., and Kleiman, F.E. (2010b). To polyadenylate or to deadenylate: that is the question. Cell Cycle Georget. Tex *9*, 4437–4449.

Zhao, Y., Yu, Y., Zhai, J., Ramachandran, V., Dinh, T.T., Meyers, B.C., Mo, B., and Chen, X. (2012a). The Arabidopsis Nucleotidyl Transferase HESO1 Uridylates Unmethylated Small RNAs to Trigger Their Degradation. Curr. Biol. *22*, 689–694.

Zhao, Y., Mo, B., and Chen, X. (2012b). Mechanisms that impact microRNA stability in plants. RNA Biol. *9*, 1218–1223.

Zimmer, S.L., Schein, A., Zipor, G., Stern, D.B., and Schuster, G. (2009). Polyadenylation in Arabidopsis and Chlamydomonas organelles: the input of nucleotidyltransferases, poly(A) polymerases and polynucleotide phosphorylase. Plant J. *59*, 88–99.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. *31*, 3406–3415.

THESE DE EMILIE FERRIER

ROLE ET MODE D'ACTION DE L'UTP:RNA URIDYLYLTRANSFERASE URT1 DANS L'URIDYLATION ET LA DEGRADATION DES ARNM CHEZ *ARABIDOPSIS THALIANA*

Résumé :

La dégradation des ARN est un mécanisme essentiel à la régulation de l'expression des génomes. L'importance de l'uridylation dans les mécanismes de dégradation des ARN commence juste à être appréciée. Cette thèse présente l'étude d'une uridylyltransférase URT1 (*UTP :RNA Uridylyltransferase 1*) et de son rôle dans la dégradation des ARN chez *Arabidopsis thaliana*. L'étude des propriétés catalytiques de URT1 montre que cette uridylyltransférase est intrinsèquement et strictement spécifique des UTP et est distributive pour les premières uridines ajoutées. URT1 est responsable *in vivo* de l'uridylation des ARNm après une étape de déadénylation. Cette uridylation des ARNm oligoadénylés protège leur extrémité 3' contre des attaques ribonucléolytiques et polarise la dégradation de 5' en 3'. Par ailleurs, cet effet protecteur est retrouvé au niveau des ARNm associés aux polysomes, ce qui pourrait être important dans le cas d'une dégradation co-traductionnelle des ARNm. Enfin, URT1 est localisée dans le cytosol au niveau des granules de stress et des *processing bodies.* Le mécanisme d'adressage de URT1 dans les *processing bodies* implique une partie de la région N terminale prédite comme intrinsèquement désorganisée, alors que le domaine nucléotidyltransférase C-terminal semble suffisant pour permettre l'adressage de URT1 au niveau des processing bodies et granules de stress en réponse à un stress thermique.

Ces travaux de thèse ont permis de mieux comprendre les mécanismes et les rôles de l'uridylation dans la dégradation des ARNm chez Arabidopsis. Ils ouvrent des perspectives dans l'étude d'autres fonctions de l'uridylation comme l'inhibition de la traduction.

Abstract :

RNA degradation is an essential mechanism for the regulation of genome expression. The importance of uridylation for RNA degradation is just emerging. This thesis presents the study of URT1 (UTP :RNA Uridylyltransferase 1) and its role in RNA degradation in *Arabidopsis thaliana*. URT1 is an uridylyltransferase intrinsically and strictly specific for UTP and is distributive for the first nucleotides added. URT1 uridylates mRNA *in vivo* after a deadenylation step. This uridylation protects mRNA's 3' end from further attacks and polarise degradation in the 5' to 3' direction. This protection of 3' ends by uridylation and its conferred polarity of 5' to 3' degradation are also detected in polysomes. Uridylation is therefore likely important in case of co-translational degradation of mRNAs. A region in URT1's N terminal region predicted to be intrinsically disorganised is required for addressing URT1 to processing bodies. However, following heat shock, the nucleotidyltransferase domain present in the C terminal region of URT1 is sufficient to address URT1 to both *processing bodies* and stress granules,

This work contributes to a better understanding of the mechanisms and roles of uridylation in RNA degradation in Arabidopsis thaliana. These results also open perspectives for studying other functions of uridylation such as translation inhibition.