



**HAL**  
open science

# Mécanismes de la dégradation des ARN dépendants ou indépendants de la déadénylation

Léna Audebert

► **To cite this version:**

Léna Audebert. Mécanismes de la dégradation des ARN dépendants ou indépendants de la déadénylation. Biochimie, Biologie Moléculaire. Sorbonne Université, 2022. Français. NNT : 2022SORUS374 . tel-03943917

**HAL Id: tel-03943917**

**<https://theses.hal.science/tel-03943917>**

Submitted on 17 Jan 2023

**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.



Sorbonne Université

ED515 - Complexité du vivant

Génétique des Interactions Macromoléculaires - UMR 3525, Institut Pasteur

## Mécanismes de la dégradation des ARN dépendants ou indépendants de la déadénylation

Par Léna Audebert

Thèse de doctorat de Biochimie et biologie structurale

Dirigée par Cosmin Saveanu

Présentée et soutenue publiquement le 16 septembre 2022

Devant un jury composé de :

Saveanu, Cosmin, Directeur de recherche Institut Pasteur, Directeur de thèse

Bousquet-Antonelli, Cécile, Directrice de recherche CNRS, Rapporteur

Libri, Domenico, Directeur de recherche CNRS, Rapporteur

Torchet, Claire, Chargée de recherche CNRS - HDR, Examinatrice

Namy, Olivier, Directeur de recherche CNRS, Examinateur, Président du jury



*pour Alexis,*



## Remerciements

J'aimerais remercier toutes les personnes qui m'ont soutenue et ont contribué de près ou de loin à mon travail de thèse.

Je remercie Cécile Bousquet-Antonelli et Domenico Libri d'avoir accepté d'être rapporteurs de mon travail. Merci à Claire Torchet d'avoir accepté mon invitation pour faire partie de mon jury de thèse. Merci à Olivier Namy d'avoir été membre de mon jury et d'avoir présidé ma soutenance. Cela a été un grand honneur et plaisir de pouvoir échanger avec vous.

Je remercie Cosmin Saveanu de m'avoir donné l'opportunité de réaliser une thèse sous sa direction. Merci notamment pour tous les conseils avisés, les nombreux partages bibliographiques ou encore l'initiation à l'analyse bio-informatique. Merci également d'avoir été disponible au cours de ces années, que ce soit pour répondre à mes questions ou discuter du projet.

Je remercie Dominique Weil, Lionel Bénard et Benoît Letallec pour leurs précieux conseils et encouragements au cours de mes comités de thèse.

Je remercie l'ANR, la Fondation ARC pour la recherche sur le cancer, Sorbonne Université, et l'Institut Pasteur pour le financement de ma thèse.

Merci à toutes les personnes du 4<sup>ème</sup> étage du bâtiment Fernbach pour leur gentillesse et leur soutien au cours de ma thèse, cela a été un plaisir de pouvoir évoluer à vos côtés.

Merci aux membres, passés et présents du laboratoire Génétique des Interactions Macromoléculaires. Merci à Alain Jacquier de m'avoir accueillie au sein de son laboratoire, merci notamment de m'avoir permis de présenter mon travail à un congrès international. Merci à Micheline Fromont-Racine pour ses encouragements. Merci à Gwenaël Badis-Bréard qui m'a permis de faire mon stage de master au sein du laboratoire et qui a co-encadré ma première année de thèse.

Merci à Varun Khanna et Emmanuelle Permal pour le soutien, la formation et l'aide à l'analyse bio-informatique qui m'ont été précieux. Merci à Abdelkader Namane et Laurence Decourty, notamment pour les données de spectrométrie de masse. Merci à Lucia Oreus pour tous les précieux milieux de culture réalisés.

Merci à Agathe Gilbert et Estelle Dacheux pour leur soutien au cours de ma thèse.

Merci à Frank Feuerbach-Fournier pour les nombreuses discussions et partages scientifiques ainsi que pour les innombrables expériences ensemble. Merci surtout de m'avoir sans relâche encouragée.

À mes amies et amis, merci. Une pensée pour Léa.

Merci à toute ma famille, une pensée pour ma mère ainsi que mon père, mon grand frère Alexis, mon petit frère Léo, mon cousin Sylvain, mes tantes Anne-Marie et Martine. À ma grande sœur Marine qui m'a sans cesse inspirée, merci.



# Abstract

RNA degradation is a critical step in maintaining the proper balance required for gene expression and its regulation. In eukaryotes, RNA decay is thought to be controlled by the speed of poly-A tail shortening. Below a threshold, the presence of an oligo-A tail leads to decapping activation and rapid RNA degradation. However, progressive deadenylation is not required for a major fast RNA degradation pathway, the nonsense mediated mRNA decay (NMD), which occurs independently of the size of a poly-A tail. Based on large-scale poly-A tail length measurements and RNA half-life estimates, we hypothesize that a deadenylation-independent degradation model, similar to NMD, applies to other unstable RNAs. This model predicts that changes in the deadenylation speed should not impact the half-life of these RNAs. Poly-A tail shortening could be thus considered as a parallel event rather than a requirement for RNA degradation. We found that, as predicted by a deadenylation-independent model of RNA degradation, an unstable reporter RNA that is insensitive to NMD, is not stabilized by an inhibition of deadenylation, even if it is highly sensitive to the decapping activity. These results were obtained in a degron system for quick and specific depletion of the proteins, which allows a minimal global perturbation of the tested cells. In line with these functional results, we identified unstable RNAs as a population of RNA molecules associated with the poly-A binding protein and conclude that long poly-A tails are a natural feature of unstable RNAs. Our results are consistent with recent large-scale estimates of poly-A tails length and RNA stability and imply that molecular mechanisms involved in the initiation of RNA degradation in eukaryotes remain to be discovered.

# Résumé

La dégradation de l'ARN est une étape essentielle pour maintenir l'équilibre nécessaire à l'expression des gènes et à leur régulation. Chez les eucaryotes, le modèle le plus courant propose que la dégradation de l'ARN est contrôlée par la vitesse de raccourcissement de la queue poly-A. En dessous d'un certain seuil, la présence d'une courte queue oligo-A entraîne l'activation du décapping et la dégradation de l'ARN. Cependant, la déadénylation progressive n'est pas nécessaire pour une voie majeure de dégradation rapide de l'ARN, la dégradation par le « nonsense-mediated mRNA decay » (NMD), qui se passe indépendamment de la taille de la queue poly-A. Sur la base de mesures à grande échelle de la longueur de la queue poly-A et d'estimations de la demi-vie de l'ARN, nous proposons qu'un modèle de dégradation indépendant de la déadénylation, similaire au NMD, puisse exister pour d'autres ARN instables. Ce modèle prédit que des changements de vitesse de déadénylation ne devraient pas avoir d'impact sur la demi-vie de ces ARN. Le raccourcissement de la queue des poly-A semble être dans ce cas un processus parallèle et non pas nécessaire pour la dégradation de l'ARN. Comme prédit par notre modèle, un ARN rapporteur instable et insensible au NMD n'est pas stabilisé par une diminution de la vitesse de déadénylation, même s'il est très sensible à l'inactivation du décapping. Ces résultats ont été obtenus dans un système de dégron de déplétion rapide et spécifique de protéines d'intérêt, qui permet d'éviter une perturbation globale du métabolisme des cellules testées. En accord avec ces résultats fonctionnels, nous avons identifié les ARN instables comme une population de molécules d'ARN associées à la protéine de liaison cytoplasmique des poly-A et conclu que de longues queues de poly-A sont une caractéristique naturelle des ARN instables. Nos résultats sont compatibles avec les estimations récentes de la taille des queues poly-A et de la vitesse de dégradation des transcrits et devraient avoir un impact important sur la découverte des mécanismes moléculaires par lesquels la dégradation des ARN est initiée chez les eucaryotes.

# Sommaire

Introduction générale.....	1
I. Transcription génique.....	1
I.1. Initiation et élongation de la transcription.....	1
I.2. Terminaison de la transcription et export des ARN.....	2
II. Rôles cytoplasmiques des queues poly-A et dégradation des ARN messagers.....	4
II.1. Découverte et fonctions décrites des queues poly-A.....	4
II.2. Evaluation des tailles des queues poly-A.....	5
II.3. Pab1, la protéine de liaison cytoplasmique des queues poly-A.....	5
II.3.1. Caractérisation de Pab1.....	5
II.3.2. Les fonctions de Pab1.....	6
II.4. La dégradation des ARN chez <i>S. cerevisiae</i> .....	7
II.4.1. La déadénylation.....	7
II.4.1.1. Le complexe Pan2-Pan3 (PAN).....	7
A. Pan2.....	8
B. Pan3.....	8
C. Reconnaissance des queues poly-A et déadénylation.....	8
II.4.1.2. Le complexe Ccr4-Not (CNOT).....	8
A. Les nucléases Ccr4 et Pop2.....	9
a. Ccr4.....	9
b. Pop2.....	9
c. Liaison de Ccr4 et Pop2.....	9
B. Les autres protéines du complexe CNOT.....	10
a. Not1.....	10
b. Les autres protéines Nots.....	10
II.4.1.3. Redondance des complexes PAN et CNOT.....	11
II.4.1.4. Les phases de déadénylation de CNOT et l'implication de Pab1.....	12
II.4.1.5. Influence de la queue poly-A sur la traduction et la stabilité des ARN.....	12
II.4.2. Exemples de régulation de la déadénylation.....	13
II.4.3. Déadénylation et décapping des ARN.....	14
II.4.3.1. Le complexe de décapping des ARN.....	14
A. Dcp2.....	14
B. Dcp1.....	15
C. Protéines associées au complexe de décapping.....	15
D. Formation du complexe de décapping.....	15
II.4.3.2. Le complexe Lsm1-7 et les protéines associées, un lien entre la queue poly-A et le décapping.....	16
A. Le complexe Lsm1-7.....	16
B. Pat1 et le complexe Lsm1-7.....	16
II.4.3.3. Connexion entre les facteurs de la déadénylation et du décapping.....	17
A. Implication des protéines Nots.....	17
B. Implication de Dhh1 dans la dégradation des ARN.....	18

II.4.3.4. Vitesse de déadénylation et décapping.....	18
II.5. Les voies de dégradation 5'-3' et 3'-5'.....	19
II.6. Régulation de la transcription par la quantité d'ARN cytoplasmique.....	20
III. Traduction des ARN et ses liens avec la dégradation.....	21
III.1. Initiation de la traduction.....	21
III.2. Elongation de la traduction.....	22
III.3. Terminaison de la traduction.....	22
III.4. Optimisation des codons.....	24
III.5. Traduction et dégradation.....	27
III.5.1. La dégradation co-traductionnelle.....	27
III.5.2. La dégradation des ARN par le NMD.....	28
III.5.3. La dégradation par les voies NSD et NGD.....	30
Introduction aux résultats.....	31
Résultats.....	32
I. Upf1 s'associe à des protéines qui lient les ARN avec des queues poly-A de différentes tailles.....	32
I.1. L'étude des complexes associés à la dégradation des ARN par spectrométrie de masse révèle de nouvelles connexions.....	32
I.1.1. Interactions de la protéine Pab1.....	32
I.1.2. Le complexe Lsm1-7/Pat1 interagit avec les facteurs du NMD.....	33
I.1.3. Le facteur de décapping Dhh1 révèle les complexes associés à la dégradation des ARN.....	33
I. 2. Etude des ARN associés à Upf1 et Lsm1-7 et Upf1 et Pab1 .....	36
I.2.1. Lsm1-TAP s'associe avec l'ARN de Dcp2.....	36
I.2.2. Pab1 s'associe avec des ARN instables.....	37
II. La déadénylation n'est pas nécessaire à la dégradation des ARN.....	43
II.1. La taille de la queue poly-A est inversement corrélée à la demi-vie des ARN.....	43
II.2. Le raccourcissement de la queue poly-A n'est pas nécessaire au déclenchement de la dégradation des ARN.....	45
II.2.1. Construction de gènes rapporteurs pour analyser le rôle de la déadénylation dans la dégradation des ARN.....	45
II.2.2. La déplétion des déadénylases.....	46
II.3. La déadénylation n'est pas requise pour la dégradation d'ARN rapporteurs.....	48
II.4. L'absence de Pab1 n'affecte pas la demi-vie des ARN mais empêche leur déadénylation.....	51
Conclusions et discussion.....	54
I. Analyse des complexes protéiques et des ARN associés au complexe Lsm1-7.....	54
II. Le lien entre Upf1 et Pab1 autour du NMD.....	55
III. Un modèle actualisé pour la dégradation des ARN.....	56
Figures supplémentaires.....	61
Matériel et méthodes.....	67
Bibliographie.....	81
Article de thèse.....	98
Abstract.....	99

Introduction.....	100
Results.....	102
The core NMD factor Upf1 is associated with RNA degradation complexes not involved in NMD.....	102
Pab1 preferentially associates with unstable RNA.....	103
Decay kinetics modelling explains the relation between poly(A) tail length and RNA stability.....	104
Rapid degradation of deadenylation factors tests the impact of the poly(A) tail shortening on RNA levels.....	107
Decapping, but not deadenylation speed, affects the degradation of a reporter RNA..	108
Discussion.....	110
Materials and methods.....	112
Yeast and bacterial strains.....	112
Media and growth conditions.....	112
Plasmids construction.....	113
Cell culture, RNA extraction and degradation kinetics.....	113
Northern blots.....	114
RNA half life measurements.....	114
Affinity purification for RNA sequencing, immunoblot and mass-spectrometry.....	115
Illumina RNA sequencing results analysis.....	115
Reverse transcription and quantitative PCR.....	115
Poly(A) tail length estimation with ePAT.....	116
Protein extracts and immunoblots.....	116
Mass spectrometry acquisition and data analysis.....	116
Nanopore sequencing and analysis.....	117
Data and software availability.....	117
Acknowledgements.....	118
References.....	118
Figure legends.....	123
Figures.....	129
Supplementary figures.....	136

## Table des figures

Figure 1. Terminaison de la transcription des ARNm et polyadénylation chez les eucaryotes.....	4
Figure 2. Représentation du complexe CNOT.....	11
Figure 3. Les phases de déadénylation.....	12
Figure 4. Mécanisme du décapping dépendant du complexe Lsm1-7/Pat1.....	17
Figure 5. Dogme actuel pour la dégradation des ARN chez <i>S. cerevisiae</i> .....	20
Figure 6. Mécanismes de l'initiation, de l'élongation et de la terminaison de la traduction.	23
Figure 7. L'usage des codons influence la traduction <i>via</i> des mécanismes dépendants des ARNt.....	26
Figure 8. Dégradation de l'ARN par le nonsense mediated mRNA decay.....	30
Figure 9. Analyse des protéines associées aux complexes de dégradation de l'ARN et à Pab1.....	35
Figure 10. Les ARN de Dcp2 et de Pat1 sont associés à Lsm1-TAP.....	37
Figure 11. Rpl28 pré-messager, substrat du NMD, est associé à Pab1 et à Upf1.....	38
Figure 12. Pab1-TAP est associée à des ARN instables.....	40
Figure 13. Comparaison des ARN purifiés par Rpl16a et Pab1.....	41
Figure 14. Validation de l'association de Pab1 à des substrats du NMD qui ont de longues queues poly-A.....	42
Figure 15. La demi-vie des ARN n'est pas corrélée à la longueur de la queue poly-A à l'état d'équilibre.....	44
Figure 16. L'optimisation des codons affecte la demi-vie d'ARN rapporteurs qui sont stabilisés en absence de la traduction.....	46
Figure 17. La déplétion des déadénylases par le système AID.....	47
Figure 18. La demi-vie des versions optimale et non-optimale de His3 n'est pas impactée par la déplétion de Ccr4-AID et Pop2-AID mais est augmentée quand le décapping est compromis.....	49
Figure 19. La déplétion de Pop2 et Ccr4 empêche la déadénylation totale des ARN rapporteurs.....	51
Figure 20. La déplétion de Pab1 ne déstabilise pas les ARN rapporteurs.....	52
Figure 21. Un rapporteur NMD est sensible à la déplétion de Pab1.....	53
Figure 22. Un modèle de dégradation des ARN indépendant de la déadénylation.....	59
Figure S1. L'analyse statistique de DESeq2 révèle de nombreux transcrits différentiellement présents dans les purifications Rpl16, Upf1 et Pab1 et peu avec Lsm1.	61
Figure S2. Les rapporteurs sont stabilisés par les déplétions de Dcp2 et Xrn1.....	62
Figure S3. La stabilité de His3 ne dépend pas de l'exosome cytoplasmique quand la voie 5' est active.....	62
Figure S4. La déplétion des déadénylases n'affecte pas la stabilité des rapporteurs.....	63
Figure S5. La déplétion de Pop2 et Ccr4 empêche la déadénylation totale des ARN rapporteurs.....	64
Figure S6. La déplétion de Pab1 empêche le raccourcissement des queues poly-A.....	65
Figure S7. Le rapporteur His3-PTC est sensible aux déplétions de Upf1, Dcp2 et Xrn1....	65



Figure S8. L'absence de Dhh1 ne stabilise pas les rapporteurs His3 optimal et non optimal .....	66
Figure 1. A network of RNA-mediated protein interactions links Upf1 to Pab1 and the Lm1-7 complex.....	129
Figure 2. Pab1 associates with unstable RNAs that have longer than average poly(A) tails. ....	130
Figure 3. An alternative kinetic model for RNA degradation explains the inverse correlation between poly(A) tail length and mRNA stability.....	131
Figure 4. An inducible degron system shows that RNA levels are independent from massive poly(A) tail changes.....	132
Figure 5. Codon biased reporters are stabilized when translation is inhibited.....	133
Figure 6. Non-optimal and optimal HIS3 reporter RNAs are stabilized by depletion of Dcp2, but not by depletion Ccr4 or Pop2.....	134
Figure 7. Deadenylation inhibition increases poly(A) tail sizes of reporter RNA, without an effect on their apparent half-life.....	135
Figure S1 related to Figure 1. Enrichment of specific factors in purified complexes.....	136
Figure S2 related to Figure 2. Pab1 preferentially binds long poly(A) tail versions of cellular RNAs.....	137
Figure S3 related to Figure 3. Robustness of the negative correlations between poly(A) tail length and half-life of RNA.....	138
Figure S4 related to Figure 5. Depletion of Dhh1 has variable effects on codon-biased reporters.....	139
Figure S5 in relation with Fig. 6. Half-life of reporter RNA as measured by RT-qPCR.....	140
Figure S6, in relation with Fig. 7. Deadenylation inhibition increases poly(A) tail sizes of reporter RNA.....	141

# Introduction générale

La synthèse de molécules d'ARN (acide ribonucléique) à partir de l'ADN (acide désoxyribonucléique), ou transcription, constitue la première étape de l'expression de l'information génétique. Puisqu'il y a production d'ARN d'un côté, il y a dégradation de l'autre. La temporalité de cette dégradation varie selon le type d'ARN. Certains ARN sont immédiatement dégradés après leur transcription dans le noyau, d'autres vont persister, être exportés dans le cytoplasme où ils y seront traduits pour produire des protéines, essentielles au fonctionnement cellulaire. Afin de perpétuer et de remettre à disposition les éléments nécessaires à ce cycle, le recyclage des ARN est une étape critique pour la balance physiologique de la cellule. La dégradation des ARN implique une diversité de facteurs dont les mécanismes et la spécificité sont seulement partiellement identifiés.

## I. Transcription génique

La transcription génique est la première étape de l'expression des gènes. Chez les cellules eucaryotes, celle-ci se déroule dans le noyau. Au cours de ce processus, les ARN polymérases vont recopier l'ADN pour synthétiser différents types d'ARN. Chez la plupart des eucaryotes, il existe trois ARN polymérases différentes, chacune étant responsable de la synthèse d'un ou de plusieurs types d'ARN. L'ARN polymérase I synthétise les ARN ribosomiques (ARNr). L'ARN polymérase II synthétise les ARN messagers<sup>1</sup> (ARNm) qui permettront la synthèse protéique et des ARN non codant. L'ARN polymérase III synthétise les ARN de transfert (ARNt) et d'autres ARN non codant.

### I.1. Initiation et élongation de la transcription

L'ARN polymérase II de la levure *Saccharomyces cerevisiae* est composée de 12 sous-unités dont une, Rpb1, qui a une extrémité C-terminale appelée CTD pour « *Carboxy Terminal Domain* ». Elle subit de nombreuses modifications biochimiques durant la transcription. Le CTD contient notamment des motifs répétés dont trois sérines pouvant être phosphorylées en position 2, 5 et 7. La phosphorylation des différentes sérines du CTD évolue au cours de la transcription : durant l'élongation, la phosphorylation de la sérine en position 5 est progressivement remplacée par la phosphorylation de la sérine 2. Ces modifications de phosphorylation sont importantes pour le recrutement de différents facteurs intervenant dans la régulation de la transcription ou les modifications co-transcriptionnelles de l'ARN, revue dans (Phatnani and Greenleaf, 2006).

L'initiation de la transcription débute par l'assemblage de l'ARN polymérase II avec les facteurs de transcription généraux, TFIIA, TFIIB, TFIID, TFIIE, TFIIIF et TFIIH et le complexe Mediator au niveau des séquences promotrices pour former le complexe de pré-initiation ou PIC. Le PIC permet le positionnement de l'ARN polymérase au site d'initiation

---

<sup>1</sup> Par la suite, « ARN » désignera les ARN messagers.

de la transcription, l'ouverture du duplex d'ADN pour former ce qui est appelé « la bulle de transcription » et le positionnement de l'ADN dans le site actif de la polymérase pour permettre la synthèse d'ARN, revue dans (Sainsbury et al., 2015).

Le départ des facteurs d'initiation de la région promotrice marque le début de l'élongation de la transcription. L'ARN polymérase II quitte la région promotrice et débute l'élongation de la transcription au niveau du site d'initiation de la transcription (aussi appelé TSS pour « *transcription start site* ») en coopération avec des facteurs d'élongation. Au cours de l'élongation deux modifications majeures interviennent sur l'ARN : l'ajout d'une coiffe qui le protège des exoribonucléases qui pourraient le dégrader et l'épissage qui permet de retirer les parties introniques, non codantes de l'ARN pré-messager. La phosphorylation de la sérine 5 du CTD de l'ARN polymérase II participe notamment au recrutement des facteurs impliqués dans l'ajout de la coiffe de l'ARN et au complexe d'épissage. Aussi dès que l'extrémité 5' de l'ARN émerge du tunnel de sortie de l'ARN polymérase II, la coiffe lui est ajoutée. Celle-ci est composée d'un N7-méthylguanosine (m7G) connectée au prochain nucléotide par une liaison 5' – 5' triphosphate. L'épissage est réalisé par le spliceosome, composé de protéines et de petits ARN nucléaires (ARNsn) qui s'assemblent au site d'épissage, revue dans (Peck et al., 2019).

## 1.2. Terminaison de la transcription et export des ARN

La terminaison de la transcription a lieu quand l'ARN et l'ARN polymérase II sont relâchés de l'ADN. Pour les ARN messagers, la terminaison est médiée par le complexe CPF–CF - pour « *cleavage and polyadenylation factor and cleavage complex* » - composé de trois sous-complexes : CPF, CFIA et CFIB. La transcription d'un signal de polyadénylation riche en adénines et uraciles et la phosphorylation de la sérine 2 du CTD de l'ARN polymérase II qui interagit avec la protéine Pcf11 *via son* domaine CID (« *CTD-interaction domain* ») provoquent le clivage de l'ARN par l'endonucléase Ysh1. Pap1 qui est la poly-A polymérase synthétise alors une chaîne d'adénosines à l'extrémité 3' de l'ARN tandis que Rat1 dégrade l'extrémité restante toujours associée à l'ARN polymérase II. Une perte des facteurs d'élongation couplée à un changement de conformation de l'ARN polymérase II conduirait à la dissociation de la polymérase selon le modèle allostérique. Un autre modèle, « Torpedo » propose que Rat1 associé à Rai1 et Rtt103 interagissent avec le CTD de l'ARN polymérase II ce qui mènerait à la dissociation du complexe d'élongation (**figure 1**), revue dans (Porrúa and Libri, 2015)

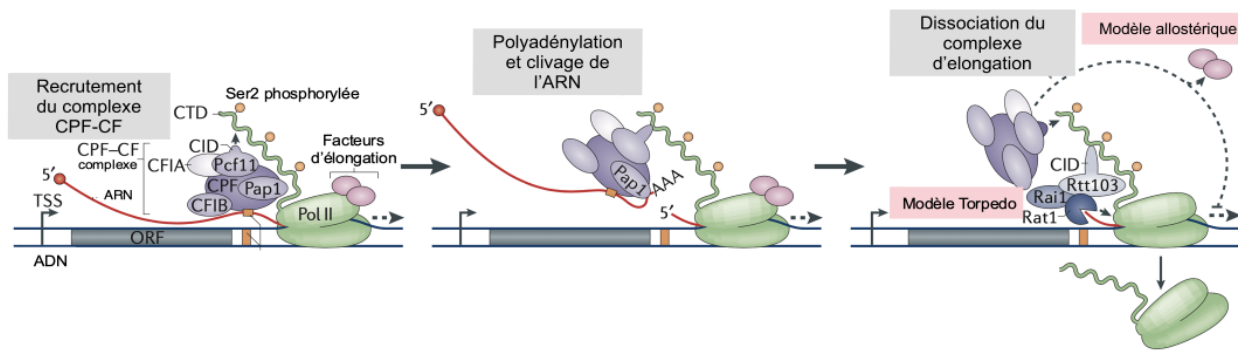
Chez *S. cerevisiae*, il existe une autre terminaison de la transcription pour certains ARN non codant tels que les CUTs (« *cryptic unstable transcript* » (Wyers et al., 2005; Davis and Ares, 2006; Houalla et al., 2006)) qui n'ont pas de sites de terminaison de la transcription et de polyadénylation canoniques. La terminaison de ces ARN générés par des événements d'initiation de transcription pervasifs dans des zones dépourvues de nucléosomes, dépend des protéines Nrd1–Nab3–Sen1 (NNS). Nrd1 et Nab3 sont des protéines liant l'ARN en reconnaissant des séquences spécifiques. La reconnaissance de ces séquences par Nrd1 et Nab3 est importante pour la terminaison médiée par le complexe NNS qui serait pré-recruté sur le complexe d'élongation de la transcription grâce

à l'interaction entre le domaine CID de Nrd1 et le CTD de l'ARN polymérase, phosphorylé sur la sérine 5. Contrairement à ce qui est observé dans le cas de la terminaison dépendante de CPF-CF, l'ARN n'est pas clivé et bien que le mécanisme de dissociation de l'ARN polymérase II ne soit encore complètement élucidé, il nécessite l'activité hélicase de Sen1. L'ARN libéré par la dissociation du complexe d'élongation est alors oligo-adenylé par le complexe TRAMP qui contient des poly-A polymérases non canoniques, Trf4 et Trf5 et les protéines Air1/2 et Mtr4. La polyadénylation de l'ARN permet la dégradation par un complexe de dégradation nucléaire : l'exosome nucléaire.

Des événements d'initiation de transcription pervasifs générant des ARN non codant instables analogues aux CUTs ont été mis en évidence dans d'autres organismes. Chez les mammifères, ces transcrits dénommés PROMPT pour « *promoter-proximal transcript* » sont également rapidement dégradés par l'exosome nucléaire après leur synthèse. Cependant, contrairement à ce qui est observé chez la levure, leur terminaison dépend des facteurs de terminaison de la transcription canoniques, potentiellement recrutés par le complexe nucléaire de liaison à la coiffe (CBC) et la protéine ARS2 qui interagissent également avec l'exosome nucléaire, revue dans (Porrua and Libri, 2015).

Une fois que la queue poly-A est synthétisée par Pap1, elle est liée par Nab2 chez *S. cerevisiae* pour « *Nuclear polyAdenylated RNA-Binding* » (Anderson et al., 1993) ou PABPN1 chez les mammifères (Wahle, 1991). La liaison de Nab2 à la queue poly-A permettrait d'une part de protéger l'ARN néosynthétisé de la dégradation par l'exosome nucléaire et d'autre part de faciliter son export. Nab2 aurait aussi un rôle dans le contrôle de la taille des queues poly-A des ARN néo-synthétisés : des mesures de taille de queues poly-A ont montré que la déplétion de Nab2 augmentait la longueur de la queue poly-A de certains transcrits chez *S. cerevisiae* (Turtola et al., 2021) et revue dans (Tudek et al., 2018).

L'export des ARN consiste en le passage des ARN du noyau au cytoplasme par le biais des pores nucléaires (NPC pour « *nuclear pore complexes* »). Les ARN sont liés par les protéines Mex67/Mtr2 à l'aide de l'hélicase Sub2 accompagnée de Yra1 chez *S. cerevisiae*, homologues de NXF1/NXT1 avec l'hélicase UAP56 chez l'humain. D'autres facteurs associés aident à l'export dont le complexe TREX (conservé de la levure à l'humain) qui se lie aux ARN au cours de la transcription. La liaison de ces différents complexes aurait pour effet d'une part, de changer la conformation des ARN remodelés grâce à Sub2 et l'hydrolyse d'ATP, et d'autre part « d'activer » l'ARN pour l'export. Une fois « compétents », les ARN sont chargés au sein du pore nucléaire pour passer dans le cytoplasme où ils sont remodelés par Dbp5 pour éviter un éventuel retour au noyau. Enfin les facteurs Mex67/Mtr2 sont relâchés et Nab2 est remplacé par Pab1, revue dans (Stewart, 2019).



**Figure 1. Terminaison de la transcription des ARNm et polyadénylation chez les eucaryotes**

La terminaison de la transcription débute par la reconnaissance de l'ARN par le complexe CPF-CF au niveau d'un signal de polyadénylation. L'ARN est clivé et Pap1 polyadényle l'ARN. Rat1 hydrolyse le fragment d'ARN généré en aval du site de clivage. Selon le modèle allostérique, un changement de conformation dissocierait l'ARN polymérase II tandis que selon le modèle « Torpedo », Rat1 interviendrait dans la dissociation de l'ARN polymérase II. Figure et légende adaptées de (Porrua and Libri, 2015).

## II. Rôles cytoplasmiques des queues poly-A et dégradation des ARN messagers

Les queues poly-A sont une modification particulière des ARN. Depuis leur découverte, de nombreuses fonctions leur ont été attribuées. Elles seraient impliquées dans l'export des ARN dans le cytoplasme, leur traduction et leur dégradation.

### II.1. Découverte et fonctions décrites des queues poly-A

Les premières études qui font état de la présence d'une succession d'adénosines à l'extrémité des ARN datent des années soixante. Elles débutent par la découverte d'une enzyme capable de synthétiser un polymère d'adénosines à l'extrémité de chaînes d'acides ribonucléiques (Edmonds and Abrams, 1960). Plus d'un demi-siècle plus tard, le rôle de la queue poly-A - pour poly-adénosine - reste ambigu. Les difficultés à comprendre ses fonctions peuvent notamment s'expliquer par ses rôles complexes et ambivalents. Cette succession d'homo-nucléotides aurait au moins, selon les modèles décrits, un double rôle au sein du cytoplasme - qu'il est possible de définir comme : « pro-traductionnel » et « dégradation-stimulante ». Il y aurait d'une part, un effet stimulant pour la traduction *via* l'interaction de facteurs d'initiation de la traduction entre la queue poly-A et la coiffe des ARN et d'autre part une fonction stimulante pour la dégradation, par l'intervention de nucléases qui utiliseraient cette extrémité particulière pour dégrader les ARN.

Un autre type de modification 3' qui peut influencer la dégradation des ARN est l'uridylation qui consiste en l'ajout d'uridines à l'extrémité 3' des ARN par des terminales uridylyltransferases (TUTases). Elles existent de *S. pombe* à l'humain sans équivalent identifié chez *S. cerevisiae*. Cette modification a des effets sur la stabilité des ARN soit *via* la liaison de complexes qui vont initier la dégradation, comme le complexe Lsm1-7 détaillé ultérieurement, soit parce qu'elle est reconnue directement par l'exosome cytoplasmique. Chez les mammifères et la plante modèle *Arabidopsis thaliana* notamment, cette modification interviendrait après un clivage interne opéré par un micro ARN *via* le complexe RISC (« *RNA-induced silencing complex* »). Alors que le fragment 3' généré est dégradé par Xrn1, le fragment 5' est polyuridylé, décappé puis dégradé dans le sens 5'-3' par Xrn1 et dans le sens 3'-5' par des exoribonucléases spécifiques et l'exosome cytoplasmique, revue dans (Zigáčková and Vaňáčová, 2018).

## II.2. Evaluation des tailles des queues poly-A

La longueur de la queue poly-A diffère selon les espèces étudiées : initialement estimées de 50 à 80 nucléotides chez la levure *S. cerevisiae* (McLaughlin et al., 1973; Brown and Sachs, 1998) et de 150 à 200 nucléotides chez les mammifères (Edmonds et al., 1971).

L'utilisation de puces à ADN (Beilharz and Preiss, 2007) ainsi que le séquençage Illumina ont permis la détermination des tailles des queues poly-A à l'échelle du génome et non plus seulement sur quelques gènes cibles. Ces études à grande échelle ont ré-évalué à la baisse la taille des queues poly-A chez *S. cerevisiae*, avec une taille médiane de 28 nucléotides mais aussi chez les mammifères, avec des tailles comprises entre 67 et 96 nucléotides. Chez la plante modèle *Arabidopsis thaliana* et la drosophile, la taille médiane a été évaluée à 50 nucléotides (Subtelny et al., 2014). Plus récemment, le séquençage par la technique Nanopore des queues poly-A chez *S. cerevisiae* a révélé que la moyenne était de 50 nucléotides pour les ARN messagers (Tudek et al., 2021).

Le séquençage massif a également permis la découverte de la composition parfois hétérogène des queues poly-A. En effet, des guanylations au sein des queues poly-A dans des cellules HeLa ont été identifiées (H. Chang et al., 2014). Les enzymes impliquées ont ensuite été caractérisées par inactivation de différentes « *terminal nucleotidyl transferases* » (TENTs). Ces modifications au sein des queues poly-A auraient pour effet de diminuer l'affinité des déadénylases et de ce fait, de ralentir la déadénylation des ARN avec pour conséquence une dégradation plus lente (Lim et al., 2018; Tang et al., 2019).

## II.3. Pab1, la protéine de liaison cytoplasmique des queues poly-A

### II.3.1. Caractérisation de Pab1

Au sein du cytoplasme la queue poly-A est liée par la protéine cytoplasmique Pab1 – pour « *poly-A binding protein* » – qui est la protéine de liaison majeure des queues poly-A chez *S. cerevisiae* (Baer and Kornberg, 1980, 1983).



Si chez la levure il n'existe qu'une seule protéine cytoplasmique de liaison aux queues poly-A, chez les métazoaires et les plantes, plusieurs protéines cytoplasmiques pouvant lier les queues poly-A des ARN (PABPCs pour « *cytoplasmic poly-A binding protein* ») ont été identifiées, revue dans (Smith et al., 2014). Les mieux caractérisées chez l'humain sont les protéines PABPC1 et PABPC4, qui auraient toutes deux un rôle dans la stabilité des ARN cytoplasmiques (Kajjo et al., 2022).

Pab1 est composée de quatre domaines RRM (« *RNA Recognition Motif* ») (Baer and Kornberg, 1983). Elle n'a besoin que d'une queue poly-A de 12 nucléotides pour se lier *via* ses deux domaines RRM1 et RRM2 (Sachs et al., 1987), mais son empreinte totale sur l'ARN serait de 27 nucléotides (Baer and Kornberg, 1980). Pab1 contient dans sa partie C-terminale un domaine Mlle, « mademoiselle », qui peut interagir avec un domaine peptidique PAM-2, « *PABP-interacting motif 2* ». Ce type de domaine, PAM-2, est retrouvé dans les protéines Pan3 et eRF3/Sup35 impliquées respectivement dans la déadénylation et la terminaison de la traduction avec lesquelles Pab1 interagit (Siddiqui et al., 2007; Xie et al., 2014; Roque et al., 2015).

Une étude récente a montré l'influence de la région 3' non traduite (*3'UTR*, « *untranslated region* ») dans la liaison de Pab1 aux queues poly-A. L'environnement nucléotidique proche de la queue poly-A varie selon le site de terminaison de la transcription et cela modifierait l'affinité de Pab1 pour les queues poly-A. Cette étude a par ailleurs montré que la liaison de Pab1 aux différentes isoformes d'un même ARN était très variable, les isoformes ayant les demi-vies les plus courtes étant les plus associées à Pab1 (Moqtaderi et al., 2018).

L'observation en microscopie électronique à transmission a permis l'observation de plusieurs Pab1 se succédant sur les queues poly-A des ARN. Cela a confirmé l'hypothèse que les queues poly-A des ARN pouvaient être liées par plusieurs Pab1 (Sawazaki et al., 2018).

### II.3.2. Les fonctions de Pab1

Pab1 est une protéine essentielle chez la levure *S. cerevisiae* qui intervient dans la régulation de nombreuses étapes du métabolisme des ARN, par l'intermédiaire des différents facteurs avec lesquels elle peut interagir.

Les premières études utilisant un mutant thermosensible de Pab1 ont mis en évidence l'importance de cette protéine dans la régulation de la taille des queues poly-A et de l'initiation de la traduction (Sachs and Davis, 1989).

Des études ultérieures ont mis en évidence des interactions directes de Pab1 avec les facteurs d'initiation et de terminaison de la traduction tels que eIF4G et eRF3 respectivement (Gallie, 1998; Wells et al., 1998; Hoshino et al., 1999; Cosson et al., 2002). L'interaction de Pab1 avec eIF4G faciliterait la formation du modèle « *closed-loop* » d'après lequel les ARN seraient circularisés par l'interaction de l'extrémité 3' *via* Pab1 à l'extrémité 5' par les facteurs d'initiation ce qui jouerait un rôle dans l'initiation de la traduction.

Certaines mutations secondaires permettent la survie de la cellule en absence de Pab1. Ces mutations se répartissent en deux catégories chez *S. cerevisiae* : celles qui affectent des gènes impliqués dans la synthèse des protéines de la sous-unité ribosomique 60S et celles affectant des gènes codant pour des protéines impliquées dans la dégradation des ARN dont Lsm1, Pat1, Dcp1, Xrn1 et la protéine peu caractérisée Pbp1, qui limiterait la déadénylation, revue dans (Brambilla et al., 2019). De même chez les mammifères, des mutations de Dcp2, Xrn1 ou Lsm1 rétablissent la viabilité cellulaire lorsque PABPC1 est inactivée (Kajjo et al., 2022). Le fait que des mutations de certaines des protéines majeures impliquées dans la dégradation des ARN restaurent la viabilité cellulaire en l'absence de Pab1 suggère que Pab1 joue un rôle dans la régulation de la dégradation des ARN.

Bien qu'essentiellement cytoplasmique, une partie de Pab1 localisée dans le noyau jouerait un rôle dans l'export des ARN par l'intermédiaire de son interaction avec le facteur d'export Crm1/Xpo1 (Brune et al., 2005). Cette interaction entre Pab1 avec l'export n'est pas conservée chez les mammifères où la protéine PABPC1 ne semble pas impliquée dans l'export au cytoplasme des ARN. Pourtant, après un stress aux rayons ultraviolets les protéines PABPC1 et PABPC4 sont relocalisées dans le noyau. Cette relocalisation serait dû au fait que les PABPCs lient les queues poly-A des ARN dont l'export est bloqué dans ces conditions de stress plutôt qu'à un rôle actif de ces protéines dans l'export (Burgess et al., 2011).

Enfin une étude récente a montré que, chez la levure, la fraction nucléaire de Pab1 pourrait jouer un rôle dans la polyadénylation des ARN dans le noyau lorsque la protéine de liaison aux queues poly-A nucléaire, Nab2, est limitante (Turtola et al., 2021).

## **II.4. La dégradation des ARN chez *S. cerevisiae***

La dégradation des ARN est un processus majeur pour la cellule, médiée par de nombreux facteurs. En plus d'être l'étape de recyclage des ARN, c'est également une étape qui permet de contrôler la quantité de protéines synthétisées. La dégradation est d'ailleurs souvent co-translationnelle. La compréhension fine de la temporalité de cette dégradation et de son déclenchement restent à élucider.

### **II.4.1. La déadénylation**

D'après le modèle actuel chez *S. cerevisiae*, pour la majorité des ARN codant pour des protéines fonctionnelles, la dégradation débute par le raccourcissement de la queue poly-A des ARN – autrement dit - la déadénylation. Celle-ci est catalysée par différentes déadénylases assemblées en complexes : le complexe Pan2-Pan3 (PAN) et le complexe Ccr4-Not-Pop2 (CNOT) (Tucker et al., 2001).

#### **II.4.1.1. Le complexe Pan2-Pan3 (PAN)**

Le complexe PAN -pour « *Pab-dependent poly-A nuclease* » - composé de Pan2 et de Pan3 est conservé chez les eucaryotes. Initialement décrit pour avoir une capacité de



déadénylation dépendante de Pab1 chez *S. cerevisiae* (Sachs and Deardorff, 1992), c'est le premier complexe cytoplasmique à déadényler les ARN (Brown and Sachs, 1998).

### A. Pan2

Pan2 contient un domaine RNase de la famille DEDD, responsable de l'activité 3'-5' exonucléase. Cette famille partage des acides aminés conservés au niveau du site catalytique et a la particularité de lier des ions métalliques (Zuo, 2001). Pan2 nécessite du magnésium et une extrémité 3'OH libre et libère comme produit des 5'-adénosines monophosphates lors de l'hydrolyse des poly-A (Boeck et al., 1996).

### B. Pan3

Pan3, caractérisée initialement par le même laboratoire (Brown et al., 1996) lie spécifiquement les queues poly-A des ARN *via* un motif en doigt de zinc mais ne possède pas d'activité exonucléase et a donc besoin de Pan2 pour déadényler efficacement les ARN (Wolf et al., 2014).

Pan3 en dimère interagit avec Pab1 *via* un motif PAM-2 situé en C-terminal de Pan3 (Siddiqui et al., 2007). Le complexe serait d'après la résolution de la structure par cryo-microscopie électronique, formé d'une unité de Pan2 et deux unités de Pan3 et il se lierait préférentiellement à des ARN qui auraient plusieurs molécules de Pab1 (Jonas et al., 2014; Schäfer et al., 2019).

### C. Reconnaissance des queues poly-A et déadénylation

La reconnaissance spécifique des queues poly-A est la particularité des déadénylases, c'est-à-dire qu'elles discriminent, *a priori*, une extrémité de poly-adénosines du reste de l'ARN.

Une chaîne de poly-adénosines adopte une conformation particulière avec un empilement des bases et la formation d'une hélice (Saenger et al., 1975). Jusqu'à récemment il n'y avait pas de données sur la reconnaissance spécifique de cette chaîne de poly-A même s'il avait été montré que les déadénylases avaient des préférences de substrat, le complexe PAN se lie aux poly-adénosines mais également à des poly-uridines (Sachs and Deardorff, 1992). Une étude récente a montré que Pan2 reconnaissait spécifiquement cet empilement d'adénosines et que la présence de guanosines interrompant les chaînes d'adénosines perturbait cet empilement et la dégradation des queues poly-A par Pan2, revue dans (Tang and Passmore, 2019).

En résumé, le recrutement du complexe PAN sur les ARN serait médié par l'interaction entre le motif PAM-2 de Pan3 et Pab1. Pan3 interagirait également avec la queue poly-A. Enfin, la déadénylation des ARN serait permise par l'activité nucléase de Pan2 qui catalyserait la réaction par la reconnaissance de ces empilements d'adénosines, revues dans (Wolf and Passmore, 2014; Passmore and Collier, 2022).

#### II.4.1.2. Le complexe Ccr4-Not (CNOT)

Le complexe CNOT est un complexe multiprotéique conservé chez les eucaryotes. Il est composé de neuf sous-unités chez *S. cerevisiae* dont deux qui possèdent les activités de déadénylation : Ccr4 pour « *carbon catabolite repression* » et Pop2 pour

« *PGK promoter directed overproduction* » (avec pour *alias*, Caf1 pour « *Ccr4 associated factor* »), (**figure 2**) (Daugeron, 2001; Tucker et al., 2001).

Le complexe a initialement été considéré comme un régulateur de la transcription de gènes impliqués dans le métabolisme chez *S. cerevisiae*. Pop2 avait un effet sur la quantité d'une enzyme impliquée dans la glycolyse tandis que des mutations de Ccr4 affectaient une enzyme impliquée dans la fermentation (Sakai et al., 1992; Denis, 1984; Denis and Malvar, 1990).

En plus d'être caractérisé comme un complexe impliqué dans la déadénylation des ARN, le complexe CNOT intervient dans d'autres mécanismes cellulaires. Il aurait notamment de multiples rôles régulateurs dans le métabolisme des protéines : en étant impliqué à la fois dans la régulation de la synthèse des ARN ribosomiques, l'ubiquitinylation de certaines protéines et l'assemblage du protéasome. Le complexe aurait également un rôle dans la régulation de la transcription, il serait présent au sein de régions promotrices de gènes, des complexes d'initiation et d'élongation de la transcription et aussi associé à des facteurs de transcription. Le complexe CNOT interviendrait également dans la résolution de structures R loop qui sont des structures ADN/ARN formées lors de la transcription pouvant entraîner un stress cellulaire. Enfin parmi ses autres rôles, ce complexe serait également impliqué dans l'export des ARN et le cycle cellulaire, revue dans (Chalabi Hagkarim and Grand, 2020).

## A. Les nucléases Ccr4 et Pop2

### a. Ccr4

Ccr4 est une enzyme qui fait partie de la famille des EEP (« *exonuclease-endonuclease-phosphatase* »). Il a d'abord été montré chez *S. cerevisiae* que Ccr4 était une exonucléase ayant besoin de magnésium et d'une extrémité d'adénosine 3'OH et qui pouvait également fonctionner sur de l'ADN simple brin (Chen et al., 2002). Il existe deux orthologues chez l'humain CNOT6 et CNOT6L (Lau et al., 2009). Les études *in vitro* chez l'humain ont montré que CNOT6 hydrolysait des homo-polymères d'adénosines, qu'elle pouvait agir sur des désoxyribonucléotides mais avait une préférence pour des ribonucléotides (Wang et al., 2010).

### b. Pop2

Pop2, comme Pan2 appartient à la famille des ribonucléases DEDD. C'est l'autre nucléase du complexe qui possède également une activité catalytique. (Daugeron, 2001). Chez *S. cerevisiae* il a été montré qu'elle pouvait, en plus d'hydrolyser des poly-adénosines, hydrolyser des poly-uridines et des poly-cytosines (Thore et al., 2003). Il existe chez l'humain deux orthologues, CNOT7 et CNOT8 (Lau et al., 2009).

### c. Liaison de Ccr4 et Pop2

La résolution de la structure du complexe Ccr4-Pop2-Not1 chez *S. cerevisiae* a montré que Pop2 interagit *via* un motif riche en leucine avec la région N-terminale de Ccr4. Les domaines nucléases des deux enzymes seraient par ailleurs particulièrement conservés (Basquin et al., 2012).

## B. Les autres protéines du complexe CNOT

Les autres protéines qui composent le complexe ont un rôle dans la structure du complexe en lui-même, la liaison à l'ARN et l'interaction avec le ribosome.

Caf40 et Caf130 sont deux autres protéines associées à Ccr4 chez *S. cerevisiae*. Elles n'ont pas encore été précisément caractérisées mais elles interagiraient toutes deux avec Not1 (Chen et al., 2001).

Les protéines Not 1-5 pour « *Negative on TATA less* » ont initialement été identifiées comme des répresseurs de la transcription de gènes dépourvus de séquence TATA (Collart and Struhl, 1994). La délétion des protéines Nots impacte la déadénylation, comme *Tucker et al.* l'ont montré sur un ARN rapporteur en délétant les protéines Not2, Not3, Not4 et No5 (Tucker et al., 2002).

### a. Not1

Not1, qui est essentielle chez *S. cerevisiae* est la protéine centrale de structure qui permet le recrutement de Ccr4 et de Pop2. Elle contient dans sa partie N-terminale cinq paires d'hélices alpha antiparallèles – répétitions HEAT - qui forment un domaine MIF4G pour « *Middle domain of eukaryotic initiation factor 4G* », similaire à un domaine retrouvé dans le facteur d'initiation de la traduction eIF4G. C'est par ce domaine que serait médiée l'interaction avec Pop2. Les autres protéines Nots n'interviendraient pas dans cette liaison, car Ccr4 et Pop2 peuvent s'associer en leur absence (Bai et al., 1999; Maillet et al., 2000; Basquin et al., 2012; Petit et al., 2012). Pop2 interagit également avec Not1 chez l'humain (Chen et al., 2021).

### b. Les autres protéines Nots

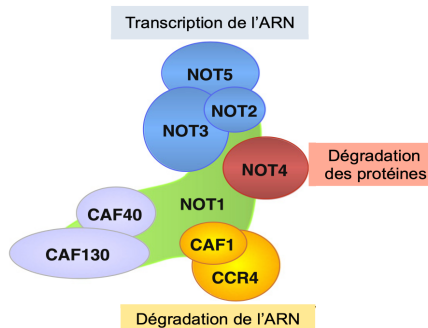
Les interactions entre les protéines Nots ont été identifiées initialement chez *S. cerevisiae* (Oberholzer and Collart, 1998). Puis la résolution de la structure du module NOT du complexe CNOT humain a montré que CNOT1, CNOT2 et CNOT3 - homologues respectivement des protéines Not1, Not2 et Not5 - interagissaient entres-elles par leurs extrémités C-terminales (Boland et al., 2013) comme cela avait été également observé chez la drosophile (Bawankar et al., 2013).

Not3 et Not5 ont de fortes similarités dans leurs extrémités N-terminales. Not3 est présente chez *S. cerevisiae* mais pas chez l'humain, sa délétion chez *S. cerevisiae* n'entraîne pas de défaut de croissance majeur (Oberholzer and Collart, 1998).

Une structure récente de Not5 a montré que cette protéine interagissait avec le ribosome chez *S. cerevisiae*. Il a par été ailleurs montré que la délétion de Not5 stabilisait spécifiquement des ARN instables. Ce rôle de Not5 lierait directement la déadénylation à la dégradation étant donné que Not5 est liée au complexe de déadénylation (Buschauer et al., 2020).

Not4, aussi appelée Mot2 chez *S. cerevisiae* est une ubiquitine ligase. Chez l'humain, un domaine RING a été caractérisé : ce type de domaine a la particularité d'être présent dans des protéines qui jouent un rôle dans l'ubiquitinylation des protéines telles que les E3 ubiquitines ligases (Hanzawa et al., 2001). La délétion de la protéine Not4 chez *S. cerevisiae*, affecte l'assemblage du protéasome ce qui a suggéré que Not4 avait un rôle

dans l'intégrité de ce complexe (Panasenko and Collart, 2011). Not4 ubiquitinye la protéine ribosomique Rps7a chez *S. cerevisiae* sans provoquer la dégradation de la protéine par le protéasome. Cette modification serait davantage importante pour la structure de la protéine en elle-même (Panasenko and Collart, 2012).



**Figure 2. Représentation du complexe CNOT.**

*Not2, Not3 et Not5 s'associeraient aux ARN pour le recrutement du complexe Ccr4-Not, Not4 est impliquée dans le contrôle de la qualité des protéines, Caf1/Pop2 et Ccr4 sont impliquées dans la dégradation des ARN. Not1 est la protéine d'échafaudage, ou structurante du complexe. Les protéines Caf40 et Caf130 n'ont pas de fonction exacte décrite. Figure et légende adaptées d'après (Collart et al., 2013).*

### II.4.1.3. Redondance des complexes PAN et CNOT

La délétion commune des deux complexes empêcherait toute déadénylation. En revanche, la délétion de l'un ou de l'autre n'empêche pas toute déadénylation même si les queues poly-A ne sont plus autant raccourcies que dans des cellules de type sauvage. Par ailleurs la délétion de Ccr4 ou de Pop2 augmente la demi-vie d'ARN rapporteurs (Tucker et al., 2001).

La présence de deux complexes de déadénylases interroge sur le rôle de chacun. Bien que catalysant la même réaction enzymatique, il semblerait que les complexes aient des préférences de substrat : selon la longueur de la queue poly-A et la présence de Pab1.

Chez les mammifères, la sur-expression de Pan2 a montré, sur un gène rapporteur, que les queues poly-A étaient raccourcies mais que ce raccourcissement s'arrêtait à 20 nucléotides. Ce n'était pas le cas quand c'était Ccr4 qui était sur-exprimée car les queues poly-A étaient complètement raccourcies. Cela a suggéré qu'en plus d'avoir une préférence pour de longues queues poly-A, le complexe PAN précède le complexe CNOT (Yamashita et al., 2005). Cela avait également été montré chez *S. cerevisiae* quelques années plus tôt par délétion des gènes codant pour Pan2 et Pan3 (Brown and Sachs, 1998) ou des gènes codant pour Ccr4 et Pop2 sur des gènes rapporteurs (Tucker et al., 2001). Plus récemment, des expériences de déadénylation *in vitro* ont confirmé que le complexe PAN intervenait plutôt sur de longues queues poly-A et le complexe CNOT sur des plus courtes queues poly-A (Schäfer et al., 2019).

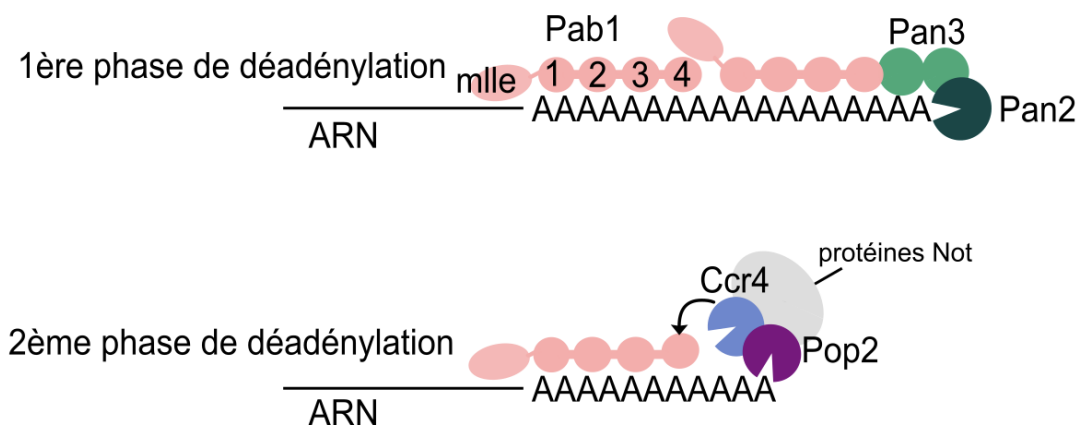
Étant donné que la déadénylation par le complexe PAN ne peut aller au-delà d'une vingtaine d'adénosines et que c'est environ l'empreinte d'une molécule de Pab1, il a été suggéré que c'était la présence de Pab1 qui empêchait un raccourcissement plus important (Tucker et al., 2001).

Chez les mammifères, une augmentation globale de la quantité d'ARN a été observée quand la sous-unité Pop2 du complexe CNOT était déplétée. Ce n'était pas le cas pour le complexe PAN et cela a suggéré que le complexe CNOT avait un effet plus important que le complexe PAN sur la demi-vie des ARN (Yi et al., 2018).

#### II.4.1.4. Les phases de déadénylation de CNOT et l'implication de Pab1

Deux études récentes publiées simultanément ont analysé les effets de Pab1 dans la déadénylation (Webster et al., 2018; Yi et al., 2018). Dans ces deux études, le raccourcissement de la queue poly-A sur des ARN synthétisés *in vitro* a été observé avec le complexe CNOT purifié.

D'après ces expériences, Ccr4 seule ne fonctionne qu'en présence de Pab1. Pop2 seule, raccourcit plus facilement des ARN dépourvus de Pab1. Lorsque les deux protéines du complexe sont présentes, les ARN sont déadénylés et encore plus rapidement quand il y a ajout de Pab1. Il semblerait donc que dans le complexe, Pop2 agisse sur les poly-A non associées avec Pab1 et Ccr4 sur les poly-A liées par Pab1. Enfin, Pab1 stimule et accélère la déadénylation par Ccr4 (**figure 3**) (Webster et al., 2018; Yi et al., 2018).



#### Figure 3. Les phases de déadénylation

*Pab1*, composée de 4 domaines RRM et d'un domaine m1-le, se lie aux queues poly-A des ARN. La déadénylation débute avec le complexe PAN. Pan3 en dimère lie les queues poly-A via son motif en doigt de zinc et interagit avec Pab1 via son domaine PAM-2. Pan2 qui possède l'activité nucléase dégrade les queues poly-A. Après un certain raccourcissement, le complexe CNOT prend le relais. Pop2 dégrade les queues poly-A dépourvues de Pab1 tandis que Ccr4 dégrade les queues poly-A liées par Pab1.

#### II.4.1.5. Influence de la queue poly-A sur la traduction et la stabilité des ARN

Le rôle de la queue poly-A dans la traduction et la stabilité des ARN est principalement médié par la protéine cytoplasmique qui les lie, Pab1.

Étant donné que Pab1 est impliquée dans la traduction des ARN *via* son interaction avec les facteurs d'initiation, il a été déterminé une relation de causalité entre la longueur

de la queue poly-A, la présence de Pab1 et la traduction. L'idée étant que plus la queue poly-A est longue, plus il y aurait de possibilités de liaisons de Pab1, plus la traduction serait importante. Dans des oocytes de drosophile. *Lim et al.* ont observé une corrélation entre l'efficacité de la traduction, mesurée par « *ribosome profiling* » - technique qui permet d'observer l'empreinte du ribosome - (Ingolia et al., 2009) et la longueur de la queue poly-A. Plus la queue poly-A était longue, plus il semblait que la traduction était importante (Lim et al., 2016).

Cependant, de récentes études tendent à montrer que la longueur de la queue poly-A n'est pas systématiquement corrélée à la traduction. Chez plusieurs organismes dont *C. elegans* et *S. cerevisiae*, *Lima et al.* ont observé que les ARN les plus exprimés – avec un taux d'expression déterminé par l'abondance des transcrits - sont ceux qui ont une queue poly-A courte, d'environ 30 nucléotides ce qui correspondrait à la liaison d'une seule molécule de Pab1. *Lima et al.* expliquent cette observation par le fait que ces ARN subiraient une étape de « *pruning* » qui consisterait en la dégradation de la queue poly-A et l'enlèvement de Pab1. Il ne resterait au final plus qu'une molécule Pab1 qui se lierait aux facteurs d'initiation de la traduction et cela aurait pour effet de promouvoir la traduction et de bloquer la dégradation (Lima et al., 2017).

Une récente étude a montré que la corrélation entre efficacité de la traduction, longueur de la queue poly-A et présence de Pab1 dépendait du stade de développement de l'organisme. Pour des oocytes de Xénope, cette dépendance serait liée au fait qu'à ce stade du développement, la quantité de Pab1 est limitante et qu'elle lierait davantage les longues queues poly-A. Cette corrélation disparaît dans des cellules matures où la quantité de Pab1 ne serait plus limitante. Aussi la présence de Pab1 sur les queues poly-A plutôt que leur longueur serait déterminante pour la traduction des ARN (Xiang and Bartel, 2021).

## II.4.2. Exemples de régulation de la déadénylation

Certaines séquences contenues dans les ARN jouent un rôle dans la régulation de la déadénylation. C'est le cas par exemple des régions AU riches (AREs) qui contiennent un fort taux d'uridines et d'adénosines. Elles pourraient recruter des protéines stimulant la déadénylation, revue dans (Garneau et al., 2007). Chez la levure *S. cerevisiae*, des régions riches en GC reconnues par les protéines Pufs permettrait une dégradation rapide des ARN *via le* recrutement du complexe CNOT. La protéine Puf5 interagit avec Pop2 et sa délétion diminue l'efficacité de la déadénylation d'un ARN rapporteur (Goldstrohm et al., 2006).

Chez les mammifères, il existe une famille de protéines dénommées BTG/Top, impliquées dans la régulation du cycle cellulaire et qui interagissent avec le complexe CNOT, revue dans (Mauxion et al., 2009). Des expériences de déadénylation *in vitro* ont montré que la protéine BTG2 du complexe BTG/Top interagissait avec PABPC1 pour stimuler la déadénylation d'un ARN rapporteur (Stupfler et al., 2016). Plus récemment, une étude a révélé que les protéines BTG1 et BTG2 pouvaient avoir un rôle dans le maintien de la quiescence des lymphocytes T : en promouvant la déadénylation des ARN et en



diminuant l'abondance des transcrits. Des mutations de ces protéines ont révélé que les queues poly-A des transcrits des lymphocytes T étaient plus longues. Cela avait pour effet d'augmenter la quantité de transcrits, ce qui serait lié à une activation des lymphocytes T (Hwang et al., 2020).

### II.4.3. Deadénylation et décapping des ARN

Le dogme actuel suggère que la déadénylation est limitante pour permettre le décapping et la dégradation *in fine* de la majorité des ARN. Cela signifie que la vitesse de déadénylation des ARN déterminerait leur vitesse de dégradation. Par exemple, un ARN instable aurait sa queue poly-A raccourcie rapidement pour initier le décapping puis la dégradation. Cela implique également que les protéines de décapping sont recrutées seulement sur des transcrits deadénylés.

Cependant il existe des exceptions à ce dogme, certains ARN étant dégradés indépendamment de la déadénylation. C'est le cas des substrats du « *nonsense mediated mRNA decay* » (NMD) détaillé ultérieurement mais également des ARN de *EDC1* et de *RPS28B* qui sont décappés et dégradés par Xrn1 sans qu'il n'y ait de déadénylation préalable. Dans le cas de l'ARN de *RPS28B*, il s'agirait d'une régulation de la quantité d'ARN produite *via* la protéine Rps28b elle-même et le facteur de décapping, Edc3 (Badis et al., 2004; Muhrad and Parker, 2005; He et al., 2014).

#### II.4.3.1. Le complexe de décapping des ARN

Le complexe de décapping Dcp1/Dcp2 est conservé chez les eucaryotes. Il hydrolyse la coiffe des ARN (m7Gppp) alors liée au premier nucléotide de l'ARN ce qui libère l'ARN avec une extrémité 5' phosphate accessible pour l'exoribonucléase Xrn1 et un m7Gpp.

L'environnement nucléotidique à proximité de la coiffe peut influencer la stabilité de l'ARN. Deux méthylations de l'adénosine suivant la coiffe ont été rapportées dans des cellules de mammifères. Ces modifications changeraient notamment la stabilité des ARN en ayant un effet protecteur contre le décapping par Dcp2 (Mauer et al., 2017).

#### A. Dcp2

Dcp2 pour « *mRNA decapping 2* » est la sous-unité catalytique du complexe de décapping. Elle a été identifiée par un crible génétique chez *S. cerevisiae* comme étant un acteur de la dégradation des ARN. Elle a notamment été caractérisée comme étant importante pour le décapping dépendant et indépendant de la déadénylation des ARN (Dunckley, 1999).

Dcp2 fait partie de la super-famille des Nudix hydrolases (Dunckley, 1999). C'est le domaine Nudix qui contient l'activité catalytique (van Dijk, 2002) avec en N-terminal ce celui-ci, un domaine régulateur (NRD) (She et al., 2006). La partie C-terminale de Dcp2 est la partie la plus divergente. La partie C-terminale, plus longue chez *S. cerevisiae* que pour la protéine humaine, contient des motifs HLMs pour « *Helical Leucine-rich Motif* » (Gaudon, 1999). Cette extrémité C-terminale pourrait avoir un rôle de régulation positive ou négative du décapping par la liaison de facteurs impliqués dans la dégradation. Par

exemple, Pat1 et Upf1 interagissent avec l'extrémité C-terminale de Dcp2 chez *S. cerevisiae* (He and Jacobson, 2015) et plus particulièrement au niveau des domaines HLMs pour Pat1 (Charenton et al., 2017).

## B. Dcp1

Dcp1 est la sous-unité régulatrice du complexe, elle permettrait un meilleur recrutement de Dcp2 aux ARN. Elle est essentielle pour permettre le décapping chez *S. cerevisiae*. Une interaction directe entre Dcp1 et Dcp2 a été révélée chez les levures *S. cerevisiae* et *Schizosaccharomyces pombe* (She et al., 2008). En absence de Dcp1, les ARN de gènes rapporteurs, s'accumulent sous forme oligo-adenylée et cappés ce qui a renforcé l'idée que Dcp1 intervenait dans le décapping des ARN après leur déadénylation (Beelman et al., 1996).

## C. Protéines associées au complexe de décapping

Des protéines Edc pour « *Enhancers of decapping* » ont été découvertes associées au décapping. Leur présence favoriserait l'hydrolyse de la coiffe des ARN. Edc1 et Edc2 sont deux protéines identifiées par crible génétique : la sur-expression des protéines avait pour effet de réduire le défaut de croissance d'une délétion de Dcp2 et leur délétion stabilisait un ARN rapporteur (Dunckley et al., 2001). La protéine Edc3 a quant à elle été identifiée par analyse *in silico* de protéines interagissant aussi bien avec le complexe Dcp1/Dcp2 qu'avec Xrn1, Dhh1 ou le complexe Lsm1-7 (Kshirsagar and Parker, 2004) Scd6/Lsm14 identifiée par crible génétique chez *S. cerevisiae* est un suppresseur du défaut de croissance induit par Edc3 (Decourty et al., 2008). Des expériences de « *tethering* » de Scd6 ont montré la stabilisation d'un ARN rapporteur chez *S. cerevisiae* (Zeidan et al., 2018). Enfin une récente étude a montré par microscopie à fluorescence que la délétion de Scd6 conduisait à une relocalisation de Dcp2 dans le noyau (Tishinov and Spang, 2021) Pby1 serait un autre activateur du décapping chez *S. cerevisiae*, elle interagirait avec le domaine Nudix de Dcp2 (Charenton et al., 2020).

## D. Formation du complexe de décapping

Des études structurales chez plusieurs organismes ont montré que le complexe de décapping pouvait adopter des conformations différentes selon la liaison des protéines Edc et que cela aurait pour effet de contribuer à l'activation du décapping.

Chez *S. pombe*, Dcp2 contient des séquences auto-inhibitrices dans sa partie C-terminale. L'auto-inhibition pourrait être levée *via* un changement de conformation du complexe par la liaison des protéines : Edc1 et Edc3 (Paquette et al., 2018).

La structure d'un complexe Dcp1/Dcp2 avec Edc3 a par ailleurs été déterminée chez la levure, *Kluyveromyces lactis*. Dans cette structure, Dcp2 lie Dcp1 par son extrémité N-terminale et Edc3 par son extrémité C-terminale. Enfin, ce serait au niveau du domaine catalytique Nudix que viendrait se placer la coiffe de l'ARN pour le décapping. Edc3 aurait un rôle stimulateur pour le décapping (Charenton et al., 2016). Une autre forme « active » du complexe de décapping a été décrite structurellement par la formation d'un complexe contenant les protéines Edc1, Dcp1, Dcp2 et Edc3 chez la levure *K. lactis* (Mugridge et al.,



2018). En parallèle, chez *S. pombe*, une structure contenant Edc1, Dcp1 et Dcp2 qui favoriserait la conformation active de Dcp2 a été décrite (Valkov et al., 2016). Chez l'humain, le décapping serait facilité par un assemblage de Dcp1 et Dcp2 avec Edc4 (Chang et al., 2014).

#### II.4.3.2. Le complexe **Lsm1-7** et les protéines associées, un lien entre la queue poly-A et le décapping

##### A. Le complexe **Lsm1-7**

Le complexe Lsm1-7 est constitué de sept protéines Lsm (Like Sm) numérotées de 1 à 7 qui s'assemblent sous la forme d'un anneau. Il lie de courtes queues poly-A et serait responsable du déclenchement du décapping. C'est un complexe cytoplasmique à l'inverse du complexe Lsm2-8 spécifique du noyau qui intervient dans l'épissage (Mayes, 1999). La caractérisation de délétions ou de mutations qui inactivaient les protéines Lsm1 à 7 a montré que des ARN déadénylés et cappés s'accumulaient dans des expériences réalisées sur des ARN rapporteurs. Il en a été déduit que le complexe Lsm1-7 participait au déclenchement du décapping des ARN. (Boeck et al., 1998; Bouveret, 2000; Tharun et al., 2000; He and Parker, 2001). Par ailleurs, les ARN qui s'accumulaient avaient des queues poly-A d'une taille comprise entre 10 et 20 nucléotides. Aussi il en a été déduit que c'était à cette taille que le décapping se déclenchait (Tharun et al., 2005). Enfin, des expériences *in vitro* avec le complexe Lsm1-7 purifié ont montré qu'il avait une préférence pour des formes oligo-adénylées ou oligo-uridylées (Chowdhury et al., 2007). Des interactions entre le complexe Lsm1-7 et le complexe de décapping ont été mises en évidence par double-hybride (Fromont-Racine et al., 2020).

L'accumulation de queues poly-A courtes associée à des ARN cappés couplée au fait que le complexe Lsm1-7 a une préférence de liaison à des ARN oligo-adénylés a suggéré que ce complexe liait les ARN en fin de déadénylation pour déclencher le décapping.

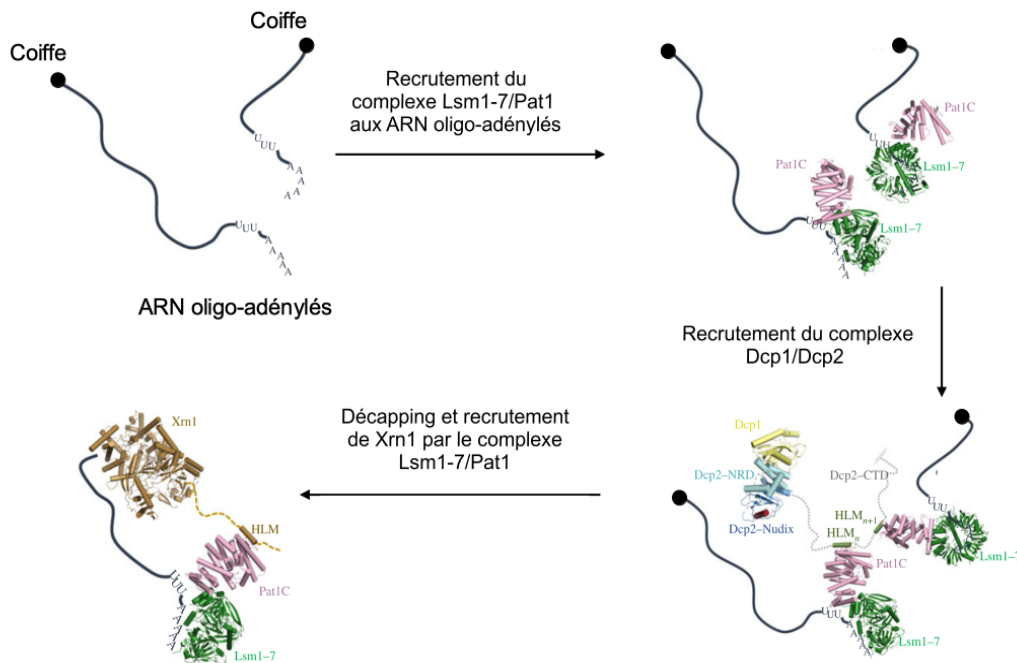
##### B. **Pat1** et le complexe **Lsm1-7**

Pat1 serait à la fois impliquée dans la liaison aux courtes queues poly-A des ARN et le décapping.

Pat1 interagit avec le complexe Lsm1-7 et sa présence semble augmenter l'affinité du complexe Lsm1-7/Pat1 pour sa liaison à l'ARN (Chowdhury et al., 2014). Chez *S. pombe*, des données de structure ont montré que Pat1 interagit également avec le complexe de décapping *via* le domaine HLMs C-terminal de Dcp2. L'interaction aurait pour effet de lever l'autoinhibition du complexe de Dcp1/Dcp2 et de promouvoir le décapping (Lobel et al., 2019). Pat1 pourrait également interagir avec Xrn1, revue dans (Charenton and Graille, 2018). La **figure 4** résume les interactions entre le complexe de décapping et le complexe Lsm1-7/Pat1.

En plus d'être associée au complexe Dcp1-2 et au complexe Lsm1-7, Pat1 a été retrouvée associée aux facteurs d'initiation de la traduction eIF4E et eIF4G et ce de manière dépendante de l'ARN (Tharun and Parker, 2001). En absence de Pat1, la

répression de la traduction induite par une carence en glucose n'est plus observée tandis que sa sur-expression inhibe la traduction suggérant que cette protéine pouvait avoir un rôle charnière dans la traduction et le déclenchement de la dégradation des ARN (Coller and Parker, 2005).



**Figure 4. Mécanisme du décappage dépendant du complexe Lsm1-7/Pat1**

Le complexe Lsm1-7/Pat1 est recruté sur un ARN déadénylé et s'associe au complexe de décappage et à Xrn1 pour la dégradation 5'-3' des ARN. Figure adaptée d'après (Charenton and Graille, 2018).

### II.4.3.3. Connexion entre les facteurs de la déadénylation et du décappage

Le lien entre la déadénylation et le décappage ne serait pas uniquement médié par le complexe Lsm1-7 qui lie des ARN déadénylés. En effet en plus de cela, des connexions entre les facteurs de décappage et de la déadénylation ont été observées.

#### A. Implication des protéines Nots

L'ARN du gène *EDC1* serait soumis à un mécanisme de dégradation indépendant de la déadénylation et a été utilisé comme substrat pour l'étude du décappage (Muhlrad and Parker, 2005). Les délétions de Not2, Not3 et Not5 stabilisent l'ARN de *EDC1*, suggérant que ces protéines participaient au déclenchement du décappage (Alhusaini and Coller, 2016).

## B. Implication de Dhh1 dans la dégradation des ARN

Dhh1 est une ARN hélicase de la famille DEAD (Strahl-Bolsinger and Tanner, 1993). Des expériences de co-immunoprécipitation ont montré que cette protéine interagissait avec Dcp1, Pop2, Lsm1, Pat1 et Not1 (Coller et al., 2001; Maillet and Collart, 2002). La résolution de la structure de Dhh1 a mis en évidence une interaction avec la partie N-terminale de Pat1 (Sharif et al., 2013).

En absence de Dhh1, l'ARN du gène *MFA2* est stabilisé contrairement à ce qui est observé pour un gène rapporteur sensible au NMD. Cela a suggéré que Dhh1 avait un rôle dans le recyclage des ARN, indépendamment du NMD. De plus, en absence de Dhh1 il y a stabilisation de formes capées et déadénylées suggérant un rôle potentiel de Dhh1 dans le déclenchement du décapping (Coller et al., 2001). *In vitro*, Dhh1 stimule le décapping médié par Dcp1 (Fischer, 2002). L'ensemble de ces observations suggèrent l'existence d'un lien entre Dhh1 et le décapping.

Dhh1 a ensuite été identifiée comme une protéine intervenant dans la reconnaissance de ribosomes ralentis. L'addition d'une succession de codons rares au sein d'un ARN provoque sa déstabilisation dans une cellule de type sauvage mais cet effet est moins marqué dans une souche présentant une délétion de Dhh1, suggérant que cette protéine reconnaît les ARN traduits moins rapidement et que cela déclencherait leur dégradation (Sweet et al., 2012). Plus récemment, Dhh1 a été identifiée comme étant la protéine qui était responsable du déclenchement de la dégradation d'ARN ayant des codons peu « optimisés » (Radhakrishnan et al., 2016). L'optimisation des codons et son impact sur la stabilité des ARN sont détaillés ultérieurement.

Dhh1 serait la principale protéine qui connecterait la déadénylation au décapping. Elle est présumée comme étant l'élément déclencheur du décapping et de la déadénylation dans le modèle dépendant du raccourcissement de la queue poly-A, revues dans (Charenton and Graille, 2018; Vidya and Duchaine, 2022).

### II.4.3.4. Vitesse de déadénylation et décapping

Chez *S. cerevisiae*, des observations en dynamique après induction puis arrêt transcriptionnel réalisées dans les années 90 pour observer le raccourcissement des queues poly-A ont mené à l'hypothèse que la déadénylation était importante pour la dégradation des ARN.

*MFA2* est considéré comme étant un ARN instable (demi-vie de 3,5 minutes). Il a été observé que des mutations dans son 3'UTR augmentaient la demi-vie du transcrit et que le profil de déadénylation changeait avec une déadénylation qui semblait plus lente. Cela a suggéré que *MFA2* devait contenir, dans sa région 3'UTR, des éléments stimulants pour la déadénylation et responsables de son instabilité (Muhlrad and Parker, 1992). En parallèle, l'accumulation d'espèces oligo-adénylées pour un autre rapporteur, *PGK1* considéré comme un ARN stable (demi-vie de 45 minutes) et l'absence de courtes queues poly-A pour *MFA2* a mené Muhlrad et al. à formuler l'hypothèse qu'un ARN instable était déadénylé plus rapidement pour initier sa dégradation qu'un ARN stable. Par ailleurs, le

déclenchement du décapping n'était observé pour ces deux rapporteurs que quand l'ARN avait ses queues poly-A complètement raccourcies (Muhlrad et al., 1994, 1995).

Chez *S. cerevisiae* et chez les mammifères, deux études ont montré que les ARN décappés présentaient des queues poly-A courtes. Pour les espèces cappées en revanche, les queues poly-A étaient plus longues et de tailles variables. L'ensemble de ces observations suggère que la déadénylation se produit sur des ARN encore cappés et que non seulement la déadénylation précède mais est aussi requise pour initier le décapping (Couttet et al., 1997; Bouveret, 2000).

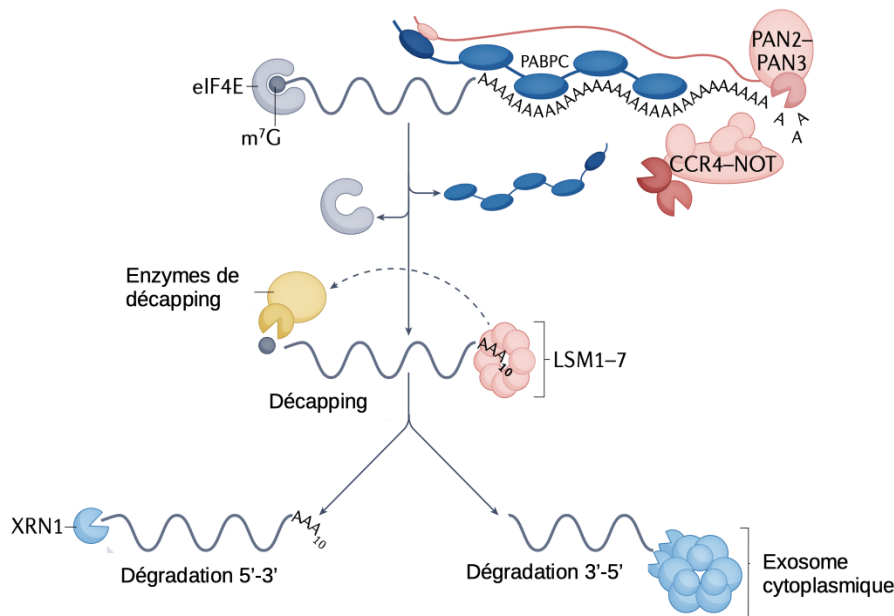
Différents modèles mathématiques ont été créés pour tenter de modéliser et de prédire l'accumulation des ARN cytoplasmiques à l'état d'équilibre ou « *steady state* ». Dans le modèle établi par *Cao et Parker*, en utilisant les données obtenues sur les gènes *MFA2* et *PGK1*, de petites variations dans la vitesse de déadénylation ont des effets majeurs sur l'accumulation des ARN dans le cytoplasme, suggérant que la déadénylation serait l'étape clé du processus de la dégradation des ARN cytoplasmiques (Cao and Parker, 2001). Plus récemment, de nouvelles simulations de dégradation des ARN ont été réalisées d'après des mesures de demi-vies et de queues poly-A à grande échelle dans des cellules humaines. Les auteurs proposent que la vitesse de la déadénylation est similaire pour les ARN stables et instables, jusqu'à 50 nucléotides. En dessous de ce seuil, la déadénylation et la dégradation sont plus rapides pour les ARN instables que pour les ARN stables (Eisen et al., 2020).

## II.5. Les voies de dégradation 5'-3' et 3'-5'

D'après le modèle canonique représenté en **figure 5**, la dégradation des ARN après la déadénylation, peut avoir lieu selon deux voies.

Dans la voie de dégradation 5'-3', la liaison du complexe Lsm1-7/Pat1 aux courtes queues poly-A générées par l'action des déadénylases permettrait le recrutement du complexe Dcp1/Dcp2 pour le décapping des ARN. Une fois décappés, les ARN qui présentent une extrémité 5' phosphate libre sont soumis à une dégradation dépendante de l'exoribonucléase Xrn1 (He and Jacobson, 2001). Xrn1 n'utiliserait comme substrat que des ARN 5' monophosphates (Jinek et al., 2011).

L'autre voie de dégradation dégrade les ARN dans le sens 3'-5'. Elle est effectuée par l'exosome cytoplasmique qui est un complexe conservé composé de plusieurs sous-unités, dont Dis3 qui possède des activités endonucléase et exonucléase. L'exosome est accompagné des protéines activatrices Ski. Même si l'exosome dégrade les ARN dans le sens 3'-5', comme les déadénylases, la déadénylation préalable serait nécessaire pour dissocier Pab1 des queues poly-A et rendre accessible l'extrémité 3' de l'ARN à l'exosome. Cette voie est considérée comme mineure car elle interviendrait pour dégrader les ARN qui ne seraient pas dégradés par la voie 5'-3', par exemple lorsque la disponibilité du complexe de décapping ou de Xrn1 est affectée, revue dans (Łabno et al., 2016).



### Figure 5. Dogme actuel pour la dégradation des ARN chez *S. cerevisiae*

La déadénylation précède le décapping des ARN et est un prérequis pour la dégradation. Les ARN sont d'abord déadénylés par le complexe PAN (Pan2 et Pan3) puis par le complexe CNOT (Ccr4, Pop2 et les protéines Nots). Les courtes queues poly-A peuvent ensuite être liées par le complexe Lsm1-7 ce qui déclencherait le décapping par Dcp1/Dcp2. La dégradation est effectuée ensuite par Xrn1 dans le sens 5'-3' ou par l'exosome cytoplasmique dans le sens 3'-5'. Figure et légende adaptées d'après (Passmore and Collier, 2022).

## II.6. Régulation de la transcription par la quantité d'ARN cytoplasmique

Depuis quelques années, les notions de rétro-contrôle, « *buffering* », et marquage transcriptionnel des ARN pour la dégradation ont émergé.

*Sun et al.* ont observé que, dans des souches de levure présentant des délétions de différentes protéines impliquées dans la dégradation et le métabolisme des ARN, les modifications induites par ces mutations sur la dégradation des ARN étaient compensées par des modifications de leur synthèse. Ces observations suggèrent que la dégradation cytoplasmique des ARN pourrait impacter leur transcription et qu'il existe un rétro-contrôle de la transcription des ARN dans le noyau par leur quantité dans le cytoplasme. Ce rétro-contrôle aurait lieu en réponse à certains stress ou infections qui impacteraient la quantité

d'ARN cellulaires. Bien que le mécanisme responsable de ce « *buffering* » ne soit pas élucidé, il semble que Xrn1 joue un rôle clé puisque contrairement aux autres protéines testées, la délétion de cette protéine empêche la dégradation des ARN et augmente leur transcription entraînant une forte accumulation d'ARN. Cet effet serait en partie médié par Nrg1, un répresseur général de la transcription (Sun et al., 2012, 2013).

En réalisant des mesures de demi-vie de certains ARN par marquage métabolique couplées à des immuno-précipitations de la chromatine (CHIP-seq) pour regarder l'occupation des promoteurs par l'ARN polymérase II, *Abernathy et al.* ont montré que chez les mammifères, l'expression d'une endonucléase virale provoquait à la fois la dégradation d'ARN cellulaires et une diminution de la transcription. Cette dégradation serait dépendante de Xrn1 (Abernathy et al., 2015). Il a également été montré qu'en réponse à ce stress, certaines protéines cytoplasmiques telles que PABPC1 étaient relocalisées dans le noyau. Cette présence nucléaire pourrait réprimer la synthèse de certains gènes. Le rôle de ce rétro-contrôle n'a pas encore été élucidé (Gilbertson et al., 2018).

### III. Traduction des ARN et ses liens avec la dégradation

La traduction des ARN est l'étape qui permet la synthèse des protéines, grâce à un complexe ribonucléoprotéique, le ribosome. Chez les cellules eucaryotes, elle se déroule au sein du cytoplasme. La traduction est un processus séquentiel mais elle peut se diviser en trois étapes : la reconnaissance de l'ARN, sa traduction et l'arrêt de la traduction (**figure 6.A**).

Chez *S. cerevisiae*, le ribosome est composé d'une petite sous-unité 40S (S pour Svedberg, qui désigne le coefficient de sédimentation de la particule) composée de protéines et d'un ARN ribosomique (ARNr) 18S et d'une grande sous-unité composée de protéines et des ARNr 5S, 5,8S et 25S. L'assemblage des deux sous-unités du ribosome est appelé 80S. La sous-unité 40S est la première à intervenir dans la traduction, elle reconnaît notamment l'ARN à traduire et le codon AUG initiateur. La grande sous-unité 60S vient s'associer dans un second temps pour ce qui marque le début de la synthèse protéique. Trois sites dans le ribosome se distinguent, dans lesquels vont se placer les ARN de transfert (ARNt) au cours de la traduction. Le site aminoacyl A où viennent initialement se placer les ARNt chargés d'un acide aminé, le site peptidyl-transférase P où est réalisée la liaison peptidique entre l'acide aminé situé au site A et celui au site P, le site de sortie (« *exit* »), E où les ARNt sont relâchés.

#### III.1. Initiation de la traduction

Plusieurs facteurs d'initiation vont se coordonner pour se lier à l'ARN et former le complexe d'initiation de la traduction eIF4F. Il est notamment composé de eIF4G, d'une hélicase, eIF4A, assistée par eIF4B, et de eIF4E/CDC33 qui est la protéine cytoplasmique



de liaison à la coiffe des ARN. Au total 11 facteurs d'initiation interviendraient pour démarrer la traduction, revue dans (Dever et al., 2016).

Pab1 *via* son domaine RRM2 participe au recrutement de l'un d'entre eux, eIF4G (Tif4631/Tif4632) chez *S. cerevisiae* (Tarun and Sachs, 1996) et chez l'humain (Safaei et al., 2012). eIF4G se lie à un autre facteur d'initiation, eIF4E (Fortes et al., 2000). La liaison de ces différents facteurs provoquerait la formation d'un complexe dit « *closed-loop* » ce qui signifie que l'ARN adopte une conformation circulaire (Wells et al., 1998).

D'autre part, se forme le complexe de pré-initiation 43S aussi appelé PIC. Ce complexe contient la petite sous-unité 40S du ribosome et un complexe dit « ternaire » composé de la protéine eIF2 associée à une molécule de GTP et d'un ARN de transfert méthionine dit « initiateur » (ARNt<sub>i</sub><sup>Met</sup>). L'ARN de transfert chargé se place plus précisément au niveau du site P de la sous-unité 40S. Enfin le PIC lie l'ARN dans sa partie 5' non traduite (5'UTR), formant alors le complexe 48S. Ensuite une étape de « *scanning* », vers le 3' de l'ARN consiste en la recherche d'un codon AUG « *start* » par le PIC grâce à l'anticodon UAC de l'ARNt<sub>i</sub><sup>Met</sup>. Les nucléotides juste en amont du codon « *start* » ont une importance dans l'efficacité de la traduction et définissent une séquence qui promeut la reconnaissance du codon « *start* » appelée séquence consensus de « Kozak ». Une fois que le codon AUG « *start* » est reconnu, eIF2 associé à du GDP est relargué du complexe et la 60S est recrutée pour former le ribosome ce qui initie la phase d'élongation de la traduction, revue dans (Dever et al., 2016).

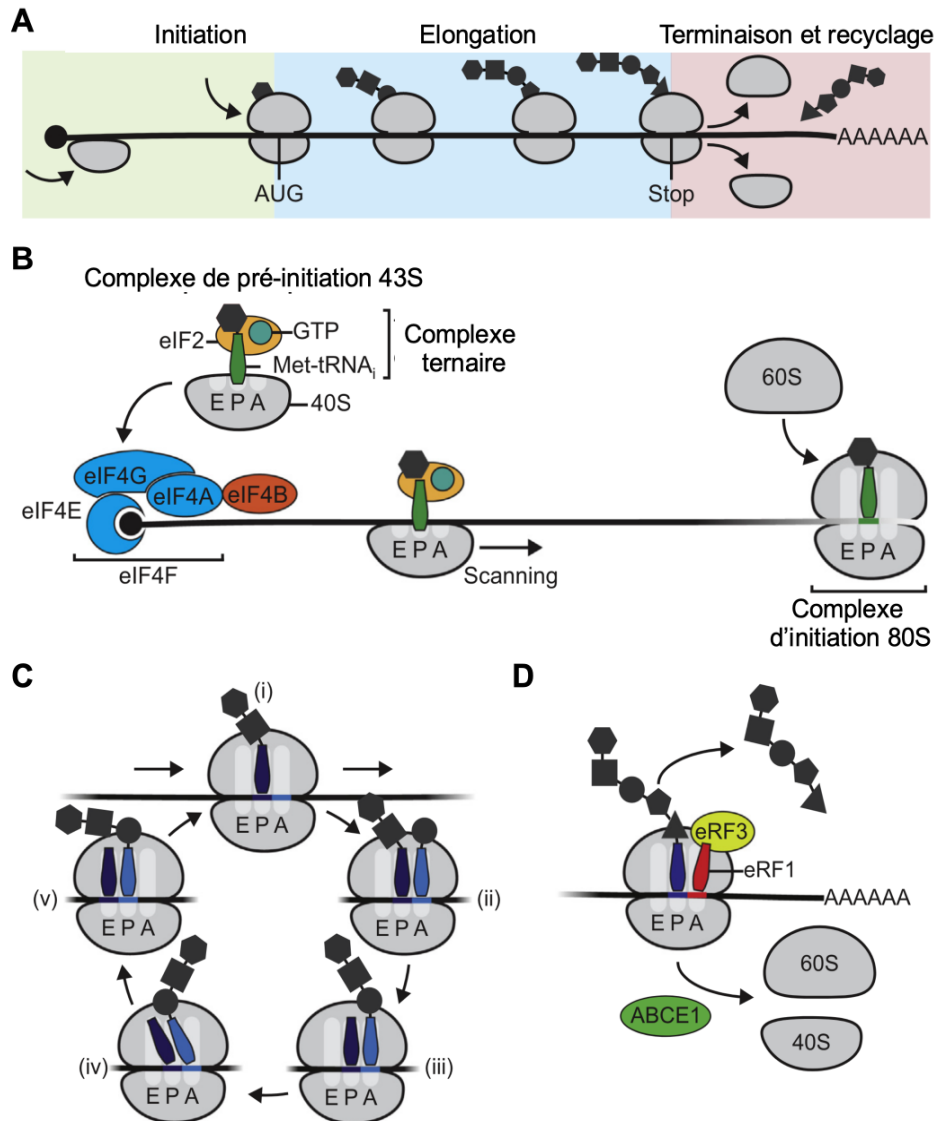
## III.2. Elongation de la traduction

L'élongation de la traduction est l'étape qui va consister en la synthèse de la protéine, composée d'une succession d'acides aminés. Alors que l'ARNt<sub>i</sub><sup>Met</sup> est positionné sur l'AUG au sein du site P du ribosome, un nouvel ARNt chargé associé au facteur d'élongation eEF1A et de GTP se positionne au niveau du site A du ribosome. Lors de la reconnaissance codon-anticodon, le GTP est hydrolysé et eEF1A est relâché de la structure puis recyclé par eEF1B pour un nouveau cycle. Durant la formation de la liaison peptidique, le peptide naissant situé au niveau du site P du ribosome est transféré au nouvel aminoacyl ARNt situé au site A. eIF5a/Hyp2 qui est un autre facteur d'élongation viendrait se lier au sein du site E du ribosome ce qui aurait pour effet de catalyser la liaison peptidique entre les deux acides aminés. Le ribosome adopterait alors une conformation « hybride ». eEIF2-GTP viendrait s'associer au site A ce qui aurait pour effet de transloquer les ARNt situés respectivement aux sites A et P aux sites P et E. Après la dissociation de eEIF2-GDP, le site A est disponible pour accepter un nouvel ARNt (**figure 6.B**), revue dans (Dever et al., 2018).

## III.3. Terminaison de la traduction

La terminaison de la traduction intervient quand un codon stop (UAA, UGA, ou UAG) se situe au site A du ribosome. eRF1/Sup45 et eRF3/Sup35 sont les deux facteurs de terminaison. Les deux protéines s'associent pour former le complexe ternaire

eRF1/eRF3-GTP. L'hydrolyse du GTP permet le positionnement de eRF1 au niveau du codon stop et eRF3 est relâché (**figure 6.C**). Enfin le facteur eRF1 en coordination avec un facteur qui intervient pour recycler le ribosome, Abce1/Rli1 permettrait l'hydrolyse de la liaison peptidique et la libération de la protéine. Enfin l'hydrolyse d'ATP par Abce1/Rli1 permet la dissociation du ribosome (**figure 6.D**), revue dans (Dever and Green, 2012).



**Figure 6. Mécanismes de l'initiation, de l'élongation et de la terminaison de la traduction.**

(A) Vue d'ensemble des différentes phases de la traduction. (B) Schéma de l'initiation, (C) de l'élongation, (D) de la fin de la traduction et du recyclage. Figure et légende adaptées d'après (Sonneveld et al., 2020).



### III.4. Optimisation des codons

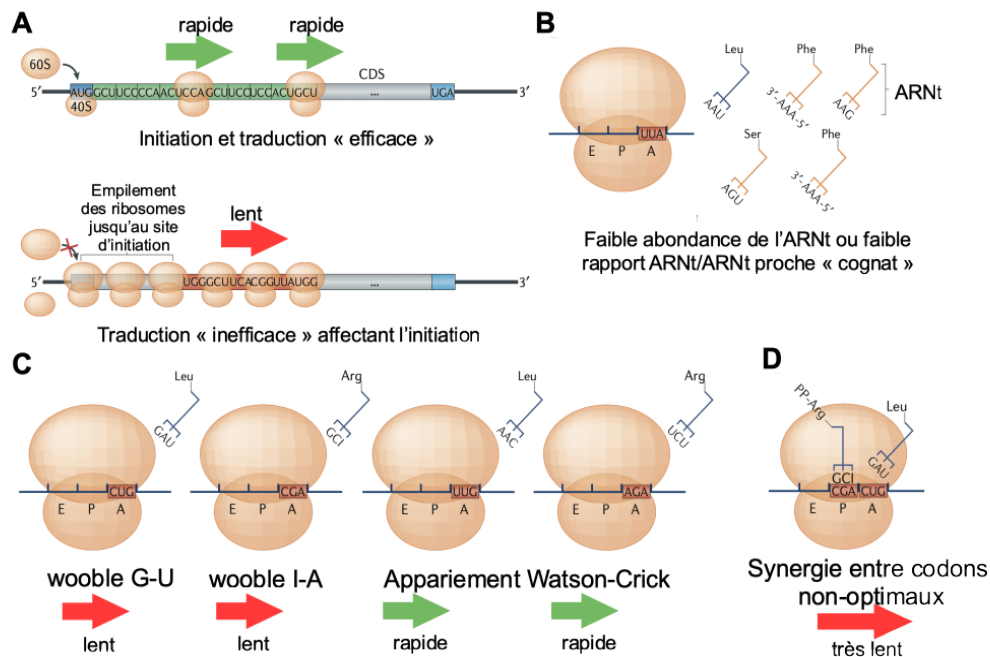
Le code génétique repose sur la correspondance entre un acide aminé et un triplet de nucléotides sur l'ARNm, appelé codon. Il existe  $4^3$  soit 64 codons différents dont trois qui correspondent aux codons stop - AUG, UGA et UAG - et les 61 autres aux 20 acides aminés. Ce code génétique est dit « dégénéré » car il peut exister plusieurs codons dits « synonymes » pour un même acide aminé. Certains codons et leurs ARNt associés sont très abondants, il s'agit du « biais de codon ». Dans des gènes très exprimés tels que les gènes codant pour les protéines ribosomales il y a une sur-représentation de codons qui seraient efficacement traduits, à l'inverse par exemple des cibles du NMD, qui ne codent pas souvent pour une protéine fonctionnelle.

Un des enjeux de la synthèse protéique est la bonne reconnaissance du codon par l'ARNt chargé correspondant. C'est ce qui permettra d'introduire dans le polypeptide l'acide aminé correct. Pour cela, plusieurs ARNt chargés vont venir se placer au sein du site A du ribosome et ce jusqu'à ce que l'ARNt *via* sa boucle anticodon soit associé au codon correspondant. L'ARNt est alors dit « *cognate* ». Lors de cette étape, des ARNt « *non-cognate* » (pas d'appariement possible avec le codon) et « *proche-cognate* » (deux appariements possibles avec le codon) peuvent se placer au sein du ribosome. Ces ARNt « *non ou proche-cognate* » peuvent se lier partiellement au codon de part les interactions entre bases. Si le codon lu est rare et que des ARNt « *proche-cognate* » sont abondants, alors la reconnaissance du codon par l'ARNt correspondant sera lente. D'où l'idée que les codons rares sont décodés moins rapidement et que le ribosome resterait plus longtemps sur ces codons (**figure 7.A**), revues dans (Hanson and Coller, 2018; Bae and Coller, 2022).

Plusieurs méthodes ont été utilisées pour déterminer le biais dans l'usage des codons. Une des premières utilisées est le « *Codon Adaptation Index* » (CAI) qui permet de déterminer pour chaque acide aminé la fréquence d'utilisation de chacun de ses codons en utilisant comme référence les gènes les plus exprimés. Le principe étant que plus un codon est présent dans un gène exprimé, plus son utilisation est importante (Sharp and Li, 1987). Le « *tRNA adaptative index* » (tAi) est un autre calcul d'usage des codons qui se base sur les appariements codons-anticodons et prend en compte à la fois les liaisons canoniques *Watson-Crick* (définies par les paires adénine/uracile et guanine/cytosine) et non canoniques, *wobble*. Les appariements non canoniques interviennent quand un ARNt peut s'apparier à plusieurs codons de par la permissivité de la troisième base de l'anticodon. Par exemple l'inosine est une base modifiée de la boucle anticodon qui peut s'apparier avec l'uracile, l'adénine et la cytosine. Le tAi reflète ainsi la disponibilité des ARNt pour un codon (**figure 7.B**) (Reis, 2004). *Pechmann et Frydman* ont par la suite défini un seuil à partir duquel un codon était considéré comme optimal avec l'idée qu'un codon optimal avait une disponibilité d'ARNt en excès par rapport à son usage (Pechmann and Frydman, 2013).

L'optimisation des codons affecte la synthèse protéique et la stabilité des ARN. L'optimalité des codons aurait un rôle sur la vitesse de synthèse des protéines : comme les

codons non optimaux sont décodés plus lentement, les ARNs sont traduits moins rapidement (**figure 7.C**). Un décodage ralenti en fonction des codons a par exemple été montré chez *E. coli* par marquage radioactif des protéines néosynthétisées (Sørensen and Pedersen, 1991). Des expériences de « *ribosome profiling* » chez *S. cerevisiae* ont montré par la suite que les ARN qui avaient des codons rares et donc moins d'ARNt correspondants étaient décodés plus lentement (Weinberg et al., 2016). Par ailleurs, la présence de paires de codons non-optimaux semble avoir un effet multiplicateur sur une baisse de l'efficacité de la traduction (**figure 7.D**). Chez *S. cerevisiae*, 17 paires de codons inhibitrices pour la traduction ont été identifiées (Gamble et al., 2016). D'autre part, des cinétiques de dégradation chez *S. cerevisiae* ont montré que la stabilité des ARN dépendait de l'usage des codons. (Hoekema et al., 1987). Enfin, une analyse bio-informatique utilisant des données de séquençage à grande échelle a mis en évidence une corrélation entre le pourcentage de codon optimaux et la demi-vie des ARN qui a ensuite été validée expérimentalement. Cette analyse a permis de définir le « *stability correlation coefficient* » (CSC) corrélant la demi-vie des ARN à leur usage de codons (Presnyak et al., 2015).



**Figure 7. L'usage des codons influence la traduction via des mécanismes dépendants des ARNt**

**(A)** Les codons optimaux (vert) permettent une translocation rapide du ribosome, libérant ainsi l'extrémité 5' de la séquence d'ARN codante (CDS). Les codons non optimaux (rouge) peuvent ralentir la translocation des ribosomes alors que ces derniers « attendent » un ARNt rare ce qui peut créer une accumulation de ribosomes. **(B)** Lorsque l'ARNt reconnu par le codon est rare dans le pool d'ARNt ou en faible quantité par rapport aux espèces d'ARNt « proche-cognat », la traduction peut être ralentie en raison de l'augmentation du temps nécessaire pour que le codon soit reconnu par l'ARNt correspondant. **(C)** Les codons qui doivent être décodés par les espèces d'ARNt qui ne forment pas de paires de bases canoniques Watson-Crick en troisième position du codon provoquent un ralentissement du ribosome, en raison du temps accru nécessaire à l'incorporation des ARNt wobble. Deux interactions wobble communes chez la levure sont illustrées, G-U et G-I en comparaison aux paires de bases traditionnelles Watson-Crick, qui sont associées à une traduction plus efficace. **(D)** Les doublets de codons non optimaux induisent un ralentissement du ribosome encore plus important que lorsqu'ils sont isolés. Légende et figure adaptées d'après (Hanson and Collier, 2018).

## III.5. Traduction et dégradation

Il est désormais établi que la dégradation des ARN est étroitement liée à la traduction. Cette dégradation dépendante de la traduction peut s'envisager par le fait qu'un ARN messager doit d'abord avoir permis la production de protéines avant d'être dégradé.

### III.5.1. La dégradation co-traductionnelle

En analysant les ARN présents dans différentes fractions de gradients de saccharose qui permettent de séparer les ARN associés aux ribosomes des ARN libres, *Hu et al.* ont observé que des ARN décappés pouvaient être associés avec des ribosomes en cours de traduction. Ils ont également observé la présence de fragments de dégradation dans ces fractions de gradient suggérant que le décapping et la dégradation intervenaient sur des ARN en cours de traduction (Hu et al., 2009). Par ailleurs, la dégradation des ARN instables dans les polysomes avait été démontrée auparavant (Zhang et al., 1997).

Quand l'élongation de la traduction est inhibée par l'ajout de cycloheximide, une drogue qui empêche la translocation du ribosome, les ARN sont stabilisés et ce sont principalement des espèces cappés qui sont observées. Bien que cette drogue affecte la traduction, il est possible que l'effet observé sur la stabilisation des ARN ne soit pas dû directement au blocage de l'élongation de la traduction mais à l'inhibition de la synthèse d'un ou de plusieurs facteurs de dégradation des ARN. Dans ce cas l'effet sur la stabilisation des ARN serait indirect (Chan et al., 2018).

Le lien entre la traduction et la dégradation a été étudié par l'analyse de fragments d'ARN en cours de dégradation. Lorsque les ARN sont décappés puis dégradés par des exoribonucléases telle que Xrn1, cela génère des extrémités 5' phosphate libres. Cette caractéristique a été utilisée pour isoler et identifier par séquençage les ARN en cours de dégradation après la ligation d'une amorce spécifique à cette extrémité (Pelechano et al., 2015).

L'utilisation de cette technique a permis de montrer que les fragments de dégradation chez *S. cerevisiae* s'accumulaient avec une périodicité de 3 nucléotides dans les parties codantes des ARN alors que ce n'était pas le cas pour les régions non traduites. Par ailleurs, il y avait une forte accumulation de fragments 17 nucléotides en amont du codon stop des séquences codantes. Ces deux observations ont conduit *Pelechano et al.* à suspecter l'implication de la traduction dans la dégradation : la périodicité de 3 nucléotides est également celle observée lors du décodage des codons et de la translocation du ribosome. L'accumulation de fragments spécifiquement avant le stop pourrait s'expliquer par la persistance du ribosome au stop. En effet, les analyses par « *ribosome profiling* » ayant montré que le ribosome couvrait environ une trentaine de nucléotides (Ingolia et al., 2009), un ribosome « bloqué » au codon stop protégerait une séquence d'environ  $15 \pm 2$  nucléotides de part et d'autre de celui-ci ce qui est compatible avec la distance de 17 nucléotides en amont du codon stop observée. Dans un mutant de délétion pour Xrn1, cette périodicité est perdue alors qu'elle est renforcée après inhibition de la traduction par ajout de cycloheximide (Pelechano et al., 2015). L'ensemble de ces

résultats suggère que la dégradation des ARN par Xrn1 est co-translationnelle, et que la périodicité de 3 nucléotides observée serait la conséquence du déplacement du ribosome, protégeant la portion d'ARN auquel il est associé de la dégradation par cette exonucléase. Une interaction physique entre Xrn1 et le site de sortie du ribosome a par ailleurs été mise en évidence par cryo-microscopie électronique (Tesina et al., 2019).

### III.5.2. La dégradation des ARN par le NMD

Si la corrélation entre la séquence de l'ARN et sa stabilité est maintenant clairement établie, la première observation de cette corrélation remonte à la fin des années 70 chez *S. cerevisiae* avec la découverte du « *nonsense mediated mRNA decay* » (NMD) par l'introduction d'un codon non-sens au sein du gène codant pour la protéine URA3 (Losson and Lacroute, 1979). Le NMD est un mécanisme de contrôle de qualité des ARN dépendant de la traduction qui dégrade spécifiquement les ARN ayant une courte séquence codante suivie d'un long 3'UTR. Outre les ARN contenant des mutations non-sens, les ARN ayant de petites phases codantes en amont de la séquence canonique (uORF pour « *upstream open-reading-frame* » (Malabat et al., 2015)), les ARN épissés qui ont conservé des introns, et certains ARN non codant issus de la transcription pervasive tels que les XUTs (« *Xrn1-sensitive unstable transcript* », (van Dijk et al., 2011)) et les SUTs (« *stable unannotated transcripts* », (Neil et al., 2009; Xu et al., 2009)) présentent la caractéristique d'avoir une courte séquence codante suivie d'un long 3' UTR et constituent des cibles du NMD, revue dans (Celik et al., 2017).

Upf1, Upf2 et Upf3 sont les trois protéines essentielles au NMD chez *S. cerevisiae*, (Leeds et al., 1991, 1992; Cui et al., 1995). Ces protéines sont dénommées ainsi car les mutations « upf » ont initialement été identifiées par crible génétique chez *S. cerevisiae* : les mutations de ces gènes augmentaient la fréquence de glissement (« *frameshift* ») du ribosome (Culbertson et al., 1980). Upf1 est impliquée dans toutes les étapes du NMD et est de ce fait considérée comme la protéine clé du NMD. Upf1 est une ARN/ADN hélicase ATP-dépendante qui possède à son extrémité N-terminale une région en doigt de zinc riche en cystéines et histidines (CH). Cette extrémité constitue une plateforme d'interaction avec d'autres protéines telles que Upf2, eRF1, eRF3, Dcp2, Ebs1 et Nmd4 chez *S. cerevisiae*. (He and Jacobson, 1995; He et al., 1996; Czaplinski et al., 1998; He and Jacobson, 2015; Dehecq et al., 2018). Le domaine C-terminal, de Upf1 contient une région riche en sérines et glutamines (SQ) qui interviendrait dans la régulation de l'activité hélicase, revue dans (Lavysh and Neu-Yilik, 2020).

Chez *S. cerevisiae*, la présence d'un codon stop précoce (PTC pour « *premature stop codon* ») suivi d'un long 3'UTR dans un ARN provoquerait le recrutement de Upf1, Upf2 et Upf3 (He and Jacobson, 2001; Muhlrاد et al., 1994; Muhlrاد and Parker, 1999). Un deuxième complexe de protéines composé de Nmd4 (homologue chez la levure de Smg6) et Ebs1 (homologue chez la levure de Smg5/7) et du complexe de décapping viendrait ensuite s'associer à Upf1 permettant la dégradation de l'ARN par Xrn1 sans qu'il n'y ait déadénylation préalable (Dehecq et al., 2018) (**figure 8**).

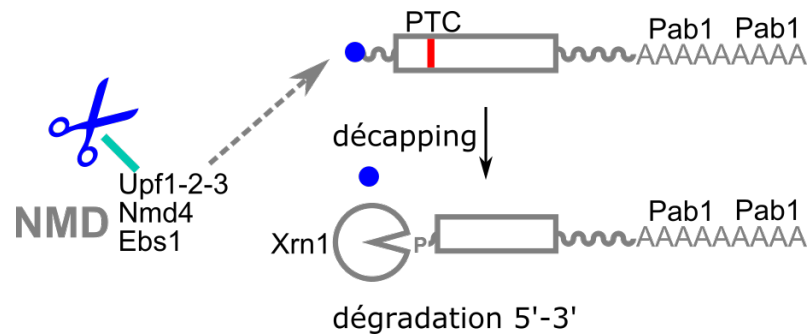
Chez les mammifères, la reconnaissance de PTC se fait par deux voies potentiellement distinctes : une similaire à celle de la levure, par l'intermédiaire d'un long 3'UTR, et une autre à laquelle participent un complexe macromoléculaire assemblé en aval des jonctions entre exons lors de l'épissage des ARN : l'EJC pour « *exon junction complex* ». Lorsqu'un codon stop est situé au sein d'une séquence codante et qu'il y a en aval un EJC, la reconnaissance du codon stop aberrant est médiée par le recrutement des facteurs du NMD (UPF2, UPF3) *via* l'EJC, ce qui permet leurs interactions avec UPF1 pour activer la dégradation de l'ARN. Contrairement à la levure où les ARN sont dégradés essentiellement par Xrn1, chez les mammifères ceux-ci seraient dégradés dans le sens 5'-3' par XRN1 et dans le sens 3'-5' par les déadénylases et l'exosome cytoplasmique. Il y aurait également un clivage endonucléolytique par la protéine SMG6 (homologue humain de Nmd4), recrutée par UPF1, revue dans (Lykke-Andersen and Jensen, 2015).

Plusieurs hypothèses ont été formulées pour expliquer comment la cellule était capable de distinguer un codon stop aberrant (PTC pour « *premature termination codon* ») d'un codon stop « normal » en fin de phase codante (NTC pour « *normal termination codon* »).

En comparaison d'un NTC, un ribosome resterait plus longtemps en attente des facteurs de terminaison de la traduction au niveau d'un PTC. La terminaison de la traduction serait donc plus lente et/ou moins efficace sur un PTC que sur un NTC, ce « retard » permettrait la reconnaissance de l'ARN cible et sa dégradation par la machinerie du NMD et l'hydrolyse d'ATP par Upf1 aiderait à prévenir l'accumulation de ribosomes au niveau des PTC (Serdar et al., 2016).

Lors de la terminaison de la traduction, Pab1 participerait au recrutement des facteurs de terminaison eRF1/eRF3 grâce à l'interaction entre Pab1 et eRF3 (Hoshino et al., 1999; Cosson et al., 2002; Roque et al., 2015; Ivanov et al., 2016). Chez les mammifères, il a été montré qu'il existe une compétition entre PABPC1 (homologue de Pab1) et Upf1 pour le recrutement des facteurs de terminaisons de la traduction eRF1/eRF3 (Ivanov et al., 2008; Singh et al., 2008). Cette compétition jouerait un rôle dans la dégradation des ARN, la liaison de Upf1 aux facteurs de terminaison sur le ribosome terminal déclencherait la dégradation de l'ARN tandis que la liaison de PABPC1 à ces facteurs de terminaison permettrait de préserver l'intégrité de l'ARN. En accord avec cette hypothèse, l'ancrage physique de Pab1 ou de eRF3 à proximité d'un PTC peut inhiber la dégradation d'ARN par le NMD (Amrani et al., 2004; Behm-Ansmant et al., 2007; Ivanov et al., 2008; Silva et al., 2008). Certaines études suggèrent que Upf1 serait recrutée sur l'ensemble des ARN, qu'ils soient substrats du NMD ou non, et serait déplacé par les ribosomes lors de la traduction vers les régions 3' non traduites où il s'accumulerait (Hogg and Goff, 2010; Zünd et al., 2013). L'enrichissement de Upf1 sur les longs 3'UTR augmenterait la probabilité que Upf1 plutôt que Pab1 interagisse avec les facteurs de terminaison de la traduction déclenchant alors la dégradation des ARN correspondants (Amrani et al., 2004; Singh et al., 2008; Eberle et al., 2008; Hogg and Goff, 2010). Les ARN cibles du NMD dépourvus de queue poly-A et donc potentiellement dépourvus de Pab1 peuvent toujours être stabilisés en absence de Upf1 (Meaux et al., 2008). Ces

observations suggèrent que le recrutement de Pab1 n'est pas nécessaire pour le mécanisme du NMD mais que cette dernière peut contribuer à la terminaison normale de la traduction. Une étude récente a montré que plus un PTC était proche de l'extrémité 3' de l'ARN, plus la terminaison de la traduction était efficace et moins il y avait de « *readthrough* » de ce codon stop. Cette étude a également montré que Pab1 jouerait un rôle important dans le contrôle de l'efficacité de terminaison (Wu et al., 2020).



**Figure 8. Dégradation de l'ARN par le nonsense mediated mRNA decay**

La dégradation par le NMD chez *S. cerevisiae* est déclenchée par la reconnaissance du PTC par le complexe composé de Upf1, Upf2 et Upf3. Un deuxième complexe composé de Upf1, Nmd4 et Ebs1 déclenche le décapping de l'ARN via Dcp1/Dcp2 ce qui initie la dégradation 5'-3' par l'exoribonucléase Xrn1.

### III.5.3. La dégradation par les voies NSD et NGD

En plus du NMD, il existe deux autres voies de dégradation des ARN qui présentent des défauts de traduction. La voie NSD (« *non-stop decay* ») cible les ARN qui n'ont pas de codon de terminaison. La voie NGD (« *no-go decay* ») dégrade les ARN dont la traduction est interrompue à cause de ribosomes bloqués (substrats « *no-go* »).

Dans le cas de la voie NSD, en l'absence de codon de terminaison, l'élongation se poursuit jusqu'au niveau des queues poly-A de l'ARN. Ski7 reconnaît le ribosome bloqué qui a son site accepteur vide, ce qui déclenche le recrutement de l'exosome cytoplasmique et l'ARN pourrait être déadénylé par l'exoribonucléase Dis3. En absence de Ski7, ces ARN pourraient être dégradés dans le sens 5'-3' par Xrn1. Les substrats NSD proviennent par exemple d'ARN dont la transcription s'est terminée à des sites de terminaison alternatifs, revue dans (Łabno et al., 2016).

La voie NGD intervient quand un ribosome est bloqué sur un ARN en cours de traduction. Ce blocage peut être dû à la présence de structures ou d'une suite de codons rares, le ribosome ne pouvant aller au-delà. L'ARN est alors clivé par une endonucléase, Cue2 (D'Orazio et al., 2019). Le fragment 3' ainsi généré est dégradé par Xrn1 et le fragment 5' par l'exosome cytoplasmique, revue dans (Łabno et al., 2016).

Dans ces deux voies, il semblerait que des protéines de recyclage du ribosome, Dom34 et Hbs1 qui ont des similarités de structure avec les facteurs de terminaison de la traduction eRF1 et eRF3, permettent la dissociation du ribosome et son recyclage, revue dans (D'Orazio and Green, 2021).



# Introduction aux résultats

Lorsque les ARN sont dans le cytoplasme, ils sont liés par la protéine de liaison des queues poly-A, Pab1, qui est impliquée dans de nombreux processus : export des ARN, traduction et déadénylation des ARN. La dégradation des ARN messagers débuterait par le raccourcissement de la queue poly-A et donc l'élimination progressive de Pab1. Une fois que les queues poly-A sont courtes (environ 10 résidus) et dépourvues de Pab1, elles seraient liées par le complexe Lsm1-7. Ce deuxième complexe jouerait un rôle dans la dégradation des ARN car il participerait au déclenchement du décapping. Revue dans (Passmore and Collier, 2022).

Étant donné que ces deux complexes semblent lier les ARN de manière séquentielle mais ont des fonctions *a priori* très différentes, nous avons souhaité étudier les ARN liés par ces deux protéines avec l'idée que nous aurions une population d'ARN stables associée à Pab1 et une population d'ARN en cours de dégradation associée au complexe Lsm1-7.

Pour cela, nous avons combiné une analyse des protéines associées à Pab1 et à Lsm1-7 d'une part, avec des immuno-précipitations des ARN par ces protéines et d'autres protéines impliquées dans la dégradation des ARN d'autre part. Dans une deuxième partie, des corrélations inattendues entre la demi-vie des ARN et la longueur de la queue poly-A nous ont amené à reconsidérer le rôle de la déadénylation dans la dégradation des ARN.



# Résultats

## I. Upf1 s'associe à des protéines qui lient les ARN avec des queues poly-A de différentes tailles

### I.1. L'étude des complexes associés à la dégradation des ARN par spectrométrie de masse révèle de nouvelles connexions

En combinant des purifications par affinité d'une protéine d'intérêt et l'identification par LC-MS/MS ( « *Liquid Chromatography coupled to tandem Mass Spectrometry* ») des protéines partenaires, nous avons analysé la composition de différents complexes protéiques impliqués dans le métabolisme des ARN. Pour cela des souches de levure contenant une protéine avec à son extrémité C-terminale une étiquette TAP ( « *Tandem Affinity Purification* ») ont été utilisées. Les extraits protéiques ont ensuite été purifiés par affinité sur des billes magnétiques couplées à des IgG (immunoglobulines G) afin de récupérer spécifiquement la protéine de fusion possédant l'étiquette TAP et les protéines associées (Rigaut et al., 1999). Les purifications sont réalisées avec ou sans une étape d'incubation avec de la ribonucléase (RNase) pour vérifier si les interactions observées sont dépendantes ou non de l'ARN.

Ce type d'approche a été réalisé avec différentes protéines impliquées dans la dégradation des ARN (Dhh1, Pat1, Lsm1 et Lsm7) ainsi qu'avec Pab1, la protéine cytoplasmique de liaison aux queues poly-A et a permis de mettre en évidence de nouvelles interactions détaillées ci-dessous.

#### I.1.1. Interactions de la protéine Pab1

Outre les interactions avec des facteurs impliqués dans la traduction (Tif4632/eIF4G) (Tarun and Sachs, 1996), la déadénylation (Pan3) (Siddiqui et al., 2007) ou l'export des ARN (Sto1/Cbp80) déjà connues, revue dans (Stewart, 2019), l'analyse des protéines co-purifiées avec Pab1 a permis de mettre en évidence une interaction avec la protéine Upf1 impliquée dans le NMD (**figure 9.A** rectangle rose). Cette interaction entre Upf1 et Pab1 a été confirmée par western blot, réalisé sur des extraits protéiques de souche de levure exprimant à la fois une version de Upf1 avec une étiquette TAP (Upf1-TAP) et une version de Pab1 possédant une étiquette HA (pour « *Hemagglutinin* »), Pab1-HA (**figure 9.B**) Un western blot avec des anticorps anti-HA sur des extraits de protéines totales ou purifiées sur des billes d'IgG a montré un enrichissement de Pab1-HA dans les extraits purifiés par rapport aux extraits totaux. Cet enrichissement n'est pas retrouvé si les extraits sont préalablement incubés avec de la RNase suggérant que l'interaction entre les deux protéines est médiée par l'ARN.

### I.1.2. Le complexe Lsm1-7/Pat1 interagit avec les facteurs du NMD

Lsm1 et Lsm7 sont associées dans le complexe Lsm1-7 qui lie de courtes queues poly-A. Les purifications de Lsm1-TAP et Lsm7-TAP ont montré qu'elles partageaient les mêmes interactions protéiques (**figure 9.A**) : le complexe Lsm1-7 est associé au complexe de décapping Dcp1/Dcp2 et aux facteurs qui lui sont associés dont Dhh1, Scd6, Pby1 et l'exoribonucléase Xrn1. Le profil d'interaction de Pat1 qui fait partie du complexe Lsm1-7/Pat1, est très similaire à ceux de Lsm1 et Lsm7 et confirme les interactions précédemment décrites de cette protéine avec Dcp2 et Xrn1, revue dans (Charenton and Graille, 2018),(Lobel et al., 2019). Seule Lsm7 qui fait également partie du complexe nucléaire Lsm2-8 participant à l'épissage des ARN, a montré des interactions avec les protéines regroupées dans la catégorie « épissage des ARN », confirmant la spécificité de la méthode utilisée.

Les purifications des différents membres du complexe Lsm1-7/Pat1 ont également mis en évidence une interaction de ces protéines avec Upf1 (**figure 9.A**) confirmant une observation précédente du laboratoire. Lors d'une analyse des complexes protéiques associés à la protéine Upf1, *Dehecq et al.* avaient mis en évidence une interaction entre Upf1 et le complexe Lsm1-7/Pat1, dépendante de l'ARN (Dehecq et al., 2018). Cette observation est étonnante étant donné que la dégradation des ARN dépendante de Upf1 par le NMD ne nécessite pas de déadénylation préalable, alors que le complexe Lsm1-7/Pat1 est connu pour se lier à des ARN partiellement déadénylés, possédant de courtes queues poly-A pour déclencher le décapping

### I.1.3. Le facteur de décapping Dhh1 révèle les complexes associés à la dégradation des ARN

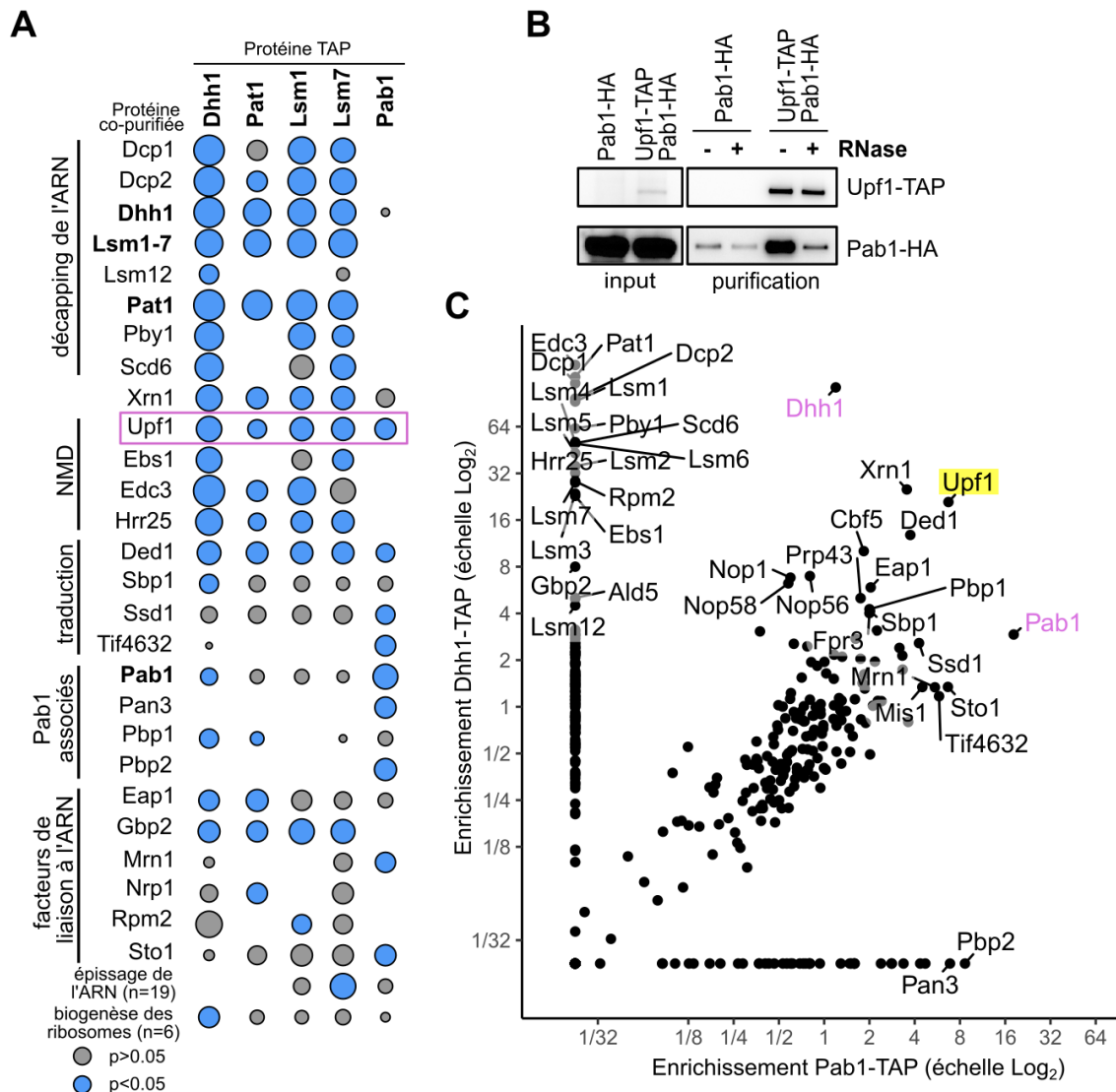
Comme attendu, les protéines enrichies après purification de Dhh1, qui est associée au décapping des ARN, étaient essentiellement des protéines impliquées dans la dégradation des ARN : le complexe de décapping (Dcp1/Dcp2) et des facteurs associés (Pby1 et Scd6), le complexe Lsm1-7/Pat1 ou encore l'exoribonucléase Xrn1 (**figure 9.A**). Des protéines impliquées dans le contrôle de la traduction (Sbp1 et Ded1) ont été également associées avec Dhh1 en accord avec le rôle potentiel de Dhh1 dans la répression traductionnelle.

Un enrichissement de Upf1 a été également observé parmi les protéines co-purifiées avec Dhh1 (**figure 9.A**). Bien qu'une interaction entre Upf1 et Dhh1 ait été déjà été observée dans certaines analyses à grande échelle (Gavin et al., 2002; Tarassov et al., 2008), aucune interaction entre ces deux protéines n'a pu être mise en évidence dans un crible double hybride (Swisher and Parker, 2011).

Dhh1-TAP montre une association avec Pab1 alors que Dhh1 n'est pas significativement enrichie lorsque l'on purifie Pab1 ce qui pourrait s'expliquer par l'existence au sein de la cellule, d'une population d'ARN décappés (potentiellement associées à Dhh1) avant d'être déadénylés (et donc encore associées à Pab1), comme c'est *a priori* le cas pour les cibles du NMD. Il n'est cependant pas exclu que d'autres ARN non cibles du NMD soient dégradés de la même façon.

A cet égard il est intéressant de noter que Upf1, acteur majeur du NMD, est associé aussi bien avec des protéines impliquées dans le décapping et la dégradation des ARN tels que Dhh1 et le complexe Lsm1-7/Pat1 qu'avec la protéine Pab1. De plus, si l'on compare les protéines présentes lorsque l'on purifie Dhh1-TAP avec celles présentes lorsque l'on purifie Pab1-TAP, Upf1 apparaît enrichie sensiblement de la même manière dans les deux purifications (**figure 9.C**) et est l'une des rares protéines à être enrichie de façon significative dans chacune des purifications réalisées ici (**figure 9.A et 9.C**).

Bien que les interactions entre Upf1 et Pab1 d'une part et entre Upf1 avec Dhh1 et Lsm1-7/Pat1 d'autre part soient probablement médiées par l'ARN (voir **figure 9.B** pour Pab1 et (Dehecq et al., 2018) pour Lsm1-7/Pat1), ces observations suggèrent l'existence de complexes indépendants liant l'ARN de manière séquentielle : un complexe associé à Pab1, des facteurs d'export et de la traduction des ARN et le complexe de décapping associé au complexe Lsm1-7/Pat1, avec Upf1 comme protéine charnière entre ces deux complexes. Afin de vérifier cette hypothèse, nous avons décidé d'analyser les ARN associés à ces différentes protéines.



**Figure 9. Analyse des protéines associées aux complexes de dégradation de l'ARN et à Pab1**

**(A)** Protéines enrichies (seuil fixé à 2) en association avec Dhh1, Pat1, Lsm1, Lsm7 et Pab1. L'enrichissement est représenté par la taille du cercle (échelle logarithmique), la couleur bleue indiquant une p-value du t-test inférieure à 0,05. Par souci de clarté, les facteurs d'épissage de l'ARN et les protéines impliquées dans la biogenèse des ribosomes trouvés dans les différentes purifications ont été regroupés. L'enrichissement moyen et les p-values ont été calculés pour chaque groupe dont la liste complète des protéines se trouve dans le matériel et méthodes. **(B)** Validation par western blot des interactions entre Upf1-TAP et Pab1-HA. Comme contrôle négatif, une souche exprimant uniquement Pab1-HA a été utilisée. Le traitement RNase (nucléase micrococcale) est indiqué. **(C)** Dhh1-TAP et Pab1-TAP sont toutes deux associées au facteur du NMD, Upf1. L'enrichissement a été calculé sur au moins trois répétitions en utilisant Dhh1-TAP ou Pab1-TAP comme appât et seules les protéines significativement enrichies ayant un facteur d'enrichissement supérieur à 4 soit avec Dhh1-TAP soit avec Pab1-TAP (p-value du t-test inférieure à 0,05) sont représentées (les protéines utilisées comme appât sont indiquées en rose).

## I. 2. Etude des ARN associés à Upf1 et Lsm1-7 et Upf1 et Pab1

Afin de déterminer quels étaient les ARN communs entre Upf1 et le complexe Lsm1-7 d'une part et Upf1 et Pab1 d'autre part, j'ai réalisé des purifications TAP de ces protéines et identifié les ARN associés par séquençage Illumina (RIP-seq). Le même type d'approche a été réalisé avec une protéine de la grande sous-unité ribosomique, Rpl16a, afin de déterminer l'état de traduction des ARN (Halbeisen et al., 2009).

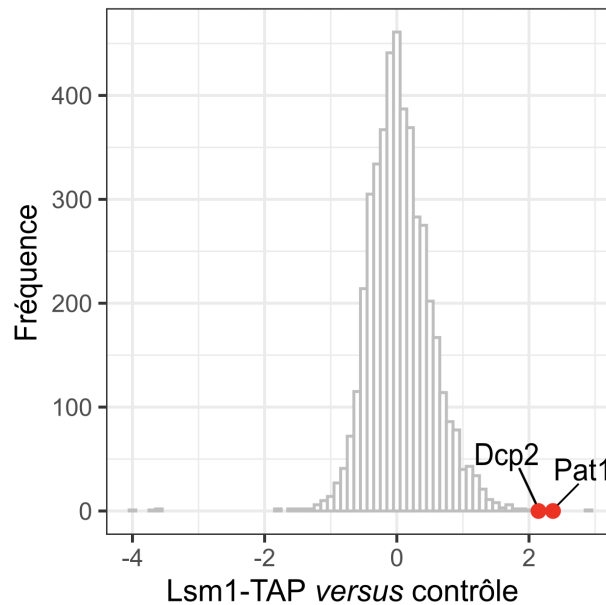
### I.2.1. Lsm1-TAP s'associe avec l'ARN de Dcp2

Une fois les fichiers de comptage obtenus pour les différentes purifications, j'ai réalisé une analyse avec le logiciel DESeq2 et une interface informatique, « Sartools » qui facilite la visualisation des données (Love et al., 2014; Varet et al., 2016). DESeq2 permet de déterminer quels sont les transcrits statistiquement différenciellement présents dans les différentes conditions expérimentales.

L'analyse a montré qu'il y avait beaucoup de gènes pour lesquels les ARN étaient différenciellement présents dans les purifications Rpl16a, Upf1 et Pab1 mais pas pour Lsm1, en comparaison d'un contrôle de la purification qui a été fait dans une souche qui n'exprime pas le TAP afin d'avoir le bruit de fond (**figure supplémentaire 1, S1**). Cela suggère que Lsm1 s'associe moins spécifiquement aux ARN ou que les conditions de purification ne sont pas optimales pour cette protéine.

Il est cependant intéressant de noter que les ARN de *DCP2* et *PAT1* sont retrouvés associés à la purification de Lsm1 (**figure 10**). Lsm1-7 étant connu pour se lier à de courtes queues poly-A, cela peut suggérer que les ARN de *DCP2* et *PAT1* à l'état d'équilibre sont majoritairement sous forme oligo-adénylée ce qui ne serait pas le cas pour l'ensemble des ARN.

Étant donné que le complexe Lsm1-7 interagit *via* Pat1 avec le complexe de décapping Dcp1/Dcp2, revue dans (Charenton and Graille, 2018), on peut aussi imaginer que cette interaction se fasse avec les protéines en cours de synthèse expliquant l'enrichissement particulier de ces ARN.



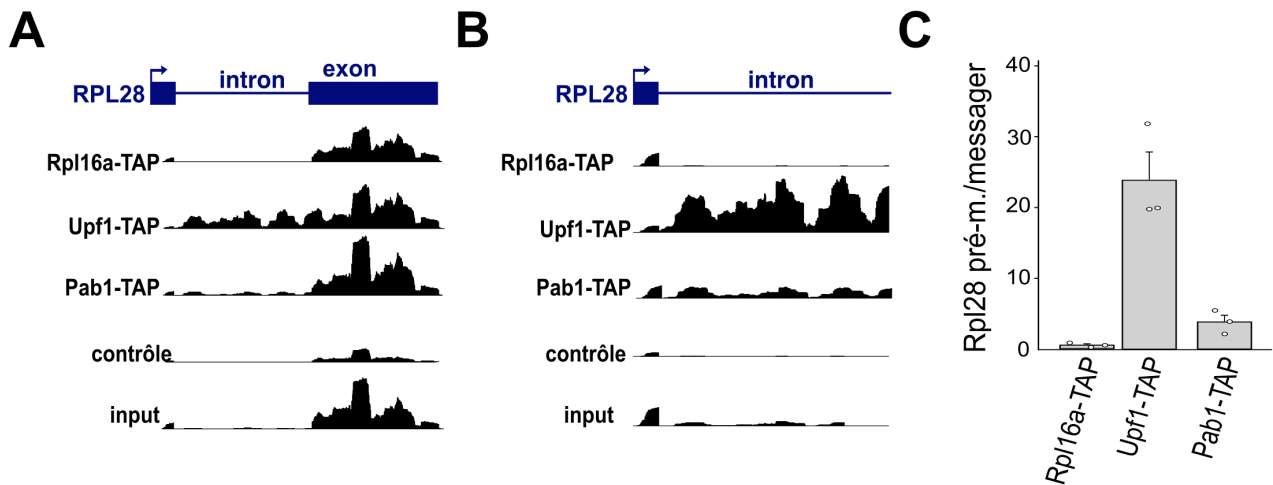
**Figure 10. Les ARN de Dcp2 et de Pat1 sont associés à Lsm1-TAP**

La fréquence représente le nombre de transcrits enrichis par Lsm1-TAP par rapport au contrôle qui a été fait dans une souche qui n'exprime pas le TAP afin d'avoir le bruit de fond.

### 1.2.2. Pab1 s'associe avec des ARN instables

À l'inverse de ce qui a été observé avec la purification de Lsm1-TAP en **figure 10**, de nombreux ARN ont été enrichis par rapport au contrôle lorsque Pab1, Upf1 ou Rpl16a ont été purifiés (**figure S1**). J'ai donc concentré les analyses sur la comparaison des ARN enrichis après purification de Pab1, Upf1 et Rpl16a.

Une première inspection visuelle des données de séquençage a été réalisée en utilisant le logiciel IGV (Robinson et al., 2011). Les **figures 11.A** et **11.B** montrent les résultats obtenus pour le gène *RPL28* dont la transcription génère un transcrit normalement épissé, insensible au NMD et un transcrit non épissé qui, une fois exporté au cytoplasme est une cible du NMD. Comme attendu, seul le transcrit mature épissé a été enrichi dans la purification de Rpl16a-TAP alors que dans la purification de Upf1-TAP les deux types de transcrits, messenger et pré-messenger sont présents (**figure 11.A** et **11.B**). L'ARN pré-messenger de *RPL28* a été également enrichi dans la purification de Pab1 (**figure 11.A** et **11.B**). Ce résultat a été validé par PCR quantitative sur les ARN associés à chacune des protéines purifiées (**figure 11.C**).



**Figure 11. Rpl28 pré-messager, substrat du NMD, est associé à Pab1 et à Upf1**  
**(A-B)** Exemples de profils de séquençage IGV dans les purifications de Rpl16a-TAP, Upf1-TAP et Pab1-TAP. Les ARN totaux ont été analysés en parallèle. Un contrôle de la purification a été réalisé dans une souche qui n'exprime pas le TAP, ceci représente le bruit de fond. **(C)** Validation par qPCR de l'association spécifique de l'ARN pré-messager de RPL28 à Upf1-TAP et Pab1-TAP mais pas à Rpl16a-TAP.

Dans un deuxième temps, nous avons analysé l'enrichissement des transcrits individuels dans chacune des purifications par rapport à leur abondance dans la cellule, en comparant les valeurs obtenues après purification à celles obtenues par séquençage d'une population d'ARN total à l'équilibre (**figure 12.A-C**).

Outre les ARN messagers codant pour les protéines, l'ARN polymérase II synthétise de nombreux ARN non codant. Ces ARN sont en général instables car ils sont rapidement dégradés dans le noyau par l'exosome nucléaire (CUTs) ou dans le cytoplasme par le NMD (XUTs/SUTs). Les protéines étudiées étant essentiellement cytoplasmiques, les ARN non codant susceptibles d'être enrichis sont *a priori* des XUTs et des SUTs, cibles du NMD, terme utilisé pour désigner cette catégorie d'ARN instables dans laquelle j'ai également inclus les ARN pré-messagers, eux aussi dégradés par le NMD.

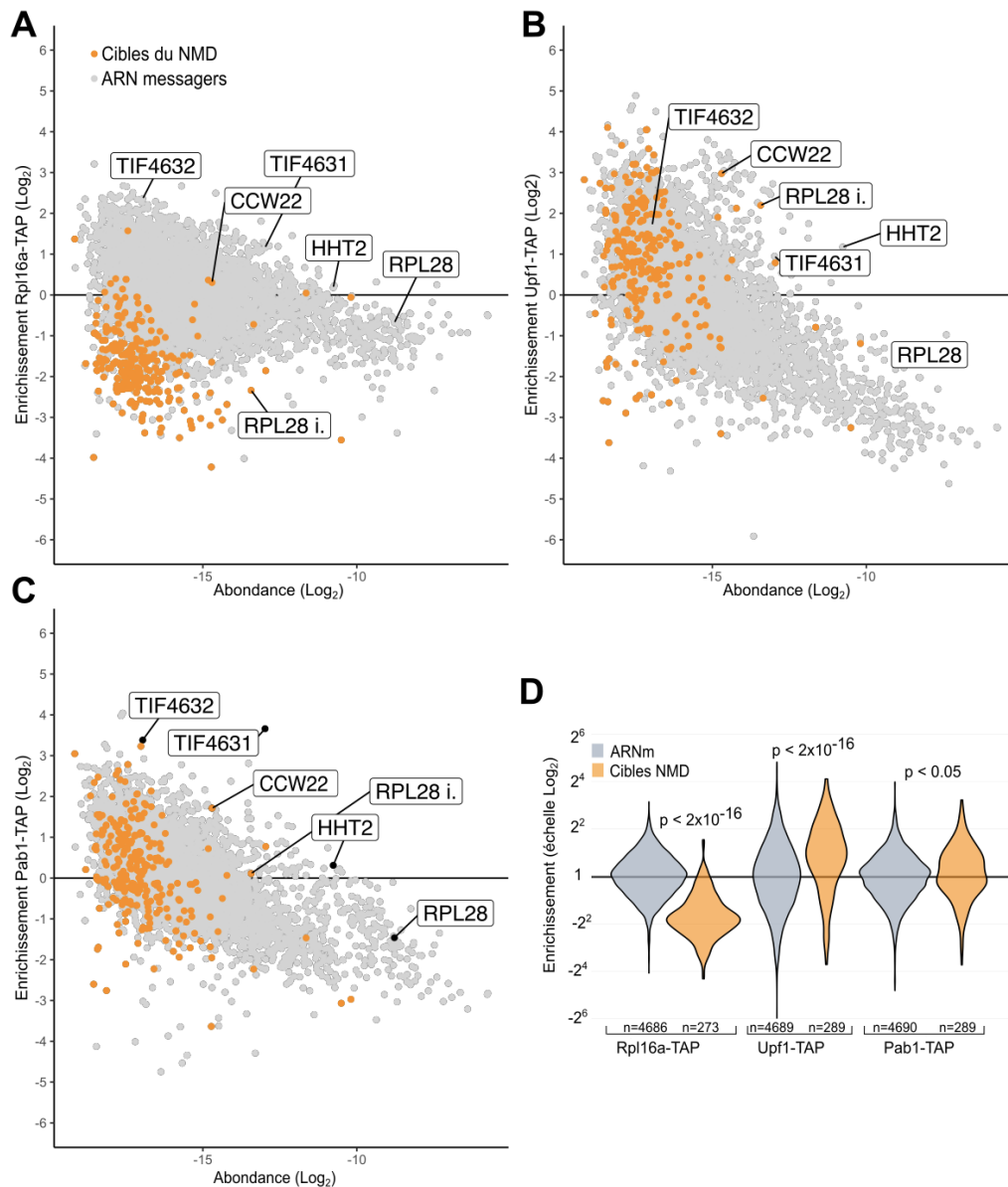
Comme attendu, les ARN enrichis après purification de Rpl16a-TAP, protéine associée à des ARN activement traduits, ont été essentiellement des ARN messagers (**figure 12.A** en gris). A l'inverse, les ARN cibles du NMD ayant des capacités codantes restreintes ont été peu associées avec Rpl16a-TAP (**figure 12.A** en orange). Ces transcrits non codant ont été en revanche enrichis lorsque l'on a immuno-précipité Upf1 (**figure 12.B** en orange). L'association des ARN messagers avec Upf1 (**figure 12.B** en gris) est plus étonnante puisque *a priori* ces ARN ne sont pas dégradés par le NMD. Une partie des ARN messagers associée à Upf1 pourrait cependant être constituée d'isoformes sensibles au NMD, issues d'évènements d'initiation de la transcription non canoniques, que notre analyse ne nous permet pas de distinguer des isoformes majeures non ciblées par le NMD (Malabat et al., 2015). Il est également possible qu'à l'instar de ce qui a été observé chez les mammifères, Upf1 s'associe de façon non spécifique et indépendamment de la

traduction à tous les ARN messagers dont il serait « délogé » lors de la traduction (Hogg and Goff, 2010; Zünd et al., 2013). Une analyse statistique comparant la distribution des différents types de transcrits dans chacune des purifications de Rpl16a (**figure 12.D** à gauche) et Upf1 (**figure 12.D** au centre) a confirmé que les ARN messagers étaient plus enrichis que les ARN cibles du NMD dans la purification de Rpl16a alors que c'était l'inverse avec Upf1. Lorsque le même type d'analyse a été réalisée pour les ARN associés à Pab1, on a observé parmi les ARN enrichis, la présence aussi bien d'ARN messagers que d'ARN instables, cibles du NMD (**figure 12.C**). L'analyse statistique indique que ces dernières ont été un peu plus enrichies avec Pab1 que ne le sont les ARN messagers (**figure 12.D** à droite).

L'association de Pab1 aux ARN messagers n'est pas surprenante car ceux-ci étant *a priori* polyadénylés, on s'attend à les enrichir lorsque l'on purifie la protéine de liaison aux queues poly-A. L'enrichissement des ARN instables, cibles du NMD parmi les ARN associés à Pab1, même s'il peut paraître surprenant, peut s'expliquer par le fait que ces ARN étant dégradés sans déadénylation préalable, ils ont de longues queues poly-A qui peuvent fixer plusieurs molécules de Pab1, facilitant leur purification.

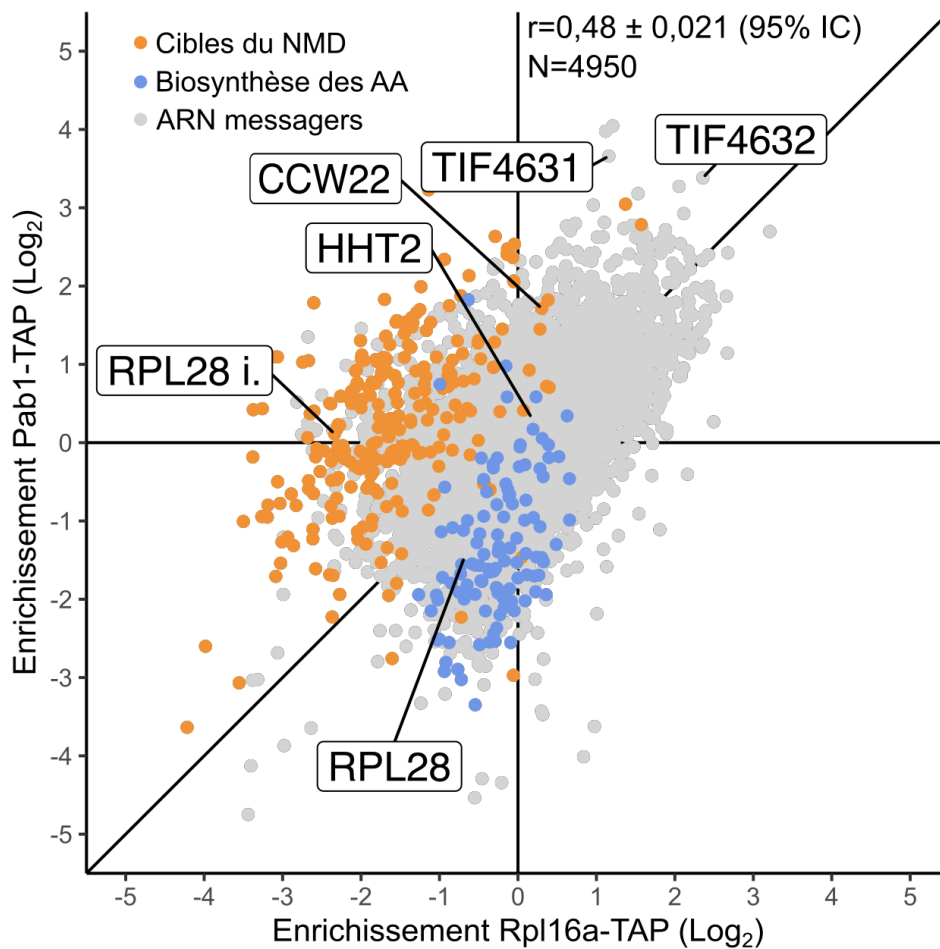
Une comparaison des ARN enrichis dans les purifications de Rpl16a et Pab1 a montré une relativement bonne corrélation entre les deux ( $r = 0,48$ ) suggérant que les ARN en cours de traduction sont plus enrichis par Pab1 que les ARN peu traduits comme cela a déjà été montré (**figure 13** et (Costello et al., 2015)). Ainsi parmi les 150 ARN messagers les moins enrichis par Pab1, 112 correspondent à des transcrits de gènes impliqués dans la biosynthèse des acides aminés (en bleu sur la **figure 13**), qui sont peu traduits dans les conditions de culture utilisées, comme le montre leur faible enrichissement avec Rpl16a-TAP. Cet enrichissement des ARN activement traduits pourrait être médié par l'interaction de Pab1 avec eIF4G (Tif4631/Tif4632) qui est impliqué dans l'initiation de la traduction et parce que le complexe tertiaire Pab1-eIF4G-ARN serait plus stable que le complexe Pab1-ARN. Il est d'ailleurs intéressant de noter que les ARN codant pour les deux protéines Tif4631 et Tif4632 sont parmi les plus enrichis dans la purification de Pab1 (**figure 12.C**). Le domaine d'interaction de eIF4G avec Pab1 étant localisé dans le premier tiers de la protéine (acides aminés 201 à 315 sur 914 pour Tif4632), il est possible que Pab1 interagisse avec eIF4G de façon co-translationnelle (Tarun and Sachs, 1996).





### Figure 12. Pab1-TAP est associée à des ARN instables

**(A-C)** Représentation des ARN associés à Rpl16a-TAP, Upf1-TAP et Pab1-TAP par rapport à l'abondance des transcrits dans l'extrait total. L'enrichissement est calculé en faisant le ratio de chaque purification sur le signal obtenu dans l'extrait total (« input »). Les ARN cibles du NMD sont en orange alors que les ARN messagers sont en gris. Certains transcrits sont annotés, RPL28 i. désigne l'ARN pré-messager de RPL28. **(D)** Distributions des valeurs d'enrichissement des ARN messagers (gris) et des cibles du NMD (orange) dans les purifications de Rpl16a-TAP (gauche), Upf1-TAP (centre) et Pab1-TAP (droite) par rapport à un extrait d'ARN total à l'état d'équilibre. Le nombre de transcrits pour chaque classe est spécifié. La p-value (p) a été déterminée par le test statistique de Wilcoxon-Mann-Whitney pour la catégorie des ARN messagers montrant des valeurs supérieures à la catégorie des cibles du NMD pour Rpl16a-TAP et inférieures à la catégorie des cibles du NMD pour Upf1-TAP et Pab1-TAP.



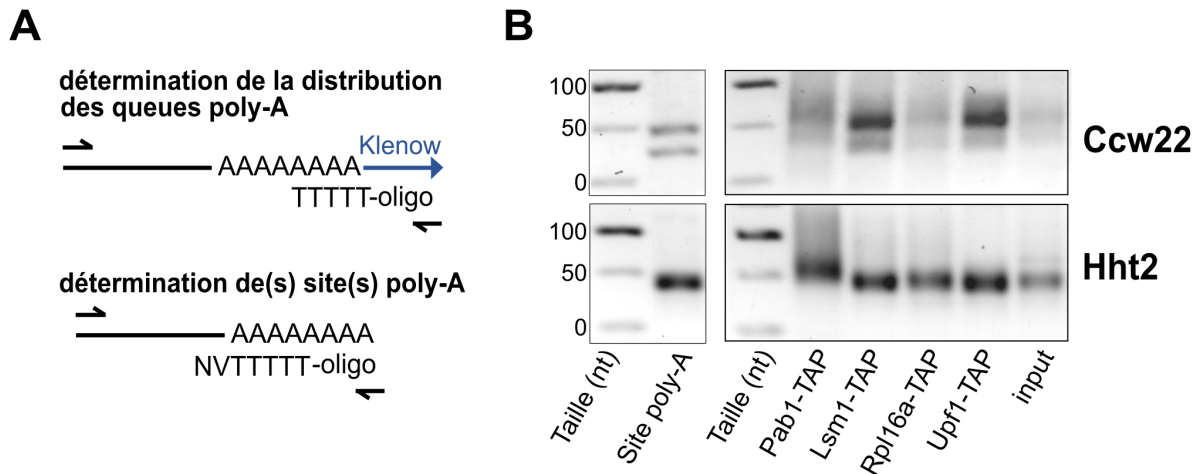
**Figure 13. Comparaison des ARN purifiés par Rpl16a et Pab1**

Représentation des ARN associés à Rpl16a-TAP en comparaison de ceux associés à Pab1-TAP. Les ARN cibles du NMD sont en orange, les transcrits appartenant à la biosynthèse des acides aminés (« AA ») sont en bleu, les ARN messagers sont en gris. Certains transcrits sont annotés, RPL28 i. désigne l'ARN pré-messager de RPL28. Un test statistique Pearson a été réalisé pour obtenir le coefficient de corrélation  $r$  de 0,48 avec une  $p$ -value inférieure à  $2 \times 10^{-16}$ . L'intervalle de confiance (IC) de 95 % pour le coefficient de corrélation est indiqué. Le nombre de transcrits ( $N$ ) est indiqué.

Étant donné que Pab1 joue un rôle complexe dans le contrôle de la taille de la queue poly-A des ARN en permettant le recrutement des déadénylases et qu'elle interagit *a priori* avec tout ARN polyadénylé ayant une queue poly-A d'au moins 12 résidus (Sachs et al., 1987), nous avons décidé d'analyser la longueur des queues poly-A des ARN enrichis avec Pab1 en utilisant la technique de l'ePAT pour « *extension Poly(A) Test* » (Jänicke et al., 2012) qui permet d'observer la distribution des queues poly-A et d'en mesurer la taille (**figure 14.A**). Cette méthode permet également de déterminer l'existence de sites de polyadénylation alternatifs.

Les résultats des expériences de ePAT réalisées pour Ccw22, un substrat du NMD associé à Upf1 et à Pab1 et pour Hht2 sont présentés sur la **figure 14.B**. Dans les deux

cas, les ARN immuno-précipités avec Pab1 possèdent des queues poly-A plus longues que celles observées dans un extrait d'ARN total ou dans les ARN immuno-précipités avec Rpl16a, Upf1 ou Lsm1 (**figure 14.B**). La même observation a été faite sur tous les ARN testés (**figure 14.B** et données non présentes) ce qui suggère que Pab1-TAP enrichit spécifiquement les ARN ayant de longues queues poly-A et pas ou peu les formes ayant de courtes queues poly-A. Par ailleurs, dans l'extrait total (« *input* »), les queues poly-A observées sont courtes et il n'y a pas les formes longues observées avec la purification Pab1-TAP ce qui suggère que les formes polyadénylées possédant de longues queues poly-A sont des formes minoritaires.



**Figure 14. Validation de l'association de Pab1 à des substrats du NMD qui ont de longues queues poly-A**

**(A)** La technique de l'ePAT de (Jänicke et al., 2012) permet d'observer la distribution des queues poly-A pour un ARN donné. Une amorce oligo-dT est hybridée à l'ARN. L'incubation en présence de Klenow (fragment de l'ADN polymérase I de *E. coli*) et de désoxyribonucléotides permet à cette dernière de recopier la séquence ADN du duplex ARN/ADN ajoutant ainsi à l'extrémité 3'OH une séquence connue qui sera utilisée, après transcription inverse de l'ARN, pour réaliser une amplification par PCR. L'utilisation, en parallèle d'une amorce oligo-dT ancree ((dT)<sub>12</sub>VN où V représente n'importe quel nucléotide sauf A et N n'importe quel nucléotide) permet la détermination des sites de polyadénylation et le calcul de la taille des queues poly-A. **(B)** Visualisation de la distribution des queues poly-A dans les purifications Pab1, Lsm1, Rpl16a et Upf1 ainsi que pour un extrait d'ARN total :input. Le « site poly-A » correspond au site de terminaison de l'ARN.

## II. La déadénylation n'est pas nécessaire à la dégradation des ARN

### II.1. La taille de la queue poly-A est inversement corrélée à la demi-vie des ARN

Pour la majorité des ARN messagers, le modèle de dégradation couramment admis indique que celle-ci commence par un raccourcissement de la queue poly-A sous l'action séquentielle des deux complexes de déadénylation, PAN et CNOT. Puis, lorsque la taille de la queue poly-A atteint environ 10 résidus, elle est liée par le complexe Lsm1-7 permettant ainsi le recrutement du complexe de décapping Dcp1/Dcp2 et la dégradation par l'exoribonucléase Xrn1, revue dans (Passmore and Collier, 2022).

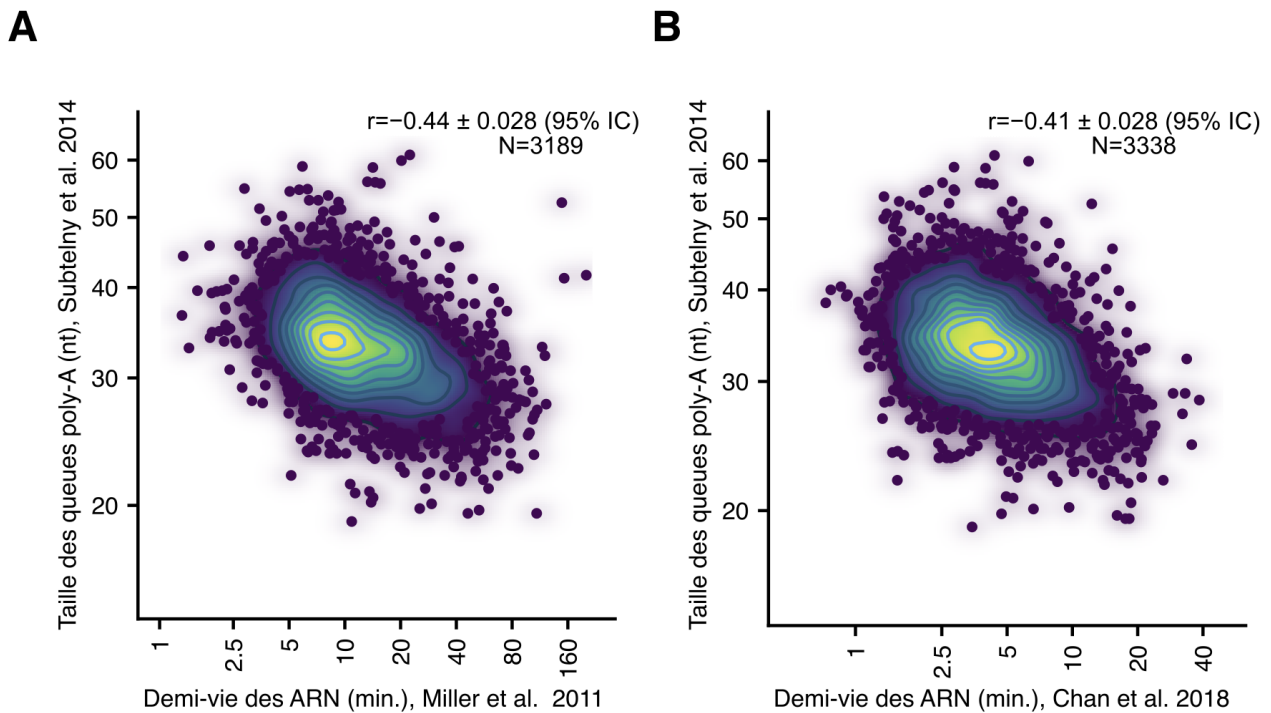
D'après ce modèle, où la déadénylation des ARN est un prérequis pour la dégradation et constitue l'étape limitante de ce processus, la stabilité des ARN dépend essentiellement de la vitesse de déadénylation. Un ARN instable étant déadénylé plus rapidement qu'un ARN stable, la longueur de la queue poly-A d'un ARN à l'équilibre serait un indicateur de la demi-vie de cet ARN.

Cependant, ce n'est pas aussi simple, puisque lorsque que l'on effectue des corrélations entre la demi-vie des ARN et la longueur de leur queue poly-A, on obtient l'image inverse de ce qui est attendu : les ARN ayant de longues queues poly-A ont des demi-vies courtes alors que les ARN ayant de courtes queues poly-A ont des demi-vies longues (**figure 15**). Bien que cette corrélation ne tienne pas compte de la dynamique de la déadénylation puisqu'il s'agit ici de la distribution moyenne des queues poly-A des ARN à l'état d'équilibre, le résultat obtenu est difficilement compatible avec un modèle dans lequel la stabilité d'un ARN dépend essentiellement de la vitesse de raccourcissement de sa queue poly-A.

En admettant que la déadénylation soit un prérequis pour la dégradation, la corrélation observée sur la **figure 15** est plus compatible avec un modèle dans lequel les ARN instables seraient très rapidement déadénylés et dégradés (après déadénylation) de sorte qu'à l'état d'équilibre seules les formes ayant des queues poly-A longues sont détectées tandis que les ARN stables seraient dégradés plus lentement après déadénylation. En d'autres termes, après déadénylation, le déclenchement du décapping et la dégradation par Xrn1 ne seraient pas identiques pour les ARN stables et les ARN instables et constituerait l'étape limitante au moins pour les ARN stables.

On pourrait cependant aussi avancer une autre hypothèse pour expliquer le fait que les ARN instables aient de longues queues poly-A : que ceux-ci soient dégradés rapidement sans qu'il n'y ait de déadénylation préalable, comme c'est le cas pour les ARN dégradés par la voie NMD. Aussi, il ne serait pas possible d'observer des queues poly-A de tailles intermédiaires ou courtes pour ces ARN. Pour les ARN stables, en revanche, les formes majoritaires observées auraient de courtes queues poly-A car ils auraient le temps

d'être déadénylés avant d'être dégradés. D'après ce modèle alternatif, la déadénylation ne serait pas nécessairement un prérequis pour la dégradation et la vitesse de raccourcissement de la queue poly-A ne serait pas critique pour la stabilité des ARN.



**Figure 15. La demi-vie des ARN n'est pas corrélée à la longueur de la queue poly-A à l'état d'équilibre**

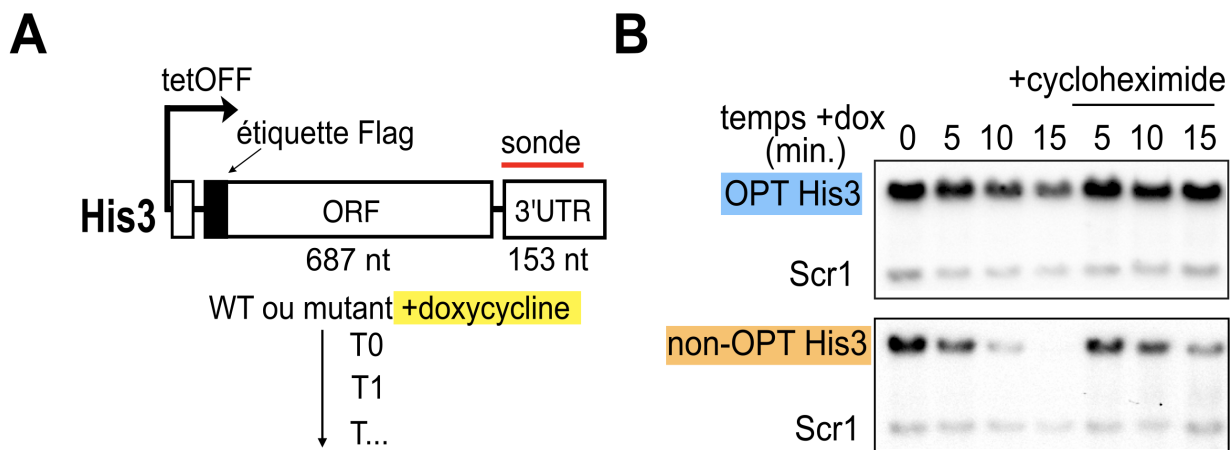
Les mesures de queue poly-A ont été moyennées et proviennent de (Subtelny et al., 2014). Les mesures de demi-vies proviennent de (Miller et al., 2011) **(A)** ou (Chan et al., 2018) **(B)**. La densité des points est indiquée par la couleur sur le graphique, de jaune à violet, jaune correspondant au maximum de densité et violet au minimum. Un test statistique Pearson a été réalisé pour obtenir le coefficient de corrélation  $r$  avec une  $p$ -value inférieure à  $2 \times 10^{-16}$ . L'intervalle de confiance (IC) de 95 % pour le coefficient de corrélation est indiqué. Le nombre de transcrits ( $N$ ) est indiqué.

## II.2. Le raccourcissement de la queue poly-A n'est pas nécessaire au déclenchement de la dégradation des ARN

### II.2.1. Construction de gènes rapporteurs pour analyser le rôle de la déadénylation dans la dégradation des ARN

Afin de pouvoir tester l'importance de la déadénylation dans la dégradation des ARN, nous voulions disposer d'ARN rapporteurs de stabilités différentes dont nous pourrions analyser la dégradation dans différents fonds génétiques ou conditions de culture. Pour déstabiliser un ARN, il est possible d'introduire un codon stop au sein d'une séquence codante. Cependant, l'ARN étant alors reconnu par Upf1, Upf2 et Upf3 et dégradé par la voie NMD indépendamment du raccourcissement de la queue poly-A, il n'était pas possible d'utiliser cette stratégie pour générer un ARN instable et tester l'importance de la déadénylation dans sa dégradation indépendamment du NMD. Aussi, afin de générer des rapporteurs exprimant un ARN stable et un ARN instable permettant d'analyser l'effet d'une inhibition de la déadénylation sur leur dégradation, j'ai utilisé le fait que le biais de codons d'un ARN a un effet majeur sur la stabilité des ARN (Presnyak et al., 2015). En utilisant les données publiées par le laboratoire de *Jeff Collier* sur l'optimisation des codons (Presnyak et al., 2015; Radhakrishnan et al., 2016), j'ai construit des plasmides permettant la synthèse d'un ARN codant pour la protéine His3 mais avec des biais de codons différents. L'un de ces plasmides permet la transcription d'un ARN dont la séquence a été déterminée comme la plus optimale pour la traduction (nommé « OPT His3 », qui correspond à la version optimisée 100% de l'article) tandis que l'autre a une séquence considérée comme non optimale pour la traduction (« non-OPT His3 », qui correspond à la version optimisée à 10% de l'article). Ces deux ARN sont sous contrôle d'un promoteur dont la transcription peut être réprimée par l'ajout de doxycycline (**figure 16.A**). Pour déterminer la demi-vie des ARN et valider des différences de stabilité en fonction de la composition en codons, des cinétiques de dégradation des ARN ont été réalisées (**figure 16.B**). Alors que l'ARN non optimal disparaît après 10 minutes, l'ARN optimal persiste plus de 15 minutes, validant l'importance de la composition des codons d'un ARN pour sa stabilité. Par ailleurs, l'ajout de cycloheximide, un inhibiteur de la traduction, stabilise les deux rapporteurs, ce qui confirme le fait que la dégradation de ces ARN est affectée par la traduction (Hu et al., 2009).





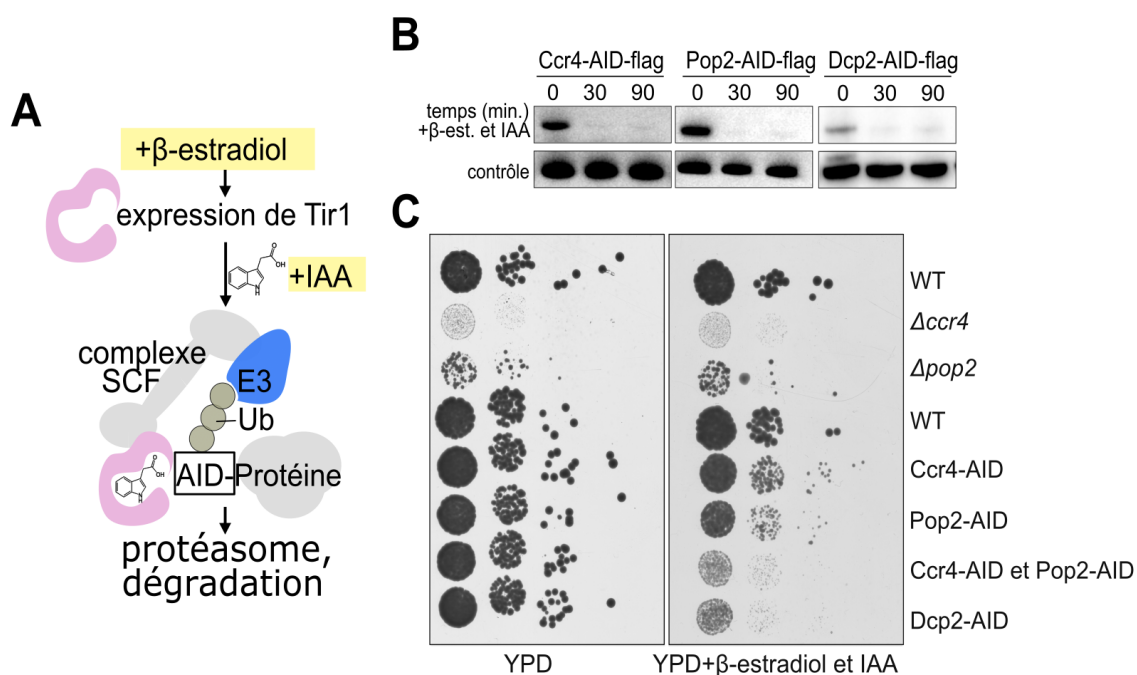
**Figure 16. L'optimisation des codons affecte la demi-vie d'ARN rapporteurs qui sont stabilisés en absence de la traduction**

**(A)** Schéma des rapporteurs sous contrôle d'un promoteur dont la transcription peut être réprimée par l'ajout de doxycycline dans le milieu de culture. La sonde pour le northern blot est indiquée en rouge. **(B)** Northern blot réalisé avec les ARN totaux extraits de cellules traitées ou non avec de la doxycycline (concentration finale de 10  $\mu\text{g/mL}$ ) pour inhiber la transcription et de la cycloheximide (concentration finale de 5  $\mu\text{g/mL}$ ) pour inhiber la traduction. Les temps de traitement avec les drogues avant l'extraction des ARN sont indiqués. Les cellules ont été transformées avec les plasmides permettant l'expression de la version optimisée ou non de l'ARN codant pour la protéine His3. Les ARN ont été séparés, après dénaturation à 65°C dans du formazole, sur un gel d'agarose à 1,5%. Après transfert des ARN sur membrane de nylon, celles-ci ont été hybridées simultanément avec des sondes marquées à la digoxigénine correspondant soit à la région 3' UTR des ARN rapporteurs soit à l'ARN non codant Scr1, qui sert de contrôle de charge.

## II.2.2. La déplétion des déadénylases

L'influence de la délétion des déadénylases sur la stabilité des ARN a déjà été testée chez *S. cerevisiae* (Tucker et al., 2001). Cependant, même si la délétion des sous-unités catalytiques des déadénylases est viable, elle affecte fortement la croissance cellulaire, revue dans (Collart, 2016). Aussi pour réduire les potentiels effets secondaires liés à la délétion de ces gènes, nous avons décidé de générer de nouvelles souches dans lesquelles les déadénylases peuvent être rapidement et temporairement déplétées.

Nous avons pour cela utilisé le système de dégradation inducible par l'auxine (AID pour « *auxin-inducible degron* ») qui permet de déclencher la dégradation rapide d'une protéine cible en présence d'auxine (Mendoza-Ochoa et al., 2019) (**figure 17.A**). Les résultats présentés sur la **figure 17.B** montrent que 30 minutes après induction du système de dégradation, par ajout des drogues, les protéines cibles étiquetées avec le domaine AID étaient quasiment indétectables. Par ailleurs, les tests de croissance ont montré qu'à l'inverse des souches délétées pour Ccr4 et Pop2, les souches exprimant les protéines de fusion Ccr4-AID et Pop2-AID ne présentaient de phénotype de croissance que lorsqu'elles étaient cultivées en présence de drogues (**figure 17.C**). L'effet de la déplétion des protéines sur la croissance concorde avec le fait que Ccr4 et Pop2 seraient importantes pour le bon fonctionnement cellulaire.



### Figure 17. La déplétion des déadénylases par le système AID

(A) En présence d'auxine (IAA : acide indole acétique), Tir1 interagit avec les protéines conservées Skp1 et Cdc53 pour former un complexe E3 ubiquitine ligase SCF-Tir1 permettant la poly-ubiquitinylation de la protéine de fusion avec le domaine AID et sa dégradation par le protéasome. Pour éviter un effet potentiellement toxique dû à une éventuelle dégradation incontrôlée de la protéine cible même en absence d'auxine, nous avons utilisé un système analogue à celui décrit par Mendoza-Ochoa et al., dans lequel l'expression de Tir1 est induite par ajout de  $\beta$ -estradiol dans le milieu (Mendoza-Ochoa et al., 2019) (B) Western blot, la quantité de protéine-fusion avec le domaine AID (Ccr4, Pop2, Dcp2) diminue après ajout de  $\beta$ -estradiol et d'auxine. Les temps correspondant à l'ajout des drogues sont indiqués. (C) Test de croissance de souches de type sauvage (WT pour « wild type ») ou mutantes après ajout ou non de  $\beta$ -estradiol et d'auxine. Des dilutions sérielles des cultures ont été déposées sur boîte et incubées pendant 48h à 30°C.



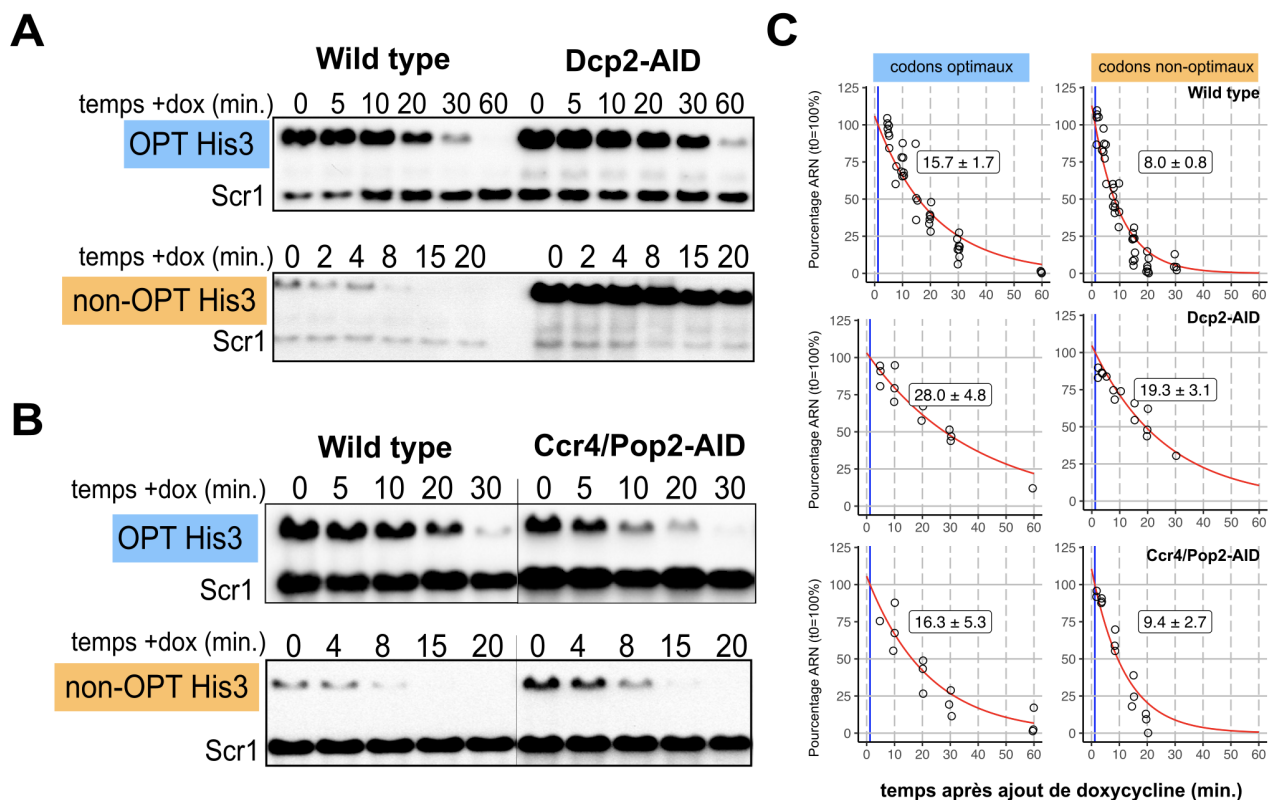
## II.3. La déadénylation n'est pas requise pour la dégradation d'ARN rapporteurs

Après avoir construit les rapporteurs permettant la synthèse d'ARN stables et instables (**figure 16**) et les souches dans lesquelles une protéine d'intérêt peut être rapidement déplétée (**figure 17**), nous avons réalisé des cinétiques de dégradation des ARN afin de tester l'effet de ces déplétions sur la stabilité des ARN. Pour cela, les différentes souches transformées avec les plasmides rapporteurs ont été traitées pendant une heure avec du  $\beta$ -estradiol et de l'auxine (IAA) pour dépléter la ou les protéines cibles, avant l'ajout de doxycycline, pour arrêter la transcription de l'ARN.

Dans une cellule de type sauvage (« *wild type* »), l'ARN ayant une séquence optimisée a une demi-vie (mesurée par PCR quantitative) d'environ 16 minutes, près de deux fois plus longue que l'ARN dont la séquence n'est pas optimale (**figure 18.A et C**). Après déplétion de Dcp2 ou de l'exoribonucléase Xrn1, les deux types d'ARN sont stabilisés (**figures 18.A et S2**), et en absence de Dcp2, leur demi-vie est augmentée d'un facteur 1,6 à 2 pour la version optimale et la version non-optimale respectivement (**figure 18.C**). En revanche, la délétion du gène *SKI2*, impliqué dans le recrutement de l'exosome cytoplasmique et la dégradation des ARN de 3' vers 5', n'a aucun effet sur la stabilité des ARN testés sauf quand le décapping est inhibé en même temps (**figure S3**).

Contrairement, à ce que nous avons observé en absence de Dcp2, la déplétion des déadénylases Ccr4, Pop2 et Pan2 seules ou en combinaison n'a pas entraîné la stabilisation de ces ARN rapporteurs (**figures 18.B et S4.A**). Aucun effet significatif sur la demi-vie des rapporteurs n'a pu être mesuré par PCR quantitative en absence de Ccr4, Pop2 ou de Ccr4 et Pop2 par rapport à une cellule de type sauvage (**figures 18.C et S4.B**).

L'ensemble de ces résultats indique que les deux rapporteurs ARN sont majoritairement dégradés dans le sens 5'-3' par l'exoribonucléase Xrn1, après décapping mais qu'ils n'auraient pas besoin d'être, au préalable, déadénylés contrairement au modèle canonique de dégradation dans lequel le décapping est initié par le recrutement du complexe Dcp1/Dcp2 sur de courtes queues poly-A *via* le complexe Lsm1-7, revue dans (Passmore and Collier, 2022).



**Figure 18.** La demi-vie des versions optimale et non-optimale de His3 n'est pas impactée par la déplétion de Ccr4-AID et Pop2-AID mais est augmentée quand le décappage est compromis

**(A-B)** Northern blot réalisés avec des ARN totaux extraits de souches wild type, Dcp2-AID ou Ccr4-AID/Pop2-AID traitées avec de l'auxine et du  $\beta$ -estradiol pendant 1h et pendant différentes périodes de temps avec de la doxycycline indiqué au dessus des northern blot. Les cellules ont été transformées avec les plasmides permettant l'expression de la version optimisée ou non de l'ARN de His3. Après dénaturation à 65°C dans du formazol, les ARN ont été séparés sur un gel d'agarose à 1,5% et transférés sur membrane de nylon. Les membranes ont été hybridées simultanément avec des sondes marquées à la digoxigénine correspondant soit à la région 3' UTR des ARN rapporteurs soit à l'ARN non codant Scr1, qui sert de contrôle de charge. **(C)** qPCR et mesures de demi-vies réalisées sur au moins 3 expériences indépendantes pour chaque souche et chaque plasmide. La demi-vie a été déterminée par une régression non linéaire et un « lag » de 1,2 minutes a été appliqué afin de corriger le délai de l'arrêt transcriptionnel. L'intervalle de confiance (95%) pour le temps de demi-vie de l'ARN basé sur une régression non-linéaire utilisant tous les points de minimum 3 répétitions de l'expérience est indiqué.

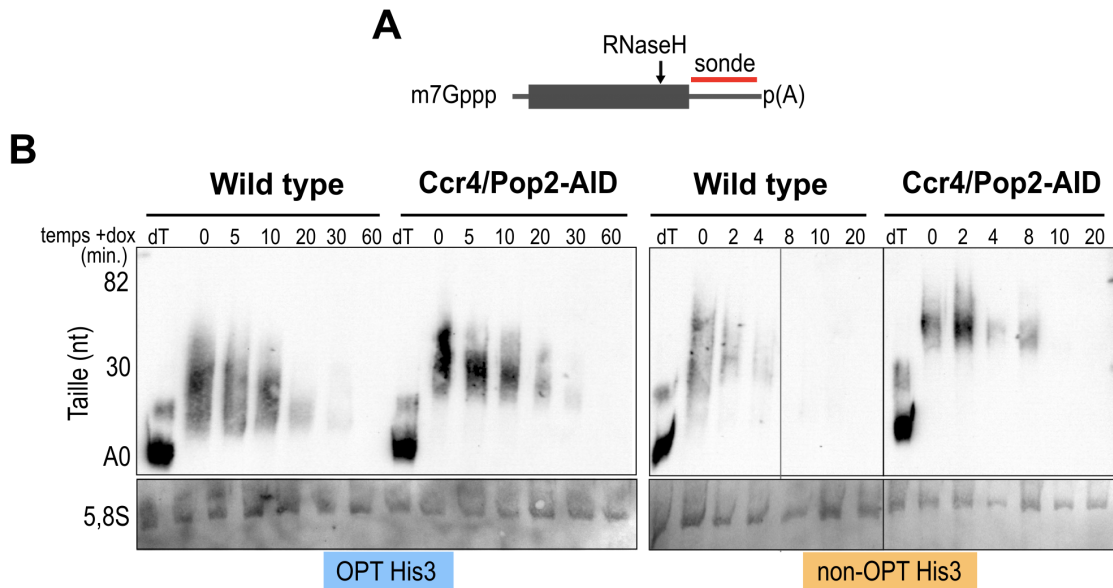
Étant donné que la déplétion des déadénylases ne montrait pas d'effet sur la stabilité des ARN, nous avons souhaité vérifier si l'absence des déadénylases avait au moins un effet sur la taille des queues poly-A. Pour cela, nous avons réalisé des digestions à la RnaseH (qui clive des duplex ARN/ADN) des ARN extraits à différents temps après

arrêt de la transcription afin de pouvoir déterminer, par northern blot, l'évolution de la taille de la queue poly-A au cours du temps (**figure 19.A**).

Les résultats de ces expériences ont montré que les queues poly-A qui, à l'état d'équilibre, avant arrêt de la transcription (temps = 0), avaient une distribution large de 0 à 55 nucléotides environ, ont évolué progressivement vers des formes courtes avec une distribution moins étendue, indiquant que les formes les plus longues sont progressivement raccourcies (**figure 19.B**). Bien que pour la version non-optimisée du rapporteur, la dégradation rapide de l'ARN n'ait pas permis d'observer un raccourcissement aussi important que celui observé pour la version optimale, nous avons pu quand même observer un raccourcissement au cours du temps (**figure 19.B**, non-OPT His3 panneau de droite *versus* OPT His3 panneau de gauche).

En absence de Ccr4 et Pop2, pour l'ARN ayant une séquence optimale, nous avons toujours observé une diminution de la taille de la queue poly-A au cours du temps, bien que celle-ci ait été moins rapide que dans une souche de type sauvage et que l'on n'ait jamais observé de formes aussi courtes (**figure 19.B**, panneau de gauche et **figure S5**). Pour l'ARN ayant une séquence non-optimale en revanche, il semble que la déplétion de Ccr4 et Pop2 inhibe le raccourcissement de la queue poly-A ou en tout cas, ait un fort effet de ralentissement de la vitesse de déadénylation (**figure 19.B**, panneau de droite et **figure S5**).

Ces résultats indiquent que Ccr4 et Pop2 ont un rôle dans la déadénylation des ARN et qu'en leur absence, les ARN ne peuvent pas être totalement déadénylés comme cela a précédemment été décrit (Tucker et al., 2001). L'inhibition du complexe CNOT *via* la déplétion de Ccr4 et Pop2 inhibant ou empêchant la déadénylation totale des ARN sans avoir d'effet sur leur demi-vie, indique que contrairement à ce qui est admis d'après le modèle canonique, la déadénylation pourrait ne pas avoir un rôle majeur dans le cas de la dégradation de ces ARN.



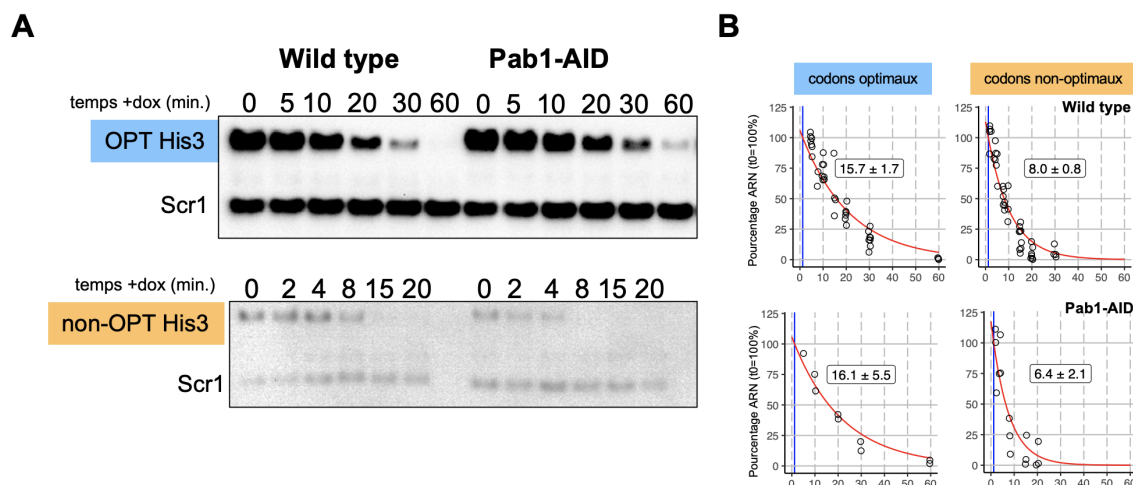
**Figure 19. La déplétion de Pop2 et Ccr4 empêche la déadénylation totale des ARN rapporteurs**

**(A)** Schéma indiquant la position du site de clivage des ARN par la RNaseH en présence de l'oligonucléotide LA121. Le rectangle gris représente l'ORF et le trait gris en aval le 3'UTR, p(A) symbolise la queue poly-A. La sonde pour le northern blot est indiquée en rouge. **(B)** Northern blot réalisés avec des ARN totaux clivés à la RNaseH avec l'oligonucléotide LA121 provenant de souches wild type, Ccr4-AID/Pop2-AID traitées avec de l'auxine et du  $\beta$ -estradiol pendant 1h et pendant différentes périodes de temps avec de la doxycycline pour arrêter la transcription du gène rapporteur. Les temps sont indiqués au dessus des northern blot. Les cellules ont été transformées avec les plasmides permettant l'expression de la version optimisée ou non de l'ARN de His3. Les ARN clivés ont été séparés, après dénaturation à 65°C dans du formazol, sur un gel d'acrylamide (6%) dénaturant. Après transfert des ARN sur membrane de nylon, celles-ci ont été hybridées avec une sonde marquée à la digoxigénine correspondant à la région 3' UTR des ARN rapporteurs. Une réaction de digestion à la RNaseH avec l'oligonucléotide LA121 et un oligo-dT (dT) permet de générer un fragment d'ARN sans queue poly-A (A0), utilisé pour estimer la taille des queues poly-A indiquées sur la gauche. L'ARN 5,8S visualisé après coloration au bleu de méthylène de la membrane est utilisé comme contrôle de charge.

## II.4. L'absence de Pab1 n'affecte pas la demi-vie des ARN mais empêche leur déadénylation

Pab1 a de multiples rôles dans le métabolisme cellulaire, elle est impliquée à la fois dans le contrôle de la taille de la queue poly-A, la traduction et la déadénylation des ARN, et leur stabilité. De récentes expériences *in vitro* ont suggéré que Pab1 stimulait la déadénylation médiée par le complexe CNOT (Webster et al., 2018; Yi et al., 2018). Aussi nous avons voulu voir *in vivo* si la déplétion de Pab1 empêchait la déadénylation des ARN et si cela affectait leur stabilité.

De la même manière que précédemment, nous avons réalisé des cinétiques de dégradation des ARN dans une condition wild type et dans une condition où nous avons déplété Pab1. Aucune différence de stabilité des ARN testés n'a été observée en absence de Pab1 par rapport à une souche wild type (**figure 20.A**) et les demi-vies pour les deux rapporteurs n'étaient pas sensiblement différentes (**figure 20.B**). Cela suggère que l'absence de Pab1 n'empêche pas la dégradation des ARN. Cependant, même si la déplétion de Pab1 n'a pas montré d'effet sur la stabilité des ARN, les queues poly-A observées étaient plus longues (**figure S6**) en accord avec ce qui a précédemment été montré (Sachs and Davis, 1989).



### Figure 20. La déplétion de Pab1 ne déstabilise pas les ARN rapporteurs

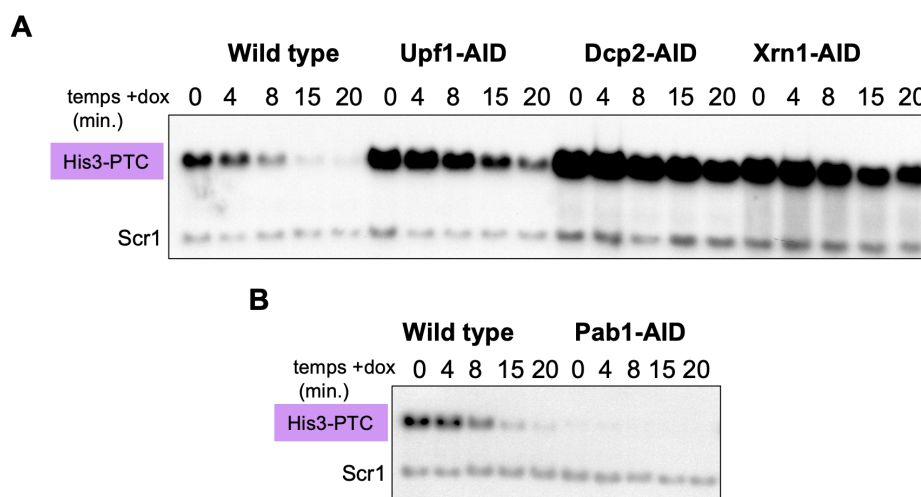
**(A)** Northern blot réalisés avec des ARN totaux extraits de souches wild type ou Pab1-AID traitées avec de l'auxine et du  $\beta$ -estradiol pendant 1h et pendant différentes périodes de temps avec de la doxycycline pour arrêter la transcription du gène rapporteur. Les temps sont indiqués au dessus des northern blot. Les cellules ont été transformées avec les plasmides permettant l'expression de la version optimisée ou non de l'ARN de His3. Après dénaturation à 65°C dans du formazol, les ARN ont été séparés sur un gel d'agarose à 1,5% et transférés sur membrane de nylon. Les membranes ont été hybridées simultanément avec des sondes marquées à la digoxigénine correspondant soit à la région 3' UTR des ARN rapporteurs soit à l'ARN non codant Scr1, qui sert de contrôle de charge. **(B)** qPCR et mesures de demi-vies réalisées sur au moins 3 expériences indépendantes pour chaque souche et chaque plasmide. La demi-vie a été déterminée par une régression non linéaire et un « lag » de 1,2 minutes a été appliqué afin de corriger le délai de l'arrêt transcriptionnel. L'intervalle de confiance (95%) pour le temps de demi-vie de l'ARN basé sur une régression non-linéaire utilisant tous les points de minimum 3 répétitions de l'expérience est indiqué.

Étant donné que parmi les ARN enrichis par immuno-précipitation de Pab1 nous avons identifié de nombreux ARN cibles du NMD (**figure 12.C**), nous avons voulu étudier l'effet de la déplétion de cette protéine sur la stabilité et la taille de la queue poly-A sur un ARN cible du NMD. Nous avons donc construit un nouveau plasmide rapporteur

permettant l'expression d'un ARN dégradé par le NMD en introduisant un codon stop prématuré à la place du 24<sup>ème</sup> codon de la séquence codant pour His3 dans le plasmide exprimant la version optimisée.

Cet ARN, dénommé His3-PTC, est rapidement dégradé dans une souche wild type et stabilisé lorsque Upf1, la protéine clé du NMD, Dcp2 ou Xrn1 ont été déplétées, confirmant qu'il est dégradé par le NMD dans la souche wild type (**figure 21.A**). De façon surprenante, lorsque Pab1 a été déplétée, cet ARN est devenu presque indétectable par northern blot (**figure 21.B**). Le fait que cet ARN cible du NMD s'accumule moins dans une souche déplétée pour Pab1 suggère qu'il est plus rapidement dégradé et que le NMD serait plus « efficace » en absence de Pab1 qu'en condition wild type.

Étant donné l'implication de Pab1 dans la terminaison de la traduction et le fait que Pab1 et Upf1 interagissent toutes les deux avec le facteur de terminaison eRF3 (Czaplinski et al., 1998; Hoshino et al., 1999; Cosson et al., 2002), on peut imaginer que lorsque Pab1 est présente, il y a une compétition avec Upf1 pour la terminaison au PTC et que Pab1 prévient de la dégradation par la voie du NMD. Quand elle est absente, il n'y a plus de compétition, Upf1 serait recruté plus efficacement au PTC par les facteurs de terminaison de la traduction et les ARN dégradés plus rapidement, comme cela a déjà été proposé (Amrani et al., 2004).



### Figure 21. Un rapporteur NMD est sensible à la déplétion de Pab1

Northern blot réalisés avec des ARN totaux extraits de souches wild type ou Upf1-AID, Dcp2-AID, Xrn1-AID (**A**) et Pab1-AID (**B**) traitées avec de l'auxine et du  $\beta$ -estradiol pendant 1h et pendant différentes périodes de temps avec de la doxycycline. Les temps sont indiqués au dessus des northern blot. Les cellules ont été transformées avec le plasmide His3-PTC. Après dénaturation à 65°C dans du formazol, les ARN ont été séparés sur un gel d'agarose à 1,5% et transférés sur membrane de nylon. Les membranes ont été hybridées simultanément avec des sondes marquées à la digoxigénine correspondant soit à la région 3' UTR des ARN rapporteurs soit à l'ARN non codant Scr1, qui sert de contrôle de charge.



# Conclusions et discussion

## I. Analyse des complexes protéiques et des ARN associés au complexe Lsm1-7

Notre analyse des complexes impliqués dans la dégradation des ARN ainsi que les données publiées par *Dehecq et al.*, ont confirmé l'interaction du complexe Lsm1-7/Pat1, avec le complexe de décapping Dcp1/Dcp2 et ses co-facteurs, dont Dhh1. Par ailleurs, une interaction de Upf1 avec le complexe Lsm1-7/Pat1 avait déjà été observée, et bien qu'elle dépende de l'ARN, il est surprenant de voir Upf1 associée au complexe Lsm1-7 qui lie de courtes queues poly-A, étant donné que Upf1 est impliquée dans la dégradation des ARN par le NMD qui est, *a priori*, indépendante du raccourcissement de la queue poly-A ((Dehecq et al., 2018) et **figure 9**). Les cibles du NMD ne sont *a priori* pas liées par le complexe Lsm1-7/Pat1 et les délétions de Lsm1, Lsm7 ou Pat1 n'affectent pas la stabilité d'un rapporteur NMD à l'inverse des ARN dégradés par la voie canonique, qui montrent une accumulation d'espèces d'ARN cappées et déadénylées (Boeck et al., 1998; Bouveret, 2000).

Aussi, nous avons souhaité identifier la population d'ARN liés à la fois par Upf1 et par le complexe Lsm1-7. Des analyses des ARN associés à Lsm1 avaient déjà été réalisées par la technique du CLIP (« *Cross-linking immunoprecipitation* »). Cette technique, qui permet, comme le RIP-seq d'identifier les ARN associés à une protéine d'intérêt, est caractérisée par une étape de crosslinking aux rayons ultraviolets, qui stabilise les interactions entre les acides nucléiques et les protéines par la création de ponts covalents et permet, d'une part de faire des purifications en conditions dénaturantes diminuant les interactions non spécifiques et d'autre part, d'identifier le site d'interaction de la protéine avec l'ARN au nucléotide près. Environ 1200 ARN messagers enrichis d'un facteur 2 dans la purification de Lsm1 ont été identifiés et une analyse des sites d'interactions ARN-protéine a montré que Lsm1 s'associait préférentiellement aux régions 3'UTR des ARN (Mitchell et al., 2013). Nos analyses n'ont montré qu'une centaine d'ARN enrichis dans la purification Lsm1 par rapport au contrôle (**figure S1**). Mais de façon intéressante, parmi les transcrits enrichis dans cette purification, on retrouve les ARN codant pour les protéines Dcp2 et Pat1 parmi les 3 transcrits les plus enrichis (**figure 10**). L'association spécifique de ces ARN peut être dû à une forte association du complexe Lsm1-7 aux formes déadénylés de ces ARN ou au fait que ces ARN seraient davantage sous forme déadénylée que la majorité des ARN. Cependant, ni l'ARN codant pour Dcp2, ni celui codant pour Pat1 ne figurent parmi les ARN messagers purifiés avec Lsm1-TAP dans les expériences de CLIP (Mitchell et al., 2013). Étant donné que les protéines Pat1 et Dcp2 interagissent avec Lsm1, on peut imaginer que les ARN correspondants soient enrichis dans nos expériences de RIP-seq parce que Lsm1 interagirait avec les protéines

en cours de synthèse. Cette association de Lsm1-7 avec les protéines Dcp2 et Pat1 lors de leur traduction permettrait la formation d'un complexe de décapping « fonctionnel » composé de Lsm1-7/Pat1/Dcp1-Dcp2. Pour tester cette dernière hypothèse, nous pourrions combiner des purifications de complexes protéiques avec un marquage métabolique des protéines néosynthétisées utilisant la méthode de « pSILAC » (« *pulse Stable Isotope Labeling by Amino acid in Cell culture* ») (Pratt et al., 2002). En analysant en parallèle des purifications de Lsm1 et d'autres partenaires de Dcp2 et Pat1, ce type de marquage métabolique pourrait permettre de déterminer si les protéines néosynthétisées Dcp2 et Pat1 sont associées à Lsm1 plutôt qu'avec une autre protéine avec laquelle elles interagissent également, par exemple Dcp1 pour Dcp2 (She et al., 2008) et Dcp2 pour Pat1 (Charenton et al., 2017; Lobel et al., 2019).

Ce type d'association co-translationnelle, déjà proposé pour l'association des protéines avec les ARN messagers de leurs partenaires protéiques par *Duncan et Mata*, a été par la suite identifié comme un mécanisme général dans la formation de complexes multi-protéiques (Duncan and Mata, 2011; Shiber et al., 2018).

Un assemblage co-translationnel peut avoir de nombreux avantages pour la formation d'un complexe multiprotéique. Cela peut par exemple permettre d'assurer un assemblage ordonné des sous-unités du complexe qui se ferait une fois que les sites de liaisons respectifs émergeraient du ribosome. De plus, les interactions qui en résultent pourraient avoir un effet sur le repliement de la chaîne polypeptidique naissante ou sur la stabilisation des domaines repliés, voire dans certains cas, d'assurer une protection contre la dégradation. Enfin, compte tenu de l'environnement encombré du compartiment cytosolique, la formation de complexes co-translationnels peut permettre de mieux contrôler l'association des « bonnes » protéines au « bon » moment.

## II. Le lien entre Upf1 et Pab1 autour du NMD

A l'inverse de l'interaction entre le complexe Lsm1-7 et Upf1, l'interaction entre Pab1 et Upf1 dépendante de l'ARN (**figure 9.A** et **9.B**), était plus attendue dans la mesure où Pab1 lie tous les ARN polyadénylés et que les cibles du NMD liées par Upf1 étant dégradées sans déadénylation préalable, sont forcément visibles sous forme polyadénylées. L'analyse des ARN associés à Pab1 a par ailleurs confirmé qu'en plus de lier les ARN messagers, elle liait les cibles du NMD (**figure 12.C** et **12.D**).

L'association d'ARN messagers à Upf1 (**figure 12.B** et **12.D**) est moins intuitive mais peut s'expliquer par le fait que les ARN qui présentent des sites alternatifs d'initiation de la transcription générant des isoformes sensibles au NMD n'ont pas été exclus de l'analyse et/ou que, comme certaines études le suggèrent, Upf1 lie tous les ARN de façon indiscriminée et que le NMD ne serait déclenché que dans le cas où il s'agit d'une cible du NMD (Hogg and Goff, 2010; Zünd et al., 2013; Malabat et al., 2015).

En l'occurrence, le mécanisme de reconnaissance d'un codon stop comme précoce et le déclenchement du NMD restent peu clairs. Il semble cependant que le recrutement



sur le ribosome terminal par les facteurs de terminaison eRF1/eRF3 de Upf1 ou de Pab1 ait un rôle important : le recrutement de Upf1 (et Upf2, Upf3) par les facteurs de terminaison déclencherait la dégradation de l'ARN tandis que le recrutement de Pab1 permettrait de préserver l'intégrité de l'ARN (Ivanov et al., 2008; Singh et al., 2008). Les résultats que nous avons obtenus sur ARN rapporteur cible du NMD en absence de Pab1 sont en accord avec l'existence d'une compétition entre Pab1 et Upf1 comme facteur déclenchant de la dégradation de ces ARN. Notre observation que l'ARN du gène rapporteur His3-PTC, cible du NMD, s'accumule moins dans la cellule lorsque Pab1 est déplétée qu'en condition wild type (**figure 21.B**) s'expliquerait par une meilleure « efficacité » du mécanisme de dégradation par la voie du NMD, conséquence d'un recrutement plus « efficace » de Upf1 par eRF3 sur le ribosome au PTC en absence de compétition avec Pab1. Cet ARN différant seulement par une base (générant un PTC) de la version optimale de His3, qui elle n'est pas affectée par la déplétion de Pab1 (**figure 20**), il est peu probable que l'effet observé sur son accumulation résulte d'un problème de transcription ou d'export de cet ARN en absence de Pab1.

Le même effet d'absence de Pab1 sur le NMD a été observé dans une souche *pab1Δ*, *spb8Δ* sur un rapporteur *PGK1* dans lequel des PTC ont été introduits à des distances variables du codon initiateur (Meaux et al., 2008). Il est intéressant de noter que dans ces expériences, la moindre accumulation du transcrit sensible au NMD en absence de Pab1 par rapport à une souche wild type est plus visible lorsque le PTC est relativement proche du codon initiateur mais pas s'il est proche de l'extrémité 3' de la séquence codante même si l'ARN correspondant est toujours substrat du NMD, suggérant que le mécanisme déclenchant le NMD est plus complexe qu'une simple compétition entre Pab1 et Upf1 pour l'interaction avec les facteurs de terminaison de la traduction (Kervestin et al., 2012).

### III. Un modèle actualisé pour la dégradation des ARN

Le modèle de dégradation des ARN couramment admis indique que celle-ci commence par un raccourcissement de la queue poly-A et que le décapping par Dcp1/Dcp2 et la dégradation par l'exoribonucléase Xrn1 sont déclenchés lorsque la taille de la queue poly-A atteint environ une dizaine de résidus, revue dans (Passmore and Collier, 2022). D'après ce modèle, la stabilité des ARN dépend essentiellement de leur vitesse de déadénylation.

Ce modèle est basé sur des expériences d'induction puis d'arrêt transcriptionnel qui ont montré sur des rapporteurs de stabilités différentes, une déadénylation accélérée pour un ARN instable par rapport à un ARN stable, et un déclenchement du décapping et de la dégradation par Xrn1 uniquement quand les queues poly-A étaient raccourcies (Muhlrad et al., 1995) ainsi que sur des simulations qui ont montré que c'est la modification de la

vitesse de déadénylation qui a le plus d'impact sur la demi-vie des ARN (Cao and Parker, 2001).

Cependant ce modèle est difficilement compatible avec notre analyse montrant qu'il existait une corrélation inverse entre la taille de la queue poly-A et la demi-vie des ARN à l'état d'équilibre (**figure 15**) ainsi qu'avec certaines études montrant que des ARN déadénylés s'accumulaient sans être dégradés (Herrick et al., 1990; Muhrad et al., 1995). Les courtes queues poly-A devraient, d'après un modèle majoritairement dépendant de la déadénylation, être transitoires et rarement observées. Leur accumulation suggère que la vitesse de dégradation d'un ARN dépend aussi du déclenchement du décapping sur ces formes oligo-adenylées. Une étude récente portant sur l'analyse et la dégradation de plusieurs milliers de gènes chez les mammifères a par ailleurs montré que la vitesse de dégradation de formes oligo-adenylées des ARN variait beaucoup d'un gène à un autre et était d'autant plus élevée que les ARN étaient déadénylés rapidement (Eisen et al., 2020).

Que la vitesse de dégradation des ARN dépende uniquement de la vitesse de déadénylation ou qu'elle dépende à la fois de la vitesse de déadénylation et de décapping, tous les modèles s'accordent sur le fait que la déadénylation est requise pour la dégradation des ARN. De plus, les délétions des gènes *CCR4* ou *POP2* codant pour les déadénylases du complexe CNOT chez la levure augmentent la demi-vie d'ARN rapporteurs (Tucker et al., 2001).

De façon surprenante, lorsque nous avons réalisé des cinétiques de dégradation sur nos rapporteurs, nous n'avons pas observé de stabilisation des rapporteurs optimisés ou non pour la traduction de His3 en absence de Ccr4 et Pop2 alors même que leurs queues poly-A n'étaient pas totalement raccourcies, indiquant que l'absence des déadénylases a un effet sur la taille des queues poly-A, sans avoir d'effet sur la stabilité des ARN (**figures 18, 19.B et S5**). Cependant, nous avons observé que ceux-ci continuaient d'être déadénylés même quand Ccr4 et Pop2 étaient déplétées (**figure 19.B et S5**) ce qui suggère que des quantités résiduelles d'enzymes sont suffisantes pour la déadénylation ou que d'autres enzymes peuvent raccourcir les queues poly-A longues de ces ARN. Un autre complexe de déadénylases, le complexe PAN composé de Pan2/Pan3 est connu pour être le premier complexe à déadényler les ARN et en particulier les formes possédant de longues queues poly-A (Brown and Sachs, 1998; Tucker et al., 2001; Schäfer et al., 2019). Aussi la déadénylation par le complexe PAN pourrait expliquer le léger raccourcissement des longues formes de queues poly-A observées sur la **figure 19.B**. Par ailleurs, quand les déadénylases étaient déplétées, la déadénylation des deux rapporteurs semblait s'arrêter à environ 15 résidus d'adénosines ce qui pourraient correspondre à la présence d'une molécule de Pab1 encore liée à la queue poly-A, en accord avec les observations récentes indiquant que seule Ccr4 est capable d'éliminer Pab1 liée à la queue poly-A (Webster et al., 2018; Yi et al., 2018). Il a été montré dans ces mêmes études que Pab1 promouvait la déadénylation des ARN *in vitro* et qu'en son absence les ARN étaient déadénylés moins efficacement. Lorsque nous avons regardé les queues poly-A des rapporteurs en absence de Pab1, nous avons effectivement observé que les queues poly-A étaient plus longues mais les ARN n'étaient pas stabilisés (**figures**

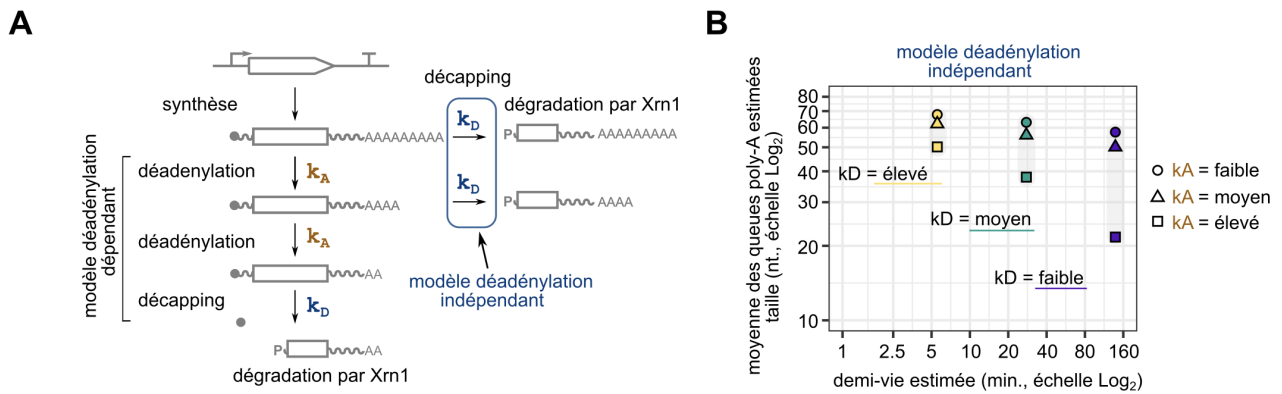
**20** et **S6**), contrairement à ce qui avait déjà été observé dans des mutants de délétion de différents domaines de Pab1 (Simón and Séraphin, 2007). Une différence majeure entre nos expériences et les données publiées est l'utilisation de mutants de délétion *versus* des déplétions temporaires et rapides (1h) des protéines par le système AID dans notre cas. Même s'il peut rester une petite quantité de la protéine non dégradée après 1h de traitement (**figures 17.B**), la déplétion semble suffisante car nous observons des effets directs sur les queues poly-A des ARN (**figures 19.B, S5 et S6**).

Les ARN rapporteurs utilisés ont été fortement stabilisés en absence de Dcp2 ou Xrn1 (**figures 18.A, 18.C et S2**) mais pas en absence des déadénylases Ccr4 et Pop2 (**figures 18.B et 18.C**). Ces résultats suggèrent que, contrairement au modèle canonique de dégradation, ces ARN peuvent être décappés et dégradés indépendamment d'un raccourcissement de la queue poly-A.

Bien qu'on ne puisse exclure que ce type de dégradation indépendante de la déadénylation ne se produise qu'en réponse à la déplétion de Ccr4 et Pop2, il est intéressant de noter que si on réalise des simulations de dégradation des ARN en considérant la déadénylation et le décapping comme deux phénomènes concomitants mais indépendants (**figure 22.A**), on peut, en jouant uniquement sur les constantes de déadénylation ( $k_A$ ) et de décapping ( $k_D$ ) (**figure 22.B**), trouver des paramètres qui expliquent la corrélation inverse entre taille des queues poly-A à l'équilibre et demi-vie des ARN de levure présentée sur la **figure 15**.

D'après cette simulation, dans laquelle la valeur de la constante de décapping ( $k_D$ ) seule détermine la demi-vie des ARN, on peut voir que pour un ARN instable ( $k_D$ =élevé), des variations de la vitesse de déadénylation ( $k_A$ ) n'affectent pas beaucoup la taille des queues poly-A observées, et leur distribution est restreinte (**figure 22.B**, jaune). En revanche, pour un ARN stable ( $k_D$ =faible) la demi-vie étant plus longue, il y aura plus de chances que l'on observe une distribution étendue des tailles de queues poly-A avec notamment, des queues poly-A courtes (**figure 22.B**, violet).

Comme souligné précédemment, ce modèle simple qui ne dépend que du décapping permet d'expliquer l'anti-corrélation observée en **figure 15** avec des ARN stables qui ont de courtes queues poly-A et des ARN instables qui ont de longues queues poly-A mais un modèle dans lequel le décapping serait dépendant de la déadénylation et constituerait l'étape limitante peut également expliquer cette anti-corrélation. Aussi on peut imaginer que selon les ARN et les conditions, ce soit l'un ou l'autre des mécanismes de dégradation qui se produise.



**Figure 22. Un modèle de dégradation des ARN indépendant de la déadénylation**

(A) Un modèle simplifié de la dégradation de l'ARN inclut deux paramètres cinétiques : un qui caractérise la vitesse de deadenylation ( $k_A$ ) et un autre qui caractérise la vitesse de decapping ( $k_D$ ). Dans le modèle canonique dépendant de la déadénylation, la vitesse de dégradation des ARN est dictée essentiellement par la vitesse de déadénylation, à condition que le  $k_D$  ait une valeur importante puisque la vitesse d'apparition de la forme qui subit le decapping dépend du  $k_A$  alors que dans un modèle alternatif, indépendant de la déadénylation, le  $k_A$  et le  $k_D$  ne sont pas liés et la vitesse de dégradation dépend du  $k_D$ .  $k_A$  et  $k_D$  désignent des constantes arbitraires de déadénylation et de dégradation respectivement. (B) Simulation de la moyenne des queues poly-A en fonction de la demi-vie en faisant varier les constantes de déadénylation et de dégradation dans un modèle de dégradation indépendant de la déadénylation.

Quel que soit le modèle de dégradation des ARN : de manière dépendante ou non de la déadénylation, les rapporteurs étudiés sont dégradés à des vitesses différentes et le ou les facteur(s) responsable(s) de la dégradation plus rapide d'un ARN dont la séquence codante n'est pas optimisée n'ont pas été identifiés.

Bien que Dhh1, impliquée dans le decapping ait été identifiée par *Radhakrishnan et al.* comme étant un facteur qui pouvait reconnaître les ARN non-optimaux et les adresser à la dégradation (Radhakrishnan et al., 2016), nous n'avons pas observé de stabilisation des rapporteurs utilisés testés en absence de Dhh1 alors que nous avons utilisé les séquences codantes publiées par ces auteurs (figure S8.A), ce qui a été validé par PCR quantitative (figure S8.C). Nous avons également testé les rapporteurs initiaux utilisés par *Radhakrishnan et al.* qui comportent des 5'UTR et 3'UTR différents de nos rapporteurs. Avec leurs rapporteurs, nous avons pu observer une stabilisation pour la version non-optimisée en condition de déplétion pour Dhh1 mais pas lorsque Dhh1 était délétée. Il n'y avait en revanche pas d'effet de la déplétion ou de la délétion de Dhh1 sur la version optimale, en accord avec les résultats publiés ((Radhakrishnan et al., 2016) et figure S8.B).

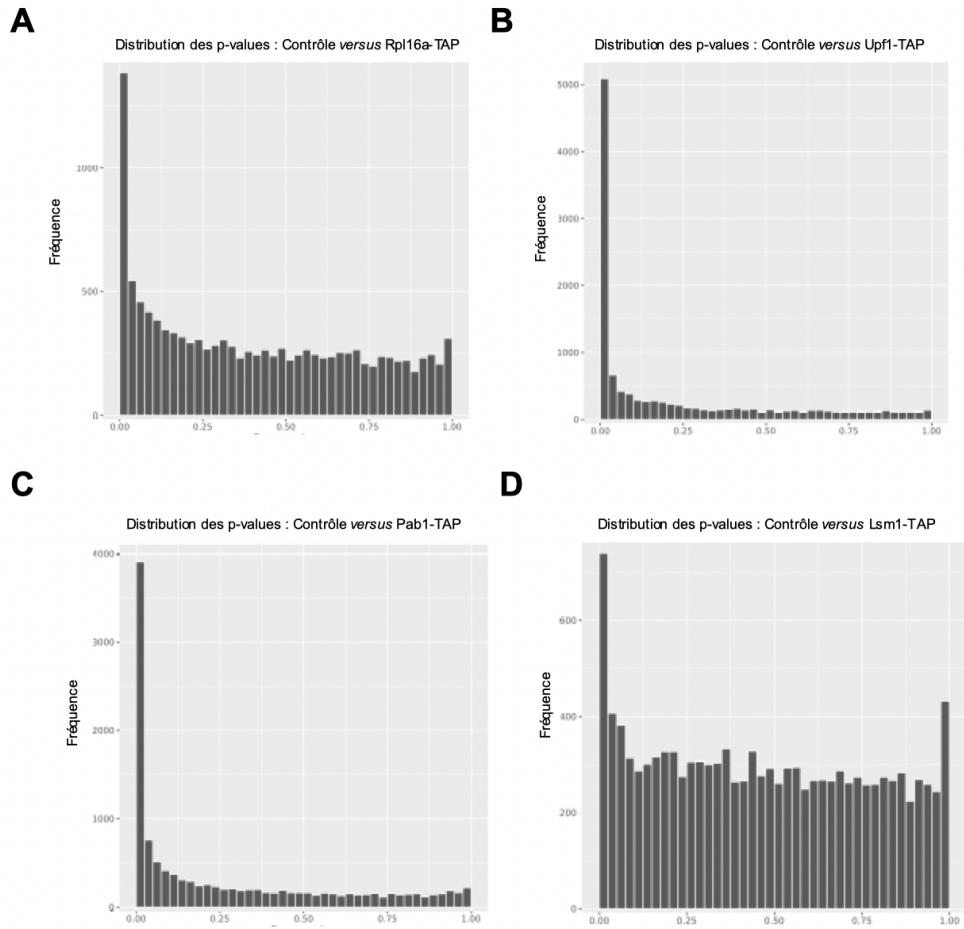
Ces différences de résultats pour des séquences identiques pourraient s'expliquer par le fait que les régions 5'UTR et 3'UTR des rapporteurs utilisés sont différentes, plusieurs études font état de différences de stabilités liée à ces séquences (Malabat et al.,

2015; Moqtaderi et al., 2018). Par ailleurs, l'arrêt transcriptionnel utilisé par *Radhakrishnan et al.* est un passage d'un milieu contenant du galactose à un milieu contenant du glucose qui peut entraîner plus de modifications dans le métabolisme cellulaire comparé à l'ajout de doxycycline pour nos rapporteurs.

L'étude de rapporteurs optimisés ou non pour la traduction nous a permis de découvrir que la déadénylation n'était pas essentielle pour la dégradation des ARN dans certains cas, et que les ARN n'avaient pas besoin d'être oligo-adénylés pour initier le décapping. Cependant nous n'avons pas pu identifier les facteurs qui favorisaient la dégradation rapide de la version non-optimisée ou les éléments qui protégeaient la version optimisée de la dégradation.

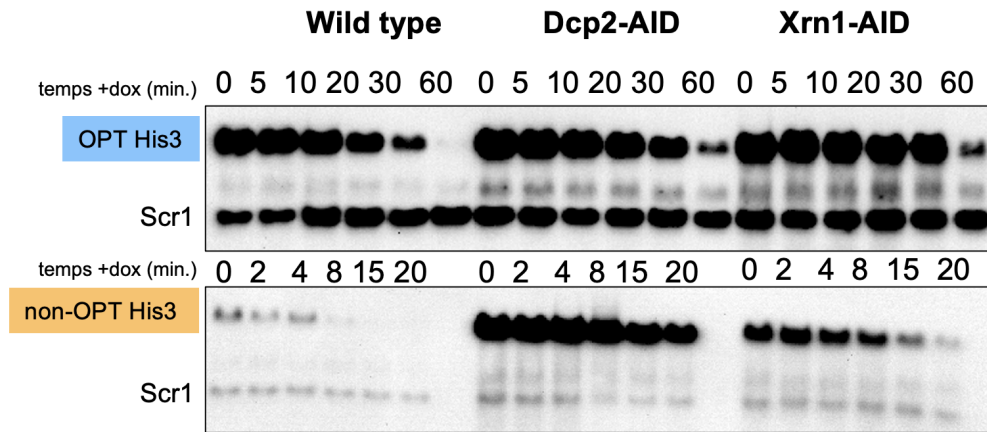
Pour identifier les facteurs impliqués dans la stabilité d'un ARN optimal ou non-optimal, il faudrait disposer d'un rapporteur pour lequel on pourrait rapidement analyser la disparition de l'ARN après arrêt transcriptionnel et qui pourrait être utilisé pour transformer une banque de mutants et tester l'effet des mutations sur la stabilité de l'ARN. Pour cela on pourrait imaginer la construction de rapporteurs analogues à ceux utilisés ici mais exprimant des ARN codant pour une protéine fluorescente dont la demi-vie reflète celle de l'ARN, par exemple une GFP (« *Green Fluorescent Protein* ») contenant une séquence déstabilisatrice PEST (Li et al., 1998) et utiliser ce système pour cribler par microscopie de fluorescence ou FACS (« *Fluorescence-activated cell sorting* ») une banque de délétions de gènes.

# Figures supplémentaires



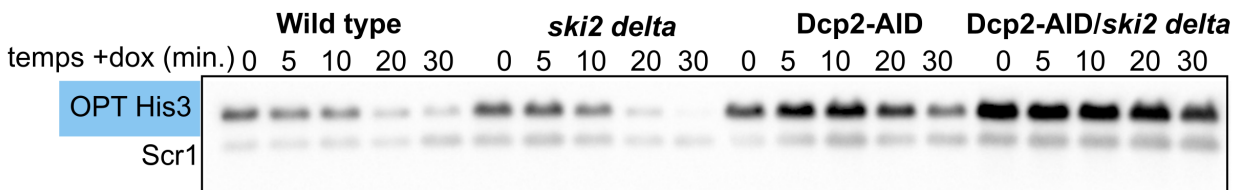
**Figure S1.** L'analyse statistique de DESeq2 révèle de nombreux transcrits différentiellement présents dans les purifications Rpl16, Upf1 et Pab1 et peu avec Lsm1

*Distribution des p-values ajustées en fonction de la fréquence des transcrits. Les transcrits dont la p-value est proche de 0 sont considérés comme plus significativement différentiellement présents que ceux proches de 1.*



**Figure S2. Les rapporteurs sont stabilisés par les déplétions de Dcp2 et Xrn1**

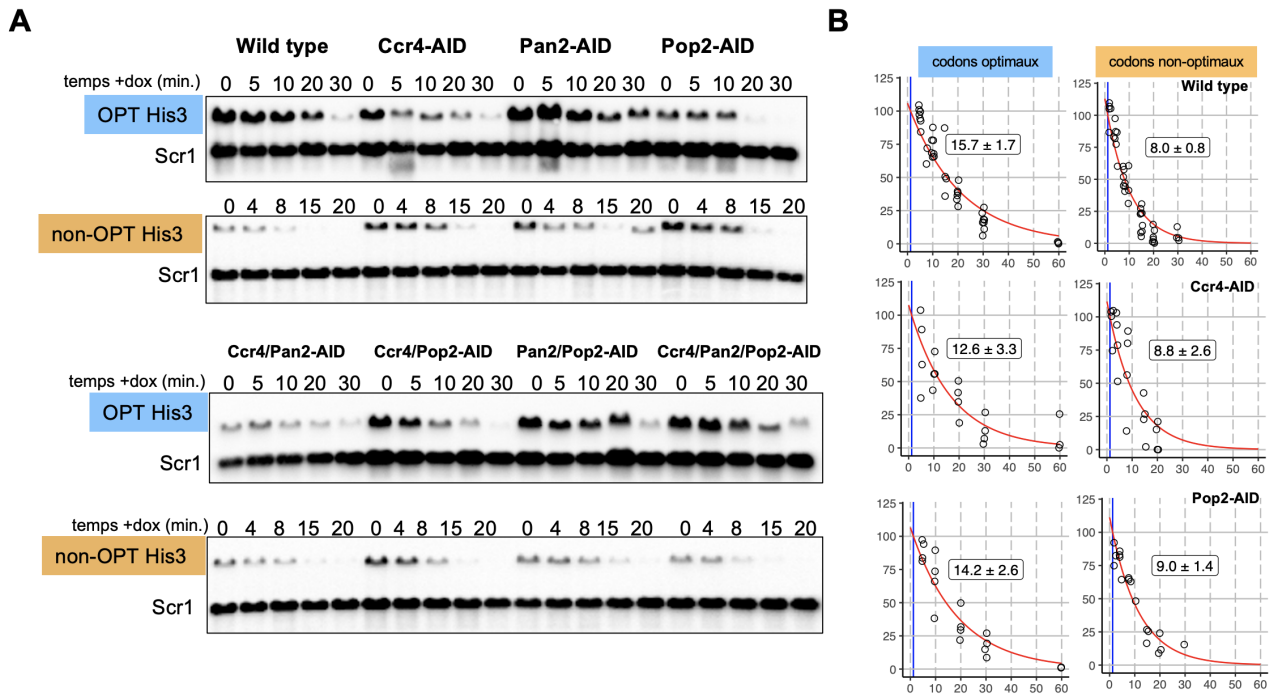
Northern blot réalisés avec des ARN totaux extraits de souches de souches *wild type*, *Dcp2-AID* ou *Xrn1-AID*, traitées avec de l'auxine et du  $\beta$ -estradiol pendant 1h et pendant différentes périodes de temps avec de la doxycycline. Les temps sont indiqués au dessus des northern blot. Les cellules ont été transformées avec les plasmides permettant l'expression de la version optimisée ou non de l'ARN de His3. Après dénaturation à 65°C dans du formazol, les ARN ont été séparés sur un gel d'agarose à 1,5% et transférés sur membrane de nylon. Les membranes ont été hybridées simultanément avec des sondes marquées à la digoxigénine correspondant soit à la région 3' UTR des ARN rapporteurs soit à l'ARN non codant *Scr1*, qui sert de contrôle de charge.



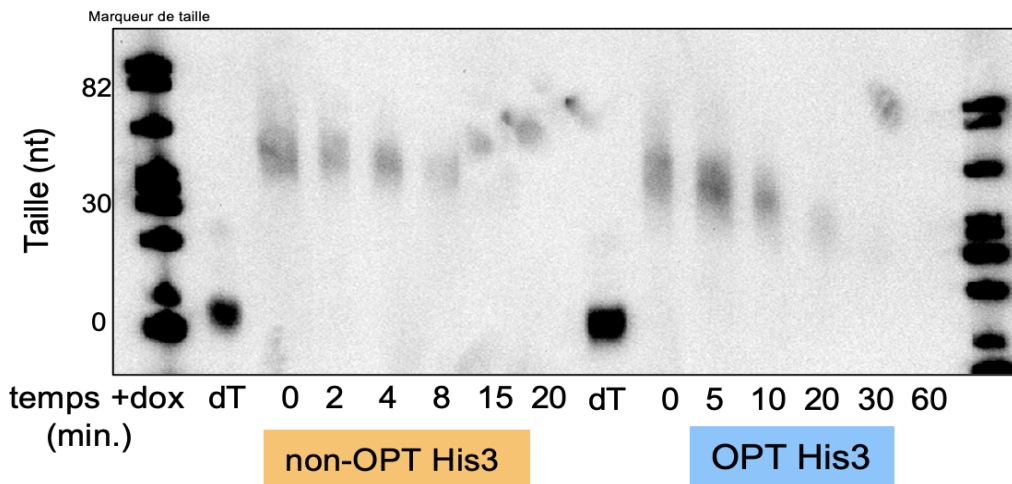
**Figure S3. La stabilité de His3 ne dépend pas de l'exosome cytoplasmique quand la voie 5' est active**

Northern blot réalisés avec des ARN totaux extraits de souches de souches *wild type*, d'une souche déléetée pour *SKI2* ou déplétée pour *Dcp2-AID*, ainsi qu'une souche *ski2 delta* déplétée pour *Dcp2-AID*, traitées avec de l'auxine et du  $\beta$ -estradiol pendant 1h et pendant différentes périodes de temps avec de la doxycycline. Les temps sont indiqués au dessus des northern blot. Les cellules ont été transformées avec les plasmides permettant l'expression de la version optimisée ou non de l'ARN de His3. Après dénaturation à 65°C dans du formazol, les ARN ont été séparés sur un gel d'agarose à 1,5% et transférés sur membrane de nylon. Les membranes ont été hybridées simultanément avec des sondes marquées à la digoxigénine correspondant soit à la région 3' UTR des ARN rapporteurs soit à l'ARN non codant *Scr1*, qui sert de contrôle de charge.



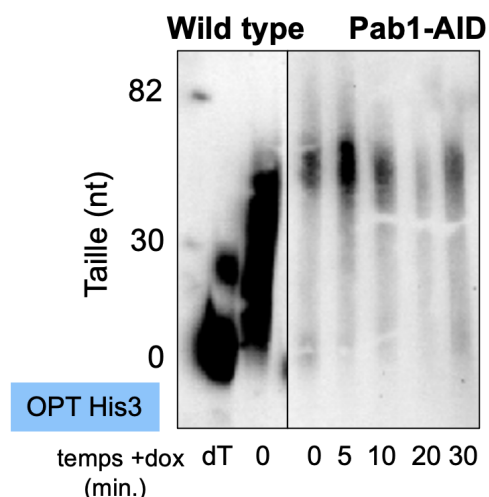


**Figure S4. La déplétion des déadénylases n'affecte pas la stabilité des rapporteurs**  
**(A)** Northern blot réalisés avec des ARN totaux extraits de souches wild type ou de souches déplétées pour les protéines AID indiquées, traitées avec de l'auxine et du  $\beta$ -estradiol pendant 1h et pendant différentes périodes de temps avec de la doxycycline. Les temps sont indiqués au dessus des northern blot. Les cellules ont été transformées avec les plasmides permettant l'expression de la version optimisée ou non de l'ARN de His3. Après dénaturation à 65°C dans du formazol, les ARN ont été séparés sur un gel d'agarose à 1,5% et transférés sur membrane de nylon. Les membranes ont été hybridées simultanément avec des sondes marquées à la digoxigénine correspondant soit à la région 3' UTR des ARN rapporteurs soit à l'ARN non codant Scr1, qui sert de contrôle de charge. **(B)** qPCR et mesures de demi-vies réalisées sur au moins 3 expériences indépendantes pour chaque souche et chaque plasmide. La demi-vie a été déterminée par une régression non linéaire et un « lag » de 1,2 minutes a été appliqué afin de corriger le délai de l'arrêt transcriptionnel. L'intervalle de confiance (95%) pour le temps de demi-vie de l'ARN basé sur une régression non-linéaire utilisant tous les points de minimum 3 répétitions de l'expérience est indiqué.



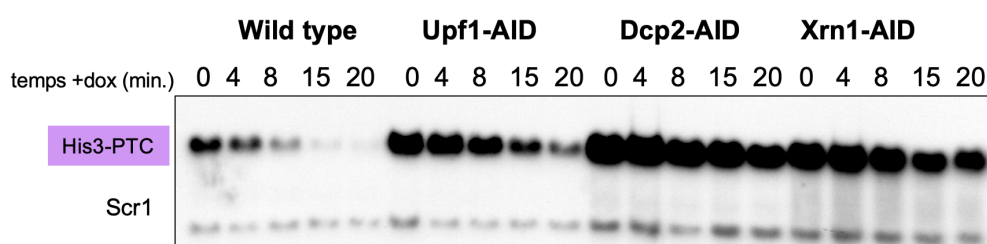
**Figure S5. La déplétion de Pop2 et Ccr4 empêche la déadénylation totale des ARN rapporteurs**

Northern blot réalisés avec des ARN totaux clivés à la RNaseH avec l'oligonucléotide LA121 provenant de la souche Ccr4-AID/Pop2-AID traitée avec de l'auxine et du  $\beta$ -estradiol pendant 1h et pendant différentes périodes de temps avec de la doxycycline pour arrêter la transcription du gène rapporteur. Les temps sont indiqués au dessus des northern blot. Les cellules ont été transformées avec les plasmides permettant l'expression de la version optimisée ou non de l'ARN de His3. Les ARN clivés ont été séparés, après dénaturation à 65°C dans du formazol, sur un gel d'acrylamide (6%) dénaturant. Après transfert des ARN sur membrane de nylon, celles-ci ont été hybridées avec une sonde marquée à la digoxigénine correspondant à la région 3' UTR des ARN rapporteurs. Une réaction de digestion à la RNaseH avec l'oligonucléotide LA121 et un oligo-dT (dT) permet de générer un fragment d'ARN sans queue poly-A (A0), utilisé pour estimer la taille des queues poly-A indiquées sur la gauche.



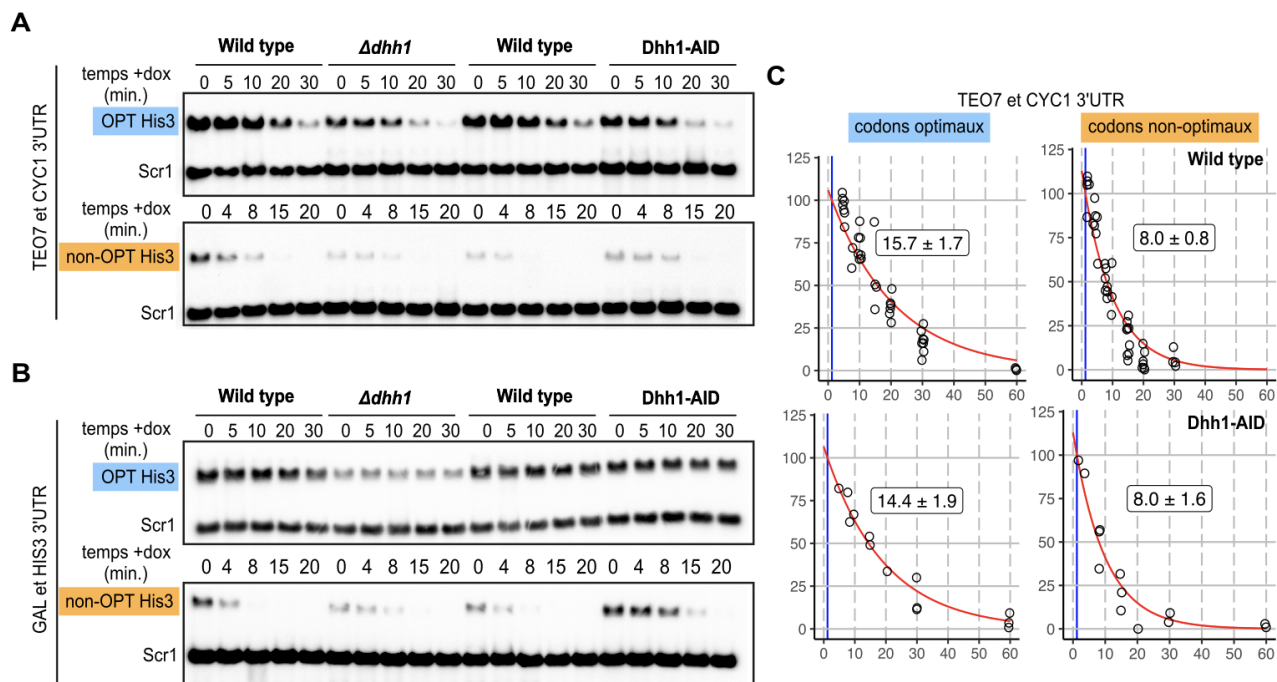
**Figure S6. La déplétion de Pab1 empêche le raccourcissement des queues poly-A**

Northern blot réalisés avec des ARN totaux clivés à la RNaseH avec l'oligonucléotide LA121 provenant de souches wild type ou Pab1-AID traitées avec de l'auxine et du  $\beta$ -estradiol pendant 1h et pendant différentes périodes de temps avec de la doxycycline pour arrêter la transcription du gène rapporteur. Les temps sont indiqués au dessus des northern blot. Les cellules ont été transformées avec les plasmides permettant l'expression de la version optimisée ou non de l'ARN de His3. Les ARN clivés ont été séparés, après dénaturation à 65°C dans du formazol, sur un gel d'acrylamide (6%) dénaturant. Après transfert des ARN sur membrane de nylon, celles-ci ont été hybridées avec une sonde marquée à la digoxigénine correspondant à la région 3' UTR des ARN rapporteurs. Une réaction de digestion à la RNaseH avec l'oligonucléotide LA121 et un oligo-dT (dT) permet de générer un fragment d'ARN sans queue poly-A (A0), utilisé pour estimer la taille des queues poly-A indiquées sur la gauche.



**Figure S7. Le rapporteur His3-PTC est sensible aux déplétions de Upf1, Dcp2 et Xrn1**

Northern blot réalisés avec des ARN totaux extraits de souches wild type ou Upf1-AID, Dcp2-AID, Xrn1-AID, traitées avec de l'auxine et du  $\beta$ -estradiol pendant 1h et pendant différentes périodes de temps avec de la doxycycline. Les temps sont indiqués au dessus des northern blot. Les cellules ont été transformées avec le plasmide His3-PTC. Après dénaturation à 65°C dans du formazol, les ARN ont été séparés sur un gel d'agarose à 1,5% et transférés sur membrane de nylon. Les membranes ont été hybridées simultanément avec des sondes marquées à la digoxigénine correspondant soit à la région 3' UTR des ARN rapporteurs soit à l'ARN non codant Scr1, qui sert de contrôle de charge.



**Figure S8. L'absence de Dhh1 ne stabilise pas les rapporteurs His3 optimal et non optimal**

**(A-B)** Northern blot réalisés avec des ARN totaux extraits de souches wild type, *dhh1* delta ou de souches déplétées *Dhh1-AID*, traitées avec de l'auxine et du  $\beta$ -estradiol pendant 1h et pendant différentes périodes de temps avec de la doxycycline. Les temps sont indiqués au dessus des northern blot. Les cellules ont été transformées avec le plasmide His3-PTC. Après dénaturation à 65°C dans du formazol, les ARN ont été séparés sur un gel d'agarose à 1,5% et transférés sur membrane de nylon. Les membranes ont été hybridées simultanément avec des sondes marquées à la digoxigénine correspondant soit à la région 3' UTR des ARN rapporteurs soit à l'ARN non codant *Scr1*, qui sert de contrôle de charge.

**(C)** qPCR et mesures de demi-vies réalisées sur au moins 3 expériences indépendantes pour chaque souche et chaque plasmide. En ordonnée est représenté le pourcentage d'ARN et en abscisse le temps après ajout de doxycycline. La demi-vie a été déterminée par une régression non linéaire et un « lag » de 1,2 minutes a été appliqué afin de corriger le délai de l'arrêt transcriptionnel. L'intervalle de confiance (95%) pour le temps de demi-vie de l'ARN basé sur une régression non-linéaire utilisant tous les points de minimum 3 répétitions de l'expérience est indiqué.

# Matériel et méthodes

## Souches de levure

Les souches de levure *Saccharomyces cerevisiae* dérivent toutes des souches de référence BY4741 (Mat a) et BY4742 (Mat alpha) (Baker Brachmann et al., 1998).

Toutes les souches exprimant une protéine fusion avec l'étiquette TAP sont issues d'une collection de souches construite par *Ghaemmaghami et al.* (Ghaemmaghami et al., 2003). La souche *dhh1Δ* provient d'une collection de souches construites par *Giaever et al.* (Giaever et al., 2002) distribuée par *EuroScarf* (<http://www.euroscarf.de>).

Une spore Mat alpha obtenue après sporulation de la souche GBy68, fournie par *G Bushkin* (Bushkin et al., 2019), a été croisée avec une souche BY4741 contenant la mutation *trp1-1* dans le gène *TRP1*. Après sporulation de la souche diploïde ainsi obtenue, nous avons sélectionné des spores ayant les génotypes suivants : Mat a, *LYS2*, *met15Δ0*, *ura3Δ0*, *leu2Δ0*, *his3Δ1*, *trp1-1*, *RME1-Δ*, *ins-308A*, *TAO3(E1493Q)* dénommée LMA5395 et Mat alpha, *lys2Δ0*, *ura3Δ0*, *leu2Δ0*, *his3Δ1*, *trp1-1*, *RME1-Δ*, *ins-308A*, *TAO3 (E1493Q)* dénommée LMA5396. La souche LMA5395 a été transformée avec le plasmide 1692 linéarisé par digestion avec *Bsu36I* et une souche prototrophe pour le tryptophane a été sélectionnée pour générer la souche LMA5419.

Pour construire les souches dans lesquelles les protéines d'intérêt peuvent être déplétées, une étiquette polyG-3Flag-miniAID amplifiée par PCR en utilisant le plasmide 1451 comme matrice (Dehecq et al., 2018) a été insérée juste avant le codon stop des gènes. Toutes ces insertions ont été faites dans la souche Mat alpha, LMA5396. Les souches obtenues ont été croisées avec la souche LMA5419. Après sporulation des cellules diploïdes obtenues, les souches ayant le génotype désiré ont été sélectionnées. Pour la souche permettant la déplétion de *Ccr4*, l'étiquette AID amplifiée par PCR en utilisant comme matrice le plasmide 1602 a été insérée juste après le codon initiateur du gène par recombinaison médiée par CRISPR/CAS9 (Mans et al., 2015).

## Milieux de culture et conditions de croissance

Les levures sont cultivées soit dans du milieu riche YPD soit du milieu synthétique complet dépourvu d'uracile pour la sélection des transformants et le maintien des plasmides réplicatifs. Les levures ont été cultivées à 30°C sauf pour les cinétiques de dégradation des ARN où elles ont été cultivées à 25°C.

## Construction de plasmides

Plasmides permettant l'expression des ARN de séquences optimisée ou non :

Les séquences codant pour la protéine His3 avec 10 % (non-opt His3) ou 100 % (opt-His3) de codons optimaux proviennent des données publiées par Radhakrishnan et ses collaborateurs (Radhakrishnan et al., 2016). Ces séquences synthétisées par l'entreprise TWIST ont été réamplifiées en utilisant les oligonucléotides LA070 et LA071 et clonés

dans les sites *Bam*HI et *Not*I du plasmide pCM189 (Garí et al., 1997) pour créer les plasmides pCM189-NFlag-HIS3-10 et pCM189-NFlag-HIS3-100.

Deux fragments du gène rapporteur opt-His3 ont été amplifiés par PCR avec les oligonucléotides MFR917 et 3890 d'une part et avec les oligonucléotides LA084 et 3889 d'autre part puis les deux produits de PCR ont été mélangés et ré-amplifiés en utilisant les oligonucléotides LA70 et LA71. Le produit de PCR ainsi obtenu a été digéré avec les enzymes de restriction *Sbf*I et *Bam*HI et cloné dans les sites *Sbf*I et *Bam*HI du plasmide pCM189\_HIS3\_100% afin de remplacer la séquence codante optimisée du gène HIS3 par une courte séquence codante de 24 acides aminés suivie d'un long 3'UTR. Le plasmide généré est dénommé pCM189-Nflag-HIS3-PTC24.

#### Plasmide pour le système de déplétion protéique induite par l'auxine (AID)

Un produit de PCR réalisé en utilisant les oligonucléotides 3765 et 3766 et le plasmide pFS461 (<https://www.addgene.org/search/catalog/plasmids/?q=pfs461>) comme matrice a été utilisé pour transformer une souche dérivée de BY4741 prototrophe pour l'uracile. Les transformants résistants au 5-FOA (acide fluoro orotique) ont été sélectionnés et le bon remplacement de la séquence codante du gène *URA3* par celle codant pour le facteur de transcription artificiel dénommé Z3EV\_Tf a été vérifié par séquençage du locus d'insertion. L'allèle (*ura3::Z3EV*) obtenu a été amplifié en utilisant les oligonucléotides 293 et 3746 et l'ADN génomique de la souche créée comme matrice. Le produit de PCR obtenu a été digéré avec l'enzyme de restriction *Sst*II et cloné dans les sites *Sst*II and *Sma*I du plasmide pRS304 pour générer le plasmide 1603.

Le promoteur Z3EV a été amplifié par PCR en utilisant comme matrice le plasmide pFS478 (<https://www.addgene.org/search/catalog/plasmids/?q=pfs478>) et les oligonucléotides 3740 et 3872. Le produit de PCR obtenu a été digéré avec les enzymes de restriction *Sal*I et *Bam*HI. La séquence codante pour la protéine Tir1 d'*O. sativa* suivie du terminateur du gène *NRD1* a été amplifiée par PCR en utilisant les oligonucléotides 3718 et 3719 et comme matrice l'ADN génomique d'une souche NRD1-AID donnée par Drice Challal du laboratoire de Domenico Libri. Le produit de PCR obtenu a été digéré avec les enzymes de restriction *Xho*I et *Bam*HI. Ces deux produits de PCR ont ensuite été clonés ensemble au site *Xho*I du plasmide 1603 pour générer le plasmide 1692.

### **Extraction d'ARN**

L'ARN total de levure a été extrait en utilisant le protocole décrit par Collart et Oliviero (Collart and Oliviero, 1993)

### **Cinétique de dégradation des ARN**

Pour les cinétiques de dégradation des ARN, les cellules transformées avec les plasmides permettant l'expression des différentes version de l'ARN codant pour His3 sous contrôle du promoteur TET-OFF ont été cultivées à 25°C dans du milieu synthétique dépourvu d'uracile jusqu'à une DO<sub>600</sub> d'environ 0,4 avant ajout de doxycycline à une concentration finale de 10µg/ml pour arrêter la transcription. Les cellules ont été récoltées à différents

temps après l'arrêt de la transcription, en général (sauf indication contraire) 0, 2, 4, 8, 10, 20, 30 minutes pour le rapporteur exprimant la version non-optimale de His3 ou la version contenant le PTC et 0, 5, 10, 20, 30, 60 minutes pour le rapporteur exprimant la version optimale de His3.

Lorsque les plasmides utilisés contenaient le gène *HIS3* sous contrôle du promoteur du gène *GAL1*, les cultures ont été réalisées dans du milieu synthétique dépourvu d'uracile et contenant du galactose (concentration finale 2% p/v) comme source de carbone et l'arrêt transcriptionnel a été réalisé par ajout à la culture de glucose à une concentration finale de 2% p/v.

Pour l'analyse et la caractérisation des souches AID, les cellules ont été cultivées à 25°C dans un milieu synthétique complet jusqu'à ce que la  $DO_{600}$  atteigne environ 0,4 puis du  $\beta$ -estradiol à une concentration finale de 1  $\mu$ M et de l'auxine à une concentration finale de 100  $\mu$ M ont été ajoutés dans le milieu pour induire la transcription de Tir1 et la dégradation de la protéine. Des aliquotes de culture ont été récupérées à différents temps après l'ajout des drogues pour extraire les protéines totales et analyser la disparition de la protéine par western blot avec un anticorps anti-Flag couplé à la peroxidase.

Pour les expériences d'arrêt de la traduction, la cycloheximide à une concentration finale de 5  $\mu$ g/ml a été ajoutée au milieu en même temps que la doxycycline.

### **Northern blot**

Les ARN totaux ont été séparés sur un gel d'agarose à 1,5 % (p/v) dans du TBE, transférés sur une membrane de nylon (Hybond N+, Amersham) puis crosslinkés aux UV à 0.120 Joule/cm<sup>2</sup> à 254 nm. Les membranes ont été hybridées avec des sondes ARN marquées avec de la digoxigénine par transcription *in vitro* avec l'ARN polymérase du phage T7 en utilisant le kit DIG RNA Labelling kit (SP6/T7) de Roche (cat. No 11175025910). Les oligonucléotides AJ529 (promoteur T7) et LA94 ou AJ529 et LA95 ont été hybridés pour générer les matrices permettant la synthèse des sondes spécifiques de la région 3' des rapporteurs His3 (His3-tCyc) et Scr1 respectivement.

Pour l'analyse de la taille des queues poly-A, les ARN totaux ont été clivés avec la RNaseH (NEB) en présence de l'oligonucléotide LA121 spécifique de la région 3'UTR des rapporteurs. Un oligo-dT a été ajouté à la réaction pour éliminer la queue poly-A et déterminer la taille du fragment généré sans A (« A0 »). Après une heure à 37°C, la réaction a été arrêtée par ajout d'EDTA et les ARN purifiés par traitement au phénol-chloroforme et précipitation à l'éthanol. Les ARN clivés ont été séparés par migration sur un gel d'acrylamide dénaturant (6% acrylamide, 8M urée), transférés sur membrane de nylon (Hybond N+, Amersham) et crosslinkés comme précédemment. L'échelle de poids moléculaire correspond au plasmide pBR322 digéré par *MspI* préalablement marqué à la digoxigénine en utilisant la terminale transférase (NEB) en présence de digoxigénin-11-ddUTP.

Après transfert, les membranes ont été colorées au bleu de méthylène (0.2% bleu de méthylène, 0.4M acétate de sodium et 0.4M acide acétique) ce qui permet de visualiser l'ARNr 5,8S utilisé comme contrôle de charge.



## Transcription inverse et PCR quantitative

Avant transcription inverse (RT) avec la Superscript II ou III (Invitrogen) que ce soit pour la préparation des banques ou les PCR quantitatives, les échantillons d'ARN extraits ont été traités à la DNase I (kit Ambion TURBO DNA-free).

Pour les quantifications par PCR quantitative du pré-messager et de l'ARN mature de *RPL28*, un mélange des oligonucléotides CS888 (*RPL28*-prematuration), CS889 (*RPL28*) avec des hexamères aléatoires ont été utilisés pour la transcription inverse. Des dilutions sérielles de l'ADNc ont été quantifiées par qPCR avec du SYBR Green (Sigma) en utilisant les amorces CS887 et CS888 pour le pré-messager de *RPL28* et CS889 et CS946 pour l'ARN mature. L'amplification a été réalisée dans un Stratagene MX3005P et le logiciel correspondant (système de PCR quantitative MXpro), avec une première étape de dénaturation de 5 minutes à 95°C suivie de 40 cycles à 95°C pendant 20 s et 60°C pendant 1 min.

Pour les quantifications avec les sondes TaqMan, la transcription inverse a été réalisée avec les oligonucléotides LA126 (*Act1*), LA102 (*His3-tCyc*) et des hexamères aléatoires. Des dilutions sérielles de l'ADNc ont été quantifiées par PCR quantitative avec des sondes TaqMan, LA122 pour *His3-tCyc*, LA125 pour *Act1* avec les amorces LA123 et LA102 pour *His3* et LA124 et LA126 pour *Act1*. L'amplification a été réalisée avec la machine Bio-Rad CFX96 et le logiciel correspondant (logiciel CFX Maestro), avec une première étape de dénaturation de 30 secondes à 95°C suivie de 40 cycles à 95°C pendant 15 secondes et 60°C pendant 30 secondes.

### Calcul de demi-vie par qPCR

Chaque valeur de cycle obtenu pour *Act1* a été retranché à celle de *His3* pour normaliser les valeurs. La moyenne de trois dilutions en série a été effectuée. Chaque point de cinétique a été retranché à la valeur du point à  $t=0$ .

Pour la détermination des demi-vies, il s'agit d'une régression non-linéaire appliquée à l'ensemble des données disponibles (tous replicats confondus). Un « lag » de 1,2 minutes a été appliqué afin de corriger le délai de répression de la transcription par la doxycycline (Baudrimont et al., 2017). Un intervalle de confiance a été appliqué en utilisant la fonction « confint » du paquet « MASS » (<https://rdrr.io/cran/MASS/man/confint.html>).

## Western blot

Des extraits de protéines totales ont été préparés à partir de 5 unités de  $DO_{600}$  d'une culture exponentielle avec la méthode utilisant un traitement alcalin (Kushnirov, 2000). Les protéines ont été dénaturées pendant 3 min à 95°C, et les débris cellulaires ont été culottés par centrifugation. 10  $\mu$ L de surnageant ont été déposés sur des gels d'acrylamide NuPAGE Novex 4-12% Bis-Tris (Life technologies). Après transfert sur membrane de nitrocellulose avec un système semi-sec rapide (Biorad trans-blot) avec tampon discontinu (BioRad technote 2134), les protéines ont été détectées par hybridation avec les anticorps appropriés.

## Purification par affinité (TAP) et séquençage (RIP-seq)

Les protéines marquées avec l'étiquette TAP ont été purifiées en utilisant une méthode de purification en une étape. Les culots cellulaires congelés provenant d'une culture de 4 litres ont été resuspendus dans un tampon de lyse (20 mM HEPES K pH 7,4, 100 mM KOAc, 0,5% Triton X100, 5mM MgCl<sub>2</sub>, inhibiteur de protéase, complexe Vanadyl-Ribonucléoside 1X, 40 unités ml<sup>-1</sup> RNasin) et lysés en utilisant un MagNAlyser (Roche) (deux passages de 90s à 4000 rpm). Le lysat a été clarifié par centrifugation à 4°C pendant 20 min à 14000 t/min. Des billes magnétiques (Dynabeads M-270 epoxy, invitrogen) couplées à des IgG ont été ajoutées à l'extrait protéique et incubées pendant 1,5 h à 4°C (Oeffinger et al, 2007). Les billes ont été recueillies à l'aide d'un aimant et lavées cinq fois avec un tampon de lavage (20 mM HEPES K pH 7,4, 100 mM KOAc 0,5% Triton X100, 5mM MgCl<sub>2</sub>). Après lavage, les protéines ont été éluées par incubation dans un tampon d'éluion (SDS 2% dans Tris-EDTA pH 7.5) 10 min à 65°C. L'éluat a ensuite été traité au phénol/chloroforme pour éliminer les protéines et les acides nucléiques ont été précipités avec de l'acétate d'ammonium et de l'éthanol. Les échantillons ont été traités avec RiboZero (Illumina) ou Ribominus (invitrogen) pour éliminer les ARN ribosomiques 25S et 18S (plus les ARN 5,8S et 5S dans le cas du RiboZero) et les banques ont été préparées à l'aide du kit TruSeq Stranded mRNA (Illumina). Nous avons strictement suivi le protocole, sauf que nous avons commencé directement à l'étape de fragmentation après le RiboZero et que nous avons ajusté le nombre de cycles requis pour amplifier les banques.

Une aliquote du lysat clarifié a été prélevée avant la purification TAP pour extraire les ARN totaux (input) et a été traitée de la même manière que les fractions purifiées pour générer la banque.

### Traitement des résultats du séquençage

Nous remercions Caroline Proux et Laurence Ma de la plateforme Biomics pour le séquençage des échantillons avec le séquenceur Illumina hiseq2500 ainsi que le démultiplexage des séquences.

Les différentes étapes pour obtenir le fichier de comptage depuis les fichiers *.fastq* ont été réalisées sur la plateforme Galaxy de l'Institut Pasteur (Mareuil et al., 2017).

Les « reads » ont été alignés sur les 16 chromosomes et le génome mitochondrial de la souche de référence S288C. Pour l'alignement, nous avons utilisé le logiciel STAR (version STAR\_2.7), qui permet d'aligner sur les introns. Les paramètres par défauts ont été appliqués sauf les suivants : `--alignIntronMax 1500 --alignMatesGapMax 1500 --alignSJoverhangMin 25`. Le génome utilisé est *Saccharomyces\_cerevisiae.R64-1-1.dna.toplevel.fa* qui provient de <https://www.ensembl.org/index.html> avec le fichier *Saccharomyces\_cerevisiae.R64-1-1.101.gtf* qui donne les positions des introns et des exons.

Pour le comptage, nous avons utilisé *feature count* du paquet *Subread* avec un fichier d'annotations de notre laboratoire (*scGFF2GTF\_final14.gtf*) qui contient les annotations pour les SUTs/XUTs/CUTs avec les paramètres supplémentaires suivants « O -M ».

Pour la visualisation IGV (Robinson et al., 2011), la fonction *bamCoverage* de *deepTools* a été utilisée en mettant le bin à 1 et sans normalisation.

Pour l'analyse statistique de la **figure S1**, DESeq2 avec l'interface de Sartools a été utilisé en laissant les paramètres par défaut (Love et al., 2014; Varet et al., 2016). L'analyse des transcrits statistiquement différentiellement présents dans les purifications a été réalisée par rapport à l'input.

#### Analyse des résultats de RIP-seq

Une fois les fichiers de comptages obtenus, deux ARN sur-représentés ont été enlevés : RDN et YLR154C. Ensuite le nombre de reads a été filtré pour chaque gène en mettant un seuil à 10 reads pour les 3 réplicats de l'input. Pour normaliser les comptages, chaque gène a été divisé par le nombre total de reads pour chaque purification. La moyenne des  $\text{Log}_2$  des 3 réplicats a ensuite été effectuée. Les valeurs d'enrichissement ont été obtenues en faisant le rapport de chaque purification sur l'input. ( $\text{Log}_2$  purif –  $\text{Log}_2$  input). Afin de centrer les résultats, chaque valeur d'enrichissement a été corrigée par la médiane du nombre de reads de chaque purification.

Pour l'analyse des **figures 12 et 13**, les transcrits suivants ont été enlevés : les CUTs, les ARNsn, les ARNsno, les ARNt, les ETS, les ITS, les transposons, *Icr1*, *Irt1*, *Lsr1*, *Nme1*, *Rna170*, *Rpr1*, *Scr1*, *Srg1* et *Tlc1*. Pour les cibles du NMD, tous les XUTs, SUTs et introns des ARN messagers ainsi que deux autres ARN cibles du NMD, *Ccw22* et *Ctr2* ont été sélectionnés. Le test statistique de Wilcoxon a été utilisé pour déterminer la significativité des différences au sein des classes de transcrits sélectionnés.

Pour la catégorie « Biosynthèse des acides aminés », les 150 gènes les moins associés à Pab1-TAP qui faisaient partis des 2750 ARN les plus abondants (dans l'input) ont été sélectionnés, en excluant les CUTs, les ARNt et d'autres ARN nucléaires. La liste générée a ensuite été chargée sur G-profiler dans le but de sortir une liste de gènes associés à une voie cellulaire (Raudvere et al., 2019).

G-profiler a trouvé une liste associée à la voie de biosynthèse des acides aminés avec une significativité de p-value ajustée =  $4,764 \times 10^{-17}$ , il s'agit de la KEGG:01230 qui regroupe les gènes suivants : *CYS3*, *CDC19*, *PRS4*, *LYS2*, *TYR1*, *PYC2*, *HIS7*, *ARO4*, *SHM1*, *ILV6*, *HIS4*, *HA1*, *CIT2*, *PGK1*, *THR4*, *IDP1*, *LYS21*, *GLT1*, *LYS20*, *TRP1*, *ARO3*, *TPI1*, *ALT2*, *ARO1*, *HOM2*, *LYS4*, *PRO1*, *TRP4*, *SAM2*, *GLY1*, *PRO3*, *HOM3*, *HIS1*, *ARG5,6*, *SER3*, *ILV1*, *TRP2*, *MET6*, *PRS2*, *HIS2*, *IRC7*, *LEU1*, *TRP5*, *PYC1*, *ARO2*, *ARO8*, *MCY1*, *ASN2*, *CYS4*, *TDH3*, *SER2*, *PFK1*, *ENO1*, *PRS3*, *ARG4*, *THR1*, *ENO2*, *BAT1*, *HIS6*, *SER33*, *LYS12*, *HIS5*, *LYS1*, *TDH1*, *ARG2*, *RPE1*, *ACO2*, *TDH2*, *ILV3*, *STR2*, *HOM6*, *BAT2*, *FBA1*, *AAT1*, *GPM1*, *PRS1*, *TRP3*, *YLL058W*, *AAT2*, *SHM2*, *ALT1*, *SAM1*, *ACO1*, *TAL1*, *ILV5*, *YML082W*, *ARG7*, *ILV2*, *PFK2*, *IDP3*, *IDH1*, *LEU4*, *MET2*, *PHA2*, *CIT1*, *LYS9*, *GPM3*, *ARG1*, *PRS5*, *ARG8*, *RKI1*, *LEU9*, *IDH2*, *SER1*, *HIS3*, *PRO2*, *PYK2*, *CAR1*, *GLN1*, *ARO7*, *TKL1*, *ASN1*.

#### **Mesures des queues poly-A**

Nous avons utilisé le protocole publié par (Jänicke et al., 2012). Pour les échantillons de RIP (Pab1-TAP, Lsm1-TAP, Rpl16a-TAP, Upf1-TAP) nous avons utilisé des quantités

d'ARN variables, de 5ng à 100ng. Pour l'input, nous avons utilisé 1 µg d'ARN. Pour le reste, le protocole est le même que celui décrit par *Jänicke et al.*, nous avons utilisé l'oligonucléotide CS1400 qui s'hybride au niveau des queues poly-A des ARN et pour le TVN-PAT, nous avons utilisé l'oligonucléotide CS1401 qui s'ancre au niveau du site de polyadénylation de l'ARN.

Pour les réactions de PCR, les oligonucléotides CS1402 et CS1455 et CS1402 et GB1399 ont été utilisés pour l'amplification des queues poly-A de *HHT2* et *CCW22* respectivement. Les réactions de PCR ont été réalisées avec la polymérase Phusion (ThermoFisher) avec 30 cycles d'amplification. Les produits de PCR ont été déposés sur un gel d'agarose 2,5% (p/v) et colorés au Bromure d'Ethidium.

### **Spectrométrie de masse et analyses des résultats**

Les expériences liées à la spectrométrie de masse et au western blot de la **figure 9** ont été réalisées selon le protocole décrit dans (Dehecq et al., 2018) excepté pour la normalisation pour la spectrométrie de masse. Pour le score LTOP2, la moyenne de l'intensité des trois plus intenses peptides pour chaque protéine (ou 2 si il n'y en avait que 2) a été réalisée. Seules les protéines pour lesquelles il y avait au moins deux valeurs indépendantes de LTOP2 ont été retenues. Les valeurs brutes de LTOP2 pour chaque purification ont été centrées sur la médiane.

Pour générer la **figure 9.A**, les protéines qui avaient un enrichissement supérieur à 4 avec une p-value inférieure à 0,05 dans au moins une purification ont été sélectionnées. Des groupes de protéines ont été sélectionnés, le groupe « épissage de l'ARN » contient les protéines suivantes : Brr2, Prp3, Prp4, Prp6, Prp8, Prp24, Prp28, Prp31, Prp38, Smd1, Snu66, Snu23, Snu114, Sad1, Smd1, Smd2, Spp381, Lsm8 et le groupe « biogénèse des ribosomes contient les protéines suivantes : Rex, Nop56, Nop1, Cbf5, Fpr3, Prp43.

### **Simulations de la dégradation de l'ARN**

Des simulations de la dégradation de l'ARN pour la **figure 22.B** qui incluent la déadénylation et le décapping ont été réalisées auparavant et impliquaient de nombreuses étapes et constantes cinétiques pour tenir compte d'un nombre important d'étapes de déadénylation (D. Cao and Parker, 2001). Des modèles plus simples ont été par la suite proposés (Wu et al., 2013; Wu and Tian, 2016) et semblent aussi adaptés à la description du processus de dégradation de l'ARN. Nous avons utilisé un modèle cinétique simplifié pour répondre à la question du lien entre la vitesse de déadénylation, la fréquence du décapping et la vitesse apparente de dégradation d'un ARN. Nous nous sommes basés sur le constat que la plupart du temps, les temps de vie des ARN sont estimés dans des systèmes dans lesquels on interrompt la transcription et on regarde le devenir, dans le temps de toutes les espèces d'ARN, polyadénylés ou oligo-adénylés. Les schémas réactionnels sont présentés dans la **figure 22.A**. Nous avons noté une constante cinétique apparente  $k_A$  pour la vitesse de passage des ARN à queues poly(A) longues à des queues

poly(A) intermédiaires et plus loin à des ARN oligoadénylés, qui sont ensuite décappés avec une constante de vitesse  $k_D$ .

Deux modèles ont été utilisés dans *tellurium* (Choi et al., 2018), un programme permettant de simuler le devenir des divers espèces moléculaires au cours du temps. Le modèle « en série » (non montré), dans lequel seule l'espèce oligo-adenylée peut être decappée est le suivant, en utilisant la syntaxe *antimony* (Smith et al., 2009) :

```
$T -> PAI; kS*T
PAI -> PAm; kA*PAI
PAm -> PAs; kA*PAm
PAs -> ; kD*PAs
```

Ici, la formation de l'ARN avec de longues queues poly-A, PAI, se fait à une vitesse constante, dictée par la constante  $k_S$  et le facteur T, représentant la transcription (le signe \$ devant T indique une concentration constante). Dans le cas où la dégradation de l'ARN peut survenir sur toutes les formes d'ARN **en figure 22.B**, le modèle cinétique est légèrement différent (les étapes supplémentaires de dégradation des ARN indépendant de la déadénylation sont les deux dernières) :

```
$T -> PAI; kS*T
PAI -> PAm; kA*PAI
PAm -> PAs; kA*PAm
PAs -> ; kD*PAs
PAI -> ; kD*PAI
PAm -> ; kD*PAm
```

Ces modèles sont adaptables à la fois à des expériences de chasse par arrêt transcriptionnel (dans ce cas le terme contenant T disparaît) et à des simulations de situations en « steady-state ». Pour comprendre comment les diverses espèces PAI, PAm et PAs varient les unes par rapport aux autres dans diverses conditions et dans des conditions stationnaires, nous avons résolu le système d'équations différentielles correspondant aux deux modèles cinétiques à l'aide de *Maxima* (<https://maxima.sourceforge.io>). Les valeurs d'abondance relative pour les espèces de poly-A de tailles différentes sont, dans le premier modèle (non montré) :  $PAI=T/k_A$ ,  $PAm=T/k_A$  et  $PAs=T/k_D$  et, pour le deuxième modèle en **figure 22.B** :  $PAI= T/(k_A+k_D)$ ,  $PAm = T/(k_D^2/k_A+2*k_D+k_A)$ ,  $PAs = T/(k_D^3/k_A^2 + 2*k_D/k_A + k_D)$ .

## Disponibilités des données

Le numéro d'accès des données de séquençage est GSE160642. Les données de protéomiques ont été déposées dans le ProteomeXchange *via* le dépôt PRIDE (Vizcaíno et al., 2016) avec l'identifiant PXD028008.

<b>Souche</b>	<b>Génotype</b>	<b>Référence</b>
LMA2154	MATa ura3Δ0 his3Δ1 leu2Δ0 met15Δ0	(Baker Brachmann et al., 1998)
LMA2194	UPF1(YMR080C)-TAP::HIS3MX6	(Ghaemmaghmi et al., 2003)
LMA2195	RPL16A(YIL133C)-TAP::HIS3MX6	(Ghaemmaghmi et al., 2003)
LMA2626	ski2(YLR398C)Δ::KANMX6	(Giaever et al., 2002)
LMA4985	LSM1(YJL124C)-TAP::HIS3MX6	(Ghaemmaghmi et al., 2003)
LMA4987	PAB1(YER165W)-TAP::HIS3MX6	(Ghaemmaghmi et al., 2003)
LMA5042	DCP2(YNL118C)-degron::NAT, ski2(YLR398C)Δ::Hygro	(Zhang et al., 2018)
LMA5060	ccr4(YAL021C)Δ::KANMX6	(Giaever et al., 2002)
LMA5062	pop2(YNR052C)Δ::KANMX6	(Giaever et al., 2002)
LMA5262	dhh1(YDL160C)Δ::KANMX6	(Giaever et al., 2002)
LMA5395	TAO3_E1493Q/RME1(YGR044C) insert A - 308/trp1(YDR007W)_D87Stop (trp1-1)	Cette étude
LMA5396	TAO3_E1493Q/RME1(YGR044c) insert A - 308/trp1(YDR007w)_D87Stop (trp1-1)	Cette étude
LMA5419	TAO3_E1493Q/RME1(YGR044C) insert A - 308/trp1-1::TRP1-prURA3-Z3EVtf-tURA3- Z3EV-TIR1	Cette étude
LMA5575	LMA5419 DCP2(YNL118c)::polyG-3Flag- AID::KANMX6	Cette étude
LMA5578	LMA5419 XRN1(YGL173C)::polyG-3Flag- AID::KANMX6	Cette étude
LMA5593	LMA5419 CCR4(YAL021C)::polyG-3Flag- AID::KANMX6	Cette étude
LMA5595	LMA5419 PAN2(YGL094C)::polyG-3Flag- AID::KANMX6	Cette étude
LMA5597	LMA5419 POP2(YNR052C)::polyG-3Flag- AID::KANMX6	Cette étude
LMA5599	LMA5419 AID-3Flag-polyG-	Cette étude

	CCR4(YAL021C)/PAN2(YGL094C)::polyG-3Flag-AID::KANMX6	
LMA5601	LMA5419 AID-3Flag-polyG-CCR4(YAL021C)/CAF1(YNR052c)::polyG-3Flag-AID::KANMX6	Cette étude
LMA5603	LMA5419 AID-3Flag-polyG-PAN2(YGL094C)/POP2(YNR052C)::polyG-3Flag-AID::KANMX6	Cette étude
LMA5605	LMA5419 AID-3Flag-polyG-PAN2(YGL094C)/POP2(YNR052C)/CCR4(YAL021C)::polyG-3Flag-AID::KANMX6	Cette étude
LMA5615	LMA5419 DHH1(YDL160C)::polyG-3Flag-AID::KANMX6	Cette étude
LMA5629	LMA5419 PAB1(YER165W)::polyG-3Flag-AID::KANMX6	Cette étude
LMA5654	LMA5419 UPF1(YMR080C)::polyG-3Flag-AID::KANMX6	Cette étude

<b>Plasmide</b>	<b>ID et description (marqueur levure)</b>	<b>Référence</b>
pCM189-NFLAG-HIS3-10	1638 (TETO7-NFLAG-HIS3-10) (URA3)	Cette étude
pCM189-NFLAG-HIS3-100	1640 (TETO7-NFLAG-HIS3-100) (URA3)	Cette étude
pCM189-NFLAG-HIS3-PTC24	1707 (TETO7-NFLAG-HIS3-PTC24) (URA3)	Cette étude
pJC858_pGAL_FLAG_10pc_HIS3	1710 (pGAL-NFLAG-HIS3-10) (URA3)	(Radhakrishnan et al., 2016)
pJC867_pGAL_FLAG_optHIS3	1711 (pGAL-NFLAG-HIS3-100) (URA3)	(Radhakrishnan et al., 2016)
pFA6a-polyG-3Flag-miniAID_opt-kanMX6 (72426)		1451 (Dehecq et al., 2018)
pRS304_pURA3-Z3EVtf-tURA3		1603 Cette étude
pRS304_pURA3-Z3EVtf-tURA3-Z3EV-TIR1		1692 Cette étude



<b>N° oligonucléotide</b>	<b>Nom</b>	<b>Sequence 5' → 3'</b>	<b>Utilisation</b>
3314	dcp2-F2	GCCCAACTTCGACGTCCGTTC GAATGGAAGCTTCAGGGTCTAA TGAATTATTAAGCATTTTGCAT AGGAAGcggatccccgggtaattaa	Souche Dcp2-AID
3692	Dcp2-R1	TATTTTACTTTACATCGCTTCG GCTGCCTTCATTTACAGTGTGT CTATAAACGTATAACACTTATT CTTgaattcgagctcgtttaaac	Souche Dcp2-AID
3891	dhh1-f2	GCAGCAGGGGTATCCTCCACA GCAGGAACATTTTCATGGCGAT GCCACCTGGTCAGTCACAACC CCAGTATcggatccccgggtaattaa	Souche Dhh1-AID
3892	dhh1-r1	aaaataatcgacgattagatcacaaaa aagcgtatctcaccacagtagtatttttctt agatattctGAATTCGAGCTCGTT TAAAC	Souche Dhh1-AID
3490	nam7-f2	CTTCGCGAGAGAGTTACAGAG AGAAGAACA AAAAGCATGAATT GTCAA AAGACTTCAGCAATTT GGGAATAcggatccccgggtaattaa	Souche Upf1-AID
MD116	UPF1_stop_ R1_longtine	GTATCACAAGCCAAGTTTAAACA TTTTATTTTAAACAGGGTTCACC GAATTAgaattcgagctcgtttaaac	Souche Upf1-AID
CS1479	PAB1_fw_F 2	CTGCCTATGAGTCTTTCAAAAA GGAGCAAGAACAACAACTG AGCAAGCTcggatccccgggtaatta a	Souche Pab1-AID
CS1480	PAB1_rv_R 1	AAGATGATAAGTTTGTGAGTA GGGAAGTAGGTGATTACATAG AGCATTAgattcgagctcgtttaaac	Souche Pab1-AID
3705	caf1-f2	GCATAAGTTCCCAAACGGTAC GGATTTGCGCAAGTACCAAGG TGTCATATACGGTATTGATGGG GACCAAcggatccccgggtaattaa	Souche Pop2-AID
3706	caf1Δ- M13rev	TCCAATCATAAGCTGATGTTGC TTTTAACTTTTTTTTTTAAAA TTGTGTATACATATAGTACATAA	Souche Pop2-AID

		ATGAggaaacagctatgacatg	
3707	pan2-f2	GAAAGTACTGAACAGCGTGTA CGAAGAAGGAAGAGCCCATA ATTTCAAAGTTCCAGAACTT CAAAGGGAcggatccccgggtaatt aa	Souche Pan2-AID
3708	pan2- M13rev	TTTTTTGTATACTTTTCCAACC AGTGGTGCGCTGGTGGCTCTT GAGATTACGTGAAAGGCACTG CACCATAggaaacagctatgacatg	Souche Pan2-AID
3836	AID-Flag- CCR4- NoNGG-fwd	AAGGGCACAGCATAAGGGAC ACCAGCAAGGGAAGTCTGACT GACGTTATCCCTGCAAACACTAC CGCTAATGaaagaaaagagtgc c	Souche AID-Ccr4
3757	pGAL-AID- Flag-CCR4- rev	TGCTGCTGTTGCTGCTGTTGC TGTTGCTGCTGCGGGCCAC ATTAGGGTAGCCTAGTAAAGA AGGGTCGTTcttgcacgtcatccttg	Souche AID-Ccr4
3834	ccr4- gRNA_fwd	TGCGCATGTTTCGGCGTTCGA AACTTCTCCGCAGTGAAAGAT AAATGATCttgcagggataacgtcag tGTTTT	Souche AID-Ccr4/ARN guide pour CRISPR/CAS9
3835	ccr4_gRNA _rev	GTTGATAACGGACTAGCCTTAT TTAACTTGCTATTTCTAGCTC TAAAACactgacgttatccctgcaaaG ATCA	Souche AID-Ccr4/ARN guide pour CRISPR/CAS9
MFR1018	MFR1018	AGTCACAAAGCAATGCTGCTG ACCGTGATAATAAAAAGACG AATCTACTcggatccccgggtaattaa	Souche Xrn1-AID
MFR1019	MFR1019	GATATACTATTAAAGTAACCTC GAATATACTTCGTTTTTAGTCG TATGTTgaattcgagctcgttaaac	Souche Xrn1-AID
3889	HIS3_100_ PTC24_fwd	TTACTAACGAAACcTAGATTCA AATTGCCATC	Construction His3 PTC
3890	HIS3_100_ PTC24_rev	CAATTTGAATCTaGGTTTCGTT AGTAATACGC	Construction His3 PTC
3718	Nrd1-AID	CATGGTACTcGAgAAGCGC	construction

			plasmide 1692
3719	Nrd1-AID	aaaggatccATGACTTATTTTCCT GAAGAAG	construction plasmide 1692
3740	Z3EV- promotor- Sall-fwd	CGGCGtcGACGAGGCAAGC	construction plasmide 1692
3872	Z3EV- promotor- BamHI-rev	aatggatccTATAGTTTTTTCTCCT TGACGTAAAG	construction plasmide 1692
293	ura3R	TGTTCTTGATTTGTGCCCGG	construction plasmide 1603
3746	ura3_SstII(-) 237+	agaaggtaCCgCggctgtgg	construction plasmide 1603
3765	ura3::TIR1- fwd	gaaattgccagttattcttaaccaactgc acagaacaaaaacctgcaggaaacgaa gataaatcATGACTTATTTTCCTG AAGAAGTTG	insertion de Z3EV_Tf au locus URA3
3766	ura3::TIR1- rev	taattaaattgaagctctaatttgagtttag tatacatgcatttactataatacagtttTC ATAAAATCTTGACGAAGTTAG GAGC	insertion de Z3EV_Tf au locus URA3
AJ529	AJ529_for_ T7_dble_str and	GCTAATACGACTCACTATAGGG	Promoteur T7
CS1400	PAT- anchor_ primer	GCGAGCTCCGCGGCCGCGTT TTTTTTTTTT	ePAT
CS1401	polyA_tail_e valuation	GCGAGCTCCGCGGCCGCGTT TTTTTTTTTTVN	Oligo-dT ancré pour ePAT
CS1402	polyA_tail_e valuation	GCGAGCTCCGCGGCCGCG	ePAT
CS1455	HHT2_fwd	CGTTTACAGACATAATTGCGG G	Amplification de Hhf2
GB1399	YDR134c_f wd	CTTAGGTTACACGCAACTAAA C	Amplification de Ccw22
LA70	His3_DNA_ synthesis_f	aaatacacacactaaattaccggatc	clonage de His3

	wd		
LA71	His3_DNA_synthesis_rv	taactaattacatgatgcggcc	clonage de His3
LA84	tCyc_rv	gggacctagacttcaggttg	clonage du terminateur Cyc1
LA94	His3-tCyc_RNA_probe	ccgctctaaccgaaaaggaaggaggttag acaacctgaagtctaggtccctatttttt tatagttatgccctatagtgagtcgtattagc	Sonde ARN His3-cyc1
LA95	Scr1_RNA_probe	GAAGTGTCCCGGCTATAATAA ATCGATCTTTGCGGGCAGCCC GTTGGCAGccctatagtgagtcgtatt agc	Sonde ARN Scr1
LA102	pCM189_rev	cttcaggtgtctaactcctcc	Transcription inverse et PCR quantitative pour His3
LA121	RnaseH_cleavage_rev	ctgcagggccctagcCTACAT	Clivage par la RNaseH
LA122	His3_5'-Cy5/TAO/3'IBRQ_fwd	taattagttatgtcacgcttacattcacgc	Sonde Taqman
LA123	pcm189_100_fwd	cctgcaggagggccgcat	PCR quantitative pour His3
LA124	Act1_fwd	CAAGGTATCATGGTCCGGTATG G	PCR quantitative pour Act1
LA125	ACT1_5'-HEX/ZEN/3'IBFQ_fwd	CCTACGTTGGTGATGAAGCTC AATCCA	Sonde Taqman
LA126	Act1_rev	CGTGTTC AATTGGGTAACGTA AAG	Transcription inverse et PCR quantitative pour Act1
MFR917	MFR917	TATGGCATGCATGTGCTCTG	Construction His3-PTC

# Bibliographie

- Abernathy, E., Gilbertson, S., Alla, R., Glaunsinger, B., 2015. Viral Nucleases Induce an mRNA Degradation-Transcription Feedback Loop in Mammalian Cells. *Cell Host & Microbe* 18, 243–253. <https://doi.org/10.1016/j.chom.2015.06.019>
- Alhusaini, N., Collier, J., 2016. The deadenylase components Not2p, Not3p, and Not5p promote mRNA decapping. *RNA* 22, 709–721. <https://doi.org/10.1261/rna.054742.115>
- Amrani, N., Ganesan, R., Kervestin, S., Mangus, D.A., Ghosh, S., Jacobson, A., 2004. A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. *Nature* 432, 112–118. <https://doi.org/10.1038/nature03060>
- Anderson, J.T., Wilson, S.M., Datar, K.V., Swanson, M.S., 1993. NAB2: a yeast nuclear polyadenylated RNA-binding protein essential for cell viability. *Mol Cell Biol* 13, 2730–2741. <https://doi.org/10.1128/mcb.13.5.2730-2741.1993>
- Badis, G., Saveanu, C., Fromont-Racine, M., Jacquier, A., 2004. Targeted mRNA Degradation by Deadenylation-Independent Decapping. *Molecular Cell* 15, 5–15. <https://doi.org/10.1016/j.molcel.2004.06.028>
- Bae, H., Collier, J., 2022. Codon optimality-mediated mRNA degradation: Linking translational elongation to mRNA stability. *Molecular Cell* 82, 1467–1476. <https://doi.org/10.1016/j.molcel.2022.03.032>
- Baer, B.W., Kornberg, R.D., 1983. The protein responsible for the repeating structure of cytoplasmic poly(A)-ribonucleoprotein. *Journal of Cell Biology* 96, 717–721. <https://doi.org/10.1083/jcb.96.3.717>
- Baer, B.W., Kornberg, R.D., 1980. Repeating structure of cytoplasmic poly(A)-ribonucleoprotein. *Proc. Natl. Acad. Sci. U.S.A.* 77, 1890–1892. <https://doi.org/10.1073/pnas.77.4.1890>
- Bai, Y., Salvadore, C., Chiang, Y.-C., Collart, M.A., Liu, H.-Y., Denis, C.L., 1999. The CCR4 and CAF1 Proteins of the CCR4-NOT Complex Are Physically and Functionally Separated from NOT2, NOT4, and NOT5. *Mol Cell Biol* 19, 6642–6651. <https://doi.org/10.1128/MCB.19.10.6642>
- Baker Brachmann, C., Davies, A., Cost, G.J., Caputo, E., Li, J., Hieter, P., Boeke, J.D., 1998. Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* 14, 115–132. [https://doi.org/10.1002/\(SICI\)1097-0061\(19980130\)14:2<115::AID-YEA204>3.0.CO;2-2](https://doi.org/10.1002/(SICI)1097-0061(19980130)14:2<115::AID-YEA204>3.0.CO;2-2)
- Basquin, J., Roudko, V.V., Rode, M., Basquin, C., Séraphin, B., Conti, E., 2012. Architecture of the Nuclease Module of the Yeast Ccr4-Not Complex: the Not1-Caf1-Ccr4 Interaction. *Molecular Cell* 48, 207–218. <https://doi.org/10.1016/j.molcel.2012.08.014>
- Baudrimont, A., Voegeli, S., Vilorio, E.C., Stritt, F., Lenon, M., Wada, T., Jaquet, V., Becskei, A., 2017. Multiplexed gene control reveals rapid mRNA turnover. *Sci. Adv.* 3, e1700006. <https://doi.org/10.1126/sciadv.1700006>
- Bawankar, P., Loh, B., Wohlbold, L., Schmidt, S., Izaurralde, E., 2013. NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain. *RNA Biology* 10, 228–244. <https://doi.org/10.4161/rna.23018>

- Beelman, C.A., Stevens, A., Caponigro, G., LaGrandeur, T.E., Hatfield, L., Fortner, D.M., Parker, R., 1996. An essential component of the decapping enzyme required for normal rates of mRNA turnover. *Nature* 382, 642–646. <https://doi.org/10.1038/382642a0>
- Behm-Ansmant, I., Gatfield, D., Rehwinkel, J., Hilgers, V., Izaurralde, E., 2007. A conserved role for cytoplasmic poly(A)-binding protein 1 (PABPC1) in nonsense-mediated mRNA decay. *EMBO J* 26, 1591–1601. <https://doi.org/10.1038/sj.emboj.7601588>
- Beilharz, T.H., Preiss, T., 2007. Widespread use of poly(A) tail length control to accentuate expression of the yeast transcriptome. *RNA* 13, 982–997. <https://doi.org/10.1261/rna.569407>
- Boeck, R., Lapeyre, B., Brown, C.E., Sachs, A.B., 1998. Capped mRNA Degradation Intermediates Accumulate in the Yeast *spb8-2* Mutant. *Mol Cell Biol* 18, 5062–5072.
- Boeck, R., Tarun, S., Rieger, M., Deardorff, J.A., Müller-Auer, S., Sachs, A.B., 1996. The Yeast Pan2 Protein Is Required for Poly(A)-binding Protein-stimulated Poly(A)-nuclease Activity. *Journal of Biological Chemistry* 271, 432–438. <https://doi.org/10.1074/jbc.271.1.432>
- Boland, A., Chen, Y., Raisch, T., Jonas, S., Kuzuoğlu-Öztürk, D., Wohlbold, L., Weichenrieder, O., Izaurralde, E., 2013. Structure and assembly of the NOT module of the human CCR4–NOT complex. *Nat Struct Mol Biol* 20, 1289–1297. <https://doi.org/10.1038/nsmb.2681>
- Bouveret, E., 2000. A Sm-like protein complex that participates in mRNA degradation. *The EMBO Journal* 19, 1661–1671. <https://doi.org/10.1093/emboj/19.7.1661>
- Brambilla, M., Martani, F., Bertacchi, S., Vitangeli, I., Branduardi, P., 2019. The *Saccharomyces cerevisiae* poly (A) binding protein (Pab1): Master regulator of mRNA metabolism and cell physiology. *Yeast* 36, 23–34. <https://doi.org/10.1002/yea.3347>
- Brown, C.E., Sachs, A.B., 1998. Poly(A) Tail Length Control in *Saccharomyces cerevisiae* Occurs by Message-Specific Deadenylation. *Mol Cell Biol* 18, 6548–6559. <https://doi.org/10.1128/MCB.18.11.6548>
- Brown, C.E., Tarun, S.Z., Boeck, R., Sachs, A.B., 1996. PAN3 encodes a subunit of the Pab1p-dependent poly(A) nuclease in *Saccharomyces cerevisiae*. *Mol Cell Biol* 16, 5744–5753. <https://doi.org/10.1128/MCB.16.10.5744>
- Brune, C., Munchel, S.E., Fischer, N., Podtelejnikov, A.V., Weis, K., 2005. Yeast poly(A)-binding protein Pab1 shuttles between the nucleus and the cytoplasm and functions in mRNA export. *RNA* 11, 517–531. <https://doi.org/10.1261/rna.7291205>
- Burgess, H.M., Richardson, W.A., Anderson, R.C., Salaun, C., Graham, S.V., Gray, N.K., 2011. Nuclear relocalisation of cytoplasmic poly(A)-binding proteins PABP1 and PABP4 in response to UV irradiation reveals mRNA-dependent export of metazoan PABPs. *Journal of Cell Science* 124, 3344–3355. <https://doi.org/10.1242/jcs.087692>
- Buschauer, R., Matsuo, Y., Sugiyama, T., Chen, Y.-H., Alhusaini, N., Sweet, T., Ikeuchi, K., Cheng, J., Matsuki, Y., Nobuta, R., Gilmozzi, A., Berninghausen, O., Tesina, P., Becker, T., Coller, J., Inada, T., Beckmann, R., 2020. The Ccr4-Not complex monitors the translating ribosome for codon optimality. *Science* 368, eaay6912. <https://doi.org/10.1126/science.aay6912>
- Bushkin, G.G., Pincus, D., Morgan, J.T., Richardson, K., Lewis, C., Chan, S.H., Bartel, D.P., Fink, G.R., 2019. m6A modification of a 3' UTR site reduces RME1 mRNA

- levels to promote meiosis. *Nat Commun* 10, 3414. <https://doi.org/10.1038/s41467-019-11232-7>
- Cao, Dan, Parker, R., 2001. Computational modeling of eukaryotic mRNA turnover. *RNA* 7, 1192–1212. <https://doi.org/10.1017/S1355838201010330>
- Cao, D., Parker, R., 2001. Computational modeling of eukaryotic mRNA turnover. *RNA* 7, 1192–1212.
- Celik, A., He, F., Jacobson, A., 2017. NMD monitors translational fidelity 24/7. *Curr Genet* 63, 1007–1010. <https://doi.org/10.1007/s00294-017-0709-4>
- Chalabi Hagkarim, N., Grand, R.J., 2020. The Regulatory Properties of the Ccr4–Not Complex. *Cells* 9, 2379. <https://doi.org/10.3390/cells9112379>
- Chan, L.Y., Mugler, C.F., Heinrich, S., Vallotton, P., Weis, K., 2018. Non-invasive measurement of mRNA decay reveals translation initiation as the major determinant of mRNA stability. *eLife* 7, e32536. <https://doi.org/10.7554/eLife.32536>
- Chang, C.-T., Bercovich, N., Loh, B., Jonas, S., Izaurralde, E., 2014. The activation of the decapping enzyme DCP2 by DCP1 occurs on the EDC4 scaffold and involves a conserved loop in DCP1. *Nucleic Acids Research* 42, 5217–5233. <https://doi.org/10.1093/nar/gku129>
- Chang, H., Lim, J., Ha, M., Kim, V.N., 2014. TAIL-seq: Genome-wide Determination of Poly(A) Tail Length and 3' End Modifications. *Molecular Cell* 53, 1044–1052. <https://doi.org/10.1016/j.molcel.2014.02.007>
- Charenton, C., Gaudon-Plesse, C., Back, R., Ulryck, N., Cosson, L., Séraphin, B., Graille, M., 2020. Pby1 is a direct partner of the Dcp2 decapping enzyme. *Nucleic Acids Research* 48, 6353–6366. <https://doi.org/10.1093/nar/gkaa337>
- Charenton, C., Gaudon-Plesse, C., Fourati, Z., Taverniti, V., Back, R., Kolesnikova, O., Séraphin, B., Graille, M., 2017. A unique surface on Pat1 C-terminal domain directly interacts with Dcp2 decapping enzyme and Xrn1 5'–3' mRNA exonuclease in yeast. *Proc. Natl. Acad. Sci. U.S.A.* 114. <https://doi.org/10.1073/pnas.1711680114>
- Charenton, C., Graille, M., 2018. mRNA decapping: finding the right structures. *Phil. Trans. R. Soc. B* 373, 20180164. <https://doi.org/10.1098/rstb.2018.0164>
- Charenton, C., Taverniti, V., Gaudon-Plesse, C., Back, R., Séraphin, B., Graille, M., 2016. Structure of the active form of Dcp1–Dcp2 decapping enzyme bound to m7GDP and its Edc3 activator. *Nat Struct Mol Biol* 23, 982–986. <https://doi.org/10.1038/nsmb.3300>
- Chen, J., Chiang, Y.-C., Denis, C.L., 2002. CCR4, a 3'-5' poly(A) RNA and ssDNA exonuclease, is the catalytic component of the cytoplasmic deadenylase. *The EMBO Journal* 21, 1414–1426. <https://doi.org/10.1093/emboj/21.6.1414>
- Chen, J., Rappsilber, J., Chiang, Y.-C., Russell, P., Mann, M., Denis, C.L., 2001. Purification and characterization of the 1.0 MDa CCR4-NOT complex identifies two novel components of the complex 1 Edited by D. Draper. *Journal of Molecular Biology* 314, 683–694. <https://doi.org/10.1006/jmbi.2001.5162>
- Chen, Y., Khazina, E., Izaurralde, E., Weichenrieder, O., 2021. Crystal structure and functional properties of the human CCR4-CAF1 deadenylase complex. *Nucleic Acids Research* 49, 6489–6510. <https://doi.org/10.1093/nar/gkab414>
- Choi, K., Medley, J.K., König, M., Stocking, K., Smith, L., Gu, S., Sauro, H.M., 2018. Tellurium: An extensible python-based modeling environment for systems and synthetic biology. *Biosystems* 171, 74–79. <https://doi.org/10.1016/j.biosystems.2018.07.006>



- Chowdhury, A., Kalurupalle, S., Tharun, S., 2014. Pat1 contributes to the RNA binding activity of the Lsm1-7–Pat1 complex. *RNA* 20, 1465–1475. <https://doi.org/10.1261/rna.045252.114>
- Chowdhury, A., Mukhopadhyay, J., Tharun, S., 2007. The decapping activator Lsm1p-7p–Pat1p complex has the intrinsic ability to distinguish between oligoadenylated and polyadenylated RNAs. *RNA* 13, 998–1016. <https://doi.org/10.1261/rna.502507>
- Collart, M.A., 2016. The Ccr4-Not complex is a key regulator of eukaryotic gene expression. *WIREs RNA* 7, 438–454. <https://doi.org/10.1002/wrna.1332>
- Collart, M.A., Oliviero, S., 1993. Preparation of Yeast RNA. *Current Protocols in Molecular Biology* 23. <https://doi.org/10.1002/0471142727.mb1312s23>
- Collart, M.A., Panasenko, O.O., Nikolaev, S.I., 2013. The Not3/5 subunit of the Ccr4-Not complex: A central regulator of gene expression that integrates signals between the cytoplasm and the nucleus in eukaryotic cells. *Cellular Signalling* 25, 743–751. <https://doi.org/10.1016/j.cellsig.2012.12.018>
- Collart, M.A., Struhl, K., 1994. NOT1(CDC39), NOT2(CDC36), NOT3, and NOT4 encode a global-negative regulator of transcription that differentially affects TATA-element utilization. *Genes Dev.* 8, 525–537. <https://doi.org/10.1101/gad.8.5.525>
- Coller, J., Parker, R., 2005. General Translational Repression by Activators of mRNA Decapping. *Cell* 122, 875–886. <https://doi.org/10.1016/j.cell.2005.07.012>
- Coller, J.M., Tucker, M., Sheth, U., Valencia-Sanchez, M.A., Parker, R., 2001. The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes. *RNA* 7, 1717–1727. <https://doi.org/10.1017/S135583820101994X>
- Cosson, B., Couturier, A., Chabelskaya, S., Kiktev, D., Inge-Vechtomov, S., Philippe, M., Zhouravleva, G., 2002. Poly(A)-Binding Protein Acts in Translation Termination via Eukaryotic Release Factor 3 Interaction and Does Not Influence [ *PSI*<sup>+</sup> ] Propagation. *Mol Cell Biol* 22, 3301–3315. <https://doi.org/10.1128/MCB.22.10.3301-3315.2002>
- Costello, J., Castelli, L.M., Rowe, W., Kershaw, C.J., Talavera, D., Mohammad-Qureshi, S.S., Sims, P.F.G., Grant, C.M., Pavitt, G.D., Hubbard, S.J., Ashe, M.P., 2015. Global mRNA selection mechanisms for translation initiation. *Genome Biol* 16, 10. <https://doi.org/10.1186/s13059-014-0559-z>
- Couttet, P., Fromont-Racine, M., Steel, D., Pictet, R., Grange, T., 1997. Messenger RNA deadenylation precedes decapping in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 94, 5628–5633. <https://doi.org/10.1073/pnas.94.11.5628>
- Cui, Y., Hagan, K.W., Zhang, S., Peltz, S.W., 1995. Identification and characterization of genes that are required for the accelerated degradation of mRNAs containing a premature translational termination codon. *Genes Dev.* 9, 423–436. <https://doi.org/10.1101/gad.9.4.423>
- Culbertson, M.R., Underbrink, K.M., Fink, G.R., 1980. FRAMESHIFT SUPPRESSION IN *SACCHAROMYCES CEREVISIAE* . II. GENETIC PROPERTIES OF GROUP II SUPPRESSORS. *Genetics* 95, 833–853. <https://doi.org/10.1093/genetics/95.4.833>
- Czaplinski, K., Ruiz-Echevarria, M.J., Paushkin, S.V., Han, X., Weng, Y., Perlick, H.A., Dietz, H.C., Ter-Avanesyan, M.D., Peltz, S.W., 1998. The surveillance complex interacts with the translation release factors to enhance termination and degrade aberrant mRNAs. *Genes Dev.* 12, 1665–1677. <https://doi.org/10.1101/gad.12.11.1665>

- Daugeron, M.-C., 2001. The yeast POP2 gene encodes a nuclease involved in mRNA deadenylation. *Nucleic Acids Research* 29, 2448–2455. <https://doi.org/10.1093/nar/29.12.2448>
- Davis, C.A., Ares, M., 2006. Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 103, 3262–3267. <https://doi.org/10.1073/pnas.0507783103>
- Decourty, L., Saveanu, C., Zemam, K., Hantraye, F., Frachon, E., Rousselle, J.-C., Fromont-Racine, M., Jacquier, A., 2008. Linking functionally related genes by sensitive and quantitative characterization of genetic interaction profiles. *Proc. Natl. Acad. Sci. U.S.A.* 105, 5821–5826. <https://doi.org/10.1073/pnas.0710533105>
- Dehecq, M., Decourty, L., Namane, A., Proux, C., Kanaan, J., Le Hir, H., Jacquier, A., Saveanu, C., 2018. Detection and Degradation of Nonsense-mediated mRNA Decay Substrates Involve Two Distinct Upf1-bound Complexes. <https://doi.org/10.1101/266833>
- Denis, C.L., 1984. IDENTIFICATION OF NEW GENES INVOLVED IN THE REGULATION OF YEAST ALCOHOL DEHYDROGENASE II. *Genetics* 108, 833–844. <https://doi.org/10.1093/genetics/108.4.833>
- Denis, C.L., Malvar, T., 1990. The CCR4 gene from *Saccharomyces cerevisiae* is required for both nonfermentative and spt-mediated gene expression. *Genetics* 124, 283–291. <https://doi.org/10.1093/genetics/124.2.283>
- Dever, T.E., Dinman, J.D., Green, R., 2018. Translation Elongation and Recoding in Eukaryotes. *Cold Spring Harb Perspect Biol* 10, a032649. <https://doi.org/10.1101/cshperspect.a032649>
- Dever, T.E., Green, R., 2012. The Elongation, Termination, and Recycling Phases of Translation in Eukaryotes. *Cold Spring Harbor Perspectives in Biology* 4, a013706–a013706. <https://doi.org/10.1101/cshperspect.a013706>
- Dever, T.E., Kinzy, T.G., Pavitt, G.D., 2016. Mechanism and Regulation of Protein Synthesis in *Saccharomyces cerevisiae*. *Genetics* 203, 65–107. <https://doi.org/10.1534/genetics.115.186221>
- D’Orazio, K.N., Green, R., 2021. Ribosome states signal RNA quality control. *Molecular Cell* 81, 1372–1383. <https://doi.org/10.1016/j.molcel.2021.02.022>
- D’Orazio, K.N., Wu, C.C.-C., Sinha, N., Loll-Krippelber, R., Brown, G.W., Green, R., 2019. The endonuclease Cue2 cleaves mRNAs at stalled ribosomes during No Go Decay. *eLife* 8, e49117. <https://doi.org/10.7554/eLife.49117>
- Duncan, C.D.S., Mata, J., 2011. Widespread Cotranslational Formation of Protein Complexes. *PLoS Genet* 7, e1002398. <https://doi.org/10.1371/journal.pgen.1002398>
- Dunckley, T., 1999. The DCP2 protein is required for mRNA decapping in *Saccharomyces cerevisiae* and contains a functional MutT motif. *The EMBO Journal* 18, 5411–5422. <https://doi.org/10.1093/emboj/18.19.5411>
- Dunckley, T., Tucker, M., Parker, R., 2001. Two Related Proteins, Edc1p and Edc2p, Stimulate mRNA Decapping in *Saccharomyces cerevisiae*. *Genetics* 157, 27–37. <https://doi.org/10.1093/genetics/157.1.27>
- Eberle, A.B., Stalder, L., Mathys, H., Orozco, R.Z., Mühlemann, O., 2008. Posttranscriptional Gene Regulation by Spatial Rearrangement of the 3’ Untranslated Region. *PLoS Biol* 6, e92. <https://doi.org/10.1371/journal.pbio.0060092>

- Edmonds, M., Abrams, R., 1960. Polynucleotide Biosynthesis: Formation of a Sequence of Adenylate Units from Adenosine Triphosphate by an Enzyme from Thymus Nuclei. *Journal of Biological Chemistry* 235, 1142–1149. [https://doi.org/10.1016/S0021-9258\(18\)69494-3](https://doi.org/10.1016/S0021-9258(18)69494-3)
- Edmonds, M., Vaughan, M.H., Nakazato, H., 1971. Polyadenylic Acid Sequences in the Heterogeneous Nuclear RNA and Rapidly-Labeled Polyribosomal RNA of HeLa Cells: Possible Evidence for a Precursor Relationship. *Proc. Natl. Acad. Sci. U.S.A.* 68, 1336–1340. <https://doi.org/10.1073/pnas.68.6.1336>
- Eisen, T.J., Eichhorn, S.W., Subtelny, A.O., Lin, K.S., McGeary, S.E., Gupta, S., Bartel, D.P., 2020. The Dynamics of Cytoplasmic mRNA Metabolism. *Molecular Cell* 77, 786–799.e10. <https://doi.org/10.1016/j.molcel.2019.12.005>
- Fischer, N., 2002. The DEAD box protein Dhh1 stimulates the decapping enzyme Dcp1. *The EMBO Journal* 21, 2788–2797. <https://doi.org/10.1093/emboj/21.11.2788>
- Fortes, P., Inada, T., Preiss, T., Hentze, M.W., Mattaj, I.W., Sachs, A.B., 2000. The Yeast Nuclear Cap Binding Complex Can Interact with Translation Factor eIF4G and Mediate Translation Initiation. *Molecular Cell* 6, 191–196. [https://doi.org/10.1016/S1097-2765\(05\)00003-1](https://doi.org/10.1016/S1097-2765(05)00003-1)
- Fromont-Racine, M., Mayes, A.E., Brunet-Simon, A., Rain, J.-C., Colley, A., Dix, I., Decourty, L., Joly, N., Ricard, F., Beggs, J.D., Legrain, P., 2000. Genome-Wide Protein Interaction Screens Reveal Functional Networks Involving Sm-Like Proteins. *Yeast* 1, 95–110. [https://doi.org/10.1002/1097-0061\(20000630\)17:2<95::AID-YEA16>3.0.CO;2-H](https://doi.org/10.1002/1097-0061(20000630)17:2<95::AID-YEA16>3.0.CO;2-H)
- Gallie, D.R., 1998. A tale of two termini: *Gene* 216, 1–11. [https://doi.org/10.1016/S0378-1119\(98\)00318-7](https://doi.org/10.1016/S0378-1119(98)00318-7)
- Gamble, C.E., Brule, C.E., Dean, K.M., Fields, S., Grayhack, E.J., 2016. Adjacent Codons Act in Concert to Modulate Translation Efficiency in Yeast. *Cell* 166, 679–690. <https://doi.org/10.1016/j.cell.2016.05.070>
- Garí, E., Piedrafita, L., Aldea, M., Herrero, E., 1997. A Set of Vectors with a Tetracycline-Regulatable Promoter System for Modulated Gene Expression in *Saccharomyces cerevisiae*. *Yeast* 13, 837–848. [https://doi.org/10.1002/\(SICI\)1097-0061\(199707\)13:9<837::AID-YEA145>3.0.CO;2-T](https://doi.org/10.1002/(SICI)1097-0061(199707)13:9<837::AID-YEA145>3.0.CO;2-T)
- Garneau, N.L., Wilusz, J., Wilusz, C.J., 2007. The highways and byways of mRNA decay. *Nat Rev Mol Cell Biol* 8, 113–126. <https://doi.org/10.1038/nrm2104>
- Gaudon, C., 1999. Role of the essential yeast protein PSU1 in transcriptional enhancement by the ligand-dependent activation function AF-2 of nuclear receptors. *The EMBO Journal* 18, 2229–2240. <https://doi.org/10.1093/emboj/18.8.2229>
- Gavin, A.-C., Bösch, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.-M., Cruciat, C.-M., Remor, M., Höfert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M.-A., Copley, R.R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G., Superti-Furga, G., 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415, 141–147. <https://doi.org/10.1038/415141a>
- Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O’Shea, E.K., Weissman, J.S., 2003. Global analysis of protein expression in yeast. *Nature* 425, 737–741. <https://doi.org/10.1038/nature02046>

- Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., Arkin, A.P., Astromoff, A., El Bakkoury, M., Bangham, R., Benito, R., Brachet, S., Campanaro, S., Curtiss, M., Davis, K., Deutschbauer, A., Entian, K.-D., Flaherty, P., Foury, F., Garfinkel, D.J., Gerstein, M., Gotte, D., Güldener, U., Hegemann, J.H., Hempel, S., Herman, Z., Jaramillo, D.F., Kelly, D.E., Kelly, S.L., Kötter, P., LaBonte, D., Lamb, D.C., Lan, N., Liang, H., Liao, H., Liu, L., Luo, C., Lussier, M., Mao, R., Menard, P., Ooi, S.L., Revuelta, J.L., Roberts, C.J., Rose, M., Ross-Macdonald, P., Scherens, B., Schimmack, G., Shafer, B., Shoemaker, D.D., Sookhai-Mahadeo, S., Storms, R.K., Strathern, J.N., Valle, G., Voet, M., Volckaert, G., Wang, C., Ward, T.R., Wilhelmy, J., Winzeler, E.A., Yang, Y., Yen, G., Youngman, E., Yu, K., Bussey, H., Boeke, J.D., Snyder, M., Philippsen, P., Davis, R.W., Johnston, M., 2002. Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418, 387–391. <https://doi.org/10.1038/nature00935>
- Gilbertson, S., Federspiel, J.D., Hartenian, E., Cristea, I.M., Glaunsinger, B., 2018. Changes in mRNA abundance drive shuttling of RNA binding proteins, linking cytoplasmic RNA degradation to transcription. *eLife* 7, e37663. <https://doi.org/10.7554/eLife.37663>
- Goldstrohm, A.C., Hook, B.A., Seay, D.J., Wickens, M., 2006. PUF proteins bind Pop2p to regulate messenger RNAs. *Nat Struct Mol Biol* 13, 533–539. <https://doi.org/10.1038/nsmb1100>
- Halbeisen, R.E., Scherrer, T., Gerber, A.P., 2009. Affinity purification of ribosomes to access the translome. *Methods, Global approaches to study gene regulation* 48, 306–310. <https://doi.org/10.1016/j.ymeth.2009.04.003>
- Hanson, G., Coller, J., 2018. Codon optimality, bias and usage in translation and mRNA decay. *Nat Rev Mol Cell Biol* 19, 20–30. <https://doi.org/10.1038/nrm.2017.91>
- Hanzawa, H., de Ruwe, M.J., Albert, T.K., van der Vliet, P.C., Timmers, H.T.M., Boelens, R., 2001. The Structure of the C4C4RING Finger of Human NOT4 Reveals Features Distinct from Those of C3HC4 RING Fingers. *Journal of Biological Chemistry* 276, 10185–10190. <https://doi.org/10.1074/jbc.M009298200>
- He, F., Brown, A.H., Jacobson, A., 1996. Interaction between Nmd2p and Upf1p is required for activity but not for dominant-negative inhibition of the nonsense-mediated mRNA decay pathway in yeast. *RNA* 2, 153–170.
- He, F., Jacobson, A., 2015. Control of mRNA decapping by positive and negative regulatory elements in the Dcp2 C-terminal domain. *RNA* 21, 1633–1647. <https://doi.org/10.1261/rna.052449.115>
- He, F., Jacobson, A., 2001. Upf1p, Nmd2p, and Upf3p Regulate the Decapping and Exonucleolytic Degradation of both Nonsense-Containing mRNAs and Wild-Type mRNAs. *Mol Cell Biol* 21, 1515–1530. <https://doi.org/10.1128/MCB.21.5.1515-1530.2001>
- He, F., Jacobson, A., 1995. Identification of a novel component of the nonsense-mediated mRNA decay pathway by use of an interacting protein screen. *Genes Dev.* 9, 437–454. <https://doi.org/10.1101/gad.9.4.437>
- He, F., Li, C., Roy, B., Jacobson, A., 2014. Yeast Edc3 Targets *RPS28B* mRNA for Decapping by Binding to a 3' Untranslated Region Decay-Inducing Regulatory Element. *Mol Cell Biol* 34, 1438–1451. <https://doi.org/10.1128/MCB.01584-13>
- He, W., Parker, R., 2001. The Yeast Cytoplasmic Lsm1/Pat1p Complex Protects mRNA 3' Termini From Partial Degradation. *Genetics* 158, 1445–1455. <https://doi.org/10.1093/genetics/158.4.1445>

- Herrick, D., Parker, R., Jacobson, A., 1990. Identification and comparison of stable and unstable mRNAs in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 10, 2269–2284. <https://doi.org/10.1128/MCB.10.5.2269>
- Hoekema, A., Kastelein, R.A., Vasser, M., de Boer, H.A., 1987. Codon replacement in the PGK1 gene of *Saccharomyces cerevisiae*: experimental approach to study the role of biased codon usage in gene expression. *Mol Cell Biol* 7, 2914–2924. <https://doi.org/10.1128/mcb.7.8.2914-2924.1987>
- Hogg, J.R., Goff, S.P., 2010. Upf1 Senses 3'UTR Length to Potentiate mRNA Decay. *Cell* 143, 379–389. <https://doi.org/10.1016/j.cell.2010.10.005>
- Hoshino, S., Imai, M., Kobayashi, T., Uchida, N., Katada, T., 1999. The Eukaryotic Polypeptide Chain Releasing Factor (eRF3/GSPT) Carrying the Translation Termination Signal to the 3'-Poly(A) Tail of mRNA. *Journal of Biological Chemistry* 274, 16677–16680. <https://doi.org/10.1074/jbc.274.24.16677>
- Houalla, R., Devaux, F., Fatica, A., Kufel, J., Barrass, D., Torchet, C., Tollervey, D., 2006. Microarray detection of novel nuclear RNA substrates for the exosome. *Yeast* 23, 439–454. <https://doi.org/10.1002/yea.1369>
- Hu, W., Sweet, T.J., Chamnongpol, S., Baker, K.E., Collier, J., 2009. Co-translational mRNA decay in *Saccharomyces cerevisiae*. *Nature* 461, 225–229. <https://doi.org/10.1038/nature08265>
- Hwang, S.S., Lim, J., Yu, Z., Kong, P., Sefik, E., Xu, H., Harman, C.C.D., Kim, L.K., Lee, G.R., Li, H.-B., Flavell, R.A., 2020. mRNA destabilization by BTG1 and BTG2 maintains T cell quiescence. *Science* 367, 1255–1260. <https://doi.org/10.1126/science.aax0194>
- Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., Weissman, J.S., 2009. Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling. *Science* 324, 218–223. <https://doi.org/10.1126/science.1168978>
- Ivanov, A., Mikhailova, T., Eliseev, B., Yeramala, L., Sokolova, E., Susorov, D., Shuvalov, A., Schaffitzel, C., Alkalaeva, E., 2016. PABP enhances release factor recruitment and stop codon recognition during translation termination. *Nucleic Acids Res* 44, 7766–7776. <https://doi.org/10.1093/nar/gkw635>
- Ivanov, P.V., Gehring, N.H., Kunz, J.B., Hentze, M.W., Kulozik, A.E., 2008. Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. *EMBO J* 27, 736–747. <https://doi.org/10.1038/emboj.2008.17>
- Jänicke, A., Vancuylenberg, J., Boag, P.R., Traven, A., Beilharz, T.H., 2012. ePAT: A simple method to tag adenylated RNA to measure poly(A)-tail length and other 3' RACE applications. *RNA* 18, 1289–1295. <https://doi.org/10.1261/rna.031898.111>
- Jinek, M., Coyle, S.M., Doudna, J.A., 2011. Coupled 5' Nucleotide Recognition and Processivity in Xrn1-Mediated mRNA Decay. *Molecular Cell* 41, 600–608. <https://doi.org/10.1016/j.molcel.2011.02.004>
- Jonas, S., Christie, M., Peter, D., Bhandari, D., Loh, B., Huntzinger, E., Weichenrieder, O., Izaurralde, E., 2014. An asymmetric PAN3 dimer recruits a single PAN2 exonuclease to mediate mRNA deadenylation and decay. *Nat Struct Mol Biol* 21, 599–608. <https://doi.org/10.1038/nsmb.2837>
- Kajjo, S., Sharma, S., Chen, S., Brothers, W.R., Cott, M., Hasaj, B., Jovanovic, P., Larsson, O., Fabian, M.R., 2022. PABP prevents the untimely decay of select mRNA populations in human cells. *The EMBO Journal* 41. <https://doi.org/10.15252/embj.2021108650>

- Kervestin, S., Li, C., Buckingham, R., Jacobson, A., 2012. Testing the faux-UTR model for NMD: Analysis of Upf1p and Pab1p competition for binding to eRF3/Sup35p. *Biochimie* 94, 1560–1571. <https://doi.org/10.1016/j.biochi.2011.12.021>
- Kshirsagar, M., Parker, R., 2004. Identification of Edc3p as an Enhancer of mRNA Decapping in *Saccharomyces cerevisiae*. *Genetics* 166, 729–739. <https://doi.org/10.1534/genetics.166.2.729>
- Kushnirov, V.V., 2000. Rapid and reliable protein extraction from yeast. *Yeast* 16, 857–860. [https://doi.org/10.1002/1097-0061\(20000630\)16:9<857::AID-YEA561>3.0.CO;2-B](https://doi.org/10.1002/1097-0061(20000630)16:9<857::AID-YEA561>3.0.CO;2-B)
- Łabno, A., Tomecki, R., Dziembowski, A., 2016. Cytoplasmic RNA decay pathways - Enzymes and mechanisms. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1863, 3125–3147. <https://doi.org/10.1016/j.bbamcr.2016.09.023>
- Lau, N.-C., Kolkman, A., van Schaik, F.M.A., Mulder, K.W., Pijnappel, W.W.M.P., Heck, A.J.R., Timmers, H.Th.M., 2009. Human Ccr4–Not complexes contain variable deadenylase subunits. *Biochemical Journal* 422, 443–453. <https://doi.org/10.1042/BJ20090500>
- Lavysh, D., Neu-Yilik, G., 2020. UPF1-Mediated RNA Decay—Danse Macabre in a Cloud. *Biomolecules* 10, 999. <https://doi.org/10.3390/biom10070999>
- Leeds, P., Peltz, S.W., Jacobson, A., Culbertson, M.R., 1991. The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon. *Genes & Development* 5, 2303–2314. <https://doi.org/10.1101/gad.5.12a.2303>
- Leeds, P., Wood, J.M., Lee, B.S., Culbertson, M.R., 1992. Gene products that promote mRNA turnover in *Saccharomyces cerevisiae*. *Mol Cell Biol* 12, 2165–2177. <https://doi.org/10.1128/mcb.12.5.2165-2177.1992>
- Li, X., Zhao, X., Fang, Y., Jiang, X., Duong, T., Fan, C., Huang, C.-C., Kain, S.R., 1998. Generation of Destabilized Green Fluorescent Protein as a Transcription Reporter. *Journal of Biological Chemistry* 273, 34970–34975. <https://doi.org/10.1074/jbc.273.52.34970>
- Lim, J., Kim, D., Lee, Y., Ha, M., Lee, M., Yeo, J., Chang, H., Song, J., Ahn, K., Kim, V.N., 2018. Mixed tailing by TENT4A and TENT4B shields mRNA from rapid deadenylation. *Science* 361, 701–704. <https://doi.org/10.1126/science.aam5794>
- Lim, J., Lee, M., Son, A., Chang, H., Kim, V.N., 2016. mTAIL-seq reveals dynamic poly(A) tail regulation in oocyte-to-embryo development. *Genes Dev.* 30, 1671–1682. <https://doi.org/10.1101/gad.284802.116>
- Lima, S.A., Chipman, L.B., Nicholson, A.L., Chen, Y.-H., Yee, B.A., Yeo, G.W., Collier, J., Pasquinelli, A.E., 2017. Short poly(A) tails are a conserved feature of highly expressed genes. *Nat Struct Mol Biol* 24, 1057–1063. <https://doi.org/10.1038/nsmb.3499>
- Lobel, J.H., Tibble, R.W., Gross, J.D., 2019. Pat1 activates late steps in mRNA decay by multiple mechanisms. *Proc. Natl. Acad. Sci. U.S.A.* 116, 23512–23517. <https://doi.org/10.1073/pnas.1905455116>
- Losson, R., Lacroute, F., 1979. Interference of nonsense mutations with eukaryotic messenger RNA stability. *Proceedings of the National Academy of Sciences* 76, 5134–5137. <https://doi.org/10.1073/pnas.76.10.5134>
- Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550. <https://doi.org/10.1186/s13059-014-0550-8>

- Lykke-Andersen, S., Jensen, T.H., 2015. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nat Rev Mol Cell Biol* 16, 665–677. <https://doi.org/10.1038/nrm4063>
- Maillet, L., Collart, M.A., 2002. Interaction between Not1p, a Component of the Ccr4-Not Complex, a Global Regulator of Transcription, and Dhh1p, a Putative RNA Helicase. *Journal of Biological Chemistry* 277, 2835–2842. <https://doi.org/10.1074/jbc.M107979200>
- Maillet, L., Tu, C., Hong, Y.K., Shuster, E.O., Collart, M.A., 2000. The essential function of not1 lies within the Ccr4-not complex. *Journal of Molecular Biology* 303, 131–143. <https://doi.org/10.1006/jmbi.2000.4131>
- Malabat, C., Feuerbach, F., Ma, L., Saveanu, C., Jacquier, A., 2015. Quality control of transcription start site selection by nonsense-mediated-mRNA decay. *eLife* 4. <https://doi.org/10.7554/eLife.06722>
- Mans, R., van Rossum, H.M., Wijsman, M., Backx, A., Kuijpers, N.G.A., van den Broek, M., Daran-Lapujade, P., Pronk, J.T., van Maris, A.J.A., Daran, J.-M.G., 2015. CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. *FEMS Yeast Research* 15, fov004. <https://doi.org/10.1093/femsyr/fov004>
- Mareuil, F., Doppelt-Azeroual, O., Ménager, H., 2017. A public Galaxy platform at Pasteur used as an execution engine for web services. <https://doi.org/10.7490/F1000RESEARCH.1114334.1>
- Mauer, J., Luo, X., Blanjoie, A., Jiao, X., Grozhik, A.V., Patil, D.P., Linder, B., Pickering, B.F., Vasseur, J.-J., Chen, Q., Gross, S.S., Elemento, O., Debart, F., Kiledjian, M., Jaffrey, S.R., 2017. Reversible methylation of m6Am in the 5' cap controls mRNA stability. *Nature* 541, 371–375. <https://doi.org/10.1038/nature21022>
- Mauxion, F., Chen, C.-Y.A., Séraphin, B., Shyu, A.-B., 2009. BTG/TOB factors impact deadenylases. *Trends in Biochemical Sciences* 34, 640–647. <https://doi.org/10.1016/j.tibs.2009.07.008>
- Mayes, A.E., 1999. Characterization of Sm-like proteins in yeast and their association with U6 snRNA. *The EMBO Journal* 18, 4321–4331. <https://doi.org/10.1093/emboj/18.15.4321>
- McLaughlin, C.S., Warner, J.R., Edmonds, M., Nakazato, H., Vaughan, M.H., 1973. Polyadenylic Acid Sequences in Yeast Messenger Ribonucleic Acid. *Journal of Biological Chemistry* 248, 1466–1471. [https://doi.org/10.1016/S0021-9258\(19\)44322-6](https://doi.org/10.1016/S0021-9258(19)44322-6)
- Meaux, S., van Hoof, A., Baker, K.E., 2008. Nonsense-Mediated mRNA Decay in Yeast Does Not Require PAB1 or a Poly(A) Tail. *Molecular Cell* 29, 134–140. <https://doi.org/10.1016/j.molcel.2007.10.031>
- Mendoza-Ochoa, G.I., Barrass, J.D., Terlouw, B.R., Maudlin, I.E., de Lucas, S., Sani, E., Aslanzadeh, V., Reid, J.A.E., Beggs, J.D., 2019. A fast and tuneable auxin-inducible degron for depletion of target proteins in budding yeast. *Yeast* 36, 75–81. <https://doi.org/10.1002/yea.3362>
- Miller, C., Schwalb, B., Maier, K., Schulz, D., Dümcke, S., Zacher, B., Mayer, A., Sydow, J., Marcinowski, L., Dölken, L., Martin, D.E., Tresch, A., Cramer, P., 2011. Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast. *Mol Syst Biol* 7, 458. <https://doi.org/10.1038/msb.2010.112>
- Mitchell, S.F., Jain, S., She, M., Parker, R., 2013. Global analysis of yeast mRNPs. *Nat Struct Mol Biol* 20, 127–133. <https://doi.org/10.1038/nsmb.2468>



- Moqtaderi, Z., Geisberg, J.V., Struhl, K., 2018. Extensive Structural Differences of Closely Related 3' mRNA Isoforms: Links to Pab1 Binding and mRNA Stability. *Molecular Cell* 0. <https://doi.org/10.1016/j.molcel.2018.08.044>
- Mugridge, J.S., Tibble, R.W., Ziemniak, M., Jemielity, J., Gross, J.D., 2018. Structure of the activated Edc1-Dcp1-Dcp2-Edc3 mRNA decapping complex with substrate analog poised for catalysis. *Nat Commun* 9, 1152. <https://doi.org/10.1038/s41467-018-03536-x>
- Muhrad, D., Decker, C.J., Parker, R., 1995. Turnover mechanisms of the stable yeast PGK1 mRNA. *Mol Cell Biol* 15, 2145–2156. <https://doi.org/10.1128/MCB.15.4.2145>
- Muhrad, D., Decker, C.J., Parker, R., 1994. Deadenylation of the unstable mRNA encoded by the yeast MFA2 gene leads to decapping followed by 5'→3' digestion of the transcript. *Genes Dev* 8, 855–866. <https://doi.org/10.1101/gad.8.7.855>
- Muhrad, D., Parker, R., 2005. The yeast EDC1 mRNA undergoes deadenylation-independent decapping stimulated by Not2p, Not4p, and Not5p. *EMBO J* 24, 1033–1045. <https://doi.org/10.1038/sj.emboj.7600560>
- Muhrad, D., Parker, R., 1999. Aberrant mRNAs with extended 3' UTRs are substrates for rapid degradation by mRNA surveillance. *RNA* 5, 1299–1307. <https://doi.org/10.1017/S1355838299990829>
- Muhrad, D., Parker, R., 1992. Mutations affecting stability and deadenylation of the yeast MFA2 transcript. *Genes Dev* 6, 2100–2111. <https://doi.org/10.1101/gad.6.11.2100>
- Neil, H., Malabat, C., d'Aubenton-Carafa, Y., Xu, Z., Steinmetz, L.M., Jacquier, A., 2009. Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature* 457, 1038–1042. <https://doi.org/10.1038/nature07747>
- Oberholzer, U., Collart, M.A., 1998. Characterization of NOT5 that encodes a new component of the Not protein complex. *Gene* 207, 61–69. [https://doi.org/10.1016/S0378-1119\(97\)00605-7](https://doi.org/10.1016/S0378-1119(97)00605-7)
- Panasenko, O.O., Collart, M.A., 2012. Presence of Not5 and ubiquitinated Rps7A in polysome fractions depends upon the Not4 E3 ligase: Not4 influence ribosome ubiquitination and assembly. *Molecular Microbiology* 83, 640–653. <https://doi.org/10.1111/j.1365-2958.2011.07957.x>
- Panasenko, O.O., Collart, M.A., 2011. Not4 E3 Ligase Contributes to Proteasome Assembly and Functional Integrity in Part through Ecm29. *Mol Cell Biol* 31, 1610–1623. <https://doi.org/10.1128/MCB.01210-10>
- Paquette, D.R., Tibble, R.W., Daifuku, T.S., Gross, J.D., 2018. Control of mRNA decapping by autoinhibition. *Nucleic Acids Research* 46, 6318–6329. <https://doi.org/10.1093/nar/gky233>
- Passmore, L.A., Collier, J., 2022. Roles of mRNA poly(A) tails in regulation of eukaryotic gene expression. *Nat Rev Mol Cell Biol* 23, 93–106. <https://doi.org/10.1038/s41580-021-00417-y>
- Pechmann, S., Frydman, J., 2013. Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding. *Nat Struct Mol Biol* 20, 237–243. <https://doi.org/10.1038/nsmb.2466>
- Peck, S.A., Hughes, K.D., Victorino, J.F., Mosley, A.L., 2019. Writing a wrong: Coupled RNA polymerase II transcription and RNA quality control. *WIREs RNA* 10. <https://doi.org/10.1002/wrna.1529>
- Pelechano, V., Wei, W., Steinmetz, L.M., 2015. Widespread Co-translational RNA Decay Reveals Ribosome Dynamics. *Cell* 161, 1400–1412. <https://doi.org/10.1016/j.cell.2015.05.008>

- Petit, A.-P., Wohlbold, L., Bawankar, P., Huntzinger, E., Schmidt, S., Izaurralde, E., Weichenrieder, O., 2012. The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4–NOT deadenylase complex. *Nucleic Acids Research* 40, 11058–11072. <https://doi.org/10.1093/nar/gks883>
- Phatnani, H.P., Greenleaf, A.L., 2006. Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev.* 20, 2922–2936. <https://doi.org/10.1101/gad.1477006>
- Porrua, O., Libri, D., 2015. Transcription termination and the control of the transcriptome: why, where and how to stop. *Nat Rev Mol Cell Biol* 16, 190–202. <https://doi.org/10.1038/nrm3943>
- Pratt, J.M., Petty, J., Riba-Garcia, I., Robertson, D.H.L., Gaskell, S.J., Oliver, S.G., Beynon, R.J., 2002. Dynamics of Protein Turnover, a Missing Dimension in Proteomics. *Molecular & Cellular Proteomics* 1, 579–591. <https://doi.org/10.1074/mcp.M200046-MCP200>
- Presnyak, V., Alhusaini, N., Chen, Y.-H., Martin, S., Morris, N., Kline, N., Olson, S., Weinberg, D., Baker, K.E., Graveley, B.R., Collier, J., 2015. Codon Optimality Is a Major Determinant of mRNA Stability. *Cell* 160, 1111–1124. <https://doi.org/10.1016/j.cell.2015.02.029>
- Radhakrishnan, A., Chen, Y.-H., Martin, S., Alhusaini, N., Green, R., Collier, J., 2016. The DEAD-Box Protein Dhh1p Couples mRNA Decay and Translation by Monitoring Codon Optimality. *Cell* 167, 122–132.e9. <https://doi.org/10.1016/j.cell.2016.08.053>
- Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., Vilo, J., 2019. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research* 47, W191–W198. <https://doi.org/10.1093/nar/gkz369>
- Reis, M. d., 2004. Solving the riddle of codon usage preferences: a test for translational selection. *Nucleic Acids Research* 32, 5036–5044. <https://doi.org/10.1093/nar/gkh834>
- Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., Séraphin, B., 1999. A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* 17, 1030–1032. <https://doi.org/10.1038/13732>
- Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., Mesirov, J.P., 2011. Integrative genomics viewer. *Nat Biotechnol* 29, 24–26. <https://doi.org/10.1038/nbt.1754>
- Roque, S., Cerciat, M., Gaugué, I., Mora, L., Floch, A.G., de Zamaroczy, M., Heurgué-Hamard, V., Kervestin, S., 2015. Interaction between the poly(A)-binding protein Pab1 and the eukaryotic release factor eRF3 regulates translation termination but not mRNA decay in *Saccharomyces cerevisiae*. *RNA* 21, 124–134. <https://doi.org/10.1261/rna.047282.114>
- Sachs, A.B., Davis, R.W., 1989. The poly(A) binding protein is required for poly(A) shortening and 60S ribosomal subunit-dependent translation initiation. *Cell* 58, 857–867. [https://doi.org/10.1016/0092-8674\(89\)90938-0](https://doi.org/10.1016/0092-8674(89)90938-0)
- Sachs, A.B., Davis, R.W., 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. <https://doi.org/10.1128/mcb.7.9.3268-3276.1987>
- Sachs, A.B., Deardorff, J.A., 1992. Translation initiation requires the PAB-dependent poly(A) ribonuclease in yeast. *Cell* 70, 961–973. [https://doi.org/10.1016/0092-8674\(92\)90246-9](https://doi.org/10.1016/0092-8674(92)90246-9)

- Saenger, W., Riecke, J., Suck, D., 1975. A structural model for the polyadenylic acid single helix. *Journal of Molecular Biology* 93, 529–534. [https://doi.org/10.1016/0022-2836\(75\)90244-2](https://doi.org/10.1016/0022-2836(75)90244-2)
- Safaei, N., Kozlov, G., Noronha, A.M., Xie, J., Wilds, C.J., Gehring, K., 2012. Interdomain Allostery Promotes Assembly of the Poly(A) mRNA Complex with PABP and eIF4G. *Molecular Cell* 48, 375–386. <https://doi.org/10.1016/j.molcel.2012.09.001>
- Sainsbury, S., Bernecky, C., Cramer, P., 2015. Structural basis of transcription initiation by RNA polymerase II. *Nat Rev Mol Cell Biol* 16, 129–143. <https://doi.org/10.1038/nrm3952>
- Sakai, A., Chibazakura, T., Shimizu, Y., Hishinuma, F., 1992. Molecular analysis of *POP2* gene, a gene required for glucose-derepression of gene expression in *Saccharomyces cerevisiae*. *Nucl Acids Res* 20, 6227–6233. <https://doi.org/10.1093/nar/20.23.6227>
- Sawazaki, R., Imai, S., Yokogawa, M., Hosoda, N., Hoshino, S., Mio, M., Mio, K., Shimada, I., Osawa, M., 2018. Characterization of the multimeric structure of poly(A)-binding protein on a poly(A) tail. *Sci Rep* 8, 1455. <https://doi.org/10.1038/s41598-018-19659-6>
- Schäfer, I.B., Yamashita, M., Schuller, J.M., Schüssler, S., Reichelt, P., Strauss, M., Conti, E., 2019. Molecular Basis for poly(A) RNP Architecture and Recognition by the Pan2-Pan3 Deadenylase. *Cell* 177, 1619-1631.e21. <https://doi.org/10.1016/j.cell.2019.04.013>
- Serdar, L.D., Whiteside, D.L., Baker, K.E., 2016. ATP hydrolysis by UPF1 is required for efficient translation termination at premature stop codons. *Nat Commun* 7, 14021. <https://doi.org/10.1038/ncomms14021>
- Sharif, H., Ozgur, S., Sharma, K., Basquin, C., Urlaub, H., Conti, E., 2013. Structural analysis of the yeast Dhh1–Pat1 complex reveals how Dhh1 engages Pat1, Edc3 and RNA in mutually exclusive interactions. *Nucleic Acids Research* 41, 8377–8390. <https://doi.org/10.1093/nar/gkt600>
- Sharp, P.M., Li, W.-H., 1987. The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. *Nucl Acids Res* 15, 1281–1295. <https://doi.org/10.1093/nar/15.3.1281>
- She, M., Decker, C.J., Chen, N., Tumati, S., Parker, R., Song, H., 2006. Crystal structure and functional analysis of Dcp2p from *Schizosaccharomyces pombe*. *Nat Struct Mol Biol* 13, 63–70. <https://doi.org/10.1038/nsmb1033>
- She, M., Decker, C.J., Svergun, D.I., Round, A., Chen, N., Muhlrad, D., Parker, R., Song, H., 2008. Structural Basis of Dcp2 Recognition and Activation by Dcp1. *Molecular Cell* 29, 337–349. <https://doi.org/10.1016/j.molcel.2008.01.002>
- Shiber, A., Döring, K., Friedrich, U., Klann, K., Merker, D., Zedan, M., Tippmann, F., Kramer, G., Bukau, B., 2018. Cotranslational assembly of protein complexes in eukaryotes revealed by ribosome profiling. *Nature* 561, 268–272. <https://doi.org/10.1038/s41586-018-0462-y>
- Siddiqui, N., Mangus, D.A., Chang, T.-C., Palermino, J.-M., Shyu, A.-B., Gehring, K., 2007. Poly(A) Nuclease Interacts with the C-terminal Domain of Polyadenylate-binding Protein Domain from Poly(A)-binding Protein. *Journal of Biological Chemistry* 282, 25067–25075. <https://doi.org/10.1074/jbc.M701256200>
- Silva, A.L., Ribeiro, P., Inácio, Â., Liebhaber, S.A., Romão, L., 2008. Proximity of the poly(A)-binding protein to a premature termination codon inhibits mammalian

- nonsense-mediated mRNA decay. *RNA* 14, 563–576.  
<https://doi.org/10.1261/rna.815108>
- Simón, E., Séraphin, B., 2007. A specific role for the C-terminal region of the Poly(A)-binding protein in mRNA decay. *Nucleic Acids Research* 35, 6017–6028.  
<https://doi.org/10.1093/nar/gkm452>
- Singh, G., Rebbapragada, I., Lykke-Andersen, J., 2008. A Competition between Stimulators and Antagonists of Upf Complex Recruitment Governs Human Nonsense-Mediated mRNA Decay. *PLoS Biol* 6, e111.  
<https://doi.org/10.1371/journal.pbio.0060111>
- Smith, L.P., Bergmann, F.T., Chandran, D., Sauro, H.M., 2009. Antimony: a modular model definition language. *Bioinformatics* 25, 2452–2454.  
<https://doi.org/10.1093/bioinformatics/btp401>
- Smith, R.W.P., Blee, T.K.P., Gray, N.K., 2014. Poly(A)-binding proteins are required for diverse biological processes in metazoans. *Biochemical Society Transactions* 42, 1229–1237. <https://doi.org/10.1042/BST20140111>
- Sonneveld, S., Verhagen, B.M.P., Tanenbaum, M.E., 2020. Heterogeneity in mRNA Translation. *Trends in Cell Biology* 30, 606–618.  
<https://doi.org/10.1016/j.tcb.2020.04.008>
- Sørensen, M.A., Pedersen, S., 1991. Absolute in vivo translation rates of individual codons in *Escherichia coli*. *Journal of Molecular Biology* 222, 265–280.  
[https://doi.org/10.1016/0022-2836\(91\)90211-N](https://doi.org/10.1016/0022-2836(91)90211-N)
- Stewart, M., 2019. Polyadenylation and nuclear export of mRNAs. *Journal of Biological Chemistry* 294, 2977–2987. <https://doi.org/10.1074/jbc.REV118.005594>
- Strahl-Bolsinger, S., Tanner, W., 1993. A yeast gene encoding a putative RNA helicase of the “DEAD”-box family. *Yeast* 9, 429–432. <https://doi.org/10.1002/yea.320090414>
- Stupfler, B., Birck, C., Séraphin, B., Mauxion, F., 2016. BTG2 bridges PABPC1 RNA-binding domains and CAF1 deadenylase to control cell proliferation. *Nat Commun* 7, 10811. <https://doi.org/10.1038/ncomms10811>
- Subtelny, A.O., Eichhorn, S.W., Chen, G.R., Sive, H., Bartel, D.P., 2014. Poly(A)-tail profiling reveals an embryonic switch in translational control. *Nature* 508, 66–71.  
<https://doi.org/10.1038/nature13007>
- Sun, M., Schwalb, B., Pirkl, N., Maier, K.C., Schenk, A., Failmezger, H., Tresch, A., Cramer, P., 2013. Global Analysis of Eukaryotic mRNA Degradation Reveals Xrn1-Dependent Buffering of Transcript Levels. *Molecular Cell* 52, 52–62.  
<https://doi.org/10.1016/j.molcel.2013.09.010>
- Sun, M., Schwalb, B., Schulz, D., Pirkl, N., Etzold, S., Larivière, L., Maier, K.C., Seizl, M., Tresch, A., Cramer, P., 2012. Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. *Genome Res.* 22, 1350–1359. <https://doi.org/10.1101/gr.130161.111>
- Sweet, T., Kovalak, C., Collier, J., 2012. The DEAD-Box Protein Dhh1 Promotes Decapping by Slowing Ribosome Movement. *PLoS Biol* 10, e1001342.  
<https://doi.org/10.1371/journal.pbio.1001342>
- Swisher, K.D., Parker, R., 2011. Interactions between Upf1 and the Decapping Factors Edc3 and Pat1 in *Saccharomyces cerevisiae*. *PLoS ONE* 6, e26547.  
<https://doi.org/10.1371/journal.pone.0026547>
- Tang, T.T.L., Passmore, L.A., 2019. Recognition of Poly(A) RNA through Its Intrinsic Helical Structure. *Cold Spring Harb Symp Quant Biol* 84, 21–30.  
<https://doi.org/10.1101/sqb.2019.84.039818>

- Tang, T.T.L., Stowell, J.A.W., Hill, C.H., Passmore, L.A., 2019. The intrinsic structure of poly(A) RNA determines the specificity of Pan2 and Caf1 deadenylases. *Nat Struct Mol Biol* 26, 433–442. <https://doi.org/10.1038/s41594-019-0227-9>
- Tarassov, K., Messier, V., Landry, C.R., Radinovic, S., Molina, M.M.S., Shames, I., Malitskaya, Y., Vogel, J., Bussey, H., Michnick, S.W., 2008. An in Vivo Map of the Yeast Protein Interactome. *Science* 320, 1465–1470. <https://doi.org/10.1126/science.1153878>
- Tarun, S.Z., Sachs, A.B., 1996. Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *The EMBO Journal* 15, 7168–7177. <https://doi.org/10.1002/j.1460-2075.1996.tb01108.x>
- Tesina, P., Heckel, E., Cheng, J., Fromont-Racine, M., Buschauer, R., Kater, L., Beatrix, B., Berninghausen, O., Jacquier, A., Becker, T., Beckmann, R., 2019. Structure of the 80S ribosome–Xrn1 nuclease complex. *Nat Struct Mol Biol* 26, 275–280. <https://doi.org/10.1038/s41594-019-0202-5>
- Tharun, S., He, W., Mayes, A.E., Lennertz, P., Beggs, J.D., Parker, R., 2000. Yeast Sm-like proteins function in mRNA decapping and decay. *Nature* 404, 515–518. <https://doi.org/10.1038/35006676>
- Tharun, S., Muhlrads, D., Chowdhury, A., Parker, R., 2005. Mutations in the *Saccharomyces cerevisiae* LSM1 Gene That Affect mRNA Decapping and 3' End Protection. *Genetics* 170, 33–46. <https://doi.org/10.1534/genetics.104.034322>
- Tharun, S., Parker, R., 2001. Targeting an mRNA for Decapping. *Molecular Cell* 8, 1075–1083. [https://doi.org/10.1016/S1097-2765\(01\)00395-1](https://doi.org/10.1016/S1097-2765(01)00395-1)
- Thore, S., Mauxion, F., Séraphin, B., 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. <https://doi.org/10.1038/sj.embor.7400020>
- Tishinov, K., Spang, A., 2021. The mRNA decapping complex is buffered by nuclear localization. *Journal of Cell Science* 134, jcs259156. <https://doi.org/10.1242/jcs.259156>
- Tucker, M., Staples, R.R., Valencia-Sanchez, M.A., Muhlrads, D., Parker, R., 2002. Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in *Saccharomyces cerevisiae*. *The EMBO Journal* 21, 1427–1436. <https://doi.org/10.1093/emboj/21.6.1427>
- Tucker, M., Valencia-Sanchez, M.A., Staples, R.R., Chen, J., Denis, C.L., Parker, R., 2001. The Transcription Factor Associated Ccr4 and Caf1 Proteins Are Components of the Major Cytoplasmic mRNA Deadenylation Complex in *Saccharomyces cerevisiae*. *Cell* 104, 377–386. [https://doi.org/10.1016/S0092-8674\(01\)00225-2](https://doi.org/10.1016/S0092-8674(01)00225-2)
- Tudek, A., Krawczyk, P.S., Mroczek, S., Tomecki, R., Turtola, M., Matylla-Kulińska, K., Jensen, T.H., Dziembowski, A., 2021. Global view on the metabolism of RNA poly(A) tails in yeast *Saccharomyces cerevisiae*. *Nat Commun* 12, 4951. <https://doi.org/10.1038/s41467-021-25251-w>
- Tudek, A., Lloret-Llinares, M., Jensen, T.H., 2018. The multitasking polyA tail: nuclear RNA maturation, degradation and export. *Phil. Trans. R. Soc. B* 373, 20180169. <https://doi.org/10.1098/rstb.2018.0169>
- Turtola, M., Manav, M.C., Kumar, A., Tudek, A., Mroczek, S., Krawczyk, P.S., Dziembowski, A., Schmid, M., Passmore, L.A., Casañal, A., Jensen, T.H., 2021. Three-layered control of mRNA poly(A) tail synthesis in *Saccharomyces cerevisiae*. *Genes Dev.* 35, 1290–1303. <https://doi.org/10.1101/gad.348634.121>

- Valkov, E., Muthukumar, S., Chang, C.-T., Jonas, S., Weichenrieder, O., Izaurralde, E., 2016. Structure of the Dcp2–Dcp1 mRNA-decapping complex in the activated conformation. *Nat Struct Mol Biol* 23, 574–579. <https://doi.org/10.1038/nsmb.3232>
- van Dijk, E., 2002. Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures. *The EMBO Journal* 21, 6915–6924. <https://doi.org/10.1093/emboj/cdf678>
- van Dijk, E.L., Chen, C.L., d’Aubenton-Carafa, Y., Gourvennec, S., Kwapisz, M., Roche, V., Bertrand, C., Silvain, M., Legoix-Né, P., Loeillet, S., Nicolas, A., Thermes, C., Morillon, A., 2011. XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. *Nature* 475, 114–117. <https://doi.org/10.1038/nature10118>
- Varet, H., Brillet-Guéguen, L., Coppée, J.-Y., Dillies, M.-A., 2016. SARTools: A DESeq2- and EdgeR-Based R Pipeline for Comprehensive Differential Analysis of RNA-Seq Data. *PLoS ONE* 11, e0157022. <https://doi.org/10.1371/journal.pone.0157022>
- Vidya, E., Duchaine, T.F., 2022. Eukaryotic mRNA Decapping Activation. *Front. Genet.* 13, 832547. <https://doi.org/10.3389/fgene.2022.832547>
- Vizcaíno, J.A., Csordas, A., del-Toro, N., Dianes, J.A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., Xu, Q.-W., Wang, R., Hermjakob, H., 2016. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* 44, D447–D456. <https://doi.org/10.1093/nar/gkv1145>
- Wahle, E., 1991. A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. *Cell* 66, 759–768. [https://doi.org/10.1016/0092-8674\(91\)90119-J](https://doi.org/10.1016/0092-8674(91)90119-J)
- Wang, H., Morita, M., Yang, X., Suzuki, T., Yang, W., Wang, J., Ito, K., Wang, Q., Zhao, C., Bartlam, M., Yamamoto, T., Rao, Z., 2010. Crystal structure of the human CNOT6L nuclease domain reveals strict poly(A) substrate specificity. *EMBO J* 29, 2566–2576. <https://doi.org/10.1038/emboj.2010.152>
- Webster, M.W., Chen, Y.-H., Stowell, J.A.W., Alhusaini, N., Sweet, T., Graveley, B.R., Collier, J., Passmore, L.A., 2018. mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases. *Molecular Cell* 70, 1089–1100.e8. <https://doi.org/10.1016/j.molcel.2018.05.033>
- Weinberg, D.E., Shah, P., Eichhorn, S.W., Hussmann, J.A., Plotkin, J.B., Bartel, D.P., 2016. Improved Ribosome-Footprint and mRNA Measurements Provide Insights into Dynamics and Regulation of Yeast Translation. *Cell Reports* 14, 1787–1799. <https://doi.org/10.1016/j.celrep.2016.01.043>
- Wells, S.E., Hillner, P.E., Vale, R.D., Sachs, A.B., 1998. Circularization of mRNA by Eukaryotic Translation Initiation Factors. *Molecular Cell* 2, 135–140. [https://doi.org/10.1016/S1097-2765\(00\)80122-7](https://doi.org/10.1016/S1097-2765(00)80122-7)
- Wolf, J., Passmore, L.A., 2014. mRNA deadenylation by Pan2–Pan3. *Biochemical Society Transactions* 42, 184–187. <https://doi.org/10.1042/BST20130211>
- Wolf, J., Valkov, E., Allen, M.D., Meineke, B., Gordiyenko, Y., McLaughlin, S.H., Olsen, T.M., Robinson, C.V., Bycroft, M., Stewart, M., Passmore, L.A., 2014. Structural basis for Pan3 binding to Pan2 and its function in mRNA recruitment and deadenylation. *EMBO J* 33, 1514–1526. <https://doi.org/10.15252/emboj.201488373>
- Wu, C., Roy, B., He, F., Yan, K., Jacobson, A., 2020. Poly(A)-Binding Protein Regulates the Efficiency of Translation Termination. *Cell Reports* 33, 108399. <https://doi.org/10.1016/j.celrep.2020.108399>

- Wu, Q., Smith-Miles, K., Zhou, T., Tian, T., 2013. Stochastic modelling of biochemical systems of multi-step reactions using a simplified two-variable model. *BMC Syst Biol* 7, S14. <https://doi.org/10.1186/1752-0509-7-S4-S14>
- Wu, Q., Tian, T., 2016. Stochastic modeling of biochemical systems with multistep reactions using state-dependent time delay. *Sci Rep* 6, 31909. <https://doi.org/10.1038/srep31909>
- Wyers, F., Rougemaille, M., Badis, G., Rousselle, J.-C., Dufour, M.-E., Boulay, J., Régnault, B., Devaux, F., Namane, A., Séraphin, B., Libri, D., Jacquier, A., 2005. Cryptic Pol II Transcripts Are Degraded by a Nuclear Quality Control Pathway Involving a New Poly(A) Polymerase. *Cell* 121, 725–737. <https://doi.org/10.1016/j.cell.2005.04.030>
- Xiang, K., Bartel, D.P., 2021. The molecular basis of coupling between poly(A)-tail length and translational efficiency. *eLife* 10, e66493. <https://doi.org/10.7554/eLife.66493>
- Xie, J., Kozlov, G., Gehring, K., 2014. The “tale” of poly(A) binding protein: The MLL domain and PAM2-containing proteins. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1839, 1062–1068. <https://doi.org/10.1016/j.bbagr.2014.08.001>
- Xu, Z., Wei, W., Gagneur, J., Perocchi, F., Clauder-Münster, S., Camblong, J., Guffanti, E., Stutz, F., Huber, W., Steinmetz, L.M., 2009. Bidirectional promoters generate pervasive transcription in yeast. *Nature* 457, 1033–1037. <https://doi.org/10.1038/nature07728>
- Yamashita, A., Chang, T.-C., Yamashita, Y., Zhu, W., Zhong, Z., Chen, C.-Y.A., Shyu, A.-B., 2005. Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nat Struct Mol Biol* 12, 1054–1063. <https://doi.org/10.1038/nsmb1016>
- Yi, H., Park, J., Ha, M., Lim, J., Chang, H., Kim, V.N., 2018. PABP Cooperates with the CCR4-NOT Complex to Promote mRNA Deadenylation and Block Precocious Decay. *Molecular Cell* 70, 1081–1088.e5. <https://doi.org/10.1016/j.molcel.2018.05.009>
- Zeidan, Q., He, F., Zhang, F., Zhang, H., Jacobson, A., Hinnebusch, A.G., 2018. Conserved mRNA-granule component Scd6 targets Dhh1 to repress translation initiation and activates Dcp2-mediated mRNA decay in vivo. *PLoS Genet* 14, e1007806. <https://doi.org/10.1371/journal.pgen.1007806>
- Zhang, E., Khanna, V., Namane, A., Doyen, A., Dacheux, E., Turcotte, B., Jacquier, A., Fromont-Racine, M., 2018. A SKI subcomplex specifically required for the degradation of ribosome-free RNA regions. <https://doi.org/10.1101/409490>
- Zhang, S., Welch, E.M., Hogan, K., Brown, A.H., Peltz, S.W., Jacobson, A., 1997. Polysome-associated mRNAs are substrates for the nonsense-mediated mRNA decay pathway in *Saccharomyces cerevisiae*. *RNA* 3, 234–244.
- Zigáčková, D., Vaňáčková, Š., 2018. The role of 3' end uridylation in RNA metabolism and cellular physiology. *Phil. Trans. R. Soc. B* 373, 20180171. <https://doi.org/10.1098/rstb.2018.0171>
- Zünd, D., Gruber, A.R., Zavolan, M., Mühlemann, O., 2013. Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3' UTRs. *Nat Struct Mol Biol* 20, 936–943. <https://doi.org/10.1038/nsmb.2635>
- Zuo, Y., 2001. Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic Acids Research* 29, 1017–1026. <https://doi.org/10.1093/nar/29.5.1017>

## **Article de thèse**



# Deadenylation rate does not affect unstable RNA degradation in yeast

Léna Audebert<sup>1,2</sup>, Frank Feuerbach<sup>1</sup>, Laurence Decourty<sup>1</sup>, Abdelkader Namane<sup>1</sup>, Emmanuelle Permal<sup>1,3</sup>, Gwenaël Badis-Breard<sup>1</sup>, Cosmin Saveanu<sup>1\*</sup>

Génétique des Interactions Macromoléculaires, Institut Pasteur, Sorbonne Université, CNRS UMR3525, Paris, France

## Affiliations

1 - Institut Pasteur, Université Paris Cité, CNRS UMR3525, Génétique des Interactions Macromoléculaires, F-75015 Paris, France.

2 – Sorbonne Université, Collège doctoral, F75005 Paris, France.

3 - Institut Pasteur, Université Paris Cité, Bioinformatics and Biostatistics Hub, F-75015 Paris, France

\* cosmin.saveanu@pasteur.fr

## Abstract

RNA turnover in eukaryotes is triggered by decapping, believed to be activated mostly on oligoadenylated RNA. Thus, the main determinant of RNA degradation is considered to be deadenylation speed. To test this prevalent model, we investigated the protein and RNA composition of degradation complexes that bind long and short poly(A) tail RNAs. Contrary to expectations, we found that the poly(A) binding protein Pab1 was strongly associated with unstable RNAs with long poly(A) tails. This observation was correlated with genome-wide studies showing that the most unstable RNAs have longer than average poly(A) tails. We modelled RNA decay and found that, to explain the available large-scale results, the decapping of oligoadenylated RNA has to occur at variable rates. Alternatively, the data can also be explained by decapping activation at various rates but independent of the poly(A) tail size. To distinguish between these two hypotheses, we measured the stability of a rapidly degraded reporter RNA after fast depletion of deadenylases. Despite a global increase of RNA poly(A) tail length, slowing down deadenylation did not affect the degradation of the unstable reporter RNA. Our analyses and results suggest that decapping activation dictates RNA stability but is independent of the deadenylation speed.

## Introduction

Addition of poly(A) tails to RNAs is a major co-transcriptional event during mRNA formation in eukaryotes. Polyadenylation has probably evolved as a catalytic tailing of the 3' end of RNAs by RNA degradation enzymes that can act in reverse of their exonucleolytic action using available triphosphorylated nucleotides (Slomovic & Schuster, 2011). This type of "degradative" poly(A) tail characterizes bacteria, archaea and the degradation of specific RNAs in the nucleus and organelles of eukaryotes (reviewed in Liudkovska & Dziembowski, 2021). However, in eukaryotes, the poly(A) tails generated by co-transcriptional action of the canonical poly(A) polymerases, assist RNA export to the nucleus, translation and have a protective effect from degradation in the cytoplasm (reviewed in Passmore & Collier, 2021). Cytoplasmic poly(A) length changes are particularly important for translational regulation of gene expression during development (reviewed in Winata & Korzh, 2018).

The effects of the poly(A) tail on RNA are mostly dependent on their association with poly(A) tail binding proteins (PABPs), present in all eukaryotes. PABPs are involved in mRNA translation through interactions with translation-associated factors (reviewed in Vicens *et al*, 2018). PABPs also have an active role in deadenylation and protection from it, by acting both as a stimulators and inhibitors of the enzymes of the major deadenylation CCR4/NOT complex (reviewed in Passmore & Collier, 2021). Changes in deadenylation speed were proposed to mediate the effects of translation, through codons usage, on RNA stability (Presnyak *et al*, 2015; Herrick *et al*, 1990). Estimates of half-life for reporter mRNAs that are not translatable indicate, however, that translation can have both a stabilizing and destabilizing effect on RNA (Jaquet *et al*, 2022). How this occurs is unclear, even if translation stalls on non-optimal coding sequences have been proposed to be detected by the action of one of the CCR4/NOT components, Not5, to activate deadenylation (Buschauer *et al*, 2020). Surprisingly, depletion of both PABP1 and PABP4 in human cells, leads to major transcriptome changes, even in the absence of obvious translation defects (Kajjo *et al*, 2022), suggesting that the links between poly(A), translation and RNA degradation might need to be re-evaluated. Moreover, despite the high abundance of cytoplasmic PABP, the protein can become limiting for the translation of oligoadenylated RNAs (Xiang & Bartel, 2021).

Accelerated deadenylation, observed for unstable RNAs, is considered an essential determinant for RNA decay, eventually triggered by decapping. This prevalent model for RNA degradation (Decker & Parker, 1993) is based on pulse-chase experiments showing that RNA decapping, the limiting step in RNA degradation, occurs only after the poly(A) tail was reduced to 10-12 As (Muhlrad *et al*, 1994, 1995). This shortening of the poly(A) tail occurs at different speeds, slower for more stable RNA, such as PGK1, and faster for unstable ones, such as MFA2 (Cao & Parker, 2001). Oligoadenylated RNAs are recognized

by the Lsm1-7/Pat1 complex, allowing recruitment and activation of decapping (Tharun *et al*, 2000; Bouveret *et al*, 2000). Similar events are believed to occur in mammalian cells as well (reviewed in Chen & Shyu, 2011), even if the involvement of an Lsm1-7 complex in activating the decapping of reporter RNAs (Yamashita *et al*, 2005) has not been yet documented.

While the deadenylation-dependent model of RNA degradation is compatible with RNA reporter experiments, a mathematical model of pulse-chase results showed that decapping speed of oligoadenylated RNA species must be about 17 times slower for a stable RNA than for an unstable one (Cao & Parker, 2001). This result suggests that deadenylation speed is only partially responsible for RNA degradation rates. Moreover, the availability of large-scale methods that estimate RNA degradation rates and RNA poly(A) length, led to the conflicting observation that, contrary to the expectations, average poly(A) lengths of the most stable RNAs are unusually short (Lima *et al*, 2017). No proposed molecular mechanism can reconcile this observation with the prevalent model of deadenylation-dependent RNA decay.

RNA degradation can occur without prior deadenylation for unstable RNA populations, such as those that are degraded through nonsense-mediated mRNA decay (NMD) (Muhlrad & Parker, 1994). While deadenylation-independent RNA degradation can also occur in specific occurrences (see Vidya & Duchaine, 2022), it is not known whether such mechanisms is important more generally and besides NMD. We became interested in deadenylation-dependent and deadenylation-independent mechanisms of RNA degradation during our analysis of NMD complexes (Dehecq *et al*, 2018). This extensive investigation of protein complexes by quantitative mass-spectrometry of purified fractions revealed the association of the core NMD factor to the poly(A) binding protein Pab1, but also to the Lsm1-7/Pat1 proteins, specific for oligoadenylated RNA. Intrigued by this observation, we analysed the populations of RNA associated with several proteins, including Pab1. Surprisingly, we found that many unstable RNA were strongly enriched in association with Pab1, in correlation with their long average poly(A) tails under steady-state conditions (Tudek *et al*, 2021).

Encouraged by the obtained results, we reconsidered and explored the properties of kinetic models for RNA degradation, both dependent and independent of deadenylation using modern modelling tools and a minimal set of parameters. This analysis led to the conclusion that either decapping speed of oligoadenylated RNA is highly variable or that deadenylation-independent RNA decay is more prevalent than previously thought. To test these models we established an inducible degron system for the rapid concomitant depletion of the major deadenylases Ccr4 and Pop2 (Caf1) and for the decapping enzyme Dcp2 and validated their global effects by Nanopore RNA sequencing. As predicted by a deadenylation-independent model, inefficient deadenylation did not change the degradation rate of reporter RNAs, despite the presence of extended poly(A) tails. These experimental and computational results suggest that, at least for fast RNA degradation, deadenylation speed is not limiting in yeast. Our results are compatible with large-scale

data and correspond to predictions of a model in which deadenylation speed and decapping activation are independent events.

## Results

### The core NMD factor Upf1 is associated with RNA degradation complexes not involved in NMD

When characterizing yeast NMD complexes by an extensive set of affinity purification experiments followed by quantitative mass-spectrometry, we observed that the core NMD factor Upf1 was associated with both Pab1 and Lsm1-7/Pat1. These proteins are supposed to bind to two different types of RNA substrates, as several copies of Pab1 bind long poly(A) tails (Schäfer *et al*, 2019) and the Lsm1-7/Pat1 has affinity for oligoadenylated RNA (Chowdhury *et al*, 2014). For a better understanding of the involved factors, we undertook the purification of tagged versions of the following proteins present in association with Upf1 (Dehecq *et al*, 2018): Lsm1, specific to the Lsm1-7 complex, Lsm7, present also in the nuclear Lsm2-8 complex (Bouveret *et al*, 2000); Pat1, strongly associated with Lsm1-7 (Tharun *et al*, 2000; Bouveret *et al*, 2000), Dhh1, an RNA helicase proposed to be important in the translation-dependent destabilization of RNAs with non-optimal codons (Radhakrishnan *et al*, 2016) and Pab1, the major cytoplasmic poly(A) binding protein in yeast. The enrichment for specific proteins in each of the purifications as compared to total protein levels was visualized as volcano plots (**Fig 1a, b** for Pab1-TAP and Lsm1-TAP and **Fig S1 a to c** for the other tested proteins), with robustly enriched protein names indicated). The results confirmed previous findings about the protein complexes of each of the tagged factor (**Data Set S1**), validating the approach and the new results. Importantly, the RNA helicase Upf1 was found significantly enriched in all these purifications, showing that the purified factors and Upf1 are robustly found together in yeast RNPs.

For a global view of the obtained protein association results, we placed them in the context of previously published results. To this end, we obtained a list of published interactions for the tagged proteins that we used for the purification from the BioGrid database (Oughtred *et al*, 2020) and combined the results with our own. This led to an overview of the protein interaction network presented in **Fig. 1c**. Upf1 was in a particular position in this network, as it was present both in Lsm1-TAP and Pab1-TAP purifications. The situation of Upf1 was clearly different from other factors, as shown by comparing the enrichment ratios for Lsm2, a component of Lsm1-7 and Lsm2-8 complexes, and Upf1 in our experiments (**Fig. S1d**). These results potentially reflect the presence of Upf1 in many different RNA-protein complexes through interactions that are probably mediated by RNA, as many of Upf1 interactions are lost when purification protocols include an RNase step (Dehecq *et al*, 2018). Previous tests of Pab1 association with Upf1 showed this interaction to be resistant to RNase A treatment (Richardson *et al*, 2012). To check the dependence of Upf1 association

with Pab1, we purified Upf1 from a strain expressing a HA tagged version of Pab1 in the presence or absence of micrococcal nuclease. The specific enrichment of Pab1-HA in association with Upf1-TAP was lost after RNase treatment (**Fig. 1d**) confirming our large-scale results (Dehecq *et al*, 2018). The discrepancy from previous data could be explained by the use in those experiments of a milder, RNase A treatment, which cleaves RNA at the 3' position of C and U, in comparison with the broader spectrum of degradation of the micrococcal nuclease used here.

The analysis of the mass spectrometry data led to several interesting conclusions. It clearly showed the expected overlap between the interactomes of Pat1, Lsm1 and Lsm7, including their association with the decapping complex and the 5' to 3' exonuclease Xrn1. While many of the factors found in the mentioned purifications were also specifically enriched with Dhh1, its interactome was broader, and included other proteins with Lsm domains, such as previously identified Lsm12 (Cary *et al*, 2015) and Scd6. These results indicate that Dhh1 binds to RNA protein complexes that are not associated with Lsm1-7 and Pat1.

With the exception of the translation-related RNA helicase Ded1, Pab1 did not share interaction partners with the other purified proteins. The presence of nuclear cap binding complex (Cbc2 and Sto1, **Fig. 1a**) in association with Pab1 is compatible with the previous identification of these interactions by large-scale approaches (Gavin *et al*, 2006) and co-immunoprecipitation (Bretes *et al*, 2014). Since cap binding complex associates with RNA co-transcriptionally and accompanies RNAs for export from the nucleus (Sen *et al*, 2019), the observed interactions with Pab1 occur probably on newly formed RNAs, endowed by long poly(A) tails. These tails are shortened in a step-wise manner, with the Pan2-Pan3 complex involved in the process and assisted by Pab1 ((Boeck *et al*, 1996; Brown *et al*, 1996)). Interestingly, we found Pan3, the RNA binding component of the Pan2-Pan3 deadenylation complex (Wolf *et al*, 2014), specifically enriched in association with Pab1, an interaction that has been identified in two-hybrid experiments (Mangus *et al*, 2004) and reconstituted *in vitro* with purified proteins (Siddiqui *et al*, 2007). This interaction is enhanced by the presence of relatively long poly(A) sequences (Schäfer *et al*, 2019), but has not been directly detected by co-purification from yeast cells until now.

Altogether, the quantitative mass-spectrometry results were compatible with a situation in which Upf1 is present in at least two different types of mRNPs in which mRNAs have opposite poly(A) tail lengths, long in association with Pab1, and short when bound by Lsm1-7 (**Fig. 1d**).

## **Pab1 preferentially associates with unstable RNA**

To explore the potentially different populations of RNA molecules present in Lsm1-7 and Pab1-associated mRNPs together with Upf1, we sequenced the RNA co-purified with Lsm1-TAP, Pab1-TAP and Upf1-TAP. A purification of Rpl16a-TAP was also done in parallel, as it allows an estimation of the levels of RNA bound to ribosomes (Halbeisen *et al*, 2009).

The enrichment profiles for RNA with the tagged proteins, were, with the exception of the Lsm1-TAP situation, different from a control purification performed with a strain that did not express any tagged protein. We thus focused on the results obtained with Upf1, Rpl16a and Pab1. To validate the obtained results, we analysed first a sub-category of yeast RNA, the Xrn1-dependent unstable transcripts (XUT, van Dijk *et al*, 2011) and the stable unannotated transcripts (SUT, Xu *et al*, 2009) that are known to be unstable and frequently targets of the NMD pathway (Malabat *et al*, 2015). Since Upf1 is known to interact with NMD substrates (Johansson *et al*, 2007), we investigated the enrichment of SUT/XUT signal in comparison with total RNA. We observed a significant enrichment of this category in the Upf1 purification, while XUT/SUT RNAs were strongly depleted from the ribosome associated fraction (**Fig. 2a**). This depletion was expected, since XUT/SUT only contain spurious short coding sequences and, like other NMD substrates (Heyer & Moore, 2016), are probably at most associated with a single translating ribosome. Surprisingly, XUT/SUT were enriched in the Pab1 associated fraction. To confirm this observation, we turned to a different category of translation-destabilized transcripts, the unspliced mRNA precursors. One of them, the precursor of the mRNA for the ribosomal protein RPL28 (CYH2) was extensively used in NMD studies as a sensitive reporter, whose levels increase when NMD is inactivated. To differentiate pre-mRNA from spliced RNA, we separated the signal from the exons and introns of ribosomal protein gene RNAs that are spliced. As expected, and similar to the XUT/SUT situation, pre-mRNAs were significantly depleted from the ribosome associated fraction, strongly enriched in association with Upf1, and also enriched in the Pab1 bound RNA population (**Fig. 2b**). These large scale results (**Data Set S2**) can be visualized for specific transcripts, such as the SUT141, in comparison with the neighbouring gene transcript PFK1 (**Fig. 2c**). The relative enrichment of the pre-L28 transcript in Upf1-TAP and Pab1-TAP purification in comparison with the spliced RPL28 mRNA measured by RNA sequencing was confirmed by RT-Q-PCR (**Fig. 2d**).

We found Pab1 to be strongly associated with unstable transcripts that are also bound by NMD. To understand why these transcripts were preferentially enriched with Pab1-TAP, we took advantage of the recent Nanopore sequencing results that included the XUT/SUT in the evaluation of poly(A) tail length (Tudek *et al*, 2021). The XUT transcripts had longer than average poly(A) tails (**Fig. 2e**), which probably reflects their fast degradation through a deadenylation-independent mechanism. Since rapid degradation does not allow deadenylation to occur, long poly(A) tails of NMD substrates are a marker of RNA instability. We verified, using ePAT (Jänicke *et al*, 2012) that the poly(A) length for specific RNAs was indeed longer in the Pab1-associated fraction (**Fig. S2**). We conclude that Pab1 preferentially binds long poly(A) tail RNAs that have relatively short half-lives.

## Decay kinetics modelling explains the relation between poly(A) tail length and RNA stability

Stimulated by the unexpected finding that Pab1 prefers to bind to unstable RNAs with relatively long poly(A) tails, we explored the relationship between RNA half life and poly(A)



tail length from published large-scale data sets. We found poly(A) tail length was strongly inversely correlated with RNA stability (**Fig. 3a, 3b, S3a, b**) as previously reported (Chan *et al*, 2018; Lima *et al*, 2017). The observed correlations were independent on the method used for measuring RNA half-life or poly(A) tail length, which were well correlated (**Fig. S3c, S3d**). Even if rapid "pruning" of the poly(A) tail for highly expressed and stable RNAs has been proposed to explain these results (Lima *et al*, 2017), the hypothetical molecular mechanism responsible is not known.

If deadenylation speed dictates RNA stability, one would expect that changes in deadenylation activity to be reflected in predictable changes in RNA poly(A) tail lengths for both stable and unstable RNA. We took advantage of the recently published Nanopore sequencing results to look at the changes in poly(A) tail length in deadenylation deletion strains affecting PAN2 and CCR4 (Tudek *et al*, 2021). Most stable 25% RNAs and most unstable 25% RNAs showed opposite relative changes in poly(A) tail length (**Fig. 3c**). While unstable RNAs showed globally a milder increase in poly(A) tail length than more stable RNAs in the absence of PAN2, they had a larger increase in poly(A) tail length in the absence of CCR4. Thus, slowing down deadenylation leads to variations in poly(A) tail length that are not clearly linked with RNA stability, leading to the important question of the real impact of deadenylation on RNA degradation in yeast.

Previous analyses of RNA degradation assumed that deadenylation is the critical step in RNA degradation, with rapid degradation of RNA once it reaches a critical size of the poly(A) tail, of about 10 to 12 nucleotides (Muhlrad *et al*, 1994, 1995; Cao & Parker, 2001). We wondered how well this deadenylation-dependent model for RNA degradation fits the observed inverse correlation between average RNA poly(A) tail length and RNA half-life. To this end, we used a simplified kinetic model for RNA degradation inspired from previous work (Cao & Parker, 2001). In this model, only three poly(A) species are considered for each RNA, starting with long poly(A) tail RNA, which is deadenylated at a speed that is proportional with  $k_A$ , a deadenylation constant (**Fig. 3d**). Deadenylation gives rise to an intermediate poly(A) tail length RNA, which is further deadenylated to reach the stage of oligoadenylation that gets degraded with a first order constant  $k_D$ . The set of differential equations for the kinetics of RNA degradation includes a synthesis phase, which includes the processes before the arrival of an RNA in the cytoplasm ( $T$  is a transcript specific RNA synthesis pool from which the longest poly(A) species is synthesized at a constant rate). Solving the set of associated differential equations with Maxima, an algebra computer system (Maxima, 2022), led to expressions for the variation with time for each of the three considered species of poly(A) RNA (**Fig. S3e**). The limit of these functions for very large values of time, corresponding to steady-state levels of the transcripts, are inversely proportional with the deadenylation rate for the long and intermediate poly(A) RNA and inversely proportional with the degradation rate for the oligoadenylated RNA. A deadenylation-independent model was also considered, in which, in addition to the final degradation step for the oligoadenylated RNA, degradation can occur for the longer poly(A) species (**Fig. 3d, Fig. S3f** for the corresponding steady-state equations).

We wondered how deadenylation speed would affect the observed total RNA degradation rate and the steady-state average poly(A) size. Tellurium, a modelling and simulation software package (Choi *et al*, 2018), was used to follow the changes in the concentration of the three poly(A) RNA species over time, using the equations represented in **Fig. 3e**. The computed steady-state levels for the three species were used as a starting point in simulated "chase" experiments, to calculate both the variation in average poly(A) tails and amount of remaining RNA (**Fig. 3f**, for an example). For a constant  $k_D$  value corresponding to a very short half-life of the oligoadenylated RNA, we varied the value of  $k_A$  by a factor of 25. This led to a change in the half-life for the simulated RNA kinetics from 5 minutes, when deadenylation was fast, to 120 minutes, when it was slow. Using arbitrary values for the three poly(A) species of 70, 40 and 10 nucleotides, we computed average sizes of the poly(A) tails at steady-state in each of the three conditions. The obtained values were inversely correlated with the apparent half-life, ranging from 46 As for the fastest decaying RNA to 55 As for the slowest ones (**Fig. 3g**). This result is contrary to the observed long poly(A) tails of the most unstable RNAs (**Fig. 3a, b**). There are two alternatives to the model that can be considered to solve this conundrum. First, the decay rate of the oligoadenylated RNA could be variable. As expected, for a fixed fast deadenylation rate, decreasing  $k_D$  by a factor of 5 led to an increase in the apparent half-life from 5 to 9 minutes and a decrease in the average length of poly(A) tails at steady-state from 46 to 30 As. Further decreasing this constant by a factor of 5 led to an increase in RNA half-life from 9 to 35 minutes, which was accompanied by a decrease in average half-life of RNAs from 30 to 16 As. Thus,  $k_D$  changes, rather than deadenylation speed might be responsible for the observed correlation between poly(A) size and RNA instability (**Fig. 3g**).

We established that deadenylation speed alone cannot explain the differences in RNA degradation speed for yeast mRNA in a manner that is satisfactory for the observed steady-stated distribution of the average size of poly(A) tails. Since many unstable RNA in yeast are degraded through NMD, not dependent on deadenylation (Muhlrad & Parker, 1994), we tested the impact of introducing deadenylation independent decay steps in the kinetic model for RNA degradation (**Fig. 3d, e**). If the  $k_D$  for the various degradation steps is similar, this model is easier to see intuitively since the observed degradation of RNA would occur at a global  $k_D$  speed. Each different  $k_D$  value led to the same apparent half-life for the simulated experiment, with a range of average poly(A) tails length that was dependent on the simulated  $k_A$ . The effect of varying  $k_D$  and  $k_A$  for the simulated average poly(A) tail and half-life is represented in **Fig. 3h**.

In conclusion, the inverse correlation between average poly(A) tail length and RNA half life can be explained by the classical deadenylation-dependent model, but only when the limiting step for RNA degradation is not deadenylation speed but rather the degradation of the oligoadenylated species. Alternatively, the observed anti-correlation can also be explained by an RNA degradation model in which decapping can occur even on long poly(A) tail RNA, as it happens for NMD substrates. The two modes of RNA degradation are not mutually exclusive.



## Rapid degradation of deadenylation factors tests the impact of the poly(A) tail shortening on RNA levels

To investigate the importance of deadenylation as a determinant for RNA degradation rates, we decided to test the effect of blocking deadenylation on RNA levels and poly(A) lengths. The Ccr4-Not complex concentrates the major cytoplasmic deadenylation activities in two catalytic subunits, Ccr4 and Pop2 (reviewed by Collart, 2016). Previous studies on deadenylation effects used deletion strains for these deadenylases. However, the strains lacking one or the other of the deadenylases are slow growing and prone to accumulating secondary compensatory mutations, as shown, for example for *pop2Δ* (Szamecz *et al*, 2014). Moreover, slow growing strains can display global and specific changes in RNA degradation rates that are unrelated with the specific action of the deleted gene (García-Martínez *et al*, 2016). A method that partially circumvents these problems is the use of regulated promoters, such as GAL1, which works well to decrease the activity of CCR4, for example (Badis *et al*, 2004). Since the block is transcriptional, such methods need relatively long periods of time for protein depletion. An alternative is the induction of targeted protein degradation by fusing proteins of interest with an auxin-inducible degron (AID) domain recognized by the F-box, auxin-sensitive *O. sativa* TIR1 (Nishimura *et al*, 2009). As noted previously (Natsume *et al*, 2016), TIR1 has a residual degradation effect on AID-tagged proteins, even in the absence of the indole-3-acetic acid (IAA) auxin, and this effect could be avoided by regulating the expression of the protein (Mendoza-Ochoa *et al*, 2019). We adapted the *O. sativa* TIR1 and mini-AID degron system (Natsume *et al*, 2016) for fast and inducible specific degradation of Ccr4, Pop2 or both deadenylases (schematics in **Fig. 4a**). Addition of  $\beta$ -estradiol, for induction of TIR1 expression and IAA, to trigger proteasome-degradation of proteins led to rapid decrease in the levels of Ccr4, Pop2 or both proteins in a double degron Ccr4/Pop2 strain (**Fig. 4b**). This decrease in protein levels was accompanied by a slow-growth phenotype that was stronger for the depletion of Ccr4, in correlation with the stronger effect on growth observed for a *ccr4Δ* strain in comparison with a *pop2Δ* strain (**Fig. 4c**). In parallel with the deadenylation defective strains, we also generated an inducible degron strain for Dcp2, derived from a degron strain previously built in our laboratory (Dehecq *et al*, 2018).

We used Nanopore sequencing to estimate, at the same time, changes in poly(A) tail length and RNA levels (Tudek *et al*, 2021) in a strain depleted for the two deadenylases of the Ccr4-Not complex and, as a control, in a strain depleted for Dcp2. Depletion of these proteins led to global changes in poly(A) tail length, when compared with a wild type strain treated in parallel with  $\beta$ -estradiol and IAA. As expected, when the deadenylases were depleted, mRNA poly(A) tails became globally longer, with an average shifting from 42.5 to 52 As (**Fig. 4d**). A similar shift was measured previously in cells in which CCR4 was deleted (Tudek *et al*, 2021). Depletion of Dcp2 had the opposite effect, with a decrease of the average poly(A) tail size to 36.3 A. The advantage of Nanopore sequencing is that, in addition to the information about poly(A) tail length, an estimate of the amount of RNA and changes in different conditions is made. As we were interested in the impact of poly(A) tail

changes on RNA stability we looked at how individual transcript poly(A) tails changed in the degron strains in relation with RNA amounts (**Data Set S3**). For example, a known NMD substrate, expected to be stabilized by inhibition of Dcp2, showed a drop in poly(A) tail length from 51 to 36 nucleotides after Dcp2 depletion. This change is expected for NMD substrates whose degradation is independent of deadenylation since their stabilization leaves more time for deadenylation. We were surprised to find a strong global negative correlation between the increase in mRNA levels and the change in poly(A) length when Dcp2 was depleted (**Fig. 4e**). RNAs whose levels increased the most, had also the largest decrease in the size of their poly(A) tails. The global negative correlation suggests that either the NMD substrates are a lot more prevalent than previously thought or, alternatively, that deadenylation-independent RNA degradation occurs on other unstable RNA, independent of NMD. This hypothesis was supported by the lack of any correlation between changes in poly(A) tail length and RNA levels in the strains depleted for the Ccr4 and Pop2 deadenylases (**Fig. 4f**).

Perturbation of deadenylation or decapping led to results that are compatible with a deadenylation-independent degradation of unstable RNAs in yeast. Alternatively, these results can also be explained by a deadenylation-dependent mechanism in which the limiting step is not deadenylation speed (**Fig. 3g**) but the decapping of oligoadenylated RNA.

## Decapping, but not deadenylation speed, affects the degradation of a reporter RNA

Our results indicated that an NMD-like deadenylation-independent degradation could affect a sizeable fraction of yeast mRNAs. To be able to link deadenylation and degradation rates for RNAs that are not NMD substrates, we used a reporter system to estimate the dynamics of RNA levels and of the poly(A) tail when decapping or deadenylation were inactivated. To compare RNAs with different stabilities, we took advantage of the previous observation that coding sequence bias affects RNA degradation (Presnyak *et al*, 2015; Herrick *et al*, 1990) and built two reporter RNAs with HIS3 coding sequences that were either optimal (OPT-HIS3) or suboptimal (non-OPT-HIS3) in terms of codon usage. Instead of using a GAL1 repression system with identical coding sequences reporter RNAs (10% and 100% optimal, Radhakrishnan *et al*, 2016) we preferred a tetOFF system to be able to block transcription for the reporter with minimal changes in the growth medium (**Fig. 5a**). Estimations of RNA half-life for the optimal and non-optimal HIS3 reporters were obtained by RT-Q-PCR, which indicated that turnover of optimal HIS3 RNA was slower than the non-optimal one. Half-life estimates were of 16 minutes in one case and 8 minutes in the other (**Fig. 5b**). To further establish if translation was affecting mRNA half-life for these new reporters, we evaluated the effect of a cycloheximide treatment on RNA degradation. Increasing cycloheximide from 5 to 50 µg/ml led to progressive stabilization of RNA, especially visible for the non-opt-HIS3 reporter in the tested short time frame (**Fig. 5d**).

We also evaluated the stability of reporter RNAs in relation with Dhh1, an RNA helicase that distinguishes non-optimal codon, fast turnover RNAs from more stable ones

(Radhakrishnan *et al*, 2016). Since the *dhh1* $\Delta$  strain was growing very slowly in our hands, we used the inducible degron strategy to rapidly deplete the protein (**Fig. S4a**). Unexpectedly, depletion of Dhh1 was not followed by the stabilisation of non-OPT-HIS3 or opt-HIS3 mRNAs, nor did the deletion of DHH1 (**Fig. S4b, c**). Intrigued by this result, we tested the kinetics of degradation of the original reporters, kindly provided by J. Collier (Radhakrishnan *et al*, 2016). We observed the stabilization of the GAL1 driven non-OPT-HIS3 mRNA both in the *dhh1* $\Delta$  and Dhh1 depletion conditions (**Fig. S4d**). Thus, the non-OPT-HIS3 reporter RNA was rapidly degraded in a translation-dependent process, albeit its turnover did not seem to depend on Dhh1. The reason for this difference from published results can be the use of a tetOFF system in our case and, especially, the difference in the 3'UTR region of the reporter, originating from the CYC1 gene in our case and from the HIS3 locus in the previously studied reporters. The difference from published data did not preclude the use of the new reporters to study the mechanisms of RNA degradation, albeit they raise the question of the potential role of 3' UTR regions in translation-dependent RNA degradation.

While the new reporters were sensitive to translation inhibition, we wondered if decapping was a limiting step in their turnover. We thus tested the levels of RNA in a chase experiment after inhibition of transcription with doxycyclin in a wild type and after Dcp2 depletion. The half-life of both reporters doubled (**Fig. 6a, b** and **Fig. S5a**) in these conditions, showing that reporter RNAs are dependent on decapping for their fast degradation. Intriguingly, the measured half life of the non-OPT-HIS3 and OPT-HIS3 was not affected in inducible degron strains after depletion of Ccr4, Pop2 or both these deadenylases (**Fig. 6c, d, e** and **Fig. S5b, c, d**). This was intriguing, particularly in light of the massive and global increase in the RNA poly(A) tail lengths in the double depletion strain measured by Nanopore sequencing under similar conditions (**Fig. 4d, f**). We thus looked at both steady-state and deadenylation dynamics for the reporter RNAs, by using RNase H digestion and Northern blotting (**Fig. 7a, Fig. S6a**). Depletion of deadenylases, either alone or combined, drastically changed the steady-state distribution of poly(A) tails for both reporters. The most striking change was the decrease in the signal corresponding to relatively short poly(A) tails (**Fig. 7b, c, Fig. S6b, c**).

Further insight into the changes in poly(A) tails length distribution was obtained by quantifying the Northern blot signal and normalizing the values to see their evolution over time (**Fig. 7d, e**). We observed, as expected, the progressive decrease in the size of poly(A) tails with time, as a result of deadenylation of the reporter RNA under wild-type conditions. At steady-state, poly(A) tails encompassed a broad range, from less than 10 to more than 80 A. However, when deadenylases were depleted, in isolation or at the same time, the signal corresponding to short poly(A) tails was strongly reduced. While short poly(A) tails could not be visible for the reporter RNAs in cells depleted for Ccr4, they could be partially detected in the Pop2 depleted cells. These differences were comparable with those observed for total RNA poly(A) tail dynamics after transcription shut-off, where CCR4 deletion led to a loss of deadenylated species shorter than 20-23 A (Webster *et al*, 2018).

In conclusion, major changes in poly(A) and deadenylation status of reporter RNAs had no impact on RNA half-life, suggesting that deadenylation-independent RNA degradation might be a more prevalent mechanism than previously thought.

## Discussion

We present here several lines of evidence in favour of a deadenylation-independent model for a large RNA degradation in yeast. Our conclusions affect the understanding of how gene expression is regulated and suggest that deadenylation is not the main cause but it rather accompanies RNA degradation. First, we found that the poly(A) binding protein Pab1 associates tightly with unstable RNAs with longer than average poly(A) tails, many of them substrates of the NMD pathway (**Fig. 2**). Interestingly, and correlated with previous studies, unstable RNAs tend to have longer than average poly(A) tails not only in yeast, but also in other species, such as *C. elegans* (Lima *et al*, 2017). Second, we propose a simple kinetic model, with only two parameters, to explain the steady-state poly(A) tail lengths as a function of RNA degradation rate. In line with the thorough examination of RNA degradation by modeling (Cao & Parker, 2001, 2003), we found that, to explain large-scale results, either deadenylation-independent degradation is prevalent, or deadenylation plays a minor role in comparison with decapping activation on oligoadenylated species (**Fig. 3**). Third, to have a better and minimally perturbed view of the importance of deadenylation in RNA degradation, we built strains for rapid degradation of the major deadenylation factors Ccr4 and Pop2 and show that, even if most RNAs bear longer poly(A) tails under these conditions, their relative levels were not affected. This was in contrast with the depletion of Dcp2, which clearly showed that the RNAs that had initially the longer poly(A) tails and were most unstable, increased in their levels and showed a downshift in poly(A) tail length (**Fig. 4**). Finally, the detailed investigation of the kinetics of RNA degradation for two reporter RNAs independent of NMD and having different degradation rates (**Fig. 5**), showed the lack of correlation between poly(A) status and rate of deadenylation and the degradation speed (**Fig. 6, 7**).

Our results might seem at odds with the widely accepted idea that deadenylation precedes decapping in the degradation of RNA in eukaryotes. However, even the results of large-scale metabolic labeling and poly(A) tail length measurements in 3T3 mouse cells (Eisen *et al*, 2020) are compatible with a deadenylation-independent mechanism for RNA degradation. Individual highly unstable transcripts, such as *Metrn1*, showed a broad distribution of poly(A) tails at steady-state, while a very stable one, such as *Eef2*, displayed a more compact distribution, with shorter poly(A) tails on average. Individual cases were not reflected in a global correlation between poly(A) tail length and RNA stability, potentially because deadenylation and decapping have different ranges and dynamics in mammalian cells compared with yeast. Importantly, fitting a deadenylation-dependent model for RNA degradation involved a highly variable parameter, the speed of decapping for oligoadenylated RNA species. This speed varied by a factor of 1000 (Eisen *et al*, 2020),

which is similar to the range of variation in the reported model for the deadenylation speed. Thus, the presented model only fits experimental data if it includes a highly variable and limiting decapping speed. Indeed, our modelling of RNA degradation kinetics identified the same requirement for a strong variation in decapping speed for the deadenylation-dependent mechanism, to explain large-scale steady-state results (**Fig. 3g**). As demonstrated here, a deadenylation-independent mechanism could also lead to a similar distribution of poly(A) tails length and half-life (**Fig. 3f**). Previous simulations of RNA degradation based on the analysis of pulse-chase results for an unstable and stable RNA in yeast led to a similar conclusion: the stable RNA decapping should happen at much lower rates than the unstable one (Cao & Parker, 2001). Thus, even if degradation of RNA occurs only after deadenylation, a slow decapping rate could ensure the stability of an RNA, independent on its deadenylation rate. Our results bring into attention this important, and often overlooked element in RNA degradation. What mechanisms are involved in such a difference is unknown.

While ours and previous modelling results are mostly correlative, it is interesting to see what happens with RNA degradation in the case of perturbing deadenylation or decapping. An important result that cannot be explained by a deadenylation-dependent model for RNA degradation, is the lack of correlation between changes in poly(A) tail length and changes in RNA levels in deadenylation mutants as shown here (**Fig. 4f**) and previously published for *ccr4Δ* and *pan2Δ* (Tudek *et al*, 2021). These results were correlated with the lack of a stabilizing effect of depleting Ccr4, Pop2 or of both deadenylases on reporter RNAs, despite important changes in their poly(A) status (**Fig. 6, 7**).

Altogether, these results tend to demonstrate that deadenylation is not a cause for RNA degradation, but rather a parallel process, that can occur at different rates for different transcripts. Further arguments for an uncoupling between deadenylation and RNA degradation can be found in published results from a seminal paper that described the importance of codon usage in RNA stability. In a chase experiment, the signal corresponding to oligo(dT) purified RNA for the highly stable PGK1 transcript was lost in 20 to 30 minutes, while the signal corresponding to poly(A)- RNA remained constant for over 60 minutes after transcription inhibition (Herrick *et al*, 1990). It is thus likely that the poly(A)+ fraction "degradation" rate was, in reality, a measure of progressive deadenylation with the loss of signal corresponding to a shift of RNA to shorter poly(A) tails. Such a discrepancy between the poly(A)+ and total RNA degradation rates was also observed on a large scale (Presnyak *et al*, 2015) and could be thus explained, at least partially, by the fact that measuring a poly(A)+ fraction signal loss corresponds to both measuring deadenylation and RNA degradation.

We believe that a re-evaluation of the deadenylation-dependent and deadenylation-independent models for RNA degradation could help in better understanding the complex relationship between the poly(A) tail, translation and RNA stability. The results presented here should encourage the use of alternative models for RNA degradation in the interpretation of large-scale results to allow the identification of the responsible molecular mechanisms.

## Materials and methods

### Yeast and bacterial strains

*Saccharomyces cerevisiae* strains were derived from BY4741 (Mat a) and BY4742 (Mat  $\alpha$ ) strains. The full list of used strains is provided in the Supplementary material. Strain GBy68 (Bushkin *et al*, 2019) was kindly provided by G. Bushkin, Whitehead Institute for Biomedical Research, Cambridge, MA, USA. An alpha haploid spore derived from this strain, with the *rme1*- $\Delta$ , *ins*-308A and the *tao3*(E1493Q) alleles was selected and crossed with a *trp1*-1 (i.e. *trp1*(D87-stop)) derivative of BY4741. After sporulation of the resulting diploid cell, spores with the following genotype Mat a, *LYS2*, *met15* $\Delta$ 0, *ura3* $\Delta$ 0, *leu2* $\Delta$ 0, *his3* $\Delta$ 1, *trp1*-1, *RME1*- $\Delta$ , *ins*-308A, *TAO3*(E1493Q) or Mat alpha, *lys2* $\Delta$ 0, *ura3* $\Delta$ 0, *leu2* $\Delta$ 0, *his3* $\Delta$ 1, *trp1*-1, *RME1*- $\Delta$ , *ins*-308A, *TAO3*(E1493Q) were selected to create strains LMA5395 and LMA5396 respectively. LMA5395 was transformed with plasmid 1692 linearized by *Bsu*36I digestion. A tryptophan prototroph transformant was selected to create strain LMA5419.

*E. coli* strain NEB 10-beta (NEB Cat# C3019) was used for construction of plasmids by ligation-less Gibson assembly (Fu *et al*, 2014) and their multiplication. All the plasmid inserts were verified by sequencing.

C-terminal TAP-tagged strains originated from the collection of systematically built strains (Ghaemmaghami *et al*, 2003). Deletion strains were part of the systematic yeast gene deletion collection (Giaever *et al*, 2002) distributed by EuroScarf (<http://www.euroscarf.de>) or were built by transformation of BY4741 strain with a cassette containing a selection marker cassette flanked by long recombination arms located upstream and downstream the open reading frame of the gene. Deletions were tested by PCR amplification of the modified locus.

To create auxin-inducible degron strains, the indicated genes were C-terminally tagged with a polyG-3Flag-miniAID-kanMX6 cassette amplified by PCR using plasmid p1451 as a template (Dehecq *et al*, 2018) in LMA5396. The resulting strains were then crossed with LMA5419 and sporulated to recover strains of the desired genotype. The only exceptions were the CCR4-degron and HYP2-degron strains for which the AID-tag amplified from plasmid p1602 has been inserted at the N-terminus of the protein by CRISPR/CAS9 engineering as described (Mans *et al*, 2015).

### Media and growth conditions

Yeast cells were grown in YPD (20g.L<sup>-1</sup> glucose, 10g.L<sup>-1</sup> yeast extract, 20g.L<sup>-1</sup> bacto-peptone, 20g.L<sup>-1</sup> bactoagar for plates only) and in synthetic media without uracil to select transformants and maintain plasmids with the URA3 marker. All yeast strains were freshly thawed from frozen stocks and grown at 30°C. Bacterial strains were grown in LB media, supplemented with antibiotics when necessary, at 37°C.



## Plasmids construction

Plasmids for the His3 codon-changed reporters. Coding sequences for His3 versions, « 10 % » and « 100 %» were recovered from published data (Radhakrishnan *et al*, 2016). The DNA fragments were synthesised by Twist Bioscience (San Francisco, CA, USA). The coding sequences were amplified with LA070 and LA071 before a Gibson assembly reaction. The final PCR product was cloned into the *Bam*HI and *Not*I sites of pCM189 (Garí *et al*, 1997).

Plasmids for the degron system. For the p1603 plasmid, the Z3EV artificial transcription factor was PCR amplified from plasmid pFS461 (Ohira *et al*, 2017) using oligonucleotides FF3765 and FF3766 and the PCR product was used to transform a URA3 derivative of BY4741. 5-FOA resistant cells were selected and checked for correct replacement of the URA3 coding sequence by the Z3EV artificial transcription factor coding sequence. The *ura3::Z3EV* allele obtained was amplified from yeast genomic DNA using oligonucleotides FF293 and FF3746. The final PCR product was digested with *Sst*II and cloned into the *Sst*II and *Sma*I sites of pRS304.

For the p1692 plasmid, the Z3EV promoter was PCR amplified from plasmid pFS478 (Ohira *et al*, 2017) using oligonucleotides FF3740 and FF3872. The obtained PCR product was digested with *Sal*I and *Bam*HI. The *O. sativa* TIR1 ORF followed by the Nrd1 terminator was PCR amplified using oligonucleotides FF3718 and FF3719 from the genomic DNA of a *nrd1-AID* strain kindly provided by D. Challal (Domenico Libri laboratory, IJM, Paris, France). The PCR product was digested with *Bam*HI and *Xho*I. Both PCR products were cloned together into the *Xho*I site of plasmid p1603.

## Cell culture, RNA extraction and degradation kinetics

For RT-qPCR and RNA sequencing, cells were first grown in YPD to log phase and collected. Total RNA was extracted using the hot phenol extraction method and precipitated using ammonium acetate and ethanol.

For the tetOFF promoter inhibition and steady state analysis, cells expressing the appropriate plasmids were grown at 25°C in synthetic media without uracil to allow expression of the reporter mRNA. For the analysis of degron mutants, when cells reached a OD600 of 0.4,  $\beta$ -estradiol for Tir1 expression at a final concentration of 1  $\mu$ M and auxin at a final concentration of 100  $\mu$ M for protein depletion were added directly to the media. Several time points were collected to verify protein depletion by western blot with an anti-flag HRP conjugated . After 1h depletion, cell were harvested for steady state analysis. For transcriptional repression, doxyxycline was added at a final concentration of 10  $\mu$ g/mL. Cells were collected at the time points indicated in the Northern Blot, generally, 0, 2, 4, 8, 10, 20, 30 minutes for the non-optimal His3 RNA and 0, 5, 10, 20, 30, 60 for the more stable, optimal His3 RNA.

For the GAL1 promoter shut-off and steady state analysis, cells expressing the appropriate plasmids were grown at 25°C in synthetic media without uracil with 2% galactose to allow expression of the reporter mRNA. For the steady state analysis, cells were harvested at OD600 of 0.4. For transcriptional inhibition, cells were diluted in 2 % raffinose, 0.4 % galactose and transcription was repressed by adding glucose to a final concentration of 1%. Cells were collected at the time points indicated in the figures.

For translation arrest, cycloheximide was added at a final concentration of 5 to 50 µg/mL together with the doxycycline.

## Northern blots

RNAs were separated on 1.5% agarose gels, transferred on nylon membrane (Hybond N+, Amersham, GE) that was UV cross-linked at 0.120 Joules and probed with DIG-labelled RNA or DNA probes. Digoxigenin-containing RNA probes were generated by in vitro transcription with T7 polymerase using the DIG RNA Labelling kit (SP6/T7) from Roche (cat. No 11175025910). The oligonucleotide AJ529 (T7 promoter) pre-annealed with a target-specific oligonucleotide composed of the T7 promoter reverse complement fused to a template sequence, LA94 for His3-tCyc and LA95 for Scr1 were used as a template. Specific DNA probes for the non-optimal and optimal His3 RNA were generated using MFR917 and LA084. After PCR amplification and purification, a single stranded amplification was done using MFR918 for the opt and the non-opt version of the probes.

For poly(A) tail visualisation, RNAs were cleaved using Rnase H (New England Biolabs, Ipswich, MA, USA) with LA121 as a specific primer for His3-tCyc. Oligo-dT were added to the reaction to remove poly(A) tails for the « A0 » reference position. RNAs were then purified with acid phenol-chloroform extraction. Cleaved RNAs were migrated on 6% acrylamide gel, transferred on a nylon membrane (Hybond N+, Amersham) and UV cross-linked. For the DNA ladder, a denatured PBR322 DNA digested with *MspI* was used. Prior to the assay, digoxigenin-11-ddUTP was added to the 3' end of the digested fragments using terminal transferase enzyme (New England Biolabs, Ipswich, MA, USA).

## RNA half life measurements

A mix of reverse qPCR oligonucleotides, LA126 (Act1), LA102 (His3-tCyc) and random hexamers were used for the reverse transcription. Serial dilutions of cDNA were quantified by TaqMan probe-based quantitative PCR. The amplification was done with the Bio-Rad CFX96 machine and the corresponding software (CFX Maestro Software), with step 1 (95°C for 30 secondes) and step 2 (40 cycles of 95°C for 15 secondes and 60°C for 30 secondes). Used probes were LA122-5'-Cy5/TAO/3'IBRQ for His3-tCyc and LA125-5'-HEX/ZEN/3'IBFQ for Act1. Primers were LA102 and LA123 for HIS3, LA124 and LA126 for ACT1.



## Affinity purification for RNA sequencing, immunoblot and mass-spectrometry

TAP-tagged proteins were purified using a one-step purification method (as detailed in, Namane & Saveanu, 2022). Frozen cell pellets of 4 L culture were resuspended in lysis buffer (20 mM HEPES K pH 7.4, 100 mM KOAc, 0.5% Triton X100, 5mM MgCl<sub>2</sub>, Protease inhibitor, 1X Vanadyl-Ribonucleoside Complex, 40 units ml<sup>-1</sup> RNasin) and lysed using a MagNA Lyser (two passages of 90s at 4000 rpm). The lysate was cleared at 4°C for 20 min at 14000 rpm. Magnetic beads (Dynabeads M-270 epoxy) coupled to IgG were added to the protein extract and incubated for 1,5 h at 4°C (Oeffinger *et al*, 2007). Beads were collected with a magnet and extensively washed five times with a washing buffer (20 mM HEPES K pH 7.4, 100 mM KOAc 0.5% Triton X100, 5mM MgCl<sub>2</sub>). After washing, proteins were eluted by incubation in elution buffer (SDS 2% in Tris-EDTA pH 7.5) 10 minutes at 65°C. After collection, RNA were extracted by hot acid phenol/chloroform method, precipitated with ammonium acetate and ethanol. Samples were treated by RiboZero (Illumina) or Ribominus (Thermo Scientific) rRNA removal reagents and libraries were prepared using TruSeq Stranded mRNA kit (Illumina). We strictly followed the protocol except that we started directly at the fragmentation step after the rRNA removal step and we adjusted the number of cycle required to amplify libraries. For RNase treatment, 1 ul of micrococcal nuclease (Biolabs M02475, 2x10<sup>6</sup> U/ml) was added to 200 µl buffer containing 1 mM CaCl<sub>2</sub> and incubated for 10 min 37°C.

An aliquot of RNA was harvested prior to TAP purification for the analysis of total RNAs (input) and was treated the same way as the purified fractions except for the purification.

## Illumina RNA sequencing results analysis

Reads were aligned along the 16 chromosomes and mitochondrial sequence of *S. cerevisiae* S288C genome. For the mapping, we used the STAR (version 2.7) with default parameters except for: `--alignIntronMax 1500 --alignMatesGapMax 1500 --alignSJoverhangMin 25`. For the introns we used the annotated GTF file for *Saccharomyces cerevisiae* (R64-1-1.104) from ENSEMBL. Read counts were obtained with feature count from the Subread package, with an annotation file that contains non-coding RNA coordinates in addition to yeast gene transcripts (Malabat *et al*, 2015). The analysed data was visualized using the Integrated Genome Viewer, IGV (Robinson *et al*, 2011) and further data processing was done using R (R Core Team, 2022). It consisted in

## Reverse transcription and quantitative PCR

The extracted RNA samples were treated with DNase I (Ambion TURBO DNA-free kit) before reverse-transcription (RT) with the Superscript II or III (Invitrogen) and library preparation. The purified and input samples were used for reverse transcription with

transcript-specific primers. A mix of reverse qPCR oligonucleotides, CS888 (RPL28-premature), CS889 (RPL28) and random hexamers were used for the RT.

Serial dilutions of cDNA were quantified by qPCR, CS887-CS888 (RPL28-premature) and CS889-CS946 (RPL28). The amplification was done in a Stratagene MX3005P and corresponding software (MXpro quantitative PCR system), with step 1 (95°C for 5 min) and step 2 (40 cycles of 95°C for 20 s and 60°C for 1 min).

## Poly(A) tail length estimation with ePAT

We used the published protocol for ePAT estimation of poly(A) tail length of specific transcripts (Jänicke *et al*, 2012). For RIP samples (Pab1-TAP, Lsm1-TAP, Rpl16a-TAP, Upf1-TAP) we used 5 to 100ng RNA for the reaction.

For the input, we used 1 µg of total RNA. For the PAT, anchor primer CS1400 that binds poly(A) tails was used, while for the TVN-PAT assay, used as a size control, we used CS1401. For the TVN-PAT reaction, the protocol was similar to ePAT except that no Klenow polymerase was added to the reaction. CS1402 was used for all genes as a reverse primer. For Hht2 we used CS1455 and GB1399 for Ccw22. For the PCR reactions, cDNA was diluted 1:6 with water. PCR reactions were done in a total volume of 20 µL using 5 µL of diluted cDNA input and a Phusion polymerase (ThermoFisher) with 30 cycles of amplification. For detection of PCR amplicons from ePAT and TVN-PAT, PCR reaction were loaded into a 2.5% high-resolution gel (MetaPhor Agarose) or 2% standard agarose and stained with ethidium bromide.

## Protein extracts and immunoblots

Total protein extracts were prepared from 5 ml of exponential culture cells using an alkaline treatment (Kushnirov, 2000). Cells were incubated in 200 µL of 0.1M NaOH for 5 min at room temperature, collected by 3 min centrifugation and resuspended in 50 µL of SDS sample buffer containing DTT (0.1M). Proteins were denatured for 3 min at 95°C, and cellular debris were pelleted by centrifugation. 10 µL of supernatant or diluted supernatant (for quantification scale) were loaded on acrylamide NuPAGE Novex 4-12% Bis-Tris gels (Life technologies). Transfer to a nitrocellulose membrane was done with a semi-dry fast system (Biorad trans-blot) with discontinuous buffer (BioRad technote 2134).

Proteins were detected by hybridization with anti-FLAG-HRP , for the detection of the FLAG tag (A8592, Sigma-Aldrich, clone M2, monoclonal), PAP (P1291, Sigma-Aldrich, peroxidase anti-peroxidase soluble complex antibody), for the detection of the protein A fragment of the TAP tag or anti-HA peroxidase 1/500 (clone 3F10, Roche).

## Mass spectrometry acquisition and data analysis

Protein samples were treated with Endoprotease Lys-C (Fujifilm Wako chemicals, Osaka, Japan) and Trypsin (Trypsin Gold Mass Spec Grade, Promega). Peptide samples were

desalted using OMIX C18 pipette tips (Agilent Technologies). The peptides mixtures were analysed by nano-LC-MS/MS using an Ultimate 3000 system (Thermo Fisher Scientific) coupled to an LTQ-Orbitrap Velos mass spectrometer. Peptides were desalted on-line using a trap column (C18 Pepmap100, 5 $\mu$ m, 300 $\mu$ m $\times$ 5mm (Thermo Scientific) and then separated using 120min RP gradient (5–45% acetonitrile/0.1% formic acid) on an Acclaim PepMap100 analytical column (C18, 3 $\mu$ m, 100., 75 $\mu$ m id x 150mm, (Thermo Scientific) with a flow rate of 0.340 $\mu$ L.min<sup>-1</sup>. The mass spectrometer was operated in standard data dependent acquisition mode controlled by Xcalibur 2.2. The instrument was operated with a cycle of one MS (in the Orbitrap) acquired at a resolution of 60,000 at m/z 400, with the top 20 most abundant multiply-charged (2+ and higher) ions subjected to CID fragmentation in the linear ion trap. An FTMS target value of 1e6 and an ion trap MSn target value of 10000 were used. Dynamic exclusion was enabled with a repeat duration of 30s with an exclusion list of 500 and exclusion duration of 60s. Lock mass of 445.12002 was enabled for all experiments. The results were analysed with MaxQuant and processed with Perseus and R, as previously described (Dehecq *et al*, 2018). Enrichment calculations were based on a standard data set of protein abundance in yeast (Ho *et al*, 2018). At least three independent experiments were performed for each purified protein.

## Nanopore sequencing and analysis

Degradation of target proteins was induced for 1 hour in exponential growth cultures in rich (YPD) medium, with  $\beta$ -estradiol (1 $\mu$ M final) and IAA (100 $\mu$ M final). Protein depletion was verified by immunoblot against the FLAG tag. Total RNA were extracted using glass beads and hot phenol protocols. RNA libraries were prepared from 250 to 500 ng of oligo-(dT)25-enriched mRNA from with a Direct RNA Sequencing Kit (catalog no. SQK-RNA002, Oxford Nanopore Technologies) according to the manufacturer's instructions. Sequencing was performed using R9.4 flow cells on a MinION device (ONT). Raw data were basecalled using Guppy (ONT). Reads were mapped to the SC288 reference genome Sacc38.fa using minimap2.1 with options -k 14 -ax map-ont --secondary = no and processed with samtools 1.9 (samtools view -b -o). The poly(A) tail lengths for each read were estimated using the Nanopolish 0.13.2 polya function (Workman *et al*, 2019). In subsequent analyses, only length estimates with the QC tag that was reported by Nanopolish as PASS were considered. For further analyses, only features with at least 20 assigned reads in two independent replicates were selected. Statistics of obtained results, poly(A) tail length and counts, were computed with R.

## Data and software availability

The accession number for the sequencing data reported in this paper is GSE160642 (GEO, Gene Expression Omnibus, NCBI). The MS proteomics data that support the findings of this study have been deposited in the ProteomeXchange repository via the

PRIDE interface\_(Perez-Riverol *et al*, 2022) with the dataset identifier PXD028008. Nanopore results were deposited in GEO, with accession number GSM6500926.

## Acknowledgements

This work was funded, in whole or in part, by the Institut Pasteur, CNRS, ANR (grant ANR-18-CE11-0003-04), the "Ministère de l'enseignement supérieur et de la recherche, France" and the "Fondation ARC" (PhD grant to LA). For the purpose of open access, the authors have applied a CC-BY public copyright license to any Author Manuscript version arising from this submission. We thank our colleagues of the Genetics of Macromolecular Interactions laboratory for fruitful discussions. We thank Mariette Matondo and Julia Chamot-Rooke for providing access to the Orbitrap Velos mass spectrometer at the Proteomics facility of the Institut Pasteur. We thank Laurence Ma (Biomics Platform, C2RT, Institut Pasteur, Paris, funded by France Génomique, ANR-10-INBS-09-09 and IBISA) for assistance in DNA sequencing. We thank G. Bushkin, J. Coller, D. Challal and D. Libri for sharing reagents and strains. The authors declare to have no conflicts of interest.

## References

- Badis G, Saveanu C, Fromont-Racine M & Jacquier A (2004) Targeted mRNA degradation by deadenylation-independent decapping. *Mol Cell* 15: 5–15
- Baudrimont A, Voegeli S, Vilorio EC, Stritt F, Lenon M, Wada T, Jaquet V & Becskei A (2017) Multiplexed gene control reveals rapid mRNA turnover. *Sci Adv* 3: e1700006
- Boeck R, Tarun S Jr, Rieger M, Deardorff JA, Müller-Auer S & Sachs AB (1996) The yeast Pan2 protein is required for poly(A)-binding protein-stimulated poly(A)-nuclease activity. *J Biol Chem* 271: 432–438
- Bouveret E, Rigaut G, Shevchenko A, Wilm M & Séraphin B (2000) A Sm-like protein complex that participates in mRNA degradation. *EMBO J* 19: 1661–1671
- Bretes H, Rouviere JO, Leger T, Oeffinger M, Devaux F, Doye V & Palancade B (2014) Sumoylation of the THO complex regulates the biogenesis of a subset of mRNPs. *Nucleic Acids Res* 42: 5043–5058
- Brown CE, Tarun SZ Jr, Boeck R & Sachs AB (1996) PAN3 encodes a subunit of the Pab1p-dependent poly(A) nuclease in *Saccharomyces cerevisiae*. *Mol Cell Biol* 16: 5744–5753
- Buschauer R, Matsuo Y, Sugiyama T, Chen Y-H, Alhusaini N, Sweet T, Ikeuchi K, Cheng J, Matsuki Y, Nobuta R, *et al* (2020) The Ccr4-Not complex monitors the translating ribosome for codon optimality. *Science* 368
- Bushkin GG, Pincus D, Morgan JT, Richardson K, Lewis C, Chan SH, Bartel DP & Fink GR (2019) m6A modification of a 3' UTR site reduces RME1 mRNA levels to promote meiosis. *Nat Commun* 10: 3414
- Cao D & Parker R (2001) Computational modeling of eukaryotic mRNA turnover. *RNA* 7: 1192–1212

- Cao D & Parker R (2003) Computational modeling and experimental analysis of nonsense-mediated decay in yeast. *Cell* 113: 533–545
- Cary GA, Vinh DBN, May P, Kuestner R & Dudley AM (2015) Proteomic Analysis of Dhh1 Complexes Reveals a Role for Hsp40 Chaperone Ydj1 in Yeast P-Body Assembly. *G3 (Bethesda)* 5: 2497–2511
- Chan LY, Mugler CF, Heinrich S, Vallotton P & Weis K (2018) Non-invasive measurement of mRNA decay reveals translation initiation as the major determinant of mRNA stability. *Elife* 7: e32536
- Chen C-YA & Shyu A-B (2011) Mechanisms of deadenylation-dependent decay. *Wiley Interdiscip Rev RNA* 2: 167–183
- Choi K, Medley JK, König M, Stocking K, Smith L, Gu S & Sauro HM (2018) Tellurium: An extensible python-based modeling environment for systems and synthetic biology. *Biosystems* 171: 74–79
- Chowdhury A, Kalurupalle S & Tharun S (2014) Pat1 contributes to the RNA binding activity of the Lsm1-7-Pat1 complex. *RNA* 20: 1465–1475
- Collart MA (2016) The Ccr4-Not complex is a key regulator of eukaryotic gene expression. *Wiley Interdiscip Rev RNA* 7: 438–454
- Decker CJ & Parker R (1993) A turnover pathway for both stable and unstable mRNAs in yeast: evidence for a requirement for deadenylation. *Genes Dev* 7: 1632–1643
- Dehecq M, Decourty L, Namane A, Proux C, Kanaan J, Le Hir H, Jacquier A & Saveanu C (2018) Nonsense-mediated mRNA decay involves two distinct Upf1-bound complexes. *EMBO J* 37: e99278
- van Dijk EL, Chen CL, d'Aubenton-Carafa Y, Gourvennec S, Kwapisz M, Roche V, Bertrand C, Silvain M, Legoix-Né P, Loeillet S, *et al* (2011) XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. *Nature* 475: 114–117
- Eisen TJ, Eichhorn SW, Subtelny AO, Lin KS, McGeary SE, Gupta S & Bartel DP (2020) The Dynamics of Cytoplasmic mRNA Metabolism. *Mol Cell* 77: 786-799.e10
- Fu C, Donovan WP, Shikapwashya-Hasser O, Ye X & Cole RH (2014) Hot Fusion: An Efficient Method to Clone Multiple DNA Fragments as Well as Inverted Repeats without Ligase. *PLoS ONE* 9: e115318
- García-Martínez J, Delgado-Ramos L, Ayala G, Pelechano V, Medina DA, Carrasco F, González R, Andrés-León E, Steinmetz L, Warringer J, *et al* (2016) The cellular growth rate controls overall mRNA turnover, and modulates either transcription or degradation rates of particular gene regulons. *Nucleic Acids Res* 44: 3643–3658
- Garí E, Piedrafita L, Aldea M & Herrero E (1997) A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in *Saccharomyces cerevisiae*. *Yeast* 13: 837–848
- Gavin A-C, Aloy P, Grandi P, Krause R, Boesche M, Marzioch M, Rau C, Jensen LJ, Bastuck S, Dümpelfeld B, *et al* (2006) Proteome survey reveals modularity of the yeast cell machinery. *Nature* 440: 631–636
- Ghaemmaghami S, Huh W-K, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK & Weissman JS (2003) Global analysis of protein expression in yeast. *Nature* 425: 737–741
- Giaever G, Chu AM, Ni L, Connelly C, Riles L, Véronneau S, Dow S, Lucau-Danila A, Anderson K, André B, *et al* (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418: 387–391
- Halbeisen RE, Scherrer T & Gerber AP (2009) Affinity purification of ribosomes to access the translome. *Methods* 48: 306–310

- Harrison PF, Powell DR, Clancy JL, Preiss T, Boag PR, Traven A, Seemann T & Beilharz TH (2015) PAT-seq: a method to study the integration of 3'-UTR dynamics with gene expression in the eukaryotic transcriptome. *RNA* 21: 1502–1510
- Herrick D, Parker R & Jacobson A (1990) Identification and comparison of stable and unstable mRNAs in *Saccharomyces cerevisiae*. *Mol Cell Biol* 10: 2269–2284
- Heyer EE & Moore MJ (2016) Redefining the Translational Status of 80S Monosomes. *Cell* 164: 757–769
- Ho B, Baryshnikova A & Brown GW (2018) Unification of Protein Abundance Datasets Yields a Quantitative *Saccharomyces cerevisiae* Proteome. *Cell Syst* 6: 192-205.e3
- Jänicke A, Vancuylenberg J, Boag PR, Traven A & Beilharz TH (2012) ePAT: a simple method to tag adenylated RNA to measure poly(A)-tail length and other 3' RACE applications. *RNA* 18: 1289–1295
- Jaquet V, Wallerich S, Voegeli S, Túrós D, Vitoria EC & Becskei A (2022) Determinants of the temperature adaptation of mRNA degradation. *Nucleic Acids Research* 50: 1092–1110
- Johansson MJO, He F, Spatrick P, Li C & Jacobson A (2007) Association of yeast Upf1p with direct substrates of the NMD pathway. *Proc Natl Acad Sci USA* 104: 20872–20877
- Kajjo S, Sharma S, Chen S, Brothers WR, Cott M, Hasaj B, Jovanovic P, Larsson O & Fabian MR (2022) PABP prevents the untimely decay of select mRNA populations in human cells. *EMBO J* 41: e108650
- Kushnirov VV (2000) Rapid and reliable protein extraction from yeast. *Yeast* 16: 857–860
- Lima SA, Chipman LB, Nicholson AL, Chen Y-H, Yee BA, Yeo GW, Collier J & Pasquinelli AE (2017) Short poly(A) tails are a conserved feature of highly expressed genes. *Nat Struct Mol Biol* 24: 1057–1063
- Liudkovska V & Dziembowski A (2021) Functions and mechanisms of RNA tailing by metazoan terminal nucleotidyltransferases. *Wiley Interdiscip Rev RNA* 12: e1622
- Malabat C, Feuerbach F, Ma L, Saveanu C & Jacquier A (2015) Quality control of transcription start site selection by Nonsense-Mediated-mRNA Decay. *Elife* 4: e06722
- Mangus DA, Evans MC, Agrin NS, Smith M, Gongidi P & Jacobson A (2004) Positive and negative regulation of poly(A) nuclease. *Mol Cell Biol* 24: 5521–5533
- Mans R, van Rossum HM, Wijsman M, Backx A, Kuijpers NGA, van den Broek M, Daran-Lapujade P, Pronk JT, van Maris AJA & Daran J-MG (2015) CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 15: fov004
- Maxima (2022) Maxima, a computer algebra system. Version 5.43.2.
- Mendoza-Ochoa GI, Barrass JD, Terlouw BR, Maudlin IE, de Lucas S, Sani E, Aslanzadeh V, Reid JAE & Beggs JD (2019) A fast and tuneable auxin-inducible degron for depletion of target proteins in budding yeast. *Yeast* 36: 75–81
- Miller C, Schwalb B, Maier K, Schulz D, Dümcke S, Zacher B, Mayer A, Sydow J, Marcinowski L, Dölken L, *et al* (2011) Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast. *Mol Syst Biol* 7: 458
- Muhlrad D, Decker CJ & Parker R (1994) Deadenylation of the unstable mRNA encoded by the yeast MFA2 gene leads to decapping followed by 5'→3' digestion of the transcript. *Genes Dev* 8: 855–866
- Muhlrad D, Decker CJ & Parker R (1995) Turnover mechanisms of the stable yeast PGK1 mRNA. *Mol Cell Biol* 15: 2145–2156

- Muhlrad D & Parker R (1994) Premature translational termination triggers mRNA decapping. *Nature* 370: 578–581
- Namane A & Saveanu C (2022) Composition and Dynamics of Protein Complexes Measured by Quantitative Mass Spectrometry of Affinity-Purified Samples. *Methods Mol Biol* 2477: 225–236
- Natsume T, Kiyomitsu T, Saga Y & Kanemaki MT (2016) Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. *Cell Rep* 15: 210–218
- Nishimura K, Fukagawa T, Takisawa H, Kakimoto T & Kanemaki M (2009) An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nat Methods* 6: 917–922
- Oeffinger M, Wei KE, Rogers R, DeGrasse JA, Chait BT, Aitchison JD & Rout MP (2007) Comprehensive analysis of diverse ribonucleoprotein complexes. *Nat Methods* 4: 951–956
- Ohira MJ, Hendrickson DG, Scott Mclsaac R & Rhind N (2017) An estradiol-inducible promoter enables fast, graduated control of gene expression in fission yeast. *Yeast* 34: 323–334
- Oughtred R, Rust J, Chang C, Breitkreutz B-J, Stark C, Willems A, Boucher L, Leung G, Kolas N, Zhang F, *et al* (2020) The BioGRID database: A comprehensive biomedical resource of curated protein, genetic and chemical interactions. *Protein Sci*
- Passmore LA & Collier J (2021) Roles of mRNA poly(A) tails in regulation of eukaryotic gene expression. *Nat Rev Mol Cell Biol*
- Perez-Riverol Y, Bai J, Bandla C, García-Seisdedos D, Hewapathirana S, Kamatchinathan S, Kundu DJ, Prakash A, Frericks-Zipper A, Eisenacher M, *et al* (2022) The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res* 50: D543–D552
- Presnyak V, Alhusaini N, Chen Y-H, Martin S, Morris N, Kline N, Olson S, Weinberg D, Baker KE, Graveley BR, *et al* (2015) Codon Optimality Is a Major Determinant of mRNA Stability. *Cell* 160: 1111–1124
- R Core Team (2022) R: A Language and Environment for Statistical Computing Vienna, Austria: R Foundation for Statistical Computing
- Radhakrishnan A, Chen Y-H, Martin S, Alhusaini N, Green R & Collier J (2016) The DEAD-Box Protein Dhh1p Couples mRNA Decay and Translation by Monitoring Codon Optimality. *Cell* 167: 122-132.e9
- Richardson R, Denis CL, Zhang C, Nielsen MEO, Chiang Y-C, Kierkegaard M, Wang X, Lee DJ, Andersen JS & Yao G (2012) Mass spectrometric identification of proteins that interact through specific domains of the poly(A) binding protein. *Mol Genet Genomics* 287: 711–730
- Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G & Mesirov JP (2011) Integrative genomics viewer. *Nat Biotechnol* 29: 24–26
- Schäfer IB, Yamashita M, Schuller JM, Schüssler S, Reichelt P, Strauss M & Conti E (2019) Molecular Basis for poly(A) RNP Architecture and Recognition by the Pan2-Pan3 Deadenylase. *Cell* 177: 1619-1631.e21
- Sen R, Barman P, Kaja A, Ferdoush J, Lahudkar S, Roy A & Bhaumik SR (2019) Distinct Functions of the Cap-Binding Complex in Stimulation of Nuclear mRNA Export. *Mol Cell Biol* 39: e00540-18



- Siddiqui N, Mangus DA, Chang T-C, Palermino J-M, Shyu A-B & Gehring K (2007) Poly(A) nuclease interacts with the C-terminal domain of polyadenylate-binding protein domain from poly(A)-binding protein. *J Biol Chem* 282: 25067–25075
- Slomovic S & Schuster G (2011) Exonucleases and endonucleases involved in polyadenylation-assisted RNA decay. *Wiley Interdiscip Rev RNA* 2: 106–123
- Smith LP, Bergmann FT, Chandran D & Sauro HM (2009) Antimony: a modular model definition language. *Bioinformatics* 25: 2452–2454
- Subtelny AO, Eichhorn SW, Chen GR, Sive H & Bartel DP (2014) Poly(A)-tail profiling reveals an embryonic switch in translational control. *Nature* 508: 66–71
- Szamecz B, Boross G, Kalapis D, Kovács K, Fekete G, Farkas Z, Lázár V, Hrtyan M, Kemmeren P, Groot Koerkamp MJA, *et al* (2014) The genomic landscape of compensatory evolution. *PLoS Biol* 12: e1001935
- Tharun S, He W, Mayes AE, Lennertz P, Beggs JD & Parker R (2000) Yeast Sm-like proteins function in mRNA decapping and decay. *Nature* 404: 515–518
- Tudek A, Krawczyk PS, Mroczek S, Tomecki R, Turtola M, Matylla-Kulińska K, Jensen TH & Dziembowski A (2021) Global view on the metabolism of RNA poly(A) tails in yeast *Saccharomyces cerevisiae*. *Nat Commun* 12: 4951
- Vicens Q, Kieft JS & Rissland OS (2018) Revisiting the closed loop model and the nature of mRNA 5'–3' communication. *Mol Cell* 72: 805–812
- Vidya E & Duchaine TF (2022) Eukaryotic mRNA Decapping Activation. *Frontiers in Genetics* 13: 832547
- Webster MW, Chen Y-H, Stowell JAW, Alhusaini N, Sweet T, Graveley BR, Collier J & Passmore LA (2018) mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases. *Mol Cell* 70: 1089-1100.e8
- Winata CL & Korzh V (2018) The translational regulation of maternal mRNAs in time and space. *FEBS Lett* 592: 3007–3023
- Wolf J, Valkov E, Allen MD, Meineke B, Gordiyenko Y, McLaughlin SH, Olsen TM, Robinson CV, Bycroft M, Stewart M, *et al* (2014) Structural basis for Pan3 binding to Pan2 and its function in mRNA recruitment and deadenylation. *EMBO J* 33: 1514–1526
- Workman RE, Tang AD, Tang PS, Jain M, Tyson JR, Razaghi R, Zuzarte PC, Gilpatrick T, Payne A, Quick J, *et al* (2019) Nanopore native RNA sequencing of a human poly(A) transcriptome. *Nat Methods* 16: 1297–1305
- Xiang K & Bartel DP (2021) The molecular basis of coupling between poly(A)-tail length and translational efficiency. *eLife* 10: e66493
- Xu Z, Wei W, Gagneur J, Perocchi F, Clauder-Münster S, Camblong J, Guffanti E, Stutz F, Huber W & Steinmetz LM (2009) Bidirectional promoters generate pervasive transcription in yeast. *Nature* 457: 1033–1037
- Yamashita A, Chang T-C, Yamashita Y, Zhu W, Zhong Z, Chen C-YA & Shyu A-B (2005) Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nat Struct Mol Biol* 12: 1054–1063



## Figure legends

### **Figure 1. A network of RNA-mediated protein interactions links Upf1 to Pab1 and the Lm1-7 complex.**

**(a)** Enrichment of proteins in association with Pab1-TAP was estimated by quantitative mass spectrometry for identified proteins in comparison with their abundance in a total protein extract. Labeled proteins correspond to those for which at least three independent measures were available, proteins were enriched by a factor higher than 2 ( $\log_2$  value of 1) and were robustly detected, as judged by a p-value lower than 0.05 for a t-test (null hypothesis). Ribosomal proteins are indicated in blue.

**(b)** Protein association to Lsm1-TAP, with results processed and displayed similar to panel (a).

**(c)** Proteins enriched with tagged versions of Lsm1, Lsm7, Pat1, Dhh1 and Pab1 at least by a factor of two, were selected. Those having at least two links with the tagged factors (this study, blue arrows) or previously published results (orange arrows), were represented using Cytoscape. The size of the nodes is proportional to the total number of edges (connectivity) and the width of the blue arrows is proportional to the enrichment factor. Tagged protein nodes have red edges. Proteins with a related function, are represented in proximity to each other and annotated.

**(d)** The interaction between Pab1 and Upf1 was tested by co-purification of Pab1-HA with Upf1-TAP in the presence or absence of micrococcal nuclease treatment. A strain lacking the TAP tag on Upf1 was used as a control.

**(e)** Hypothesis: since Pab1 and Lsm1 have different affinities for RNAs with poly(A) tails of different lengths, this suggests that Upf1 associates with both long poly(A) and oligoadenylated mRNAs.

### **Figure S1 related to Figure 1. Enrichment of specific factors in purified complexes.**

Volcano plots represent the enrichment of proteins, relative to the protein abundance in a total extract as measured by label-free mass spectrometry. Results are presented for purifications using TAP tagged versions of Pat1 **(a)**, Lsm7 **(b)**, and Dhh1 **(c)**. Ribosomal proteins are highlighted in blue.

**(d)** Examples of enrichment values across the various purifications for Lsm2, a component of Lsm1-7 and Lsm2-8 complexes and Upf1. As expected, Lsm2 was not identified in the Pab1 purification experiments.

### **Figure 2. Pab1 associates with unstable RNAs that have longer than average poly(A) tails.**

(a) Relative enrichment of NMD substrates, annotated as "XUT/SUT", in association with Rpl16a-TAP, Upf1-TAP and Pab1-TAP, showing the preference of Upf1 and Pab1 to this unstable RNA population (orange), in comparison with other cellular RNA (blue). The red horizontal line indicates the median of the enrichment values, with the first and third quartiles indicated. The p-values correspond to a Wilcoxon rank sum test with continuity correction, N is the number of values in each category.

(b) An analysis similar to (a), comparing the enrichment of pre-mRNA for ribosomal protein genes (intron signal, orange) with the enrichment of mRNA for the same category (exon signal, blue).

(c) Example of signal intensity for RNAs associated with Pab1-TAP, Rpl16a-TAP or Upf1-TAP for a region of the yeast genome that corresponds to SUT141, an unstable RNA, in comparison with the PFK1 stable mRNA (transcribed in proximity but from the opposite strand). The vertical scale is the same for the samples of the same replicated experiment. Pab1 associated RNA signal is depicted in red. The image was obtained with the Integrated Genome Viewer, IGV (Robinson *et al*, 2011).

**Figure S2 related to Figure 2. Pab1 preferentially binds long poly(A) tail versions of cellular RNAs.**

(a) Schematics of the ePAT method (Jänicke *et al*, 2012) for the estimation of poly(A) tail length, which uses the propensity of the Klenow fragment of the *E. coli* DNA polymerase I to extend an RNA molecule on a DNA template (left). As a control, an anchored oligonucleotide that allows an estimation of the size of the 3' end region and transcript end (TVN-PAT, right). "V" represents any of the four nucleotides, except T.

(b) Profiles of poly(A) tail sizes for CCW22, an NMD substrate, in a total extract or in association with Upf1-TAP, Rpl16a-TAP, Lsm1-TAP and Pab1-TAP (purified RNA). Amplification products from ePAT or TVN-PAT reactions were separated on a 2% agarose gel and stained with ethidium bromide. A second downstream transcription termination site leads to the observed profile (marked in blue, "A12'").

(c) Similar to (b), for the HHT2 mRNA. The approximate sizes of the detected poly(A) tails are indicated at the right of the gel image.

**Figure 3. An alternative kinetic model for RNA degradation explains the inverse correlation between poly(A) tail length and mRNA stability.**

**(a)** Average poly(A) tail length shows a negative correlation with RNA half-life. Published poly(A) tail size data (log<sub>2</sub> transformed, Subtelny *et al*, 2014), were represented on the vertical axis as a function of estimated RNA half-life, horizontal axis (log<sub>2</sub> transformed, Miller *et al*, 2011). The Pearson product moment correlation coefficient and its 95% confidence interval are displayed on the graph.

**(b)** Similar to panel (a), but with data for poly(A) size and RNA half-life from other publications (Harrison *et al*, 2015; Chan *et al*, 2018).

**(c)** RNAs for which half life is short, in the first quartile, half life inferior to 7.3 min, or long, in the third quartile, over 19.9 min (Miller *et al*, 2011), were selected to evaluate how their poly(A) size changed in deadenylation mutants, as measured by Nanopore sequencing (Tudek *et al*, 2021). The quartiles of the distributions are indicated, median in red, as well as the p-value of a Wilcoxon rank sum test with continuity correction for the difference between the two populations.

**(d)** Drawing of two models for RNA degradation, one in which deadenylation is required to induce decapping, and another in which decapping can occur on RNAs with any size of poly(A) tail. Three RNA with long (PAI), intermediate (PAm) and short (PAs) poly(A) tails are depicted.  $k_A$  is a kinetic constant for deadenylation and  $k_D$  represents the decapping rate constant.

**(e)** Parameters used to simulate RNA degradation measurements that include poly(A) tail information and kinetic equations using the Antimony syntax (Smith *et al*, 2009). "T" represents all the processes leading to mRNA formation. PAI, PAm and PAs represent concentrations of RNA.

**(f)** Example of simulated RNA degradation curve for the deadenylation-dependent kinetic model, starting from a steady-state situation described by a decapping constant of 0.5min<sup>-1</sup> and a deadenylation constant of 0.05 min<sup>-1</sup>. The relative variation of the amounts of the three forms of RNA is indicated.

**(g)** Examples of simulated half-life values and average poly(A) tail length for the deadenylation-dependent model for RNAs with three values of  $k_D$  (low, dark blue, intermediate, green, and high, yellow) in combination with three values of  $k_A$  (low, circles, intermediate, triangles, and high, squares). Values of poly(A) for the three species were arbitrarily set to 70, 40 and 10 nucleotides. Both axes are logarithmic.

**(h)** Similar to (g), but for the deadenylation-independent kinetic model. The  $k_D$  values were adjusted to obtain half-life and poly(A) size values compatible with published results (see panels **a** and **b**).

**Figure S3 related to Figure 3. Robustness of the negative correlations between poly(A) tail length and half-life of RNA.**

- (a), (b) Negative correlations between published results, similar to the ones shown in **Fig. 3a and 3b**, including results for poly(A) tail size (Harrison *et al*, 2015; Subtelny *et al*, 2014) and estimates of RNA half-life (Chan *et al*, 2018; Miller *et al*, 2011).
- (c) Despite the difference in the experimental strategies used, published half-life and poly(A) tail size (d) are correlated.
- (e) The set of differential equations used to estimate time changes for the three modeled poly(A) species in a deadenylation-dependent model (upper region) and the solutions (lower region). The terms that correspond to steady-state are circled in gray.
- (f) Similar to (e), for deadenylation-independent model.

**Figure 4. An inducible degron system shows that RNA levels are independent from massive poly(A) tail changes.**

- (a) Schematics of the developed inducible degron system in which *O. sativa* Tir1 expression is under the control of an estrogen sensitive promoter. In the presence of the plant auxin hormone indole-3-acetic acid (IAA), Tir1 targets a E3 ubiquitin ligase to the AID domain fused to the protein of interest (POI) for its rapid proteasome degradation.
- (b) Immunoblot estimation of the decrease in the levels of deadenylases targeted by the inducible degron system. Protein degradation was induced by addition of  $\beta$ -estradiol (1  $\mu$ M) and IAA (100  $\mu$ M) and total protein extracts were tested at the indicated time points with anti-FLAG antibodies. The "control" signal comes from a yeast protein that is detected by the antibodies in all samples and serves as a protein load estimate.
- (c) Serial dilution plate growth assay for deletion and inducible degron (ideg) strains affecting deadenylases of the CCR4/NOT complex and the decapping enzyme Dcp2. Growth on rich medium (YPD, left) was compared with growth on medium containing  $\beta$ -estradiol and IAA (right).
- (d) Comparison of the global distribution of median poly(A) tails for mRNAs as detected by Nanopore sequencing in strains depleted for Dcp2 or Ccr4 and Pop2 by treatment with  $\beta$ -estradiol and IAA for 1 hour in liquid medium. The results represent the average of two independent experiments. Red lines correspond to medians of the distributions, with black lines indicating the first and the third quartiles. Indicated p-values correspond to a paired Wilcoxon signed rank test with continuity correction for the null hypothesis.
- (e) Relationship between the measured changes in poly(A) tail length in Dcp2-depleted cells and the increase in the levels of the corresponding mRNAs, compared with a wild type strain. The Pearson product moment correlation coefficient and its 95% confidence interval are displayed on the graph.
- (f) Similar to (e) for the strain depleted for the Ccr4 and Pop4 deadenylases.

**Figure 5. Codon biased reporters are stabilized when translation is inhibited.**

(a) Two new reporters in which RNA synthesis is under the control of a tetOFF system were built using the previously published 100% optimal ("OPT-HIS3") and 10% optimal ("non-OPT-HIS3") coding regions for the HIS3 gene (Radhakrishnan *et al*, 2016). The coding region was inserted upstream the CYC1 terminator region.

Degradation rates for the codon optimal (b) and non-optimal (c) reporters were measured by RT-Q-PCR for the amount of remaining RNA after addition of doxycyclin to the culture medium. An exponential decay curve with a delay factor to account for the action of doxycyclin (Baudrimont *et al*, 2017) was fitted to the points and the estimated half-life and its 95% confidence interval is indicated.

(d) Dose-dependent effect of the translation inhibitor cycloheximide (from 0 to 50 µg/ml) was tested by Northern blotting for optimal (upper panel) and non-optimal (lower panel) HIS3 reporters. SCR1 was used as a loading control.

**Figure S4 related to Figure 5. Depletion of Dhh1 has variable effects on codon-biased reporters.**

(a) Validation of a Dhh1 inducible degron strain by immunoblotting for the FLAG tag of Dhh1 at 0, 30 and 90 minutes after addition of β-estradiol and IAA. "Control" represents a signal that is specific for the anti-FLAG antibodies in yeast extracts.

(b) Optimized and non-optimal HIS3 reporter RNA half-life was estimated by RT-Q-PCR in strains depleted for Dhh1 (3 independent replicated experiments). Estimated half-life and the 95% confidence interval are indicated.

(c) Northern blot equivalent of the experiments indicated in panel (b), with the inclusion of results for the slow growing *dhh1Δ* strain., for the non-optimal HIS3 reporter (upper panel), and the optimized one (lower panel). Please note that the incubation times with doxycyclin are longer for the more stable RNA.

(d) Northern blot estimations of the stabilizing effects of DHH1 deletion and depletion by the inducible degron system for the HIS3 reporters shown to be sensitive to DHH1 (Radhakrishnan *et al*, 2016).

**Figure 6. Non-optimal and optimal HIS3 reporter RNAs are stabilized by depletion of Dcp2, but not by depletion Ccr4 or Pop2.**

Degradation rates for optimized HIS3 (left panel) and non-optimal HIS3 (right panel) were visualized by Northern blot (HIS3r reporter signal) in comparison with an SCR1 control in a wild type strain (a) after depletion of Dcp2 (b), Ccr4 (c), Pop2 (d) or simultaneous depletion of Ccr4 and Pop2 (e). For each situation, the estimated half life and the 95% confidence interval obtained from RT-Q-PCR and the results are presented in **Suppl Fig. S5**.

**Figure S5 in relation with Fig. 6. Half-life of reporter RNA as measured by RT-qPCR.**

Impact of depletion of Dcp2 **(a)**, Pop2 **(b)**, Ccr4 **(c)**, or the double depletion of Ccr4 and Pop2 **(d)** on the estimated half-life of HIS3 reporter RNA, codon-optimized (left panel) or non-optimal (right panel). Indicated estimates correspond to half-life in minutes with a 95% confidence interval. All the experiments were performed independently at least three times. The wild-type situation is presented in **Fig. 5b** and **5c**.

**Figure 7. Deadenylation inhibition increases poly(A) tail sizes of reporter RNA, without an effect on their apparent half-life.**

**(a)** Schematics of the RNase H digestion experiment to evaluate poly(A) tail distribution of the optimal HIS3 reporter RNA by Northern blotting.

**(b)** Evolution of poly(A) tail length after transcription shut off with doxycyclin in a wild type strain (left side) or after depletion of Ccr4 and Pop2 through the inducible degron system (right side). A star indicates samples treated with oligo-dT prior to RNase H digestion. Approximate sizes of the poly(A) tails are indicated.

**(c)** Similar to **(b)**, but for depletion of Ccr4 or Pop2. A wild type steady-state RNA sample is also shown for comparison with **(b)**.

**(d)** Poly(A) tail signal for the experiments presented in **(b)**. The values were normalized to the same maximum signal and are not indicative of the amount of remaining RNA at different time points, only of the distribution. The partially digested signal corresponding to A15 was used to calibrate the plots. The horizontal axis corresponds to distance from the region of the A0 signal.

**(e)** Poly(A) tail signal for the experiments presented in **(c)**, processed as in **(d)**.

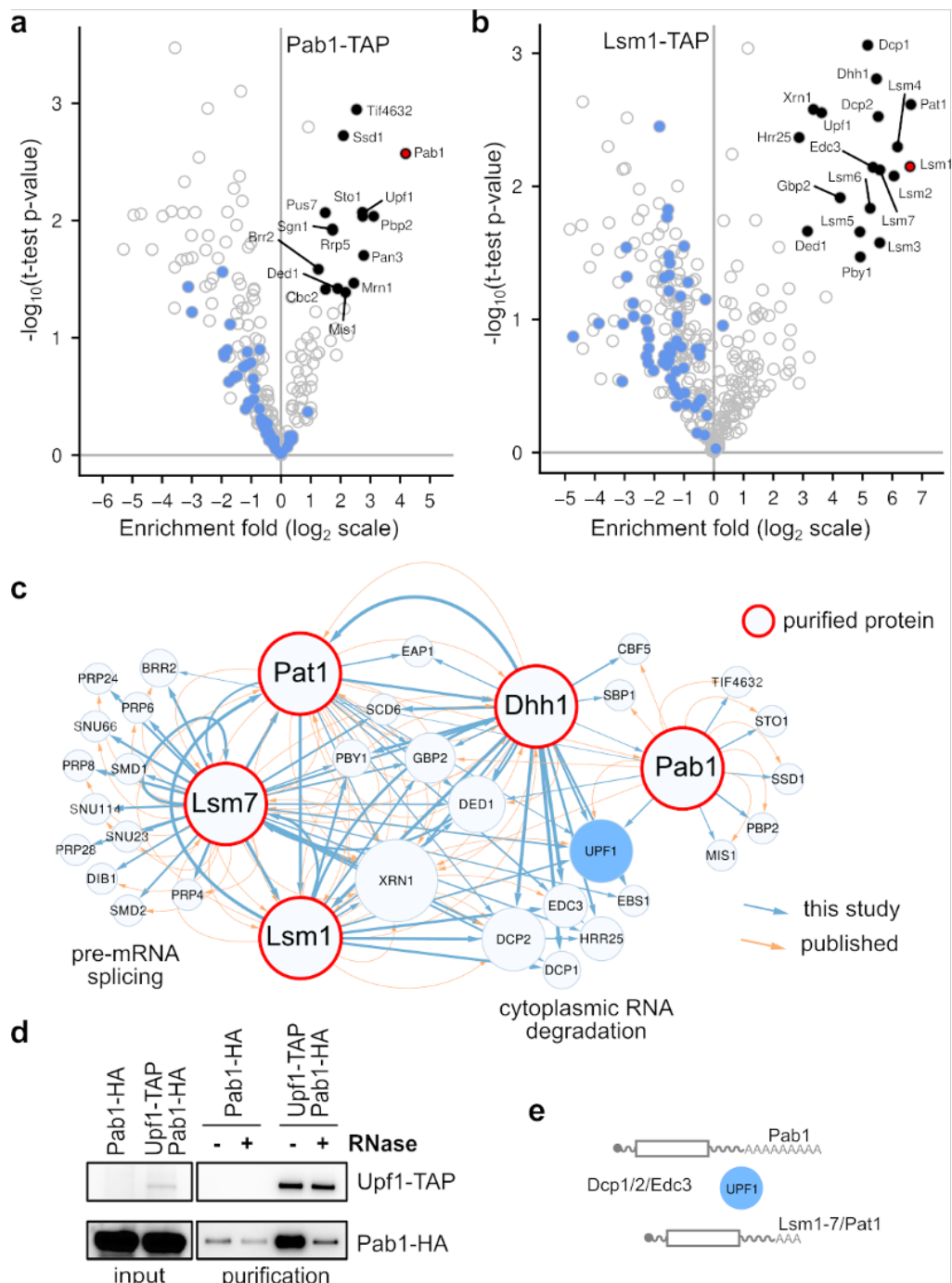
**Figure S6, in relation with Fig. 7. Deadenylation inhibition increases poly(A) tail sizes of reporter RNA.**

**(a)** Schematics of the RNase H digestion experiment to evaluate poly(A) tail distribution of the non-optimal HIS3 reporter RNA by Northern blotting.

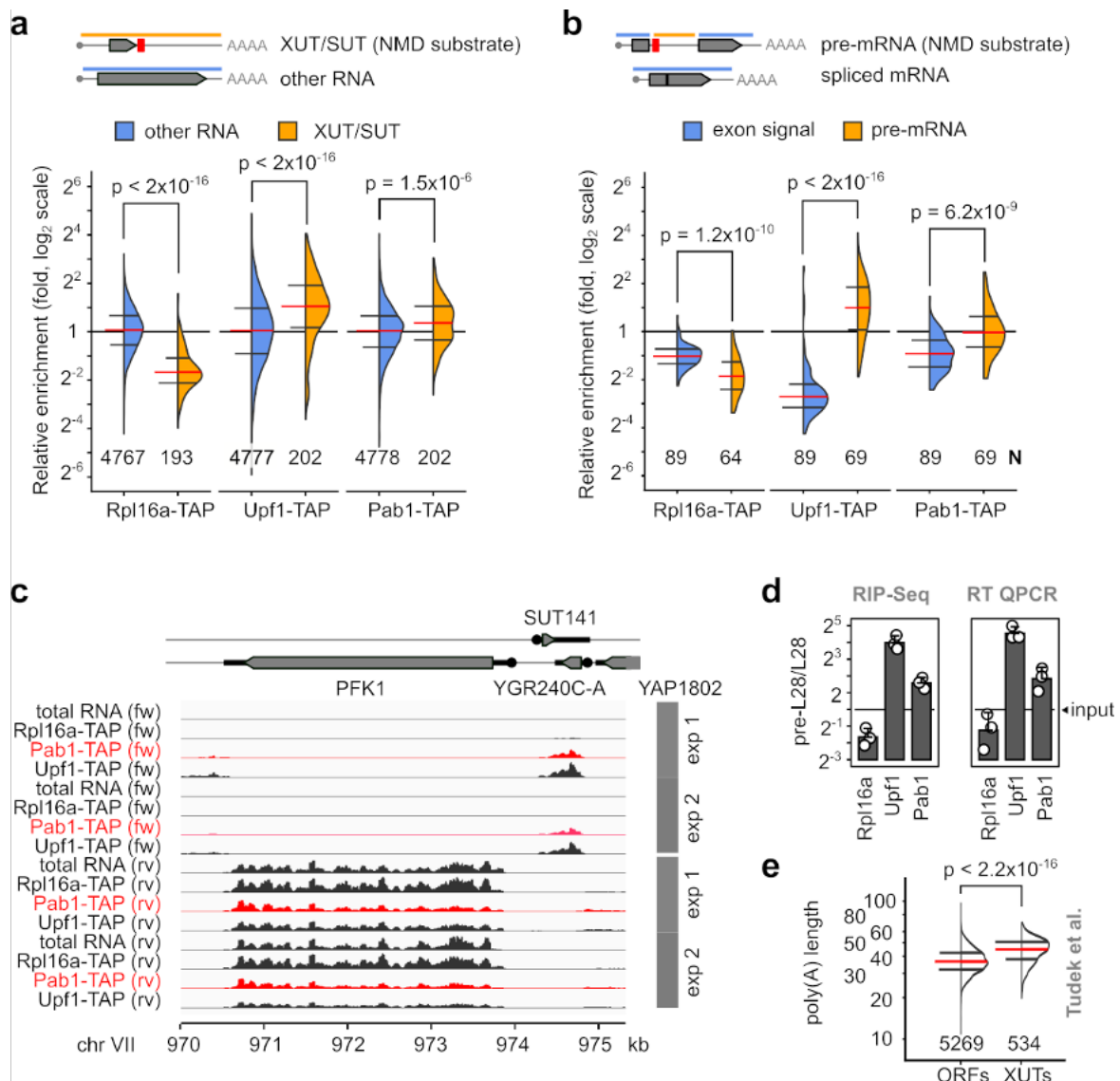
**(b)** Evolution of poly(A) tail length after transcription shut off with doxycyclin in a wild type strain (left side) or after depletion of Ccr4 and Pop2 through the inducible degron system (right side). A star indicates samples treated with oligo-dT prior to RNase H digestion. Approximate sizes of the poly(A) tails are indicated.

**(c)** Similar to **(b)**, but for depletion of Ccr4 or Pop2. A wild type steady-state and double depletion RNA samples are also shown for comparison with **(b)**.

# Figures

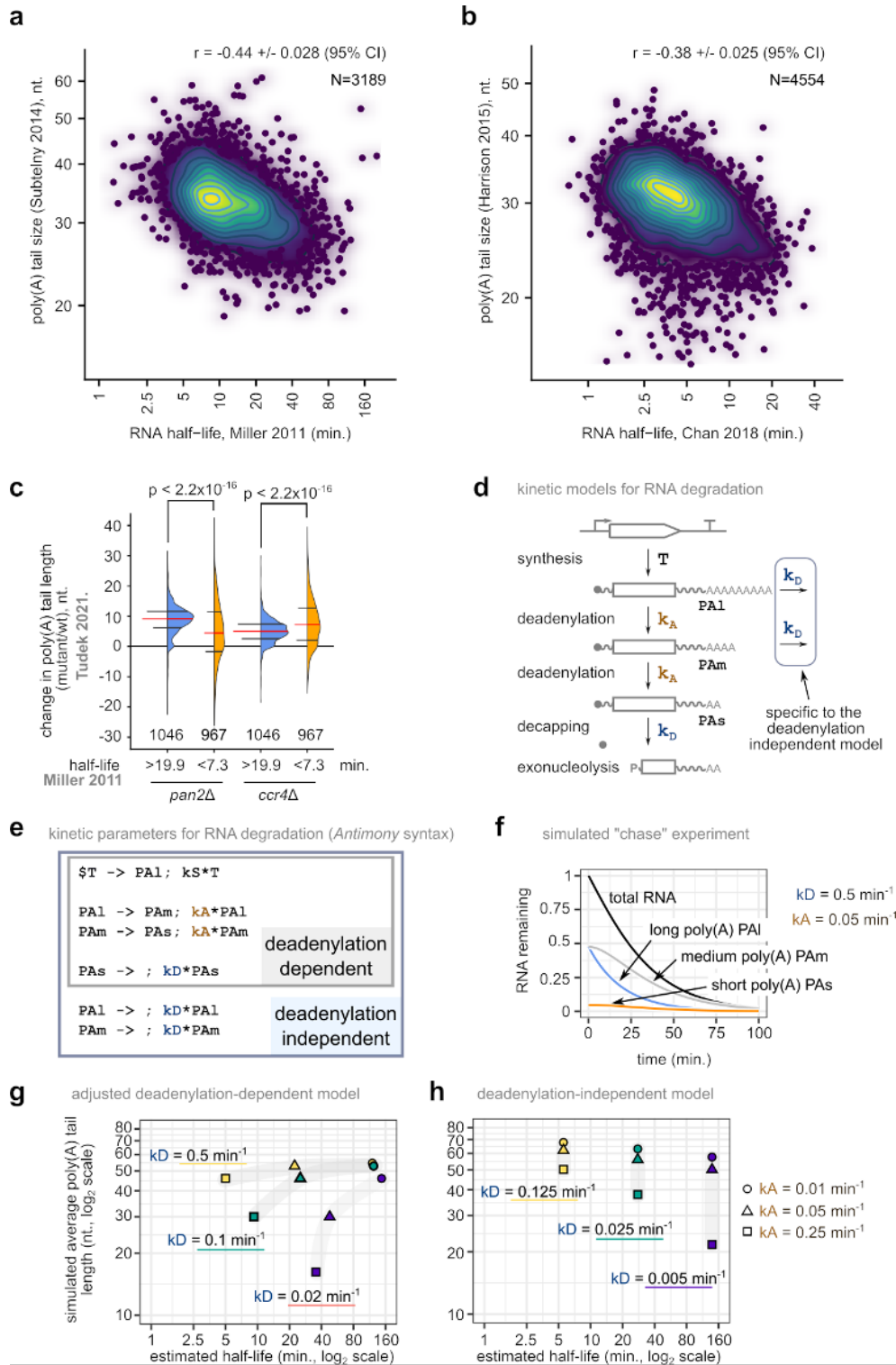


**Figure 1. A network of RNA-mediated protein interactions links Upf1 to Pab1 and the Lm1-7 complex.**

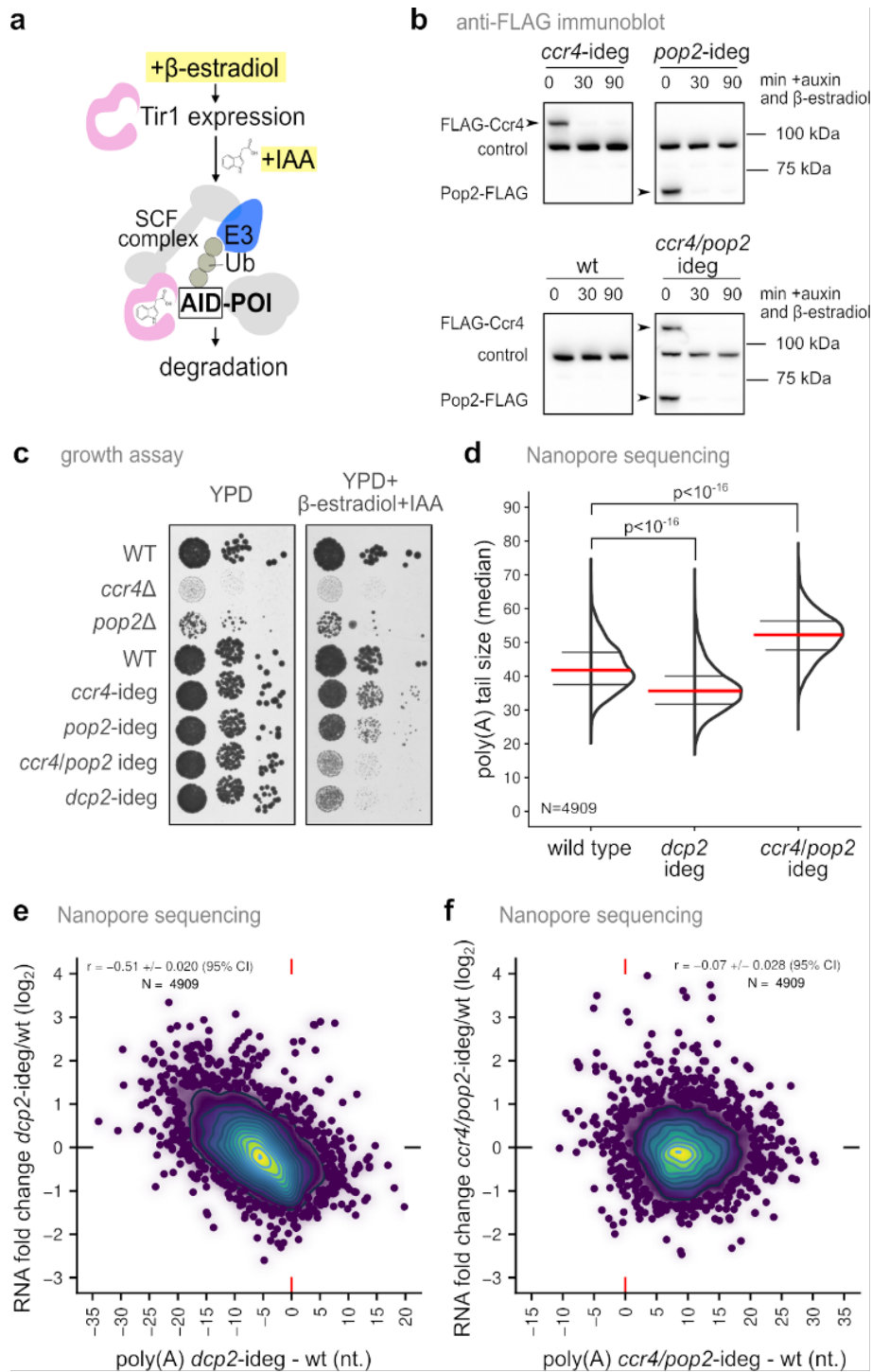


**Figure 2. Pab1 associates with unstable RNAs that have longer than average poly(A) tails.**





**Figure 3. An alternative kinetic model for RNA degradation explains the inverse correlation between poly(A) tail length and mRNA stability.**



**Figure 4. An inducible degron system shows that RNA levels are independent from massive poly(A) tail changes.**

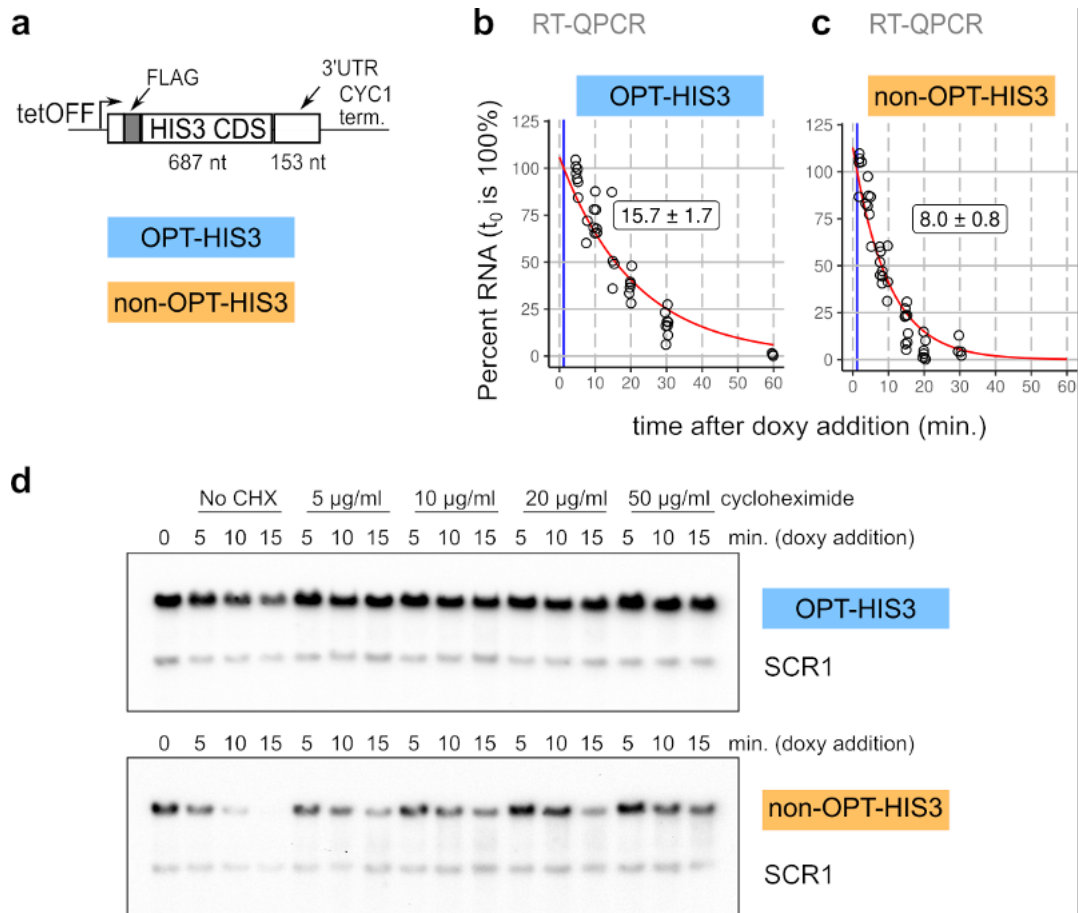
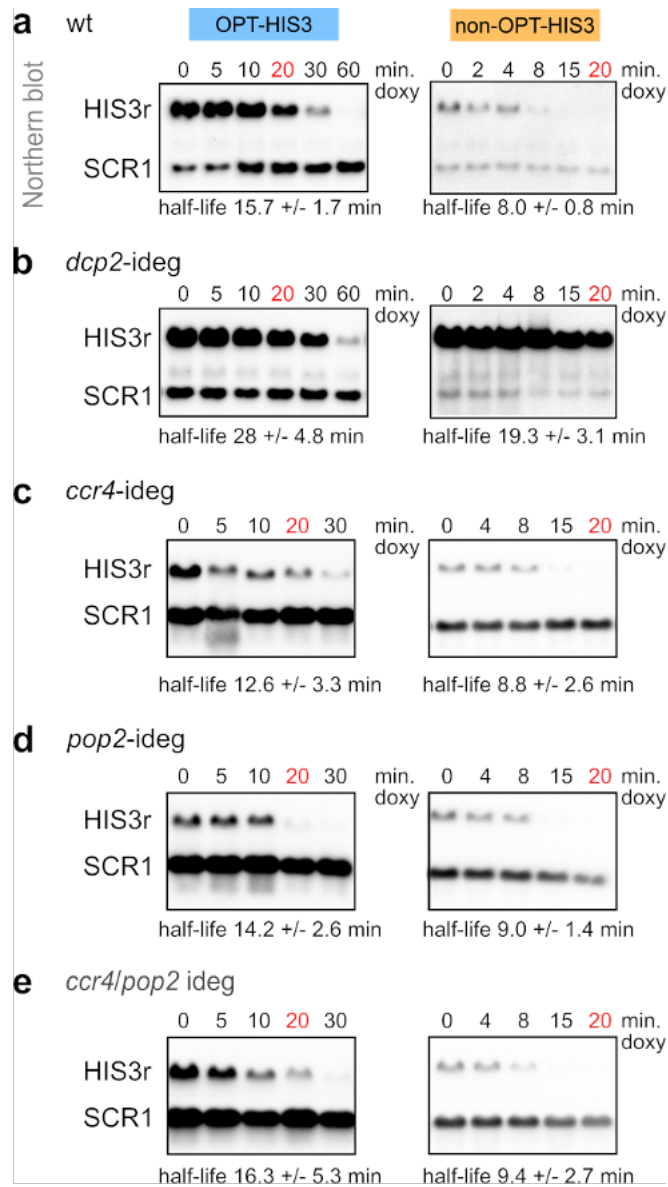
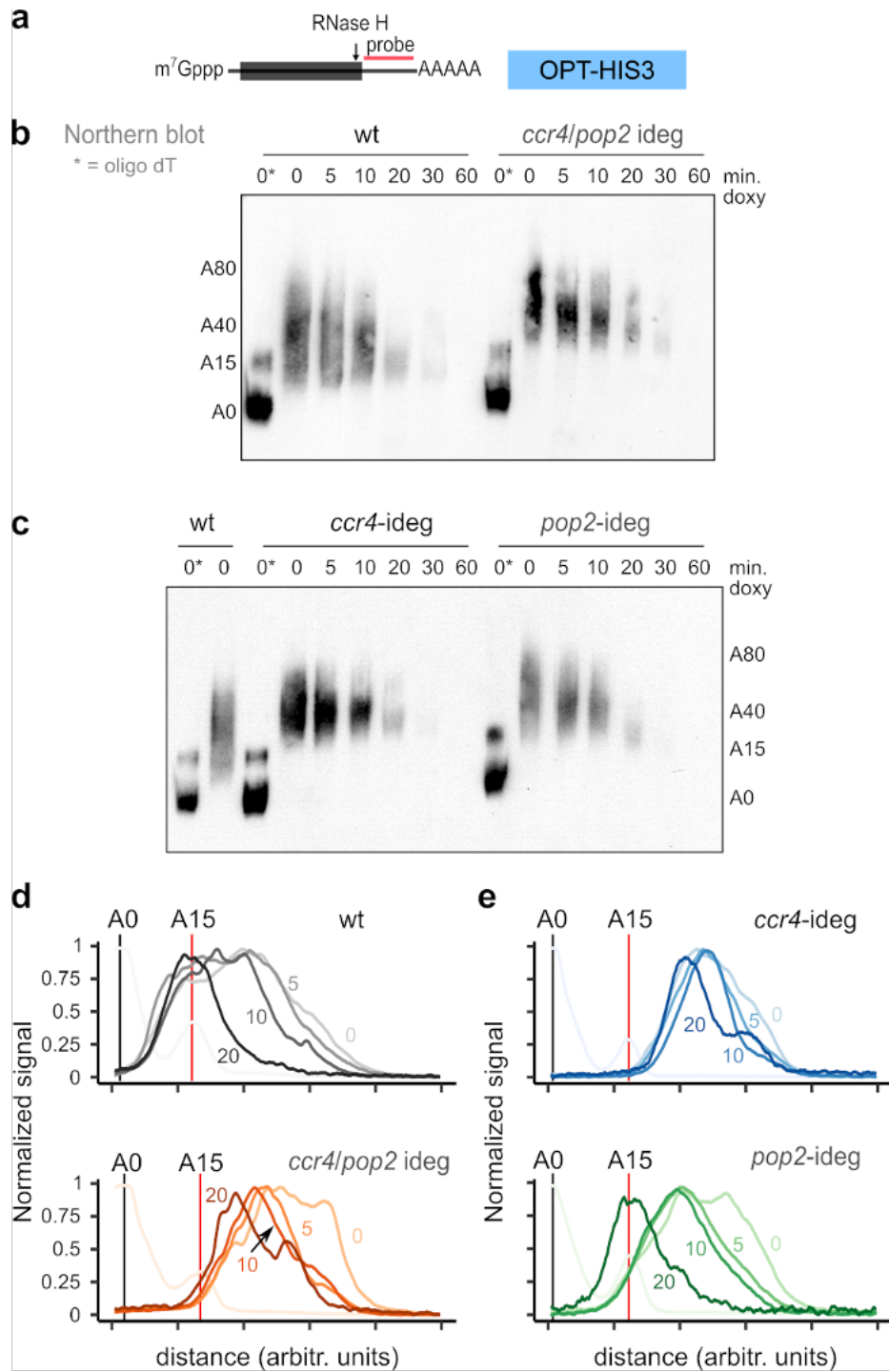


Figure 5. Codon biased reporters are stabilized when translation is inhibited.



**Figure 6. Non-optimal and optimal HIS3 reporter RNAs are stabilized by depletion of Dcp2, but not by depletion Ccr4 or Pop2.**



**Figure 7. Deadenylation inhibition increases poly(A) tail sizes of reporter RNA, without an effect on their apparent half-life.**

## Supplementary figures

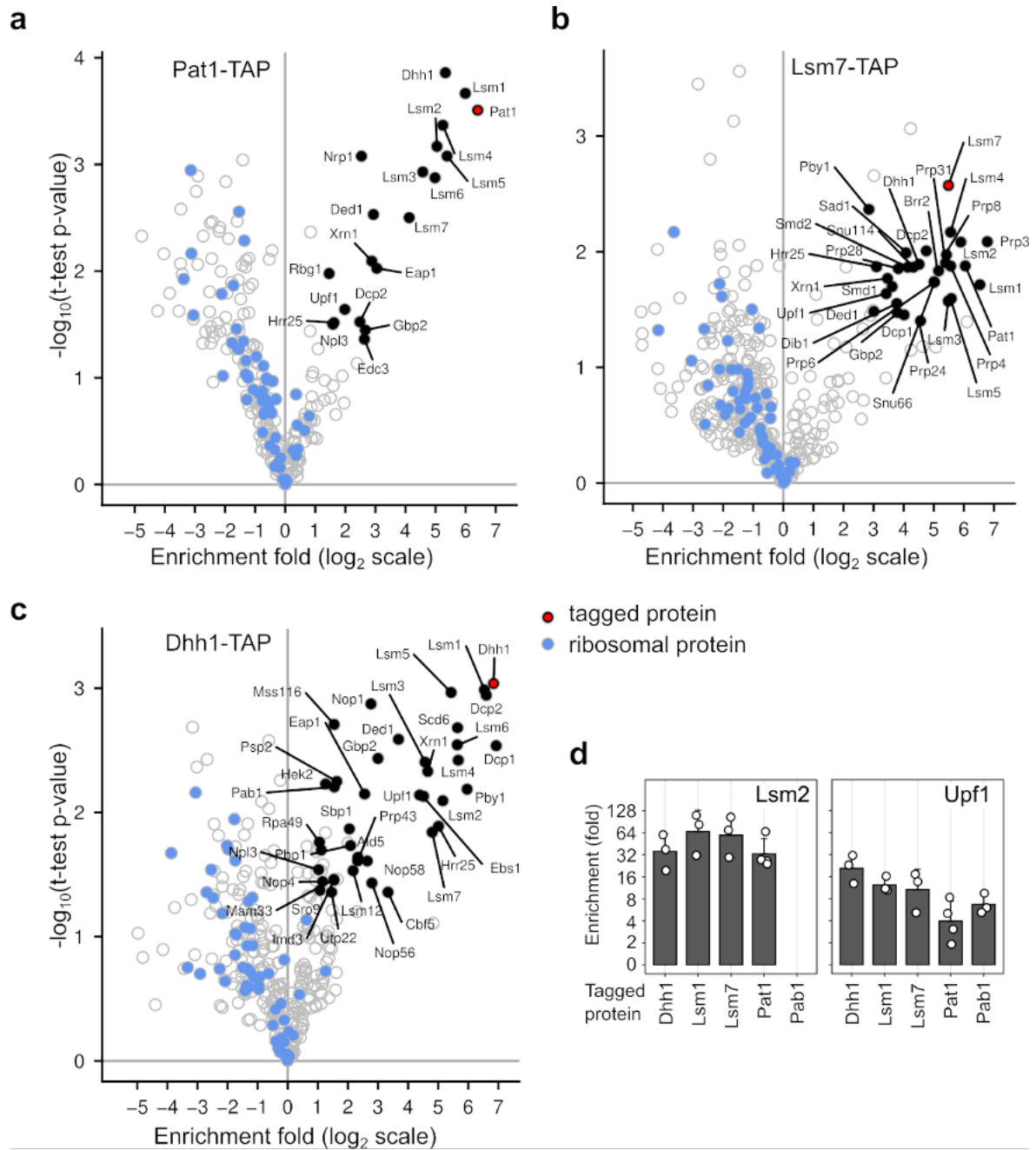


Figure S1 related to Figure 1. Enrichment of specific factors in purified complexes.

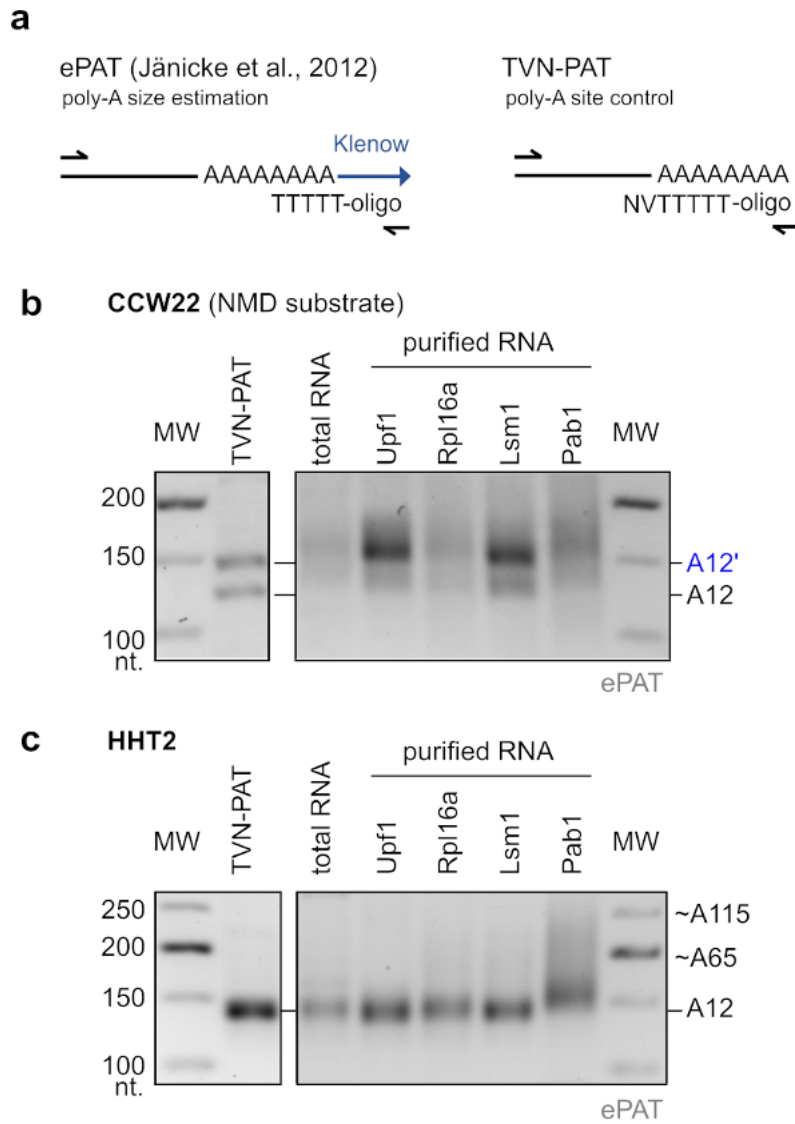
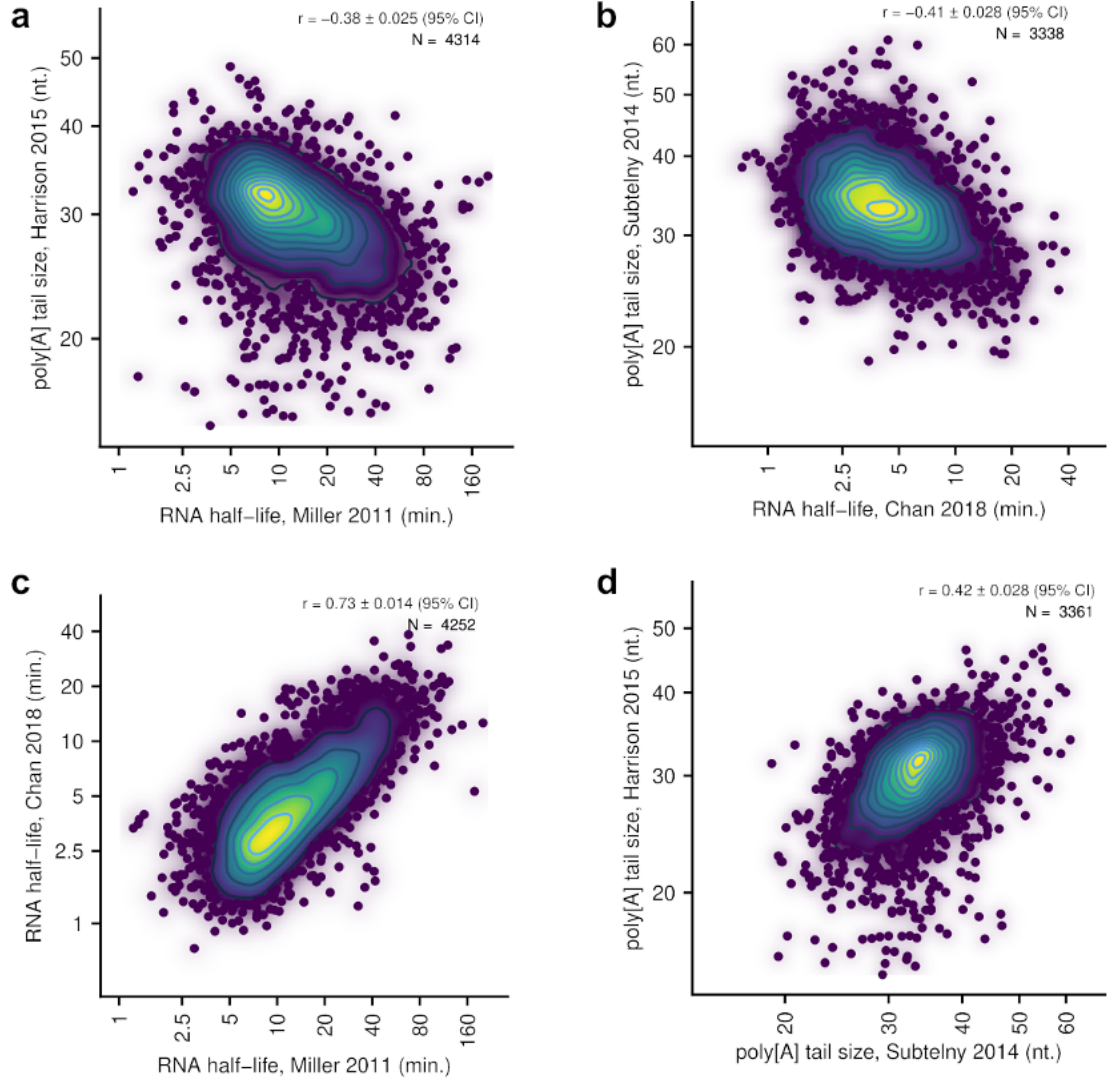


Figure S2 related to Figure 2. Pab1 preferentially binds long poly(A) tail versions of cellular RNAs.





**e** Deadenylation-dependent model

$$\begin{aligned}
 \frac{d}{dt} \text{PAI}(t) &= T - k_A \cdot \text{PAI}(t) \\
 \frac{d}{dt} \text{PAm}(t) &= k_A \text{PAI}(t) - k_A \text{PAm}(t) \\
 \frac{d}{dt} \text{PAs}(t) &= k_A \text{PAm}(t) - k_D \text{PAs}(t)
 \end{aligned}$$

differential equations

---


$$\begin{aligned}
 \text{PAI}(t) &= \left(\frac{T}{k_A}\right) - \frac{T e^{-k_A t}}{k_A} \\
 \text{PAm}(t) &= -T t e^{-k_A t} - \frac{T e^{-k_A t}}{k_A} + \left(\frac{T}{k_A}\right) \\
 \text{PAs}(t) &= -\frac{T k_A^2 e^{-k_D t}}{k_D^3 - 2k_A k_D^2 + k_A^2 k_D} - \frac{T k_A t e^{-k_A t}}{k_D - k_A} - \frac{(T k_D - 2T k_A) e^{-k_A t}}{k_D^2 - 2k_A k_D + k_A^2} + \left(\frac{T}{k_D}\right)
 \end{aligned}$$

solutions

**f** Deadenylation-independent model

$$\begin{aligned}
 \frac{d}{dt} \text{PAI}(t) &= T - k_A \text{PAI}(t) - k_D \text{PAI}(t) \\
 \frac{d}{dt} \text{PAm}(t) &= k_A \text{PAI}(t) - k_D \text{PAm}(t) - k_A \text{PAm}(t) \\
 \frac{d}{dt} \text{PAs}(t) &= k_A \text{PAm}(t) - k_D \text{PAs}(t)
 \end{aligned}$$

differential equations

---


$$\begin{aligned}
 \text{PAI}(t) &= \left(\frac{T}{k_D + k_A}\right) - \frac{T e^{-(k_D + k_A)t}}{k_D + k_A} \\
 \text{PAm}(t) &= -\frac{T k_A t e^{-(k_D + k_A)t}}{k_D + k_A} - \frac{T k_A e^{-(k_D + k_A)t}}{k_D^2 + 2k_A k_D + k_A^2} + \left(\frac{T k_A}{k_D^2 + 2k_A k_D + k_A^2}\right) \\
 \text{PAs}(t) &= \frac{T k_A t e^{-(k_D + k_A)t}}{k_D + k_A} + \frac{(T k_D + 2T k_A) e^{-(k_D + k_A)t}}{k_D^2 + 2k_A k_D + k_A^2} - \frac{T e^{-k_D t}}{k_D} + \left(\frac{T k_A^2}{k_D^3 + 2k_A k_D^2 + k_A^2 k_D}\right)
 \end{aligned}$$

solutions

**Figure S3 related to Figure 3. Robustness of the negative correlations between poly(A) tail length and half-life of RNA.**



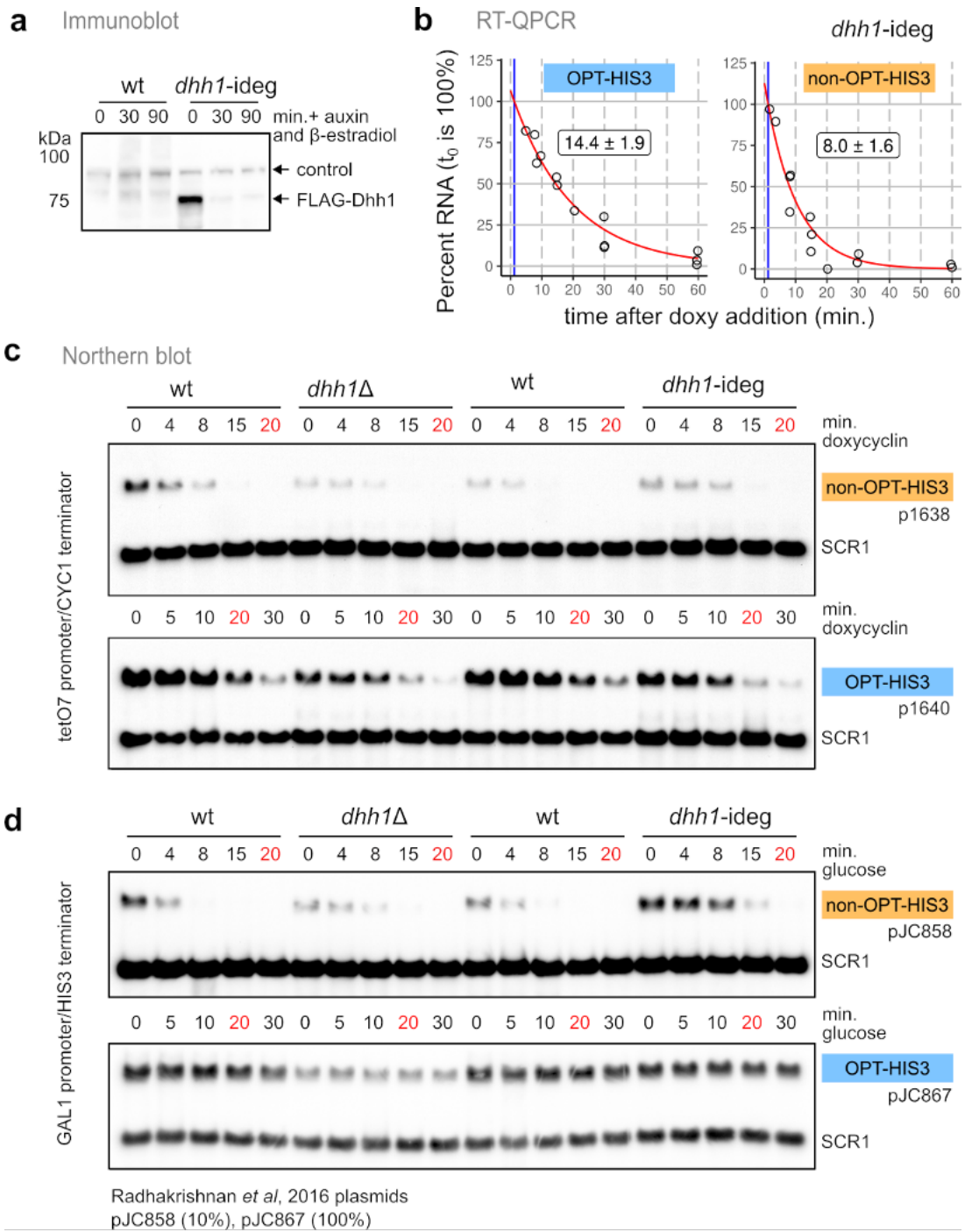


Figure S4 related to Figure 5. Depletion of Dhh1 has variable effects on codon-biased reporters.

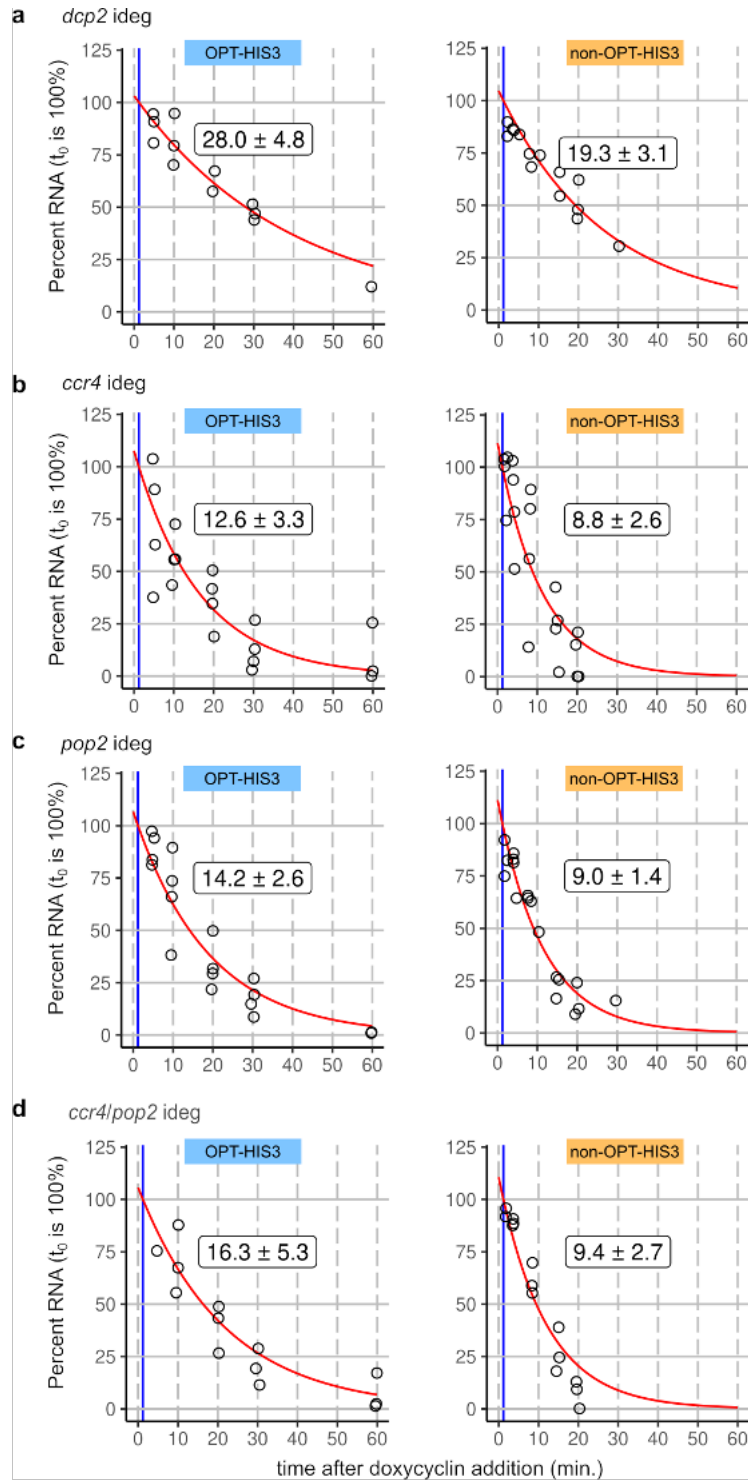
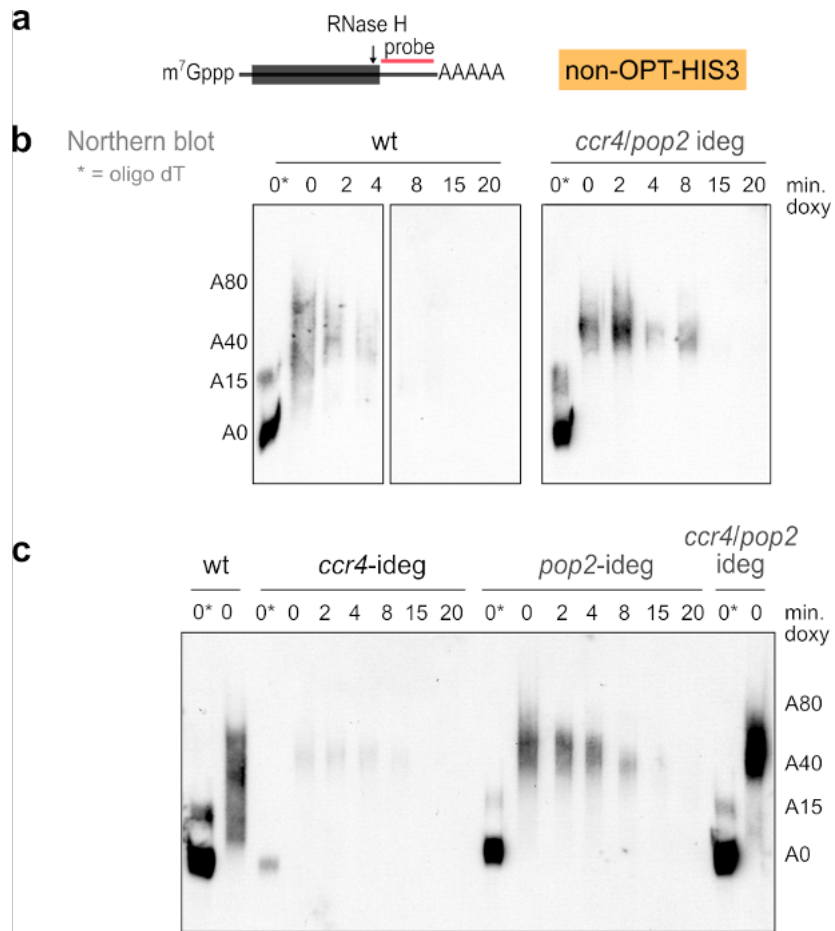


Figure S5 in relation with Fig. 6. Half-life of reporter RNA as measured by RT-qPCR.



**Figure S6, in relation with Fig. 7. Deadenylation inhibition increases poly(A) tail sizes of reporter RNA.**