

In vitro biochemical study of Nonsense-mediated mRNA Decay complexes in Saccharomyces cerevisiae

Nadia Ruiz Gutierrez

▶ To cite this version:

Nadia Ruiz Gutierrez. In vitro biochemical study of Nonsense-mediated mRNA Decay complexes in Saccharomyces cerevisiae. Structural Biology [q-bio.BM]. Sorbonne Université, 2023. English. NNT: 2023SORUS684 . tel-04591447

HAL Id: tel-04591447 https://theses.hal.science/tel-04591447

Submitted on 28 May 2024

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.



Ecole Doctorale Complexité du Vivant (ED 515) Sorbonne Université

In vitro biochemical study of Nonsense-mediated mRNA Decay complexes in *Saccharomyces cerevisiae*

Thesis prepared at the Institut de Biologie de l'Ecole Normale Supérieure and publicly defended on September 22, 2023 for the title of PhD by

Nadia RUIZ GUTIERREZ

Under the supervision of **Dr. Hervé LE HIR**

Research director (DR) at the Institut de Biologie de l'Ecole Normale Supérieure

Members of the jury:

Isabelle BEHM-ANSMANT, CR at Biopole de l'Université de Lorraine, Nancy - *Reporter* Jan KADLEC, CR at Institut de Biologie Structurale, Grenoble - *Reporter* Dominique WEIL, DR at Institut de Biologie Paris Seine, Paris - *President* Marc BOUDVILLAIN, DR at Centre de Biologie Moléculaire, Orléans - *Examiner* David GATFIELD, Professor at Université de Lausanne, Switzerland - *Examiner* Marina PINSKAYA, CR at Institut Curie, Paris - *Examiner*

Table of contents

Table of contents	3
Résumé de la thèse	6
List of Figures, Nomenclature & Abbreviations	7
Acknowledgements	10
Introduction	13
Chapter 1: About the importance of quality control systems	17
1) The cell is not as perfect as it seems	17
a) Brief overview of gene expression step by step	17
b) Sources of errors	19
2) Errors can be fatal	20
a) (Not) all errors have consequences	20
b) A hidden layer of complexity	20
3) First responders act in the nucleus	22
Chapter 2: Three strikes and you're out, the three-step process of cytoplasmic mRNA qualit	У
controls	23
1) The beginning of the end: it all starts with translation	23
a) Translation mechanism at a glance	23
b) What is "normal" and what is not: aberrant mRNA decay pathways	25
2) Recruitment of specialized factors	27
a) Recruitment by faulty termination	27
b) Recruitment of RNA binding proteins to specific RNA sequences or structures	27
3) Pulling the trigger on RNA decay	28
a) Starting from the tail: deadenylation and 3'-5' decay	29
b) Starting from the head: decapping and 5'-3' decay	30
c) Starting from the middle: endonucleolytic decay	32
d) P-bodies: mRNA degradation or translation repression?	32
Chapter 3: Diving deeper, molecular basis of NMD	35
1) A brief history of NMD research	35
2) NMD in yeast: detect, recruit and degrade	35
a) The NMD holy trinity: Upf1/Upf2/Upf3	35
b) Implication of the Upf proteins and translation termination	36
c) The Faux 3' UTR model for yeast NMD	36
3) NMD in metazoans: a higher level of complexity	38
a) A more diverse Upf core complex	38
b) Metazoan specific NMD factors	39
c) Current mechanistic view of NMD in humans	41
d) The many variations of NMD in eukaryotes	43
4) Technological advances and current approaches to study NMD	45
Chapter 4: NMD, more than just molecular housekeeping	47
1) What does NMD recognize as a target?	47

a) Beyond quality control: whole transcriptome regulation	
b) NMD-triggering characteristics	47
2) Physiological role and importance of NMD	
a) Roles in cellular metabolism	
b) NMD efficiency modulation	
3) NMD in health and disease, a double edged sword	
a) Genetic diseases	49
b) Neurodevelopmental disorders	49
c) A dual role in cancer	50
d) Therapeutics	50
e) Antiviral response	51
Chapter 5: Upf1, the cornerstone of NMD	53
1) Identification: from fortuitous genetic screens to homology	53
2) Structure: Upf1 is a conserved tiny molecular engine	53
a) Helicases are ubiquitous	53
b) Upf1: a Super Family 1 helicase	55
c) Upf1 is conserved among eukaryotes	56
3) Activity: substrate whisperer or recruitment platform?	
a) The catalytic activities of Upf1	
b) Upf1 activity is essential for NMD	58
4) Interactions and regulation: taming the master orchestrator	58
a) Upf1 is a hub for protein-protein interactions	
b) Regulation by phosphorylation	58
c) Regulations by partners	60
5) Beyond NMD: Upf1's expanding repertoire of cellular functions	61
Gaps in the knowledge	63
General questions about NMD	63
How is RNA decay triggered and modulated during NMD?	64
Scope of the thesis	64
Novel Upf1-bound complex in S. cerevisiae	64
Towards a universal NMD model: "Detector & Effector"	66
PhD project	67
Results	69
Part 1: Incompatibility of Upf1 and Upf2 interaction with Upf1-decapping compl	.ex in yeast
provides new insights into NMD dynamics	71
Introduction	71
Results	72
Dcp2 directly interacts with Upf1 N-terminal CH domain	72
Upf1 makes direct contact with both Nmd4 and Ebs1	75
Reconstitution of Upf1-containing decapping complex	76
Yeast Upf1-Upf2 interaction is homologous to the human one	78

Dcp2 and Upf2 compete for interaction with Upf1	80
Upf1 ATPase activity and RNA binding differentially regulated within the Upf1-23 a	nd
Upf1-decapping complexes	
Part 2: Single molecule characterization of yeast Upf1	85
1) Magnetic tweezers for the study of Upf1 helicase activity	
a) The intricate nature of helicase activity	85
b) Single-molecule approaches to study helicases	85
c) Magnetic tweezers experimental setup for helicase activity monitoring	
2) Single molecule studies reveal different autoregulations of human and yeast Upf1 on	RNA 89
a) Background and aim of the study	89
b) Yeast Upf1 CH domain stabilizes the helicase core grip on RNA	90
c) The behavior of the yeast Upf1 helicase core differs on DNA and RNA	93
3) Discussion	95
Part 3: RNA degradation by the Upf1-decapping complex components	97
1) Nmd4 as a potential active endonuclease	
2) Development of a quantitative decapping assay to study decapping regulation	
Discussion & Perspectives	101
1) Structure of the Upf1-decapping complex	103
a) Upf1 structure: human versus yeast	103
b) An intricate web of interactions within the Upf1-decapping complex	105
c) CryoEM for solving the structure	106
2) An evolutionary conserved mechanism?	107
a) Nmd4 as a potential active endonuclease	
b) The role of phosphorylation in yeast NMD	108
3) Revisiting the mechanistic details of yeast NMD	109
a) What is the role of Upf1?	109
b) Correlation between Upf1 speed, NMD efficiency and P-body formation	110
c) Upf2 recruitment temporality	
TL:DR (too long; didn't read)	
Materials & Methods	115
Bibliography	127
Annex: Novel approaches to study helicases using magnetic tweezers	155

Résumé de la thèse

La fidélité de l'expression génétique et le contrôle de la qualité de l'ARN sont essentiels au maintien de l'homéostasie cellulaire. Dans les cellules eucaryotes, la voie Nonsense-mediated mRNA Decay (NMD) est un système de surveillance essentiel chargé d'identifier et d'éliminer les molécules d'ARN défectueuses. Au cœur de cette voie se trouve le complexe Upf, composé des protéines Upf1, Upf2 et Upf3, qui coordonne la reconnaissance et la dégradation des transcrits ciblés par le NMD. Parmi ces composants, Upf1, une ARN hélicase, est le principal orchestrateur de la NMD. Son domaine hélicase (HD) est flangué d'un domaine régulateur riche en cystéine et en histidine (CH) et d'un domaine C-terminal désordonné. Récemment, des purifications d'affinité des complexes liés à Upf1 suivies de spectrométrie de masse quantitative ont permis l'identification du complexe Upf1-decapping chez Saccharomyces cerevisiae. Il comprend entre autres les protéines de decapping Dcp1 et Dcp2, et les protéines Nmd4, Ebs1. Ici, nous avons utilisé une approche in vitro pour étudier la dynamique des interactions complexes entre Upf1 et ses partenaires. En utilisant des protéines recombinantes purifiées des protéines d'intérêt, nous avons effectué des tests d'interaction pour identifier les interactions directes. Nous avons ainsi démontré que Dcp2 interagit directement avec le domaine CH d'Upf1, tandis que Nmd4 et Ebs1 interagissent avec des régions distinctes d'Upf1. Toutes ces interactions étaient compatibles, ce qui suggère la formation d'un complexe stable Upf1-decapping et d'un réseau complexe d'interactions. En outre, Upf2 interagit également avec le domaine CH, de manière similaire aux homologues humains, mais incompatible avec la liaison de Dcp2. Au sein des complexes liés à Upf2, Dcp2 ou Nmd4, la liaison à l'ARN et les activités ATPase d'Upf1 sont régulées différentiellement. En outre, nous avons exploré les mécanismes d'autorégulation et les propriétés fonctionnelles de l'activité d'Upf1 de la levure et de l'homme sur l'ARN et l'ADN. En utilisant une combinaison de techniques biochimiques et biophysiques, nous avons effectué une analyse comparative complète à l'échelle de la molécule unique à l'aide de pinces magnétiques. Cette étude a mis en évidence des comportements autorégulateurs contrastés de l'activité d'Upf1 sur l'ARN. Chez la levure, la présence du domaine CH renforce l'activité de l'hélicase, alors que chez l'homme, elle a des effets répressifs. Ces résultats mettent en lumière les variations spécifiques à chaque espèce dans les mécanismes de surveillance de l'ARN et soulignent l'interaction complexe entre les différents domaines de l'hélicase Upf1 et ses partenaires. Dans l'ensemble, cette étude approfondie a fourni des informations précieuses sur l'importance fonctionnelle de l'hélicase Upf1 et ses implications dans la surveillance de l'ARN, contribuant ainsi à une meilleure compréhension du métabolisme de l'ARN.

List of Figures

Introduction

Chapter 1 Figure 1: Eukaryotic gene expression steps Figure 2: Structure and characteristics of a mature mRNA Figure 3: RNA classification and pervasive transcription Chapter 2 Figure 4: Eukaryotic translation steps Figure 5: mRNA characteristics triggering translation-dependent quality control systems Figure 6: Recognition mechanisms for specific RNA decay pathways Figure 7: General cytoplasmic mRNA decay pathways Figure 8: Structural domains of yeast decapping complex Dcp1/Dcp2 Figure 9: Endonucleolytic decay of mRNA Chapter 3 Figure 10: Structural domains of the yeast Upf proteins Figure 11: Faux 3' UTR model for yeast NMD Figure 12: Structural domains and solved structures of the human Upf proteins Figure 13: Structural domains of the Smg proteins Figure 14: EJC-dependent model for NMD Figure 15: Conservation of NMD factors among eukaryotes Chapter 4 Figure 16: NMD triggering characteristics Chapter 5 Figure 17: Cellular roles of eukaryotic RNA helicases Figure 18: Yeast and human Upf1 structure Figure 19: Phylogenetic analysis of Upf1 Figure 20: Yeast and human Upf1 helicase core activity Figure 21: Upf1 interaction network Figure 22: Human Upf1 modulation by Upf2 and Stau1 Figure 23: Yeast Upf1 is involved in two mutually exclusive complexes Figure 24: Detector and Effector model for yeast NMD Results Part 1

Figure 1: Dcp2 directly interacts with Upf1 CH domain through two Upf1-Binding domains

Figure 2: Nmd4 and Ebs1 directly interact with Upf1

Figure 3: Upf1-decapping complex reconstitution

Figure 4: Direct interaction of Upf2 with Upf1 CH domain

Figure 5: Upf2 and Dcp2 compete for binding to Upf1-CH domain

Figure 6: Upf1 activity is modulated by its partners

Figure S1: UBD1 and UBD2 equally contribute to Upf1 binding

Figure S2: Ebs1 directly interacts with Nmd4 and Dcp1

Part 2

Figure 25: Magnetic tweezers experimental setup

Figure 26: Yeast and human Upf1 are oppositely regulated by the CH domain

Figure 27: Yeast Upf1 helicase activity is enhanced by the CH domain

Figure 28: The activity is not influenced by tag nor force

Figure 29: Yeast Upf1 helicase core is more processive in DNA

Part 3

Figure 30: Magnetic tweezers bead loss after Nmd4 injection

Figure 31: Bulk in vitro endonuclease assays

Figure 32: A quantitative decapping assay

Discussion & Perspectives

Figure 33: Proposed model for NMD in yeast

Materials & Methods

Figure 34: Phosphorylated primer cloning technique

Figure 35: Gibson assembly cloning technique

Figure 36: ATPase assay

<u>Nomenclature</u>

To maintain consistency throughout this manuscript, genes will be denoted in all lowercase italic letters (e.g. *upf1*), while proteins will be indicated with an initial capital letter and subsequent lowercase letters (e.g. Upf1) unless standard nomenclature has been defined (e.g. eIF4G) or it refers to a complex or group of proteins (e.g. PABP). Additionally, species names will be written in italics (e.g. *S. cerevisiae*) and abbreviations will be specified by underlining the meaningful letters (e.g. NMD <u>nonsense-mediated mRNA decay</u>). The nomenclature of recombinant proteins used in this manuscript is always as follows: TagNt-Protein_truncation-TagCt. For example, CBP-Upf1_FL-His corresponds to the Upf1 protein in Full Length version with the CBP tag in N-ter (<u>N-ter</u>minal) and the 6-His tag in C-ter (<u>C-ter</u>minal). Hexahistidine (6-His) tags (usually fused in Ct) will not be mentioned unless the protein is exclusively fused to said tag.

Abbreviations

Biological Notions

3' UTR: 3' untranslated region ARE: AU-rich element CDE: constitutive decay element CUT: cryptic unstable transcripts DNA: desoxyribonucleic acid iTSS: Internal Transcription Start Site IncRNA: long non coding RNA miRNA: micro RNA mRNA: messenger RNA mRNP: messenger ribonucleoprotein MUT: meiotic-specific noncoding transcripts NA: nucleic acids ncRNA: non coding RNA ORF: open reading frame PolII: polymerase II PROMPT: promoter upstream transcripts PTC: premature termination codon **RBP: RNA binding protein** RNA: ribonucleic acid rRNA: ribosomal RNA siRNA: small interference RNA

snoRNA: small nucleolar RNA snRNA: small nuclear RNA SUT: stable unannotated transcripts TERRA: telomeric repeat-containing RNAs tRNA: transfer RNA uaRNA: upstream antisense RNA uORF: upstream ORF WT: wild type XUT: XRN1-dependent unannotated transcripts **Biological pathways:** AMD: ARE-mediated decay GMD: glucocorticoid-receptor mediated mRNA decay HMD: replication-dependent histone mRNA decay NGD: no-go decay NMD: nonsense-mediated mRNA decay NSD: no-stop decay RIDD: regulated Ire1-dependent decay RMD: regnase1-mediated decay RNAi: RNA interference mRNA decay

RoMD: roguin-mediated decay SMD: staufen mediated mrna decay SMD: staufen-mediated decay Species: A. thaliana: Arabidopsis thaliana C. albicans: Candida albicans C. elegans: Caenorhabditis elegans D. melanogaster: Drosophila melanogaster D. rerio: Danio rerio G. lamblia: Giacardia lamblia H. sapiens: Homo sapiens S. cerevisiae: Saccharomyces cerevisiae S. pombe: Schizosaccharomyces pombe T. brucei: Trypanosoma brucei T. thermophilia: Tetrahymena thermophilia Proteins & Complexes: CBC: cap binding complex Ccr4: carbon catabolite repression Dcp: decapping proteins DECID : decay inducing complex Ebs1: Est1-like Bcy1 suppressor 1 Edc: enhancer of decapping eEF: eukaryotic elongation factor elF: eukaryotic initiation factor EJC: exon junction complex eRF1/3: eukaryotic release factors 1/3 Est1: ever shorter telomeres protein 1 HITT: histidine triad LSm: like-Sm Magoh: mago nashi human homologue Nam: nuclear accommodation of mitochondria Not: negative regulator of transcription PABP: poly(A) binding protein Pan2/3: poly(A) nuclease PIKK: phosphatidylinositol 3-kinase related kinase Pp2A: phosphatase protein 2A Rent: regulator of nonsense transcripts **RISC: RNA-induced silencing complex** RRM: RNA recognition motif SF1: super family 1 (of helicases) SKI: super killer proteins Slbp: stem loop binding protein Smg: supresors with morphogenic effect on genitalia

SURF: SMG1-Upf1-release ractors Upf: up-frameshift Xrn1: 5'-3' exoribonuclease 1 Y14/Rbm8A: RNA binding motif protein 8A Protein domains: 14-3-3: phosphopeptide binding C-ter: C-terminal CH: cysteine/histidine domain Ct: C-terminal domain EVH1: Ena/Vasp homology 1 FL: full length HD: helicase domain HI M: helical leucine rich motif N-ter: N-terminal MIF4G: middle domain of eIF4G PIN: PilT-amino-terminal domain S/T-Q: serine threonine glutamine repeats **Protein tags:** 6His: 6 histidine CBP: calmodulin binding peptide GST: glutathione S-transferase TS: twin strep Materials & Methods: ADPNP: adenylyl-3-y-imidodi-phosphate BB: binding buffer BSA: bovine serum albumin CLIP: cross-linking immunoprecipitation Cryo-EM: cryo electron microscopy LB: lysogeny broth OD: optical density PCR: polymerase chain reaction SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis TB: terrific broth **Diseases:** CF: cystic fibrosis DMD: Duchenne muscular dystrophy

Acknowledgements

When I started writing this manuscript one of the first parts I wrote were the acknowledgements. I called this my 'active procrastination moments' because even if I wasn't writing science, I was reflecting on my PhD and how special and intense this whole experience has been. However, the longer I spent writing, the harder it became to think about the end. I routinely revisited my initial acknowledgements to remind myself of the state of mind I was in when I began writing. I had clearly underestimated the difficulty of writing this manuscript and the emotional weight of finishing this chapter of my life. Alas, if you are reading this, it means that I made it.

I dedicate this manuscript to the memory of my beloved grandfather, **Tito**, who passed away during my PhD. As the original biochemist in our family, he was a source of inspiration and encouragement for me throughout my academic journey. I hope that I have made him proud with this work.

My first acknowledgements have to go to my parents, Sergio and Ana Cristina, to whom I owe everything. They worked tirelessly to make my dream of studying in Paris a reality. Despite being 10,000 kilometers away, they were always there for me, constantly checking on me when I felt lonely and stressed and always believed in me. They raised my siblings and I to never give up. To my dear brother **Nataniel**, thank you for your support and for pushing me to be the best version of myself. I admire you so much, you are my role model. To my little twin sisters, Julieta and Miranda, I love you unconditionally and cannot wait to welcome you to France next year. To my beloved grandmothers Martita and Mutti, I miss you every day, thank you for all the love you give me. Lastly, among my family, I have to thank my incredible partner, Arnaud. Your unwavering love, support, and encouragement have been a constant source of strength. Throughout our relationship, you have always gone above and beyond to make me feel cherished and appreciated. Your kindness and generosity have helped me through rough times, especially during the writing of this manuscript. I always felt supported, even though you were on the other side of the world (literally) while I was writing. Thank you for being my best friend, my home, my confidant, and my soulmate. You and Catsu have brought so much joy and happiness into my life, and I am truly blessed to have you by my side. I love you more than anything.

To my PhD supervisor **Hervé**, I am grateful for your guidance throughout this journey. Thank you for choosing me and trusting me for this project. It must not have been easy to handle two loud Spanish speaking PhD students simultaneously but despite some hard times, we found a way to communicate. Thank you for always pushing me to strive for excellence. You taught me to be patient but especially you made me a better scientist. Above all, thank you for encouraging me to think critically and to be independent, even when you made fun of my questions (which, I agree, may have been excessive and foolish sometimes) and my failed western blots (n=2). I really hope that all of our hard work will see the light.

I also thank **Cosmin**, who introduced me to Hervé back in 2019 when I was looking for an M2 internship. You were the mastermind behind the project and of great help when I had questions. Among the other collaborators who made this work possible, I thank **Marc** and **Gwenaël** for the long meetings and brain teasers which you know I adore, as well as **Irène**, **Amandine**, **Toni**, **Laurence** and **Agathe** for their work in the project. Our fruitful collaboration has taught me so much. A special thanks goes to my second lab, the physicists: **Vincent**, **Jean-François**, **Martin**, **Jess**, **Raph** and **Nicolas**. I really enjoyed spending ¹/₄ of my PhD with you, thank you for showing me alternative ways to look at science. Martin, Jess and Raph, thank you for your kind friendship.

To all of the HLH members, thank you for your kindness and support. **OB**, you have also been a great mentor and a good friend. Isa, you were the first to take me under your wings to teach me my way in the lab, no purification would have been possible without your quidance. Tommaso, bambino, I have greatly appreciated all of our discussions, personal and professional. Cindy, this PhD would not have been the same without your amazing vegan desserts. Jeanne, it was a pleasure to have you as my student, your help to the project was invaluable. To the latest member of the HLH lab, Clara, I hope your PhD is going to be great. As for past members of the lab, Quentin, Oh Sung, Toni, Maëva and Elias, I miss you dearly, thank you for the great times we spent together. Quentin, I don't have words to express my gratitude. You believed in me since the first day of my internship and to this day, you continue to be an amazing friend. Finally, I want to express my deepest gratitude to my PhD conjoined twin sister: Luchis. Your sense of humor has made every day brighter and more meaningful at the lab (and I think I don't only speak for myself). Throughout our PhD journey, you have been a constant source of support and have always offered a listening ear and wise advice whenever I needed it. I feel fortunate to have had you by my side.

This brings me to the Top 4. **Bertrand**, **Morgane**, **Lucia** and I have shared so many experiences during these past four years. From leaking toilets, deep conversations and endless parties. Thank you Mo for sewing all of our clothes. No PhD should be done without a great group of friends like you. I should maybe say Top 5 with the arrival of **Thomas Petit** to the section 2 PhD group. You all have been a great support system and I am grateful for every moment we shared together.

I am also grateful for the support of the people in the section, particularly Martine, who was crucial in ensuring the success of my (many) protein purifications. I am thankful for the camaraderie and laughs shared with members of section 2, including Flo, Vero, Alexandra, Marion, Diletta, Simon, Emma, Léa, Victoria, Alan, Julien, Alice, Terence, Hugues, Laurent, and Benoît. Without Bilel's expertise with computers and Abdul, Béatrice and Virginie's help with bureaucracy, none of us could work in the institute. I also would like to acknowledge the amazing team I had the pleasure to exchange with while teaching: Andréa, Alain, Eric, Renaud, Amaury, Gersende, Morgane TC. I am thankful to have worked alongside such supportive and humane people.

Among the other people that I have been fortunate to meet in IBENS are my dear friends and beer partners: the plant guys (Antinéa, Jerôme, Pierre, Justine, Souraya, Basile, Marc, Odon, Lucas, Thierry, Louna, Angélique), the new plant guys (Weronika, Liam, Alexandre, Fabien), the neurobiologists (Hind, Romain 1 and Romain 2, Marie 1 and Marie 2, Alan, Victor, Julie, Stefano, Aurélien, Olivier) and the teachers (Marie, Gersende and Barbara). When everything seemed awful, you were always there to share a beer with me and make me laugh (or cry), to give me advice and listen to my frustrations. Thank you, from the bottom of my heart. But most importantly, thank you Youcef and Mangala, for creating a safe space in which we all could go and decompress. You have been amazing hosts and even better friends.

Although my time at IBENS has been incredibly rewarding, I am equally grateful for the abundance of love and support I have received from so many people outside of this institution. I would like to extend my heartfelt thanks to the Belifontains, including **Délia** and **Adri** who have been like a second family to me and nothing short of amazing. I am also grateful to **Alix**, my first friend in France, with whom I have had the pleasure to share 10 years of friendship. My dear friends from university, **Amélie**, **Bruno**, **Clara**, **Téoman**, **Tom**, **Val**, and **Seb**, have also been a constant source of support and encouragement. The few friends I have left from Bolivia, **Cami** and **Alex**, and I am grateful for our enduring connection. Lastly, I want to express my appreciation to other friends I made more recently, the surfers collectively known as **La Meute** (Houda, Clem, Manu, Penny, Ninon, Nico, **Bjorn**, **Kevin**, **Bap**, **Vio** & Co) and the doctors (**Chaud**, **Schoumi**, **Valoche**, **Tom** & Co) who have welcomed me into their groups with open arms. I consider myself incredibly lucky to have all of you in my life.

Going back to the beginning, I'd also like to thank my first mentors, **Mathilde Garcia** and **Fréderic Devaux**, who saw potential in me, as well as the other lab members who taught me how to pipet: **Thierry**, **Medine** and **Antonin**. Their patience and willingness to share their knowledge have been instrumental in my development as a scientist.

Last but not least, I'd like to deeply thank the jurors who generously agreed to evaluate my work: Isabelle Behm-Ansmant, Jan Kadlec, Dominique Weil, David Gatfield, Marina Pinskaya and Marc Boudvillain. I deeply appreciate your time and effort, especially for those who traveled from far to come. I am also grateful for the members of my thesis committee Claire Torchet, Gilles Fisher and Nicolas Leulliot for the insightful conversations and advice they gave me during my PhD. Thank you again Hervé, Gwenaël, and my non biologist father Sergio, for taking the time to correct and proofread this endless manuscript.

The final miscellaneous thanks go to **Alexandra Elbakyan**, the creator of SciHub, **Catsu**, the most beautiful cat in the world, and **chatGPT** for obvious reasons. Finally, I thank my biology professor of the first year of *prépa* who told me to go back to my country because I would never succeed.

Introduction

General Introduction

Throughout its lifespan, the fate of an RNA molecule is subject to strict regulation by various cellular quality control mechanisms. These mechanisms are responsible for maintaining cellular homeostasis, a fundamental prerequisite for the survival and proper functioning of eukaryotic organisms.

The first chapter serves as an introduction to the importance of quality control systems, emphasizing the significance of gene expression regulation and the need for robust quality control mechanisms. We explore the interplay between transcription, RNA processing, and translation, highlighting how errors or defects at any stage can have detrimental effects on cellular function. By elucidating the complex network of regulatory pathways, we gain a deeper appreciation for the sophisticated machinery that maintains cellular homeostasis.

Chapter 2 provides an overview of cytoplasmic RNA decay pathways, offering insights into their shared characteristics and underlying mechanisms. By examining the intricate interplay between various decay machineries, we discuss the coordinated efforts required for efficient RNA turnover.

In Chapter 3, we narrow our focus to the mechanistic details of one particular pathway that has been the central focus of this PhD work: Nonsense-mediated mRNA decay (NMD). We explore the intricate steps involved in NMD, from recognition of premature termination codons to the degradation of faulty mRNA transcripts. By unraveling the molecular players and regulatory factors involved in NMD, we gain a deeper understanding of how this pathway safeguards the fidelity of gene expression.

Chapter 4 takes a broader perspective, zooming out to explore the importance of NMD in both health and disease. We examine how dysregulation of NMD can lead to various pathological conditions, including genetic disorders and cancer. By highlighting the pivotal role of NMD in maintaining cellular integrity, we underscore its significance as a potential therapeutic target for intervention.

Finally, in Chapter 5, we devote our attention to the main orchestrator of NMD: the helicase Upf1. We unravel the intricate functions of Upf1 in coordinating the various steps of NMD, from RNA surveillance to mRNA degradation. By dissecting the structural and functional aspects of Upf1, we gain insights into its essential role as a key regulator of RNA quality control.

Chapter 1: About the importance of quality control systems

1) The cell is not as perfect as it seems

a) Brief overview of gene expression step by step

In this manuscript we will mainly focus on mRNAs (messenger RNAs) which are transcribed by RNA PollI (polymerase II) and have protein-coding potential. Replication, transcription, and translation are commonly regarded as the central steps of gene expression, yet there are several other crucial stages that must influence gene expression prior to protein synthesis. These include co-transcriptional processes such as capping and splicing, as well as post-transcriptional steps such as polyadenylation, base modifications, and export (Figure 1). A typical mRNA has the following characteristics: a <u>*T*-methyl-Guanosine (m7G) cap structure, a regulatory untranslated region (5' UTR), an</u> open reading frame (ORF), encompassed between a start codon (AUG) and a termination or stop codon (TC, encoded by UAA, UGA or UAG), a regulatory 3' UTR region and a 3' poly(A) tail (varying between 70-80 adenosines long in yeast and up to 250 in humans) (Figure 2). Maturation of a pre-messenger RNA to a mature mRNA within the nucleus consists in addition of the cap m7G structure by capping enzymes (reviewed in Ramanathan et al., 2016), excision of introns by the spliceosome (reviewed in Rogalska et al., 2023; Will & Lührmann, 2011), base modification such as m6A, m5C, Ψ , (reviewed in Kumar & Mohapatra, 2021), and polyadenylation of the 3' end (reviewed in Passmore & Coller, 2022) (Figure 1). Capping and polyadenylation constitute critical steps to protect the mRNA from exonucleolytic degradation, among other essential roles. The mature mRNA, which can vary in length from a couple hundred to thousands of nucleotides, is then bound by multiple RBPs (RNA binding proteins) to form a mRNP (messenger ribonucleoprotein) particle. The mRNP can then be exported and translated by the ribosome in the cytoplasm (Figure 1).

Among the many RBPs that coat mRNAs, we find the <u>n</u>uclear <u>cap</u> <u>b</u>inding <u>c</u>omplex (nCBC) which protects the cap (reviewed in Topisirovic et al., 2011) and the <u>poly(A)</u> <u>b</u>inding <u>p</u>roteins (PABPs) which protect the poly(A) tail (reviewed in (Eliseeva et al., 2013; Passmore & Coller, 2022). In metazoans, the EJC (<u>e</u>xon junction <u>c</u>omplex) is deposited by the spliceosome 27 nucleotides upstream exon-exon junctions (reviewed in Le Hir et al., 2016).

Often, we keep the simplistic view of a linear and naked mRNA such as depicted in Figure 2, yet, mRNPs are most likely highly compact, have secondary or tertiary structures and are decorated by hundreds of RBPs (Gerstberger et al., 2014; Pierron & Weil, 2018; Bonneau et al., 2023). Apart from ubiquitous RBPs such as CBC, EJC and PAPBs, each mRNP is unique in terms of RBP composition (Singh et al., 2015). Around 1 500 RBPs are encoded by the human genome, corresponding to 10% of the cell proteome (Gerstberger et al., 2014; Hentze et al., 2018), which gives an early indication of how complex and diverse the RBP composition of mRNPs is, and explains why their structure and dynamics are still poorly characterized. Recent CryoEM studies of native yeast mRNPs showed how there is an intricate network of protein-protein interactions leading to high compaction of the mRNP

(Bonneau et al., 2023). Some RBPs bind the mRNA co-transcriptionally, leading to rapid packaging of the nascent mRNP (Singh et al., 2015). Moreover, mRNPs composition is also highly dynamic, undergoing remodeling during export and translation and disassembly by active mechanisms which may expose the RNA to degradation (Singh et al., 2015). These particles adopt a compact, flexible, and linear structure, which is loosened during translation and highly compacted after translation (Adivarahan et al., 2018; Metkar et al., 2018).



Figure 1: Eukaryotic gene expression steps

Schematization of the various co-transcriptional and post-transcriptional processes involved in the maturation of mRNA within the nucleus. The mRNA synthesis begins with RNA polymerase II transcription of pre-mRNA. The pre-mRNA is capped, spliced, polyadenylated and edited (m6A, m5C, Ψ). After capping, the m7G cap is bound by nuclear cap binding proteins (nCBC, beige) which is replaced by eIF4E/eIF4G (gray) in the cytoplasm. After splicing, the exon junctions are loaded with exon junction complexes (EJC, purple). After polyadenylation, the poly(A) tail is bound by polyA binding proteins (PABP, light red). Quality controls checkpoints and the associated proteins are labeled in red.



Figure 2: Structure and characteristics of a mature mRNA

The coding sequence (CDS, gray) starts at the start codon (green) and stops at the termination/stop codon (gray). It is encompassed by 3' and 5' untranslated regions (UTR, orange). In the 5' end, the mRNA is capped (m7G) and in the 3' end it is polyadenylated (poly(A)). The mRNA is coated by multiple RNA binding proteins among which, the cap binding proteins (CBC, dark gray), the exon junction complex (EJC, purple), and the poly(A) binding proteins (PABPs, light red).

b) Sources of errors

Certain gene expression steps, such as splicing or translation, exhibit greater complexity compared to others due to a higher number of involved components or increased vulnerability to errors. Yet, each step carries the potential for errors that could ultimately compromise the integrity of gene expression and its products. Errors can occur at different scales: (1) Permanent codon replacement on the DNA due to point mutations (insertions, deletions (indels), substitutions) or large-scale mutations (indels, translocation, duplication, inversion) during DNA replication, meiosis or mitosis. These errors can either substitute an amino acid (aa) codon for a stop codon (nonsense mutations) or cause frameshifting (indel mutations) which most likely generates a stop codon within the next codons (Antonarakis & Cooper, 2013). (2) Temporary codon replacement at the RNA level by uncorrected errors during transcription, splicing defects (alternative splicing, intron retention), RNA editing errors, alternative polyadenylation, physical alteration (cleavage). Again, indels or substitutions can lead to frameshifting and/or premature stop codons. (3) Misreading of the RNA sequence by the ribosome by readthrough, wrong selection of transfer RNAs (tRNAs) or defective translation initiation can lead to erroneous as incorporation or frameshifting (Reynolds et al., 2010; Joshi et al., 2019). In addition, slippery sequences can cause the ribosome to shift its reading frame by one or two nucleotides (nt), thus changing all of the subsequent reading frames (Lykke-Andersen & Jensen, 2015; Pisareva et al., 2011).

To ensure accuracy and mitigate errors, all of these processes have undergone extensive evolutionary refinement and are subject to stringent verification by cellular checkpoints or quality control systems. Given the significant implications of gene expression failures, these mechanisms must strike a delicate balance between the efficiency of synthesis and the risk of errors. In some cases, these processes have even developed intrinsic proofreading and correction mechanisms, further enhancing their precision and fidelity.

2) Errors can be fatal

a) (Not) all errors have consequences

All of the aforementioned sources of errors can lead to gene mutations which may be deleterious for the cell. Examples are missense mutations, which result in aa substitution, or nonsense mutations, which give rise to premature termination codons (PTCs). However, some are silent, thus have no impact on the polypeptide sequence due to the genetic code redundancy. As their name suggests, silent mutations have little to no impact when they occur in the coding sequence. Conversely, missense and nonsense mutations can have significant consequences depending on their location. A missense mutation that affects an important residue for folding, catalytic activity, target recognition or intracellular transport, for example, is likely to result in a non-functional protein and can have severe effects on the cell. Similarly, nonsense mutations lead to translation halt and the corresponding protein will be truncated (Hall & Thein, 1994). Therefore, the consequences depend on the location of the PTC relative to the bona fide stop. In the best-case scenario, the protein is still relatively functional, while in the worst-case scenario, in addition to wasting cell energy to produce the mutated protein, such non-functional protein is deleterious for the cell. Among the multiple impacts of this, we find dominant negative effects, haploinsufficiency (Fan et al., 2001), cytoplasm overload of misfolded proteins that may aggregate (van Leeuwen et al., 1998) or localization problems. Hence, it is safe to say that quality control surveillance is of essence. Interestingly, errors can also be beneficial: mistranslation has been associated with oxidative stress response in mammalians (Reynolds et al., 2010) and mutations are the main driver of evolution. Therefore a tightly controlled balance on the absolute number of mutations is necessary to preserve fitness.

b) A hidden layer of complexity

The development of genome wide sequencing techniques during the human genome project (Djebali et al., 2012) revealed that 75% of the genome is transcribed, among which only 25% of these transcripts are protein-coding. Several classes of <u>non-coding RNAs</u> (ncRNAs) that do not have the potential to code for proteins have been identified (reviewed in Hombach & Kretz, 2016; Zhang et al., 2019). Housekeeping ncRNAs include transfer <u>RNAs</u> (tRNAs), <u>ribosomal RNAs</u> (rRNAs), <u>small nuclear RNAs</u> (snRNAs) and <u>small nucleotar RNAs</u> (snoRNAs), which are involved in translation decoding, ribosome structure, splicing and RNA modifications respectively (Figure 3A). However, apart from housekeeping ncRNAs are generally classified by their length: long <u>non coding RNAs</u> (lncRNAs) which are longer than 200 nucleotides, and small ncRNAs (reviewed in Mattick et al., 2023). Some small ncRNAs, such as <u>micro RNAs</u> (miRNAs), <u>small interference RNAs</u> (siRNAs) and <u>PIWI-interacting RNAs</u> (piRNAs), participate in RNA interference mechanisms thus are key regulators of gene expression (Figure 3A).

In contrast to other ncRNAs, lncRNAs are transcribed by PollI so they can be capped, polyadenylated and recognized by the translation machinery like mRNAs (reviewed in Jarroux et al., 2017; Mattick et al., 2023). Currently many other characteristics than their

length contribute to their classification (Jarroux et al., 2017; Mattick et al., 2023). Although their functions had been largely unknown, today they are considered as drivers of evolution and contribute to genome complexity as some of them have been shown to have specific functions in gene expression (Hombach & Kretz, 2016; Jarroux et al., 2017; Mattick et al., 2023).



Figure 3: RNA classification and pervasive transcription

(A) RNAs are classified into two main categories: messenger RNA (mRNA) and non-coding RNA (ncRNA). Non-coding RNAs encompass various subclasses, including housekeeping ncRNA, which comprises transfer RNA (tRNA) and ribosomal RNA (rRNA), as well as regulatory ncRNA. Regulatory ncRNAs can be further divided into long ncRNA (lncRNA) and small ncRNA. The small ncRNA category is classified into several subclasses, namely microRNA, small nucleolar RNA (snoRNA), small interfering RNA (siRNA), small nuclear RNA (snRNA), and PIWI-interacting RNA (piRNA). (Figure extracted from Inamura, 2017).

(B) Pervasive transcripts are produced when RNA polymerase II (RNA PolII) initiates transcription nucleosome-free regions (NFRs). These transcription events give rise to numerous ncRNAs (red). Cryptic promoters are indicated by red arrows, while promoters of protein-coding genes are represented by black arrows. Pervasive ncRNAs are classified based on their susceptibility to various cellular ribonucleases (CUT, SUT or XUT) (Figure modified from Villa & Porrua, 2022).

Additionally, other small ncRNAs exist and are generated through what we call pervasive transcription (reviewed in Jensen et al., 2013; Villa & Porrua, 2022), which arises from nonspecific activity of RNA PollI (Neil et al., 2009; Xu et al., 2009). Pervasive transcription gives rise to unstable transcripts which were only discovered when inactivating guality control mechanisms in the cell since they are guickly degraded in wild type (WT) conditions. Cryptic unstable transcripts (CUTs) were identified in yeast when nuclear quality control mechanisms were inactivated (Wyers et al., 2005; Davis & Ares, 2006; Houalla et al., 2006) (Figure 3B). Subsequently, some transcripts were found to be less sensitive to nuclear quality control inactivation and were denominated as stable unannotated transcripts (SUTs) (Neil et al., 2009; Xu et al., 2009) (Figure 3B). Xrn1-dependent unannotated transcripts (XUTs) were discovered by inactivating the 5'-3' cytoplasmic exonuclease Xrn1 (van Dijk et al., 2011) (Figure 3B). These three types of transcripts are not mutually exclusive. CUTs and SUTs are generated by transcription in nucleosome-depleted regions within the extremities of a gene which suggest that most promoters are bidirectional (Figure 3B; Jensen et al., 2013; Villa & Porrua, 2022). In mammals, divergent transcripts, called promoter upstream transcripts (PROMPTs) or upstream antisense RNAs (uaRNAs), are targets of the nuclear exosome (Preker et al., 2008; Flynn et al., 2011; Ntini et al., 2013) and therefore reminiscent of the yeast

Saccharomyces cerevisiae (S.c.) CUTs. Similarly, more stable mammalian lncRNAs may be compared to SUTs or XUTs.

Other pervasive transcripts have been described in more specific situations such as <u>te</u>lomeric <u>repeat-containing RNAs</u> (TERRAs), which are involved in telomere homeostasis (Luke et al., 2008) and <u>meiotic-specific noncoding transcripts</u> (MUTs), which arise only during yeast sporulation (Lardenois et al., 2011). The expression of most PolII ncRNAs is tightly regulated by quality control systems which recognize them as aberrant RNAs.

3) First responders act in the nucleus

Every step of the way, quality control systems are employed to ensure correct gene expression. In order to mitigate the export of aberrant mRNAs into the cytoplasm, each mRNP has to pass checkpoints within the nucleus (Villa et al., 2008; Schmid & Jensen, 2010; Kilchert et al., 2016). This also ensures a good balance between the availability of RBPs and nuclear transcripts to avoid titration of RBPs by faulty RNAs that can disturb nascent mRNP formation (Tudek et al., 2019).

Failure to add a functional cap triggers rapid decay by the nuclear 5'-3' exonuclease Rat1 and co-factor Rai1 (Jiao et al., 2010; Xiang et al., 2009). Failure to correctly polyadenylate transcripts triggers rapid RNA decay through nuclear exosomal activity, as the polyadenylase has been shown to bind the nuclear exosome protein Rrp6 (Milligan et al., 2005, Burkard & Butler, 2000). Splicing errors or errors during mRNP formation can trigger exonucleolytic decay (Rat1 or exosome), or endonucleolytic decay through Swt1 (Bousquet-Antonelli et al., 2000; Skružný et al., 2009). Nuclear quality control systems also ensure retention of aberrant mRNAs in the nucleus to avoid cytoplasmic export and translation of aberrant proteins. mRNAs are not released from their transcription sites until the poly(A) tail is correctly added or in case of impaired nuclear export systems (Custódio et al., 1999). Furthermore, aberrant mRNAs can also be retained in the nucleus by anchoring to the nuclear pore complex (Galy et al., 2004). The endonuclease Swt1 is tightly associated with the nuclear pore proteins, ensuring rapid turnover of the anchored mRNAs (Skružný et al., 2009).

While nuclear quality control systems are designed to be highly stringent, there are cases in which aberrant transcripts can still bypass them and be exported to the cytoplasm. In yeast, certain pre-mRNA molecules that undergo incomplete splicing are transported to the cytoplasm, resulting in their presence within that cellular compartment. Subsequently, these cytoplasmic pre-mRNA molecules undergo a secondary quality control process (Sayani et al., 2008). Three main translation-dependent cytoplasmic RNA quality control mechanisms have been described thus far: <u>no-stop decay</u> (NSD), <u>no-go decay</u> (NGD) and <u>nonsense-mediated RNA <u>decay</u> (NMD), which is the main focus of this PhD project.</u>

<u>Chapter 2: Three strikes and you're out, the three-step</u> process of cytoplasmic mRNA quality controls

Multiple RNA decay mechanisms operate in the cytoplasm. The intricate web of cytosolic mRNA decay pathways is united by a tripartite process consisting of (1) translation and target recognition, (2) recruitment of specific factors, and (3) decay triggering through general decay pathways. In this chapter we will briefly review how cytoplasmic quality control systems and other specific RNA decay pathways share pivotal common steps and how they are intrinsically intertwined. We focus on the similarity of translation-dependent quality control systems (NGD, NSD and NMD) with specific transcript degradation systems.

1) The beginning of the end: it all starts with translation

When a PollI transcribed RNA escapes all nuclear quality control systems, it may be recognized by the translation machinery in the cytoplasm. Translation represents thus the first checkpoint for cytoplasmic mRNA quality control and constitutes a critical layer of regulation that ensures the accuracy and efficiency of protein synthesis. The ribosome allows crosstalks between translation and mRNA decay and is considered a regulatory hub for orchestrating homeostatic RNA degradation, RNA guality control and stress signaling (Morris et al., 2021). Beyond the relatively well-characterized surveillance pathways of NSD, NGD, and NMD that regulate aberrant mRNA decay, a diverse range of specific mRNA decay mechanisms that are also translation-dependent have emerged as critical regulators of functional mRNA expression. Remarkably, mRNA half-lives in yeast have been shown to be highly variable, spanning from a mere 3 minutes to a staggering 200 minutes (Sun et al., 2012). Such variability in mRNA stability, coupled with the fluctuating protein abundances observed under different physiological states (Melamed et al., 2008; Molina-Navarro et al., 2008), points to a widespread interplay between translation and mRNA degradation. To better understand this interplay, we will first provide a brief overview of the translation process.

a) Translation mechanism at a glance

Once a transcript is exported to the cytoplasm the nCBC is replaced by the cytoplasmic cap binding proteins eIF4F/eIF4G (<u>e</u>ukaryotic <u>i</u>nitiation <u>factor 4F/4G</u>). The mRNP was thought to adopt a circular conformation in which eIF4F interacts with the PABPs (Tarun & Sachs, 1997; Wells et al., 1998) (Figure 1). However, it is noteworthy to point out that the "closed loop model" has been recently challenged (Vicens et al., 2018).

The pre-initiation 43S complex composed of the 40S subunit of the ribosome along with the initiator tRNA assembles near the cap and scans the RNA until it reaches the AUG start codon, cognate to the initiator methionine tRNA (Figure 4). The scanning has recently been estimated to be extremely quick: 100-200 nt/s compared to the 9-15 nt/s during translation elongation (Wang et al., 2022; Paternoga & Wilson, 2023). When the 43S pre-initiation complex reaches the start codon, the large subunit 60S is recruited to form the complete

80S ribosome. Within the 80S, three sites can be distinguished: the A site (<u>a</u>minoacyl), the P site (<u>peptidyl</u>) and the E site (<u>exit</u>), located in that order from 3' to 5'.

Translation elongation is a cycle of steps in which: (1) the cognate tRNA along with the corresponding amino acid and eEF1A (<u>e</u>ukaryotic <u>e</u>longation <u>factor 1A</u>) enter site A, (2) the peptide bond between the nascent peptide and the next amino acid is formed in site P, (3) the nascent polypeptide is transferred to the P/E site and the tRNA is released, (4) the ribosome translocates one codon downstream (Figure 4).

Translation terminates once the ribosome reaches one of three stop codons (UAA, UGA or UAG) which do not have any associated tRNAs (reviewed in Hellen, 2018 and Karousis & Mühlemann, 2019). When a stop codon is in the ribosome's reading frame, instead of cognate tRNA binding to the A site, the ternary complex composed of eRF1/eRF3-GTP (eukaryotic release factors 1/3, Sup45 and Sup35 respectively in yeast) is recruited (Bertram et al., 2000; Mitkevich et al., 2006; Blanchet et al., 2015). eRF1 recognizes the



Figure 4: Eukaryotic translation steps

This figure depicts the process of translation of an mRNA by the ribosome. The pre-initiation 43S complex composed of the 40S ribosomal subunit and the initiator tRNA is assembled near the mRNA cap. The complex scans the mRNA until it reaches the AUG start codon, which is recognized by the initiator methionine tRNA. The large 60S subunit joins the 40S subunit to form the complete 80S ribosome. The 80S ribosome bears three sites: the A site (aminoacyl), the P site (peptidyl), and the E site (exit), arranged from 3' to 5' end of the mRNA. During translation elongation, the cognate tRNA carrying the corresponding amino acid enters the A site, a peptide bond is formed between the nascent peptide and the next amino acid in the P site, the nascent polypeptide is transferred to the P/E site, and the tRNA is released, the ribosome translocates one codon downstream. Translation terminates when the ribosome encounters a stop codon, which lacks associated tRNAs. The ternary eRF1/eRF3-GTP complex is recruited. eRF1 recognizes the stop codon, and GTP hydrolysis by eRF3 triggers the release of the nascent peptide and the dissociation of the ribosome. The interaction between eRF1 and Abce1 allows ribosome recycling. (Based on Schuller & Green, 2018).

stop codon within the A site, as it has a similar structure to that of a tRNA (Cheng et al., 2009). eRF3 is a GTPase protein, and the energy released by GTP hydrolysis triggers a conformational change in eRF1 which allows terminal peptidyl-tRNA bond hydrolysis, nascent peptide release and eRF3 release (Frolova et al., 1999; Salas-Marco & Bedwell, 2004; Alkalaeva et al., 2006; Jin et al., 2010). The ribosome is then dissociated by the interaction between eRF1 and Abce1 (Rli1 in yeast), an ATPase which allows eRF1 release, 60S and 40S subunit dissociation and ribosome recycling (Pisarev et al., 2010; Shoemaker & Green, 2011; Becker et al., 2012). Despite being a highly controlled mechanism, translation heterogeneity can arise by differences in primary mRNA sequence, RBP composition, RNA modifications, mRNA secondary or tertiary structure, and ribosome composition (Sonneveld et al., 2020). In addition, recent studies have shown that translation can also occur on nCBC-bound RNAs (Hoek et al., 2019).

b) What is "normal" and what is not: aberrant mRNA decay pathways

Aberrant translation occurs when the ribosome encounters one of three situations: (1) it reaches the poly(A) tail due to the lack of a termination codon, (2) it stalls due to the presence of roadblocks or (3) the termination is premature. These three scenarios correspond respectively to the targets of NSD, NGD and NMD (Shoemaker & Green, 2011; Roy & Jacobson, 2013; Powers et al., 2020; Morris et al., 2021) (Figure 5).

In NSD (Frischmeyer et al., 2002; van Hoof et al., 2002), where a termination codon is absent or misread due to various factors such as errors during gene expression or alternative polyadenylation sites (Ozsolak et al., 2010), the ribosome continues translating until it reaches the poly(A) tail (Figure 5A). This results in the synthesis of a polypeptide that is rich in positively charged lysine (AAA-Lys). The polypeptide is hindered by electrostatic interactions within the E site of the ribosome (Guydosh & Green, 2017). As a consequence, the ribosome stalls and the mRNA is cleaved (Guydosh & Green, 2017). Similarly, NGD is triggered when the ribosome stalls during translation due to the presence of a stable secondary structure within the coding region or if a cognate tRNA is missing (Doma & Parker, 2007; Simms et al., 2017) (Figure 5B). Ribosome stalling can cause ribosome collisions (De & Mühlemann, 2022). In addition to mRNA decay, NSD and NGD also trigger <u>r</u>ibosome-associated protein <u>g</u>uality <u>c</u>ontrol (RQC), a protein quality control system which allows the degradation of the nascent polypeptide when ribosome collision arises (Defenouillère & Fromont-Racine, 2017).

Aberrant translation that triggers NMD, on the other hand, is far less well defined (See Chapter 3). Two hypotheses have been formulated for signaling the presence of a PTC-containing mRNA (Figure 5C): (1) The PTC artificially creates a longer 3' UTR which increases the distance between the ribosome and the PABPs affecting termination efficiency. Long 3' UTRs are thus considered as signals for NMD triggering (Amrani et al., 2004). (2) The presence of an EJC in the 3' UTR, either by the existence of a spliced intron in the region or by the presence of a PTC before the last intron. Either way, during translation the EJC is not disassembled by the ribosome, therefore it represents a signal for NMD triggering (Kashima et al., 2006). NMD triggers rapid RNA degradation, however, the link

between NMD and RQC for protein elimination is still misunderstood. Recent results point towards no RQC triggering and no ribosome stalling nor collision in NMD given that ribosome collision sensors are not involved in NMD (Karousis et al., 2020; De & Mühlemann, 2022). On the contrary, it has been suggested that NMD polypeptides are selectively targeted to the aggresome (Park et al., 2020).



Figure 5: mRNA characteristics triggering translation- dependent quality control systems

This figure presents the characteristics that initiate No-Stop Decay (NSD), No-Go Decay (NGD) and Nonsense-Mediated mRNA Decay (NMD) in an mRNA. (A) NGD is triggered when the ribosome stalls during translation due to the presence of a stable secondary structure within the coding region or a truncated mRNA. The stalled ribosome recruits Dom34 and Hbs1 and the mRNA is degraded. (B) NSD is triggered by the absence or misreading of a termination codon which leads to continued translation until the ribosome reaches the poly(A) tail. This results in the synthesis of a polypeptide rich in positively charged lysine (purple). Consequently, the ribosome stalls, Dom34 and Hbs1 are recruited and the mRNA is cleaved. (C) Two NMD triggering characteristics are proposed: the presence of a premature termination codon (PTC) which creates an extended 3' untranslated region (3'UTR) and the presence of an Exon Junction Complex (EJC) in the 3'UTR, either through the presence of a spliced 3'UTR intron or due to the existence of a PTC. In both cases, the trimeric Upf1/2/3 complex is recruited to trigger decay either by the long 3'UTR or by the EJC.

2) Recruitment of specialized factors

a) Recruitment by faulty termination

As discussed above, the aberrant situations that ribosomes might encounter prompt the recruitment of quality control factors to trigger mRNA decay. NSD and NGD involve specialized release factors, Dom34 (Pelota in metazoans) and Hsb1 (Figure 5A, 5B). These proteins interact with the A site of the stalled ribosome by having similar structures as tRNA and eRF1 (Doma & Parker, 2006; Lee et al., 2007; Graille et al., 2008; Saito et al., 2013). In contrast, the rules governing the recruitment of the main actor of NMD called Upf1 (<u>up</u>-<u>f</u>rameshift), are still debated and will be discussed further in chapter 3. Nevertheless, the Upf1 recruitment along with its partners Upf2 and Upf3 is essential for NMD, whether they are recruited by a 3' UTR EJC or by the terminating ribosome. (Figure 5C).

b) Recruitment of RNA binding proteins to specific RNA sequences or structures

Translation and mRNA decay are coupled beyond mere RNA surveillance (Tuck et al., 2020). While NSD, NGD and NMD pathways are activated by aberrant translation, other mRNA decay pathways rely on protein/RNA interactions to initiate decay of functional mRNAs (Morris et al., 2021). These mechanisms allow tight molecular control of mRNA homeostasis in the cytoplasm. Examples of protein/RNA trans actors for selective mRNA decay are Staufen-mediated decay (SMD), ARE-mediated decay (AMD), Regnase1-mediated decay (RMD), replication-dependent histone mRNA decay (HMD), regulated Ire1-dependent decay (RIDD) and RNA interference (RNAi). All of these pathways are linked to translation. Other examples of pathways that have similar mechanisms but do not depend on translation are glucocorticoid-receptor mediated mRNA



Figure 6: Recognition mechanisms for specific RNA decay pathways Illustration of the recognition structures, motifs or sequences of RNA binding proteins involved in various examples of specific RNA decay pathways.

<u>d</u>ecay (GMD; Cho et al., 2015) and <u>Ro</u>quin-<u>m</u>ediated <u>d</u>ecay (RoMD, Leppek et al., 2013; Mino et al., 2015).

The common step for the latter mechanisms is that an RBP recognizes either specific secondary structures or sequences within the target RNA to trigger decay (Figure 6). For example, Staufen, the main actor of SMD, recognizes a 3' double stranded structure (Kim et al., 2005) (Figure 6). In HMD, the histone mRNAs bear a conserved stem loop in their 3' instead of a poly(A) tail which is recognized by Slbp (stem-loop binding protein) (Marzluff et al., 2008; Marzluff & Koreski, 2017) (Figure 6). Furthermore, other decay mechanisms rely on specific RNA motif recognition: the glucocorticoid receptor of GMD recognizes a GC-rich motif in presence of glucocorticoid (Cho et al., 2015; Park et al., 2016), Regnase 1 (also called ZC3H12A/MCPIP1) and Roquin bind constitutive decay elements (CDE) in the 3' UTRs (Leppek et al., 2013; Mino et al., 2015) (Figure 6). Similarly, ARE-containing mRNAs (<u>AU-rich elements</u>) are bound by multiple ARE-binding proteins such as Ttp, Brf-1 to trigger decay during AMD (Sanduja et al., 2011) (Figure 6). Finally, in RIDD mRNAs encoding membrane and secreted proteins are bound by Ire1 during their co-translational translocation to the endoplasmic reticulum (Hollien et al., 2009) (Figure 6).

More specific targeting of sequences for mRNA degradation and translation repression happens during RNA interference gene silencing, in which miRNAs or siRNAs complementary to the targeted sequences allow the recruitment of Ago1/Ago2 to form the <u>RNA-induced silencing complex</u> (RISC) (Behm-Ansmant et al., 2006; Sioud, 2021) (Figure 6).

In summary, the various cytoplasmic mechanisms responsible for the decay of specific mRNAs described above utilize a critical protein to determine the fate of the mRNA. This protein is essential for precise control over target selection and mRNA degradation.

Unlike the surveillance pathways of NSD, NGD and NMD, which target a large number of faulty mRNAs, these specific mRNA degradation pathways target only a subset of mRNAs which are recognized by the RBP. By employing these mechanisms, cells are able to selectively degrade certain mRNA molecules and preserve the dynamic equilibrium of the transcriptome in response to changing conditions. The presence of multiple pathways for mRNA decay underscores the adaptability of cells to different environments.

3) Pulling the trigger on RNA decay

Recruitment of general cytoplasmic RNA decay factors is a shared feature of the aforementioned mRNA decay pathways. For complete mRNA degradation, the ends of the molecule must be exposed to the action of exonucleases. The 5' cap and 3' poly(A) tail structures shield mRNA ends from exonucleases which only target mRNA ends possessing a 5' monophosphate or a 3' hydroxyl group. Consequently, these protective structures often represent the primary targets for mRNA degradation, ensuring efficient decay for clearance of aberrant transcripts or mRNA turnover.

The majority of mRNAs are degraded in a very specific temporally orchestrated mechanism. The current view of cytoplasmic general mRNA decay is that deadenylation and decapping are coupled, as shortening of the poly(A) tail was thought to be a prerequisite to stimulate decapping and exoribonucleolytic decay through the 5' end (reviewed in Łabno et al., 2016) (Figure 7). Shortening of the poly(A) tail by the deadenylases Pan2/3 and Ccr4/Not followed by the recruitment of the Lsm1-7/Pat1 complex allows the recruitment and activation of the decapping complex Dcp1/Dcp2. This in turn provides a substrate to Xrn1, a 5'-3' exonuclease and to 3'-5' decay by the cytoplasmic exosome complex (Łabno et al., 2016). While deadenylation has always been considered the limiting step of RNA degradation, recent evidence suggests that decapping itself is the limiting step and decapping activation is independent of deadenylation speed (Audebert et al., 2023).





Illustration of the different complexes involved in general cytoplasmic RNA decay. The mRNA is first deadenylated by the Ccr4/Not and Pop complexes (red). The deadenylated mRNA is then recognized by the Lsm/Pat1 complex (orange) which triggers 5' decapping by Dcp1/2 and the Edc co-factors (brown). The decapped mRNA is then degraded by Xrn1 (gray). The deadenylated mRNA is further degraded by the Exosome/SKI (blue/yellow) complex and the remaining capped nucleotide is degraded by DcpS (green).

a) Starting from the tail: deadenylation and 3'-5' decay

Deadenylases

There are two main deadenylation complexes Pan2/3 (poly(<u>A</u>)-<u>n</u>uclease <u>2/3</u>), and Ccr4/Not (<u>c</u>arbon <u>c</u>atabolite <u>r</u>epression <u>4/n</u>egative regulator <u>of</u> <u>t</u>ranscription) (Decker & Parker, 1993; Tucker et al., 2001). Pan2 belongs to the RNAse D family and is the catalytically active protein. Its cofactor Pan3, regulates Pan2 activity. Pan2/3 act on the preliminary poly(A) trimming (Wolf & Passmore, 2014). Following this preliminary step, the Ccr4/Not

multiprotein complex is recruited (Collart & Panasenko, 2017). The catalytically active proteins Ccr4 (CNOT6 in mammals) and Pop2 (Caf1 in mammals) interact with their cofactor Not1 (Basquin et al., 2012). Their activities are redundant but they target different mRNA subsets (Maryati et al., 2015), as Ccr4 mainly deadenylates PABP-bound mRNAs unlike Pop2 which prefers naked poly(A) tails (Webster et al., 2018). Other cofactors such as Not2, Not3 and Not5 link deadenylation with decapping (Alhusaini & Coller, 2016). Additional poly(A) ribonucleases exist and are involved in more specific pathways. For example, PARN is able to bind 5' caps and is involved in piRNA biogenesis, and Nocturin, Angel1 and Angel2 which are involved in circadian rhythm (Godwin et al., 2013).

Cytoplasmic 3'-5' Exosome & cofactors

Once the RNA is deadenylated, the 3' hydroxyl end is accessible for processive exonuclease degradation performed by the cytoplasmic exosomal multiprotein complex (Lubas et al., 2013; Malecki et al., 2013). This conserved core (Exo9) is composed of 9 subunits and binds the catalytically active protein Rrp44 (Dis3) to form the active exosome (Exo10) (Schaeffer et al., 2009; Zinder & Lima, 2017). Within the tunnel-like structure formed by Exo9, Rrp44 is the only catalytically active protein and it exhibits exo- and endonucleolytic activities (Lebreton et al., 2008; Schaeffer et al., 2009). The Ski7 protein bridges the main exosomal cofactor: the SKI complex (super killer). Composed of Ski2, Ski3 and Ski8, the SKI complex facilitates the mRNA's entry into the Exo9 structure through the helicase activity of Ski2 (Halbach et al., 2013). Once the totality of the RNA has been trimmed, the m7G cap is removed from the oligonucleotide by DcpS (scavenger decapping, Dcs1 in yeast) (van Dijk et al., 2002; Liu et al., 2008). Unlike the canonical decapping proteins, DcpS is part of the HITT family (histidine triad). It has higher affinity for oligonucleotides and cleaves the bond between the α and β phosphates, releasing m7GMP (Łabno et al., 2016).

Recruitment of 3'-5' degradation for specialized RNA decay pathways

As mentioned before, the third step of specific mRNA decay pathways is to recruit general mRNA decay factors. Indeed, during NSD, the Ski7 protein which has a structure similar to that of eEF1A and eRF3 is recruited to the A site of the ribosome to allow the recruitment of the SKI complex to the exosome and trigger 3'-5' decay (van Hoof et al., 2002; Graille & Séraphin, 2012). The miRNA-RISC complex recruits the Ccr4/Not complex (Orban & Izaurralde, 2005; Giraldez et al., 2006; Behm-Ansmant et al., 2006). And during NMD, Upf1 recruits Smg5/Smg7 to link the deadenylation machinery (Loh et al., 2013). Furthermore, Ccr4/Not is also recruited to AREs by Ttp *via* Not1 (Fabian et al., 2013).

b) Starting from the head: decapping and 5'-3' decay

Decapping & cofactors

For 5'-3' degradation, a decapping step by Dcp1/Dcp2 (decapping proteins 1/2) and its cofactors Edc1-2-3 (enhancers of decapping) are required to liberate a 5' monophosphate extremity (van Dijk et al., 2002). Despite early studies suggesting that Dcp1 was the active decapping enzyme, it is now known that Dcp1 and Dcp2 together form a holoenzyme in

which Dcp2 is the catalytically active protein. Dcp2 bears an N-terminal MutT/Nudix domain involved in pyrophosphatase activity (Figure 8) and Dcp1 is its essential cofactor which directly binds Dcp2 upstream the Nudix domain (Figure 8) (van Dijk et al., 2002; Steiger et al., 2003). Dcp2 hydrolyses the 5' cap by cleaving the pyrophosphate bond between phosphates β and γ , which releases m7GDP and a 5' monophosphorylated RNA (Stevens, 1988). Without Dcp1, Dcp2 has little to no activity, neither *in vitro* nor *in vivo* (Arribas-Layton et al., 2013). This essential cofactor is principally made of one domain: EVH1 (Ena/Vasp homology 1) (Figure 8), implicated in protein-protein interactions within proline-rich sequences (Ball et al., 2002; She et al., 2004). Dcp1 increases the activity of Dcp2 by 100-fold but does not influence its affinity for 7mG cap (Charenton et al., 2016).

The decapping holoenzyme has a multiprotein cofactor: the LSm1-7 complex (Like-<u>Sm</u>) which forms a ring-like structure and recruits Dcp1 (Bouveret et al., 2000; Tharun et al., 2000) along with Pat1 (Sharif & Conti, 2013). The LSm1-7/Pat1 complex is associated to oligoadenylated RNAs *in vivo* (Tharun & Parker, 2001) and Pat1 inhibits translation (Nissan et al., 2010). The conserved decapping cofactor Edc3 (enhancer of decapping 3) activates the decapping catalytic activity for only a subset of mRNAs (Badis et al., 2004; Dong et al., 2007). Other cofactors that directly influence Dcp2 catalytic activity are Edc1 and Edc2 (homologous to PNRC2 in human). Edc1/2 stabilize the active conformation of the holoenzyme (Charenton & Graille, 2018). In human, Edc4 (also called Ge-1) is another decapping activator which does not have a homologue in yeast but activates decapping *in vitro* (Arribas-Layton et al., 2013; Yu et al., 2005). Despite not being essential for its activity in yeast (Dunckley & Parker, 1999), the C-ter domain of Dcp2, which is not conserved in humans, bears multiple HLMs (helical leucine-rich motifs) allowing the binding of various cofactors (Figure 8), which link decapping to mRNA decay (Fromm et al., 2012; He & Jacobson, 2015).



Figure 8: Structural domains of *S. cerevisiae* decapping complex Dcp1/Dcp2

Dcp1 is composed by the EVH1 domain, which interacts directly with the N-terminal domain of Dcp2 named DBD (Dcp1-binding domain). Dcp2 N-terminal Nudix domain bears catalytic decapping activity. Downstream of the catalytic domain, Dcp2 binds the Edc3 co-factor through the E3BM (Edc3-binding motif). In its C-terminal domain, Dcp2 bears multiple HLM domains that bind Pat1 and two Upf1-binding domains (UBD).

Cytoplasmic 5'-3' exonuclease Xrn1

Once the RNA is irreversibly decapped, it is destined to be degraded by Xrn1 which can happen in a co-translational manner (Hu et al., 2009; Nagarajan et al., 2013; Pelechano et al., 2015; Serdar et al., 2016). This conserved enzyme is an extremely processive and rapid 5'-3' exonuclease as it degrades 26 +/- 5 nt/s and has a processivity of over 10.5 kb in yeast (Athapattu et al., 2021). It recognizes 5' monophosphate as a substrate and liberates

monophosphate nucleotides while processively degrading the RNA (Stevens, 1979, 1988; Jinek et al., 2011). In humans, Xrn1 is directly recruited to the substrates through Dcp1 and Edc4 (Braun et al., 2012). In yeast, it is recruited through the LSm1-7/Pat1 complex (Bouveret et al., 2000; Krogan et al., 2006; Nissan et al., 2010).

Recruitment of the 5'-3' degradation for specialized RNA decay pathways

Other than deadenylation, mRNA decay pathways can also recruit decapping to degrade the RNA. For example, decapping is the predominant degradation pathway of NMD substrates in *S. cerevisiae* (Muhlrad & Parker, 1994) and *A. thaliana* (Nagarajan et al., 2013). It has been proposed that in yeast Upf1 recruits Dcp2 (He & Jacobson, 2015; He et al., 2022). Indeed, two-hybrid assays showed that Dcp2 interacts with Upf1 through two UBDs (<u>Upf1-binding domains</u>) in its C-terminal domain (He & Jacobson, 1995, 2015; He et al., 2022) (Figure 8). Both domains share conserved amino acids and have redundant activities (He et al., 2022). Nevertheless, the specific biochemical aspects of the Upf1-Dcp2 interaction are still missing as no proof of it being direct has been reported and no regulation of Upf1 nor Dcp2 activities have been investigated (see Results). In humans, Upf1 directly recruits the decapping activator Pncr2 (Cho et al., 2009; Lai et al., 2012) to trigger decapping.

In HMD, once Slbp is recruited to the histone mRNA, it recruits Upf1 which is phosphorylated by ATR and DNA-PK kinases (Kaygun & Marzluff, 2005; Choe et al., 2014). The complex Slbp/phospho-Upf1 recruits PNRC2 for decapping (Choe et al., 2014). In addition, during GMD, the glucocorticoid receptor bound to the RNA recruits PNRC2 for decapping (Cho et al., 2015; Park et al., 2016). Similarly, Dcp1/Dcp2 are recruited during miRNA mediated mRNA silencing (Rehwinkel et al., 2005; Behm-Ansmant et al., 2006). Furthermore, Ttp activates decapping in AMD (Fenger-Grøn et al., 2005).

c) Starting from the middle: endonucleolytic decay

An alternative way to generate 5' monophosphate and 3' hydroxyl extremities simultaneously is to cleave the RNA. This can be achieved by the action of endonucleases (reviewed in Tomecki & Dziembowski, 2010; Matelska et al., 2017) (Figure 9). In addition to the well characterized endonucleolytic cleavages induced by RNAi mechanisms, numerous endonucleases are recruited to targeted decay pathways: NMD recruits Smg6 (Okada-Katsuhata et al., 2012; Chakrabarti et al., 2014; Nicholson et al., 2014), RIDD recruits Ire1 (Hollien et al., 2009), RMD recruits Regnase 1, and GMD recruits Hrsp12/Rida (Mistiniene et al., 2005). As for NGD and NSD, the endonucleolytic activity of Dom34 is debated (Passos et al., 2009; Tomecki & Dziembowski, 2010), but recent studies showed that Cue2 is the endonuclease involved in both these pathways (D'Orazio et al., 2019). Furthermore, the RISC complex itself presents endonucleolytic activity through the PIWI domain of the Ago proteins (Song et al., 2003; Martinez & Tuschl, 2004).



Figure 9: Endonucleolytic degradation of mRNA Illustration of the different complexes involved after endonucleolytic RNA decay (green). The exosome/SKI complex degrades the 3' end (blue/yellow) and remaining capped nucleotide is degraded by DcpS (brown). The 5' end is degraded by Xrn1 (gray).

d) P-bodies: mRNA degradation or translation repression?

P-bodies (processing bodies) are conserved dynamic membrane-free organelles found within the cytoplasm and formed by liquid-liquid phase separation (reviewed in Luo et al., 2018; Standart & Weil, 2018). They house mRNPs composed of non-translating RNAs and several factors involved in RNA metabolism and decay. Notable components of P-bodies include decapping factors (Dcp1, Dcp2, Edc3, Lsm1-7, Pat1), deadenylases (PARN, Ccr4/Not), exoribonucleases (Xrn1), RNA silencing factors (Ago1, Ago2), NMD proteins (Upf1, Smg7), and several other RBPs including translation repressing factors (reviewed in Parker & Sheth, 2007; Decker & Parker, 2012). For a long time, it was thought that P-bodies were where general mRNA decay and NMD took place, since oligoadenylated and PTC-containing transcripts accumulate in P-bodies. However, it has been demonstrated that P-bodies are not required for mRNA degradation (Eulalio et al., 2007), that mRNPs can exit P-bodies to continue translation (Brengues et al., 2005; Bhattacharyya et al., 2006) and that mRNA degradation intermediates are not enriched within P-bodies (Horvathova et al., 2017; Courel et al., 2019). Purification of these structures using fluorescence-activated particle sorting (FAPS) revealed their composition via high throughput proteomic and transcriptomic analyses (Hubstenberger et al., 2017). They showed no accumulation of 5' degraded mRNAs, confirming that in P-bodies, mRNAs are not degraded but rather their translation is repressed.

Chapter 3: Diving deeper, molecular basis of NMD

NMD is a translation-dependent mRNA quality control pathway. It targets and degrades nonsense-containing transcripts in the cytoplasm and the main orchestrators are the Upf proteins (Upf1, Upf2 and Upf3). This chapter aims to describe the main factors and molecular mechanisms of NMD in eukaryotes and highlight how, despite 40 years of research, many players and aspects are still misunderstood.

1) A brief history of NMD research

In the late 1970s - early 1980s, independent research teams discovered that nonsense gene mutations rendered unexpectedly low amounts of the corresponding mRNA in S. cerevisiae and human cells (Chang & Kan, 1979; Losson & Lacroute, 1979; Maquat et al., 1981). In the early 1990s, Upf1 along with Upf2 and Upf3 were found to be assembled into a "surveillance complex" in yeast and in C. elegans and to be the essential proteins for the mechanism (Leeds et al., 1991, 1992; Pulak & Anderson, 1993; Cui et al., 1995). NMD was subsequently discovered in most eukaryotic organisms (reviewed in Lloyd, 2018), proving it was a universally conserved mechanism. The term 'nonsense-mediated mRNA decay' was coined in the mid 1990s (Peltz et al., 1993), which coincided with the research peak of the mechanism. The link with translation was described in the late 1990s (Pulak & Anderson, 1993; Carter et al., 1995; Zhang & Maquat, 1997; Thermann et al., 1998). Since then, numerous research laboratories have made significant contributions to the NMD field. Research on NMD led to the discovery of the EJC (Le Hir, Izaurralde, et al., 2000). Today, NMD is certainly the most studied eukaryotic quality control mechanism, yet, many aspects are still debated and despite almost four decades of research in this field, no satisfactory model has been described to unify all of the findings.

2) NMD in yeast: detect, recruit and degrade

a) The NMD holy trinity: Upf1/Upf2/Upf3

The main factor of NMD is the helicase Upf1, originally discovered in *S. cerevisiae* (Culbertson et al., 1980) along with Upf2 and Upf3 (Leeds et al., 1991). All three factors are conserved across eukaryotes and co-immunoprecipitation experiments showed they can form a trimeric complex (He & Jacobson, 1995; Weng et al., 1996a, 1996b; He et al., 1997, Dehecq et al., 2018). In addition, all Upf proteins sediment together in the same polysomal fractions (Atkin et al., 1997). Interestingly, depletion of Upf3 leads to loss of Upf2 in the polysomes indicating a tight link between these two proteins (Atkin et al., 1997). Upf1 is mainly found in the cytoplasm but can also be nuclear. Upf2 is perinuclear and Upf3 is mainly nuclear (Lykke-Andersen et al., 2000). The relative abundance of Upf1 in the cell is 10-fold higher than that of Upf2 and Upf3 (Maquat & Serin, 2001).

Among the Upf proteins, only Upf1 is catalytically active as the *upf1* gene encodes for an ATP-dependent RNA helicase (Weng et al., 1996a). The helicase domain (HD) is flanked in N-terminus by a cysteine-histidine rich domain (CH) and by a disordered C-terminal tail (Figure 10). More details about Upf1 and its catalytic activity will be described in chapter 5. Upf2 encodes for an acidic protein composed of three MIF4G (middle domain of e<u>IF4G</u>) domains (Cui et al., 1995; He & Jacobson, 1995) and Upf3 encodes for a basic protein which
shuttles between the cytoplasm and the nucleus (Lee & Culbertson, 1995; Shirley et al., 1998) (Figure 10). In S. *cerevisiae*, 11 putative phosphorylation sites have been reported for Upf1, among which two tyrosine residues have been reported to regulate ATPase activity (Lasalde et al., 2014). In addition, phosphorylation of several residues in the N-ter region of Upf2 appears important for NMD (Wang et al., 2006). Yet, their role in NMD and the kinase(s) responsible for these phosphorylations have not been determined to date.



Figure 10: Structural domains of the yeast Upf proteins

Upf1 is composed of an N-terminal cysteine and histidine rich domain (CH, green) and a central helicase domain (HD, orange). Upf2 contains three MIF4G domains (blue). Upf3 contains an RNA recognition motif (RRM, light blue).

b) Implication of the Upf proteins and translation termination

Mutations on the upf genes abolish NMD which results in stabilization of PTC-containing transcripts and other NMD targets (see Chapter 4). In yeast, a second phenotype independent of NMD is associated with upf deletion: nonsense suppression. This phenotype was first observed using yeast strains (Tyr- or can1-100) unable to grow on certain media (lacking tyrosine or containing canavanine respectively), which would lose their lethality when mutating the upf genes. Indeed, when a near cognate tRNA (capable of base pairing with stop codons at two of the three codon positions) is recruited to a premature stop codon instead of the termination factors, the ribosome reads through the stop codon and an untruncated protein is synthesized (Leeds et al., 1992; Weng et al., 1996a, 1996b; Czaplinski et al., 1998; Maderazo et al., 2000; Wang et al., 2001). Synthesis of the protein allowed the survival of the yeasts in the said media. This suggested that the Upf proteins were involved in translation termination of nonsense codons. Indeed, Upf1 interacts with both release factors eRF1 (Sup45 in yeast) and eRF3 (Sup35 in yeast) (Czaplinski et al., 1998, 1999). In addition, Upf2 and Upf3 bind eRF3 in a mutually exclusive manner, compete with eRF1 for eRF3 binding and enhance translation termination (Wang et al., 2001). Furthermore, Upf1 interacts with the 40S ribosomal subunit protein Rps26 (Min et al., 2013). However, the precise role of Upf1 in translation termination is still misunderstood.

c) The Faux 3' UTR model for yeast NMD

NMD requires discriminating between a premature and a *bona fide* stop codon to recruit the Upf trinity and decay-inducing factors. However, no clear hints as to the signals of distinction between these two identical codons are. The current working model for yeast NMD is known as the "Faux 3' UTR" model (Amrani et al., 2004; Karousis & Mühlemann, 2019; Kurosaki et al., 2019; Muñoz et al., 2023) (Figure 11). It is based on the observation that translation termination at a premature termination codon (PTC) is slower compared to

Introduction Chapter 3

normal termination (Amrani et al., 2004). This inefficiency is attributed to the disruption of interactions between the PABPs, eRF1/3, and the ribosome (Ivanov et al., 2008; Ivanov et al., 2016; Wu et al., 2020). In the Faux 3' UTR model, elongation of the 3' UTR caused by a PTC hinders the efficient termination process, facilitating the recruitment of Upf proteins and initiating degradation (Amrani et al., 2004). The primary mode of RNA degradation in yeast NMD is deadenylation-independent decapping (Muhlrad & Parker, 1994) followed by Xrn1 exoribonucleolytic decay. Indeed, Upf1 would recruit the decapping heterodimer Dcp1/Dcp2 (He & Jacobson, 1995; He et al., 2022) to trigger decay. Supporting this model is the fact that tethering PABP next to a PTC suppresses NMD (Amrani et al., 2004; Silva et al., 2008).



Figure 11: Faux 3' UTR model for yeast NMD (Based on Amrani et al., 2004)

Translation termination at a premature termination codon (PTC) is less efficient compared to normal termination, leading to disrupted interactions between the poly(A) binding proteins (PABPs), eRF1/3, and the ribosome. The elongated 3' untranslated region (Faux 3' UTR) resulting from the PTC impedes efficient termination, facilitating the recruitment of Upf proteins and triggering mRNA degradation.

While the Faux 3' UTR model currently serves as the prevailing mechanistic explanation for yeast NMD, the PABPs and a poly(A) tail are not absolutely necessary for NMD triggering (Meaux et al., 2008). Thus, the Faux 3' UTR model still lacks precise understanding regarding Upf recruitment and the triggering of decay.

Altogether, yeast NMD may be a simple three-step mechanism: (1) Translation termination and recruitment of eRF1, (2) Termination is detected as premature probably due to kinetic differences in termination efficiency. If termination is premature, the Upf1-2/3 surveillance complex is recruited. (3) The surveillance complex recruits decapping proteins Dcp1/Dcp2 and triggers RNA decay independently of deadenylation.

3) NMD in metazoans: a higher level of complexity

a) A more diverse Upf core complex

In humans, the C-terminal domain of Upf1 is named SQ in account of the multiple serines and glutamines. Two Upf1 alternative splicing isoforms exist and differ in the length of a small regulatory loop within the domain 1B (Gowravaram et al., 2018; Fritz et al., 2022): isoform 1 has a 22-residue long loop, and isoform 2 has an 11-residue long loop (Figure 12A). Their relative quantity varies among cellular types (Gowravaram et al., 2018) but isoform 1 is the less abundant one.



Figure 12: Structural domains and solved structures of the human Upf proteins

(A) Schematic structure and domain boundaries of the human Upf proteins. Upf1 is composed of an N-terminal cysteine and histidine rich domain (CH, green), a central helicase domain (HD, orange) and a serine and glutamine rich C-terminal domain (SQ). Two isoforms (Upf11 and Upf12) differ from 11 amino acids within the helicase domain. Upf2 contains three MIF4G domains (blue), it directly interacts with the Upf1 CH domain through its C-terminal domain. Two isoforms of Upf3 exist (Upf3a and Upf3b), both contain an RNA recognition motif and EJC-binding domain (RRM and EBM, light blue), they interact with the third MIF4G domain of Upf2 through the RRM domain. (B) Crystal structure of Upf1 CH (green) and HD (orange) interaction with Upf2 C-terminus (blue) domain (PDB: 2WJV from Clerici et al., 2009) (C) Crystal structure of the interaction between Upf2 third MIF4G domain (blue) and Upf3 RRM domain (cyan) (PDB: 1UW4 from Kadlec et al., 2004).

Upf2 orthologs only share 21% homology between yeasts and humans (Culbertson & Leeds, 2003). The third MIF4G domain directly binds Upf3 (Kadlec et al., 2004), and the C-ter tail directly binds Upf1's N-ter domain (Kadlec et al., 2006; Clerici et al., 2009) (Figure 12A, 12B). Mutations in the first MIF4G domain abolish NMD (Fourati et al., 2014).

Among the Upf proteins, Upf3 is the less conserved one. Upf3 contains an RRM (RNA recognition motif) domain via which, instead of interacting with RNA, it interacts with Upf2 (Kadlec et al., 2004) (Figure 12A, 12C). In humans there are two paralogs of this protein issued from gene duplication: Upf3a and Upf3b (a.k.a. Upf3X) (Figure 12A). The upf3a gene, which is located in chromosome 19, encodes for a 476 amino acid protein. The upf3b gene, also named upf3X due to its chromosome X location, encodes for a 483 amino acid protein. They mainly differ in their C-terminal domains and their expression patterns in different cell types. Upf3a is highly expressed in cell types where chromosome X is silenced (Serin et al., 2001). It was previously thought that Upf3a and Upf3b had redundant activities in NMD as RNA tethering of both led to RNA destabilization (Lykee-Andersen et al., 2000). However more recent studies showed markedly different activities as Upf3a can negatively regulate NMD by sequestering Upf2 in the nucleus (Kunz et al., 2006, Shum et al., 2016). In addition, although both Upf3 proteins can bind the EJC through their EBM (EJC-binding motif) they have different affinities (Kunz et al., 2006; Buchwald et al., 2010) (Figure 12A). Upf3b is involved in interactions with eRF3 and dissociates post-termination ribosomes (Neu-Yilik et al., 2017). In vitro interaction assays showed that Upf1 and Upf3 proteins do not interact directly with each other but via their interaction with Upf2 (Chamieh et al., 2008). However, recent co-immunoprecipitation data suggest that Upf3b is able to bind Upf1 in absence of Upf2 (Neu-Yilik et al., 2017). While the isolated trimeric complex structure has not yet been solved, the interactions Upf2-Ct/Upf1-CH and Upf2-MIF4G/Upf3 have been solved independently (Figure 12B, 12C, Clerici et al., 2009; Kadlec et al., 2004) and there is a CryoEM structure of the Upf complex bound to the EJC (Melero et al., 2012).

b) Metazoan specific NMD factors

Exon Junction Complex

In metazoans, especially mammals, NMD is tightly related to mRNA splicing. Indeed, it was the study of NMD which led to the discovery of the EJC in humans (Thermann et al., 1998; Zhang et al., 1998; Le Hir, Izaurralde, et al., 2000; Le Hir, Moore, et al., 2000). During the process of pre-messenger RNA splicing in the nucleus, the spliceosome deposits the EJC at -24 to -27 nucleotides downstream of exon-exon junctions (Le Hir, Izaurralde, et al., 2000; Hocq et al., 2018). This implies that a given mRNA is most likely loaded by multiple EJCs. The core of this complex is formed by three proteins conserved only in metazoans: the DEAD-box helicase eIF4A3 (eukaryotic initiation factor 4A3), Y14 (or Rbm8A, RNA binding motif protein 8A) and Magoh (Mago nashi human homologue) (Ballut et al., 2005). This core complex stably clamps the RNA in a sequence independent manner (Ballut et al., 2005; Andersen et al., 2006; Bono et al., 2006). The EJC is involved in many processes such as pre-mRNA splicing, mRNA transport, localization and translation, in addition to being important for NMD activation (reviewed in Le Hir et al., 2016). Once it is loaded onto the

mRNA in the nucleus, the EJC accompanies the mRNP into the cytoplasm. Among the nuclear peripheral factors of the EJC, we find multiple transcription factors, spliceosomal proteins and export factors (Le Hir et al., 2016). The EJC can also bind NMD factors such as Upf3 and Smg6 (Gehring et al., 2003; Kashima et al., 2010). In the cytoplasm, Upf2 binds the Upf3-bound EJC and links Upf1 to the RNA (Chamieh et al., 2008; Buchwald et al., 2010; Melero et al., 2012). Interestingly, this is the only known example of direct recruitment of Upf1 to the RNA during NMD. Therefore, the presence of a stop codon at least 50-55 nucleotides upstream a spliced junction carrying an EJC elicits NMD in mammals (Nagy & Maquat, 1998; Thermann et al., 1998; Zhang et al., 1998). This can happen either by the presence of a PTC or if the RNA has a spliced 3' UTR intron. The EJC is then disassembled by the translating ribosome as it passes through the RNA during the pioneer round of translation (Maquat et al., 2010) and by the interaction between Y14/Magoh with PYM (partner of Y14 and Magoh) (Gehring et al., 2009).

<u>Smg1/8/9</u>

In metazoans, Smg1 and Smg8/9 proteins are required for NMD activation and responsible for the phosphorylation cycle of Upf1. Smg1 is a serine/threonine kinase of the PIKK (<u>phosphatidylinositol 3-kinase related kinase</u>) family (Denning et al., 2001). It is responsible for the phosphorylation of the serine/threonine-glutamine (S/T-Q) motifs present on the and CH (N-ter) and SQ (C-ter) domains of Upf1 (Yamashita et al., 2001; Grimson et al., 2004). This phosphorylation is regulated by Smg8 and Smg9 which form a complex with Smg1 at the RNA and inhibit its kinase activity (Yamashita et al., 2009; Fernández et al., 2011; Arias-Palomo et al., 2011; Langer et al., 2020).

Smg5/Smg7 and Smg6

Phosphorylated Upf1 recruits the phosphopeptide binding 14-3-3 like domains of Smg5/7 and Smg6 *via* its phosphorylated S/T-Q motifs in the CH and SQ domains (Ohnishi et al., 2003; Chakrabarti et al., 2014) (Figure 13A). Smg5 and Smg7 form a heterodimer through interactions between the 14-3-3 domains (Fukuhara et al., 2005; Jonas et al., 2013) (Figure 13C). The dimer is recruited primarily to the SQ domain of Upf1 and in turn recruits the Ccr4/Not complex allowing for deadenylation of the target RNA (Unterholzner & Izaurralde, 2004; Jonas et al., 2013). This dimer also contributes to the dephosphorylation of Upf1 by interacting with PP2A (protein phosphatase <u>2A</u>) (Anders et al., 2003).

Smg5 and Smg6 share a PIN (<u>Pi</u>IT <u>N</u>-terminus domain) of the RNAseH family (Figure 13A). Smg6 can weakly interact with phosphorylated Upf1 but is predominantly recruited to the HD of Upf1 regardless of its phosphorylation state (Chakrabarti et al., 2014; Nicholson et al., 2014). In addition, the EJC directly interacts with Smg6 (Kashima et al., 2010) and recent results show that Mln51, an EJC cofactor, stimulates Smg6 activity (Gerbracht et al., 2020). The PIN domain of Smg6 exhibits endonuclease activity and cleaves NMD substrates (Glavan et al., 2006; Eberle et al., 2009), unlike Smg5 which has an inactive PIN domain (Huntzinger et al., 2008) (Figure 13B). Global analyses of NMD target sequences revealed that Smg6 cleaves within a degenerate pentameric motif of single stranded RNAs (Schmidt, 2015). The degradation pathways of Smg6 and Smg5/Smg7 were thought to be



independent from each other, however, recent results suggest otherwise as depletion of Smg5/Smg7 leads to Smg6 endonucleolytic activity inhibition (Boehm et al., 2021).

Figure 13: Structural domains of the human Smg proteins

(A) Schematic structure and domain boundaries of the human Smg proteins. Smg5 (purple), Smg7 (pink), and Smg6 (green) feature a 14-3-3 protein-protein interaction domain at their N-terminus. In their C-terminal regions, both Smg5 and Smg6 possess an endonucleolytic PIN domain; of which only Smg6 is active. The catalytic residues within the PIN domain are marked in red. (B) Crystal structure of the Smg5/Smg7 dimer of *C. elegans* (PDB: 3ZHE from Jonas et al., 2013) (C) Crystal structures of the PIN domains of Smg5 (left, purple) (PDB: 2HWY from Glavan et al., 2006) and Smg6 (right, green) (PDB: 2HWW from Glavan et al., 2006) The three residues within the active sites (in red) are labelled.

c) Current mechanistic view of NMD in humans

EJC-dependent NMD: SURF/DECID

The current prevailing model for NMD triggering in humans is the SURF/DECID model (Kashima et al., 2006; Kurosaki et al., 2019; Yi et al., 2021) which is based on the fact that mRNAs with a PTC at 50-55 nucleotides upstream the last exon-exon junction are NMD targets (Nagy & Maquat, 1998) (Figure 14). The translational machinery is expected to remove all RBPs coating the mRNA as it passes during the so-called 'pioneer round of translation' (Maquat et al., 2010). This is the case of the EJCs deposited on the exonic junctions during the splicing of pre-messenger RNAs: the ribosome dissociates all EJCs in the coding region with the help of the PYM protein (Gehring et al., 2009).

When the ribosome recognizes the presence of a stop codon, the termination factors eRF1/3 are recruited, along with Upf1 and Smg1 to form the SURF complex (<u>Smg1-Upf1-Release Factors</u>) (Figure 14). If the stop codon is premature, there are downstream Upf2/Upf3b-bound EJCs that remain bound to the RNA which provide the link

to Upf1, the ribosome and the EJC. The Upf1/Smg1/Upf2 interaction promotes phosphorylation of Upf1 by Smg1 and form the DECID complex (decay inducing complex). The latter is composed of phospho-Upf1, Smg1, Upf2/Upf3 recruits the Smg5/Smg7 dimer and Smg6 which trigger RNA decay either through recruitment of Ccr4/Not or endoribonuclease activity respectively (Kashima et al., 2006) (Figure 14). Interestingly, the EJC-dependent model is the only example of Upf3-Upf2 direct recruitment to the mRNA (Chamieh et al., 2008). Note that this model is not valid in *S. cerevisiae* due to the fact that EJC proteins are absent in this organism. In addition, the fact that NMD only happens during the pioneer round of translation (Maquat et al., 2010) has been extensively challenged (Maderazo et al., 2003; Gaba et al., 2005; Durand & Lykke-Andersen, 2013; Rufener & Mühlemann, 2013; Hoek et al., 2019) by showing that NMD can happen at any round of translation of an mRNA.



Figure 14: The SURF/DECID model of EJC-dependent NMD (Based on Kashima et al., 2006)

The termination factors eRF1/3, Upf1, and Smg1 form the SURF complex upon recognition of a stop codon by the ribosome. In the presence of premature stop codons, downstream exon junction complexes (EJCs) bound by Upf2/Upf3b remain associated with the RNA, establishing a link between Upf1, the ribosome, and the EJC. The Upf1/Smg1/Upf2 interaction leads to Upf1 phosphorylation by Smg1, resulting in the dissociation of eRF1/3. The DECID (decay-inducing complex), consisting of phospho-Upf1, Smg1, and Upf2/3, recruits the Smg5/7 dimer and Smg6. Smg5/7 triggers RNA decay through the recruitment of Ccr4/Not complex, while Smg6 induces endoribonuclease activity, ultimately leading to mRNA degradation.

EJC-independent NMD

While EJC-dependent NMD is widespread, an EJC downstream a termination codon is not an absolute necessity to trigger NMD (Bühler et al., 2006). This shows that some mechanistic details of NMD are still missing. The only constant NMD feature is length of the 3' UTR which seems to be determinant for NMD triggering (Eberle et al., 2008). The length between the poly(A) tail and the stop codon could, like the Faux 3' UTR yeast model, explain EJC-independent NMD in which a longer 3' UTR prevents PABP interaction with eRF3 and promote Upf1 binding to eRF3, eliciting thus NMD (Ivanov et al., 2008). Yet, no correlation between NMD targets and 3' UTR length have been observed in humans (Karousis et al., 2021; Tani et al., 2012). Studies also showed no different ribosomal occupancy in NMD targets versus non NMD targets (Tate et al., 2018; Karousis et al., 2020), suggesting that NMD activation is not due to slow kinetics of termination. Nonetheless, given the three dimensional structure of RNA, it is possible that length itself is not the determinant factor, but more the physical distance between the PABPs and the terminating ribosome (De & Mühlemann, 2022). This model is closer to that of yeast NMD (Amrani et al., 2004) and would suggest an evolutionary conservation of a basic form of NMD and the existence of NMD enhancing pathways in metazoans (Upf1 phosphorylation and EJC-enhanced Upf1 recruitment). Another more recent model, termed the "Detection & Potentiation" model, suggests that Upf1 detects abnormally long 3' UTRs upon binding and persists on the transcript. This persistence would lead to the recruitment of other Upf proteins, activation of Upf1, and initiation of NMD (Hogg & Goff, 2010). Nevertheless, EJC-dependent NMD is considered more efficient than EJC-independent NMD.

Upf2 and UPf3 independent branches of NMD

In addition to EJC-independent NMD in humans, two studies showed that NMD can also be triggered in absence of Upf2 (Gehring et al., 2005) and of Upf3b (Chan et al., 2007). Chan and collaborators focused on the TCR (<u>T-cell receptor</u>) mRNA which generates PTCs through programmed rearrangements and is a well known NMD target (Chan et al., 2007). They showed that siRNA depletion of Upf3b had no effect on NMD triggering of these mRNAs and identified 348 other transcripts that underwent Upf3b-independent NMD degradation (Chan et al., 2007). This novel NMD pathway was dependent on Upf1, EJC core protein eIF4A3 and the Smg proteins, proving that it was indeed an NMD decay and not another <u>Upf1-dependent RNA decay</u> (UMD) pathway. Furthermore, Gehring and collaborators showed that depletion of Upf2 had no effect on the Upf3b-bound EJC NMD activation (Gehring et al., 2005), suggesting that Upf1 can also be recruited in the absence of Upf2 to trigger decay in an EJC dependent manner. However, their results may be an experimental bias given that they tether the proteins, forcing the system. This mechanism is likely different from the Upf3-independent NMD shown by Chan and collaborators since TCR NMD requires Upf2 (Chan et al., 2007).

Limitations of the models

Altogether, the various NMD pathways described above highlight the clear gaps in the knowledge in the field. Today, no universal NMD mechanism has been able to unify all of the findings. Several major shortcomings of the current models can be highlighted: (1) The

assignment of essential roles to factors not conserved among eukaryotes despite NMD being conserved, (2) Upf1 recruitment and role are still largely misunderstood, (3) The events linking PTC recognition to degradation initiation are ambiguous, (4) the timing and dynamics of ribosome recycling and translation termination are not well defined.

The many variations of NMD in eukaryotes

The mechanism of NMD is present in virtually all eukaryotic organisms studied to date, and most eukaryotes share the Upf core complex. However, many other NMD factors have been lost or acquired during evolution.

In excavates (such as parasites), which are the most basal group in eukaryotes, Upf1 and Upf2 orthologues exist (Figure 15). In *Trypanosoma brucei* an ortholog of Upf1 exists, its knockdown had no effect on an NMD reporter but tethering experiments did decrease RNA levels (Delhi et al., 2011). In *Giardia lamblia* we find orthologs of Upf1 and Upf2, and overexpressing Upf1 led to an NMD reporter degradation (Chen et al., 2008), suggesting the existence of a Upf1-dependent simplified NMD in this organism. Other non-parasitic species among excavates such as *Naegleria gruberi* have orthologs of Smg1 and Smg9, indicating that a more complex NMD mechanism could have been present in the last common ancestor of eukaryotes (Causier et al., 2017) (Figure 15). Similarly, in plants, which diverged early in the tree of life, we find orthologues of the Upf1-2/3 core complex, but also





of Smg1 and Smg5/6/7 (Causier et al., 2017), further confirming the existence of a more complex NMD early in evolution (Figure 15).

The Smg1 kinase has been lost in evolution on numerous occasions, for example in excavates, ciliates, *S. cerevisiae* and *A. thaliana* (but not in other plants) (Lloyd, 2018). In most of these examples, Smg 8, Smg9 and the phosphorylated residues of Upf1 are also lost (Causier et al., 2017) (Figure 15). These evolutionary losses suggest the existence of alternative factors and/or mechanisms for NMD activation. For example, deletion of Smg1 in *D. melanogaster* has little to no effect on NMD (Chen et al., 2005), but is important when Smg5 is absent (Nelson et al., 2018).

Moreover, the Smg5/6/7 effector family has been widely diversified. In plants, there are only Smg7 orthologs (Lloyd, 2018). In T. thermophilia, a ciliate, there is a Smg6 ortholog named Smg6L which bears a NYN domain, potentially homologous to the PIN domain (Tian et al., 2017). Furthermore, recent phylogenetic studies showed the widespread loss of magoh and y14 genes across various lineages, such as the Saccharomycetaceae family including S. cerevisiae, and C. albicans (Boisramé et al., 2019). However, EJC core components and peripheral factors have been retained in early-branching yeast species, including the Yarrowiaceae and the Phaffomycetaceae families (Boisramé et al., 2019). These findings suggest independent losses of the genes in multiple lineages, occurring at least two times. In addition, not all metazoans require the EJC for NMD (Behm-Ansmant et al., 2007; Gatfield et al., 2003). Evidence suggests the co-existence of EJC-dependent and EJC-independent NMD in invertebrates such as D. melanogaster (Gatfield et al., 2003), C. elegans (Longman et al., 2007), and plants (Kerényi et al., 2008). This suggests that EJC-dependent NMD might constitute an alternative enhanced NMD pathway. Other intron-containing species such as S. pombe show no evidence of EJC-dependent NMD (Wen & Brogna, 2010). Interestingly, in T. thermophila, NMD depends on splicing but is independent of EJC (Tian et al., 2017). Thus, the role of EJC as a PTC determinant or just as an NMD enhancer is still debated.

4) Technological advances and current approaches to study NMD

Most of the current information regarding the composition and organization of NMD complexes comes from genetic screens (Culbertson et al., 1980; Dujardin et al., 1980)), two-hybrid studies (He & Jacobson, 1995), and *in cellulo* co-immunoprecipitation experiments (Kashima et al., 2006; Ivanov et al., 2008). In the late 90s and 2000s, countless studies used individual RNA reporters for studying NMD, especially NMD-triggering characteristics. Yet, nowadays high throughput sequencing techniques have replaced virtually every use of reporters and give a global view of NMD targets. Many novel techniques have been developed in the 21st century and pave the way towards a better understanding of NMD and Upf1 function. Advances in structural biology such as Cryo-EM have allowed easier, faster and more precise structure solving of protein complexes such as the Smg1/8/9 complex (Langer et al., 2020). Development of quantitative proteomics either coupled to proximity labeling (Schweingruber et al., 2016) or metabolic labeling (Flury et al., 2014) provided a more precise list of direct or indirect

partners of Upf1. Quantitative mass spectrometry coupled to affinity purification of multiple NMD factors allowed the clear description of Upf1-bound complexes (Dehecq et al., 2018). Overall, these techniques allowed a better understanding of NMD complex dynamics and revealed more transient interactions. Furthermore, microscopy advances coupled to real time and single molecule techniques allowed more precise tracking of NMD events in real time. For example, visualization of NMD targets at the protein level (Pereverzev et al., 2015; Gurskaya et al., 2016) or at the RNA level (Hoek et al., 2019; Dave et al., 2023) allowed to demonstrate that NMD can happen at any round of translation. Furthermore, single molecule characterization of Upf1 helicase activity using magnetic tweezers showed that it is a highly processive helicase (Fiorini et al., 2015; Kanaan et al., 2018). Finally, knock down enhancement using CRISPR/Cas9 instead of siRNA has also opened the path to a better characterization of the associated phenotypes (Boehm et al., 2021; Katsioudi et al., 2023).

Chapter 4: NMD, more than just molecular housekeeping

As previously discussed, NMD targets PTC-containing RNAs. It is the first discovered and the most studied cytoplasmic translation-dependent decay pathway. In this chapter, we will explore the essential role of this mechanism for the cell and how broad its effects are.

1) What does NMD recognize as a target?

a) Beyond quality control: whole transcriptome regulation

Early research on NMD primarily focused on the PTC-containing transcripts which are upregulated when essential NMD factors are depleted. However, advancements in large-scale genomics allowed the identification of a plethora of other transcripts that were impacted by NMD abolition (Mendell et al., 2004; Wittmann et al., 2006; Yepiskoposyan et al., 2011; Tani et al., 2012; Colombo et al., 2017). Indeed, it is estimated that up to 5-10% of the transcriptome is targeted by NMD in normal physiological conditions (Mühlemann & Jensen, 2012; Lykke-Andersen & Jensen, 2015). Recent studies using better tools to deplete NMD factors estimated considerably higher proportions: 20-40% of expressed genes would be affected by NMD in metazoans (Boehm et al., 2021) and more than 50 % in yeast (Malabat et al., 2015). Although these numbers might be overestimated due to indirect effects of Upf1 depletion and the large number of UMD pathways independent from NMD (see Chapter 5), they illustrate how NMD shapes transcriptomes by adding another layer of post-transcriptional gene expression control.

b) NMD-triggering characteristics

The guest to understand what characteristics make a transcript an NMD substrate has led to the classification of three main distinct groups by He & Jacobson 2015 (Figure 16). Class I: Aberrant mRNAs (i.e. PTC-containing) resulting from genomic mutations, splicing errors (Pulak & Anderson, 1993; Sayani et al., 2008) or erroneous gene rearrangements (Li & Wilkinson, 1998; He et al., 2003). Class II: Non-coding RNAs recognized by the translation machinery such as lncRNAs and RNAs from pervasive transcription, pseudogene transcription and transposable elements (Amrani et al., 2006; Kurihara et al., 2009; Rebbapragada & Lykke-Andersen, 2009; Hurt et al., 2013; Malabat et al., 2015). Given that these transcripts have no coding potential, they most likely bear termination codons relatively early in the sequence. Class III: "Normal" mRNAs bearing NMD-triggering characteristics such as upstream ORFs (Mendell et al., 2004), long 3' UTRs (Muhlrad & Parker, 1999; Amrani et al., 2004; Hogg & Goff, 2010), short ORFs (Decourty et al., 2014), selenocysteine codons (Moriarty et al., 1998; Seyedali & Berry, 2014), alternative polyadenylation products (Kishor et al., 2020), alternative isoforms in 5' TSS (Malabat et al., 2015) and alternative splicing (Lareau et al., 2007; Jaillon et al., 2008; McGlincy & Smith, 2008).

Two studies conducted by the same team showed that NMD targets have lower translational efficiency than non-NMD substrates in yeast, mRNAs with lower ribosome densities, high rate of out of frame translation and/or low codon optimality were also targeted by NMD (Celik et al., 2017; Ganesan et al., 2022). However, these results have

been recently challenged by a study that found no difference in ribosomal density between NMD sensitive and insensitive transcripts using an *in vitro* toeprinting assay on human cell lysates (Karousis et al., 2020).

As described in chapter 3, in mammals, the presence of an EJC downstream a stop codon is an NMD-triggering feature. This can occur either due to the presence of a PTC (class I) or by a spliced intron in the 3' UTR (class III). Two interesting facts are noteworthy: (1) Most essential NMD factors are involved in an autoregulatory loop, with mRNAs coding for the main NMD factors being themselves NMD targets (Huang et al., 2011; Yepiskoposyan et al., 2011; Longman et al., 2013; Degtiar et al., 2015), and (2) The placement of a stop codon close to the start codon fails to elicit NMD (Pereira et al., 2015) meaning that there is a limit on the NMD-triggering length of a 3' UTR. The vast number of different NMD-triggering characteristics highlights the complexity of NMD targets, contributing to the current lack of understanding of the mechanistic workings of NMD.



Figure 16: NMD triggering characteristics (Based on He & Jacobson, 2015)

RNAs with premature termination codons (PTC), long 3' untranslated regions (3'UTR), upstream open reading frames (uORF), exon junction complex (EJC) downstream a stop codon or splicing defects constitute NMD targets.

2) Physiological role and importance of NMD

a) Roles in cellular metabolism

The broad spectrum of NMD targets provides an early hint of how NMD is a crucial mechanism involved in essential cellular processes (Nasif et al., 2018). Indeed, NMD is involved in a panoply of pathways, most of which require rapid RNA degradation. For instance, embryonic development and viability (Medghalchi et al., 2001; McIlwain et al., 2010; Li et al., 2015; Nelson et al., 2016), neural development (Colak et al., 2013), hematopoiesis (Weischenfeldt et al., 2008), cell cycle regulation (Azzalin & Lingner, 2006), growth and proliferation (Avery et al., 2011; Lou et al., 2014), immune response (Li & Wilkinson, 1998; Gloggnitzer et al., 2014), circadian rhythms (Ri et al., 2019; Katsioudi et al., 2023) and stress response (Karam et al., 2015). The extensive involvement of NMD in cellular metabolism is notably illustrated by its implication in disease, that we will describe

further. However, given the multiple roles of Upf1 in many essential pathways, it is difficult to attribute a clear role when depleting it.

b) NMD efficiency modulation

Given that NMD regulates a large number of transcripts and plays an important role in multiple cellular pathways, it naturally can be regulated in different developmental and environmental contexts to promote adaptation. Mechanisms that inhibit Upf1 availability or function have evolved to modulate NMD efficiency. For instance, inhibition of Upf1 production by RNAi promotes differentiation of neuronal stem cells during neurogenesis (Bruno et al., 2011; Lou et al., 2014). Cleavage of Upf1 by caspases is triggered by severe DNA damage during chemotherapy treatments, which decreases NMD efficiency and promotes apoptosis (Jia et al., 2015; Popp & Maquat, 2015). Autoregulation also occurs through NMD itself, as NMD factors are NMD targets (Yepiskoposyan et al., 2011). Furthermore, there is also a tissue-specific variability of NMD efficiency (Zetoune et al., 2008) which is still very poorly characterized. Other aspects of NMD modulation (such as viral hijacking or during cancer progression) will be discussed further.

3) NMD in health and disease, a double edged sword

a) Genetic diseases

It is estimated that 30% of the 7000 known rare genetic human diseases involve premature termination (Miller & Pearce, 2014). Nonsense-mutations account for approximately 20% of single-base pair associated diseases (Mort et al., 2008). ß-Thalassemia, Cystic Fibrosis (CF), and Duchenne Muscular Dystrophy (DMD) are among the most studied monogenic diseases caused by single protein quantity or quality defects (Kerr et al., 2001; Linde et al., 2007). In these diseases, the nature of the mutations carried by the alleles of the implicated genes (HBB, CFTR and DMD respectively) determine the nature and severity of the phenotypes (Miller & Pearce, 2014). Contrary to missense mutations, nonsense mutations give rise to truncated proteins that are most likely non functional and have potentially double-negative effects, but also trigger the degradation of the mRNA through NMD, thus exacerbating the consequences.

b) Neurodevelopmental disorders

As a master gene expression regulator, NMD is essential during brain development. Indeed, even if NMD levels are downregulated during neurodevelopment (Bruno et al., 2011), its correct activity is crucial. Numerous studies have linked *upf3b* mutations to neurological disorders and intellectual disabilities, including intellectual disability, schizophrenia, and autism spectrum disorder (Tarpey et al., 2007; Laumonnier et al., 2010; Addington et al., 2011; Nguyen et al., 2012). These disorders are linked to the X chromosome, the location of the *upf3b* locus. Copy number variations of other NMD factors have also been associated with neurodevelopmental disorders suggesting that the correct functioning of NMD is essential (Nguyen et al., 2014). The few described mutations of *upf1* and *upf2* are linked to intellectual disabilities and autism (Sun et al., 2023).

c) A dual role in cancer

Two main aspects of NMD implication in cancer are noteworthy: (1) Genetic instability in tumors leads to the accumulation of nonsense mutations and thus generates NMD substrates (Lindeboom et al., 2016; Popp & Maguat, 2018; Fernandes et al., 2019). NMD activation is important to avoid the translation of truncated proteins with cytotoxic effects. However, NMD-triggering mutations in tumor-suppressor genes are enriched in tumors, which also makes NMD pro-tumoral (Ivanov et al., 2007; Lindeboom et al., 2016; Supek et al., 2021). (2) Mutations of NMD factors (Upf and Smg proteins) have been largely described in numerous cancer types. For instance, Upf1 is downregulated in pancreatic adenosquamous carcinoma, lung adenocarcinoma, ovarian cancer and glioma among others (Popp & Maquat, 2018). Additionally, low levels of Upf1 and Smg6 expression are correlated with low survival prognostic in cancer patients (Chen et al., 2021; Cowen & Tang, 2017), which makes NMD factors potential biomarkers. NMD could therefore be considered as a double-edged sword given it simultaneously protects against nonsense mutations yet can exacerbate certain physiological conditions when overactivated. It is therefore of essence to decipher the mechanistic basis of NMD in order to deepen our understanding of its pathological implication, and aim for therapeutic innovation.

d) Therapeutics

Over the past twenty years, many strategies have emerged to tackle nonsense-mutation related diseases. Most of them aim to suppress PTCs and restore protein function by translational readthrough, others aim to inhibit NMD (reviewed in Keeling et al., 2013; Martins-Dias & Romão, 2021). Initially, PTC readthrough was achieved by the use of aminoglycosides (antibiotics), such as gentamicin, which bind the ribosome's decoding center (A site) and decrease proofreading activity at low doses (Eustice & Wilhelm, 1984). The first proof of concept was described by Burke and Mogg in 1985 (Burke & Mogg, 1985). They showed that gentamicin efficiently induced readthrough of nonsense codons in yeast and mammalian cells (Burke & Mogg, 1985), although it did not show positive results for CF or DMD patients and presented high negative drawbacks such as long term toxicity (kidney damage, hearing loss) and antibiotic resistance (Martins-Dias & Romão, 2021). Since then, a second generation of aminoglycoside derivatives have been developed and are currently being tested in clinical trials. More recently, PTC therapeutics, the pharmaceutical company founded by Stuart Peltz in 1998, made significant advances on the development of new molecules using high throughput screens of small compound libraries which are now approved by the FDA and used to treat many diseases such as spinal muscular atrophy and DMD. Other approaches for PTC readthrough have been attempted with the use of nucleic acid-based methodologies (suppressor tRNAs, pseudouridylation, RNA editing, gene editing) (Keeling, Du, et al., 2013; Martins-Dias & Romão, 2021). In addition, strategies of NMD inhibition to avoid mRNA degradation and maintain steady-state mRNA levels combined with PTC readthrough to obtain better protein yields are being explored. NMD inhibition is more challenging due to the lack of specificity and the vast number of implications the mechanism has on cellular fate, but novel drugs such as NMDI 1 (NMD inhibitor 1) (Durand et al., 2007; Keeling, Wang, et al., 2013) and amlexanox (Gonzalez-Hilarion et al., 2012) are currently being tested.

e) Antiviral response

Viruses rely on the host's cellular machinery for their replication and have developed strategies to hijack host cellular pathways for their advantage (Balistreri et al., 2017; Popp et al., 2020; May & Simon, 2021; Sun et al., 2023). Host cells and viruses have been engaged in a co-evolutionary arms race. While host cells continually develop new defense strategies against viral attacks, viruses, in turn, adapt and find ways to circumvent these defenses. Thus, viral genomes and host cells have co-evolved for millenia. As a cytoplasmic quality control system, NMD acts as an intrinsic innate immunity pathway by targeting and degrading viral RNAs, especially since most viral RNAs bear NMD-triggering features such as long 3' UTRs and polycistronic ORFs. Several positive single-stranded RNA viruses and retroviruses including Rous sarcoma virus (RSV), Semliki forest virus (SFV), Potato virus X (PVX) and Turnip crinkle virus (TCV), Hepatitis C virus (HCV), Zika virus (ZIKV), Murine hepatitis virus (MHV) (LeBlanc & Beemon, 2004; Balistreri et al., 2014; Garcia et al., 2014; Ramage et al., 2015; Fontaine et al., 2018; Wada et al., 2018) are targeted by NMD. In order to escape NMD, viruses have developed two strategies: cis-avoidance of NMD or trans-inhibition of NMD (Popp et al., 2020; Sun et al., 2023). Cis-avoidance of NMD can be achieved by the recruitment of host proteins to RNA sequences or structures which shield the transcripts from NMD by preventing Upf1 binding (Ge et al., 2016), ribosomal readthrough elements to avoid NMD (Baker & Hogg, 2017; May et al., 2018) or unstructured RNAs which are NMD resistant (Popp et al., 2020). Additionally, certain viral proteins are capable of inhibiting NMD in trans through direct protein-protein interactions. Examples include the nucleocapsid protein (N) of Coronaviruses (Gordon et al., 2020), the HTLV-1 Tax protein (Fiorini et al., 2018) and the capsid protein of SFV (Contu et al., 2021) and of SIKV (Fontaine et al., 2018), that directly interact with Upf1 and inhibit NMD.

Chapter 5: Upf1, the cornerstone of NMD

1) Identification: from fortuitous genetic screens to homology

Upf1, Smg2, Nam7, Rent1 all refer to the NMD keystone. Upf1 was almost simultaneously discovered in S. cerevisiae by two different genetic screens: nuclear genes that suppressed mitochondrial mutations to restore respiratory competence (nuclear accommodation of mitochondria - nam - Dujardin et al., 1980; Altamura et al., 1992; Leeds et al., 1992) and genes that suppressed frameshift mutations (up-frameshift - upf) (Culbertson et al., 1980). Later, it was discovered in C. elegans through genetic screens looking for informational suppressors (suppressor with morphogenetic effect on genitalia - smg - Hodgkin et al., 1989) and identified in mammalians through homology resemblance as regulator of nonsense transcripts - rent (Perlick et al., 1996; Applequist et al., 1997). Eventually, the upf1 gene was described in virtually every eukaryotic organism studied to date: chimpanzees, dogs, cows, mice, rats, chickens, zebrafish, fruit fly, mosquito, other fungi (S. pombe, K. lactis, M. oryzae, N. crassa, E. gossypii), plants and frogs (Gupta & Li, 2018). Upf1 is essential among multicellular organisms: A. thaliana (Riehs-Kearnan et al., 2012), D. melanogaster (Avery et al., 2011), D. rerio (Wittkopp et al., 2009) and mammals (Medghalchi et al., 2001). Interestingly, the deletion of Upf1 in yeast is not lethal, suggesting that it assumes an increasingly important role in multicellular organisms.

2) Structure: Upf1 is a conserved tiny molecular engine

a) Helicases are ubiquitous

Helicases are tiny molecular motors that are involved in virtually every process involving NA (nucleic acids), from DNA replication, reparation, recombination, chromatin remodeling, telomere maintenance to RNA export, transcription, splicing, ribosome biogenesis, RNA storage, and more (Jankowsky, 2011; Raney et al., 2013) (Figure 17A). Around 1% of the genome encodes helicases and there are over 70 RNA helicases that have been described thus far. Each of them have different inherent functions and spatiotemporal implications in NA metabolism (Figure 17A).

Helicases are grouped into six superfamilies (SF). SF1-2 primarily consist of monomers or dimers, whereas SF3-6 operate as hexamers, forming ring-like structures. Other classification parameters are their NA preference (DNA and/or RNA), their polarity (5'-3' or 3'-5') and the coupling between force, NTP hydrolysis and translocation (active or passive) (Singleton et al., 2007). Helicases are ubiquitous molecular engines that use the energy of NTP (nucleoside triphosphate) hydrolysis to bind NA and some can translocate along NA with specific features such as step size, speed, processivity and sequence specificity. Contrary to common belief, the main role of helicases is not solely NA unwinding. In fact, numerous helicases act as NA clamps, protein recruitment platforms, mRNP remodelers, and/or single strand translocators (Linder & Fuller-Pace, 2015; Valentini & Linder, 2021) (Figure 17B). Given their importance in gene expression, their activity is usually tightly regulated, whether they act on specific NA sequences and complexes or they have no sequence specificity.



Figure 17: Cellular roles of eukaryotic RNA helicases

(A) Eukaryotic RNA metabolism processes are represented by white circles. Gray lines indicate the links between these different processes. Colored circles indicate the number of helicases involved in each process. The RNA helicases shown in this diagram are grouped by family and color code (Figure extracted from Jankowsky, 2011). **(B)** Multiple molecular actions of RNA helicases (Figure extracted from Valentini & Linder, 2021).

b) Upf1: a Super Family 1 helicase

Upf1 is the archetype of the Upf1-like family within the SF1 helicases as classified by Fairman-Williams and collaborators (Fairman-Williams et al., 2010). SF1 helicases share the same structural characteristics in their helicase cores: two RecA-like domains (Rec1A and Rec1B) forming a NA-binding domain and an NTP binding pocket (Figure 18). Upf1-like helicases act mainly as monomers and are characterized by the capacity to unwind both DNA and RNA duplexes in a 5'-3' directional manner using the hydrolysis of ATP. Upf1 has an α -helical stalk region and two protruding domains (1B and 1C) within the helicase core (Cheng et al., 2007; Chakrabarti et al., 2011) (Figure 18B, 18C). Its helicase core is flanked at the N-ter by the CH domain which contains three Zn-finger-like structures in which conserved cysteine and histidine residues coordinate zinc ions (Altamura et al., 1992; Applequist et al., 1997) (Figure 18A, 18C). In yeast, this domain rests above the RNA-bound helicase core and directly interacts with it through a hydrophobic pocket (Chakrabarti et al., 2011) (Figure 18C). At the C-ter, Upf1 bears an unstructured and less conserved domain. In humans, it is named SQ on account of its serine/glutamine and serine/glutamine/proline repeats.



Figure 18: Yeast and human Upf1 structure

(A) Homo sapiens (H.s.) and Saccharomyces cerevisiae (S.c.) Upf1 domain boundaries. The SF1 helicase central domain (HD) is composed of two RecA domains (yellow), and two protrusions 1B (orange) within the stalk region (gray) and 1C (red) within the RecA1 domain. The N-terminal cysteine-histidine rich domain (green) and the C-terminal domain (SQ in humans) encompass the HD domain. (B) Crystal structure of the H.s. HD domain of Upf1 bound to RNA (black) and ADP:AIF⁴⁻ (gray) (PDB: 2XZO from Chakrabarti et al., 2014). (C) Crystal structure of the S.c. CH-HD Upf1 bound to RNA (black) and ADP:AIF⁴⁻ (gray) (PDB: 2XZL from Chakrabarti et al., 2014).

Like many of the SF1 helicases, Upf1 is thought to be a monomer since *in vitro* purification of yeast and human Upf1 showed that both sediment as monomers in solution (Czaplinski et al., 1995; Bhattacharya et al., 2000). Yet, two-hybrid assays showed intermolecular interactions between the CH and HD domains in yeast Upf1 (He et al., 2013), suggesting the possible existence of dimers (see Discussion).

c) Upf1 is conserved among eukaryotes

Yeast and human *upf1* genes encode for 971-aa and 1118-aa long proteins respectively. They share 51% overall identity, 74% similarity of their helicase domains (aa 230-851 in yeast and aa 295-914 in humans) and 90% similarity of their CH domains (aa 54-230 in yeast and aa 115-295 in humans) (Applequist et al., 1997). Sequence identities between human, plant, fruit fly, nematode and yeast vary between 40-62% but among zebrafish, mouse and human they are over 90% (Culbertson & Leeds, 2003) (Figure 19). Both species have an acidic stretch upstream the CH domain, and human Upf1 presents a unique glycine-rich motif (PGGXG) in the CH domain (Applequist et al., 1997). Given the high homology of these proteins, it is assumed their regulations and activities are conserved.



Figure 19: Phylogenetic analysis of Upf1 (Figure extracted from Culberson & Leeds, 2003)

Peptide alignments of Upf1 proteins were converted to DNA sequences with codon boundaries maintained. Trees were generated by the maximum likelihood method and rooted by designating the two plant phyla in each alignment as the outgroup. The length of the branches corresponds to the substitution rate per site.

3) Activity: substrate whisperer or recruitment platform?

a) The catalytic activities of Upf1

As an SF1 helicase, Upf1 can directly bind DNA and RNA, and directionally (5'-3') translocate through NAs using the energy of ATP hydrolysis (Czaplinski et al., 1995; Bhattacharya et al., 2000). Its ATPase activity is NA-binding dependent: it cannot hydrolyze ATP without binding NA, however, ATP hydrolysis decreases Upf1 RNA binding affinity (Czaplinski et al., 1995; Bhattacharya et al., 2000; Chamieh et al., 2008). On the contrary, it can bind NA without the need of ATP binding nor hydrolysis (Czaplinski et al., 1995; Bhattacharya et al., 2000; Chamieh et al., 2008). On the contrary, it can bind NA without the need of ATP binding nor hydrolysis (Czaplinski et al., 1995; Bhattacharya et al., 2000). One key aspect of helicase characterization is processivity: the capacity to translocate along NA without releasing the substrate. Our team has long been interested in the characterization of yeast and human Upf1 activity using a combination of single molecule and biochemical approaches. Using magnetic tweezers, they showed how the human Upf1 helicase core translocated slowly (\approx 1 bp/s) but very processively along RNA and DNA (Fiorini et al., 2015) (Figure 20A). Moreover, they showed that the CH and



Figure 20: Yeast and human Upf1 helicase core activity

(A) Magnetic tweezers unwinding assay of human Upf1 helicase core (*H.s.* Upf1 HD) using a 180 bp RNA hairpin shows slow constant unwinding and rezipping speeds of around 1 bp/s (Figure extracted from Fiorini et al., 2015). (B) Bulk unwinding assays on *H.s.* Upf1 CH-HD and HD-SQ truncations showing inhibition of the unwinding activity (Figures extracted from Fiorini et al., 2013 and Chakrabarti et al., 2011). (C) Magnetic tweezers unwinding assay of yeast Upf1 helicase core (S.c. Upf1 HD) using a 1.2 kb DNA hairpin shows rapid constant unwinding and rezipping speeds of around 10 bp/s (Figure extracted from Kanaan et al., 2018). (D) Magnetic tweezers single molecule binding assay using S.c. Upf1 HD and cycling forces without ATP show a steady grip of Upf1 onto the DNA (Figure extracted from Kanaan et al., 2018).

the SQ domains inhibit helicase and ATPase activities (Chamieh et al., 2008; Chakrabarti et al., 2011; Fiorini et al., 2013) (Figure 20B). The yeast Upf1 helicase core, on the other hand, presented 10-fold higher translocation rates on DNA than its human homolog (~10 bp/s) (Kanaan et al., 2018) (Figure 20C). Upf1 is remarkably processive on DNA, as it is able to translocate over 10 kb without releasing its substrate (Kanaan et al., 2018) and to maintain a strong grip on DNA when bound in absence of ATP, even under external force constraints (Figure 20D) (Kanaan et al., 2018). This property was attributed to the presence of the 1B and 1C protrusions within the helicase core, since mutating them had a high impact on processivity and RNA grip (Kanaan et al., 2018). It is noteworthy that these measures were

done with the isolated helicase cores, thus not considering intramolecular effects, and on DNA, not the *bona fide* Upf1 substrate for NMD.

The human Upf1 isoforms have different inherent catalytic activities: isoform 2 has a stiff loop which makes it less active than isoform 1 (Gowravaram et al., 2018). Nevertheless, exogenous expression of Upf1 in Hela Δ Upf1 cells showed equal rescue between the two isoforms (Nicholson et al., 2014), suggesting these differences on catalytic activity have no effect on NMD.

b) Upf1 activity is essential for NMD

The presence of Upf1 is critical for NMD, but its activity proves equally important. Indeed, mutations affecting ATPase activity, NA-binding, and processivity show NMD defects *in vivo* (Weng et al., 1996b; Kashima et al., 2006; Franks et al., 2010; Kurosaki et al., 2014; Kanaan et al., 2018; Lee et al., 2015; Chapman et al., 2022). ATPase activity is also necessary for efficient translation termination at premature stop codons and ribosome recycling (Franks et al., 2010; Serdar et al., 2016). Yet, the specific way in which Upf1 is involved in NMD and the link between ATPase activity levels, processivity, RNA binding and NMD efficiency are still unclear. Indeed, it is still debated whether Upf1 activity in NMD is crucial for unwinding double-stranded RNA structures, scanning the single-stranded RNA, for mRNP remodeling, or for complex and/or ribosome recycling/disassembly (Franks et al., 2010; Hogg & Goff, 2010; Fiorini et al., 2015).

4) Interactions and regulation: taming the master orchestrator

a) Upf1 is a hub for protein-protein interactions

One of the key characteristics of Upf1 is its extensive network of partners. Studying the Upf1 interactome has been instrumental in understanding NMD dynamics. From two-hybrid assays to cutting-edge mass spectrometry techniques, the identification of a diverse repertoire of Upf1 partners, including several RBPs, has provided valuable insights into the complex network of molecular interactions involving Upf1. However, for most of these interactions, it remains to be demonstrated if they are direct or not, especially since Upf1 is an RBP itself, many of the known interactions might be mediated by RNA.

Apart from its direct and/or indirect interactions with NMD-associated proteins such as Upf2/3 (Leeds et al., 1991, 1992; Pulak & Anderson, 1993; Cui et al., 1995; Perlick et al., 1996; Chamieh et al., 2008, Dehecq et al., 2018), Smg (Ohnishi et al., 2003; Kashima et al., 2006; Yamashita et al., 2001), and EJC (Chamieh et al., 2008; Buchwald et al., 2010; Flury et al., 2014; Schweingruber et al., 2016), Upf1 also interacts with a broad spectrum of general RNA decay factors (Figure 21).





Upf1 molecular interactions. Human interactions are indicated with black lines, yeast interactions with red lines and interactions in both organisms in green. Upf1-mediated decay pathways (UMD) are grouped by the dotted circle.

Among the decay-related proteins mentioned in chapter 2 that interact with Upf1 we find Stau1, Stau2 (Kim et al., 2005; Gong et al., 2009), Regnase 1 (Mino et al., 2015), Ago1/2 (Jin et al., 2009), Slbp (Choe et al., 2014).

In addition, Upf1 also interacts with translation factors such as eIF1/3 (Czaplinski et al., 1998; Isken et al., 2008), Rrp26 (Min et al., 2013), and eIF3/4G (Isken et al., 2008; Flury et al., 2014; Schweingruber et al., 2016).

It is noteworthy that a substantial proportion of the protein-protein interactions involving Upf1 are targeted at the CH domain. This domain appears to be a central hub for numerous entities such as Stau1, Upf2, Rrp26, eRF3, Dcp2, and even the CH domain itself, to engage in direct interactions with Upf1. Such a diverse array of molecular interactions at the CH domain highlights its crucial role in facilitating various cellular processes. Interestingly, Upf2 and Stau1, which both interact with the CH domain, are not mutually exclusive but rather cooperative (Gowravaram et al., 2019).

b) Regulation by phosphorylation

In metazoans, the CH and Ct regions are highly conserved and contain multiple Smg1 kinase phosphorylation sites that regulate Upf1 activity and NMD activation (Durand et al., 2016). It has been proposed that phosphorylation is needed for Upf1 to stay bound to the mRNA (Kurosaki et al., 2014), to release the eRFs (Kashima et al., 2006), to repress translation initiation (Isken et al., 2008) and remodel mRNPs (Kashima et al., 2006; Ivanov et al., 2008). This is supported by the fact that deletions of Smg5, Smg6 and Smg7 increase the relative quantity of phospho-Upf1 (Page et al., 1999; Okada-Katsuhata et al., 2012). These peptides are absent in *S. cerevisiae* and there is no described homolog for

Smg1 in this organism, although phosphorylations of some Upf1 tyrosine residues involved in the regulation of its helicase activity have been observed (Lasalde et al., 2014). The recent discovery of the association of the kinase Hrr25 to Upf1 makes it a good candidate for Upf1 phosphorylation (see Scope of the thesis, Dehecq et al., 2018). Phospho-Upf1 is able to recruit Smg5/Smg7 and Smg6 in mammals (Ohnishi et al., 2003; Okada-Katsuhata et al., 2012; Chakrabarti et al., 2014). As mentioned in chapter 2, the Smg5/Smg7 dimer recruits the deadenylation machinery (Loh et al., 2013) and the Smg6 endonuclease triggers endonucleolytic decay of the target RNA (Gatfield & Izaurralde, 2004; Glavan et al., 2006; Huntzinger et al., 2008; Eberle et al., 2009).

c) Regulations by partners

Among the large array of interactors, very few direct partners of Upf1 have been described. Interestingly, the regulation of Upf1 by Upf2 is the only well characterized regulation (Figure 22). Upf1 directly interacts with Upf2 through its CH domain, inducing a drastic conformational change of Upf1 (Chamieh et al., 2008; Clerici et al., 2009; Chakrabarti et al., 2011). The bipartite interaction of Upf2 does not directly activate Upf1 but rather lifts the autoinhibition of Upf1 by its CH domain (Chakrabarti et al., 2011; Fiorini et al., 2015). Upf2 also undergoes conformational changes as it clamps the CH domain of Upf1 through a beta-sheet and an alpha helix on opposite surfaces (Clerici et al., 2009). In addition, Upf2 regulates Upf1's phosphorylation as it directly interacts with Smg1 (Clerici et al., 2014; Melero et al., 2014).

The few other regulators of Upf1 are the HLV-1 Tax protein which inhibits RNA binding and translocation *in vitro* (Fiorini et al., 2018), Stau1 which inhibits ATPase activity *in vitro* (Gowravaram et al., 2019) (Figure 22), Stau2 which promotes helicase activity (Park et al., 2013) and eRF3 and eRF1 which inhibit ATPase activity (Czaplinski et al., 1998).



Figure 22: Human Upf1 modulation by Upf2 and Stau1 (Figure extracted from Gowravaran et al., 2019)

Recombinant human Upf1 CH-HD RNA-dependent ATPase activity was measured using an enzyme-coupled phosphate detection assay in the presence of either UPF2 or Stau1. The data points, accompanied by error bars, represent the average values and standard deviation derived from three separate experiments.

5) Beyond NMD: Upfl's expanding repertoire of cellular functions

Apart from NMD, Upf1 also is also involved in many other RNA degradation mechanisms we described in chapter 1 such as Staufen-mediated mRNA Decay (SMD), Replication-dependent Histone mRNA Decay (HMD), Glucocorticoid Receptor-mediated mRNA Decay (GMD), Regnase 1-mediated mRNA Decay (RMD) and TSN-mediated microRNA decay (TumiD) (reviewed in Kim & Maquat, 2019; Lavysh & Neu-Yilik, 2020). Moreover, Upf1 has been recently shown to contribute to the degradation of m6a methylated mRNAs by interacting with the m6A reader YTHDF2 and recruiting the decapping machinery (Boo et al., 2022; Gibbs & Chanfreau, 2022). In addition to RNA degradation, Upf1 also contributes to protein quality control of the synthesized protein (reviewed in Hwang et al., 2021).

Given the numerous UMD pathways that are independent from NMD, it is plausible that many of the so-called NMD-targets might actually be targets of other UMDs, particularly those with no clear NMD-triggering features. The multifaceted roles, interactions, and regulations of Upf1 contribute to the structural and functional complexity of this protein.

Other less studied roles of Upf1 include its implication in cell cycle progression and DNA damage response (Azzalin & Lingner, 2006; Addinall et al., 2011), promotion of R-loop formation (Ngo et al., 2021), its debated E3 ubiquitin ligase activity (Takahashi et al., 2008; Feng et al., 2017) and its involvement in telomere maintenance (Chawla et al., 2011) (reviewed in Imamachi et al., 2012).

Gaps in the knowledge

General questions about NMD

One of the main unanswered questions in the field is how a PTC is differentiated from a *bona fide* stop codon. Both signal the end of protein translation and are identical in nature, so what makes them different? The current models propose that this happens either through 3' UTR length, ribosome termination efficiency or Upf1 recruitment through the EJC, yet these conjectures are still ambiguous.

Moreover, the exact sequence of events that occur during the pathway are still misunderstood. The precise chronological steps of NMD events, including how Upf1 is recruited and how it triggers decay, remains elusive. While it is commonly assumed that the first step of NMD is the recruitment of Upf1/Upf2/Upf3, there is no concrete evidence to support this idea. The existence of EJC, Upf2 and Upf3 independent pathways suggests alternative recruitments of Upf1, both in yeast and humans. How are the different NMD factors coordinated in space and time to ensure accurate and efficient recognition and degradation of aberrant transcripts?

Furthermore, the binding rules that govern the recruitment of Upf1, are still unclear. It is uncertain whether Upf1 acts as a platform for protein recruitment, as a scanner for target recognition, or as an mRNP remodeler. Upf1 recruitment to the RNA was initially proposed to happen through the terminating ribosome bound to release factors eRF1/3 (Kashima et al., 2006). These results were challenged by the findings of Upf1 binds preferentially the 3' UTRs of NMD targets in a translation-independent and length-dependent manner (Hogg & Goff, 2010). Additionally, RIP-seq (RNA immunoprecipitation sequencing) and CLIP-seq (cross-linking immunoprecipitation sequencing) experiments showed that the 3' UTRs of most RNAs were bound by Upf1 and that when translation was inhibited Upf1 was accumulated in the coding sequence (Hurt et al., 2013; Zünd et al., 2013). This suggests its 3' UTR binding is only due to the displacement by the ribosome, not to a preference of 3' UTR binding. They defined Upf1 as a 'promiscuous' helicase which recognizes virtually every transcript in the cell and undergoes frequent rounds of RNA binding/release. Other studies described that Upf1 only stays on its targets when phosphorylated as a marker of NMD induction (Kurosaki et al., 2014). Altogether, the factors that determine the timing, position, and function of Upf1 binding on RNA remain poorly characterized. The only known fact is that its activity is essential for NMD, but the implications of an "activated" or "inhibited" Upf1 are not clear. Does the activity have to be high or be maintained at a constant level to trigger NMD? Moreover, what is the link between ATP hydrolysis, RNA binding, RNA translocation and decay?

Furthermore, despite a large number of interactions being described for Upf1, very few studies focused on the specific nature of these interactions nor on the effect they have on Upf1's activity, especially in yeast. For instance, its interaction with the decapping machinery had been proposed in the 90s, however no study has shown whether the interaction is direct. Additionally, the role of phosphorylation in NMD remains a subject of

debate, as in metazoans it seems to play a crucial role in regulating the activity and interactions of Upf1, yet its importance is not well characterized in yeast.

How is RNA decay triggered and modulated during NMD?

It was noted that yeast cells predominantly employ the decapping mechanism and 5' to 3' decay for NMD, while metazoans primarily rely on endonucleolytic cleavage and the deadenylation pathway for the same process. However, the underlying reasons for this mechanistic divergence between the two evolutionary lineages remain unclear. One possible explanation is the distinct regulatory requirements of NMD in yeast and metazoans, which may have resulted from their divergent biological processes and cellular environments. For instance, yeast cells typically exhibit faster growth rates and rely heavily on post-transcriptional gene regulation to adapt to changing conditions, which could necessitate the rapid degradation of aberrant mRNAs through decapping. In addition, 5'-3' degradation prevents the translation of truncated proteins. In contrast, metazoans have evolved complex developmental and physiological processes, and NMD may be required to function over longer timescales, necessitating a more gradual and controlled deadenylation-based mechanism. Would this also explain the difference in speed of the Upf1 proteins in humans and in yeast?

Moreover, Smg6 is a potent endonuclease with no sequence specificity, meaning that if it is not tightly regulated, it can wreak havoc in the cell. However, it is unclear whether Smg6 is activated only when bound to Upf1 or if other factors are involved. The regulation of NMD may be dependent on the concentration of NMD factors in proximity of the target transcript or the accessibility to the RNA. Furthermore, the regulation of decapping and deadenylation in NMD has not been explored. Decapping is an irreversible step that commits an RNA to its degradation, thus it is crucial to ensure that only aberrant or undesired transcripts are degraded. Again, it is unknown if decapping and deadenylation are regulated by NMD factors.

Is there a universal basic model which would explain NMD in all organisms?

In sum, these uncertainties underscore the necessity for further research in this field, indicating that substantial progress is still needed in order to fully understand the mechanistic details of NMD.

Scope of the thesis

NMD is unquestionably an essential mechanism for quality control and transcriptome regulation, which is illustrated by its broad implication in cellular metabolism and disease (see Chapter 4). Nevertheless, despite more than four decades of research, our understanding of NMD remains incomplete, with many aspects of this complex process still the subject of intense debate and misunderstanding.

Novel Upf1-bound complex in S. cerevisiae

Our collaborators at Institut Pasteur in the team led by Cosmin Saveanu, employed an advanced experimental approach to study the *in vivo* composition and dynamics of NMD complexes in *S. cerevisiae* (Dehecq et al., 2018).



Figure 23: Yeast Upf1 is involved in two mutually exclusive complexes (Modified from Dehecq et al., 2018) Enrichment values for proteins identified in tagged Upf1 **(A)**, Upf2 **(B)**, Upf3 **(C)**, Nmd4 **(D)**, Ebs1 **(E)** and Dcp1 **(F)** purifications. Tagged proteins are indicated in bold for each experiment. Black bars correspond to average enrichment values obtained in purifications done without RNase and orange bars with an RNase A and RNase T1 treatment. Error bars represent SD. Dots represent individual enrichment values for each protein in the six replicates for Upf1-TAP, 3 for Upf1-TAP and RNase, 6 for Upf2-TAP, 3 for Upf2-TAP and RNase, 5 for Upf3-TAP, 3 for Upf3-TAP and RNase, 4 for Nmd4-TAP, 3 for Nmd4-TAP and RNase, 3 for Ebs1-TAP, with and without RNase and 3 for Dcp1-TAP with or without RNase. Only proteins enriched by a factor of 16 or more in one of the purifications presented here are shown. For clarity, groups of related proteins were combined (the group of Lsm1 to Lsm7 is marked as "Lsm1-7" and the values correspond to the mean of all the values found in purifications; CK2 group corresponds to Cka1, Cka2, Ckb1 and Ckb2). **(G)** Visual representation of the Detector (Upf1-2/3) and Effector (Upf1-decapping) distinct complexes. **(H)** Schematic representation of the domain structure of Nmd4 and Ebs1 from Saccharomyces cerevisiae compared with human Smg6, Smg5 and Smg7. Identity percentages among the different domains of Smg proteins, Nmd4 and Ebs1, are indicated. Values represent percent of identical residues as a fraction of all the aligned residues.

They combined affinity purification of genomically tagged NMD proteins with quantitative mass spectrometry, completing a remarkable total of 112 experiments. An enhanced data analysis protocol enabled them to differentiate between enriched proteins and background noise. Upf1-TAP co-immunoprecipitation (co-IP) showed a high enrichment of the known partners Upf2 and Upf3 that was RNA-dependent and the decapping heterodimer Dcp1/Dcp2 and its cofactor Edc3, but also novel partners such as Nmd4, Ebs1 and Hrr25 in a RNA-independent manner (Figure 23A). To further characterize these interactions, they immunoprecipitated all enriched partners, with or without RNAse treatment (Figure 23B-F). Co-IPs of Upf2-TAP and Upf3-TAP were only enriched on the Upf proteins (Figure 23B-23C). Conversely, co-IP of Nmd4-TAP, Ebs1-TAP or Dcp1-TAP showed high enrichment of all other Upf1 partners except for the Upf2-3 proteins (Figure 23D-F). Thus, they identified two distinct, mutually exclusive complexes associated with Upf1: the core Upf complex (Upf1-2/3) and the Upf1-decapping complex composed of Upf1, Dcp1/Dcp2, Nmd4, Ebs1 and Hrr25. These complexes were named Detector (Upf1-2/3) and Effector (Upf1-decapping), each exhibiting unique characteristics and functionalities (Figure 23G).

The striking sequence resemblance between Nmd4/Ebs1 and the mammalian Smg5-7 proteins, led them to propose a unified model for NMD across species. In *S. cerevisiae*, there are two potential Smg7 homologs: Est1 (ever shorter telomeres protein 1) and Ebs1 (Est1-like Bcy1 suppressor 1). They both present a 14-3-3 domain characteristic of Smg5/6/7 proteins but lack the PIN domain of Smg5/6 proteins (Figure 23H).

Est1 is involved in telomere regulation but not in NMD (DeZwaan & Freeman, 2009), unlike Ebs1 which was previously shown to be involved in NMD as a putative Smg7 ortholog (Ford et al., 2006; Luke et al., 2007). Interestingly, Smg5/6/7 are also implicated in telomere maintenance (Azzalin et al., 2007).

Likewise, the Nmd4 protein had been discovered long ago by two-hybrid assays as being involved in NMD (He & Jacobson, 1995) (which explains the name) but no further studies were conducted. Among other interesting proteins within the Upf1-decapping complex, they found Hrr25, a serine-threonine kinase which could potentially be the functional homolog of the metazoan Smg1.

Towards a universal NMD model: "Detector & Effector"

According to this model, NMD functions through successive formation of Upf1-bound Detector and Effector complexes. The implications of this model are significant, as it suggests a conserved mechanism for NMD across different eukaryotic organisms. The Upf1-2/3, the 'Detector complex', would sense and bind NMD targets (Figure 24). Following a conformational switch, potentially due to Upf1 phosphorylation by Hrr25, Upf2/Upf3 would be released and Nmd4, Ebs1, Dcp1/Dcp2 and Edc3 would be recruited to form the 'Effector complex' (Figure 24). Nmd4 and Ebs1 would have molecular roles very close to those of Smg6 and Smg5/Smg7 respectively and Dcp1/Dcp2 would allow the initiation of degradation *via* decapping (Dehecq et al., 2018). This model resembles the human SURF/DECID model (Kashima et al., 2006) and suggests an evolutionary conservation of NMD mechanisms.



Figure 24: Detector and Effector model for yeast NMD (Modified from Dehecq et al., 2018) The Upf1-2/3 complex (Detector complex) recognizes and binds to NMD targets. Upon a conformational switch, potentially triggered by Upf1 phosphorylation by Hrr25 kinase, Upf2/3 is released. Subsequently, the Upf1- decapping (Effector complex) is formed, involving the recruitment of Nmd4, Ebs1, Dcp1/2, and Edc3.

PhD project

The research project that I undertook during my PhD was built upon the foundation laid by our collaborators at Institut Pasteur, who made the groundbreaking discovery of the Upf1-decapping complex in *S. cerevisiae* (Dehecq et al., 2018). Their research served as the launching point for our detailed investigation into the internal interactions and the regulation of the active proteins within this complex. Moreover, given the plausible homology and the proposed model for a conserved NMD, it is of clear interest to study it in a less complex organism such as *S. cerevisiae*.

To achieve our research goals, we opted to take an *in vitro* approach, to dig deeper into the dynamics and direct interactions within the complexes. Using purified recombinant proteins we aimed to reconstitute the Upf1-decapping complex and to get access to more mechanistic details difficult to measure *in vivo*. Our primary objectives throughout the

course of the project were threefold. First, we aimed to gain a comprehensive understanding of the interactions between the various components of the complex. Second, we sought to investigate the regulation of the main orchestrator: the Upf1 helicase. Finally, we aimed to shed new light on the functional mechanisms of the Upf1-decapping complex, building on the work of our collaborators to further our understanding of this important area of research. This approach would allow us to access more precisely to the dynamics of the Upf1-decapping complex.

(1) In vitro dissection of yeast Upf1 interactions and regulation

As part of my research project, I aimed to gain a comprehensive understanding of the protein-protein interactions that are involved in the formation of the Upf1-2/3 and Upf1-decapping complexes. Our team has long specialized in the *in vitro* reconstitution of RBP complexes such as the EJC (Barbosa et al., 2012; Fiorini et al., 2012; Chazal et al., 2013; Wang et al., 2018; Busetto et al., 2020). To study the direct interactions between Upf1 and its partners, I undertook a series of *in vitro* interaction assays, utilizing over twenty different recombinant proteins that were purified for this purpose. The purification of truncated domains of Upf1 allowed me to to identify the specific domains that are involved in these interactions.

(2) Characterization of yeast Upf1 activity using magnetic tweezers

In collaboration with Vincent Croquette's team at the physics department of the Ecole Normale Supérieure, our team has focused on the single molecule characterization of RNA helicases for a while (Fiorini et al., 2015; Hodeib et al., 2017; Fiorini et al., 2018; Gowravaram et al., 2018; Kanaan et al., 2018; Rieu et al., 2021; Ruiz-Gutierrez et al., 2022). I aimed to characterize the activity of yeast Upf1 in RNA and DNA and the influence of the N- and C-terminal domains using magnetic tweezers.

(3) RNA decay activation: decapping and endonucleolytic activities

Lastly, I aimed to gain a deeper understanding of the mechanisms that underlie the final stages of NMD, which are responsible for triggering mRNA decay. We designed an innovative *in vitro* quantitative assay, which allowed the assessment of the decapping activity of recombinant Dcp1 and Dcp2 proteins in detail and their regulation within the Upf1-decapping complex.

A collaborative work

This ambitious project was carried out in close collaboration with the teams of Cosmin Saveanu and Gwenaël Badis-Breard at Institut Pasteur who carried out *in vivo* experiments complementary to our *in vitro* approach and the team of Marc Graille at Ecole Polytechnique, who solved the structure of the yeast Nmd4/Upf1-HD complex by crystallography (Barbarin-Bocahu et al., in preparation). More recently, we established a collaboration with the team led by Elena Conti at the Max Planck Institute of Munich to attempt resolving the structure of the reconstituted Upf1-decapping complex using Cryo-EM.

Results

Part 1: Incompatibility of Upf1 and Upf2 interaction with Upf1-decapping complex in yeast provides new insights into NMD dynamics

The following part corresponds to the draft of a manuscript which is in progress.

Introduction

Nonsense-mediated mRNA decay (NMD) is a conserved translation-dependent mRNA degradation pathway that plays a critical role in mRNA quality control and gene expression. It recognizes and degrades aberrant mRNAs that contain premature termination codons (PTCs) as well as other physiological mRNAs (Malabat et al., 2015; Johanson et al., 2007; Brogna et al., 2016; Celik et al., 2017). Thereby, it also participates in whole transcriptome regulation and is implicated in a wide range of cellular processes, including differentiation (Xu et al., 2016), stress response (Karam et al., 2015), and immune surveillance (Jung et al., 2006, Li & Wilkinson, 1998; Jiang et al., 2017). Thus, dysregulation of NMD is linked to several pathologies, including cancer (Lindemboom et al., 2016), neurodegeneration (Jaffrey & Wilkinsen 2018), and various genetic disorders (Nassif & Muhlemann, 2018). At the heart of this mechanism is the conserved RNA helicase Upf1 and its core partners Upf2 and Upf3 (Leeds et al., 1992; Pulak & Anderson., 1993; Cui et al., 1995; Perlick et al., 1996). The trimeric Upf complex recruits other RNA decay factors to the target RNA that vary between species.

In yeast, NMD targets are mostly degraded through decapping by the Dcp1/Dcp2 heterodimer (Muhlrad & Parker 1994) whereas in humans, most are cleaved by the endonuclease Smg6 (Huntzinger et al., 2008; Eberle et al., 2009) but some can be degraded by decapping (Lykke-Andersen, 2002) and deadenylation (Loh et al., 2013). Broadly, NMD substrates can be categorized into two types based on their molecular features: those with long 3' untranslated regions (UTRs) and those with an exon-junction complex (EJC) located at least 50 nucleotides downstream of the termination codon. While all organisms exhibit NMD of mRNAs with long 3' UTRs, EJC-dependent NMD isn't universally conserved, especially in organisms devoid of EJC proteins, such as *S. cerevisiae*.

From these two classes of NMD targets arise two widely accepted models for NMD activation. The SURF/DECID model (Kashima et al., 2006) describes Upf1 recruitment by a Upf2/Upf3-bound EJC downstream of a premature termination codon. After Upf1 phosphorylation by Smg1/8/9, decay-inducing partners Smg5/Smg7 and Smg6 are recruited and trigger RNA decay. In the "Faux 3' UTR" model (Amrani et al., 2004) NMD is triggered by a longer 3' UTR (consequence of a PTC) which increases the distance of the ribosome with the poly(A) binding proteins and thus decreases efficiency in translation termination. The ribosome would then recruit Upf1/2/3 to trigger decay. Although both of these models recognize the critical functions of the Upf proteins, the recruitment and the role of these proteins are still widely unknown. Moreover, the SURF/DECID model assigns crucial roles to several non-conserved proteins, particularly the Smg proteins and the EJC core, despite the highly conserved nature of NMD.
Upf1 is a highly conserved RNA helicase present in all eukaryotes (Applequist et al., 1997; Serin et al., 2001), and is involved in a variety of RNA decay pathways and other cellular processes (reviewed in Kim and Maquat., 2019), providing an early indication of the large range of protein co-factors that interact with Upf1. It has two conserved structured domains: the N-terminal cysteine-histidine-rich (CH) domain and the central helicase domain (HD). The C-terminal (Ct) domain is unstructured and less conserved (Figure 1A).

Upf1 can directly bind nucleic acids (NAs) and directionally (5' - 3') translocate along DNA and RNA using the energy from ATP hydrolysis (Bhattacharya et al., 2000; Chakrabarti et al., 2011; Kanaan et al., 2018). Thus, it is capable of unwinding double-stranded NA, translocate along single-stranded NA, and remodel mRNPs (Fiorini et al., 2015). Its helicase activity, ATPase activity, and processivity are essential for efficient degradation of NMD targets (Weng et al., 1996; Kanaan et al., 2018). Nonetheless, its specific way of action in NMD and the link between ATPase activity levels, processivity, RNA binding and NMD efficiency is still unclear.

Research on Upf1's regulation during NMD is currently remarkably limited and mostly restricted to humans. Only the intramolecular inhibition (Chakrabarti et al., 2011; Fiorini et al., 2015), and regulation by Upf2 have been found to directly modulate human Upf1 helicase and ATPase activities (Weng et al., 1998; Fiorini et al., 2015; Chakrabarti et al., 2023; Xue et al., 2023). Although many different partners of Upf1 have been described (Flury et al., 2014), little is known about their mode of interaction, whether interactions are direct or not and whether these partners modulate Upf1's activities.

Quantitative mass spectrometry analysis of NMD complexes revealed the existence of two mutually exclusive complexes in yeast: the Upf1-2/3 core complex and the Upf1-decapping complex. The latter includes the decapping heterodimer Dcp1/Dcp2, Nmd4 and Ebs1 (Dehecq et al., 2018). The clear mutual exclusivity of the complexes suggests the existence of a switch between partners and a possible differential regulation of Upf1 during NMD within the different complexes. In addition, the existence of a distinct Upf1-decapping complex suggests a direct recruitment of decay-inducing partners and paves the way towards a direct modulation of RNA degradation by Upf1. Here, we aim to better characterize these late steps of yeast NMD *in vitro* by dissecting the interactions within the Upf-bound complexes and studying the regulation of Upf1's activity by its partners.

Results

Dcp2 directly interacts with Upf1 N-terminal CH domain

Saccharomyces cerevisiae (S.c.) Upf1 is a 971 amino acid (aa) protein composed of three domains: an N-terminal domain rich in cysteine and histidine (CH, aa 54-220), followed by the SF1 helicase domain (HD, aa 221-851), and a less conserved and poorly structured C-terminal region (Ct, aa 852-971). To study the interaction between Upf1 and the components of the decapping complex, we purified several versions of Upf1 recombinant proteins from *E. coli*, including the full-length (FL) protein Upf1 (Upf1-FL) and truncated versions (Figure 1A). Each version of Upf1 was fused to an N-terminal calmodulin-binding peptide (CBP). S.c. Dcp2 (970 aa) bears a decapping Nudix domain flanked by an

Results Part 1

N-terminal regulatory domain (NRD) required for Dcp1 recruitment (She et al., 2006) and a short domain that mediates the interaction with its activator Edc3 (Harigaya et al., 2010; Charenton et al., 2016) (Figure 1B). Recently, yeast two-hybrid experiments delineated two domains within the C-terminal domain important for Dcp2 association with Upf1 (He et al., 2022), named Upf1 binding domains (UBD) 1 and 2 (Figure 1B). Multiple sequence alignments revealed clearly conserved residues between the two sequences (Figure 1B). We purified a short glutathione-S-transferase (GST)-tagged version of Dcp2 encompassing the UBD domains (Dcp2-UBD, aa 434-720) and a longer His-tagged version (Dcp2-L, aa 1-720) (Figure 1B) co-purified with full-length GST-Dcp1. For binding assays, we mixed different combinations of recombinant proteins, and interactions were detected by copurification on calmodulin resin. After extensive washing, the eluted protein(s) were fractionated by SDS-PAGE and directly visualized by Coomassie staining (see Methods).



Figure 1: DCP2 directly interacts with UPF1 CH domain through two UPF1-Binding domains

(A) Schematic representation of yeast Upf1 constructs used for pulldown assays. All constructs bear an N-terminal CBP tag and a C-terminal 6-His tag. (B) Schematic representation of yeast Dcp2 constructs used for pulldown assays and sequence alignments Upf1-binding domains (UBDs). Dcp2-UBD corresponds to a truncation containing UBD1 and UBD2 (434-720) and bear a GST tag in N-terminus and 6-His in C-terminus. Dcp1 / Dcp2(1-720) was co-purified using an N-terminal GST tag on Dcp1-FL and a C-terminal 6-His tag on Dcp2(1-720). (C, D) CBP-Upf1-pulldown assays using Upf1 truncations against Dcp2-UBD (C), Dcp1-FL/Dcp2(1-720) (D). Proteins (input) were mixed together and incubated before Upf1 pulldown using calmodulin beads. After washing (250mM NaCl) proteins were eluted, loaded in an SDS-PAGE gel and revealed using Coomassie blue. Possible CH-dimers are indicated with (*) and protein contaminants with (#). (E) AlphaFold2 models of the S.c. Upf1-CH domain bound to Dcp2 UBD1 (wheat) and UBD2 (brown). The surface of the Upf1 CH domain is colored based on amino acid physicochemical properties: hydrophobic residues in gray, polar residues in cyan, positively charged residues in blue, negatively charged residues in red and Cys in yellow. The side chain of some conserved residues from the UBD1 and UBD2 residues are shown as sticks.

When CBP-Upf1 proteins were used as bait, Dcp2-UBD was efficiently co-precipitated with all CH-bearing constructs of Upf1 (Figure 1C, lanes 2-4) but not by the constructs lacking the CH domain (Figure 1C, lanes 5-6). The two isolated UBD domains, UBD1 and UBD2 (Figure S1A), equally contributed to Upf1 binding, given that both were co-precipitated as efficiently by CH-containing CBP-Upf1 constructs (Figure S1B, S1C). Consistently, the heterodimer Dcp1/Dcp2-L was also pulled down by all the CH-bearing constructs (Figure 1D, lanes 2-4). When washing at higher NaCl concentrations, the interaction was maintained (Figure S1D). Remarkably, AlphaFold2 multimer structural prediction of the Upf1-CH domain in the presence of either Dcp2 UBD1 or UBD2 showed that the conserved hydrophobic residues of UBDs would similarly interact with the hydrophobic pocket within the CH domain (Figure 1E). Both AlphaFold2 models presented





(A) Schematic representation of the different protein domains and recombinant proteins of S.c. Upf1 and Dcp2. (B-D) *In vitro* pulldown assays using CBP-Upf1 as bait. The indicated proteins (Input) were mixed together and incubated before pulldown of the indicated proteins. After washes (250 mM NaCl) proteins were eluted, loaded in an SDS-PAGE gel and revealed using Coomassie blue staining. (E, F) pLDDT scores of UBD1 (E) and UBD2 (F) AlphaFold2 simulations. pLDDT >90 in dark blue, 70 < pLDDT < 90 in cyan, 50 < pLDDT < 70 in yellow and pLDDT < 50 in orange.

high pLDDT confidence scores (Figure S1E). Collectively, these results show that Dcp2 directly interacts with the CH domain of Upf1 and that this interaction is compatible with Dcp2 binding to Dcp1.

Upf1 makes direct contact with both Nmd4 and Ebs1

To investigate the interaction of Upf1 with Nmd4 and Ebs1, both of which are components of the Upf1-decapping complex (Dehecq et al., 2018), we employed a similar approach. Nmd4 contains an N-terminal PIN domain followed by a short disordered C-terminal arm (Dehecq et al., 2018; Barbarin-Bocahu & Graille, 2023), whereas Ebs1 bears a 14-3-3 N-terminal domain flanked by a long and disordered C-terminal region (Figure 2A). We purified full-length Nmd4 (aa 1-218) fused to an N-terminal twin-strep (TS) tag, and a soluble version of Ebs1 deleted from the last 282 aa (Ebs1, aa 1-591) fused to an N-terminal GST tag (Figure 2B).

As we previously showed, (Dehecq et al., 2018; Barbarin-Bocahu et al., *in preparation*) Upf1-HD co-precipitated Nmd4 (Figure 2C, lane 4). Interestingly, all different Upf1 versions co-precipitated Nmd4, except for the isolated Ct domain (Figure 2C). These results suggest that Nmd4 interacts with two different regions of Upf1: the CH and HD domains.



Figure 2: Nmd4 and Ebs1 directly interact with Upf1.

(A) Schematic representation of Nmd4 construct used for pulldown assays. Nmd4-FL(1-218) was fused to a Twin Strep tag in N-terminal and to a 6-His tag in C-terminal. (B) Schematic representation of Ebs1 construct used for pulldown assays. Ebs1(1-591) was fused to a GST tag in N-terminal and a 6-His tag in C-terminal. (C, D) CBP-Upf1-pulldown assays using Upf1 truncations against Nmd4 (C) or Ebs1 (D). Proteins (input) were mixed together before Upf1 pulldown using calmodulin beads. After washing (250 mM NaCl) proteins were eluted, loaded in an SDS-PAGE gel and revealed by Coomassie blue. Possible CH-dimers are indicated with (*) and protein contaminants with (#).

In the case of Ebs1, we observed that it was more efficiently co-precipitated by the Upf1 constructs containing the Ct domain (Figure 2D, lanes 2 and 5). We did not test Upf1-HD-Ct and Upf1-CH-HD which co-migrated with GST-Ebs1. A weak interaction, slightly more intense than that of the negative control, was also detected with Upf1-CH (Figure 2D, lanes 1 and 3). Therefore, if Ebs1 may have several contact points with Upf1, it mainly interacts with the Ct domain, consistent with our previous data (Dehecq et al., 2018). Together, these results show that Upf1 directly and specifically interacts with at least three components of the decapping complex *via* contact points in its three domains.

Reconstitution of Upf1-containing decapping complex

We next tested whether these individual interactions were compatible. We first used CBP-Upf1-FL for CBP-pulldowns, with either one partner, a combination of two partners, or all three partners (Figure 3A). To ensure distinctive tags for each protein and avoid protein co-migration, we purified an HisZZ N-terminal tagged Ebs1(1-591). Consistent with our previous observations, Upf1-FL co-precipitated with TS-Nmd4, GST-Dcp2-UBD, and HisZZ-Ebs1 individually (Figure 3A, lanes 2-4). Upf1 precipitated its partners as efficiently as the combination of multiple partners (Figure 3A, lanes 5-7) except in the case of Nmd4 and Ebs1, for which we observed a slight negative interference (Figure 3A, lane 7). Interestingly, Ebs1 co-precipitation seemed more efficient when all four proteins were present (Figure 3A, compare lanes 4 and 8), suggesting a potential cooperative binding. Altogether, these results suggest no interaction incompatibility within the Upf1-decapping complex. However, we cannot exclude that Upf1 co-precipitates each partner individually.





(A) CBP-Upf1-pulldown assays using Upf1-FL against single, double or triple combination of Nmd4-FL, Ebs1(1-591) and Dcp2-UBD. (B) GST-Dcp2-UBD-pulldown assays using Dcp2-UBD as bait against single, double or triple combination of Nmd4-FL, Ebs1(1-591) and Upf1-FL. (C) GST-Dcp1-FL/Dcp2-L-pulldown assays using Dcp2-UBD as bait against single, double or triple combination of Nmd4-FL, Ebs1(1-591) and Upf1-FL. (C) GST-Dcp1-FL/Dcp2-L-pulldown assays using Dcp2-UBD as bait against single, double or triple combination of Nmd4-FL, Ebs1(1-591) and Upf1-FL. Proteins (input) were mixed together and incubated then pulled down using calmodulin beads (A) or glutathione resin (B, C) respectively to pulldown on Upf1, Dcp2(UBD) and Dcp1-FL/Dcp2-L. After washing (250 mM NaCl) proteins were eluted, loaded in an SDS-PAGE gel and revealed using Coomassie blue.

Therefore, we performed a similar experiment using GST-Dcp2-UBD as bait. Dcp2-UBD individually co-precipitated Upf1-FL, but not Nmd4 nor Ebs1 (Figure 3B, lanes 2-4). In agreement with this observation, Dcp2-UBD did not precipitate Nmd4 nor Ebs1 when mixed together (Figure 3B, lane 6). In contrast, Dcp2-UBD co-precipitated Nmd4 and/or Ebs1 in the presence of Upf1 (Figure 3B, lanes 5, 7, and 8), demonstrating that the four proteins can co-exist in a single tetrameric complex. In addition, we performed GST-pulldowns using the GST-Dcp1/Dcp2-L heterodimer as bait. Given the instability of the heterodimer Dcp1/Dcp2 the interactions were lower. Our results showed that Upf1, Nmd4 and Ebs1 were also specifically co-precipitated by GST-Dcp1/Dcp2-L (Figure 3C).

Despite the observation that Upf1 holds a central position in the decapping complex, we also explored potential interactions between the other components. We purified a CBP-tagged version of Ebs1 to perform CBP pulldowns (Figure S2A). When using CBP-Ebs1 as bait, the heterodimer Dcp1/Dcp2-L and the isolated Dcp1 protein were specifically co-precipitated, unlike Dcp2-UBD (Figure S2B). This suggests direct interactions between Ebs1 and Dcp1. To confirm this, we performed GST-pulldowns on GST-Dcp2-UBD, GST-Dcp1/Dcp2-L heterodimer and GST-Dcp1 against CBP-Ebs1 (Figure S2C). Consistently, both the Dcp1/Dcp2-L heterodimer as well as the isolated Dcp1 co-precipitated CBP-Ebs1, confirming a direct interaction with Dcp1 (Figure S2C).



Figure S2: Ebs1 directly interacts with Nmd4 and Dcp1

(A) Schematic representation of the different protein domains and recombinant proteins of S.c. Dcp2, Nmd4 and Ebs1. (B-G) *In vitro* pulldown assays. The proteins (Input) were mixed together and incubated before pulldown of the indicated proteins. After washes (250 mM NaCl) proteins were eluted, loaded in an SDS-PAGE gel and revealed using Coomassie blue staining. Protein contaminants are indicated with (#).

Furthermore, we performed TS-pulldowns using TS-Nmd4 as bait against the decapping constructs Dcp1/Dcp2-L2 (Figure S2A), which encompasses the first UBD, and the isolated Dcp1 protein. No direct interaction was detected between the constructs (Figure S2D), which was confirmed by performing the reverse pulldown (Figure S2E). TS-pulldowns of TS-Nmd4 against multiple constructs of Ebs1 bearing different tags showed direct interaction between the latter proteins (Figure S2F). In addition, given the plausible homology between Ebs1 and the Smg5/7 heterodimer which dimerizes *via* its 14-3-3 domain (Fukihara et al., 2005; Jonas et al., 2013), we also performed a GST pulldown on GST-Ebs1 against CBP-Ebs1 (Figure S2G). The results show co-precipitation of CBP-Ebs1 by GST-Ebs1 which suggests that Ebs1 is able to oligomerize *in vitro*.

Yeast Upf1-Upf2 interaction is homologous to the human one

Upf1, Upf2, and Upf3 together constitute the conserved complex essential for NMD (He & Jacobson, 1995; Serin et al., 2001). The overall organization of the Upf complex appears to be conserved, with Upf2 bridging Upf1 and Upf3 (He et al., 1997; Chamieh et al., 2008; Dehecq et al., 2018). Upf2 is a conserved protein composed of three MIF4G domains, the third of which binds Upf3 (Kadlec et al., 2004), followed by a disordered C-terminal region containing the Upf1 binding domain (Figure 4A). The crystal structure of human Upf1/Upf2 interaction showed that Upf2 binds the CH domain in a bipartite manner, binding one side of the CH domain *via* a beta hairpin and on the other side *via* an alpha helix (Clerici et al., 2009). The Alphafold2 model of yeast Upf1-CH interaction with Upf2-Ct highly confidently predicted a similar mode of interaction in which yeast Upf2 would also clamp opposite sides of the CH domain (Figure 4B).

To study the interaction of yeast Upf1 and Upf2 in vitro, we purified a truncated version of Upf2 corresponding to the predicted Upf1 binding domain (He et al., 1997; aa 933-1089, Figure 4A). CBP-pulldown assays with the different versions of CBP-Upf1 showed a direct interaction between Upf1-FL, CH-HD, CH, and Upf2-Ct (Figure 4C), confirming the direct interaction of Upf2 with the Upf1 CH domain. Interestingly, the coprecipitation of Upf2-Ct by CBP-Upf1-FL and CBP-Upf1-CH-HD seemed weaker than by that of the isolated CH domain, and almost disappeared when the washes were performed with higher NaCl concentrations (Figure 4D, lanes 2 and 4). This shows that Upf2-Ct interacts more efficiently with the isolated CH domain than with the CH domain flanked by the HD domain and suggests that the presence of the HD domain interferes with Upf2 binding. The yeast Upf1-CH-HD protein adopts a closed conformation in which the CH domain contacts the helicase RecA2 and stalk domains when bound to RNA in the transition state (Chakrabarti et al., 2011). In this crystal structure, the residue I693 in the RecA2 domain interacts with the F131 residue which is located within a hydrophobic pocket on the surface of the CH domain. In humans, the corresponding residue (F192), is involved in the interaction with the alpha helix of Upf2 (Chakrabarti et al., 2011). To investigate whether the yeast interactions followed the same mechanism, we produced two mutant versions of CBP-Upf1-CH-HD expected to disrupt the intramolecular interaction between the CH and the HD domains. 1693 was replaced by an arginine (1693R) and F131 was replaced by a glutamate (F131E).

Interestingly, I693R mutation restored the co-precipitation of Upf2-Ct by CBP-Upf1-CH-HD under stringent washing conditions, while the F131E completely abolished the interaction (Figure 4E). Therefore, free Upf1 most likely adopts a conformation in which the CH domain contacts the RecA2 domain, involving the hydrophobic residue F131, which is not favorable for Upf2 binding (Figure 4A, 4B).





(A) Schematic representation of yeast Upf2 construct used for pulldown assays. Upf2-Ct corresponds to the (933-1089) C-terminal domain and is tagged with 6-His in C-terminus. (B) Alphafold2 multimer prediction of the interaction between Upf1-CH domain (green) and Upf2-Ct domain. pLDDT scores of Upf2-Ct are marked. pLDDT >90 in dark blue, 70 < pLDDT < 90 in cyan, 50 < pLDDT < 70 in yellow and pLDDT < 50 in orange. (C, D) CBP-Upf1-pulldown assays using Upf1 truncations against Upf2-Ct washed at 250 mM NaCl (C) or 350 mM NaCl (D). (E) CBP-Upf1-pulldown assays using Upf1 CH-HD WT, I693R or F131E against Upf2-Ct. Proteins (input) were mixed together and incubated before Upf1 pulldown using calmodulin beads. After washing (350 mM NaCl), proteins were eluted, loaded in an SDS-PAGE gel and revealed using Coomassie blue. Possible CH-dimers are indicated with (*) and protein contaminants with (#).

Dcp2 and Upf2 compete for interaction with Upf1

Both Upf2 and Dcp2 directly interact with the CH domain of Upf1 (Figures 1 and 4). To explore whether these two Upf1 partners contact the same binding sites, we superposed the Alphafold2 predictions of Upf1-CH/Dcp2-UBD1, Upf1-CH/Dcp2-UBD2 with Upf1-CH/Upf2-Ct. Remarkably, the UDB domains of Dcp2 interact with the outer part of the CH domain in a position similar to that of Upf2's beta sheet (Figure 5A). This apparent clash suggests that Dcp2 and Upf2 cannot simultaneously bind to the Upf1 CH domain.

To test this hypothesis biochemically, we conducted competition assays by pulling down CBP-Upf1-CH from a protein mixture containing a fixed amount of Upf2-Ct with gradually increasing quantities of Dcp2-UBD (Figure 5B). The presence of two or four times more Dcp2-UBD modified neither the amount of Dcp2 nor the amount of Upf2 co-precipitated with Upf1-CH. In contrast, the reverse experiment, in which the amount of Dcp2-UBD was constant while the amount of Upf2-UBD increased, showed that high levels of Upf2-Ct





Figure 5: Upf2 and Dcp2 compete for binding to Upf1-CH domain

(A) Superposition of the AlphaFold2 models of the S.c. Upf1-CH domain (green) bound to Dcp2 UBD1 (wheat) and UBD2 (brown) onto the crystal structure of human UPF1/UPF2 complex (PDB code : 2WJV; Clerici et al. 2009). (B, C) CBP-Upf1-pulldown assays using Upf1-CH against an increasing amount of Dcp2(UBD) with a constant amount of Upf2-Ct (B) increasing amount of Upf2-Ct with a constant amount of Dcp2(UBD) (C). Proteins (input) were mixed together and incubated then Upf1 proteins were pulled down using calmodulin beads. After washing (250 mM NaCl) proteins were eluted, loaded in an SDS-PAGE gel and revealed using Coomassie blue.

almost completely prevented Dcp2 co-precipitation by Upf1-CH (Figure 5C). These experiments clearly show that Upf2 and Dcp2 compete for the same binding site on the Upf1-CH domain. The bipartite interaction of Upf2 with the CH domain seems more stable than that of Dcp2, which most likely binds to only one side of the CH domain, as predicted by Alphafold2. Consistent with our previous data showing that the Upf1-2/3 and Upf1-decapping complexes purified from yeast are mutually exclusive (Dehecq et al., 2018), our results strongly suggest that this incompatibility is, at least in part, due to the competition between Upf2 and Dcp2 for interaction with Upf1-CH.

Upf1 ATPase activity and RNA binding differentially regulated within the Upf1-23 and Upf1-decapping complexes

We previously showed that in humans, Upf2 slightly activates Upf1 ATPase activity *in vitro* (Chamieh et al., 2008). We performed similar experiments by testing the impact of yeast Upf2, Nmd4 and Dcp2 onto Upf1 ATPase activity. To favor the interaction of the different partners and avoid affinity biases, the activity of Upf1 was assayed with a 20-fold excess of each partner and in the presence of a large molar excess of both ATP and single stranded RNA (see Methods). As previously observed with human proteins, Upf2-Ct activates Upf1-CH-HD ATPase activity (Figure 6A). The activation is more pronounced here (around 6-fold) most likely because a higher excess of Upf2-Ct was used than previously (Chamieh et al., 2008). A similar impact was observed in excess of Nmd4 (Figure 6A). In contrast, the addition of Dcp2-UBD only marginally activates Upf1 activity by 2-fold under the same conditions (Figure 6A). We also observed that the effects of Nmd4 and Dcp2-UBD are not





(A) Timecourse of γ^{-32} P release by ATPase activity of S.c. Upf1_CH-HD alone (orange) or incubated with an excess of Nmd4-FL (green), Dcp2(UBD) (brown) or Upf2-Ct (blue). The data correspond to mean values of three replicates and error bars to standard deviations. (B, C) RNA-pulldown assays using Upf1-incubated with the indicated proteins with ATP (B) or ADPNP (C). Proteins (input) were mixed together in presence or not of 3'-biotinylated 33 mer RNA and incubated before RNA pulldown using magnetic streptavidin beads. After washing (300 mM NaCl) proteins were eluted, loaded in an SDS-PAGE gel and revealed using Coomassie blue.

cumulative (data not shown) and that, individually, each UBD has the same effect (data not shown).

We next evaluated whether the same partners affect the ability of Upf1-FL to bind and be retained onto a short 3'-end biotinylated RNA. As for the protein interaction assays, the proteins are first mixed in the presence of RNA and an excess of ATP, then the mixture is incubated at 30°C for 20 min before extensive washing performed at 4°C. In these conditions, CBP-Upf1-FL was poorly retained by RNA (Figure 6B, lane 4). The addition of Upf2-Ct and/or Dcp2-UBD had very little effect (Figure 6B, lanes 5-7). Interestingly, the addition of Nmd4 clearly increased the amount of Upf1 retained allowing the co-precipitation of Dcp2-UBD but not Upf2-Ct (Figure 6B, lanes 8-10). We performed a similar experiment by replacing ATP with ADPNP, a non hydrolyzable analog of ATP (Figure 6C). In the presence of ADPNP, Upf1-FL alone is more efficiently co-precipitated by RNA (Figure 6C, lane 4) showing that ATP hydrolysis by Upf1 favors its detachment from RNA. The addition of Dcp2-UBD slightly stabilized Upf1 onto the RNA, unlike Upf2 which reduced its co-precipitation (Figure 6C, lanes 5 and 6). Interestingly, Dcp2-UBD co-precipitates with the RNA-bound Upf1, unlike Upf2-Ct which is not detectable (Figure 6C, lanes 5 and 6). When adding both Upf2-Ct and Dcp2-UBD, co-precipitation of Upf1 and Dcp2 was drastically reduced (Figure 6C, lane 7). Furthermore, as previously observed with ATP, the presence of Nmd4 highly stabilized Upf1 RNA binding and thus, Dcp2 co-precipitation (Figure 6C, lanes 8 and 9) suggesting the existence of a stable Upf1/Dcp2-UBD/Nmd4/RNA association. The addition of Upf2-Ct, only partially affected RNA binding of Upf1 bound to Dcp2 and Nmd4 (Figure 6C, lane 10), suggesting that within the Upf1-decapping complex, the accessibility of Upf2 is diminished.

Taken together, these results reveal that while Upf2 and Dcp2 share binding sites on Upf1, their binding has strikingly different effects on Upf1 activities: Upf2 highly activates ATPase activity and drastically reduces RNA binding, and Dcp2 only marginally activates ATPase activity and has no visible effect on RNA binding. These observations indicate that Upf1's characteristics are regulated differently whether it is in the Upf1-2/3 or Upf1-decapping complexes.

Discussion

Our *in vitro* interaction assays using purified proteins revealed direct interactions between Upf1, the decapping heterodimer Dcp1/Dcp2, and two other components of the Upf1-decapping complex: Nmd4 and Ebs1 (Figures 1, 2). We reconstituted the complex (Figure 3) and notably showed that Nmd4 strongly enhances Upf1 RNA binding (Figure 6). Our results suggest that 1) Upf1 directly recruits the decapping machinery to the target mRNA, and 2) the Upf1-decapping complex is highly stabilized on RNA by Nmd4, unlike the Upf1-2 complex. In addition, our interaction assays revealed that additional direct contacts occur within the Upf1-decapping complex between Ebs1 and Dcp1, Nmd4, and itself, thus positioning Ebs1 as a central regulator of the Upf1-decapping complex (Figure S2). It is tempting to speculate that all physical interactions contribute to the overall stability of the Upf1-decapping complex *in vivo*. This intricate web of interactions would

explain why deletions of the UBD domains of Dcp2 and single deletions of Ebs1, and Nmd4 have no strong effect on NMD (Dehecq et al., 2018; He et al., 2022).

The hydrolysis of the 5' m7GpppG structure by the heterodimer Dcp2/Dcp1 releases a 5' monophosphate extremity and irreversibly commits the RNA to rapid degradation by the exoribonuclease Xrn1. Our compelling evidence proving the direct interaction of Dcp2 and Upf1 (Figure 1) confirms the direct recruitment of the decapping machinery to NMD substrates by Upf1 in yeast. This mechanism had already been proposed a long time ago by two hybrid experiments (He & Jacobson, 1995), however, only recently, in vivo experiments in yeast provided more concrete evidence on this recruitment (Dehecq et al., 2018, He et al., 2022; Ganesan et al., 2022). The direct interaction we demonstrated not only paves the way for structural studies, which have already provided significant insights into the mechanistics of decapping in the past (Charenton et al., 2016), but also opens the door to study if decapping is regulated within the Upf1-decapping complex. While it is possible that the decapping machinery is only recruited to the target mRNA, it also may be activated by their direct partners. Indeed, the decapping activator Edc3 was also found in the Upf1-decapping complex, which we did not include in our reconstitution given that its association was RNA-dependent (Dehecg et al., 2018). In vivo, Edc3 helps Upf1 association with Dcp2 (He et al., 2022) but it only affects the decapping of two specific RNAs (Badis et al., 2004; Dong et al., 2007), which raises questions about its general role within the Upf1-decapping complex. One possible explanation is that decapping complexes differ in their composition in a transcript -specific manner, among which only some contain Edc3. Furthermore, it is possible that Upf1 only serves to recruit Dcp2 given that the UBD domains are spatially far away from the catalytic Nudix domain. In humans, the C-terminal region of Dcp2 is not conserved, meaning that the Upf1-binding domains determined in yeast do not exist in human Dcp2. The recruitment of decapping during NMD in humans is ensured through the decapping activator PNRC2 (Cho et al., 2009; Lai et al., 2012), suggesting more elaborate dynamics and additional layers of regulation and control of this 'no-return' step.

We next delved into investigating the mutual exclusivity between the Upf1-decapping and Upf1-2/3 complexes. To accomplish this, we purified the C-terminal region of yeast Upf2, predicted to interact with the CH domain of Upf1, and demonstrated that, similar to the interaction observed in humans, Upf2 binds to the CH domain in a bipartite manner (Figure 4). Competition assays between Upf2 and Dcp2 binding to the CH domain clearly unveiled the incompatibility of these interactions (Figure 5), shedding light on the *in vivo* mutual exclusivity of the complexes (Dehecq et al., 2018). According to the original Detector-Effector model, the formation of the Upf1-Upf2-Upf3 complex is the initial step of NMD and enables the recognition of the target mRNA (Dehecq et al., 2018). Once the target is detected, an event occurs that liberates Upf2 and Upf3 while recruiting the other Upf1 partners. In this linear model, it is assumed that the Upf1-Upf2-Upf3 complex exists while Upf1 is bound to the RNA given that their interaction is not RNAse sensitive (Dehecq et al., 2018). However, recent findings in humans indicate that Upf1's interaction with Upf2 is incompatible with Upf1-RNA binding due to conformational changes upon interaction (Xue et al., 2023). When RNA was added, the Upf1/Upf2 complex dissociated, and Upf2

addition released Upf1 from RNA (Xue et al., 2023). Our data, which demonstrate that yeast Upf1 loses its RNA affinity when bound to Upf2 (Figure 6) support this counterintuitive result. It is tempting to speculate that Upf2 may have a higher affinity for the Upf1 CH domain than Dcp2 (Figure 5), thus we suggest that the Upf1-decapping complex precedes the Upf1-Upf2-Upf3 complex. These observations imply that Upf2 may be involved in Upf1 recycling rather than its recruitment. Supporting this model, Upf2 and Upf3 are necessary for efficient translation termination of 5' intermediate decay NMD targets, indicating that decapping occurs before ribosome dissociation and thus the association of Upf2 and Upf3 (Serdar et al., 2016). Interestingly, the deletion of Upf2 leads to a shift of Upf1 association to heavier polysomal fractions, which could potentially reflect defects in Upf1 recycling (Ganesan et al., 2022). The deletion of Upf2 also led to a shift on Nmd4 association to polysomes contrary to deletion of Upf1 which abolished it (Dehecq et al., 2018), reflecting that Nmd4 is recruited even in absence of Upf2. Additionally, studies of fluorescence in situ hybridization (FISH) observed aggregate formation of NMD targets upon Upf2 deletion, also indicating recycling defects (Sayani et al., 2008). Previous work by Sheth and Parker proposed that Upf1 acts prior to Upf2 and Upf3, as deletion of Upf2 and Upf3 increased the formation of P-bodies containing Dcp2, while no accumulation was observed when Upf1 was deleted (Sheth & Parker, 2006). Although it cannot be excluded that the Upf core complex is involved in both early and late steps of NMD, it is possible that Upf2 and Upf3 co-sediment with large polysome fractions only due to their interaction with release factors or the ribosome itself (Serdar et al., 2016; Ganesan et al., 2022; Wang et al., 2001; Min et al., 2013), rather than their direct association with Upf1 within an mRNP. Overall, the sequential progression of events may be more complex than initially assumed, indicating a more intricate process. It is also plausible that multiple Upf1 states coexist within the same mRNP and that a Upf1-Upf2-RNA mRNP is a mere transitory state.

Part 2: Single molecule characterization of yeast Upf1

1) Magnetic tweezers for the study of Upf1 helicase activity

a) The intricate nature of helicase activity

Enzymes are typically characterized by a single catalytic activity. However, helicases, which function as nanometric molecular motors, possess a more intricate nature. Similar to motors, they utilize the energy released from nucleoside triphosphate (NTP) hydrolysis to translocate along NAs (Wu & Spies, 2013; Patel & Donmez, 2006; Abdelhaleem, 2010). Their activity is much more complex given that they coordinate and repeat multiple different activities for their function: NTP binding and hydrolysis, NA binding, release (Patel & Donmez, 2006). Rather than functioning in isolation, these activities are tightly coordinated, forming a cohesive system within the helicase in which each activity relies on and influences the others. These different catalytic activities are translated into the properties of the helicase, which confer them their biological roles (Abdelhaleem, 2010). Helicases can bind different substrates (double stranded (ds), simple stranded (ss) or duplex DNA or RNA), translocate along their substrates (unwind or rezip a dsNA), interact with and/or remodel roadblocks (NA tertiary structures or NA-binding proteins). These activities are defined by their processivity (number of consecutive base pairs translocated), polarity (5'-3' or 3'-5'), affinity to NA, step size, and rate (number of nucleotides translocated per NTP hydrolysis cycle and number of steps per second) (Abdelhaleem, 2010). The process of helicase NA unwinding can be seen as a cyclical mechanism comprising repeated steps that act as the limiting factors for the overall rate (Wu & Spies, 2013; Patel & Donmez, 2006). Despite their structural similarities, individual helicases possess distinct characteristics as their classification, based on amino acid patterns, does not consistently align with their signature activities (Wu & Spies, 2013; Gorbalenya & Koonin, 1993), and even closely related sister helicases can exhibit markedly different behaviors (Kanaan et al., 2018). Importantly, their activities are tightly regulated either by regulatory flanking domains or by their partners, to ensure specific spatiotemporal activity.

b) Single-molecule approaches to study helicases

Bulk assays for helicase characterization (reviewed in Bianco, 2021) enable the qualitative assessment of average protein activity within a heterogeneous pool (>10⁸ molecules). Assays have progressed from demonstrating their capacity to translocate along NA and hydrolyze NTPs for strand separation to advanced single-molecule characterization. The latter not only offers visual comprehension but also provides information about the forces associated with the functioning of motor proteins (Bianco, 2021). The development of optical microscopy and micromanipulation techniques have allowed the study of helicases in nanometric and millisecond scales, which enable us to estimate the heterogeneity of protein population and to study subpopulations that can co-exist and have different activities (Bianco, 2021). Single-molecule manipulation techniques, such as optical tweezers (OT), magnetic tweezers (MT), and atomic force microscopy (AFM), involve applying mechanical forces to a NA substrate while monitoring the helicase activity (Sun & Wang, 2016; Monachino et al., 2017).

AFM records the position of a sharp tip by collecting the reflected laser beam from a micro-lever. To detect the unwinding activity of a helicase, a NA molecule is attached between the micro-lever and the surface (Fisher et al., 2000). While this approach is mainly used for surface topography imaging and nanomaterial characterization, it has been used to measure the force generated by a helicase (Marsden et al., 2006) Yet, its application in kinetic studies is limited due to its wide force range and relatively low resolution (Sun & Wang, 2016).

OT involves directing a laser beam through a high numerical aperture lens to focus the beam. This results in a concentrated light intensity at the focal point, creating a trap where particles with a higher refractive index than the surrounding medium can be confined. This laser-induced trap can act as a grip, allowing for the movement of particles by adjusting the position of the focal point. To study helicases, suitable NA substrates are attached to particles on one end and to a coverslip or a second optical trap on the other end (Comstock et al., 2015). This setup enables the application of adjustable forces ranging from 0.1 to 100 pN on the molecules. The real-time tracking of the particle's position within the trap and the exerted force allows for precise visualization of molecular motor movement (Svoboda & Block, 1994; Mallik & Gross, 2004). OT offers the advantage of high-speed data acquisition, reaching frequencies in the tens of kilohertz range, which enables the detection and analysis of rapid dynamics within biological systems. Additionally, two optical traps can be combined, reducing instrument noise and ensuring a spatial resolution of a few base pairs. This level of resolution is well-suited for precise measurements like determining the step size of a helicase. Another main advantage of OT is that it can be integrated with the detection and tracking of fluorescent particles, such as in the case of smFRET (single molecule Förster/fluorescence resonance energy transfer) experiments (Comstock et al., 2015).

MT is a micromanipulation technique initially developed by our collaborators in the physics department to measure the elastic properties of DNA (Allemand et al., 1998; Strick et al., 1998; Bustamante et al., 2000). Since, it has been used for studying several NA-related enzymes, among which, numerous helicases (Dessinges et al., 2004; Manosas et al., 2010; Hodeib et al., 2017; Valle-Orero et al., 2022; Ruiz-Gutierrez et al., 2023). Compared to OT, in the vision field of MT, multiple particles can be simultaneously tracked. This allows the parallel measurement of thousands of single molecule events in real time and confers robust statistical relevance (Sun and Wang, 2016). Consequently, average data can be associated with unique behaviors, enhancing the characterization of molecular properties. However, MT exhibits lower resolution compared to OT. The spatial and temporal resolution of measurements is constrained due to the tracking of the paramagnetic bead's position using a camera. Nevertheless, recent developments in video-based three-dimensional microscopy (Rieu et al., 2021), substrate design (Cheng et al., 2011) and experimental setups (Abbondanzieri et al., 2005) have allowed higher spatio-temporal resolution, thus achieving single base accuracy. In addition, integration of fluorescent-labeled components as in OT, has recently allowed the monitoring of more dynamic complexes (Graves et al., 2015; Fan et al., 2016).

Altogether, these assays grant access to helicase characteristics that are usually inaccessible by classical biochemical analyses. For instance, their behavior when confronted to roadblocks (NA-binding proteins or tertiary structures) (Park et al., 2010; Manosas et al., 2013; Fiorini et al., 2015; Valle-Orero et al., 2022), strand switching behaviors (Valle-Orero et al., 2022; Fiorini et al., 2015), the variability in step length (Cheng et al., 2011; Qi et al., 2013; Syed et al., 2014) or precise speed and processivity (Fiorini et al., 2015; Kanaan et al., 2018). Furthermore, contrary to single molecule techniques, bulk helicase activity assays, usually performed using (radio) labeled double stranded NAs, only allow the observation of the unwound and duplex state of the NA, overlooking the possible intermediate states where the helicase is not able to fully unwind the dsNA (Bianco, 2021).

c) Magnetic tweezers experimental setup for helicase activity monitoring

The underlying principle of MT is as that a NA molecule (usually DNA), is tethered on one end to a superparamagnetic bead (approximately 1 µm in diameter) and on the other end to the glass surface within a microfluidic home-made chamber mounted under an inverted microscope (Ruiz-Gutierrez et al., 2022, see Annex) (Figure 25A). The inlet port allows the injection of the RNA-bound beads, proteins and buffer exchanges (Figure 25B). The outlet port is connected to a home-made automatic aspirator (Figure 25B). The microchamber is illuminated to create diffraction rings around the beads, and the PlayItAgainSam® software analyzes these rings to determine the position (ΔZ) of each tracked bead with a high accuracy (≈ 1 nm). Above the sample, a pair of permanent magnets generates a horizontally oriented magnetic field (Figure 25A). The superparamagnetic bead experiences a force that controls its vertical position and a torque that can immobilize the bead in a specific orientation. The force magnitude, which can be finely controlled, increases as the magnets approach the surface, resulting in forces ranging from 0.1 pN to up to 30 pN. When subjected to high force (>20 pN for RNA and >13 pN for DNA), the NA is mechanically elongated and adopts its maximum length (Figure 25C). By adjusting the position of the magnets along the vertical axis, the attached NA molecule can be stretched, while rotating the magnets allows for twisting of the NA. The position of the paramagnetic bead is continuously monitored with nanometer precision for extended periods, allowing tracking for over thousands of minutes. In contrast to OT, MT naturally exerts a constant force on the NA without the need for a feedback system, enabling the simultaneous recording of dozens of molecules. In the case of helicases, the NA is designed to fold into a self complementary sequence (hairpin, HP), which can be unwound by the enzyme when injected in low quantities (<10 nM) (Figure 25D). A constant intermediate force (\approx 7-14 pN) needs to be applied for the HP to remain closed yet under tension without undergoing elongation (Figure 25C). Depending on the helicase's polarity, the HP can be designed to have single stranded regions for helicase loading either in 3' or in 5'. In excess of NTPs, once the helicase is loaded onto its substrate, it translocates within the double stranded HP while unwinding the NA (step 1, 2 in Figure 25D, 25E).

Results Part 2



Figure 25: Magnetic tweezers experimental setup (modified from Ruiz-Gutierrez et al., 2022) (A) Nucleic acid attachment and microscope setup. (B) Structure and assembly of a magnetic tweezers microfluidic chamber (C) Force extension and bead position (D) Schematic representation of helicase activity unwinding and rezipping a nucleic acid hairpin in magnetic tweezers and (E) corresponding hypothetical trace.

Upon reaching the apex of the hairpin, if it is still bound to the NA, it can translocate onto the single stranded NA (step 3 in Figure 25D, 25E). As the helicase travels, the HP closes behind it (step 4 in Figure 25D, 25E). We call this phase rezipping. The conversion of the beads' position relative to the elongation of the HP is made by measuring its position at maximal force and dividing it by the known length of the substrate. 1 nm approximately corresponds to 1 bp. By deducing the real-time position of the helicase along the NA, its velocity and processivity can be estimated.

2) Single molecule studies reveal different autoregulations of human and yeast Upfl on RNA

a) Background and aim of the study

Our team has extensively studied the activity of yeast (S.c.) and human (H.s.) Upf1 by employing a combination of single molecule and biochemical techniques. Bulk assays demonstrated that the CH and SQ domains of human Upf1 inhibit the helicase and ATPase activities of the HD domain (Chamieh et al., 2008; Chakrabarti et al., 2011; Fiorini et al., 2013). Using magnetic tweezers, they measured the human HD domain translocation speed (~1 bp/s) and showed that despite its slow speed, it was remarkably processive on both RNA and DNA (Fiorini et al., 2015). The inhibition by the CH domain was also confirmed using magnetic tweezers (Fiorini et al., 2015). Furthermore, they characterized the yeast Upf1 helicase core on DNA using magnetic tweezers, which exhibited a 10-fold higher translocation speed (~10 bp/s) compared to its human counterpart (Kanaan et al., 2018). Notably, it demonstrated exceptional processivity on DNA, capable of translocating over 10 kb without releasing its substrate (Kanaan et al., 2018). Despite the application of external forces, the helicase core maintained a strong grip on its substrate, tightly binding it in the absence of ATP (Kanaan et al., 2018). This robust binding to DNA was attributed to the presence of the 1B and 1C protrusions within the helicase core. Mutating these regions significantly affected processivity and the ability to grip DNA (Kanaan et al., 2018). However, S.c. Upf1 investigations focused on the isolated helicase core due to its higher structural stability, disregarding eventual intramolecular effects. Similarly, given that Upf1 is able to bind and translocate along DNA and RNA, all experiments on S.c. Upf1 were performed using DNA as a substrate due to its lower susceptibility to degradation and lower secondary structure formation.

To evaluate the potential modulation of the S.c. Upf1 activity by its CH domain we purified CBP-tagged *H.s.* and S.c. Upf1 CH-HD and HD constructs (Figure 26A) and performed bulk ATPase assays (see Methods). To limit biases due to protein batch concentration, we confirmed the measured concentrations by loading the proteins in an SDS-PAGE gel (Figure 26A). Our results confirmed what was previously shown for human Upf1 (Fiorini et al., 2015; Chakrabarti et al., 2011; Fiorini et al., 2013), showing that the isolated HD domain had 4-fold higher activity than the CH-containing construct (Figure 26B, compare red with blue). Unexpectedly, the yeast CH-containing construct had 3-fold higher ATPase activity than the isolated HD domain (Figure 26B, compare green with orange). ATPase activity level differences between yeast and human isolated HD domains were proportional to the previously measured translocation speeds and correlated with the 10-fold difference

between human and yeast Upf1 isolated helicase cores (Figure 26B, compare orange with red), suggesting a tight link between ATPase hydrolysis and translocation speed. When comparing the CH-HD proteins, a striking 37-fold difference was observed between yeasts and humans (Figure 26B, compare green with blue). While our results are not an absolute estimation of each protein's activity due to eventual difference in protein batches, they clearly show an opposite inherent autoregulation of yeast and human Upf1 activity by the CH domain, leading to the conclusion that the yeast Upf1 is much more active than the human one.



Figure 26: Yeast and human Upf1 are oppositely regulated by the CH domain (A) Schematic representation of Saccharomyces cerevisiae (S.c.) and Homo sapiens (H.s.) UPF1 truncations constructs used for ATPase assays and SDS-PAGE acrylamide gel loading (500 ng) of the recombinant proteins. (B) Time course of γ^{-32} P release by ATPase activity of H.s. UPF1 and S.c. UPF1 constructs (HD and CH-HD). The data correspond to mean values of three replicates and error bars to standard deviations.

b) Yeast Upf1 CH domain stabilizes the helicase core grip on RNA

The unforeseen autoregulation of the S.c. Upf1 ATPase activity by the CH domain (Figure 26) prompted us to employ MT to gain a better understanding of the regulatory impact of the CH domain on the helicase activity in yeast. Additionally, in order to assess the potential influence of the C-terminal domain, we purified the FL protein. In collaboration with the team lead by Vincent Croquette at the Laboratory of Statistical Physics of the Ecole Normale Supérieure (LPS ENS), we synthesized a 180 bp hairpin (HP) using a two-step protocol of *in vitro* transcription of a double stranded DNA palindromic sequence (Figure 27A) (Ruiz-Gutierrez et al., 2022). At constant force (10 pN) and with saturating concentrations of ATP, the recombinant proteins of S.c. Upf1 were injected onto the RNA bound surfaces at low concentrations to ensure single helicase events.

S.c. Upf1 FL and CH-HD displayed characteristic regular and unpaused unwinding and rezipping of the RNA HP (Figure 27B), while the isolated helicase core of S.c. Upf1 not only had an elevated rate of substrate release compared to the CH-containing proteins but also displayed more paused and slower events (Figure 27B). We thus calculated the processivity by measuring the number of translocated nucleotides (unwound and rezipped)

and dividing by the number of falling events (Figure 27C). For example, Upf1-HD translocated 64,834 nucleotides during the experiments and fell 520 times, yielding an estimated processivity of 124 nucleotides, lower than the total length of the HP (Figure 27C).



Figure 27: Yeast Upf1 helicase activity is enhanced by the CH domain (A) RNA hairpin used as substrate. (B) Traces corresponding to the translocation of yeast Upf1 FL (purple), CH-HD (green) and HD (orange). (C) Calculated processivities of the different Upf1 proteins. (D) Unwinding and rezipping speeds of yeast Upf1 proteins

However, given that very few falling events were measured for the CH-containing proteins, processivities could only be estimated since they are higher than the translocated number of nucleotides (Figure 27C). Despite this, processivities of the CH-containing proteins were significantly higher (> 100-fold) than that of the isolated HD domain. In addition, we measured the average unwinding and rezipping speeds of the proteins. Both the FL and CH-HD presented comparable speeds, calculated to be around 6 nt/s during unwinding and 9 nt/s during rezipping (Figure 27D). The isolated HD domain presented slightly lower unwinding (\approx 4 nt/s) and rezipping (\approx 3 nt/s) rates (Figure 27D).

Altogether, these results confirm that the CH domain influences positively the helicase activity of S.c. Upf1, contrary to what has been observed for humans (Chakrabarti et al., 2011; Fiorini et al., 2013; Fiorini et al., 2015).



c) The behavior of the yeast Upf1 helicase core differs on DNA and RNA

Previous magnetic tweezers assays on the helicase core of S.c. Upf1 performed on DNA showed extremely high processivity and DNA grip (Figure 20, Kanaan et al., 2018), contrary to our measures on RNA (Figure 27). To confirm that the behavior of S.c. Upf1 HD domain on RNA was not due to the presence of the N-terminal CBP tag, we tested an untagged HD construct which showed comparable speeds and processivities to that of the tagged version (Figure 28A, 28C), confirming no tag bias. In addition, we confirmed that force had no effect on Upf1 activity by applying higher forces (15 pN) during the experiments (Figure 28B, 28C), which rendered similar processivities and translocation speeds to those recorded at 7 pN (Figure 27).

Thus, our results suggest that the activity of S.c. Upf1 helicase core is noticeably different depending on the substrate. To better compare the behavior of S.c. Upf1 in the same conditions we synthesized a DNA HP substrate identical in sequence and length to that of the RNA substrate used previously (Figure 29A). The CH-bearing constructs displayed slower and more paused events than those measured on RNA (compare Figure 27B and 29B), but similar to those of the isolated helicase core on both RNA and DNA (compare Figure 27B and 29B). For the FL and CH-HD proteins, many of these characteristic saw-tooth events displayed longer pauses when the apex was reached (Figure 29B). However, most events of all constructs displayed complete unwinding and rezipping of the DNA HP (Figure 29B), which resulted in high estimated processivities (Figure 29C). Interestingly, the calculated unwinding speeds of the CH-bearing constructs (\approx 5 nt/s) were lower than that of HD (\approx 9 nt/s), contrary to what we measured for RNA (compare Figure 27D and 29D). However, all constructs presented comparable rezipping speeds (\approx 8 nt/s) (Figure 29D). Similar to the behavior on RNA, no significant difference was observed in either parameter between the FL and CH-HD constructs, suggesting that the C-terminal domain has no effect on yeast Upf1 activity (Figure 29). Interestingly, the differences of speed and processivity observed between the isolated HD domain and the CH-bearing constructs was not as important as in RNA.

Altogether, these results show how the difference of activity between the isolated helicase core and the CH-bearing construct are less marked on DNA, suggesting the helicase core may have a better affinity or a tighter grip on DNA than on RNA. The results we obtained confirmed the previous published results of high processivity and strong grip of the isolated helicase core of yeast Upf1 on DNA (Kanaan et al., 2018). However, it appears that on RNA, the helicase core requires the CH domain to maintain a tighter grip and high processivity.



Figure 29: Yeast Upf1 helicase enhancement by CH is less striking in DNA (A) DNA hairpin used as substrate. (B) Traces corresponding to the translocation of yeast Upf1 FL (dark purple), CH-HD (dark green) and HD (blue). (C) Calculated processivities of the different Upf1 proteins.

3) Discussion

While MT are a powerful tool, a great disadvantage is the low concentration of protein that can be injected into the surface. Indeed, depending on the NA binding affinity of the protein of interest it can vary but usually remains in the nM magnitude. Therefore, two problems arise from this. First, if high quantities of protein need to be injected (>1 μ M), the microfluidic chamber is quickly saturated by unspecific interactions which prevents bead hybridization and thus renders it single-use. Second, if the dissociation constant (Kd) of a protein-protein interaction is greater than a few nM, it is virtually impossible to study the complex using MT. Indeed, when Fiorini and collaborators attempted to study the regulation of *H.s.* Upf1-CH-HD helicase activity by Upf2 using MT (Fiorini et al., 2015), they injected the preformed complex into the surface and observed that only 76% of the events were affected. Although the exact concentration of protein injection was not mentioned, the results are consistent with the Kd of this interaction which is of 200 nM (Clerici et al., 2009). It is thus important to be aware of the influence of Kd when using MT.

Furthermore, recent single molecule experiments raised concerns regarding the influence of the force applied to the HP on helicase activity during our MT experiments (Chapman et al., 2022). The authors argue that Upf1 might behave like a passive helicase and thus be activated by the force applied. Our experiments conducted at higher forces (Figure 28) demonstrate that there is no bias induced. In addition, force-independence of human Upf1 was also previously demonstrated in the lab (Fiorini et al., 2015). It is also known that Upf1 is not a passive helicase but rather an active one. Indeed, it is true that among the SF1 family, some sister helicases are force-dependent, however, Upf1 is not one of them. It is important to note that all recombinant proteins utilized in this study were produced and purified from *E. coli*, meaning that they may lack post-translational modifications present in their natural context. Consequently, to validate our findings which contribute to a better understanding of the regulatory mechanisms and functional differences of Upf1 helicases in yeast and humans, further investigations employing complementary bulk and *in vivo* assays are warranted.

Part 3: RNA degradation by the Upf1-decapping complex components

1) Nmd4 as a potential active endonuclease

Nmd4 is a component of the Upf1-decapping complex (Figure 3 of Results Part 1; Dehecq et al., 2018). It directly interacts with the helicase domain of Upf1 (Figure 2 of Results Part 1; Barbarin-Bocahu et al., *in preparation*). Thus, it may regulate Upf1's helicase activity. In the quest of understanding whether Upf1 is regulated by Nmd4, and if so, how, we studied its activity using MT. We injected a mix of Upf1 and an excess of Nmd4 into an RNA HP-bound microfluidic chamber.

During the first experiment, we injected a mix of 1 mM ATP, 30 nM Upf1-HD and 1 μ M of Nmd4-FL onto a surface in which 14 HP-beads were tracked. To our surprise, after injection we observed rapid bead loss (100%), with an apparent exponential kinetic in less than 30 min (Figure 30A). This unexpected observation contrasted with the high stability demonstrated in our previous experiments using RNA HP substrates (See Results Part 2), where we could record experiments for several hours without bead loss. We hypothesized that this observation could be due to either RNAse contamination of the mixture or the buffers or to an eventual endonuclease activity of Nmd4.

To dig deeper, we injected 1 mM of ATP with lower concentrations of Nmd4 (100 nM) in absence of Upf1 into a new RNA-bound surface in which 22 beads were tracked. We also observed rapid bead loss after injection, but with an apparent linear kinetic (Figure 30B). Interestingly, when we applied high force to open the hairpins, many beads were immediately lost (Figure 30B, blue lines). This observation suggested that the RNA HP had been previously cleaved within the complementary region, either on the double stranded region or in the single stranded loop (Figure 30C), but since it was closed, the bead stayed attached at intermediate forces (Figure 30C). Given the two different kinetics observed with

С







(A, B) Magnetic tweezers measures of bead attachment to RNA surface in time at 7 pN after addition (vertical dotted red line) of 1 mM ATP, 1 nM Upf1-HD and 1 μ M of Nmd4-FL (A) or 1 mM ATP and 100 nM Nmd4-FL (B). The start and the end of the experiment are marked by the black vertical dotted lines and hairpin openings by applying higher forces (30 pN) are marked by the blue vertical dotted lines. (C) Structure of the RNA hairpin. Possible endonucleolytic cleavage sites are marked by arrows. Single stranded regions are marked in red. Blue segments correspond to double stranded regions of RNA (light blue) and DNA (dark blue). or without Upf1 it is possible that Upf1 has an effect on Nmd4 activity. However, given the low affinity constants measured for this complex (>1 μ M) (Barbarin-Bocahu et al., *in preparation*), it is difficult to know whether the complex is formed or not in the experiment. In addition, this difference could be due to the difference in concentration of Nmd4 injected in the two experiments. Therefore, it is necessary to confirm our results with further experiments.

To test whether our results were due to a nuclease contamination of the protein purification or a *bona fide* endonucleolytic activity of Nmd4, our collaborators provided us with three constructs of Nmd4: the PIN domain, the C-terminal domain and the full length (FL) protein. We performed bulk endonuclease assays using a radiolabelled linear short mRNA which we incubated at different timepoints with the Nmd4 proteins before separating the RNA species on a Urea-acrylamide gel (Figure 31). After autoradiography, no degradation was detected for the PIN and C-terminal constructs of Nmd4, however the FL construct presented a smear denoting potential nuclease activity, especially visible at the 90 minute time point, whether the buffer contained magnesium or manganese (Figure 31A). Interestingly, when EDTA was added instead of divalent cations, the smear was not observable (Figure 31B).

These preliminary results point towards an eventual endonucleolytic activity of Nmd4, which would be detectable at the single molecule level, at least *in vitro*. However, at this stage we cannot rule out the possibility of a nuclease contamination of our protein purifications or buffers, a non negligible possibility given that nucleases are ubiquitous and extremely hard to denature. Nevertheless, given that Nmd4 possesses a PIN domain, it is clearly not impossible that it retained a residual endonuclease activity.



2) Development of a quantitative decapping assay to study decapping regulation

To assess Dcp2 activity alone or in presence of its partners within the Upf1-decapping complex, we developed a quantitative in vitro assay to measure decaping activity over time. This luminescence-based assay was first attempted by our collaborators in the team of Cosmin Saveanu to assess in vivo purified complexes. They set out to purify the Upf1-decapping complex as they previously did (Dehecq et al., 2018) using a genomically tagged Dcp1-TAP yeast strain. However, they did not manage to get sufficient quantities to obtain reproducible results. Given that our recombinant protein yields were higher, we adapted their protocol to develop a quantitative in vitro decapping assay. A short (33 nt) RNA was transcribed in vitro and capped using the Vaccinia virus capping enzyme (NEB). We also tested a long RNA molecule (152 nt) co-transcriptionally capped. After purification, the RNA was treated with the decapping holoenzyme (with or without its partners) for a certain duration. The samples were then inactivated by heat and then treated by Xrn1. During the Xrn1 treatment, only the uncapped substrates can be degraded by Xrn1, which releases free nucleotides including adenosines. Adenosine quantities were then dosed by a luminescence-based commercial kit (AMP-glo, Promega, Figure 32A). By dosing the quantity of Xrn1-sensitive transcripts over time from a four-point kinetic curve, we can calculate initial reaction velocities and compare the decapping activity with or without partners. To ensure the reliability and reproducibility of our results, we made several notable modifications to the original protocol such as heat inactivation of the decapping proteins to avoid phenol purification of the RNA. In addition, we optimized the in vitro transcription and capping process of the RNA and tested different protein concentrations and buffers which needed to be compatible with each step of the protocol. As a positive control, we produced and purified a CBP-tagged version of Edc3, which is known to activate decapping (Charenton et al., 2016).

Using the long RNA, which contained 54 adenosines, in presence of the Dcp1/Dcp2(1-663) heterodimer, we managed to get a linear decapping activity (Figure 32B, brown), effectively confirming the functionality of our proteins. When we incubated the samples with a 5-fold molar excess of Edc3, we observed a slight positive activation as expected (Charenton et al., 2016) (Figure 32B, gray). In addition, a 5-fold molar excess of Upf1-CH resulted in a 2-fold increase in decapping activity (Figure 32B, orange). When using the short RNA, which contained 14 adenosines, the decapping activity of Dcp1/Dcp2(1-663) was also linear (Figure 32C, brown). However, when incubating with 5-fold excess of Upf1-CH, no significant difference on decapping activity was observed (Figure 32C, orange), and surprisingly, Upf1-CH-HD had a negative effect (Figure 32C, light blue).

While our preliminary results using the long substrate initially suggested that Upf1 may play a role in the activation of decapping, the use of the shorter substrate did not corroborate our observations. We thus performed a control using the shorter version of Dcp1/Dcp2(1-315) which lacks the UBD domains, and saw that the addition of Upf1-CH negatively impacted decapping activity, even though they should not interact (Figure 32D, compare beige with orange). This suggests that our previous results using the short RNA

are not exploitable. This inhibition may be due to the 5-fold excess of partners we added. Indeed, it is possible that there is a steric hindrance of accessibility to the cap by Dcp1/Dcp2 when using high amounts of protein and a short RNA, even if the experiments were performed with a 10-fold excess of RNA compared to Dcp1/Dcp2 (0.2 μ M of Dcp1/Dcp2, 2 μ M of RNA and 1 μ M of partner). To verify this hypothesis we need to optimize the protocol and perform more controls to validate our assay. In addition, we should also test the effect of Ebs1 on decapping.



Discussion & Perspectives

Discussion & Perspectives

The objective of this project was to explore the underlying mechanisms of NMD in yeast with an *in vitro* biochemical approach. Before I started this work, our collaborators revealed the existence of two mutually exclusive Upf1-bound complexes (Dehecq et al., 2018). The complexes were named the Upf1-detector (or Upf1-2/3) complex, comprising the core Upf proteins Upf1, Upf2, and Upf3, and the Upf1-effector (or Upf1-decapping) complex, which included the decapping holoenzyme Dcp1 and Dcp2 and novel partners such as Nmd4, Ebs1, and Hrr25. Our aim was to characterize these complexes *in vitro* using recombinant proteins in order to gain a better understanding of (1) the intramolecular protein interactions, (2) the dynamics of the complexes and (3) the activity and regulation of Upf1.

1) Structure of the Upf1-decapping complex

a) Upf1 structure: human versus yeast

Regulation of Upf1 activity by the CH domain

Serving as the main orchestrator, the multidisciplinary protein Upf1 plays a pivotal role in NMD. It is well acknowledged that the human CH and SQ domains inhibit helicase and ATPase activity. This was deduced from comparative experiments conducted on recombinant proteins of human Upf1, both in the absence and presence of the CH domain (Chamieh et al., 2008; Fiorini et al., 2013). In addition, complementary in vitro experiments showed that the addition of Upf2 activates Upf1 (Chamieh et al., 2008; Chakrabarti et al., 2011; Fiorini et al., 2015). The interpretation of this mechanism was based on several crystal structures. First, the human CH-HD Upf1 bound to Upf2 showed that Upf1 adopts an open conformation when bound to Upf2 (Figure 12, Clerici et al., 2009). Second, the structure of the yeast Upf1 CH-HD showed the close conformation of the CH domain in the absence of Upf2 (Figure 18, Chakrabarti et al., 2011). Third, the isolated helicase domain of H.s. Upf1 compared to the isolated helicase domain of S.c. Upf1 showed high structural homology and conservation when RNA-bound (Figure 18, Cheng et al., 2007; Chakrabarti et al., 2011). The latter was the basis to extrapolate that the unsolved H.s. Upf1 CH-HD structure adopts the same closed conformation as S.c. Upf1 CH-HD. Thus, the current model is that Upf2 relieves Upf1's inhibition by interacting with the inhibitory CH domain and displacing it from the helicase domain.

We observed that yeast Upf1 is not inhibited by the CH domain, on the contrary, it is activated (Figures 26, 28 and 29). In agreement with our results, the crystal structure of the yeast Upf1 CH-HD showed a larger RNA-binding surface than the isolated HD domain as seen by the larger RNAse protected fragments in presence of the CH domain (Chakrabarti et al., 2011). Indeed, the presence of the CH domain, which interacts with the RecA2 domain, pushes the protrusion 1B towards the 3' end of the RNA, leading to additional contacts between domain 1B and three nucleotides, thus strengthening the interaction and potentially also the processivity (Chakrabarti et al., 2011).

The human model was built upon biochemical experiments that were only validated for human proteins but based on yeast structures. No confirmation through the biochemical study of yeast proteins was performed. However, given that no structure of *H.s.* Upf1-CH-HD alone has been solved, it is possible that it adopts the same open conformation even in absence of Upf2 and not the closed conformation of its yeast counterpart. This would explain why it inhibits human helicase activity by the sterical hindrance of a flapping surface. Thus, more structural studies of human and yeast Upf1 alone, bound by its partners and bound to its substrates (RNA and/or ATP) are needed. It is interesting to note that even with high levels of homology, it is risky to extrapolate results from yeasts to humans or *vice versa*.

Upf1 dimerisation

Another interesting observation we made was the possibility that yeast Upf1 may form dimers through the CH domain. Indeed, when the CH-bearing constructs of yeast Upf1 were loaded on SDS-PAGE gels, they showed the presence of a band of exactly double molecular weight than that of the expected monomer. Interestingly, neither the yeast Upf1 constructs that lacked the CH domain nor the human proteins presented this band. Western Blots confirmed that the band corresponds to a CBP-tagged protein and mass spectrometry analysis revealed that it corresponds to Upf1. The band is seen despite highly denaturing conditions (SDS-PAGE), suggesting it is a very stable dimer. At first we speculated that it was a byproduct of oxidation, given that the CH domain bears multiple cysteines which could form disulfide bonds. However, even with high reducing conditions $(>10 \mu M DTT)$, the band was still present. In vitro pulldown assays mixing the CH and CH-HD yeast Upf1 constructs, revealed the apparition of a third band corresponding to the exact molecular weight of a CH/CH-HD dimer, which was also CBP tagged as confirmed by Western Blots. Again, this supplementary band was not observable when using the human proteins. Curiously, the F131E mutant of yeast Upf1 (Figure 5) did not present the 'dimer' band after purification. Given that the F131 residue is located within the CH domain, it suggests that the 'dimer' formation is not a matter of cysteine disulfide bond but may rather correspond to an extremely stable three dimensional structure. Overall, the accumulating evidence we obtained indicating the existence of yeast Upf1 dimers in vitro reflects that it may not only be due to experimental in vitro biases, but might indeed be present and have a function in vivo. The fact that none of these observations were seen when using the human homologues would indicate that no dimer formation occurs in humans. This could partially explain the differential activities we observed between yeasts and humans (Figure 26).

However, at this point, we cannot interpret a role for Upf1 dimerisation in NMD. It is possible that the dimerisation is not important for NMD or Upf1 activity but rather plays a role during translation for folding and stability, as it is the case for many other protein complexes (Levy et al., 2016). Even if the monomer is clearly stable, the dimerization could help to stabilize Upf1 co-translationnaly. Upf1 dimerisation was first suggested by the lab of Alan Jacobson through their two hybrid assays which showed intermolecular interactions between CH domains (He et al., 2013). They suggested a role in decapping recruitment,

given that there are two UBDs with redundant functions (He et al., 2013, He et al., 2022). This hypothesis has been largely contradicted given that *in vitro* purification of yeast and human Upf1 showed that it was stable as a monomer in solution (Czaplinski et al., 1995; Bhattacharya et al., 2000; Cheng et al., 2007) and no evidence of dimerization was observed *in vivo*.

To deepen our understanding of these suspected dimers, we will soon characterize the activity of the F131E mutant which does not form this species. Whether this mutant shows a specific phenotype *in vivo* would be interesting to assess, but distinguishing the effect from the lack of interaction with Upf2 is probably impossible. In addition, it would be interesting to separate both species, dimers and monomers, by gel filtration and assess their ATPase activity and/or helicase activity using MT.

b) An intricate web of interactions within the Upf1-decapping complex

By specifying the direct interactions between proteins within the Upf1-decapping complex (Figures 1, 2, S1 and S2), our results reflect how there is a much more intricate network of protein-protein interactions than we initially expected. Ebs1 interacts with three out of the five components that we included in our study, as well as multiple Upf1 domains (Figures 2 and S2). This suggests that it may act as a scaffold protein, contributing to the formation of the Upf1-decapping complex. Similarly, the redundancy of Dcp2-UBD domains (Figures 1 and S1) reflects the importance of decapping recruitment. All of these redundant interactions that we observe in vitro, suggest that in vivo, there may be multiple failsafe mechanisms to ensure the recruitment of decapping in case of NMD dysregulation. For instance, the fact that deletion of the C-terminal domain of Dcp2 has a very low effect on NMD in vivo (He et al., 2022), could be due to its indirect recruitment through the Ebs1/Dcp1 interaction (Figure S2). This would compensate for the lack of direct recruitment by Upf1. To check for this possibility, our collaborators created a yeast strain containing a C-terminal truncated genomic version of Dcp2 and deleted Ebs1 (Δ Ebs1/Dcp2(1-300)). They verified the stability of a known NMD substrate, the pre-L28 mRNA, by RT-qPCR. However, no significant increase in the levels of the mRNA were observed between the Δ Ebs1 and Δ Ebs1/Dcp2(1-300) strains, suggesting no additional impact on NMD (personal communication).

Three hypotheses emerge from these findings: (1) It is possible the truncated version of Dcp2 is more stable in the cell given that it lacks the disordered C-terminal tail, therefore it is more expressed and thus compensates for the lack of recruitment. In addition, this truncation of Dcp2 lacks its autoinhibitory domain positioned downstream the Nudix domain, therefore it could be also more active than the FL version, (2) It is also possible that other supplementary mechanisms of compensation take place *in vivo* to avoid NMD dysregulation as seen by the fact that only the double deletion of Nmd4 and Ebs1 had a significant effect on NMD (Dehecq et al., 2018), (3) Given that the NMD defects were only measured by RT-qPCR, it is possible that other targets could be affected at a transcript specific level. Therefore, it would be interesting to perform whole transcripts. The same observation goes for deletion of the other components of the Upf1-decapping complex.

The loss of Upf1 interaction with Dcp2 in absence of the UBD domains has only been observed by two-hybrid experiments (He & Jacobson, 1995, He et al., 2022). Therefore, it would be interesting to perform *in vivo* co-immunoprecipitation assays to confirm this observation. I performed *in vitro* pulldowns on Upf1 in presence of the heterodimer Dcp1/Dcp2(1-315) (data not shown), but the results were not conclusive given that the negative controls were not clean. In a short term perspective, when I manage to find better pulldown conditions, I will test whether it is possible to reconstitute the Upf1-decapping complex with the Dcp2 version lacking the UBDs in presence of Ebs1.

c) CryoEM for solving the structure

The reconstitution of the Upf1-decapping complex (Figure 3) opens the door for structural studies. Our initial project included the possibility of X-ray crystallization by our collaborators in the team of Marc Graille of the protein-protein interaction complexes we would find *in vitro*. Indeed, they recently solved the structure of Upf1-HD/Nmd4 (Bocahu-Barbarin et al., *in preparation*). As for the other interactions, one big challenge was the extremely low protein purification yield and stability of the components. Especially the disordered C-terminal domain of Upf1, Ebs1, which seems to be highly unstructured and the long region between the Nudix domain and the UBD domain of Dcp2. Nevertheless, given the high stability of the Dcp2-UBD protein, it would be conceivable to attempt the crystallization of the Upf1-CH/Dcp2-UBD complex in the future.

The recent advances achieved in protein structure prediction through deep-learning programs like AlphaFold2 or RoseTTAFold may help with this task, as they have the potential to revolutionize the field of biology for decades to come. A good example is the crystal structure of *Kluyveromyces lactis* Nmd4 protein that was recently published by our collaborators. The structure was solved using an AlphaFold2 prediction to determine the phase of the crystal they obtained (Barbarin-Bocahu & Graille, 2023). Furthermore, the AlphaFold2 predictions that we obtained for the Upf1/Dcp2 and Upf1/Upf2 interactions (Figures 1 and 4) are strikingly consistent with our *in vitro* results and clearly illustrate how helpful and accurate these models can be.

Although X-ray crystallography seems unlikely to be achieved, the Upf1-decapping complex we reconstituted is potentially a good candidate for CryoEM studies. In collaboration with the team of Elena Conti in the Max Planck Institute, we will attempt to solve the structure using this technique. However, we never attempted gel filtration assays to confirm the tetrameric nature of the complex because our protein yields were very low. Thus, the main challenge will certainly be to maintain the stability of the complex during gel filtration. It is possible that the complex will need to be crosslinked. Altogether, solving the structure of the Upf1-decapping complex would pave the way towards a better understanding of Upf1's regulation by its partners. As it has been seen with the Upf1/Upf2 crystal structure (Clerici et al., 2009), Upf1 may adopt a very specific conformation when bound to its partners and substrates. The stabilization of Upf1 RNA-binding by Nmd4 could also be observed. It may also be possible to observe the dimerization we suspect. In addition, the source of incompatibility of both Upf1-bound complexes could be established if we have access to the structure of Upf1-CH-HD bound to Dcp2.

2) An evolutionary conserved mechanism?

Nonsense-mediated mRNA decay is an essential mechanism, not only for eliminating aberrant mRNAs but also for fine-tuning the transcriptome. It is conserved from yeasts to humans, however, the main actors and the current accepted models are not all conserved. The possible functional and structural homology between Nmd4, Ebs1 and Hrr25 with the Smg proteins (Figure 23H) lays the first stones for a universally conserved mechanism.

a) Nmd4 as a potential active endonuclease

The degradation machinery recruited for mRNA decay during NMD varies between organisms. In yeast, the primary degradation process involves deadenylation-independent decapping (Muhlrad & Parker, 1994). However, in higher eukaryotes such as human and drosophila, the predominant mechanism relies on the endonucleolytic activity of the PIN domain of Smg6 (Gatfield & Izaurralde, 2004; Glavan et al., 2006; Huntzinger et al., 2008). The PIN domain endonucleolytic activity is governed by three conserved aspartic acids within the catalytic site. Interestingly, Smg5 lacks two of the residues within the triad, rendering it inactive (Figure 13; Glavan et al., 2006; Huntzinger et al., 2008). Conversely, Nmd4 retained two out of three conserved residues (Barbarin-Bocahu & Graille, 2023). Given that the mutation of the equivalent third residue in Smg6 reduces its endonucleolytic activity (Glavan et al., 2006), Nmd4 was expected to be inactive. However, our preliminary results fortuitously revealed that it may be active (Figures 30 and 31). Probably not as active as WT Smg6, but active enough to detect RNA cleavage in vitro. Interestingly, when analyzing further the previous published results for Smg6, only the double mutation showed high inactivation (Glavan et al., 2006). Mutating only one out of the three conserved catalytic residues of Smg6 only partially affected its activity (Glavan et al., 2006). This would mean that a PIN domain containing at least two aspartates such as Nmd4 could be indeed active.

The main obstacle to validate our observations will be to distinguish between a nuclease contamination and a bona fide endonucleolytic activity of Nmd4. Usually, nuclease contamination is due to human or bacterial nucleases which are ubiquitous. The loss of RNA degradation when adding EDTA (Figure 31) shows that the activity we detect is dependent on a divalent cation. However, most nucleases rely on divalent cations for their catalytic activity. It would be interesting to find a specific PIN endonuclease inhibitor to validate our hypothesis but, to our knowledge, no such inhibitor exists. To circumvent this problem, it is possible to produce a radiolabeled circular RNA that only endonucleases can cleave (Lebreton et al., 2008). Furthermore, to develop a more quantitative endonuclease assay we could take advantage of the single molecule resolution of MT. Our HP substrate is not designed to evaluate endonucleolytic activity since losing the beads on the vision field implies having only one event per bead. This requires preparing new surfaces for every experiment, which is time consuming and expensive. Thus, it would be worth exploring alternative designs of RNA substrates. For example adding a DNA 'leash' (Kostrz et al., 2019) that would allow the beads to remain attached after RNA cleavage. After cleavage, the RNA could be religated by washing and injecting an RNA ligase into the surface. Contrary to bulk experiments with linear RNAs, this MT substrate would allow us to control
the size and the number of cleavage sites within our substrate and/or specific sequences and thus have access to more precise data on the kinetics of the reaction. Nevertheless, the best confirmation will be to evaluate the endonuclease activity of single and double mutants of the catalytic site (D44A, D139N) of Nmd4 that our collaborators have already purified. In addition, they produced WT or mutated versions of the PIN domain of Smg6 as positive and negative controls.

Then, we will test whether Upf1 influences NMD4 nuclease activity. Curiously, no studies have focused on the regulation of the endonucleolytic activity of Smg6 by Upf1 and the other components of the DECID complex. Given the high impact of Nmd4 on Upf1's RNA binding and ATPase activity (Figure 6), it would be worth investigating whether human Upf1 activity may be modulated by Smg6 and *vice versa*.

While many laboratories have tried to look for signals of endonucleolytic cleavage of NMD targets in yeast, they have failed to prove that NMD targets are degraded through such mechanism (personal communication from Bertrand Seraphin and Roy Parker). Whether Nmd4 is involved in direct RNA degradation or only in Upf1 activity modulation *in vivo* will be important to elucidate in the future.

b) The role of phosphorylation in yeast NMD

Our collaborators at the Institut Pasteur focused on the phosphorylation targets of the kinase Hrr25 *in vivo*. Given that the deletion of Hrr25 is lethal, they designed a mutant which is inactive when bound to a non hydrolyzable ATP homolog called 3-MB-PP1. Using this mutant, they observed that Upf1 would migrate as a double band in Western Blots, which was lost when adding 3-MB-PP1 (personal communication). When deleting the C-ter domain of Upf1, no upper band was observed (personal communication). Together, these results suggest that the C-terminal domain of Upf1 is the target of phosphorylation by Hrr25.

Given the phospho-dependent interactions between Upf1 and Smg5/Smg7 in humans (Chakrabarti et al., 2014), the fact that Ebs1 directly interacts with the C-terminal domain of yeast Upf1 (Figure 2), and the resemblance between Ebs1 and Smg5/Smg7 (Figure 23H), we explored whether there could be a phospho-dependent interaction between Ebs1 and Upf1. Three phosphorylation sites were identified by mass spectrometry study of yeast Upf1 (personal communication of the laboratory of Pedro Beltrao): S869, S884 and S887. During the internship of Jeanne Dupas, an M1 student I supervised during the spring 2020, we produced recombinant phosphomimetic Upf1 to test whether Ebs1 interacts better with Upf1 when its C-terminus is phosphorylated. We produced simple or double phosphomimetic mutants of two of the three potential phosphorylated serines to aspartates (S884D and S887D) which mimic phosphorylated serines. No clear difference in apparent affinity was observed in our binding assays using the phosphomimetic and the WT versions of Upf1 (data not shown). However, we did not test the third serine of interest due to lack of time, but it will be important to test it before concluding. In addition, phosphomimetics are usually employed in *in vivo* experiments where an additional control can be performed: mutating the phosphorylation target to an inert residue which cannot be

phosphorylated (alanine). Therefore, it is possible that our approach is not adapted for this specific question. Thus, complementary to our approach, our collaborators in the team of Marc Graille intend to perform isothermal titration calorimetry experiments with a phosphorylated peptide to test the hypothesis.

While there has been one study examining the impact of phosphorylation on human Upf1 RNA binding footprints (Kurosaki et al., 2014), how phosphorylation affects Upf1's activity is still elusive. Overall, the implications of phosphorylation on yeast NMD are still poorly characterized despite the fact that the main components of NMD (Upf1, Upf2 and Dcp2) have been shown to be phosphorylated in yeast (Wang et al., 2006; Yoon et al., 2010); Lasalde et al., 2014). Thus, future experiments should focus on this additional layer of regulation. Moreover, we did not include Hrr25 in our study because we suspected very transient interactions since it is a kinase. However, the protein is cloned and purified so we could test its kinase activity *in vitro*.

3) Revisiting the mechanistic details of yeast NMD

a) What is the role of Upf1?

It is well-established that Upf1's ATPase activity, helicase activity and processivity are crucial for NMD (Weng et al., 1996b; Kashima et al., 2006; Franks et al., 2010; Kurosaki et al., 2014; Kanaan et al., 2018; Lee et al., 2015; Chapman et al., 2022) and that in humans, Upf1 phosphorylation is important for partner recognition and activity regulation (Ohnishi et al., 2003; Kashima et al., 2006; Cho et al., 2009; Kurosaki et al., 2014; Chakrabarti et al., 2014; Flury et al., 2014). However, despite these findings, our understanding of Upf1 remains quite limited and many questions remain unanswered. Is it implicated in processes like ribosome displacement and RNA secondary structure remodeling and if so, how? Indeed, despite depictions commonly placing it on the 3' end, the exact positioning of Upf1 relative to the ribosome is unknown. Does Upf1 remain tightly bound to the RNA or translocates along it? Altogether, many aspects of its function and behavior are unclear.

One of the most important gaps in our knowledge is the precise mechanism by which Upf1 recognizes premature termination codons and distinguishes them from *bona fide* stop codons is still not fully understood. Some studies proposed that Upf1 is bound to all accessible mRNAs (Hogg & Goff, 2010; Zünd et al., 2013; Kurosaki et al., 2014). Based on this observation, Chapman and collaborators recently proposed two models for Upf1 binding and its role in NMD: the 'butterfly model' and the 'bulldozer model' (Chapman et al., 2022). The 'butterfly model' rests upon the facts that ATP hydrolysis leads to RNA release (Czaplinski et al., 1995; Bhattacharya et al., 2000) and ATP deficient mutants remain firmly attached to their targets, losing thus target selectivity (Kurosaki et al., 2014; Lee et al., 2015). In this model, Upf1 would be a promiscuous helicase that senses the prey by skimming the RNA. On the other hand, the 'bulldozer model' is based on the fact that Upf1 is able to remodel mRNPs (Fiorini et al., 2015) and is extremely processive on DNA and RNA (Figure 26; Fiorini et al., 2015; Kanaan et al., 2018). This model represents Upf1 as a war machine that once it is bound to its target, no RBP nor secondary structure can disturb it. Only the study from Fiorini and collaborators studied the CH-HD protein but most *in vitro*

studies of Upf1's characteristics have been performed using DNA and only focused on the isolated HD domain (Fiorini et al., 2015; Kanaan et al., 2018; Chapman et al., 2022), ignoring the potential regulations by the flanking domains. Our results underscore the importance of studying the full length proteins, as isolated domains can show strikingly different activities (Figures 27 and 29).

While these opposite models are both well supported by Upf1 characteristics, none of them have been strictly proven to be true. Consequently, the fundamental question persists: how does Upf1 leverage its activity to initiate NMD? A significant challenge in unraveling this aspect likely stems from Upf1's involvement not only in NMD but also in numerous other cytoplasmic RNA decay pathways, its potential roles within the nucleus which remain relatively unexplored, and that it is difficult to follow the successive steps of NMD and to isolate complexes that most likely exist only transiently.

b) Correlation between Upf1 speed, NMD efficiency and P-body formation

It is very intriguing that despite their high homology (90%) (Applequist et al., 1997) the CH domains of human and yeast have such different effects on the helicase domain (Figure 26). However, the fact that even a single mutation can have substantial effects on the activity of Upf1 (Kanaan et al., 2018) supports this counterintuitive result, meaning that the 10% difference there is between the two domains might be enough to explain the difference. If we take a closer look at the sequences of these domains, we find that the linker between the CH and HD domains is longer in humans than in yeast. Previous work in our lab by Joanne Kanaan showed that a chimera of yeast Upf1 bearing the human linker presented no difference on its helicase activity on a DNA substrate on MT. We should produce the reverse chimera to evaluate whether it is the shorter yeast linker that has an effect and redo these experiments using our new RNA substrate given that the differences are more pronounced on RNA (Figures 27 and 29).

Furthermore, the substantial difference of activity levels between the human and yeast proteins is also intriguing (Figure 26, Fiorini et al., 2015; Kanaan et al., 2018). It is tempting to question how ATP hydrolysis and translocation speed affect NMD itself. Is there a threshold or a limit of Upf1 activity needed to trigger NMD? If so, why is it different between yeasts and humans? Could it be that given the high activity of Upf1 in yeast, NMD is more efficient in this organism? Or that in humans, a strong Upf1 trans-activator is required to reach a comparable high activity?

An intriguing connection has been established between NMD and P-bodies. Mutations affecting the ATPase and/or helicase activities of Upf1 induce the formation of P-bodies in both yeast and human. Within these P-bodies, Upf1 and other NMD factors co-localize alongside their NMD targets (Sheth & Parker, 2006; Franks et al., 2010). Similarly, deletion of Upf2 and Upf3 lead to P-body formation (Sheth & Parker, 2006). Although P-bodies are not necessary for NMD (Stalder & Mühlemann, 2009), it is possible that defaults in NMD trigger P-body formation as a way to safeguard the cell. This observation could serve as a new tool to assess NMD efficiency in cells, complementary to mRNA stabilization. Interestingly, human cells maintain P-bodies under normal physiological conditions, whereas stress induction is required to trigger P-body formation in yeast. This disparity

prompts further inquiry into a potential correlation between genome complexity and NMD efficacy. Is it possible that smaller and simpler genomes inherently entail fewer errors to rectify, making NMD activation and regulation less necessary? Alternatively, could the slower activity of human Upf1 be a deliberate adaptation that allows for better fine-tuning and substrate selection?

An interesting example of differential activity levels between highly conserved proteins are bacterial RNA polymerases (RNAP). In *B. subtilis*, RNAP is 5-fold faster than in *E. coli*, thus it outpaces the pioneer ribosome which leads to transcription and translation decoupling (Johnson et al., 2020). This example suggests that the discrepancies between yeast and human Upf1 activity levels may serve a purpose which is clearly not evident at the moment.

c) Upf2 recruitment temporality

While the Upf trinity is a relatively well-characterized complex in NMD due to its essential role in the mechanism, the dynamics of this complex remain uncertain. It has been largely assumed that the recruitment of the Upf1-2/3 complex represents the initial step in triggering NMD. However, besides their co-precipitation in the same polysomal fractions (Atkin et al., 1997), there is no other data to support this assumption. Our findings indicate that an mRNP complex containing Upf1 bound to Upf2 is either extremely transient or nonexistent in yeast given that Upf2 drastically reduces Upf1 affinity for RNA (Figure 6). These results align with recent data demonstrating the same trend in humans (Xue et al., 2023). In contrast, we observed that the Upf1-decapping complex is significantly stabilized on mRNA by Nmd4 (Figure 6). In addition, based on our observations, Upf2 seems to exhibit a stronger affinity for the Upf1 CH domain compared to Dcp2 (Figure 5). These findings suggest that Upf2 could displace Dcp2 from Upf1 and thus play a role in the recycling of Upf1 rather than its initial recruitment. While the involvement of Upf2 in both steps of NMD cannot be ruled out, it is worth considering that the co-sedimentation of the Upf core complex with large polysome fractions may be attributed to their interaction with release factors or the ribosome itself (Serdar et al., 2016; Ganesan et al., 2022; Wang et al., 2001; Min et al., 2013). This raises the possibility that their presence in these fractions may be due to indirect interactions with release factors or the ribosome, rather than direct association with Upf1 within an mRNP.

Based on these findings, we suggest revising the current model of NMD in yeast (Figure 33) in which the Upf1-decapping complex precedes the assembly of the Upf1-2/3 complex. In this model, Upf1 is first recruited alone to the terminating ribosome, either by the ribosome itself (Ganesan et al., 2022) or simply by its promiscuous binding to RNA (Hogg & Goff, 2010), without the direct involvement of Upf2/Upf3. Subsequently, Upf1 would recruit decay-inducing proteins such as Dcp1/Dcp2, Nmd4, and Ebs1 (Figure 3), leading to the stabilization of its binding to the mRNA (Figure 6). Upon decay triggering through decapping and potentially by the endonucleolytic activity of Nmd4 (Figures 30 and 31), Upf1 would release the decay-inducing proteins and subsequently recruit Upf2/Upf3.



Figure 33: Proposed model for NMD in yeast

Upf1 is recruited to the mRNA by the ribosome or by self binding, a long 3' allows maintaining Upf1 bound to the RNA. The Upf1-decapping complex is assembles and triggers decay. A conformational switch happens that allows recruitment of Upf2/3 for Upf1 recycling and liberation of eRF3 for translation termination.

The mechanism underlying this switch remains an open question. It is possible that phosphorylation of Upf1 by Hrr25 induces a conformational change, liberating the CH domain from its intramolecular interaction with the HD domain. This conformational switch could enable the binding of Upf2/Upf3 to Upf1 by displacement of Dcp2 and the subsequent release of Nmd4 and Ebs1. The interaction between Upf1 and Upf2/Upf3 would not only facilitate the liberation of Upf1 from its association with the RNA (Figure 6), but also contribute to the release of Upf2/Upf3 from their interaction with eRF3 and thus translation termination.

The precise mechanisms governing the recruitment of decay-inducing proteins are still not fully understood but may be influenced by translation termination efficiency, as proposed in the Faux 3' UTR model (Figure 11; Amrani et al., 2004). Whether decay is activated within this mRNP or simply recruited remains unknown.

While acknowledging the speculative nature of certain aspects of this model, it is important to note that the proposed model is consistent with the results obtained in our study and the existing knowledge of yeast NMD. The need for *in vivo* confirmation of these hypotheses is recognized and further experimentation is warranted to validate the model's predictions. If our proposed model is accurate, it would be expected that the Upf1 I693R mutant, which exhibits enhanced interaction with Upf2 (Figure 4), should have an impact on NMD or, at the very least, on the recycling of Upf1. To test this hypothesis, Cosmin Saveanu's team conducted experiments by complementing a Δ Upf1 strain with the mutant version and examining RNA stability of their selected mRNA NMD targets using RT-qPCR. However, no discernible differences in RNA stability were detected. To further validate the model, additional experiments could be performed. For example, microscopy could be employed to assess whether the mutant induces P-body formation. Transcriptome sequencing could be conducted to investigate potential effects on mRNA targets on a genome-wide scale. Furthermore, it would be valuable to examine whether the formation of the Upf1-decapping complex is affected by the mutant.

Moreover, our observations suggest that NMD may be more evolutionarily conserved than previously believed. The functional homology observed between the Smg proteins in humans and the components of the Upf1-decapping complex in yeast supports the notion that NMD shares common principles across different organisms. This finding suggests that the revisited model we propose for yeast NMD could potentially be adapted to the EJC-independent model of NMD in humans.

TL;DR (too long; didn't read)

- Dcp2 is directly recruited to the mRNA by Upf1 and may be modulated by Upf1
- Nmd4 highly stabilizes Upf1 onto the RNA and may be an active endonuclease
- Ebs1 contacts multiple proteins within the Upf1-decapping complex
- Dcp2 and Upf2 compete for interaction with Upf1 but Upf2 wins
- Human and yeast Upf1 autoregulations by the CH domain are opposite
- Upf2 may serve for Upf1 recycling rather than Upf1 recruitment

Materials & Methods

Materials & Methods

1) Molecular biology

a) Molecular clonings

Most of the molecular clonings carried out during this PhD were intended to produce plasmids for recombinant protein purification in *Escherichia coli*. Three main techniques were performed depending on the clonings: restriction enzyme, phosphorylated primers or Gibson assembly. The common steps of all three techniques will be described in the last section. All clonings were thoroughly designed *in silico* beforehand using Benchling.

Restriction enzyme clonings

Classical restriction enzyme clonings were performed either to fuse coding sequences with existing vectors or to insert tags into existent plasmids. In this technique the following steps were performed and will be detailed further: (1) PCR amplification of the insert, (2) digestion of insert and vector, (3) purification of insert and vector, (4) ligation, (5) bacterial transformation, (6) plasmid verification.

Directed mutagenesis through phosphorylated primers

Point mutations and sequence insertions or deletions in the coding plasmids of interest were achieved by the phosphorylated primers technique, in which a PCR is performed with phosphorylated oligonucleotides designed in the zone of interest as indicated in Figure 34.

PCR amplification was followed by ligation, MH1 transformation and plasmid verification for obtention of the final plasmid. Phosphorylation of the primers was performed with 5 μ L of Buffer A PNK 10X, 5 μ L ATP (10 μ M), 500 μ M of primer and 1 μ L of PNK (Thermofisher) for 30 min at 37°C. The product was extracted with phenol chloroform and ethanol precipitated. The oligo was resuspended in 50 μ L of water and the concentration was measured using Nanodrop before diluting to 10 μ M.



Figure 34: Phosphorylated primers cloning technique

The phosphorylated primers were then used for amplifying the desired plasmid using a classical PCR technique described further. The PCR product was then treated with 5 μ L of DpnI enzyme for 1 h at 37°C to digest all initial plasmid. After gel purification the plasmid was ligated and transformed into MH1 *E. coli* strain. The plasmid was then verified, amplified and stocked as described further.

Gibson assembly

The Gibson assembly technique (Gibson et al., 2009) (Figure 35) was used for cloning plasmids where restriction enzymes were not available.



The insert was either ordered as a gene block (IDT) or amplified from an existing plasmid. For amplification, primers containing the homology sequence (20-40 nucleotides) were ordered from Eurofins. The vector was usually linearized by restriction digestion or by PCR.

Figure 35: Gibson assembly cloning technique (Figure extracted from Gibson et al. 2009)

Isothermal	Buffer:
Sourcemat	Dunci.

PEG-8000	25%
Tris-HCl pH 7.5	500 mM
MgCl ₂	50 mM
DTT	50 mM
dNTPs	1 mM each
NAD	5 mM

Gibson mix 2X:

Isothermal Buffer	100 µL
TS exonuclease (10 U/µL)	0,2 µL
Taq DNA ligase (40 U/µL)	50 µL
Phusion (2 U/µL)	6,25 µL
Complete with H ₂ O to	250 µL

Gibson assembly was performed as follows: in a 1.5 mL Eppendorf tube mix 10 μ L of Gibson mix 2X, 1 μ L of vector (100 ng/ μ L), the equivalent of 2-fold more in mols of insert and complete with water to 20 μ L. Incubate 1 h at 50°C. Transform *E. coli* MH1 with 2 μ L (10 ng) of Gibson assembly.

Common steps

• DNA amplification by PCR:

Buffer HF or GC 5X	10 µL
dNTPs (10 mM)	1 µL
Forward primer (10 µM)	2.5 µL
Reverse primer (10 µM)	2.5 µL
Vector (10 ng/µL)	1 µL
DMSO	2.5 µL
Phusion (Thermofisher)	1 µL
Complete with H_2O to	50 µL

Program:

- 98°C, 2 min (initial denaturation)
 - 98°C, 15 sec (denaturation)
 - Tm°C, 15 sec (hybridization)
 - 72°C, X sec (extension)
- 72°C, 5 min (final extension)
- 10°C, hold

X: 30 sec per kb

Tm°C calculated using NEB Tm calculator

• <u>DNA digestion</u>: Most restriction enzyme digestions were performed using enzymes from NEB, thus compatible with CutSmart 10X buffer.

CutSmart 10X	5 µL
Vector (1 µg/µL)	1 μL-5 μL
Enzyme 1 (20 U/µL)	1.5 µL
Enzyme 2 (20 U/µL)	1.5 µL
Complete with H_2O to	50 µL

→ Incubate 30 min at 37°C. Inactivate enzymes according to the manufacturer's indication.

- <u>DNA purification</u>: DNA purification was performed according to manufacturer's indications using Wizard ® SV Gel and PCR Clean-Up System (Promega). For plasmids and PCR products presenting more than one band, gel purification was favored.
- <u>Ligation</u>:

T4 ligase buffer 10X	1.5 µL
Vector	50 ng
Insert	3 fold in mols
T4 Ligase (NEB)	0.8 µL
Complete with H ₂ O to	15 µL

→ Incubate 2 h at RT or overnight at 16° C.

<u>Bacterial transformation</u>: 100 μL of competent MH1 *E. coli* were thawed on ice, 10 ng of DNA (3 μL of ligation or 2 μL Gibson assembly) were added, samples were incubated 30 min on ice then 2 min at 42°C for thermic shock. After 10 min incubation on ice, 1mL

of LB media was added and samples were incubated for 1h at 37°C. 100 μ L were plated in a *Petri* dish containing LB agarose complemented with the desired antibiotic. The dishes were incubated overnight at 37°C or over the weekend at room temperature (RT).

• <u>Verification by colony PCR</u>: after transformation, 5-10 single colonies were selected and added to the following mix :

Green buffer 10X	1.5 µL
dNTPs (10 mM)	0.4 µL
Forward primer (10 µM)	0.4 µL
Reverse primer (10 µM)	0.4 µL
DMSO	0.6 µL
Colony	Isolated clone
Taq polymerase	0.2 µL
Complete with H_2O to	15 µL

Program:

- 94°C, 10 min (initial denaturation)
 - 94°C, 30 sec (denaturation)
 - Tm°C, 30 sec
 - (hybridization)
- 72°C, X sec (extension)
- 72°C, 10 min (final extension)
- 10°C, hold
- X: 1 min per kb
- <u>Miniprep and sequencing</u>: for positive clones minipreps were performed using PureYield ® Plasmid Miniprep System (Promega) using 2 mL of bacterial culture. Samples were sent for Illumina SupremeRun sequencing at Eurofins.
- <u>Midiprep and stock</u>: one positive clone was selected from sequencing and plasmids were prepared using NucleoBond ® Xtra Midiprep System (Macherey-Nagel). Stocks were measured using Nanodrop and stored at -20°C.

b) RNA

For RNA experiments all conditions must be RNAse free. All materials were thoroughly cleaned using RNaseZap [®] (ThermoFisher) prior to the experiments and filter cones were systematically used.

In vitro transcription of uncapped RNA

Two different RNA were synthesized: a 154 nt RNA corresponding to the *tpi1* gene sequence (Long), a known target of NMD in yeast, and a 32 nt RNA based on the *tpi1* gene but modified to avoid secondary structures and add more adenosines (Short). For the long RNA, the compatible vector for polymerization using SP6 RNA polymerase was obtained by restriction enzyme cloning performed in Cosmin Saveanu's lab (pHL 1670). The vector was linearized by enzyme digestion followed by phenol extraction and chloroform precipitation and the quantity was measured using Nanodrop. For the short RNA, two complementary oligonucleotides (500 μ M each) were ordered from Eurofins and hybridized for *in vitro* transcription.

The HiScribe ® SP6 RNA Synthesis kit (NEB) was used for *in vitro* transcription:

SP6 Reaction Buffer 10X	20 µL
ATP, UTP, CTP, GTP 50 mM	20 µL each
Vector	2 µg
SP6 polymerase mix	20 µL
Complete with H ₂ O to	200 µL

→ Incubate 2 h at 37°C and purify by TBE-Urea acrylamide gel (detailed further).

In vitro capping

In vitro capping of the transcribed and purified short RNA was performed using Vaccinia Capping System (NEB). First, 5 μ g of RNA (32 nt) were diluted in a final volume of 15 μ L, heated at 65°C for 5 minutes and placed on ice for 5 minutes. The capping reaction was performed as follows:

Denatured RNA (0.5 µg/µL)	15 µL
10X Capping Buffer	2 µL
GTP (10 mM)	1 µL
SAM (2 mM)	1μL
Vaccinia Capping Enzyme	1μL
Complete with H ₂ O to	20 µL

 \rightarrow Incubate 90 min at 37°C.

The capped RNA was then treated by 1 μ L of Xrn1 (NEB) per μ g of RNA for 1h at 37°C.

Co-transcriptional capping of RNA

Co-transcriptional capping was performed as follows:

SP6 Reaction Buffer 10X	20 µL
ATP, UTP, CTP (50 mM)	20 µL
GTP (50 mM)	4 µL
m7GpppG (100 mM)	8 µL
Linearized vector	2 µL
Complete with H_2O to	200 µL

→ Samples were mixed and incubated for 2 h at 37°C. 16 μ L of DNAse 1 were added and samples were incubated for 15 min at 37°C. RNA was extracted by acid phenol and precipitated using ethanol. The sample was resuspended in 3 μ L of H₂O, 1 μ L Buffer NEB3, 6 μ L Xrn1 (NEB) and incubated 1 h at 37°C before being purified on gel.

TBE-Urea acrylamide gel RNA purification

For a large 5% acrylamide TBE-Urea gel the following products were mixed:

TBE 1X, 7M Urea	26.6 mL
TB1 1X, 7M Urea, 15% acrylamide 19/1	13.4 mL
APS 100%	240 µL
Temed	48 µL
Final volume	40 mL

→ The mix was left to polymerize for at least 2 h. 20 μ L RNA Loading Dye 2X (NEB) were added to each sample. Samples were denatured 3 min at 95°C and thawed in ice right after. The gel was loaded and migrated in 1X TBE, 20-40 mA, 1 h.

When migration was over, the gel was placed between two sheets of Saran wrap over a TLC plate and the band under 254 nm UV light was localized and cut. The RNA was eluted in 400 μ L of elution buffer (300 nM NaCl, 60 mM AcONa, 0,2% SDS), overnight, in rotation at room temperature. The next day, 1 mL EthOH 100% (salt is already included in the elution buffer) was added and the RNA was precipitated at -80°C for 2 h. Pellets were centrifuged, washed twice with EthOH 70% and resuspended in 21 μ L of water. RNA was quantified by Nanodrop and stored at -20°C.

c) Protein visualization

Protein quantification

Three methods of protein quantification were tested during my PhD: Nanodrop measures, Bradford assay and Qubit protein quantification. All give significantly different results depending on the proteins but among the methods, Bradford was the most accurate one. Bradford assays were performed as follows: mix 1-10 μ L of protein, 790-799 μ L of water and 200 μ L of Bradford reagent (Sigma-Aldrich) in a 1mL cuvette. Vortex and measure in spectrophotometer. To estimate protein concentration, a standard curve using increasing quantities of purified BSA was first obtained. Calculations were done taking into account the slope and initial point of the standard curve, as well as protein dilution in the cuvette. To avoid protein measurement biases, all protein concentrations in this work were fine tuned by direct visualization on gel.

SDS-PAGE and Coomassie Blue revelation

The desired quantity of purified protein was mixed with 10 μ L of protein Loading Dye 4X, boiled 5 min at 98°C and loaded into a 15-well precast NuPAGE 4-12% Bis-Tris (ThermoFisher) gel. The gels migrated in MOPS 1X at 60 V - 80 V for about 2 hours. Gels were then washed with water and incubated with 15 mL of Quick Coomassie Stain (ProteinArk) for at least 2 h, then washed with milli Q water for at least 20 min before scanning.

2) Biochemistry

a) Protein expression and purification

The following protocol is based on Fiorini et al., 2012.

Protein overexpression

BL21(DE3)-CodonPlus competent *Escherichia coli* bacteria (100 μ L, Agilent) were transformed with 500 ng of plasmid of interest by 2 min at 42°C heat shock. After addition of 1 mL liquid LB media and 1 hour incubation at 37°C, cells were pelleted by 4 min centrifugation at 3500 g. The pellet was resuspended in 50 μ L LB media, plated in antibiotic containing LB plates and left to grow overnight at 37°C. 2-3 colonies were resuspended in 25 mL LB liquid media containing antibiotic in a 250 mL erlenmeyer and incubated at 37°C, 200 rpm for 4 hours. The bacteria were inoculated in 1 L of LBAI or TBAI containing metals (Formedium) added with the plasmid antibiotic resistance and chloramphenicol in a 5 L erlenmeyer and incubated at 18°C, 200 rpm overnight.

Other culture media may be used depending on the expressed proteins. For non-inducible media, induction is performed by addition of 0.2-0.5 mM of IPTG when optic density reaches the exponential phase. Although I highly recommend the use of auto-inducible media (Luria Broth or Terrific Broth), with or without metals, protein yields may vary upon culture conditions.

Bacterial cultures were pelleted by 10 min, 4000 g centrifugation at 4°C. Pellets can be stored up to a week at -20°C. However I recommend using fresh cultures for purification.

First protein purification (Example of CBP-tagged proteins)

Every following step is performed at 4°C. The pellet was resuspended in 25 mL of Lysis buffer (1.5X PBS, 1 mM MgAc2, 0.1% NP-40, 20 mM imidazole, 10% glycerol, 2 mM DTT). 1X of protease inhibitors (Apoprotinine, Leupeptine, PMSF, Pepstatin) were added before sonication in a metal beaker at 30% amplitude and 1 second pulse for 4 min. In order to keep low temperatures the beaker was placed on a mix of dry ice and ice. The bacterial debris was pelleted by centrifugation for 30 min at 18 000 rpm.

The supernatant was retrieved and mixed in a Falcon tube with 1 mL of Nickel-NTA Agarose slurry (Qiagen) equilibrated with Lysis Buffer. The sample was incubated in a rotator for 1 hour at 4°C before centrifugation at 4 000 rpm for 5 min at 4°C to pellet the beads. The supernatant (Flow Through) was removed and the beads were washed with 25mL of Lysis Buffer. Beads were transfered into a Poly-Prep chromatography column (Bio-Rad) pre-equilibrated with 1 mL of Lysis Buffer. The column was washed with 5 mL of Lysis Buffer then with 5 mL of Wash Buffer (1.5X PBS, 250 mM NaCl, 1 mM MgAc2, 0.1% NP-40, 50 mM imidazole, 10% glycerol, 2 mM DTT) then again with 5 mL of Lysis Buffer.

The column was closed and the resin was incubated with 1 mL of Elution Buffer (1.5X PBS, 1 mM MgAc2, 0.1% NP-40, 250 mM imidazole, 10% glycerol, 2 mM DTT) for 10 minutes at 4°C. Elution was performed successively with 1 mL of Elution Buffer and protein concentrations were measured by Bradford Assay. The fractions with highest concentration were pooled together and dialysed overnight against 1 L of Calmodulin Binding Buffer

(CBB) (10mM Tris-HCl pH 7.5, 250 mM NaCl, 1 mM MgAc2, 4 mM CaCl2, 0.05% NP-40, 10% glycerol, 2 mM DTT) at 4°C.

Second protein purification

The dialysed fractions were retrieved and mixed with 500 μ L of Calmodulin Sepharose 4B resin (50% Slurry, GE Healthcare) washed with CBB. The samples were incubated in a rotor for 2 hours at 4°C then transferred into a Poly-Prep chromatography column (Bio-Rad) pre-equilibrated with 1 mL of CBB. Wash with 5 mL of CBB. The column was closed and the resin was incubated with 500 μ L of Calmodulin Elution Buffer (10 mM Tris-HCl pH 7.5, 250 mM NaCl, 1 mM MgAc2, 0.05% NP-40, 20 mM EGTA, 10% glycerol, 2 mM DTT) for 10 min at 4°C. Elution was performed successively with 500 μ L of Calmodulin Elution Buffer and protein concentrations were measured by Bradford Assay. The highest concentrated fractions were pooled together and dialysed overnight against 1 L of Storage Buffer B1.5 (1.5X PBS, 1 mM MgAc2, 10% glycerol, 2 mM DTT). The dialysed proteins were aliquoted in small volumes (10-100 μ L, depending on the final protein concentration) and stored at -80°C.

For each step, 10µL may be kept for loading a 10% SDS-PAGE gel, revealed by Coomassie Blue coloring in order to visualize the different purification steps. Reducing agents must be added to the buffers extemporaneously. Buffers may have to be adapted for each purified protein depending on their stability. Some proteins may need more or less salt, no glycerol, or different pH. If the protein precipitates during dialysis, centrifuge and measure the concentration again. For higher reproducibility during experiments, aliquots should be only thawed once before each experiment. Leftovers should be discarded after thawing.

b) Pulldown assays

Affinity and RNA-pulldown assays

The proteins of interest were mixed together in a Protein LoBind tube (Eppendorf) and completed to 6X μ L (usually 30 μ L) with B1.5. 1/6th μ L was retrieved from each tube for Input deposition. For RNA free pulldowns, 1/6th μ L of H₂O was added. For RNA-pulldown assays or affinity assays in presence of RNA, 2 μ L of 10mM 3' biotinylated 30 mer RNA and 2 μ L of 50 μ M ADPNP were added and the rest was completed to 1/6th μ L with H₂O. The samples were incubated for 20 minutes at 30°C. For CBP precipitation, 15 μ L of Calmodulin Sepharose 4B resin (50% Slurry, GE Healthcare) pre-blocked beads were added. For RNA-precipitation, 5 μ L of preblocked Streptavidine Dynabeads MyOne (Thermofisher) were added. The samples were completed with 1X BB 120/10 to make a total volume of 200 μ L and incubated for 1 hour at 4°C with rotation. The beads were washed three times with 800 μ L of 1X BB 125/10, 1X BB 250/10, or 1X BB 500/10. After each wash, samples were rotated for 2 minutes at 4°C, centrifuged for 30 seconds at 2000 rpm and 4°C. The supernatant was removed using a double (P1000+P10) tip, while leaving the beads. For elution 10 μ L of LD 4X were added, and loaded into a 15-well 4-16% acrylamide SDS-page gel.

Pre-blocking of beads

1 mL of beads were centrifuged for 1 minute at 3000 rpm and 4°Cand washed with 900 μ L of Blocking Buffer (20 mM Hepes pH 7.5, 150 mM NaCl, 0.1% NP-40). The beads were then resuspend in 500 μ L of Blocking Buffer with 40 μ L of 5M NaCl, 2.5 μ L of Glycogen Carrier (2 γ), 5 μ L of tRNA yeast (10 γ), 50 μ L of BSA (10 γ). Beads were incubated for at least 2 hours at 4°C with rotation before being washed three times with 900 μ L of Blocking Buffer at 4°C. The pre-blocked beads were resuspended in 500 μ L of 1X BB 250/10 and kept at 4°C.

3) Enzymology

All proteins for enzymatic activity tests were dialysed in their corresponding buffer, stored in small aliquots (<10 μ L), *fast frozen*^{*} and stored at -80°C. Only one thawing cycle per protein can be used. *Fast freezing was not systematically performed.

a) ATPase test



Figure 36: ATPase assay

The following protocol is based on Fiorini et al., 2012. Proteins (dialysed in B1.5) were thawed on ice and diluted in ATPase buffer 1X (50 mM MES pH 6.5, 50 mM KAc, 5 mM MgAc₂; 2 mM DTT, 0.1 mg/mL BSA). For each sample, Upf1 (0,5-1 pmol) was mixed with its partner (5-10 pmol) in a final volume of 8 μ L completed with B1.5 and incubated for 1 hour at 4°C. For each reaction, 12 μ L of ATP premix (0.1 μ L of ATP γ^{32} P (800 Ci/mmol, Perkin Elmer), ATPase Buffer 1X, 2 mM ATP, 0,4 mg/mL RNA polyU) was added and the reaction tubes are sequentially placed at 30°C every 30 s. For each time point, 4 μ L were collected and quenched in 400 μ L of Charcoal mix (10% activated charcoal (Sigma: C5510); 10 mM EDTA pH 8.0), vortexed and kept on ice. At the end of the time-course, all quenched reactions were centrifuged 15 minutes at 14 000 rpm, 4°C. 140 μ L of supernatant were transferred to fresh tubes and scintillation was counted using the Cerenkov method (Figure 36).

b) Decapping test

Preliminary preparations were performed for the decapping test: complete decapping buffer (CDB): 199 μ L DB10X (100 mM Tris HCl pH8, 1 M KCl, 20 mM MgCl2) + 1 μ L MnCl2 1 M. (This step is done extemporaneously to avoid manganese precipitation), n x 4 tubes for timepoints (RT) and (n) reactions were labeled and placed on ice and a mixture of dry ice + ethanol 100% with a rack inside was prepared. For each sample (n), 0.2 μ M of decapping proteins, 0.2-2 μ M of partner were mixed together and the sample was completed with the Dialysis Buffer (van Dijk et al., 2002) to 10 μ L. Samples were incubated

1 h on ice. The RNA premix for n+1 samples was prepared as follows: for 1 sample 2 μ M capped RNA, 2.25 μ L CDB 10X and add H₂O to a final volume of 12.5 μ L 12.5 μ L of RNA mix were added to each sample. Samples were mixed, centrifuged and 5 µL were retrieved for timepoint 0'. The retrieved sample was frozen in the ethanol/dry ice mixture and the reaction tube was placed at 30°C. For timepoint, 5 μ L were retrieved and frozen. When all time points were finished, all tubes were incubated at 65°C for 10 min to inactivate decapping enzymes then placed 5 min on ice to cool down. The Xrn1 premix was prepared for n+1 samples (for 1 sample 0.2 μ L Xrn1 (NEB) + 0.5 μ L CDB 10X + 4.3 μ L H₂O; Vf = 5 μ L) and 5 μ L Xrn1 premix were added to the samples. Samples were incubated 1 h at 37°C. In a 96-well white skirted low volume plate, 10 µL Reactive 1 were placed in each well using an automatic pipet. When Xrn1 incubation was finished, samples were transferred to the plate which was shaken for 4 min at 600 rpm, 25°C, in a block adapted to a 96-well plate. The plate was centrifuged for 1 min and incubated for 1 h at 25°C. The mix of Reaction 2 for (n+1) samples was prepared as follows: for 1 sample mix 0.2 μ L Reactive 2 and 19.8 μ L Kinase Glo. When the incubation was finished, 20 μ L of Reaction 2 mix were added in each well using an automatic pipet. The plate was shaken for 4 min, incubated for 1 h in the 25° C block. Values were read in a Berthold machine using the program for kinetic measurement during 0.1 s.

4) Biophysics

All the methods for magnetic tweezers experiments have been published in Methods in Enzymology (Ruiz-Gutierrez et al., 2022) and are available in the Annex.

Bibliography

Bibliography

- Abbondanzieri, E. A., Greenleaf, W. J., Shaevitz, J. W., Landick, R., & Block, S. M. (2005). Direct observation of base-pair stepping by RNA polymerase. *Nature*, 438(7067), Article 7067. https://doi.org/10.1038/nature04268
- Addinall, S. G., Holstein, E.-M., Lawless, C., Yu, M., Chapman, K., Banks, A. P., Ngo, H.-P., Maringele, L., Taschuk, M., Young, A., Ciesiolka, A., Lister, A. L., Wipat, A., Wilkinson, D. J., & Lydall, D. (2011). Quantitative Fitness Analysis Shows That NMD Proteins and Many Other Protein Complexes Suppress or Enhance Distinct Telomere Cap Defects. *PLOS Genetics*, 7(4), e1001362. https://doi.org/10.1371/journal.pgen.1001362
- Addington, A. M., Gauthier, J., Piton, A., Hamdan, F. F., Raymond, A., Gogtay, N., Miller, R., Tossell, J., Bakalar, J., Inoff-Germain, G., Gochman, P., Long, R., Rapoport, J. L., & Rouleau, G. A. (2011). A novel frameshift mutation in UPF3B identified in brothers affected with childhood onset schizophrenia and autism spectrum disorders. *Molecular Psychiatry*, 16(3), 238-239. https://doi.org/10.1038/mp.2010.59
- Adivarahan, S., Livingston, N., Nicholson, B., Rahman, S., Wu, B., Rissland, O. S., & Zenklusen, D. (2018). Spatial Organization of Single mRNPs at Different Stages of the Gene Expression Pathway. *Molecular Cell*, 72(4), 727-738.e5. https://doi.org/10.1016/j.molcel.2018.10.010
- Alhusaini, N., & Coller, J. (2016). The deadenylase components Not2p, Not3p, and Not5p promote mRNA decapping. RNA (New York, N.Y.), 22(5), 709-721. https://doi.org/10.1261/rna.054742.115
- Alkalaeva, E. Z., Pisarev, A. V., Frolova, L. Y., Kisselev, L. L., & Pestova, T. V. (2006). In vitro reconstitution of eukaryotic translation reveals cooperativity between release factors eRF1 and eRF3. Cell, 125(6), 1125-1136. https://doi.org/10.1016/j.cell.2006.04.035
- Allemand, J. F., Bensimon, D., Lavery, R., & Croquette, V. (1998). Stretched and overwound DNA forms a Pauling-like structure with exposed bases. *Proceedings of the National Academy of Sciences of the United States of America*, 95(24), 14152-14157. https://doi.org/10.1073/pnas.95.24.14152
- Altamura, N., Groudinsky, O., Dujardin, G., & Slonimski, P. P. (1992). NAM7 nuclear gene encodes a novel member of a family of helicases with a Zn-ligand motif and is involved in mitochondrial functions in Saccharomyces cerevisiae. *Journal of Molecular Biology*, 224(3), 575-587. https://doi.org/10.1016/0022-2836(92)90545-U
- Amrani, N., Dong, S., He, F., Ganesan, R., Ghosh, S., Kervestin, S., Li, C., Mangus, D. A., Spatrick, P., & Jacobson, A. (2006). Aberrant termination triggers nonsense-mediated mRNA decay. *Biochemical Society Transactions*, 34(Pt 1), 39-42. https://doi.org/10.1042/BST20060039
- 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(7013), 112-118. https://doi.org/10.1038/nature03060
- Anders, K. R., Grimson, A., & Anderson, P. (2003). SMG-5, required for C.elegans nonsense-mediated mRNA decay, associates with SMG-2 and protein phosphatase 2A. *The EMBO Journal*, 22(3), 641-650. https://doi.org/10.1093/emboj/cdg056
- Andersen, C. B. F., Ballut, L., Johansen, J. S., Chamieh, H., Nielsen, K. H., Oliveira, C. L. P., Pedersen, J. S., Séraphin, B., Le Hir, H., & Andersen, G. R. (2006). Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA. *Science (New York, N.Y.)*, 313(5795), 1968-1972. https://doi.org/10.1126/science.1131981
- Antonarakis, S. E., & Cooper, D. N. (2013). Chapter 7 Human Gene Mutation in Inherited Disease: Molecular Mechanisms and Clinical Consequences. In D. Rimoin, R. Pyeritz, & B. Korf (Éds.), Emery and Rimoin's Principles and Practice of Medical Genetics (Sixth Edition) (p. 1-48). Academic Press. https://doi.org/10.1016/B978-0-12-383834-6.00007-0
- Applequist, S. E., Selg, M., Raman, C., & Jäck, H. M. (1997). Cloning and characterization of HUPF1, a human homolog of the Saccharomyces cerevisiae nonsense mRNA-reducing UPF1 protein. *Nucleic Acids Research*, 25(4), 814-821. https://doi.org/10.1093/nar/25.4.814
- Arias-Palomo, E., Yamashita, A., Fernández, I. S., Núñez-Ramírez, R., Bamba, Y., Izumi, N., Ohno, S., & Llorca, O. (2011). The nonsense-mediated mRNA decay SMG-1 kinase is regulated by large-scale conformational changes controlled by SMG-8. *Genes & Development*, 25(2), 153-164. https://doi.org/10.1101/gad.606911
- Arribas-Layton, M., Wu, D., Lykke-Andersen, J., & Song, H. (2013). Structural and functional control of the eukaryotic mRNA decapping machinery. *Biochimica et Biophysica Acta (BBA) Gene Regulatory*

Mechanisms, 1829(6-7), 580-589. https://doi.org/10.1016/j.bbagrm.2012.12.006

- Athapattu, U. S., Amarasekara, C. A., Immel, J. R., Bloom, S., Barany, F., Nagel, A. C., & Soper, S. A. (2021). Solid-phase XRN1 reactions for RNA cleavage: Application in single-molecule sequencing. *Nucleic Acids Research*, 49(7), e41. https://doi.org/10.1093/nar/gkab001
- Atkin, A. L., Schenkman, L. R., Eastham, M., Dahlseid, J. N., Lelivelt, M. J., & Culbertson, M. R. (1997). Relationship between Yeast Polyribosomes and Upf Proteins Required for Nonsense mRNA Decay*. *Journal of Biological Chemistry*, 272(35), 22163-22172. https://doi.org/10.1074/jbc.272.35.22163
- Audebert, L., Feuerbach, F., Decourty, L., Namane, A., Permal, E., Badis, G., & Saveanu, C. (2023). Deadenylation rate is not a major determinant of RNA degradation in yeast (p. 2023.01.16.524186). bioRxiv. https://doi.org/10.1101/2023.01.16.524186
- Avery, P., Vicente-Crespo, M., Francis, D., Nashchekina, O., Alonso, C. R., & Palacios, I. M. (2011). Drosophila Upf1 and Upf2 loss of function inhibits cell growth and causes animal death in a Upf3-independent manner. RNA (New York, N.Y.), 17(4), 624-638. https://doi.org/10.1261/rna.2404211
- Azzalin, C. M., & Lingner, J. (2006). The human RNA surveillance factor UPF1 is required for S phase progression and genome stability. *Current Biology: CB*, 16(4), 433-439. https://doi.org/10.1016/j.cub.2006.01.018
- Azzalin, C. M., Reichenbach, P., Khoriauli, L., Giulotto, E., & Lingner, J. (2007). Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science (New York, N.Y.)*, 318(5851), 798-801. https://doi.org/10.1126/science.1147182
- Badis, G., Saveanu, C., Fromont-Racine, M., & Jacquier, A. (2004). Targeted mRNA degradation by deadenylation-independent decapping. *Molecular Cell*, 15(1), 5-15. https://doi.org/10.1016/j.molcel.2004.06.028
- Baker, S. L., & Hogg, J. R. (2017). A system for coordinated analysis of translational readthrough and nonsense-mediated mRNA decay. *PloS One*, 12(3), e0173980. https://doi.org/10.1371/journal.pone.0173980
- Balistreri, G., Bognanni, C., & Mühlemann, O. (2017). Virus Escape and Manipulation of Cellular Nonsense-Mediated mRNA Decay. *Viruses*, 9(1), 24. https://doi.org/10.3390/v9010024
- Balistreri, G., Horvath, P., Schweingruber, C., Zünd, D., McInerney, G., Merits, A., Mühlemann, O., Azzalin, C., & Helenius, A. (2014). The host nonsense-mediated mRNA decay pathway restricts Mammalian RNA virus replication. *Cell Host & Microbe*, 16(3), 403-411. https://doi.org/10.1016/j.chom.2014.08.007
- Ball, L. J., Jarchau, T., Oschkinat, H., & Walter, U. (2002). EVH1 domains: Structure, function and interactions. FEBS Letters, 513(1), 45-52. https://doi.org/10.1016/S0014-5793(01)03291-4
- Ballut, L., Marchadier, B., Baguet, A., Tomasetto, C., Séraphin, B., & Le Hir, H. (2005). The exon junction core complex is locked onto RNA by inhibition of elF4AIII ATPase activity. *Nature Structural & Molecular Biology*, 12(10), 861-869. https://doi.org/10.1038/nsmb990
- Barbosa, I., Haque, N., Fiorini, F., Barrandon, C., Tomasetto, C., Blanchette, M., & Le Hir, H. (2012). Human CWC22 escorts the helicase eIF4AIII to spliceosomes and promotes exon junction complex assembly. *Nature Structural & Molecular Biology*, 19(10), 983-990. https://doi.org/10.1038/nsmb.2380
- 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(2), 207-218. https://doi.org/10.1016/j.molcel.2012.08.014
- Becker, T., Franckenberg, S., Wickles, S., Shoemaker, C. J., Anger, A. M., Armache, J.-P., Sieber, H., Ungewickell, C., Berninghausen, O., Daberkow, I., Karcher, A., Thomm, M., Hopfner, K.-P., Green, R., & Beckmann, R. (2012). Structural basis of highly conserved ribosome recycling in eukaryotes and archaea. *Nature*, 482(7386), 501-506. https://doi.org/10.1038/nature10829
- Behm-Ansmant, I., Kashima, I., Rehwinkel, J., Saulière, J., Wittkopp, N., & Izaurralde, E. (2007). mRNA quality control: An ancient machinery recognizes and degrades mRNAs with nonsense codons. FEBS Letters, 581(15), 2845-2853. https://doi.org/10.1016/j.febslet.2007.05.027
- Behm-Ansmant, I., Rehwinkel, J., & Izaurralde, E. (2006). MicroRNAs silence gene expression by repressing protein expression and/or by promoting mRNA decay. Cold Spring Harbor Symposia on Quantitative Biology, 71, 523-530. https://doi.org/10.1101/sqb.2006.71.013
- Bertram, G., Bell, H. A., Ritchie, D. W., Fullerton, G., & Stansfield, I. (2000). Terminating eukaryote translation: Domain 1 of release factor eRF1 functions in stop codon recognition. *RNA*, 6(9), 1236-1247.
- Bhattacharya, A., Czaplinski, K., Trifillis, P., He, F., Jacobson, A., & Peltz, S. W. (2000). Characterization of the biochemical properties of the human Upf1 gene product that is involved in nonsense-mediated mRNA decay. RNA, 6(9), 1226-1235.
- Bhattacharyya, S. N., Habermacher, R., Martine, U., Closs, E. I., & Filipowicz, W. (2006). Relief of

microRNA-mediated translational repression in human cells subjected to stress. *Cell*, 125(6), 1111-1124. https://doi.org/10.1016/j.cell.2006.04.031

- Bianco, P. R. (2021). Insight into the biochemical mechanism of DNA helicases provided by bulk-phase and single-molecule assays. *Methods*. https://doi.org/10.1016/j.ymeth.2021.12.002
- Blanchet, S., Rowe, M., Von der Haar, T., Fabret, C., Demais, S., Howard, M. J., & Namy, O. (2015). New insights into stop codon recognition by eRF1. *Nucleic Acids Research*, 43(6), 3298-3308. https://doi.org/10.1093/nar/gkv154
- Boehm, V., Kueckelmann, S., Gerbracht, J. V., Kallabis, S., Britto-Borges, T., Altmüller, J., Krüger, M., Dieterich, C., & Gehring, N. H. (2021). SMG5-SMG7 authorize nonsense-mediated mRNA decay by enabling SMG6 endonucleolytic activity. *Nature Communications*, 12(1), 3965. https://doi.org/10.1038/s41467-021-24046-3
- Boisramé, A., Devillers, H., Onésime, D., Brunel, F., Pouch, J., Piot, M., & Neuvéglise, C. (2019). Exon junction complex components Y14 and Mago still play a role in budding yeast. *Scientific Reports*, *9*(1), 849. https://doi.org/10.1038/s41598-018-36785-3
- Bonneau, F., Basquin, J., Steigenberger, B., Schäfer, T., Schäfer, I. B., & Conti, E. (2023). Nuclear mRNPs are compact particles packaged with a network of proteins promoting RNA-RNA interactions. *Genes & Development*, 37(11-12), 505-517. https://doi.org/10.1101/gad.350630.123
- Bono, F., Ebert, J., Lorentzen, E., & Conti, E. (2006). The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA. *Cell*, *126*(4), 713-725. https://doi.org/10.1016/j.cell.2006.08.006
- Boo, S. H., Ha, H., Lee, Y., Shin, M.-K., Lee, S., & Kim, Y. K. (2022). UPF1 promotes rapid degradation of m6A-containing RNAs. *Cell Reports*, 39(8), 110861. https://doi.org/10.1016/j.celrep.2022.110861
- Bousquet-Antonelli, C., Presutti, C., & Tollervey, D. (2000). Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell*, 102(6), 765-775. https://doi.org/10.1016/s0092-8674(00)00065-9
- Bouveret, E., Rigaut, G., Shevchenko, A., Wilm, M., & Séraphin, B. (2000). A Sm-like protein complex that participates in mRNA degradation. *The EMBO Journal*, 19(7), 1661-1671. https://doi.org/10.1093/emboj/19.7.1661
- Braun, J. E., Truffault, V., Boland, A., Huntzinger, E., Chang, C.-T., Haas, G., Weichenrieder, O., Coles, M., & Izaurralde, E. (2012). A direct interaction between DCP1 and XRN1 couples mRNA decapping to 5' exonucleolytic degradation. *Nature Structural & Molecular Biology*, 19(12), 1324-1331. https://doi.org/10.1038/nsmb.2413
- Brengues, M., Teixeira, D., & Parker, R. (2005). Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies. *Science (New York, N.Y.)*, 310(5747), 486-489. https://doi.org/10.1126/science.1115791
- Bruno, I. G., Karam, R., Huang, L., Bhardwaj, A., Lou, C. H., Shum, E. Y., Song, H.-W., Corbett, M. A., Gifford, W. D., Gecz, J., Pfaff, S. L., & Wilkinson, M. F. (2011). Identification of a microRNA that activates gene expression by repressing nonsense-mediated RNA decay. *Molecular Cell*, 42(4), 500-510. https://doi.org/10.1016/j.molcel.2011.04.018
- Buchwald, G., Ebert, J., Basquin, C., Sauliere, J., Jayachandran, U., Bono, F., Le Hir, H., & Conti, E. (2010). Insights into the recruitment of the NMD machinery from the crystal structure of a core EJC-UPF3b complex. *Proceedings of the National Academy of Sciences of the United States of America*, 107(22), 10050-10055. https://doi.org/10.1073/pnas.1000993107
- Bühler, M., Steiner, S., Mohn, F., Paillusson, A., & Mühlemann, O. (2006). EJC-independent degradation of nonsense immunoglobulin-mu mRNA depends on 3' UTR length. Nature Structural & Molecular Biology, 13(5), 462-464. https://doi.org/10.1038/nsmb1081
- Burkard, K. T. D., & Butler, J. S. (2000). A Nuclear 3'-5' Exonuclease Involved in mRNA Degradation Interacts with Poly(A) Polymerase and the hnRNA Protein Npl3p. *Molecular and Cellular Biology*, 20(2), 604-616.
- Burke, J. F., & Mogg, A. E. (1985). Suppression of a nonsense mutation in mammalian cells in vivo by the aminoglycoside antibiotics G-418 and paromomycin. *Nucleic Acids Research*, 13(17), 6265-6272. https://doi.org/10.1093/nar/13.17.6265
- Busetto, V., Barbosa, I., Basquin, J., Marquenet, É., Hocq, R., Hennion, M., Paternina, J. A., Namane, A., Conti, E., Bensaude, O., & Le Hir, H. (2020). Structural and functional insights into CWC27/CWC22 heterodimer linking the exon junction complex to spliceosomes. *Nucleic Acids Research*, 48(10), 5670-5683. https://doi.org/10.1093/nar/gkaa267
- Bustamante, C., Smith, S. B., Liphardt, J., & Smith, D. (2000). Single-molecule studies of DNA mechanics.

 Current
 Opinion
 in
 Structural
 Biology,
 10(3),
 279-285.

 https://doi.org/10.1016/s0959-440x(00)00085-3
 10(3),
 279-285.
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),
 10(3),

- Carter, M. S., Doskow, J., Morris, P., Li, S., Nhim, R. P., Sandstedt, S., & Wilkinson, M. F. (1995). A regulatory mechanism that detects premature nonsense codons in T-cell receptor transcripts in vivo is reversed by protein synthesis inhibitors in vitro. *The Journal of Biological Chemistry*, 270(48), 28995-29003. https://doi.org/10.1074/jbc.270.48.28995
- Causier, B., Li, Z., De Smet, R., Lloyd, J. P. B., Van de Peer, Y., & Davies, B. (2017). Conservation of Nonsense-Mediated mRNA Decay Complex Components Throughout Eukaryotic Evolution. *Scientific Reports*, 7(1), 16692. https://doi.org/10.1038/s41598-017-16942-w
- Celik, A., Baker, R., He, F., & Jacobson, A. (2017). High-resolution profiling of NMD targets in yeast reveals translational fidelity as a basis for substrate selection. *RNA (New York, N.Y.)*, 23(5), 735-748. https://doi.org/10.1261/rna.060541.116
- Chakrabarti, S., Bonneau, F., Schüssler, S., Eppinger, E., & Conti, E. (2014). Phospho-dependent and phospho-independent interactions of the helicase UPF1 with the NMD factors SMG5-SMG7 and SMG6. *Nucleic Acids Research*, *42*(14), 9447-9460. https://doi.org/10.1093/nar/gku578
- Chakrabarti, S., Jayachandran, U., Bonneau, F., Fiorini, F., Basquin, C., Domcke, S., Le Hir, H., & Conti, E. (2011). Molecular Mechanisms for the RNA-Dependent ATPase Activity of Upf1 and Its Regulation by Upf2. *Molecular Cell*, 41(6), 693-703. https://doi.org/10.1016/j.molcel.2011.02.010
- Chamieh, H., Ballut, L., Bonneau, F., & Hir, H. L. (2008). NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. *Nature Structural & Molecular Biology*, 15(1), 85-93. https://doi.org/10.1038/nsmb1330
- Chan, W.-K., Huang, L., Gudikote, J. P., Chang, Y.-F., Imam, J. S., MacLean, J. A., & Wilkinson, M. F. (2007). An alternative branch of the nonsense-mediated decay pathway. *The EMBO Journal*, 26(7), 1820-1830. https://doi.org/10.1038/sj.emboj.7601628
- Chang, J. C., & Kan, Y. W. (1979). Beta 0 thalassemia, a nonsense mutation in man. Proceedings of the National Academy of Sciences of the United States of America, 76(6), 2886-2889. https://doi.org/10.1073/pnas.76.6.2886
- Chapman, J. H., Craig, J. M., Wang, C. D., Gundlach, J. H., Neuman, K. C., & Hogg, J. R. (2022). UPF1 mutants with intact ATPase but deficient helicase activities promote efficient nonsense-mediated mRNA decay. *Nucleic Acids Research*, 50(20), 11876-11894. https://doi.org/10.1093/nar/gkac1026
- Charenton, C., & Graille, M. (2018). mRNA decapping: Finding the right structures. *Philosophical Transactions of* the Royal Society B: Biological Sciences, 373(1762), 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. *Nature Structural & Molecular Biology*, 23(11), 982-986. https://doi.org/10.1038/nsmb.3300
- Chawla, R., Redon, S., Raftopoulou, C., Wischnewski, H., Gagos, S., & Azzalin, C. M. (2011). Human UPF1 interacts with TPP1 and telomerase and sustains telomere leading-strand replication. *The EMBO Journal*, 30(19), 4047-4058. https://doi.org/10.1038/emboj.2011.280
- Chazal, P.-E., Daguenet, E., Wendling, C., Ulryck, N., Tomasetto, C., Sargueil, B., & Le Hir, H. (2013). EJC core component MLN51 interacts with eIF3 and activates translation. *Proceedings of the National Academy* of Sciences of the United States of America, 110(15), 5903-5908. https://doi.org/10.1073/pnas.1218732110
- Chen, B.-L., Wang, H.-M., Lin, X.-S., & Zeng, Y.-M. (2021). UPF1: A potential biomarker in human cancers. *Frontiers in Bioscience (Landmark Edition)*, 26(5), 76-84. https://doi.org/10.52586/4925
- Chen, Y.-H., Su, L.-H., & Sun, C.-H. (2008). Incomplete nonsense-mediated mRNA decay in Giardia lamblia. International Journal for Parasitology, 38(11), 1305-1317. https://doi.org/10.1016/j.ijpara.2008.02.006
- Chen, Z., Smith, K. R., Batterham, P., & Robin, C. (2005). Smg1 nonsense mutations do not abolish nonsense-mediated mRNA decay in Drosophila melanogaster. *Genetics*, 171(1), 403-406. https://doi.org/10.1534/genetics.105.045674
- Cheng, W., Arunajadai, S. G., Moffitt, J. R., Tinoco, I., & Bustamante, C. (2011). Single–Base Pair Unwinding and Asynchronous RNA Release by the Hepatitis C Virus NS3 Helicase. *Science*, 333(6050), 1746-1749. https://doi.org/10.1126/science.1206023
- Cheng, Z., Muhlrad, D., Lim, M. K., Parker, R., & Song, H. (2007). Structural and functional insights into the human Upf1 helicase core. *The EMBO Journal*, 26(1), 253-264. https://doi.org/10.1038/sj.emboj.7601464

- Cheng, Z., Saito, K., Pisarev, A. V., Wada, M., Pisareva, V. P., Pestova, T. V., Gajda, M., Round, A., Kong, C., Lim, M., Nakamura, Y., Svergun, D. I., Ito, K., & Song, H. (2009). Structural insights into eRF3 and stop codon recognition by eRF1. Genes & Development, 23(9), 1106-1118. https://doi.org/10.1101/gad.1770109
- Cho, H., Kim, K. M., & Kim, Y. K. (2009). Human Proline-Rich Nuclear Receptor Coregulatory Protein 2 Mediates an Interaction between mRNA Surveillance Machinery and Decapping Complex. *Molecular Cell*, 33(1), 75-86. https://doi.org/10.1016/j.molcel.2008.11.022
- Cho, H., Park, O. H., Park, J., Ryu, I., Kim, J., Ko, J., & Kim, Y. K. (2015). Glucocorticoid receptor interacts with PNRC2 in a ligand-dependent manner to recruit UPF1 for rapid mRNA degradation. Proceedings of the National Academy of Sciences of the United States of America, 112(13), E1540-E1549. https://doi.org/10.1073/pnas.1409612112
- Choe, J., Ahn, S. H., & Kim, Y. K. (2014). The mRNP remodeling mediated by UPF1 promotes rapid degradation of replication-dependent histone mRNA. *Nucleic Acids Research*, 42(14), 9334-9349. https://doi.org/10.1093/nar/gku610
- Clerici, M., Deniaud, A., Boehm, V., Gehring, N. H., Schaffitzel, C., & Cusack, S. (2014). Structural and functional analysis of the three MIF4G domains of nonsense-mediated decay factor UPF2. *Nucleic Acids Research*, 42(4), 2673-2686. https://doi.org/10.1093/nar/gkt1197
- Clerici, M., Mourão, A., Gutsche, I., Gehring, N. H., Hentze, M. W., Kulozik, A., Kadlec, J., Sattler, M., & Cusack, S. (2009). Unusual bipartite mode of interaction between the nonsense-mediated decay factors, UPF1 and UPF2. *The EMBO Journal*, 28(15), 2293-2306. https://doi.org/10.1038/emboj.2009.175
- Colak, D., Ji, S.-J., Porse, B. T., & Jaffrey, S. R. (2013). Regulation of axon guidance by compartmentalized nonsense-mediated mRNA decay. *Cell*, 153(6), 1252-1265. https://doi.org/10.1016/j.cell.2013.04.056
- Collart, M. A., & Panasenko, O. O. (2017). The Ccr4-Not Complex: Architecture and Structural Insights. In J. R. Harris & J. Marles-Wright (Éds.), *Macromolecular Protein Complexes: Structure and Function* (p. 349-379). Springer International Publishing. https://doi.org/10.1007/978-3-319-46503-6_13
- Colombo, M., Karousis, E. D., Bourquin, J., Bruggmann, R., & Mühlemann, O. (2017). Transcriptome-wide identification of NMD-targeted human mRNAs reveals extensive redundancy between SMG6- and SMG7-mediated degradation pathways. RNA (New York, N.Y.), 23(2), 189-201. https://doi.org/10.1261/rna.059055.116
- Comstock, M. J., Whitley, K. D., Jia, H., Sokoloski, J., Lohman, T. M., Ha, T., & Chemla, Y. R. (2015). Direct observation of structure-function relationship in a nucleic acid processing enzyme. *Science (New York, N.Y.)*, 348(6232), 352-354. https://doi.org/10.1126/science.aaa0130
- Contu, L., Balistreri, G., Domanski, M., Uldry, A.-C., & Mühlemann, O. (2021). Characterisation of the Semliki Forest Virus-host cell interactome reveals the viral capsid protein as an inhibitor of nonsense-mediated mRNA decay. *PLoS Pathogens*, 17(5), e1009603. https://doi.org/10.1371/journal.ppat.1009603
- Courel, M., Clément, Y., Bossevain, C., Foretek, D., Vidal Cruchez, O., Yi, Z., Bénard, M., Benassy, M., Kress, M., Vindry, C., Ernoult-Lange, M., Antoniewski, C., Morillon, A., Brest, P., Hubstenberger, A., Roest Crollius, H., Standart, N., & Weil, D. (2019). GC content shapes mRNA storage and decay in human cells. *ELife*, 8, e49708. https://doi.org/10.7554/eLife.49708
- Cowen, L. E., & Tang, Y. (2017). Identification of nonsense-mediated mRNA decay pathway as a critical regulator of p53 isoform β. *Scientific Reports*, 7(1), Article 1. https://doi.org/10.1038/s41598-017-17283-4
- 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 & Development, 9(4), 423-436. https://doi.org/10.1101/gad.9.4.423
- Culbertson, M. R., & Leeds, P. F. (2003). Looking at mRNA decay pathways through the window of molecular evolution. Current Opinion in Genetics & Development, 13(2), 207-214. https://doi.org/10.1016/S0959-437X(03)00014-5
- Culbertson, M. R., Underbrink, K. M., & Fink, G. R. (1980). Frameshift Suppression in SACCHAROMYCES CEREVISIAE. II. Genetic Properties of Group II Suppressors. *Genetics*, 95(4), 833-853.
- Custódio, N., Carmo-Fonseca, M., Geraghty, F., Pereira, H. S., Grosveld, F., & Antoniou, M. (1999). Inefficient processing impairs release of RNA from the site of transcription. *The EMBO Journal*, 18(10), 2855-2866. https://doi.org/10.1093/emboj/18.10.2855
- Czaplinski, K., Ruiz-Echevarria, M. J., González, C. I., & Peltz, S. W. (1999). Should we kill the messenger? The role of the surveillance complex in translation termination and mRNA turnover. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology,* 21(8), 685-696. https://doi.org/10.1002/(SICI)1521-1878(199908)21:8<685::AID-BIES8>3.0.CO;2-4
- 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 & Development*, 12(11), 1665-1677.

- Czaplinski, K., Weng, Y., Hagan, K. W., & Peltz, S. W. (1995). Purification and characterization of the Upf1 protein: A factor involved in translation and mRNA degradation. *RNA*, 1(6), 610-623.
- Dave, P., Roth, G., Griesbach, E., Mateju, D., Hochstoeger, T., & Chao, J. A. (2023). Single-molecule imaging reveals translation-dependent destabilization of mRNAs. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2023.01.013
- Davis, C. A., & Ares, M. (2006). Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in Saccharomyces cerevisiae. Proceedings of the National Academy of Sciences of the United States of America, 103(9), 3262-3267. https://doi.org/10.1073/pnas.0507783103
- De, S., & Mühlemann, O. (2022). A comprehensive coverage insurance for cells: Revealing links between ribosome collisions, stress responses and mRNA surveillance. RNA Biology, 19(1), 609-621. https://doi.org/10.1080/15476286.2022.2065116
- Decker, C. J., & Parker, R. (1993). A turnover pathway for both stable and unstable mRNAs in yeast: Evidence for a requirement for deadenylation. Genes & Development, 7(8), 1632-1643.
- Decker, C. J., & Parker, R. (2012). P-Bodies and Stress Granules: Possible Roles in the Control of Translation and mRNA Degradation. *Cold Spring Harbor Perspectives in Biology*, 4(9), a012286. https://doi.org/10.1101/cshperspect.a012286
- Decourty, L., Doyen, A., Malabat, C., Frachon, E., Rispal, D., Séraphin, B., Feuerbach, F., Jacquier, A., & Saveanu, C. (2014). Long Open Reading Frame Transcripts Escape Nonsense-Mediated mRNA Decay in Yeast. *Cell Reports*, 6(4), 593-598. https://doi.org/10.1016/j.celrep.2014.01.025
- Defenouillère, Q., & Fromont-Racine, M. (2017). The ribosome-bound quality control complex: From aberrant peptide clearance to proteostasis maintenance. *Current Genetics*, 63(6), 997-1005. https://doi.org/10.1007/s00294-017-0708-5
- Degtiar, E., Fridman, A., Gottlieb, D., Vexler, K., Berezin, I., Farhi, R., Golani, L., & Shaul, O. (2015). The feedback control of UPF3 is crucial for RNA surveillance in plants. *Nucleic Acids Research*, 43(8), 4219-4235. https://doi.org/10.1093/nar/gkv237
- 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. *The EMBO Journal*, 37(21), e99278. https://doi.org/10.15252/embj.201899278
- Delhi, P., Queiroz, R., Inchaustegui, D., Carrington, M., & Clayton, C. (2011). Is There a Classical Nonsense-Mediated Decay Pathway in Trypanosomes? PLOS ONE, 6(9), e25112. https://doi.org/10.1371/journal.pone.0025112
- Denning, G., Jamieson, L., Maquat, L. E., Thompson, E. A., & Fields, A. P. (2001). Cloning of a novel phosphatidylinositol kinase-related kinase: Characterization of the human SMG-1 RNA surveillance protein. The Journal of Biological Chemistry, 276(25), 22709-22714. https://doi.org/10.1074/jbc.C100144200
- Dessinges, M.-N., Lionnet, T., Xi, X. G., Bensimon, D., & Croquette, V. (2004). Single-molecule assay reveals strand switching and enhanced processivity of UvrD. *Proceedings of the National Academy of Sciences*, 101(17), 6439-6444. https://doi.org/10.1073/pnas.0306713101
- DeZwaan, D. C., & Freeman, B. C. (2009). The conserved Est1 protein stimulates telomerase DNA extension activity. Proceedings of the National Academy of Sciences, 106(41), 17337-17342. https://doi.org/10.1073/pnas.0905703106
- Dijk, E. van, Cougot, N., Meyer, S., Babajko, S., Wahle, E., & Séraphin, B. (2002). Human Dcp2: A catalytically active mRNA decapping enzyme located in specific cytoplasmic structures. *The EMBO Journal*, 21(24), 6915-6924. https://doi.org/10.1093/emboj/cdf678
- Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., Xue, C., Marinov, G. K., Khatun, J., Williams, B. A., Zaleski, C., Rozowsky, J., Röder, M., Kokocinski, F., Abdelhamid, R. F., ... Gingeras, T. R. (2012). Landscape of transcription in human cells. *Nature*, 489(7414), Article 7414. https://doi.org/10.1038/nature11233
- Doma, M. K., & Parker, R. (2006). Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature*, 440(7083), 561-564. https://doi.org/10.1038/nature04530
- Doma, M. K., & Parker, R. (2007). RNA Quality Control in Eukaryotes. *Cell*, 131(4), 660-668. https://doi.org/10.1016/j.cell.2007.10.041
- Dong, S., Li, C., Zenklusen, D., Singer, R. H., Jacobson, A., & He, F. (2007). YRA1 autoregulation requires nuclear

export and cytoplasmic Edc3p-mediated degradation of its pre-mRNA. *Molecular Cell*, 25(4), 559-573. https://doi.org/10.1016/j.molcel.2007.01.012

- D'Orazio, K. N., Wu, C. C.-C., Sinha, N., Loll-Krippleber, 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
- Dujardin, G., Pajot, P., Groudinsky, O., & Slonimski, P. P. (1980). Long range control circuits within mitochondria and between nucleus and mitochondria. I. Methodology and phenomenology of suppressors. *Molecular* & General Genetics: MGG, 179(3), 469-482. https://doi.org/10.1007/BF00271736
- Dunckley, T., & Parker, R. (1999). The DCP2 protein is required for mRNA decapping in Saccharomyces cerevisiae and contains a functional MutT motif. *The EMBO Journal*, *18*(19), 5411-5422. https://doi.org/10.1093/emboj/18.19.5411
- Durand, S., Cougot, N., Mahuteau-Betzer, F., Nguyen, C.-H., Grierson, D. S., Bertrand, E., Tazi, J., & Lejeune, F. (2007). Inhibition of nonsense-mediated mRNA decay (NMD) by a new chemical molecule reveals the dynamic of NMD factors in P-bodies. *The Journal of Cell Biology*, 178(7), 1145-1160. https://doi.org/10.1083/jcb.200611086
- Durand, S., Franks, T. M., & Lykke-Andersen, J. (2016). Hyperphosphorylation amplifies UPF1 activity to resolve stalls in nonsense-mediated mRNA decay. *Nature Communications*, 7, 12434. https://doi.org/10.1038/ncomms12434
- Durand, S., & Lykke-Andersen, J. (2013). Nonsense-mediated mRNA decay occurs during eIF4F-dependent translation in human cells. *Nature Structural & Molecular Biology*, 20(6), 702-709. https://doi.org/10.1038/nsmb.2575
- Eberle, A. B., Lykke-Andersen, S., Mühlemann, O., & Jensen, T. H. (2009). SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nature Structural & Molecular Biology*, 16(1), 49-55. https://doi.org/10.1038/nsmb.1530
- 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 Biology*, 6(4), e92. https://doi.org/10.1371/journal.pbio.0060092
- Eliseeva, I. A., Lyabin, D. N., & Ovchinnikov, L. P. (2013). Poly(A)-binding proteins: Structure, domain organization, and activity regulation. *Biochemistry. Biokhimiia*, 78(13), 1377-1391. https://doi.org/10.1134/S0006297913130014
- Eulalio, A., Behm-Ansmant, I., Schweizer, D., & Izaurralde, E. (2007). P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. *Molecular and Cellular Biology*, 27(11), 3970-3981. https://doi.org/10.1128/MCB.00128-07
- Eustice, D. C., & Wilhelm, J. M. (1984). Fidelity of the eukaryotic codon-anticodon interaction: Interference by aminoglycoside antibiotics. *Biochemistry*, 23(7), 1462-1467. https://doi.org/10.1021/bi00302a019
- Fabian, M. R., Frank, F., Rouya, C., Siddiqui, N., Lai, W. S., Karetnikov, A., Blackshear, P. J., Nagar, B., & Sonenberg, N. (2013). Structural basis for the recruitment of the human CCR4-NOT deadenylase complex by tristetraprolin. Nature Structural & Molecular Biology, 20(6), 735-739. https://doi.org/10.1038/nsmb.2572
- Fairman-Williams, M. E., Guenther, U.-P., & Jankowsky, E. (2010). SF1 and SF2 helicases: Family matters. *Current Opinion in Structural Biology*, 20(3), 313-324. https://doi.org/10.1016/j.sbi.2010.03.011
- Fan, J., Leroux-Coyau, M., Savery, N. J., & Strick, T. R. (2016). Reconstruction of bacterial transcription-coupled repair at single-molecule resolution. *Nature*, 536(7615), 234-237. https://doi.org/10.1038/nature19080
- Fan, S., Yuan, R., Ma, Y. X., Meng, Q., Goldberg, I. D., & Rosen, E. M. (2001). Mutant BRCA1 genes antagonize phenotype of wild-type BRCA1. Oncogene, 20(57), 8215-8235. https://doi.org/10.1038/sj.onc.1205033
- Feng, Q., Jagannathan, S., & Bradley, R. K. (2017). The RNA Surveillance Factor UPF1 Represses Myogenesis via Its E3 Ubiquitin Ligase Activity. *Molecular Cell*, 67(2), 239-251.e6. https://doi.org/10.1016/j.molcel.2017.05.034
- Fenger-Grøn, M., Fillman, C., Norrild, B., & Lykke-Andersen, J. (2005). Multiple processing body factors and the ARE binding protein TTP activate mRNA decapping. *Molecular Cell*, 20(6), 905-915. https://doi.org/10.1016/j.molcel.2005.10.031
- Fernandes, R., Nogueira, G., da Costa, P. J., Pinto, F., & Romão, L. (2019). Nonsense-Mediated mRNA Decay in Development, Stress and Cancer. In L. Romão (Éd.), *The mRNA Metabolism in Human Disease* (p. 41-83). Springer International Publishing. https://doi.org/10.1007/978-3-030-19966-1_3

- Fernández, I. S., Yamashita, A., Arias-Palomo, E., Bamba, Y., Bartolomé, R. A., Canales, M. A., Teixidó, J., Ohno, S., & Llorca, O. (2011). Characterization of SMG-9, an essential component of the nonsense-mediated mRNA decay SMG1C complex. *Nucleic Acids Research*, 39(1), 347-358. https://doi.org/10.1093/nar/gkq749
- Fiorini, F., Bagchi, D., Hir, H. L., & Croquette, V. (2015). Human Upf1 is a highly processive RNA helicase and translocase with RNP remodelling activities. *Nature Communications*, 6, 7581. https://doi.org/10.1038/ncomms8581
- Fiorini, F., Bonneau, F., & Le Hir, H. (2012). Chapter Twelve—Biochemical Characterization of the RNA Helicase UPF1 Involved in Nonsense-Mediated mRNA Decay. In E. Jankowsky (Éd.), Methods in Enzymology (Vol. 511, p. 255-274). Academic Press. https://doi.org/10.1016/B978-0-12-396546-2.00012-7
- Fiorini, F., Boudvillain, M., & Le Hir, H. (2013). Tight intramolecular regulation of the human Upf1 helicase by its N- and C-terminal domains. *Nucleic Acids Research*, 41(4), 2404-2415. https://doi.org/10.1093/nar/gks1320
- Fiorini, F., Robin, J.-P., Kanaan, J., Borowiak, M., Croquette, V., Hir, H. L., Jalinot, P., & Mocquet, V. (2018). HTLV-1 Tax plugs and freezes UPF1 helicase leading to nonsense-mediated mRNA decay inhibition. *Nature Communications*, 9(1), 431. https://doi.org/10.1038/s41467-017-02793-6
- Fisher, T. E., Marszalek, P. E., & Fernandez, J. M. (2000). Stretching single molecules into novel conformations using the atomic force microscope. *Nature Structural Biology*, 7(9), 719-724. https://doi.org/10.1038/78936
- Flury, V., Restuccia, U., Bachi, A., & Mühlemann, O. (2014). Characterization of phosphorylation- and RNA-dependent UPF1 interactors by quantitative proteomics. *Journal of Proteome Research*, 13(6), 3038-3053. https://doi.org/10.1021/pr5002143
- Flynn, R. A., Almada, A. E., Zamudio, J. R., & Sharp, P. A. (2011). Antisense RNA polymerase II divergent transcripts are P-TEFb dependent and substrates for the RNA exosome. *Proceedings of the National Academy of Sciences*, 108(26), 10460-10465. https://doi.org/10.1073/pnas.1106630108
- Fontaine, K. A., Leon, K. E., Khalid, M. M., Tomar, S., Jimenez-Morales, D., Dunlap, M., Kaye, J. A., Shah, P. S., Finkbeiner, S., Krogan, N. J., & Ott, M. (2018). The Cellular NMD Pathway Restricts Zika Virus Infection and Is Targeted by the Viral Capsid Protein. *mBio*, 9(6), 10.1128/mbio.02126-18. https://doi.org/10.1128/mbio.02126-18
- Ford, A. S., Guan, Q., Neeno-Eckwall, E., & Culbertson, M. R. (2006). Ebs1p, a Negative Regulator of Gene Expression Controlled by the Upf Proteins in the Yeast Saccharomyces cerevisiae. *Eukaryotic Cell*, 5(2), 301-312. https://doi.org/10.1128/EC.5.2.301-312.2006
- Fourati, Z., Roy, B., Millan, C., Coureux, P.-D., Kervestin, S., van Tilbeurgh, H., He, F., Usón, I., Jacobson, A., & Graille, M. (2014). A highly conserved region essential for NMD in the Upf2 N-terminal domain. *Journal* of Molecular Biology, 426(22), 3689-3702. https://doi.org/10.1016/j.jmb.2014.09.015
- Franks, T. M., Singh, G., & Lykke-Andersen, J. (2010). Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense-mediated mRNA decay. Cell, 143(6), 938-950. https://doi.org/10.1016/j.cell.2010.11.043
- Frischmeyer, P. A., van Hoof, A., O'Donnell, K., Guerrerio, A. L., Parker, R., & Dietz, H. C. (2002). An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science (New York, N.Y.)*, 295(5563), 2258-2261. https://doi.org/10.1126/science.1067338
- Fritz, S. E., Ranganathan, S., Wang, C. D., & Hogg, J. R. (2022). An alternative UPF1 isoform drives conditional remodeling of nonsense-mediated mRNA decay. *The EMBO Journal*, 41(10), e108898. https://doi.org/10.15252/embj.2021108898
- Frolova, L. Y., Tsivkovskii, R. Y., Sivolobova, G. F., Oparina, N. Y., Serpinsky, O. I., Blinov, V. M., Tatkov, S. I., & Kisselev, L. L. (1999). Mutations in the highly conserved GGQ motif of class 1 polypeptide release factors abolish ability of human eRF1 to trigger peptidyl-tRNA hydrolysis. RNA (New York, N.Y.), 5(8), 1014-1020. https://doi.org/10.1017/s135583829999043x
- Fromm, S. A., Truffault, V., Kamenz, J., Braun, J. E., Hoffmann, N. A., Izaurralde, E., & Sprangers, R. (2012). The structural basis of Edc3- and Scd6-mediated activation of the Dcp1:Dcp2 mRNA decapping complex. *The EMBO Journal*, 31(2), 279-290. https://doi.org/10.1038/emboj.2011.408
- Fukuhara, N., Ebert, J., Unterholzner, L., Lindner, D., Izaurralde, E., & Conti, E. (2005). SMG7 Is a 14-3-3-like Adaptor in the Nonsense-Mediated mRNA Decay Pathway. *Molecular Cell*, 17(4), 537-547. https://doi.org/10.1016/j.molcel.2005.01.010
- Gaba, A., Jacobson, A., & Sachs, M. S. (2005). Ribosome occupancy of the yeast CPA1 upstream open reading frame termination codon modulates nonsense-mediated mRNA decay. *Molecular Cell*, 20(3), 449-460.

https://doi.org/10.1016/j.molcel.2005.09.019

- Galy, V., Gadal, O., Fromont-Racine, M., Romano, A., Jacquier, A., & Nehrbass, U. (2004). Nuclear retention of unspliced mRNAs in yeast is mediated by perinuclear Mlp1. Cell, 116(1), 63-73. https://doi.org/10.1016/s0092-8674(03)01026-2
- Ganesan, R., Mangkalaphiban, K., Baker, R. E., He, F., & Jacobson, A. (2022). Ribosome-bound Upf1 forms distinct 80S complexes and conducts mRNA surveillance. *RNA*, 28(12), 1621-1642. https://doi.org/10.1261/rna.079416.122
- Garcia, D., Garcia, S., & Voinnet, O. (2014). Nonsense-mediated decay serves as a general viral restriction mechanism in plants. *Cell Host & Microbe*, 16(3), 391-402. https://doi.org/10.1016/j.chom.2014.08.001
- Gatfield, D., & Izaurralde, E. (2004). Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila. *Nature*, 429(6991), 575-578. https://doi.org/10.1038/nature02559
- Gatfield, D., Unterholzner, L., Ciccarelli, F. D., Bork, P., & Izaurralde, E. (2003). Nonsense-mediated mRNA decay in Drosophila: At the intersection of the yeast and mammalian pathways. *The EMBO Journal*, 22(15), 3960-3970. https://doi.org/10.1093/emboj/cdg371
- Ge, Z., Quek, B. L., Beemon, K. L., & Hogg, J. R. (2016). Polypyrimidine tract binding protein 1 protects mRNAs from recognition by the nonsense-mediated mRNA decay pathway. *eLife*, 5, e11155. https://doi.org/10.7554/eLife.11155
- Gehring, N. H., Kunz, J. B., Neu-Yilik, G., Breit, S., Viegas, M. H., Hentze, M. W., & Kulozik, A. E. (2005). Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements. *Molecular Cell*, 20(1), 65-75. https://doi.org/10.1016/j.molcel.2005.08.012
- Gehring, N. H., Lamprinaki, S., Kulozik, A. E., & Hentze, M. W. (2009). Disassembly of exon junction complexes by PYM. Cell, 137(3), 536-548. https://doi.org/10.1016/j.cell.2009.02.042
- Gehring, N. H., Neu-Yilik, G., Schell, T., Hentze, M. W., & Kulozik, A. E. (2003). Y14 and hUpf3b form an NMD-activating complex. *Molecular Cell*, 11(4), 939-949. https://doi.org/10.1016/s1097-2765(03)00142-4
- Gerbracht, J. V., Boehm, V., Britto-Borges, T., Kallabis, S., Wiederstein, J. L., Ciriello, S., Aschemeier, D. U., Krüger, M., Frese, C. K., Altmüller, J., Dieterich, C., & Gehring, N. H. (2020). CASC3 promotes transcriptome-wide activation of nonsense-mediated decay by the exon junction complex. *Nucleic Acids Research*, 48(15), 8626-8644. https://doi.org/10.1093/nar/gkaa564
- Gerstberger, S., Hafner, M., & Tuschl, T. (2014). A census of human RNA-binding proteins. *Nature Reviews Genetics*, 15(12), Article 12. https://doi.org/10.1038/nrg3813
- Gibbs, M. R., & Chanfreau, G. F. (2022). UPF1 adds an m6A feather to its (de)cap. *Cell Reports*, 39(8), 110898. https://doi.org/10.1016/j.celrep.2022.110898
- Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, 6(5), Article 5. https://doi.org/10.1038/nmeth.1318
- Giraldez, A. J., Mishima, Y., Rihel, J., Grocock, R. J., Van Dongen, S., Inoue, K., Enright, A. J., & Schier, A. F. (2006). Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science (New York, N.Y.)*, 312(5770), 75-79. https://doi.org/10.1126/science.1122689
- Glavan, F., Behm-Ansmant, I., Izaurralde, E., & Conti, E. (2006). Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. The EMBO Journal, 25(21), 5117-5125. https://doi.org/10.1038/sj.emboj.7601377
- Gloggnitzer, J., Akimcheva, S., Srinivasan, A., Kusenda, B., Riehs, N., Stampfl, H., Bautor, J., Dekrout, B., Jonak, C., Jiménez-Gómez, J. M., Parker, J. E., & Riha, K. (2014). Nonsense-Mediated mRNA Decay Modulates Immune Receptor Levels to Regulate Plant Antibacterial Defense. *Cell Host & Microbe*, 16(3), 376-390. https://doi.org/10.1016/j.chom.2014.08.010
- Godwin, A. R., Kojima, S., Green, C. B., & Wilusz, J. (2013). Kiss your tail goodbye: The role of PARN, Nocturnin, and Angel deadenylases in mRNA biology. *Biochimica et biophysica acta*, 1829(0), 571-579. https://doi.org/10.1016/j.bbagrm.2012.12.004
- Gong, C., Kim, Y. K., Woeller, C. F., Tang, Y., & Maquat, L. E. (2009). SMD and NMD are competitive pathways that contribute to myogenesis: Effects on PAX3 and myogenin mRNAs. *Genes & Development*, 23(1), 54-66. https://doi.org/10.1101/gad.1717309
- Gonzalez-Hilarion, S., Beghyn, T., Jia, J., Debreuck, N., Berte, G., Mamchaoui, K., Mouly, V., Gruenert, D. C., Déprez, B., & Lejeune, F. (2012). Rescue of nonsense mutations by amlexanox in human cells. Orphanet Journal of Rare Diseases, 7(1), 58. https://doi.org/10.1186/1750-1172-7-58

- Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., White, K. M., O'Meara, M. J., Rezelj, V. V., Guo, J. Z., Swaney, D. L., Tummino, T. A., Huettenhain, R., Kaake, R. M., Richards, A. L., Tutuncuoglu, B., Foussard, H., Batra, J., Haas, K., Modak, M., ... Krogan, N. J. (2020). A SARS-CoV-2 Protein Interaction Map Reveals Targets for Drug-Repurposing. *Nature*, 583(7816), 459-468. https://doi.org/10.1038/s41586-020-2286-9
- Gowravaram, M., Bonneau, F., Kanaan, J., Maciej, V. D., Fiorini, F., Raj, S., Croquette, V., Le Hir, H., & Chakrabarti, S. (2018). A conserved structural element in the RNA helicase UPF1 regulates its catalytic activity in an isoform-specific manner. *Nucleic Acids Research*, 46(5), 2648-2659. https://doi.org/10.1093/nar/gky040
- Gowravaram, M., Schwarz, J., Khilji, S. K., Urlaub, H., & Chakrabarti, S. (2019). Insights into the assembly and architecture of a Staufen-mediated mRNA decay (SMD)-competent mRNP. *Nature Communications*, 10(1), 1-14. https://doi.org/10.1038/s41467-019-13080-x
- Graille, M., Chaillet, M., & van Tilbeurgh, H. (2008). Structure of yeast Dom34: A protein related to translation termination factor Erf1 and involved in No-Go decay. *The Journal of Biological Chemistry*, 283(11), 7145-7154. https://doi.org/10.1074/jbc.M708224200
- Graille, M., & Séraphin, B. (2012). Surveillance pathways rescuing eukaryotic ribosomes lost in translation. Nature Reviews. Molecular Cell Biology, 13(11), 727-735. https://doi.org/10.1038/nrm3457
- Graves, E. T., Duboc, C., Fan, J., Stransky, F., Leroux-Coyau, M., & Strick, T. R. (2015). A dynamic DNA-repair complex observed by correlative single-molecule nanomanipulation and fluorescence. *Nature Structural & Molecular Biology*, 22(6), 452-457. https://doi.org/10.1038/nsmb.3019
- Grimson, A., O'Connor, S., Newman, C. L., & Anderson, P. (2004). SMG-1 is a phosphatidylinositol kinase-related protein kinase required for nonsense-mediated mRNA Decay in Caenorhabditis elegans. *Molecular and Cellular Biology*, 24(17), 7483-7490. https://doi.org/10.1128/MCB.24.17.7483-7490.2004
- Gupta, P., & Li, Y.-R. (2018). Upf proteins: Highly conserved factors involved in nonsense mRNA mediated decay. *Molecular Biology Reports*, 45(1), 39-55. https://doi.org/10.1007/s11033-017-4139-7
- Gurskaya, N. G., Pereverzev, A. P., Staroverov, D. B., Markina, N. M., & Lukyanov, K. A. (2016). Analysis of Nonsense-Mediated mRNA Decay at the Single-Cell Level Using Two Fluorescent Proteins. *Methods in Enzymology*, 572, 291-314. https://doi.org/10.1016/bs.mie.2016.02.008
- Guydosh, N. R., & Green, R. (2017). Translation of poly(A) tails leads to precise mRNA cleavage. RNA (New York, N.Y.), 23(5), 749-761. https://doi.org/10.1261/rna.060418.116
- Halbach, F., Reichelt, P., Rode, M., & Conti, E. (2013). The yeast ski complex: Crystal structure and RNA channeling to the exosome complex. *Cell*, 154(4), 814-826. https://doi.org/10.1016/j.cell.2013.07.017
- Hall, G. W., & Thein, S. (1994). Nonsense codon mutations in the terminal exon of the beta-globin gene are not associated with a reduction in beta-mRNA accumulation: A mechanism for the phenotype of dominant beta-thalassemia. *Blood*, 83(8), 2031-2037.
- Harigaya, Y., Jones, B. N., Muhlrad, D., Gross, J. D., & Parker, R. (2010). Identification and analysis of the interaction between Edc3 and Dcp2 in Saccharomyces cerevisiae. *Molecular and Cellular Biology*, 30(6), 1446-1456. https://doi.org/10.1128/MCB.01305-09
- He, F., Brown, A. H., & Jacobson, A. (1997). Upf1p, Nmd2p, and Upf3p are interacting components of the yeast nonsense-mediated mRNA decay pathway. *Molecular and Cellular Biology*, 17(3), 1580-1594. https://doi.org/10.1128/mcb.17.3.1580
- He, F., Ganesan, R., & Jacobson, A. (2013). Intra- and Intermolecular Regulatory Interactions in Upf1, the RNA Helicase Central to Nonsense-Mediated mRNA Decay in Yeast. *Molecular and Cellular Biology*, 33(23), 4672-4684. https://doi.org/10.1128/MCB.01136-13
- 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 & Development, 9(4), 437-454. https://doi.org/10.1101/gad.9.4.437
- He, F., & Jacobson, A. (2015). Control of mRNA decapping by positive and negative regulatory elements in the Dcp2 C-terminal domain. RNA, 21(9), 1633-1647. https://doi.org/10.1261/rna.052449.115
- He, F., Li, X., Spatrick, P., Casillo, R., Dong, S., & Jacobson, A. (2003). Genome-Wide Analysis of mRNAs Regulated by the Nonsense-Mediated and 5' to 3' mRNA Decay Pathways in Yeast. *Molecular Cell*, 12(6), 1439-1452. https://doi.org/10.1016/S1097-2765(03)00446-5
- He, F., Wu, C., & Jacobson, A. (2022). Dcp2 C-terminal cis-binding elements control selective targeting of the decapping enzyme by forming distinct decapping complexes. *eLife*, 11, e74410. https://doi.org/10.7554/eLife.74410

- Hellen, C. U. T. (2018). Translation Termination and Ribosome Recycling in Eukaryotes. Cold Spring Harbor Perspectives in Biology, 10(10), a032656. https://doi.org/10.1101/cshperspect.a032656
- Hentze, M. W., Castello, A., Schwarzl, T., & Preiss, T. (2018). A brave new world of RNA-binding proteins. Nature Reviews Molecular Cell Biology, 19(5), Article 5. https://doi.org/10.1038/nrm.2017.130
- Hir, H. L., Saulière, J., & Wang, Z. (2016). The exon junction complex as a node of post-transcriptional networks. Nature Reviews Molecular Cell Biology, 17(1), 41-54. https://doi.org/10.1038/nrm.2015.7
- Hocq, R., Paternina, J., Alasseur, Q., Genovesio, A., & Le Hir, H. (2018). Monitored eCLIP: High accuracy mapping of RNA-protein interactions. *Nucleic Acids Research*, 46(21), 11553-11565. https://doi.org/10.1093/nar/gky858
- Hodeib, S., Raj, S., Manosas, M., Zhang, W., Bagchi, D., Ducos, B., Fiorini, F., Kanaan, J., Le Hir, H., Allemand, J., Bensimon, D., & Croquette, V. (2017). A mechanistic study of helicases with magnetic traps. *Protein Science: A Publication of the Protein Society*, 26(7), 1314-1336. https://doi.org/10.1002/pro.3187
- Hodgkin, J., Papp, A., Pulak, R., Ambros, V., & Anderson, P. (1989). A new kind of informational suppression in the nematode Caenorhabditis elegans. *Genetics*, *123*(2), 301-313.
- Hoek, T. A., Khuperkar, D., Lindeboom, R. G. H., Sonneveld, S., Verhagen, B. M. P., Boersma, S., Vermeulen, M., & Tanenbaum, M. E. (2019). Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2019.05.008
- Hogg, J. R., & Goff, S. P. (2010). Upf1 senses 3'UTR length to potentiate mRNA decay. *Cell*, 143(3), 379-389. https://doi.org/10.1016/j.cell.2010.10.005
- Hollien, J., Lin, J. H., Li, H., Stevens, N., Walter, P., & Weissman, J. S. (2009). Regulated Ire1-dependent decay of messenger RNAs in mammalian cells. *The Journal of Cell Biology*, 186(3), 323-331. https://doi.org/10.1083/jcb.200903014
- Hombach, S., & Kretz, M. (2016). Non-coding RNAs: Classification, Biology and Functioning. In O. Slaby & G. A. Calin (Éds.), Non-coding RNAs in Colorectal Cancer (p. 3-17). Springer International Publishing. https://doi.org/10.1007/978-3-319-42059-2_1
- Horvathova, I., Voigt, F., Kotrys, A. V., Zhan, Y., Artus-Revel, C. G., Eglinger, J., Stadler, M. B., Giorgetti, L., & Chao, J. A. (2017). The Dynamics of mRNA Turnover Revealed by Single-Molecule Imaging in Single Cells. *Molecular Cell*, 68(3), 615-625.e9. https://doi.org/10.1016/j.molcel.2017.09.030
- 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 (Chichester, England), 23(6), 439-454. https://doi.org/10.1002/yea.1369
- Hu, W., Sweet, T. J., Chamnongpol, S., Baker, K. E., & Coller, J. (2009). Co-translational mRNA decay in Saccharomyces cerevisiae. *Nature*, 461(7261), 225-229. https://doi.org/10.1038/nature08265
- Huang, L., Lou, C.-H., Chan, W., Shum, E. Y., Shao, A., Stone, E., Karam, R., Song, H.-W., & Wilkinson, M. F. (2011). RNA homeostasis governed by cell type-specific and branched feedback loops acting on NMD. *Molecular Cell*, 43(6), 950-961. https://doi.org/10.1016/j.molcel.2011.06.031
- Hubstenberger, A., Courel, M., Bénard, M., Souquere, S., Ernoult-Lange, M., Chouaib, R., Yi, Z., Morlot, J.-B., Munier, A., Fradet, M., Daunesse, M., Bertrand, E., Pierron, G., Mozziconacci, J., Kress, M., & Weil, D. (2017). P-Body Purification Reveals the Condensation of Repressed mRNA Regulons. *Molecular Cell*, 68(1), 144-157.e5. https://doi.org/10.1016/j.molcel.2017.09.003
- Huntzinger, E., Kashima, I., Fauser, M., Saulière, J., & Izaurralde, E. (2008). SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. RNA (New York, N.Y.), 14(12), 2609-2617. https://doi.org/10.1261/rna.1386208
- Hurt, J. A., Robertson, A. D., & Burge, C. B. (2013). Global analyses of UPF1 binding and function reveal expanded scope of nonsense-mediated mRNA decay. *Genome Research*, 23(10), 1636-1650. https://doi.org/10.1101/gr.157354.113
- Hwang, H. J., Park, Y., & Kim, Y. K. (2021). UPF1: From mRNA Surveillance to Protein Quality Control. Biomedicines, 9(8), Article 8. https://doi.org/10.3390/biomedicines9080995
- Imamachi, N., Tani, H., & Akimitsu, N. (2012). Up-frameshift protein 1 (UPF1): Multitalented entertainer in RNA decay. Drug Discoveries & Therapeutics, 6(2), 55-61.
- Isken, O., Kim, Y. K., Hosoda, N., Mayeur, G. L., Hershey, J. W. B., & Maquat, L. E. (2008). Upf1 phosphorylation triggers translational repression during nonsense-mediated mRNA decay. *Cell*, 133(2), 314-327. https://doi.org/10.1016/j.cell.2008.02.030
- 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 Research*, 44(16), 7766-7776.

https://doi.org/10.1093/nar/gkw635

- Ivanov, I., Lo, K. C., Hawthorn, L., Cowell, J. K., & Ionov, Y. (2007). Identifying candidate colon cancer tumor suppressor genes using inhibition of nonsense-mediated mRNA decay in colon cancer cells. Oncogene, 26(20), 2873-2884. https://doi.org/10.1038/sj.onc.1210098
- 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. *The EMBO Journal*, 27(5), 736-747. https://doi.org/10.1038/emboj.2008.17
- Jaillon, O., Bouhouche, K., Gout, J.-F., Aury, J.-M., Noel, B., Saudemont, B., Nowacki, M., Serrano, V., Porcel, B. M., Ségurens, B., Le Mouël, A., Lepère, G., Schächter, V., Bétermier, M., Cohen, J., Wincker, P., Sperling, L., Duret, L., & Meyer, E. (2008). Translational control of intron splicing in eukaryotes. *Nature*, 451(7176), 359-362. https://doi.org/10.1038/nature06495
- Jankowsky, E. (2011). RNA helicases at work: Binding and rearranging. *Trends in Biochemical Sciences*, 36(1), 19-29. https://doi.org/10.1016/j.tibs.2010.07.008
- Jarroux, J., Morillon, A., & Pinskaya, M. (2017). History, Discovery, and Classification of lncRNAs. In M. R. S. Rao (Éd.), Long Non Coding RNA Biology (Vol. 1008, p. 1-46). Springer Singapore. https://doi.org/10.1007/978-981-10-5203-3_1
- Jensen, T. H., Jacquier, A., & Libri, D. (2013). Dealing with Pervasive Transcription. *Molecular Cell*, 52(4), 473-484. https://doi.org/10.1016/j.molcel.2013.10.032
- Jia, J., Furlan, A., Gonzalez-Hilarion, S., Leroy, C., Gruenert, D. C., Tulasne, D., & Lejeune, F. (2015). Caspases shutdown nonsense-mediated mRNA decay during apoptosis. *Cell Death and Differentiation*, 22(11), 1754-1763. https://doi.org/10.1038/cdd.2015.18
- Jiao, X., Xiang, S., Oh, C., Martin, C. E., Tong, L., & Kiledjian, M. (2010). Identification of a quality-control mechanism for mRNA 5'-end capping. *Nature*, 467(7315), 608-611. https://doi.org/10.1038/nature09338
- Jin, H., Kelley, A. C., Loakes, D., & Ramakrishnan, V. (2010). Structure of the 70S ribosome bound to release factor 2 and a substrate analog provides insights into catalysis of peptide release. *Proceedings of the National Academy of Sciences*, 107(19), 8593-8598. https://doi.org/10.1073/pnas.1003995107
- Jin, H., Suh, M. R., Han, J., Yeom, K.-H., Lee, Y., Heo, I., Ha, M., Hyun, S., & Kim, V. N. (2009). Human UPF1 Participates in Small RNA-Induced mRNA Downregulation. *Molecular and Cellular Biology*, 29(21), 5789-5799. https://doi.org/10.1128/MCB.00653-09
- Jinek, M., Coyle, S. M., & Doudna, J. A. (2011). Coupled 5' Nucleotide Recognition and Processivity in Xrn1-Mediated mRNA Decay. *Molecular Cell*, 41(5), 600-608. https://doi.org/10.1016/j.molcel.2011.02.004
- Johnson, G. E., Lalanne, J.-B., Peters, M. L., & Li, G.-W. (2020). Functionally uncoupled transcription–translation in Bacillus subtilis. *Nature*, 585(7823), Article 7823. https://doi.org/10.1038/s41586-020-2638-5
- Jonas, S., Weichenrieder, O., & Izaurralde, E. (2013). An unusual arrangement of two 14-3-3-like domains in the SMG5-SMG7 heterodimer is required for efficient nonsense-mediated mRNA decay. *Genes & Development*, 27(2), 211-225. https://doi.org/10.1101/gad.206672.112
- Joshi, K., Cao, L., & Farabaugh, P. J. (2019). The problem of genetic code misreading during protein synthesis. Yeast (*Chichester, England*), 36(1), 35-42. https://doi.org/10.1002/yea.3374
- Kadlec, J., Guilligay, D., Ravelli, R. B., & Cusack, S. (2006). Crystal structure of the UPF2-interacting domain of nonsense-mediated mRNA decay factor UPF1. RNA (New York, N.Y.), 12(10), 1817-1824. https://doi.org/10.1261/rna.177606
- Kadlec, J., Izaurralde, E., & Cusack, S. (2004). The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3. *Nature Structural & Molecular Biology*, 11(4), 330-337. https://doi.org/10.1038/nsmb741
- Kanaan, J., Raj, S., Decourty, L., Saveanu, C., Croquette, V., & Hir, H. L. (2018). UPF1-like helicase grip on nucleic acids dictates processivity. *Nature Communications*, 9(1), 3752. https://doi.org/10.1038/s41467-018-06313-y
- Karam, R., Lou, C.-H., Kroeger, H., Huang, L., Lin, J. H., & Wilkinson, M. F. (2015). The unfolded protein response is shaped by the NMD pathway. *EMBO Reports*, 16(5), 599-609. https://doi.org/10.15252/embr.201439696
- Karousis, E. D., Gurzeler, L.-A., Annibaldis, G., Dreos, R., & Mühlemann, O. (2020). Human NMD ensues independently of stable ribosome stalling. *Nature Communications*, 11(1), Article 1. https://doi.org/10.1038/s41467-020-17974-z
- Karousis, E. D., Gypas, F., Zavolan, M., & Mühlemann, O. (2021). Nanopore sequencing reveals endogenous

NMD-targeted isoforms in human cells. *Genome Biology*, 22(1), 223. https://doi.org/10.1186/s13059-021-02439-3

- Karousis, E. D., & Mühlemann, O. (2019). Nonsense-Mediated mRNA Decay Begins Where Translation Ends. *Cold* Spring Harbor Perspectives in Biology, 11(2), a032862. https://doi.org/10.1101/cshperspect.a032862
- Kashima, I., Jonas, S., Jayachandran, U., Buchwald, G., Conti, E., Lupas, A. N., & Izaurralde, E. (2010). SMG6 interacts with the exon junction complex via two conserved EJC-binding motifs (EBMs) required for nonsense-mediated mRNA decay. *Genes & Development*, 24(21), 2440-2450. https://doi.org/10.1101/gad.604610
- Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., Ohno, M., Dreyfuss, G., & Ohno, S. (2006). Binding of a novel SMG-1–Upf1–eRF1–eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes & Development*, 20(3), 355-367. https://doi.org/10.1101/gad.1389006
- Katsioudi, G., Dreos, R., Arpa, E. S., Gaspari, S., Liechti, A., Sato, M., Gabriel, C. H., Kramer, A., Brown, S. A., & Gatfield, D. (2023). A conditional Smg6 mutant mouse model reveals circadian clock regulation through the nonsense-mediated mRNA decay pathway. *Science Advances*, 9(2), eade2828. https://doi.org/10.1126/sciadv.ade2828
- Kaygun, H., & Marzluff, W. F. (2005). Translation termination is involved in histone mRNA degradation when DNA replication is inhibited. *Molecular and Cellular Biology*, 25(16), 6879-6888. https://doi.org/10.1128/MCB.25.16.6879-6888.2005
- Keeling, K. M., Du, M., & Bedwell, D. M. (2013). Therapies of Nonsense-Associated Diseases. In Madame Curie Bioscience Database [Internet]. Landes Bioscience. https://www.ncbi.nlm.nih.gov/books/NBK6183/
- Keeling, K. M., Wang, D., Dai, Y., Murugesan, S., Chenna, B., Clark, J., Belakhov, V., Kandasamy, J., Velu, S. E., Baasov, T., & Bedwell, D. M. (2013). Attenuation of Nonsense-Mediated mRNA Decay Enhances In Vivo Nonsense Suppression. PLOS ONE, 8(4), e60478. https://doi.org/10.1371/journal.pone.0060478
- Kerényi, Z., Mérai, Z., Hiripi, L., Benkovics, A., Gyula, P., Lacomme, C., Barta, E., Nagy, F., & Silhavy, D. (2008). Inter-kingdom conservation of mechanism of nonsense-mediated mRNA decay. *The EMBO Journal*, 27(11), 1585-1595. https://doi.org/10.1038/emboj.2008.88
- Kerr, T. P., Sewry, C. A., Robb, S. A., & Roberts, R. G. (2001). Long mutant dystrophins and variable phenotypes: Evasion of nonsense-mediated decay? *Human Genetics*, 109(4), 402-407. https://doi.org/10.1007/s004390100598
- Kilchert, C., Wittmann, S., & Vasiljeva, L. (2016). The regulation and functions of the nuclear RNA exosome complex. Nature Reviews. Molecular Cell Biology, 17(4), 227-239. https://doi.org/10.1038/nrm.2015.15
- Kim, Y. K., Furic, L., DesGroseillers, L., & Maquat, L. E. (2005). Mammalian Staufen1 Recruits Upf1 to Specific mRNA 3'UTRs so as to Elicit mRNA Decay. Cell, 120(2), 195-208. https://doi.org/10.1016/j.cell.2004.11.050
- Kim, Y. K., & Maquat, L. E. (2019). UPFront and center in RNA decay: UPF1 in nonsense-mediated mRNA decay and beyond. RNA, rna.070136.118. https://doi.org/10.1261/rna.070136.118
- Kishor, A., Fritz, S. E., Haque, N., Ge, Z., Tunc, I., Yang, W., Zhu, J., & Hogg, J. R. (2020). Activation and inhibition of nonsense-mediated mRNA decay control the abundance of alternative polyadenylation products. *Nucleic Acids Research*, 48(13), 7468-7482. https://doi.org/10.1093/nar/gkaa491
- Kostrz, D., Wayment-Steele, H. K., Wang, J. L., Follenfant, M., Pande, V. S., Strick, T. R., & Gosse, C. (2019). A modular DNA scaffold to study protein-protein interactions at single-molecule resolution. *Nature Nanotechnology*, 14(10), 988-993. https://doi.org/10.1038/s41565-019-0542-7
- Krogan, N. J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A. P., Punna, T., Peregrín-Alvarez, J. M., Shales, M., Zhang, X., Davey, M., Robinson, M. D., Paccanaro, A., Bray, J. E., Sheung, A., ... Greenblatt, J. F. (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature*, 440(7084), 637-643. https://doi.org/10.1038/nature04670
- Kumar, S., & Mohapatra, T. (2021). Deciphering Epitranscriptome: Modification of mRNA Bases Provides a New Perspective for Post-transcriptional Regulation of Gene Expression. Frontiers in Cell and Developmental Biology, 9. https://www.frontiersin.org/articles/10.3389/fcell.2021.628415
- Kunz, J. B., Neu-Yilik, G., Hentze, M. W., Kulozik, A. E., & Gehring, N. H. (2006). Functions of hUpf3a and hUpf3b in nonsense-mediated mRNA decay and translation. RNA, 12(6), 1015-1022. https://doi.org/10.1261/rna.12506
- Kurihara, Y., Matsui, A., Hanada, K., Kawashima, M., Ishida, J., Morosawa, T., Tanaka, M., Kaminuma, E., Mochizuki,

Y., Matsushima, A., Toyoda, T., Shinozaki, K., & Seki, M. (2009). Genome-wide suppression of aberrant mRNA-like noncoding RNAs by NMD in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, 106(7), 2453-2458. https://doi.org/10.1073/pnas.0808902106

- Kurosaki, T., Li, W., Hoque, M., Popp, M. W.-L., Ermolenko, D. N., Tian, B., & Maquat, L. E. (2014). A post-translational regulatory switch on UPF1 controls targeted mRNA degradation. Genes & Development, 28(17), 1900-1916. https://doi.org/10.1101/gad.245506.114
- Kurosaki, T., Popp, M. W., & Maquat, L. E. (2019). Quality and quantity control of gene expression by nonsense-mediated mRNA decay. Nature Reviews Molecular Cell Biology, 1. https://doi.org/10.1038/s41580-019-0126-2
- Łabno, A., Tomecki, R., & Dziembowski, A. (2016). Cytoplasmic RNA decay pathways—Enzymes and mechanisms. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 1863(12), 3125-3147. https://doi.org/10.1016/j.bbamcr.2016.09.023
- Lai, T., Cho, H., Liu, Z., Bowler, M. W., Piao, S., Parker, R., Kim, Y. K., & Song, H. (2012). Structural basis of the PNRC2-mediated link between mrna surveillance and decapping. *Structure (London, England: 1993)*, 20(12), 2025-2037. https://doi.org/10.1016/j.str.2012.09.009
- Langer, L. M., Gat, Y., Bonneau, F., & Conti, E. (2020). Structure of substrate-bound SMG1-8-9 kinase complex reveals molecular basis for phosphorylation specificity. *ELife*, 9, e57127. https://doi.org/10.7554/eLife.57127
- Lardenois, A., Liu, Y., Walther, T., Chalmel, F., Evrard, B., Granovskaia, M., Chu, A., Davis, R. W., Steinmetz, L. M., & Primig, M. (2011). Execution of the meiotic noncoding RNA expression program and the onset of gametogenesis in yeast require the conserved exosome subunit Rrp6. *Proceedings of the National Academy of Sciences of the United States of America*, 108(3), 1058-1063. https://doi.org/10.1073/pnas.1016459108
- Lareau, L. F., Brooks, A. N., Soergel, D. A. W., Meng, Q., & Brenner, S. E. (2007). The coupling of alternative splicing and nonsense-mediated mRNA decay. Advances in Experimental Medicine and Biology, 623, 190-211. https://doi.org/10.1007/978-0-387-77374-2_12
- Lasalde, C., Rivera, A. V., León, A. J., González-Feliciano, J. A., Estrella, L. A., Rodríguez-Cruz, E. N., Correa, M. E., Cajigas, I. J., Bracho, D. P., Vega, I. E., Wilkinson, M. F., & González, C. I. (2014). Identification and functional analysis of novel phosphorylation sites in the RNA surveillance protein Upf1. *Nucleic Acids Research*, 42(3), 1916-1929. https://doi.org/10.1093/nar/gkt1049
- Laumonnier, F., Shoubridge, C., Antar, C., Nguyen, L. S., Van Esch, H., Kleefstra, T., Briault, S., Fryns, J. P., Hamel, B., Chelly, J., Ropers, H. H., Ronce, N., Blesson, S., Moraine, C., Gécz, J., & Raynaud, M. (2010). Mutations of the UPF3B gene, which encodes a protein widely expressed in neurons, are associated with nonspecific mental retardation with or without autism. *Molecular Psychiatry*, 15(7), 767-776. https://doi.org/10.1038/mp.2009.14
- Lavysh, D., & Neu-Yilik, G. (2020). UPF1-Mediated RNA Decay—Danse Macabre in a Cloud. *Biomolecules*, 10(7). https://doi.org/10.3390/biom10070999
- Le Hir, H., Izaurralde, E., Maquat, L. E., & Moore, M. J. (2000). The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *The EMBO Journal*, 19(24), 6860-6869. https://doi.org/10.1093/emboj/19.24.6860
- Le Hir, H., Moore, M. J., & Maquat, L. E. (2000). Pre-mRNA splicing alters mRNP composition: Evidence for stable association of proteins at exon–exon junctions. *Genes & Development*, 14(9), 1098-1108.
- LeBlanc, J. J., & Beemon, K. L. (2004). Unspliced Rous sarcoma virus genomic RNAs are translated and subjected to nonsense-mediated mRNA decay before packaging. *Journal of Virology*, 78(10), 5139-5146. https://doi.org/10.1128/jvi.78.10.5139-5146.2004
- Lebreton, A., Tomecki, R., Dziembowski, A., & Séraphin, B. (2008). Endonucleolytic RNA cleavage by a eukaryotic exosome. *Nature*, 456(7224), 993-996. https://doi.org/10.1038/nature07480
- Lee, B. S., & Culbertson, M. R. (1995). Identification of an additional gene required for eukaryotic nonsense mRNA turnover. Proceedings of the National Academy of Sciences of the United States of America, 92(22), 10354-10358. https://doi.org/10.1073/pnas.92.22.10354
- Lee, H. H., Kim, Y.-S., Kim, K. H., Heo, I., Kim, S. K., Kim, O., Kim, H. K., Yoon, J. Y., Kim, H. S., Kim, D. J., Lee, S. J., Yoon, H. J., Kim, S. J., Lee, B. G., Song, H. K., Kim, V. N., Park, C.-M., & Suh, S. W. (2007). Structural and functional insights into Dom34, a key component of no-go mRNA decay. *Molecular Cell*, 27(6), 938-950. https://doi.org/10.1016/j.molcel.2007.07.019
- Lee, S. R., Pratt, G. A., Martinez, F. J., Yeo, G. W., & Lykke-Andersen, J. (2015). Target Discrimination in Nonsense-Mediated mRNA Decay Requires Upf1 ATPase Activity. *Molecular Cell*, 59(3), 413-425.

https://doi.org/10.1016/j.molcel.2015.06.036

- 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(12A), 2303-2314.
- Leeds, P., Wood, J. M., Lee, B. S., & Culbertson, M. R. (1992). Gene products that promote mRNA turnover in Saccharomyces cerevisiae. *Molecular and Cellular Biology*, 12(5), 2165-2177. https://doi.org/10.1128/mcb.12.5.2165-2177.1992
- Leppek, K., Schott, J., Reitter, S., Poetz, F., Hammond, M. C., & Stoecklin, G. (2013). Roquin promotes constitutive mRNA decay via a conserved class of stem-loop recognition motifs. *Cell*, 153(4), 869-881. https://doi.org/10.1016/j.cell.2013.04.016
- Levy, N., Eiler, S., Pradeau-Aubreton, K., Maillot, B., Stricher, F., & Ruff, M. (2016). Production of unstable proteins through the formation of stable core complexes. *Nature Communications*, 7, 10932. https://doi.org/10.1038/ncomms10932
- Li, S., & Wilkinson, M. F. (1998). Nonsense surveillance in lymphocytes? *Immunity*, 8(2), 135-141. https://doi.org/10.1016/s1074-7613(00)80466-5
- Li, T., Shi, Y., Wang, P., Guachalla, L. M., Sun, B., Joerss, T., Chen, Y.-S., Groth, M., Krueger, A., Platzer, M., Yang, Y.-G., Rudolph, K. L., & Wang, Z.-Q. (2015). Smg6/Est1 licenses embryonic stem cell differentiation via nonsense-mediated mRNA decay. *The EMBO Journal*, 34(12), 1630-1647. https://doi.org/10.15252/embj.201489947
- Linde, L., Boelz, S., Nissim-Rafinia, M., Oren, Y. S., Wilschanski, M., Yaacov, Y., Virgilis, D., Neu-Yilik, G., Kulozik, A. E., Kerem, E., & Kerem, B. (2007). Nonsense-mediated mRNA decay affects nonsense transcript levels and governs response of cystic fibrosis patients to gentamicin. *The Journal of Clinical Investigation*, 117(3), 683-692. https://doi.org/10.1172/JCI28523
- Lindeboom, R. G. H., Supek, F., & Lehner, B. (2016). The rules and impact of nonsense-mediated mRNA decay in human cancers. *Nature Genetics*, 48(10), 1112-1118. https://doi.org/10.1038/ng.3664
- Linder, P., & Fuller-Pace, F. (2015). Happy birthday: 25 years of DEAD-box proteins. *Methods in Molecular Biology (Clifton, N.J.)*, 1259, 17-33. https://doi.org/10.1007/978-1-4939-2214-7_2
- Liu, S.-W., Rajagopal, V., Patel, S. S., & Kiledjian, M. (2008). Mechanistic and Kinetic Analysis of the DcpS Scavenger Decapping Enzyme *. *Journal of Biological Chemistry*, 283(24), 16427-16436. https://doi.org/10.1074/jbc.M800341200
- Lloyd, J. P. B. (2018). The evolution and diversity of the nonsense-mediated mRNA decay pathway. F1000Research, 7. https://doi.org/10.12688/f1000research.15872.2
- Loh, B., Jonas, S., & Izaurralde, E. (2013). The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. Genes & Development, 27(19), 2125-2138. https://doi.org/10.1101/gad.226951.113
- Longman, D., Hug, N., Keith, M., Anastasaki, C., Patton, E. E., Grimes, G., & Cáceres, J. F. (2013). DHX34 and NBAS form part of an autoregulatory NMD circuit that regulates endogenous RNA targets in human cells, zebrafish and Caenorhabditis elegans. *Nucleic Acids Research*, 41(17), 8319-8331. https://doi.org/10.1093/nar/gkt585
- Longman, D., Plasterk, R. H. A., Johnstone, I. L., & Cáceres, J. F. (2007). Mechanistic insights and identification of two novel factors in the C. elegans NMD pathway. Genes & Development, 21(9), 1075-1085. https://doi.org/10.1101/gad.417707
- Losson, R., & Lacroute, F. (1979). Interference of nonsense mutations with eukaryotic messenger RNA stability. *Proceedings of the National Academy of Sciences of the United States of America*, 76(10), 5134-5137. https://doi.org/10.1073/pnas.76.10.5134
- Lou, C. H., Shao, A., Shum, E. Y., Espinoza, J. L., Huang, L., Karam, R., & Wilkinson, M. F. (2014). Posttranscriptional Control of the Stem Cell and Neurogenic Programs by the Nonsense-Mediated RNA Decay Pathway. Cell Reports, 6(4), 748-764. https://doi.org/10.1016/j.celrep.2014.01.028
- Lubas, M., Damgaard, C. K., Tomecki, R., Cysewski, D., Jensen, T. H., & Dziembowski, A. (2013). Exonuclease hDIS3L2 specifies an exosome-independent 3'-5' degradation pathway of human cytoplasmic mRNA. *The EMBO Journal*, 32(13), 1855-1868. https://doi.org/10.1038/emboj.2013.135
- Luke, B., Azzalin, C. M., Hug, N., Deplazes, A., Peter, M., & Lingner, J. (2007). Saccharomyces cerevisiae Ebs1p is a putative ortholog of human Smg7 and promotes nonsense-mediated mRNA decay. *Nucleic Acids Research*, 35(22), 7688-7697. https://doi.org/10.1093/nar/gkm912
- Luke, B., Panza, A., Redon, S., Iglesias, N., Li, Z., & Lingner, J. (2008). The Rat1p 5' to 3' exonuclease degrades telomeric repeat-containing RNA and promotes telomere elongation in Saccharomyces cerevisiae.
Molecular Cell, 32(4), 465-477. https://doi.org/10.1016/j.molcel.2008.10.019

- Luo, Y., Na, Z., & Slavoff, S. A. (2018). P-Bodies: Composition, Properties, and Functions. *Biochemistry*, 57(17), 2424-2431. https://doi.org/10.1021/acs.biochem.7b01162
- Lykke-Andersen, S., & Jensen, T. H. (2015). Nonsense-mediated mRNA decay: An intricate machinery that shapes transcriptomes. *Nature Reviews Molecular Cell Biology*, 16(11), 665-677. https://doi.org/10.1038/nrm4063
- Maderazo, A. B., Belk, J. P., He, F., & Jacobson, A. (2003). Nonsense-containing mRNAs that accumulate in the absence of a functional nonsense-mediated mRNA decay pathway are destabilized rapidly upon its restitution. *Molecular and Cellular Biology*, 23(3), 842-851. https://doi.org/10.1128/MCB.23.3.842-851.2003
- Maderazo, A., He, F., Mangus, D., & Jacobson, A. (2000). Upf1p Control of Nonsense mRNA Translation Is Regulated by Nmd2p and Upf3p. *Molecular and Cellular Biology*, 20. https://doi.org/10.1128/MCB.20.13.4591-4603.2000
- 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
- Malecki, M., Viegas, S. C., Carneiro, T., Golik, P., Dressaire, C., Ferreira, M. G., & Arraiano, C. M. (2013). The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. *The EMBO Journal*, 32(13), 1842-1854. https://doi.org/10.1038/emboj.2013.63
- Mallik, R., & Gross, S. P. (2004). Molecular motors: Strategies to get along. *Current Biology: CB*, 14(22), R971-982. https://doi.org/10.1016/j.cub.2004.10.046
- Manosas, M., Meglio, A., Spiering, M. M., Ding, F., Benkovic, S. J., Barre, F.-X., Saleh, O. A., Allemand, J. F., Bensimon, D., & Croquette, V. (2010). Magnetic tweezers for the study of DNA tracking motors. *Methods in Enzymology*, 475, 297-320. https://doi.org/10.1016/S0076-6879(10)75013-8
- Maquat, L. E., Kinniburgh, A. J., Rachmilewitz, E. A., & Ross, J. (1981). Unstable β -globin mRNA in mRNA-deficient $\beta 0$ thalassemia. *Cell*, 27(3, Part 2), 543-553. https://doi.org/10.1016/0092-8674(81)90396-2
- Maquat, L. E., Tarn, W.-Y., & Isken, O. (2010). The pioneer round of translation: Features and functions. *Cell*, 142(3), 368-374. https://doi.org/10.1016/j.cell.2010.07.022
- Marsden, S., Nardelli, M., Linder, P., & McCarthy, J. E. G. (2006). Unwinding single RNA molecules using helicases involved in eukaryotic translation initiation. *Journal of Molecular Biology*, 361(2), 327-335. https://doi.org/10.1016/j.jmb.2006.06.016
- Martinez, J., & Tuschl, T. (2004). RISC is a 5' phosphomonoester-producing RNA endonuclease. Genes & Development, 18(9), 975-980. https://doi.org/10.1101/gad.1187904
- Martins-Dias, P., & Romão, L. (2021). Nonsense suppression therapies in human genetic diseases. *Cellular and* Molecular Life Sciences, 78(10), 4677-4701. https://doi.org/10.1007/s00018-021-03809-7
- Maryati, M., Airhihen, B., & Winkler, G. S. (2015). The enzyme activities of Caf1 and Ccr4 are both required for deadenylation by the human Ccr4–Not nuclease module. *Biochemical Journal*, 469(Pt 1), 169-176. https://doi.org/10.1042/BJ20150304
- Marzluff, W. F., & Koreski, K. P. (2017). Birth and Death of Histone mRNAs. *Trends in Genetics: TIG*, 33(10), 745-759. https://doi.org/10.1016/j.tig.2017.07.014
- Marzluff, W. F., Wagner, E. J., & Duronio, R. J. (2008). Metabolism and regulation of canonical histone mRNAs: Life without a poly(A) tail. *Nature Reviews Genetics*, 9(11), Article 11. https://doi.org/10.1038/nrg2438
- Matelska, D., Steczkiewicz, K., & Ginalski, K. (2017). Comprehensive classification of the PIN domain-like superfamily. *Nucleic Acids Research*, 45(12), 6995-7020. https://doi.org/10.1093/nar/gkx494
- Mattick, J. S., Amaral, P. P., Carninci, P., Carpenter, S., Chang, H. Y., Chen, L.-L., Chen, R., Dean, C., Dinger, M. E., Fitzgerald, K. A., Gingeras, T. R., Guttman, M., Hirose, T., Huarte, M., Johnson, R., Kanduri, C., Kapranov, P., Lawrence, J. B., Lee, J. T., ... Wu, M. (2023). Long non-coding RNAs: Definitions, functions, challenges and recommendations. *Nature Reviews Molecular Cell Biology*, 24(6), Article 6. https://doi.org/10.1038/s41580-022-00566-8
- May, J. P., & Simon, A. E. (2021). Targeting of viral RNAs by Upf1-mediated RNA decay pathways. *Current Opinion in Virology*, 47, 1-8. https://doi.org/10.1016/j.coviro.2020.11.002
- May, J. P., Yuan, X., Sawicki, E., & Simon, A. E. (2018). RNA virus evasion of nonsense-mediated decay. *PLoS Pathogens*, 14(11), e1007459. https://doi.org/10.1371/journal.ppat.1007459
- McGlincy, N. J., & Smith, C. W. J. (2008). Alternative splicing resulting in nonsense-mediated mRNA decay: What is the meaning of nonsense? *Trends in Biochemical Sciences*, 33(8), 385-393. https://doi.org/10.1016/j.tibs.2008.06.001

- McIlwain, D. R., Pan, Q., Reilly, P. T., Elia, A. J., McCracken, S., Wakeham, A. C., Itie-Youten, A., Blencowe, B. J., & Mak, T. W. (2010). Smg1 is required for embryogenesis and regulates diverse genes via alternative splicing coupled to nonsense-mediated mRNA decay. *Proceedings of the National Academy of Sciences of the United States of America*, 107(27), 12186-12191. https://doi.org/10.1073/pnas.1007336107
- 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(1), 134-140. https://doi.org/10.1016/j.molcel.2007.10.031
- Medghalchi, S. M., Frischmeyer, P. A., Mendell, J. T., Kelly, A. G., Lawler, A. M., & Dietz, H. C. (2001). Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. *Human Molecular Genetics*, 10(2), 99-105. https://doi.org/10.1093/hmg/10.2.99
- Melamed, D., Pnueli, L., & Arava, Y. (2008). Yeast translational response to high salinity: Global analysis reveals regulation at multiple levels. *RNA* (*New York*, *N.Y.*), 14(7), 1337-1351. https://doi.org/10.1261/rna.864908
- Melero, R., Buchwald, G., Castaño, R., Raabe, M., Gil, D., Lázaro, M., Urlaub, H., Conti, E., & Llorca, O. (2012). The cryo-EM structure of the UPF-EJC complex shows UPF1 poised toward the RNA 3' end. *Nature Structural & Molecular Biology*, 19(5), 498-505, S1-2. https://doi.org/10.1038/nsmb.2287
- Melero, R., Uchiyama, A., Castaño, R., Kataoka, N., Kurosawa, H., Ohno, S., Yamashita, A., & Llorca, O. (2014). Structures of SMG1-UPFs complexes: SMG1 contributes to regulate UPF2-dependent activation of UPF1 in NMD. Structure (London, England: 1993), 22(8), 1105-1119. https://doi.org/10.1016/j.str.2014.05.015
- Mendell, J. T., Sharifi, N. A., Meyers, J. L., Martinez-Murillo, F., & Dietz, H. C. (2004). Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nature Genetics*, 36(10), Article 10. https://doi.org/10.1038/ng1429
- Metkar, M., Ozadam, H., Lajoie, B. R., Imakaev, M., Mirny, L. A., Dekker, J., & Moore, M. J. (2018). Higher-Order Organization Principles of Pre-translational mRNPs. *Molecular Cell*, 72(4), 715-726.e3. https://doi.org/10.1016/j.molcel.2018.09.012
- Miller, J. N., & Pearce, D. A. (2014). Nonsense-mediated decay in genetic disease: Friend or foe? *Mutation Research/Reviews in Mutation Research*, 762, 52-64. https://doi.org/10.1016/j.mrrev.2014.05.001
- Milligan, L., Torchet, C., Allmang, C., Shipman, T., & Tollervey, D. (2005). A Nuclear Surveillance Pathway for mRNAs with Defective Polyadenylation. *Molecular and Cellular Biology*, 25(22), 9996-10004. https://doi.org/10.1128/MCB.25.22.9996-10004.2005
- Min, E. E., Roy, B., Amrani, N., He, F., & Jacobson, A. (2013). Yeast Upf1 CH domain interacts with Rps26 of the 40S ribosomal subunit. RNA (New York, N.Y.), 19(8), 1105-1115. https://doi.org/10.1261/rna.039396.113
- Mino, T., Murakawa, Y., Fukao, A., Vandenbon, A., Wessels, H.-H., Ori, D., Uehata, T., Tartey, S., Akira, S., Suzuki, Y., Vinuesa, C. G., Ohler, U., Standley, D. M., Landthaler, M., Fujiwara, T., & Takeuchi, O. (2015). Regnase-1 and Roquin Regulate a Common Element in Inflammatory mRNAs by Spatiotemporally Distinct Mechanisms. *Cell*, 161(5), 1058-1073. https://doi.org/10.1016/j.cell.2015.04.029
- Mistiniene, E., Pozdniakovaite, N., Popendikyte, V., & Naktinis, V. (2005). Structure-based ligand binding sites of protein p14.5, a member of protein family YER057c/YIL051c/YjgF. International Journal of Biological Macromolecules, 37(1), 61-68. https://doi.org/10.1016/j.ijbiomac.2005.08.008
- Mitkevich, V. A., Kononenko, A. V., Petrushanko, I. Y., Yanvarev, D. V., Makarov, A. A., & Kisselev, L. L. (2006). Termination of translation in eukaryotes is mediated by the quaternary eRF1*eRF3*GTP*Mg2+ complex. The biological roles of eRF3 and prokaryotic RF3 are profoundly distinct. *Nucleic Acids Research*, 34(14), 3947-3954. https://doi.org/10.1093/nar/gkl549
- Molina-Navarro, M. M., Castells-Roca, L., Bellí, G., García-Martínez, J., Marín-Navarro, J., Moreno, J., Pérez-Ortín, J. E., & Herrero, E. (2008). Comprehensive transcriptional analysis of the oxidative response in yeast. The Journal of Biological Chemistry, 283(26), 17908-17918. https://doi.org/10.1074/jbc.M800295200
- Monachino, E., Spenkelink, L. M., & van Oijen, A. M. (2017). Watching cellular machinery in action, one molecule at a time. *The Journal of Cell Biology*, 216(1), 41-51. https://doi.org/10.1083/jcb.201610025
- Moriarty, P. M., Reddy, C. C., & Maquat, L. E. (1998). Selenium deficiency reduces the abundance of mRNA for Se-dependent glutathione peroxidase 1 by a UGA-dependent mechanism likely to be nonsense codon-mediated decay of cytoplasmic mRNA. *Molecular and Cellular Biology*, 18(5), 2932-2939. https://doi.org/10.1128/MCB.18.5.2932
- Morris, C., Cluet, D., & Ricci, E. P. (2021). Ribosome dynamics and mRNA turnover, a complex relationship under constant cellular scrutiny. *Wiley Interdisciplinary Reviews. RNA*, 12(6), e1658.

https://doi.org/10.1002/wrna.1658

- Mort, M., Ivanov, D., Cooper, D. N., & Chuzhanova, N. A. (2008). A meta-analysis of nonsense mutations causing human genetic disease. *Human Mutation*, 29(8), 1037-1047. https://doi.org/10.1002/humu.20763
- Mühlemann, O., & Jensen, T. H. (2012). MRNP quality control goes regulatory. *Trends in Genetics: TIG*, 28(2), 70-77. https://doi.org/10.1016/j.tig.2011.11.001
- Muhlrad, D., & Parker, R. (1994). Premature translational termination triggers mRNA decapping. *Nature*, 370(6490), 578-581. https://doi.org/10.1038/370578a0
- Muhlrad, D., & Parker, R. (1999). Aberrant mRNAs with extended 3' UTRs are substrates for rapid degradation by mRNA surveillance. *RNA (New York, N.Y.)*, 5(10), 1299-1307.
- Muñoz, O., Lore, M., & Jagannathan, S. (2023). The long and short of EJC-independent nonsense-mediated RNA decay. *Biochemical Society Transactions*, BST20221131. https://doi.org/10.1042/BST20221131
- Nagarajan, V. K., Jones, C. I., Newbury, S. F., & Green, P. J. (2013). XRN 5'→3' exoribonucleases: Structure, mechanisms and functions. *Biochimica* et *biophysica* acta, 1829(0), 590-603. https://doi.org/10.1016/j.bbagrm.2013.03.005
- Nagy, E., & Maquat, L. E. (1998). A rule for termination-codon position within intron-containing genes: When nonsense affects RNA abundance. *Trends in Biochemical Sciences*, 23(6), 198-199. https://doi.org/10.1016/s0968-0004(98)01208-0
- Nasif, S., Contu, L., & Mühlemann, O. (2018). Beyond quality control: The role of nonsense-mediated mRNA decay (NMD) in regulating gene expression. Seminars in Cell & Developmental Biology, 75, 78-87. https://doi.org/10.1016/j.semcdb.2017.08.053
- 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(7232), 1038-1042. https://doi.org/10.1038/nature07747
- Nelson, J. O., Förster, D., Frizzell, K. A., Luschnig, S., & Metzstein, M. M. (2018). Multiple Nonsense-Mediated mRNA Processes Require Smg5 in Drosophila. Genetics, 209(4), 1073-1084. https://doi.org/10.1534/genetics.118.301140
- Nelson, J. O., Moore, K. A., Chapin, A., Hollien, J., & Metzstein, M. M. (2016). Degradation of Gadd45 mRNA by nonsense-mediated decay is essential for viability. *eLife*, 5, e12876. https://doi.org/10.7554/eLife.12876
- Neu-Yilik, G., Raimondeau, E., Eliseev, B., Yeramala, L., Amthor, B., Deniaud, A., Huard, K., Kerschgens, K., Hentze, M. W., Schaffitzel, C., & Kulozik, A. E. (2017). Dual function of UPF3B in early and late translation termination. *The EMBO Journal*, 36(20), 2968-2986. https://doi.org/10.15252/embj.201797079
- Ngo, G. H. P., Grimstead, J. W., & Baird, D. M. (2021). UPF1 promotes the formation of R loops to stimulate DNA double-strand break repair. *Nature Communications*, 12(1), 3849. https://doi.org/10.1038/s41467-021-24201-w
- Nguyen, L., Jolly, L., Shoubridge, C., Chan, W., Huang, L., Laumonnier, F., Raynaud, M., Hackett, A., Field, M., Rodriguez, J., Srivastava, A., Lee, Y., Long, R., Addington, A., Rapoport, J., Suren, S., Hahn, C., Gamble, J., Wilkinson, M., ... Gecz, J. (2012). Transcriptome profiling of UPF3B/NMD-deficient lymphoblastoid cells from patients with various forms of intellectual disability. *Molecular psychiatry*, *17*(11), 1103-1115. https://doi.org/10.1038/mp.2011.163
- Nguyen, L. S., Wilkinson, M. F., & Gecz, J. (2014). Nonsense-mediated mRNA decay: Inter-individual variability and human disease. *Neuroscience and Biobehavioral Reviews*, 46 Pt 2, 175-186. https://doi.org/10.1016/j.neubiorev.2013.10.016
- Nicholson, P., Josi, C., Kurosawa, H., Yamashita, A., & Mühlemann, O. (2014). A novel phosphorylation-independent interaction between SMG6 and UPF1 is essential for human NMD. *Nucleic Acids Research*, 42(14), 9217-9235. https://doi.org/10.1093/nar/gku645
- Nissan, T., Rajyaguru, P., She, M., Song, H., & Parker, R. (2010). Decapping activators in Saccharomyces cerevisiae act by multiple mechanisms. *Molecular cell*, 39(5), 773-783. https://doi.org/10.1016/j.molcel.2010.08.025
- Ntini, E., Järvelin, A. I., Bornholdt, J., Chen, Y., Boyd, M., Jørgensen, M., Andersson, R., Hoof, I., Schein, A., Andersen, P. R., Andersen, P. K., Preker, P., Valen, E., Zhao, X., Pelechano, V., Steinmetz, L. M., Sandelin, A., & Jensen, T. H. (2013). Polyadenylation site-induced decay of upstream transcripts enforces promoter directionality. *Nature Structural & Molecular Biology*, 20(8), 923-928. https://doi.org/10.1038/nsmb.2640
- Ohnishi, T., Yamashita, A., Kashima, I., Schell, T., Anders, K. R., Grimson, A., Hachiya, T., Hentze, M. W.,

Anderson, P., & Ohno, S. (2003). Phosphorylation of hUPF1 induces formation of mRNA surveillance complexes containing hSMG-5 and hSMG-7. *Molecular Cell*, 12(5), 1187-1200. https://doi.org/10.1016/s1097-2765(03)00443-x

- Okada-Katsuhata, Y., Yamashita, A., Kutsuzawa, K., Izumi, N., Hirahara, F., & Ohno, S. (2012). N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. *Nucleic Acids Research*, 40(3), 1251-1266. https://doi.org/10.1093/nar/gkr791
- Orban, T. I., & Izaurralde, E. (2005). Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. *RNA*, 11(4), 459-469. https://doi.org/10.1261/rna.7231505
- Ozsolak, F., Kapranov, P., Foissac, S., Kim, S. W., Fishilevich, E., Monaghan, A. P., John, B., & Milos, P. M. (2010). Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation. *Cell*, *143*(6), 1018-1029. https://doi.org/10.1016/j.cell.2010.11.020
- Page, M. F., Carr, B., Anders, K. R., Grimson, A., & Anderson, P. (1999). SMG-2 is a phosphorylated protein required for mRNA surveillance in Caenorhabditis elegans and related to Upf1p of yeast. *Molecular* and Cellular Biology, 19(9), 5943-5951. https://doi.org/10.1128/MCB.19.9.5943
- Park, E., Gleghorn, M. L., & Maquat, L. E. (2013). Staufen2 functions in Staufen1-mediated mRNA decay by binding to itself and its paralog and promoting UPF1 helicase but not ATPase activity. Proceedings of the National Academy of Sciences of the United States of America, 110(2), 405-412. https://doi.org/10.1073/pnas.1213508110
- Park, J., Myong, S., Niedziela-Majka, A., Lee, K. S., Yu, J., Lohman, T. M., & Ha, T. (2010). PcrA helicase dismantles RecA filaments by reeling in DNA in uniform steps. *Cell*, 142(4), 544-555. https://doi.org/10.1016/j.cell.2010.07.016
- Park, O. H., Park, J., Yu, M., An, H.-T., Ko, J., & Kim, Y. K. (2016). Identification and molecular characterization of cellular factors required for glucocorticoid receptor-mediated mRNA decay. *Genes & Development*, 30(18), 2093-2105. https://doi.org/10.1101/gad.286484.116
- Parker, R., & Sheth, U. (2007). P bodies and the control of mRNA translation and degradation. *Molecular Cell*, 25(5), 635-646. https://doi.org/10.1016/j.molcel.2007.02.011
- Passmore, L. A., & Coller, J. (2022). Roles of mRNA poly(A) tails in regulation of eukaryotic gene expression. Nature reviews. Molecular cell biology, 23(2), 93-106. https://doi.org/10.1038/s41580-021-00417-y
- Passos, D. O., Doma, M. K., Shoemaker, C. J., Muhlrad, D., Green, R., Weissman, J., Hollien, J., & Parker, R. (2009). Analysis of Dom34 and Its Function in No-Go Decay. *Molecular Biology of the Cell*, 20(13), 3025-3032. https://doi.org/10.1091/mbc.E09-01-0028
- Paternoga, H., & Wilson, D. N. (2023). Ready, steady, go: Rapid ribosomal scanning to reach start codons. Molecular Cell, 83(1), 9-11. https://doi.org/10.1016/j.molcel.2022.12.008
- Pelechano, V., Wei, W., & Steinmetz, L. M. (2015). Widespread Co-translational RNA Decay Reveals Ribosome Dynamics. Cell, 161(6), 1400-1412. https://doi.org/10.1016/j.cell.2015.05.008
- Peltz, S. W., Brown, A. H., & Jacobson, A. (1993). MRNA destabilization triggered by premature translational termination depends on at least three cis-acting sequence elements and one trans-acting factor. Genes & Development, 7(9), 1737-1754. https://doi.org/10.1101/gad.7.9.1737
- Pereira, F. J. C., Teixeira, A., Kong, J., Barbosa, C., Silva, A. L., Marques-Ramos, A., Liebhaber, S. A., & Romão, L. (2015). Resistance of mRNAs with AUG-proximal nonsense mutations to nonsense-mediated decay reflects variables of mRNA structure and translational activity. *Nucleic Acids Research*, 43(13), 6528-6544. https://doi.org/10.1093/nar/gkv588
- Pereverzev, A. P., Gurskaya, N. G., Ermakova, G. V., Kudryavtseva, E. I., Markina, N. M., Kotlobay, A. A., Lukyanov, S. A., Zaraisky, A. G., & Lukyanov, K. A. (2015). Method for quantitative analysis of nonsense-mediated mRNA decay at the single cell level. *Scientific Reports*, 5, 7729. https://doi.org/10.1038/srep07729
- Perlick, H. A., Medghalchi, S. M., Spencer, F. A., Kendzior, R. J., & Dietz, H. C. (1996). Mammalian orthologues of a yeast regulator of nonsense transcript stability. *Proceedings of the National Academy of Sciences of* the United States of America, 93(20), 10928-10932.
- Pierron, G., & Weil, D. (2018). Re-viewing the 3D Organization of mRNPs. *Molecular Cell*, 72(4), 603-605. https://doi.org/10.1016/j.molcel.2018.10.044
- Pisarev, A. V., Skabkin, M. A., Pisareva, V. P., Skabkina, O. V., Rakotondrafara, A. M., Hentze, M. W., Hellen, C. U. T., & Pestova, T. V. (2010). The role of ABCE1 in eukaryotic posttermination ribosomal recycling. *Molecular Cell*, 37(2), 196-210. https://doi.org/10.1016/j.molcel.2009.12.034
- Pisareva, V. P., Skabkin, M. A., Hellen, C. U. T., Pestova, T. V., & Pisarev, A. V. (2011). Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. *The EMBO Journal*, 30(9), 1804-1817. https://doi.org/10.1038/emboj.2011.93

- Popp, M. W., & Maquat, L. E. (2015). Attenuation of nonsense-mediated mRNA decay facilitates the response to chemotherapeutics. *Nature Communications*, 6(1), Article 1. https://doi.org/10.1038/ncomms7632
- Popp, M. W., & Maquat, L. E. (2018). Nonsense-mediated mRNA Decay and Cancer. Current opinion in genetics & development, 48, 44-50. https://doi.org/10.1016/j.gde.2017.10.007
- Popp, M. W.-L., Cho, H., & Maquat, L. E. (2020). Viral subversion of nonsense-mediated mRNA decay. *RNA* (*New York, N.Y.*), 26(11), 1509-1518. https://doi.org/10.1261/rna.076687.120
- Powers, K. T., Szeto, J.-Y. A., & Schaffitzel, C. (2020). New insights into no-go, non-stop and nonsense-mediated mRNA decay complexes. *Current Opinion in Structural Biology*, 65, 110-118. https://doi.org/10.1016/j.sbi.2020.06.011
- Preker, P., Nielsen, J., Kammler, S., Lykke-Andersen, S., Christensen, M. S., Mapendano, C. K., Schierup, M. H., & Jensen, T. H. (2008). RNA exosome depletion reveals transcription upstream of active human promoters. *Science (New York, N.Y.)*, 322(5909), 1851-1854. https://doi.org/10.1126/science.1164096
- Pulak, R., & Anderson, P. (1993). MRNA surveillance by the Caenorhabditis elegans smg genes. Genes & Development, 7(10), 1885-1897.
- Qi, Z., Pugh, R. A., Spies, M., & Chemla, Y. R. (2013). Sequence-dependent base pair stepping dynamics in XPD helicase unwinding. *ELife*, 2, e00334. https://doi.org/10.7554/eLife.00334
- Ramage, H. R., Kumar, G. R., Verschueren, E., Johnson, J. R., Von Dollen, J., Johnson, T., Newton, B., Shah, P., Horner, J., Krogan, N. J., & Ott, M. (2015). A combined proteomics/genomics approach links hepatitis C virus infection with nonsense-mediated mRNA decay. *Molecular Cell*, 57(2), 329-340. https://doi.org/10.1016/j.molcel.2014.12.028
- Ramanathan, A., Robb, G. B., & Chan, S.-H. (2016). mRNA capping: Biological functions and applications. Nucleic Acids Research, 44(16), 7511-7526. https://doi.org/10.1093/nar/gkw551
- Raney, K. D., Byrd, A. K., & Aarattuthodiyil, S. (2013). Structure and Mechanisms of SF1 DNA Helicases. Advances in Experimental Medicine and Biology, 767. https://doi.org/10.1007/978-1-4614-5037-5_2
- Rebbapragada, I., & Lykke-Andersen, J. (2009). Execution of nonsense-mediated mRNA decay: What defines a substrate? *Current Opinion in Cell Biology*, 21(3), 394-402. https://doi.org/10.1016/j.ceb.2009.02.007
- Rehwinkel, J., Behm-Ansmant, I., Gatfield, D., & Izaurralde, E. (2005). A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. RNA (New York, N.Y.), 11(11), 1640-1647. https://doi.org/10.1261/rna.2191905
- Reynolds, N. M., Lazazzera, B. A., & Ibba, M. (2010). Cellular mechanisms that control mistranslation. *Nature Reviews Microbiology*, 8(12), Article 12. https://doi.org/10.1038/nrmicro2472
- Ri, H., Lee, J., Sonn, J. Y., Yoo, E., Lim, C., & Choe, J. (2019). Drosophila CrebB is a Substrate of the Nonsense-Mediated mRNA Decay Pathway that Sustains Circadian Behaviors. *Molecules and Cells*, 42(4), 301-312. https://doi.org/10.14348/molcells.2019.2451
- Riehs-Kearnan, N., Gloggnitzer, J., Dekrout, B., Jonak, C., & Riha, K. (2012). Aberrant growth and lethality of Arabidopsis deficient in nonsense-mediated RNA decay factors is caused by autoimmune-like response. *Nucleic Acids Research*, 40(12), 5615-5624. https://doi.org/10.1093/nar/gks195
- Rieu, M., Vieille, T., Radou, G., Jeanneret, R., Ruiz-Gutierrez, N., Ducos, B., Allemand, J.-F., & Croquette, V. (2021). Parallel, linear, and subnanometric 3D tracking of microparticles with Stereo Darkfield Interferometry. *Science Advances*, 7(6), eabe3902. https://doi.org/10.1126/sciadv.abe3902
- Rogalska, M. E., Vivori, C., & Valcárcel, J. (2023). Regulation of pre-mRNA splicing: Roles in physiology and disease, and therapeutic prospects. *Nature Reviews Genetics*, 24(4), Article 4. https://doi.org/10.1038/s41576-022-00556-8
- Roy, B., & Jacobson, A. (2013). The intimate relationships of mRNA decay and translation. *Trends in Genetics: TIG*, 29(12), 691-699. https://doi.org/10.1016/j.tig.2013.09.002
- Rufener, S. C., & Mühlemann, O. (2013). EIF4E-bound mRNPs are substrates for nonsense-mediated mRNA decay in mammalian cells. *Nature Structural & Molecular Biology*, 20(6), 710-717. https://doi.org/10.1038/nsmb.2576
- Ruiz-Gutierrez, N., Rieu, M., Ouellet, J., Allemand, J.-F., Croquette, V., & Le Hir, H. (2022). Novel approaches to study helicases using magnetic tweezers. In *Methods in Enzymology* (p. S0076687922001264). Elsevier. https://doi.org/10.1016/bs.mie.2022.03.035
- Saito, S., Hosoda, N., & Hoshino, S. (2013). The Hbs1-Dom34 protein complex functions in non-stop mRNA decay in mammalian cells. The Journal of Biological Chemistry, 288(24), 17832-17843. https://doi.org/10.1074/jbc.M112.448977
- Salas-Marco, J., & Bedwell, D. M. (2004). GTP hydrolysis by eRF3 facilitates stop codon decoding during eukaryotic translation termination. *Molecular and Cellular Biology*, 24(17), 7769-7778.

https://doi.org/10.1128/MCB.24.17.7769-7778.2004

- Sanduja, S., Blanco, F. F., & Dixon, D. A. (2011). The roles of TTP and BRF proteins in regulated mRNA decay. Wiley Interdisciplinary Reviews. RNA, 2(1), 42-57. https://doi.org/10.1002/wrna.28
- Sayani, S., Janis, M., Lee, C. Y., Toesca, I., & Chanfreau, G. F. (2008). Widespread Impact of Nonsense-Mediated mRNA Decay on the Yeast Intronome. *Molecular* cell, 31(3), 360-370. https://doi.org/10.1016/j.molcel.2008.07.005
- Schaeffer, D., Tsanova, B., Barbas, A., Reis, F. P., Dastidar, E. G., Sanchez-Rotunno, M., Arraiano, C. M., & van Hoof, A. (2009). The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. *Nature Structural & Molecular Biology*, 16(1), 56-62. https://doi.org/10.1038/nsmb.1528
- Schmid, M., & Jensen, T. H. (2010). Nuclear quality control of RNA polymerase II transcripts: Nuclear quality control of RNA polymerase II transcripts. Wiley Interdisciplinary Reviews: RNA, 1(3), 474-485. https://doi.org/10.1002/wrna.24
- Schweingruber, C., Soffientini, P., Ruepp, M.-D., Bachi, A., & Mühlemann, O. (2016). Identification of Interactions in the NMD Complex Using Proximity-Dependent Biotinylation (BioID). PLoS ONE, 11(3). https://doi.org/10.1371/journal.pone.0150239
- Serdar, L. D., Whiteside, D. L., & Baker, K. E. (2016). ATP hydrolysis by UPF1 is required for efficient translation termination at premature stop codons. *Nature Communications*, 7, 14021. https://doi.org/10.1038/ncomms14021
- Serin, G., Gersappe, A., Black, J. D., Aronoff, R., & Maquat, L. E. (2001). Identification and characterization of human orthologues to Saccharomyces cerevisiae Upf2 protein and Upf3 protein (Caenorhabditis elegans SMG-4). Molecular and Cellular Biology, 21(1), 209-223. https://doi.org/10.1128/MCB.21.1.209-223.2001
- Seyedali, A., & Berry, M. J. (2014). Nonsense-mediated decay factors are involved in the regulation of selenoprotein mRNA levels during selenium deficiency. RNA, 20(8), 1248. https://doi.org/10.1261/rna.043463.113
- Sharif, H., & Conti, E. (2013). Architecture of the Lsm1-7-Pat1 Complex: A Conserved Assembly in Eukaryotic mRNA Turnover. *Cell Reports*, 5(2), 283-291. https://doi.org/10.1016/j.celrep.2013.10.004
- She, M., Decker, C. J., Sundramurthy, K., Liu, Y., Chen, N., Parker, R., & Song, H. (2004). Crystal structure of Dcp1p and its functional implications in mRNA decapping. *Nature Structural & Molecular Biology*, 11(3), Article 3. https://doi.org/10.1038/nsmb730
- Sheth, U., & Parker, R. (2006). Targeting of Aberrant mRNAs to Cytoplasmic Processing Bodies. *Cell*, 125(6), 1095-1109. https://doi.org/10.1016/j.cell.2006.04.037
- Shirley, R. L., Lelivelt, M. J., Schenkman, L. R., Dahlseid, J. N., & Culbertson, M. R. (1998). A factor required for nonsense-mediated mRNA decay in yeast is exported from the nucleus to the cytoplasm by a nuclear export signal sequence. *Journal of Cell Science*, 111(21), 3129-3143. https://doi.org/10.1242/jcs.111.21.3129
- Shoemaker, C. J., & Green, R. (2011). Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. *Proceedings of the National Academy of Sciences*, 108(51), E1392-E1398. https://doi.org/10.1073/pnas.1113956108
- Shum, E. Y., Jones, S. H., Shao, A., Chousal, J. N., Krause, M. D., Chan, W.-K., Lou, C.-H., Espinoza, J. L., Song, H.-W., Phan, M. H., Ramaiah, M., Huang, L., McCarrey, J. R., Peterson, K. J., De Rooij, D. G., Cook-Andersen, H., & Wilkinson, M. F. (2016). The Antagonistic Gene Paralogs Upf3a and Upf3b Govern Nonsense-Mediated RNA Decay. *Cell*, 165(2), 382-395. https://doi.org/10.1016/j.cell.2016.02.046
- 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(3), 563-576. https://doi.org/10.1261/rna.815108
- Simms, C. L., Yan, L. L., & Zaher, H. S. (2017). Ribosome Collision Is Critical for Quality Control during No-Go Decay. Molecular Cell, 68(2), 361-373.e5. https://doi.org/10.1016/j.molcel.2017.08.019
- Singh, G., Pratt, G., Yeo, G. W., & Moore, M. J. (2015). The Clothes Make the mRNA: Past and Present Trends in mRNP Fashion. *Annual Review of Biochemistry*, 84(1), 325-354. https://doi.org/10.1146/annurev-biochem-080111-092106
- Singleton, M. R., Dillingham, M. S., & Wigley, D. B. (2007). Structure and Mechanism of Helicases and Nucleic Acid Translocases. Annual Review of Biochemistry, 76(1), 23-50. https://doi.org/10.1146/annurev.biochem.76.052305.115300

- Sioud, M. (2021). RNA Interference: Story and Mechanisms. Methods in Molecular Biology (Clifton, N.J.), 2282, 1-15. https://doi.org/10.1007/978-1-0716-1298-9_1
- Skružný, M., Schneider, C., Rácz, A., Weng, J., Tollervey, D., & Hurt, E. (2009). An Endoribonuclease Functionally Linked to Perinuclear mRNP Quality Control Associates with the Nuclear Pore Complexes. PLOS Biology, 7(1), e1000008. https://doi.org/10.1371/journal.pbio.1000008
- Song, J.-J., Liu, J., Tolia, N. H., Schneiderman, J., Smith, S. K., Martienssen, R. A., Hannon, G. J., & Joshua-Tor, L. (2003). The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nature Structural Biology*, 10(12), 1026-1032. https://doi.org/10.1038/nsb1016
- Sonneveld, S., Verhagen, B. M. P., & Tanenbaum, M. E. (2020). Heterogeneity in mRNA Translation. *Trends in Cell Biology*, 30(8), 606-618. https://doi.org/10.1016/j.tcb.2020.04.008
- Standart, N., & Weil, D. (2018). P-Bodies: Cytosolic Droplets for Coordinated mRNA Storage. Trends in Genetics, 34(8), 612-626. https://doi.org/10.1016/j.tig.2018.05.005
- Steiger, M., Carr-Schmid, A., Schwartz, D. C., Kiledjian, M., & Parker, R. (2003). Analysis of recombinant yeast decapping enzyme. RNA, 9(2), 231-238. https://doi.org/10.1261/rna.2151403
- Stevens, A. (1979). Evidence for a 5' → 3' direction of hydrolysis by a 5' mononucleotide-producing exoribonuclease from Saccharomyces cerevisiae. Biochemical and Biophysical Research Communications, 86(4), 1126-1132. https://doi.org/10.1016/0006-291X(79)90234-1
- Stevens, A. (1988). mRNA-decapping enzyme from Saccharomyces cerevisiae: Purification and unique specificity for long RNA chains. *Molecular and Cellular Biology*, 8(5), 2005-2010. https://doi.org/10.1128/mcb.8.5.2005-2010.1988
- Strick, T. R., Croquette, V., & Bensimon, D. (1998). Homologous pairing in stretched supercoiled DNA. Proceedings of the National Academy of Sciences of the United States of America, 95(18), 10579-10583. https://doi.org/10.1073/pnas.95.18.10579
- Sun, B., & Wang, M. D. (2016). Single-molecule perspectives on helicase mechanisms and functions. *Critical Reviews in Biochemistry and Molecular Biology*, 51(1), 15-25. https://doi.org/10.3109/10409238.2015.1102195
- Sun, L., Mailliot, J., & Schaffitzel, C. (2023). Nonsense-Mediated mRNA Decay Factor Functions in Human Health and Disease. *Biomedicines*, 11(3), 722. https://doi.org/10.3390/biomedicines11030722
- 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 Research*, 22(7), 1350-1359. https://doi.org/10.1101/gr.130161.111
- Supek, F., Lehner, B., & Lindeboom, R. G. H. (2021). To NMD or Not To NMD: Nonsense-Mediated mRNA Decay in Cancer and Other Genetic Diseases. *Trends in Genetics*, 37(7), 657-668. https://doi.org/10.1016/j.tig.2020.11.002
- Svoboda, K., & Block, S. M. (1994). Force and velocity measured for single kinesin molecules. *Cell*, 77(5), 773-784. https://doi.org/10.1016/0092-8674(94)90060-4
- Syed, S., Pandey, M., Patel, S. S., & Ha, T. (2014). Single-molecule fluorescence reveals the unwinding stepping mechanism of replicative helicase. *Cell Reports*, 6(6), 1037-1045. https://doi.org/10.1016/j.celrep.2014.02.022
- Takahashi, S., Araki, Y., Ohya, Y., Sakuno, T., Hoshino, S.-I., Kontani, K., Nishina, H., & Katada, T. (2008). Upf1 potentially serves as a RING-related E3 ubiquitin ligase via its association with Upf3 in yeast. RNA (New York, N.Y.), 14(9), 1950-1958. https://doi.org/10.1261/rna.536308
- Tani, H., Imamachi, N., Salam, K. A., Mizutani, R., Ijiri, K., Irie, T., Yada, T., Suzuki, Y., & Akimitsu, N. (2012). Identification of hundreds of novel UPF1 target transcripts by direct determination of whole transcriptome stability. RNA Biology, 9(11), 1370-1379. https://doi.org/10.4161/rna.22360
- Tarpey, P. S., Raymond, F. L., Nguyen, L. S., Rodriguez, J., Hackett, A., Vandeleur, L., Smith, R., Shoubridge, C., Edkins, S., Stevens, C., O'Meara, S., Tofts, C., Barthorpe, S., Buck, G., Cole, J., Halliday, K., Hills, K., Jones, D., Mironenko, T., ... Gécz, J. (2007). Mutations in UPF3B, a member of the nonsense-mediated mRNA decay complex, cause syndromic and nonsyndromic mental retardation. *Nature Genetics*, 39(9), 1127-1133. https://doi.org/10.1038/ng2100
- Tarun, S. Z., & Sachs, A. B. (1997). Binding of eukaryotic translation initiation factor 4E (eIF4E) to eIF4G represses translation of uncapped mRNA. *Molecular and Cellular Biology*, 17(12), 6876-6886.
- Tate, W. P., Cridge, A. G., & Brown, C. M. (2018). « Stop » in protein synthesis is modulated with exquisite subtlety by an extended RNA translation signal. *Biochemical Society Transactions*, 46(6), 1615-1625. https://doi.org/10.1042/BST20180190

- 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(6777), 515-518. https://doi.org/10.1038/35006676
- Tharun, S., & Parker, R. (2001). Targeting an mRNA for decapping: Displacement of translation factors and association of the Lsm1p-7p complex on deadenylated yeast mRNAs. *Molecular Cell*, 8(5), 1075-1083. https://doi.org/10.1016/s1097-2765(01)00395-1
- Thermann, R., Neu-Yilik, G., Deters, A., Frede, U., Wehr, K., Hagemeier, C., Hentze, M. W., & Kulozik, A. E. (1998). Binary specification of nonsense codons by splicing and cytoplasmic translation. *The EMBO Journal*, 17(12), 3484-3494. https://doi.org/10.1093/emboj/17.12.3484
- Tian, M., Yang, W., Zhang, J., Dang, H., Lu, X., Fu, C., & Miao, W. (2017). Nonsense-mediated mRNA decay in Tetrahymena is EJC independent and requires a protozoa-specific nuclease. *Nucleic Acids Research*, 45(11), 6848-6863. https://doi.org/10.1093/nar/gkx256
- Tomecki, R., & Dziembowski, A. (2010). Novel endoribonucleases as central players in various pathways of eukaryotic RNA metabolism. RNA, 16(9), 1692-1724. https://doi.org/10.1261/rna.2237610
- Topisirovic, I., Svitkin, Y. V., Sonenberg, N., & Shatkin, A. J. (2011). Cap and cap-binding proteins in the control of gene expression. Wiley Interdisciplinary Reviews. RNA, 2(2), 277-298. https://doi.org/10.1002/wrna.52
- Tuck, A. C., Rankova, A., Arpat, A. B., Liechti, L. A., Hess, D., Iesmantavicius, V., Castelo-Szekely, V., Gatfield, D., & Bühler, M. (2020). Mammalian RNA Decay Pathways Are Highly Specialized and Widely Linked to Translation. *Molecular Cell*, 77(6), 1222-1236.e13. https://doi.org/10.1016/j.molcel.2020.01.007
- 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 deadenylase in Saccharomyces cerevisiae. Cell, 104(3), 377-386.
- Tudek, A., Schmid, M., & Jensen, T. H. (2019). Escaping nuclear decay: The significance of mRNA export for gene expression. *Current Genetics*, 65(2), 473-476. https://doi.org/10.1007/s00294-018-0913-x
- Unterholzner, L., & Izaurralde, E. (2004). SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. *Molecular Cell*, 16(4), 587-596. https://doi.org/10.1016/j.molcel.2004.10.013
- Valentini, M., & Linder, P. (2021). Happy Birthday: 30 Years of RNA Helicases. *Methods in Molecular Biology* (Clifton, N.J.), 2209, 17-34. https://doi.org/10.1007/978-1-0716-0935-4_2
- Valle-Orero, J., Rieu, M., Tran, P. L. T., Joubert, A., Raj, S., Allemand, J.-F., Croquette, V., & Boulé, J.-B. (2022). Strand switching mechanism of Pif1 helicase induced by its collision with a G-quadruplex embedded in dsDNA. Nucleic Acids Research, 50(15), 8767-8778. https://doi.org/10.1093/nar/gkac667
- 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(7354), 114-117. https://doi.org/10.1038/nature10118
- van Hoof, A., Frischmeyer, P. A., Dietz, H. C., & Parker, R. (2002). Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science (New York, N.Y.)*, 295(5563), 2262-2264. https://doi.org/10.1126/science.1067272
- van Leeuwen, F. W., de Kleijn, D. P., van den Hurk, H. H., Neubauer, A., Sonnemans, M. A., Sluijs, J. A., Köycü, S., Ramdjielal, R. D., Salehi, A., Martens, G. J., Grosveld, F. G., Peter, J., Burbach, H., & Hol, E. M. (1998). Frameshift mutants of beta amyloid precursor protein and ubiquitin-B in Alzheimer's and Down patients. *Science (New York, N.Y.)*, 279(5348), 242-247. https://doi.org/10.1126/science.279.5348.242
- Vicens, Q., Kieft, J. S., & Rissland, O. S. (2018). Revisiting the closed loop model and the nature of mRNA 5'–3' communication. *Molecular cell*, 72(5), 805-812. https://doi.org/10.1016/j.molcel.2018.10.047
- Villa, T., & Porrua, O. (2022). Pervasive transcription: A controlled risk. *The FEBS Journal*. https://doi.org/10.1111/febs.16530
- Villa, T., Rougemaille, M., & Libri, D. (2008). Nuclear quality control of RNA polymerase II ribonucleoproteins in yeast: Tilting the balance to shape the transcriptome. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1779(9), 524-531. https://doi.org/10.1016/j.bbagrm.2008.06.009
- Wada, M., Lokugamage, K. G., Nakagawa, K., Narayanan, K., & Makino, S. (2018). Interplay between coronavirus, a cytoplasmic RNA virus, and nonsense-mediated mRNA decay pathway. Proceedings of the National Academy of Sciences of the United States of America, 115(43), E10157-E10166. https://doi.org/10.1073/pnas.1811675115
- Wang, J., Shin, B.-S., Alvarado, C., Kim, J.-R., Bohlen, J., Dever, T. E., & Puglisi, J. D. (2022). Rapid 40S scanning and its regulation by mRNA structure during eukaryotic translation initiation. *Cell*, 185(24), 4474-4487.e17. https://doi.org/10.1016/j.cell.2022.10.005
- Wang, W., Cajigas, I. J., Peltz, S. W., Wilkinson, M. F., & González, C. I. (2006). Role for Upf2p phosphorylation in

Saccharomyces cerevisiae nonsense-mediated mRNA decay. *Molecular and Cellular Biology*, 26(9), 3390-3400. https://doi.org/10.1128/MCB.26.9.3390-3400.2006

- Wang, W., Czaplinski, K., Rao, Y., & Peltz, S. W. (2001). The role of Upf proteins in modulating the translation read-through of nonsense-containing transcripts. *The EMBO Journal*, 20(4), 880-890. https://doi.org/10.1093/emboj/20.4.880
- Wang, Z., Ballut, L., Barbosa, I., & Le Hir, H. (2018). Exon Junction Complexes can have distinct functional flavours to regulate specific splicing events. *Scientific Reports*, 8(1), 9509. https://doi.org/10.1038/s41598-018-27826-y
- Webster, M. W., Chen, Y.-H., Stowell, J. A. W., Alhusaini, N., Sweet, T., Graveley, B. R., Coller, J., & Passmore, L.
 A. (2018). MRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not
 Nucleases. *Molecular Cell*, 70(6), 1089-1100.e8. https://doi.org/10.1016/j.molcel.2018.05.033
- Weischenfeldt, J., Damgaard, I., Bryder, D., Theilgaard-Mönch, K., Thoren, L. A., Nielsen, F. C., Jacobsen, S. E. W., Nerlov, C., & Porse, B. T. (2008). NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements. *Genes & Development*, 22(10), 1381-1396. https://doi.org/10.1101/gad.468808
- Wells, S. E., Hillner, P. E., Vale, R. D., & Sachs, A. B. (1998). Circularization of mRNA by Eukaryotic Translation Initiation Factors. *Molecular Cell*, 2(1), 135-140. https://doi.org/10.1016/S1097-2765(00)80122-7
- Wen, J., & Brogna, S. (2010). Splicing-dependent NMD does not require the EJC in Schizosaccharomyces pombe. The EMBO Journal, 29(9), 1537-1551. https://doi.org/10.1038/emboj.2010.48
- Weng, Y., Czaplinski, K., & Peltz, S. W. (1996a). Genetic and biochemical characterization of mutations in the ATPase and helicase regions of the Upf1 protein. *Molecular and Cellular Biology*, 16(10), 5477-5490. https://doi.org/10.1128/mcb.16.10.5477
- Weng, Y., Czaplinski, K., & Peltz, S. W. (1996b). Identification and characterization of mutations in the UPF1 gene that affect nonsense suppression and the formation of the Upf protein complex but not mRNA turnover. *Molecular and Cellular Biology*, 16(10), 5491-5506.
- Will, C. L., & Lührmann, R. (2011). Spliceosome structure and function. Cold Spring Harbor Perspectives in Biology, 3(7), a003707. https://doi.org/10.1101/cshperspect.a003707
- Wittkopp, N., Huntzinger, E., Weiler, C., Saulière, J., Schmidt, S., Sonawane, M., & Izaurralde, E. (2009). Nonsense-Mediated mRNA Decay Effectors Are Essential for Zebrafish Embryonic Development and Survival. *Molecular and Cellular Biology*, 29(13), 3517-3528. https://doi.org/10.1128/MCB.00177-09
- Wittmann, J., Hol, E. M., & Jäck, H.-M. (2006). HUPF2 silencing identifies physiologic substrates of mammalian nonsense-mediated mRNA decay. *Molecular and Cellular Biology*, 26(4), 1272-1287. https://doi.org/10.1128/MCB.26.4.1272-1287.2006
- Wolf, J., & Passmore, L. A. (2014). MRNA Deadenylation by Pan2/Pan3. *Biochemical Society transactions*, 42(1), 184-187. https://doi.org/10.1042/BST20130211
- Wu, C., Roy, B., He, F., Yan, K., & Jacobson, A. (2020). Poly(A)-Binding Protein Regulates the Efficiency of Translation Termination. Cell Reports, 33(7), 108399. https://doi.org/10.1016/j.celrep.2020.108399
- 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(5), 725-737. https://doi.org/10.1016/j.cell.2005.04.030
- Xiang, S., Cooper-Morgan, A., Jiao, X., Kiledjian, M., Manley, J. L., & Tong, L. (2009). Structure and function of the 5'—>3' exoribonuclease Rat1 and its activating partner Rai1. Nature, 458(7239), 784-788. https://doi.org/10.1038/nature07731
- 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(7232), 1033-1037. https://doi.org/10.1038/nature07728
- Xue, G., Maciej, V. D., Amorim, A. M. D., Pak, M., Jayachandran, U., & Chakrabarti, S. (s. d.). Modulation of RNA-binding properties of the RNA helicase UPF1 by its activator UPF2.
- Yamashita, A., Izumi, N., Kashima, I., Ohnishi, T., Saari, B., Katsuhata, Y., Muramatsu, R., Morita, T., Iwamatsu, A., Hachiya, T., Kurata, R., Hirano, H., Anderson, P., & Ohno, S. (2009). SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay. *Genes & Development*, 23(9), 1091-1105. https://doi.org/10.1101/gad.1767209
- Yamashita, A., Ohnishi, T., Kashima, I., Taya, Y., & Ohno, S. (2001). Human SMG-1, a novel phosphatidylinositol

3-kinase-related protein kinase, associates with components of the mRNA surveillance complex and is involved in the regulation of nonsense-mediated mRNA decay. *Genes & Development*, 15(17), 2215-2228. https://doi.org/10.1101/gad.913001

- Yepiskoposyan, H., Aeschimann, F., Nilsson, D., Okoniewski, M., & Mühlemann, O. (2011). Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. RNA, 17(12), 2108-2118. https://doi.org/10.1261/rna.030247.111
- Yi, Z., Sanjeev, M., & Singh, G. (2021). The Branched Nature of the Nonsense-Mediated mRNA Decay Pathway. *Trends in Genetics: TIG*, 37(2), 143-159. https://doi.org/10.1016/j.tig.2020.08.010
- Yoon, J.-H., Choi, E.-J., & Parker, R. (2010). Dcp2 phosphorylation by Ste20 modulates stress granule assembly and mRNA decay in Saccharomyces cerevisiae. The Journal of Cell Biology, 189(5), 813-827. https://doi.org/10.1083/jcb.200912019
- Yu, J. H., Yang, W.-H., Gulick, T., Bloch, K. D., & Bloch, D. B. (2005). Ge-1 is a central component of the mammalian cytoplasmic mRNA processing body. RNA (New York, N.Y.), 11(12), 1795-1802. https://doi.org/10.1261/rna.2142405
- Zetoune, A. B., Fontanière, S., Magnin, D., Anczuków, O., Buisson, M., Zhang, C. X., & Mazoyer, S. (2008). Comparison of nonsense-mediated mRNA decay efficiency in various murine tissues. BMC Genetics, 9(1), 83. https://doi.org/10.1186/1471-2156-9-83
- Zhang, J., & Maquat, L. E. (1997). Evidence that translation reinitiation abrogates nonsense-mediated mRNA decay in mammalian cells. *The EMBO Journal*, *16*(4), 826-833. https://doi.org/10.1093/emboj/16.4.826
- Zhang, J., Sun, X., Qian, Y., & Maquat, L. E. (1998). Intron function in the nonsense-mediated decay of β-globin mRNA: Indications that pre-mRNA splicing in the nucleus can influence mRNA translation in the cytoplasm. *RNA*, *4*(7), 801-815. https://doi.org/10.1017/S1355838298971849
- Zhang, P., Wu, W., Chen, Q., & Chen, M. (2019). Non-Coding RNAs and their Integrated Networks. *Journal of Integrative Bioinformatics*, 16(3), 20190027. https://doi.org/10.1515/jib-2019-0027
- Zinder, J. C., & Lima, C. D. (2017). Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. *Genes & Development*, 31(2), 88-100. https://doi.org/10.1101/gad.294769.116
- Zünd, D., Gruber, A. R., Zavolan, M., & Mühlemann, O. (2013). Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3' UTRs. *Nature Structural & Molecular Biology*, 20(8), 936-943. https://doi.org/10.1038/nsmb.2635



Novel approaches to study helicases using magnetic tweezers

Nadia Ruiz-Gutierrez^{a,†} , Martin Rieu^{a,b,†} , Jimmy Ouellet^c , Jean-François Allemand^{a,b} , Vincent Croquette^{a,b,d,*} ,

and Hervé Le Hir^{a,*} 💿

^aInstitut de Biologie de l'Ecole Normale Supérieure (IBENS), Ecole normale supérieure, CNRS, INSERM, PSL Research University, Paris, France

^bLaboratoire de Physique de L'Ecole Normale Supérieure de Paris, CNRS, ENS, Université PSL, Sorbonne Université, Université de Paris, Paris, France

^cDepixus SAS, 3-5 impasse Reille, Paris, France

^dESPCI Paris, Université PSL, Paris, France

*Corresponding authors: e-mail address: vincent.croquette@espci.psl.eu; lehir@bio.ens.psl.eu

Contents

1.	. Introduction			
2.	Advances in synthesis, attachment and calibration of nucleic acid hairpins			
	for magnetic tweezer experiments			
	2.1 Construction of hairpins	7		
	2.2 Surface preparation	12		
	2.3 Bead preparation	15		
	2.4 Bead selection and calibration	17		
	2.5 Fast in situ three-way junction assembly by hairpin invasion	22		
	2.6 Recombinant protein production by double affinity purification	22		
3.	Advanced protocols for helicase characterization using magnetic tweezers	26		
	3.1 Mitigating the impact of force on helicase translocation	27		
	3.2 Peeling assay	27		
	3.3 Bubble assay	29		
	3.4 Estimating NA binding dynamics of helicases	31		
4.	Discussion			
References				

Abstract

Helicases form a universal family of molecular motors that bind and translocate onto nucleic acids. They are involved in essentially every aspect of nucleic acid metabolism: from DNA replication to RNA decay, and thus ensure a large spectrum of functions in the cell, making their study essential. The development of micromanipulation techniques

1

 $^{^{\}dagger}$ These authors participated equally in the writing of this paper.

such as magnetic tweezers for the mechanistic study of these enzymes has provided new insights into their behavior and their regulation that were previously unrevealed by bulk assays. These experiments allowed very precise measures of their translocation speed, processivity and polarity. Here, we detail our newest technological advances in magnetic tweezers protocols for high-quality measurements and we describe the new procedures we developed to get a more profound understanding of helicase dynamics, such as their translocation in a force independent manner, their nucleic acid binding kinetics and their interaction with roadblocks.

1. Introduction

Helicases are ubiquitous nanometric molecular motors involved in every step of nucleic acid (NA) metabolism and are present in nearly every living branch. They are NA-dependent NTPases that convert chemical energy from nucleoside triphosphates (NTP, mainly ATP) binding and hydrolysis to mechanical energy. This induces conformational changes that lead not only to translocation on single-stranded NA (ssNA) or unwinding double-stranded NA (dsNA) as previously thought, but also to remodeling NA-bound protein structures, or simply binding to NA and serving as protein recruitment scaffolds or NA clamps (Gao & Yang, 2020; Linder & Jankowsky, 2011). Initially, the role of helicases had largely been restricted to duplex unwinding and translocation during DNA replication, repair and recombination (Abdelhaleem, 2010; Brosh & Matson, 2020; Mendoza, Bourdoncle, Boulé, Brosh, & Mergny, 2016), yet, helicases are also required for RNA metabolism such as splicing, maturation, export, translation and decay (Cordin & Beggs, 2013; Khemici & Linder, 2018; Linder & Fuller-Pace, 2015).

It is estimated that around 1% of eukaryotic genes encode a panoply of RNA and DNA helicases, that are tightly regulated by intra- and intermolecular interactions (Lohman, Tomko, & Wu, 2008; Patel & Donmez, 2006; Umate, Tuteja, & Tuteja, 2011; Wu, 2012). Host and viral-encoded helicases play important roles in virus life cycle and infection, and given their vital roles in all cellular aspects, these enzymes are also linked to numerous genetic diseases and cancers. Thus, helicases constitute excellent targets for drug therapies and diagnostic tools (Abdelkrim, Banroques, & Kyle Tanner, 2021; Brosh & Matson, 2020; Datta & Brosh, 2018; Dhar, Datta, & Brosh, 2020; Steimer & Klostermeier, 2012).

The substantial coverage of helicase functions illustrates the importance of studying their mechanisms and regulation. Helicases are classified into six

superfamilies (SF1-SF6) differing from their primary sequence, NA preference and polarity, cellular function and structure (Gorbalenya & Koonin, 1993; Singleton, Dillingham, & Wigley, 2007). Superfamilies 1 and 2 share the same characteristic sequence motifs in their helicase core and generally act as monomers or dimers, unlike SF3-SF6 helicases that form hexameric toroidal ring-like structures. SF1 and SF2 include helicases with 5'-3' or 3'-5' polarities, and some subfamilies recognize either exclusively DNA, RNA or both NA.

Despite their structural homologies, each helicase dynamic is unique. Indeed, when studying helicase dynamics many questions arise regarding their behavior: does it act locally or does it translocate onto NA and if so, at which speed, processivity (number of consecutive base pairs translocated) and polarity? What is their affinity and how strong is their grip on doublestranded (ds), single-stranded (ss) or duplex NAs? How are their properties coupled to their function? How are they regulated and how do they interact with roadblocks such as NA-binding proteins or NA tertiary structures?

In order to answer these questions, a combination of bulk assays with single-molecule micromanipulation techniques such as optical tweezers and magnetic tweezers have been employed (Bockelmann, 2004; Greenleaf, Woodside, & Block, 2007; Neuman, Lionnet, & Allemand, 2007; Neuman & Nagy, 2008). Contrary to bulk assays, single-molecule approaches consist in observing in real time the unwinding of individual dsNA by individual helicases. In magnetic tweezers, a NA is tethered on both sides and stretched by applying an external force through a paramagnetic bead subjected to the gradient field of magnets (Fig. 1A). The NA's extension is measured by optically tracking the distance between the surface and the bead. When a helicase translocates along the NA, a portion of dsNA is unwound (Fig. 1B). The real-time position of the helicase along the NA can then be deduced, and thus its velocity, its processivity, and its directionality can be estimated.

Useful practical handbooks for these methods (Bustamante, Chemla, Liu, & Wang, 2021; Vilfan, Lipfert, Koster, Lemay, & Dekker, 2009) as well as several reviews of their application to helicase study have been published (Bianco, 2021; Hodeib et al., 2016; Manosas et al., 2010; Sun et al., 2010; Yodh, Schlierf, & Ha, 2010). Here, we describe new protocols that we designed to investigate further properties of helicases such as their translocation independently of the applied force, their binding grip, and the impact of roadblocks.

The article is organized as follows: we will first provide some tips for successful synthesis of DNA/RNA hairpins and focus on general progress



Fig. 1 Standard setup for the single-molecule characterization of helicases with magnetic tweezers. (A) A nucleic acid hairpin is tethered between a coverslip and a superparamagnetic bead. Magnets create a magnetic field gradient applying a force on the molecule through the bead. (B) Progression of a helicase (in orange) through the tethered hairpin at constant force (F = 8 pN) and bead position variation. *Panel (A): Image created with BiorRender.com*.

regarding surface chemistry, NA tethering techniques and *in vitro* purification of recombinant helicases before we detail magnetic tweezers force protocols and the associated data analysis procedures that underlie our most recent advances in helicase properties characterization.

2. Advances in synthesis, attachment and calibration of nucleic acid hairpins for magnetic tweezer experiments

As mentioned before, bulk assays and single-molecule micromanipulations have been largely used for the study of helicases. The latter allow unique measurements on individual molecules, accounting for population diversity in protein batch purifications, contrary to bulk experiments which result in average measurements. Here, we focus on magnetic tweezers experiments for helicase characterization, where a NA hairpin is attached on one side to a glass surface (coverslip) and on the other side to a paramagnetic bead subjected to a force imposed by magnets. The bead's position (Δz , in nanometers) reflects the molecule's extension under the constraint (Fig. 1A). Thus, magnetic tweezers are force-clamps while optical tweezers are position clamps. Many DNA/RNA synthesis and attachment strategies have been developed to study single NA mechanics under force (Janissen, Oberbarnscheidt, & Oesterhelt, 2009; Papini, Seifert, & Dulin, 2019; Riener et al., 2003; Walsh, Wang, & Weimer, 2001; Wildling et al., 2011; Yang et al., 2019) and are now routinely used in the single-molecule community. All of the helicase experiments that we describe here are based on NA hairpins, self-complementary sequences that unfold at forces between 13 and 20 pN (depending on the NA nature and sequence) and that close between 8 and 14 pN. While short hairpins can now be directly ordered from various providers (up to ~40 base pairs), longer hairpins need to be constructed *in vitro* by attaching tethers and a loop to a dsNA segment. Short synthetic hairpins should be preferred for low-noise experiments while long constructed hairpins can be used when a precision of 2 nm is sufficient. Indeed they allow a better statistical yield, since the larger distance of the bead from the surface reduces the risk of unspecific attachment and allows better identification of well-folded hairpins. As we explain in (Valle-Orero et al., 2022), short hairpins should also be used when the biological question requires that the analysis of a low-processivity enzyme is performed during the closing of the hairpin. In this case the hairpin should be smaller than the enzyme's processivity.

The NA is generally attached to the bead through streptavidin-biotin interactions, due to their biocompatibility, fast association and extremely high dissociation time at zero force ($\tau_0 = 10^5$ to 10^6 s at neutral pH and room temperature (Chilkoti & Stayton, 1995; Wilchek & Bayer, 1990)). This can be estimated using a simple Arrhenius model that gives τ_F the dissociation time under a force F as function of its value τ_0 at 0 pN force, and the distance Δx to the transition state: $\tau_F = \tau_0 e^{-F\Delta x/kBT}$, $\Delta x \sim 0.2-1$ nm, kBT being the thermal energy (Merkel, Nassoy, Leung, Ritchie, & Evans, 1997; Pincet & Husson, 2005; Williams et al., 2000). This dissociation time can even be made virtually infinite for this range of force compared to experimental time-scales by inserting two or three biotins in series (Guo, Ray, Kirkpatrick, Lad, & Akhremitchev, 2008).

For the surface-binding end, there are two choices of NA tethering: a fast and relatively weak digoxygenin/anti-digoxygenin (DIG/AntiDIG) binding (Fig. 2A and B) (fluorescein/anti-fluorescein have also been used, (Bryant et al., 2003)) or a slow (~1 h) but strong covalent binding, the most biocompatible and specific being dibenzylcyclooctyne (DBCO)-azide Click chemistry (Eeftens, van der Torre, Burnham, & Dekker, 2015) (Fig. 2C and D). DIG/AntiDIG binding can be reinforced by functionalizing the NA with several DIG groups but this is not compatible with high-resolution experiments such as helicase stepping. Indeed, the transient binding/unbinding of single groups at the timescale of a few tens of seconds create spurious steps in the nanometer range that pollute the discrete signal of helicases. For faster attachment using Click chemistry covalent bonds, the surface



Click-covalently bound oligonucleotide

Fig. 2 Different configurations for tethering a NA hairpin between a surface and a magnetic bead in magnetic tweezers. (A, B) DIG-AntiDIG hairpin tethering. The surface is first coated with Anti-DIG antibodies and the adaptor (A) or hairpin (B) are labeled with multiple DIG groups. The hairpin may also be indirectly (A) or directly (B) tethered to the bead through dual-biotinylation. (C, D) Click chemistry hairpin tethering. The azide surface is first coated with a DBCO-labeled ssDNA that is complementary to the adaptor (C) or the hairpin (D). The hairpin may also be directly (C) or indirectly (D) tethered to the bead through dual-biotinylation. In configuration (C), a ligation of the hairpin and the surface oligonucleotide can be performed *in situ* to covalently attach the hairpin to the surface.

6

can be coated with a DBCO-azide bound oligonucleotide to which the NA of interest can be attached with a complementary sequence of more than 30 base pairs (bp). This allows performing the attachment step only once per microfluidic chamber and then rapidly hybridizing hairpins (Fig. 2D) (or adaptors, Fig. 2C) and removing them by applying denaturing conditions. This solution is tractable for forces smaller than the shearing force of NA duplexes (~45–60 pN (Hatch, Danilowicz, Coljee, & Prentiss, 2008; Li et al., 2021)). For higher forces, the hairpin can also be ligated *in situ* if it is attached through a NA adaptor to the surface oligonucleotide (Fig. 2C). The different attachment options are summarized in Fig. 2.

If the hairpin is indirectly attached by hybridization, the adaptors should be long enough (>50 bp) so as they close behind the helicase passage, avoiding the molecule's detachment by the helicase. When high spatiotemporal resolution is required, such as for helicase stepping experiments or binding assays based on fluctuating hairpins, the tethers should be integrally made of dsNA so their higher stiffness increases the spatio-temporal resolution of the experiment.

Table 1 summarizes the optimal choices of hairpin lengths and attachment modes while the following section describes the protocols to follow to synthesize them.

2.1 Construction of hairpins

All NA sequences used in this section are referenced in Supporting Information Data 1 in the online version at https://doi.org/10.1016/bs. mie.2022.03.035 and their names are indicated in italic.

2.1.1 DNA hairpins

DNA hairpins are built from dsDNA obtained either by direct synthesis of gene fragments by a suitable provider, PCR or plasmid amplification. For applications where it is important to ensure the absence of epigenetic modifications, the first two methods shall be preferred. The dsDNA is then digested by Esp3I (BsmBI), a type IIS restriction enzyme that cleaves outside its recognition site leaving a 5' overhang, generating two different overhangs with a single cutting enzyme. The digested dsDNA is then ligated to the loop structure and the 5' biotinylated Y-shape adaptor which bears the complementary sequences needed for tethering. The obtained DNA hairpin is then ready to be coupled to the streptavidin-coated beads and hybridized to the Click-bound DNA coated on the surface (Fig. 3A).

helicase experiment	Preferred hairpin lengths	Preferred attachment mode	
Good statistics of velocity and processivity	200 bp \sim several kbp	Hybridized to ssDNA coated on the surface through azide-DBCO covalent binding	
High spatio-temporal resolution	10–100 bp	Fully hybridized to ssDNA coated on the surface (no ssDNA left after hybridization). Avoid DIG-antiDIG since it can produce spurious steps in the data	
Good control of helicase concentration	Not relevant	DIG-antiDIG attachment of the hairpin to the surface or direct covalent bond of the hairpin to the surface (Eeftens et al., 2015) (to avoid DNA-coating effects on local concentration of enzyme)	
Screening of a large number of different sequences (sequence effect on helicase translocation)	10–50 bp: library of synthetic hairpins can be directly bought from providers, preventing time-consuming lab constructions	Hybridized to ssDNA coated on the surface through azide-DBCO covalent binding. (no need of hairpin functionalization and reusable surface for each sequence)	

 Table 1
 Preferred hairpin lengths and attachment methods as a function of the main objective behind the single-molecule helicase experiment.

Materials

- Plasmid of interest or synthesized gene fragment (pPS006).
- Two primers designed for PCR amplification (*PrimerHP1*, *PrimerHP2*) containing the Esp3I recognition site and leaving 5' overhangs compatible with the Y-shape (*OliBiotin + OliCompAdaptor*) and the loop (*OliLoop*).
- Q5 PCR kit.
- Wizard [®] SV Gel and PCR Clean-Up System (Promega) or any other DNA purification kit
- 5' Dual-biotin oligonucleotide (OliBiotin, chemically synthesized).
- 5' Phosphorylated oligonucleotide complementary to the adaptor to the surface (*OliCompAdaptor*)
- Hairpin loop (OliLoop, chemically synthesized).
- T4 DNA ligase kit.

ARTICLE IN PRESS

Novel approaches to study helicases using magnetic tweezers



RNA hairpin

Fig. 3 Representation of the biochemical steps for hairpin construction. (A) Construction of a DNA hairpin. (B) DIG-tailing of a DNA hairpin for DIG/AntiDIG surfaces. (C) Construction of an RNA hairpin.

Protocol

- 1. Perform PCR protocol according to the primers' characteristics using Q5 polymerase and gel purify the PCR product.
- 2. Digest the purified fragment with Esp3I one hour at 37 °C.
- 3. Purify the digestion by column (no gel purification needed).
- 4. Ligate 100nM of the product with 1μM of Y-shape (OliBiotin + OliCompAdaptor) and 2μM of OliLoop (final concentrations) with T4 DNA Ligase in its buffer in a final volume of 20μL for one hour at 25 °C.

- Run the ligation product on a 1.5% agarose gel containing EtBr at 100 V in TAE 1X
- 6. Cut out the bands corresponding to the product using a blue light table.
- Extract the DNA from the agarose band either by using an extraction kit or by electroelution to maximize elution efficiency (Davis, Dibner, & Battey, 1986) and resuspend in water.
- 8. Measure concentration by spectrophotometry.
- Tips
- Primers must be designed to have a melting temperature close to 60 °C and to avoid primer dimers.
- For time saving, digestion and ligation can be coupled in the digestion mix using NEB CutSmart Buffer: after digestion at 37 °C, add 1 mM ATP and $0.5\,\mu$ L T4 DNA Ligase and incubate 2h at room temperature.
- Other types IIS restriction enzymes such as BsaI can be used.
- Avoid any heating steps during the purification since the melted hairpins may hybridize head-to-tail as dimers, especially when they are highly concentrated.

2.1.2 DIG-tailing DNA hairpins

The following protocol allows the preparation of hairpins compatible with DIG/AntiDIG attachment in microfluidic chambers (Fig. 2C). Briefly, an oligonucleotide complementary to the ss 3' end of the hairpin is hybridized to the hairpin. Its long tail lacking guanosines is polymerized using a Klenow Exo-, a polymerase without strand-displacement activity, in presence of DIG-coupled uridines and results in a poly-DIG-tailed hairpin (Fig. 3B).

Materials

- DNA hairpin synthesized as described in Section 2.1.1.
- Filling DIG oligonucleotide (*Fill-in-Dig-NoG*, chemically synthesized) designed to be complementary to the 3' single-stranded end of the DNA hairpin and a long 5'-end flap containing regularly spaced adenines and no guanosines. The 3' of the oligonucleotide may be blocked with a 3' phosphate in order to avoid 3' polymerization.
- Klenow Exo- and NEBuffer2 (NEB).
- dATP, dGTP and dUTP-Digoxigenin (ThermoFisher).

Protocol

1. Mix 10–100 ng of hairpin with 5 nM of Filling DIG oligonucleotide, 25 nM dATP, dGTP and dUTP-Digoxygenin, 1X NEBuffer 2, with 5U Klenow Exo- in a final volume of $20 \,\mu$ L.

Note: No dCTP nor dUTP are added to the mix.

- **2.** Incubate 1 h at 37 °C.
- **3.** Run the product in a 1.5% agarose gel containing EtBr, cut the band and purify the hairpin by gel extraction column kit.

2.1.3 DNA fluctuating probe

Short DNA hairpins (*HPKinLock*) used for kinetic locking assays (Section 3.4.2) were ordered from Eurogentec in order to have a 10 bp hairpin with a 4 bp loop, making a total of 24 bp. For more information refer to (Rieu, Valle-Orero, Ducos, Allemand, & Croquette, 2021).

2.1.4 RNA hairpin

Here, we describe a fast and reliable construction of an RNA hairpin inspired from (Desai et al., 2019). It consists of the *in vitro* transcription of a dsDNA palindromic sequence. Since a palindromic sequence cannot be easily chemically synthesized, it is obtained by the ligation of two gene blocks, both containing the hairpin sequence. One of the gene blocks contains the T7 promoter for RNA synthesis and the gene block junction corresponds to the tetra-thymidine loop (Fig. 3C). This one-step protocol reduces risks of RNAse contamination.

Materials

- Gene block 1 (*GbforRNA1*, chemically synthesized) designed in order to have a T7 promoter in 5' of the 180 bp hairpin sequence followed by a type IIS restriction enzyme recognition site (BsaI) that leaves a tetrathymidine overhang.
- Gene block 2 (*GbforRNA2*, chemically synthesized) designed to have the BsaI recognition site in 5' of the 180 bp hairpin sequence.
- BsaI-HF enzyme and CutSmart buffer.
- T7 transcription kit.
- Monarch RNA purification kit (NEB, T2040L).
- Murine RNAse inhibitor (NEB, M0314L).

Protocol

- 1. Mix 2pmol of each gene block with 0.5 U of the restriction enzyme, 200 U T4 DNA ligase in 1 mM ATP, 1x CutSmart Buffer at a final volume of $12 \mu L$.
- 2. Incubate 2h at 37 °C.
- 3. Transcribe $2\mu L$ of ligated palindromic DNA according to the T7 Transcription kit manufacturer's protocol in a final volume of $20\mu L$.

- 4. Purify the product using by column.
- **5.** Elute with $20 \mu L$ RNAse free water.
- 6. Measure concentration by spectrophotometry.
- 7. Add 1 μ L of the murine RNAse inhibitor before storing at -80 °C.

Tips

- Avoiding RNAse contamination is key to perform long single-molecule experiments with the RNA hairpin. We recommend using gloves and RNAse free water.
- The palindromic DNA and the RNA hairpin can be loaded in a 1.5% Agarose gel with EtBr in order to check their size and integrity.
- Enzymes are not heat-inactivated to avoid RNA/DNA duplexes during renaturation.

2.2 Surface preparation

Here, we describe the two main ways to tether hairpins (or adaptors) to the surface: through Click chemistry or DIG/AntiDIG interaction. Click chemistry corresponds to the attachment of DNA oligonucleotides coupled to a DCBO group on an azide-functionalized surface by producing a stable triazole link (Eeftens et al., 2015). The hairpins (or adaptors) can be then indirectly tethered to the DNA-coated surface by hybridization through complementary sequences and washed away with NaOH, rendering reusable and long-lasting surfaces. The two main disadvantages of Click attachment are that hairpins may dehybridize after helicase passage and, when working with ssDNA-binding helicases, the latter are titrated by the surface DNA resulting in protein concentration underestimation. On the other hand, hairpin attachment to the surface through DIG/Anti-DIG antibody is weaker but can be improved by using poly DIG-tailed hairpins that are directly tethered to anti-DIG antibody coated surfaces (Manosas et al., 2010).

For both preparations a 2.5 mL syringe, a $80 \times 120 \text{ mm}$ Mylar sheet with two 2 mm holes (50 µm thick, 48-2F-OC, CSHyde), a $35 \times 15 \text{ mm}$ double-sided channel-containing tape 50 µm thick (AR care 92712, Adhesive Research), two $10 \times 10 \text{ mm}$ double-sided tapes with 2 mm holes (AR care 92712, Adhesive Research), an injection sink and an output connector (homemade in Altuglass) are needed (Fig. 4).

2.2.1 Preparation of ssDNA-coated surfaces through Click chemistry Materials

- Azide functionalized coverslip (PolyAn 2D Azide).
- 3'-DBCO oligonucleotide 100 nM (*OliDBCO*, chemically synthesized, resuspended in 500 mM NaCl, 10 mM PBS).

ARTICLE IN PRESS

Novel approaches to study helicases using magnetic tweezers



Fig. 4 Assembly of a microfluidic chamber for magnetic tweezers experiments. Double-sided tapes allow the assembly of the microfluidic chamber between a coverslip and a Mylar sheet, bearing a channel through which a constant flow passes. The buffers and enzymes are injected through the input and a syringe connects the output to an outside pump. *Image created with BiorRender.com*.

 Passivation buffer 1X (PB) (140 mM NaCl, 3 mM KCl, 10 mM Na₂HPO₄, 1.76 mM KH₂PO₄, 2% BSA, 2% Pluronic F-127, 5 mM EDTA, 10 mM NaN₃ pH 7.4).

Protocol

- 1. Thoroughly wash the injection sink, the output connector, the Mylar sheet and the double-sided tapes with water then with Ethanol 70%. Rinse the output tube with water to ensure it is not blocked.
- **2.** Stick one face of the double-sided tape directly onto the coverslip without removing the other side's protection (Fig. 4).
- **3.** Put a 20–40 µL drop of 3'-DBCO oligonucleotide in the middle of the channel. Incubate at room temperature for 2h until dry.
- 4. Rinse thoroughly the channel with PB then with water.
- Once everything is dry, stick the Mylar sheet on the other side of the double-sided tape, making sure the holes are aligned with the channel (Fig. 4).
- **6.** Stick the two double-sided tapes with 2 mm holes aligned with the holes in the Mylar sheet then stick the input sink and the output connector (Fig. 4).

Microfluidic chambers can be used immediately or up to a month if stored at 4°C and protected from light.

Tips

 To avoid RNase contamination, do not add BSA in the PB as we noted it often introduces RNases, and beware of dust. Buffers may be filtered, and gloves must be used at all times.

- To wash away hybridized molecules from the surface it is possible to rinse with 15 mM NaOH.
- DNA-binding protein concentrations may be underestimated due to binding to the surface DNA.
- Working with helicases can cause dehybridization of the attached molecule. This can be avoided by adding longer hybridization sequences (>50 bp) that re-hybridize and encircle the helicase after its passage.
- Packs of PolyAn 2D Azide coverslips should be used within a week. If opened, the pack should be stocked at 4 °C, protected from light.
- Never use a sodium azide (NaN₃) containing solution to dilute the DBCOcontaining oligonucleotide or it will reduce the yield of the Click-chemistry reaction between the azide surface and the DBCO-labeled DNA.

2.2.2 Preparation of AntiDIG-coated surfaces

Materials

- Non-functionalized coverslip (Schott, NEXTERION[®] Coverslip custom, #1.5H).
- Sigmacote (Sigma-Aldrich).
- Monoclonal mouse Anti-Digoxin IgG (Jackson Immuno Research, 200– 002-156).
- Passivation buffer 1X (PB) (140 mM NaCl, 3 mM KCl, 10 mM Na₂HPO₄, 1.76 mM KH₂PO₄, 2% BSA, 2% Pluronic F-127, 5 mM EDTA, 10 mM NaN₃ pH 7.4).

Protocol

- 1. Thoroughly wash the injection sink, the output connector, the Mylar sheet and the double-sided tapes with water then with Ethanol 70%. Rinse the output tube with water to ensure it is not blocked.
- 2. Put a drop of Sigmacote on the coverslip and spread it with a pipette tip. Wait until dry in a covered place to avoid dust from falling on the coverslip.
- **3.** Stick one side of the double-sided tape directly onto the coverslip. Stick the Mylar sheet on the other side. The holes in the Mylar sheet should be aligned in the middle of the tape's channel (Fig. 4).
- 4. Stick the two double-sided tapes with 2 mm holes aligned with the holes in the Mylar sheet then stick the input sink and the output connector. Check if there are any leaks by flowing water through the channel with the syringe (Fig. 4).
- 5. Once the microfluidic chamber is ready, inject $40\,\mu\text{L}$ of the antiDIG in the input and gently flow it through the channel with the syringe in

order to cover the whole channel with the antibody. Incubate 3h at room temperature and cover the input sink to avoid evaporation.

6. Rinse the channel with $200 \,\mu\text{L}$ PB and let it incubate overnight at $4 \,^{\circ}\text{C}$ to saturate unspecific sites.

Note: Microfluidic chambers can be used for up to a couple of weeks if kept with PB at 4 °C.

Tips

- When sticking the tape, apply slight pressure to ensure proper adhesion of the tape and avoid leaks, but beware of breaking the coverslip.
- Bubbles are an indicator of a leak. If there is a leak, locate it and plug it with glue.
- To avoid evaporation, always cover the input sink with a 1.5 mL Eppendorf tube lid.
- To covalently attach the anti-DIG antibodies to the surface, an epoxycoated coverslip (Schott, NEXTERION[®] Coverslip E @1.5H) may be used.

2.3 Bead preparation

Tethering the hairpin to the streptavidin-coated bead can be achieved either indirectly through a dual-biotin labeled oligonucleotide (chemically synthesized) complementary to the hairpin's end, or directly by biotinylating the hairpin itself. The use of dual-biotin increases the interaction affinity in order to stabilize the hairpin's tethering.

2.3.1 Pre-hybridization of hairpin to adaptors

Materials

- Non-biotinylated hairpin (100 nM) produced as described in Section 2.1.
- Biotinylated bead adaptor (OliBiotin, 100 nM, chemically synthesized)
- NaCl 1M or KCl 1M.
- Passivation buffer 1X (PB) (140 mM NaCl, 3 mM KCl, 10 mM Na₂HPO₄, 1.76 mM KH₂PO₄, 2% BSA, 2% Pluronic F-127, 5 mM EDTA, 10 mM NaN₃ pH 7.4).

Protocol

- Mix 1µL of the hairpin with 1µL of the adaptor and 1µL of 1M salt (NaCl or KCl).
- 2. Incubate 10 min at room temperature for hybridization.
- 3. Add 97 μ L of PB to dilute the concentration to 1 nM.
- 4. The pre-hybridized hairpin can be immediately hybridized to the beads or stored at -20 °C.

Tips

- Biotinylated adaptors must be added at the same concentration as the hairpin since they cannot be washed after bead hybridization.

2.3.2 Hairpin/bead hybridization

Materials

- Dynabeads MyOne T1 (Thermofisher).
- Hairpin (pre-hybridized to adaptor or not) (10 pM).
- Passivation buffer 1X (PB) (140 mM NaCl, 3 mM KCl, 10 mM Na₂HPO₄, 1.76 mM KH₂PO₄, 2% BSA, 2% Pluronic F-127, 5 mM EDTA, 10 mM NaN₃ pH 7.4).
- If needed (Fig. 2C) adaptor between the hairpin and the surface (OliAdaptor, 100 nM)

Protocol

- 1. Wash $5 \mu L$ of beads three times with $200 \mu L$ of PB.
- 2. Resuspend the beads in $19 \mu L$ of PB.
- **3.** Add 2μL of the diluted pre-hybridized hairpin (10 pM), and if needed, 1μL of *OliAdaptor*.
- 4. Incubate at least 10 min at room temperature in rotation.
- 5. Wash three times with $200\,\mu\text{L}$ of PB in order to remove unbound NA.
- 6. Resuspend in 20 µL of PB (final concentration of DNA, 1 pM).
- 7. Stock beads on a rotator (10 rpm) at room temperature.

Note: Beads can be used for up to several months if sedimentation does not occur. For RNA beads this estimation may be shorter.

Tips

- The concentration of pre-hybridized hairpin and incubation time may need to be adjusted. Too high attachment to the surface without the possibility to open the hairpin indicates that the beads have more than one NA tethered. On the contrary, low attachment indicates insufficient hairpin tethering to the beads. We sometimes increase the final concentration of hairpins from 1 pM up to 25 pM (stock concentration up to 250 pM).
- Washing away the adaptor oligonucleotide is crucial or free adaptors will saturate the surface. To ensure thorough washing, make sure that less than $5\,\mu$ L of liquid are left in the tube between each washing step.

2.3.3 Bead injection onto the surface

Materials

- Beads tethered to the hairpin as prepared in Section 2.3.
- Click or DIG surface prepared as described in Section 2.2.

 Passivation buffer 1X (PB) (140 mM NaCl, 3 mM KCl, 10 mM Na₂HPO₄, 1.76 mM KH₂PO₄, 2% BSA, 2% Pluronic F-127, 5 mM EDTA, 10 mM NaN₃ pH 7.4).

Protocol

- 1. Lift the magnets to remove any magnetic force.
- 2. Put 200 µL of PB into the input inlet.
- 3. Apply a flow of $10 \,\mu\text{L}$ per minute with the microscope pump.
- 4. Inject 1 µL of the beads directly onto the cell.
- 5. When the beads are seen in the field of view, quickly lower the magnets to 10 pN and stop the flow.
- 6. When the flow is fully stopped, lift the magnets and let the beads sediment for 10 min.
- 7. Wash the surface with PB at $10\,\mu$ L/s until all the beads seen in the screen are attached to the surface.

Tips

- Before injection, verify that the syringe of the microscope pump is empty. Avoid emptying it during the experiment which induces flows that detach most of the molecules.
- To avoid bubbles, when injecting the beads, do not press until the second stop of the pipette and regularly check that the input sink is not dry when the flow is on.
- The incubation time of sedimented beads may have to be adjusted if too few or too many beads are attached.
- Once the beads are attached, lowering the magnets to a higher force (20 pN) for a few seconds may help the release of untethered beads.

2.4 Bead selection and calibration

Once the beads are prepared and injected onto the surface, their correct folding must be checked and their position calibrated in order to begin the experimental process.

2.4.1 Selection of properly folded hairpins

As described in (Manosas et al., 2010), beads tethered by a single NA hairpin must be distinguished from ones either non-specifically bound to the surface, not properly folded in hairpins or bound by two or more molecules. The latter need twice as much force to be opened than a single bound hairpin. Typically, DNA hairpins rapidly open around 15 pN and RNA hairpins at 18 pN, so the double-bound beads will not be opened even at

the microscope's maximal force. Improperly folded hairpins will show progressive opening and closing, contrary to properly folded ones.

Protocol

- 1. Select a field of view that includes between 30 and 60 beads that are not aggregated.
- 2. While alternating the force between 18 and 7 pN for example, visualize the beads that show a z-position change between high and low forces.
- **3.** If the z-position variation of the bead corresponds to the size of the hairpin (around 1 nm for 1 bp) and the opening is abrupt, the bead is bound by a properly folded hairpin and should be selected.
- 4. If the bead doesn't show movement with force variation it should not be selected.

2.4.2 Bead-to-bead force correction using hairpin opening forces

The force applied on the NA hairpins is crucial for the reproducibility of helicase single-molecule experiments, especially if the helicase's behavior is strongly force-dependent, as seen in the case of the Nsp13 Sars-CoV-2 helicase (Mickolajczyk et al., 2021). Force variation can be calibrated as a function of the distance between the magnets and the microfluidic chamber using the fluctuation of a bead tethered to a relatively long NA (Allemand, 1997; Gosse & Croquette, 2002; Ostrofet, Papini, & Dulin, 2018; Yu et al., 2014). For closed hairpins, where the bead is closer to the surface, these calibration protocols are not accurate due to the change of viscous drag near the surface (Brenner, 1961) and the random attachment points of the molecule with regard to their magnetic axis (Klaue & Seidel, 2009).

Thus, magnets need to be calibrated with long dsDNA molecules which then serve for subsequent experiments with shorter hairpin molecules. However, an uncertainty remains, due to the variation of magnetization from bead to bead, which depends on the heterogeneity of the magnetic bead preparation. This variation of magnetization induces a multiplicative factor between the true force applied on the bead and the one that is estimated through the calibration which is calculated as follows.

Protocol

- 1. The calibration with long molecules is performed on a large sample of different molecules and beads in order to get an accurate average of the force applied on a typical bead of the preparation.
- 2. Each subsequent experiment with hairpins should start with a few (\sim 20) cycles of opening and closing of the hairpins at a slow magnet speed (typ-ically 0.2 mm/s) (Fig. 5A).



Fig. 5 Force adjustment protocol to correct for the bead-to-bead magnetic heterogeneity. (A) Force extension curve over several force cycles of a 1 kb hairpin tethered to a Dynabead MyOne microbead. The red (blue) points represent the points with the highest time derivative of the extension change during the opening (closing). The average of the abscissa of these points over several cycles gives the estimated (i.e. before the bead-to-bead correction) opening (closing) force for that bead. (B) The estimated opening forces are compared for all beads tracked during a particular experiment. They follow a Gaussian distribution (inset), here of average 15 pN and standard deviation 1.6 pN. The average of the force applied over all beads gives the true opening force F_{true} (dashed orange line) and allows deducing the corrective multiplicative factor for each bead. Buffer: 70 mM KCl, 5 mM MgCl₂, 20 mM Tris–HCl pH 7.6.

- **3.** Draw the extension as a function of force during these opening cycles for each bead (Fig. 5B), which follows hysteresis cycles. Then, measure and average the magnet position for which the opening and the closing of the hairpin happen. This value is then translated into estimated opening and closing forces using the force calibration of step 1.
- **4.** Draw the distribution of opening forces for the whole assembly of beads (Fig. 5B inset). The average of the distribution gives the true opening force of the hairpin in the buffer used and at this magnet speed.
- 5. The multiplicative correction factor is assessed independently for each bead and corresponds to the ratio between the estimated opening force of each individual hairpin and the average opening force of the whole sample. For Dynabeads MyOne, it is typically 10% but sometimes reaches 20% (Fig. 5B). Taking this factor into account is important to avoid force errors.

2.4.3 Nanometers to base pairs conversion

The elasticity of ssDNA is involved in many aspects of the energetic interactions with enzymes. For single-molecule helicase experiments, it also allows the conversion from nanometers (nm) to base pairs (bp). Contrary to dsDNA where this elasticity curve is in complete agreement with the Worm Like Chain model (Bouchiat et al., 1999) and where only two parameters, which have been accurately measured, fully characterize it. ssDNA elasticity is not universal, it depends on sequence and salt (McIntosh, Duggan, Gouil, & Saleh, 2014) and there are fewer experimental characterizations with somewhat diverging parameters (Jacobson, McIntosh, Stevens, Rubinstein, & Saleh, 2017; Viader-Godoy, Pulido, Ibarra, Manosas, & Ritort, 2021). The measurement of the extension of single molecules is made harder by the propensity of ssDNA to unspecifically attach to the bead or the surface. As a consequence, when the extension of a ssDNA is measured, it is possible that a shorter molecule is being analyzed and thus, less bases than expected are being stretched.

To overcome this difficulty, we use a thorough protocol for the measurement of ssDNA force-extension curves (Fig. 6A). This consists of measuring the extension of a hairpin as a function of the force while preventing its closure with a small oligonucleotide complementary to the loop (Fig. 6B). This allows removing the contribution of the handles to the total extension and thus getting rid of the surface effects (Fig. 6C). The protocol shown in Fig. 6 needs to be performed before any single-molecule helicase experiment, in the enzyme's activity buffer, to ensure a thorough estimation of the nm-to-bp conversion coefficient.



Fig. 6 Protocol for nanometer to base pair conversion. (A) Sketch of the force protocol and the DNA behavior. (B) Superposition of curves representing the extension of a 2129 bp hairpin vs force in a series of cycles with force sweeping from 2 pN to 22 pN. Light gray curves with dots showcase the raw data with typically 100 cycles. The molecule is placed in a 50mM NaCl buffer in the presence of 100nM of a 12 bp oligonucleotide complementary to the apex region of the hairpin. The blue curve from (1) to (2) is the average extension of the closed hairpin which reflects the elasticity of the handles used to attach the hairpin to the bead and the surface. The blue curve from (2) to (3) represents the averaged opening behavior of the hairpin. The green curve from (4) to (8) pN corresponds to the elastic behavior of ssDNA, while the red curve represents the same behavior until 2 pN. The red curve corresponds to the situation where the oligonucleotide is hybridized and prevents the hairpin from refolding before the force is low enough to encircle the oligonucleotide or to eject it. (C) The final elasticity curve extracted from the curve in (A). The extension shown in the blue averaged curve from (1) to (2) was subtracted from the red averaged curve and normalized by the number of base pairs in the hairpin (2129 bp). This curve thus corresponds to the extension of two ssDNA nucleotides that were fitted with the Langevin model used for Freely Jointed Chain (Smith, Finzi, & Bustamante, 1992).

2.5 Fast in situ three-way junction assembly by hairpin invasion

Assembling three-way junctions directly in the microfluidic chamber has been proven useful to study helicases or helicase-like complexes that are able to regress stalled replication forks into Holliday junctions (Amit, Gileadi, & Stavans, 2004; Bétous et al., 2013; Bianco & Lu, 2021; Dawid, Croquette, Grigoriev, & Heslot, 2004; Manosas et al., 2013; Tan et al., 2017). However, the study of the processivity of such proteins is limited by the size of the synthetic oligonucleotides used to form the junctions. In order to overcome this limitation, we developed a protocol for the fast and simple assembly of three-way junctions of arbitrary size. The protocol is described in Fig. 7A. It consists in using a modified primer (*PrimerSpacer*) to amplify by PCR a segment of dsDNA with a flap. The flap and ds segments are complementary to the hairpin that is tethered between the surface and the bead. Force cycles are applied to the hairpin with the presence in solution of the flapped dsDNA. When the force is lowered below 10 pN, the DNA hairpin can close and the flap of the PCR product can hybridize to the remaining ss region on the handle of the hairpin. Then, the PCR product can invade the tethered hairpin. Indeed, while the hairpin opens and the PCR product hybridizes, the number of hybridized bases stays constant but the bead moves in the direction of the force, which is energetically favored. Thus, the invasion spontaneously occurs, as shown in Fig. 7A, resulting in the formation of the three-way junction. Fig. 7B shows a typical trace of hairpin invasion and the three way junction resolution by RecG as an example.

2.6 Recombinant protein production by double affinity purification

Recombinant proteins used in single-molecule experiments usually go through column chromatography steps (gel filtration, size exclusion, ion exchange...), in order to ensure RNAse free, pure and properly folded active proteins (Bianco, 2021). Yet, due to low yields, this strategy is hard to implement for unstable and lowly expressed proteins bearing unstructured domains. Even if low amounts of protein are needed (10–100 nM) to obtain single events in magnetic tweezers, the double affinity purification strategy is a good compromise between yield and purity. Here, we describe an optimized protocol for this strategy, based on our previous work (Fiorini, Bonneau, & Le Hir, 2012).

Briefly, coding sequences of interest were cloned into a pET28a vector allowing fusing a Calmodulin Binding Peptide (CBP) sequence in 5' and a



Fig. 7 Protocol for a fast *in situ* assembly of arbitrary long three-way junctions onto a tethered hairpin. (A) Sketch of the experiment. (B) Corresponding experimental traces of RecG. Force cycles (1)–(2) are applied to repeatedly open and close a hairpin in presence of 18 nM of a flapped 600 bp dsDNA complementary to the hairpin, until the dsDNA binds to the hairpin (3) and spontaneously invades the hairpin, at time 510 s, to form a three-way junction (4)–(5). The substrate is then ready to study fork regression by an enzyme such as RecG (1750 s) (Manosas et al., 2013). 5 nM RecG, 25 mM TrisOAc (pH 7.50), 150 mM KOAc, 5 mM MgOAc, 1 mM dithiothreitol and 1 mM ATP.

hexahistidine sequence in 3' of the expressed recombinant proteins. The plasmids were then transformed into BL21(DE3)-CodonPlus *Escherichia coli* competent bacteria by heat shock and selected on kanamycin containing Luria Broth (LB) plates. Bacteria were then pre-cultivated in LB liquid media containing kanamycin for 4h at 37 °C and then used for inoculation of 1L of LB Broth Auto Inducible (LBAI) media containing metals.
Bacteria were then left to grow overnight at $18 \,^{\circ}$ C. After bacterial lysis, the double purification was performed first on Nickel-NTA resin then on Calmodulin-coupled resin. Proteins were dialyzed against a storage buffer and frozen at $-80 \,^{\circ}$ C. One advantage of this technique is that, by purifying by both extremities, only full-length proteins are selected and thus, no size exclusion chromatography is needed. This protocol can be adapted for the use of other tags such as GST or Twin-Strep instead of CBP, the fusion of highly structured tags may help the solubility of unstructured proteins.

2.6.1 Protein overexpression

1. Transform 100 μ L of BL21(DE3)-CodonPlus competent *Escherichia coli* bacteria (Agilent) with 500 ng of plasmid by 2 min at 42 °C heat shock. Add 1 mL liquid LB media and incubate 1 h at 37 °C. Pellet the cells by 4 min centrifugation at 3500 × g. Resuspend the pellet in 50 μ L LB media, plate in kanamycin containing LB plates and leave to grow overnight at 37 °C.

Tip: Other competent bacteria designed for protein overexpression such as BL21(DE3)-Rosetta strain may be tested depending on the expressed proteins.

- 2. Resuspend 2–3 colonies in 25 mL LB liquid media containing kanamycin in a 250 mL Erlenmeyer flask and incubate at 37 °C, 200 rpm for 4 h.
- **3.** Inoculate the culture in 1L of LBAI containing metals (Formedium) added with kanamycin and chloramphenicol in a 5L Erlenmeyer flask and incubate at 18 °C, 200 rpm overnight.

Tip: Other culture media may be used depending on the expressed proteins. For non-inducible media, induction is performed by addition of 0.2–0.5 mM of IPTG when optic density reaches the exponential phase. Although we highly recommend the use of auto-inducible media (Luria Broth or Terrific Broth), with or without metals, protein yields may vary upon culture conditions.

4. Pellet the culture by $10 \min$, $4000 \times g$ centrifugation at 4 °C.

Tip: Pellets can be stored at -20 °C. However, we recommend using fresh cultures for purification.

2.6.2 First affinity purification

All following steps must be performed at 4 °C.

1. Resuspend the pellet in 25 mL of Lysis buffer (1.5X PBS, 1 mM MgAc₂, 0.1% NP-40, 20 mM imidazole, 10% glycerol, 2 mM DTT).

- **2.** Add 1X of protease inhibitors (Apoprotinine, Leupeptine, PMSF, Pepstatin).
- 3. Sonicate in a metal beaker at 30% amplitude and 1s pulse for 4min. Tip: So temperature does not exceed 15°C, the beaker can be placed on a mix of dry ice and ice.
- 4. Pellet the bacterial debris for 30 min at 18,000 rpm.
- Retrieve the supernatant and mix in a Falcon tube with 1mL of Nickel-NTA Agarose slurry (Qiagen) equilibrated with Lysis Buffer. Incubate in a rotator for 1 h at 4°C.
- 6. Centrifuge at 4,000 rpm for 5 min at 4 °C to pellet the beads.
- **7.** Remove the supernatant (Flow Through) and wash with 25 mL of Lysis Buffer.
- 8. Transfer into a Poly-Prep chromatography column (Bio-Rad) preequilibrated with 1 mL of Lysis Buffer.
- Wash the column with 5 mL of Lysis Buffer then with 5 mL of Wash Buffer (1.5X PBS, 250 mM NaCl, 1 mM MgAc₂, 0.1% NP-40, 50 mM imidazole, 10% glycerol, 2 mM DTT) then again with 5 mL of Lysis Buffer.
- Close the column and incubate the resin with 1 mL of Elution Buffer (1.5X PBS, 1 mM MgAc₂, 0.1% NP-40, 250 mM imidazole, 10% glycerol, 2 mM DTT) for 10 min at 4 °C.
- **11.** Elute successively with 1 mL of Elution Buffer and measure protein concentration by Bradford Assay.
- 12. Pool the fractions with highest concentration and dialyze overnight against 1L of Calmodulin Binding Buffer (CBB) (10mM Tris-HCl pH 7.5, 250mM NaCl, 1mM MgAc₂, 4mM CaCl₂, 0.05% NP-40, 10% glycerol, 2mM DTT) at 4°C.

2.6.3 Second affinity purification

- 1. Retrieve the dialyzed fractions and mix with $500\,\mu\text{L}$ of Calmodulin Sepharose 4B resin (50% Slurry, GE Healthcare) washed with CBB. Incubate in a rotor for 2h at 4°C.
- 2. Transfer into a Poly-Prep chromatography column (Bio-Rad) preequilibrated with 1 mL of CBB.
- 3. Wash with $5 \,\mathrm{mL}$ of CBB.
- 4. Close the column and incubate with 500 μL of Calmodulin Elution Buffer (10 mM Tris–HCl pH 7.5, 250 mM NaCl, 1 mM MgAc₂, 0.05% NP-40, 20 mM EGTA, 10% glycerol, 2 mM DTT) for 10 min at 4 °C.

- 5. Elute successively with $500\,\mu$ L of Calmodulin Elution Buffer and measure protein concentration by Bradford Assay.
- 6. Pool the highest concentrated fractions and dialyze overnight against 1 L of Storage Buffer (1.5X PBS, 1 mM MgAc₂, 10% glycerol, 2 mM DTT).
- 7. Retrieve the dialyzed proteins and aliquot them in small volumes $(10-100\,\mu\text{L}, \text{ depending on the final protein concentration})$. Store at $-80\,^{\circ}\text{C}$.

Tips

- 10 µL of Supernatant, Pellet, Flow Through Histidine, Wash Histidine, Elution Histidine, Dialysis Histidine, Flow Through Calmodulin, Elution Calmodulin and Dialysis Calmodulin may be kept for loading a 10% SDS-PAGE gel, revealed by Coomassie Blue coloring in order to visualize the different purification steps.
- Reducing agents must be added to the buffers extemporaneously.
- Buffers may have to be adapted for each purified protein depending on their stability. Some proteins may need more or less salt, no glycerol, or different pH.
- If the protein precipitates during dialysis, centrifuge and measure the concentration again.
- For higher reproducibility during single molecule experiments, aliquots should be only thawed once before each experiment. Leftovers should be discarded after thawing.

3. Advanced protocols for helicase characterization using magnetic tweezers

In previous reviews (Hodeib et al., 2016; Manosas et al., 2010), we described classical magnetic tweezers protocols for helicase studies and explained how the processivity, velocity and directionality can be extracted by dynamically measuring the extension of a hairpin unwound by the helicase. Here, we will focus on recent experimental advances we developed for measuring other helicase properties.

First, we will describe two assays for measuring translocation of a helicase in a force-independent manner: the peeling assay and the bubble assay. Then, we will focus on the binding capacities of helicases by describing a qualitative binding assay using force cycles and a more quantitative assay using a fluctuating short hairpin. The latter allows measuring binding kinetics independently of their unwinding activity and in a fluorescence-free manner. Lastly, we will focus on the impact of roadblocks, either NA-binding proteins or NA tertiary structures, on the translocation of helicases by embedding the roadblock inside a hairpin and monitoring the time they take to either resolve or bypass the obstacle.

3.1 Mitigating the impact of force on helicase translocation

In classical magnetic tweezer assays, a constant force is applied. Typically, in order to observe a good quality signal, the applied force lies between 6 pN and 12/18 pN (DNA/RNA). At higher forces, the hairpin spontaneously opens, preventing the observation of helicase unwinding activity and at lower forces, the signal-to-noise ratio is too low. This assay draws criticism concerning passive (force-dependent activity) and active (force-independent activity) helicases. Indeed, if the helicase is passive (such as polymerase-coupled helicases) it is possible that at zero force, no activity would be measured. Active helicases, on the other hand, which would display activity even at zero force, may also be influenced by the applied force. Hence, the observed activity in these assays may differ from the helicase's behavior as measured in bulk assays. In order to go beyond this assumption, unwinding activity can be probed in other configurations. Here, we propose two assays to mitigate the impact of force on helicase assays in magnetic tweezers: the peeling assay and the bubble assay.

3.2 Peeling assay

The peeling assay consists of hybridizing a lateral oligonucleotide on the side of the main hairpin, on which the helicase translocates (Fig. 8A). In this configuration, an increasing force hinders the unwinding, instead of favoring it since the bead moves opposite to the force when the enzyme unwinds the hairpin. The unwinding helicase activity is thus in an energetic context closer to zero-force bulk experiments. It is noteworthy that this configuration induces a specific biochemical context in which the helicase is confronted with two forks, which might affect its translocation. Fig. 8B (inset, 1235s) shows an example of peeling assay of the catalytic core of *E.coli* RecQ, RecQ- Δ C. Note that after the oligonucleotide is peeled, the helicase is able to translocate directly onto the hairpin as in classical constant-force experiments (1425–1432s).

Protocol

1. Using a microfluidic chamber (prepared as described in Section 2.2), run a force cycle for successive hairpin openings and closings (for example 2s at 18 pN, 4s at 8 pN) and start the recording of the beads' positions.



Fig. 8 Peeling assay: an alternative experiment to mitigate the impact of the force on helicase translocation. (A) Sketch of the experiment. (B) Corresponding experimental traces of ReqQ- Δ C. Force cycles (1)–(2) are applied to repeatedly open and close a hairpin in presence of 100 nM of a 3' flapped ssDNA complementary to the hairpin, until the ssDNA binds to the hairpin (3). The hairpin is then closed (4) and RecQ- Δ C injected at a concentration of 100 pM (5). The helicase binds to the flap and unwinds the oligonucleotide, resulting in the progressive closing of the hairpin (6 and inset of the trace), until it completely removes the oligonucleotide at 1235s with an unwinding velocity of 97 ± 10 bp/s (7). At 1425s, another enzyme binds and unwinds the hairpin in the standard configuration with an unwinding velocity of 95 ± 0.3 bp/s (8–9). Tris–HCl (pH 7.60), 50 mM NaCl, 3 mM MgCl₂, 1 mM DTT and 1 mM ATP.

- 2. Remove the buffer from the microfluidic chamber inlet with a pipette. Rinse the inlet three times with $200 \,\mu\text{L}$ of reaction buffer (no ATP, no enzyme).
- 3. Flow $160 \,\mu\text{L}$ of reaction buffer at a constant speed of $10 \,\mu\text{L/min}$ to rinse the microfluidic channel.
- 4. Flow $100 \,\mu\text{L}$ of reaction buffer + $100 \,\text{nM}$ of complementary oligonucleotide.

Novel approaches to study helicases using magnetic tweezers

- **5.** Wait for the oligonucleotide to be hybridized to all hairpins. This results in a partial blocking of the hairpins at 7 pN in opening/closing cycles.
- 6. Repeat steps 2-3 to remove the excess oligonucleotides.
- 7. Stop the force cycles and measure the beads' positions at constant medium force (F = 8 pN).
- 8. Add the buffer mixed with the enzyme and 1 mM ATP to the inlet of the chamber. Flow at a constant speed of $5\,\mu$ L/min.
- 9. Record data during the desired period of time.

Tips

- Start the thermalization of your setup early enough before the recording, at the desired temperature to avoid slow measurement drifts due to temperature changes.
- In addition to adding reducing agents such as TCEP or DTT to the solution, you may degas the buffer to minimize protein oxidation: leave the tube containing the buffer without ATP nor enzyme in a vacuum desiccator at room temperature for 10 min.
- Between each rinsing step, carefully remove most of the buffer from the inlet. Only a few μ L should remain in the inlet so that each rinsing step allows the dilution of the previous buffer by a factor ~100.
- When preparing the enzyme mix, gently stir with the pipette tip. Avoid the insertion of air and bubble formation when mixing in order to avoid the buffer's fast re-oxygenation as well as the concentration of proteins at air-water interfaces.

Data analysis

- **1.** Find the unwinding events in the data: they consist of a progressive decrease of the bead extension followed by an abrupt rezipping of the hairpin once the full oligonucleotide has been removed.
- 2. Measure the speed of the enzyme in nm/s for each event by dividing the total loss of extension during the progressive phase by the duration of this phase.
- 3. Convert the speed from nm/s to bp/s. The conversion factor is $l_{ssNA}(F) + l_{dsNA}(F)$, since one unwound base results in removing one ds base and one ss base from the extended molecule.

3.3 Bubble assay

This assay is performed using a long hairpin (>1 kb) that is able to rehybridize after the helicase's passage and thus encircle the helicase in a "bubble." By periodically changing the force as shown in Fig. 9A, the hairpin is open at high force (>15 pN), then at medium force (\sim 7–8 pN), it starts



Fig. 9 Bubble assay: an alternative protocol to measure helicase progression inside a dsDNA at zero force. (A) Sketch of the experiment. (B) Corresponding experimental traces of human UPF1 (hUPF1) helicase core. Force cycles are applied to repeatedly open (1), close (3) and encircle the helicase (4). The translocation velocity of the helicase stays approximately constant over the whole experiment (blue dashed line, linear fit of steps (3) where the position of the bead depends on the position of the helicase). Thus, hUPF1 keeps translocating roughly at the same speed when it is encircled. Note that the time it takes for hUPF1 to be encircled increases with the encircling force. Buffer: 20mM Tris–HCl pH 7.5, 75mM KOAc, 3mM MgCl₂, 1% BSA, 1mM DTT, 1mM ATP, 30nM yUPF1-HD.

to close. When the enzyme is bound, the hairpin will only partially close until the enzyme's position. By imposing a low force (3-4 pN), the hairpin completely closes and encircles the helicase (if it's not ejected, see Section 3.4.1). At this step, no force is applied inside the hairpin, meaning that the helicase's behavior while it is encircled is measured at zero force. By comparing the bead's position over successive cycles, it is possible to

extrapolate the helicase's translocation while encircled. Fig. 9B, illustrates an example of peeling assay on the human helicase core of Upf1.

Protocol

- **1.** Follow steps 1–3 from protocol in Section 3.2.
- 2. Stop the force cycles and choose the force cycles for the experiment. For example 2s at 15pN, 5s at 8pN, 5s at 3pN.
- 3. Add the buffer mixed with the enzyme and 1 mM ATP to the inlet of the chamber. Flow at a constant speed of $5 \,\mu$ L/min.
- 4. When the inlet is almost empty, repeat step 3.
- 5. Record data during the desired period of time.

Data analysis

- 1. Measure the binding rate k_{on} of the enzyme to the substrate by drawing the distribution of times elapsed between the injection of the enzyme and the first blockage or the first unwinding event. These times will typically follow exponential distributions, of parameter $\tau = 1/k_{on}$.
- 2. If the binding rate is large compared to the time spent at high force, where the presence of the enzyme cannot be probed, decrease the enzyme concentration to ensure that a signal observed in two successive cycles is the same enzyme.
- **3.** Compute the probability of encircling the enzyme during re-hybridization by dividing the number of events where the enzyme is present both before and after the re-hybridization step by the total number of events where the enzyme is present before the re-hybridization step.

3.4 Estimating NA binding dynamics of helicases

3.4.1 Qualitative single-molecule binding assay (SMBA)

One important aspect of helicase dynamics is their binding to NA and their residence time. Standard single-molecule unwinding assays allow measuring the dissociation time of the helicase from its NA substrate in the presence of ATP, when the enzyme translocates. However, in order to decouple this dissociation from the conformational changes induced by ATP hydrolysis, it is sometimes useful to be able to measure and compare the grip of the helicase on its NA substrate in absence of ATP.

In Kanaan et al. (2018), we described an ATP-free qualitative singlemolecule binding assay (SMBA, Fig. 10B and C), to estimate helicase binding to NA and their residence time. In this assay, the protein is injected without ATP allowing binding but not translocation. Force cycles are then applied to test the persistence of the helicase despite the mechanical constraints of a closing hairpin. First, a high force opening the hairpin (>15pN for DNA and



Fig. 10 Sketch of the single molecule binding assay (SMBA) protocol allowing qualitative estimation of a helicase's grip onto nucleic acids. (A) Sketch of the experiment. (B, C) Corresponding experimental traces. The numbers corresponding to the hairpin state in panel (A) were added to the traces. Yeast UPF1 catalytic helicase core (yUpf1 HD) and IGHMBP2 helicase domain were subjected to force cycles without the addition of ATP. yUPF1 HD blocks the hairpins closure with its tight grip, unlike IGHMBP2 HD that slides and is ejected by the hairpins closure. *Panels (A) and (B): Reproduced from Kanaan, J., Raj, S., Decourty, L., Saveanu, C., Croquette, V., & Hir, H. L. (2018). UPF1-like helicase grip on nucleic acids dictates processivity.* Nature Communications, 9(1), 3752. doi:10.1038/s41467-018-06313-y, in agreement with the Creative Commons Attribution 4.0 International License (http:/creativecommons.org/licenses/by/4.0/)

>18 pN for RNA) is applied for a couple of seconds, the helicase is thus able to bind randomly anywhere in the single-stranded regions. Then, a medium force (~7–8 pN) induces the hairpin's closure. The closing fork pressing against the bound helicase can have two outcomes: if the helicase's grip is strong enough, the hairpin is blocked at the helicase's position, else, the helicase is ejected and the hairpin completely closes (Fig. 10A). The ejection time can be almost immediate or last a few tenths of seconds. Thus, this step is held for 20–30 s or more. Fig. 10B and C shows a qualitative comparison of the grip of two different catalytic cores from Superfamily 1 helicases: yeast UPF1 and yeast IGHMBP2 (Kanaan et al., 2018).

Protocol

- **1.** Follow steps 1–3 from protocol in Section 3.2.
- 2. Stop the force cycles and choose the force cycles for the experiment: 1–3s at 15–18 pN, 20–30s at 7–8 pN.
- 3. Add the buffer mixed with the enzyme without ATP to the inlet of the chamber. Flow at a constant speed of $5\,\mu$ L/min.
- 4. When the inlet is almost empty, repeat step 3.
- 5. Record data during the desired period of time.

Tips

 In order to avoid multiple helicases binding to the same substrate, the lowest concentration possible is injected and can vary from 1 to 20 nM, depending on protein activity.

Data analysis

- 1. The binding time corresponds to the inverse of the ejection rate. Sum the blocking time of the closing hairpin at intermediate force and divide it by the binding events during the recording.
- **2.** Relative errors are calculated by computing the square root of the number of unbinding events.

3.4.2 Quantitative binding kinetics using kinetic locking assay

In some cases, the association of a helicase with its substrate does not cause changes of NA extension that are sufficiently large to be measured with single-molecule micromanipulation techniques. Measurement of helicases' binding kinetics is thus often performed with bulk stopped-flow assays (Kocsis, Sarlós, Harami, Martina, & Kovács, 2014). However, such bulk assays involve complex fitting procedures, and may require to label the proteins and/or to introduce chemical competitors to evaluate dissociation kinetics. We recently showed (Rieu, Valle-Orero, et al., 2021) that it is possible to use a fluctuating hairpin to perform single-molecule measurements of helicases' binding kinetics at a few tens of milliseconds timescales. Indeed, small hairpins (10–15 bases) fluctuate rapidly when they are pulled at the force that equilibrates the free energy level of the open and closed states ($F_{1/2}$) (Woodside et al., 2006). When the helicase binds to the hairpin, it keeps the hairpin open and prevents its fluctuations (Fig. 11A). The association and dissociation rates of the enzymes from the substrates can be extracted from the duration and the frequencies of these long pauses. Fig. 11B and C shows the binding kinetics of the catalytic core of RecQ helicase in this configuration (Rieu, Valle-Orero, et al., 2021).

Protocol

- 1. Using a microfluidic chamber (prepared as described in Section 2, run a force ramp (5 s per force level, steps of 0.5 pN).
- **2.** Find beads which fluctuate rapidly between the open and closed states. Change the force to $F_{1/2}$, the force where the duration of the closed and open states are equal.
- 3. Increase the camera frequency up to a few kHz.
- **4.** Rinse the microfluidic chamber with the reaction buffer as explained in previous protocols.
- 5. Add the buffer mixed with the enzyme to the inlet of the chamber. Flow at a constant speed of $5 \mu L/min$.
- 6. Record data during the desired period of time.
- 7. Repeat step 5 with increasing helicase concentrations.

Tips

- 1. Since the fluctuations only happen in a limited range of force ($\sim 0.5-1 \, \text{pN}$ around $F_{1/2}$) and beads have some variability in their magnetization ($\sim 10-20\%$), it is hard to find many beads that fluctuate at the same magnet position, especially at high acquisition frequencies where the field of view is limited by the camera's transfer frequency. However, since the fluctuations are fast, one or two beads are usually sufficient to sample several hundreds of events.
- 2. It is important to unambiguously distinguish between the two states of the hairpins (change of 10 nm at the millisecond timescale). A magnetic tweezer setup with high-spatio-temporal resolution is crucial (Dulin et al., 2015; Huhle et al., 2015; Kim & Saleh, 2009; Rieu et al., 2021). A signal of poor quality may be improved by removing the spurious correlations between the axial and transverse movements of the beads, as described in the supplementary materials of Rieu, Vieille, et al. (2021).



Fig. 11 See figure legend on next page.

- **3.** The temperature of the sample must be precisely controlled (~0.01 °C) since the kinetics of the hairpin fluctuations are sensitive to it.
- 4. Maintaining a constant concentration of helicase is critical to assess the association time. For this reason, keeping a constant flow during the experiments is important. If the enzyme has a strong affinity to DNA, we recommend not using DNA-coated azide surfaces but other attachment chemistry, such as DIG, to avoid biasing the local concentration of helicase close to the surface.

Data analysis

Analysis of the freely fluctuating hairpin (no helicase):

- 1. Compute the standard deviation σ of the measured position in both closed states and open states. At 1 kHz of acquisition frequency and a 10 bp hairpin with 120 bp dsDNA handles, they typically lie between 1 and 2 nm.
- 2. Automatically detect the transitions between the closed and open states. A criterion of $3-4 \sigma$ (~6-8 nm) can be used by detecting a jump if the difference of position measured between two successive frames is larger than 3σ . To avoid biases caused by the finite velocity of the bead during the transition, sum up all differences of position among successive frames as long as the movement is monotonous (series of frames with the position of the bead either increasing or decreasing).
- **3.** Compute the times spent in the closed states (difference between two successive closing and opening events) and the times spent in the open states (difference between two successive closing and opening events).

Fig. 11 Single-molecule measurement of the binding kinetics of a helicase without ATP through kinetic locking. (A) Sketch of the experiment. A rapidly fluctuating 10 bp hairpin allows measuring short binding times of RecQ-ΔC through the alteration of the folding/ unfolding rates induced by the association of the helicase. (B) Corresponding force-extension curve of RecQ-ΔC. When the helicase binds to the hairpin, it blocks it in its open state for a longer time (red). (C) Distributions of open times as a function of the concentration of helicases. Fitting to double-exponential distributions allows retrieving the association and dissociation kinetic rates. *Panels (B) and (C): Reproduced and unmodified from Rieu, M., Valle-Orero, J., Ducos, B., Allemand, J.-F., & Croquette, V. (2021). Single-molecule kinetic locking allows fluorescence-free quantification of protein/nucleic-acid binding. Communications Biology, 4(1), 1083. doi:10.1038/s42003-021-02606-z, in agreement with the Creative Commons Attribution 4.0 International License (http:/creativecommons.org/licenses/by/4.0/).*

- 4. Draw the histograms of multiple (>100) closed and open times (the bin size should be a multiple of the acquisition frequency to avoid spurious irregularities). Fit the histograms by an exponential distribution and infer the typical times spent in the close and open states.
- 5. Average the closing and opening times by chunks of N points (N > 100) and draw the averages as a function of the acquisition time. Check that the averages are constant within the statistical error (square root of N). If an important drift is observed over time, this might be due to a flow that is not constant or a change of temperature. Wait until this parameter is stabilized and restart the measurement.

Derivation of the binding kinetics

- 1. When the helicase is present in solution and can bind to the substrate, the association and dissociation processes of the helicase of rates k_{on} and k_{off} adds up to the internal fluctuating rates of the hairpins. This results in the distribution of the open times of the hairpin changing from single exponential to double exponential. By fitting the time spent in the open times at the different concentrations of helicases as described in (Rieu, Valle-Orero, et al., 2021), one can deduce k_{on} and k_{off} .
- 2. Noteworthy, if k_{off} is much smaller (~10 fold) than the spontaneous closing rate of the hairpin, k_{off} will be very close to the inverse of the long-time of the double-exponential distribution (either the hairpin closes spontaneously—short time—or it stays open due to the presence of the helicase—a long time). In this case, the distributions are particularly easy to interpret.

3.4.3 Helicase interaction with roadblocks

We have previously described how some helicases are capable of remodeling ribonucleoprotein complexes (Fiorini, Bagchi, Le Hir, & Croquette, 2015) by covering a ssDNA hairpin with NA-binding proteins and analyzing the progression of the helicase in this context. In Fiorini et al. (2015), we demonstrated that the human UPF1 helicase core is able to translocate onto a Gp32-covered DNA at a normal rate, by remodeling NA-protein interactions. Recently, we developed an encirclement-based protocol to embed roadblocks, such as tertiary NA structures, in the NA hairpin (Tran et al., 2021) shown in Fig. 12A. The dynamics of a helicase interaction with such roadblocks as well as their probability of removing or bypassing them can be studied by injecting the enzyme and analyzing its stalling at the roadblock's position. The time spent at this position can be calculated in order to analyze



Time

Fig. 12 Protocol for the single-molecule observation of a helicase's interaction with a roadblock. (A) Sketch of the experiment. (B) Typical extension traces and interpretation. (1) A hairpin is opened by applying a high force on the magnetic bead (20 pN). (2) The roadblock is formed. It can consist in the deposition of a NA-binding protein present in solution or in the formation of a NA-tertiary structure. (3) At a force of \sim 8 pN, the hairpin can only partially close due to the presence of the roadblock. This allows probing the presence of the roadblock. (4) The force is then lowered to a force (generally 2–5 pN) allowing encircling the roadblock inside a double-stranded hairpin. (5) The force is then

how the helicase interacts with the roadblock. This protocol has been recently used to study the interaction of the *Saccharomyces cerevisiae* helicase Pif1 with G-quadruplexes (Valle-Orero et al., 2022). Fig. 12B shows a schematic trace of the roadblock embedding and a helicase's translocation in this configuration.

Protocols

- 1. Follow steps 1–3 from protocol in Section 3.1.
- 2. Add the roadblock in solution and observe its binding/formation at the 7–8 pN phase (it will cause a transient blockage of the closing of the hairpin). Check that the roadblock is not expelled during the low force phase and is stable when encircled in the hairpin.
- 3. Rinse the chamber with the reaction buffer.
- 4. Stop the force cycles.
- **5.** Apply a high force to open the hairpin, then go to 7 pN to close the hairpin until the roadblock position and finally to your working force at which the helicase activity is analyzed.
- 6. Apply a low force (2-4 pN) to encircle the roadblock.
- **7.** From the low force, go to the working force, where the hairpin should be closed.
- 8. Add the buffer mixed with the enzyme to the inlet of the chamber. Flow at a constant speed of $5 \,\mu$ L/min.
- 9. Record data during the desired period of time.

Data analysis

- Detect unwinding events corresponding to molecules whose extensions increase.
- 2. Check whether the helicase stalls at the roadblock position by comparing the time spent by the helicase at the roadblock position at the working force (checked at step 6 of the protocol) with the average time spent at any other position during unwinding.

increased to the working force of the helicase. The helicase is injected in solution. While the helicase unwinds the hairpin, the extension of the molecule increases. If the roadblock prevents the translocation of the helicase, a pause is observed in this increase of extension. If the helicase stalls and passes through the roadblock, one records the bypass time (in purple), which probability distribution is the main quantitative parameter provided by the assay. After the passage of the helicase, either the roadblock is still present (blue) or it has been removed (green). If it is still here, provided that the working force is too high to allow the encirclement of the roadblock, the hairpin will only partially close. Measuring several events thus allows accessing the probability of resolution.

- **3.** If the helicase stalls and then resumes translocation, draw the distribution of the stalling duration over multiple events. It will typically follow an exponential distribution, whose characteristic time will be the "bypass time" i.e. the time needed for the helicase to go through the roadblock.
- 4. Check whether the helicase removes the roadblock. For this purpose, choosing a working force that prevents the full closing of the hairpin in the presence of the roadblock is recommended. In the absence of a roadblock, the hairpin will close completely after the passage of the helicase. If the roadblock remained and was just bypassed by the enzyme, the hairpin will only partially close after the passage of the helicase.

4. Discussion

The various elements that we have described demonstrate the wealth of possibilities that have emerged to study helicases in a single molecule configuration. This panel of approaches is in part the result of the richness of NA structures (ssDNA, dsDNA, three-way or four-way junction). The only limitation seems to be our imagination. As a result, depending upon the question investigated, it is possible to choose among various possibilities to either achieve the strongest signal of the simplest experimental design. What we have described for magnetic tweezers can also be used in most situations with optical tweezers.

We have not discussed nanopore-related technologies applied to study helicases, although they offer very interesting features: thanks to their very high spatial resolution, single-molecule measurements of helicase stepping are extremely promising (Caldwell & Spies, 2017; Craig et al., 2019; Laszlo, Derrrington, & Gundlach, 2017). Their main advantage is the ability to study protein translocation on hairpin-free ssNA while measuring rates of translocation at a nucleotide resolution. Micromanipulation techniques require a hairpin that converts translocation into a change of extension and recent studies have achieved the accuracy of a single base (Cheng, Arunajadai, Moffitt, Tinoco, & Bustamante, 2011; Qi, Pugh, Spies, & Chemla, 2013; Rieu, Vieille, et al., 2021), but they are not yet as precise as nanopores.

Altogether, the single molecule approaches described here nicely complement bulk biological tools to study helicases' activities which are not simple to characterize since their unwinding effect can easily be masked by the re-hybridization of NA if nothing is done to prevent it. Novel approaches to study helicases using magnetic tweezers

References

- Abdelhaleem, M. (2010). Helicases: An overview. Methods in Molecular Biology (Clifton, N.J.), 587, 1–12. https://doi.org/10.1007/978-1-60327-355-8_1.
- Abdelkrim, Y. Z., Banroques, J., & Kyle Tanner, N. (2021). Known inhibitors of RNA helicases and their therapeutic potential. *Methods in Molecular Biology (Clifton, N.J.)*, 2209, 35–52. https://doi.org/10.1007/978-1-0716-0935-4_3.
- Allemand, J.-F. (1997). Micromanipulation d'une molécule individuelle d'ADN. Université Pierre et Marie Curie Paris VI.
- Amit, R., Gileadi, O., & Stavans, J. (2004). Direct observation of RuvAB-catalyzed branch migration of single Holliday junctions. Proceedings of the National Academy of Sciences of the United States of America, 101(32), 11605–11610. https://doi.org/10.1073/pnas.0404332101.
- Bétous, R., Couch, F. B., Mason, A. C., Eichman, B. F., Manosas, M., & Cortez, D. (2013). Substrate-selective repair and restart of replication forks by DNA translocases. *Cell Reports*, 3(6), 1958–1969.
- Bianco, P. R. (2021). Insight into the biochemical mechanism of DNA helicases provided by bulk-phase and single-molecule assays. *Methods*. https://doi.org/10.1016/j.ymeth.2021. 12.002.
- Bianco, P. R., & Lu, Y. (2021). Single-molecule insight into stalled replication fork rescue in Escherichia coli. Nucleic Acids Research, 49(8), 4220–4238. https://doi.org/10.1093/nar/ gkab142.
- Bockelmann, U. (2004). Single-molecule manipulation of nucleic acids. *Current Opinion in Structural Biology*, 14(3), 368–373.
- Bouchiat, C., Wang, M. D., Allemand, J.-F., Strick, T., Block, S. M., & Croquette, V. (1999). Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophysical Journal*, 76(1), 409–413. https://doi.org/10.1016/S0006-3495 (99)77207-3.
- Brenner, H. (1961). The slow motion of a sphere through a viscous fluid towards a plane surface. *Chemical Engineering Science*, 16(3), 242–251. https://doi.org/10.1016/0009-2509(61)80035-3.
- Brosh, R. M., & Matson, S. W. (2020). History of DNA helicases. Genes, 11(3), 255. https:// doi.org/10.3390/genes11030255.
- Bryant, Z., Stone, M. D., Gore, J., Smith, S. B., Cozzarelli, N. R., Bustamante, C., et al. (2003). Structural transitions and elasticity from torque measurements on DNA. *Nature*, 424(6946), 338–341.
- Bustamante, C. J., Chemla, Y. R., Liu, S., & Wang, M. D. (2021). Optical tweezers in single-molecule biophysics. *Nature Reviews Methods Primers*, 1(1), 1–29. https://doi.org/ 10.1038/s43586-021-00021-6.
- Caldwell, C. C., & Spies, M. (2017). Helicase SPRNTing through the nanopore. Proceedings of the National Academy of Sciences of the United States of America, 114(45), 11809–11811. https://doi.org/10.1073/pnas.1716866114.
- Cheng, W., Arunajadai, S. G., Moffitt, J. R., Tinoco, I., & Bustamante, C. (2011). Single-base pair unwinding and asynchronous RNA release by the hepatitis C virus NS3 helicase. *Science*, 333(6050), 1746–1749. https://doi.org/10.1126/science.1206023.
- Chilkoti, A., & Stayton, P. S. (1995). Molecular origins of the slow streptavidin-biotin dissociation kinetics. *Journal of the American Chemical Society*, 117(43), 10622–10628. https:// doi.org/10.1021/ja00148a003.
- Cordin, O., & Beggs, J. D. (2013). RNA helicases in splicing. RNA Biology, 10(1), 83–95. https://doi.org/10.4161/rna.22547.
- Craig, J. M., Laszlo, A. H., Nova, I. C., Brinkerhoff, H., Noakes, M. T., Baker, K. S., et al. (2019). Determining the effects of DNA sequence on Hel308 helicase translocation along single-stranded DNA using nanopore tweezers. *Nucleic Acids Research*, 47(5), 2506–2513.

- Datta, A., & Brosh, R. M. (2018). New insights into DNA helicases as druggable targets for cancer therapy. *Frontiers in Molecular Biosciences*, 5, 59. https://doi.org/10.3389/finolb. 2018.00059.
- Davis, L. G., Dibner, M. D., & Battey, J. F. (1986). Section 9-3—Electroelution. In L. G. Davis, M. D. Dibner, & J. F. Battey (Eds.), *Basic methods in molecular biology* (pp. 112–114). Elsevier. https://doi.org/10.1016/B978-0-444-01082-7.50035-7.
- Dawid, A., Croquette, V., Grigoriev, M., & Heslot, F. (2004). Single-molecule study of RuvAB-mediated Holliday-junction migration. *Proceedings of the National Academy of Sciences of the United States of America*, 101(32), 11611–11616.
- Desai, V. P., Frank, F., Lee, A., Righini, M., Lancaster, L., Noller, H. F., et al. (2019). Co-temporal force and fluorescence measurements reveal a ribosomal gear shift mechanism of translation regulation by structured mRNAs. *Molecular Cell*, 75(5), 1007–1019, e5. https://doi.org/10.1016/j.molcel.2019.07.024.
- Dhar, S., Datta, A., & Brosh, R. M. (2020). DNA helicases and their roles in cancer. DNA Repair, 96, 102994. https://doi.org/10.1016/j.dnarep.2020.102994.
- Dulin, D., Cui, T. J., Cnossen, J., Docter, M. W., Lipfert, J., & Dekker, N. H. (2015). High spatiotemporal-resolution magnetic tweezers: Calibration and applications for DNA dynamics. *Biophysical Journal*, 109(10), 2113–2125.
- Eeftens, J. M., van der Torre, J., Burnham, D. R., & Dekker, C. (2015). Copper-free click chemistry for attachment of biomolecules in magnetic tweezers. *BMC Biophysics*, 8(1), 9. https://doi.org/10.1186/s13628-015-0023-9.
- Fiorini, F., Bagchi, D., Le Hir, H., & Croquette, V. (2015). Human Upf1 is a highly processive RNA helicase and translocase with RNP remodelling activities. *Nature Communications*, 6, 7581. https://doi.org/10.1038/ncomms8581.
- Fiorini, F., Bonneau, F., & Le Hir, H. (2012). Chapter twelve—Biochemical characterization of the RNA helicase UPF1 involved in nonsense-mediated mRNA decay. In E. Jankowsky (Ed.), 511. Methods in enzymology (pp. 255–274). Academic Press. https://doi.org/10.1016/B978-0-12-396546-2.00012-7.
- Gao, Y., & Yang, W. (2020). Different mechanisms for translocation by monomeric and hexameric helicases. *Current Opinion in Structural Biology*, 61, 25–32. https://doi.org/ 10.1016/j.sbi.2019.10.003.
- Gorbalenya, A., & Koonin, E. V. (1993). Helicases: Amino acid sequence comparisons and structure-function relationships. *Current Opinion in Structural Biology*. https://doi.org/ 10.1016/S0959-440X(05)80116-2.
- Gosse, C., & Croquette, V. (2002). Magnetic tweezers: Micromanipulation and force measurement at the molecular level. *Biophysical Journal*, 82(6), 3314–3329. https://doi.org/ 10.1016/S0006-3495(02)75672-5.
- Greenleaf, W. J., Woodside, M. T., & Block, S. M. (2007). High-resolution, single-molecule measurements of biomolecular motion. *Annual Review of Biophysics and Biomolecular Structure*, 36, 171–190.
- Guo, S., Ray, C., Kirkpatrick, A., Lad, N., & Akhremitchev, B. B. (2008). Effects of multiple-bond ruptures on kinetic parameters extracted from force spectroscopy measurements: Revisiting biotin-streptavidin interactions. *Biophysical Journal*, 95(8), 3964–3976. https://doi.org/10.1529/biophysj.108.133900.
- Hatch, K., Danilowicz, C., Coljee, V., & Prentiss, M. (2008). Demonstration that the shear force required to separate short double-stranded DNA does not increase significantly with sequence length for sequences longer than 25 base pairs. *Physical Review E*, 78(1), 011920. https://doi.org/10.1103/PhysRevE.78.011920.
- Hodeib, S., Raj, S., Manosas, M., Zhang, W., Bagchi, D., Ducos, B., et al. (2016). Single molecule studies of helicases with magnetic tweezers. *Methods*, 105, 3–15.

- Huhle, A., Klaue, D., Brutzer, H., Daldrop, P., Joo, S., Otto, O., et al. (2015). Camera-based three-dimensional real-time particle tracking at kHz rates and Ångström accuracy. *Nature Communications*, 6(1). https://doi.org/10.1038/ncomms6885.
- Jacobson, D. R., McIntosh, D. B., Stevens, M. J., Rubinstein, M., & Saleh, O. A. (2017). Single-stranded nucleic acid elasticity arises from internal electrostatic tension. *Proceedings* of the National Academy of Sciences of the United States of America, 114(20), 5095–5100. https://doi.org/10.1073/pnas.1701132114.
- Janissen, R., Oberbarnscheidt, L., & Oesterhelt, F. (2009). Optimized straight forward procedure for covalent surface immobilization of different biomolecules for single molecule applications. *Colloids and Surfaces B: Biointerfaces*, 71(2), 200–207. https://doi.org/ 10.1016/j.colsurfb.2009.02.011.
- Kanaan, J., Raj, S., Decourty, L., Saveanu, C., Croquette, V., & Hir, H. L. (2018). UPF1-like helicase grip on nucleic acids dictates processivity. *Nature Communications*, 9(1), 3752. https://doi.org/10.1038/s41467-018-06313-y.
- Khemici, V., & Linder, P. (2018). RNA helicases in RNA decay. Biochemical Society Transactions, 46(1), 163–172. https://doi.org/10.1042/BST20170052.
- Kim, K., & Saleh, O. A. (2009). A high-resolution magnetic tweezer for single-molecule measurements. *Nucleic Acids Research*, 37(20), e136. https://doi.org/10.1093/nar/ gkp725.
- Klaue, D., & Seidel, R. (2009). Torsional stiffness of single superparamagnetic microspheres in an external magnetic field. *Physical Review Letters*, 102(2), 028302.
- Kocsis, Z. S., Sarlós, K., Harami, G. M., Martina, M., & Kovács, M. (2014). A nucleotidedependent and HRDC domain-dependent structural transition in DNA-bound RecQ helicase. *The Journal of Biological Chemistry*, 289(9), 5938–5949. https://doi.org/10.1074/ jbc.M113.530741.
- Laszlo, A. H., Derrrington, I. M., & Gundlach, J. H. (2017). Subangstrom measurements of enzyme function using a biological nanopore, SPRNT. *Methods in Enzymology*, 582, 387–414. Elsevier.
- Li, H., Zhang, C., Hu, Y., Liu, P., Sun, F., Chen, W., et al. (2021). A reversible shearing DNA probe for visualizing mechanically strong receptors in living cells. *Nature Cell Biology*, 23(6), 642–651. https://doi.org/10.1038/s41556-021-00691-0.
- Linder, P., & Fuller-Pace, F. (2015). Happy birthday: 25 years of DEAD-box proteins. Methods in Molecular Biology (Clifton, N.J.), 1259, 17–33. https://doi.org/10.1007/978-1-4939-2214-7_2.
- Linder, P., & Jankowsky, E. (2011). From unwinding to clamping—The DEAD box RNA helicase family. *Nature Reviews Molecular Cell Biology*, 12(8), 505–516. https://doi.org/ 10.1038/nrm3154.
- Lohman, T. M., Tomko, E. J., & Wu, C. G. (2008). Non-hexameric DNA helicases and translocases: Mechanisms and regulation. *Nature Reviews. Molecular Cell Biology*, 9(5), 391–401. https://doi.org/10.1038/nrm2394.
- Manosas, M., Meglio, A., Spiering, M. M., Ding, F., Benkovic, S. J., Barre, F.-X., et al. (2010). Magnetic tweezers for the study of DNA tracking motors. *Methods in Enzymology*, 475, 297–320. https://doi.org/10.1016/S0076-6879(10)75013-8.
- Manosas, M., Perumal, S. K., Bianco, P. R., Ritort, F., Benkovic, S. J., & Croquette, V. (2013). RecG and UvsW catalyse robust DNA rewinding critical for stalled DNA replication fork rescue. *Nature Communications*, 4(1), 2368. https://doi.org/10.1038/ ncomms3368.
- McIntosh, D. B., Duggan, G., Gouil, Q., & Saleh, O. A. (2014). Sequence-dependent elasticity and electrostatics of single-stranded DNA: Signatures of base-stacking. *Biophysical Journal*, 106(3), 659–666. https://doi.org/10.1016/j.bpj.2013.12.018.

- Mendoza, O., Bourdoncle, A., Boulé, J.-B., Brosh, R. M., & Mergny, J.-L. (2016). G-quadruplexes and helicases. Nucleic Acids Research, 44(5), 1989–2006. https://doi. org/10.1093/nar/gkw079.
- Merkel, R., Nassoy, P., Leung, A., Ritchie, K., & Evans, E. (1997). Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy. *Nature*, 397, 50–53.
- Mickolajczyk, K. J., Shelton, P. M. M., Grasso, M., Cao, X., Warrington, S. E., Aher, A., et al. (2021). Force-dependent stimulation of RNA unwinding by SARS-CoV-2 nsp13 helicase. *Biophysical Journal*, 120(6), 1020–1030. https://doi.org/10.1016/j.bpj.2020. 11.2276.
- Neuman, K. C., Lionnet, T., & Allemand, J.-F. (2007). Single-molecule micromanipulation techniques. Annual Review of Materials Research, 37, 33–67.
- Neuman, K. C., & Nagy, A. (2008). Single-molecule force spectroscopy: Optical tweezers, magnetic tweezers and atomic force microscopy. *Nature Methods*, 5(6), 491–505.
- Ostrofet, E., Papini, F. S., & Dulin, D. (2018). Correction-free force calibration for magnetic tweezers experiments. *Scientific Reports*, 8(1), 15920. https://doi.org/10.1038/s41598-018-34360-4.
- Papini, F. S., Seifert, M., & Dulin, D. (2019). High-yield fabrication of DNA and RNA constructs for single molecule force and torque spectroscopy experiments. *Nucleic Acids Research*, 47(22), e144. https://doi.org/10.1093/nar/gkz851.
- Patel, S. S., & Donmez, I. (2006). Mechanisms of helicases. Journal of Biological Chemistry, 281(27), 18265–18268. https://doi.org/10.1074/jbc.R600008200.
- Pincet, F., & Husson, J. (2005). The solution to the streptavidin-biotin paradox: The influence of history on the strength of single molecular bonds. *Biophysical Journal*, 89(6), 4374–4381. https://doi.org/10.1529/biophysj.105.067769.
- Qi, Z., Pugh, R. A., Spies, M., & Chemla, Y. R. (2013). Sequence-dependent base pair stepping dynamics in XPD helicase unwinding. *ELife*. https://doi.org/10.7554/eLife. 00334.
- Riener, C. K., Kienberger, F., Hahn, C. D., Buchinger, G. M., Egwim, I. O. C., Haselgrübler, T., et al. (2003). Heterobifunctional crosslinkers for tethering single ligand molecules to scanning probes. *Analytica Chimica Acta*, 497(1), 101–114. https://doi.org/ 10.1016/j.aca.2003.08.041.
- Rieu, M., Valle-Orero, J., Ducos, B., Allemand, J.-F., & Croquette, V. (2021). Single-molecule kinetic locking allows fluorescence-free quantification of protein/ nucleic-acid binding. *Communications Biology*, 4(1), 1083. https://doi.org/10.1038/ s42003-021-02606-z.
- Rieu, M., Vieille, T., Radou, G., Jeanneret, R., Ruiz-Gutierrez, N., Ducos, B., et al. (2021). Parallel, linear, and subnanometric 3D tracking of microparticles with Stereo Darkfield Interferometry. *Science Advances*, 7(6), eabe3902.
- Singleton, M. R., Dillingham, M. S., & Wigley, D. B. (2007). Structure and mechanism of helicases and nucleic acid translocases. *Annual Review of Biochemistry*, 76(1), 23–50. https://doi.org/10.1146/annurev.biochem.76.052305.115300.
- Smith, S. B., Finzi, L., & Bustamante, C. (1992). Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science*, 258(5085), 1122–1126. https://doi.org/10.1126/science.1439819.
- Steimer, L., & Klostermeier, D. (2012). RNA helicases in infection and disease. RNA Biology, 9(6), 751–771. https://doi.org/10.4161/rna.20090.
- Sun, Y., Sato, O., Ruhnow, F., Arsenault, M. E., Ikebe, M., & Goldman, Y. E. (2010). Single-molecule stepping and structural dynamics of myosin X. *Nature Structural & Molecular Biology*, 17(4), 485–491. https://doi.org/10.1038/nsmb.1785.
- Tan, H. Y., Wilczek, L. A., Pottinger, S., Manosas, M., Yu, C., Nguyenduc, T., et al. (2017). The intrinsically disordered linker of E. coli SSB is critical for the release from single-stranded DNA. *Protein Science*, 26(4), 700–717. https://doi.org/10.1002/pro.3115.

- Tran, P. L. T., Rieu, M., Hodeib, S., Joubert, A., Ouellet, J., Alberti, P., et al. (2021). Folding and persistence times of intramolecular G-quadruplexes transiently embedded in a DNA duplex. *Nucleic Acids Research*, 49(9), 5189–5201. https://doi.org/10.1093/ nar/gkab306.
- Umate, P., Tuteja, N., & Tuteja, R. (2011). Genome-wide comprehensive analysis of human helicases. *Communicative & Integrative Biology*, 4(1), 118–137. https://doi.org/10.4161/ cib.4.1.13844.
- Valle-Orero, J., Rieu, M., Tran, P. L. T., Joubert, A., Allemand, J.-F., Croquette, V., et al. (2022). Strand-switching mechanism of Pif1 helicase induced by its collision with a G-quadruplex embedded in dsDNA (p. 2022.01.11.475363). *bioRxiv*. https://doi.org/ 10.1101/2022.01.11.475363.
- Viader-Godoy, X., Pulido, C. R., Ibarra, B., Manosas, M., & Ritort, F. (2021). Cooperativity-dependent folding of single-stranded DNA. *Physical Review X*, 11(3), 031037. https://doi.org/10.1103/PhysRevX.11.031037.
- Vilfan, I. D., Lipfert, J., Koster, D. A., Lemay, S. G., & Dekker, N. H. (2009). Magnetic tweezers for single-molecule experiments. In P. Hinterdorfer, & A. Oijen (Eds.), *Handbook of single-molecule biophysics* (pp. 371–395). US: Springer. https://doi.org/ 10.1007/978-0-387-76497-9_13.
- Walsh, M. K., Wang, X., & Weimer, B. C. (2001). Optimizing the immobilization of single-stranded DNA onto glass beads. *Journal of Biochemical and Biophysical Methods*, 47(3), 221–231. https://doi.org/10.1016/S0165-022X(00)00146-9.
- Wilchek, M., & Bayer, E. A. (1990). Introduction to avidin-biotin technology. Methods in Enzymology, 184, 5–13. https://doi.org/10.1016/0076-6879(90)84256-g.
- Wildling, L., Unterauer, B., Zhu, R., Rupprecht, A., Haselgrübler, T., Rankl, C., et al. (2011). Linking of sensor molecules with amino groups to amino-functionalized AFM tips. *Bioconjugate Chemistry*, 22(6), 1239–1248. https://doi.org/10.1021/bc200099t.
- Williams, P. M., Moore, A., Stevens, M. M., Allen, S., Davies, M. C., Roberts, C. J., et al. (2000). On the dynamic behaviour of the forced dissociation of ligand–receptor pairs. *Journal of the Chemical Society, Perkin Transactions*, 2(1), 5–8. https://doi.org/10.1039/ A907750B.
- Woodside, M. T., Anthony, P. C., Behnke-Parks, W. M., Larizadeh, K., Herschlag, D., & Block, S. M. (2006). Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. *Science*, 314(5801), 1001–1004.
- Wu, Y. (2012). Unwinding and rewinding: Double faces of helicase? Journal of Nucleic Acids, 2012, 140601. https://doi.org/10.1155/2012/140601.
- Yang, Y.-J., Song, L., Zhao, X.-C., Zhang, C., Wu, W.-Q., You, H.-J., et al. (2019). A universal assay for making DNA, RNA, and RNA-DNA hybrid configurations for single-molecule manipulation in two or three steps without ligation. ACS Synthetic Biology, 8(7), 1663–1672. https://doi.org/10.1021/acssynbio.9b00241.
- Yodh, J. G., Schlierf, M., & Ha, T. (2010). Insight into helicase mechanism and function revealed through single-molecule approaches. *Quarterly Reviews of Biophysics*, 43(2), 185–217. https://doi.org/10.1017/S0033583510000107.
- Yu, Z., Dulin, D., Cnossen, J., Köber, M., van Oene, M. M., Ordu, O., et al. (2014). A force calibration standard for magnetic tweezers. *The Review of Scientific Instruments*, 85(12), 123114. https://doi.org/10.1063/1.4904148.

ABSTRACT

Gene expression fidelity and RNA quality control are crucial for cellular homeostasis. In eukaryotes, the Nonsense-mediated mRNA Decay (NMD) pathway serves as an essential surveillance process that detects and eliminates aberrant mRNAs harboring premature termination codons (PTCs). At the heart of this pathway lies the conserved RNA helicase Upf1 encompassing a helicase domain flanked by a cysteine- and histidine-rich regulatory domain (CH) and a disordered C-terminal domain. Upf1 orchestrates the recognition and the degradation of PTC-containing transcripts, although the precise mechanism by which Upf1 triggers mRNA decay remains elusive. Recent biochemical studies have uncovered two mutually exclusive complexes in Saccharomyces cerevisiae: the Upf1-detector complex consisting of Upf2 and Upf3 and the Upf1-decapping complex, including Dcp1, Dcp2, Nmd4 and Ebs1. This thesis project aimed to reconstitute the Upf1-decapping complex in vitro using recombinant proteins to characterize the role of Upf1 and its partners in the late steps of NMD. Notably, we identified a direct interaction between Dcp2 and the CH domain of Upf1, as well as direct interactions between different domains of Upf1 and Nmd4 and Ebs1, forming a stable Upf1-decapping complex. Similar to humans, yeast Upf2 also interacts with the CH domain. Interestingly, this interaction is incompatible with Dcp2 binding providing insight into the switch between the Upf1-detector and the Upf1-decapping complexes. Additionally, our investigations revealed that Upf1's RNA binding is nearly abolished in the presence of Upf2 but significantly enhanced within the Upf1-decapping complex. Thus, Upf1 facilitates the association of the decapping proteins on its target mRNAs, ensuring proper degradation. Furthermore, we explored the autoregulatory mechanisms and functional properties of yeast and human Upf1 helicase in RNA and DNA. Comparative single-molecule analyses using magnetic tweezers revealed contrasting autoregulatory behaviors of Upf1 helicase activity on RNA despite their high conservation. Altogether, this comprehensive study provides valuable insights into the functional importance of the Upf1 helicase to recruit the decay machinery to faulty mRNAs, contributing to our understanding of RNA metabolism.

Keywords: RNA surveillance, NMD pathway, Upf1 helicase, decapping, biochemistry, single molecule

RESUME

La fidélité de l'expression génétique et le contrôle qualité des ARNm sont essentiels à l'homéostasie cellulaire. Chez les eucaryotes, la voie Nonsense-mediated mRNA Decay (NMD) est un processus de surveillance essentiel qui détecte et élimine les ARNm aberrants contenant des codons de terminaison prématurés (PTC). Au cœur de cette voie se trouve l'hélicase à ARN Upf1, qui comprend un domaine hélicase flanqué d'un domaine régulateur riche en cystéine et en histidine (CH) et d'un domaine C-terminal désordonné. Upf1 orchestre la reconnaissance et la dégradation des transcrits contenant un PTC. Cependant, le mécanisme précis par lequel Upf1 déclenche la dégradation des ARNm reste mal compris. Des études biochimiques récentes ont mis en évidence deux complexes mutuellement exclusifs chez Saccharomyces cerevisiae : le complexe Upf1-détecteur composé d'Upf2 et Upf3 et le complexe Upf1-decapping, comprenant Dcp1, Dcp2, Nmd4 et Ebs1. Ce projet de thèse visait à reconstituer le complexe Upf1-decapping in vitro en utilisant des protéines recombinantes afin de caractériser le rôle d'Upf1 et de ses partenaires dans les étapes tardives du NMD. Nous avons notamment identifié une interaction directe entre Dcp2 et le domaine CH d'Upf1, ainsi que des interactions directes entre différents domaines d'Upf1, Nmd4 et Ebs1, formant un complexe stable Upf1-decapping. Comme chez l'homme, Upf2 interagit avec le domaine CH. De manière intéressante, cette interaction est incompatible avec la liaison de Dcp2, ce qui explique l'incompatibilité entre les complexes Upf1-détecteur et Upf1-decapping. En outre, nos recherches ont révélé que la liaison de d'Upf1 à l'ARN est presque abolie en présence d'Upf2, mais qu'elle est considérablement renforcée au sein du complexe Upf1-decapping. Ainsi, Upf1 faciliterait l'association des protéines de decapping sur ses ARNm cibles, assurant ainsi une dégradation ciblée. En outre, nous avons exploré les mécanismes d'autorégulation et les propriétés fonctionnelles des hélicases Upf1 de la levure et de l'homme sur l'ARN et l'ADN. Des analyses comparatives à l'échelle de la molécule unique à l'aide de pinces magnétiques ont révélé des comportements autorégulateurs opposés de l'activité de l'hélicase Upf1 sur l'ARN malgré leur grande conservation. Dans l'ensemble, cette étude fournit des informations précieuses sur l'importance fonctionnelle de l'hélicase Upf1 pour recruter la machinerie de dégradation des ARNm défectueux, contribuant ainsi à notre compréhension du métabolisme de l'ARN.

Mots clés: Surveillance de l'ARN, voie NMD, hélicase Upf1, decapping, biochimie, molécule unique