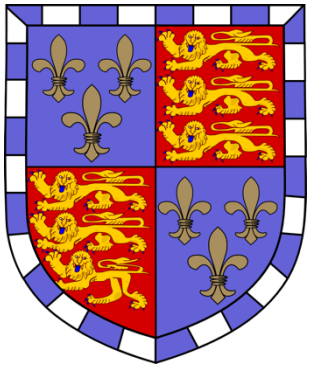
Structural and Biochemical Investigation of the Pan2-Pan3 Deadenylase Complex

Tsz Long Terence Tang



Christ's College

July 2020

This dissertation is submitted for the degree of Doctor of Philosophy

MRC Laboratory of Molecular Biology

UNIVERSITY OF CAMBRIDGE

To my parents

for all of their sacrifices

The surface is fine and powdery

I can pick it up loosely

with my toe

It does adhere in fine layers

like powdered charcoal

to the sole and sides of my boots

I only go in a small fraction of an inch

maybe an eighth

of an inch

But I can see the footprints

of my boots

and the treads

in the fine, sandy particles

Preface

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text. Work carried out as a result of collaborations is specifically declared before each Results Chapter.

This thesis is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text. The majority of work presented in Chapter 2 has been published:

1. **Tang, T.T.L.**, Stowell, J.A.W., Hill, C.H., and Passmore, L.A. (2019). The intrinsic structure of poly(A) RNA determines the specificity of Pan2 and Caf1 deadenylases. Nat Struct Mol Biol *26*, 433-442.
2. **Tang, T.T.L.**, and Passmore, L.A. (2019). Recognition of Poly(A) RNA through Its Intrinsic Helical Structure. Cold Spring Harb Symp Quant Biol *84*, 21-30.

This thesis does not exceed the prescribed word limit for the Degree Committee for Biology.

Tsz Long Terence Tang

Submitted for examination: July 2020

Summary

Structural and Biochemical Investigation of the Pan2-Pan3 Deadenylase Complex

This thesis is submitted for the degree of Doctor of Philosophy by Tsz Long Terence Tang.

Almost all mature eukaryotic mRNAs contain a 3' polyadenosine (poly(A)) tail, which promotes nuclear export of mRNAs, protects transcripts from exonucleolytic decay, and increases translation efficiency. As the poly(A) tail is involved in many steps of gene expression, its length is highly regulated. Poly(A) tail shortening, known as deadenylation, represses gene expression by initiating mRNA decay and inhibiting translation. Deadenylation is required for mRNA homeostasis; moreover, specific transcripts can be targeted for deadenylation, enabling gene expression to respond to external stimuli. In eukaryotes, cytoplasmic deadenylation is mainly carried out by two conserved multiprotein complexes: Pan2-Pan3 and Ccr4-Not.

This thesis focuses on the *in vitro* characterisation of Pan2-Pan3. While Pan2-Pan3 and Ccr4-Not show poly(A) preference, it was unknown how the conserved exonucleases Pan2 and Caf1 recognise poly(A). In Chapter 2, structural and biochemical studies reveal that these DEDD-family deadenylases recognise the intrinsic stacked, helical structure of single-stranded poly(A) RNA. Recognition of this conformation in different biological contexts, enabled by the unique physicochemical properties of adenine, suggests a possible reason for poly(A) conservation. Chapter 3 examines the *in vitro* activity of Pan2-Pan3 and how it is regulated, particularly by the cytoplasmic poly(A)-binding protein Pab1. These studies elucidate domain requirements and species-specific features of Pab1-dependent regulation of Pan2-Pan3. My experiments further examine how Pan2-Pan3 may be recruited by sequence-specific RNA-binding proteins. Together, these results provide insights into Pan2-Pan3 deadenylation and how this dynamic process could be regulated and thereby influence post-transcriptional gene expression.

Acknowledgements

Above all, I would like to thank my doctoral supervisor Lori Passmore for providing the opportunity to work on this project. I applied for a PhD at the LMB under some uncertainty; it was through Lori's ideas and enthusiasm that the science and project shone through. My sincerest appreciation for providing the freedom and support to allow me to develop as an independent scientist and explore interesting experimental avenues.

Huge thanks to my second and university supervisors, Manu Hegde and Phil Zegerman, for their invaluable scientific guidance, particularly during the early days of the PhD. Thanks to Chris Russo and members of his lab for providing feedback during our joint group meetings.

I would like to extend my gratitude to all members of the Passmore Lab – those who came before and set the groundwork, those who helped laying bricks along the way, and those who will continue building upon my experimental foundations. My deepest gratitude to James Stowell and Michael Webster, who inducted me into the field of deadenylation. I will remember the deadenylation assays, protein purifications, and the whirr of the Typhoon scanner with great fondness. Thanks to other members who worked on deadenylation, for providing helpful advice with experiments, the paper, and my thesis along the way: Anne Kelley, Max Seidel, Eva Absmeier, and Cemre Manav. Thank you to all members of the lab, past and present: Ananth Kumar, Vytautė Boreikaitė, Eleanor Sheekey, Manuel Carminati, Juan Rodríguez-Molina, Conny Yu, Tamara Sijacki, Shabih Shakeel, Pablo Alcón, Ana Casañal, and Eeson Rajendra. (Thank you to random.org for randomising the above list.) More specifically, I would like to thank Jana Wolf for establishing a detailed and helpful basis upon which my work is built. Thank you to Chris Hill for teaching me much of what I know about crystallography. Finally, I would like to thank Alžběta Roeselová for her huge contributions to the MEX3 portion of the project and for being a breeze to supervise.

Thank you to everyone in the LMB who has made my work possible: Stephen McLaughlin and Chris Johnson for their knowledge and support with a variety of biophysical techniques; Minmin Yu and Fabrice Gorrec for crystallisation screening and data collection; Jianguo Shi for his help with the insect cell facility; Christos Savva, Giuseppe Cannone, and the EM facility for their tutelage and troubleshooting; Jake Grimmett, Toby Darling, and the scientific computing facility. In particular, I would like to acknowledge Vish Chandrasekaran, Danguolė Ciziene, Chris Oubridge, and Kiyoshi Nagai for helpful scientific discussions. Thank you to Diamond Light Source and its staff for precious beam time. Further thanks to the Structural Studies divisional office, the postgraduate office, the stores team, the media kitchen, and glass-washing facility for their support. Thank you to the Carter, Russo, and Nguyen labs, as well as my PhD cohort, for much-needed social respite inside and outside the LMB.

My research was made possible, and my day-to-day life made more comfortable, by funding from the Herchel Smith Fund and their PhD studentships. They have also provided numerous opportunities for networking with like-minded students and postdoctoral fellows. Thank goodness for oral contraceptives. Christ's College and the EMBO YIP programme have provided valuable communities for scientific and non-scientific exchange. Thank you to researchers I have encountered along the way for their questions and advice.

Thank you to my ever-supportive friends and family, who patiently yet ceaselessly ask me what I am working on and what I plan to do with my life. Particular gratitude for the company of the London GSIS crew and Somerville gang. Thanks to my brother Tim and my sister Tiff for being a family away from home. I am deeply sorry for not being more present for HK friends and particularly for my parents. Mum and Dad, none of this would have been possible without you. I hope I have made you proud.

Finally, thank you to Catherine. I am *always* listening to you.

Table of Contents

1. Introduction 1

1.1 The Poly(A) Tail 2

1.2 mRNA Decay and Deadenylation 5

1.2.1 Deadenylation 7
1.2.2 Pan2-Pan3 9
1.2.3 Ccr4-Not 15
1.2.4 PARN 19
1.2.5 Minor Deadenylases 21

1.3 Properties of the Poly(A) Tail 23

1.3.1 Poly(A) Tail Length 23
1.3.2 Modifications of the Poly(A) Tail 25
1.3.3 The Structure of Poly(A) 26

1.4 RNA-binding Proteins 29

1.4.1 Poly(A)-binding Protein 30
1.4.2 RNA-binding Proteins and Deadenylation 37
1.4.3 Identification of Putative Binding Partners of Pan2-Pan3 39

1.5 Current Model of Pan2-Pan3 Function 41

1.5.1 Overall Structural Model of Pan2-Pan3 41
1.5.2 Interplay Between Pan2-Pan3 and Ccr4-Not 44

1.6 Aims 47

2. Poly(A) Specificity of DEDD Deadenylases 49

2.1 Pan2 Has Intrinsic Nucleotide Specificity 51

2.1.1 Expression and Purification of Pan2-Pan3 Constructs 51
2.1.2 Nucleotide Specificity of Pan2-Pan3 53

2.2 Intrinsic Nucleotide Specificity of Ccr4 and Caf1 56

2.3 Guanosine Prevalence in the 3' UTR 58

2.4 Structure of the Pan2-Poly(A) Complex 59

2.4.1 Overall Structure of Pan2UCH-Exo-A_7 60
2.4.2 Pan2 Contacts the Ribophosphate Backbone of Oligo(A) 62

2.5 Guanosines Disrupt Poly(A) Structure 63

2.5.1 Structures of Pan2UCH-Exo-AAGGA and -AAGGAA 63
2.5.2 Y975 is Important for Pan2 Activity and Specificity 65
2.5.3 Pan2 Binds Similarly to Oligo(A) and G-containing RNA 66

2.6 Guanosines Disrupt the Conformation of Oligo(A) in Solution 67

2.7 Cytosines and Uracils Can Stack in the Pan2 Active Site 68

2.8 Base Stacking is Required for Pan2 Activity 70

2.9 Caf1 May Recognise the Poly(A) Structure 72

2.9.1 Structural Superposition of Pan2UCH-Exo-A_7 and Caf1 72
2.9.2 Caf1 Likely Requires Base Stacking for Activity 74

2.10 Effect of Adenine Modifications 75
2.10.1 Effect of Purine Substituents on Poly(A) Structure 75
2.10.2 Effect of Adenine Modifications on Poly(A) Structure 76
2.11 Discussion 78
2.11.1 Investigation of Nucleotide Specificity by *In Vitro* Deadenylation Assays 78
2.11.2 Deadenylase Specificity Prevents 3' UTR Degradation 79
2.11.3 Guanosines Inhibit Deadenylation 81
2.11.4 Mechanism of Pan2 Recognition of Poly(A) 82
2.11.5 Poly(A) Structure in Other Biological Processes 86

3. Regulation of Pan2-Pan3 Deadenylation 91

3.1 N-terminal Truncations of Pan3 Do Not Greatly Affect Deadenylation 93
3.2 Upstream Sequence Specificity of Pan2-Pan3 96
3.3 Pab1 Binds Poly(A) with High Affinity 99
3.3.1 Purification of Poly(A)-Binding Protein 99
3.3.2 Pab1 Binds Oligo(A) with Nanomolar Affinity 100
3.4 PABPC1 Binds Poly(A) Cooperatively 102
3.5 Pab1 Stimulation of Pan2-Pan3 Depends on Poly(A) Tail Length 105
3.6 Stimulation of PAN2-PAN3 Depends on PABPC1 Stoichiometry 110
3.7 Domain Requirements of Pab1 Regulation of Pan2-Pan3 114
3.7.1 The Pan2 Exonuclease is Inhibited by Pab1 114
3.7.2 The C-terminus of Pab1 Stimulates Yeast Pan2-Pan3 115
3.7.3 PABPC1 RRM Domains Can Stimulate PAN2-PAN3 117
3.8 The PAN2-PAN3-PABPC1-poly(A) Complex 119
3.8.1 Pab1 C-terminus is Required to Interact with Pan2-Pan3 119
3.8.2 PAN2-PAN3 Interacts with PABPC1-Poly(A) Independent of Poly(A) Tail Length and Stoichiometry 121
3.8.3 The Minimal PAN2-PAN3-PABPC1-Poly(A) Complex 122
3.9 Crosslinking of Pan2-Pan3 127
3.10 Negative Stain EM of Apo Pan2-Pan3 129
3.11 Cryo-EM of Apo PAN2-PAN3 131
3.12 MEX3 is a Sequence-specific RBP 135
3.13 MEX3 Stimulates PAN2-PAN3 Activity 138
3.14 MEX3 Interacts with PAN2-PAN3 141
3.15 Discussion 143
3.15.1 Studying Pan2-Pan3 Deadenylation by *In Vitro* Deadenylation Assays 143
3.15.2 RNA-binding Characteristics of PABPC1 145
3.15.3 Regulation of Pan2-Pan3 by Pab1 147
3.15.4 Mechanism of MEX3 Stimulation of PAN2-PAN3 155

4. Conclusions and Future Perspectives 157

4.1 Recognition of the Intrinsic Structure of Poly(A) 159
4.2 Updated Model of Pan2-Pan3 Function 162
4.3 A Global View of Deadenylation 165

4.4 Conclusion 168

5. Materials and Methods 169

5.1 Common Reagents 170
5.1.1 Buffers 170
5.1.2 Liquid Media 170
5.1.3 Plate Media 170
5.1.4 Antibiotics and Additives 171

5.2 *E. coli* 171
5.2.1 Strains 171
5.2.2 Competent *E. coli* 172
5.2.3 Heat-shock Transformation 172
5.2.4 Growth for Plasmid/Bacmid Preparation 172
5.2.5 Growth for Protein Overexpression 173

5.3 *S. cerevisiae* 173
5.3.1 Strains 173
5.3.2 Competent *S. cerevisiae* 174
5.3.3 Transformation 174
5.3.4 Growth for Protein Overexpression 174

5.4 Sf9 cells 175
5.4.1 Strain 175
5.4.2 Maintenance of Sf9 Cells in Suspension 175
5.4.3 Bacmid Transfection of Adherent Sf9 Cells 175
5.4.4 Virus Amplification 176
5.4.5 Protein Overexpression 176

5.5 Molecular Biology Techniques and Cloning 177
5.5.1 Gene Synthesis 177
5.5.2 Nucleic Acid Quantification 178
5.5.3 Polymerase Chain Reaction (PCR) 178
5.5.4 Restriction Enzyme Cloning 179
5.5.5 Gibson Assembly 180
5.5.6 Multigene Expression Plasmids (pBIG1) 180
5.5.7 QuikChange Site-directed Mutagenesis 183
5.5.8 DNA Analysis by Horizontal Agarose Gel Electrophoresis 183
5.5.9 Gel Extraction of DNA 183
5.5.10 Plasmid DNA Extraction from *E. coli* 184
5.5.11 Bacmid DNA Extraction from *E. coli* 184
5.5.12 Sanger Sequencing 184

5.6 Protein Purification 185
5.6.1 Analysis of Protein Purification 187
5.6.2 3C Protease 188
5.6.3 TEV Protease 188
5.6.4 *S. cerevisiae* Pan2-Pan3 and Variants 189
5.6.5 *S. cerevisiae* and *C. thermophilum* Pan2UCH-Exo and Variants 189
5.6.6 *H. sapiens* PAN2(E980A)-PAN3-SII 190
5.6.7 *H. sapiens* PAN2(E980A)-PAN3ΔN278-SII, PAN2(E980A)-PAN3PKC-SII 191
5.6.8 *S. cerevisiae* Pab1 (RRM1-4) 192
5.6.9 *H. sapiens* PABPC1 (RRMH, RRM1-4) 192

5.6.10 *H. sapiens* MEX3A-SII, MEX3B-SII, MEX3C-SII 193
5.7 RNA Preparation 194
5.7.1 RNA Synthesis 194
5.7.2 *In vitro* Transcription 195
5.8 Biophysical Experiments 196
5.8.1 Fluorescence Polarisation Anisotropy (FPA) 196
5.8.2 Microscale Thermophoresis (MST) 196
5.8.3 Differential Scanning Fluorimetry (DSF) 197
5.8.4 Circular Dichroism (CD) Spectroscopy 197
5.9 *In vitro* Biochemical Experiments 198
5.9.1 Deadenylation Assays 198
5.9.2 Electrophoretic Mobility Shift Assay (EMSA) 200
5.9.3 Radioactive EMSA 200
5.10 *In vitro* Biochemical Experiments (Protein) 201
5.10.1 Assessment of *E. coli* Overexpression 201
5.10.2 Assessment of Sf9 Overexpression 201
5.10.3 Pull-down Assays 202
5.10.4 Analytical Size Exclusion Chromatography (SEC) 203
5.10.5 BS3 Crosslinking 203
5.11 X-ray Crystallography 204
5.11.1 Initial Crystallisation Screens 204
5.11.2 Optimisation Screens 204
5.11.3 RNA Soaking 205
5.11.4 Data Collection 206
5.11.5 Structure Determination and Refinement 207
5.11.6 Structural Representation 207
5.12 Negative Stain Electron Microscopy (EM) 208
5.12.1 Sample Preparation 208
5.12.2 Data Collection 208
5.12.3 Data Processing and Analysis 209
5.13 Cryo-electron Microscopy (Cryo-EM) 210
5.13.1 Sample Preparation 210
5.13.2 Data Collection 211
5.13.3 Data Processing and Analysis 211
5.14 Bioinformatics 212
5.14.1 Sequence Alignment 212
5.14.2 Poly(A) Site Sequence Analysis 212
5.14.3 Secondary Structure Prediction 212
5.14.4 RNA Structure Prediction 213
5.14.5 Prediction of Protein Properties 213
5.15 Others 213
5.15.1 Schematic Diagrams 213
5.15.2 Gel Representation 213
6. References 215
7. Supplementary Figures 243

List of Figures

Figure 1.1 Biological roles of the poly(A) tail. 4
Figure 1.2 Eukaryotic mRNA decay. 6
Figure 1.3 Domain schematics for *H. sapiens* PAN2 (blue) and PAN3 (orange). 10
Figure 1.4 Structural model for Pan2-Pan3. 14
Figure 1.5 Structure of *H. sapiens* CNOT6L (purple) bound to single-stranded oligo(A) DNA (green sticks) (PDB: 3NGO). 17
Figure 1.6 Structure of *S. pombe* Caf1 (PDB: 3G0Z). 18
Figure 1.7 Structure of *H. sapiens* PARN (yellow) bound to single-stranded oligo(A) RNA (green sticks) (PDB: 2A1R). 20
Figure 1.8 Structural conservation of deadenylases and related proteins. 22
Figure 1.9 Molecular basis of poly(A)-binding protein (Pab1) function. 34
Figure 1.10 (A) Putative protein binding partners of PAN2-PAN3. 39
Figure 1.11 Structural model of Pan2-Pan3 in complex with Pab1-poly(A). 43
Figure 2.1 Purification of *S. cerevisiae* Pan2-Pan3. 52
Figure 2.2 Pan2 exhibits position-dependent 3' nucleotide specificity. 54
Figure 2.3 Single guanosines inhibit Pan2 deadenylation. 55
Figure 2.4 Ccr4-Not deadenylases Ccr4 and Caf1 display nucleotide specificity. 57
Figure 2.5 Crystallisation of *S. cerevisiae* Pan2UCH-Exo E912A (blue). 59
Figure 2.6 Overall structure of Pan2UCH-Exo-A_7. 61
Figure 2.7 Interactions between Pan2UCH-Exo (blue) and oligo(A) RNA (green). 62
Figure 2.8 Structures of Pan2UCH-Exo bound to G-containing RNAs. 63
Figure 2.9 Guanosines disrupt base stacking. 64
Figure 2.10 Y975 is important for Pan2 activity and specificity. 65
Figure 2.11 Pan2 binds oligo(A) and G-containing RNAs with similar affinity. 66
Figure 2.12 Circular dichroism of oligo(A) RNA. 67
Figure 2.13 Poly(A) forms a unique intrinsic structure by circular dichroism. 68
Figure 2.14 Structures of Pan2UCH-Exo-AACCAA and -AAUUAA RNAs. 69
Figure 2.15 Base stacking is required for Pan2 activity. 71
Figure 2.16 Superposition of Pan2UCH-Exo (blue)–A_7 (green) and *S. pombe* Caf1 (pink). 73
Figure 2.17 Base stacking is required for Caf1 activity. 74
Figure 2.18 Electronic distribution of functional groups is important for poly(A) structure. 76
Figure 2.19 Effect of adenine modifications on oligo(A) structure. 77
Figure 2.20 Stringency of poly(A) specificity correlates with the hypothesised *in vivo* roles of Pan2-Pan3 and Ccr4-Not according to the biphasic model of deadenylation. 80
Figure 2.21 Pan2 recognises the intrinsic helical structure of poly(A) RNA. 85
Figure 2.22 Poly(A) structure in other biological processes. 88
Figure 3.1 N-terminal Pan3 truncations do not greatly affect Pan2-Pan3 deadenylation. 94

Figure 3.2 N-terminal PAN3 truncations do not greatly affect PAN2-PAN3 deadenylation.95
Figure 3.3 *S. cerevisiae* Pan2-Pan3 displays preference for UTR sequences. 97
Figure 3.4 *H. sapiens* PAN2-PAN3 has similar sequence preference to yeast Pan2-Pan3. 98
Figure 3.5 Purification of recombinantly expressed Pab1. 100
Figure 3.6 Pab1 binds poly(A) with nanomolar affinity. 101
Figure 3.7 The C-terminal half of PABPC1 aids cooperative RNA binding. 104
Figure 3.8 Pab1 stimulation of Pan2-Pan3 depends on poly(A) tail length. 107
Figure 3.9 PABPC1 stimulation of PAN2-PAN3 depends on poly(A) tail length. 108
Figure 3.10 Dependence of Pab1 stoichiometry on Pan2-Pan3 stimulation. 111
Figure 3.11 Effect of PABPC1 stoichiometry on deadenylation by 100 nM *H. sapiens* PAN2-PAN3ΔN278. 113
Figure 3.12 The Pan2 exonuclease alone is inhibited by Pab1. 114
Figure 3.13 The Pab1 C-terminus is required to stimulate Pan2-Pan3. 116
Figure 3.14 PABPC1 RRM1-4 can partially stimulate PAN2-PAN3. 118
Figure 3.15 The Pab1 C-terminus is required for a stable interaction with Pan2-Pan3. 120
Figure 3.16 PAN2-PAN3 interacts with PABPC1 in a poly(A) tail length- and stoichiometry-independent manner. 122
Figure 3.17 Defining the minimal complex of PAN2-PAN3-PABPC1-poly(A). 125
Figure 3.18 RRM4 binds the 3' UTR during PAN2-PAN3 deadenylation. 126
Figure 3.19 Crosslinking trials of *S. cerevisiae* and *H. sapiens* Pan2-Pan3. 128
Figure 3.20 Negative stain EM of *S. cerevisiae* and *H. sapiens* Pan2-Pan3. 130
Figure 3.21 Uncrosslinked PAN2-PAN3 aggregates during cryo-EM grid preparation. 132
Figure 3.22 Cryo-EM of PAN2-PAN3. 134
Figure 3.23 Purification of *H. sapiens* MEX3 proteins. 135
Figure 3.24 Recombinant MEX3 is a *bona fide* RNA-binding protein. 137
Figure 3.25 MEX3 stimulates full-length PAN2-PAN3. 138
Figure 3.26 MEX3 stimulates N-terminal truncations of PAN2-PAN3. 139
Figure 3.27 MEX3 does not stimulate the PAN2 catalytic unit. 140
Figure 3.29 Pull-down assays show that *H. sapiens* MEX3 directly interacts with *H. sapiens* PAN2-PAN3. 142
Figure 3.30 Model of inhibition of Pan2-Pan3 by Pab1. 148
Figure 3.31 Model for interactions between PAN2 (blue cartoon)-PAN3 (orange cartoon) and PABPC1 (grey cartoon). 151
Figure 3.32 Pan2-Pan3 and Pab1 recapitulate steady-state poly(A) tails. 153
Figure 4.1 Updated model of Pan2-Pan3 function. 164
Figure 5.1 Schematic of biGBac protocol. 182

List of Tables

Table 2.1 X-ray crystallography data collection and refinement statistics. 89
Table 5.1 List of antibiotics and additives used in liquid and plate media 171
Table 5.2 List of *E. coli* strains used in this dissertation. 171
Table 5.3 List of yeast strains used in this dissertation. 173
Table 5.4 Description of insect cell strain used in this dissertation. 175
Table 5.5 Gene synthesised constructs used in this dissertation. 177
Table 5.6 Phusion PCR Reaction Setup 178
Table 5.7 Phusion PCR Thermocycling Parameters 179
Table 5.8 Primers used for Gibson Assembly into pBIG1. 181
Table 5.9 List of purified proteins and protein complexes. 185
Table 5.10 List of synthesised RNA sequences used in this dissertation. 194
Table 5.11 List of *in vitro* transcribed RNAs used in this dissertation. 195
Table 5.12 List of crystals used in final data collection and RNA soak conditions. 206

List of Abbreviations

2×TY	2× tryptone yeast
2AP	2-aminopurine
Å	Ångstrom
A_λ	Absorbance at wavelength λ
Amp	Ampicillin
APS	Ammonium persulphate
Ar	Argon
ARE	AU-rich element
ATP	Adenosine triphosphate
β-ME	2-Mercaptoethanol
Bis-Tris	2,2-Bis(hydroxymethyl)-2,2',2''-nitrilotriethanol
bp	Base pair
BS3	Bissulfosuccinimidyl suberate
BSA	Bovine serum albumin
$CaCl_2$	Calcium chloride
CCD	Charge-coupled device
CD	Circular dichroism
CHAPSO	3-[(3-cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate
CHES	*N*-Cyclohexyl-2-aminoethanesulfonic acid
CLIP	Crosslinking immunoprecipitation
CPF	Cleavage and polyadenylation factor
CPSF	Cleavage and polyadenylation specificity factor
Cryo-EM	Cryo-electron microscopy
CTD	C-terminal domain
CTF	Contrast transfer function
CV	Column volume(s)
deg	Degrees (unit of ellipticity)
DEPC	Diethylpyrocarbonate
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleoside triphosphate
ds	Double-stranded
DSB	D-desthiobiotin
DSF	Differential scanning fluorimetry
DTT	Dithiothreitol
e^-	Electron
ε_λ	Extinction coefficient at wavelength λ
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
EMSA	Electrophoretic mobility shift assay
Exo	Exonuclease
FAM	Carboxyfluorescein
FBS	Fetal bovine serum
FEM	Feature-enhanced map
FPA	Fluorescence polarisation assay
g	gram

g	Standard acceleration due to gravity
g/L	grams dm^{-3}
GB	Gibson buffer
Gen	Gentamycin
GOI	Gene of interest
GST	Glutathione S-transferase
h	Hour(s)
H. sapiens	*Homo sapiens*
HCl	Hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	Histidine
Hz	Hertz
I	Inosine
IAB	Isothermal assembly buffer
IPTG	Isopropyl β-D-1-thiogalactopyranoside
K	Kelvin
Kan	Kanamycin
kb	Kilobases
KCl	Potassium chloride
K_d	Dissociation constant
kDa	Kilodalton
KH_2PO_4	Potassium dihydrogenphosphate
KOAc	Potassium acetate
L	Litre; dm^3
LB	Miller lysogeny broth
LDS	Lithium dodecyl sulphate
LED	Light-emitting diode
m	Metre
M	mol dm^{-3}
M^{2+}	Divalent metal cation
m^6A	*N*6-methyladenosine
m^7G	7-methylguanosine
MBP	Maltose binding protein
MES	2-(*N*-morpholino)ethanesulfonic acid
$Mg(OAc)_2$	Magnesium acetate
$MgCl_2$	Magnesium chloride
min	Minute(s)
miRNA	MicroRNA
$MnCl_2$	Manganese (II) chloride
MOPS	3-(*N*-morpholino)propanesulfonic acid
MQ H_2O	Milli-Q water
MRE	MEX3 response element
mRNA	Messenger RNA
MST	Microscale thermophoresis
N_2	Nitrogen
Na_2HPO_4	Sodium hydrogen phosphate
NaCl	Sodium chloride
NAD	Nicotinamide adenine dinucleotide
NaOAc	Sodium acetate
NaOH	Sodium hydroxide
$NH_4H_2PO_4$	Ammonium dihydrogen phosphate

NH_4HCO_3	Ammonium hydrogen carbonate
Ni^{2+}-NTA	Nickel^{2+}-Nitrilotriacetic acid
NOβG	N-octyl- β-D-glucoside
nt	Nucleotide
NTP	Nucleoside triphosphate
O_2	Oxygen
OD_λ	Optical density at wavelength λ
P	Fluorescence polarisation units
P	Purine
PAGE	Polyacrylamide gel electrophoresis
PAM2	PABP-interacting motif 2
PAR-CLIP	Photoactivatable ribonucleoside crosslinking and immunoprecipitation
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
pI	Isoelectric point
PIPES	Piperazine-*N*,*N*'-bis(2-ethanesulfonic acid
piRNA	Piwi RNA
PK	Pseudokinase domain
PKC	Pseudokinase-C-terminal domain
PMSF	Phenylmethylsulfonyl fluoride
PMT	Photomultiplier tube
PNK	Polynucleotide 5'-hydroxyl-kinase
POI	Protein of interest
Polh	Polyhedrin
Pol II	RNA Polymerase II
PRE	Pumilio-response element
psi	Pounds per square inch
RB	Running buffer
$RbCl_2$	Rubidium chloride
RBD	RNA-binding domain
RBP	RNA-binding protein
RMSD	Root mean squared deviation
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
rpm	Revolutions per minute
RRM	RNA recognition motif domain(s)
RT	Room temperature
s	Second(s)
S. cerevisiae	*Saccharomyces cerevisiae*
S. pombe	*Schizosaccharomyces pombe*
SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
*Sf*9	*Spodoptera frugiperda Sf*21 ovarian cells
SII	StrepTag II
snoRNA	Small nucleolar RNA
snRNA	Small nuclear RNA
ss	Single-stranded
SV40	Simian virus 40
$t_{1/2}$	Half-life
TAE	Tris acetate EDTA

TB	Terrific broth
TBE	Tris borate EDTA
TCEP	Tris(2-carboxyethyl)phosphine
TE	Tris-EDTA
TEM	Transmission electron microscope
TEMED	Tetramethylethylenediamine
TEV	Tobacco etch virus protease or TEV protease cleavage site (-ENLYFQG-)
T_m	Melting temperature
Tris	Tris(hydroxymethyl)aminomethane
Trp	Tryptophan
TYE	Tryptone yeast extract
UCH	Ubiquitin C-terminal hydrolase
Ura	Uracil
UTR	Untranslated region
UV	Ultraviolet
V	Volt(s)
v/v	Volume/volumen percentage
W	Watt(s)
w/v	Mass/volume percentage
WD40	Trp-Asp 40 repeat
wt	wild-type
YEPD	Yeast extract peptone D-glucose
YFP	Yellow fluorescent protein
YM4	Yeast minimal media 4
YPD	Yeast peptone D-glucose
ZnF	Zinc finger domain

Standard 3-letter and 1-letter abbreviations for amino acids are used.

Standard 1-letter abbreviations for nucleotides are used.

1. Introduction

1.1 The Poly(A) Tail

The central dogma of molecular biology postulates that the flow of genetic information occurs by transcription of DNA to form messenger RNA (mRNA), which is translated into functional protein. Biological systems stringently regulate these fundamental processes to control the timing, location, and amount of gene product. This enables cells and organisms to respond to stimuli, maintain homeostasis, develop and differentiate, and be capable of evolution. As intermediates between DNA and protein, mRNAs must be highly regulated. Eukaryotic mRNA regulation is multifaceted and can occur by its synthesis, degradation, translation by ribosomes, chemical modifications, higher order structure formation, complex formation with RNA-binding proteins or other RNAs, and subcellular localisation. Correctly produced and processed mRNAs possess unique marks to enable their proper function; these also allow *bona fide* mRNAs to be distinguished from erroneous transcripts and enable quality control.

One such mark is the poly(A) tail (Sachs, 1990; Zhao et al., 1999). Almost all mature eukaryotic mRNAs, except those encoding metazoan replication-dependent histones, possess a poly(A) tail (Adesnik et al., 1972). By nuclease resistance, filter retention, and cosedimentation studies, poly(A) tails were discovered 50 years ago in rabbit reticulocytes (Lim and Canellakis, 1970), vaccinia viral mRNA (Kates and Beeson, 1970), and murine S180 cells (Lee et al., 1971). The poly(A) tail was soon found to be associated with the 3' end (Mendecki et al., 1972; Molloy et al., 1972) and conserved in eukaryotes from yeast (McLaughlin et al., 1973) to humans (Darnell et al., 1971). Radioactive labelling and transcription inhibition suggested that poly(A) was added after transcription (Darnell et al., 1971; Mendecki et al., 1972). Decades of studies have revealed that poly(A) addition (polyadenylation) is carried out by a multiprotein complex (yeast: CPF, metazoan: CPSF) containing the canonical poly(A) polymerase Pap1, following endonucleolytic cleavage of pre-mRNA by CPF/CPSF (Kumar et al., 2019; Proudfoot, 2011).

Alternative cleavage leads to isoforms with varied 3' untranslated regions (3' UTRs), increasing the repertoire of mature transcripts derivable from a pre-mRNA (Tian and Manley, 2017).

The biological role of the 3' poly(A) tail has been extensively investigated (Figure 1.1). Firstly, the poly(A) tail is required for efficient nuclear export of mRNAs. The adenosine analogue cordycepin, which terminates poly(A) elongation, inhibited mRNA export to the cytoplasm (Penman et al., 1970). Further studies showed that the poly(A) tail is a determinant of nuclear export of mature mRNAs (Fuke and Ohno, 2008) by binding to the nuclear poly(A) binding protein Nab2, recruiting mRNA to nuclear export factors (Soucek et al., 2012). Additionally, the poly(A) tail, originally found to be polysome-associated (Baglioni et al., 1972; Darnell et al., 1971), stimulates translation efficiency. In rabbit reticulocyte extract (Munroe and Jacobson, 1990), yeast extract (Preiss et al., 1998), *Xenopus* oocytes (Huez et al., 1974; Nudel et al., 1976), and CHO cells (Gallie, 1991), a minimum poly(A) tail length was required for efficient translation, though it was not absolutely required (Proweller and Butler, 1994; Searfoss and Wickner, 2000). Finally, the poly(A) tail enhances transcript stability (Marbaix et al., 1975; Wilson et al., 1978) by protecting the mRNA body from exonucleolytic decay. Interestingly, although no poly(A) tail was initially found in prokaryotic mRNAs (Perry et al., 1972), oligo(A) tailing was later identified to stimulate degradation (Rorbach et al., 2014), raising the question of the role of poly(A) tails in mitochondria and chloroplasts.

Poly(A) tails regulate biological function through proteins which bind and act on poly(A). In the cytoplasm, poly(A) is bound by the conserved cytoplasmic poly(A) binding protein Pab1, which is thought to confer increased stability and translation efficiency. Regulation of gene expression can also be effectuated through poly(A) degradation by conserved cytoplasmic deadenylase enzymes and enzymatic complexes. Their activities are further influenced by RNA-binding proteins *in cis*. Thus, the poly(A) tail is intimately involved in many stages of the mRNA life cycle and regulates gene expression in eukaryotes.

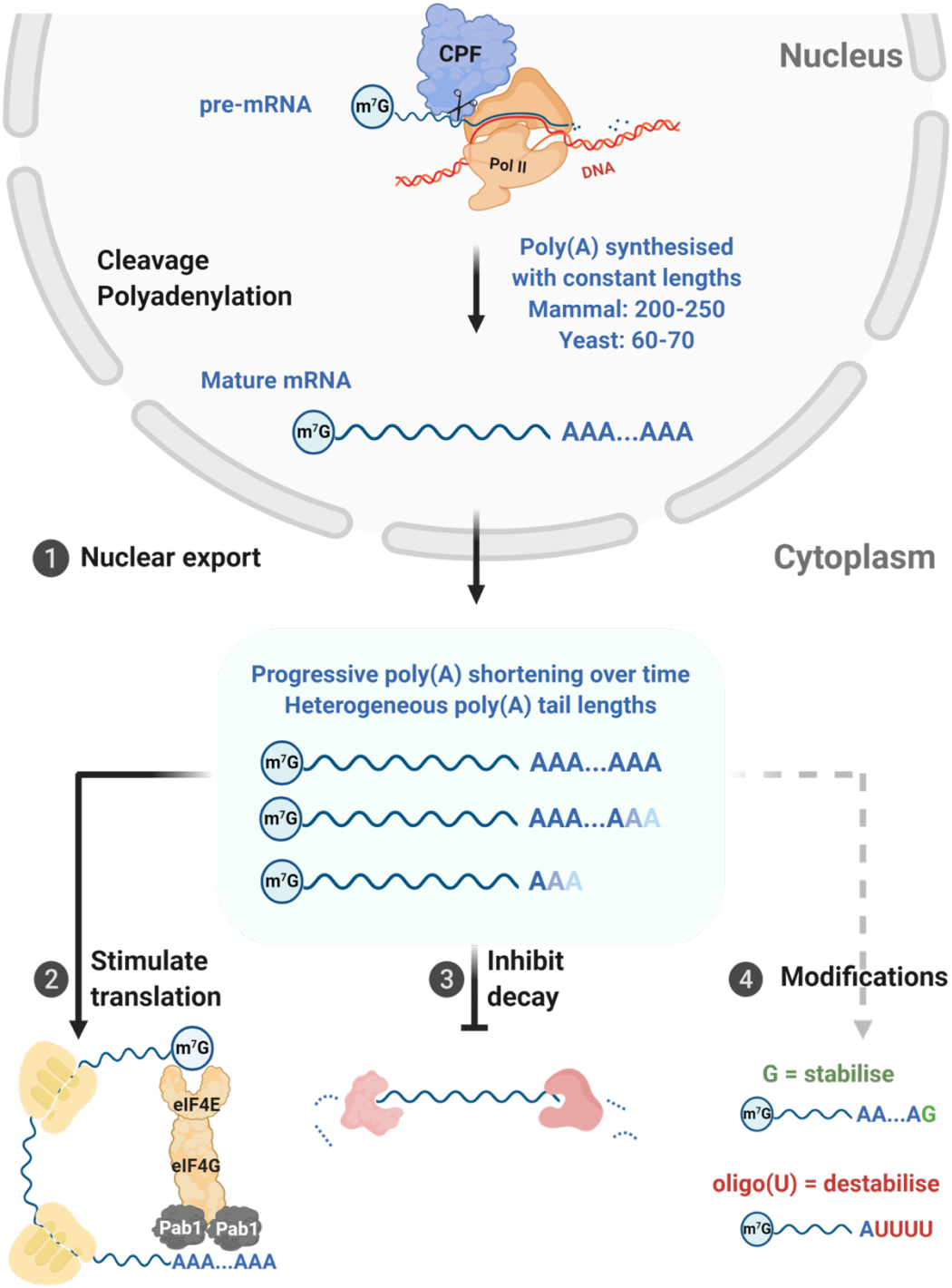


Figure 1.1 Biological roles of the poly(A) tail. Poly(A) tails are synthesised with a relatively narrow range of lengths, and are shortened in the cytoplasm to generate a transcript pool with heterogeneous poly(A) tail lengths. The poly(A) tail: (**1**) is required for nuclear export; (**2**) stimulates translation efficiency; (**3**) stabilises the transcript; (**4**) can be modified to affect mRNA stability.

1.2 mRNA Decay and Deadenylation

For mRNAs, regulation of gene expression largely occurs by controlling cytoplasmic transcript abundance, thereby affecting the amount of gene product. Transcript abundance is a function of its synthesis and decay. Most mRNAs are thought to be constitutively degraded to maintain homeostasis; moreover, decay of specific transcripts is regulated to allow cells to adapt to changing conditions (Garneau et al., 2007). mRNA decay also removes aberrant transcripts to prevent deleterious protein accumulation or resource consumption (Doma and Parker, 2007).

mRNA decay determines its half-life ($t_{½}$), affecting gene expression. A short-lived mRNA may be translated less before its degradation; high transcript turnover also enables a rapid response to stimuli. Thus, eukaryotic mRNAs and their isoforms display a range of $t_{½}$ from minutes to hours (Friedel et al., 2009; Geisberg et al., 2014; Wang et al., 2002b). Moreover, transcriptomics and proteomics have revealed that regulation of mRNA $t_{½}$ involves the interplay between proteins and *cis*- and *trans*-acting RNAs (Rabani et al., 2017; Sohrabi-Jahromi et al., 2019; Tuck and Tollervey, 2013). Computational models have been developed to simulate mRNA decay system dynamics, but were limited by the achievable complexity (Cao and Parker, 2001, 2003).

While there are many pathways of eukaryotic mRNA turnover, canonical mRNA decay is initiated by poly(A) tail shortening (deadenylation), reducing Pab1 affinity for mRNA (Figure 1.2) (Parker and Song, 2004). Pab1 dissociation destabilises cap binding by eIF4E, inhibits translation initiation, and allows decapping by the decapping complex Dcp1-Dcp2 (Coller and Parker, 2004). Alternatively, Pab1 dissociation leaves a 3' oligo(A) which can be oligouridylated (Section 1.3.2). Oligo(U) or oligo(A) is bound by the Lsm1-7-Pat1 complex, recruiting Dcp1-Dcp2 (Chowdhury and Tharun, 2009). Deadenylation and decapping generate unprotected 5' and 3' ends, which are rapidly acted on by the 5'-3' exoribonuclease Xrn1 (Nagarajan et al., 2013) and the 3'-5' cytoplasmic exosome (Houseley et al., 2006) to fully degrade the mRNA.

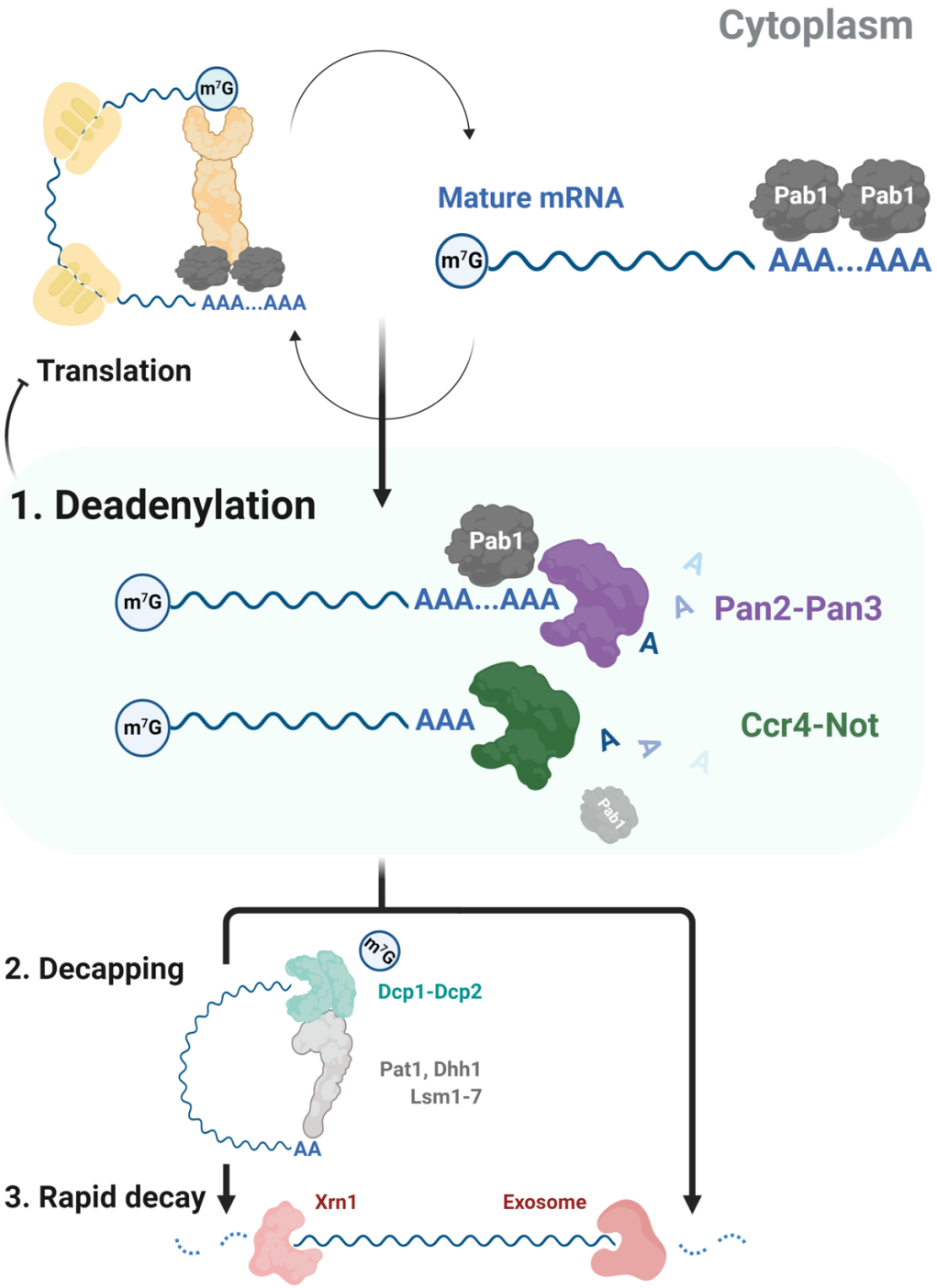


Figure 1.2 Eukaryotic mRNA decay. Canonical transcript degradation is initiated by poly(A) tail shortening (deadenylation). Dissociation of poly(A) binding protein and binding of the Lsm1-7-Pat1 complex are thought to enable the decapping complex Dcp1-Dcp2 to be recruited to remove the m^{7}G cap. This process creates free 5' and 3' ends which are rapidly degraded by exoribonucleases (Xrn1 and the cytoplasmic exosome), removing the remainder of the transcript.

The canonical pathway of mRNA decay is complicated by observations that different transcripts undergo unique degradation pathways. For some *S. cerevisiae* and *S. pombe* transcripts, decapping occurs independently of deadenylation to initiate RNA decay (Badis et al., 2004; Muhlrad and Parker, 2005; Rissland and Norbury, 2009). Alternatively, short interfering RNAs (siRNAs) can cause endonucleolytic cleavage of mRNA to form exonuclease-susceptible fragments (Carthew and Sontheimer, 2009; Gill et al., 2004). Finally, mRNA decay feeds back to transcription and vice versa via poorly understood mechanisms. Yeast Pol II subunits Rpb4/Rpb7 shuttle into the cytoplasm to affect mRNA decay and translation (Harel-Sharvit et al., 2010; Lotan et al., 2005; Lotan et al., 2007), transcription initiation and elongation affect mRNA decay (Bregman et al., 2011; Slobodin et al., 2020), and depletion of cytoplasmic degradation factors leads to a decrease in transcription (Sun et al., 2013; Sun et al., 2012). These mechanisms allow transcript abundance to be buffered under ever-changing cellular conditions.

1.2.1 Deadenylation

Cytoplasmic poly(A) tail shortening was first observed through radioactive pulse-chase studies (Sheiness and Darnell, 1973). Deadenylation was soon linked to mRNA decay, as poly(A) tail shortening generated transient deadenylated intermediates which were rapidly degraded (Brewer and Ross, 1988; Mercer and Wake, 1985; Shyu et al., 1991; Wilson and Treisman, 1988). A seminal study showed that deadenylation was the first and rate-limiting step of mRNA decay in yeast; full decay only occurred when the tail was shortened to a threshold length. Accordingly, when a destabilising 3' UTR was introduced into an otherwise stable transcript, the rate of deadenylation and subsequent onset of decay were accelerated, affecting the half-life of the chimeric transcript (Decker and Parker, 1993). The requirement of deadenylation for mRNA decay was further confirmed in mammals (Couttet et al., 1997). Finally, the canonical pathway of degradation was determined by showing that deadenylated intermediates had a m^7G cap, and that depletion of the 5'-3' Xrn1 exonuclease stabilised deadenylated and

decapped intermediates (Hsu and Stevens, 1993; Muhlrad et al., 1994, 1995). Thus, deadenylation represses gene expression by poly(A) tail shortening to reduce translation efficiency; moreover, it commits mRNAs to further decay by being the first and rate-limiting step of degradation. The importance of deadenylation in regulating gene expression is further demonstrated by its inhibition and modulation upon viral infection, thereby stabilising viral transcripts (Dougherty et al., 2011; Garneau et al., 2008; Mitton-Fry et al., 2010).

Given the central role of deadenylation in gene expression, it is unsurprisingly important in other cellular processes. Deadenylation is involved in monitoring codon optimality; transcripts with lower codon optimality are more rapidly deadenylated and have shorter half-lives (Narula et al., 2019; Presnyak et al., 2015; Webster et al., 2018). A recent paper has identified the molecular basis of how a translating ribosome can be linked to the deadenylation machinery (Buschauer et al., 2020). Deadenylation is also involved in nonsense-mediated decay of transcripts with premature terminating codons (Chen and Shyu, 2003; Lejeune et al., 2003; Loh et al., 2013). Deadenylation is further important in cellular conditions which require rapid changes in gene expression. For instance, deadenylation is globally inhibited during cell stress, suggesting that reduced degradation is important to maintain mRNA homeostasis (Gowrishankar et al., 2006; Hilgers et al., 2006). Finally, in oocytes and embryos, targeted deadenylation is antagonised by cytoplasmic readenylation, maintaining transcript dormancy until early development, when maternal mRNAs are selectively degraded (Charlesworth et al., 2013; Fox and Wickens, 1990; Huarte et al., 1992; Varnum and Wormington, 1990).

Eukaryotic deadenylation in the cytoplasm is mainly carried out by two highly conserved multiprotein complexes: Ccr4-Not and Pan2-Pan3. In certain cellular contexts and organisms, other deadenylase enzymes may predominate. Deadenylase enzymes, which generally belong to the DEDD or EEP exonuclease superfamilies (Zuo and Deutscher, 2001), will be further discussed below, with a particular focus on their mechanisms of poly(A) recognition.

1.2.2 Pan2-Pan3

Pan2-Pan3 (Poly(A) Nuclease) was identified as responsible for Pab1-dependent poly(A) tail shortening in yeast (Sachs and Davis, 1989a; Sachs and Deardorff, 1992). Accordingly, depletion of Pab1 or the Pan complex resulted in longer global poly(A) tail lengths. The complex comprises two non-essential components: Pan2 and Pan3 (Boeck et al., 1996; Brown et al., 1996). Pan2 is a distributive, 3'-5', metal-ion-dependent exoribonuclease (Lowell et al., 1992), whereas Pan3 is the constitutive positive regulator of Pan2 activity (Brown and Sachs, 1998a). Both proteins are highly conserved across eukaryotes, suggesting a significant cellular role (Uchida et al., 2004). In humans, the importance of PAN2-PAN3 is disputed. PAN2 deletion led to no change in global poly(A) tail length (Yi et al., 2018). On the other hand, PAN3 knockdown led to transcriptome-wide changes in transcript stability and reduced deadenylation of model transcripts (Chen et al., 2018).

In addition to its involvement in general mRNA deadenylation, Pan2-Pan3 may play distinct roles *in vivo*. It regulates the stability HIF1A and RAD5 transcripts, which are involved in the hypoxic response and replicative stress respectively (Bett et al., 2013; Hammet et al., 2002). Thus, Pan2-Pan3 may display transcript specificity, particularly during cell stress. Furthermore, Pan2-Pan3 is a key component of cytoplasmic phase-separated droplets known as P-bodies, as knockdown of Pan3 leads to reduced P-body formation (Zheng et al., 2008). During normal or stressed conditions, mRNAs distribute between P-bodies, the cytoplasm, and stress granules, dynamically regulating their decay and fate (Decker and Parker, 2012). mRNA association with P-bodies is globally associated with down-regulation due to the condensation of decay factors and inaccessibility to polysomes (Cougot et al., 2004; Hubstenberger et al., 2017). However, whether a P-body transcript is subjected to constitutive decay is context-dependent; P-bodies may also stabilise or store mRNAs (Aizer et al., 2014; Wang et al., 2018a). Finally, pull-downs with nuclear extract have implicated Pan2-Pan3 in trimming newly synthesised poly(A) tails in

the nucleus (Brown and Sachs, 1998a; Mangus et al., 2004b). The nuclear role of Pan2-Pan3 is corroborated by slight mRNA accumulation in the nucleus upon its deletion as well as genetic interactions with the mRNA export receptor Mex67, suggesting that its deletion may cause incorrect post-transcriptional mRNA processing (Dunn et al., 2005; Estruch et al., 2009). However, its putative role in nuclear poly(A) trimming is inconsistent with observations that the majority of Pan2-Pan3 is evenly localised to the cytoplasm in mammals (Zheng et al., 2008) and in yeast (Dheur et al., 2005; Huh et al., 2003). Therefore, the dominant role of Pan2-Pan3 is likely to be bulk cytoplasmic deadenylation.

Pan2-Pan3 comprises one molecule of Pan2 and two molecules of Pan3 (Jonas et al., 2014; Schafer et al., 2014; Wolf et al., 2014); both are multidomain proteins with numerous functionally distinct regions (Figure 1.3). The structured domains are largely conserved across different species (Supplementary Figure 1-2), with the greatest variation observed in low-complexity regions and loop regions within domains.

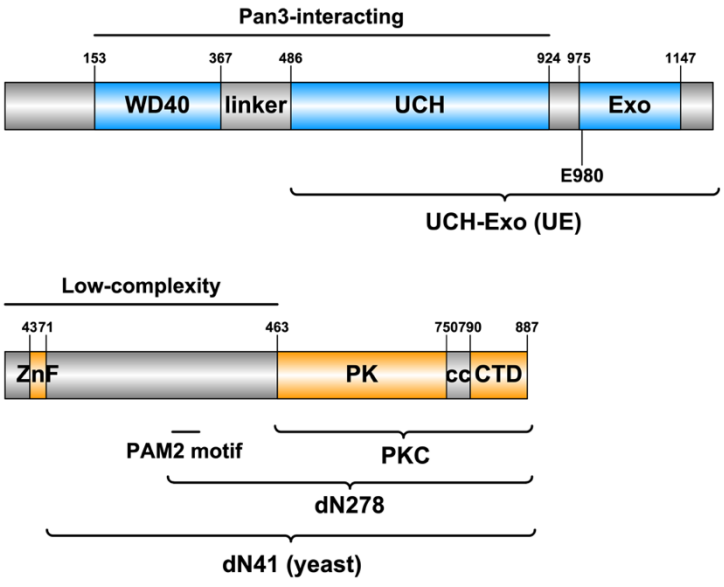


Figure 1.3 Domain schematics for *H. sapiens* PAN2 (blue) and PAN3 (orange). Domain boundaries are marked by residue number. Horizontal lines show functionally significant regions of the proteins. Curved brackets denote boundaries of constructs used in this dissertation.

1.2.2.1 Pan2

Pan2 has an N-terminal WD40 domain adopting the canonical β-propeller fold (Jonas et al., 2014). A lateral surface of the WD40 domain interacts with the Pan3 C-terminal domain (CTD) to mediate complex formation (Figure 1.4A). Pull-down assays showed that WD40 is required for Pan3 interaction in humans and flies, but is dispensable in yeast (Jonas et al., 2014; Schafer et al., 2014). One of the canonical interaction surfaces of the WD40 domain is exposed, suggesting it may interact with binding partners (Figure 1.4D, inset) (Schafer et al., 2019).

A linker connecting the Pan2 N- and C-terminal domains is required to interact with the Pan3 dimer, enforcing the 1:2 stoichiometry of Pan2-Pan3 (Schafer et al., 2014; Wolf et al., 2014). While the linker is predicted to be disordered (Supplementary Figure 3), it forms a defined structure when bound to Pan3, interacting with the Pan3 CTD and coiled coil by forming short secondary structural elements and inserting residues into surface grooves (Figure 1.4B).

The Pan2 C-terminus comprises a structurally contiguous module (UCH-Exo), where the ubiquitin hydrolase (UCH) and exonuclease (Exo) domains extensively interact (Jonas et al., 2014; Schafer et al., 2014). UCH-Exo is an active deadenylase alone but requires Pan3 for full activity (Schafer et al., 2014). UCH belongs to the herpesvirus-associated ubiquitin-specific protease (HAUSP) family, but lacks conserved catalytic residues. Thus, Pan2 cannot hydrolyse ubiquitin *in vitro* (Quesada et al., 2004). In Pan2-Pan3, UCH binds the "closed" side of the asymmetric Pan3 dimer (Figure 1.4C). The Exo domain is a canonical DEDD exonuclease (Zuo and Deutscher, 2001) and is the most conserved domain of Pan2 among different species (Uchida et al., 2004). In the Pan2 active site, conserved acidic side chains and a histidine residue (DEDDh) coordinate two metal ions, which likely catalyse cleavage by activating the attacking nucleophile and stabilising the leaving group, via the canonical two-metal-ion catalytic mechanism (Figure 1.4D, inset) (Schafer et al., 2014; Steitz and Steitz, 1993). Accordingly, mutating DEDD side chains or chelating metal ions by EDTA ablated Pan2 activity (Lowell

et al., 1992; Schafer et al., 2014). Nonetheless, before the work presented in this dissertation, there was no structure of Pan2-oligo(A); thus, whether the exonuclease has intrinsic poly(A) specificity and its mechanistic basis were unknown.

1.2.2.2 Pan3

Pan3 contains a low-complexity N-terminus and a structured C-terminal half (Supplementary Figure 4). At the N-terminus, Pan3 contains a CCCH zinc finger domain (ZnF) which specifically binds poly(A) with micromolar affinity (Wolf et al., 2014). The structure of the Pan3 ZnF is similar to that of the Tristetraprolin family (TTP, Section 1.4.2), which binds AU-rich sequences, suggesting a potential common evolutionary origin. Accordingly, deletion of the Pan3 zinc finger leads to reduced affinity for polyadenylated RNA, implying that ZnF confers a degree of poly(A) specificity. However, as ZnF is tethered to the remainder of the complex via a low-complexity region, it is unlikely to influence poly(A) specificity in the active site, but rather confers higher affinity of the full Pan2-Pan3 complex for poly(A).

The Pan3 low-complexity region contains a PABP-interacting motif 2 (PAM2). PAM2 interacts with the C-terminal domain (CTD) of Pab1 with micromolar affinity, recruiting Pan2-Pan3 to the Pab1-poly(A) RNP (Kozlov et al., 2004; Kozlov et al., 2010). PAM2 adopts an extended conformation in a groove in the Pab1 CTD (Figure 1.9C). However, yeast and human Pab1 CTDs are structurally divergent and bind PAM2 differently; this may result in species specificity for Pab1-Pan3 interactions (Siddiqui et al., 2007). Mutation of a key phenylalanine in PAM2 recapitulates the longer poly(A) tails observed in pan2Δ and pan3Δ strains, suggesting that the PAM2-Pab1 interaction regulates poly(A) tail length (Siddiqui et al., 2007). The importance of PAM2 is highlighted by predicted neighbouring phosphorylation sites, which can influence PAM2 affinity for the Pab1 CTD (Huang et al., 2013). Moreover, human PAN3 has two isoforms which differ in the addition of a sequence upstream of PAM2. These two

isoforms bind the Pab1 CTD with different affinities, potentially affecting steady-state poly(A) tail length and P-body distribution (Chen et al., 2018). Thus, the interaction between PAM2 and Pab1 CTD may regulate poly(A) tail shortening under different conditions.

The C-terminal half of Pan3 consists of a pseudokinase domain (PK) and a small, uniquely-folded C-terminal domain (CTD), linked by a coiled coil. Crystal structures have shown that PKC forms a constitutive asymmetric dimer (Figure 1.4D), with asymmetry enforced by aspartate and glycine amino acids, causing the coiled coil to be kinked (Christie et al., 2013). The coiled coil mediates dimerisation and is stabilised by hydrophobic interactions. This leads to "open" (CTD does not contact PK) and "closed" (CTD contacts PK) sides of the dimer, which bind Pan2 WD40 and UCH domains respectively. Interactions with Pan2, particularly the linker region, reinforce Pan3 dimer asymmetry (Schafer et al., 2014; Wolf et al., 2014). Interestingly, the metazoan Pan3 dimer pseudokinase interface contains a putative tryptophan binding site, commonly found in GW182-family proteins involved in miRNA-mediated post-transcriptional mRNA regulation (Figure 1.4D, inset). This is consistent with studies showing that RNA-induced silencing complexes directly recruit Pan2-Pan3 (Section 1.4.2).

The pseudokinase domain adopts a characteristic bilobed structure, but lacks the conserved motifs and activation loop required for activity. Thus, PK is likely inactive. Nonetheless, a crystal structure showed that PK can bind ATP in an Mg^{2+}-dependent manner, stabilising the Pan3 dimer (Figure 1.4D, inset) (Christie et al., 2013). As the adenine moiety is common to ATP and poly(A) RNA, this raised the possibility that PK contributes to poly(A) binding. Accordingly, mutation of the ATP-binding pocket reduced Pan2-Pan3 activity *in vitro* (Christie et al., 2013; Schafer et al., 2014). However, PK does not contact substituents which are specific to the adenine moiety, and extension of 3' OH would sterically clash with Pan3, raising the question of how PK could bind poly(A). An alternative explanation could be that constitutive ATP binding stabilises folding of the Pan3 dimer core.

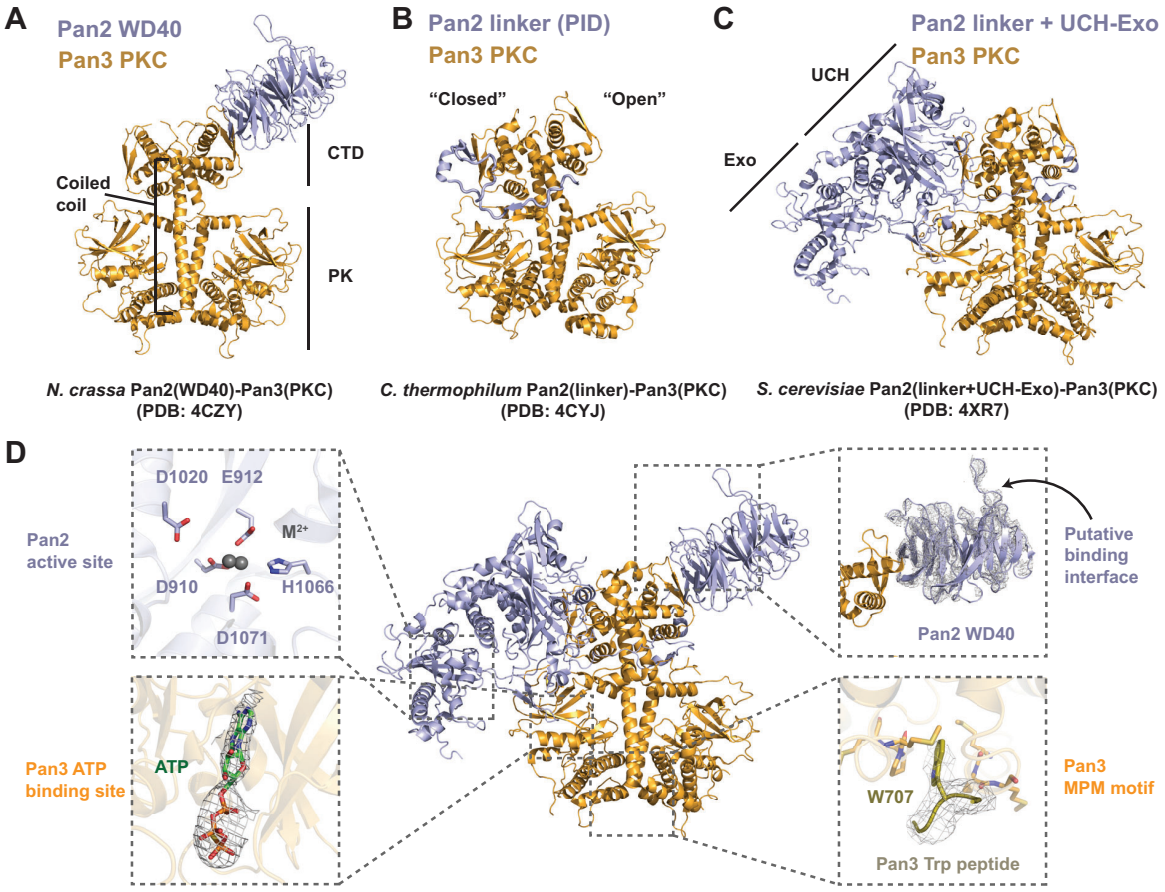


Figure 1.4 Structural model for Pan2-Pan3. Pan2 is blue, Pan3 is orange. (**A**) Structure of *N. crassa* Pan2(WD40)-Pan3(PKC). (**B**) Structure of *C. thermophilum* Pan2(linker)-Pan3(PKC). (**C**) Structure of *S. cerevisiae* Pan2(linker+UCH-Exo)-Pan3(PKC). (**D**) Structural model of the of the Pan2-Pan3 core, predicted by superposition of the crystal structures in panels A-C. Inset (clockwise from bottom left): the ATP binding site in the PK cleft; the Pan2 Exo active site (catalytic side chains are shown as sticks, and M^{2+} ions are shown as grey spheres); a putative interaction surface on the Pan3 WD40 domain; the putative interaction site of GW182 GW-rich peptides at the Pan3 dimerisation interface. The mesh shows the 2mFo-DFc maps contoured to 1.5σ.

1.2.3 Ccr4-Not

Ccr4-Not proteins were first identified as transcription regulators. Ccr4 (carbon catabolite repression 4) was implicated in regulating non-fermentative growth (Denis, 1984; Draper et al., 1994), whereas the Not proteins (negative on TATA) were involved in negatively regulating TATA-dependent HIS3 transcription (Collart and Struhl, 1993, 1994). They were soon found to form a constitutive complex (Liu et al., 1998), which could be purified endogenously or through recombinant expression and reconstitution (Chen et al., 2001; Raisch et al., 2019; Stowell et al., 2016a). Functionally distinct modules centred around the large Not1 scaffold have also been identified and biochemically and structurally characterised (Basquin et al., 2012; Bawankar et al., 2013; Bhaskar et al., 2013; Boland et al., 2013; Mathys et al., 2014).

In the context of mRNA decay, Ccr4-Not is thought to be the main eukaryotic multiprotein deadenylase complex. Ccr4-Not contains two deadenylase enzymes: Ccr4 and Caf1 (Zuo and Deutscher, 2001). These deadenylases, together with the central MIF4G domain of the Not1 scaffold, form a structurally contiguous U-shaped nuclease module (Basquin et al., 2012). In this module, Ccr4 is tethered to the Ccr4-Not complex through Caf1; disrupting the Not1-Caf1 interface prevented Ccr4 association with the complex (Basquin et al., 2012; Petit et al., 2012). Although their active sites are distant in the nuclease module, Ccr4 and Caf1 are both active (Maryati et al., 2015; Stowell et al., 2016a). There is growing evidence that Ccr4 and Caf1 are functionally different: for example, Ccr4 is the only deadenylase in Ccr4-Not that can remove Pab1 (Webster et al., 2018; Yi et al., 2018), whereas Caf1 can interact with the antiproliferative protein Tob (Horiuchi et al., 2009) and appears more important in degradation of transcripts with non-optimal codons (Webster et al., 2018).

1.2.3.1 Ccr4

Ccr4 (metazoa: CNOT6/CNOT6L) is an EEP-family exonuclease specific for single-stranded poly(A) RNA (Chen et al., 2002; Viswanathan et al., 2003). In yeast, it is considered the major deadenylase as its deletion led to longer global poly(A) tails (Tucker et al., 2002). However, Ccr4 is non-essential, suggesting functional redundancy with other deadenylases or compensation from other decay pathways. *In vivo*, Ccr4 is important in cell proliferation and cell cycle progression, though its global role in regulating transcript stability means that pleiotropic effects cannot be ruled out (Mittal et al., 2011; Morita et al., 2007).

Ccr4 consists of two domains: its LRR domain interacts with Caf1 and is thus important for incorporation into Ccr4-Not (Basquin et al., 2012; Clark et al., 2004), and its exonuclease domain carries out deadenylation. A model of poly(A) recognition by Ccr4 has been proposed based on a crystal structure of CNOT6L in complex with single-stranded oligo(A) DNA (Wang et al., 2010). The nucleic acid substrate was bound at the base of the active site cleft of the heart-shaped exonuclease domain (Figure 1.5A). The scissile phosphate faced two active site Mg^{2+} ions, coordinated by conserved residues (Figure 1.5B, top inset). Oligo(A) DNA formed van der Waals and hydrogen bonding interactions with active site amino acid side chains, but most adenine bases were poorly resolved (A_{+2}) or showed no significant density (A_{-1}). The Ccr4 active site appears to specify for adenines via A_{+1}. A_{+1} stacked against Phe484, and its *N*6 amine formed a hydrogen bond with the carbonyl oxygen of Asn412 (Figure 1.5B, bottom inset). Guanines would have a carbonyl oxygen at C6, precluding hydrogen bond formation; cytosines and uracils would provide insufficient stacking energy with Phe484 and may sterically clash with the active site due to the pyrimidine ring.

It must be noted that the structure, as modelled, predicts that an oligo(A) substrate would be cleaved endonucleolytically to release an A_2 product. Furthermore, 5' extension of the modelled substrate would result in significant steric clashes with the active site. The model is

therefore contradictory to biochemical studies showing that Ccr4 is a 3'-5' exonuclease. Thus, it is likely that oligo(A) ssDNA is either wrongly modelled in the Ccr4 active site, or the lack of a 2' OH group of oligo(A) DNA caused different binding in the Ccr4 active site.

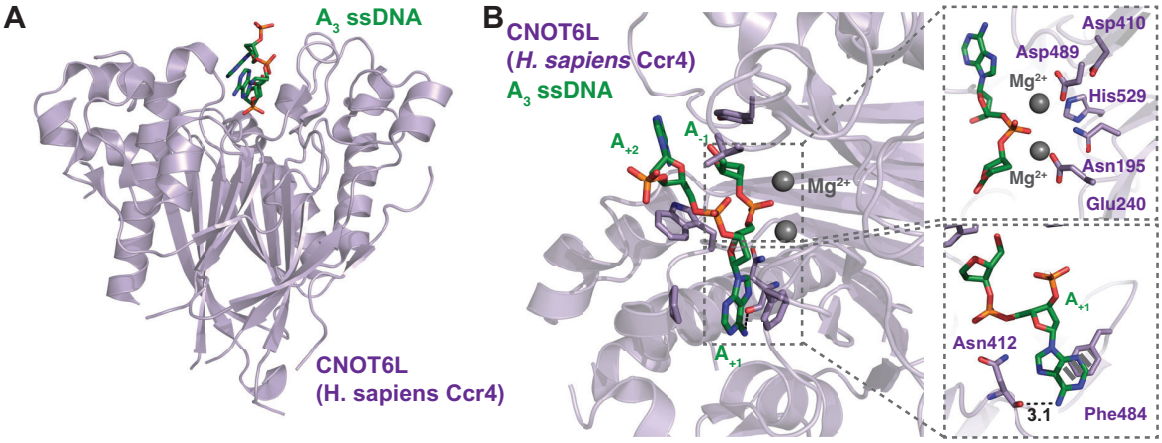


Figure 1.5 Structure of *H. sapiens* CNOT6L (purple) bound to single-stranded oligo(A) DNA (green sticks) (PDB: 3NGO). (**A**) Overall structure of CNOT6L-A_3 ssDNA. ssDNA is bound in an active site cleft in CNOT6L. (**B**) CNOT6L active site. Inset, top: active site residues (purple sticks) of CNOT6L coordinate two metal ions (grey spheres) for catalysis. Inset, bottom: specificity for A_{+1} is determined by a hydrogen bond (black line) between Asn412 carbonyl oxygen with *N*6 amine and stacking (grey dash) on Phe484.

1.2.3.2 Caf1

Caf1 (Ccr4-associated factor 1; metazoa: CNOT7/CNOT8) is a DEDD-family exonuclease which was found to act in a similar pathway to Ccr4 (Sakai et al., 1992). It was soon identified as an active 3'-5' poly(A)-specific exonuclease. However, its deadenylase activity was deemed less important than that of Ccr4 as its deletion did not greatly affect poly(A) tail length, and Caf1 catalytic mutants could complement its deletion. This indicated that Caf1 mainly bridges Ccr4 with the remainder of Ccr4-Not (Bianchin et al., 2005; Tucker et al., 2001; Viswanathan et al., 2004). In contrast, mutagenesis and synthetic lethality studies suggest that the deadenylation activity of Caf1 is important (Ohn et al., 2007). *In vivo*, Caf1 is involved in eukaryotic development (Berthet et al., 2004; Molin and Puisieux, 2005; Mostafa et al., 2020) and cell proliferation (Aslam et al., 2009; Doidge et al., 2012). Nonetheless, these studies are complicated by the requirement of Caf1 for Ccr4 association with Ccr4-Not.

Multiple studies have determined Caf1 crystal structures, alone (*S. cerevisiae* and *S. pombe* Caf1; PDB: 1U0C, 2P51, 3G0Z, 3G10) (Andersen et al., 2009; Jonstrup et al., 2007; Thore et al., 2003), in complex with Tob (*H. sapiens* CNOT7, PDB: 2D5R) (Horiuchi et al., 2009), or in complex with the Not1 MIF4G domain (*S. cerevisiae* Caf1, PDB: 4B8A) (Basquin et al., 2012). The structure of Caf1 is conserved across species and does not undergo significant conformational change upon partner binding (Figure 1.6). Caf1 adopts the canonical DEDD exonuclease fold and contains an active site cleft, where two metal ions are coordinated by conserved DEDD side chains (Figure 1.6). In *S. cerevisiae*, the canonical DEDD motif is replaced by SEDQ; however, yeast Caf1 is active *in vitro* (Thore et al., 2003). Under physiological metal ion concentrations, the Caf1 active site preferentially bound Mn^{2+} and Zn^{2+} ions (Figure 1.6, inset). The identity of the metal ions in the Caf1 active site also affected its deadenylation activity (Andersen et al., 2009). Nonetheless, the aforementioned structures did not contain RNA. Thus, the basis of Caf1 specificity and how poly(A) binds to the Caf1 active site were unknown prior to this dissertation.

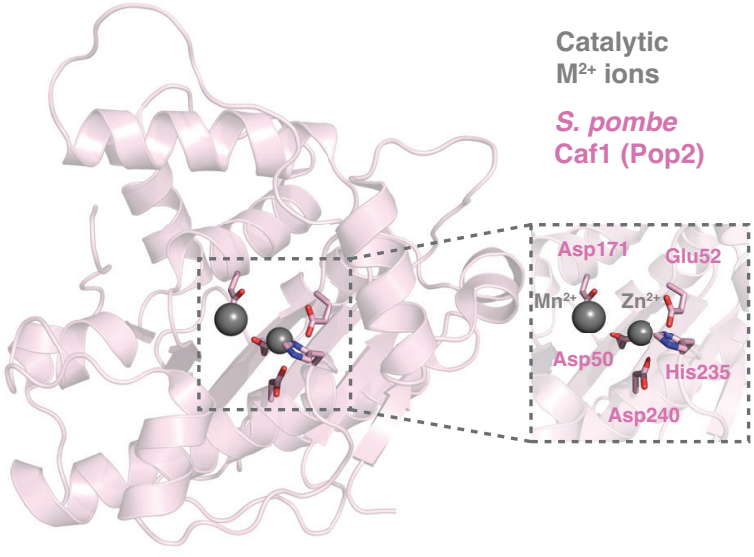


Figure 1.6 Structure of *S. pombe* Caf1 (PDB: 3G0Z). Caf1 is a pink cartoon, and active site metal ions are grey spheres. Inset: active site residues (pink sticks) coordinate two metal ions for catalysis.

1.2.4 PARN

PARN (Poly(A)-specific Ribonuclease) was identified as a DEDD-family poly(A)-selective exonuclease in mammals (Astrom et al., 1991, 1992; Korner and Wahle, 1997). Cytoplasmic deadenylation by PARN was implicated in early metazoan (Kim and Richter, 2006; Korner et al., 1998) and *Arabidopsis* development (Chiba et al., 2004; Reverdatto et al., 2004). In contrast with known deadenylases, PARN is stimulated by the 5' m^7G cap (Gao et al., 2000; Martinez et al., 2001). The mechanism of cap binding has been elucidated by NMR solution and crystal structures, showing that m^7G is recognised by the C-terminal RRM domain via a non-canonical mechanism (Monecke et al., 2008; Nagata et al., 2008; Niedzwiecka et al., 2016). The RRM domain can also bind poly(A), which may confer specificity to PARN (Nilsson et al., 2007). However, the lack of effect on poly(A) tail length and shortening upon knockdown (Yamashita et al., 2005; Yi et al., 2018) suggests a more limited role of PARN in cytoplasmic deadenylation.

The mechanism of poly(A) recognition by PARN has been elucidated by a crystal structure of its DEDD exonuclease domain bound to oligo(A) RNA (Figure 1.7A) (Wu et al., 2005). In this structure, PARN was homodimeric and RNA was bound in each active site cleft, with the scissile phosphate facing the conserved DEDD motif (Figure 1.7B, inset). Oligo(A) RNA contacted several amino acids, with the terminal adenine (A_{+1}) stacking against the F115 aromatic ring (Figure 1.7B). However, there were no base-specific interactions between adenine and the active site. Furthermore, examination of the electron density map revealed ambiguities in the modelled RNA, complicating our understanding of how RNA is bound in the PARN active site. Thus, while the PARN active site is specific for adenines (Henriksson et al., 2010), how PARN distinguishes between poly(A) and other ribonucleotides is unknown. Finally, a more recent structure of RRM and exonuclease domains in complex with the m^7G cap (PDB: 3D45) suggested that, in one monomer, m^7G binding may clash with poly(A) in the exonuclease active site (closed), whereas in the other, m^7G binding is compatible with poly(A)

binding (open) (Wu et al., 2009). Thus, in the PARN dimer, it is plausible that one monomer is responsible for cap-binding while the other binds and shortens the poly(A) tail.

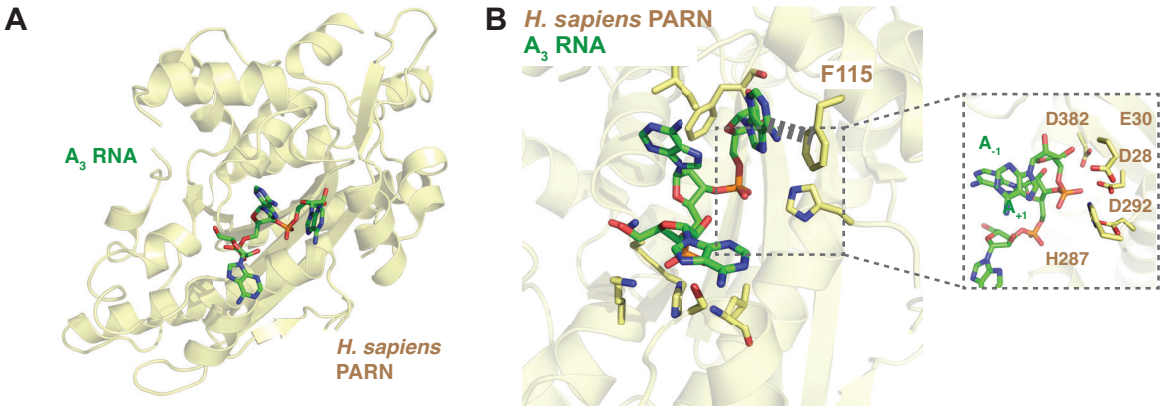


Figure 1.7 Structure of *H. sapiens* PARN (yellow) bound to single-stranded oligo(A) RNA (green sticks) (PDB: 2A1R). (**A**) Overall structure of PARN-A_3. A_3 is bound in the active site cleft. (**B**) Active site of PARN. Amino acids contacting A_3 are yellow sticks. The key stacking interaction with F115 is a grey dash. Inset: zoomed in view of the active site. Active site residues are pointed towards the scissile phosphate group.

The predominantly nuclear localisation of PARN suggests it plays an alternative physiological role to cytoplasmic deadenylation. Studies have revealed that PARN trims the oligo(A) 3'-end during non-coding RNA maturation. These include microRNAs (miRNAs) (Lee et al., 2019; Yoda et al., 2013), small nucleolar RNAs (snoRNAs) (Berndt et al., 2012; Son et al., 2018), piwi RNAs (piRNAs) (Tang et al., 2016), and telomerase *TERC* RNA (Moon et al., 2015). This agrees with the role of PARN mutations in developmental disorders such as dyskeratosis congenita, with patients showing aberrant *TERC* and snoRNA maturation (Dhanraj et al., 2015; Tummala et al., 2015). Thus, *in vivo*, PARN is likely more important in processing of small, non-coding RNAs.

1.2.5 Minor Deadenylases

In contrast to the major deadenylases, minor deadenylases play functionally distinct roles, particularly in higher eukaryotes. These generally belong to the DEDD and EEP exonuclease families, suggesting that they may have arisen through gene duplication (Figure 1.8).

Minor deadenylases are often functionally specialised, participating, for instance, in noncoding RNA processing. The EEP exonuclease ANGEL (hCCR4d), a Ccr4 homologue, forms a complex with the DEDD-family Caf1 homologue TOE1 (hCaf1z). This complex does not interact with CCR4-NOT subunits, and instead localises to nuclear Cajal bodies and trims snRNA 3'-ends (Deng et al., 2019; Faber et al., 2002; Lardelli et al., 2017; Wagner et al., 2007). Interestingly, TOE1 and PARN significantly overlap in their snRNA specificity, suggesting they may be partially redundant (Son et al., 2018). In addition, PNLDC1, a DEDD-family exonuclease, is thought to carry out 3'-end trimming and maturation of piwi RNAs (piRNAs), which silence transposons during spermatogenesis in many organisms (Anastasakis et al., 2016; Ding et al., 2017; Izumi et al., 2016; Nishimura et al., 2018; Zhang et al., 2017).

The roles of other putative deadenylases are less clear. PDE12, which degrades the immune signalling molecule 2'-5' oligo(A), is structurally homologous to Ccr4 (Wood et al., 2015). PDE12 also deadenylates mitochondrial oligo(A) and poly(A) tails, affecting mitochondrial transcript stability, their translation, and mitochondrial tRNA maturation (Pearce et al., 2017; Poulsen et al., 2011; Rorbach et al., 2011). Finally, nocturnin was a predicted EEP deadenylase by Ccr4 homology (Baggs and Green, 2003) and by its effect on poly(A) tails and transcript stability (Kojima et al., 2015; Niu et al., 2011). A recent study showed that nocturnin is an $NADP^+$/NADPH phosphatase (Estrella et al., 2019), consistent with its mitochondrial localisation and role in lipid metabolism (Green et al., 2007). The adenine moiety common to NADP(H) and poly(A) may suggest a common evolutionary origin of nocturnin and Ccr4.

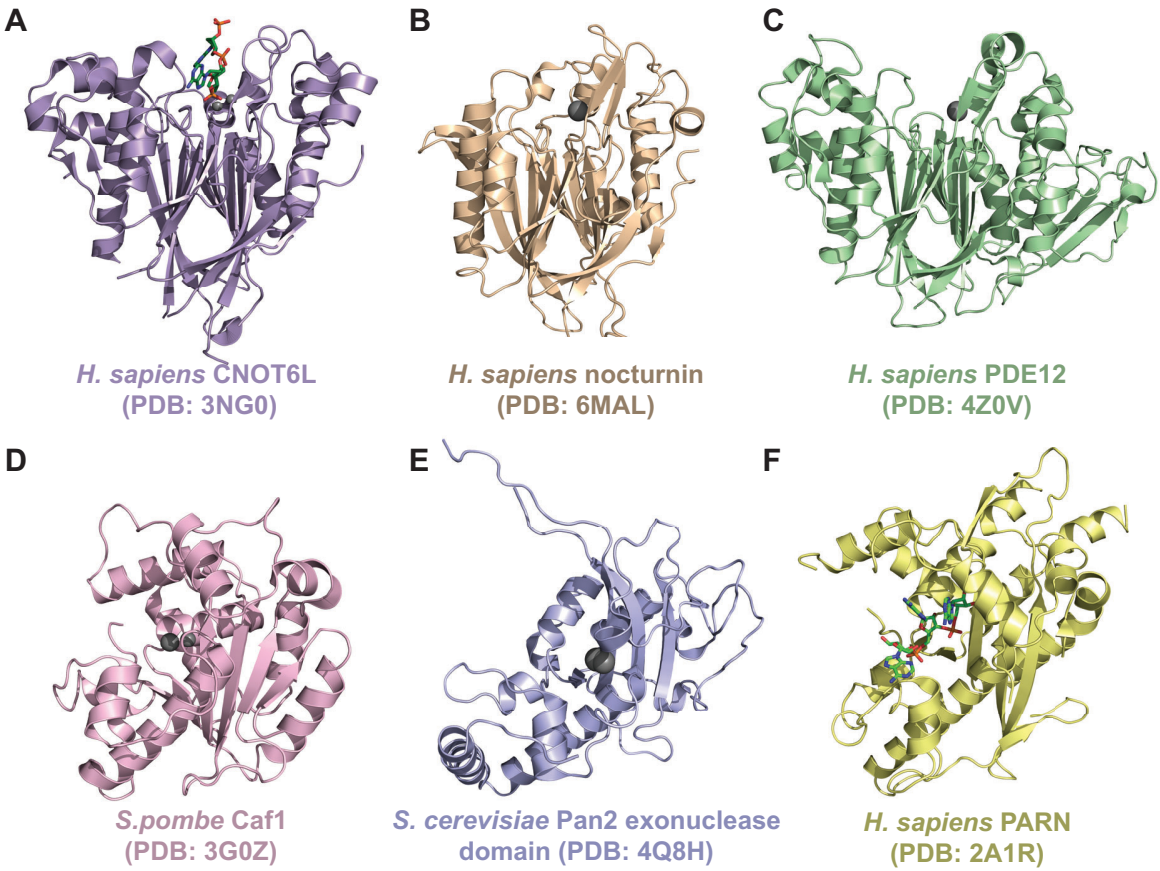


Figure 1.8 Structural conservation of deadenylases and related proteins. Nucleic acids, if present, are shown as green sticks. Metal ions, if present, are shown as grey spheres. The top row corresponds to the EEP-superfamily deadenylases. The bottom row corresponds to the DEDD-superfamily deadenylases. (**A**) Structure of *H. sapiens* CNOT6L bound to oligo(A) ssDNA. (**B**) Structure of *H. sapiens* nocturnin. (**C**) Structure of *H. sapiens* PDE12. (**D**) Structure of *S. pombe* Caf1. (**E**) Structure of *S. cerevisiae* Pan2 exonuclease domain. (**F**) Structure of *H. sapiens* PARN exonuclease domain bound to oligo(A) ssRNA.

Overall, the large repertoire of deadenylases with unique localisation and substrate specificities enables gene expression to be regulated in multifaceted ways. Nonetheless, it remains unclear how these deadenylases are regulated and display such tight substrate specificity. These questions can hopefully be addressed with further biochemical and structural investigations.

1.3 Properties of the Poly(A) Tail

1.3.1 Poly(A) Tail Length

Given the importance of the poly(A) tail in gene expression, its length is highly regulated. Early studies showed that poly(A) tail lengths vary within (Gorski et al., 1974; Mansbridge et al., 1974; Woo et al., 1975) and between mRNAs (Groner et al., 1974; Groner and Phillips, 1975). Fractionation and pulse-chase studies showed that newly synthesised poly(A) tails are 200-250 A's in mammals and 60-70 A's in yeast (Brawerman and Diez, 1975; Brown and Sachs, 1998b; Groner et al., 1974; Merkel et al., 1975; Sheiness et al., 1975). In the cytoplasm, poly(A) tails are shorter and more broadly distributed in length, implying that poly(A) tails are shortened after nuclear export (Ahlquist and Kaesberg, 1979; Groner and Phillips, 1975). However, these studies were limited by electrophoretic resolution and inability to isolate short poly(A) tails. PCR-based methods have interrogated the poly(A) tail length of model transcripts. PAT assays use oligo(dT) primers, which estimate tail lengths when combined with gene-specific primers. However, they are limited by throughput (Bazzini et al., 2012; Janicke et al., 2012; Salles et al., 1999). Microarray analyses (PASTA) can correlate poly(A) tail length with function on a transcriptome-wide level, but cannot measure precise tail lengths (Beilharz and Preiss, 2011).

Due to inaccuracies in sequencing long homopolymers, next-generation sequencing has only recently been used for poly(A) tail length determination. These transcriptome-wide studies are also less biased for long poly(A) tails. One method, TAIL-seq, uses Hidden Markov Models to monitor the fluorescence at the poly(A)-non-A transition to measure poly(A) tail lengths (Chang et al., 2014). Another technique, PAL-seq, calculates poly(A) tail length by the fluorescent signal of labelled streptavidin, which binds to biotinylated dUTP spike-ins during Illumina sequencing (Subtelny et al., 2014). These were later adapted for specific biological problems, including PacBio long-read sequencing to identify splice variants (Legnini et al.,

2019; Liu et al., 2019) and metabolic pulses to monitor poly(A) tail shortening dynamics (Eisen et al., 2020). Nanopore sequencing has also been used to measure poly(A) tail length by persistence of the adenine current, but is limited by precision (Krause et al., 2019).

Poly(A) tail sequencing confirmed that poly(A) tails are highly heterogeneous in length (Chang et al., 2014; Subtelny et al., 2014) and median tail lengths vary greatly depending on species (Lima et al., 2017; Subtelny et al., 2014). Metabolic pulse labelling in human cells showed that while global tail lengths are heterogeneous, mean poly(A) tails approach a steady-state length with longer labelling (Eisen et al., 2020). This is in line with studies showing that some species exhibit clear peaks of steady-state poly(A) tail length (Lima et al., 2017; Subtelny et al., 2014). Global or transcript poly(A) tails can be distributed with a 30-A periodicity (Kelly and Cox, 1982; Krowczynska and Brawerman, 1986; Lima et al., 2017; Stoeckle and Guan, 1993; Wilson and Treisman, 1988), which correlates with one Pab1 RNA-binding footprint (Baer and Kornberg, 1983). Thus, Pab1 and its distribution on the poly(A) tail may play a role in determining the steady-state length of poly(A) tails. However, it is unknown why this periodicity is only observed in some transcripts or species.

Transcriptome-wide measurements of poly(A) tail length have elucidated its biological role. In some species and cellular contexts, poly(A) tail length positively correlates with translation efficiency (Beilharz and Preiss, 2007; Eichhorn et al., 2016; Lackner et al., 2007; Lim et al., 2016; Morgan et al., 2017; Subtelny et al., 2014) and mRNA half-life (Chang et al., 2014). However, in HeLa cells (Legnini et al., 2019; Park et al., 2016), mice (Subtelny et al., 2014), and *C. elegans* (Lima et al., 2017), mRNAs with shorter poly(A) tails were associated with higher abundance and higher translation efficiency. This discrepancy suggests that poly(A) tail length may be regulated when large-scale changes in transcript stability and translation efficiency are required (e.g. during development or the cell cycle), and that somatic cells may have significant poly(A) homeostasis. Alternatively, it is possible that certain transcripts have differences in

poly(A) requirement for stability and translation efficiency, and these requirements are masked in global analyses. Indeed, representative mRNAs with different half-lives show variable deadenylation kinetic profiles (Eisen et al., 2020). Future work may further elucidate the complex relationship between poly(A) tail length and regulation of gene expression.

1.3.2 Modifications of the Poly(A) Tail

The poly(A) tail was long thought to be a homogeneous stretch of adenosines. However, sequencing techniques have identified non-A nucleotides in the poly(A) tail. These non-A nucleotides are conserved in higher eukaryotes (Chang et al., 2014; Legnini et al., 2019; Liu et al., 2019) and are mainly added by the non-canonical poly(A) polymerases TENT4A and TENT4B (Lim et al., 2018). Notably, guanosines near the 3'-end of the poly(A) tail correlated with increased transcript half-life, and transcripts displayed different guanylation frequencies, suggesting that guanylation (or non-A addition) may regulate mRNA stability (Chang et al., 2014). Nonetheless, the regulation of non-A nucleotide addition and the mechanism by which these stabilise mRNAs were unknown prior to the PhD.

Poly(A) tails can also be modified by a 3' oligo(U) tail. Oligouridylation by the non-canonical *S. pombe* poly(A) polymerase Cid1 was shown to stimulate mRNA decay by targeting uridylated transcripts to the Dis3L2 exonuclease for degradation (Malecki et al., 2013; Rissland et al., 2007; Rissland and Norbury, 2009). In metazoa, oligouridylation by TUT4 and TUT7 occurs on oligo(A) tails (Lim et al., 2014; Zuber et al., 2016). Oligo(U) can recruit Lsm1-7, which in turn binds the decapping complex to decap RNAs (Song and Kiledjian, 2007). However, before my dissertation, it was unknown whether oligouridylation alone affects deadenylation.

Chemical modifications of adenine can also regulate gene expression. For example, adenine can be methylated to form *N*6-methyladenine (m^6A), which is enriched in the 3' UTR (Meyer et al., 2012). This recruits RNA-binding proteins or affects local secondary structure to regulate

RNA degradation or splicing (Shi et al., 2019). Moreover, cap-proximal m^6A can be methylated on the 2' hydroxyl group to form m^6A_m, affecting mRNA stability and translation efficiency (Mauer et al., 2017; Sendinc et al., 2019). Finally, adenosines in brain-specific and Alu-derived transcripts can be deaminated to inosine (Shi et al., 2019; Tajaddod et al., 2016), which can cause translational stalling (Licht et al., 2019). However, it is unknown if these modifications occur in the poly(A) tail and how these might affect poly(A) processing.

1.3.3 The Structure of Poly(A)

Due to its biological prevalence and ease of synthesis, poly(A) RNA has been used as a model to study single-stranded RNA structure. Early studies revealed that, unlike the random coil thought to be formed by single-stranded RNA, poly(A) forms unique structures. At acidic pH, poly(A) can dimerise to form a double helix, which is maintained by inter-strand hydrogen bonds involving the protonated *N*6 amine (Holcomb and Tinoco, 1965; Rich et al., 1961; Safaee et al., 2013; Suck et al., 1976). Conversely, at physiological pH, poly(A) RNA can adopt a single-stranded helix, as shown by light scattering (Steiner and Beers, 1957), hypochromicity (Hashizume and Imahori, 1967), specific viscosity (Fresco and Klemperer, 1959), SAXS (Luzzati et al., 1964), and circular dichroism (CD) studies (Brahms, 1964; Brahms et al., 1966; Warshaw and Tinoco, 1965). Thermal denaturation and hypochromicity studies have estimated the thermodynamics of helix formation, suggesting that this process is entropically unfavoured (Applequist and Damie, 1966; Leng and Felsenfeld, 1966). Moreover, poly(A) helix formation is non-cooperative, as shown by CD studies suggesting low length dependence (Brahms et al., 1966; Poland et al., 1966) and gradual denaturation (as opposed to the rapid transition expected of cooperative folding) in temperature jump experiments (Dewey and Turner, 1979; Eisenberg and Felsenfeld, 1967; Leng and Felsenfeld, 1966; Porschke, 1973). Recently, poly(A) RNAs were studied by protein channels (Lin et al., 2010) and AFM (Seol et al., 2007), confirming previous observations of helix-coil kinetics and non-cooperativity of helix formation.

These experiments, particularly CD spectroscopy, suggested that the planar adenines are stacked perpendicularly to the helical axis. Indeed, adenine-adenine stacking is important for the helical conformation observed by CD (Van Holde et al., 1965) and NMR studies (Chan and Nelson, 1969; Ts'o et al., 1969). Later reports revealed that purine-purine stacking is more thermodynamically favourable than purine-pyrimidine or pyrimidine-pyrimidine stacking, providing an explanation for why adenines stack in the poly(A) helix (Brown et al., 2015; Friedman and Honig, 1995; Norberg and Nilsson, 1995; Takaya et al., 2008). Importantly, molecular dynamics simulations and NMR demonstrated that the favourable adenine-adenine stacking configuration was compatible with the constraints of the ribophosphate backbone in a poly(A) helix (Isaksson et al., 2004; Kondo and Danyluk, 1976; Lee et al., 1976; Mignon et al., 2005). Models of the single-stranded poly(A) helix were proposed based on a crystal structure of an A_3 dimer at low pH, where two bases formed the aforementioned duplex, but overhanging adenosines were stacked on the base-paired adenines, providing a two-nucleotide stretch of the poly(A) helix (Suck et al., 1976). In the absence of a crystal structure of single-stranded poly(A), analysis of NOESY footprints of an A-rich oligonucleotide in solution enabled another structural model to be devised, suggesting that the pyrimidine ring mainly stacks on top of the imidazole ring (Isaksson et al., 2004). From these models, the poly(A) helix has predicted 9-fold helical symmetry with a pitch of 25.4 Å and a rise of 2.8 Å per residue (Saenger et al., 1975). However, given the dynamics of oligo(A) and lack of fibre diffraction even at low temperatures (Pinamonti et al., 2017), it is likely that poly(A) rapidly interconverts between a random coil and a stacked, helical conformation at physiological pH, with most studies observing an ensemble of the two predominant conformations.

Interestingly, poly(A) is not the only structured polyribonucleotide. Poly(C), while less biologically prevalent, can form a base-stacked, right-handed helix with 6-fold symmetry at physiological pH (Arnott et al., 1976). Moreover, poly(C) displays significant CD signal and

forms a double-stranded helix at low pH (Fasman et al., 1964). Thus, single-stranded polyribonucleotides may form a higher order structure depending on the sequence of the RNA.

Despite comprehensive studies on the chemistry and physical properties of poly(A), its structure is rarely discussed or examined in a biological context. Indeed, the biological significance of adenine stacking has only been demonstrated in the context of viral kissing complexes, where adenine stacking stabilises a two-nucleotide RNA base pair (Chen and Garcia, 2012; Stephenson et al., 2013). Therefore, despite its prevalence, especially in eukaryotes, the biological significance of the intrinsic conformation of poly(A) was unknown.

1.4 RNA-binding Proteins

In cells, RNAs are associated with RNA-binding proteins (RBPs) to form ribonucleoprotein (RNP) complexes to carry out their function. RBPs bind and regulate mRNAs throughout their lifetime, enabling the cell to achieve quality control, to respond to specific stimuli, to maintain homeostasis for particular transcripts, and to increase the complexity and fine-tuning of gene expression. The functional units of RBPs are RNA-binding domains (RBDs), which are extraordinarily structurally diverse (Corley et al., 2020). In addition, these are often modular: multiple domains within the same protein can act independently or cooperatively to confer specificity and affinity. RBDs are often tethered to effector domains, enabling RNA-binding to be coupled with downstream biological effects. Finally, RBPs are also enriched in low-complexity regions, allowing for regulated subcellular, often membrane-free, localisation (Hofweber and Dormann, 2019), as well as interactions with protein partners.

Within the context of mRNA decay, RNA-binding proteins can be stabilising (e.g. the AU-rich element binding protein HuR (Peng et al., 1998)) or destabilising. Furthermore, the poly(A) tail can be bound by poly(A) binding proteins, affecting poly(A)-dependent processes such as deadenylation (Mangus et al., 2003). Below, I discuss the conserved cytoplasmic poly(A) binding protein Pab1 and its role in mRNA decay. Additionally, the mechanism of recruitment of deadenylase complexes by RBPs will be explored, culminating in the examination of a novel study which can be used to identify novel RBP binding partners of Pan2-Pan3.

1.4.1 Poly(A)-binding Protein

Cytoplasmic poly(A)-binding protein (yeast: Pab1, human: PABPC1) was discovered as the protein protecting a periodic 27-adenosine pattern upon poly(A) RNP digestion (Baer and Kornberg, 1980, 1983; Greenberg, 1980). The importance of Pab1 is shown by its conservation and essential nature in eukaryotes (Belostotsky and Meagher, 1993; Grange et al., 1987; Lefrere et al., 1990; Nietfeld et al., 1990; Sachs et al., 1986; Wang et al., 1992). Given the high abundance of Pab1 and its nanomolar affinity for poly(A), it is thought that poly(A) is completely coated with Pab1 following nuclear export (Ho et al., 2018). Pab1 was soon found to be important in the regulation of mRNA translation and decay (Bernstein et al., 1989; Grossi de Sa et al., 1988), accounting for previous results where polyadenylated transcripts were more efficiently translated (Jacobson and Favreau, 1983) and addition of exogenous poly(A) (presumably sequestering Pab1) can reduce transcript stability (Bergmann and Brawerman, 1977). The significance of Pab1 is reflected in frequent changes of its stability, localisation, and complex formation upon viral infection (Smith and Gray, 2010).

Pab1 contains three functional modules (Figure 1.9A): four tandem N-terminal RNA recognition motif domains (RRM) bind RNA with high affinity; the proline-rich linker (P-linker) mediates cooperative multimerisation; and the C-terminal domain (CTD) binds PAM2 to facilitate interactions with other proteins (Kozlov et al., 2002; Kozlov et al., 2001; Kuhn and Pieler, 1996; Melo et al., 2003; Riback et al., 2017). All domains play distinct roles in post-transcriptional regulation of gene expression and will be discussed below.

1.4.1.1 Effect of Pab1 on Translation

Pab1 is thought to stimulate translation initiation by interacting with eIF4G, an essential multidomain protein which binds the cap-binding eIF4E. eIF4G recruits the 40S ribosomal subunit to an mRNA (Prevot et al., 2003). The Pab1-eIF4G interaction was mapped to the eIF4G N-terminus and Pab1 RRM2 (Imataka et al., 1998; Kessler and Sachs, 1998; Otero et al., 1999; Tarun and Sachs, 1996; Tarun et al., 1997). The interaction of eIF4G with RRM1-2 has been elucidated by a crystal structure of Pab1 RRM1-2-oligo(A) bound to eIF4G(178-203) (Safaee et al., 2012b). The eIF4G peptide forms two antiparallel β-strands and a short α-helix upon binding to the RRM2 helical bundle opposite to the RNA-binding surface. This binding further increases the affinity of RRM1-2 for RNA.

Regulation of translation initiation at the 5'-end by the 3' Pab1-poly(A) RNP implies cross-talk between 5'- and 3'-ends of mRNAs (Vicens et al., 2018). Pab1 simultaneously and cooperatively binds poly(A) and eIF4G, suggesting that transcripts may be circularised (Figure 1.2). This would facilitate reinitiation of translation and allow fine-tuning of translation initiation by, for example, protein phosphorylation (Fraser et al., 1999; Le et al., 2000). Nonetheless, mRNA circularisation remains controversial due to contradictory evidence from atomic force microscopy, cryo-electron tomography, and single-molecule FISH studies (Adivarahan et al., 2018; Afonina et al., 2014; Brandt et al., 2010; Khong and Parker, 2018; Wells et al., 1998). It is plausible that mRNA only transiently circularises during translation initiation.

Finally, Pab1 regulates translation termination by binding eRF3, a GTPase which stimulates the release activity of eRF1. This interaction was elucidated by crystal structures of the Pab1 CTD bound to PAM2 of eRF3 (Figure 1.9C) (Kozlov and Gehring, 2010). This structure showed that PAM2 adopts an extended conformation in the Pab1 CTD cleft and binds mainly through hydrophobic interactions. Accordingly, mutations of the interaction interface caused changes in polysome profiles and read-through frequency *in vivo* (Roque et al., 2015).

1.4.1.2 RNA Binding by Pab1

The tandem N-terminal RRMs of Pab1 interact with high-affinity with poly(A). The RRM domains can compensate for Pab1 deletion, demonstrating the importance of RNA binding (Burd et al., 1991; Sachs et al., 1987). High-affinity interactions with poly(A) are mostly supported by RRM1-2 as they display similar affinity for oligo(A) compared to full-length protein (Burd et al., 1991; Khanam et al., 2006; Kuhn and Wahle, 2004). Poly(A) specificity has been revealed by crystal structures of human PABPC1 RRM1-2 bound to oligo(A) (Figure 1.9B) (Deo et al., 1999; Safaee et al., 2012b). Each RRM comprises a four-stranded antiparallel β-sheet backed by two short α-helices, and are connected by linkers which form short α-helices upon poly(A) binding. RRM1-2 bind poly(A) in an antiparallel manner; RRM1 binds the 3'-end of oligo(A) and RRM2 binds the 5'-end. Poly(A) adopts an extended conformation when bound to RRM1-2; the crystallographic footprint of RRM1-2 is nine adenosines, roughly agreeing with the minimum length of A_{12} required for high-affinity interaction. RRM1-2 forms extensive contacts with oligo(A), including hydrogen bonds between amino acid side chains and base-specific functional groups, as well as stacking of adenine on the aromatic rings of tyrosine or phenylalanine (Figure 1.9B, inset). Interestingly, adenines which do not form base-specific interactions with the RRM domains stack on neighbouring adenines.

On the other hand, no structure of RRM3-4 (or RRM1-4) has been determined, and thus, the molecular basis of RRM3-4 interaction with RNA is unknown. However, there is evidence that RRM3-4 play a distinct role compared to RRM1-2. While RRM3-4 retain high-affinity for poly(A), they display lower specificity *in vitro* and can bind to AU-rich or non-A RNAs (Deardorff and Sachs, 1997; Kuhn and Pieler, 1996; Mullin et al., 2004; Sladic et al., 2004). Because RRM3-4 bind RNA 5' to RRM1-2, this lower specificity raises the possibility that RRM3-4 could bind to the 3' UTR, especially when the poly(A) tail is shortened to a length which cannot support all four RRMs. Indeed, *in vivo* PAR-CLIP and CLIP-seq showed that

Pab1 is enriched at the 3'-terminus of the 3'-UTR, suggesting that RRM3-4 could bind to sequences adjacent to the poly(A) tail (Baejen et al., 2014; Kini et al., 2016). Moreover, Pab1 occupancy at the 3' UTR is disfavoured by RNA structure, suggesting naked, single-stranded RNA in the 3' UTR is required for correct Pab1 localisation (Moqtaderi et al., 2018). RRM3-4 binding to the 3' UTR would provide a potential mechanism by which the 5' Pab1 molecule is anchored in a fixed register relative to the poly(A) tail, causing poly(A) phasing of subsequent Pab1 molecules rather than a random distribution of Pab1 along the poly(A) tail (Section 1.3.1).

Despite the lack of high-resolution structural information of how all four RRMs recognise the poly(A) tail, the overall conformation of Pab1 on poly(A) has been investigated. Upon full-length PABPC1 binding, a 5' and 3'-labelled poly(A) substrate showed an increased FRET signal, in contrast with the decreased FRET signal when RRM1-2 were bound. Combined with mutational analysis, this suggested that full-length PABPC1 binding causes poly(A) bending due to the linker between RRM2 and RRM3 (Hong et al., 2017). Negative stain EM of poly(A) coated by Pab1 showed multiple Pab1 molecules bound end-to-end on poly(A); however, no significant RNA bending was visible (Sawazaki et al., 2018). Finally, Pab1 is thought to cooperatively multimerise on the poly(A) tail. Single-molecule nanopore detachment studies and mutational analysis of Pab1 showed that inter-subunit interactions are largely mediated by the P-linker, though controversy remains as to which region is responsible for cooperative multimerisation (Lin et al., 2012; Melo et al., 2003; Yao et al., 2007).

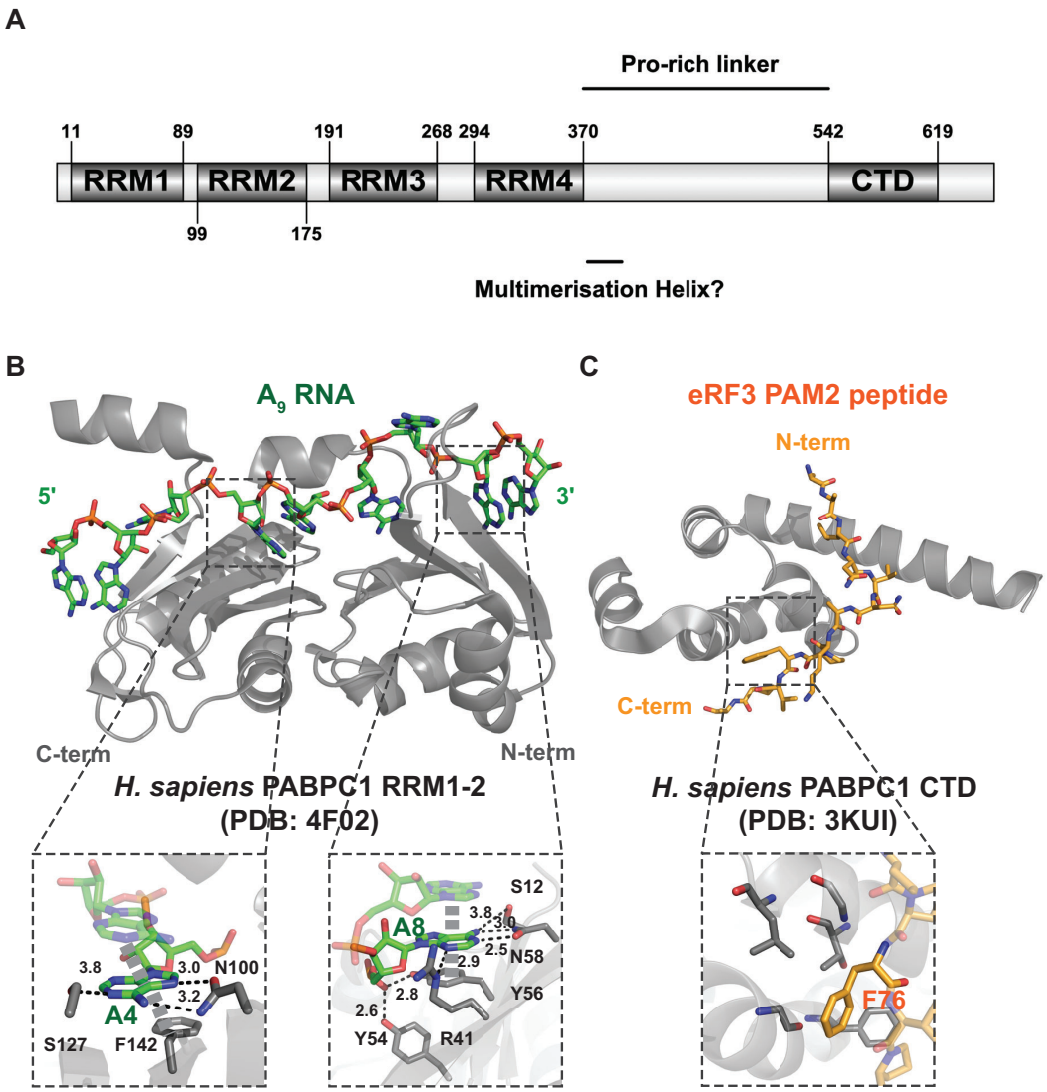


Figure 1.9 Molecular basis of poly(A)-binding protein (Pab1) function. (**A**) Domain schematic of Pab1. Its N-terminus comprises four tandem RRM domains; its C-terminus consists of a proline-rich linker followed by an α-helical C-terminal domain. A helix following RRM4 is thought to mediate multimerisation. (**B**) Molecular basis of RRM1-2 (grey cartoon) binding to poly(A) RNA (green). Insets: base-specific interactions with A_4 and A_8. Putative hydrogen bonds are black lines; stacking interactions are grey dashes. (**C**) Molecular basis of recognition of the eRF3 PAM2 peptide (orange) by the Pab1 CTD (grey cartoon). The peptide adopts an extended conformation. Inset: the key F76 side chain is buried in a hydrophobic pocket in the PABPC1 CTD.

1.4.1.3 Effect of Pab1 on Deadenylation

As Pab1 is thought to coat the poly(A) tail, it is unsurprising that Pab1 affects deadenylation. A temperature-sensitive Pab1 mutant strain has longer global poly(A) tails (Sachs and Davis, 1989b). Similarly, a pab1Δ/spb2Δ (suppressor) yeast strain displayed slower deadenylation and longer poly(A) tails for model transcripts (Caponigro and Parker, 1995). These suggested that Pab1 is required for poly(A) tail shortening. Conversely, Pab1 also stabilises the poly(A) tail by inhibiting deadenylation and preventing mRNA decay. Its depletion accelerated the rate of degradation of model transcripts in cell-free extracts (Bernstein et al., 1989). In *Xenopus* oocytes, Pab1 overexpression inhibited maturation-specific deadenylation (Wormington et al., 1996). Similarly, in yeast extract, exogenous addition of poly(A) (sequestering Pab1) stimulated poly(A) tail shortening, while Pab1 addition inhibited deadenylation (Wilusz et al., 2001).

What are the effects of Pab1 on deadenylases *in vitro*? PARN, which carries out cytoplasmic deadenylation in some contexts (Section 1.2.4), is inhibited by Pab1 (Korner et al., 1998). The effect of Pab1 on Ccr4-Not is more complex. Ccr4 is stimulated by Pab1 and can displace the RRMs, but Caf1 is inhibited by Pab1 (Webster et al., 2018; Yi et al., 2018). Ccr4 can also directly interact with the C-terminal half of Pab1 by pull-down studies, potentially accounting for the stimulatory effect (Webster et al., 2018). However, the molecular basis of this interaction is unknown, and it is unclear how Ccr4 is able to displace the tightly bound Pab1 from poly(A).

In contrast, Pan2-Pan3 had been thought to require Pab1 for its activity (Boeck et al., 1996; Brown et al., 1996). Later studies suggested that the complex was active alone, but was stimulated by Pab1 (Schafer et al., 2014; Wolf et al., 2014). Binding of Pab1 by Pan2-Pan3 was thought to be mediated by an interaction between the PAM2 motif of Pan3 and the Pab1 CTD (Kozlov et al., 2004). A recent structure of yeast Pan2-Pan3-Pab1-poly(A) revealed that distinct domains recognise the Pab1-Pab1 interface on the poly(A) tail (Figure 1.11C) (Schafer et al., 2019). Nonetheless, the straightforward model of Pan2-Pan3 stimulation by Pab1 appears to

be incomplete. In *in vitro* deadenylation assays, Pan2-Pan3 was stalled at 30-A intervals upon incubation with a long poly(A) tail coated with Pab1, suggesting that it would be transiently inhibited upon encountering Pab1. Furthermore, it is unknown how Pan2-Pan3 would displace Pab1 from the poly(A) tail. Thus, the precise effects of Pab1 stoichiometry and individual domains on Pan2-Pan3 remained unresolved before this dissertation.

The *in vitro* effects of Pab1 on deadenylases are corroborated by *in vivo* studies. Firstly, in agreement with recognition of the multimerisation interface by Pan2-Pan3 and the potential role of the Pab1 P-linker in recruiting Ccr4, P-linker deletion showed deadenylation defects *in vivo* (Yao et al., 2007). Similarly, replacement of the P-linker with other sequences or mutation of RRM1 RNA-binding residues affected the initial phase of deadenylation, likely due to reduced Pan2-Pan3 activity (Section 1.5.2) (Zhang et al., 2013). Finally, removal of either the P-linker and/or the CTD from Pab1 slowed RNA decay and deadenylation following transcriptional shutoff *in vivo*, consistent with the requirement of these regions to interact with Ccr4 and Pan3 respectively (Simon and Seraphin, 2007). Nonetheless, these *in vivo* studies are complicated by the highly intricate feedback mechanisms maintaining transcript homeostasis.

Despite the above data, numerous questions remain. Firstly, as poly(A) tails are cytoplasmically shortened to steady-state lengths (Section 1.3.1) and are bound by Pab1, the interplay between Pab1 and deadenylase complexes likely determine, or at least influence, the steady-state poly(A) tail length. The mechanistic basis of this interplay remained unclear. Secondly, Pab1 interacts with numerous components involved in translation (Section 1.4.1.1). A previous study has shown that translation termination may be linked to deadenylation through eRF3 (Funakoshi et al., 2007). However, the effect of transcript circularisation or eIF4G binding to Pab1 on deadenylation was unclear. Finally, the mechanistic basis and molecular requirements for the interaction between Pab1 and deadenylase complexes, such as Pan2-Pan3, remained unresolved prior to this dissertation.

1.4.2 RNA-binding Proteins and Deadenylation

mRNA decay and deadenylation rates account for a large part of variability in gene expression (Jacobson and Peltz, 1999). Variation of $t_{1/2}$ is largely attributable to *cis*- RNA elements (Cheng et al., 2017), many of which are RBP binding sites. Sequence-specific RBPs or RNPs can affect deadenylation by binding RNA motifs in the 3' UTR, concomitantly recruiting enzymes to deadenylate and thus reduce $t_{1/2}$ of target transcripts. Deadenylase recruitment often involves intrinsically disordered regions on RBPs, which may become structured upon binding. Examples of RBP modulation of deadenylation are discussed below.

The AU-rich element (ARE), when introduced into an otherwise stable 3' UTR, caused rapid deadenylation and degradation of the target transcript (Caput et al., 1986; Shaw and Kamen, 1986; Shyu et al., 1991). ARE-binding proteins such as RHAU and KSRP can recruit PARN and the exosome to promote mRNA decay (Gherzi et al., 2004; Tran et al., 2004), but the best characterised ARE-BP is TTP (metazoan: ZFP36) (Brooks and Blackshear, 2013). TTP binds AREs via its zinc finger domain, and recruits Ccr4-Not via a C-terminal low-complexity region which forms an amphipathic helix upon Not 1 binding (Fabian et al., 2013; Sandler et al., 2011).

Micro-RNAs (miRNAs) are well-studied 22-nucleotide RNAs containing a "seed sequence" complementary to the 3'-UTR of target transcripts (Ameres and Zamore, 2013). Functional miRNAs bind auxiliary proteins such as Ago and GW182 (mammals: TNRC6) to form the RNA-induced silencing complex (RISC) (Pratt and MacRae, 2009), which recruits Pan2-Pan3 or Ccr4-Not to repress gene expression (Chen et al., 2009; Eulalio et al., 2009; Wakiyama et al., 2007; Wu et al., 2006). Pan2-Pan3 is recruited through interaction of key tryptophans in GW repeats (and/or M2 and C domains) with Pan3 (Braun et al., 2011; Chekulaeva et al., 2011; Huntzinger et al., 2013; Kuzuoglu-Ozturk et al., 2012). Ccr4-Not binding is likely multipartite

and involves tryptophans binding to CNOT9, and M1, M2, and C domains interacting with Not1 (Braun et al., 2011; Chen et al., 2014; Huntzinger et al., 2013; Mathys et al., 2014).

Pumilio proteins (yeast: Puf, metazoa: PUM) contain a conserved PUM domain which modularly interacts with the Pumilio-response element (PRE) in RNA (Goldstrohm et al., 2018; Wang et al., 2002a). Puf deletions affected poly(A) tail lengths by recruiting Ccr4-Not to rapidly deadenylate target mRNAs (Goldstrohm et al., 2007; Olivas and Parker, 2000). Targeted deadenylation by Puf can be reconstituted *in vitro* with recombinantly purified proteins (Webster et al., 2019). Ccr4-Not appears to be recruited by the Puf N-terminus, but the binding mechanism remains unclear (Arvola et al., 2020; Webster et al., 2019).

YTH domain RBPs preferentially recognise the modified nucleotide *N*6-methyladenosine (m^6A), and numerous structures of the YTH domain bound to m^6A-containing RNAs have been determined (Li et al., 2014; Xu et al., 2015; Xu et al., 2014). YTH domains can also bind non-m^6A RNAs (e.g. DSR motif) in a different mode in organisms where *N*6-methylation is not known to occur (Stowell et al., 2018; Wang et al., 2016; Wu et al., 2017). YTH domain proteins can recognise sequence motifs in the 3'-UTR and recruit deadenylase complexes for targeted deadenylation. In particular, the low-complexity N-terminal region of YTHDF2 or Mmi1 directly interacts with Not1 to recruit Ccr4-Not (Du et al., 2016; Stowell et al., 2016b).

The examples detailed above demonstrate how the RNA-binding activity of RBPs can affect deadenylation and thus the half-life of their cognate transcripts. Generally, this occurs through direct recruitment of deadenylase complexes, such as Ccr4-Not, for targeted deadenylation.

1.4.3 Identification of Putative Binding Partners of Pan2-Pan3

Despite many instances of Ccr4-Not recruitment by sequence-specific RBPs to carry out targeted deadenylation, such examples are rare for Pan2-Pan3; only GW182 has been shown to directly interact with Pan3 (Huntzinger et al., 2013). Nonetheless, Pan3 contains an extensive low-complexity region (Figure 1.3, Supplementary Figure 4) and Pan2-Pan3 contains many available surfaces for interaction (Figure 1.4). Thus, Pan2-Pan3 may interact with RBPs.

No study directly investigating Pan2-Pan3 binding partners has been performed. However, a study of protein-protein interactomes in human cytoplasmic RNA granules has identified possible binding partners of human PAN2-PAN3 (Youn et al., 2018). In this study, proximity labelling was carried out with known P-body or stress granule proteins (Sears et al., 2019), which were tagged with the biotin ligase BirA. Biotinylation of proximal prey proteins enabled them to be pulled down and identified by mass spectrometry. PAN2 and PAN3 were pulled down by multiple bait proteins (Figure 1.10).

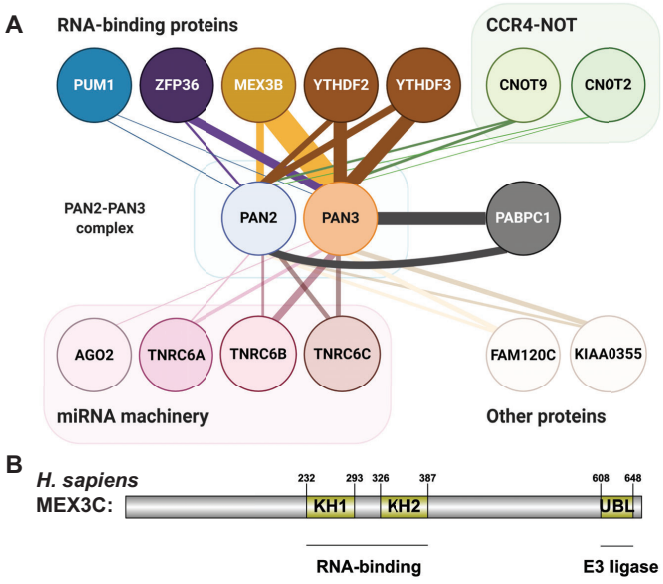


Figure 1.10 (**A**) Putative protein binding partners of PAN2-PAN3. BirA-tagged bait proteins are grouped by function. The width of lines connecting proteins to PAN2 and PAN3 are proportional to the fold enrichment over the negative control, suggesting *bona fide* interactions. (**B**) Domain schematic of *H. sapiens* MEX3C. Domain boundaries are marked by residue number. Horizontal lines show functionally significant regions of the proteins.

From this study, many known interactors of PAN2-PAN3 were identified, including PABPC1 and the RISC proteins AGO2 and TNRC6. Low-enrichment bait proteins included CCR4-NOT and PUM1, suggesting that PAN2-PAN3 may be proximal to other deadenylases and RBPs bound to the 3'-UTR. However, intriguingly, PAN2 and PAN3 were highly enriched for several bait RBPs, including ZFP36, YTHDF2/3, and MEX3B. Furthermore, prey-prey coupling analysis identified MEX3A to be a protein which was pulled-down at similar frequencies as PAN2-PAN3, suggesting that MEX3 may interact with PAN2-PAN3.

The MEX3 proteins are especially interesting as they play a role in numerous developmental and metabolic processes and are implicated in growth factor regulation and cancer (Ariz et al., 2009; Chao et al., 2019; Draper et al., 1996; Huang and Hunter, 2015; Jiang et al., 2012; Jiao et al., 2012a; Jiao et al., 2012b; Le Borgne et al., 2014; Xue et al., 2018). MEX3 is conserved in higher eukaryotes; humans contain four MEX3 paralogues which have divergent linker sequences and may play distinct roles *in vivo* (Supplementary Figure 5) (Buchet-Poyau et al., 2007). MEX3 binds RNA via tandem KH domains with ~100 nM affinity. The sequence specificity and molecular basis of RNA-binding by the KH domains have been determined by SELEX and a crystal structure of the RNA-protein complex (Buchet-Poyau et al., 2007; Draper et al., 1996; Pagano et al., 2009; Yang et al., 2017). The MEX3 response element (MRE) RNA motif is bound in a basic groove on the KH domain surface, with specificity conferred by a variable loop. In agreement with Youn and colleagues, MEX3 localises to RNA granules in a regulated manner (Courchet et al., 2008). Intriguingly, in higher eukaryotes, MEX3 contains an additional C-terminal RING domain, whose E3 ubiquitin ligase activity is important for degradation of the HLA-A2 transcript (Cano et al., 2012; Moududee et al., 2018).

Because the relatively under-characterised MEX3 proteins repressed gene expression and were proximal to PAN2 and PAN3, they were selected for further study.

1.5 Current Model of Pan2-Pan3 Function

1.5.1 Overall Structural Model of Pan2-Pan3

At the start of the work presented in this dissertation, extensive structural studies of different combinations of Pan2-Pan3 domains had been carried out (Christie et al., 2013; Jonas et al., 2014; Schafer et al., 2014; Wolf et al., 2014) (PDB: 4BWK, 4CZY, 4XR7, 4CYJ). These have allowed a structural model of the Pan2-Pan3 core complex to be predicted (Figure 1.4D). Because binding of individual Pan2 domains (WD40, UCH-Exo, linker) did not cause large rearrangements in the structure of the Pan3 C-terminus or the Pan2 domains, Pan2-Pan3 was predicted to act as a rigid complex. Pan2-Pan3 was thought to adopt a tri-lobed conformation, with a 16 nm diameter in its maximum dimension.

During my thesis, a cryo-EM study was published, detailing how yeast Pan2-Pan3 recognises Pab1-poly(A) (Schafer et al., 2019). Pan2-Pan3 adopts a structure similar to the predicted model, except WD40 is more rotated towards the UCH-Exo domain (Figure 1.11A). Upon binding Pab1-poly(A), Pan2-Pan3 does not undergo significant conformational changes. Although half of the Pan3 low-complexity N-terminus was present, no density corresponding to the low-complexity region, or the Pab1 P-linker and CTD, was observed.

A_{90} RNA, which could accommodate three Pab1 molecules, was used in the study. The 3:1 stoichiometry of Pab1:RNA was confirmed by size exclusion chromatography. When bound to Pan2-Pan3, Pab1-poly(A) is extended, contacting a large interface of the Pan2-Pan3 core (Figure 1.11A). Of the three Pab1 molecules, the RRM domains (Section 1.4.1.2) of two were modelled; the third displays only weak density for RRM1 (Figure 1.11B). The lack of available structural models for RRM3-4 (alone or with poly(A)) hampered the modelling of these domains into the overall structure. Given the low local resolution of Pab1-poly(A), poly(A)

RNA was not readily visible and was instead traced as a continuous stretch of density. Poly(A) RNA was modelled with dramatic turns in the absence of protein binding (Figure 1.11B); this is difficult to justify without biochemical data or a higher-resolution structure. Thus, how Pab1 RRMs bind poly(A), both alone and when interacting with Pan2-Pan3, remains unclear.

Pan2-Pan3 contacts the Pab1 multimer at various interfaces. Near the active site, a conserved Pan2 C-terminal α-helix contacts RRM1, and poly(A) appears to exit RRM1 and directly enter the Pan2 active site. The lack of high-resolution information prevented the understanding of how the Pan2 active site specifies for poly(A). Notably, Pan2-Pan3 contacts Pab1-Pab1 oligomerisation interfaces, which consist of a conserved α-helix following RRM4 (Supplementary Figure 6) interacting with the RRM1 helical bundle. These interfaces are recognised by the Pan3 PK and the Pan2 WD40 (opposite the putative protein binding interface) domains (Figure 1.11C). The importance of Pab1 RRM oligomerisation provides a potential explanation of the effects of Pab1 deletion mutants on deadenylation *in vivo*; for example, deletion of the P-rich linker following RRM4 resulted in deadenylation defects (Section 1.4.1.3). Deletion of the Pan2 WD40 domain also results in slower deadenylation, reinforcing its importance in recognising the Pab1-Pab1 interface. Together with biochemical assays, these data suggest that Pan2-Pan3 preferentially acts on longer poly(A) tails by specifically binding Pab1 multimers and is less efficient on poly(A) alone, or poly(A) which is bound by only one Pab1 molecule.

While previous structures of Pan2-Pan3 and the novel structure of Pan2-Pan3 bound to the Pab1-poly(A) RNP served as key studies for understanding Pan2-Pan3 function, several important questions remain. Firstly, the limited resolution prevented elucidation of the molecular basis of interactions between the Pan2-Pan3 core, Pab1, and poly(A). Moreover, the study did not clarify the domain requirements and thus relative importance of different surfaces for the interaction with Pab1. Finally, it is unknown whether the architecture of Pan2-

Pan3-Pab1-poly(A) is conserved across eukaryotes, and whether Pan2-Pan3 from different species recognise the Pab1-poly(A) RNP via similar interfaces. Nonetheless, the high degree of conservation of Pan2-Pan3 and Pab1 implies that the overall architecture is likely conserved.

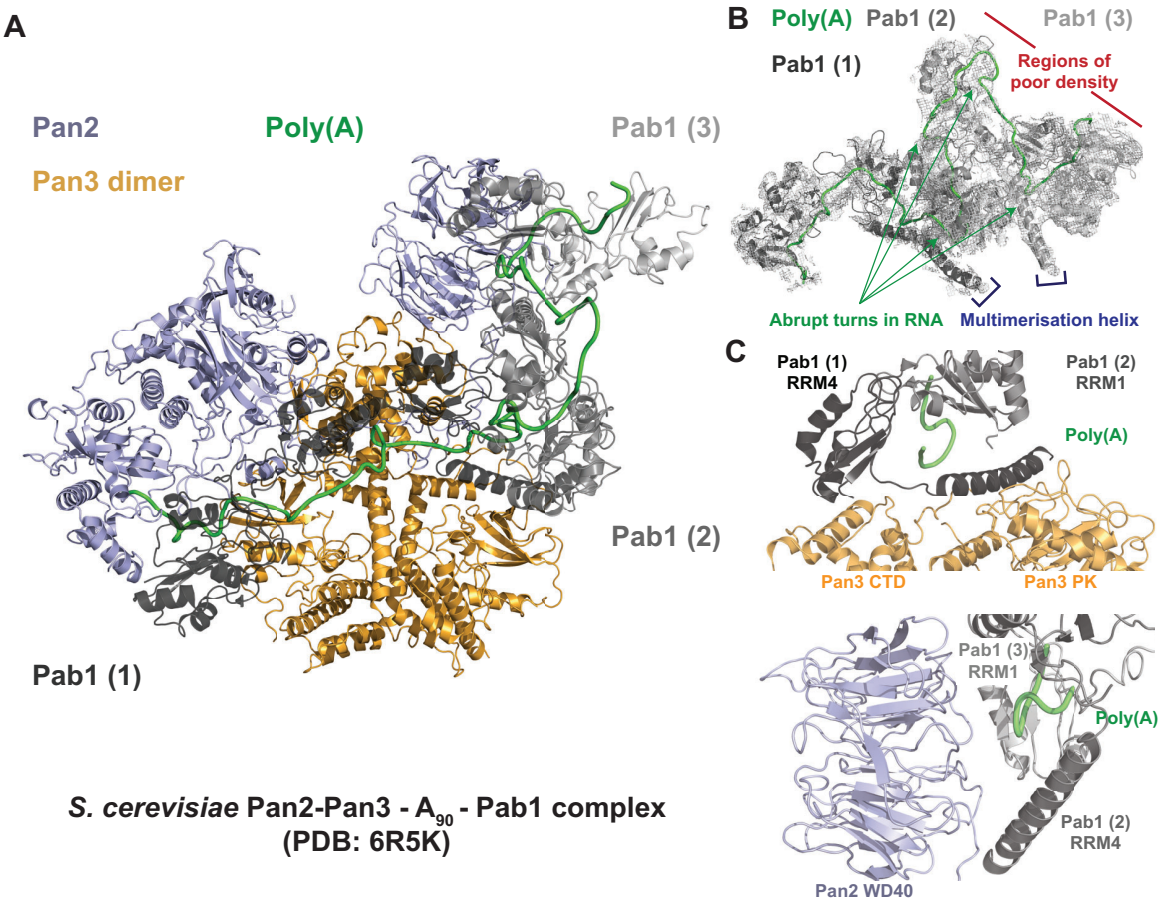


Figure 1.11 Structural model of Pan2-Pan3 in complex with Pab1-poly(A). Pan2 (blue), Pan3 (orange), poly(A) (green), Pab1 (grey) are depicted in cartoon. (**A**) Overall structural model of the complex. Pab1 molecules are shaded from dark to light depending on their proximity to the 3'-end of poly(A). (**B**) Structural model of the Pab1-poly(A) mRNP. The RNA, as modelled, contains dramatic turns in the absence of protein binding (green). The map is poorer towards the 5'-end, possibly due to flexibility or unphased binding of Pab1 (red). A long α-helix mediates Pab1-Pab1 multimerisation between RRM4 and RRM1. (**C**) Pan2-Pan3 recognises Pab1-Pab1 multimerisation interfaces. The 3' interface is recognised by the interaction of Pan3 PKC on the open side of the dimer. The 5' interface is bound by the "bottom" face of the WD40 domain.

1.5.2 Interplay Between Pan2-Pan3 and Ccr4-Not

An outstanding question in deadenylation is the existence of two functionally redundant multiprotein complexes which carry out cytoplasmic deadenylation. Yeast Pan2-Pan3 and Ccr4-Not are constitutively expressed to similar levels (Ho et al., 2018), and both complexes are predominantly cytoplasmic (Huh et al., 2003; Tucker et al., 2002). None of the deadenylase enzymes are, on their own, essential. Deletion of *pan2* or *ccr4* alone only led to a slight increase in global poly(A) tail lengths in yeast (Boeck et al., 1996; Tucker et al., 2001). However, knockdown of CNOT7 (Caf1) in humans (thereby depleting both CNOT7 and CNOT6 from CCR4-NOT) led to a greater increase in poly(A) tail lengths (Yi et al., 2018), and simultaneous deletion of *ccr4* and *pan2* in yeast resulted in almost no observable deadenylation of a model transcript (Tucker et al., 2002; Tucker et al., 2001). Furthermore, the *pan2*Δ/*ccr4*Δ yeast strain displayed growth defects, suggesting reduced cell fitness (Tucker et al., 2002). Therefore, the constitutive deadenylase complexes appear to be functionally redundant, but removal of two deadenylases greatly affects poly(A) tail lengths and cellular fitness.

Although Ccr4-Not and Pan2-Pan3 activities mutually compensate, they may play distinct roles *in vivo* (Yamashita et al., 2005). Pulse-chase studies of a model β-globin reporter transcript suggested that deadenylation is biphasic; the first step is slow and distributive, whereas the second step is more processive and is followed by rapid decay of the transcript body. Overexpression of mutant Pan2 or Ccr4 suggests that Pan2 is responsible for the first phase, whereas Ccr4 acts in the second phase. Thus, this study led to the prevailing view that Pan2-Pan3 carries out deadenylation first, followed by Ccr4-Not. Nonetheless, it is unknown how the deadenylase complexes are targeted to poly(A) tails of different lengths. Moreover, while transcripts undergo poly(A) tail shortening to a steady-state length upon entry into the cytoplasm (Eisen et al., 2020), the involvement of Pan2-Pan3 and/or Ccr4-Not in this process and how the steady-state poly(A) tail length is related to gene expression, were unknown.

Historically, the stimulatory and inhibitory effects of Pab1 on Pan2-Pan3 and on recombinantly purified Ccr4 respectively suggested that Pab1 could account for biphasic deadenylation (Kozlov et al., 2004; Tucker et al., 2002). In this model, poly(A) tail removal is mainly carried out by Pan2-Pan3 when recruited by Pab1, and once Pab1 dissociates due to poly(A) shortening, the remaining free poly(A) is removed by Ccr4-Not. New results, however, complicate this model. Firstly, Pan2-Pan3 may require multiple Pab1 molecules for efficient deadenylation as it recognises the interface between adjacent Pab1 RRMs (Schafer et al., 2019). Moreover, Ccr4-Not is stimulated by poly(A)-binding protein *in vitro* (Webster et al., 2019; Yi et al., 2018). The exact mechanistic effects of Pab1 on either complex and how these are consistent with the biphasic model of deadenylation remain unclear.

Finally, Pan2-Pan3 and Ccr4-Not may be differentially localised and recruited to transcripts. Pan2-Pan3 is a constitutive component of P-bodies (Section 1.2.2). While Ccr4-Not is also found in P-bodies, it is more distributed in the cytoplasm (Zheng et al., 2008). As transcripts are differentially sorted to P-bodies, particularly under cell stress, Pan2-Pan3 may be involved in the deadenylation of P-body-localised mRNAs (Aizer et al., 2014; Hubstenberger et al., 2017). Conversely, Ccr4-Not directly interacts with sequence-specific RBPs such as YTHDF2 (Du et al., 2016), Mmi1 (Stowell et al., 2016a), Zfs1, TTP, and Puf3 (Webster et al., 2019). These recruit Ccr4-Not to their target transcripts for rapid deadenylation, suggesting that Ccr4-Not is more important in directed mRNA turnover. On the other hand, it was unknown whether Pan2-Pan3 displays analogous direct interactions with RBPs to target it to different (or overlapping) subsets of transcripts. In addition, even in the absence of RBPs, there is evidence that Pan2-Pan3 and Ccr4-Not differentially deadenylate mRNAs based on the sequence of the 3' UTR (Lowell et al., 1992; Webster et al., 2019). Thus, Pan2-Pan3 and Ccr4-Not may have different sequence- and RBP-dependent transcript specificities.

1.6 Aims

Biochemical and structural studies characterising Pan2-Pan3 and its regulation by Pab1 had been published prior to my dissertation. Biochemical experiments have shown that Pan2-Pan3 and Ccr4-Not display poly(A) specificity. However, it was largely unknown how individual deadenylases, particularly Pan2 and Caf1, recognise the poly(A) tail. Understanding the molecular basis of poly(A) specificity would provide insight into whether deadenylases continue into the 3' UTR and how the nucleotide composition of the poly(A) tail affects deadenylation. In addition, while the structures of yeast Pan2-Pan3 and Pan2-Pan3-Pab1-poly(A) have been elucidated, the involvement of domains in deadenylation and in the interaction between Pan2-Pan3 and Pab1-poly(A) have not been thoroughly investigated. Furthermore, yeast and human mRNAs are synthesised with different initial poly(A) tail lengths and reach different steady-state poly(A) tail lengths (Section 1.3.1). Studies on human PAN2-PAN3 and how it contacts PABPC1-poly(A) may elucidate species-specific differences. Finally, no sequence-specific RBPs were reported to affect Pan2-Pan3 deadenylation. However, the recent study by Youn and colleagues provided the rationale to test RBPs such as MEX3, for their effect on *H. sapiens* PAN2-PAN3. While MEX3 RNA binding specificity and affinity has been studied, a direct interaction between PAN2-PAN3 and MEX3 and the effect of MEX3 on PAN2-PAN3 activity remained to be investigated. The aims of my thesis were:

1. Elucidation of the molecular basis of nucleotide specificity of the DEDD-superfamily deadenylases Pan2 and Caf1.
2. Expression and purification of recombinant *H. sapiens* PAN2-PAN3 and PABPC1. Structural and *in vitro* biochemical characterisation of yeast and human Pan2-Pan3 alone and in complex with poly(A)-binding protein.
3. Expression and purification of recombinant *H. sapiens* MEX3. *In vitro* characterisation and reconstitution of the effect of MEX3 on PAN2-PAN3 activity.

2. Poly(A) Specificity of DEDD Deadenylases

Declaration

The work presented in Chapter 2 was previously published:

1. **Tang, T.T.L.**, Stowell, J.A.W., Hill, C.H., and Passmore, L.A. (2019). The intrinsic structure of poly(A) RNA determines the specificity of Pan2 and Caf1 deadenylases. Nat Struct Mol Biol *26*, 433-442.
2. **Tang, T.T.L.**, and Passmore, L.A. (2020). Recognition of Poly(A) RNA through Its Intrinsic Helical Structure. Cold Spring Harb Symp Quant Biol. *84*, 21-30.

J.A.W.S., L.A.P., and I conceived the study. Ccr4-Not constructs were cloned and purified by J.A.W.S. and Michael Webster. Crystallisation attempts of Caf1 with RNA were carried out by J.A.W.S. Crystallisation, X-ray crystallography data collection, and data analysis were carried out with the help of J.A.W.S. and C.H.H. I performed all other experiments in this chapter.

2.1 Pan2 Has Intrinsic Nucleotide Specificity

Previous studies had shown that Pan2-Pan3 is preferentially active on A_{15} relative to other polyribonucleotides, but can act on short stretches of non-poly(A) RNA (Lowell et al., 1992; Schafer et al., 2014). The preference for poly(A) was attributed to Pan3, which directly binds poly(A) through its N-terminal zinc finger (Wolf et al., 2014), and binds to the Pab1 CTD through its PAM2 motif, recruiting Pan2-Pan3 to the poly(A) tail (Siddiqui et al., 2007). However, it was unknown if the exonuclease alone shows intrinsic specificity for adenosines.

2.1.1 Expression and Purification of Pan2-Pan3 Constructs

Precise determination of poly(A) specificity of Pan2-Pan3 required the *in vitro* reconstitution of purified protein and RNA components. Thus, human and *S. cerevisiae* Pan2-Pan3, as well as the yeast catalytic module (Pan2UCH-Exo) (Figure 1.3), were expressed and purified. Yeast Pan2-Pan3 was expressed and purified according to methods developed by Jana Wolf (Wolf et al., 2014). Briefly, a plasmid containing Pan2 and Strep II (SII)-tagged Pan3 was transformed into yeast. Pan2-Pan3 overexpression was induced by galactose addition. SII-tagged Pan2-Pan3 was purified by affinity purification, cation exchange chromatography, and size exclusion chromatography (Figure 2.1A). Purification of Pan2UCH-Exo was carried out according to a published protocol with modifications (Schafer et al., 2014). His-tagged Pan2UCH-Exo was overexpressed in *Sf9* insect cells and purified by IMAC, anion exchange chromatography, and size exclusion chromatography (Figure 2.1B). Finally, I developed a protocol to express and purify *H. sapiens* PAN2-PAN3ΔN278, which was more stably expressed than the complex with full-length PAN3. This truncation mutant retains PAM2, which is required to interact with the PABPC1 CTD. Individual proteins were cloned into a pBIG vector, which was used to generate a bacmid for baculovirus-mediated overexpression in *Sf9* insect cells (Section 5.4.5).

SII-tagged complex was purified by affinity purification, anion exchange chromatography, and size exclusion chromatography (Figure 2.1C).

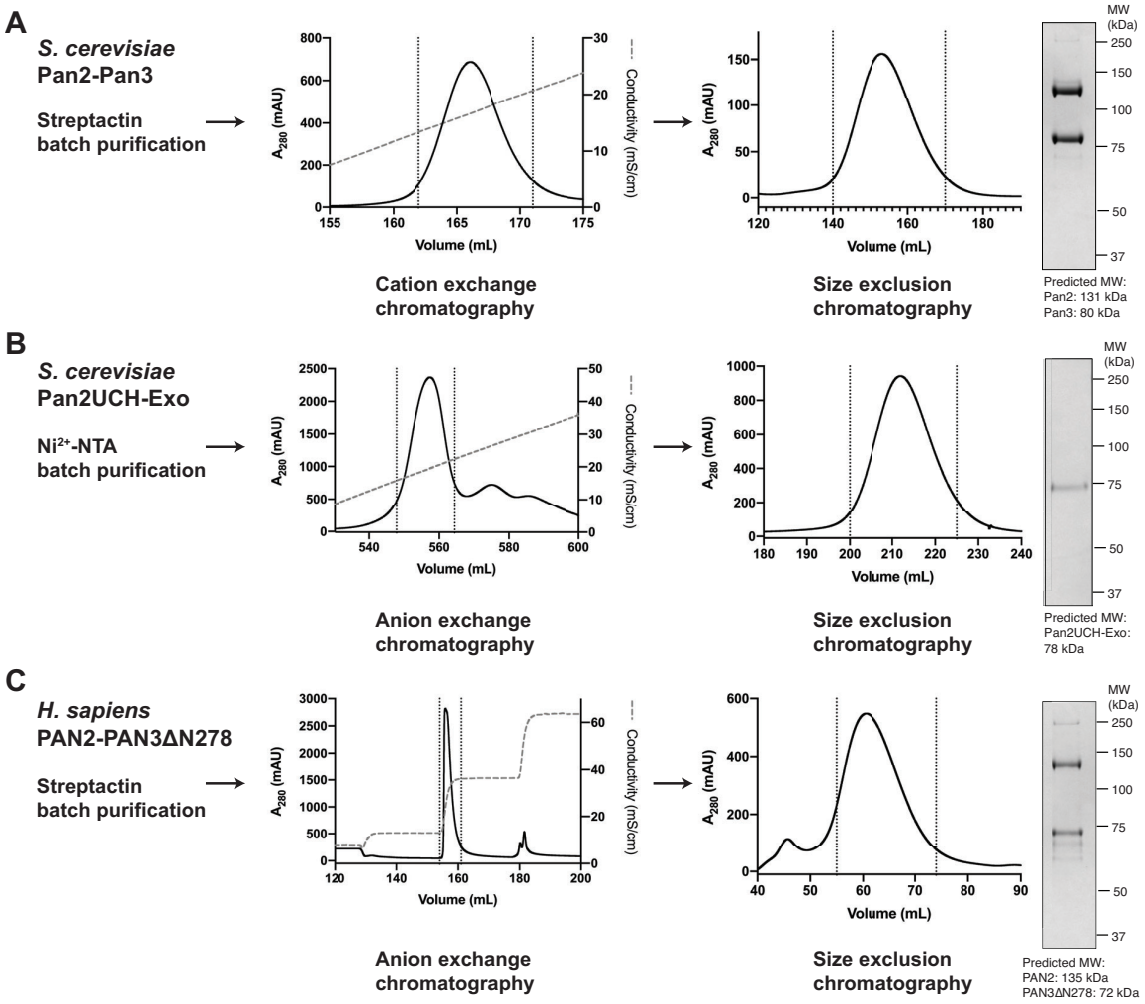


Figure 2.1 Purification of *S. cerevisiae* Pan2-Pan3. Chromatograms from the purification of: (**A**) *S. cerevisiae* Pan2-Pan3; (**B**) *S. cerevisiae* Pan2UCH-Exo; and (**C**) *H. sapiens* PAN2-PAN3ΔN278 and SDS polyacrylamide gels of the purified complexes. Pooled fractions corresponding to the protein/protein complex of interest are outlined by the dashed lines.

2.1.2 Nucleotide Specificity of Pan2-Pan3

Pan2-Pan3 nucleotide specificity was tested by *in vitro* deadenylation assays on 5' 6-FAM-labelled RNAs. These contain an upstream 20-mer non-A sequence (X_{20}), an A_{30} tail, and different 3' termini (-C_3, -U_3, -G_3). The upstream 20-mer sequence was designed to mimic a 3' UTR and is predicted to be unstructured (Jonstrup et al., 2007). In these assays, deadenylation of the substrate RNA shortens it from the 3'-end, resulting in increased electrophoretic mobility when analysed by denaturing PAGE. To quantify deadenylation, the band corresponding to intact RNA (lowest mobility) was boxed and its intensity was measured by densitometry. This reflects Pan2-Pan3 activity, as inhibition of deadenylation would result in a high intensity of intact RNA over time (Figure 2.2A).

Relative to an RNA with 3' poly(A), 3' guanosines, but not cytosines or uracils, strongly inhibited deadenylation by *S. cerevisiae* Pan2-Pan3 (Figure 2.2B) and *H. sapiens* PAN2-PAN3 (Figure 2.2C). Thus, Pan2-Pan3 nucleotide specificity is conserved among different species. To assess if Pan3 is required for specificity, assays were repeated with Pan2UCH-Exo, which only contain the contiguous catalytic unit. UCH was included because a loop extending from the UCH domain may contact RNA (Schafer et al., 2014). Pan2UCH-Exo had the same specificity as Pan2-Pan3, suggesting that the exonuclease is intrinsically specific for 3' nucleotides (Figure 2.2D). In line with the lack of strict specificity and the designed X_{20} sequence, Pan2-Pan3 also readily degraded X_{20}, as shown by the disappearance of the band corresponding to fully deadenylated RNA substrate at longer time points (Figure 2.2A). Due to the relatively low specificity of Pan2-Pan3 deadenylation, the fluorescently-labelled substrate was not easily tracked, leading to the potential ambiguity that non-specific RNases may be responsible for RNA degradation. Future experiments may be improved by observation of all 3'-5' degradation intermediates until the fluorescent label, or by inhibition of deadenylation once the X_{20} sequence is reached, for example by the introduction of a phosphorothioate bond.

As guanosines strongly inhibited Pan2-Pan3, I investigated the relative importance of the three terminal positions on deadenylation (Figure 2.2E). Exonucleolytic cleavage occurs between N_{+1} (3') and N_{-1} (5') residues. Of the three terminal nucleotides, a strong inhibitory effect was observed when both N_{-1} and N_{-2} were G's (-**GG**G, -**GG**A) compared to when only one was G (-**G**AA, -A**G**A. -A**G**G). Furthermore, the N_{+1} nucleotide was less important than upstream nucleotides in inhibition: for example, RNA terminating in -GG**A** was deadenylated more slowly than that ending in -AG**G** despite both containing two guanosines at the 3' end. These results demonstrate that Pan2 specifies against guanosines in a position-dependent manner.

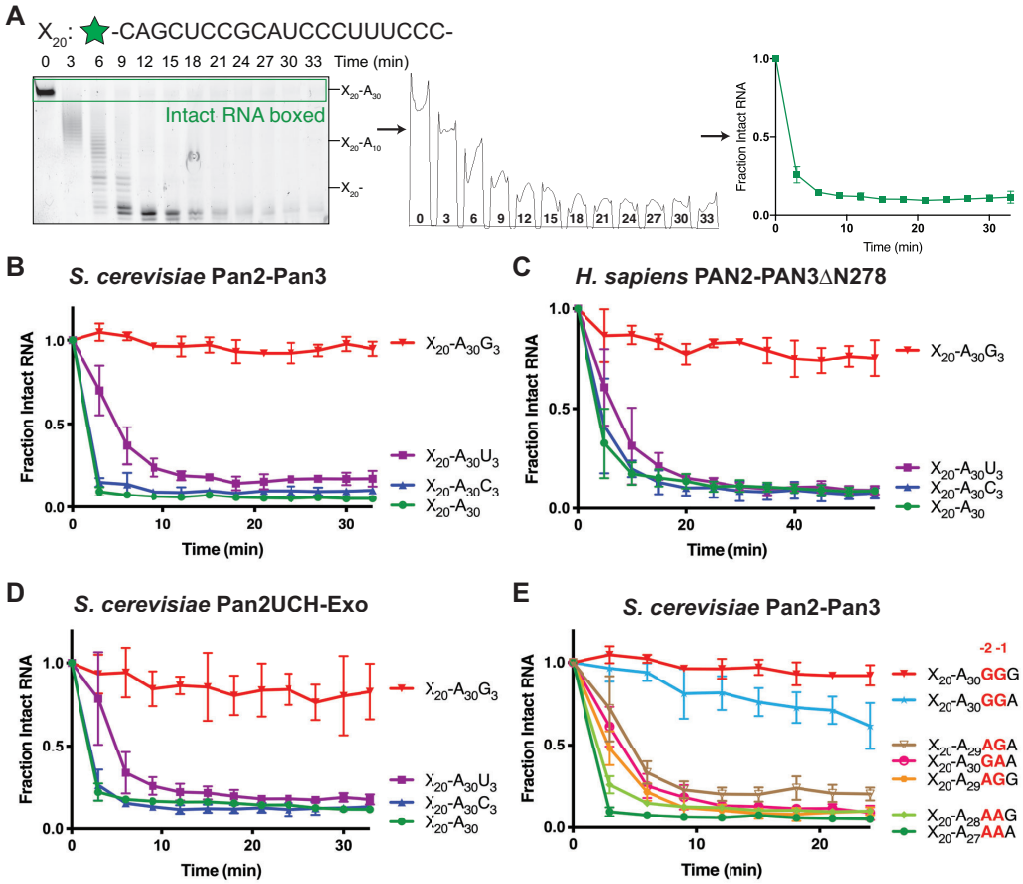


Figure 2.2 Pan2 exhibits position-dependent 3' nucleotide specificity. Reactions were performed in triplicate, data points are the mean, and error bars are standard deviation. (**A**) Analysis of deadenylation assays by densitometry of intact RNA. Intensities are normalised to that at time = 0 and plotted against time. Straight lines connect each point for clarity. Analysis of deadenylation by (**B**) 50 nM *S. cerevisiae* Pan2-Pan3 on 200 nM RNA; (**C**) 500 nM *H. sapiens* PAN2-PAN3ΔN278 on 500 nM RNA; and (**D**) 100 nM *S. cerevisiae* Pan2UCH-Exo on 200 nM RNA. Analysis of deadenylation on 200 nM RNA substrates with guanosines at different positions by (**E**) 50 nM *S. cerevisiae* Pan2-Pan3.

A recent study showed that TENT4A and TENT4B can intermittently incorporate single guanosines into poly(A) tails in human cells (up to 5% of certain transcript classes) (Lim et al., 2018). To test the effect of a single interrupting guanosine on Pan2 activity, I compared deadenylation of two differentially labelled RNAs: one contained an A_{30} tail (5' Alexa647-labelled, false-coloured red), whereas the other contained the same poly(A) tail interrupted by a single guanosine (5' 6-FAM-labelled, false-coloured blue). Strikingly, this assay showed that Pan2 was inhibited upon encountering a -GAAA sequence, compared to uninhibited, distributive activity on the non-guanylated substrate (Figure 2.3). These results suggest that, within a pool of RNAs containing guanylated and pure poly(A) tails, physiologically relevant single guanosine incorporation into the poly(A) tail can inhibit Pan2.

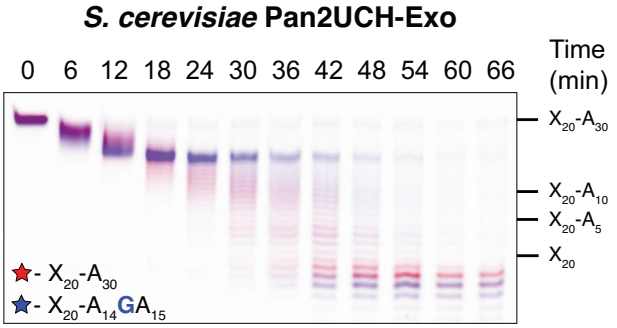


Figure 2.3 Single guanosines inhibit Pan2 deadenylation. Two-colour deadenylation assay on 100 nM of each RNA with different 5' labels (Alexa-647-X_{20}-A_{30}; 6-FAM-X_{20}-$A_{14}GA_{15}$) by 50 nM *S. cerevisiae* Pan2UCH-Exo shows that Pan2 preferentially deadenylates an RNA with a poly(A) tail compared to an RNA containing an interrupting guanosine. RNAs are represented with false colouring.

2.2 Intrinsic Nucleotide Specificity of Ccr4 and Caf1

Previous studies had shown that the Ccr4-Not deadenylases Ccr4 and Caf1 display poly(A) specificity alone and cannot degrade sequences such as -C_{10} (Bianchin et al., 2005; Chen et al., 2002). However, deadenylase specificity has not been systematically studied in the full Ccr4-Not complex. To determine Ccr4 and Caf1 specificity, I carried out deadenylation assays with recombinant *S. pombe* Ccr4-Not on RNA substrates with different 3' termini (-C_3, -U_3, -G_3) as above. Wild-type Ccr4-Not was slightly affected by uracils, moderately inhibited by cytosines, and strongly inhibited by guanosines (Figure 2.4A). Thus, Ccr4-Not has different nucleotide specificity compared to Pan2-Pan3, which is only strongly inhibited by guanosines (Figure 2.2).

To test the specificity of individual nucleases, I repeated deadenylation assays with recombinant Ccr4-Not containing catalytic point mutations in either Caf1 or Ccr4. Ccr4-inactive Ccr4-Not (only Caf1 active) displayed strict poly(A) stringency and was strongly inhibited by substrates with terminal non-A nucleotides (Figure 2.4B). Conversely, Caf1-inactive Ccr4-Not (only Ccr4 active) was strongly inhibited by terminal guanosines and cytosines but not by uracils (Figure 2.4C). Thus, Ccr4 likely accounts for the ability of the wild-type complex to remove terminal uracils (Figure 2.4A). To confirm that observed specificities are not due to other Ccr4-Not components, assays were repeated with recombinant *S. pombe* Caf1. Caf1 had the same specificity as Ccr4-inactive Ccr4-Not (Figure 2.4D), suggesting that Caf1 specificity is intrinsic to the deadenylase. *S. pombe* Ccr4 did not display sufficient activity to assess its specificity. Thus, Ccr4 and Caf1 have distinct intrinsic specificities.

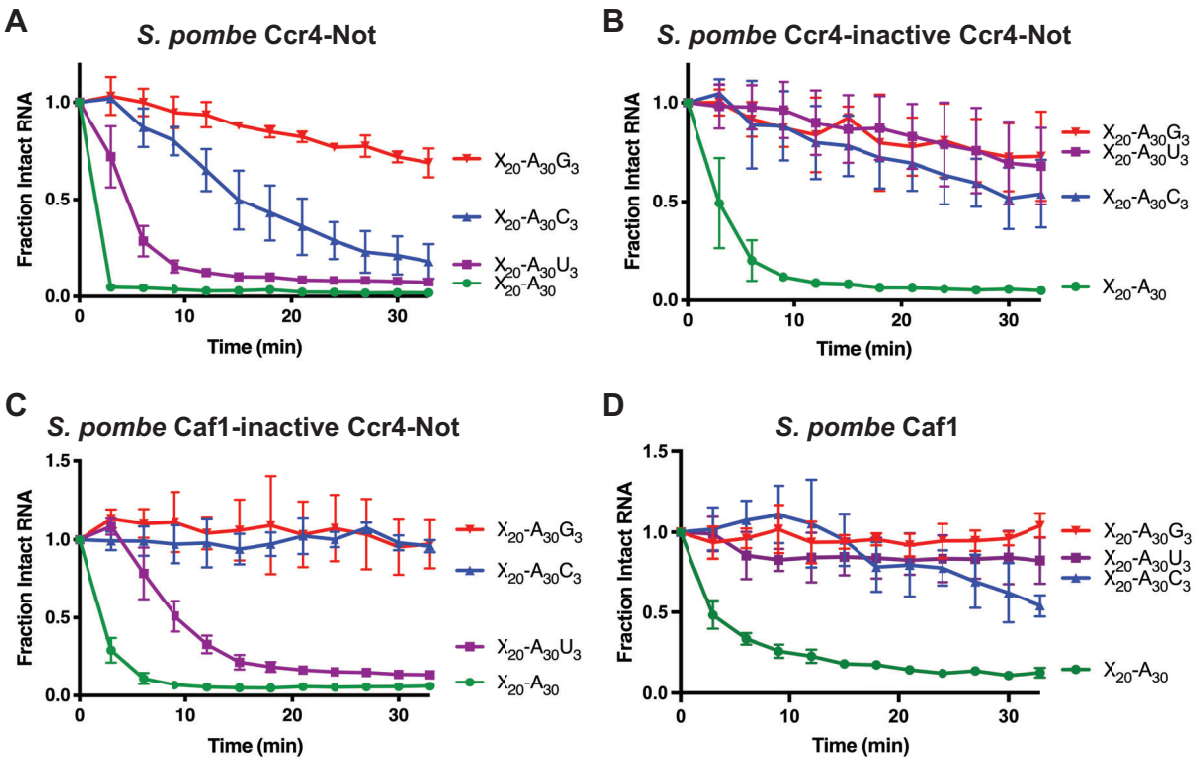


Figure 2.4 Ccr4-Not deadenylases Ccr4 and Caf1 display distinct nucleotide specificities. Reactions were performed in triplicate, data points are the mean, and error bars are standard deviation. Analysis of deadenylation on 200 nM RNA substrates with different 3' termini by (**A**) 50 nM *S. pombe* wild-type Ccr4-Not; (**B**) 50 nM *S. pombe* Ccr4-inactive Ccr4-Not; (**C**) 50 nM *S. pombe* Caf1-inactive Ccr4-Not; and (**D**) 2.5 μM *S. pombe* Caf1.

2.3 Guanosine Prevalence in the 3′ UTR

While other non-A nucleotides such as uracils have a relatively small effect on deadenylation, tandem guanosines strongly inhibit both Pan2-Pan3 and Ccr4-Not. Accordingly, given the X_{20} sequence used in deadenylation assays, which is relatively deficient in guanosines, Pan2-Pan3 was able to degrade the UTR-like sequence (Figure 2.2A). Given the strong inhibitory effect of guanosines, I hypothesised that they may be enriched at the 3' UTR proximal to the poly(A) tail in order to prevent deadenylases from degrading the 3' UTR.

I sought to determine if tandem G's (-GG- or -GGG-) are enriched near the 3' termini of the 3' UTR above their expected occurrence if nucleotide composition were random. Poly(A) sites of human transcripts were obtained from the comprehensive poly(A) database from the Tian laboratory (Wang et al., 2018c). The 20 nucleotides upstream of the predicted poly(A) site were isolated and the prevalence of tandem G's in this sequence was calculated. This revealed that -GG- occurs in 41.4% of 3' UTRs, whereas -GGG- occurs in 9.6% of 3' UTRs. These frequencies are substantially lower than the predicted 70.7% ($1 - \left(\frac{15}{16}\right)^{19}$) and 24.7% ($1 - \left(\frac{63}{64}\right)^{18}$) within a random 20-nucleotide sequence respectively.

This result suggests that tandem G's are globally less prevalent than expected in the 3' UTR proximal to the poly(A) site, and thus are not used as a mechanism to prevent deadenylases from degrading the 3' UTR in human cells. However, it cannot be ruled out that this is used by certain transcripts or transcript classes for regulating degradation of the 3' UTR by deadenylase complexes. It is also plausible that different transcript isoforms may include tandem guanosines in order to enable differential stability. Finally, it is possible that different species may have different frequencies of tandem G's to inhibit degradation of the 3' UTR.

2.4 Structure of the Pan2-Poly(A) Complex

Previous structures of the Pan2 catalytic unit had shown that the exonuclease adopts a canonical DEDD-family fold, and conserved acidic residues coordinate M^{2+} ions for catalysis (Jonas et al., 2014; Schafer et al., 2014). However, as these structures lacked RNA, the mechanism of Pan2 specificity was unknown. To elucidate the molecular basis of Pan2 specificity, a catalytically inactive mutant (E912A) of *S. cerevisiae* Pan2UCH-Exo was crystallised (Figure 2.5A-B), which diffracted to 3.0 Å (Table 2.1). In this crystal form, the Pan2 active site faced large solvent channels and was not blocked by crystal packing (Figure 2.5C). Superposition of apo Pan2UCH-Exo with the previous structure of the yeast Pan2 catalytic unit (PDB: 4Q8H) showed a virtually identical structure, with a backbone RMSD of 0.51 Å (Figure 2.5D).

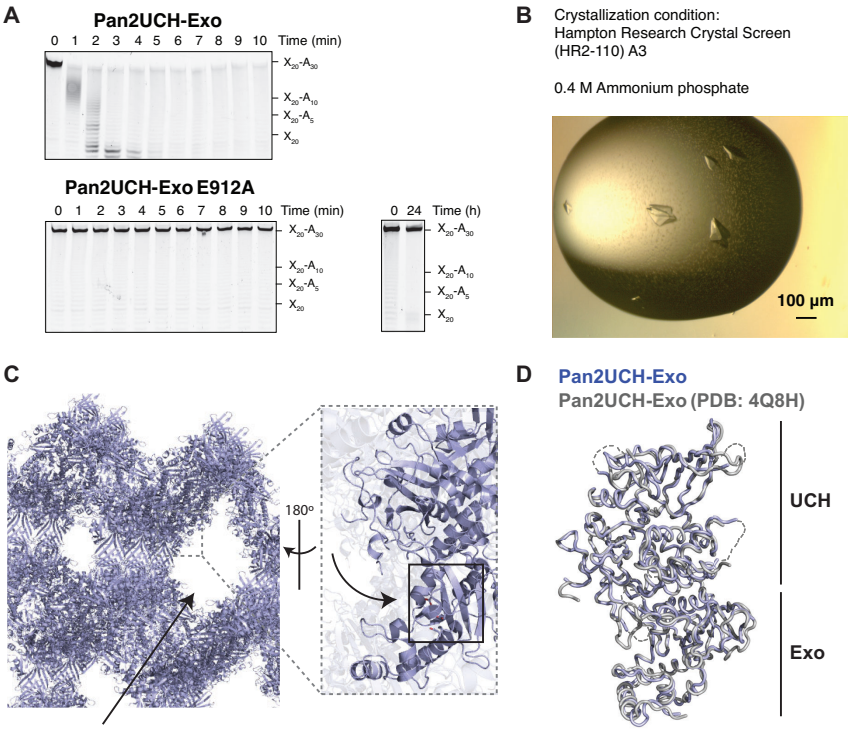


Figure 2.5 Crystallisation of *S. cerevisiae* Pan2UCH-Exo E912A (blue). (**A**) Deadenylation of 200 nM $X_{20}A_{30}$ RNA by 100 nM Pan2UCH-Exo or the catalytic mutant Pan2UCH-Exo E912A, demonstrating that the mutant is inactive *in vitro*. (**B**) Pan2UCH-Exo E912A crystals. (**C**) Crystal packing of Pan2UCH-Exo E912A show large solvent channels (black arrow) and the accessible Pan2 active site (black box). (**D**) Superposition of apo Pan2UCH-Exo E912A (blue) with the previously determined structure of apo Pan2UCH-Exo (grey, PDB: 4Q8H).

2.4.1 Overall Structure of Pan2UCH-Exo-A_7

As the active site was accessible to solvent, Pan2UCH-Exo crystals were soaked in the crystallisation buffer containing A_7 RNA, using times and concentrations which were optimised through crystal morphology (dissolution and other defects) and maximum resolution. These crystals diffracted to 3.3 Å (Figure 2.6A, Table 2.1). No conformational changes could be observed in Pan2UCH-Exo compared to the apo protein (Schafer et al., 2014), suggesting that the catalytic unit is rigid and does not undergo RNA-dependent rearrangements. Electron density for five nucleotides were observed in the active site cleft (Figure 2.6A, inset). Soaking longer oligo(A)s (up to A_{10}) did not result in the observation of density for more nucleotides, suggesting that oligo(A) is only locally ordered in the Pan2 active site. The 3'-most scissile phosphate of RNA, between A_{-1} and A_{+1}, faced the catalytic DEDD amino acid side chains, but no catalytic metal ions could be seen, consistent with soak buffer conditions and the E912A mutation (Figure 2.6B). Moreover, the achieved resolution was insufficient to observe any water molecules in the active site, which are predicted to be required for catalysis. Further soaks with Pan2UCH-Exo in a buffer containing RNA and Ca^{2+} (which inhibited Pan2 in *in vitro* deadenylation assays) did not yield density for metal ions. In this case, it is possible that the larger ionic radius of Ca^{2+} prevents its binding to the Pan2 active site. Superposition with a Caf1 structure with two divalent metal ions (PDB: 3G0Z) showed that RNA binding is compatible with the predicted position of two ions in the active site (Figure 2.6C). The observed A_7 is thus likely bound in a catalytically competent conformation.

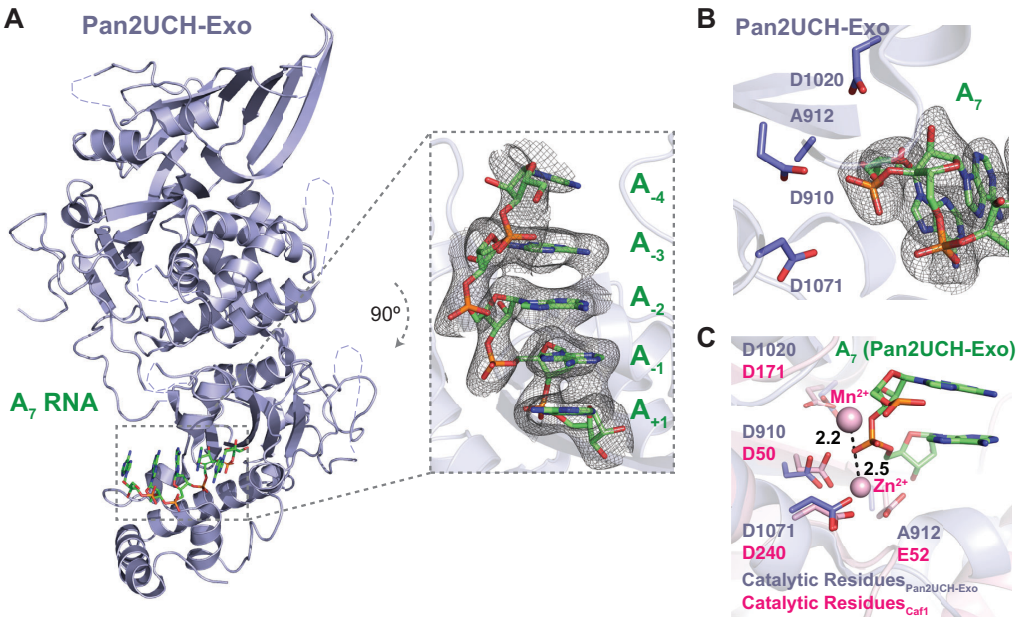


Figure 2.6 Overall structure of Pan2UCH-Exo-A_7. (**A**) A_7 RNA (green sticks) is bound in the Pan2 (blue cartoon) active site. Inset: close-up view of the A_7 RNA, numbered relative to scissile phosphate. The mesh is a feature-enhanced map contoured to 1.3σ. (**B**) Catalytic acidic side chains and the scissile phosphate. The mesh is a feature-enhanced map contoured to 1.8σ. (**C**) Superposition of Pan2UCH-Exo (blue) with *S. pombe* Caf1 (PDB: 3G0Z, pink). Black numbers are distance in Å. Metal ions are pink spheres and are positioned for productive catalysis.

2.4.2 Pan2 Contacts the Ribophosphate Backbone of Oligo(A)

Interactions between Pan2UCH-Exo and A_7 were elucidated by the structure. The terminal A_{+1} was π-stacked against the Y975 aromatic ring in the active site (Figure 2.7A). Further 3' nucleotides could not be sterically accommodated, consistent with Pan2 exonucleolytic activity. RNA 2'-hydroxyl, 3'-hydroxyl, and phosphate groups formed hydrogen bonds with Pan2 main chain (F913, Y1046, and L1049) or side chain (N1019 and S0148) atoms (Figure 2.7A).

Strikingly, Pan2 interacted with oligo(A) mainly via the ribophosphate backbone; adenines pointed into solvent and were not within interaction distances with Pan2UCH-Exo (Figure 2.7B). The lack of base-specific interactions was surprising as Pan2 could discriminate between adenines and guanines (Figure 2.2D). Upon closer examination, oligo(A) formed a single-stranded A-form-like helix in the Pan2 active site. In this conformation, adenine bases are π-stacked in an offset parallel manner, and riboses adopt C3'-endo sugar pucker. As there were no base-specific contacts, these results suggest that Pan2 may recognise oligo(A) conformation.

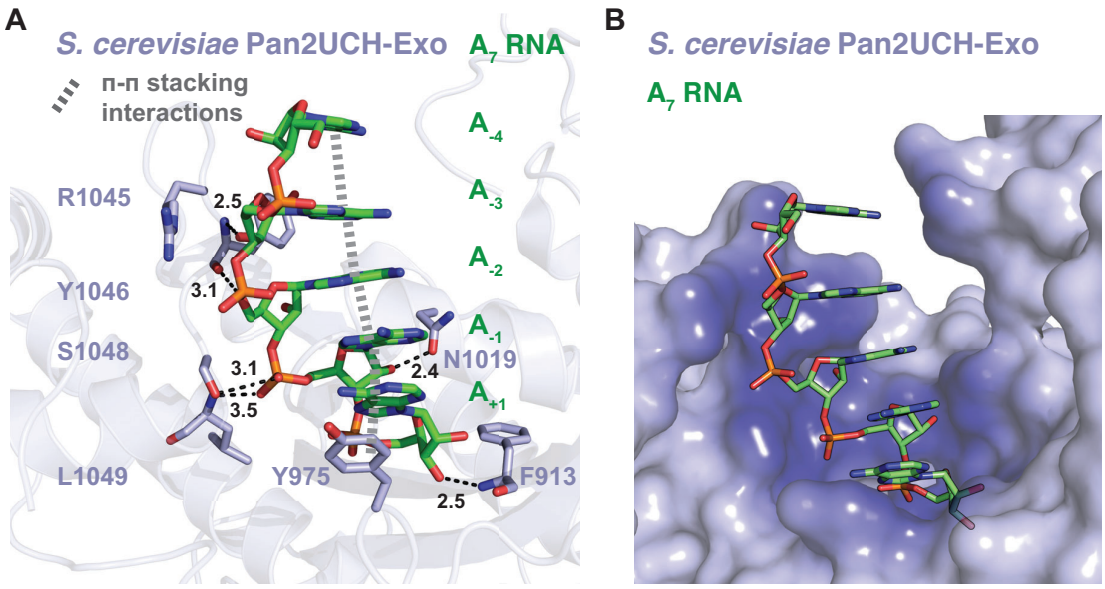


Figure 2.7 Interactions between Pan2UCH-Exo (blue) and oligo(A) RNA (green). (**A**) Summary of interactions between Pan2UCH-Exo and A_7. Putative hydrogen bonds are in black dashes; interatomic distances are in Å. π-stacking interactions are grey dashes. (**B**) Surface of Pan2UCH-Exo bound to A_7. The surface is coloured by RNA proximity from dark (<3 Å) to light (>7 Å).

2.5 Guanosines Disrupt Poly(A) Structure

2.5.1 Structures of Pan2UCH-Exo-AAGGA and -AAGGAA

Pan2 discriminated between A's and G's *in vitro* (Section 2.1.2), but did not form adenine-specific contacts (Section 2.4.2). To elucidate how guanosines inhibit Pan2, AAGGA and AAGGAA RNAs were soaked into Pan2UCH-Exo crystals, which diffracted to 3.3 Å (Table 2.1). These RNAs were chosen as tandem G's were sufficient to strongly inhibit Pan2 (Figure 2.2). Electron density for four residues was observed (Figure 2.8A, C). Both RNAs formed similar interactions with Pan2UCH-Exo compared to A_7 (Figure 2.8B, D), suggesting that Pan2UCH-Exo affinity for G-containing RNAs may be similar to oligo(A).

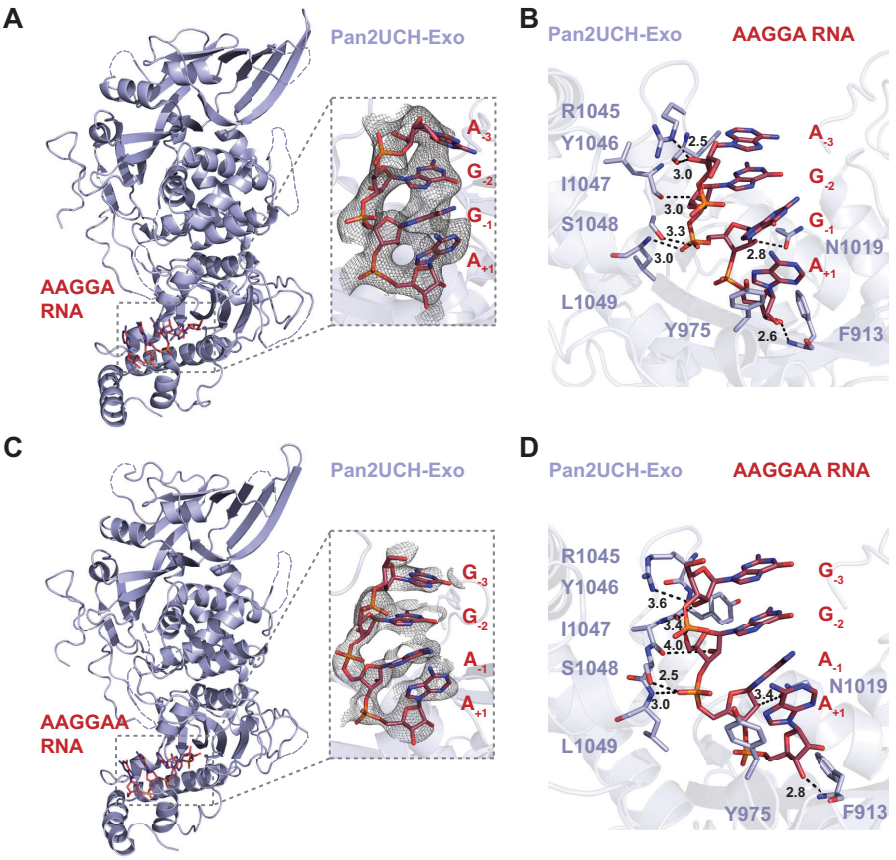


Figure 2.8 Structures of Pan2UCH-Exo bound to G-containing RNAs. Overall structures of Pan2UCH-Exo (blue cartoon) bound to (**A**) AAGGA or (**C**) AAGGAA RNAs (red sticks). Meshes are feature-enhanced maps contoured to 1.8σ. Summary of interactions between Pan2UCH-Exo and (**B**) AAGGA and (**D**) AAGGAA. Putative hydrogen bonds are in black dashed lines; interatomic distances are in Å.

Despite forming similar interactions with Pan2UCH-Exo compared to A_7, AAGGA and AAGGAA adopted dramatically different conformations. In contrast with oligo(A), inter-base parallel π-stacking was disrupted for AAGGA and AAGGAA, such that the bases were at an angle relative to each other (Figure 2.9A-B). Furthermore, the helical conformation adopted by the ribophosphate backbone of G-containing RNAs was distorted compared to that of oligo(A) (Figure 2.9A-B). This may lead to a positional change of the scissile phosphate relative to the coordinated metal ions, leading to inhibition of cleavage. Nonetheless, the resolution of Pan2-RNA complexes was insufficient to elucidate this positional change if it indeed occurs. In agreement with the formation of a less stable conformation, B-factors for G-containing RNAs are higher than that for oligo(A) (Table 2.1). This could reflect either lower occupancy or increased flexibility of G-containing RNAs in the Pan2 active site. Thus, these structures suggest that G-containing RNAs form a distorted helical structure with disrupted base-stacking in the Pan2 active site, in contrast with the stacked, helical conformation formed by oligo(A) RNAs.

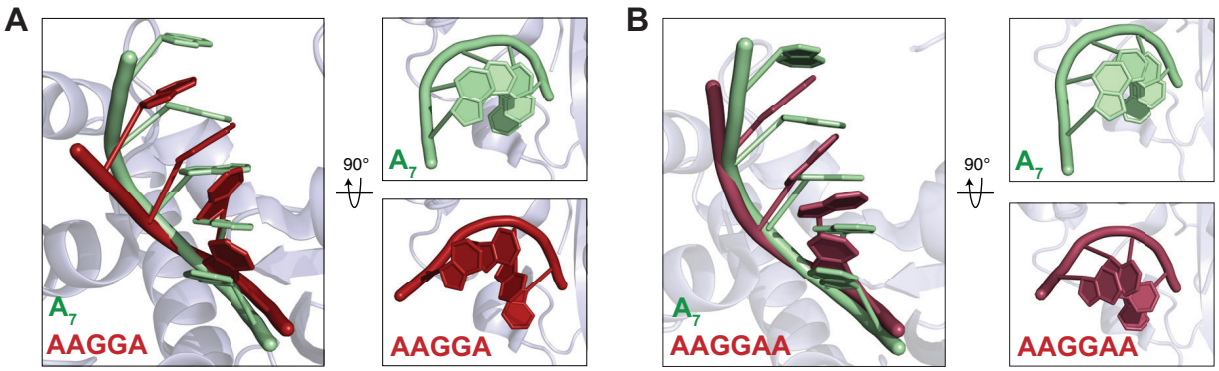


Figure 2.9 Guanosines disrupt base stacking. Superposition with A_7 RNA (green) shows disrupted base-stacking and distorted ribophosphate backbones of (**A**) AAGGA and (**B**) AAGGAA RNAs (red) when bound in the Pan2UCH-Exo (blue) active site.

2.5.2 Y975 is Important for Pan2 Activity and Specificity

In addition to having distorted conformations, G-containing RNAs also stacked more poorly on the Y975 aromatic ring at the active site base (Figure 2.8B, D; compared to Figure 2.7A). Y975 is a conserved aromatic residue in all Pan2 orthologues (Figure 2.10A). To test the significance of this residue, Y975 was mutated to alanine and the resulting catalytic unit was tested for activity and nucleotide specificity. Pan2UCH-Exo Y975A was much less active compared to wild-type, suggesting that Y975 is important in RNA binding and/or catalytic activity (Figure 2.10B). Moreover, while the nucleotide specificity of Pan2UCH-Exo Y975A was the same as wild-type (Figure 2.10C), Pan2UCH-Exo Y975A showed a reduced ability to discriminate against guanosines; wild-type Pan2UCH-Exo is 118× inhibited by -G_3, whereas Pan2UCH-Exo Y975A is only 11.3× inhibited by -G_3 (Figure 2.10D). Thus, the oligo(A) stacking interaction on Y975 is important for Pan2 activity and specificity.

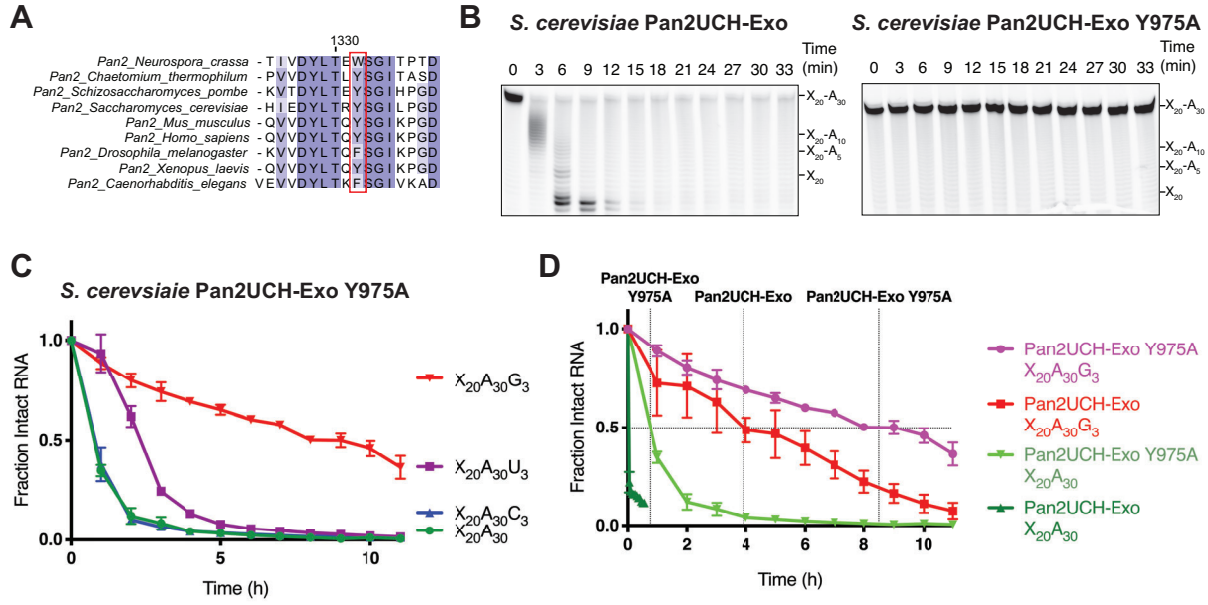


Figure 2.10 Y975 is important for Pan2 activity and specificity. (**A**) Pan2 sequence alignment from representative eukaryotes. The conserved aromatic residue is boxed in red. (**B**) Comparison of deadenylation on 200 nM 5' 6-FAM-labelled $X_{20}A_{30}$ by 100 nM Pan2UCH-Exo and Pan2UCH-Exo Y975A. (**C**) Analysis of deadenylation on RNA substrates with different 3' termini by *S. cerevisiae* Pan2UCH-Exo Y975A using conditions in (B). (**D**) Comparison of deadenylation by Pan2UCH-Exo and Pan2UCH-Exo Y975A on $X_{20}A_{30}$ and $X_{20}A_{30}G_3$ substrates using conditions in (B). The horizontal line represents half the intensity of that at time = 0. The vertical line is the time when half of RNA has been deadenylated.

2.5.3 Pan2 Binds Similarly to Oligo(A) and G-containing RNA

One hypothesis for Pan2 discrimination between A's and G's is by having significantly different binding affinities for these substrates. Thus, I attempted to measure the affinity of Pan2UCH-Exo for A- and G-containing RNAs. FPA, MST, and EMSA assays using fluorescently labelled A_5/A_2G_3 and $A_{15}/A_{12}G_3$ were unsuccessful as Pan2UCH-Exo bound fluorophores (6-FAM, Alexa-647) with micromolar affinity, similar to its affinity for RNA. Thus, EMSAs were carried out using radiolabelled or unlabelled A_5/A_2G_3. Studies with radiolabelled RNAs showed no significant difference in Pan2 affinity between oligo(A) and G-containing substrates (data not shown). EMSA experiments using unlabelled RNAs estimated that Pan2UCH-Exo bound A_{15} with 2× higher affinity relative to $A_{12}G_3$ (Figure 2.11A-B). It must be noted that binding curves are not complete in these experiments due to ligand depletion and high sample requirements for measuring a micromolar interaction. Moreover, as 100 nM Pan2 can efficiently act on 200 nM oligo(A), it is unclear if the micromolar affinities observed in EMSAs represent productive RNA binding in the Pan2 active site. Nonetheless, this small affinity difference, as assessed by EMSAs, is unlikely to account for the dramatic difference in activity, suggesting that discrimination between oligo(A) and G-containing RNAs is not attributable to differential affinities.

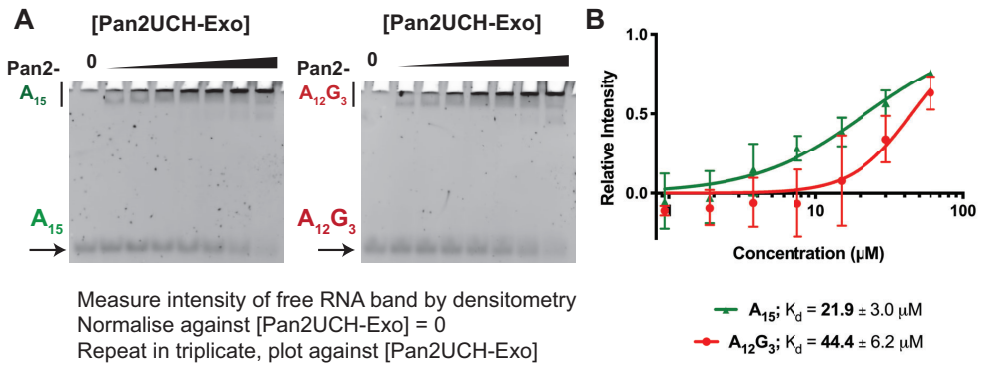


Figure 2.11 Pan2 binds oligo(A) and G-containing RNAs with similar affinity. Pan2UCH-Exo affinity for A_{15} (green) and $A_{12}G_3$ (red) RNAs as estimated by unlabelled EMSAs. (**A**) Free RNA disappears and protein-bound RNA has reduced mobility with increasing protein. (**B**) Normalised intensity plot of the free RNA band against concentration. Each point is a mean of triplicate, and error bars are standard deviation.

2.6 Guanosines Disrupt the Conformation of Oligo(A) in Solution

Circular dichroism (CD) assays were performed to test if guanosines also disrupted oligo(A) structure in solution without Pan2. For nucleic acids, CD signals are sensitive to the chiral higher order geometry, such as the right-handed A-form-like helix predicted to be formed by poly(A). Oligo(A) displayed a characteristic solution spectrum (Figure 2.12A; maximum: 264 nm; minimum: 249 nm), consistent with previous CD studies (Section 1.3.3). This is due to the base-stacked helical structure, which decreases with increased temperature and fewer A's (Figure 2.12A-B). Introducing two G's into oligo(A) RNA disrupted the CD signal and thus the helical conformation (Figure 2.12C). Structural disruption occurred both 5' and 3' to the tandem G's, as signal reduction was greatest when G's are introduced in the middle compared to 5' or 3' ends of oligo(A) (Figure 2.12D). Hence, in the absence of protein binding, oligo(A) forms an intrinsic stacked, helical structure which is disrupted by guanosines.

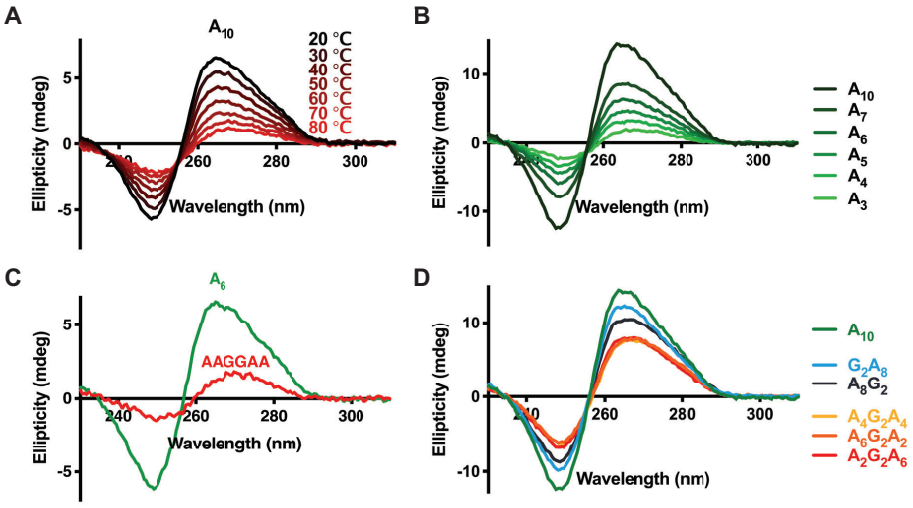


Figure 2.12 Circular dichroism of oligo(A) RNA. (**A**) The CD spectrum of 5 μM A_{10} was monitored from 20-80 °C. Characteristic peak and trough signals are reduced with increasing temperature (dark to light). (**B**) Characteristic signal of 9 μM oligo(A) increases with greater RNA length (light to dark green). (**C**) Guanosines disrupt oligo(A) structure. Introducing two G's (red) into oligo(A) (green) reduces its signal. RNAs are at 9 μM. (**D**) Guanosines disrupt oligo(A) structure in 5' and 3' directions as their introduction into the middle of oligo(A) (orange-red) cause greater signal reduction than to 5' or 3' ends (blue). RNAs are at 25 μM.

2.7 Cytosines and Uracils Can Stack in the Pan2 Active Site

Unlike G's, cytosines and uracils have a relatively modest effect on Pan2 activity (Figure 2.2). Therefore, I studied the effect of cytosines and uracils on poly(A) structure. To understand their effect on the solution conformation of poly(A), I repeated CD studies on oligo(A) RNAs interrupted with two cytosines or uracils. Introduction of C's or U's into oligo(A) also disrupted the signature of a stacked, helical structure of oligo(A) to a degree similar to guanosines (Figure 2.13A). Thus, non-A nucleotides are not compatible with the intrinsic solution conformation of oligo(A). The CD signature of oligo(A) was further compared to other polyribonucleotides. Interestingly, the tendency of poly(A) to form the characteristic single-stranded helix was unique, as other polyribonucleotides had either no significant chiral conformation (poly(U) and poly(G)) or a different conformation (poly(C), maximum: 276 nm). The poly(C) CD signal likely represents the previously-determined single-stranded helical structure (Arnott et al., 1976), but is notably different to that formed by poly(A). Thus, the intrinsic structure of poly(A) is unique and is not formed by other polyribonucleotides.

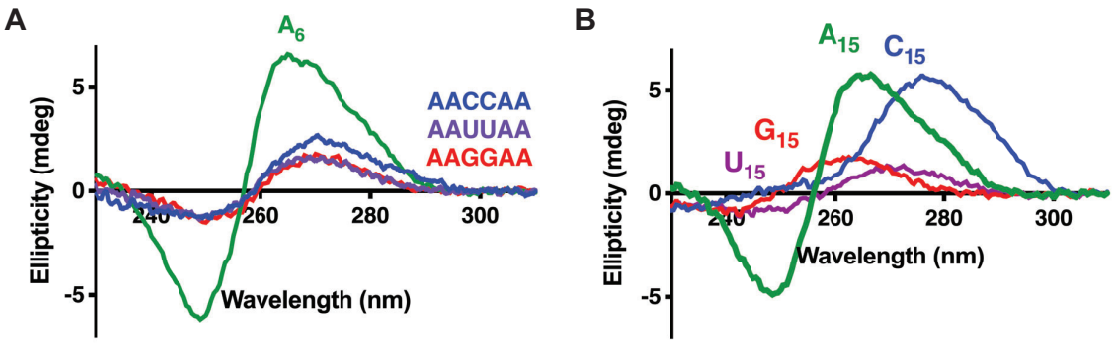


Figure 2.13 Poly(A) forms a unique intrinsic structure by circular dichroism. (**A**) CD spectra of 9 μM A_6, AAUUAA, AACCAA, and AAGGAA. Introducing two uracils (purple) and cytosines (blue) disrupts characteristic signal and thus poly(A) structure, to a similar extent as guanosines (red). (**B**) CD spectra of 5 μM A_{15} (green), U_{15} (purple), C_{15} (blue), and G_{15} (red). The characteristic signature of poly(A) is unique among polyribonucleotides.

To study the conformation of cytosine- and uracil-containing RNAs in the Pan2 active site, I soaked AACCAA and AAUUAA RNAs into Pan2UCH-Exo crystals. These crystals diffracted to 3.1 and 3.0 Å respectively (Table 2.1). Similar to guanosine-containing RNAs, Pan2UCH-Exo did not undergo significant conformational changes and four 3' nucleotides were observed. In contrast to AAGGA and AAGGAA RNAs, however, the bases were well-stacked in an offset parallel manner, and were thus conformationally similar to oligo(A) (Figure 2.14A-B). Furthermore, the ribophosphate backbone of C- and U-containing RNAs followed a similar helical trajectory to that of oligo(A) (Figure 2.14C). The conformation adopted by RNAs in the active site therefore correlates with their ability to be removed by Pan2. Thus, Pan2 is likely able to act on C- and U-containing RNAs because cytosines and uracils are permissive for single-stranded helical formation, despite C- and U-containing RNAs not forming a poly(A)-like conformation *in vitro*. The position of the scissile phosphate is therefore expected to be minimally affected by C's or U's.

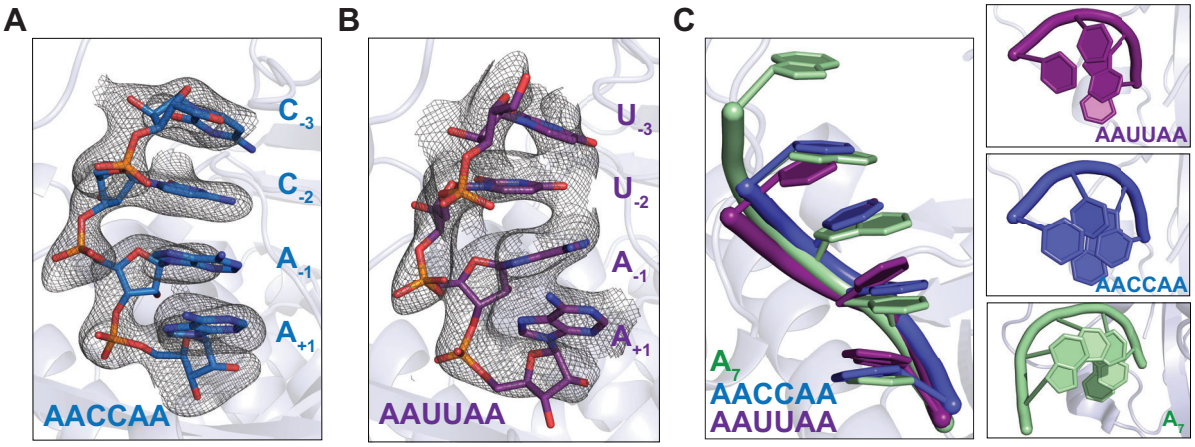


Figure 2.14 Structures of Pan2UCH-Exo-AACCAA and -AAUUAA RNAs. Pan2UCH-Exo is in light blue, AACCAA is dark blue, and AAUUAA is purple. Active site electron density of (**A**) AACCAA or (**B**) AAUUAA RNAs. The grey mesh corresponds to feature-enhanced maps contoured to 1.8σ. (**C**) Superposition of A_7 (green), AAUUAA (purple), and AACCAA (blue). The bases in AACCAA and AAUUAA are relatively well-stacked, and the helical conformation of the ribophosphate backbone is roughly the same as oligo(A).

2.8 Base Stacking is Required for Pan2 Activity

Guanosines disrupt the stacked, helical conformation of oligo(A) both in solution and in the crystal structure. As RNA is required to form local structure in the Pan2 active site for activity, and G's disrupt the oligo(A) helix in both 5' and 3' directions, I hypothesised that Pan2 would be inhibited before guanosines reach the active site base. To determine the range of guanosine influence on Pan2 deadenylation, deadenylation assays were performed with the $X_{20}A_{30}$ substrate interrupted by two G's ($X_{20}A_{14}GGA_{14}$). Compared to the substrate with the poly(A) tail (Figure 2.15A), two interrupting guanosines strongly inhibited Pan2 (Figure 2.15B), consistent with earlier results (Section 2.1.2). The single-nucleotide resolution enabled the observation that Pan2 was inhibited when it encountered -GGAAA and -GGAAAA sequences (Figure 2.15B). At the point of inhibition, as guanosines were distal to the scissile phosphate, G's likely inhibit Pan2 by altering the local 3' conformation.

To further test the effect of disrupted base stacking on Pan2, an RNA with two interrupting dihydrouracils (DHU) was designed. DHU is a uracil analogue which contains the same functional groups as U's except for a C-C single bond instead of a C=C double bond between C5 and C6 (Figure 2.15C). Thus, DHU is non-planar and disrupts oligonucleotide stacking (Dalluge et al., 1996). We previously observed that 3' uracils moderately inhibited Pan2 activity (Section 2.1.2). While the introduction of two uracils had a mild inhibitory effect on Pan2 activity and caused a pause in deadenylation (Figure 2.15D), this inhibition was rapidly overcome. Conversely, the introduction of two interrupting DHUs strongly inhibited Pan2 (Figure 2.15E). As U's and DHUs contain similar functional groups, these results suggest that base stacking, rather than the identity of functional groups, are key to Pan2 recognition of RNA. Hence, Pan2 activity requires the RNA substrate to adopt a local poly(A)-like stacked, helical conformation, and disruption of stacking by guanosines or dihydrouracils strongly inhibit deadenylation by Pan2.

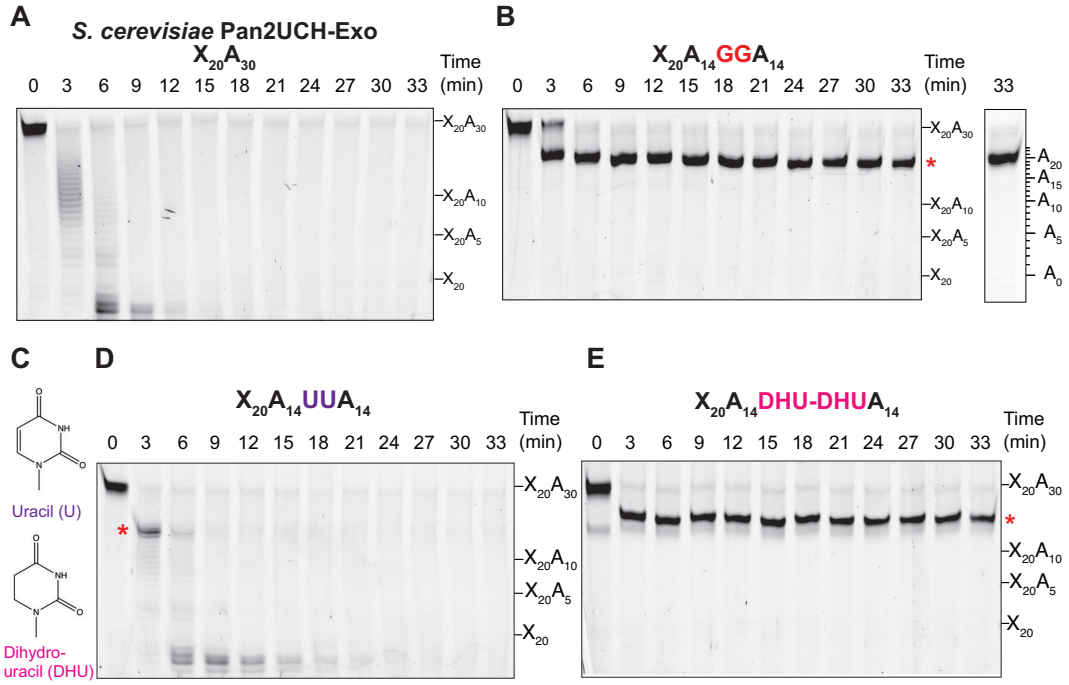


Figure 2.15 Base stacking is required for Pan2 activity. Denaturing polyacrylamide gels show deadenylation by 100 nM *S. cerevisiae* Pan2UCH-Exo on 200 nM (**A**) $X_{20}A_{30}$; (**B**) $X_{20}A_{14}GGA_{14}$; (**D**) $X_{20}A_{14}UUA_{14}$; (**E**) $X_{20}A_{14}DHU\text{-}DHUA_{14}$. Panel B additionally shows the RNA length at which inhibition occurs, relative to a partially digested RNA ladder (not shown). Red asterisks indicate points of inhibition. Panel **C** shows the chemical structures of uridine and dihydrouridine bases.

2.9 Caf1 May Recognise the Poly(A) Structure

2.9.1 Structural Superposition of Pan2UCH-Exo-A_7 and Caf1

As Pan2 requires its substrate to be a stacked, oligo(A)-like, single-stranded helix in its active site, I sought to determine whether the same held true for Caf1. Similar to Pan2, Caf1 is a DEDD-family deadenylase which showed strict poly(A) specificity (Figure 2.4). However, the lack of a Caf1-oligo(A) structure had prevented the molecular mechanism of poly(A) recognition by the Caf1 active site to be determined. To obtain a Caf1-oligo(A) structure, we carried out extensive co-crystallisation and soaking experiments. However, in contrast with Pan2, crystal packing of human and *S. pombe* Caf1 was much more compact, which obstructed the Caf1 active site. This, together with the low expected affinity of the exonuclease domain for RNA, likely precluded RNA binding in the active site of Caf1 crystals.

To determine whether the single-stranded RNA helix formed by oligo(A) in the Pan2 active site is compatible with Caf1, I carried out structural superposition of Pan2UCH-Exo and *S. pombe* Caf1. Pan2 Exo and Caf1 are structurally homologous (Figure 2.16A), allowing accurate superposition. This revealed that the key Y975 residue at the Pan2 active site base is replaced by a leucine (L125), which has a hydrophobic but non-aromatic side chain (Figure 2.16A). This suggests that the stacking interaction may be less important in RNA binding by Caf1 and its deadenylation activity. Oligo(A) from the Pan2UCH-Exo-A_7 structure did not significantly clash with the Caf1 active site, but notably, adenine bases were not freely exposed to solvent and were instead proximal to the active site surface (Figure 2.16B). Finally, structural superposition allowed putative interactions to be hypothesised, suggesting that Caf1 may form base-specific interactions with adenines, consistent with the observation that Caf1 was specific for poly(A) and was inhibited by all other non-A nucleotides (Figure 2.16C). Further experiments will be required to test the model of oligo(A) binding in the Caf1 active site.

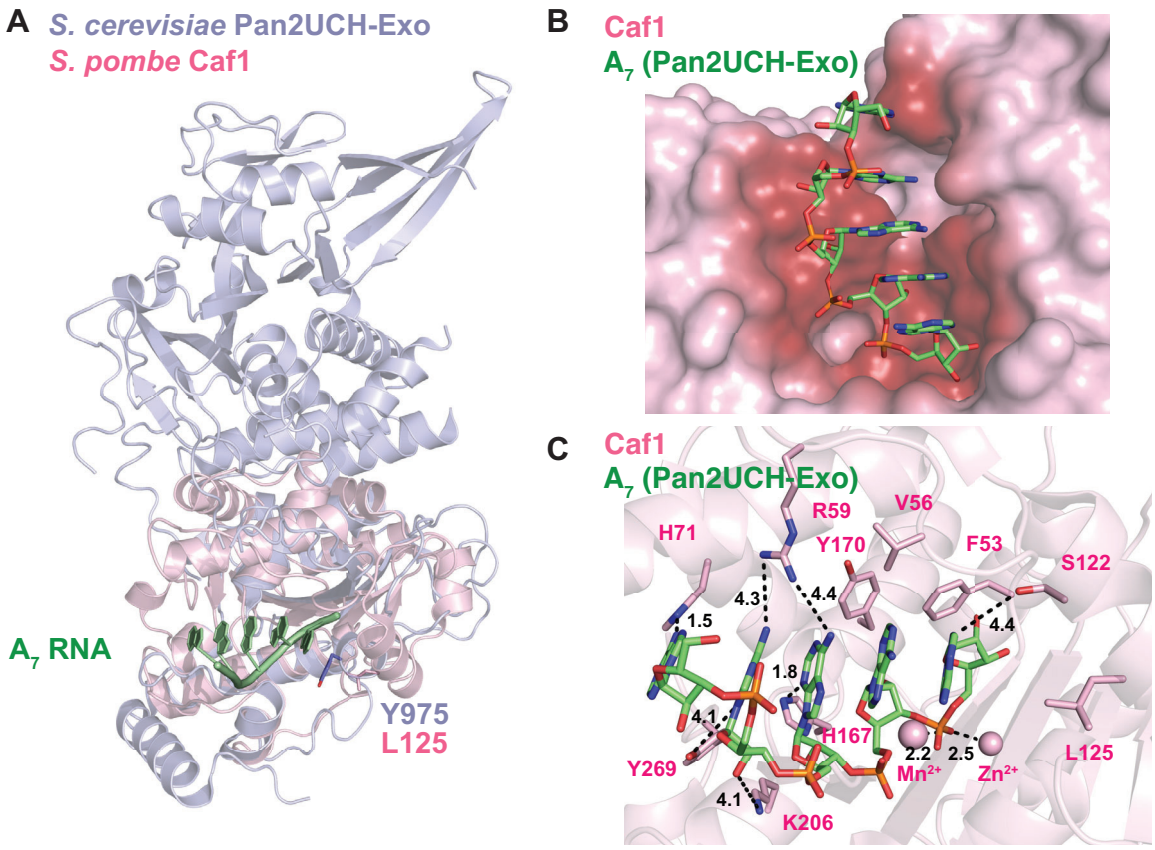


Figure 2.16 Superposition of Pan2UCH-Exo (blue)–A_7 (green) and *S. pombe* Caf1 (pink, PDB: 3G0Z). (**A**) Overall structures of the Pan2 exonuclease domain and Caf1 show that the enzymes are structurally homologous. At the active site base, leucine (L125, pink sticks) replaces the conserved aromatic residue (Y975, blue sticks). (**B**) The Caf1 active site cavity can accommodate A_7. Caf1 surface is coloured by RNA proximity from dark (<3 Å) to light (>7 Å). (**C**) Putative interactions between Caf1 and stacked, helical oligo(A). Amino acid side chains near the modelled substrate are pink sticks. Possible hydrogen bonds are black dashes. Distances are in Å.

2.9.2 Caf1 Likely Requires Base Stacking for Activity

To test if base-stacking is required for Caf1 activity, deadenylation assays with interrupting non-A nucleotides were repeated with Ccr4-inactive Ccr4-Not (only Caf1 active). Similar to Pan2, Caf1 was strongly inhibited by guanosines relative to pure poly(A) (Figure 2.17A-B). In addition, Caf1 was inhibited by uracils to a stronger degree than Pan2 as it was more strongly stalled as it encounters the tandem uracils (Figure 2.17C). Nonetheless, Caf1 could eventually remove uracils over the time course. These results are in line with previous observations of strict poly(A) specificity (Figure 2.4) as well as structural superposition showing possible base-specific interactions with the substrate (Figure 2.16C). However, DHUs, which contain the same functional groups as U's but disrupt stacking, more strongly inhibited Caf1 than uracils (Figure 2.17D). Therefore, Caf1 likely also requires base stacking for poly(A) specificity. Consistent with structural superposition, the stacked, helical poly(A) conformation and base-specific contacts together contribute to Caf1 nucleotide specificity.

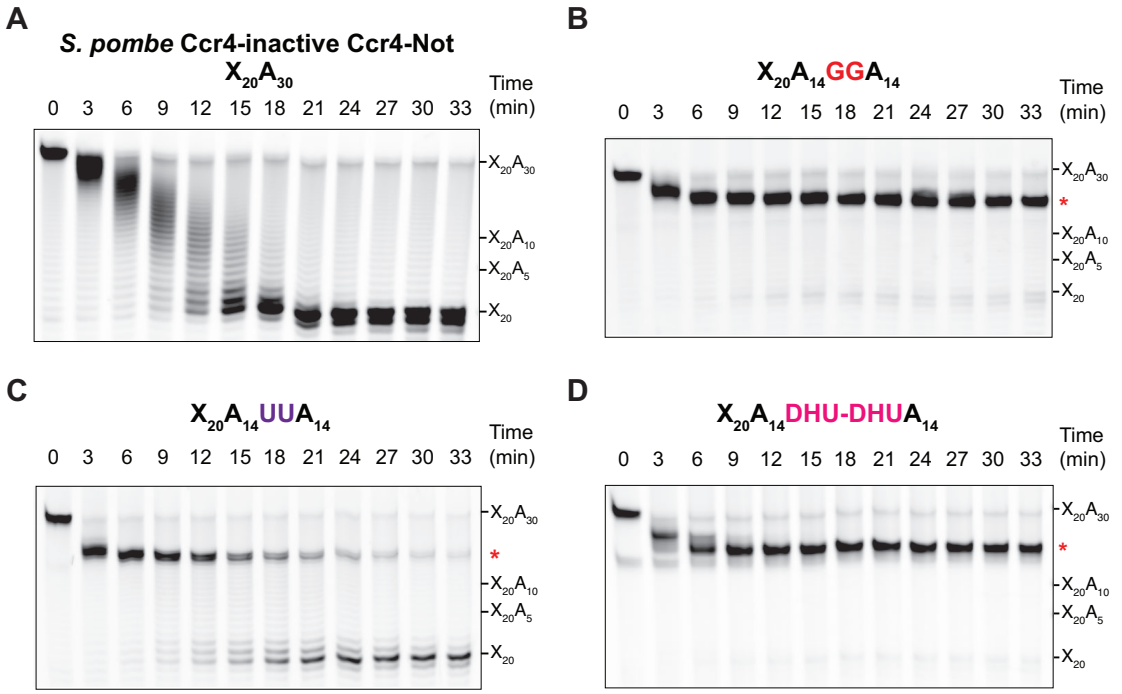


Figure 2.17 Base stacking is required for Caf1 activity. Denaturing RNA gels show deadenylation by 50 nM *S. pombe* Ccr4-inactive Ccr4-Not on 200 nM (**A**) $X_{20}A_{30}$; (**B**) $X_{20}A_{14}GGA_{14}$; (**C**) $X_{20}A_{14}UUA_{14}$; (**D**) $X_{20}A_{14}$DHU-DHUA_{14}. Red asterisks indicate points of inhibition.

2.10 Effect of Adenine Modifications

2.10.1 Effect of Purine Substituents on Poly(A) Structure

Guanine, despite being a planar aromatic purine, disrupts stacking of a poly(A) helix both alone in solution and in the Pan2 active site. This is in seeming contradiction with previous studies, which showed more thermodynamically favourable stacking by G-G, G-A, and A-G than A-A dinucleotides (Brown et al., 2015; Friedman and Honig, 1995). Interestingly, a molecular dynamics study showed that the stacked state of A-A is a steeper thermodynamic minimum compared to that of G-G in their free energy profiles, suggesting that stacked A-A may be a more stable state compared to stacked G-G (Norberg and Nilsson, 1995). These results suggest that simple thermodynamics of dinucleotide stacking are not predictive of the ability to form the single-stranded A-form-like helix of oligo(A).

Guanines and adenines differ in the position of amine (*N*6 in adenine, *N*2 in guanine) and an additional carbonyl (*C*6 in guanine) around the purine moiety. While these substituents maintain heterocyclic aromaticity, they affect the electronic distribution within the aromatic ring system through inductive and resonance effects. These functional groups also affect the overall dipole moment of the bases (Sponer et al., 2001). I hypothesised that the distribution of functional groups in adenine facilitates the single-stranded RNA helix. To test this, oligo(A) RNA substrates interrupted by two purines (P), lacking the *C*6 amine, or two 2-aminopurines (2AP), which lack the *C*6 amine but contain an additional *C*2 amine, were compared to adenine (Figure 2.18A). Introduction of purines or 2APs disrupted the characteristic signal of oligo(A) by CD; thus, they do not form the intrinsic helical conformation of poly(A) (Figure 2.18B). Additionally, purine- or 2AP-containing oligo(A) RNAs had shifted maxima and minima, suggesting a subtly different conformation. $X_{20}A_{30}$ substrates with interrupting purines or 2APs were used in deadenylation assays to test their effect on Pan2. Consistent with the disruption

of the intrinsic structure of oligo(A), these nucleotides inhibited Pan2 activity (Figure 2.18C-E). While these bases are not entirely permissive for the stacked, helical conformation, they inhibit Pan2 less than guanosines, suggesting that guanosines more strongly disrupt the stacked poly(A) helix compared to purines or 2-aminopurines. The distribution of functional groups around purine is thus important for the intrinsic poly(A) structure. These results suggest that guanosines disrupt stacking as the electrostatic distribution of functional groups is incompatible with the helical constraints of poly(A). The distribution of functional groups around the purine moiety in adenine likely enables electrostatic complementarity upon stacking, thereby facilitating poly(A) helix formation.

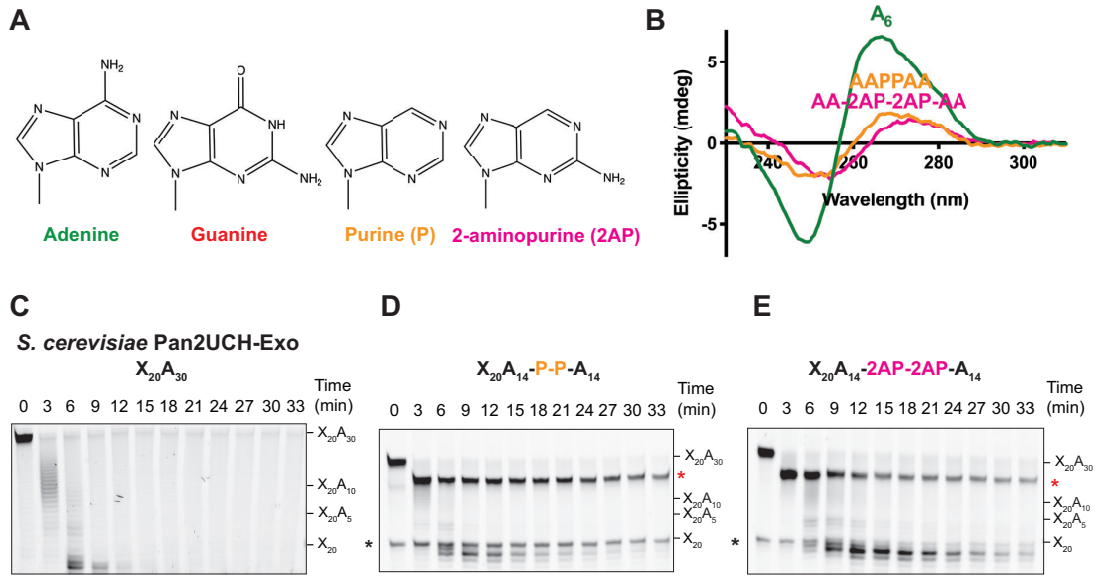


Figure 2.18 Electronic distribution of functional groups is important for poly(A) structure. (**A**) Chemical structures of adenine (green), guanine (red), purine (orange), and 2-aminopurine (pink). (**B**) CD spectra of 9 μM A_6, AAPPAA, and AA-2AP-2AP-AA. Introducing P's and 2AP's disrupts oligo(A) structure and causes a shift in peak and trough positions. (**C-E**) Deadenylation assays by 100 nM Pan2UCH-Exo of 200 nM (**C**) $X_{20}A_{30}$, (**D**) $X_{20}A_{14}$-P-P-A_{14}, and (**E**) $X_{20}A_{14}$-2AP-2AP-A_{14}. Red asterisks indicate points of inhibition; black asterisks indicate non-specific RNA. Pan2 is inhibited by interrupting P's or 2AP's.

2.10.2 Effect of Adenine Modifications on Poly(A) Structure

Adenines can be modified by *N*6-methylation or deamination to form *N*6-methyladenine (m^6A) or inosine (I) respectively (Figure 2.19A). These modifications are commonly found in mRNAs, but it was not known if they occur in the poly(A) tail (Section 1.3.2). These

modifications can also cause changes in the electronic distribution of adenine. CD and deadenylation experiments with oligo(A) containing two m^{6}A's or I's were performed. m^{6}A did not have a large effect on the signature CD spectrum of oligo(A); however, the modified RNA showed slightly shifted peak and trough wavelengths, demonstrating the sensitivity of higher-order structure to even small changes of electronic distribution (Figure 2.19B). In contrast, inosines, similar to other modified purines, disrupted the poly(A)-like conformation when introduced into oligo(A) (Figure 2.19B). Deadenylation assays with $X_{20}A_{30}$ substrates with interrupting nucleotides were tested for their effect on Pan2 (Figure 2.19C-E). Consistent with CD experiments, m^{6}A did not inhibit deadenylation, whereas inosines slightly inhibited Pan2. Interestingly, inosines appeared to inhibit Pan2 activity less than purines or 2-aminopurines, possibly due to the more similar electronic distribution between inosines and adenines. These reinforce the above results that the intrinsic helical conformation of poly(A) is dependent on the electronic distribution around the purine moiety of adenine.

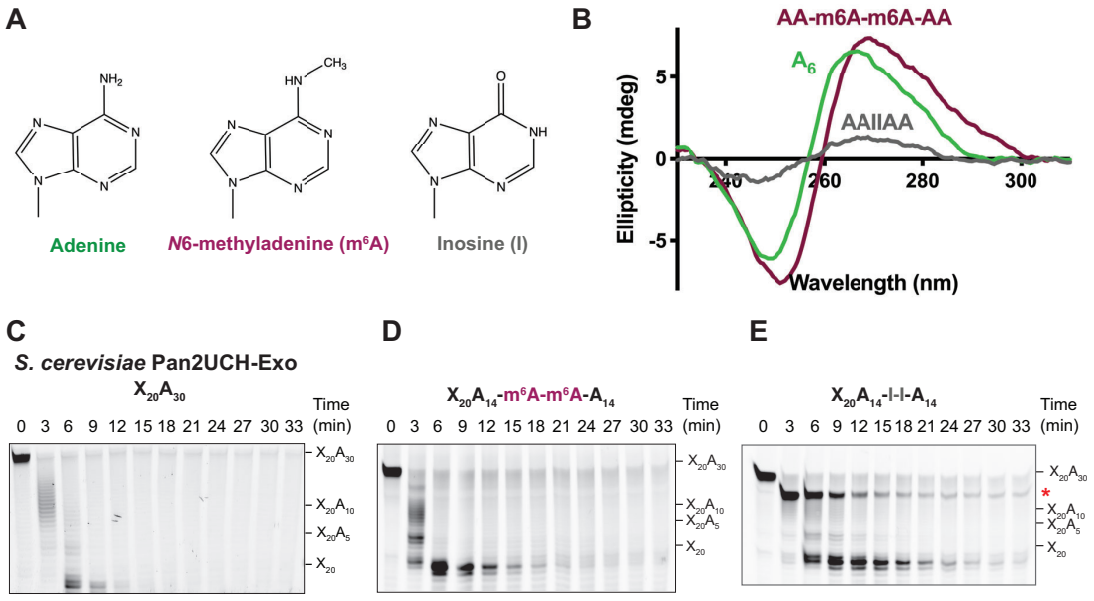


Figure 2.19 Effect of adenine modifications on oligo(A) structure. (**A**) Chemical structures of adenine (green), *N*6-methyladenine (purple), and inosine (grey). (**B**) CD spectra of 9 μM A_6, AA-m^{6}A-m^{6}A-AA, and AA-I-I-AA. *N*6-methylation does not affect oligo(A) structure but causes a slight shift in peak and trough positions; inosine disrupts the spectrum of oligo(A). (**C-E**) Deadenylation assays by 100 nM Pan2UCH-Exo of 200 nM (**C**) $X_{20}A_{30}$, (**D**) $X_{20}A_{14}$- m^{6}A-m^{6}A -A_{14}, and (**E**) $X_{20}A_{14}$-I-I-A_{14}. Red asterisks indicate points of inhibition. Pan2 is unaffected by m^{6}A's, but is inhibited by I's.

2.11 Discussion

2.11.1 Investigation of Nucleotide Specificity by *In Vitro* Deadenylation Assays

In Sections 2.1, 2.2, and 2.5 , I investigated the nucleotide specificity of deadenylase enzymes both alone and in the context of the full multiprotein complexes. This was carried out by monitoring the disappearance of the fluorescently-labelled intact substrate. However, given the relatively low specificity of Pan2 deadenylation and the designed RNA (Figure 2.2), Pan2 degrades the upstream UTR-like sequence. Due to electrophoresis conditions and image processing, not all of the initial signal can be accounted for by the total intensity of degradation products, for example in the later time points of Figure 2.2A. This raises the possibility that the assays shown may have been contaminated by other non-specific nucleases. Future experiments may be improved by altering electrophoresis conditions in order to be able to resolve all degradation products, including the fluorophore alone. Alternatively, a non-cleavable bond, such as a phosphorothioate linkage (Strzelecka et al., 2020), can be introduced between the upstream RNA and poly(A) tail, allowing the observation of fully deadenylated, fluorescently-labelled RNA species.

The assays shown above were carried out under specific substrate and enzyme concentrations, where the substrate is in slight excess over the enzyme (quasi-steady state). These conditions were chosen due to the practicality of carrying out deadenylation assays, including detectability of fluorescently-labelled RNA, a reasonable experimental time, and ability to distinguish between differences in enzyme behaviour. Nonetheless, these assay conditions prevent the quantitative characterisation of kinetic parameters of the association, cleavage, and dissociation steps, as well as distinguishing between productive and non-productive RNA binding. Moreover, from these assays alone, we cannot determine whether substrate specificity at the

RNA binding or degradation steps within the catalytic cycle. These limitations can be partially resolved by carrying out *in vitro* deadenylation assays under different conditions. For example, pre-steady-state conditions, where the enzyme concentration is significantly greater than substrate concentration, combined with the resolution of reaction products at an individual nucleotide level, enable binding and catalytic parameters for different substrates to be modelled (Jankowsky and Putnam, 2010; Licatalosi et al., 2020). For example, such analysis characterising a consecutive multi-step reaction has been carried out for the Rrp6 exonuclease, enabling its specificity and the role of terminal nucleotides in substrate recognition to be elucidated (Axhemi et al., 2020). Similar experiments with deadenylases can resolve productive/non-productive binding as well as catalytic rate constants for individual deadenylation steps for the full assessment of deadenylation specificity.

2.11.2 Deadenylase Specificity Prevents 3′ UTR Degradation

The above experiments have investigated the specificity of the major eukaryotic deadenylases Pan2, Ccr4, and Caf1 alone and in the context of the full Pan2-Pan3 and Ccr4-Not complexes. Poly(A) RNA conformation contributes to the specificity of the DEDD superfamily deadenylases Pan2 and Caf1, with Caf1 more stringent against non-A nucleotides than Pan2 (Section 2.1.2, Section 2.22.2). Consistent with this, Pan2 can degrade the upstream X_{20} sequence *in vitro*, which mimics the 3' UTR, whereas Ccr4-Not stops at the end of the poly(A) tail. The higher stringency of Caf1 can be explained by the proximity of amino acid side chains to the modelled oligo(A), which may form adenine-specific interactions (Section 2.9.1, Figure 2.16), in addition to the requirement for the base-stacked, helical conformation. Ccr4 also has higher adenine specificity compared to Pan2, likely mediated by base-specific contacts in the Ccr4 active site that select against pyrimidines and guanosines (Section 1.2.3.1). However, *in vitro* deadenylation assays show that Ccr4 is able to remove uracils (Section 2.2); the molecular basis of how the Ccr4 active site can recognise uracils is yet to be tested.

The different stringencies of Pan2, Caf1, and Ccr4 may be linked to their *in vivo* roles (Figure 2.20). A biphasic model of deadenylation postulates that Pan2-Pan3 removes the distal poly(A) tail, whereas Ccr4-Not removes A's closer to the 3' UTR. Thus, Ccr4 and Caf1 are more likely to encounter non-A nucleotides, and may require more stringent adenine specificity to ensure they only degrade poly(A). In contrast, in this model, Pan2 would likely only encounter poly(A). Thus, it may rely on the less specific recognition of poly(A) structure. The biphasic model is also consistent with the below-expected frequency of tandem guanosines in the 3' UTR (Section 2.3), which would not be required to prevent spurious degradation if Ccr4-Not is the primary complex removing the proximal poly(A) tail. Nonetheless, it remains unclear whether and how deadenylases would be stopped upon encountering the 3' UTR *in vivo*.

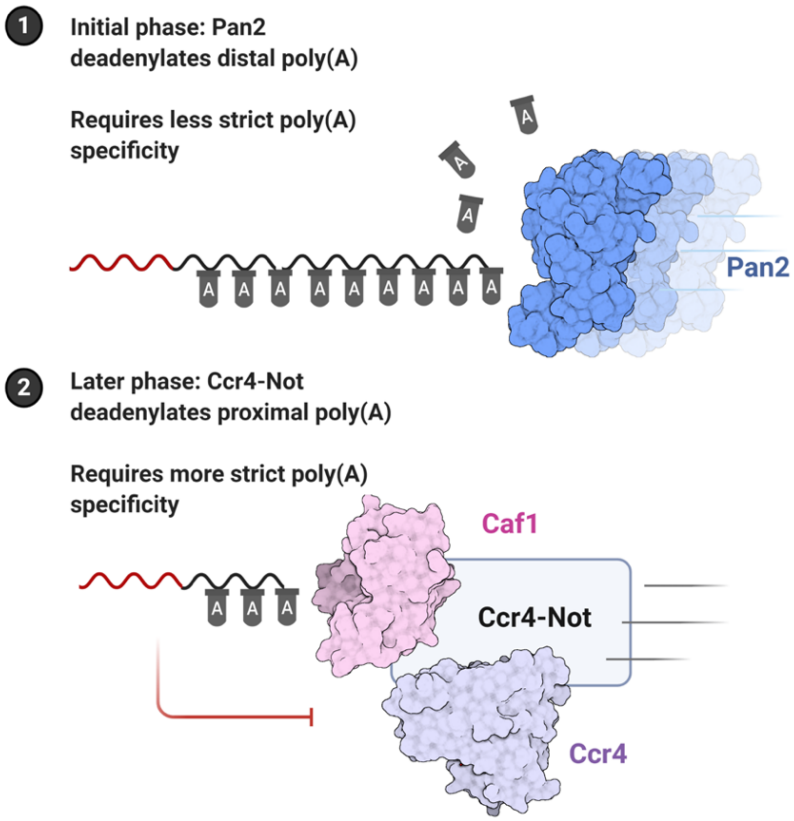


Figure 2.20 Stringency of poly(A) specificity correlates with the hypothesised *in vivo* roles of Pan2-Pan3 and Ccr4-Not according to the biphasic model of deadenylation. (**1**) Pan2-Pan3 is thought to remove the UTR-distal poly(A) tail, and would likely only encounter poly(A). It would thus require lower poly(A) specificity; *in vitro*, it is able to remove uracils and cytosines. (**2**) Ccr4-Not is thought to remove the more UTR-proximal poly(A) tail, and thus would encounter both poly(A) and the non-A UTR sequence. It would thus require higher poly(A) specificity; *in vitro*, both Caf1 and Ccr4 are inhibited by non-A nucleotides.

2.11.3 Guanosines Inhibit Deadenylation

Next-generation sequencing has revealed that non-A nucleotides can be incorporated into the poly(A) tails of higher eukaryotic mRNAs (Chang et al., 2014; Legnini et al., 2019; Liu et al., 2019). 3' uracils, added by Cid1 in fission yeast and TUT4/TUT7 in higher eukaryotes, have a relatively small effect on Ccr4 and Pan2 activity (Section 2.1.2, Section 2.2). Thus, 3' uridylation alone may not inhibit deadenylation. Instead, poly(U)-selective RNA-binding proteins such as the Lsm1-7/Pat1 complex may be required to prevent deadenylase activity on oligo(U) tails.

In contrast, guanosines strongly inhibit deadenylation by Pan2-Pan3 and Ccr4-Not (Section 2.1.2, Section 2.2). A recent report identified TENT4A and TENT4B as non-canonical polymerases responsible for non-adenosine addition into the poly(A) tail (Lim et al., 2018). Additionally, more frequently guanylated transcripts show increased half-lives (Chang et al., 2014). *In vivo*, poly(A) transcripts would be in vast excess over guanylated transcripts. Thus, single guanosines would likely be sufficient to inhibit deadenylation by Pan2 (Figure 2.3) and Ccr4-Not (Lim et al., 2018). Taken together, these data suggest that the increased stability of transcripts with guanylated poly(A) tails is due to inhibition of both major deadenylases, thereby slowing the rate-limiting step of canonical mRNA decay.

To date, adenine modifications have only been found in trypanosomal poly(A) tails (Viegas et al., 2020). In this study, m^6A's were identified in the VSG transcript and contributed to slowed deadenylation. However, the lack of inhibition by m^6A's on Pan2 deadenylation (Figure 2.19) suggests that methylation alone is insufficient to inhibit deadenylases. Regardless, whether non-A nucleotides are found in the poly(A) tail in other species remains unknown, as limitations in sequencing techniques may have precluded their detection. Increased sensitivity and improvements in sequencing techniques, such as long-read nanopore sequencing, may detect such modifications if they occur in the poly(A) tail (Krause et al., 2019).

2.11.4 Mechanism of Pan2 Recognition of Poly(A)

The structure of poly(A) RNA has been studied by numerous methods (Section 1.3.3). From these studies, a single-stranded helical structure of poly(A) had been proposed, but its biological significance was unknown. Extension of the observed adenosines within the crystal structure by superposition (Section 2.4) forms a helix whose parameters are similar to the previously proposed model (Figure 2.21A). The observed adenosines are also sterically compatible with the Caf1 active site (Section 2.9.1). Thus, Pan2, and possibly Caf1, exploits the intrinsic helical structure of poly(A) in poly(A) tail recognition. As poly(A) does not display fibre diffraction even at low temperatures, poly(A) likely interconverts between random coil and a stacked, helical structure in solution, which is stabilised when the ribophosphate backbone of oligo(A) RNA is bound in the Pan2 active site.

These results also represent a novel paradigm of adenine recognition and discrimination from other nucleotides. A metastudy of A-containing structures showed that adenines are frequently buried in hydrophobic cavities due to the hydrophobic purine moiety (Moodie et al., 1996). In contrast, in the Pan2UCH-Exo-A_7 structure, adenine bases are relatively unburied and point outwards into solvent. In the same study, some structures contained aromatic side chains which are stacked against adenine in an offset parallel manner. Together with the conserved π-stacking interaction between the terminal A and Y975 (Figure 2.7), this demonstrates the importance of stacking interactions in adenine recognition. In the context of poly(A), its tendency to form the base-stacked, helical structure must also be considered. Our experiments demonstrate that this unique conformation is due to the electronic distribution around the central purine, as changing the position of functional groups disrupts stacking (Figure 2.18). Future experiments may focus on elucidating the importance of different chemical moieties in this stacked, helical structure. For example, the importance of sugar pucker can be tested using locked nucleic acids (LNA), which enforce the 3'-*endo* conformation, and the role of the 2'-OH

group can be examined using poly(A) DNA or 2'-methylated nucleic acids. These modified nucleotides could provide insight into formation of the intrinsic poly(A) structure, and why poly(A) may be conserved throughout evolutionary history.

This indirect readout of single-stranded RNA structure is reminiscent of the mechanism where DNA-binding proteins such as the Trp repressor utilise indirect readout to recognise specific DNA sequences (Otwinowski et al., 1988). In this and many other protein-DNA complexes such as the papillomavirus E2 DNA-binding domain and TATA binding protein (Hegde et al., 1992; Kim et al., 1993), the target DNA sequence facilitates a specific DNA structure, which deviates from the canonical B-form DNA duplex, which allows its recognition by DNA-binding proteins. Analogous to this, sequences which are permissive for formation of the stacked, helical structure of poly(A) can be recognised by the Pan2 active site.

The intrinsic structure of poly(A), which is disrupted by G's, allows guanosines to be distinguished from adenosines as guanosines disrupt the base-stacked, helical structure. This mechanism of conformational disruption contrasts with conventional views of how adenines and guanines, both similarly-shaped purines, are discriminated on the molecular level through different patterns of hydrogen bond donors and acceptors (Nobeli et al., 2001). An example is the cytoplasmic poly(A)-binding protein Pab1, which recognises poly(A) by base-specific contacts and not its intrinsic structure (Deo et al., 1999; Safaee et al., 2012a). Although Pan2 binds guanosine-containing RNAs with a similar affinity to oligo(A) (Figure 2.11), G-containing RNAs are unstacked in the Pan2 active site relative to oligo(A) (Figure 2.9). Pan2 is also inhibited before G's reach the active site base, suggesting that guanosines disrupt the local structure to inhibit Pan2 (Figure 2.15). This provides further support that Pan2 requires an oligo(A)-like structure in its active site for deadenylation. However, it remains unclear how cleavage of the terminal phosphate can be inhibited by the disruption of stacking. It is plausible that unstacking leads to a subtle change in the position of the scissile phosphate group, which

cannot be observed in the crystal structure due to limited resolution. Attempts to improve the resolution by crystallisation of the catalytic unit (including deletion of flexible termini and loop regions) from the thermophilic yeast *Chaetomium thermophilum* did not yield crystals.

As C- and U-containing RNAs do not strongly inhibit deadenylation by Pan2 (Section 2.1.2) and are stacked in a helical conformation when bound to Pan2 (Section 2.7), a model of Pan2 activity can be proposed (Figure 2.21B). In this model, bases permissive for the oligo(A)-like structure (adenines, cytidines, uridines) can be removed by Pan2. Cytosines and uracils slightly inhibit deadenylation by Pan2, likely due to the entropic cost of ordering to form the single-stranded A-form-like helix. In contrast, bases which disrupt stacking (guanines, dihydrouridines) strongly inhibit Pan2.

The preferred substrate of the Pan2 exonuclease is poly(A); however, it can readily degrade C- and U-containing RNAs. Its preference for adenosines and relative lack of specificity is reminiscent of the polyadenylation activity of polynucleotide phosphorylase (PNPase), a reversible phosphorolytic enzyme which is conserved in prokaryotes and organelles (Mohanty and Kushner, 2011). PNPase synthesises heteropolymeric tails on prokaryotic and organellar transcripts which are thought to be committed for degradation (Mohanty and Kushner, 2000; Mohanty et al., 2004; Rott et al., 2003; Yehudai-Resheff et al., 2001; Yehudai-Resheff et al., 2003). These mixed tails contain approximately 50% adenosines, suggesting that tail addition by PNPase displays a preference for A's but is relatively non-specific. This pattern of specificity is similar to the deadenylation specificity of Pan2, raising the possibility that PNPase specifies for adenosines through its intrinsic helical structure. Elucidation of the mechanism of PNPase polymerase specificity could resolve this possibility.

Finally, Caf1 likely recognises the stacked, helical structure of poly(A) in a similar manner, as stacking inhibitors such as DHU also strongly inhibit Caf1 activity. However, the lack of a conserved aromatic residue at the base of the active site suggests that poly(A) stacking may

not be the predominant mechanism for specificity. This is further shown by the inhibition of Caf1 deadenylation by cytosines and uracils, which are permissive for stacking in the single-stranded poly(A) helical conformation (Figure 2.4). The additional specificity of Caf1 may be due to additional base-specific contacts with poly(A), as observed in structural superposition of Pan2UCH-Exo-A_7 with Caf1 (Figure 2.16). Elucidation of the molecular mechanism of Caf1 recognition of poly(A) may enable the relative contributions of base-stacking and base-specific interactions to be clarified.

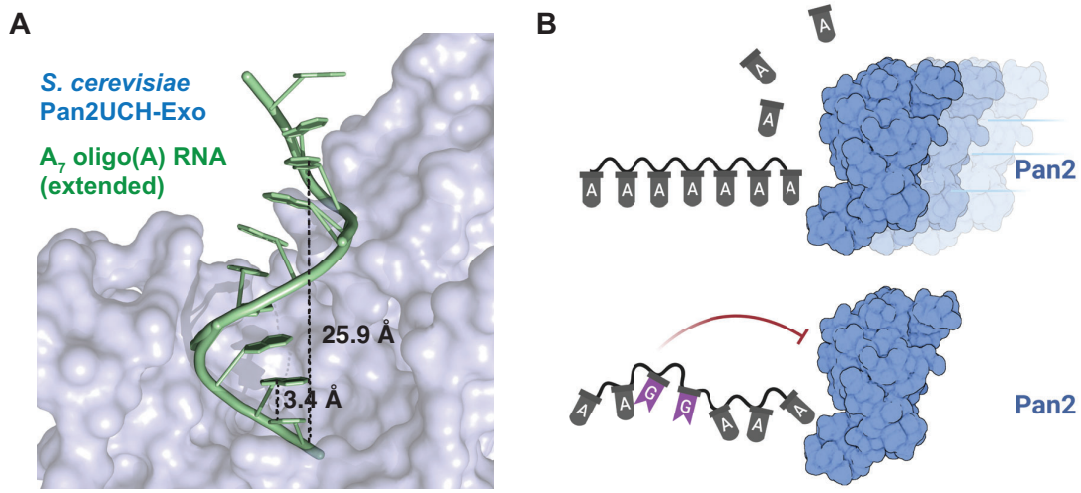


Figure 2.21 Pan2 recognises the intrinsic helical structure of poly(A) RNA. (**A**) Extension of the oligo(A) helix bound to the Pan2 active site, modelled by duplication and superposition of the observed A_5. Distances are shown in Å. (**B**) Proposed model for poly(A) RNA recognition by Pan2 through its stacked, single-stranded, A-form-like helical conformation. Bases which are permissive to this structure can be removed. Disruption of this conformation inhibits Pan2.

2.11.5 Poly(A) Structure in Other Biological Processes

Given the numerous roles of poly(A) in post-transcriptional regulation of gene expression, the intrinsic single-stranded poly(A) helix is unsurprisingly implicated in other biological processes.

Firstly, adenine stacking has been observed in polyadenylation, the processive addition of adenosines to the 3'-end of a nascent transcript by poly(A) polymerase. A structure of yeast poly(A) polymerase in complex with ATP and oligo(A) has been determined, providing a model of adenosine addition to an elongating poly(A) tail (Balbo and Bohm, 2007). In this structure, most adenines flip out to form base-specific contacts with the active site channel (e.g. A_{-2}, A_{-3}, A_{-4}, A_{-5}). Interestingly, the incoming ATP π-stacks against the terminal adenine of the poly(A) tail (A_{-1}). This mode of stacking is reminiscent of the helical structure observed in the Pan2 active site, suggesting that the stacking geometry of an incoming ATP against existing poly(A) contributes to the specificity of adenosines addition by Pap1 (Figure 2.22A).

Recent studies have identified the role of the single-stranded oligo(A) helix in translation quality control (Chandrasekaran et al., 2019; Tesina et al., 2019). Translation is normally terminated when the ribosome encounters a stop codon before the poly(A) tail. However, if the ribosome encounters poly(A) RNA, the ribosome stalls and a quality control pathway results in mRNA degradation and ribosome disassembly (Doma and Parker, 2006; Juszkiewicz and Hegde, 2017; Sundaramoorthy et al., 2017). In the A-site of rabbit and yeast ribosomes, poly(A) forms a single-stranded RNA helix, which stacks between 18S rRNA bases A1825 and C1698 (Figure 2.22B). This leads to a structural rearrangement in the decoding centre, contributing to ribosome stalling. Analogous to poly(A) tail recognition by Pan2, guanosine introduction (when poly(A) is replaced with $(AAG)_n$) disrupts the poly(A) helix, and thus the ribosome does not stall (Arthur et al., 2015; Juszkiewicz and Hegde, 2017). A recent study showed that this mechanism may be used by endogenous substrates to tune gene expression

(Matsuo et al., 2020): for the yeast *sdd1* transcript, an A-rich sequence in the coding region causes ribosome stalling. However, the sequence which causes stalling is not poly(A), but is able to form the stacked, A-form helix in the A-site. This is again analogous to Pan2, where sequences permissive for the oligo(A)-like structure can be exploited for biological function.

Finally, the intrinsic base-stacked helix of oligo(A) is observed in the active sites of exosome exonucleases and a helicase involved in nuclear RNA quality control. Oligo(A) tailing by the nuclear TRAMP complex can mark aberrant RNAs for degradation in eukaryotes (Glaunsinger and Lee, 2010). Together with TRAMP, the oligo(A) tail recruits the exosome for RNA decay. Structured RNAs can also be resolved by the Mtr4 helicase of TRAMP. Furthermore, the exosome is implicated in rapid cytoplasmic 3'-5' mRNA degradation following deadenylation. Structures of archaeal (Lorentzen and Conti, 2005) and eukaryotic (Lorentzen et al., 2008; Wasmuth et al., 2014; Zinder et al., 2016) exosome nucleases demonstrate that oligo(A) can form a single-stranded helix in their active sites (Figure 2.22C). Similar to Pan2, these exonucleases contact their RNA substrates mainly via the ribophosphate backbone, enabling this base-stacked conformation to form. In the case of Rrp44, the single-stranded oligo(A) helix is promoted by stacking on a tyrosine residue in the active site base. The oligo(A) helix is further observed in the core RecA domains of the Mtr4 helicase (Weir et al., 2010). However, given the broad specificity of unwinding and degradation required of TRAMP and the exosome, it is unlikely that oligo(A) stacking significantly contributes to nucleotide specificity. Instead, these structures demonstrate the tendency for the single-stranded helix to form when the ribophosphate backbone of an oligo(A) RNA is constrained by the protein active site.

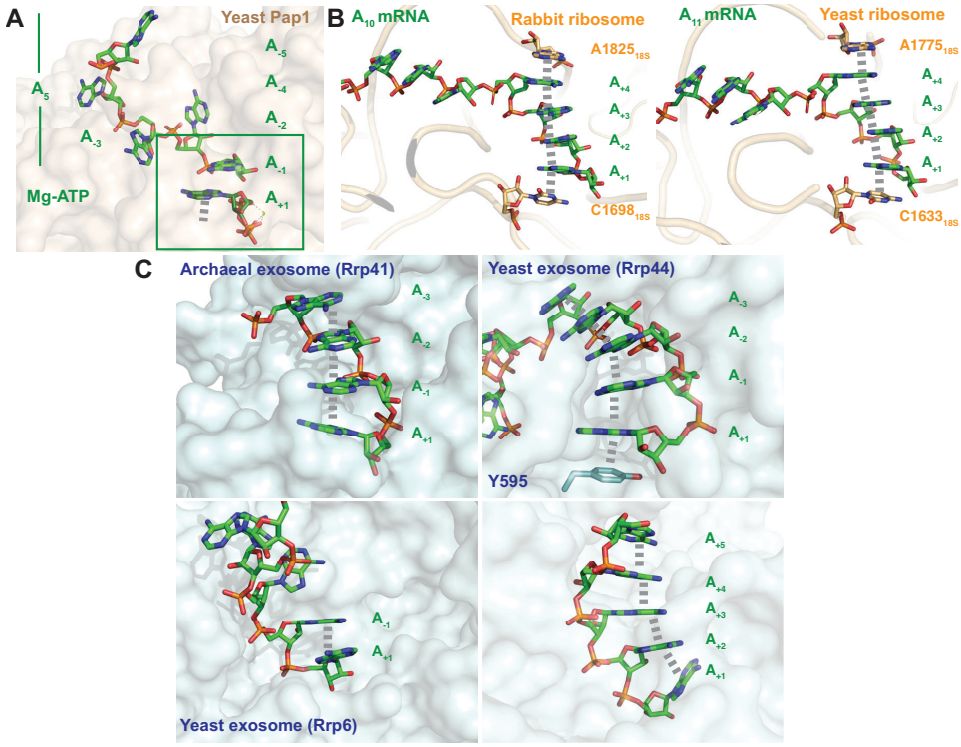


Figure 2.22 Poly(A) structure in other biological processes. π-stacking interactions are grey dashes. (**A**) Structure of the yeast poly(A) polymerase Pap1 (khaki) bound to poly(A) RNA and an incoming ATP (green sticks) (PDB: 2Q66). The 3'-terminal adenine stacks against the ATP adenine. (**B**) Structures of an RNA containing a poly(A) stretch (green sticks) bound to the rabbit ribosome (*left,* orange cartoon; PDB: 6SGC) or yeast ribosome (*right*, orange cartoon; PDB: 6T7T). Oligo(A) RNA is stacked between an adenine and cytidine (orange sticks) of 18S rRNA. (**C**) Structures of the exosome exonucleases Rrp41 (*top left*, blue; PDB: 2C38), Rrp44 (*top right,* blue; PDB: 2VNU), Rrp6 (*bottom left*, blue; PDB: 4OO1), and the exosome-associated helicase Mtr4 (*bottom right*, blue; PDB: 2XGJ) bound to an oligo(A) substrate (green sticks). Oligo(A) adopts a stacked, helical structure in their active sites.

The intrinsic helical conformation of poly(A) is important in its recognition by the Pan2 and Caf1 deadenylases. In addition, it plays a role in translation quality control and polyadenylation, and is observed in exosome exonucleases. Together, the experiments here uncover a novel paradigm of recognising the characteristic structure adopted by single-stranded RNA. The ubiquity of the stacked oligo(A) helix in biology suggests that, because of the unique electrostatic distribution of functional groups in adenine and the resulting helical structure, single-stranded poly(A) has been selected as a mark for a correctly processed mRNA. This mode of recognition also raises the possibility that other structured single-stranded RNAs, such as poly(C), may also be indirectly recognised through their structure. Future studies may yet uncover the importance of the structure of single-stranded RNA in indirect RNA binding.

Table 2.1 X-ray crystallography data collection and refinement statistics.

	Pan2UE (6R9I)	**Pan2UE-A7 (6R9J)**	**Pan2UE-AAGGAA (6R9M)**	**Pan2UE-AAGGA (6R9O)**	**Pan2UE-AAUUAA (6R9P)**	**Pan2UE-AACCAA (6R9Q)**
Data collection						
Space group	*I222*	*I222*	*I222*	*I222*	*I222*	*I222*
Cell dimensions						
a, *b*, *c* (Å)	91.37, 117.07, 255.32	91.48, 116.41, 257.83	91.41, 118.10, 257.28	90.76, 117.59, 256.57	91.88, 117.09, 255.89	92.19, 117.32, 256.66
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	106.41-3.00 (3.08-3.00)	71.95-3.33 (3.42-3.33)	86.17-3.33 (3.42-3.33)	85.50-3.32 (3.41-3.32)	106.45-2.98 (3.06-2.98)	72.49-3.08 (3.16-3.08)
R_{merge}*	0.093 (1.761)	0.095 (1.791)	0.097 (1.527)	0.085 (1.512)	0.065 (1.207)	0.086 (1.733)
$I/\sigma I$*	12.4 (1.3)	8.3 (1.0)	7.3 (1.2)	12.0 (1.0)	13.5 (1.1)	12.7 (1.1)
$CC_{1/2}$*	0.998 (0.774)	0.990 (0.590)	0.983 (0.649)	0.998 (0.624)	0.999 (0.726)	0.917 (0.638)
Completeness (%)*	100.0 (99.9)	99.5 (98.6)	100.0 (99.9)	100.0 (99.8)	100.0 (99.1)	100.0 (100.0)
Redundancy*	6.6 (6.6)	4.6 (4.5)	6.5 (6.0)	6.5 (6.3)	6.6 (6.2)	6.6 (6.3)
Refinement						
Resolution (Å)	3.00	3.33	3.33	3.32	2.98	3.08
Observed reflections	184847	94264	135553	135414	188815	172231
Unique reflections	27902	20507	20875	20744	28662	26217
R_{work} / R_{free}	0.2434/ 0.2895	0.2472/ 0.2982	0.2369/ 0.2953	0.2545/ 0.3092	0.2610/ 0.3086	0.2605/ 0.2970
No. atoms						
Protein	4542	4458	4586	4457	4548	4590
Ligand	/	A_7 RNA: 107	AAGGAA RNA: 87	AAGGA RNA: 87	AAUUAA RNA: 81	AACCAA RNA: 81
B-factors						
Protein	113.69	148.9	150.6	156.5	131.1	126.6
Ligand	/	A_7 RNA: 241.3	AAGGAA RNA: 254.9	AAGGA RNA: 296.6	AAUUAA RNA: 258.6	AACCAA RNA: 275.8
RMSD						
Bond lengths (Å)	0.003	0.001	0.002	0.002	0.004	0.002
Bond angles (°)	0.651	0.419	0.507	0.483	0.804	0.484
Ramachandran Plot						
Favoured (%)	92.4	94.7	95.2	94.8	91.4	94.8
Allowed (%)	7.6	5.3	4.8	5.2	8.6	5.2
Outliers (%)	0.0	0.0	0.0	0.0	0.0	0.0

* indicates collection statistics in the highest resolution shell.

3. Regulation of Pan2-Pan3 Deadenylation

Declaration

Lori Passmore and I conceived the work presented in this chapter. Sections 3.2, 3.12, 3.13, 0 were carried out by Alžběta Roeselová under my supervision. Collection and processing of negative stain EM data of *S. cerevisiae* Pan2-Pan3 was carried out with the help of Michael Webster. Collection and processing of cryo-EM data was carried out with the help of Shabih Shakeel. I performed all other experiments in this chapter.

In this section, "Pan2-Pan3" will refer to the general and yeast poly(A) nuclease complexes, whereas the human complex will be "PAN2-PAN3". The general and yeast cytoplasmic poly(A)-binding proteins will be "Pab1", whereas the human protein will be "PABPC1".

3.1 N-terminal Truncations of Pan3 Do Not Greatly Affect Deadenylation

Since its discovery, the regulation of deadenylation by Pan2-Pan3 have not been widely studied. Structural studies have shown that Pan2-Pan3 forms a compact core complex (Jonas et al., 2014; Schafer et al., 2014; Schafer et al., 2019; Wolf et al., 2014). Thus, deletions within the Pan2-Pan3 core are likely detrimental to protein folding or complex formation. Pan3 also contains an extensive N-terminal low-complexity region, which is dispensable for complex formation but may be required for full activity. Thus, the Pan3 N-terminus was truncated to remove functional regions such as the zinc finger and PAM2 motif (Section 1.2.2, Figure 3.1A). Exclusion of PAM2 also allowed its requirement in Pab1-dependent deadenylation to be tested.

Recombinantly expressed and purified Pan2-Pan3 was tested for activity on $X_{20}A_{30}$ RNA. *S. cerevisiae* Pan2-Pan3 distributively removed poly(A) until it encountered the non-A region (Figure 3.1B, D). In agreement with previous studies and the lack of tandem guanosines in the UTR-like sequence (Chapter 2), Pan2-Pan3 readily degraded non-A sequences after a brief stall. Interestingly, Pan2-Pan3 stalled 2-3 nucleotides before it encountered the non-A sequence, and 2-3 nucleotides after the poly(A)-non(A) junction. These transient stalls are likely the result of the helical ordering required of non-A sequences as they enter the active site (Chapter 2).

Consistent with a previous study (Wolf et al., 2014), deletion of the Pan3 zinc finger (ΔZnF) resulted in slightly lower activity (Figure 3.1C, compare 8, 12 min). As the zinc finger domain specifically binds poly(A), its deletion likely resulted in lower overall affinity of Pan2-Pan3 for poly(A) and thus lower activity. Further truncation of the low-complexity region (PKC) did not cause deadenylation to be significantly slower (Figure 3.1D-E, compare 30, 35 min time points). N-terminal truncations of Pan3 did not affect the stall pattern as Pan2-Pan3 reaches the non-A sequence, suggesting the stall pattern is independent of the Pan3 N-terminus.

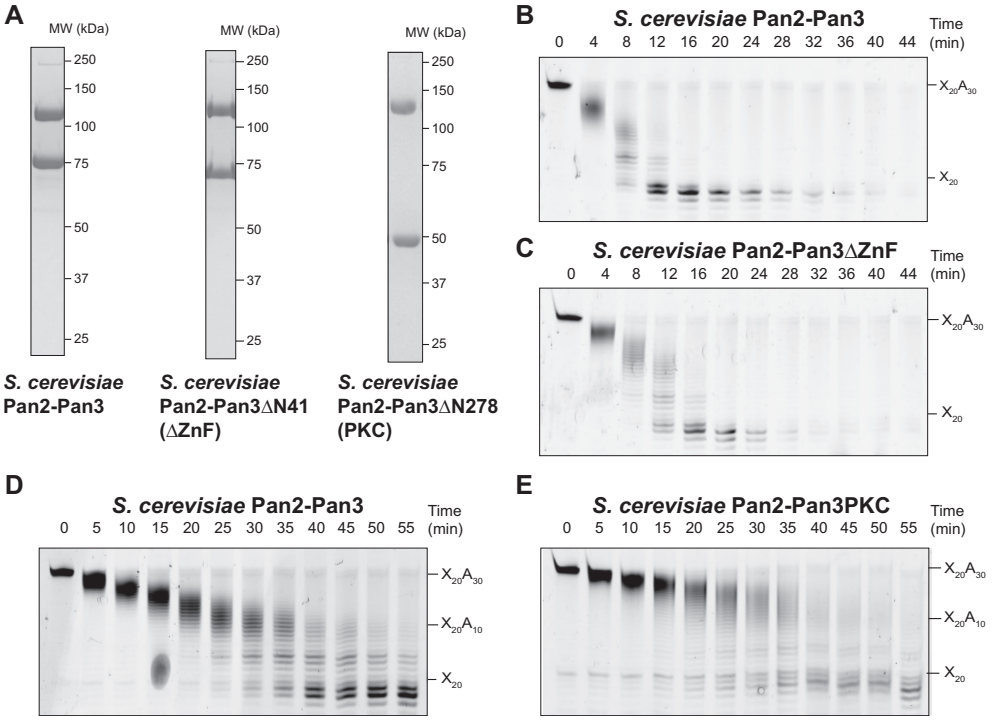


Figure 3.1 N-terminal Pan3 truncations do not greatly affect Pan2-Pan3 deadenylation. (**A**) SDS PAGE of recombinantly expressed and purified *S. cerevisiae* Pan2-Pan3 with N-terminal truncation variants. (**B-E**) Deadenylation assays on 200 nM 5' 6-FAM-labelled $X_{20}A_{30}$ RNA with: (**B**) 100 nM *S. cerevisiae* Pan2-Pan3; (**C**) 100 nM *S. cerevisiae* Pan2-Pan3ΔZnF; (**D**) 50 nM *S. cerevisiae* Pan2-Pan3; and (**E**) 50 nM *S. cerevisiae* Pan2-Pan3PKC. The discrepancy in activity between *S. cerevisiae* Pan2-Pan3 in (**B**) and (**D**) is the result of using 2× lower Pan2-Pan3 concentration in (**D**).

To test if these features are conserved in *H. sapiens* PAN2-PAN3, the N-terminus of human PAN3 was similarly truncated. Complexes were recombinantly expressed by baculovirus-mediated Sf9 insect cell overexpression and purified. However, *H. sapiens* PAN3 was prone to degradation during purification, likely due to the extensive low-complexity region (Section 1.2.2). Therefore, PAN3 was cloned to contain a bulky N-terminal GST tag and I developed a new purification protocol (Figure 3.2A). This notably improved the final yield of human PAN2-PAN3 complexes (Figure 3.2B), which were tested for activity on $X_{20}A_{60}$ RNA. Characteristics of yeast Pan2-Pan3, including stall patterns at the poly(A)-non(A) junction (Figure 3.2C), as well as slightly lower activity upon deletion of the zinc finger or the N-terminal low-complexity region (Figure 3.2D-E), were conserved for this RNA substrate. Unexpectedly, full-length PAN2-PAN3 deadenylated distributively, whereas N-terminal

truncations, which are expected to have lower affinity for poly(A), were more processive, as fully deadenylated product appeared when intact RNA was still present (Figure 3.2D-E, e.g. 21, 24 min). The mechanism of the higher processivity of N-terminal truncations is unknown. Together, these experiments show that Pan3 N-terminal truncations result in slightly slower deadenylation activity compared to wild-type, likely attributable to the deletion of the poly(A)-specific ZnF domain. Patterns of deadenylation upon encountering the junction between poly(A) and the upstream sequence are also conserved between yeast and human complexes, suggesting that Pan2-Pan3 behaviour is dependent on the upstream non-A RNA sequence.

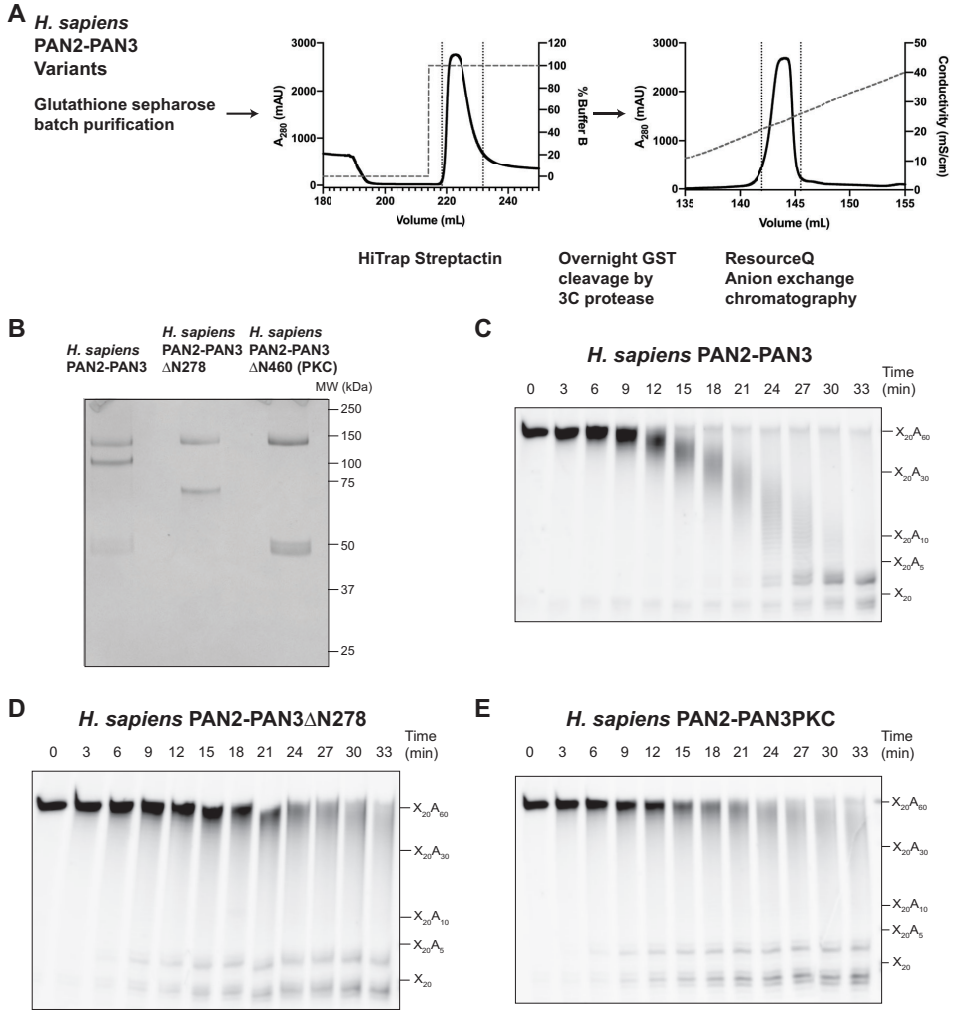


Figure 3.2 N-terminal PAN3 truncations do not greatly affect PAN2-PAN3 deadenylation. (**A**) Purification protocol of *H. sapiens* PAN2-PAN3 and variants thereof. Peak fractions are outlined with a dotted line. (**B**) SDS polyacrylamide gels of recombinantly expressed and purified *H. sapiens* PAN2-PAN3 N-terminal truncations. (**C-E**) Deadenylation assays on 200 nM 5' 6-FAM-labelled $X_{20}A_{60}$ RNA with 100 nM: (**C**) *H. sapiens* PAN2-PAN3; (**D**) *H. sapiens* PAN2-PAN3ΔN278; and (**E**) *H. sapiens* PAN2-PAN3PKC.

3.2 Upstream Sequence Specificity of Pan2-Pan3

Previous studies had suggested that yeast Pan2-Pan3 deadenylates model transcripts at different rates *in vitro* depending on the 3' UTR sequence (Lowell et al., 1992). However, evidence for differential activities was limited by the long 3' UTRs and thus difficulty of isolating sequence features which may give rise to differential deadenylation. Furthermore, previous studies did not use highly purified Pan2-Pan3. Interpretation of results were thus complicated by other components which may have been co-purified.

To determine if yeast Pan2-Pan3 displays sequence preference, its activity was tested on RNAs containing varying upstream sequences, including some which form hairpins of different stabilities (Supplementary Figure 7). These RNAs contain 30-A tails and UTRs of similar lengths; therefore, any difference in activity will be solely due to the UTR sequence.

These *in vitro* experiments demonstrated that yeast Pan2-Pan3 displayed different activities depending on the upstream sequence (Figure 3.3A-C, E). For example, the A3-A_{30} RNA containing the A-rich upstream sequence (Figure 3.3B) was deadenylated more rapidly than less A-rich P2-A_{30} and Z2-A_{30} RNAs (Figure 3.3C, E). Upstream hairpins appeared to slow deadenylation, as H2-A_{30}, which contains a thermodynamically stable hairpin, was deadenylated more slowly than single-stranded RNAs (Figure 3.3F). Intriguingly, V2-A_{30} was strongly resistant to deadenylation (Figure 3.3D); however, it is unclear why the V2- sequence would cause significantly slower deadenylation by Pan2-Pan3. Moreover, the pattern of stalling at the poly(A)-non(A) junction differed, suggesting that the stall pattern is dependent on the 3' UTR sequence. Thus, Pan2-Pan3 may preferentially act on RNAs with a predominantly single-stranded upstream 3' UTR, with a further preference for A-rich sequences. This could be due to the Pan3 zinc finger domain, which preferentially binds single-stranded oligo(A) sequences, thereby increasing the overall affinity of the complex for single-stranded, A-rich RNAs.

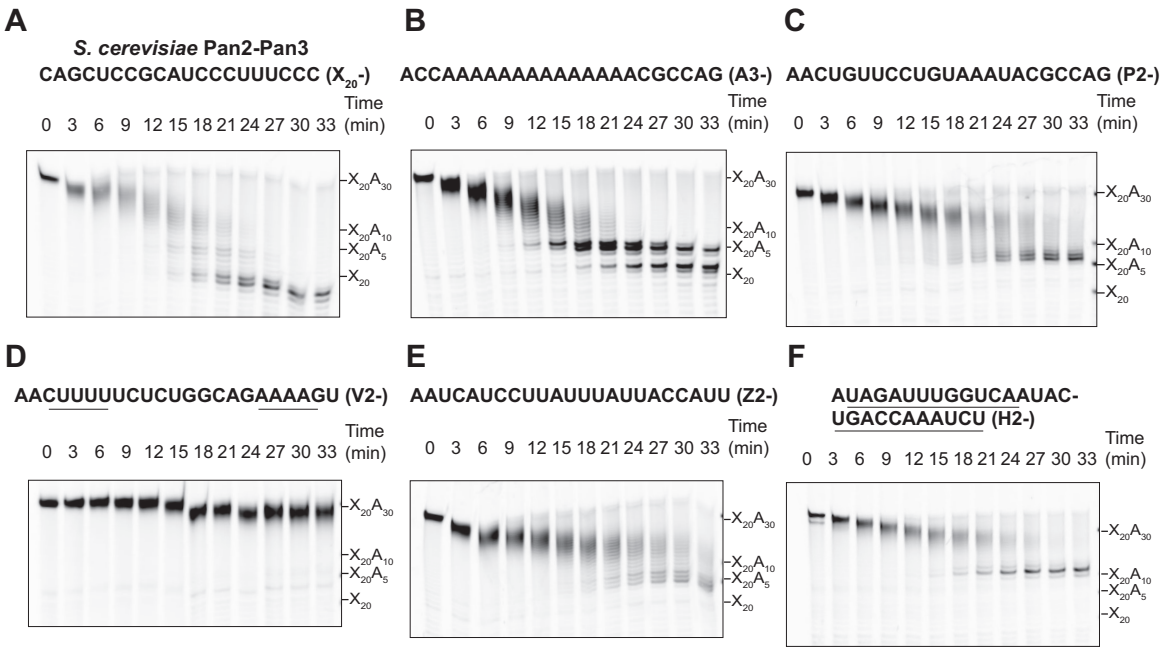


Figure 3.3 *S. cerevisiae* Pan2-Pan3 displays preference for UTR sequences. (**A-F**) Deadenylation assays by 50 nM *S. cerevisiae* Pan2-Pan3 on 200 nM (**A**) X_{20}-A_{30}; (**B**) A3-A_{30}; (**C**) P2-A_{30}; (**D**) V2-A_{30}; (**E**) Z2-A_{30}; and (**F**) H2-A_{30}. All RNAs were 5' 6-FAM-labelled; upstream UTR sequences are shown above the figure panels. Regions which are predicted to form complementary hairpins are underlined.

To determine if *H. sapiens* PAN2-PAN3 displays similar sequence preference, PAN2-PAN3ΔN278 activity was tested on the same RNAs. Similar to S. cerevisiae *Pan2-Pan3*, *H. sapiens* PAN2-PAN3 deadenylated RNAs with single-stranded UTR sequences most rapidly (Figure 3.4A-C, E). The A-rich A3-A_{30} was also deadenylated relatively rapidly by the human complex (Figure 3.4B), demonstrating a conserved preference of Pan2-Pan3 for A-rich UTR sequences. In contrast with yeast Pan2-Pan3, however, Z2-A_{30} appeared to be the preferred substrate and was fully deadenylated within 12 minutes (Figure 3.4E). Consistent with the yeast complex, non-A sequences containing hairpins were deadenylated more slowly (Figure 3.4D, F), and V2-A_{30} was deadenylated most slowly (Figure 3.4D).

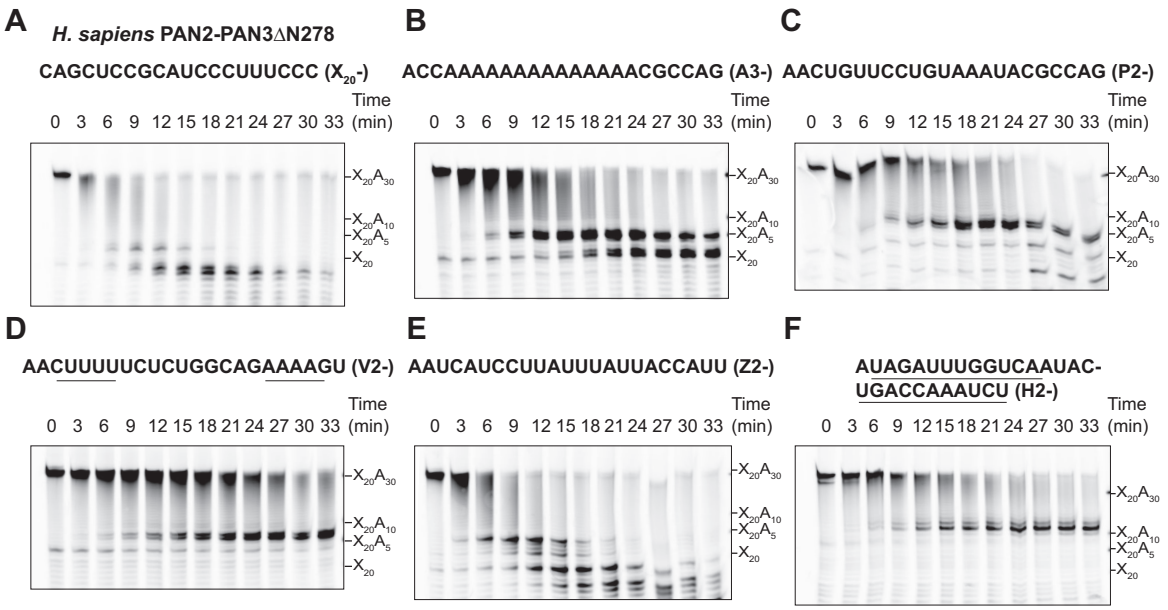


Figure 3.4 *H. sapiens* PAN2-PAN3 has similar sequence preference to yeast Pan2-Pan3. (**A-F**) Deadenylation assays by 100 nM *H. sapiens* PAN2-PAN3ΔN278 on 200 nM (**A**) X_{20}-A_{30}; (**B**) A3-A_{30}; (**C**) P2-A_{30}; (**D**) V2-A_{30}; (**E**) Z2-A_{30}; and (**F**) H2-A_{30}. All RNAs were 5' 6-FAM-labelled; upstream UTR sequences are shown above the figure panels. Regions which are predicted to form complementary hairpins are underlined.

These results confirm that Pan2-Pan3 has substrate preference which is dependent on the non-A sequence directly upstream of the poly(A) tail. Certain sequence preferences, such as for A-rich single-stranded upstream RNA, appear conserved between yeast and human Pan2-Pan3 complexes. Furthermore, as the PAN2-PAN3ΔN278 construct lacks the PAN3 N-terminal zinc finger, these sequence preferences cannot be explained by binding by the poly(A)-specific zinc finger domain alone. The relative domain contributions to the observed intrinsic sequence preference would provide insight into the relative transcript preference of Pan2-Pan3 complexes in yeast and humans. Finally, it remains unknown whether these *in vitro* sequence preferences correspond to the 3' UTR sequences of transcripts whose decay rates are more dependent on Pan2-Pan3 activity *in vivo* (Sun et al., 2013). Further investigations of the intrinsic sequence specificity of Pan2-Pan3 could partially explain the observed transcript dependence.

3.3 Pab1 Binds Poly(A) with High Affinity

3.3.1 Purification of Poly(A)-Binding Protein

In the cell, cytoplasmic poly(A) is thought to be coated with Pab1, which can regulate deadenylation by Pan2-Pan3 (Section 1.4.1.3). To study the precise effects of Pab1 on Pan2-Pan3, I developed a purification protocol for full-length yeast and human Pab1. A Pab1 variant containing the N-terminal RNA recognition motif (RRM) domains, but not the proline-rich linker (P-linker) or CTD, was also purified using the same protocol (Section 1.4.1). This construct was designed to test the effect of the Pab1 C-terminus on RNA binding and Pan2-Pan3 activity. In light of the recent study showing that the helix immediately following RRM4 may be involved in Pab1 multimerisation (Schafer et al., 2019), an additional construct of human PABPC1 was cloned, expressed, and purified. This construct contains the long α-helix following RRM4 as determined by secondary structural prediction (Supplementary Figure 6), and would allow the role of this helix in PAN2-PAN3 interaction to be determined.

Yeast Pab1 was recombinantly expressed in *E. coli* with an N-terminal GST-tag which was cleavable with 3C protease. Purification was performed in four steps, including overnight 3C cleavage of the GST tag (Figure 3.5A). I further developed an overexpression and purification protocol for *H. sapiens* PABPC1. 3C proteases from picorna- and caliciviruses have been known to cleave the PABPC1 P-linker (Kuyumcu-Martinez et al., 2004; Rivera and Lloyd, 2008; Zhang et al., 2007). Thus, PABPC1 was cloned into a pET28 vector with a C-terminal SII tag with a TEV cleavage site. Human PABPC1 was recombinantly expressed in *E. coli*. Purification was performed similarly to yeast Pab1, except with a cation exchange chromatography step instead of anion exchange chromatography (Figure 3.5B). This was required due to the different predicted isoelectric points of *S. cerevisiae* Pab1 (pI: 5.71) and *H. sapiens* PABPC1 (pI: 9.43). The SII tag was cleaved overnight by TEV protease.

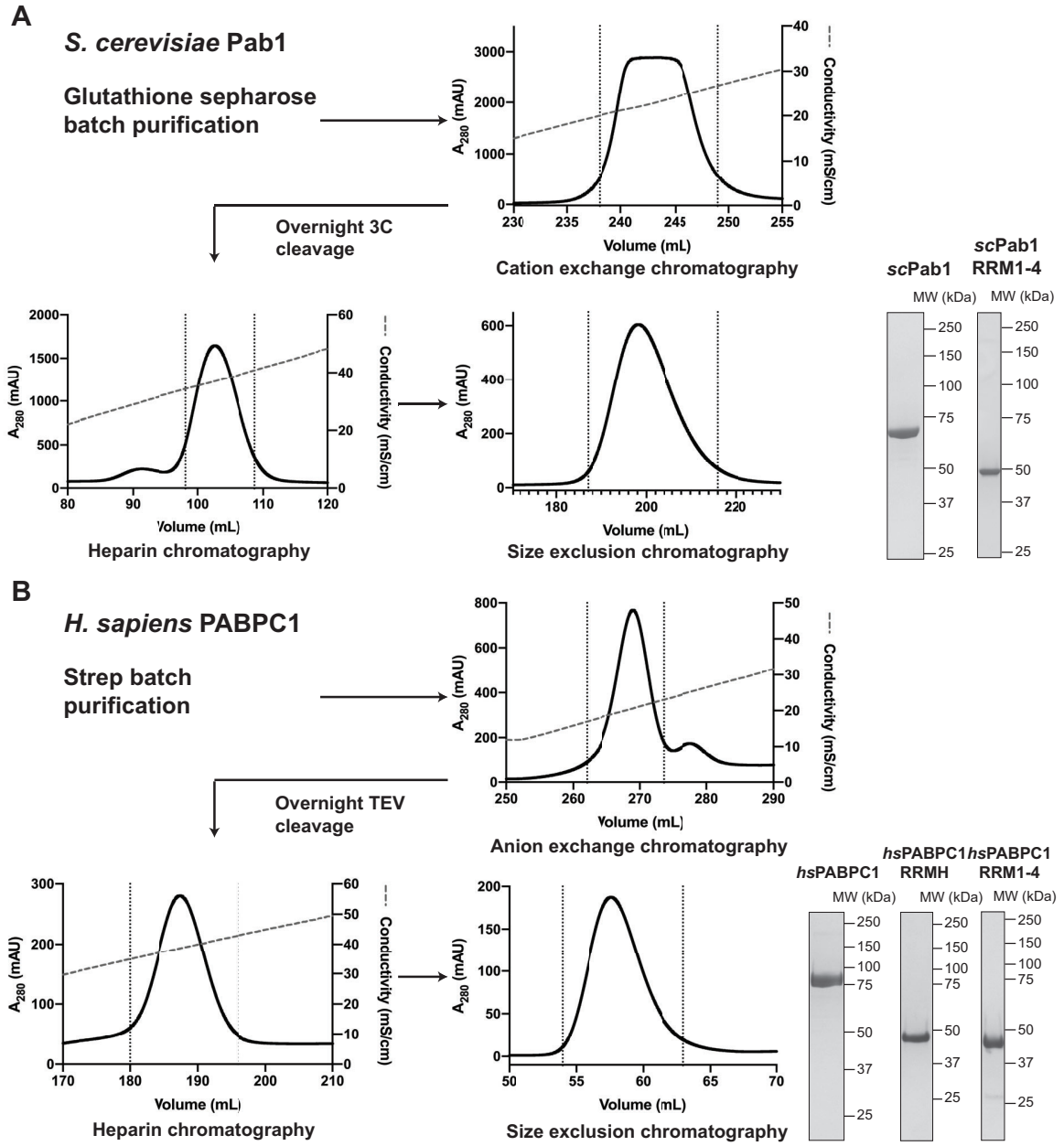


Figure 3.5 Purification of recombinantly expressed Pab1. Purification of (**A**) Pab1 and (**B**) PABPC1. SDS polyacrylamide gels of the constructs show their homogeneity following purification.

3.3.2 Pab1 Binds Oligo(A) with Nanomolar Affinity

To verify that recombinantly expressed and purified Pab1 binds poly(A) *in vitro*, fluorescence polarisation assays (FPAs) were performed. Pab1 and PABPC1 were titrated against a fixed concentration of 6-FAM-labelled oligo(A) of different lengths to assess their affinities (Figure 3.6A-B). For all oligo(A) RNAs, affinity increased with oligonucleotide length. Consistent with previous observations, both *S. cerevisiae* and *H. sapiens* Pab1 bound oligo(A) with high affinity,

with dissociation constants (K_d, Figure 3.6A-B) in the nanomolar range. Interestingly, human PABPC1 bound poly(A) more strongly than yeast Pab1 (compare A_{12}: ~6×; A_{10}: ~7×; A_7: ~50×), particularly for shorter oligo(A) RNA. The affinity estimates of human PABPC1 for poly(A) are suboptimal due to the lack of data points near the inflection point of the binding curves; future studies may measure fluorescence polarisation in this narrower concentration range to provide better estimates of binding affinity. These results suggest that there may be different requirements for strong oligo(A) binding. For example, yeast Pab1 may require two RRM domains for high-affinity interactions (thus nanomolar affinity is not reached until the oligo(A) > 10), whereas human PABPC1 may require only one (high affinity is reached when oligo(A) = 7). These results show that there may be species-specific differences in poly(A) binding affinity, though the mechanistic basis and physiological importance of these affinity differences are unknown. Further understanding of how Pab1 and PABPC1 differ in their mechanism of poly(A) binding could provide insight into these binding differences.

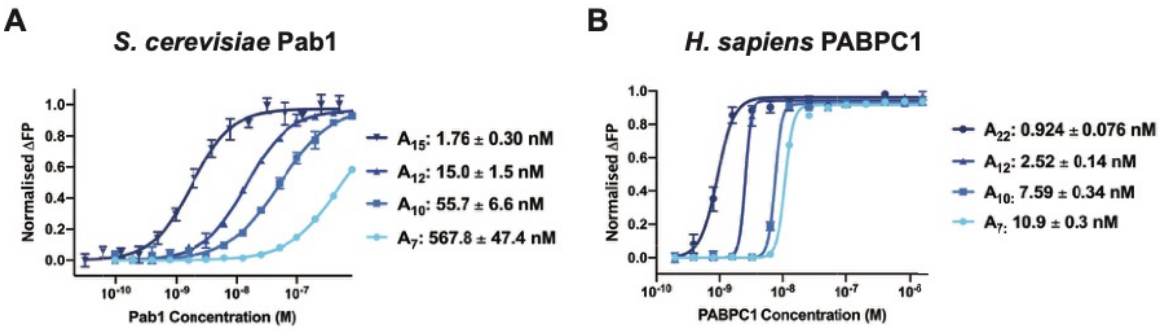


Figure 3.6 Pab1 binds poly(A) with nanomolar affinity. Binding affinity of (**A**) *S. cerevisiae* Pab1 and (**B**) *H. sapiens* PABPC1 for 0.05 nM oligo(A) RNAs by fluorescence polarisation assays. Oligo(A) RNAs are coloured from light to dark by oligonucleotide length. Estimated K_d values are shown with 95% confidence intervals next to the respective RNAs. Because saturation and minimum binding are not reached for Pab1-A_7 and PABPC1-A_{22} binding curves, K_d estimates are unlikely to be accurate. Assays were carried out in triplicate, points are the mean polarisation, and error bars are standard deviation.

3.4 PABPC1 Binds Poly(A) Cooperatively

The dramatically different slopes at the inflection points of FP binding curves suggest that Pab1 and PABPC1 bind oligo(A) RNAs with different cooperativities (Figure 3.6). To confirm the different cooperativities, Hill coefficients of RNA binding were estimated from curve-fitting parameters (Figure 3.7A). Yeast Pab1 was non-cooperative ($h \sim 1$) or slightly positively cooperative ($h > 1$) on longer poly(A). In contrast, human PABPC1 was very positively cooperative ($h >> 1$) on all tested RNAs. Thus, human PABPC1 and yeast Pab1 show different cooperativity upon RNA binding. The high positive cooperativity of PABPC1 suggests that it has a greater tendency to multimerise on a poly(A) tail, likely originating from intermolecular PABPC1-PABPC1 interactions. The Hill coefficients are likely to be higher with increased poly(A) tail length as each oligo(A) is more likely to accommodate more than one Pab1.

To validate these results, electrophoretic mobility shift assays (EMSAs) were carried out under the same buffer conditions as in deadenylation assays, titrating yeast and human Pab1 against fluorescently-labelled $X_{20}A_{30}$. EMSAs enabled multimeric (1×, 2×, 3×+) Pab1 to be resolved under native conditions. Consistent with FPAs, *S. cerevisiae* Pab1 was minimally cooperative when binding $X_{20}A_{30}$ (Figure 3.7B). 2× Pab1-bound species only appeared when RNA was saturated with 1× Pab1; that is, Pab1 is not more likely to bind a 30-A poly(A) tail already bound by Pab1. In comparison, *H. sapiens* PABPC1 showed slight positive cooperativity on the same substrate (Figure 3.7C), as 2× PABPC1-bound species appeared even when unbound RNA was present. This suggests that, despite A_{30} only being able to accommodate ~1 PABPC1 molecule, PABPC1 prefers to bind PABPC1-poly(A).

As PABPC1 showed positive cooperativity, I repeated EMSAs using a longer RNA which can accommodate more than one PABPC1 to study the domain requirements for cooperativity. In agreement with FPA assays, *H. sapiens* PABPC1 bound $X_{20}A_{60}$ cooperatively, as multimeric (2×,

3×+) species were observed at the lowest tested protein concentration (Figure 3.7D). As the protein concentration was increased, PABPC1 populated higher multimeric states, concomitant with disappearance of the 1× PABPC1-RNA band. Thus, full-length PABPC1 binds RNA with positive cooperativity. To test the effect of the P-linker and CTD on cooperativity, PABPC1 containing only the RRM domains was assayed for RNA binding as above (Figure 3.7E). $X_{20}A_{60}$ RNA binding by PABPC1 RRM1-4 was less cooperative and more independent. For example, at a pre-determined concentration of 1× protein-RNA (light blue arrow), PABPC1 RRM1-4 showed equal distribution between 0×, 1×, and 2×, whereas full-length PABPC1 had mainly 0× and 2× species. Similarly, when protein is 2× RNA (dark blue arrow), PABPC1 RRM1-4 was equally distributed between 1×, 2×, and 3×, whereas full-length PABPC1 had 2× and 3× species. The PABPC1 C-terminus is thus important in cooperative poly(A) binding, consistent with previous studies implicating the P-linker in cooperativity (Lin et al., 2012; Melo et al., 2003; Yao et al., 2007). Interestingly, these results suggest that, at non-saturating PABPC1 concentrations and under conditions used for deadenylation assays, a mixture of different PABPC1 stoichiometries bound to longer poly(A) tails would be observed.

While the results suggest that human PABPC1 is more cooperative than yeast Pab1, interpretation is complicated by the lack of experiments with yeast Pab1 RRM1-4. Experiments should hence be repeated with equivalent yeast Pab1 truncations. Moreover, EMSAs of PABPC1 RRM1-4 should be corroborated by Hill coefficients from FPAs. Further truncations of the P-linker could also identify the region required for cooperativity. Finally, comparisons of the precise mechanism of multimerisation may allow species-specific differences to be elucidated, allowing the different cooperativities to be rationalised on a molecular basis.

A

Hill coefficients (From Figure 3.10)

	S. cerevisiae Pab1	*H. sapiens* PABPC1
A_7	1.04	6.58
A_{10}	1.02	8.02
A_{12}	1.22	9.31
A_{15}	1.32	/
A_{22}	/	3.29

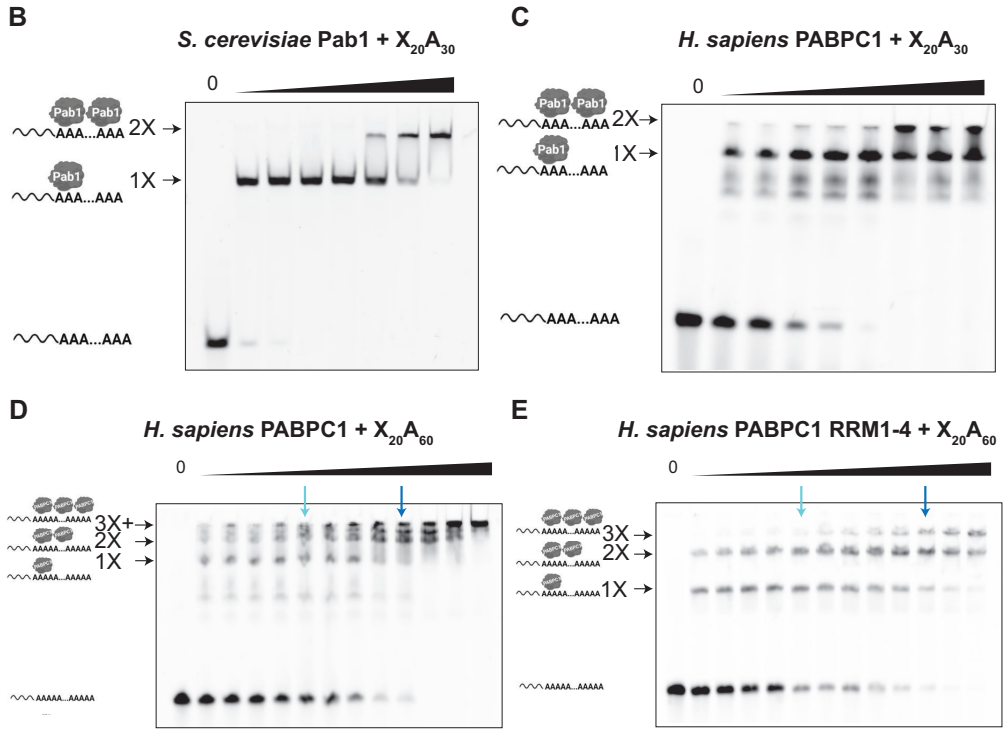


Figure 3.7 The C-terminal half of PABPC1 aids cooperative RNA binding. (**A**) Hill coefficients obtained from curve fitting parameters from Figure 3.6. EMSAs of binding to 200 nM 6-FAM-labelled $X_{20}A_{30}$ by (**B**) *S. cerevisiae* Pab1 (concentration: 200-500 nM); and (**C**) *H. sapiens* PABPC1 (concentration: 150-450 nM). EMSAs of (**D**) *H. sapiens* PABPC1 and (**E**) PABPC1 RRM1-4 binding to 200 nM 6-FAM-labelled $X_{20}A_{60}$. Protein concentrations range from 200-1000 nM. Schematics show putative Pab1/PABPC1 multimeric states. Light blue: 350 nM PABPC1; dark blue: 700 nM PABPC1.

3.5 Pab1 Stimulation of Pan2-Pan3 Depends on Poly(A) Tail Length

In the proposed biphasic deadenylation model (Section 1.5.2), Pan2-Pan3 is stimulated by poly(A)-binding protein until Pab1 dissociates from the poly(A) tail. As seen above, Pab1 had high affinity for short stretches of poly(A) (Section 3.3.2), suggesting that Pan2-Pan3 would be stimulated by Pab1 even on short poly(A) tails. This hypothesis is inconsistent with the biphasic model, where Pan2-Pan3 is thought to only remove the distal poly(A) tail (Yamashita et al., 2005). Having purified homogeneous yeast and human Pab1 and characterised their RNA binding behaviour, I aimed to obtain an *in vitro* reconstituted system of Pan2-Pan3 deadenylation of a Pab1-bound poly(A) tail to clarify the model of Pan2-Pan3 stimulation.

Pan2-Pan3 deadenylation of $X_{20}A_{30}$ (200 nM) was tested in the presence or absence of 1× Pab1. As determined by EMSAs, a slight excess (250 nM) of Pab1 was required for 1× binding of RNA. This slightly higher amount may be the result of a proportion of inactive purified Pab1 or Pab1 dissociation on native gels (due to salt concentration or application of an electric field). In *in vitro* deadenylation assays, buffers were altered so that buffer and salt concentrations were constant when an additional component was added. Surprisingly, overall, the presence of Pab1 inhibited Pan2-Pan3 activity on an A_{30} poly(A) tail (Figure 3.8A-B). While deadenylation was mostly complete by 20 min in the absence of Pab1, Pan2-Pan3 was strongly stalled at ~10 A's when Pab1 was present. This result was surprising as Pab1 was thought to stimulate Pan2-Pan3. Closer examination of deadenylation with Pab1 revealed distinct tail lengths at which Pan2-Pan3 activity was stalled (Figure 3.8B). Two stall ranges (17-21 A's; 8-11 A's) were observed; between these regions, deadenylation by Pan2-Pan3 proceeded rapidly. The difference between the two stalls corresponds roughly to the binding footprint of one RRM domain, suggesting that Pan2-Pan3 may be stalled when it encounters a Pab1 RRM. As RRM1-

2 have the highest affinity for poly(A), it is likely that Pan2-Pan3 is stalled by RRM1 and Pab1 slides (or peels off and rebinds) along the poly(A) RNA.

Because this result contrasts with previous studies showing that Pan2-Pan3 is stimulated by Pab1, an RNA with a longer A_{60} tail was used in deadenylation assays. This substrate more closely models the length of a newly-exported transcript as yeast mRNAs are initially synthesised with a 60-70-A tail (Brown and Sachs, 1998a). Similar to the A_{30} substrate, Pan2-Pan3 alone deadenylated distributively and at a constant rate (Figure 3.8C-D). However, in contrast with $X_{20}A_{30}$, Pan2-Pan3 displayed biphasic behaviour on a longer poly(A) tail when Pab1 was present. By 8 min, 30 A's were removed in the presence of Pab1, but only 10-20 A's had been removed when Pab1 is absent. Thus, the initial phase of deadenylation is accelerated in the presence of Pab1. To more closely observe this initial phase, deadenylation assays were repeated over a shorter time course. Relative to when Pab1 is absent, deadenylation was stimulated in the initial phase by Pab1. Pan2-Pan3 was then stalled at ~30 A's and deadenylation proceeded more slowly compared to when Pab1 was present (Figure 3.8E-F). Therefore, Pab1 stimulates Pan2-Pan3 in an *in vitro* reconstituted system, but this stimulation depends on poly(A) tail length. Deadenylation is only stimulated until 30 A's; subsequently, Pan2-Pan3 is inhibited. This result is similar to experiments carried out in a recent study, showing that the initial phase of deadenylation of poly(A)-Pab1 to ~30 A's is rapid, but is subsequently slowed (Schafer et al., 2019).

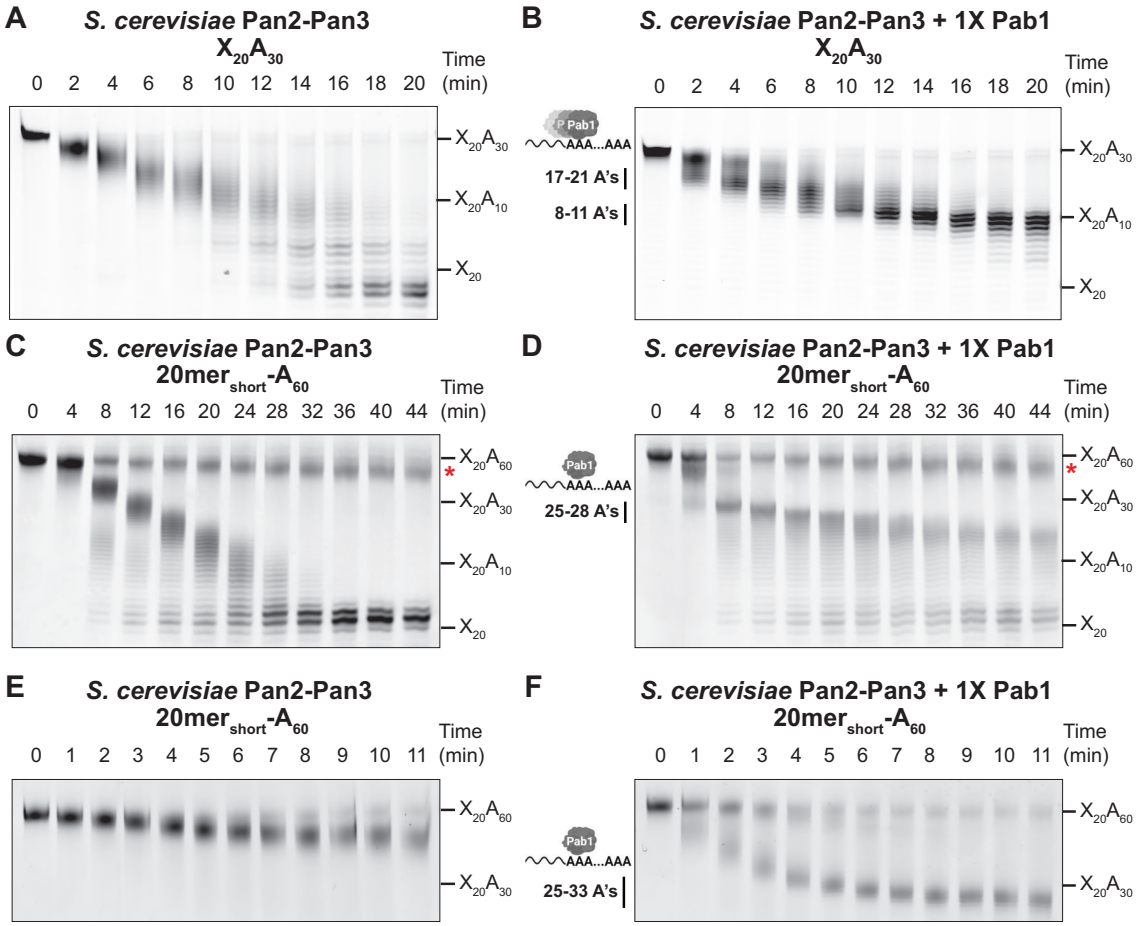


Figure 3.8 Pab1 stimulation of Pan2-Pan3 depends on poly(A) tail length. (**A-B**) Deadenylation assays of 200 nM 6-FAM-labelled $X_{20}A_{30}$ by 30 nM *S. cerevisiae* Pan2-Pan3 with (**A**) no Pab1 and (**B**) 1× Pab1 (250 nM). A model of Pab1 sliding/peeling shows how individual RRMs may determine stall points. (**C-F**) Deadenylation assays of 200 nM unlabelled $20mer_{short}$-A_{60} by 50 nM *S. cerevisiae* Pan2-Pan3 with (**C, E**) no Pab1 and (**D, F**) 1× Pab1 (250 nM). Panels **E-F** are a shorter time course to show Pab1 stimulation of deadenylation at early time points. Stall regions are indicated by lines. The red asterisk is likely a contaminating RNA from *in vitro* transcription.

To test if tail length-dependent stimulation by PABPC1 is conserved for human PAN2-PAN3, assays were repeated with PAN2-PAN3ΔN278 and 1× PABPC1. PAN3ΔN278 retains the PAM2 motif and should interact with the PABPC1 CTD. 350 nM PABPC1 was required to achieve 1× binding to 200 nM A_{30} on native EMSAs, possibly due to the higher pI of PABPC1, which may affect RNP behaviour when an electric field is applied during EMSAs. Similar to yeast Pan2-Pan3, deadenylation of an A_{30} tail by PAN2-PAN3 was inhibited by 1× PABPC1 relative to no PABPC1 (Figure 3.9A-B). In contrast with yeast proteins, no stall points corresponding to RRMs were observed; this may be elucidated with longer time points.

Deadenylation assays were repeated with $X_{20}A_{60}$ RNA with no or 1× PABPC1. Without PABPC1, similar processive deadenylation was observed compared to the A_{30} substrate. In the presence of 1× PABPC1, PAN2-PAN3 deadenylation was initially stimulated (compare 6-12 min time points), and inhibition by PABPC1 again occurred at 30 A's, similar to the PABPC1 footprint (Figure 3.9C-D). Interpretation of deadenylation of $X_{20}A_{60}$ was complicated by cooperative binding of PABPC1 on RNA, resulting in 0×, 1× and 2× PABPC1-bound RNAs at a given PABPC1 concentration (Figure 3.7D, light blue arrow). These species could display distinct behaviours in deadenylation (Figure 3.9D). Unbound RNA was deadenylated at a similar rate to when no PABPC1 was added. 1× PABPC1-poly(A) was initially stimulated until a tail length of 30 A's, corresponding to the footprint of one PABPC1, is reached. 2× PABPC1-poly(A) was not stimulated and was stalled at ~60 A's. Nonetheless, further experiments to isolate different PABPC1-bound species are likely required to confirm this hypothesis.

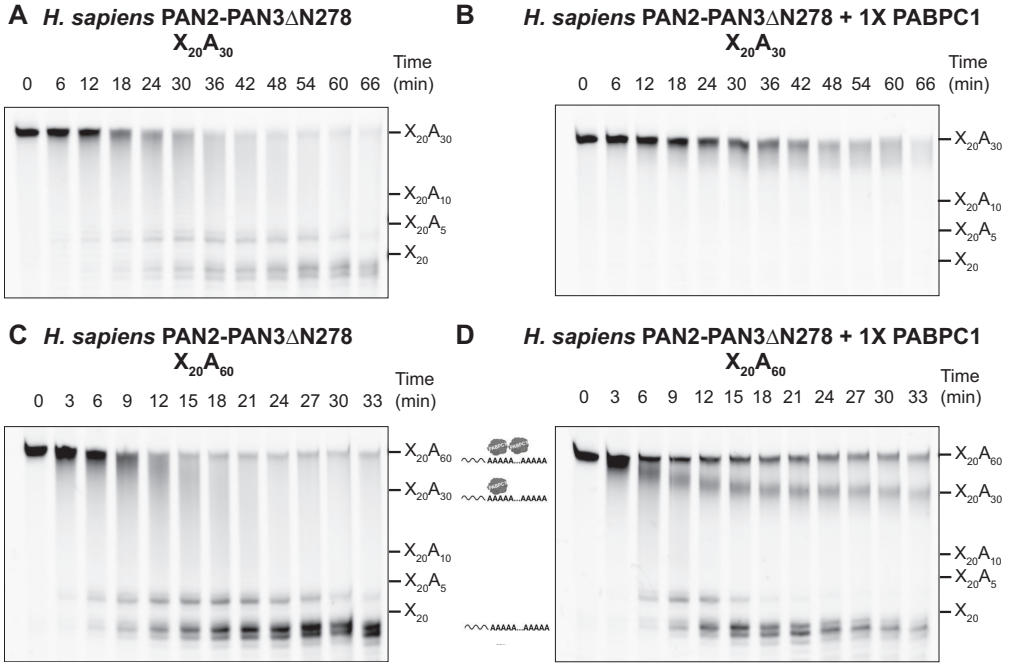


Figure 3.9 PABPC1 stimulation of PAN2-PAN3 depends on poly(A) tail length. (**A-B**) Deadenylation assays of 200 nM 6-FAM-labelled $X_{20}A_{30}$ by 200 nM *H. sapiens* PAN2-PAN3ΔN278 with (**A**) no PABPC1 and (**B**) 1× PABPC1 (350 nM). (**C-D**) Deadenylation assays of 200 nM 6-FAM-labelled $X_{20}A_{60}$ by 200 nM *H. sapiens* PAN2-PAN3ΔN278 with (**C**) no PABPC1 and (**D**) 1× PABPC1 (350 nM).

Our results show that Pab1 stimulation of yeast and human Pan2-Pan3 depends on the poly(A) tail length. *In vitro*, Pab1 stimulates Pan2-Pan3 to a tail length of ~30 A's. When Pab1 is bound to a poly(A) tail of 30 A's or fewer, Pan2-Pan3 is inhibited. Interestingly, these results recapitulate the biphasic model, which suggests that initial deadenylation is carried out by Pan2-Pan3 *in vivo*. As this stall tail length is similar to the footprint of Pab1, Pan2-Pan3 may be inhibited upon encountering the RRM domains of Pab1. These results are in contrast with previous studies from our laboratory, which had shown that Pan2-Pan3 was stimulated by Pab1 on an A_{80} substrate with the CYC1 3' UTR (Wolf et al., 2014). However, as the 3' UTR sequence is different, the effects of the non-A sequence on Pab1-dependent stimulation of Pan2-Pan3 cannot be ruled out.

3.6 Stimulation of PAN2-PAN3 Depends on PABPC1 Stoichiometry

The observation that different PABPC1-poly(A) ratios may undergo different deadenylation suggests that Pab1 stoichiometry may influence poly(A) tail shortening. To test this, yeast Pan2-Pan3 was incubated with 0×, 1×, or 2× Pab1 on short A_{30} poly(A) tails (Figure 3.10A-C). As previously, 1× Pab1 inhibited deadenylation and caused strong stalling at approximately 10 A's (Figure 3.10B). Additional Pab1 did not change the inhibitory nature of Pab1; however, the stall regions at ~A_{10} and ~A_{18} became stronger (Figure 3.10B-C). Because at the concentrations used, A_{30} is unlikely to support simultaneous binding of two Pab1 molecules, the higher concentration of Pab1 likely allows more rapid rebinding by Pab1, preventing opportunistic poly(A) removal by Pan2-Pan3. If this were true, it suggests that Pan2-Pan3 cannot actively remove poly(A) bound by Pab1; instead, it can only opportunistically remove unbound poly(A) when Pab1 dissociates. This model is roughly consistent with the $t_{1/2}$ of 16 min for *S. pombe* Pab1 on an A_{10} tail (Webster et al., 2018); approximately half of $X_{20}A_{10}$ was deadenylated in 16 min in the presence of 1× Pab1 (compare 20 and 36 min).

Assays were repeated with 0×, 1×, 2×, and 4× Pab1 on A_{60} substrates (Figure 3.10D-G). In the presence of Pab1, Pan2-Pan3 was initially stimulated (0-8 min for each Pab1 concentration), and subsequently inhibited at A_{30}. Between 30-0 A's, stronger inhibition was observed with higher Pab1 concentrations, as estimated by the higher intensity of the stalled band and slower appearance of full deadenylated product compared to 1× Pab1. This stronger stalling is consistent with the above results. Nonetheless, the similar degree of initial stimulation with higher Pab1 concentrations was surprising. At superstoichiometric Pab1 concentrations, the poly(A) tail would be expected to be coated with Pab1; if Pan2-Pan3 cannot remove Pab1, it would be expected to stall at a longer poly(A) tail length. The observation that Pan2-Pan3 is

stimulated down to 30 A's regardless of Pab1 concentration thus suggests that Pan2-Pan3 is either able to readily remove distally bound Pab1, or that the proximal Pab1 is more stably anchored than distal Pab1 molecules and has a lower off rate from poly(A) RNA.

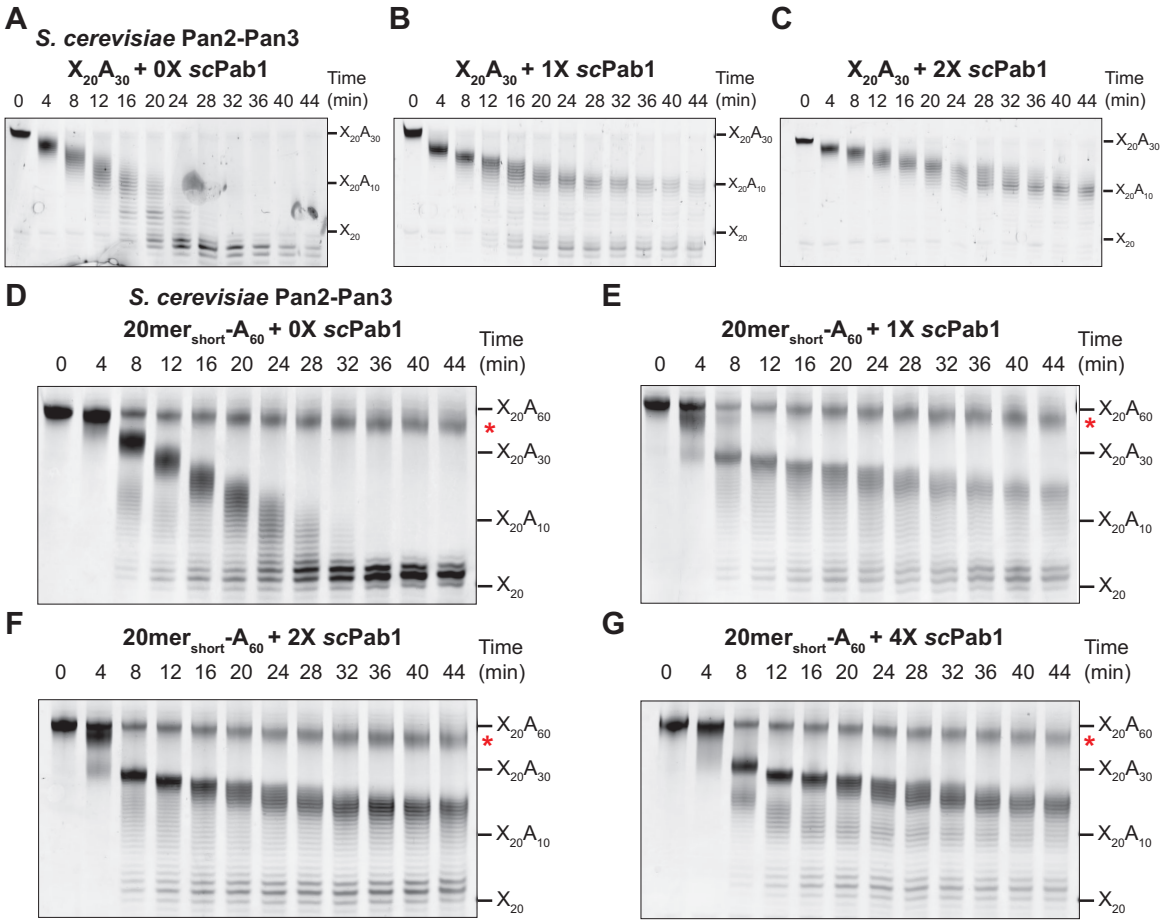


Figure 3.10 Dependence of Pab1 stoichiometry on Pan2-Pan3 stimulation. (**A-C**) Deadenylation assays of 200 nM 6-FAM-labelled $X_{20}A_{30}$ by 50 nM *S. cerevisiae* Pan2-Pan3 with (**A**) no Pab1; (**B**) 1× Pab1 (250 nM); and (**C**) 2× Pab1 (500 nM). Inhibition is stronger with higher Pab1 concentration. (**D-G**) Deadenylation assays of 200 nM unlabelled $20mer_{short}$-A_{60} by 50 nM *S. cerevisiae* Pan2-Pan3 with (**D**) no Pab1; (**E**) 1× Pab1 (250 nM); (**F**) 2× Pab1 (500 nM); and (**G**) 4× Pab1 (1 μM). Stimulation of the initial phase of deadenylation is the same regardless of Pab1 concentration, whereas inhibition between 30-0 A's is stronger with higher Pab1 concentration. The red asterisk is a likely contaminating RNA from *in vitro* transcription.

To test if these properties are conserved for human PAN2-PAN3, assays were repeated with RNAs with longer poly(A) tails ($X_{20}A_{60}$ and $X_{20}A_{90}$) and varied titrations of PABPC1. Human mRNAs are synthesised with longer poly(A) tails (150-200 A's) (Brawerman and Diez, 1975) compared to yeast; deadenylation is therefore initiated on longer poly(A) tails. Similar to

previous observations (Section 3.5), PABPC1 initially stimulated PAN2-PAN3, which was then stalled at A_{30} or A_{60}, likely corresponding to 1× and 2× PABPC1-bound RNA respectively (Figure 3.11A-B). In contrast with yeast, however, higher PABPC1 concentrations caused stalling at longer poly(A) tails. At 2× and 3× PABPC1, almost no deadenylation was observed over the time course (Figure 3.11C-D). At these concentrations (Figure 3.7D, dark blue arrow), poly(A) is saturated with 2-3 PABPC1 molecules. This result is thus consistent with the inability of PAN2-PAN3 to remove PABPC1.

Repeating these assays with an A_{90} poly(A) tail revealed a similar phenomenon (Figure 3.11E-H). At sub-saturating concentrations of PABPC1 (1×, 2×), PAN2-PAN3 was initially stimulated. Deadenylation was then inhibited until PAN2-PAN3 encountered the footprint of PABPC1 (at 30 A's and 60 A's). As the concentration of PABPC1 was further increased and more poly(A) tails became completely bound by PABPC1, deadenylation was increasingly inhibited. The results suggest that general inhibition of human PAN2-PAN3 by PABPC1 is independent of poly(A) tail length, and human PAN2-PAN3 cannot actively remove PABPC1.

Overall, these results suggest that *S. cerevisiae* and *H. sapiens* Pan2-Pan3 complexes display different behaviours on a Pab1-bound poly(A) tail. Both complexes are stimulated by Pab1 during initial deadenylation. *S. cerevisiae* Pan2-Pan3 can readily remove (or not be inhibited by) Pab1 which is distal to the 3' UTR, and is inhibited by proximal Pab1 at ~30 A's, regardless of Pab1-poly(A) stoichiometry. On the other hand, *H. sapiens* PAN2-PAN3 is inhibited upon encountering distal or proximal PABPC1, and increasing PABPC1 stoichiometry causes stalling at longer poly(A) tails, as poly(A) is saturated with PABPC1. This discrepancy could be the result of higher affinity and cooperativity of PABPC1 binding to poly(A) tails compared to Pab1, which would increase the stability of distally bound PABPC1. Structural and biochemical experiments are required to elucidate the mechanistic basis of this difference.

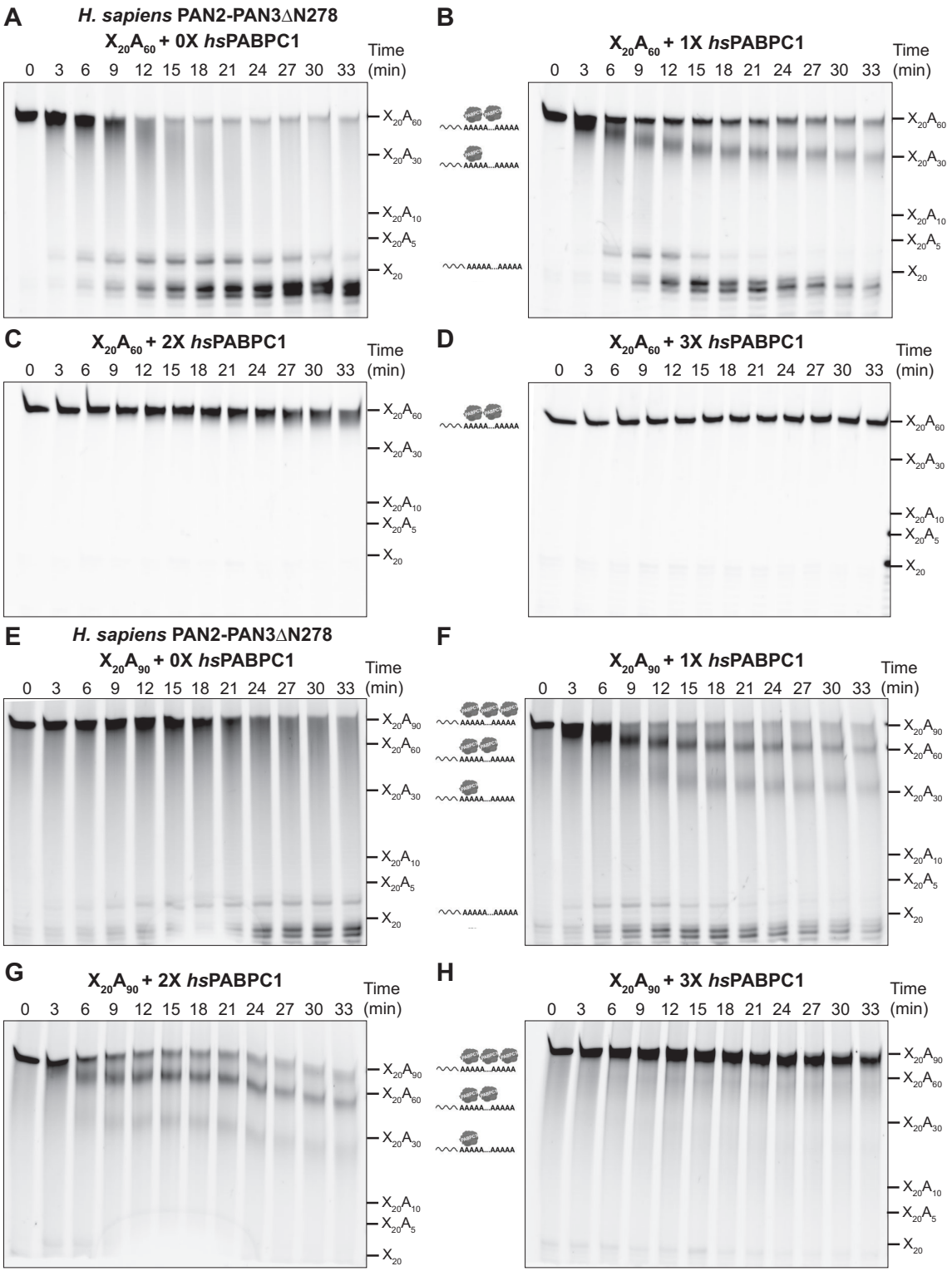


Figure 3.11 Effect of PABPC1 stoichiometry on deadenylation by 100 nM *H. sapiens* PAN2-PAN3ΔN278. (**A-D**) Deadenylation of 200 nM 6-FAM-labelled $X_{20}A_{60}$ with (**A**) no PABPC1; (**B**) 1× PABPC1 (350 nM); (**C**) 2× PABPC1 (700 nM); and (**D**) 3× PABPC1 (1.05 μM). (**E-H**) Deadenylation of 200 nM 6-FAM-labelled $X_{20}A_{90}$ with (**E**) no PABPC1; (**F**) 1× PABPC1 (350 nM); (**G**) 2× PABPC1 (700 nM); and (**H**) 3× PABPC1 (1.05 μM). Stimulation is ablated at higher Pab1 concentration and PAN2-PAN3 is inhibited upon encountering PABPC1. The poly(A) tail length at which this occurs varies depending on PABPC1 stoichiometry. RNAs with putative PABPC1 stoichiometries is shown next to corresponding stall points.

3.7 Domain Requirements of Pab1 Regulation of Pan2-Pan3

3.7.1 The Pan2 Exonuclease is Inhibited by Pab1

Further *in vitro* deadenylation assays were carried out to understand the mechanism of Pab1-mediated regulation of Pan2-Pan3 activity. Yeast Pan2-Pan3 was inhibited upon encountering the 3'-UTR-proximal Pab1; however, it was unknown whether Pan3 or non-catalytic domains of Pan2 are required for this inhibitory effect. To test if the Pan2 exonuclease is also inhibited by Pab1, Pan2UCH-Exo (Chapter 2) was used in deadenylation assays to compare the effect of Pab1 on the exonuclease alone. On the $X_{20}A_{30}$ substrate, Pan2UCH-Exo was inhibited by the presence of 1× Pab1 (Figure 3.12). Furthermore, stall points at A_{10} and A_{18}, which likely correspond to individual RRM domains, were observed during deadenylation by Pan2UCH-Exo. Therefore, the Pan2 exonuclease alone is inhibited by Pab1, and Pan3 is not required for Pab1-mediated inhibition. This supports a model where Pan2-Pan3 is inhibited when the Pan2 active site encounters the tightly-bound Pab1-poly(A). It remains to be seen if the human PAN2 exonuclease is also inhibited by PABPC1; however, the high degree of conservation of Pan2 and the higher affinity of PABPC1 for poly(A) suggests that this effect is likely conserved.

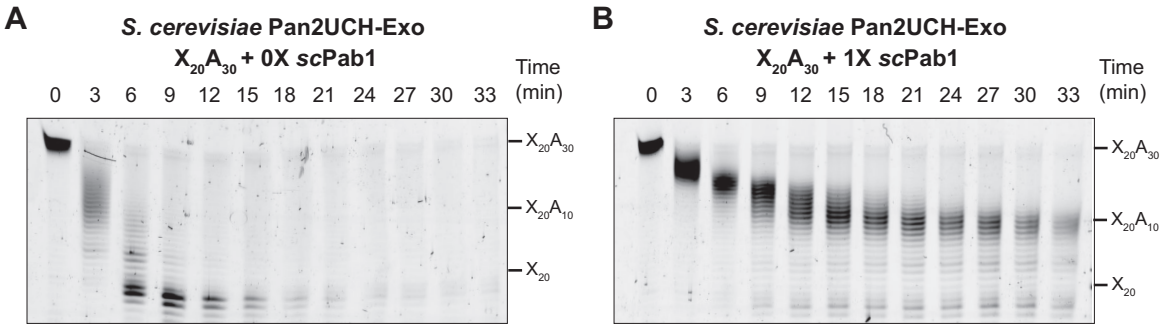


Figure 3.12 The Pan2 exonuclease alone is inhibited by Pab1. (**A-B**) Deadenylation of 200 nM 6-FAM-labelled $X_{20}A_{30}$ by 100 nM *S. cerevisiae* Pan2UCH-Exo with (**A**) no Pab1; and (**B**) 1× Pab1 (250 nM). Similar to Pan2-Pan3, Pan2UCH-Exo is strongly inhibited by Pab1.

3.7.2 The C-terminus of Pab1 Stimulates Yeast Pan2-Pan3

To test how Pab1 stimulates Pan2-Pan3, truncations of yeast and human Pan2-Pan3 and Pab1 were cloned, overexpressed, and purified as above (Figure 3.1, Figure 3.2). These truncations were designed to test the importance of the PAM2-CTD interaction in stimulation, as this interaction was thought to recruit Pan2-Pan3 to the poly(A) tail prior to the PhD (Siddiqui et al., 2007). This recruitment to poly(A) was thought to stimulate Pan2-Pan3 activity. However, the recent structure of yeast Pan2-Pan3 in complex with Pab1-poly(A) showed that Pan2-Pan3 binds Pab1 multimerisation interfaces, which consist of an α-helix following RRM4 and RRM1 of the adjacent Pab1 (Schafer et al., 2019). Thus, these surfaces could also be important in recruiting Pan2-Pan3 to the poly(A) tail and stimulating its activity.

To test if the PAM2-Pab1 CTD interaction is required for stimulation of activity, a mutation was introduced in the Pan3 PAM2 motif. Mutation of a key phenylalanine to alanine reduces the affinity of interaction between PAM2 and the CTD by >6000× for PAIP2 (Kozlov et al., 2004); thus, the same mutation was introduced into Pan3 (F156A), effectively disrupting the PAM2-Pab1 interaction. Compared to when Pab1 was absent (Supplementary Figure 8), the F156A mutant could still be stimulated by full-length Pab1; however, notably, the proportion of rapidly deadenylated RNAs was much lower compared to wild-type Pan2-Pan3 (Figure 3.13A-B). To confirm the lack of absolute requirement of the PAM2 motif, Pan3 lacking the entire N-terminal low-complexity region (Pan3PKC) was tested in deadenylation assays with Pab1 (Figure 3.13C-D). Compared to when Pab1 was absent (Supplementary Figure 8), Pan2-Pan3PKC was stimulated by Pab1, but the stimulation was attenuated compared to full-length Pan2-Pan3. These results suggest that while the PAM2-Pab1 interaction is not absolutely required for stimulation, it aids full stimulation of Pan2-Pan3 activity. The lack of absolute requirement of the PAM2 motif suggests that there is an alternative mechanism of stimulation.

Deadenylation assays were repeated using a C-terminal truncation of Pab1 containing only the RNA-binding domains (RRM1-4). This construct did not stimulate either full-length Pan2-Pan3 (Figure 3.13E, compare with Figure 3.13C and Supplementary Figure 8C) or Pan2-Pan3PKC (Figure 3.13F, compare with Figure 3.13D and Supplementary Figure 8D), but caused stalling once the complexes reached A_{30} (Figure 3.13E). Thus, the Pab1 C-terminus is important in Pan2-Pan3 stimulation, whereas the RRM domains cannot stimulate Pan2-Pan3 but are sufficient to inhibit the complex. This is in agreement with the structural and biochemical studies published during my PhD, suggesting that the helix immediately following RRM4 is important for contacting Pan2-Pan3 and thus stimulation of its activity (Schafer et al., 2019). Pab1 constructs including this helix will be required to test its requirement in stimulating Pan2-Pan3. Overall, my results suggest that the Pab1 C-terminus is required to stimulate Pan2-Pan3. However, the PAM2-Pab1 interaction is not absolutely required for, but greatly aids, stimulation of Pan2-Pan3 deadenylation.

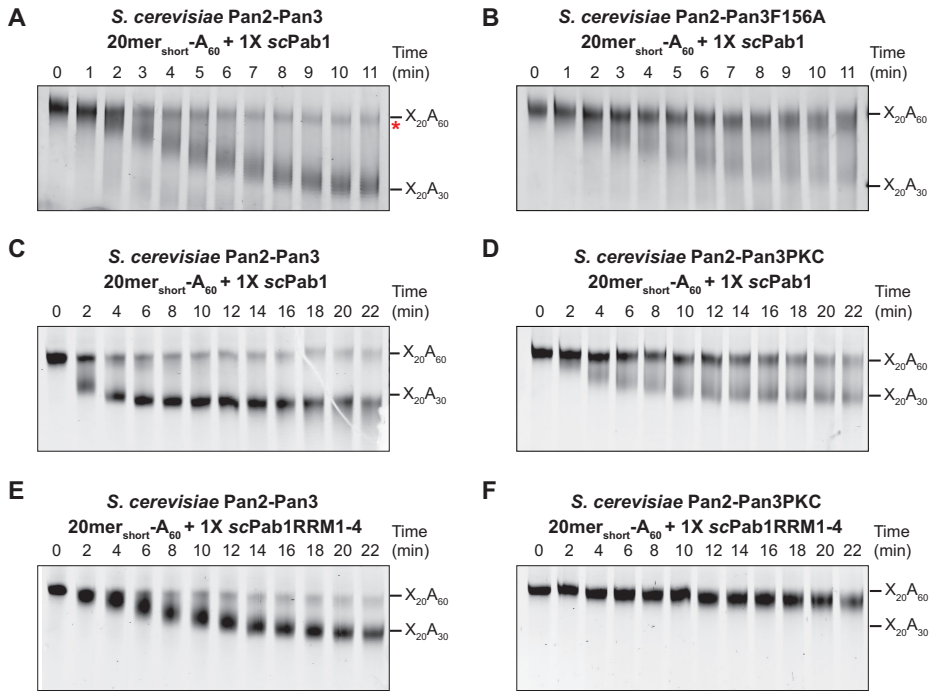


Figure 3.13 The Pab1 C-terminus stimulates Pan2-Pan3. (**A-D**) Deadenylation of 200 nM unlabelled 20mer$_{short}$-A_{60} bound by 1× Pab1 (250 nM) by 50 nM (**A, C**) *S. cerevisiae* Pan2-Pan3; (**B**) *S. cerevisiae* Pan2-Pan3F156A; and (**D**) *S. cerevisiae* Pan2-Pan3PKC. (**E-F**) Deadenylation of 200 nM unlabelled 20mer$_{short}$-A_{60} bound by 1× Pab1RRM1-4 (250 nM) by 50 nM (**E**) *S. cerevisiae* Pan2-Pan3; and (**F**) *S. cerevisiae* Pan2-Pan3PKC.

3.7.3 PABPC1 RRM Domains Can Stimulate PAN2-PAN3

To test if the domain requirements of stimulation are conserved for human PAN2-PAN3 and PABPC1, deadenylation assays were carried out with N-terminal truncations of PAN3. Similar to yeast Pan2-Pan3, systematic truncation of the N-terminus did not ablate PABPC1-mediated stimulation of human PAN2-PAN3 (Figure 3.14A-C). Nonetheless, PAN2-PAN3PKC (Figure 3.14C) was less stimulated compared to PAN3 constructs containing the PAM2 motif (Figure 3.14A-B), suggesting that the PABPC1 CTD is important in the full stimulation of the complex. Thus, analogous to yeast Pan2-Pan3 and Pab1, the PAM2-PABPC1 interaction is not absolutely required for stimulation, but aids full stimulation of PAN2-PAN3 deadenylation.

Human PABPC1 truncations were tested for their ability to stimulate PAN2-PAN3. Compared to full-length PABPC1 (Figure 3.14A-C), a construct where the CTD and P-linker were truncated but the multimerisation helix was retained (RRMH) retained the ability to stimulate both full-length PAN2-PAN3 (Figure 3.14D) and PAN2-PAN3PKC (Figure 3.14F). Thus, the multimerisation interface may also be important in recruitment of the PAN2-PAN3 core complex to the poly(A) tail. Nonetheless, a construct lacking the multimerisation helix was also able to stimulate PAN2-PAN3 (Figure 3.14G-I, compared to Supplementary Figure 9). Notably, both C-terminal PABPC1 truncations partially stimulated PAN2-PAN3 less than full-length PABPC1, as judged by the lower intensity of the species corresponding to stimulated RNA (Figure 3.14A-C). This confirms the importance of the PABPC1 CTD in full stimulation of PAN2-PAN3. The ability of RRM1-4 to stimulate PAN2-PAN3 contrasts with yeast Pab1, where the same construct does not stimulate yeast Pan2-Pan3. This suggests that PABPC1 RRM1-4 may form additional contacts with the PAN2-PAN3 core. My results show that while PABPC1 is similar to Pab1 in that the C-terminal domain is important for full stimulation of PAN2-PAN3, its RRM domains may form different contacts with PAN2-PAN3 such that PABPC1 RRM1-4 alone can stimulate deadenylation.

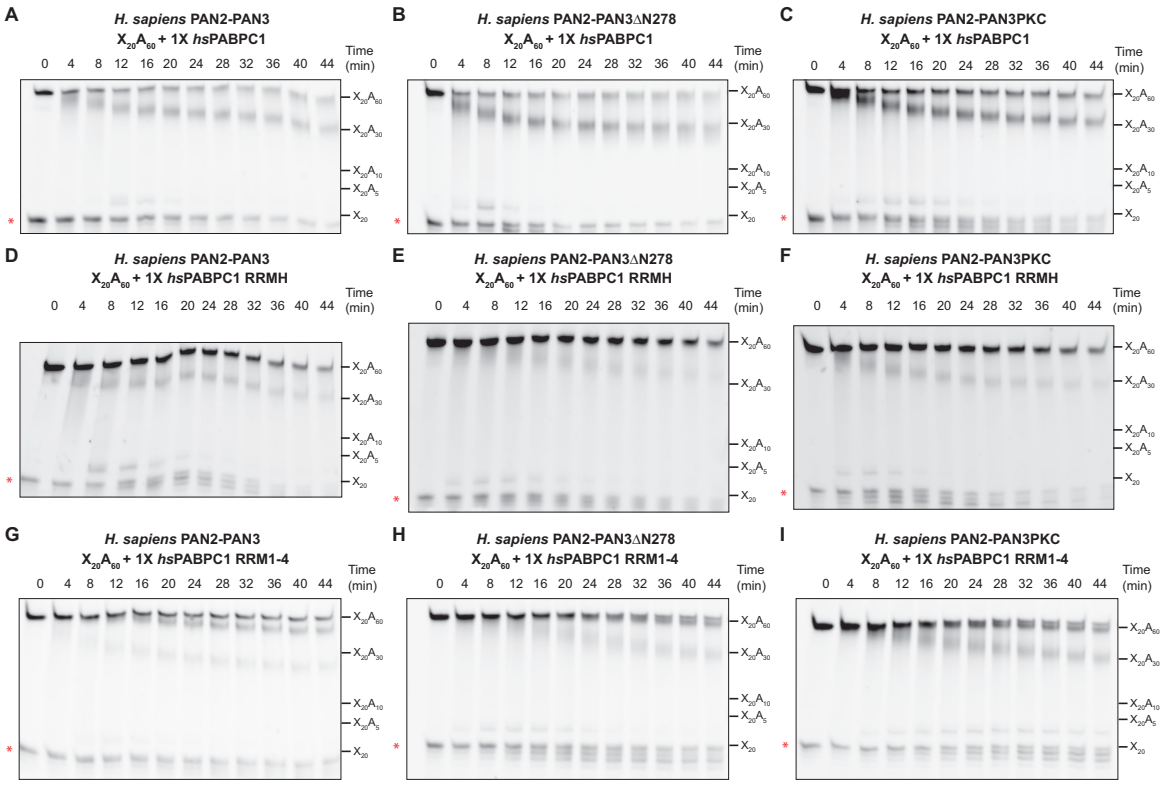


Figure 3.14 PABPC1 RRM1-4 can partially stimulate PAN2-PAN3. (**A-C**) Deadenylation of 200 nM 6-FAM-labelled $X_{20}A_{60}$ bound by 1× PABPC1 (350 nM) by 100 nM (**A**) *H. sapiens* PAN2-PAN3; (**B**) *H. sapiens* PAN2-PAN3ΔN278; and (**C**) *H. sapiens* PAN2-PAN3PKC. (**D-F**) Deadenylation of 200 nM 6-FAM-labelled $X_{20}A_{60}$ bound by 1× PABPC1 RRMH (350 nM) by 100 nM (**D**) *H. sapiens* PAN2-PAN3; (**E**) *H. sapiens* PAN2-PAN3ΔN278; and (**F**) *H. sapiens* PAN2-PAN3PKC. (**G-I**) Deadenylation of 200 nM 6-FAM-labelled $X_{20}A_{60}$ bound by 1× PABPC1 RRM1-4 (350 nM) by 100 nM (**G**) *H. sapiens* PAN2-PAN3; (**H**) *H. sapiens* PAN2-PAN3ΔN278; and (**I**) *H. sapiens* PAN2-PAN3PKC. C-terminal truncations of PABPC1 retain the ability to stimulate PAN2-PAN3, but full stimulation of PAN2-PAN3 requires the C-terminus of PABPC1.

3.8 The PAN2-PAN3-PABPC1-poly(A) Complex

3.8.1 Pab1 C-terminus is Required to Interact with Pan2-Pan3

In vitro deadenylation assays suggest that the interaction between yeast Pan2-Pan3 and Pab1-poly(A) may be different to that between human PAN2-PAN3 and PABPC1-poly(A). For yeast, the RRM domains alone cannot stimulate Pan2-Pan3. This is likely due to recognition of the Pab1 multimerisation interface, suggesting that at least two Pab1 molecules must be simultaneously bound to a >60-A poly(A) tail for a stable interaction with Pan2-Pan3.

To test if the Pab1-Pan2-Pan3 interaction depends on poly(A) length, pull-down assays were performed. SII-tagged Pan2E912A-Pan3 (Pan2 catalytic mutant) was immobilised on magnetic StrepTactin beads. Beads were incubated with a constant amount of excess Pab1, which was pre-mixed with RNAs with different poly(A) tail lengths. The poly(A) tail length correlated with the amount of Pab1 pulled down by immobilised Pan2-Pan3 (Figure 3.15A). Specifically, Pan2-Pan3 bound more Pab1-A_{30} compared to shorter poly(A) tails (lanes 1-3 vs lanes 4-5); lengthening poly(A) tails from A_{30} to A_{60} and A_{80} further increased the amount of Pab1 pulled down (lanes 4-5 vs lanes 6-7). Thus, consistent with previous studies, longer poly(A) tails (allowing Pab1-Pab1 multimerisation interfaces to form on poly(A)) facilitate more stable interactions with Pan2-Pan3.

To assess domain requirements for a stable Pan2-Pan3-Pab1-poly(A) interaction, truncations used in deadenylation assays were tested in their ability to interact in pull-down assays (Figure 3.15B). Full-length Pan2-Pan3 interacted stoichiometrically with Pab1 (lane 1); however, Pan2-Pan3PKC only pulled down slightly more Pab1 (lane 2) than background (lane 3). In contrast, Pab1 RRM1-4 was not efficiently pulled down by any Pan2-Pan3 variants above background (lanes 4-5 vs lane 6). This reinforces previous observations that the Pab1 C-terminus is

important for the interaction with Pan2-Pan3 and thus stimulation of its deadenylation activity (Figure 3.13). In summary, the ability of Pab1 truncations to interact with Pan2-Pan3 constructs corresponds to the ability of Pab1 to stimulate Pan2-Pan3. More specifically, full interaction and stimulation are observed with full-length proteins, but the Pan3PKC construct retains partial ability to interact with Pab1 and is thus stimulated. The RRM1-4 construct cannot stably interact with any Pan2-Pan3 construct tested and so is not stimulatory. These results support the idea that stimulation of Pan2-Pan3 activity is due to its direct recruitment to the poly(A) tail by Pab1 and not by auxiliary effects.

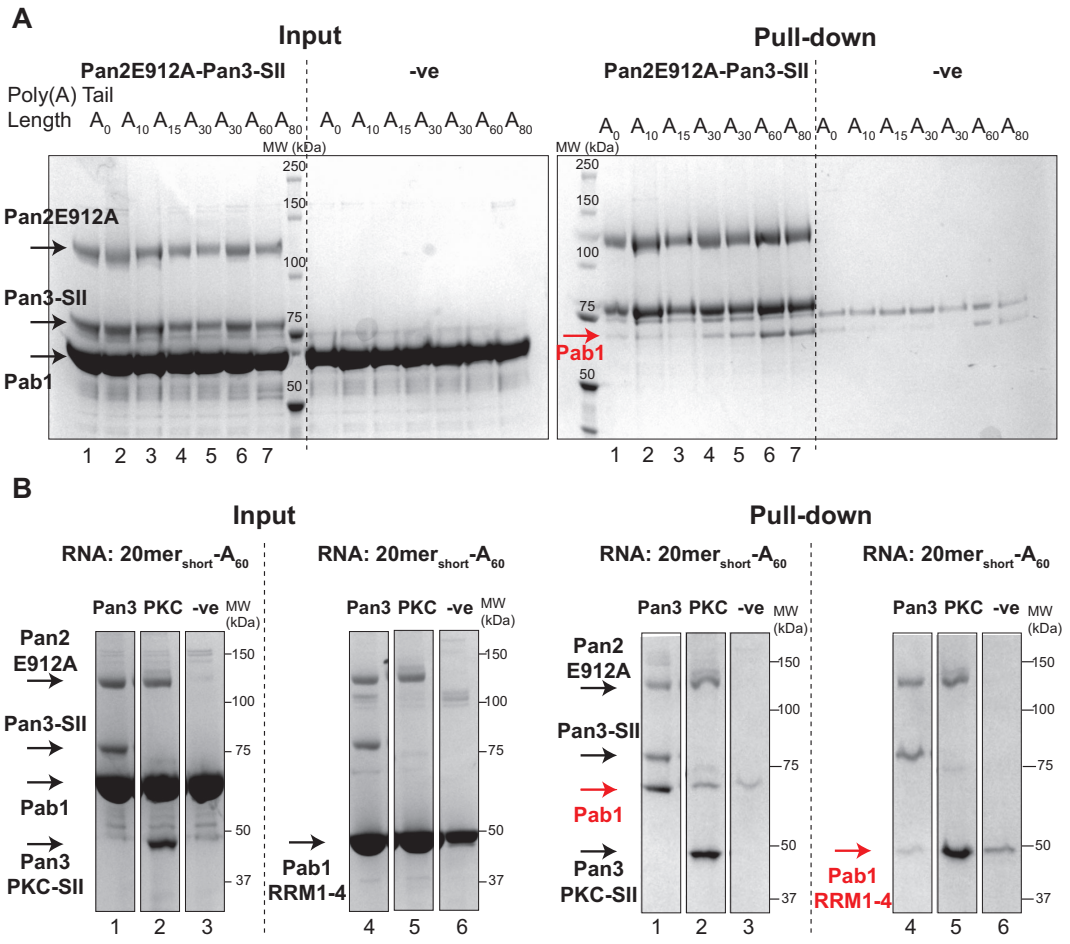


Figure 3.15 The Pab1 C-terminus is required for a stable interaction with Pan2-Pan3. Pull-down assays to test (**A**) poly(A) tail length and (**B**) domain requirements for Pan2-Pan3 interaction with Pab1-poly(A). Input proteins are indicated with black arrows; bait protein is indicated with a red arrow. (**A**) 0.05 nmol Pan2E912A-Pan3-SII was incubated with 0.25 nmol Pab1 + 0.25 nmol RNA. RNAs used were: X_{20}, $X_{20}A_{10}$, A_{15}, A_{30}, $X_{20}A_{30}$, 20mer$_{short}$-A_{60}, CYC1-A_{80}. (**B**) 0.07 nmol Pan2E912A-Pan3(PKC)-SII was incubated with 0.35 nmol Pab1(RRM1-4) + 0.35 nmol 20mer$_{short}$-A_{60}.

3.8.2 PAN2-PAN3 Interacts with PABPC1-Poly(A) Independent of Poly(A) Tail Length and Stoichiometry

To test if PAN2-PAN3 displays poly(A) length-dependence in its interaction with PABPC1, pull-down assays were repeated with immobilised PAN2E980A-PAN3ΔN278-SII (catalytic mutant) and PABPC1 (Figure 3.16A). Surprisingly, human PAN2-PAN3 stoichiometrically interacted with PABPC1 independent of tail-length. The lack of RNA likely caused non-specific PABPC1 binding to beads, accounting for the higher PABPC1 amount when no RNA was present (lane 1). Increasing poly(A) length (from lane 2-5) did not result in substantially more PABPC1 being pulled down. Thus, the interaction between PAN2-PAN3 and PABPC1 is not strongly dependent on poly(A) tail length, which contrasts with the poly(A) tail length-dependence of the interaction between yeast Pan2-Pan3 and Pab1-poly(A) (Figure 3.16A).

One possibility for the apparent independence of poly(A) tail length is that the cooperativity of poly(A) binding by PABPC1 meant that not all RNAs were bound by one PABPC1. For example, on an A_{60} tail, addition of 1× PABPC1 caused a mixture of 0×, 1×, and 2× PABPC1-bound species (Section 3.4). To test the stoichiometry dependence of the interaction, pull-down assays were repeated with constant PABPC1, but decreasing RNA concentrations as a proportion of poly(A) tail length (Figure 3.16B). Thus, the stoichiometry of PABPC1:RNA was 1:1 for RNAs that can accommodate one PABPC1 (lanes 1-3), 2:1 for A_{60} which can bind two (lane 4), and 3:1 for A_{90} (lane 5). At these stoichiometries, poly(A) tails should be saturated with PABPC1. These pull-down assays showed that the interaction between PAN2-PAN3 and PABPC1 was independent of PABPC1:poly(A) stoichiometry, as a similar amount of PABPC1 was pulled down in each lane. The independence of PABPC1 stoichiometry and poly(A) tail length is strikingly different to that of yeast Pan2-Pan3 and Pab1-poly(A) (Schafer et al., 2019), and potentially suggests a different mechanism of recognition of the PABPC1-poly(A) RNP.

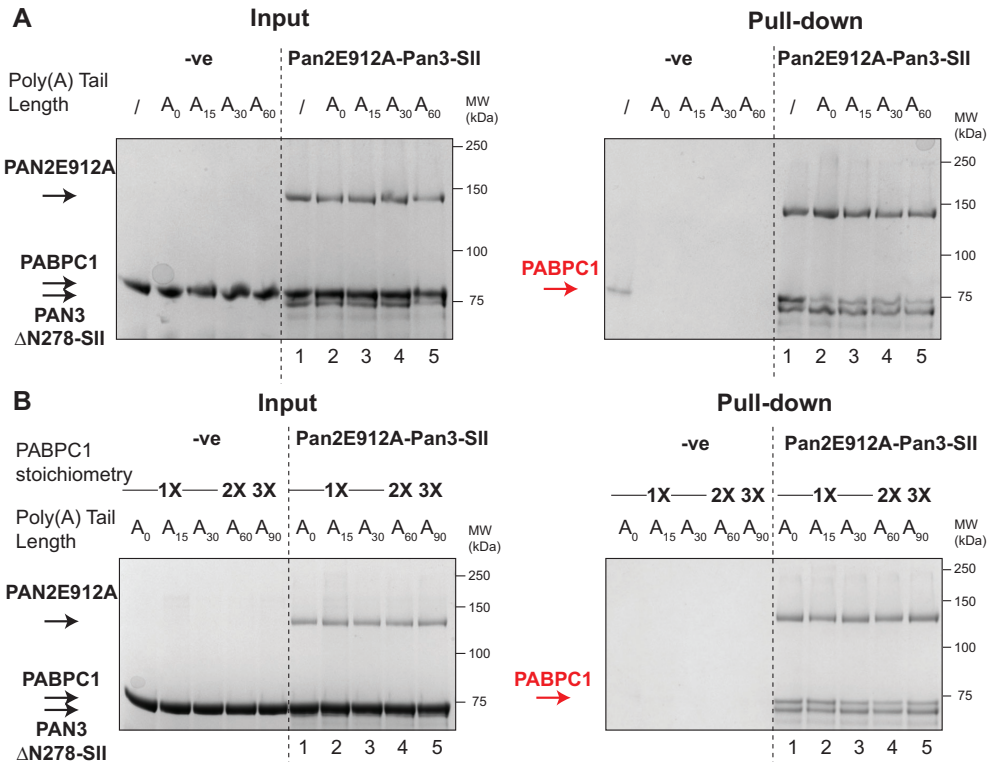


Figure 3.16 PAN2-PAN3 interacts with PABPC1 in a poly(A) tail length- and stoichiometry-independent manner. Pull-down assays to test the effects of (**A**) poly(A) tail length and (**B**) PABPC1-poly(A) stoichiometry on PAN2-PAN3 interaction with PABPC1-poly(A). Input proteins are indicated by black arrows; bait protein is shown by the red arrow. (**A**) 0.04 nmol PAN2E980A-PAN3ΔN278-SII was incubated with 0.1 nmol PABPC1 + 0.1 nmol RNA. RNAs used were: X_{20}, A_{15}, $X_{20}A_{30}$, $X_{20}A_{60}$. (**B**) 0.05 nmol PAN2E980A-PAN3ΔN278-SII was incubated with 0.3 nmol PABPC1 + 0.3 nmol X_{20}, A_{15}, $X_{20}A_{30}$, or + 0.15 nmol $X_{20}A_{60}$, or + 0.1 nmol $X_{20}A_{90}$.

3.8.3 The Minimal PAN2-PAN3-PABPC1-Poly(A) Complex

To understand the molecular basis of recognition of PABPC1-poly(A) by PAN2-PAN3, and thereby elucidate species-specific differences which can account for the different *in vitro* behaviours observed above, high-resolution cryo-EM studies would likely be required. Cryo-EM is more likely to succeed in structural investigations compared to other methods due to the large size of the ternary complex (MW: ~500 kDa) and the likely higher flexibility of different components, such as PABPC1-poly(A). To obtain a stable complex which is amenable to cryo-EM studies, size exclusion chromatography (SEC) was performed to define the minimal PAN2-PAN3, PABPC1, and RNA complex which remains stably bound. A stable complex would prevent on-grid dissociation and minimise the amount of low-complexity

region, which is hypothesised to be detrimental to cryo-EM imaging. Finally, given the lower local resolution of Pab1-poly(A) in the previous structure (Schafer et al., 2019), additional strategies may be required to stabilise PABPC1 binding on the poly(A) tail.

First, as PAN2-PAN3PKC is stimulated by PABPC1 (Figure 3.14C), PAN2-PAN3 lacking the low-complexity N-terminus of PAN3 was tested for interaction with PABPC1-poly(A). Incubation with PABPC1-$X_{20}A_{60}$ caused a higher-molecular weight species with stoichiometric PAN2-PAN3PKC and PABPC1 to appear (magenta), compared to PAN2-PAN3PKC alone (black) (Figure 3.17A). The presence of RNA in this peak was confirmed by the high A_{260}/A_{280} ratio. Thus, PAN2-PAN3PKC stably interacts with PABPC1-poly(A), and this ternary complex can be separated from apo PAN2-PAN3PKC by SEC.

As RRM1-4 could stimulate PAN2-PAN3 activity *in vitro* (Figure 3.14G), albeit weakly, PABPC1 RRM1-4 and PAN2-PAN3PKC were tested for their ability to stably interact (blue). A clear shift of the peak to a higher molecular weight species was observed with PAN2-PAN3-PABPC1 RRM1-4 (blue, 1.45 mL) compared to unbound PAN2-PAN3PKC (grey, 1.51 mL) (Figure 3.17B). RNA presence was confirmed by the high A_{260}/A_{280} ratio and stoichiometric binding was verified by SDS PAGE. Thus, the RRM domains alone can stably interact with PAN2-PAN3PKC. As the multimerisation helix immediately following RRM4 is involved in interacting with Pan2-Pan3 in the yeast complex, the same predicted helix (Supplementary Figure 6) was included in the RRMH construct, which was used in further studies.

The effect of poly(A) tail length on complex formation was also tested. When PABPC1 RRMH was mixed with A_0 (magenta) and A_{30} (green) and incubated with PAN2-PAN3PKC, no higher molecular-weight species appeared, and PABPC1 RRMH-RNA and PAN2-PAN3PKC remained as two distinct peaks (Figure 3.17C; magenta and green). When PABPC1 RRMH was mixed with A_{60}, a peak corresponding to the ternary complex appeared, whose stoichiometry was confirmed by SDS-PAGE (Figure 3.17C, blue). Therefore, SEC

experiments suggest that two PABPC1 molecules must be simultaneously bound to a longer (~A_{60}) poly(A) tail for a stable interaction with PAN2-PAN3. This is in contrast with pull-down studies, as PAN2-PAN3PKC was able to pull-down PABPC1-RNA independent of stoichiometry. This is possibly due to the ability of pull-down experiments to detect less stable and more transient interactions.

Finally, in the structure of the *S. cerevisiae* Pan2-Pan3-Pab1-poly(A) complex, Pab1-poly(A) is only visible at low local resolution, hindering molecular understanding of how the RRM domains contact poly(A) and how Pan2-Pan3 interacts with Pab1-poly(A) (Schafer et al., 2019). I suspected that the low resolution might be due to Pab1, which could slide or bind in any register along the length of poly(A) RNA. Thus, an RNA was designed to contain a short upstream UTR-like sequence (X_{10}) and a poly(A) tail slightly shorter than that which can accommodate 3× PABPC1 molecules (A_{83}). Combined with the predicted lower RNA binding specificity of PABPC1 RRM4 (Section 1.4.1.2), this could, in effect, force RRM4 to bind the UTR-like sequence, ensuring a fixed PABPC1 pattern on the poly(A) tail. Consistent with the minimum tail length of A_{60} required for a stable interaction, this RNA was able to form a stable ternary complex with PAN2-PAN3PKC and PABPC1RRMH (Figure 3.17D).

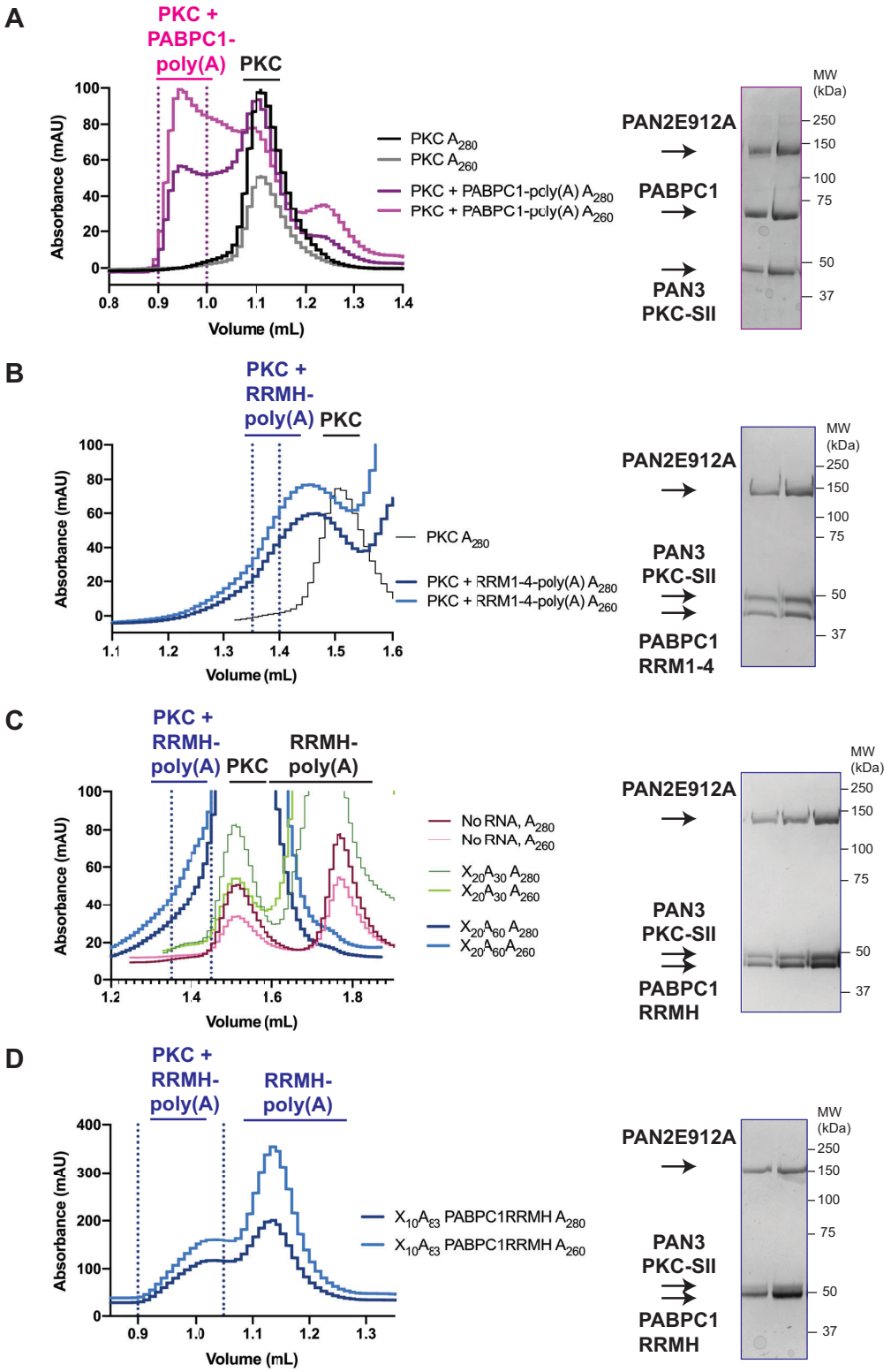


Figure 3.17 Defining the minimal complex of PAN2-PAN3-PABPC1-poly(A). Size exclusion chromatography assays and SDS PAGE of the marked fractions show that (**A**) PAN2-PAN3PKC interacts with PABPC1-poly(A); (**B**) PABPC1RRM1-4-poly(A) interacts with PAN2-PAN3PKC; (**C**) length-dependence of RNA required for a stable ternary complex; and (**D**) the optimised PAN2-PAN3-PABPC1-poly(A) complex for initial structural investigations. The dotted lines denote the fractions shown in SDS PAGE. The different elution volumes of panels B-C compared to panels A, D are due to the use of a Superose6 3.2/300 column instead of a S200 3.2/300 column.

Previous studies suggested that Pab1 RRM3-4, and particularly RRM4, have lower poly(A) specificity and may bind single-stranded RNA upstream of the poly(A) tail (Baejen et al., 2014; Kini et al., 2016; Mullin et al., 2004; Sladic et al., 2004). RRM4 binding to the UTR-like sequence was demonstrated *in vitro* by Ccr4-Not deadenylation of $X_{20}A_{30}$ (Webster et al., 2018). Deadenylation was stalled at 22 adenosines with a UTR sequence (instead of the 28-A Pab1 footprint), suggesting that RRM4 was bound to the UTR and anchored distal Pab1. To test if PABPC1 is similarly arranged during PAN2-PAN3 deadenylation, deadenylation assays were performed to elucidate stall points. PABPC1 RRM1-4 inhibited PAN2-PAN3ΔN278 (Figure 3.18A-B). In addition to the overall inhibition, a strong stall was visible at $A_{21\text{-}22}$, suggesting that RRM4 was bound to the upstream sequence during PAN2-PAN3 deadenylation (Figure 3.18B). Similarly, for an A_{60} tail saturated with PABPC1 RRM1-4, deadenylation was inhibited at ~50 A's, though the resolution was insufficient to determine the exact stall point (Figure 3.18C-D). Thus, RRM4 binding to the non-A sequence can be recapitulated in PAN2-PAN3 deadenylation, and $N_{10}A_{83}$ RNA may fix PABPC1 on poly(A) for structural investigation.

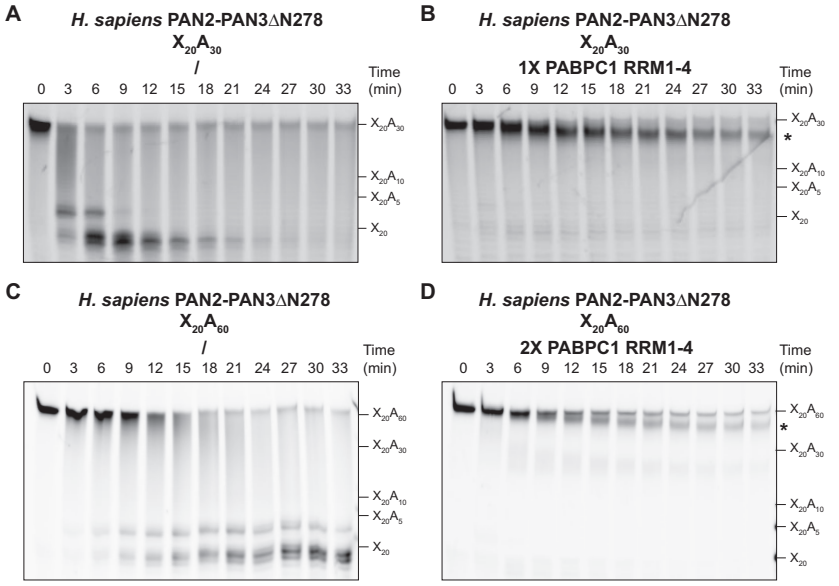


Figure 3.18 RRM4 binds the 3' UTR during PAN2-PAN3 deadenylation. Deadenylation assays of 200 nM $X_{20}A_{30}$ by 100 nM PAN2-PAN3ΔN278 in the (**A**) absence or (**B**) presence of 1× PABPC1 RRM1-4 (350 nM). Deadenylation assays of 200 nM $X_{20}A_{60}$ by 100 nM PAN2-PAN3ΔN278 in the (**C**) absence or (**D**) presence of 2× PABPC1 RRM1-4 (700 nM). Asterisks denote the stall points at approximately (**B**) 22 A's and (**D**) 50 A's. These suggest that RRM4 binds to the upstream non-A sequence.

3.9 Crosslinking of Pan2-Pan3

To gain mechanistic insight into deadenylation by Pan2-Pan3, I aimed to perform cryo-EM studies first on apo Pan2-Pan3, followed by the ternary complex with Pab1 and poly(A). While yeast Pan2-Pan3 appears to be stable in solution, previous preliminary cryo-EM studies have shown that it is difficult to obtain individual particles of the predicted size (Jana Wolf). Observable particles appeared to be denatured during grid preparation. Thus, a crosslinking approach was carried out in order to covalently stabilise Pan2-Pan3, which could be compared to uncrosslinked sample in preliminary cryo-EM studies. BS3 crosslinking of accessible lysine amines may prevent Pan2-Pan3 dissociation during grid preparation or alter surface properties by neutralising positively charged amines and by introducing a hydrophobic alkyl moiety to potentially prevent its interaction with the air-water interface. However, crosslinking may also stabilise non-physiological conformations of the complex. Alternatively, flexible regions may be crosslinked in multiple conformations, precluding high-resolution structure determination.

Crosslinking of yeast and human Pan2-Pan3 was optimised by BS3 titration (Figure 3.19A, C). PAN2-PAN3ΔN278 was used due to its higher yield in overexpression and purification. Ideal BS3 concentration (yeast: 0.5 mM; human: 0.75 mM) was determined by the minimum concentration at which uncrosslinked proteins disappeared. Crosslinked products were analysed by size exclusion chromatography (Figure 3.19B, D). Crosslinked Pan2-Pan3 at the predicted molecular weight could be separated from larger crosslinked products. This condition was used for further structural studies of Pan2-Pan3 containing (yeast: full-length, human: ΔN278) and lacking (PKC) the low-complexity regions.

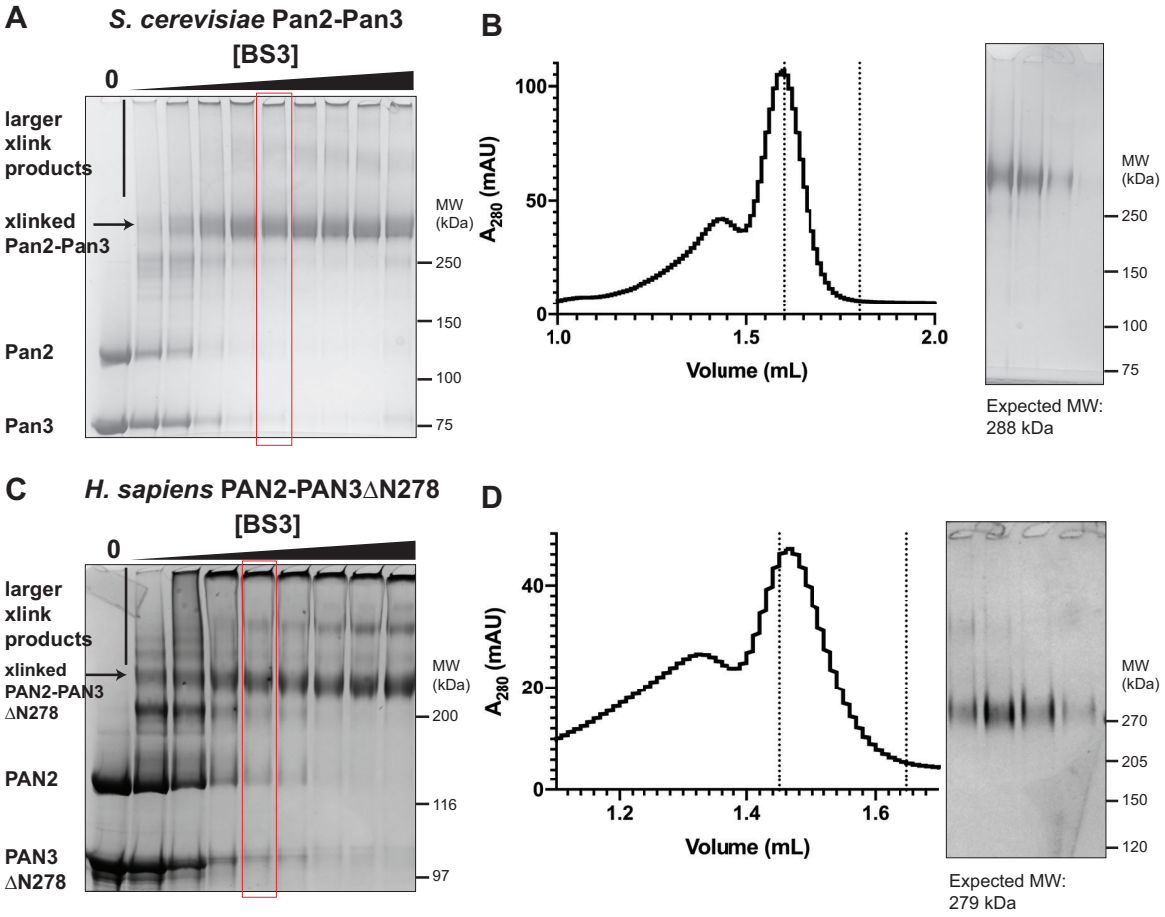


Figure 3.19 Crosslinking trials of *S. cerevisiae* and *H. sapiens* Pan2-Pan3. BS3 titration of (**A**) Pan2-Pan3 and (**C**) PAN2-PAN3ΔN278. Separation of crosslinked (**B**) Pan2-Pan3 and (**D**) PAN2-PAN3ΔN278 by size exclusion chromatography, and SDS PAGE of fractions. The pooled fractions are denoted by dotted lines and were selected to avoid higher-order aggregates. As observed in SDS PAGE, no significant degradation products are observed.

3.10 Negative Stain EM of Apo Pan2-Pan3

To assess sample homogeneity and determine the overall shape of yeast and human Pan2-Pan3, negative stain electron microscopy was performed using 30 nM crosslinked *S. cerevisiae* Pan2-Pan3 or 40 nM uncrosslinked *H. sapiens* PAN2-PAN3ΔN278. Compared to the uncrosslinked complex, crosslinking did not dramatically alter the particle distribution or final 3D map obtained for *S. cerevisiae* Pan2-Pan3, suggesting that BS3 crosslinking did not cause large-scale conformational changes (data not shown). However, qualitatively, crosslinking led to a more even particle distribution and less visible complex disassembly; thus, crosslinking may alter surface properties to improve sample preparation.

Both *S. cerevisiae* and *H. sapiens* Pan2-Pan3 displayed good staining and particle distribution (Figure 3.20A, D). SEC-MALS experiments have estimated that *S. cerevisiae* Pan2-Pan3 is ~17 nm (Jana Wolf), and the composite structure of the Pan2-Pan3 core (Section 1.2.2) suggests the particle is ~16 nm in its maximum dimension. These estimates are consistent with the ~20 nm stain-excluded particles observed in negative stain EM micrographs. The similar size of *H. sapiens* PAN2-PAN3ΔN278 suggests that the human complex does not form a significantly larger or different particle despite the longer primary sequence (Supplementary Figure 1, Supplementary Figure 2). 2D classification and 3D maps from selected 2D classes (Figure 3.20B, E) revealed that Pan2-Pan3 forms a tri-lobed structure, which is relatively flat from "side" and "top" views, in line with the putative structural model of the Pan2-Pan3 core complex (Figure 3.20C, F; yellow). The larger size of the 3D map from negative stain EM compared to the rescaled composite structural model may be due to inaccurate pixel size measurements or the effect of stain exclusion from the particle. In the maximum dimension, 3D maps of yeast Pan2-Pan3 are ~16 nm, whereas those of human PAN2-PAN3ΔN278 are ~18 nm. Finally, consistent with structural predictions which suggest that the Pan3 N-terminus is largely disordered (Supplementary Figure 3, Supplementary Figure 4), no significant density

additional to the core complex is visible in the low-resolution reconstruction. Thus, these results reveal that the yeast and human Pan2-Pan3 cores form a conserved tri-lobed structure with a maximum dimension of 16-18 nm, in line with previous structural predictions.

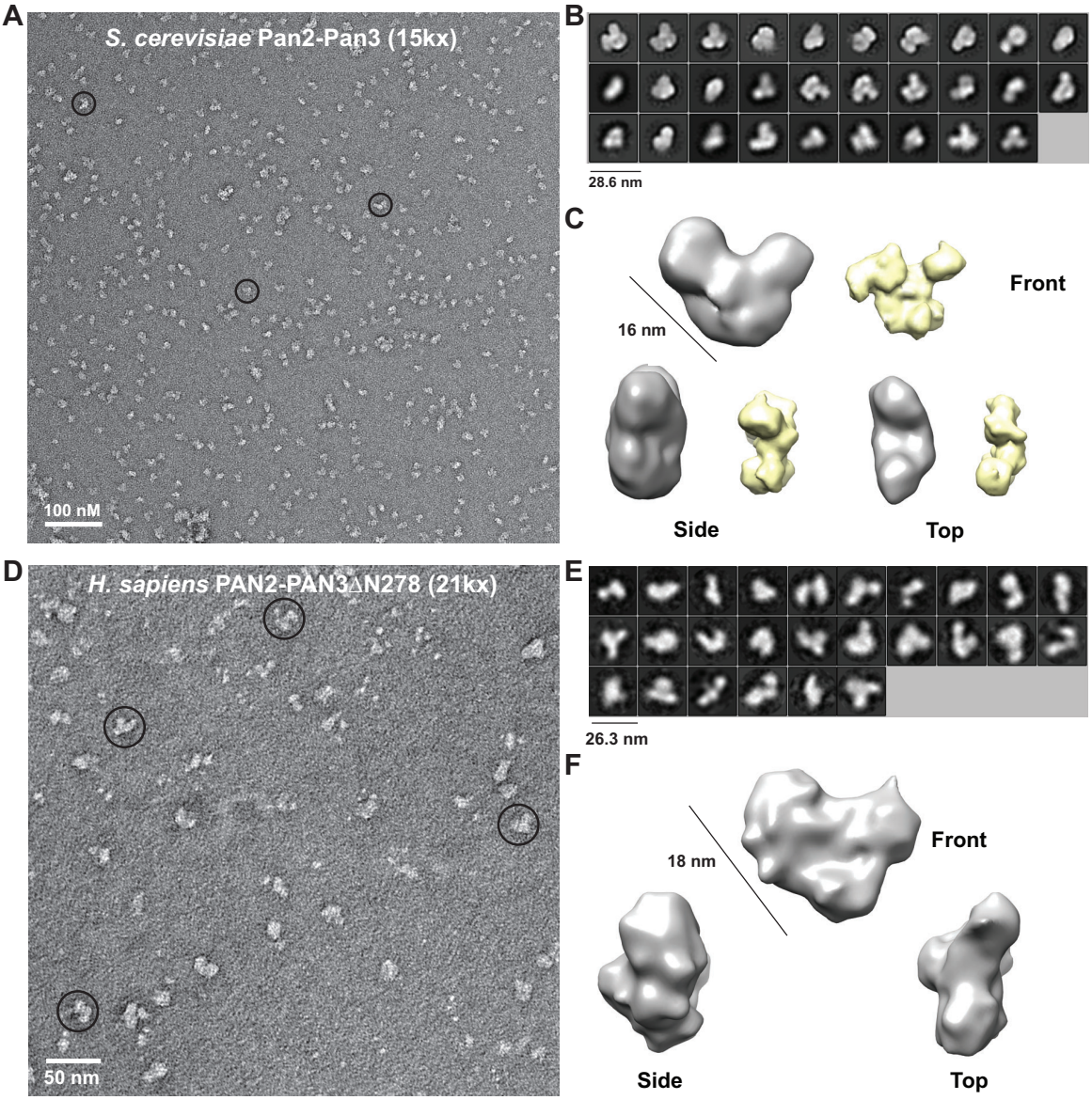


Figure 3.20 Negative stain EM of *S. cerevisiae* and *H. sapiens* Pan2-Pan3. Representative micrographs of (**A**) 30 nM BS3-crosslinked *S. cerevisiae* Pan2-Pan3 and (**D**) 40 nM *H. sapiens* PAN2-PAN3ΔN278. Scale bars are shown in the bottom left. Particles of the complex are circled in black. 2D class averages of (**B**) *S. cerevisiae* Pan2-Pan3 and (**D**) *H. sapiens* PAN2-PAN3. Scale bars indicate the box size used in classification. The averages shown are a selection from two rounds of classification with 50 classes. Grey 3D maps of (**E**) *Sc*Pan23 and (**F**) *Hs*PAN23, showing "front", "side", and "top" views of the core complex. Lines indicate the maximum particle dimension of the reconstruction. The yellow map is a 20-Å low-pass filtered map of the predicted composite structure shown in Section 1.2.2.

3.11 Cryo-EM of Apo PAN2-PAN3

During my PhD, a cryo-EM study detailed the 4.5 Å structure of *S. cerevisiae* Pan2-Pan3 (Schafer et al., 2019). In that study, truncation of the Pan3 low-complexity N-terminus resulted in higher resolution, suggesting that intrinsically disordered regions may be detrimental to cryo-EM. Pan2-Pan3 is tri-lobed and has a similar structure as the proposed model (Section 1.2.2). However, as there was no structural information of human PAN2-PAN3, how PAN2-PAN3 contacted PABPC1-poly(A) and how this may be related to the different poly(A) tail shortening properties between human and yeast, were unknown. To obtain higher-resolution information on human PAN2-PAN3, cryo-EM was performed on two complexes: PAN2-PAN3ΔN278, which retains significant low-complexity region and contains the PAM2 motif; and PAN2-PAN3ΔN460 (PKC), which lacks the N-terminal low-complexity region.

Uncrosslinked PAN2-PAN3ΔN278 was imaged on unmodified UltrAuFoil grids. At 1.5 μM, PAN2-PAN3ΔN278 partially aggregated (Figure 3.21A, red circles). At lower concentrations, most holes were empty and I observed aggregation at hole edges (Figure 3.21A, inset). This suggests that PAN2-PAN3ΔN278 was prone to aggregation, consistent with previous studies of uncrosslinked *S. cerevisiae* Pan2-Pan3 on UltrAuFoil grids. Several strategies were attempted to overcome sample aggregation. Firstly, a buffer differing in salt concentration (150 mM vs. 250 mM), salt identity (50 mM NaCl + 100 mM KCl vs. 250 mM NaCl), and pH (7.5 vs. 8.0), was used. This buffer was used in the study which yielded the cryo-EM structure of *S. cerevisiae* Pan2-Pan3 (Schafer et al., 2019). These changes did not affect overall particle distribution, with significant aggregation observed in grid holes. In addition, two detergents were separately added to protein buffer (0.04%, 0.25% (v/v) n-octyl β-glucoside; 0.25% (v/v) CHAPSO). Detergent addition caused the formation of larger aggregates at various protein concentrations, suggesting that it was not useful in relieving protein aggregation (Figure 3.21B, red circle).

Moreover, changes in surface tension of the protein buffer caused uneven ice distribution and grid holes with empty centres (Figure 3.21B, inset). Thus, for uncrosslinked PAN2-PAN3, no tested condition provided micrographs of sufficient quality for data collection.

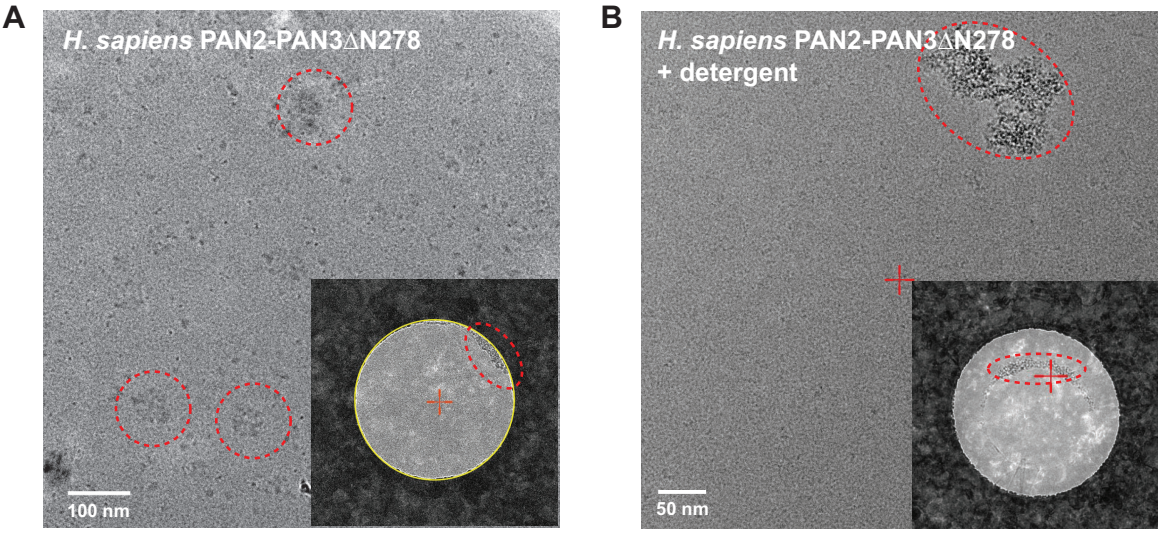


Figure 3.21 Uncrosslinked PAN2-PAN3 aggregates during cryo-EM grid preparation. (**A**) 1.5 μM uncrosslinked PAN2-PAN3ΔN278. Red circles indicate protein aggregation in grid holes. **Inset:** low-magnification view showing protein aggregation on the hole edge. (**B**) 700 nM uncrosslinked PAN2-PAN3ΔN278 in buffer containing 0.25% n-octyl-β-glucoside. Red circles indicate protein aggregation. **Inset:** low-magnification view shows altered ice distribution in grid holes and protein aggregation at the air-water interface.

To alleviate protein aggregation during sample preparation, homogeneously crosslinked PAN2-PAN3ΔN278 (Section 3.9) was imaged on unmodified UltrAuFoil grids. Compared to un-crosslinked sample, crosslinking improved particle distribution (Figure 3.22A). Firstly, at a lower concentration (1.2 μM), grid holes contained individual particles (green circles). Secondly, fewer aggregates were observed in grid holes. These results suggest that BS3 crosslinking altered surface properties of PAN2-PAN3 such that it was less prone to aggregation during grid preparation. A dataset was collected on this sample (1951 micrographs). Owing to its small size and flat architecture (Section 3.10), high defocus (-3.0 to -5.5 μm) was required to observe particles. Three rounds of 2D classification of Cryolo-picked particles (Wagner et al., 2019) resulted in 2D class averages (Figure 3.22B). These contained representative "front" views with the characteristic tri-lobed shape, as well as elongated class averages which could

represent "top" or "side" views. 2D class averages were approximately 16-18 nm, consistent with 3D maps obtained from negative stain EM (Figure 3.21F). Thus, the particles likely represent *bona fide* PAN2-PAN3 core complexes. However, 2D class averages contained no higher-resolution information (on the secondary structural level). Thus, it is possible that PAN2-PAN3 was crosslinked in multiple similar but non-homogeneous states, PAN2-PAN3 retained structural flexibility even when covalently crosslinked, or portions of the complex were denatured during sample preparation, for example at the air-water interface.

To assess sample homogeneity and improve particle picking accuracy, a dataset was collected on BS3-crosslinked PAN2-PAN3PKC using a phase plate (Danev et al., 2014). As the Pan3 low-complexity N-terminus was detrimental to *S. cerevisiae* Pan2-Pan3 cryo-EM analysis, I reasoned that further truncation of PAN3 may improve the sample. In micrographs, in addition to particles of the correct size, small polypeptides were observed (Figure 3.22C). As PAN2-PAN3 is covalently linked, these may represent individual domains, suggesting local denaturation during sample preparation. This would give rise to structural heterogeneity, which could explain the lack of higher-resolution features. 2D class averages were of similar quality to the previous sample and secondary structure could not be resolved (Figure 3.22D).

Another way of altering surface properties to improve a sample is by coating the grid with a thin carbon layer. UltrAuFoil grids were coated with graphene oxide using standard methods. Compared to unmodified grids, much less PAN2-PAN3PKC (200 nM) was required for good particle distribution (Figure 3.22E), suggesting that PAN2-PAN3PKC adsorbed to graphene oxide, potentially preventing denaturation by interaction with the air-water interface. However, larger data collections have not yet been performed due to the difficulty of grid reproduction. A very small dataset resulted in 2D class averages which did not display high-resolution information (Figure 3.22F). Thus, graphene oxide deposition may not be useful for data collection, but more data must be collected with/without crosslinking for conclusive evidence.

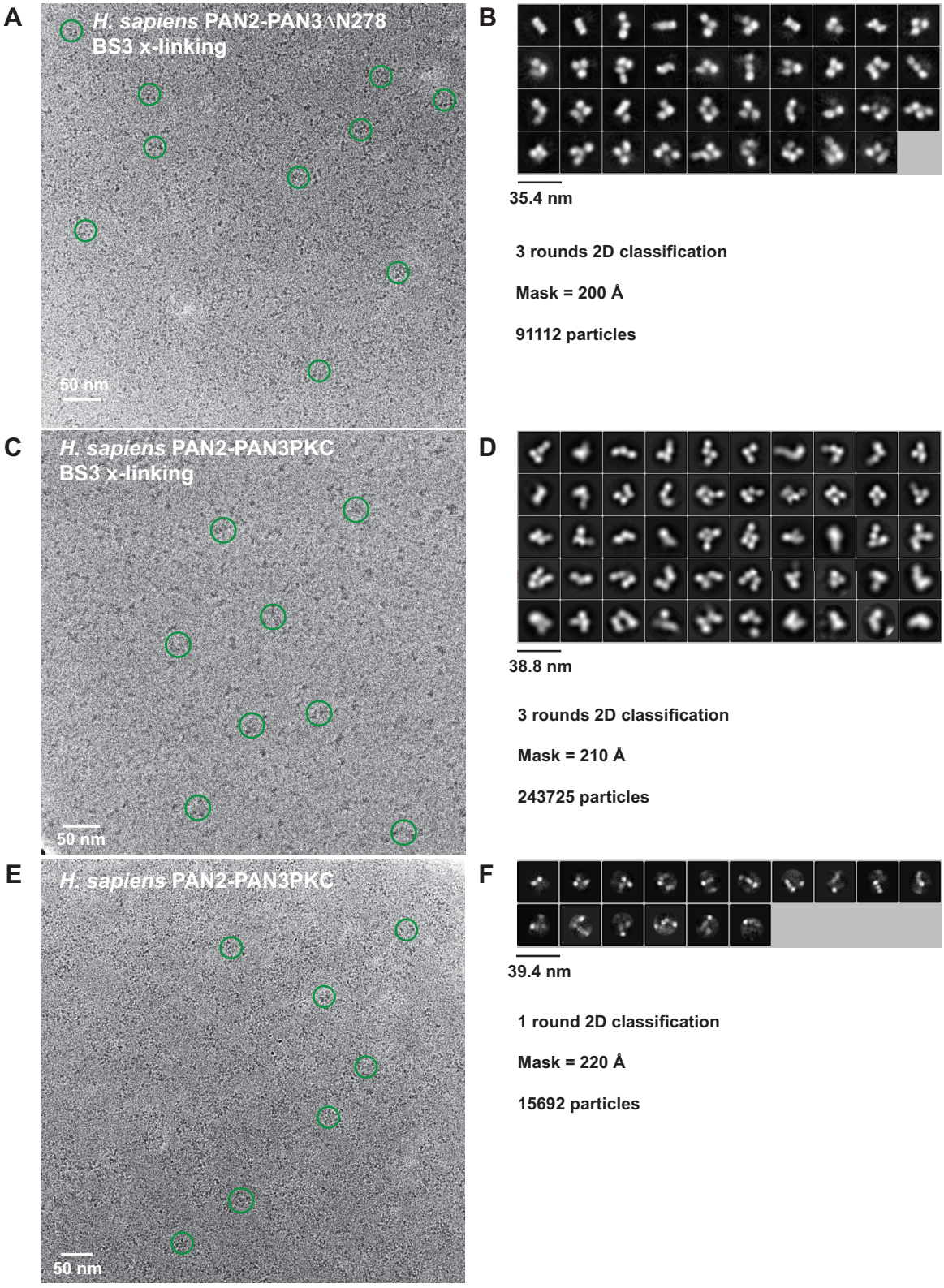


Figure 3.22 Cryo-EM of PAN2-PAN3. Representative micrographs of (**A**) 1.2 μM BS3-crosslinked PAN2-PAN3ΔN278; (**C**) 1.0 μM BS3-crosslinked PAN2-PAN3PKC (phase plate); and (**E**) 200 nM uncrosslinked PAN2-PAN3PKC (graphene oxide). Green circles indicate individual PAN2-PAN3 particles. Representative 2D class averages of (**B**) 1.2 μM BS3-crosslinked PAN2-PAN3ΔN278; (**D**) 1.0 μM BS3-crosslinked PAN2-PAN3PKC; and (**F**) 200 nM uncrosslinked PAN2-PAN3PKC.

3.12 MEX3 is a Sequence-specific RBP

Proteins and protein complexes, including PAN2-PAN3, generally do not act in isolation. Instead, they often require other binding partners to carry out their function. Despite its extensive biochemical characterisation, PAN2-PAN3 binding partners have not been studied in detail, with the notable exception of PABPC1. Thus, I sought to identify interacting partners of PAN2-PAN3 which may regulate its function. A recent proximal labelling study identified the RBP MEX3 as a potential interactor of PAN2-PAN3 (Youn et al., 2018). As MEX3 is a repressor of gene expression and binds specifically to the MRE sequence, it could recruit PAN2-PAN3 for deadenylation of targeted transcripts. Therefore, I investigated the effect of MEX3 proteins on PAN2-PAN3 in *in vitro* deadenylation assays. A recombinant expression and purification protocol for the four paralogous human MEX3 proteins was developed. SII-tagged MEX3 was overexpressed by baculovirus-infected Sf9 cells. Protein was purified by a four-step protocol, including overnight cleavage of the SII tag by 3C protease (Figure 3.23). MEX3D was not purified due to the difficulty of overexpression in Sf9 cells.

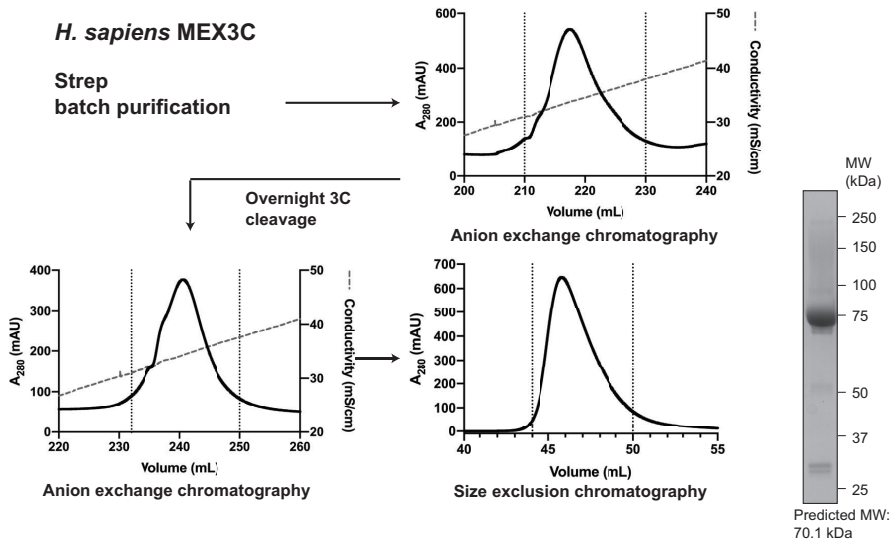


Figure 3.23 Purification of *H. sapiens* MEX3 proteins. Purification was carried out by a four-step protocol with overnight 3C cleavage. An SDS polyacrylamide gel of the purified protein is shown. Collected fractions are denoted by dotted lines.

To test if purified MEX3A-C are *bona fide* RBPs, binding assays were performed using a 6-FAM-labelled RNA with the consensus MEX3 binding site (MRE) AGAGUUUA (Pagano et al., 2009). EMSAs using a 6-FAM-labelled MRE-$X_{20}A_{30}$ RNA showed that MEX3A, MEX3B, and MEX3C bound these RNAs with roughly 100 nM affinity (Figure 3.24A, C, E). This is slightly lower affinity than the previously measured 36 nM for *C. elegans* MEX-3 (Pagano et al., 2009). To confirm the binding affinities measured by EMSAs, fluorescence polarisation assays (FPAs) were performed in triplicate. Notably, at higher protein:RNA ratios, MEX3 bound non-specifically to RNA; MEX3A-C had ~1 μM affinity for an A_{22} RNA (data not shown). Binding curves could be fit to fluorescence polarisation data if points for high MEX3 concentrations were treated as outliers, resulting in affinities of between 30-90 nM, in line with EMSA observations and the previous measurement (Figure 3.24B, D, F). Notably, the affinity for MRE-X_{20} is similar to that for MRE-$X_{20}A_{30}$, suggesting that MEX3 bound to its target site in the upstream non-A sequence with approximately 100 nM affinity. Thus, our results demonstrate that recombinantly expressed and purified MEX3A-C proteins are *bona fide* sequence-specific RNA-binding proteins.

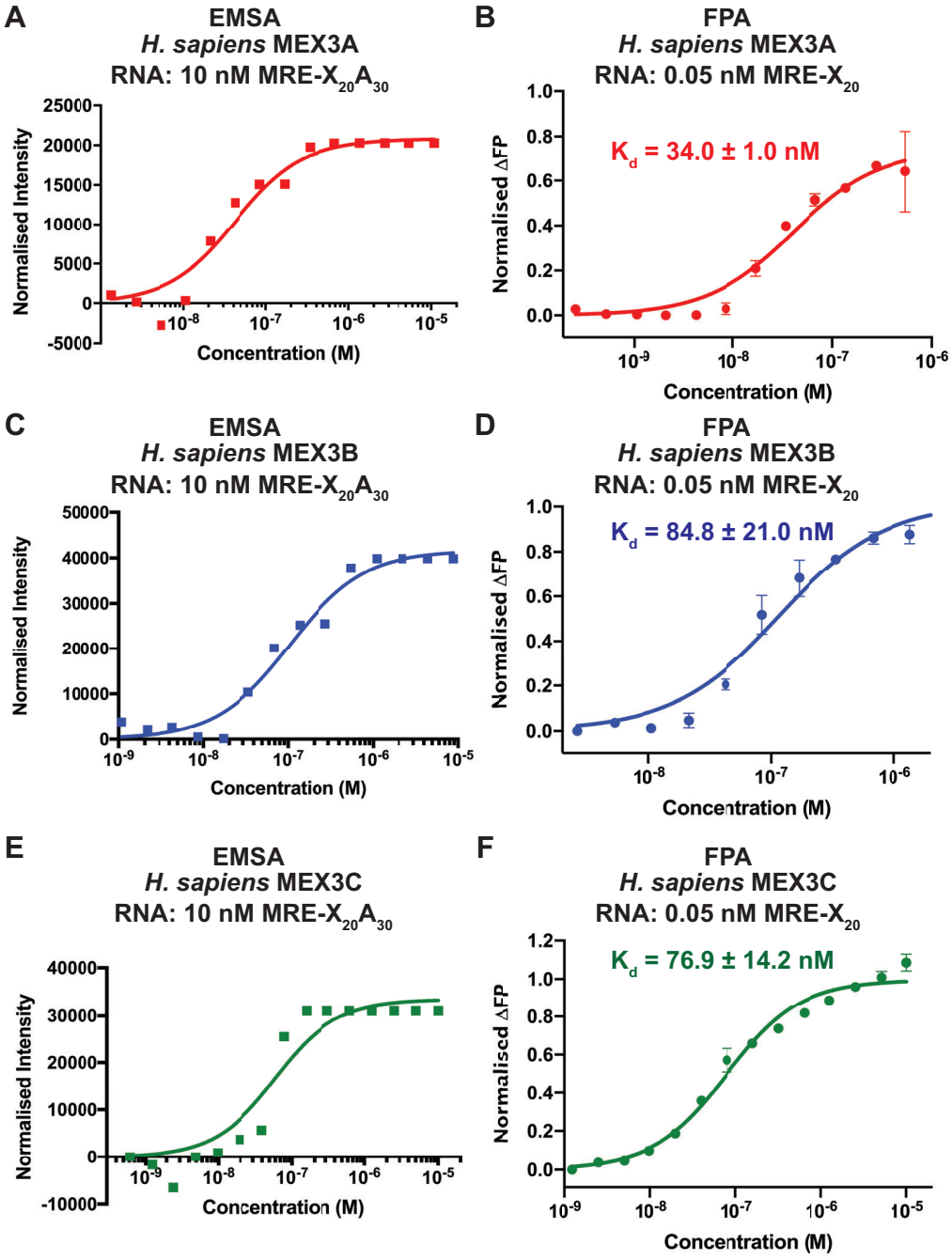


Figure 3.24 Recombinant MEX3 is a *bona fide* RNA-binding protein. RNA binding affinities as measured by EMSAs (**A, C, E**) and FPAs (in triplicate; **B, D, F**). Binding to MRE-X_{20} RNA by (**A-B**) MEX3A; (**C-D**) MEX3B; and (**E-F**) MEX3C. All points represent means and error bars are standard deviation. K_d estimates are shown next to the curves, which were fit to the points using a quadratic function. The number following K_d represents 95% confidence intervals of the K_d estimate.

3.13 MEX3 Stimulates PAN2-PAN3 Activity

To test the effect of MEX3 on PAN2-PAN3, purified MEX3 was pre-incubated with MRE-$X_{20}A_{30}$ RNA and used in *in vitro* deadenylation assays. Compared to when MEX3 was absent (Figure 3.25A), MEX3A, MEX3B, and MEX3C modestly stimulated full-length PAN2-PAN3 (Figure 3.25B-D). PAN2-PAN3 was stimulated approximately 1.5× (compare 70-80 min of +MEX3 to 110 min of -MEX3). However, MEX3 did not increase PAN2-PAN3 processivity, which would be expected if MEX3 strongly bound PAN2-PAN3 and recruited it to the target RNA. This modest stimulation may have been due to the relatively low purity of MEX3 used in assays due to suboptimal purification, the lack of saturation of MRE-$X_{20}A_{30}$ RNAs by MEX3, unknown post-translational modifications on MEX3 or PAN2-PAN3, or low affinity between MEX3 and PAN2-PAN3. Thus, full understanding of stimulation can only be performed when purification and RNA-binding characterisation are improved.

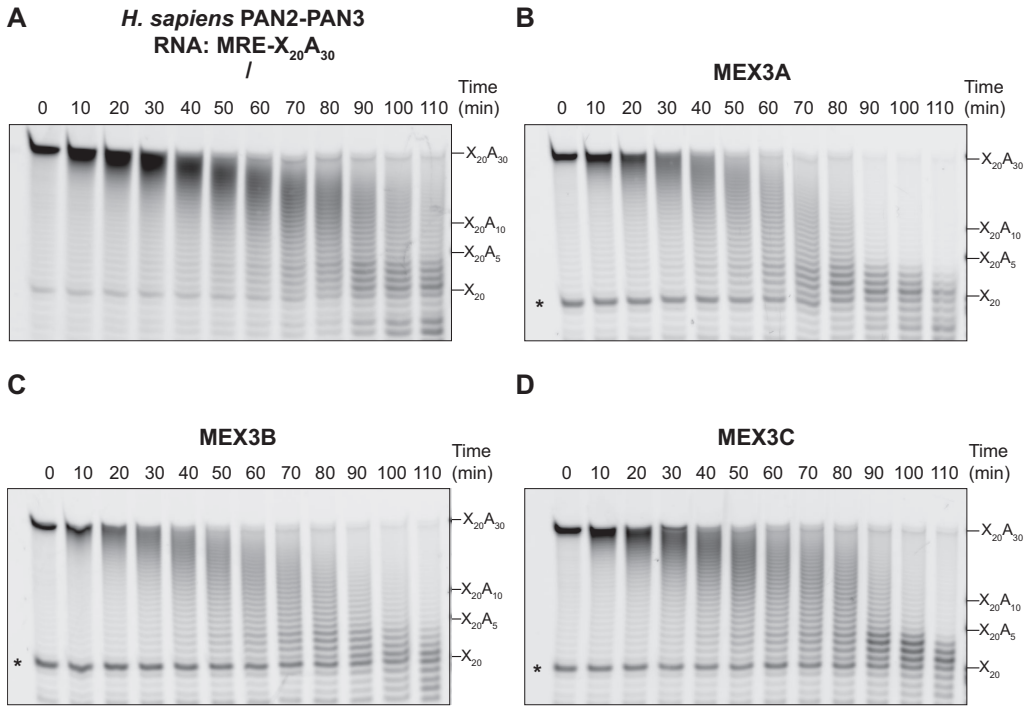


Figure 3.25 MEX3 stimulates full-length PAN2-PAN3. Deadenylation of 200 nM MRE-$X_{20}A_{30}$ by 200 nM *H. sapiens* PAN2-PAN3 with (**A**) no MEX3; (**B**) 200 nM MEX3A; (**C**) 200 nM MEX3B; and (**D**) 200 nM MEX3C. Black asterisks denote a contaminating RNA.

To determine if the low-complexity N-terminus of PAN3 is required for stimulation, I tested the effect of MEX3C on deadenylation by PAN2-PAN3ΔN278 and PAN2-PAN3PKC, N-terminal truncations of PAN3 which remove half or all of the low-complexity region. Compared to when MEX3 was absent (Figure 3.26A, C), MEX3C strongly stimulated both PAN2-PAN3ΔN278 and PAN2-PAN3PKC activity (Figure 3.26B, D). Notably, the stimulatory effect was 4-5× for N-terminal truncated constructs, compared to the 1.5× stimulation for full-length PAN2-PAN3. There are several potential explanations for this discrepancy. For example, inclusion of the PAN3 N-terminal zinc finger may compete for MRE binding, attenuating the stimulatory effect; alternatively, MEX3C may functionally substitute for the PAN3 zinc finger, tethering the N-terminally truncated PAN2-PAN3 complexes to the poly(A) substrate. These results suggest that the N-terminal half of the low-complexity region of PAN3 is not required for stimulation of PAN2-PAN3 activity by MEX3.

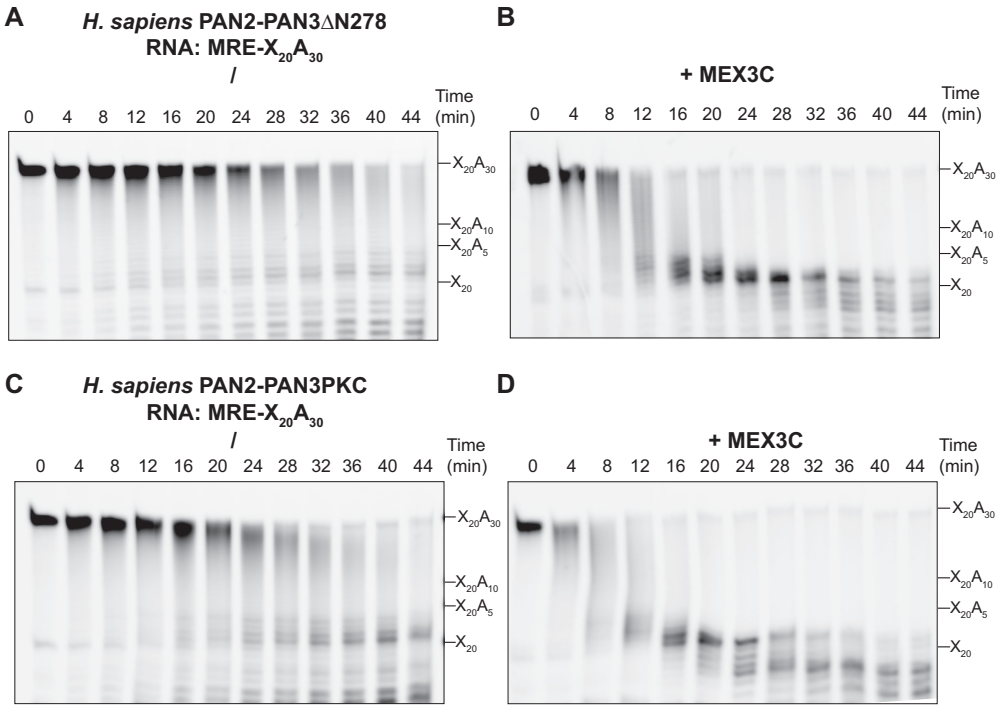


Figure 3.26 MEX3 stimulates N-terminal truncations of PAN2-PAN3. Deadenylation of 200 nM MRE-$X_{20}A_{30}$ by 100 nM *H. sapiens* PAN2-PAN3ΔN278 with (**A**) no MEX3; or (**B**) 100 nM MEX3C. Deadenylation of 200 nM MRE-$X_{20}A_{30}$ by 100 nM *H. sapiens* PAN2-PAN3ΔPKC with (**A**) no MEX3; or (**B**) 100 nM MEX3C. MEX3C stimulates deadenylation by N-terminal truncations of PAN2-PAN3 by 4-5×.

Finally, to test if MEX3 can stimulate the *H. sapiens* PAN2 catalytic unit, deadenylation assays were repeated with MEX3C and the *H. sapiens* PAN2UCH-Exo. The addition of MEX3C had a negligible effect on deadenylation by the PAN2 catalytic unit. This suggests that MEX3-dependent stimulation of PAN2-PAN3 activity requires either PAN3 or the N-terminus (the WD40 domain and PAN3-interacting linker) of PAN2. It remains to be seen if other MEX3 paralogues are able to stimulate the catalytic unit.

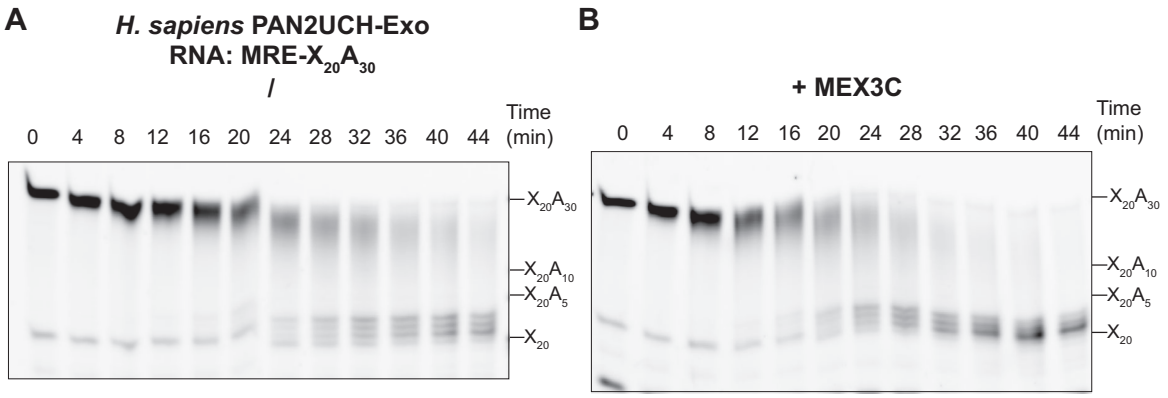


Figure 3.27 MEX3 does not stimulate the PAN2 catalytic unit. Deadenylation of 200 nM MRE-X20A30 by 100 nM H. sapiens PAN2UCH-Exo with (A) no MEX3; or (B) 100 nM MEX3C.

3.14 MEX3 Interacts with PAN2-PAN3

One possible mechanism by which MEX3 can stimulate PAN2-PAN3 is by directly recruiting the deadenylase complex to target RNAs. To test if PAN2-PAN3 and MEX3 directly interact, pull-down assays were performed. In these assays, SII-tagged PAN2-PAN3 was immobilised on magnetic StrepTactin beads and incubated with MEX3 proteins. As MEX3B was not visible as a clearly defined band, it was omitted from pull-down assays. Pull-down assays showed that full-length PAN2-PAN3 could interact with MEX3A and MEX3C, albeit in sub-stoichiometric amounts (Figure 3.28A, right panel, lanes 3 & 5). Given the sub-stoichiometric amount of MEX3 pulled down, this interaction is likely weak. As RNAs were not included in pull-down assays, MEX3 binding to PAN2-PAN3 could be partially mediated by RNA.

To determine if the truncated PAN2-PAN3ΔN278 retains the ability to interact with MEX3, pull-down assays were repeated with MEX3C (Figure 3.28B). MEX3A overlaps with degradation products (asterisk) from PAN2-PAN3 purification and thus pull-down assays were inconclusive. These assays demonstrated that PAN2-PAN3ΔN278 directly interacts with MEX3C in sub-stoichiometric amounts (Figure 3.28B, right panel, lane 3). Thus, at least for MEX3C, the N-terminal half of the PAN3 low-complexity region is dispensable for interaction. This is in line with the observation that deadenylation by the PAN2-PAN3ΔN278 construct can be stimulated by MEX3. Taken together, these results of direct interaction correlate with the ability of MEX3 RBPs to stimulate PAN2-PAN3 activity. This agrees with the model whereby PAN2-PAN3 activity is stimulated by direct recruitment of the deadenylase complex to target RNAs. The domain requirements of both PAN2-PAN3 and MEX3 of this interaction remain to be investigated, and further truncations of both species will be required to identify the minimal interacting complex.

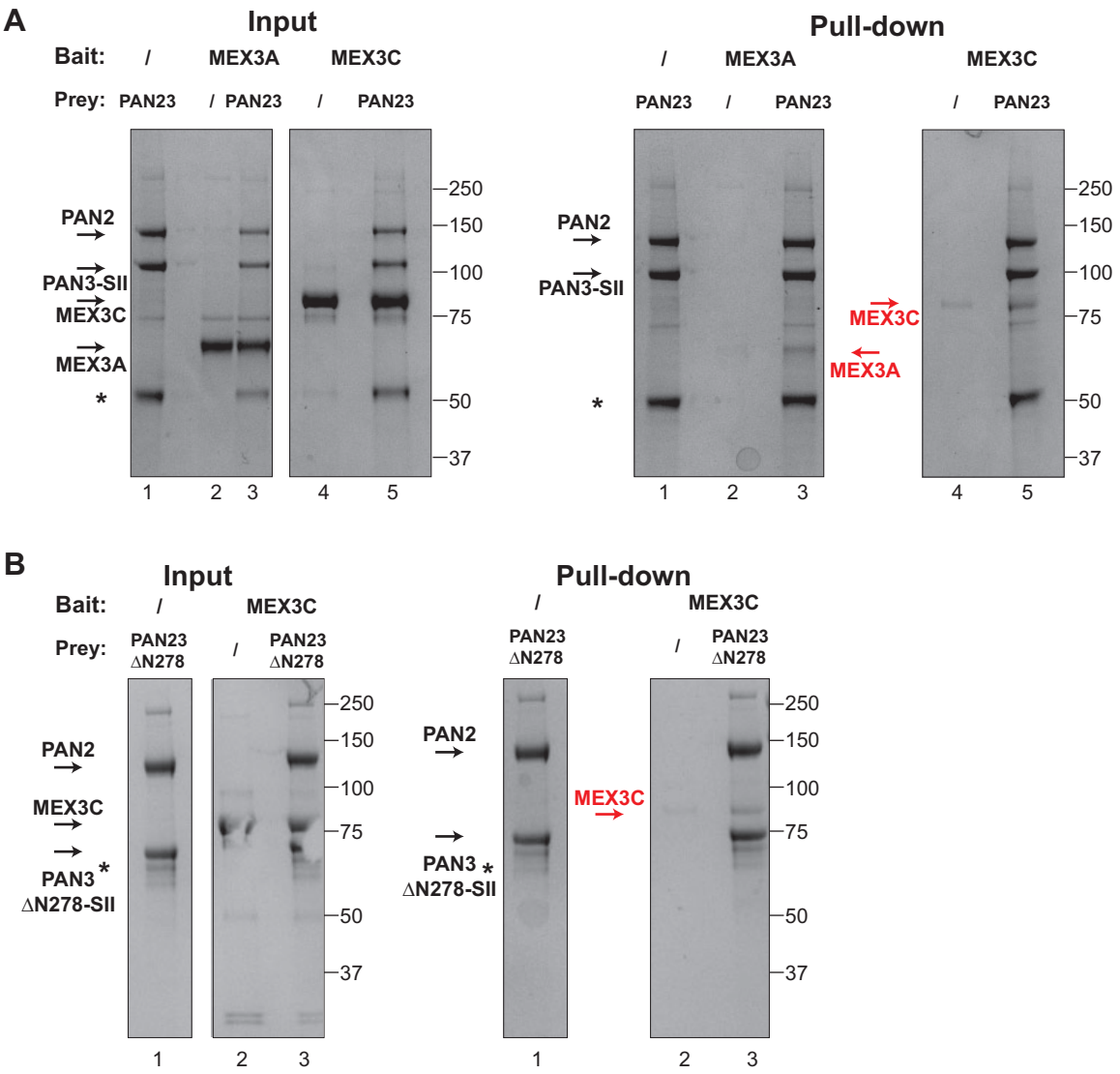


Figure 3.28 Pull-down assays show that *H. sapiens* MEX3 directly interacts with *H. sapiens* PAN2-PAN3. (**A**) Immobilised full-length PAN2-PAN3 (lane 1) was tested for interaction with MEX3A (lane 2-3) and MEX3C (lane 4-5). Pull-downs show that MEX3A and MEX3C can be pulled down in sub-stoichiometric amounts (lane 3, 5) by PAN2-PAN3 compared to control (lane 2, 4). (**B**) Immobilised PAN2-PAN3ΔN278 (lane 1) was tested for interaction with MEX3C (lane 2-3). Pull-downs show that MEX3C can be pulled down in sub-stoichiometric amounts (lane 3) by PAN2-PAN3ΔN278 compared to control (lane 2). The asterisks indicate contaminating PAN3 degradation products.

3.15 Discussion

3.15.1 Studying Pan2-Pan3 Deadenylation by *In Vitro* Deadenylation Assays

In this chapter, I have extensively investigated deadenylation by yeast and human Pan2-Pan3 complexes, as well as their regulation by RNA-binding proteins, using *in vitro* deadenylation assays. Conditions for *in vitro* assays were largely selected due to practical reasons, including the yield of purified enzyme, assay time, and detection of the fluorescently labelled substrate. While most assays were carried out comparatively (and absolute kinetic parameters were not measured), their interpretation were nonetheless limited in several ways.

One limitation in the reaction setup is the usage of only one enzyme concentration, and thus a fixed enzyme:substrate ratio. It is thus difficult to draw conclusions regarding enzymatic processivity (which, in a multi-turnover system, is determined by the ratio of the rate constant for degradation to the probability of dissociation at each nucleotide) and how this is affected by additional components. One possible experiment to test processivity is a pulse-chase experiment under pre-steady-state conditions (enzyme > substrate). The deadenylase would be incubated with a fluorescently labelled "hot" substrate, and would be chased later by an unlabelled "cold" substrate; a processive enzyme would continue deadenylation of the hot substrate due to its low probability of dissociation. These experiments have been carried out for biological systems such as the synthesis of telomeric repeats by the telomerase (Chen and Greider, 2003), and can be adapted for deadenylation assays to assess changes in processivity.

Another limitation in these assays is the inability to fully analyse kinetic parameters of a single deadenylation event, that is, the kinetics of removing one A from the poly(A) tail. For an enzyme to act on a particular substrate, it must bind the substrate, catalysis must occur, and

the product must dissociate from the enzyme (reinitiating the catalytic cycle). In the current assay setup, these steps and their individual rate constants cannot be deconvoluted. A further complication is that the product generated (an RNA with one fewer A) remains a substrate and is further processed. Thus, characteristics such as the rate-limiting step of catalysis, or whether the protein alone displays different kinetic regimes dependent on poly(A) tail length and/or upstream sequence (Bisaria et al., 2017), cannot be determined. A pre-steady-state approach to follow product formation during the first turnover event, combined with large analytical denaturing gels for single-nucleotide resolution and possibly a phosphorothioate linkage at the penultimate ribophosphate bond to prevent further deadenylation, may elucidate these complex enzymatic parameters (Sassa et al., 2013). Precise elucidation of the specificity and kinetic parameters of deadenylation of a complex substrate mixture, emulating what deadenylase complexes would encounter *in vivo*, would provide valuable insight into deadenylation behaviour and preferences. Similar studies have been carried out for enzymes such as RNase P and could be applied to deadenylase complexes (Yandek et al., 2013).

Finally, deadenylation is a complex multi-turnover enzymatic process. As the poly(A) tail is shortened, there is an increasing diversity of potential competing substrates in the reaction, dramatically increasing the complexity of binding/dissociation and catalytic rates. This observation is further complicated by the presence of differently deadenylated species, particularly when RBPs such as Pab1 are present (for example in Figure 3.8C or Figure 3.11B). One possible way to approach this complexity is to distinguish between different species by plotting RNA tail length distributions against time and thereby characterise each population separately under distributive assay conditions (steady-state Michaelis-Menten conditions). Alternatively, improvements in modelling of enzyme behaviour, such as that carried out for the exosome exonuclease Rrp6 (Axhemi et al., 2020) or the TRAMP complex (Jia et al., 2011) may further our understanding of deadenylase function (Choi et al., 2017).

3.15.2 RNA-binding Characteristics of PABPC1

Recombinant overexpression and purification of Pab1 has allowed species-specific differences in RNA binding to be elucidated. The RNA-binding properties of *H. sapiens* PABPC1 contrast with those of *S. cerevisiae* Pab1. Firstly, PABPC1 binds oligo(A) RNAs with 5-50× higher affinity than *S. cerevisiae* Pab1 (Section 3.3.2). This difference is smaller (~2×) for longer oligo(A). Thus, there may be additional, or different, interactions of human PABPC1 with oligo(A) RNA compared to that of yeast Pab1 and oligo(A), giving rise to the higher affinity.

In addition, *H. sapiens* PABPC1 binds RNA with significant positive cooperativity compared to yeast Pab1, which can be attributed to the PABPC1 C-terminus (Section 3.4). Together with the recent structure of Pab1-poly(A) RNP, these results suggest that the helix following RRM4 may be involved in cooperative RNA binding (Schafer et al., 2019). However, as the PABPC1 P- linker also contains low-complexity regions, often implicated in multimerisation, multiple regions could be required for cooperativity. Further PABPC1 constructs with different C-terminal truncations could elucidate the requirements of cooperativity.

Finally, it is unknown if different RNA binding characteristics are related to the different *in vivo* characteristics of yeast and metazoan Pab1. Firstly, metazoa have longer steady-state median poly(A) tails compared to yeast (Section 1.3.1). The higher cooperativity of PABPC1 may allow multiple molecules to be simultaneously bound on a newly synthesised poly(A) tail, thereby supporting and stabilising longer poly(A) tails. Additionally, multiple PABPC1 molecules would offer the opportunity of each individual molecule being differentially post-translationally modified or bind to different proteins. Thus, PABPC1 multimerisation would allow processes such as translation and RNA decay to be more finely regulated, consistent with the more complex mechanisms of regulating gene expression in higher eukaryotes.

Binding of single-stranded poly(A) by the highly conserved Pab1 and its transient dissociation during poly(A) processing are reminiscent of the mechanism of single-stranded DNA (ssDNA) binding by the highly conserved protein RPA. RPA is required for processing of ssDNA intermediates found during DNA replication, repair, and recombination (Chen and Wold, 2014). Similar to Pab1, RPA binds ssDNA via tandem DNA-binding domains with subnanomolar affinity (Kim et al., 1994; Kim and Wold, 1995). In line with the high affinity for ssDNA and low overall k_{off} (Patrick and Turchi, 2001), single-molecule studies have demonstrated that RPA remains stably bound to ssDNA and rarely dissociates; however, when other ssDNA-binding proteins or processing machinery are present, RPA can readily dissociate via an exchange-dependent model (Gibb et al., 2014). Moreover, FRET studies demonstrate that RPA is able to diffuse along single-stranded DNA without full dissociation due to its modular DNA-binding domains (Nguyen et al., 2014). Thus, despite its high affinity for ssDNA, RPA binding is highly dynamic and can be remodelled in the presence of other machinery or DNA-binding proteins. It would be intriguing to study if the tandem RRMs of Pab1 bind poly(A) RNA analogously and can be similarly remodelled or maintained by poly(A)-processing proteins or other poly(A)-binding proteins.

3.15.3 Regulation of Pan2-Pan3 by Pab1

3.15.3.1 Inhibition of Pan2-Pan3 by Pab1

Biochemical dissection of Pab1 regulation of Pan2-Pan3 using yeast and human proteins has revealed some common features. Firstly, Pan2-Pan3 is inhibited by stably bound Pab1. For yeast, this occurs when the poly(A) tail is ~30 A's, which can stably accommodate one Pab1 molecule (Figure 3.8). For humans, however, the point of inhibition depends on PABPC1:poly(A) stoichiometry; stably-bound PABPC1, in phases of 27-28 A's, cannot be removed by PAN2-PAN3 (Figure 3.10). Stability of binding may be related to cooperativity of RNA binding; however, PABPC1 RRM1-4, which does not bind cooperatively to long poly(A) RNA, also exhibits stoichiometry-dependent inhibitory effects on PAN2-PAN3.

The mechanism of inhibition can be rationalised by superposition of crystal structures of Pan2UCH-Exo-poly(A) (Chapter 2) and PABPC1 RRM1-2–A_{10} RNA (Safaee et al., 2012b). Structural superposition was carried out at the 3' end of the RNA, which is bound by Pab1 RRM1 and is present in the Pan2 exonuclease active site (Figure 3.29A). Superposition showed steric clashes between Pab1 RRM1 and the Pan2 active site. Thus, when poly(A) is tightly bound by Pab1 RRMs, it cannot be removed by Pan2. These results are consistent with *in vitro* deadenylation assays showing that the RRMs are sufficient to inhibit Pan2-Pan3 or the Pan2 exonuclease domain (Figure 3.12, Figure 3.18). This model can be confirmed using RNA-binding mutants of RRM1, which should inhibit Pan2-Pan3 activity less than wild-type, as weaker RRM1 binding should allow Pan2-Pan3 to more readily remove the poly(A) tail.

Biochemical evidence also suggests that Pab1 may be differently bound on the poly(A) tail during deadenylation (Figure 3.29B). Under non-saturating Pab1 concentrations and poly(A) tail lengths that can support one or more Pab1 molecules, Pab1 could freely bind along the poly(A) tail. As the poly(A) tail is opportunistically shortened, only RRM4 can bind to the

upstream non-A sequence, such that only RRM1-3 is bound to the poly(A) tail. This would give rise to the shorter than expected footprints observed during PAN2-PAN3 deadenylation of PABPC1-poly(A) (Figure 3.18), consistent with previous observations of *S. pombe* Ccr4-Not deadenylation of Pab1-poly(A) (Webster et al., 2018). At this point, PABPC1 can no longer freely bind along the poly(A) tail, causing a strong stall of PAN2-PAN3, and the poly(A) tail can only likely be removed if the RRM domains transiently dissociate. This model could be tested, for example, by RNA-binding mutants of RRM4, which should cause a weaker stall at A_{22} compared to the wild-type protein. Moreover, whether RRM4 of PABPC1 displays any sequence preference, and thus may be enriched on particular transcripts and cause their poly(A) tails to be accordingly phased or stabilised, requires further investigation. Sequencing and biochemical experiments may elucidate the relationship between Pab1 occupancy and transcript stability *in vivo*. Finally, whether Pab1 is randomly bound along the poly(A) tail could be tested by single-molecule studies of poly(A)-binding properties of Pab1.

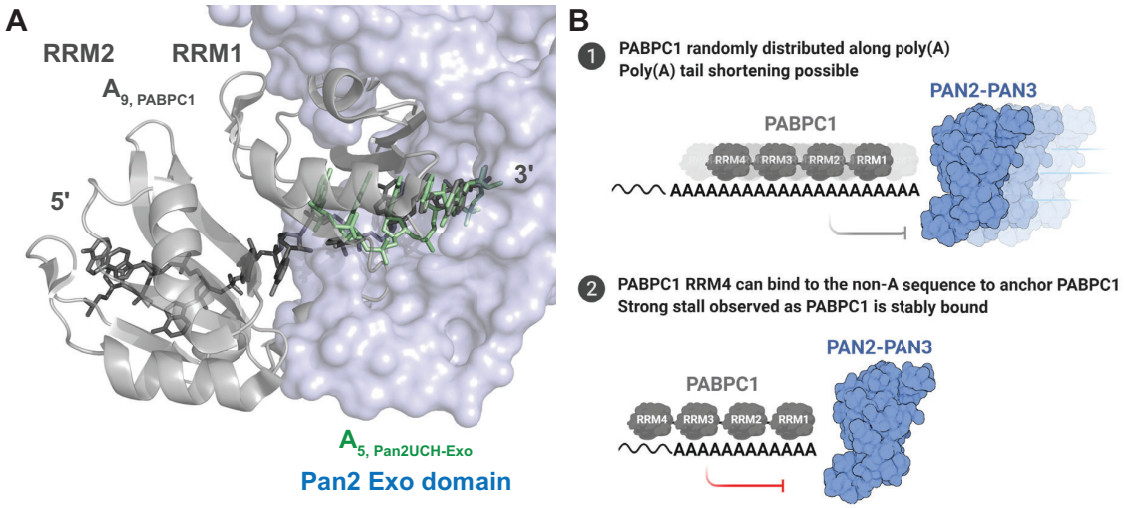


Figure 3.29 Model of inhibition of Pan2-Pan3 by Pab1. (**A**) Structural superposition of PABPC1 RRM1-2 (grey cartoon) – A_{10} (grey sticks) (PDB: 4F02) with Pan2UCH-Exo (light blue surface) – A_7 (green sticks) (PDB: 6R9J). The superposed structures show that PABPC1 RRM1 would sterically clash with Pan2UCH-Exo. (**B**) Pab1 may rearrange on the poly(A) tail during deadenylation. (**1**) PABPC1 RRMs can bind freely along a poly(A) tail when PABPC1 is at sub-saturating concentrations. Opportunistic poly(A) tail shortening by PAN2-PAN3 can occur. (**2**) Once the poly(A) tail is sufficiently shortened, the less poly(A)-specific RRM4 can bind in the UTR non-A sequence, causing strong PAN2-PAN3 stalling and subsequent PABPC1 to be phased on the poly(A) tail.

3.15.3.2 Stimulation of Pan2-Pan3 by Pab1

Pab1 can also stimulate Pan2-Pan3. Prior to my PhD, stimulation of Pan2-Pan3 activity was thought to be solely mediated by the interaction between the Pab1 CTD and the PAM2 motif of Pan3 (Mangus et al., 2004a). In both yeast and humans, full stimulation is observed only with full-length Pab1 and constructs where the Pan3 PAM2 motif is present. However, my results surprisingly demonstrate that removal of PAM2 attenuates, but does not ablate, Pab1 stimulation (Figure 3.13, Figure 3.14). Furthermore, the ability of Pab1 to stimulate Pan2-Pan3 (Figure 3.16) correlates with their ability to interact by pull-downs or by SEC (Figure 3.17). As yeast Pan2-Pan3 does not undergo significant conformational changes upon binding to the Pab1-poly(A) RNP (Schafer et al., 2019), Pab1 likely stimulates Pan2-Pan3 through a direct interaction, recruiting Pan2-Pan3 to the poly(A) substrate, rather than through an allosteric mechanism. Notably, Pan2-Pan3 and Pab1 are both components of cytoplasmic phase-separated condensates (Riback et al., 2017; Zheng et al., 2008). It is not known whether or not phase separation behaviour plays a role in Pab1 modulation of Pan2-Pan3 activity.

Nonetheless, Pab1 stimulation of Pan2-Pan3 can be different depending on species. In yeast, stimulation requires the Pab1 C-terminus; whether the helix involved in Pab1 multimerisation is sufficient to form interfaces to stably interact with and stimulate Pan2-Pan3 is yet to be tested. The RRM domains alone cannot stimulate Pan2-Pan3 (Figure 3.8). In humans, on the other hand, RRM1-4 could stimulate and stably interact with PAN2-PAN3 (Figure 3.9, Figure 3.14, Figure 3.17). This result enables a model of stimulation of Pan2-Pan3 to be devised (Figure 3.30). In this model, there are three interacting interfaces between Pab1 and Pan2-Pan3. The first is between RRMs and the Pan2-Pan3 core. In yeast, this is insufficient for a stable interaction; in humans, this interaction may involve different contacts, enabling stable complex formation. The second interface is between the Pan2-Pan3 core and multimerisation interfaces between adjacent Pab1 molecules, which have been shown to be important in

recruiting Pan2-Pan3 to the poly(A) tail (Schafer et al., 2019). Secondary structure predictions suggest that the α-helix following RRM4 is conserved between yeast and humans (Supplementary Figure 6), but the molecular basis of multimerisation and its importance in binding human PAN2-PAN3 are unknown. Finally, for full stimulation of Pan2-Pan3 activity, the Pan3 PAM2 motif binds the Pab1 CTD. While there is no structure of a complex of Pab1 CTD and Pan3 PAM2, the high conservation of PAM2 sequences suggests that the structure is likely to be similar to previously determined structures of Pab1 CTD-PAM2 (Kozlov et al., 2004; Kozlov and Gehring, 2010). The relative importance of the PAM2-CTD interaction in forming a stable ternary complex between Pan2-Pan3 and Pab1-poly(A) may be related to structural differences between yeast and human Pab1 CTDs (Section 1.4.1).

During my PhD, Schafer and colleagues hypothesised an alternative model for how Pan2-Pan3 activity is stimulated (Schafer et al., 2019). In their model, the key interface required for Pan2-Pan3 stimulation is between the Pan2-Pan3 core and the multimerisation surface between adjacent Pab1 molecules. This would provide an explanation for why Pan2-Pan3 deadenylation is slowed at 30 A's, as an A_{30} tail can only accommodate one Pab1 molecule. However, my experimental results demonstrate that the PABPC1 RRM domains alone can partially stimulate PAN2-PAN3. Furthermore, yeast Pan2-Pan3 activity is stimulated until 30 A's regardless of Pab1 stoichiometry. Finally, PAN2-PAN3 is stalled at longer tail lengths when an increased concentration of PABPC1 is incubated with the poly(A) tail. These discrepancies favour a model where multiple contacts are important in recruiting Pan2-Pan3 to the poly(A) tail, and that deadenylation is slowed when Pan2-Pan3 encounters the RRM domains of stably-bound Pab1, rather than the inability of a Pab1 monomer to stimulate Pan2-Pan3 deadenylation.

In this model, the molecular basis of these interactions could be resolved by structural investigation of the human PAN2-PAN3-PABPC1-poly(A) complex. As a minimal, stably-interacting ternary complex has been isolated (Section 3.8.3), my results establish a platform

for structural studies, particularly by cryo-EM. Preliminary cryo-EM studies of apo *H. sapiens* PAN2-PAN3 (Section 3.11) may aid grid preparation, though the surfaces of the ternary complex may be sufficiently different to avoid denaturation at the air-water interface.

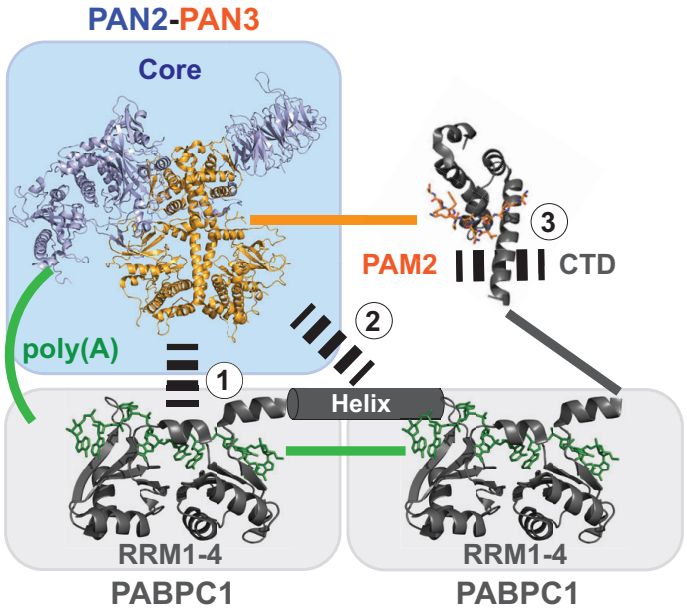


Figure 3.30 Model for interactions between PAN2 (blue cartoon)-PAN3 (orange cartoon) and PABPC1 (grey cartoon). The first interaction interface (**1**) is between RRM domains (RRM1-2 shown; PDB: 4F02) and the PAN2-PAN3 core (composite model from PDB: 4XR7, 4CYJ, and 4CZY), and has been shown by interaction studies. The second (**2**) is inferred from the structure of the yeast Pan2-Pan3-Pab1-poly(A) complex (PDB: 6R5K) and is between the PABPC1 multimerisation interface (helix-RRM1 intermolecular interaction) and the PAN2-PAN3 core. The third interaction (**3**) is between the PAM2 motif of the low-complexity PAN3 N-terminus and the PABPC1 CTD (PDB: 3KUI).

3.15.3.3 Correlation with Steady-State Poly(A) Tail Lengths

One outstanding question in the field is how steady-state cytoplasmic poly(A) tail lengths are determined, and why different species display a wide range of steady-state poly(A) tail lengths. Biochemical characterisation of the effect of Pab1 on Pan2-Pan3 activity has revealed features which may relate to the steady-state poly(A) tail lengths in different organisms.

In yeast (*S. cerevisiae* and *S. pombe*), the median steady-state poly(A) tail length is approximately 30 A's (Subtelny et al., 2014). Although poly(A) tails are shortened from their initial lengths to a steady state in the cytoplasm, the mechanistic basis of this deadenylation is unclear. In my assays, deadenylation of Pab1-poly(A) by Pan2-Pan3 is stimulated until ~30 A's, after which deadenylation is inhibited (Figure 3.8, Figure 3.10). This is independent of Pab1:poly(A) stoichiometry. It had previously been observed that Caf1 deletion (depleting both Caf1 and Ccr4 from Ccr4-Not) in yeast does not greatly affect the steady-state poly(A) tail length of ~30 A's (Webster et al., 2018). Taken together, these results suggest that the steady-state poly(A) tail length in yeast is determined by the interplay between stimulation of Pan2-Pan3 by Pab1 down to ~30 A's and inhibition of Pan2-Pan3 by Pab1 on shorter poly(A) tails (Figure 3.31A). This may be related to the low cooperativity of RNA binding by yeast Pab1 (Section 3.4), maintaining an even Pab1 stoichiometry on abundant cytoplasmic poly(A) tails.

In some species, global poly(A) tail lengths are phased in 30-A increments, similar to the binding footprint of Pab1, suggesting that Pab1 determines steady-state poly(A) tail length (Lima et al., 2017). This phasing effect can be reproduced *in vivo* when poly(A) is limiting, suggesting that poly(A) tail phasing may be a consequence of either high Pab1 concentrations or low amounts of mature, polyadenylated mRNA (Eisen et al., 2020). Using human PAN2-PAN3 and sub-saturating PABPC1 concentrations, this periodicity can be recapitulated from a constant initial poly(A) tail length, as PAN2-PAN3 is stimulated until encountering PABPC1 RRM domains (Figure 3.11). These species of different poly(A) tail lengths are likely due to

different PABPC1 stoichiometries on poly(A). However, whether Pab1 is phased on the poly(A) tail and their correlation with steady-state poly(A) tail lengths are unknown.

In some eukaryotes, median steady-state poly(A) tails are often longer than 30 A's, but are approximately multiples of 30 (Chang et al., 2014; Eisen et al., 2020; Subtelny et al., 2014). In human 3T3 cells, for example, mean mRNA poly(A) tail lengths are distributed with a ~90-A median. Poly(A) tail lengths which are multiples of PABPC1 footprints can be reproduced by PAN2-PAN3 at saturating PABPC1:poly(A) stoichiometries, where PAN2-PAN3 deadenylation is inhibited upon encountering PABPC1 (Figure 3.11). These correlations suggest that steady-state poly(A) tail length may depend on the stoichiometry of Pab1 to poly(A) and cooperativity of RNA binding, which determine the number of Pab1 molecules bound to each transcript (Figure 3.31B). Nonetheless, it is unclear if transcripts show different Pab1 occupancies, and if this correlates with transcript-specific median poly(A) tail lengths. Finally, it is unknown if Pab1:poly(A) stoichiometry is dynamically regulated, particularly in response to global shifts in gene expression such as cellular stress.

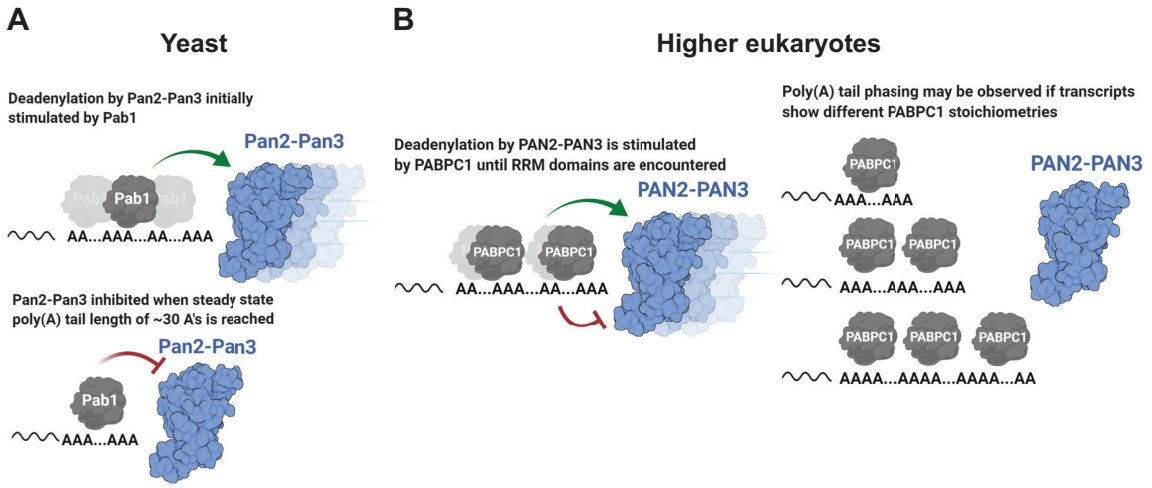


Figure 3.31 Pan2-Pan3 and Pab1 recapitulate steady-state poly(A) tails. Correlations of *in vivo* steady-state poly(A) tail lengths in (**A**) yeast and (**B**) higher eukaryotes with biochemical results obtained *in vitro* with Pan2-Pan3 and Pab1.

3.15.4 Mechanism of MEX3 Stimulation of PAN2-PAN3

Based on a previous proximity labelling study, MEX3 was identified as a potential interactor of PAN2-PAN3 (Youn et al., 2018). My results show that the MEX3 can directly interact with PAN2-PAN3 and stimulate its activity (Figure 3.25, Figure 3.26, Figure 3.28). Furthermore, this stimulation does not depend on the low-complexity N-terminus of PAN3, but requires either the PAN2 WD40 domain or the PAN3 C-terminus (Figure 3.26, Figure 3.27). I speculate that the exposed surface of PAN2 WD40 or the tryptophan binding pocket formed at the PAN3 dimer interface may be involved in the interaction with MEX3. Analogous to other RBPs which bind Ccr4-Not to mediate targeted deadenylation and mRNA decay (Section 1.4.2), stimulation likely occurs through direct recruitment of PAN2-PAN3 to target transcripts. This is the first instance where a sequence-specific RNA-binding protein has been shown to stimulate PAN2-PAN3 deadenylation activity.

Despite the direct interaction between PAN2-PAN3 and MEX3, the mechanism of interaction and its domain requirements remain unknown. Future studies using different truncations of PAN2-PAN3 and MEX3 may elucidate the minimal regions required for a stable interaction, which could aid in understanding how MEX3 and other potential RNA-binding proteins recruit PAN2-PAN3 to their cognate transcripts for deadenylation. Overall, the results that MEX3 stimulates removal of the poly(A) tail are consistent with the role of MEX3 as a repressor of post-transcriptional gene expression (Pereira et al., 2013). By stimulating removal of the poly(A) tail by PAN2-PAN3, MEX3 may target its RNA for further decay and thus shorter half-lives, as well as inhibiting translation by removal of the poly(A) tail.

4. Conclusions and Future Perspectives

In this dissertation, I sought to characterise the yeast and human Pan2-Pan3 deadenylase complexes. Two main questions were investigated in each of the results chapters:

1. What is the molecular mechanism of poly(A) recognition by the Pan2 and Caf1 deadenylase enzymes? (Chapter 2)
2. How is the deadenylation activity of Pan2-Pan3 regulated, particularly by RNA binding proteins such as poly(A)-binding protein or MEX3? (Chapter 3)

To address the questions above, I carried out *in vitro* reconstitution of deadenylation using individual proteins and enzymatic complexes, which were recombinantly expressed and purified. This was combined with biochemical, biophysical, and structural methods to understand the mechanistic basis of my observations. Below, I summarise the major findings of the work described in this dissertation, and propose future experiments which could better elucidate the roles of Pan2-Pan3 deadenylation and intrinsic single-stranded RNA structure in the regulation of gene expression.

4.1 Recognition of the Intrinsic Structure of Poly(A)

In Chapter 2, I investigated the nucleotide specificity of the individual deadenylases of Pan2-Pan3 and Ccr4-Not (Figure 4.1). Using *in vitro* deadenylation assays on RNAs containing different 3' non-A nucleotides, I found that the exonucleases Pan2, Ccr4, and Caf1 have distinct nucleotide specificities (Figure 2.2, Figure 2.4). This difference in intrinsic specificity may correlate with their proposed *in vivo* roles in the biphasic model of deadenylation, and may explain why deadenylase complexes do not continue degradation into the 3' UTR.

Non-A nucleotides had been shown to be added to poly(A) tails in a non-templated manner (Chang et al., 2014; Legnini et al., 2019; Lim et al., 2018; Liu et al., 2019). More specifically, guanosine incorporation, while infrequent, correlates with transcript stability (Chang et al., 2014). *In vitro*, single guanosines inhibited Pan2 (Figure 2.3) and isolated CNOT6L and CNOT7 (Lim et al., 2018). Tandem guanosines also inhibited the recombinantly expressed and purified deadenylase complexes (Section 2.1.2, 2.2). Guanylation thus selectively stabilises transcripts by inhibiting deadenylation, the rate-limiting step of canonical mRNA decay. Interestingly, transcripts display different guanylation frequencies, suggesting that poly(A) tail modification is regulated. While the non-canonical polymerases TENT4A and TENT4B carry out non-A addition, the mechanistic basis of their regulation is unknown. A recent study demonstrated that the RNA-binding protein ZCCHC14 binds specifically to viral RNAs, recruiting TENT4 to add 3' non-A nucleotides and thereby increase viral RNA stability (Kim et al., 2020). This reinforces the biological significance of non-A modifications in regulating gene expression. Future proteomic or RIP-seq studies may identify factors, such as ZCCHC14, which facilitate non-A modifications on specific transcripts, either constitutively or in response to different stimuli. The addition of non-A nucleotides to the poly(A) tail may be a mechanism to further increase the repertoire of post-transcriptional regulation of gene expression.

To elucidate the molecular basis of poly(A) recognition by Pan2, I determined the crystal structure of Pan2UCH-Exo bound to oligo(A), revealing that active site adenosines adopt a stacked, helical structure, with similar parameters to the predicted structure of single-stranded poly(A) (Figure 2.7). Combined with the lack of base-specific interactions, this suggests that Pan2 recognises the intrinsic structure of poly(A). My results have revealed a new paradigm of single-stranded poly(A) RNA recognition by its intrinsic structure. This structure is the result of electrostatic complementarity upon base-stacking and constraints of the ribophosphate backbone in the poly(A) helix, as changing the identity or positions of functional groups can disrupt intrinsic stacking and Pan2 activity (Figure 2.18). Additionally, the poly(A) helix appears to be important in numerous biological systems (Figure 2.22). To expand on this indirect readout model, structures of poly(A) binding macromolecules should be surveyed. Comparison of poly(A) structure across different systems could provide insight into the conservation of poly(A) through evolutionary history. Finally, other single-stranded RNAs such as poly(C) are also intrinsically structured; characterisation of these RNAs bound to cognate partners may unveil the exploitation of their structure in binding specificity. These structural investigations could reveal the fundamental role of the intrinsic conformations of single-stranded RNA in biology and the importance of indirect readout in RNA binding.

As Pan2 recognises the intrinsic poly(A) structure, disruption of this conformation inhibited Pan2 activity (Figure 2.15). Structural investigations revealed that guanylated RNAs adopted an unstacked structure in the Pan2 active site (Figure 2.9). However, as the G-containing RNA could bind to the active site, it remains unclear why guanylated RNAs cannot be cleaved by Pan2. Comparison of higher resolution structures of Pan2 in complex with oligo(A) and G-containing RNAs, preferably in the presence of catalytic metal ions, may reveal subtle changes in the position of the scissile phosphate, accounting for the inhibitory effect.

In contrast, C- and U-containing RNAs had relatively little effect on Pan2 activity, and formed the required stacked, helical structure in the Pan2 active site (Figure 2.14). However, these RNAs did not have the characteristic CD signature of oligo(A). Poly(A) likely exists as an equilibrium between a stacked helix and a random coil in solution; the stacked, helical conformation is only stabilised when the ribophosphate backbone of oligo(A) is contacted by the Pan2 active site. These results led to the hypothesis that nucleotides which are permissive for the helical conformation would be recognised by the Pan2 active site and could thus be removed. A key experiment required to further support this model is the observation of formation of the stacked, helical conformation of RNA (oligo(A) and C-/U-containing RNAs) upon Pan2 binding. Methods such as NMR or CD accompanied with titration of Pan2 to saturate RNA binding may be carried out to provide direct evidence for this model.

Superposition of the Pan2-A_7 structure with a structure of *S. pombe* Caf1 showed that the oligo(A) helix can be accommodated in the Caf1 active site (Figure 2.16). Accordingly, Caf1 was inhibited by the stacking inhibitor DHU, suggesting it may also recognise the intrinsic conformation of poly(A) (Figure 2.17). Modelling of the oligo(A) substrate revealed amino acid side chains which could form base-specific interactions with oligo(A). This prediction agrees with the higher nucleotide specificity of Caf1, as nucleotides which are permissive for stacking such as cytosines and uracils inhibited Caf1 activity (Figure 2.4). A crystal structure of Caf1-poly(A) could demonstrate whether the stacked, helical structure of poly(A) is also recognised by Caf1. Nonetheless, extensive screening of Caf1 from different species or mutagenesis of crystal interaction surfaces would likely be required, as the active site is occluded by crystal contacts in existing Caf1 structures (Andersen et al., 2009; Jonstrup et al., 2007; Thore et al., 2003). Finally, structural and biochemical characterisation of different poly(A)-binding mechanisms of deadenylase enzymes, such as those of the EEP exonuclease family, could elucidate the significance and conservation of oligo(A) structure in deadenylation.

4.2 Updated Model of Pan2-Pan3 Function

In Chapter 3, I reconstituted deadenylation *in vitro* with recombinantly expressed and purified proteins to address how Pan2-Pan3 deadenylation is regulated by the UTR sequence, the cytoplasmic poly(A)-binding protein Pab1, and the sequence-specific RBP MEX3.

Using RNAs containing different upstream non-A sequences but the same poly(A) tail length, I found that Pan2-Pan3 activity is affected by the non-A sequence (Figure 3.3, Figure 3.4). This suggests that Pan2-Pan3 may preferentially deadenylate mRNAs with different 3'-UTRs. Nonetheless, the basis of how upstream sequences affect Pan2-Pan3 activity is unknown. A previous study had identified transcripts whose overall abundance and decay rates were affected when Pan2 was mutated *in vivo* (Sun et al., 2013); nonetheless, gene ontology and sequence analysis failed to reveal any clear patterns of Pan2-Pan3 transcript preference. Sequencing, biochemical, and structural studies of the mechanistic basis of the UTR sequence preference of Pan2-Pan3 may provide an explanation for its *in vivo* substrate dependence.

I tested how Pan2-Pan3 activity is regulated by recombinantly expressed and purified yeast Pab1 and human PABPC1 (Figure 4.1). Yeast Pab1 and human PABPC1 bound poly(A) RNA with different affinities and cooperativities (Figure 3.6, Figure 3.7). Biophysical and structural studies could provide a molecular explanation for potentially different mechanisms of RNA binding and multimerisation by Pab1. Intriguingly, the higher cooperativity of RNA binding by PABPC1 correlates with longer *in vivo* tail lengths in human cells. This may suggest a mechanism of ensuring a higher stoichiometry of PABPC1 on longer poly(A) tails, providing greater stability and the potential of differential regulation by PABPC1 modification.

Pab1 had antagonising effects on Pan2-Pan3 deadenylation. Using *in vitro* deadenylation assays, I found that stably-bound Pab1 inhibited Pan2-Pan3 in yeast (Figure 3.8) and humans (Figure 3.9). For both complexes, inhibition was likely due to encountering the poly(A)-bound RRM

domains (Figure 3.29). However, Pab1 also stimulated Pan2-Pan3 activity. Full stimulation required multiple interaction interfaces between Pab1-poly(A) and Pan2-Pan3 (Figure 3.13, Figure 3.14, Figure 3.30). However, while the RRM domains of yeast Pab1 did not stimulate Pan2-Pan3 (Figure 3.13), RRM1-4 of PABPC1 partially stimulated human PAN2-PAN3 (Figure 3.14). This suggests that there may be differences in the interaction mechanism between yeast and human complexes. Structural investigations of human PAN2-PAN3 in complex with PABPC1 and poly(A) could reveal the molecular basis of interaction, and thus elucidate the mechanism of PAN2-PAN3 stimulation. Furthermore, the yeast and human ternary complexes could give insight into species-specific differences in steady-state poly(A) tail lengths. Finally, the published structure of yeast Pan2-Pan3-Pab1-poly(A) was limited by resolution, hindering modelling and thus the interpretation of molecular interactions within the Pab1-poly(A) RNP. A higher resolution structure, made possible by improved sample preparation, data collection, and processing, may provide biological insight into the molecular mechanism of Pab1-poly(A) engagement by Pan2-Pan3. Having identified the minimal human PAN2-PAN3 and PABPC1-poly(A) complex required for a stable interaction (Figure 3.17), I aim to structurally characterise this ternary complex. These investigations will be aided by ongoing efforts to crystallise portions of PABPC1-poly(A) and PAN2-PAN3 in our laboratory.

Finally, a proximity labelling study identified MEX3 as a potential interactor of PAN2-PAN3 *in vivo* (Youn et al., 2018). *In vitro* reconstitution of deadenylation using a substrate RNA containing a MEX3 binding site showed that MEX3 stimulates deadenylation by human PAN2-PAN3 (Figure 3.25, Figure 3.26), likely by directly recruiting the deadenylase complex (Figure 3.28). To my knowledge, this is the first time a sequence-specific RNA-binding protein has been shown to stimulate PAN2-PAN3 activity, suggesting that PAN2-PAN3 can be involved in transcript-targeted deadenylation. However, the regions of MEX3 and PAN2-PAN3 which mediate this interaction are unknown. The domains required for interaction and

its mechanistic basis can be investigated through biochemical and structural methods using different truncations of MEX3 and PAN2-PAN3.

In the study by Youn and colleagues, other sequence-specific RBPs were identified as proximal to PAN2-PAN3 (Youn et al., 2018). These included YTHDF2 and YTHDF3, components of the miRNA machinery (TNRC6 proteins and Ago2), and ZFP36 (Figure 1.10). These proteins could be expressed, purified, and tested in *in vitro* assays for their effect on PAN2-PAN3 deadenylation. Surveying the mechanisms of interaction or stimulation may elucidate universal rules regarding regulation of PAN2-PAN3 activity by RBPs.

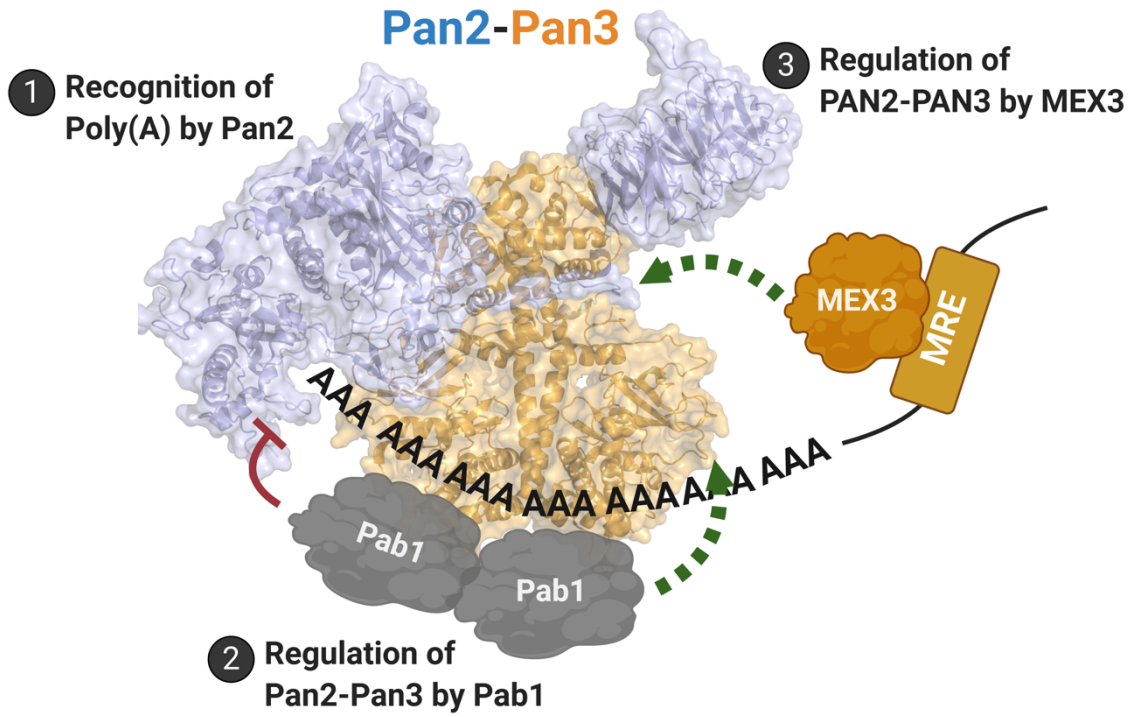


Figure 4.1 Updated model of Pan2-Pan3 function. Pan2 and Pan3 are shown as blue and orange cartoons respectively. (**1**) In Chapter 2, I demonstrated that the Pan2 exonuclease recognised poly(A) through its intrinsic base-stacked, helical structure. (**2**) In Chapter 3, I investigated the mechanism of Pan2-Pan3 regulation by Pab1 (grey schematic). The RRM domains of stably-bound Pab1 could inhibit Pan2 activity, likely through steric effects; on the other hand, interaction of Pan2-Pan3 with Pab1 through multiple interfaces stimulated Pan2-Pan3 deadenylation. (**3**) In Chapter 3, I found that the RBP MEX3 (orange schematic) stimulated PAN2-PAN3 activity through a direct interaction, representing the first demonstration of an RBP facilitating targeted deadenylation by Pan2-Pan3.

4.3 A Global View of Deadenylation

mRNA deadenylation plays a key role in post-transcriptional regulation of gene expression. Deadenylation by Pan2-Pan3 and Ccr4-Not determines steady-state poly(A) tail lengths and represses gene expression by initiating mRNA decay and inhibiting translation. However, important questions remain regarding their *in vivo* roles, correlations of *in vitro* deadenylation and *in vivo* measurements of poly(A) tail length and its shortening, and their interplay with other mRNP-associated macromolecules.

An open question in the field of RNA decay is why there are two conserved cytoplasmic deadenylase complexes: Pan2-Pan3 and Ccr4-Not. Despite the importance of deadenylation, deleting individual deadenylases does not impair cellular fitness (Giaever et al., 2002). This may be due to robust feedback loops to maintain mRNA homeostasis by increasing the expression of other RNA decay components or by globally reducing transcription. Thus, future *in vivo* studies would benefit from transient depletion, such as the auxin-inducible degron system (Barrass et al., 2019), or expression of dominant negative catalytic mutants. These may attenuate feedback mechanisms and can be combined with poly(A) sequencing and metabolic labelling. Such studies may elucidate the role of deadenylases in different stages of poly(A) tail shortening and their effects on the dynamics of deadenylation *in vivo*.

Several hypotheses for the conservation of both Pan2-Pan3 and Ccr4-Not have been proposed. Firstly, our laboratory has observed that Pan2-Pan3 (Figure 3.3, Figure 3.4) and Ccr4-Not (Webster et al., 2019) have preference for the sequence upstream of the poly(A) tail. It is thus possible that the two complexes are important for deadenylation of different transcripts. Nonetheless, as both complexes can deadenylate all tested RNA substrates *in vitro*, intrinsic sequence specificity is unlikely to fully explain observed differences in transcript dependence (Sun et al., 2013). Secondly, different RBPs may recruit either Pan2-Pan3 or Ccr4-Not to target

transcripts for deadenylation. Ccr4-Not interacts with sequence-specific RBPs to destabilise their cognate transcripts (Du et al., 2016; Stowell et al., 2018; Webster et al., 2019), whereas my experiments showed that PAN2-PAN3 binds and is stimulated by MEX3 (Figure 3.25, Figure 3.28). However, it remains unclear if RBPs recruit deadenylases to different transcript subsets, and how RNAs which are not bound by deadenylase-recruiting RBPs are targeted. As the interactome of Pan2-Pan3 is poorly clarified, future experiments may identify binding partners of Pan2-Pan3 by pull-down or proximity-based labelling studies (Trinkle-Mulcahy, 2019). Finally, Pan2-Pan3 and Ccr4-Not could be important under different cellular contexts. For example, RNPs localise to phase condensates during cellular stress (Decker and Parker, 2012). Common regulatory RBPs such as PABPC1 and Ded1 have also been shown to phase separate (Iserman et al., 2020; Riback et al., 2017). As Pan2-Pan3 can localise to P-bodies, deadenylation by Pan2-Pan3 may predominate during cellular stress (Zheng et al., 2008). This possibility could be investigated by poly(A) tail length studies during different cellular conditions, combined with the perturbation of individual deadenylase enzymes.

In vivo, different species exhibit a wide range of steady-state poly(A) tail lengths. In some studies, the poly(A) tail is phased by 30-A increments, equivalent to the footprint of one Pab1 molecule (Lima et al., 2017). This periodicity implies that Pab1 molecules are phased according to their distance from the 3'-UTR-poly(A) junction. In other studies, steady-state poly(A) tails are broadly distributed, but median lengths are frequently multiples of 30, suggesting that Pab1 stoichiometry may determine the steady-state poly(A) tail length (Chang et al., 2014; Eisen et al., 2019; Subtelny et al., 2014). My *in vitro* experiments with Pan2-Pan3 and Pab1 could recapitulate steady-state poly(A) tail lengths, as Pan2-Pan3 is generally stimulated by Pab1 until it encounters stably-bound Pab1 RRMs, when it becomes inhibited. The steady-state tail length of 30 A's in yeast can be reproduced with Pan2-Pan3 deadenylation of a Pab1-bound poly(A) tail (Figure 3.10). Furthermore, phased poly(A) tails and longer tail lengths are observed with

PAN2-PAN3 deadenylation of PABPC1-poly(A), depending on PABPC1 stoichiometry (Figure 3.11). The hypothesis that Pan2-Pan3 may regulate the steady-state poly(A) tail is further supported by the observation that C-terminal truncations of yeast Pab1, which greatly diminish its interaction with Pan2-Pan3, increased steady-state poly(A) tail lengths comparable to when Pan2 or Pan3 is individually deleted (Brown and Sachs, 1998a; Mangus et al., 2004a). Thus, the interplay between Pab1 stoichiometry and its regulation of Pan2-Pan3 could affect steady-state poly(A) tail lengths.

Work presented in this dissertation and elsewhere has thus far reconstituted deadenylation using minimal components. However, mRNAs, RBPs, and poly(A) processing enzymes do not exist in isolation *in vivo*. Thus, more complex reaction mixtures should be reconstituted to dissect mechanisms of deadenylation. A previous study had postulated that deadenylation is biphasic: Pan2-Pan3 slowly removes the distal poly(A) tail, whereas Ccr4-Not rapidly degrades proximal poly(A) (Yamashita et al., 2005). However, how deadenylase complexes are targeted to different tail lengths, and the tail length at which deadenylation transitions from Pan2-Pan3 to Ccr4-Not, are unclear. Reconstitution of biphasic deadenylation *in vitro* could complement poly(A) sequencing studies in elucidating the role of Pan2-Pan3 and Ccr4-Not in different phases of poly(A) tail shortening. Moreover, more intricate mRNP complexes can be tested in deadenylation assays. One example is the RNA-induced silencing complex (Ameres and Zamore, 2013). Using our minimal system, the requirement of RISC components for stimulation of deadenylation can be tested. Finally, deadenylation enzymes can be incorporated into complex pathways of mRNA decay. For example, while mRNAs are thought to circularise *in vivo* via the eIF4G-Pab1 interaction (Vicens et al., 2018), it is unknown how the 5'-cap and the eIF4F complex affects deadenylation by Pan2-Pan3 or Ccr4-Not. Reconstitution of mRNA circularisation and downstream decay pathways allows deadenylation to be integrated into the bigger picture of RNA decay in a biochemically manipulable *in vitro* system.

4.4 Conclusion

The field of deadenylation and RNA decay has grown immensely since the discovery of the poly(A) tail. Owing to improvements in genetics, biochemistry, biophysics, structural biology, and sequencing techniques, regulation of gene expression at the post-transcriptional level has been gradually uncovered, revealing highly complex pathways involving numerous factors. By investigating biochemical and structural mechanisms of deadenylation by Pan2-Pan3, this dissertation has uncovered new paradigms of single-stranded RNA binding, Pan2-Pan3 function, and its regulation by RNA-binding proteins. My findings, along with other recent work, can serve as a platform for future experiments and collaborative efforts to elucidate the pathways of post-transcriptional regulation of gene expression.

5. Materials and Methods

5.1 Common Reagents

5.1.1 Buffers

All buffers are made up with MQ H_2O. All other buffers are listed in their respective sections.

DEPC H_2O: Add 1 mL 0.1% DEPC to 1 L MQ H_2O. Autoclave and cool to RT.
MES running buffer (MES-RB): 25 mM MES, 25 mM Tris, 1.75 mM SDS, 0.5 mM EDTA, pH 7.3
MOPS running buffer (MOPS-RB): As above (MES-RB), with 25 mM MOPS, pH 7.7
Phosphate-buffered saline (PBS): 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4
Tris acetate EDTA (TAE): 40 mM Tris-acetate, 1 mM EDTA, pH 8.5
Tris borate EDTA (TBE): 90 mM Tris-borate, 1 mM EDTA, pH 8.3
Tris EDTA (TE): 10 mM Tris pH 8.0, 1 mM EDTA

5.1.2 Liquid Media

All media are made to 1 L with MQ H_2O, and autoclaved and cooled to RT before use.

Miller lysogeny broth (LB): 5 g yeast extract, 10 g tryptone, 10 g NaCl. Adjust to pH 7.5.
Terrific broth (TB): 24 g yeast extract, 12 g tryptone, 4 mL glycerol. Make up to 900 mL with MQ H_2O. Autoclave and cool to RT; add 100 mL KH_2PO_4 to 0.1 M and pH 7.5.
2× Tryptone Yeast (2×TY): 10 g yeast extract, 16 g tryptone, 5 g NaCl. Adjust to pH 7.5.
YM4: 8 g yeast nitrogen base, 11 g casamino acids, 55 mg adenine, 55 mg tyrosine.
YM4 + tryptophan (YM4+Trp): As above (see YM4) + 55 mg tryptophan.
YM4 + uracil (YM4+Ura): As above (see YM4) + 20 mg uracil.
Yeast peptone D-glucose (YPD): 20 g peptone, 20 g D-glucose, 10 g yeast extract.

5.1.3 Plate Media

Recipes are made up to 1 L with MQ H_2O, autoclaved, poured into 90 mm Petri dishes, and cooled to RT or 4 °C. Antibiotics are added to their final concentrations while cooling.

Bacmid: 5 g yeast extract, 10 g tryptone, 8 g NaCl, 15 g agar. Add 50 μg/mL Kan, 10 μg/mL tetracycline, 7 μg/mL Gen, 40 μg/mL IPTG, 100 μg/mL Bluo-Gal (ThermoFisher).
Tryptone yeast extract (TYE): 5 g yeast extract, 10 g tryptone, 8 g NaCl, 15 g agar, pH 7.0.
Yeast extract peptone D-glucose (YEPD): YPD + 25 g agar, pH 7.0.
-Ura, -Trp, -Ura -Trp: 8 g yeast nitrogen base, 55 mg Tyr. Make up to 800 mL with MQ H_2O. Cool to RT and add 100 mL 20% w/v glucose and 100 mL 10× amino acid mix (minus drop-outs).

5.1.4 Antibiotics and Additives

Antibiotic and IPTG stocks were made up in MQ H_2O, filter-sterilised, and stored at -20 °C. 20% raffinose and 20% galactose were used immediately.

Table 5.1 List of antibiotics and additives used in liquid and plate media

Antibiotic/Additive	Stock Concentration	Final Concentration
Ampicillin	50 mg/mL	50 μg/mL
Gentamycin	50 mg/mL	50 μg/mL
IPTG	1 M	1 mM
Kanamycin	50 mg /mL	25 μg/mL
20% Galactose	20% (w/v)	2% (w/v)
20% Raffinose	20% (w/v)	Varies

5.2 *E. coli*

5.2.1 Strains

Table 5.2 List of *E. coli* strains used in this dissertation.

Strain	Genotype	Source
BL21 STAR (DE3)	F- *ompT hsdSB*(rB- mB-) *gal dcm lon rne131* (DE3)	ThermoFisher
TOP10	F- *mcrA* Δ(*mrr-hsd*RMS-*mcrBC*) Φ80d*lacZ*ΔM15 Δ*lacX*74 *recA1 araD139* Δ(*ara-leu*)7697 *galK rpsL150*(StrR) *endA1 nupG*	ThermoFisher
DH10 EMBacY	F- *mcrA* Δ(*mrr-hsd*RMS-*mcrBC*) Φ80d*lacZ*ΔM15 Δ*lacX*74 Δ(*ara-leu*)7697 *recA1 relA1 araD139 galE15 galK16* Tn10.10 *rpsL150*(StrR) *endA1 nupG spoT1* λ– (bEMBacY, pTn7helper)	Imre Berger

5.2.2 Competent *E. coli*

Cells were made competent in collaboration with James Stowell, Eva Absmeier, and Tamara Sijacki. DH10 EMBacY cells were made competent by Jianguo Shi.

E. coli was streaked from a frozen stock onto a TYE plate and incubated overnight at 37 °C. One colony was used to inoculate 5 mL warm LB and shaken overnight at 200 rpm, 37 °C. 2.5 mL was transferred to 250 mL warm LB in sterile Erlenmeyer flasks and shaken at 200 rpm, 37 °C to an OD_{600} of 0.45. Cells were decanted into 50 mL tubes, incubated on ice for 15 min, and centrifuged at 1150 × g, 15 min, 4 °C. Pellets were gently resuspended in 80 mL chilled TFB1 buffer (10 mM MES, 10 mM $RbCl_2$, 50 mM $MnCl_2$, 10 mM $CaCl_2$, pH 5.8 with acetic acid, filter-sterilised), incubated on ice for 15 min, and centrifuged as above. Pellets were gently resuspended in 10 mL chilled TFB2 buffer (10 mM MOPS, 100 mM $CaCl_2$, 10 mM $RbCl_2$, 15% (v/v) glycerol, pH 6.6 with NaOH, filter-sterilised) and incubated on ice for 15 min. 25 μL cells were aliquoted into chilled 1.5 mL tubes and frozen at -80 °C.

5.2.3 Heat-shock Transformation

Plasmid DNA (100-500 ng) or Gibson reactions (5 μL, Section 5.5.5) were incubated with chemically competent *E. coli* (Section 5.2.2) for 20 min on ice, heat shocked at 42 °C for 45 s, chilled on ice before adding 200 μL 2×TY, and incubated at 37 °C, 200 rpm, 1 h for recovery. DH10 EMBacY cells were incubated for 2.5 h for recovery. Cultures were spread on correct plates and incubated overnight at 37 °C. Bacmid plates were incubated over two nights.

5.2.4 Growth for Plasmid/Bacmid Preparation

One TOP10 colony was inoculated in 5 mL 2×TY or TB in a sterile 50 mL tube with the correct antibiotic. For bacmid preparations, one white colony of DH10 EMBacY (signifying

lacZ disruption and thus insert integration) from bacmid plates was used to inoculate 5 mL 2×TY in a sterile 50 mL tube with 1× Gen and Kan. The culture was shaken at 37 °C, 200 rpm overnight, and centrifuged at 3000 × g. The supernatant was discarded; pellets were frozen at -20 °C or immediately used for DNA preparation (Section 5.5.10, Section 5.5.11).

5.2.5 Growth for Protein Overexpression

One BL21 (DE3) STAR colony was picked and inoculated in ~100-150 mL TB in a sterile Erlenmeyer flask with the correct antibiotic. Alternatively, the frozen glycerol stock (from a colony overexpressing the POI, Section 5.10.1) was scraped and inoculated. This was shaken at 37 °C, 200 rpm, overnight. 10 mL was transferred to 1 L warm TB in 2 L Erlenmeyer flasks. Expression cultures were shaken at 37 °C, 200 rpm until $OD_{600} \approx 0.5\text{-}1.0$. Overexpression was induced by adding IPTG to 1 mM. Cultures were shaken at 18 °C overnight and centrifuged at 3876 × g. The supernatant was discarded; pellets were resuspended in chilled PBS, centrifuged at 3000 × g, flash frozen in liquid N_2, and stored at -80 °C.

5.3 *S. cerevisiae*

5.3.1 Strains

Table 5.3 List of yeast strains used in this dissertation.

Strain	Genotype	Source
BCY123	*pep4::HIS3 prb::LEU2 bar1:HISG lys2::GAL1/10-GAL4 can1 ade2 ura3 leu2–3,112*	Kiyoshi Nagai

5.3.2 Competent *S. cerevisiae*

S. cereivisiae was made competent using the Frozen-EZ Yeast Transformation II Kit (Zymo Research). One colony of pre-streaked BCY123 was used to inoculate 5 mL YPD and grown at 30 °C, 200 rpm overnight. 0.5 mL pre-culture was used to inoculate 10 mL YPD. Cells were grown at 30 °C, 200 rpm to $OD_{600} \approx 0.8$-1.0 and centrifuged at 500 × g, 3 min at 21 °C. The cell pellet was gently resuspended in 10 mL EZ solution 1 and centrifuged as above. Cells were gently resuspended in 500 μL EZ solution 2 and 25 μL aliquots were slowly frozen at -20 °C.

5.3.3 Transformation

S. cereivisiae was transformed using the Frozen-EZ Yeast Transformation II Kit (Zymo Research). Competent yeast (Section 5.3.2) was mixed with 0.25-1.0 μg plasmid DNA and 250 μL EZ solution 3, and incubated at 30 °C, 200 rpm, 45 min. Cultures were plated on the appropriate single/double auxotrophic selection plate and incubated at 30 °C for 2 days.

5.3.4 Growth for Protein Overexpression

S. cerevisiae overexpression protocols were developed by Jana Wolf. Cells from a transformed plate (Section 5.3.3) were used to inoculate 500 mL YM4 (+Trp/Ura if necessary) pre-cultures with 2% (w/v) raffinose and shaken at 30 °C, 200 rpm overnight. Pre-cultures were added to 1 L warm YM4 (+Trp/Ura if necessary) to $OD_{600} \approx 0.2$ in 2 L Erlenmeyer flasks with 1% (w/v) raffinose. Cells were shaken at 30 °C, 200 rpm until $OD_{600} \approx 1.0$-1.2. Overexpression was induced by 2% (w/v) galactose and the culture was shaken at 30 °C, 200 rpm overnight. Cells were centrifuged at 3876 × g. Pellets were resuspended in chilled PBS, centrifuged at 3000 × g, flash frozen in liquid N_2, and stored at -80 °C.

5.4 Sf9 cells

All steps were performed in a sterile tissue culture hood. Cell density and viability were checked using a Countess™ cell counter (ThermoFisher) with 0.2% (w/v) Trypan Blue.

5.4.1 Strain

Table 5.4 Description of insect cell strain used in this dissertation.

Strain	Description	Source
Sf9	Clonal isolate of *Spodoptera frugiperda* Sf21 ovarian cells	Jianguo Shi

5.4.2 Maintenance of Sf9 Cells in Suspension

Maintenance was carried out by Jianguo Shi. Low passage Sf9 stocks (+DMSO, -80 °C) were recovered in small volumes. Suspension cultures were maintained in Insect-XPRESS™ media (Lonza) at 2×10^6 cells/mL at 27 °C, 140 rpm for 25-35 passages.

5.4.3 Bacmid Transfection of Adherent Sf9 Cells

Sf9 cells (Section 5.4.2) were diluted to 0.5×10^6 cells/mL with Insect-XPRESS™ media (Lonza). 2 mL was added to each well of a 6-well tissue culture plate (Corning) and left to adhere for ~30 min at 27 °C. 50 μg purified bacmid (Section 5.5.11) was mixed with 1 mL Insect-XPRESS™ media (Lonza). 25 μL FuGENE® HD transfection reagent (Promega) was slowly pipetted into the DNA solution, vortexed, incubated at 20 °C for ~30 min and divided between five wells of adherent Sf9. The sixth well was kept as a negative control. The plate was placed in a sealed box with wet tissue and incubated at 27 °C for 3-5 days. Cells were checked for fluorescence after 3 days, until fluorescence from YFP was visible in ~10% of Sf9 cells. Cells were discarded if the negative control was fluorescent, cells were not fluorescent

six days post-transfection, or the culture was contaminated. The supernatant was removed and filter-sterilised; this (P1) was mixed 1:1 with FBS (Gibco, ThermoFisher) and stored at 4 °C in the dark for a maximum of one year.

5.4.4 Virus Amplification

1-5 mL of P1 virus (depending on fluorescence upon harvest, Section 5.4.3) was transferred to 50 mL 2×10^6 cells/mL Sf9 cells (Section 5.4.2). Cell density and fluorescence were monitored daily. Cell division should arrest approximately 48 hours post-transfection; if not, cells were maintained at 2×10^6 cells/mL by dilution in Insect-XPRESS™ media (Lonza). The supernatant was harvested one day after 90-100% high-viability (>90%) cells were fluorescent. The culture was centrifuged at 2000 × g; the supernatant was removed and filter-sterilised. This (P2) was stored at 4 °C in the dark for a maximum of three months.

5.4.5 Protein Overexpression

4-6 L expression cultures were diluted to 1.5 or 2×10^6 cells/mL and split into 500 mL in 2 L roller bottles. 50 mL P2 (Section 5.4.4) was divided evenly among the bottles. Cell division should arrest and 90-100% of cells should be fluorescent by 24 h. Cells were harvested at a predetermined optimal time (Section 5.10.2) by centrifugation at 3876 × g. The supernatant was discarded; pellets were resuspended in pre-chilled PBS and centrifuged at 3000 × g. Pellets were frozen in liquid N_2 and stored at -80 °C.

5.5 Molecular Biology Techniques and Cloning

5.5.1 Gene Synthesis

S. cerevisiae Pan2 and Pan3-SII were cloned into yeast expression vectors (pRS-) by Jana Wolf. Other genes were synthesised and cloned by Epoch Life Science into vectors as below (Table 5.5). SII and His_8 tags were LEVLFQGPWSHPQFEKGSAGSAAGSGAGWSHPQFEK and MHHHHHHHHHLEVLFQGP respectively. Gene sequences were codon-optimised for *E. coli*.

Table 5.5 Gene synthesised constructs used in this dissertation.

Protein Name	RefSeq	Uniprot	Vector	Notes
***sc*Pan2**	NP_011421.1	P53010	pACEBac1	5' leader: BamHI/XhoI, polh promoter (GGATCC CTCGAG TCAAC). 3' leader: XbaI/KpnI, stop. (GGATCC CTCGAG TCAAC).
***sc*Pan3-SII**	NP_011421.1	P36102		
His_8-*sc*Pan2 UCH-Exo	NP_011421.1	P53010		
His_8-*ct*Pan2 UCH-Exo	XM_006694602.1	G0SAK8		
***hs*PAN2 (Iso 1)**	NP_001120932.1	Q504Q3		
***hs*PAN3-SII (Iso 1)**	NP_787050.6	Q58A45		
***sc*Pab1**	NP_011092.1	P04147	pGEX-6P1	Restriction cloning: BamHI and XbaI sites
***hs*PABPC1**	NP_002559.2	P11940		
***hs*MEXA-SII**	NP_001087194.1	A1L020	pACEBac1	As above (pACEBac1).
***hs*MEXB-SII**	NP_115622.2	Q6ZN04		
***hs*MEXC-SII**	NP_057710.3	Q5U5Q3		
***hs*MEXD-SII (Iso 1)**	NP_976049.3	Q86XN8		

5.5.2 Nucleic Acid Quantification

Nucleic acid concentration was determined using Nanodrop ND-1000 and Nanodrop One spectrophotometers (ThermoFisher) using 1.5 μL sample. DNA concentrations in ng/μL were obtained using $\varepsilon = 50$ ng/μL cm. RNA extinction coefficients were obtained from the Oligo Calculator tool by atdbio. Sample purity was estimated by the A_{260}/A_{280} ratio.

5.5.3 Polymerase Chain Reaction (PCR)

PCR reactions were carried out in 0.2 mL PCR tubes (Axygen). Components were added to the reaction in the below order (Table 5.6). Primers were designed with 18-27 nt complementarity to the target DNA. The reaction was cycled in a Veriti 96-well thermal cycler (Applied Biosystems) as below (Table 5.7). The reaction was used immediately in QIAquick PCR purification kits according to manufacturers' instructions (Qiagen), directly analysed by horizontal agarose gel electrophoresis (Section 5.5.8), or frozen until later use.

Table 5.6 Phusion PCR Reaction Setup

Component	Volume (μL)
MQ H_2O	32.0
5× Phusion HF Buffer (NEB)	10.0
10 μM forward primer (Sigma Aldrich)	2.5
10 μM reverse primer (Sigma Aldrich)	2.5
0.5-10 ng/μL DNA template, diluted in MQ H_2O	1.0
10 mM dNTP mix (GE)	1.0
Phusion DNA polymerase (NEB)	1.0

Table 5.7 Phusion PCR Thermocycling Parameters

	Temperature (°C)	Time
1×	98	1 min
35×	98	20 s
	50-65 (5-10 °C lower than predicted T_m)	30 s
	72	30 s per kb
1×	72	2.5 min per kb
	4	Hold

5.5.4 Restriction Enzyme Cloning

Plasmids (1-5 μg) were digested with restriction enzymes (NEB) at temperatures and in buffers according to manufacturers' instructions for ~2 h. Completely digested plasmids were checked by horizontal agarose gel electrophoresis as a single band migrating higher than an undigested negative control (Section 5.5.8). The digested plasmid was gel-purified (Section 5.5.9).

Primers for insert amplification were designed such that a few random nucleotides followed by the desired restriction site were introduced. PCR was carried out and purified as above (Section 5.5.3). The entire volume of PCR product was digested and purified as above.

5.5.4.1 DNA Fragment Ligation

Ligation of DNA with complementary sticky ends was performed with T4 DNA ligase (NEB) in 10 μL according to manufacturers' instructions. Vectors were diluted to 10 ng/μL; inserts were diluted to 3× the vector. 1 μL of each was added to the ligation reaction and incubated at 20 °C for 20 min. The reaction was transformed into TOP10 cells (Section 5.2.3).

5.5.5 Gibson Assembly

Vectors and inserts were amplified by PCR as above (Section 5.5.3). Alternatively, shorter inserts (e.g. a TEV-SII tag) were ordered as a gene block (Sigma Aldrich). For large deletions, vectors containing the desired insert were amplified by PCR, omitting the deleted sequence. Primers and gene blocks were designed to contain at least 15 overlapping nucleotides at the 5' and 3' ends. The PCR products were gel-purified (Section 5.5.9) before Gibson Assembly.

5× Isothermal Assembly Buffer (IAB): 500 mM Tris-HCl pH 7.5, 20 mM MgCl2, 1 mM dNTP mix, 50 mM DTT, 1.5 g PEG-8000, 5 mM NAD. Dissolve at 4 °C. Make up to 6 mL in MQ H2O and store at -20 °C in 320 µL aliquots.

Gibson Buffer (GB): 1.2 µL T5 exonuclease (NEB), 20 µL Phusion polymerase (NEB), 160 µL Taq ligase (NEB), 700 µL MQ H2O, 320 µL 5× IAB. Store 15 µL aliquots at -20 °C.

Vector was diluted to 10 ng/µL; inserts were diluted to 50 ng/µL. 1 µL of each was added to one GB aliquot and topped up to 20 µL using MQ H_2O. The reaction was incubated at 50 °C for 30 min. 5 or 10 µL of this reaction was transformed into TOP10 *E. coli* (Section 5.2.3).

5.5.6 Multigene Expression Plasmids (pBIG1)

Multigene expression plasmids (pBIG1) were made according to a modified biGBac protocol (Weissmann et al., 2016). The GOI was cloned into pACEBac1, which contains a polyhedrin promoter (constitutive overexpression in Sf9 cells) and an SV40 terminator. Gene expression cassettes were amplified by PCR (Section 5.5.3) using specific primers, designed to minimise incorrect assembly (Table 5.8). The first cassette had CasI_F as the forward primer and the last cassette had CasV_R as the reverse primer.

Table 5.8 Primers used for Gibson Assembly into pBIG1. α and ϖ are complementary to the 3' and 5' ends of SwaI-digested pBIG1 vector respectively. Sequences in bold represent the overlapping sequences for Gibson Assembly.

Primer	Sequence
CasI_F (α)	**AACGCTCTATGGTCTAAAGATTT**-AAATCGACCTACTCCGGAATATTAATAGATC
CasI_R (β)	**AAACGTGCAATAGTATCCAGTTT**-ATTTAAATGGTTATGATAGTTATTGCTCAGCG
CasII_F (β)	**AAACTGGATACTATTGCACGTTT**-AAATCGACCTACTCCGGAATATTAATAGATC
CasII_R (γ)	**AAACATCAGGCATCATTAGGTTT**-ATTTAAATGGTTATGATAGTTATTGCTCAGCG
CasIII_F (γ)	**AAACATCAGGCATCATTAGGTTT**-ATTTAAATGGTTATGATAGTTATTGCTCAGCG
CasV_R (ϖ)	**AACCCCGATTGAGATATAGATTT**-ATTTAATGGTTATGATAGTTATTGCTCAGCG

pBIG1 was modified by Matthias Girbig (Andrew Carter lab, MRC LMB) to include necessary Gibson overhangs, spacer sequences, and SwaI restriction sites and to confer Gen rather than Amp resistance. 1 μg pBIG1A was linearised by overnight incubation with SwaI according to manufacturers' instructions (NEB) and gel-purified (Section 5.5.9). Gibson Assembly was carried out as above (Section 5.5.5) with pBIG1A at 10 ng and inserts each at 50 ng.

A schematic of the biGBac protocol is shown in Figure 5.1.

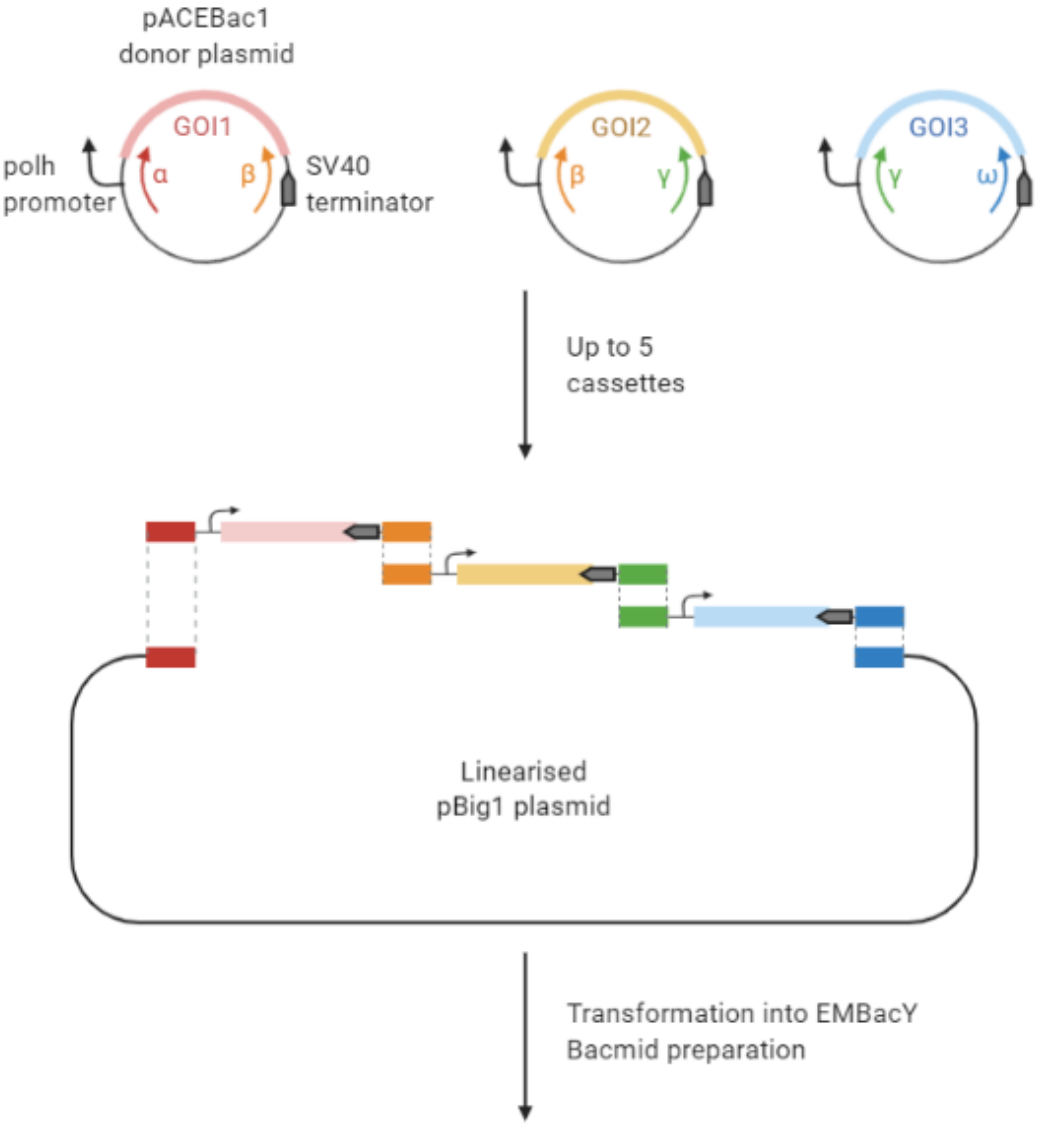


Figure 5.1 Schematic of biGBac protocol. Gene cassettes and the flanking polh promoter and SV40 terminator are amplified. Purified cassettes are used in Gibson assembly with linearised pBIG1, which is transformed into EMBacY for bacmid preparation.

5.5.6.1 Confirmation of Correct Gene Insertion

Correct gene insertion was confirmed by SwaI digest of the assembled pBIG plasmid, which contains unique SwaI sites in linker sequences between each insert. The plasmid was digested by SwaI (NEB) for 2 h according to manufacturers' instructions. The digestion products were analysed by horizontal agarose gel electrophoresis (Section 5.5.8). Insertion of the correct genes was confirmed by single bands (vector and inserts) at the predicted lengths.

5.5.7 QuikChange Site-directed Mutagenesis

Point mutations were made according to a modified protocol from QuikChange II. The PCR was as above (Section 5.5.3) except with 25 ng DNA template and 25 thermocycles. Template DNA was digested with 2 μL DpnI for 2 h at 37 °C; DpnI was heat-inactivated by incubation at 65 °C for 10 min. 5 or 10 μL was used to transform TOP10 *E. coli* (Section 5.2.3).

5.5.8 DNA Analysis by Horizontal Agarose Gel Electrophoresis

DNA analysis was performed using a Mini-Sub Cell GT system (BioRad). 1.5/3.0 g Hi-Pure low EEO agarose (BioGene, for 1/2% agarose gels respectively) was added to 150 mL TAE and dissolved by microwaving for 90 s. 15 μL SYBR-Safe (Life Technologies) was added. Gels were cooled in 15×10 cm UV-transparent trays with the appropriate comb inserted.

DNA samples were mixed with 6× purple loading dye (NEB) and added to wells. Gels were run at 100 V in TAE for 45-75 min depending on the desired resolution and were visualised using UV transillumination with a GelDoc XR+ system (BioRad).

5.5.9 Gel Extraction of DNA

Gels were prepared and run as above (Section 5.5.8). Bands were visualised using a DR46B blue LED transilluminator (Clare Chemical Research) with an orange filter. Bands were excised using a scalpel and transferred to a 2 mL tube. DNA from the band was recovered using the Zymoclean Del DNA Recovery Kit (Zymo Research) according to manufacturers' instructions, except that EB was pre-heated to 50 °C to maximise elution. DNA concentration and purity were quantified as above (Section 5.5.2). The elution step was repeated for higher yields.

5.5.10 Plasmid DNA Extraction from *E. coli*

DNA was extracted from *E. coli* pellets using the QIAprep Spin Miniprep Kit (Qiagen) according to manufacturers' instructions, except elution was carried out with 20 μL EB heated to 50 °C and columns were incubated at 37 °C for 10 min to maximise DNA yields. DNA concentration and purity were quantified as above (Section 5.5.2).

5.5.11 Bacmid DNA Extraction from *E. coli*

DNA was extracted using buffers from the QIAprep Spin Miniprep Kit (Qiagen). The pellet was resuspended in 300 μL buffer P1 and lysed in 300 μL buffer P2. Lysis was stopped by adding 400 μL buffer N3, and precipitate was removed by centrifugation at 15000 × g, 20 min in a tabletop centrifuge. 1 mL supernatant was slowly added to 1 mL pre-chilled isopropanol, and incubated on ice for 2 h. Subsequent centrifugation steps were carried out at 15000 × g, 10 min in a tabletop centrifuge. The DNA was pelleted by centrifugation and washed 3× with 500 μL 70% ethanol, centrifuging and discarding the supernatant between each wash step. After the final wash step, residual ethanol was briefly centrifuged and removed using a gel-loading tip. The pellet was air-dried under sterile conditions (taking care not to over-dry) and was gently re-dissolved in 50 μL sterile EB at 37 °C by tapping.

5.5.12 Sanger Sequencing

Sanger sequencing was carried out to ensure correct gene insertion into a vector, correct site-directed mutagenesis, or confirmation of plasmid identity. Sequencing was performed by Source Bioscience according to platform instructions. Primers were designed to be ~15-20 nt, have a T_m of ~55 °C, and contain no significant secondary structure. Sequencing results were analysed using SnapGene.

5.6 Protein Purification

A list of purified proteins and protein complexes are shown in Table 5.9.

Buffers and salts were purchased as powder (>99% grade, Sigma Aldrich). 1 M buffer stocks were pH adjusted with 10 M NaOH or 38% HCl, filter-sterilised, and stored at RT. Purification buffers were pH adjusted and filter-sterilised before immediate use. 1 M DTT (Formedium) was stored at -20 °C. 0.5 M pH 7 TCEP (Soltec Ventures) was stored at -20 °C. cOmplete™ Mini EDTA-free protease inhibitors (Roche) were freshly dissolved at 1 tablet/50 mL buffer. 100 mM PMSF (Sigma Aldrich) was stored at 4 °C. 1 M DNase I and RNase A stocks (Sigma Aldrich) were made up in PBS and stored at -20 °C. BioLock biotin blocking solution (IBA Life Sciences) was stored at -20 °C.

Yeast was lysed at 35 kpsi in a C Series Cell Disruptor (Constant Systems) at 4 °C. Sonication was carried out in a VCX750 Vibra-cell processor (Sonics). Samples were ultracentrifuged in an Optima XPN-100 centrifuge (Beckman Coulter) with Ti45 rotors and 70 mL tubes. Purification was carried out in an AKTA Pure 25 (GE) with an S9 sample pump and an F-9C fraction collector. Sample purity and homogeneity was monitored at each stage by protein SDS PAGE (Section 5.6.1.1). Protein was concentrated with 15 mL Amicon® Ultra-15 centrifugal filter units (Millipore) or 20 mL Vivaspin concentrators (Sartorius). Protein at small volumes was concentrated with 500 µL Vivaspin concentrators (Sartorius). Molecular weight cut-offs were chosen to be smaller than half the MW of the desired complex to minimise protein loss.

Table 5.9 List of purified proteins and protein complexes. Abbreviations of *sc* (*S. cerevisiae*), *ct* (*Chaetomium thermophilum*), and *hs* (*H. sapiens*) are used in the Materials and Methods section.

Protein	Plasmid(s)	Cells
*sc*Pan2-Pan3	pRS426 *sc*Pan2-*sc*Pan3-SII	BCY123

*sc*Pan2E912A-Pan3	pRS426 *sc*Pan2; pRS424 *sc*Pan3E912A-SII	
*sc*Pan2-Pan3ΔZnF	pRS426 *sc*Pan2-*sc*Pan3ΔN41-SII	
*sc*Pan2-Pan3F156A	pRS426 *sc*Pan2; pRS424 *sc*Pan3F156A-SII	
*sc*Pan2-Pan3PKC	pRS426 *sc*Pan2; pRS424 *sc*Pan3ΔN278-SII	
His_8-*sc*Pan2UCH-Exo (UE)	pACEBac1 His_8-*sc*Pan2UCH-Exo (460-1115)	Sf9
His_8-*sc*Pan2UE E912A	pACEBac1 His_8-*sc*Pan2UCH-Exo E912A	
His_8-*sc*Pan2UE Y975A	pACEBac1 His_8-*sc*Pan2UCH-Exo Y975A	
His_8-*ct*Pan2UE E899A	pACEBac1 His_8-*ct*Pan2UCH-Exo E899A	
His_8-*ct*Pan2UE E899A LΔCΔ	pACEBac1 His_8-*ct*Pan2UCH-Exo E899A Δ(175-196) ΔC70	
*hs*PAN2-PAN3	pBIG1A *hs*PAN2 GST-*hs*PAN3-SII	
*hs*PAN2E980A-PAN3	pBIG1A *hs*PAN2E980A GST-*hs*PAN3-SII	
*hs*PAN2-PAN3ΔN278	pBIG1A *hs*PAN2 *hs*PAN3ΔN278-SII	
*hs*PAN2E980A-PAN3ΔN278	pBIG1A *hs*PAN2E980A *hs*PAN3ΔN278-SII	
*hs*PAN2-PAN3PKC	pBIG1A *hs*PAN2 *hs*PAN3ΔN460-SII	
*hs*PAN2E980A-PAN3PKC	pBIG1A *hs*PAN2E980A *hs*PAN3ΔN460-SII	
*sc*Pab1	pGEX-6P-2 *sc*Pab1	BL21 STAR (DE3)
*sc*Pab1 RRM1-4	pGEX-6P-2 *sc*Pab1ΔC169	
His_6-*hs*PABPC1	pET28a His_6-*hs*PABPC1-TEV-SII	
His_6-*hs*PABPC1 RRM1-4	pET28a His_6-*hs*PABPC1ΔC266-TEV-SII	
His_8-*hs*PABPC1 RRMH	pET28a His_6-*hs*PABPC1ΔC242-TEV-SII	

*hs*MEX3A-SII	pACEBac1 *hs*MEX3A-SII	Sf9
*hs*MEX3B-SII	pACEBac1 *hs*MEX3B-SII	
*hs*MEX3C-SII	pACEBac1 *hs*MEX3C-SII	

5.6.1 Analysis of Protein Purification

5.6.1.1 SDS Polyacrylamide Gel Electrophoresis

Protein samples were mixed with 4× LDS buffer (ThermoFisher) and denatured at 95 °C, 5 min. Samples were separated by vertical electrophoresis on pre-cast 4-12% Bis-Tris gradient gels or 4-12% Bolt gels (Invitrogen) in XCell SureLock Mini-Cell Systems or Bolt Mini Gel Tanks (ThermoFisher) respectively. For large proteins/complexes, MOPS-RB was used; for smaller proteins, MES-RB was used. For Bolt gels, Bolt MOPS- or MES-RB (ThermoFisher) were used. Gels were run at 200 V for 45-80 min depending on the resolution required.

5.6.1.2 Protein Quantification

Sample absorbance at 280 nm (A_{280}) was determined with Nanodrop ND-1000 and Nanodrop One spectrophotometers (ThermoFisher) using 1.5 μL sample, blanked against the relevant buffer. Extinction coefficients at 280 nm (ε_{280}) for single proteins were estimated using the tryptophan and tyrosine content of the POI (ProtParam). ε_{280} for protein complexes were estimated as the sum of individual proteins. The path length (l) is 1 cm. The protein concentration was derived from the rearranged Beer-Lambert Law:

$$Concentration\ (M) = \frac{A_{280}}{\varepsilon_{280}\ l}$$

5.6.2 3C Protease

GST-3C protease was purified by James Stowell. *E. coli* pellets were resuspended in 4 °C PBS and cells were lysed by sonication (3s ON, 6s OFF, 70% amplitude, 3 × 1 min). Lysate was clarified by centrifugation at 235 418 × g for 30 min, 4 °C. Clarified lysate was incubated with 5 mL glutathione sepharose 4B resin (GE) for 2 h, 4 °C, washed with 10 CV PBS, and eluted with PBS + 20 mM reduced glutathione. Protein was buffer-exchanged into storage buffer (20 mM HEPES pH 7.5, 300 mM NaCl, 0.5 mM TCEP, 20% glycerol) using a HiPrep Sephadex G-25 desalting column (GE). 1 mg/mL aliquots were frozen in liquid N_2 and stored at -80 °C.

5.6.3 TEV Protease

His-TEV protease was overexpressed as an N-terminal MBP fusion. A basic linker between MBP and His tags allows MBP cleavage during *E. coli* expression. Pellets were resuspended, lysed, and clarified as above (Section 5.6.2). Clarified lysate was incubated with 5 mL Ni^{2+}-NTA resin (GE) for 2 h, 4 °C, washed with 50 mL PBS + 20 mM imidazole, and eluted with 25 mL PBS + 500 mM imidazole. Eluate was loaded onto a 5 mL HiTrap S HP column (GE) and protein was eluted with a 15 CV gradient from 150-1000 mM NaCl. Fractions containing TEV protease were diluted to 2.5 mg/mL in PBS with 20% glycerol. Aliquots were frozen in liquid N_2 and stored at -80 °C.

5.6.4 *S. cerevisiae* Pan2-Pan3 and Variants

Purification protocols for *sc*Pan2-Pan3 were developed by Jana Wolf. Yeast pellets were resuspended 1:2 (w/v) in Buffer A (50 mM PIPES pH 6.5, 300 mM NaCl, 5 mM β-ME, cOmplete protease inhibitors, 500 μM PMSF, 20 μg/mL DNase I) and lysed by passing the suspension 3× through a cell disruptor. Lysate was centrifuged at 235 418 × g for 30 min, 4 °C. Clarified supernatant was incubated with 5 mL equilibrated StrepTactin resin (IBA) for 2 h, 4 °C. Resin was washed with 50 mL Buffer A, 50 mL Buffer B (50 mM PIPES pH 6.5, 1 M NaCl, 2 mM β-ME), and protein was eluted with 30 mL Buffer C (50 mM PIPES pH 6.5, 300 mM NaCl, 2 mM β-ME, 20 mM DSB). Eluate was pooled and diluted 6× with Buffer D (25 mM PIPES pH 6.5, 1 mM β-ME) and loaded onto a 5 mL Resource S column (GE) equilibrated in 5% Buffer E (buffer D + 1 M NaCl). The column was washed with 5 CV 10% Buffer E and protein was eluted with a 20 CV gradient to 100% Buffer E. Fractions were pooled, concentrated, and loaded onto a HiLoad Superdex200 26/60 column (GE) equilibrated in Buffer F (20 mM CHES pH 9.0, 150 mM NaCl, 1 mM TCEP). Peak fractions were pooled and concentrated at 4 °C. Aliquots were frozen in liquid N_2 and stored at -80 °C.

5.6.5 *S. cerevisiae* and *C. thermophilum* Pan2UCH-Exo and Variants

Sf9 pellets were resuspended in 200 mL Buffer A (100 mM HEPES pH 8.0, 300 mM NaCl, cOmplete protease inhibitors, 2 mM DTT, 500 μM PMSF, 20 μg/mL DNase I, 20 μg/mL RNase A) and lysed by sonication (4 s ON, 8s OFF, 3× 1.5 min, 70% amplitude). Lysate was centrifuged at 235 418 × g for 30 min, 4 °C. Clarified supernatant was incubated with 4 mL Ni^{2+}-NTA resin (GE) for 1.5 h, 4 °C. Resin was washed with 50 mL Buffer A, 50 mL Buffer B (50 mM HEPES pH 8.0, 200 mM NaCl, 30 mM Imidazole, 1 mM DTT) and protein was

eluted with 30 mL Buffer C (50 mM HEPES pH 8.0, 200 mM NaCl, 500 mM Imidazole, 1 mM DTT). Eluate was diluted 4× with Buffer D (25 mM HEPES pH 8.0, 1 mM DTT) and loaded onto a 5 mL HiTrap Q HP column (GE), equilibrated in 5% Buffer E (Buffer D + 1 M NaCl). The column was washed with 7.5% Buffer E and protein was eluted with a 30 CV 7.5-100% Buffer E gradient. For crystallisation trials, the His_8 tag of *ct*Pan2UCH-Exo was cleaved overnight at 4 °C with 200 μL 20 μM 3C protease before repeating the HiTrap Q step. Protein fractions were pooled, concentrated, and loaded onto a Superdex200 26/60 column (GE), equilibrated in Buffer F (20 mM HEPES pH 8.0, 200 mM NaCl, 1 mM TCEP, optionally with 5 mM $CaCl_2$). Peak protein fractions were pooled and concentrated. For crystallography, protein was immediately used to set up crystallisation trays (Section 5.11.1, Section 5.11.2). Otherwise, protein was frozen in liquid N_2 and stored at -80 °C.

5.6.6 *H. sapiens* PAN2(E980A)-PAN3-SII

Sf9 pellets were resuspended in 200 mL Buffer A (100 mM HEPES pH 8.0, 300 mM NaCl, cOmplete protease inhibitors, 2 mM DTT, 500 μM PMSF, 20 μg/mL DNase I), lysed and centrifuged as above (Section 5.6.5), and filtered using 0.8 μm filtration cartridges. Clarified supernatant was incubated with 4 mL equilibrated glutathione sepharose 4B resin (GE) for 2 h, 4 °C. Resin was washed with 50 mL Buffer A, 50 mL Buffer B (50 mM HEPES pH 8.0, 1 M NaCl, 2 mM DTT) and protein was eluted with 30 mL Buffer C (50 mM HEPES pH 8.0, 300 mM NaCl, 2 mM DTT, 20 mM reduced glutathione). Eluate was pooled and diluted 5× with Buffer D (25 mM HEPES pH 8.0, 500 mM NaCl, 2 mM DTT) and loaded onto a 5 mL HiTrap StrepTrap HP column (GE) equilibrated in Buffer D. The column was washed with 5 CV Buffer D and eluted with 5 CV Buffer E (Buffer D + 10 mM DSB). Fractions containing GST-PAN2-PAN3-SII were pooled and digested with 200 μL 20 μM GST-3C protease at 4 °C, overnight. The reaction was incubated with 1 mL glutathione sepharose 4B resin (GE)

for 1 h, 4 °C. Flowthrough was concentrated to 20 mL, diluted 5× in Buffer F (25 mM HEPES pH 8.0, 1 mM TCEP), and loaded onto a 1 mL Resource Q column (GE), which was washed with 10% Buffer G (Buffer F + 1 M NaCl). Protein was eluted by a 20 CV 10-100% Buffer G gradient. Because PAN2-PAN3-SII tended to precipitate during centrifugal concentration, concentration steps were minimised. Peak fractions were pooled and minimally concentrated at 4 °C. Aliquots were frozen in liquid N_2 and stored at -80 °C.

5.6.7 *H. sapiens* PAN2(E980A)-PAN3ΔN278-SII, PAN2(E980A)-PAN3PKC-SII

Sf9 pellets were resuspended in 200 mL Buffer A (100 mM HEPES pH 8.0, 300 mM NaCl, cOmplete protease inhibitors, 2 mM DTT, 500 μM PMSF, 20 μg/mL DNase I, 2 mL Biolock), lysed, centrifuged, and filtered as above (Section 5.6.5). Clarified supernatant was incubated with 4 mL equilibrated StrepTactin resin (IBA) for 2 h, 4 °C. Resin was washed with 50 mL Buffer A and 50 mL Buffer B (50 mM HEPES pH 8.0, 1 M NaCl, 1 mM DTT, 100 μL Biolock). Protein was eluted with 30 mL Buffer C (50 mM HEPES pH 8.0, 300 mM NaCl, 1 mM DTT, 20 mM DSB). Eluate was diluted 6× with Buffer D (25 mM HEPES pH 8.0, 1 mM DTT) and loaded onto a 5 mL HiTrap Q HP column (GE), equilibrated in 5% Buffer E (Buffer D + 1 M NaCl). The column was washed with 10 CV 7.5% Buffer E and protein was eluted with a 25 CV 7.5-100% Buffer E gradient. Protein fractions were concentrated and loaded onto a Superdex200 26/60 column (GE) equilibrated in Buffer F (20 mM HEPES pH 8.0, 250 mM NaCl, 1 mM TCEP). Peak fractions were pooled, concentrated at 4 °C, frozen in liquid N_2, and stored at -80 °C.

5.6.8 *S. cerevisiae* Pab1 (RRM1-4)

E. coli pellets were resuspended in Buffer A (100 mM HEPES pH 8.0, 300 mM NaCl, 2 mM DTT, 20 μg/mL DNase I, 20 μg/mL RNase A, cOmplete protease inhibitors, 500 μM PMSF). Cells were lysed by sonication (4s ON, 8s OFF, 70% amplitude, 5× 1 min) and centrifuged at 235 418 × g for 30 min, 4 °C. Clarified lysate was incubated with 4 mL glutathione sepharose 4B resin (GE) for 2 h, 4 °C. Resin was washed with 50 mL Buffer A, 50 mL Buffer B (50 mM HEPES pH 8.0, 1 M NaCl, 2 mM DTT), and protein was eluted with 30 mL Buffer C (50 mM HEPES pH 8.0, 300 mM NaCl, 2 mM DTT, 20 mM reduced glutathione). Eluate was diluted 6× in Buffer D (50 mM HEPES pH 8.0, 2 mM DTT) and loaded onto a 5 mL HiTrap Q HP column (GE), equilibrated in 5% Buffer E (Buffer D + 1 M NaCl). The column was washed with 10 CV 7.5% Buffer E and protein was eluted with a 25 CV 7.5-100% Buffer E gradient. Protein was cleaved overnight with 200 μL 20 μM 3C protease at 4 °C. The reaction was diluted 3× in Buffer D and loaded onto a 5 mL HiTrap Heparin HP column equilibrated in 10% Buffer E. The column was washed with 6 CV 15% Buffer E and protein was eluted with a 20 CV 15-100% Buffer E gradient. Concentrated fractions were loaded onto a Superdex75 26/60 column (GE), equilibrated in Buffer F (25 mM HEPES pH 8.0, 150 mM NaCl, 1 mM TCEP). Fractions were concentrated at 4 °C, frozen in liquid N_2, and stored at -80 °C.

5.6.9 *H. sapiens* PABPC1 (RRMH, RRM1-4)

E. coli pellets were resuspended in Buffer A (100 mM HEPES pH 8.0, 500 mM NaCl, 2 mM DTT, 20 μg/mL DNase I, 10 μg/mL RNase A, cOmplete protease inhibitors, 1 mM PMSF, 2 mL Biolock). Cells were lysed and centrifuged as above (Section 5.6.8). Clarified lysate was incubated with 4 mL StrepTactin resin (IBA) for 6 h, 4 °C to liberate PABPC1 from bound RNA. Resin was washed with 100 mL Buffer A and 100 mL Buffer B (50 mM HEPES pH 8.0, 1 M NaCl, 1 mM DTT). Protein was eluted with 30 mL Buffer C (50 mM HEPES pH 8.0,

300 mM NaCl, 1 mM DTT, 20 mM DSB). Eluate was diluted 5× in Buffer D (25 mM HEPES pH 8.0, 1 mM DTT) and loaded onto a 5 mL HiTrap S HP column, equilibrated in 5% Buffer E (Buffer D + 1 M NaCl). The column was washed with 10% Buffer E, and protein was eluted with a 25 CV 10-100% Buffer E gradient. Pooled fractions were digested with 200 μL 20 μM TEV protease overnight at 4 °C. This was diluted 8× in Buffer D and loaded onto a 5 mL HiTrap Heparin HP column equilibrated in 5% Buffer E. The column was washed with 10% Buffer E, and protein was eluted with a 25 CV 10-100% Buffer E gradient. Fractions were concentrated and loaded onto a Superdex75 26/60 column (GE). Protein fractions were concentrated at 4 °C, frozen in liquid N_2, and stored at -80 °C.

5.6.10 *H. sapiens* MEX3A-SII, MEX3B-SII, MEX3C-SII

Sf9 pellets were resuspended in Buffer A (100 mM HEPES pH 8.0, 300 mM NaCl, 2 mM DTT, 20 μg/mL DNase I, cOmplete protease inhibitors, 1 mM PMSF, 2 mL Biolock), lysed, and clarified as above (Section 5.6.5). Clarified supernatant was incubated with 4 mL StrepTactin resin (IBA) for 2 h, 4 °C. The resin was washed with 50 mL Buffer A and 50 mL Buffer B (50 mM HEPES pH 8.0, 1 M NaCl, 2 mM DTT). Protein was eluted with 30 mL Buffer C (50 mM HEPES pH 8.0, 200 mM NaCl, 2 mM DTT, 20 mM DSB). Eluate was diluted 4× in Buffer D (25 mM HEPES pH 8.0, 1 mM DTT) and loaded onto a 5 mL HiTrap Q HP column (GE), equilibrated with 5% Buffer E (25 mM HEPES pH 8.0, 1 mM DTT, 1 M NaCl). The column was washed with 10% Buffer E and protein was eluted with a 40 CV 10-100% Buffer E gradient. Fractions were pooled and cleaved overnight with 200 μL 3C protease. The digestion reaction was diluted in Buffer D to 50 mM NaCl and the HiTrap Q step was repeated. Fractions were pooled, concentrated, and loaded onto an S75 16/60 column (GE), equilibrated in Buffer F (20 mM HEPES pH 8.0, 200 mM NaCl, 1 mM TCEP). Peak fractions were pooled, concentrated, flash-frozen in liquid N_2, and stored at -80 °C.

5.7 RNA Preparation

5.7.1 RNA Synthesis

Fluorescently-labelled RNA for *in vitro* deadenylation assays, fluorescence polarisation assays, and analytical size-exclusion chromatography were synthesised by IDT with a 5' 6-FAM or 5' Alexa647 label. Longer RNAs or RNAs containing 2-aminopurine (2AP), purine (P), *N6*-methyladenosine (m^6A) or inosine (I) were synthesised by Horizon Discovery with a 5' 6-FAM label or unlabelled. Unlabelled RNA for circular dichroism spectroscopy or crystallography were synthesized by IDT.

Sequences with specific nomenclature are shown in Table 5.10. Sequences with standard one-letter abbreviations are otherwise used in the text.

Table 5.10 List of synthesised RNA sequences used in this dissertation.

RNA	Sequence
X_{20}-	CAGCUCCGCAUCCCUUUCCC-
X_{10}-	CAGCUCCGCU-
MRE-X_{20}-	CCCAGAGCCCUUUAGUCUUC-
MRE-	CAGAGUUUAG-
P2-	AACUGUUCCUGUAAAUACGCCAG-
Z2-	AAUCAUCCUUAUUUAUUACCAUU-
V2-	AACUUUUUCUCUGGCAGAAAAGU-
H2-	AUAGAUUUGGUCAAUACUGACCAAAUCU-
A3-	ACCAAAAAAAAAAAAAACGCCAG-

5.7.2 *In vitro* Transcription

A pUC57 vector with a T7 transcription start site, the $20mer_{short}$ sequence, a poly(A) tail, and flanking 5' BamHI and 3' BsaI restriction sites was designed by Michael Webster. BsaI cleaves at a defined site upstream of the recognition sequence, enabling an RNA terminating in poly(A) to be transcribed. The vector was transformed into TOP10 *E. coli* and colonies were used to inoculate 200 mL LB + Amp. The culture was shaken at 200 rpm, 37 °C, overnight. The culture was centrifuged at 6000 × g, 30 min and plasmid DNA was purified using the Plasmid Plus Mega kit (Qiagen) according to manufacturers' instructions. Plasmid was linearised overnight, 37 °C by BsaI (NEB) and checked by agarose gel electrophoresis for completion (Section 5.5.8). Linearised plasmid was used in a 5 mL transcription reaction with 250 μL 20× transcription buffer (800 mM Tris pH 8.0, 20 mM spermidine, 0.2% Triton-X-100, 100 mM DTT), 1.6 mL NTPs (ratio depending on sequence), 150 μL 1 M $MgCl_2$, 62.5 μL pyrophosphatase, and 50 μL T7 polymerase. The reaction was incubated at 37 °C, 4 h, mixed with 5 mL loading dye (80% formamide, 10 mM EDTA pH 8.0, 1 mg/mL bromophenol blue), and loaded onto a pre-warmed 10% 20×40 cm denaturing gel (Section 5.9.1.1), run at 37 W until the dye front reached the bottom. The RNA band was visualised by UV shadowing and excised. RNA was electroeluted with the Whatman Elutrap Electroelution System Kit according to manufacturers' instructions (ThermoFisher Scientific). RNA was concentrated by 15 mL centrifugal filtration units (Millipore) and stored at -80 °C.

Table 5.11 List of *in vitro* transcribed RNAs used in this dissertation.

RNA	Sequence
$20mer_{short}$-	GGGAGGGAUCCCCCUUUCCC-

5.8 Biophysical Experiments

5.8.1 Fluorescence Polarisation Anisotropy (FPA)

FPA assays were carried out in triplicate in flat-bottomed 384-well black polystyrene low-binding microplates (100 μL for high-affinity (K_d < 50 nM) interactions; 20 μL for low-affinity (K_d > 50 nM) interactions) (Corning). 5' 6-FAM/Alexa488-labelled RNAs (final concentration: 0.05 nM), a two-fold dilution series of protein (lowest concentration ≈ 0.1 nM), and 1× FPA buffer (10 mM HEPES pH 7.5, 100 mM NaCl, 1 mM $Mg(OAc)_2$, 0.1 mM TCEP) were mixed in wells and topped up with MQ H_2O. The last well of each column was a negative control (no protein). The plate was incubated at RT, 1 h and scanned in a PHERAstar FSX plate reader (BMG Labtech) with correct lasers (495 nm) and filters (520 nm) for stimulation and detection. The focal plane was adjusted to be at maximum fluorescence (~0.5× the plate height) and gain reference was run on control wells to detect 40 mP. Data analysis was carried out using GraphPad Prism 8 to plot mean polarisation against protein concentration. Error bars represent standard deviation. Polarisation was normalised by subtracting the negative control, and dissociation constants (K_d) were estimated using Hill-slope non-linear regression for low-affinity or highly cooperative interactions, or a quadratic function for high-affinity interactions.

5.8.2 Microscale Thermophoresis (MST)

MST assays were carried out using standard capillaries with Monolith NT.115 (NanoTemper). 100 nM (or 15 nM) 5' 6-FAM-labelled A_{15} RNA was mixed with a two-fold dilution series of *sc*Pan2UCH-Exo E912A and 1× MST buffer (10 mM HEPES pH 7.5, 100 mM NaCl, 1 mM $Mg(OAc)_2$, 0.1 mM TCEP, 0.05 mg/mL BSA) in Protein LoBind tubes (Eppendorf). These were incubated at RT for 1 h. Capillaries were scanned by the MST laser at 75% power and fluorescence was monitored using the correct lasers (495 nm) and filters (520 nm).

5.8.3 Differential Scanning Fluorimetry (DSF)

DSF was carried out in a Prometheus NT.48 (NanoTemper). 0.1 or 0.2 mg/mL *sc*Pan2UCH-Exo E912A (1.32 or 2.64 μM) was incubated with/without 10 μM unlabelled A_{15} RNA in 1× DSF buffer (10 mM HEPES pH 7.5, 120 mM NaCl, 1 mM $Mg(OAc)_2$, 0.1 mM TCEP). The mixture was incubated at RT for 1 h. Absorbance at 330 nm (tryptophans buried) and 355 nm (tryptophans solvent-exposed) was monitored across a temperature gradient of 18-95 °C. First derivatives of melting curves were plotted against temperature in GraphPad Prism for analysis.

5.8.4 Circular Dichroism (CD) Spectroscopy

CD assays were carried out on a J-815 spectrometer (Jasco). Unlabelled RNA was incubated in 1× CD buffer (20 mM HEPES pH 8.0, 150 mM NaCl, 1 mM $Mg(OAc)_2$, 1 mM TCEP) in DNA LoBind 0.5 mL tubes (Eppendorf) at RT for 1 h. RNA concentration was chosen to give a minimum peak signal of 5 mdeg. 200 μL of sample was added to 1 mm glass cuvettes (Jasco) which were washed 3× with 70% ethanol, MQ H_2O, and CD buffer between runs. Samples were scanned at 340-200 nm at 0.5 nm intervals and 50 nm/min. The instrument was adjusted to have 100 mdeg sensitivity with 1 s response speed. Resulting ellipticity (in mdeg) was plotted against wavelength in GraphPad Prism for qualitative comparison. Data were normalised by subtraction of a buffer only control. Assays were carried out at least in duplicate.

5.9 *In vitro* Biochemical Experiments

5.9.1 Deadenylation Assays

Deadenylation assays were designed based on published methods (Webster et al., 2017). The specific activity of the same protein/protein complex across different preparations was assumed to be identical. 50 µL assays were performed at RT in Protein LoBind tubes (Eppendorf) in deadenylation buffer (20 mM PIPES pH 6.8, 10 mM KCl, 2 mM $Mg(OAc)_2$, 0.1 mM TCEP). RNA was at 200 nM. The deadenylase was diluted to 10× the required concentration in 20 mM HEPES pH 8.0, 300 mM NaCl, 1 mM $Mg(OAc)_2$, 1 mM TCEP and kept on ice before addition. If two components were added, the deadenylase and additional protein were diluted to 10× the indicated concentration in 10 mM HEPES pH 8.0, 150 mM NaCl, 0.5 mM $Mg(OAc)_2$, 0.5 mM TCEP. This ensured that salt concentration and pH of the reaction were kept constant across different runs. Additional proteins were incubated with the RNA for 1 h prior to the reaction. The reaction was topped up to 45 µL with DEPC H_2O.

Deadenylation reactions were started by adding 5 µL deadenylase. 4 µL aliquots were removed at indicated time points and mixed immediately with formamide loading dye (96% formamide, 10 mM EDTA pH 7.5, 0.3% bromophenol blue, 1% SDS). Samples were resolved on a 20% urea denaturing polyacrylamide gel (Section 5.9.1.1), run at 350 V for 50-75 min in 1× TBE. Fluorescently-labelled RNA was imaged on a Typhoon FLA 7000 (GE) (stimulation: 473 nm, emission filter: 520 nm) or Typhoon 5 Biomolecular Imaging System (GE) (stimulation: 488 nm, emission filter: 520 nm for 6-FAM; stimulation: 635 nm, emission filter: 685 nm for Alexa647). Unlabelled RNA was stained (Section 5.9.1.2) and imaged on a Typhoon 5 System (GE) (stimulation: 488 nm; emission filter: 520 nm). Gels were scanned at 50 µm pixel size. Images were analysed in Adobe Photoshop with adjustments to increase contrast.

5.9.1.1 Denaturing Polyacrylamide Gel Electrophoresis

Gel mixes were prepared by dissolving 7 M urea in 20% (w/v) 19:1 gas-stabilised acrylamide:bis-acrylamide mix (Accugel, National Diagnostics) and 1× TBE. The solution was degassed, filter-sterilised with 0.22 μm vacuum filtration units (Millipore), and stored at 4 °C. 7 mL of gel mix was mixed with 35 μL 10% (v/v) APS (Sigma-Aldrich) and 3.5 μL TEMED (Sigma-Aldrich), and left to polymerise in 1 mm-spacer glass plates in Mini-PROTEAN casting systems (Bio-Rad) for 3-4 h with the appropriate Mini-PROTEAN gel combs inserted. Gels were assembled in Mini-PROTEAN Tetra cells (Bio-Rad) before running. For 20 × 40 cm gels, 200 mL of gel mix was used with the same APS and TEMED ratios and left to polymerise in 20 × 40 cm glass plates with 2 mm spacers. Gels were assembled in custom electrophoresis systems (Cambridge Electrophoresis) before running.

5.9.1.2 SYBR Staining

2 μL SYBR-Green II (Bio-Rad) was diluted in 20 mL 1× TBE and mixed in a dark container. The gel was transferred to the container and incubated on an orbital shaker for 10 min. The stain was decanted and the gel was destained for 10 min in deionised water before imaging.

5.9.1.3 Quantification of Deadenylation Assays

Gels were run and scanned as above (Section 5.9.1). In FIJI (Schindelin et al., 2012), the band intensity of non-deadenylated substrate was measured by boxing the band and integrating the resulting peak from the intensity profile. At each time point, intensity was normalised by dividing by that at time = 0. These measurements were repeated in triplicate. Normalised intensities were plotted against time in GraphPad Prism 8, with each point representing the mean normalised intensity and error bars representing standard deviation. Adjacent points were connected by straight lines to aid analysis.

5.9.2 Electrophoretic Mobility Shift Assay (EMSA)

Gels were prepared as above (Section 5.9.1.1); urea was omitted and acrylamide concentration was 6-10% (w/v). In 10 μL, 5' 6-FAM-labelled or unlabelled RNAs (FAM: 10 nM (affinity), 200 nM (stoichiometry); unlabelled: 5 μM) were incubated with 1 μL 10× protein in EMSA buffer (10 mM HEPES pH 7.5, 100 mM NaCl, 1 mM $Mg(OAc)_2$, 0.1 mM TCEP) and 1 μL 10× loading dye (30% glycerol 0.2% (w/v) Orange G) at RT, 1 h. Stoichiometry assays were carried out in deadenylation buffer (20 mM PIPES pH 6.8, 10 mM KCl, 2 mM $Mg(OAc)_2$, 0.1 mM TCEP). Reactions were run on native TBE polyacrylamide gels at 100 V at RT in Mini-PROTEAN Tetra cells (Bio-Rad) until the dye reached the bottom. Gels were imaged as above (Section 5.9.1). For quantitation, EMSAs were repeated 3×; band intensity of unbound RNA was measured using FIJI (Schindelin et al., 2012) and plotted against protein concentration in GraphPad Prism 8. K_d was estimated using Hill-slope non-linear regression.

5.9.3 Radioactive EMSA

Oligoribonucleotides were radiolabelled with γ^{32}P-ATP (Sigma-Aldrich) and T4 PNK (NEB) according to manufacturers' instructions. Free radiolabel was removed by denaturing polyacrylamide gel electrophoresis (Section 5.9.1.1) on a 20% gel at 200 V, 1 h and bands were visualised using CL-Xposure Film (ThermoFisher) by 1 min exposure. Radioactive bands were excised and radioactivity was measured by a Geiger Counter at 10 cm. Gel pieces were incubated in 10 mM NaOAc, pH 5.5 overnight. RNA was precipitated by adding 3× ethanol and 1 μL GlycoBlue (ThermoFisher) and incubating on dry ice, 1 h. RNA was redissolved in 50 μL TE, and 1 μL was used in EMSAs as above (Section 5.9.2). Bands were detected as above with 30 min exposure. Nucleic acid concentration was estimated by excising the lane with no protein and measuring the radioactivity as above.

5.10 *In vitro* Biochemical Experiments (Protein)

5.10.1 Assessment of *E. coli* Overexpression

Two/four single colonies from a freshly transformed BL21 (DE3) STAR plate (Section 5.2.3) were picked and inoculated in 5 mL TB with the correct antibiotic at 37 °C, overnight. 1 mL of this culture was mixed with 1 mL 100% glycerol and stored at -80 °C. 500 μL of this culture was used to inoculate 50 mL TB media with the correct antibiotic at 37 °C until $OD_{600} \approx 0.5$. Overexpression was induced by adding 1 mM IPTG and the culture was incubated at 18 °C overnight. This was centrifuged at 3000 × g, 4 °C and the pellet was resuspended in 2 mL chilled PBS. *E. coli* was lysed by sonication (VCX750 Vibra-cell processor (Sonics)) with a small tip at 10% for 1 min. The supernatant was incubated with 100 μL resin (glutathione sepharose 4B (GE) for GST-tagged proteins, StrepTactin (IBA) for SII-tagged proteins) at 4 °C, 2 h. The resin was centrifuged at 15,000 × g, 10 min. Resin was washed twice with 200 μL PBS and protein was eluted with 20 μL PBS with 10 mM reduced glutathione or 10 mM DSB at 4 °C, 20 min. 2 μL eluate was analysed by SDS-PAGE (Section 5.6.1.1). Overexpression was assessed by a dominant band at the predicted molecular weight and the glycerol stock was used in subsequent *E. coli* overexpression (Section 5.2.5).

5.10.2 Assessment of Sf9 Overexpression

Two colonies from a freshly transformed bacmid plate (Section 5.2.3) were picked and bacmid preparation, transfection, and virus amplification were carried out as above (Section 5.2.4, Section 5.4.3, Section 5.4.4). Each P2 virus was used to inoculate a 50 mL culture at the same virus:media ratio as in overexpression. Cultures were incubated at 27 °C for four days. At 24 h intervals, the cell density, viability, and fluorescence were monitored as above (Section 5.4). 5 mL of culture was removed, centrifuged at 2000 × g, 10 min, frozen in liquid N_2, and stored

at -80 °C. Cell pellets were resuspended in 2 mL pre-chilled PBS and lysed by rotating with sterile glass beads at 4 °C for 10 min. Supernatant was incubated with 50 µL resin (Ni^{2+}-NTA resin (GE) for His-tagged proteins; StrepTactin (IBA) for SII-tagged proteins) at 4 °C, 2 h. The resin was centrifuged at 15,000 × g, 10 min. Resin was washed twice with 200 µL PBS and protein was eluted with 20 µL PBS with 100 mM imidazole or 10 mM DSB at 4 °C, 20 min. 2 µL eluate was analysed by SDS-PAGE as above (Section 5.6.1.1). Overexpression was assessed by a dominant band at the predicted molecular weight. The ideal time of harvest was assessed as a compromise between cell density (time of arrest), high viability, high fluorescence, and high protein expression. Maximum expression was usually achieved 1-1.5 days after all cells were fluorescent and arrested. This time of harvest was used in Sf9 overexpression cultures.

5.10.3 Pull-down Assays

1 µL MagStrep XT resin (IBA) was equilibrated in 50 µL pull-down buffer (25 mM HEPES pH 7.5, 100 mM NaCl, 2 mM $Mg(OAc)_2$, 0.01% Tween-20, 0.5 mM TCEP). 50 µL of 1 µM bait protein (*sc*Pan2E912A-SII-Pan3; *hs*PAN2E980A-PAN3-SII or variants thereof) diluted in pull-down buffer was incubated with the resin on a rotating platform at 4 °C, 1 h. Supernatant was removed with a gel-loading pipette tip on a magnetic rack. The resin was washed with 50 µL pull-down buffer. 2-6× excess (*sc*Pab1, *hs*PABPC1, *hs*MEX3) prey protein (optionally with RNA) was incubated with the resin on a rotating platform at 4 °C, 1 h. 10 µL slurry was removed (INPUT). The slurry was washed 3× with 50 µL pull-down buffer to ensure removal of non-specifically bound prey protein. Resin was incubated with 30 µL pull-down buffer + 20 mM DSB at 4 °C, 1 h and 10 µL eluate was removed (PULL-DOWN). The input and pull-down fractions were analysed by SDS-PAGE (Section 5.6.1.1). Negative controls of prey proteins only were included to ensure no non-specific binding to the resin.

5.10.4 Analytical Size Exclusion Chromatography (SEC)

SEC was used to assay complex formation (*hs*PAN2E980A-PAN3-SII with *hs*PABPC1 and variants thereof) or BS3 crosslinking efficiency. To assay complex formation, 10 µL of 10 µM PAN2-PAN3 was incubated with the desired PABPC1 stoichiometry (±RNA) in SEC buffer (20 mM HEPES pH 8.0, 250 mM NaCl, 1 mM TCEP) in 25 µL. This was incubated at 4 °C, 1 h and was injected onto a Superose6 3.2/300 (GE) or Superdex200 3.2/300 column (GE). SEC buffer was applied at 0.06 mL/min and 50 µL fractions were collected in a 96-well plate (Corning). Complex formation was assayed by coelution by SDS-PAGE (Section 5.6.1.1), incorporation of nucleic acid in the complex ($A_{260} > A_{280}$), and shift of eluted complexes to higher molecular weights. To assay crosslinking, the reaction was concentrated to 50 µL and injected via a 50 µL loop onto a Superose6 3.2/300 (GE) column. SEC buffer was applied at 0.06 mL/min and 50 µL fractions were collected in a 96-well plate (Corning). Crosslinking efficiency was assayed by SDS-PAGE (Section 5.6.1.1). 3-8% Tris-acetate gels (Invitrogen) and 1× Tris-acetate buffer were used to separate high-molecular weight protein samples.

5.10.5 BS3 Crosslinking

Optimal BS3 concentration for crosslinking a specific complex was carried out by mixing 5 µL 15 µM complex with 10× BS3 dissolved in MQ H_2O (final concentration: 0-4 mM) in crosslinking buffer (20 mM HEPES pH 8.0, 200 mM NaCl, 1 mM TCEP) in 50 µL. The reaction was incubated on ice for 30 min and was quenched with 100 mM NH_4HCO_3. The reaction was further incubated on ice for 5 min. 10 µL of the crosslinking reaction was analysed by SDS-PAGE (Section 5.6.1.1). 3-8% Tris-acetate gels (Invitrogen) and 1× Tris-acetate buffer were used to separate high-molecular weight proteins. For crosslinking for EM, the procedure was repeated as above with 0.75 mM BS3 (final concentration) and scaled up to 250 µL.

5.11 X-ray Crystallography

5.11.1 Initial Crystallisation Screens

Crystallisation screens were from Hampton Research, Molecular Dimensions, Rigaku, and Jena Bioscience. 80 μL was transferred to the reservoir of MRC2 crystallisation microplates (SWISSCI), sealed with a Freedom Evo plate sealer (Tecan), and stored at 10 °C. Plates were centrifuged at 1000 × g, 1 min to remove condensation. 70-100 μM freshly purified protein (Section 5.6) was used to set up crystallisation screens using a Mosquito Crystal Nanolitre Liquid Handling Robot (TTP Labtech). 100 nL protein was mixed with 100 nL mother liquor in sitting drops; plates were sealed and stored at 20 °C. Crystals of *S. cerevisiae* Pan2UCH-Exo appeared in several conditions between 24-96 h incubation and were harvested using CrystalCap SPINE HT cryoloops (Hampton Research) in mother liquor + 20% glycerol.

5.11.2 Optimisation Screens

Based on crystal morphology and initial data collection on crystals from initial screens, conditions with easily handled crystals and high-resolution information were selected for optimisations screens. 0.4 M $NH_4H_2PO_4$ (Hampton Research Crystal Screen 1, A3) was found to be best. Four-corner screens varying $NH_4H_2PO_4$ concentration (Sigma Aldrich) were set up using a Dragonfly Crystal Screen Optimiser (TTP Labtech). Optimisation screens for other conditions were also carried out as above. Notably, condition 1 is derived from the crystallisation condition of *sc*Pan2UE from Schafer and colleagues (Schafer et al., 2014).

1. 100 mM sodium acetate (pH: 3.5-5.6); PEG-200 (concentration: 0-40%)

2. 100 mM sodium citrate (pH: 3.5-5.9); ammonium sulphate (concentration: 0.1-1.5 M)

3. Sodium/potassium phosphate mix (pH: 5.0-8.2; concentration: 0.8-1.8 M)

200 nL protein was mixed with 200 nL mother liquor in sitting drops. For RNA co-crystallisation trials, 100 nL 120-140 μM protein + 100 nL 2 mM RNA in TE (or 100 nL TE) was mixed with 200 nL mother liquor.

Crystals from these optimisation screens were also optimised for cryoprotectant. In brief, mother liquor + 25% glycerol/25% ethylene glycol/30% PEG-200/30% PEG-400 (Sigma-Aldrich) was made as a 10 μL stock. Crystals were transferred to 1 μL on a coverslip before harvesting. RNA was insoluble in low-molecular weight PEGs; 25% ethylene glycol appeared to provide higher resolution and was thus selected for further data collection. Crystals were harvested and handled using CrystalCap SPINE HT cryoloops (Hampton Research).

5.11.3 RNA Soaking

DNA oligonucleotides have an estimated maximum solubility of ~10 mM in aqueous solutions (IDT). RNA oligonucleotides (A_5, A_7, A_{10}, AAGGAA, AAGGA, AACCAA, AAUUAA) were thus dissolved in TE to 10 mM. Because cocrystallisation efforts were limited by throughput and low occupancy of the RNA substrate (low maximum RNA concentration), crystals from four-corner screens were instead used in RNA soak experiments.

RNA soaks were optimised for time and concentration. In brief, 20 μL solutions of mother liquor with different RNA concentrations (1-5 mM) were made up. Cryoprotectant contained the mother liquor, RNA, and 25% ethylene glycol. For time courses, crystals from optimisation plates were fished and transferred to 0.5 μL drops containing RNA (2 or 5 mM) in clean MRC 2 crystallisation plates (SWISSCI). Wells were sealed and incubated for 2 h, 5 h, or overnight. Crystals were fished, transferred to the cryoprotectant (1 μL on a glass cover slip), and rapidly frozen in liquid N_2. For RNA concentration trials, crystals from optimisation plates were transferred to 1-5 mM RNA solutions and incubated for 5 h, before being fished, transferred to cryoprotectant, and frozen as above. Final soak conditions are shown below in Table 5.12.

Table 5.12 List of crystals used in final data collection and RNA soak conditions.

Name	Soak conditions
*sc*Pan2UCH-ExoE912A-A_5	5 mM, overnight
*sc*Pan2UCH-ExoE912A-A_7	5 mM, overnight
*sc*Pan2UCH-ExoE912A-A_{10}	5 mM, overnight
*sc*Pan2UCH-ExoE912A-AAGGA	5 mM, 5 h
*sc*Pan2UCH-ExoE912A-AAGGAA	5 mM, 2 h
*sc*Pan2UCH-ExoE912A-AAUUAA	5 mM, 6 h
*sc*Pan2UCHExoE912A-AACCAA	5 mM, 6 h

5.11.4 Data Collection

Diffraction data for *sc*Pan2UE, *sc*Pan2UE-A_5, -A_7, and -A_{10} were recorded at 100 K at Diamond Light Source beamline I03 with a Pilatus3 6M detector (Dectris) at 0.9763 Å. Diffraction data for *sc*Pan2UE-A_7 + Ca^{2+}, -AAGGA, -AAGGAA, -AACCAA, and -AAUUAA were recorded at 100 K at Diamond Light Source beamline I04-1 with a Pilatus 6M-F (25 Hz) detector (Dectris) at 0.9159 Å. Screening was carried out by 0.5 s exposure at 70-100% transmission and 0.5° oscillations at 0° and 45°. The data collection strategy was determined according to iMOSFLM (Battye et al., 2011) for maximum redundancy, depending on the determined crystal symmetry. Diffraction data were collected by 0.2 s exposure at the transmission used for screening, with 0.2° oscillations.

5.11.5 Structure Determination and Refinement

Diffraction data was truncated according to R_{merge} against batch number. Diffraction data were indexed and integrated using DIALS within the Xia2 pipeline (Winter, 2010). Datasets were scaled and merged with AIMLESS as part of CCP4i2 (Potterton et al., 2018). Resolution cut-off was determined by $CC_{1/2} > 0.5$ and $I/\sigma I > 1.0$ in the outer resolution shell (Karplus and Diederichs, 2012). The apo *S. cerevisiae* Pan2UCH-Exo structure was solved by molecular replacement using Phaser (McCoy et al., 2007) with *Neurospora crassa* Pan2UCH-Exo (PDB: 4CZW) (Jonas et al., 2014) as a search model followed by manual building. Subsequent ligand-bound structures were refined against the apo structure. RNA in the active site was visualised first by mFo-DFc maps, built using RCrane (Keating and Pyle, 2012), and further positioned in Coot (Emsley et al., 2010) with the help of feature-enhanced maps (FEMs) (Afonine et al., 2015). The structures were iteratively refined using PHENIX.refine, and maps were calculated in PHENIX (Adams et al., 2011).

5.11.6 Structural Representation

Structural figures were rendered in PyMOL (Schrodinger, 2015). The FFT tool in CCP4i was used to generate FEM and mFo-DFc maps of the asymmetric unit. These were imported into and displayed in PyMOL using the *isomesh* feature. Superposition analysis for calculation of backbone RMSD (against a previously solved structure) (Schafer et al., 2014), as well as superposition analysis between different RNA-bound Pan2UCH-Exo structures, were carried out with PyMOL. Structure-based alignments were carried out using SSM and LSQ superposition implemented in Coot.

5.12 Negative Stain Electron Microscopy (EM)

5.12.1 Sample Preparation

2% (w/v) uranyl acetate (Sigma-Aldrich) was dissolved in MQ H_2O, filter-sterilised, and stored at RT in the dark. 200 mesh carbon-coated copper grids (CF400-Cu, EMS) were glow-discharged in a PELCO easiGlow machine (Ted Pella) for 45 s at 0.30 mBar and 15 mA with the holey carbon side facing up. 2 × 30 μL drops of uranyl acetate (per grid) were pipetted onto parafilm. Sample (*sc*Pan23 and *bs*PAN23) was diluted to 25-50 nM in 20 mM HEPES pH 8.0, 200 mM NaCl, 1 mM TCEP. 3 μL purified protein (Section 5.6) was pipetted onto the carbon side of the grid, held by Dumoxel N5 tweezers (Dumont), and incubated at RT, 1 min. The sample was blotted with a sharp-tipped filter paper (Whatman) perpendicular to the tweezer. The sample was stained and washed in each drop of uranyl acetate for 30 s. The sample was re-blotted as above and allowed to dry for 1 min before storage at 20 °C.

5.12.2 Data Collection

Data collection was carried out on a Tecnai Spirit TEM (FEI) at 120 kV. Grids were transferred to a room-temperature side-entry holder and inserted into the TEM. Regions of relatively thin stain and good contrast between particles and background stain were identified. 150-200 micrographs were collected with an applied flux density (dose) of ~50 $e^-/Å^2$. Collection was carried out with an Ultrascan 1000 CCD camera (Gatan) at 15,000× or 21,000× magnification with 0.1-1.0 μm underfocus and a 40 nm objective aperture inserted.

5.12.3 Data Processing and Analysis

Particle homogeneity and sample quality were qualitatively assessed by comparing particles to their predicted size (~20 nm). Swarm particle picking was carried out using E2boxer (Tang et al., 2007). Picked coordinates were imported into RELION 3.0 (Zivanov et al., 2018). Particles were extracted with a box size ~2× the predicted particle size and classified (25 iterations) using reference-free 2D classification. Classes were selected and 2D classification was repeated. Particles from 2D classes which displayed sufficient features were used to carry out 3D refinement, using Pan2-Pan3 merged and superposed from several crystal structures as an initial model (Jonas et al., 2014; Schafer et al., 2014; Wolf et al., 2014). In brief, the merged crystal structures were converted from .pdb to .mrc, resized according to negative stain data processing parameters, and low-pass filtered by 50 Å prior to refinement (T = 4, 25 iterations). The 3D map obtained was qualitatively assessed using Chimera (Pettersen et al., 2004).

5.13 Cryo-electron Microscopy (Cryo-EM)

5.13.1 Sample Preparation

Uncrosslinked *hs*PAN2-PAN3 (and variants thereof) were diluted to 0.75-1.5 μM in 20 mM HEPES pH 8.0, 250 mM NaCl, 1 mM TCEP (optionally with 0.04-0.25% NOβG or CHAPSO) or in 20 mM HEPES pH 7.5, 50 mM NaCl, 100 mM KOAc, 5 mM $Mg(OAc)_2$, 2 mM DTT. For graphene oxide grids, uncrosslinked complexes were diluted to 150-400 nM in 20 mM HEPES pH 8.0, 250 mM NaCl, 1 mM TCEP. Crosslinking of *hs*PAN2-PAN3 was carried out as above (Section 5.10.5) and separated by size exclusion chromatography (Section 5.10.4). Peak fractions corresponding to crosslinked Pan2-Pan3 (and not higher-order aggregates) were pooled, concentrated in 500 μL Vivaspin concentrators (Sartorius), and diluted to 0.5-1.0 μM in the above buffer.

UltraAuFoil R1.2/1.3 300 mesh gold grids (QuantiFoil) (Russo and Passmore, 2014) were glow-discharged in a Model 1070 NanoClean Plasma Cleaner (Fischione) in 9:1 Ar:O_2 at 35 W power for 30-45 s. For graphene oxide deposition, gold grids were glow discharged in a S150B glow discharger (Edwards) for 1 min. 2% (w/v) graphene oxide suspension (Sigma-Aldrich) was diluted to 0.2% in MQ H_2O and centrifuged at 300 × g to remove large aggregates. 3 μL of this suspension was applied to the above gold grids and incubated at 4 °C for 1 min. Grids were washed 3× with MQ H_2O and allowed to dry overnight.

Grids were immediately used for vitrification. 3 μL of protein sample was applied to grids in a Vitrobot Mark IV cryo-plunger (FEI), whose blotting chamber had been pre-incubated at 4 °C and 100% humidity. Grids were blotted using homemade filter paper (Whatman) at -10 blot force for 2.0-4.5 s (dependent on previous optimisation) and immediately plunged into liquid ethane. Grids were stored in liquid nitrogen until data collection.

5.13.2 Data Collection

For screening, grids were transferred to a Model 626 cryo-holder (Gatan) at ~-180 °C and examined in a F20 Tecnai-G2 TEM (FEI) with a Falcon II detector (FEI). Grids were screened for ice thickness, particle homogeneity/density, and graphene oxide coverage. Micrographs were collected at 62,000× (1.64 Å/pixel) in linear mode at ~50 $e^-/Å^2$ at 3.0-5.0 μm underfocus.

Data collection was carried out on a Titan Krios (FEI) at 300 kV. Movies were collected on a Falcon III direct electron detector (FEI) in linear mode (flux density: 70 $e^-/Å^2$) or counting mode (flux density: 50 $e^-/Å^2$). Data was collected at a nominal magnification of 59,000× for linear mode (1.33 Å/pixel on Krios 1, 1.38 Å/pixel on Krios 2) or 75,000× (1.04 Å/pixel on Krios 1) for counting mode. A defocus of 2.5-5.0 μm was used. For one dataset, Volta phase plate (Danev et al., 2017) was used; for this dataset, defocus values of 0.5-0.9 μm were used.

5.13.3 Data Processing and Analysis

Cryo-EM movies were corrected using MotionCor2 (Zheng et al., 2017). Motion-corrected micrographs had CTF estimated using CTFFIND4 (Rohou and Grigorieff, 2015) or Gctf (Zhang, 2016). At each step, micrographs were checked for quality (empty holes, ice contamination, stage drift, astigmatism) and discarded accordingly. CTF-corrected micrographs were used in particle picking using either a template-based method in RELION-3.0 (Zivanov et al., 2018) or a machine-learning method in Cryolo (Wagner et al., 2019). All subsequent processing steps were carried out in RELION-3.0. Particles were extracted with a box size ~50% greater than the particle diameter (~18 nm). To eliminate poorly centred selections or particles which are too small to account for the full complex, two to three rounds of 2D classification were carried out with 25 iterations and a regularisation parameter (*T*) of 2. The mask size to display 2D class averages is 18 nm.

5.14 Bioinformatics

5.14.1 Sequence Alignment

Amino acid sequence alignments were carried out using the T-Coffee suite (Notredame et al., 2000). Protein sequences were submitted in the .fasta format, and alignments were output in the .clustalw_aln format. Display of sequence alignments were carried out using ESPript 3 (Robert and Gouet, 2014) using shading by percentage equivalence.

5.14.2 Poly(A) Site Sequence Analysis

Poly(A) sites in humans were obtained from PolyA_DB3 from Bin Tian's laboratory (Wang et al., 2018b). To assess if GG or GGG motifs were more prevalent proximal to the poly(A) site, 20 nucleotides 5' to the poly(A) site were isolated and motif prevalence was analysed. The percentage prevalence of these motifs was calculated and was compared to the expected percentage if these motifs were randomly distributed.

5.14.3 Secondary Structure Prediction

Protein secondary structure prediction for domain boundary prediction and construct design, as well as structural based alignment of proteins across different species, was carried out in JPred 4 (Drozdetskiy et al., 2015). Prediction of protein disorder was carried out using DISOPRED3 (Jones and Cozzetto, 2015). Visualisation of secondary structure and disorder predictions were carried out in JalView 2 (Waterhouse et al., 2009).

5.14.4 RNA Structure Prediction

RNA secondary structure prediction was carried out to design RNA substrates using the RNAstructure web server (https://rna.urmc.rochester.edu/RNAstructureWeb/) from the Mathews Lab (Reuter and Mathews, 2010). Default settings were used for RNA structure prediction.

5.14.5 Prediction of Protein Properties

The molecular weight, molar extinction coefficients at 280 nm (ε_{280}), and pK_a of individual proteins or protein complexes were calculated and estimated from their amino acid composition using ProtParam (https://web.expasy.org/protparam/).

5.15 Others

5.15.1 Schematic Diagrams

Schematic diagrams and models were created in BioRender (https://biorender.com/). Domain diagrams were created in Illustrator for Biological Sequences (IBS) (Liu et al., 2015). Chemical structures of nucleotide bases were created in ChemDraw (PerkinElmer).

5.15.2 Gel Representation

DNA and Protein gels were imaged using a ChemiDoc XRS+ imaging system (Bio-Rad). Gel cropping, contrast adjustments, and further analysis were carried out using Adobe PhotoShop. Intensity measurements of gels were carried out in Fiji (Schindelin et al., 2012). RNA gels were imaged using a Typhoon FLA 7000 or a Typhoon 5 Gel and Blot Imaging System (GE).

6. References

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Echols, N., Headd, J.J., Hung, L.W., Jain, S., Kapral, G.J., Grosse Kunstleve, R.W., *et al.* (2011). The Phenix software for automated determination of macromolecular structures. Methods *55*, 94-106.

Adesnik, M., Salditt, M., Thomas, W., and Darnell, J.E. (1972). Evidence that all messenger RNA molecules (except histone messenger RNA) contain Poly (A) sequences and that the Poly(A) has a nuclear function. J Mol Biol *71*, 21-30.

Adivarahan, S., Livingston, N., Nicholson, B., Rahman, S., Wu, B., Rissland, O.S., and Zenklusen, D. (2018). Spatial Organization of Single mRNPs at Different Stages of the Gene Expression Pathway. Mol Cell *72*, 727-738 e725.

Afonina, Z.A., Myasnikov, A.G., Shirokov, V.A., Klaholz, B.P., and Spirin, A.S. (2014). Formation of circular polyribosomes on eukaryotic mRNA without cap-structure and poly(A)-tail: a cryo electron tomography study. Nucleic Acids Res *42*, 9461-9469.

Afonine, P.V., Moriarty, N.W., Mustyakimov, M., Sobolev, O.V., Terwilliger, T.C., Turk, D., Urzhumtsev, A., and Adams, P.D. (2015). FEM: feature-enhanced map. Acta Crystallogr D Biol Crystallogr *71*, 646-666.

Ahlquist, P., and Kaesberg, P. (1979). Determination of the length distribution of poly(A) at the 3' terminus of the virion RNAs of EMC virus, poliovirus, rhinovirus, RAV-61 and CPMV and of mouse globin mRNA. Nucleic Acids Res *7*, 1195-1204.

Aizer, A., Kalo, A., Kafri, P., Shraga, A., Ben-Yishay, R., Jacob, A., Kinor, N., and Shav-Tal, Y. (2014). Quantifying mRNA targeting to P-bodies in living human cells reveals their dual role in mRNA decay and storage. Journal of Cell Science *127*, 4443-4456.

Ameres, S.L., and Zamore, P.D. (2013). Diversifying microRNA sequence and function. Nat Rev Mol Cell Biol *14*, 475-488.

Anastasakis, D., Skeparnias, I., Shaukat, A.N., Grafanaki, K., Kanellou, A., Taraviras, S., Papachristou, D.J., Papakyriakou, A., and Stathopoulos, C. (2016). Mammalian PNLDC1 is a novel poly(A) specific exonuclease with discrete expression during early development. Nucleic Acids Research *44*, 8908-8920.

Andersen, K.R., Jonstrup, A.T., Van, L.B., and Brodersen, D.E. (2009). The activity and selectivity of fission yeast Pop2p are affected by a high affinity for Zn2+ and Mn2+ in the active site. RNA *15*, 850-861.

Applequist, J., and Damie, V. (1966). Thermodynamics of the one-stranded helix-coil equilibrium in polyadenylic acid. J Am Chem Soc *88*, 3895-3900.

Ariz, M., Mainpal, R., and Subramaniam, K. (2009). C. elegans RNA-binding proteins PUF-8 and MEX-3 function redundantly to promote germline stem cell mitosis. Dev Biol *326*, 295-304.

Arnott, S., Chandrasekaran, R., and Leslie, A.G. (1976). Structure of the single-stranded polyribonucleotide polycytidylic acid. J Mol Biol *106*, 735-748.

Arthur, L., Pavlovic-Djuranovic, S., Smith-Koutmou, K., Green, R., Szczesny, P., and Djuranovic, S. (2015). Translational control by lysine-encoding A-rich sequences. Sci Adv *1*.

Arvola, R.M., Chang, C.T., Buytendorp, J.P., Levdansky, Y., Valkov, E., Freddolino, P.L., and Goldstrohm, A.C. (2020). Unique repression domains of Pumilio utilize deadenylation and decapping factors to accelerate destruction of target mRNAs. Nucleic Acids Res *48*, 1843-1871.

Aslam, A., Mittal, S., Koch, F., Andrau, J.C., and Winkler, G.S. (2009). The Ccr4-NOT deadenylase subunits CNOT7 and CNOT8 have overlapping roles and modulate cell proliferation. Mol Biol Cell *20*, 3840-3850.

Astrom, J., Astrom, A., and Virtanen, A. (1991). In vitro deadenylation of mammalian mRNA by a HeLa cell 3' exonuclease. EMBO J *10*, 3067-3071.

Astrom, J., Astrom, A., and Virtanen, A. (1992). Properties of a HeLa cell 3' exonuclease specific for degrading poly(A) tails of mammalian mRNA. J Biol Chem *267*, 18154-18159.

Axhemi, A., Wasmuth, E.V., Lima, C.D., and Jankowsky, E. (2020). Substrate selectivity by the exonuclease Rrp6p. Proc Natl Acad Sci U S A *117*, 982-992.

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

Baejen, C., Torkler, P., Gressel, S., Essig, K., Soding, J., and Cramer, P. (2014). Transcriptome maps of mRNP biogenesis factors define pre-mRNA recognition. Mol Cell *55*, 745-757.

Baer, B.W., and Kornberg, R.D. (1980). Repeating structure of cytoplasmic poly(A)-ribonucleoprotein. Proc Natl Acad Sci U S A *77*, 1890-1892.

Baer, B.W., and Kornberg, R.D. (1983). The protein responsible for the repeating structure of cytoplasmic poly(A)-ribonucleoprotein. J Cell Biol *96*, 717-721.

Baggs, J.E., and Green, C.B. (2003). Nocturnin, a deadenylase in Xenopus laevis retina: a mechanism for posttranscriptional control of circadian-related mRNA. Curr Biol *13*, 189-198.

Baglioni, C., Pemberton, R., and Delovitch, T. (1972). Presence of polyadenylic acid sequences in RNA of membrane-bound polyribosomes. FEBS Lett *26*, 320-322.

Balbo, P.B., and Bohm, A. (2007). Mechanism of poly(A) polymerase: structure of the enzyme-MgATP-RNA ternary complex and kinetic analysis. Structure *15*, 1117-1131.

Barrass, J.D., Mendoza-Ochoa, G.I., Maudlin, I.E., Sani, E., and Beggs, J.D. (2019). Tuning Degradation to Achieve Specific and Efficient Protein Depletion. J Vis Exp.

Basquin, J., Roudko, V.V., Rode, M., Basquin, C., Seraphin, B., and Conti, E. (2012). Architecture of the nuclease module of the yeast Ccr4-not complex: the Not1-Caf1-Ccr4 interaction. Mol Cell *48*, 207-218.

Battye, T.G., Kontogiannis, L., Johnson, O., Powell, H.R., and Leslie, A.G. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr D Biol Crystallogr *67*, 271-281.

Bawankar, P., Loh, B., Wohlbold, L., Schmidt, S., and Izaurralde, E. (2013). NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain. RNA Biol *10*, 228-244.

Bazzini, A.A., Lee, M.T., and Giraldez, A.J. (2012). Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish. Science *336*, 233-237.

Beilharz, T.H., and Preiss, T. (2007). Widespread use of poly(A) tail length control to accentuate expression of the yeast transcriptome. RNA *13*, 982-997.

Beilharz, T.H., and Preiss, T. (2011). Polyadenylation state microarray (PASTA) analysis. Methods Mol Biol *759*, 133-148.

Belostotsky, D.A., and Meagher, R.B. (1993). Differential organ-specific expression of three poly(A)-binding-protein genes from Arabidopsis thaliana. Proc Natl Acad Sci U S A *90*, 6686-6690.

Bergmann, I.E., and Brawerman, G. (1977). Control of breakdown of the polyadenylate sequence in mammalian polyribosomes: role of poly(adenylic acid)-protein interactions. Biochemistry *16*, 259-264.

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

Bernstein, P., Peltz, S.W., and Ross, J. (1989). The poly(A)-poly(A)-binding protein complex is a major determinant of mRNA stability in vitro. Mol Cell Biol *9*, 659-670.

Berthet, C., Morera, A.M., Asensio, M.J., Chauvin, M.A., Morel, A.P., Dijoud, F., Magaud, J.P., Durand, P., and Rouault, J.P. (2004). CCR4-associated factor CAF1 is an essential factor for spermatogenesis. Mol Cell Biol *24*, 5808-5820.

Bett, J.S., Ibrahim, A.F.M., Garg, A.K., Kelly, V., Pedrioli, P., Rocha, S., and Hay, R.T. (2013). The P-body component USP52/PAN2 is a novel regulator of HIF1A mRNA stability. Biochemical Journal *451*, 185-194.

Bhaskar, V., Roudko, V., Basquin, J., Sharma, K., Urlaub, H., Seraphin, B., and Conti, E. (2013). Structure and RNA-binding properties of the Not1-Not2-Not5 module of the yeast Ccr4-Not complex. Nat Struct Mol Biol *20*, 1281-1288.

Bianchin, C., Mauxion, F., Sentis, S., Seraphin, B., and Corbo, L. (2005). Conservation of the deadenylase activity of proteins of the Caf1 family in human. RNA *11*, 487-494.

Bisaria, N., Jarmoskaite, I., and Herschlag, D. (2017). Lessons from Enzyme Kinetics Reveal Specificity Principles for RNA-Guided Nucleases in RNA Interference and CRISPR-Based Genome Editing. Cell Syst *4*, 21-29.

Boeck, R., Tarun, S., Rieger, M., Deardorff, J.A., MullerAuer, S., and Sachs, A.B. (1996). The yeast Pan2 protein is required for poly(A)-binding protein-stimulated poly(A)-nuclease activity. Journal of Biological Chemistry *271*, 432-438.

Boland, A., Chen, Y., Raisch, T., Jonas, S., Kuzuoglu-Ozturk, D., Wohlbold, L., Weichenrieder, O., and Izaurralde, E. (2013). Structure and assembly of the NOT module of the human CCR4-NOT complex. Nat Struct Mol Biol *20*, 1289-1297.

Brahms, J. (1964). Circular Dichroism Investigations of 2 Conformations of Polyriboadenylic Acid. Nature *202*, 797-&.

Brahms, J., Michelson, A.M., and Van Holde, K.E. (1966). Adenylate oligomers in single- and double-strand conformation. J Mol Biol *15*, 467-488.

Brandt, F., Carlson, L.A., Hartl, F.U., Baumeister, W., and Grunewald, K. (2010). The three-dimensional organization of polyribosomes in intact human cells. Mol Cell *39*, 560-569.

Braun, J.E., Huntzinger, E., Fauser, M., and Izaurralde, E. (2011). GW182 Proteins Directly Recruit Cytoplasmic Deadenylase Complexes to miRNA Targets. Molecular Cell *44*, 120-133.

Brawerman, G., and Diez, J. (1975). Metabolism of the polyadenylate sequence of nuclear RNA and messenger RNA in mammalian cells. Cell *5*, 271-280.

Bregman, A., Avraham-Kelbert, M., Barkai, O., Duek, L., Guterman, A., and Choder, M. (2011). Promoter elements regulate cytoplasmic mRNA decay. Cell *147*, 1473-1483.

Brewer, G., and Ross, J. (1988). Poly(A) shortening and degradation of the 3' A+U-rich sequences of human c-myc mRNA in a cell-free system. Mol Cell Biol *8*, 1697-1708.

Brooks, S.A., and Blackshear, P.J. (2013). Tristetraprolin (TTP): interactions with mRNA and proteins, and current thoughts on mechanisms of action. Biochim Biophys Acta *1829*, 666-679.

Brown, C.E., and Sachs, A.B. (1998a). Poly(A) tail length control in Saccharomyces cerevisiae occurs by message-specific deadenylation. Molecular and Cellular Biology *18*, 6548-6559.

Brown, C.E., and Sachs, A.B. (1998b). Poly(A) tail length control in Saccharomyces cerevisiae occurs by message-specific deadenylation. Mol Cell Biol *18*, 6548-6559.

Brown, C.E., Tarun, S.Z., Boeck, R., and Sachs, A.B. (1996). PAN3 encodes a subunit of the Pab1p-dependent poly(A) nuclease in Saccharomyces cerevisiae. Molecular and Cellular Biology *16*, 5744-5753.

Brown, R.F., Andrews, C.T., and Elcock, A.H. (2015). Stacking Free Energies of All DNA and RNA Nucleoside Pairs and Dinucleoside-Monophosphates Computed Using Recently Revised AMBER Parameters and Compared with Experiment. Journal of Chemical Theory and Computation *11*, 2315-2328.

Buchet-Poyau, K., Courchet, J., Le Hir, H., Seraphin, B., Scoazec, J.Y., Duret, L., Domon-Dell, C., Freund, J.N., and Billaud, M. (2007). Identification and characterization of human Mex-3 proteins, a novel family of evolutionarily conserved RNA-binding proteins differentially localized to processing bodies. Nucleic Acids Res *35*, 1289-1300.

Burd, C.G., Matunis, E.L., and Dreyfuss, G. (1991). The Multiple Rna-Binding Domains of the Messenger-Rna Poly(a)-Binding Protein Have Different Rna-Binding Activities. Molecular and Cellular Biology *11*, 3419-3424.

Buschauer, R., Matsuo, Y., Sugiyama, T., Chen, Y.H., Alhusaini, N., Sweet, T., Ikeuchi, K., Cheng, J., Matsuki, Y., Nobuta, R.*, et al.* (2020). The Ccr4-Not complex monitors the translating ribosome for codon optimality. Science *368*.

Cano, F., Bye, H., Duncan, L.M., Buchet-Poyau, K., Billaud, M., Wills, M.R., and Lehner, P.J. (2012). The RNA-binding E3 ubiquitin ligase MEX-3C links ubiquitination with MHC-I mRNA degradation. EMBO J *31*, 3596-3606.

Cao, D., and Parker, R. (2001). Computational modeling of eukaryotic mRNA turnover. RNA *7*, 1192-1212.

Cao, D., and Parker, R. (2003). Computational modeling and experimental analysis of nonsense-mediated decay in yeast. Cell *113*, 533-545.

Caponigro, G., and Parker, R. (1995). Multiple functions for the poly(A)-binding protein in mRNA decapping and deadenylation in yeast. Genes Dev *9*, 2421-2432.

Caput, D., Beutler, B., Hartog, K., Thayer, R., Brownshimer, S., and Cerami, A. (1986). Identification of a Common Nucleotide-Sequence in the 3'-Untranslated Region of Messenger-Rna Molecules Specifying Inflammatory Mediators. Proceedings of the National Academy of Sciences of the United States of America *83*, 1670-1674.

Carthew, R.W., and Sontheimer, E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. Cell *136*, 642-655.

Chan, S.I., and Nelson, J.H. (1969). Proton magnetic resonance studies of ribose dinucleoside monophoshates in aqueous solution. I. The nature of the base-stacking interaction in adenylyl 3'--5')adenosine. J Am Chem Soc *91*, 168-183.

Chandrasekaran, V., Juszkiewicz, S., Choi, J., Puglisi, J.D., Brown, A., Shao, S., Ramakrishnan, V., and Hegde, R.S. (2019). Mechanism of ribosome stalling during translation of a poly(A) tail. Nat Struct Mol Biol *26*, 1132-1140.

Chang, H., Lim, J., Ha, M., and Kim, V.N. (2014). TAIL-seq: genome-wide determination of poly(A) tail length and 3' end modifications. Mol Cell *53*, 1044-1052.

Chao, H., Deng, L., Xu, F., Yu, Z., Xu, X., Huang, J., and Zeng, T. (2019). MEX3C regulates lipid metabolism to promote bladder tumorigenesis through JNK pathway. Onco Targets Ther *12*, 3285-3294.

Charlesworth, A., Meijer, H.A., and de Moor, C.H. (2013). Specificity factors in cytoplasmic polyadenylation. Wiley Interdiscip Rev RNA *4*, 437-461.

Chekulaeva, M., Mathys, H., Zipprich, J.T., Attig, J., Colic, M., Parker, R., and Filipowicz, W. (2011). miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs. Nature Structural & Molecular Biology *18*, 1218-U1262.

Chen, A.A., and Garcia, A.E. (2012). Mechanism of enhanced mechanical stability of a minimal RNA kissing complex elucidated by nonequilibrium molecular dynamics simulations. Proc Natl Acad Sci U S A *109*, E1530-1539.

Chen, C.Y., and Shyu, A.B. (2003). Rapid deadenylation triggered by a nonsense codon precedes decay of the RNA body in a mammalian cytoplasmic nonsense-mediated decay pathway. Mol Cell Biol *23*, 4805-4813.

Chen, C.Y.A., Zhang, Y.Q., Xiang, Y., Han, L., Chang, J.T., and Shyu, A.B. (2018). Antagonistic actions of two human Pan3 isoforms on global mRNA turnover (vol 19, pg 1639, 2014). Rna *24*, 1607-1612.

Chen, C.Y.A., Zheng, D.H., Xia, Z.F., and Shyu, A.B. (2009). Ago-TNRC6 triggers microRNA-mediated decay by promoting two deadenylation steps. Nature Structural & Molecular Biology *16*, 1160-U1166.

Chen, J., Chiang, Y.C., and Denis, C.L. (2002). CCR4, a 3'-5' poly(A) RNA and ssDNA exonuclease, is the catalytic component of the cytoplasmic deadenylase. EMBO J *21*, 1414-1426.

Chen, J., Rappsilber, J., Chiang, Y.C., Russell, P., Mann, M., and Denis, C.L. (2001). Purification and characterization of the 1.0 MDa CCR4-NOT complex identifies two novel components of the complex. J Mol Biol *314*, 683-694.

Chen, J.L., and Greider, C.W. (2003). Determinants in mammalian telomerase RNA that mediate enzyme processivity and cross-species incompatibility. EMBO J *22*, 304-314.

Chen, R., and Wold, M.S. (2014). Replication protein A: single-stranded DNA's first responder: dynamic DNA-interactions allow replication protein A to direct single-strand DNA intermediates into different pathways for synthesis or repair. Bioessays *36*, 1156-1161.

Chen, Y., Boland, A., Kuzuoglu-Ozturk, D., Bawankar, P., Loh, B., Chang, C.T., Weichenrieder, O., and Izaurralde, E. (2014). A DDX6-CNOT1 Complex and W-Binding Pockets in CNOT9 Reveal Direct Links between miRNA Target Recognition and Silencing. Molecular Cell *54*, 737-750.

Cheng, J., Maier, K.C., Avsec, Z., Rus, P., and Gagneur, J. (2017). Cis-regulatory elements explain most of the mRNA stability variation across genes in yeast. Rna *23*, 1648-1659.

Chiba, Y., Johnson, M.A., Lidder, P., Vogel, J.T., van Erp, H., and Green, P.J. (2004). AtPARN is an essential poly(A) ribonuclease in Arabidopsis. Gene *328*, 95-102.

Choi, B., Rempala, G.A., and Kim, J.K. (2017). Beyond the Michaelis-Menten equation: Accurate and efficient estimation of enzyme kinetic parameters. Sci Rep *7*, 17018.

Chowdhury, A., and Tharun, S. (2009). Activation of decapping involves binding of the mRNA and facilitation of the post-binding steps by the Lsm1-7-Pat1 complex. RNA *15*, 1837-1848.

Christie, M., Boland, A., Huntzinger, E., Weichenrieder, O., and Izaurralde, E. (2013). Structure of the PAN3 Pseudokinase Reveals the Basis for Interactions with the PAN2 Deadenylase and the GW182 Proteins. Molecular Cell *51*, 360-373.

Clark, L.B., Viswanathan, P., Quigley, G., Chiang, Y.C., McMahon, J.S., Yao, G., Chen, J., Nelsbach, A., and Denis, C.L. (2004). Systematic mutagenesis of the leucine-rich repeat (LRR) domain of CCR4 reveals specific sites for binding to CAF1 and a separate critical role for the LRR in CCR4 deadenylase activity. J Biol Chem *279*, 13616-13623.

Collart, M.A., and Struhl, K. (1993). CDC39, an essential nuclear protein that negatively regulates transcription and differentially affects the constitutive and inducible HIS3 promoters. EMBO J *12*, 177-186.

Collart, M.A., and Struhl, K. (1994). NOT1(CDC39), NOT2(CDC36), NOT3, and NOT4 encode a global-negative regulator of transcription that differentially affects TATA-element utilization. Genes Dev *8*, 525-537.

Coller, J., and Parker, R. (2004). Eukaryotic mRNA decapping. Annu Rev Biochem *73*, 861-890.

Corley, M., Burns, M.C., and Yeo, G.W. (2020). How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms. Molecular Cell *78*, 9-29.

Cougot, N., Babajko, S., and Seraphin, B. (2004). Cytoplasmic foci are sites of mRNA decay in human cells. Journal of Cell Biology *165*, 31-40.

Courchet, J., Buchet-Poyau, K., Potemski, A., Bres, A., Jariel-Encontre, I., and Billaud, M. (2008). Interaction with 14-3-3 adaptors regulates the sorting of hMex-3B RNA-binding protein to distinct classes of RNA granules. J Biol Chem *283*, 32131-32142.

Couttet, P., Fromont-Racine, M., Steel, D., Pictet, R., and Grange, T. (1997). Messenger RNA deadenylylation precedes decapping in mammalian cells. Proc Natl Acad Sci U S A *94*, 5628-5633.

Dalluge, J.J., Hashizume, T., Sopchik, A.E., McCloskey, J.A., and Davis, D.R. (1996). Conformational flexibility in RNA: the role of dihydrouridine. Nucleic Acids Res *24*, 1073-1079.

Danev, R., Buijsse, B., Khoshouei, M., Plitzko, J.M., and Baumeister, W. (2014). Volta potential phase plate for in-focus phase contrast transmission electron microscopy. Proc Natl Acad Sci U S A *111*, 15635-15640.

Danev, R., Tegunov, D., and Baumeister, W. (2017). Using the Volta phase plate with defocus for cryo-EM single particle analysis. Elife *6*.

Darnell, J.E., Philipson, L., Wall, R., and Adesnik, M. (1971). Polyadenylic acid sequences: role in conversion of nuclear RNA into messenger RNA. Science *174*, 507-510.

Deardorff, J.A., and Sachs, A.B. (1997). Differential effects of aromatic and charged residue substitutions in the RNA binding domains of the yeast poly(A)-binding protein. J Mol Biol *269*, 67-81.

Decker, C.J., and Parker, R. (1993). A turnover pathway for both stable and unstable mRNAs in yeast: evidence for a requirement for deadenylation. Genes Dev *7*, 1632-1643.

Decker, C.J., and 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*.

Deng, T.T., Huang, Y., Weng, K., Lin, S., Li, Y.J., Shi, G., Chen, Y.L., Huang, J.J., Liu, D., Ma, W.B.*, et al.* (2019). TOE1 acts as a 3 exonuclease for telomerase RNA and regulates telomere maintenance. Nucleic Acids Research *47*, 391-405.

Denis, C.L. (1984). Identification of new genes involved in the regulation of yeast alcohol dehydrogenase II. Genetics *108*, 833-844.

Deo, R.C., Bonanno, J.B., Sonenberg, N., and Burley, S.K. (1999). Recognition of polyadenylate RNA by the poly(A)-binding protein. Cell *98*, 835-845.

Dewey, T.G., and Turner, D.H. (1979). Laser temperature-jump study of stacking in adenylic acid polymers. Biochemistry *18*, 5757-5762.

Dhanraj, S., Gunja, S.M., Deveau, A.P., Nissbeck, M., Boonyawat, B., Coombs, A.J., Renieri, A., Mucciolo, M., Marozza, A., Buoni, S.*, et al.* (2015). Bone marrow failure and developmental delay caused by mutations in poly(A)-specific ribonuclease (PARN). J Med Genet *52*, 738-748.

Dheur, S., Nykamp, K.R., Viphakone, N., Swanson, M.S., and Minvielle-Sebastia, L. (2005). Yeast mRNA poly(A) tail length control can be reconstituted in vitro in the absence of Pab1p-dependent poly(A) nuclease activity. Journal of Biological Chemistry *280*, 24532-24538.

Ding, D.Q., Liu, J.L., Dong, K.Z., Midic, U., Hess, R.A., Xie, H.R., Demireva, E.Y., and Chen, C. (2017). PNLDC1 is essential for piRNA 3 ' end trimming and transposon silencing during spermatogenesis in mice. Nature Communications *8*.

Doidge, R., Mittal, S., Aslam, A., and Winkler, G.S. (2012). The anti-proliferative activity of BTG/TOB proteins is mediated via the Caf1a (CNOT7) and Caf1b (CNOT8) deadenylase subunits of the Ccr4-not complex. PLoS One *7*, e51331.

Doma, M.K., and Parker, R. (2006). Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. Nature *440*, 561-564.

Doma, M.K., and Parker, R. (2007). RNA quality control in eukaryotes. Cell *131*, 660-668.

Dougherty, J.D., White, J.P., and Lloyd, R.E. (2011). Poliovirus-mediated disruption of cytoplasmic processing bodies. J Virol *85*, 64-75.

Draper, B.W., Mello, C.C., Bowerman, B., Hardin, J., and Priess, J.R. (1996). MEX-3 is a KH domain protein that regulates blastomere identity in early C. elegans embryos. Cell *87*, 205-216.

Draper, M.P., Liu, H.Y., Nelsbach, A.H., Mosley, S.P., and Denis, C.L. (1994). CCR4 is a glucose-regulated transcription factor whose leucine-rich repeat binds several proteins important for placing CCR4 in its proper promoter context. Mol Cell Biol *14*, 4522-4531.

Drozdetskiy, A., Cole, C., Procter, J., and Barton, G.J. (2015). JPred4: a protein secondary structure prediction server. Nucleic Acids Res *43*, W389-394.

Du, H., Zhao, Y., He, J.Q., Zhang, Y., Xi, H.R., Liu, M.F., Ma, J.B., and Wu, L.G. (2016). YTHDF2 destabilizes m(6)A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. Nature Communications *7*.

Dunn, E.F., Hammell, C.M., Hodge, C.A., and Cole, C.N. (2005). Yeast poly(A)-binding protein, Pab1, and PAN, A poly(A) nuclease complex recruited by Pab1, connect mRNA biogenesis to export. Genes & Development *19*, 90-103.

Eichhorn, S.W., Subtelny, A.O., Kronja, I., Kwasnieski, J.C., Orr-Weaver, T.L., and Bartel, D.P. (2016). mRNA poly(A)-tail changes specified by deadenylation broadly reshape translation in Drosophila oocytes and early embryos. Elife *5*.

Eisen, T.J., Eichhorn, S.W., Subtelny, A.O., Lin, K.S., McGeary, S.E., Gupta, S., and Bartel, D.P. (2019). The Dynamics of Cytoplasmic mRNA Metabolism. Mol Cell.

Eisen, T.J., Eichhorn, S.W., Subtelny, A.O., Lin, K.S., McGeary, S.E., Gupta, S., and Bartel, D.P. (2020). The Dynamics of Cytoplasmic mRNA Metabolism. Mol Cell *77*, 786-799 e710.

Eisenberg, H., and Felsenfeld, G. (1967). Studies of the temperature-dependent conformation and phase separation of polyriboadenylic acid solutions at neutral pH. J Mol Biol *30*, 17-37.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501.

Estrella, M.A., Du, J., Chen, L., Rath, S., Prangley, E., Chitrakar, A., Aoki, T., Schedl, P., Rabinowitz, J., and Korennykh, A. (2019). The metabolites NADP(+) and NADPH are the targets of the circadian protein Nocturnin (Curled). Nat Commun *10*, 2367.

Estruch, F., Peiro-Chova, L., Gomez-Navarro, N., Durban, J., Hodge, C., del Olmo, M., and Cole, C.N. (2009). A genetic screen in Saccharomyces cerevisiae identifies new genes that interact with mex67-5, a temperature-sensitive allele of the gene encoding the mRNA export receptor. Molecular Genetics and Genomics *281*, 125-134.

Eulalio, A., Huntzinger, E., Nishihara, T., Rehwinkel, J., Fauser, M., and Izaurralde, E. (2009). Deadenylation is a widespread effect of miRNA regulation. Rna *15*, 21-32.

Faber, A.W., Van Dijk, M., Raue, H.A., and Vos, J.C. (2002). Ngl2p is a Ccr4p-like RNA nuclease essential for the final step in 3'-end processing of 5.8S rRNA in Saccharomyces cerevisiae. RNA *8*, 1095-1101.

Fabian, M.R., Frank, F., Rouya, C., Siddiqui, N., Lai, W.S., Karetnikov, A., Blackshear, P.J., Nagar, B., and Sonenberg, N. (2013). Structural basis for the recruitment of the human CCR4-NOT deadenylase complex by tristetraprolin. Nat Struct Mol Biol *20*, 735-739.

Fasman, G.D., Lindblow, C., and Grossman, L. (1964). The Helical Conformations of Polycytidylic Acid: Studies on the Forces Involved. Biochemistry *3*, 1015-1021.

Fox, C.A., and Wickens, M. (1990). Poly(A) removal during oocyte maturation: a default reaction selectively prevented by specific sequences in the 3' UTR of certain maternal mRNAs. Genes Dev *4*, 2287-2298.

Fraser, C.S., Pain, V.M., and Morley, S.J. (1999). Cellular stress in Xenopus kidney cells enhances the phosphorylation of eukaryotic translation initiation factor (eIF)4E and

the association of elF4F with poly(A)-binding protein. Biochemical Journal *342*, 519-526.

Fresco, J.R., and Klemperer, E. (1959). Polyriboadenylic Acid, a Molecular Analogue of Ribonucleic Acid and Desoxyribonucleic Acid. Annals of the New York Academy of Sciences *81*, 730-741.

Friedel, C.C., Dolken, L., Ruzsics, Z., Koszinowski, U.H., and Zimmer, R. (2009). Conserved principles of mammalian transcriptional regulation revealed by RNA half-life. Nucleic Acids Res *37*, e115.

Friedman, R.A., and Honig, B. (1995). A free energy analysis of nucleic acid base stacking in aqueous solution. Biophys J *69*, 1528-1535.

Fuke, H., and Ohno, M. (2008). Role of poly (A) tail as an identity element for mRNA nuclear export. Nucleic Acids Research *36*, 1037-1049.

Funakoshi, Y., Doi, Y., Hosoda, N., Uchida, N., Osawa, M., Shimada, I., Tsujimoto, M., Suzuki, T., Katada, T., and Hoshino, S.I. (2007). Mechanism of mRNA deadenylation: evidence for a molecular interplay between translation termination factor eRF3 and mRNA deadenylases. Genes & Development *21*, 3135-3148.

Gallie, D.R. (1991). The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. Genes Dev *5*, 2108-2116.

Gao, M., Fritz, D.T., Ford, L.P., and Wilusz, J. (2000). Interaction between a poly(A)-specific ribonuclease and the 5' cap influences mRNA deadenylation rates in vitro. Mol Cell *5*, 479-488.

Garneau, N.L., Sokoloski, K.J., Opyrchal, M., Neff, C.P., Wilusz, C.J., and Wilusz, J. (2008). The 3' untranslated region of sindbis virus represses deadenylation of viral transcripts in mosquito and Mammalian cells. J Virol *82*, 880-892.

Garneau, N.L., Wilusz, J., and Wilusz, C.J. (2007). The highways and byways of mRNA decay. Nat Rev Mol Cell Biol *8*, 113-126.

Geisberg, J.V., Moqtaderi, Z., Fan, X., Ozsolak, F., and Struhl, K. (2014). Global analysis of mRNA isoform half-lives reveals stabilizing and destabilizing elements in yeast. Cell *156*, 812-824.

Gherzi, R., Lee, K.Y., Briata, P., Wegmuller, D., Moroni, C., Karin, M., and Chen, C.Y. (2004). A KH domain RNA binding protein, KSRP, promotes ARE-directed mRNA turnover by recruiting the degradation machinery. Mol Cell *14*, 571-583.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B.*, et al.* (2002). Functional profiling of the Saccharomyces cerevisiae genome. Nature *418*, 387-391.

Gibb, B., Ye, L.F., Gergoudis, S.C., Kwon, Y., Niu, H., Sung, P., and Greene, E.C. (2014). Concentration-dependent exchange of replication protein A on single-stranded DNA revealed by single-molecule imaging. PLoS One *9*, e87922.

Gill, T., Cai, T., Aulds, J., Wierzbicki, S., and Schmitt, M.E. (2004). RNase MRP cleaves the CLB2 mRNA to promote cell cycle progression: novel method of mRNA degradation. Mol Cell Biol *24*, 945-953.

Glaunsinger, B.A., and Lee, Y.J. (2010). How tails define the ending: divergent roles for polyadenylation in RNA stability and gene expression. RNA Biol *7*, 13-17.

Goldstrohm, A.C., Hall, T.M.T., and McKenney, K.M. (2018). Post-transcriptional Regulatory Functions of Mammalian Pumilio Proteins. Trends Genet *34*, 972-990.

Goldstrohm, A.C., Seay, D.J., Hook, B.A., and Wickens, M. (2007). PUF protein-mediated deadenylation is catalyzed by Ccr4p. J Biol Chem *282*, 109-114.

Gorski, J., Morrison, M.R., Merkel, C.G., and Lingrel, J.B. (1974). Size heterogeneity of polyadenylate sequences in mouse globin messenger RNA. J Mol Biol *86*, 363-371.

Gowrishankar, G., Winzen, R., Dittrich-Breiholz, O., Redich, N., Kracht, M., and Holtmann, H. (2006). Inhibition of mRNA deadenylation and degradation by different types of cell stress. Biol Chem *387*, 323-327.

Grange, T., de Sa, C.M., Oddos, J., and Pictet, R. (1987). Human mRNA polyadenylate binding protein: evolutionary conservation of a nucleic acid binding motif. Nucleic Acids Res *15*, 4771-4787.

Green, C.B., Douris, N., Kojima, S., Strayer, C.A., Fogerty, J., Lourim, D., Keller, S.R., and Besharse, J.C. (2007). Loss of Nocturnin, a circadian deadenylase, confers resistance to hepatic steatosis and diet-induced obesity. Proc Natl Acad Sci U S A *104*, 9888-9893.

Greenberg, J.R. (1980). Proteins crosslinked to messenger RNA by irradiating polyribosomes with ultraviolet light. Nucleic Acids Res *8*, 5685-5701.

Groner, B., Hynes, N., and Phillips, S. (1974). Length heterogeneity in the poly (adenylic acid) region of yeast messenger ribonucleic acid. Biochemistry *13*, 5378-5383.

Groner, B., and Phillips, S.L. (1975). Polyadenylate metabolism in the nuclei and cytoplasm of Saccharomyces cerevisiae. J Biol Chem *250*, 5640-5646.

Grossi de Sa, M.F., Standart, N., Martins de Sa, C., Akhayat, O., Huesca, M., and Scherrer, K. (1988). The poly(A)-binding protein facilitates in vitro translation of poly(A)-rich mRNA. Eur J Biochem *176*, 521-526.

Hammet, A., Pike, B.L., and Heierhorst, J. (2002). Posttranscriptional regulation of the RAD5 DNA repair gene by the Dun1 kinase and the Pan2-Pan3 poly(A)-nuclease complex contributes to survival of replication blocks. Journal of Biological Chemistry *277*, 22469-22474.

Harel-Sharvit, L., Eldad, N., Haimovich, G., Barkai, O., Duek, L., and Choder, M. (2010). RNA polymerase II subunits link transcription and mRNA decay to translation. Cell *143*, 552-563.

Hashizume, H., and Imahori, K. (1967). Circular dichroism and conformation of natural and synthetic polynucleotides. J Biochem *61*, 738-749.

Hegde, R.S., Grossman, S.R., Laimins, L.A., and Sigler, P.B. (1992). Crystal structure at 1.7 A of the bovine papillomavirus-1 E2 DNA-binding domain bound to its DNA target. Nature *359*, 505-512.

Henriksson, N., Nilsson, P., Wu, M.S., Song, H.W., and Virtanen, A. (2010). Recognition of Adenosine Residues by the Active Site of Poly(A)-specific Ribonuclease. Journal of Biological Chemistry *285*, 163-170.

Hilgers, V., Teixeira, D., and Parker, R. (2006). Translation-independent inhibition of mRNA deadenylation during stress in Saccharomyces cerevisiae. RNA *12*, 1835-1845.

Ho, B., Baryshnikova, A., and Brown, G.W. (2018). Unification of Protein Abundance Datasets Yields a Quantitative Saccharomyces cerevisiae Proteome. Cell Syst *6*, 192-205 e193.

Hofweber, M., and Dormann, D. (2019). Friend or foePost-translational modifications as regulators of phase separation and RNP granule dynamics. Journal of Biological Chemistry *294*, 7137-7150.

Holcomb, D.N., and Tinoco, I. (1965). Conformation of Polyriboadenylic Acid - Ph and Temperature Dependence. Biopolymers *3*, 121-&.

Hong, K.Y., Lee, S.H., Gu, S., Kim, E., An, S., Kwon, J., Lee, J.B., and Jang, S.K. (2017). The bent conformation of poly(A)-binding protein induced by RNA-binding is required for its translational activation function. Rna Biology *14*, 370-377.

Horiuchi, M., Takeuchi, K., Noda, N., Muroya, N., Suzuki, T., Nakamura, T., Kawamura-Tsuzuku, J., Takahasi, K., Yamamoto, T., and Inagaki, F. (2009). Structural basis for the antiproliferative activity of the Tob-hCaf1 complex. J Biol Chem *284*, 13244-13255.

Houseley, J., LaCava, J., and Tollervey, D. (2006). RNA-quality control by the exosome. Nat Rev Mol Cell Biol *7*, 529-539.

Hsu, C.L., and Stevens, A. (1993). Yeast cells lacking 5'-->3' exoribonuclease 1 contain mRNA species that are poly(A) deficient and partially lack the 5' cap structure. Mol Cell Biol *13*, 4826-4835.

Huang, K.L., Chadee, A.B., Chen, C.Y.A., Zhang, Y.Q., and Shyu, A.B. (2013). Phosphorylation at intrinsically disordered regions of PAM2 motif-containing proteins modulates their interactions with PABPC1 and influences mRNA fate. Rna-a Publication of the Rna Society *19*, 295-305.

Huang, N.N., and Hunter, C.P. (2015). The RNA binding protein MEX-3 retains asymmetric activity in the early Caenorhabditis elegans embryo in the absence of asymmetric protein localization. Gene *554*, 160-173.

Huarte, J., Stutz, A., O'Connell, M.L., Gubler, P., Belin, D., Darrow, A.L., Strickland, S., and Vassalli, J.D. (1992). Transient translational silencing by reversible mRNA deadenylation. Cell *69*, 1021-1030.

Hubstenberger, A., Courel, M., Benard, M., Souquere, S., Ernoult-Lange, M., Chouaib, R., Yi, Z., Morlot, J.B., Munier, A., Fradet, M.*, et al.* (2017). P-Body Purification Reveals the Condensation of Repressed mRNA Regulons. Molecular Cell *68*, 144-+.

Huez, G., Marbaix, G., Hubert, E., Leclercq, M., Nudel, U., Soreq, H., Salomon, R., Lebleu, B., Revel, M., and Littauer, U.Z. (1974). Role of the polyadenylate segment in the translation of globin messenger RNA in Xenopus oocytes. Proc Natl Acad Sci U S A *71*, 3143-3146.

Huh, W.K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., and O'Shea, E.K. (2003). Global analysis of protein localization in budding yeast. Nature *425*, 686-691.

Huntzinger, E., Kuzuoglu-Ozturk, D., Braun, J.E., Eulalio, A., Wohlbold, L., and Izaurralde, E. (2013). The interactions of GW182 proteins with PABP and deadenylases are required for both translational repression and degradation of miRNA targets. Nucleic Acids Research *41*, 978-994.

Imataka, H., Gradi, A., and Sonenberg, N. (1998). A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. Embo Journal *17*, 7480-7489.

Isaksson, J., Acharya, S., Barman, J., Cheruku, P., and Chattopadhyaya, J. (2004). Single-stranded adenine-rich DNA and RNA retain structural characteristics of their respective double-stranded conformations and show directional differences in stacking pattern. Biochemistry *43*, 15996-16010.

Iserman, C., Desroches Altamirano, C., Jegers, C., Friedrich, U., Zarin, T., Fritsch, A.W., Mittasch, M., Domingues, A., Hersemann, L., Jahnel, M.*, et al.* (2020). Condensation of Ded1p Promotes a Translational Switch from Housekeeping to Stress Protein Production. Cell *181*, 818-831 e819.

Izumi, N., Shoji, K., Sakaguchi, Y., Honda, S., Kirino, Y., Suzuki, T., Katsuma, S., and Tomari, Y. (2016). Identification and Functional Analysis of the Pre-piRNA 3 ' Trimmer in Silkworms. Cell *164*, 962-973.

Jacobson, A., and Favreau, M. (1983). Possible involvement of poly(A) in protein synthesis. Nucleic Acids Res *11*, 6353-6368.

Jacobson, A., and Peltz, S.W. (1999). Regulation of mRNA stability - Tools for turnover: Methods for analysis of mRNA stability in eukaryotic cells. Methods-a Companion to Methods in Enzymology *17*, 1-2.

Janicke, A., Vancuylenberg, J., Boag, P.R., Traven, A., and Beilharz, T.H. (2012). ePAT: a simple method to tag adenylated RNA to measure poly(A)-tail length and other 3' RACE applications. RNA *18*, 1289-1295.

Jankowsky, E., and Putnam, A. (2010). Duplex unwinding with DEAD-box proteins. Methods Mol Biol *587*, 245-264.

Jia, H., Wang, X., Liu, F., Guenther, U.P., Srinivasan, S., Anderson, J.T., and Jankowsky, E. (2011). The RNA helicase Mtr4p modulates polyadenylation in the TRAMP complex. Cell *145*, 890-901.

Jiang, H., Zhang, X., Luo, J., Dong, C., Xue, J., Wei, W., Chen, J., Zhou, J., Gao, Y., and Yang, C. (2012). Knockdown of hMex-3A by small RNA interference suppresses cell proliferation and migration in human gastric cancer cells. Mol Med Rep *6*, 575-580.

Jiao, Y., Bishop, C.E., and Lu, B. (2012a). Mex3c regulates insulin-like growth factor 1 (IGF1) expression and promotes postnatal growth. Mol Biol Cell *23*, 1404-1413.

Jiao, Y., George, S.K., Zhao, Q., Hulver, M.W., Hutson, S.M., Bishop, C.E., and Lu, B. (2012b). Mex3c mutation reduces adiposity and increases energy expenditure. Mol Cell Biol *32*, 4350-4362.

Jonas, S., Christie, M., Peter, D., Bhandari, D., Loh, B., Huntzinger, E., Weichenrieder, O., and Izaurralde, E. (2014). An asymmetric PAN3 dimer recruits a single PAN2 exonuclease to mediate mRNA deadenylation and decay. Nat Struct Mol Biol *21*, 599-608.

Jones, D.T., and Cozzetto, D. (2015). DISOPRED3: precise disordered region predictions with annotated protein-binding activity. Bioinformatics *31*, 857-863.

Jonstrup, A.T., Andersen, K.R., Van, L.B., and Brodersen, D.E. (2007). The 1.4-A crystal structure of the S. pombe Pop2p deadenylase subunit unveils the configuration of an active enzyme. Nucleic Acids Res *35*, 3153-3164.

Juszkiewicz, S., and Hegde, R.S. (2017). Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination. Mol Cell *65*, 743-750 e744.

Karplus, P.A., and Diederichs, K. (2012). Linking crystallographic model and data quality. Science *336*, 1030-1033.

Kates, J., and Beeson, J. (1970). Ribonucleic acid synthesis in vaccinia virus. II. Synthesis of polyriboadenylic acid. J Mol Biol *50*, 19-33.

Keating, K.S., and Pyle, A.M. (2012). RCrane: semi-automated RNA model building. Acta Crystallogr D Biol Crystallogr *68*, 985-995.

Kelly, J.M., and Cox, R.A. (1982). Periodicity in the length of 3'-poly(A) tails from native globin mRNA of rabbit. Nucleic Acids Res *10*, 4173-4179.

Kessler, S.H., and Sachs, A.B. (1998). RNA recognition motif 2 of yeast Pab1p is required for its functional interaction with eukaryotic translation initiation factor 4G. Molecular and Cellular Biology *18*, 51-57.

Khanam, T., Muddashetty, R.S., Kahvejian, A., Sonenberg, N., and Brosius, J. (2006). Poly(A)-binding protein binds to A-rich sequences via RNA-binding domains 1+2 and 3+4. RNA Biol *3*, 170-177.

Khong, A., and Parker, R. (2018). mRNP architecture in translating and stress conditions reveals an ordered pathway of mRNP compaction. Journal of Cell Biology *217*, 4124-4140.

Kim, C., Paulus, B.F., and Wold, M.S. (1994). Interactions of human replication protein A with oligonucleotides. Biochemistry *33*, 14197-14206.

Kim, C., and Wold, M.S. (1995). Recombinant human replication protein A binds to polynucleotides with low cooperativity. Biochemistry *34*, 2058-2064.

Kim, D., Lee, Y.S., Jung, S.J., Yeo, J., Seo, J.J., Lee, Y.Y., Lim, J., Chang, H., Song, J., Yang, J., *et al.* (2020). Viral hijacking of the TENT4-ZCCHC14 complex protects viral RNAs via mixed tailing. Nat Struct Mol Biol.

Kim, J.H., and Richter, J.D. (2006). Opposing polymerase-deadenylase activities regulate cytoplasmic polyadenylation. Mol Cell *24*, 173-183.

Kim, J.L., Nikolov, D.B., and Burley, S.K. (1993). Co-crystal structure of TBP recognizing the minor groove of a TATA element. Nature *365*, 520-527.

Kini, H.K., Silverman, I.M., Ji, X.J., Gregory, B.D., and Liebhaber, S.A. (2016). Cytoplasmic poly(A) binding protein-1 binds to genomically encoded sequences within mammalian mRNAs. Rna *22*, 61-74.

Kojima, S., Gendreau, K.L., Sher-Chen, E.L., Gao, P., and Green, C.B. (2015). Changes in poly(A) tail length dynamics from the loss of the circadian deadenylase Nocturnin. Scientific Reports *5*.

Kondo, N.S., and Danyluk, S.S. (1976). Conformational properties of adenylyl-3' leads to 5'-adenosine in aqueous solution. Biochemistry *15*, 756-768.

Korner, C.G., and Wahle, E. (1997). Poly(A) tail shortening by a mammalian poly(A)-specific 3'-exoribonuclease. J Biol Chem *272*, 10448-10456.

Korner, C.G., Wormington, M., Muckenthaler, M., Schneider, S., Dehlin, E., and Wahle, E. (1998). The deadenylating nuclease (DAN) is involved in poly(A) tail removal during the meiotic maturation of Xenopus oocytes. EMBO J *17*, 5427-5437.

Kozlov, G., De Crescenzo, G., Lim, N.S., Siddiqui, N., Fantus, D., Kahvejian, A., Trempe, J.F., Elias, D., Ekiel, I., Sonenberg, N.*, et al.* (2004). Structural basis of ligand recognition by PABC, a highly specific peptide-binding domain found in poly(A)-binding protein and a HECT ubiquitin ligase. Embo Journal *23*, 272-281.

Kozlov, G., and Gehring, K. (2010). Molecular Basis of eRF3 Recognition by the MLLE Domain of Poly(A)-Binding Protein. Plos One *5*.

Kozlov, G., Menade, M., Rosenauer, A., Nguyen, L., and Gehring, K. (2010). Molecular Determinants of PAM2 Recognition by the MLLE Domain of Poly(A)-Binding Protein. Journal of Molecular Biology *397*, 397-407.

Kozlov, G., Siddiqui, N., Coillet-Matillon, S., Trempe, J.F., Ekiel, I., Sprules, T., and Gehring, K. (2002). Solution structure of the orphan PABC domain from Saccharomyces cerevisiae poly(A)-binding protein. Journal of Biological Chemistry *277*, 22822-22828.

Kozlov, G., Trempe, J.F., Khaleghpour, K., Kahvejian, A., Ekiel, I., and Gehring, K. (2001). Structure and function of the C-terminal PABC domain of human poly(A)-binding protein. Proceedings of the National Academy of Sciences of the United States of America *98*, 4409-4413.

Krause, M., Niazi, A.M., Labun, K., Torres Cleuren, Y.N., Muller, F.S., and Valen, E. (2019). tailfindr: alignment-free poly(A) length measurement for Oxford Nanopore RNA and DNA sequencing. RNA *25*, 1229-1241.

Krowczynska, A., and Brawerman, G. (1986). Structural features in the 3'-terminal region of polyribosome-bound rabbit globin messenger RNAs. J Biol Chem *261*, 397-402.

Kuhn, U., and Pieler, T. (1996). Xenopus poly(A) binding protein: functional domains in RNA binding and protein-protein interaction. J Mol Biol *256*, 20-30.

Kuhn, U., and Wahle, E. (2004). Structure and function of poly(A) binding proteins. Biochimica Et Biophysica Acta-Gene Structure and Expression *1678*, 67-84.

Kumar, A., Clerici, M., Muckenfuss, L.M., Passmore, L.A., and Jinek, M. (2019). Mechanistic insights into mRNA 3'-end processing. Curr Opin Struct Biol *59*, 143-150.

Kuyumcu-Martinez, M., Belliot, G., Sosnovtsev, S.V., Chang, K.O., Green, K.Y., and Lloyd, R.E. (2004). Calicivirus 3C-like proteinase inhibits cellular translation by cleavage of poly(A)-binding protein. J Virol *78*, 8172-8182.

Kuzuoglu-Ozturk, D., Huntzinger, E., Schmidt, S., and Izaurralde, E. (2012). The Caenorhabditis elegans GW182 protein AIN-1 interacts with PAB-1 and subunits of the PAN2-PAN3 and CCR4-NOT deadenylase complexes. Nucleic Acids Research *40*, 5651-5665.

Lackner, D.H., Beilharz, T.H., Marguerat, S., Mata, J., Watt, S., Schubert, F., Preiss, T., and Bahler, J. (2007). A network of multiple regulatory layers shapes gene expression in fission yeast. Mol Cell *26*, 145-155.

Lardelli, R.M., Schaffer, A.E., Eggens, V.R., Zaki, M.S., Grainger, S., Sathe, S., Van Nostrand, E.L., Schlachetzki, Z., Rosti, B., Akizu, N.*, et al.* (2017). Biallelic mutations in the 3' exonuclease TOE1 cause pontocerebellar hypoplasia and uncover a role in snRNA processing. Nat Genet *49*, 457-464.

Le Borgne, M., Chartier, N., Buchet-Poyau, K., Destaing, O., Faurobert, E., Thibert, C., Rouault, J.P., Courchet, J., Negre, D., Bouvard, D.*, et al.* (2014). The RNA-binding protein Mex3b regulates the spatial organization of the Rap1 pathway. Development *141*, 2096-2107.

Le, H., Browning, K.S., and Gallie, D.R. (2000). The phosphorylation state of poly(A)-binding protein specifies its binding to poly(A) RNA and its interaction with eukaryotic initiation factor (eIF) 4F, eIFiso4F, and eIF4B. Journal of Biological Chemistry *275*, 17452-17462.

Lee, C.H., Ezra, F.S., Kondo, N.S., Sarma, R.H., and Danyluk, S.S. (1976). Conformational properties of dinucleoside monophosphates in solution: dipurines and dipyrimidines. Biochemistry *15*, 3627-3639.

Lee, D., Park, D., Park, J.H., Kim, J.H., and Shin, C. (2019). Poly(A)-specific ribonuclease sculpts the 3' ends of microRNAs. RNA *25*, 388-405.

Lee, S.Y., Mendecki, J., and Brawerman, G. (1971). A polynucleotide segment rich in adenylic acid in the rapidly-labeled polyribosomal RNA component of mouse sarcoma 180 ascites cells. Proc Natl Acad Sci U S A *68*, 1331-1335.

Lefrere, V., Vincent, A., and Amalric, F. (1990). Drosophila melanogaster poly(A)-binding protein: cDNA cloning reveals an unusually long 3'-untranslated region of the mRNA, also present in other eukaryotic species [corrected]. Gene *96*, 219-225.

Legnini, I., Alles, J., Karaiskos, N., Ayoub, S., and Rajewsky, N. (2019). FLAM-seq: full-length mRNA sequencing reveals principles of poly(A) tail length control. Nat Methods *16*, 879-886.

Lejeune, F., Li, X., and Maquat, L.E. (2003). Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. Mol Cell *12*, 675-687.

Leng, M., and Felsenfeld, G. (1966). A study of polyadenylic acid at neutral pH. J Mol Biol *15*, 455-466.

Li, F., Zhao, D., Wu, J., and Shi, Y. (2014). Structure of the YTH domain of human YTHDF2 in complex with an m(6)A mononucleotide reveals an aromatic cage for m(6)A recognition. Cell Res *24*, 1490-1492.

Licatalosi, D.D., Ye, X., and Jankowsky, E. (2020). Approaches for measuring the dynamics of RNA-protein interactions. Wiley Interdiscip Rev RNA *11*, e1565.

Licht, K., Hartl, M., Amman, F., Anrather, D., Janisiw, M.P., and Jantsch, M.F. (2019). Inosine induces context-dependent recoding and translational stalling. Nucleic Acids Res *47*, 3-14.

Lim, J., Ha, M., Chang, H., Kwon, S.C., Simanshu, D.K., Patel, D.J., and Kim, V.N. (2014). Uridylation by TUT4 and TUT7 Marks mRNA for Degradation. Cell *159*, 1365-1376.

Lim, J., Kim, D., Lee, Y.S., Ha, M., Lee, M., Yeo, J., Chang, H., Song, J., Ahn, K., and Kim, V.N. (2018). Mixed tailing by TENT4A and TENT4B shields mRNA from rapid deadenylation. Science *361*, 701-704.

Lim, J., Lee, M., Son, A., Chang, H., and Kim, V.N. (2016). mTAIL-seq reveals dynamic poly(A) tail regulation in oocyte-to-embryo development. Genes Dev *30*, 1671-1682.

Lim, L., and Canellakis, E.S. (1970). Adenine-rich polymer associated with rabbit reticulocyte messenger RNA. Nature *227*, 710-712.

Lima, S.A., Chipman, L.B., Nicholson, A.L., Chen, Y.H., Yee, B.A., Yeo, G.W., Coller, J., and Pasquinelli, A.E. (2017). Short poly(A) tails are a conserved feature of highly expressed genes. Nat Struct Mol Biol *24*, 1057-1063.

Lin, J., Kolomeisky, A., and Meller, A. (2010). Helix-coil kinetics of individual polyadenylic acid molecules in a protein channel. Phys Rev Lett *104*, 158101.

Lin, J.X., Fabian, M., Sonenberg, N., and Meller, A. (2012). Nanopore Detachment Kinetics of Poly(A) Binding Proteins from RNA Molecules Reveals the Critical Role of C-Terminus Interactions. Biophysical Journal *102*, 1427-1434.

Liu, H.Y., Badarinarayana, V., Audino, D.C., Rappsilber, J., Mann, M., and Denis, C.L. (1998). The NOT proteins are part of the CCR4 transcriptional complex and affect gene expression both positively and negatively. EMBO J *17*, 1096-1106.

Liu, W., Xie, Y., Ma, J., Luo, X., Nie, P., Zuo, Z., Lahrmann, U., Zhao, Q., Zheng, Y., Zhao, Y.*, et al.* (2015). IBS: an illustrator for the presentation and visualization of biological sequences. Bioinformatics *31*, 3359-3361.

Liu, Y., Nie, H., Liu, H., and Lu, F. (2019). Poly(A) inclusive RNA isoform sequencing (PAIso-seq) reveals wide-spread non-adenosine residues within RNA poly(A) tails. Nat Commun *10*, 5292.

Loh, B., Jonas, S., and Izaurralde, E. (2013). The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. Genes Dev *27*, 2125-2138.

Lorentzen, E., Basquin, J., Tomecki, R., Dziembowski, A., and Conti, E. (2008). Structure of the active subunit of the yeast exosome core, Rrp44: diverse modes of substrate recruitment in the RNase II nuclease family. Mol Cell *29*, 717-728.

Lorentzen, E., and Conti, E. (2005). Structural basis of 3' end RNA recognition and exoribonucleolytic cleavage by an exosome RNase PH core. Mol Cell *20*, 473-481.

Lotan, R., Bar-On, V.G., Harel-Sharvit, L., Duek, L., Melamed, D., and Choder, M. (2005). The RNA polymerase II subunit Rpb4p mediates decay of a specific class of mRNAs. Genes Dev *19*, 3004-3016.

Lotan, R., Goler-Baron, V., Duek, L., Haimovich, G., and Choder, M. (2007). The Rpb7p subunit of yeast RNA polymerase II plays roles in the two major cytoplasmic mRNA decay mechanisms. J Cell Biol *178*, 1133-1143.

Lowell, J.E., Rudner, D.Z., and Sachs, A.B. (1992). 3'-Utr-Dependent Deadenylation by the Yeast Poly(a) Nuclease. Genes & Development *6*, 2088-2099.

Luzzati, V., Mathis, A., Masson, F., and Witz, J. (1964). Structure Transitions Observed in DNA + Poly a in Solution as Function of Temperature + Ph. Journal of Molecular Biology *10*, 28-&.

Malecki, M., Viegas, S.C., Carneiro, T., Golik, P., Dressaire, C., Ferreira, M.G., and Arraiano, C.M. (2013). The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. EMBO J *32*, 1842-1854.

Mangus, D.A., Evans, M.C., Agrin, N.S., Smith, M., Gongidi, P., and Jacobson, A. (2004a). Positive and negative regulation of poly(A) nuclease. Mol Cell Biol *24*, 5521-5533.

Mangus, D.A., Evans, M.C., and Jacobson, A. (2003). Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. Genome Biol *4*, 223.

Mangus, D.A., Smith, M.M., McSweeney, J.M., and Jacobson, A. (2004b). Identification of factors regulating poly(A) tail synthesis and maturation. Molecular and Cellular Biology *24*, 4196-4206.

Mansbridge, J.N., Crossley, J.A., Lanyon, W.G., and Williamson, R. (1974). The poly(adenylic acid) sequence of mouse globin messenger RNA. Eur J Biochem *44*, 261-269.

Marbaix, G., Huez, G., Burny, A., Cleuter, Y., Hubert, E., Leclercq, M., Chantrenne, H., Soreq, H., Nudel, U., and Littauer, U.Z. (1975). Absence of polyadenylate segment in globin

messenger RNA accelerates its degradation in Xenopus oocytes. Proc Natl Acad Sci U S A *72*, 3065-3067.

Martinez, J., Ren, Y.G., Nilsson, P., Ehrenberg, M., and Virtanen, A. (2001). The mRNA cap structure stimulates rate of poly(A) removal and amplifies processivity of degradation. J Biol Chem *276*, 27923-27929.

Maryati, M., Airhihen, B., and Winkler, G.S. (2015). The enzyme activities of Caf1 and Ccr4 are both required for deadenylation by the human Ccr4-Not nuclease module. Biochem J *469*, 169-176.

Mathys, H., Basquin, J., Ozgur, S., Czarnocki-Cieciura, M., Bonneau, F., Aartse, A., Dziembowski, A., Nowotny, M., Conti, E., and Filipowicz, W. (2014). Structural and biochemical insights to the role of the CCR4-NOT complex and DDX6 ATPase in microRNA repression. Mol Cell *54*, 751-765.

Matsuo, Y., Tesina, P., Nakajima, S., Mizuno, M., Endo, A., Buschauer, R., Cheng, J.D., Shounai, O., Ikeuchi, K., Saeki, Y.*, et al.* (2020). RQT complex dissociates ribosomes collided on endogenous RQC substrate SDD1. Nature Structural & Molecular Biology *27*, 323-+.

Mauer, J., Luo, X., Blanjoie, A., Jiao, X., Grozhik, A.V., Patil, D.P., Linder, B., Pickering, B.F., Vasseur, J.J., Chen, Q.*, et al.* (2017). Reversible methylation of m(6)Am in the 5' cap controls mRNA stability. Nature *541*, 371-375.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J Appl Crystallogr *40*, 658-674.

McLaughlin, C.S., Warner, J.R., Edmonds, M., Nakazato, H., and Vaughan, M.H. (1973). Polyadenylic acid sequences in yeast messenger ribonucleic acid. J Biol Chem *248*, 1466-1471.

Melo, E.O., Dhalia, R., Martins de Sa, C., Standart, N., and de Melo Neto, O.P. (2003). Identification of a C-terminal poly(A)-binding protein (PABP)-PABP interaction domain: role in cooperative binding to poly (A) and efficient cap distal translational repression. J Biol Chem *278*, 46357-46368.

Mendecki, J., Lee, S.Y., and Brawerman, G. (1972). Characteristics of the polyadenylic acid segment associated with messenger ribonucleic acid in mouse sarcoma 180 ascites cells. Biochemistry *11*, 792-798.

Mercer, J.F., and Wake, S.A. (1985). An analysis of the rate of metallothionein mRNA poly(A)-shortening using RNA blot hybridization. Nucleic Acids Res *13*, 7929-7943.

Merkel, C.G., Kwan, S., and Lingrel, J.B. (1975). Size of the polyadenylic acid region of newly synthesized globin messenger ribonucleic acid. J Biol Chem *250*, 3725-3728.

Meyer, K.D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C.E., and Jaffrey, S.R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell *149*, 1635-1646.

Mignon, P., Loverix, S., Steyaert, J., and Geerlings, P. (2005). Influence of the pi-pi interaction on the hydrogen bonding capacity of stacked DNA/RNA bases. Nucleic Acids Res *33*, 1779-1789.

Mittal, S., Aslam, A., Doidge, R., Medica, R., and Winkler, G.S. (2011). The Ccr4a (CNOT6) and Ccr4b (CNOT6L) deadenylase subunits of the human Ccr4-Not complex contribute to the prevention of cell death and senescence. Mol Biol Cell *22*, 748-758.

Mitton-Fry, R.M., DeGregorio, S.J., Wang, J., Steitz, T.A., and Steitz, J.A. (2010). Poly(A) tail recognition by a viral RNA element through assembly of a triple helix. Science *330*, 1244-1247.

Mohanty, B.K., and Kushner, S.R. (2000). Polynucleotide phosphorylase functions both as a 3' right-arrow 5' exonuclease and a poly(A) polymerase in Escherichia coli. Proc Natl Acad Sci U S A *97*, 11966-11971.

Mohanty, B.K., and Kushner, S.R. (2011). Bacterial/archaeal/organellar polyadenylation. Wiley Interdiscip Rev RNA *2*, 256-276.

Mohanty, B.K., Maples, V.F., and Kushner, S.R. (2004). The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in Escherichia coli. Mol Microbiol *54*, 905-920.

Molin, L., and Puisieux, A. (2005). C. elegans homologue of the Caf1 gene, which encodes a subunit of the CCR4-NOT complex, is essential for embryonic and larval development and for meiotic progression. Gene *358*, 73-81.

Molloy, G.R., Sporn, M.B., Kelley, D.E., and Perry, R.P. (1972). Localization of polyadenylic acid sequences in messenger ribonucleic acid of mammalian cells. Biochemistry *11*, 3256-3260.

Monecke, T., Schell, S., Dickmanns, A., and Ficner, R. (2008). Crystal structure of the RRM domain of poly(A)-specific ribonuclease reveals a novel m(7)G-cap-binding mode. J Mol Biol *382*, 827-834.

Moodie, S.L., Mitchell, J.B., and Thornton, J.M. (1996). Protein recognition of adenylate: an example of a fuzzy recognition template. J Mol Biol *263*, 486-500.

Moon, D.H., Segal, M., Boyraz, B., Guinan, E., Hofmann, I., Cahan, P., Tai, A.K., and Agarwal, S. (2015). Poly(A)-specific ribonuclease (PARN) mediates 3'-end maturation of the telomerase RNA component. Nat Genet *47*, 1482-1488.

Moqtaderi, Z., Geisberg, J.V., and Struhl, K. (2018). Extensive Structural Differences of Closely Related 3' mRNA Isoforms: Links to Pab1 Binding and mRNA Stability. Mol Cell *72*, 849-861 e846.

Morgan, M., Much, C., DiGiacomo, M., Azzi, C., Ivanova, I., Vitsios, D.M., Pistolic, J., Collier, P., Moreira, P.N., Benes, V.*, et al.* (2017). mRNA 3' uridylation and poly(A) tail length sculpt the mammalian maternal transcriptome. Nature *548*, 347-351.

Morita, M., Suzuki, T., Nakamura, T., Yokoyama, K., Miyasaka, T., and Yamamoto, T. (2007). Depletion of mammalian CCR4b deadenylase triggers elevation of the p27Kip1 mRNA level and impairs cell growth. Mol Cell Biol *27*, 4980-4990.

Mostafa, D., Takahashi, A., Yanagiya, A., Yamaguchi, T., Abe, T., Kureha, T., Kuba, K., Kanegae, Y., Furuta, Y., Yamamoto, T.*, et al.* (2020). Essential functions of the CNOT7/8 catalytic subunits of the CCR4-NOT complex in mRNA regulation and cell viability. RNA Biol *17*, 403-416.

Moududee, S.A., Jiang, Y., Gilbert, N., Xie, G., Xu, Z., Wu, J., Gong, Q., Tang, Y., and Shi, Y. (2018). Structural and functional characterization of hMEX-3C Ring finger domain as an E3 ubiquitin ligase. Protein Sci *27*, 1661-1669.

Muhlrad, D., Decker, C.J., and Parker, R. (1994). Deadenylation of the unstable mRNA encoded by the yeast MFA2 gene leads to decapping followed by 5'-->3' digestion of the transcript. Genes Dev *8*, 855-866.

Muhlrad, D., Decker, C.J., and Parker, R. (1995). Turnover mechanisms of the stable yeast PGK1 mRNA. Mol Cell Biol *15*, 2145-2156.

Muhlrad, D., and Parker, R. (2005). The yeast EDC1 mRNA undergoes deadenylation-independent decapping stimulated by Not2p, Not4p, and Not5p. EMBO J *24*, 1033-1045.

Mullin, C., Duning, K., Barnekow, A., Richter, D., Kremerskothen, J., and Mohr, E. (2004). Interaction of rat poly(A)-binding protein with poly(A)- and non-poly(A) sequences is preferentially mediated by RNA recognition motifs 3+4. Febs Letters *576*, 437-441.

Munroe, D., and Jacobson, A. (1990). mRNA poly(A) tail, a 3' enhancer of translational initiation. Mol Cell Biol *10*, 3441-3455.

Nagarajan, V.K., Jones, C.I., Newbury, S.F., and Green, P.J. (2013). XRN 5'-->3' exoribonucleases: structure, mechanisms and functions. Biochim Biophys Acta *1829*, 590-603.

Nagata, T., Suzuki, S., Endo, R., Shirouzu, M., Terada, T., Inoue, M., Kigawa, T., Kobayashi, N., Guntert, P., Tanaka, A., *et al.* (2008). The RRM domain of poly(A)-specific ribonuclease has a noncanonical binding site for mRNA cap analog recognition. Nucleic Acids Res *36*, 4754-4767.

Narula, A., Ellis, J., Taliaferro, J.M., and Rissland, O.S. (2019). Coding regions affect mRNA stability in human cells. RNA *25*, 1751-1764.

Nguyen, B., Sokoloski, J., Galletto, R., Elson, E.L., Wold, M.S., and Lohman, T.M. (2014). Diffusion of human replication protein A along single-stranded DNA. J Mol Biol *426*, 3246-3261.

Niedzwiecka, A., Nilsson, P., Worch, R., Stepinski, J., Darzynkiewicz, E., and Virtanen, A. (2016). Molecular recognition of mRNA 5' cap by 3' poly(A)-specific ribonuclease (PARN) differs from interactions known for other cap-binding proteins. Biochim Biophys Acta *1864*, 331-345.

Nietfeld, W., Mentzel, H., and Pieler, T. (1990). The Xenopus laevis poly(A) binding protein is composed of multiple functionally independent RNA binding domains. EMBO J *9*, 3699-3705.

Nilsson, P., Henriksson, N., Niedzwiecka, A., Balatsos, N.A., Kokkoris, K., Eriksson, J., and Virtanen, A. (2007). A multifunctional RNA recognition motif in poly(A)-specific ribonuclease with cap and poly(A) binding properties. J Biol Chem *282*, 32902-32911.

Nishimura, T., Nagamori, I., Nakatani, T., Izumi, N., Tomari, Y., Kuramochi-Miyagawa, S., and Nakano, T. (2018). PNLDC1, mouse pre-piRNA Trimmer, is required for meiotic and post-meiotic male germ cell development. Embo Reports *19*.

Niu, S., Shingle, D.L., Garbarino-Pico, E., Kojima, S., Gilbert, M., and Green, C.B. (2011). The Circadian Deadenylase Nocturnin Is Necessary for Stabilization of the iNOS mRNA in Mice. Plos One *6*.

Nobeli, I., Laskowski, R.A., Valdar, W.S.J., and Thornton, J.M. (2001). On the molecular discrimination between adenine and guanine by proteins. Nucleic Acids Research *29*, 4294-4309.

Norberg, J., and Nilsson, L. (1995). Stacking Free-Energy Profiles for All 16 Natural Ribodinucleoside Monophosphates in Aqueous-Solution. Journal of the American Chemical Society *117*, 10832-10840.

Notredame, C., Higgins, D.G., and Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol *302*, 205-217.

Nudel, U., Soreq, H., and Littauer, U.Z. (1976). Globin mRNA species containing poly(A) segments of different lengths. Their functional stability in Xenopus oocytes. Eur J Biochem *64*, 115-121.

Ohn, T., Chiang, Y.C., Lee, D.J., Yao, G., Zhang, C., and Denis, C.L. (2007). CAF1 plays an important role in mRNA deadenylation separate from its contact to CCR4. Nucleic Acids Res *35*, 3002-3015.

Olivas, W., and Parker, R. (2000). The Puf3 protein is a transcript-specific regulator of mRNA degradation in yeast. EMBO J *19*, 6602-6611.

Otero, L.J., Ashe, M.P., and Sachs, A.B. (1999). The yeast poly(A)-binding protein Pab1p stimulates in vitro poly(A)-dependent and cap-dependent translation by distinct mechanisms. Embo Journal *18*, 3153-3163.

Otwinowski, Z., Schevitz, R.W., Zhang, R.G., Lawson, C.L., Joachimiak, A., Marmorstein, R.Q., Luisi, B.F., and Sigler, P.B. (1988). Crystal structure of trp repressor/operator complex at atomic resolution. Nature *335*, 321-329.

Pagano, J.M., Farley, B.M., Essien, K.I., and Ryder, S.P. (2009). RNA recognition by the embryonic cell fate determinant and germline totipotency factor MEX-3. Proceedings of the National Academy of Sciences of the United States of America *106*, 20252-20257.

Park, J.E., Yi, H., Kim, Y., Chang, H., and Kim, V.N. (2016). Regulation of Poly(A) Tail and Translation during the Somatic Cell Cycle. Molecular Cell *62*, 462-471.

Parker, R., and Song, H. (2004). The enzymes and control of eukaryotic mRNA turnover. Nat Struct Mol Biol *11*, 121-127.

Patrick, S.M., and Turchi, J.J. (2001). Stopped-flow kinetic analysis of replication protein A-binding DNA: damage recognition and affinity for single-stranded DNA reveal differential contributions of k(on) and k(off) rate constants. J Biol Chem *276*, 22630-22637.

Pearce, S.F., Rorbach, J., Van Haute, L., D'Souza, A.R., Rebelo-Guiomar, P., Powell, C.A., Brierley, I., Firth, A.E., and Minczuk, M. (2017). Maturation of selected human mitochondrial tRNAs requires deadenylation. Elife *6*.

Peng, S.S.Y., Chen, C.Y.A., Xu, N.H., and Shyu, A.B. (1998). RNA stabilization by the AU-rich element binding protein, HuR, an ELAV protein. Embo Journal *17*, 3461-3470.

Penman, S., Rosbash, M., and Penman, M. (1970). Messenger and heterogeneous nuclear RNA in HeLa cells: differential inhibition by cordycepin. Proc Natl Acad Sci U S A *67*, 1878-1885.

Pereira, B., Le Borgne, M., Chartier, N.T., Billaud, M., and Almeida, R. (2013). MEX-3 proteins: recent insights on novel post-transcriptional regulators. Trends Biochem Sci *38*, 477-479.

Perry, R.P., Kelley, D.E., and LaTorre, J. (1972). Lack of polyadenylic acid sequences in the messenger RNA of E. coli. Biochem Biophys Res Commun *48*, 1593-1600.

Petit, A.P., Wohlbold, L., Bawankar, P., Huntzinger, E., Schmidt, S., Izaurralde, E., and Weichenrieder, O. (2012). The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4-NOT deadenylase complex. Nucleic Acids Res *40*, 11058-11072.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem *25*, 1605-1612.

Pinamonti, G., Zhao, J.B., Condon, D.E., Paul, F., Noe, F., Turner, D.H., and Bussi, G. (2017). Predicting the Kinetics of RNA Oligonucleotides Using Markov State Models. Journal of Chemical Theory and Computation *13*, 926-934.

Poland, D., Vournakis, J.N., and Scheraga, H.A. (1966). Cooperative interactions in single-strand oligomers of adenylic acid. Biopolymers *4*, 223-235.

Porschke, D. (1973). The dynamics of nucleic-acid single-strand conformation changes. Oligo- and polyriboadenylic acids. Eur J Biochem *39*, 117-126.

Potterton, L., Agirre, J., Ballard, C., Cowtan, K., Dodson, E., Evans, P.R., Jenkins, H.T., Keegan, R., Krissinel, E., Stevenson, K.*, et al.* (2018). CCP4i2: the new graphical user interface to the CCP4 program suite. Acta Crystallogr D Struct Biol *74*, 68-84.

Poulsen, J.B., Andersen, K.R., Kjaer, K.H., Durand, F., Faou, P., Vestergaard, A.L., Talbo, G.H., Hoogenraad, N., Brodersen, D.E., Justesen, J.*, et al.* (2011). Human 2'-phosphodiesterase localizes to the mitochondrial matrix with a putative function in mitochondrial RNA turnover. Nucleic Acids Research *39*, 3754-3770.

Pratt, A.J., and MacRae, I.J. (2009). The RNA-induced silencing complex: a versatile gene-silencing machine. J Biol Chem *284*, 17897-17901.

Preiss, T., Muckenthaler, M., and Hentze, M.W. (1998). Poly(A)-tail-promoted translation in yeast: implications for translational control. RNA *4*, 1321-1331.

Presnyak, V., Alhusaini, N., Chen, Y.H., Martin, S., Morris, N., Kline, N., Olson, S., Weinberg, D., Baker, K.E., Graveley, B.R.*, et al.* (2015). Codon optimality is a major determinant of mRNA stability. Cell *160*, 1111-1124.

Prevot, D., Darlix, J.L., and Ohlmann, T. (2003). Conducting the initiation of protein synthesis: the role of eIF4G. Biology of the Cell *95*, 141-156.

Proudfoot, N.J. (2011). Ending the message: poly(A) signals then and now. Genes Dev *25*, 1770-1782.

Proweller, A., and Butler, S. (1994). Efficient translation of poly(A)-deficient mRNAs in Saccharomyces cerevisiae. Genes Dev *8*, 2629-2640.

Quesada, V., Diaz-Perales, A., Gutierrez-Fernandez, A., Garabaya, C., Cal, S., and Lopez-Otin, C. (2004). Cloning and enzymatic analysis of 22 novel human ubiquitin-specific proteases. Biochemical and Biophysical Research Communications *314*, 54-62.

Rabani, M., Pieper, L., Chew, G.L., and Schier, A.F. (2017). A Massively Parallel Reporter Assay of 3' UTR Sequences Identifies In Vivo Rules for mRNA Degradation. Mol Cell *68*, 1083-1094 e1085.

Raisch, T., Chang, C.T., Levdansky, Y., Muthukumar, S., Raunser, S., and Valkov, E. (2019). Reconstitution of recombinant human CCR4-NOT reveals molecular insights into regulated deadenylation. Nat Commun *10*, 3173.

Reuter, J.S., and Mathews, D.H. (2010). RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinformatics *11*, 129.

Reverdatto, S.V., Dutko, J.A., Chekanova, J.A., Hamilton, D.A., and Belostotsky, D.A. (2004). mRNA deadenylation by PARN is essential for embryogenesis in higher plants. RNA *10*, 1200-1214.

Riback, J.A., Katanski, C.D., Kear-Scott, J.L., Pilipenko, E.V., Rojek, A.E., Sosnick, T.R., and Drummond, D.A. (2017). Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response. Cell *168*, 1028-+.

Rich, A., Davies, D.R., Crick, F.H., and Watson, J.D. (1961). The molecular structure of polyadenylic acid. J Mol Biol *3*, 71-86.

Rissland, O.S., Mikulasova, A., and Norbury, C.J. (2007). Efficient RNA polyuridylation by noncanonical poly(A) polymerases. Mol Cell Biol *27*, 3612-3624.

Rissland, O.S., and Norbury, C.J. (2009). Decapping is preceded by 3' uridylation in a novel pathway of bulk mRNA turnover. Nat Struct Mol Biol *16*, 616-623.

Rivera, C.I., and Lloyd, R.E. (2008). Modulation of enteroviral proteinase cleavage of poly(A)-binding protein (PABP) by conformation and PABP-associated factors. Virology *375*, 59-72.

Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res *42*, W320-324.

Rohou, A., and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J Struct Biol *192*, 216-221.

Roque, S., Cerciat, M., Gaugue, I., Mora, L., Floch, A.G., de Zamaroczy, M., Heurgue-Hamard, V., and Kervestin, S. (2015). Interaction between the poly(A)-binding protein Pab1 and the eukaryotic release factor eRF3 regulates translation termination but not mRNA decay in Saccharomyces cerevisiae. Rna-a Publication of the Rna Society *21*, 124-134.

Rorbach, J., Bobrowicz, A., Pearce, S., and Minczuk, M. (2014). Polyadenylation in bacteria and organelles. Methods Mol Biol *1125*, 211-227.

Rorbach, J., Nicholls, T.J.J., and Minczuk, M. (2011). PDE12 removes mitochondrial RNA poly(A) tails and controls translation in human mitochondria. Nucleic Acids Research *39*, 7750-7763.

Rott, R., Zipor, G., Portnoy, V., Liveanu, V., and Schuster, G. (2003). RNA polyadenylation and degradation in cyanobacteria are similar to the chloroplast but different from Escherichia coli. J Biol Chem *278*, 15771-15777.

Russo, C.J., and Passmore, L.A. (2014). Electron microscopy: Ultrastable gold substrates for electron cryomicroscopy. Science *346*, 1377-1380.

Sachs, A. (1990). The role of poly(A) in the translation and stability of mRNA. Curr Opin Cell Biol *2*, 1092-1098.

Sachs, A.B., Bond, M.W., and Kornberg, R.D. (1986). A single gene from yeast for both nuclear and cytoplasmic polyadenylate-binding proteins: domain structure and expression. Cell *45*, 827-835.

Sachs, A.B., and Davis, R.W. (1989a). The poly(A) binding protein is required for poly(A) shortening and 60S ribosomal subunit-dependent translation initiation. Cell *58*, 857-867.

Sachs, A.B., and Davis, R.W. (1989b). The Poly(a) Binding-Protein Is Required for Poly(a) Shortening and 60s Ribosomal Subunit-Dependent Translation Initiation. Cell *58*, 857-867.

Sachs, A.B., Davis, R.W., and Kornberg, R.D. (1987). A Single Domain of Yeast Poly(a)-Binding Protein Is Necessary and Sufficient for Rna-Binding and Cell Viability. Molecular and Cellular Biology *7*, 3268-3276.

Sachs, A.B., and Deardorff, J.A. (1992). Translation Initiation Requires the Pab-Dependent Poly(a) Ribonuclease in Yeast. Cell *70*, 961-973.

Saenger, W., Riecke, J., and Suck, D. (1975). A structural model for the polyadenylic acid single helix. J Mol Biol *93*, 529-534.

Safaee, N., Kozlov, G., Noronha, A.M., Xie, J., Wilds, C.J., and Gehring, K. (2012a). Interdomain allostery promotes assembly of the poly(A) mRNA complex with PABP and eIF4G. Mol Cell *48*, 375-386.

Safaee, N., Kozlov, G., Noronha, A.M., Xie, J.W., Wilds, C.J., and Gehring, K. (2012b). Interdomain Allostery Promotes Assembly of the Poly(A) mRNA Complex with PABP and eIF4G. Molecular Cell *48*, 375-386.

Safaee, N., Noronha, A.M., Rodionov, D., Kozlov, G., Wilds, C.J., Sheldrick, G.M., and Gehring, K. (2013). Structure of the parallel duplex of poly(A) RNA: evaluation of a 50 year-old prediction. Angew Chem Int Ed Engl *52*, 10370-10373.

Sakai, A., Chibazakura, T., Shimizu, Y., and Hishinuma, F. (1992). Molecular analysis of POP2 gene, a gene required for glucose-derepression of gene expression in Saccharomyces cerevisiae. Nucleic Acids Res *20*, 6227-6233.

Salles, F.J., Richards, W.G., and Strickland, S. (1999). Assaying the polyadenylation state of mRNAs. Methods *17*, 38-45.

Sandler, H., Kreth, J., Timmers, H.T., and Stoecklin, G. (2011). Not1 mediates recruitment of the deadenylase Caf1 to mRNAs targeted for degradation by tristetraprolin. Nucleic Acids Res *39*, 4373-4386.

Sassa, A., Beard, W.A., Shock, D.D., and Wilson, S.H. (2013). Steady-state, pre-steady-state, and single-turnover kinetic measurement for DNA glycosylase activity. J Vis Exp, e50695.

Sawazaki, R., Imai, S., Yokogawa, M., Hosoda, N., Hoshino, S., Mio, M., Mio, K., Shimada, I., and Osawa, M. (2018). Characterization of the multimeric structure of poly(A)-binding protein on a poly(A) tail. Scientific Reports *8*.

Schafer, I.B., Rode, M., Bonneau, F., Schussler, S., and Conti, E. (2014). The structure of the Pan2-Pan3 core complex reveals cross-talk between deadenylase and pseudokinase. Nat Struct Mol Biol *21*, 591-598.

Schafer, I.B., Yamashita, M., Schuller, J.M., Schussler, S., Reichelt, P., Strauss, M., and Conti, E. (2019). Molecular Basis for poly(A) RNP Architecture and Recognition by the Pan2-Pan3 Deadenylase. Cell *177*, 1619-1631 e1621.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B.*, et al.* (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods *9*, 676-682.

Schrodinger, L. (2015). The PyMOL Molecular Graphics System, Version 2.0

Searfoss, A.M., and Wickner, R.B. (2000). 3' poly(A) is dispensable for translation. Proc Natl Acad Sci U S A *97*, 9133-9137.

Sears, R.M., May, D.G., and Roux, K.J. (2019). BioID as a Tool for Protein-Proximity Labeling in Living Cells. Enzyme-Mediated Ligation Methods *2012*, 299-313.
Sendinc, E., Valle-Garcia, D., Dhall, A., Chen, H., Henriques, T., Navarrete-Perea, J., Sheng, W., Gygi, S.P., Adelman, K., and Shi, Y. (2019). PCIF1 Catalyzes m6Am mRNA Methylation to Regulate Gene Expression. Mol Cell *75*, 620-630 e629.
Seol, Y., Skinner, G.M., Visscher, K., Buhot, A., and Halperin, A. (2007). Stretching of homopolymeric RNA reveals single-stranded helices and base-stacking. Phys Rev Lett *98*, 158103.
Shaw, G., and Kamen, R. (1986). A Conserved Au Sequence from the 3' Untranslated Region of Gm-Csf Messenger-Rna Mediates Selective Messenger-Rna Degradation. Cell *46*, 659-667.
Sheiness, D., and Darnell, J.E. (1973). Polyadenylic acid segment in mRNA becomes shorter with age. Nat New Biol *241*, 265-268.
Sheiness, D., Puckett, L., and Darnell, J.E. (1975). Possible relationship of poly(A) shortening to mRNA turnover. Proc Natl Acad Sci U S A *72*, 1077-1081.
Shi, H., Wei, J., and He, C. (2019). Where, When, and How: Context-Dependent Functions of RNA Methylation Writers, Readers, and Erasers. Mol Cell *74*, 640-650.
Shyu, A.B., Belasco, J.G., and Greenberg, M.E. (1991). Two distinct destabilizing elements in the c-fos message trigger deadenylation as a first step in rapid mRNA decay. Genes Dev *5*, 221-231.
Siddiqui, N., Mangus, D.A., Chang, T.C., Palermino, J.M., Shyu, A.B., and Gehring, K. (2007). Poly(A) nuclease interacts with the C-terminal domain of polyadenylate-binding protein domain from poly(A)-binding protein. J Biol Chem *282*, 25067-25075.
Simon, E., and Seraphin, B. (2007). A specific role for the C-terminal region of the Poly(A)-binding protein in mRNA decay. Nucleic Acids Research *35*, 6017-6028.
Sladic, R.T., Lagnado, C.A., Bagley, C.J., and Goodall, G.J. (2004). Human PABP binds AU-rich RNA via RNA-binding domains 3 and 4. European Journal of Biochemistry *271*, 450-457.
Slobodin, B., Bahat, A., Sehrawat, U., Becker-Herman, S., Zuckerman, B., Weiss, A.N., Han, R., Elkon, R., Agami, R., Ulitsky, I.*, et al.* (2020). Transcription Dynamics Regulate Poly(A) Tails and Expression of the RNA Degradation Machinery to Balance mRNA Levels. Mol Cell.
Smith, R.W., and Gray, N.K. (2010). Poly(A)-binding protein (PABP): a common viral target. Biochem J *426*, 1-12.
Sohrabi-Jahromi, S., Hofmann, K.B., Boltendahl, A., Roth, C., Gressel, S., Baejen, C., Soeding, J., and Cramer, P. (2019). Transcriptome maps of general eukaryotic RNA degradation factors. Elife *8*.
Son, A., Park, J.E., and Kim, V.N. (2018). PARN and TOE1 Constitute a 3 ' End Maturation Module for Nuclear Non-coding RNAs. Cell Reports *23*, 888-898.
Song, M.G., and Kiledjian, M. (2007). 3' Terminal oligo U-tract-mediated stimulation of decapping. RNA *13*, 2356-2365.
Soucek, S., Corbett, A.H., and Fasken, M.B. (2012). The long and the short of it: the role of the zinc finger polyadenosine RNA binding protein, Nab2, in control of poly(A) tail length. Biochim Biophys Acta *1819*, 546-554.
Sponer, J., Leszczynski, J., and Hobza, P. (2001). Electronic properties, hydrogen bonding, stacking, and cation binding of DNA and RNA bases. Biopolymers *61*, 3-31.
Steiner, R.F., and Beers, R.F. (1957). Polynucleotides .2. Physical Properties of Solutions of Some Polynucleotides. Biochimica Et Biophysica Acta *26*, 336-348.
Steitz, T.A., and Steitz, J.A. (1993). A general two-metal-ion mechanism for catalytic RNA. Proc Natl Acad Sci U S A *90*, 6498-6502.

Stephenson, W., Asare-Okai, P.N., Chen, A.A., Keller, S., Santiago, R., Tenenbaum, S.A., Garcia, A.E., Fabris, D., and Li, P.T. (2013). The essential role of stacking adenines in a two-base-pair RNA kissing complex. J Am Chem Soc *135*, 5602-5611.

Stoeckle, M.Y., and Guan, L. (1993). High-resolution analysis of gro alpha mRNA poly(A) shortening: regulation by interleukin-1 beta. Nucleic Acids Res *21*, 1613-1617.

Stowell, J.A.W., Wagstaff, J.L., Hill, C.H., Yu, M., McLaughlin, S.H., Freund, S.M.V., and Passmore, L.A. (2018). A low-complexity region in the YTH domain protein Mmi1 enhances RNA binding. J Biol Chem *293*, 9210-9222.

Stowell, J.A.W., Webster, M.W., Kogel, A., Wolf, J., Shelley, K.L., and Passmore, L.A. (2016a). Reconstitution of Targeted Deadenylation by the Ccr4-Not Complex and the YTH Domain Protein Mmi1. Cell Rep *17*, 1978-1989.

Stowell, J.A.W., Webster, M.W., Kogel, A., Wolf, J., Shelley, K.L., and Passmore, L.A. (2016b). Reconstitution of Targeted Deadenylation by the Ccr4-Not Complex and the YTH Domain Protein Mmi1. Cell Reports *17*, 1978-1989.

Strzelecka, D., Smietanski, M., Sikorski, P.J., Warminski, M., Kowalska, J., and Jemielity, J. (2020). Functional and LC-MS/MS analysis of in vitro transcribed mRNAs carrying phosphorothioate or boranophosphate moieties reveal polyA tail modifications that prevent deadenylation without compromising protein expression. bioRxiv, 2020.2007.2002.184598.

Subtelny, A.O., Eichhorn, S.W., Chen, G.R., Sive, H., and Bartel, D.P. (2014). Poly(A)-tail profiling reveals an embryonic switch in translational control. Nature *508*, 66-+.

Suck, D., Manor, P.C., and Saenger, W. (1976). The structure of a trinucleoside diphosphate: adenylyl-(3',5')-adenylyl-(3',5')-adenosine hexahydrate. Acta Crystallographica *32*, 1727-1737.

Sun, M., Schwalb, B., Pirkl, N., Maier, K.C., Schenk, A., Failmezger, H., Tresch, A., and Cramer, P. (2013). Global Analysis of Eukaryotic mRNA Degradation Reveals Xrn1-Dependent Buffering of Transcript Levels. Molecular Cell *52*, 52-62.

Sun, M., Schwalb, B., Schulz, D., Pirkl, N., Etzold, S., Lariviere, L., Maier, K.C., Seizl, M., Tresch, A., and Cramer, P. (2012). Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. Genome Research *22*, 1350-1359.

Sundaramoorthy, E., Leonard, M., Mak, R., Liao, J., Fulzele, A., and Bennett, E.J. (2017). ZNF598 and RACK1 Regulate Mammalian Ribosome-Associated Quality Control Function by Mediating Regulatory 40S Ribosomal Ubiquitylation. Mol Cell *65*, 751-760 e754.

Tajaddod, M., Jantsch, M.F., and Licht, K. (2016). The dynamic epitranscriptome: A to I editing modulates genetic information. Chromosoma *125*, 51-63.

Takaya, T., Su, C., de La Harpe, K., Crespo-Hernandez, C.E., and Kohler, B. (2008). UV excitation of single DNA and RNA strands produces high yields of exciplex states between two stacked bases. Proc Natl Acad Sci U S A *105*, 10285-10290.

Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: an extensible image processing suite for electron microscopy. J Struct Biol *157*, 38-46.

Tang, W., Tu, S., Lee, H.C., Weng, Z., and Mello, C.C. (2016). The RNase PARN-1 Trims piRNA 3' Ends to Promote Transcriptome Surveillance in C. elegans. Cell *164*, 974-984.

Tarun, S.Z., Jr., and Sachs, A.B. (1996). Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. EMBO J *15*, 7168-7177.

Tarun, S.Z., Wells, S.E., Deardorff, J.A., and Sachs, A.B. (1997). Translation initiation factor eIF4G mediates in vitro poly(A) tail-dependent translation. Proceedings of the National Academy of Sciences of the United States of America *94*, 9046-9051.

Tesina, P., Lessen, L.N., Buschauer, R., Cheng, J., Wu, C.C., Berninghausen, O., Buskirk, A.R., Becker, T., Beckmann, R., and Green, R. (2019). Molecular mechanism of translational stalling by inhibitory codon combinations and poly(A) tracts. EMBO J, e103365.

Thore, S., Mauxion, F., Seraphin, B., and Suck, D. (2003). X-ray structure and activity of the yeast Pop2 protein: a nuclease subunit of the mRNA deadenylase complex. EMBO Rep *4*, 1150-1155.

Tian, B., and Manley, J.L. (2017). Alternative polyadenylation of mRNA precursors. Nat Rev Mol Cell Biol *18*, 18-30.

Tran, H., Schilling, M., Wirbelauer, C., Hess, D., and Nagamine, Y. (2004). Facilitation of mRNA deadenylation and decay by the exosome-bound, DExH protein RHAU. Mol Cell *13*, 101-111.

Trinkle-Mulcahy, L. (2019). Recent advances in proximity-based labeling methods for interactome mapping. F1000Res *8*.

Ts'o, P.O., Kondo, N.S., Schweizer, M.P., and Hollis, D.P. (1969). Studies of the conformation and interaction in dinucleoside mono- and diphosphates by proton magnetic resonance. Biochemistry *8*, 997-1029.

Tuck, A.C., and Tollervey, D. (2013). A transcriptome-wide atlas of RNP composition reveals diverse classes of mRNAs and lncRNAs. Cell *154*, 996-1009.

Tucker, M., Staples, R.R., Valencia-Sanchez, M.A., Muhlrad, D., and Parker, R. (2002). Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in Saccharomyces cerevisiae. EMBO J *21*, 1427-1436.

Tucker, M., Valencia-Sanchez, M.A., Staples, R.R., Chen, J.J., Denis, C.L., and Parker, R. (2001). The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in Saccharomyces cerevisiae. Cell *104*, 377-386.

Tummala, H., Walne, A., Collopy, L., Cardoso, S., de la Fuente, J., Lawson, S., Powell, J., Cooper, N., Foster, A., Mohammed, S.*, et al.* (2015). Poly(A)-specific ribonuclease deficiency impacts telomere biology and causes dyskeratosis congenita. J Clin Invest *125*, 2151-2160.

Uchida, N., Hoshino, S., and Katada, T. (2004). Identification of a human cytoplasmic poly(A) nuclease complex stimulated by poly(A)-binding protein. Journal of Biological Chemistry *279*, 1383-1391.

Van Holde, K.E., Brahms, J., and Michelson, A.M. (1965). Base interactions of nucleotide polymers in aqueous solution. J Mol Biol *12*, 726-739.

Varnum, S.M., and Wormington, W.M. (1990). Deadenylation of maternal mRNAs during Xenopus oocyte maturation does not require specific cis-sequences: a default mechanism for translational control. Genes Dev *4*, 2278-2286.

Vicens, Q., Kieft, J.S., and Rissland, O.S. (2018). Revisiting the Closed-Loop Model and the Nature of mRNA 5'-3' Communication. Mol Cell *72*, 805-812.

Viegas, I.J., de Macedo, J.P., De Niz, M., Rodrigues, J.A., Aresta-Branco, F., Jaffrey, S.R., and Figueiredo, L.M. (2020). N6-methyladenosine in poly(A) tails stabilize VSG transcripts. bioRxiv, 2020.2001.2030.925776.

Viswanathan, P., Chen, J., Chiang, Y.C., and Denis, C.L. (2003). Identification of multiple RNA features that influence CCR4 deadenylation activity. J Biol Chem *278*, 14949-14955.

Viswanathan, P., Ohn, T., Chiang, Y.C., Chen, J., and Denis, C.L. (2004). Mouse CAF1 can function as a processive deadenylase/3'-5'-exonuclease in vitro but in yeast the deadenylase function of CAF1 is not required for mRNA poly(A) removal. J Biol Chem *279*, 23988-23995.

Wagner, E., Clement, S.L., and Lykke-Andersen, J. (2007). An unconventional human Ccr4-Caf1 deadenylase complex in nuclear cajal bodies. Mol Cell Biol *27*, 1686-1695.

Wagner, T., Merino, F., Stabrin, M., Moriya, T., Antoni, C., Apelbaum, A., Hagel, P., Sitsel, O., Raisch, T., Prumbaum, D.*, et al.* (2019). SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. Commun Biol *2*, 218.

Wakiyama, M., Takimoto, K., Ohara, O., and Yokoyama, S. (2007). Let-7 microRNA-mediated mRNA deadenylation and translational repression in a mammalian cell-free system. Genes & Development *21*, 1857-1862.

Wang, C., Zhu, Y., Bao, H., Jiang, Y., Xu, C., Wu, J., and Shi, Y. (2016). A novel RNA-binding mode of the YTH domain reveals the mechanism for recognition of determinant of selective removal by Mmi1. Nucleic Acids Res *44*, 969-982.

Wang, C.W., Schmich, F., Srivatsa, S., Weidner, J., Beerenwinkel, N., and Spang, A. (2018a). Context-dependent deposition and regulation of mRNAs in P-bodies. Elife *7*.

Wang, H., Morita, M., Yang, X., Suzuki, T., Yang, W., Wang, J., Ito, K., Wang, Q., Zhao, C., Bartlam, M.*, et al.* (2010). Crystal structure of the human CNOT6L nuclease domain reveals strict poly(A) substrate specificity. EMBO J *29*, 2566-2576.

Wang, M.Y., Cutler, M., Karimpour, I., and Kleene, K.C. (1992). Nucleotide sequence of a mouse testis poly(A) binding protein cDNA. Nucleic Acids Res *20*, 3519.

Wang, R., Nambiar, R., Zheng, D., and Tian, B. (2018b). PolyA_DB 3 catalogs cleavage and polyadenylation sites identified by deep sequencing in multiple genomes. Nucleic Acids Res *46*, D315-D319.

Wang, R.J., Nambiar, R., Zheng, D.H., and Tian, B. (2018c). PolyA_DB 3 catalogs cleavage and polyadenylation sites identified by deep sequencing in multiple genomes. Nucleic Acids Research *46*, D315-D319.

Wang, X., McLachlan, J., Zamore, P.D., and Hall, T.M. (2002a). Modular recognition of RNA by a human pumilio-homology domain. Cell *110*, 501-512.

Wang, Y., Liu, C.L., Storey, J.D., Tibshirani, R.J., Herschlag, D., and Brown, P.O. (2002b). Precision and functional specificity in mRNA decay. Proc Natl Acad Sci U S A *99*, 5860-5865.

Warshaw, M.M., and Tinoco, I., Jr. (1965). Absorption and optical rotatory dispersion of six dinucleoside phosphates. J Mol Biol *13*, 54-64.

Wasmuth, E.V., Januszyk, K., and Lima, C.D. (2014). Structure of an Rrp6-RNA exosome complex bound to poly(A) RNA. Nature *511*, 435-439.

Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M., and Barton, G.J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189-1191.

Webster, M.W., Chen, Y.H., Stowell, J.A.W., Alhusaini, N., Sweet, T., Graveley, B.R., Coller, J., and Passmore, L.A. (2018). mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases. Mol Cell *70*, 1089-1100 e1088.

Webster, M.W., Stowell, J.A.W., and Passmore, L.A. (2019). RNA-binding proteins distinguish between similar sequence motifs to promote targeted deadenylation by Ccr4-Not. Elife *8*.

Webster, M.W., Stowell, J.A.W., Tang, T.T.L., and Passmore, L.A. (2017). Analysis of mRNA deadenylation by multi-protein complexes. Methods *126*, 95-104.

Weir, J.R., Bonneau, F., Hentschel, J., and Conti, E. (2010). Structural analysis reveals the characteristic features of Mtr4, a DExH helicase involved in nuclear RNA processing and surveillance. Proc Natl Acad Sci U S A *107*, 12139-12144.

Weissmann, F., Petzold, G., VanderLinden, R., Huis In 't Veld, P.J., Brown, N.G., Lampert, F., Westermann, S., Stark, H., Schulman, B.A., and Peters, J.M. (2016). biGBac enables rapid gene assembly for the expression of large multisubunit protein complexes. Proc Natl Acad Sci U S A *113*, E2564-2569.

Wells, S.E., Hillner, P.E., Vale, R.D., and Sachs, A.B. (1998). Circularization of mRNA by eukaryotic translation initiation factors. Mol Cell *2*, 135-140.

Wilson, M.C., Sawicki, S.G., White, P.A., and Darnell, J.E., Jr. (1978). A correlation between the rate of poly(A) shortening and half-life of messenger RNA in adenovirus transformed cells. J Mol Biol *126*, 23-36.

Wilson, T., and Treisman, R. (1988). Removal of poly(A) and consequent degradation of c-fos mRNA facilitated by 3' AU-rich sequences. Nature *336*, 396-399.

Wilusz, C.J., Gao, M., Jones, C.L., Wilusz, J., and Peltz, S.W. (2001). Poly(A)-binding proteins regulate both mRNA deadenylation and decapping in yeast cytoplasmic extracts. Rna *7*, 1416-1424.

Winter, G. (2010). xia2: an expert system for macromolecular crystallography data reduction. Journal of Applied Crystallography *43*, 186-190.

Wolf, J., Valkov, E., Allen, M.D., Meineke, B., Gordiyenko, Y., McLaughlin, S.H., Olsen, T.M., Robinson, C.V., Bycroft, M., Stewart, M.*, et al.* (2014). Structural basis for Pan3 binding to Pan2 and its function in mRNA recruitment and deadenylation. EMBO J *33*, 1514-1526.

Woo, S.L., Rosen, J.M., Liarakos, C.D., Choi, Y.C., Busch, H., Means, A.R., and O'Malley (1975). Physical and chemical characterization of purified ovalbumin messenger RNA. J Biol Chem *250*, 7027-7039.

Wood, E.R., Bledsoe, R., Chai, J., Daka, P., Deng, H.F., Ding, Y., Harris-Gurley, S., Kryn, L.H., Nartey, E., Nichols, J.*, et al.* (2015). The Role of Phosphodiesterase 12 (PDE12) as a Negative Regulator of the Innate Immune Response and the Discovery of Antiviral Inhibitors. Journal of Biological Chemistry *290*, 19681-19696.

Wormington, M., Searfoss, A.M., and Hurney, C.A. (1996). Overexpression of poly(A) binding protein prevents maturation-specific deadenylation and translational inactivation in Xenopus oocytes. EMBO J *15*, 900-909.

Wu, B.X., Xu, J.H., Su, S.C., Liu, H.H., Gan, J.H., and Ma, J.B. (2017). Structural insights into the specific recognition of DSR by the YTH domain containing protein Mmi1. Biochemical and Biophysical Research Communications *491*, 310-316.

Wu, L.G., Fan, J.H., and Belasco, J.G. (2006). MicroRNAs direct rapid deadenylation of mRNA. Proceedings of the National Academy of Sciences of the United States of America *103*, 4034-4039.

Wu, M., Nilsson, P., Henriksson, N., Niedzwiecka, A., Lim, M.K., Cheng, Z., Kokkoris, K., Virtanen, A., and Song, H. (2009). Structural basis of m(7)GpppG binding to poly(A)-specific ribonuclease. Structure *17*, 276-286.

Wu, M.S., Reuter, M., Lilie, H., Liu, Y.Y., Wahle, E., and Song, H.W. (2005). Structural insight into poly(A) binding and catalytic mechanism of human PARN. Embo Journal *24*, 4082-4093.

Xu, C., Liu, K., Ahmed, H., Loppnau, P., Schapira, M., and Min, J. (2015). Structural Basis for the Discriminative Recognition of N6-Methyladenosine RNA by the Human YT521-B Homology Domain Family of Proteins. J Biol Chem *290*, 24902-24913.

Xu, C., Wang, X., Liu, K., Roundtree, I.A., Tempel, W., Li, Y., Lu, Z., He, C., and Min, J. (2014). Structural basis for selective binding of m6A RNA by the YTHDC1 YTH domain. Nat Chem Biol *10*, 927-929.

Xue, M., Chen, L.Y., Wang, W.J., Su, T.T., Shi, L.H., Wang, L., Zhang, W., Si, J.M., Wang, L.J., and Chen, S.J. (2018). HOTAIR induces the ubiquitination of Runx3 by interacting with Mex3b and enhances the invasion of gastric cancer cells. Gastric Cancer *21*, 756-764.

Yamashita, A., Chang, T.C., Yamashita, Y., Zhu, W., Zhong, Z., Chen, C.Y., and Shyu, A.B. (2005). Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. Nat Struct Mol Biol *12*, 1054-1063.

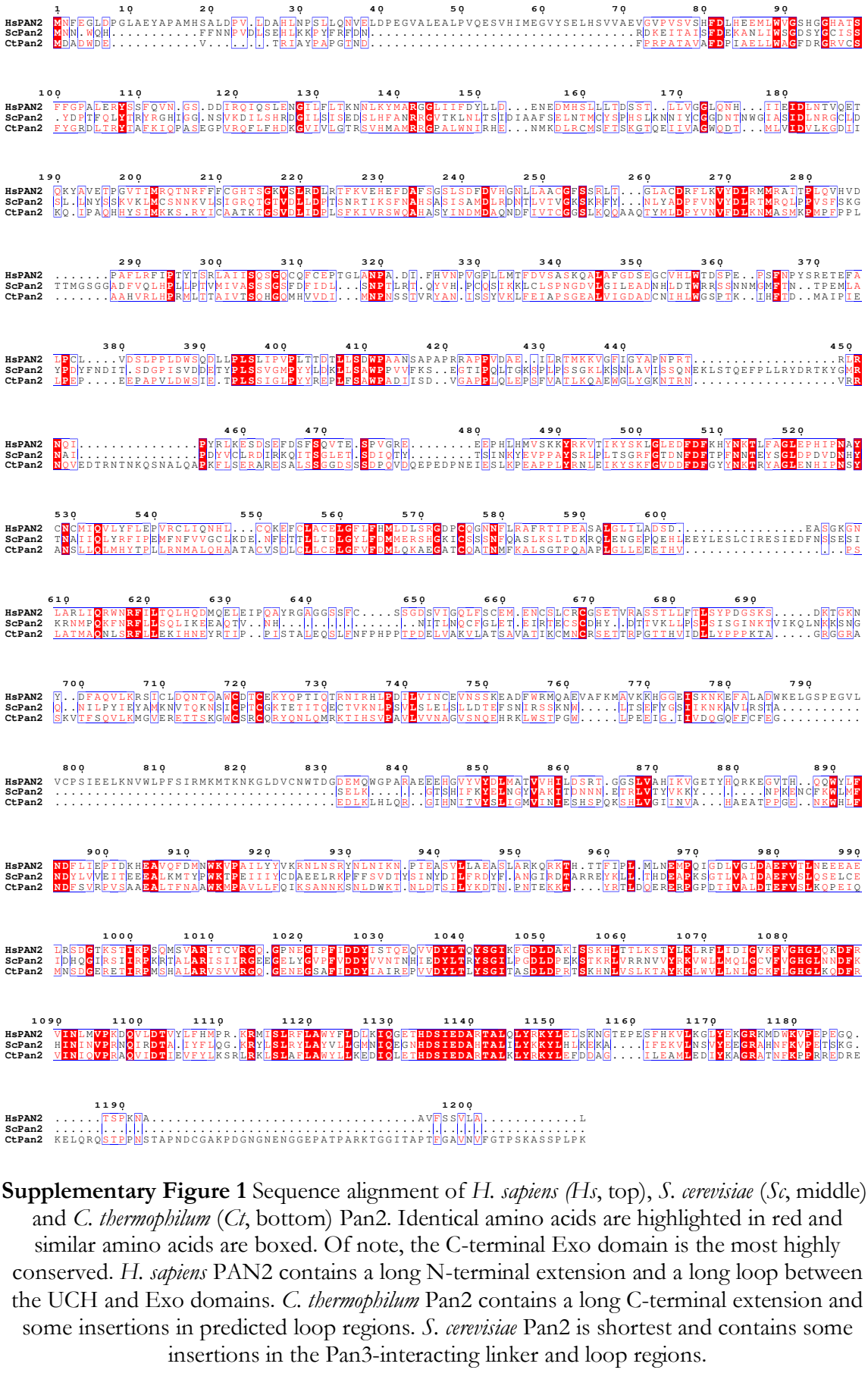
Yandek, L.E., Harris, M., Merrick, W., OhioLINK Electronic Theses and Dissertations Center., Case Western Reserve University. Biochemistry, and Case Western Reserve

University (2013). Multiple substrate kinetics of ribonuclease P : relative rate constant determination through internal competition.
Yang, L., Wang, C., Li, F., Zhang, J., Nayab, A., Wu, J., Shi, Y., and Gong, Q. (2017). The human RNA-binding protein and E3 ligase MEX-3C binds the MEX-3-recognition element (MRE) motif with high affinity. J Biol Chem *292*, 16221-16234.
Yao, G., Chiang, Y.C., Zhang, C.X., Lee, D.J., Laue, T.M., and Denis, C.L. (2007). PAB1 self-association precludes its binding to Poly(A), thereby accelerating CCR4 deadenylation in vivo. Molecular and Cellular Biology *27*, 6243-6253.
Yehudai-Resheff, S., Hirsh, M., and Schuster, G. (2001). Polynucleotide phosphorylase functions as both an exonuclease and a poly(A) polymerase in spinach chloroplasts. Mol Cell Biol *21*, 5408-5416.
Yehudai-Resheff, S., Portnoy, V., Yogev, S., Adir, N., and Schuster, G. (2003). Domain analysis of the chloroplast polynucleotide phosphorylase reveals discrete functions in RNA degradation, polyadenylation, and sequence homology with exosome proteins. Plant Cell *15*, 2003-2019.
Yi, H., Park, J., Ha, M., Lim, J., Chang, H., and Kim, V.N. (2018). PABP Cooperates with the CCR4-NOT Complex to Promote mRNA Deadenylation and Block Precocious Decay. Mol Cell *70*, 1081-1088 e1085.
Yoda, M., Cifuentes, D., Izumi, N., Sakaguchi, Y., Suzuki, T., Giraldez, A.J., and Tomari, Y. (2013). Poly(A)-specific ribonuclease mediates 3'-end trimming of Argonaute2-cleaved precursor microRNAs. Cell Rep *5*, 715-726.
Youn, J.Y., Dunham, W.H., Hong, S.J., Knight, J.D.R., Bashkurov, M., Chen, G.I., Bagci, H., Rathod, B., MacLeod, G., Eng, S.W.M.*, et al.* (2018). High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies. Mol Cell *69*, 517-532 e511.
Zhang, B., Morace, G., Gauss-Muller, V., and Kusov, Y. (2007). Poly(A) binding protein, C-terminally truncated by the hepatitis A virus proteinase 3C, inhibits viral translation. Nucleic Acids Res *35*, 5975-5984.
Zhang, C.X., Lee, D.J., Chiang, Y.C., Richardson, R., Park, S., Wang, X., Laue, T.M., and Denis, C.L. (2013). The RRM1 domain of the poly(A)-binding protein from Saccharomyces cerevisiae is critical to control of mRNA deadenylation. Molecular Genetics and Genomics *288*, 401-412.
Zhang, K. (2016). Gctf: Real-time CTF determination and correction. J Struct Biol *193*, 1-12.
Zhang, Y., Guo, R., Cui, Y.Q., Zhu, Z.P., Zhang, Y.W., Wu, H., Zheng, B., Yue, Q.L., Bai, S., Zeng, W.T.*, et al.* (2017). An essential role for PNLDC1 in piRNA 3 ' end trimming and male fertility in mice. Cell Research *27*, 1392-1396.
Zhao, J., Hyman, L., and Moore, C. (1999). Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. Microbiol Mol Biol Rev *63*, 405-445.
Zheng, D.H., Ezzeddine, N., Chen, C.Y.A., Zhu, W.M., He, X.W., and Shyu, A.B. (2008). Deadenylation is prerequisite for P-body formation and mRNA decay in mammalian cells. Journal of Cell Biology *182*, 89-101.
Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods *14*, 331-332.
Zinder, J.C., Wasmuth, E.V., and Lima, C.D. (2016). Nuclear RNA Exosome at 3.1 A Reveals Substrate Specificities, RNA Paths, and Allosteric Inhibition of Rrp44/Dis3. Mol Cell *64*, 734-745.
Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife *7*.

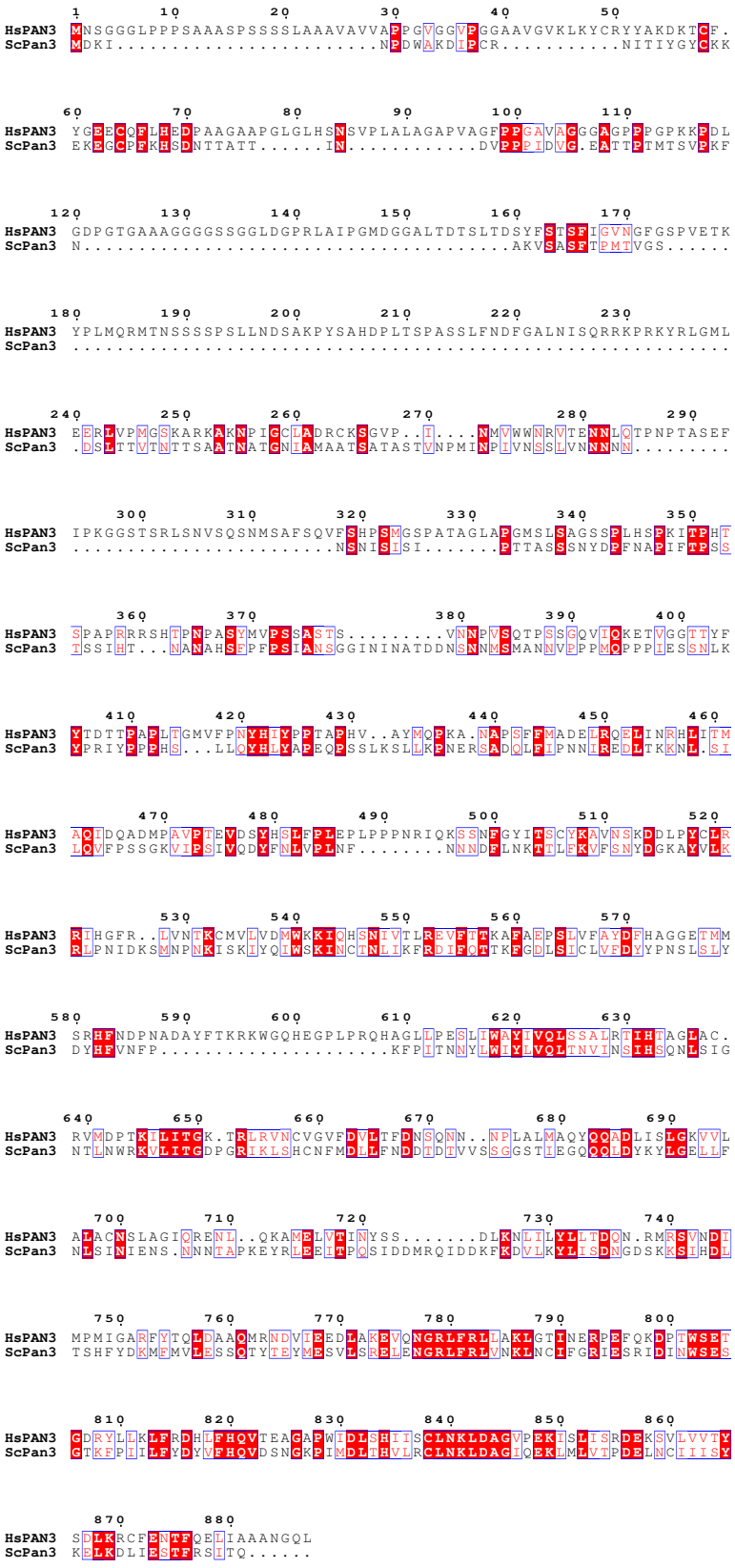
Zuber, H., Scheer, H., Ferrier, E., Sement, F.M., Mercier, P., Stupfler, B., and Gagliardi, D. (2016). Uridylation and PABP Cooperate to Repair mRNA Deadenylated Ends in Arabidopsis. Cell Rep *14*, 2707-2717.

Zuo, Y., and Deutscher, M.P. (2001). Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. Nucleic Acids Res *29*, 1017-1026.

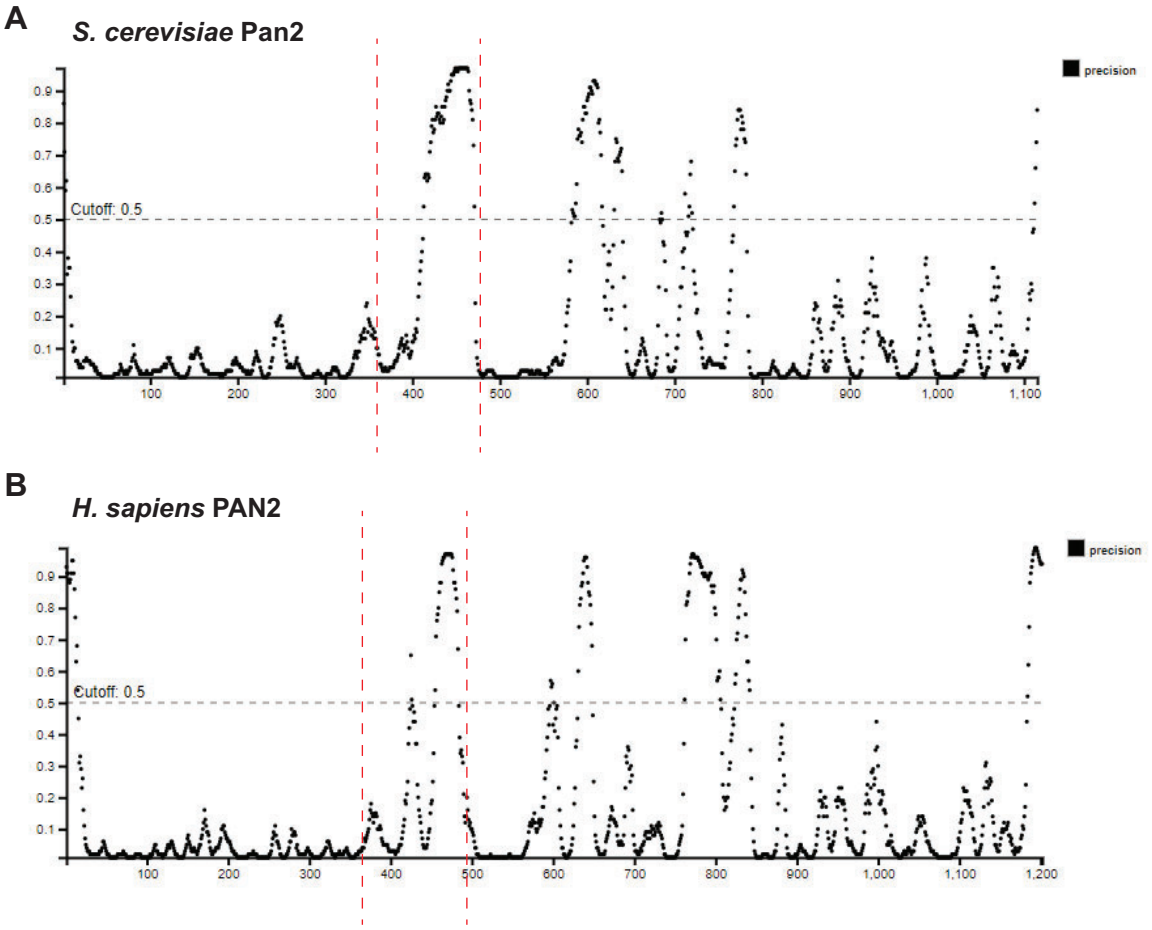
7. Supplementary Figures



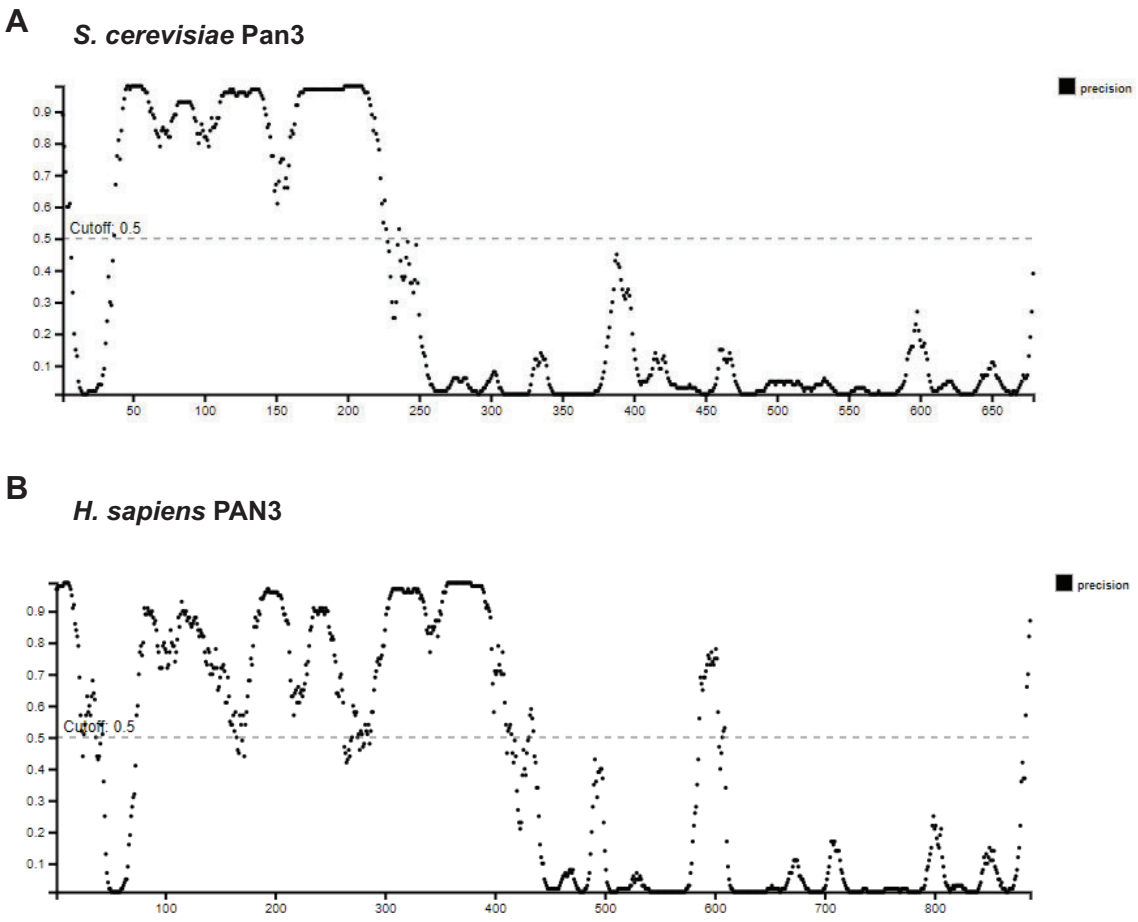
Supplementary Figure 1 Sequence alignment of *H. sapiens (Hs*, top), *S. cerevisiae* (*Sc*, middle), and *C. thermophilum* (*Ct*, bottom) Pan2. Identical amino acids are highlighted in red and similar amino acids are boxed. Of note, the C-terminal Exo domain is the most highly conserved. *H. sapiens* PAN2 contains a long N-terminal extension and a long loop between the UCH and Exo domains. *C. thermophilum* Pan2 contains a long C-terminal extension and some insertions in predicted loop regions. *S. cerevisiae* Pan2 is shortest and contains some insertions in the Pan3-interacting linker and loop regions.



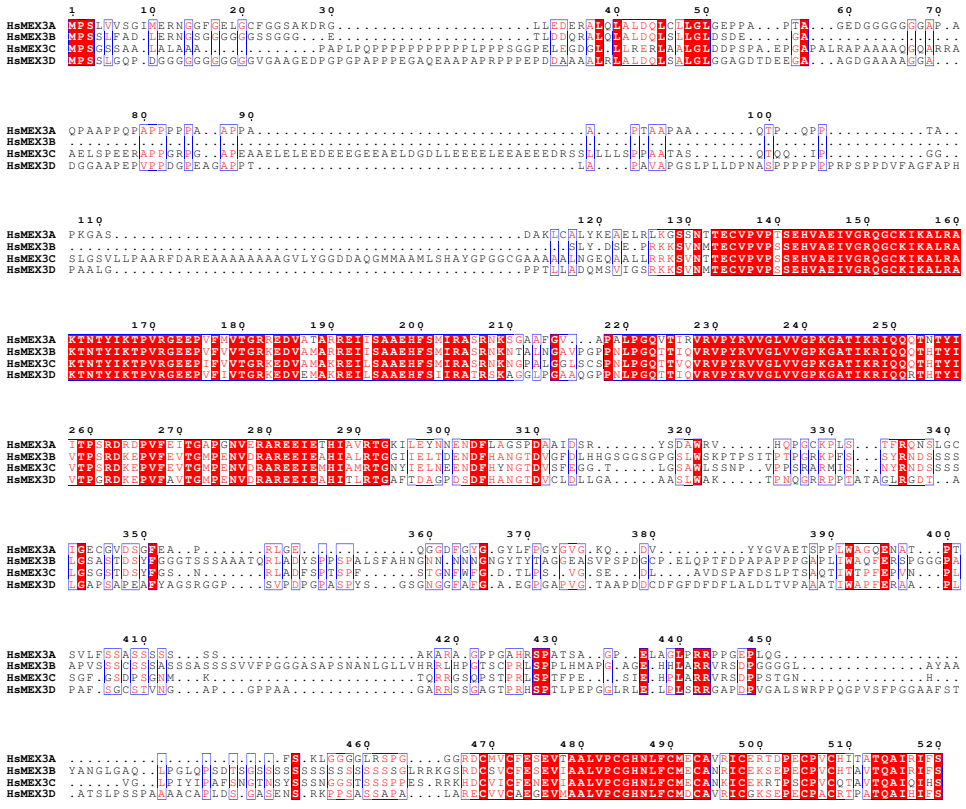
Supplementary Figure 2 Sequence alignment of *H. sapiens (Hs*, top) and *S. cerevisiae* (*Sc*, bottom) Pan3. Identical amino acids are highlighted in red and similar amino acids are boxed. The C-terminal half is most highly conserved. *H. sapiens* PAN3 contains a long N-terminal extension in its low-complexity region and additional insertions in loop regions.



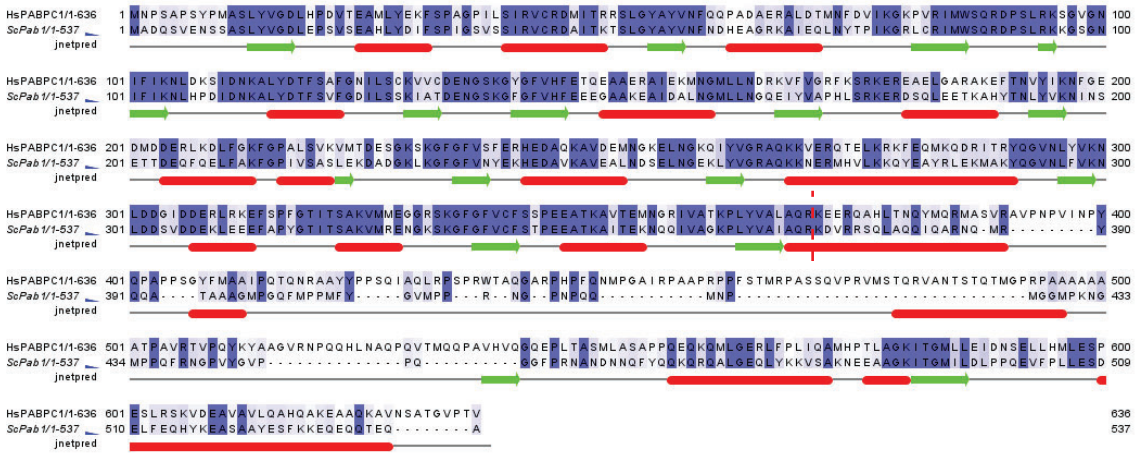
Supplementary Figure 3 DISOPRED plots of (**A**) *S. cerevisiae* Pan2 and (**B**) *H. sapiens* PAN2. The threshold for disorder is at 0.5; amino acid sequences above the threshold are predicted to be disordered. The red dash lines delineate the predicted linker regions between the N-terminal WD40 domain and UCH domain. This region is predicted to be disordered in both species.



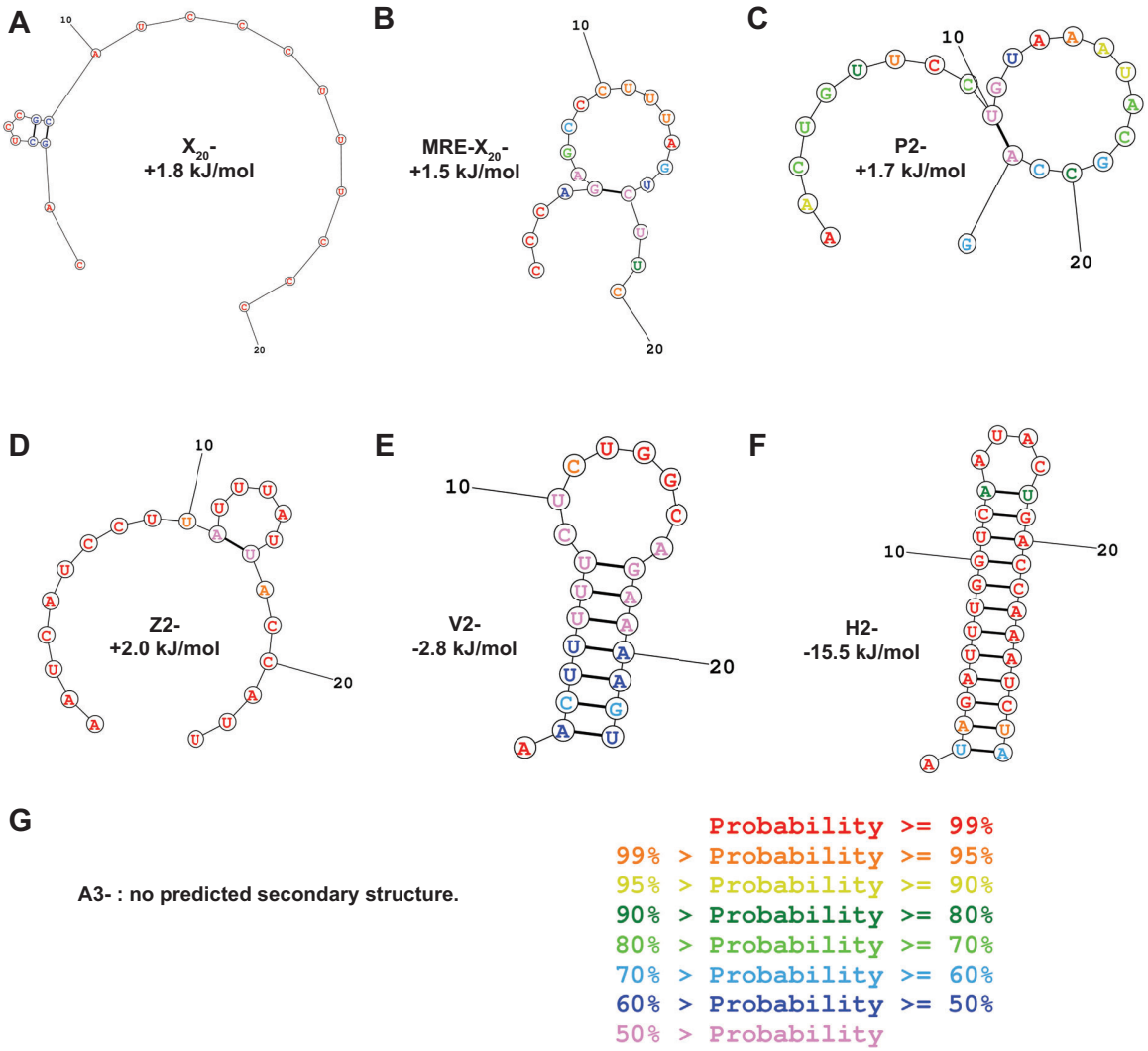
Supplementary Figure 4 DISOPRED plots of (**A**) *S. cerevisiae* Pan3 and (**B**) *H. sapiens* PAN3. The threshold for disorder is at 0.5; amino acid sequences above the threshold are predicted to be disordered. The N-termini of both *S. cerevisiae* and *H. sapiens* Pan3 are both predicted to be significantly disordered. Notably, the *H, sapiens* disordered N-terminus is predicted to be ~2× longer.



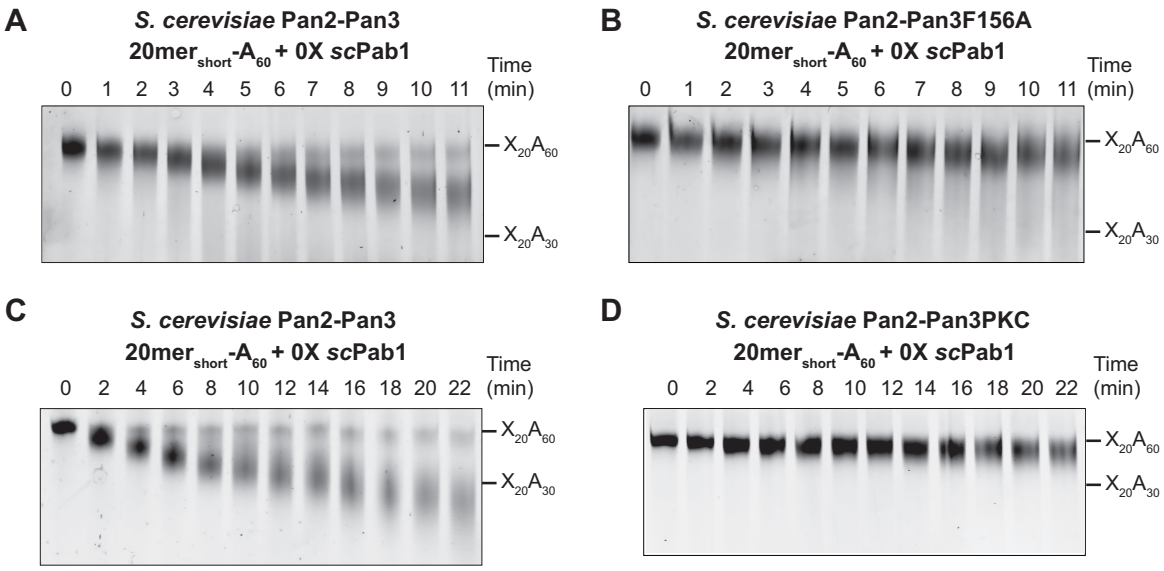
Supplementary Figure 5 Sequence alignment of *H. sapiens* MEX3A-D (top to bottom). Identical amino acids are highlighted in red and similar amino acids are boxed. The central KH domains and C-terminal RING domain are highly conserved; intervening loop regions are less conserved.



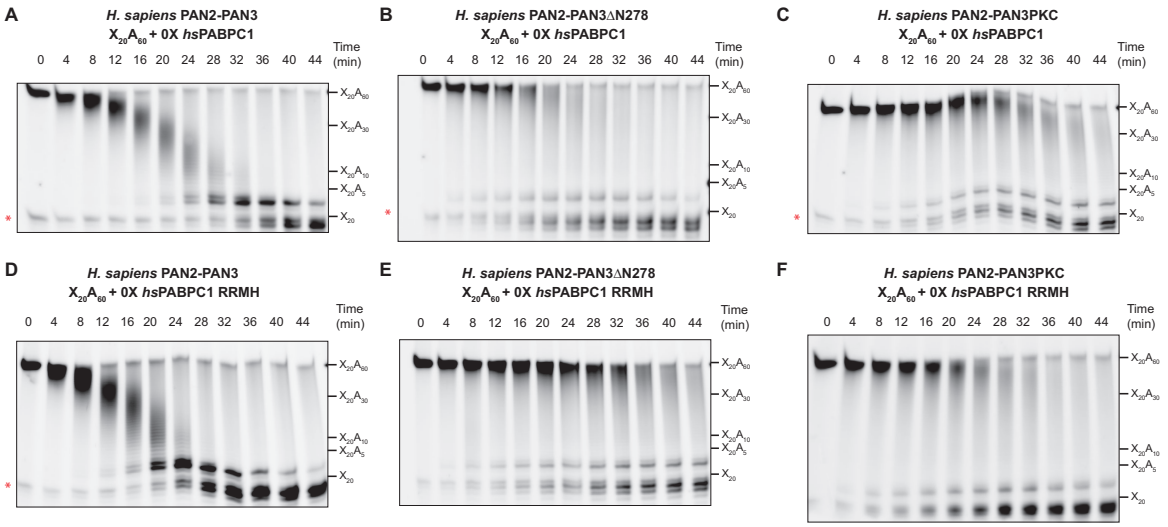
Supplementary Figure 6 Sequence alignment and secondary structure prediction of *H. sapiens* PABPC1 (top) and *S. cerevisiae* Pab1 (bottom). The predicted secondary structure is of a multiple sequence alignment of PABPC1 and Pab1. Red cylinders represent predicted alpha helices and green arrows represent predicted beta strands. The red dotted line marks the conventional domain boundary of RRM4. A long helical region is predicted to immediately follow RRM4; that region is also partially conserved between humans and yeast.



Supplementary Figure 7 Predicted secondary structures formed by RNAs used in deadenylation assays. (**A**) X_{20}-; (**B**) MRE-X_{20}-; (**C**) P2-; (**D**) Z2-; (**E**) V2-; (**F**) H2-; (**G**) A3-. Predicted base pairs are shown as bold lines between bases. The number next to each sequence is the predicted Gibson free energy for formation of the structure shown. The probability of base pairs is shown as a rainbow legend as indicated. RNAs A-D, G are unlikely to form any secondary structure, as given by the positive Gibson free energy of structure formation. RNAs E-F are likely to form stable hairpins as given by the negative Gibson free energy, and high probability of base-pairing.



Supplementary Figure 8 Negative controls (0× Pab1) for Figure 3.13A-D. (**A-D**) Deadenylation of 200 nM unlabelled $20mer_{short}$-A_{60} by 50 nM (**A**) *S. cerevisiae* Pan2-Pan3; (**B**) *S. cerevisiae* Pan2-Pan3F156A; (**C**) *S. cerevisiae* Pan2-Pan3 (activity difference vs panel A is due to the different time points); and (**D**) *S. cerevisiae* Pan2-Pan3PKC.



Supplementary Figure 9 Negative controls (0× PABPC1, PABPC1 RRMH) for Figure 3.14A-I. (**A-C**) Deadenylation of 200 nM 5' 6-FAM-labelled X_{20}-A_{60} by 100 nM (**A**) *H. sapiens* PAN2-PAN3; (**B**) *H. sapiens* PAN2-PAN3ΔN278; (**C**) *H. sapiens* PAN2-PAN3PKC. (**D-F**) Deadenylation of 200 nM 5' 6-FAM-labelled X_{20}-A_{60} by 100 nM (**D**) *H. sapiens* PAN2-PAN3; (**E**) *H. sapiens* PAN2-PAN3ΔN278; (**F**) *H. sapiens* PAN2-PAN3PKC. Red asterisks denote a contaminating RNA.